ACS chemical biology
ARTICLE

Reprogramming the Translation Initiation for the Synthesis of Physiologically Stable Cyclic Peptides

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

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

ABSTRACT The initiation codon dictates that the translation initiation event exclusively begins with methionine. We report here a new technology to reprogram the initiation event, where various amino acids and those bearing N^{α}-acyl groups can be used as an initiator for peptide synthesis. The technology is built upon the concept of genetic code reprogramming, where methionine is depleted from the translation system and the initiation codon is reassigned to the desired amino acid. We have applied this technology to the synthesis of an antitumor cyclic peptide, G7−18NATE, closed by a physiologically stable bond, and it is also extended to the custom synthesis of its analogues with various ring sizes. Significantly, cyclization occurs spontaneously upon translation of the precursor linear peptides. To demonstrate the practicality of this methodology, we also prepared a small cyclic peptide library designated by 160 distinct mRNAs. Thus, this technology offers a new means to prepare a wide array of *in vivo* compatible cyclic peptide libraries for the discovery of peptidic drug candidates against various therapeutic targets.

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

Received for review November 8, 2007
and accepted December 18, 2007.
Published online January 24, 2008
10.1021/cb700233t CCC: $40.75

© 2008 American Chemical Society

Polypeptide synthesis in prokaryotic translation is generally initiated with N^{α}-formyl methionine (fMet) (*1*). The initiation process involves methionylation of the initiator tRNA ($tRNA^{fMet}_{CAU}$) catalyzed by methionyl-tRNA synthetase (MetRS) followed by formylation of the α-amino group on Met by methionyl-tRNA formyltransferase (MTF); the resulting fMet-$tRNA^{fMet}_{CAU}$ subsequently binds to initiation factors (IFs) and the complex enters the peptidyl-tRNA site of the ribosome to initiate translation (*2, 3*). Although strict governance by the above mechanism controls the initiation event starting with fMet at the N-terminus exclusively, two methods have been devised to circumvent this requirement (*4–7*).

The first method involves engineering of the initiator tRNA where the CAU anticodon is altered to a triplet that assigns another amino acid (*4, 5*). Critical tRNA recognition elements in some aminoacyl-tRNA synthetases (aaRSs) reside in the anticodon interaction (*8*). Therefore, the transplantation of an appropriately chosen anticodon into $tRNA^{fMet}$ is able to deceive the recognition by aaRS, leading to a mischarging of the cognate amino acid onto the engineered initiator $tRNA^{fMet}_{XXX}$ (XXX indicates the implanted anticodon). When such a tRNA, together with a mRNA containing the countercodon to the tRNA anticodon, is overexpressed in *Escherichia coli* the translation starts with the mischarged N^{α}-formylated amino acid. Since not all aaRSs can be deceived by this type of tRNA engineering, only a limited number of aaRSs (Val, Phe, Gln, Ile, and Lys) were able to mischarge the cognate amino acids onto $tRNA^{fMet}_{XXX}$ and lead to initiation.

The second method also involves engineering of the initiator tRNA. The initiator tRNA anticodon is altered to CUA and the mRNA AUG codon is replaced with the UAG amber codon (*6, 7*), in a fashion similar to the incorpo-

ration of nonproteinogenic amino acids during the elongation event (*9–11*). When such a tRNA is *chemically* aminoacylated through multistep synthesis (*12*) and subjected to *in vitro* translation, the designated amino acid initiates peptide expression. Unfortunately, this approach suffers from the contamination of polypeptides initiated by other amino acids since the deacylated $tRNA^{fMet}{}_{CUA}$ can be recharged with the amino acids by cognate endogenous aaRSs (*6, 7*). This method has thus far demonstrated the initiation of translation with only a limited number of amino acids (Met, Val, and Lys) tagged with fluorescent or biotin groups for the purpose of expressing N-end-labeled polypeptides *in vitro*.

Taken together, the currently available technologies do not allow for the synthesis of polypeptides with a well-controlled N-terminal modification, and it remains unclear whether more diverse types of amino acids including those with various N^{α}-acyl groups are tolerated in the initiation event. We here have used the concept of genetic code reprogramming to engineer the initiation event. This new technology allows us to expand the repertoire of usable α-amino acids and their N^{α}-acyl groups for initiation.

As an application, we have demonstrated the ribosomal synthesis of an antitumor cyclic peptide, G7–18NATE, closed by a physiologically stable thioether bond. Significantly, the cyclization occurs spontaneously upon the translation of its precursor linear peptide so that no additional manipulation is necessary after translation. Moreover, this technology is applicable to the synthesis of cyclic peptides with a wide range of ring sizes, and thus it offers a novel tool to construct mRNA-programmed cyclic peptide libraries for the discovery of peptidic drug candidates with *in vivo* compatibility.

RESULTS AND DISCUSSION

Genetic Code Reprogramming for the Initiation Event. Reprogramming the genetic code is a technique in which the codons assigned to proteinogenic amino acids are reassigned to nonproteinogenic amino acids (*13*). This technique has been utilized for the reprogramming of the elongation event where multiple nonproteinogenic amino acids have been successfully incorporated into the nascent peptide chain (*13–20*). In principle, it is also applicable to the alteration of initiation, but it has not been yet extensively used for such an aim.

To perform reprogramming of the translation initiation, we have employed two key systems to facilitate reprogramming of the initiation event. The first one is an *Escherichia coli* reconstituted cell-free translation system, referred to as the PURE system (*21*). By *withdrawing* certain amino acids from the translation system (referred to as *w*PURE), vacant codons can be created in the genetic table. Met is depleted to make the initiation codon vacant, and therefore an alternative amino acid (Xaa) can be reassigned to initiation (Figure 1, panel a). The second key system is a *de novo* tRNA acylation system consisting of an artificially evolved ribozyme, termed flexizyme (Figure 1, panel b) (*20, 22, 23*). Flexizyme allows for virtually any amino acid to be charged onto tRNA. Thus, we expect that by filling the vacant start codon in *w*PURE system with Xaa-$tRNA^{fMet}{}_{CAU}$ prepared with the flexizyme system, *i.e.*, through the integration of these two systems, the initiator Met can be readily reassigned to any noncanonical amino acid (Figure 1, panel c).

To test whether these two systems could be integrated to reprogram the initiation event, a 14-mer peptide bearing a Flag tag sequence (*24*) at the C-terminus was expressed in the *w*PURE system lacking Met in the absence or presence of Met-$tRNA^{fMet}{}_{CAU}$ prepared by the flexizyme system (Figure 2, panel a). As a positive control, the wild-type peptide was expressed in the normal PURE system containing the full sets of amino acids (Figure 2, panel b, lane 1). As a negative control, when only the *w*PURE system was used, initiation did not occur and thereby no peptide was produced (Figure 2, panel b, lane 2). In contrast, when Met-$tRNA^{fMet}{}_{CAU}$ prepared by the flexizyme system was included in the *w*PURE system, initiation proceeded and yielded the expected band with the same intensity and mobility as the wild type (Figure 2, panel b, lane 3 *vs* lane 1). Matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) analysis of the product gave the expected molecular mass of the peptide bearing fMet at its N-terminus (Figure 2, panel c, Met). Thus, Met-$tRNA^{fMet}{}_{CAU}$ synthesized by the flexizyme system was formylated by endogenous MTF and able to initiate translation as efficiently as the native initiation event in the normal PURE system. These results demonstrated that the concept of genetic code reprogramming would be applicable to engineering of the initiation event as well as the elongation event aiming to express peptides containing various amino acids at the N-terminus.

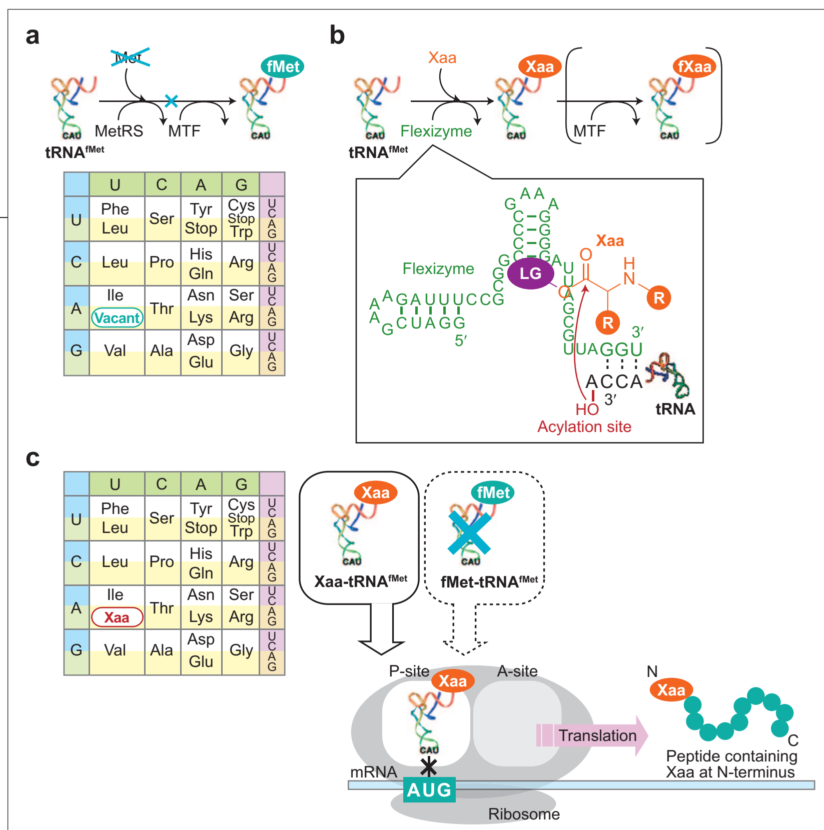

Figure 1. Engineering the initiation event by genetic code reprogramming. a) Codon table in the absence of Met. Upon removing Met from the translation mixture, tRNA$^{fMet}_{CAU}$ is no longer available for methionylation by MetRS, consequently making the start codon vacant, *i.e.*, *w*PURE system. b) Synthesis of Xaa-tRNA$^{fMet}_{CAU}$ by the flexizyme system. Xaa indicates various kinds of α-amino acid. The secondary structure of a representative flexizyme is shown in cyan letters. When Xaa-tRNA$^{fMet}_{CAU}$ is added to the translation apparatus, it can be formylated by endogenous MTF. c) Reassignment of the start codon from Met to Xaa. When Xaa-tRNA$^{fMet}_{CAU}$ prepared by the flexizyme system is added to the *w*PURE system, the translation is exclusively initiated with Xaa, expressing a peptide containing Xaa at the N-terminus.

Tolerance of Amino Acids in Initiation. We next tested initiation using each of the 19 proteinogenic amino acids by charging each amino acid to tRNA$^{fMet}_{CAU}$ with the flexizyme system. Tricine−SDS PAGE (soduim dodecyl sulfate polyacrylamide gel electrophoresis) analysis of the translation products showed that 11 amino acids out of 19 were able to initiate translation with greater than 50% efficiency relative to wild type (Figure 2, panel b, highlighted in orange). In addition, two amino acids exhibited the ability to initiate translation with approximately 30−50% efficiency relative to wild type (Figure 2, panel b, highlighted in pink). MALDI-TOF analysis of these translated peptides gave a single peak with the expected molecular mass of the N-terminal-formylated form initiated with the programmed amino acid (Figure 2, panel c, highlighted in the same color code as Figure 2, panel b). This indicates that the initiation occurs exclusively with the N^{α}-formylated amino acid as programmed by the flexizyme system.

The remaining six amino acids (Glu, Lys, Gly, Asp, Arg, and Pro) showed less than 30% translation efficiency compared to wild-type expression, with Arg and Pro producing yields in the single digits (Figure 2, panel b, highlighted in cyan). Despite their poor yields, the Flag-purified peptides initiated by Glu, Lys, Gly, and Asp gave the expected molecular mass of the formylated form, indicating that the full-length peptides were correctly initiated with the designated amino acids. Albeit with low efficiency, even Arg and Pro were able to initiate translation giving peaks corresponding to peptides initiated with programmed amino acid in the respective mass spectrum, although Arg gave a mixture of formylated and nonformylated peptides while Pro only produced the nonformylated peptide (Figure 2, panel c).

Interestingly, the majority of "good" initiators have hydrophobic side chains. On the other hand, six amino acids exhibited less than 30% efficiency in initiation (Figure 2, panel b, highlighted in cyan); among them, four have charged side chains (Lys, Glu, Asp, and Arg), one has no side chain (Gly), and the least efficient initiator Pro has the secondary amine. We assumed that inefficient initiation with these amino acids was caused by poor formylation of the α-amino group on the Xaa-tRNA$^{fMet}_{CAU}$ catalyzed by MTF. In the crystal structure of the complex between fMet-tRNA$^{fMet}_{CAU}$ and MTF, the side chain of Met fits into a binding pocket surrounded by hydrophobic amino acids (*25*). Presumably, this hydrophobic environment might dismiss the charged or hydrophilic side chains resulting in poor formylation on the α-amino group, while Pro, the only amino acid bearing the secondary amine, is likely enzymatically incompatible with MTF. Indeed, the two least efficient initiators, Arg and Pro, afforded nonformylated peptide in each mass spectrum of the product.

The above results indicated that the integration of the *w*PURE and flexizyme systems enabled us to readily reassign the AUG codon from Met to other amino acids and express the desired peptide initiated with the assigned amino acid. Although traces of undesirable peptides with slower or/and faster mobility were detected in the tricine-SDS PAGE analysis, these bands also appeared in the control experiment in the absence of aa-tRNA$^{fMet}_{CAU}$ (Figure 2, panel b, lane 2). This suggested that these bands appeared from aa-tRNA$^{fMet}_{CAU}$-independent expressions of the mRNA template, likely

due to in-frame mis-initiations. However, the expression levels of these peptides were the negligible background, and indeed we observed only a single desirable full-length product in the MALDI-TOF analysis. This was a sharp contrast to the previous method that suffered from contaminations of undesirable full-length peptides mis-initiated with undesignated amino acids.

We believe that the observed nearly exclusive expression of the desirable peptides has been achieved by two critical strategies we took for reprogramming the translation initiation event. First, depletion of Met in the *w*PURE system enabled us to suppress the background initiation reaction as low as possible. Second, the flexizyme system enabled us to use the natural initiator tRNA$^{fMet}_{CAU}$ as an aminoacyl-carrier for noncognate amino acids, and thereby we achieved a maximal efficiency for initiating the translation with each amino acid. By the combination of these two strategies, we were able to express the full-length peptide initiated with the designated noncanonical amino acid as a sole product.

Enhancing the Initiation Efficiency by Preacylation of α-Amino Group. As described above, an intriguing observation was made regarding the Arg-initiated peptide in which the N-terminus was not fully formylated (Figure 2, panel c, Arg). Likewise, the N-terminus of Pro-initiated peptide was not formylated at all (Figure 2, panel c, Pro). This result implies that these inefficient initiators are poor substrates for MTF (Figure 3, panel a). This observation led us to postulate that the efficiency of formylation on the α-amino group of Xaa-tRNA$^{fMet}_{CAU}$ by MTF might play an important role in determining the initiation efficiency, since the expression levels of these peptides were significantly lower than peptides initiated with other amino acids. RajBhandary and coworkers (*26*) have also proposed

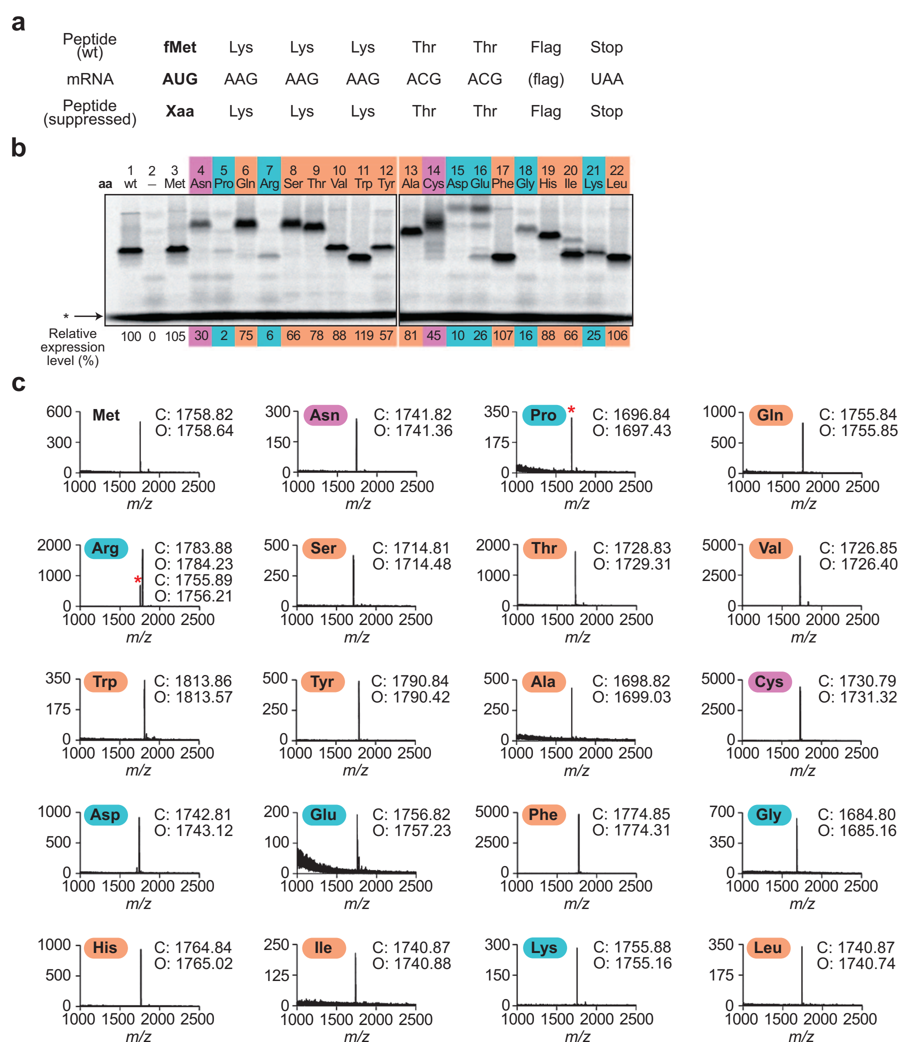

Figure 2. Tolerance of various amino acids in initiation. a) The mRNA sequence that expresses peptides initiated by various amino acids. Flag in parentheses indicates the RNA sequence encoding a Flag peptide (DYKDDDDK). b) Tricine-SDS PAGE analysis of the translation products. Lane 1, expression of wild type; lane 2, in the absence of Met; lane 3, initiated by Met-tRNA$^{fMet}_{CAU}$; lanes 4–22, initiated by Xaa-tRNA$^{fMet}_{CAU}$. Each expression level relative to wild type is determined by a mean score of triplicates or more. The band indicated by asterisk corresponds to the remaining [^{14}C]-Asp that was not incorporated into the peptide. The amino acids giving >50%, 30–50% and <30% of the wild-type expression level are highlighted in orange, pink, and cyan, respectively. Observed minor bands were likely attributed to peptides originated from the aa-tRNA$^{fMet}_{CAU}$-independent background expression, *i.e.*, random starts of translation of the RNA template. c) MALDI-TOF MS spectra of the translated peptides. The noncanonical amino acids are highlighted in the same color code as used in panel b. The red asterisk indicates a peak corresponding to the nonformylated peptide, while the unlabeled peak corresponds to the formylated peptides. The calculated mass (C) and observed mass (O) are shown in each spectrum.

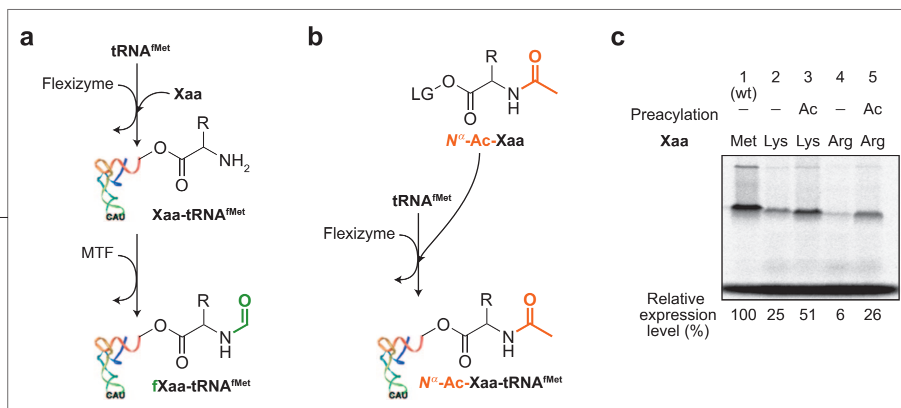

Figure 3. Increase in expression level by N^{α}-acylation. a) Formylation of Xaa-tRNA$^{fMet}_{CAU}$ in the *w*PURE system. When Xaa-tRNA$^{fMet}_{CAU}$ is used in the initiation, it is generally formylated. b) Aminoacylation of tRNA$^{fMet}_{CAU}$ with preacetylated amino acids by the flexizyme system. This method ensures N^{α}-acylation on the α-amino group. c) Tricine−SDS PAGE analysis of the translated peptide. Lane 1, expression of wild type; lane 2, initiated with Lys-tRNA$^{fMet}_{CAU}$; lane 3, initiated with N^{α}-Ac-Lys-tRNA$^{fMet}_{CAU}$; lane 4, initiated with Arg-tRNA$^{fMet}_{CAU}$; lane 5, initiated with N^{α}-Ac-Arg-tRNA$^{fMet}_{CAU}$. Each expression level relative to wild type is determined by a mean score of triplicates or more.

such a correlation based on experiments using an alternative codon for initiation. We therefore considered the possibility that preacylation on the α-amino group might increase the expression level of these peptides (Figure 3, panel b).

In order to verify this hypothesis, we chose two amino acids, Lys and Arg, both of which showed poor expression levels (25% and 6%, respectively) and prepared N^{α}-Ac-Lys-tRNA$^{fMet}_{CAU}$ and N^{α}-Ac-Arg-tRNA$^{fMet}_{CAU}$ by using the flexizyme system. It was investigated whether the expression level initiated with these preacylated amino acids increased compared with the corresponding α-amino-free amino acids. In both cases, peptide expression levels increased compared to those initiated with the corresponding α-amino-free amino acids (51% *vs* 25% for Lys and 26% *vs* 6% for Arg, Figure 3, panel c) and initiation proceeded exclusively with the designated N^{α}-Ac-amino acids (see Figure S1 for the MALDI-TOF molecular weight confirmation). It should be noted that even though the α-amino group on Arg was preacylated, the expression level of the peptide initiated with N^{α}-Ac-Arg did not exceed 50% efficiency compared with that of Met-initiated peptide. This suggests that other components in the *w*PURE system, such as IFs and ribosome, also play some roles in determining the initiation efficiency. Nonetheless, the preacylation on the α-amino group enhances the expression level for "modest" initiators.

Tolerance of N^{α}-Acyl Groups in Initiation. Biologically active peptides often have unique N-terminal-acyl groups. For instance, the N-terminus of some peptides synthesized by nonribosomal peptide synthetases are modified with fatty acids (*27*), while some peptide hormones have an acetyl or pyroglutamate group at their N-terminus (*28, 29*). Thus, it is of interest to see if the reprogrammed initiation in the *w*PURE system is able to tolerate amino acids with a variety of N^{α}-acyl groups. We chose Phe as an N^{α}-acyl carrier and expressed peptides initiated with Phe having alkyl or fatty acids (Ac, Hex, Mhe, Oct), pyroglutamate (pyE), and functionalities accessible to post-translational modification (Cba, Pen, Pyl, N_3Ac, oxP, ClAc; Figure 4, panel a).

Translation was initiated with N^{α}-acyl-Phe-tRNAs prepared by the flexizyme system, and the *N*-acyl-peptides were expressed in the *w*PURE system (Figure 4, panel b). Remarkably, the expression levels were generally higher than those of fMet- and fPhe-initiated control peptides; particularly the use of Phe modified with hydrophobic fatty acids gave significantly higher expression levels. MALDI-TOF analysis of these peptides agreed with the expected products without contamination of other misinitiated peptides (Figure 4, panel c). Thus, the results clearly indicated that a wide variety of N^{α}-acyl groups are tolerated in the initiation event. The N^{α}-acyl groups incorporated in this study involved various chemical functionalities such as fatty acid chains, pyroglutamate, and those that are applicable to post-translational modification. Hence, this approach would allow us to directly translate naturally occurring bioactive peptides containing unique N-terminal groups.

Ribosomal Synthesis of Cyclic Peptides *via* a Thioether Linkage. G7−18NATE is a potential peptidic anticancer agent, which has been shown to inhibit the interaction between the SH2 domain of Grb7 and receptor tyrosine kinases (*30–32*). The peptide consists of 11 natural amino acids and a non-natural cyclic structure with a nonreducible thioether bond linking the N-terminal acetyl group to the C-terminal Cys residue (see Figure S2a). Due to the fact that G7−18NATE contains a noncanonical bond, ordinary translation has been inapplicable to its synthesis; therefore this peptide could only be produced through chemical synthesis. However, we envisaged that our platform technology would enable us to perform the ribosomal synthesis of G7−18NATE.

We expected that reprogramming initiation with N^{α}-chloroacetyl (ClAc)-Trp would allow us to synthesize the linear N^{α}-ClAc-(G7−18NATE) precursor peptide, and subsequently the sulfhydryl group on the terminal Cys would spontaneously attack the α-carbon of the N^{α}-ClAc group, closing the ring (Figure 5, panel a) (*33–35*). We

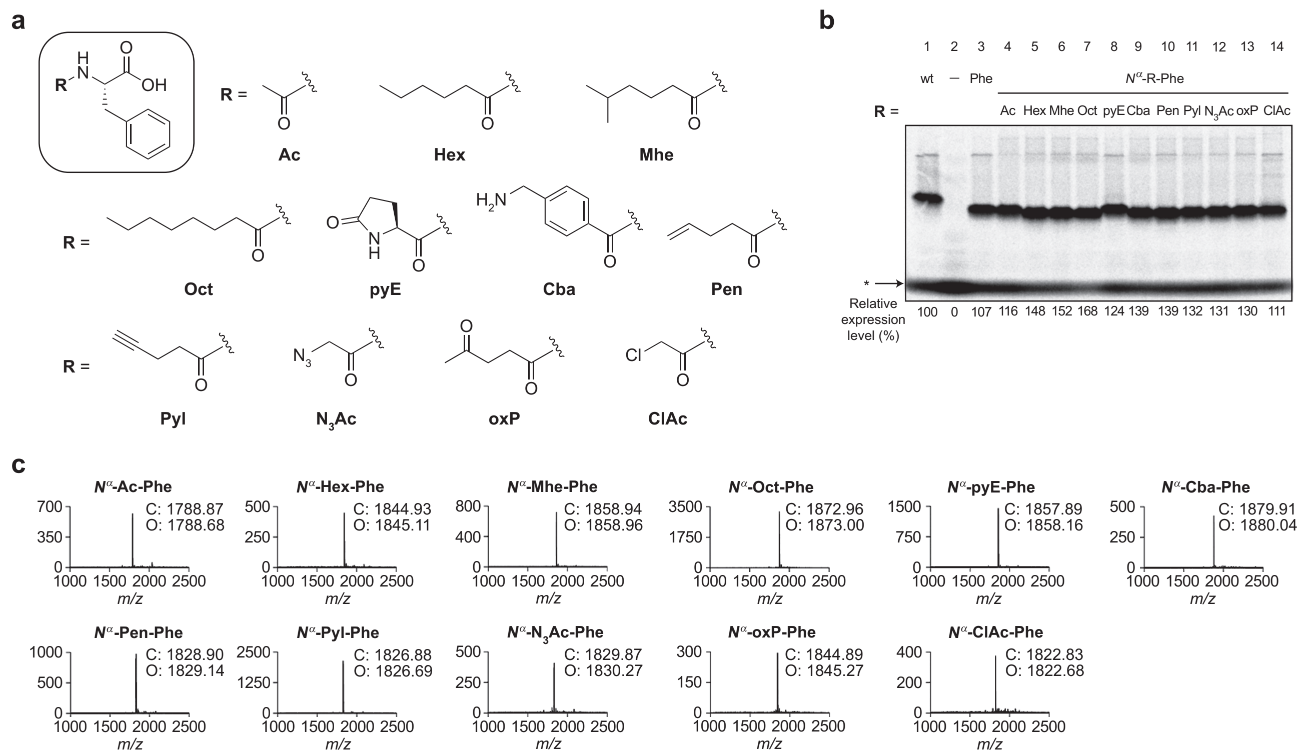

Figure 4. Tolerance of N^{α}-acyl groups in initiation. a) N^{α}-acyl phenylalanine derivatives used in this study. b) Tricine−SDS PAGE analysis of the translation products initiated with various N^{α}-acyl phenylalanine derivatives. Lane 1, expression of wild type; lane 2, in the absence of Met; lane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

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

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

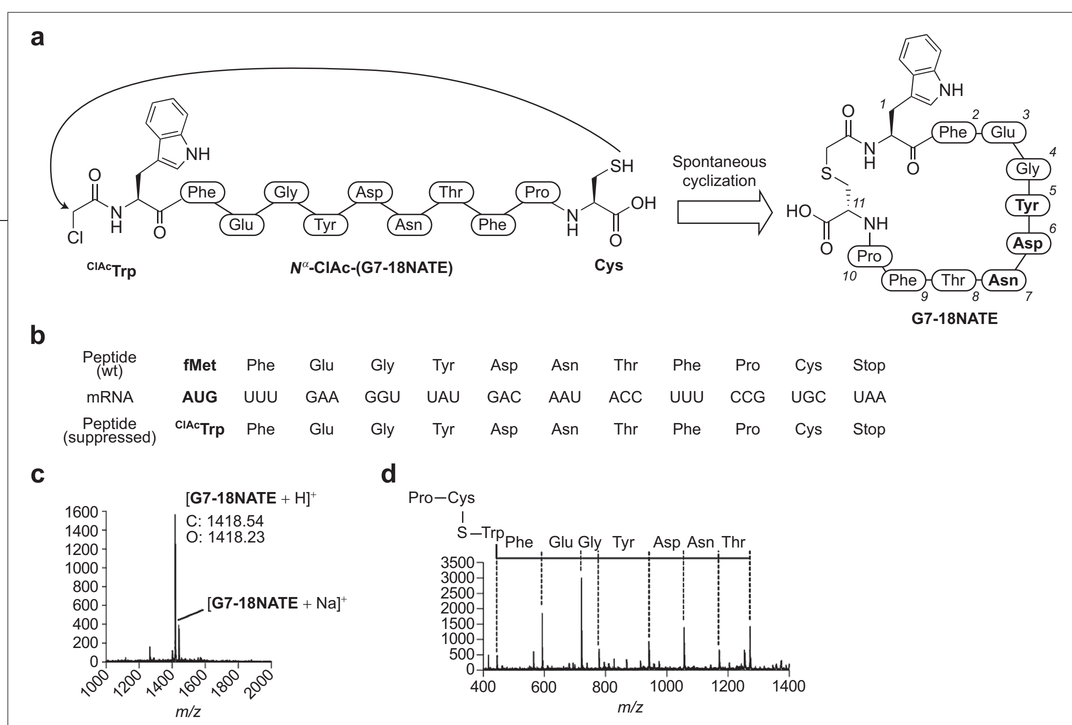

Figure 5. Ribosomal synthesis of G7−18NATE. a) The spontaneous cyclization affording G7−18NATE. The precursor linear peptide, N^{α}-ClAc-G7−18NATE, translated in the *w*PURE system spontaneously cyclizes into G7-18NATE by intramolecuar attack of Cys sulfhydryl group to the α-carbon of the N^{α}-ClAc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

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

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

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

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

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

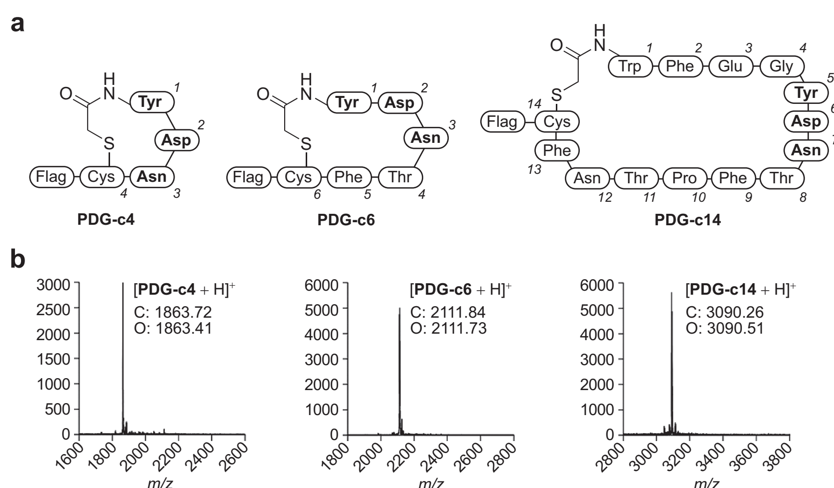

Figure 6. Custom synthesis of the cyclic peptides with various ring sizes. a) Structure of the expressed cyclic peptides. Three amino acids known to be critical residues in G7−18NATE are shown in bold, and "Flag-tag" indicates a modified Flag peptide sequence (LTTDYKDDDDK). b) MALDI-TOF mass spectrum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

CONCLUSION

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

METHODS

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

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

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

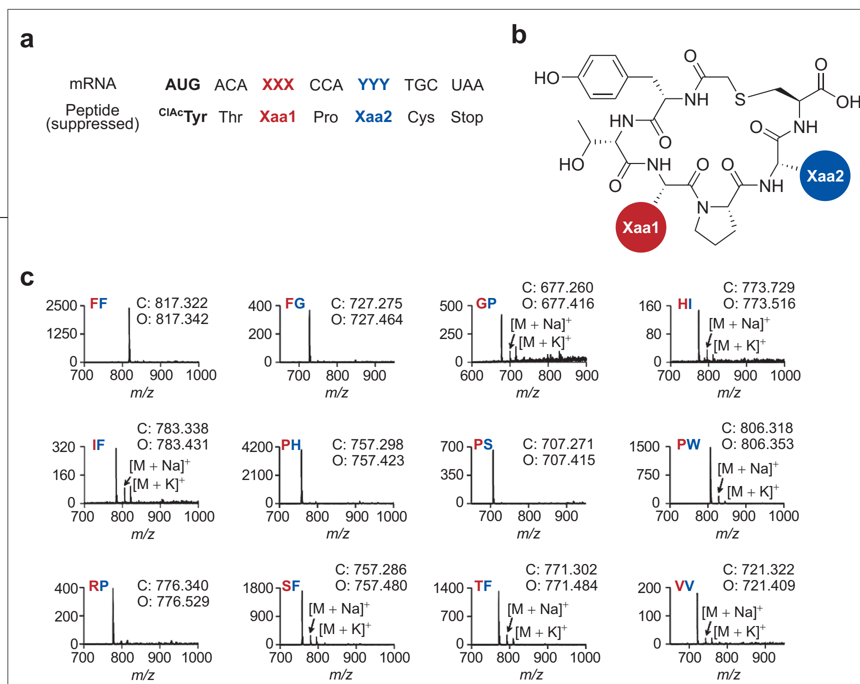

Figure 7. Construction of a cyclic peptide library. a) The sequence of mRNA used in the library synthesis. The codons indicated in red (Xaa1) and blue (Xaa2) were changed in each expression. The start codon is suppressed with N^{α}-ClAc-Tyr. b) The structure of cyclic peptide contained in the peptide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

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

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

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

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

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

REFERENCES

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

19. Seebeck, F. P., and Szostak, J. W. (2006) Ribosomal synthesis of dehydroalanine-containing peptides, *J. Am. Chem. Soc. 128*, 7150–7151.
20. Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis, *Nat. Methods 3*, 357–359.
21. Shimizu, Y., Inoue, A., Tomari, Y., Suzuki, T., Yokogawa, T., Nishikawa, K., and Ueda, T. (2001) Cell-free translation reconstituted with purified components, *Nat. Biotechnol. 19*, 751–755.
22. Murakami, H., Saito, H., and Suga, H. (2003) A versatile tRNA aminoacylation catalyst based on RNA, *Chem. Biol. 10*, 655–662.
23. Murakami, H., Kourouklis, D., and Suga, H. (2003) Using a solid-phase ribozyme aminoacylation system to reprogram the genetic code, *Chem. Biol. 10*, 1077–1084.
24. Brizzard, B. L., Chubet, R. G., and Vizard, D. L. (1994) Immunoaffinity purification of FLAG epitope-tagged bacterial alkaline phosphatase using a novel monoclonal antibody and peptide elution, *Biotechniques 16*, 730–735.
25. Schmitt, E., Panvert, M., Blanquet, S., and Mechulam, Y. (1998) Crystal structure of methionyl-tRNA(f)(Met) transformylase complexed with the initiator formylmethionyl-tRNA(f)(Met), *EMBO J. 17*, 6819–6826.
26. Li, Y., Ramesh, V., Mangroo, D., Taneja, C., and RajBhandary, U. L. (2000) Suppressor mutations in *Escherichia coli* methionyl-tRNA formyltransferase that compensate for the formylation defect of a mutant tRNA aminoacylated with lysine, *Biochemistry 39*, 8039–8046.
27. Grunewald, J., and Marahiel, M. A. (2006) Chemoenzymatic and template-directed synthesis of bioactive macrocyclic peptides, *Microbiol. Mol. Biol. Rev. 70*, 121–146.
28. Harris, J. I., and Roos, P. (1956) Amino-acid sequence of a melanophore-stimulating peptide, *Nature 178*, 90.
29. Takayasu, S., Sakurai, T., Iwasaki, S., Teranishi, H., Yamanaka, A., Williams, S. C., Iguchi, H., Kawasawa, Y. I., Ikeda, Y., Sakakibara, I., Ohno, K., Ioka, R. X., Murakami, S., Dohmae, N., Xie, J., Suda, T., Motoike, T., Ohuchi, T., Yanagisawa, M., and Sakai, J. (2006) A neuropeptide ligand of the G protein-coupled receptor GPR103 regulates feeding, behavioral arousal, and blood pressure in mice, *Proc. Natl. Acad. Sci. U.S.A. 103*, 7438–7443.
30. Pero, S. C., Oligino, L., Daly, R. J., Soden, A. L., Liu, C., Roller, P. P., Li, P., and Krag, D. N. (2002) Identification of novel non-phosphorylated ligands, which bind selectively to the SH2 domain of Grb7, *J. Biol. Chem. 277*, 11918–11926.
31. Tanaka, S., Pero, S. C., Taguchi, K., Shimada, M., Mori, M., Krag, D. N., and Arii, S. (2006) Specific peptide ligand for Grb7 signal transduction protein and pancreatic cancer metastasis, *J. Natl. Cancer. Inst. 98*, 491–498.
32. Pero, S. C., Shukla, G. S., Cookson, M. M., Flemer, S., Jr., and Krag, D. N. (2007) Combination treatment with Grb7 peptide and Doxorubicin or Trastuzumab (Herceptin) results in cooperative cell growth inhibition in breast cancer cells, *Br. J. Cancer 96*, 1520–1525.
33. Roberts, K. D., Lambert, J. N., Ede, N. J., and Bray, A. M. (1998) Efficient synthesis of thioether-based cyclic peptide libraries, *Tetrahedron Lett. 39*, 8357–8360.
34. Lung, F. D. T., King, C. R., and Roller, P. P. (1999) Development of non-phosphorylated cyclic thioether peptide binding to the Grb2-SH2 domain, *Lett. Pept. Sci. 6*, 45–49.
35. Yu, L., Lai, Y. H., Wade, J. V., and Coutts, S. M. (1998) A simple and efficient method for the syntheses of thioether cyclic peptides, *Tetrahedron Lett. 39*, 6633–6636.
36. Hanes, J., and Pluckthun, A. (1997) *In vitro* selection and evolution of functional proteins by using ribosome display, *Proc. Natl. Acad. Sci. U.S.A. 94*, 4937–4942.
37. Nemoto, N., Miyamoto-Sato, E., Husimi, Y., and Yanagawa, H. (1997) *In vitro* virus: bonding of mRNA bearing puromycin at the 3′-terminal end to the C-terminal end of its encoded protein on the ribosome *in vitro*, *FEBS Lett. 414*, 405–408.
38. Roberts, R. W., and Szostak, J. W. (1997) RNA-peptide fusions for the *in vitro* selection of peptides and proteins, *Proc. Natl. Acad. Sci. U.S.A. 94*, 12297–12302.