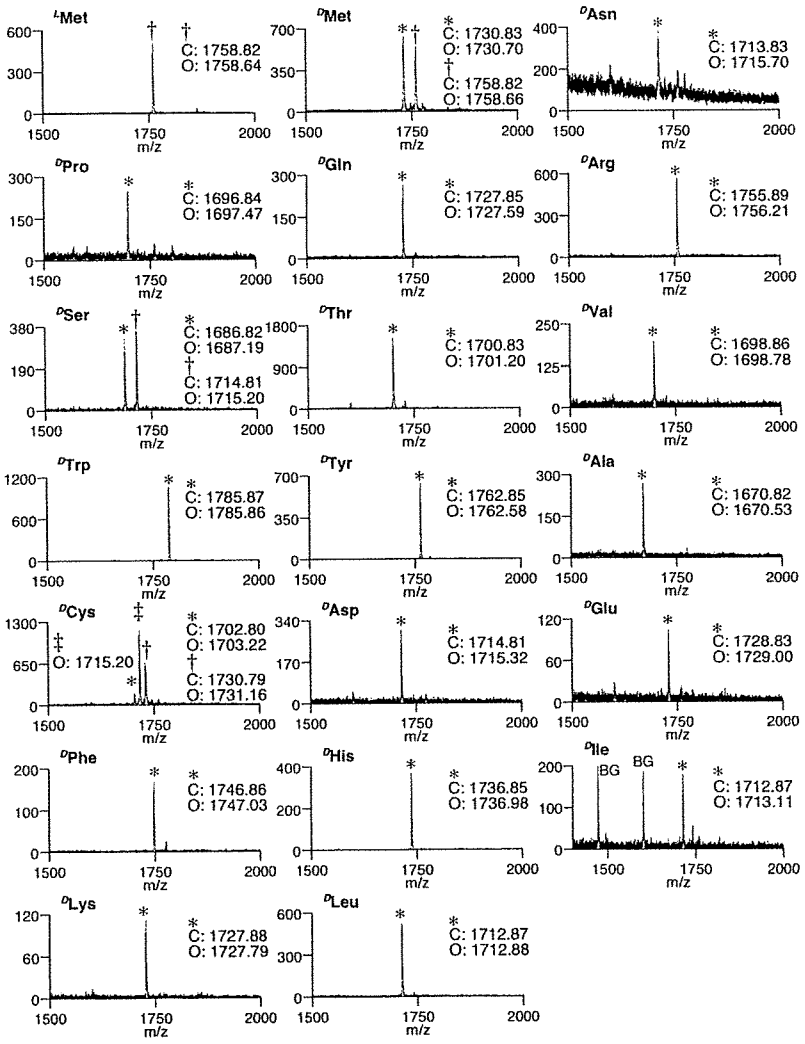


**FIGURE 3.** MALDI-TOF mass spectra of the translated peptides. * and † indicate a peak corresponding to the H- and f-peptide, respectively. "BG" indicates peaks corresponding to the background expression originated in-frame misinitiation. ‡ indicates the peak corresponding to f-$^{D}$Ser-peptide proposed in this work. The calculated mass (C) and observed mass (O) are shown in each spectrum.

tricine-SDS PAGE analysis (Fig. 2B). Three different ratios, 99:1, 95:5, and 80:20, of the mixtures of the $^{D}$aa-tRNA$^{fMet}_{CAU}$ and $^{L}$aa-tRNA$^{fMet}_{CAU}$ were prepared for the expression of peptides and each translation product was analyzed by MALDI-TOF mass spectrometry (Fig. 4). A 1% *L*-contamination to the sample of $^{D}$Tyr and $^{D}$Trp gave a single major peak corresponding to H-peptide, but a very tiny peak of f-peptide was accompanied in the case of $^{D}$Trp-initiation. On the other hand, even with a 1% contamination of $^{L}$Leu, $^{D}$Leu suffered from a nearly 1:1 mixture of intense peaks of H- and f-peptides. A 5% *L*-contamination gave a significant increase in the relative intensity of the peaks corresponding to f-peptides in the cases of both $^{D}$Trp and $^{D}$Leu, while $^{D}$Tyr-initiation yielded only a minor peak of f-peptide. In the case of 20% L-contamination, f-$^{L}$Trp- and f-$^{L}$Leu-initiations almost competed out the corresponding $^{D}$aa-initiation, and even for $^{D}$Tyr the peak ratio of H- and f-peptides became nearly 1:1. These results clearly indicated that even if a small amount, as little as 1%, of $^{L}$aa-tRNA$^{fMet}_{CAU}$ is contaminated in $^{D}$aa-tRNA$^{fMet}_{CAU}$, the peptide product would suffer from the formation of f-peptide. Since we observed a single peak of H-peptide in the 16 cases in Figure 3, these products should be assigned to H-$^{D}$aa-peptides. We thus concluded that the racemization of $^{D}$aa-tRNA$^{fMet}_{CAU}$ did not occur during the translation in these cases; even if it occurred, its degree should be far less than 1%, which would be negligible.

## $^{D}$Met-, $^{D}$Ser-, and $^{D}$Cys-tRNA$^{fMet}_{CAU}$ act as modest substrates of MTF

The above studies so far supported that racemization of $^{D}$aa-tRNA$^{fMet}_{CAU}$ does not generally occur during the translation. However, in the case of initiation with $^{D}$Met-, $^{D}$Cys-, or $^{D}$Ser-tRNA$^{fMet}_{CAU}$ we did observe two peaks corresponding to H-peptide and f-peptide (for $^{D}$Cys there was an additional peak). Therefore, it was still necessary to verify the possibility of their racemization by another method. Accordingly, we decided to use the well-known fact that $^{D}$aa are generally poor substrates in the elongation event using the amber suppression method (Noren et al. 1989; Bain et al. 1991; Ellman et al. 1992; Starck et al. 2003; Tan et al. 2004; Murakami et al. 2006). Since the flexizyme system is able to afford the $^{D}$aa-tRNA molecules for initiation and elongation (charged onto tRNA$^{fMet}_{CAU}$ and tRNA$^{AsnE-2}_{CUA}$ [Ohta et al. 2007], respectively) with exactly the same quality, we should be able to verify the occurrence of racemization by running the peptide synthesis with the $^{D}$aa- or $^{L}$aa-tRNA$^{fMet}_{CAU}$ initiation and $^{D}$aa- or $^{L}$aa-tRNA$^{AsnE-2}_{CUA}$ elongation side by side. Thus, in parallel to the translation of the mRNA previously designed for initiating with the $^{D}$aa-tRNA$^{fMet}_{CAU}$, we designed an mRNA template containing the amber codon and expressed the peptide in the presence of $^{D}$aa-tRNA$^{AsnE-2}_{CUA}$ (Fig. 5A).

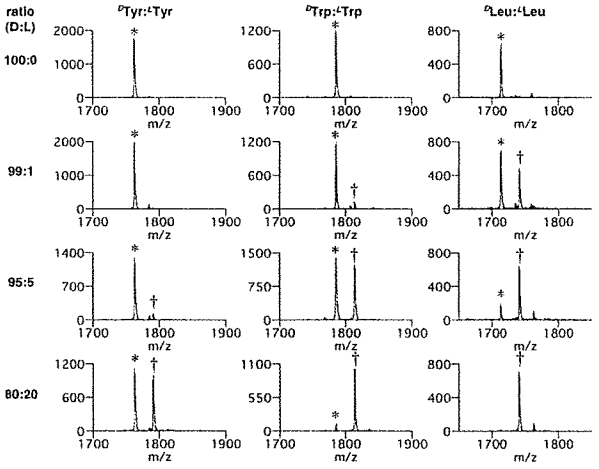


**FIGURE 4.** MALDI-TOF mass spectra of the peptides initiated with the mixture of $^{D}$aa-tRNA$^{fMet}_{CAU}$ and $^{L}$aa-tRNA$^{fMet}_{CAU}$. * and † indicate a peak corresponding to the H- and f-peptide, respectively.

Both $^{L}$Met-tRNA$^{fMet}_{CAU}$ and $^{L}$Met-tRNA$^{AsnE\text{-}2}_{CUA}$ functioned as the translation initiator and elongator, respectively, for the cognate mRNA templates, giving the expected full-length peptides (Fig. 5B, lanes 1,3). On the other hand, $^{D}$Met-tRNA$^{fMet}_{CAU}$ could initiate the translation whereas $^{D}$Met-tRNA$^{AsnE\text{-}2}_{CUA}$ could not suppress the amber codon (Fig. 5B, lanes 2,4). This clearly indicated that $^{D}$Met-tRNA$^{AsnE\text{-}2}_{CUA}$ was not racemized during the translation, and therefore it was reasonable to assume that $^{D}$Met-tRNA$^{fMet}_{CAU}$ was not either. Likewise, $^{D}$Cys was not incorporated into the nascent peptide chain by amber suppression (Fig. 5B, lanes 5–8), suggesting that $^{D}$Cys-tRNA$^{AsnE\text{-}2}_{CUA}$ was not racemized. This result also indicated that $^{D}$Cys-tRNA$^{fMet}_{CAU}$ was not racemized and primed the translation with $^{D}$Cys. To our surprise, the amber suppression by $^{D}$Ser-tRNA$^{AsnE\text{-}2}_{CUA}$ gave an evident band as intense as that by $^{L}$Ser-tRNA$^{AsnE\text{-}2}_{CUA}$; however, their mobility were slightly different (Fig. 5B, cf. lanes 11 and 12). MALDI-TOF analysis of the respective peptide products expressed in the presence of $^{D}$Ser-tRNA$^{AsnE\text{-}2}_{CUA}$ and $^{L}$Ser-tRNA$^{AsnE\text{-}2}_{CUA}$ showed the same expected molecular weight, implying that these peptides have the same compositions of sequence (Supplemental Fig. S2). These results suggested that the observed difference in mobility between lanes 11 and 12 (Fig. 5B) could be attributed to the difference in chirality of the Ser residue, i.e., $^{D}$Ser-tRNA$^{AsnE\text{-}2}_{CUA}$ elongated the peptide chain without racemization. This result assured that $^{D}$Ser-tRNA$^{fMet}_{CAU}$ initiated the translation without racemization.

To this end, we concluded that neither $^{D}$Met-, $^{D}$Cys-, nor $^{D}$Ser-tRNA$^{fMet}_{CAU}$ were racemized, but they were partially formylated by MTF and competitively initiated the translation. Even though the f-$^{D}$aa-tRNA$^{fMet}_{CAU}$ was formed presumably in only a small amount, its higher initiation efficiency compared to $^{D}$aa-tRNA$^{fMet}_{CAU}$ could affect the outcome of translation. Consequently, we obtained a mixture of the respective H-$^{D}$aa-peptide and f-$^{D}$aa-peptide.

MALDI-TOF analysis of the $^{D}$Cys-primed peptides showed two peaks, of which the observed molecular weights were consistent with H-$^{D}$Cys- and f-$^{D}$Cys-peptides like $^{D}$Met and $^{D}$Ser. However, there is an additional peak with the molecular weight of 1715.20 Da (Fig. 3, $^{D}$Cys, see peak indicated by ‡). Although it is difficult to define what exactly this peptide is, the molecular weight is consistent with f-$^{D}$Ser-peptide (the calculated molecular weight is 1714.81 Da). We therefore suggest the following mechanism as a possible scenario to convert $^{D}$Cys-primed peptide to f-$^{D}$Ser-peptide (Supplemental Fig. S3). The α-amino group of $^{D}$Cys-tRNA$^{fMet}_{CAU}$ could be formylated by MTF similar to $^{D}$Met and $^{D}$Ser, while the sulfhydryl side-chain group could also be formylated because of its inherent low pKa and high nucleophilicity. The formylated sulfhydryl group then might become a good leaving group so that the oxygen of $N^{\alpha}$-formyl group likely could attack the β-carbon of the side chain to yield an oxazoline-containing peptide. Hydrolysis of the oxazoline ring consequently might yield f-$^{D}$Ser-peptide. Note that this unusual event occurred in only the case of initiation with $^{D}$Cys, not $^{L}$Cys. Therefore, the *D*-configuration of $^{D}$Cys likely played a critical role in processing this unusual and interesting chemistry.

## Preacylation of the α-amino group on $^{D}$aa enhances initiation efficiency

In our previous work on reprogramming the initiation using various $^{L}$aa, we have found that preacylation on $^{L}$aa-tRNA$^{fMet}_{CAU}$, e.g., $N^{\alpha}$-acetylation, significantly enhances the initiation efficiency (Goto et al. 2008). Therefore, we wondered whether the same trend could be observed for $^{D}$aa-tRNA$^{fMet}_{CAU}$ initiation. Three $^{D}$aa with moderate initiation efficiencies ($^{D}$Met, $^{D}$Trp, and $^{D}$Phe; 28%, 19%, and 38%, respectively) and two $^{D}$aa with low efficiencies ($^{D}$Asn and $^{D}$Ala; 0.4% and 3%, respectively) were chosen for the synthesis of the corresponding Ac-$^{D}$aa substrates and charged onto tRNA$^{fMet}_{CAU}$ using the flexizyme system. The resulting Ac-$^{D}$aa-tRNA$^{fMet}_{CAU}$ was used to initiate the translation to produce each peptide in parallel to the initiation with the corresponding $^{D}$aa-tRNA$^{fMet}_{CAU}$ (Fig. 6).

In all cases, the initiation with Ac-$^{D}$aa-tRNA$^{fMet}_{CAU}$ dramatically increased the expression level compared to that with $^{D}$aa-tRNA$^{fMet}_{CAU}$; particularly, the preacetylation of $^{D}$Asn and $^{D}$Ala increased the expression level from nearly invisible band intensity to a clearly visible intensity (Fig. 6, lanes 8–11, 0.4%–31% and 3%–56%, respectively). MALDI-TOF analyses of the respective peptides were also consistent with the expected molecular weights of the Ac-$^{D}$aa-peptides (Supplemental Fig. S4), indicating that Ac-$^{D}$aa-tRNA$^{fMet}_{CAU}$ exclusively initiated the translation.

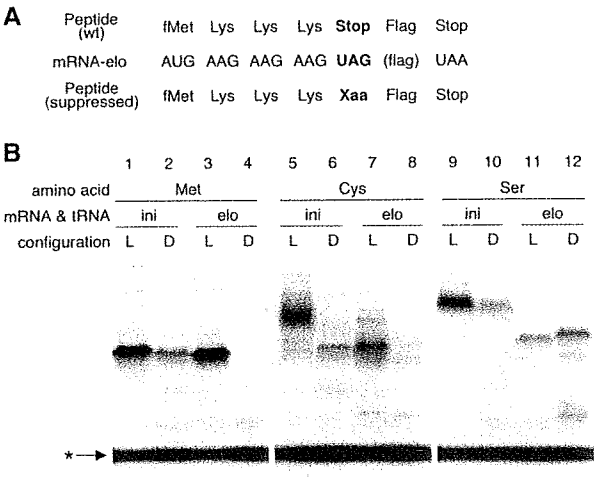


**FIGURE 5.** Determination of the configuration of the N-terminal residue. (*A*) The mRNA sequence used for the expression with amber suppression. UAG codon indicated in bold was suppressed with various amino acids. Flag in parentheses indicates the RNA sequence encoding the Flag peptide sequence (DYKDDDDK). (*B*) Tricine-SDS PAGE analysis of the translation products. (Lanes *1,5,9*) reprogrammed initiation with $^{L}$aa; (lanes *2,6,10*) reprogrammed initiation with $^{D}$aa; (lanes *3,7,11*) amber suppression with $^{L}$aa; (lanes *4,8,12*) amber suppression with $^{D}$aa. The types of side chain are Met in lanes 1–4, Cys in lanes 5–8, and Ser in lanes 9–12. The band indicated by an asterisk corresponds to free [$^{14}$C]-Asp that remained unincorporated into the Flag peptide.

## DISCUSSION

By means of the genetic code reprogramming for initiation, we have shown that the translation apparatus is able to use $^{D}$aa in the initiation event even though its efficiency varies depending upon the type of side chain. The most intriguing finding is that in most cases the peptide product had a free N terminus, i.e., without a formyl group, indicating that MTF does not generally formylate the $^{D}$aa-tRNA$^{fMet}_{CAU}$ (Fig. 7). Moreover, the use of Ac-$^{D}$aa-tRNA$^{fMet}_{CAU}$, an analog of formylated initiator, is able to significantly increase the initiation efficiency, giving the Ac-$^{D}$aa-peptide with a higher yield than that initiated with $^{D}$aa-tRNA$^{fMet}_{CAU}$. This suggests that the preacylation of $^{D}$aa-tRNA$^{fMet}_{CAU}$ is able to overcome the inherent modest initiation efficiency of $^{D}$aa controlled by MTF.

We have shown that the translation apparatus tolerates $^{D}$aa for initiation to afford H-$^{D}$aa-peptide or Ac-$^{D}$aa-peptide when $^{D}$aa-tRNA$^{fMet}_{CAU}$ or Ac-$^{D}$aa-tRNA$^{fMet}_{CAU}$ is given to the Met-withdrawn *w*PURE system. This is in sharp contrast to the little success achieved in the incorporation of $^{D}$aa into the nascent peptide chain via amber-suppression elongation (Fig. 7). Why so? To provide a definitive answer(s) to this question, more detailed mechanistic investigations are certainly required; but at the present stage we are able to suggest the following three potential reasons.

The first reason can be attributed to the nature of the mechanism for the fidelity control, which relies upon the function of initiation factor 2 (IF-2). Despite the fact that MetRS charges $^{L}$Met onto both initiator tRNA$^{fMet}_{CAU}$ and elongator tRNA$^{Met}_{CAU}$, IF-2 recognizes only f-$^{L}$Met-tRNA$^{fMet}_{CAU}$ over $^{L}$Met-tRNA$^{Met}_{CAU}$ (Schmitt et al. 1996; Boelens and Gualerzi 2002). The differences between these two $^{L}$Met-tRNAs lies in the tRNA's body sequence (particularly that the 5′-terminal nucleotide is unpaired in the initiator whereas it is paired in the elongator) (Mayer et al. 2001) and the formyl group on $^{L}$Met. Our previous and current studies of the reprogrammed initiation have shown that the formylated aminoacyl-tRNA$^{fMet}_{CAU}$ very likely recruits IF-2 more efficiently than nonformylated, yet the formyl group is not the essential selection element. Apparently, a more critical selection element is the structural features of the initiator tRNA$^{fMet}_{CAU}$ distinct from the elongator tRNA$^{Met}_{CAU}$ (Mayer et al. 2003). Our results show that IF-2 likely tolerates not only various $^{L}$aa but also even $^{D}$aa (particularly when $N^{\alpha}$-acylated) as far as they are charged onto tRNA$^{fMet}_{CAU}$, suggesting that its recognition of side chains and the chirality of amino acids are less strict compared with that of the tRNA's body sequence. On the other hand, in the elongation event, it has been firmly established that EF-Tu discriminates cognate pairs of $^{L}$aa and elongator tRNA against noncognate pairs upon recruiting them to the ribosome elongation complex with the strict control of a nearly uniform affinity toward the cognate pairs (Stanzel et al. 1994; Ibba and Söll 1999; LaRiviere et al. 2001). This mechanism is also applicable to pairs of nonproteinogenic amino acids and amber or other possible elongator tRNAs; particularly, the opposite chirality in $^{D}$aa is forcefully rejected by this EF-Tu selection filter, resulting in the

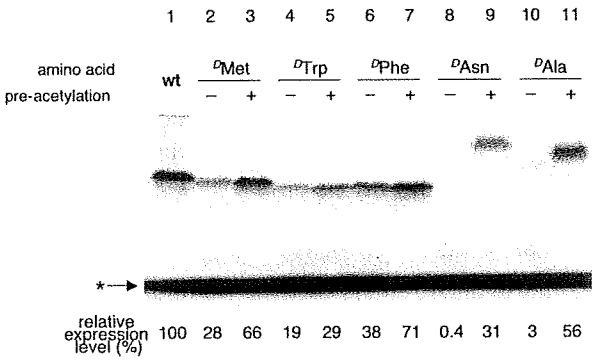


**FIGURE 6.** Tricine PAGE analysis of the translation products with $N^{\alpha}$-acylated initiators. (Lane *1*) expression of wild type; (lanes 2–11) reprogrammed initiation with Na-free amino acids or Na-acetyl amino acids. Each expression level relative to wild type is determined by a mean score of duplicates or more. The band indicated by an asterisk corresponds to free [$^{14}$C]-Asp that remained unincorporated into the Flag peptide.

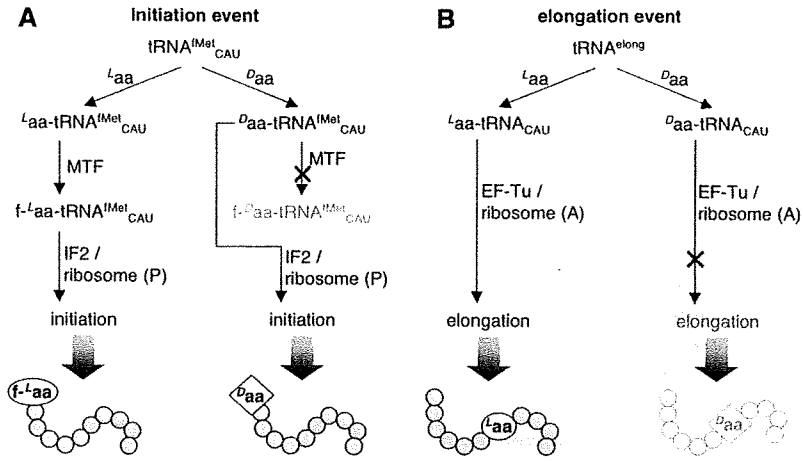


**FIGURE 7.** Initiation versus elongation with $^{L}$aa and $^{D}$aa. (*A*) $^{L}$aa charged onto tRNA$^{fMet}_{CAU}$ is formylated by MTF to give f-$^{L}$aa-tRNA$^{fMet}_{CAU}$; then IF2 brings it to the P site of the ribosome initiation complex and the translation is started to yield f-$^{L}$aa-peptide. On the other hand, $^{D}$aa charged onto tRNA$^{fMet}_{CAU}$ is not generally formylated, i.e., it bypasses the formylation step; yet IF2 is able to bind and bring $^{D}$aa-tRNA$^{fMet}_{CAU}$ to the P site of the ribosome initiation complex; thus the translation reaction can be initiated with $^{D}$aa, generally giving H-$^{D}$aa-peptide. (*B*) $^{L}$aa charged onto amber tRNA$_{CUA}$ binds to EF-Tu and goes to the ribosome P site, and $^{L}$aa is incorporated into the peptide nascent chain. $^{D}$aa charged onto amber tRNA$_{CUA}$ presumably binds to EF-Tu poorly and also is incompatible with the ribosome A site and in most cases fails to suppress the amber codon.

observed poor incorporation efficiency of $^{D}$aa-elongation in general.

As the second reason, we suggest that the recognition of initiator $^{D}$aa-tRNA$^{fMet}_{CAU}$ by P site of ribosome is less strict than that of elongator $^{D}$aa-tRNAs by the A site. Presumably, since the P site needs to accommodate peptidyl-tRNA, it has a more spacious pocket than the A site (Noller et al. 2005). In fact, our recent studies using various peptidyl-tRNAs for initiation imply that the P site has a surprising tolerance toward the nonproteinogenic peptidyl groups (Y. Goto, H. Murakami, and H. Suga, in prep.). Moreover, in the P site, the aminoacyl carbonyl carbon of initiator f-$^{L}$Met-tRNA$^{fMet}_{CAU}$ or peptidyl-tRNAs acts as an electrophile to the α-amino group of elongator $^{L}$aa-tRNAs, so that it is only necessary to set its corresponding carbonyl group at the appropriate position; this positioning would be unlikely to be influenced by the chirality of the α-carbon of electrophile amino acids. In the A site, on the other hand, in order to efficiently process the nucleophilic attack of the α-amino group of elongator $^{L}$aa-tRNA to the carbonyl group of the initiator or peptidyl-tRNAs, precise positioning of the α-amino group is critical. This positioning should be largely influenced by the chirality of the α-carbon of nucleophile amino acids, in contrast to that of electrophile in the P site. We thus propose that the spacious and functional differences in ribosome's P and A sites play a critical role in accepting and discriminating $^{D}$aa in the initiation and elongation events, respectively.

The last reason involves a technical issue; since the background initiation was nearly completely suppressed by depleting Met in the *w*PURE system, undesired competition of f-Met-tRNA$^{fMet}_{CAU}$ against $^{D}$aa-tRNA$^{fMet}_{CAU}$ does not occur, leading to the exclusive initiation of $^{D}$aa-tRNA$^{fMet}_{CAU}$. On the other hand, the background of the amber suppression cannot be completely repressed due to the inclusion of RFs in the present methodology. This may suggest that if the background level of competing elongations were controlled by genetic code reprogramming, $^{D}$aa can be incorporated into nascent peptide chain.

Although, as stated earlier, knowing the exact mechanism in the $^{D}$aa-initiation requires more detailed investigations, particularly at the molecular level, the present work has given us an intriguing question for a general mechanism of the fidelity controls at initiation. It has been established that some aaRSs, such as TyrRS, TrpRS, and AspRS, mischarge the corresponding $^{D}$aa onto the cognate elongator tRNAs (Calendar and Berg 1966; Soutourina et al. 2000). This mischarging event gives a negative impact on cell growth via two possible mechanisms. A small portion of such a mischarged $^{D}$aa may be incorporated into a nascent peptide chain even though $^{D}$aa is generally a poor substrate for elongation (Roesser et al. 1989; Bain et al. 1991; Ellman et al. 1992; Dedkova et al. 2003; Murakami et al. 2006). Alternatively, undesirable generation of the mischarged elongator tRNAs may decrease the concentration of available elongator tRNAs for the innate function. At least in bacteria and yeast, it has been found that $^{D}$Tyr-tRNA deacylase plays a role in discharging such mischarged $^{D}$aa-tRNAs, and thus the harmful accumulation of $^{D}$aa-tRNAs are avoided (Calendar and Berg 1967; Soutourina et al. 2004). On the other hand, we found in the present work that $^{D}$Met-tRNA$^{fMet}_{CAU}$ could be formylated by MTF and the resulting f-$^{D}$Met-tRNA$^{fMet}_{CAU}$ could rather efficiently initiate the translation. Therefore, even if MetRS mischarged only a small fraction of tRNA$^{fMet}_{CAU}$ with $^{D}$aa, the resulting f-$^{D}$Met-tRNA$^{fMet}_{CAU}$ would compete with f-$^{L}$Met-tRNA$^{fMet}_{CAU}$ at the initiation of protein synthesis. Hence, it is intriguing to study whether MetRS mischarges $^{D}$Met onto tRNA$^{fMet}_{CAU}$. If so, it would be important to test whether $^{D}$Tyr-tRNA deacylase is able to discharge $^{D}$Met-tRNA$^{fMet}_{CAU}$, since it has been shown that this enzyme is able to discharge not only $^{D}$Tyr-tRNA$^{Tyr}$ but also $^{D}$Trp-tRNA$^{Trp}$ and $^{D}$Asp-tRNA$^{Asp}$ (Soutourina et al. 2000). If MetRS is able to mischarge $^{D}$Met and if $^{D}$Tyr-tRNA deacylase is unable to discharge

$^{D}$Met-tRNA$^{fMet}_{CAU}$, it will be interesting to investigate the mechanism of how the generation of f-$^{D}$Met-peptide is avoided or how f-$^{D}$Met-peptide is processed, e.g., whether peptide deformylase and methionine aminopeptidase are able to remove the f-$^{D}$Met group from the f-$^{D}$Met-peptide.

Nevertheless, the technical merit of our work is apparent. The reprogrammed initiation with $^{D}$aa and acyl-$^{D}$aa in translation enables us to synthesize a variety of H-$^{D}$aa and acyl-$^{D}$aa-peptides, respectively. The $^{D}$aa-capping would grant resistance against proteolytic degradation to the peptide as demonstrated in previous works using chemical synthesis (Hong et al. 1999; Tugyi et al. 2005). We have reported here a new method for the ribosomal synthesis of $^{D}$aa-capped or acyl-$^{D}$aa-capped peptides. The mRNA-programmed synthesis of such peptide libraries should provide a new avenue to discover novel physiologically stable peptidic drug candidates against various therapeutic targets, and such investigations are underway.

## MATERIALS AND METHODS

### Materials

Chemicals were purchased from Watanabe Chemical Industries, Nacalai Tesque, Kanto Chemical, Sigma-Aldrich Japan, or Wako Pure Chemical Industries unless noted otherwise and used without further purification. All oligonucleotides were purchased from Operon Biotechnologies. Flexizyme and tRNA molecules were synthesized using the same procedure as previously described (Murakami et al. 2006; Ohta et al. 2007; Goto et al. 2008).

### General protocol of translation

$^{D}$aa-tRNA$^{fMet}_{CAUU}$ was prepared by the following procedure. We heated 40 μM tRNA$^{fMet}_{CAU}$ in 0.2 M HEPES-K buffer (pH 7.5), 0.2 M KCl (7.5 μL) 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 Murakami et al. 2006) (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 3 μL of 25 mM $^{D}$aa substrate in DMSO and incubated on ice for the optimized times, generally 2–6 h (Murakami et al. 2006). 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 at 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, Lys, 50 μM [$^{14}$C]-Asp, and 120 μM of various $^{D}$aa-tRNA$^{fMet}_{CAU}$ molecules. The wild-type expression was carried out with *w*PURE system with 0.04 μM mDNA1 containing 200 μM each Met, Thr, Tyr, 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).

### Analysis of peptides by MALDI-TOF

For MALDI-TOF analysis, the 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 washing the resin with 30 μL of W buffer (50 mM Tris-HCl at pH 8.0, 150 mM NaCl), the immobilized peptides were eluted with 10 μL of 0.2% TFA. The purified peptide was desalted with ZipTip$_{\mu\text{-C18}}$ (Millipore), and eluted with 1 μL of a 50% MeCN, 0.1% TFA solution saturated with the matrix (*R*)-cyano-4-hydroxycinnamic acid. MALDI-TOF mass spectrometry was performed using an autoflex II TOF/TOF (BRUKER DALTONICS) under the linear/positive mode and externally calibrated with peptide calibration standard II (BRUKER DALTONICS).

### Competition of $^{D}$aa-tRNA$^{fMet}_{CAU}$ and $^{L}$aa-tRNA$^{fMet}_{CAU}$

$^{D}$aa-tRNA$^{fMet}_{CAU}$ and $^{L}$aa-tRNA$^{fMet}_{CAU}$ (aa = Tyr, Trp, and Leu) were prepared by the flexizyme system with the same procedure as described in the general protocol of translation. After completing the flexizyme reaction, the reaction mixture for $^{D}$aa-tRNA$^{fMet}_{CAU}$ and that for the corresponding $^{L}$aa-tRNA$^{fMet}_{CAU}$ were mixed with three different ratio (*D*:*L* = 99:1, 95:5, and 80:20). Then the reaction was stopped by addition of 45 μL of 0.6 M sodium acetateat pH 5, and the RNA mixture 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. Translation reaction was carried out with the same procedure as described above, except for adding the mixture of $^{D}$aa-tRNA$^{fMet}_{CAU}$ and $^{L}$aa-tRNA$^{fMet}_{CAU}$ to be 120 μM as total concentration of tRNAs in the translation mixture. MALDI-TOF analysis of the products was carried out in the same manner as above.

## SUPPLEMENTAL DATA

Supplemental material can be found at http://www.rnajournal.org.

## ACKNOWLEDGMENTS

We thank Dr. Patrick C. Reid for 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., Grants-in-Aid for JSPS Fellows (18-10526) to Y.G., a research and development projects of the Industrial Science and Technology Program in the New Energy and Industrial Technology Development Organization (NEDO) to H.S., and the Industrial Technology Research Grant Program in NEDO (05A02513a) to H.M.

*Received January 28, 2008; accepted March 30, 2008.*

## REFERENCES

Bain, J.D., Diala, E.S., Glabe, C.G., Wacker, D.A., Lyttle, M.H., Dix, T.A., and Chamberlin, A.R. 1991. Site-specific incorporation of nonnatural residues during in vitro protein biosynthesis with semisynthetic aminoacyl-tRNAs. *Biochemistry* **30:** 5411–5421.

Boelens, R. and Gualerzi, C.O. 2002. Structure and function of bacterial initiation factors. *Curr. Protein Pept. Sci.* **3:** 107–119.

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.

Calendar, R. and Berg, P. 1966. The catalytic properties of tyrosyl ribonucleic acid synthetases from *Escherichia coli* and *Bacillus subtilis*. *Biochemistry* **5:** 1690–1695.

Calendar, R. and Berg, P. 1967. D-Tyrosyl RNA: Formation, hydrolysis and utilization for protein synthesis. *J. Mol. Biol.* **26:** 39–54.

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.

Gold, L. 1988. Posttranscriptional regulatory mechanisms in *Escherichia coli*. *Annu. Rev. Biochem.* **57:** 199–233.

Goto, Y., Ohta, A., Sako, Y., Yamagishi, Y., Murakami, H., and Suga, H. 2008. Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chem. Biol.* **3:** 120–129.

Gualerzi, C.O. and Pon, C.L. 1990. Initiation of mRNA translation in prokaryotes. *Biochemistry* **29:** 5881–5889.

Hong, S.Y., Oh, J.E., and Lee, K.H. 1999. Effect of D-amino acid substitution on the stability, the secondary structure, and the activity of membrane-active peptide. *Biochem. Pharmacol.* **58:** 1775–1780.

Ibba, M. and Söll, D. 1999. Quality control mechanisms during translation. *Science* **286:** 1893–1897.

Kozak, M. 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* **47:** 1–45.

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.

Mayer, C., Stortchevoi, A., Kohrer, C., Varshney, U., and RajBhandary, U.L. 2001. Initiator tRNA and its role in initiation of protein synthesis. *Cold Spring Harb. Symp. Quant. Biol.* **66:** 195–206.

Mayer, C., Kohrer, C., Kenny, E., Prusko, C., and RajBhandary, U.L. 2003. Anticodon sequence mutants of *Escherichia coli* initiator tRNA: Effects of overproduction of aminoacyl-tRNA synthetases, methionyl-tRNA formyltransferase, and initiation factor 2 on activity in initiation. *Biochemistry* **42:** 4787–4799.

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 nonnatural polypeptide synthesis. *Nat. Methods* **3:** 357–359.

Noller, H.F., Hoang, L., and Fredrick, K. 2005. The 30S ribosomal P site: A function of 16S rRNA. *FEBS Lett.* **579:** 855–858.

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., Higashimura, E., and Suga, H. 2007. Synthesis of polyester by means of genetic code reprogramming. *Chem. Biol.* **14:** 1315–1322.

Roesser, J.R., Xu, C., Payne, R.C., Surratt, C.K., and Hecht, S.M. 1989. Preparation of misacylated aminoacyl-tRNA$^{Phe}$s useful as probes of the ribosomal acceptor site. *Biochemistry* **28:** 5185–5195.

Sankaranarayanan, R. and Moras, D. 2001. The fidelity of the translation of the genetic code. *Acta Biochim. Pol.* **48:** 323–335.

Schagger, H. and von Jagow, G. 1987. Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.* **166:** 368–379.

Schmitt, E., Guillon, J.M., Meinnel, T., Mechulam, Y., Dardel, F., and Blanquet, S. 1996. Molecular recognition governing the initiation of translation in *Escherichia coli*. A review. *Biochimie* **78:** 543–554.

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.

Söll, D. 1990. The accuracy of aminoacylation—Ensuring the fidelity of the genetic code. *Experientia* **46:** 1089–1096.

Soutourina, J., Plateau, P., and Blanquet, S. 2000. Metabolism of D-aminoacyl-tRNAs in *Escherichia coli* and *Saccharomyces cerevisiae* cells. *J. Biol. Chem.* **275:** 32535–32542.

Soutourina, O., Soutourina, J., Blanquet, S., and Plateau, P. 2004. Formation of D-tyrosyl-tRNA$^{Tyr}$ accounts for the toxicity of D-tyrosine toward *Escherichia coli*. *J. Biol. Chem.* **279:** 42560–42565.

Stanzel, M., Schon, A., and Sprinzl, M. 1994. Discrimination against misacylated tRNA by chloroplast elongation factor Tu. *Eur. J. Biochem.* **219:** 435–439.

Starck, S.R., Qi, X., Olsen, B.N., and Roberts, R.W. 2003. The puromycin route to assess stereo- and regiochemical constraints on peptide bond formation in eukaryotic ribosomes. *J. Am. Chem. Soc.* **125:** 8090–8091.

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.

Tugyi, R., Uray, K., Ivan, D., Fellinger, E., Perkins, A., and Hudecz, F. 2005. Partial D-amino acid substitution: Improved enzymatic stability and preserved Ab recognition of a MUC2 epitope peptide. *Proc. Natl. Acad. Sci.* **102:** 413–418.

Chemistry & Biology

# Synthesis of Polyester by Means of Genetic Code Reprogramming

Atsushi Ohta,[1] Hiroshi Murakami,[2] Eri Higashimura,[1] and Hiroaki Suga[1,2,*]
[1]Department of Chemistry and Biotechnology, Graduate School of Engineering, The University of Tokyo, 113-8656 Tokyo, Japan
[2]Research Center for Advanced Science and Technology, The University of Tokyo, 153-8904 Tokyo, Japan
*Correspondence: hsuga@rcast.u-tokyo.ac.jp
DOI 10.1016/j.chembiol.2007.10.015

## SUMMARY

**Here we report the ribosomal polymerization of α-hydroxy acids by means of genetic code reprogramming. The flexizyme system, a ribozyme-based tRNA acylation tool, was used to reassign individual codons to seven types of α-hydroxy acids, and then polyesters were synthesized under controls of the reprogrammed genetic code using a reconstituted cell-free translation system. The sequence and length of the polyester segments were specified by the mRNA template, indicating that high-fidelity ribosome expression of polyesters was possible. This work opens a door for the mRNA-directed synthesis of backbone-altered biopolymers.**

## INTRODUCTION

The translation system polymerizes amino acids to polypeptides using the genetic information present in the form of trinucleotide codons in the mRNA sequences. In this manner, the ribosome machinery acts as a template-directed synthesizer of polymers consisting of amino acids. Each of the codons directs the incorporation of 1 of 20 proteinogenic α-amino acids into the polypeptide chain, and therefore the ribosome is generally used for the synthesis of polypeptides, not for other biopolymers. If the codons are reassigned to nonproteinogenic amino acids, and if the translation system is adaptable to such alterations, mRNA-directed synthesis of nonproteogenic polypeptides by means of genetic code reprogramming is possible [1–6]. However, we have not yet witnessed the ribosomal synthesis of biopolymers with a nonpeptide backbone by such a methodology.

An α-hydroxy acid is chemically analogous to an α-amino acid, where the hydroxy group acts as a nucleophile. Despite this atomic difference, the ribosome is able to accept an α-hydroxy acid as a substrate to form an ester bond when hydroxyacyl-tRNA (ha-tRNA) is supplied to the translation system [7–14]. Previously, the specific incorporation of a single α-hydroxy acid into the polypeptide chain using amber stop codon suppression has been reported [7, 8, 10, 13]. Unfortunately, this strategy allows for the assignment of only a single kind of α-hydroxy acid, and therefore it is not suitable for the synthesis of complex polyesters composed of several different α-hydroxy acids.

In another approach reported over 30 years ago, the ribosome polymerizes phenyllactic acid ($F^{lac}$) on polyuridylic acid (poly-U) via random initiation and termination upon the addition of $F^{lac}$-$tRNA^{Phe}$ [12]. However, due to the methodology for the preparation of $F^{lac}$-$tRNA^{Phe}$ where Phe-$tRNA^{Phe}$ was chemically deaminated by nitrous acid, it was difficult to deplete the Phe-$tRNA^{Phe}$ contaminant completely from the translation mixture; therefore, the resulting polyester was actually composed of a heterogeneous random mixture of $F^{lac}$ and Phe with a ratio of approximately 7:3. Due to the fact that the poly-U template did not specify either a start or stop site, the ribosome could randomly initiate or terminate the synthesis of polyesters, and thus the length of the synthesized polyester could not be regulated.

Collectively, these previous findings indicate that the ribosome is capable of catalyzing the polymerization of α-hydroxy acids, but it remains unknown whether the mRNA-directed polymerization of several different α-hydroxy acids is possible. Here we report the synthesis of polyester composed of various α-hydroxy acids using a reprogrammed genetic code where the codons are reassigned from proteinogenic α-amino acids to α-hydroxy acids.

## RESULTS

### Genetic Code Reprogramming for the Incorporation of α-Hydroxy Acids

To demonstrate programmed synthesis of polyesters, we took advantage of two recently developed technologies (Figure 1). The first one is a de novo tRNA acylation system consisting of artificially evolved ribozymes, called flexizymes [14–17]. This system enables us to charge various hydroxy acids onto tRNAs bearing any desired anticodons and body sequences. The second one is a reconstituted *Escherichia coli* cell-free translation system, called the PURE (protein synthesis using recombinant elements) system [18, 19]. This translation system lets us create vacant codon boxes by withdrawing the corresponding aminoacyl-tRNA synthetases (ARSs) and amino acids from the reconstituted translation mixture (referred to as

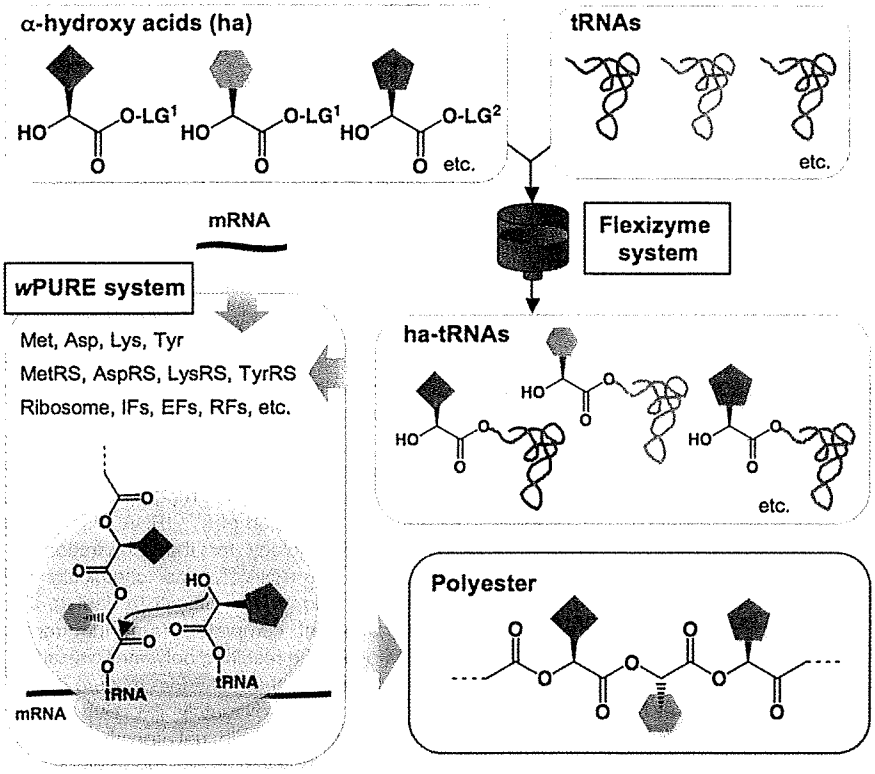


Figure 1. Ribosomal Polyester Synthesis by Means of the Flexizyme (eFlexiresin or dFlexiresin) and *w*PURE Systems
$LG^1$ and $LG^2$ indicate cyanomethyl and 3,5-dinitrobenzyl leaving groups, respectively (see Figure 2A). IFs, initiation factors; EFs, elongation factors; RFs, release factors.

the *w*PURE system). Vacant codons can then be reassigned to any desired α-hydroxy acids by using the flexizyme system to charge the α-hydroxy acids to the vacant codon's cognate tRNA. By combining these two technologies, the genetic code can be reprogrammed for the mRNA-directed synthesis of polyesters.

We chose seven α-hydroxy acids, of which four were phenyllactic acid derivatives ($F^{lac}$, $mF^{lac}$, and $cF^{lac}$, $pF^{lac}$) and three were nonaromatic α-hydroxy acids ($G^{lac}$, $L^{lac}$, and $A^{lac}$), to investigate mRNA template-directed ester polymerization (Figure 2A). We arbitrarily selected seven codons, and their corresponding anticodons were implanted into $tRNA^{Asn\text{-}E1}_{NNN}$ and $tRNA^{Asn\text{-}E2}_{NNN}$ sequences (NNN indicates anticodon sequence; Figures 3A and 3B), with each of the seven codons assigned to one of the seven α-hydroxy acids (Figure 2B). These tRNA sequences were chosen because we have previously shown that they could act as orthogonal tRNAs in the PURE system [14]. We next designed the mRNA template to initiate the polyester synthesis with fMet. For purification and detection purposes, we also incorporated a modified FLAG peptide (KKDYKDDDDK) at the C terminus (vide infra). Accordingly, the *w*PURE system for the polyester synthesis included only four amino acids (Met [M], Lys [K], Asp [D], and Tyr [Y]) and their cognate ARSs. It should be noted that although the translated polyesters would be conjugated with the above polypeptide sequence (i.e., polyester-polypeptide hybrids would be translated), the polypeptide segment was embedded in the polymers strictly for detection and purification purposes. Because the genetic code reprogramming was applied to the polyester segments only, we referred to such polymers as polyesters for the sake of simplicity of our experimental significance.

We first tested whether each tRNA bearing the designated anticodon retained its orthogonality and thereby could incorporate an α-hydroxy acid at a single template-directed position in a model polypeptide, without significant competition from other amino acids present in the translation mixture. Our conventional assay using a microhelix RNA indicated that the flexizyme system was able to efficiently charge all of the α-hydroxy acids onto either $tRNA^{Asn\text{-}E1}_{NNN}$ or $tRNA^{Asn\text{-}E2}_{NNN}$ (see Figure S1 in the Supplemental Data available with this article online). All of the α-hydroxy acid-charged tRNAs retained their orthogonality in the *w*PURE system, unfortunately with one exception: neither $tRNA^{Asn\text{-}E1}_{CUG}$ nor $tRNA^{Asn\text{-}E2}_{CUG}$ bearing a CUG anticodon assigning $A^{lac}$ possessed the expected orthogonality, resulting in the minor incorporation of Lys into the targeted CAG position in the

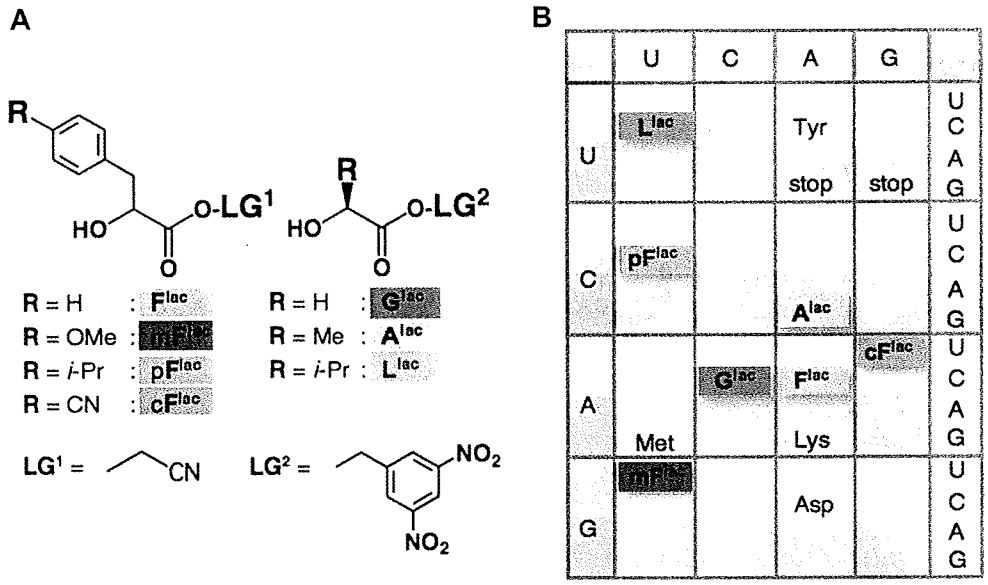


**Figure 2. Genetic Code Reprogramming for mRNA-Directed Polyester Synthesis**

(A) Chemical structure and abbreviation of α-hydroxy acids used in this study. $F^{lac}$, phenyllactic acid; $mF^{lac}$, *p*-methoxyphenyllactic acid; $pF^{lac}$, *p*-isopropylphenyllactic acid; $cF^{lac}$, *p*-cyanophenyllactic acid; $G^{lac}$, glycolic acid; $A^{lac}$, lactic acid; $L^{lac}$, isopropyllactic acid. The abbreviations used are based on structurally similar amino acids. The stereochemistry of $pF^{lac}$, $mF^{lac}$, and $cF^{lac}$ is racemic, whereas that of other α-hydroxy acids is *S* configuration. All phenyllactic acid derivatives bear a cyanomethyl ester group ($LG^1$) and are charged onto tRNAs shown in Figure 3 using eFlexiresin, while others bear a 3,5-dinitrobenzyl ester group ($LG^2$) and are charged onto the tRNAs using dFlexiresin.

(B) The reprogrammed genetic code used in this study. α-hydroxy acids assigned to the respective triplets are color-coded as shown.

polyester-reading frame. Presumably, these tRNAs were susceptible to Lys misacylation by LysRS included in the *w*PURE system, so that the resulting Lys-$tRNA^{Asn\text{-}E1/E2}_{CUG}$ competed with the $A^{lac}$-$tRNA^{Asn\text{-}E1/E2}_{CUG}$ for incorporation into the polypeptide chain (data not shown). To circumvent this problem, we developed an orthogonal tRNA derived from a *Mycobacteriophage L5* $tRNA^{Asn}_{CUG}$ ($tRNA^{MLAsn}$ in Figure 3C). We found that $tRNA^{MLAsn}_{CUG}$ was not aminoacylated with Lys by LysRS, and thus the misincorporation of Lys was not observed. To this end, we engineered each of the seven anticodons, each assigned to a different α-hydroxy acid, into three different tRNA body sequences to explore the mRNA-directed expression of polyesters (Figures 2B and 3A–3C).

### mRNA-Directed Synthesis of Polyesters Containing Four Consecutive Ester Linkages

We designed four nucleotide templates (hereafter referred to as T1–T4) to express polyesters (E1–E4) consisting of four consecutive ester bonds in the *w*PURE system (Figure 4A). As a control, these templates were also translated in the conventional PURE system, and thereby the expression level of polyesters could be compared with that of polypeptides (P1–P4). To detect the expressed products containing polyester, [$^{14}$C]Asp was included in the *w*PURE system. [$^{14}$C]Asp was incorporated into the Asp residues in the C-terminal FLAG peptide sequence, and the expressed polyester-[$^{14}$C]FLAG products were monitored by tricine-SDS-PAGE and autoradiography.

Tricine-SDS-PAGE analysis revealed that expression of E1 was observed only when all ha-tRNAs were present in the *w*PURE system and its expression level was comparable to that of the control peptide, P1 (Figure 4B, lanes 1–5). The molecular mass (ms) of E1 determined by MALDI-TOF analysis was consistent with the calculated ms of the full-length E1 (Figure 5A, E1). Likewise, the respective T2–T4 templates expressed the E2–E4 polyesters with the expected ms (Figure 4B, lanes 6–11; Figure 5A, E2–E4). The expression levels of P1, E1, E2, E3, and E4 were estimated by the incorporation of [$^{14}$C]Asp by tricine-SDS-PAGE autoradiography analysis considering that five aspartic acids were included in the polymers, giving approximately 6 pmol/5 μl, 9.5 pmol/5 μl, 15 pmol/5 μl, 12 pmol/5 μl, and 5.3 pmol/5 μl, respectively (data not shown). These results provide solid evidence that the length and sequence of the tetrapolyesters are strictly controlled by the mRNA template sequence.

### mRNA-Directed Synthesis of Longer Polyesters

We next explored the capability of ribosomes for the synthesis of various lengths of polyesters. In addition to T1 expressing tetrapolyester E1, five templates T5–T9 were designed to express tri-, penta-, hexa-, octa-, and dodecapolyesters E5–E9 (Figure 4A). We were able to detect tripolyester E5 with the expected ms similar to that of tetrapolyester E1 (Figure 5B, E5 and E1), whereas the full-length peak of pentapolyester E6 was barely detectable exhibiting a significantly poor signal/noise ratio, while the

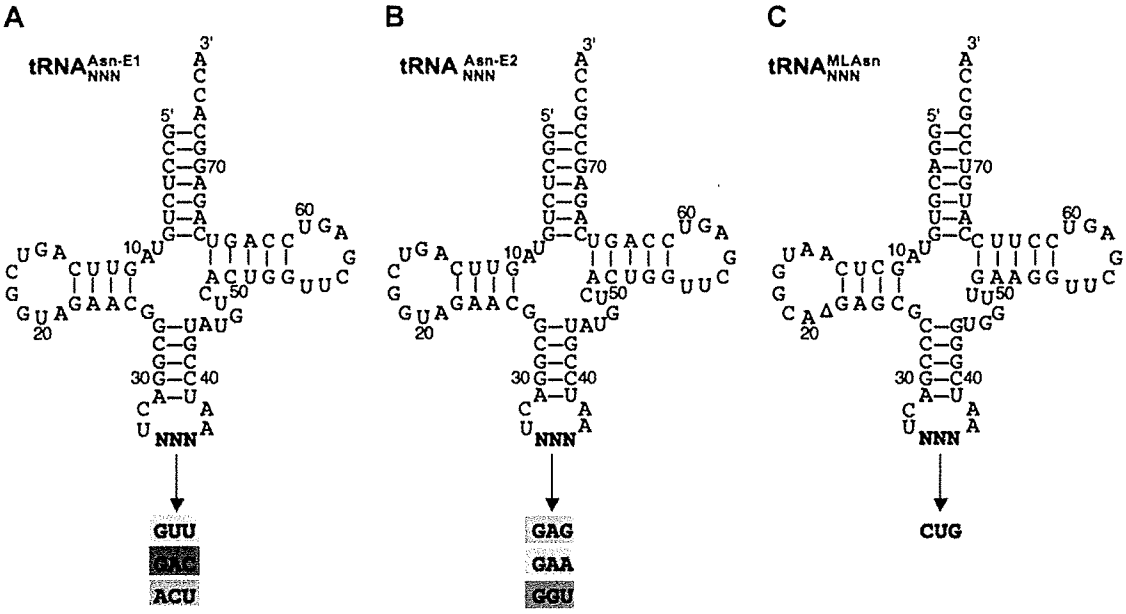


Figure 3. Three tRNA Body Sequences Used in This Study
(A) $tRNA^{Asn\text{-}E1}_{NNN}$.
(B) $tRNA^{Asn\text{-}E2}_{NNN}$.
(C) $tRNA^{MLAsn}_{NNN}$. NNN indicates anticodon and each anticodon is color-coded to pair with the corresponding codon shown in Figure 2B.

full-length peaks of E7–E9 were not detected (Figure S2). We wondered whether the lack of a full-length peak of E7–E9 was due to a significant reduction in polyester expression. Tricine-SDS-PAGE analysis showed that the expression level was slightly diminished as the polyester became longer, yet a detectable level of expression was observed for E6–E9 (Figure 4B, lanes 12–19). Due to the fact that [$^{14}$C]Asp was incorporated into the C-terminal FLAG sequence, we concluded that the polyester-FLAG could be expressed in all cases. We thus speculated that the full-length E6–E9 might be difficult to ionize under ordinary MALDI-TOF conditions. Unfortunately, attempts under more stringent ionization conditions, such as longer/stronger irradiation or the use of other matrices, resulted in fragmentation of the polyesters (data not shown).

Due to the difficulty in ionization of polyesters longer than a pentapolyester, we modified the strategy to obtain evidence for longer polyesters by MALDI-TOF analysis. The full-length polyester was hydrolyzed under basic conditions to generate shortened fragments. Because the FLAG peptide region could not be hydrolyzed under such conditions, we expected that $(ha)_n$-FLAG (n indicates the number of α-hydroxy acyl groups [ha] remaining after the base hydrolysis) could be detected by MALDI-TOF. Indeed, upon analysis of the hydrolyzed products, we were able to observe the $(ha)_2$-FLAG product as a major peak for E6–E9, occasionally with a minor peak of ha-FLAG (Figure 5B, E6–E9). Most importantly, the observed ms for each of the $(ha)_2$-FLAG products hydrolytically generated from the full-length E6–E9 was consistent with the expected ms of $(ha)_2$-FLAG encoded by the respective mRNA sequence (Figure 4A). Taken together, our findings show that polyesters can be continuously synthesized up to dodecapolyester in accordance with the mRNA template.

## DISCUSSION

The flexizyme system is a highly flexible acylation tool that enables us to charge various acids onto any desired tRNAs. By means of this system, we were able to readily prepare tRNAs charged with a variety of α-hydroxy acids (ha-tRNAs). The various ha-tRNAs were then utilized in a special reconstituted cell-free translation system in which unneeded amino acids and their corresponding ARSs are withdrawn from the translation mixture (referred to as the *w*PURE system). Removal of such components creates vacant codons which can then be reassigned to specific α-hydroxy acids using the flexizyme system to generate ha-tRNAs that designate the respective codons. Thus, by combining these two technologies, we could use ribosomes for the synthesis of mRNA template-designed polyesters. Our data clearly show that polyesters of up to 12 α-hydroxy acids in length could be synthesized with the length and sequence defined by the mRNA template.

Currently, it is not clear what hindered the synthesis of polyesters longer than 12 α-hydroxy acids, yet based on this study and our other unpublished work on the incorporation of exotic amino acids into the nascent peptide

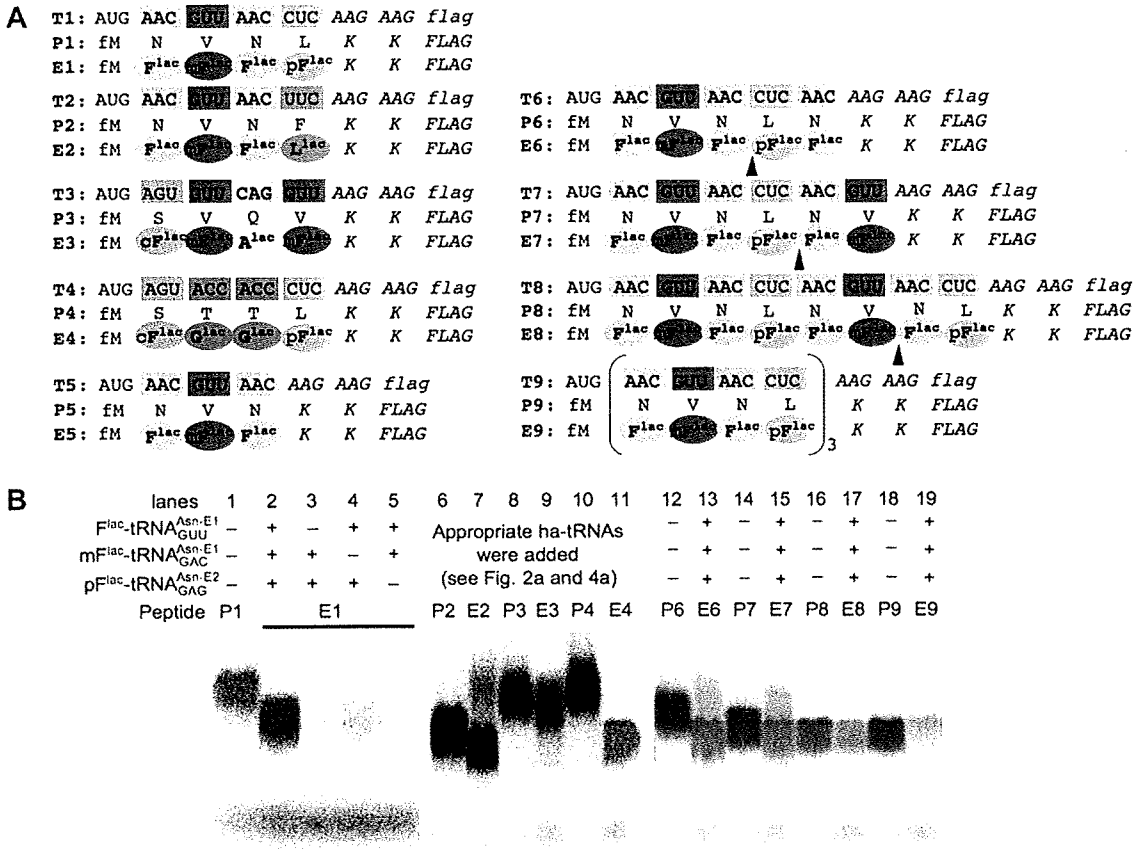


**Figure 4. Polyester Synthesis**

(A) Sequences of mRNA templates (T1–T9), polypeptides (P1–P9), and polyesters (E1–E9). Reprogrammed codons in mRNA (rectangles) and α-hydroxy acids (circles) are highlighted in colors matching those in Figure 2B. Black arrowheads indicate the positions of hydrolyzed ester bonds that give $(ha)_2$-FLAGs observed in Figure 5B.

(B) Tricine-SDS-PAGE analysis of the expressed products. The products, labeled with [$^{14}$C]Asp in the C-terminal FLAG peptide, were detected by autoradiography. Due to the basic conditions (pH 8.5) of the tricine-SDS-PAGE running buffer, each observed band derived from E1–E9 was likely the corresponding $(ha)_n$-FLAG peptide(s) generated by the hydrolysis of full-length polyester. Note that the mobility of the peptide or polyester-peptide hybrid was dependent upon its composition, that is, the net charge or hydrophobicity, and therefore the peptide mobility did not accurately reflect the peptide length. The combination of ha-tRNAs used in each lane for polyester synthesis was as follows: lane 7, $F^{lac}$-$tRNA^{Asn\text{-}E1}_{GUU}$, $mF^{lac}$-$tRNA^{Asn\text{-}E1}_{GAC}$, and $L^{lac}$-$tRNA^{Asn\text{-}E2}_{GAA}$; lane 9, $cF^{lac}$-$tRNA^{Asn\text{-}E1}_{ACU}$, $mF^{lac}$-$tRNA^{Asn\text{-}E1}_{GAC}$, and $A^{lac}$-$tRNA^{MLAsn}_{CUG}$; lane 11, $cF^{lac}$-$tRNA^{Asn\text{-}E1}_{ACU}$, $G^{lac}$-$tRNA^{Asn\text{-}E2}_{GGU}$, and $pF^{lac}$-$tRNA^{Asn\text{-}E2}_{GAG}$.

chain, we propose the following possibility. According to various experiments examining the kinetics of a single elongation of an α-hydroxy acid, the ester bond transfer rate is estimated to be at least one order of magnitude slower than its peptide counterpart [20, 21]. Moreover, the dissociation constant of $F^{lac}$-$tRNA^{Phe}$ ($K_d$ = 30 μM) with EF-Tu-GTP is about two orders of magnitude higher than that of Phe-$tRNA^{Phe}$ ($K_d$ = 100 nM) [22]. Even though we attempted to compensate for the poor affinity of ha-tRNA for EF-Tu by increasing the concentration of ha-tRNA, these efforts were insufficient. Therefore, this combination of a slow esteryl transfer rate for α-hydroxy acids with the poor affinity of ha-tRNA for EF-Tu likely causes the ribosome to stall during the translation of long polyesters, possibly resulting in the frequent dissociation of the translation complex and polyesteryl-tRNA molecule from the mRNA template. Our observation that polyester yield correlated with polyester length appears to support this hypothesis. Detailed studies in the future should provide not only more insights into the mechanism but also insights into how to engineer the translation system toward a more efficient synthesis of longer polyesters.

One obvious strategy toward generating a more efficient translation system for polyester synthesis would be to engineer EF-Tu or orthogonal tRNAs to have a tighter binding affinity [23, 24]. Another possible strategy is to generate a mutant ribosome capable of catalyzing the esteryl transfer reaction more efficiently [25–27]. Thus, if we were able to optimize the affinity of ha-tRNA for EF-Tu and ribosomes by selection or increase the

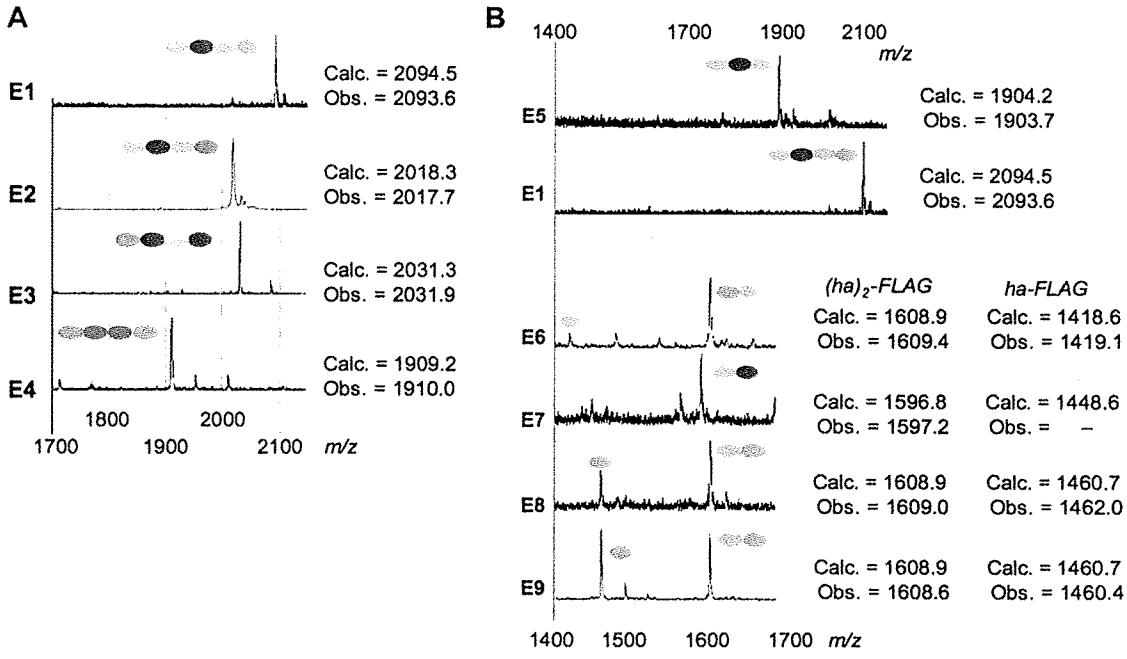


Figure 5. MALDI-TOF Analysis of Polyesters

(A) MALDI-TOF spectra of E1–E4. Calculated mass (M+H) and observed mass (M+H) are presented with the polyester moiety color-coded based on the reprogrammed genetic table shown in Figure 2A.

(B) MALDI-TOF spectra of E1 and E5–E9. Calculated mass (M+H) and observed mass (M+H) of full-length (E1 and E5) or $(ha)_2$-FLAG and ha-FLAG (E6–E9) are shown for each spectrum. The fragmented N-terminal products were removed during the FLAG-purification step, and therefore only the C-terminal fragments appear in the spectra.

efficiency of ribosomes for the esteryl transfer reaction, the translation system could be turned into a more efficient template-directed synthesizer for polyesters or even other nonpeptide biopolymers.

In light of the fact that polyesters have significantly different properties compared to peptides in terms of plasticity and rigidity, it is intriguing to investigate polyesters and polyester-polypeptide hybrid biopolymers engineered with a wide variety of side chains in strictly controlled sequences and lengths. However, because it has been difficult to synthesize such biopolymers with a variety of side chains, such biopolymers have not been well studied. The methodologies described in this report should open the door for the synthesis of a wide variety of polyesters and hybrid polymers, and thus will enable us to screen them for new possible functions and applications in the future.

## SIGNIFICANCE

**We have reported the mRNA-directed synthesis of polyesters, composed of several different α-hydroxy acids, whose sequence and length are fully controlled through reprogramming of the genetic code. This approach should be extendable to programmed synthesis of other types of backbone-modified biopolymers, such as polypeptide-ester hybrids and N-methyl-polypeptides.**

## EXPERIMENTAL PROCEDURES

### α-Hydroxy Acid Substrate Synthesis

L-phenyllactic acid ($F^{lac}$) and lactic acid ($A^{lac}$) were purchased from Sigma-Aldrich and glycolic acid ($G^{lac}$) was purchase from Wako, Japan. *p*-isopropylphenyllactic acid ($pF^{lac}$), *p*-methoxyphenyllactic acid ($mF^{lac}$), and *p*-cyanophenyllactic ($cF^{lac}$) acid were synthesized from *p*-isopropylbenzaldehyde, *p*-methoxybenzaldehyde, and *p*-cyanobenzaldehyde, respectively, as previously reported [28]. *p*-isopropyllactic acid ($L^{lac}$) was synthesized as previously described [29]. All of the phenyllactic acid derivatives were converted to cyanomethyl ester ($LG^1$), and $G^{lac}$, $A^{lac}$, and $L^{lac}$ were converted to dinitrobenzyl ester ($LG^2$) by the procedure reported elsewhere [14, 17].

### Hydroxyacyl-tRNA Synthesis

For the hydroxyacylation of phenyllactic acid-$LG^1$ derivatives ($F^{lac}$, $mF^{lac}$, $pF^{lac}$, and $cF^{lac}$) eFlexiresin was used, whereas for other α-hydroxy acid-$LG^2$ derivatives ($G^{lac}$, $L^{lac}$, and $A^{lac}$) dFlexiresin was used. Due to the fact that tRNA hydroxyacylation was performed in the same manner regardless of either eFlexiresin or dFlexiresin, we refer to these resins commonly as Flexiresin in the protocol described below. Note that a general protocol for the preparation of Flexiresin is described in Supplemental Experimental Procedures. Hydroxyacylation was carried out as follows: a solution of 20 μl of 50 μM tRNA and 3 μl of 1 M Tris-HCl (pH 8.0) was heated at 95°C for 2 min and cooled to room temperature over 5 min. Six microliters of 3 M $MgCl_2$ was added to this mixture and the resulting solution was then loaded onto 10 μl (resin volume) of Flexiresin. After a suspension was made, the mixture was placed on ice and incubated for 5 min. The charging reaction was initiated by the addition of 1 μl of 300 mM α-hydroxy acid activated with an appropriate leaving group ($LG^1$ or $LG^2$) in

DMSO, and the mixture was incubated for 3 hr on ice. When *p*-isopropyllactic acid-LG[2] ($L^{lac}$) was used as a substrate, the reaction was incubated for 16 hr. Following the incubation, the supernatant was removed and Flexiresin-bound ha-tRNA was eluted with 30 μl of elution buffer (50 mM HEPES-K [pH 7.5], 10 mM EDTA) four times. To the pooled eluate 13.3 μl of 3 M NaOAc (pH 5.2) was added, and the resulting solution was precipitated with 280 μl of ethanol. After centrifugation, the resulting pellet was washed with 70% ethanol and dried under vacuum. Typically, the resulting ha-tRNA was dissolved in 0.5 μl of water to perform the expression of polyesters.

### mRNA-Directed Polyester Synthesis

The *w*PURE system contains all the necessary components for translation except for AlaRS, ArgRS, AsnRS, CysRS, GlnRS, GluRS, GlyRS, HisRS, IleRS, LeuRS, PheRS, ProRS, SerRS, ThrRS, TrpRS, ValRS, and all 20 standard amino acids. To 12.6 μl of the above *w*PURE system was added 3.0 μl of 0.4 μM template DNA, 0.5 μl of 180 mM EDTA (pH 8.0), 3.0 μl of the mixture of Met, Tyr, and Lys (2 mM each), and 2.4 μl of 2 mM Asp (for MALDI-TOF analysis) or 500 μM [$^{14}$C]Asp (for tricine-SDS-PAGE assay). The mixture of ha-tRNAs (approximately 1000 pmol each) dissolved in 1.5 μl of 1 mM NaOAc (pH 5.2) was added to 3.5 μl of the above stock solution (total 5 μl translation volume) and the resulting mixture was incubated for 3 hr at 37°C. The synthesis of polyesters was then analyzed by the following procedures.

### SDS-PAGE Analysis

The gel dimensions were as follows. The lengths of separation and stacking gels were approximately 55 and 10 mm, respectively, and the width and thickness were 83 and 0.75 mm, respectively. Separation gels contained 15% acrylamide (acrylamide:bisacrylamide = 19:1), 1 M Tris-HCl (pH 8.5), 0.1% SDS, and 13% glycerol, whereas the stacking gels contained 4% acrylamide (acrylamide:bisacrylamide = 29:1), 0.75 M Tris-HCl (pH 8.5), and 0.075% SDS. Electrophoresis was performed for 80 min under 20 mA constant mode. The anode running buffer was 200 mM Tris-HCl (pH 8.9), whereas the cathode running buffer contained 100 mM Tris, 100 mM tricine, and 0.1% SDS (the resulting buffer was approximately pH 8.3). The products labeled with [$^{14}$C]Asp in the C-terminal FLAG peptide were quantified by autoradiography using an image analyzer (FLA-5100, Fuji, Japan).

### MALDI-TOF Analysis

The expressed product was incubated with FLAG-M2 agarose (Sigma) and the resin was washed with 50 μl of washing buffer (50 mM MOPS-K [pH 7.0], 150 mM NaCl). The immobilized product was eluted with 10 μl of 0.2% TFA. The purified product was then desalted with $ZipTip_{\mu\text{-}C18}$ (Millipore), and eluted with 1 μl of a 50% MeCN, 0.1% TFA solution saturated with the matrix *R*-cyano-4-hydroxycinnamic acid. MALDI-MS measurement of each product was performed using Autoflex II TOF/TOF (Bruker Daltonics) under the linear/positive mode and externally calibrated with Substance P (1347.5354 Da), Bombesin (1619.8223 Da), ACTH clip (1–17) (2093.0862 Da), and ACTH clip (18–39) (2465.1983 Da) as standards.

### Supplemental Data

Supplemental Data include two figures and Supplemental Experimental Procedures and can be found with this article online at http://www.chembiol.com/cgi/content/full/14/12/1315/DC1/.

### ACKNOWLEDGMENTS

We thank Professor M. Komiyama for the MALDI instrumentation and Dr. Patrick C. Reid for proofreading. This work was supported by grants from the Japan Society for the Promotion of Science Grants-in-Aid for Scientific Research (S) (16101007) and the U.S. National Institutes of Health (GM59159).

Received: May 30, 2007
Revised: September 11, 2007
Accepted: October 22, 2007
Published: December 26, 2007

### REFERENCES

1. Josephson, K., Hartman, M.C., and Szostak, J.W. (2005). Ribosomal synthesis of unnatural peptides. J. Am. Chem. Soc. *127*, 11727–11735.
2. 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.
3. Frankel, A., Millward, S.W., and Roberts, R.W. (2003). Encodamers: unnatural peptide oligomers encoded in RNA. Chem. Biol. *10*, 1043–1050.
4. Merryman, C., and Green, R. (2004). Transformation of aminoacyl tRNAs for the in vitro selection of "drug-like" molecules. Chem. Biol. *11*, 575–582.
5. Kwon, I., Kirshenbaum, K., and Tirrell, D.A. (2003). Breaking the degeneracy of the genetic code. J. Am. Chem. Soc. *125*, 7512–7513.
6. Seebeck, F.P., and Szostak, J.W. (2006). Ribosomal synthesis of dehydroalanine-containing peptides. J. Am. Chem. Soc. *128*, 7150–7151.
7. Ellman, J.A., Mendel, D., and Schultz, P.G. (1992). Site-specific incorporation of novel backbone structures into proteins. Science *255*, 197–200.
8. Koh, J.T., Cornish, V.W., and Schultz, P.G. (1997). An experimental approach to evaluating the role of backbone interactions in proteins using unnatural amino acid mutagenesis. Biochemistry *36*, 11314–11322.
9. 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.
10. England, P.M., Lester, H.A., and Dougherty, D.A. (1999). Mapping disulfide connectivity using backbone ester hydrolysis. Biochemistry *38*, 14409–14415.
11. Fahnestock, S., and Rich, A. (1971). Synthesis by ribosomes of viral coat protein containing ester linkages. Nat. New Biol. *229*, 8–10.
12. Fahnestock, S., and Rich, A. (1971). Ribosome-catalyzed polyester formation. Science *173*, 340–343.
13. Millward, S.W., Takahashi, T.T., and Roberts, R.W. (2005). A general route for post-translational cyclization of mRNA display libraries. J. Am. Chem. Soc. *127*, 14142–14143.
14. 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.
15. 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.
16. Murakami, H., Saito, H., and Suga, H. (2003). A versatile tRNA aminoacylation catalyst based on RNA. Chem. Biol. *10*, 655–662.
17. Saito, H., Kourouklis, D., and Suga, H. (2001). An in vitro evolved precursor tRNA with aminoacylation activity. EMBO J. *20*, 1797–1806.
18. 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.
19. Forster, A.C., Weissbach, H., and Blacklow, S.C. (2001). A simplified reconstitution of mRNA-directed peptide synthesis: activity of

the ε enhancer and an unnatural amino acid. Anal. Biochem. *297*, 60–70.

20. Fahnestock, S., Neumann, H., Shashoua, V., and Rich, A. (1970). Ribosome-catalyzed ester formation. Biochemistry *9*, 2477–2483.
21. Bieling, P., Beringer, M., Adio, S., and Rodnina, M.V. (2006). Peptide bond formation does not involve acid-base catalysis by ribosomal residues. Nat. Struct. Mol. Biol. *13*, 423–428.
22. Derwenskus, K.H., and Sprinzl, M. (1983). Interaction of cinnamyl-tRNAPhe with *Escherichia coli* elongation factor Tu. FEBS Lett. *151*, 143–147.
23. 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.
24. Asahara, H., and Uhlenbeck, O.C. (2002). The tRNA specificity of *Thermus thermophilus* EF-Tu. Proc. Natl. Acad. Sci. USA *99*, 3499–3504.
25. Cochella, L., and Green, R. (2004). Isolation of antibiotic resistance mutations in the rRNA by using an in vitro selection system. Proc. Natl. Acad. Sci. USA *101*, 3786–3791.
26. 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.
27. 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.
28. Dahn, H., and Rotzler, G. (1991). Unstable 1,1,2-enetriols as (probable) intermediates in the decarboxylation of α,β-diketo acids. J. Org. Chem. *56*, 3080–3082.
29. Deechongkit, S., You, S.L., and Kelly, J.W. (2004). Synthesis of all nineteen appropriately protected chiral α-hydroxy acid equivalents of the α-amino acids for Boc solid-phase depsi-peptide synthesis. Org. Lett. *6*, 497–500.

 

Biochemical and Biophysical Research Communications 381 (2009) 533–536

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

journal homepage: www.elsevier.com/locate/ybbrc

# TRAIL inhibited the cyclic AMP responsible element mediated gene expression

Yasuhito Tokumoto [a,*], Katsuhisa Horimoto [b], Jun Miyake [a,c]

[a] Department of Bioengineering, Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
[b] Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST), 2-42 Aomi, Koto-ku, Tokyo135-0064, Japan
[c] Research Institute for Cell Engineering (RICE), National Institute of Advanced Industrial Science and Technology (AIST), 1-8-31 Midorigaoka, Ikeda, Osaka 563-8577, Japan

ARTICLE INFO

*Article history:*
Received 14 February 2009
Available online 20 February 2009

*Keywords:*
CRE
CREB
TRAIL
Signal transduction
Transcription

ABSTRACT

Tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) not only causes apoptotic cell death in tumor cells, but also activates some transcription factors and affects several other cellular functions. In this study, we observed the effect of administration of TRAIL on gene expression downstream of the cyclic AMP responsive element (CRE) enhancer by using the signal transduction reporter *cis*-element plasmid pCRE-d2EGFP. Western blotting showed that after administration of TRAIL, the expression level of reporter protein d2EGFP was down-regulated in NIH3T3 cells. To confirm the TRAIL-induced down-regulation of CRE enhancer controlled gene expression, DNA Chip time series analysis of the intrinsic genes expressed in NIH3T3 cells was carried out. As a result, the expression levels of six genes, which have CRE sequence in their promoter region, were slightly down-regulated within three hours after administration of TRAIL.



## Introduction

Tumor necrosis factor (TNF)-related apoptosis-inducing ligand (TRAIL) is a member of the structurally related TNF cytokine family [1,2]. TRAIL has apoptotic anticancer activity which is not accompanied by general toxicity in most normal cells and tissues. Unfortunately, the mechanism of this tumor specific toxicity is unknown. TRAIL binds to death receptors (DR4 or DR5), which are expressed on their target cell surface, as trimers. The binding of a TRAIL trimer is followed by a trimer formation of DRs and aggregation of Fas-associated death domain (FADD) and procaspase-8 to the cytoplasmic region of DRs. The complex of these molecules is called the death inducing signal complex (DISC). In DISC, procaspase-8 is digested to become the activated caspase-8 that is the initiator of the apoptotic caspase cascade. In addition to this proteolytic cascade, the TRAIL-DR complex also activates transcription factors via TNF receptor type 1 associated death domain protein (TRADD) to the TNF receptor-associated factor 2 (TRAF2) pathway [1]. The activation of nuclear factor kappa B (NF-κB) and c-Jun (cellular oncogene ju-nana) by TRAF2 is well established. Under TRAIL stimulation, activated NF-κB induces the transcription of the genes of the cellular form of Fas-associated death domain like interleukin-1beta-converting enzyme inhibitory protein (cFLIP) and inhibitor of apoptosis protein (IPA); both proteins inhibit caspase-8. Compared with NF-κB, the function of c-Jun in TRAIL stimulation is more controversial. Moderate and long-term activation of c-Jun is pro-apoptotic, but strong and short-term activation of c-Jun is anti-apoptotic [2].

Recently, two independent groups reported that cAMP-responsive element binding protein (CREB), another transcription factor, is also phosphorylated by TRAIL stimulation [3,4]. CREB was initially identified as a target of the cyclic AMP (cAMP) signaling pathway, but studies on the activation of the immediate early genes revealed that CREB is a substrate for other than cAMP-dependent protein kinase A [5]. All signaling pathways that activate CREB lead to the phosphorylation of a particular residue, serine 133 (Ser133). Phosphorylated CREB specifically binds to the genomic DNA at a conserved cAMP-responsive element (CRE), which consists of an eight-base pair palindrome (TGACGTCA) and is typically found within 100 nucleotides from the TATA box [6,7]. Phosphorylated CREB also recruits another nuclear protein, CREB-binding protein (CBP). After binding to phosphorylated CREB, CBP serves as a molecular bridge that allows upstream transcription factors to recruit and stabilize the RNA polymerase II transcription complex at the TATA box. Although TRAIL-dependent phosphorylation of CREB has been shown in two tumor cell lines, SK-N-MC (neuroblastoma) and Jurkat (T cell leukemia), both are TRAIL sensitive and die within a few hours after administration of TRAIL, so it was difficult to identify the target genes of the phosphorylated CREB. In this study, we search for target genes of phosphorylated CREB after TRAIL stimulation in the TRAIL-resistant non-tumor cell line NIH3T3 [8]. First, by using the CRE containing signal transduction reporter plasmid pCRE-d2EGFP, we estimated the transcriptional activity of CREB by western blotting. Second, by a DNA Chip time series analysis, we evaluated the effect of TRAIL administration on the gene

* Corresponding author. Fax: +81 3 5841 1792.
*E-mail address:* y.tokumoto@will.dpc.u-tokyo.ac.jp (Y. Tokumoto).

0006-291X/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2009.02.076

expression patterns of the intrinsic genes, which possess CRE sequences in their promoter regions.

## Materials and methods

*Materials.* All chemicals without annotation were purchased from Wako Pure Chemical Industries. Dulbecco's modified Eagle's medium (DMEM), Opti-MEM1, Lipofectamine 2000, fetal bovine serum (FBS), Trypsin–EDTA, Antibiotic–Antimycotic mixture (penicillin, streptomycin, amphotericin B) were purchased from Invitrogen. Soybean trypsin inhibitor and Complete Mini proteinase inhibitor mix were purchased from Roche. 25 $cm^2$ Tissue culture flasks (T25) and 6 well cell culture plates were purchased from Nunc. Twenty-four well glass bottom culture plate (EzView) was purchased from Asahi Techno Glass. Cell strainer (40 mm nylon mesh) was purchase from Becton–Dickinson. Human recombinant TRAIL was purchased from PeproTech EC. Forskolin was purchased from Sigma–Aldrich.

*Transfection of reporter plasmid to NIH3T3.* The destabilized fluorescence protein d2EGFP gene containing signal transduction *cis*-element repoter plasmid pCRE-d2EGFP DNA (Clontech, #6034-1) and the control plasmid pd2EGFP-N1 DNA (Clontech, #6009-1) were purchased from Clontech. Plasmid DNA purification was carried out by using EndoFree Plasmid Maxi kit (QIAGEN #12362). Mouse embryonic fibroblast cell line NIH3T3 (RCB 1862) was provided by RIKEN Bio-Resource Center (Ibaraki, Japan). Cells were cultured in DMEM-10% FBS at 37 °C in 5% $CO_2$. Transfection was carried out by using a Lipofectamine 2000 transfection reagent according to the supplier's instruction. In short, for $3.0 \times 10^6$ cells of NIH3T3 inoculated in T25 culture flask, 12 μg plasmid DNA and 30 μl of Lipofectamine 2000 transfection reagent mixture diluted in 1.5 ml of OPTI-MEM1 was added over night. The transfection efficiency was examined by using an Olympus IX 70 fluorescence inverted microscope.

*Western blotting.* NIH3T3 pCRE-d2EGFP transient transfectant cells and NIH3T3 pd2EGFP-N1 transient transfectant cells were inoculated in the 6 well cell culture plate ($2.0 \times 10^6$ cells/well in 3 ml of DMEM-10% FBS). Cells were cultured over night at 37 °C in 5% $CO_2$. After then they were treated with TRAIL (200 ng/ml) or forskolin (10 μM). Harvest of the cells was carried out every hour. Cells were washed twice by ice-cold PBS and lysed in 100 μl of RIPA buffer (50 mM Tris–HCl, pH 7.5, 150 mM NaCl, 1% NP40, 0.5% sodium deoxycholate, 10 mM EDTA) containing Complete Mini proteinase inhibitor mix. The protein concentration of each crude extract was estimated by using BCR protein assay kit (PIERCE) according to the supplier's instructions. Ten micrograms of protein containing each samples loaded on individual well of 10% polyacrylamide gel and SDS–PAGE was carried out. After the electrophoresis, proteins were electro-transferred to Immun-blot PVDF membrane (Bio-Rad). After blocking by skim milk, PVDF membrane blotted with rabbit polyclonal anti-GFP antibody (MBL) for 1 h. Membrane was washed three times by PBS-T, then the second antibody, HRP conjugated goat anti-rabbit IgG antibody (Cell Signaling) was added. Existing chemiluminescent detection was carried out by using an ECL plus western blotting detection system (GE Healthcare). The pattern of blotting on PVDF membrane was imaged by a Typhoon 9410 Variable Mode Imager (GE Healthcare). The intensity of each band was estimated by ImageQuant TL software (GE Healthcare). For the internal control assay, the PVDF membrane was stripped by Restore PLUS western Blot Stripping Buffer (Thermo Scientific). The blocking procedure was carried out again and followed by the blotting of mouse monoclonal anti-β-tubulin antibody and HRP conjugated goat anti-mouse Ig polyclonal antibody (GE Healthcare).

*DNA Chip time series analysis.* NIH3T3 pCRE-d2EGFP transient transfectant cells were inoculated in the 6 well cell culture plate ($2.0 \times 10^6$ cells/well in 3 ml of DMEM-10% FBS). Cells were cultured over night at 37 °C in 5% $CO_2$. After then they were treated with TRAIL (200 ng/ml) and harvested at 0, 1, 2, 3 h after the administration of TRAIL. They were washed twice with ice-cold PBS. The preparations of total RNA were carried out by using the RNeasy Plus Mini kit (QIAGEN). After the estimation of the quality and the quantity, 10 μg of total RNA of each sample was used for Affymetrix Mouse gene 1.0 ST Array assay (Affymetrix). Array data were analyzed by using the GeneSpring GX 7.3.1 software (Affymetrix).

## Results and discussion

For the assay of TRAIL-induced CREB activation, we used the signal transduction *cis*-element reporter plasmid pCRE-d2EGFP. This plasmid contains three copies of the CRE enhancer consensus sequence (TGACGTCA) fused to a TATA-like promoter region from the Herpes simplex virus thymidine kinase promoter. In this system the activation of CREB, which binds to CRE, induces transcription of d2EGFP mRNA. After transfection of pCRE-d2EGFP into NIH3T3, we administrated TRAIL (200 ng/ml) and observed cell death and expression of d2EGFP using fluorescence microscopy. During a 24-h period, NIH3T3/pCRE-d2EGFP cells did not show either the cell death or an apparent increase in fluorescence of the d2EGFP protein (data not shown). It is said that the slow rate of chromophore formation and the stability of EGFP preclude the use of EGFP as a reporter to monitor fast changes in promoter activity [9,10]. To solve this problem, we tried to use a destabilized form of EGFP, d2EGFP, as a reporter protein, but it did not work well.

Since the cell biological assay might not suitable for this experiment, we tried to evaluate the protein expression level of d2EGFP by western blotting, a biochemical assay. First 200 ng/ml of TRAIL was administered to NIH3T3/pCRE-d2EGFP transient transfectant cells. At that time, for the positive control of CREB–CRE transcriptional activation, these cells were also treated with 10 μM of forskolin [11]. The cells were harvested every hour and the time-course samples were assessed by western blotting. The result was contrary to expectations. As shown in Fig. 1, the expression level of the reporter protein d2EGFP in NIH3T3 was clearly down-regulated from 4 h after administration of TRAIL. Because the d2EGFP protein increased from 8 h after administration of forskolin (positive control), the CREB–CRE transcription activation system worked well in NIH3T3 cells. Quantitative analysis showed that after 10 h of TRAIL administration, the amount of d2EGFP de-

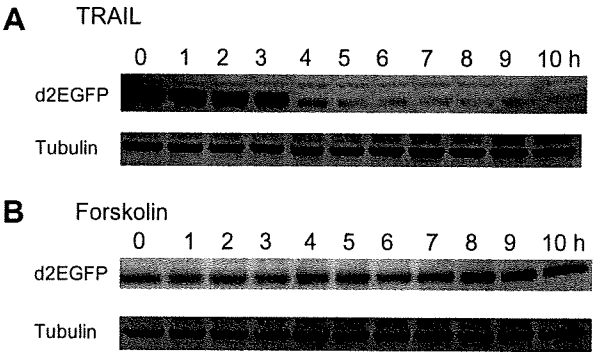


**Fig. 1.** A time-course analysis of reporter protein d2EGFP expression in NIH3T3 cells by western blotting. NIH3T3 cells harboring pCRE-d2EGFP plasmid DNA were treated by TRAIL (200 ng/ml) or forskolin (10 μM) and harvested every hour (0–10 h). (A) TRAIL-treated cells. (B) Forskolin-treated cells. The upper panel is anti-GFP antibody blotting (d2EGFP) and the lower panel is anti-β-tubulin antibody blotting (tubulin, as an internal control).

creased to less than 5% (Fig. 2). It was possible that the TRAIL-activated proteolysis cascade (caspase-8, caspase-3) might preferentially degrade d2EGFP protein. Therefore we transfected pd2EGFP-N1 plasmid DNA into NIH3T3 cells and used them in the TRAIL administration experiment. pd2EGFP-N1 is a cell-labeling plasmid containing a d2EGFP gene driven by the CMV promoter. NIH3T3 cells harboring pd2EGFP-N1 plasmid DNA constitutively expressed d2EGFP protein without any stimulation. As shown in Fig. 3, the administration of TRAIL to pd2EGFP-N1 transfected NIH3T3 cells did not reduce the amount of d2EGFP protein, so the hypothesis of accelerated degradation of d2EGFP protein was disproved.

To observe if this TRAIL-induced transcriptional inhibition downstream of the CRE enhancer occured only in the ex-chromosomal transient gene expression system or not, we examined the NIH3T3 cells' intrinsic gene expression pattern after administration of TRAIL by an exhaustive DNA Chip analysis. The reporter protein d2EGFP is a destabilized variant of the EGFP protein, and its half-life is around 2 h in a mammalian cell [9]. Because the down-regulation of d2EGFP protein in NIH3T3 cells occurred at 4 h after administration of TRAIL (Fig. 1), we estimated that the inhibition of the gene transcription of d2EGFP had to happen within 3 h after administration of TRAIL. Thus our study was focused on this time period. In previous studies, 23 genes were confirmed as being under the control of the CREB–CRE system [6,7]. Depending on the DNA Chip analysis data, we examined the transcription level of these genes from 1 to 3 h after administration of TRAIL (Table 1). Although six genes (26%) were down-regulated, but only one gene (4%) was up-regulated. The other 16 genes (70%) were unaffected. These data showed that TRAIL signaling works as a repressor of the CREB–CRE transcription system in NIH3T3 cells.

In previous studies, two independent groups reported the TRAIL induces phosphorylation of CREB at its Ser133. However the phosphorylation of CREB did not always elevate its transcriptional activity. $Ca^{2+}$/calmodulin-dependent protein kinase type II (CaMKII) is a member of the calmodulin serine/threonine kinase family. CaMKII phosphorylates CREB at its two of serine residues, one being the authentic Ser133, and the other Ser142. Phosphorylation of Ser142 inhibits the CREB transcriptional activity even when Ser133 is phosphorylated [12,13].

It is worth noting that the only gene that was up-regulated by TRAIL stimulation was c-fos (Gene ID 14281). c-Fos is a transcription factor that forms the heterodimer complex AP-1 with c-Jun [14]. As described previously, c-Jun is the major player in TRAIL-induced transcriptional control. Therefore the up-regulation of c-fos transcription may contribute to the resistance of NIH3T3 against TRAIL-induced apoptosis.

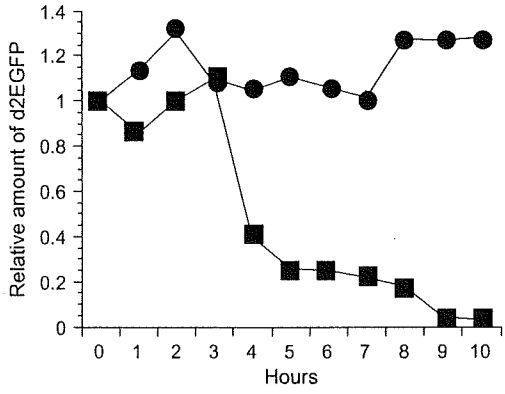


**Fig. 2.** Quantitative analysis of the band intensity of the western blots shown in Fig. 1. The intensity of each band was estimated by ImageQuant TL software and normalized by the band intensity of tubulin (internal control) in the same lane. The expression level of the d2EGFP protein is shown as the relative amount against time 0. Closed squares represent TRAIL-stimulated cells and closed circles represent forskolin-stimulated cells.

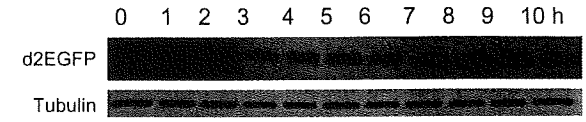


**Fig. 3.** TRAIL did not affect the stability of the d2EGFP protein. NIH3T3 cells harboring pd2EGFP-N1 plasmid DNA were treated by TRAIL (200 ng/ml) and harvested every hour (0–10 h). Western blotting was carried out. The upper panel is anti-GFP antibody blotting (d2EGFP) and the lower panel is anti-β-tubulin antibody blotting (tubulin, as an internal control).

**Table 1**
The effect of TRAIL administration to the intrinsic gene expressions.

| Name | Gene ID | Time 0 | 1 h | 2 h | 3 h |
|---|---|---|---|---|---|
| Cryaa | 12954 | 1 | 0.900 | 0.913 | 0.796 |
| Cyp19a1 | 13075 | 1 | 0.945 | 0.884 | 0.842 |
| Penk1 | 18619 | 1 | 0.771 | 0.968 | 0.724 |
| Pth | 19226 | 1 | 0.806 | 0.917 | 0.893 |
| Polb | 18970 | 1 | 0.915 | 0.848 | 0.819 |
| Tat | 234724 | 1 | 0.780 | 0.752 | 0.708 |
| Fos | 14281 | 1 | 1.133 | 1.157 | 1.216 |
| Adrb2 | 11555 | 1 | 1.064 | 1.054 | 1.035 |
| Chgb | 12653 | 1 | 1.070 | 0.966 | 0.968 |
| Cycs | 13063 | 1 | 0.844 | 1.039 | 0.968 |
| Fn1 | 14268 | 1 | 0.962 | 0.958 | 0.966 |
| Gcg | 14526 | 1 | 0.985 | 0.935 | 0.960 |
| Hsd11b1 | 15483 | 1 | 0.930 | 1.084 | 1.198 |
| H2-D1 | 14964 | 1 | 0.951 | 1.022 | 0.997 |
| Ldha | 16828 | 1 | 1.054 | 0.988 | 1.023 |
| Pck1 | 18534 | 1 | 1.130 | 1.133 | 0.946 |
| Plau | 18792 | 1 | 0.876 | 0.915 | 0.967 |
| Pou1f1 | 18736 | 1 | 0.925 | 0.944 | 0.914 |
| Scg2 | 20254 | 1 | 0.929 | 0.991 | 0.917 |
| Sst | 20604 | 1 | 0.928 | 0.879 | 1.023 |
| Syn1 | 20964 | 1 | 1.054 | 0.953 | 1.079 |
| Th | 21823 | 1 | 0.967 | 0.951 | 1.045 |
| Vip | 22353 | 1 | 1.018 | 0.991 | 0.944 |

The genes listed above possess at least one CRE site that is required for CREB-binding in their promoter region (6, 7). Relative amount of their transcripts after the TRAIL administration (0–3 h) estimated by DNA chip analysis was shown.

## Acknowledgments

This study was supported by a project grant from New Energy and Industrial Technology Development Organization (NEDO; Japan). We thanks to Qiaofan Li and Shinpei Tamaki for their technical supports.

## References

[1] F.C. Kimberley, G.R. Screaton, Following a TRAIL: update on a ligand and its five receptors, Cell Res. 14 (2004) 359–372.
[2] C. Falschlehner, C.H. Emmerich, B. Gerlach, H. Walczak, TRAIL signaling: decisions between life and death, Int. J. Biochem. Cell Biol. 39 (2007) 1462–1475.
[3] D. Milani, G. Zauli, E. Rimondi, C. Celeghini, S. Marmiroli, P. Narducci, S. Capitani, P. Secchiero, Tumor necrosis factor-related apoptosis-inducing ligand sequentially activates pro-survival and pro-apoptotic pathways in SK-N-MC neuronal cells, J. Neurochem. 86 (2003) 126–135.
[4] L. Caravatta, S. Sancilio, V. Di Giacomo, R. Rana, A. Cataldi, R. Di Pietro, PI3-K/AKT-dependent activation of cAMP-response element-binding (CREB) protein in Jurkat T Leukemia cells treated with TRAIL, J. Cell. Physiol. 214 (2008) 192–200.
[5] A.J. Shaywitz, M.E. Greenberg, CREB: a stimulus-induced transcription factor activated by a diverse array of extracellular signals, Annu. Rev. Biochem. 68 (1999) 821–861.
[6] M. Montminy, Transcriptional regulation by cyclic AMP, Annu. Rev. Biochem. 66 (1997) 807–822.

[7] K.A.W. Lee, N. Masson, Transcriptional regulation by CREB and its relatives, Biochim. Biophys. Acta 1174 (1993) 221–233.
[8] B.J. Kim, M.-S. Kim, K.-B. Kim, K.-W. Kim, Y.-M. Hong, I.-K. Kim, H.-W. Lee, Y.-K. Jung, Sensitizing effects of cadmium on TNF-α-and TRAIL-mediated apoptosis of NIH3T3 cells with distinct expression pattern of p53, Carcinogenesis 23 (2002) 1411–1417.
[9] R. Heim, D.C. Prasher, R.Y. Tsien, Wavelength mutations and posttranslational autoxidation of green fluorescent protein, Proc. Natl. Acad. Sci. USA 91 (1994) 12501–12504.
[10] I. Davis, C.H. Girdham, P.H. O'Farrell, A nuclear GFP that marks nuclei in living *Drosophila* embryos; maternal supply overcomes a delay in the appearance of zygotic fluorescence, Dev. Biol. 170 (1995) 726–729.
[11] K.B. Seamon, W. Padgett, J.W. Daly, Forskolin: unique diterpene activator of adenylate cyclase in membrane and in intact cells, Proc. Natl. Acad. Sci. USA 78 (1981) 3363–3367.
[12] P.K. Dash, K.A. Karl, M.A. Colicos, R. Prywes, E.R. Kandel, CAMP response element-binding protein is activated by $Ca^{2+}$/calmodulin- as well as cAMP-dependent protein kinase, Proc. Natl. Acad. Sci. USA 88 (1991) 5061–5065.
[13] P. Sun, H. Enslen, P.S. Myung, Differential activation of CREB by $Ca^{2+}$/calmodulin-dependent protein kinases type II and type IV involves phosphorylation of a site that negatively regulates activity, Genes Dev. 8 (1994) 2527–2539.
[14] L.A. Tibbles, J.R. Woodgett, The stress-activated protein kinase pathways, Cell. Mol. Life Sci. 55 (1999) 1230–1254.

Math.comput.sci. 1 (2008), 473–485

1661-8270/030473-13, *published online* March 5, 2008
DOI 10.1007/s11786-007-0037-z

**Mathematics in Computer Science**

# Inference of Probabilities over a Stochastic IL-System by Quantifier Elimination

Hiroshi Yoshida, Katsuhisa Horimoto, and Hirokazu Anai

**Abstract.** An algebraic approach based on quantifier elimination is proposed for the inference of probabilistic parameters over stochastic Lindenmayer systems with interaction, IL-systems. We are concerned with a multi-cellular organism as an instance of a stochastic IL system. The organism starts with one or a few cells, and develops different types of cells with distinct functions. We have constructed a simple model with cell-type order conservation and have assessed conditions for high cell-type diversity. This model is based on the stochastic IL-system for three types of cells. The cell-type order conservation corresponds to interaction terms in the IL-system. In our model, we have successfully inferred algebraic relations between the probabilities for cell-type diversity by using a symbolic method, quantifier elimination (QE). Surprisingly, three modes for the proliferation and transition rates emerged for various ratios of the initial cells to the developed cells. Furthermore, we have found that the high cell-type diversity pattern originates from the cell-type order conservation. Thus, QE has yielded analysis of the IL-system, which has revealed that, during the developing process of multi-cellular organisms, complex but explicit relations exist between cell-type diversity patterns and developmental rates.

**Mathematics Subject Classification (2000).** 03C10, 62M99, 65H10, 03D05.

**Keywords.** Lindenmayer system with interactions, cell-type diversity, quantifier elimination, algebraic approach.

## 1. Introduction

Inference problems for Lindenmayer (L) systems have been studied over decades. For instance, Feliciangeli, Herman and Walker gave a number of algorithms, which, for a finite set of finite sequences of strings of symbols, effectively produce grammars deriving these sequences [7, 9]. Much attention has been paid to grammars

useful in modelling developmental processes in biology and to algorithms to produce models that represent the developmental rules. However, the numbers of inferred grammars are so large that the essential parts of the developmental processes are obscure.

L-system models for developmental systems with cell lineages [11] and for biological tissues such as blood vessels of the eye [12, 17] have been inferred. In these L-systems, interactions between cells have been ignored, forming 0L-systems, because of the complexity or difficulty of analysis and inference. Such complexity or difficulty in inference problems appears even when the modelled systems are not biological. Indeed, many studies have limited their consideration on inference to 0L-systems [4, 18, 19].

Inference studies for L-systems have been vitalized by the development of effective methods for analysis and synthesis of these systems. In this paper, an algebraic approach based on quantifier elimination is presented for the derivation of probabilities over stochastic L-systems with interactions from the models. These derivation procedures have been performed exactly; that is, exact relations between probabilities and the desired pattern have been obtained. Such relations yield feasibility of inference for IL. The stochastic IL-system adopted in this paper is used for modelling a multi-cellular organism.

In a multi-cellular organism, a single cell – an egg – or a group of cells develops into a certain pattern with a variety of cell types [8]. These different cell types are created through cell differentiation and cell-to-cell interaction. The cell starts with an initial type, and then changes into several intermediate types before differentiating into the final type. In 1968, such a process of filamentous multicells was modelled by Lindenmayer [13, 14] and in 2005, relations between cell-to-cell interaction and multi-cellularity were surveyed [23]. Furthermore, stochastic aspects were introduced into L-systems [5, 6]. The stochastic L-system can take account of the influences of proliferation and transition rates, depending on the cell types.

Our aim in this paper is to infer algebraic relations between proliferation and transition rates for high cell-type diversity patterns with cell-type order conservation rules. For this purpose, we have constructed a model based on the stochastic IL-system, and have analyzed it by quantifier elimination (QE). The relations between parameters of the IL-system have allowed us to understand explicit relations between the cell-type order conservation rule and high cell-type diversity pattern. By using a symbolic approach, we are able to obtain the new knowledge of various models [1, 15, 20, 21]. In this paper, for instance, some modes are discovered as shown in Section 4.2. Such knowledge may shed some light on diversity of primitive multi-cells such as *Anabaena* [2] which can be described as an IL-system.

This paper is organized as follows. In Section 2, we give a brief overview of our previous model and results [22], wherein the cell-type order conservation rule appeared spontaneously. In Section 3, we introduce a model of a multi-cellular organism consisting of one-dimensional cells. This model postulates the cell-type order conservation rule as interactions of the L-system. The results of algebraic