J. Biochem. 2010;148(2):179–187 doi:10.1093/jb/mvq051

JB THE JOURNAL OF BIOCHEMISTRY

Site-specific incorporation of 4-Iodo-L-phenylalanine through opal suppression

Received April 15, 2010; accepted April 29, 2010; published online May 21, 2010

Koichiro Kodama[1,*,†], Hiroshi Nakayama[2], Kensaku Sakamoto[1,3], Seketsu Fukuzawa[4], Takanori Kigawa[3], Takashi Yabuki[1], Makoto Kitabatake[5], Koji Takio[6] and Shigeyuki Yokoyama[1,3,‡]

[1]Department of Biophysics and Biochemistry, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033; [2]Biomolecular Characterization Team, Advanced Development and Supporting Center, RIKEN Wako Institute, 2-1 Hirosawa, Wako, Saitama 351-0033; [3]RIKEN Systems and Structural Biology Center, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045; [4]Department of Chemistry, School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-0033; [5]Institute for Virus Research, Kyoto University, Kyoto 606-8507; and [6]Highthroughput Factory, RIKEN Harima Institute, 1-1-1 Kouto, Mikazuki, Sayo, Hyogo 679-5148, Japan

*Present address: Koichiro Kodama, Chemical Biology Institute, Waseda University, 3-4-1, Okubo, Shinjuku-ku, Tokyo, 169-8555, Japan.

[†]Koichiro Kodama. email: k.kodama@kurenai.waseda.jp

[‡]Shigeyuki Yokoyama. email: yokoyama@biochem.s.u-tokyo.ac.jp

A variety of unique codons have been employed to expand the genetic code. The use of the opal (UGA) codon is promising, but insufficient information is available about the UGA suppression approach, which facilitates the incorporation of non-natural amino acids through suppression of the UGA codon. In this study, the UGA codon was used to incorporate 4-iodo-L-phenylalanine into position 32 of the Ras protein in an *Escherichia coli* cell-free translation system. The undesired incorporation of tryptophan in response to the UGA codon was completely repressed by the addition of indolmycin. The minor amount (3%) of contaminating 4-bromo-L-phenylalanine in the building block 4-iodo-L-phenylalanine led to the significant incorporation of 4-bromo-L-phenylalanine (21%), and this problem was solved by using a purified 4-iodo-L-phenylalanine sample. Optimization of the incubation time was also important, since the undesired incorporation of free phenylalanine increased during the cell-free translation reaction. The 4-iodo-L-phenylalanine residue can be used for the chemoselective modification of proteins. This method will contribute to advancements in protein engineering studies with non-natural amino acid substitutions.

Keywords: cell-free protein synthesis/indolmycin/non-natural amino acid/opal codon/Transfer RNA.

Abbreviations: A294G-PheRS, the PheRS mutant containing glycine residue at position 294 of the α subunit; brF, 4-bromo-L-phenylalanine residue; brF32-Ras, a mutant Ras protein containing 4-bromo-L-phenylalanine residue at position 32; CAT, chloramphenicol acetyltransferase; ESI, electrospray ionization; F32-Ras, a mutant Ras protein containing a phenylalanine residue at position 32; FRET, fluorescence resonance energy transfer; HPLC, high performance liquid chromatography; iF-tRNA$^{Phe}_{UCA}$, 4-iodo-L-phenylalanyl-tRNA$^{Phe}_{UCA}$; iF32, 4-iodo-L-phenylalanine residue; iF32-Ras, a mutant Ras protein containing 4-iodo-L-phenylalanine residue at position 32; LC-MS, liquid chromatography-mass spectrometry; Lys-C, *Achromobacter* protease I; MALDI-MS, matrix-assisted laser desorption/ionization-mass spectrometry; ODS, octadecyl silyl; PAGE, polyacrylamide gel electrophoresis; RF-2, release factor 2; PheRS, phenylalanyl-tRNA synthetase; pK7-Ras-32op, an expression plasmid that encodes the *ras* gene carrying a UGA mutation at position 32; SDS, sodium dodecyl sulfate; TLC, thin layer chromatography; TrpRS, tryptophanyl-tRNA synthetase; Y32-Ras, the wild-type Ras protein (1-171).

Advances in genetic engineering are now allowing researchers to use non-natural amino acids for protein engineering studies. These amino acids have been incorporated into proteins site-specifically in not only living cells but also in cell-free protein synthesis systems. As compared with the cell-based systems, cell-free translation systems can be easily controlled or modified by the addition of chemical compounds and translation factors (*e.g.* aminoacylated tRNA), and they have facilitated the incorporation of a greater variety of non-natural amino acids into proteins (*1*–*9*). Non-natural amino acids provide a new set of building blocks to produce highly functional proteins with capabilities beyond those of the natural proteins obtained through evolutionary selection using the common 20 amino acids [*e.g.* alloprotein (*10*), artificial protein (*11*)]. Non-natural amino acids have been used for chemoselective modification (*12*, *13*), fluorescent labeling (*14*), *in vitro* selection of mRNA display libraries (*5*), and so on. For instance, the aryl iodide of a 4-iodo-L-phenylalanine residue, a chemoselective tag for palladium-catalysed reactions, was employed for the site-specific functionalization of proteins (*12*, *13*). A donor and acceptor pair for fluorescence resonance energy transfer (FRET) was introduced in response to two 4 base codons, into different positions of calmodulin, and the conformation of the protein was monitored by FRET changes upon binding with a calmodulin-binding peptide (*14*).

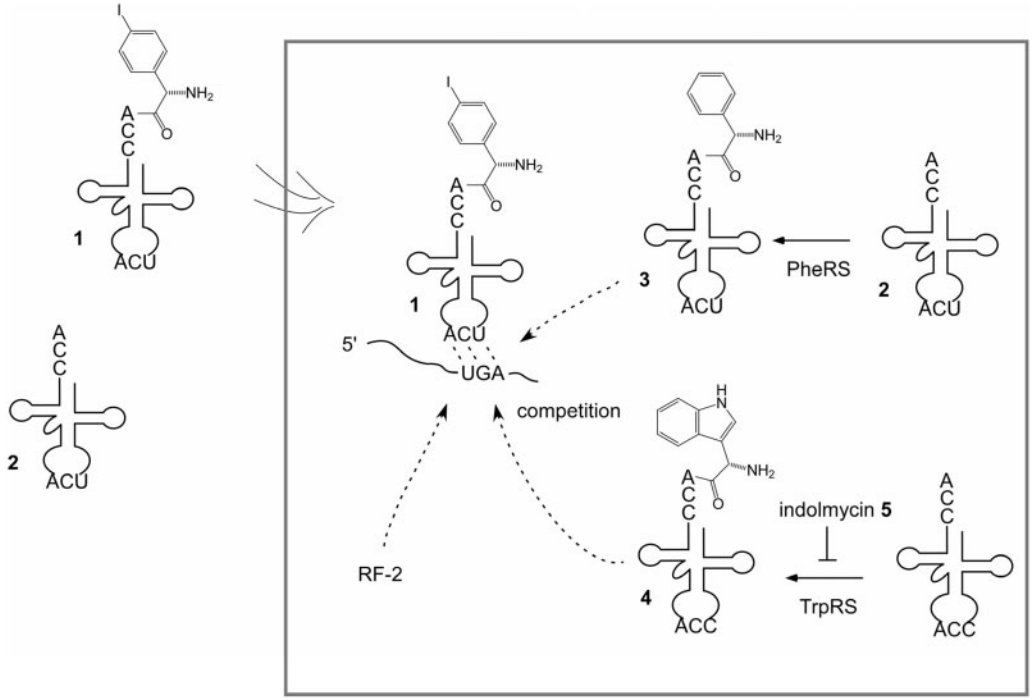

Fig. 1 Scheme for the site-specific incorporation of 4-iodo-L-phenylalanine using pre-acylated tRNA$^{Phe}_{UCA}$ 1 in the cell-free translation system. The arrow indicates that a mixture of **1** and uncharged tRNA$^{Phe}_{UCA}$ **2** was added to the cell-free translation system. **1** competes with phenylalanyl-tRNA$^{Phe}_{UCA}$ **3**, RF-2 and tryptophanyl-tRNA$^{Trp}_{CCA}$ **4**.

Several techniques for pre-acylation, the *in vitro* aminoacylation of tRNAs with non-natural amino acids, have been developed. For instance, non-natural amino acids are recognized by altered-specificity mutants of the aminoacyl-tRNA synthetase (*e.g.* phenylalaninyl-tRNA synthetase, PheRS) for attachment to the tRNA variants (suppressor tRNAs) (*15–17*). In addition, pre-acylation using ribozymes (*8*) and PNA thioesters (*18*) has been reported. Another approach is the chemical acylation of a dinucleotide, followed by its ligation to a truncated tRNA (*1–3*, *5–7*, *19–21*). The aminoacylated suppressor tRNAs are added to a cell-free translation system. These tRNAs suppress a non-sense mutation within the gene of a target protein, which results in translational readthrough by the incorporation of the non-natural amino acids at the target sites.

A variety of unique codons have been used to expand the genetic code, including UAG (*1*, *2*), four base codons (*22*), and unnatural base codons (*16*, *23*). In addition to these unique codons, the opal (UGA) codon may be used for site-specific incorporation (*24*). The UGA suppression approach was applied to incorporate 5-hydroxytryptophan, using an orthogonal tryptophanyl-tRNA synthetase (TrpRS) and tRNA^{Trp} pair in human 293T cells (*25*). However, this is the only report of the UGA suppression approach, and information about the site-specific incorporation of non-natural amino acids is limited. In theory, the UGA suppression approach may be used for multiple incorporations of non-natural amino acids (*21*).

We previously reported the UAG suppression approach, using an *Escherichia coli* PheRS variant (A294G-PheRS) (*26*, *27*) and the UAG suppressor tRNA$^{Phe}_{CUA}$ (*17*). The suppressor tRNA$^{Phe}_{CUA}$ has critical mutations in the anticodon (*28*) and possesses minimal amino acid-accepting activity under physiological conditions in cell-free protein synthesis (*e.g.* 10 mM $MgCl_2$, pH 7.5). We found low-fidelity conditions (*e.g.* 70 mM $MgCl_2$, 2.5 M DMSO, pH 8.1) that enhanced the misacylation of the suppressor tRNA$^{Phe}_{CUA}$. The UAG suppressor tRNA$^{Phe}_{CUA}$ was pre-acylated with 4-iodo-L-phenylalanine under the low-fidelity conditions and was used for cell-free protein synthesis. 4-Iodo-L-phenylalanine was incorporated in response to the UAG codon, which was confirmed by liquid chromatography-mass spectrometry (LC-MS). However, the undesired incorporation of phenylalanine in response to the UAG codon was also observed. This was mediated by tRNA$^{Phe}_{CUA}$ through the following mechanism (*17*). 4-Iodo-L-phenylalanyl-tRNA$^{Phe}_{CUA}$ is gradually deacylated by either an enzymatic or non-enzymatic process during the incubation (*29*), and free tRNA$^{Phe}_{CUA}$ is aminoacylated by the endogenous PheRS in the cell-free translation system. The resulting phenylalanyl-tRNA$^{Phe}_{CUA}$ is re-used for cell-free protein synthesis again. In addition, phenylalanine could be incorporated in response to not only the UAG codon but also the UGA codon (Fig. 1).

In this study, the UGA suppression approach was employed to incorporate 4-iodo-L-phenylalanine into position 32 of the Ras protein, in an *E. coli* cell-free translation system. The UGA codon is reportedly read by tRNA$^{Trp}_{CCA}$ (*30*); the mismatch of the UGA codon with the CCA anticodon leads to the misincorporation of tryptophan (natural readthrough). The undesired incorporation of an amino acid can be repressed by

addition of an inhibitor that is specific to the cognate aminoacyl-tRNA synthetase (*7*). We selected indolmycin, a specific inhibiator of *E. coli* TrpRS (*31*, *32*) (Fig. 1). The undesired incorporation of amino acids in response to the UGA codon was completely repressed by optimizing the reaction conditions.

Materials and Methods

General

4-Iodo- and 4-bromo-L-phenylalanine were purchased from Sigma-Aldrich (St Louis, MO, USA). The PD-10 desalting column and the Resource Q and Superdex 75 resins were obtained from Amersham Pharmacia (GE Healthcare) Bio-Sciences (Piscataway, NJ, USA). The high performance liquid chromatography (HPLC) column (COSMOSIL 5C-18 AR-II, 20 mm × 250 mm) was purchased from Nacalai Tesque (Kyoto, Japan). The matrix-assisted laser desorption/ionization-mass spectrometry (MALDI-MS) analysis was performed by Dr Naoshi Dohmae (RIKEN Wako). The quantitative amino acid analysis was performed with a model 835 Amino Acid Analyzer (Hitachi, Tokyo, Japan) by Masao Chijimatsu (RIKEN Wako). The expression plasmid with the opal mutation at position 32 of the *ras* gene (pK7-Ras-32op) was constructed from the expression plasmid pK7-Ras (*33*). Indolmycin was a gift from Prof. Dieter Söll (Yale University). The S30 extract (*34*) was prepared from *E. coli* strain A19 (*metB*, *rna*) (*29*). The preparation of the tRNA transcript, the cell-free protein synthesis, and the LC-MS analysis were performed as described previously (*17*). The synthesized Ras protein was purified by anion exchange (Resource Q) and size exclusion (Superdex 75) chromatography. The fluorescent imaging scanner FLA-2000 and the proprietary software program (Image Gauge version 3.01) were products of Fuji Photo Film (Tokyo, Japan).

Purification of 4-iodo-L-phenylalanine

4-Iodo-L-phenylalanine (50 mg/ml) was dissolved in acetic acid/water (8 : 2), and was purified via preparative ODS-HPLC using a linear gradient of 0–35% (*v*/*v*) acetonitrile/water over 30 min. 4-Iodo- and 4-bromo-L-phenylalanine were eluted at 27 and 24 min, respectively. Fractions were collected and lyophilized to obtain a pure 4-iodo-L-phenylalanine.

Pre-acylation of tRNA$^{Phe}_{UCA}$

Escherichia coli tRNA$^{Phe}_{UCA}$ (G2G3G4) was incubated at RT for 25 min in the presence of $MgCl_2$ (60 mM), HEPES–KOH (pH 8.1, 40 mM), ATP–KOH (pH 7, 10 mM), DTT (10 mM), Tween-40 [0.25% (*v*/*v*)], dimethylsulfoxide (2.6 M), 4-iodo-L-phenylalanine (5 mg/ml) and the Ala294Gly mutant of *E. coli* phenylalanyl-tRNA synthetase (*26*, *27*), A294G-PheRS (1 μM). The pre-acylated tRNA (iF-tRNA$^{Phe}_{UCA}$) was treated with phenol/chloroform (5 : 1, pH 4.5), desalted on a PD-10 desalting column equilibrated with ammonium acetate buffer (50 mM, pH 5.0), and precipitated with 2-propanol. The resulting pellet was rinsed twice with ethanol/water (7 : 3, 5 ml) and ethanol (1 ml). It was dried with a water–aspirator vacuum, and was stored at −20°C.

Acidic polyacrylamide gel electrophoresis analysis

The aminoacylation mixture was extracted with phenol/chloroform (5 : 1, pH 4.5), and was mixed with an equal volume of the sample buffer containing NaOAc (0.2 M, pH 4.5) and glycerol [4.4 M, 50% (*v*/*v*)]. This sample was separated by polyacrylamide gel electrophoresis (PAGE) under acidic conditions [NaOAc (0.2 M, pH 5.2), urea (7.9 M) and polyacrylamide (9%T, 5%C)] at 4°C in an acidic buffer containing NaOAc (0.1 M, pH 5.2). The gel was stained with GelStar nucleic acid stain. The aminoacylated tRNA was detected with a fluorescent imaging analyzer, model FLA2000, and the aminoacylation level of the tRNA was estimated by using the proprietary software program (Image Gauge version 3.01). The effective digit for the calculated results was 1.

Results

We tried to incorporate 4-iodo-L-phenylalanine in response to the UGA codon at position 32 of the Ras protein, in a similar manner to the site-specific incorporation utilizing the UAG codon (*17*), in order to compare these two suppression approaches. First, the tRNA$^{Phe}_{UCA}$ transcript was prepared by *in vitro* transcription of the tRNA$^{Phe}_{UCA}$ gene, in which the original 5′-terminal sequence (C2C3C4) was changed to increase the transcription yield (Fig. 2). These mutations are not essential for tRNA^{Phe} to be recognized by *E. coli* PheRS (*28*), but the tRNA$^{Phe}_{UCA}$ transcript

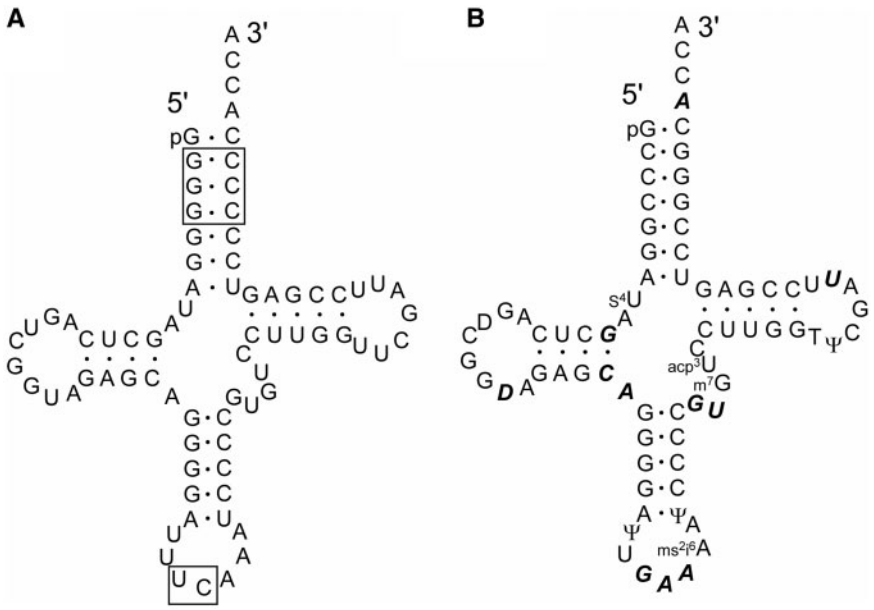

Fig. 2 Comparisions of the tRNAs. (A) *Escherichia coli* opal suppressor tRNA$^{Phe}_{UCA}$ transcript and (B) *E. coli* native tRNA$^{Phe}_{GAA}$. The boxed nucleotides in the suppressor tRNA$^{Phe}_{UCA}$ differ from those in the wild-type tRNA$^{Phe}_{GAA}$ transcript. The nucleotides in italic and bold styles are essential for the recognition of the native tRNA$^{Phe}_{CCA}$ by *E. coli* PheRSs.[4] U, 4-thiouridine; D, dihydrouridine; ψ, pseudouridine; ms^{2}i^{6}A, 2-methylthio-N^6-isopentenyladenosine; m^{7}G, 7-methylguanosine; acp^{3}U, 3-(3-amino-3-carboxypropyl) uridine; T, 5-methyluridine.

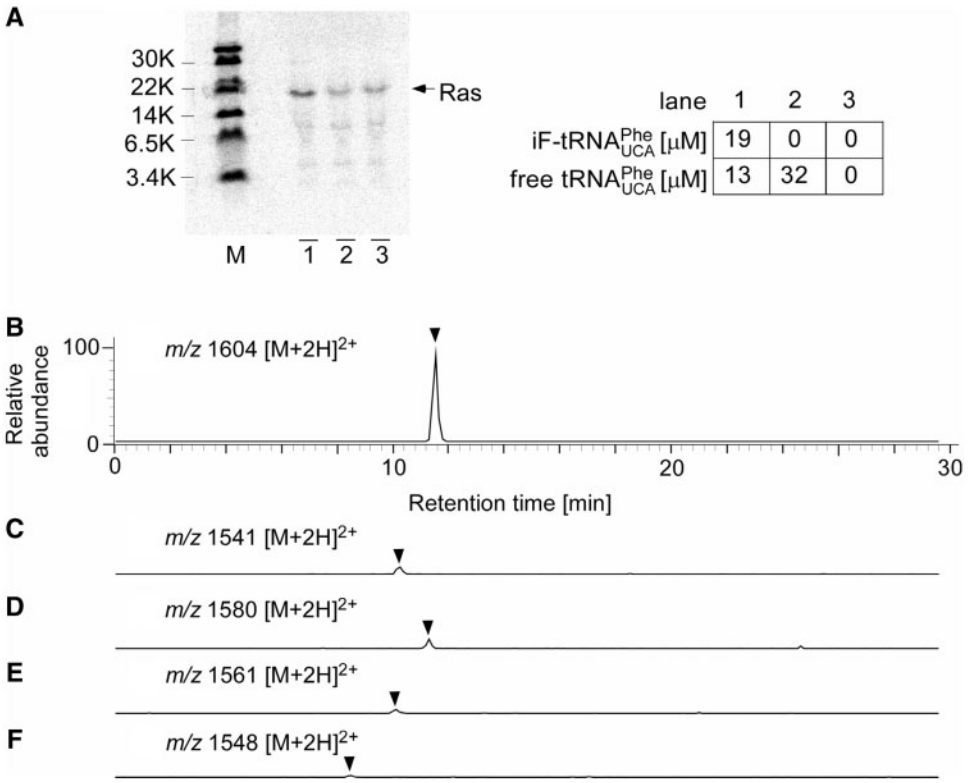

Fig. 3 Analyses of a cell-free translation product. (A) SDS–PAGE analysis of [^{14}C]-leucine-labeled translation products. The reaction mixture was incubated at 37°C for 40 min with the mixture (lane 1) of **1** (19 μM) and **2** (13 μM), **2** (32 μM, lane 2) or without tRNA$^{Phe}_{UCA}$ (lane 3). The arrow indicates the 20 kDa band for the full-length Ras protein. The selected ion chromatograms of *m*/*z* (B) 1604, (C) 1541, (D) 1580, (E) 1561 and (F) 1548 correspond to the K3 peptide (residues 17-42) containing 4-iodo-L-phenylalanine, phenylalanine, 4-bromo-L-phenylalanine, tryptophan and a carboxymethylated cysteine residue, respectively, at position 32.

has critical mutations in the first and second positions of the anticodon (G34U, A35C). Therefore, the transcript was aminoacylated under low-fidelity conditions to increase the misacylation with 4-iodo-L-phenylalanine (*17*), and the results were monitored by acidic PAGE (*35*). The misacylation was enhanced when the $MgCl_2$ concentration was increased to 60 mM. Another important factor affecting the aminoacylation reaction was the pH condition. Misacylation of the tRNA$^{Phe}_{UCA}$ was observed between pH 6.4 and 8.1, but the optimum value was pH 8.1. The ratio of 4-iodo-L-phenylalanyl-tRNA$^{Phe}_{UCA}$ (iF-tRNA$^{Phe}_{UCA}$) **1** to free tRNA$^{Phe}_{UCA}$ **2** was estimated to be 6 : 4 after aminoacylation under the pH 8.1 conditions (Supplementary Fig. S1).

The mixture of **1** and **2** was added to an *E. coli* cell-free translation system with pK7-Ras-32op, an expression plasmid for the *ras* gene containing a UGA codon at position 32. This experiment was performed according to the published procedure for UAG suppression (*17*), with a slight modification. The wild-type Ras protein (residues 1–171, Y32-Ras) has no tryptophan residue (*36*), and therefore all of the proteogenic amino acids, except for tryptophan, were added to the reaction mixture with L-[^{14}C]-leucine. The reaction mixture was incubated at 37°C for 40 min, to synthesize the Ras mutant containing a 4-iodo-L-phenylalanine residue at position 32 (iF32-Ras). The translation product was analyzed by sodium dodecyl sulfate (SDS)–PAGE and autoradiography to confirm the expression of the Ras protein. The 20 kDa band corresponding to the readthrough product, the full-length Ras protein, was detected on the gel (Fig. 3A, lane 1). The amount of the synthesized Ras protein depended on the amount of iF-tRNA$^{Phe}_{UCA}$ **1** that was added to the cell-free translation system. The yield of the readthrough product was 55% relative to Y32-Ras under the optimum conditions (Supplementary Fig. S2). In addition, the possibility of natural readthrough at the UGA codon was investigated. The 3.3 kDa band corresponding to a truncated product, in which translation was terminated at position 32, was scarcely observed on the gel (Fig. 3A, lane 1). A translation product was observed at 20 kDa, even when no pre-acylated tRNA$^{Phe}_{UCA}$ **1** was added to the cell-free translation reaction (Fig. 3A, lanes 2 and 3). These results showed that natural UGA suppression occurred at a significant level in the *E. coli* cell-free translation system.

To identify the amino acid residue at position 32 of the synthesized protein, iF32-Ras was carboxymethylated and subjected to in-gel digestion using *Achromobacter* protease I (Lys-C) (*37*). The K3 peptide, liberated from position 17 through position 42, was analysed by LC-MS, as described earlier (*15*). The abundance of amino acid residues at position 32 was estimated from the ratio of the ion-peak intensities in the selected ion chromatograms, assuming that the

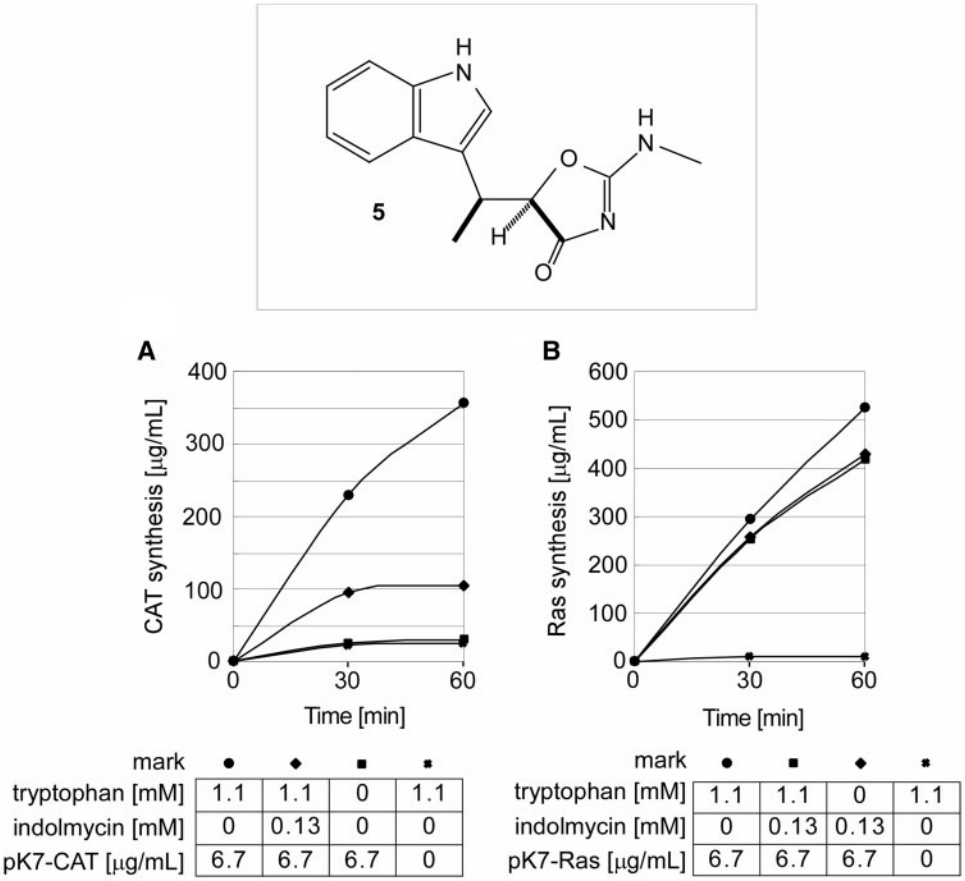

Fig. 4 The effect of indolmycin on the *E. coli* cell-free translation system. Time courses of cell-free protein synthesis of (A) CAT and (B) Y32-Ras in the presence/absence of tryptophan and/or indolmycin. Inset: chemical structure of indolmycin. Abbreviation: pK7-CAT, an expression plasmid containing the *CAT* gene (*52*).

ionization efficiency of each peptide in electrospray ionization (ESI) was identical. The LC-MS analysis revealed that the UGA codon at position 32 was translated as 4-iodo-L-phenylalanine (80%), phenylalanine (7%), 4-bromo-L-phenylalanine (9%) and tryptophan (4%) (Fig. 3B–E), and these results were also supported by the MALDI-MS analysis (Supplementary Fig. S3A). The incorporation of cysteine was negligible (<2%) (Fig. 3F).

The addition of a specific inhibitor of bacterial TrpRS is a strategy to prevent tryptophan incorporation. The effect of indolmycin **5** (*31*, *32*) on the *E. coli* cell-free translation system was examined, as a preliminary experiment. In the presence of **5**, two proteins, chloramphenicol acetyltransferase (CAT), which has four tryptophan residues, and Y32-Ras, were synthesized. The results indicated that the cell-free synthesis of CAT was suppressed by the addition of **5** (0.13 mM) (Fig. 4A), but the cell-free synthesis of Y32-Ras was hardly affected by 5 (Fig. 4B). Next, **5** (0.13 mM) was added to the cell-free translation system to synthesize iF32-Ras, and this mixture was incubated at 37°C for 40 min. The translation product was digested with Lys-C and subjected to an LC-MS analysis, which indicated that no tryptophan was incorporated into position 32 (Fig. 5D). In addition, the amino acid residues at position 32 were identified (Fig. 5A–C and Supplementary Fig. S3B): 4-iodo-L-phenylalanine (71%), phenylalanine (8%) and 4-bromo-L-phenylalanine (21%).

4-Bromo-L-phenylalanine is a 4-halogenated phenylalanine analog, which could be a potential contaminant of the 4-iodo-L-phenylalanine preparation. Therefore, the purity of the 4-iodo-L-phenylalanine was assessed by thin layer chromatography (TLC) and quantitative amino acid analysis (*38*). Although only a single spot was observed on the TLC plate, the quantitative amino acid analysis revealed that the commercially available 4-iodo-L-phenylalanine sample contained a small amount (3%) of 4-bromo-L-phenylalanine. On the other hand, the analysis did not detect the presence of phenylalanine (Supplementary Fig. S4A). The 4-iodo-L-phenylalanine was purified by ODS-HPLC, and the purified sample was subjected to quantitative amino acid analysis once again. After this purification, the absence of 4-bromo-L-phenylalanine was confirmed (Supplementary Fig. S4B).

The purified 4-iodo-L-phenylalanine was added to the cell-free translation reaction, which was performed under the same conditions, except that the incubation time was decreased to 30 min, in order to minimize phenylalanylation of the uncharged UGA suppressor $tRNA^{Phe}_{UCA}$. The translation product was analysed to confirm that no 4-bromo-L-phenylalanine was incorporated into Ras (Supplementary Fig. S3C). The translation product was further subjected to SDS–PAGE,

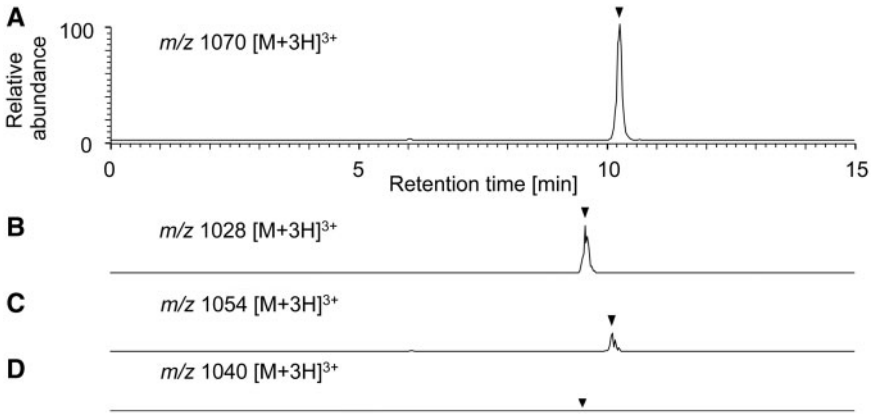

Fig. 5 An LC-MS analysis of a translation product. The selected ion chromatograms of *m*/*z* (A) 1070, (B) 1028, (C) 1054 and (D) 1040 corresponding to the K3 peptide containing 4-iodo-L-phenylalanine, phenylalanine, 4-bromo-L-phenylalanine and tryptophan, respectively, at position 32.

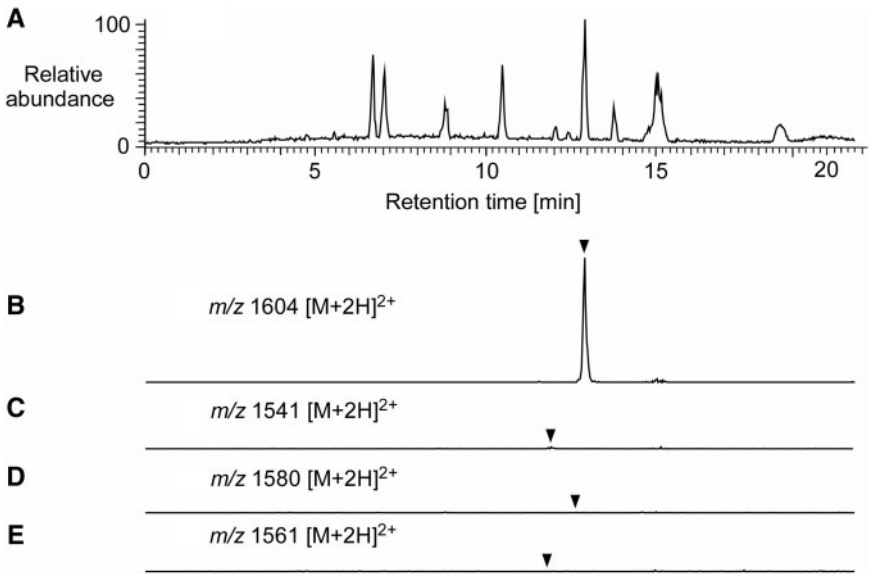

Fig. 6 An LC-MS analysis of a translation product. (A) Total ion chromatogram of Lys-C digested peptides of cell-free translation products. The selected ion chromatograms of *m*/*z* (B) 1604, (C) 1541, (D) 1580 and (E) 1561 correspond to the K3 peptide containing 4-iodo-L-phenylalanine, phenylalanine, 4-bromo-L-phenylalanine and tryptophan, respectively, at position 32.

and the excised protein band was digested with Lys-C. An LC-MS analysis of the K3 peptide indicated that 98% of the amino acid residues at position 32 were 4-iodo-L-phenylalanine, and that the incorporation of phenylalanine and 4-bromo-L-phenylalanine was negligible (<2%, Fig. 6B–D).

Discussion

The UGA codon at position 32 was read by not only the UGA suppressor $tRNA^{Phe}_{UCA}$ but also the endogenous $tRNA^{Trp}_{CCA}$. Even when no UGA suppressor $tRNA^{Phe}_{UCA}$ was added to the *E. coli* cell-free translation system, translational readthrough was observed at the UGA codon (Fig. 3A, lane 3; Supplementary Fig. S2, lane 5). An the LC-MS analysis revealed that the UGA codon was read as tryptophan (4%) (Fig. 3E). This tryptophan was not delivered by the UGA suppressor $tRNA^{Phe}_{UCA}$, since the $tRNA^{Phe}_{UCA}$ lacks some of the important identity nucleotides for aminoacylation by *E. coli* TrpRS under physiological conditions (Supplementary Fig. S5) (*39*). Natural readthrough was not detected when the UAG codon was at the same position (*17*).

The undesired incorporation of tryptophan was effectively prevented by the addition of indolmycin **5** (Fig. 4D). The tryptophan that had been incorporated into position 32 was derived from the *E. coli* S30 extract, since no tryptophan was added to the cell-free translation reaction. The residual tryptophan was attached to its cognate $tRNA^{Trp}_{CCA}$ by the endogenous TrpRS in the cell-free translation system, and this is the reason why the addition of a TrpRS inhibitor was effective. Another strategy for preventing the misincorporation of tryptophan is to knock out the $tRNA^{Trp}$ present in the cell-free translation system with antisense DNA oligonucleotides (*40*), but optimal conditions were not established. In addition, the finding that

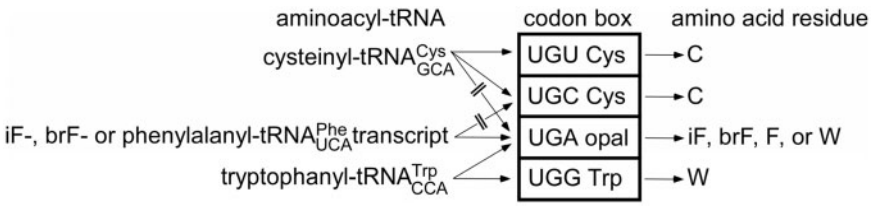

Fig. 7 The results of codon-anticodon recognition in the cell-free system.

no tryptophan was incorporated in the presence of 5 implies that it is possible to incorporate iodo-L-phenylalanine instead of tryptophan. The UGG codon may be read by iF-tRNA$^{Phe}_{CCA}$ but not recognized by release factors, resulting in a higher yield of translation products.

The addition of TrpRS inhibitors is not suitable for synthesizing alloproteins containing tryptophan residues, which seriously limits the scope of the application of our method. Tryptophan is the least abundant amino acid found in proteins, and accounts for only 1.3% of the codons (*41*). Many tryptophan residues can be replaced with another hydrophobic amino acid (*e.g.* tyrosine), without impairing protein function (*42*). However, a tryptophan residue is essential for the stability and activity of some proteins (*43*, *44*). A solution to this drawback is the addition of an engineered tRNA$^{Trp}_{CCA}$, pre-acylated with tryptophan, to the cell-free translation system with **5**. The engineered tRNA$^{Trp}_{CCA}$ may deliver tryptophan in response to the UGG codon. The level of misreading of UGA codons was reportedly extremely low for a suppressor tRNA$^{Trp}_{CCA}$ (Su9) with a purine at position 33 (*45*). More extensive engineering of the anticodon loop and stem may produce a tRNA$^{Trp}_{CCA}$ strictly specific to the UGG codons. For instance, the modification of cytidine to 2′-*O*-methylcytidine (Cm) at position 34 prevents unorthodox pairing with A in the third position of the codon (*46*). The Cm modification would increase the efficiency of reading G-ending codons (*47*).

The tRNA$^{Phe}_{UCA}$ transcript prepared with T7 RNA polymerase lacks post-transcriptional modifications (Fig. 2A). The unmodified U at position 34 could form a wobble base pair (*48*) with the nucleosides at the third position of cysteine codons. In addition, due to the lack of the ms^{2}i^{6}A modification at position 37, the tRNA$^{Phe}_{UCA}$ transcript could cause third letter misreading of the codons (*49*). Therefore, the possibility of misincorporation in response to the UGC/UGU codon was investigated. The synthesized protein was carboxymethylated to detect the cysteine residue by LC-MS. The LC-MS analysis of the K8 peptide (C118 to K147) revealed that only cysteine was incorporated into position 118 (Supplementary Fig. S6), suggesting that the cysteine codon (UGC) was not read by the tRNA$^{Phe}_{UCA}$ transcript. On the other hand, the K4 peptide (Q43 to K88), which contained two cysteine residues at positions 51 (UGU) and 80 (UGC), was not identified on ion chromatograms. However, the mass value of the full-length iF32-Ras protein agreed with that of the calculated mass value (calc. +4 Da.), which suggests that the amino acid residues at positions 51 and 80 are cysteines (Supplementary Fig. S6C). Thus, cysteine codons were not recognized by the tRNA$^{Phe}_{UCA}$ transcript in the cell-free translation system (Fig. 7). In addition, the LC-MS analysis showed that the endogenous tRNA$^{Cys}_{GCA}$ in the cell-free translation system did not recognize the UGA codon, since the incorporation of cysteine into position 32 was negligible (<2%, Fig. 3F).

The commercially available 4-iodo-L-phenylalanine contained 4-bromo-L-phenylalanine (3%, Supplementary Fig. S4A), and the Ras protein synthesized with the unpurified preparation included a significant amount of the 4-bromo-L-phenylalanine residue (brF) (21%) (Fig. 5C). The minor impurity (3%) resulted in the major incorporation of 4-bromo-L-phenylalanine (21%). This phenomenon may have been caused by the difference in the van der Waals radii of iodine (I, 1.98 Å) and bromine (Br, 1.85 Å). 4-Bromo-L-phenylalanine might be preferred over 4-iodo-L-phenylalanine by A294G-PheRS during the pre-acylation step (*26*, *27*). Thus, the purity of the building blocks was shown to be important for preventing the incorporation of undesired non-natural amino acids. In addition, these results suggested that the site-specific incorporation of 4-bromo-L-phenylalanine is possible. 4-Bromo-L-phenylalanine is a useful non-natural amino acid residue, since the heavy-atom bromine within crystallized proteins is used for phase determination in the multi-wavelength anomalous diffraction (MAD) method, a technique for the determination of protein structures in X-ray crystallography (*50*, *51*).

The undesired incorporation of free phenylalanine was mediated by not only tRNA$^{Phe}_{CUA}$ (4%, *17*) but also tRNA$^{Phe}_{UCA}$ (2%). Reducing the incubation time (from 40 to 30 min) effectively decreased the phenylalanine incorporation from 8% (Fig. 5B) to <2% (Fig. 6C). This suggested that the amino acid that was incorporated in response to the UGA codon after 30 min was phenylalanine, which was mainly derived from the free amino acids added to the cell-free translation system (0.5 mM each). Thus, optimization of the incubation time is important for cell-free protein synthesis using not only the UAG suppressor tRNA$^{Phe}_{CUA}$ but also the UGA suppressor tRNA$^{Phe}_{UCA}$.

In conclusion, a non-natural amino acid, 4-iodo-L-phenylalanine, was incorporated into position 32 of the Ras protein through UGA suppression in an *E. coli* cell-free protein synthesis system, as confirmed

by LC-MS. The abundance of the 4-iodo-L-phenylalanine residue at position 32 was improved, from 71% to >98%, by optimizing the reaction conditions. The UGA suppression approach has the potential for the site-specific incorporation of multiple non-natural amino acids, and will contribute to advancements in protein engineering research.

Supplementary Data

Supplementary data are available at *JB* online.

Acknowledgements

We are grateful to Dr Naoshi Dohmae (RIKEN Wako Institute), Prof. Dieter Söll (Yale University), Prof. Kazuyuki Takai (Ehime University) and Prof. Yoichi Nakao (Waseda University) for valuable discussions and suggestions. We also thank Mr Masao Chijimatsu (RIKEN Wako Institute) and Ms Emi Yamaguchi-Nunokawa (RIKEN GSC) for advice and excellent support.

Conflict of interest
None declared.

References

1. 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
2. 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
3. Hohsaka, T., Sato, K., Sisido, M., Takai, K., and Yokoyama, S. (1994) Site-specific incorporation of photofunctional nonnatural amino acids into a polypeptide through in vitro protein biosynthesis. *FEBS Lett.* **344**, 171–174
4. Mamaev, S.V., Laikhter, A.L., Arslan, T., and Hecht, S.M. (1996) Firefly Luciferase: alteration of the color of emitted light resulting from substitutions at position 286. *J. Am. Chem. Soc.* **118**, 7243–7244
5. Li, S., Millward, S., and Roberts, R. (2002) In vitro selection of mRNA display libraries containing an unnatural amino acid. *J. Am. Chem. Soc.* **124**, 9972–9973
6. Ahn, J.H., Kim, N.Y., Kim, T.W., Son, J.M., Kang, T.J., Park, C.G., Choi, C.Y., and Kim, D.M. (2005) Effect of temperature and ATP supply on the efficiency of programmed nonsense suppression. *Anal. Biochem.* **345**, 110–115
7. Sando, S., Kanatani, K., Sato, N., Matsumoto, H., Hohsaka, T., and Aoyama, Y. (2005) A small-molecule-based approach to sense codon-templated natural-unnatural hybrid peptides. Selective silencing and reassignment of the sense codon by orthogonal reacylation stalling at the single-codon level. *J. Am. Chem. Soc.* **127**, 7998–7999
8. 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
9. Hartman, M.C., Josephson, K., Lin, C.W., and Szostak, J.W. (2007) An expanded set of amino acid analogs for the ribosomal translation of unnatural peptides. *PLoS One* **2**, e972
10. Koide, H., Yokoyama, S., Kawai, G., Ha, J.M., Oka, T., Kawai, S., Miyake, T., Fuwa, T., and Miyazawa, T. (1988) Biosynthesis of a protein containing a nonprotein amino acid by *Escherichia coli*: L-2-aminohexanoic acid at position 21 in human epidermal growth factor. *Proc. Natl Acad. Sci. USA* **85**, 6237–6241
11. van Hest, J.C. and Tirrell, D.A. (1998) Efficient introduction of alkene functionality into proteins *in vivo*. *FEBS Lett.* **428**, 68–70
12. Kodama, K., Fukuzawa, S., Nakayama, H., Kigawa, T., Sakamoto, K., Yabuki, T., Matsuda, N., Shirouzu, M., Takio, K., Tachibana, K., and Yokoyama, S. (2006) Regioselective carbon-carbon bond formation in proteins with palladium catalysis; new protein chemistry by organometallic chemistry. *Chem. Bio. Chem.* **7**, 134–139
13. Kodama, K., Fukuzawa, S., Nakayama, H., Sakamoto, K., Kigawa, T., Yabuki, T., Matsuda, N., Shirouzu, M., Takio, K., Yokoyama, S., and Tachibana, K. (2007) Site-specific functionalization of proteins by organopalladium reactions. *Chem. Bio. Chem.* **8**, 232–238
14. Kajihara, D., Abe, R., Iijima, I., Komiyama, C., Sisido, M., and Hohsaka, T. (2006) FRET analysis of protein conformational change through position-specific incorporation of fluorescent amino acids. *Nat. Methods* **3**, 923–929
15. Kiga, D., Sakamoto, K., Kodama, K., Kigawa, T., Matsuda, T., Yabuki, T., Shirouzu, M., Harada, Y., Nakayama, H., Takio, K., Hasegawa, Y., Endo, Y., Hirao, I., and Yokoyama, S. (2002) An engineered *Escherichia coli* tyrosyl-tRNA synthetase for site-specific incorporation of an unnatural amino acid into proteins in eukaryotic translation and its application in a wheat germ cell-free system. *Proc. Natl Acad. Sci. USA* **99**, 9715–9720
16. Hirao, I., Ohtsuki, T., Fujiwara, T., Mitsui, T., Yokogawa, T., Okuni, T., Nakayama, H., Takio, K., Yabuki, T., Kigawa, T., Kodama, K., Yokogawa, T., Nishikawa, K., and Yokoyama, S. (2002) An unnatural base pair for incorporating amino acid analogs into proteins. *Nat. Biotechnol.* **20**, 177–182
17. Kodama, K., Fukuzawa, S., Sakamoto, K., Nakayama, H., Kigawa, T., Yabuki, T., Matsuda, N., Shirouzu, M., Takio, K., Tachibana, K., and Yokoyama, S. (2006) A new protein engineering approach combining chemistry and biology, part I; site-specific incorporation of 4-iodo-L-phenylalanine *in vitro* by using misacylated suppressor $tRNA^{Phe}$. *Chem. Bio. Chem.* **7**, 1577–1581
18. Ninomiya, K., Minohata, T., Nishimura, M., and Sisido, M. (2004) *In situ* chemical aminoacylation with amino acid thioesters linked to a peptide nucleic acid. *J. Am. Chem. Soc.* **126**, 15984–15989
19. Heckler, T.G., Chang, L.H., Zama, Y., Naka, T., Chorghade, M.S., and Hecht, S.M. (1984) T4 RNA ligase mediated preparation of novel "chemically misacylated" $tRNA^{Phe}$s. *Biochemistry* **23**, 1468–1473
20. Short, G.F. 3rd, Golovine, S.Y., and Hecht, S.M. (1999) Effects of release factor 1 on *in vitro* protein translation and the elaboration of proteins containing unnatural amino acids. *Biochemistry* **38**, 8808–8819
21. Ohtsuki, T., Manabe, T., and Sisido, M. (2005) Multiple incorporation of non-natural amino acids into a single protein using tRNAs with non-standard structures. *FEBS Lett.* **579**, 6769–6774
22. Hohsaka, T., Ashizuka, Y., Sasaki, H., Murakami, H., and Sisido, M. (1999) Incorporation of two different nonnatural amino acids independently into a single

protein through extension of the genetic code. *J. Am. Chem. Soc.* **121**, 12194–12195
23. Bain, J.D., Switzer, C., Chamberlin, A.R., and Benner, S.A. (1992) Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code. *Nature* **356**, 537–539
24. Köhrer, C., Sullivan, E.L., and RajBhandary, U.L. (2004) Complete set of orthogonal 21st aminoacyl-tRNA synthetase-amber, ochre and opal suppressor tRNA pairs: concomitant suppression of three different termination codons in an mRNA in mammalian cells. *Nucleic Acids Res.* **32**, 6200–6211
25. Zhang, Z., Alfonta, L., Tian, F., Bursulaya, B., Uryu, S., King, D.S., and Schultz, P.G. (2004) Selective incorporation of 5-hydroxytryptophan into proteins in mammalian cells. *Proc. Natl Acad. Sci. USA* **101**, 8882–8887
26. Kast, P. and Hennecke, H. (1991) Amino acid substrate specificity of *Escherichia coli* phenylalanyl-tRNA synthetase altered by distinct mutations. *J. Mol. Biol.* **222**, 99–124
27. Ibba, M., Kast, P., and Hennecke, H. (1994) Substrate specificity is determined by amino acid binding pocket size in *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry* **33**, 7107–7112
28. Peterson, E.T. and Uhlenbeck, O.C. (1992) Determination of recognition nucleotides for *Escherichia coli* phenylalanyl-tRNA synthetase. *Biochemistry* **31**, 10380–10389
29. Yabuki, T., Kigawa, T., Dohmae, N., Takio, K., Terada, T., Ito, Y., Laue, E.D., Cooper, J.A., Kainosho, M., and Yokoyama, S. (1998) Dual amino acid-selective and site-directed stable-isotope labeling of the human c-Ha-Ras protein by cell-free synthesis. *J. Biomol. NMR* **11**, 295–306
30. Geller, A.I. and Rich, A. (1980) A UGA termination suppression $tRNA^{Trp}$ active in rabbit reticulocytes. *Nature* **283**, 41–46
31. Werner, R.G., Thorpe, L.F., Reuter, W., and Nierhaus, K.H. (1976) Indolmycin inhibits prokaryotic tryptophanyl-tRNA ligase. *Eur. J. Biochem.* **68**, 1–3
32. Kitabatake, M., Ali, K., Demain, A., Sakamoto, K., Yokoyama, S., and Söll, D. (2002) Indolmycin resistance of *Streptomyces coelicolor* A3(2) by induced expression of one of its two tryptophanyl-tRNA synthetases. *J. Biol. Chem.* **277**, 23882–23887
33. Kigawa, T., Muto, Y., and Yokoyama, S. (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. *J. Biomol. NMR* **6**, 129–134
34. Pratt, J.M. (1984) Coupled transcription-translation in prokaryotic cell-free systems. *Transcription and Translation: A Practical Approach*. (Hames, B.D. and Higgins, S.J., eds.), pp. 179–209, Oxford University Press, USA.
35. Varshney, U., Lee, C.P., and RajBhandary, U.L. (1991) Direct analysis of aminoacylation levels of tRNAs in vivo. Application to studying recognition of Escherichia coli initiator tRNA mutants by glutaminyl-tRNA synthetase. *J. Biol. Chem.* **266**, 24712–24718
36. Ha, J.M., Ito, Y., Kawai, G., Miyazawa, T., Miura, K., Ohtsuka, E., Noguchi, S., Nishimura, S., and Yokoyama, S. (1989) Conformation of guanosine 5′-diphosphate as bound to a human c-Ha-Ras mutant protein: a nuclear Overhauser effect study. *Biochemistry* **28**, 8411–8416
37. Masaki, T., Tanabe, M., Nakamura, K., and Soejima, M. (1981) Studies on a new proteolytic enzyme from *A. chromobacter lyticus* M497–1. I. Purification and some enzymatic properties. *Biochim. Biophys. Acta.* **660**, 44–50
38. Moore, S., Spackman, D.H., and Stein, W.H. (1958) Chromatography of amino acids on sulfonated polystyrene resins. *Anal. Chem.* **30**, 1185–1190
39. Himeno, H., Hasegawa, T., Asahara, H., Tamura, K., and Shimizu, M. (1991) Identity determinants of E. coli tryptophan tRNA. *Nucleic Acids Res.* **19**, 6379–6382
40. Kanda, T., Takai, K., Yokoyama, S., and Takaku, H. (1998) Knocking out a specific tRNA species within unfractionated *Escherichia coli* tRNA by using antisense (complementary) oligodeoxyribonucleotides. *FEBS Lett.* **440**, 273–276
41. Jukes, T.H., Holmquist, R., and Moise, H. (1975) Amino acid composition of proteins: selection against the genetic code. *Science* **189**, 50–51
42. Akanuma, S., Kigawa, T., and Yokoyama, S. (2002) Combinatorial mutagenesis to restrict amino acid usage in an enzyme to a reduced set. *Proc. Natl Acad. Sci. USA* **99**, 13549–13553
43. Sierks, M.R., Ford, C., Reilly, P.J., and Svensson, B. (1989) Site-directed mutagenesis at the active site Trp120 of *Aspergillus awamori* glucoamylase. *Protein Eng.* **2**, 621–625
44. Campanini, B., Raboni, S., Vaccari, S., Zhang, L., Cook, P.F., Hazlett, T.L., Mozzarelli, A., and Bettati, S. (2003) Surface-exposed tryptophan residues are essential for *O*-acetylserine sulfhydrylase structure, function, and stability. *J. Biol. Chem.* **278**, 37511–37519
45. Goldman-Levi, R., Kopelowitz, J., and Engelberg-Kulka, H. (1994) The nature of the nucleotide at the 5′ side of the tRNA Su9 anticodon affects UGA suppression in *Escherichia coli*. *DNA Res.* **1**, 123–127
46. Horie, N., Yamaizumi, Z., Kuchino, Y., Takai, K., Goldman, E., Miyazawa, T., Nishimura, S., and Yokoyama, S. (1999) Modified nucleosides in the first positions of the anticodons of $tRNA^{Leu}_4$ and $tRNA^{Leu}_5$ from *Escherichia coli*. *Biochemistry* **38**, 207–217
47. Satoh, A., Takai, K., Ouchi, R., Yokoyama, S., and Takaku, H. (2000) Effects of anticodon 2′-*O*-methylations on tRNA codon recognition in an *Escherichia coli* cell-free translation. *RNA* **6**, 680–686
48. Takai, K. and Yokoyama, S. (2003) Roles of 5-substituents of tRNA wobble uridines in the recognition of purine-ending codons. *Nucleic Acids Res.* **31**, 6383–6391
49. Bouadloun, F., Srichaiyo, T., Isaksson, L.A., and Bjork, G.R. (1986) Influence of modification next to the anticodon in tRNA on codon context sensitivity of translational suppression and accuracy. *J. Bacteriol.* **166**, 1022–1027
50. Xie, J., Wang, L., Wu, N., Brock, A., Spraggon, G., and Schultz, P.G. (2004) The site-specific incorporation of *p*-iodo-L-phenylalanine into proteins for structure determination. *Nat. Biotechnol.* **22**, 1297–1301
51. Sakamoto, K., Murayama, K., Oki, K., Iraha, F., Kato-Murayama, M., Takahashi, M., Ohtake, K., Kobayashi, T., Kuramitsu, S., Shirouzu, M., and Yokoyama, S. (2009) Genetic encoding of 3-iodo-L-tyrosine in *Escherichia coli* for single-wavelength anomalous dispersion phasing in protein crystallography. *Structure* **17**, 335–344
52. Kim, D.M., Kigawa, T., Choi, C.Y., and Yokoyama, S. (1996) A highly efficient cell-free protein synthesis system from *Escherichia coli*. *Eur. J. Biochem.* **239**, 881–886