

Engineering of D- and T-arms of tRNA for enhancing incorporation of D-amino acid and β -amino acid

Technical Field

[0001] The present invention relates to engineering of D- and T-arms of tRNA for enhancing the incorporation of non-proteinogenic amino acids.

Background Art

[0002] In nature, a ribosomal translation system makes use of 19 proteinogenic L-amino acids and glycine, but non-proteinogenic amino acids such as D-amino acids and β -amino acids are extruded from peptide synthesis or protein synthesis by the ribosomal translation system. D-amino acids and β -amino acids are found in various natural peptides and they are usually synthesized in a ribosomal-independent manner.

Various methods for introducing these non-proteinogenic amino acids are therefore being developed.

For example, non-proteinogenic amino acids such as D-amino acids, N-methylamino acids, β -amino acids, and α -hydroxy acids are introduced by genetic code reprogramming combined with a reconstructed in vitro translation system such as Flexible in vitro translation (FIT) system (Non-Patent Documents 1 to 9).

[0003] In spite of success in introduction of some D-amino acids into a peptide chain as disclosed in Non-Patent Documents 4, 5 and from 10 to 12, consecutive introduction of D-amino acids has still a large problem (Non-Patent Documents 4 and 12).

Non-Patent Document 5 discloses use of tRNA^{GluE2} pre-charged with D amino acids prepared using a flexizyme to achieve the consecutive incorporation of the D-amino acids by means of the FIT system. In the method disclosed in Non-Patent Document 5, a higher concentration of EF-Tu is presumed to contribute to the enhancement of accommodation of D-aminoacyl-tRNA. However, the overall expression level of a peptide having two consecutive D-Alas was less than 0.3 μ M and was still lower than that of total L-type peptides, that is, about 1.4 μ M.

[0004] With regards to β -amino acids, in spite of success in introducing a specific β -amino acid into a peptide chain, there remains a large problem in consecutive introduction of β -amino acids (Non-Patent Documents 8, and from 13 to 18).

Citation List

Non-Patent Documents

[0005] Non-Patent Document 1: Murakami, H. et al. Nat. Methods 3, 357-359 (2006).

Non-Patent Document 2: Goto, Y. et al. Nat. Protoc. 6, 779-790 (2011).

Non-Patent Document 3: Goto, Y. et al. RNA 14, 1390-1398 (2008).

Non-Patent Document 4: Fujino, T. et al. J. Am. Chem. Soc. 135, 1830-1837 (2013).

Non-Patent Document 5: Katoh, T. et al. Cell Chem. Biol. 24, 46-54 (2017).

Non-Patent Document 6: Kawakami, T. et al. Chem Biol 15, 32-42 (2008).

Non-Patent Document 7: Kawakami, T. et al. J Am Chem Soc 135, 12297-12304 (2013).

Non-Patent Document 8: Fujino, T. et al. J Am Chem Soc 138, 1962-1969 (2016).

Non-Patent Document 9: Ohta, A. et al. Chem Biol 14, 1315-1322 (2007).

Non-Patent Document 10: Dedkova, L.M. et al. J. Am. Chem. Soc. 125, 6616-6617 (2003).

Non-Patent Document 11: Dedkova, L.M. et al. Biochemistry 45, 15541-15551 (2006).

Non-Patent Document 12: Achenbach, J. et al. Nucleic Acids Res 43, 5687-5698 (2015).

Non-Patent Document 13: Heckler, T. G. et al. J. Biol. Chem. 258, 4492-4495 (1983).

Non-Patent Document 14: Maini, R. et al. Biochemistry 54, 3694-3706 (2015).

Non-Patent Document 15: Iqbal, E. S. et al. Org. Biomol. Chem. 16, 1073-1078 (2018).

Non-Patent Document 16: Dedkova, L. M. et al. Biochemistry 51, 401-415 (2011).

Non-Patent Document 17: Maini, R. et al. Bioorg. Med. Chem. 21, 1088-1096 (2013).

Non-Patent Document 18: Czekster, C. M. et al. J. Am. Chem. Soc. 138, 5194-5197 (2016).

Summary

Technical Problem

[0006] The problem to be solved by the present invention is to provide, for translation in a cell-free translation system, a novel translation system capable of synthesizing peptides having therein consecutive non-proteinogenic amino acids.

Solution to Problem

[0007] As a result of intensive investigation, the present inventors have found that they can solve the above-described problem, paying attention to the D- and T-arms of aminoacyl-tRNA and completed the present invention.

[0008] The present invention is as follows:

(1) A tRNA containing a base sequence represented by SEQ ID NO: 1 and encoding a non-proteinogenic amino acid.

N₁N₂GCN₃N₄N₅N₆N₇N₈N₉N₁₀N₁₁GCN₁₂N₁₃ (SEQ ID NO: 1)

(in SEQ ID NO: 1,

N₁ to N₁₃ each represents an arbitrary base, N₃ to N₁₁ form a D-loop, and N₁N₂GC forms a base pair with N₁₃N₁₂CG).

(2) The tRNA as described above in (1), further containing a base sequence represented by SEQ ID NO: 2.

AGGGG(N₁₄)_mCCCCU (SEQ ID NO: 2)

(in SEQ ID NO: 2,

N₁₄ represents an arbitrary base, m stands for an integer of 1 or more, (N₁₄)_m form a T-loop, and AGGGG forms a base pair with UCCCC).

(3) The tRNA as described above in (1) or (2), wherein a non-proteinogenic amino acid charged at a 3' terminal is a D-amino acid, a β-amino acid, or an α,α-disubstituted amino acid.

(4) A translation system for the synthesis of a peptide having two or more consecutive non-proteinogenic amino acids therein, containing the tRNA described above in any of (1) to (3).

(5) A method of constructing a peptide library, including a step of translating in a cell-free translation system by using the tRNA as described above in any of (1) to (3).

(6) A method of constructing a library of a complex between a peptide and an mRNA encoding the peptide, including a step of translating in a cell-free translation system by using the tRNA as described above in any of (1) to (3).

(7) A peptide library constructed by the method as described above in (5) or a peptide-mRNA complex library constructed by the method as described above in (6).

(8) The peptide library or peptide-mRNA complex library as described above in (7), including a peptide having two or more consecutive non-proteinogenic amino acids.

(9) The peptide library or peptide-mRNA complex library as described above in (8), including a peptide having four or more consecutive non-proteinogenic amino acids and being intramolecularly crosslinked between the non-proteinogenic amino acids.

Advantageous Effects of Invention

[0009] The present invention makes it possible to provide, for translation in a cell-free translation system, a novel translation system capable of synthesizing a peptide having consecutive non-proteinogenic amino acids.

Brief Description of Drawings

[0010] Fig. 1 is a schematic view showing promotion of incorporation of proline and D-amino acid by EF-P. Fig. 1A shows EF-P, a prokaryotic translation factor that promotes peptide bond formation between two prolines (Pro). EF-P has a role of promoting incorporation of consecutive prolines. EF-P binds to ribosome in between an E site and a P site, interacts with the peptidyl-prolyl-tRNA at P site, and promotes

peptidyl transfer between the peptidyl-prolyl-tRNA at P site and the prolyl-tRNA at A site. Fig. 1B shows the structure of tRNA^{Pro1E2} having a chimeric structure of D-arm of tRNA^{Pro1} and T-stem of tRNA^{GluE2}.

Fig. 2 shows promotion of D-amino acid incorporation by EF-P. Fig. 2A shows the mRNA sequence of mR1 and mR2 and the peptide sequence of rP1 and rP2 respectively corresponding thereto, each used in Examples. An arbitrarily selected codon is introduced into NNN and is used for incorporation of D-amino acid (X) by using a pre-charged D-aminoacyl-tRNA. The amino acid sequence of flag is DYKDDDDK. Fig. 2B shows the structure of tRNA used for incorporation of D-amino acid at CCG codon. A wild type tRNA^{PRO1}_{CCG} (WT) has a C/G base pair at positions 13 and 22. A tRNA^{Pro1}_{CCG} (C13G) mutant has C-to-G mutation at position 13 to break formation of the base pair. The sequence of an anticodon loop can be arbitrarily changed for decoding different codons. Fig. 2C shows an expression level of rP1 peptide having two D-alanines introduced therein. Codons used for D-alanine incorporation are shown. Black bars indicate EF-P(+) translation system and white outlined bars show EF-P(-) translation system. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-). The EF-P concentration in the EF-P(+) translation system is 5 μ M. Translation time is 15 minutes and error bars are s. d. (n = 3). Fig. 2D and Fig. 2E show the expression level of rP1 peptide (Fig. 2D) and rP2 peptide (Fig. 2E) containing two or three consecutive D-amino acids or Aibs (2-aminobutyric acids). Wild-type tRNA^{PRO1}_{CCG} (WT) (plain bars) or a C13G mutant of tRNA^{Pro1}_{CCG} (dotted bars) were used for incorporation of D-amino acids or Aibs at CCG codon. Black bars indicate EF-P(+) translation system and white outlined bars indicate EF-P(-) translation system. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-). The concentration of EF-P in the EF-P(+) translation system was 5 μ M. Translation time is 15 minutes

and error bars are s. d. ($n = 3$). An enlarged graph showing incorporation of D-Histidine and Aib is shown in the upper right side of Fig. 2D. Fig. 2D and Fig. 2E show that EF-P recognizes P-site peptidyl D-aminoacyl-tRNA as well as recognizing peptidyl-L-prolyl-tRNA.

Fig. 3A shows the results of tricine SDS-PAGE analysis of rP1 peptide having two consecutive D-alanines (refer to Fig. 2C). Fig. 3B shows an expression level of rP1 peptide and rP2 peptide containing 2 or 3 consecutive L-prolines. Wild type $\text{tRNA}^{\text{Pro1}}_{\text{CCG}}$ (WT) (normal bar) or a C13G mutant of $\text{tRNA}^{\text{Pro1}}_{\text{CCG}}$ (dot bar) was used for L-proline incorporation at CCG codon. Black bars indicate EF-P(+) translation system and white outlined bars show EF-P(-) translation system. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-). The EF-P concentration in the EF-P(+) translation system was 5 μM . Translation time is 15 minutes and error bars are s. d. ($n = 3$). Fig. 3C and Fig. 3D show the results of tricine SDS-PAGE analysis of rP1 peptide and rP2 peptide having two or three consecutive L-prolines (Fig. 3C) or D-amino acids or Aibs (Fig. 3D). Translation was performed in a 2.5 μL of a reaction solution at 37°C for 15 minutes in the presence or absence of EF-P (refer to Figs. 2D and 2E). Fig. 3E shows the MALDI-TOF-MS spectrum of rP1 peptide and rP2 peptide. Wild type $\text{tRNA}^{\text{Pro1}}$ was used for incorporation of L-proline, D-amino acid, and Aib. The arrows show peaks corresponding to the intended products. The calculated value and the observed value are indicated by m/z .

Fig. 4 shows the titration of EF-P and time-course analysis in the translation of a D-Ala-containing peptide. Fig. 4A shows the titration of an EF-P concentration in the translation of rP1 peptide containing two consecutive D-alanines at CCG codons. The translation time was set at 40 minutes (black circle) or 15 minutes (black square).

Error bars are s. d. (n=3). Fig. 4B shows the time-course analysis of the translation of rP1 peptide containing two consecutive D-alanines. The black circle and black square show the results in the EF-P(+) and EF-P(-) translation systems, respectively. The concentration of EF-P in the EF-P(+) translation system was 5 μ M. Error bars are s. d. (n=3).

Fig. 5 shows the effect of a tRNA structure on the EF-P-dependent promotion of D-amino acid incorporation. Fig. 5A shows the structure of tRNAs used for D-amino acid incorporation. The pale letters in tRNA^{Pro1E1}_{CCG} and tRNA^{Pro1E2}_{CCG} indicate a mutation from a wild-type tRNA^{Pro1}_{CCG} sequence (An Escherichia coli prolyl-tRNA synthase cannot recognize tRNAs having such mutation introduced therein). Fig. 5B shows an expression level of rP1 peptide containing two consecutive D-alanines at CCG codon. tRNA used for D-Ala incorporation is indicated. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-). Translation time is 15 minutes and error bars are s. d. (n = 3).

Fig. 6A shows the tricine SDS-PAGE analysis results of rP1 peptide synthesized using various tRNAs for incorporating two consecutive D-alanines therein. Translation was performed in a 2.5 μ L of a reaction solution at 37°C for 15 minutes in the presence or absence of EF-P (refer to Fig. 5B). Fig. 6B shows the tricine SDS-PAGE analysis results of rP3, rP4, rP5, rP6, and rP7 peptides. tRNA^{Pro1E2} was used for D-amino acid incorporation. Translation was performed in a 2.5 μ L of a reaction solution at 37°C for 15 minutes in the presence or absence of EF-P (refer to Fig. 7B). Fig. 6C shows the MALDI-TOF-MS spectrum of rP3, rP4, rP5, rP6, and rP7 peptides synthesized using 5 μ M EF-P and tRNA^{Pro1E2}. The arrows show peaks corresponding to the intended products. The calculated value and the observed value are indicated by m/z. It is to be noted that rP7 peptide was detected mainly as

a 2-mercaptoethanol adduct to D-cysteine.

Fig. 7 shows D-amino acid incorporation into various peptide sequences promoted by EF-P. Fig. 7A shows mRNA sequences (mR3, mR4, mR5, mR6, and mR7) and peptide sequences (rP3, rP4, rP5, rP6, and rP7) corresponding thereto. Fig. 7B shows the expression level of peptides rP3, rP4, rP5, rP6, and rP7 for which a tRNA^{Pro1E2} mutant was used for incorporation of a D-amino acid. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-). Translation time is 15 minutes and error bars are s. d. (n = 3).

Fig. 8 shows translation of a cyclic D-peptide formed by a disulfide bond or a thioether bond. Fig. 8A shows mRNA sequences (mR8 and mR9) and peptide sequences (rP8 and rP9) corresponding thereto. Two D-cysteines contained in rP8 peptide form a disulfide bond to form a macrocyclic structure. The sulfhydryl group of the D-Cysteine in rP9 peptide attacks the α -carbon of the N-terminal chloroacetyl (ClAc) group to form a thioether bond and thereby form a macrocyclic structure. Fig. 8B and Fig. 8C show the expression level of rP8 peptide (Fig. 8B) and rP9 peptide (Fig. 8C) for which tRNA^{Pro1E2} or tRNA^{GluE2} was used for D-amino acid incorporation. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-). Translation time is 15 minutes and error bars are s. d. (n = 3).

Fig. 9A and Fig. 9B show the tricine SDS-PAGE analysis results of rP8 peptide (Fig. 9A) and rP9 peptide (Fig. 9B). For D-amino acid incorporation, tRNA^{Pro1E2} or tRNA^{GluE2} was used. Translation was performed in a 2.5 μ L of a reaction solution at 37°C for 15 minutes in the presence or absence of EF-P (refer to Figs. 8B and 8C). Fig. 9C shows the MALDI-TOF-MS spectrum of rP8 and rP9 peptides synthesized using 5 μ M EF-P and tRNA^{Pro1E2}. The arrows show peaks corresponding to the intended products. The calculated value and the observed value are indicated by

m/z.

Fig. 10 shows promotion of β -amino acid incorporation by EF-P. Fig. 10a shows the structure of tRNA used for β -amino acid incorporation at CCG codon. Fig. 10b shows the structure of a β -amino acid. Fig. 10c shows, similar to Fig. 2A, the mRNA sequence of mR1 used in Examples for β -amino acid incorporation and the peptide sequence of rP1 corresponding thereto. The CCG codon is used for β -amino acid (Xaa) incorporation by using a pre-charged β -aminoacyl-tRNA. The amino acid sequence of flag is DYKDDDDK. Fig. 10d shows the expression level of rP1 peptide containing two consecutive β -amino acids or D-alanine. Wild-type tRNA^{PRO1}_{CCG} (WT) (plain bars) or tRNA^{Pro1}_{CCG}, that is, a C13G mutant (dotted bars) were used for incorporation of a β -amino acid or D-alanine at CCG codon. Black bars indicate an EF-P(+) translation system and white outlined bars indicate an EF-P(–) translation system. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(–). The concentration of EF-P in the EF-P(+) translation system was 5 μ M. Translation time is 15 minutes and error bars are s. d. (n = 3). Fig. 10e and Fig. 10f show the titration of an EF-P concentration in the translation of rP1 peptide containing two consecutive β -amino acids. For the titration, lysylated (Lysylated) EF-P (black circle) or unmodified EF-P (white circle) were used. Error bars are s. d. (n=3).

Fig. 11 shows the effect of a tRNA structure on the EF-P dependent promotion of β -amino acid incorporation. Fig. 11a shows the structure of tRNAs used for β -amino acid incorporation. The dot lines show a T-stem motif or a D-arm motif optimized for EF-Tu and EF-P binding. Fig. 11b shows the expression level of a rP1 peptide containing two consecutive β hM. tRNAs used for β hM incorporation are shown. Numbers above the bars mean a translation yield calculated as a ratio of

EF-P(+) to EF-P(-). Translation time is 15 minutes and error bars are s. d. (n = 3). Fig. 11c shows the structure of tRNA^{Pro1E2} having a D-arm motif and a T-stem motif optimized for β -amino acid incorporation.

Fig. 12 shows incorporation of up to 7 consecutive β -amino acids. Fig. 12a shows, similar to Fig. 2A, the mRNA sequence of the mR2 used in Examples for β -amino acid incorporation and the peptide sequence of rP2 corresponding thereto. From 2 to 7 consecutive β hMs were incorporated at CCG codons (n=2 to 7) by tRNA^{Pro1E2}_{CCG}. Fig. 12b shows the expression level of rP2 peptide containing consecutive β hMs. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-) or an expression level (μ M). N.D. means “not detected”. Translation time is 15 minutes and error bars are s. d. (n = 3). Fig. 12c shows the MALDI-TOF-MS spectrum of rP2 peptide having from 2 to 7 consecutive β hMs incorporated therein. In the translation system, 5 μ M EF-P was present. The arrows show peaks corresponding to the intended products. The calculated value and the observed value are indicated by m/z.

Fig. 13 shows incorporation of two kinds of β -amino acids. Fig. 13a shows, similar to Fig. 2A, the mRNA sequence of from mR3-3 to mR3-7 used in Examples for β -amino acid incorporation and the peptide sequence of from rP3-3 to rP3-7 corresponding thereto. β hM and β hF were incorporated at AUU codon and CAU codon by using tRNA^{Pro1E2}_{Gau} and tRNA^{Pro1E2}_{GUG}, respectively. Fig. 13b shows the expression level of from rP3-3 peptide to rP3-7 peptide. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-) or an expression level (μ M). N.D. means “not detected”. Translation time is 15 minutes and error bars are s. d. (n = 3). Fig. 13c shows the MALDI-TOF-MS spectrum of rP3-3 peptide to rP3-7 peptide having β hM and β hF incorporated therein. In the translation system,

5 μ M EF-P was present. The arrows show peaks corresponding to the intended products. The calculated value and the observed value are indicated by m/z.

Fig. 14 shows incorporation of a plurality of D-amino acids and β -amino acids. Fig. 14a shows, similar to Fig. 2A, the mRNA sequence of from mR4 to mR11 used in Examples for incorporation of D-amino acids and β -amino acids and the peptide sequence of from rP4 to rP11 corresponding thereto. β hM and β hF were incorporated at AUU codon and CAU codon by using tRNA^{Pro1E2}_{GAU} and tRNA^{Pro1E2}_{GUG}, respectively whereas D-amino acids were incorporated at ACU codon by using tRNA^{Pro1E2}_{GGU}. Fig. 14b shows the structure of a macrocyclic peptide rP11. The thiol group of D-cysteine reacts with the N-terminal chloroacetyl group to form a thioether bond and thereby produce a macrocyclic structure. Fig. 14c and Fig. 14d show the expression level of from rP4 peptide to rP9 peptide and that of from rP10 peptide to rP11 peptide, respectively. Numbers above the bars mean a translation yield calculated as a ratio of EF-P(+) to EF-P(-) or an expression level (μ M). N.D. means “not detected”. Translation time is 15 minutes and error bars are s. d. (n = 3).

Fig. 15a shows the results of tricine SDS-PAGE analysis of a rP1 peptide having β -amino acid(s) or D-alanine(s) incorporated therein. tRNAs used for the incorporation of these amino acids are shown. Refer to Fig. 10d about quantification of each band. Fig. 15b shows the MALDI-TOF-MS spectrum of rP1 peptide. Wild type tRNA^{Pro1} was used for the incorporation of these amino acids. In the translation system, 5 μ M EF-P was present. The arrows show peaks corresponding to the intended products. The calculated value and the observed value are indicated by m/z.

Fig. 16a to Fig. 16d respectively show the tricine SDS-PAGE analysis results of rP1 peptide (Fig. 16a), rP2 peptide (Fig. 16b), from rP3-3 peptide to rP3-7 peptide

(Fig. 16c), and from rP4 peptide to rP11 peptide (Fig. 16d). Refer to Fig. 11b, Fig. 12b, Fig. 13b, and Fig. 14c and Fig. 14d about quantification of each band. Fig. 16e shows the MALDI-TOF-MS spectrum of from rP4 peptide to rP11 peptide. Wild type tRNA^{Pro1} was used for the incorporation of these amino acids. In the translation system, 5 μ M EF-P was present. The arrows show peaks corresponding to the intended products. The calculated value and the observed value are indicated by m/z.

Description of Embodiments

[0011] The tRNA of the present invention is a tRNA containing a base sequence represented by SEQ ID NO: 1 and encoding a non-proteinogenic amino acids.

N₁N₂GCN₃N₄N₅N₆N₇N₈N₉N₁₀N₁₁GCN₁₂N₁₃ (SEQ ID NO: 1)

In SEQ ID NO: 1,

N₁ to N₁₃ each represents an arbitrary base, N₃ to N₁₁ form a D-loop, and N₁N₂GC forms base pairs with N₁₃N₁₂CG.

[0012] An EF-P protein which is a translation factor has an important role in promoting elongation of proline-proline (Pro-Pro) bond formation and alleviates stalling of translation elongation (ribosomal stalling) caused by peptidyl-tRNA drop-off at the Pro-Pro sequence.

The EF-P protein recognizes a 9-nt D-loop having a structure closed with a stable 4-bp D-stem sequence, which is a specific D-arm motif found in tRNA^{Pro} isoacceptors, for promotion of proline-selective peptidyl transfer and depends on tRNA^{Pro} in the promotion of a Pro-Pro transfer reaction rate.

In the present invention, the base sequence represented by SEQ ID NO: 1 is a D-arm composed of a D-loop and D-stem and by introducing the base sequence

represented by SEQ ID NO: 1 into tRNA, peptidyl transfer by EF-P peptide can be promoted.

[0013] Another structure in the tRNA of the present invention containing a base sequence represented by SEQ ID NO: 1, that is, an acceptor stem, an anticodon stem, an anticodon loop, a variable loop, and T arm may have any base sequences. Of these, the anticodon loop may have, as needed, a base sequence corresponding to a codon to which a non-proteinogenic amino acid is assigned.

The tRNA, at the 3' terminal thereof, has a CCA sequence and binds to an arbitrary amino acid. The tRNA of the present invention binds to a non-proteinogenic amino acid.

The term "tRNA encodes a non-proteinogenic amino acid" as used herein means that the tRNA binds to the non-proteinogenic amino acid via the CCA sequence of the tRNA.

[0014] In SEQ ID NO: 1, any base can be selected insofar as N₁N₂ and N₁₃N₁₂ form a base pair, but N¹ and N¹³ are preferably G and C, respectively and N₂ and N₁₂ are preferably C and G, respectively.

The base sequence represented by SEQ ID NO: 1 is preferably a base sequence represented by SEQ ID NO: 3.

GCGCN₃N₄N₅N₆N₇N₈N₉N₁₀N₁₁GCGC (SEQ ID NO: 3)

In SEQ ID NO: 3

N₃ to N₁₁ have the same meanings as described above and GCGC forms a base pair with CGCG.

[0015] The base sequence of from N₃ to N₁₁ in SEQ ID NO: 1 and SEQ ID NO: 3 is preferably a base sequence represented by SEQ ID NO: 4.

AGCCUGGUA (SEQ ID NO: 4)

The base sequence represented by SEQ ID NO: 4 forms a D-loop in tRNA.

In SEQ ID NO: 4, one base or some bases may be substituted.

The term “some bases may be substituted” means that in 9 bases forming the D-loop, 2, 3, 4, 5, 6, 7, 8, or 9 bases may be substituted. In the 9 bases forming the D-loop, from 1 to 4 bases may be substituted, from 1 to 3 bases may be substituted, from 1 to 2 bases may be substituted, or one base may be substituted.

[0016] The base sequence represented by SEQ ID NO: 1 is preferably a base sequence represented by SEQ ID NO: 5 and is preferably a base sequence represented by SEQ ID NO: 6.

$N_1N_2GCGCAGCCUGGUAGCGCN_{12}N_{13}$ (SEQ ID NO: 5)

In SEQ ID NO: 5,

N_1 , N_2 , N_{12} , and N_{13} have the same meanings as described above and N_1N_2GC forms a base pair with $N_{13}N_{12}CG$.

GCGCGCAGCCUGGUAGCGCGC (SEQ. ID. NO: 6)

In SEQ ID NO: 5 and SEQ ID NO: 6, one base or some bases may be substituted.

The term “in SEQ ID NOS: 5 and 6, some bases may be substituted” means that in SEQ ID NO: 4 showing the 9 bases forming the D-loop, 2, 3, 4, 5, 6, 7, 8, or 9 bases may be substituted. In SEQ ID NO: 4 showing the 9 bases forming the D-loop, from 1 to 4 bases may be substituted, from 1 to 3 bases may be substituted, from 1 to 2 bases may be substituted, or one base may be substituted.

[0017] The T-stem of the aminoacyl-tRNA regulates the interaction with an EF-Tu protein and enhances incorporation of a non-proteinogenic amino acid such as N-methylamino acid.

In the present invention, the base sequence represented by SEQ ID NO: 2 is a T-arm composed of a T-loop and a T-stem and by introducing the base sequence represented by SEQ ID NO: 2 into the tRNA, accommodation by an EF-Tu protein is accelerated when a non-proteinogenic amino acid is introduced into a peptide or a protein.

[0018] The tRNA of the present invention is preferably a tRNA further containing the base sequence represented by SEQ ID NO: 2.

AGGGG(N₁₄)_mCCCCU (SEQ ID NO: 2)

In SEQ ID NO: 2,

N₁₄ represents an arbitrary base, m stands for an integer of 1 or more, (N₁₄)_m form a T-loop, and

AGGGG forms a base pair with UCCCC.

[0019] m is not particularly limited insofar as (N₁₄)_m form a T-loop, and it may be an integer of 2, 3, 4, 5, 6, 7, 8, 9, or 10 and is preferably an integer of from 6 to 8.

(N₁₄)_m are preferably a base sequence represented by SEQ ID NO: 7 derived from the T-loop of tRNA^{GluE2}.

UUCGAAU (SEQ ID NO: 7)

In SEQ ID NO: 7, one base or some bases may be substituted.

The term "some bases may be substituted" means that in 7 bases forming the T-loop, 2, 3, 4, 5, 6, or 7 bases may be substituted. In the 7 bases forming the T-loop,

from 1 to 3 bases may be substituted, from 1 to 2 bases may be substituted, or one base may be substituted.

[0020] In SEQ ID NO: 2, AGGGG and UCCCC which form a base pair form a T-stem.

In the base sequence forming a T-stem, one base or some bases may be substituted insofar as a base pair is formed.

The term “in the T-stem of SEQ ID NO: 2, some bases may be substituted” means that in 10 bases forming the T-stem, 2, 3, 4, 5, 6, 7, 8, 9, or 10 bases may be substituted. In the 10 bases forming the T-stem, 2 bases may be substituted, 4 bases may be substituted, 6 bases may be substituted, or 1, 3, or 5 bases may be substituted insofar as a base pair is formed.

[0021] A tRNA preferred in the present invention contains both the base sequence represented by SEQ ID NO: 1 and the base sequence represented by SEQ ID NO: 2 in order to exhibit a promotion of peptidyl transfer by an EF-P protein and a promotion of accommodation by an EF-Tu protein.

The tRNA of the present invention is preferably a chimera of tRNA^{Pro1} and tRNA^{GluE2} and such a tRNA will hereinafter be called “tRNA^{Pro1E2}”.

More specifically, tRNA^{Pro1} has preferably the T-stem of tRNA^{GluE2} introduced therein.

Using tRNA^{Pro1E2} pre-charged with a non-proteinogenic amino acid such as D-amino acid or β -amino acid and an EF-P protein in combination enables enhancement of the expression level of not only a linear peptide but also a peptide containing such non-proteinogenic amino acids consecutively, for example, four or five consecutive non-proteinogenic amino acids.

[0022] In the tRNA to be used in the present invention, a base sequence other than SEQ ID NO: 1 and/or SEQ ID NO: 2 may be a sequence derived from a wild type tRNA or a sequence derived from a wild type tRNA derived from an Escherichia coli, or it may be an artificial tRNA prepared by in vitro transcription.

[0023] In the present invention, the translation system contains the tRNA of the present invention and because of containing the tRNA, this system is capable of synthesizing a peptide having amino acids generally difficult to be consecutively introduced consecutively bonded to each other. More specifically, it is a translation system capable of synthesizing a peptide having two or more non-proteinogenic amino acids consecutively bound to each other.

The translation system is not particularly limited insofar as it contains the tRNA of the present invention and it contains a component to be used in a cell-free translation system.

The translation system preferably contains an EF-P protein, and more preferably it further contains an EF-Tu protein.

[0024] In the translation system of the present invention, codon reassignment making use of a flexizyme is performed by assigning a non-proteinogenic amino acid to a codon selected as needed as NNN.

In the translation system of the present invention, a natural translation system may be used. In the natural translation system, a tRNA having an anticodon corresponding to each amino acid is present and the tRNA has an intrinsic sequence even in a region other than an anticodon loop.

In the present invention, insofar as the base sequence represented by SEQ ID NO: 1 and/or SEQ ID NO: 2 is introduced, base sequences of an acceptor stem, an anticodon stem, an anticodon loop, or a variable loop may be arbitrary in the sense

that a sequence intrinsic to tRNA in such natural translation system can be assigned as another sequence.

[0025] In the translation system of the present invention, an arbitrary amino acid may be reassigned to all the NNNs (N represents an arbitrary base) by making use of a flexizyme and in this case, all the tRNAs may be artificial. In this case, elongator tRNAs corresponding to each NNN to be added to the translation system may have, as 85% or more or 90% or more of the full length thereof, the same base sequence. Alternatively, a group of elongator tRNAs having almost the same sequences except the sequence of an anticodon can be used.

The term “anticodon loop” as used herein means a loop portion of a single strand of tRNA including an anticodon. The sequence of the anticodon loop can be determined as needed by those skilled in the art so as to complement the codon-anticodon interaction.

[0026] Translation of a peptide in a known cell-free translation system by using the tRNA of the present invention enables synthesis of a peptide having two or more consecutive non-proteinogenic amino acids therein.

The cell-free translation system used above is not particularly limited.

The method of constructing a peptide library can provide a library more abundant in variety because peptides containing a non-proteinogenic amino acid encoded by an appropriate codon and having two or more consecutive non-proteinogenic amino acids can be produced or peptides having two or more non-proteinogenic amino acids can be produced efficiently.

[0027] Provided in the present invention is a method of constructing a peptide library including:

a step of preparing an mRNA library containing mRNAs encoding each peptide of the peptide library, the mRNAs each containing one or more NNNs encoding a non-proteinogenic amino acid; and

a step of translating the mRNAs of the mRNA library in a cell-free translation system having therein a tRNA having an anticodon corresponding to any of the codons of NNN and charged with a non-proteinogenic amino acid corresponding to the codon.

[0028] In the present invention, a method of constructing a peptide-mRNA complex library is performed by binding puromycin to the downstream region of ORF (Open reading frame) of each of the mRNAs upon preparation of the mRNA library in the above-described method of constructing a peptide library. Puromycin may be bound to the mRNA via a linker composed of a peptide or a nucleic acid. By binding puromycin to the downstream region of ORF of the mRNA, a ribosome that has translated the ORF of the mRNA incorporates puromycin therein to form an mRNA-peptide complex. Such a peptide-mRNA complex can link genotype to phenotype and can be applied to in vitro display.

[0029] In the present invention, first, NNN to which a non-proteinogenic amino acid is assigned is selected and a tRNA having an anticodon of the NNN in an anticodon loop and encoding the non-proteinogenic amino acid is prepared using a flexizyme or the like.

Then, by preparing and translating an mRNA library in which each mRNA contains one or more NNNs encoding a non-proteinogenic amino acid, a peptide containing the assigned non-proteinogenic amino acid(s) can be expressed.

[0030] The term “NNN” as used herein means a codon specifying an amino acid and three Ns forming the codon are each independently selected from adenine (A),

guanine (G), cytosine (C), and uracil (U). A mRNA may contain a plurality of NNNs encoding a non-proteinogenic amino acid.

[0031] In the present invention, an arbitrary amino acid may be reassigned to NNN not encoding a non-proteinogenic amino acid or a codon-amino acid relation based on a natural genetic code may be used.

In reassignment, an amino acid having a codon-amino acid relation different from that in a natural genetic code table or that having the same relation may be assigned.

The term “natural genetic code table” means a table showing amino acids assigned to an mRNA triplet (codon) in a living body and it is shown in the following Table 1.

[0032] [Table 1]

Second letter base →									
	U		C		A		G		
	Codon	Amino Acid	Codon	Amino Acid	Codon	Amino Acid	Codon	Amino Acid	
U	UUU	Phenylalanine	UCU	Serine	UAU	Tyrosine	UGU	Cysteine	U
	UUC	Phenylalanine	UCC	Serine	UAC	Tyrosine	UGC	Cysteine	C
	UUA	Leucine	UCA	Serine	UAA	Stop	UGA	Stop	A
	UUG	Leucine	UCG	Serine	UAG	Stop	UGG	Tryptophan	G
C	CUU	Leucine	CCU	Proline	CAU	Histidine	CGU	Arginine	U
	CUC	Leucine	CCC	Proline	CAC	Histidine	CGC	Arginine	C
	CUA	Leucine	CCA	Proline	CAA	Glutamine	CGA	Arginine	A
	CUG	Leucine	CCG	Proline	CAG	Glutamine	CGG	Arginine	G
A	AUU	Isoleucine	ACU	Threonine	AAU	Asparagine	AGU	Serine	U
	AUC	Isoleucine	ACC	Threonine	AAC	Asparagine	AGC	Serine	C
	AUA	Isoleucine	ACA	Threonine	AAA	Lysine	AGA	Arginine	A
	AUG	Methionine	ACG	Threonine	AAG	Lysine	AGG	Arginine	G
G	GUU	Valine	GCU	Alanine	GAU	Aspartic acid	GGU	Glycine	U
	GUC	Valine	GCC	Alanine	GAC	Aspartic acid	GGC	Glycine	C
	GUA	Valine	GCA	Alanine	GAA	Glutamic acid	GGA	Glycine	A
	GUG	Valine	GCG	Alanine	GAG	Glutamic acid	GGG	Glycine	G

[0033] Assignment of an amino acid different from that of the natural genetic code table to each codon is achieved by codon reassignment making use of, for example, an artificial aminoacylation ribozyme flexizyme (Flexizyme). Using the flexizyme

enables aminoacylation of a desired amino acid on a tRNA having an arbitrary anticodon, so that an arbitrary amino acid can be assigned to an arbitrary codon.

[0034] The term “amino acid” as used herein embraces, in addition to proteinogenic amino acids, artificial amino acid mutants or derivatives. Examples include proteinogenic L-amino acids and compounds obtained by chemical synthesis and having properties known per se in the art as characteristics of amino acids.

When the proteinogenic amino acids (proteinogenic amino acids) are represented by a three-letter code known per se in the art, they are Arg, His, Lys, Asp, Glu, Ser, Thr, Asn, Gln, Cys, Gly, Pro, Ala, Ile, Leu, Met, Phe, Trp, Tyr, and Val.

The term “non-proteinogenic amino acids” (non-proteinogenic amino acids) means natural or unnatural amino acids other than the proteinogenic amino acids.

Examples of the unnatural amino acids include α,α -disubstituted amino acids (such as α -methylalanine), N-alkyl- α -amino acids, D-amino acids, β -amino acids, and α -hydroxy acids, each of which has a main chain structure different from that of natural amino acids; amino acids having a side chain structure different from that of natural amino acids (such as norleucine and homohistidine); amino acids having extra methylene in a side chain thereof (such as "homo"amino acids, homophenylalanine, and homohistidine); and amino acids obtained by substituting a carboxylic acid functional group in a side chain thereof by a sulfonic acid group (such as cysteic acid). Specific examples of the unnatural amino acid include amino acids described in WO2015/030014.

The non-proteinogenic amino acid is preferably a D-amino acid, a β -amino acid, or an α,α -disubstituted amino acid, with a D-amino acid or a β -amino acid being more preferred.

[0035] In the present invention, a peptide library including 1×10^6 or more kinds of peptides can be constructed.

The number of amino acids contained in each peptide and encoded by NNN is not particularly limited insofar as a non-proteinogenic amino acid is contained and it can be set at, for example, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, or the like. The position of NNN encoding a non-proteinogenic amino acid in each mRNA is not particularly limited and by using the tRNA of the present invention, a peptide having non-proteinogenic amino acids consecutively arranged therein is translated at the same level with library synthesis of proteinogenic amino acids, even from an mRNA in which NNNs encoding a non-proteinogenic amino acid are consecutively arranged.

[0036] Codon reassignment can make use of a translation system constructed by arbitrarily removing a component of a translation system according to the purpose and reconstituting only the necessary components. For example, when a translation system excluding a specific amino acid is reconstituted, a codon corresponding to the amino acid becomes an empty codon encoding no amino acid. When an arbitrary amino acid is linked to a tRNA having an anticodon complementary to the empty codon by making use of a flexizyme or the like, the linked product is added and translation is performed, the arbitrary amino acid is encoded by the codon and a peptide having the arbitrary amino acid introduced therein instead of the excluded amino acid is translated.

[0037] The term "cell-free translation system" as used herein means a translation system not containing cells. As the cell-free translation system, for example, an *Escherichia coli* extract, a wheat germ extract, a rabbit reticulocyte extract, or an insect cell extract can be used. A re-constituted type cell-free translation system may be used, which is constructed by reconstituting a ribosomal protein,

aminoacyl-tRNA synthetase (aaRS), ribosomal RNA, amino acid, rRNA, GTP, ATP, translation initiation factor (IF), elongation factor (EF), release factor (RF), and ribosome recycling factor (RRF), each of which is purified, and other factor(s) necessary for translation.

The system may contain RNA polymerase for performing transcription from DNA simultaneously. Examples of a commercially available cell-free translation system include Escherichia-coli derived systems such as "RTS-100" (Registered trade mark) of Roche Diagnostics, reconstituted translation systems such as "PURESYSTEM" (Registered trade mark) of PGI and "PURExpressR In Vitro Protein Synthesis Kit" and the like of New England Biolabs, and systems using a wheat germ extract such as those of ZOEGENE Corporation and CellFree Sciences.

As a system using a ribosome of Escherichia coli, for example, a technology described in the following documents is known: H. F. Kung et al., 1977, The Journal of Biological Chemistry Vol. 252, No. 19, 6889-6894; M. C. Gonza et al., 1985, Proceeding of National Academy of Sciences of the United States of America Vol. 82, 1648-1652; M. Y. Pavlov and M. Ehrenberg, 1996, Archives of Biochemistry and Biophysics Vol. 328, No. 1, 9-16; Y. Shimizu et al., 2001, Nature Biotechnology Vol. 19, No. 8, 751-755; H. Ohashi et al., 2007, Biochemical and Biophysical Research Communications Vol. 352, No. 1, 270-276.

The cell-free translation system can provide a high-purity expression product without purification.

It is to be noted that the cell-free translation system of the present invention may be used not only for the translation but also for transcription after addition of a necessary component for transcription.

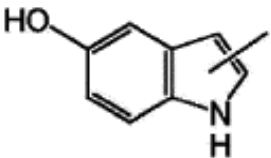
[0038] The peptide obtained in the present invention may be a cyclic peptide. It may be, for example, a cyclic peptide cyclized in combination of an amino acid having a functional group 1 and an amino acid having a functional group 2 corresponding thereto, each shown below in Table 1.

Either of the functional group 1 or 2 may be placed on the N terminal side; they may be placed at the N terminal and the C terminal, respectively; one of them may be a terminal amino acid and the other one may be a non-terminal amino acid; or both may be a non-terminal amino acid.

A linkage formed between the functional group 1 and the functional group 2 is said to be a chemical crosslinking structure for forming a molecular cyclic structure of a cyclic peptide.

[0039]

[Table 2]

	Functional group 1	Functional group 2
(A)	$\begin{array}{c} \text{O} \\ \parallel \\ \text{---C---C---X}_1 \\ \quad \quad \text{H}_2 \end{array} \quad (\text{A-1})$	$\text{HS---} \quad (\text{A-2})$
(B)	$\text{---C}\equiv\text{C---H} \quad (\text{B-1})$	$\text{N}_3\text{---} \quad (\text{B-2})$
(C)	$\text{---Ar---CH}_2\text{NH}_2 \quad (\text{C-1})$	 $\quad (\text{C-2})$
(D)	$\text{---C}\equiv\text{C---CH}_2\text{---X}_1 \quad (\text{D-1})$	$\text{HS---} \quad (\text{D-2})$
(E)	$\text{---Ar---CH}_2\text{---X}_1 \quad (\text{E-1})$	$\text{HS---} \quad (\text{E-2})$

In the formulas, X₁ represents a leaving group and examples of the leaving group include halogen atoms such as Cl, Br, and I and Ar represents a substituted or unsubstituted aromatic ring.

[0040]

As an amino acid having the functional group of (A-1), for example, a chloroacetylated amino acid can be used. Examples of the chloroacetylated amino acid include N-chloroacetyl-L-alanine, N-chloroacetyl-L-phenylalanine, N-chloroacetyl-L-tyrosine, N-chloroacetyl-L-tryptophan, N-3-(2-chloroacetamido)benzoyl-L-phenylalanine, N-3-(2-chloroacetamido)benzoyl-L-tyrosine, N-3-(2-chloroacetamido)benzoyl-L-tryptophan, β -N-chloroacetyl-L-diaminopropanoic acid, γ -N-chloroacetyl-L-diaminobutyric acid, δ -N-chloroacetyl-L-ornithine, and ϵ -N-chloroacetyl-L-lysine, and D-amino acid derivatives corresponding thereto.

As the amino acid having the functional group of (A-1), N-chloroacetyl-L-tryptophan and N-chloroacetyl-L-tyrosine are preferred, with D-forms being more preferred.

The present specification sometimes clearly describes that the amino acid used is an L-form, but it may be either an L-form or a D-form, or it may be a mixture, at any ratio, of an L-form and a D-form. Even when the present specification does not clearly describe that the amino acid used is an L-form or a D-form, it means that it may be either an L-form or a D-form, or it may be a mixture, at any ratio, of an L-form and a D-form.

[0041] Examples of an amino acid having the functional group (A-2) include cysteine, homocysteine, mercaptonorvaline, mercaptonorleucine, 2-amino-7-mercaptoheptanoic acid, and 2-amino-8-mercaptooctanoic acid.

As the amino acid having the functional group (A-2), cysteine is preferred.

[0042] Examples of a cyclization method with the amino acid having the functional group (A-1) and the amino acid having the functional group (A-2) include the methods described, for example, in Kawakami, T. et al., Nature Chemical Biology

5, 888-890 (2009); Yamagishi, Y. et al., ChemBioChem 10, 1469-1472 (2009); Sako, Y. et al., Journal of American Chemical Society 130, 7932-7934 (2008); Goto, Y. et al., ACS Chemical Biology 3, 120-129 (2008); and Kawakami T. et al, Chemistry & Biology 15, 32-42 (2008), and WO2008/117833 .

[0043] As an amino acid having the functional group (B-1), for example, propargylglycine, homopropargylglycine, 2-amino-6-heptynoic acid, 2-amino-7-octynoic acid, and 2-amino-8-nonynoic acid can be used.

A 4-pentynoylated or 5-hexynoylated amino acid may also be used.

Examples of the 4-pentynoylated amino acid include N-(4-pentenoyl)-L-alanine, N-(4-pentenoyl)-L-phenylalanine, N-(4-pentenoyl)-L-tyrosine, N-(4-pentenoyl)-L-tryptophan, N-3-(4-pentynoylamido)benzoyl-L-phenylalanine, N-3-(4-pentynoylamido)benzoyl-L-tyrosine, N-3-(4-pentynoylamido)benzoyl-L-tryptophan, β -N-(4-pentenoyl)-L-diaminopropanoic acid, γ -N-(4-pentenoyl)-L-diaminobutyric acid, σ -N-(4-pentenoyl)-L-ornithine, and ϵ -N-(4-pentenoyl)-L-lysine, and D-amino acid derivatives corresponding thereto.

Examples of the 5-hexynolated amino acid include amino acids obtained by substituting the 4-pentynol group, of the compounds exemplified as the 4-pentynolated amino acid, by a 5-hexynol group.

[0044] Examples of an amino acid having the functional group (B-2) include azidoalanine, 2-amino-4-azidobutanoic acid, azidovaline, azidonorleucine, 2-amino-7-azidoheptanoic acid, and 2-amino-8-azidooctanoic acid.

An azidoacetylated or 3-azidopentanoylated amino acid can also be used.

Examples of the azidoacetylated amino acid include N-azidoacetyl-L-alanine, N-azidoacetyl-L-phenylalanine, N-azidoacetyl-L-tyrosine, N-azidoacetyl-L-tryptophan,

N-3-(4-pentynoylamido)benzoyl-L-phenylalanine,
N-3-(4-pentynoylamido)benzoyl-L-tyrosine,
N-3-(4-pentynoylamido)benzoyl-L-tryptophan, β -N-azidoacetyl-L-diaminopropanoic acid, γ -N-azidoacetyl-L-diaminobutyric acid, σ -N-azidoacetyl-L-ornithine, and ϵ -N-azidoacetyl-L-lysine, and D-amino acid derivatives corresponding thereto.

Examples of the 3-azidopentanoylated amino acid include amino acids obtained by substituting the azidoacetyl group of, the compounds exemplified as the azidoacetylated amino acid, by a 3-azidopentanoyl group.

[0045] Examples of a cyclization method with the amino acid having the functional group (B-1) and the amino acid having the functional group (B-2) include the methods described, for example, in Sako, Y. et al., Journal of American Chemical Society 130, 7932-7934 (2008) and WO2008/117833 .

[0046] Examples of an amino acid having the functional group (C-1) include N-(4-aminomethyl-benzoyl)-phenylalanine (AMBF) and 3-aminomethyltyrosine.

Examples of an amino acid having the functional group (C-2) include 5-hydroxytryptophan (WOH).

Examples of a cyclization method with the amino acid having the functional group (C-1) and the amino acid having the functional group (C-2) include the method described, for example, in Yamagishi, Y. et al., ChemBioChem 10, 1469-1472 (2009) and WO2008/117833 .

[0047] Examples of an amino acid having the functional group (D-1) include 2-amino-6-chloro-hexynoic acid, 2-amino-7-chloro-heptynoic acid, and 2-amino-8-chloro-octynoic acid.

Examples of an amino acid having the functional group (D-2) include cysteine, homocysteine, mercaptonorvaline, mercaptonorleucine, 2-amino-7-mercaptoheptanoic acid, and 2-amino-8-mercaptooctanoic acid.

Examples of a cyclization method with the amino acid having the functional group (D-1) and the amino acid having the functional group (D-2) include the method described, for example, in WO2012/074129.

[0048] Examples of an amino acid (E-1) include N-3-chloromethylbenzoyl-L-phenylalanine, N-3-chloromethylbenzoyl-L-tyrosine, and N-3-chloromethylbenzoyl-L-tryptophan, and D-amino acid derivatives corresponding thereto.

Examples of an amino acid (E-2) include cysteine, homocysteine, mercaptonorvaline, mercaptonorleucine, 2-amino-7-mercaptoheptanoic acid, and 2-amino-8-mercaptooctanoic acid.

A cyclization method with the amino acid having the functional group (E-1) and the amino acid having the functional group (E-2) can be performed with reference to, for example, the cyclization method with the (A-1) and (A-2) or the cyclization method with the (D-1) and (D-2).

[0049] The ring-forming amino acid is preferably a combination of the amino acid having the functional group (A-1) and the amino acid having the functional group (A-2), more preferably a combination of N-acetyltryptophan having a leaving group instead of H thereof and cysteine, much more preferably a combination of N-haloacetyl-D-tyrosine or N-haloacetyl-D-tryptophan, with N-chloroacetyl-D-tyrosine or N-chloroacetyl-D-tryptophan being more preferred, and cysteine (Cys).

[0050] The present invention also provides a peptide library or a peptide-mRNA complex library.

Construction of a peptide library with the tRNA of the present invention facilitates synthesis of a peptide containing a non-proteinogenic amino acid on almost an equal expression level to that of a peptide composed of an L-amino acid. Therefore, compared with a conventional peptide library, a peptide library abundant in variety in terms of the introduction of a non-proteinogenic amino acid can be obtained.

Above all, a library including a peptide having two or more consecutive non-proteinogenic amino acids can be obtained.

Construction of a peptide-mRNA complex library with the tRNA of the present invention facilitates synthesis of a peptide containing a non-proteinogenic amino acid on almost an equal expression level to that of a peptide composed of an L-amino acid. Therefore, compared with a conventional peptide-mRNA complex library, a peptide-mRNA complex library abundant in variety in terms of the introduction of a non-proteinogenic amino acid can be obtained.

Above all, a library including a peptide having two or more consecutive non-proteinogenic amino acids can be obtained.

[0051] In the present invention, a peptide library and a peptide-mRNA complex library having, as a peptide structure thereof, a structure having four or more consecutive non-proteinogenic amino acids and intramolecularly crosslinked between non-proteinogenic amino acids can be obtained.

For intramolecular crosslinking between non-proteinogenic amino acids in a structure having four or more consecutive non-proteinogenic amino acids, the

intramolecular crosslink can be constructed by making use of non-proteinogenic amino acids used upon translation in a cell-free translation system and cyclization.

Preferred examples include a disulfide bond between D-cysteine and D-cysteine and a thioether bond between a non-proteinogenic amino acid having a CIAc group introduced therein and D-cysteine.

[0052] The present invention also provides a screening method using the peptide library constructed using the tRNA of the present invention, for identifying a peptide which binds to a target substance.

The screening method includes a step of bringing the peptide library constructed using the tRNA of the present invention into contact with a target substance and incubating the resulting mixture.

The target substance is not particularly limited herein and low molecular compounds, high molecular compounds, nucleic acids, peptides, proteins, sugars, and lipids can be used.

The screening method further includes a step of selecting a peptide bound to the target substance. The peptide bound to the target substance is selected, for example, by labeling peptides detectably by a known method and after the above-described step of bringing them into contact, washing a surface of a solid phase carrier with a buffer, and detecting a compound bound to the target substance.

Examples of the detectable label include enzymes such as peroxidase and alkaline phosphatase, radioactive substances such as ¹²⁵I, ¹³¹I, ³⁵S, and ³H, fluorescent substances such as fluorescein isothiocyanate, rhodamine, dansyl chloride, phycoerythrin and tetramethyl rhodamine isothiocyanate, and near infrared fluorescent materials, light-emitting substances such as luciferase, luciferin, and aequorin, and nanoparticles such as gold colloid and quantum dot. When the

enzyme is used, detection can be achieved by adding a substrate of the enzyme to develop a color. Detection can also be achieved by binding biotin to a peptide and then binding avidin or streptavidin labeled with the enzyme or the like to the biotin-bound peptide.

[0053] Screening using the peptide-mRNA complex library can be carried out by applying the TRAP display method.

In this case, after the peptide-mRNA complex library is subjected to a reverse transcription reaction, the resulting library is brought into contact with a target substance. A complex that binds to the target substance is selected and its DNA is amplified by PCR. By adding the resulting DNA to a TRAP reaction system, a peptide-mRNA complex library is constructed again. A similar operation is repeated.

Since this concentrates the peptide-mRNA complex having high affinity for the target substance, a peptide that binds to the target substance can be identified efficiently by analyzing the DNA sequence of the concentrated complex.

[0054] Disclosure of all the Non-Patent Documents and Reference Literatures cited herein is incorporated herein by reference in their entirety.

Examples

[0055] The present invention will hereinafter be described in detail by Examples but the present invention is not limited to them. The present invention can be changed by those skilled in the art into various modes without departing from the significance of the present invention and such a change is also embraced within the scope of the present invention.

[0056] Preparation of flexizyme and tRNA

A flexizyme (dFx or eFx) and tRNA (shown in Table 3) used for pre-charging D-amino acid, β -amino acid, 2-aminoisobutyric acid, and L-proline were transcribed by T7 RNA polymerase in vitro.

A template DNA having a T7 promoter before a flexizyme (dFx or eFx) or tRNA sequence was prepared by extension of a pair of forward and reverse extension primers (shown in Table 4) and then, PCR was performed using a pair of forward and reverse extension primers (shown in Table 5).

[0057]

[Table 3]

RNA name	RNA sequence
dFx	GGAUCGAAAGAUUUCGCAUCCCCGAAAGGGUACAUGGCGUUAGGU
eFx	GGAU CGAAAGAUUUCGCGGCCCGAAAGGGGAUUAGCGUUAAGGU
tRNA, Pro1(CG G)	CGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGUUCGGGACGAAGGGGUUCGAGGUUCGAAUCCUCUAUACACCGACCA
tRNA, Pro1(GGU)	CGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGUUGGUUACGAAGGGGUUCGAGGUUCGAAUCCUCUAUACACCGACCA
tRNA, Pro1(GUG)	CGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGUUGUGAACGAAGGGGUUCGAGGUUCGAAUCCUCUAUACACCGACCA
tRNA, Pro1C13G(CG G)	CGGUGAUUUGGCGGAGCCUGGUAGCGCACUUCGUUCGGGACGAAGGGGUUCGAGGUUCGAAUCCUCUAUACACCGACCA
tRNA, Pro1E1(CG G)	CGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGUUCGGGACGAAGGGGUUCGAGGUUCGAAUCCUCUAUACACCGACCA
tRNA, Pro1E2(CG G)	GGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGUUCGGGACGAAGGGGUUCGAGGUUCGAAUCCUCUAUACACCGGCCA
tRNA, Pro1E2(GAU)	GGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGCGUGAUUACGAAGGGGUUCAGGGGUUCGAAUCCUCUAUACACCGGCCA
tRNA, Pro1E2(GGU)	GGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGUUGGUUACGAAGGGGUUCAGGGGUUCGAAUCCUCUAUACACCGGCCA
tRNA, Pro1E2(GUG)	GGGUGAUUUGGCGCAGCCUGGUAGCGCACUUCGUUGUGUACGAAGGGGUUCAGGGGUUCGAAUCCUCUAUACACCGGCCA
tRNA, GluE2(CG G)	GUCCCCUUCGUUCAGAGGCCACAGCACACCGCCUUCGGGAGGCGGUUACAGGGGUUCGAAUCCUCUAUAGGGGACGCCA
tRNA, AsnE2(CG G)	GGCUCUGUAGUUCAGUCGGUAGAACGGCGGAUUCGGGAUCCGUUUGUACUUGGUUCGAGUCCAGAGCGCGCCA
tRNA, fMet(CAU)	CGCGGGGUGGAGCAGCCUGGUAGCUCUGCGGGCUCUUAACCCGAAGAUCGUGGUGUCAAUCCGGCCCCCGCAACCA

[0058]

[Table 4]

[0059]

RNA name	Extension primer (forward)	Extension primer (reverse)
JFX	GTAATACGACTACTATAGGATCGAAAGATTTCGCG	ACCTAAGCCATGTAACCTTTTCGGGGATGCGGAAATCTTTCGATCC
JFX	GTAATACGACTACTATAGGATCGAAAGATTTCGCG	ACCTAAGCCATATCCCTTTTCGGGGCGCGGAAATCTTTCGATCC
RNA, ProI (CGG)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGTCGGGTGATAGAGGATTCGAAACCTCGACCCCTTCGTCGGAACGGAAGTGC
RNA, ProI (GGU)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGTCGGGTGATAGAGGATTCGAAACCTCGACCCCTTCGTTACCAACGAAGTGCCTACCAAG
RNA, ProI (GU G)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGTCGGGTGATAGAGGATTCGAAACCTCGACCCCTTCGTTCCACAACGAAGTGCCTACCAAG
RNA, ProI C13G (CGG)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGTCGGGTGATAGAGGATTCGAAACCTCGACCCCTTCGTCGCGAACGAAGTGC
RNA, ProI E1 (CGG)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGTCGGGTGATAGGAGGATTCGAAACCCCTCGACCCCTTCGTCGCGAACGAAGTGCCTACCAAG
RNA, ProI E2 (CGG)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGCGGGGTGATAGAGGATTCGAAACCCCTTCGACCCCTTCGTCGCGAACGAAGTGCCTACCAAG
RNA, ProI E2 (GUU)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGCGGGGTGATAGAGGATTCGAAACCTTCGACCCCTTCGTTATCCACAACGAAGTGCCTACCAAG
RNA, ProI E2 (GU G)	GTAATACGACTACTATAGGATGATTGGCGAGGCTGGTAGCGGCACTTCG	TGGCGGGGTGATAGAGGATTCGAAACCTTCGACCCCTTCGTTTTCACAACGAAGTGCCTACCAAG
RNA, GlnE2 (CGG)	GTAATACGCTTCCTGCTAGAGGCCCAAGGACACCCGGTUCGGGAGGGCGGTAAACAGGGGTTTCG	TGGCGTCCCTTCAGGGGATTCGGAACCCCTGTTTACCCGCC
RNA, AsnE2 (CGG)	CCTATACGACTACTATAGGCTGGTAGGACGCTGGTGGTAGAAGGCGCGA	TGGCGGCTCTGACTGGATTCGGAACCGAGTGATACATACGGAATCCCGCGGCTTCTACCGACTG
RNA, MetE (CAU)	GTAATACGACTACTATAGCGGGGTGGTAGGACGCTGGTAGCTCGTTCGGGCTCATACACCGCAAGATCG	TGGTTTCGGCGGGCGGATTTGGAACCGAGGACTCTTCGGGGTTATGAGC

[Table 5]

RNA name	PCR primer (forward)	PCR primer (reverse)
dFx	GGCGTAATACGACTCACTATAG	ACCTAACGCCATGTACCCT
eFx	GGCGTAATACGACTCACTATAG	ACCTAACGCTAATCCCCCT
tRNA, Pro1(CGG)	GGCGTAATACGACTCACTATAC	TGmGTCGGGTGATAGAGGATTC
tRNA, Pro1(GGU)	GGCGTAATACGACTCACTATAC	TGmGTCGGGTGATAGAGGATTC
tRNA, Pro1(GUG)	GGCGTAATACGACTCACTATAC	TGmGTCGGGTGATAGAGGATTC
tRNA, Pro1C13G(CGG)	GGCGTAATACGACTCACTATAC	TGmGTCGGGTGATAGAGGATTC
tRNA, Pro1E1(CGG)	GGCGTAATACGACTCACTATAC	TGmGTCGGGTGATAGGGGATTC
tRNA, Pro1E2(CGG)	GGCGTAATACGACTCACTATAG	TGmGCGGGGTGATAGGGGATTC
tRNA, Pro1E2(GAU)	GGCGTAATACGACTCACTATAG	TGmGCGGGGTGATAGGGGATTC
tRNA, Pro1E2(GGU)	GGCGTAATACGACTCACTATAG	TGmGCGGGGTGATAGGGGATTC
tRNA, Pro1E2(GUG)	GGCGTAATACGACTCACTATAG	TGmGCGGGGTGATAGGGGATTC
tRNA, GluE2(CGG)	GTAATACGACTCACTATAGTCCCCCTTCGTCTAG	TGmGCGTCCCCCTAGGGGATTC
tRNA, AsnE2(CGG)	GTAATACGACTCACTATAG	TGmGCGGCTCTGACTGGAC
tRNA, fMet(CAU)	GGCGTAATACGACTCACTATAC	TGmGTTGCGGGGGCCGGA

[0060] In Table 5, Gm represents 2'-O-methylguanosine.

[0061] The PCR product was subjected to phenol/chloroform extraction and ethanol precipitation, and then used for transcription at 37°C for 16 hours in 250 µL of a reaction mixture containing 40 mM Tris-HCl (pH 8.0), 22.5 mM MgCl₂, 1 mM DTT, 1 mM spermidine, 0.01% Tripton X-100, 120 nM T7 RNA polymerase, 0.04 U/µL RNasin RNase inhibitor (Promega) and 3.75 mM NTP mix.

In preparing tRNA, 5 mM GMP or CMP was added to the above-described solution, depending on the nucleotide at the 5' end (G or C). The resulting RNA transcript was treated with RQ1 DNase (Promega) at 37°C for 30 minutes and then purified by 8% (tRNA) or 12% (flexizyme) polyacrylamide gel containing 6 M urea.

[0062] Aminoacylation of tRNA

After a D-amino acid and a β-amino acid were pre-activated as 3,5-dinitrobenzyl ester (DBE) or cyanomethyl ester (CME), the tRNA was charged with the resulting amino acid by using an appropriate flexizyme (dFx for DBE activated amino acid, eFx for CME activated amino acid).

D-alanine, D-serine, and D-cysteine, L-β-homoglutamine (βhQ), L-β-homomethionine (βhM), L-β-homophenylglycine (βhF), L-proline, and 2-aminobutyric acid (Aib) were DBE-activated, while chloroacetyl D-phenylalanine (D-ClAcPhe) was CME activated. The above-described activated amino acids were synthesized according to the method described in Non-Patent Document 1 and Reference Literature 18.

The flexizyme reaction (aminoacylation) was performed at 0°C in a reaction mixture containing 50 mM HEPES-KOH (pH 7.5 for the D-amino acid, pH 8.7 for the β-amino acid), 600 mM MgCl₂, 20% DMSO, 25 µM dFx or eFx, 25 µM tRNA, and 5 mM activated amino acid.

The eFx was used for D-^{ClAc}Phe-CME, while the dFx was used for the other amino acids. The reaction was carried out for 2 hours for D-Ala, D-^{ClAc}Phe, L-Pro, and Aib; 6 hours for D-Ser and D-Cys; and 22 hours for the β -amino acid. The aminoacyl-tRNA was collected by ethanol precipitation. The pellet was then washed twice with 70% ethanol containing sodium acetate (pH 5.2) and once with 70% ethanol and dissolved in 1 mM sodium acetate (pH 5.2).

[0063] Preparation of EF-P

An E. coli EF-P gene was cloned into a modified pET28a vector having a PreScission protease recognition site instead of a thrombin site. E. coli EpmA and EpmB genes were cloned into a pETDuet-1 vector. The resulting vectors were co-introduced into E. coli Rosetta2 (DE3). The cells were cultured at 37°C for 2 hours in an LB medium containing 0.5 mM IPTG and lysed by ultrasonication. The cell lysate was applied to a Ni-NTA column and histidine labeled EF-P was purified. The column was washed with a buffer A (20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 2 mM imidazole, and 1 mM 2-mercaptoethanol) and the histidine-labeled EF-P was eluted by the buffer A containing 300 mM imidazole. Turbo3C protease was added to the eluate to cleave a histidine tag, followed by dialysis overnight at 4°C against the buffer A. The sample was applied to a Ni-NTA column and the flow-through and the washing fraction were collected as EF-P without a histidine tag. Then, the protein was concentrated through Amicon ultra 10 k centrifugal filter (Merck Millipore). The modification of EF-P was confirmed by mass spectrometric analysis described in Reference Literature 3.

[0064] Translation and quantification of model peptide

The translation of a model peptide was performed in a modified FIT (Flexible in vitro translation) system in which the concentrations of IF2, EF-G, and EF-Tu/Ts (3,

0.1, and 20 μ M, respectively) were optimized for incorporation of a D-amino acid and a β -amino acid (refer to Non-Patent Document 5).

The modified FIT system has the following constitution:

50mM HEPES-KOH (pH7.6), 100 mM potassium acetate, 12.3 mM magnesium acetate, 2 mM ATP, 2 mM GTP, 1 mM CTP, 1 mM UTP, 20 mM creatine phosphoric acid, 0.1 mM 10-formyl-5,6,7,8-tetrahydrofolic acid, 2 mM spermidine, 1 mM DTT, 1.5 mg/mL E.coli total tRNA, 1.2 μ M E.coli ribosome, 0.6 μ M methionyl-tRNA formyltransferase, 2.7 μ M IF1, 3 μ M IF2, 1.5 μ M IF3, 0.1 μ M EF-G, 20 μ M EF-Tu/Ts, 0.25 μ M RF2, 0.17 μ M RF3, 0.5 μ M RRF, 4 μ g/mL creatine kinase, 3 μ g/mL myokinase, 0.1 μ M inorganic pyrophosphatase, 0.1 μ M nucleotide diphosphate kinase, 0.1 μ M T7 RNA polymerase, 0.13 μ M AspRS, 0.11 μ M LysRS, 0.03 μ M MetRS, 0.02 μ M TyrRS, 0.05 mM [14 C] aspartic acid, 0.5 mM lysine, 0.5 mM methionine, 0.5 mM tyrosine, 25 μ M each pre-charged aminoacyl-tRNA, and 0.5 μ M DNA template. For the translation of rP1 and rP2 or β -amino acid-containing peptide, 0.09 μ M GlyRS and 0.5 mM glycine were added to the above solution.

The translation reaction was made at 37°C in 2.5 μ L of a solution. An equal volume of a stop solution (0.9M Tris-HCl (pH 8.45), 8% SDS, 30% glycerol, and 0.01% xylene cyanol) was added to stop the reaction, followed by incubation at 95°C for 2 or 3 minutes. Next, the sample was analyzed by autoradiography using 15% tricine SDS-PAGE and Typhoon FLA 7000 (GE Healthcare). A peptide yield was normalized by the intensity of [14 C]-Asp band.

[0065] For MALDI-TOF mass spectrometric analysis, 0.5 mM non-radioactive aspartic acid was added to the above solution instead of [14 C] aspartic acid.

Translation was performed at 37°C for 20 minutes or 40 minutes, followed by dilution with an equal volume of 2×HBS buffer (100 mM HEPES-KOH (pH 7.6), 300 mM

NaCl). Then, the peptide was purified by anti-FLAG M2 affinity gel (Sigma). The reaction mixture was added to 5 μ L of gel beads and the mixture was incubated at 25°C for 30 minutes. The peptide on the gel beads was washed with 25 μ L of 1 $\times$ HBS buffer (50 mM HEPES-KOH (pH 7.6), 150 mM NaCl) and was eluted from the beads by 15 μ L of 0.2% trifluoroacetic acid. Then, the peptide was desalted with SPE C-tip (Nikkyo Technos) and eluted by 1.2 μ L of 80% acetonitrile which was a 0.5% acetic acid solution containing 50% saturated (R) cyano-4-hydroxycinnamic acid. MALDI-TOF-mass spectrometric analysis was performed in a reflector/positive mode by using ultrafleXtreme (Bruker Daltonics). Peptide Mass Standard II (Bruker Daltonics) was used for external mass calibration.

[0066] Essential D-arm structure of tRNA^{Pro1} recognized by EF-P for promoting consecutive D-amino acid incorporation

The following test was carried out to confirm that the ability of EF-P enhancing a peptidyl transfer (PT) rate in Pro-Pro elongation promoted the slow incorporation of some D-amino acids in a nascent peptide chain.

First, two mRNA constructs (mR1 and mR2) each of which encoded two and three consecutive D-amino acids were prepared (Fig. 2A).

The D-amino acids were pre-charged on a tRNA^{Pro1} derivative having a corresponding anticodon by the flexizyme technology and they were used in the FIT system containing 5 μ M EF-P protein in order to synthesize an rP1 peptide or rP2 peptide.

In this particular FIT system, [¹⁴C]Asp, cold Met, Tyr, Lys, and Gly and ARSs corresponding thereto were added and other amino acids/ARSs pair not encoded by the mRNAs (mR1 and mR2) was omitted. Instead, D-aminoacyl-tRNA was added to assign the D-amino acid at NNN codon. The concentrations of IF-2, EF-Tu, and

EF-G were set at 3 μ M, 20 μ M, and 0.1 μ M, respectively, which were values optimized in advance for the consecutive incorporation of the D-amino acids (Non-Patent Documents 2 and 5). A wild type (WT) tRNA^{Pro1}_{CGG} has a C/G base pair at positions 13 and 22 located at the end of the D-stem, whereas a mutant tRNA^{Pro1}_{CGG} has a C-to-G point mutation at position 13 and disrupts base pair formation (Fig. 2B). EF-P strictly recognized the 9-nt D-loop closed with the stable D-stem structure of 4 base pairs of WT tRNA^{Pro1} and therefore, the C13G mutation decreased the recognition by EF-P. To decode the NNN codon designating a D-amino acid in the mR1 and mR2, tRNA^{Pro1}_{CGG} and the C13G mutant as a negative control were used. The anticodon of tRNA^{Pro1} was changed into another one for decoding the target codon (Fig. 2C).

Three respectively different codons (NNN=CCG, ACU, or CAC) for introducing two consecutive D-alanines into an rP1 peptide (rP1-D-Ala₂) were tested in the mR1 by using WT D-Ala- tRNA^{Pro1} (Fig. 2C, Fig. 3A).

The peptide thus obtained was separated by tricine-SDS-PAGE and quantified by autoradiography. Addition of EF-P brought an important improvement on the expression level of the rP1 peptide (rP1-D-Ala₂) within a similar range (from 0.4 to 0.6 μ M) without depending on codons. The apparent level of the enhancement effect however differed depending on codons due to a difference in the background level of EF-P-free D-Ala incorporation (2.6-, 4.7- and 12-fold for CCG, ACU and CAC, respectively). As a result of EF-P enhancement of the consecutive incorporation of L-Pro present at the position of X residues in rP1-L-Pro₂ and rP2-L-Pro₃ peptides by using L-Pro-tRNA^{Pro1} as a positive control, 1.4- and 22-fold improvement effects were observed, respectively (Figs. 3B and 3C).

Next, it was studied whether or not EF-P was able to enhance the incorporation of D-Ser, D-Cys and D-His charged onto tRNA^{Pro1}_{CGG} suppressing the CCG codon on mR1 (Fig. 2D and Fig. 3D). The enhancement effect was 5.4-, 8.4-, and 2.1-fold, respectively, indicating that EF-P enhanced two consecutive incorporation of these D-amino acids within a range of from 2- to 10-fold. On the other hand, when the C13G mutant of tRNA^{Pro1}_{CGG} was used, the EF-P enhancement effect was 2.1-, 2.4-, 3.5- and 0.9-fold for D-Ala, D-Ser, D-Cys and D-His in rP1-D-X₂ peptide, respectively and it decreased. Since the absolute expression level was enhanced by 2-fold or more by the use of WT/EF-P over C13G/EF-P, the EF-P effect was definitely observed in each case. Accurate expression of these peptides was confirmed by MALDI-TOF mass spectrometric analysis (Fig. 3E).

[0067] Then, an achiral α , α -dimethyl substituted amino acid (Aib, 2-aminobutyric acid) was used (Fig. 2D and Figs. 3D and 3E).

It has been confirmed that the α , α -disubstitution set a similar situation in the ribosome PT center like a D-amino acid and very low efficiency of consecutive incorporation of Aib was observed. The EF-P enhancement for the incorporation of Aib into rP1-Aib₂ peptide was 20-fold by the use of WT tRNA^{Pro1}, while that by the C13G mutant was 5-fold. The expression level was enhanced 3-fold or more by the WT/EF-P combination.

Next, three consecutive incorporation of D-Alas and Aibs into the nascent chain of rP2 peptide was studied (Fig. 2E and Figs. 3D and 3E). The enhancement effect on the incorporation of D-Ala into rP2-D-Ala₃ peptide by EF-P was 1.9-fold using WT tRNA^{Pro1}_{CGG}, while the expression using the C13G mutant was undetectable level even in the presence of EF-P. The enhancement effect by EF-P was 3-fold when rP2-Aib₃ was expressed using WT tRNA^{Pro1}_{CGG}, while no effect was observed when

the C13G mutant was used. The above-described results show that EF-P stimulates the expression level of consecutive D-amino acid incorporation.

[0068] In addition, the EF-P concentration was titrated as a function of the expression level of rP1-D-Ala₂ peptide in the presence of WT D-Ala-tRNA^{Pro1}_{CGG} (Fig. 4A). In the titration experiments at 40- and 15-min time points, 5–7 μ M EF-P showed the maximum level of rP1-D-Ala₂ expression, whereas higher concentrations of EF-P (10 and 15 μ M) declined the EF-P effect. In consideration that the concentration of ribosome in the translation system of Examples was 1.2 μ M, the optimum concentration of EF-P was about 5-fold of the ribosome concentration.

Time course analysis of the translation of rP1-D-Ala₂ peptide was carried out in the presence of EF-P (5 μ M) and absence of EF-P. In both conditions, the yield of full-length peptide plateaued in 50 minutes (Fig. 4B). In 50 min, the expression level of rP1-D-Ala₂ peptide reached approximately 1 μ M in the presence of 5 μ M EF-P. Such a concentration was comparable to the expression level of a full-L-peptide containing L-Ala₂ (Reference Literature 3).

[0069] Engineering of T-stem structure of tRNA^{Pro1} for further enhancing consecutive D-alanine incorporation

The T-stem of tRNA^{GluE2} was introduced into tRNA^{Pro1} to form a chimeric tRNA called herein tRNA^{Pro1E1} (Fig. 5A, tRNA^{Pro1E1}_{CGG}).

Further, another chimeric tRNA called herein tRNA^{Pro1E2}_{CGG} having additional mutation at a discriminator base recognized by ProRS (Reference Literature 9) was formed. Due to introduction of such mutation, tRNA^{Pro1E2}_{CGG} was prevented from being charged with Pro by ProRS. This means that it has orthogonality to ProRS.

Ability of incorporating two consecutive D-alanines into rP1-D-Ala₂ was compared among WT tRNA^{Pro1}_{CGG}, tRNA^{Pro1E1}_{CGG}, tRNA^{Pro1E2}_{CGG}, tRNA^{GluE2}_{CGG}, and tRNA^{AsnE2}_{CGG} (Fig. 5A) (Fig. 5B and Fig. 6A).

As a result, it was confirmed from comparison with other tested tRNAs that a peptide expression level was enhanced by 4-fold or more by using EF-P with tRNA^{Pro1E1}_{CGG} (4.1-fold) or with tRNA^{Pro1E2}_{CGG} (5.0-fold) in combination. This clearly means that the chimeric tRNA^{Pro1E1} and tRNA^{Pro1E2} derived from the T-stem of chimeric tRNA^{GluE2} and the D-arm of tRNA^{Pro1} enhance a total PT rate, assisted by EF-P and EF-Tu.

[0070] Incorporation of a plurality of D-amino acids enhanced by EF-P

The combination of D-Ala-tRNA^{Pro1E2} and EF-P enables enhancement of the expression level of rPA-D-Ala₂.

Five respectively different template mRNAs were prepared (Fig. 7A, mR3 to mR7) and peptides corresponding thereto containing, in the sequence thereof, one to three respectively different D-amino acids were synthesized (Fig. 7A, rP3 to rP7 peptides).

In these mRNAs, D-Ala-tRNA^{Pro1E2}_{GAU}, D-Cys-tRNA^{Pro1E2}_{GUG}, and D-Ser-tRNA^{Pro1E2}_{GGU} were assigned to AUU, CAU and ACU codons, respectively. It was confirmed that respective rP3 and rP4 expression was enhanced in the presence of 5 μM EF-P (Fig. 6B). On the other hand, no band was detected in the absence of EF-P. This means that markedly enhanced expression was observed in the rP3 and rP4 containing two kinds of two-consecutive D-amino acids (D-Ala₂-YY-D-Ala₂ or D-Ala₂-YY-D-Cys₂; Y = L-Tyr).

The rP5 peptide containing two kinds of two-consecutive D-amino acids (D-Ala₂-YY-D-Ser₂) expressed a detectable level of a rP5 peptide even in the

absence of EF-P, but addition of EF-P enhanced the expression level by 13-fold (Fig. 6B and Fig. 7B).

The expression of rP6 and rP7 peptides alternately containing L- and D-amino acids was low but detectable in the presence of EF-P. On the other hand, the addition of EF-P enhanced the expression level by 3.3- and 3.0-fold, respectively (Fig. 6B and Fig. 7B).

MALDI-TOF mass spectrometric analysis showed that the peptides each had an expected molecular weight in the presence of EF-P (Fig. 6C).

The above-described results suggest that the combination of EF-P and D-aminoacyl-tRNA^{Pro1E2} has a remarkable impact on the enhancement of the expression level of a peptide containing multiple kinds of D-amino acids.

[0071] Ribosomal synthesis of cyclic D-peptide stimulated by EF-P

As a result of demonstration of the expression of a macrocyclic D-peptide using a tRNA^{GluE2} derivative by the FIT system, the expression level of the macrocyclic D-peptide was 4-fold lower than that of a macrocyclic L-peptide.

tRNA^{Pro1E2} was therefore used, which enhanced the expression level of the macrocyclic D-peptide. More specifically, use of EF-P led to a success in enhancing the expression level of two model D-peptides of rP8 and rP9 peptides (Fig. 8).

The rP8 contains D-Cys-D-Ser-D-Ala-D-Ser-D-Cys.

MALDI-TOF mass spectrometric analysis of rP8 peptide expressed using tRNA^{Pro1E2} as an elongator carrier with EF-P showed that the above D-amino acids were introduced correctly and two D-Cys residues formed a disulfide bond to give a macrocyclic structure (Fig. 9C).

Tricine-SDS-PAGE analysis of rP8 peptide showed that EF-P paired with D-aminoacyl-tRNA^{Pro1E2} enhanced the expression level by about 8-fold compared with EF-P-inactive D-aminoacyl-tRNA^{GluE2} (Fig. 8B and Fig. 9A).

In a similar manner, rP9 peptide was designed as another model peptide containing a D-^{ClAc}Phe-D-Ser-D-Ser-D-Cys continuous stretch. The N-terminal chloroacetyl group of D-^{ClAc}Phe reacted with the sulfhydryl group of D-Cys and formed a thioether macrocyclic structure (Fig. 8A). EF-P paired with D-aminoacyl-tRNA^{Pro1E2} enhanced rP9 peptide expression by about 10-fold compared with tRNA^{GluE2} (Fig. 8C), showing that a thioether macrocyclic D-peptide was produced efficiently (Fig. 9C).

[0072] Incorporation of consecutive β -amino acids enhanced by EF-P

Incorporation of D-amino acids was conducted as described below according to the above-described method.

Wild type (WT) tRNA^{Pro1CGG} or C13G mutant tRNA^{Pro1CGG} (Fig. 10a) was pre-charged with β hQ, β hM, and β hF (Fig. 10b) by the flexizyme technology and was used for the translation of a template mRNA of mR1 (Fig. 10c) in the FIT system containing 5 μ M EF-protein to synthesize an rP1 peptide.

The Tyr-Lys-Lys-Tyr-Lys-Lys-Tys-Lys sequence before the β -amino acid is a sequence introduced for the detection of the expressed peptide in tricine SDS-PAGE or MALDI-TOF mass spectrometric analysis.

The peptide was expressed in the presence of [¹⁴C]Asp, subjected to 15% tricine SDS-PAGE, and quantified using autoradiography by the intensity of [¹⁴C]Asp introduced into the flag-tag region at the C-terminal (Fig. 10d and Fig. 15a).

In the presence of cold Asp, expression of the same peptide was confirmed by MALDI-TOF mass spectrometric analysis even in the absence of EF-P (Fig. 15b). When the WT tRNA^{Pro1}_{CGG} was used, the incorporation of β -amino acids increased due to the presence of EF-P (Fig. 10d, 1.7-fold for β hQ, 4.6-fold for β hM, and 1.3-fold for β hF).

The above-described results clearly show that EF-P recognizes the D-arm structure of tRNA^{Pro1}. It promotes the incorporation of β -amino acids responding thereto.

Then, the optimized concentration of EF-P in the incorporation of β hQ and β hM into rP1 peptide was studied (Reference Literatures 1 and 2) and the EF-P concentration was preferably from about 5 to 10 μ M in the incorporation of β -amino acids. A similar tendency was confirmed in the incorporation of D-alanine.

Based on the above-described results, it has been presumed that EF-P occupies the E site of ribosome and inhibits the translocation of deacyl tRNA from P site to E site.

[0073] Engineering of T-stem structure of tRNA^{Pro1} for enhancing EF-Tu binding in consecutive β -alanine incorporation

tRNAs shown in Fig. 11a were charged with β hM by the flexizyme technology and whether or not further incorporation enhancement occurred was compared among the tRNAs charged with β hM.

In contrast with tRNA^{AsnE2} based on Escherichia coli tRNA^{Asn} and having a T-stem and a D-arm each of which was not optimized, tRNA^{GluE2} derived from tRNA^{Glu} having a T-stem optimized for EF-Tu binding, and tRNA^{Pro1}, β hM-tRNA^{Pro1E2} showed a 7.3-fold incorporation enhancement effect in the presence of RF-P and showed

almost a 2-fold incorporation enhancement effect over $\beta\text{hM-tRNA}^{\text{Pro1}}$ in the presence of EF-P (Fig. 11b).

[0074] Incorporation of 7 consecutive β -amino acids

βhM was introduced into from 2 to 7 consecutive CGG codons of mR2 by using $\text{tRNA}^{\text{GluE2}}_{\text{CGG}}$ charged with βhM to synthesize rP2- βhM_n peptides corresponding to mR2 (Fig. 12 and Fig. 16b).

Although the translation efficiency of rP2- βhM_n peptides decreased with an increase in the number of consecutive βhMs , a full-length rP2 peptide containing 7-consecutive βhMs was detected in the presence of EF-P (Fig. 12b and Fig. 12c).

In addition, a marked incorporation enhancement effect by EF-P was confirmed in rP2- βhM_2 peptide and rP2- βhM_3 peptide (5.9-fold and 16.2-fold, respectively).

Similarly, $\text{tRNA}^{\text{Pro1E2}}_{\text{GAU}}$ and $\text{tRNA}^{\text{Pro1E2}}_{\text{GUG}}$ were pre-charged with βhM and βhF respectively and peptides of from rP3-3 to rP3-7 having each of them incorporated therein by AUU codon and CAU codon were synthesized (Fig. 13a).

Expression of the peptides of from rP3-3 to rP3-7 was confirmed in the presence of EF-P (Fig. 13b, Fig. 13c, and Fig. 16c).

In the peptides of from rP3-3 to rP3-5, a marked incorporation enhancement effect by EFP was confirmed (7.4-fold, 21.7-fold, and 30.3-fold, respectively). In addition, in the peptides of from rP3-6 to rP3-7, an expression enhancement effect by EF-P from N.D. to a detectable level was confirmed.

[0075] Incorporation of a plurality of D-amino acids and β -amino acids

To show the versatility of the present method, eight kinds of peptides, that is, rP4 to rP11 having some of three kinds of β -amino acids and four kinds of D-amino

acids incorporated at various positions thereof were synthesized (Fig. 14a and Fig. 14b).

The peptides of from rP4 to rP5 showed an incorporation enhancement effect by EF-P (1.9-fold and 1.6-fold, respectively in Fig. 14c and Fig. 16d). The peptides of from rP6 to rP7, which had two sets of consecutive β -amino acids incorporated therein, showed an incorporation enhancement effect by EF-P (12.3-fold and 19.3-fold, respectively in Fig. 14c and Fig. 16d). The peptides of from rP8 to rP9, which had a combination of β -amino acids and D-amino acids incorporated therein, showed an incorporation enhancement effect by EP to elevate an expression level to a detectable one (9.2-fold for rP8 peptide and 0.04 μ M for rP8 peptide in Fig. 14c and Fig. 16d).

Macrocyclic peptides formed by a thioether bond (from rP10 peptide to rP11 peptide) having consecutive β -amino acids incorporated therein were synthesized (Fig. 14a). An initiator codon was assigned to D-^{ClAc}Phe and the chloroacetyl group of the D-^{ClAc}Phe reacted with a side-chain thiol group of D-cysteine downstream thereof to form a thioether bond. rP10 peptide and rP11 peptide also showed an incorporation enhancement effect by EF-P (3.2-fold and 2.3-fold, respectively, in Fig. 14d and Fig. 16d).

A by-product was produced due to an incorporation mistake and synthesis of the peptides was confirmed by MALDI-TOF mass spectrometric analysis (Fig. 16e).

[0076] [Reference Literatures]

1. Ude, S. et al. Translation elongation factor EF-P alleviates ribosome stalling at polyproline stretches. *Science* 339, 82-5 (2013).
2. Doerfel, L.K. et al. EF-P is essential for rapid synthesis of proteins containing consecutive proline residues. *Science* 339, 85-8 (2013).

3. Katoh, T., Wohlgemuth, I., Nagano, M., Rodnina, M.V. & Suga, H. Essential structural elements in tRNA(Pro) for EF-P-mediated alleviation of translation stalling. *Nat Commun* 7, 11657 (2016).
4. LaRiviere, F.J., Wolfson, A.D. & Uhlenbeck, O.C. Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. *Science* 294, 165-8 (2001).
5. Dale, T., Sanderson, L.E. & Uhlenbeck, O.C. The affinity of elongation factor Tu for an aminoacyl-tRNA is modulated by the esterified amino acid. *Biochemistry* 43, 6159-66 (2004).
6. Becker, H.D. & Kern, D. *Thermus thermophilus*: a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc Natl Acad Sci U S A* 95, 12832-7 (1998).
7. Stanzel, M., Schon, A. & Sprinzl, M. Discrimination against misacylated tRNA by chloroplast elongation factor Tu. *Eur J Biochem* 219, 435-9 (1994).
8. Schrader, J.M., Chapman, S.J. & Uhlenbeck, O.C. Tuning the affinity of aminoacyl-tRNA to elongation factor Tu for optimal decoding. *Proc Natl Acad Sci U S A* 108, 5215-20 (2011).
9. Hasegawa, T. & Yokogawa, T. *Escherichia coli* proline tRNA: structure and recognition sites for prolyl-tRNA synthetase. *Nucleic Acids Symp. Ser.* 44, 7-8 (2000).
10. Pavlov, M.Y. et al. Slow peptide bond formation by proline and other N-alkylamino acids in translation. *Proc Natl Acad Sci U S A* 106, 50-4 (2009).
11. Doerfel, L.K. et al. Entropic Contribution of Elongation Factor P to Proline Positioning at the Catalytic Center of the Ribosome. *J Am Chem Soc* 137, 12997-3006 (2015).

12. Wohlgemuth, I., Brenner, S., Beringer, M. & Rodnina, M.V. Modulation of the rate of peptidyl transfer on the ribosome by the nature of substrates. *J Biol Chem* 283, 32229-35 (2008).
13. Johansson, M. et al. pH-sensitivity of the ribosomal peptidyl transfer reaction dependent on the identity of the A-site aminoacyl-tRNA. *Proc Natl Acad Sci U S A* 108, 79-84 (2011).
14. Melnikov, S. et al. Molecular insights into protein synthesis with proline residues. *EMBO Rep* 17, 1776-1784 (2016).
15. Driggers, E.M., Hale, S.P., Lee, J. & Terrett, N.K. The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat Rev Drug Discov* 7, 608-24 (2008).
16. Hipolito, C.J. & Suga, H. Ribosomal production and in vitro selection of natural product-like peptidomimetics: the FIT and RaPID systems. *Curr Opin Chem Biol* 16, 196-203 (2012).
17. Passioura, T., Katoh, T., Goto, Y. & Suga, H. Selection-based discovery of druglike macrocyclic peptides. *Annu Rev Biochem* 83, 727-52 (2014).
18. Saito, H., Kourouklis, D. & Suga, H. An in vitro evolved precursor tRNA with aminoacylation activity. *EMBO J* 20, 1797-806 (2001).

What is claimed is:

1. A tRNA comprising a base sequence represented by SEQ ID NO: 1 and encoding a non-proteinogenic amino acid.

$N_1N_2GCN_3N_4N_5N_6N_7N_8N_9N_{10}N_{11}GCN_{12}N_{13}$ (SEQ ID NO: 1)

(in SEQ ID NO: 1,

N_1 to N_{13} each represents an arbitrary base, N_3 to N_{11} form a D-loop, and

N_1N_2GC forms a base pair with $N_{13}N_{12}CG$).

2. The tRNA according to Claim 1, further comprising a base sequence represented by SEQ ID NO: 2.

$AGGGG(N_{14})_mCCCCU$ (SEQ ID NO: 2)

(in SEQ ID NO: 2,

N_{14} represents an arbitrary base, m stands for an integer of 1 or more, $(N_{14})_m$ form a T-loop, and

$AGGGG$ forms a base pair with $UCCCC$).

3. The tRNA according to Claim 1 or 2, wherein the non-proteinogenic amino acid is a D-amino acid, a β -amino acid, or an α,α -disubstituted amino acid.
4. A translation system for the synthesis of a peptide having two or more consecutive non-proteinogenic amino acids, comprising the tRNA as claimed in any one of Claim 1 to 3.
5. A method of constructing a peptide library, comprising a step of translating in a cell-free translation system by using the tRNA as claimed in any one of Claims 1 to 3.

6. A method of constructing a library of a complex between a peptide and an mRNA encoding the peptide, comprising a step of translating in a cell-free translation system by using the tRNA as claimed in any one of Claims 1 to 3.
7. A peptide library constructed by the method as claimed in Claim 5 or a peptide-mRNA complex library constructed by the method as claimed in Claim 6.
8. The peptide library or peptide-mRNA complex library as claimed in Claim 7, comprising a peptide having two or more consecutive non-proteinogenic amino acids.
9. The peptide library or peptide-mRNA complex library as claimed in Claim 8, comprising a peptide having four or more consecutive non-proteinogenic amino acids and being intramolecularly crosslinked between the non-proteinogenic amino acids.