

明 細 書

発明の名称：凝集の少ない P P R タンパク質及びその利用

技術分野

[0001] 本発明は、意図した核酸に結合可能なタンパク質を用いる核酸操作技術に関する。本発明は、医療（創薬支援、疾患の治療）、農業（農産物生産、育種）、化学（生物学的物質生産）などの幅広い分野で有用である。

背景技術

[0002] PPRタンパク質は、約35アミノ酸長からなるPPRモチーフの繰り返しを含むタンパク質で、PPRモチーフ1つが1つの塩基と特異的に結合することができる。PPRモチーフ内の1番目、4番目、ii番目(次のモチーフの2つ前)のアミノ酸の組み合わせによってアデニン、シトシン、グアニン、ウラシル(又はチミン)のうちのどれに結合するかが決まる（特許文献1、2）。

[0003] 天然に存在するRNA結合PPRモチーフのうち、もっとも多く出現するそれぞれの塩基に対応した組み合わせは、アデニンは、1番目がバリリン、4番目がトレオニン、ii番目がアスパラギン、シトシンは、1番目がバリリン、4番目がアスパラギン、ii番目がセリン、グアニンは、1番目がバリリン、4番目がトレオニン、ii番目がアスパラギン酸、ウラシルは、1番目が、バリリン、4番目がアスパラギン、ii番目がアスパラギン酸である（非特許文献1～5）。これらのアミノ酸の組み合わせを利用することで任意の配列に特異的に結合できるPPRタンパク質の設計が可能である。

先行技術文献

特許文献

[0004] 特許文献1：国際公開W02013/058404

特許文献2：国際公開W02014/175284

特許文献3：特願2019-100551

非特許文献

[0005] 非特許文献1：Coquille, S. et al. An artificial PPR scaffold for progr

amenable RNA recognition. Nature Communications 5, Article number: 5729(2014)

非特許文献2: Shen, C. et al. Specific RNA Recognition by Designer Pentatricopeptide Repeat Protein. Molecular Plant 8, 667-670(2015)

非特許文献3: Shen, C. et al. Structural basis for specific single-stranded RNA recognition by designer pentatricopeptide repeat proteins. Nature Communications volume 7, Article number: 11285 (2016)

非特許文献4: Miranda, R. G. et al. RNA-binding specificity landscapes of designer pentatricopeptide repeat proteins elucidate principles of PPR-RNA interactions. Nucleic Acids Research, 46(5), 2613-2623(2018)

非特許文献5: Yan, J. et al. Delineation of pentatricopeptide repeat codes for target RNA prediction. Nucleic Acids Research, gkz075(2019)

発明の概要

発明が解決しようとする課題

- [0006] 本発明者らは、上記のアミノ酸の組み合わせを利用し、高い性能を有し、また多くの、例えば15以上のPPRモチーフを連結したPPRタンパク質を作製することを検討してきた（特許文献3）。一方で、本発明者らの検討によると、このような方法で作製したPPRタンパク質の一部には、凝集性を示すものがあることが分かった。特に、PPRタンパク質を動物培養細胞で発現させた場合に、凝集が見られることがあった。

課題を解決するための手段

- [0007] そこで、この点をPPRモチーフ内のアミノ酸変異により解決することを検討した。そしてPPRタンパク質の1モチーフ目（N末端側）の、6番目、好ましくは6番目と9番目のアミノ酸を親水性のアミノ酸にすることでPPRの凝集性を改善することができることを見出し、本発明を完成した。

- [0008] 本発明は以下を提供する。

[1] 下記のいずれか1つの、PPRモチーフ：

(C-1)配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2)配列番号：4～7のいずれか1の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(A-2)(A-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつアデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつアデニン結合性であるPPRモチーフ；

(G-1)配列番号：9の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつグアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(U-2)(U-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつウラシル結合性であるPPRモチーフ。

[2] 下記のいずれか1つの、PPRモチーフ：

(C-1)配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2)配列番号：4～7のいずれか1の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換したPPRモチーフ；

(A-2)(A-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつアデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつアデニン結合性であるPPRモチーフ；

(G-1)配列番号：9の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換した配列からなるPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつグアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換した配列からなるPPRモチーフ；

(U-2)(U-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミ

ノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつウラシル結合性であるPPRモチーフ。

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミン酸である組み合わせ

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミンである組み合わせ

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がリジンである組み合わせ

- ・位置6のアミノ酸アスパラギン酸であり、かつ位置9アミノ酸がグリシンである組み合わせ

[3] 下記のいずれか1つである、1又は2に記載のPPRモチーフ：

(C-4)配列番号：4の配列からなるPPRモチーフ；

(A-4)配列番号：58の配列からなるPPRモチーフ；

(G-4)配列番号：59の配列からなるPPRモチーフ；

(U-4)配列番号：60の配列からなるPPRモチーフ。

[4] 1～3のいずれか1項に記載のPPRモチーフの、PPRタンパク質におけるN末端から1番目のPPRモチーフとしての使用。

[5] PPRタンパク質の凝集性を減少させるための、4に記載の使用。

[6] 下記の式1で表されるPPRモチーフを1～30個含む、特定の塩基配列を有する標的核酸と結合可能なタンパク質において、N末端から1番目のPPRモチーフ (M_1) の A_6 アミノ酸が、親水性アミノ酸である、タンパク質。

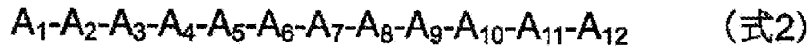
[化1]

(Helix A)-X-(Helix B)-L (式1)

(式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[化2]



式2中、 $A_1 \sim A_{12}$ はそれぞれ独立にアミノ酸を表し；

Xは、存在しないか又は1～9アミノ酸長からなる部分であり；

Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[化3]



式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $L_{iii} \sim L_{vii}$ は存在しない場合がある。）

[7] M_1 の A_9 アミノ酸が、親水性アミノ酸又はグリシンである、6に記載のタンパク質。

[8] M_1 の A_6 アミノ酸が、アスパラギン又はアスパラギン酸である、6又は7に記載のタンパク質。

[9] M_1 の A_9 アミノ酸が、グルタミン、グルタミン酸、リジン、又はグリシンである、6～8のいずれか1項に記載のタンパク質。

[10] M_1 の A_6 アミノ酸、及び M_1 の A_9 アミノ酸が、下記のいずれかの組み合わせである、6～9のいずれか1項に記載のタンパク質。

・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミン酸である組み合わせ

・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミンである組み合わせ

・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がリジンである組み合わせ

せ

・ A_6 アミノ酸アスパラギン酸であり、かつ A_9 アミノ酸がグリシンである組み合わせ

[11] 蛍光タンパク質、核移行シグナルペプチド、及びタグタンパク質からなる群より選択される少なくとも一つと、1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は6～10のいずれか1項に記載のタンパク質との、融合タンパク質。

[12] 6に定義したPPRモチーフを含む、特定の塩基配列を有する標的核酸と結合可能なPPRタンパク質の、改質方法であって、N末端から1番目のPPRモチーフ (M_1) の A_6 アミノ酸をより親水性とする、方法。

[13] 1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、6～10のいずれか1項に記載のタンパク質、又は11に記載の融合タンパク質を用いることを特徴とする、核酸の検出方法。

[14] 1～3のいずれか1項に記載のPPRモチーフ、1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は6～10のいずれか1項に記載のタンパク質をコードする核酸。

[15] 14に記載の核酸を含む、ベクター。

[16] 15に記載のベクターを含む、細胞（ヒト個体は除く。）。

[17] 1～3のいずれか1項に記載のPPRモチーフ、1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は6～10のいずれか1項に記載のタンパク質、又は15に記載のベクターを用いる、核酸の操作方法（ヒト個体での実施を除く。）。

[18] 17に記載の操作方法を含む、生物の生産方法。

[0009] 本発明はまた以下を提供する。

[1] 下記のいずれか1つの、PPRモチーフ：

(C-1) 配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2) 配列番号：4～7のいずれか1の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、

かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(A-2)(A-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつアデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつアデニン結合性であるPPRモチーフ；

(G-1)配列番号：9の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつグアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(U-2)(U-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつウラシル結合性であるPPRモチーフ。

[2] 下記のいずれか1つの、PPRモチーフ；

(C-1)配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2)配列番号：4～7のいずれか1の配列において、位置1、4、6、9、及び34

のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換したPPRモチーフ；

(A-2)(A-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつアデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつアデニン結合性であるPPRモチーフ；

(G-1)配列番号：9の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換した配列からなるPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつグアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換した配列からなるPPRモチーフ；

(U-2)(U-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつウラシル結合性であるPPRモチーフ。

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミン酸である組み合わせ

- ・位置6アミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミンである組み合わせ

- ・位置6アミノ酸がアスパラギンであり、かつ位置9のアミノ酸がリジンである組み合わせ

- ・位置6アミノ酸アスパラギン酸であり、かつ位置9アミノ酸がグリシンである組み合わせ

[3] 1又は2に記載のPPRモチーフの、PPRタンパク質におけるN末端から1番目のPPRモチーフとしての使用。

[4] PPRタンパク質の凝集性を減少させるための、3に記載の使用。

[5] 下記の式1で表されるPPRモチーフを1〜30個含む、特定の塩基配列を有する標的核酸と結合可能なタンパク質において、N末端から1番目のPPRモチーフ (M_1) の A_6 アミノ酸が、親水性アミノ酸である、タンパク質。

[化4]



(式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[化5]



式2中、 $A_1 \sim A_{12}$ はそれぞれ独立にアミノ酸を表し；

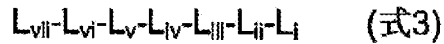
Xは、存在しないか又は1〜9アミノ酸長からなる部分であり；

Helix Bは、11〜13アミノ酸長からなる、 α ヘリックス構造を形成可能な部

分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[化6]



式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $L_{iii} \sim L_{vii}$ は存在しない場合がある。）

[6] M_1 の A_9 アミノ酸が、親水性アミノ酸又はグリシンである、5に記載のタンパク質。

[7] M_1 の A_6 アミノ酸が、アスパラギン又はアスパラギン酸である、5又は6に記載のタンパク質。

[8] M_1 の A_9 アミノ酸が、グルタミン、グルタミン酸、リジン、又はグリシンである、5～7のいずれか1項に記載のタンパク質。

[9] M_1 の A_6 アミノ酸、及び M_1 の A_9 アミノ酸が、下記のいずれかの組み合わせである、5～8のいずれか1項に記載のタンパク質。

- ・ A_6 のアミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミン酸である組み合わせ

- ・ A_6 のアミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミンである組み合わせ

- ・ A_6 のアミノ酸がアスパラギンであり、かつ A_9 アミノ酸がリジンである組み合わせ

- ・ A_6 のアミノ酸アスパラギン酸であり、かつ A_9 アミノ酸がグリシンである組み合わせ

[10] 蛍光タンパク質、核移行シグナルペプチド、及びタグタンパク質からなる群より選択される少なくとも一つと、1若しくは2に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は5～9のいずれか1項に記載のタンパク質との、融合タンパク質。

[11] 3に定義したPPRモチーフを含む、特定の塩基配列を有する標的核酸と

結合可能なPPRタンパク質の、改質方法であって、N末端から1番目のPPRモチーフ (M_1) のA₆アミノ酸をより親水性とする、方法。

[12] 1若しくは2に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、5～9のいずれか1項に記載のタンパク質、又は10に記載の融合タンパク質を用いることを特徴とする、核酸の検出方法。

[13] 1若しくは2に記載のPPRモチーフ、1若しくは2に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は5～9のいずれか1項に記載のタンパク質をコードする核酸。

[14] 13に記載の核酸を含む、ベクター。

[15] 14に記載のベクターを含む、細胞（ヒト個体は除く。）。

[16] 1若しくは2に記載のPPRモチーフ、1若しくは2に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は5～9のいずれか1項に記載のタンパク質、又は14に記載のベクターを用いる、核酸の操作方法（ヒト個体での実施を除く。）。

[17] 16に記載の操作方法を含む、生物の生産方法。

図面の簡単な説明

[0010] [図1]PPRモチーフの設計方法。A：1モチーフ目の6番目及び9番目のアミノ酸は外側に露出する。B：シトシンを認識する1モチーフ目の6番目、9番目のアミノ酸に関しては、典型的な組み合わせとして、ロイシンとグリシン (C₆L9G)、変異型として、ロイシンとグルタミン酸(C₆L9E)、アスパラギンとグルタミン(C₆N9Q)、アスパラギンとグルタミン酸(C₆N9E)、アスパラギンとトリジン(C₆N9K)、アスパラギン酸とグリシン(C₆D9G)を選択した。

[図2]各PPRタンパク質の凝集性及び核内への移行。GFP及び核移行シグナル配列を融合したPPRの細胞内で発現を蛍光顕微鏡画像で確認した。EGFPを融合した場合、PPRcag_1 (6L9G)とPPRcag_2(6L9E)においては、核へ局在せず、核の周りで強く凝集する様子が確認された。一方で、PPRcag_3 (6N9Q), PPRcag_4 (6N9E), PPRcag_5(6N9K), PPRcag_6(6D9G)においては、凝集性は低いものの核へ局在しなかった。mClover3を融合した場合、PPRcag_1 (6L9G)とPPRcag_2

(6L9E)においては、核へ局在するものの、核の中で凝集する様子が確認された。PPRcag_3 (6N9Q), PPRcag_4(6N9E), PPRcag_5(6N9K), PPRcag_6(6D9G)においては、核へ局在し、かつ凝集性が見られなかった。

[図3]PPRタンパク質とRNAとの結合実験。6番目、及び9番目のアミノ酸変異を含む全てのPPRにおいて、標的であるCAGx6へ特異的に結合することがわかった。標的配列への結合力は、PPRcag_1と比較してPPRcag_2は同程度、PPRcag_3は80%程度、PPRcag_4は60%程度、PPRcag_5は120%程度、PPRcag_6は130%程度であった。

[図4]N末から第1番目のPPRモチーフの凝集への影響。それぞれのPPRタンパク質を大腸菌発現系で作成し精製し、ゲルろ過クロマトグラフィーで分離した。溶出画分 (Elution vol.) が少ないほど分子サイズが大きい。V2では、8から10mLの溶出画分で溶出された一方で、v3.2では、12から14mLの溶出画分にピークが見られた。このことから、v2では、タンパク質サイズが大きくなっていることから凝集している可能性が示唆され、その凝集はv3.2において改善されていることが分かった。

発明を実施するための形態

[0011] [PPRモチーフ、PPRタンパク質]

(定義)

本発明でPPRモチーフというときは、特に記載した場合を除き、Web上のタンパク質ドメイン検索プログラムでアミノ酸配列を解析した際に、PfamにおいてPF01535、PrositeにおいてPS51375で得られるE値が所定値以下（望ましくはE-03）のアミノ酸配列をもつ30～38アミノ酸で構成されるポリペプチドをいう。本発明で定義するPPRモチーフを構成するアミノ酸の位置番号は、PF01535とほぼ同義である一方で、PS51375のアミノ酸の場所から2引いた数（例；本発明の1番→PS51375の3番）に相当する。ただし、“ii” (-2)番のアミノ酸というときは、PPRモチーフを構成するアミノ酸の後ろ（C末端側）から2番目のアミノ酸、又は次のPPRモチーフの1番アミノ酸に対して2個N末端側、すなわち-2番目のアミノ酸とする。次のPPRモチーフが明確に同定されない場合

、次のヘリックス構造の1番目のアミノ酸に対して、2コ前のアミノ酸を“ii”とする。Pfamについては<http://pfam.sanger.ac.uk/>、Prositateについては<http://www.expasy.org/prosite/>を参照することができる。

[0012] PPRモチーフの保存アミノ酸配列は、アミノ酸レベルでの保存性は低い、2次構造上で2つの α ヘリックスはよく保存されている。典型的なPPRモチーフは35アミノ酸で構成されるが、その長さは30～38アミノ酸と可変的である。

[0013] 本発明でいうPPRモチーフは、より具体的には、式1で表される、30～38アミノ酸長のポリペプチドからなる。

[0014] [化7]



式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[0015] [化8]



式2中、 $\text{A}_1 \sim \text{A}_{12}$ はそれぞれ独立にアミノ酸を表し；

Xは、存在しないか又は1～9アミノ酸長からなる部分であり；

Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[0016] [化9]



式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $\text{L}_{\text{iii}} \sim \text{L}_{\text{vii}}$ は存在しない場合がある。

[0017] 本発明でPPRタンパク質というときは、特に記載した場合を除き、上述のPP

Rモチーフを、1個以上、好ましくは2個以上有するPPRタンパク質をいう。本明細書でタンパク質というときは、特に記載した場合を除き、ポリペプチド（複数のアミノ酸がペプチド結合した鎖）からなる物質全般をいい、比較的 low molecular weight のポリペプチドからなるものも含まれる。本発明でアミノ酸という場合、通常のアミノ酸分子を指すことがあるほか、ペプチド鎖を構成しているアミノ酸残基を指すことがある。いずれを指しているかは、文脈から、当業者には明らかである。

[0018] 本発明で、PPRモチーフの標的核酸における塩基との結合性に関し、特異性／特異的というときは、特に記載した場合を除き、4種類のうちのいずれか一つの塩基に対する結合活性が、他の塩基に対する結合活性より高いことをいう。

[0019] 本発明で核酸というときは、RNA又はDNAを指す。なおPPRタンパク質は、RNA又はDNA中の塩基に対して特異性を有しうるが、核酸モノマーに結合するわけではない。

[0020] PPRモチーフは、1、4、ii番の3つのアミノ酸の組み合わせが、塩基との特異的な結合のために重要であり、これらの組み合わせにより、結合する塩基がいずれであるかを決定できる（前掲特許文献1、2）。

[0021] 具体的には、RNA結合性のPPRモチーフに関しては、1、4、ii番の3つのアミノ酸の組み合わせと結合可能な塩基との関係は、下記のとおりである（前掲特許文献1参照）。

(3-1) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン及びアスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にCに、その次にA又はGに対して結合するという、選択的なRNA塩基結合能を有する。

(3-2) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合し、次にGに、その次にCに対して結合するが、Uには結合しないという、選択的なRNA塩基結合能を有する。

(3-3) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Cに強く結合し、次にA又はUに対して結合するが、Gには結合しないという、選択的なRNA塩基結合能を有する。

(3-4) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、グルタミン酸、グリシン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合するが、A、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-5) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Cに強く結合し、次にUに、その次にAに対して結合するが、Gには結合しないという、選択的なRNA塩基結合能を有する。

(3-6) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合し、次にUに対して結合するが、AとCには結合しないという、選択的なRNA塩基結合能を有する。

(3-7) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、リジン、トレオニン、アスパラギン酸、の場合、そのPPRモチーフは、Gに強く結合し、次にAに対して結合するが、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-8) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、セリン、アスパラギン酸の場合、そのPPRモチーフは、Aに強く結合し、次にCに、その次にG及びUに対して結合するという、選択的なRNA塩基結合能を有する。

(3-9) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、セリンの場合、そのPPRモチーフは、Cに強く結合し、次にUに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(3-10) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルア

ラニン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合するが、G、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-11) A_i 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にAに対して結合するが、G及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-12) A_i 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、トレオニン、トレオニン、アスパラギンの場合、そのPPRモチーフは、Aに強く結合するが、G、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-13) A_i 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、メチオニン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合し、次にCに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(3-14) A_i 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、プロリン、アスパラギン酸の場合PPR、そのモチーフは、Uに強く結合し、次にCに対して結合するが、A及びGには結合しないという、選択的なRNA塩基結合能を有する。

(3-15) A_i 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、チロシン、プロリン、アスパラギン酸の場合、そのPPRモチーフは、Uに強く結合するが、A、G及びCには結合しないという、選択的なRNA塩基結合能を有する。

(3-16) A_i 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、ロイシン、トレオニン、アスパラギン酸の場合、そのPPRモチーフは、Gに強く結合するが、A、U及びCには結合しないという、選択的なRNA塩基結合能を有する。

[0022] 具体的には、DNA結合性のPPRモチーフに関しては、1、4、ii番の3つのアミノ酸の組み合わせと結合可能な塩基との関係は、下記のとおりである（前掲特許文献2参照）。

(2-1) A_i 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミ

ノ酸、グリシン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合する；

(2-2) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、グルタミン酸、グリシン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合する；

(2-3) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、グリシン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-4) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、グルタミン酸、グリシン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-5) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、グリシン、セリンであるとき、そのPPRモチーフは、Aに選択的に結合し、次にCに対して結合する；

(2-6) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、イソロイシン、任意のアミノ酸であるとき、そのPPRモチーフは、T及びCに選択的に結合する；

(2-7) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、イソロイシン、アスパラギンであるとき、そのPPRモチーフは、Tに選択的に結合し、次にCに対して結合する；

(2-8) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、ロイシン、任意のアミノ酸であるとき、そのPPRモチーフは、T及びCに選択的に結合する；

(2-9) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、ロイシン、アスパラギン酸であるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-10) A_i 、 A_4 、及び L_{ij} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、ロイシン、リシンであるとき、そのPPRモチーフは、Tに選択的に結合

する；

(2-11) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、メチオニン、任意のアミノ酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-12) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、メチオニン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-13) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、メチオニン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合し、次にCに対して結合する；

(2-14) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、アスパラギン、任意のアミノ酸であるとき、そのPPRモチーフは、C及びTに選択的に結合する；

(2-15) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-16) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-17) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、グリシン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-18) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-19) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、トレオニン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-20) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合し、次にCに対して結合する；

(2-21) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、チロシン、アスパラギン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合し、次にCに対して結合する；

(2-22) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、アスパラギン、アスパラギンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-23) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、アスパラギンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-24) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、セリン、アスパラギン、アスパラギンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-25) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、アスパラギンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-26) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、アスパラギン、セリンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-27) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、アスパラギン、セリンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-28) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、アスパラギン、トレオニンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-29) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、ア

スパラギン、トレオニンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-30) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、アスパラギン、トリプトファンであるとき、そのPPRモチーフは、Cに選択的に結合し、次にTに対して結合する；

(2-31) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、アスパラギン、トリプトファンであるとき、そのPPRモチーフは、Tに選択的に結合し、次にCに対して結合する；

(2-32) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、プロリン、任意のアミノ酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-33) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、プロリン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-34) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、プロリン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-35) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、チロシン、プロリン、アスパラギン酸であるとき、そのPPRモチーフは、Tに選択的に結合する；

(2-36) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、セリン、任意のアミノ酸であるとき、そのPPRモチーフは、A及びGに選択的に結合する；

(2-37) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、セリン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-38) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、セリン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

に結合する；

(2-39) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、セリン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-40) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、トレオニン、任意のアミノ酸であるとき、そのPPRモチーフは、A及びGに選択的に結合する；

(2-41) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、トレオニン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合する；

(2-42) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギン酸であるとき、そのPPRモチーフは、Gに選択的に結合する；

(2-43) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-44) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、フェニルアラニン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-45) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-46) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、バリン、トレオニン、アスパラギンであるとき、そのPPRモチーフは、Aに選択的に結合する；

(2-47) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、バリン、任意のアミノ酸であるとき、そのPPRモチーフは、A、C及びTに結合するが、Gには結合しない；

(2-48) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、イソロイシン、バリン、アスパラギン酸であるとき、そのPPRモチーフは、Cに選択的に結合し、次にAに対して結合する；

(2-49) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、バリン、グリシンであるとき、そのPPRモチーフは、Cに選択的に結合する；

(2-50) A_1 、 A_4 、及び L_{ii} の3つのアミノ酸の組み合わせが、順に、任意のアミノ酸、バリン、トレオニンであるとき、そのPPRモチーフは、Tに選択的に結合する；に基づいて決定される、タンパク質であって、選択的なDNA塩基結合能を有する。

[0023] (3つのアミノ酸の特に好ましい組み合わせ)

RNA結合性のPPRモチーフにおいては、各塩基を認識して特異的に結合可能な、代表的な1番目、4番目、ii番目のアミノ酸の組み合わせがある。具体的には、アデニンを認識する組み合わせは、1番目がバリン、4番目がトレオニン、ii番目がアスパラギン、シトシンを認識する組み合わせは、1番目がバリン、4番目がアスパラギン、ii番目がセリン、グアニンを認識する組み合わせは、1番目がバリン、4番目がトレオニン、ii番目がアスパラギン酸、ウラシルを認識する組み合わせは、1番目がバリン、4番目がアスパラギン、ii番目がアスパラギン酸である（前掲非特許文献1～5）。本発明の好ましい態様の一つでは、これらの組み合わせが用いられる。

[0024] (凝集性の改善)

本発明者らは、天然に存在する既存のPPRモチーフのアミノ酸情報から、PPRモチーフの6番目の位置のアミノ酸は疎水性（特にロイシン）、9番目の位置のアミノ酸は非親水性のアミノ酸（特にグリシン）である場合が非常に多いことを見出した。すでに結晶構造が得られているPPRタンパク質の構造（非特許文献6：Coquille et al., 2014 Nat. Commun.; PDB ID: 4PJQ, 4WN4, 4WSL, 4PJR; 非特許文献7：Shen et al., 2015 Nat. Commun., PDB ID: 5I9D, 5I9F, 5I9G, 5I9H）から、1モチーフ目(N末側)のそれら6番目、9番目は外側に

露出するため、その露出した疎水性アミノ酸が原因で凝集性を示すと想像した（図1A）。一方で、2モチーフ目以降においては、6番目、9番目のアミノ酸は、タンパク質内に埋もれ、疎水性コアを形成するため、すべてのモチーフの6番目、9番目に親水性残基を入れるとタンパク質構造が崩壊する可能性があると考えた。そこで1モチーフ目のみ、6番目、好ましくは6番目及び9番目のアミノ酸を親水性のアミノ酸（アスパラギン、アスパラギン酸、グルタミン、グルタミン酸、リジン、アルギニン、セリン、トレオニン）にすることでPPRの凝集性を減少させることとした。

[0025] 具体的には、次のようにする。

特定の塩基配列を有する標的核酸と結合可能なタンパク質において、N末端から1番目のPPRモチーフ（M₁）において：

(1) A₆アミノ酸を、親水性アミノ酸、好ましくはA₆アミノ酸を、アスパラギン又はアスパラギン酸とする。

(2) さらに、A₉アミノ酸を、親水性アミノ酸又はグリシン、好ましくはグルタミン、グルタミン酸、リジン、又はグリシンとする。

(3) あるいは、A₆アミノ酸及びA₉アミノ酸を、下記のいずれかの組み合わせとする。

・ A₆アミノ酸がアスパラギンであり、かつA₉アミノ酸がグルタミン酸である組み合わせ

・ A₆アミノ酸がアスパラギンであり、かつA₉アミノ酸がグルタミンである組み合わせ

・ A₆アミノ酸がアスパラギンであり、かつA₉アミノ酸がリジンである組み合わせ

・ A₆アミノ酸アスパラギン酸であり、かつA₉アミノ酸がグリシンである組み合わせ

[0026] （新規PPRモチーフ）

本発明は、上記により見出された、凝集性の改善された新規なPPRモチーフ、及びそれを含む新規なPPRタンパク質を提供する。

[0027] 本発明により提供される新規なPPRモチーフは、下記のものである。

(C-1)配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2)配列番号：4～7のいずれか1の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(A-2)(A-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつアデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつアデニン結合性であるPPRモチーフ；

(G-1)配列番号：9の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつグアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(U-2)(U-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6

、及び34のアミノ酸は同一であり、かつウラシル結合性であるPPRモチーフ。

[0028] このようなPPRモチーフのうち、特に好ましいものは、下記のものである。

(C-1)配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2)配列番号：4～7のいずれか1の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換したPPRモチーフ；

(A-2)(A-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつアデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつアデニン結合性であるPPRモチーフ；

(G-1)配列番号：9の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換した配列からなるPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつグアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換した配列からなるPPRモチーフ；

フ；

(U-2)(U-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつウラシル結合性であるPPRモチーフ。

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミン酸である組み合わせ

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミンである組み合わせ

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がリジンである組み合わせ

- ・位置6のアミノ酸アスパラギン酸であり、かつ位置9アミノ酸がグリシンである組み合わせ

[0029] 配列番号：4～10の具体的な配列は、図1、及び配列表に示されている。

[0030] このようなPPRモチーフのうち、さらに好ましいものは、下記のものである。

(C-4)配列番号：4の配列からなるPPRモチーフ；

(A-4)配列番号：58の配列からなるPPRモチーフ；

(G-4)配列番号：59の配列からなるPPRモチーフ；

(U-4)配列番号：60の配列からなるPPRモチーフ。

[0031] 配列番号：58～60の配列は、下記、及び配列表に示されている。

配列番号：58の配列 VTYTTNIDQLCKAGKVDEALELFKEMRSKGVKPNV

配列番号：59の配列 VTYTTNIDQLCKAGKVDEALELFDKEMKRGKIPDV

配列番号：60の配列 VTYNTNIDQLCKAGRLDEAEELLEEMEEKGKIPDV

[0032] (凝集性が改善されたPPRタンパク質)

本発明はまた、上記により見出された、凝集性が改善されたPPRタンパク質

を提供する。

[0033] 好ましい態様の一つでは、 M_1 の A_9 アミノ酸は、 M_1 の他のアミノ酸がいずれの場合であっても、また M_1 以外のモチーフのアミノ酸配列がいずれの場合であっても、非疎水性アミノ酸又はグリシンである。非疎水性アミノ酸は、親水性アミノ酸、又はシステイン、若しくはヒスチジンであり；好ましくは親水性アミノ酸、すなわちアルギニン、アスパラギン、アスパラギン酸、グルタミン酸、グルタミン、リジン、セリン、又はトレオニンであり；より好ましくはグルタミン、グルタミン酸、リジンである。

[0034] 好ましい態様の一つでは、 M_1 の A_9 アミノ酸は、 M_1 の他のアミノ酸がいずれの場合であっても、また M_1 以外のモチーフのアミノ酸配列がいずれの場合であっても、グルタミン、グルタミン酸、リジン、又はグリシンである。

[0035] 好ましい態様の一つでは、 M_1 の A_6 アミノ酸は、 M_1 の他のアミノ酸がいずれの場合であっても、また M_1 以外のモチーフのアミノ酸配列がいずれの場合であっても、非疎水性アミノ酸である。非疎水性アミノ酸は、例えば親水性アミノ酸、又はシステイン、若しくはヒスチジンであり；好ましくは親水性アミノ酸、すなわちアルギニン、アスパラギン、アスパラギン酸、グルタミン酸、グルタミン、リジン、セリン、又はトレオニンであり；より好ましくはアスパラギン、又はアスパラギン酸である。

[0036] 特に好ましい態様の一つでは、 M_1 の A_6 アミノ酸及び A_9 アミノ酸は、 M_1 の他のアミノ酸がいずれの場合であっても、また M_1 以外のモチーフのアミノ酸配列がいずれの場合であっても、下記のいずれかの組み合わせである：

- ・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミン酸である組み合わせ
- ・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミンである組み合わせ
- ・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がリジンである組み合わせ
- ・ A_6 アミノ酸アスパラギン酸であり、かつ A_9 アミノ酸がグリシンである組み合わせ

わせ

[0037] RNA結合性のタンパク質の好ましい態様の一つでは、 M_1 の A_6 アミノ酸及び A_9 アミノ酸が上記の条件を満たし、かつ含まれるPPRモチーフの少なくとも一つ、好ましくは半数以上、より好ましくはすべてが、下記のいずれかを満たす：

- ・ 結合する塩基がシトシンである場合は、 A_1 がバリンであり、 A_4 がアスパラギンであり、かつ A_{11} がセリンである
- ・ 結合する塩基がアデニンである場合は、 A_1 がバリンであり、 A_4 がトレオニンであり、かつ A_{11} がアスパラギンである
- ・ 結合する塩基がグアニンである場合は、 A_1 がバリンであり、 A_4 がトレオニンであり、 A_{11} がアスパラギン酸である
- ・ 結合する塩基がウラシル又はチミンである場合は、 A_1 がバリンであり、 A_4 がアスパラギンであり、 A_{11} がアスパラギン酸である

[0038] RNA結合性のタンパク質の好ましい態様の一つでは、 M_1 は上述した新規なPPRモチーフである。

[0039] 特に好ましい態様の一つでは、 M_1 が、下記のいずれか1つのポリペプチドからなるPPRモチーフであり、

- ・ 結合する塩基がシトシンである場合は、SEQ ID NOs:4-7のいずれか1の配列からなるポリペプチド
- ・ 結合する塩基がアデニンである場合は、SEQ ID NO:8の配列において、位置6及び9のアミノ酸を、次の段落に定義した組み合わせのいずれか1つを満たすように置換したポリペプチド
- ・ 結合する塩基がグアニンである場合は、SEQ ID NO:9の配列において、位置6及び9のアミノ酸を、次の段落に定義した組み合わせのいずれか1つを満たすように置換したポリペプチド
- ・ 結合する塩基がウラシルである場合は、SEQ ID NO:10の配列において、位置6及び9のアミノ酸を、次の段落に定義した組み合わせのいずれか1つを満たすように置換したポリペプチド

M_1 以外のPPRモチーフの少なくとも一つが、下記のいずれか1つのポリペプチ

ドからなるPPRモチーフである：

- ・ 結合する塩基がシトシンである場合は、SEQ ID NO:2の配列からなるポリペプチド
- ・ 結合する塩基がアデニンである場合は、SEQ ID NO:8の配列からなるポリペプチド
- ・ 結合する塩基がグアニンである場合は、SEQ ID NO:9の配列からなるポリペプチド
- ・ 結合する塩基がウラシルである場合は、SEQ ID NO:10の配列からなるポリペプチド

[0040] 上記の段落でいう組み合わせは、下記のいずれかである：

- ・ A₆アミノ酸がアスパラギンであり、かつA₉アミノ酸がグルタミン酸である組み合わせ
- ・ A₆アミノ酸がアスパラギンであり、かつA₉アミノ酸がグルタミンである組み合わせ
- ・ A₆アミノ酸がアスパラギンであり、かつA₉アミノ酸がリジンである組み合わせ
- ・ A₆アミノ酸アスパラギン酸であり、かつA₉アミノ酸がグリシンである組み合わせ

[0041] 特に好ましい態様の一つでは、M₁が、下記のいずれか1つのポリペプチドからなるPPRモチーフであり、

- ・ 結合する塩基がシトシンである場合は、SEQ ID NOs:4の配列からなるポリペプチド
- ・ 結合する塩基がアデニンである場合は、SEQ ID NO:58の配列からなるポリペプチド
- ・ 結合する塩基がグアニンである場合は、SEQ ID NO:59の配列からなるポリペプチド
- ・ 結合する塩基がウラシルである場合は、SEQ ID NO:60の配列からなるポリペプチド

M₁以外のPPRモチーフの少なくとも一つが、下記のいずれか1つのポリペプチドからなるPPRモチーフである：

- ・ 結合する塩基がシトシンである場合は、SEQ ID NO:2の配列からなるポリペプチド
- ・ 結合する塩基がアデニンである場合は、SEQ ID NO:8の配列において、位置15のアミノ酸のリジンへの置換を行った配列からなるポリペプチド
- ・ 結合する塩基がグアニンである場合は、SEQ ID NO:9の配列からなるポリペプチド
- ・ 結合する塩基がウラシルである場合は、SEQ ID NO:10の配列からなるポリペプチド

[0042] (性能の高いPPRモチーフの骨格の利用)

本発明の好ましい態様の一つでは、シトシン、アデニン、グアニン、ウラシル（又はチミン）それぞれに対するPPRモチーフにおいて、1、4、6、9、ii番目以外のアミノ酸を、特定のアミノ酸とすることができる。詳細には、シロイヌナズナのPPRモチーフ配列のうち、位置1、4、及びiiのアミノ酸の組み合わせがアデニンを認識するPPRモチーフのためにはVTN、シトシンを認識するPPRモチーフのためにはVSN、グアニンを認識するPPRモチーフのためにはVTD、ウラシルを認識するPPRモチーフのためにはVNDであるものを収集し、各位置に出現するアミノ酸の種類及びその数をまとめたときに、各位置において高頻度で出現するアミノ酸を選択することで、PPRモチーフの性能を高めることができる。

[0043] 1、4、6、9、ii番目以外のアミノ酸を、上述のように高頻度でみられるアミノ酸とするとの観点からは、RNA結合性のPPRタンパク質を得るためには、下記のPPRモチーフのアミノ酸配列を参考にすることができる。

シトシンに対応するPPRモチーフとしては、SEQ ID NOs:4-7のいずれか1の配列からなるPPRモチーフ；

アデニンに対応するPPRモチーフとしては、SEQ ID NO:8のいずれか1の配列からなるPPRモチーフ；

グアニンに対応するPPRモチーフとしては、SEQ ID NO:9のいずれか1の配列からなるPPRモチーフ；

ウラシルに対応するPPRモチーフとしては、SEQ ID NO:10のいずれか1の配列からなるPPRモチーフ。

[0044] (用語の説明等)

本発明で塩基配列（ヌクレオチド配列ということもある。）又はアミノ酸配列に関し「同一性」というときは、特に記載した場合を除き、2つの配列を最適の態様で整列させた場合に、2つの配列間で共有する一致した塩基又はアミノ酸の個数の百分率を意味する。すなわち、同一性＝（一致した位置の数／位置の全数）×100で算出でき、市販されているアルゴリズムを用いて計算することができる。また、このようなアルゴリズムは、Altschul et al., J. Mol. Biol. 215(1990) 403-410に記載されるNBLAST及びXBLASTプログラム中に組み込まれている。より詳細には、塩基配列又はアミノ酸配列の同一性に関する検索・解析は、当業者には周知のアルゴリズム又はプログラム（例えば、BLASTN、BLASTP、BLASTX、ClustalW）により行うことができる。プログラムを用いる場合のパラメーターは、当業者であれば適切に設定することができ、また各プログラムのデフォルトパラメーターを用いてもよい。これらの解析方法の具体的な手法もまた、当業者には周知である。

[0045] 本明細書において、塩基配列又はアミノ酸配列に関し、同一性がある（又は同一性が高い）というときは、特に記載した場合を除き、いずれの場合も少なくとも、70%、好ましくは80%以上、より好ましくは85%以上、さらに好ましくは90%以上、さらに好ましくは95%以上、さらに好ましくは97.5%以上さらに好ましくは99%以上の同一性がある場合を指す。

[0046] また、本発明でPPRモチーフ又はタンパク質に関し、「置換、欠失、又は付加した配列」というときの置換等されるアミノ酸の個数は、特に記載した場合を除き、いずれのモチーフ又はタンパク質においても、そのアミノ酸配列からなるモチーフ又はタンパク質が所望の機能を有する限り特に限定されないが、1～9個又は1～4個程度であるか、性質の似たアミノ酸への置換であれ

ば、さらに多くの個数の置換等がありうる。このようなアミノ酸配列に係るポリヌクレオチド又はタンパク質を調製するための手段は、当業者にはよく知られている。

[0047] 性質の似たアミノ酸とは、ハイドロパシー、荷電、pKa、溶解性等の物性が似ているアミノ酸をいい、例えば、次のようなものを指す。

疎水性アミノ酸；アラニン、バリン、グリシン、イソロイシン、ロイシン、フェニルアラニン、プロリン、トリプトファン、チロシン

非疎水性アミノ酸；アルギニン、アスパラギン、アスパラギン酸、グルタミン酸、グルタミン、リジン、セリン、トレオニン、システイン、ヒスチジン；

親水性アミノ酸；アルギニン、アスパラギン、アスパラギン酸、グルタミン酸、グルタミン、リジン、セリン、トレオニン；

酸性アミノ酸：アスパラギン酸、グルタミン酸；

塩基性アミノ酸：リジン、アルギニン、ヒスチジン；

中性アミノ酸：アラニン、アスパラギン、システイン、グルタミン、グリシン、イソロイシン、ロイシン、メチオニン、フェニルアラニン、プロリン、セリン、トレオニン、トリプトファン、チロシン、バリン；

含硫アミノ酸：メチオニン、システイン；

含芳香環アミノ酸：チロシン、トリプトファン、フェニルアラニン。

[0048] 遺伝子、核酸、ポリヌクレオチド、タンパク質、及びモチーフ等に関し、「作製」は、「生産」又は「製造」と言い換えることができる。また遺伝子等に関し、パーツを組み合わせて作製する場合に、「構築」ということがあるが、「構築」も、「生産」又は「製造」と言い換えることができる。

[0049] 本発明のPPRモチーフ、それを含むタンパク質、又はそれらをコードする核酸は、当業者であれば、従来技術、及び本明細書の実施例の項の記載を利用して製造することができる。

[0050] [PPRタンパク質の特徴、用途]
(PPRタンパク質凝集性の改善)

本発明の新規 PPRモチーフを用いて作製したPPRタンパク質は、細胞内での凝集性が減少されている。PPRタンパク質の凝集性は、当業者であれば、PPRタンパク質を細胞内で発現させ、凝集の有無を確認することにより、評価できる。確認は、PPRタンパク質を蛍光タンパク質と融合させて発現すれば、より容易である。本発明者らの検討によると、PPRタンパク質の一番目のモチーフにおけるアミノ酸の適切な改変により、PPRタンパク質の細胞内での凝集性が改善され、核への移行性が向上している。

[0051] (結合力)

本発明の新規 PPRモチーフを用いて作製したPPRタンパク質は、細胞内での凝集性が減少されているのみならず、同じ標的RNAに対する既存のPPRモチーフを用いて作製したPPRタンパク質と同等、又はそれより高いRNA結合性能を有する。同等とは、55%以上であることをいい、好ましくは75%程度であることをいう。

[0052] 標的配列に対する結合力は、EMSA(Electrophoretic Mobility Shift Assay)やBiacoreを用いた方法によって評価することができる。EMSAは、タンパク質と核酸が結合したサンプルを電気泳動した際、核酸分子の移動度が結合していない場合と比較して変化する性質を利用する方法である。Biacoreに代表される分子間相互作用解析機器は、反応速度論的解析ができるため、詳細なタンパク質-核酸結合解析が可能である。

[0053] 標的配列に対する結合力はまた、固相化された標的核酸に、候補タンパク質を含む溶液を供し、標的核酸に結合したタンパク質を検出又は定量することによっても解析できる。この方法は、ELISA (Enzyme-Linked Immuno Sorbent Assay)を応用したものであることから、RPB-ELISA (RNA-protein binding ELISA)法と称することがある。固相化された標的核酸に、候補タンパク質を含む溶液を供する工程は、具体的には、プレートに固定された標的核酸分子に、対象結合タンパク質を含む溶液を流すことにより実施できる。標的核酸分子の固定化は、既存の種々の固定化方法が利用でき、例えばストレプトアビジンがコーティングされたウェルプレートに対して、ビオチン修飾した標

的核酸分子を含む核酸プローブを与えることにより達成できる。詳細な実験条件は、本発明の実施例の項に詳述した実験方法を参考にすることができる。RPB-ELISAでは、対象PPRタンパク質とその標的RNAを加えたサンプルの発光量から、バックグラウンドシグナル（標的RNAを加えずに対象PPRタンパク質を加えた際の発光シグナル値）を差し引いた値を対象PPRタンパク質とその標的RNAとの結合力とすることができる。

[0054] [PPRタンパク質の利用]

（複合体、融合タンパク質）

本発明により提供されるPPRモチーフ又はPPRタンパク質は、機能性領域を連結し、複合体とすることができる。また、タンパク質性の機能性領域を連結し、融合タンパク質とすることができる。機能性領域とは、生体内又は細胞内で特定の生物学的機能、例えば酵素機能、触媒機能、阻害機能、亢進機能などの機能を有する部分、又は標識としての機能を有する部分をいう。そのような領域は、例えば、タンパク質、ペプチド、核酸、生理活性物質、薬剤からなる。なお、以下では、本発明を、複合体に関し、融合タンパク質を例に説明することがあるが、当業者であれば、その説明に準じて、融合タンパク質以外の複合体の場合についても理解することができる。

[0055] 好ましい態様の一つでは、機能性領域はリボヌクレアーゼ（RNase）である。RNaseの例は、RNase A（例えば、bovine pancreatic ribonuclease A: PDB 2AAS）、RNase Hである。

[0056] 好ましい態様の一つでは、機能性領域は蛍光タンパク質である。蛍光タンパク質の例は、mCherry、EGFP、GFP、Sirius、EBFP、ECFP、mTurquoise、TagCFP、AmCyan、mTFP1、MidoriishiCyan、CFP、TurboGFP、AcGFP、TagGFP、Azami-Green、ZsGreen、EmGFP、HyPer、TagYFP、EYFP、Venus、YFP、PhiYFP、PhiYFP-m、TurboYFP、ZsYellow、mBanana、KusabiraOrange、mOrange、TurboRFP、DsRed-Express、DsRed2、TagRFP、DsRed-Monomer、AsRed2、mStrawberry、TurboFP602、mRFP1、JRed、KillerRed、HcRed、KeimaRed、mRaspberry、mPlum、PS-CFP、Dendra2、Kaede、EosFP、KikumeGRである。融合タンパク質とし

ての、凝集性の改善、及び／又は核への効率的な移行の観点から、好ましい例はmClover3である。

[0057] 好ましい態様の一つでは、機能性領域は、標的がmRNAである場合に、標的mRNAからのタンパク質発現量を向上させる機能ドメインである (W02017/209122)。mRNAからのタンパク質発現量を向上させる機能ドメインの例は、例えば、mRNAの翻訳を直接的又は間接的に促進することが知られているタンパク質の機能ドメインの全部又は機能的な一部であってよい。より具体的には、mRNAヘリボソームを誘導するドメイン、mRNAの翻訳開始又は翻訳促進に関連するドメイン、mRNAの核外への輸送に関連するドメイン、小胞体膜への結合に関連するドメイン、小胞体保留シグナル (ER retention signal) 配列を含むドメイン、又は、小胞体シグナル配列を含むドメインであってよい。さらに具体的には、上記のmRNAヘリボソームを誘導するドメインは、DENR (Density-regulated protein)、MCT-1 (Malignant T-cell amplified sequence 1)、TPT1 (Translationally-controlled tumor protein)、及び、Lerepo4 (Zinc finger CCCH-domain) からなる群から選択されるポリペプチドの全部又は機能的な一部を含むドメインであってよい。また、上記のmRNAの翻訳開始又は翻訳促進に関連するドメインは、eIF4E及びeIF4Gからなる群から選択されるポリペプチドの全部又は機能的な一部を含むドメインであってよい。また、上記のmRNAの核外への輸送に関連するドメインは、SLBP (Stem-loop binding protein) の全部又は機能的な一部を含むドメインであってよい。また、上記の小胞体膜への結合に関連するドメインは、SEC61B、TRAP-alpha (Translocon associated protein alpha)、SR-alpha、Dia1 (Cytochrome b5 reductase 3)、及び、p180からなる群から選択されるポリペプチドの全部又は機能的な一部を含むドメインであってよい。また、上記の小胞体保留シグナル (ER retention signal) 配列は、KDEL (SEQ ID NO:55) またはKEEL (SEQ ID NO:56) 配列を含むシグナル配列であってよい。また、前記小胞体シグナル配列は、MGWSCIIILFVATATGAHS (SEQ ID NO:57) を含むシグナル配列であってよい。

[0058] 本発明において、機能性領域は、PPRタンパク質のN末端側に融合されてよく、C末端側に融合されてもよく、N末端側とC末端側の両方に融合されてもよい。また、複合体又は融合タンパク質は、複数の機能性領域（例えば、2～5個）を含んでよい。さらに、本発明の複合体又は融合タンパク質は、機能性領域とPPRタンパク質とがリンカー等を介して間接的に融合されていてもよい。

[0059] （PPRタンパク質等をコードする核酸、ベクター、細胞）

本発明は、上述の、PPRモチーフ、PPRタンパク質又は融合タンパク質をコードする核酸、核酸を含むベクター（例えば増幅のためのベクター、発現ベクター）も提供する。増幅のためのベクターは、大腸菌や酵母を宿主として用いる。本明細書において、発現ベクターとは、例えば上流から、プロモーター配列を有するDNA、所望のタンパク質をコードするDNA、及びターミネーター配列を有するDNAを含むベクターを意味するが、所望の機能を発揮する限り、必ずしもこの順に配列されている必要はない。本発明においては、当業者が通常使用し得る様々なベクターを組み換えて使用することができる。

[0060] 本発明のPPRタンパク質、又は融合タンパク質は、真核生物（例えば、動物、植物、微生物（酵母、等）、原生生物）の細胞で機能し得る。本発明の融合タンパク質は、特に、動物細胞内（*in vitro*又は*in vivo*）で機能し得る。本発明のPPRタンパク質、又は融合タンパク質、又はそれを発現するベクターを導入し得る動物細胞としては、例えば、ヒト、サル、ブタ、ウシ、ウマ、イヌ、ネコ、マウス、ラット由来の細胞を挙げることができる。また、本発明のPPRタンパク質、又は融合タンパク質、又はそれを発現するベクターを導入し得る培養細胞としては、例えば、チャイニーズハムスター卵巣（CHO）細胞、COS-1細胞、COS-7細胞、VERO（ATCC CCL-81）細胞、BHK細胞、イヌ腎由来MDCK細胞、ハムスターAV-12-664細胞、HeLa細胞、WI38細胞、293細胞、293T細胞、PER、C6細胞を挙げることができるが、これらに限定されない。

[0061] （用途）

本発明のPPRタンパク質、又は融合タンパク質は、生体内又は細胞内に、核

酸配列特異的に機能性領域をデリバリーし、機能させることができる可能性がある。GFP等の標識部分を連結した複合体は、所望のRNAを生体内で可視化するために用いる。

[0062] また本発明のPPRタンパク質、又は融合タンパク質は、細胞内又は生体内において、核酸配列特異的に改変・破壊を行うことができ、また新たな機能を付与できる可能性がある。特にRNA結合性のPPRタンパク質は、オルガネラで見られるすべてのRNA加工のステップ、切断、RNA編集、翻訳、スプライシング、RNA安定化に関与している。したがって、本発明により提供されるPPRタンパク質の改質に関わる方法、及び本発明により提供されるPPRモチーフ及びPPRタンパク質は、様々な分野で、以下のような利用が期待できる。

[0063] (1)医療

・特定の疾患に関連した特定のRNAを認識し、結合するPPRタンパク質を作製する。また、特定のRNAに関し、標的配列を解析し、及び付随するタンパク質を解析する。それらの解析結果は、疾患の治療のための化合物の探索に用いる。

[0064] 例えば、動物では、LRPPRCと同定されるPPRタンパク質の異常がLeigh syndrome French Canadian (LSFC; リー症候群、亜急性壊死性脳脊髄症)を引き起こすことが知られている。本発明は、LSFCの処置（予防、治療、進行の抑制）に寄与しうる。既存のPPRタンパク質の多くは、RNA操作（RNA上での遺伝情報の変換；多くの場合、C→U）の編集部位の指定に働く。このタイプのPPRタンパク質は、RNA編集酵素と相互作用すると示唆される付加モチーフがC末端側に存在する。このような構造を有するPPRタンパク質により、塩基多型を導入すること、又は塩基多型に起因した疾患又は状態を処置することが期待できる。

[0065] ・RNAの抑制・発現をコントロールした細胞を作製する。このような細胞には、分化・未分化状態をモニタリングした幹細胞（例えば、iPS細胞）、化粧品の評価用モデル細胞、創薬のメカニズム解明や薬理試験を目的として、機能性RNAの発現をON/OFFできる細胞が含まれる。

[0066] ・特定の疾患に関連した特定のRNAに対して特異的に結合するPPRタンパク質を作製する。このようなPPRタンパク質を、プラスミド、ウイルスベクター、mRNA、精製タンパク質を用いて細胞へ導入し、そのPPRタンパク質が細胞内でその標的RNAと結合することで、疾患の原因であるRNA機能を変化（改善）させることができる。機能変化する手段は、例えば、結合によるRNA構造の変化、分解することによるノックダウン、スプライシングによるスプライシング反応の変化、塩基置換などが挙げられる。

[0067] (2) 農林水産業

- ・農作物、林産物、水産物等において、収量や品質を改善する。
- ・耐病性の向上、環境耐性の向上、向上された又は新たな機能性を有した生物を育種する。

[0068] 例えば、雑種第一代（F1）作物に関し、PPRタンパク質によるミトコンドリアRNAの安定化や翻訳制御を用いて人工的にF1作物を作出し、収率や品質を改善できる可能性がある。PPRタンパク質によるRNA操作及びゲノム編集は、従来技術よりも正確かつ迅速に、生物の品種改良、育種（生物を遺伝的に改良すること）が可能である。また、PPRタンパク質によるRNA操作及びゲノム編集は、遺伝子組み換えのように外来遺伝子により形質を転換するものではなく、元来動植物が有するRNAやゲノムを扱う技術である点で、変異体の選抜や戻し交雑という旧来の育種の手法に近いといえる。そのため、地球規模の食糧問題、環境問題にも、確実かつ迅速に対応しうる。

[0069] (3) 化学

- ・微生物、培養細胞、植物体、動物体（例えば昆虫体）を利用した有用物質生産において、DNA、RNAの操作により、タンパク発現量を制御する。これにより、有用物質の生産性を向上することができる。有用物質の例は、抗体、ワクチン、酵素等のタンパク質性の物質のほか、医薬品の中間体、香料、色素等の比較的低分子の化合物である。

[0070] ・藻類や微生物の代謝経路の改変により、バイオ燃料の産生効率を改善する。

実施例

[0071] [実施例1：蛍光タンパク質融合PPRタンパク質の細胞内解析]

(モチーフの設計)

標的配列は、CAG配列が6回繰り返したCAGCAGCAGCAGCAGCAG(SEQ ID NO:1)とした。PPRモチーフは、1番目、4番目、ii番目のアミノ酸配列によって認識される塩基が決定される。シトシンを認識するPPRモチーフには、1番目にバリン、4番目にアスパラギン、ii番目にセリン、アデニンを認識するPPRモチーフには、1番目にバリン、4番目にトレオニン、ii番目にアスパラギン、グアニンを認識するPPRモチーフには、1番目にバリン、4番目にトレオニン、ii番目にアスパラギン酸をそれぞれ配置した。なお、ウラシルを認識するPPRモチーフには、1番目はバリン、4番目はアスパラギン、ii番目はアスパラギン酸をそれぞれ配置すればよい。

[0072] さらに、シトシンを認識する1モチーフ目(図1Aの、Mutated motif)の6番目、9番目のアミノ酸に関しては、典型的な組み合わせとして、ロイシンとグリシン(C₆L9G, PPRcag₁、前掲SEQ ID NO:2)、変異型として、ロイシンとグルタミン酸(C₆L9E, PPRcag₂、SEQ ID NO:3)、アスパラギンとグルタミン(C₆N9Q, PPRcag₃、SEQ ID NO:4)、アスパラギンとグルタミン酸(C₆N9E, PPRcag₄、SEQ ID NO:5)、アスパラギンとリジン(C₆N9K, PPRcag₅、SEQ ID NO:6)、アスパラギン酸とグリシン(C₆D9G, PPRcag₆、SEQ ID NO:7)を選択した(図1B)。これらのPPRモチーフ配列を用いて、CAGCAGCAGCAGCAGCAG配列(前掲SEQ ID NO:1)に結合するように並べ、PPR遺伝子を作製した(SEQ ID NOs:11-16)。なお、各々のPPRモチーフをコードする18個のDNAを効率よく正確に連結させるために、シトシン、アデニン、グアニンそれぞれに対するPPRモチーフにおいて、1、4、6、9、ii番目以外のアミノ酸を、上述のように高頻度で見られるアミノ酸とした(SEQ ID NOs:8-9、前掲特許文献1参照)。

[0073] (プラスミドの作製)

PPR遺伝子を含むプラスミドは、Golden Gate法を利用して構築した。より詳細には、順番にシームレスで連結されるように設計された10種類の中間ベ

クターDest-a, b, c, d, e, f, g, h, i, jを準備し、1モチーフ及び2モチーフ（A、C、G、Uに対応するPPRモチーフ、AA、AC、AG、AU、CA、CC、CG、CU、GA、GC、GG、GU、UA、UC、UG、UUそれぞれの塩基組み合わせを認識する2つのPPRモチーフ）の20種類のモチーフ各々を、10種類のベクターに挿入することにより、200種類のパーツを作製した。

- [0074] Dest-aは、gaagacataaactccgtggtcacATACagagaccaaggtctcaGTGGtcacatacatgtcttc(SEQ ID N0:43)、
- Dest-bは、gaagacatATACagagaccaaggtctcaGTGGtgacataatgtcttc(SEQ ID N0:44)、
- Dest-cは、gaagacatcATACagagaccaaggtctcaGTGGttacatatgtcttc(SEQ ID N0:45)、
- Dest-dは、gaagacatacATACagagaccaaggtctcaGTGGttacaatgtcttc(SEQ ID N0:46)、
- Dest-eは、gaagacattacATACagagaccaaggtctcaGTGGtgacatgtcttc(SEQ ID N0:47)、
- Dest-fは、gaagacattgacATACagagaccaaggtctcaGTGGttaatgtcttc(SEQ ID N0:48)、
- Dest-gは、gaagacatgttacATACagagaccaaggtctcaGTGGtcatgtcttc(SEQ ID N0:49)、
- Dest-hは、gaagacatggtcacATACagagaccaaggtctcaGTGGtatgtcttc(SEQ ID N0:50)、
- Dest-iは、gaagacattggttacATACagagaccaaggtctcaGTGGatgtcttc(SEQ ID N0:51)、
- Dest-jは、gaagacatgtggtgacATACagagaccaaggtctcaGTGGtcttc(SEQ ID N0:52)
- を、遺伝子合成技術により作製し、pUC57-kanへクローニングすることにより準備した。

- [0075] 標的塩基配列に沿ってDest-aからDest-jを選択し、Golden Gate反応によりベクターへクローニングした。ここで使用するベクターは、18個連結されたP

PR配列のN末端にMGNSV(SEQ ID NO:53)、C末端にELTYNTLISGLGKAGRARDPPV(SEQ ID NO:54)のアミノ酸配列が付加されるように設計した。正しいサイズの遺伝子がクローニングできていることを確認し、またクローニングされた遺伝子の配列をシーケンシングにて確認した。

[0076] (細胞内での発現の検出)

動物培養細胞での発現プラスミドpcDNA3.1は、CMVプロモーターとSV40 poly Aシグナル配列を含み、それらの間に発現させたい遺伝子を挿入することができる。細胞内でのPPRタンパク質の発現を検出するため、蛍光タンパク質を融合したPPRタンパク質を発現させ、その蛍光画像から細胞内での凝集性及び核への移行性を解析することとした。N末側から、EGFP、核移行シグナル配列、PPRタンパク質、FLAGエピトープタグの順番で融合したタンパク質遺伝子をpcDNA3.1へ挿入した(SEQ ID NOs:17-22)。また、N末側からmClover3、PPRタンパク質、核移行シグナル配列、FLAGエピトープタグの順番で融合したタンパク質遺伝子をpcDNA3.1へ挿入した(SEQ ID NOs:23-28)。コントロールとしてPPRを含まないプラスミドも作製した(SEQ ID NOs:35-36)。

[0077] HEK293T細胞を9 mL DMEM, 1 mL FBS入りの10cm ディッシュに 1×10^6 cells/well播種した。37°C、5%CO₂環境下で2日間培養した後、細胞を回収した。回収した細胞をPLLコーティングされた96ウェルプレートへ1ウェルあたり、 4×10^4 cells/wellで細胞播種し、37°C、5% CO₂環境下で1日間培養した。200 ng プラスミドDNA、0.6 μ L Fugene(登録商標)-HD (Promega, E2311)、200 μ L Opti-MEMを混ぜ合わせ、全量ウェルへ加え、37°C、5%CO₂環境下で1日間培養した。培養後、培地を取り除き、50 μ L PBSで1回洗浄後、1 μ L Hoechst (1mg/mL、同仁化学、346-07951)、50 μ L PBSを加え、37°C、5%CO₂環境下で10分間置き、その後、50 μ L PBSで1回洗浄した。洗浄後、50 μ L PBSを加えて、各ウェルのGFP蛍光及びHoechst蛍光画像を、蛍光顕微鏡 DMi8 (Leica) を用いて取得した。

[0078] 結果を図2に示す。EGFP及び核移行シグナル配列を融合したPPRの細胞内で発現を確認した結果、PPRcag_1 (6L9G)とPPRcag_2(6L9E)においては、核へ局

在せず、核の周りで強く凝集する様子が確認された。一方で、PPRcag_3 (6N9Q), PPRcag_4(6N9E), PPRcag_5(6N9K), PPRcag_6(6D9G)においては、凝集性は低いものの核へ局在しなかった。mClover3を融合した場合、PPRcag_1 (6L9G)とPPRcag_2(6L9E)においては、核へ局在するものの、核の中で凝集する様子が確認された。PPRcag_3 (6N9Q), PPRcag_4(6N9E), PPRcag_5(6N9K), PPRcag_6(6D9G)においては、核へ局在し、かつ凝集性が見られなかった。したがって、凝集性を改善するためには、6N9E, 6N9Q, 6N9K, 6D9G変異が良く、また核へ効率的に局在させるためには、EGFPよりもmClover3を用いる方が良いことがわかった。

[0079] [実施例2：CAG結合PPRタンパク質のRNA結合解析]

PPRcag_1、PPRcag_2、PPRcag_3、PPRcag_4、PPRcag_5、PPRcag_6について、標的RNAとの結合を確認するため、組み換えタンパク質を作製し、結合実験を行った。

[0080] それぞれのPPRタンパク質のN末側にルシフェラーゼ、C末側に6 x ヒスチジンタグ配列を融合したタンパク質遺伝子を設計し、大腸菌発現プラスミドへクローニングした (SEQ ID NOs:29-34)。またコントロールとしてPPRタンパク質を含まないNluc-Hisx6タンパク質遺伝子も作製した (SEQ ID NO:37)。

[0081] 完成したプラスミドを大腸菌Rosetta(DE3)株へ形質転換した。この大腸菌を、2 mLの100 μ g/mLアンピシリン入りLB培地で37°C、12時間培養し、OD₆₀₀が、0.5から0.8に到達した時に、15°Cのインキュベーターに培養液を移し、30分間静置させた。その後、100 μ L (終濃度0.1 mM IPTG) を加え、15°C、16時間培養を行った。5,000 x g, 4°C, 10分間遠心によって、大腸菌ペレットを回収し、1.5 mLの溶解バッファー (20 mM Tris-HCl, pH8.0, 150 mM NaCl, 0.5% NP-40, 1 mM MgCl₂, 2mg/ml リゾチーム, 1 mM PMSF, 2 μ LのDNase) を加え、-80°Cで20分間凍結させた。25°C、30分間浸透しながら細胞の凍結破碎をおこなった。続いて3700 rpm, 4°C, 15分間遠心操作を行い、可溶性のPPRタンパク質を含む上清(大腸菌ライセート)を回収した。

[0082] PPRタンパク質とRNAとの結合実験は、ストレプトアビジンプレート上でのP

PRタンパク質とビオチン化RNAの結合実験方法により行った。

[0083] 標的であるCAGx6配列、標的としないCGGx6、CUGx6、CCGx6、及びD1b(UGGUG UAUCUUGUCUUUA)配列 (SEQ ID NO:42の8-25番) を含む30塩基のRNA(それぞれ順に、SEQ ID NOs:38-42)に5'末端をビオチン修飾したRNAプローブを合成した (Grainer)。2.5 pmolビオチン化RNAプローブをストレプトアビジンコーティングプレート (Cat No. 15502, Thermo fisher) に加え、30分間、室温で反応させ、プローブ洗浄バッファー(20mM Tris-HCl (pH7.6)、150 mM NaCl、5 mM MgCl₂、0.5% NP-40、1mM DTT、0.1% BSA)で洗浄した。バックグラウンド測定のため、ビオチン化RNAを加えずに溶解bufferを加えたウェルも準備した(-Probe)。その後、ブロッキングバッファー(20 mM Tris-HCl (pH 7.6)、150 mM NaCl、5 mM MgCl₂、0.5% NP-40、1mM DTT、1% BSA)を加え、30分間、室温でプレート表面のブロッキングを行った。そこへ、 1.5×10^8 LU/ μ L 発光量を有するルシフェラーゼ融合PPRタンパク質が含まれる大腸菌ライセートを、100 μ L 加え、30分間、室温で結合反応を行った。その後、200 μ Lの洗浄バッファー(20 mM Tris -HCl (pH7.6)、150 mM NaCl、5 mM MgCl₂、0.5% NP-40、1mM DTT)で5回洗浄を行った。洗浄バッファーで、2500倍希釈したルシフェラーゼ基質 (Promega, E151A) 40 μ Lをウェルへ加えて5分間反応させた後、発光量をプレートリーダー (PerkinElmer, Cat No. 5103-35) で測定した。

[0084] その結果を図3に示す。全てのPPRにおいて、標的であるCAGx6へ特異的に結合することがわかった。標的配列への結合力は、PPRcag_1と比較してPPRcag_2は同程度、PPRcag_3は80%程度、PPRcag_4は60%程度、PPRcag_5は120%程度、PPRcag_6は130%程度であった。これらのことからPPRcag_4以外はほとんど変異による結合性能の変化は見られないことがわかった。

[0085] [実施例3：PPRタンパク質の凝集の制御]

V2モチーフを用いたPPRタンパク質 (塩基配列はSEQ ID NO:61、アミノ酸配列はSEQ ID NO:62) と、v3.2モチーフを用いたPPRタンパク質 (塩基配列はSEQ ID NO:63、アミノ酸配列はSEQ ID NO:64) それぞれを大腸菌発現系で作成

し精製し、ゲルろ過クロマトグラフィーで分離した。なお、v2モチーフとは、SEQ ID NO:2、SEQ ID NOs:8~10の配列を有するPPRモチーフを指し、v3.2モチーフとは、N末端から1番目の場合はSEQ ID NO:4、SEQ ID NOs:58~60のいずれかの配列を有するPPRモチーフであり、それ以外の場合はアデニンに対してはSEQ ID NO:8において15番目のアスパラギン酸をリジンに置換したものであり、アデニン以外の塩基に対してはSEQ ID NOs:2, 9, 10から選択した配列を有するPPRモチーフを指す。

[0086] (タンパク質の発現・精製)

目的のPPRをコードしたDNA配列を含むpE-SUMOpro Kanプラスミドを用いて、大腸菌Rosetta株を形質転換し、37度で培養後、OD₆₀₀が0.6に達した時に20度に温度を下げ、終濃度0.5mMになるようにIPTGを加え、目的PPRタンパク質をSUMO融合タンパク質として、大腸菌内で発現させた。1晩、培養後、菌体を遠心によって集菌し、Lysis Buffer (50mM Tris-HCl pH8.0, 500mM NaCl)で再懸濁した。超音波破碎によって、大腸菌を破碎し、17000g, 30minの遠心後、上清画分をNi-Agaroseカラムに供与し、20mM imidazoleを含むLysis Bufferでカラム洗浄後、400mM imidazoleを含むLysis Bufferで、SUMO融合目的PPRタンパク質を溶出させた。溶出後、Ulp1によるSUMOタンパク質を目的PPRタンパク質から切り離すと同時に、透析により、イオン交換Buffer(50mM Tris-HCl pH8.0, 200mM NaCl)にタンパク溶液を置換した。その後、SPカラムを用いた陽イオン交換クロマトグラフィーを行った。カラム供与後、NaCl濃度を200mMから1Mまで徐々に上げることによって、タンパク質を溶出させた。目的PPRタンパク質を含む画分をSuperdex200カラムを用いたゲルろ過クロマトグラフィーによって最終精製を行った。ゲルろ過Buffer (25mM HEPES pH7.5, 200mM NaCl, 0.5mM tris(2-carboxyethyl)phosphine (TCEP)) で平衡化したゲルろ過カラムに、イオン交換から溶出した目的PPRタンパク質を供与した。最後に目的PPRタンパク質を含む画分を濃縮し、液体窒素で凍結し、次の分析に使用するまで-80度で保存した。

[0087] (ゲルろ過クロマトグラフィー)

精製したリコンビナントPPRタンパク質を濃度1mg/mlに調整した。ゲルろ過クロマトグラフィーは、Superdex 200 increase 10/300 GL (GE Helthcare)を用いた。25mM HEPES pH7.5, 200mM NaCl, 0.5mM tris(2-carboxyethyl)phosphine (TCEP)で平衡化したゲルろ過カラムに、調整したタンパク質を供与し、ゲルろ過カラムより溶出してくる溶液を280nmの吸光度を測定することにより、タンパク質の性質を分析した。

[0088] (結果)

結果を図4に示す。溶出画分 (Elution vol.) が少ないほど分子サイズが大きい。V2では、8から10mLの溶出画分で溶出された一方で、v3.2では、12から14mLの溶出画分にピークが見られた。このことから、v2では、タンパク質サイズが大きくなっていることから凝集している可能性が示唆され、その凝集はv3.2において改善されていることが分かった。

請求の範囲

[請求項1]

下記のいずれか1つの、PPRモチーフ：

(C-1)配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2)配列番号：4～7のいずれか1の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(A-2)(A-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつアデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつアデニン結合性であるPPRモチーフ；

(G-1)配列番号：9の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつグアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列からなるPPRモチーフにおいて、位置6のアミノ酸をアスパラギン又はアスパラギン酸に置換したPPRモチーフ；

(U-2)(U-1)の配列において、位置1、4、6、及び34のアミノ酸以外の

アミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウ
ラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置
1、4、6、及び34のアミノ酸は同一であり、かつウラシル結合性であ
るPPRモチーフ。

[請求項2]

下記のいずれか1つの、PPRモチーフ：

(C-1)配列番号：4～7のいずれか1の配列からなるPPRモチーフ；

(C-2)配列番号：4～7のいずれか1の配列において、位置1、4、6、9、
及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加し
た配列からなり、かつシトシン結合性であるPPRモチーフ；

(C-3)配列番号：4～7のいずれか1の配列と少なくとも80%の配列同一
性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、
かつシトシン結合性であるPPRモチーフ；

(A-1)配列番号：8の配列において、位置6及び9のアミノ酸を、下記に
定義した組み合わせのいずれか1つを満たすように置換したPPRモチー
フ；

(A-2)(A-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外
のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつ
アデニン結合性であるPPRモチーフ；

(A-3)(A-1)の配列と少なくとも80%の配列同一性を有し、ただし位置
1、4、6、9、及び34のアミノ酸は同一であり、かつアデニン結合性で
あるPPRモチーフ；

(G-1)配列番号：9の配列において、位置6及び9のアミノ酸を、下記に
定義した組み合わせのいずれか1つを満たすように置換した配列から
なるPPRモチーフ；

(G-2)(G-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外
のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつ
グアニン結合性であるPPRモチーフ；

(G-3)(G-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつグアニン結合性であるPPRモチーフ；

(U-1)配列番号：10の配列において、位置6及び9のアミノ酸を、下記に定義した組み合わせのいずれか1つを満たすように置換した配列からなるPPRモチーフ；

(U-2)(U-1)の配列において、位置1、4、6、9、及び34のアミノ酸以外のアミノ酸の1～9個を置換、欠失、又は付加した配列からなり、かつウラシル結合性であるPPRモチーフ；

(U-3)(U-1)の配列と少なくとも80%の配列同一性を有し、ただし位置1、4、6、9、及び34のアミノ酸は同一であり、かつウラシル結合性であるPPRモチーフ。

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミン酸である組み合わせ

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がグルタミンである組み合わせ

- ・位置6のアミノ酸がアスパラギンであり、かつ位置9のアミノ酸がリジンである組み合わせ

- ・位置6のアミノ酸アスパラギン酸であり、かつ位置9アミノ酸がグリシンである組み合わせ

[請求項3] 下記のいずれか1つである、請求項1又は2に記載のPPRモチーフ：

(C-4)配列番号：4の配列からなるPPRモチーフ；

(A-4)配列番号：58の配列からなるPPRモチーフ；

(G-4)配列番号：59の配列からなるPPRモチーフ；

(U-4)配列番号：60の配列からなるPPRモチーフ。

[請求項4] 請求項1～3のいずれか1項に記載のPPRモチーフの、PPRタンパク質におけるN末端から1番目のPPRモチーフとしての使用。

[請求項5] PPRタンパク質の凝集性を減少させるための、請求項4に記載の使用

。

[請求項6] 下記の式1で表されるPPRモチーフを1～30個含む、特定の塩基配列を有する標的核酸と結合可能なタンパク質において、N末端から1番目のPPRモチーフ (M_1) の A_6 アミノ酸が、親水性アミノ酸である、タンパク質。

[化1]

(Helix A)-X-(Helix B)-L (式1)

(式中：

Helix Aは、12アミノ酸長の、 α ヘリックス構造を形成可能な部分であって、式2で表され、

[化2]

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

式2中、 $A_1 \sim A_{12}$ はそれぞれ独立にアミノ酸を表し；

Xは、存在しないか又は1～9アミノ酸長からなる部分であり；

Helix Bは、11～13アミノ酸長からなる、 α ヘリックス構造を形成可能な部分であり；

Lは、2～7アミノ酸長の、式3で表される部分であり；

[化3]

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

式3中、各アミノ酸は、“i” (-1)、“ii” (-2)、とC末端側からナンバリングされ、

ただし、 $L_{iii} \sim L_{vii}$ は存在しない場合がある。)

[請求項7] M_1 の A_9 アミノ酸が、親水性アミノ酸又はグリシンである、請求項6に記載のタンパク質。

[請求項8] M_1 の A_6 アミノ酸が、アスパラギン又はアスパラギン酸である、請求項6又は7に記載のタンパク質。

[請求項9] M_1 の A_9 アミノ酸が、グルタミン、グルタミン酸、リジン、又はグリ

シンである、請求項6～8のいずれか1項に記載のタンパク質。

- [請求項10] M_1 の A_6 アミノ酸、及び M_1 の A_9 アミノ酸が、下記のいずれかの組み合わせである、請求項6～9のいずれか1項に記載のタンパク質。
- ・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミン酸である組み合わせ
 - ・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がグルタミンである組み合わせ
 - ・ A_6 アミノ酸がアスパラギンであり、かつ A_9 アミノ酸がリジンである組み合わせ
 - ・ A_6 アミノ酸アスパラギン酸であり、かつ A_9 アミノ酸がグリシンである組み合わせ
- [請求項11] 蛍光タンパク質、核移行シグナルペプチド、及びタグタンパク質からなる群より選択される少なくとも一つと、請求項1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は請求項6～10のいずれか1項に記載のタンパク質との、融合タンパク質。
- [請求項12] 請求項6に定義したPPRモチーフを含む、特定の塩基配列を有する標的核酸と結合可能なPPRタンパク質の、改質方法であって、N末端から1番目のPPRモチーフ (M_1) の A_6 アミノ酸をより親水性とする、方法。
- [請求項13] 請求項1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、請求項6～10のいずれか1項に記載のタンパク質、又は請求項11に記載の融合タンパク質を用いることを特徴とする、核酸の検出方法。
- [請求項14] 請求項1～3のいずれか1項に記載のPPRモチーフ、請求項1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は請求項6～10のいずれか1項に記載のタンパク質をコードする核酸。
- [請求項15] 請求項14に記載の核酸を含む、ベクター。

- [請求項16] 請求項15に記載のベクターを含む、細胞（ヒト個体は除く。）。
- [請求項17] 請求項1～3のいずれか1項に記載のPPRモチーフ、請求項1～3のいずれか1項に記載のPPRモチーフをN末端から1番目のPPRモチーフとして含むPPRタンパク質、又は請求項6～10のいずれか1項に記載のタンパク質、又は請求項15に記載のベクターを用いる、核酸の操作方法（ヒト個体での実施を除く。）。
- [請求項18] 請求項17に記載の操作方法を含む、生物の生産方法。