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.

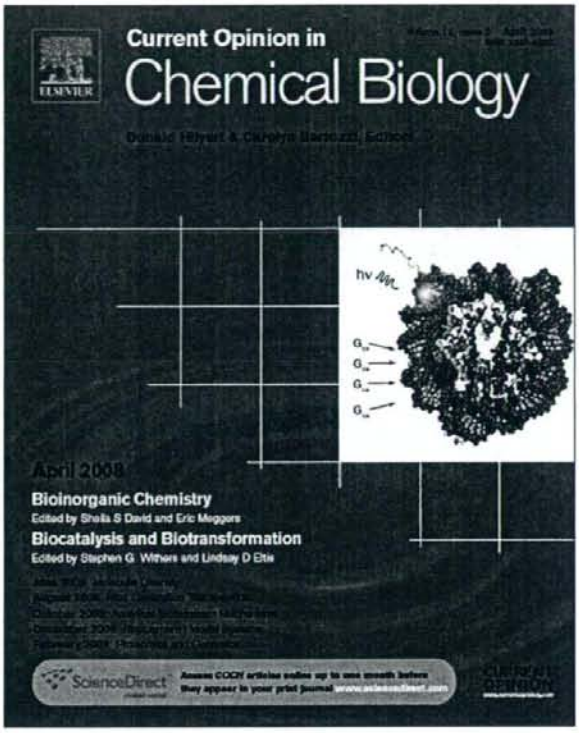
Acknowledgment: We thank Dr. P. C. Reid for critical proofread. This work was supported by grants of Japan Society for the promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) to H.S., Grants-in-Aid for JSPS Fellows to Y.S. (17-11918) and Y.G. (18-10526), and a research and development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

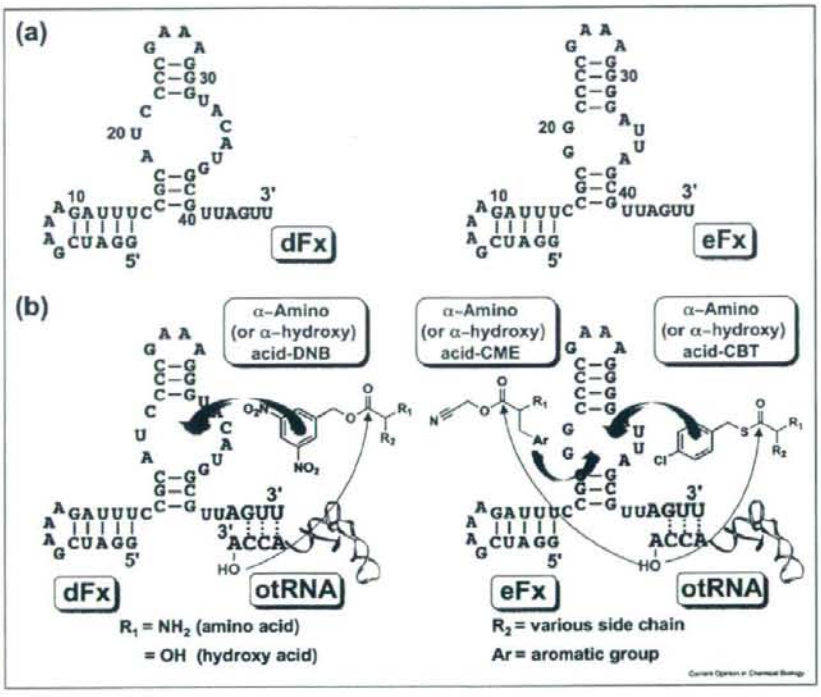
Genetic code reprogramming coupled with PURE system supplemented with non-proteinogenic amino acids

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 ϖPURE 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 ϖPURE 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 ϖPURE 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 ΨPURE 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

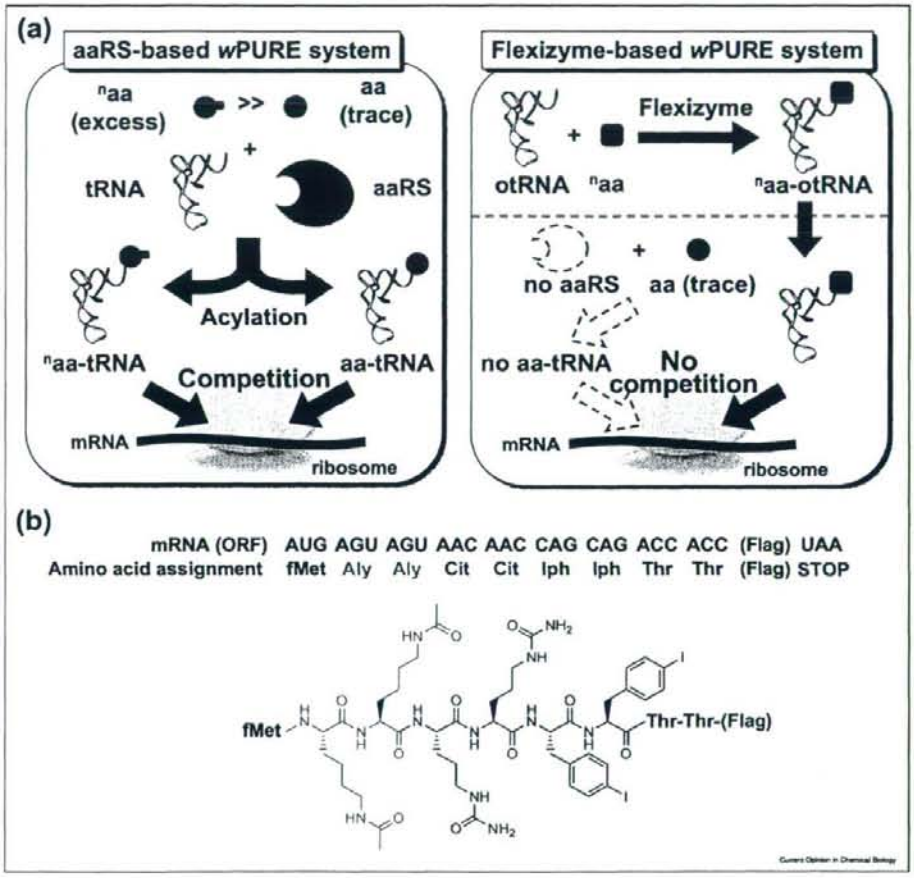
We, Suga and co-workers, took a completely different approach from Szostak's one. We have a long history in the development of ribozymes capable of aminoacylating tRNAs [36–41]. The latest such ribozymes, called dinitro-flexizyme (dFx) and enhanced-flexizyme (eFx), constitute a highly flexible tool for 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 ⁿaa-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 ⁿaa-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 ⁿaa-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 ⁿaa-tRNA (ⁿaa, 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 ⁿaa-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.

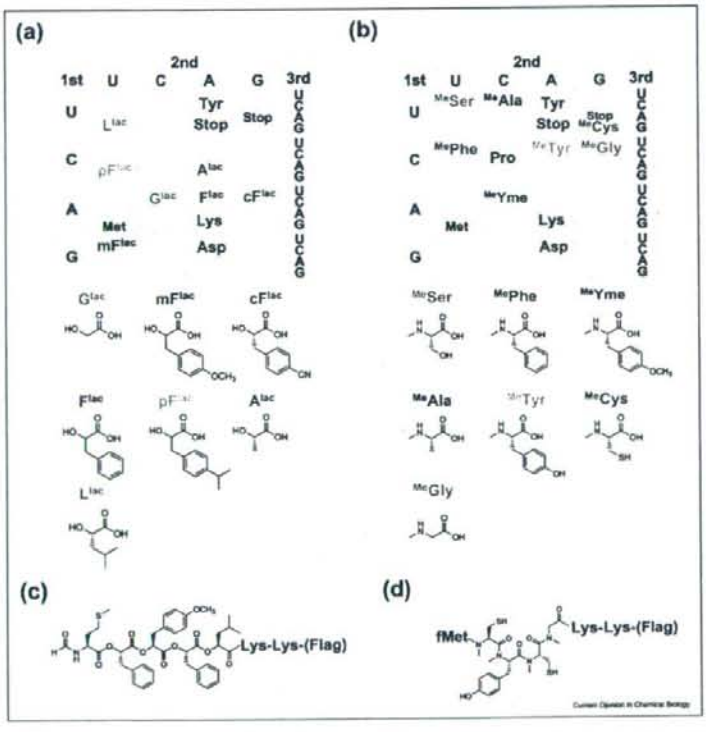
consistent with the expected molecular mass. These results provided concrete evidence of the successful expression of a desirable peptide containing multiple non-proteinogenic amino acids by genetic code reprogramming using the flexizyme method.

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

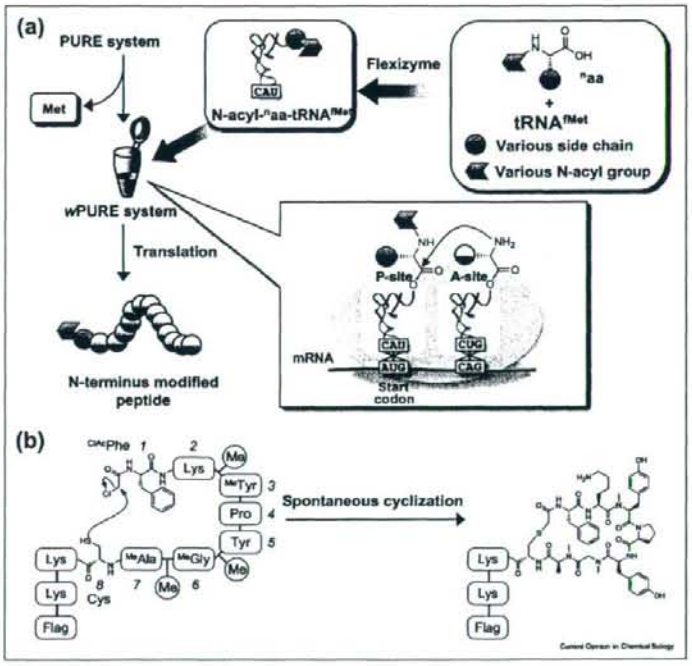


Synthesis of biopolymers containing α-hydroxy acid or α-*N*-methyl-amino acid by means of genetic code reprogramming. (a) Reprogrammed genetic code for polyester and chemical structures of α-hydroxy acids. (b) Reprogrammed genetic code for *N*-methyl-peptide and chemical structures of α-*N*-methyl-amino 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 ɯPURE and flexizyme systems. In this ɯPURE 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 wPURE 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.

Ribosomal synthesis of drug-like non-standard peptides

Natural product peptides often contain multiple *N*-methylations on their backbones, *N*-acylated amino acids or cyclic structures closed by a linkage other than a disulfide bond. These structural features contribute to increases in structural rigidity, proteolytic stability, and membrane permeability. With this in mind, we next applied our methodology to the synthesis of natural product-like non-standard peptides.

To achieve this goal, we recently developed a method to reprogram the translation initiation by depleting Met from ΔPURE system and replacing it with an aminoacyl-$tRNA^{fMet}_{CUA}$ prepared by flexizymes (Figure 4a). Surprisingly, we found that virtually any proteinogenic amino acids could be used as an initiator for translation (Goto *et al.*, *ACS Chem Biol*, in press). Particularly, we were pleased to find that a variety of *N*-acyl groups, such as fatty acids and pyroglutamate, could be added to the initiator amino acids designated by the reprogramming. When one of such acyl groups, 2-chloroacyl group (ClAc), attached to an amino acid was used for the initiation and a single cysteine residue (Cys) was included in the elongating chain, spontaneous cyclization occurred through reaction between the ClAc and sulfhydryl group on Cys, forming a physiologically stable thioether bond (Figure 4b). The combination of this cyclization technology with the incorporation of α-*N*-methyl-amino acids into the nascent peptide chain enabled us to synthesize a wide array of macrocyclic peptides bearing *N*-methylated backbones (Kawakami *et al.*, *Chem Biol*, in press), one such example being shown in Figure 4b.

Conclusions and perspectives

Owing to the fact that the technology is still under development, conclusive answers to the questions raised earlier (vide supra) cannot be given yet; however, at the current stage we are able to offer some insights. Regarding the first question, we generally observe that the expression level of non-standard peptides containing up to four non-proteinogenic amino acids is comparable to that of wildtype expression, generally in the range of 5–10 μM (nmol/mL). However, when the number of non-proteinogenic amino acids is increased, particularly in the use of exotic amino acids, the expression level is reduced to sub-μM range (sub-nmol/mL). Regarding the second question, it is possible to reassign 8–10 codons to non-proteinogenic amino acids by our method, and up to 13 codons by the Szostak's method [35••]. Unfortunately, the expression level of non-standard peptides becomes lower with an increase in their lengths. Although we have witnessed significant improvements in these two areas from recent studies, there remain certain limitations, and it is clear that improvements in the expression level by further engineering of the translation system are required for future technical development [51–54]. Regarding the last question, the incorporation efficiency of non-proteinogenic amino acids such as α-*N*-methyl-amino acids tends to be better for the genetic code reprogramming method than the amber suppression method, probably because the competition that causes the reduction in efficiency is repressed to a minimum [11,44]. However, more thorough investigations of exotic amino acids such as β-, γ- and D-amino acids, which have been thus far largely unsuccessful, are awaited [15,44,35••,42••].

It has become clear now that genetic code reprogramming enables us to synthesize non-standard peptides in a template-dependent manner. This technology is currently being extended to the screening of non-standard peptides against therapeutic targets. Particularly, its coupling with suitable *in vitro* display methods, such as ribosome or mRNA display, would become a powerful screening strategy to accelerate the discovery of peptidic aptamers [55•,56]. We expect that exciting results of this type should be forthcoming from those laboratories pursuing genetic code reprogramming.

Acknowledgements

We thank Dr PC Reid for discussion and proof reading. This work was supported by grants of National Institute of Health (GM59159), Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) and a research and development project of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO) to HS, and AO was supported by Grants-in-Aid for JSPS Fellows (19-1722).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Srinivasan G, James CM, Krzycki JA: **Pyrrolysine encoded by UAG in Archaea: charging of a UAG-decoding specialized tRNA.** *Science* 2002, **296**:1459-1462.
2. Stadtman TC: **Selenocysteine.** *Annu Rev Biochem* 1996, **65**:83-100.
3. Grunewald J, Marahiel MA: **Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides.** *Microbiol Mol Biol Rev* 2006, **70**:121-146.
4. Kohli RM, Walsh CT, Burkart MD: **Biomimetic synthesis and optimization of cyclic peptide antibiotics.** *Nature* 2002, **418**:658-661.
5. Kopp F, Mahlert C, Grunewald J, Marahiel MA: **Peptide macrocyclization: the reductase of the nostocyclopeptide synthetase triggers the self-assembly of a macrocyclic imine.** *J Am Chem Soc* 2006, **128**:16478-16479.
6. Bain JD, Glabe CG, Dix TA, Chamberlin AR: **Biosynthetic site-specific incorporation of a non-natural amino acid *in vitro* a polypeptide.** *J Am Chem Soc* 1989, **111**:8013-8014.
7. Noren CJ, Anthony-Cahill SJ, Griffith MC, Schultz PG: **A general method for site-specific incorporation of unnatural amino acids into proteins.** *Science* 1989, **244**:182-188.
8. Bain JD, Diala ES, Glabe CG, Wacker DA, Lyttle MH, Dix TA, Chamberlin AR: **Site-specific incorporation of nonnatural residues during *in vitro* protein biosynthesis with semisynthetic aminoacyl-tRNAs.** *Biochemistry* 1991, **30**:5411-5421.

9. Chung HH, Benson DR, Schultz PG: **Probing the structure and mechanism of Ras protein with an expanded genetic-code.** *Science* 1993, **259**:806-809.

10. Ellman J, Mendel D, Anthony-Cahill S, Noren CJ, Schultz PG: **Biosynthetic method for introducing unnatural amino acids site-specifically into proteins.** *Methods Enzymol* 1991, **202**:301-336.

11. Ellman JA, Mendel D, Schultz PG: **Site-specific incorporation of novel backbone structures into proteins.** *Science* 1992, **255**:197-200.

12. Hendrickson TL, de Crecy-Lagard V, Schimmel P: **Incorporation of nonnatural amino acids into proteins.** *Annu Rev Biochem* 2004, **73**:147-176.

13. Mendel D, Ellman J, Schultz PG: **Protein-biosynthesis with conformationally restricted amino-acids.** *J Am Chem Soc* 1993, **115**:4359-4360.

14. Wang L, Xie J, Schultz PG: **Expanding the genetic code.** *Annu Rev Biophys Biomol Struct* 2006, **35**:225-249.

15. Sando S, Abe K, Sato N, Shibata T, Mizusawa K, Aoyama Y: **Unexpected preference of the *E. coli* translation system for the ester bond during incorporation of backbone-elongated substrates.** *J Am Chem Soc* 2007, **129**:6180-6186.

16. Heckler TG, Chang LH, Zama Y, Naka T, Chorghade MS, Hecht SM: **T4 RNA ligase mediated preparation of novel 'chemically misacylated' tRNAPheS.** *Biochemistry* 1984, **23**:1468-1473.

17. Heckler TG, Roesser JR, Xu C, Chang PI, Hecht SM: **Ribosomal binding and dipeptide formation by misacylated tRNA(Phe)S.** *Biochemistry* 1988, **27**:7254-7262.

18. Link AJ, Tirrell DA: **Reassignment of sense codons *in vivo*.** *Methods* 2005, **36**:291-298.

19. Wang L, Brock A, Herberich B, Schultz PG: **Expanding the genetic code of *Escherichia coli*.** *Science* 2001, **292**:498-500.

20. Xie JM, Schultz PG: **An expanding genetic code.** *Methods* 2005, **36**:227-238.

21. Liu DR, Magliery TJ, Pasternak M, Schultz PG: **Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins *in vivo*.** *Proc Natl Acad Sci U S A* 1997, **94**:10092-10097.

22. Rodriguez EA, Lester HA, Dougherty DA: **Improved amber and opal suppressor tRNAs for incorporation of unnatural amino acids *in vivo*. Part 1: minimizing misacylation.** *Rna—Publ Rna Soc* 2007, **13**:1703-1714.

23. Rodriguez EA, Lester HA, Dougherty DA: **Improved amber and opal suppressor tRNAs for incorporation of unnatural amino acids *in vivo*. Part 2: evaluating suppression efficiency.** *Rna—Publ Rna Soc* 2007, **13**:1715-1722.

24. Hohsaka T, Ashizuka Y, Murakami H, Sisido M: **Incorporation of nonnatural amino acids into streptavidin through *in vitro* frame-shift suppression.** *J Am Chem Soc* 1996, **118**:9778-9779.

25. Ninomiya K, Minohata T, Nishimura M, Sisido M: ***In situ* chemical aminoacylation with amino acid thioesters linked to a peptide nucleic acid.** *J Am Chem Soc* 2004, **126**:15984-15989.

26. Zhang Z, Alfonta L, Tian F, Bursulaya B, Uryu S, King DS, Schultz PG: **Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells.** *Proc Natl Acad Sci U S A* 2004, **101**:8882-8887.

27. Kohrer C, Yoo JH, Bennett M, Schaack J, RajBhandary UL: **A possible approach to site-specific insertion of two different unnatural amino acids into proteins in mammalian cells via nonsense suppression.** *Chem Biol* 2003, **10**:1095-1102.

28. Forster AC, Tan Z, Nalam MN, Lin H, Qu H, Cornish VW, Blacklow SC: **Programming peptidomimetic syntheses by translating genetic codes designed *de novo*.** *Proc Natl Acad Sci U S A* 2003, **100**:6353-6357.

29. Tan Z, Blacklow SC, Cornish VW, Forster AC: ***De novo* genetic codes and pure translation display.** *Methods* 2005, **36**:279-290.

30. Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, Ueda T: **Cell-free translation reconstituted with purified components.** *Nat Biotechnol* 2001, **19**:751-755.

31. Shimizu Y, Kanamori T, Ueda T: **Protein synthesis by pure translation systems.** *Methods* 2005, **36**:299-304.

32. Hartman MC, Josephson K, Szostak JW: **Enzymatic aminoacylation of tRNA with unnatural amino acids.** *Proc Natl Acad Sci U S A* 2006, **103**:4356-4361.

33. • Josephson K, Hartman MC, Szostak JW: **Ribosomal synthesis of unnatural peptides.** *J Am Chem Soc* 2005, **127**:11727-11735.
Using the PURE system supplemented with non-proteinogenic amino acids, the authors demonstrated mRNA-directed non-standard polypeptide synthesis containing 10 different non-proteinogenic amino acids.

34. Seebeck FP, Szostak JW: **Ribosomal synthesis of dehydroalanine-containing peptides.** *J Am Chem Soc* 2006, **128**:7150-7151.

35. •• Hartman MC, Josephson K, Lin CW, Szostak JW: **An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides.** *PLoS ONE* 2007, **2**:e972.
Ninety-nine non-proteinogenic amino acids were examined for their ability to be added onto tRNAs by aaRS in the wPURE system and incorporated into a specific site. It was shown that non-proteinogenic amino acids with various non-natural sidechains were incorporated, while backbone-modifying ones showed modest incorporation efficiencies.

36. Saito H, Kourouklis D, Suga H: **An *in vitro* evolved precursor tRNA with aminoacylation activity.** *EMBO J* 2001, **20**:1797-1806.

37. Hodgson DR, Suga H: **Mechanistic studies on acyl-transferase ribozymes and beyond.** *Biopolymers* 2004, **73**:130-150.

38. Kourouklis D, Murakami H, Suga H: **Programmable ribozymes for mischarging tRNA with nonnatural amino acids and their applications to translation.** *Methods* 2005, **36**:239-244.

39. Murakami H, Kourouklis D, Suga H: **Using a solid-phase ribozyme aminoacylation system to reprogram the genetic code.** *Chem Biol* 2003, **10**:1077-1084.

40. Murakami H, Saito H, Suga H: **A versatile tRNA aminoacylation catalyst based on RNA.** *Chem Biol* 2003, **10**:655-662.

41. Ramaswamy K, Saito H, Murakami H, Shiba K, Suga H: **Designer ribozymes: programming the tRNA specificity into flexizyme.** *J Am Chem Soc* 2004, **126**:11454-11455.

42. •• Murakami H, Ohta A, Ashigai H, Suga H: **A highly flexible tRNA acylation method for non-natural polypeptide synthesis.** *Nat Methods* 2006, **3**:357-359.
This article reports a highly flexible *de novo* tRNA acylation system consisting of ribozymes (flexizymes, eFx and dFx). This system enables the charging of virtually any non-proteinogenic amino acids and hydroxy acids. This system was also applied to mRNA-programmed non-standard polypeptide synthesis.

43. Ohuchi M, Murakami H, Suga H: **The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus.** *Curr Opin Chem Biol* 2007.

44. Tan Z, Forster AC, Blacklow SC, Cornish VW: **Amino acid backbone specificity of the *Escherichia coli* translation machinery.** *J Am Chem Soc* 2004, **126**:12752-12753.

45. Zhang B, Tan Z, Dickson LG, Nalam MN, Cornish VW, Forster AC: **Specificity of translation for *N*-alkyl amino acids.** *J Am Chem Soc* 2007, **129**:11316-11317.

46. Merryman C, Green R: **Transformation of aminoacyl tRNAs for the *in vitro* selection of 'drug-like' molecules.** *Chem Biol* 2004, **11**:575-582.

47. Frankel A, Millward SW, Roberts RW: **Encodamers: unnatural peptide oligomers encoded in RNA.** *Chem Biol* 2003, **10**:1043-1050.

48. England PM, Zhang Y, Dougherty DA, Lester HA: **Backbone mutations in transmembrane domains of a ligand-gated ion channel: implications for the mechanism of gating.** *Cell* 1999, **96**:89-98.

49. Fahnestock S, Rich A: **Ribosome-catalyzed polyester formation.** *Science* 1971, **173**:340-343.

50. Ohta A, Murakami H, Higashimura E, Suga H: **Synthesis of polyester by means of genetic code reprogramming.** *Chem Biol* 2007, **14**:1315-1322.

51. Doi Y, Ohtsuki T, Shimizu Y, Ueda T, Sisido M: **Elongation factor Tu mutants expand amino acid tolerance of protein biosynthesis system.** *J Am Chem Soc* 2007, **129**:14458-14462.

52. Dedkova LM, Fahmi NE, Golovine SY, Hecht SM: **Enhanced D-amino acid incorporation into protein by modified ribosomes.** *J Am Chem Soc* 2003, **125**:6616-6617.

53. Dedkova LM, Fahmi NE, Golovine SY, Hecht SM: **Construction of modified ribosomes for incorporation of D-amino acids into proteins.** *Biochemistry* 2006, **45**:15541-15551.

54. Wang K, Neumann H, Peak-Chew SY, Chin JW: **Evolved orthogonal ribosomes enhance the efficiency of synthetic genetic code expansion.** *Nat Biotechnol* 2007, **25**:770-777.

55. • Millward SW, Fiacco S, Austin RJ, Roberts RW: **Design of cyclic peptides that bind protein surfaces with antibody-like affinity.** *ACS Chem Biol* 2007, **2**:625-634.
Using a bis-NHS cross-linking reagent to join the N-terminus to an internal lysine residue of peptide, a trillion-member library of cyclic peptides was prepared in the mRNA display format. After *in vitro* selection against signaling protein Gαi1, a high affinity cyclic peptide ligand with K_d = 2.1 nM was identified.

56. Ja WW, West AP Jr, Delker SL, Bjorkman PJ, Benzer S, Roberts RW: **Extension of Drosophila melanogaster life span with a GPCR peptide inhibitor.** *Nat Chem Biol* 2007, **3**: 415-419.

MINIREVIEW / MINISYNTHÈSE

Ribosomal synthesis of nonstandard peptides[1]

Taek Jin Kang and Hiroaki Suga

Abstract: It is well known that standard peptides, which comprise proteinogenic amino acids, can act as specific chemical probes to target proteins with high affinity. Despite this fact, a number of peptide drug leads have been abandoned because of their poor cell permeability and protease instability. On the other hand, nonstandard peptides isolated as natural products often exhibit remarkable pharmaco-behavior and stability in vivo. Although it is likely that numerous nonstandard therapeutic peptides capable of recognizing various targets could have been synthesized, enzymes for nonribosomal peptide syntheses are complex; therefore, it is difficult to engineer such modular enzymes to build nonstandard peptide libraries. Here we describe an emerging technology for the synthesis of nonstandard peptides that employs an integrated system of reconstituted cell-free translation and flexizymes. We summarize the historical background of this technology and discuss its current and future applications to the synthesis of nonstandard peptides and drug discovery.

Key words: nonstandard peptide, misaminoacylation, therapeutic peptide, translation.

Résumé : Il est bien connu que les peptides standards, qui comprennent les acides aminés protéinogènes, peuvent agir comme sondes chimiques spécifiques pour cibler des protéines avec une haute affinité. Malgré cela, plusieurs prototypes de drogues peptidiques ont été abandonnés à cause de leur faible perméabilité cellulaire et à leur sensibilité vis-à-vis les protéases. Cependant, des peptides non standards isolés comme produits naturels font souvent preuve de qualités pharmacologiques uniques et de stabilité *in vivo*. Quoiqu'il soit probable que plusieurs peptides thérapeutiques non standards capables de reconnaître différentes cibles puissent être synthétisés, les enzymes de synthèse de peptides non ribosomiques sont complexes; il est alors difficile de concevoir de telles enzymes modulaires pour construire des banques de peptides non standards. Dans cet article, nous décrivons une technologie en émergence qui permet de synthétiser des peptides non standards à l'aide d'un système intégré de traduction acellulaire reconstituée et de flexizymes. Nous résumons ici la base historique de cette technologie et ses applications actuelles et futures pour la synthèse de peptides non standards et pour la découverte de médicaments.

Mots-clés : peptide non standard, mis-aminoacylation, peptide thérapeutique, traduction.

[Traduit par la Rédaction]

Introduction: standard peptides versus nonstandard peptides

The recent successful development of therapeutic proteins has made a significant impact on the pharmaceutical industry (Leader et al. 2008), yet most drug-development initiatives focus on small organic molecules, because of their target flexibility (intra- and extra-cellular target proteins), stability, and oral bioavailability. Despite such advantages, the binding of small organic molecules to proteins mainly relies on the existence of the narrow and deep binding pockets of their target proteins, where they can fit in. Unfortunately, not all therapeutically relevant target proteins have such a characteristic site. Instead, it is known that protein–protein interactions often occur on a wide and shallow protein interface, of which the standard size spans about 1600 Å^2 (Lo Conte et al. 1999). Additionally, small, deep cavities that can serve as binding sites for small organic molecules are rarely found at the interface of protein–protein interaction pairs (Arkin and Wells 2004). Therefore, the development of small organic molecules capable of disrupting protein–protein interactions may be intrinsically difficult. In this sense, therapeutic antibodies and antibody-like proteins are ideal molecules to disrupt these interactions (Leader et al. 2008); however, they are costly, and their poor membrane permeability restricts their use to extracellular targets (Arkin and Wells 2004). Thus, it is desirable to develop therapeutic agents that are big enough to cover the interface of protein targets but still able to penetrate into cells so that the agents can function against both extracellular and intracellular targets.

Received 22 October 2007. Revision received 25 January 2008. Accepted 28 January 2008. Published on the NRC Research Press Web site at bcb.nrc.ca on 20 March 2008.

T.J. Kang and H. Suga.[2] Research Center for Advanced Science and Technology, University of Tokyo, 153-8904 Tokyo; and Department of Chemistry and Biotechnology, Graduate School of Engineering, University of Tokyo, 113-8656 Tokyo, Japan.

[1]This paper is one of a selection of papers published in this Special Issue, entitled CSBMCB — Systems and Chemical Biology, and has undergone the Journal's usual peer review process.

[2]Corresponding author (e-mail: hsuga@rcast.u-tokyo.ac.jp).

 doi:10.1139/O08-009

In the last few decades, a number of short, naturally occurring peptides possessing a variety of biological activities have been discovered. Potencies of these peptides are remarkably high, often exhibiting activities at low concentrations. For instance, human urotensin-II (hU-II), consisting of 11 amino acids, is a potent vasoconstrictor that strongly binds to one of the G-protein-coupled receptors and induces calcium mobilization at subnanomolar concentrations (Ames et al. 1999; Coulouarn et al. 1998). Moreover, peptide fragments artificially designed from substrates of target proteins or their receptors have been shown to retain their biological activities. For instance, a short peptide derived from the erythropoietin receptor (EPO-R) was able to bind to EPO-R and activate the signaling pathway (Naranda et al. 1999). Furthermore, phage display technology has enabled us to screen artificial peptides from random peptide libraries (Parmley and Smith 1988). Along the same lines, phage display selection against EPO-R has given rise to a peptide sequence that has no sequence homology to EPO yet exhibits activity both in vitro and in vivo (Johnson and Jolliffe 2000; Wrighton et al. 1996). Even though the phage display method has been successfully used to generate a variety of peptide sequences that bind to target proteins, such peptides rarely exhibit high therapeutic potencies in vivo. This is because peptides consisting of the 20 proteinogenic (standard) amino acids are generally susceptible to proteases and are often digested before exhibiting their expected biological activities. Thus, to devise protease-resistant peptides based on the phage-selected peptides, each of the analogs must be chemically synthesized and nonproteinogenic amino acids incorporated, follwed by tedious rescreening against the target to optimize such sequences.

Nature overcomes this limitation by synthesizing nonstandard peptides containing unusual monomers, e.g., N^{α}-methylated amino acids, hydroxy acids, and amino acids with D-configuration or non-proteinogenic side-chains. Whereas standard peptides are synthesized by the mRNA-directed polymerization of amino acids by the translation machinery, nonstandard peptides are generally synthesized by the template-independent synthesis machinery consisting of clusters of modular protein enzymes, called nonribosomal peptide sysnthetases (NRPSs) (Fischbach and Walsh 2006). Remarkably, these nonstandard peptides, which had been isolated as natural products, exhibit a wide range of biological activities not only against microorganisms but also in human cells (Schwarzer et al. 2003). Engineering appropriate modules in NRPSs is expected to generate new machineries capable of synthesizing novel kinds of nonstandard peptides. In a preliminary study, a small library of nonstandard peptides was prepared by NRPSs containing engineered donor and acceptor communication domains (Hahn and Stachelhaus 2006). However, current knowledge about the generality and portability of communication domains is yet insufficient to generate randomly shuffled enzyme domains. Thus, it is still a demanding task to further engineer the clusters that produce a variety of nonstandard peptide libraries for the discovery of novel therapeutic molecules.

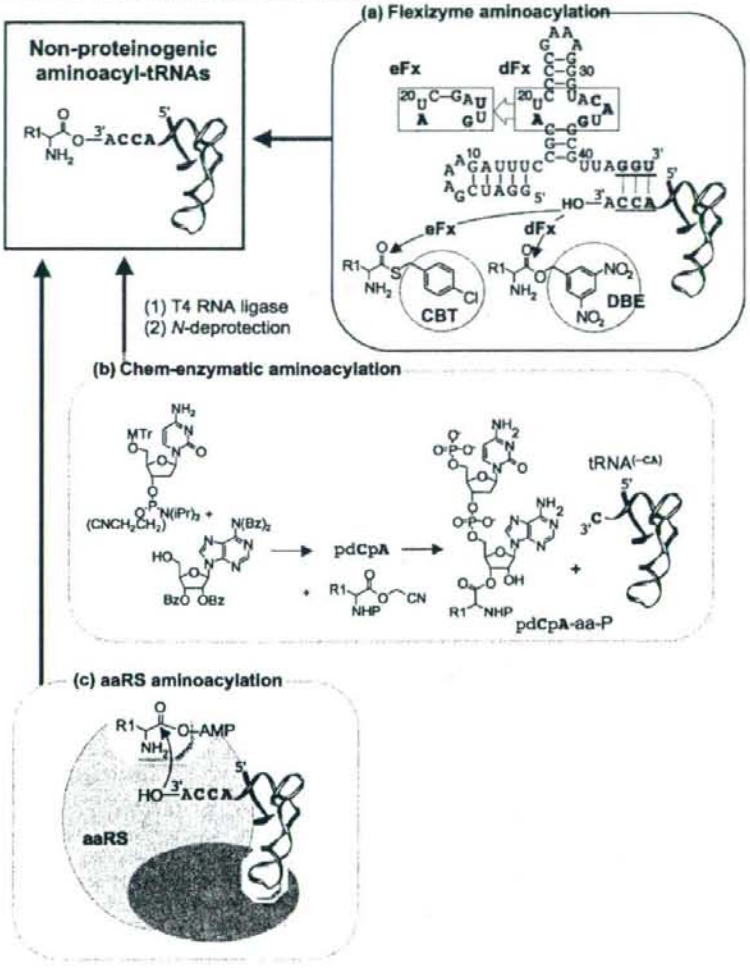
Chemical synthesis has been the only alternative method to generate nonstandard peptide libraries, but it is not necessarily ideal for handling diverse libraries against various protein targets. When the library size reaches one million, it is impractical to test all the individual peptides for their binding properties (Weiss and Chamberlin 2003); thus, diverse libraries are usually subject to various selection strategies to reduce the number of selectants (Turk and Cantley 2003). Additionally, selectants are not usually amplifiable in the chemically synthesized libraries; thus, the absolute amount necessary to identify positive hits restricts practical diversity. Instead, a new technology, termed genetic code reprogramming, has recently been devised for the synthesis of nonstandard peptides in a template-directed format using a reconstituted cell-free translation system. Here, we discuss the development of this emerging technology, as well as its strategies for the synthesis of nonstandard peptides and their libraries and future applications.

Nonsense and 4-base codon suppressions versus genetic code reprogramming

Although the ordinary translation system strictly incorporates 20 proteinogenic amino acids into the nascent peptide chain, some organisms use the same codon for different purposes. For example, UGA and UAG codons occasionally encode selelnocysteine and pyrrolysine, respectively, in a variety of organisms, whereas these codons denote the translation stop signal in general (Zhang and Gladyshev 2007). It is also known that an appropriate manipulation of the translation system enables us to incorporate nonproteinogenic amino acids into peptides. A classical example is that chemically generated misaminoacylated tRNA (alanyl-$tRNA^{Cys}$) supported translation similar to normal enzymatically prepared aminoacyl-tRNAs in vitro (Chapeville et al. 1962; Ehrenstein et al. 1963), indicating that the tRNA aminoacylation event by aminoacyl-tRNA synthetases (aaRSs), and not the ribosome, is the major player controlling translation fidelity. This created the possibility of using the ribosome for making peptides containing nonproteinogenic amino acid(s), and later efforts were devoted to devising methods for assigning nonproteinogenic amino acids to a single codon by using appropriate misacylation technologies.

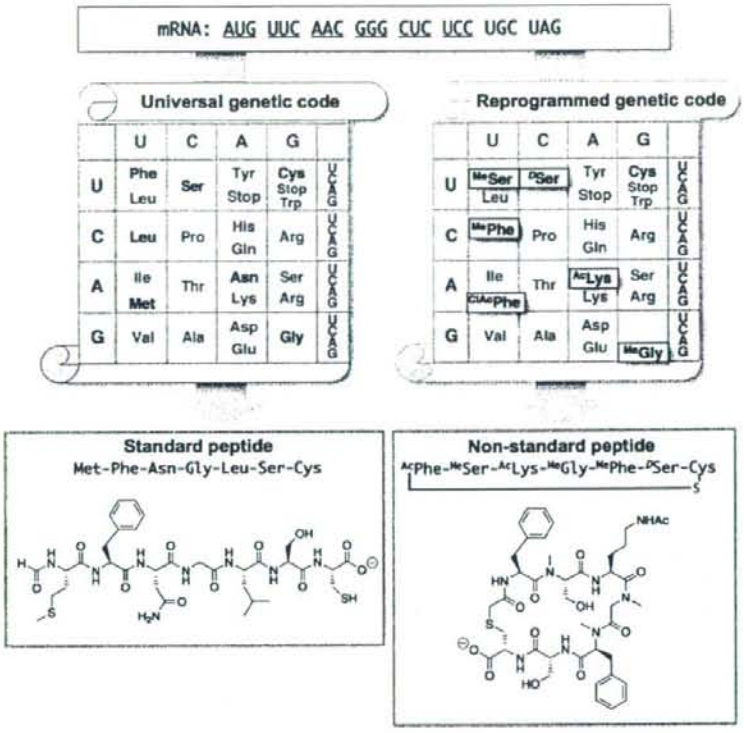
There are 64 possible combinations of triplet nucleotides (codons), all of which are used to code for the 20 amino acids, with the exception of 3 codons, which denote the translation stop signal. These stop codons (or nonsense codons) can be re-assigned to nonproteinogenic amino acid(s) (Noren et al. 1989). However, this method, generally called nonsense suppression, suffers from inherent competition with release factor(s), occasionally yielding low incorporation efficiency, depending upon downstream or upstream codon sets. In addition, this method is suitable only for the incorporation of a single type of nonproteinogenic amino acid into a peptide chain at the specific site assigned by one of the stop codons (usually a UAG stop codon), which is not diverse enough for the construction of a nonstandard peptide library. Complementary to this nonsense suppression, nucleotide quadruplet codons (4-base codons) have been used to assign non-proteinogenic amino acids (Hohsaka et al. 1996), where the programmed frameshift occurs at the 4-base codon as a correct reading frame. This method enabled the incorporation of two, and occasionally three, nonproteinogenic amino acids charged onto

Fig. 1. Preparation of misaminoacyl-tRNAs. (*a*) Flexizyme method (left). Two ribozymes, dFx and eFx, show different substrate specificities, as shown. Flexizyme recognizes only the 3 nucleotides at the 3′-end of tRNA (N_{73}-C_{75}) and the active ester of amino acid substrates, thus this method is compatible with a wide range of acids and tRNAs. (*b*) Chem-enzymatic method (top). The terminal dinucleotide (pdCpA) is chemically synthesized using the standard phosphoramidite method (1); the substrate amino acid (aa) protected by an appropriate protective group is chemically linked to the 3′ end of pdCpA (2); the resulting pdCpA-aa is subsequently ligated to the truncated tRNA (tRNA–3′CA) lacking the 3′-terminal CA by the action of T4 RNA ligase (3): deprotection of the protective group to generate aa-tRNA (4). MTr, methoxytrityl; Bz, benzoyl. (*c*) Enzymatic method using engineered aminoacyl-tRNA synthetase (aaRS, bottom). Both the rational design and the molecular evolution were reported to render the extremely high specificity of some aaRSs. Unlike the previous two methods, this approach enables the large scale syntheses of nonstandard amino acids by expressing engineered aaRS in the host cell. However, because of high specificity, each engineered aaRS must be prepared for each nonstandard amino acid and for each anticodon. So far, the choice of nonstandard amino acids is limited and the anticodon is also limited to the anti-UAG codon.



tRNAs bearing 4-base anticodons into the peptide chain (Ohtsuki et al. 2005). However, because 4-base codons must be designed based on rarely used codons in *Escherichia coli*, the number of usable codons is still restricted, particularly upon using exotic nonproteinogenic acids that are often difficult to incorporate into a peptide chain, and this method suffers from incomplete synthesis of the peptide as a result of undesired reading of the 4-base codon(s) by com-

Fig. 2. Genetic reprogramming approach for the preparation of nonstandard peptides. When the mRNA sequence shown is translated to a peptide, the translation system faithfully follows the universal code book (shown on the left) in the usual in vitro translation system (PURE system). The resulting heptapeptide (fMet-Phe-Asn-Gly-Leu-Ser-Cys) is shown below. In this example, by removing Met, Phe, Asn, Gly, Leu, and Ser (and (or) corresponding aaRSs) from the translation system (wPURE system), codons for those amino acids can be re-assigned to nonstandard amino acids. For this genetic reprogramming, the key component is misaminoacyl-tRNA (^{ClAc}Phe-tRNA$_{ini}$, ^{Me}Ser-tRNA$_{GAA}$, ^{Ac}Lys-tRNA$_{GUU}$, ^{Me}Gly-tRNA$_{CCC}$, ^{Me}Phe-tRNA$_{GAG}$, and ^{D}Ser-tRNA$_{CGA}$, in this example; ^{ClAc}Phe, N^{α}-(2-chloroacetyl)phenylalanine; ^{Ac}Lys, N^{ε}-acetyl-lysine; ^{Me}Gly and ^{Me}Phe, N^{α}-methyl-glycine and N^{α}-methyl-phenylalanine, repectively). The resulting nonstandard peptide is shown below.

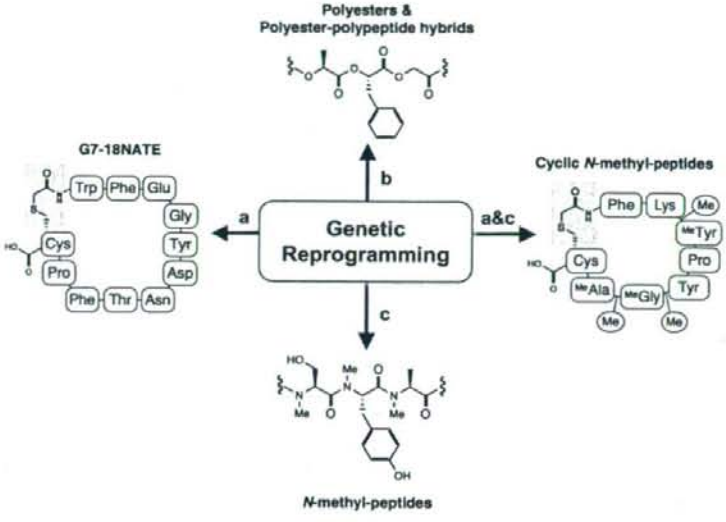


peting aminoacyl-tRNA(s) bearing the normal 3-base anticodon(s).

In addition to the above limitations, the step of tRNA aminoacylation with nonproteinogenic amino acids relies on laborious and technically demanding chem-enzymatic processes (Hecht et al. 1978; Heckler et al. 1984; Noren et al. 1989; Robertson et al. 1989). This technical barrier made this method so specialized that only a subset of researchers was able to use it (Fig. 1). More recently, Schultz and other groups have succeeded in developing mutant aaRSs that charge a certain family of nonproteinogenic amino acids (mostly aromatic amino acid analogs) onto orthogonal tRNA (Kiga et al. 2002; Wang and Schultz 2005; Yoo and Tirrell 2007; Zhang et al. 2004). However, this method is limited to the incorporation of a single nonproteinogenic amino acid, and also the choice of the amino acids remains limited (Fig. 1).

More recently, independent efforts were made to overcome such limitations in conventional nonproteinogenic amino acid mutagenesis by Foster et al., followed by those of several other groups (Forster et al. 2003; Josephson et al. 2005; Murakami et al. 2006). In contrast to the aforementioned methods involving nonsense or 4-base codon suppression, this method involves the reassignment of arbitral codons from proteinogenic to nonproteinogenic amino acids. However, nucleotide triplet codons in the mRNA pass genetic information to the peptide through the tightly regulated aminoacylation of cognate tRNAs, making it very difficult to reassign the genetic code in the usual translation system. A key technology that makes it possible to break the tight regulation of the correspondence is a reconstituted *E. coli* cell-free translation system, often referred to as PURE (protein synthesis using recombinant elements) system (Shimizu et al. 2001). The most important feature of this translation

Fig. 3. Genetic code reprogramming and its applications. Genetic code reprogramming made it possible to incorporate the key amino acids into a peptide for various purposes. (*a*) N^a-(2-chloroacetyl)tyrosine was incorporated into the sequence of G7-18NATE via initiation codon reprogramming. Subsequent cyclization through thioether yielded a cyclized G7-18NATE analog. (*b*) α-Hydroxy acids, lactic acid, phenyllactic acid, and glycolic acid, were incorporated into a peptide structure consecutively, yielding an ester-backbone-containing peptide. (*c*) N^a-Methyl-serine, N^a-methyl-phenylalanine, and N^a-methyl-alanine are incorporated into a peptide. The resulting peptide has a nonstandard backbone that can be found in many therapeutically important natural products. This approach can be combined with the initiation codon reprogramming to ribosomally synthesize cyclized N^a-methylated peptides, yielding a structure that has great potential in peptide therapeutics.



system is that certain amino acids and aaRSs can be withdrawn from the translation elements. For example, when the translation system is deprived of leucine and its corresponding aaRS, leucinyl-tRNA cannot be synthesized in the system, and thus is practically withdrawn, even though corresponding tRNA is present in the system. Using such a withdrawn PURE system, termed wPURE, vacant codons can be created at a researcher's will (Fig. 2). The significant differences between this method and the nonsense or 4-base codon suppression are (*i*) because natural aminoacyl-tRNAs are removed from the translation system, competition between natural and nonstandard aminoacyl-tRNAs would not occur, and (*ii*) through appropriate selection of codons for the nonstandard amino acid, the maximal sense suppression efficiency can be obtained. A major barrier that remains for this method is how nonproteinogenic amino acids can be assigned to these vacant codons, i.e., how can they be attached onto the desired tRNAs capable of reading the vacant codons. Clearly, a nonlaborious and reliable method is necessary to perform this process.

Methods for tRNA aminoacylation, the key technology

The Szostak group showed a pioneering example of codon table redesigning (Josephson et al. 2005). Using nonproteinogenic amino acids compatible with aaRSs, they reassigned the universal genetic code to one that contains 12 nonproteinogenic amino acids, and also showed that this translation system was compatible with the newly assigned genetic code. The translation system resulted in peptides containing nonproteinogenic amino acids, as designated by the mRNA sequence. This technology, together with their proprietary technology, mRNA display, had opened a possibility for in vitro screening of nonstandard peptide aptamers (peptides capable of binding to targets). However, since this system utilizes aaRSs for mischarging tRNAs with nonproteinogenic amino acids, only those that structurally resemble proteinogenic amino acids (and thus, are compatible with aaRSs) can be used, thereby limiting the variety of peptides that can be used. A more serious problem of this approach is contamination of proteinogenic amino acids carried over into the wPURE system with purified ribosome or recombinant factors. When exotic nonproteinogenic amino acids, which are usually inefficiently incorporated into the nascent peptide chain are used for the suppressions, a small amount of such unavoidable proteinogenic contaminants readily outcompete the nonproteinogenic amino acids, leading to peptides composed primarily of these standard amino acids.

Obviously, a general method for the preparation of nonstandard aminoacyl-tRNAs is a prerequisite for genetic reprogramming. Although the chem-enzymatic method is suitable to this application, and indeed, was used in the earlier work by Foster et al. (Forster et al. 2003), its technical difficulties and laboriousness hinder its widespread use in creating peptide libraries with diverse kinds of non-proteino-

Table 1. Yields of acyl-tRNAs.

Acid substrate	Flexizyme	Yield (%)
Standard amino acids		
Ala-DBE	dFx	36[a]
Asn-DBE	dFx	22[b]
Asp-DBE	dFx	52[b]
Cys-DBE	dFx	46[b]
Gln-DBE	dFx	46[b]
Glu-DBE	dFx	17[a]
Gly-DBE	dFx	39[a]
His-DBE	dFx	29[a]
Leu-DBE	dFx	37[a]
Lys-DBE	dFx	36[a]
Met-DBE	dFx	35[a]
Phe-CME	eFx	47[a]
Pro-DBE	dFx	37[b]
Ser-DBE	dFx	38[a]
Trp-CME	eFx	36[a]
Tyr-CME	eFx	34[a]
Val-DBE	dFx	13[a]
Amino acids with nonstandard side-chains		
Aly-DBE	dFx	33[a]
Bly-DBE	dFx	30[a]
Cit-DBE	dFx	35[a]
α-Hydroxy acids		
Hbi-DBE	dFx	25[a]
Hle-DBE	dFx	51[b]
Hph-CME	eFx	82[b]
β-Amino acid and N^{α}-methyl amino acid		
Bal-DBE	dFx	17[a]
Mle-DBE	dFx	55[b]

Note: Yields were calculated using either a streptavidin-dependent gel-shift assay or acid PAGE, depending on the substrate (Murakami et al. 2006). Amino acids are represented in standard three-letter abbreviations except as follows: Aly, ε-*N*-acetyl-lysine; Bly, ε-*N*-biotinyl-lysine; Cit, L-citrulline; Hbi, δ-*N*-biotinyl-(*S*)-hydroxybutanoic acid; Hle, (*S*)-3-isoproylactic acid; Hph, (*S*)-3-phenyllactic acid; Bal, β-alanine; Mle, α-*N*-methyl-leucine. Substrates were activated by either 3,5-dinitrobenzyl (DBE) or cyanomethyl ester (CME).

[a]Yield calculated using a streptavidin-dependent gel-shift assay.

[b]Yield calculated using acid PAGE.

genic amino acids. Instead, an artificial RNA enzyme (ribozyme) capable of catalyzing aminoacylation of tRNAs was generated from a random sequence pool of RNA by our group (Saito et al. 2001), and later, this ribozyme turned into a highly flexible tool for tRNA aminoacylation and was named the flexizyme system (Murakami et al. 2006). The system consists of two kinds of flexizymes, called dFx and eFx, which can be used depending upon the choice of a leaving group on the substrates: dFx for 3,5-dinitobenzyl ester (DBE), and eFx for 4-chlorobenzyl thioester (CBT) or cyanomethyl ester (CME) (Fig. 1). Because the flexizymes recognize their cognate leaving group, and not the side-chain or the free amino group of a given substrate, they are able to charge a wide variety of α-amino acids with nonproteinogenic side-chains. However, the most remarkable feature is that the combination of these two flexizymes gives virtually no restriction of substrates such as N^{α}-methylated amino acids, D-α-amino acids, β-amino acids, and even α-hydroxy acids (Table 1). Moreover, flexizymes recognize N_{73}–C_{75} of tRNA by base pairings with G43–U45, where C_{74} and C_{75} are common to all tRNAs and N_{73} can be any of A, G, and U (even C can be accepted if the incubation time is prolonged); thus virtually any tRNA can be a substrate for flexizymes (Fig. 1).

Thus, we combined the flexizyme system with a wPURE system to demonstrate the genetic code reprogramming for the synthesis of nonstandard peptides. In the first demonstration of the combined systems, three codons (AGU, AAC, and CAG for serine, asparagines, and glutamine, respectively) were reassigned to three nonstandard amino acids (ε-*N*-acetyl-lysine, citrulline, and *p*-iodo-phenylalnine, respectively) so that a 17-mer peptide possessing 6 nonproteinogenic amino acids could be synthesized by the ribosome as efficiently as the original sequence with standard amino acids (Murakami et al. 2006). Typical yields range from several to 20 μmol/L in the translation mixture. With this foundation, our efforts have been made to prepare a wide variety of nonstandard peptides with highly altered chemical structures. We will describe some of our most recent results to exemplify the versatility of this new method.

Ribosomal synthesis of nonstandard polypeptides and polyesters

The SH2 domain of Grb7 plays a role in signaling by binding to an intracellular phosphorylated tyrosine of several receptor tyrosine kinases. Pero et al. performed a selection of peptides against the SH2 domain using phage display and isolated novel high affinity peptides (Pero et al. 2002). G7-18, one of the most potent peptides selected, has the characteristics of a nonphosphorylated and cyclic form, closed by a disulfide bond between two internal cysteine residues. Unfortunately, this disulfide bond in G7-18 was reducible so that its linear form not only failed to exhibit activity but also was susceptible to proteases in vivo. The same team ingeniously substituted the disulfide bond with a thioether bond by chemical synthesis; the N terminus of the cysteine residue of G7-18 was substituted with a 2-chloroacetyl group (ClAc), resulting in the intramolecular attack of the C-terminal cysteine side-chain to the 2-position of the acetyl group. Despite the fact that this new peptide, called G7-18NATE, lost approximately 10-fold of its affinity to Grb7 compared with that of G7-18, it still exhibited antitumor activity in an animal study. We recently succeeded in the ribosomal synthesis of G7-18NATE using initiation codon reprogramming, in which ^{ClAc}Trp was used to prime the translation in a methionine-depleted wPURE system (Goto et al. 2008). Initiator tRNA was charged with ^{ClAc}Trp by the action of flexizyme, and the resulting aminoacyl-tRNA was used for the translation initiation instead of initiator tRNA charged with α-*N*-formyl-methione. Remarkably, the peptide was spontaneously and only intramolecularly cyclized through the thioether bond upon the completion of translation, and therefore no extra treatment was necessary. We envision that our method will allow us to re-investigate the sequence of G7-18NATE by coupling initiation codon reprogramming with an appropriate in vitro display system.