Chapter 3

Incorporation of unnatural amino acid into mRNA display libraries by amber codon suppression

Abstract: The incorporation of unnatural amino acids into selectable, amplifiable peptide and protein libraries expands the chemical diversity of such libraries, thus considerably facilitating the process of obtaining ligands with improved properties (affinity, specificity, and function), particularly against therapeutically interesting targets. Biocytin, a biotin derivative of lysine, can be inserted into a mRNA-protein fusion molecule through amber stop codon suppression. The templates containing the codon corresponding to biocytin tRNA (a UAG stop codon) can be enriched by iterative cycles of selection against a streptavidin agarose matrix.

Publication: Li SW, Millward S, Roberts RW; *J. Am. Chem. Soc.*, 124 (34): 9972-9973 AUG 2002

3.1 Introduction

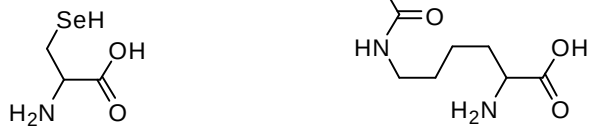
1. Amber stop codon suppression

Site-specific mutagenesis is an important approach for modifying the amino acid residues of proteins for structural and functional analysis. However, only 20 naturally occurring residues are convertible to one another in that the translational apparatus of living cells can only assemble these naturally occurring amino acids. The incorporation of unnatural amino acids into proteins is highly desired because it will add more available tools for the determination of various properties of proteins. For instance, the inclusion of a fluorescent probe in a protein can make this protein traceable for its role in a metabolite pathway or other cellular processes[1]. The introduction of an unnatural ketone group provides a specific site in a protein for selective post-translational modification[2,3]. The replacement of a key residue with an unnatural analog can also facilitate the determination of its functions[4].

The *in vivo* protein translational machinery is highly conserved in both prokaryotic and eukaryotic cells and consists of a large ribosomal complex and numerous proteins. Usually, the biosynthesis of a protein in a living cell involves initiation, elongation and termination. In the initiation stage, the ribosome scans through a messenger RNA (mRNA) to search for a starting point (AUG codon) in a suitable context. Once a start codon is found, a methionine-charged transfer RNA (tRNA) moves into the ribosome, recognizes this AUG codon, and serves as the first residue for protein synthesis. During the elongation stage, the ribosome processes along the mRNA template while various charged tRNA enter the ribosome to add more amino acids to the extended polypeptide chain dependent on the sequence of the mRNA template. Finally, when the

ribosome meets one of stop codons (UAG, UGA, or UAA), the nascent polypeptide chain leaves the ribosome, which itself also dissociates from the mRNA template. Other than mRNA and various tRNAs, this complex process also requires many accessory enzymes in every stage, such as initiation factors, elongation factors, and termination factors. The accuracy of protein translation from the genetic information contained in mRNA is controlled by aminoacyl-tRNA synthetases, enzymes that acylate tRNAs with different amino acids based on tRNA recognition properties. For example, isoleucine-tRNA synthetase is able to charge cognate tRNAs with isoleucine but not valine, whose structure differs from that of isoleucine by only a methylene group[5,6]. Such a high specificity provided by tRNA synthetases can keep the newly translated proteins from those catastrophic mutations due to the misacylation of tRNA. On the other hand, the editing function of tRNA synthetases also makes it extremely difficult to express proteins containing unnatural residues by simply supplying an unnatural amino acid to the growth medium.

In some rare cases, however, proteins newly translated *in vivo* do contain unnatural amino acids. For example, during the translation of formate dehydrogenase in *E. coli*, selenocysteine, which is not one of the twenty primary amino acids, is incorporated[7]. A more recent discovery is that the methylamine methyltransferases isolated from *M. barkeri* includes another unnatural residue, L-pyrrolysine[8]. The



selenocysteine

X=Me, NH_2, OH

pyrrolysine

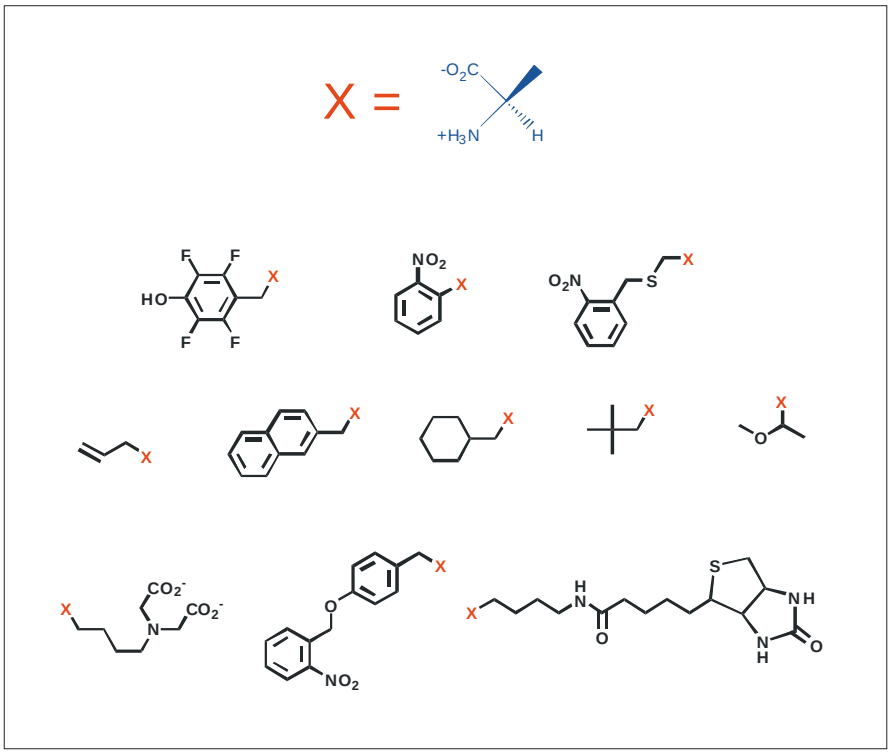
Two naturally inserted non-canonical amino acids

programming of unnatural amino acids into these proteins results from the existence of an in-frame UAG stop codon in their gene, a UAG codon tRNA suppressor, and a dedicated tRNA-synthetase that charges UAG stop codon suppressor with unnatural amino acids. Schultz and others have taken advantage of this observation and developed a semi-synthetic strategy to translate proteins carrying unnatural residues site-specifically that are compatible with ribosome[9].

The original semi-synthetic strategy was reported to translate proteins containing unnatural side chains in bacterial lysate[10]. A combination of chemical synthesis and run-off transcription was employed to prepare a misacylated-tRNA whose 3-base anticodon in the loop was replaced with an anti-UAG codon. Supplement of this semi-synthetic tRNA and mRNA template containing a UAG stop codon in the desirable position resulted in incorporation of the non-canonical residue. Many amino acids, including fluorescent or spectroscopic probes[1], affinity tags[11], and analogues for detailed mechanistic analysis[4], have been incorporated by this method (Figure 3.1). In addition to non-native side chains, various molecules different from standard amino acids can also be introduced. The acylation of tRNA suppressor with 2-hydroxyl acid, for instance, can replace the amide bond with a ester linkage that is acid-labile[12]. The efficient incorporation of N-methyl amino acids can generate proteins that are protease-resistant[12]. The semi-synthetic tRNAs were also injected into oocytes for *in vivo* expression of proteins containing non-natural residues[13].

Recently, Wang and his coworkers demonstrated that they were able to incorporate a O-methyl-L-tyrosine into Dihydrofolate Reductase (DHFR) in *E. coli.*, without the requirement of semi-synthetic tRNA[14]. Their approach starts with generating

Figure 3.1 Examples of unnatural amino acids incorporated into protein by amber stop codon suppression method[15].



an orthogonal tRNA and tRNA synthetase pair, which does not interact with other pairs existing in *E. coli*. Then the orthogonal synthetase was engineered so that it only charges the orthogonal tRNA with an unnatural amino acid. The orthogonal tRNA delivers the attached novel amino acid into proteins in response to a UAG codon inserted at any position of interest. Using this method, they have incorporated O-methyl-L-tyrosine into proteins with purity higher than 99 percent, which is close to the translation fidelity of natural amino acids. The application of this strategy in eukaryotic cells has been under investigation.

2. Combinatorial peptide libraries

Combinatorial peptide libraries have been extensively applied to select for novel peptides with desirable properties, including agonists and antagonists of receptors, inhibitors of enzymes, and hormones that control various metabolic pathways. Techniques such as phage-[16], ribosome-[17], and mRNA-display[18-20] can generate very large peptide and protein libraries and sieve them for functional molecules. These libraries share the same feature that the individual peptide is associated with its own encoding DNA or RNA template, which can be amplified easily by polymerase chain reaction (PCR) or the combination of reverse transcription and PCR (RT-PCR). This makes it possible to amplify the small fraction of peptides enriched with desirable properties for further improvement. For example, phage-display libraries, by far the most commonly used method for creating peptide libraries biologically, are based on the expression of coat proteins on the surface of bacteriophage, virus that infect bacteria. One of the coat proteins of the phage M13, termed gp13, is fused with N-terminal

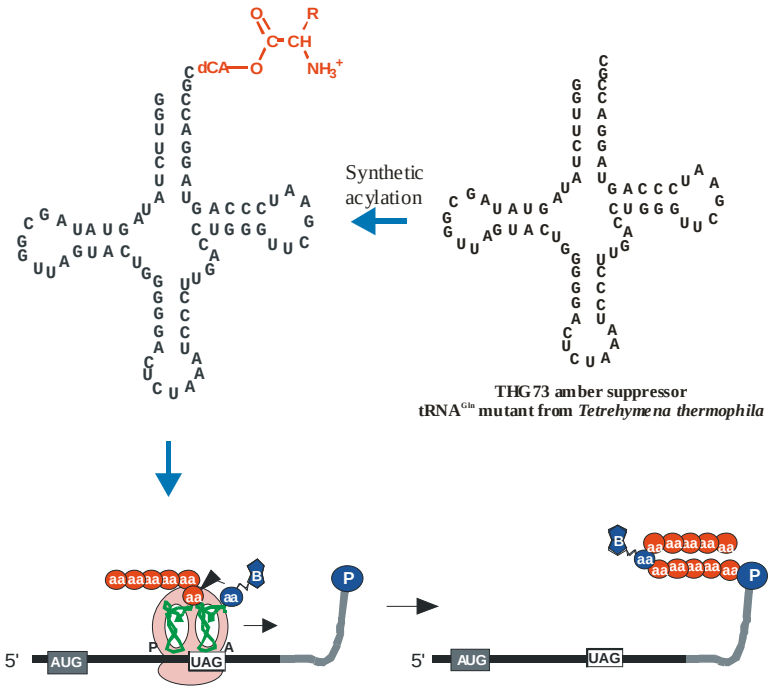
recombinant peptides that are encoded by synthetic randomized oligonucleotides inserted to the 5'-end of the gene (gIII). In ribosome-display libraries, the nascent peptides or proteins remain bound on mRNA through the non-covalent mRNA·ribosome·peptide complex, which is stabilized by a high concentration of magnesium ion and low temperature. The mRNA part of complex in enriched libraries against targets of interest contains the necessary genetic information for the iterative cycles. In mRNA-display libraries, translation extracts of rabbit reticulocyte lysate are used to generate combinatorial libraries of peptides and proteins covalently fused to their own mRNA via a 3' puromycin. These libraries are strictly monovalent and provide for the synthesis of more than 10^{13} independent peptide or protein sequences in a selectable format[21-25].

The chemical diversity that may be programmed into these libraries, however, is limited to the 20 naturally occurring amino acid side chains. Recently, there has been great interest in extending the unnatural strategy to systems where the novel residue may be selected for its function. Here, a combination of mRNA-display and UAG stop codon suppression has been developed to perform *in vitro* selection on mRNA display libraries containing the unnatural amino acid biocytin, a biotin derivative of lysine (Figure 3.2).

3.2 Results and discussion

First, mRNA-peptide fusions containing the unnatural residue biocytin were constructed. Biocytin, the biotin derivative of lysine, represented an excellent choice for the target residue as this has been inserted into proteins previously, and the biotin moiety could be readily used to select peptides that have incorporated this amino acid. The mRNA display libraries were prepared in the rabbit reticulocyte translation extract due to

Figure 3.2 Incorporation of unnatural amino acids into mRNA-display library by amber stop codon suppression. The semi-synthetic charged THG-73 tRNA can insert the unnatural amino acid in responsive to the UAG stop codon on mRNA template[13].



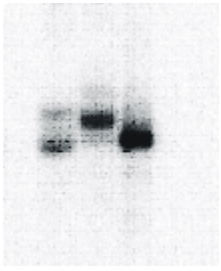
the excellent stability of the template in this media and the efficiency of fusion formation. The amber suppressor tRNA THG73 (a modified *Tetrehymena thermophila* Gln tRNA) was chosen to insert the unnatural residues by nonsense UAG suppression as this construct has high efficiency in eukaryotic translation systems (Figure 3.3).

Next, two templates were constructed to test insertion of the unnatural residue. The first template (Pep1) is a control containing all 20 amino acids, but no stop codon, while the second template (Pep2) contains a similar amino acid composition and a single UAG stop codon at the third position. For Pep2, fusion formation occurs only when the suppressor tRNA is added, consistent with incorporation of biocytin into the Pep2 mRNA-peptide fusion (Figure 3.4).

In order to demonstrate that this approach could be used to select peptides from libraries based on the function of the unnatural residue, the TTG codon that encoded the Trp residue in position 8 of the template pep1 was replaced with an NNS saturation cassette containing 32 possible codons encoding all 20 possible amino acids and the UAG stop (Lib1). Two rounds of *in vitro* selection were performed using streptavidin agarose as an affinity matrix.

Sequencing after one round of selection indicated that UAG stop codons were being enriched at both the randomized position and elsewhere in the open reading frame via point mutations. After a second round of selection vs. streptavidin agarose, nine clones were sequenced from the library. Eight out of nine (88%) contain a UAG stop codon at the randomized position or elsewhere, including two that contain a GAG to UAG transversion at position 5 (Figure 3.5).

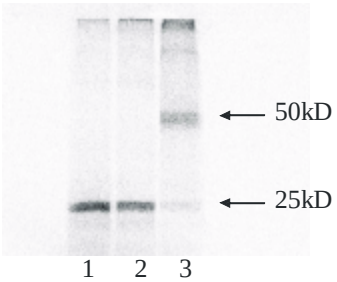
Figure 3.3 Ligation of truncated tRNA with dCA-biocytin gives functional amber codon suppressor. [A]. The ligated tRNA with pdCpA-biocytin is resolved from unligated tRNA. [B]. Supplement of biocytin-charged tRNA can suppress UAG stop codon to give full length 50 kD protein in rabbit retic lysate.



1 2 3

1: tRNA + T4 RNA ligase
2: tRNA + dCA-biocytin + T4 RNA ligase
3: ^{32}P labeled truncated tRNA (THG73)

[A]



1 2 3

1: nACHRα198 mRNA in lysate
2: nACHRα198+ truncated tRNA
3: nACHRα198 + suppressor

[B]

Figure 3.4 Dependence of fusion formation on addition of biocytin-charged THG73. [A]. Sequences of fusion templates tested. [B]. Gel showing formation of fusion products, labeled with 35S-methionine, in the presence (+) or absence (-) of biocytin-charged THG73. Fusion formation on the stop codon-containing template (Pep2) occurs only in the presence of the suppressor tRNA.

[A]

```
Sd29: MG.LDYKDEDKRQEIHLADLCKPFWVYTSGGGG
Sd25: MGRQEIHWANDLCKPFWVYTSGGGG
```

[B]

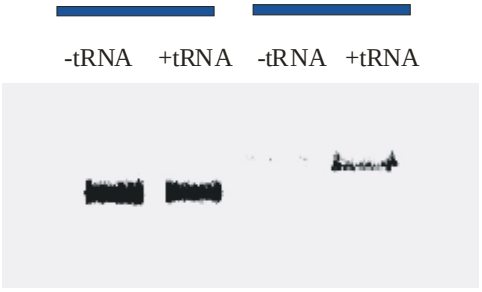


Figure 3.5 Sequences present in Lib1 before and after selection vs. streptavidin agarose. Amino acids sequence in 3-letter format of Lib1 before selection is shown in the second line and its DNA sequence is shown in the third line. Xxx represents all 20 amino acid residues plus UAG stop codon. N is equal amount of all 4 nucleotides and S is 50% G plus 50% C in that position. The sequences in NNS saturation region are highlighted in green and the emerging UAG stop codons are highlighted in red. The sequences same as the original template are labeled with dash (-).

```
 1          5             10             15             20             25
MetGlyArgGlnGluIleHisXxxAlaAsnAspLeuCysLysProPheTrpValTyrThrSerGlyGlyGlyGly
ATGGGCCGCCAGGAGATCCACNNSGCCAACGACCTGTGCAAGCCCTTCTGGGTGTACACCTCCGGCGGCGGCGGC
                  Sequences of 9 clones before selection
--------------------AAG----------------------------------------------------
--------------------AGC----------------------------------------------------
--------------------CCG----------------------------------------------------
--------------------TGC----------------------------------------------------
--------------------GTC----------------------------------------------------
--------------------TGG----------------------------------------------------
--------------------CAG----------------------------------------------------
--------------------AGC----------------------------------------------------
--------------------ATC----------------------------------------------------
              Sequences of 9 clones after one round selection
--------------------TAG----------------------------------------------------
--------------------TAG----------------------------------------------------
--------------------TTG----------------------------------------------------
--------------------AAG----------------------------------------------------
--------------------GAC----------------------------------------------------
--------------------GTG----------------------------------------------------
--------------------CAG----------------------------------------------------
--------------------AGG----------------------------------------------------
------------TAG------CAG----------------------------------------------------
             Sequences of 9 clones after two rounds selection
--------------------TAG----------------------------------------------------
--------------------TAG----------------------------------------------------
--------------------TAG----------------------------------------------------
--------------------TAG----------------------------------------------------
------------TAG------CAG----------------------------------------------------
------------TAG------CAG----------------------------------------------------
--------------------TAG----------------------------------------------------
--------------------TAG----------------------------------------------------
--------------------GAG----------------------------------------------------
```

3.3 Conclusion

This experiment represents the first combination of an *in vitro* selection experiment and nonsense suppression. It now allows selectable peptide and protein libraries to be constructed containing any non-natural amino acid that is compatible with the translation apparatus. These libraries should facilitate the discovery of novel ligands with functionalities beyond those provided by the 20 naturally occurring residues. For example, peptides containing N-substituted amino acids are protease resistant[26] and can show enhanced affinity for natural protein interaction modules such as SH3 domains[27]. This approach should also aid physical organic dissection of protein interfaces, particularly where molecular interactions display context dependence. Additionally, this approach provides a convenient way to construct new molecular tools based on known pharmacophores, in that therapeutically useful small molecules may be presented adjacent to the chemical diversity present in a 10^{13}-member peptide or protein library. Finally, translation systems that allow insertion of two or more unnatural amino acids[28] now provide the intriguing possibility for construction of wholly unnatural libraries in a selectable mRNA display format.

3.4 Materials and methods

1. Synthesis of biocytin-tRNA suppressor (Figure 3.6)

a. Synthesis of NVOC-biocytin cyanomethyl ester

Biocytin (100 mg, 0.26 mmol, Molecular Probes) and sodium carbonate (56 mg, 0.54 mmol) were dissolved in a mixture of water (15 ml) and THF (10 ml). A solution of

6-nitroveratryloxycarbonyl chloride (NVOC-Cl) (74 mg, 0.26 mmol, Sigma) in 10 ml THF was added slowly. After 3 hours, solvents were removed *in vacuo*. Then, 3 ml of dry DMF and 3 ml of chloroacetonitrile, as well as 800 μl triethylamine was added into remaining residues. After overnight stirring, solvents were removed *in vacuo* and the remaining solid was purified by flash chromatography (silica gel, 10% MeOH in CH_2Cl_2).

b. Preparation of biocytin-dpCpA

Tetrabutylammonium salt of dinucleotide (dpCpA) was a gift from Prof. Dennis Dougherty's group (Division of Chemistry and Chemical Engineering, California Institute of Technology). At room temperature, 10 mg dpCpA (8.3 μmol) and 16 mg NVOC-biocytin cyanomethyl ester (25 μmol) was mixed in dry DMF under argon. 20 μl ammonium acetate (25 mM, pH 4.5) was added to quench the reaction after 1 hour stirring. The crude product was purified by reversed phase semi-preparative HPLC using a gradient from 25 mM NH_4OAc (pH 4.5) to CH_3CN. The appropriate fractions were combined and lyophilized. The resulting solid was redissolved in 10 mM acetic acid/CH_3CN and lyophilized to give 3 mg biocytin-dpCpA as a pale yellow solid. The product was confirmed by mass spectrum as it gives a peak ($[M+H]^+$) at MW=1230.4 (Figure 3.7).

c. In vitro transcription of tRNA

THG73 tRNA was synthesized *in vitro* from FokI linearized plasmid harboring THG73 tRNA gene (gift from Prof. Dennis Dougherty's group) using T7 MEGAshortscript kit (Ambion). The product was purified by polyacrylamide gel electrophoresis and dissolved in water.

d. Ligation of biocytin to THG73 tRNA

The mixture of THG73 tRNA (25 μg in 10 μl water) and HEPES (20 μl, 10 mM, pH 7.5) was heated at 94ºC for 3 minutes and cooled down to 37ºC slowly. 8 μl biocytin-dpCpA (3 mM in DMSO), 32 μl 2.5X reaction buffer (25 μl 400 mM pH 7.5 HEPES; 10 μl 100 mM DTT; 25 μl 200 mM $MgCl_2$; 3.75 μl 10 mM ATP; 10 μl 5 mg/ml BSA; 26.25 μl water; 1 μl RNasin (Promega)), 5 μl water, as well as 5 μl T4 RNA ligase (New England Biolabs) was added. After 1 hour incubation at 37ºC, the reaction mixture was extracted once with an equal volume of phenol (saturated with 300 mM sodium acetate, pH 5.0):$CHCl_3$:isoamyl alcohol (25:24:1), then precipitated with 3 volume of cold ethanol at -20ºC. The precipitate was washed with cold 70% (v/v) ethanol, dried under vacuum, and resuspended in 5 μl 1mM sodium acetate, pH 5.0. The amount of biocytin-tRNA was quantified by measuring A_{260} and the concentration was adjusted to 1 μg/μl with 1 mM sodium acetate (pH 5.0). Prior to the suppression reaction, the biocytin-tRNA solution was deprotected by xenon lamp equipped with a 315 nm cut-off filter for 5 minutes.

2. General procedure to make mRNA-peptide fusions

a. Construction of fusion template

Synthetic DNA templates including Pep1, Pep2, and Lib1 were purified by preparative polyacrylamide gel electrophoresis. Polymerase chain reaction (PCR) of these templates with two synthetic primers, sd2 and sd26, generates double stranded DNA. mRNA was produced by T7 runoff transcription of these templates in the presence

of RNAsecure (Ambion) followed by size exclusion column purification (NAP25 column, Amersham Pharmacia Biotech). The flexible DNA linker containing puromycin, F30P (5'-$dA_{21}[C_9]_3$dACdCP; C_9=triethylene glycol phosphate, Glen Research; P = CPG-puromycin, Glen Research), was synthesized using standard chemistry. The oligonucleotide was chemically phosphorylated using phosphorylation reagent II (Glen Research) and purified by OPC cartridge. Ligation of pF30P to transcribed mRNA was done by mixing mRNA, pF30P, a splint in a 1:0.5:1.2 ratio with 2 Units of T4 DNA ligase (New England Biolabs) per picomole of template mRNA. After ligation, the fusion template was gel-purified, electroeluted and desalted by ethanol precipitation.

b. Translation and fusion formation

The fusion template was translated in reticulocyte lysate (Novagen) using standard conditions (800 nM template) with the addition of ^{35}S-methionine as the labeling reagent. In the case of templates containing UAG stop codon, 2 μg of deprotected biocytin-tRNA suppressor was also added. On completion of translation, fusion formation was stimulated by addition of $MgCl_2$ and KCl to 50 mM and 0.6 M, respectively, and incubated at -20ºC overnight. The resulting ^{35}S-labeled mRNA peptide fusions were directly loaded to 15% tricine SDS-PAGE for separation. After running, the gel was dried and exposed to phosphor screen (Molecular Dynamics) for several hours. The phosphor screen was then scanned to give image shown on Figure 3.4 B.

3. Enrichment of UAG stop codon by selection against streptavidin-agarose matrix

a. Template-base (dT) purification

To isolate fusion, the lysate was diluted in binding buffer (1M NaCl, 20 mM Tris pH 8.0, 1 mM DTT, 10 mM EDTA, 0.2% Triton X-100) and incubated with dT-cellulose at 4ºC for 1 hour. Bound fusions were washed with washing buffer (0.3M NaCl, 20 mM Tris pH 8.0) and eluted by ddH_2O.

b. Reverse transcription and selective step

Fusion after dT purification was concentrated and used for reverse transcription with Superscript II RNase H^- reverse transcriptase (BRL, life Technologies) following standard conditions recommended by the manufacturer. The reaction mixture (50 μl) was directly added into 1 ml phosphate buffer (50 mM, pH 7) and streptavidin-agarose matrix (Pierce). After a 1-hour incubation at 4ºC, the matrix was washed with washing buffer (50 mM phosphate pH 8.0, 100 mM NaCl, 0.1% SDS) 500 μl × 6 times. The matrix then was used for PCR amplification with sd2 and sd26. The PCR product was cloned with TOPO Clone kit (Invitrogen) for sequencing.

Figure 3.6 General scheme to prepare unnatural amino acid charged tRNA. The amine group of amino acid is protected by 4,5-methoxyl-2-nitrobenzyle group first, then its carboxylic group is activated by chloroacetonitrile. The activated amino acid is coupled to dinucleotide pdCpA and ligated to truncated tRNA by T4 RNA ligase.

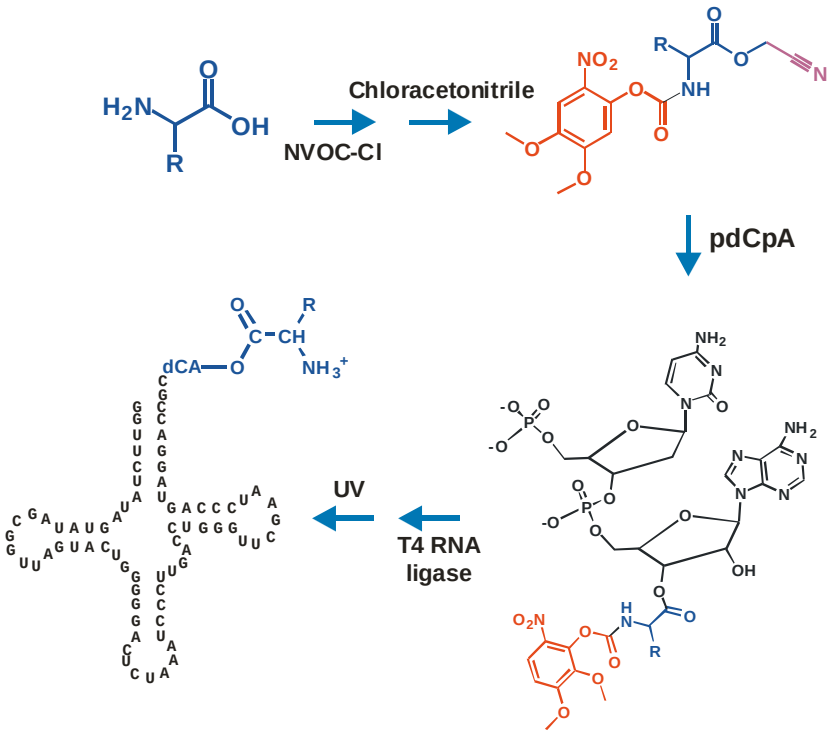
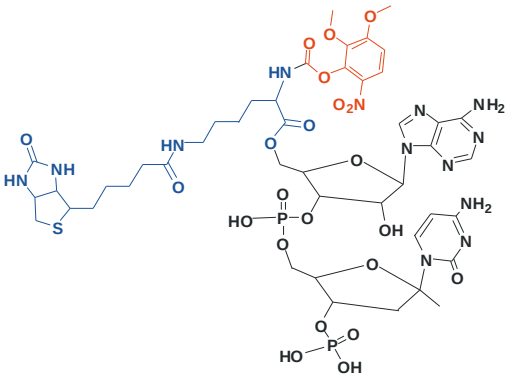
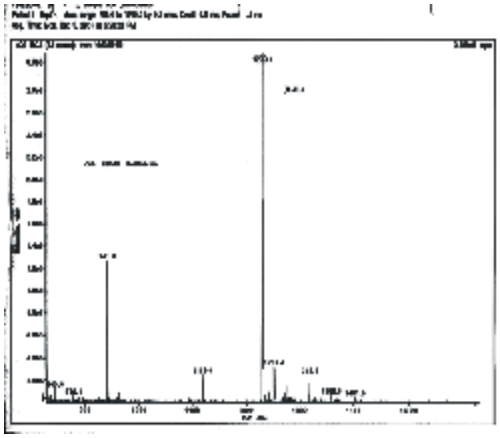


Figure 3.7 Confirmation of biocytin-pdCpA by MALDI-MS.



$C_{45}H_{61}N_{13}O_{22}P_2S$
Exact Mass: 1229.33
Mol. Wt.: 1230.05



Oligo Sequences:

Pep1

5’-ACTATTTACAACCACCATGGGCCGCCAGGAGATCCACTGGGCCAACGA
CCTGTGCAAGCCCTTCTGGGTGTACACCTCC

Pep2

5’-CTATTTACAACCACCATGGGCTAGCTTGACTACAAGGACGAGGACAAGCG
CCAGGAGATCCACTTGGCCAACGACCTGTGCAAGCCCTTCTGGGTGTACACC
TCC

Lib1

5’-ACTATTTACAACCACCATGGGCCGCCAGGAGATCCACNNSGCCAACGACC
TGTGCAAGCCCTTCTGGGTGTACACCTCC
N: 30% A, 30% C, 20% G, 20% T

Sd2

5’-GGATTCTAATACGACTCACTATAGGGACAATTACTATTTACAACCACCATG

Sd4

5’-TTTTTTTTTTTTNGCCGCCGCCGCC
N: 75% T; 9.4%A, 9.4% G; 6.2% C

Sd26

5’-GCCGCCGCCGCCGGAGGTGTACACCCAGAAG

3.5 References

1. Cornish, V. W. et al. Site-specific incorporation of biophysical probes into proteins. *Proc. Natl. Acad. Sci. USA* **91**, 2910-2914 (1994).

2. Datta, D., Wang, P., Carrico, I. S., Mayo, S. L. and Tirrell, D. A. A designed phenylalanyl-tRNA synthetase variant allows efficient *in vivo* incorporation of aryl ketone functionality into proteins. *J. Am. Chem. Soc.* **124**, 5652-5653 (2002).

3. Liu, H., Wang, L., Brock, A., Wong, C. H. and Schultz, P. G. A method for the generation of glycoprotein mimetics. *J. Am. Chem. Soc.* **125**:1702-3 (2003).

4. Arslan, T., Mamaev, S. V., Mamaeva, N. V. and Hecht, S. M. Structurally modified firefly luciferase, effects of amino acid substitution at position 286. *J. Am. Chem. Soc.* **119**, 10877-10887 (1997).

5. Nureki, O. et al. Enzyme structure with two catalytic sites for double-sieve selection of substrate. *Science* **280**, 578-582 (1998).

6. Hopfield, J. J. Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci. USA* **71**, 4135-4139 (1974).

7. Berry, M. J. et al. Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3' untranslated region. *Nature* **353**, 273-276 (1991).

8. Srinivasan, G., James, C. M. and Krzycki, J. A. Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA. *Science* **296**, 1459-1462 (2002).

9. Ellman, J., Mendel, D., Anthony-Cahill, S., Noren, C. J. and Schultz, P. G. Biosynthetic method for introducing unnatural amino acids site-specifically into proteins. *Methods Enzymol.* **202**, 301-336 (1991).

10. Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C. and Schultz, P. G. A general method of site-specific incorporation of unnatural amino acids into proteins. *Science* **244**, 182-187 (1989).

11. Gallivan, J. P., Lester, H. A. and Dougherty, D. A. Site-specific incorporation of biotinylated amino acids to identify surface-exposed residues in integral membrane proteins. *Chem. Biol.* **4**, 739-749 (1997).

12. Bain, J. D., Glabe, C. G., Dix, T. A. and Chamberlin, A. R. Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. *J. Am. Chem. Soc.* **111**, 8013-8014 (1989).

13. Saks, M. E. et al. An engineered *Tetrahymena* $tRNA^{Gln}$ for *in vivo* incorporation of unnatural amino acids into proteins by nonsense suppression. *J. Biol. Chem.* **271**, 23169-23175 (1996).

14. Wang, L., Brock, A., Herberich, B. and Schultz, P. G. Expanding the genetic code of *Escherichia coli*. *Science* **292**, 498-500 (2001).

15. Nowak, M. W. etal. *In vivo* incorporation of unnatural amino acids into ion channels in *xenopus* oocyte expression system. *Methods Enzymol.* **293**, 504-529 (1998).

16. Smith, G. P. and Petrenko, V. A. Phage Display. *Chem. Rev.* **97**, 391-410 (1997).

17. Hanes, J. and Pluckthun, A. *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl. Acad. Sci. USA* **94**, 4937-4942 (1997).

18. Roberts, R. W. & Szostak, J. W. RNA-peptide fusions for the *in vitro* selection of peptides and proteins. *Proc. Natl. Acad. Sci. USA* **94**, 12297-12302 (1997).

19. Nemoto, N., Miyamoto-Sato, E., Husimi, Y. and Yanagawa, H. *In vitro* virus: bonding of mRNA bearing puromycin at the 3'-terminal end to the C-terminal end of its encoded protein on the ribosome *in vitro*. *FEBS Lett.* **414**, 405-408 (1997).

20. Roberts, R. W. Totally *in vitro* protein selection using mRNA-protein fusions and ribosome display. *Curr. Opin. Chem. Biol.* **3**, 268-273 (1999).

21. Keefe, A. D. and Szostak, J. W. Functional proteins from a random-sequence library. *Nature* **410**, 715-718 (2001).

22. Barrick, J. E., Takahashi, T. T., Ren, J., Xia, T. and Roberts, R. W. Large libraries reveal diverse solutions to an RNA recognition problem. *Proc. Natl. Acad. Sci. USA* **98**, 12374-12378 (2001).

23. Tabuchi, I., Soramoto, S., Nemoto, N. and Husimi, Y. An *in vitro* DNA virus for *in vitro* protein evolution. *FEBS Lett.* **508**, 309-312 (2001).

24. Liu, R., Barrick, J. E., Szostak, J. W. and Roberts, R. W. Optimized synthesis of RNA-protein fusions for *in vitro* protein selection. *Methods Enzymol.* **318**, 268-293 (2000).

25. Cho, G., Keefe, A. D., Liu, R., Wilson, D. S. and Szostak, J. W. Constructing high complexity synthetic libraries of long ORFs using *in vitro* selection. *J. Mol. Biol.* **297**, 309-319 (2000).

26. Miller, S. et al. Comparison of the proteolytic susceptibilities of homologous l-amino-acid, d-amino-acid, and n-substituted glycine peptide and peptoid oligomers. *Drug Dev. Res.* **35**, 20-32 (1995).

27. Nguyen, J. T., Turck, C. W., Cohen, F. E., Zuckermann, R. N. and Lim, W. A. Exploiting the basis of proline recognition by SH3 and WW domains: design of N-substituted inhibitors. *Science* **282**, 2088-2092 (1998).

28. Hohsaka, T., Ashizuka, Y., Sasaki, H., Murakami, H. and Sisido, M. Incorporation of two different nonnatural amino acids independently into a single protein through extension of the genetic code. *J. Am. Chem. Soc.* **121**, 12194-12195 (1999).