

# Four-base codon mediated mRNA display to construct peptide libraries that contain multiple nonnatural amino acids

Norihito Muranaka<sup>1,2</sup>, Takahiro Hohsaka<sup>2,3,\*</sup> and Masahiko Sisido<sup>1</sup>

<sup>1</sup>Department of Bioscience and Biotechnology, Faculty of Engineering, Okayama University, 3-1-1 Tsushimanaka, Okayama 700-8530, Japan, <sup>2</sup>School of Materials Science, Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan and <sup>3</sup>PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho, Kawaguchi, Saitama 332-0012, Japan

Received May 2, 2005; Revised September 22, 2005; Accepted December 7, 2005

## ABSTRACT

***In vitro* selection and directed evolution of peptides from mRNA display are powerful strategies to find novel peptide ligands that bind to target biomolecules. In this study, we expanded the mRNA display method to include multiple nonnatural amino acids by introducing three different four-base codons at a randomly selected single position on the mRNA. Another nonnatural amino acid may be introduced by suppressing an amber codon that may appear from a (NNK)<sub>n</sub> nucleotide sequence on the mRNA. The mRNA display was expressed in an *Escherichia coli in vitro* translation system in the presence of three types of tRNAs carrying different four-base anticodons and a tRNA carrying an amber anticodon, the tRNAs being chemically aminoacylated with different nonnatural amino acids. The complexity of the starting mRNA-displayed peptide library was estimated to be  $1.1 \times 10^{12}$  molecules. The effectiveness of the four-base codon mediated mRNA display method was demonstrated in the selection of biocytin-containing peptides on streptavidin-coated beads. Moreover, a novel streptavidin-binding nonnatural peptide containing benzoylphenylalanine was obtained from the nonnatural peptide library. The nonnatural peptide library from the four-base codon mediated mRNA display provides much wider functional and structural diversity than conventional peptide libraries that are constituted from 20 naturally occurring amino acids.**

## INTRODUCTION

*In vitro* selection and directed evolution of peptide library are useful techniques for finding novel ligands that bind to target biomolecules. Several strategies for linking genotypes (genes) and phenotypes (peptides) have been developed, including phage display (1,2), ribosome display (3–5) and mRNA display (6–9). In particular, mRNA display strategy, in which covalently linked conjugates of an mRNA and its peptide product are synthesized in an *in vitro* translation system, is advantageous for its large library size and the stable mRNA–peptide linkage. Peptide ligands that bind to target molecules, such as ATP (10), streptavidin (11) and human  $\alpha$ -thrombin (12) have been successfully obtained by using the mRNA display.

Because these directed evolution processes are essentially based on the biochemical translation of mRNA library, the diversity of the amino acids is limited within the 20 naturally occurring amino acids. Recent findings on peptide drugs, however, revealed that introduction of nonnatural amino acids is one of essential points for successful drug discovery (13,14). Since the mRNA display strategy is more advantageous than the synthetic peptide library in its large library size and its potentiality for directed evolution, introduction of nonnatural amino acids into the mRNA display strategy is a promising approach towards a widely applicable and highly flexible technique for finding useful peptide ligands and drugs.

Incorporation of nonnatural amino acids into biochemically expressed proteins has been first carried out on the basis of the stop-codon suppression in the presence of a chemically aminoacylated suppressor tRNA (15,16). Recently, the stop-codon suppression strategy has been introduced to an mRNA display to construct a peptide library containing a single nonnatural amino acid (17). The stop-codon suppression method, however, is limited to incorporate only a single nonnatural

\*To whom correspondence should be addressed. Tel: +81 761 51 1681; Fax: +81 761 51 1149; Email: hohsaka@jaist.ac.jp

amino acid along a peptide, since the types of the stop codons that can be assigned to nonnatural amino acids are practically limited only to an amber (UAG) codon or, in limited cases, to other stop codons.

We have been developing a four-base codon strategy for incorporating multiple nonnatural amino acids into specific positions of proteins (18–20). Actually, two different nonnatural amino acids have been incorporated independently into single proteins (20). Introduction of the four-base codon strategy into the mRNA display method will, therefore, be a promising approach to construct a peptide library that contains multiple nonnatural amino acids.

In this study, we constructed a peptide library containing multiple nonnatural amino acids from an mRNA that contained three four-base codons AGGU, CGGU and GGGU, at a randomly selected single position. The mRNA may also contain an amber codon in the  $(\text{NNK})_n$  sequence at random positions. The effectiveness of the expanded peptide library was demonstrated by the recovery of biocytin-containing peptides from the peptide library by selection with streptavidin-coated beads. Moreover, we have isolated a novel streptavidin-binding peptide that contains a nonnatural amino acid. These results indicate that the four-base codon mediated mRNA display to construct a library of nonnatural peptides provides a widely applicable and highly flexible method for finding effective peptide ligands and drugs.

## MATERIALS AND METHODS

### Library construction

A DNA library (167mer) encoding  $\text{H}_2\text{N-MASMTGGQQ-MGZX}_9\text{-MGMSGs-COOH}$  (Z indicates one of three types of nonnatural amino acids encoded by three four-base codons, and X indicates 1 of 20 naturally occurring amino acids plus a nonnatural amino acid encoded by UAG codon) was constructed by PCR using two primers. A 5' primer T7proSD (5'-GAAATTAATACGACTCACTAT-AGGGAGACCAC-AACGGTTTCCCTCTAGAAATAATTTTGTT-TAACTT-TAAGAAGGAGATATACATATGGCTAGCATGAC-3') contained a T7 promoter, a 5' stem-loop and a Shine-Dalgarno ribosome binding site. A 3' primer T7tagZX<sub>9</sub>CR (5'-AGATCCAGA-CATTCCCAT(MNN)<sub>9</sub>ACCBACCCAT-TTGCTGTCCACCAGTCA-TGCTAGCCATATGTATAT-CTCCTTCTT-3', where M indicates A or C, N indicates A, C, G or T, and B indicates C, G or T) contained a T7 tag sequence, a random sequence VGGT(NNK)<sub>9</sub>, and a 3' constant region. The PCR product was purified by using QIAquick PCR purification kit (Qiagen), and amplified by PCR using a 5' primer T7pro (5'-GAAATTAATACGACTCACTA-3') and a 3' primer CR (5'-AGATCCAGACATTCCCAT-3'). The PCR product was purified by using the PCR purification kit, and transcribed to mRNA by using T7 RNA polymerase. The mRNA was then ligated with a 3'-puromycin-oligonucleotide 5'-AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCP-3' (prepared by Japan Bio Service, P indicates 3'-puromycin) by using a DNA splint oligonucleotide and T4 DNA ligase (New England Biolabs). The resulting product was purified by 4% PAGE with 7 M urea.

### Formation of mRNA-displayed peptide library

The puromycin-linked mRNA was translated in an *Escherichia coli* *in vitro* translation system in the presence of tRNAs chemically aminoacylated with nonnatural amino acids. Aminoacyl-tRNAs were prepared by a chemically aminoacylation technique as reported previously (18,19). In the case of biocytin-containing peptide library, one of tRNAs with ACCC, ACCG, ACCU and CUA anticodons was aminoacylated with biocytin, and the others were aminoacylated with nonnatural amino acids listed in Table 1 (Run A–D). Translation was carried out in 30  $\mu\text{l}$  of a reaction mixture containing 55 mM HEPES–KOH, pH 7.5, 210 mM potassium glutamate, 6.9 mM ammonium acetate, 12 mM magnesium acetate, 1.7 mM DTT, 1.2 mM ATP, 0.28 mM GTP, 26 mM phosphoenolpyruvate, 1 mM spermidine, 1.9% polyethyleneglycol-8000, 35  $\mu\text{g/ml}$  folic acid, 0.1 mM of each amino acid, 10  $\mu\text{M}$  anti-tmRNA oligonucleotide (5), 30 U of ribonuclease inhibitor (Takara Bio), 4.8  $\mu\text{g}$  of mRNA-dA<sub>27</sub>dCdCP, 3.3  $\mu\text{M}$  of each nonnatural aminoacyl-tRNA and 6  $\mu\text{l}$  of *E. coli* S-30 extract (Promega). The reaction mixture was incubated at 25°C for 10 min, then cooled on ice for 40 min. Peptide transfer to puromycin moiety of the mRNA-dA<sub>27</sub>dCdCP was facilitated by additional incubation at 10°C for 1 h in the presence of 500 mM KCl. After incubation, EDTA (final concentration 20 mM) was added to dissociate the ribosome complex. Then the mixture was incubated at 37°C for 20 min after addition of 1 vol of saturated NaHCO<sub>3</sub>. This treatment hydrolyzed an ester bond of the remaining aminoacyl-tRNAs without degradation of mRNA portion of the mRNA-displayed peptide. The resulting solution was neutralized with 2  $\mu\text{l}$  of acetic acid, and gel-filtered with MicroSpin G-25 column (Amersham Bioscience) to remove free nonnatural amino acids and to exchange to TBK buffer (50 mM Tris–HCl, pH 7.4, 500 mM KCl, 5 mM EDTA, 0.5  $\mu\text{g/ml}$  BSA and 0.05% Tween 20). Formation of the mRNA-displayed peptide was confirmed on 15% SDS–PAGE followed by western blotting using an anti-T7 tag antibody (Novagen) and an alkaline phosphatase-labeled anti-mouse IgG (Promega).

### MS measurement of peptide containing nonnatural amino acids

The nonnatural amino acids were incorporated into T7tag-XSNHHHHHHK-FABP (FABP; fatty acid binding protein) by using CGGU codon, in which X indicates nonnatural amino acid. Cell-free translation mixture (50  $\mu\text{l}$ ) was loaded

**Table 1.** The amount of collected mRNA-displayed nonnatural peptide

Run	Nonnatural aminoacyl-tRNA	Collected mRNA-displayed peptide	(molecules/30 $\mu\text{l}$ of <i>in vitro</i> translation)
A	biocytin-tRNA <sub>ACCU</sub> azoAla-tRNA <sub>ACCC</sub>	bzoPhe-tRNA <sub>ACCG</sub> napAla-tRNA <sub>CUA</sub>	$3.9 \times 10^7$
B	bphAla-tRNA <sub>ACCU</sub> azoAla-tRNA <sub>ACCC</sub>	biocytin-tRNA <sub>ACCG</sub> napAla-tRNA <sub>CUA</sub>	
C	bphAla-tRNA <sub>ACCU</sub> biocytin-tRNA <sub>ACCC</sub>	bzoPhe-tRNA <sub>ACCG</sub> napAla-tRNA <sub>CUA</sub>	$4.1 \times 10^8$
D	bphAla-tRNA <sub>ACCU</sub> azoAla-tRNA <sub>ACCC</sub>	bzoPhe-tRNA <sub>ACCG</sub> biocytin-tRNA <sub>CUA</sub>	
E	bphAla-tRNA <sub>ACCU</sub> azoAla-tRNA <sub>ACCC</sub>	bzoPhe-tRNA <sub>ACCG</sub> napAla-tRNA <sub>CUA</sub>	$5.1 \times 10^6$

onto 20  $\mu$ l of MagneHis Ni-Particles (Promega). The beads were washed five times with 200  $\mu$ l wash buffer that contained 10 mM Tris-HCl, pH 7.5 and 100 mM NaCl, and then eluted with 25  $\mu$ l elution buffer that contained 10 mM Tris-HCl, pH 7.5 and 500 mM imidazole. The eluate was mixed 0.1 AU of lysylendopeptidase (Wako Chemicals), then incubated at 37°C for 6 h. The resulting peptide fragments were desalted and concentrated by using ZipTipC18 (Millipore) and eluted with a matrix solution that contained saturated  $\alpha$ -cyano-4-hydroxycinnamic acid in 50% acetonitrile and 0.1% TFA. The mass measurement was performed by MALDI-TOF MS (Voyager, Applied Biosystems) in the positive mode using angiotensin II as an external calibrant.

### In vitro selection

Dynabeads M-280 Streptavidin (Dyna) (20  $\mu$ l) were washed five times with 100  $\mu$ l TBS buffer (50 mM Tris-HCl, pH 7.5 and 150 mM NaCl). The beads were gently agitated in 50  $\mu$ l of 50 mg/ml biotin-BSA (Sigma) in TBS at room temperature for 30 min. The beads were washed five times with 100  $\mu$ l TBS buffer, and gently agitated in 50  $\mu$ l of 0.4 mg/ml streptavidin (Vector) in TBS at room temperature for 30 min. After washing five times with 100  $\mu$ l TBS buffer, the beads were gently agitated in the mRNA-displayed peptide solution at 4°C for 1 h. The resulting beads were washed four times with 100  $\mu$ l TBK at 4°C for 10 min, four times with 100  $\mu$ l of 4 M urea, and twice with 100  $\mu$ l water (21). Then, mRNA-displayed peptides bound to the beads were eluted with elution buffer (10 mM EDTA and 95% formamide) by heating at 65°C for 5 min. The eluate was ethanol-precipitated, and the pellet was dissolved in 2  $\mu$ l water.

### RT-PCR and quantification of mRNA

Collected mRNA-displayed peptides were reverse-transcribed by using ReverTra Ace (TOYOBO). A reaction mixture (10  $\mu$ l), which contained 50 mM Tris-HCl, pH 7.5, 75 mM

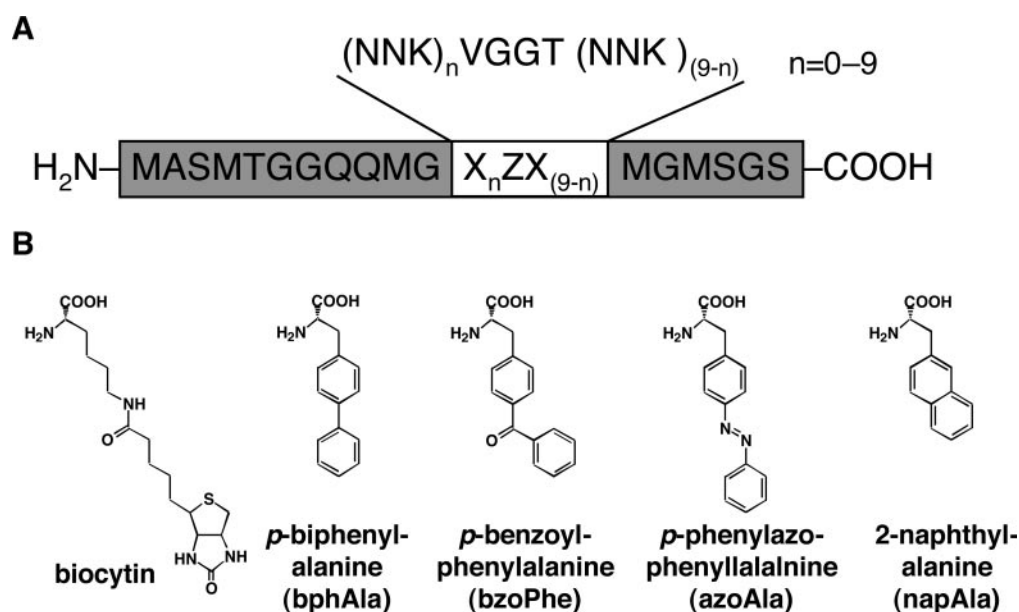
KCl, 3 mM MgCl<sub>2</sub>, 10 mM DTT, 1 mM each dNTP, 0.5  $\mu$ M 3' primer CR and 1  $\mu$ l of collected mRNA-displayed peptide solution, were heated at 70°C for 5 min. After cooling to 55°C, 50 U of ReverTra Ace was added, then the resulting mixture was incubated at 55°C for 5 min, and heated at 99°C for 5 min.

Real-time quantitative PCR using a 5'-BODIPY FL-labeled primer (22) was carried out on Smart Cycler (Cepheid) with an initial denaturation at 98°C for 10 s, primer annealing at 55°C for 2 s, and extension at 74°C for 20 s. The reaction mixture (25  $\mu$ l) contained 20 mM Tris-HCl, pH 8.2, 2 mM MgSO<sub>4</sub>, 10 mM (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 10 mM KCl, 0.1% Triton X-100, 0.1 mg/ml BSA, 0.1  $\mu$ M BODIPY FL-labeled primer 5'loop-7 (5'-BODIPY FL-CACAACGGTTTCCCTCTAGAAAT-3'), 0.25 U of KOD Dash (TOYOBO) and 5  $\mu$ l of the reverse-transcribed products. Fluorescence was measured after extension in each cycle, and the PCR was stopped after 60 cycles. The amount of collected mRNA-displayed peptides was determined from a standard curve for serial dilutions (10<sup>6</sup>-10<sup>11</sup> molecules/ $\mu$ l) of the mRNA. The real-time PCR product (138 bp) was purified by 8% PAGE, and amplified by PCR using a 5' primer 5'loop (5'-GGGAGACCACAA-CGGTTTCCCTCT-3') and 3' primer CR.

For sequencing, the PCR product was cloned into pGEM-T Vector (Promega), and individual colonies were picked for direct PCR amplification and sequencing.

### In vitro selection of streptavidin-binding nonnatural peptides

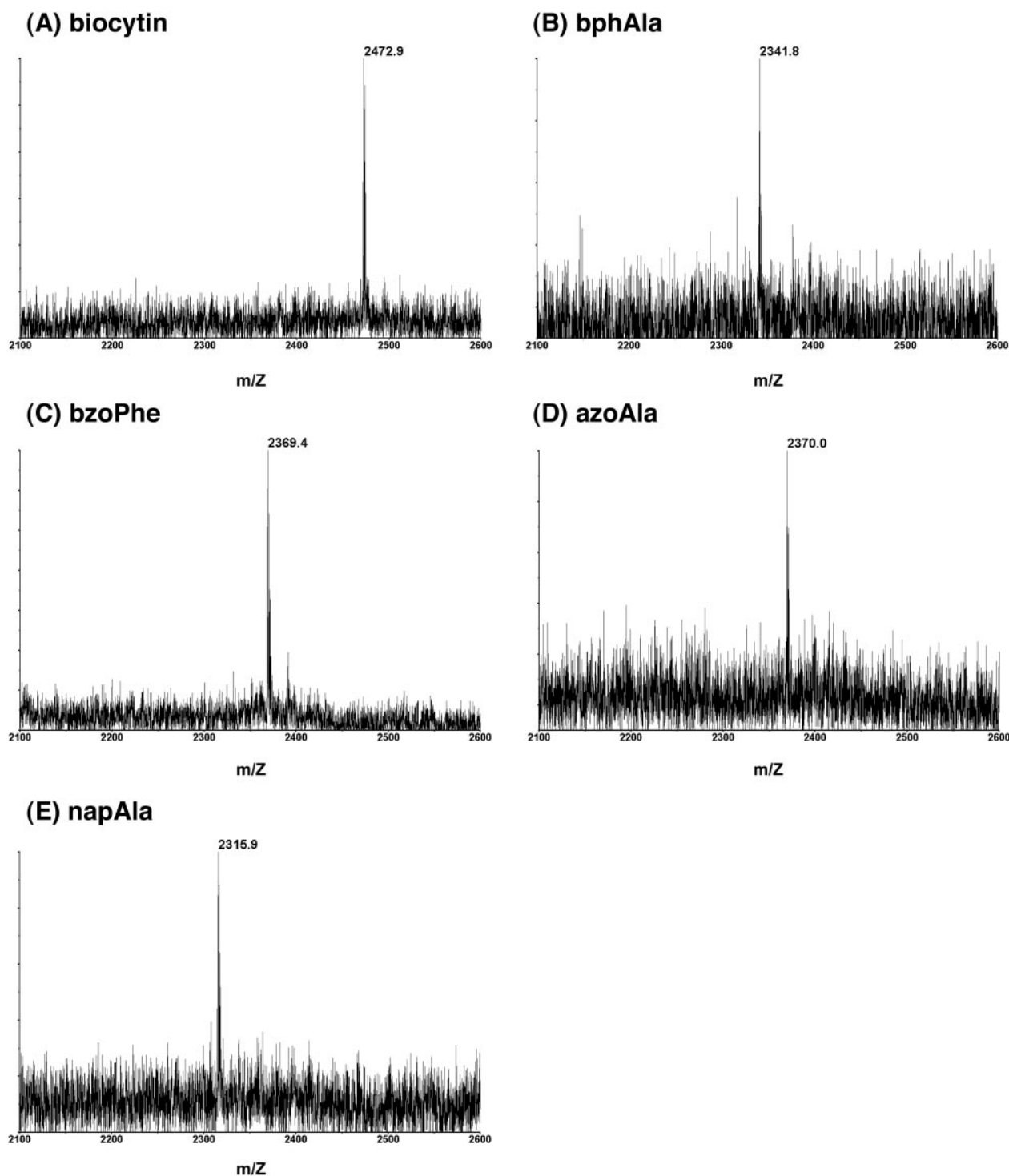
In order to prepare DNA library encoding H<sub>2</sub>N-MASMTGGQQMGX<sub>n</sub>ZX<sub>(9-n)</sub>MGMSGG-COOH ( $n = 0-9$ ), 10 PCR processes were carried out by using a 5' primer T7proSD and 3' primer T7tagX<sub>n</sub>ZX<sub>(9-n)</sub>CR (5'-AGATCCAG-ACATTCCCAT-(MNN)<sub>(9-n)</sub>ACCB(MNN)<sub>n</sub>ACCCATTTGCTGTCCACCAGTCATG-CTAGCCATATGTATATCTCCT-TCTT-3'), respectively. The PCR products were combined,



**Figure 1.** (A) Nucleotide and amino acid sequence of a nonnatural peptide library. Z indicates one of three types of nonnatural amino acids encoded by four-base codons VGGT, and X indicates 1 of 20 naturally occurring amino acids and TAG-encoded nonnatural amino acid. (B) Structures of nonnatural amino acids used in this study.

and puromycin-linked mRNA library was prepared as described above. *In vitro* translation was done in the presence of tRNAs aminoacylated with nonnatural amino acids as shown in Table 1 (Run E). Collection of mRNA-displayed peptides that bind to streptavidin was performed as described

above, and the resulting RT-PCR products were amplified by PCR using a 5' primer +T7pro (5'-GAAATTAATACGACTCACTATAGGGAGACCACAACGGTT-3') and a 3' primer CR, followed by transcription to mRNA. After eight rounds of selection, PCR products were cloned and sequenced.



**Figure 2.** MALDI-TOF MS measurements of digested peptide ASMTGGQQMGXSNHHHHHHK obtained from T7tag-XSNHHHHHHK-FABP containing biocytin (A), bphAla (B), bzoPhe (C), azoAla (D) and napAla (E) at X position. The MS spectra were externally calibrated with angiotensin II (1046.5).

## RESULTS AND DISCUSSION

### Library design

A DNA library encoding 27mer polypeptides containing 10 randomized natural and nonnatural amino acids,  $(\text{NNK})_n\text{VGGT}(\text{NNK})_{(9-n)}$ , was designed as shown in Figure 1A. The randomized region consists of a single four-base codon VGGT (V indicates A, C or G), and nine NNK codons (K indicates G or T). Four-base codons AGGT, CGGT and GGGT were chosen to encode nonnatural amino acids, because these four-base codons were found to be successfully decoded by the tRNAs containing the complementary four-base anticodons in an *E. coli in vitro* translation system (20). Amber codon TAG that may appear in the randomized  $(\text{NNK})_n$  region was also utilized to encode a nonnatural amino acid by using an amber suppressor tRNA. The use of amber suppressor tRNA was also effective to decrease the possibility of terminating of peptide elongation. The theoretical diversity of the library amounts  $21^9$  (nonapeptide of 20 natural amino acids plus 1 TAG-encoded nonnatural amino acid) times 30 (10 positions of 3 nonnatural amino acids decoded by four-base codons) =  $2.4 \times 10^{13}$  different peptide sequences. Moreover, the nonnatural peptide library adds high degree of functional and structural diversity to the peptide sequences.

The randomized region was linked with a T7 tag at the N-terminal for efficient translation in the *E. coli in vitro* translation system and for easy detection by an anti-T7 tag antibody, and with a 3' constant region at the C-terminal to ligate a DNA linker containing puromycin at the 3' end.

### Selective enrichment of mRNA-displayed peptides containing biocytin

To demonstrate the feasibility of the four-base codon mediated mRNA display method to construct a nonnatural peptide library, selective enrichment of peptides containing a particular nonnatural amino acid from a DNA library  $\text{VGGT}(\text{NNK})_9$  ( $n = 0$  in Figure 1A) was investigated. We have chosen biocytin (Figure 1B), a strong binder to streptavidin, as the nonnatural component and examined if the biocytin-containing peptides are enriched in the selection processes on streptavidin-coated beads. Biocytin was incorporated into peptides by one of three four-base codons AGGU, CGGU and GGGU, or by an amber codon UAG. A tRNA that carries one of the three complementary four-base anticodons or a CUA anticodon was aminoacylated with biocytin by chemical aminoacylation method. Codons other than biocytin-encoding one were used to introduce three of the four nonnatural amino acids, *p*-biphenylalanine (bphAla), *p*-benzoylphenylalanine (bzoPhe), *p*-phenylazophenylalanine (azoAla) and 2-naphthylalanine (napAla) (Figure 1B).

Incorporation of the nonnatural amino acids was confirmed by mass spectrometry as shown in Figure 2. The nonnatural amino acids were incorporated into T7tag-XSNHHHHHHK-FABP (X indicates nonnatural amino acid). The cell-free translation product was purified by Ni-affinity beads and digested by lysylendopeptidase, then the resulting peptide fragments were analyzed by MALDI-TOF MS. The mass peaks corresponding to the Met-processed peptide ASMT-GGQMGXSNHHHHHHK were observed for biocytin

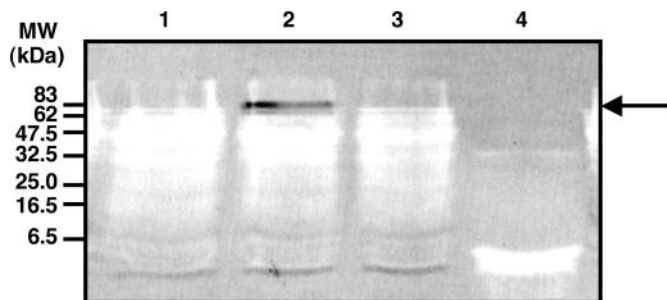
(calcd, 2473.1; found, 2472.9), bphAla (calcd, 2342.0; found, 2341.8), bzoPhe (calcd, 2370.0; found, 2369.4), azoAla (calcd, 2370.0; found, 2370.0) and napAla (calcd, 2316.0; found, 2315.9). These results support the incorporation of the nonnatural amino acids into the mRNA-displayed peptide library.

A DNA library  $\text{VGGT}(\text{NNK})_9$  was transcribed using T7 RNA polymerase, and the resulting mRNA was ligated to a DNA linker containing 3' puromycin (8). Then,  $3.6 \times 10^{13}$  molecules of puromycin-ligated mRNA were added to an *E. coli in vitro* translation system with four nonnatural aminoacyl-tRNAs to generate mRNA-displayed nonnatural peptide library (5,8,18).

The formation of mRNA-displayed peptide was confirmed by western blotting using an anti-T7 tag antibody (Figure 3). When puromycin was linked to the 3' end of mRNA, mRNA-displayed peptide was observed at around 60 kDa, which disappeared after treatment with either RNase A or Proteinase K. On the other hand, no product was detected in the absence of the 3' puromycin. This result indicates that the mRNA and translated peptide were successfully linked together through the 3' puromycin.

When an mRNA-displayed peptide was synthesized using a streptavidin mRNA containing  $\text{A}_{12}\text{dU}(\text{FITC})\text{A}_{14}\text{CCP}$  at the 3'-terminal, the yield of the mRNA-displayed peptide was estimated to be 6% of the input mRNA by fluorescence image of PAGE (data not shown). The same experiment was carried out using a mRNA- $\text{A}_{12}\text{dU}(\text{FITC})\text{A}_{14}\text{CCP}$  encoding  $\text{VGGT}(\text{NNK})_9$  peptide, but the yield could not be determined because of a broad smearing band of the mRNA with the random  $\text{VGGT}(\text{NNK})_9$  sequence. When the yield of the mRNA-displayed peptide for the nonnatural peptide library was assumed to be similar to that for the streptavidin mRNA and the average decoding efficiency of the four-base and amber codons was assumed to be 50% (20), the number of the mRNA-displayed peptide was expected to be  $1.1 \times 10^{12}$  (3% of  $3.6 \times 10^{13}$ ) in 30  $\mu\text{l}$  of cell-free translation. Although this value was lower than that for rabbit reticulocyte ( $9 \times 10^{13}/\text{ml}$ ) and wheat germ ( $1 \times 10^{14}/\text{ml}$ ) cell-free systems, it was much higher than that for phage display ( $1 \times 10^9$ ) (2,8,9).

Then, the mRNA-displayed peptide library containing biocytin was collected by using streptavidin-coated magnetic



**Figure 3.** Western blot analysis of mRNA-displayed nonnatural peptide library generated in an *E. coli in vitro* translation system in the presence of bphAla-tRNA<sub>ACCU</sub>, biocytin-tRNA<sub>ACCG</sub>, azoAla-tRNA<sub>ACCC</sub>, napAla-tRNA<sub>CUA</sub>. Lane 1 contained a mRNA-dA<sub>27</sub> without puromycin, and lanes 2–4 contained a mRNA-dA<sub>27</sub>dCdc-puromycin. The mRNA-displayed nonnatural peptide library was treated with RNase A (lane 3) and Proteinase K (lane 4).

beads. The mRNA on the beads was eluted with EDTA and formamide, and reverse-transcribed to cDNA. The resulting cDNA was amplified by a real-time quantitative PCR using a 5'-BODIPY FL-labeled primer (22). The fluorescence intensity of BODIPY FL decreased with the increase of the amount of PCR product, and the amount of the collected mRNA-displayed peptide was estimated by using a standard curve. As shown in Table 1,  $3.9 \times 10^7$  to  $8.5 \times 10^8$  molecules of mRNA-displayed peptide were collected when one of the tRNAs aminoacylated with biocytin was added to the *in vitro* translation system. On the contrary,  $5.1 \times 10^6$  molecules of mRNA-displayed peptide were recovered

when no biocytin-tRNA was added. These results indicate that the mRNA-displayed peptides containing biocytin were collected in a highly specific manner.

The real-time PCR product was amplified, cloned and sequenced (Figure 4). When the peptide library was constructed in the presence of biocytin-tRNA<sub>ACCU</sub>, 8 out of 10 clones (80%) contained AGGT four-base codon. The lacks of base in some DNA sequences may be generated during PCR. In these DNA sequences, the peptide synthesis may be continued with frameshifting at any positions. Ten out of ten clones (100%) contained CGGT in the case of biocytin-tRNA<sub>ACCG</sub>, and ten out of eleven clones (91%) contained GGGT in the case of

**(A) biocytin-tRNA<sub>ACCU</sub>**

AGGT	CTT	ATA	ATT	GGT	TGG	AGA	CGC	TTG	AT
BIO	L	I	I	G	W	R	R	L	*
AGGT	CAT	AAT	GGG	GTG	ACT	CGG	CAG	GCG	ACT
BIO	H	N	G	V	T	R	Q	A	T
AGGT	ACT	TCG	CTG	CCG	GAG	CGG	GAG	CTT	TGT
BIO	T	S	L	P	E	R	E	L	C
AGGT	AGG	CAT	GAT	GGT	TCT	ACT	CCT	ACT	CGG
BIO	R	H	D	G	S	T	P	T	R
AGGT	GGT	GCT	CAG	GTT	GCG	AAG	CCT	CCG	AGG
BIO	G	A	Q	V	A	K	P	P	R
AGGT	GCT	CTG	ACG	CTT	AAT	AGG	CCG	CCG	TAT
BIO	A	L	T	L	N	R	P	P	Y
AGGT	AAG	ACT	CAG	TTG	GAT	CTT	ACG	GAG	TGT
BIO	K	T	Q	L	D	L	T	E	C
AGGT	CTT	AGT	ACG	CAG	ACT	ACG	CCT	CCT	GCT
BIO	L	S	T	Q	T	T	P	P	A
GGGT	CGG	ATT	ACT	GCT	CGT	CCG	TCT	GGT	CGT
AZO	R	I	T	A	R	P	S	G	R
GGGT	GGT	CTT	GTT	AGG	ATA	GGC	GTA	CTC	GG
AZO	G	L	V	R	I	G	V	L	*

**(B) biocytin-tRNA<sub>ACCG</sub>**

CGGT	TCG	GGG	GAG	GAT	GTT	TCT	CAG	CCT	GAT
BIO	S	G	E	D	V	S	Q	P	D
CGGT	TTG	CAG	TCT	ACT	ATG	AAG	ATT	CAT	TTG
BIO	L	Q	S	T	M	K	I	H	L
CGGT	GGT	CCT	ATT	TTG	TGT	GAT	TGT	GTG	ACT
BIO	G	P	I	L	C	D	C	V	T
CGGT	ATT	AAT	AAT	CGG	ACA	AGG	CCT	AAG	CAG
BIO	I	N	N	R	T	R	P	K	Q
CGGT	ATG	ATT	CAT	ACA	CCT	CCT	CAT	TTT	ACG
BIO	M	I	H	T	P	P	H	F	T
CGGT	CAG	GTG	TAG	ACG	AGT	TTG	CGT	AGT	ACG
BIO	Q	V	NAP	T	S	L	R	S	T
CGGT	CAT	ATT	ACT	CAG	CTT	CAG	TAT	CCT	CCT
BIO	H	I	T	Q	L	Q	Y	P	P
CGGT	GCT	CTT	TCT	CGT	TCT	GCT	CAT	CCT	TGT
BIO	A	L	S	R	S	A	H	P	C
CGGT	AAG	CGT	ACT	GCT	ATT	CGT	CCG	CGG	GA
BIO	K	R	T	A	I	R	P	R	*
CGGT	TGT	CAG	GGT	ATT	TGT	TTT	CAG	CGT	CTT
BIO	C	Q	G	I	C	F	Q	R	L

**(C) biocytin-tRNA<sub>ACCC</sub>**

GGGT	GCT	AAG	CCT	TCT	CCT	CCT	GAT	CGT	CGT
BIO	A	K	P	S	P	P	D	R	R
GGGT	ACT	GGT	CTG	GAG	CCT	TCT	TCT	CAT	CTT
BIO	T	G	L	E	P	S	S	H	L
GGGT	TAT	TCT	CGG	GTT	GCG	ACT	TTT	GCT	ACT
BIO	Y	S	R	V	A	T	F	A	T
GGGT	GAT	AAG	TCG	CTT	CAG	ACT	CCT	CGT	CAT
BIO	D	K	S	L	Q	T	P	R	H
GGGT	TGT	AAG	CTT	CTG	GGT	AGG	AAG	AGG	CCT
BIO	C	K	L	L	G	R	K	R	P
GGGT	GGT	CCG	GCT	CGG	CTG	GGT	TGT	CTT	GC
BIO	G	P	A	R	L	G	C	L	*
GGGT	GTG	GTA	TTT	ACG	CCT	CGT	TGT	CCT	CGG
BIO	V	V	F	T	P	R	C	P	R
GGGT	GCT	TGT	CTT	CTC	GTG	TTA	TTT	GGA	CT
BIO	A	C	L	L	V	L	F	G	*
GGGT	GTT	TGT	TCT	TCG	GAT	AAG	CCT	CTT	TCG
BIO	V	C	S	S	D	K	P	L	S
GGGC	ACG	TCA	TAG	GCT	CCG	GCT	ATT	CAT	GTG
BIO	T	S	NAP	A	P	A	I	H	V
AGGT	CGT	TTT	TGT	GGT	AAG	CGG	ACT	CCG	AAG
BPH	R	F	C	G	K	R	T	P	K

**(D) biocytin-tRNA<sub>CUA</sub>**

GGGT	CCT	GAG	TGT	CGG	GCT	CCT	ACG	CTT	TAG
AZO	P	E	C	R	A	P	T	L	BIO
CGGT	TCG	CCT	AGT	TAG	GTT	CTC	AGC	TTG	CT
BZO	S	P	S	BIO	V	L	S	L	*
CGGT	CTG	CAT	TGG	ACG	AGT	TAG	AGT	TGG	GCG
BZO	L	H	W	T	S	BIO	S	W	A
CGGT	CCT	GCT	CCT	TAG	CAT	ACT	AAT	CCG	GCT
BZO	P	A	P	BIO	H	T	N	P	A
AGGT	TAG	TCT	AGT	GGT	CCT	ACG	GTT	CGT	CCG
BPH	BIO	S	S	G	P	T	V	R	P
GGGT	CTG	TAG	CCT	CGG	CCT	GGT	ACG	CCT	CGT
AZO	L	BIO	P	R	P	G	T	P	R
AGGT	GCT	ATG	ACT	AGT	ACT	CCT	ACT	TCG	TAG
BPH	A	M	T	S	T	P	T	S	BIO
CGGT	GTT	CTG	TGG	TAG	CTG	CAA	CCT	CAT	
BZO	V	L	W	BIO	L	Q	P	H	*

**Figure 4.** Sequence analyses of collected mRNA-displayed peptides from a DNA library VGGT(NNK)<sub>9</sub> with biocytin-tRNAs. Biocytin was introduced into peptides by using tRNA<sub>ACCU</sub> (A), tRNA<sub>ACCG</sub> (B), tRNA<sub>ACCC</sub> (C) and tRNA<sub>CUA</sub> (D). BIO, BPH, BZO, AZO and NAP in amino acid sequence indicate biocytin, bphAla, bzoPhe, azoAla and napAla, respectively. Asterisks indicate that the DNA sequences that contain deletion mutations, which may be suppressed by frameshifting at any positions.

biocytin-tRNA<sub>ACC</sub>. One GGGC clone obtained in the latter case may be accidentally generated during RT-PCR after *in vitro* selection. In addition, eight out of eight clones (100%) contained an amber codon at various positions in the case of biocytin-tRNA<sub>CUA</sub>. The eight clones also contained unbiased mixture of four-base codons at the first position. There is no apparent trend and no similarity to other streptavidin-binding peptides reported previously. These results clearly demonstrate that the mRNA-displayed peptides containing biocytin were specifically recovered from the mRNA-displayed peptide library containing four nonnatural amino acids including a single biocytin, and that these nonnatural amino acids were independently incorporated under the direction of four-base codons and amber codon. In addition, the results indicate a successful generation of nonnatural peptide library containing two types of nonnatural amino acids in single peptides by using both the four-base codon and the amber codon. This suggests the possibility of an mRNA-displayed peptide library containing more than two nonnatural amino acids in single peptides by adding another four-base codons to the genetic code.

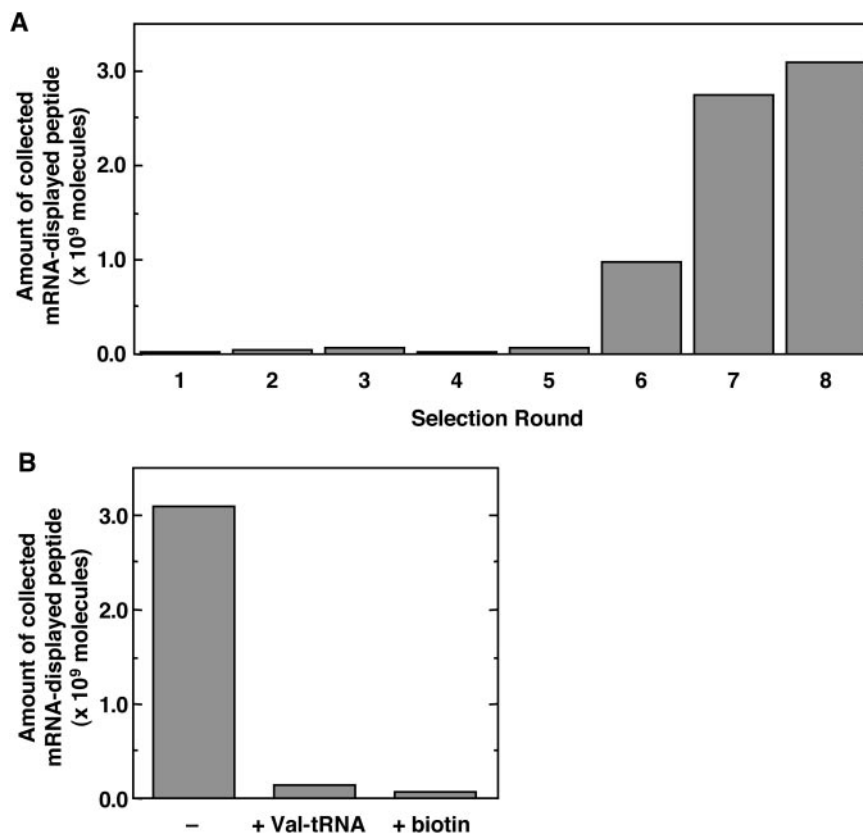
#### Selection of streptavidin-binding nonnatural peptide

As the second demonstration of the effectiveness of the four-base codon mediated mRNA display, selection of nonnatural peptides that bind to streptavidin was constructed. DNA

library of (NNK)<sub>n</sub>VGGT(NNK)<sub>(9-n)</sub> ( $n = 0-9$ ) was prepared by combining products of 10 oligonucleotide-directed PCR processes. The mRNA library was expressed in an *E. coli* *in vitro* translation system in the presence of tRNAs aminoacylated with nonnatural amino acids (Run E in Table 1). The theoretical diversity of the peptide library was  $2.4 \times 10^{13}$ , and the number of the mRNA-displayed peptide containing nonnatural amino acid was estimated to be  $1.1 \times 10^{12}$ . Thus, the coverage of the synthesized library was estimated to be 5% of the theoretical complexity.

The selection of the peptides was performed as in the case of biocytin-containing peptides. The amount of mRNA-displayed peptide eluted from streptavidin-coated magnetic beads was quantified at each selection cycle. The cDNA library obtained after the RT-PCR of the mRNA-displayed peptides was used as the template for transcription to prepare mRNA library of the next cycle. After six rounds of selection cycle, the amount of mRNA-displayed peptide increased significantly (Figure 5A). At the eighth round, the amount of mRNA-displayed peptide was 120-fold relative to that of the first round, indicating that streptavidin-binding nonnatural peptides were enriched through the repeated selection cycles. By cloning and sequencing cDNA after the eighth round, 12 out of 12 clones were found to contain a single peptide sequence, APCSSA(bzoPhe)DDV.

The incorporation of benzoylphenylalanine is interpreted neither by high decoding efficiency of the nonnatural amino



**Figure 5.** Selection of streptavidin-binding nonnatural peptides from an mRNA-displayed nonnatural peptide library containing bphAla, bzoPhe, azoAla and napAla. (A) The amount of collected mRNA-displayed peptides (molecules/30  $\mu$ l of *in vitro* translation) was quantified by real-time PCR at each selection cycle. (B) The amount of collected mRNA-displayed peptide at the eighth round of selection. As control experiments, four nonnatural aminoacyl-tRNAs were replaced by Val-tRNAs (+Val-tRNA), and the selection procedure was performed under the presence of 500  $\mu$ M biotin (+biotin).

acid nor by the high decoding efficiency of the CGGU codon. Actually, the decoding efficiency of CGGG four-base codon by benzoylphenylalanine is lower than those by biphenylalanine and 2-naphthylalanine and a little higher than that of phenylazophenylalanine (19). Similarly the decoding efficiency of CGGU codon by tRNA<sub>ACCG</sub> is lower than that of GGGU codon by tRNA<sub>ACCC</sub> (20). Therefore, the recovery of the nonnatural peptide containing benzoylphenylalanine was not due to biased decoding of the four-base codons, but a result of selection processes.

Incorporation of valine instead of nonnatural amino acids into the eighth round clone significantly decreased the amount of mRNA-displayed peptide (Figure 5B). In addition, when the selection procedure was performed in the presence of 500  $\mu$ M biotin, a very low amount of mRNA-displayed peptide was recovered. These results indicate that the selected peptide containing a single benzoylphenylalanine specifically binds to streptavidin. Benzoylphenylalanine itself does not have affinity to streptavidin, but it may interact with a hydrophobic biotin-binding site of streptavidin in association with the surrounding amino acids. Any peptide containing two nonnatural amino acids was not obtained in the present experiment, possibly because two hydrophobic residues may not be needed for the binding to streptavidin.

A bzoPhe containing 10mer peptide APCSSA(bzoPhe)-DDV labeled with FITC at the N-terminal was chemically synthesized and its binding activity to streptavidin in solution was evaluated by fluorescence polarization measurement. The fluorescence polarization of the FITC-labeled peptide was, however, not changed upon addition of streptavidin. This result suggests that the 10mer peptide does not have the binding activity and the adjacent constant region may be necessary. Taken together with the result that the amount of the collected mRNA-displayed peptide was decreased when bzoPhe was substituted by Val (Figure 5B), the bzoPhe residue binds to streptavidin in association with the surrounding amino acids including constant region.

## CONCLUSIONS

Four-base codon mediated mRNA display method was developed to construct a peptide library that contained multiple nonnatural amino acids at a randomly selected single position along a peptide chain. Four-base codons AGGU, CGGU and GGGU were used to assign three different nonnatural amino acids. Another nonnatural amino acid was also introduced by suppressing UAG amber codon that appeared in the (NNK)<sub>n</sub> nucleotide sequence on the mRNA.

The most important advantage of our strategy is that four different types of nonnatural amino acids are encoded in peptide library by three four-base codons and one stop codon. In addition, it may be possible to obtain ligand peptides containing two types of nonnatural amino acids in the single peptide chain. A bzoPhe containing peptide was, in fact, selected from the nonnatural peptide library containing four types of nonnatural amino acids in Figure 5. Moreover, peptides containing biocytin and another nonnatural amino acid were obtained in Figure 4, although the latter one would not contribute to the binding to streptavidin. These results demonstrate the advantage of our nonnatural peptide library over previously reported library containing only one nonnatural

amino acid (17). The present nonnatural peptide library has a wide functional and structural diversity and will make a powerful tool to discover new artificial ligands to target proteins or other biomolecules.

## ACKNOWLEDGEMENTS

We thank Professor Kenzo Fujimoto for help with MALDI-TOF MS measurement. M.S. wishes to thank a Grand-in-Aid for Scientific Research (S) from the Ministry of Education, Science, Sports and Culture, Japan (No.15101008). N.M. is also grateful to a VBL grant from Okayama University. Funding to pay the Open Access publication charges for this article was provided by the Ministry of Education, Science, Sports and Culture, Japan.

*Conflict of interest statement.* None declared.

## REFERENCES

- Smith, G.P. (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science*, **228**, 1315–1317.
- Smith, G.P. and Petrenko, V.A. (1997) Phage display. *Chem. Rev.*, **97**, 391–410.
- Mattheakis, L.C., Bhatt, R.R. and Dower, W.J. (1994) An *in vitro* polysome display system for identifying ligands from very large peptide libraries. *Proc. Natl Acad. Sci. USA*, **91**, 9022–9026.
- Hanes, J. and Plückthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display. *Proc. Natl Acad. Sci. USA*, **94**, 4939–4942.
- Hanes, J., Jermutus, L. and Plückthun, A. (2000) Selecting and evolving functional proteins *in vitro* by ribosome display. *Methods Enzymol.*, **328**, 404–430.
- Robert, R.W. and Szostak, J.W. (1997) RNA-peptide fusions for the *in vitro* selection of peptides and proteins. *Proc. Natl Acad. Sci. USA*, **94**, 12297–12302.
- Nemoto, N., Miyamoto-Sato, E., Husimi, Y. and Yanagawa, H. (1997) *In vitro* virus: bound 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.
- Liu, R., Barrick, J.E., Szostak, J.W. and Roberts, R.W. (2000) Optimized synthesis of RNA-protein fusion for *in vitro* protein selection. *Methods Enzymol.*, **318**, 268–293.
- Miyamoto-Sato, E., Takashima, H., Fuse, S., Sue, K., Ishizaka, M., Tateyama, S., Horisawa, K., Sawasaki, T., Endo, Y. and Yanagawa, H. (2003) Highly stable and efficient mRNA templates for mRNA-protein fusion and C-terminally labeled proteins. *Nucleic Acids Res.*, **31**, e78.
- Keef, A.D. and Szostak, J.W. (2001) Functional proteins from a random-sequence library. *Nature*, **410**, 715–718.
- Wilson, D.S., Keefe, A.D. and Szostak, J.W. (2001) The use of mRNA display to select high-affinity protein-binding peptides. *Proc. Natl Acad. Sci. USA*, **98**, 3750–3755.
- Raffler, N.A., Schneider-Mergener, J. and Famulok, M. (2003) A novel class of small functional peptides that bind and inhibit human  $\alpha$ -thrombin isolated by mRNA display. *Chem. Biol.*, **10**, 69–79.
- Burke, T.R., Jr, Smyth, M.S., Otaka, A., Nomizu, M., Roller, P.P., Wolf, G., Case, R. and Shoelson, S.E. (1994) Nonhydrolyzable phosphotyrosyl mimetics for the preparation of phosphatase-resistant SH2 domain inhibitors. *Biochemistry*, **33**, 6490–6494.
- Tamamura, H., Omagari, A., Hiramoto, K., Oishi, S., Habashita, H., Kanamoto, T., Gotoh, K., Yamamoto, N., Nakashima, H., Otaka, A. *et al.* (2002) Certification of the critical importance of L-3-(2-naphthyl)alanine at position 3 of a specific CXCR4 inhibitor, T140, leads to an exploratory performance of its downsizing study. *Bioorg. Med. Chem.*, **10**, 1417–1426.
- Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C. and Schultz, P.G. (1989) A general method for site-specific incorporation of unnatural amino acids into proteins. *Science*, **244**, 182–188.



16. Bain, J.D., Glabe, C.G., Dix, T.A., Chamberlin, A.R. and Diala, E.S. (1989) Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. *J. Am. Chem. Soc.*, **111**, 8013–8014.
17. Li, S., Millward, S. and Roberts, R.W. (2002) *In vitro* selection of mRNA display containing an unnatural amino acid. *J. Am. Chem. Soc.*, **124**, 9972–9973.
18. Hohsaka, T., Ashizuka, Y., Murakami, H. and Sisido, M. (1996) Incorporation of nonnatural amino acids into streptavidin through *in vitro* frameshift suppression. *J. Am. Chem. Soc.*, **118**, 9778–9779.
19. Hohsaka, T., Kajihara, D., Ashizuka, Y., Murakami, H. and Sisido, M. (1999) Efficient incorporation of nonnatural amino acids with large aromatic groups into streptavidin in *in vitro* protein synthesizing systems. *J. Am. Chem. Soc.*, **121**, 34–40.
20. Hohsaka, T., Ashizuka, Y., Taira, H., Murakami, H. and Sisido, M. (2001) Incorporation of nonnatural amino acids into protein by using various four-base codons in an *Escherichia coli in vitro* translation system. *Biochemistry*, **40**, 11060–11064.
21. Bessho, Y., Hodgson, D.R.W. and Suga, H. (2002) A tRNA aminoacylation system for nonnatural amino acids based on a programmable ribozyme. *Nat. Biotechnol.*, **20**, 723–728.
22. Kurata, S., Kanagawa, T., Yamada, K., Torimura, M., Yokomasu, T., Kamagata, Y. and Kurane, R. (2001) Fluorescent quenching-based quantitative detection of specific DNA/RNA using a BODIPY FL-labeled probe or primer. *Nucleic Acids Res.*, **29**, e34.