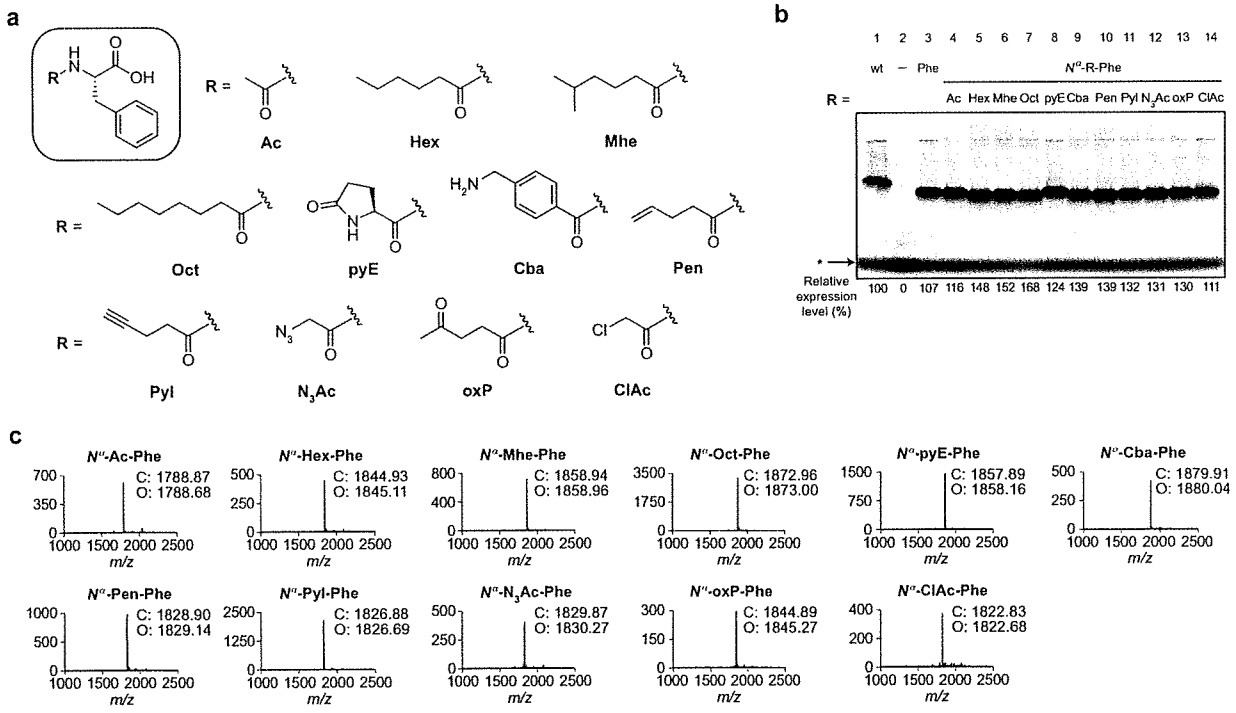


Figure 4. Tolerance of N^{α}-acyl groups in initiation. a) N^{α}-acyl phenylalanine derivatives used in this study. b) Tricine–SDS PAGE analysis of the translation products initiated with various N^{α}-acyl phenylalanine derivatives. Lane 1, expression of wild type; lane 2, in the absence of Met; lane 3, initiated with Phe-tRNA$^{fMet}_{CAU}$; lanes 4–14, initiated with various N^{α}-acyl-Phe-tRNA$^{fMet}_{CAU}$ molecules. Each expression level relative to wild type is determined by a mean score of duplicates. The band indicated by an asterisk corresponds to the remaining [^{14}C]-Asp that was not incorporated into the peptide. c) MALDI-TOF MS spectra of the translated peptides. The calculated mass (C) and observed mass (O) are shown in each spectrum.

thus designed a mRNA sequence to express the linear peptide initiated withN^{α}-ClAc-Trp in the *w*PURE system (Figure 5, panel b). After expression, the resulting peptide was purified by molecular weight cut and desalting filters. MALDI-TOF analysis of the peptide gave peaks of the cyclic G7–18NATE with the expected molecular masses of $[M + H]^+$ and $[M + Na]^+$ (Figure 5, panel c), and MALDI-TOF/TOF analysis revealed peaks corresponding to the peptide fragments bearing the intact thioether linkage (Figure 5, panel d). This was a sharp contrast to the observation that no cyclization occurred in a control peptide in which the C-terminal Cys was substituted with Ser; *i.e.*, the N^{α}-ClAc group did not react potential other sulfhydryl reagents such as mercaptoethanol (Figure S3). Remarkably, such a simple purification protocol yielded the desired cyclic peptide as the main product, while no peak corresponding to the unreacted linear N^{α}-ClAc-(G7–18NATE) peptide was observed, suggesting that cyclization of the linear peptide occurred rapidly. Furthermore, although MALDI-TOF analysis does not permit the quantitative discussion, it should be noted that the undesired minor peaks, which could be caused by side reactions, *e.g.*, intermolecular reactions with other thiol molecules, such as DTT, mercaptoethanol, or free Cys, included in the *w*PURE system, were not observed in the mass spectrum. This indicated that the cyclization did not substantially suffer from such side reactions. In fact, because the intramolecular cyclization took place without undesirable side reactions, we were able to isolate the desirable full-length peptide through a simple purification procedure using a molecular weight cut filter.

Moreover, the expression level of G7–18NATE was quantified by the incorporated radioisotope counts of [^{14}C]-Asp calibrated against known [^{14}C]-Asp concentrations (see Figure S4). G7-18NATE was expressed at a

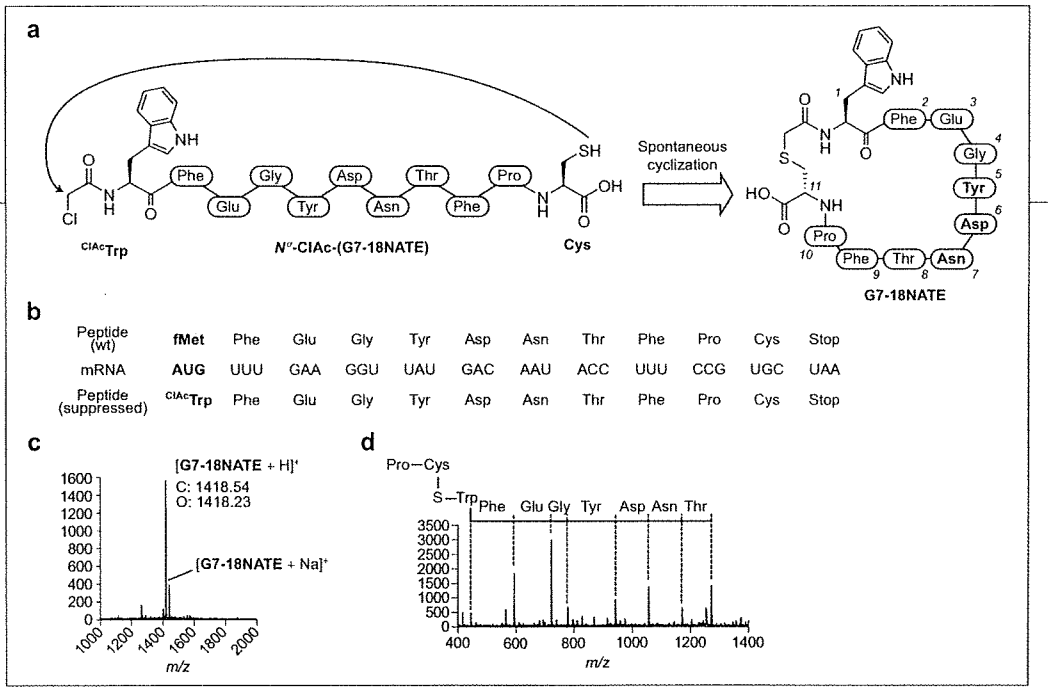


Figure 5. Ribosomal synthesis of G7−18NATE. a) The spontaneous cyclization affording G7−18NATE. The precursor linear peptide, N^{α}-ClAc-G7−18NATE, translated in the *w*PURE system spontaneously cyclizes into G7-18NATE by intramolecuar attack of Cys sulfhydryl group to the α-carbon of the N^{α}-ClAc group. b) The sequence of mRNA used in the ribosomal synthesis of G7−18NATE. For the synthesis of G7−18NATE, the start codon is suppressed with N^{α}-ClAc-Trp. c) MALDI-TOF mass spectrum of the expressed G7−18NATE. The calculated mass (C) and observed mass (O) are shown in the figure. The peaks at 1263.97 and 1401.23 *m/z* were occasionally observed when the molecular weight cut purification protocol was utilized, regardless of template sequence. Therefore, it is likely that these peaks originated from components in the *w*PURE system. d) MALDI-TOF/TOF analysis of G7−18NATE expressed in the *w*PURE system. The representative peaks that correspond to the peptide fragments containing the thioether linkage are labeled in the spectrum.

concentration of approximately 3.7 μM (5.3 μg/mL), which is comparable to the control expression initiated with Met generated by the normal PURE system (~5 μM), indicating that the reprogrammed initiation was achieved without significant reduction in expression level.

To further examine the flexibility of this strategy for the synthesis of cyclic peptides, we designed three peptide sequences based on the structure of G7−18NATE with arbitrarily chosen ring sizes (Figure 6, panel a). Three amino acids (Tyr, Asp, and Asn), which are the most critical residues for the binding activity to Grb7, were kept in the structural designs of these cyclic peptides (*30*), and the ring size was varied from 11 (as in the original G7−18NATE) to 4, 6, or 14 residues. These cyclic peptides were designed to include a modified Flag peptide at the C-terminus in order to ensure MALDI-TOF ionization. MALDI-TOF analysis of the translation products showed that each of the expressed peptides spontaneously cyclized upon translation of the corresponding linear peptide (Figure 6, panel b). This result suggested that the *in situ* cyclization process is independent of peptide sequence and ring size. A limitation of this technology is likely that the peptide can have only a single Cys residue in order to avoid undesirable competing cyclization reactions, but yet it is widely applicable to a variety of sequences of cyclic peptides.

It should be noted that G7−18NATE was an engineered synthetic peptide designed based on the parental peptide (G7−18NA) (*30*). G7−18NA was originally selected by phage display and thus composed of proteinogenic amino acids linked with a Cys−Cys disulfide bond (see Figure S2b). Despite the fact that the engineering of G7−18NA to G7−18NATE reduced the binding activity to Grb7, its *in vivo* stability resulted in successful suppression of pancreatic cancer metastasis in mice (*30*–*32*). Ribosomal synthesis of G7−18NATE would enable us to construct a mRNA-programmed G7−18NATE-like peptide library for reoptimizing the binding to Grb7. Particularly, the coupling of this technology with an *in vitro* display system (*36*–*38*) would allow us to rapidly screen the above library without sacrificing the peptide's *in vivo* compatibility.

Construction of a Combinatorial Cyclic Peptide Library. We have demonstrated the generality of the *in situ* cyclization of peptides with various sequences and ring sizes. This indicates that a variety of *in vivo* compatible cyclic peptide libraries can be readily prepared by simply designing mRNA sequence libraries. As such a demonstration, we have constructed a peptide library consisting of 160 distinct cyclic peptides. We prepared 160 DNA templates, coding 6-mer peptides bearing Cys at the C-terminus, each of which has two varying codons assigning Xaa1 and Xaa2 (Figure 7, panels a and b, and Figure S4); then, peptides were expressed in parallel from the respective DNA templates in the presence of N^{α}-ClAc-Tyr-$tRNA^{fMet}_{CAU}$ using the *w*PURE sys-

tem. To confirm the quality of products, 12 samples were randomly chosen from the library and were passed through a desalting filter. Remarkably, even without vigorous purifications, we were able to detect the desired cyclic peptide with the expected molecular weight as a sole product in all samples (Figure 7, panel c). This demonstration proves the practicality of this technology for the synthesis of cyclic peptide libraries closed by the nonreducible thioether bond. We are currently pursuing the screening of extended cyclic peptide libraries prepared by this platform technology against chosen therapeutic targets.

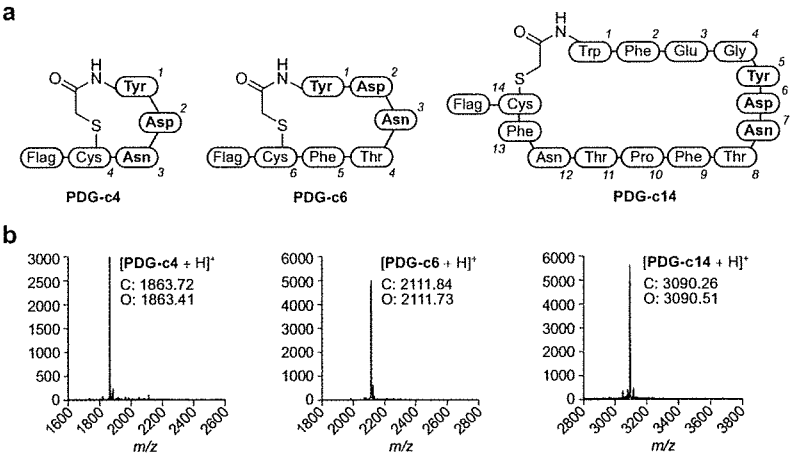


Figure 6. Custom synthesis of the cyclic peptides with various ring sizes. a) Structure of the expressed cyclic peptides. Three amino acids known to be critical residues in G7–18NATE are shown in bold, and "Flag-tag" indicates a modified Flag peptide sequence (LTTDYKDDDDK). b) MALDI-TOF mass spectrum of the expressed cyclic peptides. The calculated mass (C) and observed mass (O) are shown in the spectra.

CONCLUSION

In conclusion, we have applied the concept of genetic code reprogramming to the engineering of the initiation event in translation by means of the *w*PURE and flexizyme systems. The translation apparatus surprisingly tolerates all proteinogenic amino acids in the initiation event, particularly those having hydrophobic side chains with high efficiencies. It also accepts a wide array of N^{α}-acyl groups containing various functionalities. This technology has been applied to the ribosomal synthesis of cyclic peptides linked *via* a physiologically stable thioether bond, where the cyclization occurs spontaneously upon translation of the precursor linear peptide. Indeed, we have demonstrated the construction of a high-quality library consisting of 160 distinct cyclic peptides by simply designing a mRNA library. Because this new platform technology is widely applicable to the synthesis of various cyclic peptide libraries containing not only natural but also nonproteinogenic amino acids using reprogramming of the elongation event, it is a powerful tool to accelerate the discovery of peptidic drug candidates against various therapeutic targets.

METHODS

General Protocol of Translation. Xaa-tRNA$^{fMet}_{CAU}$ was prepared by the following procedure. tRNA$^{fMet}_{CAU}$ (40 μM) in 0.2 M Hepes-K (Hepes, 2-[4-(2-hydroxyethyl)-1-piperadinyl]ethansulfonic acid) buffer pH 7.5, 0.2 M KCl (7.5 μL) was heated at 95 °C for 3 min and cooled to 25 °C for 5 min. $MgCl_2$ (3 M, 3 μL) and flexizyme (dFx or eFx, see ref (*20*)) (200 μM, 1.5 μL) were added, and the mixture was incubated at 25 °C for 5 min. The reaction was initiated by addition of 3μL of 25 mM substrate (N^{α}-acyl-amino acid 3,5-dinitrobenzyl ester or cyanomethyl ester) in dimethyl sulfoxide and incubated on ice for the optimized times, generally 2–6 h (*20*). After acylation, the reaction was stopped by addition of 45 μL of 0.6 M sodium acetate at pH 5, and the RNA was recovered by ethanol precipitation. The pellet was rinsed twice with 70% ethanol with 0.1 M sodium acetate, pH 5.0, and once with 70% ethanol. The Xaa-tRNA$^{fMet}_{CAU}$ was dissolved in 0.5 μL of 1 mM sodium acetate just before adding to translation mixture.

The *w*PURE system containing all necessary components for translation except for all 20 standard amino acids was used in this study. Translation was carried out using *w*PURE system with 0.04 μM mDNA1 containing 200 μM each Thr, Tyr, and Lys 50 μM [^{14}C]-Asp, and 120 μM of various Xaa-tRNA$^{fMet}_{CAU}$ molecules. The wild-type expression was carried out with *w*PURE system with 0.04 μM mDNA1 and 3 mM ethylenediamine tetraacetic acid containing 200 μM each Thr, Tyr, and Lys and 50 μM [^{14}C]-Asp. The translation mixture (2.5 μL) was incubated at 37 °C for 1 h and analyzed by Tricine–SDS-PAGE and autoradiography (FLA-5100, Fuji, Japan).

Mass Spectra Measurements of Peptides. For mass spectra analysis, translation reaction (5 μL) was performed in the presence of Asp, instead of [^{14}C]-Asp. The translation product from mDNA1 was immobilized with FLAG-M2 agarose (Sigma). After the resin was washed with 30 μL of W buffer (50 mM Tris-HCl (Tris, tris(hydroxymethyl)aminomethane) pH 8.0, 150 mM NaCl), the immobilized peptides were eluted with 10 μL of 0.2% trifluoroacetic acid (TFA). The purified peptide was desalted with ZipTip$_{\mu\text{-C18}}$ (Millipore), and eluted with 1 μL of a 50% acetonitrile, 0.1% TFA solution saturated with the matrix (*R*)-cyano-4-hydroxycinnamic acid. MALDI MS measurement was performed

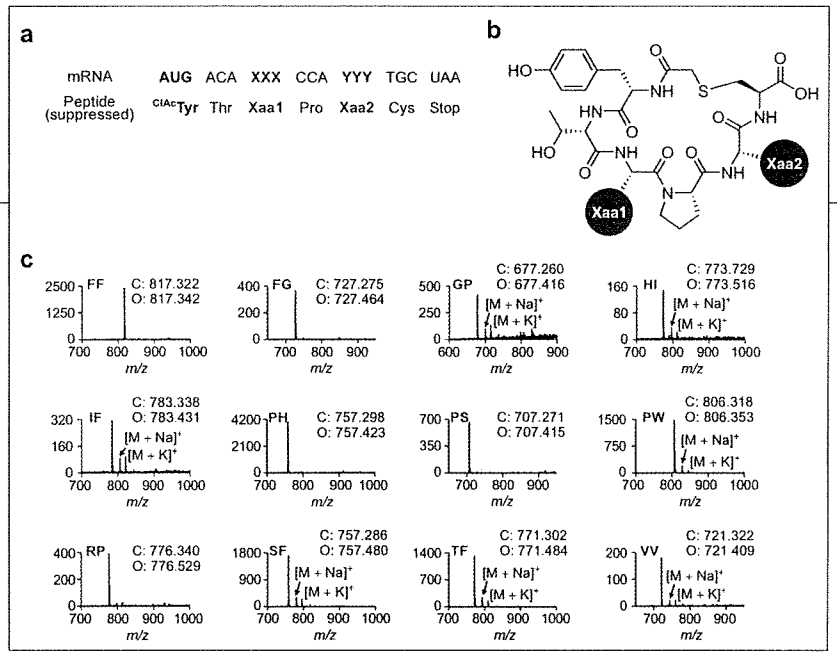


Figure 7. Construction of a cyclic peptide library. a) The sequence of mRNA used in the library synthesis. The codons indicated in red (Xaa1) and blue (Xaa2) were changed in each expression. The start codon is suppressed with N^{α}-ClAc-Tyr. b) The structure of cyclic peptide contained in the peptide library. c) MALDI-TOF analysis of the components of the peptide library. The amino acids corresponding to Xaa1 and Xaa2 were shown in red and blue in each spectrum, respectively. A one-letter amino acid abbreviation was used. The peaks labeled by $[M + Na]^+$ and $[M + K]^+$ correspond to sodium adduct and potassium adduct of the desirable product, respectively. The calculated mass (C) and observed mass (O) are shown in each spectrum.

using autoflex II TOF/TOF (Bruker Daltonics) under the linear/positive mode and externally calibrated with peptide calibration standard II (Bruker Daltonics).

Ribosomal Synthesis of G7–18NATE. Translation was carried out using *w*PURE system with 0.04 μM mDNA2 containing 200 μM each Phe, Glu, Gly, Tyr, Asn, Thr, Pro, Asp, and Cys and 120 μM N^{α}-ClAc-Trp-$tRNA^{fMet}_{CAU}$. The translation mixture (5 μL) was incubated at 37 °C for 1 h. The product was diluted three times with water and purified through Microcon YM-10 (Millipore) (10,000 MW cut filter). The purified peptides were desalted with $ZipTip_{\mu\text{-}C18}$ (Millipore) and eluted with 1 μL of a 50% acetonitrile, 0.1% TFA solution saturated with the matrix (*R*)-cyano-4-hydroxycinnamic acid. MALDI-MS measurements were performed using autoflex II TOF/TOF (Bruker Daltonics) under the linear/positive mode and externally calibrated using peptide calibration standard II (Bruker Daltonics). MALDI-TOF/TOF analysis was also carried out using autoflex II TOF/TOF under the lift mode.

Construction of Combinatorial Cyclic Peptide Library. Translation reactions were carried out in parallel using 96-well plates. 0.04 μM DNA templates were mixed with *w*PURE system containing 22 μM N^{α}-ClAc-Trp-$tRNA^{fMet}_{CAU}$. The expressions of samples #1–144 were performed with aa mix A (Ile, Ala, Gly, Leu, Val, Pro, Asn, Ser, Thr, Tyr, Phe, and Trp; 500 μM each), and the expressions of sample #145–160 were performed with aa mix B (Pro, Arg, Asp, Glu, Gln, His, Ile, Ala, Thr, Tyr, Asn; 500 μM each). The translation mixture (40 μL) was incubated at 37 °C for 1 h. For characterization of the product, randomly selected samples (#11, 30, 53, 68, 72, 95, 107, 123, 131, 149, 150, and 156) were analyzed by MALDI-TOF mass spectrometry. The translation product (3 μL) was acidified by adding 5 μL of 1% TFA and desalted with $ZipTip_{\mu\text{-}C18}$ (Millipore). The sample was eluted with 1 μL of a 0.1% TFA solution containing 50% acetonitrile and 2,5-dihydroxybenzoic acid (10 mg mL^{-1}) onto MTP 384 target plate ground steel T F (Bruker Daltonics). MALDI-MS measurements were performed using autoflex II TOF/TOF under the linear/positive mode and externally calibrated using peptide calibration standard II (Bruker Daltonics).

Acknowledgment: We thank Y. Ohshiro for providing N^{α}-ClAc-tyrosine CME and Dr. P. C. Reid for the discussion and proofreading. This work was supported by grants from the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) to H.S., a Grants-in-Aid for JSPS Fellows (18-10526) to Y.G., a research and development project of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO) to H.S., and the Industrial Technology Research Grant Program in NEDO (05A02513a) to H.M.

Supporting Information Available: This material is available free of charge *via* the Internet.

REFERENCES

1. Kozak, M. (1983) Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles, *Microbiol. Rev. 47*, 1–45.
2. Gualerzi, C. O., and Pon, C. L. (1990) Initiation of mRNA translation in prokaryotes, *Biochemistry 29*, 5881–5889.
3. Gold, L. (1988) Posttranscriptional regulatory mechanisms in *Escherichia coli*, *Annu. Rev. Biochem. 57*, 199–233.
4. Varshney, U., and RajBhandary, U. L. (1990) Initiation of protein synthesis from a termination codon, *Proc. Natl. Acad. Sci. U.S.A. 87*, 1586–1590.
5. Chattapadhyay, R., Pelka, H., and Schulman, L. H. (1990) Initiation of *in vivo* protein synthesis with non-methionine amino acids, *Biochemistry 29*, 4263–4268.
6. Mamaev, S., Olejnik, J., Olejnik, E. K., and Rothschild, K. J. (2004) Cell-free N-terminal protein labeling using initiator suppressor tRNA, *Anal. Biochem. 326*, 25–32.
7. Olejnik, J., Gite, S., Mamaev, S., and Rothschild, K. J. (2005) N-terminal labeling of proteins using initiator tRNA, *Methods 36*, 252–260.
8. Giege, R., Sissler, M., and Florentz, C. (1998) Universal rules and idiosyncratic features in tRNA identity, *Nucleic Acids Res. 26*, 5017–5035.
9. Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code, *Annu. Rev. Biophys. Biomol. Struct. 35*, 225–249.
10. Hecht, S. M. (1992) Probing the Synthetic Capabilities of a Center of Biochemical Catalysis, *Acc. Chem. Res. 25*, 545–552.
11. Link, A. J., Mock, M. L., and Tirrell, D. A. (2003) Non-canonical amino acids in protein engineering, *Curr. Opin. Biotechnol. 14*, 603–609.
12. Hecht, S. M., Alford, B. L., Kuroda, Y., and Kitano, S. (1978) "Chemical aminoacylation" of tRNAs, *J. Biol. Chem. 253*, 4517–4520.
13. Forster, A. C., Tan, Z., Nalam, M. N., Lin, H., Qu, H., Cornish, V. W., and Blacklow, S. C. (2003) Programming peptidomimetic syntheses by translating genetic codes designed *de novo*, *Proc. Natl. Acad. Sci. U.S.A. 100*, 6353–6357.
14. Tan, Z., Forster, A. C., Blacklow, S. C., and Cornish, V. W. (2004) Amino acid backbone specificity of the *Escherichia coli* translation machinery, *J. Am. Chem. Soc. 126*, 12752–12753.
15. Tan, Z., Blacklow, S. C., Cornish, V. W., and Forster, A. C. (2005) *De novo* genetic codes and pure translation display, *Methods 36*, 279–290.
16. Frankel, A., Millward, S. W., and Roberts, R. W. (2003) Encodamers: unnatural peptide oligomers encoded in RNA, *Chem. Biol. 10*, 1043–1050.
17. Frankel, A., and Roberts, R. W. (2003) *In vitro* selection for sense codon suppression, *RNA 9*, 780–786.
18. Josephson, K., Hartman, M. C., and Szostak, J. W. (2005) Ribosomal synthesis of unnatural peptides, *J. Am. Chem. Soc. 127*, 11727–11735.

19. Seebeck, F. P., and Szostak, J. W. (2006) Ribosomal synthesis of dehydroalanine-containing peptides, *J. Am. Chem. Soc. 128*, 7150–7151.
20. Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis, *Nat. Methods 3*, 357–359.
21. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components, *Nat. Biotechnol. 19*, 751–755.
22. Murakami, H., Saito, H., and Suga, H. (2003) A versatile tRNA aminoacylation catalyst based on RNA, *Chem. Biol. 10*, 655–662.
23. Murakami, H., Kourouklis, D., and Suga, H. (2003) Using a solid-phase ribozyme aminoacylation system to reprogram the genetic code, *Chem. Biol. 10*, 1077–1084.
24. Brizzard, B. L., Chubet, R. G., and Vizard, D. L. (1994) Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution, *Biotechniques 16*, 730–735.
25. Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (1998) Crystal structure of methionyl-tRNA(f)(Met) transformylase complexed with the initiator formylmethionyl-tRNA(f)(Met), *EMBO J. 17*, 6819–6826.
26. Li, Y., Ramesh, V., Mangroo, D., Taneja, C., and RajBhandary, U. L. (2000) Suppressor mutations in *Escherichia coli* methionyl-tRNA formyltransferase that compensate for the formylation defect of a mutant tRNA aminoacylated with lysine, *Biochemistry 39*, 8039–8046.
27. Grunewald, J., and Marahiel, M. A. (2006) Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides, *Microbiol. Mol. Biol. Rev. 70*, 121–146.
28. Harris, J. I., and Roos, P. (1956) Amino-acid sequence of a melanophore-stimulating peptide, *Nature 178*, 90.
29. Takayasu, S., Sakurai, T., Iwasaki, S., Teranishi, H., Yamanaka, A., Williams, S. C., Iguchi, H., Kawasawa, Y. I., Ikeda, Y., Sakakibara, I., Ohno, K., Ioka, R. X., Murakami, S., Dohmae, N., Xie, J., Suda, T., Motoike, T., Ohuchi, T., Yanagisawa, M., and Sakai, J. (2006) A neuropeptide ligand of the G protein-coupled receptor GPR103 regulates feeding, behavioral arousal, and blood pressure in mice, *Proc. Natl. Acad. Sci. U.S.A. 103*, 7438–7443.
30. Pero, S. C., Oligino, L., Daly, R. J., Soden, A. L., Liu, C., Roller, P. P., Li, P., and Krag, D. N. (2002) Identification of novel non-phosphorylated ligands, which bind selectively to the SH2 domain of Grb7, *J. Biol. Chem. 277*, 11918–11926.
31. Tanaka, S., Pero, S. C., Taguchi, K., Shimada, M., Mori, M., Krag, D. N., and Arii, S. (2006) Specific peptide ligand for Grb7 signal transduction protein and pancreatic cancer metastasis, *J. Natl. Cancer. Inst. 98*, 491–498.
32. Pero, S. C., Shukla, G. S., Cookson, M. M., Flemer, S., Jr., and Krag, D. N. (2007) Combination treatment with Grb7 peptide and Doxorubicin or Trastuzumab (Herceptin) results in cooperative cell growth inhibition in breast cancer cells, *Br. J. Cancer 96*, 1520–1525.
33. Roberts, K. D., Lambert, J. N., Ede, N. J., and Bray, A. M. (1998) Efficient synthesis of thioether-based cyclic peptide libraries, *Tetrahedron Lett. 39*, 8357–8360.
34. Lung, F. D. T., King, C. R., and Roller, P. P. (1999) Development of non-phosphorylated cyclic thioether peptide binding to the Grb2-SH2 domain, *Lett. Pept. Sci. 6*, 45–49.
35. Yu, L., Lai, Y. H., Wade, J. V., and Coutts, S. M. (1998) A simple and efficient method for the syntheses of thioether cyclic peptides, *Tetrahedron Lett. 39*, 6633–6636.
36. Hanes, J., and Pluckthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display, *Proc. Natl. Acad. Sci. U.S.A. 94*, 4937–4942.
37. Nemoto, N., Miyamoto-Sato, E., Husimi, Y., and Yanagawa, H. (1997) *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.
38. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the *in vitro* selection of peptides and proteins, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12297–12302.

acs chemical biology

Ribosomal Synthesis of Peptidase-Resistant Peptides Closed by a Nonreducible Inter-Side-Chain Bond

Yusuke Sako[†,‡], Yuki Goto[†,§], Hiroshi Murakami[†], and Hiroaki Suga[†,‡,§,*]

[†]Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1, Komaba, Meguro, Tokyo, 153-8904, Japan, [‡]Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 7-3-1, Hongo, Bunkyo, Tokyo, 113-8656, Japan, and [§]Department of Advanced Interdisciplinary Studies, Graduate School of Engineering, The University of Tokyo, 4-6-1, Komaba, Meguro, Tokyo, 153-8904, Japan

ABSTRACT Here we report a new enabling technology for the synthesis of peptidase-resistant cyclic peptides by means of genetic code reprogramming involving the flexizyme (a tRNA acylation ribozyme) and PURE (a reconstituted cell-free translation) systems. In this work, we have developed a new nonproteinogenic amino acid bearing a chloroacetyl group in the side chain, which forms a physiologically stable thioether bond by intramolecular reaction with the sulfhydryl group of a Cys residue in the peptide chain upon translation. Significantly, this chemistry takes place spontaneously *in situ* of the translation solution, giving the corresponding cyclic peptides independent of ring sizes. We have used this method to convert human urotensin II, known as a potent vasoconstrictor, to its analogue containing a thioether bond, showing that this new analogue retains biological activity. Moreover, this peptide exhibits remarkable resistance against peptidases under reducing conditions. Thus, this technology offers a new means to accelerate the discovery of therapeutic peptidic drugs.

A large number of naturally occurring peptides are known to exhibit biological activity. Many of these peptides are constrained by a cyclic scaffold (*1–3*), which not only increases their affinity to the biological targets but also improves their physiological stability against peptidases (*4, 5*). Therefore, such a structure would be an important feature for developing therapeutic agents. In nature, cyclic peptides are generally synthesized in two distinct ways. Microbial macrocyclic peptides are synthesized by nonribosomal peptide synthetase (NRPS) comprised of multiple enzymes, among which a carboxy-terminal thioesterase generally catalyzes the cyclization (*5, 6*). NRPS can synthesize cyclic peptides with a variety of cyclization formats, such as N–C-termini amide bond closure and side chain–C-terminus depsi bond closure (*7*). Moreover, those synthesized by NRPS often contain unique nonproteinogenic amino acids (*8*). The combination of these two features of macrocyclic peptides confers their stability against peptidases. A serious drawback of the NRPS system is the difficulty of its engineering; several studies have been reported to generate novel peptides by manipulating NRPS, but it is still laborious to design and synthesize cyclic peptides by this method (*9–11*).

Another way of generating cyclic peptides is mRNA-directed synthesis using the translation apparatus. When a linear peptide containing two or more cysteine (Cys) residues is expressed in accordance with its mRNA template, a disulfide bond(s) is formed by oxidation to yield the corresponding cyclic peptide (*12*). Since any sequence of template DNA can be readily prepared by solid-phase chemistry, its transcription followed by translation produces the desirable cyclic peptide, thus

*Corresponding author, hsuga@rcast.u-tokyo.ac.jp.

Received for review January 13, 2008 and accepted February 13, 2008.

Published online March 14, 2008

10.1021/cb800010p CCC: $40.75

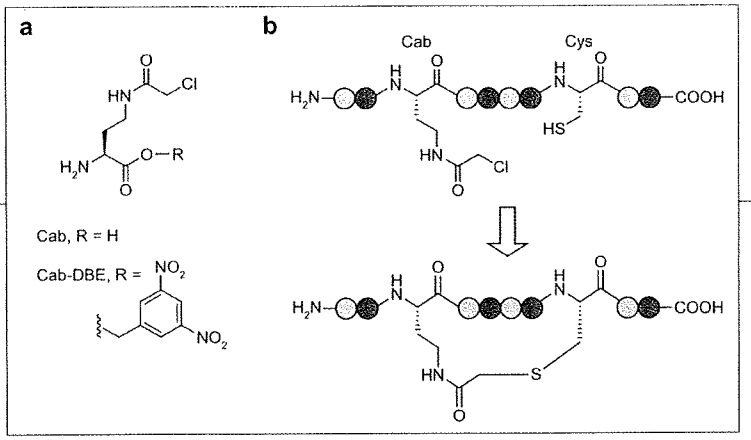


Figure 1. Incorporation of Cab into a peptide chain followed by inter-side-chain cyclization. (a) Chemical structure of N^{γ}-(2-chloroacetyl)-α,γ-diamino-butylic acid (Cab) and 3,5-dinitrobenzyl ester (DBE) form of Cab as a substrate for dFx (dinitro-Flexizyme). (b) Schematic representation of peptide cyclization *via* the spontaneous intramolecular reaction between Cab and Cys side chains.

significantly simplifying the process of designing cyclic peptides. Moreover, when a template DNA library containing fully or partially randomized sequences in the open-reading frame is prepared, a peptide library consisting of billions to trillions of sequences can be synthesized. By means of phage display (*13*), for instance, peptide ligands against a target protein can be screened. Importantly, such a constrained cyclic structure is often crucial to obtain high affinity ligands to the target. A disadvantage of the ribosomal synthesis is that usable building blocks are generally limited to proteinogenic amino acids, and thus the cyclization format is restricted to only the disulfide bond. Unfortunately, the disulfide bond is reducible under physiological conditions, and thereby such a peptide often loses its biological activity upon linearization and also becomes susceptible to peptidases (*14*).

To overcome the above drawbacks of cyclic peptides formed by a disulfide bond, several methods were devised but their application had been limited to only chemical synthesis of such peptides until recently (*15*); however, three research teams independently reported methods applicable to peptides expressed in cell-free translation systems. The first method reported by Szostak and coworkers (*16*) uses α,α′-dibromo-*m*-xylene as a cross-linker to cyclize a peptide containing two Cys residues expressed in a reconstituted *E. coli* cell-free translation (PURE) system (*17*). This methodology originally developed by Meloen *et al.* (*18*) generates two thioether bonds between two Cys residues and the cross-linker agent, inserting five carbons (approximately 8 Å) between the Cys–Cys bond. The second method reported by Roberts and coworkers utilizes disuccinimidyl glutarate as a cross-linker (*19, 20*). In this work, rabbit reticulocyte lysate translation system was used to express desired peptides containing an N-terminal free amine of Met and a single Lys in the nascent peptide chain, and the above cross-linker agent ligates the N-terminal amine and the Lys ϵ-amine. This linkage also inserts five carbons but is significantly longer than the Cys–Cys bond length due to the additional length of the side chain of Lys (by more than 11 Å).

In contrast to the above two methods requiring cross-linker agents for the conversion from a linear peptide to the cyclic form, our group has recently devised a new method that does not require such a post-translational manipulation. In this method, an N^{α}-(2-chloroacetyl)-amino acid (ClAc-aa) charged onto *E. coli* initiator $tRNA^{fMet}_{CAU}$ is used to initiate the translation using a Met-withdrawn PURE system (called *w*PURE system), yielding a peptide containing ClAc-aa at the N-terminus and a single Cys in the nascent chain (*21*). The translated peptide is cyclized by nucleophilic attack of the Cys side chain to the α-carbon of ClAc group at the N-terminus, resulting in the formation of a nonreducible thioether bond. Notably, this "N-terminus-to-side-chain closure" chemistry takes place spontaneously and nearly quantitatively *in situ* in the translation mixture independent of the peptide ring size. This linkage is approximately 3 Å shorter than the parental cyclic peptide formed by Cys–Cys linkage, thus giving a more constrained structure.

Here we expand this methodology to an "inter-side-chain" cyclization using N^{γ}-(2-chloroacetyl)-α,γ-diaminobutylic acid (Cab) as a new family of nonproteinogenic amino acids (Figure 1, panel a). Incorporation of Cab followed by Cys into the nascent peptide chain enables us to spontaneously generate a cyclic peptide closed by a thioether bond between these two side chains (Figure 1, panel b). Distinct from the previous "N-terminus-to-side-chain" linkage, this method leaves N- and C-terminal sequences untouched, so that this linkage is a substitute of the Cys–Cys bond with an insertion of the acetylamide moiety (approximately 5 Å longer than that of Cys–Cys). Using this new method, we have synthesized an analogue of human urotensin II (hU-II), a potent vasoconstrictor peptide (*22, 23*), closed by a thioether bond and demonstrated its biological activity. Remarkably, this analogue exhibits high peptidase resistance, in contrast to the Cys–Cys-linked peptide highly susceptible to proteolysis. Thus, this represents a novel method for generating nonstandard cyclic peptides and provides a powerful means to develop drug-like peptides.

RESULTS AND DISCUSSION

Incorporation of Cab into a Linear Peptide and Its Spontaneous Cyclization. To incorporate Cab into the nascent peptide chain, we used two technologies. The first is flexizyme, a *de novo* ribozyme tRNA aminoacylation catalyst (*24*). This artificial enzyme is able to charge virtually any amino acid assisted by 3,5-dinitrobenzyl ester (DBE) leaving group onto any desired tRNA; a very flexible tool for these type of studies. The second is a so-called PURE system (*17*). Since this *E. coli* cell-free translation system is reconstituted from recombinant protein factors, ribosome, and essential organic and inorganic sources, we can withdraw desired protein factors and/or amino acids from the components. In this work, we prepared a PURE system in which Leu was withdrawn (see Methods), and this PURE system was referred to as the *w*PURE system.

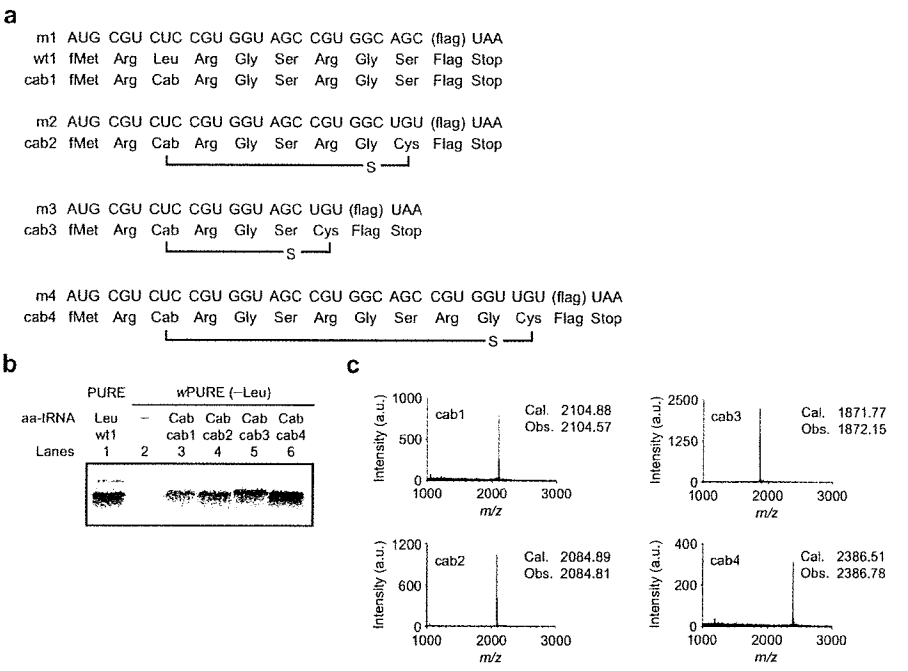


Figure 2. Incorporation of Cab into a peptide chain and its spontaneous cyclization. (a) Sequence of mRNAs (m1–4) and corresponding peptides (wt and cab1–4). Flag in parentheses indicates the RNA sequence encoding the Flag peptide. (b) Tricine SDS-PAGE analysis of the translated products labeled with [^{14}C]-Asp: lane 1, *in vitro* translation with the ordinary PURE system to yield wild type; lane 2, a negative control using *w*PURE system in the absence of aa-tRNA; lanes 3–6, incorporation of Cab into the respective peptides in the presence of Cab-tRNA$^{Asn\text{-}E1}_{GAG}$ prepared by the flexizyme system. In the *w*PURE system, Cab was coded by CUC codon. (c) MALDI-TOF MS spectra of the translated peptides containing Cab (cab1–4). The calculated molecular mass (Cal.) and observed molecular mass (Obs.) for singly charged species, $[M + H]^+$ are shown in the spectra.

We synthesized Cab-DBE and attempted to charge Cab onto a tRNA analogue, microhelix RNA, using flexizyme. This conventional analytical method was reliable to judge the flexizyme aminoacylation ability toward the chosen aminoacyl-DBE by separating the charged RNA from uncharged RNA using denaturing-acid-PAGE. As expected, flexizyme was able to charge Cab onto microhelix RNA with over a 50% yield (Supplementary Figure S1). We then prepared Cab-tRNA$^{Asn\text{-}E1}_{GAG}$ that suppresses Leu codon (CUC) by using flexizyme and Cab-DBE. This particular tRNA was chosen because of its orthogonal property against *E. coli* aaRS (*25*); even if Cab-free tRNA$^{Asn\text{-}E1}_{GAG}$ were generated by deacylation during the translation, it would remain uncharged with proteinogenic amino acids by aaRSs present in the *w*PURE system, thus repressing any undesirable background incorporations.

To evaluate the efficiency of single incorporation of Cab into a nascent peptide chain, we designed an open-reading frame (ORF) in mRNA that expresses a model peptide consisting of fMet-Arg-Leu-Arg-Gly Ser-Arg-Gly Ser followed by a Flag peptide (DYKDDDDK: D, [^{14}C]-Asp or Asp; Y, Tyr; K, Lys) as a wild-type peptide in the PURE system (Figure 2, panel a, wt1 and Figure 2, panel b, lane 1). In the *w*PURE system, the Leu codon (CUC) was suppressed by Cab-tRNA$^{Asn\text{-}E1}_{GAG}$, generating a Cab-containing peptide (Figure 2, panel a, cab1). Tricine SDS-PAGE analysis revealed that the peptide was visible only when Cab-tRNA$^{Asn\text{-}E1}_{GAG}$ was present in the *w*PURE system, giving an approximately 40% yield compared with that of wild type (Figure 2, panel b, lanes 2 and 3). To confirm the incorporation of Cab, cab1 peptide was analyzed by MALDI-TOF mass spectrometry, showing

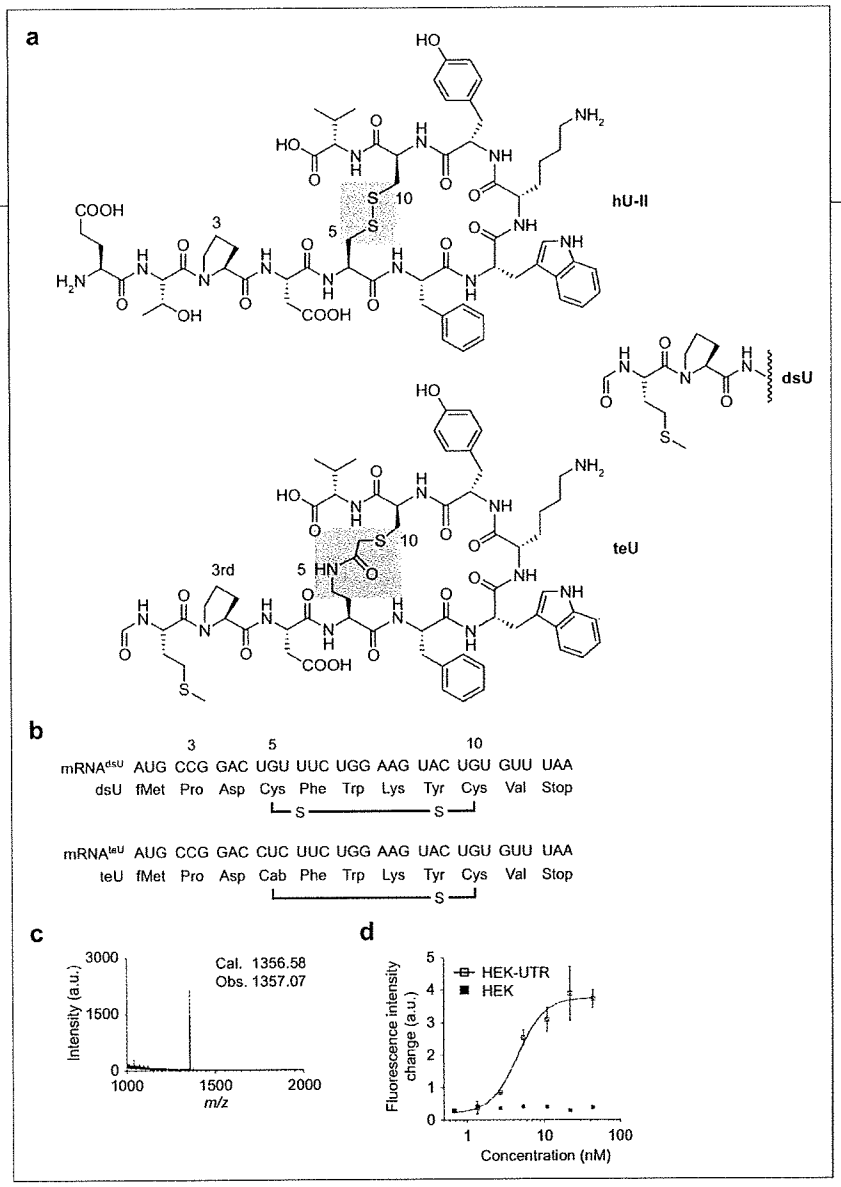


Figure 3. Expression of human urotensin II-like peptides. (a) Molecular structures of human urotensin II (hU-II), dsU, and teU. The inter-side-chain bond in each peptide is highlighted in a gray rectangle. (b) Sequences of mRNA templates encoding human urotensin II-like peptide, dsU (disulfide containing urotensin II-like peptide) and teU (thioether containing urotensin II-like peptide), and the corresponding peptides. (c) MALDI-TOF MS spectrum of teU. (d) Titration of calcium mobilization as a function of teU concentration in HEK-UTR and HEK cells. All experimental points were performed in triplicate.

that its observed molecular mass was consistent with the calculated mass (Figure 2, panel c, cab1). It should be noted that only a single peak corresponding to cab1 was observed, indicating that the 2-chloroacetyl group was intact during the translation, *i.e.* no side reaction occurred with ingredients such as mercaptoethanol and DTT (see supplementary note in the Supporting Information).

Although the 2-chloroacetyl group of Cab was intact in cab1 peptide, we expected that incorporation of a single Cys into the nascent peptide chain would result in selective intramolecular ring closure between Cab and Cys residues. To see whether the expected intramolecular reaction could take place, we designed an mRNA template and expressed a peptide (Figure 2, panel a, m2 and cab2) in which Ser in cab1 peptide was replaced with Cys. Upon the intramolecular cyclization of cab2, it was expected to form a 7-residue sized ring. We also designed two mRNA templates expressing shorter and longer peptides than the cab2 peptide (Figure 2, panel a, m3 and cab3; m4 and cab4), giving 5- and 10-residue ring sizes, respectively. All three peptides were expressed well, giving an expression level comparable to or even better than that of cab1 (Figure 2, panel b, lanes 4–6). MALDI-TOF analysis of the respective peptide showed a clean single product (Figure 2, panel c, cab2–cab4), and the observed molecular mass of the respective peptides was consistent with their expected molecular mass of the cyclic form closed by the thioether bond. These results indicated that the cyclization between Cab and Cys side chains took place spontaneously and nearly quantitatively *in situ* in the translation mixture. Importantly, the cyclization occurred independent of the ring sizes, proving this methodology of its general applicability.

Expression and Biological Activity of Human Urotensin II-Like Peptides. Human urotensin II (hU-II) is a potent vasoconstrictor that binds to a receptor, called the urotensin-receptor (UTR), that belongs to a family of G-protein coupled receptors (*22, 23*). hU-II consists of 11 amino acids and possesses a single disulfide bridge between two Cys residues in positions 5 and 10 (Figure 3, panel a, hU-II). In hU-II, not only endocyclic residues but also exocyclic residues play a critical role in exhibiting the biological activity *in vivo* (*26*). For instance, truncation of either Glu^1–Asp^4 at the N-terminus or Val^{11} at the C-terminus led to greater than 200-fold loss, and both truncations led to nearly completely loss of the vasoconstrictor activity. Considering the importance of the exocyclic residues of hU-II, our previous method "N-terminus-to-side-chain closure" is apparently not appli-

cable to this peptide. However, the "inter-side-chain closure" method developed in this work should be suited for this purpose. Here, the objective of our proof-of-concept study is 2-fold; we (1) demonstrate the synthesis of a hU-II analogue containing the nonreducible thioether bond (Figure 3, panel a, teU) and (2) show the retention of biological activity and its increased proteolytic stability under physiological conditions.

According to previous studies, the N-terminal residues of hU-II, Glu^1, and Thr^2 are not essential for the activity (*26, 27*). We therefore designed a template $mRNA^{dsU}$ expressing a hU-II-like peptide consisting of the minimal activity domain, where the translation would be initiated with fMet, elongated with Pro^3 and the rest of sequence, and eventually closed by the Cys^5-Cys^{10} disulfide bond (Figure 3, panels a and b, dsU). The dsU peptide expressed in PURE system was isolated by a conventional ultrafilteration and desalting purification protocol reported elsewhere (*21*). The MALDI-TOF analysis of dsU gave a clean single peak with the expected molecular mass (see Figure 4, panel b, pK–), suggesting that the peptide isolated by our simple protocol should be sufficient for biological assay. To estimate the peptide concentration, dsU was labeled with [^{14}C]-Asp and analyzed by a liquid scintillation counter calibrated against known [^{14}C]-Asp concentrations. We observed an approximately 1.3 μg/mL (=1.0 μM) expression level (data not shown).

To examine if dsU peptide retained the hU-II-like activity, we performed a calcium mobilization assay. HEK-293 cells stably expressing UTR, referred to as HEK-UTR, were preincubated with a fluorescent calcium indicator dye, Fluo-4 a.m., and then dsU was added to the medium. The fluorescent intensity change was monitored using a microtiter-plate reader. We titrated Ca^{2+} mobilization responses as a function of the dsU concentrations, giving a value of an effective concentration for the half-maximal response (EC_{50}) with 0.47 ± 0.11 nM (Table 1, Supplementary Figure S2a). Under the same conditions HEK-293 did not show any response to dsU even at greater than 2 orders of magnitude higher concentrations (data not shown), indicating that the response is strictly dependent upon the expression of UTR. Importantly, the observed value was consistent with that determined for a chemically synthesized authentic hU-II (EC_{50} = 0.55 ± 0.18 nM, Table 1, Supplementary Figure S2b). These results firmly established that our procedures involving the cell-free expression of the peptide, purification, concentration estimation, and biological assay are reliable to evaluate hU-II-like peptides.

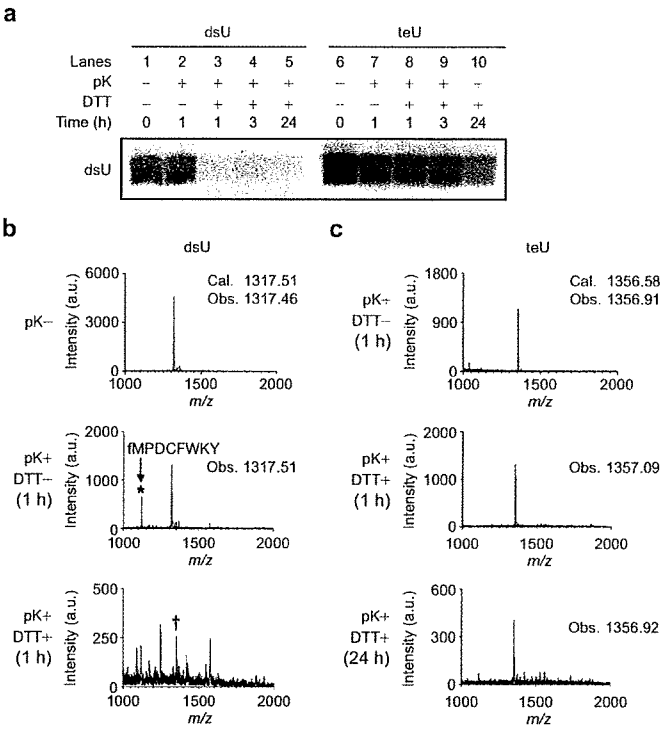


Figure 4. Analysis of peptide stability by proteinase K (pK). (a) Tricine SDS-PAGE analysis of dsU and teU incubated with pK. Peptides were labeled with [^{14}C]-Val and detected by autoradiography. dsU and teU were analyzed without pK (lanes 1 and 6) or with pK (lane 2–5 and 7–10), in the absence of DTT (lanes 2 and 7) or presence of DTT (lane 3–5 and 8–10) for indicated time periods. Note that the intensity of the teU band in lane 10 slightly decreased after 24 h, and this phenomenon was observed repeatedly in triplicate experiments. This could be attributed to a fraction of peptide being stuck to the tube surface and difficult to elute by buffer wash; alternatively, a minor degradation of teU could occur by the pK incubation although such fragments were not observed in the following MALDI-TOF analysis. (b, c) MALDI-TOF MS spectra of dsU and teU incubated with pK. In the absence of pK, both dsU (pK–) and teU (Figure 3, panel c) showed a single peak corresponding to the intact peptide. dsU and teU were then incubated with pK in the absence of DTT or presence of DTT for 1 h (pK+/DTT– 1 h, and pK+/DTT+ 1 h). In the spectrum of dsU without DTT, an additional peak (*) was observed, which corresponds to the hydrolyzed peptide between Tyr^9 and Cys^{10} (Cal. = 1117.45, Obs. = 1117. 45). In the spectrum of dsU with DTT, we could not detect any peak corresponding to the intact peptide (the peak corresponding to † near 1351 Da was a noise peak since the ms did not match to the intact dsU peak). teU was further incubated with pK in the presence of DTT for 24 h (pK+/DTT+ 24 h), showing the intact peptide peak.

We then prepared a template $mRNA^{teU}$ (Figure 3, panel b), where the Leu codon was suppressed with

TABLE 1. Calcium mobilization assay in HEK-UTR cells

Name	Sequence	EC_{50} (nM)[a]
hU-II	⌐S—S¬ ETPDCFWKYCV	0.55 ± 0.18
dsU	⌐S—S¬ fMPDCFWKYCV	0.47 ± 0.11
teU	⌐——S¬ fMPDCabFWKYCV	4.6 ± 0.58

[a] EC_{50} indicates the effective concentration that gives the half-maximal response. The standard deviation was determined by experiments in triplicate.

Cab-$tRNA^{Asn\text{-}E1}_{GAG}$ giving the substitution of Cys^5 in dsU with Cab^5 (Figure 3, panels a and b, teU). Indeed, the translation of this template in *w*PURE system in the presence of Cab-$tRNA^{Asn\text{-}E1}_{GAG}$ expressed teU in a comparable yield to the dsU in a 1.2 μg/mL expression level (data not shown), and MALDI-TOF analysis of the product purified by our protocol showed a clean single product (Figure 3, panel c). Most importantly, the observed molecular mass of the peptide was matched with the expected molecular mass of the cyclic form closed by the thioether bond, indicating that the cyclization between Cab^5 and Cys^{10} side chains took place spontaneously and nearly quantitatively.

Having the cyclic teU in hand, we tested its ability for the induction of Ca^{2+} mobilization using HEK-UTR and HEK-293 cells under the same titration conditions performed for hU-II and dsU. HEK-UTR exhibited a Ca^{2+} mobilization response as a function of concentration of teU, and EC_{50} value was determined to be 4.6 ± 0.58 nM (Figure 3, panel d). In contrast, HEK-293 did not respond to teU at all, indicating that the response was clearly UTR-dependent. The EC_{50} value of teU was, however, approximately 10-fold higher than that for dsU, suggesting that the additional atomic length of the acetylamide group introduced to the linkage somewhat perturbed the original cyclic constrain. Yet it was remarkable that teU was able to induce Ca^{2+} mobilization at a low concentration (nM). Thus, our next question was whether teU possesses improved proteolytic stability.

Peptidase Resistance of teU. Considering the clinical use of peptides, their stability against peptidases (proteases) existing *in vivo* is crucial. Our goal was to show that the nonreducible thioether linkage implanted into the hU-II-like-peptide, teU, conferred peptidase resistance. Before testing this, we first treated three peptides with human plasma (HP) to see how stable the synthetic hU-II and *in vitro* translated dsU and teU were against peptidases. After incubation for 24 h with HP, hU-II was digested into two peptide fragments observed in MALDI-TOF; the cleavage occurred at two sites, between the first and second residues and between the third and forth residues (Supplementary Figure S3a). Interestingly, when dsU was treated with HP for 1 h or even 24 h, it exhibited notable stability for HP digestion (Supplementary Figure S3b). Likewise, teU also exhibited strong resistance against HP (Supplementary Figure S4a). Because the obvious difference of hU-II from dsU/teU was the absence or presence of a formyl group at the N-terminus, we hypothesized that the N-formy group of dsU/teU protected against peptidases in HP. Indeed, when nonformylated forms of deU and teU were prepared by translation, both peptides became susceptible to HP digestion (Supplementary Figures S3c and S4b). We concluded that the N-terminal formylation granted HP resistance.

Proteinase K (pK) is a nonspecific and highly active protease that can cleave various sequences of peptides. We considered that treatment of disulfide-bridged cyclic peptides with pK under reducing conditions would mimic extreme physiological situations. When dsU was treated with pK in the absence of DTT for 1 h, most dsU seemed intact as analyzed by tricine-PAGE radioautoradiography (Figure 4, panel a, lane 2), but MALDI-TOF analysis of this peptide showed a fragment corresponding to the C-terminal digestion at Tyr^9 and Cys^{10} residues (Figure 4, panel b, pK+/DTT–). Upon treatment of dsU with pK in the presence of DTT, the full-length peptide band completely disappeared after 1 h (Figure 4, panel a, lanes 3–5). This was consistent with the MALDI-TOF data where no corresponding full-length peak was observed (Figure 4, panel b, pK+/DTT+). These results clearly imply that the reduction of disulfide bond in dsU significantly increases pK susceptibility.

Substitution of the disulfide bond with the Cab^{5}–Cys^{10} thioether bond made a dramatic change in the pK susceptibility under reducing conditions. When teU was treated with pK, no degradation was detected under not only nonreducing conditions but also reducing conditions for 1 or 3 h (Figure 4, panel a, lanes 7–9). The band corresponding to the full-length peptide could be also observed even after 24 h, showing a remarkable pK-resistance (Figure 4, panel a, lane 10). This observation also well agreed with the observation by their MALDI-TOF analysis where the intact peak was cleanly observed after 1 and 24 h (Figure 4, panel c, pK+/DTT–, and pK+/DTT+ 1 or 24 h). Thus, the thioether bond implanted into teU along with the N-terminal formyl group conferred peptidase-resistance under reduced conditions.

CONCLUSION

We have demonstrated that a new member of nonproteinogenic amino acids containing the 2-chloroacetyl group in its side chain, named Cab, can be efficiently incorporated into a designated site in a peptide assigned by the genetic code reprogramming using the combination of flexizyme and *w*PURE systems. This unique chemical group on the side chain selectively reacts with the sulfhydryl group of Cys in the peptide chain in an intramolecular fashion, giving an inter-side-chain thioether bond. Remarkably, this chemistry takes place spontaneously *in situ* in the translation mixture independent of the ring sizes. We have applied this platform technology to the synthesis of a human urotensin II (hU-II) analogue closed by a nonreducible Cab-Cys thioether bond that replaces the Cys–Cys disulfide bond in hU-II. This substitution in this peptide, called teU, results in a mild reduction of its calcium mobilization activity but yet it exhibits the effective concentration for the half-maximal response (EC_{50}) with less than 5 nM. Most significantly, teU is highly peptidase-resistant against human plasma and proteinase K under reduced conditions in contrast to the parental hU-II (as well as dsU), which is susceptible to proteolysis.

Our data suggest that the cyclic peptide structure closed by the inter-side-chain thioether bond can be a generic scaffold that exhibits peptidase resistance. Also, this strategy is generally applicable to the synthesis of cyclic peptides independent of the ring sizes. These two features grant us a new opportunity to construct a new type of peptide library in the mRNA-encoding format. Particularly, the simplicity and reliability of our cyclization strategy enable us to readily couple this system with an appropriate *in vitro* display system, such as ribosome (*28*) or mRNA display (*29, 30*), to screen peptide aptamers from a large diverse library. Moreover, the expression level observed in our study (>1 μg/mL) also allows us to perform parallel synthesis of a focused cyclic peptide library for cell-based screenings. We expect these approaches to accelerate the discovery of novel peptidic drugs against various therapeutic targets.

METHODS

Translation of Wild-Type and Cab-Containing Peptides. The PURE system was prepared as described previously (*31*), which was reconstituted with ribosome, 20 aaRSs, protein factors, necessary organic and inorganic components including 1 mM DTT (note that an unknown concentration of mercaptoethanol would be also present due to carrying-over from the purification processes of aaRSs and protein factors), and 20 proteinogenic amino acids (0.2 mM each; for tricine SDS-PAGE analysis, 50 μM [^{14}C]-Asp was used in place of Asp), and all wild-type peptides were translated using this PURE system in the presence of a 40 nM DNA template (mDNA^{m1} and mDNA^{dsU}, see Supplementary Table S2) at 37 °C for 1 h at a scale of 5 μL total volume. *w*PURE system was reconstituted with the same components as the PURE system except that only necessary proteinogenic amino acids (0.2 mM each) and Cab-tRNA$^{Asn\text{-}E1}_{GAG}$ (50 μM) were added to the mixture depending upon the peptide kinds stated below: To express cab1 peptide, *w*PURE system in the presence of 40 nM mDNA^{m1}, 0.2 mM each of Met, Arg, Gly, Ser, Asp (this was replaced with 50 μM [^{14}C]-Asp for tricine SDS-PAGE analysis), Tyr, and Lys was used. To express cab2–cab4, the above *w*PURE system plus 0.2 mM Cys was used in the presence of 40 nM of the corresponding DNA template (mDNA^{m2-m4}). To express teU, *w*PURE system containing 0.2 mM each of Met, Pro, Asp, Phe, Trp, Lys, Tyr, Cys, and Val was used in the presence of 40 nM mDNA^{teU}. For tricine SDS-PAGE analysis, 2.5 μL of reaction mixture was taken and mixed with 2.5 μL of loading buffer (0.9 M pH 8.5 Tris-HCl, 30% glycerol, and 8% SDS) and analyzed by 15% tricine SDS-PAGE.

Mass Spectra Measurements of Peptides. For the mass analysis, peptides were purified by two different protocols. For FLAG-containing peptides (cab1–cab4), 5 μL of the reaction mixture was mixed with 5 μL of 2 × TBS (100 mM pH 8.0 Tris-HCl, 300 mM NaCl) and incubated in prewashed FLAG-M2 agarose (Sigma) for 1 h. The resin was then washed with 10 μL of TBS (50 mM pH 8.0 Tris-HCl, 150 mM NaCl) once, and the immobilized peptide was eluted with 2 μL of 0.2% TFA. The resulting peptide was desalted by C18 micro ZipTip (Millipore), and eluted with 1 μL of a 50% acetonitrile, 0.1% TFA solution saturated with the matrix (*R*)-cyano-4-hydroxycinnamic acid. The peptide lacking FLAG (dsU or teU) was purified with Microcon YM-10 centrifugal filter devices (Millipore). After the translation, 5 μL of the reaction mixture was diluted with 45 μL of water and added into the sample reservoir that was prewashed with water. The

reservoir was put into a vial, and centrifuged at 14000*g* for 20 min. The flow-through was collected and acidified with 1% TFA, followed by the ZipTip purification described above. Mass measurements were performed using MALDI-TOF (autoflex TOF/TOF, BRUKER).

Cell Culture and Transfection. HEK293 cells were grown in Minimum Essential Medium Eagle supplemented with 10% FCS, 100 units/mL of penicillin, and 100 μg/mL streptomycin (Sigma-Aldrich) at 37 °C in a 5% CO_2 atmosphere. For the preparation of the cell line which stably expresses the U-II receptor, the cells were transfected with the human U-II-receptor cDNAs (UMR cDNA Resource Center, www.cdna.org) using the Effectene transfection reagent (QIAGEN) according to the manufacturer's protocol. Transfected cells were selected with G418, and a clonal cell line, designated as HEK-UTR, was obtained.

Calcium Mobilization Assay. After the translation, 5 μL of the reaction mixture was diluted with an equal amount of 1% TFA and purified with ZipTip. Peptides were eluted with 4 μL of 50% acetonitrile, 0.1% TFA solution. The eluent was concentrated by MV-100 Micro Vac (TOMY) for 10 min with heat and dissolved in 10 μL of assay buffer (Hanks Balanced Salts Solution (HBSS), 10 mM HEPES, 200 μM of $CaCl_2$, 0.1% BSA, and 2.5 mM probenecid). To determine the peptide concentration, 5 μL of translated peptides labeled with [^{14}C]-Asp were purified by ZipTIp as above, and the eluent was mixed with 10 mL of a liquid scintillation cocktail (Ultima Gold, Perkin-Elmer) and analyzed by a liquid scintillation counter (Wallac 1409, Wallac). The standard curve was created with 1 μL of various concentrations of monomer [^{14}C]-Asp. For the calcium mobilization assay, 2×10^5 HEK-UTR cells were seeded into poly(D-lysine)-coated 96-well black-wall, clear-bottom microtiter plates (Becton Dickinson) 12–16 h before the assay. Cells were loaded with 1 μM of the fluorescent calcium indicator Fluo-4 a.m. (Dojindo) in assay buffer for 1 h at 37 °C, 5% CO_2. After the incubation, cells were washed three times with the assay buffer and further incubated with 90 μL of the assay buffer for 10 min at 37 °C. The plate was then transferred to a microplate reader (Varioskan, Thermo Scientific) and the baseline fluorescence was measured for the first 10 s at a 2-s interval. Then 10 μL of the peptide solution with various concentrations was injected manually with a 10 μL scale microsyringe and the observed fluorescence was recorded for 2 min. The Ca^{2+} response was determined as a subtraction of maximum fluorescence intensity and average fluorescence intensity of the first 10 s.

Proteinase K Protection Assay. For the proteinase K (pK) protection assay, dsU and teU were translated in *w*PURE system containing 50 μM of [^{14}C]-Val, and 0.2 mM each of Met, Pro, Asp, Phe, Trp, Lys, Tyr, and Cys in the presence of 40 nM $mDNA^{dsU}$ or $mDNA^{teU}$ and 50 μM of Cab-$tRNA^{Asn\text{-}E1}_{GAG}$ only for teU synthesis. The respective peptide was purified by ZipTip as described in the calcium mobilization assay section. Eluent was concentrated in vacuo and dissolved in 2.5 μL of pK buffer (10 mM Tri-HCl buffer pH 7.5, 10 mM EDTA, 5 μg/mL pK) in the presence or absence of 1 mM DTT. The peptide was incubated at 37 °C for indicated time periods, and the peptidase digestion was stopped by heating the sample tube at 95 °C for 5 min. Samples were then analyzed by 15% Tricine-SDS PAGE. For MALDI-TOF MS analysis, translated peptide was first purified with Microcon, and flow-through was collected, concentrated and then dissolved in 2.5 μL of pK buffer with or without 1 mM DTT, followed by incubation at 37 °C for indicated time periods. Reaction was stopped by heating the sample tube at 95 °C for 5 min and analyzed by MALDI-TOF MS as described above.

Acknowledgment: We thank Dr. P. C. Reid for critical proofread. This work was supported by grants of Japan Society for the promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) to H.S., Grants-in-Aid for JSPS Fellows to Y.S. (17-11918) and Y.G. (18-10526), and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO).

Supporting Information Available: This material is available free of charge *via* the Internet.

REFERENCES

1. Fliri, H., Baumann, G., Enz, A., Kallen, J., Luyten, M., Mikol, V., Movva, R., Quesniaux, V., Schreier, M., and Walkinshaw, M. (1993) Cyclosporins. structure-activity relationships, *Ann. N.Y. Acad. Sci. 696*, 47–53.
2. Hubbard, B. K., and Walsh, C. T. (2003) Vancomycin assembly: nature's way, *Angew. Chem., Int. Ed. 42*, 730–765.
3. Chatterjee, C., Paul, M., Xie, L., and van der Donk, W. A. (2005) Biosynthesis and mode of action of lantibiotics, *Chem. Rev. 105*, 633–684.
4. Khan, A. R., Parrish, J. C., Fraser, M. E., Smith, W. W., Bartlett, P. A., and James, M. N. (1998) Lowering the entropic barrier for binding conformationally flexible inhibitors to enzymes, *Biochemistry 37*, 16839–16845.
5. March, D. R., Abbenante, G., Bergman, D. A., Brinkworth, R. I., Wickramasinghe, W., Begun, J., Martin, J. L., and Fairlie, D. P. (1996) Substrate-based cyclic peptidomimetics of Phe-Ile-Val that inhibit HIV-1 protease using a novel enzyme-binding mode, *J. Am. Chem. Soc. 118*, 3375–3379.
6. Sieber, S. A., and Marahiel, M. A. (2005) Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics, *Chem. Rev. 105*, 715–738.
7. Schwarzer, D., Finking, R., and Marahiel, M. A. (2003) Nonribosomal peptides: from genes to products, *Nat. Prod. Rep. 20*, 275–287.
8. Marahiel, M. A., Stachelhaus, T., and Mootz, H. D. (1997) Modular peptide synthetases involved in nonribosomal peptide synthesis, *Chem. Rev. 97*, 2651–2674.
9. Hahn, M., and Stachelhaus, T. (2006) Harnessing the potential of communication-mediating domains for the biocombinatorial synthesis of nonribosomal peptides, *Proc. Natl. Acad. Sci. U.S.A. 103*, 275–280.
10. Fischbach, M. A., and Walsh, C. T. (2006) Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms, *Chem. Rev. 106*, 3468–3496.
11. Baltz, R. H. (2006) Molecular engineering approaches to peptide, polyketide and other antibiotics, *Nat. Biotechnol. 24*, 1533–1540.
12. Andreu, D., Albericio, F., Sole, N. A., Munson, M. C., Ferrer, M., and Barany, G. (1994) Formation of disulfide bonds in synthetic peptides and proteins, *Methods Mol. Biol. 35*, 91–169.
13. Dennis, M. S., Eigenbrot, C., Skelton, N. J., Ultsch, M. H., Santell, L., Dwyer, M. A., O'Connell, M. P., and Lazarus, R. A. (2000) Peptide exosite inhibitors of factor VIIa as anticoagulants, *Nature 404*, 465–470.
14. Janecka, A., Zubrzycka, M., and Janecki, T. (2001) Somatostatin analogs, *J. Pept. Res. 58*, 91–107.
15. Li, P., and Roller, P. P. (2002) Cyclization strategies in peptide derived drug design, *Curr. Top. Med. Chem. 2*, 325–341.
16. Seebeck, F. P., and Szostak, J. W. (2006) Ribosomal synthesis of dehydroalanine-containing peptides, *J. Am. Chem. Soc. 128*, 7150–7151.
17. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components, *Nat. Biotechnol. 19*, 751–755.
18. Timmerman, P., Beld, J., Puijk, W. C., and Meloen, R. H. (2005) Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces, *Chembiochem 6*, 821–824.

19. Millward, S. W., Takahashi, T. T., and Roberts, R. W. (2005) A general route for post-translational cyclization of mRNA display libraries, *J. Am. Chem. Soc. 127*, 14142–14143.
20. Millward, S. W., Fiacco, S., Austin, R. J., and Roberts, R. W. (2007) Design of cyclic peptides that bind protein surfaces with antibody-like affinity, *ACS Chem. Biol. 2*, 625–634.
21. Goto, Y., Ohta, A., Sako, Y., Yamagishi, Y., Murakami, H., and Suga, H. (2008) Reprograming the translation initiation for the synthesis of physiologically stable cyclic peptides, *ACS Chem. Biol. 3*, 120–129.
22. Ames, R. S., Sarau, H. M., Chambers, J. K., Willette, R. N., Aiyar, N. V., Romanic, A. M., Louden, C. S., Foley, J. J., Sauermelch, C. F., Coatney, R. W., Ao, Z., Disa, J., Holmes, S. D., Stadel, J. M., Martin, J. D., Liu, W. S., Glover, G. I., Wilson, S., McNulty, D. E., Ellis, C. E., Elshourbagy, N. A., Shabon, U., Trill, J. J., Hay, D. W., Ohlstein, E. H., Bergsma, D. J., and Douglas, S. A. (1999) Human urotensin-II is a potent vasoconstrictor and agonist for the orphan receptor GPR14, *Nature 401*, 282–286.
23. Mori, M., Sugo, T., Abe, M., Shimomura, Y., Kurihara, M., Kitada, C., Kikuchi, K., Shintani, Y., Kurokawa, T., Onda, H., Nishimura, O., and Fujino, M. (1999) Urotensin II is the endogenous ligand of a G-protein-coupled orphan receptor, SENR (GPR14), *Biochem. Biophys. Res. Commun. 265*, 123–129.
24. Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis, *Nat. Methods 3*, 357–359.
25. Ohta, A., Murakami, H., Higashimura, E., and Suga, H. (2007) Synthesis of polyester by means of genetic code reprogramming, *Chem. Biol. 14*, 1315–1322.
26. Brkovic, A., Hattenberger, A., Kostenis, E., Klabunde, T., Flohr, S., Kurz, M., Bourgault, S., and Fournier, A. (2003) Functional and binding characterizations of urotensin II-related peptides in human and rat urotensin II-receptor assay, *J. Pharmacol. Exp. Ther. 306*, 1200–1209.
27. Rossowski, W. J., Cheng, B. L., Taylor, J. E., Datta, R., and Coy, D. H. (2002) Human urotensin II-induced aorta ring contractions are mediated by protein kinase C, tyrosine kinases and Rho-kinase: inhibition by somatostatin receptor antagonists, *Eur. J. Pharmacol. 438*, 159–170.
28. Hanes, J., and Pluckthun, A. (1997) In vitro selection and evolution of functional proteins by using ribosome display, *Proc. Natl. Acad. Sci. U.S.A. 94*, 4937–4942.
29. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the in vitro selection of peptides and proteins, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12297–12302.
30. Nemoto, N., Miyamoto-Sato, E., Husimi, Y., and Yanagawa, H. (1997) 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.
31. Kawakami, T., Murakami, H., and Suga, H. (2008) Messenger RNA-Programmed Incorporation of Multiple N-Methyl-Amino Acids into Linear and Cyclic Peptides, *Chem. Biol. 15*, 32–42.

Available online at www.sciencedirect.com
ScienceDirect
Current Opinion in Chemical Biology
ELSEVIER

Synthesis of biopolymers using genetic code reprogramming

Atsushi Ohta[1], Yusuke Yamagishi[1] and Hiroaki Suga[1,2]

Genetic code reprogramming is a new emerging methodology that enables us to synthesize non-standard peptides containing multiple non-proteinogenic amino acids using translation machinery. This review describes the historical background of this methodology and what distinguishes it from the classical 'nonsense suppression' methodology, followed by a discussion of recent developments in combining this methodology with other compatible technologies. Specifically, we discuss in detail the combination of genetic code reprogramming with flexizymes, *de novo* tRNA acylation ribozymes that facilitate the charging process of a variety of non-proteinogenic amino acids onto tRNAs bearing designated anticodons, and summarize some of the recent demonstrations of the synthesis of non-standard peptides with cyclic structure or/and altered backbones employing this technology.

Addresses
[1] Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 113-8656 Tokyo, Japan
[2] Research Center for Advanced Science and Technology, The University of Tokyo, 153-8904 Tokyo, Japan

Corresponding author: Suga, Hiroaki (hsuga@rcast.u-tokyo.ac.jp)

Current Opinion in Chemical Biology 2008, **12**:159–167

This review comes from a themed issue on
Biocatalysis and Biotransformation
Edited by Stephen G. Withers and Lindsay Eltis

Available online 29th February 2008

1367-5931/$ – see front matter

DOI 10.1016/j.cbpa.2007.12.009

Introduction

It is known that many bacteria, fungi, and other organisms have two distinct sets of machinery for the synthesis of peptides. The first one is the translation system consisting of the ribosome and its associated factors. The ribosome polymerizes amino acids in accordance with the mRNA sequence consisting of a start codon that assigns translation initiation, followed by a stretch of elongation codons, ending with a stop codon that assigns translation termination. These features make this system highly versatile for the synthesis of a variety of peptide sequences by simply designing mRNA (or the corresponding DNA) templates. The only limitation is that each codon is specifically paired with one of the 20 proteinogenic amino acids by the strict recognition of the aminoacyl-tRNA synthetases (aaRSs). Therefore, synthesized peptides can only be made from combinations of the 20 proteinogenic amino acids, unless a special mechanism is employed to incorporate a non-proteinogenic amino acid [1,2].

The second set of machinery for the synthesis of peptides is that of the so-called non-ribosomal peptide synthetases (NRPSs) consisting of clusters of modular protein enzymes. In contrast to the translation system, this system synthesizes peptides in a template-independent manner, that is, the polymerization of amino acids is dictated by each enzyme in the clusters. The NRPS enzymes include epimerases, *N*-methylases, lactonases, and so on; therefore, the final products often contain non-proteinogenic amino acids and/or a variety of cyclic structures [3] (in this review such peptides are referred to as non-standard peptides). Such features give rise to structural rigidity that increases affinity to biological targets and peptidase resistance; thus, they are clearly important for the development of therapeutic peptides. Unlike the translation system, however, production of novel kinds of non-standard peptides using known NRPSs requires swapping and engineering of their modules. Such a task is still technically demanding and only a limited number of successes (i.e. the preparation of new kinds of non-standard peptides) have been reported in the literature [4,5].

In order to realize the full potential of non-standard peptide therapeutics, there exists a significant need for a new platform technology capable of synthesizing a variety of non-standard peptides. Here we summarize a new emerging methodology, called 'genetic code reprogramming' that is able to potentially fulfill such desires and expand the repertoire of the sequences and structures of non-standard peptides.

Genetic code reprogramming distinguished from expansion

Over two decades ago, several groups independently devised a method for the incorporation of non-proteinogenic amino acids into a specific site in a target protein by assignment using the amber codon (UAG) [6,7]. The amber codon is generally recognized as a termination codon by release factor-1 (RF-1), but it is possible to compete it out by the addition of an appropriate suppressor $tRNA_{CUA}$ (the subscript bases indicate its anticodon) pre-charged with non-proteinogenic amino acids, the so-called nonsense suppression [8–17]. The tRNA pre-charging step had long been executed by laborious chemoenzymatic chemistry, but more recently mutants of a certain family of aaRSs have been developed to charge non-proteinogenic amino acids onto the suppressor $tRNA_{CUA}$ [18–20]. The elegance of this method is in the use of a

suppressor $tRNA_{CUA}$ that is inert (orthogonal) to endogenous aaRSs [21]. This orthogonality of $tRNA_{CUA}$ is crucial to avoiding the charging of the (deacylated) $tRNA_{CUA}$ with proteinogenic amino acids, thereby ensuring that proteinogenic amino acids do not compete with the non-proteinogenic amino acid for insertion at the amber codon [22,23].

This 'genetic code expansion' strategy using nonsense suppression was applied to many non-proteinogenic amino acids using the chemoenzymatic and mutant aaRSs methods [14,18,20,21], and the designated amino acids were successfully incorporated into a specific site in the translated protein. The major drawback of this methodology is that since RF-1 competes with aminoacyl-$tRNA_{CUA}$ for the designated amber codon, background termination is unavoidable and sometimes dominates over the suppression, thus giving a very low production of the translated protein. In principle, another stop codon or 4-base codon that programs a frame-shift can be used to assign the second non-proteinogenic amino acid [24]. Unfortunately, only a limited number of successful examples of such an application have been reported [24–27]; presumably, the expression level is significantly diminished as a result of two termination codons.

In 2003, Forster *et al.* reported a landmark experiment where three non-proteinogenic amino acids were assigned to three different codons and incorporated into a peptide in succession by sense suppression [28]. To achieve this, a special translation system lacking proteinogenic amino acids, aaRSs, RFs, and ribosome recycling factor (RRF) was used and supplemented with the above three pre-charged tRNAs carrying non-proteinogenic amino acids for sense suppression and fMet – $tRNA^{fMet}_{CAU}$ for the translation initiation [29]. Although this demonstration was performed under single turnover conditions where the RFs and RRF were not included in the translation system resulting in the loss of ribosome recycling, this significant work opened a new door for the concept of genetic code reprogramming toward the synthesis of non-standard peptides.

This work raised three crucial questions. Since competing proteinogenic amino acids are removed from the translation system (note that their complete depletion is difficult, and this issue is discussed later in this review), the sense suppressions in principle do not compete with the amino acids for the elongation. This would result in achieving high incorporation efficiencies of non-proteinogenic amino acids into the designated sites. Unfortunately, Forster's translation conditions did not allow the ribosome to turn over as mentioned above. Thus, the first question is what increase in expression level is achievable compared with wildtype peptide expression if the multiple turnover conditions is utilized for the expression of non-standard peptides. If satisfactory levels of expression are achievable, the second question is how many codons can be reprogrammed. The last question is whether the genetic code reprogramming allows us to incorporate 'exotic' non-proteinogenic amino acids previously known to be difficult to incorporate into the peptide chain by nonsense suppression.

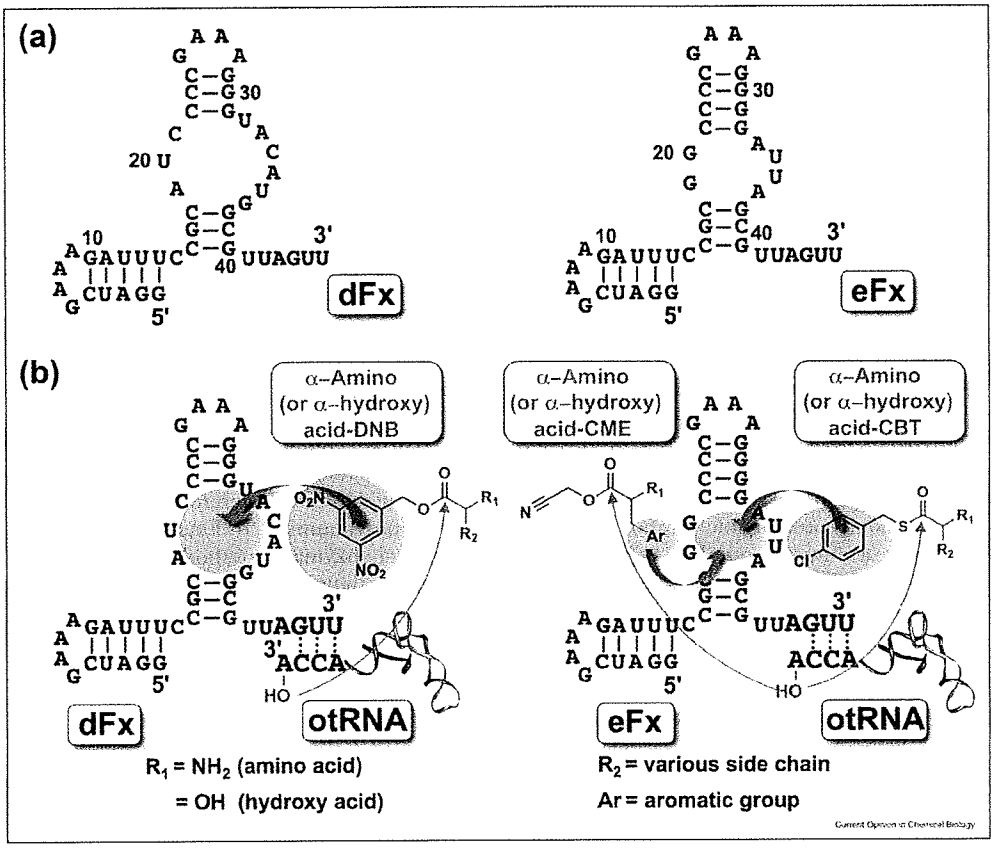
Genetic code reprogramming coupled with PURE system supplemented with non-proteinogenic amino acids

In 2001, Shimizu *et al.* reported a reconstituted *E. coli* cell-free translation, referred to as PURE system (PURE is the abbreviation of Protein synthesis Using Recombinant Elements) [30,31]. This cell-free system contains all of the essential components for translation, and therefore the ribosome can turnover similar to ordinary translation systems. Importantly, this system also allows us to *withdraw* certain components (referred to as *w*PURE system), meaning that any desirable vacant codons can be created for reprogramming; therefore, it is suitable for genetic code reprogramming. In Forster's experiment, the classical chemoenzymatic method was used to prepare tRNAs charged with non-proteinogenic amino acids (abbreviated as ^{n}aa-tRNAs), yet this method is laborious and technically demanding to prepare a wide array of ^{n}aa-tRNAs. Clearly, this process needed to be improved in order to facilitate more extensive studies.

Szostak *et al.* recently found that naturally occurring aaRSs were able to mischarge a variety of non-proteinogenic amino acids onto cognate tRNAs when high concentrations of such amino acids and aaRSs were present [32]. Thus, they prepared a *w*PURE system lacking some proteinogenic amino acids, and to this was added an excess amount of non-proteinogenic amino acids to replace them [33•,34,35••]. It turned out that many of such amino acids could be mischarged onto cognate tRNAs by natural aaRSs under such conditions and incorporated into a single position of a peptide chain. Despite such promiscuous properties of aaRSs, those incorporated into the nascent peptide chain with high efficiency (over 75%) were somewhat proteinogenic-like. Most 'exotic' amino acids disappointingly exhibited less than 25% incorporation efficiency [35••]. It should be noted that such non-proteinogenic amino acids with poor efficiency often suffered from competing incorporations by proteinogenic amino acids, resulting in their contamination into the synthesized peptide. This is because, even though proteinogenic amino acids were removed from the *w*PURE system, a trace amount of contamination from those carried over with ribosome, translation factors, and tRNAs is unavoidable; and thereby there is a risk that those charged onto cognate tRNAs could compete out non-proteinogenic amino acids.

It has to be emphasized, however, that the virtue of this methodology is the simplicity. Addition of non-proteinogenic amino acids with high incorporation efficiency into

Figure 1



Flexizyme catalyzed aminoacylation. **(a)** The secondary structures of dFx and eFx. **(b)** Acylation of tRNA by dFx and eFx. dFx recognizes the 3,5-dinitrobenzylalcohol leaving group (highlighted in green) in α-amino (α-hydroxy) acid–DBE, while eFx recognizes the aromatic side chain (highlighted in green) in α-amino (α-hydoxy) acid–CME (cyanomethyl ester) or 4-chlorobenzylthiol leaving group (highlighted in green) in α-amino (α-hydroxy) acid–CBT. Both flexizymes recognize a 5′-NCC-3′ (*N* = G, A, U) sequence at the 3′-terminus of tRNA and specifically acylate at the 3′-hydroxyl group of tRNA.

the ϖPURE system lets us synthesize non-standard peptides according to the mRNA sequence with the reprogrammed genetic code. In fact, a non-standard peptide containing 13 non-proteinogenic amino acids has been achieved, and its molecular mass was confirmed by MALDI-TOF mass spectrometry [35••]. In terms of the number of non-standard incorporations, this remains the current record high.

Genetic code reprogramming coupled with flexizymes, a *de novo* tRNA acylation tool

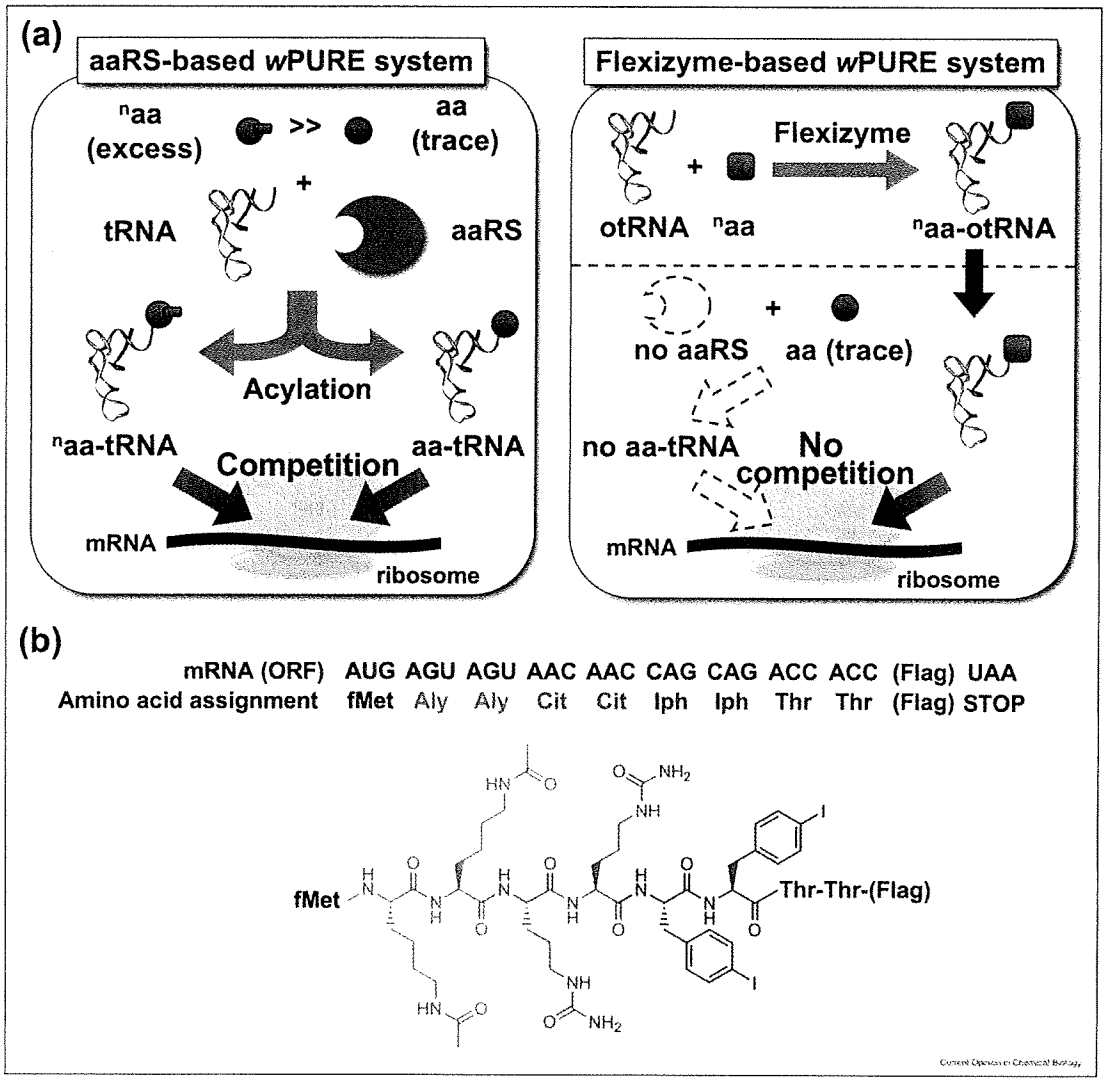
We, Suga and co-workers, took a completely different approach from Szostak's one. We have a long history in the development of ribozymes capable of aminoacylating tRNAs [36–41]. The latest such ribozymes, called dinitro-flexizyme (dFx) and enhanced-flexizyme (eFx), constitute a highly flexible tool for tRNA aminoacylation [42••,43] (Figure 1a). dFx charges amino and hydroxy acids esterified with a 3,5-dinitrobenzyl (DBE) group independent from the type of side chain, while eFx charges those activated with 4-chloro-benzylthioester (CBT) with a variety of side chains and those with cyanomethyl ester (CME) with aromatic side chains. Both flexizymes form three base pairs with the tRNA 3′-end sequence that is most common among tRNAs (Figure 1b). By means of flexizymes, virtually any amino or hydroxy acid can be charged onto any desired tRNAs.

An advantage of the flexizyme technology is the ability to charge a wide variety of non-proteinogenic amino acids onto orthogonal tRNAs (otRNA) with designated anticodons (abbreviated as ^{n}aa-otRNAs). This means that even if such ^{n}aa-otRNAs were deacylated in the translation mixture, they would not be re-charged with the carry-over proteinogenic amino acids by cognate aaRSs (Figure 2a). Upon using the flexizyme method, vacant codons for the reassignment of non-proteinogenic amino acids can be created by removing not only proteinogenic amino acids

but also cognate aaRSs. The combination of these features significantly diminishes the undesirable background of proteinogenic amino acid competing incorporations into the peptide chain.

To demonstrate the compatibility of ^{n}aa-otRNAs prepared by flexizymes, we designed a model peptide sequence containing three non-proteinogenic amino acids in six consecutive positions followed by a FLAG peptide sequence, in which the non-proteinogenic amino acids were assigned to three codons, AGU, ACC, and CAG [42••] (Figure 2b). Appropriate control experiments revealed that the presence of all three ^{n}aa-otRNAs was essential for expression of the full-length peptide, implying that background proteinogenic amino acid incorporation is negligible as expected. Most importantly, the observed molecular mass of the full-length peptide expressed in the presence of the three ^{n}aa-otRNAs was

Figure 2



Genetic code reprogramming. **(a)** Non-standard polypeptide synthesis in the aaRS-based system versus flexizyme-based system. In the aaRS-based *w*PURE system, an unavoidable trace amount of proteinogenic amino acids can be charged onto cognate tRNAs (aa-tRNA), which compete with ^{n}aa-tRNA (^{n}aa, non-proteinogenic amino acid) for the suppression site. In the flexizyme-based *w*PURE system, on the contrary, aaRSs are absent and therefore no competing aa-tRNA is generated. Thus, non-proteinogenic amino acid can be exclusively incorporated into the suppression site. **(b)** Messenger RNA-programmed synthesis of a non-standard peptide. Three ^{n}aa-otRNAs charged with Aly, Cit, and Iph (colored in orange, magenta and blue, respectively) prepared by flexizymes were added to the PURE system in place of Ser, Asn, and Gln in order to reassign three codons, AGU, ACC and CAG, to Aly, Cit, and Iph, respectively.

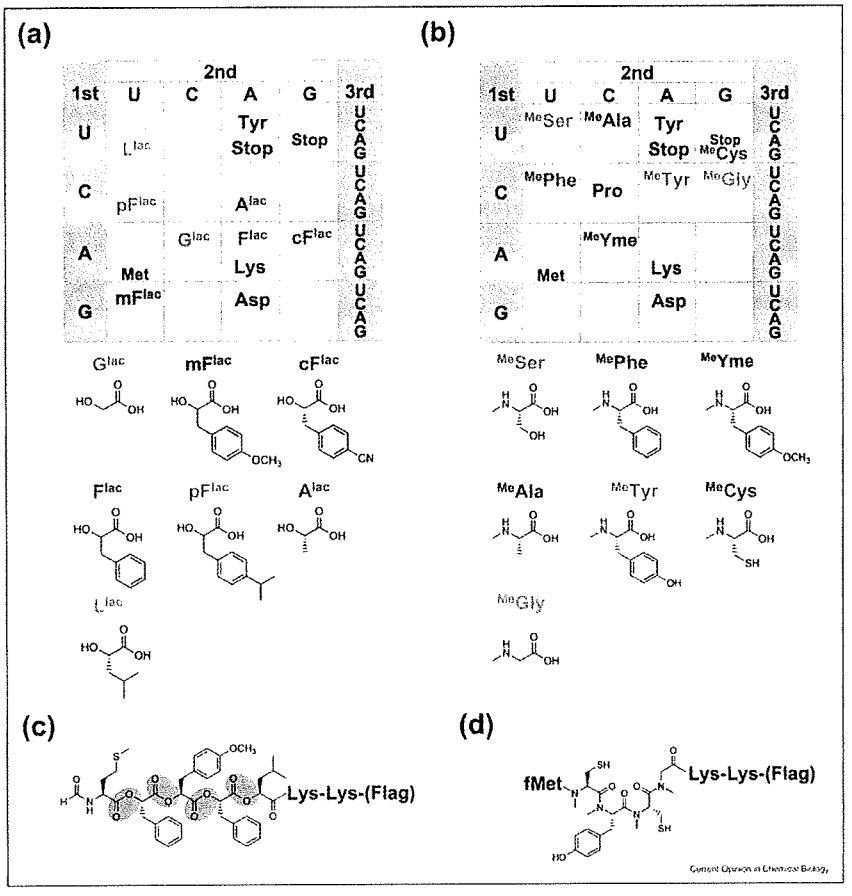
consistent with the expected molecular mass. These results provided concrete evidence of the successful expression of a desirable peptide containing multiple non-proteinogenic amino acids by genetic code reprogramming using the flexizyme method.

The ribosome as a biopolymer synthesizer

The above success in the proof-of-concept experiment encouraged us to extend our method to the synthesis of non-standard peptides and other types of biopolymers. Although under normal translation conditions the ribosome polymerizes only α-amino acids, it has been shown that the ribosome also accepts certain α-hydroxy acids and α-*N*-methyl-amino acids as monomer substrates [8,9,11, 42••,44–48]. For instance, the nonsense suppression method was used to incorporate a designated α-hydroxy acid or α-*N*-methyl-amino acid into a specific site in proteins. Their incorporation yields were dependent upon the kind of monomer structures, ranging from good to poor [8,11]; the phenylalanine counterparts, phenyllactic acid and *N*-methyl-Phe, were exceptionally good monomers for amber suppression (~70% suppression efficiency), whereas other types of α-hydroxy acids and α-*N*-methyl-amino acids were incorporated rather modestly or poorly [8,9,11,42••]. More recently, Cornish and co-workers showed that the sense suppression method increases incorporation efficiency for certain types of α-*N*-methyl-amino acids into a single site of the nascent peptide chain [44,45].

On the contrary, little knowledge was available for the polymerization or multiple incorporations of these monomers. Fahnestock and Rich reported over 30 years ago that phenyllactic acid (F^{lac}) charged onto $tRNA^{Phe}$ was

Figure 3

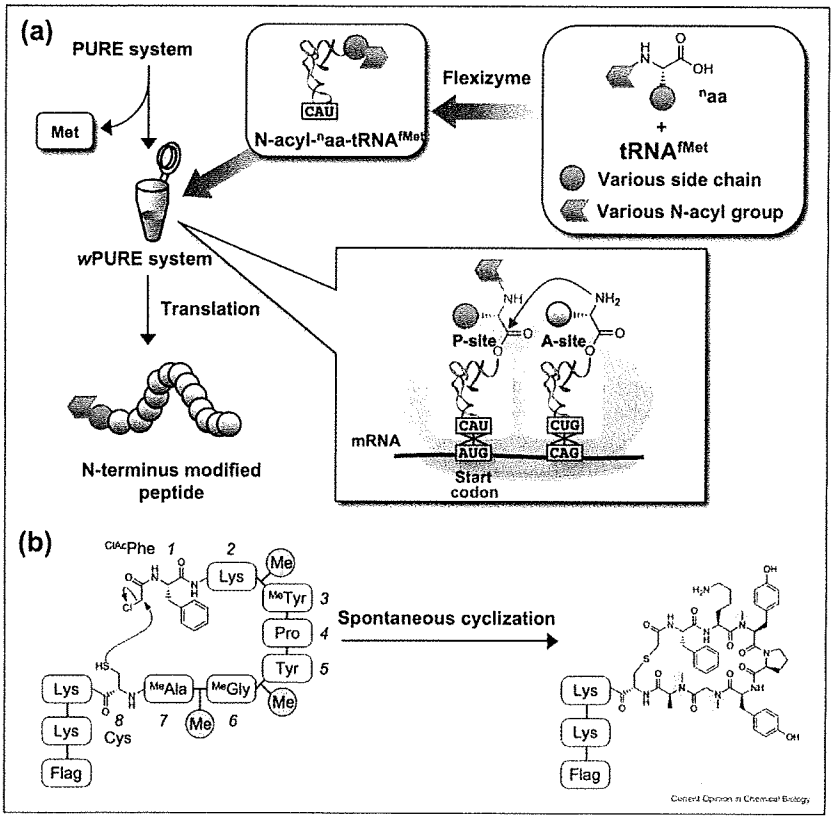


Synthesis of biopolymers containing α-hydroxy acid or α-*N*-methyl-amino acid by means of genetic code reprogramming. **(a)** Reprogrammed genetic code for polyester and chemical structures of α-hydroxy acids. **(b)** Reprogrammed genetic code for *N*-methyl-peptide and chemical structures of α-*N*-methyl-amino acids. **(c)** A representative example for the sequence composition of a polyester-polypeptide hybrid biopolymer. Ester bonds are highlighted in blue. **(d)** A representative example for the sequence composition of *N*-methyl-peptide. The methyl-peptide bonds are highlighted in yellow.

polymerized onto polyuridylic acid via random initiation and termination [49]. Because of the methodology for the preparation of F^{lac}-$tRNA^{Phe}$ where Phe-$tRNA^{Phe}$ was chemically deaminated by nitrous acid, it was difficult to deplete the Phe-$tRNA^{Phe}$ contaminant completely from the translation mixture. Thus the product ended up being not only a heterogeneous random mixture of F^{lac} and Phe but also various lengths of polyesters. Another example more recently reported by Frankel *et al.* was the polymerization of α-*N*-methyl-Phe (^{Me}Phe) charged on $tRNA^{Val}$ by the sense suppression of GUA (Val) codon using rabbit reticulocyte lysate [47]. Despite the fact that ^{Me}Phe could be polymerized according to poly(GUA)$_n$ within the mRNA open-reading frame up to a decamer, the synthesized polymers suffered from contamination of certain proteinogenic amino acids.

We attempted to resolve these problems with our methodology using ϖPURE and flexizyme systems. In this ϖPURE system, we removed certain cognate pairs of amino acids and aaRSs, creating vacant codons as discussed earlier. The flexizyme system enables us to charge a wide variety of α-hydroxy acids and α-*N*-methyl-amino acids onto tRNAs bearing desirable anticodons, and therefore new acid–codon pairs can be created to fill in the above vacant codons. The combination of these two systems readily allowed us to reassign chosen codons to α-hydroxy/α-*N*-methyl-amino acids (Figure 3a and b). Indeed, we succeeded in the mRNA-directed expression of polyesters up to dodecamers bearing three different side chains [50]; one of the examples, a tetramer, being shown in Figure 3c. We also successfully executed the synthesis of poly-*N*-methyl-peptides up to a decamer bearing three or four different side chains (Kawakami *et al.*, *Chem Biol*, in press); one of the examples, a tetramer, being shown in Figure 3d. It should be noted that these isolated full-length peptides or polyesters did not contain undesirable side products, suggesting that the competing background translation was suppressed to a negligible level.

Figure 4



Reprogramming of the translation initiation and its application to the synthesis of cyclic peptide. (**a**) Schematic presentation of the peptide synthesis initiated with an *N*-acyl amino acid. Translation initiation is reprogrammed by Met-depleted *w*PURE system and supplemented with *N*-acyl-^{n}aa-$tRNA^{fMet}$ prepared by flexizyme. Side chain and *N*-acyl group in ^{n}aa are highlighted in pink and orange. (**b**) The synthesis of natural product-like cyclic *N*-methyl peptide. After the translation, linear peptide is spontaneously closed by the reaction occurring between the N-terminal chloroacetyl group and the sulfhydryl group on Cys. *N*-methyl peptide bonds are highlighted in yellow.