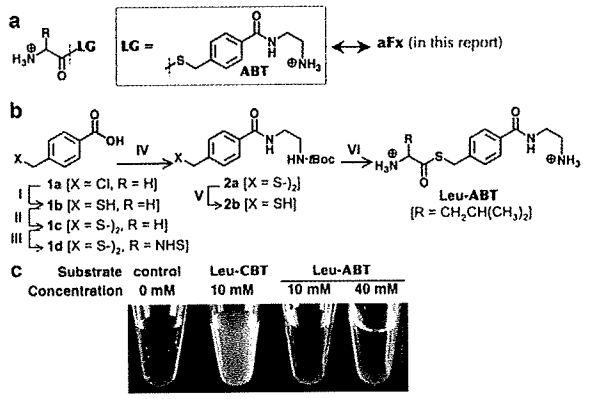


Figure 2. A new substrate design, synthesis, and solubility in the reaction buffer. (a) Chemical structures of acyl-donor substrates. R represents amino acid sidechains including non-proteinogenic ones. Abbreviations are as follows: LG, leaving group; ABT, amino-modified benzyl thioester; aFx, amino flexizyme. (b) Synthesis of aminoacyl-ABT. An example of Leu derivative is shown. Reagents and conditions: (I) thiourea, H_2O, 1 h, reflux, then 10% NaOH, 1 h, reflux, quant.; (II) I_2, EtOH, 0 °C, 1.5 h, quant.; (III) *N*-hydroxysuccinimide (NHS), EDC·HCl, 1,4-dioxane, CH_2Cl_2, rt, 1 h, quant.; (IV) mono-*t*Boc-ethylendiamine, H_2O, 1,4-dioxane, 0 °C, 2 h, 91%; (V) DTT, DMF, 95 °C, 1 h, 50%; (VI) *t*Boc-L-leucine, EDC·HCl, DMAP, CH_2Cl_2, rt, 1 h, then HCl, AcOEt, CH_2Cl_2, rt, 0.5 h, 30%. (c) Solubility of Leu-CBT and Leu-ABT. Each compound was dissolved in 0.1 M Hepes-K (pH 7.5), 20% DMSO at the concentration as shown.

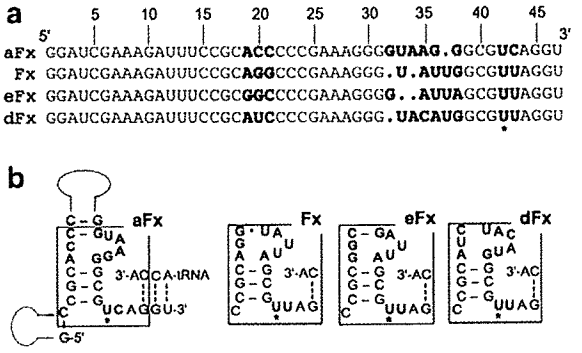


Figure 3. Sequence alignment and secondary structures of aFx and other flexzymes. (a) Sequence alignment. Bases originating from the random regions were shown in bold, and the conserved bass (U42) was shown with asterisk. (b) The secondary structure of aFx compared with the catalytic domain in other flexizymes.

hampers the intrinsic ability of the flexizymes. Therefore we have a dilemma between poor solubility of the substrates and loss of flexizyme activity by the addition of DMSO. Here we report a new pair of flexizyme and substrate activated by a benzyl thioester in which *p*-position is derived to an amino group that is protonated in the reaction buffer to give the ammonium salt. By using this pair, we have demonstrated aminoacylation of tRNA with amino acids bearing hydrophobic sidechains or the α-*N*-octanoyl modification on the sidechain.

In order to make the leaving group more water-soluble, we derived a commercially available 4-chlorobenzoic acid to have a primary amino group and a thiol group for activation of the acyl group (Fig. 2a). Synthesis of this leaving group and its derivatization of amino acid was straightforward as shown in Figure 2b. We referred this amino-derivatized benzyl thioester aminoacyl-donor to as aminoacyl-ABT (Fig. 2a).

First, we compared the solubility of 10 mM Leu-CBT and Leu-ABT in the flexizyme reaction buffer containing 20% (v/v) of DMSO. Despite the solution of Leu-CBT with a cloudy suspension, Leu-ABT gave a completely clear solution. The solution remained clear even with 40 mM Leu-ABT (Fig. 2c). Thus, the ABT leaving group significantly improved the water solubility of the Leu substrate. However, the alteration of the leaving group turned out to be detrimental for the acylation activity of the available flexizymes, dFx and eFx, yielding less than 1% leucylation onto a microhelix RNA (Fig. S1; regarding this analytical method see below). This observation imposed us to perform in vitro selection of a new fleixyzme capable of reacting the ABT-derived substrate.

The method of in vitro selection and the flexizyme pool containing three random regions (Fig. S2a) were the same as our previous report[10] except that the active sequences for self-aminoacylation were enriched by the reaction with Leu-ABT followed by the selective biotinylation on the α-amino group of the charged Leu. After six rounds of selection of active sequences by capturing on a streptavidin-agarose resin (Fig. S2b), we observed an enrichment of the pool. The 3′-terminal specific activity of the pool round 6 was confirmed by the activity loss upon the periodate oxidation (Fig. S2c). 23 clones isolated from the pool were sequenced, and their alignment revealed that 19 clones fell in a single class (Fig. S2d). We picked 5 representative clones and tested for self-aminoacylation; and found that all are active toward Leu-ABT. Among them, clone 11 exhibited the desired 3′-terminal specific activity in the highest efficiency (Fig. S3). Therefore, we focused on this clone for further studies. We referred this new flexizyme to as aFx as oppose to eFx and dFx.

The sequence alignment and drawing the individual secondary structure of aFx, eFx, dFx and the parental Fx revealed the difference in the composition of catalytic domain (Fig. 3a and b). The only conserved base is U42; this base is known to form a unique U-turn structure and plays a critical role in presenting the catalytic domain to the 3′-hydroxyl group at the tRNA 3′-end.[24] Because of lacking our knowledge of the 3-dimensional structure of the catalytic domain in the individual flexizyme except for Fx, it is yet unclear how the sequence variation is able to adapt to the respective leaving group (Fig. 3b). However, even for Fx we have shown that the catalytic domain is susceptible to mutagenesis,[24] it seems that the catalytic domain can be readily adapted to various benzyl leaving groups by the sequence alterations. Nonetheless, the present result indicates that the selection from the doped pool of flexizyme against substrates with different characteristic leaving groups readily generates new flexizyme variants.

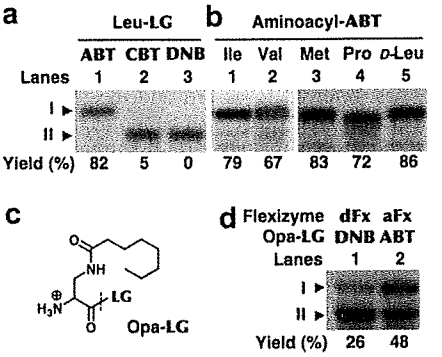


Figure 4. Aminoacylation ability of aFx. (a) Selectivity of aFx to cognate Leu-ABT against non-cognate substrates. Bands of I and II indicate aminoacyl-microhelix RNA and microhelix RNA, respectively. Conditions: 0.1 M Hepes-K (pH 7.5), 50 mM $MgCl_2$, 20% DMSO, 20 μM microhelix RNA, 20 μM aFx, and 5 mM substrate on ice for 2 h; (b) Flexibility of aFx toward sidechain in the aminoacyl-ABT structures. Conditions for Ile-ABT (lane 1): 0.1 M Hepes-K (pH 7.5) and 10 mM substrate on ice for 24 h. Conditions for Val-ABT (lane 2): 0.1 M Bicine-K (pH 9.0) and 10 mM Val-ABT on ice for 6 h. Reactions in lanes 3–5 were carried out by the same conditions described in Figure 4a. (c) Chemical structure of Opa-LG. Opa, 2-amino-3-(octanamido)propanoic acid; LG, leaving group. (d) Aminoaylation of Opa onto a microhelix RNA using aFx or dFx against the cognate Opa-LG. The reaction conditions were the same as those described in Figure 4a, lane 1.

To verify the ability of aFx for trans-activity, we prepared the aFx domain alone by in vitro transcription from the corresponding DNA template, and examined the aminoacylation onto a microhelix RNA using denaturing acid PAGE (poly-acrylamide gel electrophoresis), a previously reported method for the conventional and reliable activity assay for trans-acting flexizymes.[10] The trans-acting aFx was able to charge Leu using Leu-ABT onto the microhelix RNA in 82% yield (Fig. 4a, lane 1). Interestingly, it was virtually inactive toward Leu-CBT and Leu-DNB (Fig. 4a, lanes 2 and 3), indicating that aFx exhibited high specificity to the ABT leaving group.

To investigate the tolerance of aFx toward amino acid sidechains and α-chirality, we synthesized five additional amino acids, Ile, Val, Met, Pro, and D-Leu, activated by ABT (Fig. 4b). The CBT or DNB derivatives of these amino acids were known to have modest or poor solubility at 5 or 10 mM in the reaction buffer unless including above 20% (v/v) DMSO. The ABT derivatives of these amino acids were, however, efficiently charged onto the microhelix RNA in high yields (Fig. 4b). This result shows that aFx is able to charge the ABT-derived amino acids independent from their sidechain structure and α-chirality.

Lastly, we performed aminoacylation using a non-proteinogenic amino acid bearing a fatty acid, 2-amino-3-(octanamido)propanoic acid (Opa, Fig. 4c). We found that eFx was able to charge Opa-CBT onto the microhelix RNA in 26% yield (Fig. 4d, lane 1), whereas aFx was able to charge Opa-ABT in 48% yield (Fig. 4d, lane 2). Clearly, improving the water-solubility of the acyl-donor substrate offered a better yield.

In conclusion, we have devised a new pair of flexizyme and substrate activating group for tRNA aminoacylation, referred to as aFx and ABT. aFx selectively charges ABT-activated amino acids onto tRNA independent from the sidechain kinds. Importantly, the ABT leaving group makes hydrophobic amino acid substrates more water-friendly than CBT and DBE. We now have three pairs of flexizyme and acyl-donor, allowing us to choose the most appropriate pair of flexizyme and acyl-donor depending upon the substrate properties for the genetic code reprogramming.

Acknowledgments

This paper is dedicated to the recognition of the 2009 Tetrahedron Young Investigator Award in Bioorganic and Medicinal Chemistry for Professor Carlos F. Barbas. 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 (7734) to Y.Y., and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO).

Supplementary data

Supplementary data (synthetic procedures, a method of in vitro selection, and analysis for aminoacylation were reported in supplementary data) associated with this article can be found, in the online version, at doi:10.1016/j.bmcl.2009.03.114.

References and notes

1. Hendrickson, T. L.; de Crecy-Lagard, V.; Schimmel, P. *Annu. Rev. Biochem.* **2004**, *73*, 147.
2. Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225.
3. Link, A. J.; Tirrell, D. A. *Methods* **2005**, *36*, 227.
4. Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 6616.
5. Forster, A. C.; Tan, Z.; Nalam, M. N.; Lin, H.; Qu, H.; Cornish, V.; Blacklow, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *1000*, 6353.
6. Merryman, C.; Green, R. *Chem. Biol.* **2004**, *11*, 575.
7. Tan, Z.; Forster, A. C.; Blacklow, S. C.; Cornish, V. W. *J. Am. Chem. Soc.* **2004**, *126*, 12752.
8. Tan, Z.; Blacklow, S. C.; Cornish, V. W.; Forster, A. C. *Methods* **2005**, *36*, 279.
9. Josephson, K.; Hartman, M. C. T.; Szostak, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 11727.
10. Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. *Nat. Methods* **2006**, *3*, 357.
11. Zhang, B.; Tan, Z.; Dickson, L. G.; Nalam, M. N.; Cornish, V. W.; Forster, A. C. *J. Am. Chem. Soc.* **2007**, *129*, 11316.
12. Ohta, A.; Murakami, H.; Higashimura, E.; Suga, H. *Chem. Biol.* **2007**, *14*, 1315.
13. Kawakami, T.; Murakami, H.; Suga, H. *Chem. Biol.* **2008**, *15*, 32.
14. Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 120.
15. Sako, Y.; Goto, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 241.
16. Ohta, A.; Yamagishi, Y.; Suga, H. *Curr. Opin. Chem. Biol.* **2008**, *12*, 159.
17. Kang, T. J.; Suga, H. *Biochem. Cell Biol.* **2008**, *86*, 92.
18. Sako, Y.; Morimoto, J.; Murakami, H.; Suga, H. *J. Am. Chem. Soc.* **2008**, *130*, 7232.
19. Goto, Y.; Murakami, H.; Suga, H. *RNA* **2008**, *14*, 1390.
20. Ohta, A.; Murakami, H.; Suga, H. *ChemBioChem* **2008**, *9*, 2773.
21. Kang, T. J.; Yuzawa, S.; Suga, H. *Chem. Biol.* **2008**, *15*, 1166.
22. Murakami, H.; Saito, H.; Suga, H. *Chem. Biol.* **2003**, *10*, 655.
23. Ohuchi, M.; Murakami, H.; Suga, H. *Curr. Opin. Chem. Biol.* **2007**, *11*, 537.
24. Xiao, H.; Murakami, H.; Suga, H.; Ferre-D'Amare, A. R. *Nature* **2008**, *454*, 358.

Ribosomal synthesis of dehydrobutyrine- and methyllanthionine-containing peptides†

Yuki Goto,[ab] Kazuhiro Iwasaki,[ac] Kohei Torikai,[a] Hiroshi Murakami[a] and Hiroaki Suga*[abc]

Received (in Cambridge, UK) 3rd March 2009, Accepted 31st March 2009
First published as an Advance Article on the web 30th April 2009
DOI: 10.1039/b904314d

We report here the ribosomal synthesis of methyllanthionine-containing cyclic peptides involving a site-specific incorporation of vinylglycine under the reprogrammed genetic code, followed by the isomerization of the vinylglycine to dehydrobutyrine, and the subsequent intramolecular Michael addition of a cysteine residue placed at a downstream position of the vinylglycine.

Dehydrobutyrine (Dhb), one of the family of dehydroamino acids, is frequently found in naturally occurring peptides.[1] For instance, lantibiotics, a class of bacteriocin, have Dhb residues that are generated by enzymatic dehydration of a threonine residue in the precursor peptides expressed by the translation machinery.[2] In addition, lantibiotics also contain thioether-bridged cyclic residues, referred to as methyllanthionine (MeLn). These unique residues are generated from Dhb residues by nucleophilic attack of the sulfhydryl group in cysteine (Cys) residues *via* enzymatic intramolecular Michael addition, forming cyclic structures.[2] The thioether bonds formed at MeLn residues are non-reducible, thereby providing physiological stability of their active structures. This suggests that MeLn-bridged cyclic structures provide excellent peptidic frameworks for the discovery of new drug candidates.[3] However, since the formation of Dhb and MeLn residues from the precursor peptides relies on the actions of specific enzymes,[2] the biosynthetic method would not be synthetically versatile enough to generate diverse kinds of such peptides. On the other hand, direct incorporation of Dhb into the nascent peptide chain is very likely intractable due to the hydrolytic instability of the 2-amino group on Dhb itself (decomposing to the 2-oxoacyl group) and its inefficient elongation by the translation machinery. Hence, a generic and simple method that facilitates the synthesis of Dhb- and MeLn-containing peptides, particularly which enables the construction of a library is needed.

We report here a novel methodology for the ribosomal synthesis of Dhb-containing peptides and MeLn-containing cyclic peptides without enzymatic assistance by expressing the precursor peptides bearing vinylglycine (Vgl) residue. We conceived that Vgl could be useful as a key non-proteinogenic amino acid based on the following considerations: (1) Vgl could be site-specifically incorporated into the desired peptide sequence by means of the genetic code reprogramming,[4] and Cys could also be incorporated into a downstream position of the Vgl residue in the peptide chain. (2) Heating of the Vgl-containing peptide at nearly neutral pH could induce isomerization from Vgl to Dhb, and simultaneously, the resulting Dhb residue could accept the intramolecular Michael addition by Cys side-chains. Our report herein includes the ribosomal synthesis of MeLn-containing cyclic peptides that mimic the B- and C-ring segments of nisin,[2c] one of the members of lantibiotics.

Although it was known that treatment of *N*-benzoyl-Vgl esters under acidic or basic conditions in non-aqueous media, resulted in isomerization from Vgl to Dhb residue,[5] the chemical behaviour of Vgl at neutral or near-neutral pH in aqueous media was unknown in the literature. Therefore, at the first stage of this series of experiments, we synthesized two esters, *N*-acetyl-Vgl methyl ester and Vgl 3,5-dinitrobenzyl ester (**1** and **2** in Fig. 1A and B, respectively), and their chemical behaviours in D_2O buffered at pH 7.4 were monitored by ^{1}H-NMR. It should be noted that **1** was a mimic of Vgl in a peptidic structure[6] and **2** was a substrate for our acylation RNA catalyst (flexible tRNA acylation ribozyme; flexizyme)[4c] that would be used in the subsequent experiment. The summary of the observations was as follows (see more detailed observations and discussions in the ESI†): (1) The α-proton of **1** and **2** was deuterated at 37 °C over 2.5 h to yield **3a** and **4**, respectively, but no isomerization to Dhb was observed. (2) Upon heating **1** at 95 °C for 0.5 h, the isomerization of **1** occurred to yield **5** with a single deuteration on the γ-methyl group, along with a byproduct **3b** generated through the ester hydrolysis. The characteristic triplet at δ 6.83 ppm assigned to the β-olefinic proton of Dhb has a similar chemical shift to that of (*Z*)-Dhb found in lantibiotics.[7] Based on the above results, we predicted that **2** would

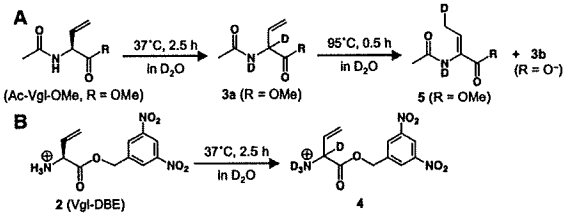


Fig. 1 Model reactions for the isomerization of Vgl to Dhb. Each reaction was monitored by ^{1}H-NMR. See ESI for more details.†

[a] *Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1, Komaba, Meguro, Tokyo 153-8904, Japan. E-mail: hsuga@rcast.u-tokyo.ac.jp; Fax: +81 3 5452 5495; Tel: +81 3 5452 5495*
[b] *Department of Advanced Interdisciplinary Studies, Graduate School of Engineering, The University of Tokyo, 4-6-1, Komaba, Meguro, Tokyo 153-8904, Japan*
[c] *Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 4-6-1, Komaba, Meguro, Tokyo 153-8904, Japan*

† Electronic supplementary information (ESI) available: Discussion about the model reactions for isomerization of Vgl and the configuration of MeLn, NMR spectra of the model reactions, and the experimental methods. See DOI: 10.1039/b904314d

remain intact under mild conditions, *i.e.* Vgl would be likely intact during flexizyme-catalyzing aminoacylation without isomerization to Dhb, while at elevated temperatures Vgl in peptide chain would isomerize to (*Z*)-Dhb.

At the second stage, the ribosomal synthesis of a model peptide containing Vgl at a specific site was performed (Fig. 2A, P-vgl). We used the genetic code reprogramming by means of a withdrawn PURE (protein translation using recombinant elements)[8] (*w*PURE) system integrated with the flexizyme.[4c–e] In this particular experiment, a *w*PURE system lacking tryptophan (Trp) was prepared to make the cognate codon (UGG) vacant; Vgl-tRNA$^{AsnE\text{-}2}_{CCA}$[4d] was prepared by the flexizyme and added to the *w*PURE system for the reassignment of UGG to Vgl. Tricine-SDS-PAGE analysis of the peptide expressed under three different conditions (Fig. 2B) revealed that P-vgl was expressed only in the presence of Vgl-tRNA$^{AsnE-2}_{CCA}$ in the *w*PURE system (lanes 2 *vs.* 3) and its expression level was 44% relative to the wildtype peptide (P-wt) expressed in the *w*PURE system plus Trp (lanes 1 *vs.* 3). MALDI-TOF mass spectrometry of P-vgl showed a single major peak consistent with the desired molecular mass (Fig. 2C). Upon heating P-vgl at 95 °C, we expected that the Vgl would isomerize to Dhb residue. To confirm this Vgl → Dhb isomerization, P-vgl was incubated with an excess amount of 2-mercaptoethanol (ME), by which the generated Dhb was trapped to yield the corresponding Michael-adduct (P-vgl-ME). In fact, the molecular weight of the resulting peptide increased by 78 Da that was consistent with that of ME (Fig. 2D). This result clearly demonstrates that the isomerization of Vgl to Dhb occurs under heating and the resulting Dhb residue is readily reactive with a thiol *via* intermolecular Michael addition to afford thioethers.

At the final stage of our experiments, we synthesized MeLn-containing cyclic peptides. We chose two cyclic

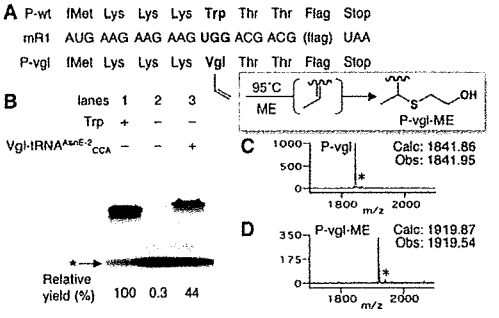


Fig. 2 Ribosomal synthesis of a Dhb-containing peptide. (A) mRNA and the expressed peptide sequences used in this study. Conversions from Vgl to Dhb, and to the ME adduct was illustrated in the grey box. Flag in parentheses indicates the RNA sequence encoding a Flag peptide (DYKDDDDK). (B) Tricine-SDS PAGE analysis of the translation products. Lane 1, wild type expression in *w*PURE system plus Trp; lane 2, *w*PURE system with tRNA$^{AsnE\text{-}2}_{CCA}$, lanes 3, *w*PURE system with Vgl-tRNA$^{AsnE\text{-}2}_{CCA}$. The band indicated by the asterisk corresponds to the remaining [^{14}C]-Asp that was not incorporated into the peptide. (C) MALDI-TOF mass spectrum of the expressed peptide in *w*PURE system before ME treatment, and (D) that after. The calculated (Calc) and observed (Obs) mass values are shown in each spectrum. Each peak labelled with asterisk denotes a sodium adduct of the major product.

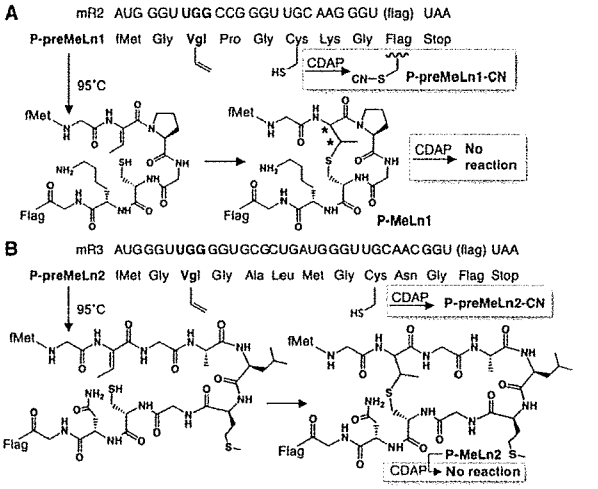


Fig. 3 Synthesis of B-ring (A) and C-ring (B) segments of nisin. CDAP reaction of each peptide is shown in the grey box. Asterisks indicate stereogenic centers newly produced by the intramolecular Michael addition. See ESI for the MALDI-TOF mass spectra of the respective peptides and the stereochemistry.†

peptides as synthetic targets based on the sequences of B- and C-ring segments of nisin (Fig. 3A and B, P-MeLn1 and P-MeLn2). We designed two mRNAs encoding the MeLn-containing cyclic peptide sequences along with the C-terminal Flag-tag appendix, in which the MeLn residue was replaced with Vgl and Cys assigned by UGG and UGC codons (mR2 and mR3). MALDI-TOF mass spectrometry of the expressed peptides showed the consistent mass values with those of Vgl-containing peptides (P-preMeLn1 and P-preMeLn2, respectively; ESI Fig. S1A and E).† We then performed thermal isomerization of the Vgl to Dhb residue and simultaneous intramolecular Michael addition of the Cys thiol. Unlike the intermolecular reaction demonstrated in Fig. 2, the molecular weight of the peptide before and after the tandem isomerization-cyclization would not change; therefore we verified if the Cys thiol in the products reacted with 1-cyano-4-dimethylaminopyridinium tetrafluoroborate (CDAP), which is known as a thiol-selective cyanylation agent[9] suitable for monitoring a loss of the free thiol group. As expected, when both precursor peptides were treated with CDAP, their respective mass values increased by 25 Da corresponding to the cyano group (ESI Fig. S1B and F).† In contrast, upon the CDAP treatment of the heated peptides (MnLn1 and MnLn2), their mass values did not change (ESI Fig. S1D and H).† This concluded that the desired cyclization took place smoothly to afford the MeLn-containing cyclic peptides.[10,11]

Here we have reported a novel methodology for the ribosomal synthesis of MeLn-containing cyclic peptides without assistance of modification enzymes. The key chemistry utilized in this method is the incorporation of Vgl into a specific site designated by a reprogrammed genetic code and the thermal isomerization of the Vgl residue to Dhb followed by simultaneous intramolecular Michael addition of the sulfhydryl group of a Cys residue present in the peptide chain. The virtue of this method requires no reagents that potentially damage the amino acid residues after the incorporation of Vgl.

Moreover, the peptide sequence is readily programmable by the mRNA template sequence. Seebeck *et al.* recently reported a method to generate dehydroalanine-containing peptides from the selenalysine-containing precursor peptides expressed by translation.[12] This elegant method relies on oxidative elimination of selenium by hydrogen peroxide to afford a dehydroalanine residue. However, strong oxidative reagents such as hydrogen peroxide could also oxidize the thiol in Cys (as well as the sulfide group in methionine) exisiting in the peptide chain, so that Cys may be no longer reactive with Dha as a Michael donor. In contrast, the method reported here resulted in a clean formation of peptides bearing intact Dhb and Cys residues, which can readily be transformed into MeLn residues by heating. Hence, a coupling of ribosomal peptide library synthesis with this methodology opens a new avenue for the discovery of novel MeLn-containing cyclic peptides with not only antimicrobial but also other biological activities when it is integrated with an appropriate screening technique.

This work was supported by grants of the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) to H. S., Grants-in-Aid for JSPS Fellows (18-10526) to Y. G., and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO).

Notes and references

1 U. Schmidt, A. Lieberknecht and J. Wild, *Synthesis*, 1988, 159–172.

2 (*a*) R. W. Jack and H. G. Sahl, *Trends Biotechnol.*, 1995, **13**, 269–278; (*b*) C. Chatterjee, M. Paul, L. Xie and W. A. van der Donk, *Chem. Rev.*, 2005, **105**, 633–684; (*c*) J. Lubelski, R. Rink, R. Khusainov, G. N. Moll and O. P. Kuipers, *Cell. Mol. Life Sci.*, 2008, **65**, 455–476, and references cited therein.

3 G. Osapay, L. Prokai, H. S. Kim, K. F. Medzihradszky, D. H. Coy, G. Liapakis, T. Reisine, G. Melacini, Q. Zhu, S. H. H. Wang, R. H. Mattern and M. Goodman, *J. Med. Chem.*, 1997, **40**, 2241–2251.

4 (*a*) A. C. Forster, Z. Tan, M. N. Nalam, H. Lin, H. Qu, V. W. Cornish and S. C. Blacklow, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 6353–6357; (*b*) K. Josephson, M. C. Hartman and J. W. Szostak, *J. Am. Chem. Soc.*, 2005, **127**, 11727–11735; (*c*) H. Murakami, A. Ohta, H. Ashigai and H. Suga, *Nat. Methods*, 2006, **3**, 357–359; (*d*) A. Ohta, H. Murakami, E. Higashimura and H. Suga, *Chem. Biol.*, 2007, **14**, 1315–1322; (*e*) Y. Sako, Y. Goto, H. Murakami and H. Suga, *ACS Chem. Biol.*, 2008, **3**, 241–249.

5 (*a*) T. Bretschneider, W. Miltz, P. Munster and W. Steglich, *Tetrahedron*, 1988, **44**, 5403–5414; (*b*) D. B. Berkowitz, J. M. McFadden and M. K. Sloss, *J. Org. Chem.*, 2000, **65**, 2907–2918.

6 Although we also attempted the synthesis of *N*-acetyl-Vgl-NH-Et and the model reaction using it, we ended up isolating only *N*-acetyl-Dhb-NH-Et, which made our isomerization studies impossible. Presumably, under basic conditions the isomerization is readily promoted.

7 (*a*) M. Slijper, C. W. Hilbers, R. N. H. Konings and F. J. M. Vandeven, *FEBS Lett.*, 1989, **252**, 22–28; (*b*) H. Zhou and W. A. van der Donk, *Org. Lett.*, 2002, **4**, 1335–1338; (*c*) Y. T. Zhu, M. D. Gieselman, H. Zhou, O. Averin and W. A. van der Donk, *Org. Biomol. Chem.*, 2003, **1**, 3304–3315.

8 Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa and T. Ueda, *Nat. Biotechnol.*, 2001, **19**, 751–755.

9 J. Wu and J. T. Watson, *Protein Sci.*, 1997, **6**, 391–398.

10 The estimated yield of P-MeLn1 and P-MeLn2 were 1.1 and 0.8 pmol μL^{-1}, respectively. See Methods in the ESI for more details†.

11 For the discussion about the configuration of MeLn, see ESI†.

12 F. P. Seebeck and J. W. Szostak, *J. Am. Chem. Soc.*, 2006, **128**, 7150–7151.

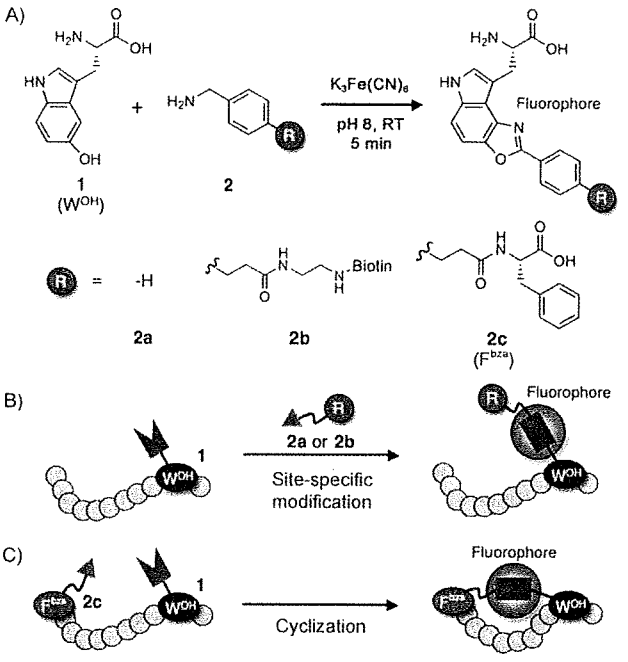
DOI: 10.1002/cbic.200900021

Ribosomal Synthesis of Cyclic Peptides with a Fluorogenic Oxidative Coupling Reaction

Yusuke Yamagishi,[a] Hiroshi Ashigai,[a] Yuki Goto,[c] Hiroshi Murakami,[b] and Hiroaki Suga *[a, b, c]

Cyclic peptides often have better physiological properties than linear peptides due to their increased structural rigidity and peptidase resistance.[1–3] Despite the fact that translation provides an attractive platform for the synthesis of cyclic peptides, those expressed by the translation apparatus have cysteine–cysteine (Cys–Cys) disulfide bridges, and this limits their use. However, because substitution of the disulfide bond with a nonreducible bond is expected to give better physiological stability, it has been of interest to develop novel methodologies for generating such nonstandard cyclic peptides expressed by translation. For instance, a non-natural amide linkage was introduced between the N-terminal α-amino and the lysine ε-amino groups by the treatment of peptides with a crosslinking reagent that had been activated by *N*-hydroxysuccinimide.[4,5] Likewise, the sulfhydryl groups of two Cys residues existing in peptides were crosslinked with α,α′-dibromo-*m*-xylene to generate a unique cyclic peptide.[6] In contrast to these approaches, we have devised unique methods that use nonproteinogenic amino acids with sidechains that selectively form linkages to the cognate pair.[7–9] Here, we extend such methodology to incorporate a pair of nonproteinogenic amino acids containing benzylamine and 5-hydroxyindole functionalities into designated sites assigned by reprogrammed codons. Under oxidative conditions, these groups instantly react with each other and form a fluorescent heterocyclic moiety. Thus, this method enables the conversion of a nonfluorescent linear peptide to a fluorescent cyclic peptide.

5-Hydroxytryptamine (serotonin) is known to react rapidly and under mild conditions with benzylamine (**2a**) in the presence of potassium ferricyanide, $K_3Fe(CN)_6$, to yield a benzoxazole derivative (**14**, Figures S1 and S5 in the Supporting Information).[10–12] It is noteworthy that the resulting adduct is fluorescent with an excitation maximum at 345 nm and an emission maximum at 460 nm. We have envisioned that this oxidative reaction would allow us to develop a new class of orthogonal pairs for the intermolecular and intramolecular reactions of polypeptides involving 5-hydroxytryptophan (W^{OH}, **1**) and three benzylamine derivatives (**2a–c**, Scheme 1A). As a



Scheme 1. A) Fluorogenic oxidative coupling reactions between 5-hydroxytryptophan **1** (W^{OH}) and benzylamine derivatives **2**. Reaction between **1** and **2** is achieved by treatment with K_3FeCN_6 in aqueous media to form the fluorophore structure. B) Schematic presentation of the site-specific modification of peptides bearing W^{OH} by intermolecular reaction with benzylamine derivatives (**2a** or **2b**). Blue symbol and red triangle represent 5-hydroxyindole and benzylamine groups, respectively. C) Schematic presentation of the cyclization of peptides by the intramolecular reaction between **1** and **2c** (F^{bza}).

preliminary study, W^{OH} was incorporated into a model peptide at a specific site, and the resulting peptide could be site-selectively modified with a benzylamine (**2a**) or its biotin-bearing derivative (**2b**) (Scheme 1B). Upon successful demonstration of the above intermolecular coupling reaction, we attempted cyclization of peptides between the benzylamine attached to the α-amino group of phenylalanine (F^{bza}, **2c**) and W^{OH} in the nascent chain to give fluorescent cyclic peptides (Scheme 1C). In both cases, incorporation of the above nonproteinogenic amino acids into peptides was achieved by genetic code reprogramming by using the flexizyme system,[13,14] and a modified peptide-translation by using the recombinant elements (PURE) system.[15–17] The former system consists of *flexi*ble tRNA acylation ribo*zymes*, and the latter system is a reconstituted

[a] *Y. Yamagishi, H. Ashigai, Prof. Dr. H. Suga*
Department of Chemistry and Biotechnology,
The University of Tokyo
7-3-1, Hongo, Bunkyo, Tokyo (Japan)
Fax: (+81) 03-5452-5495
E-mail: hsuga@rcast.u-tokyo.ac.jp

[b] *Dr. H. Murakami, Prof. Dr. H. Suga*
Research Center for Advanced Science and Technology,
The University of Tokyo
4-6-1, Komaba, Meguro, Tokyo (Japan)

[c] *Y. Goto, Prof. Dr. H. Suga*
Department of Advanced Interdisciplinary Studies
Graduate School of Engineering, The University of Tokyo
4-6-1, Komaba, Meguro, Tokyo (Japan)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200900021: further details of all experiments.

cell-free translation system in which isoleucine (Ile) and/or methionine (Met) were withdrawn, referred to as *w*PURE system.

To reassign W^{OH} to AUC (one of the vacant Ile codons), W^{OH}-$tRNA^{Asn\text{-}E1}_{GAU}$ prepared by the flexizyme was added to a *w*PURE system lacking Ile. An mRNA template expressing a dodecamer model peptide containing the Ile codon at position 11 (**3**, Figure S5) was subjected to translation. The expressed peptide was desalted by using a conventional C_{18}-resin filter, and the isolated peptide was analyzed by using a MALDI-TOF mass spectrometer. The observed molecular mass (ms) was consistent with the expected ms of peptide **3** containing W^{OH}11 (Figure S5B). The expression level of **3** was approximately 51% relative to that of the wild-type peptide containing Ile11 expressed in the ordinary PURE system (Figures S3 and S4); this value translated to approximately 0.46 pmol μL^{-1} of **3** compared with 0.91 pmol μL^{-1} of the wild-type as determined by the quantification method reported elsewhere.[18]

For the conjugation, the translation solution containing crude peptide **3** was treated with the benzylamine derivatives **2a** or **2b** in the presence of $K_3Fe(CN)_6$ at room temperature for 5 min (Figure S5A) and then desalted. MALDI-TOF analysis revealed that **3** was modified by the respective reagents to yield **3a** and **3b** without undesirable side-reactions (Figure S5C and S5D). To further confirm the formation of the conjugation through a heterocyclic structure, we next examined the fluorescent properties of peptides **3**, **3a**, and **3b** (Figure S5E). The unmodified peptide **3** showed nearly the same emission profile as a blank sample; this indicates that it does not exhibit any fluorescent properties under such conditions (Figure S5E, **3**). In contrast, both modified peptides **3a** and **3b** exhibited similar fluorescent profiles with an excitation maximum at 345 nm and an emission maximum at 460 nm (Figure S5E, **3a** and **3b**). The above fluorescent profiles were nearly identical to that of fluorophore **14** (Figures S1and S5E), which is derived from 5-hydroxyindole and **2a**; this implies that the same fluorophore was generated. Because the molecular absorption coefficient and fluorescent quantum yield were not available in the literature, we determined these parameters for **14** under similar conditions to which the peptides were monitored. These values were determined to be $1.54\times10^4\ cm^{-1}\,M^{-1}$ and 0.54 at pH 8.0 in 250 mM borate buffer. The above data encouraged us to perform the same chemistry in intramolecular fashion and apply it to the directed cyclization of peptides.

We synthesized a phenylalanine derivative bearing a benzylamine group on the α-amino group (Scheme 1A, **2c**) and charged it onto $tRNA^{fMet}_{CAU}$ by means of flexizyme.[7,8,19,20] A *w*PURE system lacking Met and Ile was prepared, to which F^{bza}-$tRNA^{fMet}_{CAU}$ and W^{OH}-$tRNA^{Asn\text{-}E1}_{GAU}$ were added in order to initiate the translation with F^{bza} instead of *N*-formylmethionine (fMet); this was followed by elongation with the corresponding amino acid sequence including W^{OH}11 (Figure 1A, **4**). Oxidation of dodecamer **4** in the presence of $K_3Fe(CN)_6$ resulted in a decrease

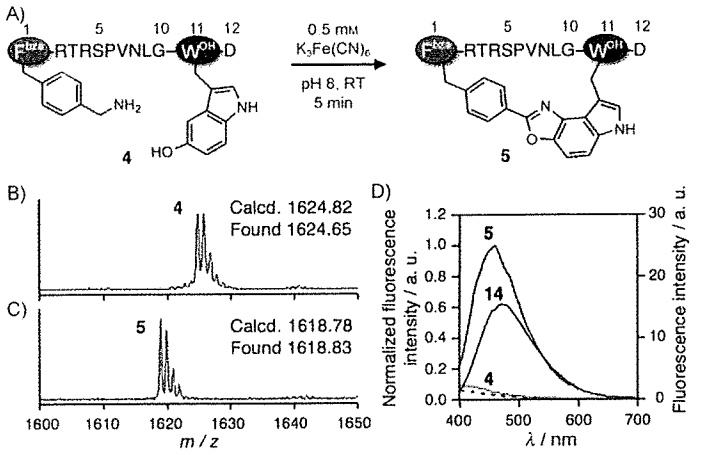


Figure 1. A) Intramolecular cyclization of peptide **4** through an oxidative coupling reaction. B) MALDI-TOF spectra of linear peptide **4** and C) cyclic peptide **5**. The calculated molecular mass (Calcd.) and found molecular mass (Found) for singly charged species, $[M+H]^+$ are shown in the spectra. D) Emission spectra of **4**, **5**, and **14**. The spectra were recorded in 250 mM borate buffer and DMSO 10% (*v*/*v*) adjusted pH to 8.0. The green line indicates the fluorescent profile of 300 nM **14** at an excitation wavelength of 326 nm (right axis), and other lines show the emission profiles at an excitation wavelength of 345 nm (left axis). The dashed line shows a blank profile obtained in the above buffer, while red and blue lines indicate the profiles of **4** and **5**, respectively.

in molecular weight that corresponded to a loss of six protons observed in MALDI-TOF analysis of the product; this suggested that **5** was being produced (Figure 1B and C). The excitation and emission profiles of **5** showed the respective maxima at 345 and 460 nm, whereas such a profile was not observed for **4** (Figure 1D). The fluorescent profile of **5** was similar to that of **14** as well as the modified peptides **3a** and **3b**. This result solidified the idea that the linear peptide **4** was cyclized to afford **5** with the expected eleven-membered ring structure closed by the fluorescent linker (Figure 1A).

To survey the orthogonality of the oxidative coupling reaction with various sidechains of proteinogenic amino acids, we designed peptide sequences containing F^{bza}, W^{OH} and those amino acids that are potentially reactive under oxidative conditions, such as Lys, Cys, Tyr and Trp (as opposed to an unreactive Leu) at position 13 (Figure 2A, **6–10**). Expression of peptides **7–10** took place smoothly with efficiencies of the range of 51–75% relative to the control Leu13-containing peptide **6** (Figure 2B, lanes 1–5). MALDI-TOF analysis of the resulting peptides indicated that all were converted to the desired cyclic peptides based on the observation of the same mass change corresponding to the loss of six protons (Figure S6A–E) with one exception; peptide **8**, which contains Cys13, seemed to have an extra set of peaks with an average of −8.07 mass shift along with the set of peaks with a −6.04 shift (Figure S6C, **8**). To the best of our knowledge, the sulfhydryl group of glutathione intermolecularly reacts with a dione derived from the oxidation of W^{OH}; this results in a 7-*S*-glutathionyl-tryptophan-4,5-dione adduct.[21] Thus, we attributed the above observation to a similar side-reaction between Cys13 and W^{OH}7 that competed with the reaction between F^{bza} and W^{OH}7. To further confirm the oxidative coupling between F^{bza} and W^{OH}7, we mea-

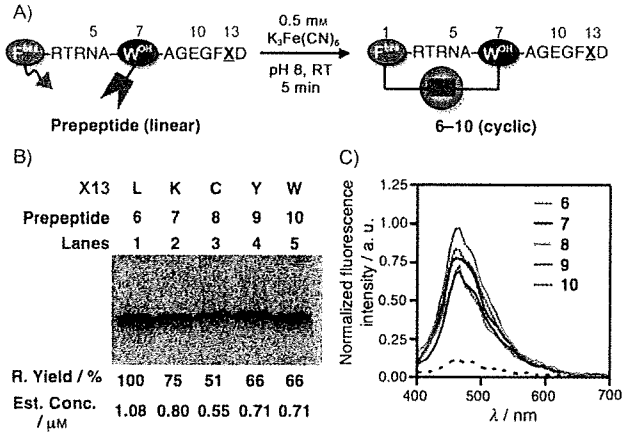


Figure 2. Competitive reaction between W^{OH} and potentially reactive proteinogenic amino acids against W^{OH} and F^{bza}. A) Peptide sequences used for the verification. The proteinogenic residue X13 was altered to L, K, C, Y, and W. B) Tricine SDS-PAGE analysis of the linear peptide labeled with [^{14}C]-Asp. The band of peptide was quantified by radio-autoradiography. Lane 1, **Pre-6** (X13=L); lane 2, **Pre-7** (K); lane 3, **Pre-8** (C); lane 4, **Pre-9** (Y); lane 5, **Pre-10** (W). The peptide concentration was estimated by our standard quantification method in which the radioisotope (RI) counts of [^{14}C]-Asp were plotted against its known concentration to determine the calibration line. Then, the RI counts of the expressed peptides were fitted to the calibration line to estimate their respective concentration. R. Yield and Est. Conc. denote relative yield and estimated concentration. C) Emission spectra of cyclic **6–10**. The spectra were recorded in 250 mM borate buffer and DMSO 10% (*v/v*) adjusted pH at 8.0. Solid lines show the emission profiles at an excitation wavelength of 345 nm. The dashed line shows a blank profile obtained in translation without DNA template, while red, blue, orange, green and cyan lines indicate those of **6**, **7**, **8**, **9**, and **10**, respectively.

sured the fluorescence spectra of the product (Figure 2C). Fluorescence intensities were normalized according to the concentration of the respective linear peptides **7–10** against the control peptide **6** and quantified by tricine SDS-PAGE analysis (Figure 2B). All peptides exhibited similar fluorescent properties; this is consistent with the idea that they were cyclized. It should be noted that even though production of peptide **8** was accompanied by side-product formation, its fluorescent intensity was similar to that of **6** (Figure 2C, spectrum **8**). Because we expected that the side-product would not show a similar fluorescent property, we concluded that the side-reaction between Cys13 and $W^{OH}7$ was fairly minor; this provided evidence that the intramolecular coupling reaction between F^{bza} and $W^{OH}7$ took place in a fairly selective manner.

Lastly, we verified that this cyclization method would be applicable to the synthesis of peptides with various ring sizes (Figure 3). We prepared three different mRNAs capable of expressing decamer, octamer, and hexamer peptides in the presence of F^{bza}-$tRNA^{fMet}_{CAU}$ and W^{OH}-$tRNA^{Asn\text{-}E1}_{GAU}$ in the Met/Ile-lacking *w*PURE system. Treatment of the expressed peptides with $K_3Fe(CN)_6$ was expected to produce the corresponding cyclic peptides rings containing W^{OH} at positions 9, 7, and 5 in peptides **11–13**, respectively (Figure 3A–C). In all cases, the linear peptides were cleanly converted to the cyclic peptides with the expected loss of six protons detected by MALDI-TOF analysis (Figure 3A–C, ms data in the middle panel). Moreover, all of the cyclic peptides **11–13** exhibited nearly identical fluorescent profiles to **14**, whereas the linear ones showed no fluorescence (Figure 3A–C, the bottom panel). These data clearly demonstrate that this method is applicable to the cyclization of peptides with a wide range of ring sizes (from five- to eleven-membered rings).

In conclusion, we have developed a new method for the cyclization of peptides that involves oxidative coupling of W^{OH} and F^{bza} attached to their N terminus by means of reprogramming the genetic code. The selective coupling of the above pairs takes place rapidly upon addition of $K_3Fe(CN)_6$ and generates the desired fluorescent heterocyclic linkage. This chemistry enables us to convert nonfluorescent linear peptides to fluorescent cyclic peptides. Notably, the reaction takes place independently from the forming ring sizes, at least in the range of five- to eleven-membered rings as demonstrated in the current study. The chemistry demonstrated here by using the *w*PURE system is readily applicable to the platform of solid-phase or

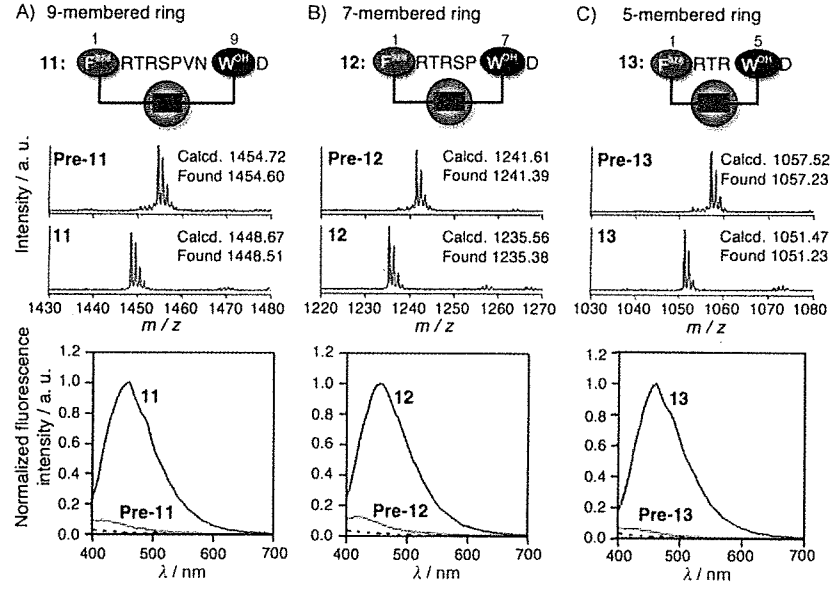


Figure 3. Synthesis of cyclic peptides with various ring sizes. Peptides **11–13** consist of A) 9, B) 7, and C) 5-membered rings. The upper panel of each Figure shows the structure of cyclic peptide, the middle panel exhibits MALDI-TOF spectra of precursor (**Pre-11–13**) and cyclic peptides (**11–13**), and bottom panel shows the excitation and emission profiles of **Pre-11–13** and **11–13**. The dashed line shows a blank profile. The conditions and designations of lines are the same as Figure 1.

solution-phase chemical synthesis, and thereby allows us to access to the large-scale production of the fluorescent cyclic peptide. On the other hand, the ribosomal expression of the peptides in the mRNA-encoding manner enables us to couple with an appropriate in vitro display technique that facilitates the screening of a large sequence space of cyclic peptides. Thus, the synthetic methodology reported herein offers us a new means for the discovery of cyclic peptide probes containing a unique fluorescent heterocyclic structure against various biological targets.

Experimental Section

Chemical synthesis of the amino acid derivatives, biotinyl benzylamine, and the benzoxazole fluorophore is described in the Supporting Information. Flexizyme, $tRNA^{Asn\text{-}E1}_{GAU}$, $tRNA^{fMet}_{CAU}$, and DNA templates coding peptides were prepared as previously reported.[13] Aminoacyl-tRNAs (W^{OH}-$tRNA^{Asn\text{-}E1}_{GAU}$ and F^{bza}-$tRNA^{fMet}_{CAU}$) were prepared by using the flexizyme system. The *w*PURE system was reconstituted with the same components as the PURE system except that the necessary proteinogenic amino acids were added. Peptide **3**, which contained W^{OH}, was synthesized by using W^{OH}-$tRNA^{Asn\text{-}E1}_{GAU}$ and the *w*PURE system without Ile. Linear peptides **4**, and **6–13** containing W^{OH} and F^{bza} were synthesized by using W^{OH}-$tRNA^{Asn\text{-}E1}_{GAU}$, F^{bza}-$tRNA^{fMet}_{CAU}$ and the *w*PURE system without Ile and Met. Translation mixtures were incubated at 37 °C for 30 min. For MALDI-TOF analysis, 5.0 μL of reaction mixture was acidified with TFA (1% (*v*/*v*)) in H_2O and purified by using a PerfectPure C-18 Tip (Eppendorf, Hamburg, Germany). Peptide product on the resin was eluted with an acetonitrile (50% (*v*/*v*))/TFA (0.1% (*v*/*v*)) solution (2.0 μL total) saturated with the matrix α-cyano-4-hydroxycinnamic acid (Bruker Daltonics, Bremen, Germany).

For MALDI-TOF analysis of oxidation coupling reactions, modification was performed as follows: borate buffer (1.0 M, 2.5 μL, pH 8.0), **2a** or **2b** in DMSO (100 mM, 1.0 μL), and $K_3Fe(CN)_6$ in H_2O (13.3 mM, 1.5 μL) were added to the translation mixture (5.0 μL) containing **3** (0.46 μM). Components and concentrations of translation mixture were described in elsewhere.[22] On the other hand, cyclization was carried out as follows: borate buffer (500 mM, 2.5 μL, pH 8.0), and $K_3Fe(CN)_6$ in H_2O (2.5 mM, 2.5 μL) were added to the translation mixture (5.0 μL). The resulting mixture was left for 5 min at room temperature. The translation mixture was acidified with TFA in H_2O (1% (*v*/*v*)). The product was purified by using a C-18 Tip and eluted with acetonitrile (50% (*v*/*v*))/ TFA (0.1% (*v*/*v*)) solution (2.0 μL total) saturated with the matrix α-cyano-4-hydroxycinnamic acid.

For fluorescence analysis of oxidation coupling reactions, each peptide was purified by using a C-18 Tip from translation mixture (5.0 μL), and eluted with an acetonitrile (50% (*v*/*v*))/TFA (0.1% (*v*/*v*)) solution (10 μL total). The eluent was evaporated for 30 min at room temperature, and the resulting residue was dissolved in H_2O (1.25 μL). Borate buffer (1.0 M, 1.25 μL, pH 8.0), **2a** or **2b** in DMSO (100 mM, 0.5 μL) were added to modified peptide **3**, and then $K_3Fe(CN)_6$ in H_2O (5 mM, 2.0 μL) was added. To generate cyclic peptide **5** and **6–13**, the peptide residue was dissolved in H_2O (1.25 μL), borate buffer (1.0 M, 1.25 μL, pH 8.0), and DMSO (0.5 μL), and then added to $K_3Fe(CN)_6$ in H_2O (1.25 mM, 2.0 μL). The resulting mixture was incubated for 5 min at room temperature. The emission and excitation spectra were measured by using a 1536-well micro plate (BD Falcon, Franklin Lakes, USA) at 25 °C. Excitation and emission slit-widths were kept constant at 5.0 nm.

Acknowledgements

We thank Dr. M. Komiyama for the use of UV-Vis spectrophotometer and spectrofluorometer and Dr. P. C. Reid for critical proofreading. We thank Drs. I. Hirao and T. Mitsui in the RIKEN Genomic Sciences Center for the use of a FAB mass spectrometer. This work was supported by grants of Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO) to H.S.

Keywords: cyclization · fluorescent probes · oxidation · peptides · protein modifications

[1] P. Li, P. P. Roller, *Curr. Top. Med. Chem.* **2002**, *2*, 325–341.
[2] J. Grunewald, M. A. Marahiel, *Microbiol. Mol. Biol. Rev.* **2006**, *70*, 121–146.
[3] C. Chatterjee, M. Paul, L. Xie, W. A. van der Donk, *Chem. Rev.* **2005**, *105*, 633–684.
[4] S. W. Millward, T. T. Takahashi, R. W. Roberts, *J. Am. Chem. Soc.* **2005**, *127*, 14142–14143.
[5] S. W. Millward, S. Fiacco, R. J. Austin, R. W. Roberts, *ACS Chem. Biol.* **2007**, *2*, 625–634.
[6] F. P. Seebeck, J. W. Szostak, *J. Am. Chem. Soc.* **2006**, *128*, 7150–7151.
[7] Y. Goto, A. Ohta, Y. Sako, Y. Yamagishi, H. Murakami, H. Suga, *ACS Chem. Biol.* **2008**, *3*, 120–129.
[8] Y. Sako, Y. Goto, H. Murakami, H. Suga, *ACS Chem. Biol.* **2008**, *3*, 241–249.
[9] Y. Sako, J. Morimoto, H. Murakami, H. Suga, *J. Am. Chem. Soc.* **2008**, *130*, 7232–7234.
[10] J. Ishida, M. Yamaguchi, M. Nakamura, *Analyst* **1991**, *116*, 301–304.
[11] R. Sundaramoorthi, V. K. Kansal, B. C. Das, P. Potier, *J. Chem. Soc. Chem. Commun.* **1986**, *5*, 371–372.
[12] D. L. Boger, L. R. Cerbone, D. Yohannes, *J. Org. Chem.* **1988**, *53*, 5163–5166.
[13] H. Murakami, A. Ohta, H. Ashigai, H. Suga, *Nat. Methods* **2006**, *3*, 357–359.
[14] M. Ohuchi, H. Murakami, H. Suga, *Curr. Opin. Chem. Biol.* **2007**, *11*, 537–542.
[15] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat. Biotechnol.* **2001**, *19*, 751–755.
[16] A. Ohta, Y. Yamagishi, H. Suga, *Curr. Opin. Chem. Biol.* **2008**, *12*, 159–167.
[17] T. J. Kang, H. Suga, *Biochem. Cell Biol.* **2008**, *86*, 92–99.
[18] Our general method for the peptide quantification was as follows: radioisotope (RI) counts of [^{14}C]-Asp were plotted against its known concentrations to determine the calibration line. Then, RI counts of the expressed wild-type peptide and cyclic peptides were fitted to the calibration line to estimate the individual concentration of peptides. We found that this quantification method was reliable to assess the quantity of peptide previously proven by performing activity assays by using a known bioactive peptide (see refs. [7] and [8]).
[19] S. Mamaev, J. Olejnik, E. K. Olejnik, K. J. Rothschild, *Anal. Biochem.* **2004**, *326*, 25–32.
[20] J. Olejnik, S. Gite, S. Mamaev, K. J. Rothschild, *Methods* **2005**, *36*, 252–260.
[21] Z. Wu, G. Dryhurst, *Bioorg. Chem.* **1996**, *24*, 127–149.
[22] T. Kawakami, H. Murakami, H. Suga, *Chem. Biol.* **2008**, *15*, 32–42..

Received: January 12, 2009
Published online on May 26, 2009

CHEMBIOCHEM

DOI: 10.1002/cbic.200900058

Ribosomal Synthesis of Peptides with C-Terminal Lactams, Thiolactones, and Alkylamides

Eiji Nakajima,[a] Yuki Goto,[a] Yusuke Sako,[a] Hiroshi Murakami,[b] and Hiroaki Suga*[a, b]

The C terminus of a peptide expressed by the translation apparatus generally ends in a carboxylate group. On the other hand, the C termini of some naturally occurring peptides have amide moieties instead of carboxylates, which are believed to give better biostability. Here, we describe a new strategy for the ribosomal synthesis of peptides featuring C-terminal lactam, thiolactone, and alkylamide units. The method was based on the concept of genetic code reprogramming involving the flexizymes (flexible tRNA acylation ribozymes) and the PURE (peptide synthesis using recombinant elements) system, in which vacant codons are reassigned to nonproteinogenic amino acids; this enabled us to convert the C termini of peptides into the above functionalities. We have also applied this method to the synthesis of a macrocyclic peptide closed by an amide bond formed between a lysine side chain and the peptide C terminus. This method thus offers us new opportunities to express various peptides with C-terminal modifications as well as macrocyclic peptides using the translation apparatus, and potentially to accelerate the discovery of peptidic drugs designed for various therapeutic targets.

Introduction

The C terminus of a peptide expressed by the translation apparatus generally ends in a carboxylate group, which is generated by ribosome-catalyzed hydrolysis of the carboxy ester bond between the C-terminal amino acid and the 3′-OH of peptidyl-tRNA and triggered by binding of the release factors to the ribosome A site.[1,2] On the other hand, it is known that the C termini of some naturally occurring peptides are modified by an amide instead of the carboxylate,[3–5] and potentially contribute to their biostability.[6–8] These C-terminal amide groups are generated by enzymatic oxidative cleavage at the α positions of the subsequent amino acids in the parent peptides.[9] More recently, intein, known as a protein-splicing enzyme, has been engineered to trap the thioester intermediate, which subsequently reacts with an externally added amine to form the C-amidated peptide.[10] Notably, this methodology has been applied to C-terminal modifications featuring not only simple amides, but also alkylamides.[11,12]

We have been engaged in a research program directed towards the development of new methodologies for the ribosomal synthesis of nonstandard peptides. Since the translation apparatus is restricted to the use of the 20 standard proteinogenic amino acids assigned by triplet nucleotides (codons) on mRNA, the expressed peptides generally consist of combinations of these amino acids. However, by breaking the "natural" assignments between codons and proteinogenic amino acids, it is possible to express nonstandard peptides containing multiple nonproteinogenic amino acids. We refer to this "artificial" codon reassignment as genetic code reprogramming.[13–22] Despite the simplicity of the concept behind this approach, there have been two major technical obstacles to its achievement: 1) how the natural assignment of the codon or codon box to the proteinogenic amino acid can be broken, that is, how can the codon or codon box be made vacant for the reassignment; and 2) how the nonproteinogenic amino acid(s) can be charged with a tRNA bearing the anticodon corresponding to the vacant codon.

We have utilized two systems to overcome these inconveniences. To assist with point (1), we have taken advantage of the manipulatability of a reconstituted *E. coli* cell-free translation system, often referred to as the PURE (peptide synthesis using recombinant elements) system.[23–25] The most important feature of this translation system is that certain amino acids, and, if necessary, aminoacyl-tRNA synthetases, can be withdrawn from the translation elements. Use of such a withdrawn PURE system, termed *w*PURE, allows us to create desired vacant codons.

To assist with point (2), we have developed an enzyme system, named flexizyme system, to facilitate the preparation of tRNAs charged with various nonproteinogenic amino acids. It consists of two flexizymes, dFx and eFx, that can be used depending on the choice of leaving group in the substrates: dFx for a 3,5-dinitrobenzyl ester (DBE) and eFx either for a 4-chlorobenzyl thioester (CBT) or for a cyanomethyl ester (CME). The use of a combination of these two systems has thus allowed us to reprogram the genetic code, and enabled ribosomal

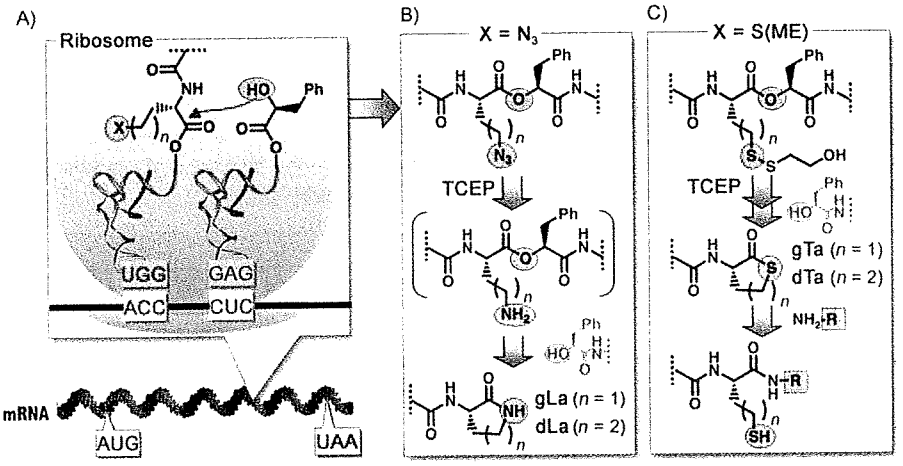
[a] *E. Nakajima, Dr. Y. Goto, Dr. Y. Sako, Prof. H. Suga*
Department of Chemistry and Biotechnology
Graduate School of Engineering, The University of Tokyo
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656 (Japan)
Fax: (+81) 35-452-5495
E-mail: hsuga@rcast.u-tokyo.ac.jp

[b] *Dr. H. Murakami, Prof. H. Suga*
Research Center for Advanced Science and Technology
The University of Tokyo
4-6-1 Komaba, Meguro-ku, Tokyo 153-8904 (Japan)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.200900058.

synthesis of nonstandard peptides, the sequences and compositions of which are encoded in the mRNA sequences.[26–36]

As a new application of this genetic code reprogramming, here we report a methodology for the ribosomal synthesis of peptides featuring C-terminal modifications involving the site-specific incorporations of γ-azidohomoalanine (Aha) or 2-mercaptoethanol-masking homocysteine ((ME)Hcy) followed by phenyllactic acid (F^{lac}) under the reprogrammed genetic code (Scheme 1 A). The resulting peptides contain a peptide bond originating from the above nonproteinogenic amino acid, followed by a single ester bond originating from F^{lac}. Subsequent treatment of such peptides with tris(2-carboxyethyl)phosphine (TCEP) triggers the conversion of the terminal residue either into a γ-lactam (Scheme 1 B, gLa) or into a γ-thiolactone (gTa) that can be converted into a variety of alkylamides (Scheme 1 C). Moreover, when a lysine residue is incorporated upstream of the C-terminal gTa, macrocyclization takes place to afford a cyclic peptide. This methodology thus enables us to prepare a wide variety of C-terminal-modified peptides using the mRNA-directed translation platform.



Scheme 1. A strategy for ribosomal synthesis of nonstandard peptides containing lactam, thiolactone, and alkylamide moieties. A) Schematic presentation of successive incorporation of a nonproteinogenic amino acid (Xaa) and phenyllactic acid (F^{lac}) in a ribosome. Xaa-$tRNA^{Asn-E1}_{GGU}$ and F^{lac}-$tRNA^{Asn-E1}_{GAG}$ decode ACC and CUC codons, respectively, on the mRNA and result in the Xaa-F^{lac} sequence, linked by an ester bond. B) Post-translational TCEP reduction of the azide side chain, which triggers the formation of γ/δ-lactams at the C terminus of the peptide; TCEP: tris(2-carboxyethyl)phosphine. C) Post-translational TCEP reduction of the ME-protected sulfhydryl side chain, which triggers the formation of a γ/δ-thiolactone, followed by inter- or intramolecular amidation at the C terminus of the peptide; ME: 2-mercaptoethanol.

Results and Discussion

Incorporation of the unnatural substrates and synthesis of C-terminally cyclized peptides

To construct the foundation for the above methodology, we first examined the incorporation of a nonproteinogenic amino acid bearing a nucleophilic side chain, such as a γ-amino or thio group, followed by F^{lac} into a peptide chain at specific sites; the expressed peptide should then self-rearrange to afford the corresponding γ-lactam or thiolactone. Unfortunately, initial attempts making direct use either of γ-aminohomoalanine or of Hcy did not succeed in activation of the carboxyl group by DBE for tRNA aminoacylation (vide infra) because of the occurrence of rapid self-cyclization. We therefore revised the strategy and masked the amino group in γ-aminohomoalanine by replacing it with a γ-azide moiety (Figure 1 A, Aha). In the case of Hcy, its sulfhydryl group was masked by a disulfide bond with 2-mercaptoethanol (ME; Figure 1 A, (ME)Hcy). Conveniently, both masking groups could be converted into the corresponding nucleophiles by exposure to TCEP; thereby the system was primed to trigger the desired spontaneous C-terminal cyclization.

To incorporate Aha or (ME)Hcy (for convenience, we refer to these nonproteinogenic amino acids as Xaa) and F^{lac} successively into a peptide chain at the designated positions, threonine (T; ACC) and leucine (L; CUC) codons were chosen for the reassignment of Xaa and F^{lac}, respectively. The reassignment was achieved by the combination of the appropriate flexizyme and *w*PURE systems. Xaa and F^{lac} were esterified with DBE and CME groups, respectively, and thus were made suitable substrates for dFx and eFx (Figure 1 A). To make the ACC and CUC codons vacant, we prepared a *w*PURE system in which amino acids T and L were withdrawn. The reassignment of Xaa and F^{lac} to these vacant codons was achieved by preparation of the corresponding acyl-tRNAs bearing the anticodons. Specifically, $tRNA^{Asn-E1[27]}_{GGU}$ was charged with Aha and (ME)Hcy by using dFx, while $tRNA^{Asn-E1}_{GAG}$ was charged with F^{lac} by using eFx. Similarly, $tRNA^{Asn-E1}_{GGU}$ was charged with δ-azidonorvaline (Anv) and δ-(ME)mercaptonorvaline ((ME)Mnv; Figure 1 A, $n=2$).[27] The acylation efficiency of each substrate was confirmed by using a tRNA analogue: microhelix RNA (Figure S1 in the Supporting Information). These acyl-tRNAs were thus used to express the desired nonstandard peptides.

To observe the successive incorporations of Xaa followed by F^{lac} into the peptide chain, we designed an mRNA template expressing a model peptide (Figure 1 B, mR1 and P1-Xaa-F^{lac}). Note that in this model peptide the FLAG sequence was upstream of Xaa-F^{lac}. This allowed us to purify the peptide using anti-FLAG antibody. In this way, the whole translated products could be isolated for the detection not only of the expected full-length peptide but also of truncated peptides generated by potential failures of the codon reassignments. When

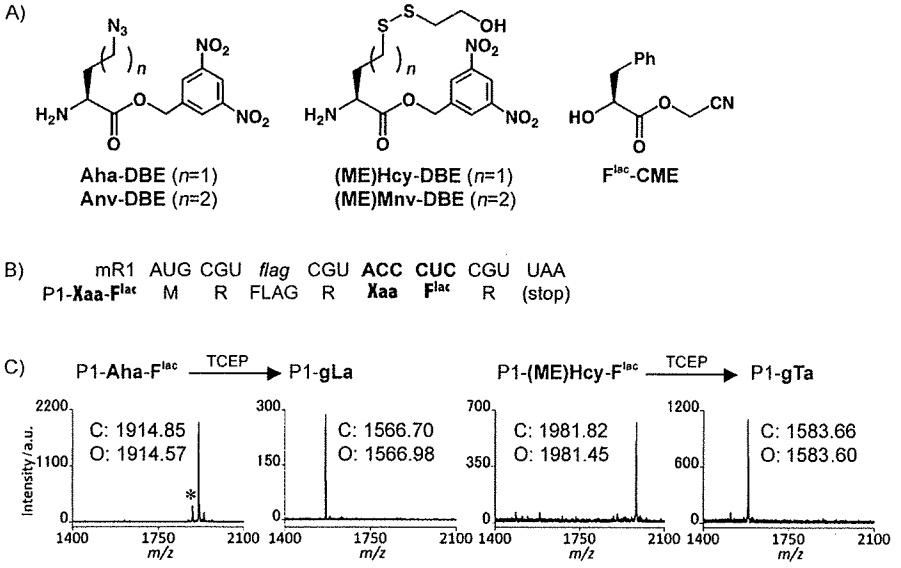


Figure 1. Ribosomal synthesis of the C-terminal γ/δ-lactam- and γ/δ-thiolactone-containing peptides, P1-gLa and P1-gTa, respectively. A) Structures of the nonproteinogenic amino acid DBEs and phenyllactic acid CME; Aha: azidohomoalanine; Anv: azidonorvaline; (ME)Hcy: 2-mercaptoethanol-protected homocysteine; (ME)Mnv: 2-mercaptoethanol-protected mercaptonorvaline; DBE: 3,5-dinitrobenzyl ester; CME: cyanomethyl ester. B) Sequences of mR1 and P1-Xaa-F^{lac}; *flag* indicates the RNA sequence coding the FLAG peptide (DYKDDDDK). Xaa and F^{lac} are assigned to codons ACC and CUC, respectively. C) MALDI-TOF analysis of the expressed peptides. The calculated (C) and observed (O) molecular masses for singly charged species, $[M+H]^+$, of the peptide are shown in each spectrum. TCEP reduction of P1-Aha-F^{lac} and P1-(ME)Hcy-F^{lac} resulted in the formation of γ-lactam and γ-thiolactone, respectively, at the peptide C terminus. The asterisk (*) denotes a peak of P1-Aha-F^{lac} in which Aha was presumably reduced by DTT present in the translation mixture; calcd: 1890.02, found: 1890.71.

translation was executed in the presence of Xaa-$tRNA^{Asn\text{-}E1}_{GGU}$ and F^{lac}-$tRNA^{Asn\text{-}E1}_{GAG}$ in the *w*PURE system, and the FLAG-isolated product obtained from the experiment was analyzed by MALDI-TOF, a single major peak was observed (Figure 1C, P1-Aha-F^{lac} and P1-(ME)Hcy-F^{lac}). In all cases, the major peak was consistent with the expected molecular mass of the full-length peptide, whereas a minor peak observed in the P1-Aha-F^{lac} spectrum (Figure 1C, indicated by asterisk) was assigned to the full-length peptide with the amino group, and was presumably generated by the dithiothreitol reduction of the azide side chain.[37] The peptide expression directed by the reprogrammed codons thus successfully yielded the expected full-length peptide.

Subsequent TCEP reduction of the masking group triggered the self-cyclization of the Xaa residue, and released the F^{lac}-R dipeptide. MALDI-TOF analysis of the product generated from the P1-Aha-F^{lac} peptide gave a single major peak that was consistent with the expected MS value for P1-gLa (Figure 1C). Similarly, TCEP reduction of P1-(ME)Hcy-F^{lac} afforded the corresponding thiolactone, P1-gTa. The same procedure was also employed for P1-Anv-F^{lac} and P1-(ME)Mnv-F^{lac} to yield the corresponding C-terminal δ-lactam and thiolactone peptides (Figure S2 in the Supporting Information). These data clearly demonstrate that the C-terminal rearrangement from Xaa-F^{lac} to the γ/δ-lactam or thiolactone cleanly took place in these model peptides expressed with the reprogrammed codons.

To verify the versatility of this methodology, we translated another mRNA (mR2) encoding a longer and more complex peptide sequence: P2-Xaa-F^{lac} (Figure 2A). For this 16-mer peptide, mR2 was translated in the presence of [^{14}C]-Asp, so that the expression of the corresponding radiolabeled product could be monitored by appropriate PAGE analysis (Figure S3 in the Supporting Information). Notably, high expression levels were observed only in the presence of both Xaa-$tRNA^{Asn\text{-}E1}_{GGU}$ and F^{lac}-$tRNA^{Asn\text{-}E1}_{GAG}$, and the observed expression level of P2-Xaa-F^{lac} fell to 33–58% of that of wild-type P2-Thr-Leu, the quantity of which was determined to be approximately 3 pmol μL^{-1}. As in the case of the P1-Xaa-F^{lac}, C-terminal γ-lactam and γ-thiolactone formation of P2-Xaa-F^{lac} was triggered by TCEP reduction. Because the purification tag was not implanted in this set of peptides, the product was simply desalted by using a resin-modified micropipette tip. Even after such a simple workup, MALDI-TOF analysis both of

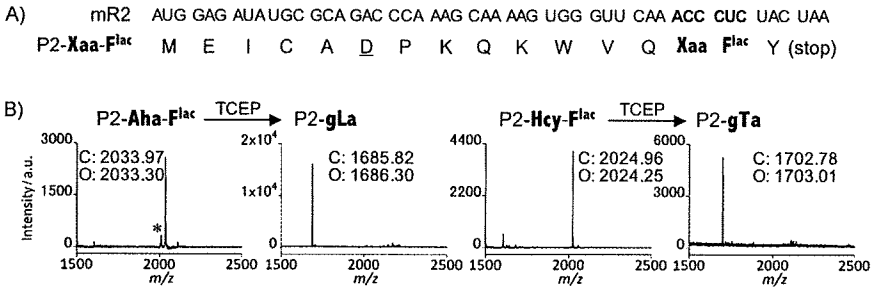


Figure 2. Ribosomal synthesis of the C-terminal γ/δ-lactam- and γ/δ-thiolactone-containing peptides, P2-gLa and P2-gTa, respectively. A) Sequences of mR2 and P2-Xaa-F^{lac}. Xaa and F^{lac} are assigned to codons ACC and CUC, respectively. B) MALDI-TOF analysis of the peptides. Calculated (C) and observed (O) molecular masses for the singly charged species, $[M+H]^+$, of the peptide are shown in each spectrum. TCEP reduction of P2-Aha-F^{lac} and P2-(ME)Hcy-F^{lac} resulted in the formation of γ-lactam and γ-thiolactone, respectively, at the peptide C termini. The asterisk (*) denotes a peak of P2-Aha-F^{lac} in which Aha was presumably reduced by DTT present in the translation mixture; calcd: 2008.02, found: 2008.54.

the precursor peptide (Figure 2B, P2-Aha-F^{lac} or P2-Hcy-F^{lac}) and of the peptide containing C-terminal γ-lactam or γ-thiolactone moieties (Figure 2B, P2-gLa and P2-gTa) showed clean single major peaks consistent with the expected MS values. Likewise, expression and TCEP conversion of the peptides into the corresponding C-terminal δ-lactam and δ-thiolactone peptides was also successfully executed (Figure S4 in the Supporting Information, P2-dLa and P2-dTa).

Synthesis of C-terminally modified peptides and of a macrocyclized peptide with the γ-thiolactones

Because a C-terminal thiolactone might be reactive toward amine nucleophiles,[38,39] we were interested in converting it into various C-terminal amide groups (Figure 3A). P1-gTa was accordingly treated with thirteen different amines, from simple ammonium hydroxide to a fluorescent-labeled alkylamine (Figure 3B, **1**–**13**). On addition of an excess amount of the individual amine to the P1-gTa, the desired C-terminal amidation of P1-gTa occurred cleanly, and occasionally along with oxidation of the liberated sulfhydryl group with ME (Figure 3C). It should be noted that ME was one of the standard reagents included both in the ordinary PURE and in the *w*PURE systems, and such ME adducts were presumably generated by oxidation by air during the amidation process. In addition, P2-gTa was also subjected to C-terminal amidation with the three chosen amines, **1**, **9**, and **13**, to yield the desired C-amide peptides.

Lastly, we attempted macrocyclization of a peptide, through linking between the C terminus and an amino group in a lysine (K) side chain from the corresponding γ-thiolactone peptide intermediate. We designed a peptide, P3-Hcy-F^{lac}, containing a single K residue upstream of Hcy-F^{lac}, and expressed it from the parent mRNA (mR3; Figure 4A). The expressed peptide was treated with TCEP to induce the C-terminal γ-thiolactone formation, and this was followed by incubation with ME under basic conditions to trigger the desired macrocyclization (Figure 4B). MALDI-TOF analyses both of the intermediate peptide (P3-gTa) and of the macrocyclic peptide (mcP3-(ME)Hcy) confirmed the desired conversion (Figure 4C). This demonstrates that the amidation methodology developed here is applicable to both intermolecular and intramolecular reactions, and has enabled us to prepare a wide array of C-terminal peptide-alkylamides.

Conclusions

We have developed a new methodology for the ribosomal synthesis of peptides incorporating C-terminal lactam, thiolactone, and alkylamide moieties. The key technical development involves the ribosomal incorporation of a nonproteinogenic amino acid (Xaa), in which the amino or thio nucleophile in the side chain is masked (e.g., Aha or (ME)Hcy), followed by F^{lac}, at the C-terminal region of a peptide by means of genetic code reprogramming. TCEP reduction of the masking group exposes the amino or sulfhydryl nucleophile, and triggers C-terminal γ-lactam (gLa) or thiolactone (gTa) formation. Moreover, the C-terminal γ-thiolactone can be amidated by treatment with various alkylamines. Intramolecular amidation between a lysine residue and the C-terminal gTa yields a macrocyclic peptide. Since such a modification confers biostability on the peptide, the bioactivity of the peptide can be enhanced.[6–8] Most importantly, the ribosomal synthesis of the peptides is mRNA encoded, which facilitates the preparation of peptide libraries.[29] Because C-terminal amide modification of linear and cyclic peptides should enhance their physiological stability, this methodology gives us an opportunity to explore the sequence space of such peptides and could potentially yield new peptidic drugs in the future.

Experimental Section

Preparation of aminoacyl-tRNAs with flexizymes: Aminoacylation reactions were performed by the following procedure: $tRNA^{Asn\text{-}E1}$ (40 μM) in Tris-HCl (0.2 M, pH 8.0, 6.25 μL) was heated at 95 °C for 1 min, and the system was then allowed to cool to room temperature over 5 min. $MgCl_2$ (3 M, 2.5 μL) and dFx (200 μM, 1.25 μL) were then added and the solution was incubated at room temperature for 5 min. Substrate in DMSO (25 mM, 2.5 μL) was then added and the mixture was incubated on ice for 2 h. The acylation reaction was quenched by addition of sodium acetate (0.6 M, pH 5.0, 40 μL), and the RNA was recovered by ethanol precipitation. The pellet was rinsed twice with ethanol (70%) containing sodium acetate (0.1 M, pH 5.0), and once with ethanol (70%). The pellet was dried and stored at −80 °C.

Ribosomal synthesis of wild-type peptides and of peptides containing nonproteinogenic amino acids/hydroxy acids: The ordinary PURE system was reconstituted with ribosome, 20 aminoacyl RNA synthetases, protein factors, necessary organic and inorganic components, and the 20 proteinogenic amino acids (0.2 mM each); for tricine SDS-PAGE analysis, [^{14}C]-Asp (50 μM) was used in place of Asp, as reported elsewhere.[29] Control wild-type peptide was translated with this PURE system in the presence of DNA template (40 nM) at 37 °C for 1 h, in a total volume of 5 μL. The *w*PURE system was reconstituted with the same components as the PURE system, except that only the necessary proteinogenic amino acids (0.2 mM each) and acyl-tRNAs (50 μM each) were added to the mixture, depending upon the peptide being translated. For the synthesis of P1, which contained the nonproteinogenic amino acids Aha, Anv, (ME)Hcy or (ME)Mnv, and phenyllactic acid (F^{lac}), the *w*PURE system in the presence of DNA template D1 (40 nM), Met, Arg, Tyr, Lys, and Asp (0.2 mM each), and Aha-$tRNA^{Asn\text{-}E1}_{GGU}$, Anv-$tRNA^{Asn\text{-}E1}_{GGU}$, (ME)Hcy-$tRNA^{Asn\text{-}E1}_{GGU}$, or (ME)Mnv-$tRNA^{Asn\text{-}E1}_{GGU}$ (50 μM each), and F^{lac}-$tRNA^{Asn\text{-}E1}_{GAG}$ (50 μM) were used. For the synthesis of P2, which contained the nonproteinogenic amino acids Aha, Anv, (ME)Hcy, or (ME)Mnv, and phenyllactic acid (F^{lac}), the above *w*PURE system in the presence of DNA template D2 (40 nM), Met, Glu, Ile, Cys, Ala, Pro, Lys, Gln, Trp, Val, Arg, Tyr, Asp (0.2 mM each; Asp was replaced with 50 μM [^{14}C]-Asp for tricine SDS-PAGE analysis), and Aha-$tRNA^{Asn\text{-}E1}_{GGU}$ or Anv-$tRNA^{Asn\text{-}E1}_{GGU}$, (ME)Hcy-$tRNA^{Asn\text{-}E1}_{GGU}$, or (ME)Mnv-$tRNA^{Asn\text{-}E1}_{GGU}$ (50 μM each), and F^{lac}-$tRNA^{Asn\text{-}E1}_{GAG}$ (50 μM) were used. For tricine SDS-PAGE analysis the reaction mixture (2.5 μL) was mixed with loading buffer (2.5 μL; 0.9 M Tris-HCl, pH 8.5, 30% glycerol, and 8% SDS) and analyzed by 15% tricine SDS-PAGE.

Mass spectra measurements of peptides: For the mass analysis a sample (5 μL) of the reaction mixture was mixed with 2×TBS (5 μL; 100 mM Tris-HCl, pH 8.0, 300 mM NaCl) and incubated in prewashed FLAG-M2 agarose (Sigma) for 1 h. The resin was then washed once with TBS (50 mM Tris-HCl, pH 8.0, 150 mM NaCl; 10 μL), and

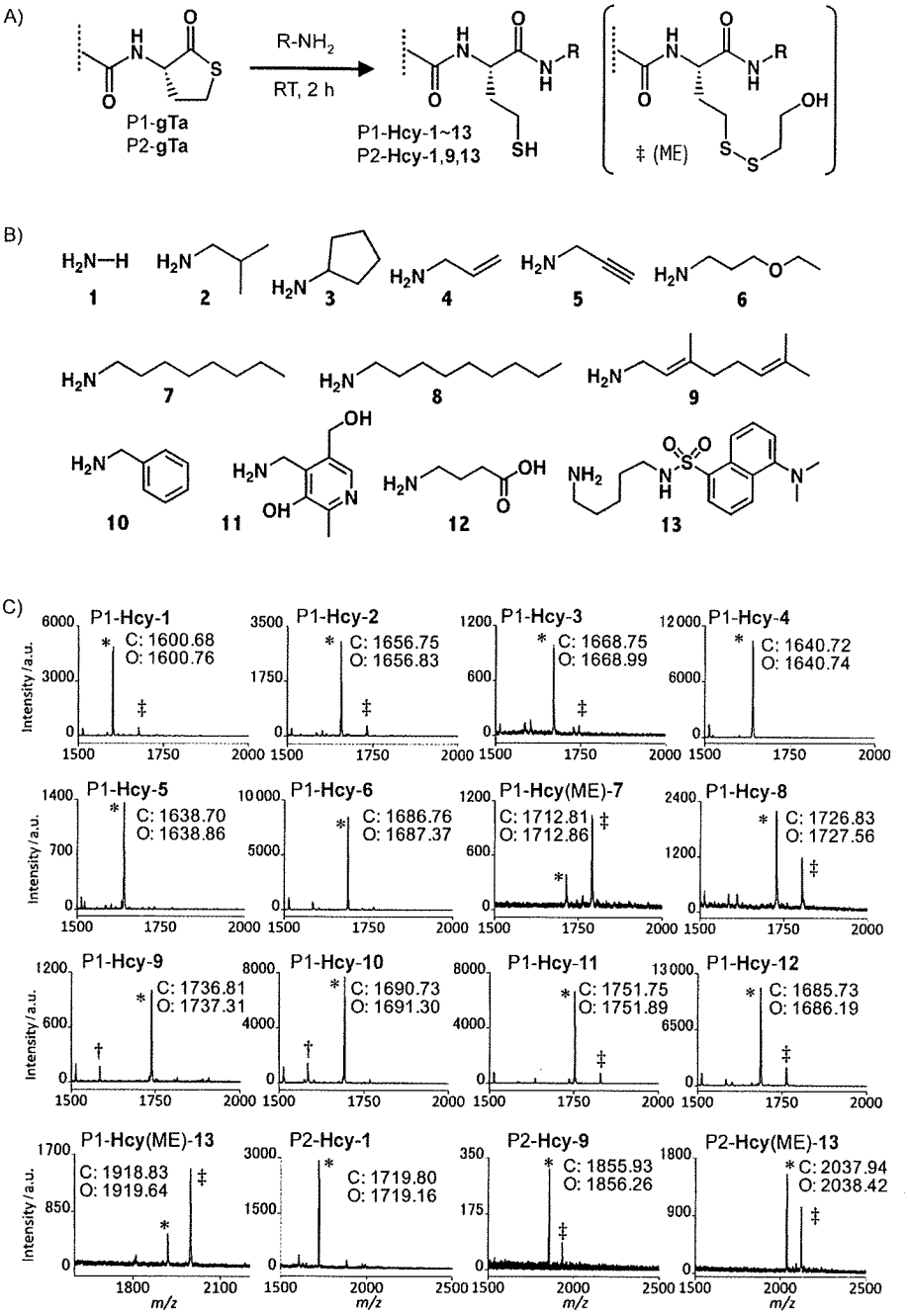


Figure 3. Amidation of P1-gTa and P2-gTa with a variety of amines. A) Schematic representation of C-terminal amidation with γ-thiolactones. P1-gTa and P2-gTa were incubated with amines that opened their γ-thiolactone moieties to yield the alkylamides. The products in parentheses represent the oxidation of the sulfhydryl group of Hcy with ME. B) Primary amines used in this study: ammonia (**1**), isobutylamine (**2**), cyclopentylamine (**3**), allylamine (**4**), propargylamine (**5**), 3-ethoxypropylamine (**6**), octylamine (**7**), nonylamine (**8**), geranylamine (**9**), benzylamine (**10**), pyridoxine (**11**), 4-aminobutanoic acid (**12**), and dansylcadaverine (**13**). The C terminus of P1-gTa was subjected to amidation with all of the amines, whereas that of P2-gTa was subjected to **1**, **9**, or **13**. C) MALDI-TOF MS spectra of the C-terminally modified P1 and P2. Calculated (C) and observed (O) molecular masses for the singly charged species, $[M+H]^+$, of the peptide are shown in each spectrum. Asterisks (*) indicate peaks corresponding to the peptide modified with each amine at the C terminus. Daggers (†) indicate unreacted P1-gTa, and double daggers (‡) indicate products with sulfhydryl groups oxidized with ME, in which the molecular mass of P1-Hcy-X (X = **1**–**13** is increased by approximately 76.

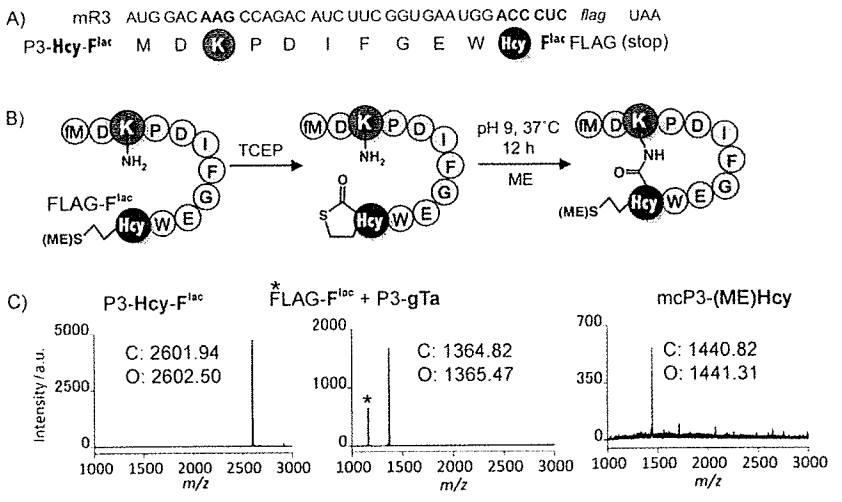


Figure 4. Macrocyclization of P3 with γ-thiolactone. A) Sequences of mR3 and P3-Xaa-F^{lac}. (ME)Hcy and F^{lac} are assigned to codons ACC and CUC, respectively. B) Schematic representation of the macrocyclization with P3-gTa. TCEP reduction of P3-Hcy-F^{lac} induced the formation of P3-gTa, which was incubated with ME (10 mM) at pH 9.0, 37 °C for 12 h. These procedures afforded mcP3-(ME)Hcy through linkage between the C terminus and the side chain of the K residue. C) MALDI-TOF analysis of the corresponding peptides. Calculated (C) and observed (O) molecular masses for the singly charged species, $[M+H]^+$, of the peptide are shown in each spectrum. The asterisk (*) denotes a peak corresponding to liberated F^{lac}-FLAG from P3-Hcy-F^{lac}; calcd: 1162.13, found: 1163.44.

the immobilized peptide was eluted with TFA (0.2%, 5 μL). The resulting peptide was desalted with a C18 micro-ZipTip (Millipore), and eluted with aq. acetonitrile (50%, 1 μL, 0.1% TFA), saturated with the matrix (*R*)-cyano-4-hydroxycinnamic acid. Mass measurements were performed by MALDI-TOF (autoflex TOF/TOF, Bruker).

Synthesis of C-terminal γ/δ-lactams or thiolactones: For the C-terminally cyclized peptides, translation product (5 μL) was mixed with TCEP (100 mM, 0.5 μL) and bicine-KOH (1 M, pH 9.0, 0.5 μL) at room temperature for 2 h. For the confirmation of cyclization, the mixture was purified on FLAG resin, eluted with TFA (0.2%, 5 μL), and analyzed by MALDI-TOF.

C-terminal modification with various alkylamides: For C-terminal modification with various alkylamides, the C-terminal γ-thiolactone peptide was prepared as described above, in 5 μL total volume. The mixture was incubated with alkylamine (2 M, 0.5 μL, ammonia, isobutylamine, cyclopentylamine, allylamine, propargylamine, 3-ethoxypropylamine, octylamine, nonylamine, geranylamine, benzylamine, pyridoxine, 4-aminobutanoic acid, or dansylcadaverine, as shown in Figure 3B) at 25°C for 2 h. These alkylamines were purchased from Kanto Chemicals (Tokyo, Japan), Sigma–Aldrich (Japan), or Nacalai Tesque (Kyoto, Japan). For the confirmation of modification, the mixture was purified on FLAG resin, eluted with TFA (0.2%, 2 μL), and analyzed by MALDI-TOF. Note that ME (1 mM) was included in the *w*PURE system as one of the standard reagents, and occasionally formed a disulfide bond with the free sulfhydryl group in residues such as Hcy and Cys by air oxidation.

Peptide macrocyclization: For the macrocyclization of the peptide, the C-terminal γ-thiolactone peptide was synthesized as described above in a total volume of 5 μL, and purified on FLAG resin and eluted with TFA (0.2%, 5 μL). The eluent was incubated with 2-mercaptoethanol (100 mM, 0.5 μL) and bicine-KOH (1 M, pH 9.0, 0.5 μL) at 37 °C for 12 h to induce the macrocyclization between the C terminus and the lysine side chain.

Acknowledgements

This work was supported by Technology Program grants from the New Energy and Industrial Technology Development Organization (NEDO) and the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) to H.S., and from the Global COE Program "Chemistry Innovation through Cooperation of Science and Engineering", MEXT, Japan to E.N.

Keywords: genetic code · peptidomimics · protein engineering · ribosomal synthesis · ribozymes

[1] L. L. Kisselev, R. H. Buckingham, *Trends Biochem. Sci.* **2000**, *25*, 561.
[2] S. Petry, A. Weixlbaumer, V. Ramakrishnan, *Curr. Opin. Struct. Biol.* **2008**, *18*, 70.
[3] D. J. Merkler, *Enzyme Microb. Technol.* **1994**, *16*, 450.
[4] A.-J. Ren, Z.-F. Guo, Y.-K. Wang, L. Lin, X. Zheng, W.-J. Yuan, *Peptides* **2009**, *30*, 439.
[5] V. Sanchez-Margalet, M. Lucas, R. Goberna, *J. Mol. Endocrinol.* **1996**, *16*, 1.
[6] L. Hilsted, K. Hint, J. Christiansen, J. F. Rehfeld, *Gastroenterology* **1988**, *94*, 96.
[7] K. R. Shoemaker, P. S. Kim, D. N. Brems, S. Marqusee, E. J. York, I. M. Chaiken, J. M. Stewart, R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 2349.
[8] A. Wettergren, L. Pridal, M. Wojdemann, J. J. Holst, *Regul. Pept.* **1998**, *77*, 83.
[9] A. F. Bradbury, D. G. Smyth, *Biosci. Rep.* **1987**, *7*, 907.
[10] I. R. Cottingham, A. Millar, E. Emslie, A. Colman, A. E. Schnieke, C. McKee, *Nat. Biotechnol.* **2001**, *19*, 974.
[11] T. J. Tolbert, C.-H. Wong, *J. Am. Chem. Soc.* **2000**, *122*, 5421.
[12] J. Mukhopadhyay, A. N. Kapanidis, V. Mekler, E. Kortkhonjia, Y. W. Ebright, R. H. Ebright, *Cell* **2001**, *106*, 453.
[13] B. Zhang, Z. Tan, L. G. Dickson, M. N. Nalam, V. W. Cornish, A. C. Forster, *J. Am. Chem. Soc.* **2007**, *129*, 11316.
[14] Z. Tan, A. C. Forster, S. C. Blacklow, V. W. Cornish, *J. Am. Chem. Soc.* **2004**, *126*, 12752.
[15] Z. Tan, S. C. Blacklow, V. W. Cornish, A. C. Forster, *Methods* **2005**, *36*, 279.
[16] A. O. Subtelny, M. C. Hartman, J. W. Szostak, *J. Am. Chem. Soc.* **2008**, *130*, 6131.
[17] C. Merryman, R. Green, *Chem. Biol.* **2004**, *11*, 575.
[18] K. Josephson, M. C. Hartman, J. W. Szostak, *J. Am. Chem. Soc.* **2005**, *127*, 11727.
[19] M. C. Hartman, K. Josephson, C. W. Lin, J. W. Szostak, *PLoS ONE* **2007**, *2*, e972.
[20] A. Frankel, S. W. Millward, R. W. Roberts, *Chem. Biol.* **2003**, *10*, 1043.
[21] A. C. Forster, H. Weissbach, S. C. Blacklow, *Anal. Biochem.* **2001**, *297*, 60.
[22] A. C. Forster, Z. Tan, M. N. Nalam, H. Lin, H. Qu, V. W. Cornish, S. C. Blacklow, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 6353.
[23] Y. Shimizu, Y. Kuruma, B. W. Ying, S. Umekage, T. Ueda, *FEBS J.* **2006**, *273*, 4133.
[24] Y. Shimizu, T. Kanamori, T. Ueda, *Methods* **2005**, *36*, 299.
[25] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, *Nat. Biotechnol.* **2001**, *19*, 751.

[26] H. Murakami, A. Ohta, H. Ashigai, H. Suga, *Nat. Methods* **2006**, *3*, 357.
[27] A. Ohta, H. Murakami, E. Higashimura, H. Suga, *Chem. Biol.* **2007**, *14*, 1315.
[28] Y. Goto, H. Murakami, H. Suga, *RNA* **2008**, *14*, 1390.
[29] Y. Goto, A. Ohta, Y. Sako, Y. Yamagishi, H. Murakami, H. Suga, *ACS Chem. Biol.* **2008**, *3*, 120.
[30] T. J. Kang, H. Suga, *Biochem. Cell Biol.* **2008**, *86*, 92.
[31] T. J. Kang, S. Yuzawa, H. Suga, *Chem. Biol.* **2008**, *15*, 1166.
[32] T. Kawakami, H. Murakami, H. Suga, *J. Am. Chem. Soc.* **2008**, *130*, 16861.
[33] A. Ohta, H. Murakami, H. Suga, *ChemBioChem* **2008**, *9*, 2773.
[34] Y. Sako, Y. Goto, H. Murakami, H. Suga, *ACS Chem. Biol.* **2008**, *3*, 241.
[35] Y. Sako, J. Morimoto, H. Murakami, H. Suga, *J. Am. Chem. Soc.* **2008**, *130*, 7232.
[36] T. Kawakami, H. Murakami, H. Suga, *Chem. Biol.* **2008**, *15*, 32.
[37] L. Cartwright, D. W. Hutchinson, V. W. Armstrong, *Nucleic Acids Res.* **1976**, *3*, 2331.
[38] C. E. Hop, R. Bakhtiar, *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1049.
[39] J. Perła-Kaján, T. Twardowski, H. Jakubowski, *Amino Acids* **2007**, *32*, 561.

Received: February 4, 2009
Published online on April 15, 2009

J|A|C|S COMMUNICATIONS

Published on Web 03/20/2009

Translation Initiation with Initiator tRNA Charged with Exotic Peptides

Yuki Goto[†,‡] and Hiroaki Suga[*,†,‡]

Research Center for Advanced Science and Technology and Department of Advanced Interdisciplinary Studies, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8904, Japan

Received January 24, 2009; E-mail: hsuga@rcast.u-tokyo.ac.jp

The translation system evolved to polymerize 20 specific kinds of proteinogenic L-α-amino acids with extremely high accuracy according to the sequence information encoded in mRNA. Exclusion of nonproteinogenic amino acids from the polymerization is achieved by sophisticated mechanisms involving the multistep selection of correctly charged aminoacyl-tRNAs.[1] Despite the fact that techniques allowing researchers to manipulate the genetic code, so-called genetic code expansion[2] or genetic code reprogramming,[3] have been developed and many successes in incorporating non-proteinogenic L-α-amino acids into nascent peptide chain have been reported, nonproteinogenic amino acids containing more drastically altered structures (referred to as exotic amino acids), such as D- and β-amino acids, are notoriously difficult or often impossible to be elongated.[2a−c,3c,4] On the other hand, recent investigations have shown that the initiation apparatus is more tolerant to exotic amino acids than elongation event. In fact, it has been reported that "nonmethionine" amino acids, for example, various L-α-amino acids[5a] and some D-α-amino acids[5b] with or without *N*-acyl groups, were able to initiate translation, whereas formylmethionine (fMet) is a sole initiator in the ordinary prokaryotic translation system (Figure 1A).[6] The observed tolerance of initiation has prompted

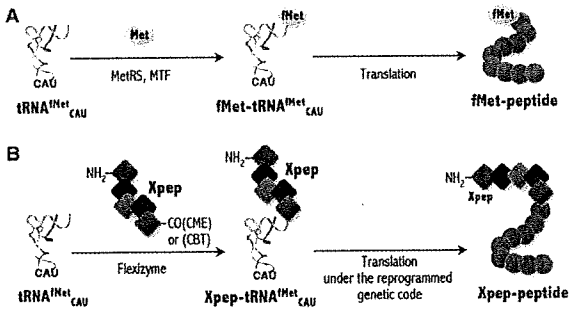


Figure 1. Reprogramming of the initiation event with exotic peptides. (A) Initiation in an ordinary prokaryotic translation system. The initiator tRNA (tRNA$^{fMet}_{CAU}$) is methionylated by methionyl-tRNA synthetase (MetRS) and its α-amino group is formylated by methionyl-tRNA formyltransferase (MTF) to afford fMet-tRNA$^{fMet}_{CAU}$. Initiation with fMet-tRNA$^{fMet}_{CAU}$ yields a peptide containing fMet at the N-terminus (fMet-peptide). (B) Outline of translation initiation with exotic peptides reported here. Flexizyme charges a short exotic peptide (Xpep) onto the tRNA$^{fMet}_{CAU}$. Reprogramming of the initiation event allows us to prime the translation reaction with Xpep-tRNA$^{fMet}_{CAU}$, yielding a peptide bearing several exotic amino acids at the N-terminus (Xpep-peptide).

us to further explore the repertoire expansion. Here we report the reprogramming of translation initiation with "exotic peptides" containing several exotic amino acids (Figure 1B).

To facilitate the reprogramming of initiation event, we utilized our original methodology in the combination of flexizyme and *w*PURE systems. The former system consists of a tRNA aminoacylation ribozyme, referred to as flexizyme, enabling us to charge a wide variety of nonproteinogenic amino acids activated with certain ester or thioester groups onto the 3′ hydroxyl group of any desired tRNAs, including the initiator tRNA$^{fMet}_{CAU}$.[3c] The latter system is based on a reconstituted *E. coli* cell-free translation system, so-called protein-translation using recombinant elements (PURE),[7] in which methionine (Met) is *with-drawn* to make the start codon vacant and thus it is called a *w*PURE system.[5,8] By using the integrated systems, that is, by adding an acylated tRNA$^{fMet}_{CAU}$ prepared by flexizyme to the *w*PURE system, the start codon AUG was reassigned from fMet to desirable non-fMet initiators.[5]

Although the flexizyme system has provided us a nearly unlimited opportunity for the synthesis of tRNA charged with nonproteinogenic amino acids (Xaa-tRNA),[3c,5,8] it is unknown if it can be applicable to tRNA peptidylation. We therefore first verified whether exotic peptides (Xpep) could be charged onto tRNA$^{fMet}_{CAU}$ by means of flexizyme. We designed eleven short peptides containing various combinations of proteinogenic amino acids and exotic amino acids including D-phenylalanine (^{D}Phe), D-glutamic acid (^{D}Glu), D-lysine (^{D}Lys), *N*-methyl-L-phenylalanine (^{Me}Phe), *N*-methyl-L-glutamic acid (^{Me}Glu), 4-aminobenzoic acid (^{4}Abz) and β-alanine (βAla) (the individual structures are shown in Supporting Information, Figure S1). To make the peptides to be accessible to the flexizyme system, those containing the C-terminal ^{L}Phe, ^{D}Phe, ^{Me}Phe, or ^{L}Met were derived to the cyanomethyl esters (CME) (Table 1, entries 1−8, 10, and 11), while one containing the C-terminal ^{L}Gln was derived to 4-cholorobenzyl thioester (CBT)

Table 1. The Sequences of Peptide Initiators Used in This Study and Their Acylation Yields and Translation Efficiencies

entry	initiator sequence[a]	activating group for flexizyme[b]	acylation yield (%)[c]	translation efficiency (%)[d]
1	^{D}Phe−^{L}Phe	CME	85	88
2	^{D}Phe−^{D}Phe−^{L}Phe	CME	44	68
3	^{D}Phe−^{4}Abz−^{L}Phe	CME	30	133
4	^{D}Phe−βAla−^{L}Phe	CME	69	115
5	^{D}Phe−^{Me}Glu−^{L}Phe	CME	47	39
6	^{D}Phe−βAla−^{D}Phe	CME	68	26
7	^{D}Phe−βAla−^{Me}Phe	CME	66	88
8	^{D}Phe−βAla−^{L}Met	CME	48	91
9	^{D}Phe−βAla−^{L}Gln	CBT	38	19
10	^{D}Glu−^{D}Lys−^{L}Phe	CME	55	25
11	^{D}Glu−^{D}Lys−^{D}Glu−^{D}Lys−^{L}Phe	CME	55	12

[a] ^{D}Xaa, ^{L}Xaa, ^{Me}Xaa, ^{4}Abz, and βAla denote D-amino acids, L-amino acids, *N*-methyl-*L*-amino acids, 4-aminobenzoic acid, and β-alanine, respectively. [b] CME and CBT denote cyanomethyl ester and 4-chlorobenzyl thioester activating groups, respectively. [c] Yields of the acylation were calculated based on the band intensity in acid PAGE (see Figure S2 for the detailed description). [d] Relative translation efficiencies of Xpep-peptides were determined by comparing with the band intensity of fMet-peptide in tricine-SDS PAGE (see Figure 2B). The expression quantity of the fMet-peptide was determined to be 7.5 pmol/μL based on the method reported elsewhere.[5a,8b]

† Research Center for Advanced Science and Technology.
‡ Department of Advanced Interdisciplinary Studies.

10.1021/ja900597d CCC: $40.75

(Table 1, entry 9). For verification of the peptidylation ability of flexizyme, we used our conventional assay system using a tRNA analogue, microhelix RNA.[3c] The di-, tri-, and pentapeptide substrates were incubated with the microhelix RNA in the presence of the flexizyme, and the products were analyzed by denaturing acid polyacrylamide gel electrophoresis (PAGE). In all cases, a single mobility-shifted band corresponding to the individual Xpep-RNA was observed (Figure S2), indicating that the flexizyme system is compatible to the peptidyl substrates for tRNA peptidylation. Importantly, the observed yields were in the range of 30−69% (Table 1), which were sufficient to carry out the translation experiment based on our previous experimental knowledge.[3c] We thus next investigated whether the translation apparatus could accept these Xpep-tRNA$^{fMet}_{CAU}$ as an initiator.

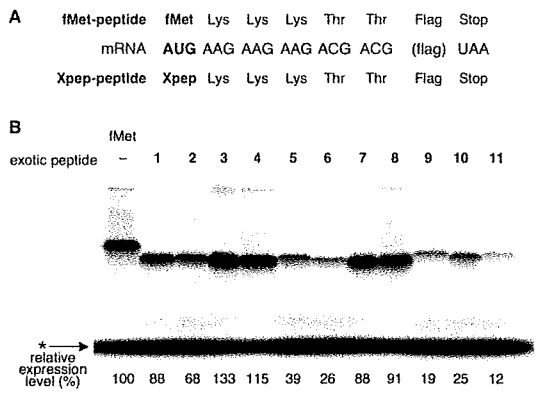


Figure 2. Translation initiation with exotic peptides. (A) Sequences of a mRNA template and expressed fMet/Xpep-peptides. Flag in the parentheses indicates the RNA sequence encoding a Flag peptide (DYKDDDDK). (B) Tricine-SDS PAGE of the translation products initiated with various exotic peptides. The product of the *w*PURE translation reaction containing each Xpep-tRNA$^{fMet}_{CAU}$ and [^{14}C]-Asp was analyzed by tricine-SDS PAGE (lanes 1−11 corresponding the exotic peptides shown in Figure S1). "fMet" indicates the wild-type expression in the ordinary PURE system in which fMet acted as the initiator. Each expression level relative to the fMet-peptide is determined by a mean score of triplicates or more. The band indicated by the asterisk corresponds to the remaining [^{14}C]-Asp that was not incorporated into the peptide.

We designed a mRNA template that encoded a 14-mer peptide, in which the C-terminus contained a Flag peptide sequence for the convenience of radiolabeling-detection using [^{14}C]-Asp as well as the peptide isolation (Figure 2A). Each peptide was expressed in the presence of the respective Xpep-tRNA$^{fMet}_{CAU}$ in the *w*PURE system, where the initiator was reassigned to the short exotic peptide. The translation product was analyzed by tricine-SDS PAGE (Figure 2B) and its relative translation efficiency was determined by the comparison with the wild-type expression where fMet acted as the initiator in the ordinary PURE system (Table 1). In addition, the expressed peptide was subjected to mass spectrometry for identification of the translation product (Figure S3). To our surprise, all Xpep-tRNAs could initiate the translation. The expression efficiencies were depending upon the peptides, ranging from 12% to 133% relative to the fMet-peptide expression. Although we could not define a general trend that correlated the sequence compositions to the observed expression levels, the longest peptide **11** was least efficient so that the long length of peptide might hamper the initiation event. Given the lower efficiency in relatively hydrophilic peptides (**5**, **9**, **10**, and **11**) and the previous observation that hydrophilic amino acids could initiate translation less efficiently than hydrophobic ones,[5] the hydrophilicity of initiator might also affect the initiation efficiency. Nonetheless, it is clear that exotic peptides containing consecutive D-amino acids or the combination of exotic amino acids charged onto tRNA$^{fMet}_{CAU}$ could initiate the translation to afford peptides containing a stretch of exotic amino acids at the N-terminus.

Here we have reported that ribosome is able to accept tRNA$^{fMet}_{CAU}$charged with nonstandard peptides as an initiator, and thus peptides containing a variety of exotic amino acids at the N-terminus could be expressed. To the best of our knowledge, this work represents the first demonstration of translation initiation with peptidyl molecules regardless of natural or non-natural building blocks. In contrast to the fact that elongation of exotic amino acids has often encountered difficulties, the initiation reprogramming approach enables us to synthesize peptides containing exotic amino acids at the N-terminus. Particularly, we envisage that the integration of this approach with appropriate peptide cyclization techniques[5a,8c] will open a new avenue for the ribosomal synthesis of unusual peptides. Moreover, the reported method facilitating the synthesis of desired peptidyl-tRNAs with proteinogenic or exotic amino acids by means of flexizyme system would provide a new tool for studying the mechanism of initiation and/or elongation in translation.

Acknowledgment. We thank Dr. Hiroshi Murakami for valuable discussion and suggestion. We thank Professor Kazuhiko Nakatani for the use of ESI-TOF instrumentation. We also thank Mr. Yusuke Yamagishi for assistance with manuscript preparation. 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 (18-10526) to Y.G., 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: Detailed structure of short peptides used here, acid PAGE analysis, mass spectra, and experimental methods. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) (a) Söll, D. *Experientia* **1990**, *46*, 1089–1096. (b) Sankaranarayanan, R.; Moras, D. *Acta Biochim. Pol.* **2001**, *48*, 323–335.
(2) (a) Roesser, J. R.; Xu, C.; Payne, R. C.; Surratt, C. K.; Hecht, S. M. *Biochemistry* **1989**, *28*, 5185–5195. (b) Bain, J. D.; Diala, E. S.; Glabe, C. G.; Wacker, D. A.; Lyttle, M. H.; Dix, T. A.; Chamberlin, A. R. *Biochemistry* **1991**, *30*, 5411–5421. (c) Ellman, J. A.; Mendel, D.; Schultz, P. G. *Science* **1992**, *255*, 197–200. (d) Link, A. J.; Mock, M. L.; Tirrell, D. A. *Curr. Opin. Biotechnol.* **2003**, *14*, 603–609. (e) Wang, L.; Xie, J.; Schultz, P. G. *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225–249.
(3) (a) Forster, A. C.; Tan, Z.; Nalam, M. N.; Lin, H.; Qu, H.; Cornish, V. W.; Blacklow, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 6353–6357. (b) Josephson, K.; Hartman, M. C.; Szostak, J. W. *J. Am. Chem. Soc.* **2005**, *127*, 11727–11735. (c) Murakami, H.; Ohta, A.; Ashigai, H.; Suga, H. *Nat. Methods* **2006**, *3*, 357–359.
(4) (a) Starck, S. R.; Qi, X.; Olsen, B. N.; Roberts, R. W. *J. Am. Chem. Soc.* **2003**, *125*, 8090–8091. (b) Tan, Z.; Forster, A. C.; Blacklow, S. C.; Cornish, V. W. *J. Am. Chem. Soc.* **2004**, *126*, 12752–12753. (c) Dedkova, L. M.; Fahmi, N. E.; Golovine, S. Y.; Hecht, S. M. *Biochemistry* **2006**, *45*, 15541–15551.
(5) (a) Goto, Y.; Ohta, A.; Sako, Y.; Yamagishi, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 120–129. (b) Goto, Y.; Murakami, H.; Suga, H. *RNA* **2008**, *14*, 1390–1398.
(6) (a) Kozak, M. *Microbiol. Rev.* **1983**, *47*, 1–45. (b) Gold, L. *Annu. Rev. Biochem.* **1988**, *57*, 199–233. (c) Gualerzi, C. O.; Pon, C. L. *Biochemistry* **1990**, *29*, 5881–5889.
(7) Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. *Nat. Biotechnol.* **2001**, *19*, 751–755.
(8) (a) Ohta, A.; Murakami, H.; Higashimura, E.; Suga, H. *Chem. Biol.* **2007**, *14*, 1315–1322. (b) Sako, Y.; Goto, Y.; Murakami, H.; Suga, H. *ACS Chem. Biol.* **2008**, *3*, 241–249. (c) Sako, Y.; Morimoto, J.; Murakami, H.; Suga, H. *J. Am. Chem. Soc.* **2008**, *130*, 7232–7234. (d) Kawakami, T.; Murakami, H.; Suga, H. *Chem. Biol.* **2008**, *15*, 32–42.

JA900597D

ARTICLES

nature
structural &
molecular biology

Bases in the anticodon loop of $tRNA^{Ala}_{GGC}$ prevent misreading

Hiroshi Murakami[1], Atsushi Ohta[2] & Hiroaki Suga[1,2]

The bases at positions 32 and 38 in the tRNA anticodon loop are known to have a specific conservation depending upon the anticodon triplets. Here we report that evolutionarily conserved pairs of bases at positions 32 and 38 in $tRNA^{Ala}_{GGC}$ prevent misreading of a near-cognate valine codon, GUC. The $tRNA^{Ala}_{GGC}$ molecules with the conserved A32-U38 and C32-G38 pairs do not read GUC, whereas those with three representative nonconserved pairs, U32-U38, U32-A38 and C32-A38, direct the misincorporation of alanine at this valine codon into the peptide chain. Overexpression of the nonconserved $tRNA^{Ala}_{GGC}$ in *Escherichia coli* is toxic and prevents cell growth. These results suggested that the bases at positions 32 and 38 in $tRNA^{Ala}_{GGC}$ evolved to preserve the fidelity of the cognate codon reading.

Decoding fidelity of translation relies on accurate selection of an aminoacyl-tRNA (aa-tRNA) whose anticodon base-pairs with the cognate codon encoded in the mRNA. It is well known that the third base pair in the codon-anticodon interaction tolerates a wobble pair represented by a U·G interaction, in addition to the canonical Watson-Click base pair, however, this decoding does not generally alter the identity of the amino acid in the peptide, and thus its fidelity is maintained[1]. On the other hand, if a U·G mispair at the first or second base pair occurs, such a codon-anticodon interaction does accompany an amino acid alteration, and therefore this kind of miscoding is generally prohibited. Still, some examples of misreading of the first codon have been reported in literature. For instance, in *Saccharomyces cerevisiae*, amber (UAG) mutations generated by UV irradiation are read by Gln-$tRNA^{Gln}_{CUG}$, causing termination to be suppressed by glutamine incorporation[2–6]. In *E. coli*, mutation of the AGC (serine) codon to GGC (glycine) codon at the catalytic Ser68 residue in β-lactamase is suppressed by the endogenous Ser-$tRNA^{Ser}_{GCU}$, although the probability of misreading resulting in glycine incorporation was estimated to be less than 1 in 1,000 (ref. 7).

It is also known that a base mutation or mutations near the junction of arms in the tRNA cloverleaf structure diminish decoding fidelity. One of the well-known cases is the G24A mutation in the D-stem of $tRNA^{Trp}_{CCA}$, the so-called Hirsh suppressor tRNA, which misreads CGG (arginine) and UGA (see **Fig. 1a** for a reference of the base position in a tRNA structure, $tRNA^{Ala}_{GGC}$)[8,9]. It was recently shown that the Hirsh suppressor $tRNA^{Trp}_{CCA}$ elevates the rates of both GTP hydrolysis and accommodation independently from the codon-anticodon interaction, and thus the misreading described above occurs[9]. These experiments suggest that, remotely, this base in the tRNA body has a crucial role in controlling the decoding event. Similarly, artificial mutations introduced into the C27-G43 Watson-Crick base pair in the anticodon stem of $tRNA^{Trp}_{CUG}$ increased the frequency of misreading of the first position wobble[10,11]. For instance, $tRNA^{Trp}_{CUG}$ bearing the G27-A43 mispair misread the UAG amber codon 40 times more frequently than the wild-type pair. Taken together with other biochemical data, it was postulated that such mutations possibly alter the angle of the junction of the anticodon stem and the central tRNA L-shaped structure, increasing the frequency of wobble reading[10].

Some bases in tRNA anticodon loop are also known to contribute to the maintenance of decoding fidelity. Although a typical example is base modifications in the anticodon loop that disrupt codon recognition[12,13], here we focus on sequence variations in the anticodon loop. For instance, *E. coli* has $tRNA^{Gly}$ with three isoacceptors for GGN (N can be any base) codons, whereas *Mycoplasma mycoides* has only $tRNA^{Gly}_{UCC}$ for reading these codons. It turns out that the difference in the sequence of the anticodon loop between *E. coli* $tRNA^{Gly}_{UCC}$ and *M. mycoides* $tRNA^{Gly}_{UCC}$ is a base at position 32, in which the former has U32 whereas the latter has C32, both pairing with A38. Notably, the U32C mutation introduced into *E. coli* $tRNA^{Gly}_{UCC}$ made it capable of reading all four glycine codons[14,15]. This suggests that the base at position 32 in the anticodon loop influences the tolerance of the U34·U and U34·C mispairs in codon-anticodon recognition. As described earlier, however, this misreading does not accompany an amino acid alteration. Hence, the study described above does not explain the importance of these bases at positions 32 and 38 in decoding fidelity. Nevertheless, this work prompted us to investigate whether the conservation of positions 32 and 38 contributes to the ability of tRNAs to correctly decode cognate codons in *E. coli*.

[1]Research Center for Advanced Science and Technology, University of Tokyo, Meguro-ku, Tokyo, Japan. [2]Department of Chemistry and Biotechnology, Graduate School of Engineering, University of Tokyo, Bunkyo-ku, Tokyo, Japan. Correspondence should be addressed to H.M. (hmura@rcast.u-tokyo.ac.jp) or H.S. (hsuga@rcast.u-tokyo.ac.jp).

Received 29 October 2008; accepted 13 February 2009; published online 22 March 2009; doi:10.1038/nsmb.1580

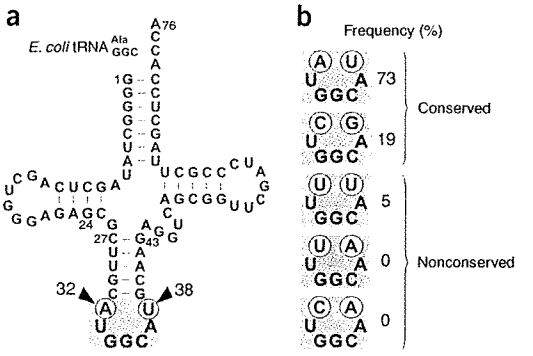


Figure 1 Structure of wild-type $tRNA^{Ala}_{GGC}$ and its variants (**a**) wild-type *E. coli* $tRNA^{Ala}_{GGC}$. (**b**) Frequency in the occurrence of the 32-38 pair in 84 nonredundant sequences of bacterial $tRNA^{Ala}_{GGC}$. The $tRNA^{Ala}_{GGC}$ with the A32-U38 and C32-G38 pairs are referred to as conserved $tRNA^{Ala}_{GGC}$, whereas those with the U32-U38, U32-A38 and C32-A38 pairs are referred to as nonconserved $tRNA^{Ala}_{GGC}$.

RESULTS

An evolutionary bias of the 32-38 pair in $tRNA^{Ala}_{GGC}$

First, we used the tDNA database to look for evolutionary bias in the 32-38 pair. Prior to our study, a 1982 report on the comparison of 42 kinds of *E. coli* and bacterial phage tRNA sequences focusing on the anticodon stem-loop region, proposed that some of the base pairs in the anticodon stem and the bases at positions 37 and 38 might show a preference for certain nucleotides depending upon the base at position 36, which forms the first base pair in the codon-anticodon interaction[16]. This finding has led to 'the extended anticodon hypothesis', which posits that these bases evolved to optimize translation efficiency and, possibly, decoding fidelity. Furthermore, this hypothesis was experimentally verified by suppression of the amber codon by mutant $tRNA^{Trp}_{CUA}$, and of the ochre codon by mutant $tRNA^{Glu}_{UUA}$, demonstrating that tRNAs with an extended anticodon sequence showed the highest suppression efficiency[17–20].

More recently, 5,601 bacterial tRNA sequences were extracted from the tDNA database and used to analyze the statistical conservation of bases at the 32 and 38 positions[21]. Certain tRNAs have a specific subset of combinations that differ from those of other tRNAs. For instance, 99% of $tRNA^{Ala}_{GGC}$ contain either A32-U38 (77%) or C32-G38 (22%), whereas the bases contained in bacterial tRNAs in general have frequencies of 52% for C32-A38, 17% for U32-A38, 11% for U32-U38 and 8% for C32-C38. Notably, $tRNA^{Ala}_{GGC}$ derivatives with nonconserved pairs such as U32-U38 and U32-A38 dissociate from the A site of the *E. coli* ribosome four to ten times more slowly than those containing A32-U38 (ref. 22). Moreover, the U32C mutation of $tRNA^{Gly}_{CCC}$, which is 98% conserved with the U32-A38 pair, increases the affinity of the tRNA not only to the cognate codon but also to the near-cognate codons involving third position mismatches[21]. These results imply that the 32-38 pair influences the affinity of tRNAs in the A site; however, again, these third position mismatches in the codon-anticodon interaction do not alter the amino acid, so it is unclear whether the evolutionary force driving this bias in the 32-38 pair, which depends on the anticodon triplet, arises from the need to control efficiency in translation or decoding fidelity.

We also independently searched the bacterial tDNA database[23] to assess the sequence bias in the 32-38 pair in 84 nonredundant sequences of bacterial $tRNA^{Ala}_{GGC}$. We found a trend similar to that previously described[21] (**Fig. 1**). Note that no bacterial $tRNA^{Ala}_{GGC}$ contains U32-A38 and C32-A38 pairs, whereas archeal $tRNA^{Ala}_{GGC}$ has the U32-A38 pair (18% out of 17 nonredundant sequences) but, again, no C32-A38 pair. Because the above sequence bias in the 32-38 pair possibly determines the translation efficiency of Ala-$tRNA^{Ala}_{GGC}$, we analyzed the difference in translation efficiency of *in vitro* transcripts between *E. coli* $tRNA^{Ala}_{GGC}$ (**Fig. 1a**) with the conserved A32-U38 pair (wild type) or the C32-G38 pair, as well as the nonconserved U32-U38, U32-A38 and C32-A38 pairs (**Fig. 1b**). For simplicity, we refer the former and latter sets of $tRNA^{Ala}_{GGC}$ as conserved and nonconserved $tRNA^{Ala}_{GGC}$, respectively.

No change in the decoding efficiency of the GCC cognate codon

To assess the translation efficiency of each $tRNA^{Ala}_{GGC}$ variant, we used an *E. coli* cell-free translation system that was specially reconstituted for this experiment. In this system, the native tRNAs were entirely substituted with *in vitro* transcripts of four tRNAs ($tRNA^{fMet}_{CAU}$, $tRNA^{Tyr}_{GUA}$, $tRNA^{Asp}_{GUC}$ and $tRNA^{Lys}_{CUU}$; we refer to the mixture of these tRNAs as 'tRNA mix') along with a $tRNA^{Ala}_{GGC}$, referred to as the *w*PURE system (w stands for 'withdrawn'). To validate whether this

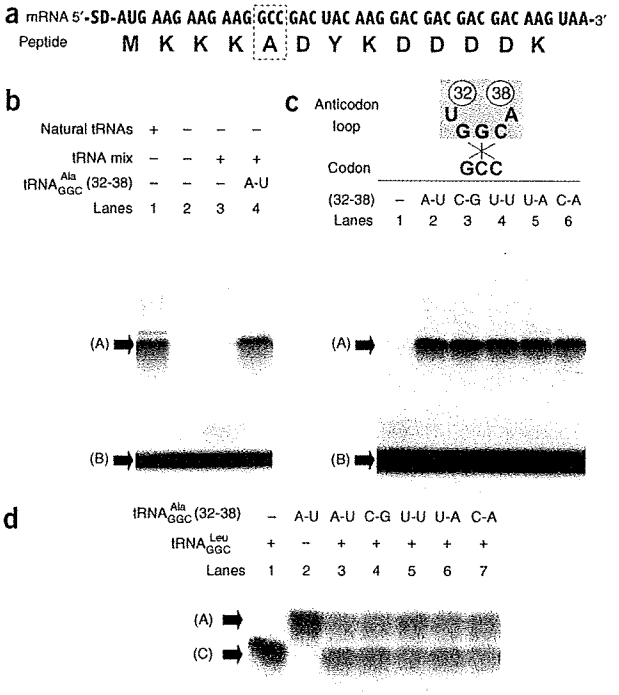


Figure 2 Decoding efficiency of the GCC codon by $tRNA^{Ala}_{GGC}$ with the conserved or nonconserved 32-38 pair. (**a**) Sequences of mRNA and peptide used in this study. The GCC (alanine) codon was placed at the fifth position. (**b**) Tricine SDS-PAGE analysis of the peptide expressed in the presence of tRNA mix and wild-type $tRNA^{Ala}_{GGC}$ in the *w*PURE system. The tRNA mix consists of *in vitro* transcripts of $tRNA^{fMet}_{CAU}$, $tRNA^{Tyr}_{GUA}$, $tRNA^{Asp}_{GUC}$ and $tRNA^{Lys}_{CUU}$. The peptide was expressed at 37 °C for 15 min in the presence of 0.2 mM proteinogenic amino acids (except aspartate) and 50 μM [^{14}C]aspartate. Arrows indicate alanine-containing peptide (A) and [^{14}C]aspartate (B). (**c**) Tricine SDS-PAGE analysis of the peptide in the presence of tRNA mix and each $tRNA^{Ala}_{GGC}$ variant in the *w*PURE system. (**d**) Tricine SDS-PAGE analysis of the competitive decoding of the GCC codon by Ala-$tRNA^{Ala}_{GGC}$ and Leu-$tRNA^{Leu}_{GGC}$ in the *w*PURE system. The competition contained 3 μM $tRNA^{Leu}_{GGC}$ and each $tRNA^{Ala}_{GGC}$ variant to a concentration of 3 μM. Arrows indicate Ala-peptide (A) and Leu-peptide (C).