

## In Vitro Selection of mRNA Display Libraries Containing an Unnatural Amino Acid

Shuwei Li, Steven Millward, and Richard Roberts\*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received May 3, 2002

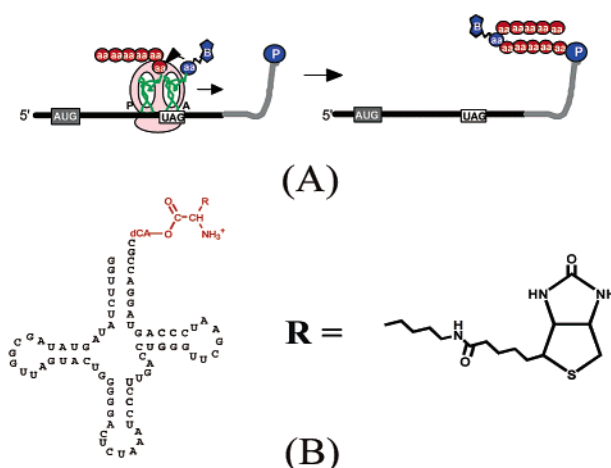
Techniques such as phage<sup>1</sup> and mRNA display<sup>2–4</sup> now make it possible to generate very large peptide and protein libraries and sieve them for functional molecules. Presently, the chemical diversity that may be programmed into these libraries is limited to the 20 naturally occurring amino acid side chains. Over 10 years ago, Schultz and co-workers demonstrated that unnatural amino acids, residues that do not occur in normal proteins, may be inserted into proteins if they are escorted to the ribosome by an orthogonal suppressor tRNA recognizing a stop codon.<sup>5,6</sup> This unnatural strategy allows proteins to be constructed that contain a novel residue at a desired location, enabling insertion of affinity tags,<sup>7</sup> spectroscopic probes,<sup>8</sup> and analogues for detailed mechanistic analysis.<sup>9</sup> This suppression strategy has also been extended to eukaryotic systems both in vitro<sup>10,11</sup> and in vivo.<sup>12,13</sup>

Recently, there has been great interest in extending the unnatural strategy to systems where the new residue may be selected for its function. Here, we demonstrate that in vitro selection experiments can be performed on mRNA display libraries containing the unnatural amino acid biocytin (Figure 1A). In mRNA display, translation extracts 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<sup>13</sup> independent peptide or protein sequences in a selectable format.<sup>14–18</sup>

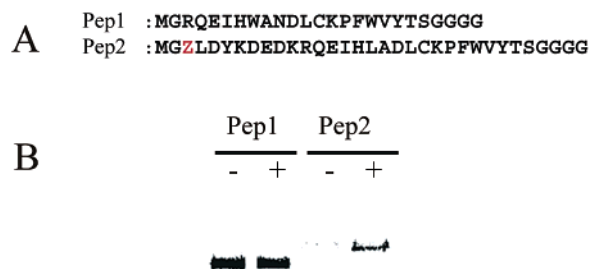
We began by working to synthesize mRNA–peptide fusions containing an unnatural residue. Biocytin, a biotin derivative of lysine, represented an excellent choice for our target residue as this has been inserted into proteins previously,<sup>7</sup> and the biotin moiety could be readily used to select sequences that have incorporated this amino acid. We normally construct mRNA display libraries in the rabbit reticulocyte translation extract due to the excellent stability of the template in this media and the efficiency of fusion formation. We chose the amber suppressor tRNA THG73 (a modified *Tetrahymena thermophila* Gln tRNA) to insert our unnatural residues by nonsense UAG suppression (Figure 1B) as this construct has high efficiency in eukaryotic translation systems.<sup>12</sup>

We next constructed two templates 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 (Figure 2A). For Pep2, fusion formation occurs only when the suppressor tRNA is added, consistent with incorporation of biocytin into the Pep2 mRNA–peptide fusion (Figure 2B).

We wished to demonstrate that this approach could be used to select peptides from libraries based on the function of the unnatural residue. To do this, we replaced the TTG codon that encoded the



**Figure 1.** (A) Scheme for insertion of unnatural amino acids into mRNA display libraries via amber suppression. (B) Biocytin-charged THG73.



**Figure 2.** 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 <sup>35</sup>S-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.

Trp residue in position 8 of the template pep1 with an NNS saturation cassette containing 32 possible codons encoding all 20 possible amino acids and the UAG stop (Lib1). We then performed two rounds of in vitro selection using streptavidin agarose as our affinity matrix (see Supporting Information for details of the experiments).

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 versus 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).

\* To whom correspondence should be addressed. E-mail: rroberts@cco.caltech.edu.

```

1       5       10      15      20      25
MetGlyArgGlnGluIleHisXXXAlaAsnAspLeuCysLysProPheTrpValTyrThrSerGlyGlyGlyGly
ATGGCCGCCAGGAGATCCACNNSGCCAACGACCTGTGCAAGCCCTTCTGGGTGTACACCTCCGGCGCGCCGC
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-----
-----CAG-----
-----TAG-----
-----TAG-----
-----GAG-----

```

**Figure 3.** Sequences present in Lib1 before and after selection vs streptavidin agarose. Amino acids sequence in three-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 four 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. Sequences the same as the original template are labeled with a dash (—).

Our experiments represent the first combination of an in vitro selection experiment and nonsense suppression of which we are aware. Our approach now allows selectable peptide and protein libraries to be constructed containing any nonnatural 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<sup>19</sup> and can show enhanced affinity for natural protein interaction modules such as SH3 domains.<sup>20</sup> Our 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 therapeu-

peutically 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<sup>21</sup> now provide the intriguing possibility for construction of wholly unnatural libraries in a selectable mRNA display format.

**Acknowledgment.** We thank Lintong Li for her help in preparing biocytin-charged-THG73. We thank Professor Dennis Dougherty's (Caltech) group for providing the dinucleotide dCpA and the plasmid for THG73 tRNA transcription. This work was supported by NIH Grant GM60416 (R.W.R).

**Supporting Information Available:** Experimental procedures (PDF). This material is available free of charge via the Internet at <http://pubs.acs.org>.

## References

- (1) Smith, G. P.; Petrenko, V. A. *Chem. Rev.* **1997**, *97*, 391–410.
- (2) Roberts, R. W.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12297–12302.
- (3) Nemoto, N.; Miyamoto-Sato, E.; Husimi, Y.; Yanagawa, H. *FEBS Lett.* **1997**, *414*, 405–408.
- (4) Roberts, R. W. *Curr. Opin. Chem. Biol.* **1999**, *3*, 268–273.
- (5) Noren, C. J.; Anthony-Cahill, S. J.; Griffith, M. C.; Schultz, P. G. *Science* **1989**, *244*, 182–187.
- (6) Ellman, J.; Mendel, D.; Anthony-Cahill, S.; Noren, C. J.; Schultz, P. G. *Methods Enzymol.* **1991**, *202*, 301–336.
- (7) Gallivan, J. P.; Lester, H. A.; Dougherty, D. A. *Chem. Biol.* **1997**, *4*, 739–749.
- (8) Cornish, V. W.; Benson, D. R.; Altenbach, C. A.; Hideg, K.; Hubbell, W. L.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2910–2914.
- (9) Arslan, T.; Mamaev, S. V.; Mamaeva, N. V.; Hecht, S. M. *J. Am. Chem. Soc.* **1997**, *119*, 10877–10887.
- (10) Bain, J. D.; Glabe, C. G.; Dix, T. A.; Chamberlin, A. R. *J. Am. Chem. Soc.* **1989**, *111*, 8013–8014.
- (11) Kargino, V. A.; Mamaev, S. V.; Hecht, S. M. *Nucleic Acids Res.* **1997**, *25*, 3912–3916.
- (12) Saks, M. E.; R., S. J.; Nowak, M. W.; Kearney, P. C.; Du, F.; Abelson, J. N.; Lester, H. A.; Dougherty, D. A. *J. Biol. Chem.* **1996**, *271*, 23169–23175.
- (13) Wang, L.; Brock, A.; Herberich, B.; Schultz, P. G. *Science* **2001**, *292*, 498–500.
- (14) Keefe, A. D.; Szostak, J. W. *Nature* **2001**, *410*, 715–718.
- (15) Barrick, J. E.; Takahashi, T. T.; Ren, J.; Xia, T.; Roberts, R. W. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 12374–12378.
- (16) Tabuchi, I.; Soramoto, S.; Nemoto, N.; Husimi, Y. *FEBS Lett.* **2001**, *508*, 309–312.
- (17) Liu, R.; Barrick, J. E.; Szostak, J. W.; Roberts, R. W. *Methods Enzymol.* **2000**, *318*, 268–293.
- (18) Cho, G.; Keefe, A. D.; Liu, R.; Wilson, D. S.; Szostak, J. W. *J. Mol. Biol.* **2000**, *297*, 309–319.
- (19) Miller, S.; Simon, R.; Ng, S.; Zuckermann, R.; Kerr, J.; Moos, W. *Drug Dev. Res.* **1995**, *35*, 20–32.
- (20) Nguyen, J. T.; Turck, C. W.; Cohen, F. E.; Zuckermann, R. N.; Lim, W. A. *Science* **1998**, *282*, 2088–2092.
- (21) Hohsaka, T.; Ashizuka, Y.; Sasaki, H.; Murakami, H.; Sisido, M. *J. Am. Chem. Soc.* **1999**, *121*, 12194–12195.

JA026789Q

# Supporting material

## *Synthesis of Biocytin-tRNA Suppressor*

### 1. 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  $\mu$ l 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<sub>2</sub>Cl<sub>2</sub>).

### 2. 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  $\mu$ mol) and 16 mg NVOC-biocytin cyanomethyl ester (25  $\mu$ mol) was mixed in dry DMF under argon. 20  $\mu$ l 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<sub>4</sub>OAc (pH 4.5) to CH<sub>3</sub>CN. The appropriate fractions were combined and lyophilized. The resulting solid was redissolved in 10 mM acetic acid/CH<sub>3</sub>CN 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.

### 3. *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.

### 4. Ligation of Biocytin to THG73 tRNA

The mixture of THG73 tRNA (25  $\mu$ g in 10  $\mu$ l water) and HEPES (20  $\mu$ l, 10 mM, pH 7.5) was heated at 94 °C for 3 minutes and cooled down to 37 °C slowly. 8  $\mu$ l biocytin-dpCpA (3 mM in DMSO), 32  $\mu$ l 2.5X reaction buffer (25  $\mu$ l 400 mM pH 7.5 HEPES; 10  $\mu$ l 100 mM DTT; 25  $\mu$ l 200 mM MgCl<sub>2</sub>; 3.75  $\mu$ l 10 mM ATP; 10  $\mu$ l 5 mg/ml BSA; 26.25  $\mu$ l water; 1  $\mu$ l RNasin (Promega)), 5  $\mu$ l water, as well as 5  $\mu$ l T4 RNA ligase (New England Biolabs) was added. After 1 hour incubation at 37 °C, the reaction mixture was extracted once with an equal volume of phenol (saturated with 300 mM sodium acetate, pH 5.0):CHCl<sub>3</sub>:isoamyl alcohol (25:24:1), then precipitated with 3 volume of cold ethanol at -20 °C. The precipitate was washed with cold 70% (v/v) ethanol, dried under vacuum, and resuspended in 5  $\mu$ l 1mM sodium acetate, pH 5.0. The amount of biocytin-tRNA was quantified by measuring A<sub>260</sub> and the concentration was adjusted to 1  $\mu$ g/ $\mu$ l 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.

## ***General procedure to make mRNA-peptide fusions***

### 1. Construction of Fusion Template

Synthetic DNA templates including Pep1 (5'-ACTATTTACAACCACCATGGGCCGCC AGGAGATCCACTGGGCCAACGACCTGTGCAAGCCCTTCTGGGTGTACACCTCC-3'), Pep2 (5'-ACTATTTACAACCACCATGGGGCTAGCTTGACTACAAGGACGAGGACAAGCG CCAGGAGATCCACTTGGCCAACGACCTGTGCAAGCCCTTCTGGGTGTACACCTCC-3'), Lib1 (5'-ACTATTTACAACCACCATGGGCCGCCAGGAGATCCACNNSGCCAACGACCTG TGCAAGCCCTTCTGGGTGTACACCTCC-3') were purified by preparative polyacrylamide gel electrophoresis. Polymerase chain reaction (PCR) of these templates with two synthetic primers, sd2 (5'-GGATTCTAATACGACTACTATAGGGACAATTACTATTTACAACCACCATG-3') and sd26 (5'-GCCGCCGCCGCCGAGGTGTACACCCAGAAG-3'), 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<sub>21</sub>[C<sub>9</sub>]<sub>3</sub>dACdCP; C<sub>9</sub>=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 (5'-TTTTTTTTTTTTTGCCGCCGCC-3') 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.

### 2. Translation and Fusion Formation

The fusion template was translated in reticulocyte lysate (Novagen) using standard conditions (800 nM template) with the addition of <sup>35</sup>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<sub>2</sub> and KCl to 50 mM and 0.6 M, respectively, and incubated at -20 °C overnight. The resulting <sup>35</sup>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 2B.

## ***Enrichment of UAG Stop Codon by Selection against Streptavidin-Agarose Matrix***

### 1. 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 4 °C for 1 hour. Bound fusions were washed with washing buffer (0.3M NaCl, 20 mM Tris pH 8.0) and eluted by ddH<sub>2</sub>O.

### 2. Reverse Transcription and Selective Step

Fusion after dT purification was concentrated and used for reverse transcription with Superscript II RNase H<sup>-</sup> 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 4 °C, 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.