

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

Initiating protein synthesis with noncanonical monomers in vitro and in vivo

### Permalink

<https://escholarship.org/uc/item/1rp902wf>

### ISBN

9780128212530

### Authors

Tharp, Jeffery M  
Walker, Joshua A  
Söll, Dieter  
et al.

### Publication Date

2021

### DOI

10.1016/bs.mie.2021.05.002

Peer reviewed



Published in final edited form as:

*Methods Enzymol.* 2021 ; 656: 495–519. doi:10.1016/bs.mie.2021.05.002.

## Initiating protein synthesis with noncanonical monomers *in vitro* and *in vivo*

Jeffery M. Tharp<sup>#a</sup>, Joshua A. Walker<sup>#c</sup>, Dieter Söll<sup>a,b,\*</sup>, Alanna Schepartz<sup>c,d,\*</sup>

<sup>a</sup>Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, CT, United States

<sup>b</sup>Department of Chemistry, Yale University, New Haven, CT, United States

<sup>c</sup>Department of Chemistry, University of California, Berkeley, CA, United States

<sup>d</sup>Department of Molecular & Cell Biology, University of California, Berkeley, CA, United States

# These authors contributed equally to this work.

### Abstract

With few exceptions, ribosomal protein synthesis begins with methionine (or its derivative *N*-formyl-methionine) across all domains of life. The role of methionine as the initiating amino acid is dictated by the unique structure of its cognate tRNA known as tRNA<sup>fMet</sup>. By mis-acylating tRNA<sup>fMet</sup>, we and others have shown that protein synthesis can be initiated with a variety of canonical and noncanonical amino acids both *in vitro* and *in vivo*. Furthermore, because the  $\alpha$ -amine of the initiating amino acid is not required for peptide bond formation, translation can be initiated with a variety of structurally disparate carboxylic acids that bear little resemblance to traditional  $\alpha$ -amino acids. Herein, we provide a detailed protocol to initiate *in vitro* protein synthesis with substituted benzoic acid and 1,3-dicarbonyl compounds. These moieties are introduced at the N-terminus of peptides by mis-acylated tRNA<sup>fMet</sup>, prepared by flexizyme-catalyzed tRNA acylation. In addition, we describe a protocol to initiate *in vivo* protein synthesis with aromatic noncanonical amino acids (ncAAs). This method relies on an engineered chimeric initiator tRNA that is acylated with ncAAs by an orthogonal aminoacyl-tRNA synthetase. Together, these systems are useful platforms for producing N-terminally modified proteins and for engineering the protein synthesis machinery of *Escherichia coli* to accept additional nonproteinogenic carboxylic acid monomers.

## 1. Introduction

Ribosomal protein synthesis begins with the amino acid methionine across all domains of life. This unique ability of methionine to initiate protein synthesis is derived from the structure of its cognate tRNA, which exists in two forms: elongator and initiator methionine tRNA. The elongator tRNA (tRNA<sup>Met</sup>) exclusively installs methionine at elongating AUG codons; whereas, the initiator tRNA (tRNA<sup>fMet</sup>) exclusively installs methionine at AUG codons downstream of a Shine-Dalgarno sequence, thereby initiating synthesis of a new

\* Corresponding authors: dieter.soll@yale.edu; schepartz@berkeley.edu.

polypeptide chain (Fig. 1A and B) (Bhattacharyya & Varshney, 2016; Rodnina, 2018). Despite almost universal conservation in Nature, methionine is not the only amino acid that can initiate protein synthesis. Several decades ago, it was shown that mutants of tRNA<sup>fMet</sup>, that were mis-acylated with non-methionine amino acids, can initiate protein synthesis *in vivo* (Chattopadhyay, Pelka, & Schulman, 1990; Varshney & RajBhandary, 1990). More recently, it was shown that mis-acylated tRNA<sup>fMet</sup> can initiate *in vitro* translation with any of the 20 canonical amino acids, including both the D- and L-stereoisomers (Goto, Ohta, et al., 2008; Goto, Murakami, & Suga, 2008). Initiation is also not limited to the canonical amino acids. Using native and engineered variants of the methionyl-tRNA synthetase, tRNA<sup>fMet</sup> can be mis-acylated and initiate translation with numerous structural analogs of methionine *in vivo* (Fig. 2, 1–7) (Budisa et al., 1995; Ngo, Schuman, & Tirrell, 2013). Moreover, we have developed a chimeric variant of tRNA<sup>fMet</sup> that is acylated with noncanonical amino acids (ncAAs) by an orthogonal aminoacyl-tRNA synthetase (aaRS, *vide infra*). Using this system we have initiated translation with a variety of aromatic ncAAs (Fig. 2, 8–13) in *Escherichia coli* to produce proteins containing these unnatural moieties at their N-terminus (Tharp et al., 2020; Tharp, Vargas-Rodriguez, Schepartz, & Söll, 2021).

A unique feature of the initiating (N-terminal) amino acid is that it is the only residue whose  $\alpha$ -amine is not involved in peptide bond formation. Therefore,  $N^\alpha$ -modified amino acids can be installed into proteins at the initiating position. For example, in prokaryotes translation initiates with the  $N^\alpha$ -modified variant of methionine, *N*-formyl-L-methionine (Fig. 1A) (Bhattacharyya & Varshney, 2016). Similarly, using synthetically mis-acylated tRNA<sup>fMet</sup>, translation has been initiated with various  $N^\alpha$ -acyl and  $N^\alpha$ -alkyl amino acids *in vitro* (Fig. 2, 14–17) (Goto, Ohta, et al., 2008; Tadayoshi et al., 2010). In addition to these  $N^\alpha$ -modified amino acids, synthetically mis-acylated initiator tRNAs have been used to initiate translation with structurally disparate non- $\alpha$ -amino acids that bear little resemblance to traditional proteinogenic amino acids. These include, for example, long-chain carboxylic acids, substituted benzoic acids, 1,3-dicarbonyls, peptides, and helical aromatic oligomer-peptide hybrids (Fig. 2, 18–23) (Ad et al., 2019; Goto & Suga, 2009; Lee, Schwarz, Kim, Moore, & Jewett, 2020; Rogers et al., 2018). Taken together, these studies demonstrate exceptional versatility of wild type ribosomes to catalyze peptide bond formation with diverse carboxylic acids at the initiating position.

Some non- $\alpha$ -amino acids that have been introduced into proteins at the initiating position are precursors of polyamide polymers. For example, substituted benzoic acids (*e.g.*, 18) are precursors of aramids; a class of aromatic polyamides with diverse properties. Similarly, long-chain amino acids (*e.g.*, 22) are monomeric precursors of Nylon. Using wild type ribosomes, these molecules are primarily limited to incorporation at the initiating position; however, we hypothesize that mutant ribosomes could be engineered to catalyze peptide bond formation between consecutive polyamide monomers. Such ribosomes would enable DNA-encoded synthesis of sequence-defined chemical polymers with unprecedented properties. Furthermore, we envision that platforms that enable noncanonical initiation can provide a starting point to engineer translational components (ribosomes, aaRSs, elongation factors, *etc.*) toward this end. In this chapter, we provide detailed methods for initiating *in vitro* translation with substituted benzoic acid and 1,3-dicarbonyl monomers *via* flexizyme-

catalyzed acylation of tRNA<sup>fMet</sup>. In addition, we also describe methods for applying orthogonal aaRSs to initiate translation with various aromatic ncAAs *in vivo*.

## 2. Initiating translation with substituted benzoic acid and 1,3-dicarbonyl monomers *in vitro*

*In vivo* ribosomal translation of chemical polymers is the holy grail of polymer science. However, biosynthesis in living cells is multifaceted and requires each monomer to be cell-permeable (bioavailable), as well as compatible with an aaRS and the ribosomal translation machinery (Chin, 2017). In this way, *in vivo* experiments can hinder our ability to evaluate the capacity of ribosomal translation to tolerate new chemical moieties. Therefore, *in vitro* techniques that enable direct interrogation of ribosomal translation are a powerful complement to *in vivo* biosynthesis efforts. *In vitro* translation (IVT) systems, such as the commercially available *Protein Synthesis Using Purified Recombinant Elements* (PURE) system and extract-based systems, are cell-free reconstitutions of all of the necessary components for ribosomal translation (Hammerling, Krüger, & Jewett, 2020; Shimizu et al., 2001). The absence of a cell membrane enables researchers to circumvent issues related to monomer bioavailability and to make user-defined adjustments to reaction conditions (translation factors, template concentration, mis-acylated tRNA concentration, *etc.*). A complementary technique to IVT is flexizyme-mediated tRNA acylation. Flexizymes are aaRS-like ribozymes designed to charge diverse non-natural amino acids onto a variety of tRNAs (Lee, Bessho, Wei, Szostak, & Suga, 2000; Murakami, Ohta, Ashigai, & Suga, 2006; Murakami, Saito, & Suga, 2003; Niwa, Yamagishi, Murakami, & Suga, 2009; Ramaswamy, Saito, Murakami, Shiba, & Suga, 2004). We and others have shown that flexizymes are capable of mis-acylating tRNA with diverse unnatural monomers such as  $\alpha$ -hydroxy, *N*-alkyl-amino, D- $\alpha$ -amino,  $\alpha$ -thio, and  $\alpha$ -amino(carbothio) acids,  $\beta$ -amino acids, long chain- and cyclic-amino acids, aramids, and 1,3-dicarbonyls (Ad et al., 2019; Katoh & Suga, 2018, 2020a, 2020b; Katoh, Tajima, & Suga, 2017; Kawakami, Murakami, & Suga, 2008; Lee et al., 2019, 2020; Maini et al., 2019; Ohta, Murakami, Higashimura, & Suga, 2007; Takatsuji et al., 2019). This eliminates the need to engineer an aaRS for each monomer of interest (Chin, 2017). Using flexizyme-mediated tRNA mis-acylation together with IVT, we investigated the tolerance of wild type *E. coli* ribosomes for diverse aramids and 1,3-dicarbonyls (Fig. 3) (Ad et al., 2019). We found that wild type *E. coli* ribosomes accepted tRNA<sup>fMet</sup> mis-acylated with both aramids and 1,3-dicarbonyls. These substrates were elongated to create diverse aramid-peptide and polyketide-peptide hybrid molecules. This work expanded the scope of reactions catalyzed by flexizymes and wild type ribosomes as well as established a starting point for *in vivo* translation of these novel materials.

### 2.1 Protocol

#### 2.1.1 Equipment

1. Analytical balance
2. Vortex mixer
3. Centrifuge

4. Thermocycler
5. PAGE gel casting apparatus
6. PAGE gel electrophoresis chamber and power supply
7. Gel imaging system
8. Liquid chromatography-mass spectrometer (LC-MS)

### 2.1.2 Materials—*DNA Oligonucleotide Sequences* (Integrated DNA Technologies)

|   |                 |   |
|---|-----------------|---|
| 1 | <i>MetF-F</i>   | ATTCTGCAGTAATACGACTCACTATACGC<br>GGGGTGGAGCAGCTGGTAGC<br>TCGTCGGGCTCATA   |
| 2 | <i>MetR-R*</i>  | TmGGTTGCGGGGGCCGGATTGAACCGACG<br>ACCTTCGGGTTATGAGCCCGA<br>CGAGCTA<br><i>*mG represents 2' O-methyl-deoxymethylguanosine</i> |
| 3 | <i>MVFlag-1</i> | TAATACGACTCACTATAGGGTAACTTTAA<br>CAAGGAGAAAAACATGGTATTTG<br>ACTACAAGG   |
| 4 | <i>MVFlag-2</i> | CGAAGCTTACTGTGTCGTCGTCCTTTGTA<br>GTCAAATACCATGTTTTTCCT<br>TGTTAAAG  |
| 5 | <i>MVFlag-3</i> | GCGAATTAATACGACTCACTATAGGGTTAA<br>CTTTAACA  |
| 6 | <i>MVFlag-4</i> | AAACCCCTCCGTTTAGAGAGGGGTTATGCT<br>AGTTACTTGTGTCGTCGTCCTC<br>TTG   |

### *RNA Oligonucleotide Sequences* (Integrated DNA Technologies)

|   |                            |  |
|---|----------------------------|--|
| 1 | <i>Microhelix (MH) RNA</i> | GGCUCUGUUCGCAGAGCCGCCA                           |
| 2 | <i>aFx</i>                 | GGUAGGAAAUGUUCCGCACCCCGAAAGG<br>GGUAAGGGCGUCAGGU |
| 3 | <i>dFx</i>                 | GGAUCGAAAGAUUCCGCAUCCCGAAAGG<br>GUACAUGGCGUUAGGU |
| 4 | <i>eFx</i>                 | GGAUCGAAAGAUUCCGCGGCCCGAAAGG<br>GGAUUAGCGUUAGGU  |

#### 2.1.2.1 Reagents

1. RNase-free water (Invitrogen, Catalog # 10977-015)
2. HEPES Buffer: 500 mM HEPES, pH 7.5
3. Bicine Buffer: 500 mM Bicine, pH 9.0
4. Magnesium chloride, 1 M
5. Acrylamide/bis-acrylamide (19:1 crosslinker), 40% (w/v) (Bio-Rad, Catalog # 1610144)

6. Urea, 72% (w/v)
7. Sodium Acetate Buffer: 3 M sodium acetate, pH 5.2 (AmericanBio, Catalog # AB13168–01000)
8. Ammonium persulfate, 10 mgmL<sup>-1</sup>
9. Tetramethylethylenediamine (TEMED)
10. RNA Loading Buffer: 95% (v/v) formamide, 150 mM sodium acetate pH 5.2, 10 mM EDTA, and 200 µg mL<sup>-1</sup> bromophenol blue
11. TBE Buffer: 10.8 gL<sup>-1</sup> tris, 5.5 gL<sup>-1</sup> boric acid, 10 mM EDTA, pH 8.0
12. Ethidium bromide
13. RNase A: 1.5 U µL<sup>-1</sup> in 200 mM sodium acetate, pH 5.2 (Millipore Sigma, Catalog # 10109142001)
14. Trichloroacetic acid, 50% (w/v)
15. Basic phenol solution (10 mM Tris, 1 mM EDTA, pH 8.0) (Millipore Sigma, Catalog # P4557–100 ML)
16. Acidic phenol solution (saturated with citric acid, pH 4.5) (Millipore Sigma, Catalog # 6702–400 ML)
17. Chloroform
18. Ethanol, 200 proof (AmericanBio, Catalog # AB00515–00500)
19. RNase-free DNase I (Fisher Scientific, Catalog # EN0521)
20. L-methionine, 33 mM
21. L-valine, 33 mM
22. L-tyrosine, L-phenylalanine, and L-lysine stock, 33 mM each
23. L-aspartic acid, 7 mM (adjusted to pH 7.0)
24. Ni-NTA Agarose (Qiagen, Catalog # 30210)

#### 2.1.2.2 Kits and consumables

1. Hiscribe™T7 Quick High Yield RNA Synthesis Kit (New England Biolabs, Catalog # E2050S)
2. PURExpress® ( tRNA, aa) *In Vitro* Translation (IVT) Kit (New England Biolabs, Catalog # E6840)
3. RNase-free Micro Bio-spin™ P-30 Columns, Tris Buffer (Bio-Rad, Catalog # 7326250)

#### 2.1.3 Selection of flexizyme and synthesis of flexizyme-compatible

**monomers**—Flexizymes are aaRS-like ribozymes that recognize various tRNAs *via* base pairing interactions between the 3'-end CCA of the tRNA and the 3'-end GGU of flexizyme

(Morimoto, Hayashi, Iwasaki, & Suga, 2011). Originally, flexizyme (Fx) was designed to recognize aromatic amino acids that contain an activated cyanomethyl ester (CME) leaving group (Lee et al., 2000). Later, efforts to create flexizyme variants that tolerate a variety of amino acids yielded dFx and eFx, which recognize their substrates *via* the aromatic leaving groups 3,5-dinitrobenzyl ester (DNB) and *p*-chlorobenzyl thioester (CBT) respectively (Murakami et al., 2006). A third flexizyme was later developed to aid in the acylation of poorly soluble monomers (Niwa et al., 2009). This flexizyme, aFx, recognizes its substrates *via* the hydrophilic leaving group (2-aminoethyl) amidocarboxybenzyl thioester (ABT). The structure of flexizyme leaving groups and the primary sequence of flexizyme variants are provided in Fig. 4.

Flexizymes are commonly synthesized by *in vitro* transcription; however, in our studies, we exclusively use flexizymes synthesized (*via* solid-phase techniques) by Integrated DNA Technologies. Standard “desalting” purification removes small truncation products and organic contaminants such that there is no noticeable impact on flexizyme activity. The synthesis of flexizyme compatible substrates is unique to each monomer of interest. Within the work presented here we synthesized greater than 20 flexizyme compatible substrates. Thus, a detailed description of the synthetic methodology is beyond the scope of this protocol. However, our lab and others have previously reported the synthesis of dozens of flexizyme compatible monomers. (Ad et al., 2019; Katoh & Suga, 2018, 2020a, 2020b; Katoh et al., 2017; Kawakami et al., 2008; Lee et al., 2020; Lee et al., 2019; Maini et al., 2019; Ohta et al., 2007; Takatsuji et al., 2019).

Identification of the optimal flexizyme variant and leaving group for charging tRNA with a given monomer is somewhat empirical. Generally, eFx and the leaving group CME may be used with monomers that contain aromatic side chains or backbones. However, monomers devoid of an aromatic moiety require the leaving groups ABT, DNB, or CBT along with the corresponding flexizymes aFx, dFx, and eFx.

**2.1.4 Optimizing conditions for flexizyme-catalyzed acylation**—Microhelix (MH) RNA, a 22-nucleotide surrogate for the acceptor stem of full-length tRNA, is often used to optimize conditions for flexizyme-mediated tRNA acylation. The small size of MH RNA makes it accessible *via* solid-phase synthesis and enables characterization of acylation *via* acid-urea PAGE gel-shift analysis. In addition to gel-shift analysis we adapted a highly sensitive RNase A/LC-MS assay to directly observe the 2' and 3' acyl-adenosine of the acylated MH RNA (McMurry & Chang, 2017). A summary of MH RNA acylation optimization is provided in Fig. 5.

#### **2.1.4.1 Flexizyme-catalyzed acylation of microhelix (MH) RNA**

1. Mix 1  $\mu$ L of appropriate flexizyme (250  $\mu$ M stock in RNase-free water), 1  $\mu$ L of MH RNA (250  $\mu$ M stock in RNase-free water), and 1  $\mu$ L of buffer (500 mM HEPES pH 7.5 or 500 mM Bicine pH 9.0)
2. Incubate sample at 95°C for 2 min
3. Allow sample to cool at room temperature for 5 min

4. Add 6  $\mu\text{L}$  of magnesium chloride (1 M stock in RNase-free water).
5. Add 1  $\mu\text{L}$  of appropriate monomer (50 mM stock in dimethyl sulfoxide)
6. Pipette mix and incubate at 4°C for desired amount of time.

*Note: Typically, incubation for 48h is sufficient to observe acylation.*

#### 2.1.4.2 Acid-urea PAGE gel-shift analysis

1. Prepare unpolymerized acid-urea PAGE gel mixture as follows: 5 mL of 40% (w/v) acrylamide/bis-acrylamide (19:1), 5 mL of 72% (w/v) urea, and 167  $\mu\text{L}$  of 3 M sodium acetate (pH 5.2).
2. To initiate polymerization, add 100  $\mu\text{L}$  of 10 mg  $\mu\text{L}^{-1}$  ammonium persulfate (freshly prepared) and 8  $\mu\text{L}$  of TEMED. The final gel composition is 20% acrylamide/bis-acrylamide, 36% (w/v) urea, 50 mM sodium acetate (pH 5.2), 0.1% (w/v) ammonium persulfate, and 0.08% (v/v) TEMED
3. Quickly mix and transfer acid-urea PAGE gel mixture to an appropriate gel casting apparatus. Reserve a small amount of unpolymerized gel mixture to monitor the polymerization status of the PAGE gel
4. Allow the gel to polymerize until the reserved portion has solidified—typically approximately 30 min
5. Add 1  $\mu\text{L}$  of acylation reaction mixture to 1  $\mu\text{L}$  of denaturing RNA loading buffer (95% (v/v) formamide, 150 mM sodium acetate (pH 5.2), 10 mM EDTA, and 200  $\mu\text{g}$   $\mu\text{L}^{-1}$  bromophenol blue).
6. Run the denatured sample on an acid-urea PAGE gel with 50 mM sodium acetate (pH 5.2) as the running buffer at 120V until the dye has run to the bottom of the gel—approximately 3–4h. *Note: It is advisable to run acylation reaction samples against MH RNA and flexizyme controls. These controls can help detect minor changes in sample migration*
7. Stain the gel with ethidium bromide (1  $\mu\text{g}$   $\text{mL}^{-1}$  in TBE Buffer) for 2 min. *Note: MH RNA may exhibit poor staining with other detection reagents. In our hands, ethidium bromide has provided the strongest signal for these samples*
8. Destain the gel with TBE Buffer for 1 min
9. Image with an appropriate gel imaging system and quantify UV densitometry using ImageJ

#### 2.1.4.3 RNase A/LC-MS analysis

1. Add 1  $\mu\text{L}$  of acylation reaction mixture to 1.1  $\mu\text{L}$  of RNase A (1.5 U  $\mu\text{L}^{-1}$  in 200 mM sodium acetate, pH 5.2) and incubate at room temperature for 5 min
2. Precipitate RNase A by adding 0.25  $\mu\text{L}$  of 50% (w/v) trichloroacetic acid and incubating the mixture at room temperature for 5 min
3. Remove insoluble material by centrifugation (20,000  $\times g$ ) for 2 min at 4°C



4. Dilute 1  $\mu\text{L}$  of the sample with water to 20  $\mu\text{L}$  and freeze by incubation at  $-80^{\circ}\text{C}$  for 5 min
5. Before analysis, thaw the sample and remove insoluble material by centrifugation at  $(20,000 \times g)$  for 10 min at  $4^{\circ}\text{C}$
6. Analyze the sample *via* LC-MS to detect the exact mass of the 2' and 3' acyl-adenosine

#### 2.1.5 Preparation of mis-acylated tRNA<sup>fMet</sup>

##### 2.1.5.1 *In vitro* transcription and purification of tRNA<sup>fMet</sup>

1. Prepare DNA template for transcribing *E. coli* tRNA<sup>fMet</sup> using polymerase chain reaction (PCR) by annealing and extending the oligonucleotides *MetT-F* and *MetT-R*. The oligonucleotides are mixed 1:1 at 10  $\mu\text{M}$ . *MetT-R* contains a 2'-*O*-methyl-deoxymethylguanosine at the penultimate nucleotide to prevent run-off transcription
2. Extract the extended template 1 $\times$  with 1 volume of a 1:1 (v/v) basic phenol (pH 8.0)/chloroform mixture
3. Precipitate the extracted template with 3 volumes of 95% (v/v) ethanol
4. Pellet the precipitated template by centrifugation  $(20,000 \times g)$  for 30 min at  $4^{\circ}\text{C}$
5. Remove the supernatant and allow the pellet to dry at room temperature for 15 min
6. Suspend the purified template in RNase-free water and quantify by measuring the absorbance at 260 nm
7. Set up *in vitro* transcription of tRNA<sup>fMet</sup> using HiScribe<sup>TM</sup> T7 Quick High Yield RNA Synthesis Kit from NEB using the manufacturer's protocol for short transcripts at a total reaction scale of 200  $\mu\text{L}$  (6.7 mM ATP, GTP, UTP, and CTP and 50 ng  $\mu\text{L}^{-1}$  DNA template).
8. Incubate *in vitro* transcription reaction at  $37^{\circ}\text{C}$  for 6h
9. Treat reaction with 10U of RNase-free DNase at  $37^{\circ}\text{C}$  for 2h
10. Add sodium acetate buffer, pH 5.2 to a final concentration of 200 mM
11. Extract the reaction 1 $\times$  with one volume of a 1:1 (v/v) acidic phenol (pH 4.5)/chloroform mixture and 2 $\times$  with one volume of chloroform
12. Precipitate the extracted product with 3 volumes of 95% (v/v) ethanol. If necessary, incubate at  $-80^{\circ}\text{C}$  for 1h or  $-20^{\circ}\text{C}$  overnight to encourage precipitation
13. Pellet the precipitated tRNA<sup>fMet</sup> by centrifugation  $(20,000 \times g)$  for 30 min at  $4^{\circ}\text{C}$
14. Remove the supernatant and wash the pellet 2 $\times$  with ice cold 70% (v/v) ethanol. Centrifuge  $(20,000 \times g)$  for 5 min at  $4^{\circ}\text{C}$
15. Allow the pellet to dry at room temperature for 15 min

16. Suspend the product in RNase-free water and remove excess nucleotides using RNase-free Micro Bio-Spin<sup>®</sup> P-30 columns, Tris Buffer from Bio-Rad using the manufacturer's protocol
17. Quantify the recovery of tRNA<sup>fMet</sup> by measuring the absorbance at 260 nm
18. Store purified tRNA<sup>fMet</sup> at  $-80^{\circ}\text{C}$  for optimal stability. For short-term (<1 month),  $-20^{\circ}\text{C}$  can be used for storage

#### 2.1.5.2 Flexizyme-catalyzed Mis-acylation of tRNA<sup>fMet</sup>

1. Mix 10  $\mu\text{L}$  of appropriate flexizyme (250  $\mu\text{M}$  stock in RNase-free water), 10  $\mu\text{L}$  of tRNA<sup>fMet</sup> (250  $\mu\text{M}$  stock in RNase-free water), and 10  $\mu\text{L}$  of buffer (500 mM HEPES pH 7.5 or 500 mM Bicine pH 9.0).
2. Incubate sample at  $95^{\circ}\text{C}$  for 2 min
3. Allow sample to cool at room temperature for 5 min
4. Add 60  $\mu\text{L}$  of magnesium chloride (1 M stock in RNase-free water).
5. Add 10  $\mu\text{L}$  of appropriate monomer (50 mM stock in dimethyl sulfoxide)
6. Pipette mix and incubate at  $4^{\circ}\text{C}$  for amount of time determined by MH RNA optimization experiments
7. If desired, monitor the reaction by removing a 1  $\mu\text{L}$  aliquot of the tRNA<sup>fMet</sup> acylation reaction mixture for analysis *via* the RNase A digestion assay described above
8. Add sodium acetate buffer (pH 5.2) to a final concentration of 300 mM and precipitate the mis-acylated tRNA<sup>fMet</sup> by adding ethanol to a final concentration of 70% (v/v).
9. To encourage precipitation, incubate sample at  $-80^{\circ}\text{C}$  for 1h
10. Pellet the mis-acylated tRNA<sup>fMet</sup> by centrifugation ( $20,000 \times g$ ) for 30 min at  $4^{\circ}\text{C}$
11. Wash the pellet 1 $\times$  with 500  $\mu\text{L}$  of ice cold 70% (v/v) ethanol. Centrifuge ( $20,000 \times g$ ) for 5 min at  $4^{\circ}\text{C}$
12. Allow the pellet to dry at room temperature for 5 min
13. If mis-acylated tRNA<sup>fMet</sup> will be used immediately, suspend the pellet in 1 mM sodium acetate (pH 5.2). Otherwise, store the dry pellet at  $-80^{\circ}\text{C}$  for later use

**2.1.6 In vitro translation reactions**—The commercially available *Protein Synthesis Using Purified Recombinant Elements* (PURE) system is an *in vitro* assembly of translation factors, tRNAs, RNAP polymerase, and ribosomes that have been individually purified and supplemented with amino acids, nucleotides, and co-factors (Hammerling et al., 2020; Shimizu et al., 2001). Supplementing this mixture with mis-acylated tRNA enables direct interrogation of ribosomal translation. A schematic overview of *in vitro* translation (IVT)

with the PURE system and subsequent analysis of translation products *via* LC-MS is provided in Fig. 6.

#### 2.1.6.1 Synthesis of MVFDYKDDDDK (MVF-FLAG) template DNA

1. Prepare DNA template for expression of MVFDYKDDDDK by annealing and extending the oligonucleotides *MVFflag-1* and *MVFflag-2*. For extension, the oligonucleotides are mixed 1:1 at 10  $\mu$ M
2. Extract the extended template 1 $\times$  with 1 volume of a 1:1 (v/v) basic phenol (pH8.0)/chloroform mixture
3. Precipitate the extracted template with 3 volumes of 95% (v/v) ethanol
4. Pellet the precipitated template by centrifugation (20,000  $\times$  *g*) for 30 min at 4°C
5. Remove the supernatant and allow the pellet to dry at room temperature for 15 min
6. Suspend the purified template in RNase-free water and quantify by measuring the absorbance at 260 nm
7. Amplify the extension product of *MVFflag-1* and *MVFflag-2* by PCR using primers *MVFflag-3* and *MVFflag-4*
8. Extract the extended template 1 $\times$  with 1 volume of a 1:1 (v/v) basic phenol (pH 8.0)/chloroform mixture
9. Precipitate the extracted template with 3 volumes of 95% (v/v) ethanol
10. Pellet the precipitated template by centrifugation (20,000  $\times$  *g*) for 30 min at 4°C
11. Wash the pellet 2 $\times$  with 500  $\mu$ L 70% (v/v) ethanol. Centrifuge (20,000  $\times$  *g*) for 5 min at 4°C
12. Remove the supernatant and allow the pellet to dry at room temperature for 15 min
13. Suspend the purified template in RNase-free water and quantify by measuring the absorbance at 260 nm

#### 2.1.6.2 *In vitro* translation with Mis-acylated tRNA<sup>fMet</sup>

1. The model FLAG-tagged peptide MVFDYKDDDDK (MVF-FLAG) was expressed using the PureExpress® ( tRNA, aa) kit by NEB with some modifications
2. To initiate ribosomal translation with unnatural monomers, combine the following:
  - a. 5 $\mu$ L of solution A ( tRNA, aa)
  - b. 0.25  $\mu$ L of a valine stock solution at 33 mM
  - c. 0.25  $\mu$ L of a stock solution containing 33 mM tyrosine, 33 mM phenylalanine, and 33 mM lysine

- d. 1  $\mu$ L of a stock solution containing 7 mM aspartic acid, pH 7.0
  - e. 2.5  $\mu$ L of tRNA solution
  - f. 7.5  $\mu$ L of solution B
  - g. 500–1000ng of MVF-FLAG DNA template
  - h. 1.25–2.5 nmols of mis-acylated tRNA<sup>fMet</sup> freshly prepared in 1 mM sodium acetate, pH 5.2
  - i. RNase-free water to bring the total reaction volume to 25  $\mu$ L
3. To perform a control expression of the MVF-FLAG template, omit the mis-acylated tRNA<sup>fMet</sup> and add 0.25  $\mu$ L of a methionine stock solution at 33 mM
  4. To demonstrate that the translation product is template-dependent, omit the MVF-FLAG DNA template
  5. Incubate reactions for 6h at 37°C
  6. Quench reactions by placing them on ice and adding 25  $\mu$ L of dilution buffer (10 mM magnesium acetate and 100 mM sodium chloride).
  7. To remove most proteins and nucleic acids, add 5  $\mu$ L of Ni-NTA slurry and incubate on ice for 1h
  8. To remove the Ni-NTA resin, centrifuge the sample (20,000  $\times g$ ) for 10 min at 4°C
  9. Freeze the supernatant at –80°C for 5 min
  10. Remove insoluble material by centrifugation (20,000  $\times g$ ) for 10 min at 4°C
  11. Analyze the supernatant *via* LC-MS to detect the exact mass of the translation product

### 3. Initiating translation with noncanonical amino acids *in vivo*

*In vitro*, mis-acylated tRNA<sup>fMet</sup> can be prepared using various chemical acylation strategies, or by flexizyme-catalyzed acylation. These complementary methods are highly adaptable and can be used to produce tRNA<sup>fMet</sup> acylated with virtually any desired acid. However, these methods are not compatible with protein synthesis *in vivo*. In living cells, tRNAs are acylated by the aaRSs, an ancient class of enzymes that have evolved high efficiency and fidelity toward recognition of a single amino acid substrate. Genetic code expansion (GCE) is a technique that uses orthogonal aaRS and tRNA pairs to co-translationally install ncAAs into proteins, typically in response to the amber (UAG) nonsense codon (Fig. 7A) (Chin, 2017; Young & Schultz, 2018). One aaRS•tRNA pair that is frequently used for this purpose is the tyrosyl-tRNA synthetase (MjTyrRS) and tRNA<sup>Tyr</sup> pair from *Methanocaldococcus jannaschii*. Over the past two decades, the MjTyrRS has been engineered to recognize dozens of aromatic ncAAs to facilitate their co-translational incorporation into protein in *E. coli* (Dumas, Lercher, Spicer, & Davis, 2015). The MjTyrRS recognizes its cognate tyrosine tRNA using a small set of identity elements including the discriminator base (A73) and the

first base pair (C1–G72) (Fechter, Rudinger-Thirion, Tukalo, & Giegé, 2001; Sakamoto & Hayashi, 2019). By transplanting these nucleotides into *E. coli* tRNA<sup>fMet</sup> we developed a chimeric tRNA (*tRNA<sup>Ty2</sup>*) that is a substrate for MjTyrRS-catalyzed aminoacylation with ncAAs (Fig. 7B). Because this chimeric tRNA was designed using tRNA<sup>fMet</sup> as a scaffold, once acylated by MjTyrRS, *tRNA<sup>Ty2</sup>* can initiate translation with ncAAs at UAG codons downstream of a Shine-Dalgarno sequence (Fig. 7C).

Using *tRNA<sup>Ty2</sup>*, together with engineered MjTyrRS variants, we have demonstrated efficient ncAA initiation of several genes. These include the common reporter genes *chloramphenicol acetyltransferase*, *superfolder green fluorescent protein*, and the bacterial *luciferase* gene. In addition, we have found that this system is compatible with several commonly employed laboratory strains of *E. coli*. However, while *tRNA<sup>Ty2</sup>* can be used to initiate translation with ncAAs in wild type strains, the expression yield is quite low. This is likely due to competition with endogenous tRNA<sup>fMet</sup> for binding to ribosomal P-sites (Kapoor, Das, & Varshney, 2011). To overcome this limitation, we developed a strain of *E. coli* in which three of the four genomic copies of tRNA<sup>fMet</sup> were deleted from the genome (strain DH10B *metZ WV*). Studies have shown that *metZ WV* strains show improved initiation with mutant initiator tRNAs (Kapoor et al., 2011; Samhita, Virumäe, Remme, & Varshney, 2013). Using strain DH10B *metZ WV* initiation with ncAAs is at least as efficient as elongation with the same ncAA.

### 3.1 Protocol

#### 3.1.1 Equipment

1. Stationary and shaking incubators
2. Water bath
3. Cell density meter
4. Analytical balance
5. 2 × 1L shaker flasks
6. Petri plates (100 mm × 20 mm)
7. Culture tubes (14 mL; 17 mm × 100 mm)
8. Polypropylene gravity flow columns (*e.g.*, Bio-Rad Catalog # 7311550)
9. Centrifuge
10. Autoclave
11. Mass spectrometer

#### 3.1.2 Materials

1. Plasmid pBAD sfGFP[1UAG] *tRNA<sup>Ty2</sup>* (Tharp et al., 2020) (Plasmids are available from the corresponding authors upon request)
2. Plasmid pMW pCNFRS (Tharp et al., 2020)
3. 2 × YT media

4. LB agar
5. Ampicillin
6. Spectinomycin
7. *para*-acetyl-L-phenylalanine (**8**, pAcF, Chem-Impex, Catalog # 24756) or desired ncAA
8. Glycerol solution, 10% v/v (autoclaved)
9. 50× salts solution (autoclaved) (Hammill, Miyake-Stoner, Hazen, Jackson, & Mehl, 2007)
10. MgSO<sub>4</sub>, 1 M (filtered)
11. L-aspartate solution, 5% w/v (pH adjusted to 7.5 and filtered)
12. L-leucine solution, 30 mM (autoclaved)
13. 25× amino acid solution (filtered) (Tharp et al., 2020)
14. 5000× Trace metal solution (filtered) (Hammill et al., 2007)
15. Autoclaved MilliQ water
16. Isopropyl β-D-1 thiogalactopyranoside, 1 M (filtered)
17. L-arabinose, 20% w/v (filtered)
18. 10× BugBuster<sup>®</sup> protein extraction reagent (MilliporeSigma, Catalog # 70921)
19. Benzonase<sup>®</sup> nuclease (MilliporeSigma, Catalog # E1014–25KU)
20. Lysozyme
21. Talon<sup>®</sup> metal affinity resin (Takara Bio, Catalog # 635501)
22. IMAC Buffer A: 50 mM Tris pH 8.0, 500 mM NaCl, 10 mM Imidazole
23. IMAC Buffer B: 50 mM Tris pH 8.0, 300 mM NaCl, 250 mM Imidazole

### 3.1.3 Expression and purification of sfGFP-1pAcF

#### 3.1.3.1 Transformation

1. Prepare LB agar plates containing 100 μg mL<sup>-1</sup> ampicillin and 50 μg mL<sup>-1</sup> spectinomycin
2. Co-transform DH10B *metZ<sub>WV</sub>* by adding 2 μL of the plasmid pBAD sfGFP[1UAG] *tRNA<sup>Ty2</sup>* (~100 ng μL<sup>-1</sup>) and 3 μL of plasmid pMW pCNFRS (~20 ng μL<sup>-1</sup>) to a 100 μL aliquot of chemically competent cells in a 1.5 mL Eppendorf<sup>™</sup> tube. Gently flick the tube to mix and incubate the cell/plasmid mixture on ice for 10 min
3. Transfer the tube containing the cell/plasmid mixture to a 42°C water bath and incubate for 45s. Place the tube back on ice and incubate an additional 2 min

4. Transfer the cell/plasmid mixture to a culture tube containing 900  $\mu\text{L}$  of  $2 \times \text{YT}$  (without antibiotics) and incubate in a shaking incubator at  $37^\circ\text{C}$  for 45 min
5. Pipette 100  $\mu\text{L}$  of the culture onto the surface of a dry LB agar plate (containing ampicillin and spectinomycin). Using a cell spreader or glass beads, evenly spread the solution across the entire surface of the agar plate until the liquid is absorbed. Incubate the plate upside down, overnight at  $37^\circ\text{C}$

### 3.1.3.2 Protein expression

1. The day before protein expression, use a sterile toothpick or disposable pipette tip to pick a single colony of co-transformed DH10B *metZWV*. Grow the cells overnight at  $37^\circ\text{C}$  in 10 mL  $2 \times \text{YT}$  supplemented with  $100 \mu\text{g mL}^{-1}$  ampicillin and  $50 \mu\text{g mL}^{-1}$  spectinomycin
2. Prepare 200 mL of chemically defined media (CDM) by mixing the following components in an autoclaved 1 L shaker flask: 20 mL glycerol (10%, v/v), 4 mL  $50\times$  salts, 400  $\mu\text{L}$   $\text{MgSO}_4$  (1 M), 10 mL L-aspartate (5% w/v), 4 mL L-leucine (30 mM), 8 mL amino acid solution ( $25\times$ ), 40  $\mu\text{L}$  trace metals ( $5000\times$ ), and autoclaved MilliQ water to 200 mL ( $\sim 153$  mL). Supplement the prepared media with ampicillin and spectinomycin to 100 and  $50 \mu\text{g mL}^{-1}$ , respectively. *Note: Protein expression can be carried out entirely in  $2 \times \text{YT}$  or other rich media, however, this will result in higher background (protein expressed in the absence of an ncAA) and, depending on which MjTyrRS variant is used, may produce significant amounts of protein containing canonical amino acids*
3. Add the 10 mL overnight culture of co-transformed DH10B *metZWV* to the prepared CDM and incubate in a shaking incubator at  $37^\circ\text{C}$  until the  $\text{OD}_{600}$  reaches 0.3–0.6. *Note: DH10B *metZWV* grow considerably slower in CDM than in rich media. Under these conditions the culture will require  $\sim 4\text{--}5\text{h}$  to reach  $\text{OD}_{600}=0.3$*
4. Prepare a fresh solution (100 mM) of pAcF by dissolving 0.041g of pAcF in 2 mL of 0.3 M NaOH
5. Once the culture reaches an  $\text{OD}_{600}$  of 0.3–0.6 induce protein expression by adding to the culture flask 20  $\mu\text{L}$  IPTG (1 M, 0.1 mM final concentration) and 2 mL of L-arabinose (20%, 0.2% final concentration). Divide the culture by pouring half of the volume ( $\sim 100$  mL) into a second shaker flask. To one flask add 2 mL pAcF solution (100 mM in 0.3 M NaOH, 2 mM final concentration). To the second flask add 2 mL of 0.3M NaOH. The second culture will serve as a negative (no ncAA) control to gauge the degree of canonical amino acid incorporation
6. Return both flasks to the shaking incubator and incubate overnight ( $\sim 18\text{h}$ ) at  $37^\circ\text{C}$
7. Pellet the cells by centrifugation ( $3500 \times g$ , 20 min) and discard the supernatant. Store cell pellets at  $-80^\circ\text{C}$  until purification

### 3.1.3.3 Protein purification

1. Prepare 1× Lysis Buffer by combining the following components: 18 mL IMAC Buffer A, 2 mL BugBuster<sup>®</sup> protein extraction reagent, 20 mg lysozyme (1 mg mL<sup>-1</sup> final concentration), and 500 units Benzonase<sup>®</sup> nuclease
2. Transfer 1 mL of Talon<sup>®</sup> metal affinity resin to a polypropylene gravity flow column. Wash the resin with 5 column volumes of MilliQ water followed by 10 column volumes of IMAC Buffer A
3. Thaw cell pellets at room temperature and then resuspend each pellet in 10 mL 1× Lysis Buffer and incubate at room temperature for 30 min with gentle agitation
4. Pellet insoluble material by centrifugation (10,000 × *g*, 45 min) and transfer the soluble lysate to the column containing equilibrated Talon<sup>®</sup> metal affinity resin. Allow the lysate to slowly flow through the column
5. Wash the resin with 20 column volumes of IMAC Buffer A and then elute bound proteins with ~5 mL of IMAC Buffer B collecting ~0.5 mL fractions
6. Analyze the fractions by gel electrophoresis and combine fractions containing purified sfGFP-1pAcF. *Recommended: Further analysis of the protein by mass spectroscopy is recommended to confirm ncAA incorporation*

## 4. Summary

In this chapter we provided a detailed protocol to initiate *in vitro* protein synthesis with substituted benzoic acids and 1,3-dicarbonyls. The method described here relies on flexizyme-catalyzed acylation to generate tRNA<sup>fMet</sup> that is mis-acylated with the desired acid. The mis-acylated tRNA<sup>fMet</sup> is then used in cell-free protein synthesis reactions to initiate mRNA translation and introduce the benzoic acid or 1,3-dicarbonyl at the N-terminus of peptides. In addition to those described here, the *E. coli* ribosome has been shown to accept many diverse non- $\alpha$ -amino acids at the initiating position. More recently, some of these acids, including substituted benzoic acids, have been shown to be compatible with ribosomal elongation (Kato & Suga, 2020a, 2020b; Lee et al., 2020). Thus, the present represents an exciting time in the history of synthetic biology, as we are on the cusp of engineering ribosomes that can catalyze bond formation between consecutive non- $\alpha$ -amino acid monomers. Such technology will open the door to the eventual ribosomal synthesis of DNA-encoded chemical polymers with precisely defined sequences.

Herein we also described a protocol to initiate *in vivo* protein synthesis with ncAAs. This method relies on an engineered chimeric initiator tRNA ( $\dot{\lambda}$ tRNA<sup>Ty2</sup>) that is acylated by the MjTyrRS—an archaeal aaRS that is orthogonal in *E. coli* and has been engineered to recognize numerous aromatic ncAAs (Dumas et al., 2015). When co-expressed along with MjTyrRS, in media containing a desired ncAA,  $\dot{\lambda}$ tRNA<sup>Ty2</sup> can initiate protein synthesis at UAG codons to produce proteins containing the ncAA at their N-terminus. Although not described here, this system is compatible with the pyrrolysyl-tRNA synthetase (PylRS)•tRNA<sup>Pyl</sup> pair and can be used along with mutually orthogonal PylRS•tRNA<sup>Pyl</sup> pairs



to co-translationally install two or three distinct ncAA monomers into a single protein in *E. coli* (Tharp et al., 2020, 2021). Although MjTyrRS is currently limited to  $\alpha$ -amino and  $\alpha$ -hydroxy acid substrates, using *tRNA*<sup>Tyr2</sup> with further engineered MjTyrRS variants could enable genetic encoding of more disparate non-  $\alpha$ -amino acids at the N-terminus of proteins *in vivo*.

## Acknowledgments

This work was funded by the Center for Genetically Encoded Materials, an NSF Center for Chemical Innovation (NSF CHE-2002182 to A.S.). During the writing of this review, J.M.T was supported by NIGMS (R35GM122560 to D.S.).

## References

- Ad O, Hoffman KS, Cairns AG, Featherston AL, Miller SJ, Söll D, et al. (2019). Translation of diverse aramid- and 1,3-dicarbonyl-peptides by wild type ribosomes *in vitro*. *ACS Central Science*, 5, 1289–1294. [PubMed: 31403077]
- Bhattacharyya S, & Varshney U (2016). Evolution of initiator tRNAs and selection of methionine as the initiating amino acid. *RNA Biology*, 13, 810–819. [PubMed: 27322343]
- Budisa N, Steipe B, Demange P, Eckerskorn C, Kellermann J, & Huber R (1995). High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *European Journal of Biochemistry*, 230, 788–796. [PubMed: 7607253]
- Chattapadhyay R, Pelka H, & Schulman LH (1990). Initiation of *in vivo* protein synthesis with non-methionine amino acids. *Biochemistry*, 29, 4263–4268. [PubMed: 2112406]
- Chin JW (2017). Expanding and reprogramming the genetic code. *Nature*, 550, 53–60. [PubMed: 28980641]
- Dumas A, Lercher L, Spicer CD, & Davis BG (2015). Designing logical codon reassignment—Expanding the chemistry in biology. *Chemical Science*, 6, 50–69. [PubMed: 28553457]
- Fechter P, Rudinger-Thirion J, Tukalo M, & Giegé R (2001). Major tyrosine identity determinants in *Methanococcus jannaschii* and *Saccharomyces cerevisiae* *tRNA*<sup>Tyr</sup> are conserved but expressed differently. *European Journal of Biochemistry*, 268, 761–767. [PubMed: 11168416]
- Goto Y, Murakami H, & Suga H (2008). Initiating translation with D-amino acids. *RNA*, 14, 1390–1398. [PubMed: 18515548]
- Goto Y, Ohta A, Sako Y, Yamagishi Y, Murakami H, & Suga H (2008). Reprogramming the translation initiation for the synthesis of physiologically stable cyclic peptides. *ACS Chemical Biology*, 3, 120–129. [PubMed: 18215017]
- Goto Y, & Suga H (2009). Translation initiation with initiator tRNA charged with exotic peptides. *Journal of the American Chemical Society*, 131, 5040–5041. [PubMed: 19301866]
- Hammerling MJ, Krüger A, & Jewett MC (2020). Strategies for *in vitro* engineering of the translation machinery. *Nucleic Acids Research*, 48, 1068–1083. [PubMed: 31777928]
- Hammill JT, Miyake-Stoner S, Hazen JL, Jackson JC, & Mehl RA (2007). Preparation of site-specifically labeled fluorinated proteins for 19F-NMR structural characterization. *Nature Protocols*, 2, 2601–2607. [PubMed: 17948003]
- Kapoor S, Das G, & Varshney U (2011). Crucial contribution of the multiple copies of the initiator tRNA genes in the fidelity of *tRNA*<sup>Met</sup> selection on the ribosomal P-site in *Escherichia coli*. *Nucleic Acids Research*, 39, 202–212. [PubMed: 20798174]
- Katoh T, & Suga H (2018). Ribosomal incorporation of consecutive  $\beta$ -amino acids. *Journal of the American Chemical Society*, 140, 12159–12167. [PubMed: 30221942]
- Katoh T, & Suga H (2020a). Ribosomal elongation of aminobenzoic acid derivatives. *Journal of the American Chemical Society*, 142, 16518–16522. [PubMed: 32946698]
- Katoh T, & Suga H (2020b). Ribosomal elongation of cyclic  $\gamma$ -amino acids using a reprogrammed genetic code. *Journal of the American Chemical Society*, 142, 4965–4969. [PubMed: 32129615]

- Katoh T, Tajima K, & Suga H (2017). Consecutive elongation of D-amino acids in translation. *Cell Chemical Biology*, 24, 46–54. [PubMed: 28042044]
- Kawakami T, Murakami H, & Suga H (2008). Messenger RNA-programmed incorporation of multiple *N*-methyl-amino acids into linear and cyclic peptides. *Chemistry & Biology*, 15, 32–42. [PubMed: 18215771]
- Lee N, Bessho Y, Wei K, Szostak JW, & Suga H (2000). Ribozyme-catalyzed tRNA aminoacylation. *Nature Structural Biology*, 7, 28–33. [PubMed: 10625423]
- Lee J, Schwarz KJ, Kim DS, Moore JS, & Jewett MC (2020). Ribosome-mediated polymerization of long chain carbon and cyclic amino acids into peptides *in vitro*. *Nature Communications*, 11, 4304.
- Lee J, Schwieter KE, Watkins AM, Kim DS, Yu H, Schwarz KJ, et al. (2019). Expanding the limits of the second genetic code with ribozymes. *Nature Communications*, 10, 5097.
- Maini R, Kimura H, Takatsuji R, Katoh T, Goto Y, & Suga H (2019). Ribosomal formation of thioamide bonds in polypeptide synthesis. *Journal of the American Chemical Society*, 141, 20004–20008. [PubMed: 31815469]
- McMurry JL, & Chang MCY (2017). Fluorothreonyl-tRNA deacylase prevents mistranslation in the organofluorine producer *Streptomyces cattleya*. *Proceedings of the National Academy of Sciences of the United States of America*, 114, 11920–11925. [PubMed: 29078362]
- Morimoto J, Hayashi Y, Iwasaki K, & Suga H (2011). Flexizymes: Their evolutionary history and the origin of catalytic function. *Accounts of Chemical Research*, 44, 1359–1368. [PubMed: 21711008]
- Murakami H, Ohta A, Ashigai H, & Suga H (2006). A highly flexible tRNA acylation method for non-natural polypeptide synthesis. *Nature Methods*, 3, 357–359. [PubMed: 16628205]
- Murakami H, Saito H, & Suga H (2003). A versatile tRNA aminoacylation catalyst based on RNA. *Chemistry & Biology*, 10, 655–662. [PubMed: 12890539]
- Ngo JT, Schuman EM, & Tirrell DA (2013). Mutant methionyl-tRNA synthetase from bacteria enables site-selective N-terminal labeling of proteins expressed in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America*, 110, 4992–4997. [PubMed: 23479642]
- Niwa N, Yamagishi Y, Murakami H, & Suga H (2009). A flexizyme that selectively charges amino acids activated by a water-friendly leaving group. *Bioorganic & Medicinal Chemistry Letters*, 19, 3892–3894. [PubMed: 19364647]
- Ohta A, Murakami H, Higashimura E, & Suga H (2007). Synthesis of polyester by means of genetic code reprogramming. *Chemistry & Biology*, 14, 1315–1322. [PubMed: 18096500]
- Ramaswamy K, Saito H, Murakami H, Shiba K, & Suga H (2004). Designer ribozymes: Programming the tRNA specificity into flexizyme. *Journal of the American Chemical Society*, 126, 11454–11455. [PubMed: 15366888]
- Rodnina MV (2018). Translation in prokaryotes. *Cold Spring Harbor Perspectives in Biology*, 10, a032664. 10.1101/cshperspect.a032664. [PubMed: 29661790]
- Rogers JM, Kwon S, Dawson SJ, Mandal PK, Suga H, & Huc I (2018). Ribosomal synthesis and folding of peptide-helical aromatic foldamer hybrids. *Nature Chemistry*, 10, 405–412.
- Sakamoto K, & Hayashi A (2019). Synthetic tyrosine tRNA molecules with noncanonical secondary structures. *International Journal of Molecular Sciences*, 20, 92.
- Samhita L, Virumäe K, Remme J, & Varshney U (2013). Initiation with elongator tRNAs. *Journal of Bacteriology*, 195, 4202–4209. [PubMed: 23852868]
- Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K, et al. (2001). Cell-free translation reconstituted with purified components. *Nature Biotechnology*, 19, 751–755.
- Tadayoshi E, Kaori S, Takuya C, Yasunori T, Issei I, & Takahiro H (2010). Incorporation of non-natural amino acids with two labeling groups into the N-terminus of proteins. *Bulletin of the Chemical Society of Japan*, 83, 176–181.
- Takatsuji R, Shinbara K, Katoh T, Goto Y, Passioura T, Yajima R, et al. (2019). Ribosomal synthesis of backbone-cyclic peptides compatible with *in vitro* display. *Journal of the American Chemical Society*, 141, 2279–2287. [PubMed: 30648857]
- Tharp JM, Ad O, Amikura K, Ward FR, Garcia EM, Cate JHD, et al. (2020). Initiation of protein synthesis with non-canonical amino acids *in vivo*. *Angewandte Chemie International Edition*, 59, 3122–3126. [PubMed: 31828898]

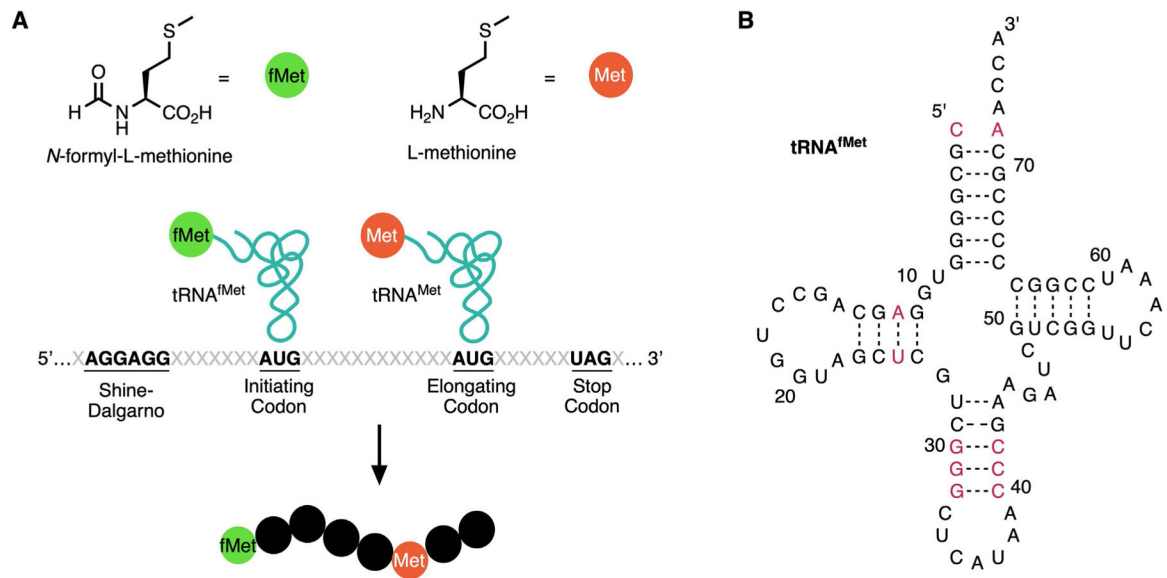
- Tharp JM, Vargas-Rodriguez O, Schepartz A, & Söll D (2021). Genetic encoding of three distinct noncanonical amino acids using reprogrammed initiator and nonsense codons. *ACS Chemical Biology*, 16, 766–774. 10.1021/acscchembio.1c00120. [PubMed: 33723984]
- Varshney U, & RajBhandary UL (1990). Initiation of protein synthesis from a termination codon. *Proceedings of the National Academy of Sciences of the United States of America*, 87, 1586–1590. [PubMed: 2406724]
- Young DD, & Schultz PG (2018). Playing with the molecules of life. *ACS Chemical Biology*, 13, 854–870. [PubMed: 29345901]

Author Manuscript

Author Manuscript

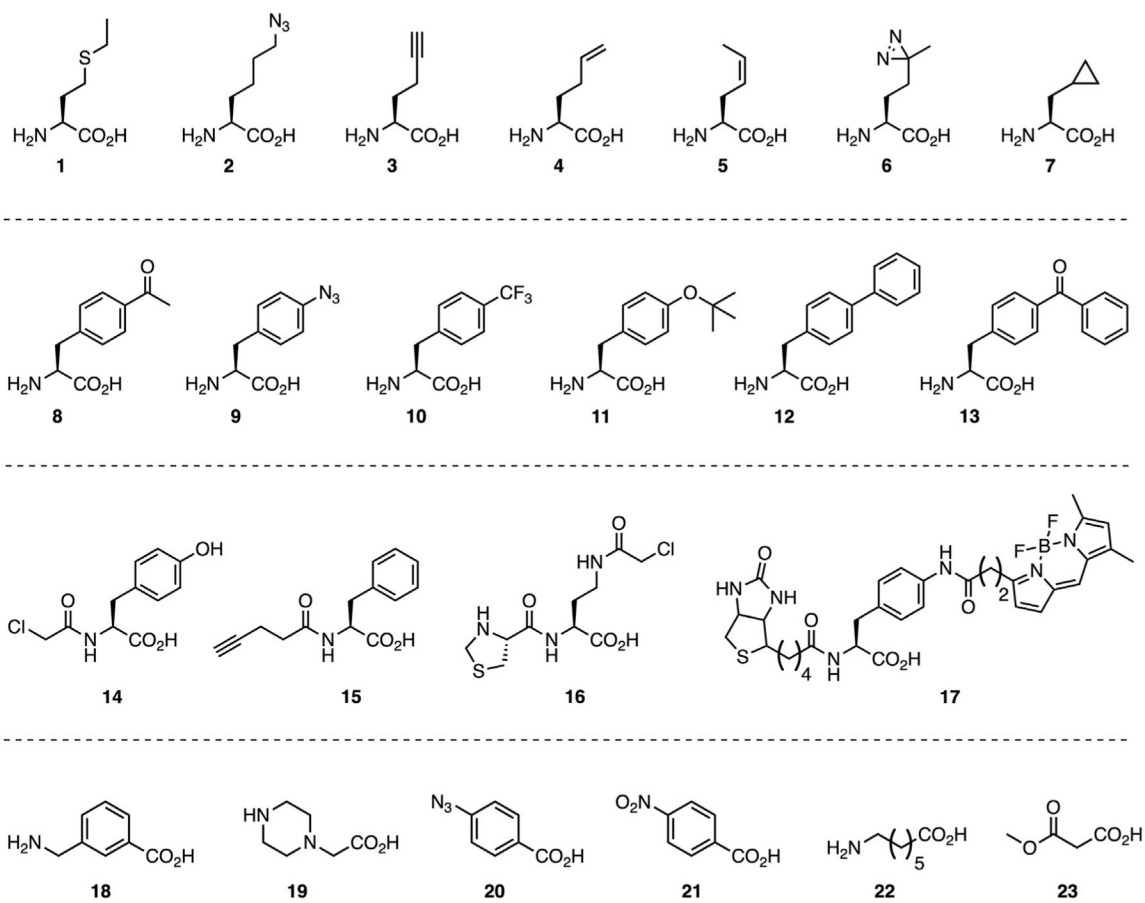
Author Manuscript

Author Manuscript

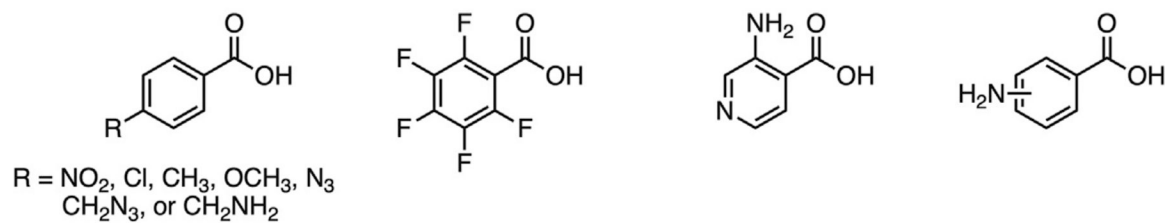
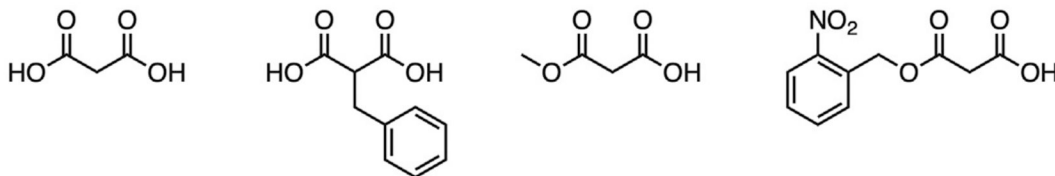


**Fig. 1.**

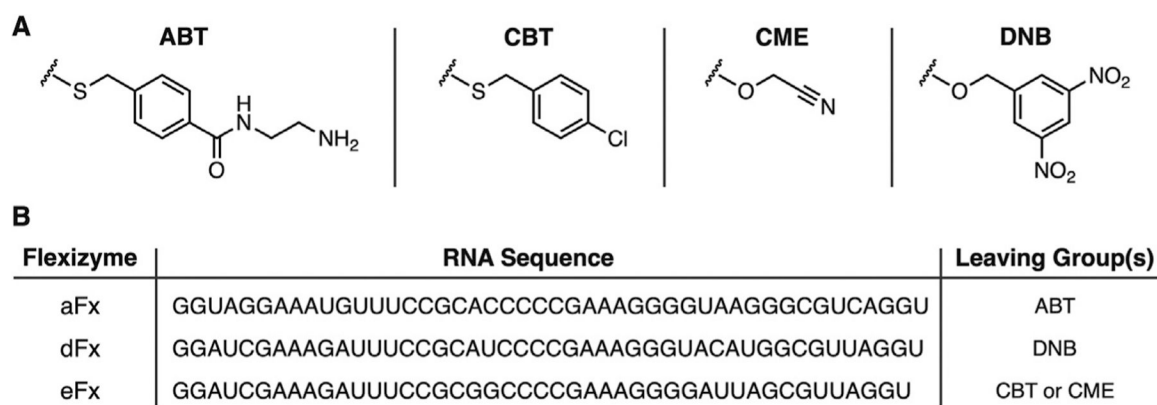
(A) Genetic encoding of *N*-formyl-L-methionine and L-methionine in response to initiating and elongating AUG codons. *N*-formyl-L-methionine is selectively incorporated at the N-terminus by tRNA<sup>fMet</sup> while L-methionine is selectively incorporated at elongating positions by tRNA<sup>Met</sup>. (B) The cloverleaf structure of tRNA<sup>fMet</sup> from *E. coli*. Highlighted residues are conserved among initiator tRNAs and contribute to the ability of tRNA<sup>fMet</sup> to initiate translation.



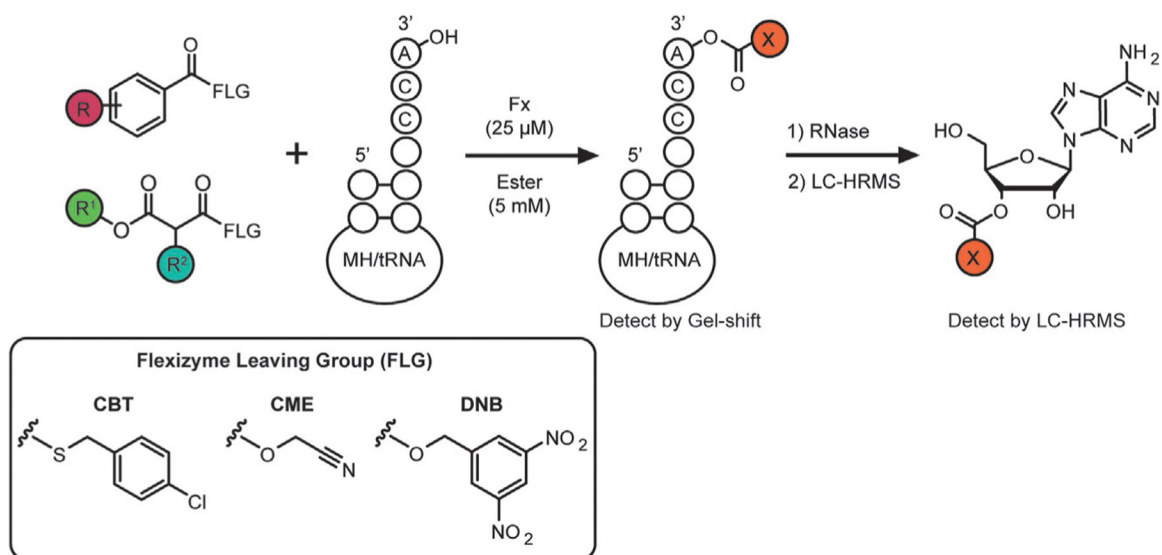
**Fig. 2.**  
Representative amino and non- $\alpha$ -amino acids that have been used to initiate translation *in vitro* and *in vivo*.

**Benzoic Acid Aramids****1,3-Dicarbonyls**

**Fig. 3.**  
Structures of benzoic acid, aramid, and malonate monomers used to initiate translation in this work.

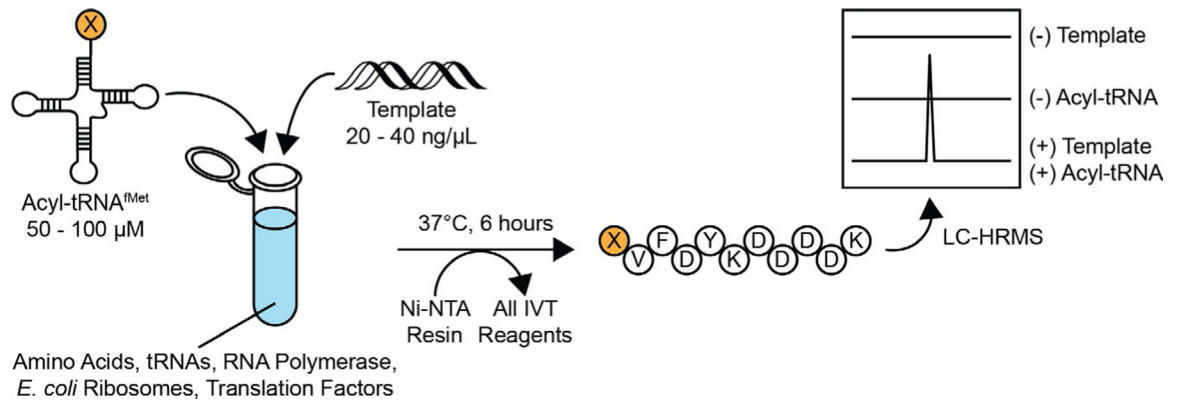


**Fig. 4.** (A) Structures of Leaving groups for flexizyme-compatible substrates. (B) Summary of flexizyme sequences and leaving groups used to identify substrates.

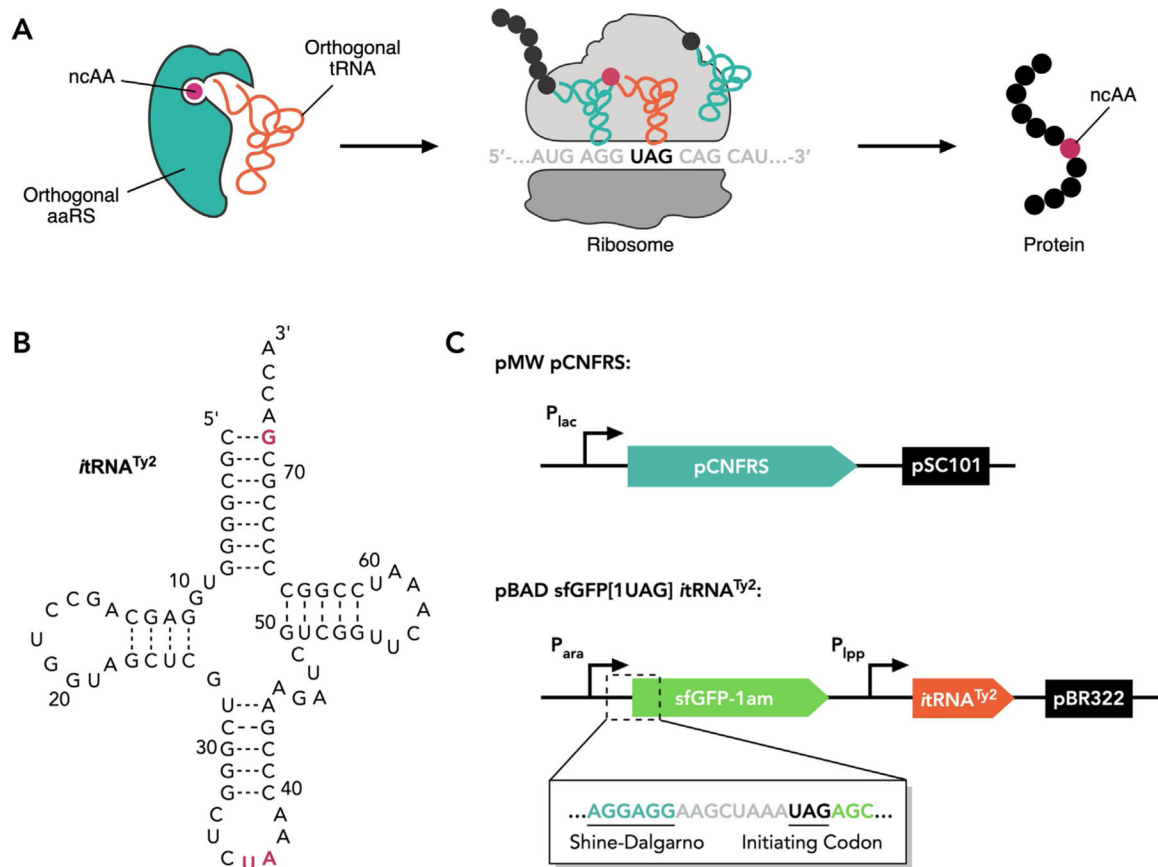


**Fig. 5.**  
Overview of microhelix acylation and analysis.





**Fig. 6.** Overview of reprogrammed ribosomal translation via the PURE system and subsequent LC-MS analysis of the translation products.

**Fig. 7.**

(A) Overview of genetic code expansion. Orthogonal aminoacyl-tRNA synthetase and tRNA pairs are used to introduce ncAAs into proteins during normal ribosomal protein synthesis. Typically, orthogonal suppressor tRNAs are used to introduce ncAAs in response to the UAG nonsense codon. (B) The cloverleaf structure of  $tRNA^{Ty2}$ . Positions highlighted in magenta differ from the sequence of  $tRNA^{fMet}$ . (C) Plasmid maps of the two-plasmid system used to introduce ncAAs at the N-terminus of superfolder green fluorescent protein (sfGFP). sfGFP with an M1  $\rightarrow$  UAG mutation (sfGFP-1 am) is co-expressed along with  $tRNA^{Ty2}$  and a polyspecific MjTyrRS mutant (pCNFRS) that recognizes *para*-substituted phenylalanine derivatives.