Organic & Biomolecular Chemistry

Dynamic Article Links

Cite this: *Org. Biomol. Chem.*, 2011, **9**, 8495

www.rsc.org/obc

PAPER

Improvement of *in vitro*-transcribed amber suppressor tRNAs toward higher suppression efficiency in wheat germ extract†

Atsushi Ogawa,* Yasunori Doi and Nobuto Matsushita

Received 9th August 2011, Accepted 27th September 2011
DOI: 10.1039/c1ob06351k

In vitro-transcribed, unmodified, and non-aminoacylated amber suppressor tRNAs that are recognized by natural aminoacyl-tRNA synthetase were improved toward higher suppression efficiency in batch-mode cell-free translation in wheat germ extract. The suppression efficiency of the suppressor obtained through four sequence optimization steps (anticodon alteration of natural tRNAs (the first generation); chimerization of the efficient suppressors in the first generation; investigation and optimization of the effective parts in the second generation; combination of the optimized parts in the third generation) and by the terminal tuning was approximately 60%, which was 2.4-fold higher than that of the best suppressor in the first generation. In addition, an eRF1 aptamer further increased the efficiency up to 85%. This highly efficient suppression system also functioned well in a dialysis-based large-scale protein synthesis.

Introduction

Nonsense suppressor tRNAs (sup-tRNAs) are unique tRNAs that correspond to nonsense codons (UAA: ochre, UGA: opal, and UAG: amber) on mRNA. An aminoacylated sup-tRNA is incorporated into the ribosome at the corresponding nonsense codon on the mRNA in translation, competing with the release factor(s) ((e)RFs), to produce a 'read-through' protein. While some natural tRNAs are known to function as sup-tRNAs without anticodon-mutations, largely depending on the mRNA codon context,[1] many types of anticodon-adjusted sup-tRNAs have been artificially constructed.[2] Among them, pre-aminoacylated sup-tRNAs with a non-natural amino acid (naa)[3a] or orthogonal sup-tRNA/engineered aminoacyl-tRNA synthetase (aaRS) pairs[3b] have been widely used for syntheses of naa-incorporated proteins in various organisms' translation systems, both *in vitro* and *in vivo*. On the other hand, anticodon-adjusted sup-tRNAs that are aminoacylated with a normal amino acid by the cognate natural aaRS have been used *in vivo* for genetic analysis of chain terminating mutations and for elucidating various aspects of tRNA function.[2a–c] Recently, some *in vitro*-transcribed anticodon-adjusted sup-tRNAs have been reported that lack modified bases but are also aminoacylated with a normal amino acid even *in vitro*.[4] These sup-tRNAs, named here as 'IVT-sup-tRNAs', meaning *in vitro*-transcribed, unmodified, and non-aminoacylated sup-tRNAs, have the advantage that they are easily prepared and are readily introduced into the ribosome even *in vitro* without special treatment such as pre-aminoacylation, the addition of engineered aaRS, or posttranscriptional base modification. They enable the facile *in vitro* suppression of nonsense codons, so that they are easy to be applied to a read-through ribosome display method with natural mRNAs[4a] and to be used as an element of biosensors in combination with functional RNAs such as aptamers and aptazymes.[4b] However, to date, only bacterial IVT-sup-tRNAs have been found to show high suppression efficiencies *in vitro* (~100% or ~50%, in an RFs-free reconstituted[4a,b] or bacterial extract-based translation system,[4c] respectively). We report herein the amber IVT-sup-tRNA obtained by improvement of natural plant tRNA sequences through four generations and the terminal tuning functions efficiently in a cell-free plant translation system (wheat germ extract).

Results and discussion

Screening of natural tRNAs: searching for a basis of amber IVT-sup-tRNA (the first generation)

The simplest method for constructing amber IVT-sup-tRNAs is to just alter the anticodon of natural tRNAs into CUA for the amber codon, UAG. Although most of the sup-tRNAs constructed in this way are not supposed to be aminoacylated because natural aaRS, in general, strictly recognize the corresponding tRNA at various parts, including the anticodon loop, some aaRS are expected to be insensitive to anticodon alteration of their cognate tRNAs.[5a] SerRS and LeuRS in *E. coli* and humans recognize not the anticodon but dominantly the long variable arm,[5] though the anticodon of $tRNA^{Leu}$ in yeast is a major identity determinant.[6] In plant cells (tobacco protoplasts), an anticodon-adjusted amber sup-$tRNA^{Leu}_{CAA\rightarrow CUA}$ transcribed *in vivo*

Senior Research Fellow Center, Ehime University, 3 Bunkyo-cho, Matsuyama, Ehime, 790-8577, Japan. E-mail: a-ogawa@ccr.ehime-u.ac.jp; Fax: +81 89 927 8450; Tel: +81 89 927 8450

† Electronic supplementary information (ESI) available: Sequences of templates and primers for PCR, tRNA sequences in the first generation, doubly chimeric IVT-sup-tRNAs, effects of the 31–39 base pair in S1-2AC[a]. See DOI: 10.1039/c1ob06351k

is successfully aminoacylated and incorporated into the amber codon (suppression efficiency: 20–25%).[7] Moreover, SerRS and LeuRS in *E. coli* and yeast aminoacylate *in vitro* transcripts of tRNAs without base modifications as efficiently as native tRNAs extracted from cells.[5b,c,6,8] In fact, known bacterial IVT-sup-tRNAs is derived from natural $tRNA^{Ser}$ or $tRNA^{Leu}$.[4] It has also been reported that the anticodon of $tRNA^{Ala}$ in *E. coli* and humans is not a primary determinant.[5a,9] Therefore, we first narrowed down candidate tRNAs for the original basis of IVT-sup-tRNAs to these three kinds of tRNAs: $tRNA^{Ser}$; $tRNA^{Leu}$; $tRNA^{Ala}$.

tRNA gene sequences in *Oryza sativa* (rice), the same family (Poaceae) as wheat, were chosen as the tRNA framework.[10] The *Oryza sativa* genome has 61 $tRNA^{Ser}$, 58 $tRNA^{Leu}$, and 43 $tRNA^{Ala}$ genes. These genes can be divided into isodecoders, which have the same anticodon, as shown in Fig. 1A. We first picked up the representative(s) from each isodecoder and then altered their anticodons into CUA to prepare 18 kinds of amber IVT-sup-tRNA (the first generation).[11] Their relative suppression efficiencies were evaluated with 1-h cell-free translation in batch mode in wheat germ extract[12] by using a mRNA template encoding YPet (yellow fluorescent protein)[13] that has the amber codon in the upstream of the YPet gene (amber template, Fig. 1B).[14] As a result, S1, S2 and S3 derived from $tRNA^{Ser}_{CGA}$ or $tRNA^{Ser}_{UGA}$ relatively efficiently suppressed the amber codon (Fig. 1C). 24 ± 2% was found to be the absolute suppression efficiency of the best suppressor (S2), which is the percentage of the fluorescence intensity of YPet translated from the amber template with the IVT-sup-tRNA compared to that from an amber-free mRNA template (UCU for Ser instead of the amber codon, Fig. 1B) in the presence of the IVT-sup-tRNA.[15] Given the fact that sup-tRNAs must compete with eRFs, this suppression efficiency is modestly high. It is comparable to the *in vivo* suppression efficiency (20–25%) of the bean-derived sup-$tRNA^{Leu}_{CAA\rightarrow CUA}$ whose bases should be modified in tobacco protoplasts,[7] though bases in the IVT-sup-tRNA are expected to be modified only rarely in wheat germ extract with an only 1 h incubation period. In fact, whereas L6 has almost the same sequence as the sup-$tRNA^{Leu}_{CAA\rightarrow CUA}$ used in tobacco protoplasts,[7] it showed a much lower suppression efficiency, probably because bases in L6 are not completely modified. Therefore, the high suppression efficiency of some IVT-sup-$tRNA^{Ser}_{CUA}$ indicates that base modification of these IVT-sup-tRNAs is not necessarily essential, though, of course, it would be better for the tRNAs. These results also show that wheat SerRS does not recognize the anticodon of $tRNA^{Ser}_{CGA}$ and $tRNA^{Ser}_{UGA}$ as well as SerRS in *E. coli* and humans.[5b,d] On the other hand, IVT-sup-tRNAs from $tRNA^{Ser}_{AGA}$, $tRNA^{Ser}_{GGA}$, and $tRNA^{Ser}_{GCU}$ hardly suppressed the amber codon. These results suggest that these tRNAs require base modification or that the mechanism of wheat SerRS for recognizing $tRNA^{Ser}$ is different among isodecoders. Other IVT-sup-tRNAs from $tRNA^{Leu}$ and $tRNA^{Ala}$ are also poor suppressors, except for L7 from $tRNA^{Leu}_{UAA}$, which is still poorer than S1, S2, and S3. Their low suppression efficiencies can be explained in the same way as for some poor $tRNAs^{Ser}_{CUA}$.

Chimerization of S1 and S2 to construct the second generation

Among the three efficient IVT-sup-tRNAs in the first generation, S1 is nearly identical to S3, with only a three-base difference,

A)

isotype	anticodon	gene copy	IVT-sup-tRNA
Ser	CGA	8	S1(6)
	UGA	17	S2(7), S3(6)
	AGA	12	S4(9)
	GGA	4	S5(4)
	GCU	20	S6(10), S7(7)
Leu	CAG	10	L1(8)
	UAG	10	L2(5), L3(4)
	AAG	15	L4(14)
	CAA	19	L5(9), L6(7)
	UAA	4	L7(4)
Ala	CGC	12	A1(10)
	UGC	10	A2(10)
	AGC	20	A3(19)
	GGC	1	A4(1)

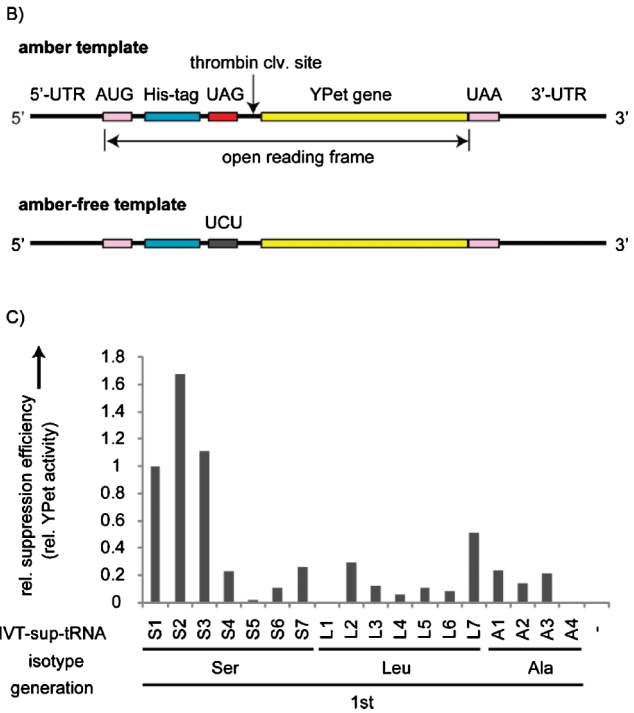


Fig. 1 (A) Classification of $tRNA^{Ser}$, $tRNA^{Leu}$, $tRNA^{Ala}$ in the *Oryza sativa* genome by their natural anticodon, which was converted into CUA for the amber codon to construct representative IVT-sup-tRNAs (first generation). The number in parenthesis in the 'IVT-sup-tRNA' column indicates the gene copy number of tRNAs having sequences (nearly) identical to the representative. The sequences of these tRNAs are summarized in the ESI†. (B) Schematic diagram of mRNA templates used. (C) Relative suppression efficiency of IVT-sup-tRNAs to S1.[14]

despite them being derived from different isodecoders (Fig. 2A). In contrast, uniform parts between S1 and S2 are only a T-loop and an anticodon-loop (Fig. 2A). We therefore speculated that S1 and S2 have individual and different attributes as suppressors, *e.g.* the accurate active tRNA structure or the avidity for SerRS and elongation factors. If so, replacement of their parts with each other may improve the suppression efficiency. To confirm this assumption, we partitioned S1 and S2 into several parts, as shown in Fig. 2A (acceptor arm (Aarm); D-arm (Darm) & T-arm (Tarm);[16] anticodon arm (ACarm); variable arm (Varm)) and replaced a part with the other's to produce chimeric

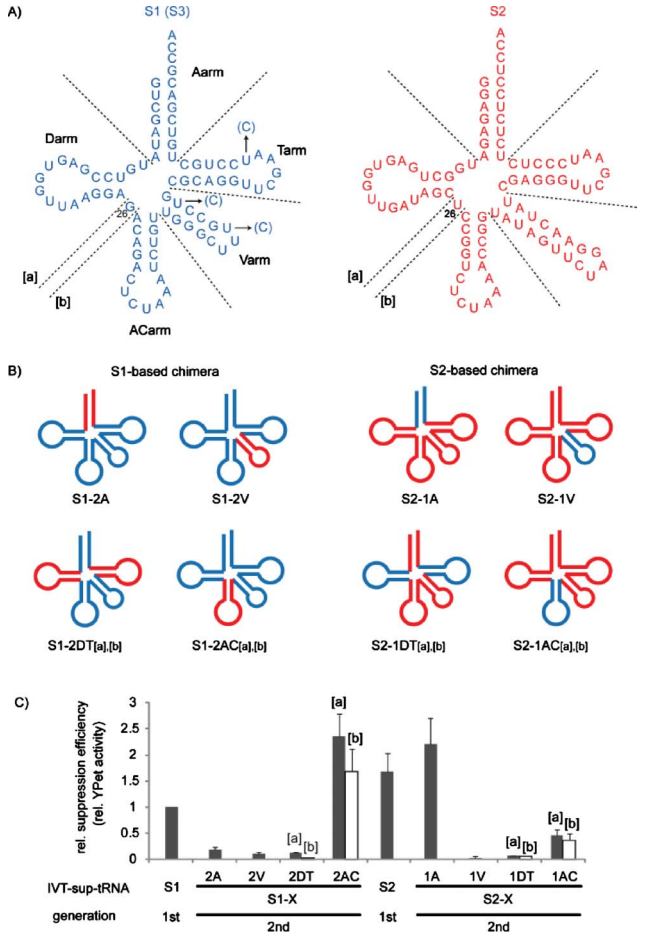


Fig. 2 (A) The sequences of S1 (left) and S2 (right) in the first generation. The bases in parentheses in S1 represent the differences between S1 and S3. The dotted lines indicate the partition sites. (B) Schematic diagrams of chimeric IVT-sup-tRNAs (second generation). The blue and red lines show bases derived from S1 and S2, respectively. (C) Relative suppression efficiency of chimeric IVT-sup-tRNA to S1.[14]

IVT-sup-tRNAs (the second generation, Fig. 2B). Because the base between Darm and ACarm is different between S1 (G26) and S2 (U26),[17] the borderline in this site was set in two ways ([a] and [b] in Fig. 2A). The suppression abilities of these chimeric IVT-sup-tRNAs were evaluated as described above (Fig. 2C). As expected, two kinds of chimeras (S1-2AC$_{[a]}$ and S2–1A) showed higher suppression efficiencies than their original IVT-sup-tRNAs (S1 and S2, respectively). Notably, S1-2AC$_{[a]}$, which is the S1-based chimera whose ACarm including the borderline base is replaced with that of S2, suppressed the amber codon more than twice as well as S1. Meanwhile, the suppression efficiency of S2–1A, which is based on S2 and whose Aarm is derived from S1, was slightly higher than that of S2 and as high as that of S1-2AC$_{[a]}$. These results indicate that ACarm in S2 and Aarm in S1 are independently effective for suppression. In addition, the results in Fig. 2C suggest the following features of IVT-sup-tRNA$^{Ser}_{CUA}$: Varm should be grouped with Darm and Tarm, meaning that these three parts interact with each other; the borderline between Darm and ACarm should be set in [a], indicating that the borderline base at the 26th position is related to ACarm. Incidentally, although doubly chimeric IVT-sup-tRNAs, in which two parts were replaced, were also prepared, none overcame the suppression efficiency of the original IVT-sup-tRNAs (see Fig. S1 in the ESI†).

Comparison between S1-2AC$_{[a]}$ and S1 leading to the third generation

The different parts between S1-2AC$_{[a]}$ and its parent S1 are the anticodon stem and the borderline base because the anticodon-loop is identical between S1 and S2 (Fig. 3A). As regards the anticodon stem, its strength in S1-2AC$_{[a]}$ (*i.e.* in S2) is larger due to being GC-rich. We therefore presumed that the stronger stem probably contributes to formation of a rigid active structure of tRNA and thus to higher suppression ability. To verify this, the C27–G43 and G29–C41 pairs in S1-2AC$_{[a]}$, both of which are A–U pairs in S1, were replaced with G27–C41 and/or C29–G41 pairs, U–A pairs, or A–U pairs (Fig. 3A). As a result, whereas the substitution with G–C and/or C–G pairs increased suppression efficiencies, both the U–A and A–U pairs gave lower suppression efficiencies (Fig. 3B). However, given the fact that S1-2AC$_{[a]}$-$A_{27}U_{43}$-$A_{29}U_{41}$ (can be named S1-U_{26}-$U_{31}A_{39}$) is different from S1 only in the borderline base and the 31–39 base pair, its suppression ability was relatively higher than that of S1. Thus, the borderline base G26 or the A31–U39 base pair in S1 was converted to U or the

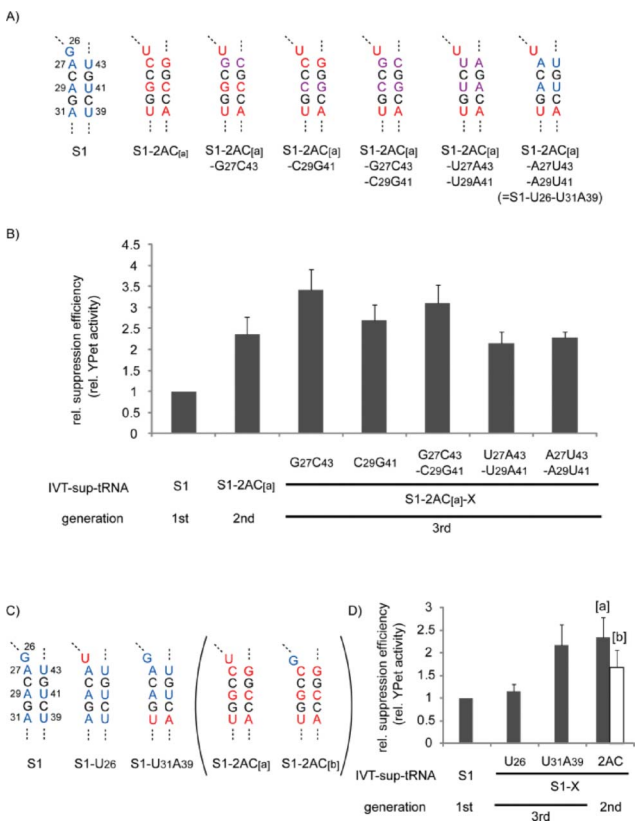


Fig. 3 A comparison between S1 and S1-2AC$_{[a]}$ leading to the third generation based on S1. (A, C) The sequences of the anticodon-stem and the 26th base, which are the different parts between S1 and S1-2AC$_{[a]}$, in various S1-based IVT-sup-tRNAs. The numbers adjacent to the bases indicate widely used nucleotide positions in tRNAs. The blue, red, black, and purple letters represent bases specific to S1 or S2, common to them, or in neither of them, respectively. (B, D) Relative suppression efficiency of IVT-sup-tRNA to S1.[14]

U–A pair in S1-2AC$_{[a]}$-A$_{27}$U$_{43}$-A$_{29}$U$_{41}$, respectively, to investigate which part plays a more important role in suppression (Fig. 3C). As in Fig. 3D, the conversion of the borderline base from G26 to U26 hardly showed an improvement in the suppression ability (S1-U$_{26}$). Although this result seems to be contradictory to the difference in suppression between S1-2AC$_{[a]}$ and S1-2AC$_{[b]}$, G26 perhaps somewhat distorts the anticodon stem only in the case of a G–C rich anticodon stem. On the other hand, the U31–A39 pair dramatically increased the suppression efficiency up to that comparable to S1-2AC$_{[a]}$ (Fig. 3D, S1-U$_{31}$A$_{39}$). By contrast, when the U31–A39 pair in S1-2AC$_{[a]}$ was replaced with A–U, the suppression efficiency decreased to that comparable to S1 (Fig. S2 in the ESI†). In addition, replacement of this base pair in S1-2AC$_{[a]}$ with G31–C39, C31–G39, G31–U39, and U31–G39 pairs on the whole diminished the suppression efficiency (Fig. S2 in the ESI†). These results clearly suggest that not the strength of the anticodon stem but the U31–A39 pair makes the larger contribution to the higher suppression ability of S1-2AC$_{[a]}$ than S1. It is probably because this pair is the root end of the anticodon-loop and thus is sensitive to alteration of the anticodon. The U31–A39 pair may be suitable for flipping out the anticodon CUA, which is consistent with the fact that one of the most efficient amber IVT-sup-tRNA in *E. coli*, tRNA$^{SerU}_{CUA}$, has the U31–A39 pair with the same anticodon-loop sequence as that of S1 and S2.[4a] In this manner, a series of alterations of the anticodon stem in S1-2AC$_{[a]}$ or S1 revealed the effective part for amber suppression. In addition, it generated various kinds of IVT-sup-tRNAs as in Fig. 3A and 3C (the S1-based third generation), among which a better suppressor, S1-2AC$_{[a]}$-G$_{27}$C$_{43}$, was discovered. This suppressor showed 3.4-fold and 1.5-fold higher suppression efficiency than S1 and S1-2AC$_{[a]}$, respectively.

Comparison between S2–1A and S2 leading to the other third generation

We also investigated the other promising suppressor in the second generation, S2–1A. It is mainly composed of S2, and the different parts between them are the acceptor stem and the 73th base, called a discriminator base (Fig. 4A). Because the sequence in the acceptor stem is known to be crucial for recognition by aaRS and is too complex to be deciphered, the discriminator base was focused on in the comparison. In tRNAs for Ser from various organisms, G73 is well-conserved. However, in contrast to the name 'discriminator', yeast and *E. coli* SerRS, unlike mammalian

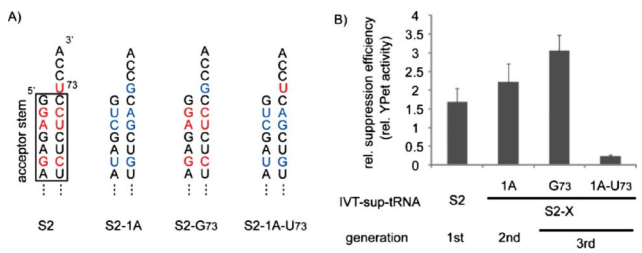


Fig. 4 A comparison between S2 and S2–1A leading to the third generation based on S2. (A) The sequences of Aarm, which is the different part between S2 and S2–1A, in various S2-based IVT-sup-tRNAs. The letter colours have the same meaning as in Fig. 3. The box indicates the acceptor stem. (B) Relative suppression efficiency of IVT-sup-tRNA to S1.[14]

SerRS, do not predominantly recognize this base, at least *in vitro*.[8] In *Oryza sativa*, the discriminator base is almost G, but U in the S2 family. Therefore, the discriminator base in Poaceae tRNAs for Ser appears to not be an identity determinant. Nonetheless, because G73 is major, we attempted to substitute U73 in S2 with G, as in S1 (and S2–1A), to examine the effect of G73 (Fig. 4A). As a result, surprisingly, G73 substitution in S2 approximately doubled the suppression efficiency (Fig. 4B, S2-G$_{73}$). In contrast, when G73 in S2–1A was replaced with U, suppression hardly occurred (Fig. 4 B, S2–1A-U$_{73}$). These results indicate that the increase in suppression efficiency from S2 to S2–1A can be attributed to G73, which is probably better recognized by SerRS. In addition, they show that U73 is available only with the acceptor stem in S2, and that G73 is better suited, even with the acceptor stem in S2. In fact, G73U-mutated S1 also induced almost no suppression (suppression efficiency = 0.18 ± 0.1 relative to S1). Therefore, wheat SerRS seems to recognize Aarm in S1-based tRNAs and S2-based tRNAs in different ways: it primarily recognizes the discriminator base only in the former. Moreover, the suppression efficiency of S2-G$_{73}$, regarded as one of the S2-based third generation, corresponds to a 1.4-fold higher level than that of S2–1A (Fig. 4B), indicating that the acceptor stem in S2 was found to be more favorable than that in S1, at least in S2-based tRNAs.

Combination of the optimized parts in S1-2AC$_{[a]}$-G$_{27}$C$_{43}$ and S2-G$_{73}$ to construct the fourth generation

We have now obtained two kinds of highly efficient amber suppressors, S1-2AC$_{[a]}$-G$_{27}$C$_{43}$ and S2-G$_{73}$, from different routes by optimizing IVT-sup-tRNA$^{Ser}_{CUA}$ sequences. As shown in Fig. 5A, S1 and S2, efficient IVT-sup-tRNAs in the first generation, were improved by chimerization to S1-2AC$_{[a]}$ and S2–1A in the second generation, respectively, which was further improved to S1-2AC$_{[a]}$-G$_{27}$C$_{43}$ and S2-G$_{73}$ in the third generation, respectively. We next combined the optimized parts of these two suppressors in the third generation to construct S1-2AC$_{[a]}$-G$_{27}$C$_{43}$-2A-G$_{73}$ and S2-G$_{27}$C$_{43}$-G$_{73}$ as the fourth generation (Fig. 5A (far right) and B). Although both of them have the anticodon stem with U26 in S1-2AC$_{[a]}$-G$_{27}$C$_{43}$ and the acceptor stem with G73 in S2-G$_{73}$, other parts in the former and the latter are based on S1 and S2, respectively. Unfortunately, the former (S1-2AC$_{[a]}$-G$_{27}$C$_{43}$-2A-G$_{73}$) showed a slightly lower suppression efficiency than S1-2AC$_{[a]}$-G$_{27}$C$_{43}$ (Fig. 5C, blue, far right). This was, however, as expected because dc2$_{[a]}$, which is nearly identical to S1-2AC$_{[a]}$-G$_{27}$C$_{43}$-2A-G$_{73}$ except for the C27–G43 pair and the U73, was a very poor suppressor (Fig. S1 in the ESI†). Nonetheless, the relatively higher suppression efficiency of S1-2AC$_{[a]}$-G$_{27}$C$_{43}$-2A-G$_{73}$ than dc2$_{[a]}$ revealed again the large positive effects of the G27–C43 pair and the G73. In contrast, S2-G$_{27}$C$_{43}$-G$_{73}$ was expected to further improve the suppression efficiency, taking into account the fact that it is identical to S2-G$_{73}$ with the G27–C43 pair instead of C27–G43 and that S1-2AC$_{[a]}$-G$_{27}$C$_{43}$ obtained from S1-2AC$_{[a]}$ by this substitution showed 1.5-fold higher suppression efficiency than its parent, S1-2AC$_{[a]}$. In fact, as expected, it suppressed the amber codon 1.2-fold more than S2-G$_{73}$ (Fig. 5C, red, far right). In total through S2-based four generations, it was twice more efficient than the great-grandparent (S2). As described above, in the case of S1-based generations, the suppression efficiency increased 3.4-fold from the first (S1) to the third (S1-2AC$_{[a]}$-G$_{27}$C$_{43}$). Although, it was surprising to us that

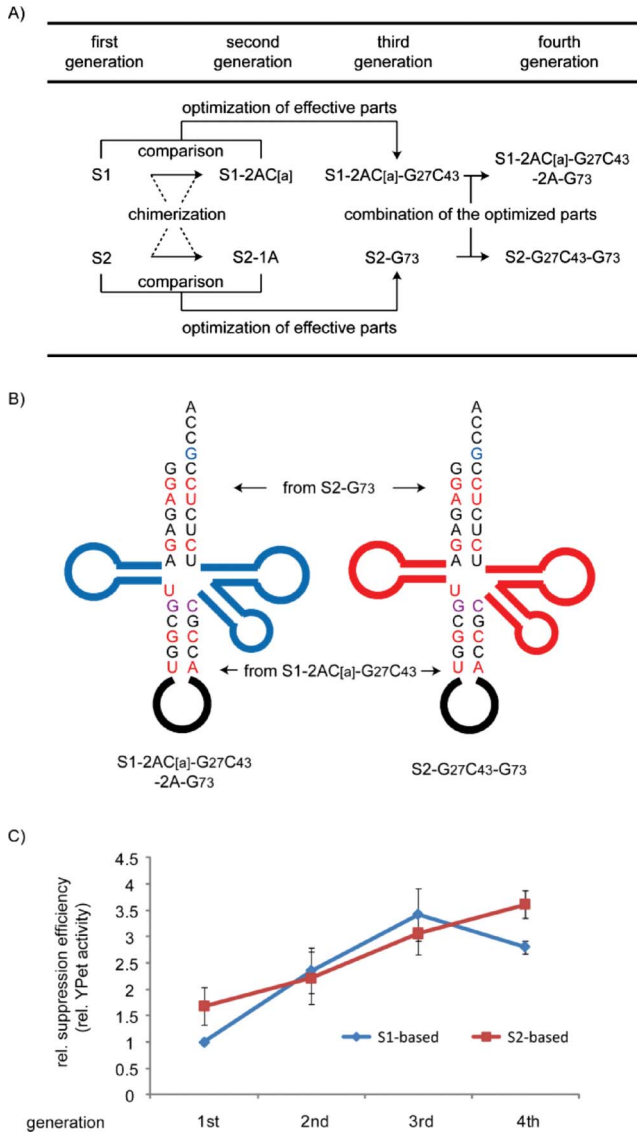


Fig. 5 (A) Schematic diagram of IVT-sup-tRNA sequence optimization through four generations. The S1-based (above) and S2-based (below) most efficient IVT-sup-tRNAs in each generation and methods for improvement between generations are summarized. (B) The sequences of the S1-based (left) and S2-based (right) fourth generation, both of which have the anticodon-stem and the U26 from S1-2AC$_{[a]}$-G$_{27}$C$_{43}$ and the acceptor-arm from S2-G$_{73}$. The line and letter colours have the same meaning as in Fig. 2 and 3, respectively. (C) Relative suppression efficiency of the best IVT-sup-tRNA in each generation to that of S1.[14] The blue and red lines represent S1-based and S2-based IVT-sup-tRNAs, respectively.

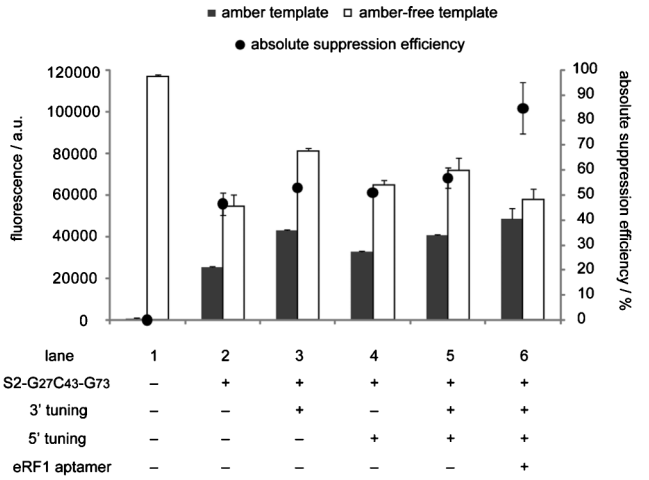


Fig. 6 Fluorescence of YPet translated from the amber template (closed bar) or the amber-free template (open bar) in the absence or presence of the best IVT-sup-tRNA, S2-G$_{27}$C$_{43}$-G$_{73}$, with termini tuning and/or an eRF1 aptamer. The filled circles show the absolute suppression efficiency (right Y-axis), which is the percentage of the fluorescence intensity of YPet translated from the amber template with the IVT-sup-tRNA compared to that from the amber-free template in the presence of the IVT-sup-tRNA.

anticodon-adjusted natural tRNAs (first generation) have room to be so considerably improved, it is probably because anticodon alteration and incomplete base modification of natural tRNAs quite reduce their superiority in sequence, or perhaps because natural tRNA sequences are originally not optimal for high translation efficiency due to the balance with other tRNAs *in vivo*.

Fine tuning of the termini in the best suppressor, S2-G$_{27}$C$_{43}$-G$_{73}$

S2-G$_{27}$C$_{43}$-G$_{73}$ is now the best suppressor. The absolute suppression efficiency of S2-G$_{27}$C$_{43}$-G$_{73}$ was 46 ± 4% (Fig. 6, lane 2), which is approximately twice as high as that of S2. It is also comparable to the suppression efficiency of one of the best prokaryotic IVT-suppressors, tRNA$^{Leu5}_{CUA}$, in *E. coli* extract (44%).[4c] To further improve the suppression efficiency, we next fine-tuned the termini of S2-G$_{27}$C$_{43}$-G$_{73}$. In terms of the 3′ terminus, DNA templates modified with 2′-methoxy-dG at the second base from the 5′ terminus were used to reduce non-templated 3′ nucleotide addition by T7 RNA polymerase during *in vitro* transcription of tRNAs.[18] As far as the 5′ terminus is concerned, guanosine monophosphate (GMP) was used to produce 5′ monophosphoryl tRNAs, whose 5′ ends are different from 5′ triphosphate generated by normal *in vitro* transcription but are similar to those of cellular tRNAs processed by RNase P.[19] As a result, both of the 3′ and 5′ tuning slightly increased the suppression efficiency (Fig. 6, lanes 3 and 4). Notably, the 3′-tuned tRNA reduced the inhibition effect on translation both in the amber and the amber-free template, although the reason why is not yet clear. The 5′-3′-doubly tuned tRNA showed an absolute suppression efficiency of 57 ± 4%, which is 1.2-fold higher than that of the non-tuned original tRNA, S2-G$_{27}$C$_{43}$-G$_{73}$ (Fig. 6, lane 5). It is therefore concluded that terminal tuning of tRNAs is somewhat significant for *in vitro* activation of tRNAs.

Improvement of the suppression efficiency by an aptamer bound to eRF1

One possible reason for the absolute suppression efficiency of the optimized IVT-sup-tRNA still being much lower than 100% is that the IVT-sup-tRNA competes with eRFs. In *E. coli* systems, a reconstituted translation system without RFs[4a,4b,20] or an RNA aptamer tightly binding to and inhibiting RF1[4c] have been used to reduce sup-tRNA/RF competition and to increase suppression efficiencies. In eukaryotic systems, an RNA aptamer for human eRF1 has also been selected[21] and it works well to facilitate pre-aminoacylated tRNA to be incorporated at nonsense codons, even in wheat germ extract.[3a] Thus, we used the human eRF1-aptamer to investigate whether it further increases the suppression

efficiency of the 5′-3′-doubly tuned S2-$G_{27}C_{43}$-G_{73}. As a result, despite the eRF1 aptamer slightly inhibiting translation termination on the genuine stop codon (UAA) (Fig. 6, lanes 5 *vs.* 6, open bar) because eRF1 recognizes all of the stop codons, the translation efficiency of the amber template increased (Fig. 6, lanes 5 *vs.* 6, closed bar). Consequently, the absolute suppression efficiency was up to 85 ± 10%, which is comparable to that of the *E. coli* IVT-sup-$tRNA^{Leu5}_{CUA}$ with an RF1 aptamer (89%).[4c] Incidentally, the relative translation efficiency of the amber template in the presence of the IVT-sup-tRNA and the aptamer (Fig. 6, lane 6, closed bar) in comparison to that of the amber-free template in the absence of any exogenous RNA (Fig. 6, lane 1, open bar), which is named here as 'expression efficiency', was 41 ± 4% and thus the absolute amount of the expressed protein was relatively large.

Suppression in large-scale protein synthesis

Although the experiments described above were performed in batch mode, that is, small-scale protein synthesis, the wheat germ cell-free translation system has another translation mode based on dialysis that can produce a large amount of protein.[12] Because some applications of IVT-sup-tRNAs possibly require a volume of read-through protein, we next examined the suppression efficiency of the 5′-3′-doubly tuned S2-$G_{27}C_{43}$-G_{73} in the large-scale protein synthesis. Large-scale translation was performed for three days under the conditions such that the concentrations of the IVT-sup-tRNA and the aptamer inside a dialysis cup were identical to those in batch mode. As shown in Fig. 7, the IVT-sup-tRNA suppressed the amber codon in the 3-d large-scale synthesis as efficiently as in the 1 h small-scale protein synthesis, indicating that the IVT-sup-tRNA is recycled as well as natural tRNAs. The absolute suppression efficiency and expression efficiency were calculated from the fluorescence to be 78% and 33%, respectively, which are comparable to those in the small-scale synthesis (85% and 41%, respectively). The amount of protein expressed through suppression was approximately 100-fold greater (approx. 5 μg/10 μL extract) than that in the small-scale synthesis, which is large enough to further analyze the read-through protein not only in the enzymatic activity but also in various ways including amino acid sequence analyses and mass analyses (*vide infra*).

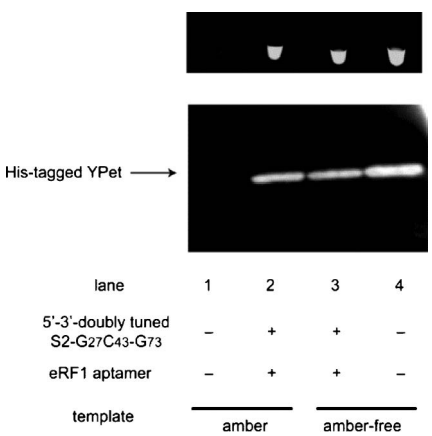


Fig. 7 Fluorescence images (above) and Western blot (below) of His–tagged YPet translated from the amber template or the amber-free template in the absence or presence of the 5′-3′-doubly tuned S2-$G_{27}C_{43}$-G_{73} with the eRF1 aptamer in a large-scale protein synthesis.

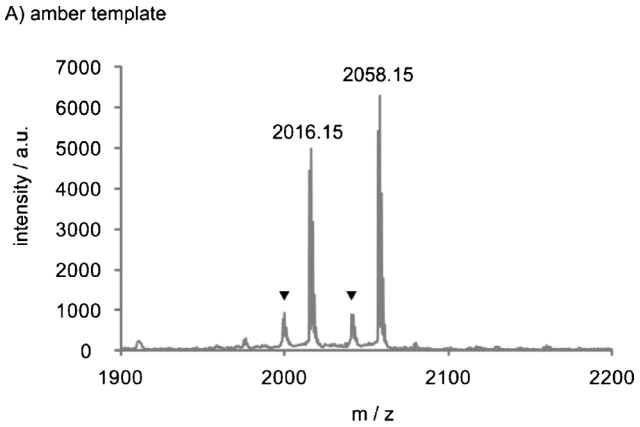
Verification of Ser incorporation to the amber codon with MALDI-TOF-MS

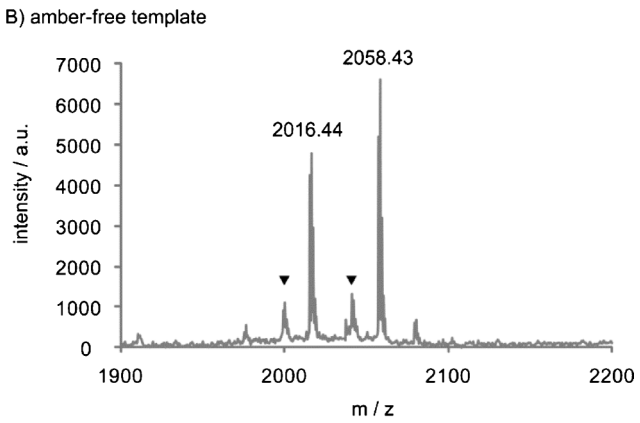
The most efficient suppressor S2-$G_{27}C_{43}$-G_{73} is originally based on the natural $tRNA^{Ser}$. However, some base mutations on the tRNA may disrupt the orthogonality of aaRS/tRNA pairs. In particular, alteration of the discriminator base is likely to cause mis-aminoacylation, although Aarm in the IVT-sup-tRNA is identical to that in S5 and S7 families, except for C69, and is different from those in all other tRNAs in *Oryza sativa*. We thus finally verified which amino acid is incorporated into the amber codon in suppression by the 5′-3′-doubly tuned S2-$G_{27}C_{43}$-G_{73}. His-tagged YPet expressed through suppression with the IVT-sup-tRNA on the amber template in a large-scale synthesis was purified and cleaved by thrombin (Fig. 1B). The N-terminal cleaved fragment including the amino acid X corresponding to the amber codon (^{N}AHHHHHHXRYSSGLVPR^{C}, calculated mass for $(M+H)^+$: 2016.17 for X = S; 1986.14 for X = the smallest natural amino acid G; 2115.30 for X = the largest natural amino acid W) was then analyzed by MALDI-TOF-MS (Fig. 8).[22] The obtained mass spectrum (Fig. 8A) was nearly identical to that of a positive control (an N-terminal fragment cleaved from His-tagged protein expressed on the amber-free template, Fig. 8B). Two main peaks 2016.15 and 2058.15 were consistent with the calculated mass for Ser-incorporated peptide (2016.17) and for the N-acetylated peptide (2058.20), respectively (Fig. 8C).[23] Therefore, it is undoubtedly true that Ser was attached to the IVT-sup-tRNA by SerRS and then incorporated at the amber codon.

Conclusions

In conclusion, we obtained a highly efficient amber IVT-sup-$tRNA^{Ser}$, that functions in wheat germ extract, through four sequence optimization steps: (1) anticodon alteration of natural tRNAs, (2) chimerization of the efficient IVT-sup-tRNAs in the first generation, (3) investigation and optimization of the effective parts through comparison between the first and second generations, and (4) combination of the optimized parts in the third generation. The best of the youngest, fourth generation (S2-$G_{27}C_{43}$-G_{73}) showed suppression approximately twice as high as that of the best original IVT-sup-tRNA in the oldest, first generation (S2), which was constructed just by alteration of the anticodon of natural tRNAs for Ser. Moreover, the terminal tuning of the IVT-sup-tRNA and the aptamer for eRF1 were found to be effective for suppression. Eventually, the absolute suppression efficiency of the best IVT-sup-tRNA with the aptamer was quite high, 85% and 78% for a small-scale and a large-scale cell-free protein synthesis, respectively, despite incomplete base modifications in its body. In addition, it was verified with mass spectrometry that Ser was incorporated on the amber codon by the IVT-sup-tRNA, suggesting that Ser was attached to the IVT-sup-tRNA by SerRS.

The best IVT-sup-tRNA obtained here is considerably highly efficient sup-tRNA that suppresses the amber codon in a cell-free eukaryotic translation system without pre-aminoacylation and posttranscriptional base modification. The absolute suppression efficiency was 85% with the eRF1 aptamer and nearly 60% even without the aptamer, which is comparable to that of bacterial efficient IVT-sup-tRNAs.[4] Because the IVT-sup-tRNA can easily





C) calculated mass

amino acid sequence	mass for $(M+H)^+$
AHHHHHHSRYSSGLVPR	2016.17
Acetyl-AHHHHHHSRYSSGLVPR	2058.20

Fig. 8 MALDI-TOF-MS analyses of the N-terminal fragment of YPet including the amino acid incorporated on the amber codon. (A, B) Mass spectra of the fragment from YPet translated on the amber template (A) and the amber-free template (B) in the presence of the 5′-3′-doubly tuned S2-$G_{27}C_{43}$-G_{73} with the eRF1 aptamer in a large-scale protein synthesis. See the note 23 for the arrow heads.[23] (C) Calculated mass of the fragment. The underlined Ser is the amino acid corresponding to the amber codon.

be prepared and can effectively switch on complete translation of amber-mutated template mRNA, it can be available in various applications. It can be used as an element of biofunction-assisted biosensors in combination with the amber-mutated reporter gene such as luciferase and GFP.[4b] At the same time, it is available for an *in vitro* read-through ribosome display method, which can display proteins on their template mRNAs with nonsense condons without altering the mRNAs,[4a] with natural mRNAs in not only plants but also other eukaryotes because various eukaryotic mRNAs can be translated in the wheat germ extract used here.[12,24] Although efficient IVT-sup-tRNAs for other nonsense codons (opal and ochre) are required for this method, they can probably be obtained by the optimization strategy described herein. In addition, because it is assumed that bases in the IVT-sup-tRNA were completely modified *in vivo* unlike *in vitro*, the sup-tRNA may function *in vivo* more efficiently than or at least as efficiently as *in vitro*. Therefore, it may also be used for gene regulators in plants and be a useful tool for research regarding plant tRNA processing and function.[25]

The optimization strategy used in the present study and the knowledge obtained here are also useful for the design of amber sup-tRNAs for naa. Although an efficient orthogonal naaRS/sup-tRNA pair or naa-misacylated sup-tRNA is required for naa incorporation into the protein of interest, it is often that the optimization of the sup-tRNA is insufficient in comparison to that of naaRS. In addition, efficient incorporation of naa '*in vitro*' is desirable because naa is generally toxic to cells.[3a] The strategy, in particular partition and chimerization of efficient sup-tRNAs, and the knowledge to achieve higher suppression efficiency (*e.g.* significance of the 31–39 base pair in the anticodon alteration and the terminal tuning *in vitro*) would probably improve sup-tRNAs for naa to better ones.

Experimental

Preparation of DNA templates for mRNAs and tRNAs

DNA templates for *in vitro* transcription were prepared with standard polymerase chain reactions (PCRs) by using PrimeSTAR Max DNA Polymerase (Takara Bio) according to the manufacturer's protocol. The pMK-RQ-based plasmids encoding the YPet gene under the SP6 promoter sequence, pHis-TAG-RY-YPet for the amber template and pHis-SRY-YPet for the amber-free template, which are PCR templates for mRNAs, were artificially synthesized by Mr. Gene (Regensburg, German). PCR templates for tRNAs were synthesized by Invitrogen. Normal primers for PCRs were purchased from Invitrogen or Operon Biotechnologies. The forward primers for tRNA templates included the T7 promoter sequence for *in vitro* transcription. The reverse primers for tRNA templates were purified by oligonucleotide purification cartridge (OPC) to remove short primers generated during primer syntheses. An OPC-purified reverse primer including a 2′-methoxy base was from Tsukuba Oligo Service. The sequences of templates and primers are summarized in the ESI†.

Preparation of mRNAs

Run-off transcription of the obtained DNA templates having the SP6 promoter sequence was performed with an AmpliScribe SP6 High Yield Transcription Kit (Epicentre, Wisconsin, USA) according to the manufacturer's protocol. The transcribed mRNA was purified with an RNeasy MinElute Cleanup Kit (QIAGEN) and quantified by the absorbance at 260 nm.

Preparation of tRNAs

tRNAs were constructed by *in vitro* run-off transcription of the DNA templates including the T7 promoter sequence with an AmpliScribe T7 High Yield Transcription Kit (Epicentre) according to the manufacturer's protocol. In preparing 5′-monophosporyl tRNA, GMP (Wako) was added to the reaction mixture in such a way that the final concentration was 75 mM, which was 10-fold higher than each NTP. The transcribed tRNA was purified with a QIAquick Nucleotide Removal Kit (QIAGEN) and quantified by the absorbance at 260 nm.

Translation in wheat germ extract

Cell-free translation was carried out by using a WEPRO1240 Expression Kit (CellFree Sciences), as described previously, with slight modifications.[12,26] For batch-mode translation, a mixture (10 μL) of mRNA (3 pmol), tRNA (0 or 50 pmol), *in vitro* transcribed eRF1 aptamer-12 (0 or 10 pmol),[21] WEPRO1240 (wheat germ extract, 2 μL), creatine kinase (final concentration: 40 ng μL^{-1}), and SUB-AMIX (final concentration: 1×) containing materials (amino acids, ATP, GTP *etc.*), the latter three of which were attached to the kit, was incubated at 26 °C for 1 h. For large-scale translation, a mixture (60 μL) of unpurified mRNA (10 μL) prepared with the RNA polymerase attached to the kit according to the manufacturer's protocol, tRNA (0 or 300 pmol), eRF1 aptamer-12 (0 or 60 pmol), WEPRO1240 (10 μL), creatine kinase (final concentration: 100 ng μL^{-1}), and 1× SUB-AMIX (37 μL) was poured inside a dialysis cup (MWCO = 3500) that outside was filled with 1× SUB-AMIX (750 μL), and was incubated at 16 °C for 3 days.

YPet assay

The fluorescence intensities of translation solution diluted 5-fold with water were measured by using an excitation wavelength of 485 nm and an emission wavelength of 535 nm on a Wallac ARVO MX (Perkin-Elmer) with a black 96-well plate. Fluorescence images were acquired using a Visirays-B (ATTO) and a Printgraph AE-6981 FXCP (ATTO) equipped with an appropriate filter (SCF515).

Western blot

An aliquot (1 μL) of the translation solution in a large-scale synthesis was resolved by 15% SDS-PAGE. The resultant gel was transferred onto PVDF membrane (Invitrogen), and Western blot analysis was performed. His-tagged proteins were visualized with a Penta-His HRP conjugate (QIAGEN) and ECL Plus Western Blotting Detection Reagent (GE Healthcare). Chemiluminescence images were acquired using a Light-Capture (ATTO).

MALDI-TOF-MS analyses

The His-tagged YPet synthesized on a large scale was purified with a MagExtractor –His-tag– kit (TOYOBO) and quantified by the absorbance at 517 nm. 1 U of Thrombin (Novagen) was added to the purified protein (approx. 15 pmol) to cleave out the N terminus containing the amino acid corresponding to the amber codon. The cleavage was verified with SDS-PAGE. The cleaved protein (approx. 10 pmol) was desalted and concentrated with a ZipTip (Millipore) and then analyzed with a Voyager DE Pro Biospectrometry Workstation (Applied Biosystems) with a positive mode by using CHCA as a matrix.

Acknowledgements

This work was supported by 'Special Coordination Funds for Promoting Science and Technology' from the Ministry of Education, Culture, Sports, Science and Technology, the Japanese Government.

Notes and references

1 H. Beier and M. Grimm, *Nucleic Acids Res.*, 2001, **29**, 4767–4782; and references therein.
2 (*a*) G. Eggertsson and D. Söll, *Microbiol. Rev.*, 1988, **52**, 354–374; (*b*) J. P. Capone, P. A. Sharp and U. L. RajBhandary, *EMBO J.*, 1985, **4**, 213–221; (*c*) S. Franklin, T. Y. Lin and W. R. Folk, *Plant. J.*, 1992, **2**, 583–584; (*d*) J. Xie and P. G. Schultz, *Nat. Rev. Mol. Cell Biol.*, 2006, **7**, 775–782.
3 (*a*) J. Gubbens, S. J. Kim, Z. Yang, A. E. Johnson and W. R. Skach, *RNA*, 2010, **16**, 1660–1672; (*b*) J. W. Chin, T. A. Cropp, J. C. Anderson, M. Mukherji, Z. Zhang and P. G. Schultz, *Science*, 2003, **301**, 964–967.
4 (*a*) A. Ogawa, S. Sando and Y. Aoyama, *ChemBioChem*, 2006, **7**, 249–252; (*b*) A. Ogawa and M. Maeda, *ChemBioChem*, 2008, **9**, 2204–2208; (*c*) S. Sando, A. Ogawa, T. Nishi, M. Hayami and Y. Aoyama, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 1216–1220.
5 (*a*) P. J. Beuning and K. Musier-Forsyth, *Biopolymers*, 1999, **52**, 1–28; (*b*) J. R. Sampson and M. E. Saks, *Nucleic Acids Res.*, 1993, **21**, 4467–4475; (*c*) G. Tocchini-Valentini, M. E. Saks and J. Abelson, *J. Mol. Biol.*, 2000, **298**, 779–793; (*d*) T. Achsel and H. J. Gross, *EMBO J.*, 1993, **12**, 3333–3338; (*e*) K. Breitschopf, T. Achsel, K. Busch and H. J. Gross, *Nucleic Acids Res.*, 1995, **23**, 3633–3637; (*f*) R. Geslain and T. Pan, *J. Mol. Biol.*, 2010, **396**, 821–831.
6 A. Soma, R. Kumagai, K. Nishikawa and H. Himeno, *J. Mol. Biol.*, 1996, **263**, 707–714.
7 V. T. C. Carneiro, G. Pelletier and I. Small, *Plant Mol. Biol.*, 1993, **22**, 681–690.
8 H. Himeno, S. Yoshida, A. Soma and K. Nishikawa, *J. Mol. Biol.*, 1997, **264**, 704–711.
9 (*a*) C. Francklyn and P. Schimmel, *Nature*, 1989, **337**, 478–481; (*b*) K. Shiba, T. Ripmaster, N. Suzuki, R. Nichols, P. Plotz, T. Noda and P. Schimmel, *Biochemistry*, 1995, **34**, 10340–10349.
10 tRNA sequences are highly homologous in the same family; *e.g.* tRNA sequences in *Oryza sativa* and *Zea mays*, both of which are Poaceae, are nearly identical.
11 We chose as the representative(s) the major type(s) among all isodecoder genes determined by structural alignment, which was referred to the genomic tRNA database site (http://gtrnadb.ucsc.edu/).
12 (*a*) K. Madin, T. Sawasaki, T. Ogasawara and Y. Endo, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 559–564; (*b*) T. Sawasaki, T. Ogasawara, R. Morishita and Y. Endo, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 14652–14657; (*c*) N. Kamura, T. Sawasaki, Y. Kasahara, K. Takai and Y. Endo, *Bioorg. Med. Chem. Lett.*, 2005, **15**, 5402–5406.
13 A. W. Nguyen and P. S. Daugherty, *Nat. Biotechnol.*, 2005, **23**, 355–360.
14 Because the effects of IVT-sup-tRNAs on the amber-free template were comparable among some examined IVT-sup-tRNAs except for 3′ tuned IVT-sup-tRNAs, the relative suppression efficiency, which is equivalent to the relative fluorescence of YPet translated from the amber template in the presence of IVT-sup-tRNAs, was used for evaluation of IVT-sup-tRNAs during sequence optimization steps.
15 Incidentally, the absolute suppression efficiency of S1 was 13 ± 1%.
16 Because Darm and Tarm are known to interact with each other in L-shaped tertiary structures, these two parts were regarded as one part.
17 The bases (UG) between the 5′ side of Aarm and Darm and the base (C) between Varm and Tarm are identical between S1 and S2.
18 C. Kao, M. Zheng and S. Rüdisser, *RNA*, 1999, **5**, 1268–1272.
19 S. Arends and A. Schön, *Eur. J. Biochem.*, 1997, **244**, 635–645.
20 Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa and T. Ueda, *Nat. Biotechnol.*, 2001, **19**, 751–755.
21 J. Carnes, L. Frolova, S. Zinnen, G. Drugeon, M. Phillippe, J. Justesen, A. L. Haenni, L. Leinwand, L. L. Kisselev and M. Yarus, *RNA*, 2000, **6**, 1468–1479.
22 Incidentally, it is known that N-terminal Met is completely removed in wheat germ extract in the case that the penultimate amino acid is Ala, see: T. Kanno, M. Kitano, R. Kato, A. Omori, Y. Endo and Y. Tozawa, *Protein Expression Purif.*, 2007, **52**, 59–65.
23 Although another two smaller peaks indicated by arrow heads were detected, these peaks were also observed in the positive control. Each of these peaks was 16 smaller than their right main peak. These peaks are therefore likely peptides derived from the main peptides

involved in some reactions in wheat germ extract or during mass measurement.
24 Although rare codons would appear in the 3′ originally-untranslated region of natural mRNAs, it has been reported that there is almost no codon preference in the wheat germ extract used here, see: M. Madono, T. Sawasaki, R. Morishita and Y. Endo, *New Biotechnol.*, 2011, **28**, 211–217.
25 (*a*) N. Choisne, A. Martin-Canadell and I. Small, *Plant J.*, 1997, **11**, 597–604; (*b*) B. Ulmasov and W. Folk, *Plant Cell*, 1995, **7**, 1723–1734; (*c*) N. Choisne, V. T. C. Carneiro, G. Pelletier and I. Small, *Plant Mol. Biol.*, 1997, **36**, 113–123; (*d*) K. Akama and H. Beier, *Nucleic Acids Res.*, 2003, **31**, 1197–1207.
26 (*a*) A. Ogawa, *ChemBioChem*, 2009, **10**, 2465–2468; (*b*) A. Ogawa, *RNA*, 2011, **17**, 478–488; (*c*) A. Ogawa, *ChemBioChem*, 2011, **12**, 881–885.