(19) DANMARK

(10) **DK/EP 3466978 T3**

Patent- og
Varemærkestyrelsen

(12) Oversættelse af
europæisk patentskrift

(51) Int.Cl.: ***C 07 K 19/00 (2006.01)*** ***C 07 K 14/415 (2006.01)*** ***C 07 K 14/435 (2006.01)***
C 12 N 15/09 (2006.01) ***C 12 N 15/62 (2006.01)***

(45) Oversættelsen bekendtgjort den: **2021-04-26**

(80) Dato for Den Europæiske Patentmyndigheds
bekendtgørelse om meddelelse af patentet: **2021-04-07**

(86) Europæisk ansøgning nr.: **17806675.9**

(86) Europæisk indleveringsdag: **2017-05-30**

(87) Den europæiske ansøgnings publiceringsdag: **2019-04-10**

(86) International ansøgning nr.: **JP2017020076**

(87) Internationalt publikationsnr.: **WO2017209122**

(30) Prioritet: **2016-06-03 US 201662345252 P** **2016-06-17 JP 2016120524**

(84) Designerede stater: **AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR**

(73) Patenthaver: **Kyushu University, National University Corporation, 744, Motooka, Nishi-ku, Fukuoka-shi, Fukuoka 819-0395, Japan**

(72) Opfinder: **NAKAMURA, Takahiro, c/o KYUSHU UNIVERSITY, NATIONAL UNIVERSITY CORPORATION, 744, Motooka, Nishi-ku, Fukuoka-shi, Fukuoka 819-0395, Japan**
YAGI, Yusuke, c/o KYUSHU UNIVERSITY, NATIONAL UNIVERSITY CORPORATION, 744, Motooka, Nishi-ku, Fukuoka-shi, Fukuoka 819-0395, Japan

(74) Fuldmægtig i Danmark: **Zacco Denmark A/S, Arne Jacobsens Allé 15, 2300 København S, Danmark**

(54) Benævnelse: **Fusionsprotein til forbedring af proteinekspression fra mål-mRNA**

(56) Fremdragne publikationer:
EP-A1- 2 784 157
WO-A1-2011/111829
WO-A1-2013/058404
WO-A1-2014/175284
JP-A- 2006 217 871
JP-A- 2009 544 296
JP-A- 2015 221 026
JP-A- 2016 039 810
US-A1- 2016 075 744
WOODSON J D et al.: "Coordination of gene expression between organellar and nuclear genomes", NATURE REVIEWS GENETICS, vol. 9, no. 5, May 2008 (2008-05), pages 383-395, XP055551958,

Description

[Technical Field]

[0001] The present invention relates to fusion proteins for improving protein expression levels from target mRNAs.

[Background Art]

[0002] Techniques of binding nucleic acid-binding protein factors revealed by a variety of analyses to sequences of interest are established and used in recent years. Use of this sequence-specific binding enables removal of a target DNA sequence or regulation (activation or inactivation) of expression of a protein coding gene present downstream of the target DNA sequence in some extent.

[0003] While zinc finger nuclease (ZFN), TAL effector nuclease (TALEN), Crispr-cas9, and the like are known as techniques using protein factors which act on DNA, the development of techniques using protein factors which act specifically to RNA is still limited.

[0004] The present inventors have proposed a method of designing a protein which can specifically bind to a target RNA sequence using the properties of PPR proteins (protein having one or more pentatricopeptide repeat (PPR) motifs), which are proteins mainly found in plants (Patent Literature 1).

[Citation List]

[Patent Literature]

[0005] [Patent Literature 1]

(1) WO2013/058404
(2) EP 2 784 157
(3) US 2016/075744

[Summary of Invention]

[Technical Problem]

[0006] In the disclosure according to Patent Literature 1, the amino acids which function when a PPR motif demonstrates RNA-binding properties were identified, and the relation between the structure of the PPR motif and the target base was revealed, thereby enabling the construction of proteins which have one or more PPR motifs and can bind to RNAs having any sequence and length. However, no method has ever been found which actually regulates target RNAs using the techniques according to Patent Literature 1.

[0007] Patent literature 2 relates to DNA binding proteins using PPR motif. The document teaches a complex comprising a functional region bound to a PPR motif-containing protein. However, there is no experimental data on a complex comprising a PPR motif region linked to a functional region. In addition, patent literature 2 only suggests "RNAse" as a proteinic functional region but does not at all teach or suggest a domain which guides ribosome to the mRNA, a domain which initiates or promotes translation of the mRNA, a domain associated with nuclear export of the mRNA, a domain binding to an endoplasmic reticulum membrane, a domain containing an endoplasmic reticulum retention signal (ER retention signal) sequence or a domain containing an endoplasmic reticulum signal sequence.

[0008] Patent literature 3 relates to DNA binding proteins using PPR motif but it does not teach or suggest improving protein expression from a target RNA by using the PPR protein

[Solution to Problem]

[0009] As a result of extensive research on a method of improving a protein expression level from a target mRNA using a PPR protein, the present inventors have found that a fusion protein of a predetermined functional domain and a PPR protein improves the protein expression level from the target mRNA, and have completed the present invention.

[0010] Specifically, an embodiment of the present invention relates to a fusion protein for improving a protein expression level from a target mRNA, the fusion protein comprising:

(A) one or more functional domains which improve a protein expression level from an mRNA; and
(B) a polypeptide moiety which can bind to a target mRNA in an RNA base-selective or RNA base sequence-specific

manner,

wherein polypeptide moiety (B) is a polypeptide moiety comprising one or more PPR motifs, each PPR motif comprising a polypeptide consisting of 30 to 38 amino acids in length and being represented by Formula 1:

[Formula 1]

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

where

Helix A is a moiety which consists of 12 amino acids in length and can form an α-helix structure, and is represented by Formula 2:

[Formula 2]

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

where A_1 to A_{12} each independently represent an amino acid;
X is not present, or is a moiety consisting of 1 to 9 amino acids in length;
Helix B is a moiety which consists of 11 to 13 amino acids in length and can form an structure;
L is a moiety consisting of 2 to 7 amino acids in length and represented by Formula 3:

[Formula 3]

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

where the amino acids are numbered from the C-terminal as "i" (-1), "ii" (-2), ... and
L_{iii} to L_{vii} may not be present, and
a combination of three amino acids A_1, A_4, and L_{ii} or a combination of two amino acids A_4 and Lii corresponds to a base or base sequence of the target mRNA,
wherein one or more functional domains (A) are selected from the group consisting of a domain which guides ribosome to the mRNA, a domain which initiates or promotes translation of the mRNA, a domain associated with nuclear export of the mRNA, a domain which binds to an endoplasmic reticulum membrane, a domain containing an endoplasmic reticulum retention signal (ER retention signal) sequence, and a domain containing an endoplasmic reticulum signal sequence.

[0011] In an embodiment according to the present invention, polypeptide moiety (B) comprises 2 to 30 PPR motifs, and the plurality of PPR motifs is arranged so as to specifically bind to the base sequence of the target mRNA.

[0012] Moreover, in an embodiment according to the present invention, polypeptide moiety (B) comprises 5 to 25 PPR motifs.

[0013] Moreover, in an embodiment according to the present invention, one or more functional domains (A) each bind to an N-terminal side and/or a C-terminal side of polypeptide moiety (B).

[0014] Moreover, in an embodiment according to the present invention, the domain which guides ribosome to the mRNA is a domain containing all or functional part of a polypeptide selected from the group consisting of DENR (Density-regulated protein), MCT-1 (Malignant T-cell amplified sequence 1), TPT1 (Translationally-controlled tumor protein), and Lerepo4 (Zinc finger CCCH-domain),
the domain initiating or promoting translation of the mRNA is a domain containing all or functional part of a polypeptide selected from the group consisting of eIF4E and eIF4G,
the domain associated with nuclear export of the mRNA is a domain containing all or functional part of SLBP (Stem-loop binding protein),
the domain binding to an endoplasmic reticulum membrane is a domain containing all or functional part of a polypeptide selected from the group consisting of SEC61B, TRAP-alpha (Translocon associated protein alpha), SR-alpha, Dial (Cytochrome b5 reductase 3), and p180,
the endoplasmic reticulum retention signal (ER retention signal) sequence is a signal sequence containing a KDEL (KEEL) sequence, or
the endoplasmic reticulum signal sequence is a signal sequence containing MGWSCIILFLVATATGAHS (SEQ ID NO: 22).

[0015] Moreover, in an embodiment according to the present invention, the combination of the three amino acids A_1, A_4, and L_{ii} in each of the PPR motifs is:

(valine, threonine, asparagine), (phenylalanine, serine, asparagine), (phenylalanine, threonine, asparagine), (isoleucine, asparagine, aspartic acid), or (threonine, threonine, asparagine) in order of (A_1, A_4, L_{ii}) if a target base for the PPR motif is A (adenine);
(glutamic acid, glycine, aspartic acid), (valine, threonine, aspartic acid), (lysine, threonine, aspartic acid), or (leucine, threonine, aspartic acid) in order of (A_1, A_4, L_{ii}) if the target base for the PPR motif is G (guanine);
(valine, asparagine, aspartic acid), (isoleucine, asparagine, asparagine), (isoleucine, asparagine, aspartic acid), (isoleucine, methionine, aspartic acid), (phenylalanine, proline, aspartic acid), or (tyrosine, proline, aspartic acid) in order of (A_1, A_4, L_{ii}) if the target base for the PPR motif is U (uracil); or
(valine, asparagine, asparagine), (isoleucine, asparagine, asparagine), (valine, asparagine, serine), or (isoleucine, methionine, aspartic acid) in order of (A_1, A_4, L_{ii}) if the target base for the PPR motif is C (cytosine).

[0016] Moreover, in an embodiment according to the present invention, the combination of the two amino acids A_4 and L_{ii} in each of the PPR motifs is:

(threonine, asparagine), (serine, asparagine), or (glycine, asparagine) in order of (A_4, Lii) if a target base for the PPR motif is A (adenine);
(threonine, aspartic acid) or (glycine, aspartic acid) in order of (A_4, L_{ii}) if the target base for the PPR motif is G (guanine);
(asparagine, aspartic acid), (proline, aspartic acid), (methionine, aspartic acid), or (valine, threonine) in order of (A_4, Lii) if the target base for the PPR motif is U (uracil); or
(asparagine, asparagine), (asparagine, serine), or (leucine, aspartic acid) in order of (A_4, L_{ii}) if the target base for the PPR motif is C (cytosine).

[0017] Another embodiment according to the present invention relates to a nucleic acid which encodes the fusion protein according to the present invention.

[0018] Still another embodiment according to the present invention relates to a vector (preferably an expression vector) comprising the nucleic acid according to the present invention.

[0019] Further still another embodiment according to the present invention relates to an *in vitro* method of improving a protein expression level from a target mRNA within a cell, the method comprising:

a step of providing the fusion protein according to the present invention or the vector according to the present invention; and
a step of introducing the fusion protein or the vector into the cell.

[0020] Moreover, in an embodiment according to the present invention, the cell is a eukaryotic cell.

[0021] Moreover, in an embodiment according to the present invention, the cell is an animal cell.

[0022] Moreover, in an embodiment according to the present invention, the animal cell is a human cell.

[0023] Inventions having any combination of one or more features of the present invention described above are also included in the scope of the present invention.

[Brief Description of Drawings]

[0024]

[Figure 1] Figure 1 illustrates a schematic view of an effector plasmid and a reporter plasmid used in Examples, and a schematic view of an experimental outline. Figure 1A illustrates a schematic view of the effector plasmid and the reporter plasmid used in Examples. A fusion protein of PPR motifs and eIF4G expresses from the effector plasmid. In Examples, a CRR4 protein was used, whose target sequence is well researched. From the reporter plasmid, renilla luciferase (RLuc) and firefly luciferase (FLuc) are transcribed in the form of a dicistronic mRNA. A PPR-binding sequence (here, CRR4-binding sequence) was inserted into a site on the 5' end of FLuc. Figure 1B illustrates a schematic view of an experimental outline of Examples. Irrespective of the presence/absence of the PPR-binding sequence, RLuc is translated at a similar level. For this reason, the activity value of RLuc can be treated as a control in transfection in this reporter system. The translation of Fluc is started only when PPR-eIF4G binds to the PPR-binding sequence and translation factors can be attracted by the effects of eIF4G. In contrast, the translation of FLuc remains at a low level if the PPR-binding sequence is not present and thus, PPR-eIF4G cannot bind to the PPR-binding sequence.
[Figure 2] Figure 2 illustrates an experimental procedure of a reporter assay using HEK293T cells.
[Figure 3] Figure 3 shows the experimental results of Example 1. The activation of sequence-specific translation depends on CRR4-eIF4G and the PPR-binding sequence. This experiment was performed using an effector plasmid,

into which CRR4-Flag (without translation activating factor, in white) or CRR4-eIF4G (with translation activating factor, in gray) was inserted, and a reporter vector with or without an inserted PPR-binding sequence. From the results, it was verified that specific translation activity increased 2.75 times in the presence of both PPR-eIF4G and the PPR-binding sequence. The value represents the average and the standard deviation (N = 3).

[Figure 4] Figure 4 illustrates an outline of the experiment in Example 2.

[Figure 5] Figure 5 illustrates the experimental results in Example 2 and the functions of the domains.

[Figure 6] Figure 6 illustrates the experimental results in Example 2 and the functions of the domains.

[Description of Embodiment]

[PPR motifs and PPR proteins]

[0025] Unless otherwise specified, the term "PPR motif" as used herein indicates a polypeptide which is composed of 30 to 38 amino acids and has an amino acid sequence having an E value equal to or less than a predetermined value (desirably E-03), the E value being obtained at PF01535 in Pfam and PS51375 at Prosite during the analysis of the amino acid sequence with a protein domain search program on the Web. The position number of an amino acid forming the PPR motif defined in the present invention is substantially as defined as PF01535 while it corresponds to the number obtained by subtracting 2 from the location of the amino acid in PS51375 (for example, position 1 in the present invention corresponds to position 3 in PS51375). Note that the term "ii" (-2)-th amino acid refers to the second amino acid from the tail end (C-terminal side) of the amino acids forming one PPR motif or the amino acid close to the N-terminal by two amino acids from the first amino acid of the next PPR motif (that is, -2 amino acid). If the next PPR motif is not clearly identified, the forward amino acid by two amino acids from the first amino acid of the next helix structure is defined as "ii". See http://pfam.sanger.ac.uk/ for Pfam and http://www.expasy.org/prosite/ for Prosite.

[0026] Although the conserved amino acid sequence of the PPR motif has low conservation properties at the amino acid level, two α-helices are well conserved on the secondary structure. Although a typical PPR motif is composed of 35 amino acids, its length is variable from 30 to 38 amino acids.

[0027] According to the invention, the term PPR motif used in the present invention is composed of a polypeptide having 30 to 38 amino acids in length and being represented by Formula 1:

[Formula 4]

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

where

Helix A is a moiety which consists of 12 amino acids in length and can form an structure, and is represented by Formula 2:

[Formula 5]

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

where A_1 to A_{12} each independently represent an amino acid;

X is not present, or is a moiety consisting of 1 to 9 amino acids in length;

Helix B is a moiety which consists of 11 to 13 amino acids in length and can form the structure; and

L is a moiety consisting of 2 to 7 amino acids in length and represented by Formula 3:

[Formula 6]

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

where the amino acids are numbered from the C-terminal side as "i" (-1), "ii" (-2), ... and

L_{iii} to L_{vii} may not be present, and

a combination of three amino acids A_1, A_4, and L_{ii} or a combination of two amino acids A_4 and Lii corresponds to a base or base sequence of the target mRNA,

wherein one or more functional domains (A) are selected from the group consisting of a domain which guides ribosome to the mRNA, a domain which initiates or promotes translation of the mRNA, a domain associated with nuclear export of the mRNA, a domain which binds to an endoplasmic reticulum membrane, a domain containing an endoplasmic reticulum retention signal (ER retention signal) sequence, and a domain containing an endoplasmic reticulum signal sequence.

[0028] Unless otherwise specified, the term "PPR protein" used in the present invention indicates a PPR protein comprising one or more PPR motifs described above, preferably two or more PPR motifs described above. Unless otherwise specified, the term "protein" used herein generally indicates substances consisting of polypeptides (chains of several amino acids bound through peptide bond), also including those consisting of relatively low molecular weight polypeptides. The term "amino acid" used in the present invention may indicates a usual amino acid molecule, or otherwise may indicate an amino acid residue forming a peptide chain in some cases. Persons skilled in the art clearly understand from contexts which case the term indicates.

[0029] Unless otherwise specified, the "selective" used in the present invention about the binding properties of the PPR motif to the RNA bases indicates that the binding activity of a PPR motif to one of the RNA bases is higher than the binding activity thereof to other bases. Persons skilled in the art can plan the experiment for this selectivity and verify it, and can also determine through calculation.

[0030] Unless otherwise specified, the term "RNA base" used in the present invention indicates a base of a ribonucleotide forming an RNA, specifically adenine (A), guanine (G), cytosine (C), or uracil (U). Note that although the PPR protein can have selectivity to the base in the RNA, it does not bind to a nucleic acid monomer.

[0031] PPR protein is present in many plants, and 500 proteins, about 5000 motifs can be found in Arabidopsis thaliana. PPR motifs and PPR proteins having a variety of amino acid sequences are also present in many land plants such as Oryza, Populus, and Selaginella tamariscina. In the present invention, PPR motifs and PPR proteins present in the natural world may be used, or PPR motifs and PPR proteins designed based on the method disclosed in WO2013/058404, for example, may be used. Specifically, desired PPR motifs and PPR proteins can be designed based on the following information disclosed in WO2013/058404.

(I) Information on the position of the amino acid essential for selective binding

[0032] The combination (A1, A4, Lii) of three, i.e., 1st, 4th, and "ii" (-1)-th amino acids of a PPR motif or the combination (A4, Lii) of two, i.e., 4th and "ii" (-1)-th amino acids is essential for selective binding to the RNA base, and the target RNA base for binding can be determined by these combinations.

[0033] The present invention can use the findings about the combination of three amino acids A1, A4, and Lii, and/or the combination of two amino acids A4 and Lii disclosed in WO2013/058404.

(II) Information about the correspondence of the combination of three amino acids A1, A4, and Lii to RNA bases

[0034]

(3-1) If the combination of three amino acids A1, A4, and Lii is valine, asparagine, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to U is the strongest, and binding to C is the second strongest, followed by binding to A or G.

(3-2) If the combination of three amino acids A1, A4, and Lii is valine, threonine, and asparagine in this order, the PPR motif has a selective RNA base binding ability as follows: binding to A is the strongest, and binding to G is the second strongest, followed by binding to C without binding to U.

(3-3) If the combination of three amino acids A1, A4, and Lii is valine, asparagine, and asparagine in this order, the PPR motif has a selective RNA base binding ability as follows: binding to C is the strongest and binding to A or U is the second strongest, without binding to G.

(3-4) If the combination of three amino acids A1, A4, and Lii is glutamic acid, glycine, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to G is strong, without binding to A, U, or C.

(3-5) If the combination of three amino acids A1, A4, and Lii is isoleucine, asparagine, and asparagine in this order, the PPR motif has a selective RNA base binding ability as follows: binding to C is the strongest, and binding to U is the second strongest, followed by binding to A, without binding to G.

(3-6) If the combination of three amino acids A1, A4, and Lii is valine, threonine, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to G is the strongest and binding to U is the second strongest, without binding to A or C.

(3-7) If the combination of three amino acids A1, A4, and Lii is lysine, threonine, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to G is the strongest and binding to A is the second strongest, without binding to U or C.

(3-8) If the combination of three amino acids A1, A4, and Lii is phenylalanine, serine, and asparagine in this order, the PPR motif has a selective RNA base binding ability as follows: binding to A is the strongest, and binding to C is the second strongest, followed by binding to G and U.

(3-9) If the combination of three amino acids A1, A4, and Lii is valine, asparagine, and serine in this order, the PPR motif has a selective RNA base binding ability as follows: binding to C is the strongest and binding to U is the second

strongest, without binding to A or G.
(3-10) If the combination of three amino acids A1, A4, and Lii is phenylalanine, threonine, and asparagine in this order, the PPR motif has a selective RNA base binding ability as follows: binding to A is strong, without binding to G, U, or C.
(3-11) If the combination of three amino acids A1, A4, and Lii is isoleucine, asparagine, aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to U is the strongest and binding to A is the second strongest, without binding to G or C.
(3-12) If the combination of three amino acids A1, A4, and Lii is threonine, threonine, and asparagine in this order, the PPR motif has a selective RNA base binding ability as follows: binding to A is strong, without binding to G, U, or C.
(3-13) If the combination of three amino acids A1, A4, and Lii is isoleucine, methionine, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to U is the strongest and binding to C is the second strongest, without binding to A or G.
(3-14) If the combination of three amino acids A1, A4, and Lii is phenylalanine, proline, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to U is the strongest and binding to C is the second strongest, without binding to A or G.
(3-15) If the combination of three amino acids A1, A4, and Lii is tyrosine, proline, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to U is strong, without binding to A, G, or C.
(3-16) If the combination of three amino acids A1, A4, and Lii is leucine, threonine, and aspartic acid in this order, the PPR motif has a selective RNA base binding ability as follows: binding to G is strong, without binding to A, U, or C.

(II) Information about the correspondence of the combination of two amino acids A4 and Lii to the RNA bases

[0035]

(2-1) If A4 and Lii in this order are asparagine and aspartic acid, the PPR motif has a selective RNA base binding ability as follows: binding to U is the strongest, and binding to C is the second strongest, followed by binding to A and G.
(2-2) If A4 and Lii in this order are asparagine and asparagine, the PPR motif has a selective RNA base binding ability as follows: binding to C is the strongest, binding to U is the second strongest, followed by binding to A and G.
(2-3) If A4 and Lii in this order are threonine and asparagine, the PPR motif has a selective RNA base binding ability with strong binding to A and weak binding to G, U, and C.
(2-4) If A4 and Lii in this order are threonine and aspartic acid, the PPR motif has a selective RNA base binding ability with strong binding to G and weak binding to A, U, and C.
(2-5) If A4 and Lii in this order are serine and asparagine, the PPR motif has a selective RNA base binding ability as follows: binding to A is the strongest and binding to G, U, and C is the second strongest.
(2-6) If A4 and Lii in this order are glycine and aspartic acid, the PPR motif has a selective RNA base binding ability as follows: binding to G is the strongest, and binding to U is the second strongest, followed by binding to A, without binding to C.
(2-7) If A4 and Lii in this order are asparagine and serine, the PPR motif has a selective RNA base binding ability as follows: binding to C is the strongest, and binding to U is the second strongest, followed by binding to A and G.
(2-8) If A4 and Lii in this order are proline and aspartic acid, the PPR motif has a selective RNA base binding ability as follows: binding to U is the strongest, and binding to G, C, and C is the second strongest, without binding to A.
(2-9) If A4 and Lii in this order are glycine and asparagine, the PPR motif has a selective RNA base binding ability as follows: binding to A is the strongest, and binding to G is the second strongest, without binding to C or U.
(2-10) If A4 and Lii in this order are methionine and aspartic acid, the PPR motif has a selective RNA base binding ability with strong binding to U and weak binding to A, G, and C.
(2-11) If A4 and Lii in this order are leucine and aspartic acid, the PPR motif has a selective RNA base binding ability as follows: binding to C is the strongest, and binding to U is the second strongest, without binding to A or G.
(2-12) If A4 and Lii in this order are valine and threonine, the PPR motif has a selective RNA base binding ability as follows: binding to U is the strongest, and binding to A is the second strongest, without binding to G or C.

[Use of PPR motifs and PPR proteins]

Identification and design:

[0036] One PPR motif can recognize a specific base of an RNA. According to the present invention, PPR motifs selective to A, U, G, or C can be selected or designed by disposing appropriate amino acids in specific positions of a PPR motif. Furthermore, a protein containing an appropriate series of such PPR motifs can recognize its corresponding specific sequence. Moreover, according to the findings described above, a PPR motif which can selectively bind to a

desired RNA base and a protein having a plurality of PPR motifs which can sequence-specifically bind to a desired RNA can be designed. In design, the sequence information of a naturally occurring PPR motif may be referred with respect to moieties other than the amino acids disposed in the important positions of the PPR motif. Alternatively, a PPR motif may be designed by using a naturally occurring PPR motif as a whole and replacing only the amino acids in the important positions with other amino acids. The repetition number of the PPR motif can be appropriately determined according to the target sequence; for example, the repetition number can be 2 or more, or 2 to 30.

[0037] The PPR motif or PPR protein thus designed can be prepared by a method well known to persons skilled in the art. For example, a nucleic acid sequence encoding an amino acid sequence of the designed PPR motif or PPR protein can be determined from the amino acid sequence, and may be cloned to prepare a transformant (such as an expression vector) which produces a desired PPR motif or PPR protein.

Preparation and use of fusion protein:

[0038] The present invention relates to a fusion protein of the PPR motif or PPR protein described above as defined in the claims (i.e., a polypeptide which can bind RNA base-selectively or RNA base sequence-specifically to the target mRNA) and one or more functional domains which improve a protein expression level from an mRNA.

[0039] The "functional domain which improves a protein expression level from an mRNA" which can be used in the present invention may be all or functional part of a functional domain of a known protein which directly or indirectly promotes the translation of the mRNA, for example. According to the invention, the functional domain which can be used in the present invention is selected from a domain which guides ribosome to the mRNA, a domain initiating or promoting translation of the mRNA, a domain associated with nuclear export of the mRNA, a domain binding to an endoplasmic reticulum membrane, a domain containing an endoplasmic reticulum retention signal (ER retention signal) sequence, and a domain containing an endoplasmic reticulum signal sequence, for example.

[0040] More specifically, the domain which guides ribosome to the mRNA may be a domain containing all or functional part of a polypeptide selected from the group consisting of DENR (Density-regulated protein), MCT-1 (Malignant T-cell amplified sequence 1), TPT1 (Translationally-controlled tumor protein), and Lerepo4 (Zinc finger CCCH-domain). The domain initiating or promoting translation of the mRNA may be a domain containing all or functional part of a polypeptide selected from the group consisting of eIF4E and eIF4G. The domain associated with nuclear export of the mRNA may be a domain containing all or functional part of SLBP (Stem-loop binding protein). The domain binding to an endoplasmic reticulum membrane may be a domain containing all or functional part of a polypeptide selected from the group consisting of SEC61B, TRAP-alpha (Translocon associated protein alpha), SR-alpha, Dial (Cytochrome b5 reductase 3), and p180. The endoplasmic reticulum retention signal (ER retention signal) sequence may be a signal sequence containing a KDEL (KEEL) sequence. The endoplasmic reticulum signal sequence may be a signal sequence containing MGWSCIIL-FLVATATGAHS (SEQ ID NO: 22).

[0041] In the fusion protein according to the present invention, the functional domain may be fused to the N-terminal side of the PPR protein, may be fused to the C-terminal side of the PPR protein, or may be fused to both of the N-terminal side and the C-terminal side thereof. Moreover, the fusion protein according to the present invention may include several functional domains (for example, 2 to 5 functional domains). Furthermore, in the fusion protein according to the present invention, the functional domain and the PPR protein may be indirectly fused via a linker, for example.

[0042] The present invention also relates to a nucleic acid encoding the fusion protein described above, and a vector (such as an expression vector) comprising the nucleic acid. The expression vector herein refers to, for example, a vector comprising a DNA having a promoter sequence, a DNA encoding a desired protein, and a DNA having a terminator sequence, in this order from upstream. The expression vector may not have these DNAs in this order as long as it demonstrates desired functions. A variety of expression vectors which can be usually used by persons skilled in the art can be used in the present invention.

[0043] Because the fusion protein according to the present invention uses the RNA translation mechanism of eukaryotes, it can function in cells of eukaryotes (such as animals, plants, microorganisms (e.g., yeasts), and protists). The fusion protein according to the present invention can function within animal cells (in vitro or in vivo) in particular. Examples of animal cells into which the fusion protein according to the present invention or a vector which expresses the fusion protein according to the present invention can be introduced can include cells derived from human, monkey, pig, cow, horse, dog, cat, mouse, and rat. Examples of cultured cells into which the fusion protein according to the present invention or a vector which expresses the fusion protein according to the present invention can be introduced can include, but should not be limited to, Chinese hamster ovarian (CHO) cells, COS-1 cells, COS-7 cells, VERO (ATCC CCL-81) cells, BHK cells, dog kidney-derived MDCK cells, hamster AV-12-664 cells, HeLa cells, Wl38 cells, 293 cells, 293T cells, and PER.C6 cells.

[0044] The terms used herein excluding those particularly defined are used for illustration of the specific embodiments, and are not intended to be limitative to the invention.

[0045] The term "comprise" used herein, unless contexts clearly require different understandings, is intended to express

that a described entry (such as a member, a step, a component, or a number) is present, and is intended not to exclude the presence of other entries (such as a member, a step, a component, or a number).

[0046] Unless otherwise defined, all the terms used herein (including technical terms and scientific terms) have the same meanings as those broadly understood by persons skilled in the art to which the present invention belongs. Unless otherwise clearly defined, the terms used herein should be interpreted as having the meanings consistent to those herein and its related technical field, and should not be interpreted as idealized or excessively formal meanings.

[0047] Hereinafter, the present invention will be described more in detail with reference to Examples. However, the present invention can be implemented with a variety of aspects, and should not be construed as limitative to Examples described below.

[Examples]

Example 1: Improvement in protein expression level from target mRNA by fusion protein of PPR motif and eIF4G

Materials

(Equipment)

[0048]

- Basic facility for molecular biological experiment (for construction of plasmids, for example)
- Inverted microscope (DM IL S40, Leica Microsystems, Wetzlar, Germany)
- CO_2 incubator (KM-CC17RH2, Panasonic Healthcare, Tokyo, Japan)
- Clean bench (MHE-S1300A2, Panasonic Healthcare, Tokyo, Japan)
- Aspirator (SP-30, Air Liquide Medical Systems, Bovezzo BS, Italy)
- Centrifuge (swing rotor) (LC-200, Tomy Seiko, Tokyo, Japan)
- Ultra-low temperature freezer (-80°C) (MDF-C8V, Panasonic Healthcare, Tokyo, Japan)
- plate reader (EnSight Kaleido, PerkinElmer, Waltham, MA, USA)

(Cell culturing)

[0049]

- HEK293T cell line (see note 1)
- Dulbecco's modified Eagle's culture medium (DMEM, glucose-rich) (see note 2)
- 100× penicillin-streptomycin solution
- Fetal bovine serum (FBS) (see note 3)
- EDTA-NaCl solution: 10 mM EDTA and 0.85% (w/v) NaCl, pH adjusted to 7.2 to 7.4, autoclave sterilized, stored at room temperature
- 100 × 20 mm cell culture petri dish (Greiner bio one, Frickenhausen, Germany)
- 10 mL disposable sterilized pipette
- 15 mL and 50 mL plastic centrifuge tubes
- 1.8 mL cryotube (Nunc; Thermo Fisher Scientific, Waltham, MA, USA)
- Freeze container (Nalgene; Thermo Fisher Scientific, Waltham, MA, USA)
- Bambanker (Lymphotec, Tokyo, Japan)

(Transfection)

[0050]

- Effector plasmid: pcDNA3.1 (Thermo Fisher Scientific, Waltham, MA, USA) was used as a basic vector. A fusion gene of PPR and eIF4G is inserted into an expression cassette (100 ng/μL) (see note 4).
- Reporter plasmid: pcDNA3.1 (Thermo Fisher Scientific, Waltham, MA, USA) was used as a basic vector. Luciferase genes are inserted into an expression cassette, and a PPR-binding sequence is inserted into its 5'-UTR (100 ng/μL).
- 96-well plate coated with poly-L-lysine (AGC Techno glass, Shizuoka, Japan)
- 1x phosphate-buffered saline, PBS(-): 1.47 mM KH_2PO_4, 8.1 mM Na_2HPO_4, 137 mM NaCl, and 2.7 mM KCl. pH adjusted to 7.4, autoclave sterilized, stored at room temperature
- Hemocytometer (for counting the number of cells) (Improved Neubauer Type Cell counter plate, Watson, Hyogo,

Japan)
- Transfection reagent (HilyMax, Dojindo Molecular Technologies, Kumamoto, Japan)

(Luciferase assay)

[0051]

- Dual-Glo Luciferase Assay System (Promega, Madison, WI, USA.)
- 96-well luminometer plate (PerkinElmer, Waltham, MA, USA).

Experimental method

(Construction of vector)

[0052] The reporter assay requires an effector plasmid and a reporter plasmid. These two plasmids both are constructed based on pcDNA3.1. The effector plasmid includes a fusion gene encoding a PPR protein and a partial domain of human eIF4G (SEQ ID NO: 1) (Figure 1A). The PPR protein moiety used was CRR4 (SEQ ID NO: 2). The reporter plasmid includes two open reading frames (ORFs), specifically, renilla luciferase (RLuc) and firefly luciferase (FLuc), which are dicistronically transcribed (Figure 1A). The RLuc gene is located on the side of the 5'-end of the FLuc gene, and was used as a control of gene expression. The PPR-binding region is inserted into the 5'-UTR of the ORF of FLuc, and consists of three repetitions of a CRR4-recognizing sequence (5'-UAUCUUGUCUUUA-3') (SEQ ID NO: 3) interrupted with four-base sequences (ATCG and GATC). To express both of the fused effector gene and the reporter gene, a cytomegalovirus promoter (CMV) and a bovine growth hormone gene-derived polyadenylation signal were used. For a control experiment, an effector plasmid having no eIF4G was constructed by fusing a FLAG epitope tag to the PPR. A control reporter plasmid without a PPR-binding region was also constructed.

[0053] The outline of the procedures from cell culturing to the reporter assay in Examples is shown in Figure 2.

(Cell culturing from frozen stock)

[0054] This step is aseptically performed. All the tools are preliminarily antisepticized with 70% ethanol.

1. A 9 mL DMEM culture medium is placed into a 15 mL centrifuge tube (sterilized).
2. 1 mL of frozen HEK293T cells in a cryotube is incubated within a water bath at 37°C to quickly melt the cells.
3. The cells are placed into the 15 mL centrifuge tube containing 9 mL DMEM.
4. The centrifuge tube is centrifuged at room temperature and 1100 ×g for two minutes, and the supernatant is removed.
5. The cells are resuspended in 10 mL DMEM (FBS is added such that the final concentration is 10%).
6. The suspended cells are transferred into a 100 mm petri dish. The petri dish was left to stand in an incubator at 37°C and under a 5% CO_2 condition. If the culturing was started from the frozen stock, the cultured cells were subcultured after 24 hours.

[0055] To keep the cells healthy (see note 5), the cell density on the surface of the petri dish is maintained between 10% and 80%. The passage is basically performed every three days (two times a week), or is performed according to the growth rate of the cells. Furthermore, to keep the number of passages small, cells are freshly cultured from the frozen stock once a month. Keeping the number of passages small and thus keeping the cells healthy are important for efficient DNA transfection.

(Passages to maintain cells)

[0056]

1. New 100 mm petri dishes are provided as required. 8 mL DMEM and 1 mL FBS are preliminarily placed onto each of the petri dishes.
2. The culture medium on a petri dish containing the cultured cells is removed with an aspirator (see note 6).
3. 2 mL EDTA-NaCl solution is gently added onto adhering cells on the surface of the petri dish so as not to peel off the cells. The petri dish is turned around to evenly distribute the solution across the entire surface of the petri dish. The EDTA-NaCl solution is removed with an aspirator. The petri dish is tapped to peel off the cells.
4. 10 mL DMEM is added to the cells in the petri dish, and the cells are suspended by gently pipetting.

5. 1 mL suspended cells (10% cultured cells) are added to the petri dishes preliminarily provided and each containing 9 mL culture medium. Each of the petri dishes is turned around to distribute the cells across the entire surface thereof.

(Freeze storage of cells)

[0057] A frozen stock is prepared with Bambanker reagent and cultured cells in a logarithmic growth phase at a cell density up to 50%. Use of Bambanker provides a high recovery rate and facilitates long-term storage.

1. The cells on the second day since the passage are peeled off according to the procedure for passage. 5 to 10 mL DMEM is added, and the cells are recovered in a 50 mL centrifuge tube.
2. The centrifuge tube is centrifuged at room temperature and 1100 ×g for two minutes, and the supernatant is removed.
3. 1 mL Bambanker per petri dish is added to suspend the cells.
4. The suspended cells are quickly dispensed into cryotubes, and the cryotubes are covered with their lids.
5. The cryotubes are placed in a dedicated freeze container, and are left to stand at -80°C for 12 hours (see note 7).
6. The cryotubes are transferred into a standard sample box, and are stored at -80°C or in liquid nitrogen.

(Transient gene introduction (transfection))

[0058]

1. Before starting transfection, petri dishes each containing the cells on the second day since the passage are provided as required, and the cells are checked whether they are healthy (normal) or not (see note 8). About 96 assays can be performed with one petri dish as an estimate.
2. The cells on the second day since the passage are peeled off according to the procedure for passage, and the suspended cells are transferred into a 50 mL centrifuge tube.
3. The centrifuge tube is centrifuged at room temperature and 1100 ×g for two minutes, and the supernatant is removed.
4. Cell clusters are completely dispersed in 10 mL DMEM (FBS is added such that the final concentration is 10%) .
5. The number of cells is counted with a hemocytometer and an inverted microscope. The cells are suspended in an appropriate amount of DMEM (FBS is added such that the final concentration is 10%) such that the number of cells is 1 to 2 × 10^5 cells/mL.
6. A 96-well plate is provided. 200 μL (2 to 4 × 10^4 cells/mL) per well of suspended cultured cells is placed into each well, and the plate is left to stand overnight in an incubator at 37°C under a 5% CO_2 condition. One well is used for one assay.
7. On the next day, the culture medium is carefully removed from each well, and is replaced with 100 μL of new DMEM (FBS is added such that the final concentration is 10%).
8. 400 ng effector plasmid (4 μL of 100 ng/μL) and 100 ng reporter plasmid (1 μL of 100 ng/μL) are placed into a single well on a new 96-well PCR plate (or a 0.2 mL tube).
9. For one assay, 1 μL HilyMAX is diluted with 10 μL serum-free DMEM.
10. 11 μL diluted solution is placed into each of the wells containing the plasmids. The solution is well mixed with the plasmids by pipetting.
11. The solution is left to stand at room temperature for 15 minutes. The total amount of the mixture is placed into the wells containing the cultured cells. The plate is left to stand in an incubator at 37°C under a 5% CO_2 condition for 24 hours.

(Luciferase assay)

[0059] The dual luciferase assay is performed using Dual-Glo Luciferase Assay System according to the usage instruction from the manufacturer except for a few modifications.

1. After 24 hours from the transfection, the culture medium of each well is replaced with 40 μL 1× PBS(-).
2. 40 μL of Dual-Glo luciferase reagent is placed into each well, and is well mixed with the culture medium by pipetting.
3. The mixture is left to stand at room temperature for 10 minutes, and the total amount thereof is transferred into a 96-well luminometer plate.
4. The light emission from firefly luciferase related with expression of FLuc gene is measured with a plate reader.
5. A Stop & Glo substrate is 100-fold diluted with a Dual-Glo Stop & Glo buffer. 40 μL of the diluted solution is added into each well.

6. The plate is left to stand at least at room temperature for 10 minutes, and then the light emission from renilla luciferase related with expression of RLuc gene is measured.

(Data analysis)

[0060]

1. The value of FLuc/RLuc is calculated to correct a difference in transfection efficiency between the assays or experimental errors.
2. An increase in activity of reporter gene expression is determined in the presence of the PPR-binding region and in the absence thereof by dividing an experimental value obtained using the plasmid according to the present invention (plasmid encoding a fusion protein of CRR4 and a translation activation domain eIF4G) by an experimental value obtained using a control plasmid (plasmid encoding a fusion protein of CRR4 and FLAG-tag).

Experimental results

[0061] The results of the luciferase assay are shown in Figure 3. As shown in Figure 3, 2.75-fold translation activity was specifically verified in the presence of both of the PPR-eIF4G and the PPR-binding sequence. That is, it is demonstrated that the fusion protein of the PPR protein and the functional domain which improve a protein expression level from an mRNA improves the protein expression level from the target mRNA.

Notes

[0062] (Note 1) HEK293T is a human fetus-derived kidney cell line which expresses an SV40 large T antigen. The cell line is readily cultured, and can be transfected with high efficiency by a variety of methods. HEK293T cells are available from RIKEN BRC (ja.brc.riken.jp) or ATCC (www.atcc.org).
[0063] (Note 2) 1× penicillin-streptomycin solution is added to DMEM to avoid contamination with microorganisms.
[0064] (Note 3) Before use, FBS is inactivated at 56°C for 30 minutes, and is stored at 4°C.
[0065] (Note 4) The purity of the plasmid is significantly important to the transfection efficiency. The plasmid should be isolated using a kit of a transfection grade.
[0066] (Note 5) A daily growth rate is an index indicating that the cells are healthy. To avoid suppression of cell growth, the cells should be always cultured in a sufficient space under a sufficient nutritional condition.
[0067] (Note 6) HEK293T cells should be gently treated when the culture medium is replaced because the cells are readily peeled off from the culturing petri dish.
[0068] (Note 7) The dedicated freeze container is a box whose freezing speed can be adjusted (about -1°C per minute at -80°C), and enables the cells to be freeze stored in a non-programmable -80°C freezer.
[0069] (Note 8) In transfection, cells are used at a culture density of 50 to 80%. However, an appropriate cell density depends on the transfection reagent. Additionally, the ratio of the transfection reagent (μL) to the plasmid DNA (μg) should be also optimized according to the usage instructions from the manufacturer. The procedure described herein is optimized for a condition where a 96-well plate, HEK293T cells, and HilyMAX as a transfection reagent are used.

Example 2: Improvement in protein expression level from target mRNA by fusion protein of PPR and another functional domain

[0070] In the case where useful substances are produced using cells, the amounts of protein synthesized by endogenous genes and exogenous genes should be precisely controlled. The final amount of the synthesized protein is determined by the insertion positions of genes, the mRNA transcription amount, post-transcriptional regulation (regulation at an RNA level), post-translational modification, and the like. For these reasons, the present inventors have devised a method of enhancing the translation of mRNAs taking advantage of the fact that a PPR protein sequence-specifically binds to a target RNA molecule (Figure 4). In the translation of mRNAs in eukaryotes, an mRNA undergoes action of a translation initiation factor (eukaryotic initiation factor; eIF). As a result, the ribosome is recruited near the translation starting point, and then the translation of the mRNA is started. In other words, the present inventors have considered that the translation of the mRNA can be artificially enhanced if the ribosome can only be recruited onto the mRNA. Moreover, the translation of an mRNA into a protein is usually performed in the ER. For this reason, the present inventors have considered that the translation of the mRNA can be enhanced by intentionally localizing the target mRNA into the ER.

Verification by experiment

[0071] To verify the idea above, a reporter assay system using animal cultured cells (HEK293T) was prepared (the experiment was performed by the same method as that in Example 1 except that different functional domains were used). The system was constructed using CRR4 protein (one of Arabidopsis thaliana PPR proteins), which is known to bind to a specific RNA sequence (UAUCUUGUCUUUA) (SEQ ID NO: 3). First, a fusion protein expression vector (effector plasmid) of CRR4 and a candidate protein functional domain was prepared. The selected candidate domains were (a) eIF proteins (eIF4E and eIF4G), (b) ribosome-bound proteins (DENR, MCT-1, TPT1, and Lerepo4), (c) translational regulation factors (SLBPs) of Histone which promote transport of the transcribed mRNA from the nucleus to the cytoplasm, (d) ER anchor proteins (SEC61B, TRAP-alpha, SR-alpha, Dial, and p180), (e) ER retention signal (KDEL), and (f) ER signal peptide. The fusion proteins were cloned so as to express in the form of HA-CRR4-XX or XX-CRR4-HA (HA: epitope tag (SEQ ID NO: 4); XX: candidate domain).

[0072] The reporter plasmid included an expression cassette where renilla luciferase (RLuc) and firefly luciferase (FLuc) are transcribed in the form of a dicistronic mRNA under the control of a CMV promoter. Three PPR-binding sequences (UAUCUUGUCUUUA) (SEQ ID NO: 3) are inserted into a site on the 5'-end of Fluc.

[0073] The effector plasmid and the reporter plasmid were transfected into HEK293T cells, and the intensities of light emission from RLUC and FLUC were measured. The intensity of light emission from RLUC was treated as a transfection control, and the value of the intensity of light emission from FLUC/the intensity of light emission from RLUC was treated as a translation activity amount.

Results

[0074] The results shown in Figures 5 and 6 were examined using the following indices (A) and (B).

(A) Comparison between the absence of and the presence of the target

[0075] The comparison shows an amount of sequence-specific change in translation.

(B) Comparison to the presence of the target and the absence of the effector (empty) (black dashed line)

[0076] The comparison shows an amount of change in translation caused by addition of the domain.

1. eIF4E was fused to the C-terminal side of CRR4.

(A) 2.7 times
(B) 1.6 times

2. eIF4G was fused to the C-terminal side of CRR4.

(A) 4.5 times
(B) 3.3 times

3. DENR was fused to the N-terminal side of CRR4.

(A) 1.7 times
(B) 1.3 times

4. DENR was fused to the C-terminal side of CRR4.

(A) 2.4 times
(B) 1.7 times

5. MCT-1 was fused to the N-terminal side of CRR4.

(A) 1.3 times
(B) 1.0 time

6. MCT-1 was fused to the C-terminal side of CRR4.

(A) 2.0 times
(B) 1.2 times

7. TPT-1 was fused to the N-terminal side of CRR4.

(A) 1.4 times
(B) 1.0 time

8. TPT-1 was fused to the C-terminal side of CRR4.

(A) 2.4 times
(B) 1.9 times

9. Lerepo4 was fused to the N-terminal side of CRR4.

(A) 3.0 times
(B) 1.8 times

10. Lerepo4 was fused to the C-terminal side of CRR4.

(A) 3.3 times
(B) 2.6 times

11. SLBP was fused to the C-terminal side of CRR4.

(A) 4.1 times
(B) 3.3 times

12. Sec61B was fused to the C-terminal side of CRR4.

(A) 1.6 times
(B) 1.6 times

13. Sec61BTM was fused to the C-terminal side of CRR4.

(A) 2.4 times
(B) 1.9 times

14. TRAP-alpha was fused to the C-terminal side of CRR4.

(A) 3.5 times
(B) 4.5 times

15. TRAPTM was fused to the C-terminal side of CRR4.

(A) 2.3 times
(B) 1.6 times

16. SR-alpha was fused to the N-terminal side of CRR4.

(A) 1.7 times
(B) 1.5 times

17. DialTM was fused to the N-terminal side of CRR4.

(A) 1.8 times
(B) 1.2 times

18. P180TM2R was fused to the N-terminal side of CRR4.

(A) 2.1 times
(B) 1.5 times

19. P180TMH was fused to the N-terminal side of CRR4.

(A) 2.3 times
(B) 2.5 times

20. P180TM2 was fused to the N-terminal side of CRR4.

(A) 3.0 times
(B) 2.1 times

21. KDEL was fused to the C-terminal side of CRR4.

(A) 1.8 times
(B) 1.4 times

22. KEEL was fused to the C-terminal side of CRR4.

(A) 2.3 times
(B) 2.1 times

23. Signal peptide (SP) was fused to the N-terminal side of CRR4.

(A) 1.4 times
(B) 2.0 times

[0077] As shown above, an increase in translation was found in all the functional domains in both of the indices (A) and the targets (B). Namely, it was clearly shown that the fusion protein according to the present invention can enhance the translation of the target mRNA.

[0078] The amino acid sequences of the functional domains used in Examples are listed below:

[Table 1-1]

Domain	Sequence
eIF4E	MATVEPETTPTPNPPTTEEEKTESNQEVANPEHYIKHPLQNRWALW FFKNDKSKTWQANLRLISKFDTVEDFWALYNHIQLSSNLMPGCDYS LFKDGIEPMLEDEKNKRGGRWLITLNKQQRRSDLDRFWLETLLCLI GESFDDYSDDVCGAVVNVRAKGDKIAIWTTECENREAVTHIGRVYK ERLGLPPKIVIGYQSHADTATKSGSTTKNRFVVGRY (SEQ ID NO: 5)

(continued)

Domain	Sequence
eIF4G	EEKKRYDREFLLGFQFIFASMQKPEGLPHISDVVLDKANKTPLRPL DPTRLQGINCGPDFTPSFANLGRTTLSTRGPPRGGPGGELPRGPQA GLGPRRSQQGPRKEPRKIIATVLMTEDIKLNKAEKAWKPSSKRTAA DKDRGEEDADGSKTQDLFRRVRSILNKLTPQMFQQLMKQVTQLAID TEERLKGVIDLIFEKAISEPNFSVAYANMCRCLMALKVPTTEKPTV TVNFRKLLLNRCQKEFEKDKDDDEVFEKKQKEMDEAATAEERGRLK EELEEARDIARRRSLGNIKFIGELFKLKMLTEAIMHDCVVKLLKNH DEESLECLCRLLTTIGKDLDFEKAKPRMDQYFNQMEKIIKEKKTSS RIRFMLQDVLDLRGSNWVPRRGDQGPKTIDQIHKEAEMEEHREHIK VQQLMAKGSDKRRGGPPGPPISRGLPLVDDGGWNTVPISKGSRPID TSRLTKITKPGSIDSNNQLFAPGGRLSWGKGSSGGSGAKPSDAASE AARPATSTLNRFSALQQAVPTESTDNRRVVQRSSLSRERGEKAGDR GDRLERSERGGDRGDRLDRARTPATKRSFSKEVEERSRERPSQPEG LRKAASLTEDRDRGRDAVKREAALPPVSPLKAALSEEELEKKSKAI IEEYLHLNDMKEAVQCVQELASPSLLFIFVRHGVESTLERSAIARE HMGQLLHQLLCAGHLSTAQYYQGLYEILELAEDMEIDIPHVWLYLA ELVTPILQEGGVPMGELFREITKPLRPLGKAASLLLEILGLLCKSM GPKKVGTLWREAGLSWKEFLPEGQDIGAFVAEQKVEYTLGEESEAP GQRALPSEELNRQLEKLLKEGSSNQRVFDWIEANLSEQQIVSNTLV RALMTAVCYSAIIFETPLRVDVAVLKARAKLLQKYLCDEQKELQAL YALQALVVTLEQPPNLLRMFFDALYDEDVVKEDAFYSWESSKDPAE QQGKGVALKSVTAFFKWLREAEEESDH (SEQ ID NO: 1)
DENR	MAADISESSGADCKGDPRNSAKLDADYPLRVLYCGVCSLPTEYCEY MPDVAKCRQWLEKNFPNEFAKLTVENSPKQEAGISEGQGTAGEEEE KKKQKRGGRGQIKQKKKTVPQKVTIAKIPRAKKKYVTRVCGLATFE IDLKEAQRFFAQKFSCGASVTGEDEIIIQGDFTDDIIDVIQEKWPE VDDDSIEDLGEVKK (SEQ ID NO: 6)
MCT-1	MFKKFDEKENVSNCIQLKTSVIKGIKNQLIEQFPGIEPWLNQIMPK KDPVKIVRCHEHIEILTVNGELLFFRQREGPFYPTLRLLHKYPFIL PHQQVDKGAIKFVLSGANIMCPGLTSPGAKLYPAAVDTIVAIMAEG

[Table 1-2]

	KQHALCVGVMKMSAEDIEKVNKGIGIENIHYLNDGLWHMKTYK (SEQ IND NO: 7)
TPT-1	MIIYRDLISHDEMFSDIYKIREIADGLCLEVEGKMVSRTEGNIDDS LIGGNASAEGPEGEGTESTVITGVDIVMNHHLQETSFTKEAYKKYI KDYMKSIKGKLEEQRPERVKPFMTGAAEQIKHILANFKNYQFFIGE NMNPDGMVALLDYREDGVTPYMIFFKDGLEMEKC (SEQ ID NO: 8)

(continued)

Leropo4	**PPKKQAQAGGSKKAEQKKKEKIIEDKTFGLKNKKGAKQQKFIKAVT HQVKFGQQNPRQVAQSEAEKKLKKDDKKKELQELNELFKPVVAAQK ISKGADPKSVVCAFFKQGQCTKGDKCKFSHDLTLERKCEKRSVYID ARDEELEKDTMDNWDEKKLEEVVNKKHGEAEKKKPKTQIVCKHFLE AIENNKYGWFWVCPGGGDICMYRHALPPGFVLKKDKKKEEKEDEIS LEDLIERERSALGPNVTKITLESFLAWKKRKRQEKIDKLEQDMERR KADFKAGKALVISGREVFEFRPELVNDDDEEADDTRYTQGTGGDEV DDSVSVNDIDLSLYIPRDVDETGITVASLERFSTYTSDKDENKLSE ASGGRAENGERSDLEEDNEREGTENGAIDAVPVDEKSFHWRGFG** (SEQ ID NO: 9)
SLBP	**ACRPRSPPRHQSRCDGDASPPSPARWSLGRKRRADGRRWRPEDAEE AEHRGAERRPESFTTPEGPKPRSRCSDWASAVEEDEMRTRVNKEMA RYKRKLLINDFGRERKSSSGSSDSKESMSTVPADFETDESVLMRRQ KQINYGKNTIAYDRYIKEVPRHLRQPGIHPKTPNKFKKYSRRSWDQ QIKLWKVALHFWDPPAEEGCDLQEIHPVDLESAESSSEPQTSSQDD FDVYSGTPTKVRHMDSQVEDEFDLEACLTEPLRDFSAMS** (SEQ ID NO: 10)
Sec61B	**PGPTPSGTNVGSSGRSPSKAVAARAAGSTVRQRKNASCGTRSAGRT TSAGTGGMWRFYTEDSPGLKVGPVPVLVMSLLFIASVFMLHIWGKY TRS** (SEQ ID NO: 11)
sec61B-TM	**VGPVPVLVMSLLFIASVFMLHIW** (SEQ IND NO: 12)
TRAP-alpha	**RLLPRLLLLLLLVFPATVLFRGGPRGLLAVAQDLTEDEETVEDSII EDEDDEAEVEEDEPTDLVEDKEEEDVSGEPEASPSADTTILFVKGE DFPANNIVKFLVGFTNKGTEDFIVESLDASFRYPQDYQFYIQNFTA LPLNTVVPPQRQATFEYSFIPAEPMGGRPFGLVINLNYKDLNGNVF QDAVFNQTVTVIEREDGLDGETIFMYMFLAGLGLLVIVGLHQLLES RKRKRPIQKVEMGTSSQNDVDMSWIPQETLNQINKASPRRLPRKRA QKRSVGSDE** (SEQ ID NO: 13)
TRAP-TM	**TIFMYMFLAGLGLLVIVGLHQLL** (SEQ IND NO: 14)
SR-alpha	**LDFFTIFSKGGLVLWCFQGVSDSCTGPVNALIRSVLLQVGFQKILT**

[Table 1-3]

	LTYVDKLIDDVHRLFRDKYRTEIQQQSALSLLNGTFDFQNDFLRLL REAEESSKIRAPTTMKKFEDSEKAKKPVRSMIETRGEKPKEKAKNS KKKGAKKEGSDGPLATSKPVPAEKSGLPVGPENGVELSKEELIRRK REEFIQKHGRGMEKSNKSTKSDAPKEKGKKAPRVWELGGCANKEVL DYSTPTTNGTPEAALSEDINLIRGTGSGGQLQDLDCSSSDDEGAAQ NSTKPSATKGTLGGMFGMLKGLVGSKSLSREDMESVLDKMRDHLIA KNVAADIAVQLCESVANKLEGKVMGTFSTVTSTVKQALQESLVQIL QPQRRVDMLRDIMDAQRRQRPYVVTFCGVNGVGKSTNLAKISFWLL ENGFSVLIAACDTFRAGAVEQLRTHTRRLSALHPPEKHGGRTMVQL FEKGYGKDAAGIAMEAIAFARNQGFDVVLVDTAGRMQDNAPLMTAL AKLITVNTPDLVLFVGEALVGNEAVDQLVKFNRALADHSMAQTPRL IDGIVLTKFDTIDDKVGAAISMTYITSKPIVFVGTGQTYCDLRSLN AKAVVAALMKA (SEQ ID NO: 15)
DiaTM	**STLGHMVLFPVWFLYSLL** (SEQ ID NO: 16)
P180TMR2	**DIYDTQTLGVVVFGGFMVVSAIGIFLVSTFSMKETSYEEALANQRK EMAKTHHQKVEKKKKEKTVEKKGKTKKKEEKPNGKIPDHDPAPNVT VLLREPVRAPAVAVAPTPVQPPIIVAPVATVPAMPQEKLASSPKDK KKKEKKVAKVEPAVSSVVNSIQVLTSKAAILETAPKEGRNTDVAQS PEAPKQEAPAKKKSGSKKKGPPDADGPLYLPYKTLVSTVGSMVFNE GEAQRLIEILSEKAGIIQDTWHKATQKGDPV**(SEQ ID NO: 17)
P180TMH	**LGVVVFGGFMVVSAIGIFLVSTF** (SEQ ID NO: 18)
P180TM2	**DIYDTQTLGVVVFGGFMVVSAIGIFLVSTF** (SEQ ID NO: 19)
KDEL	**KDEL** (SEQ ID NO: 20)
KEEL	**KEEL** (SEQ ID NO: 21)
ER signal peptide	**MGWSCIILFLVATATGAHS** (SEQ ID NO: 22)

SEQUENCE LISTING

[0079]

<110> KYUSHU UNIVERSITY, NATIONAL UNIVERSITY CORPORATION

<120> **Fusion Protein For Increasing Protein Expression From A Target mRNA**

<130> EDFP1601F

<150> US62/345252
<151> 2016-06-03

<150> JP2016-120524
<151> 2016-06-17

<160> 22

<170> PatentIn version 3.5

<210> 1
<211> 993
<212> PRT
<213> Homo sapiens

<400> 1

```
Glu Glu Lys Lys Arg Tyr Asp Arg Glu Phe Leu Leu Gly Phe Gln Phe
1               5                  10                  15

Ile Phe Ala Ser Met Gln Lys Pro Glu Gly Leu Pro His Ile Ser Asp
            20                  25                  30

Val Val Leu Asp Lys Ala Asn Lys Thr Pro Leu Arg Pro Leu Asp Pro
        35                  40                  45

Thr Arg Leu Gln Gly Ile Asn Cys Gly Pro Asp Phe Thr Pro Ser Phe
    50                  55                  60

Ala Asn Leu Gly Arg Thr Thr Leu Ser Thr Arg Gly Pro Pro Arg Gly
65                  70                  75                  80

Gly Pro Gly Gly Glu Leu Pro Arg Gly Pro Gln Ala Gly Leu Gly Pro
                85                  90                  95

Arg Arg Ser Gln Gln Gly Pro Arg Lys Glu Pro Arg Lys Ile Ile Ala
            100                 105                 110

Thr Val Leu Met Thr Glu Asp Ile Lys Leu Asn Lys Ala Glu Lys Ala
        115                 120                 125

Trp Lys Pro Ser Ser Lys Arg Thr Ala Ala Asp Lys Asp Arg Gly Glu
    130                 135                 140

Glu Asp Ala Asp Gly Ser Lys Thr Gln Asp Leu Phe Arg Arg Val Arg
```

145 150 155 160

Ser Ile Leu Asn Lys Leu Thr Pro Gln Met Phe Gln Gln Leu Met Lys
165 170 175

Gln Val Thr Gln Leu Ala Ile Asp Thr Glu Glu Arg Leu Lys Gly Val
180 185 190

Ile Asp Leu Ile Phe Glu Lys Ala Ile Ser Glu Pro Asn Phe Ser Val
195 200 205

Ala Tyr Ala Asn Met Cys Arg Cys Leu Met Ala Leu Lys Val Pro Thr
210 215 220

Thr Glu Lys Pro Thr Val Thr Val Asn Phe Arg Lys Leu Leu Leu Asn
225 230 235 240

Arg Cys Gln Lys Glu Phe Glu Lys Asp Lys Asp Asp Asp Glu Val Phe
245 250 255

Glu Lys Lys Gln Lys Glu Met Asp Glu Ala Ala Thr Ala Glu Glu Arg
260 265 270

Gly Arg Leu Lys Glu Glu Leu Glu Glu Ala Arg Asp Ile Ala Arg Arg
275 280 285

Arg Ser Leu Gly Asn Ile Lys Phe Ile Gly Glu Leu Phe Lys Leu Lys
290 295 300

Met Leu Thr Glu Ala Ile Met His Asp Cys Val Val Lys Leu Leu Lys
305 310 315 320

Asn His Asp Glu Glu Ser Leu Glu Cys Leu Cys Arg Leu Leu Thr Thr
325 330 335

Ile Gly Lys Asp Leu Asp Phe Glu Lys Ala Lys Pro Arg Met Asp Gln
340 345 350

Tyr Phe Asn Gln Met Glu Lys Ile Ile Lys Glu Lys Lys Thr Ser Ser
355 360 365

Arg Ile Arg Phe Met Leu Gln Asp Val Leu Asp Leu Arg Gly Ser Asn
370 375 380

Trp Val Pro Arg Arg Gly Asp Gln Gly Pro Lys Thr Ile Asp Gln Ile
385 390 395 400

```
His Lys Glu Ala Glu Met Glu Glu His Arg Glu His Ile Lys Val Gln
                405                 410                 415

Gln Leu Met Ala Lys Gly Ser Asp Lys Arg Arg Gly Gly Pro Pro Gly
            420                 425                 430

Pro Pro Ile Ser Arg Gly Leu Pro Leu Val Asp Asp Gly Gly Trp Asn
        435                 440                 445

Thr Val Pro Ile Ser Lys Gly Ser Arg Pro Ile Asp Thr Ser Arg Leu
    450                 455                 460

Thr Lys Ile Thr Lys Pro Gly Ser Ile Asp Ser Asn Asn Gln Leu Phe
465                 470                 475                 480

Ala Pro Gly Gly Arg Leu Ser Trp Gly Lys Gly Ser Ser Gly Gly Ser
                485                 490                 495

Gly Ala Lys Pro Ser Asp Ala Ala Ser Glu Ala Ala Arg Pro Ala Thr
            500                 505                 510

Ser Thr Leu Asn Arg Phe Ser Ala Leu Gln Gln Ala Val Pro Thr Glu
        515                 520                 525

Ser Thr Asp Asn Arg Arg Val Val Gln Arg Ser Ser Leu Ser Arg Glu
    530                 535                 540

Arg Gly Glu Lys Ala Gly Asp Arg Gly Asp Arg Leu Glu Arg Ser Glu
545                 550                 555                 560

Arg Gly Gly Asp Arg Gly Asp Arg Leu Asp Arg Ala Arg Thr Pro Ala
                565                 570                 575

Thr Lys Arg Ser Phe Ser Lys Glu Val Glu Glu Arg Ser Arg Glu Arg
            580                 585                 590

Pro Ser Gln Pro Glu Gly Leu Arg Lys Ala Ala Ser Leu Thr Glu Asp
        595                 600                 605

Arg Asp Arg Gly Arg Asp Ala Val Lys Arg Glu Ala Ala Leu Pro Pro
    610                 615                 620

Val Ser Pro Leu Lys Ala Ala Leu Ser Glu Glu Glu Leu Glu Lys Lys
625                 630                 635                 640

Ser Lys Ala Ile Ile Glu Glu Tyr Leu His Leu Asn Asp Met Lys Glu
                645                 650                 655
```

Ala Val Gln Cys Val Gln Glu Leu Ala Ser Pro Ser Leu Leu Phe Ile
660 665 670

Phe Val Arg His Gly Val Glu Ser Thr Leu Glu Arg Ser Ala Ile Ala
675 680 685

Arg Glu His Met Gly Gln Leu Leu His Gln Leu Leu Cys Ala Gly His
690 695 700

Leu Ser Thr Ala Gln Tyr Tyr Gln Gly Leu Tyr Glu Ile Leu Glu Leu
705 710 715 720

Ala Glu Asp Met Glu Ile Asp Ile Pro His Val Trp Leu Tyr Leu Ala
725 730 735

Glu Leu Val Thr Pro Ile Leu Gln Glu Gly Gly Val Pro Met Gly Glu
740 745 750

Leu Phe Arg Glu Ile Thr Lys Pro Leu Arg Pro Leu Gly Lys Ala Ala
755 760 765

Ser Leu Leu Leu Glu Ile Leu Gly Leu Leu Cys Lys Ser Met Gly Pro
770 775 780

Lys Lys Val Gly Thr Leu Trp Arg Glu Ala Gly Leu Ser Trp Lys Glu
785 790 795 800

Phe Leu Pro Glu Gly Gln Asp Ile Gly Ala Phe Val Ala Glu Gln Lys
805 810 815

Val Glu Tyr Thr Leu Gly Glu Glu Ser Glu Ala Pro Gly Gln Arg Ala
820 825 830

Leu Pro Ser Glu Glu Leu Asn Arg Gln Leu Glu Lys Leu Leu Lys Glu
835 840 845

Gly Ser Ser Asn Gln Arg Val Phe Asp Trp Ile Glu Ala Asn Leu Ser
850 855 860

Glu Gln Gln Ile Val Ser Asn Thr Leu Val Arg Ala Leu Met Thr Ala
865 870 875 880

Val Cys Tyr Ser Ala Ile Ile Phe Glu Thr Pro Leu Arg Val Asp Val
885 890 895

Ala Val Leu Lys Ala Arg Ala Lys Leu Leu Gln Lys Tyr Leu Cys Asp
900 905 910

Glu Gln Lys Glu Leu Gln Ala Leu Tyr Ala Leu Gln Ala Leu Val Val
915 920 925

Thr Leu Glu Gln Pro Pro Asn Leu Leu Arg Met Phe Phe Asp Ala Leu
930 935 940

Tyr Asp Glu Asp Val Val Lys Glu Asp Ala Phe Tyr Ser Trp Glu Ser
945 950 955 960

Ser Lys Asp Pro Ala Glu Gln Gln Gly Lys Gly Val Ala Leu Lys Ser
965 970 975

Val Thr Ala Phe Phe Lys Trp Leu Arg Glu Ala Glu Glu Glu Ser Asp
980 985 990

<210> 2
<211> 561
<212> PRT
<213> Arabidopsis thaliana

<400> 2

Ala Phe Ala Ser Ser Arg Arg Pro Tyr Leu Ala Asp Phe Ala Arg Cys
1 5 10 15

Val Phe His Glu Tyr His Val Cys Ser Phe Ser Phe Gly Glu Val Glu
20 25 30

Asp Pro Phe Leu Trp Asn Ala Val Ile Lys Ser His Ser His Gly Lys
35 40 45

Asp Pro Arg Gln Ala Leu Leu Leu Leu Cys Leu Met Leu Glu Asn Gly
50 55 60

Val Ser Val Asp Lys Phe Ser Leu Ser Leu Val Leu Lys Ala Cys Ser
65 70 75 80

Arg Leu Gly Phe Val Lys Gly Gly Met Gln Ile His Gly Phe Leu Lys
85 90 95

Lys Thr Gly Leu Trp Ser Asp Leu Phe Leu Gln Asn Cys Leu Ile Gly
100 105 110

Leu Tyr Leu Lys Cys Gly Cys Leu Gly Leu Ser Arg Gln Met Phe Asp
115 120 125

Arg Met Pro Lys Arg Asp Ser Val Ser Tyr Asn Ser Met Ile Asp Gly
130 135 140

Tyr Val Lys Cys Gly Leu Ile Val Ser Ala Arg Glu Leu Phe Asp Leu
145 150 155 160

Met Pro Met Glu Met Lys Asn Leu Ile Ser Trp Asn Ser Met Ile Ser
165 170 175

Gly Tyr Ala Gln Thr Ser Asp Gly Val Asp Ile Ala Ser Lys Leu Phe
180 185 190

Ala Asp Met Pro Glu Lys Asp Leu Ile Ser Trp Asn Ser Met Ile Asp
195 200 205

Gly Tyr Val Lys His Gly Arg Ile Glu Asp Ala Lys Gly Leu Phe Asp
210 215 220

Val Met Pro Arg Arg Asp Val Val Thr Trp Ala Thr Met Ile Asp Gly
225 230 235 240

Tyr Ala Lys Leu Gly Phe Val His His Ala Lys Thr Leu Phe Asp Gln
245 250 255

Met Pro His Arg Asp Val Val Ala Tyr Asn Ser Met Met Ala Gly Tyr
260 265 270

Val Gln Asn Lys Tyr His Met Glu Ala Leu Glu Ile Phe Ser Asp Met
275 280 285

Glu Lys Glu Ser His Leu Leu Pro Asp Asp Thr Thr Leu Val Ile Val
290 295 300

Leu Pro Ala Ile Ala Gln Leu Gly Arg Leu Ser Lys Ala Ile Asp Met
305 310 315 320

His Leu Tyr Ile Val Glu Lys Gln Phe Tyr Leu Gly Gly Lys Leu Gly
325 330 335

Val Ala Leu Ile Asp Met Tyr Ser Lys Cys Gly Ser Ile Gln His Ala
340 345 350

Met Leu Val Phe Glu Gly Ile Glu Asn Lys Ser Ile Asp His Trp Asn
355 360 365

Ala Met Ile Gly Gly Leu Ala Ile His Gly Leu Gly Glu Ser Ala Phe

370 375 380

Asp Met Leu Leu Gln Ile Glu Arg Leu Ser Leu Lys Pro Asp Asp Ile
385 390 395 400

Thr Phe Val Gly Val Leu Asn Ala Cys Ser His Ser Gly Leu Val Lys
405 410 415

Glu Gly Leu Leu Cys Phe Glu Leu Met Arg Arg Lys His Lys Ile Glu
420 425 430

Pro Arg Leu Gln His Tyr Gly Cys Met Val Asp Ile Leu Ser Arg Ser
435 440 445

Gly Ser Ile Glu Leu Ala Lys Asn Leu Ile Glu Glu Met Pro Val Glu
450 455 460

Pro Asn Asp Val Ile Trp Arg Thr Phe Leu Thr Ala Cys Ser His His
465 470 475 480

Lys Glu Phe Glu Thr Gly Glu Leu Val Ala Lys His Leu Ile Leu Gln
485 490 495

Ala Gly Tyr Asn Pro Ser Ser Tyr Val Leu Leu Ser Asn Met Tyr Ala
500 505 510

Ser Phe Gly Met Trp Lys Asp Val Arg Arg Val Arg Thr Met Met Lys
515 520 525

Glu Arg Lys Ile Glu Lys Ile Pro Gly Cys Ser Trp Ile Glu Leu Asp
530 535 540

Gly Arg Val His Glu Phe Phe Val Asp Ser Ile Glu Val Ser Ser Thr
545 550 555 560

Leu

<210> 3
<211> 13
<212> RNA
<213> Artificial Sequence

<220>
<223> Synthetic RNA

<400> 3 uaucuugucu uua 13

<210> 4
<211> 12

<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic peptide

<400> 4

Met Ala Gly Tyr Pro Tyr Asp Val Pro Asp Tyr Ala
1 5 10

<210> 5
<211> 220
<212> PRT
<213> Homo sapiens

<400> 5

```
Met Ala Thr Val Glu Pro Glu Thr Thr Pro Thr Pro Asn Pro Pro Thr
1               5                  10                  15

Thr Glu Glu Glu Lys Thr Glu Ser Asn Gln Glu Val Ala Asn Pro Glu
            20                  25                  30

His Tyr Ile Lys His Pro Leu Gln Asn Arg Trp Ala Leu Trp Phe Phe
        35                  40                  45

Lys Asn Asp Lys Ser Lys Thr Trp Gln Ala Asn Leu Arg Leu Ile Ser
    50                  55                  60

Lys Phe Asp Thr Val Glu Asp Phe Trp Ala Leu Tyr Asn His Ile Gln
65                  70                  75                  80

Leu Ser Ser Asn Leu Met Pro Gly Cys Asp Tyr Ser Leu Phe Lys Asp
                85                  90                  95

Gly Ile Glu Pro Met Leu Glu Asp Glu Lys Asn Lys Arg Gly Gly Arg
            100                 105                 110

Trp Leu Ile Thr Leu Asn Lys Gln Gln Arg Arg Ser Asp Leu Asp Arg
        115                 120                 125

Phe Trp Leu Glu Thr Leu Leu Cys Leu Ile Gly Glu Ser Phe Asp Asp
    130                 135                 140

Tyr Ser Asp Asp Val Cys Gly Ala Val Val Asn Val Arg Ala Lys Gly
145                 150                 155                 160

Asp Lys Ile Ala Ile Trp Thr Thr Glu Cys Glu Asn Arg Glu Ala Val
                165                 170                 175

Thr His Ile Gly Arg Val Tyr Lys Glu Arg Leu Gly Leu Pro Pro Lys
            180                 185                 190

Ile Val Ile Gly Tyr Gln Ser His Ala Asp Thr Ala Thr Lys Ser Gly
        195                 200                 205

Ser Thr Thr Lys Asn Arg Phe Val Val Gly Arg Tyr
    210                 215                 220
```

<210> 6
<211> 198
<212> PRT
<213> Homo sapiens

<400> 6

Met Ala Ala Asp Ile Ser Glu Ser Ser Gly Ala Asp Cys Lys Gly Asp
1 5 10 15

Pro Arg Asn Ser Ala Lys Leu Asp Ala Asp Tyr Pro Leu Arg Val Leu
20 25 30

Tyr Cys Gly Val Cys Ser Leu Pro Thr Glu Tyr Cys Glu Tyr Met Pro
35 40 45

Asp Val Ala Lys Cys Arg Gln Trp Leu Glu Lys Asn Phe Pro Asn Glu
50 55 60

Phe Ala Lys Leu Thr Val Glu Asn Ser Pro Lys Gln Glu Ala Gly Ile
65 70 75 80

Ser Glu Gly Gln Gly Thr Ala Gly Glu Glu Glu Glu Lys Lys Lys Gln
85 90 95

Lys Arg Gly Gly Arg Gly Gln Ile Lys Gln Lys Lys Lys Thr Val Pro
100 105 110

Gln Lys Val Thr Ile Ala Lys Ile Pro Arg Ala Lys Lys Lys Tyr Val
115 120 125

Thr Arg Val Cys Gly Leu Ala Thr Phe Glu Ile Asp Leu Lys Glu Ala
130 135 140

Gln Arg Phe Phe Ala Gln Lys Phe Ser Cys Gly Ala Ser Val Thr Gly
145 150 155 160

Glu Asp Glu Ile Ile Ile Gln Gly Asp Phe Thr Asp Asp Ile Ile Asp
165 170 175

Val Ile Gln Glu Lys Trp Pro Glu Val Asp Asp Asp Ser Ile Glu Asp
180 185 190

Leu Gly Glu Val Lys Lys
195

<210> 7
<211> 181
<212> PRT
<213> Homo sapiens

<400> 7

Met Phe Lys Lys Phe Asp Glu Lys Glu Asn Val Ser Asn Cys Ile Gln
1 5 10 15

Leu Lys Thr Ser Val Ile Lys Gly Ile Lys Asn Gln Leu Ile Glu Gln
20 25 30

Phe Pro Gly Ile Glu Pro Trp Leu Asn Gln Ile Met Pro Lys Lys Asp
35 40 45

Pro Val Lys Ile Val Arg Cys His Glu His Ile Glu Ile Leu Thr Val
50 55 60

Asn Gly Glu Leu Leu Phe Phe Arg Gln Arg Glu Gly Pro Phe Tyr Pro
65 70 75 80

Thr Leu Arg Leu Leu His Lys Tyr Pro Phe Ile Leu Pro His Gln Gln
85 90 95

Val Asp Lys Gly Ala Ile Lys Phe Val Leu Ser Gly Ala Asn Ile Met
100 105 110

Cys Pro Gly Leu Thr Ser Pro Gly Ala Lys Leu Tyr Pro Ala Ala Val
115 120 125

Asp Thr Ile Val Ala Ile Met Ala Glu Gly Lys Gln His Ala Leu Cys
130 135 140

Val Gly Val Met Lys Met Ser Ala Glu Asp Ile Glu Lys Val Asn Lys
145 150 155 160

Gly Ile Gly Ile Glu Asn Ile His Tyr Leu Asn Asp Gly Leu Trp His
165 170 175

Met Lys Thr Tyr Lys
180

<210> 8
<211> 172
<212> PRT
<213> Homo sapiens

<400> 8

```
Met Ile Ile Tyr Arg Asp Leu Ile Ser His Asp Glu Met Phe Ser Asp
1               5                  10                  15

Ile Tyr Lys Ile Arg Glu Ile Ala Asp Gly Leu Cys Leu Glu Val Glu
            20                  25                  30

Gly Lys Met Val Ser Arg Thr Glu Gly Asn Ile Asp Asp Ser Leu Ile
        35                  40                  45

Gly Gly Asn Ala Ser Ala Glu Gly Pro Glu Gly Glu Gly Thr Glu Ser
    50                  55                  60

Thr Val Ile Thr Gly Val Asp Ile Val Met Asn His His Leu Gln Glu
65                  70                  75                  80

Thr Ser Phe Thr Lys Glu Ala Tyr Lys Lys Tyr Ile Lys Asp Tyr Met
                85                  90                  95

Lys Ser Ile Lys Gly Lys Leu Glu Glu Gln Arg Pro Glu Arg Val Lys
            100                 105                 110

Pro Phe Met Thr Gly Ala Ala Glu Gln Ile Lys His Ile Leu Ala Asn
        115                 120                 125

Phe Lys Asn Tyr Gln Phe Phe Ile Gly Glu Asn Met Asn Pro Asp Gly
    130                 135                 140

Met Val Ala Leu Leu Asp Tyr Arg Glu Asp Gly Val Thr Pro Tyr Met
145                 150                 155                 160

Ile Phe Phe Lys Asp Gly Leu Glu Met Glu Lys Cys
                165                 170
```

<210> 9
<211> 412
<212> PRT
<213> Homo sapiens

<400> 9

```
Pro Pro Lys Lys Gln Ala Gln Ala Gly Gly Ser Lys Lys Ala Glu Gln
1               5                  10                  15
```

Lys Lys Lys Glu Lys Ile Ile Glu Asp Lys Thr Phe Gly Leu Lys Asn
20 25 30

Lys Lys Gly Ala Lys Gln Gln Lys Phe Ile Lys Ala Val Thr His Gln
35 40 45

Val Lys Phe Gly Gln Gln Asn Pro Arg Gln Val Ala Gln Ser Glu Ala
50 55 60

Glu Lys Lys Leu Lys Lys Asp Asp Lys Lys Lys Glu Leu Gln Glu Leu
65 70 75 80

Asn Glu Leu Phe Lys Pro Val Val Ala Ala Gln Lys Ile Ser Lys Gly
85 90 95

Ala Asp Pro Lys Ser Val Val Cys Ala Phe Phe Lys Gln Gly Gln Cys
100 105 110

Thr Lys Gly Asp Lys Cys Lys Phe Ser His Asp Leu Thr Leu Glu Arg
115 120 125

Lys Cys Glu Lys Arg Ser Val Tyr Ile Asp Ala Arg Asp Glu Glu Leu
130 135 140

Glu Lys Asp Thr Met Asp Asn Trp Asp Glu Lys Lys Leu Glu Glu Val
145 150 155 160

Val Asn Lys Lys His Gly Glu Ala Glu Lys Lys Lys Pro Lys Thr Gln
165 170 175

Ile Val Cys Lys His Phe Leu Glu Ala Ile Glu Asn Asn Lys Tyr Gly
180 185 190

Trp Phe Trp Val Cys Pro Gly Gly Gly Asp Ile Cys Met Tyr Arg His
195 200 205

Ala Leu Pro Pro Gly Phe Val Leu Lys Lys Asp Lys Lys Lys Glu Glu
210 215 220

Lys Glu Asp Glu Ile Ser Leu Glu Asp Leu Ile Glu Arg Glu Arg Ser
225 230 235 240

Ala Leu Gly Pro Asn Val Thr Lys Ile Thr Leu Glu Ser Phe Leu Ala
245 250 255

Trp Lys Lys Arg Lys Arg Gln Glu Lys Ile Asp Lys Leu Glu Gln Asp
260 265 270

Met Glu Arg Arg Lys Ala Asp Phe Lys Ala Gly Lys Ala Leu Val Ile
275 280 285

Ser Gly Arg Glu Val Phe Glu Phe Arg Pro Glu Leu Val Asn Asp Asp
290 295 300

Asp Glu Glu Ala Asp Asp Thr Arg Tyr Thr Gln Gly Thr Gly Gly Asp
305 310 315 320

Glu Val Asp Asp Ser Val Ser Val Asn Asp Ile Asp Leu Ser Leu Tyr
325 330 335

Ile Pro Arg Asp Val Asp Glu Thr Gly Ile Thr Val Ala Ser Leu Glu
340 345 350

Arg Phe Ser Thr Tyr Thr Ser Asp Lys Asp Glu Asn Lys Leu Ser Glu
355 360 365

Ala Ser Gly Gly Arg Ala Glu Asn Gly Glu Arg Ser Asp Leu Glu Glu
370 375 380

Asp Asn Glu Arg Glu Gly Thr Glu Asn Gly Ala Ile Asp Ala Val Pro
385 390 395 400

Val Asp Glu Lys Ser Phe His Trp Arg Gly Phe Gly
405 410

<210> 10
<211> 269
<212> PRT
<213> Homo sapiens

<400> 10

Ala Cys Arg Pro Arg Ser Pro Pro Arg His Gln Ser Arg Cys Asp Gly
1 5 10 15

Asp Ala Ser Pro Pro Ser Pro Ala Arg Trp Ser Leu Gly Arg Lys Arg
20 25 30

Arg Ala Asp Gly Arg Arg Trp Arg Pro Glu Asp Ala Glu Glu Ala Glu
35 40 45

His Arg Gly Ala Glu Arg Arg Pro Glu Ser Phe Thr Thr Pro Glu Gly
50 55 60

Pro Lys Pro Arg Ser Arg Cys Ser Asp Trp Ala Ser Ala Val Glu Glu
65 70 75 80

```
Asp Glu Met Arg Thr Arg Val Asn Lys Glu Met Ala Arg Tyr Lys Arg
                85                  90                  95

Lys Leu Leu Ile Asn Asp Phe Gly Arg Glu Arg Lys Ser Ser Ser Gly
            100                 105                 110

Ser Ser Asp Ser Lys Glu Ser Met Ser Thr Val Pro Ala Asp Phe Glu
        115                 120                 125

Thr Asp Glu Ser Val Leu Met Arg Arg Gln Lys Gln Ile Asn Tyr Gly
    130                 135                 140

Lys Asn Thr Ile Ala Tyr Asp Arg Tyr Ile Lys Glu Val Pro Arg His
145                 150                 155                 160

Leu Arg Gln Pro Gly Ile His Pro Lys Thr Pro Asn Lys Phe Lys Lys
                165                 170                 175

Tyr Ser Arg Arg Ser Trp Asp Gln Gln Ile Lys Leu Trp Lys Val Ala
            180                 185                 190

Leu His Phe Trp Asp Pro Pro Ala Glu Glu Gly Cys Asp Leu Gln Glu
        195                 200                 205

Ile His Pro Val Asp Leu Glu Ser Ala Glu Ser Ser Ser Glu Pro Gln
    210                 215                 220

Thr Ser Ser Gln Asp Asp Phe Asp Val Tyr Ser Gly Thr Pro Thr Lys
225                 230                 235                 240

Val Arg His Met Asp Ser Gln Val Glu Asp Glu Phe Asp Leu Glu Ala
                245                 250                 255

Cys Leu Thr Glu Pro Leu Arg Asp Phe Ser Ala Met Ser
            260                 265
```

<210> 11
<211> 95
<212> PRT
<213> Homo sapiens

<400> 11

```
Pro Gly Pro Thr Pro Ser Gly Thr Asn Val Gly Ser Ser Gly Arg Ser
1               5                   10                  15

Pro Ser Lys Ala Val Ala Ala Arg Ala Ala Gly Ser Thr Val Arg Gln
            20                  25                  30
```

Arg Lys Asn Ala Ser Cys Gly Thr Arg Ser Ala Gly Arg Thr Thr Ser
35 40 45

Ala Gly Thr Gly Gly Met Trp Arg Phe Tyr Thr Glu Asp Ser Pro Gly
50 55 60

Leu Lys Val Gly Pro Val Pro Val Leu Val Met Ser Leu Leu Phe Ile
65 70 75 80

Ala Ser Val Phe Met Leu His Ile Trp Gly Lys Tyr Thr Arg Ser
85 90 95

<210> 12
<211> 23
<212> PRT
<213> Homo sapiens

<400> 12

Val Gly Pro Val Pro Val Leu Val Met Ser Leu Leu Phe Ile Ala Ser
1 5 10 15

Val Phe Met Leu His Ile Trp
20

<210> 13
<211> 285
<212> PRT
<213> Homo sapiens

<400> 13

Arg Leu Leu Pro Arg Leu Leu Leu Leu Leu Leu Leu Val Phe Pro Ala
1 5 10 15

Thr Val Leu Phe Arg Gly Gly Pro Arg Gly Leu Leu Ala Val Ala Gln
20 25 30

Asp Leu Thr Glu Asp Glu Glu Thr Val Glu Asp Ser Ile Ile Glu Asp
35 40 45

Glu Asp Asp Glu Ala Glu Val Glu Glu Asp Glu Pro Thr Asp Leu Val
50 55 60

Glu Asp Lys Glu Glu Glu Asp Val Ser Gly Glu Pro Glu Ala Ser Pro
65 70 75 80

Ser Ala Asp Thr Thr Ile Leu Phe Val Lys Gly Glu Asp Phe Pro Ala
85 90 95

```
Asn Asn Ile Val Lys Phe Leu Val Gly Phe Thr Asn Lys Gly Thr Glu
            100                 105                 110

Asp Phe Ile Val Glu Ser Leu Asp Ala Ser Phe Arg Tyr Pro Gln Asp
        115                 120                 125

Tyr Gln Phe Tyr Ile Gln Asn Phe Thr Ala Leu Pro Leu Asn Thr Val
    130                 135                 140

Val Pro Pro Gln Arg Gln Ala Thr Phe Glu Tyr Ser Phe Ile Pro Ala
145                 150                 155                 160

Glu Pro Met Gly Gly Arg Pro Phe Gly Leu Val Ile Asn Leu Asn Tyr
                165                 170                 175

Lys Asp Leu Asn Gly Asn Val Phe Gln Asp Ala Val Phe Asn Gln Thr
            180                 185                 190

Val Thr Val Ile Glu Arg Glu Asp Gly Leu Asp Gly Glu Thr Ile Phe
        195                 200                 205

Met Tyr Met Phe Leu Ala Gly Leu Gly Leu Leu Val Ile Val Gly Leu
    210                 215                 220

His Gln Leu Leu Glu Ser Arg Lys Arg Lys Arg Pro Ile Gln Lys Val
225                 230                 235                 240

Glu Met Gly Thr Ser Ser Gln Asn Asp Val Asp Met Ser Trp Ile Pro
                245                 250                 255

Gln Glu Thr Leu Asn Gln Ile Asn Lys Ala Ser Pro Arg Arg Leu Pro
            260                 265                 270

Arg Lys Arg Ala Gln Lys Arg Ser Val Gly Ser Asp Glu
        275                 280                 285
```

<210> 14
<211> 23
<212> PRT
<213> Homo sapiens

<400> 14

```
Thr Ile Phe Met Tyr Met Phe Leu Ala Gly Leu Gly Leu Leu Val Ile
1               5                   10                  15

Val Gly Leu His Gln Leu Leu
            20
```

<210> 15
<211> 609
<212> PRT
<213> Homo sapiens

<400> 15

```
Leu Asp Phe Phe Thr Ile Phe Ser Lys Gly Gly Leu Val Leu Trp Cys
1               5                   10                  15

Phe Gln Gly Val Ser Asp Ser Cys Thr Gly Pro Val Asn Ala Leu Ile
            20                  25                  30

Arg Ser Val Leu Leu Gln Val Gly Phe Gln Lys Ile Leu Thr Leu Thr
        35                  40                  45

Tyr Val Asp Lys Leu Ile Asp Asp Val His Arg Leu Phe Arg Asp Lys
    50                  55                  60

Tyr Arg Thr Glu Ile Gln Gln Gln Ser Ala Leu Ser Leu Leu Asn Gly
65                  70                  75                  80

Thr Phe Asp Phe Gln Asn Asp Phe Leu Arg Leu Leu Arg Glu Ala Glu
                85                  90                  95

Glu Ser Ser Lys Ile Arg Ala Pro Thr Thr Met Lys Lys Phe Glu Asp
            100                 105                 110

Ser Glu Lys Ala Lys Lys Pro Val Arg Ser Met Ile Glu Thr Arg Gly
        115                 120                 125

Glu Lys Pro Lys Glu Lys Ala Lys Asn Ser Lys Lys Lys Gly Ala Lys
    130                 135                 140

Lys Glu Gly Ser Asp Gly Pro Leu Ala Thr Ser Lys Pro Val Pro Ala
145                 150                 155                 160

Glu Lys Ser Gly Leu Pro Val Gly Pro Glu Asn Gly Val Glu Leu Ser
                165                 170                 175

Lys Glu Glu Leu Ile Arg Arg Lys Arg Glu Glu Phe Ile Gln Lys His
            180                 185                 190

Gly Arg Gly Met Glu Lys Ser Asn Lys Ser Thr Lys Ser Asp Ala Pro
        195                 200                 205

Lys Glu Lys Gly Lys Lys Ala Pro Arg Val Trp Glu Leu Gly Gly Cys
    210                 215                 220
```

```
Ala Asn Lys Glu Val Leu Asp Tyr Ser Thr Pro Thr Thr Asn Gly Thr
225                 230                 235                 240

Pro Glu Ala Ala Leu Ser Glu Asp Ile Asn Leu Ile Arg Gly Thr Gly
                245                 250                 255

Ser Gly Gly Gln Leu Gln Asp Leu Asp Cys Ser Ser Ser Asp Asp Glu
            260                 265                 270

Gly Ala Ala Gln Asn Ser Thr Lys Pro Ser Ala Thr Lys Gly Thr Leu
        275                 280                 285

Gly Gly Met Phe Gly Met Leu Lys Gly Leu Val Gly Ser Lys Ser Leu
    290                 295                 300

Ser Arg Glu Asp Met Glu Ser Val Leu Asp Lys Met Arg Asp His Leu
305                 310                 315                 320

Ile Ala Lys Asn Val Ala Ala Asp Ile Ala Val Gln Leu Cys Glu Ser
                325                 330                 335

Val Ala Asn Lys Leu Glu Gly Lys Val Met Gly Thr Phe Ser Thr Val
            340                 345                 350

Thr Ser Thr Val Lys Gln Ala Leu Gln Glu Ser Leu Val Gln Ile Leu
        355                 360                 365

Gln Pro Gln Arg Arg Val Asp Met Leu Arg Asp Ile Met Asp Ala Gln
    370                 375                 380

Arg Arg Gln Arg Pro Tyr Val Val Thr Phe Cys Gly Val Asn Gly Val
385                 390                 395                 400

Gly Lys Ser Thr Asn Leu Ala Lys Ile Ser Phe Trp Leu Leu Glu Asn
                405                 410                 415

Gly Phe Ser Val Leu Ile Ala Ala Cys Asp Thr Phe Arg Ala Gly Ala
            420                 425                 430

Val Glu Gln Leu Arg Thr His Thr Arg Arg Leu Ser Ala Leu His Pro
        435                 440                 445

Pro Glu Lys His Gly Gly Arg Thr Met Val Gln Leu Phe Glu Lys Gly
    450                 455                 460

Tyr Gly Lys Asp Ala Ala Gly Ile Ala Met Glu Ala Ile Ala Phe Ala
465                 470                 475                 480
```

Arg Asn Gln Gly Phe Asp Val Val Leu Val Asp Thr Ala Gly Arg Met
485 490 495

Gln Asp Asn Ala Pro Leu Met Thr Ala Leu Ala Lys Leu Ile Thr Val
500 505 510

Asn Thr Pro Asp Leu Val Leu Phe Val Gly Glu Ala Leu Val Gly Asn
515 520 525

Glu Ala Val Asp Gln Leu Val Lys Phe Asn Arg Ala Leu Ala Asp His
530 535 540

Ser Met Ala Gln Thr Pro Arg Leu Ile Asp Gly Ile Val Leu Thr Lys
545 550 555 560

Phe Asp Thr Ile Asp Asp Lys Val Gly Ala Ala Ile Ser Met Thr Tyr
565 570 575

Ile Thr Ser Lys Pro Ile Val Phe Val Gly Thr Gly Gln Thr Tyr Cys
580 585 590

Asp Leu Arg Ser Leu Asn Ala Lys Ala Val Val Ala Ala Leu Met Lys
595 600 605

Ala

<210> 16
<211> 18
<212> PRT
<213> Homo sapiens

<400> 16

Ser Thr Leu Gly His Met Val Leu Phe Pro Val Trp Phe Leu Tyr Ser
1 5 10 15

Leu Leu

<210> 17
<211> 261
<212> PRT
<213> Homo sapiens

<400> 17

Asp Ile Tyr Asp Thr Gln Thr Leu Gly Val Val Val Phe Gly Gly Phe
1 5 10 15

Met Val Val Ser Ala Ile Gly Ile Phe Leu Val Ser Thr Phe Ser Met
20 25 30

Lys Glu Thr Ser Tyr Glu Glu Ala Leu Ala Asn Gln Arg Lys Glu Met
35 40 45

Ala Lys Thr His His Gln Lys Val Glu Lys Lys Lys Lys Glu Lys Thr
50 55 60

Val Glu Lys Lys Gly Lys Thr Lys Lys Lys Glu Glu Lys Pro Asn Gly
65 70 75 80

Lys Ile Pro Asp His Asp Pro Ala Pro Asn Val Thr Val Leu Leu Arg
85 90 95

Glu Pro Val Arg Ala Pro Ala Val Ala Val Ala Pro Thr Pro Val Gln
100 105 110

Pro Pro Ile Ile Val Ala Pro Val Ala Thr Val Pro Ala Met Pro Gln
115 120 125

Glu Lys Leu Ala Ser Ser Pro Lys Asp Lys Lys Lys Lys Glu Lys Lys
130 135 140

Val Ala Lys Val Glu Pro Ala Val Ser Ser Val Val Asn Ser Ile Gln
145 150 155 160

Val Leu Thr Ser Lys Ala Ala Ile Leu Glu Thr Ala Pro Lys Glu Gly
165 170 175

Arg Asn Thr Asp Val Ala Gln Ser Pro Glu Ala Pro Lys Gln Glu Ala
180 185 190

Pro Ala Lys Lys Lys Ser Gly Ser Lys Lys Lys Gly Pro Pro Asp Ala
195 200 205

Asp Gly Pro Leu Tyr Leu Pro Tyr Lys Thr Leu Val Ser Thr Val Gly
210 215 220

Ser Met Val Phe Asn Glu Gly Glu Ala Gln Arg Leu Ile Glu Ile Leu
225 230 235 240

Ser Glu Lys Ala Gly Ile Ile Gln Asp Thr Trp His Lys Ala Thr Gln
245 250 255

Lys Gly Asp Pro Val
260

<210> 18
<211> 23
<212> PRT
<213> Homo sapiens

<400> 18

Leu Gly Val Val Val Phe Gly Gly Phe Met Val Val Ser Ala Ile Gly
1 5 10 15

Ile Phe Leu Val Ser Thr Phe
20

<210> 19
<211> 30
<212> PRT
<213> Homo sapiens

<400> 19

Asp Ile Tyr Asp Thr Gln Thr Leu Gly Val Val Val Phe Gly Gly Phe
1 5 10 15

Met Val Val Ser Ala Ile Gly Ile Phe Leu Val Ser Thr Phe
20 25 30

<210> 20
<211> 4
<212> PRT
<213> Homo sapiens

<400> 20

Lys Asp Glu Leu
1

<210> 21
<211> 4
<212> PRT
<213> Homo sapiens

<400> 21

Lys Glu Glu Leu
1

<210> 22
<211> 19
<212> PRT
<213> Homo sapiens

<400> 22

Met Gly Trp Ser Cys Ile Ile Leu Phe Leu Val Ala Thr Ala Thr Gly
1 5 10 15

Ala His Ser

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 2013058404 A **[0005] [0031] [0033]**
- EP 2784157 A **[0005]**
- US 2016075744 A **[0005]**
- US 62345252 B **[0079]**
- JP 2016120524 A **[0079]**

Patentkrav

1. Fusionsprotein til forbedring af et proteinekspressionsniveau fra en mål-mRNA, hvilket fusionsprotein omfatter:

(A) et eller flere funktionelle domæner, som forbedrer et proteinekspressionsniveau fra en mRNA; og

(B) en polypeptidenhed, som kan binde til en mål-mRNA på en RNA-base-selektiv eller RNA-basesekvens-specifik måde,

hvor polypeptidenheden (B) er en polypeptidenhed omfattende en eller flere PPR (pentatricopeptide repeat)-motiver, hvor hvert PPR-motiv omfatter et polypeptid bestående af 30 til 38 aminosyrer i længden og betegnes ved formel 1:

[Formel 1]

(Helix A)-X-(Helix B)-L (Formel 1),

hvori

Helix A er en enhed, som består af 12 aminosyrer i længden og kan danne en α-helix-struktur og er betegnet ved formel 2:

[Formel 2]

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

hvori A_1 til A_{12} hver uafhængigt betegner en aminosyre;

X ikke er til stede eller er en enhed bestående af 1 til 9 aminosyrer i længden;

Helix B er en enhed, som består af 11 til 13 aminosyrer i længden og kan danne en α-helix-struktur;

L er en enhed, som består af 2 til 7 aminosyrer i længden og er betegnet ved formel 3:

[Formel 3]

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

hvori aminosyrerne er nummererede fra C-terminalen som "i" (-1), "ii" (-2), ... og
L_{iii} til L_{vii} eventuelt ikke er til stede, og
en kombination af tre aminosyrer A_1, A_4 og L_{ii} eller en kombination af to aminosyrer A_4 og L_{ii} svarer til en base eller basesekvens af mål-mRNA'en,
hvor et eller flere funktionelle domæner (A) er udvalgt fra gruppen bestående af et domæne, som leder ribosom til mRNA'en, et domæne, som initierer eller fremmer translation af mRNA'en, et domæne, der er associeret med nukleær eksport af mRNA'en, et domæne, som binder til en endoplasmatisk reticulum-membran, et domæne, der indeholder en endoplasmatisk reticulum-retentionssignal (ER-retentionssignal)-sekvens, og et domæne, der indeholder en endoplasmatisk reticulum-signalsekvens.

2. Fusionsprotein ifølge krav 1, hvor polypeptidenheden (B) omfatter 2 til 30 PPR-motiver, og flerheden af PPR-motiver er indrettet, så de specifikt binder til basesekvensen af mål-mRNA'en.

3. Fusionsprotein ifølge krav 2, hvor polypeptidenheden (B) omfatter 5 til 25 PPR-motiver.

4. Fusionsprotein ifølge et hvilket som helst af kravene 1 til 3, hvor et eller flere funktionelle domæner (A) hver binder til en N-terminal og/eller en C-terminal af polypeptidenheden (B).

5. Fusionsprotein ifølge et hvilket som helst af de foregående krav,
hvor domænet, som leder ribosom til mRNAen, er et domæne, der indeholder hele eller en funktionel del af et polypeptid udvalgt fra gruppen bestående af DENR (tæthedsreguleret protein), MCT-1 (malign T-celle-amplificeret sekvens 1), TPT1 (translationalt styret tumorprotein) og Lerepo4 (zinkfinger-CCCH-domæne),
domænet, som initierer eller fremmer translation af mRNA'en, er et domæne, der indeholder hele eller en funktionel del af et polypeptid udvalgt fra gruppen bestående af eIF4E og eIF4G,

domænet, der er associeret med nukleær eksport af mRNA'en, er et domæne, der indeholder hele eller en funktionel del af SLBP (Stem-loop binding protein), domænet, der binder til en endoplasmatisk reticulum-membran, er et domæne, der indeholder hele eller en funktionel del af et polypeptid udvalgt fra gruppen bestående af SEC61B, TRAP-alpha (translocon-associeret protein alpha), SR-alpha, Dial (cytochrom b5-reduktase 3) og p180, endoplasmatisk reticulum-retentionssignal (ER-retentionssignal)-sekvensen er en signalsekvens, der indeholder en KDEL (KEEL)-sekvens, eller endoplasmatisk reticulum-signalsekvensen er en signalsekvens, der indeholder MGWSCIILFLVATATGAHS.

6. Fusionsprotein ifølge et hvilket som helst af kravene 1 til 5, hvor kombinationen af de tre aminosyrer A_1, A_4 og L_{ii} i hvert af PPR-motiverne er:
(valin, threonin, asparagin), (phenylalanin, serin, asparagin), (phenylalanin, threonin, asparagin), (isoleucin, asparagin, asparaginsyre) eller (threonin, threonin, asparagin) i rækkefølgen (A_1, A_4, L_{ii}), hvis en målbase for PPR-motivet er A (adenin);
(glutaminsyre, glycin, asparaginsyre), (valin, threonin, asparaginsyre), (lysin, threonin, asparaginsyre) eller (leucin, threonin, asparaginsyre) i rækkefølgen (A_1, A_4, L_{ii}), hvis målbasen for PPR-motivet er G (guanin);
(valin, asparagin, asparaginsyre), (isoleucin, asparagin, asparagin), (isoleucin, asparagin, asparaginsyre), (isoleucin, methionin, asparaginsyre), (phenylalanin, prolin, asparaginsyre) eller (tyrosin, prolin, asparaginsyre) i rækkefølgen (A_1, A_4, L_{ii}), hvis målbasen for PPR-motivet er U (uracil); eller
(valin, asparagin, asparagin), (isoleucin, asparagin, asparagin), (valin, asparagin, serin) eller (isoleucin, methionin, asparaginsyre) i rækkefølgen (A_1, A_4, L_{ii}), hvis målbasen for PPR-motivet er C (cytosin).

7. Fusionsprotein ifølge et hvilket som helst af kravene 1 til 5, hvor kombinationen af de to aminosyrer A_4 og L_{ii} i hvert af PPR-motiverne er:
(threonin, asparagin), (serin, asparagin) eller (glycin, asparagin) i rækkefølgen (A_4, L_{ii}), hvis en målbase for PPR-motivet er A (adenin);
(threonin, asparaginsyre) eller (glycin, asparaginsyre) i rækkefølgen (A_4, L_{ii}), hvis målbasen for PPR-motivet er G (guanin);

(asparagin, asparaginsyre), (prolin, asparaginsyre), (methionin, asparaginsyre) eller (valin, threonin) i rækkefølgen (A_4, L_{ii}), hvis målbasen for PPR-motivet er U (uracil); eller
(asparagin, asparagin), (asparagin, serin) eller (leucin, asparaginsyre) i rækkefølgen (A_4, L_{ii}), hvis målbasen for PPR-motivet er C (cytosin).

8. Nukleinsyre, der koder for fusionsproteinet ifølge et hvilket som helst af kravene 1 til 7.

9. Vektor omfattende nukleinsyren ifølge krav 8.

10. Vektor ifølge krav 9, hvor vektoren er en ekspressionsvektor.

11. In vitro-fremgangsmåde til forbedring af et proteinekspressionsniveau fra en mål-mRNA i en celle, hvilken fremgangsmåde omfatter:
et trin med at tilvejebringe fusionsproteinet ifølge et hvilket som helst af kravene 1 til 8 eller vektoren ifølge krav 9 eller 10; og
et trin med at indføre fusionsproteinet eller vektoren i en celle.

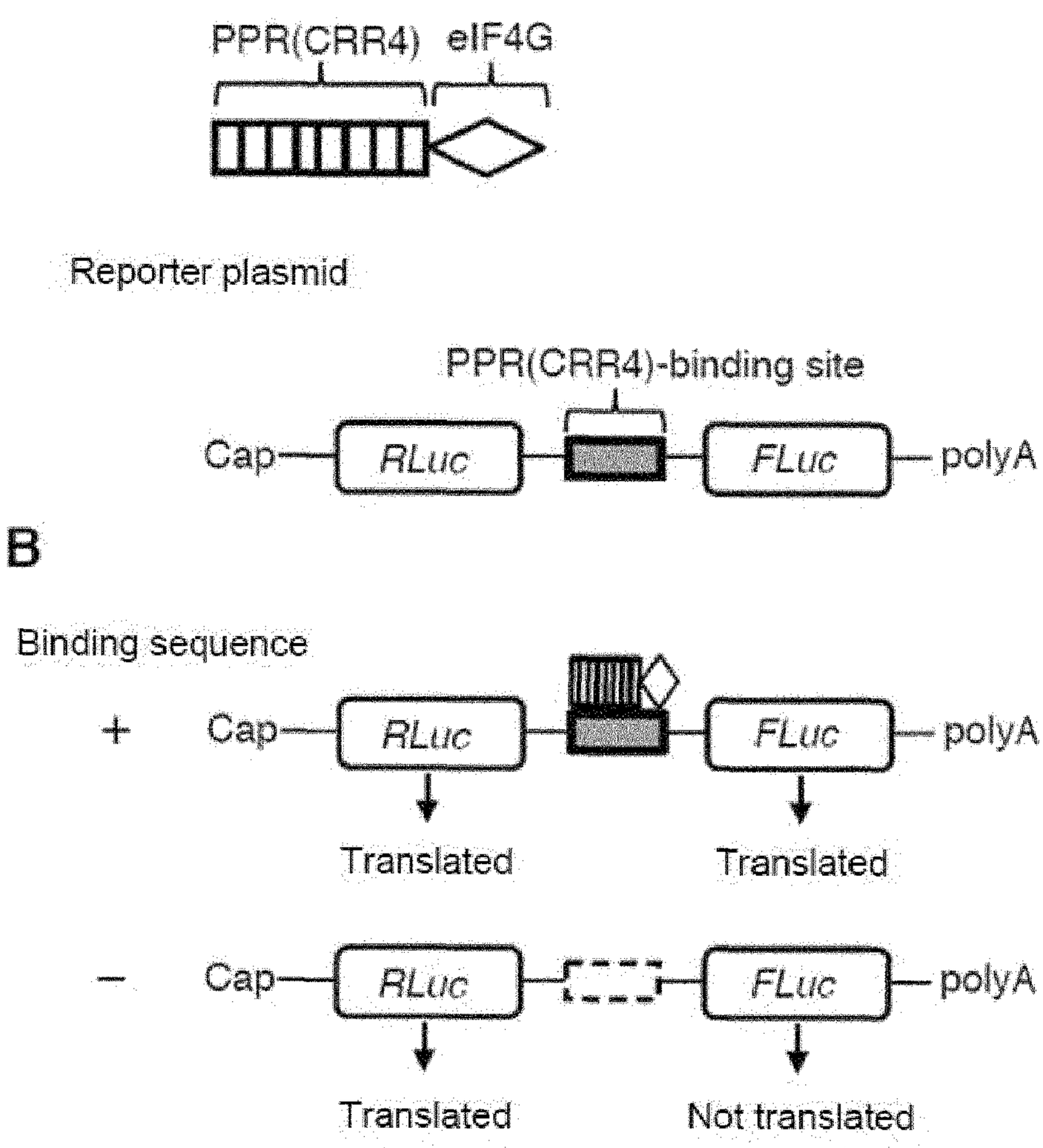
12. Fremgangsmåde ifølge krav 11,
hvor cellen er en eukaryotisk celle.

13. Fremgangsmåde ifølge krav 12,
hvor cellen er en dyrecelle.

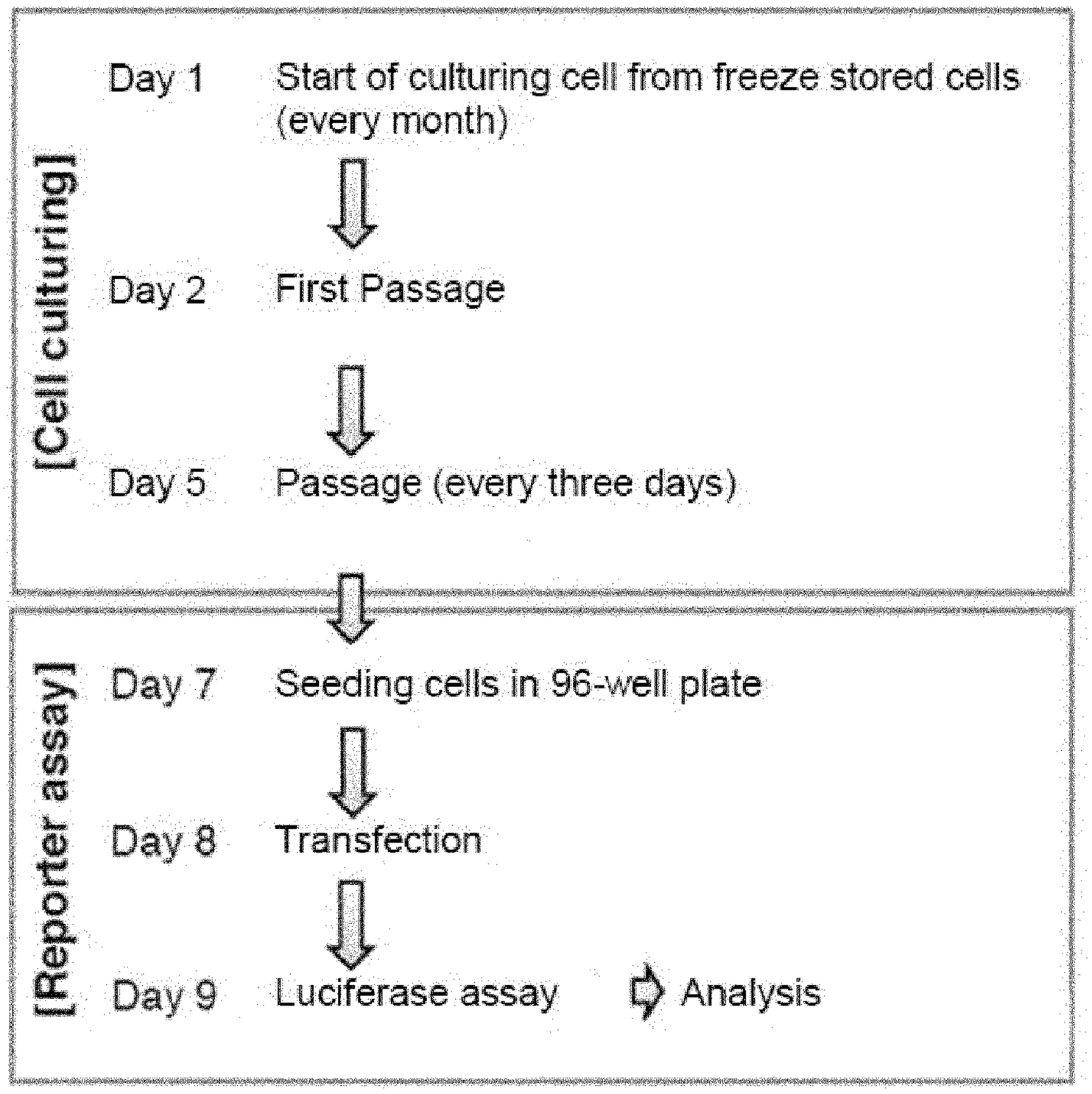
14. Fremgangsmåde ifølge krav 13,
hvor dyrecellen er en human celle.

DRAWINGS

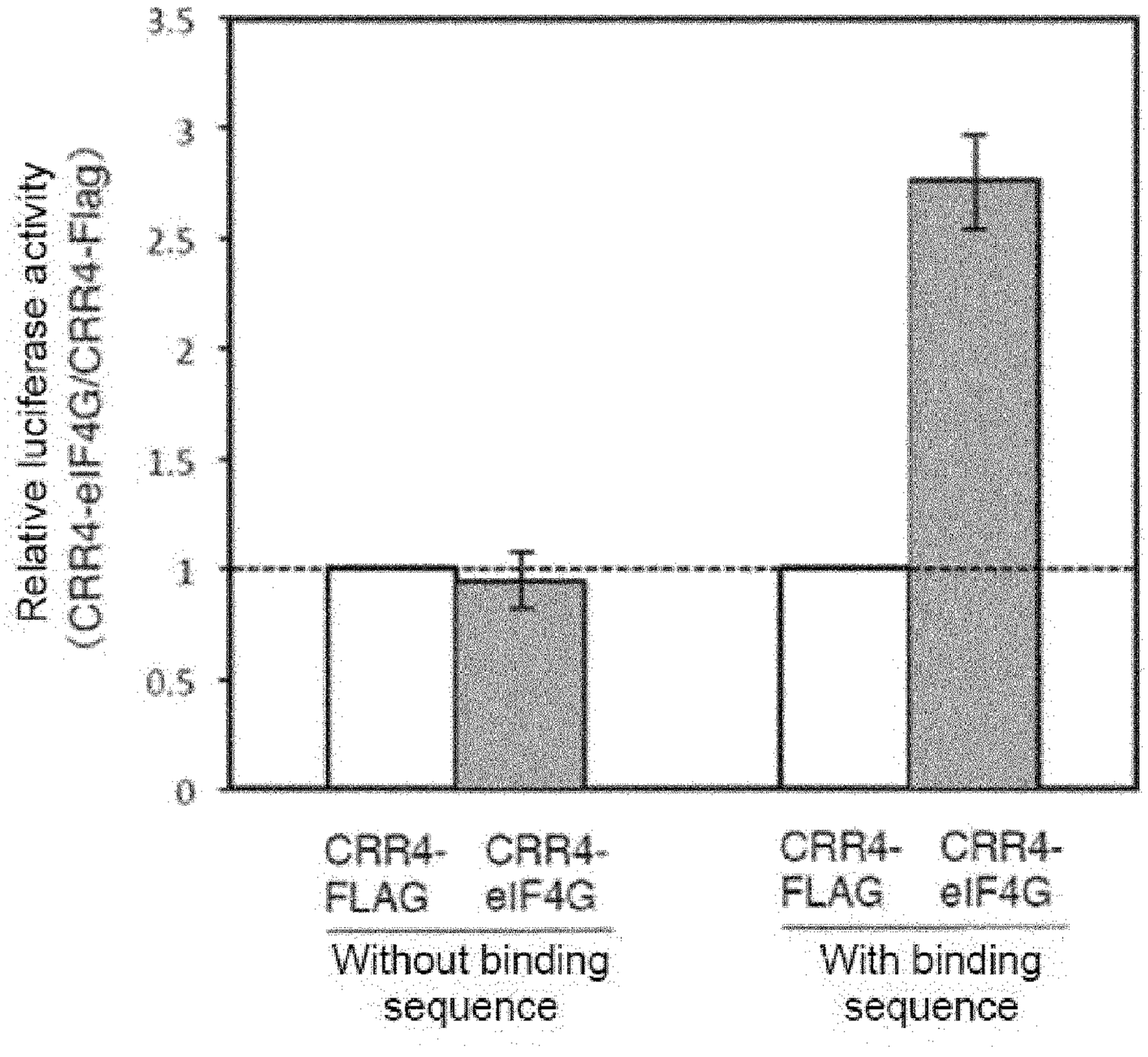
[Figure 1]



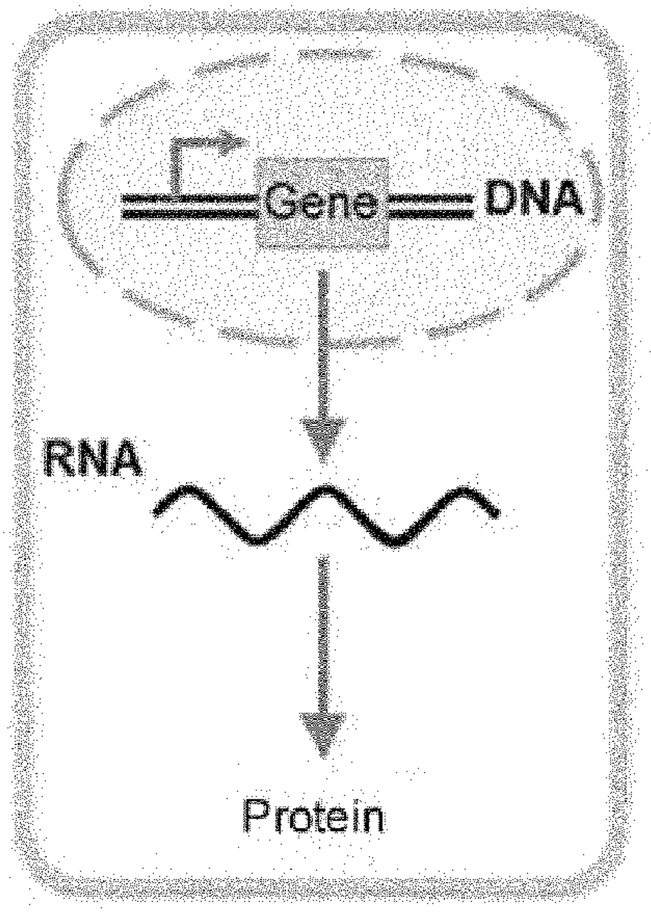
[Figure 2]


[Cell culturing]
Day 1
Start of culturing cell from freeze stored cells (every month)
Day 2
First Passage
Day 5
Passage (every three days)
[Reporter assay]
Day 7
Seeding cells in 96-well plate
Day 8
Transfection
Day 9
Luciferase assay
Analysis

[Figure 3]

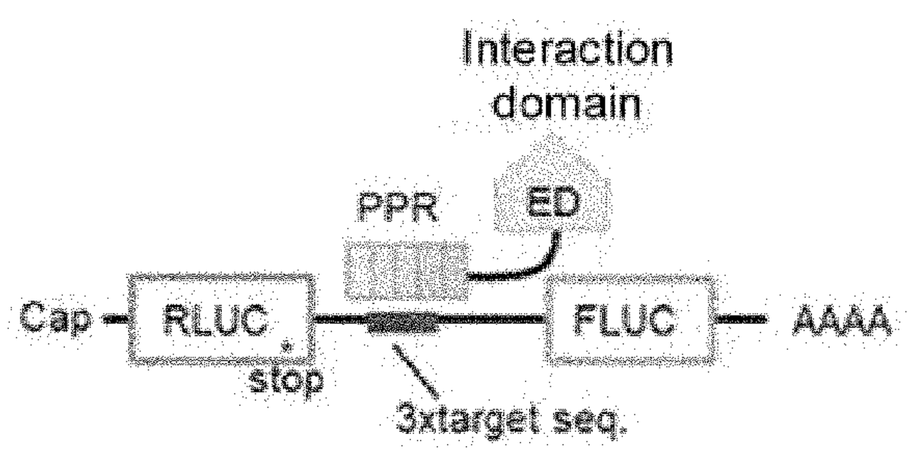

Relative luciferase activity
(CRR4-eIF4G/CRR4-Flag)
3.5
3
2.5
2
1.5
1
0.5
0
CRR4-FLAG
CRR4-eIF4G
CRR4-FLAG
CRR4-eIF4G
Without binding sequence
With binding sequence

[Figure 4]



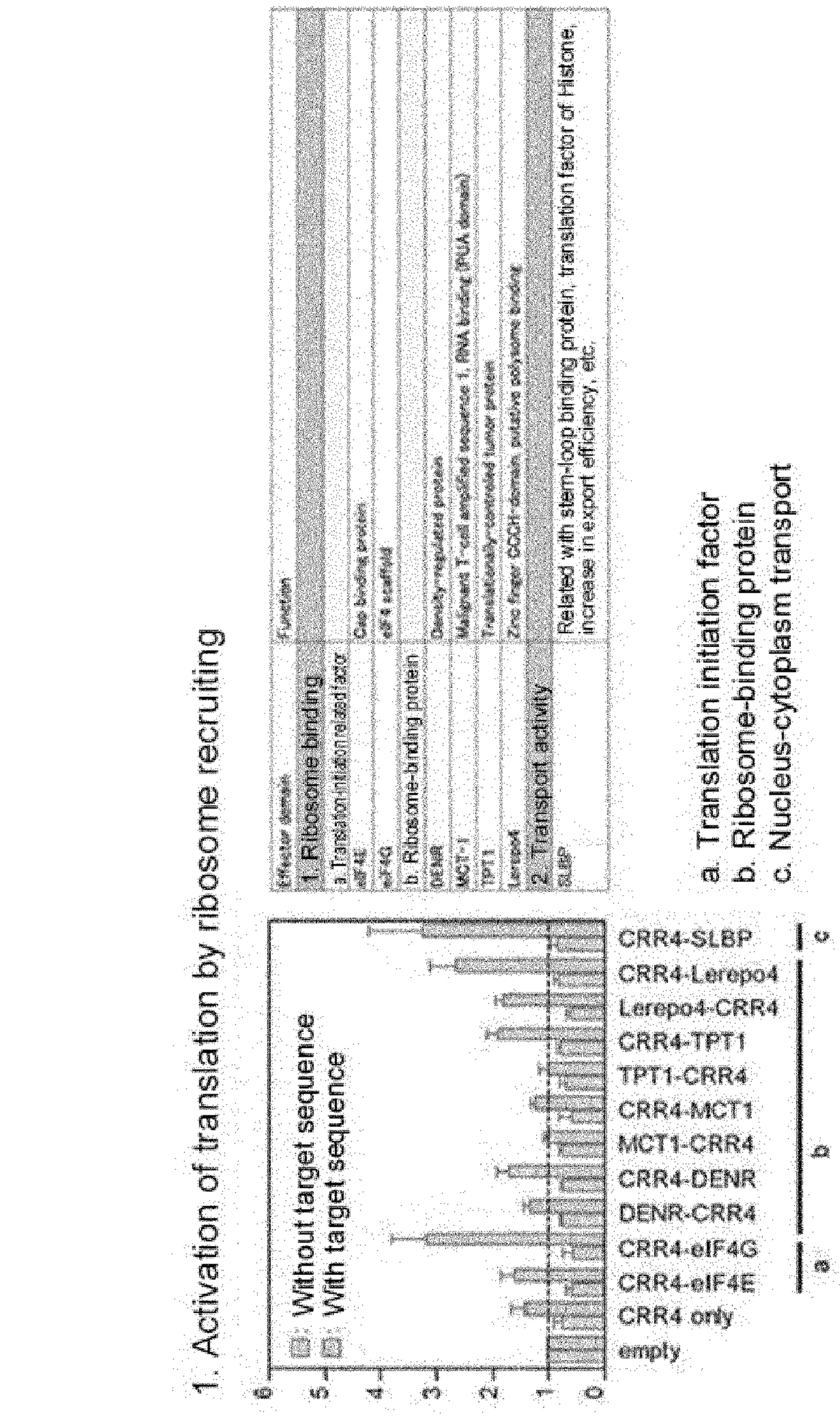
Factors to determine amount protein produced
1. Gene insertion position
2. Transcription activity
3. Regulation at RNA level
4. Post-translational modification

- Translation
- Transport from nucleus to cytoplasm
- Transport to ER membrane etc.

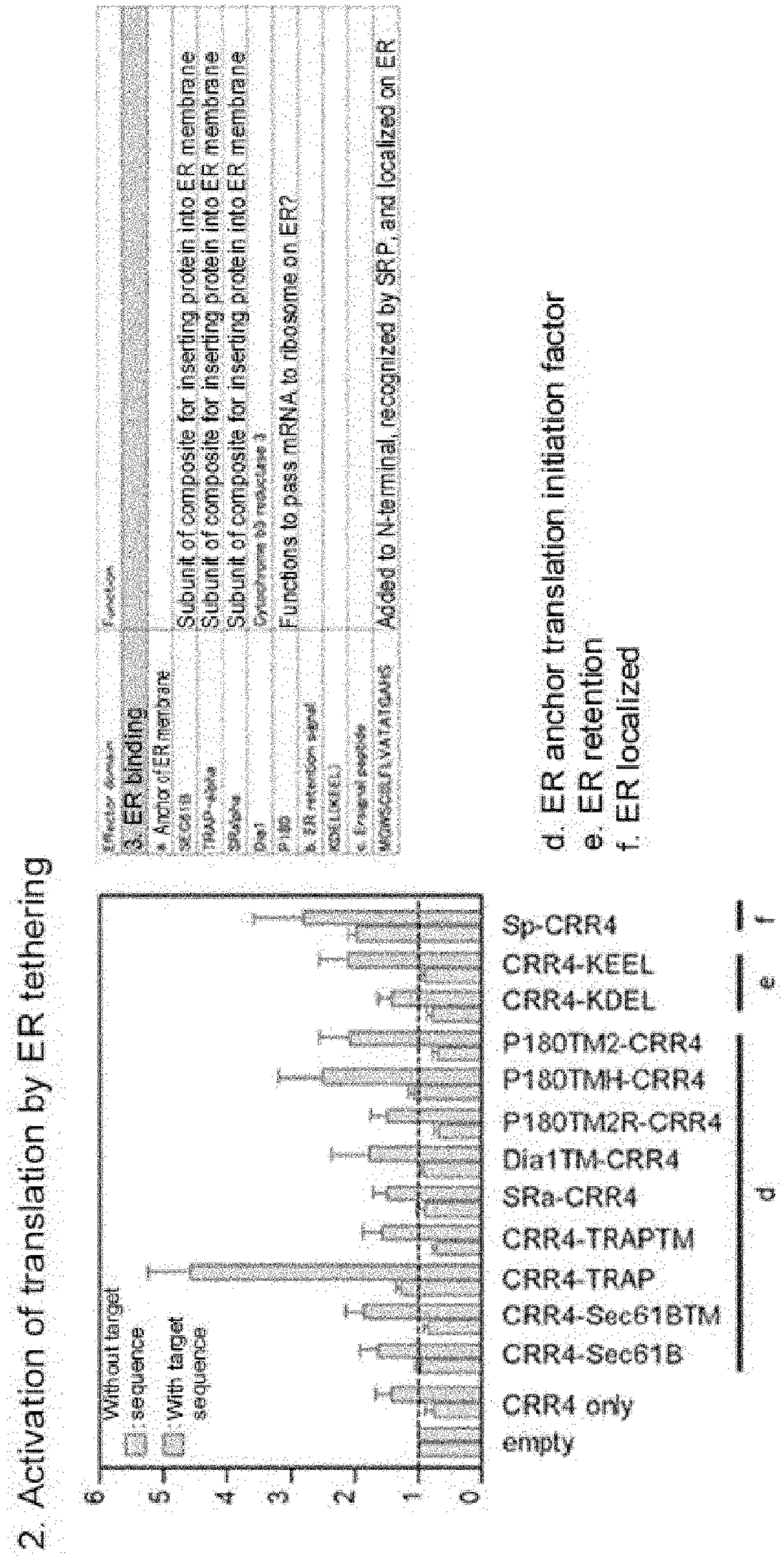


Such a reporter system is introduced into animal cultured cells to examine whether the amount of the FLUC reporter protein increases or not.
The amount of the FLUC protein increases only when PPR binds and provides an effect in the interaction domain.

[Figure 5]


1. Activation of translation by ribosome recruiting
Effector domain
Function
1. Ribosome binding
a. Translation initiation factor
eIF4E
eIF4G
b. Ribosome-binding protein
PABP
MCT-1
TPT1
Lerepo4
2. Transport activity
SLBP
Cap binding protein
eIF4 scaffold
Developmentally regulated protein
Malignant T-cell amplified sequence 1, RNA binding (RNA domain)
Translationally controlled tumor protein
Zinc finger CCCH domain, putative polysome bridge
Related with stem-loop binding protein, translation
increase in export efficiency, etc,
a. Translation initiation factor
b. Ribosome-binding protein
c. Nucleus-cytoplasm transport
Without target sequence
With target sequence
CRR4-SLBP
CRR4-Lerepo4
Lerepo4-CRR4
CRR4-TPT1
TPT1-CRR4
CRR4-MCT1
MCT1-CRR4
CRR4-DENR
DENR-CRR4
CRR4-eIF4G
CRR4-eIF4E
CRR4 only
empty
c
b
a
0
1
2
3
4
5
6

[Figure 6]


3. ER binding
Anchor of ER membrane
SEC61B
TRAP-delta
SRalpha
Dia1
P180
ER retention signal
KDEL(KEEL)
Fragment peptide
MGWKSSIFLVATATGAHS
Function
Subunit of composite for inserting protein into ER membrane
Subunit of composite for inserting protein into ER membrane
Subunit of composite for inserting protein into ER membrane
Cytoplasm or membrane ?
Functions to pass mRNA to ribosome on ER?
Added to N-terminal, recognized by SRP, and localized on ER
d. ER anchor translation initiation factor
e. ER retention
f. ER localized
2. Activation of translation by ER tethering
Without target sequence
With target sequence
6
5
4
3
2
1
0
Sp-CRR4
CRR4-KEEL
CRR4-KDEL
P180TM2-CRR4
P180TMH-CRR4
P180TM2R-CRR4
Dia1TM-CRR4
SRa-CRR4
CRR4-TRAPTM
CRR4-TRAP
CRR4-Sec61BTM
CRR4-Sec61B
CRR4 only
empty
f
e
d