Available online at www.sciencedirect.com
SCIENCE DIRECT®
Methods 36 (2005) 279–290
METHODS
www.elsevier.com/locate/ymeth

De novo genetic codes and pure translation display

Zhongping Tan [a], Stephen C. Blacklow [b], Virginia W. Cornish [a], Anthony C. Forster [c,*]

[a] Department of Chemistry, Columbia University, 3000 Broadway, New York, NY 10027, USA
[b] Department of Pathology, Brigham and Women's Hospital, Harvard Medical School, 77 Avenue Louis Pasteur, Boston, MA 02115, USA
[c] Department of Pharmacology and Vanderbilt Institute of Chemical Biology, Vanderbilt University Medical center, 23rd Avenue South at Pierce, Nashville, TN 37232, USA

Accepted 28 April 2005

Abstract

It is appealing to envision engineering translation for the genetically encoded synthesis of new classes of molecules. The complete reassignment of codons to unnatural amino acids at one or two non-adjacent sites per protein has already found wide utility (see other papers in this volume). This has been achieved by suppression at stop codons or rarely used sense codons in crude systems and in vivo. However, competing aminoacyl-tRNAs, aminoacyl-tRNA synthetases, and release factors limit efficiencies and generalization. We maximize flexibility by omitting the competing components and by reconstituting translation from His-tagged initiation and elongation factors. This approach opens up all 64 codons to amino acid reassignment and has allowed incorporation of several adjacent unnatural amino acids for the study of translation mechanism. One potential application is "peptidomimetic evolution" for ligand discovery. Toward this goal, we have demonstrated the display of polypeptides on their mRNAs in a purified translation system, termed "pure translation display."

Keywords: Unnatural amino acid; tRNA; Peptide analog; Protein synthesis; Ribosome; *Escherichia coli*; Pure; Translation; Display; Selection

1. Introduction

Translation of nucleotide sequence into amino acid sequence is dictated by an almost invariant genetic code. Crick's adaptor hypothesis predicted that this invariance was not due to direct contact between each triplet codon and its cognate amino acid, raising the possibility that recognition of aminoacyl-tRNAs by the translation apparatus is independent of the amino acid side chains [1]. Indeed, changing the amino acid attached to a tRNA did not affect the codon specificity of the tRNA [2]. However, subsequent experiments revealed that changing the amino acid on one tRNA usually altered the yield of incorporation in translation [3]. So, every tRNA/anticodon might have a preferred amino acid for optimal translation efficiency after all, and conversely, every amino acid might have a preferred tRNA(s) [4].

Much of the data on substrate recognition by the translation apparatus involve suppressor tRNAs charged with unnatural amino acids that recognize stop codons, synthesized chemoenzymatically in vitro and tested in crude translation systems [3]. Suppression has also been extended to rarely used sense codons via tRNA mutants with four-base anticodons (Sisido et al., this volume). Recently, cells have been engineered to synthesize both types of these suppressor substrates in vivo (Xie and Schultz, this volume). In effect, codon reassignment by these approaches is quantitative, since the suppression products can be purified from the non-suppression 'failure' products by virtue of their longer length. However, a different approach is needed for quantitative sense codon reassignment with three-base anticodons because the failure products have the same number of residues. Thus, alternative approaches were

* Corresponding author. Fax: +1 615 936 5555.
E-mail address: a.forster@vanderbilt.edu (A.C. Forster).

1046-2023/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymeth.2005.04.011

developed that use specially selected codons in crude systems or in vivo to enable highly biased, though incomplete, codon reassignment ([5]; Link and Tirrell, this volume). Studies in crude or in vivo translation systems are invariably complicated by competition from endogenous translation components [6].

Our strategy, initiated in collaboration with Herbert Weissbach, is to maximize specificity and flexibility. This is done by omitting the competing translation components by reconstituting translation from ribosomes, initiation and elongation factors, mRNA and custom-synthesized aminoacyl-tRNAs [7] (Fig. 1). This approach extends the synthetic capability of purified translation systems with unnatural aminoacyl-tRNAs beyond dipeptide synthesis (without translocation) from poly(U) templates [8]. Total aminoacyl-tRNA substrates are direct competitors and therefore must be removed. The 20 aminoacyl-tRNA synthetases are omitted because they recharge most tRNA mutants with natural amino acids after the tRNAs have delivered their unnatural amino acid cargoes, and they also proofread certain unnatural aminoacyl-tRNAs. Release factors were omitted as release of the peptide from the mRNA is undesirable for "pure translation display" (see below [7,9]), but they could be incorporated if fast mRNA turnover is desired. Our strategy has opened up all 64 codons to reassignment, enabled template-directed synthesis of defined unnatural polymers [10] (Fig. 2), and facilitated investigation of translation mechanism [7,11].

The reconstitution was based on a purified translation system from *Escherichia coli* [12]. We chose this system because it is the simplest and best-characterized translation system. We correctly assumed that *E. coli* initiation factors would be viable in His-tagged form, that *E. coli* ribosomes could be purified from abundant contaminating aminoacyl-tRNA synthetases and ribonucleases, that the notoriously labile aminoacyl-tRNA substrates could be stored stably for years, and that neither the aminoacyl-tRNA synthetases themselves nor translation factors EF-P, W, and Rescue would be required for synthesis of long polypeptides, despite some contrary publications (e.g., [13]). The reconstitution was initially limited to only chemically derivatized or un-derivatized natural aminoacyl-tRNA substrates ([7]; see also [14,15]), and later expanded to chemoenzymatically synthesized aminoacyl-tRNAs [10] because of the ability to alter the amino acid building block and anticodon in modular fashion irrespective of the tRNA sequences linking them (Fig. 3; see Lodder et al., this volume). For the latter system, we made some additional assumptions (other than that an unnatural amino acid would be tolerated): that substitution of the terminal CA with dCA would be acceptable (for ease of dinucleotide synthesis), that the translation apparatus would tolerate base substitutions in the aminoacyl stem (designed for optimal transcription in vitro) and anticodon (for changing codon specificity), and that omission of the tRNA nucleoside modifications would be unimportant (for ease of tRNA synthesis). Interestingly, we found that these additional assumptions hold for single incorporations of unnatural L-amino acids, but at least one of these assumptions appears to be invalid for translations requiring several adjacent unnatural L-amino acid incorporations. Unnatural polymer synthesis proceeds less efficiently than predicted, despite the absence of competitors.

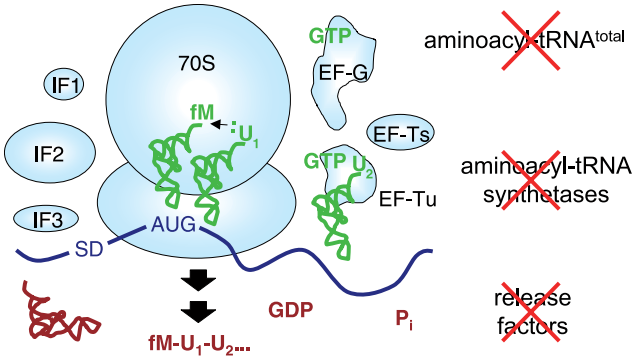

Fig. 1. Our purified His-tagged translation system (light blue) that is programmable with exogenously synthesized mRNA (dark blue) and substrates (green). Competitors were omitted (red crosses) to potentially allow complete reassignment of all 64 sense and nonsense codons. Multiple unnatural amino acids (U_1, U_2, ...) are depicted being incorporated into peptidomimetic product. SD, Shine and Dalgarno ribosome binding site. Regeneration of GTP from GDP is catalyzed by pyruvate kinase using phosphoenolpyruvate substrate (not shown). After translation, peptide products are released from the peptidyl-tRNAs by base-catalyzed hydrolysis.

The primary motivation for these studies was to develop "peptidomimetic evolution" for drug discovery [7]. High-affinity, protease-resistant, cell-permeable peptide ligands might be evolved in vitro using *N*-methyl amino acid building blocks like those prevalent in the orally available cyclosporin A. The idea was based on ribosome display, an in vitro crude translation system for displaying peptide products on the mRNAs that encoded them for directed evolution of peptide ligands [16]. Since a purified system is the only method demonstrated to allow complete reassignment of several codons in an mRNA to unnatural amino acids, translation and display in a purified system was required. We have recently demonstrated "pure translation display" by selecting from an mRNA library only those mRNAs that encode a selectable unnatural amino acid upstream of a peptide spacer long enough to span the ribosome tunnel (Fig. 4). Its efficiency is comparable with ribosome display, but it should be more amenable to optimization because of the absence of ribonucleases, proteases, and contaminating mRNAs.

Here, we detail methods for constructing genetic codes de novo and for pure translation display. We conclude by discussing what we are learning about substrate recognition, such as the amino acid backbone specificity

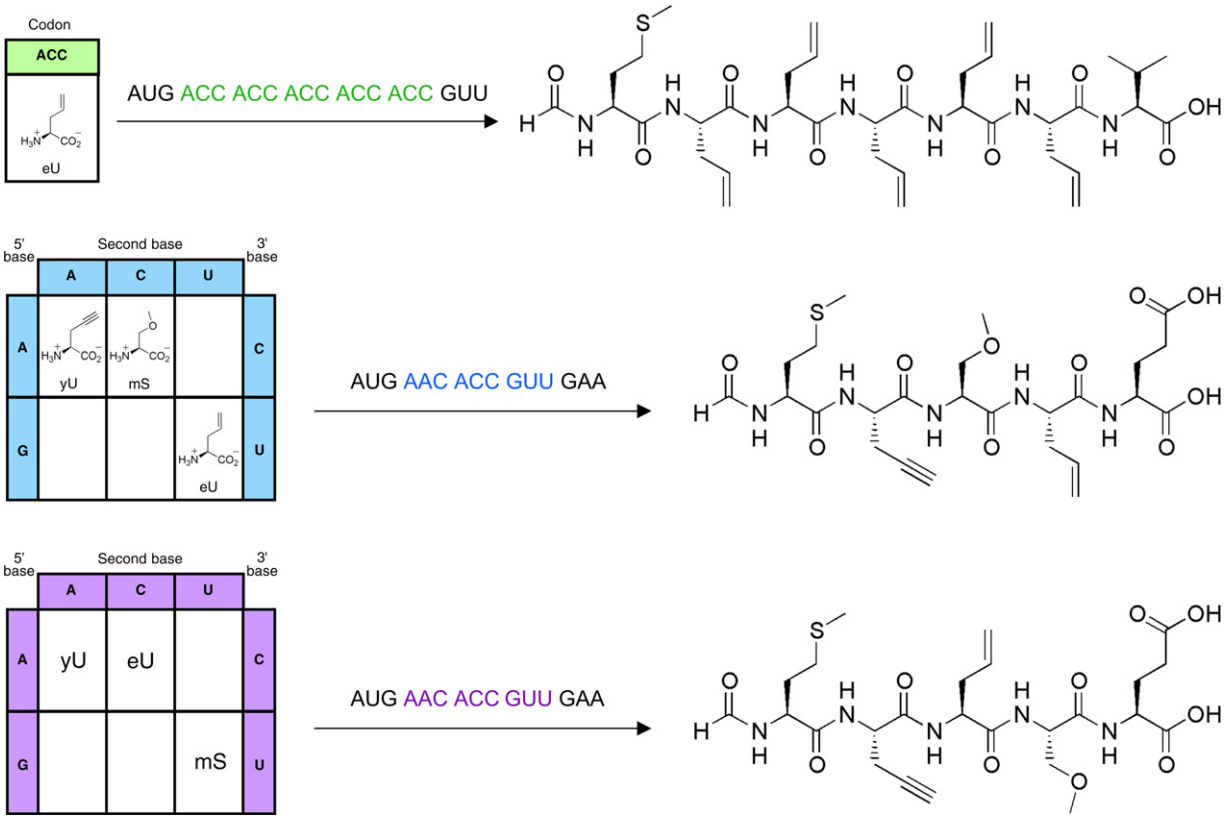

Fig. 2. De novo genetic codes. Summary of all published syntheses of unique full-length translation products containing multiple adjacent unnatural amino acids [10]. mRNAs containing the coding sequences shown above the arrows were translated according to rudimentary de novo genetic codes (left) for the colored codons in the mRNAs. The products were confirmed by complete dependence on the addition of each unnatural aminoacyl-tRNA for their syntheses and by their comigration with authentic marker peptides on HPLC.

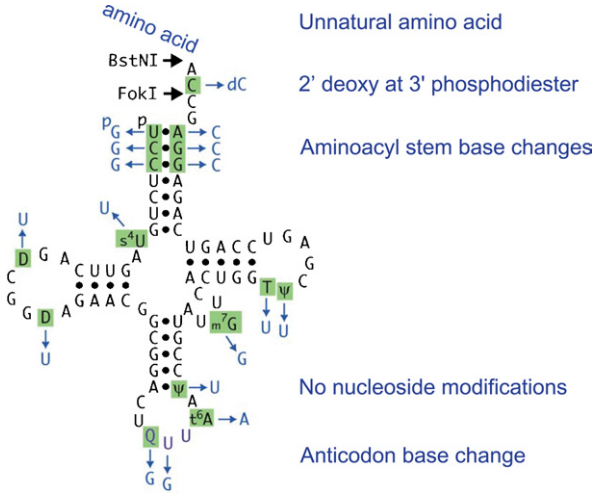

Fig. 3. Natural *E. coli* $tRNA^{Asn}$ (black; the anticodon is purple) and our synthetic ligated modular adaptor, $tRNA^{AsnB}$ (differences from the natural $tRNA^{Asn}$ are in blue). The final step in synthesis is removal of the NVOC protecting group on the amino acid by photolysis. Variants differing in their anticodon sequences and in their chemically charged amino acids were constructed for the syntheses in Fig. 2.

of the translation apparatus (Fig. 5). This knowledge and our ongoing studies should facilitate optimizing and integrating our methods to enable "peptidomimetic

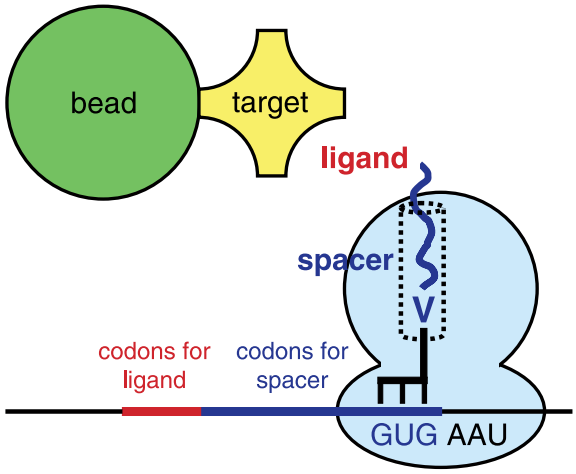

Fig. 4. Pure translation display. In contrast to ribosome display, where libraries of mRNAs lacking stop codons in crude translation systems yield complexes with ribosomes paused at the 3′ ends of the mRNAs [16], pausing in pure translation display occurs within the mRNAs due to omission of cognate aminoacyl-tRNA(s). A spacer sequence is used to extrude the ligand out of the ribosome tunnel. Though the published selection used a biotin ligand encoded by an AUG initiation codon [9], the goal is to display protease-resistant, membrane-permeable "peptidomimetic" ligands for drug discovery.

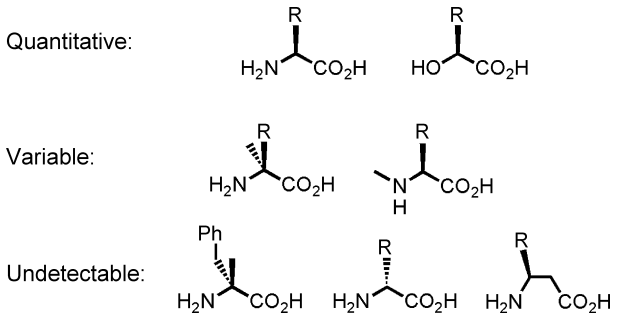

Fig. 5. Amino acid backbone specificity of the translational machinery. Summary of efficiencies measured in our purified system of single incorporations into the middle of peptides, thereby requiring activity of the chemoenzymatic aminoacyl-tRNA at both the ribosomal A and P sites and in translocation [11]. The sensitivity limits of the assays are ~5% of the yield with wild-type substrates. In general, the efficiencies are higher than in other systems, but they are in qualitative agreement [3].

evolution,'' and should improve translation engineering in general.

2. Preparation of *E. coli* translation factor proteins

2.1. Materials

Lysis buffer. 50 mM Tris–HCl, pH 8.0, 300 mM NaCl, 1 mM $MgCl_2$, 10 mM imidazole, 10 mM β-mercaptoethanol.
Wash buffer. 50 mM Tris–HCl, pH 8.0, 300 mM NaCl, 1 mM $MgCl_2$, 20 mM imidazole, 10 mM β-mercaptoethanol.
Elution buffer. 50 mM Tris–HCl, pH 8.0, 300 mM NaCl, 1 mM $MgCl_2$, 250 mM imidazole, 10 mM β-mercaptoethanol.
Dialysis buffer. 10 mM Tris–HCl, pH 7.4, 1 mM $MgCl_2$, 1 mM DTT.
E. coli BL21(DE3) (Novagen, Catalog No. 69387-3).
E. coli BL21(DE3)pLysS (Novagen, Catalog No. 69388-3).
Ni–NTA His-Binding resin (Novagen, Catalog No. 70666).
Slide-A-Lyzer dialysis cassette (Pierce).

2.2. Plasmids for over-expression of His-tagged E. coli initiation and elongation factor proteins

Escherichia coli IF1, IF2, and IF3 coding sequences, each containing an insertion of six histidines immediately after the N-terminal methionine, were synthesized by PCR from published plasmids and sub-cloned into a vector derived from pET24a (Novagen) [7]. His-tagged *E. coli* EF-Tu and EF-Ts were kindly supplied by Hwang and Miller [17]. His-tagged *E. coli* EF-G was kindly supplied by Savelsbergh and Wintermeyer [18].

2.3. Over-expression and purification of His-tagged E. coli translation factor proteins [7]

Expression of three initiation factor subclones in *E. coli* BL21(DE3)pLysS was induced with 0.1 mM IPTG. All the factors were expressed predominantly in the soluble cellular fractions and purified by step elution from Ni–NTA columns using standard protocols (Qiagen). They were dialyzed in Slide-A-Lyzer dialysis cassette against dialysis buffer. Precipitated IF3 was recovered by redissolving in 5 M urea, diluted, and then dialyzed against dialysis buffer containing 100 mM NH_4Cl. Expression and purification of three elongation factors was the same as for initiation factors, except that 10 μM GDP was included up to the last dialysis step for EF-Tu. All factors were stored at −80 °C, and all but EF-Tu could be thawed many times without loss of activity. EF-Tu was stored at 4 °C after thawing and used within a few weeks. Assays of individual factors were as described [7] High specific activity Met is incompatible with Dowex translation assays.

3. Ribosome purification

3.1. Materials

Buffer BIII. 10 mM Tris–acetate, pH 8.2, 14 mM $MgOAc_2$, 60 mM KOAc, 1 mM DTT.
Wash buffer. 10 mM Tris–acetate, pH 8.2, 14 mM $MgOAc_2$, 60 mM KOAc, 1 mM DTT, 1 M NH_4Cl.
Storage buffer. 1 mM Tris–HCl, pH 7.4, 10 mM $MgOAc_2$, 1 mM DTT.
SOLR E. coli host strain (Stratagene, Catalog No. 200298).
Thick-wall polycarbonate centrifuge tubes (Beckman Coulter, Catalog No. 355631, 38 ml).

3.2. Purification of E. coli ribosomes [7,19]

SOLR cells grown to mid-log phase in 50 μg/ml kanamycin were resuspended in buffer BIII (1 ml/g), sonicated, and the cell debris spun down at 10,000*g* for 15 min at 4 °C. The supernatant was centrifuged at 30,000*g* for 30 min at 4 °C. The volume of the resulting supernatant (S30) was adjusted with 5× buffer BIII to fill the polycarbonate centrifuge tube, and the solution centrifuged at 23,600 rpm (100,000*g*) for 15 h at 4 °C (Beckman L-70 ultracentrifuge, rotor Sw28). The supernatant (S100) was stored at −80 °C. The ribosome pellet was washed by stirring in wash buffer (0.1 ml × weight of cell pellet) at 4 °C overnight and then repelleting at 23,600 rpm. The washing was repeated twice more to give 3× washed ribosomes. The 3× ribosome pellet was washed again by stirring in wash buffer overnight and

centrifuged at 23,600 rpm for 1 min to remove the particulate contaminants. The supernatant was then spun at 23,600 rpm for 15 h and the pellet was resuspended in storage buffer to give 1 OD_{260} U/μl 4× ribosomes solution, which was stored at −80 °C. Translation assays using ribosomes washed three to four times had good dependencies on translation factors.

Note 1. An alternative method performed in other laboratories yields ribosomes that are substantially more active (see Note 13).

4. mRNA preparation

4.1. Materials

5× Milligan transcription buffer. 200 mM Tris–acetate, pH 8.0, 5 mM spermidine, 0.25 μg/μl, BSA, 100 mM $MgCl_2$.
Elution buffer. 500 mM NaOAc, pH 5.0, 0.1 mM EDTA.
T7 RNA polymerase (USB, Catalog No. 70001Y, 100 U/μl, 6000 U).
Synthetic oligodeoxynucleotides (Invitrogen).

4.2. Short mRNAs [20]

1. Synthetic oligos were purified by electrophoresis on 8% polyacrylamide/7 M urea gels 1.5 mm thick with 80 μg/well. The DNAs were stained by 0.05% toluidine blue, excised, and then eluted in 0.5 ml elution buffer overnight at 37 °C. The gel slice supernatants were removed and, in the case of very short oligos (e.g., an 18-mer), mixed with 5 μl of 1 M $MgCl_2$. The samples were then precipitated with EtOH.
2. To prepare a hybridized transcription template [20], 29 μl of 0.053 μg/μl 56-mer DNA oligo and 21 μl of 0.026 μg/μl 18-mer oligo were mixed, heated at 90 °C for 3 min, then chilled on ice.
3. A typical 100 μl of transcription reaction mixture was made by adding 6 μl of hybridized oligos and 2 μl of T7 RNA polymerase to 92 μl of transcription mixture prepared at room temperature containing 20 μl of 5× Milligan transcription buffer, 1 μl of 0.5 M DTT, 5 μl of 0.2 % Triton X-100, 27 μl of 30% PEG 8000, 16 μl of 25 mM each NTP, pH 8.0, and 23 μl H_2O. The transcription mixture was incubated at 37 °C for 2–4 h.
4. The transcript was purified by gel electrophoresis as described for the oligos.

Note 2. Extended oligo deprotection times by the manufacturer are sometimes necessary for optimal transcription of long oligos.

4.3. Long mRNAs

The DNA templates were prepared by cloning oligos or PCR-amplified oligos into plasmid vectors [9], followed by plasmid digestion. Run-off transcription and transcript purification were as described above.

5. Aminoacyl-tRNA preparation

5.1. Materials

5× Charging premix (500 mM Hepes–KOH, pH 7, 75 mM $MgCl_2$, 75 mM DTT, 25 mM neutralized ATP).
5× standard buffer (750 mM KCl, 100 mM imidazole–HCl, pH 7.6, 50 mM β-mercaptoethanol, 35 mM $MgCl_2$, 0.5 mM EDTA).
Ligation buffer (86 mM $MgCl_2$, 290 mM Hepes, pH 7.5, and 1.4 mM ATP).
tRNA isoacceptors (Subriden RNA (now defunct) and Sigma).
Folinic acid, Ca salt (Sigma Catalog No. F-7878).
FokI restriction enzyme (NEB, Catalog No. R0109S, 4000 U/ml, 1000 U).
T4 RNA ligase (NEB, Catalog No. M0204L, 20,000 U/ml, 5000 U).
Synthetic oligodeoxynucleotides (Invitrogen).
dC^{Bz} *phosphoramidite* (Applied Biosystems, Catalog No. 401160).
Bis(2-cyanoethyl)-N,N-diisopropyl-phosphoramidite (Toronto Research Chemicals, Inc., Catalog No. B42000).
6-Nitroveratryl chloroformate (Fluka, Catalog No. 23245).
Amino acid and unnatural analogs (Sigma–Aldrich–Fluka, Bachem, Advanced ChemTech, Chem-Inpex International and NovaBiochem).
Photolysis equipment (ACE-Glass, Immersion lamp, Catalog No. 7883-14; Power supply, 777830-60; Reflector, 7883-02).

5.2. Natural aminoacyl-tRNA preparation [7]

Natural aminoacyl-tRNAs were prepared from pure isoacceptors on the large scale after pilot small-scale studies. The source of the enzymes was either a tRNA-free preparation of total *E. coli* synthetases partially purified from a 150,000*g* supernatant by step elution with 0.3 M KCl from DEAE–Sepharose [7] or individual His-tagged aminoacyl-tRNA synthetases kindly prepared by Nalam from plasmids kindly provided by Ueda [21]. A typical large-scale synthesis combined 16 μM of tRNA isoacceptor, 60 μM of non-radioactive amino acid (or 16 μM if radioactive), and 3 μM of pure aminoacyl-tRNA synthetase in 1 ml of 1× charging pre-

mix at 37 °C for 30 min. The reaction was immediately extracted with phenol saturated with NaOAc, pH 5, then precipitated with EtOH at −80 °C. After resuspension in 1 ml of 0.5 mM KOAc, pH 5.1, the aminoacyl-tRNA was concentrated and rinsed with the same buffer by centrifuging in a tube containing a 10,000 MW cut off membrane (Vivaspin), then stored at −80 °C.

Note 3. Aminoacyl-tRNAs are notoriously unstable to hydrolysis, especially under physiological or more alkaline conditions in the presence of divalent metal ions, with stability varying roughly according to the size of amino acid side chain (see Fig. 7 of [22]). However, we find that the aminoacyl-tRNA stock solutions we have worked with can be stored for years with many freeze–thaw cycles without substantial deacylation, as measured by acid/urea gel electrophoresis [23] and staining with toluidine blue.

One natural aminoacyl-tRNA, fMet-tRNA$_i^{Met}$, requires additional steps for its preparation: formylation of Met-tRNA$_i^{Met}$.

1. Methionyl-tRNA$_i^{Met}$ formyltransferase was over-expressed from a plasmid kindly provided by Ueda [21].
2. *N*5,10-Methenyl-tetrahydrofolate was prepared from folinic acid and stored as described [24]. This was converted into the unstable formylation substrate, 10-formyl-tetrahydrofolate, in situ as described below.
3. *N*5,10-Methenyl-tetrahydrofolate (6 mM, 165 μl) was neutralized with 5× standard buffer (448 μl) in 2.4 ml at 25 °C for 10 min. Immediately following, a 730 μl aliquot was mixed with 58 μl of 1× standard buffer, 112 μl of 25 μM Met-tRNA$_i^{Met}$ (charged and stored as above in this section), and 100 μl of methionyl-tRNA$_i^{Met}$ formyltransferase (300 ng, freshly diluted in 1× standard buffer) and the reaction incubated at 25 °C for 10 min. The reaction was immediately extracted with phenol and the product recovered as described above in this section.

Note 4. The extent of formylation cannot be assayed by acid/urea gel electrophoresis. If the fMet-tRNA$_i^{Met}$ contained radioactive methionine, formylation was assayed by hydrolysis and passage through Dowex (as described in Section 6.3 below). If the fMet-tRNA$_i^{Met}$ was non-radioactive, it was either prepared in parallel with radioactive fMet-tRNA$_i^{Met}$ or assayed in initiation or translation assays [7]. High specific activity Met is incompatible with Dowex translation assays.

5.3. Preparation of truncated unmodified tRNA transcripts

The DNA templates were prepared by cloning oligos or PCR-amplified oligos into plasmid vectors [10,25]. Templates for tRNAs lacking their 3′-terminal CA sequence were prepared by digestion with *Fok*I, while control templates for production of full-length tRNAs were prepared by digestion of the same plasmids with *Bst*NI. Transcriptions were as in Section 4.2 above, except that 20 mM neutralized GMP was included to ensure that virtually all transcripts had a 5′ monophosphate.

Note 5. The sense strand of the tRNA sequence was cloned in the orientation of the vector not transcribed by *E. coli* RNA polymerase as a precautionary measure aimed at preventing expression of mutant tRNA precursor RNA, which might be processed by tRNA processing enzymes into toxic tRNA.

5.4. NVOC-aminoacyl-pdCpA synthesis

1. The dinucleotide was synthesized using standard nucleotide chemistry as described [26]. Briefly, 6-*N*, 6-*N*,2′-*O*,3′-*O*-tetrabenzoyl adenosine was prepared by transiently protecting the 5′ hydroxyl of adenosine with dimethoxytrityl. The protected adenosine was then coupled to the 2′-deoxycytidinylphosphoramidite and oxidized under standard conditions. A phosphoramidite was then added to the 5′ hydroxyl group of deoxycytidine and subsequently oxidized. Finally, the benzoyl and cyanoethyl protecting groups were removed by concentrated ammonium hydroxide in a sealed flask at 55 °C for 25 h. The reaction mixture was concentrated under vacuum and the residue was dissolved in H_2O. After removing the insoluble material by centrifugation, the pdCpA was purified by HPLC and ‘‘activated’’ by ion-exchange chromatography to obtain 2.2 equivalents of tetrabutylammonium counter ion per dinucleotide.

Note 6. If the ratio was lower than 2.2, the partially activated pdCpA was redissolved in H_2O and the appropriate amount of tetrabutylammonium hydroxide was added.The resulting solution was lyophilized to total dryness, which normally takes a day. The dryness of this pdCpA product controls the reactivity of the dinucleotide. A representative ^1H NMR (in D_2O) of the tetrabutylammonium salt of pdCpA is shown in Fig. S2 of [11].

2. The natural amino acids and their analogs were prepared as the *N*-nitroveratryloxycarbonyl (NVOC) cyanomethyl active ester form. Briefly, the NVOC protecting group was installed using 6-nitroveratryl chloroformate under standard conditions, and then the cyanomethyl group was introduced using chloroacetonitrile and triethylamine as the base. Typically, the side chains were protected as nitroveratryl ether for hydroxyls using 6-nitroveratryl bromide (or *t*-butyldimethylsilyl ether), nitroveratryl ester for

carboxylic groups using 6-nitroveratryl alcohol and nitroveratryl carbamate for amine groups using 6-nitroveratryl chloroformate.
3. Five equivalents of cyanomethyl active ester was used to selectively acylate pdCpA at the 2′/3′ hydroxyl group using the tetrabutylammonium salt of pdCpA in anhydrous DMF at room temperature. The reaction was monitored by comparing the pdCpA peak and the NVOC-acyl-pdCpA peak (absorptions at 260 nm) on analytical HPLC (Waters Delta 600 using a LiChroCART 250-4, RP18 column). After the ratio of these two peaks was 1:1, the reaction was quenched by adding 4:1 NH_4Ac (50 mM, pH 4.5)/acetonitrile, and the aminoacyl-pdCpA was purified using preparative HPLC (Waters Delta 600 using a Whatman Partisil 10 ODS-2 (C18) column). After desalting, the NVOC-acyl-pdCpA was dissolved in H_2O and its concentration was calculated by UV absorption at 260 and 350 nm (ε260 for pdCpA is 25,000 $cm^{-1}\,M^{-1}$; ε260 and ε350 for 6-nitroveratryl group is 2140 and 6336 $cm^{-1}\,M^{-1}$, respectively).

Note 7. Standard purities of starting materials and reaction products are sometimes insufficient to prevent significant incorporation of contaminants into translation products if the amino acid analog fails to incorporate (see Note 12).

5.5. NVOC-aminoacyl-tRNA preparation [10,11]

NVOC-acyl-pdCpA derivatives of natural amino acids and their analogs were ligated to $tRNA^{-CA}$ species.

1. Eleven nanomoles of NVOC-aminoacyl-pdCpA was lyophilized to dryness and dissolved in 4 μl DMSO followed by addition of 8.5 μl H_2O. Nine micrograms of $tRNA^{-CA}$ transcript in 11.5 μl of 0.1 mM EDTA was denatured at 85 °C for 2 min and cooled to 35 °C over 45 min. Renatured $tRNA^{-CA}$ was spun down, mixed with 7 μl ligation buffer, and the 12.5 μl of NVOC-aminoacyl-pdCpA solution was then added followed by 9 μl of T4 RNA ligase. The resulting 40 μl solution was incubated at 37 °C for 30 min.
2. Ligation was immediately terminated by adding 16 μl of 3 M NaOAc, pH 4.5, and 40 μl H_2O. The mixture was extracted with phenol saturated with 0.3 M NaOAc, pH 5, then phenol/$CHCl_3$, and then $CHCl_3$. NVOC-aminoacyl-tRNA was precipitated with EtOH, redissolved in 12 μl of 0.5 mM KOAc, pH 5.1, and stored at −80 °C.

Note 8. A sevenfold higher concentration of T4 RNA ligase than recommended in standard methods was used routinely because some derivatives needed more ligase than others for efficient ligation. The efficiencies of ligation and recovery should always be checked by acid/urea gel electrophoresis [23] and staining with toluidine blue.

3. In a pilot study, the efficiency of photolytic removal of the NVOC protecting group was confirmed using a 5 μl aqueous solution of 450 μM *N*-NVOC-L-allyl-Gly-pdCpA [26]. After photolysis for 25 min at 10 °C, deprotection was 90% as measured by HPLC.

Note 9. Since protected aminoacyl-tRNAs are more stable to storage than deprotected ones, deprotection is generally performed just before translation. However, deprotection does not preclude storage for as long as a year.

6. In vitro translation

6.1. Materials

5× Dipeptide premix, pH 6.73: 180 mM Tris–acetate, pH 7.5, 50 mM sodium 3,3-dimethyl-glutarate, pH 5.69, 180 mM NH_4OAc, 10 mM DTT, 140 mM potassium phosphoenolpyruvate, pH 6.55, 200 mM KOAc, 4 mM spermidine–HCl.
Pyruvate kinase (Sigma, Catalog No. P1506, 2.6 U/μl, prepared from rabbit muscle).
Dowex 50X8-200, ion-exchange resin (Acros, Catalog No. 20307-5000).
Ultima Gold Scintillation Cocktail (Perkin-Elmer, Catalog No. 6013329).
FMOC-protected amino acids (Sigma–Aldrich–Fluka, Bachem, Advanced ChemTech, NovaBiochem., and Chem-Impex International).
Scintillation counter (Perkin-Elmer, 2200 CA TRI-CARB Liquid Scintillation Analyzer, dual label counting program).

6.2. In vitro translations

1. Photolysis of NVOC-aminoacyl-tRNAs was carried out at 15 μM in about 1 μl in a 0.5 ml plastic tube in a large ice water bath on ice for 25 min (the bath keeps the final temperature below 25 °C).
2. Translations were performed as described [7,10]. As an example of one protocol, translation mix minus elongator tRNAs/mRNA was prepared on ice by mixing in order 10 μl of 5× dipeptide premix, 0.96 μl of 0.5 M $MgOAc_2$, 0.72 μl of diluted pyruvate kinase (0.48 U/μl, diluted with 25 mM Hepes–KOH, pH 7.0), 0.34 μl of 0.7 μg/μl IF1, 0.43 μl of 5.6 g/μl IF2 0.34 μl of 1.6 μg/μl IF3, 0.5 μl of 0.1 M GTP, pH 7.1, 0.98 μl of 4× washed ribosomes (1.0 A_{260} U/μl), 4.80 μl of 3.2 μg/μl EF-Tu, 0.30 μl of 2.6 μg/μl EF-Ts, 1.12 μl of 1.7 μg/μl EF-G, and 1.32 μl of 19 μM

fMet-tRNA$_i^{Met}$. Translations on the 5 μl scale were prepared by mixing 2.18 μl of translation mix minus elongator tRNAs/mRNA, H_2O, natural and unnatural elongator aminoacyl-tRNAs (to ~0.5 μM final concentration each), and, finally mRNA (0.5 μM final concentration). The translation was performed at 37 °C for 30 min without preincubation.

Note 10. Problems due to loss of activity in the translation assay were generally solved faster by going back to stocks of most of the components rather than by trouble-shooting the many components individually.

6.3. Analysis using cation-exchange mini-columns [27]

One microliter of 1 M NaOH was added to the translation reaction and the mixture incubated at 37 °C for 30 min to release the amino acids and peptides from the tRNAs. After acidification with 0.5 ml of 0.5 M HCl, the reaction was analyzed by passage through a cation-exchange column (Dowex 50X8-200 ion-exchange resin) to separate formylated peptides (and formyl-methionine) from unformylated amino acids. The translation products were eluted with 2 ml H_2O (additional elution volumes do not improve the peptide yield). The radioactivity of the translation products was counted in 10 ml of Ultima Gold Scintillation Cocktail. The translation efficiency was calculated by subtracting background counts, then comparing with the positive control peptide synthesized from just natural aminoacyl-tRNAs. There are two standard background controls which should yield the same d.p.m.: dependence on mRNA, defined by translations omitting mRNA, and dependence on elongator aminoacyl-tRNA, defined by translations omitting an elongator aminoacyl-tRNA corresponding to the codon of interest (with the proviso Note 11).

Note 11. With respect to control translations omitting a cognate elongator tRNA, there are certain combinations of codons and non-cognate elongator aminoacyl-tRNAs that yield an incomplete dependence on cognate elongator (Forster et al., unpublished data). Contamination of a non-cognate elongator by a cognate one does not seem to be responsible because the illegitimate codon reading has been traced to individual tRNA transcripts from several different plasmid clones. This implies that efficient readthrough occurred in vitro under these control omission conditions, although it is unclear if readthrough occurs in the corresponding test translations that contain cognate elongator. One possible explanation is that our ionic conditions, though suitable for synthesis of active proteins in vitro [28], are not as stringent as in vivo. However, complete starvation at a codon is not a typical physiological condition, and in vitro conditions generally used for measuring the accuracy of translation entail competition between cognate and non-cognate aminoacyl-tRNAs. When there is doubt about the amino acid inserted at a particular codon for any reason, e.g., due to a control translation exhibiting a poor elongator dependency on Dowex assay, it is necessary to use an alternative assay for product analysis, such as HPLC (see below) or mass spectrometry. In practice, we avoid combinations of codons and non-cognate aminoacyl-tRNAs that exhibit incomplete dependence on the cognate elongator by Dowex assay because the alternative HPLC assay is slower and less sensitive.

6.4. HPLC comigration

HPLC analysis in the presence of peptide markers [7] was used as a more specific test than the Dowex assay. It also has the advantage of very low background signal, but there is substantial product loss during analysis.

1. The marker peptides were synthesized by Research Genetics or by using a solid phase peptide synthesizer [ACT 396–5000 MPS (multiple peptide synthesizer)], Wang-resin as the resin and HOBt, HBTU, and DIEPA or HOAt, HATU, and DIEPA as the coupling reagents. Peptides were cleaved from the resin using TFA/H_2O/TIS = 95:2.5:2.5 and precipitated with ether. The precipitates were dried, redissolved in 1:1 acetonitrile/H_2O and purified by preparative HPLC. All marker peptides were confirmed by mass spectrometry. Some short peptides were confirmed by ^{1}H NMR.
2. Five microliters of radiolabeled translation reaction was treated with 1 μl of 1 M NaOH at 37 °C for 30 min, mixed with 16 μl solution of authentic marker peptide and acidified by 2.3 μl acetic acid. This translation mixture was then filtered through a Microcon 10 and analyzed by reverse-phase analytical HPLC on a C_{18} column using isocratic gradients at 1 ml/min. Eluant was collected at 1 min intervals in scintillation vials. The marker peptides were followed by absorbance at 229 nm and the translation peptide products by d.p.m.

Note 12. HPLC assay is necessary to confirm key Dowex results, such as incorporations that established new length records for defined unnatural polymer synthesis (Fig. 2), incorporations of backbone analogs that were in conflict with previous literature [11], and very inefficient incorporations. Translations with inactive chemoenzymatic substrates favor incorporation of contaminants (see Note 7) [11] because our assays generally use vast excess of substrate over detectable levels of product. One trace contaminant that incorporated has been identified by HPLC comigration: L-Ala contaminates D-Ala (Forster et al., unpublished data).

7. Pure translation display

7.1. Materials

Wash buffer (WB: 50 mM $Mg(OAc)_2$, 150 mM NaCl, 50 mM Hepes–KOH, pH 7.0, 10 mM DTT, 0.1% Tween 20).
Elution buffer (EB: 100 mM EDTA, 150 mM NaCl, 50 mM Hepes–KOH, pH 7.0, 10 mM DTT, 0.1% Tween 20).

7.2. Translations

mRNAs encoding poly(V,T) spacer sequences up to 100 amino acids long were used to ensure that the N-terminal biotinyl residue was extruded out of the ribosome tunnel for interaction with avidin target (Fig. 4). Translations were performed as above without preincubation so that initiation and recycling should be rate-limiting, thereby preventing premature termination due to exhaustion of elongator substrates. Aminoacyl-tRNA concentrations were ~0.2 μM fMet-$tRNA_i^{fMet}$ or biotinylMet-$tRNA_i^{fMet}$, ~1 μM Thr-$tRNA^{Thr}$ and ~1.5 μM Val-$tRNA^{Val}$. Products were analyzed by binding to TetraLink avidin beads (Promega) or by trichloroacetic acid (TCA) precipitation on nitrocellulose filters.

Note 13. Some 20% of the elongator substrates were incorporated, assuming quantitative precipitation and recovery for analysis, and higher incorporations were obtained by substitution of the ribosomes with a more active preparation [29,30] kindly supplied by Olsson and Ehrenberg.

7.3. Pure translation display

The washing and elution steps built upon those used for ribosome display [16].

Target. The avidin beads were first equilibrated with WB. As a precautionary measure, the beads were then blocked at 4 °C for ~5 min by addition of a similar volume of a translation that was identical in composition to the test translations, except that the initiator was fMet-$tRNA_i^{fMet}$ and the mRNAs were non-radioactive. The blocked bead suspension was then divided into 10 μl aliquots. For biotin-blocked controls, translation-blocked bead aliquots were further blocked with a quarter volume of biotin (5 mM) at 37 °C for 5 min.
Library. A library of mRNA/ribosome/peptidyl-tRNA complexes was prepared by translating in 5 μl a mixture of ^{32}P-labeled mRNAs using biotinylMet-$tRNA_i^{fMet}$ and the most active ribosome preparation for 30 min at 37 °C. The complexes were stabilized by chilling at 0 °C for 2 min, adjusting the Mg^{2+} concentration to 48 mM by addition of 95 μl ice-cold WB (using pre-cooled pipette tips), and performance of all remaining steps in a 4 °C room. The tubes were incubated for 5 min on ice and the solutions were then transferred (to leave behind any plastic-binding complexes) into the tubes containing the blocked beads and mixed by tapping every 10 min over a 40 min period at 4 °C. The beads were spun down and the supernatants were removed and the beads were washed three times with 200 μl WB. mRNAs were eluted from bound complexes by addition of 40 μl EB and incubation for 5 min, the supernatants were recovered, and any residual contaminating beads were removed by filtration.

Note 14. The efficiency of recovery of the 101-amino acid-encoding mRNA was estimated by gel electrophoresis and autoradiography to be only ~0.2%, similar to that of crude translation displays. In contrast to crude systems, optimization to virtually 100% should be possible in a purified system [30]. One possible explanation for our low yield may be aggregation due to the use of a highly hydrophobic poly(V,T) spacer sequence. This presumably explains why analysis of the translation products by various gel systems proved impractical.

8. Concluding remarks

The vast majority of published incorporations of unnatural amino acids have been at a single UAG stop codon, with efficiencies typically below 50%. We have established that the amino acid specificities encoded by several sense codons can be completely reassigned, as predicted a half century ago by Crick. Our method of codon reassignment, the exclusion of competitors, is potentially generalizable to all 64 trinucleotide codons, and we expect that every codon can be reassigned in this manner. Indeed, we have completely reassigned about a third of the codons to date using various chemoenzymatic aminoacyl-tRNAs without finding any recalcitrant codons or any inactive tRNA anticodon mutants (Forster et al., unpublished data). Of the more than 10 unnatural L-amino acids tested in this system, all except the bulky biotinyl-lysine can essentially saturate incorporation at a sense codon using a $tRNA^{Asn}$-based transcript that is free of nucleoside modifications (Forster et al., unpublished data).

Notwithstanding these successes, truly modular codon reassignment for efficient synthesis of unnatural polymers has yet to be attained. For certain amino acids with unnatural backbones, single incorporations are poor in our purified system (Fig. 5; see also [15]). Syntheses requiring several adjacent incorporations of unnatural amino acids are inefficient, even with L-amino acids, either because multiple cycles multiply small losses that are difficult to detect in single cycles, or because

the presence of adjacent unnatural aminoacyl-tRNAs on the ribosome is problematic. Decoding specificities of unnatural aminoacyl-tRNAs can be unpredictable as, in controls lacking the cognate aminoacyl-tRNA for a codon, violation of the in vivo requirement of Watson–Crick base pairing in the first two codon positions and wobble pairing in the third codon position is possible (detailed in Note 11; Forster et al., unpublished data). Thus, we cannot yet program the ribosome to synthesize long defined peptidomimetic polymers at will. The removal of competitors and provision of much longer incorporation times have not completely cured the inefficiencies of unnatural amino acid incorporation associated with the crude or in vivo systems, implying that one source of the difficulty lies in the unnatural substrates themselves. Indeed, designing genetic codes de novo in a purified system will clearly be a useful method for investigating translation mechanism, as detailed below.

First, several of our results summarized above were unexpected, implying that understanding of substrate recognition by the translation apparatus is more rudimentary than originally thought.

Second, at least one of the five types of alterations in our unnatural substrates compared with the natural ones (Fig. 3) affects an important substrate determinant for translation. It is eminently feasible to use this highly modular system to dissect out the important feature(s) by making one alteration at a time, and then to fix the deficit(s) in future substrate designs (e.g., see the strategy for mischarging native tRNAs by Sisido et al., this volume), thereby improving translation engineering in general. Whatever the deficit, it is not specific to our $tRNA^{AsnB}$ sequence. Similar deficiencies in polymer synthesis are also observed using several in vitro-synthesized derivatives of $tRNA^{Phe}$ and $tRNA^{Ala}$ (Forster et al., unpublished data).

Third, because there are no aminoacyl-tRNAs or release factors to compete with our active unnatural substrates, the surprisingly low yields of polymer synthesis by our purified system suggest another competing reaction(s). This is not simply rapid hydrolysis of our chemoenzymatic substrates or mRNAs because incorporation is linear with time over tens of minutes, and ribonuclease activity is undetectable [7,10]. Consider the steps in elongation (Fig. 6) and their kinetics measured in purified systems [31,32]. Using wild-type substrates in vitro, elongation occurs at tens of amino acid incorporations per second [32], similar to the in vivo rate (with the rate of aminoacyl-tRNA docking in the A site much higher still due to the excess of non-cognate over cognate substrates). The only reactions that compete significantly with elongation are peptidyl-tRNA drop-off from both the A and P sites, both of which take more than a minute [31,32]. Thus, our system should provide an unnatural substrate the luxury of more than a thousand times longer than the time needed by a natural substrate to achieve quantitative incorporation.

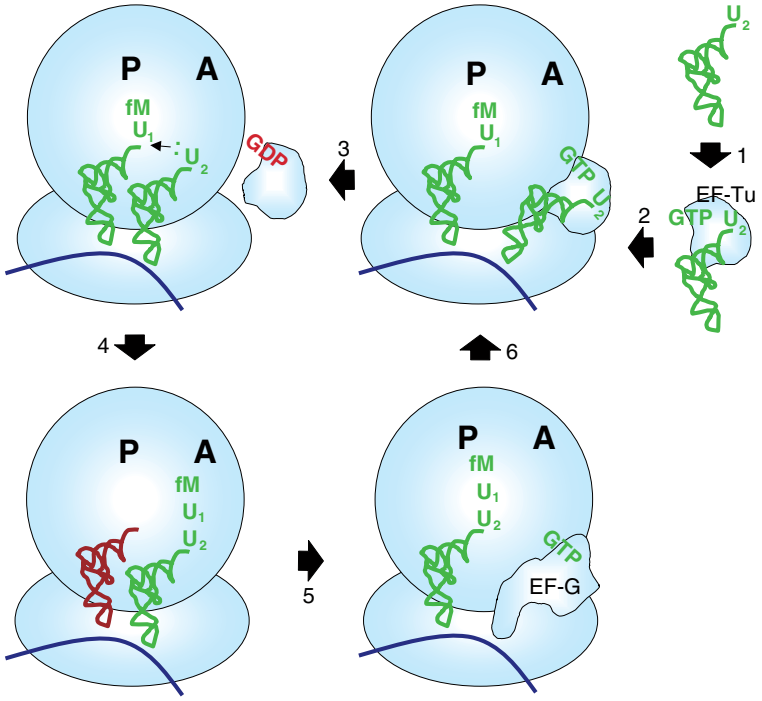

Fig. 6. Schematic illustrating steps in ribosomal elongation. U is an unnatural amino acid, and all of the tRNAs shown are unnatural. P, peptidyl site; A, aminoacyl site. For simplicity, the exit site and its bound tRNA on the left sides of the ribosomes are not shown. The positioning of the peptide is approximate and may represent either the hybrid or semi-classical model [38].

With respect to steps 1–3 (Fig. 6), mis-aminoacylation of tRNA can cause 100-fold weaker or tighter binding to EF-Tu than normal [33]. While such a perturbation should be compensated in steps 1 and 2 by our use of excess EF-Tu with L-aminoacyl-tRNAs, discrimination at step 3 [34] may be affected by either this perturbation or another unnatural feature of the substrate. Peptide bond formation is not expected to be slowed dramatically by changes limited to the amino acid side chain (top left of Fig. 5), but the amino group may be mis-positioned by an unnatural tRNA. If translocation is substantially slower, there would be more time for drop-off. Any problem might be compounded when there are two or three adjacent unmodified tRNAs bound to a single ribosome. Unmodified aminoacyl-tRNAs have higher rates of dissociation than their modified counterparts from ribosomal A or P sites, but the effect is only several-fold [4]. Drop-off could be examined by repeating our earlier experiments with a radiolabeled N-terminal amino acid instead of a C-terminal one [10]. Because long peptides bind strongly to the ribosome tunnel, thereby decreasing drop-off, it is possible that incorporation of unnatural amino acids could be improved if they were incorporated downstream of a leader sequence. For pure translation display, the leader sequence codons would be inserted at the 5′ end of the red sequence in Fig. 4.

Finally, it is interesting to compare pure translation display and our peptidomimetic evolution goal with other new technologies that accomplish related goals also by directed evolution. A pure protein-synthesis system (Shimizu et al., this volume) has been used to select for mRNA aptamers containing upstream mutations that confer binding to a downstream translation product [35]. Peptide ligands have been selected using ''DNA display,'' a method that ''translates'' DNA into polymers of potentially any small monomers by solid phase chemical syntheses involving multiple splitting and pooling steps [36] instead of ribosome-catalyzed translation in a single step. Variants of DNA display and pure translation display are possible where genotype and phenotype are not physically linked, but rather by compartments such as emulsions [37].

Acknowledgments

This work was supported by a grant from the National Institutes of Health (5 K08 CA80833 to A.C.F.), an Army Idea Award from the Department of Defense (DAMD17-00-1-0163 to S.C.B.), and a Sloan Fellowship (to V.W.C.). Z.T. is a Arun Guthikonda Memorial Graduate Fellow. S.C.B. is a Pew Scholar in the biomedical sciences and an Established Investigator of the American Heart Association.

References

[1] F.H.C. Crick, Symp. Soc. Exp. Biol. 12 (1958) 138–163.
[2] F. Chapeville, F. Lipmann, G.V. Ehrenstein, B. Weisblum, W.J. Ray, S. Benzer, Proc. Natl. Acad. Sci. USA 48 (1962) 1086–1092.
[3] V.W. Cornish, D. Mendel, P.G. Schultz, Angew. Chem. Int. Ed. Engl. 34 (1995) 621–633.
[4] R.P. Fahlman, T. Dale, O.C. Uhlenbeck, Mol. Cell 16 (2004) 799–805.
[5] A. Frankel, S.W. Millward, R.W. Roberts, Chem. Biol. 10 (2003) 1043–1050.
[6] S. Fahnestock, A. Rich, Nat. New Biol. 229 (1971) 8–10.
[7] A.C. Forster, H. Weissbach, S.C. Blacklow, Anal. Biochem. 297 (2001) 60–70.
[8] J.R. Roesser, M.S. Chorghade, S.M. Hecht, Biochemistry 25 (1986) 6361–6365.
[9] A.C. Forster, V.W. Cornish, S.C. Blacklow, Anal. Biochem. 333 (2004) 358–364.
[10] A.C. Forster, Z. Tan, M.N.L. Nalam, H. Lin, H. Qu, V.W. Cornish, S.C. Blacklow, Proc. Natl. Acad. Sci. USA 100 (2003) 6353–6357.
[11] Z. Tan, A.C. Forster, S.C. Blacklow, V.W. Cornish, J. Am. Chem. Soc. 126 (2004) 12752–12753.
[12] Y. Cenatiempo, N. Robakis, B.R. Reid, H. Weissbach, N. Brot, Arch. Biochem. Biophys. 218 (1982) 572–578.
[13] M.C. Ganoza, H. Aoki, N. Burkhardt, B.J. Murphy, Biochimie 78 (1996) 51–61.
[14] K. Stade, N. Junke, R. Brimacombe, Nucleic Acids Res. 23 (1995) 2371–2380.
[15] C. Merryman, R. Green, Chem. Biol. 11 (2004) 575–582.
[16] L.C. Mattheakis, R.R. Bhatt, W.J. Dower, Proc. Natl. Acad. Sci. USA 91 (1994) 9022–9026.
[17] Y.W. Hwang, A. Sanchez, M.C. Hwang, D.L. Miller, Arch. Biochem. Biophys. 348 (1997) 157–162.
[18] Y.P. Semenkov, M.V. Rodnina, W. Wintermeyer, Proc. Natl. Acad. Sci. USA 93 (1996) 12183–12188.
[19] H.F. Kung, C. Spears, T. Schulz, H. Weissbach, Arch. Biochem. Biophys. 162 (1974) 578–584.
[20] J.F. Milligan, O.C. Uhlenbeck, Methods Enzymol. 180 (1989) 51–62.
[21] Y. Shimizu, A. Inoue, Y. Tomari, T. Suzuki, T. Yokogawa, K. Nishikawa, T. Ueda, Nat. Biotechnol. 19 (2001) 751–755.
[22] J.H. Matthaei, H.P. Voigt, G. Heller, R. Neth, G. Schoch, H. Kubler, F. Amelunxen, G. Sander, A. Parmeggiani, Cold Spring Harb. Symp. Quant. Biol. 31 (1966) 25–38.
[23] U. Varshney, C.P. Lee, U.L. RajBhandary, J. Biol. Chem. 266 (1991) 24712–24718.
[24] S. Blanquet, P. Dessen, D. Kahn, Methods Enzymol. (1984) 141–152.
[25] J.D. Bain, C.G. Glabe, T.A. Dix, A.R. Chamberlin, E.S. Diala, J. Am. Chem. Soc. 111 (1989) 8013–8014.
[26] J. Ellman, D. Mendel, S. Anthony-Cahill, C.J. Noren, P.G. Schultz, Methods Enzymol. 202 (1991) 301–336.
[27] S. Peacock, H. Weissbach, H.A. Nash, Proc. Natl. Acad. Sci. USA 81 (1984) 6009–6013.
[28] H.F. Kung, F. Chu, P. Caldwell, C. Spears, B.V. Treadwell, B. Eskin, N. Brot, H. Weissbach, Arch. Biochem. Biophys. 187 (1978) 457–463.
[29] M.V. Rodnina, W. Wintermeyer, Proc. Natl. Acad. Sci. USA 92 (1995) 1945–1949.
[30] A.V. Zavialov, R.H. Buckingham, M. Ehrenberg, Cell 107 (2001) 115–124.
[31] R. Karimi, M. Ehrenberg, EMBO J. 15 (1996) 1149–1154.
[32] R. Karimi, M.Y. Pavlov, V. Heurgue-Hamard, R.H. Buckingham, M. Ehrenberg, J. Mol. Biol. 281 (1998) 241–252.

[33] T. Dale, L.E. Sanderson, O.C. Uhlenbeck, Biochemistry 43 (2004) 6159–6166.
[34] K.B. Gromadski, M.V. Rodnina, Mol. Cell 13 (2004) 191–200.
[35] B.W. Ying, T. Suzuki, Y. Shimizu, T. Ueda, J. Biochem. 133 (2003) 485–491.
[36] D.R. Halpin, P.B. Harbury, PLoS Biol. 2 (2004) 1022–1030.
[37] A. Aharoni, A.D. Griffiths, D.S. Tawfik, Curr. Opin. Chem. Biol. 9 (2005) 210–216.
[38] A.V. Zavialov, M. Ehrenberg, Cell 114 (2003) 113–122.