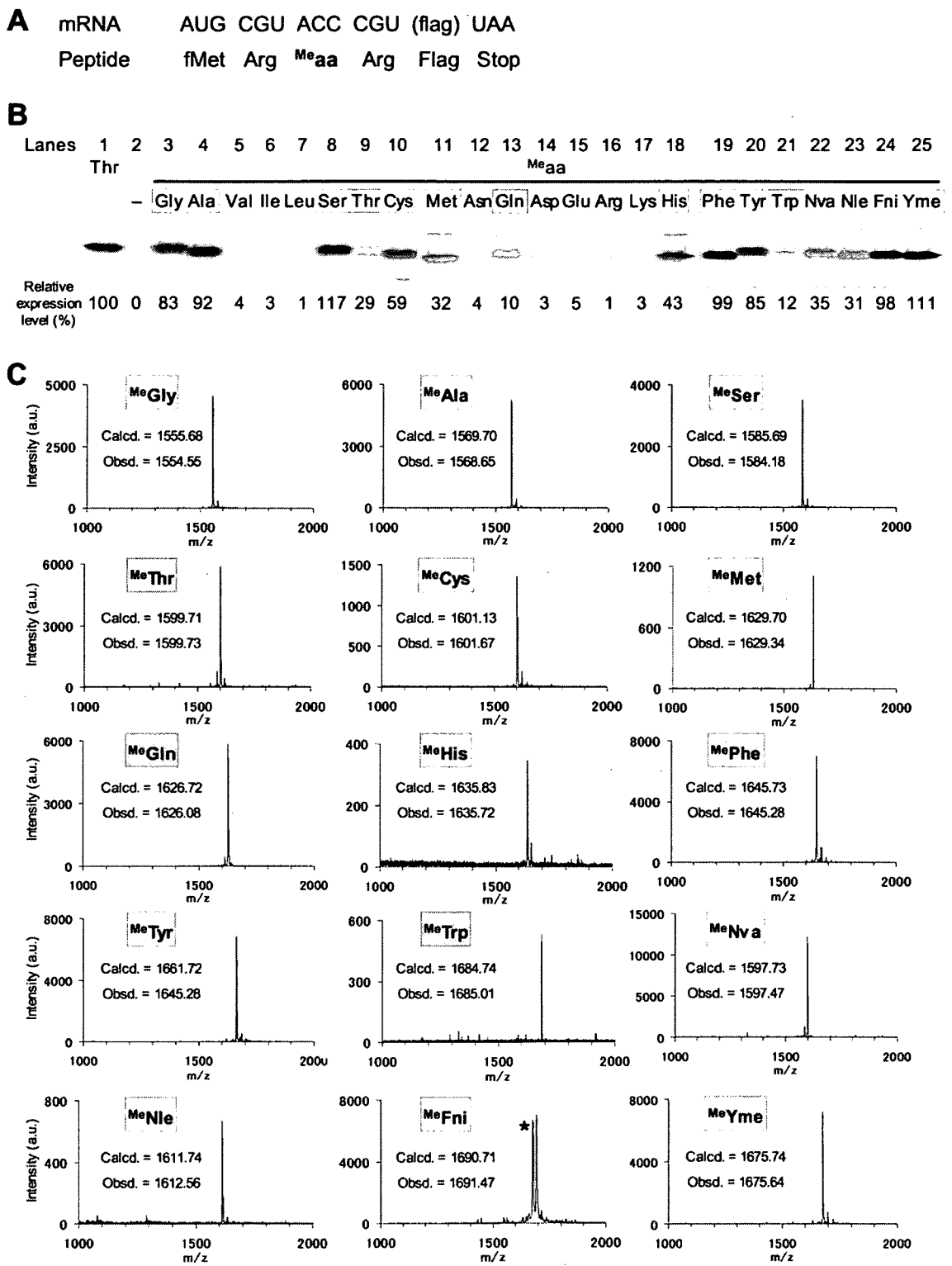

Figure 2. Single Incorporations of N^{α}-Methylated Amino Acids

(A) Sequences of mRNA and peptide. Flag in parentheses indicates the RNA sequence coding the Flag peptide (DYKDDDDK).

(B) Tricine-SDS-PAGE analysis of the expressed peptides labeled with [^{14}C]-Asp detected by autoradiography. Lane 1, the wild-type peptide expressed in the ordinary PURE system where ACC assigns Thr; lane 2, a negative control using *w*PURE system in the presence of uncharged $tRNA^{Asn\text{-}E1}_{GGU}$; lanes 3–25, expression of the peptide containing a single ^{Me}aa in the presence of designated ^{Me}aa-$tRNA^{Asn\text{-}E1}_{GGU}$ prepared by the flexizyme system. Each expression level relative to wild-type was determined by a mean score of triplicates. The amino acids giving over 30% and 10% expression levels relative to the wild-type are highlighted in cyan and orange, respectively.

reconstituted *Escherichia coli* cell-free translation system in which transcription and translation are coupled (Shimizu et al., 2001). The most important feature of this translation apparatus is that some components, such as amino acids and aminoacyl-tRNA synthases (aaRSs), can be withdrawn from the translation elements (Figure 1B). By means of such a *withdrawn* PURE system, named *w*PURE, we are able to vacant certain codons at our will. Combining these two systems enables us to readily reassign the vacant codons to ^{Me}aa and express peptides as designed on mRNA sequences according to the newly designated genetic table (Figure 1C).

In this report, we initially screened N^{α}-methylated proteinogenic and nonproteinogenic amino acids to evaluate the efficiency of their single incorporation into a peptide backbone. Then, we investigated the mRNA-programmed synthesis of peptides containing multiple ^{Me}aa, achieving the compositions of continuous stretches of up to ten ^{Me}aa residues. Finally, we investigated the synthesis of cyclic N-methyl-peptides closed by a physiologically stable thioether bond, showing its potential for the generation of in vivo compatible peptide libraries therapeutically targeting various proteins.

RESULTS

Single Incorporations of N^{α}-Methylated Amino Acids

To investigate the incorporation of multiple ^{Me}aa into a peptide backbone, various N^{α}-methylated proteinogenic and nonproteinogenic amino acids charged onto a tRNA by the flexizyme system were comprehensively evaluated for the efficiency of their single incorporation into a peptide backbone. As was discussed earlier, Merryman and Green performed a similar experiment with a dipeptide synthesis format (Merryman and Green, 2004). In contrast, we here incorporated ^{Me}aa into a 12-mer model peptide, thus involving the incorporation of ^{Me}aa into a specific site followed by elongation and termination of peptide synthesis as a format of normal peptide synthesis (Figure 2A, more details are discussed below).

For this study, we chose 23 ^{Me}aa substrates, 19 of which were derived from proteinogenic amino acids, and 4 of which were derived from nonproteinogenic amino acids (Figure S1, see the Supplemental Data available with this article online). The flexizyme-catalyzed aminoacylation efficiency for each ^{Me}aa was verified by our standard protocol (Figure S2), showing that the majority of ^{Me}aa could be charged onto tRNA with over 30% yields. Three amino acids, ^{Me}Val, ^{Me}Ile, and ^{Me}Asn, showed less than 30% yields. In the former two amino acids, their steric hindrance resulted from the combination of the β-branched side chain and the N^{α}-methyl-amino group presumably made these particular substrates much less reactive to flexizyme compared with other ^{Me}aa. In the case of ^{Me}Asn, based on our past experience working on Asn-tRNA, its tRNA-adduct might be labile due to the intramolecular attack of the amide side chain (Lee and Suga, 2001; Murakami et al., 2006); thereby the yield might appear to be low in PAGE analysis. However, we have previously observed that such aa-tRNAs prepared in similar ranges of quality (Val-, Ile-, and Asn-tRNAs in 13%–30% yield) could still be used for the incorporation into the nascent peptide chain with satisfactory yields (30%–70%) (Murakami et al., 2006). Thus, we pursued to survey all ^{Me}aa-tRNAs for the single incorporation into the model peptide.

The *w*PURE system used in our survey was composed of a limited member of amino acids and aaRSs. We designed an open-reading frame (ORF) in mRNA (Figure 2A) that expresses a short peptide consisting of fMet-Arg-^{Me}aa-Arg followed by a Flag peptide (DYKDDDDK; D, Asp; Y, Tyr; K, Lys). We selected a Thr codon (ACC) to assign ^{Me}aa and an orthogonal $tRNA^{Asn\text{-}E1}_{GGU}$ (Ohta et al., 2007) for its suppression. When the model peptide was expressed in the presence of uncharged $tRNA^{Asn\text{-}E1}_{GGU}$ in *w*PURE system, no product band appeared on tricine-SDS-PAGE (Figure 2B, lane 2), suggesting that $tRNA^{Asn\text{-}E1}_{GGU}$ could act as an orthogonal tRNA inert against the endogenous aaRSs added to the *w*PURE system.

We next tested the single incorporation of ^{Me}aa charged onto $tRNA^{Asn\text{-}E1}_{GGU}$ into the model peptide sequence. Eight out of 19 ^{Me}aa derived from proteinogenic amino acids were incorporated into the peptide with more than 30% efficiencies relative to wild-type expression (Figure 2B, lanes 1 versus those highlighted in cyan). The incorporation of the designated ^{Me}aa was also confirmed by MALDI-TOF analysis, giving the expected molecular mass as a sole peak in all cases (Figure 2C). Moreover, ^{Me}Thr, ^{Me}Gln, and ^{Me}Trp were modestly incorporated yet gave the single desired product (Figures 2B and 2C, those highlighted in orange). We called these ^{Me}aa molecules that could be incorporated into the peptide chain with more than 10% efficiencies as "accessible" ^{Me}aa. Interestingly, all of these "accessible" ^{Me}aa shared similar structural features of their side chains; either aromatic side chains (^{Me}Phe, ^{Me}Tyr, ^{Me}His, and ^{Me}Trp) or noncharged and nonbulky side chains (^{Me}Gly, ^{Me}Ala, ^{Me}Ser, ^{Me}Cys, ^{Me}Met, ^{Me}Thr, and ^{Me}Gln).

It should be noted that our results were mostly consistent with Merryman's observations, with some contradictions (Merryman and Green, 2004). In their experiments, ^{Me}Val, ^{Me}Ile, and ^{Me}Leu were ranked in the "efficient," or what we referred to as the "accessible," ^{Me}aa group. However, our experiments showed that ^{Me}Val, ^{Me}Ile, and ^{Me}Leu were grouped as rather "inaccessible" ^{Me}aa for incorporation (Figure 2B, lanes 5–7). The difference between the in vitro-transcribed $tRNA^{Asn\text{-}E1}$ and the native cognate tRNAs that Merryman et al. used as ^{Me}aa-carriers may have an effect on their incorporation efficiencies (Dale and Uhlenbeck, 2005; LaRiviere et al., 2001). However, it should be noted that in the Merryman's TLC electrophoresis assay, the mobility of these branched aliphatic ^{Me}aa were indistinguishable from that of the cognate natural ones. Therefore, it was not clearly defined if the observed product spot originated from fMet-^{Me}aa or fMet-aa. Because the reductive alkylation of some aa-tRNAs to generate ^{Me}aa-tRNAs possibly left a small amount of unreacted aa-tRNAs as Merryman et al. discussed (Merryman and Green, 2004), it could not be ruled out that such aa-tRNA contamitants competed out ^{Me}aa-tRNAs for incorporation into the nascent

(C) MALDI-TOF-MS spectra of the Flag-purified N-methyl-peptides. The calculated molecular mass (Calcd.) and observed molecular mass (Obsd.) for singly charged species, $[M+H]^{+1}$ are shown in each spectrum. In the spectrum of peptide with ^{Me}Fni, an additional peak (asterisk) was detected since the additional peak corresponds to N^{α}-methyl-*p*-nitroso-Phe (Calcd. = 1674.72, Obsd. = 1675.45) generated by photodecomposition of NO_2 to NO during the MALDI-TOF-MS analysis (Ho and Chow, 1996).

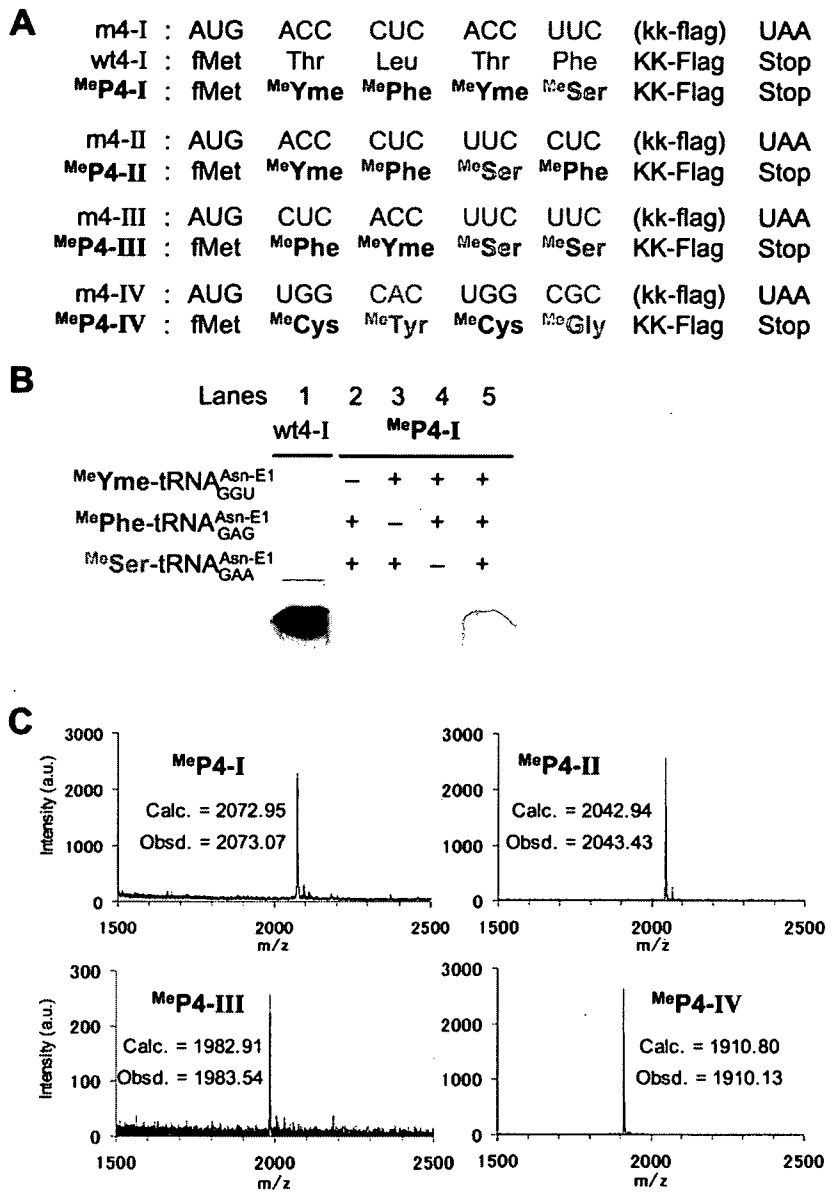

Figure 3. Four Successive Incorporations of Multiple N^{α}-Methylated Amino Acids

(A) Sequences of mRNA templates (m4-I–IV), tetra-N-methyl-peptides (^{Me}P4-I–IV), and the control wild-type peptide (wt4-I). The kk-flag in parentheses indicates the RNA sequence coding a KK-Flag peptide (KKDYKDDDDK). Arabic number denotes the number of N-methylated peptide bonds. The codons and amino acids are colored according to the reprogrammed genetic code table in Figure 1C.

(B) Tricine-SDS-PAGE analysis of the peptides expressed from m4-I labeled with [^{14}C]-Asp detected by autoradiography. Lane 1, the wild-type peptide expressed in the ordinary PURE system; lanes 2–4, negative controls in the presence of two of ^{Me}aa-tRNA^{Asn-E1} as shown; lane 5, in the presence of all three ^{Me}aa-tRNA^{Asn-E1}.

(C) MALDI-TOF-MS spectra of ^{Me}P4-I–IV. The calculated mass (Calcd.) and observed mass (Obsd.) are shown in each spectrum.

peptide chain. In contrast, in our studies, because the flexizyme system ensures the purity of each ^{Me}aa-tRNA, i.e., no contamination of aa-tRNA as confirmed by MALDI-TOF analysis (Figure 2C), we are able to avoid such issues and judge its intrinsic incorporation efficiency more reliably and quantitatively. To this end, we propose that ^{Me}Val, ^{Me}Ile, and ^{Me}Leu are inaccessible substrates due to the combination of bulkiness in the branched side chain and the N^{α}-methylated α-amino group.

Four ^{Me}aa derived from nonproteinogenic amino acids (Figure S1) were also incorporated into the corresponding peptide chain (Figure 2B, lanes 21–25), each of which the MALDI-TOF data was consistent with the expected molecular mass (Figure 2C). Again, the same trend for the incorporation efficiency seemed to apply to these nonproteinogenic ^{Me}aa, where the aromatic ones were incorporated more efficiently than the aliphatic ones. All data taken together, we concluded that ^{Me}aa having "aromatic" or "nonbulky and noncharged" side chains were preferable for the incorporation into the peptide backbone. This information is invaluable for us to select appropriate ^{Me}aa and design specific compositions of N-methylated peptides or peptide libraries while avoiding the risk of potential failure in the synthesis.

Four Successive Incorporations of Multiple N^{α}-Methylated Amino Acids

We next attempted the consecutive incorporations of multiple ^{Me}aa into the peptide backbone via genetic code reprogramming. Five ^{Me}aa derived from proteinogenic amino acids (^{Me}Ser, ^{Me}Phe, ^{Me}Tyr, ^{Me}Cys, ^{Me}Gly) and one derived from a nonproteinogenic amino acid (^{Me}Yme) were selected and assigned arbitrarily to six codons as shown in Figure 1C. The flexizyme system was utilized to charge the respective ^{Me}aa on tRNA^{Asn-E1} bearing the anticodon that reads the assigned codon. We also designed four mRNA templates containing ORFs (Figure 3A, m4-I–IV) that express the corresponding peptides comprised of three different ^{Me}aa in four successive residues (Figure 3A, ^{Me}P4-I–IV). The C terminus of the respective peptides was also designed to contain a modified Flag sequence (KK-Flag = KKDYKDDDDK; D, Asp; Y, Tyr; K, Lys) to facilitate its isolation as well as ionization upon MALDI-TOF analysis.

We first used m4-I to monitor the expression level of ^{Me}P4-I by tricine-SDS-PAGE with a series of appropriate controls (Figure 3B). ^{Me}P4-I was expressed only when all designated ^{Me}aa-tRNA^{Asn-E1}s were present in the *w*PURE system (lanes 2–5), suggesting that no competing background expression of undesignated peptides occurred. The expression level of ^{Me}P4-I was an approximately 20% relative to that of wild-type (wt4-I) expressed in the ordinary PURE system (lanes 1 versus 5). Despite the modest expression level of ^{Me}P4-I, MALDI-TOF analysis of the Flag-purified ^{Me}P4-I gave a single peak with the expected molecular mass, indicating that the assigned codons on mRNA were

correctly read by these ^{Me}aa-tRNA$^{Asn\text{-}E1}$s and the designated tetra-N-methyl-peptide was expressed.

Likewise, ^{Me}P4-II–IV peptides were expressed at levels approximately 15%–20% of the corresponding wild-type peptides (data not shown). Again, MALDI-TOF data of ^{Me}P4-II–IV peptides were consistent with the expected mass (Figure 3C). Thus, all mRNA ORFs tested in this study could be correctly translated to the tetra-N-methyl-peptides according to the assigned codons with high fidelities. This study represents, to our knowledge, the first proof-of-concept experiment for the ribosomal synthesis of N-methyl-peptides containing multiple kinds of ^{Me}aa by the genetic code reprogramming.

Expression of Longer N-Methyl-Peptides

The achieved high fidelity control in the ribosomal synthesis of the tetra-N-methyl-peptides encouraged us to attempt the expression of longer N-methyl-peptides. We designed four mRNA templates that express N-methyl-peptides containing five, six, eight, or ten consecutive N-methyl-peptide bonds (Figure 4A) (m5/^{Me}P5, m6/^{Me}P6, m8/^{Me}P8, m10/^{Me}P10, respectively). These peptides consisted of three ^{Me}aa in a repetition of the sequence ^{Me}Yme, ^{Me}Phe, and ^{Me}Ser for the corresponding length. To compare the expression level of these N-methyl-peptides, the respective wild-type peptides (wt5–10) were also expressed by using the ordinary PURE system.

All mRNA templates successfully expressed the peptides in the presence of the corresponding ^{Me}aa-tRNA$^{Asn\text{-}E1}$s in *w*PURE system (Figure 4B, gel figure). The absolute expression level for each N-methyl-peptide, determined by its radioisotope intensity originating from the incorporation of five [^{14}C]-Asp residues in the Flag peptide, gradually decreased with increasing peptide length (Figure 4B, upper graph). However, since wild-type also exhibited an inverse-correlation between expression level and peptide length, the relative expression levels of N-methyl-peptides to the corresponding wild-types remained in a similar range (10%–20%), suggesting that the relative expression level was independent of length (Figure 4B, lower graph). Despite their modest yields, the respective N-methyl-peptide had a single peak with the expected molecular mass in MALDI-TOF analysis, indicating that all peptides were expressed with high fidelities according to the corresponding mRNA templates assigned with the reprogrammed genetic code (Figure 4C).

In conclusion, our work has clearly demonstrated up to ten successive incorporations of ^{Me}aa into the peptide backbone via genetic code reprogramming with high fidelity control of both the peptide sequence and length. This technology offers a new means for the sequence-controlled synthesis of peptides containing multiple N-methyl-modifications on their backbone.

Synthesis of Cyclic N-Methylated Peptides Closed by a Thioether Bond

With the above method in our hands, we next applied it to the synthesis of cyclic peptides closed by a physiologically stable thioether bond. The foundation of this peptide ring-closing method has been established during the course of our studies on initiation reprogramming (Goto et al., 2008). Briefly, we prepared a *w*PURE system where Met was depleted so that the initiation codon became vacant. This vacant initiation codon was then filled with initiator tRNA$^{fMet}_{CUA}$ charged with an arbitrarily chosen proteinogenic or nonproteinogenic amino acid by the flexizyme system, so that the translation starts from this newly assigned initiator amino acid. When N^{α}-(α-chloroacetyl)-amino acid (N^{α}-ClAc-aa) and Cys were assigned to the initiation and elongation codons, respectively, the expressed full-length linear peptide spontaneously cyclized via an intramolecular nucleophilic attack of the sulfhydryl group of the Cys side chain to the N-terminal α-carbon of ClAc group, resulting in the formation of a thioether bond. This technology has enabled us to synthesize various cyclic peptides with any sequences and ring sizes. We here attempted to combine initiation reprogramming with elongation reprogramming using ^{Me}aa, to enable us to devise a flexible and reliable methodology for the mRNA-programmed synthesis of cyclic N-methyl-peptides.

We designed three mRNA templates that express peptides containing multiple N-methylated peptide bonds (Figure 5A, mc-I–III). Six ^{Me}aa (^{Me}Ala, ^{Me}Yme, ^{Me}Tyr, ^{Me}Gly, ^{Me}Ser, ^{Me}Phe) and six proteinogenic amino acids (Phe, Lys, Tyr, Pro, Asp, Cys) were chosen as the components of the cyclic N-methyl-peptides in this study (Figures 5A and 5B, ^{Me}cP-I–III). The respective peptides ^{Me}cP-I–III have two, three, and four N-methylated peptide bonds, respectively, implanted by the newly assigned ^{Me}aa in the codon table shown in Figure 1C. It should also be noted that ^{Me}cP-I and ^{Me}cP-II contain Pro as a proteinogenic component that also gives an additional secondary amide bond. These peptides were expressed in the presence of initiator tRNA$^{fMet}_{CAU}$ charged with N^{α}-ClAc-Phe and the corresponding ^{Me}aa-tRNAs in the above *w*PURE system. The MALDI-TOF analysis of the Flag-purified ^{Me}cP-I–III showed that all peptides were spontaneously cyclized via a thioether bond upon completion of translation, giving only the programmed cyclic N-methyl-peptides (Figure 5C). This study indicates that this technology is applicable to a wide array of cyclic peptides with multiple N-methyl-peptide bonds.

DISCUSSION

In the first stage of our proof-of-concept studies, we evaluated 19 ^{Me}aa derived from proteinogenic amino acids and 4 ^{Me}aa derived from nonproteinogenic amino acids, to determine the accessible ^{Me}aa. This study has revealed that the ribosome is accessible to ^{Me}aa with aromatic or nonbulky and noncharged side chains for the incorporation into the growing peptide. Most likely the N^{α}-methyl group added a steric bulkiness to the already bulky structure of certain amino acids, thereby reducing their incorporation levels. It is not yet clear what is the most important filter for the selection of accessible or inaccessible ^{Me}aa, but either EF-Tu or ribosome (or possibly both) is very likely the determinant(s) for their selection (Dale and Uhlenbeck, 2005). Therefore, it is of importance in future experiments to define the selection filter and engineer it to increase the tolerance toward inaccessible ^{Me}aa (Dedkova et al., 2003, 2006).

In this study, we depleted both aaRSs and cognate amino acids from the translation components in order to create vacant codon boxes for ^{Me}aa assignments. In previous studies, the genetic code reprogramming for nonproteinogenic amino acids bearing nonstandard side chains could be achieved by using a PURE system in which only proteinogenic amino acids, not aaRSs, were withdrawn (Murakami et al., 2006). We found,

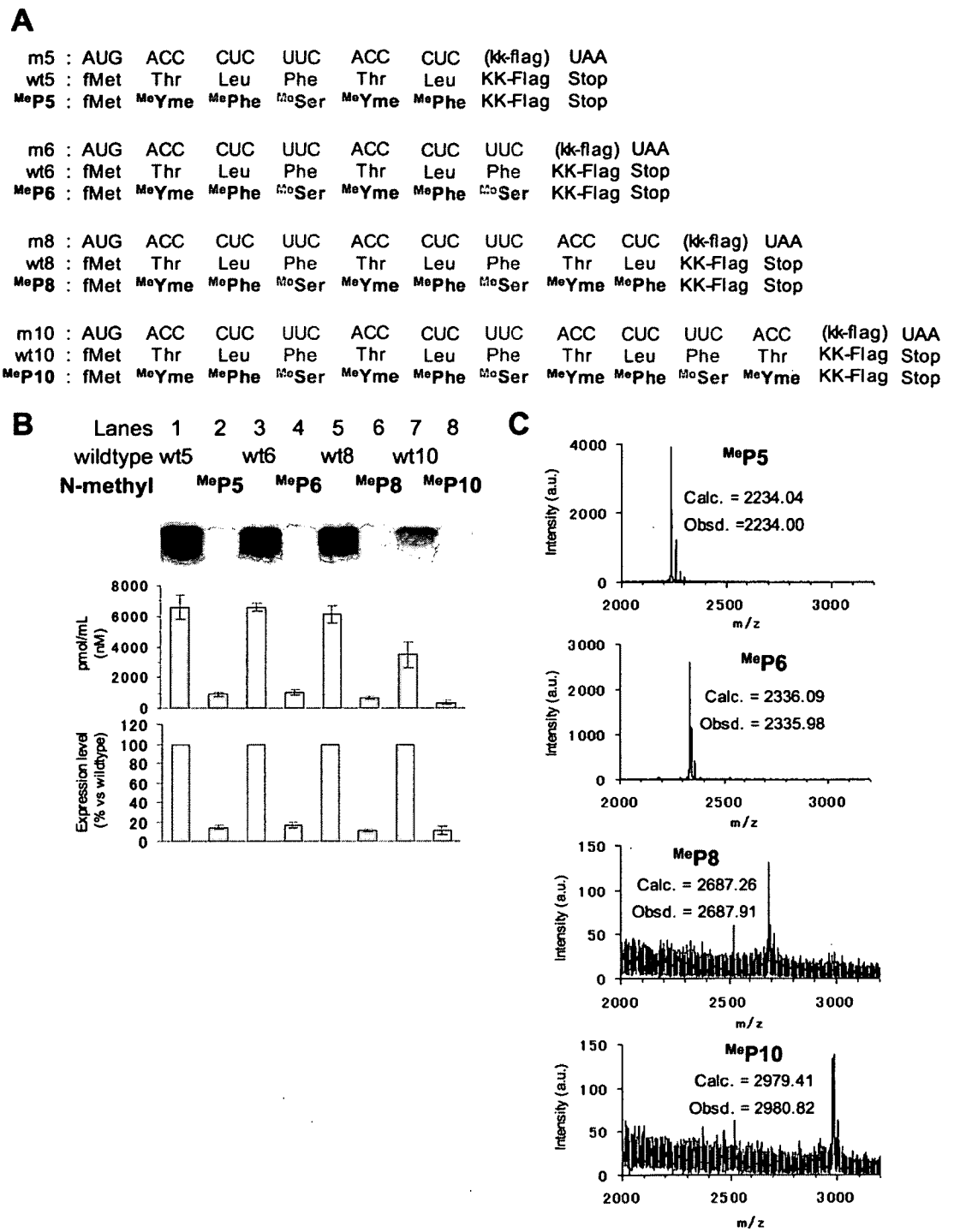

Figure 4. Expression of Longer N-Methyl-Peptides

(A) Sequences of mRNA templates (m5–10), control wild-type peptides (wt5–10), and peptides containing five, six, eight, or ten consecutive N-methyl-peptide bonds (^{Me}P5–10).

(B) Tricine-SDS-PAGE analysis of control wild-type peptides and N-methyl-peptides expressed from the respective mRNA. The peptides were labeled with [^{14}C]-Asp and detected by autoradiography. The wild-type peptides (lanes with odd numbers) were expressed in the ordinary PURE system, while N-methyl-peptides (lanes with even numbers) were expressed in the *w*PURE system containing ^{Me}Yme-tRNA$^{Asn\text{-}E1}_{GGU}$, ^{Me}Phe-tRNA$^{Asn\text{-}E1}_{GAG}$, and ^{Me}Ser-tRNA$^{Asn\text{-}E1}_{GAA}$. Absolute expression level of each peptide based on its observed radioisotope counts is shown in the upper graph, and relative expression level of each N-methyl-peptide against the corresponding wild-type peptide is shown in the lower graph. Each expression level was determined by a mean score of triplicates.

(C) MALDI-TOF-MS spectra of ^{Me}P5–10. The calculated mass (Calcd.) and observed mass (Obsd.) are shown in each spectrum.

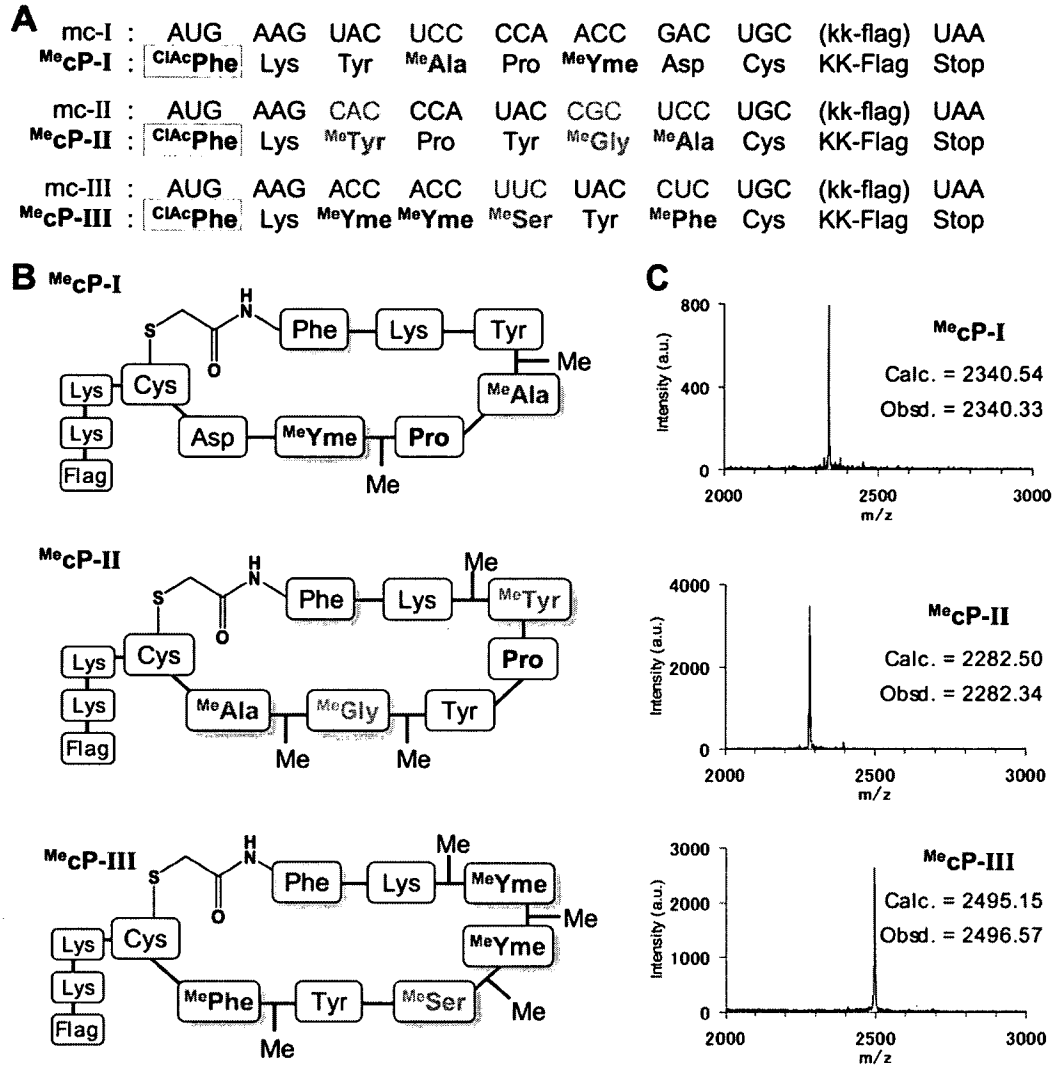

Figure 5. Synthesis of Cyclic N-Methyl-Peptides Closed by a Thioether Bond

(A) Sequences of mRNA templates (mc-I–III) and the precursor N-methyl-peptides (the linear version of ^{Me}cP-I–III). The codons and amino acids are colored according to reprogrammed genetic code table in Figure 1C.

(B) Structures of cyclic N-methyl-peptides. The locations of N-methyl groups on the peptide bond are shown in yellow.

(C) MALDI-TOF-MS spectra of cyclized ^{Me}cP-I–III. The calculated mass (Calcd.) and observed mass (Obsd.) are shown in each spectrum.

however, that in such a *w*PURE system, the incorporations of ^{Me}aa often suffered from contaminations caused by the competing incorporations of the proteinogenic amino acids to the assigned sites (data not shown). This was because that even though proteinogenic amino acids were supposedly depleted from the translation components, the recombinant translation factors and ribosome were often contaminated with small amounts of proteinogenic amino acids; therefore, they were charged onto the endogenous cognate tRNAs and readily competed out ^{Me}aa for the incorporation at the reprogrammed codon sites. Thus, it was critical to remove both aaRSs and proteinogenic amino acids from the translation components in order to repress the background incorporations of proteinogenic amino acids into the peptide. Moreover, we used orthogonal tRNAs as ^{Me}aa carriers aiming at eliminating potential mischarging events with proteinogenic amino acids catalyzed by aaRSs left in the *w*PURE system. Combining all of these controlled elements in the *w*PURE system, we were able to wipe out undesirable backgrounds and thus succeeded in synthesizing N-methyl-peptides as programmed by the mRNA.

Another key technology used in our proof-of-concept study is the flexizyme system. This system enables us to readily prepare a wide variety of ^{Me}aa-tRNAs with high purities, i.e., no contamination of the competing aa-tRNAs, so that we can ensure ^{Me}aa reassignments to the reprogrammed codons and thus determine the accessible ^{Me}aa in a quantitative manner. This technology has given us advantage over Merryman's method involving reductive alkylations on the precharged aa-tRNAs, where certain amounts of competing aa-tRNAs could remain unreacted and therefore contaminate the ^{Me}aa-tRNAs (Merryman and Green, 2004). As was discussed earlier, this might have led them to a different conclusion from ours on the accessible ^{Me}aa.

In a similar context, it would be difficult to perform ^{Me}aa assignments by the Szostak's method, where endogenous aaRSs or their mutants are used to mischarge nonproteinogenic amino acids onto endogenous tRNAs (Hartman et al., 2006; Josephson et al., 2005; Seebeck and Szostak, 2006). The major advantage of the Szostak's method is to in situ generate the mischarged aa-tRNAs and thus express the nonnatural peptides by simply adding nonproteinogenic amino acids to the *w*PURE system that lacks the corresponding proteinogenic amino acids. This system has been proven to be well suited for the incorporation of "highly efficient" nonproteinogenic amino acids into the target peptide, generally those structurally similar to proteinogenic amino acids that can readily compete out the background incorporations of amino acids originated by their contaminations (Seebeck and Szostak, 2006). Szostak et al. have also surveyed the aminoacylation ability of aaRSs toward ^{Me}aa, suggesting that ^{Me}Leu, ^{Me}Asp, ^{Me}Lys, ^{Me}His, and ^{Me}Trp could be charged onto the cognate tRNAs (Hartman et al., 2006). Unfortunately, among them only ^{Me}His and ^{Me}Trp were the "accessible" ^{Me}aa according to our experimental results shown in Figure 2C. As was discussed earlier, because of the intrinsic poor incorporation efficiency of ^{Me}aa, such background incorporations would give impure products containing certain or occasionally significant amounts of partially non-N-methylated peptides. In fact, by means of in situ aaRS's aminoacylation we found that ^{Me}His could be incorporated to the designated site of the same model peptide with competing incorporation of a smaller amount of His, whereas no incorporation of ^{Me}Trp, i.e., only incorporation of the competing Trp, was observed (T.K. and H.S., unpublished data).

Under the optimal conditions for the *w*PURE and flexizyme systems, we performed the mRNA-dependent synthesis of N-methyl-peptides containing more than four successive ^{Me}aa utilizing three different kinds of ^{Me}aa (Figures 3A and 4A). We monitored the expression level of each N-methyl-peptide by tricine-SDS-PAGE compared with the wild-type expression level based on their radioisotope intensities originating from the incorporation of [^{14}C]-Asp in the Flag peptide (Figures 3B and 4B). Using this strategy, we were able to directly compare the expression levels of the N-methyl-peptides in a quantitative manner.

Because we chose only accessible ^{Me}aa for the polymerization, we expected that the expression level would not be drastically reduced compared to the single incorporation. However, the expression of a tetra-N-methyl-peptide (^{Me}P4-I) gave only 20% of the wild-type expression level (Figure 3B, lanes 1 versus 5). Moreover, when longer N-methyl-peptides (from 5-mer to 10-mer) were expressed, their relative expression levels did not significantly change (Figure 4B). We wondered why the observed relative expression level of the tetra-N-methyl-peptides drastically dropped from that observed for the mono-N-methyl-peptide. To investigate this, we expressed mono-, di-, tri-, and tetra-N-methyl-peptides (Figure S3) (^{Me}P1-II, ^{Me}P2-II, ^{Me}P3-II, and ^{Me}P4-II) and compared their expression levels with those of the corresponding wild-type peptides (wt1-II, wt2-II, wt3-II, and wt4-II). Interestingly, the absolute expression levels of wt3-II and wt4-II were approximately three and four times higher, respectively, than that of wt1-II. On the other hand, the absolute expression levels of all N-methyl-peptides remained nearly the same. Thus, it became apparent that the relative expression levels of ^{Me}P3-II and ^{Me}P4-II to wt3-II and wt4-II appeared lower because the expression levels of wt3-II and wt4-II increased from wt1-II, whereas the absolute expression level of all N-methyl-peptides including 5- to 10-mers did not change significantly. Most importantly, MALDI-TOF analysis of the respective N-methyl-peptide gave the corresponding single peak with the correct molecular mass regardless of the length of the N-methyl-peptide. These findings led us to attempt the synthesis of cyclic-N-methyl-peptides.

In this application, we combined our technology for the ^{Me}aa incorporation into peptides with the method previously developed for the initiation reprogramming with N^{α}-ClAc-aa where a linear peptide containing Cys is expressed to afford the corresponding cyclic-peptide closed by a thioether bond between the N-terminus and the sulfhydryl group of the Cys. Because the thioether bond is nonreducible and thereby physiologically stable, it has been implanted into some therapeutic peptides in order to make them compatible for in vivo examination. Although such cyclic peptides were previously generated by only solid-phase chemistry, this method has enabled us to synthesize such cyclic peptides by means of the translation apparatus. Thus, the combination of the above two methods has offered us a platform technology for the synthesis of cyclic N-methyl-peptides by designing the corresponding mRNA templates. Remarkably, the expressed full-length cyclic N-methyl-peptides were highly pure in all three examples presented in this work. Thus, this technology is well suited for the preparation of libraries of cyclic N-methyl-peptides. Particularly, when this is coupled with in vitro display systems, such as ribosome (Hanes and Pluckthun, 1997) or mRNA displays (Nemoto et al., 1997; Roberts and Szostak, 1997), it becomes a very powerful means of screening large libraries and selecting in vivo-compatible cyclic N-methyl-peptides against therapeutic targets. Such experimental examples are shortly up coming from our laboratory.

SIGNIFICANCE

Reprogramming the genetic code allows for the simultaneous incorporation of multiple amino acid monomers with a wide variety of side chains into peptides through the ribosomal translation reaction. In this study, we achieved the mRNA-directed synthesis of various sequences and lengths of peptides containing multiple N-methylated amino acid residues upon reprogramming the genetic code by using the *w*PURE and flexizyme systems. Our demonstration shows the potential of this technology that leads us to a new strategy to construct libraries of linear and cyclic N-methyl-peptides simply by designing mRNA sequence libraries. Upon combining with appropriate screening strategies, this technology will yield a novel class of therapeutic peptides with protease resistance and cell permeability.

EXPERIMENTAL PROCEDURES

Materials and Synthesis of N^{α}-Methyl-Amino Acid Substrates

All N^{α}-methylated amino acids except N^{α}-methyl-L-cysteine and N^{α}-methyl-L-methionine were purchased from Watanabe Chemical, Japan. N^{α}-methyl-L-cysteine was synthesized by the procedure reported before from L-cysteine as a starting material (Park and Kim, 2002). N^{α}-tBoc-N^{α}-methyl-L-methionine was synthesized by the procedure reported before from N^{α}-tBoc-L-methionine as a starting material (Lee et al., 1990). N^{α}-methylated amino acids with the

aromatic side-chain (^{Me}Phe, ^{Me}Tyr, ^{Me}Trp, ^{Me}Fni, and ^{Me}Yme) were converted to cyanomethyl esters (CMEs), and the others were converted to 3,5-dinitrobenzyl esters (DBEs) by the procedure reported elsewhere (Murakami et al., 2006).

Synthesis of N^{α}-Methyl-Aminoacyl-tRNA

Aminoacylation reactions were generally performed as follows: 12.5 μl of 20 μM tRNA in 0.2 M HEPES-KOH (pH 7.5) was heated at 95°C for 1 min and cooled to room temperature over 5 min. Five microliters of 0.1 M $MgCl_2$ and 2.5 μl of 0.1 mM dFx or 5 μl of 3 M $MgCl_2$ and 2.5 μl of 0.1 mM eFx were added to the solution. Five microliters of 25 mM N^{α}-methylated amino acid substrate (DBE or CME) in DMSO was then added to the mixture and incubated on ice for 2–24 hr. For ^{Me}Thr, ^{Me}Tyr, ^{Me}Fni, and ^{Me}Yme, 5 μl of 200 mM substrate was used. The acylation reaction was quenched by addition of 150 μl of 0.6 M sodium acetate (pH 5.0), and the tRNA 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.

Ribosomal Synthesis of Linear N-Methylated Peptides and Wild-Type Peptides

The following conditions were used for the peptide synthesis without [^{14}C]-labeling. For the synthesis of N-methylated peptides (^{Me}P4-I, ^{Me}P4-II, ^{Me}P4-III, ^{Me}P4-IV, ^{Me}P5, ^{Me}P6, ^{Me}P8, ^{Me}P10, ^{Me}P1-II, ^{Me}P2-II, and ^{Me}P3-II), translation reaction was carried out by using the *w*PURE system (see the Supplemental Data) in the presence of 40 nM DNA template (one of the following templates: m4-I, m4-II, m4-III, m4-IV, m5, m6, m8, m10, m1-II, m2-II, or m3-II), 0.1 mM of the respective ^{Me}aa-tRNA(s), and 0.2 mM each of Met, Asp, Tyr, and Lys. For the synthesis of wild-type peptides (wt4-I, wt5, wt6, wt8, wt10, wt1-II, wt2-II, wt3-II, and wt4-II), translation reaction was carried out by using the ordinary PURE system in the presence of 40 nM DNA template (one of the following templates: m4-I, m5, m6, m8, m10, m1-II, m2-II, m3-II, or m4-II) and 0.2 mM each of Met, Thr, Leu, Phe, Asp, Tyr, and Lys. The translation reactions were assembled on ice and started by incubating at 37°C for 3 hr. For the detection of the peptides using autoradiography, see the section of tricine-SDS-PAGE analysis.

Tricine-SDS-PAGE Analysis

For tricine-SDS-PAGE analysis, peptides were synthesized by using 50 μM [^{14}C]-Asp in the place of 0.2 mM Asp. The translation reaction (2.5 μl) was stopped by the addition of 2.5 μl loading buffer consisting of 0.9 M Tris-HCl (pH 8.5), 30% glycerol, and 8% SDS for loading to 15% tricine-SDS-PAGE. Separation gel contained 15% acrylamide (acrylamide: bisacrylamide = 19:1), 1 M Tris•HCl (pH 8.5), 0.1% SDS, and 13% Glycerol, while the stacking gel contained 4% acrylamide (acrylamide: bisacrylamide = 29:1), 0.75 M Tris•HCl (pH 8.5), 0.075% SDS. The anode running buffer was 200 mM Tris•HCl (pH 8.9), while the cathode running buffer contains 100 mM Tris, 100 mM tricine, 0.1% SDS (the resulting buffer was approximately pH 8.3). The products, labeled with [^{14}C]-Asp in the C terminus FLAG peptide, were quantified by autoradiography with an image analyzer, FLA-5100 (Fuji, Japan), and the specific activity of [^{14}C]-Asp.

Ribosomal Synthesis of Cyclic N-Methylated Peptides

For the synthesis of cyclic N-methylated peptide (^{Me}Pc1, ^{Me}Pc2, or ^{Me}Pc3), translation reaction was carried out by using the *w*PURE system with 40 nM DNA template (mc1, mc2, or mc3) in the presence of 0.1 mM N^{α}-ClAc-Phe-tRNA$^{fMet}_{CAU}$, 0.1 mM of the respective ^{Me}aa-tRNAs, and 0.2 mM each of Met, Lys, Tyr, Pro, Asp, and Cys. The translation reactions were assembled on ice and started by incubation at 37°C for 3 hr.

MALDI-TOF-MS Analysis

The respective peptide was incubated in prewashed Flag-M2 agarose (Sigma) for an hour. After the resin was washed with TBS (50 mM Tris-HCl [pH 8.0], 150 mM NaCl), the peptide was eluted with 0.2% TFA. Eluted peptide was desalted with C18 zip tips (Millipore) and eluted with 50% acetonitrile, 0.1% TFA solution saturated with the matrix (*R*)-cyano-4-hydroxycinnamic acid. MALDI-TOF-MS analysis was performed by using Autoflex TOF/TOF (Bruker Daltonics) operated in the linear positive mode and externally calibrated with Substance P (1348.66 Da), Bombesin (1620.88 Da), ACTH clip 1–17 (2094.46 Da), and Somatostatin 28 (3149.61 Da) standards.

Supplemental Data

Supplemental Data include Supplemental Experimental Procedures for RNA synthesis, tRNA aminoacylation, and preparation of the translation systems and are available at http://www.chembiol.com/cgi/content/full/15/1/32/DC1/.

ACKNOWLEDGMENTS

We thank Y. Goto for providing N^{α}-ClAc-Phe-CME. We thank T.J. Kang, A. Ohta, and P.C. Reid for proofreading of this manuscript. 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. and a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO).

Received: September 19, 2006
Revised: November 25, 2007
Accepted: December 6, 2007
Published: January 25, 2008

REFERENCES

Bain, J.D., Diala, E.S., Glabe, C.G., Dix, T.A., and Chamberlin, A.R. (1989). Biosynthetic site-specific incorporation of a non-natural amino acid into a polypeptide. J. Am. Chem. Soc. *111*, 8013–8014.

Bain, J.D., Wacker, D.A., Kuo, E.E., and Chamberlin, A.R. (1991). Site-specific incorporation of non-natural residues into peptides: effect of residue structure on suppression and translation efficiencies. Tetrahedron *47*, 2389–2400.

Baltz, R.H. (2006). Molecular engineering approaches to peptide, polyketide and other antibiotics. Nat. Biotechnol. *24*, 1533–1540.

Billich, A., and Zocher, R. (1990). Formation of N-methylated peptide bonds in peptides and peptidols. In Biochemistry of Peptide Antibiotics, H. Kleinkauf and H. Von Doehren, eds. (Berlin: Walter de Gruyter & Co.), pp. 57–79.

Chung, H.H., Benson, D.R., and Schultz, P.G. (1993). Probing the structure and mechanism of Ras protein with an expanded genetic code. Science *259*, 806–809.

Dale, T., and Uhlenbeck, O.C. (2005). Amino acid specificity in translation. Trends Biochem. Sci. *30*, 659–665.

Dedkova, L.M., Fahmi, N.E., Golovine, S.Y., and Hecht, S.M. (2003). Enhanced D-amino acid incorporation into protein by modified ribosomes. J. Am. Chem. Soc. *125*, 6616–6617.

Dedkova, L.M., Fahmi, N.E., Golovine, S.Y., and Hecht, S.M. (2006). Construction of modified ribosomes for incorporation of d-amino acids into proteins. Biochemistry *45*, 15541–15551.

Ellman, J.A., Mendel, D., and Schultz, P.G. (1992). Site-specific incorporation of novel backbone structures into proteins. Science *255*, 197–200.

Fischbach, M.A., and Walsh, C.T. (2006). Assembly-line enzymology for polyketide and nonribosomal peptide antibiotics: logic, machinery, and mechanisms. Chem. Rev. *106*, 3468–3496.

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. USA *100*, 6353–6357.

Frankel, A., Millward, S.W., and Roberts, R.W. (2003). Encodamers: unnatural peptide oligomers encoded in RNA. Chem. Biol. *10*, 1043–1050.

Gilmore, M.A., Steward, L.E., and Chamberlin, A.R. (1999). Incorporation of noncoded amino acids by in vitro protein biosynthesis. Top. Curr. Chem. *202*, 77–99.

Goto, Y., Ohta, A., Sako, Y., Yamagishi, Y., Murakami, H., and Suga, H. (2008). Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. ACS Chem. Biol., in press.

Hahn, M., and Stachelhaus, T. (2006). Harnessing the potential of communication-mediating domains for the biocombinatorial synthesis of nonribosomal peptides. Proc. Natl. Acad. Sci. USA *103*, 275–280.

Hanes, J., and Pluckthun, A. (1997). In vitro selection and evolution of functional proteins by using ribosome display. Proc. Natl. Acad. Sci. USA *94*, 4937–4942.

Hartman, M.C., Josephson, K., and Szostak, J.W. (2006). Enzymatic aminoacylation of tRNA with unnatural amino acids. Proc. Natl. Acad. Sci. USA *103*, 4356–4361.

Hecht, S.M., Alford, B.L., Kuroda, Y., and Kitano, S. (1978). "Chemical aminoacylation" of tRNA's. J. Biol. Chem. *253*, 4517–4520.

Hendrickson, T.L., de Crecy-Lagard, V., and Schimmel, P. (2004). Incorporation of nonnatural amino acids into proteins. Annu. Rev. Biochem. *73*, 147–176.

Ho, T.-I., and Chow, Y.L. (1996). Photochemistry of nitro and nitroso compounds. In The Chemistry of Amino, Nitroso, Nitro and Related Groups. Supplement F2, S. Patai, ed. (New York: Wiley), pp. 757–821.

Hohsaka, T., and Sisido, M. (2002). Incorporation of non-natural amino acids into proteins. Curr. Opin. Chem. Biol. *6*, 809–815.

Hornbogen, T., and Zocher, R. (2005). Biosynthesis of N-methylated peptides in fungi. In Handbook of Industrial Mycology, Z. An, ed. (New York: Marcel Dekker), pp. 449–477.

Josephson, K., Hartman, M.C., and Szostak, J.W. (2005). Ribosomal synthesis of unnatural peptides. J. Am. Chem. Soc. *127*, 11727–11735.

Karginov, V.A., Mamaev, S.V., An, H., Van Cleve, M.D., Hecht, S.M., Komatsoulis, G.A., and Abelson, J.N. (1997). Probing the role of an active site aspartic acid in dihydrofolate reductase. J. Am. Chem. Soc. *119*, 8166–8176.

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

LaRiviere, F.J., Wolfson, A.D., and Uhlenbeck, O.C. (2001). Uniform binding of aminoacyl-tRNAs to elongation factor Tu by thermodynamic compensation. Science *294*, 165–168.

Lee, N., and Suga, H. (2001). A minihelix-loop RNA acts as a trans-aminoacylation catalyst. RNA *7*, 1043–1051.

Lee, J.P., Dunlap, B., and Rich, D.H. (1990). Synthesis and immunosuppressive activities of conformationally restricted cyclosporin lactam analogues. Int. J. Pept. Protein Res. *35*, 481–494.

Link, A.J., Mock, M.L., and Tirrell, D.A. (2003). Non-canonical amino acids in protein engineering. Curr. Opin. Biotechnol. *14*, 603–609.

Mendel, D., Cornish, V.W., and Schultz, P.G. (1995). Site-directed mutagenesis with an expanded genetic code. Annu. Rev. Biophys. Biomol. Struct. *24*, 435–462.

Merryman, C., and Green, R. (2004). Transformation of aminoacyl tRNAs for the in vitro selection of "drug-like" molecules. Chem. Biol. *11*, 575–582.

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

Murakami, H., Saito, H., and Suga, H. (2003b). A versatile tRNA aminoacylation catalyst based on RNA. Chem. Biol. *10*, 655–662.

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.

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.

Noren, C.J., Anthony-Cahill, S.J., Griffith, M.C., and Schultz, P.G. (1989). A general method for site-specific incorporation of unnatural amino acids into proteins. Science *244*, 182–188.

Ohta, A., Murakami, H., Eri, H., and Suga, H. (2007). Synthesis of polyester using genetic code reprogramming. Chem. Biol. *14*, 1315–1322.

Ohuchi, M., Murakami, H., and Suga, H. (2007). The flexizyme system: a highly flexible tRNA aminoacylation tool for the translation apparatus. Curr. Opin. Chem. Biol. *11*, 537–542.

Park, J.D., and Kim, D.H. (2002). Cysteine derivatives as inhibitors for carboxypeptidase A: synthesis and structure-activity relationships. J. Med. Chem. *45*, 911–918.

Roberts, R.W., and Szostak, J.W. (1997). RNA-peptide fusions for the in vitro selection of peptides and proteins. Proc. Natl. Acad. Sci. USA *94*, 12297–12302.

Robertson, S.A., Ellman, J.A., and Schultz, P.G. (1991). A general and efficient route for chemical aminoacylation of transfer RNAs. J. Am. Chem. Soc. *113*, 2722–2729.

Sagan, S., Karoyan, P., Lequin, O., Chassaing, G., and Lavielle, S. (2004). N- and Calpha-methylation in biologically active peptides: synthesis, structural and functional aspects. Curr. Med. Chem. *11*, 2799–2822.

Saito, H., Kourouklis, D., and Suga, H. (2001). An in vitro evolved precursor tRNA with aminoacylation activity. EMBO J. *20*, 1797–1806.

Seebeck, F.P., and Szostak, J.W. (2006). Ribosomal synthesis of dehydroalanine-containing peptides. J. Am. Chem. Soc. *128*, 7150–7151.

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.

Short, G.F., 3rd, Laikhter, A.L., Lodder, M., Shayo, Y., Arslan, T., and Hecht, S.M. (2000). Probing the S1/S1′ substrate binding pocket geometry of HIV-1 protease with modified aspartic acid analogues. Biochemistry *39*, 8768–8781.

Sieber, S.A., and Marahiel, M.A. (2005). Molecular mechanisms underlying nonribosomal peptide synthesis: approaches to new antibiotics. Chem. Rev. *105*, 715–738.

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.

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.

Walsh, C.T., Chen, H., Keating, T.A., Hubbard, B.K., Losey, H.C., Luo, L., Marshall, C.G., Miller, D.A., and Patel, H.M. (2001). Tailoring enzymes that modify nonribosomal peptides during and after chain elongation on NRPS assembly lines. Curr. Opin. Chem. Biol. *5*, 525–534.

Wang, L., and Schultz, P.G. (2004). Expanding the genetic code. Angew. Chem. Int. Ed. Engl. *44*, 34–66.

ACS chemical biology

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

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.

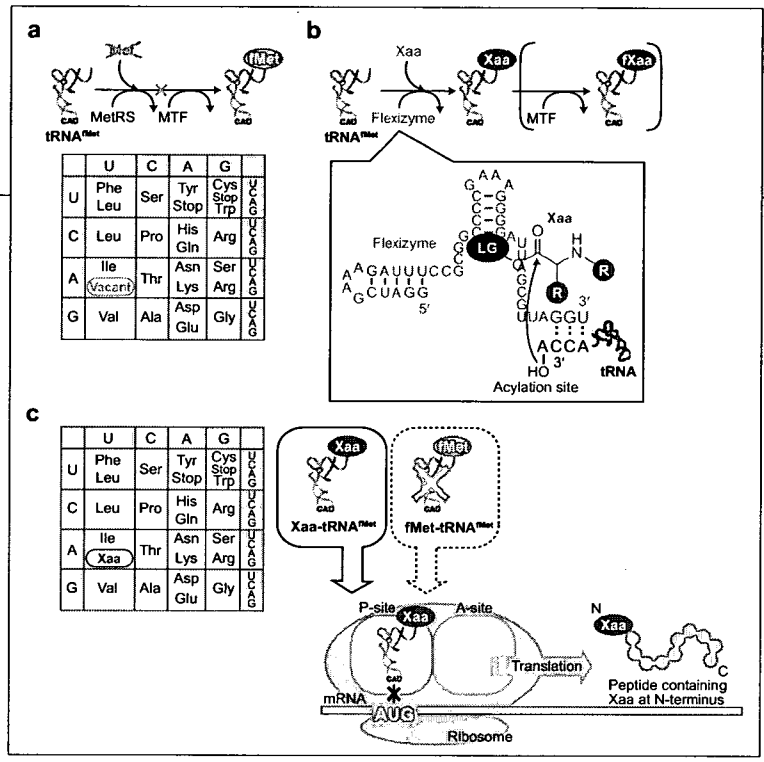

Figure 1. Engineering the initiation event by genetic code reprogramming. a) Codon table in the absence of Met. Upon removing Met from the translation mixture, $tRNA^{fMet}_{CAU}$ is no longer available for methionylation by MetRS, consequently making the start codon vacant, *i.e.*, *w*PURE system. b) 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

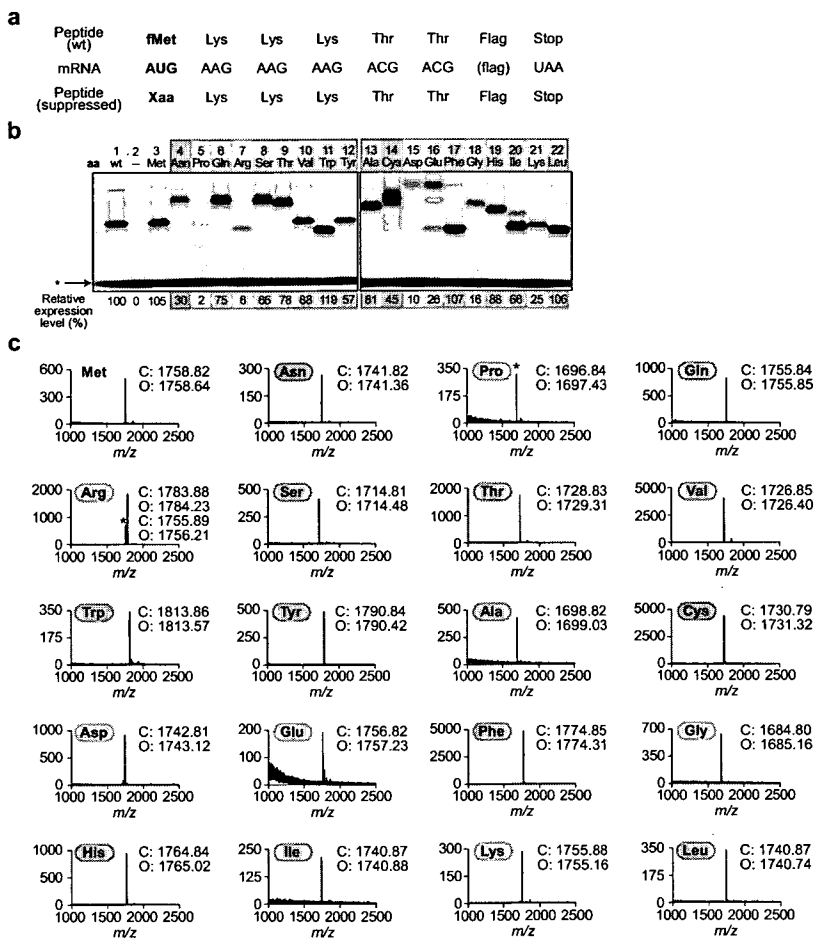

Figure 2. Tolerance of various amino acids in initiation. a) The mRNA sequence that expresses peptides initiated by various amino acids. Flag in parentheses indicates the RNA sequence encoding a Flag peptide (DYKDDDDK). b) Tricine-SDS PAGE analysis of the translation products. Lane 1, expression of wild 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.

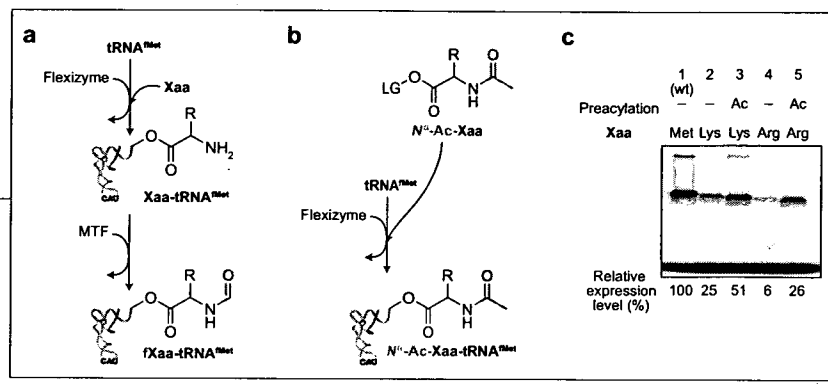

Figure 3. Increase in expression level by N^{α}-acylation. a) Formylation of Xaa-tRNA$^{fMet}_{CAU}$ in the *w*PURE system. When Xaa-tRNA$^{fMet}_{CAU}$ is used in the initiation, it is generally formylated. b) Amino-acylation of tRNA$^{fMet}_{CAU}$ with preacetylated amino acids by the flexizyme system. 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

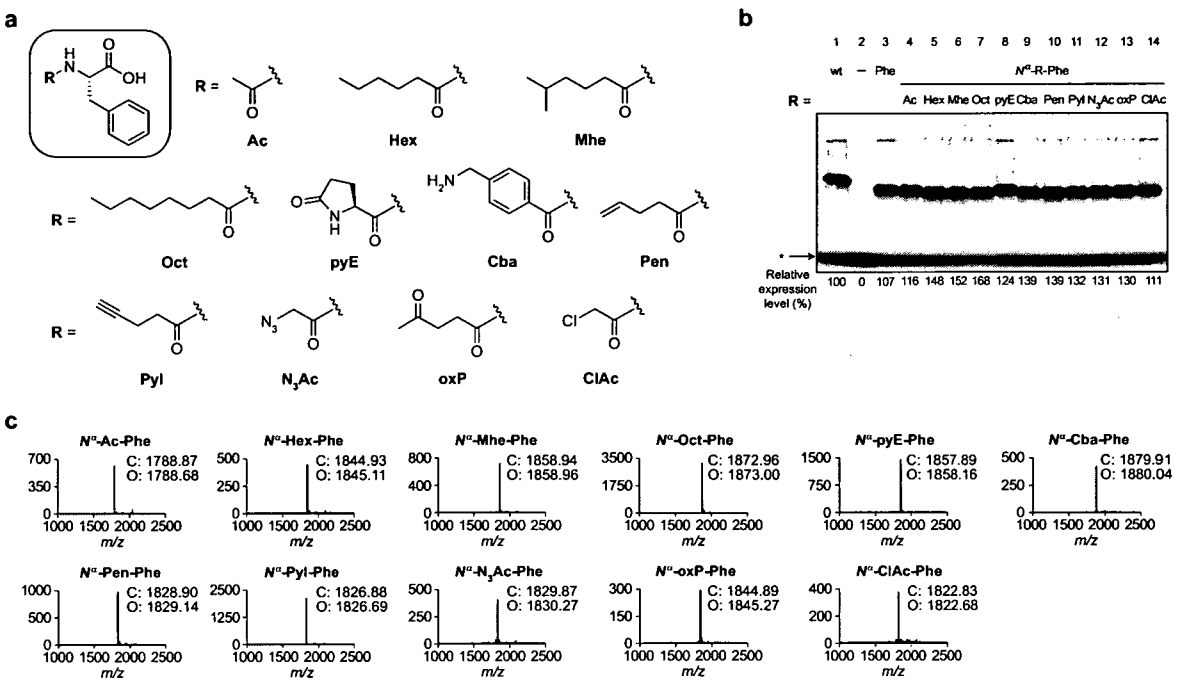

Figure 4. Tolerance of N^{α}-acyl groups in initiation. a) N^{α}-acyl phenylalanine derivatives used in this study. b) Tricine−SDS PAGE analysis of the translation products initiated with various N^{α}-acyl phenylalanine derivatives. Lane 1, expression of wild type; lane 2, in the absence of Met; lane 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 with N^{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

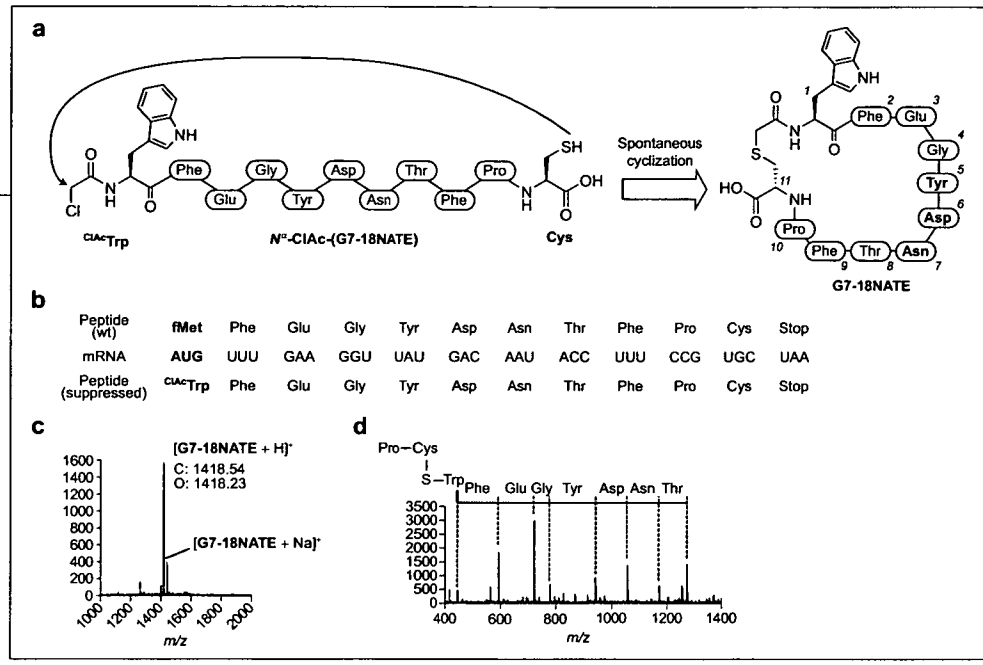

Figure 5. Ribosomal synthesis of G7−18NATE. a) The spontaneous cyclization affording G7−18NATE. The precursor linear peptide, N^{α}-ClAc-G7−18NATE, translated in the *w*PURE system spontaneously cyclizes into G7-18NATE by intramolecular attack of Cys sulfhydryl group to the α-carbon of the N^{α}-ClAc 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.

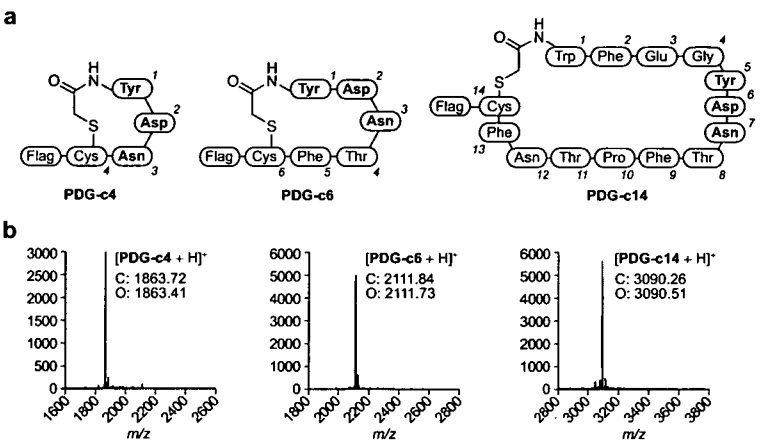

Figure 6. Custom synthesis of the cyclic peptides with various ring sizes. a) Structure of the expressed cyclic peptides. Three amino acids known to be critical residues in G7−18NATE are shown in bold, and "Flag-tag" indicates a modified Flag peptide sequence (LTTDYKDDDDK). b) MALDI-TOF mass spectrum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

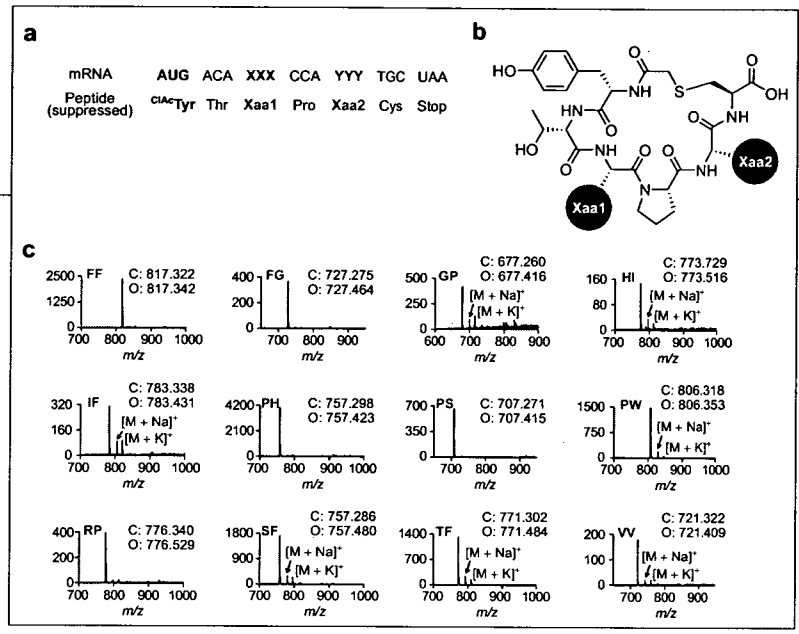

Figure 7. Construction of a cyclic peptide library. a) The sequence of mRNA used in the library synthesis. The codons indicated in red (Xaa1) and blue (Xaa2) were changed in each expression. The start codon is suppressed with N^{α}-ClAc-Tyr. b) The structure of cyclic peptide contained in the peptide 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.

Hindawi Publishing Corporation
EURASIP Journal on Bioinformatics and Systems Biology
Volume 2007, Article ID 47214, 11 pages
doi:10.1155/2007/47214

Research Article

Gene Systems Network Inferred from Expression Profiles in Hepatocellular Carcinogenesis by Graphical Gaussian Model

Sachiyo Aburatani,[1] Fuyan Sun,[1] Shigeru Saito,[2] Masao Honda,[3] Shu-ichi Kaneko,[3] and Katsuhisa Horimoto[1]

[1] *Biological Network Team, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-42 Aomi, Koto-ku, Tokyo 135-0064, Japan*
[2] *Chemo & Bio Informatics Department, INFOCOM CORPORATION, Mitsui Sumitomo Insurance Surugadai Annex Building, 3-11, Kanda-Surugadai, Chiyoda-ku, Tokyo 101-0062, Japan*
[3] *Department of Gastroenterology, Graduate School of Medical Science, Kanazawa University, 13-1 Takara-machi, Kanazawa, Ishikawa 920-8641, Japan*

Received 28 June 2006; Revised 27 February 2007; Accepted 1 May 2007

Recommended by Paul Dan Cristea

Hepatocellular carcinoma (HCC) in a liver with advanced-stage chronic hepatitis C (CHC) is induced by hepatitis C virus, which chronically infects about 170 million people worldwide. To elucidate the associations between gene groups in hepatocellular carcinogenesis, we analyzed the profiles of the genes characteristically expressed in the CHC and HCC cell stages by a statistical method for inferring the network between gene systems based on the graphical Gaussian model. A systematic evaluation of the inferred network in terms of the biological knowledge revealed that the inferred network was strongly involved in the known gene-gene interactions with high significance ($P < 10^{-4}$), and that the clusters characterized by different cancer-related responses were associated with those of the gene groups related to metabolic pathways and morphological events. Although some relationships in the network remain to be interpreted, the analyses revealed a snapshot of the orchestrated expression of cancer-related groups and some pathways related with metabolisms and morphological events in hepatocellular carcinogenesis, and thus provide possible clues on the disease mechanism and insights that address the gap between molecular and clinical assessments.

1. INTRODUCTION

Hepatitis C virus (HCV) is the major etiologic agent of non-A non-B hepatitis, and chronically infects about 170 million people worldwide [1–3]. Many HCV carriers develop chronic hepatitis C (CHC), and finally are afflicted with hepatocellular carcinoma (HCC) in livers with advanced-stage CHC. Thus, the CHC and HCC cell stages are essential in hepatocellular carcinogenesis.

To elucidate the mechanism of hepatocellular carcinogenesis at a molecular level, many experiments have been performed from various approaches. In particular, recent advances in techniques to monitor simultaneously the expression levels of genes on a genomic scale have facilitated the identification of genes involved in the tumorigenesis [4]. Indeed, some relationships between the disease and the tumor-related genes were proposed from the gene expression analyses [5–7]. Apart from the relationship between tumor-related genes and the disease at the molecular level, the information about the pathogenesis and the clinical characteristics of hepatocellular carcinogenesis has accumulated steadily [8, 9]. However, there is a gap between the information about hepatocellular carcinogenesis at the molecular level and that at more macroscopic levels, such as the clinical level. Furthermore, the relationships between tumor-related genes and other genes also remain to be investigated. Thus, an approach to describe the perspective of carcinogenesis from measurements at the molecular level is desirable to bridge the gap between the information at the two different levels.

Recently, we have developed an approach to infer a regulatory network, which is based on graphical Gaussian modeling (GGM) [10, 11]. Graphical Gaussian modeling is one of the graphical models that includes the Boolean and Bayesian models [12, 13]. Among the graphical models, GGM has the simplest structure in a mathematical sense; only the inverse of the correlation coefficient between the variables is needed,