An Amino Acid-Swapped Genetic Code

Tomoshige Fujino, Masahiro Tozaki, and Hiroshi Murakami*

KEYWORDS: Genetic code, Cell-free translation, tRNA, Aminoacyl-tRNA synthetase, Genetic firewall

ABSTRACT: Preventing the escape of hazardous genes from genetically modified organisms (GMOs) into the environment is one of the most important issues in biotechnology research. Various strategies were developed to create "genetic firewalls" that prevent the leakage of GMOs; however, they were not specially designed to prevent the escape of genes. To address this issue, we developed amino acid (AA)-swapped genetic codes orthogonal to the standard genetic code, namely SL (Ser and Leu were swapped) and SLA genetic codes (Ser, Leu, and Ala were swapped). From mRNAs encoded by the AA-swapped genetic codes, functional proteins were only synthesized in translation systems featuring the corresponding genetic codes. These results clearly demonstrated the orthogonality of the AA-swapped genetic codes against the standard genetic code and their potential to function as "genetic firewalls for gene". Furthermore, we propose "a codon-bypass strategy" to develop a GMO with an AA-swapped genetic code.

Preventing the escape of hazardous genes from genetically modified organisms (GMOs) into the environment is one of the most important issues in biotechnology research[1-5]. Many groups have reported the development of "genetic firewalls for GMOs," such as biocontainment strategies utilizing auxotrophy[6-8] and kill switches[9-11]. These strategies can prevent GMOs from escaping into the environment by blocking their survival in the absence of unnatural compounds. However, because these strategies were not designed as "genetic firewalls for genes," their ability to prevent the escape of genes from GMOs is limited (Figure 1A). For example, even if GMOs were killed by biocontainment strategies, their genes

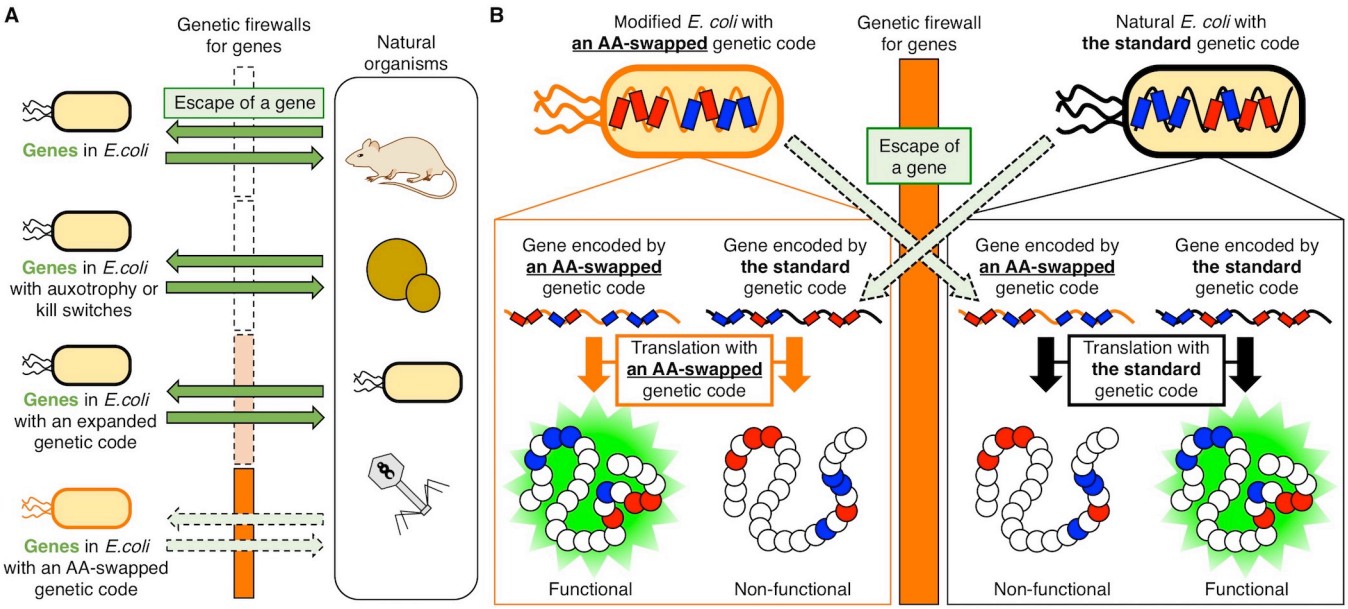

Figure 1. A genetic firewall for genes and the concept of amino acid (AA)-swapped genetic codes proposed in this study. (A) Comparison of genetic firewalls for genes in *Escherichia coli* (*E. coli*) strains with auxotrophy or kill switches, the expanded genetic code, or the AA-swapped genetic code. (B) The concept of AA-swapped genetic codes. Modified *E. coli* cells with an AA-swapped genetic code have AA-swapped genes in which some codons are swapped with each other. A functional protein is expressed from an AA-swapped gene because of reswapped decoding in the modified cells. When gene escape occurs, an AA-swapped gene is translated into a nonfunctional protein because no reswapped decoding occurs in natural *E. coli* cells; thus, the protein would have a large number of AA substitutions. In modified *E. coli* cells, natural genes would also be translated into nonfunctional proteins because of reswapped decoding. This feature would make the modified *E. coli* cells resistant against viruses.

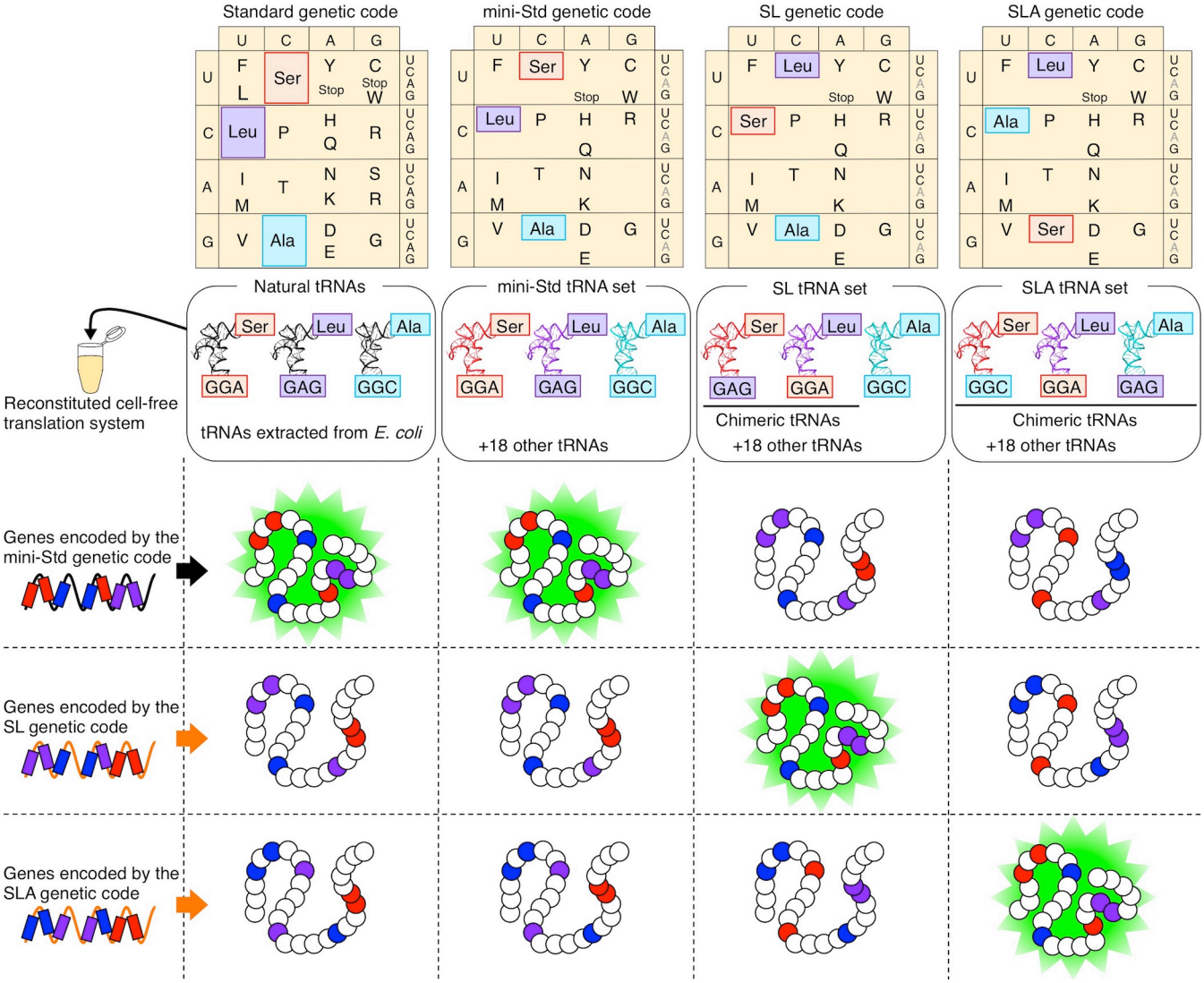

Figure 2. Schematic illustration of translation of genes encoded by the mini-Std, SL, or SLA genetic codes in cell-free translation systems with the standard, mini-Std, SL, and SLA genetic codes. The mini-Std genetic code is analogous to the standard genetic code. The SL and SLA genetic codes have swapped Ser/Leu codons and swapped Ser/Leu/Ala codons, respectively. One of the four tRNA sets were used to construct a translation system with each genetic code: the standard genetic code, tRNAs extracted from *E. coli* (natural tRNAs); the mini-Std genetic code, a mixture of 21 different T7 RNA polymerase-transcribed tRNAs (mini-Std tRNA set); the SL genetic code, the SL tRNA set constructed by replacing $tRNA^{Ser}_{GGA}$ and $tRNA^{Leu}_{GAG}$ with two chimeric tRNAs ($tRNA^{Ser}_{GAG}$ and $tRNA^{Leu}_{GGA}$) in the mini-Std tRNA set; and the SLA genetic code, the SLA tRNA set constructed by replacing $tRNA^{Ser}_{GGA}$, $tRNA^{Leu}_{GAG}$, and $tRNA^{Ala}_{GGC}$ with three chimeric tRNAs ($tRNA^{Ser}_{GGC}$, $tRNA^{Leu}_{GGA}$, and $tRNA^{Ala}_{GAG}$) in the mini-Std tRNA set. Model genes encoded by mini-Std, SL, and SLA genetic codes were prepared and were translated in each translation system. Functional proteins would be synthesized only in the reaction mixture with correct combinations of genes and tRNA sets.

would be released into the environment, and some of them could be transferred into natural organisms through horizontal gene transfer[12-14]. Recently, an expanded genetic code utilizing an unnatural amino acid-assigned UAG codon was also used to develop a genetic firewall for GMOs[15-18]. This strategy only prevents the escape of UAG-containing genes, which would produce truncated proteins in natural organisms. To date, no "genetic firewall for genes" is presently available for preventing the escape of most genes into the environment.

In this study, we proposed a new strategy for "a genetic firewall for genes" by utilizing an amino acid (AA)-swapped genetic code as an orthogonal genetic code against the standard genetic code. In an AA-swapped genetic code, some AAs in the standard genetic code are swapped with each other (Figure 1B). Functional proteins are produced from genes encoded by an AA-swapped genetic code because of the reswapped decoding of the swapped AA codons in modified organisms. Even if a gene escapes from modified organisms, only nonfunctional proteins with many AA substitutions would be produced because no reswapped decoding occurs in natural organisms. In principle, this would ensure the orthogonality of an AA-swapped genetic code; thus, it would work as a "genetic firewall for genes."

To present the proof of concept, we demonstrated the

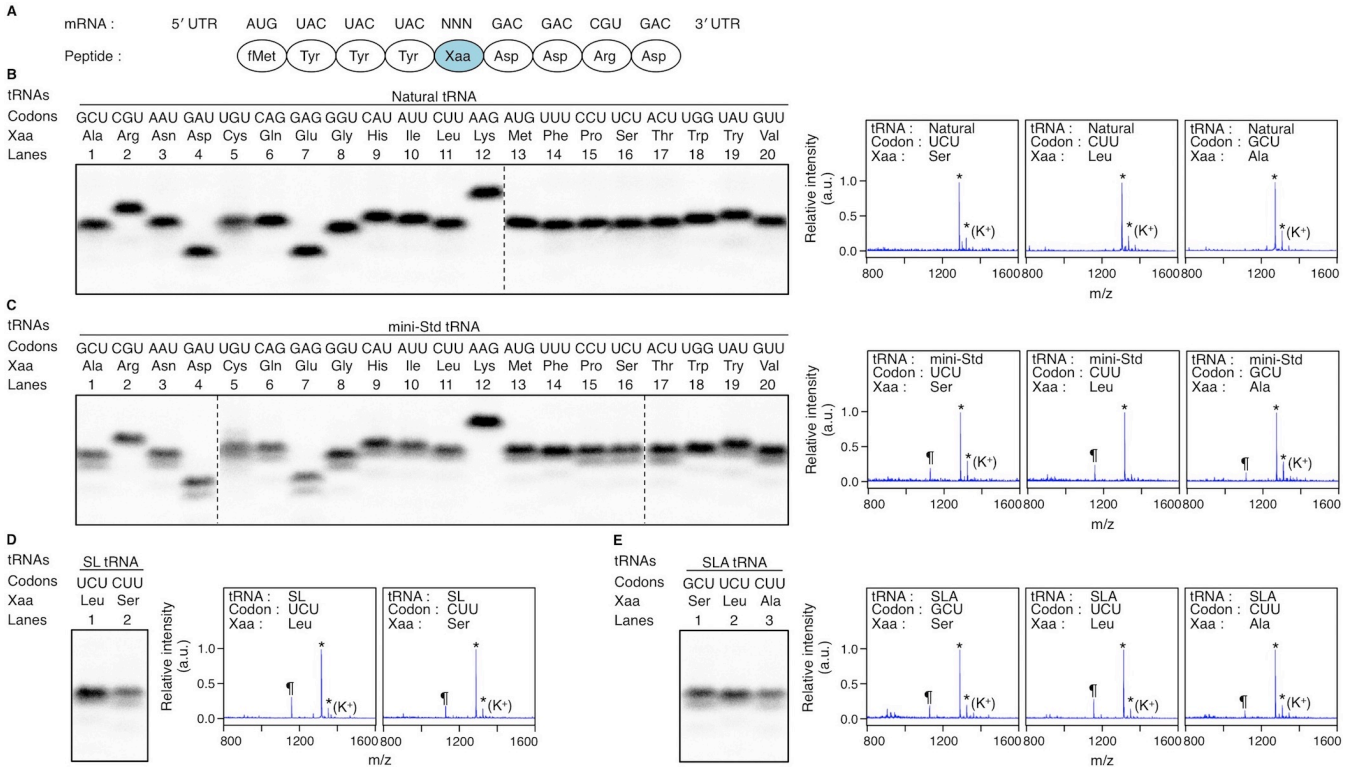

Figure 3. Tricine SDS-PAGE and MALDI-TOF-MS analyses of synthesized model peptides in cell-free translation systems containing natural tRNAs or T7-tRNA sets. (A) The model peptides and the corresponding mRNAs used in this study. The NNN and Xaa represent 20 codons and the corresponding 20 natural AAs, respectively. (B) Tricine SDS-PAGE and MALDI-TOF-MS analyses of model peptides synthesized in the cell-free translation system with natural tRNAs; (C) the mini-Std tRNA set; (D) the SL tRNA set and (E) the SLA tRNA set . For tricine SDS-PAGE analysis, [^{14}C]-Asp was added to the reaction mixture to label synthesized peptides. The peaks of MS spectra are labeled as follows: full-length model peptides containing a desired AA in the position Xaa (*), and truncated peptides lacking fMet (Tyr-Tyr-Tyr-Xaa-Asp-Asp-Arg-Asp) (¶). The calculated and observed masses of the peptides are shown in Table S2. MS spectra of other model peptides are shown in Figure S2.

orthogonality of an AA-swapped genetic code against the standard genetic code by constructing cell-free translation systems with two different AA-swapped genetic codes: the SL genetic code, in which Ser and Leu residues are swapped with each other; and the SLA genetic code, in which the Ser, Leu, and Ala residues are swapped with each other (Figure 2). Considering the application of an AA-swapped genetic code to *in vivo* studies in the future, we used an anticodon swapping strategy (*vide infra*) and avoided methods available only in *in vitro* studies, such as genetic code reprogramming utilizing the chemoenzymatic[19] or flexizyme-catalyzed[20] aminoacylation reaction.

Escherichia coli (*E. coli*) aminoacyl-tRNA synthetases (aaRSs) for Ser, Leu, and Ala attach the AAs to the corresponding tRNAs without recognizing the anticodons[21-24]. This feature would enable AA-swapping by swapping the anticodons of tRNAs because the aminoacylation reaction proceeds according to the body sequence of the tRNA while decoding occurs according to its anticodon. Such chimeric tRNAs have been used in fundamental studies[25-29], and they led to an idea partially similar to ours ("a codon-swapped organism") in which Arg or Ser codons were overwritten by Leu codons[30, 31]. Yeasts *Candida cylindracea* and *Pachysolen tannophilus* possess $tRNA^{Ser}$ or $tRNA^{Ala}$ with mutated-CAG anticodons, thus decoding standard-Leu CUG as Ser or Ala[32-35]. This reassignment of amino acids occurred in a single direction, and other standard Leu (CUH, UUR) and Arg (CGN) codons still exist. A mRNA without standard Arg (AGR) or Leu (CUG) codons will be translated to produce the same protein sequence with both standard and the above alternative genetic codes.

Instead of reassigning an amino acid in a single direction, we designed AA-swapped genetic codes where two or three AAs are swapped with each other (Figure 2). All standard-Ser/Leu codons in the SL genetic code are overwritten, or all standard-Ser/Leu/Ala codons are overwritten in the SLA genetic code. An mRNA with these codons will be translated to produce different protein sequences between the standard genetic code and AA-swapped genetic codes. We also proposed "a codon-bypass strategy" to construct an organism with an AA-swapped genetic code by avoiding a transition genetic code with codons assigned to two AAs that are, thus, lethal.

RESULTS and DISCUSSION

Development of cell-free translation systems with an AA-swapped genetic code. To construct a translation system with an AA-swapped genetic code, we prepared a reconstituted cell-free translation system[36-38] with T7 RNA polymerase-transcribed tRNA (T7-tRNA) sets (Table S1). We used three tRNA sets (Figure 2): a mini-standard (mini-Std) tRNA set (21

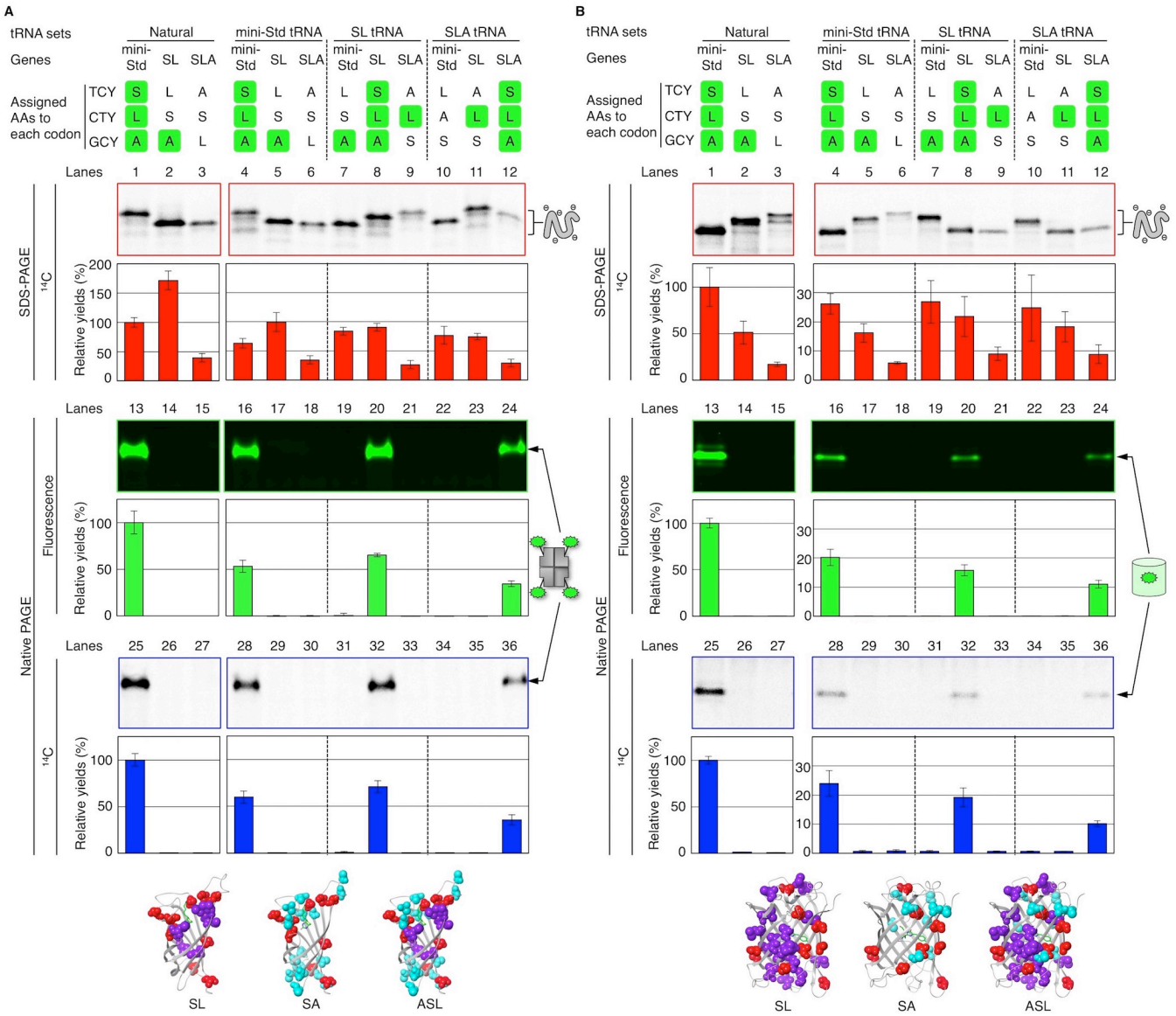

Figure 4. Orthogonality of AA-swapped genetic codes against the standard genetic code confirmed through synthesis of two model proteins. (A) Analysis of proteins from streptavidin mRNA encoded by the mini-Std, SL, or SLA genetic codes in four types of cell-free translation systems constructed with natural tRNAs, the mini-Std tRNA set, the SL tRNA set, or the SLA tRNA set. The upper panel shows the results of SDS-PAGE and subsequent autoradiography of synthesized proteins labeled with [^{14}C]-Asp. The middle panel shows the results of native PAGE and subsequent fluorescence imaging of synthesized proteins. Fluorescently labeled biotin was added to the reaction mixture before native PAGE to visualize the active fraction of streptavidin. The bottom panel shows the results of autoradiography of the dried gel used for the aforementioned fluorescence imaging. The relative band intensities were normalized to the intensity observed for the combination of native tRNAs and mini-Std mRNA and shown under each gel. Error bars show the standard deviation of each experiment in triplicates. Whole images of the gels are shown in Figure S4A. (B) Analysis of proteins from GFP mRNA encoded by the mini-Std, SL, or SLA genetic codes in four types of cell-free translation systems, each constructed with natural tRNAs, the mini-Std tRNA set, the SL tRNA set, or the SLA tRNA set. The analyses were performed as indicated previously, except without the addition of fluorescently labeled biotin. Whole image of the gels are shown in Figure S4B. Ser (red), Leu (purple), and Ala (cyan) residues are shown on the structures of a streptavidin monomer (PDB entry 1STP) and a GFP monomer (PDB entry 1B9C) at the bottom of the Figure.

T7-tRNAs corresponding to the 20 AAs and f-Met), an SL tRNA set (tRNA^{Ser} and tRNA^{Leu} in the mini-Std tRNA set were replaced with two chimeric tRNAs, namely tRNA$^{Ser}_{GAG}$ and tRNA$^{Leu}_{GGA}$), and an SLA tRNA set (tRNA^{Ala}, tRNA^{Ser}, and tRNA^{Leu} in the mini-Std tRNA set were replaced with three chimeric tRNAs, namely tRNA$^{Ser}_{GGC}$, tRNA$^{Leu}_{GGA}$, and tRNA$^{Ala}_{GAG}$). The entire anticodon loops in the chimeric tRNAs were swapped because the nucleotides adjacent to an anticodon could be important for maintaining their decoding fidelity[25, 26]. Other mutations were also introduced into the RNAs to increase aminoacylation efficiency (Table S1). Because some T7-tRNAs have lower activities than the native tRNAs[39-41], we further increased the concentrations of the aaRSs and EF-Tu to promote the aminoacylation of tRNAs and recruitment of aminoacyl-tRNAs to the ribosome, respectively (Figure S1).

Peptide synthesis in cell-free translation systems carrying T7-tRNAs. To confirm genetic code-dependent peptide synthesis, we first tested the translation of model peptides from mRNAs (MYYY-X-DDRD; X represents one of the 20 AAs; Figure 3A). The peptide sequence contains three Asp residues after X; thus, only full-length peptides could be detected via tricine-SDS-PAGE and subsequent autoradiography (^{14}C-Asp). As we expected, the produced peptides from each mRNA were identical between the translation systems carrying natural tRNAs and the mini-Std T7-tRNA set (Figure 3B and C). MALDI-TOF-MS analysis of the reaction mixture also revealed that the corresponding AAs were incorporated in each model peptide (Figure 3B and C, Figure S2, and Table S2). According to the MALDI-TOF-MS spectra, a small proportion of peptides synthesized using the mini-Std T7-tRNA set was missing the *N*-terminal fMet, and these species could represent the lower mobility band observed via tricine-SDS-PAGE (Figure 3C). This might be caused by inefficient first peptide bond formation due to unmodified $tRNA^{ini}$/$tRNA^{Tyr}$, thus leading to unusual translocation prior to the peptidyl transfer reaction. More importantly, when we used the SL or SLA T7-tRNA sets, the corresponding SL- or SLA-swapped peptides were synthesized without contamination by the original AAs (Figure 3D and E and Table S2). These data confirmed genetic code-dependent peptide synthesis in the translation system carrying the different T7-tRNA sets.

Orthogonality of AA-swapped genetic codes against the standard genetic code. The most useful feature of the AA-swapped genetic codes is their orthogonality against the standard genetic code. To demonstrate this principle, we studied the synthesis of the model functional proteins streptavidin and green fluorescent protein (GFP) in the translation systems. We prepared three types of mRNA, which were encoded by the mini-Std genetic code (corresponding to both the standard and the mini-Std genetic code), the SL genetic code, or the SLA genetic code, for each protein (Figure S3A and B). The mRNAs were added to one of the four translation systems constructed with the natural, mini-Std, SL, or SLA tRNA sets. SDS-PAGE and subsequent autoradiography illustrated that proteins with the same mobility were only synthesized in reaction mixtures with the correct combinations of mRNAs and tRNA sets (Figure 4 and Figure S4, lanes 1, 4, 8, and 12; A for streptavidin, B for GFP). The mobilities of the proteins synthesized in translation systems with incorrect combinations differed from those for proteins with the correct combinations. These mobility shifts were likely caused by the large number of resulting AA substitutions in the proteins because many Ala, Ser, and Leu residues are present in the streptavidin (14 Ser, 8 Leu, and 25 Ala residues) and GFP (12 Ser, 19 Leu, and 9 Ala residues) sequences. The mobility shifts appear to correlate with the number of Leu residues rather than the molecular weights of the proteins because the branched hydrophobic side chain of Leu interacts with SDS favorably[42], thereby increasing the negative charge on the proteins and the mobility of the proteins (Figure S3C).

These large numbers of AA substitutions should result in the loss of protein function, thereby guaranteeing the orthogonality of the SL and SLA genetic codes against the standard genetic code. To test this, we assessed the formation of the GFP fluorophore in the synthesized protein using native PAGE and subsequent fluorescence analysis. To analyze the biotin-binding activity of streptavidin, we added fluorescently labeled biotin into the reaction mixture prior to native PAGE. The results illustrated that active proteins were only synthesized in reaction mixtures with the correct combinations of mRNAs and tRNA sets (Figure 4 and Figure S4, lanes 13, 16, 20, and 24; A for streptavidin, B for GFP). Conversely, no active proteins were synthesized in the translation systems using the incorrect combinations of mRNAs and tRNA sets. These results clearly demonstrated the orthogonality of each genetic code against other genetic codes. In particular, the orthogonality of the SL and SLA genetic codes against the standard genetic code (Figure 4 and Figure S4, lanes 14 and 15) provides a new strategy to prevent the escape of genes into the environment because genes using these codes would, in most cases, not produce functional proteins when decoded by the standard genetic code. For example, using all of the protein sequences from *E. coli* MDS42 (a total of 3547 proteins), we calculated the fraction of Ser and Leu residues in each protein (Figure S5A). Although it is unclear what degree of SL swapping is sufficient to induce protein dysfunction, we predict that a large fraction of *E. coli* proteins would lose their functions, as the fraction of Ser and Leu residues in streptavidin was 13.2%, which ranked 2885th among the 3547 proteins, and the fraction for GFP was 12.4%, which is similar to that of streptavidin. The median fraction of AA replacement can be further increased from 16.2 to 25.9% by including Ala residues (Figure S5B), likely increasing the range of proteins that would lose their function following AA-swapping**.**

Effects of AA-swapped genetic codes on the folding and activities of model proteins. We also quantified the fraction of folded proteins using native PAGE and subsequent autoradiography (Figure 4 and Figure S4). Misfolded proteins aggregated and remained in the wells, or they had slower mobilities than the corresponding folded proteins. All of the proteins synthesized in the translation systems with incorrect combinations of mRNAs and tRNA sets were misfolded because of the large number of AA substitutions. Interestingly, the ratios of folded proteins to the total yields of protein synthesized in the translation systems with different correct combinations of mRNAs and tRNA sets were similar (Figure S6A). This indicated that tRNA modification had no observable influence on the folding of streptavidin or GFP; however, it did decrease protein production by half (for streptavidin) or three-quarters (for GFP) compared with that of the natural tRNAs. This result also illustrated that there was no observable influence of the genetic code on the folding of streptavidin or GFP; however, the observed effect on protein production might result from the local structural difference in the mRNA. When AA-swapping is applied to a whole genome in the future, an appropriate codon choice would be required to promote folding of some proteins as synonymous codon substitution sometimes lowered the efficiency of cotranslational protein folding[43, 44].

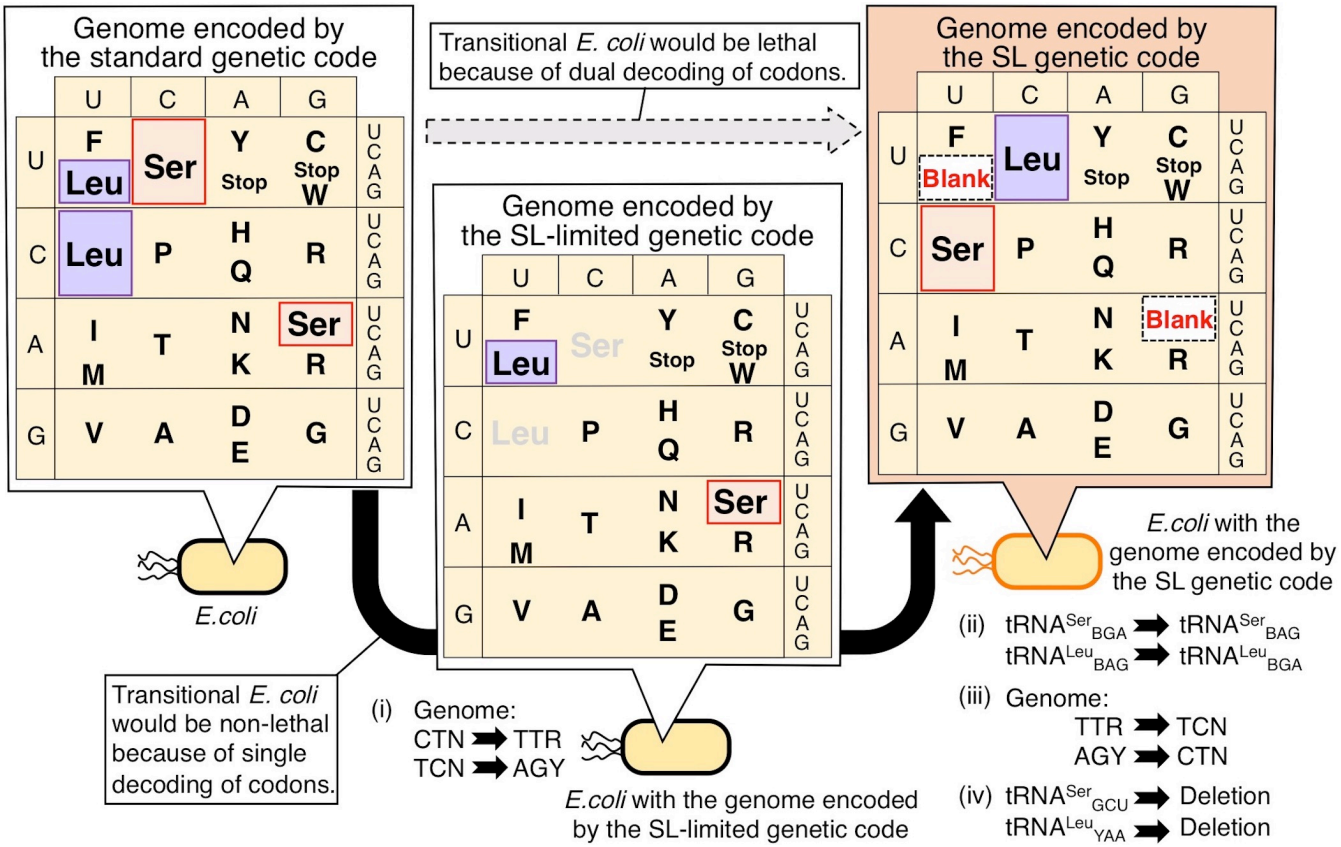

Figure 5. Codon-bypass strategy to create *Escherichia coli* (*E. coli*) strains with the SL genetic code. In the codon-bypass strategy, the *E. coli* genome is modified to the SL-limited genome and then to the SL genome. (i) Replacement of Leu CTN and Ser TCN codons with Leu TTR and Ser AGY codons, respectively, in the genome. (ii) Swapping anticodons of the corresponding $tRNA^{Ser}_{BGA}$ and $tRNA^{Leu}_{BAG}$. Because the resulted SL-limited genome has no Leu CTN and Ser TCN codons, this modification would have no effect on *E. coli* cells. (iii) Replacement of Leu TTR and Ser AGY codons with TCN and CTN codons in the genome, which are newly assigned to Leu and Ser in the genome, respectively. (iv) Deletion of tRNAs corresponding to Leu TTR codons and Ser AGY codons to complete a creation of E. coli strains with SL genetic code. This modification would have no effect on *E. coli* cells because no Leu TTR codons and Ser AGY codons are used in the genome. If the *E. coli* genome is modified to the SL genome directly by skipping the SL-limited genome, a transitional genome would be lethal for *E. coli* because it would contain codons assigned to both Ser and Leu.

Moreover, the ratios of relative fluorescence intensity to the relative amounts of folded protein were similar among the codes (Figure S6B). This indicates that streptavidin and GFP synthesized in the translation systems constructed with mini-Std, SL, or SLA tRNA sets had comparable activities to those synthesized in the translation system using natural tRNAs.

Proposal of "a codon-bypass strategy" to create GMOs with an AA-swapped genetic code. In this paper, we proposed two different AA-swapped genetic codes and demonstrated their orthogonality against the standard genetic code using cell-free translation systems carrying chimeric tRNAs. If the concept of AA-swapped genetic code is applied to living cells, it would give us "a genetic firewall for genes" that can prevent the escape of most genes in cells into the environment. To create GMOs with the AA-swapped genetic code, we must avoid a transition genetic code, which is lethal for the cells because of dual decoding of a codon to two different AAs. For example, there is no technology that can synthesize the entire genome of *E. coli* using the SL genetic code and then reboot the cells, as has been performed for *Mycoplasma mycoides*[45, 46]. Even if such a rebooting process becomes possible, a transition genetic code cannot be avoided because of the coexistence of natural tRNAs derived from the recipient cell and chimeric tRNAs derived from the synthesized genome. Similarly, even if we replace the genome step-by-step using reported technologies[47-49], there would still be a mixture of genes encoded by the standard genetic code and the SL genetic code, which would also be lethal. As a possible solution, we proposed "a codon-bypass strategy" (Figure 5) that would involve the construction of a genome with limited SL codons. In the SL-limited genome, all Leu CTN codons (92515 codons in 3547 CDSs of *E. coli* MDS42) and all Ser TCN codons (37515 codons) would be replaced with TTR codons and AGY codons, respectively. After replacing the codons, the anticodon loops of $tRNA^{Ser}_{BGA}$ and $tRNA^{Leu}_{BAG}$ would be swapped with each other. At that juncture, no genes would contain CTN or TCN codons; thus, tRNA anticodon swapping would not affect cell viability. Finally, all Leu TTR codons and all Ser AGY codons would be replaced with TCN and CTN codons, respectively, and $tRNA^{Ser}_{GCU}$ and $tRNA^{Leu}_{YAA}$ would be deleted to construct the SL genetic code. The blank TTR and AGY codons could be used for unnatural AA survival strategies to create *E. coli* strains that have a genetic firewall for GMOs as well as genes.

Recent efforts to generate blank codons in bacterial cells have strongly progressed to the point at which multiple unnatural AAs have been introduced into an expanded genetic code[16, 17,

[50-54]. In 2019, Chin's group reported the development of an altered genetic code consisting of only 61 codons by replacing two Ser codons (TCR) with two synonymous codons (AGY). Therefore, we believe the aforementioned strategy has a good possibility of creating *E. coli* strains carrying the SL genetic code. Because cells with an AA-swapped genetic code are interesting from the viewpoints of both science and engineering, the creation of *E. coli* strains with the SL genetic code could be one of the next important research targets.

MATERIAL AND METHODS

Expression and purification of C5 protein. The C5 protein gene was amplified by PCR from the *E. coli* JM109 genome using C5.F29 (5'-GCTTGCCATA TGGTTAAGCT CGCATTTCC-3') and C5.R32 (5'- GCAACGCTCG AGGGACCCGC GAGCCAGGCG AC-3') primers. After digestion with *Nde*I and *Xho*I, the resulting DNA product was cloned into the *Nde*I–*Xho*I site of the pET21a(+) vector. The plasmid was transformed into *E. coli BL21(DE3)LysS,* and cells were grown on LB plates with 100 μg/mL ampicillin, 20 μg/mL chloramphenicol, and 2% glucose. The resulting colonies were inoculated to 3 mL of LB with 100 μg/mL ampicillin, 20 μg/mL chloramphenicol, and 5% glucose and grown at 37 °C overnight. The resulting cultures were added to 300 mL of LB with 100 μg/mL ampicillin, 20 μg/mL chloramphenicol, and 5% glucose and grown at 37 °C. At A_{600} = 0.7, protein expression was induced with 0.5 mM IPTG at 25 °C overnight. Cells were pelleted, resuspended in lysis buffer (20 mM Hepes-K pH 7.8, 10 mM imidazole pH 7.8, 300 mM KOAc, 0.2 mM DTT, and 1 mM PMSF), and lysed using a sonicator. After ammonium sulfate precipitation, pellets were dissolved in Buffer A (20 mM Hepes-K pH 7.8, 300 mM KOAc, 0.2 mM DTT, 10 mM $MgCl_2$, 1 M NH_4Cl, and 0.25% Tween20). The mixture was clarified by centrifugation followed by filtration. For affinity purification, an IMAC column (BioRad, CA, USA) was connected to an NGC chromatography system (BioRad) equilibrated with Buffer A. After loading the solution, the column was washed with Buffer A, followed by Buffer B (20 mM Hepes-K pH 7.8, 10 mM imidazole pH 7.8, 300 mM KOAc, 10 mM $MgCl_2$, and 0.2 mM DTT). The protein was then eluted with Buffer C (20 mM Hepes-K pH 7.8, 300 mM AcOK, 10 mM $MgCl_2$, 250 mM Imidazole, and 0.2 mM DTT) and stored at –80 °C in 20% glycerol. The protein concentrations were determined by A_{280} according to a molar extinction coefficient estimated from the amino acid composition.

Preparation of RNase P. The M1 RNA gene was amplified by PCR from the *E. coli* JM109 genome using T7M1-RNA.F37 (5'-TAATACGACT CACTATAGAA GCTGACCAGA CAGTCGC-3') and M1RNA.R22 (5'-AGGTGAAACT GACCGATAAG CC-3') primers. The resulting DNA product was cloned into the pCR4 vector. The DNA template was amplified by two rounds of PCR using pGEM-Tseq (5'-GGAAACAGCT ATGACCATGA-3') and M1RNA.R22 primers. First PCR (100 μL) was performed under the following conditions: 1 × PfSH buffer [10 mM Tris-HCl (pH 8.4), 100 mM KCl, 0.1% (v/v) Triton X-100, 2 mM $MgSO_4$], 0.2 mM each dNTP, 0.5 μM each forward and reverse primer, 0.2 ng/μL pCR4-M1, and 2 nM of *Pfu-S* DNA polymerase[55]. DNA was amplified by 12 cycles of 95°C for 20 s, 55°C for 20 s, and 72°C for 30 s. The second PCR (1600 μL) was performed under the following conditions: 1 × PfSH buffer, 0.2 mM each dNTP, 0.5 μM each forward and reverse primer, amplified DNA (4 μL of first PCR mixture), and 2 nM of *Pfu-S* DNA polymerase. DNA was amplified by 12 cycles of 95°C for 20 s, 55°C for 20 s, and 72°C for 30 s. The amplified DNA was used directly, without purification, for runoff *in vitro* transcription. Transcription of M1 RNA (8000 μL) was performed using the following conditions: 1 × T7 buffer [40 mM Tris-HCl pH 8.0, 1 mM spermidine, 0.01% Triton X-100] 10 mM DTT, 30 mM $MgCl_2$, 5 mM each NTP, amplified DNA (1600 μL of PCR mixture), and 0.3 μM T7 RNA polymerase at 37 °C for 5 h. The resulting M1 RNA was purified by extraction with phenol/chloroform and two rounds of isopropanol precipitation. RNase P was prepared as follows: The solution (8 μL) containing 1 × T7 buffer, 10 mM DTT, 120 mM NH_4Cl, and 20 μM M1-RNA was heated at 95 °C for 5 min followed by cooling to 25 °C. 0.32 μL of 0.3 M $MgCl_2$ and 1.2 μL of 130 μM C5 protein were added to reconstitute RNase P. The product was immediately added to the transcription reaction mixture to avoid self-cleavage.

Preparation of T7-tRNA sets. DNA templates for the tRNAs were prepared by extension of the oligonucleotides listed in Table S1 (see Table S3 for sequences) followed by two rounds of PCR using the primers listed in Table S1 (see Table S3 for sequences). Extension (20 μL) was performed under the following condition: 1 × PfSH buffer, 0.2 mM of each dNTP, 1 μM each forward and reverse oligonucleotides, and 2 nM of *Pfu-S* DNA polymerase. DNA was extended by heating the mixture at 95°C for 1 min, and 5 cycles of 50°C for 1 min and 72 °C for 1 min. First round PCR (100 μL) was performed under the following conditions: 1 × PfSH buffer, 0.2 mM each dNTP, 2% DMSO, 1 μM each forward and reverse primer, extended DNA (2.5 μL of extension mixture), and 2 nM of *Pfu-S* DNA polymerase. DNA was amplified by 5 cycles of 95°C for 20 s, 50°C for 20 s, and 72°C for 30 s. Second round PCR (200 μL) was performed under the following conditions: 1 × PfSH buffer, 0.2 mM each dNTP, 2% DMSO, 1 μM each forward and reverse primer, amplified DNA (5% v/v of first PCR mixture), and 2 nM of *Pfu-S* DNA polymerase. DNA was amplified by 12 cycles of 95°C for 20 s, 50°C for 20 s, and 72°C for 30 s. The amplified DNA was used directly, without purification, for run-off *in vitro* transcription. Transcription of the tRNAs except $tRNA^{Trp}$, $tRNA^{Gln}$, and $tRNA^{Pro}$ (2 mL) was performed under the following conditions: 1 × T7 buffer, 10 mM DTT, 22.5 mM $MgCl_2$, 5 mM GMP, 3.75 mM each NTP, amplified DNA (200 μL of second PCR mixture) and 0.12 μM T7 RNA polymerase at 37 °C overnight. DNase I (0.005 U/μL) and $MnCl_2$ (2 mM) were added to the transcription mixture, and then the mixture was incubated at 37 °C for 2 h. After DNase I treatment, the synthesized tRNAs were purified by extraction with phenol/chloroform and two rounds of isopropanol precipitation. The tRNAs were dissolved in ultrapure water, and their concentrations were determined using a NanoDrop-spectrophotometer ND-1000. Transcription of $tRNA^{Trp}$, $tRNA^{Gln}$, and $tRNA^{Pro}$ (2 mL) was performed under the following conditions: 1 × T7 buffer, 10 mM dithiothreitol (DTT), 22.5 mM $MgCl_2$, 3.75 mM each NTP, amplified DNA (200 μL of second PCR mixture), RNase P (0.4 μM RNase P for $tRNA^{Gln}$ and $tRNA^{Pro}$, 0.8 μM RNase P for $tRNA^{Trp}$), and 0.12 μM T7 RNA polymerase at 37 °C overnight. DNase I treatment and purification of tRNAs was performed as described above. The resulting tRNA sequences are listed in Table S1. Twenty-one different T7-tRNAs ($tRNA^{Ala}_{GGC}$, $tRNA^{Arg}_{GCG}$, $tRNA^{Asn}_{GUU}$, $tRNA^{Asp}_{GUC}$, $tRNA^{Cys}_{GCA}$, $tRNA^{Glu}_{CUC}$, $tRNA^{Gly}_{GCC}$, $tRNA^{His}_{GUG}$, $tRNA^{Ile}_{GAU}$, $tRNA^{Leu}_{GAG}$, $tRNA^{Lys}_{CUU}$, $tRNA^{ini}_{CAU}$, $tRNA^{Met}_{CAU}$, $tRNA^{Phe}_{GAA}$, $tRNA^{Ser}_{GGA}$,

$tRNA^{Thr}_{GGU}$, $tRNA^{Tyr}_{GUA}$, $tRNA^{Val}_{GAC}$, $tRNA^{Gln}_{CUG}$, $tRNA^{Pro}_{GGG}$, and $tRNA^{Trp}_{CCA}$) were combined to prepare the mini-Std T7-tRNA set. For the SL T7-tRNA set, $tRNA^{Leu}_{GGA}$ and $tRNA^{Ser}_{GAG}$ were used instead of $tRNA^{Leu}_{GAG}$ and $tRNA^{Ser}_{GGA}$ from the mini-Std T7-tRNA set. For the SLA T7-tRNA set, $tRNA^{Leu}_{GGA}$, $tRNA^{Ser}_{GGC}$, and $tRNA^{Ala}_{GAG}$ were used instead of $tRNA^{Leu}_{GAG}$, $tRNA^{Ser}_{GGA}$, and $tRNA^{Ala}_{GGC}$ from the mini-Std T7-tRNA set. The concentration of each tRNA was adjusted to 25 μM. The tRNA mix solutions were heated at 95 °C for 5 min and then cooled to 25 °C. The refolded-tRNA mixtures were added to the translation mixtures within 1 h of refolding.

Preparation of peptide-encoding mRNAs. The DNA templates for the peptides were prepared by extension of annealed oligonucleotides, T7esD6MYYY.F55 and Y3xDDRD_nnnXaa.R44 (see Table S3 for sequences), followed by amplification by PCR using T7ex5.F22 and DRDuaaAS.R20 as primers (see Table S3 for sequences). Extension (20 μL) was performed as described above. PCR (200 μL) was performed under the following conditions: 1 × PfSH buffer, 0.2 mM each dNTP, 2% DMSO, 1 μM each forward and reverse primer, extended DNA (1 μL of extension mixture), and 2 nM *Pfu-S* DNA polymerase. DNA was amplified by 12 cycles at 95°C for 20 s, 50°C for 20 s, and 72°C for 30 s. The amplified DNA was used directly, without purification, for runoff *in vitro* transcription. Transcription of the mRNAs (1000 μL) was performed under the following conditions: 1 × T7 buffer, 10 mM DTT, 25 mM $MgCl_2$, 5 mM GMP, 5 mM each NTP, amplified DNA (200 μL of PCR mixture), and 0.12 μM T7 RNA polymerase at 37 °C overnight. The produced mRNAs were purified by extraction with phenol/chloroform and isopropanol precipitation followed by ethanol precipitation. The concentration of mRNA was determined using a NanoDrop-spectrophotometer ND-1000.

Preparation of GFP and streptavidin mRNAs. The template DNAs for GFP were amplified by PCR (1000 μL) using the following conditions: 1 × PfSH buffer, 2% DMSO, 0.2 mM each dNTP, 0.375 μM T7Aex.F24, 0.375 μM GFPuv.R30 (see Table S3 for sequences), synthesized GFP genes (see Figure S3B; GFP_mini-Std, GFP_SL, and GFP_SLA) purchased from Genscript (Piscataway, USA) and Integrated DNA Technologies, Inc. (Coralville, USA), and 2 nM of *Pfu-S* DNA polymerase. DNA was amplified by 12 cycles at 95°C for 20 s, 50°C for 20 s, and 72°C for 1 min. The amplified DNA was purified by extraction with phenol/chloroform and isopropanol precipitation. DNA was dissolved in 100 μL of 10 mM Tris-AcOH pH7.8. The DNA template was transcribed by runoff *in vitro* transcription (1000 μL) using the following conditions: 1 × T7 buffer, 10 mM DTT, 30 mM $MgCl_2$, 5 mM each NTP, template DNA (100 μL of the above DNA solution), and 0.12 μM T7 RNA polymerase at 37 °C overnight. Purification and measurement of the concentration of mRNAs was performed as described above. The mRNAs of streptavidin were prepared with a similar procedure except using template DNAs (Figure S3A; streptavidin_mini-Std, streptavidin_SL, and streptavidin_SLA) and the corresponding primers (StvWT.R20, StvSL.R20, or StvSLA.R20 instead of GFPuv.R30; see Table S3 for sequences).

Cell-free translation. Creatine kinase, creatine phosphate, and *E. coli* tRNA extract were purchased from *Roche* Diagnostics (Tokyo, *Japan*). The translation reactions were similar to those from previously published papers[36-38, 56-59], aside from the concentrations of aaRSs and EF-Tu, which were used to promote aminoacylation of tRNAs and recruitment of aminoacyl-tRNAs to ribosome, respectively[39-41]. The reaction mixture (4 μL) of a cell-free translation system containing 50 mM Hepes-K (pH 7.6), 100 mM potassium acetate, 17.6 mM magnesium acetate, 2 mM spermidine, 1 mM DTT, 2 mM ATP, 2 mM GTP, 1 mM CTP, 1 mM UTP, 20 mM creatine phosphate, 1.5 μM AlaRS, 0.06 μM ArgRS, 0.76 μM AsnRS, 0.26 μM AspRS, 0.04 μM CysRS, 0.12 μM GlnRS, 2.8 μM GluRS, 0.18 μM GlyRS, 0.04 μM HisRS, 2.8 μM IleRS, 0.08 μM LeuRS, 0.22 μM LysRS, 0.06 μM MetRS, 1.4 μM PheRS, 0.32 μM ProRS, 0.08 μM SerRS, 0.18 μM ThrRS, 0.06 μM TrpRS, 0.04 μM TyrRS, 0.04 μM ValRS, 0.6 μM methionyl-tRNA formyltransferase, 0.5 μM ribosome recycling factor, 4 μg/mL creatine kinase, 0.1 μM adenosine kinase, 0.1 μM inorganic pyrophosphatase, 0.1 μM nucleoside-diphosphate kinase, 2.7 μM initiation factor 1, 0.4 μM initiation factor 2, 1.5 μM initiation factor 3, 0.25 μM release factor 2, 0.17 μM release factor 3, 0.26 μM elongation factor G, 70 μM elongation factor Tu/elongation factor Ts, 0.1 μM peptidyl-tRNA hydrolase, 1.2 μM ribosome, 0.5 mM each amino acid (except Asp), 50 μM [^{14}C]Asp or Asp, 10 μM each T7-tRNA or 1.5 mg/mL *E. coli* tRNA extract, 2 μM mRNA coding for a protein, or 2 μM mRNA coding for a peptide was incubated at 37 °C for 3 h.

Analysis of products from the translation reactions. The synthesized peptides in the reaction mixtures were analyzed by tricine SDS-PAGE followed by autoradiography (Pharos FX imager, BioRad). Alternatively, the reactions were performed with Asp instead of [^{14}C]Asp, and the products were analyzed using MALDI-TOF-MS (AXIMA confidence, Shimadzu). Synthesized proteins in the reaction mixtures were analyzed by glycine SDS-PAGE followed by autoradiography. Synthesized proteins were also analyzed by native PAGE followed by autoradiography and fluorescence detection (λ_{ex} = 473 nm, LPB filter; Typhoon FLA 9000, GE Healthcare). For streptavidin analysis, Atto 488-biotin (Sigma-Aldrich) was added to the reaction mixture to label active streptavidin with fluorophore. The excess Atto 488-biotin was removed using streptavidin M280 magnetic beads (Thermo Fisher) before subjecting the samples to native PAGE.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Supplementary Figure 1: Optimization of aaRSs and EF-Tu concentrations in the reconstituted cell-free translation system carrying T7-tRNAs; Supplementary Figure 2: MS analysis of model peptides synthesized in cell-free translation systems containing natural tRNAs or the mini-Std tRNA set; Supplementary Figure 3: Sequences of template DNAs encoding model proteins and the translation products; Supplementary Figure 4: Full image of the gels shown in Figure 4; Supplementary Figure 5: Fraction of Ser and Leu or Ser, Leu, and Ala residues over all AA residues in each protein encoded in 3547 CDSs of E. *coli* MDS42; Supplementary Figure 6: Folding efficiency and fluorescence activity of proteins synthesized with the correct combination of tRNA sets and genes; Supplementary Table 1: Sequences of T7-tRNAs and primer sets used to prepare T7-tRNA templates; Supplementary Table 2: The calculated and observed masses of the peptides; Supple-

mentary Table 3: Primer sequences used in this study (PDF)

AUTHOR INFORMATION

Corresponding Author

Hiroshi Murakami - Department of Biomolecular Engineering, Graduate School of Engineering, Nagoya University; Institute of Nano-Life-Systems, Institutes of Innovation for Future Society, Nagoya University, Nagoya, 464-8603, Japan. Tel: +81-52-789-3327; Fax: +81-52-789-3327; Email: murah@chembio.nagoya-u.ac.jp.

Authors

Tomoshige Fujino - Department of Biomolecular Engineering, Graduate School of Engineering, Nagoya University, Nagoya, 464-8603, Japan.

Masahiro Tozaki - Department of Biomolecular Engineering, Graduate School of Engineering, Nagoya University, Nagoya, 464-8603, Japan.

Author Contributions

Experiments were designed by T.F. and H.M. Samples were prepared by T.F., M.T., and H.M. All data were acquired by T.F. Figure design and manuscript writing were performed by T.F. and H.M. The project was supervised by H.M.

Funding Sources

This work was supported by Grant-in-Aid for Exploratory Research [Grant Number 15K12741 to H.M.]; Grant-in-Aid for Early-Career Scientists [Grant Number 18K14332 to T.F.] from the Japan Society for the Promotion of Science; and a donation from Dr. Hiroshi Murakami.

Conflict of Interest

H.M. has filed patent applications (PCT/JP2016/078921 and US 15/764,015) for the construction of AA-swapped genetic codes. All other authors declare no competing interests.

ACKNOWLEDGMENT

We thank Dr. Daisuke Kiga for his helpful comments on AA-swapped genetic codes. We thank the deceased Mr. Takahiro Ishizawa for his contributions to the preliminary research.

ABBREVIATIONS

SDS, sodium dodecyl sulfate; PAGE, polyacrylamide gel electrophoresis; MALDI-TOF-MS, matrix assisted laser desorption/ionization time-of-flight mass spectrometry; fMet, formyl methionine; PCR, polymerase chain reaction; LB, lysogeny broth; IPTG, isopropyl β-D-1-thiogalactopyranoside; KOAc, potassium acetate; DTT, dithiothreitol; PMFS, phenylmethylsulfonyl fluoride; NH_4Cl, ammonium chloride; $MgCl_2$, magnesium chloride; Tris, tris (hydroxymethyl) aminomethane; HCl, hydrochloric acid; KCl, potassium chloride; $MgSO_4$, magnesium sulfate; dNTP, deoxy nucleoside triphosphate; NTP, nucleoside triphosphate; DMSO, dimethyl sulfoxide; AcOH, acetic acid; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; ATP, adenosine triphosphate; GTP, guanosine triphosphate; CTP, cytidine triphosphate; UTP, uridine triphosphate.

REFERENCES

(1) Diwo, C., and Budisa, N. (2018) Alternative biochemistries for alien life: basic concepts and requirements for the design of a robust biocontainment system in genetic isolation, *Genes 10*, DOI: 10.3390/genes10010017.

(2) Torres, L., Kruger, A., Csibra, E., Gianni, E., and Pinheiro, V. B. (2016) Synthetic biology approaches to biological containment: pre-emptively tackling potential risks, *Essays Biochem. 60*, 393-410.

(3) Whitford, C. M., Dymek, S., Kerkhoff, D., Marz, C., Schmidt, O., Edich, M., Droste, J., Pucker, B., Ruckert, C., and Kalinowski, J. (2018) Auxotrophy to xeno-DNA: an exploration of combinatorial mechanisms for a high-fidelity biosafety system for synthetic biology applications, *J. Biol. Eng. 12*, DOI: 10.1186/s13036-13018-10105-13038.

(4) Moe-Behrens, G. G., Davis, R., and Haynes, K. A. (2013) Preparing synthetic biology for the world, *Front. Microbiol. 4*, DOI: 10.3389/fmicb.2013.00005.

(5) Davison, J. (1999) Genetic exchange between bacteria in the environment, *Plasmid 42*, 73-91.

(6) Steidler, L., Neirynck, S., Huyghebaert, N., Snoeck, V., Vermeire, A., Goddeeris, B., Cox, E., Remon, J. P., and Remaut, E. (2003) Biological containment of genetically modified Lactococcus lactis for intestinal delivery of human interleukin 10, *Nat. Biotechnol. 21*, 785-789.

(7) Gallagher, R. R., Patel, J. R., Interiano, A. L., Rovner, A. J., and Isaacs, F. J. (2015) Multilayered genetic safeguards limit growth of microorganisms to defined environments, *Nucleic Acids Res. 43*, 1945-1954.

(8) Molin, S., Boe, L., Jensen, L. B., Kristensen, C. S., Givskov, M., Ramos, J. L., and Bej, A. K. (1993) Suicidal genetic elements and their use in biological containment of bacteria, *Annu. Rev. Microbiol. 47*, 139-166.

(9) Chan, C. T. Y., Lee, J. W., Cameron, D. E., Bashor, C. J., and Collins, J. J. (2016) 'Deadman' and 'Passcode' microbial kill switches for bacterial containment, *Nat. Chem. Biol. 12*, 82-86.

(10) Caliando, B. J., and Voigt, C. A. (2015) Targeted DNA degradation using a CRISPR device stably carried in the host genome, *Nat. Commun. 6*, DOI: ARTN 6989 6910.1038/ncomms7989.

(11) Wright, O., Delmans, M., Stan, G. B., and Ellis, T. (2015) GeneGuard: A modular plasmid system designed for biosafety, *ACS Synth. Biol. 4*, 307-316.

(12) Droge, M., Puhler, A., and Selbitschka, W. (1998) Horizontal gene transfer as a biosafety issue: A natural phenomenon of public concern, *J. Biotechnol. 64*, 75-90.

(13) Ma, N. J., and Isaacs, F. J. (2016) Genomic recoding broadly obstructs the propagation of horizontally transferred genetic elements, *Cell Syst. 3*, 199-207.

(14) Thomas, C. M., and Nielsen, K. M. (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria, *Nat. Rev. Microbiol. 3*, 711-721.

(15) Tack, D. S., Ellefson, J. W., Thyer, R., Wang, B., Gollihar, J., Forster, M. T., and Ellington, A. D. (2016) Addicting diverse bacteria to a noncanonical amino acid, *Nat. Chem. Biol. 12*, 138-140.

(16) Mandell, D. J., Lajoie, M. J., Mee, M. T., Takeuchi, R., Kuznetsov, G., Norville, J. E., Gregg, C. J., Stoddard, B. L., and Church, G. M. (2015) Biocontainment of genetically modified organisms by synthetic protein design, *Nature 518*, 55-60.

(17) Rovner, A. J., Haimovich, A. D., Katz, S. R., Li, Z., Grome, M. W., Gassaway, B. M., Amiram, M., Patel, J. R., Gallagher, R. R., Rinehart, J., and Isaacs, F. J. (2015) Recoded organisms engineered to depend on synthetic amino acids, *Nature 518*, 89-93.

(18) Xuan, W., and Schultz, P. G. (2017) A strategy for creating organisms dependent on noncanonical amino acids, *Angew. Chem. Int. Ed. Engl. 56*, 9170-9173.

(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" tRNAPheS, *Biochemistry 23*, 1468-1473.

(20) Murakami, H., Ohta, A., Ashigai, H., and Suga, H. (2006) A highly flexible tRNA acylation method for non-natural polypeptide synthesis, *Nat. Methods 3*, 357-359.

(21) Francklyn, C., and Schimmel, P. (1989) Aminoacylation of RNA minihelices with alanine, *Nature 337*, 478-481.

(22) Sampson, J. R., and Saks, M. E. (1993) Contributions of discrete tRNA(Ser) domains to aminoacylation by E.coli seryl-tRNA

synthetase: a kinetic analysis using model RNA substrates, *Nucleic Acids Res. 21*, 4467-4475.

(23) Asahara, H., Himeno, H., Tamura, K., Hasegawa, T., Watanabe, K., and Shimizu, M. (1993) Recognition nucleotides of Escherichia coli tRNA(Leu) and its elements facilitating discrimination from tRNASer and tRNA(Tyr), *J. Mol. Biol. 231*, 219-229.

(24) Asahara, H., Himeno, H., Tamura, K., Nameki, N., Hasegawa, T., and Shimizu, M. (1994) Escherichia coli seryl-tRNA synthetase recognizes tRNA(Ser) by its characteristic tertiary structure, *J. Mol. Biol. 236*, 738-748.

(25) Murakami, H., Ohta, A., and Suga, H. (2009) Bases in the anticodon loop of tRNA(Ala)(GGC) prevent misreading, *Nat. Struct. Mol. Biol. 16*, 353-358.

(26) Ledoux, S., Olejniczak, M., and Uhlenbeck, O. C. (2009) A sequence element that tunes Escherichia coli tRNA(GGC)(Ala) to ensure accurate decoding, *Nat. Struct. Mol. Biol. 16*, 359-364.

(27) Geslain, R., Cubells, L., Bori-Sanz, T., Alvarez-Medina, R., Rossell, D., Martí, E., and Ribas, d. P., L. (2010) Chimeric tRNAs as tools to induce proteome damage and identify components of stress responses, *Nucleic Acids Res. 38*, DOI: 10.1093/nar/gkp1083.

(28) Dorazi, R., Lingutla, J. J., and Humayun, M. Z. (2002) Expression of mutant alanine tRNAs increases spontaneous mutagenesis in Escherichia coli, *Mol. Microbiol. 44*, 131-141.

(29) Kawahara-Kobayashi, A., Masuda, A., Araiso, Y., Sakai, Y., Kohda, A., Uchiyama, M., Asami, S., Matsuda, T., Ishitani, R., Dohmae, N., Yokoyama, S., Kigawa, T., Nureki, O., and Kiga, D. (2012) Simplification of the genetic code: restricted diversity of genetically encoded amino acids, *Nucleic Acids Res. 40*, 10576-10584.

(30) Esvelt, K. M., and Wang, H. H. (2013) Genome-scale engineering for systems and synthetic biology, *Mol. Syst. Biol. 9*, DOI: 10.1038/msb.2012.1066.

(31) Church, G. M. & Regis, E. (2012) Regenesis: how synthetic biology will reinvent nature and ourselves, Basic Books, New York.

(32) Kawaguchi, Y., Honda, H., Taniguchimorimura, J., and Iwasaki, S. (1989) The Codon Cug Is Read as Serine in an Asporogenic Yeast Candida-Cylindracea, *Nature 341*, 164-166.

(33) Ohama, T., Suzuki, T., Mori, M., Osawa, S., Ueda, T., Watanabe, K., and Nakase, T. (1993) Non-universal decoding of the leucine codon CUG in several Candida species, *Nucleic Acids Res. 21*, 4039-4045.

(34) Muhlhausen, S., Findeisen, P., Plessmann, U., Urlaub, H., and Kollmar, M. (2016) A novel nuclear genetic code alteration in yeasts and the evolution of codon reassignment in eukaryotes, *Genome Res. 26*, 945-955.

(35) Riley, R., Haridas, S., Wolfe, K. H., Lopes, M. R., Hittinger, C. T., Goker, M., Salamov, A. A., Wisecaver, J. H., Long, T. M., Calvey, C. H., Aerts, A. L., Barry, K. W., Choi, C., Clum, A., Coughlan, A. Y., Deshpande, S., Douglass, A. P., Hanson, S. J., Klenk, H. P., LaButti, K. M., Lapidus, A., Lindquist, E. A., Lipzen, A. M., Meier-Kolthoff, J. P., Ohm, R. A., Otillar, R. P., Pangilinan, J. L., Peng, Y., Rokas, A., Rosa, C. A., Scheuner, C., Sibirny, A. A., Slot, J. C., Stielow, J. B., Sun, H., Kurtzman, C. P., Blackwell, M., Grigoriev, I. V., and Jeffries, T. W. (2016) Comparative genomics of biotechnologically important yeasts, *Proc. Natl. Acad. Sci. U. S. A. 113*, 9882-9887.

(36) 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.

(37) Iwane, Y., Hitomi, A., Murakami, H., Katoh, T., Goto, Y., and Suga, H. (2016) Expanding the amino acid repertoire of ribosomal polypeptide synthesis via the artificial division of codon boxes, *Nat. Chem. 8*, 317-325.

(38) Cui, Z. L., Stein, V., Tnimov, Z., Mureev, S., and Alexandrov, K. (2015) Semisynthetic tRNA complement mediates in vitro protein synthesis, *J. Am. Chem. Soc. 137*, 4404-4413.

(39) Fukunaga, J., Ohno, S., Nishikawa, K., and Yokogawa, T. (2006) A base pair at the bottom of the anticodon stem is reciprocally preferred for discrimination of cognate tRNAs by Escherichia coli lysyl- and glutaminyl-tRNA synthetases, *Nucleic Acids Res. 34*, 3181-3188.

(40) Sekine, S., Nureki, O., Sakamoto, K., Niimi, T., Tateno, M., Go, M., Kohno, T., Brisson, A., Lapointe, J., and Yokoyama, S. (1996) Major identity determinants in the "augmented D helix" of tRNA(Glu) from Escherichia coli, *J. Mol. Biol. 256*, 685-700.

(41) Nureki, O., Niimi, T., Muramatsu, T., Kanno, H., Kohno, T., Florentz, C., Giege, R., and Yokoyama, S. (1994) Molecular recognition of the identity-determinant set of isoleucine transfer-RNA from Escherichia-coli, *J. Mol. Biol. 236*, 710-724.

(42) Shaw, B. F., Schneider, G. F., Arthanari, H., Narovlyansky, M., Moustakas, D., Durazo, A., Wagner, G., and Whitesides, G. M. (2011) Complexes of native ubiquitin and dodecyl sulfate illustrate the nature of hydrophobic and electrostatic interactions in the binding of proteins and surfactants, *J. Am. Chem. Soc. 133*, 17681-17695.

(43) Buhr, F., Jha, S., Thommen, M., Mittelstaet, J., Kutz, F., Schwalbe, H., Rodnina, M. V., and Komar, A. A. (2016) Synonymous Codons Direct Cotranslational Folding toward Different Protein Conformations, *Mol. Cell 61*, 341-351.

(44) Walsh, I. M., Bowman, M. A., Soto Santarriaga, I. F., Rodriguez, A., and Clark, P. L. (2020) Synonymous codon substitutions perturb cotranslational protein folding in vivo and impair cell fitness, *Proc. Natl. Acad. Sci. U. S. A. 117*, 3528-3534.

(45) Gibson, D. G., Glass, J. I., Lartigue, C., Noskov, V. N., Chuang, R. Y., Algire, M. A., Benders, G. A., Montague, M. G., Ma, L., Moodie, M. M., Merryman, C., Vashee, S., Krishnakumar, R., Assad-Garcia, N., Andrews-Pfannkoch, C., Denisova, E. A., Young, L., Qi, Z. Q., Segall-Shapiro, T. H., Calvey, C. H., Parmar, P. P., Hutchison, C. A., Smith, H. O., and Venter, J. C. (2010) Creation of a bacterial cell controlled by a chemically synthesized genome, *Science 329*, 52-56.

(46) Hutchison, C. A., Chuang, R. Y., Noskov, V. N., Assad-Garcia, N., Deerinck, T. J., Ellisman, M. H., Gill, J., Kannan, K., Karas, B. J., Ma, L., Pelletier, J. F., Qi, Z. Q., Richter, R. A., Strychalski, E. A., Sun, L. J., Suzuki, Y., Tsvetanova, B., Wise, K. S., Smith, H. O., Glass, J. I., Merryman, C., Gibson, D. G., and Venter, J. C. (2016) Design and synthesis of a minimal bacterial genome, *Science 351*, DOI: 10.1126/science.aad6253.

(47) Wang, K. H., Fredens, J., Brunner, S. F., Kim, S. H., Chia, T. S., and Chin, J. W. (2016) Defining synonymous codon compression schemes by genome recoding, *Nature 539*, 59-64.

(48) Lau, Y. H., Stirling, F., Kuo, J., Karrenbelt, M. A. P., Chan, Y. A., Riesselman, A., Horton, C. A., Schafer, E., Lips, D., Weinstock, M. T., Gibson, D. G., Way, J. C., and Silver, P. A. (2017) Large-scale recoding of a bacterial genome by iterative recombineering of synthetic DNA, *Nucleic Acids Res. 45*, 6971-6980.

(49) Isaacs, F. J., Carr, P. A., Wang, H. H., Lajoie, M. J., Sterling, B., Kraal, L., Tolonen, A. C., Gianoulis, T. A., Goodman, D. B., Reppas, N. B., Emig, C. J., Bang, D., Hwang, S. J., Jewett, M. C., Jacobson, J. M., and Church, G. M. (2011) Precise manipulation of chromosomes in vivo enables genome-wide codon replacement, *Science 333*, 348-353.

(50) Ostrov, N., Landon, M., Guell, M., Kuznetsov, G., Teramoto, J., Cervantes, N., Zhou, M., Singh, K., Napolitano, M. G., Moosburner, M., Shrock, E., Pruitt, B. W., Conway, N., Goodman, D. B., Gardner, C. L., Tyree, G., Gonzales, A., Wanner, B. L., Norville, J. E., Lajoie, M. J., and Church, G. M. (2016) Design, synthesis, and testing toward a 57-codon genome, *Science 353*, 819-822.

(51) Napolitano, M. G., Landon, M., Gregg, C. J., Lajoie, M. J., Govindarajan, L., Mosberg, J. A., Kuznetsov, G., Goodman, D. B., Vargas-Rodriguez, O., Isaacs, F. J., Soll, D., and Church, G. M. (2016) Emergent rules for codon choice elucidated by editing rare arginine codons in Escherichia coli, *P. Natl. Acad. Sci. U. S. A. 113*, E5588-E5597.

(52) Lajoie, M. J., Rovner, A. J., Goodman, D. B., Aerni, H. R., Haimovich, A. D., Kuznetsov, G., Mercer, J. A., Wang, H. H., Carr, P. A., Mosberg, J. A., Rohland, N., Schultz, P. G., Jacobson, J. M., Rinehart, J., Church, G. M., and Isaacs, F. J. (2013) Genomically recoded organisms expand biological functions, *Science 342*, 357-360.

(53) Mukai, T., Yamaguchi, A., Ohtake, K., Takahashi, M., Hayashi, A., Iraha, F., Kira, S., Yanagisawa, T., Yokoyama, S., Hoshi, H., Kobayashi, T., and Sakamoto, K. (2015) Reassignment of a rare sense

codon to a non-canonical amino acid in Escherichia coli, *Nucleic Acids Res. 43*, 8111-8122.

(54) Fredens, J., Wang, K. H., de la Torre, D., Funke, L. F. H., Robertson, W. E., Christova, Y., Chia, T. S., Schmied, W. H., Dunkelmann, D. L., Beranek, V., Uttamapinant, C., Llamazares, A. G., Elliott, T. S., and Chin, J. W. (2018) Total synthesis of Escherichia coli with a recoded genome, *Nature 569*, 514-518.

(55) Wang, Y., Prosen, D. E., Mei, L., Sullivan, J. C., Finney, M., and Vander Horn, P. B. (2004) A novel strategy to engineer DNA polymerases for enhanced processivity and improved performance in vitro, *Nucleic Acids Res. 32*, 1197-1207.

(56) Baggott, J. E., Johanning, G. L., Branham, K. E., Prince, C. W., Morgan, S. L., Eto, I., and Vaughn, W. H. (1995) Cofactor role for 10-formyldihydrofolic acid, *Biochem. J. 308*, 1031-1036.

(57) Clemons, W. M., Brodersen, D. E., McCutcheon, J. P., May, J. L. C., Carter, A. P., Morgan-Warren, R. J., Wimberly, B. T., and Ramakrishnan, V. (2001) Crystal structure of the 30 S ribosomal subunit from Thermus thermophilus: Purification, crystallization and structure determination, *J. Mol. Biol. 310*, 827-843.

(58) Kawakami, T., Murakami, H., and Suga, H. (2008) Messenger RNA-programmed incorporation of multiple N-methyl-amino acids into linear and cyclic peptides, *Chem. Biol. 15*, 32-42.

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