2014-03-01

Nascent Peptides That Induce Translational Arrest

Christopher J. Woolstenhulme
Brigham Young University - Provo

Nascent Peptides That Induce Translational Arrest

Christopher J. Woolstenhulme

A dissertation submitted to the faculty of
Brigham Young University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Allen R. Buskirk, Chair
Barry M. Willardson
Joel S. Griffitts
Julianne H. Grose
John T. Prince

Department of Chemistry and Biochemistry

Brigham Young University

March 2014

ABSTRACT

Nascent Peptides That Induce Translational Arrest

Christopher J. Woolstenhulme
Department of Chemistry and Biochemistry, BYU
Doctor of Philosophy

Although the ribosome is a very general catalyst, it cannot synthesize all protein sequences equally well. Certain proteins are capable of stalling the ribosome during their own synthesis. Stalling events are used by both prokaryotic and eukaryotic cells to regulate gene expression. Characterization of natural stalling peptides shows that only a few strategically placed amino acids are needed to inactivate the ribosome. These motifs share little sequence similarity suggesting that there are more stalling motifs yet to be discovered.

Here we use two genetic selections in *E. coli* to discover novel stalling peptides and detail their subsequent characterization. Kinetic studies show that some of these nascent peptides dramatically inhibit rates of peptide release by release factors. We find that residues upstream of the minimal stalling motif can either enhance or suppress this effect. In other stalling motifs, such as polyproline sequences, peptidyl transfer to a subset of aminoacyl-tRNAs is inhibited. Translation factor EF-P alleviates pausing of the polyproline motifs, but has little or no effect on other stalling sequences. The EF-P ortholog eIF5A also alleviates pausing of polyproline sequences in yeast.

Our studies show that short peptides sequences are capable of stalling the ribosome during elongation and termination through different mechanisms. These sequences are underrepresented in bacterial proteomes and show evidence of stalling on endogenous *E. coli* proteins.

Keywords: ribosome, translation, stalling, EF-P, eIF5A

ACKNOWLEDGEMENTS

As with most things in life, this was a group effort. Thank you Dad, Mom, Josh, Courtney, Lee, Natalie, Jess, Jim, Olivia and Abby, I couldn't have done any of this without your love and support. A massive thank you goes to Dr. Allen Buskirk for recruiting me from Montana, giving me some great projects and sending me to some amazing labs; you are a great friend, excellent mentor and you made this a first class experience! Finally, I'd like to thank my God for being extra generous with His blessings and guidance.

TABLE OF CONTENTS

Chapter 1. Introduction

Overview of Translation

The Ribosome

The ribosome is a large macromolecular complex that is responsible for protein synthesis in eukaryotic, archaeal and bacterial cells. It consists of two subunits whose names are derived from analytical sedimentation studies performed in the 1950s and 60s. In bacteria, the smaller subunit (or 30S subunit) (Figure 1-1B) is composed of a single ribosomal RNA (rRNA) strand called the 16S rRNA and ~20 proteins of varying size (Figure 1-1C). This subunit is primarily responsible for selecting the correct amino acids to add to the growing peptide chain. The large subunit (or 50S subunit) is composed of two rRNA strands, the 23S and 5S rRNA, as well as ~30 proteins and is responsible for peptide bond formation. The 30S and 50S subunits come together to form the 70S ribosome during active protein synthesis (Figure 1-1A). The rRNA/protein ratio varies between different organisms, but is approximately two-thirds rRNA to one-third protein by mass. The catalytic and functional core of the complex is composed of rRNA, making the ribosome a ribozyme [1]. The finding that RNA possesses catalytic activity lends support to the RNA world hypothesis, which postulates that self-replicating, catalytic RNA were the precursors to current living things based on DNA, RNA and proteins [2]. The catalytic function of the ribosomes is to facilitate peptide bond formation in a process called translation.

Translation

Translation is the intricate process of synthesizing proteins as encoded by the genetic information; it occurs in four stages: initiation, elongation, termination and recycling, as outlined

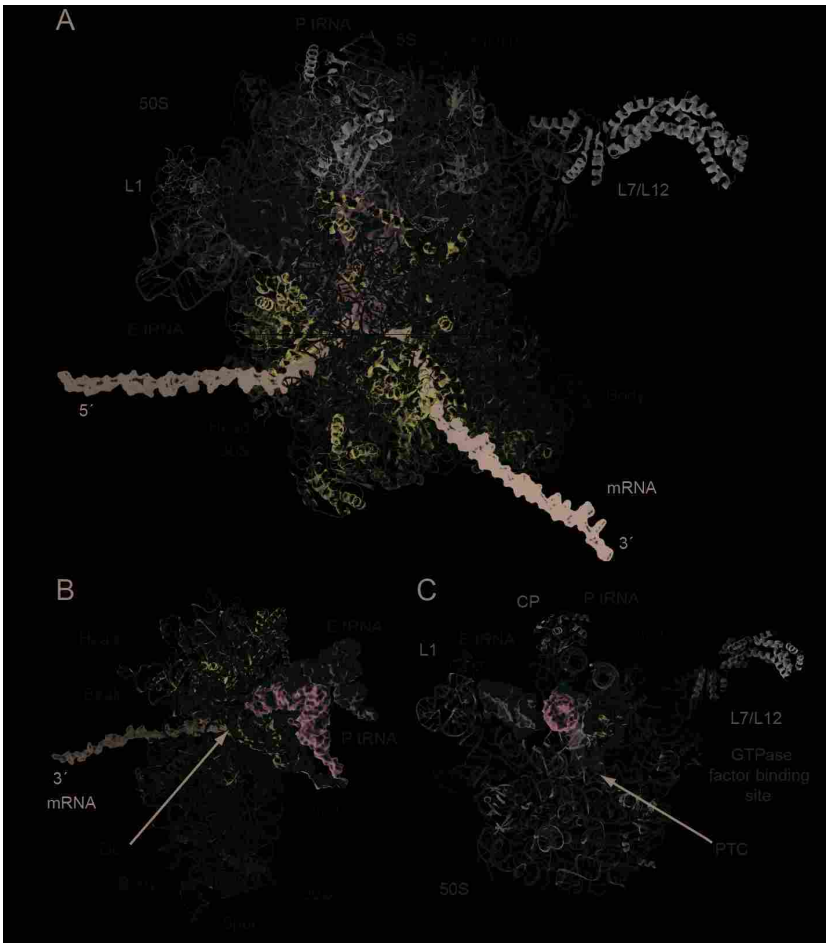


Figure 1-1. The Structure of the Ribosome. (A) 'Top' view of the 70S ribosome with mRNA and A, P and E site tRNAs. B, C, Exploded view of the 30S subunit (B) and 50S subunit (C). The structure of the L7/L12 arm10 was fit onto the 70S ribosome, with mRNA elongated by modeling. This and all other figures were made with Pymol (Delano Scientific) and Photoshop (Adobe). Reprinted by permission from Macmillan Publishers Ltd: (from Ref. 3).

in Figure 1-2 (for excellent reviews, see [3–5]). In its simplest terms, initiation involves the two subunits of the ribosome assembling on a messenger-RNA (mRNA) encoding the amino acid sequence that makes up the protein. In elongation, the ribosome decodes the mRNA by pairing the correct transfer-RNA (tRNA) with its corresponding three nucleotide codon in the mRNA. Each tRNA has already had its corresponding amino acid covalently attached to its 3'-end by enzymes called tRNA-synthetases, a process known as aminoacylation. The amino acids are then linked together as new tRNAs enter the ribosome and are moved through three ribosomal sites: the A site (aminoacyl-tRNA site), P site (peptidyl-tRNA site), and E site (exit site). Translation is completed at the termination step as the newly synthesized protein and tRNAs are released from the ribosome. Recycling then occurs where the

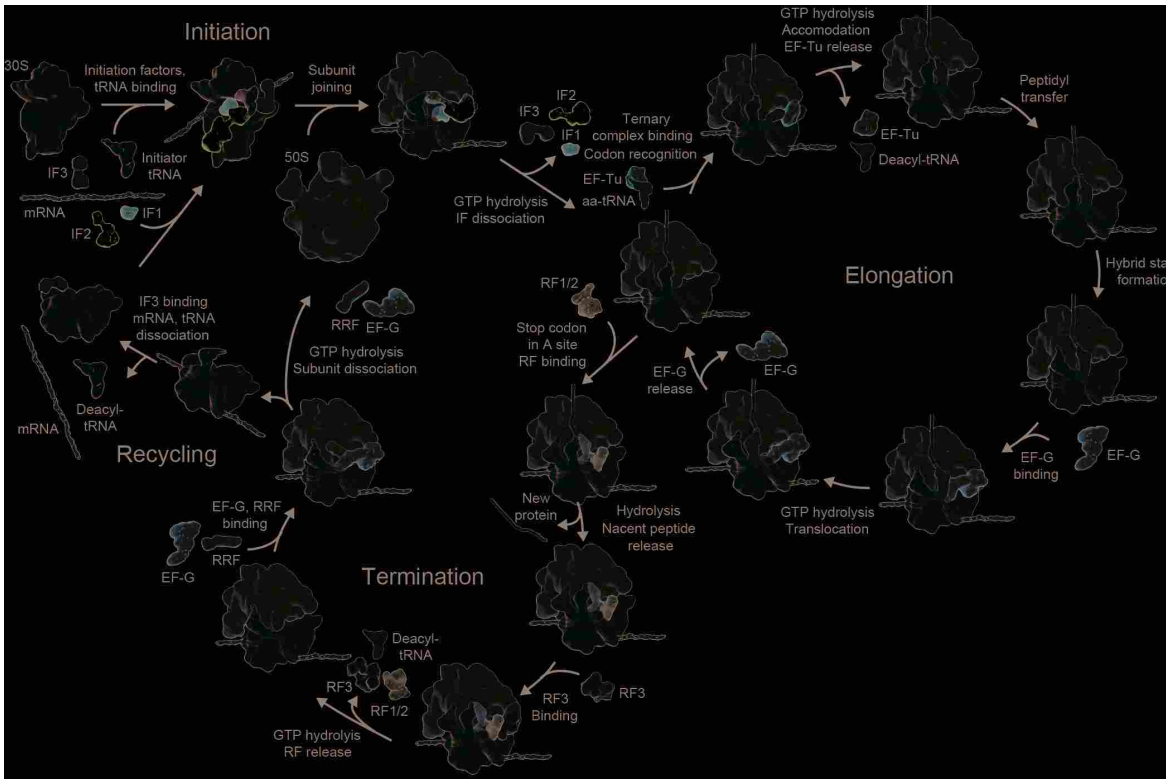


Figure 1-2. Overview of bacterial translation. For simplicity, not all intermediate steps are shown. aa-tRNA, aminoacyl-tRNA; EF elongation factor; IF, initiation factor; RF, release factor. Reprinted by permission from Macmillan Publishers Ltd: (from ref. (3))

ribosomal subunits are separated and readied for the next round of translation. We will discuss these steps in greater detail.

Initiation

Initiation starts on the 30S subunit with the co-operation of initiation factors 1, 2 and 3 (IF1, IF2 and IF3) (for reviews see [6,7]). At the end of a translation cycle, IF3 binds to the 30S E site and keeps the subunits separate after peptide release and subunit recycling (see Figure 1-3). IF1 and IF2 associate with the 30S subunit and facilitate the binding of the specially modified initiator methionine-tRNA (*N*-formyl-Met-tRNA) to the ribosome. IF1 binds at the A site and also contributes to the binding and positioning of fMet-tRNA in the P site [8]. IF2 recruits fMet-tRNA through recognition of the formylated α-amino group on the tRNA and subsequently

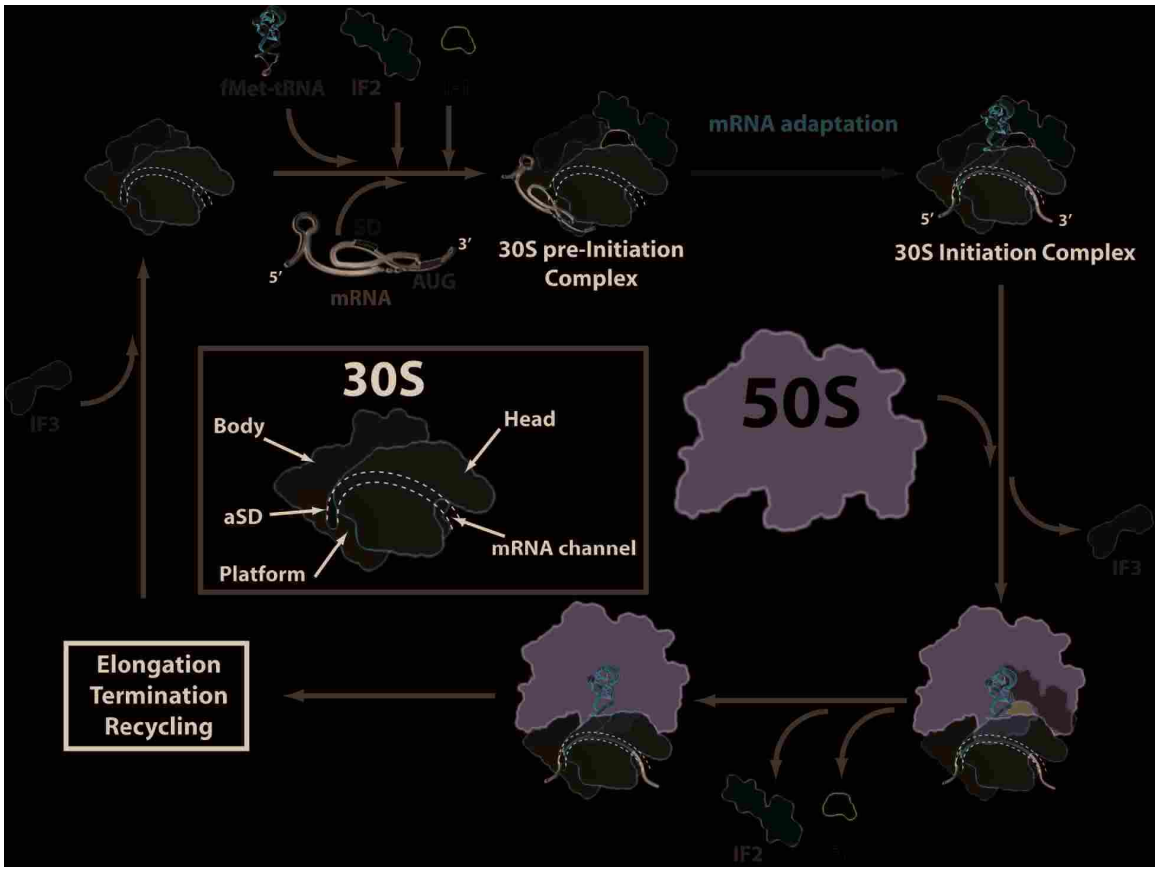


Figure 1-3. Schematic view of the initiation process**.** Formation of 30S (30SIC) and 70S (70SIC) translation initiation complexes, containing ribosomes (30S subunit in orange, 50S in brown), initiator fMet-tRNAfMet,mRNA and initiation factors IF1 (in blue), IF2 (in green) and IF3 (in light blue). View of 30S ribosomal subunit and ribosome from the top. The platform of the 30S is in red with the anti-Shine-Dalgarno (aSD) sequence in cyan. Structured mRNA binds to 30S in two distinct steps: the docking of the mRNA on the platform of the 30S subunit forms the pre-initiation complex that is followed by the accommodation of the mRNA into the normal path to promote the codon anticodon interaction in the P site (9). The resulting 30SIC engages the 50S subunit to form the 70SIC from which the initiation factors are expelled and the synthesis of the encoded protein can proceed through the elongation, termination and ribosome recycling phases (adapted from ref. 9). Reprinted from ref. 7 with kind permission from Springer Science and Business Media.

directs it to the P site [7]. These interactions increase the on-rate of P site binding of fMet-tRNA and decrease its dissociation rate. As the mRNA moves into its channel on the ribosome, the fMet-tRNA is brought to pair with the start codon. If pairing is correct, IF3 stabilizes the 30S initiation complex allowing the 50S subunit to associate (Figure 1-3). If the interaction is incorrect, IF3 destabilizes the complex causing the 30S subunit and initiation factors to dissociate before the 50S subunit can associate. In this way, IF3 acts as a kinetic fidelity control

mechanism [6]. Association of the 50S subunit ejects IF1 and IF3 from the complex allowing 70S formation. This triggers IF2 (a GTPase that binds the ribosome in its GTP bound form) to hydrolyze GTP into GDP changing its conformation. IF2 is then released from the fMet-tRNA and dissociates from the ribosome. The resulting complex (30S+50S+mRNA+tRNA) is termed the 70S initiation complex.

As mentioned earlier, binding of the mRNA to the 30S subunit is critically important for initiation and occurs in two stages. The first stage, binding of the mRNA, is a rapid step resulting in inactive 30S initiation complexes where the mRNA and fMet-tRNA do not interact (see Figure 1-3: 30S pre-Initiation Complex). A purine rich region in the 5' untranslated region (UTR) of the mRNA, termed the Shine-Dalgarno (SD) sequence, pairs with a complementary pyrimidine rich region at the 3'-end of the 16S rRNA called the anti-Shine-Dalgarno region (anti-SD) [9,10]. If the SD-anti-SD interaction is weak, the complex lingers in the inactive form and dissociates. A strong SD-anti-SD interaction prolongs the mRNA/ribosome association, allowing the mRNA to proceed to the second stage: adaptation. In this step, the mRNA moves into its proper channel on the 30S subunit (between the head and platform) to position the start codon (typically AUG) in the P site (see Figure 1-3: 30S Initiation Complex). This step is slow and requires the aid of IF1 and IF2 to align the start codon with the fMet-tRNA anti-codon. Correct pairing leads to an active 30S initiation complex that is stabilized by IF3 and binds with the 50S subunit as previously described. Once formed, the 70S initiation complex is ready for the second stage of translation: elongation.

Elongation

Elongation is the process by which the next codon on the mRNA is recognized by its corresponding aminoacyl-tRNA and the encoded amino acid is added to the peptide chain (Figure 1-2: Elongation) (for review see [4]). This process begins when aminoacyl-tRNA is brought from a pool of cellular tRNAs to the ribosome by the GTPase elongation factor Tu (EF-Tu). The tRNA-EF-Tu complex is directed to the A site of the ribosome through interactions with multimeric ribosomal protein L7/L12 (Figure 1-4A) [11]. The ribosome then determines if the tRNA is the correct, or cognate, match for the mRNA codon in a process called decoding.

The ribosome is able to discriminate between cognate and non-cognate tRNAs with a high degree of accuracy, making only one mistake per 1,000-10,000 amino acids incorporated. To achieve this high level of fidelity, the ribosome uses more than just correct base pairing of the mRNA codon/tRNA anticodon (Figure 1-4B,C). When the first two nucleotides of the mRNA/tRNA duplex correctly pair, a conformational change in the 16S rRNA is induced [12]. Universally conserved nucleotides A1492, A1493 and G530 recognize correct base-pairing by binding in the minor groove of the mRNA/tRNA helix, greatly stabilizing the mRNA/tRNA interaction (Figure 1-4C) [12]. The free energy generated from the combination of correct base-pairing and minor groove interactions is more than adequate to explain the high fidelity of the ribosome [13,14]. Residual energy from this interaction is then used to facilitate conformational changes in the ribosome including a large-scale domain closure of the 30S subunit (Figure 1-4C-E) [14–16].

The tRNA body itself is also important for proper decoding. When the tRNA-EF-Tu complex enters the ribosome, correct pairing causes a distortion in the tRNA body that allows the tRNA-EF-Tu complex to interact with both the mRNA and the 50S-binding site. This position is

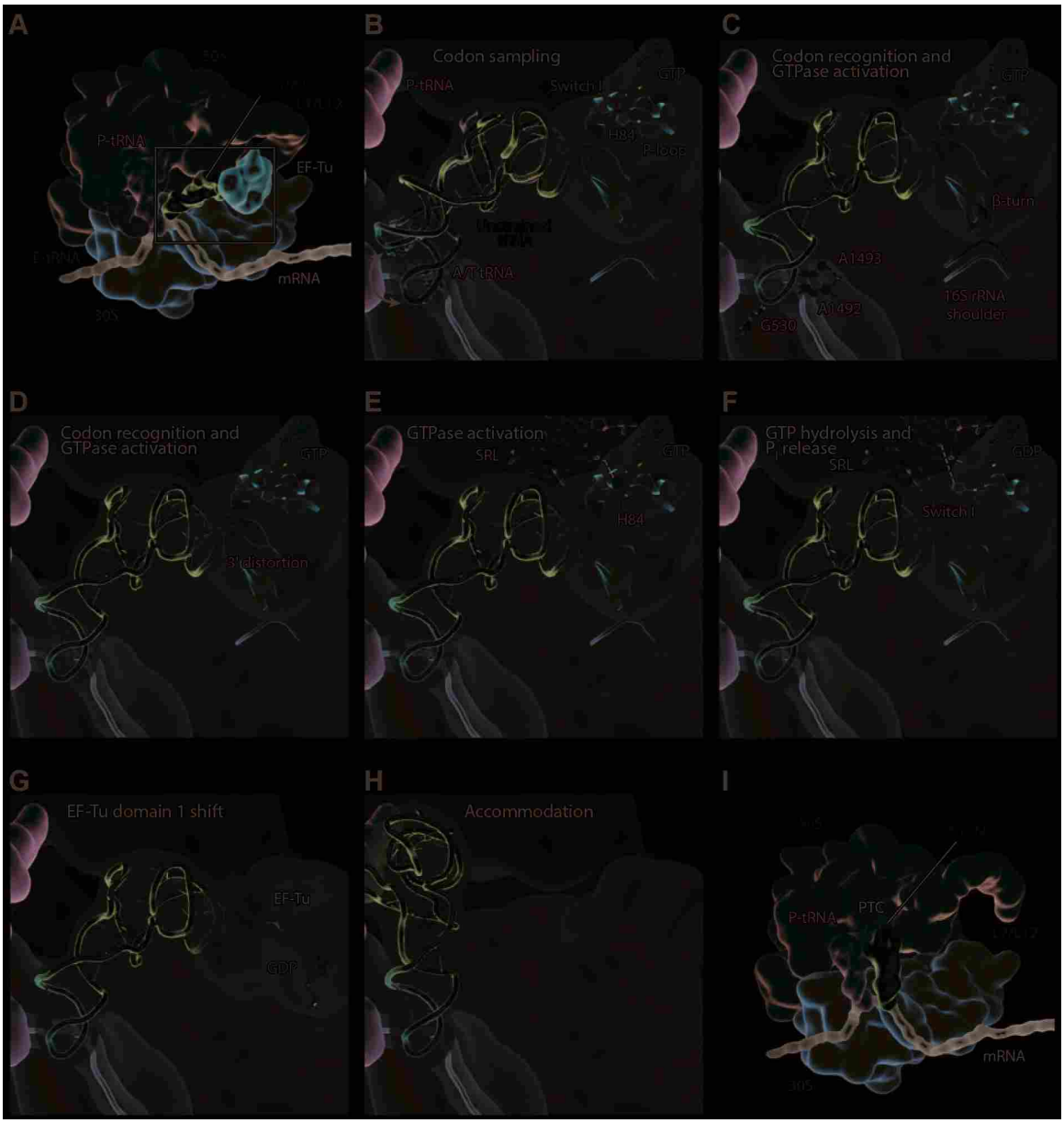


Figure 1-4. Schematic representation of the decoding pathway**.** (A) The L7/L12 stalk recruits the ternary complex to the ribosome. Deacylated transfer RNA (tRNA) may be bound in the exit (E) site (yellow) and peptidyl tRNA is in the peptidyl (P) site (green). The black rectangle represents the enlarged area in panels B-H. (B) The tRNA (purple) samples codon:anticodon pairing until a match (C) is sensed, by decoding center nucleotides G530 and A1492-A1493. Codon recognition triggers domain closure of the 30S subunit, bringing the shoulder domain into contact with elongation factor Tu (EF-Tu) (red), and shifting regions in domain 2 of the GTPase. (D) This results in a distortion of the acceptor arm of the aminoacyl tRNA. These conformational changes are all critical for properly positioning EF-Tu on the ribosome to allow GTPase activation. (E) GTPase activation does not require a large opening of the hydrophobic gate. Instead, residue A2662 of the sarcin-ricin loop (SRL) of the 23S ribosomal RNA (rRNA) positions His84 into the GTPase center, resulting in rapid GTP hydrolysis. (F) Release of Pi results in the disordering of the switch I loop and (G) a domain rearrangement of EF-Tu. (H,I) This leads to dissociation of EF-Tu from the ribosome, accommodation of aminoacyl tRNA, and peptidyl transfer. Abbreviations: A, aminoacyl; PTC, peptidyl transferase center. Republished with permission of Annual Review, Inc, from ref. (4); permission conveyed through Copyright Clearance Center, Inc.

known as the "A/T" state because the tRNA is bound in the 30S A site and to EF-Tu (Figure 1-4B) [17–19]. EF-Tu also undergoes a conformational rearrangement as it contacts the ribosome, reconfiguring its GTPase active site for hydrolysis (Figure 1-4C-D) [18]. This distortion and conformational change are energetically offset by the excess energy from decoding earlier described [13,14]. The ability the tRNA/EF-Tu to contact both the mRNA and 50S subunit is critical for GTPase activation of EF-Tu.

Activation of the GTPase activity of EF-Tu, and subsequent GTP hydrolysis, results in a large domain rearrangement in EF-Tu causing its release from tRNA and the ribosome. Prior to activation, EF-Tu prevents hydrolysis by protecting GTP from catalytic residue His84 [20] with a hydrophobic gate composed of residues Val20 and Ile60 (Figures 1-4D and 1-5) [18,21]. When GTPase activation occurs, the previously mentioned conformational rearrangements draw the GTPase center into contact A2662 of the sarcin-ricin Loop (SRL) of the 23S rRNA (Figure 1-4E-F) [21]. This pushes His84 past the hydrophobic gate where it can coordinate with a water molecule for an in-line attack on the γ-phosphate of GTP (Figure 1-5) [18,21]. This results in a massive 100° rotation of the nucleotide binding domain in EF-Tu (Figure 1-4G) [22] that causes its release from the tRNA and the ribosome [18,23]. While non-cognate tRNAs are able to illicit the hydrolysis response, the reaction is much slower, more often resulting in tRNA-EF-Tu dissociation from the ribosome than GTP hydrolysis [16]. This induced-fit mechanism of A site tRNA discrimination during initial selection acts as part of a kinetic control for translation, favoring cognate tRNAs over non-cognate tRNAs [16]. The other kinetic control of fidelity involves the next step: accommodation.

When EF-Tu leaves, the tRNA is held to the ribosome primarily by interactions at the decoding site, which are stronger for cognate vs. non-cognate tRNA. The structural stress from

the distortions in the tRNA body causes a rotation of the tRNA-CCA end similar to the unwinding of a coiled spring (Figure 1-4H) [24]. This places the tRNA-CCA end and amino acid in the peptidyl-transferase center (PTC) of the 50S subunit where it can be added to the peptide chain [24]. Movement of the tRNA-CCA end into the PTC is termed accommodation and, together with initial selection, are rate-limiting steps in translation [16]. This step is greatly accelerated with cognate tRNAs; non-cognate tRNAs are much slower and fall off more often [14,25]. Typically, non-cognate tRNAs will dissociate during initial selection; however, they may also fall off after GTP hydrolysis in order to dissipate torsional stress in the tRNA [25]. In this manner, the ribosome is able to achieve proofreading up to the point of peptide-bond formation.

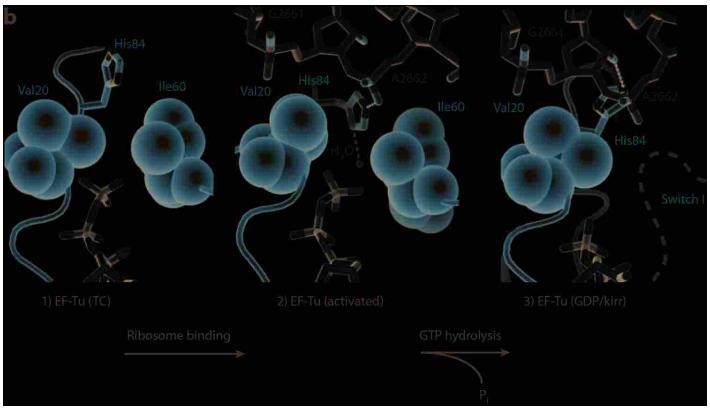


Figure 1-5. EF-Tu Hydrophobic Gate. Prior to binding of the ternary complex to the ribosome, the catalytic histidine is in an inactive conformation, rotated away from GTP (23). Binding to the ribosome and GTPase activation allows the phosphate of residue A2662 of the sarcin-ricin loop (SRL) (cyan) to position His84 into the active site. Finally, after GTP hydrolysis (18) and Pi release, the switch I loop is disordered (dashed line) and His84 has returned to an inactive conformation, contacting residue G2661 of the SRL. Republished with permission of Annual Review, Inc, from ref. (4); permission conveyed through Copyright Clearance Center, Inc.

After accommodation, the A and P site tRNA substrates are positioned at the PTC to facilitate peptide-bond formation (Figure 1-4I). The CCA ends of the A site and P site tRNAs pair with conserved nucleotides in the A-loop and P-loop of the ribosome, respectively [26]. This positions the carbonyl carbon of the P-site tRNA for nucleophilic attack by the α-amino group of the A-site tRNA, forming the peptide bond [27]. Peptide bond formation breaks the 3' ester linkage

of the peptide to the P site tRNA, transferring the peptide to the A site tRNA (Figure 1-6). This reaction occurs 10^7 times faster in the ribosome than in uncatalyzed reactions [3].

Structural studies show that the PTC is comprised entirely of RNA, suggesting that the ribosome functions as a ribozyme (RNA enzyme) [1,28,29]. The ribosome is thought to act primarily as an entropy trap to facilitate peptide bond formation. This means that the ribosome positions the substrates (CCA ends with A-loop and P-loop) and excludes water from the active site, rather than participating in conventional catalytic strategies such as acid/base catalysis [30]. Recent studies suggest that it may also play a small role in transition state stabilization [31,32].

Many mechanisms for ribosome catalyzed peptide bond formation have been proposed. Among these mechanisms, the consensus is that the peptide bond formation and 3' ester break are step-wise and not concerted [32,33], that at least two protons are in flight during catalysis [33], and that the 2'-OH of A76 of the A site tRNA acts in some capacity as a proton shuttle [34,35]. The basic mechanism is as follows: once in the PTC, the α-amine of the A site amino acid is positioned as a nucleophile to attack the C-terminal carbonyl of the peptide attached to the P site tRNA. This leads to the formation of a peptide bond and the addition of a new amino acid to the nascent peptide. The 3' ester linkage then breaks, transferring of the growing peptide to the A site tRNA and leaving a deacylated P site tRNA (Figure 1-6).

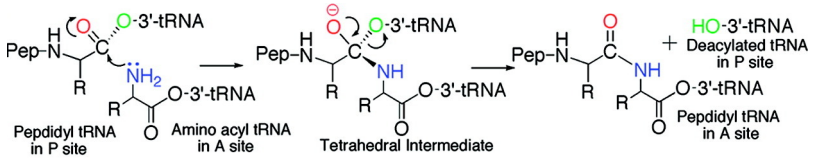


Figure 1-6. The peptidyl transferase reaction**.** In the tetrahedral intermediate the α-amine is drawn in the deprotonated form based on resent biochemical results (31). Reprinted with permission from ref. (31). Copyright 2011 American Chemical Society.

After the peptide is transferred to the A site tRNA, both the A site tRNA and the mRNA must move to P site to allow space for the next codon and cognate aminoacyl-tRNA to enter

(Figure 1-7A). This process is termed translocation. Once the P site tRNA is deacylated, its CCA end naturally moves into the E site on the 50S portion of the ribosome. This tRNA is now in a "hybrid" state called "P/E" state, where its 3' end now occupies the E site while the anticodon portion still occupies the P site (Figure 1-7B) [17,36]. The 3' end of the peptidyl tRNA then moves into the P site and occupies a hybrid position known as the "A/P" state because its anticodon portion still occupies the A site (Figure 1-7B). These movements are facilitated by a factor induced rotation of the 30S subunit relative to the 50S subunit [37].

The binding of GTPase elongation factor G (EF-G), in its GTP bound form, facilitates rotation of the 30S and 50S subunits relative to each other [38–41] as well as internal "swiveling" motions in the 30S subunit [42]. Binding near the A site, EF-G hydrolyzes GTP to GDP in a manner similar to EF-Tu [21] causing a structural rearrangement in EF-G (Figure 1-7C-D) [43]. This change in structure places a portion of EF-G directly into the A site, stabilizing the hybrid state,

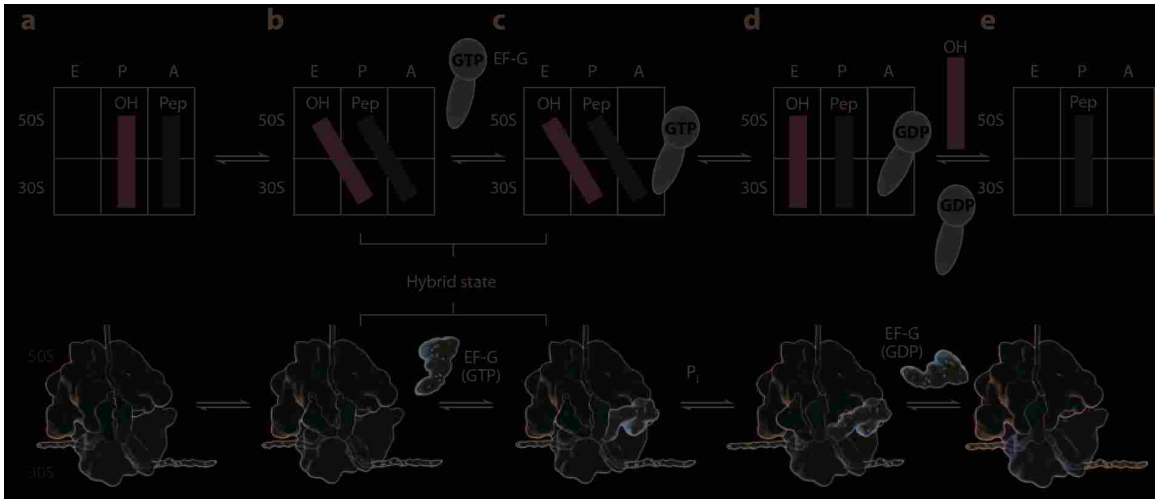


Figure 1-7. Schematic of translocation**.** Translocation can be divided into two steps: the first in which the transfer RNAs (tRNAs) move from their canonical conformations (panel a) relative to the 50S subunit (panel b), and the second, catalyzed by the GTPase elongation factor G (EF-G) (brown), in which the messenger RNA (mRNA) and tRNAs move relative to the 30S subunit (panels c and d). Although a high-resolution crystal structure of EF-G bound to the ribosome in the post-GTP hydrolysis state (panel d) was recently determined (see ref (4)), the detailed structures of A/P tRNA and EF-G bound in its GTP state to the ribosome are not known (panels b and c). Abbreviations: A, aminoacyl; E, exit; P, peptidyl. Republished with permission of Annual Review, Inc, from ref. (4); permission conveyed through Copyright Clearance Center, Inc.

providing forward movement of the mobile tRNA/mRNA duplex and also preventing back translocation of the tRNA [44]. Two nucleotides of the 16S rRNA are also used as "pawls" in a "ratchet" to fix the position of the mRNA, preventing its backward movement [44]. The codon/anti-codon interactions between the tRNAs and the mRNA cause the mRNA to be pulled one codon into the ribosome (Figure 1-7D). After GTP hydrolysis, the E site tRNA leaves the ribosome and the small subunit ratchets back into place (Figure 1-7E). EF-G-GDP also leaves the A site and the ribosome is ready for another round of elongation (Figure 1-7E) [45].

While elongation is occurring, the growing peptide moves past the PTC and enters a tunnel in the 50S subunit known as the exit tunnel [28] (Figure 1-8). This channel is approximately 100 Å long, 10-20 Å wide and can hold approximately 40 amino acids of the nascent peptide before the protein enters the cytoplasm. It is comprised primarily of the 23S rRNA, making it mostly hydrophilic [46]. Though the tunnel was originally believed to be non-interactive [28], electron microscopy experiments have

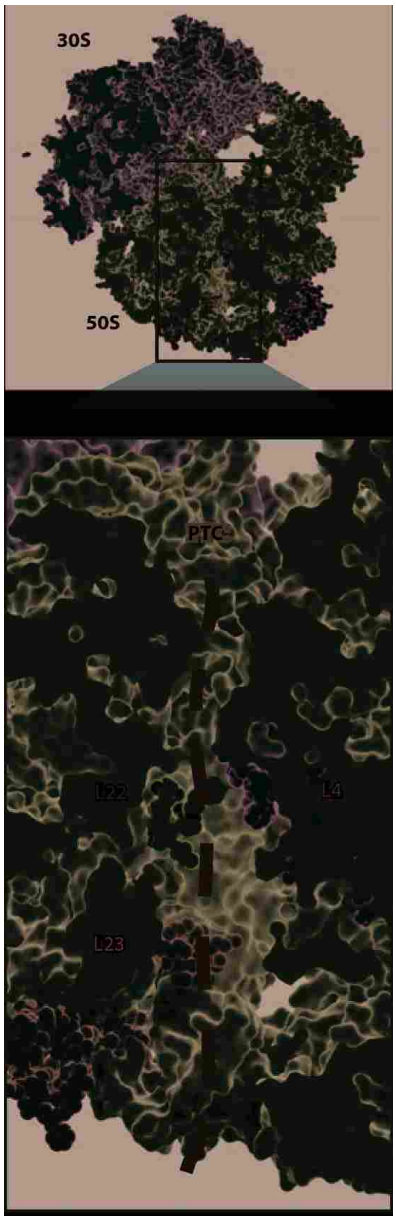


Figure 1-8. Sectional views of the bacterial and eukaryotic ribosomes**.** Images from Protein Data Bank (PDB) 2AW4/2AVY (Escherichia coli) were processed by the MacPyMOL program to show surface representations of ribosomal vertical sections, with proteins contributing or close to the exit tunnel represented as spheres. Nascent polypeptide chains are shown schematically by broken orange lines. Republished with permission of Annual Review, Inc, from ref. (50); permission conveyed through Copyright Clearance Center, Inc.

demonstrated the nascent peptide chain can adopt both α-helical and extended chain conformations in contact with the tunnel walls [47,48]. While the tunnel allows for some secondary structural formation, it is unlikely that the nascent peptide can adopt any higher order of folding due to the limited space in the exit tunnel [49]. One of the most notable features of the exit tunnel is a kink in the tunnel approximately 30 Å from the PTC. Structures of the ribosome indicate that this kink is formed by protrusions of ribosomal proteins L22 and L4 (Figure 1-8). This constriction plays a key role in the methods many stalling peptides use to arrest translation [50].

Termination

Translation is terminated when a stop codon enters the A site of the ribosome and is recognized by a class I release factor [51,52] (see reviews [53,54]). Class I release factors include release factor 1 (RF1), which recognizes stop codons UAG and UAA, and release factor 2 (RF2), which recognizes UGA and UAA. These release factors are tRNA mimics [55]) and enter the A site of ribosome in a closed form (Figure 1-9A-B) [56]. RF1 and RF2 are able to recognize stop codons using a "tripeptide anticodon" that forms extensive interactions with the mRNA stop codons and nearby rRNA bases [57]. Once the stop codon is recognized, the RF moves into an open conformation [56] moving its highly conserved GGQ motif into the peptidyl-transferase center (Figures 1-9B and 9D) [58]. This GGQ motif mimics the 3'-end of a tRNA and participates in the activation of a water molecule for nucleophilic attack on the peptidyl carbonyl [59–61]. Subsequent cleavage of the 3' ester releases the peptide from the tRNA (Figure 1-9E). The ribosome-RF1/2 complex then acts as a guanine exchange factor for the class II release factor RF3 [62], catalyzing the replacement of bound GDP with GTP. Binding of RF3-GTP to the ribosome accelerates the release of RF1/2 by stabilizing the ratcheted state of the 30S subunit [63]. Once RF1/2 leaves, RF3

hydrolyzes GTP to GDP and dissociates from the ribosome. This leaves the 70S in complex with mRNA and deacylated tRNA in P site.

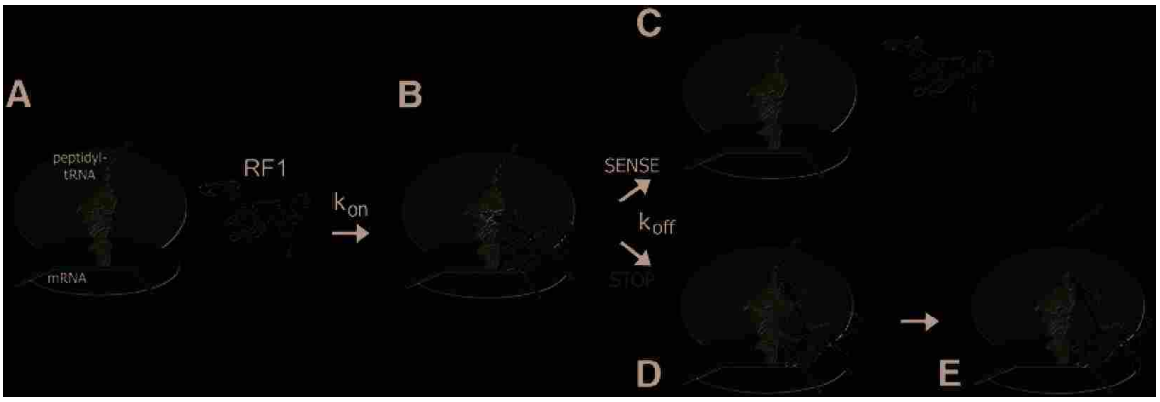


Figure 1-9. Proposed mechanism for coordination of peptidyl-tRNA hydrolysis with stopcodon recognition via a conformational switch in class I release factors**.** (A,B) Initially, the release factor binds to the ribosome in a catalytically inactive conformation. (C) If a sense codon is located in the A site, the release factor quickly dissociates (see ref. (53)). (D) If the release factor recognizes a stop codon in the A site, its switch loop along with domain 3 and the decoding center rearrange. Interaction between the switch loop and the switch-loop binding pocket in the decoding center results in tight binding of the release factor to the ribosome. (E) In this catalytically competent conformation, the GGQ motif is inserted in the peptidyl-transferase center and is capable of contributing to catalysis of peptidyl-tRNA ester bond hydrolysis. This figure is used with permission from ref. (53).

Recycling

Ribosome recycling is the final step in translation, where the subunits are separated and the translation cycle reset (for review see [64]). To begin, ribosome-recycling factor (RRF) enters the empty A site of the ribosome (Figure 1-2). This factor recruits EF-G bound with GTP (the same factor from the elongation stage) which acts the same as it would if tRNA were bound in the A site. EF-G hydrolyzes GTP, changes conformation and pushes RRF into the P site. This facilitates the separation of the subunits. IF3 then binds to the 30S E site prompting the dissociation of the tRNA and mRNA and remains bound to prevent re-association of the subunits [65,66]. In this manner, a peptide is made, released and the ribosome components made ready for the next round of translation.

Translational Arrest

Protein synthesis can pause or halt prematurely for several reasons. The ribosome can be paused by starvation or stress, low concentrations of substrates such as charged tRNAs and GTP, or the presence of antibiotics. In general, however, these pauses are reversible when the metabolic state of the cell improves. A more serious threat to translation is when the ribosome comes to the 3'-end of an mRNA without encountering a stop codon. Truncated mRNAs lacking stop codons arise from incomplete transcription reactions or mRNA decay pathways. Non-stop mRNAs are problematic because, in the absence of a stop codon, release factors are not recruited and the stalled ribosome remains locked an inactive state at the 3'-end of the non-stop mRNA. Cells have evolved rescue mechanisms to cope with non-stop mRNAs and other long-lived stalling events.

Rescue Machinery: tmRNA, ArfA and ArfB

There are three mechanisms used by bacteria to recognize and recycle ribosomes stalled on non-stop mRNAs [67]. The best characterized of these is tmRNA which is encoded by the *ssrA* gene (for reviews see [68,69]). tmRNA is an RNA molecule that contains both a tRNA-like domain (TLD) charged with alanine and an mRNA-like open reading frame. tmRNA enters the stalled ribosome and replaces the problematic mRNA with its own ORF. This allows the ribosome to resume translation and terminate normally using tmRNA as a template. The tmRNA ORF encodes a short peptide tag that is recognized by proteases, leading to degradation of the defective protein once translation is complete (Figure 1-10). Because it results in the synthesis of a chimeric protein from two RNA templates, this process is called trans-translation.

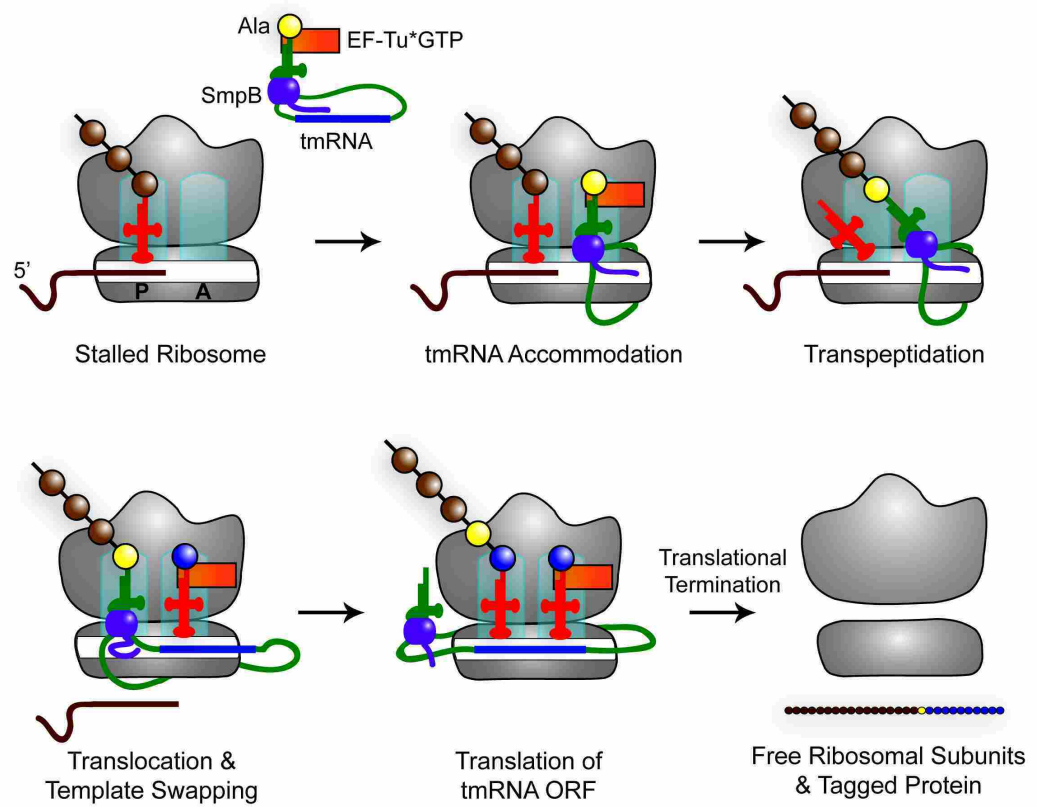


Figure 1-10. Diagram of tmRNA-mediated ribosome rescue**.** Acting as a tRNA, Alanyl-tmRNA (green) recognizes ribosomes stalled on nonstop mRNAs and adds alanine to the nascent polypeptide. Following mRNA swapping, tmRNA acts as a template (blue), encoding a 10 amino acid tag. RF1/RF2-mediated termination releases the tagged protein for degradation by cellular proteases. This also liberates the 30S and 50S subunits for new rounds of protein synthesis.

tmRNA identifies stalled ribosomes and performs trans-translation with the help of its protein partner SmpB [70]. SmpB is a small protein (160 residues) comprised of a core domain and a C-terminal tail. The core of SmpB contains an oligonucleotide binding (OB) fold that creates two RNA binding sites, one for binding tmRNA [71] and the other for binding ribosomal RNA in the A and P sites. The 30 residue long C-terminal tail is unstructured in solution but adopts an α-helical structure upon ribosome binding [72,73].

To avoid prematurely aborting translation, the tmRNA/SmpB complex must discriminate between stalled and actively translating ribosomes. Actively translating ribosomes have mRNA in the A site and the downstream mRNA channel. When a ribosome stalls, cellular

endonucleases cleave the mRNA, leaving the mRNA channel and A site empty [74]. The empty A site is recognized by SmpB, which binds in the decoding center in the A site and places its C-terminal tail into the empty mRNA channel [72]. The tail follows the path of the mRNA, making extensive contacts with the 16S rRNA and the S3, S4, and S5 proteins [72]. These interactions with the tail are essential for accommodation and peptidyl transfer to tmRNA/SmpB, as discussed below. In this manner, the competition between tmRNA/SmpB and mRNA for the A site allows discrimination between stalled and active ribosomes.

The tmRNA/SmpB complex is delivered to the A site by EF-Tu, as canonical tRNAs are, and triggers GTP hydrolysis by EF-Tu through structural mimicry of a cognate tRNA [75]. When bound to tmRNA, the SmpB body mimics the anticodon stem of a tRNA [75] and facilitates decoding through numerous contacts with the ribosomal decoding center. Conserved residue His136 (*E. coli* numbering) of the SmpB tail makes one of the most critical interactions by base stacking with G530 [72], also a key nucleotide in canonical decoding. Loss or mutation of this residue greatly diminishes activation of EF-Tu and GTP hydrolysis rates [76]. Neighboring conserved residues Lys138 and Arg139 also bind to the sugar phosphate of rRNA near G530, stabilizing these stacking interactions [72]. As with canonical decoding, recognition of the tmRNA/SmpB complex causes domain closure in the 30S subunit head, leading to GTPase activation and GTP hydrolysis. While the mechanism of EF-Tu activation resembles the canonical mechanism in some aspects, it was also observed that conditions that block EF-Tu activation may have little or no effect on the downstream peptidyl transfer step, suggesting that tmRNA dissociates from EF-Tu rapidly, even without GTP hydrolysis [76].

After decoding and GTP hydrolysis, the aminoacylated tRNA-like domain of the tmRNA moves into the PTC and facilitates peptidyl transfer. The peptide chain is transferred to tmRNA

and canonical translation resumes. During translation, tmRNA and SmpB remain bound together as they move through the ribosome, translocating from the A site to the P site and then out through the E site [77]. At the end of translation, the nascent peptide is released by binding of either release factor to the UAA stop codon in tmRNA, after which tmRNA is presumably recycled along with the ribosome subunits [69]. Prior to translocation, tmRNA exchanges its ORF for the non-stop mRNA [78]. The tmRNA ORF encodes a 10 amino acid tag (ANDENYALAA) that is added to the C-terminus of the peptide [79]. This sequence is recognized by the adaptor protein SspB, which enhances recruitment of cellular proteases like ClpXP that subsequently degrade the tagged proteins [80,81]. The sequence of the tmRNA ORF destines the protein for degradation but is not essential for ribosome rescue and recycling and can be changed without impairing tmRNA function.

Two other bacterial proteins, ArfA and ArfB, are engaged in translational rescue mechanisms similar to that catalyzed by tmRNA/SmpB [82,83]. All three systems recognize and release ribosomes stalled on non-stop mRNAs [67]. ArfA is a short peptide (72 residues in *E. coli*) that recognizes stalled ribosomes by binding the large subunit and entering the A site through an unknown mechanism [82–84]. Once in the ribosome, ArfA recruits RF2 to hydrolyze the peptidyl-tRNA and resolve arrest [84]. ArfB is a reduced paralog of translational RFs that has retained the GGQ domain [83]. It does not have an anticodon region, but it does have an unstructured tail that may function in a similar manner as the SmpB C-terminal tail [83]. Once in the A site, ArfB hydrolyzes the peptidyl-tRNA using its GGQ domain [85,86]. Together these three systems form a highly redundant rescue network necessary for cell survival.

Ribosomes stalled on non-stop mRNAs necessitate rescue because translation is irreversibly halted; however, not all stalling events are irreversible. In some instances the

ribosome is only temporarily stalled and resumes translation before mRNA degradation occurs. Known as ribosome pausing, this phenomenon often results from interference by nascent peptides that are programmed to stall their own translation. This manner of stalling is transient, reversible and serves a functional role within the cell.

The Scope and Significance of Ribosome Stalling

Certain proteins interfere with their own translation, arresting the ribosome due to a specific combination of amino acids in the nascent chain. Why would a protein evolve to inhibit its own synthesis? Many known biological examples demonstrate that ribosome stalling is effective at rapidly altering expression of genes (for review see [87]). For example, in *Bacillus subtilis*, stalling of the protein MifM is used to up-regulate translation of YidC2, a functional copy of the essential gene SpoIIIJ/YidC1, which is involved in membrane protein insertion [88,89]. *yidC2* encodes a homolog of SpoIIIJ and is located downstream of small ORF named *mifM*. When levels of SpoIIIJ are normal, translation of *yidC2* is repressed through sequestration of its SD region in mRNA secondary structure. When SpoIIIJ is defective or its cellular levels drop, MifM stalls the ribosome using the consensus sequence R^{69}IxxWIxxxxxMNxxxxDEED^{89} [88]. Prolonged ribosome occupancy on *mifM* disrupts downstream mRNA secondary structure unlocking the SD sequence of *yidC2* [88]. This facilitates *yidC2* translation, allowing the bacterial cell to quickly up-regulate expression of this gene, restoring insertion of membrane proteins like the c subunit of the F_1F_0 ATP synthetase [89,90].

A second example of gene regulation by ribosome stalling is in *Neurospora crassa,* which uses arginine-induced ribosome stalling to down-regulate arginine biosynthesis [91,92]. The *arg-2* mRNA encodes a subunit of an enzyme required for Arg biosynthesis in eukaryotes. This

mRNA also contains an upstream ORF encoding short leader peptide called the arginine attenuator peptide (AAP). AAP stalls the ribosome at termination in the presence of high levels of Arg using the consensus sequence T^{9}xxDYLxxxxWR[20] [93]. This blocks other eukaryotic ribosomes from pre-initiation scanning and thus effectively down-regulates arginine biosynthesis [92].

Comparative analysis of stalling sequences reveals that stalling peptides have little sequence similarity and rely on a few critical interactions with the ribosome to cause stalling (see Table 1-1) [50]. Despite the diversity of these motifs, many share the same three patterns of interactions with the ribosome. First, conserved residues near the N-terminus of several stalling motifs interact with the constriction site of the exit tunnel between proteins L4 and L22. Consequently, mutations in L4, L22, or nearby rRNA nucleotides were shown in genetic screens to reduce stalling [94,95]. Second, conserved residues near the C-terminus of a motif are positioned close to the PTC and may interact with nearby nucleotides [96]. Third, some motifs encode a specific aminoacyl-tRNA that binds in the A site but fails to undergo peptidyl transfer [97,98]. A variation of this rule involves stalling at termination where release factors are unable to facilitate peptide release [99–101]. All three of these interactions play a role in stalling in the SecM, ErmCL and TnaC motifs and make them excellent candidates for in-depth investigation.

SecM Stalling

In *E. coli*, the Sec protein translocase is responsible for transporting proteins across the inner plasma membrane and into the periplasm [102]. In this system, nascent proteins are trapped in their unfolded state and brought to the translocase machinery by the chaperone SecB [103]. These proteins are then transported across the membrane though the SecYEG protein channel, a process

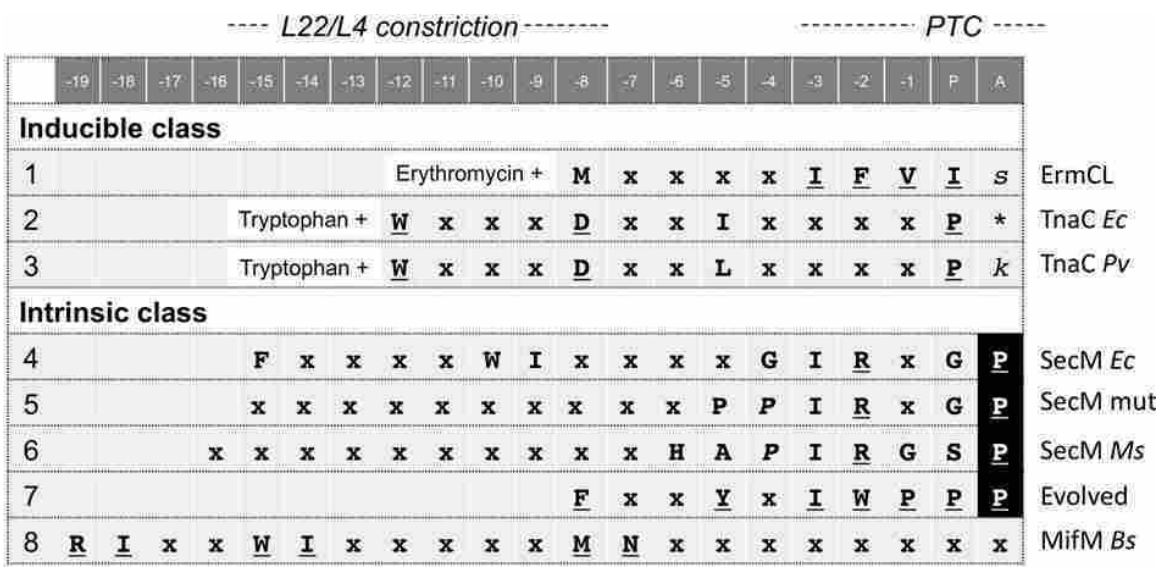


Table 1-1. Stalling Sequences**.** Stalling sequences that have been subjected to comprehensive mutational analysis. Sequences that arrest translation elongation are aligned based on their likely positions in the stalled ribosome, with numbering starting inversely from −1 for the position immediately preceding the P-site amino acid. Approximate locations of amino acids in the ribosome are indicated at the top, on the basis of the structure of the extended TnaC peptide–ribosome complex (118). Note that in the cases of SecM and others, the intraribosomal peptide may be more compacted. Translation ends with the P-site amino acid as the last amino acid of the nascent peptidyl-tRNA (note, however, that ribosomal occupancy has not been determined for MifM). Residues essential for the elongation arrest are underlined. Residues denoted x are less important as they can be changed to one or more different amino acid(s) without affecting the arrest. In all cases that have been examined, the arrest-essential amino acids need to be separated with the exact spacings shown. The A-site amino acids shown in lower case italics are not required for the arrest. The A-site prolines shown in reverse upper case are essential for the arrest. In the case of *E. coli* TnaC, the A-site codon is a UGA stop (shown by asterisk). 1, Leader peptide of the erythromycin resistance gene, *ermC* (96); 2, and 3, arrest sequence of the tryptophanase operon of *E. coli* (135) and *Proteus vulgaris* (139), respectively; 4, arrest sequence of SecM from *E. coli* (95); 5, a mutant form of the SecM arrest sequence having proline at −4 and −5 positions, of which the one at −4 (italicized) alleviates the specificity of the constriction-proximal residues (108); 6, arrest sequence of SecM from *Mannheimia succiniciproducens* (108); 7, an experimentally evolved arrest sequence obtained by genetic screening (98); 8, arrest sequence of MifM from *B. subtilis* (88). Used by permission from ref. (50).

driven by ATPase SecA [104]. What little we know of the regulation of these genes suggests that their expression is coordinated with housekeeping functions [105]. However, the regulation of *secA* by leader peptide SecM has been extensively characterized.

Expression of *secA* is regulated through a controlled feedback loop involving cellular levels of SecA and translation of the 170 amino acid long leader peptide SecM [95,106]. Both *secM* and *secA* are transcribed on the same bicistronic mRNA, with *secM* located upstream of *secA*.

Secondary structure in the mRNA sequesters the SD region of *secA*, effectively repressing its translation unless disrupted (Figure 1-11A). This does not affect *secM*, however, which is efficiently translated until the ribosome encounters an encoded peptide-stalling motif. When cellular levels of SecA are low, the stalling event in SecM translation is prolonged, causing the downstream mRNA secondary structure to dissolve. This frees the SD region of *secA* for ribosome recognition, thus activating its translation (Figure 1-11A) [95,106]. When SecA levels are normal, the SecM peptide is quickly pulled from the stalled complex by SecA and translocated across the membrane where it is degraded [106,107]. In the absence of stalling, the mRNA structure refolds and blocks initiation and translation of SecA.

SecM causes ribosome stalling using the consensus sequence F^{150}xxxxWIxxxxGIRAGP[166] to interact with the exit tunnel and PTC of the ribosome [95,108] (Figures 1-11B and 11C). During translation, SecM residues Phe150, Trp155 and Ile156 compact at the constriction site in the exit tunnel, causing interactions with residue A751 of the 23S rRNA and residues Gly91 and Ala93 of the L22 beta-hairpin loop [109,110]. These SecM residues are important for stalling and require strategic placement to interact with exit tunnel elements in order to facilitate translational arrest [95,111]. Additional exit tunnel interactions have also been observed between SecM, L22 (Gln72), L23 (Lys84), and 23S rRNA nucleotide A1321[109]; however, these contacts are not essential for stalling, and can be compensated for by the substitution of a Pro residue at position 161 [95,108].

Compaction of SecM at the constriction site causes Arg163 to interact with the exit tunnel and inactivate the PTC. Placement of Arg163 within the exit tunnel is critical and will not tolerate even a +/− 1 residue shift [108]. Once positioned, Arg163 presses against rRNA nucleotide A2062, which otherwise is freely rotating [112], pushing it into A2503 (Figure 1-11C) [29,109,113].

Arg163, A2062 and A2503 are all highly conserved and are essential for SecM stalling [114,95,108]. Interactions between these elements initiate a relay mechanism through adjoining rRNA residues resulting in the inactivation of the PTC (Figure 1-11D) [113–115] Part of this inactivation includes movement of the P-site tRNA-peptide ester-linkage 2 Å away from the PTC [109]. As previously discussed, the P-site tRNA carbonyl carbon requires proper positioning in order to facilitate

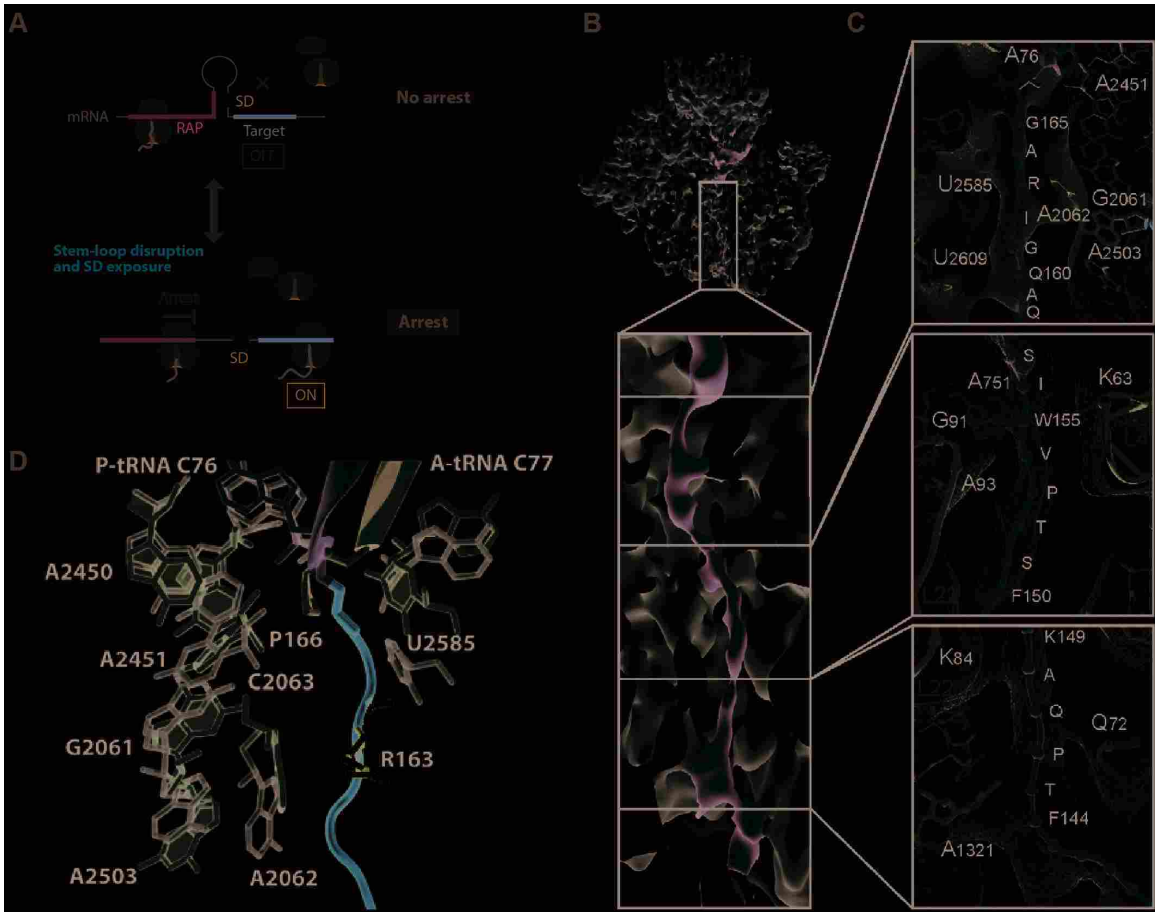


Figure 1-11. Ribosome Stalling by SecM**.** (A) Nascent polypeptide-ribosome-mRNA complexes are schematically depicted to show the effects of a ribosome stalled by an arrest sequence on the mRNA secondary structure (B) Cross-section of the large ribosomal subunit of the SecM-stalled RNC revealing the sites of interaction between the SecM nascent chain (green) and the ribosomal tunnel (gray). (C) Close-up of the upper, middle, and lower regions of the ribosomal tunnel with density (gray mesh) and molecular models for SecM nascent chain (green, with balls marking the Ca of the labeled residues; blue indicates the residue is important for stalling), the 23S rRNA (gray, except for selected colored nucleotides), and ribosomal proteins L4 (purple), L22 (orange), and L23 (cyan). (D) Positions of key relay residues after fitting compared with canonical positions in 3WDK/3WDL. Residues are shown as colored sticks (fitted structure) or in dark gray (2WDK/2WDL). (A) is used with permission from ref. (87)). (B) and (C) used with permission from ref. (109). (D) is reprinted from ref. (113). Copyright 2012, with permission from Elsevier.

nucleophilic attack by the A-site α-amino group. A 2 Å shift would dramatically reduce translational efficiency [27,109].

Altered PTC geometry also depends on Pro166, which contributes to stalling from the A site [95,116]. Pro166 is highly conserved in bacterial SecM species and is essential for stalling [95,108]. Pro166 is not added to the growing peptide chain, indicating that its action is effected through the A site as prolyl-tRNA [116]. Its effect can be explained by the unique properties of proline that make it both a poor peptidyl donor and acceptor [116–118]. The presence of prolyl-tRNA in the A site prevents both puromycin and cellular rescue-machinery from accessing the stalled ribosome complex [117]. Pro166 and Arg163 are considered the most important contributors to SecM mediated stalling as is indicated by their high sensitivity to mutation [108].

ErmCL Stalling

Erythromycin (ERY) is a macrolide antibiotic that stops bacterial cell growth by inhibiting translation and causing the accumulation of peptidyl-tRNAs [119–121]. Bacteria have evolved resistance to these drugs by using methyltransferases to dimethylate rRNA residue A2058. This residue is located in the drug binding site in the peptide exit tunnel and prevents antibiotic binding when methylated [122]. However, methylation of this base also decreases overall cell fitness by deregulating translation of certain proteins [123]. Therefore, expression of the *erm* genes that encode these methyltransferases is inducible and highly regulated [122,123].

Induction of the most extensively studied *erm* gene, *ermC,* occurs when ERY binds the ribosome and induces translational arrest. The *ermC* mRNA contains an upstream ORF that codes for a 19-amino acid long leader peptide known as ErmCL (Figure 1-12A) [124,125]. Under normal conditions, the SD site of *ermC* is sequestered by mRNA secondary structures

prohibiting its translation. Its leader peptide, *ermCL*, is constitutively expressed. If, however, ERY binds the ribosome during ErmCL synthesis, it results in translational arrest, dissolving the downstream mRNA structure and activating *ermC* expression (Figure 1-11A) (for a review see [126]). In this manner, significant ERY concentrations within the cell induce the expression of the resistance gene *ermC*.

The ErmCL leader peptide induces stalling by the strategic positioning of the conserved stalling motif $IVFI^9$ within the exit tunnel. Mutational studies have demonstrated that this sequence is critical for stalling, with Ile9 stalling in the P site [96]. The identity of A site aminoacyl-tRNA is inconsequential, as studies demonstrate that peptidyl transfer from Ile9-tRNA to any A site amino acid or even puromycin is inhibited [96].

The stalling event is initiated through the binding of ERY within the peptide exit tunnel, near the PTC, partially blocking the exit tunnel entrance (Figure 1-12B and 12C) [127,128]. The C3-cladinose ring of ERY directly interacts with the I^6VFI^9 motif forcing the ErmCL peptide to compact against the exit tunnel. This compaction pushes the peptide into contact with A2062 of the 23S rRNA and Met82-Arg84 of L22. This causes A2602 to interact with A2503, a highly conserved, modified rRNA nucleotide that resides in the exit tunnel near the PTC (Figure 1-12C and 12D) [114]. Interactions between A2602 and A2503 are believed to initiate a signal relay ending with the inactivation of the PTC [96,114,115]. ERY binding may also directly contribute to PTC inactivation through contacts with rRNA residues like C2160 [129].

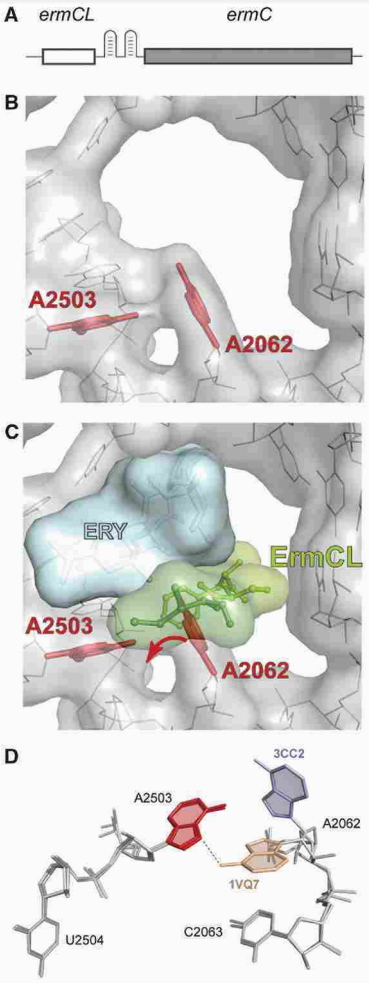


Figure 1-12. Ribsome Stalling by ErmCL. **(**A**)** The structure of the inducible *ermC* operon where the *ermC* gene is preceded by a regulatory ORF *ermCL*. Drug- and nascent-peptide-dependent ribosome stalling at ermCL ORF changes the conformation of the mRNA intergenic region (schematically shown as a two-hairpin structure), thereby releasing translational attenuation of *ermC*. **(**B**,** C**)** Erythromycin and the ErmCL nascent peptide in the ribosome exit tunnel (viewed from the PTC down the tunnel). In the vacant tunnel **(**B**)**, the nascent-peptide sensor, A2062, is free to rotate into the tunnel lumen. Binding of antibiotic ('ERY') narrows the tunnel (C). In the constricted tunnel, the ErmCL nascent peptide drives A2062 toward the tunnel wall, where it comes into close proximity to A2503. (D) Conformational flexibility of A2062. The orientations of the A2062 base are shown for the *apo* structure of the *Haloarcula marismortui* 50S ribosomal subunit (blue) (PDB accession number 3CC2) (see ref. (114)) and for the 50S subunit complexed with a transition state analog (beige) (1VQ7) (27). The A2503 base is colored red. A possible hydrogen bond between A2062 and A2503 is indicated by a dashed line. Used by permission from ref. (114).

TnaC Stalling

The final case study of gene regulation through ribosome stalling involves genes that degrade tryptophan in *E. coli* and *P. vulgaris*. The *tnaCAB* operon includes a tryptophanase (*tnaA*) that degrades Trp to indole, pyruvate and ammonia, effectively making Trp a source of carbon and nitrogen [130] as well as the tryptophan transporter *tnaB* [131]. Upstream of *tnaA* is a separate ORF encoding a 24 (*E. coli*) or 36 (*P. vulgaris*) amino acid long leader peptide called *tnaC* [131].

Bacteria are able to modulate Trp levels through a combination of transcriptional and translational controls. After transcribing *tnaC*, RNA polymerase pauses at the intergenic spacer

region between *tnaC* and *tnaA* (Figure 1-13A) [132]. Co-transcriptional translation is also occurring during this time period. When cellular levels of Trp are low, translating ribosomes reach the end of *tnaC*, terminate translation normally, and dissociate from the mRNA. This allows the Rho termination factor to bind to the mRNA and catch up to the transcriptional machinery and induce Rho-dependent termination before *tnaA* or *tnaB* can be transcribed [133]. When cellular levels of Trp are high, Trp binding to the ribosome during TnaC translation stalls the termination reaction [99]. The stalled ribosome blocks the Rho termination factor from binding the mRNA, allowing transcription of *tnaA* and *tnaB* to proceed (Figure 1-13A) [133,134]. In this way, bacteria are able to rapidly modulate levels of tryptophan within the cell.

Successful arrest of translation is dependent on free L-Trp binding and TnaC interactions with the exit tunnel. The binding site of L-Trp is proposed to overlap with the sparsomycin binding site and aminoacyl portion of the A-site tRNA [99,135–137]. Despite recent studies that implicated Ile19 of the TnaC peptide and A2058 (a PTC residue that sits at the entrance of the exit tunnel) in the creation of the binding pocket [138], the exact location of L-Trp could not be visualized in the cryo-EM structures, making the exact binding site still unknown [115].

Binding of free L-Trp causes the TnaC leader peptide to arrest translation. In *E. coli*, studies have determined that three conserved residues (Trp12, Asp16 and Pro24) are responsible for TnaC mediated arrest [135]. Another semi-conserved residue Ile19 also contributes to stalling making the consensus sequence in *E. coli* $W^{12}xxxD^{16}xxI^{19}xxxxP^{24}$stop [135]. With P^{24} in the P site, stalling occurs at termination. A similar conserved sequence, $W^{20}xxxD^{24}xxI^{27}xxxxP^{32}$, is found in *P. vulgaris* where two additional lysine residues after P^{32} indicate that stalling occurs during elongation [139]. It is interesting to note that despite having different peptide lengths, the spacing

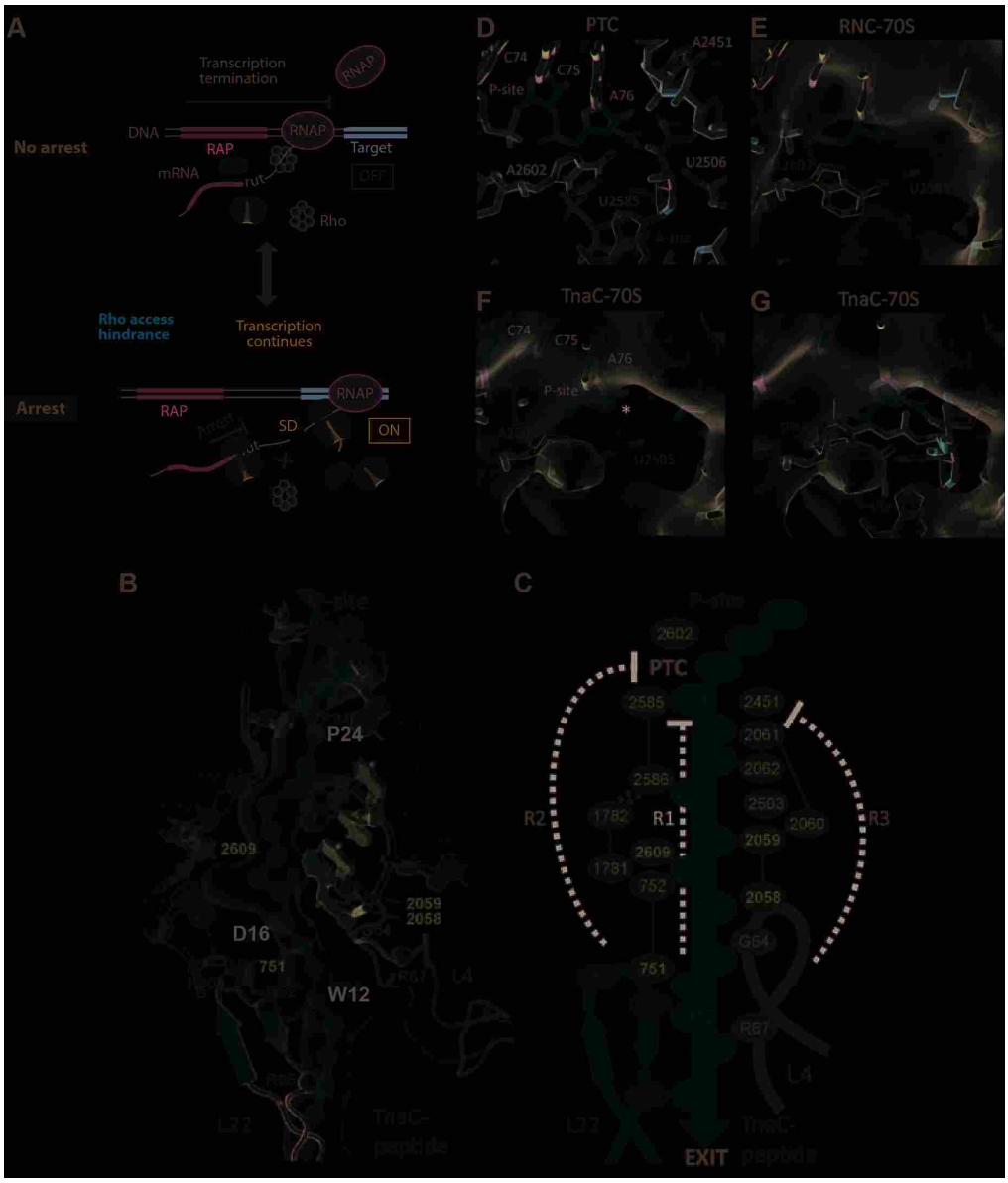


Figure 1-13. Ribosome Stalling by TnaC**.** (A) Nascent polypeptide-ribosome-mRNA complexes are schematically depicted to show access of a bacterial transcription termination factor. (B) Ribosomal components potentially involved in a relay mechanism to inactivate the PTC, with those implicated in stalling in bold. The TnaC nascent chain is in green, with residues essential for stalling colored yellow. The isolated TnaC-tRNA density is shown as a transparent gray surface. (C) Schematic indicating potential relay pathways from Trp^{12} (W12) of TnaC to the PTC, either through the nascent chain itself (R1) or through networks of interconnected 23*S* rRNA nucleotides (R2 and R3). (D) Conformation of 23*S* rRNA nucleotides at the PTC when tRNA CCA-end mimics are bound to A (cyan) and P sites (green) (PDB1VQN). (E) View into the PTC of 70*S*-RNC complex, with fitted models as in (D). Note the lack of density (gray) for nucleotide A2602. (F) View into the PTC of the TnaC-70*S* complex, with the MDFF model of the TnaC-tRNA (green) and nucleotides of the 23*S* rRNA (blue). The cryo-EM density is shown as a transparent gray surface, with an asterisk indicating the connection between P-tRNA and nascent chain. (G) As in (D), but with the antibiotic sparsomycin (SPAR, red; PDB1VQ9) and the terminal A76 and aminoacyl moiety of an A-tRNA (cyan, PDB1VQN) included. (A) Used with permission from ref. (87)). (B)-(G) From ref. (115). Reprinted with permission from AAAS.

between the stalling elements in *E. coli* and *P. vulgaris* is conserved. This spacing is essential for stalling in both organisms [139].

In *E. coli*, biochemical, structural and molecular modeling evidence suggests that consensus sequence residues form extensive contacts with amino acids and rRNA residues within the exit tunnel and near the PTC (Figure 1-13B) [94,115,135,140,141]. These include Trp12 interactions with R92 of L22, Asp16 interactions with K90 of L22, Ile19 interactions with A2058, A2059 and U2609 and Pro24 interactions with U2585 [141]. Interestingly, exit tunnel interactions by Trp12 prevent puromycin release of Pro24, suggesting a relay mechanism exists that transmits signals from the exit tunnel to inactivate the PTC (Figure 1-13C) [115,136]. This theory is supported by mutational studies showing that substitution of exit tunnel relay elements A748-A752 and U2609 of the 23S rRNA greatly diminish stalling capability [140].

Cryo-EM structures of the stalled TnaC-ribosome complex offer some insight as to how the PTC is inactivated in *E. coli*. Residues A2602 and U2585 of the PTC adopt altered conformations in the TnaC-ribosome complex (Figure 1-13E) [115] when compared to other ribosome structures (Figure 1-13F) [27]. During stalling with TnaC, A2602 adopts a rigid conformation as apposed its typical flaccid state [27,115]. This rigid conformation is reminiscent of the structures showing the ribosome bound with sparsomycin, a translational inhibitor (Figure 1-13G) [142]. U2585 is also seen to shift position to directly interact with Pro24, presumably an effect of L-Trp binding (compare Figures 13E and 13F). These conformational changes are incompatible with proper positioning of the GGQ motif of RF2, potentially explaining why stalling occurs at termination [115,143]. These changes must also prevent proper aminoacyl-tRNA positioning, since stalling also occurs during elongation in *P. vulgaris* [139] and substitution of a

Trp codon UGG for the UGA stop codon after Pro24 inhibits elongation and puromycin release in *E. coli* [99].

Overall, the stalling mechanisms used by SecM, ErmCL and TnaC have many similarities. First, both SecM and ErmCL regulate gene expression by sequestering downstream SD elements and use stalling as a means of dissolving those elements. All three facilitate stalling through strategic positioning of their residues to induce peptide compaction in the exit tunnel. Both TnaC and SecM target the constriction site in the exit tunnel using aromatic residues. TnaC and ErmCL both rely on small molecule binding near the PTC to initiate stalling. Finally, all three facilitate PTC inactivation via signal relay through peptide constriction near A2062. Despite these similarities, however, SecM, ErmCL and TnaC require different elements in the exit tunnel for stalling, specifically in portions of the proposed relay mechanisms, suggesting that multiple relay mechanisms likely exist [108,113–115].

Even though SecM, ErmCL and TnaC and the other known stalling peptide sequences have been studied thoroughly, much remains to be explored. What is the scope of gene regulation by peptide stalling? What are the combinations or patterns of amino acids that cause stalling? Are there universally conserved motifs? How do programmed stalling events evade tmRNA and other rescue machinery? Our efforts to answer these questions and to better understand peptide mediated stalling have led to the identification of novel stalling motifs and their characterization in both bacteria and yeast. The following chapters detail some of these results and highlight some of the new techniques that have been developed to identify novel stalling motifs.

Chapter 2. Genetic Identification of Nascent Peptides that Induce Ribosome Stalling

Author's Note: This chapter details the discovery and characterization of novel stalling peptides, for which I performed the toeprinting and peptidyl-tRNA release assays. The results of this study were published in the Journal of Biological Chemisty in 2009 [98]**.**

Abstract

Several nascent peptides stall ribosomes during their own translation in both prokaryotes and eukaryotes. Leader peptides that induce stalling can regulate downstream gene expression. Interestingly, stalling peptides show little sequence similarity and interact with the ribosome through distinct mechanisms. To explore the scope of regulation by stalling peptides and to better understand the mechanism of stalling, we identified and characterized new examples from random libraries. We created a genetic selection that ties the life of *E. coli* cells to stalling at a specific site. This selection relies on the natural bacterial system that rescues arrested ribosomes. We altered tmRNA, a key component of this rescue system, to direct the completion of a necessary protein if and only if stalling occurs. We identified three classes of stalling peptides: C-terminal Pro residues, SecM-like peptides, and the novel stalling sequence FxxYxIWPP. Like the leader peptides SecM and TnaC, the FxxYxIWPP peptide induces stalling efficiently by inhibiting peptidyl transfer. The nascent peptide exit tunnel and peptidyl-transferase center are implicated in this stalling event, although mutations in the ribosome affect stalling on SecM and FxxYxIWPP differently. We conclude that ribosome stalling can be caused by numerous sequences and is more common than previously believed.

Introduction

The ribosome efficiently synthesizes an enormous diversity of peptide sequences without regard to their chemical properties. This generality is not universal, however. Several polypeptides interact with the ribosome to stall their own translation, either in the elongation or termination steps [144,145]. Programmed stalling events regulate gene expression in both prokaryotes and eukaryotes [144–146], and may affect the folding of nascent polypeptides [147–149].

In two well-characterized examples from *E. coli*, ribosome stalling on a leader peptide increases the expression of a gene further downstream on the same mRNA. The secretion monitor peptide SecM, for example, regulates *secA* in response to changes in protein translocation activity [102]. If translocation activity is low, ribosome stalling on the SecM peptide alters the secondary structure of the mRNA and upregulates the translation of *secA*, a key component of the secretory machinery [106]. When activity is high, the SRP-Sec system binds the signal peptide sequence in SecM and pulls it from the stalled ribosome [106,107]. A second example is the regulation of *tnaA*, a gene required to break down tryptophan, by the leader peptide TnaC in response to Trp levels. When Trp concentrations are high, ribosome stalling on TnaC blocks a transcriptional terminator upstream of *tnaA*, increasing its expression [99,150]. Low tryptophan levels do not support ribosome stalling and lead to attenuation of the transcript.

Stalling on the SecM and TnaC peptides is the result of three interactions: the binding of the nascent peptide to the exit tunnel and the peptidyl-transferase center, and the binding of an effector in the ribosomal A site. The peptide exit tunnel in the large ribosomal subunit is 100 Å long and 15 Å wide on average [28]. Mostly made of RNA, it provides very few hydrophobic surfaces for elongating proteins to bind, accounting for their ability to pass through unhindered. A significantly constricted portion of the tunnel is formed by loops in proteins L4 and L22 [151,152].

SecM and TnaC interact with the tunnel near this constriction, using critical Trp residues 10-12 amino acids upstream of the stalling site. Ribosomal mutations that reduce stalling map to the exit tunnel, implicating A751, A2058, and U2609 in the 23S rRNA and specific residues in the L22 protein in the stalling mechanism [94,95]. A cryo-electron microscopy study of the SecM-stalled ribosome revealed a network of conformational changes in 23S rRNA emanating from the exit tunnel [153].

Nascent peptides also interact directly with the peptidyl-transferase center (PTC) to induce stalling. In the case of SecM, the identity of the final six residues is critical for stalling on the FxxxxWIxxxxGIRAGP$_{166}$ sequence [95]. Likewise, the C-terminal Pro residue in TnaC is essential for stalling on the sequence WxxxDxxxxxxxP* [99]. These amino acids must be acting within the PTC to inhibit its catalytic activity, either peptidyl transfer for SecM or peptidyl hydrolysis for TnaC. In some cases, the peptide sequence in the PTC is sufficient to induce stalling without exit-tunnel interactions. A C-terminal Pro residue in the YbeL protein inhibits termination. Peptide release is especially inefficient when Pro-stop is preceded an Asp, Glu, or Pro residue [100].

In addition to nascent-peptide interactions with the exit tunnel and the PTC, stalling on SecM and TnaC requires a specific effector molecule to bind in the A site. This binding event is thought to create a PTC conformation that is inactive. For example, SecM stalls during elongation with unreacted Pro-tRNA bound in the A site [116,154]. Mutation of this Pro codon to Ala alleviates stalling. Likewise, TnaC stalling requires the binding of free tryptophan at an unknown site near the PTC [136,137]. The action of free tryptophan can be mimicked by Trp-tRNA if the *tnaC* stop codon is mutated to a Trp codon. Other aminoacyl-tRNAs (Phe, Met, Pro) do not induce stalling on TnaC [99].

Although stalling peptides interact with the exit tunnel and PTC, they do so differently and share little sequence similarity. This led us to hypothesize that there are additional, unknown peptide sequences that might inhibit peptidyl transfer or hydrolysis. Here, we report the genetic identification and characterization of peptides that stall at high efficiency during elongation.

Results

A genetic selection for novel stalling sequences.

We set out to systematically identify peptide sequences like SecM and TnaC that interfere with peptidyl transfer or hydrolysis and induce ribosome stalling. To identify stalling peptides from random libraries, we modified a genetic selection that we previously developed to link ribosome stalling and rescue to the life of the cell [155]. In this selection, stalled ribosomes are recognized by transfer-messenger RNA (tmRNA), a small, stable RNA found in eubacteria and a key component of a quality control system for protein synthesis; for a review, see [69]. tmRNA's natural function is to release stalled ribosomes and tag the aborted nascent peptide for destruction. Acting as a transfer RNA, tmRNA enters the empty A site of the ribosome and adds Ala to the nascent polypeptide chain. tmRNA then serves as a template, encoding a short peptide tag that is recognized by cellular proteases. After this tag is translated, the ribosome is released at a stop codon within tmRNA and the aborted protein product is degraded. For the purposes of our selection, it is important to note that although tmRNA was first characterized as rescuing ribosomes stalled on mRNAs lacking stop codons [81], it can also act on ribosomes stalled by nascent peptides [100,156].

To create a genetic selection for ribosome stalling based on this ribosome rescue machinery, we altered tmRNA so that instead of tagging proteins for proteolysis, it completes the

synthesis of an essential protein, linking stalling to the life of the cell (Figure 2-1). The kanamycin resistance protein (KanR) from Tn10 has a C-terminal helix of 15 amino acids that is structurally critical [157]; truncation of this helix leads to loss of activity. To complement the truncated KanR protein, we changed the tmRNA template to encode the last 14 residues of KanR (ANKLQFHMLDEFF), referred to hereafter as *tmRNA-K1*. Together with the Ala from aminoacylated-tmRNA, these residues complete the KanR protein and restore KanR activity—but only if the ribosome stalls at exactly the right site. This serves as the basis for our selection: peptide sequences that stall the ribosome at the end of a truncated KanR protein can be easily identified in random libraries because they recruit tmRNA, complete KanR, and confer resistance to kanamycin.

How can stalling be induced at the end of the KanR protein without interfering with the final structure and activity of KanR? Two mutations, Asn255Glu and Asp257Opal, create a Glu-Pro-(Stop) sequence that induces stalling during translational termination in a *kanR* gene lacking the C-terminal helix. We previously showed that expression of this truncated *kanR-EP* construct and the altered tmRNA, *tmRNA-K1*, allows cells to survive equally well on selective (15 μg/mL kanamycin) or non-selective plates at 37 °C [155]. Under the same conditions, bacteria lacking the modified *tmRNA-K1* gene survive at the rate of 5 cfu in 10^7 plated. These results demonstrate that the introduction of the Glu-Pro-Ala "scar" from the stalling and tagging process does not destroy KanR activity. Analysis of the crystal structure of the homologous Aph(3')-IIa protein [157] suggests that the C-terminal helix in KanR is preceded by a surface-exposed loop of poorly conserved residues (Ile253 through Pro256). We anticipated that this loop region might tolerate a variety of sequences that induce stalling while maintaining robust KanR function.

Nucleotide sequences that induce stalling and tagging were isolated from a randomized library fused to a truncated *kanR* gene. 18 amino acids were deleted from KanR, including the C-terminal helix and three residues in the preceding loop. 18 random nucleotides (six codons) were cloned downstream of this truncated *kanR* gene, beginning with residue 254. No stop codon was specified. We generated a library of 5 x 10^6 mutants and introduced it, together with *tmRNA-K1*, into an *E. coli* strain lacking wild-type tmRNA. We selected for survival on plates containing 15 µg/mL kanamycin at 37 °C. Roughly 1 in 10^4 colonies survived, suggesting that a substantial fraction of the sequences induce ribosome stalling.

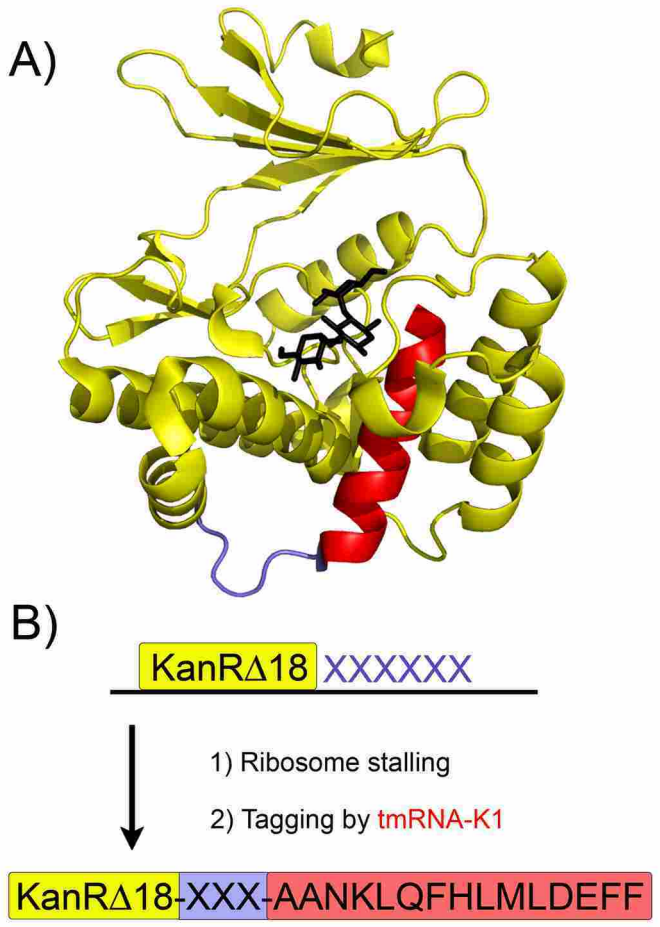


Figure 2-1. Genetic selection for sequences that induce ribosome stalling. (A) Structure of the Aph(3')-IIa protein, homologous to KanR (see ref. 98). Kanamycin in shown in black. The C-terminal helix (14 residues), shown in red, is essential for KanR activity. A four-residue surface-exposed loop prior to this helix is highlighted in blue. (B) 18 random nucleotides (six codons) were introduced at the C-terminus of a KanR protein lacking its last 18 amino acids (yellow). If the random sequence induces stalling, tmRNA-K1 rescues the stalled ribosome and directs the synthesis of the remaining KanR residues (red). Cellular survival on kanamycin plates is therefore tied to stalling on the C-terminus of KanR.

Three classes of stalling peptides.

Although we were interested primarily in peptides that induce stalling, our selection identifies any nucleic acid sequences that elicit tagging by tmRNA. In principle, nucleotide sequences containing rare codon clusters [158,159], secondary

structures, transcriptional terminators [81], or other novel mechanisms might also survive the KanR selection. Analysis of the surviving clones, however, revealed sequences that share common features at the amino acid level. These were grouped into three classes (Table 2-1).

The most common cause of stalling, found in over 90% of the clones, is inefficient termination at the sequence Pro-Stop. The Pro residue is found almost exclusively at position three of the six random codons, corresponding to native KanR residue Pro256. While there is no significant codon bias for any particular Pro codon, the opal stop codon (UGA) is highly overrepresented (23/29 clones). There is also selection for the residue just upstream of Pro-Stop: Glu is overrepresented (16/29) and Asp, Pro, and Gly are each seen several times in the –2 position.

A second class of peptides must induce stalling during elongation, not termination. These clones contain two consecutive Pro codons, most commonly at codons three and four in the random sequence with no nearby stop codon. The majority of these clones were found by performing the selection at lower stringency, lowering the temperature to 25 °C. When tested individually, they showed poor survival at 37 °C, roughly 1-10%, much weaker than the 100% survival seen with the Pro-Stop sequences above.

Class I: Pro-stop	G I **S E P *** (opal)
	G I **K D P *** (amber)
	G I **W P P *** (opal)
Class II: Pro-Pro	G I **R A P P H C** G T
	G I **R S P P N S** G T
	G I **L D P P G M** G T
Class III: Trp-Pro-Pro	G I **W P P W Y R** G T
	G I **W P P D V *** (opal)
	G I **W P P P S I** G T

Table 2-1. Stalling peptides isolated from the KanR selection. The variable region (6 codons) is shown in bold. Class I contains Pro-Stop residues at the third and fourth variable position. Class II contains a Pro-Pro sequence at positions three and four with no nearby stop codon. Class III contains Trp-Pro-Pro at the first three positions without an adjacent stop codon.

A third class of clones contain the sequence Trp-Pro-Pro without a nearby stop codon (Table 2-1). Like the other Pro-Pro sequences, these peptides also stall during elongation rather than termination. Unlike the second class of clones, however, where the two consecutive Pro codons appear at positions three and four, here they occur at positions two and three (e.g. WPPWYR). Another difference is that WPP-containing clones survive robustly (100%) in the KanR selection at 37 °C when characterized individually. Further experiments on these sequences are described below.

The sequence Pro-Stop occurs commonly and elicits tagging at high levels; to prevent such clones from overwhelming other novel sequences, we created a second library of 18 nt (six codons) in which codons four through six could not be stop codons. This was done by allowing only C, G, and A at the first nucleotide of these codons; this eliminates Phe, Tyr, Cys and Trp as well. We screened an 8 x 10^6-member library at high stringency, obtaining colonies at rates of 0.01% survival. 21/23 sequenced clones contained the sequence WPPP at the first four positions (data not shown). This result confirms that WPP-containing sequences are robust inducers of stalling, particularly when coupled with a third Pro codon.

Selection of this second library at low stringency yielded higher levels of survival (0.25%). Nearly all of surviving clones fall into the second class of stalling peptides, with two consecutive Pro codons at codons three and four of the random sequence. An alignment of 46 of these sequences reveals that Arg or His are strongly preferred at the first position, with Ala, Asp, Ser, and Pro at the second position (Figure 2-2). Including the constant Gly-Ile upstream, the consensus sequence becomes GI(R/H)xPPxx. These appear to be weaker versions of the SecM C-terminal sequence GIRAGP. A clone closely resembling this sequence (GIRAPP) is more

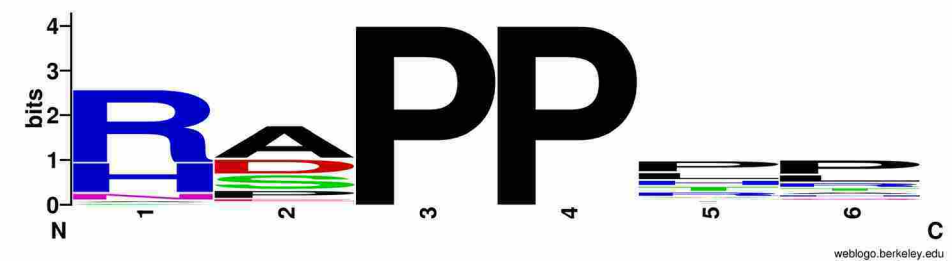


Figure 2-2. Alignment of the variable region of clones surviving the KanR selection at low stringency. Created by Weblogo. Position 1: R > H > n. Position 2: A > D > S > P. Positions 3 and 4 were Pro only. Positions 5 and 6 showed little conservation

active than the other members of this class and survives the KanR assay 100% at high stringency.

Stalling and tagging occur following WPP.

The peptide sequences in class three (containing WPP) show high levels of activity in the KanR assay and stall translation during the elongation step. We chose to further characterize three sequences: WPPPSI, WPPDV*, and WPPWYR. Where does stalling occur in these sequences? Where is the tag added by tmRNA? To analyze the tagged proteins by mass spectrometry, we first transferred the stalling sequence to the C-terminus of the GST protein. This full-length, stable protein served as a scaffold enabling overexpression of the stalling peptide. Some of the KanR protein context was fused to GST as well, from 12 amino acids upstream of the random hexamer through the stop codon 27 codons downstream. To isolate proteins tagged by tmRNA, we used a modified tmRNA encoding six His residues in its template sequence (*tmRNA-H*). GST-fusions tagged by *tmRNA-H* were purified by affinity chromatography using Ni-NTA resin and digested with trypsin. From this tryptic digest, the C-terminal tagged peptide was purified again with Ni-NTA resin.

The C-terminal peptide contains both the stall sequence from KanR and the tmRNA tag; determining its mass by MALDI-MS revealed the site of stalling and tagging by tmRNA. A

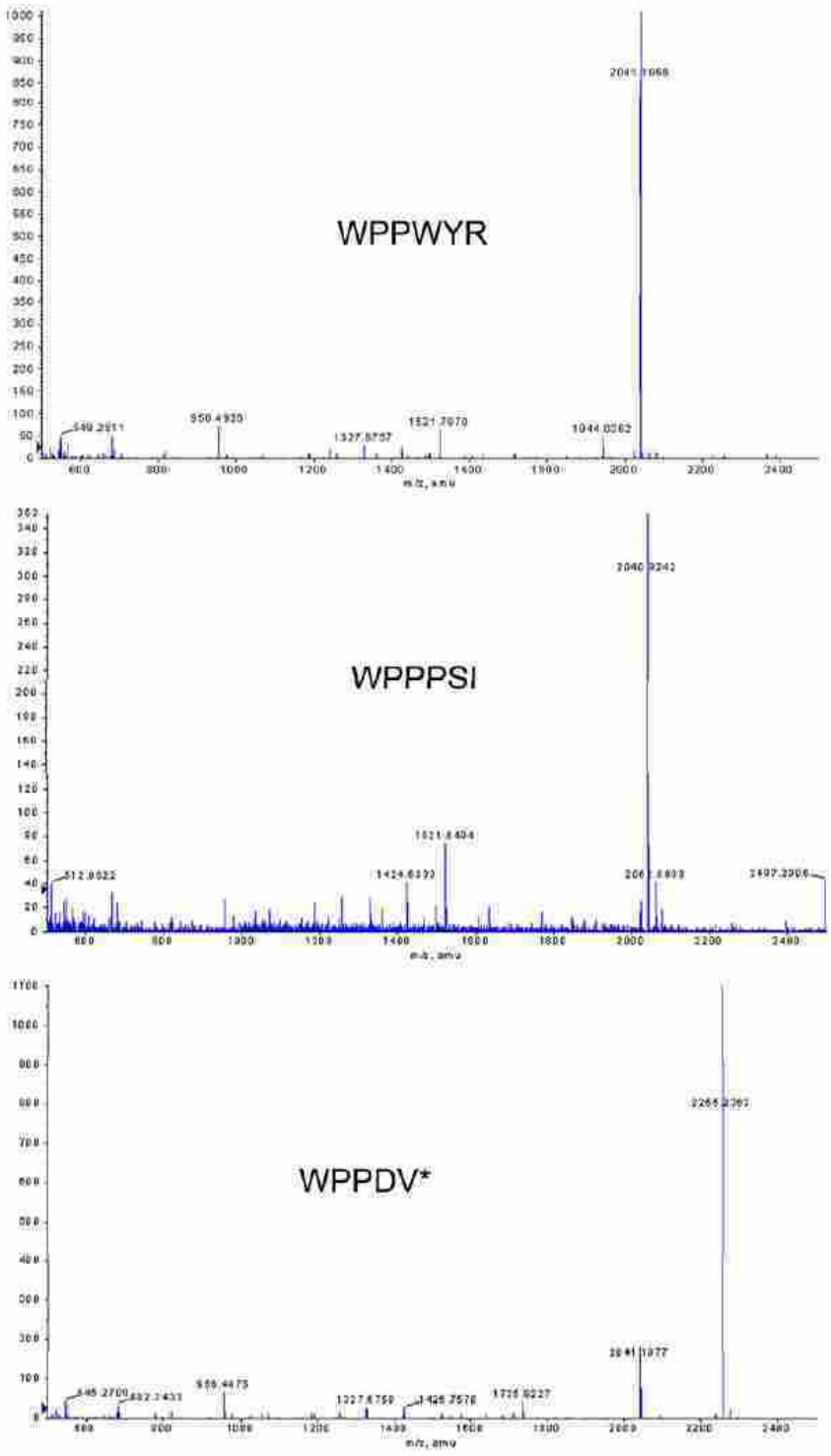


Figure 2-3. Identification of the site of stalling by MS**.** WPPWYR, WPPPSI, and WPPDV* were cloned with 12 upstream residues onto the C-terminus of GST. Stalled peptides were tagged by tmRNA-H encoding a His_6-tag. Following Ni-NTA resin purification, the tagged proteins were digested with trypsin and the C-terminal tagged peptide was repurified on Ni-NTA resin. These peptides were ionized by MALDI and analyzed by MS. All three stalling sequences produce a peak at m/z 2041, corresponding to YGIWPPAANDH_6D. The WPPDV* sequence also contains a strong peak at m/z 2255 corresponding to YGIWPPDVAANDH_6D.

single large peak in the mass spectra for the WPPPSI and WPPWYR C-terminal tagged peptides corresponded to a m/z of 2041 (Figure 2-3). This is the predicted result if the tmRNA tag is added after the second Pro (YGIWPPAANDH$_6$D). The mass spectrum of the WPPDV* peptide fragment contained the same peak at 2041 together with a more abundant peak at 2255, corresponding to the peptide YGIWPPDVAANDH$_6$D. In the WPPDV* clone, stalling occurs both after WPP and during termination at the stop codon. Peptide fingerprinting by tandem MS/MS was performed on all four of these peptides to confirm the amino acid sequence directly.

Determination of residues necessary and sufficient for stalling and tagging.

The MS data indicate that tagging occurs immediately after WPP in these three clones. What amino acids cause this stalling event? In the case of SecM and TnaC, residues essential for the highest levels of stalling are found upstream and interact with the exit tunnel. For this reason we included 12 upstream amino acids (SLQKRLFQKYGI) from KanR along with the hexamers in making the GST-fusions. To assay for stalling and tagging in the GST-fusions, we detected the tag added by *tmRNA-H* with anti-His$_6$ antibodies (Figure 2-4). High levels of tagging were detected for the full-length GST-WPPPSI fusion, referred to hereafter as 1-18 (i.e. 12 residues from KanR followed by the hexamer, Figure 2-4A). Deletion of the first four amino acids had little or no effect (5-18), but removal of the first eight nearly eliminated tagging (9-18, Figure 2-4B). We conclude that residues upstream of the WPPPSI sequence play a critical role in high-efficiency tagging. Interestingly, some minimal activity resides in the hexamer sequence alone (13-18) with no KanR upstream sequence.

To identify how each residue contributes individually to stalling, we performed alanine scanning on the full-length stalling peptide, 1-18. Residues 1 to 16 were individually mutated to

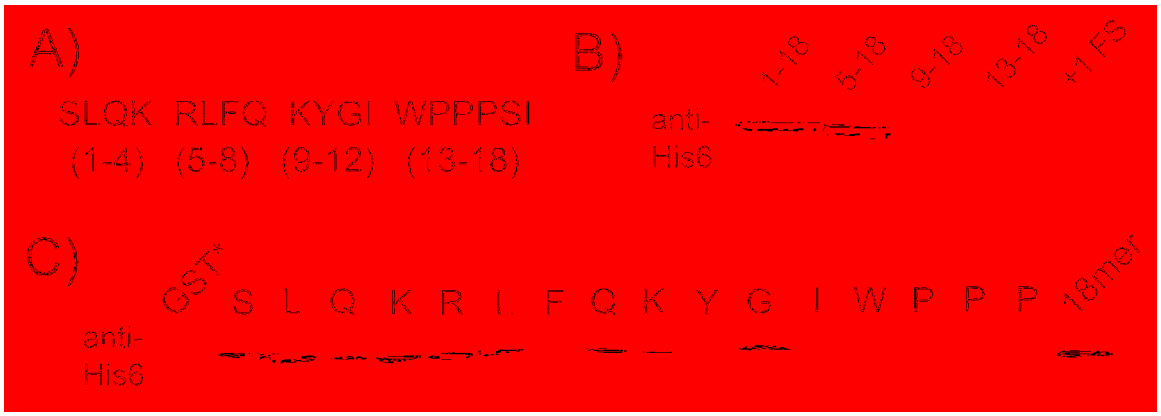


Figure 2-4. Essential elements of a stalling peptide**.** (A) WPPPSI and 12 upstream residues were divided into four groups for analysis. (B) N-terminal deletions of the 18-mer stalling peptide were fused to the C-terminus of GST. The fusions were analyzed by immunoblot for tagging by a modified tmRNA (tmRNA-H) that encodes a His_6-epitope. The +1FS vector contains the same 18-mer sequence shifted into the +1 frame to test if translation of WPPPSI is necessary for stalling. (C) Each residue of the 18-mer was mutated individually to Ala and assayed as above. A GST-Stop construct served as a negative control and the intact full-length WPPPSI 18-mer as a positive control.

alanine and assayed by immunoblot (Figure 2-4C). Consistent with the truncation results, mutating residues 1-4 has little or no effect on tagging. Alanine substitutions for Arg5, Leu6, Gln8 and Lys9 likewise make little difference in the level of tagging. In contrast, mutation of Phe7, Tyr10, Ile12, or the WPP sequence dramatically decreases tagging levels. Notably, tagging is also strongly reduced by the Pro16Ala mutation. This is surprising because the MS data shows that the third Pro in WPPPSI is not incorporated into the stalled peptide. In summary, residues in the consensus sequence FxxYxIWPPP are required for tagging.

tmRNA rescues ribosomes stalled on broken mRNA templates; perhaps these tagging events arise from RNA synthesis defects in the *kanR* mRNA or from nucleolytic cleavage. To prove that tagging requires translation of the peptide sequence, we created a mutant of the GST-WPPPSI fusion in which a single nucleotide is added upstream of the full-length stall peptide. The resulting +1 frameshift changes the identity of every amino acid in the stalling sequence except for Phe7 and Lys9 while retaining the same nucleotide sequence. Immunoblot analysis of this mutant revealed that tagging was completely abolished, demonstrating that tagging of the

GST-WPPPSI fusion is due to the amino acid sequence and not the nucleotide sequence (Figure 2-4B).

Tagging at termination in WPPDV.*

The MS data show that tagging occurs in the WPPDV* sequence both immediately after WPP and during termination. To further understand the effect on termination, we measured tagging levels for a series of GST-WPPDV* variants in the immunoblot assay (Figure 2-5). Mutation of the opal stop codon (UGA) to the more efficient ochre codon (UAA) reduced tagging slightly; replacing the stop codon altogether with an Ala codon reduced it even further. We propose that the substantial tagging that remains in the WPPDVA variant represents stalling directly after the WPP as seen in the MS data.

If the WPPDV sequence is interfering with termination, how far downstream does this effect carry? An opal stop codon immediately following WPPD tagged at the same level as the original WPPDV* sequence (Figure 2-5). Moving the stop codon one or two codons downstream by inserting Ala residues, however, reduces the tagging levels to those lacking a stop codon altogether (WPPDVA). These results show that the stop codon must be only one or two codons downstream of WPPD for stalling to occur during termination.

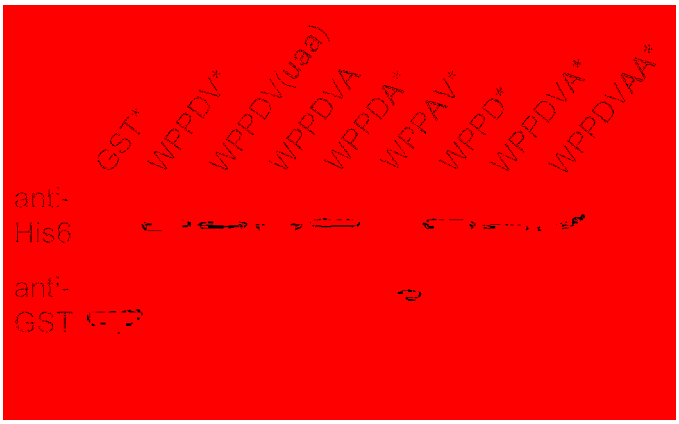


Figure 2-5. Stalling during termination at WPPDV*. Tagging of the GST-WPPDV* fusion by tmRNA-H was monitored by anti-His_6 antibodies. Mutations in bold were introduced to determine the role of spacing, stop codons, and the DV residues in this stalling event.

We next examined the role of the Asp and Val

residues. Val is not known to inhibit termination when found at the C-terminus of proteins; indeed, the Val17Ala mutant showed no loss of tagging. The Asp16Ala mutation, however, completely alleviated tagging (WPPAV*). The Asp residue must therefore be critical for tagging after WPP as well as after WPPDV during termination. This role is consistent with the critical nature of the third Pro residue in the WPPPSI clone.

The residue after WPP is critical for tagging by tmRNA.

Alanine-scanning mutagenesis of the WPPDV* and WPPPSI clones demonstrates that the residue after FxxYxIWPP plays a key role in stalling ribosomes. What amino acids besides Asp and Pro can fulfill this role? We created a library of peptide trimers following WPP in the KanR selection (WPPXXX), constrained as above to exclude stop codons. Of the clones surviving at high stringency, ~80% contain the sequence WPPPxx and another ~20% the sequence WPPDxx (data not shown). No selection was apparent for the final two amino acids. To perform a more quantitative analysis, we created mutants of the GST-WPPPSI fusion expressing all 20 amino acids in the position immediately following WPP. These were subjected to immunoblot analysis with *tmRNA-H*. Confirming the genetic data, the Pro, Asp, and Trp mutants showed high levels of tagging. The Asn mutant tagged moderately, while the other 16 amino acids showed much lower levels of tagging (Figure 2-6).

How does the residue after FxxYxIWPP contribute to stalling and tagging? The incoming aminoacyl-tRNA does not react, yet its identity is critical. It must therefore contribute to stalling by interacting with the ribosome in the A site. During ribosome stalling on SecM, Pro-tRNA performs exactly this function [154].

But if Pro-tRNA is bound in the A site, how is the GST-WPPPSI fusion tagged by tmRNA? It should block tmRNA from entering the ribosome at all. One possibility is that ribosomes stalled on FxxYxIWPPP deplete Pro-tRNA from the available cellular pool, leading to some stalling events with empty A sites. Depletion of Pro-tRNA leads to tagging of the SecM peptide by tmRNA [154]. If this is also the case with FxxYxIWPPP, then overexpression of $tRNA^{Pro}$ should alleviate tagging. To test this hypothesis, we altered the GST-WPPPSI fusion to include one or more CCC codons. CCC is decoded by only one tRNA, Pro2, which also recognizes CCU. The original WPPPSI sequence contains neither CCC nor CCU; we altered it to include CCC at the first two Pro codons (WppPSI) or the third (WPPpSI).

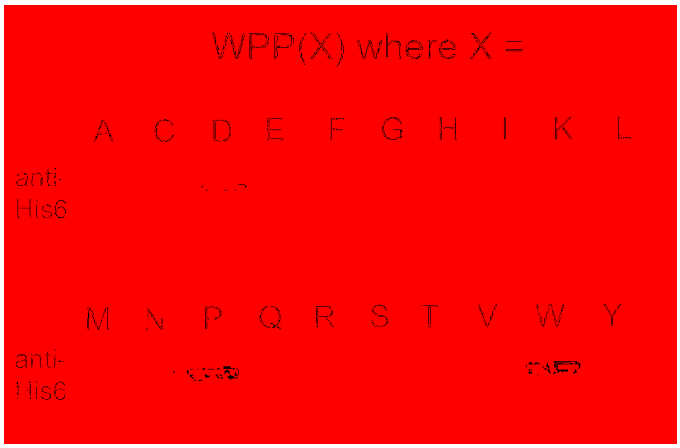


Figure 2-6. Identification of the site of stalling by MS. WPPWYR, WPPPSI, and WPPDV* were cloned with 12 upstream residues onto the C-terminus of GST. Stalled peptides were tagged by tmRNA-H encoding a His_6-tag. Following Ni-NTA resin purification, the tagged proteins were digested with trypsin and the C-terminal tagged peptide was repurified on Ni-NTA resin. These peptides were ionized by MALDI and analyzed by MS. All three stalling sequences produce a peak at m/z 2041, corresponding to YGIWPPAANDH$_6$D. The WPPDV* sequence also contains a strong peak at m/z 2255 corresponding to YGIWPPDVAANDH$_6$D.

The immunoblot assay was used to visualize the tagging levels of these GST-fusions with or without overexpression of Pro2 tRNA from the pRARE plasmid (Figure 2-7A). Tagging of the GST-WPPPSI fusion lacking CCC codons was unaffected by overexpression of Pro2. Likewise, little or no change in tagging occurred when the first two Pro residues were encoded by CCC (WppPSI). In contrast, when the third Pro codon was CCC, tagging was sharply reduced by Pro2 tRNA overexpression. In addition to the loss of tagging, the overall expression of the

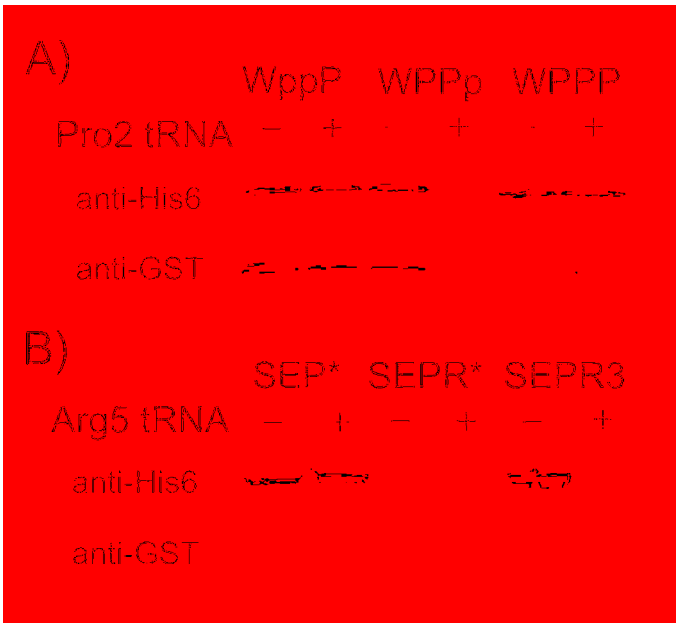


Figure 2-7. The effect of tRNA levels on stalling and tagging. (A) One or more Pro codons in the WPPPSI sequence was switched to CCC (lower case *p*) so that it would only be recognized by Pro2 tRNA. An upper case *P* represents a Pro codon not recognized by Pro2 tRNA. Tagging of the GST-WPPPSI fusion was monitored in the presence or absence of a plasmid (pRARE) overexpressing the Pro2 tRNA. pRARE also expresses several other rare tRNAs. (B) Stalling sequences containing one or three rare Arg codons (AGG) were fused to the C-terminus of GST. Tagging by tmRNA-H was monitored with anti-His_6 antibodies in the presence or absence of a plasmid (pRARE) overexpressing the cognate tRNA, Arg5.

GST-WPPpSI fusion was dramatically reduced. Pro2 overexpression had no effect on GST levels in WPPPSI or WppPSI fusions. These results show that depletion of the tRNA decoding the third Pro codon is necessary for tagging.

The role of codon usage.

We anticipated at the outset of our KanR selection experiments that we might isolate stalling sequences with rare codons. Overexpression of proteins containing consecutive rare codons induces high levels of stalling and tagging by tmRNA [158,159]. The three $tRNA^{Arg}$ isoacceptors decoding the CGG, AGA, and AGG codons are present at low levels in *E. coli* [160]. Why do such sequences not survive the KanR selection? To address this question, we measured tmRNA tagging levels for a GST-fusion construct containing SEPR* and SEPRRR encoded by the rare Arg codon AGG. SEPR* tagging was barely detectable, much lower than SEP*, while SEPRRR tagged at very high levels in the immunoblot assay (Figure 2-7B). Tagging at both sequences was completely alleviated by overexpression of the cognate tRNA (Arg5) from the pRARE plasmid. The same SEPR* and SEPRRR sequences were then cloned in place of the randomized cassette of the KanR selection

plasmid. The sequence SEP in the first three positions is known to be compatible with KanR activity; in the SEP* context it conveys 100% survival. Cells expressing SEPR* or SEPRRR sequences survived no better than an empty vector control under low stringency conditions (data not shown). These results show that tagging activity at rare codons is either insufficient or incompatible with restoring KanR function.

Direct detection of stalled ribosome complexes.

Both the KanR and the *tmRNA-H* immunoblot assays rely on tmRNA tagging to measure levels of ribosome stalling. To analyze stalling directly, we performed in vitro translation reactions and detected stalled ribosome complexes with toeprinting assays. Peptides corresponding to the C-terminal 64 residues of the GST fusions described above were expressed in a cell free transcription and translation system. The protein sequence includes the full 18-mer stalling peptide, 22 residues of upstream GST sequence, and 24 residues downstream of the predicted stalling site. A radiolabeled primer was annealed to the 3'-end of the transcript and extended by reverse transcriptase. Analysis of the FxxYxIWPPP peptide translation reaction revealed that reverse transcriptase is blocked 15-16 nt downstream of the first nucleotide in the second Pro codon (Figure 2-8). In contrast, no toeprint was seen in the translation of the FxxYxIWPAP peptide, consistent with the finding that mutation of the second Pro codon dramatically reduces tagging (Figure 2-4C). As a control, the antibiotic thiostrepton was added to trap ribosomes in the initiation stage. The disappearance of the toeprint in the FxxYxIWPPP peptide reaction when thiostrepton is added demonstrates that the block in reverse transcription is due to stalled ribosomes and not an artifact of mRNA sequence or structure. Together with the mass spectrometry and immunoblot analyses of tmRNA tagging above, the toeprinting data

demonstrate that ribosomes stall at the FxxYxIWPPP sequence with the second Pro codon in the ribosomal P site and the third Pro codon in the A site.

Ribosome stalling leads to the accumulation of peptidyl-tRNA within the ribosome. We detected this trapped peptidyl-tRNA by including [^{35}S]-methionine in the translation reaction. To prevent hydrolysis of the peptidyl-tRNA, we analyzed the products by gel electrophoresis in an acidic buffer system. Three lines of evidence support the identification of stalled peptidyl-tRNA in the FxxYxIWPPP peptide translation reaction. First, the high molecular weight band disappears upon treatment with RNase and a far smaller band appears (lanes 1 and 3). Secondly, the peptidyl-tRNA band remains in the aqueous layer following phenol extraction, while the other peptide bands disappear (lane 5). Finally, the stalled peptidyl-tRNA is less

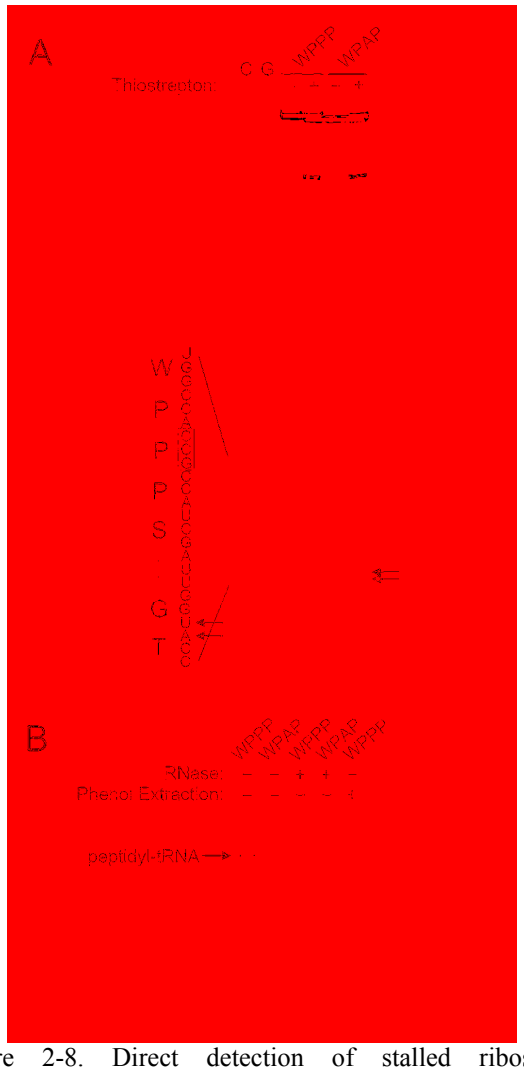


Figure 2-8. Direct detection of stalled ribosome complexes**.** (A) Stalled ribosome complexes were formed by cell-free translation of a template encoding the FxxYxIWPPP sequence in a larger (64-mer) peptide. The non-stalling Ala mutant FxxYxIWPAP served as a negative control. The position of the ribosome was determined by reverse transcription of the mRNA template and C and G sequencing lanes were run alongside. Thiostrepton was added in lanes four and six to trap the ribosome in the initiation stage, demonstrating the observed toe-print signal in lane three (marked by arrows) requires translation of the stalling site. The nucleotide and peptide sequence of the stalling site is shown at left. (B) The [^{35}S]-Met labeled products of the cell-free translation of the FxxYxIWPPP peptide or the non-stalling FxxYxIWPAP control were analyzed by Tricine-SDS PAGE. Under these conditions, the peptidyl-tRNA linkage is not hydrolyzed during electrophoresis. The stalled peptidyl-tRNA disappears when treated with RNase (lane 3) but remains in the aqueous layer upon extraction with phenol (lane 5).

abundant in the FxxYxIWPAP peptide reaction (lane 2), where it is not expected to accumulate as stalling is dramatically reduced.

Ribosomal interactions necessary for stalling.

To better understand the interactions between the nascent peptide and the ribosome that lead to stalling, we quantified stalling levels with a series of ribosome mutants. This was done by inserting the WPPPSI 18-mer after residue nine of *lacZ* and assaying for the activity of β-galactosidase. Our stalling peptides were compared to SecM and a non-stalling SecM control that has an Ala substitution of the C-terminal Pro. As shown by Nakatogawa and Ito [95], the SecM peptide dramatically inhibits *lacZ* expression; β-galactosidase activity is 1300-fold higher in the non-stalling Pro166Ala mutant (Table 2-2). The FxxYxIWPPP peptide also reduced *lacZ* expression, though not as well as SecM (116 versus 8 Miller Units, respectively). Mutation of the second Pro residue in FxxYxIWPPP results in 96-fold higher LacZ activity (Table 2-2), as expected by the reduction in tagging observed above. These results show that this selected peptide sequence induces stalling with high efficiency in a tmRNA-independent assay.

Stalling Peptide	LacZ activity
FSTPVWISQAQGIRAGP	8.2 +/– 2.0
FSTPVWISQAQGIRAG**A**	10812 +/– 1288
SLQKRLFQKYGIWPPPSI	117 +/– 14
SLQKRLFQKYGIWP**A**PSI	11172 +/– 3580

Table 2-2. Efficiency of stalling of the SecM and FxxYxIWPPP peptides**.** The peptide sequences shown were inserted into the full-length *lacZ* gene following the ninth codon. β-galactosidase activity is shown in Miller Units along with the standard deviation. The Ala substitutions in bold are known to prevent stalling in SecM (top) and FxxYxIWPPP (bottom).

Ribosomal RNA mutations that map to the exit tunnel have been shown to affect stalling on SecM and TnaC. Does the FxxYxIWPPP peptide interact with the same ribosomal RNA

nucleotides? Using β-galactosidase assays, we measured the effect of several 23S rRNA mutations on stalling on this peptide. 23S rRNA mutants were overexpressed in the presence of wild-type ribosomes. Stalling was reduced seven-fold by both the U2609A and U2609C mutations (Figure 2-9), first studied in connection with TnaC. The U754A and A751 insertion mutations, in contrast, showed no significant effect. Surprisingly, the A2058G mutation actually increases stalling eight-fold.

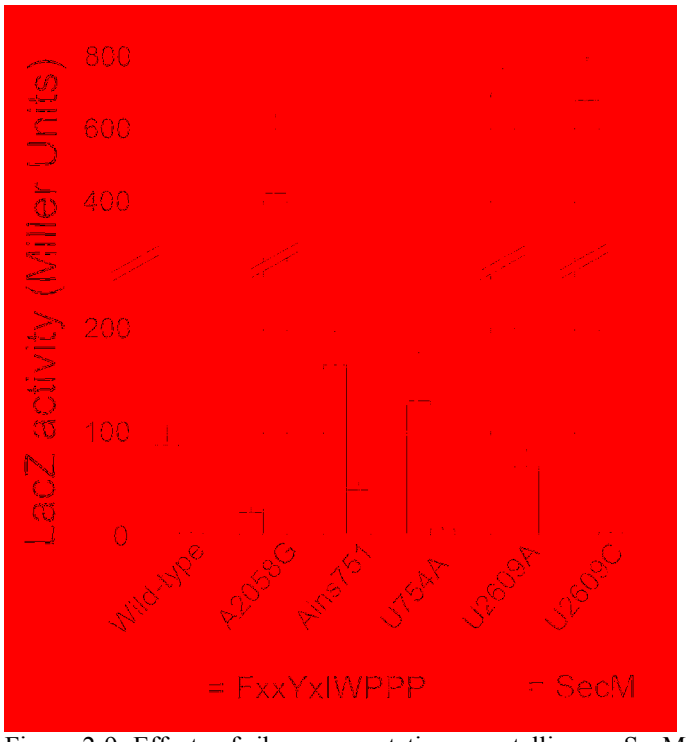


Figure 2-9. Effects of ribosome mutations on stalling on SecM and FxxYxIWPPP. The SecM stalling sequence or the WPPPSI 18-mer (see Table 2-2) were inserted after residue nine of the full-length *lacZ* gene. β-galactosidase activity was measured for the resulting SecM (white) and FxxYxIWPPP (grey) lacZ fusions in a strain overexpressing mutant 23S ribosomal RNA. The activity is reported in Miller Units. The data represent at least three independent experiments. Error bars represent the standard deviation.

Analysis of SecM-mediated stalling with the same set of rRNA mutants yielded quite a different picture. Although the U2609A mutation reduced stalling moderately (eight-fold), as it did with FxxYxIWPPP, the U2609C mutation had little or no effect. Although the A751 insertion had no effect on FxxYxIWPPP, this mutation decreased stalling on SecM six-fold. The most striking difference, however, is that the A2058 mutation increases stalling on FxxYxIWPPP but is the most effective at reducing stalling on SecM (78-fold), consistent with the findings of Nakatogawa and Ito [95]. These results show that while stalling on FxxYxIWPPP involves some of the same rRNA

nucleotides as SecM or TnaC, a unique pattern of exit tunnel interactions is required for each peptide.

Discussion

We performed a genetic selection to identify novel peptides that inhibit their own synthesis. The selection is based on the ability of tmRNA to recognize and rescue stalled ribosomes. When stalling occurs at the C-terminus of a truncated KanR protein, tmRNA encodes the missing amino acids to complete the protein and restore KanR activity.

Our library covered roughly 10% of the theoretical diversity of a library of random peptide hexamers. We recognize that some peptides that induce stalling were missed in our selection because they were either too long or incompatible with the structure and activity of KanR. We were surprised that consecutive rare codons, known to induce tagging [158,159], were not isolated in the selection. We demonstrated that tagging does occur at SEPR* and SEPRRR by immunoblot (Figure 2-7), but these sequences do not support KanR rescue by tmRNA. In the case of SEPR*, tagging is probably at too low a level to support robust KanR activity. While SEPRRR induces higher levels of tagging, the tag is probably not added at precisely the necessary site to restore the KanR protein sequence properly. Alternatively, depletion of low abundance tRNAs may be too taxing for cells. Immunoblot analysis of tagging is performed after a brief period of strong overexpression. In contrast, our genetic selection requires overexpression and tagging of KanR over long periods of cell growth and division.

The simplest and most common cause of stalling that we identified is inefficient termination at Pro-Stop sequences. Several components need to be present to cause high-efficiency stalling during termination. First, the opal stop codon (UGA) was strongly preferred

over the other two stop codons in the selection. UGA is the least efficient stop codon, leading to recoding events such as the programmed frameshift in the *prfB* gene encoding RF2 [161,162]. Pro-opal sequences cause strong +1 frameshifting at CCC_UGA [163] and significant levels of stalling and tagging by tmRNA [100]. As seen in previous studies, the residue upstream of Pro also affects the efficiency of termination [100,164]. In particular, Glu, Asp, and Pro were overrepresented in the –2 position (e.g. Glu-Pro-opal) in our selectants. These results validate our selection and demonstrate that survival in the KanR assay requires high levels of ribosome stalling.

A second set of sequences with the consensus GI(R/H)xPP show weaker activity (surviving only at low stringency). It is interesting to note that the GI residues were not part of the random hexamer library; by chance, these were the two amino acids immediately upstream. These peptides appear to be subtle variants of the SecM sequence $GIRAGP_{166}$. The GIRAPP clone that matches SecM the most closely survives even at high stringency. This suggests that some alterations in this critical SecM sequence are tolerated. Mutation of Arg163 to His or replacing Ala164 with Asp, Ser, or Pro yields substantial though weaker stalling activity. These results agree with the recent findings of Yap and Bernstein, who showed that the GIRAGP sequence in SecM exhibits significant plasticity, with only Arg and Pro residues playing key roles [108].

Our third class of selectants (containing FxxYxIWPP) stall with peptidyl-tRNA in the P site. In the case of the WPPPSI clone, for example, the mass spectrometry data show that the tmRNA tag is added after WPP. Yet the Ala scanning data show that the next residue (the third Pro) is required for tagging, even though it does not react with the nascent peptide. We propose that the aminoacyl-tRNA binds and remains unreacted in the A site, and that peptidyl-transferase

activity is inhibited by FxxYxIWPP-containing peptides. This implicates changes in the conformation of the PTC in the stalling mechanism.

The amino acid Pro plays two different roles in FxxYxIWPPP stalling. First, Pro-tRNA acts as a poor peptidyl acceptor in the A site. It fails to react with the nascent peptide. *N*-alkylamino acids such as Pro have been shown to act as slow nucleophiles in the peptidyl-transferase reaction [165]. Using full-length tRNAs, Pavlov *et al.* demonstrated that the unnatural Pro-tRNA^{Phe} reacts 23-fold slower than Phe-tRNA^{Phe} with initiation complexes containing fMet-tRNA. They speculate that this is due to steric constraints and lower nucleophilicity. Interestingly, the rate of Pro reactivity is accelerated by the natural tRNA^{Pro} isoacceptor; Pro-tRNA^{Pro} only has a three to six-fold defect. A-site bound Pro-tRNA plays a role in stalling on SecM and the 2A peptides found in viral genomes that stall at the Gly residue in the sequence D(V/I)ExNPGP [166]. One possible explanation for the necessity of Pro-tRNA is that the reduced rate of peptidyl transfer to Pro gives the nascent peptide time to interact with the exit tunnel and PTC, shifting the 23S rRNA to an inactive conformation [167].

Our results suggest that aminoacyl-tRNAs other than Pro-tRNA can induce stalling by binding in the A site. WPPD and WPPW sequences were isolated from our selections, and immunoblot analysis revealed that efficient tagging only occurs if the residue following WPP is Pro, Asp, or Trp. While we cannot say for certain, it is probably binding of the amino acid that is critical, not the codon or tRNA. The amino acid is the key component of Pro-tRNA in SecM; the Pro analog azetidine dramatically reduces stalling [106]. Likewise, the binding of free tryptophan causes stalling on TnaC [99]. A second explanation for the role of the A site aminoacyl-tRNA in stalling is that amino acid-binding near the PTC changes the ribosome or peptide conformation.

This possibility is supported by the finding that the SecM peptide undergoes a conformational change in the tunnel upon Pro-tRNA binding in the A site [110].

If aminoacyl-tRNA is bound in the A site, how can tmRNA enter the stalled ribosomes to release them and tag the nascent peptide? SecM, TnaC, and ErmCL (another stalling leader peptide) are not tagged by tmRNA because the A site of the ribosome is occupied [94,96,116,154]. Stalling on these peptides must be determined by cellular conditions to regulate gene expression; tmRNA would interfere with their biological function. But our selection and immunoblot assays rely on tagging by tmRNA to detect stalling events. We propose that overexpression of FxxYxIWPPP leads to depletion of Pro-tRNA by stalled ribosomes, creating a subset of ribosomes stalled with empty A sites that are acted on by tmRNA. Overexpression of SecM results in high levels of tagging [156,168] for exactly this reason [154]. For both SecM and FxxYxIWPPP, increasing $tRNA^{Pro}$ levels abolishes tagging (Figure 2-7). At the same time, we see that tRNA overexpression actually lowers GST levels, perhaps because stalling is more robust with the tRNA in the A site and no GST-stalled ribosomes are released by tmRNA.

The second role of Pro in FxxYxIWPPP stalling is that of a poor peptidyl donor. Peptides ending in Pro react with puromycin far slower than peptides ending in other amino acids [117,118]. It is the amino acid that inhibits peptidyl transfer, not the codon or tRNA—incorporation of Pro analogs azetidine or thiaproline restore rapid reactivity [117]. In uncatalyzed reactions, however, Pro-tRNA is as reactive as other aminoacyl-tRNAs [169], suggesting that the reduced rate is not purely due to the chemistry of prolyl-esters but the interaction of the Pro residue with the ribosome. These findings suggest that the conformationally strained Pro side chain inhibits ribosome activity [118]. In nature, C-terminal Pro residues inhibit termination in proteins in TnaC and in the UL4 gene of the mammalian virus CMV [170]. It appears that the cyclic Pro residue

interferes with conformational changes in the PTC that are required for both elongation and termination.

The WPP-containing peptides stall ribosomes robustly: in the *lacZ* assay, the WPPPSI 18-mer sequence reduced activity nearly 100-fold over the WPAPSI mutant. The sequence context has a great effect—upstream peptide sequences are required for high-efficiency tagging. Tagging in the WPPPSI clone requires the consensus peptide sequence FxxYxIWPP. Phe7 and Tyr10 are aromatic residues that may bind rRNA in the exit tunnel. At nine residues, the FxxYxIWPP is the same length as the ErmCL peptide (MGIFSIFVI) when it stalls upon binding of erythromycin in the nascent peptide exit tunnel [96]. Shortening the ErmCL peptide by deleting N-terminal residues reduces stalling significantly; this length may allow interaction with elements farther into the tunnel, such as the L22 loop [96]. The sequences of these four stalling peptides are different, the only commonality being an Ile four residues from the P site in FxxYxIWPP, ErmCL, and SecM.

Analysis of stalling levels with mutant ribosomes reveals nucleotides that are required for efficient stalling on FxxYxIWPP peptides. A2058 is near the L4 / L22 constriction; the A2058G mutation reduces stalling on SecM by nearly 80-fold, but it actually increases stalling on the WPPPSI clone by eight-fold. Likewise, mutation of U2609 has different effects on these three peptides. In TnaC, the U2609C mutant completely abolished stalling while U2609A only affected it partially [94]. SecM stalling is more reduced by the A mutant, while stalling on FxxYxIWPPP is reduced by either the C or A mutant equally. These data show that stalling on FxxYxIWPPP involves the same players as these other peptides, though the specifics of each interaction vary, suggesting that the peptides bind differently in the tunnel.

The plasticity of nascent peptide interactions with the PTC and exit tunnel is further highlighted by our finding that stalling on the WPPDV* clone occurs both after WPP and during termination. Presumably stalling requires upstream amino acids so the active sequence is FxxYxIWPPDx*. If specific interactions with the tunnel and PTC are lined up properly with the second Pro codon in the P site, it is difficult to imagine how they are aligned again to inhibit termination after the peptide has moved two amino acids farther into the tunnel. We believe that the simplest explanation for this is that the FxxYxIWPP sequence engenders a constrained peptide conformation that promotes interaction with the tunnel at several possible sites. This speculation is supported by the finding of Yap and Bernstein that SecM mutants containing Pro-Pro dipeptides (e.g. PPIRAGP) induce stalling even in the absence of the upstream arrest motif elements [108].

Like SecM and TnaC, FxxYxIWPP-containing sequences stall due to nascent peptide interactions in the PTC and exit tunnel, with an effector bound in the A site. In spite of this common mechanism, these peptides rely on different ligands in the A site and different interactions with the exit tunnel to inhibit peptidyl transfer or hydrolysis. The lack of sequence similarity in these peptides argues that many solutions exist and that regulation of gene expression by nascent peptides may be more common than the few examples characterized so far. Further characterization of the mechanism of stalling on WPP-containing sequences and its biological significance is ongoing.

Experimental Procedures

Library creation

The initial 18 nt library was created by amplifying the truncated *kanR* gene by PCR with the forward primer CATATGGCTAGCATGAGCCATATTCAACGGGAAAC and the degenerate reverse primer CGAAAGGGTACCN$_{18}$ATTACCATATTTTTGAAAAAGCCGTTTCTG. 18 random nt (six codons) were added to the 3'-end of *kanR* following residue 253. To create a second library lacking stop codons in codons four through six, the random region (N$_{18}$) in the degenerate primer was replaced by (NNB)$_3$N$_9$ where B is a mix of C, G, and T phosphoramidites. The PCR products were cloned into pBAD-KT2 [155] with NheI and BamHI and the resulting plasmids were amplified in DH10B. The libraries were then selected in X90 ssrA::cat as described [155] in media containing 15 μg/mL kanamycin at either 25 or 37 °C. The KanR fusion sequences from surviving colonies were amplified, sequenced, and recloned to verify their activities.

Mass spectrometry

The pGEX-3X vector was amplified with inverse PCR to create new restriction sites using the primers AGAGTAGCTAGCACGACCTTCGATCAGATCCG and AGAGTAGCATGCTTGACTGACGATCTGCCTCG. The library cassette was amplified by PCR from 12 residues upstream of the six random codons to the stop codon downstream, including the nucleotide sequence encoding SLQKRLFQKYGIxxxxxxGYRGSRVDRQAWLFWRMREDFQPDTD*, using the primers AGAGTAGCTAGCTCATTACAGAAACGGCTTTTTC and AGAGTAGCATGCTTTAATCTGTATCAGGCTGAAAATC. The resulting PCR products were

digested with NheI and SphI and ligated to create the pGEX-WPPPSI, WPPDV*, and WPPWYR vectors.

These plasmids and pCH201 (expressing tmRNA-H encoding a His_6 tag [101]) were introduced into X-90 ssrA::cat. Cultures were grown to mid-log phase and induced with 1 mM IPTG for 2.5 h. The cells were then harvested and lysed in B-PER reagent (Thermo-scientific) and the cell lysate was cleared in an SS-34 rotor at 15,000 RPM for 20 minutes. His-tagged GST was purified on a Ni-NTA agarose resin from the supernatant. 50 μg of protein was acetone precipitated and digested with trypsin for 14 h at 37 °C. Tryptic fragments were purified again in a Ni-NTA slurry and the peptides were loaded on a reverse-phase ZipTip column, spotted on a MALDI-plate and overlayed with an alpha-CHC matrix (Agilent Technology). Samples were analyzed with a QSTAR Pulsar QqTOF mass spectrometer.

Immunoblot assays

Ala mutations were introduced by PCR into the pGEX-WPPPSI vector described above. X-90 ssrA::cat cells were transformed with a GST vector together with pCH201 (expressing tmRNA-H). Overnight cultures were diluted to OD_{600} of 0.1, grown to OD_{600} = 0.5, induced for 2 h with 1 mM IPTG, and pelleted. The pellets were resuspended, lysed in SDS-lysis buffer, and quantitated. Equal amounts of protein were separated by 12% SDS-PAGE, transferred to a PVDF membrane, and analyzed with mouse anti-His_6 and rabbit anti-GST antibodies (Cell Signaling Technology) as detected by fluorescently labeled anti-mouse and goat anti-rabbit secondary antibodies (LICOR biosciences). Images were taken on a Licor Odyssey IR scanner.

Cell-free translation

Templates were prepared by PCR with the following primers: CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGCA GGGCTGGCAAGCCAC and GGTTATAATGAATTTTGCTTATTAACCAGCCAAGCTTGCCGGTC, adding the T7 promoter and a binding site for the NV1 primer. The PURExpress cell-free transcription-translation system (New England Biolabs) was used for in vitro protein synthesis. Briefly, 0.2 pmol template was combined on ice with 2.5 µl Solution A and 1 µl Solution B along with either 0.5 µl DMSO (5%) or thiostrepton (0.5 mM in 5% DMSO) and then incubated at 37 ºC for 30 min. 1 pmol of [^{32}P]-ATP labeled NV1 primer (GGTTATAATGAATTTTGCTTATTAAC) was added and reverse transcription performed as described [96]. Samples were then extracted with phenol and chloroform, precipitated with ethanol, separated by 6% denaturing PAGE with C and G sequencing lanes, and visualized with a phosphoimager. To detect peptidyl-tRNA in stalled complexes, [^{35}S]-Met was added to the translation reactions. Samples were analyzed by Tricine-SDS PAGE and visualized with a phosphorimager.

Miller Assays

The reporter plasmid was created by inserting the WPPPSI stalling sequence plus 12 upstream residues after the ninth codon of full-length *lacZ* (derived from pNH122 [95]). The SecM sequence FSTPVWISQAQGIRAGP was used as a control. Cells bearing a *lacZ* plasmid and a ribosomal mutant plasmid were induced with 1 mM IPTG at mid-log phase and analyzed for β-galactosidase activity using ONPG as described [171].

Chapter 3. Nascent peptides that block protein synthesis in bacteria.

Author's Note: This chapter details a study in the discovery and characterization of novel stalling peptides and was published in the Proceedings of the National Academy of Sciences in 2013 [172].

Abstract

Although the ribosome is a very general catalyst, it cannot synthesize all protein sequences equally well. Ribosomes stall on the SecM leader peptide, for example, in order to regulate expression of a downstream gene. Using a genetic selection in *E. coli*, we identified novel nascent peptide motifs that stall ribosomes. Kinetic studies show that some nascent peptides dramatically inhibit rates of peptide release by release factors. We find that residues upstream of the minimal stalling motif can either enhance or suppress this effect. In other stalling motifs, peptidyl transfer to certain aminoacyl-tRNAs is inhibited. In particular, three consecutive Pro codons pose a challenge for elongating ribosomes. Translation factor EF-P, which alleviates pausing at polyproline sequences, has little or no effect on other stalling peptides. The motifs that we identified are underrepresented in bacterial proteomes and show evidence of stalling on endogenous *E. coli* proteins.

Introduction

We commonly think of the ribosome as capable of synthesizing any protein, regardless of its sequence. But it turns out that some nascent peptides contain stalling motifs that inhibit core functions of the ribosome [50,126]. Why would a protein evolve to arrest its own synthesis? In one of the best characterized examples, stalling in the SecM leader peptide up-regulates translation of

SecA, a protein encoded downstream on the same mRNA [95]. Known stalling motifs have typically been identified based on their function as genetic switches, regulating gene expression in response to levels of protein translocation factors [88,95] or changes in the concentration of small molecule metabolites [99,126]. Further understanding of the scope and mechanism of ribosome stalling may yield additional insight into programmed ribosome stalling events that regulate gene expression in organisms from bacteria to humans [144,173]. In addition, by fine-tuning the rate of protein synthesis, stalling peptides may affect protein folding and function, as reported previously with rare codons [174,175].

Analyses of natural motifs have identified three sites of interaction within the ribosome that lead to stalling. First, conserved residues at a motif's N-terminus often interact with the ribosome near a constriction in the exit tunnel between proteins L4 and L22. Ribosomal mutations near this site were isolated in genetic screens for reduced levels of stalling [94,95]. Second, conserved residues near a motif's C-terminus interact with nucleotides surrounding the ribosomal active site, the peptidyl-transferase center or PTC [96]. Third, some motifs encode a specific aminoacyl-tRNA that acts as a poor peptidyl acceptor when bound in the A site [97,98]. The SecM consensus motif (FxxxxWIxxxxGIRAGP) showcases all three of these interactions, each of which contributes to stalling. A Trp side chain eleven residues upstream of the stall site binds near the constriction in the tunnel, the Arg residue three residues upstream is positioned close to the PTC [95], and Pro-tRNA binds in the A site, but does not react [116]. A recent cryo-EM structure of the stalled SecM complex directly visualized these interactions and begins to provide some molecular rationale for how stalling is induced [109].

A full understanding of the molecular mechanism underlying ribosome stalling has not been forthcoming because of the complexity of natural stalling motifs. In each case, stalling is

reversible and is controlled by changes in the cellular environment. Stalling in TnaC, for example, is induced by high tryptophan concentrations through the binding of free tryptophan at an unknown site within the ribosome [99]. Given that a single stalling motif interacts with the ribosome at multiple sites, deconvoluting the role of the conserved residues in the peptide is difficult enough without the added complexity of small molecule binding. Moreover, even at well-validated sites of interaction, such as the L4/L22 constriction, different stalling motifs appear to work via different mechanisms. Ribosomal mutations that reduce stalling by one motif may have no effect or even increase stalling by another [98,114]. These complexities make it challenging to obtain general conclusions about the mechanism of ribosome stalling by natural stalling peptides.

To better characterize the scope and mechanism of ribosome stalling, we set out to find a series of artificial motifs that inhibit translation during their own synthesis. We reasoned that by selecting directly for stalling peptides, we might find new motifs that are simpler than natural ones because they are not required to stall reversibly or to regulate downstream genes. The fact that only a few residues are essential for stalling by SecM, TnaC, and ErmCL [50], and the fact that these motifs share little or no sequence similarity, led us to believe that more stalling motifs exist but have not yet been identified. To this end, we developed a powerful genetic selection in *E. coli* that ties stalling on a reporter protein to cellular survival.

Here we report several new stalling motifs, some that block peptide release during translational termination and others that block peptidyl transfer. Because these motifs are short, we are able for the first time to recapitulate the stalling phenomenon using pre-steady state kinetic assays, an important step towards achieving mechanistic insights. Of particular interest, we show that polyproline sequences induce ribosome stalling. The translation factor EF-P, which

alleviates stalling at polyproline stretches, does not affect the other short motifs we identified, further defining its scope of action. Finally, our analysis of bacterial proteomes reveals that stalling motifs are underrepresented, implying that they have been selected against. Where they do occur in endogenous *E. coli* proteins, pausing is detectable by ribosome profiling. These findings argue that these short motifs have an impact on protein synthesis in bacteria.

Results

A genetic selection for stalling motifs

We previously reported the first systematic search for nascent polypeptide motifs that induce ribosome stalling [98]. In that study, we identified stalling motifs from random libraries using a genetic selection based on tmRNA, part of the machinery that rescues stalled ribosomes in bacteria [68,69]. tmRNA recognizes stalled ribosomes and directs the addition of a short peptide tag to the nascent polypeptide; this allowed us to detect stalling events in living cells. We linked tmRNA tagging of the KanR protein to cellular survival and identified a novel stalling motif, FxxYxIWPPP [98]. Because the selection depended on functional KanR, we suspect that many motifs were missed. To survive the selection, a motif within the KanR sequence had to induce ribosome stalling but not interfere with the enzyme's folding and activity.

To overcome these limitations, we developed a new selection that allows more variability in the motif length and sequence than was possible in our earlier study. We established a variant of the bacterial two-hybrid system [176,177] that links stalling with cellular survival (Figure 3-1A). In this system, cells cannot synthesize histidine unless transcription of a *HIS3* reporter gene is activated. Driven from a weak promoter, *HIS3* expression is insufficient for cellular survival unless RNA polymerase is recruited to the transcriptional start site by a DNA-binding protein, in

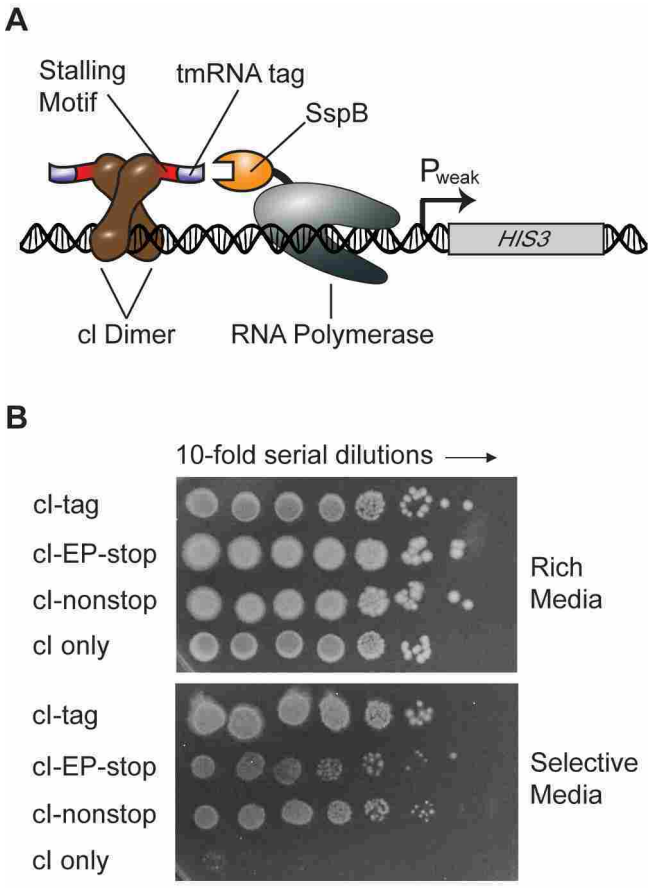


Figure 3-1. Two-hybrid selection for nascent peptides that stall ribosomes**.** A) The lambda cI protein binds to DNA upstream of the *HIS3* gene. When ribosomes stall during cI synthesis, a short tag is added to the protein by tmRNA, recruiting an SspB-RNA polymerase fusion protein. In this way, ribosome stalling in cI activates transcription of *HIS3*, restoring the cell's ability to synthesize histidine and survive on minimal media. B) Four cI constructs were tested for their ability to recruit SspB-RpoA, activate *HIS3*, and grow on media lacking histidine: cI translationally fused to the tmRNA tag; cI fused to two known inducers of stalling, Glu-Pro-stop and an mRNA lacking a stop codon; and cI with no stalling motif.

this case a modified form of lambda cI. We encoded stalling motifs at the C-terminus of the full-length cI protein, where they have little or no effect on protein structure or function. When ribosomes stall during cI synthesis, the protein is tagged by tmRNA and can therefore recruit RNA polymerase fused to SspB. The SspB protein binds specifically to the peptide tag encoded by tmRNA [178]; the resulting interaction between tagged cI and RpoA-SspB leads to recruitment of RNA polymerase to the *HIS3* gene, transcriptional activation, and survival of cells on selective media lacking histidine. Note that the natural function of SspB is to deliver tagged proteins to the ClpXP protease for destruction [178]. By including only the tag-binding domain of SspB (residues 1-117), not the ClpXP binding domain [179], and by changing the last two residues of the tmRNA tag to DD [159], we prevent degradation of tagged cI.

We validated the selection with four control cI constructs. First, expression of cI alone, with no stalling motif, does not support growth on minimal media lacking histidine, because cI is not tagged by tmRNA and thus does not bind RpoA-SspB. When the modified tmRNA tag is translationally fused to the C-terminus of cI in a second construct, the resulting protein recruits RpoA-SspB and activates *HIS3* transcription, leading to robust survival on selective media (cI-tag, Figure 3-1B). Furthermore, we showed that known stalling motifs survive the selection by generating sufficient levels of tmRNA-tagged cI. We added Glu-Pro-stop to the C-terminus of cI; this short sequence was previously shown to induce high levels of stalling and tagging *in vivo* [100]. In another construct, we expressed cI on a non-stop mRNA, where ribosomes translate to the 3'-end of the message because there is no stop codon. These ribosomes are effectively stalled and are known to be rescued by tmRNA [81]. Both of these constructs support cellular survival on selective media that is as robust as the cI-tag translational fusion (Figure 3-1B). These data demonstrate that the selection successfully ties ribosome stalling during cI synthesis with cellular survival on selective media.

To identify novel stalling motifs, we fused twenty random codons onto the C-terminus of cI and subjected the resulting library to the two-hybrid selection. Although we could only sample a tiny sliver of sequence space (~10^8 clones out of ~10^{26} possibilities), we found that a significant fraction of the library survived the selection, roughly 1 in 10^4 colonies plated. This high rate of survival suggests that it is remarkably easy to find sequences that induce stalling and tagging by tmRNA. Because two-hybrid systems are notoriously rich in false positives, we performed a secondary screen in which lysates from 150 colonies were immunoblotted with antibodies against the tmRNA-DD tag [180]. Over a quarter (41/150) showed significant levels of tmRNA-tagged cI protein, suggesting that they survived the selection because ribosome stalling

occurred during cI synthesis. Known causes of stalling could be attributed to 21 of these 41 clones. The most common motif was Pro-stop at the C-terminus of the cI protein (18 clones). Pro is known to induce stalling during termination [100]. In contrast, clusters of rare codons, another known cause of stalling, were found in only three clones. For the remaining 20, no known cause of stalling was apparent, though tagged cI was clearly being produced. Presumably something in the variable region led to tagging, but the site of stalling and the residues responsible for stalling had to be identified for each clone by additional experiments.

Determination of the site of stalling and the consensus stalling motifs

We used two different approaches to determine the site of stalling in clones that survived the two-hybrid selection. First, we selected ten clones that showed the highest levels of tagging in the immunoblot assay above and performed toeprinting assays to directly detect stalling in these clones during *in vitro* translation. A radiolabeled primer was annealed to the 3'-end of the cI transcript and extended by reverse transcriptase. When it encounters a stalled ribosome, reverse transcriptase stops 15–16 nt downstream of the first nucleotide in the P site codon. This reveals which codon is in the P site when stalling occurs. As a control, the antibiotic thiostrepton was added to trap ribosomes in the initiation stage. The disappearance of the toeprint band when thiostrepton is added demonstrates that the block in reverse transcription is due to stalled ribosomes and not an artifact of mRNA sequence or structure.

The toeprinting results show that the ten clones fall into two classes. Some stall with a stop codon in the A site, indicating that termination is inhibited; toeprinting data for three examples of this class, T1-T3, are shown in Figure 3-2A. Others stall with a sense codon in the A site, indicating that elongation is inhibited; toeprinting data are also shown for two such motifs,

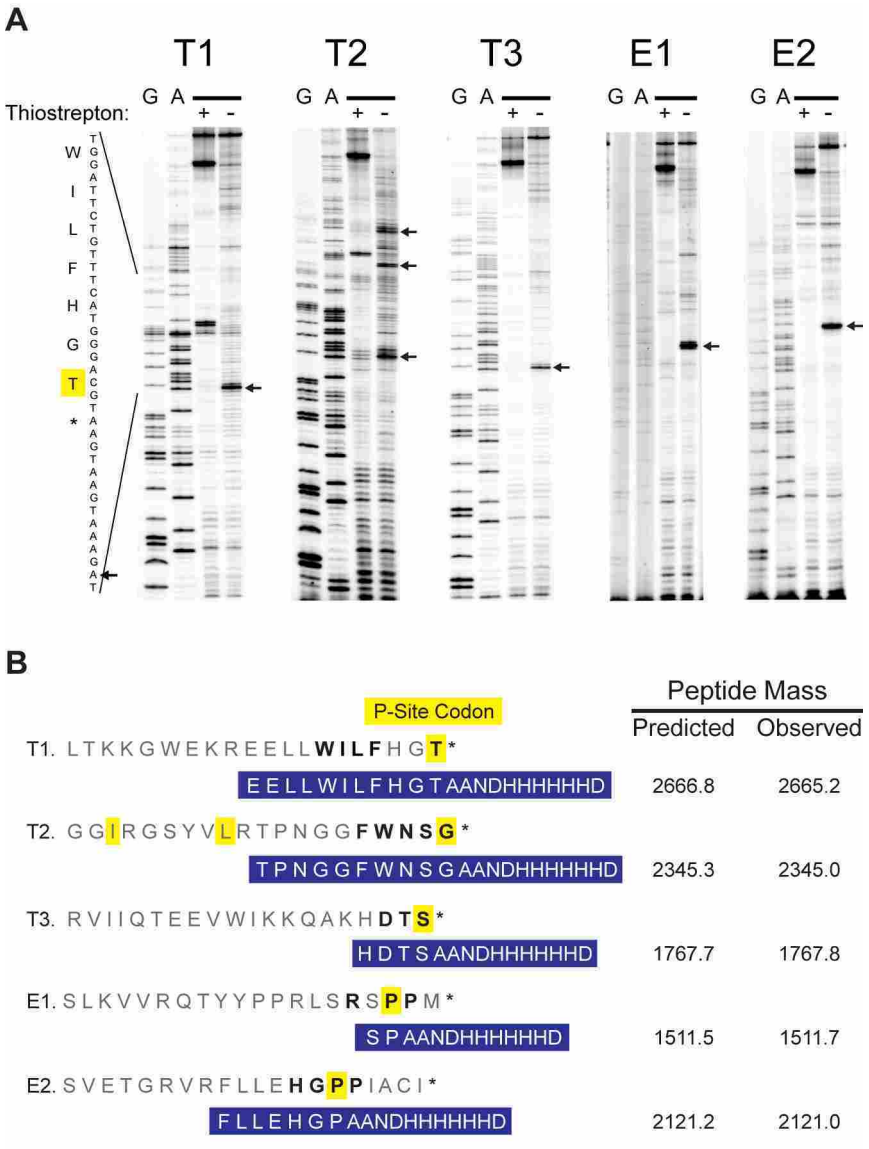


Figure 3-2. Stalling motifs that block termination (T1-T3) or elongation (E1, E2). A) Stalling sites for five motifs were determined by toeprinting assays. Blocks in cDNA synthesis are marked with arrows. The antibiotic thiostrepton traps ribosomes in initiation complexes; bands seen in both treated and untreated lanes are reverse transcriptase artifacts. B) Key amino acids in each motif are highlighted in bold. The codon labeled in yellow is positioned in the P site of the stalled ribosome. The site of stalling *in vivo* was confirmed by purifying tagged cI, digesting the protein with trypsin, and determining the mass of the C-terminal peptide (highlighted in blue) by MALDI-MS.

E1 and E2. All ten sequences and their stalling sites are given in Table 3-1. The fact that stalling was recapitulated in a reconstituted translation system (the PURE system) rules out alternate explanations such as mRNA cleavage or degradation that might have led to tagging by tmRNA and survival in the two-hybrid selection.

Motif																			
	Elongation																		
E1	S	L	K	V	V	R	Q	T	Y	Y	P	P	R	L	S	R	S	P	P	M	*
E2	S	V	E	T	G	R	V	R	F	L	L	E	H	G	P	P	I	A	C	I	*
E3	E	Q	V	I	N	L	G	P	D	E	E	W	G	A	T	R	K	C	V	H	*
E4	C	E	I	K	G	Y	L	L	P	L	K	I	A	P	Y	S	S	L	A	K	*
E5	K	F	G	G	T	I	S	C	M	Q	S	L	R	D	I	L	E	L	A	A	*
E6	F	Y	G	L	L	S	D	G	G	G	K	K	R	V	N	I	P	W	S	L	*
	Termination																		
T1	L	T	K	K	G	W	E	K	R	E	E	L	L	W	I	L	F	H	G	T	*
T2	G	G	I	R	G	S	Y	V	L	R	T	P	N	G	G	F	W	N	S	G	*
T3	R	V	I	I	Q	T	E	E	V	W	I	K	K	Q	A	K	H	D	T	S	*
T4	R	P	H	Q	R	F	V	I	P	H	V	G	F	D	*						

Table 3-1. Sequence of ten stalling motifs**.** The underlined codon (red) is found in the P site in toeprinting analyses. The full sequence is shown but it is likely that not all of these residues are essential for stalling.

We performed a second series of experiments to detect where stalling and tmRNA tagging occurred *in vivo*. This was done by analyzing tagged cI protein by mass spectrometry; the residue upstream of the tag corresponds to the codon positioned in the P site in the stalling event. We analyzed the five clones depicted in Figure 3-2A because they gave strong toeprints. We expressed each cI clone with a modified tmRNA that encodes six His residues in its tag sequence [181]. Tagged cI was purified over Ni-NTA resin and digested with trypsin. The peptide fragment corresponding to the C-terminus of the tagged cI protein contains the altered tmRNA tag and a few upstream residues, depending on where trypsin cleavage occurs. This peptide was enriched on an Ni-NTA resin and its mass was determined by MALDI-MS. The amino acid sequence of the peptide was also confirmed by tandem MS/MS. The five peptides and their masses are shown in Figure 3-2B. In each case, the residue immediately before the tag

corresponds to the codon in the P site in the stalled complex in the toeprinting assays, confirming that stalling occurs at the same site *in vitro* and *in vivo*.

We determined which residues in these five stalling motifs are required for stalling through a process of mutagenesis and reselection. We introduced mutations into the 20 codons of a given motif at a frequency of 30% per nucleotide. The resulting library was subjected to the two-hybrid selection and a consensus sequence of surviving clones was determined. Taken together, the toeprinting, mass spectrometry, and reselection experiments define five new stalling motifs, their stalling sites, and consensus sequences. Data from the motifs that inhibit termination (T1-T3) will be discussed first, followed by the data for the elongation motifs (E1 and E2).

Stalling motifs that inhibit termination

The T1 motif contains the consensus residues WILFxxT-stop, where x is any amino acid. When the T1 sequence was subjected to mutagenesis and reselection, the highest enrichment was seen at the C-terminal residue, Thr, with $P < 10^{-5}$ (Figure 3-3A). We found that mutation of this Thr to Ala abolishes stalling in the toeprinting assay (Figure 3-3B). The Trp residue seven codons back from the stop codon is also selected for ($P < 10^{-4}$); it lies in a stretch of hydrophobic residues. Ala substitutions have little or no effect on stalling by this motif, with the exception of the Phe residue at the –4 position. In contrast, Glu substitutions often reduce stalling efficiency, presumably by blocking hydrophobic interactions between these residues and the ribosomal exit tunnel. It appears that stalling in the T1 motif is induced by the C-terminal Thr residue and this stretch of upstream hydrophobic and aromatic residues.

The T2 motif is FWNSG-stop. Although two additional sites of stalling are detectable in the toeprinting assay (Figure 3-2A), only trace amounts of the corresponding peptides were

detected in the MS analysis. This suggests that stalling during termination was the predominant cause of tmRNA tagging *in vivo*. The mutagenesis and reselection experiments show conservation of the final four codons WNSG (each with $P < 0.01$), though some flexibility is evident. Asn and Asp occur at the –3 position and Ser and Thr occur at the –2 position (Figure 3-3A). We found that the Asn to Asp mutation and the Ser to Thr mutation both retain significant stalling activity, whereas replacing any residue in the FWNSG motif with Ala abolishes stalling in the toeprinting assay (Figure 3-3B). These mutations can also occur in combination: changing the final three amino acids from NSG to DTG had little or no effect on stalling efficiency.

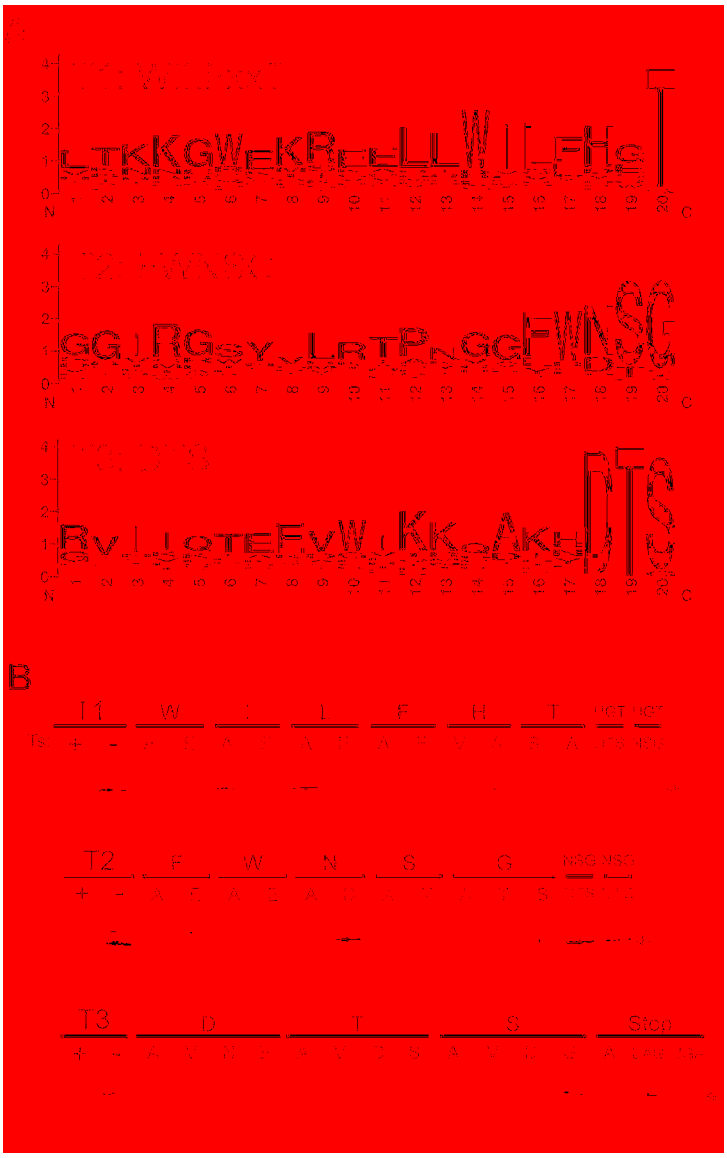


Figure 3-3. Key residues in motifs that block termination. A) The consensus sequence of each motif was determined by randomizing all twenty codons at 30% per nucleotide and subjecting the resulting library to the two-hybrid selection. B) Toeprinting analyses of a series of mutants for each motif. The original residues are shown above the line and the substitutions are given below. A thiostrepton-treated control reaction (TS) is shown in the left lane.

The T3 motif, DTS-stop, is sharply defined by the consensus sequence from the mutagenesis and reselection data. Each codon is conserved with a P value < 0.001 (Figure 3-3A). While the Asp and Thr codons are invariant, in a few surviving clones the final Ser codon is replaced with Gly. Toeprinting analyses confirm that the Ser to Gly mutation has little or no effect, whereas mutating any residue in the DTS sequence to Ala abolishes stalling (Figure 3-3B). As the T2 motif was able to accommodate either Asp or Asn at the first codon and Thr or Ser at the second, we reasoned that T3 might also exhibit some sequence plasticity. We found, however, that both the Asp to Asn and the Thr to Ser mutations strongly inhibit stalling. We conclude that DTS-stop or DTG-stop is sufficient for stalling ribosomes but that NSG-stop stalls poorly without the upstream FW residues found in T2. In support of this conclusion, we found that replacing the last three residues of T1 with DTS yields robust stalling, whereas replacing them with NSG does not (Figure 3-3B). Finally, we note that DTS only stalls during the termination step, not during elongation. Although stalling occurs at both UAG (recognized by RF1) and UGA (recognized by RF2) stop codons, mutating the stop codon to an Ala sense codon abolishes stalling (Figure 3-3B). Taken together, our analysis of the T1-T3 motifs show that short, polar peptides can induce ribosome stalling at stop codons.

Stalling motifs that inhibit elongation

Characterization of the two elongation motifs with the strongest toeprints, E1 and E2, revealed that the amino acid Pro plays a dual role in both motifs. The toeprinting data indicate that stalling occurs with the first Pro codon in the P site and the second Pro codon in the A site (Figure 3-2A). Analysis of tmRNA-tagged cI protein by mass spectrometry confirmed that ribosomes stall in this same position *in vivo* (Figure 3-2B). Randomization of the E1 motif and

reselection define the key residues as RxPP, where x can be Ser, Ala, Gly, and Pro (Figure 3-4A). Similarly, the consensus E2 motif is HGPP. In both cases, the two Pro codons are invariant, and mutation of either to Ala completely abolishes stalling in toeprinting assays (Figure 3-4B). An analysis of the nucleotide sequence of surviving clones in the E2 library shows that there is strong selection for the first two nucleotides of each Pro codon, but not for the third (Figure 3-4A). This observation suggests that the amino acid is critical, not the nucleotide sequence, or even a particular tRNA.

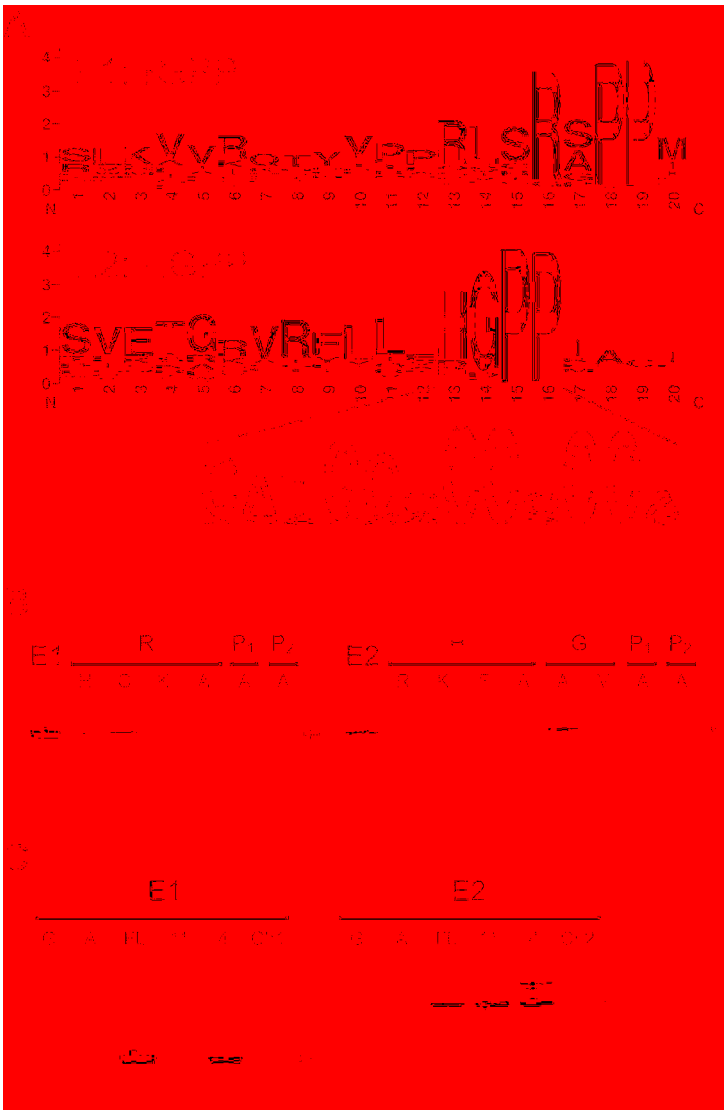


Figure 3-4. Key residues in motifs that block elongation**.** A) The consensus sequence of the E1 and E2 motifs was determined through random mutagenesis and reselection. In addition, the nucleotide consensus for the HGPP codons is shown below E2. B) Toeprinting assays reveal the effects of substitutions in the E1 motif (RxPP), left, and the E2 motif (HGPP), right. C) Toeprinting assays of E1 and E2, including the minimal motif alone (4), the minimal motif plus seven additional upstream codons (11), and the full-length construct (FL). Ch1 is a chimera containing the minimal motif of E1 (RSPP) with seven upstream codons taken from E2. Ch2 corresponds to the minimal motif of E2 (HGPP) with seven upstream codons from E1.

While these two motifs have clear similarities, a subtle context dependence arises from residues upstream of the consensus motif. At first glance, it seems that E1 and E2 are both examples of the same family, R/HxPP, where x is a small amino

acid. The first residue (R or H) is critical for both motifs: stalling is lost when it is mutated to Ala. However, replacing the Arg residue in E1 with His significantly reduced stalling levels, as did replacing the His residue of E2 with Arg (Figure 3-4B). We thought that the requirement for a particular residue (Arg or His) might arise from the peptide context; perhaps R/HxPP is sufficient for stalling, but the upstream sequence imposes a specific requirement for either Arg or His. In support of this idea, we found that the minimal motifs of E1 (RSPP) and E2 (HGPP) induce robust stalling, as strong as the original 20-codon sequence (compare the lanes labeled 4 and FL in Figure 3-4C). This demonstrates that no upstream residues are necessary. We then created chimeras containing elements of both motifs. In Ch1, the seven residues upstream of the E2 motif were put upstream of the minimal E1 motif, RxPP. As shown in Figure 3-4C, the toeprint was dramatically reduced, indicating that these additional seven residues block stalling at the minimal E1 motif. Likewise, stalling was inhibited in the Ch2 chimera, where the seven residues upstream of the E1 motif were added to the N-terminus of the minimal E2 motif, HGPP. These findings show that secondary elements in the extra seven residues somehow restrict which residues work in the minimal motif, making the sequence specific for either Arg or His.

Not only is stalling by the E1 motif sensitive to the upstream peptide sequence, it is also sensitive to the length of the nascent polypeptide. Although ribosomes stall robustly on the minimal E1 sequence RSPP, stalling is dramatically reduced when a longer truncated form of the E1 motif is translated, YYPPRLSRSPP (compare lanes labeled 4 and 11, Figure 3-4C). Note that we have not altered the motif sequence; this truncated form is part of the full motif. The explanation for this paradox may be that the RSPP residues cannot achieve the right conformation in the context of the longer YYPPRLSRSPP sequence. In the context of the full

motif, however, the peptide conformation permits productive stalling; in effect, the stalling suppressor is suppressed.

The nascent peptide dramatically influences rates of peptide release and peptidyl transfer

One limitation in the study of ribosome stalling has been the lack of biochemical tools to probe the mechanism. Because they are short, simple, and robust, the motifs defined above enabled us to recapitulate stalling in pre-steady state kinetic assays [60,182,183]. For example, we observed inhibition of peptide release on the T3 motif, DTS-stop. From purified components, we assembled ribosome complexes with MDTS peptidyl-tRNA bound in the P site and the UAA stop codon poised in the A site. We reacted this complex with saturating concentrations of RF1 and measured the rate of release of the peptide from its tRNA. We found that this rate was quite slow, 0.006 s^{-1}.

Peptides containing single Ala substitutions (MATS, MDAS, MDTA) were released about 10-fold faster than the original MDTS peptide (Figure 3-5A). This corroborates the consensus sequence and toeprinting data, indicating that each of these three residues is important for stalling by the T3 motif. In addition, we found that although Phe substitutions at the first two positions yielded a similar 10-fold decrease in stalling, changing the final residue to Phe decreased stalling by 70-fold. We speculate that the larger Phe side chain prevents Asp and Thr from making necessary contacts due to steric effects or conformational changes in the peptide.

In considering the importance of the size and geometry of the final residue, we wondered about the effect of Pro on termination rates. In this selection and in previous studies [98,100], proteins ending in Glu-Pro, Asp-Pro, and Pro-Pro were found to induce high levels of stalling and tagging by tmRNA *in vivo*. We measured the rate of release of the MEP peptide by RF1 and

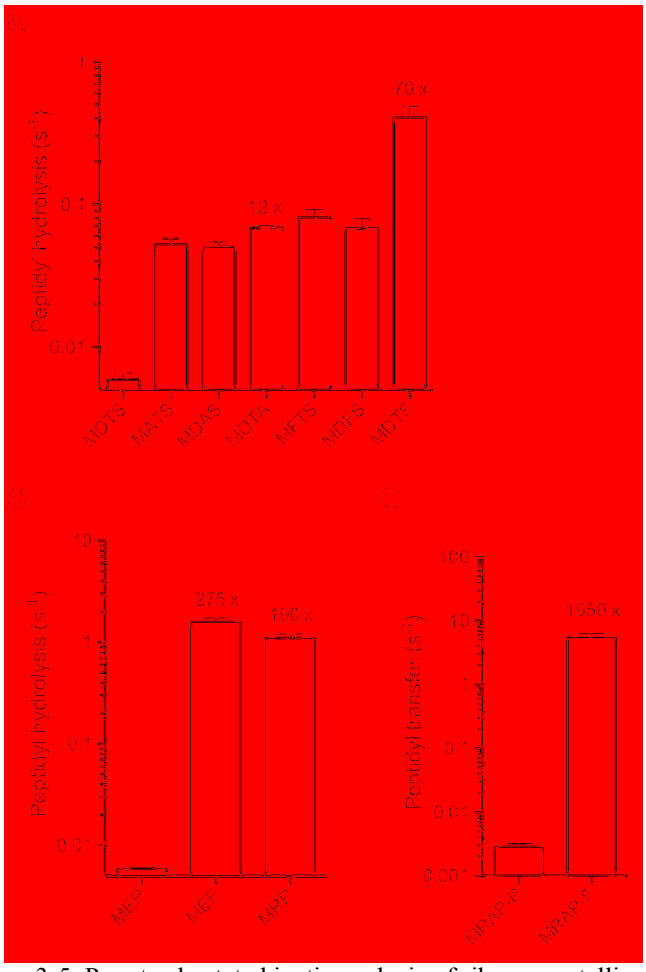


Figure 3-5. Pre-steady state kinetic analysis of ribosome stalling**.** A) A ribosome complex containing MDTS-peptidyl-tRNA in the P site and the UAA stop codon in the A site was reacted with RF1 to determine the rate of peptidyl-tRNA hydrolysis. Ala (grey) and Phe (black) mutants were characterized to test the contribution of each residue in the DTS motif. B) A ribosome complex containing MEP-peptidyl-tRNA or derivatives was reacted with RF1. C) Peptidyl-transfer rates for the E1 motif were obtained by reacting ribosome nascent chain complexes with excess ternary complex composed of Pro- or Phe-tRNA, EF-Tu, and GTP.

found it to be 0.006 s^{-1}, about the same as MDTS. Changing the last residue from Pro to Phe increased the rate of release 275-fold (MEF, Figure 3-5B), highlighting the importance of Pro at the final position. The –2 position is also critical: replacing Glu with Arg led to a 190-fold rate increase (MRP, Figure 3-5B). These data are consistent with earlier findings that when a protein ends in Pro, tagging by tmRNA is lowest when Arg is found in the –2 position [100]. Taken together, these data show that the peptide sequence, and in particular the final three amino acids of the protein, can have a profound effect on the rate of release by RF1. The effect is primarily on catalysis, not binding, as saturating concentrations of RF1 were used (representative rate data are shown in Figure 3-6).

In addition to these studies on peptide release, we also recapitulated stalling during peptidyl transfer in a pre-steady state kinetic assay [183]. To further characterize the E1 motif, RxPP, we assembled ribosome complexes on an mRNA encoding MRAPP. A ribosome complex

containing MRA peptidyl-tRNA was formed, purified, and reacted with a ternary complex of Pro-tRNA, EF-Tu, and GTP. Products were resolved by electrophoretic TLC, and, importantly, the appearance of MRAP and MRAPP could be distinguished. The rate of peptidyl transfer of MRAP to Pro-tRNA was determined to be 0.003 s^{-1}. In contrast, we found that the rate of peptidyl transfer of MRAP to Phe-tRNA was 5.5 s^{-1}, nearly 2,000-fold faster (Figure 3-5C). These data show that the incoming aminoacyl-tRNA in the A site can have dramatic effects on peptidyl-transfer rates. Taken together, these studies of the kinetics of peptide release and peptidyl transfer demonstrate the value of these motifs in providing tools for future mechanistic studies.

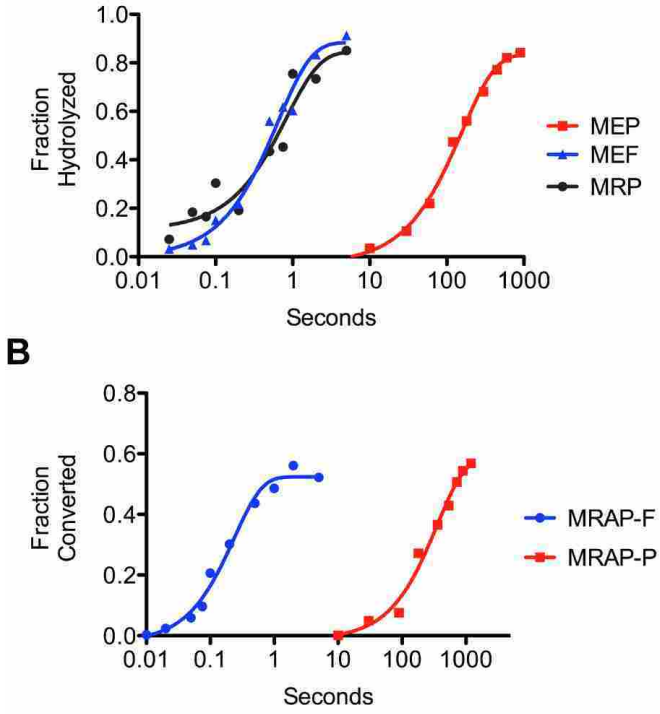


Figure 3-6. Representative rate profiles from peptide release (top) and peptidyl transfer reactions (bottom). A) Ribosome complexes containing a tripeptidyl-tRNA in the P site and a UAA stop codon in the A site were reacted with release factor 1. The fraction of the peptidyl-tRNA that is hydrolyzed is plotted against time. Curves for MEP (red), MEF (blue), and MRP (black) are shown. B) Ribosome complexes containing a peptidyl-tRNA in the P site were reacted with a ternary complex of EF-Tu, GTP, and aminoacyl-tRNA. The fraction of tetrapeptide converted to pentapeptide is plotted against time. The reaction with Phe-tRNA is shown in blue and Pro-tRNA is shown in red.

Stalling occurs at various PPX motifs in vitro

It is striking that peptidyl transfer is blocked in the E1 and E2 motifs by such short peptide sequences. This led us to ask if a motif we had identified previously, FxxYxIWPPP,

might also contain a shorter minimal motif capable of inducing stalling [98]. To test this, we monitored stalling using toeprinting assays with various derivatives of the original FxxYxIWPPP motif. As reported previously, stalling in the full-length sequence occurs with the second Pro codon in the P site and the third Pro codon in the A site. Mutation of the second Pro to Ala abolishes stalling (Figure 3-7A).

Additional analysis of the FxxYxIWPPP motif established that three Pro codons are sufficient to induce a high level of stalling in toeprinting assays. Two strong pauses are detected during translation of WPPP and APPP; these correspond to either the first or second Pro codon positioned in the P site (Figure 3-7A). Pausing at these two sites is interdependent: mutation of any of the three Pro codons to Ala reduces stalling at both sites dramatically. Although a low level of stalling is detectable with two consecutive Pro codons in the WAPP and WPPA mutants, these data argue that three Pro codons are necessary and sufficient for robust stalling.

Our previous *in vivo* studies on the FxxYxIWPPP motif indicated that, in addition to Pro-tRNA, other aminoacyl-tRNAs can also act as poor peptidyl acceptors when bound in the A site, prohibiting peptidyl transfer [98]. When the third Pro codon is mutated to Trp or Asp (e.g. FxxYxIWPPW), robust ribosome stalling and tmRNA tagging occur *in vivo*. We revisited this phenomenon in the context of the shorter, core PPP motif, testing all twenty amino acids as peptidyl acceptors in toeprinting assays. We found that robust stalling by PP(X) occurs when X is an Asn, Asp, Glu, Gly, Pro, or Trp codon (Figure 3-7B). These aminoacyl-tRNAs act as poor peptidyl acceptors for peptides ending in Pro-Pro, further defining the scope of this phenomenon.

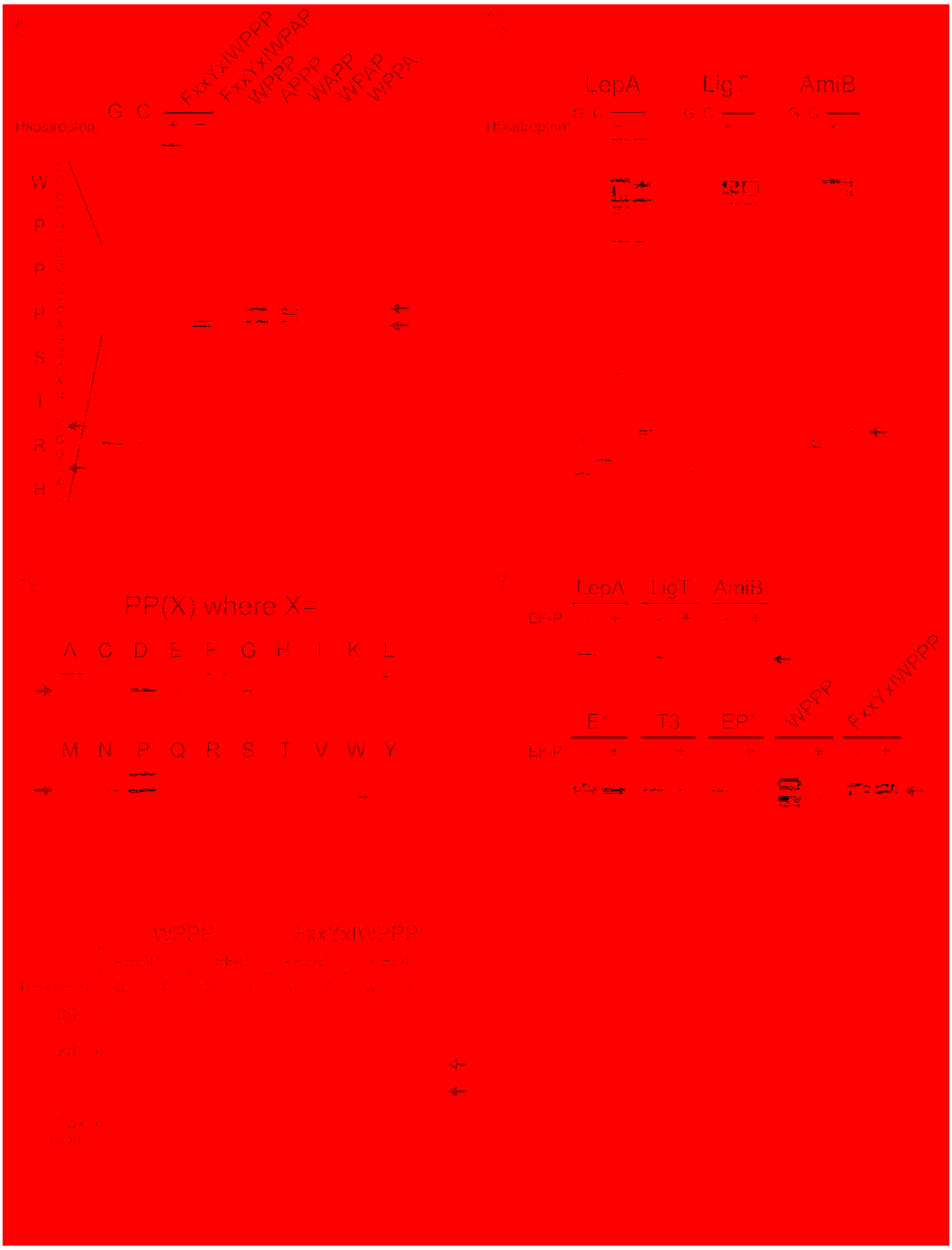


Figure 3-7. EF-P alleviates stalling at polyproline stretches but not other stalling motifs. A) Toeprinting analyses of Ala substitutions in the FxxYxIWPPP motif and a truncated derivative, WPPP. B) Analysis of pausing in the 20 PPX motifs. The arrow points to the toeprint with the second Pro codon in the P site and the X codon in the A site. C) Toeprinting analyses of three endogenous *E. coli* genes with polyproline stretches: *lepA*, *ligT*, and *amiB*. D) Various motifs were translated in the presence or absence of purified EF-P; the relevant toeprint is labeled with an arrow. E) MWPPP and MFQKYGIWPPP were translated *in vitro* with [^{35}S]-methionine in the presence or absence of EF-P. Peptidyl-tRNA accumulates in stalled ribosomes; it was visualized by SDS-PAGE and autoradiography after 3 or 7 minute reactions.

Stalling at polyproline stretches in endogenous E. coli proteins

Due to the simple nature of these motifs, they occur often in endogenous proteins; we next asked how these proteins can be translated *in vivo* given the strong pauses that occur *in vitro*. The *E. coli* MG1655 genome encodes ~100 proteins containing three or more consecutive Pro codons, and even more if Asn, Asp, Glu, Gly, or Trp are allowed at the third position. Perhaps the context of the motif within the proteins prohibits stalling, as we have seen with the E1 and E2 when upstream sequences are swapped. To address this question, we performed *in vitro* translation of LepA and LigT, both of which contain PPP motifs, and AmiB, a protein with eight consecutive Pro codons. As shown in Figure 3-7C, toeprinting analyses reveal that protein synthesis of LepA and LigT does stall dramatically at the PPP motif, with the second Pro codon in the P site. These are the clear pause sites in these genes. In AmiB, where there are eight Pro codons in a row, the stalling occurs primarily with the second Pro codon in the P site, but also to a lesser extent at downstream Pro codons. These data show that endogenous *E. coli* proteins stall at PPP motifs *in vitro*; their sequence context does not effectively suppress stalling.

EF-P alleviates stalling at polyproline stretches but not other stalling motifs

During the course of these studies, it was discovered that the translation factor EF-P relieves translational stalling at polyproline stretches, thus explaining how proteins with polyproline sequences are translated efficiently in living cells [184,185]. Elongation factor P (EF-P) was initially characterized 40 years ago as a ribosome bound translation factor that facilitated the first peptide bond [186,187]. Structural studies show that EF-P is comprised of three domains that mimic the structure of tRNA [188]. This allows EF-P to bind the ribosome between the E and P sites, contacting both the mRNA and PTC [189] (Figure 3-8A). Once bound, EF-P inserts domain I

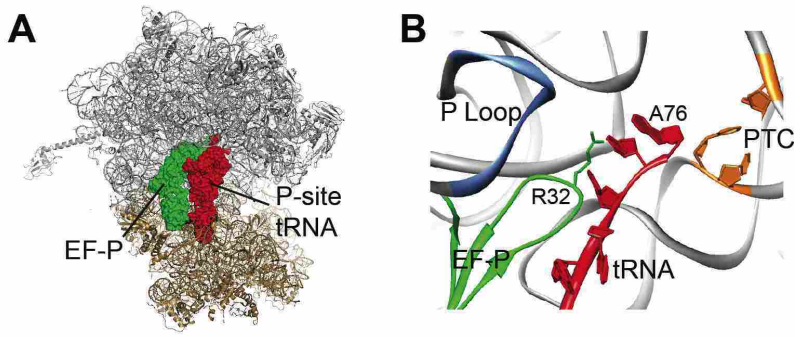


Figure 3-8. EF-P binds the ribosome near the PTC**.** A) Representation of EF-P bound between the E and P sites of the 70S ribosome in *T. thermophilus.* B) Conserved residue Arg32 (Lys34 in *E.coli*) is positioned in the PTC near the peptidyl-tRNA. Both A) and B) were prepared using MacPyMOL and structure coordinates from ref. (189) (PDB codes 3HUW & 3HUX).

(mimic of 3' end of tRNA) into the ribosome's active site, placing conserved lysine residue, Lys34, near the peptidyl-tRNA (Figure 3-8B). Lys34 is modified with β-lysine by the enzymes YjeK and YjeA [190,191]. This creates a long, flexible cofactor that can reach into the ribosome's active site and alter the PTC geometry or the chemistry of peptide bond formation. Studies have shown that this modification is essential for EF-P activity [190].

As expected, addition of EF-P to the translation reactions abolishes stalling at the polyproline stretches present in LepA, LigT, and AmiB as seen in the toeprinting assays (Figure 3-7D). Furthermore, although a peptidyl-tRNA intermediate is the primary product in the absence of EF-P, synthesis of the full-length protein is strongly enhanced by addition of EF-P (Figure 3-9).

We next asked whether EF-P functions universally to alleviate stalling. Toeprinting assays revealed that EF-P has little or no effect on stalling by the E1 (RxPP), T3 (DTS-stop), and Glu-Pro-stop motifs (Figure 3-7D). Likewise, EF-P had no effect on stalling by the full-length FxxYxIWPPP motif, though it abolished stalling by the minimal WPPP motif (Figure 3-7D). Analysis of the products of *in vitro* translation further confirmed the ability of EF-P to prevent stalling at WPPP but not at the full-length FxxYxIWPPP motif (Figure 3-7E). Translation of the MWPPP peptide led to the synthesis of a ~20 kD intermediate, consistent with the size of the

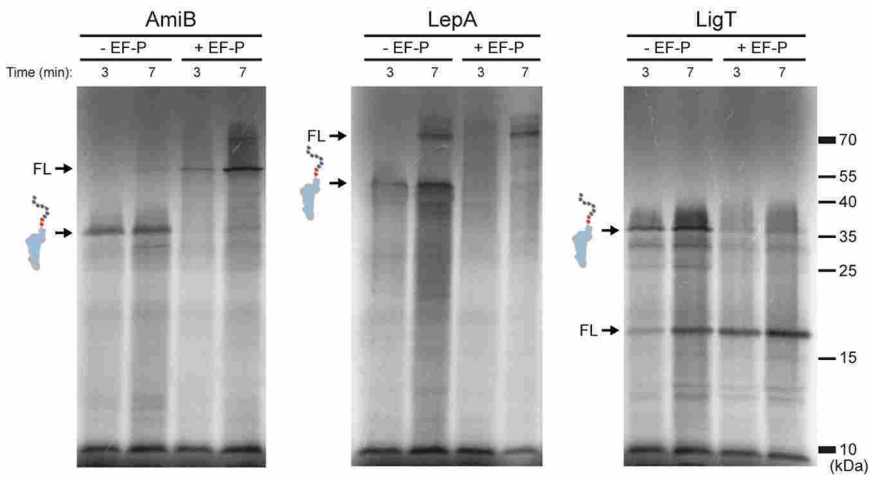


Figure 3-9. Endogenous genes prevent peptide release**.** Three endogenous *E. coli* genes, *amiB*, *lepA*, and *ligT,* were translated *in vitro* with [^{35}S]-methionine in the presence or absence or presence of EF-P. The reactions were stopped after 3 or 7 minutes. Peptidyl-tRNAs are labeled with a cartoon and the full-length, released protein is labeled 'FL'.

tRNA with a tetrapeptide attached. This intermediate does not accumulate, however, when EF-P is added to the translation reaction (Figure 3-7E). In contrast, during MFQKYGIWPPP synthesis, peptidyl-tRNA accumulates regardless of whether EF-P is added or not. Taken together, these data show that our motifs identified by genetic selection are resistant to EF-P, further defining its scope of action.

Stalling motifs have been selected against

The motifs we identified from random libraries induce ribosome stalling both in vivo and in vitro.What effect would these motifs have on the synthesis of endogenous proteins *in vivo*? Are they found in endogenous bacterial proteomes? Or have they been selected against?

Analysis of bacterial proteomes supports the idea that some of these motifs have in fact been selected against. We searched for stalling motifs in ~13.7 million bacterial proteins in the RefSeq database from 4,277 organisms in the bacterial kingdom. We found that the DTS motif occurred sixfold less at the C-terminus of proteins than one would expect based on amino acid frequencies in this dataset. Glu-Pro-stop and Pro-Pro stop were underrepresented by 1.7- and 1.8-

fold, respectively, and Asp-Pro-stop was underrepresented by 3.6-fold. Elongation motifs RSPP and HGPP were also underrepresented, by 3.3- and 2.8-fold, respectively. All of these changes were significant with $P < 10^{-6}$. These findings suggest that stalling motifs may have been broadly selected against during the course of evolution. In contrast, the six PP(X) motifs that stall *in vitro* occur at expected levels in bacterial proteomes; when analyzed together, they were slightly enriched at < 0.1% above the expected frequency ($P > 0.35$). We conclude that there is no evidence of selection against the six stalling PP(X) motifs.

Stalling occurs at several motifs in endogenous E. coli proteins

Ribosome-profiling data published by Weissman and co-workers allow us to quantify ribosome occupancy at potential stalling sites in endogenous *E. coli* proteins [192]. We define a pause score as the number of reads at the pause site divided by the median number of reads for the entire open reading frame. Sites where ribosomes are enriched have higher pause scores than sites of low ribosome occupancy. We calculated pause scores for all 8,000 tripeptide combinations. The highest reliable pause score was 5.9 +/– 0.3 for the tripeptide GGT. Indeed, many of the top 100 tripeptide motifs are rich in Gly residues (Table 3-2). In contrast, PPP has a pause score of 4.4 +/– 0.4 and other PPX stalling motifs have lower scores (PPD = 4.3 +/– 0.5; PPG = 3.9 +/– 0.4; PPE = 3.1 +/– 0.3).

Stalling motifs identified in our genetic selection have higher pause scores in the ribosome profiling data. We first looked at motifs that block termination. No proteins in the *E. coli* MG1655 genome end in Pro-Pro-stop. At the Asp-Pro-stop sequence in the SgrR gene, the ribosome pauses robustly with a score of 84, meaning that a ribosome is 84 times more likely to be found at this site than a typical position in the SgrR open reading frame. The Glu-Pro-stop

Rank	motif	pause score	error	% error	# of hits	Rank	motif	pause score	error	% error	# of hits
1	GGT	5.88	0.33	5.7	302	51	GTD	4.30	0.30	6.9	166
2	GGS	5.68	0.24	4.2	311	52	GGQ	4.27	0.31	7.4	194
3	GAG	5.44	0.26	4.7	521	53	GGH	4.25	0.38	9.0	103
4	WGP	5.38	1.09	20.2	25	54	CGY	4.24	0.91	21.5	34
5	GTC	5.31	0.81	15.2	25	55	GTN	4.23	0.56	13.2	94
6	MGT	5.30	0.78	14.7	85	56	MGI	4.23	0.40	9.5	126
7	GGA	5.28	0.30	5.6	436	57	GPT	4.21	0.41	9.7	118
8	GGG	5.28	0.24	4.6	444	58	GTY	4.21	0.36	8.6	79
9	GTI	5.21	0.47	8.9	221	59	GAI	4.20	0.21	5.1	402
10	GSD	5.13	0.57	11.1	186	60	TGT	4.19	0.33	7.8	235
11	GTT	5.09	0.66	13.0	225	61	GAS	4.19	0.17	4.1	375
12	GCG	5.05	0.42	8.3	128	62	DGP	4.19	0.38	9.1	111
13	GDG	5.05	0.46	9.1	233	63	GYG	4.16	0.31	7.4	193
14	GGV	4.96	0.24	4.8	429	64	EGY	4.16	0.59	14.3	151
15	GPN	4.94	0.59	12.1	79	65	GTV	4.16	0.25	6.1	299
16	GGF	4.93	0.35	7.1	221	66	GSS	4.16	0.29	7.0	230
17	GDC	4.91	0.65	13.2	33	67	GVG	4.16	0.19	4.6	418
18	GAT	4.87	0.27	5.6	370	68	GTA	4.15	0.42	10.1	328
19	GGD	4.85	0.31	6.3	244	69	WGA	4.15	0.49	11.9	78
20	GGI	4.78	0.20	4.3	357	70	VYG	4.13	0.86	20.8	158
21	CTF	4.73	1.12	23.6	22	71	GKS	4.13	0.21	5.2	299
22	GTG	4.71	0.24	5.1	362	72	APP	4.12	0.41	10.0	107
23	GVT	4.70	0.24	5.2	338	73	GTW	4.11	0.44	10.7	46
24	GWT	4.70	0.69	14.7	57	74	GSV	4.10	0.35	8.5	325
25	GST	4.68	0.35	7.5	222	75	GAW	4.09	0.67	16.4	95
26	GPG	4.68	0.30	6.3	167	76	GPI	4.08	0.34	8.4	101
27	GGM	4.64	0.41	8.9	182	77	GAD	4.08	0.23	5.7	356
28	GCI	4.64	0.86	18.5	71	78	RR*	4.08	0.83	20.3	33
29	VA*	4.55	0.93	20.5	26	79	GTR	4.07	0.33	8.2	174
30	GSG	4.54	0.20	4.3	352	80	GAA	4.07	0.19	4.6	570
31	GGY	4.53	0.34	7.6	172	81	GCT	4.07	0.45	11.1	56
32	AGP	4.53	0.94	20.8	158	82	GWG	4.07	0.52	12.8	64
33	GAK	4.52	0.53	11.8	270	83	GFG	4.06	0.32	7.9	289
34	GTS	4.52	0.29	6.4	200	84	AGT	4.06	0.47	11.5	321
35	GSY	4.51	0.41	9.0	112	85	GIC	4.05	0.41	10.1	53
36	LS*	4.50	0.95	21.2	21	86	GSI	4.04	0.26	6.4	224
37	GTH	4.49	0.65	14.5	76	87	GVM	4.03	0.70	17.4	167
38	GDT	4.46	0.48	10.8	234	88	GWA	4.02	0.68	17.0	79
39	GGK	4.45	0.25	5.7	266	89	GSP	4.02	0.41	10.3	138
40	WGI	4.44	0.72	16.1	46	90	GIV	4.02	0.21	5.3	397
41	GIT	4.44	0.26	5.8	314	91	GSA	4.02	0.28	6.9	303
42	GIF	4.43	0.73	16.5	151	92	GTL	4.01	0.16	4.0	477
43	**PPP**	4.39	0.42	9.6	81	93	GIQ	4.01	0.30	7.5	138
44	GVY	4.38	0.57	13.0	154	94	DGS	4.00	0.29	7.4	257
45	GGC	4.37	0.40	9.2	74	95	GGN	4.00	0.44	11.0	207
46	GTF	4.36	0.40	9.1	138	96	KGT	3.99	0.36	9.1	132
47	GPQ	4.35	1.05	24.1	90	97	DGA	3.99	0.62	15.5	250
48	**PPD**	4.33	0.49	11.3	67	98	GEG	3.99	0.25	6.3	270
49	WGT	4.32	0.67	15.5	39	99	GCE	3.99	0.43	10.8	69
50	EPP	4.32	0.77	17.8	56	100	GKT	3.97	0.18	4.4	386

Table 3-2. Pause scores of stalling peptides**.** Pause scores were calculated for all 8,000 tripeptides using published ribosome profiling data from *E. coli* MG1655 (see ref. (193)). The pause score is the ribosome density at the three codons of the tripeptide motif divided by the median density for the opening reading frame. Tripeptides were excluded if they had fewer than 20 occurrences (hits) or pause scores with higher than 25% error.

motif in TreF has a pause score of 28. None of our termination motifs (T1-T3) are found in this strain, but the RxPP elongation motif in the RecG protein has a pause score of 14, and the HGPP motif in the YaaX protein has a pause score of 10. Ribosome profiles for these four genes are shown in Figure 3-10. These data argue that these motifs indeed delay ribosomes *in vivo*, even in endogenous proteins.

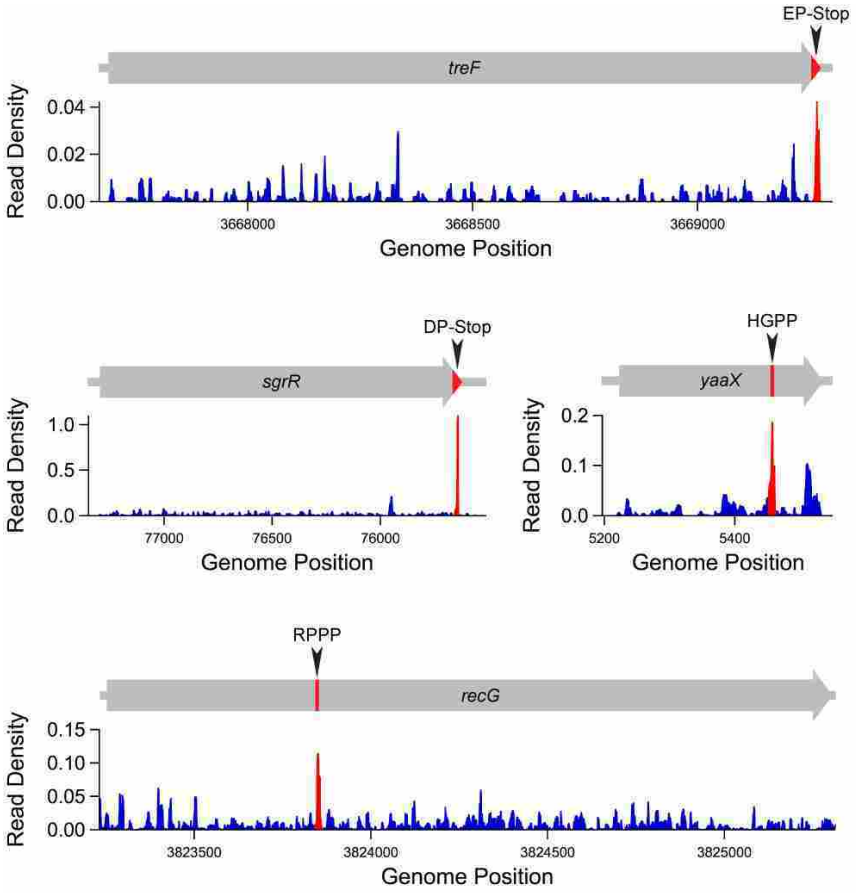


Figure 3-10. Ribosome profiling data highlight pauses at stalling motifs in vivo. Ribosome density is shown across four genes that contain stalling motifs highlighted in red. Read densities are reported in units of reads per million mapped reads (rpM). These analyses are of published datasets of *E. coli* MG1655 ribosome profiling experiments (193).

Discussion

Our selection yielded novel nascent peptide motifs that induce ribosome stalling and tmRNA tagging. Several lines of evidence support the claim that the peptide is the primary cause of stalling. First, stalling occurs in a reconstituted *in vitro* translation system, where cleavage or degradation of the mRNA is ruled out. Second, the consensus sequences show a high degree of conservation of the first two nucleotides in a codon but considerable variation at the wobble position. Third, the nascent peptide and incoming amino acid seem to play a key role in slowing reaction rates in kinetic assays. We cannot rule out a role for associated tRNAs, however, since many aspects of tRNA structure and function are common between the isoacceptors for a given amino acid.

Several of the peptides we identified inhibit translational termination with small, polar residues at the C-terminus. For example, the sequence DTS-stop is sufficient for stalling. This motif expands to D/N, T/S, and S/G if additional aromatic residues are present upstream. With an even longer stretch of hydrophobic residues, termination is slowed by a single Thr residue at the C-terminus of the WILFxxT-stop motif. Perhaps the potential of Thr to block peptide release explains why Thr is the most underrepresented residue at the C-terminus of bacterial proteins (2.1-fold). These findings define a new class of stalling peptides and argue that release factors are sensitive to the sequence of the polypeptide being released.

In these examples, upstream aromatic or hydrophobic residues probably enhance stalling through increasing the binding of the peptide to the ribosomal exit tunnel. Conserved upstream residues in natural stalling motifs such as SecM and TnaC work in the same way [109,115]. But in some motifs we identified, the upstream sequence has the opposite effect, suppressing stalling by a motif that otherwise stalls effectively. In the R/HxPP motif, for example, upstream residues

impose sequence specificity for either Arg or His. In addition, although the minimal motif R/HxPP is sufficient to block peptidyl transfer, stalling is suppressed if the peptide is of intermediate length (~11 amino acids), perhaps because specific upstream residues induce a peptide conformation that disrupts interactions between the ribosome and the minimal motif. We speculate that longer versions of this motif (~20 amino acids) stall efficiently because the N-terminus of the peptide moves past the L4/L22 constriction, restricting which conformations are available to the nascent peptide.

Proline poses a particular challenge for the ribosome, serving as both a poor peptidyl donor and a poor peptidyl acceptor [117,118]. These roles are combined in the PPP motif, where the ribosome stalls with the second Pro codon in the P site and the third in the A site, such that transfer of peptidyl-tRNA to Pro-tRNA is blocked. We found that peptidyl transfer to Asn-, Asp-, Glu-, Gly-, and Trp-tRNA is also inhibited after two Pro residues. Two of these tripeptide motifs, PPD and PPE, were recently shown to pause ribosomes in mammalian cells [193], suggesting that stalling at Pro-rich motifs may be a general phenomenon, not limited to bacteria.

Stalling by the PPP motif is abolished by EF-P, explaining how ~100 *E. coli* proteins containing this motif are translated *in vivo* [184,185]. Proteins with PPD, PPE, PPG, PPN, or PPW probably also require EF-P for their synthesis. The fact that the PP(X) motifs are not underrepresented in bacterial proteomes suggests that EF-P alleviates stalling by these motifs efficiently. In contrast, EF-P has no effect on our newly identified motifs, even those including proline residues, such as FxxYxIWPPP and RxPP. This makes sense as these were selected for their ability to induce tmRNA tagging *in vivo*, where EF-P is present. We speculate that the aromatic residues in FxxYxIWPPP interact with the ribosomal exit tunnel to stabilize the stalled conformation, blocking the activity of EF-P, which could otherwise resolve stalling by the PPP

motif. In addition, it seems that EF-P primarily affects peptidyl transfer and not peptide release, given that it does not alleviate stalling at Glu-Pro-stop, even though stalling occurs with a Pro codon in the P site. These findings begin to define the scope of EF-P's ability to relieve translational pauses.

The requirement in the PP(X) motifs for specific aminoacyl-tRNAs bound in the A site is intriguing. In a similar manner, peptidyl transfer to Asp-, Glu-, Gly-, and Trp-tRNA were also found to act as poor peptidyl acceptors in the analysis of the ErmAL1 stalling peptide by Mankin and co-workers [97]. In contrast, the strongest stalling in the ErmAL1 study was with Lys-, Arg-, and His-tRNA, but these did not stall robustly after two Pro codons. It appears that different nascent peptides modulate the reactivity of different subsets of aminoacyl-tRNAs.

Insight into the mechanism of ribosome stalling has been limited by the lack of biochemical assays to directly measure individual steps in the synthesis of stalling peptides. In many studies, stalling is inferred from changes in gene expression or tagging of the protein by tmRNA. Although stalling can be detected directly in toeprinting assays, this method cannot always pinpoint the precise step that is blocked nor determine reaction rates. In theory, many steps in the translational cycle could be inhibited during ribosome stalling, including tRNA or factor binding, translocation, peptide-bond formation, and peptide release. We were able to recapitulate stalling in pre-steady state kinetic assays using our new motifs, showing that peptide release and peptidyl transfer are inhibited by roughly 100- and 1000-fold, respectively. The fact that our motifs are short makes interpretation of their interactions with the ribosome far simpler than interpreting the function of natural motifs like SecM. These new stalling peptides cannot contact the L4/L22 constriction; we speculate that they interact directly with rRNA nucleotides near the peptidyl-transferase center to induce conformational changes that inhibit chemistry at

the active site. Going forward, these assays will allow us to determine the role of the peptide, tRNA, and rRNA sequence in the stalling mechanism.

What do our data tell us about the likelihood of finding more stalling motifs in the future? On one hand, well-characterized motifs containing Pro residues appeared often in the two-hybrid selection. The Pro-stop motif, for example, was found in half of the clones that induce tmRNA tagging. On the other hand, several new motifs were discovered that block termination with polar residues, and the cause of stalling remains unknown for several new elongation motifs. Given the tiny fraction of sequence space we accessed in our library, and the high rate of survival in the selection, it seems likely that more motifs remain undiscovered.

Ribosome profiling has potential for shedding light on variation in translational rates and for identifying new stalling motifs. A recent ribosome profiling study showed that most strong pauses could be explained by Shine-Dalgarno-like sequences 6 to 12 nt upstream of the pause site [192]. In the E5 motif (Table 3-1), stalling occurs with the nucleotides GGAGGA within 6 to 12 nucleotides of the A site codon, consistent with this report. Our analysis of the same profiling dataset revealed that Gly-rich tripeptide motifs have the highest pause scores in *E. coli*. Since Gly is encoded by GGN, it may be tempting to attribute this result to SD-like sequences. However, the middle codon of the tripeptide is positioned in the P site in our analysis, making it unlikely that these codons bind strongly to the anti-SD sequence in 16S rRNA during pausing. Consistent with the high pause score for the GGG tripeptide (Table 3-2), we note that the E6 motif stalls on three Gly codons with the second Gly in the P site (Table 3-1).

Our data argue that bacterial proteomes have been shaped by the demand for translational efficiency. The motifs that we identified were 3- to 6-fold underrepresented in a dataset including ~13.7 million bacterial proteins. For those stalling motifs that are retained in the *E. coli*

genome, we find substantial accumulation of ribosome density at the motif in ribosome profiling data sets. This shows that short motifs stall on endogenous proteins *in vivo* and have the potential to modulate protein synthesis rates in a biologically relevant way.

Experimental Procedures

Two-Hybrid Selection

The selection for stalling motifs is based on Bacteriomatch II (Agilent). Residues 1-117 of SspB were fused to the C-terminus of RNA polymerase alpha. To generate the library, 20 random codons were added to the 3'-end of the cI coding sequence. A modified tmRNA encoding the ANDENYALDD tag was expressed in the reporter strain, which lacks *hisB* and expresses *HIS3* from a weak promoter downstream of cI binding sites. A library of 3 x 10^8 cI mutants was introduced into the reporter strain and transformants were plated at 30 °C on M9 minimal media with His dropout supplement, 10 μM IPTG, and 5 mM 3-aminotriazole. The cI gene was PCR amplified from the pool of surviving clones, inserted into fresh expression vector, and passaged through the selection again. About 10% of clones survived in this second round; 150 were sequenced from this enriched pool. Additional details on plasmid construction and the cI controls are given in the Experimental Procedures.

In Vitro Translation Assays

In vitro translation and toeprinting assays were performed in the PURExpress translation system (New England Biolabs) as described previously [98]. Each set of experiments was repeated at least twice. The DNA constructs are described in the Experimental Procedures. Where

indicated, 1 μM EF-P was added to the translation reaction. Modified EF-P was purified as described previously [194].

Mass Spectrometry

cI clones were expressed in pET-15b in BL21 (DE3) cells together with a modified tmRNA encoding the ANDHHHHHHD tag. Tagged cI was purified over Ni-NTA resin, digested with trypsin, and the C-terminal tagged peptide purified and analyzed as previously described [98].

Kinetics

Ribosome nascent chain complexes were assembled and reacted with excess release factor 1 or excess ternary complex containing EF-Tu, GTP, and aminoacyl-tRNA at 23 °C in polymix buffer. Reported rates are the average of three independent experiments and standard error is given. Details of the reaction conditions and materials are discussed in the Experimental Procedures.

Statistical Analysis

We developed a likelihood ratio test to detect conserved motifs in sequences of surviving clones from the two-hybrid selection. Under the null hypothesis, we assume that the mutation rate is uniform across all the bases in the sequence. Enriched motifs are therefore the locations that deviate from the overall frequencies across the entire 20-codon sequence. Details are provided in the Experimental Procedures.

Pause scores for specific stalling motifs were obtained from *E. coli* MG1655 ribosome profiling data from the Weissman lab [192]. The pause score for known motifs was computed by

averaging the number of reads per base over the P-site codon and two codons downstream, normalized to the median number of reads per base for the given open reading frame. This approach was used because it effectively captured the increased read density in the SecM motif. Pause scores for tripeptides (Table 3-2) were computed with the average of density at all three codons and were calculated using only well-translated genes (> 10 rpkm). Motifs with fewer than 20 occurrences in the genome were discarded, as were motifs with higher than 25% error in the pause score. The error in the distribution in the pause scores was computed by bootstrapping.

Plasmid Construction

The two-hybrid selection uses three plasmids. The first, pSP100, has a pCDF origin and expresses tmRNA-DD and β-lactamase. The second, pTRG, contains the ColE1 origin and encodes the alpha subunit of RNA polymerase (RpoA) and TetR (Agilent Technologies). We amplified the first 117 codons of the *sspB* gene using the primers CCGCAAGAATTCA-GATGGATTTGTCACAGCTAACACCACGTCG and GATCTCACTAGTTTACAT-GATGCTGGTATCTTCATCGTAGGCAGC. This PCR product was subsequently cloned at the 3'-end of the *rpoA* gene in pTRG with EcoRI and SpeI.

The cI control constructs used in Figure 3-1B (cI-tag, cI-Glu-Pro-stop, and cI-nonstop) were expressed from the pBT vector (CamR and p15A origin, Agilent Technologies). The cI-alone construct is simply the original pBT plasmid. The primers TTGGCGCGGCCGCAGGG-GAGCCAGCCGCAAACGACGAAAACTACGCTTTAGACGACTAAGATCTTAGGCG and CGCCTAAGATCTTAGTCGTCTAAAGCGTAGTTTTCGTCGTTTGCGGCTGGCTCCCCTG-CGGCCGCGCCAA were used to amplify the cI gene and add the tmRNA tag (AANDENYALDD) to the C-terminus of cI. The primers TTGGCGCGGCCGCATCTG-

AACCGTGACTAAGATCTTAGGCG and CGCCTAAGATCTTAGTCACGGTTCAGATGCG-GCCGCGCCAA were used to add the Glu-Pro-Stop sequence. The primers GATGATCGGCCGGCAGCCCGCCTAATGAGCGGGCTTTTTTTAGATCTGATGAT and ATCATCAGATCTAAAAAAAGCCCGCTCATTAGGCGGGCTGCCGGCCGATCATC were used to add the *trpA* terminator after the cI gene, in order to create an mRNA transcript of defined length without a stop codon (cI-nonstop). These three PCR products were ligated into the pBT plasmid following digestion with EagI and BglII.

cI clones were overexpressed for purification for MS analysis. The gene encoding the cI-stalling motif fusion was PCR amplified from the pBT plasmid using the following primers: GATATACCATGGGCAGCACAAAAAAGAAACCATTAACACAAG and GCAGCCGGATCCCCGGCGCGCCTAAGATCT. These PCR products were digested with NcoI and BamHI and cloned into the pET15b vector. The resulting plasmids and tmRNA-His_6 plasmid pCH201 [101] were used to transform BL21 (DE3) *E. coli* cells. The cI protein was expressed, purified, and analyzed as described previously [98].

Library Construction

20 random codons were added to the 3'-end of the lambda cI protein by PCR with the primers GATAAAATATTTCTAGATTTCAGTGCAATTTATCTCT and the reverse primer TTATGCAGATCTTTACTTACTTAN_{60}TGCGGCCGCGCCAAACGTCTC where N is an equal mixture of all four bases. The PCR product was cloned into pBT using XbaI and BglII, and the resulting plasmids were amplified in XL1-Blue. 30% mutagenesis libraries of individual clones were created with the same scheme, except that the reverse primer contained the motif

sequence, not N_{60}. The motif sequence was synthesized with phosphoramidite mixtures containing 70% original nucleotide and 10% each of the other three nucleotides.

In Vitro Translation Constructs

All toeprinting DNA templates start with the following 5'-sequence, including a T7 promoter, ribosome binding site, and start codon which is underlined: CTGTACAT-TAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATG. The 3'-end of all templates includes the following primer-binding site; the DNA primer in all the experiments was NV1, GTTAATAAGCAAAATTCATTATAACC. To PCR amplify the cI clones for toeprinting, including 18 amino acids of cI and the 20-codon variable region, we used the 3'-primer GGTTATAATGAATTTTGCTTATTAACGGTAGCCAGCAGCATCCT. This places the end of the random sequence 40 nt from the NV1 primer-binding site. For the toeprints shown in Figures 3-2 and 3-4, the 5'-primer was CTGTACATTAATACGACTCACTATAGGGAGATTTTA-TAAGGAGGAAAAAATATGTGTTCCGTTGTGGGGAAAGTTAT.

For toeprinting analysis of PPP stalling, three endogenous genes were PCR amplified from genomic DNA. The constructs were synthesized with the upstream sequence shown above, such that translation starts with the natural ATG codon. The NV1 primer-binding site was added 51 nt downstream of the second Pro codon in the PPP motif in the gene. The primers used were: *lepA*,

AGCTACCGGCCGTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATA TGAAGAATATACGTAACTTTTCGATCATAGCTCAC and

AGCTACCTCGAGGGTTATAATGAATTTTGCTTATTAACCCATGAGTCGATAATTAGT GCCTGCAACGG;

ligT,

AGCTACCGGCCGTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATA

TGTCTGAACCGCAACGTCTGTTCTTTGCT and

AGCTACCTCGAGGGTTATAATGAATTTTGCTTATTAACGGAGGCGTAAAGGGTGAAC

TCCGTCACCGC;

amiB,

AGCTACCGGCCGTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATA

TGATGTATCGCATCAGAAATTGGTTGGTAGC and

AGCTACCTCGAGGGTTATAATGAATTTTGCTTATTAACAACCGCAGGCGTTTCAACG

CGTTTCGCAAC.

Analysis of WPPP and FxxYxIWPPP was performed with the following dsDNA constructs: for WPPP,

CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGAC

CATGATTACGAATTCGAGCTCATGGCCACCGCCATCGATTCGGCATGCAAGCTTGGC

ACTGGCCGTCGTTTTACAACGTCGTGTTAATAAGCAAAATTCATTATAACC encodes

MTMITNSSSWPPPSIRHASLALAVVLQRRVNKQNSL*. For the whole motif, the construct

CTGTACATTAATACGACTCACTATAGGGAGATTTTATAAGGAGGAAAAAATATGAC

CATGATTACGAATTCGAGCTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTTG

GCCACCGCCATCGATTCGGCATGCAAGCTTGGCACTGGCCGTCGTTTTACAACGTCG

TGTTAATAAGCAAAATTCATTATAACC encodes

MTMITNSSSLQKRLFQKYGIWPPPSIRHASLALAVVLQRRVNKQNSL*.

Kinetics

Materials: E. coli MRE600 tightly coupled 70S ribosomes were prepared as described previously [60]. Overexpressed native IF1 and IF3 and His-tagged IF2 were purified as described [195]. Amino-terminally His-tagged RF1 and aminoacyl-tRNA synthetases were expressed and purified as previously described [196]. His-tagged EF-Tu and EF-G were purified over Ni-NTA resin and the His-tag was removed by tobacco etch virus protease, followed by a second passage over a Ni-NTA column [197]. $tRNA^{fMet}$, $tRNA^{Phe}$, $tRNA^{Arg}$, $tRNA^{Glu}$ were purchased from Chemical Block (Russia). $tRNA^{Pro}$ and $tRNA^{Ala}$ were purified from bulk *E. coli* tRNA (Roche) with biotinylated oligos Pro1 CCGAACGAAGTGCGCTACCAGGCTG3BioTEG, Pro2 CCCATGACGGTGCGCTACCAGGCTG3BioTEG, and Ala1 GCAAAGCAGGCGCTCTCCCAGCTGA3BioTEG as described [198]. mRNA templates were synthesized by T7 RNA polymerase using DNA templates annealed to a short primer corresponding to the minimal promoter sequence. The mRNA transcripts have the following sequence: GGGUGUCUUGCGAGGAUAAGUGCAUU-AUG(X)UUUGCCCUUCUGUAGCCA where the start codon is underlined and additional codons are inserted at the X site. UAA stop codons are used in termination motifs.

tRNA aminoacylation: Initiator $tRNA^{fMet}$ was aminoacylated with formylated radiolabelled [^{35}S]-methionine using MetRS and methionyl-tRNA formyltransferase as described [199]. Pure tRNAs were charged by incubating the tRNA at 5 μM with the corresponding synthetase (~ 1 μM) in the presence of 2 mM ATP and 50 μM amino acid for 30 min at 37 °C in the following buffer: 100 mM HEPES-KOH pH 7.6, 1 mM DTT, 10 mM KCl, 20 mM $MgCl_2$. The aminoacyl-tRNA was purified by extraction by phenol and $CHCl_3$, precipitated with ethanol, and resuspended in 2 mM NaOAc pH 5.0. The following tRNAs were aminoacylated as purified

tRNAs: Pro-tRNA, Ala-tRNA, Arg-tRNA, Glu-tRNA, and Phe-tRNA. Bulk *E. coli* tRNA (Roche) was charged with a similar procedure in making the DTSA and DTSF aminoacyl-tRNA mixes for the DTS complex and related mutants, except the tRNA concentration was 100 μM and the corresponding aminoacyl-tRNA synthetases and amino acids were added.

Ribosome complex formation: Initiation complexes were prepared by incubating 70S ribosomes (2 μM) with IF1, IF2, IF3, fMet-tRNA (3 μM each), and mRNA (6 μM) in polymix buffer with 2 mM GTP at 37 °C for 45 min. 100 μL of the resulting complex was reacted with 200 μL of pre-incubated mixture containing EF-Tu (15 μM), charged tRNA (2 μM), EF-G (2 μM), and GTP (2 mM) in polymix buffer for 37 °C for 5 min. The complexes were then purified over a 1 mL sucrose cushion (1.1 M sucrose, 20 mM Tris-HCl, pH 7.5, 500 mM NH_4Cl, 10 mM $MgCl_2$, 0.5 mM EDTA) and spun at 260,000 *g* in a TLA100.3 rotor for 2 h. The pellet was resuspended in polymix buffer (95 mM KCl, 5 mM NH_4Cl, 5 mM magnesium acetate, 0.5 mM $CaCl_2$, 8 mM putrescine, 1 mM spermidine, 5 mM potassium phosphate pH 7.5, 1 mM DTT), aliquoted, and stored at –80 °C.

Release assays: ~25 nM RNCs were incubated with 5 μM RF1 in polymix buffer at 23 °C. The reaction was stopped at various time points by addition of formic acid to a final concentration of 1%. Reactions with fast rate constants (> 0.1 s^{-1}) were performed an RQF-3 quench-flow instrument (KinTek). Released peptides were separated from unreacted peptidyl-tRNA on cellulose TLC plates using electrophoretic TLC in pyridine-acetate buffer pH 2.8 [183]. The fraction of released peptide at each time point was quantified and plotted against time and the data fit with a single exponential equation. The reported rates are the averages of three separate experiments, with the standard error given.

Peptidyl transfer assays: 40 μM EF-Tu was incubated with 2 mM GTP in polymix buffer for 15 min at 37 °C to exchange GDP for GTP. This reaction was then combined with aminoacyl-tRNA to form a ternary complex, with final concentrations of 20 μM EF-Tu, 2 μM charged tRNA, and 2 mM GTP in polymix buffer. The complex was incubated for 10 min on ice, then reacted at 23 °C in polymix buffer with an equal volume of ~50 nM RNC. The reaction was stopped by addition of KOH to a final concentration of 100 mM. Peptides were resolved on electrophoretic TLC and analyzed as above. The reported rates are the averages of three separate experiments, with the standard error given.

Statistical Analysis of Peptide Libraries

We developed a likelihood ratio test to detect conserved motifs in sequences of surviving clones from the two-hybrid selection. Under the null hypothesis, we assume that the mutation rate is uniform across all the bases in the sequence. Enriched motifs are therefore the locations that deviate from the overall frequencies across the entire 20-codon sequence. Specifically, we let $X_{i(k)}$ be the *k*th ordered base count at the *i*th position. We then define:

$$p_{(k)} = \frac{\sum_{i=1}^{N} X_{i(k)}}{\sum_{i=1}^{N}\sum_{i=1}^{4} X_{i(i)}} \text{ and } r_{i(k)} = \frac{X_{i(k)}}{\sum_{i=1}^{4} X_{i(i)}},$$

where $p_{(k)}$ represents the estimate of the base frequency of the *k*th ordered base across all positions under the null hypothesis, and $r_{i(k)}$ estimates the *k*th ordered frequency at the *i*th at position. Assuming that the base positions are independent of each other and that the base occurrences across the sequences follow a multinomial distribution, we define our likelihood ratio test statistic for position *i* as:

$$\Lambda_i = \frac{p_{(4)}^{X_{i(4)}} p_{(3)}^{X_{i(3)}} p_{(2)}^{X_{i(2)}} p_{(1)}^{X_{i(1)}}}{r_{i(4)}^{X_{i(4)}} r_{i(3)}^{X_{i(3)}} r_{i(2)}^{X_{i(2)}} r_{i(1)}^{X_{i(1)}}}$$

We note that $-2\log(\Lambda_i)$ approximately follows a chi-squared distribution with three degrees of freedom. We evaluated the significance of each codon, by summing the logged ratio statistics for all three positions in the codon, and then appropriately comparing the codon level statistic to a chi-squared distribution with nine degrees of freedom to obtain a p-value.

Chapter 4. eIF5A Promotes Translation of Polyproline Motifs

Author's Note: This chapter details the characterization of translation factor eIF5a and its ability to relieve ribosome stalling, for which I performed the toeprinting assays and aided with the kinetic assays. The results of this study were published in Molecular Cell in 2013 [200].

Abstract

Translation factor eIF5A, containing the unique amino acid hypusine, was originally shown to stimulate Met-puromycin synthesis, a model assay for peptide bond formation. More recently, eIF5A was shown to promote translation elongation; however, its precise requirement in protein synthesis remains elusive. We use in vivo assays in yeast and in vitro reconstituted translation assays to reveal a specific requirement for eIF5A to promote peptide bond formation between consecutive Pro residues. Addition of eIF5A relieves ribosomal stalling during translation of three consecutive Pro residues in vitro, and loss of eIF5A function impairs translation of polyproline-containing proteins in vivo. Hydroxyl radical probing experiments localized eIF5A near the E site of the ribosome with its hypusine residue adjacent to the acceptor stem of the P site tRNA. Thus, eIF5A, like its bacterial ortholog EFP, is proposed to stimulate the peptidyl transferase activity of the ribosome and facilitate the reactivity of poor substrates like Pro.

Introduction

Ribosomes catalyze protein synthesis with the assistance of transfer RNAs (tRNAs) and translation factors. There are three tRNA binding sites on the ribosome: a centrally located peptidyl-tRNA (P) binding site, an aminoacyl-tRNA (A) binding site, and an exit (E) site that

binds deacylated tRNA following its transfer from the P site. Translation initiation factors aid in the assembly of an 80S ribosome in eukaryotes in which the initiator methionyl-tRNA (Met-$tRNA_i^{Met}$) is bound in the P site with its anticodon base-paired to the start codon on the messenger RNA (mRNA). To extend the polypeptide, the eukaryotic elongation factor eEF1A (EFTu in bacteria) delivers aminoacyl-tRNA to the A site in a codon-dependent manner. Following accommodation of the tRNA, the amino acid attached to the A site tRNA is juxtaposed to the peptidyl portion of the P site tRNA in the active site (peptidyl transferase center [PTC]) of the large (60S) ribosomal subunit. Peptide bond formation links the extended polypeptide to the A site tRNA, leaving a deacylated tRNA in the P site. Next, elongation factor eEF2 (EFG in bacteria) promotes translocation of the tRNAs and mRNA such that the A site is vacant and ready to accept the next aminoacyl-tRNA (reviewed in [18]). A common misconception is that the ribosome is a monolithic machine that catalyzes all peptide bonds at equivalent rates regardless of the amino acid. In fact, certain residues, including the imino acid Pro, are poor substrates for peptide bond formation [118,165] Recently, it was shown that the translation elongation factor EFP is essential for translation of polyproline (polyPro) sequences by bacterial ribosomes [184,185]; however, it is currently unclear how eukaryotic ribosomes manage to synthesize peptide bonds with poor substrates.

In addition to the canonical elongation factors eEF1A and eEF2, eIF5A has also been linked to translation elongation. eIF5A was initially discovered and characterized based on its ability to stimulate Met-puromycin synthesis [201–203] a reaction analogous to the synthesis of the first peptide bond. The aminoacyl analog puromycin reacts with Met-$tRNA_i^{Met}$ bound in the P site of 80S initiation complexes. Based on its ability to stimulate Met-puromycin synthesis, eIF5A was initially thought to function as a translation initiation factor, stimulating formation or

reactivity of 80S initiation complexes. However, since the assay involves formation of a peptide bond between puromycin and methionyl-tRNA, puromycin reactivity also reports on the peptidyl transferase activity of the ribosome, a key component of translation elongation. Depletion of eIF5A in vivo or inactivation of a temperature-sensitive mutant of yeast eIF5A impaired translation elongation and stabilized polysomes in the absence of cycloheximide [204] and increased the average ribosomal transit time in vivo [204,205]. Moreover, addition of eIF5A resulted in a 2-fold stimulation in the rate of tripeptide synthesis using a reconstituted yeast in vitro translation system. Taken together, these data revealed a role for eIF5A in translation elongation. However, it is difficult to rationalize the essential requirement for eIF5A in yeast with the modest 2-fold stimulation of tripeptide synthesis, suggesting that eIF5A may have a more specialized and critical requirement in translation elongation.

eIF5A is of particular interest because it is the only protein that contains the modified amino acid hypusine and because eIF5A and hypusine have been linked to tumorigenesis and cancer [206,207]. The hypsuine modification is present in all archaea and eukaryotes that have been examined, and it is formed by the transfer of an *n*-butylamine moiety from spermidine to the ε-amino group of a specific lysine side chain (K51 in yeast eIF5A), followed by addition of a hydroxyl group. The hypusine modification is essential for eIF5A function: deoxyhypusine synthase, which catalyzes the first step in hypusine formation, is essential for yeast viability, and derivatives of eIF5A lacking hypusine fail to stimulate in vitro Met-puromycin [208,209] and tripeptide synthesis [204]. Interestingly, bacterial EFP and eIF5A are orthologs; in some bacteria, a lysine side chain in EFP corresponding to the site of hypusine modification in eIF5A is posttranslationally modified by the addition of a β-lysine residue [190,191,194,210]. Like eIF5A, EFP was found to stimulate Met-puromycin synthesis, and this activity was dependent on the β-lysine

modification [211]. Earlier studies revealed that the impact of EFP on dipeptide synthesis varied for different aminoacyl analogs [187,212], suggesting that EFP, and eIF5A by extension, may facilitate the reactivity of certain amino acids in peptide bond synthesis. Consistent with these findings, recent reports showed that EFP enhances the synthesis of proteins containing stretches of consecutive Pro residues [184,185].

Results

eIF5A Stimulates Translation through PolyPro Sequences In Vivo

To further define the role of eIF5A in translation elongation and to determine whether eIF5A, like EFP, stimulates translation of specific amino acid motifs, we monitored the expression of a set of dual-luciferase reporters in isogenic yeast strains expressing wild-type eIF5A or the temperature-sensitive eIF5A-S149P mutant [204,213]. The dual-luciferase reporters, developed by Beth Grayhack and colleagues to examine codon bias in translation [214], express a single mRNA in which the 5′ *Renilla* luciferase and 3′ firefly luciferase open reading frames (ORFs) are joined in frame by sequences encoding repeats of 10 identical codons for each of the 20 amino acids (Figure 4-1A). For the initial analysis, the inserted sequences repeated the optimal codon for each amino acid [214]. As shown in Figure 4-1B (upper panel) and as previously observed [214], the ratio of firefly to *Renilla* luciferase activity varied depending on the repeated codon. Whereas the ratios for most constructs were similar to the no-insert control, low ratios were observed for the ArgAGA and CysUGU reporters (Figure 4-1B, upper panel) and high ratios were observed with GluGAA and PheUUC codon insertions (see Figure 4-2D). These eIF5A-independent effects might reflect codon or aminoacyl-tRNA abundance or the impact of

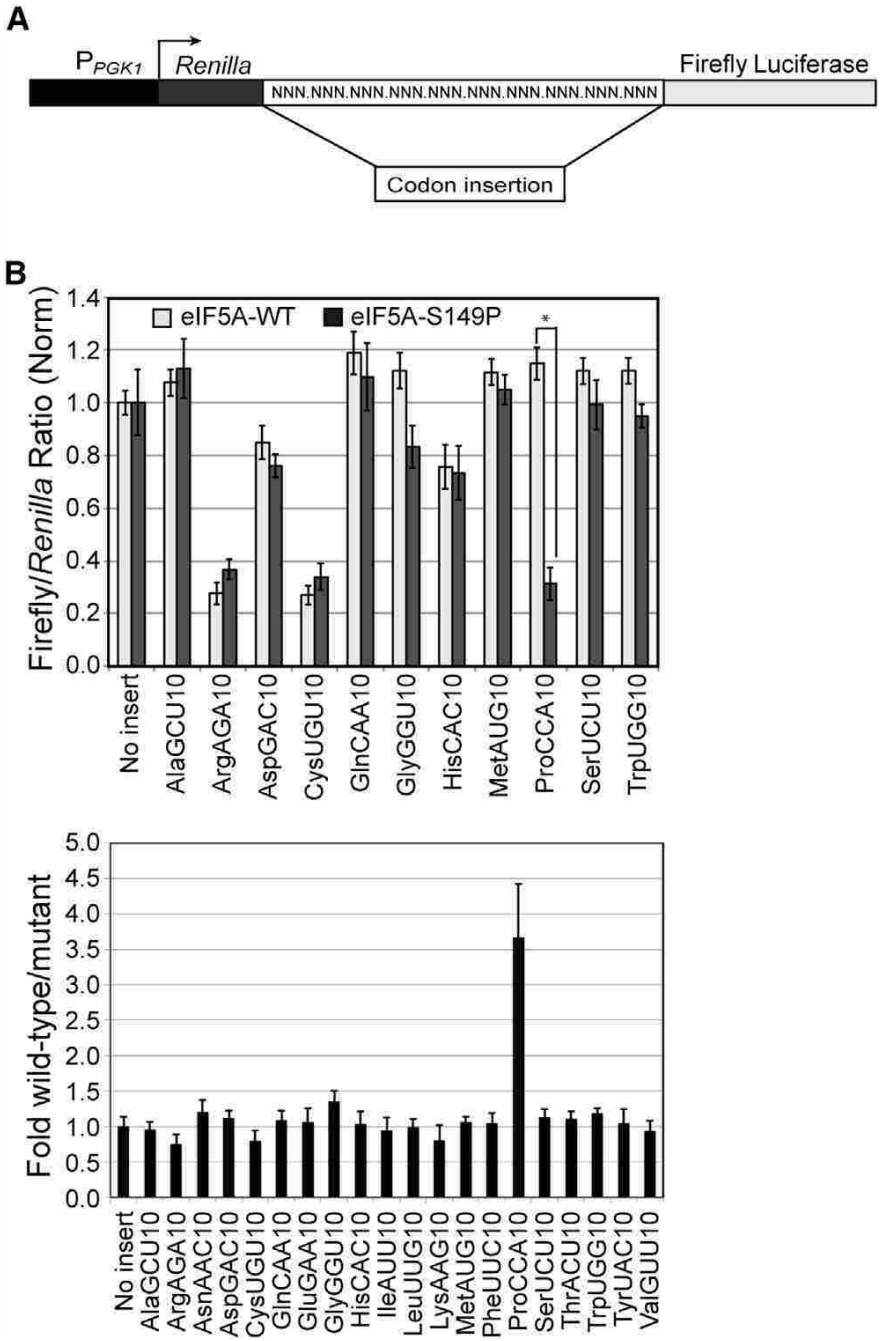
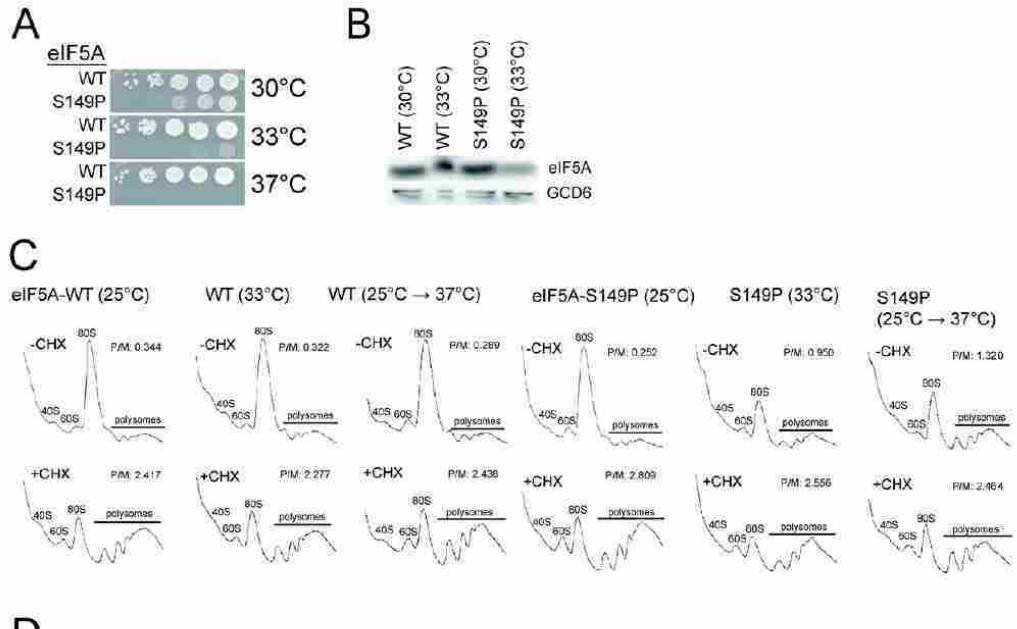


Figure 4-1. eIF5A Stimulates Translation of Polyproline Motifs in vivo**.** (A) Schematic of *Renilla*-firefly luciferase reporter construct. Codon repeats were inserted in-frame between the *Renilla* and firefly luciferase ORFs (214). (B) Dual luciferase reporter constructs containing ten repeats of the indicated codon were introduced into isogenic yeast strains expressing wild type eIF5A or temperature-sensitive eIF5A-S149P. (Top panel) Following growth at semi-permissive 33°C, luciferase activities were determined, and the firefly-to-*Renilla* luciferase ratio for each construct was normalized to the ratio obtained from controls in which the reporter contained no insert between the ORFs. (Bottom panel) The fold difference in luciferase ratios between cells expressing wild-type eIF5A and eIF5A-S149P was quantitated and then normalized to the values obtained from the no insert control. *Statistical significance for ProCCA(10) was measured by student's t-test with a p-value <0.05. Error bars were calculated as propagated standard errors of the mean for three independent transformants.



D

	Firefly/*Renilla* Ratio	
Amino Acid(10)	**eIF5A-WT (SD)**	**eIF5A-S149P (SD)**
no insert	3.09 (0.10)	6.91 (0.61)
SerUCU	3.47 (0.11)	6.85 (0.23)
HisCAC	2.34 (0.25)	5.07 (0.54)
GlnCAA	3.68 (0.22)	7.59 (0.58)
ThrACU	3.66 (0.13)	7.42 (0.05)
AsnAAC	3.97 (0.33)	7.38 (0.53)
AspGAC	3.01 (0.27)	7.17 (0.63)
GlyGGU	3.47 (0.18)	5.74 (0.23)
ProCCA	3.55 (0.15)	2.16 (0.39)
no insert	8.45 (0.27)	11.46 (0.27)
PheUUC	23.26 (0.97)	30.42 (4.11)
TrpUGG	9.46 (0.28)	10.86 (0.44)
ArgAGA	2.34 (0.35)	4.21 (0.42)
MetAUG	9.42 (0.28)	12.01 (0.58)
LysAAG	2.03 (0.50)	3.43 (0.34)
CysUGU	2.27 (0.29)	3.89 (0.57)
no insert	2.25 (0.08)	2.66 (0.12)
LeuUUG	7.22 (0.48)	8.69 (0.78)
TyrUAC	7.71 (0.81)	8.77 (1.38)
IleAUU	8.12 (1.03)	10.15 (1.35)
ValGUU	12.70 (0.82)	16.03 (2.13)
AlaGCU	2.42 (0.07)	3.00 (0.27)
GluGAA	228.98 (16.00)	254.67 (38.64)

Figure 4-2. The Temperature-sensitive eIF5A-S149P Mutant Impairs Yeast Cell Growth, Translation Elongation, and Synthesis of Polyproline Sequences, Related to Figure 4-1. (A) Isogenic wild-type and eIF5A-S149P mutants strains were grown to saturation, and 4-µl volumes of serial dilutions (OD_{600} = 1.0, 0.1, 0.01, 0.001, and 0.0001) were spotted on YPD medium and incubated for 2 days at 30°C, 33°C, and 37°C. (B) Whole cell extracts (WCEs) from yeasts strains expressing wild-type eIF5A or eIF5A-S149P and grown at permissive (30°C) and semi-permissive (33°C) temperatures were subject to immunoblot analysis using antisera specific for eIF5A or eIF2Bε (GCD6). (C) Polysome profiles were analyzed from wild type and eIF5A-S149P mutant strains grown under permissive (25°C) or semi-permissive (33°C) conditions, or following a temperature-shift from 25°C to 37°C for 2 h, were treated with 50 µg/ml cycloheximide (+CHX) or left untreated (-CHX), and WCEs were separated on sucrose gradients and fractionated to visualize polysomes and the indicated ribosomal species. Polysome/monosome (P/M) ratios were calculated by comparing the areas under the polysome and 80S peaks. (D) Summary of the firefly:*Renilla* luminescence ratios obtained following introduction of dual luciferase reporters containing 10 repeats of the indicated codon (see Fig. 4-1) into wild-type eIF5A and eIF5A-S149P mutant yeast strains and growth at 33°C. Ratios of firefly-to-*Renilla* luciferase activity were determined for three independent transformants of each construct, and the average ratio (x 10^3) and standard deviation (SD) are presented. The values in Figure 4-1 were obtained by normalizing the ratios to the respective no insert control for each of the three sets of constructs.

the inserted amino acids on luciferase activity in the bifunctional *Renilla*-firefly luciferase fusion protein.

If eIF5A stimulates the translation of specific amino acids, then the ratio of firefly to *Renilla* luciferase activity is expected to decrease when these reporters are analyzed in the strain containing eIF5A-S149P when grown at the semipermissive temperature (33°C). As shown in Figure 4-2A, the slow-growth phenotype of the eIF5A-S149P mutant at 30°C is exacerbated at 33°C, and the mutant strain fails to grow at 37°C. The impaired growth at 33°C is marked by reduced levels of eIF5A (Figure 4-2B) and by retention of polysomes in the absence of cycloheximide (Figure 4-2C), indicative of a general translation elongation defect in the strain. Analysis of all 20 luciferase reporter constructs revealed that only the Pro codon insertions revealed a strong dependence on eIF5A (Figure 4-1B, upper panel). For the ProCCA reporter, the ratio of firefly to *Renilla* luciferase in the strain expressing wild-type eIF5A was ~3.7-fold greater than the ratio observed in the strain expressing eIF5A-S149P (Figure 4-1B, lower panel), whereas this normalized ratio ranged from 0.75 (ArgAGA) to 1.35 (GlyGGU) for reporters containing any of the other 19 codon insertions.

To test whether the impaired expression of firefly luciferase from the construct containing the ProCCA codon repeats was specific to the mutation of eIF5A, two other translation elongation factors were evaluated. No significant differences in firefly:*Renilla* luciferase ratios were observed when constructs containing Pro or Ala codon insertions were examined in strains expressing temperature-sensitive mutants of translation elongation factors eEF2 or eEF3 (Figures 4-3A and 3B). Thus, polyPro peptide bond formation shows a unique dependence on eIF5A. Alternatively, this result could reflect a specific requirement for eIF5A to promote peptide bond formation by Pro-tRNA. Consistent with this hypothesis, reporters

containing 10 repeats of the Pro codons CCA, CCG, or CCU displayed a strong requirement for eIF5A, whereas no Ala codon insertions conferred a dependence on eIF5A (Figure 4-3C). While these data are not definitive, they suggest that the imino acid Pro, rather than the tRNA, likely determines the requirement for eIF5A.

To define the number of consecutive Pro residues needed to impose a requirement for eIF5A, the dual-luciferase reporters were modified to contain one, two, three, four, six, eight, or ten consecutive ProCCA or PheUUC codons. As shown in Figure 4-4, luciferase ratios for the Phe codon insertion constructs were the same in the wild-type and eIF5A-S149P mutant strains (wild-type/mutant ≈ 1.0). Likewise, insertion of one or two Pro codons did not significantly impact luciferase ratios in the eIF5A mutant strain compared to the wild-type control. In contrast, insertion of four Pro codons resulted in reduction of the luciferase ratio in the eIF5A-S149P mutant strain; some reduction may be evident with insertion of three Pro codons as well. Insertion of six, eight, or ten Pro codons further exacerbated the defect, and the normalized ratio of firefly to *Renilla* luciferase in the strain expressing wild-type eIF5A was ~3- to 4.5-fold greater than the ratio observed in cells expressing eIF5A-S149P. These results indicate that at least four (or perhaps three) consecutive Pro codons are needed to impose an eIF5A dependency on protein synthesis.

Expression of Yeast PolyPro-Containing Proteins Requires eIF5A In Vivo

Analysis of the *Saccharomyces cerevisiae* genome identified 549 proteins (out of 5,886 ORFs) that contain polyPro motifs with at least three consecutive Pro residues. To test whether expression of yeast proteins containing polyPro motifs is dependent on eIF5A, selected plasmids from the Yeast ORF Collection (Open Biosystems) were introduced into isogenic strains

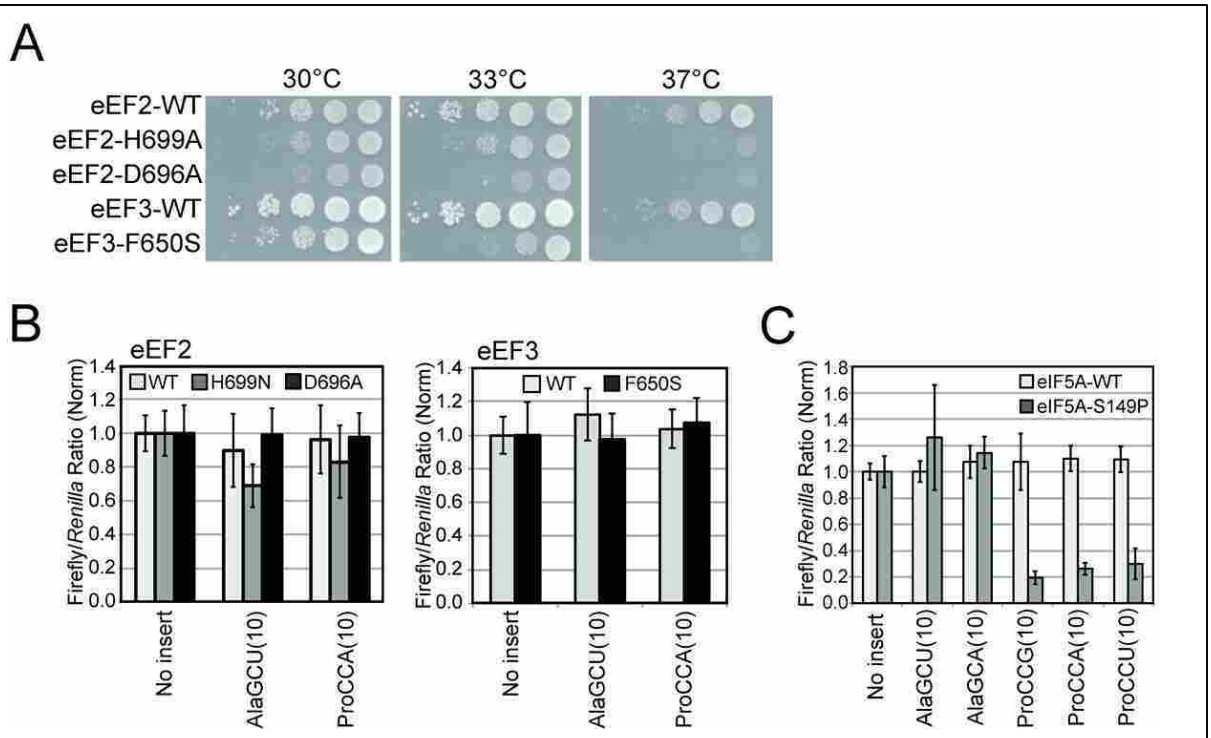


Figure 4-3. eEF2 and eEF3 Mutants do not Specifically Impair Polyproline Synthesis, Related to Figure 4-4. (A) Yeast strains expressing the indicated temperature-sensitive mutant of eEF2 or eEF3, or their respective isogenic strain expressing the wild type factor, were grown to saturation, and 4-µl volumes of serial dilutions (OD_{600} = 1.0, 0.1, 0.01, 0.001, and 0.0001) were spotted on YPD medium and incubated for 2 days at 30°C, 33°C, and 37°C. (B) Dual luciferase reporter constructs containing 10 repeats of AlaGCU or ProCCA codons were introduced into the respective wild type and eEF2 (left panel) or eEF3 (right panel) mutant strains, and luciferase activities were determined following growth at semi-permissive 33°C. Results were quantitated as described for Figure 4-1B, and error bars were calculated as propagated standard deviations (SD) for three independent transformants. (C) Dual luciferase reporters containing ten consecutive repeats of the indicated Ala or Pro codons were assayed in wild type or eIF5A-S149P mutant strains and the data was normalized to the no insert control as described in Fig. 4-1. Error bars were calculated as propagated standard deviations (SD) for three independent transformants. While these data are not definitive, as the yeast tRNAPro UGG has been reported to decode all four proline codons (see ref. (200)), they suggest that the amino acid proline rather than the tRNA likely determines the requirement for eIF5A.

expressing wild-type eIF5A or the temperature-sensitive mutant eIF5A-S149P. Transformants were grown at the semipermissive temperature of 33°C to partially inactivate eIF5A-S149P and in galactose medium to induce the *GAL1* promoter used to drive ORF expression. Protein expression was monitored by western analysis using antibodies to detect the hemagglutinin (HA)-tag incorporated at the C terminus of each ORF [215]. As shown in Figure 4-5A, expression of Ldb17 (one motif of nine consecutive Pro residues), Eap1 (two motifs of six Pro residues; one motif of three Pro residues), and Vrp1 (multiple polyPro sequences, including one motif of nine

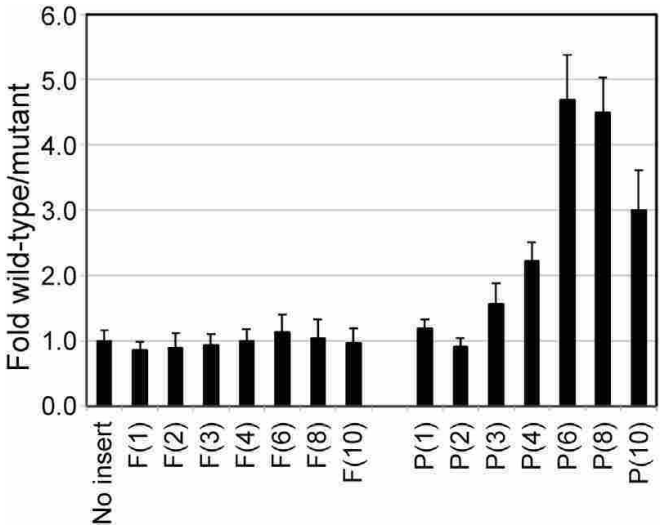


Figure 4-4. Translation of Three or More Consecutive Proline Codons Reveals eIF5A Dependency**.** Dual luciferase reporters containing 1, 2, 3, 4, 6, 8, or 10 consecutive PheUUC (F) or ProCCA (P) codons were assayed in wild type or eIF5A-S149P mutant strains and the fold difference in luciferase ratios were quantitated and normalized to the no insert control as described in Figure 4-1.

Pro residues, one motif of eight Pro residues, one motif of six Pro residues, four motifs of five Pro residues, three motifs of four Pro residues, and two motifs of three Pro residues) was dramatically reduced in the eIF5A mutant strain relative to the wild-type eIF5A strain and the loading control eIF2α (no polyPro motifs). Stable expression of Tif11 (eIF1A, no polyPro motifs) from a Yeast ORF Collection plasmid in the wild-type and mutant eIF5A strains indicates that the eIF5A-sensitive expression of the polyPro proteins is not due to impacts on the expression system (e.g., the *GAL1* promoter, growth at 33°C) (Figure 4-5A). In addition, substituting Ala in place of the nine Pro residues in the C-terminal Pro motif restored Ldb17 expression in the eIF5A-S149P mutant (Figure 4-5B), directly linking eIF5A function to the synthesis of polyPro motifs.

eIF5A Plays an Essential Role in PolyPro Peptide Synthesis

An in vitro reconstituted yeast translation assay was used to directly examine the eIF5A requirement for polyPro synthesis. As shown in Figure 4-6A, minimal translation initiation (48S) complexes encoding polyPro or polyPhe were assembled using unstructured model mRNAs to avoid the requirement for initiation factors that function in mRNA recruitment. Following ribosomal subunit joining and assembly of an 80S initiation complex with [^{35}S]Met-tRNA$_i^{Met}$ in

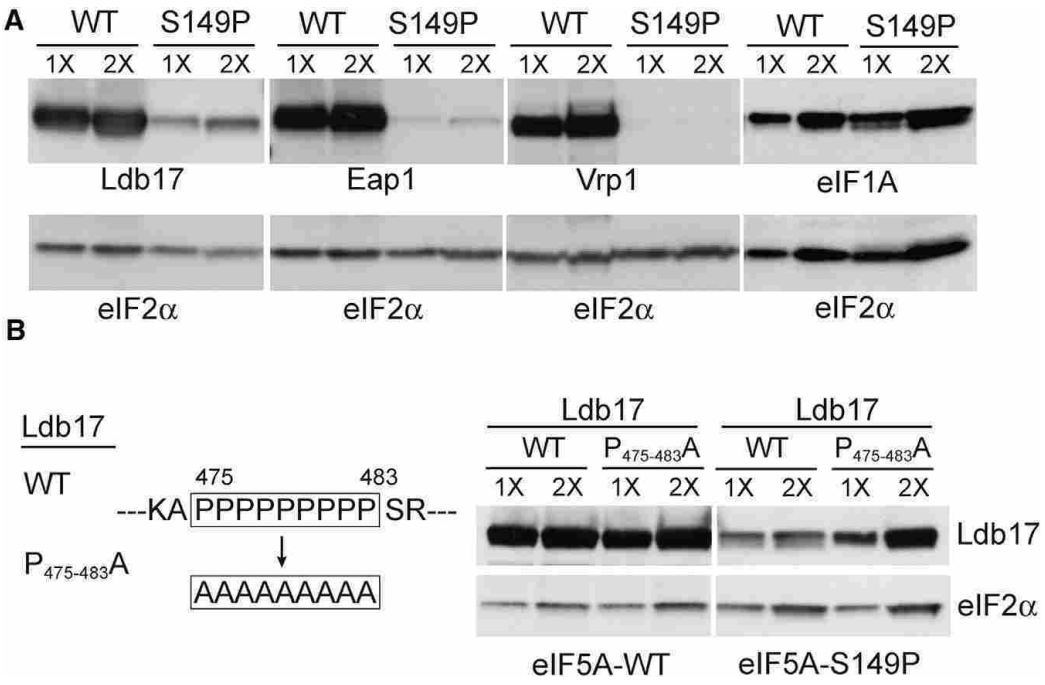


Figure 4-5. Expression of Polyproline-containing Proteins Requires eIF5A in vivo. (A) Plasmids expressing HA-tagged forms of the yeast proteins Ldb17, Eap1, Vrp1 or eIF1A under the control of the yeast *GAL1* promoter were introduced into isogenic strains expressing wild-type eIF5A or eIF5A-S149P. Cells were grown at semi-permissive 33°C in galactose medium, broken with glass beads in the presence of 10% trichloroacetic acid (TCA), and two different amounts of each extract differing by a factor of two were loaded in successive lanes and subject to immunoblot analysis using monoclonal anti-HA or polyclonal anti-yeast eIF2α antiserum. (B) The experiment in (A) was repeated using an Ldb17 construct in which Ala codons were substituted for the nine Pro codons in the polyproline motif.

the P site of the ribosome, the complex was pelleted through a sucrose cushion to remove initiation factors and unbound Met-tRNA$_i^{Met}$. Next, elongation factors and the necessary aminoacyl-tRNAs were added to the purified 80S complexes in the absence or presence of excess recombinant eIF5A. The recombinant eIF5A was prepared from *E. coli* that coexpresses the hypusine formation enzymes (Figure 4-7A), and the presence of hypusine in the recombinant eIF5A was confirmed by electrospray ionization quadrupole time-of-flight mass spectrometry (ESI-QTOF-MS) analysis (Figure 4-7B; see Experimental Procedures). Peptide synthesis was monitored by electrophoretic thin-layer chromatography (TLC) [183,216].

Synthesis of MF, MFF, and MFFF peptides progressed well in the absence and presence of eIF5A (Figures 4-6B and 4-7D), with less than 2-fold stimulation in both the observed rate

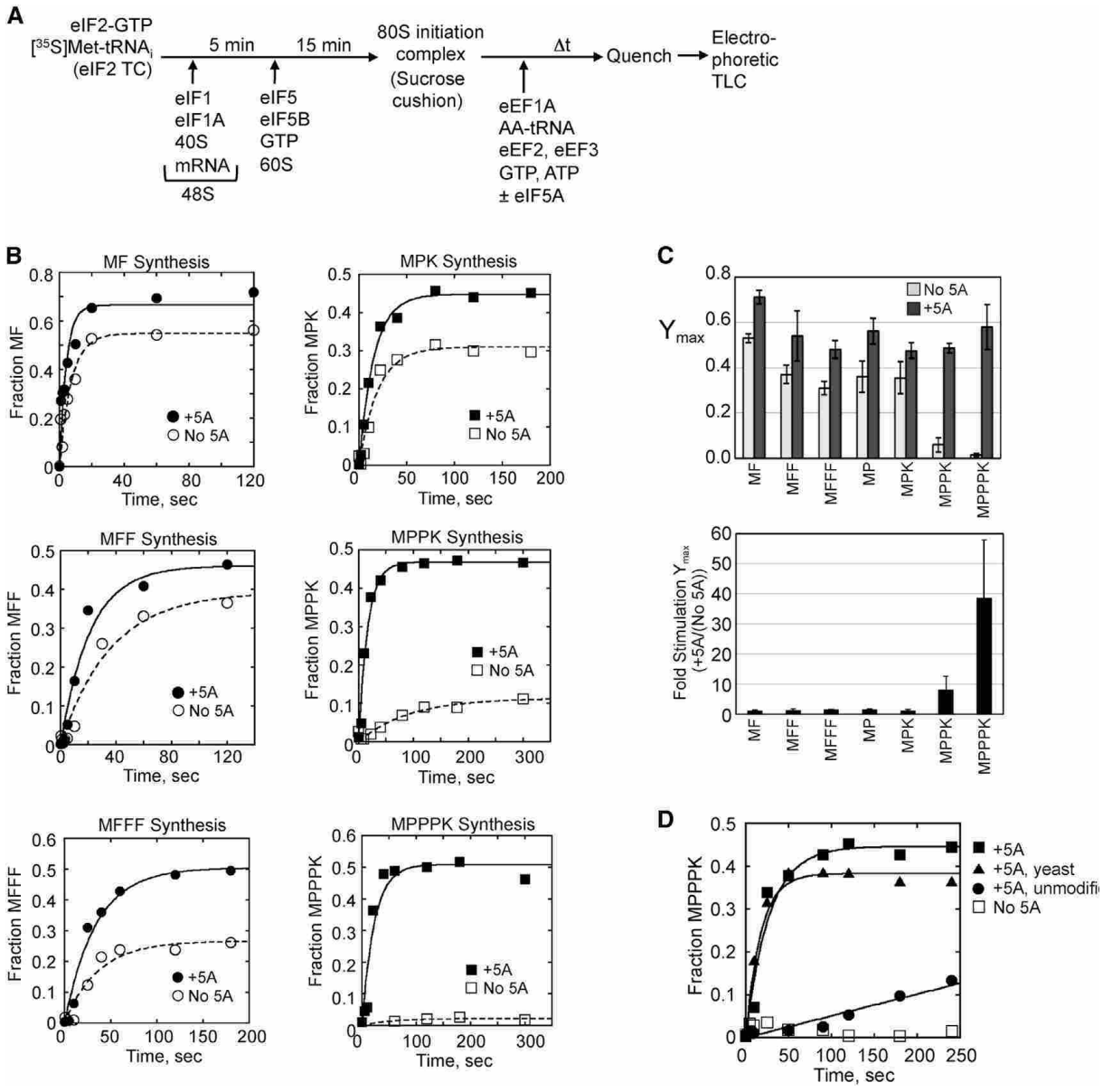


Figure 4-6. eIF5A Stimulates Synthesis of Polyproline Peptides**.** (A) Scheme for *in vitro* reconstituted translation elongation assay. (B) Fractions of MF, MFF, MFFF (left column) or MPK, MPPK, and MPPPK (right) synthesis in elongation assays (Fig. 4-7D and 4E) performed in the absence (open symbols) or presence of eIF5A (closed symbols) were plotted and fit to a single exponential equation. (C) Summary of maximum fractions of peptide synthesis (Y_{max}, top) and fold stimulation of Y_{max} by adding eIF5A (bottom) calculated from the data in panel B. Error bars are (upper) standard deviations from at least three independent experiments and (lower) calculated propagated errors. (D) Effect of eIF5A hypusine modification on peptide synthesis. Fraction of MPPPK synthesis (Fig. 4-7F) in reactions lacking eIF5A, containing unmodified eIF5A (no hypusine), or containing hypusinated eIF5A prepared from *E. coli* (+5A, see Experimental Procedures) or purified from yeast (+5A, yeast) was plotted and fit to a single exponential equation.

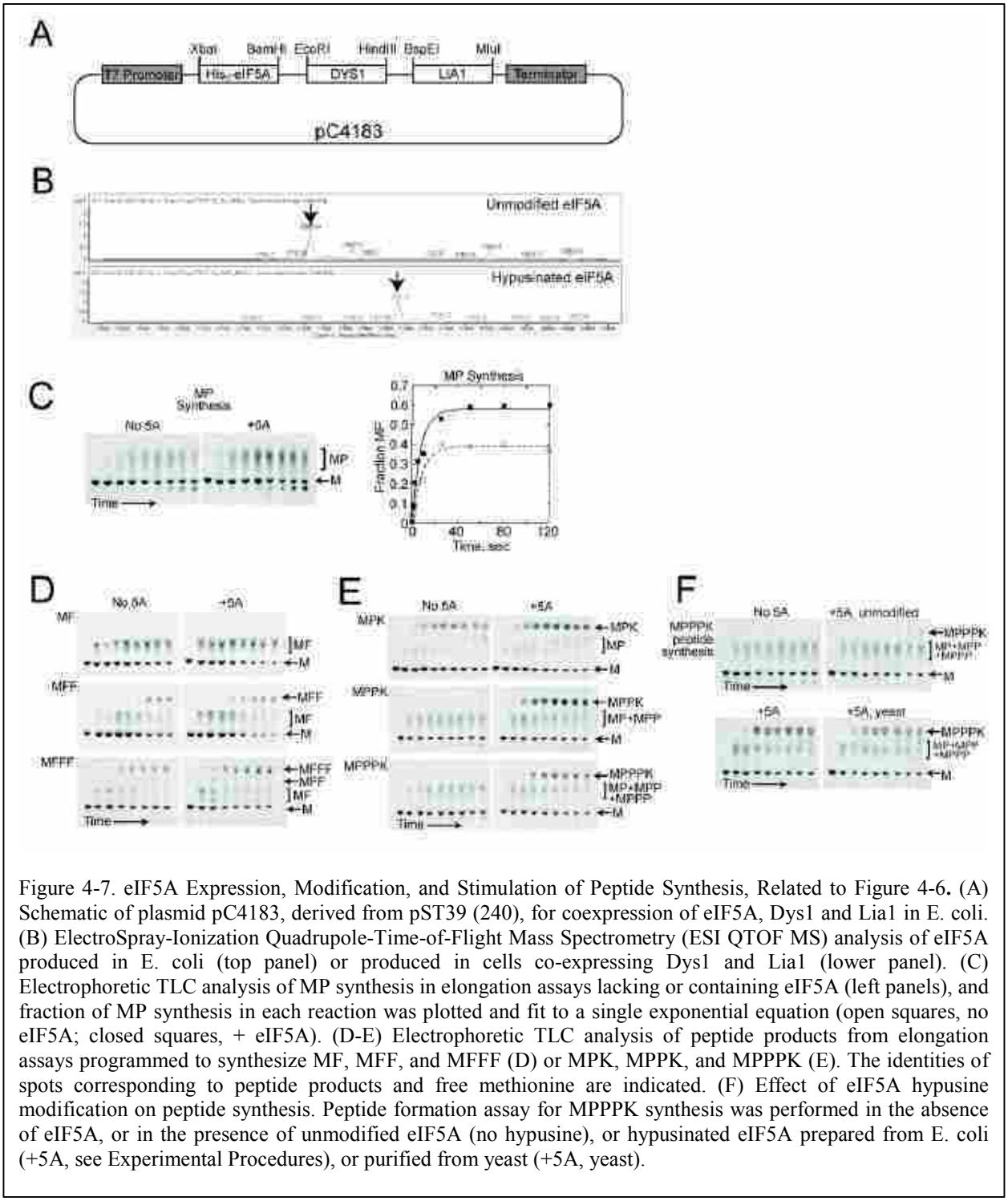


Figure 4-7. eIF5A Expression, Modification, and Stimulation of Peptide Synthesis, Related to Figure 4-6**.** (A) Schematic of plasmid pC4183, derived from pST39 (240), for coexpression of eIF5A, Dys1 and Lia1 in E. coli. (B) ElectroSpray-Ionization Quadrupole-Time-of-Flight Mass Spectrometry (ESI QTOF MS) analysis of eIF5A produced in E. coli (top panel) or produced in cells co-expressing Dys1 and Lia1 (lower panel). (C) Electrophoretic TLC analysis of MP synthesis in elongation assays lacking or containing eIF5A (left panels), and fraction of MP synthesis in each reaction was plotted and fit to a single exponential equation (open squares, no eIF5A; closed squares, + eIF5A). (D-E) Electrophoretic TLC analysis of peptide products from elongation assays programmed to synthesize MF, MFF, and MFFF (D) or MPK, MPPK, and MPPPK (E). The identities of spots corresponding to peptide products and free methionine are indicated. (F) Effect of eIF5A hypusine modification on peptide synthesis. Peptide formation assay for MPPPK synthesis was performed in the absence of eIF5A, or in the presence of unmodified eIF5A (no hypusine), or hypusinated eIF5A prepared from E. coli (+5A, see Experimental Procedures), or purified from yeast (+5A, yeast).

constant and the fraction of maximal yield (Y_{max}) for formation of the peptides in the presence of eIF5A (Figure 4-6C). These results are consistent with the previously reported ~2-fold stimulation of MFF synthesis upon adding eIF5A to the reconstituted system [204]. In preliminary

experiments, the Pro-containing peptides MP, MPP, and MPPP failed to resolve as discrete spots on the TLC (see Figures 4-7C, 7E, and 7F). However, incorporation of a Lys residue at the C terminus of the Pro peptides enabled resolution of the peptides by TLC and facilitated their quantitation. The fraction of maximum peptide yield for MPK peptide synthesis was stimulated ~1.3-fold by adding eIF5A (Y_{max} = 0.36 in the absence of eIF5A and 0.48 in the presence of eIF5A) (Figures 4-6B, 6C, and 4-7E, top panels). Thus, the presence of a single Pro residue conferred a modest eIF5A dependency for peptide synthesis. In contrast, synthesis of the MPPK peptide containing two Pros was significantly impaired in the absence of eIF5A (Y_{max} = 0.06 ± 0.03). An 8.3-fold stimulation of Y_{max} was observed upon adding eIF5A (Y_{max} = 0.49 ± 0.02) (Figures 4-6B, 6C, and 4-7E). The large difference in reaction endpoints for the Pro-containing peptides in the presence versus the absence of eIF5A suggests that competing reactions are likely occurring (e.g., peptidyl-tRNA drop-off). Since the observed rates reflect both peptide bond formation and these competing reactions, we have limited our analysis to the reaction endpoint differences and not to the observed rates. Remarkably, no detectable formation of the MPPPK peptide containing three consecutive Pro residues occurred in the absence of eIF5A during the time course of experiments. The addition of eIF5A efficiently restored MPPPK synthesis, stimulating the Y_{max} at least 39-fold (Y_{max} = 0.58 ± 0.1) (Figure 4-6C). Thus, consistent with the results of the in vivo assays, eIF5A is required for synthesis of peptides containing consecutive Pro residues.

To assess the importance of the hypusine modification on eIF5A, MPPPK synthesis was analyzed using different forms of the factor. As shown in Figures 4-6D and 4-7F, no MPPPK synthesis was detected in the absence of eIF5A and very little synthesis was detected in assays that included unmodified eIF5A prepared from *E. coli* (see Experimental Procedures). In

contrast, hypusine-modified eIF5A, prepared either from yeast or from *E. coli* coexpressing the hypusine modification enzymes (see Experimental Procedures), readily stimulated MPPPK synthesis (Y_{max} = 0.38 ± 0.02 for yeast eIF5A; Y_{max} = 0.45 ± 0.02 for recombinant eIF5A). Thus, the hypusine modification of eIF5A is necessary for efficient polyPro synthesis in vitro.

eIF5A Prevents Ribosome Stalling on Consecutive Pro Codons

The in vitro peptide synthesis assays revealed defects in synthesizing MPPK and MPPPK, suggesting that ribosomes stall when translating polyPro sequences. To directly detect ribosome stalling, toeprinting assays were performed to determine the position of the ribosome on mRNAs encoding MPPPPP and MFFFFF in the reconstituted in vitro translation system described above. A ^{32}P-labeled primer was annealed to the 3′ end of the mRNA and extended by reverse transcription. The ribosome blocks reverse transcriptase 15–16 nucleotides downstream of the first nucleotide of the P site codon. As shown in Figure 4-8 (lanes 3 and 8), toeprinting confirms that the AUG start codon is positioned in the P site in 80S initiation complexes. When elongation factors and Phe-tRNA^{Phe} were added to the ribosomal complexes translating the MFFFFF mRNA, this toeprint was diminished and a different toeprint was observed, corresponding to the final Phe codon in the P site (Figure 4-8, lanes 4 and 5). The position of this latter toeprint is consistent with the ribosome translating to the end of the ORF and arresting with the stop codon in the A site; note that no release factors were included in the translation reaction. Consistent with the robust synthesis of the MFFF peptide in the presence and absence of eIF5A (Figures 4-6B, 6C, and 4-7D), production of the MFFFFF complex in the toeprinting assays was unaffected by the presence or absence of eIF5A (Figure 4-8, lanes 4 and 5).

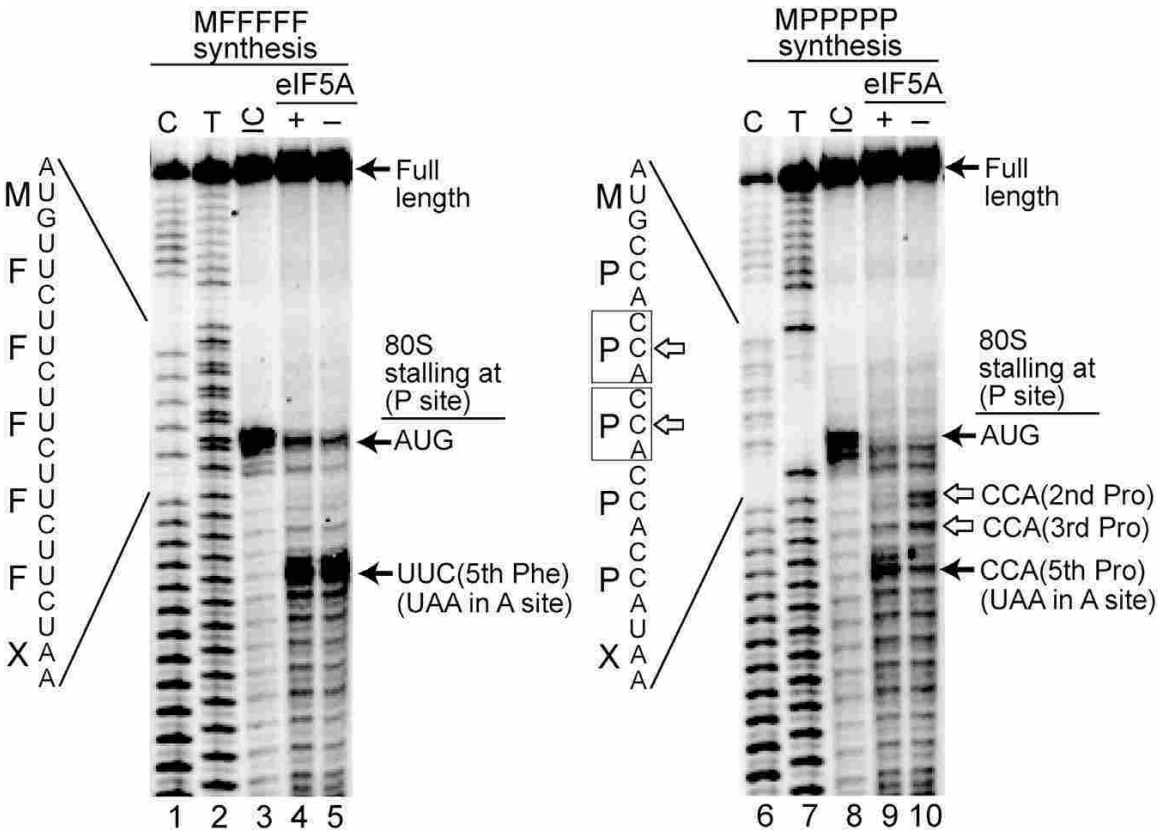


Figure 4-8. eIF5A Prevents Ribosome Stalling on Consecutive Proline Codons**.** Reconstituted peptide synthesis assays were performed in the absence or presence of eIF5A using mRNAs encoding the peptides MFFFFF (left panel) or MPPPPP (right panel). The position of the 80S ribosome was determined by reverse transcription of the mRNA template using a [^{32}P]-labeled primer, and C and T sequencing reactions were run alongside. Reactions lacking elongation factors were performed to identify 80S initiation complexes (IC) on the AUG codon (lanes 3 and 8). The identity of the 80S toeprint signals is indicated on the right; and the sequences of the mRNA and the corresponding amino acids are shown on the left with the sites of ribosome stalls at the 2nd and 3rd proline codons boxed.

Toeprinting analysis of elongation complexes revealed ribosomal stalling on the MPPPPP mRNA. In reactions lacking eIF5A (Figure 4-8, lane 10), toeprints were observed corresponding to the ribosome stalling with the second or third Pro codon in the P site. Importantly, addition of eIF5A diminished the abundance of these stalled complexes and increased the yield of ribosome complexes with the final Pro codon in the P site and the stop codon in the A site (Figure 4-8, lane 9). These results indicate that, in the absence of eIF5A, the P site tRNA is linked to an MPP or MPPP peptide and that a Pro codon, and presumably Pro-tRNA, is in the A site. The positions of the stalled ribosome complexes are consistent with the in vivo assays (Figure 4-4) that showed

that at least three consecutive Pro residues are necessary to observe an eIF5A dependency and indicate that peptide bond formation between diPro-tRNA and Pro-tRNA is particularly dependent on eIF5A. Moreover, these results are consistent with reports of *E. coli* ribosomes stalling on triPro motifs with the peptidyl-tRNA linked to consecutive C-terminal Pro residues and the third Pro codon in the A site of the ribosome [98,172].

eIF5A Binds Near the E and P sites of the 80S Ribosome

Directed hydroxyl radical mapping was used to identify the binding site for eIF5A on the yeast 80S ribosome. It is notable that EFP, which contains an extra C-terminal domain not found in eIF5A, binds between the P and E sites and contacts both the large and small subunits of the bacterial 70S ribosome [189]. A Cys-less derivative of yeast eIF5A was generated by mutating the native C23 and C39 residues to Ala and Thr, respectively. Yeast expressing the eIF5A-C23A,C39T mutant (eIF5A-ΔC) as the sole source of eIF5A grew as well as cells expressing the wild-type protein (Figure 4-9A), indicating that the mutations do not affect eIF5A function. Next, single Cys residues were introduced at four surface-exposed sites generating eIF5A-ΔC-S36C, eIF5A-ΔC-K48C, eIF5A-ΔC-M105C, and eIF5A-ΔC-T126C (Figure 4-10A). All four mutant proteins supported yeast cell growth at wild-type rates (Figure 4-9A), suggesting that the mutations did not interfere with either the essential hypusine modification of eIF5A or the function of the protein on the ribosome. The eIF5A-ΔC and four single Cys mutant proteins were purified from yeast, derivatized with Fe(II)-1-(*p*-bromoacetamidobenzyl)-EDTA (Fe[II]-BABE), which links the ferrous iron to the Cys residue, and then added to assembled 80S complexes containing 5′ end-labeled Met-[^{32}P]tRNA$_i^{Met}$ in the P site (see Figure 4-10B).

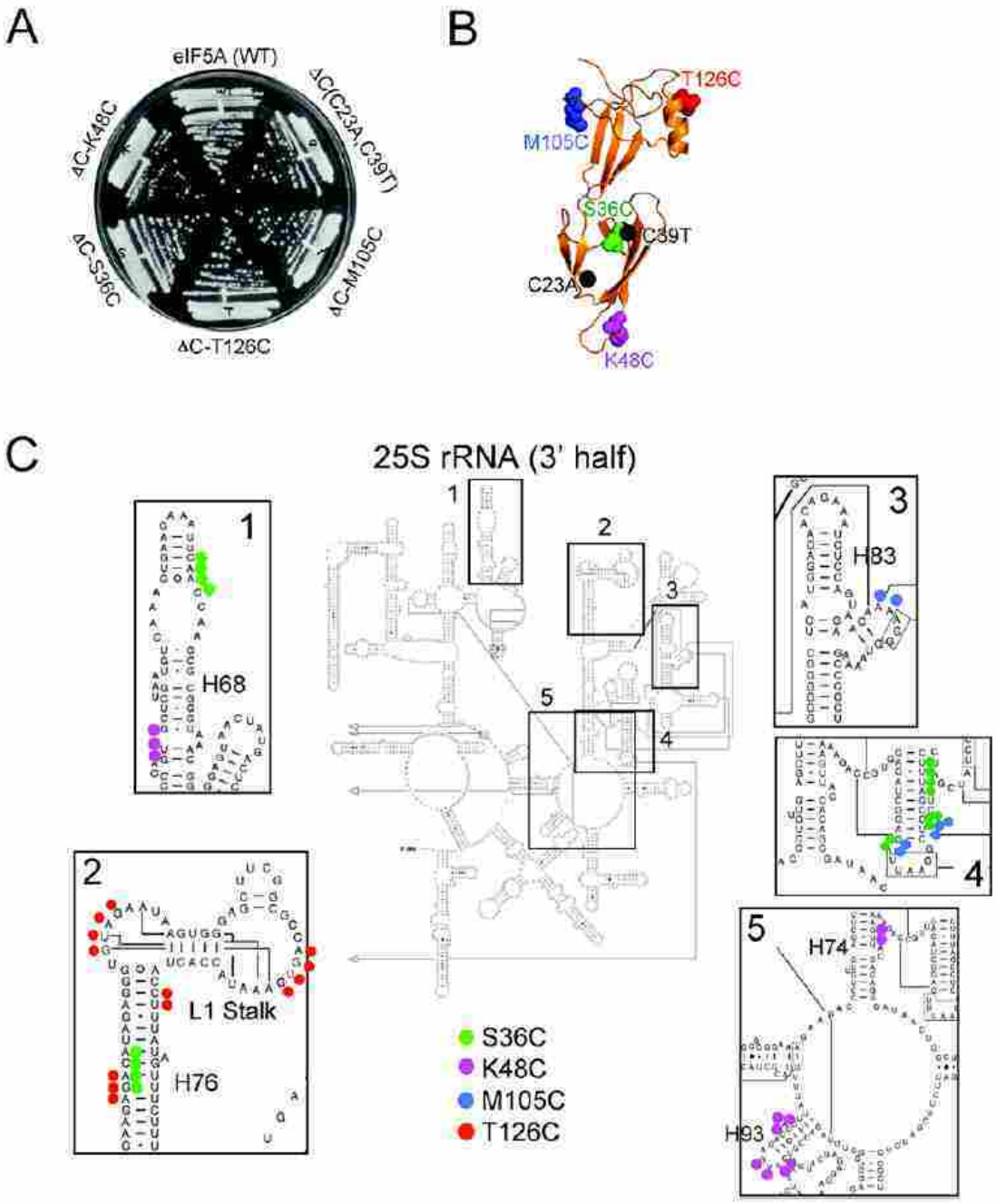


Figure 4-9. Analysis of eIF5A Cys Mutants and Sites of 25S rRNA Cleavage, Related to Figure 4-10**.** (A) Yeast growth analysis of eIF5A Cys mutants. Derivatives of yeast strain J697 expressing wild type (WT) eIF5A or the indicated eIF5A mutants were streaked on YPD medium and incubated at 30°C for 3 days. (B) Ribbon representation of yeast eIF5A (PDB 3ER0) showing the positions of the C23A and C39T mutations (black dots) that removed the native Cys residues in eIF5A, and the sites (Spheres representation) of Cys mutations for tethering Fe(II): Ser36 (green), Lys48 (magenta), Met105 (blue) and Thr126 (red). (C) Sites of directed hydroxyl radical cleavage by Fe(II)-BABE derivatives of eIF5A are mapped on the secondary structure of yeast 25S rRNA. Sites of cleavage are color-coded to match the site of Fe(II)-BABE modification on eIF5A (see panel B).

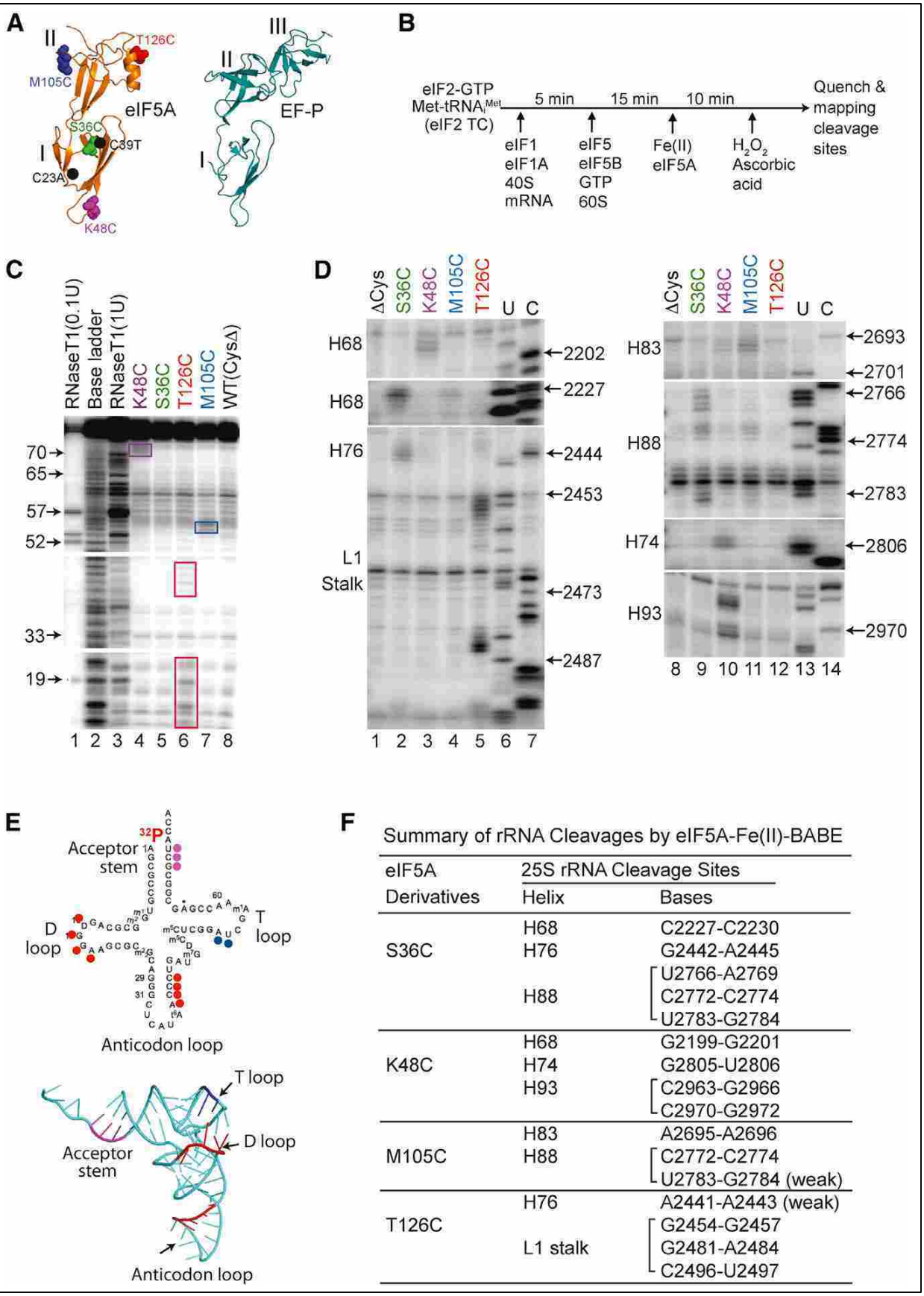

A
II
T126C
M105C
eIF5A
S36C
C39T
I
C23A
K48C
III
II
EF-P
I
B
eIF2-GTP
Met-tRNA$_i^{Met}$
(eIF2 TC)
5 min
15 min
10 min
Quench &
mapping
cleavage
sites
eIF1
eIF1A
40S
mRNA
eIF5
eIF5B
GTP
60S
Fe(II)
eIF5A
H_2O_2
Ascorbic
acid
C
RNaseT1(0.1U)
Base ladder
RNaseT1(1U)
K48C
S36C
T126C
M105C
WT(CysΔ)
70
65
57
52
33
19
1 2 3 4 5 6 7 8
D
ΔCys
S36C
K48C
M105C
T126C
U C
H68
2202
H68
2227
H76
2444
2453
L1
Stalk
2473
2487
1 2 3 4 5 6 7
H83
2693
2701
H88
2766
2774
2783
H74
2806
H93
2970
8 9 10 11 12 13 14
E
^{32}P
Acceptor
stem
D
loop
T
loop
Anticodon loop
T loop
D loop
Acceptor
stem
Anticodon loop
F
Summary of rRNA Cleavages by eIF5A-Fe(II)-BABE
eIF5A Derivatives
25S rRNA Cleavage Sites
Helix
Bases
S36C
H68
C2227-C2230
H76
G2442-A2445
H88
U2766-A2769
C2772-C2774
U2783-G2784
K48C
H68
G2199-G2201
H74
G2805-U2806
H93
C2963-G2966
C2970-G2972
M105C
H83
A2695-A2696
H88
C2772-C2774
U2783-G2784 (weak)
T126C
H76
A2441-A2443 (weak)
L1 stalk
G2454-G2457
G2481-A2484
C2496-U2497

Figure 4-10. Directed Hydroxyl Radical Probing of eIF5A Binding to 80S Ribosomal Complexes. (A) Ribbon representation of *T. thermophilus* EFP (right panel, PDB ID code 3HUW (189)) and yeast eIF5A (left panel, PDB ID code 3ER0) showing the protein domains (Roman numerals), the positions of the C23A and C39T mutations (black dots) that removed the native Cys residues in eIF5A, and the sites (Spheres representation) of Cys mutations for tethering Fe(II): Ser36 (green), Lys48 (magenta), Met105 (blue) and Thr126 (red). (B) Scheme for directed hydroxyl radical cleavage by Fe(II)-BABE modified forms of eIF5A in 80S complexes. (C) Directed hydroxyl radical cleavage of Met-[^{32}P]tRNA$_i^{Met}$ by Fe(II)-BABE-derivatized eIF5A in 80S complexes. Cleavage products were resolved on 10% (w/v) denatured polyacrylamide gels, and cleavage sites on [^{32}P]tRNA$_i^{Met}$ were determined by comparing them to samples containing eIF5A-ΔC [WT(CysΔ), lane 8]. The tRNA ladders were prepared by digesting Met-[^{32}P]tRNA$_i^{Met}$ with RNase T1 (cleaves 3′ of G residue) or by base cleavage (lane 2). The tRNA residue numbers are shown at the left, and cleavage fragments are boxed. (D) Primer extension analysis of 25S rRNA cleavage fragments produced by Fe(II)-tethered to the indicated positions in eIF5A. U and C: 25S rRNA sequencing reactions using reverse transcriptase and dideoxynucleotides ddATP and ddGTP, respectively. 25S rRNA helices and the position of the L1 stalk are indicated on the left. (E) Sites of eIF5A-Fe(II)-BABE cleavages are shown on the secondary (left) and three-dimensional (PDB ID code 1YFG (218)) structures of tRNA$_i^{Met}$. Cleavage sites are color-coded according to the site where Fe(II) was tethered on eIF5A (see A). (F) Summary of 25S rRNA cleavages by eIF5A-Fe(II)-BABE derivatives.

Cleavage of Met-[^{32}P]tRNA$_i^{Met}$ by hydroxyl radicals generated in the vicinity of the ferrous iron by the Fenton reaction was monitored by 10% denaturing gel electrophoresis. Compared to reactions employing eIF5A-ΔC, which, due to the absence of Cys residues, is not modified by Fe(II)-BABE, hydroxyl radicals generated by Fe(II)-BABE tethered to eIF5A-ΔC-K48C yielded cleavages in the tRNA$_i^{Met}$ acceptor stem bases G70–U72 (Figure 4-10C [compare lanes 4 and 8] and Figure 4-10E). As this K48C site of Fe(II)-BABE modification is located only three residues from the site of hypusine modification (K51) on eIF5A, these results place the hypusine side chain in the vicinity of the amino acid attached to the 3′-CCA end of the P site tRNA. In contrast, hydroxyl radicals generated using Fe(II)-modified eIF5A-ΔC-M105C cleaved bases A54 and U55 in the T stem region of Met-tRNA$_i^{Met}$ (Figure 4-10C, lane 7, and Figure 4-10E), while hydroxyl radicals generated using Fe(II)-modified eIF5A-ΔC-T126C cleaved two different regions in tRNA$_i^{Met}$: bases U16–A20 in the D stem-loop region and bases A38–C41 in the anticodon stem region (Figure 4-10C [lane 6] and Figure 4-10E). No noticeable tRNA$_i^{Met}$ cleavages were observed using the iron-modified form of eIF5A-ΔC-S36C (Figure 4-10C, lane 5). From these data, we conclude that the eIF5A binds alongside the P site tRNA on the 80S ribosome, a position similar to the EFP binding site on the bacterial 70S ribosome [189], with the

hypusine residue at the top of eIF5A near the aminoacyl end of the tRNA and the eIF5A domain II residue T126 near the anticodon stem of the tRNA.

In order to further define the eIF5A-binding site on the ribosome, hydroxyl radicals were generated in 80S complexes containing Fe(II)-BABE-modified forms of eIF5A, and ribosomal RNA (rRNA) cleavages were analyzed by primer extension using ^{32}P-labeled primers. Whereas Fe(II)-modified eIF5A-ΔC-S36C did not generate cleavages in $tRNA_i^{Met}$, this eIF5A derivative yielded cleavages in helices H68, H76, and H88 of 25S rRNA that were not seen with the eIF5A-ΔC control (Figure 4-10D, compare lanes 1 and 2 and lanes 8 and 9; see summaries of cleavage sites in Figures 4-10F and 4-9C). These cleavage sites are clustered around the E site tRNA binding region of the 60S subunit (Figure 4-11A, green color). Primer extension analysis of rRNA cleavages generated using Fe(II)-BABE-tethered eIF5A-ΔC-K48C revealed enhanced cleavages in helices H68, H74, and H93 (Figures 4-10D [lanes 3 and 1], 10F, and 4-9C). Noticeably, these cleavage sites map near the PTC of the 60S subunit (Figure 4-11A, magenta color), consistent with the idea that the hypusine residue (modified K51 side chain) is close to the PTC active site of the ribosome.

The rRNA cleavages generated using iron-modified eIF5A-ΔC-M105C mapped to helices H83 and H88 (Figures 4-10D [lanes 4 and 11] and 6F). Interestingly, these helix H88 cleavages partially overlap with the cleavages generated by eIF5A-ΔC-S36C (Figures 4-10D, 10F, and 4-9C), consistent with the presentation of these two residues on the same surface of eIF5A. Finally, the Fe(II)-BABE-modified form of eIF5A-ΔC-T126C generated cleavages near the L1 stalk region of the 60S subunit (Figures 4-10D [lanes 5 and 12], 10F, and 4-9C).

It is interesting to note that all of the eIF5A-generated cleavages map to the 3′ half of 25S rRNA (Figure 4-9C), and there are no detectable cleavage sites in 18S rRNA (data not shown).

Thus, in contrast to EFP, which contacts both the small and large ribosomal subunits [189], eIF5A appears principally to contact the 60S subunit when binding to 80S ribosomal complexes. To generate a model of eIF5A binding to the ribosome, the hydroxyl radical cleavage data were used to orient yeast eIF5A (Protein Data Bank [PDB] ID code 3ER0) on the yeast 60S subunit (PDB ID code 3O58; [217] with $tRNA_i^{Met}$ (PDB ID code 1YFG; [218]) bound in the P site. To initiate the docking process, the eIF5A residue K48 was positioned near the PTC and centered among its 25S rRNA and $tRNA_i^{Met}$ acceptor stem cleavage sites. Next, the residue T126 was positioned between the ribosomal L1 stalk and the D loop and anticodon stem of the P site $tRNA_i^{Met}$. After fixing the K48 and T126 locations in the model, the eIF5A structure was rotated so that residues S36 and M105 were oriented toward their respective cleavage sites (Figure 4-11A). In the 60S-$tRNA_i^{Met}$-eIF5A complex model, eIF5A binds between the P site $tRNA_i^{Met}$ and the E site such that the N-terminal domain of eIF5A (residues 1–82) and the C-terminal domain (residues 87–157) are close to the acceptor stem and D loop region of $tRNA_i^{Met}$, respectively (Figure 4-11B). This eIF5A binding position, which is based on the hydroxyl radical mapping studies presented here, overlaps with the position of EFP domains I and II as found in the cocrystal structure of EFP bound to the bacterial 70S ribosome (Figure 4-11C, [189]). It is notable that domain III of EFP, which is missing from eIF5A (Figure 4-10A), contacts the small ribosomal subunit adjacent to the anticodon stem region of fMet-$tRNA_i^{fMet}$, which is bound in the P site [189]. Thus, despite their structural and functional similarities, the C-terminal truncation of eIF5A relative to EFP apparently limits eIF5A ribosomal contacts to the large subunit and may confer a functional distinction between the two factors.

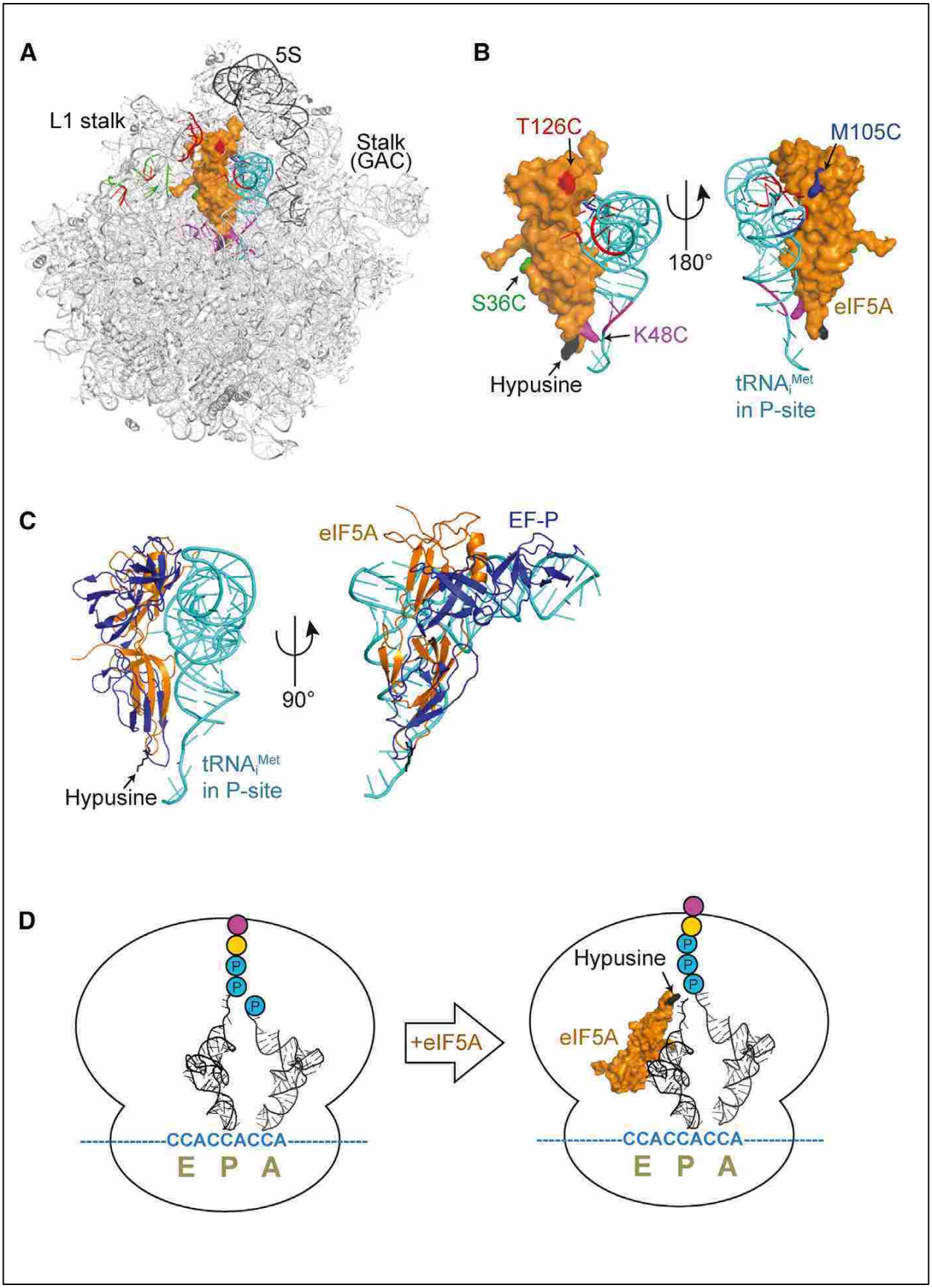

A
5S
L1 stalk
Stalk
(GAC)
B
T126C
M105C
180°
S36C
K48C
eIF5A
Hypusine
tRNAiMet
in P-site
C
eIF5A
EF-P
90°
Hypusine
tRNAiMet
in P-site
D
P
P
P
CCACCACCA
E P A
+eIF5A
Hypusine
eIF5A
CCACCACCA
E P A

Figure 4-11. Models of 60S–Met-tRNAiMet–eIF5A Complex and eIF5A Stimulating Polyproline Synthesis**.** (A) Docking model of a surface representation of yeast eIF5A (orange, PDB ID code 3ER0) and ribbons representation of $tRNA_i^{Met}$ (cyan, PDB ID code 1YFG (218)) on the ribbons structure of the yeast 60S ribosome (PDB ID code 3O58 (217) as viewed from the subunit interface. The position of $tRNA_i^{Met}$ was modeled by alignment with P-site tRNA on the bacterial ribosome (PDB ID code 2J00 (see ref. 200)), and eIF5A was docked on the 60S subunit according to the cleavage data for Met-$tRNA_i^{Met}$ and 25S rRNA. Cleavage sites in 25S rRNA and $tRNA_i^{Met}$ are color-coded according to the sites of Fe(II) attachment on eIF5A (see Figure 4-10A). Positions of L1 stalk, 5S rRNA (black), and GTPase activating center (GAC) Stalk on the 60S subunit are indicated. (B) Magnified view of docked eIF5A and P-site $tRNA_i^{Met}$ structure as shown in panel A (left) and rotated 180° (right). Lys51, the site of hypusine modification, is colored black. (C) Magnified view of docked eIF5A and P-site $tRNA_i^{Met}$ (from A) overlaid on the structure of EFP (blue) from the EFP–70S structure (PDB ID code 3HUW (189)) oriented as shown in panel A (left) and rotated 90° (right). (D) Model of ribosome stalled on polyproline sequence with di-proline attached to the P-site tRNA and Pro-$tRNA^{Pro}$ in the A site (left). (Right) Binding of eIF5A near the E site places the hypusine side chain (Lys51, black) adjacent to the peptidyl-tRNA in the peptidyl transferase center (PTC) of the ribosome where it can help promote peptide bond formation with the amino acid attached to the A-site tRNA (right).

Discussion

In addition to the universally conserved translation factors, eEF1A/EFTu and eEF2/EFG, three other factors have been implicated in translation elongation. However, these latter factors are not universally conserved. The factor eEF3 is proposed to coordinate E site tRNA release with eEF1A-aminoacyl-tRNA binding to the A site of the ribosome in some fungi including *S. cerevisiae* [219,220]. Bacterial EF4 (LepA) is proposed to maintain rapid protein synthesis under stress conditions such as high ionic strength and low temperature [221], and SelB/eEFsec is an EFTu ortholog required for the delivery of selenocysteinyl-tRNA to the ribosome [222]. In this chapter we demonstrate that eIF5A is required for the translation of polyPro sequences. As our results concur with the findings of the recent studies on EFP, the bacterial ortholog of eIF5A [184,185], we conclude that eIF5A/EFP is the third universally conserved translation elongation factor.

Our data demonstrating that eIF5A promotes the translation of polyPro sequences are consistent with the recent reports on EFP [184,185]. Partial inactivation of eIF5A-S149P in yeast (Figure 4-1, Figure 4-4 and Figure 4-5), like deletion of the *efp* gene in *E. coli* [185], impaired

expression of reporters or native proteins containing polyPro sequences in vivo. Moreover, peptide synthesis assays demonstrated that eIF5A (Figure 4-6) and EFP are critical for the in vitro synthesis of polyPro peptides [184]. Finally, toeprinting analyses revealed that, in the absence of eIF5A, translating ribosomes stall on polyPro motifs with the second or third Pro codon in the P site (Figure 4-8). Thus, diPro or triPro will be attached to the 3′-CCA end of the peptidyl-tRNA in the P site, and Pro-tRNA will be bound in the A site. These results suggest that eIF5A and EFP are required to promote synthesis of the Pro-Pro peptide bonds needed to convert diPro to triPro and higher-order polyPro sequences.

In addition to amino acid sequence and structural similarities, both eIF5A and EFP are posttranslationally modified. As described earlier, the ε-amino group of a conserved Lys residue in eIF5A from all archaea and eukaryotes is modified by deoxyhypusine synthase and deoxyhypusine hydroxylase to generate hypusine (reviewed in [223]). Similarly, the *E. coli* gene products YjeA, YjeK, and YfcM attach a β-lysine residue to the ε-amino group of Lys34 in *E. coli* EFP and then hydroxylate the side chain [190,191,194,210]. Both the hypusine modification of eIF5A and the β-lysine modification of EFP are required for these factors to stimulate polyPro synthesis (Figure 4-6D, and [184,185]). Loss of β-lysylation in *E. coli* impairs translation of a polyPro motif in *CadC*, the transcriptional activator of the *CadBA* operon, whereas loss of β-lysylation in *Salmonella enterica* alters the expression of a variety of cellular proteins, impairs virulence in mice, and alters resistance to antibiotics [190], presumably due to impaired translation of polyPro sequences. In a similar manner, perhaps the assorted genes identified in suppressor and synthetic enhancement screens with eIF5A mutants in yeast [224,225], as well as the connections between eIF5A, hypusine, cancer, and tumorigenesis in humans and other mammals [206,207], reflect altered expression of proteins containing polyPro motifs.

eIF5A and EFP were originally identified based on their ability to stimulate Met-puromycin synthesis [186,201–203]. Whereas eIF5A was thought to function as a translation initiation factor, it is notable that puromycin, an aminoacyl analog that reacts well with most peptidyl-tRNA substrates, reacts poorly with fMet-tRNA^{fMet} and with peptidyl-tRNA substrates with a C-terminal Pro [118]. In contrast, these latter substrates react well with authentic aminoacyl-tRNAs [118]. The poor reactivity with puromycin has been attributed to poor substrate positioning in the active site of the ribosome [118,183]. Thus, the Met-puromycin synthesis assay is likely not a good mimic of first peptide bond synthesis. Consistent with this notion, the k_{obs} (data not shown) and the Y_{max} for MP or MF synthesis were only modestly affected by adding eIF5A (Figures 4-6 and 4-7C); similarly, recent studies show that EFP does not stimulate dipeptide formation [184,226].

Consistent with a function in translation elongation, inactivation of eIF5A (Figure 4-2B and[204]) or of EFP, or its β-lysine modification, mimics the effects of elongation inhibitors and causes polysome retention [226]. It is unclear at present whether the polysome retention upon inactivation of eIF5A observed in Figure 4-2B reflects impaired translation elongation on the majority of cellular mRNAs or if it could be due to impaired translation of just the mRNAs containing polyPro motifs (549 of 5,889 ORFs in *S. cerevisiae* contain a polyPro tract consisting of three or more Pro residues). This prevalence of polyPro motifs in yeast (95 proteins with motifs containing 4 or more consecutive Pro residues) is consistent with eIF5A being essential in yeast, whereas the *efp* gene can be deleted in *E. coli* in which only 9 out of ~4,000 proteins contain motifs of four or more Pro residues [184]. Taken together, the puromycin, dipeptide, and polysome analyses indicate that eIF5A and EFP do not substantially stimulate first peptide bond formation, consistent with the notion that the primary function of these factors is to promote peptide bond formation, especially for poor substrates like polyPro.

Previous studies revealed that peptidyl-Pro-tRNA in the ribosomal P site reacts poorly with puromycin [117] and that Pro is an inefficient A site substrate for peptide bond formation [165]. Our data, as well as the recent studies with EFP [172,184,185], demonstrate that combining peptidyl-Pro-tRNA in the P site with Pro-tRNA in the A site dramatically impairs protein synthesis and establishes a dependency on eIF5A/EFP. At present, it is not clear why polyPro is such a poor substrate for protein synthesis; however, it may reflect the imino acid nature of Pro, the geometrical or steric constraints of a cyclic side chain, or the unique ability of Pro to readily sample both *cis* and *trans* conformations of peptide bonds [227]. Perhaps insertion of the extended hypusine (or β-lysine) side chain into the PTC (Figure 4-11D) stabilizes the proper conformation of the PTC or restricts the conformation of Pro in the P site, enabling a favorable geometry for peptide bond formation with the A site amino acid. While the data reported here and the recent studies on EFP establish that translation of homopolyPro motifs requires eIF5A/EFP, additional studies, including genome-wide ribosomal profiling [228] of wild-type and eIF5A mutant cells, will be needed to define the spectrum of amino acids and motifs that rely on eIF5A for their efficient translation.

Experimental Procedures.

Dual-Luciferase Assay

Dual-luciferase reporter constructs were obtained from Elizabeth Grayhack [214]. Whole cell extracts from yeast transformants were assayed for firefly and *Renilla* luciferase activity.

Peptide Formation and Toeprinting Assays

Initiation complexes were prepared as described previously [229] using [^{35}S]Met-tRNA$_i^{Met}$ and purified translation initiation factors. Limited amounts of initiation complexes were mixed with purified eEF1A, eEF2, eEF3, Phe (or Pro)-tRNA, guanosine triphosphate (GTP), and ATP in the presence or absence of eIF5A in buffer containing 30 mM HEPES-KOH (pH 7.4), 100 mM potassium acetate, 1 mM magnesium acetate, 2 mM magnesium chloride, 1 mM spermidine, and 2 mM dithiothreitol (DTT). All reactions were performed at 26°C. Peptide formation was monitored by electrophoretic TLC (as described previously; [216]), and the fractional yields of the peptides and free [^{35}S]Met in each reaction at different times were quantified and fit to the single exponential equation $y = Y_{max} (1-\exp [-k_{obs} \times t])$, where Y_{max} is the maximum fraction of peptide formed and k_{obs} is the observed rate constant. Toeprinting assays were performed as described [230] with minor variations.

Preparation of Initiation and Elongation Factors

Recombinant initiation factors eIF1, eIF1A, and eIF5 were purified as described previously [231]. Initiation factor eIF5B [232], native elongation factor eEF1A [216], and polyHis-tagged versions of elongation factors eEF2 and eEF3 [233,234] were purified from yeast using published protocols with some modifications. Recombinant, hypusinated eIF5A was prepared by coexpressing His$_6$-eIF5A, Dys1, and Lia1 in *E. coli*, and the hypusine modification was analyzed by MS. Cys mutants of eIF5A used for hydroxyl radical cleavage studies were purified from yeast as previously described [204].

Preparation of mRNA, tRNA, and Ribosomes

Unstructured model mRNAs based on the template 5′-GGAA$(UC)_7$U-peptide-coding-sequence-$(CU)_{10}$C-3′ with codons Met(AUG), Pro(CCA), Phe(UUC), and stop(UAA) were prepared by T7 in vitro transcription or purchased. The UGG isoacceptor of $tRNA^{Pro}$ was purified from bulk *S. cerevisiae* tRNA using a biotinylated oligonucleotide [198], and $tRNA^{Phe}$ (Chemical Block) and $tRNA^{Lys}$ (tRNA Probes) were purchased. The tRNAs were aminoacylated using *S. cerevisiae* His_6-tagged ProRS or yeast postribosomal supernatant (S100). Ribosomal subunits were prepared from the yeast strain YRP840 as described previously [235].

Directed Hydroxyl Radical Cleavage Analysis

Following addition of Fe(II)-BABE-modified eIF5A to 80S initiation complexes, hydroxyl radicals were generated by the Fenton reaction, and primer extension analyses were used to monitor rRNA cleavage sites as described previously [235].

Yeast Strains Used

Yeast Strains

Strain		Description	Source
H1511	*MAT□ura3-52 leu2-3 leu2-112 trp1-Δ63*	[236]	
J697	*MATα trp1-Δ63 ura3-52 leu2-3 leu2-112 gcn2Δ tif51b::NAT tif51a::KANMX4* p[*TIF51A, LEU2*]	[204]	
J699	*MATα trp1-Δ63 ura3-52 leu2-3 leu2-112 gcn2Δ tif51b::NAT tif51a::KANMX4* p[*tif51a-S149P, LEU2*]	[204]	
J828	*MAT□trp1-Δ63 ura3-52 leu2-3 leu2-112 GAL2+ gcn2Δ tif51a::KanMX4 tif51b::NAT* p[*LEU2*, *TIF51A*]	This study	
J832	*MAT□trp1-Δ63 ura3-52 leu2-3 leu2-112 GAL2+*	This study	

	gcn2Δ tif51a::*KanMX4 tif51b*::*NAT*	
	p[*LEU2*, *tif51a-ΔC(C23A,C39T)*]	
J836	*MAT□trp1-Δ63 ura3-52 leu2-3 leu2-112 GAL2+*	This study
	gcn2Δ tif51a::*KanMX4 tif51b*::*NAT*	
	p[*LEU2*, *tif51a-ΔC-M105C*]	
J838	*MAT□trp1-Δ63 ura3-52 leu2-3 leu2-112 GAL2+*	This study
	gcn2Δ tif51a::*KanMX4 tif51b*::*NAT*	
	p[*LEU2*, *tif51a-ΔC-T126C*]	
J917	*MAT□trp1-Δ63 ura3-52 leu2-3 leu2-112 GAL2+*	This study
	gcn2Δ tif51a::*KanMX4 tif51b*::*NAT*	
	p[*LEU2*, *tif51a-ΔC-S36C*]	
J920	*MAT□trp1-Δ63 ura3-52 leu2-3 leu2-112 GAL2+*	This study
	gcn2Δ tif51a::*KanMX4 tif51b*::*NAT*	
	p[*LEU2*, *tif51a-ΔC-K48C*]	
TKY597	*MATα leu2-3 leu2-112 ura3-52 trp1-7 yef3*::*LEU2*	(Anand et al., 2003)
	lys2 his4-713 met2-1 p[*YEF3, TRP1*]	
TKY599	*MATα leu2-3 leu2-112 ura3-52 trp1-7 yef3*::*LEU2*	(Anand et al., 2003)
	lys2 his4-713 met2-1 p[*yef3-F650S, TRP1*]	
TKY675	*MATa ade2 leu2 ura3 his3 trp1 eft1*::*HIS3*	[237]
	eft2::*TRP1* p[*EFT2-6xHis, LEU2*]	
TKY702	*MATα leu2-3 leu2-112 ura3-52 trp1-7 yef3*::*LEU2*	(Anand et al., 2003)
	lys2 his4-713 met2-1 p[*YEF3-6xHis, TRP1*]	
TKY742	*MATa ade2 leu2 ura3 his3 trp1 eft1*::*HIS3 eft2*::*TRP1*	[234]
	p[*eft2-6xHis-H699N*, *LEU2*]	
TKY825	*MATa ade2 leu2 ura3 his3 trp1 eft1*::*HIS3 eft2*::*TRP1*	[234]
	p[*eft2-6xHis-H696A*, *LEU2*]	
YRP840	*MATa leu2-3 leu2-112 his4-539 trp1 ura3-52*	[238]

cup1::*LEU2/PGK1pG/MFA2pG*

Dual Luciferase Assay

Dual luciferase reporter constructs containing 10 codon repeats at amino acid 314 between the in-frame *Renilla* and firefly luciferase open reading frames (ORFs) were obtained from Elizabeth Grayhack [214]. Constructs with insertions of variable numbers of Pro or Phe codons between the two luciferase ORFs were generated by cloning PCR products between the unique *Sal*I and *Pst*I restriction sites of the parental dual-luciferase plasmid pDL202. Individual yeast transformants were grown in SD medium containing required nutrients to OD_{600} = 0.8–1.0, harvested, and the cell pellets were resuspended in 300 µl Breaking Buffer L (20 mM Tris-HCl [pH 7.5], 100 mM KCl, 5 mM $MgCl_2$, 0.1 mM EDTA, Complete Protease Inhibitor cocktail [Roche]), and vigorously mixed with 1 vol glass beads on a vortex for 1 min at 4°C. Lysates were cleared by centrifugation at 9,400 x *g* for 5 min at 4°C, and aliquots of the supernatant were assayed for firefly luciferase activity using a microplate luminometer (Berthold) and the Dual Luciferase Reporter 1000 Assay System (Promega). Next, *Renilla* luciferase activity was measured following addition of Stop and Glo reagent (Promega).

Peptide Formation Assay

Initiation complexes were prepared in 1X Recon Buffer A (30 mM HEPES-KOH [pH 7.4], 100 mM potassium acetate, 2.5 mM magnesium acetate, 2 mM DTT) based on the protocol described by Acker et al. [229] and contained the following components: 4 nM [^{35}S]Met-$tRNA_i^{Met}$, 0.4 µM eIF2, 1 µM eIF1, µM eIF1A, 0.4 µM 40S, 1 µM mRNA, 1 µM eIF5 and 0.5 µM eIF5B. All reactions were performed at 26°C. Following assembly of initiation complexes, reactions were layered on 0.8 ml 1M sucrose cushion in 1X Recon Buffer A and then pelleted by

centrifugation at 260,000 x *g* for 1 h. Ribosomal pellets were dissolved in 1X Recon Buffer B (30 mM HEPES-KOH [pH 7.4], 100 mM potassium acetate, 1 mM magnesium acetate, 2 mM magnesium chloride, 1 mM spermidine, 2 mM DTT), and aliquots were stored at -80°C.

Peptide formation assays contained 2 nM initiation complex, 2 μM eEF1A, 1 μM eEF2, 1 μM eEF3, 5 μM eIF5A, 1 μM Phe- (or Pro) tRNA, 1 mM GTP and 1 mM ATP in 1X Recon Buffer B. The elongation assay components were pre-incubated for 15 min on ice before adding the initiation complex, and then reactions were incubated at 26°C. Progress of peptide formation was examined by electrophoretic TLC as described previously [216]. Briefly, elongation reactions were quenched at different times by mixing with an equivalent vol of 0.2 N KOH, and then 0.5 μl was spotted on a cellulose TLC plate (EMD Chemicals). The spot was dried using a heat gun, then the TLC plate was briefly equilibrated with pyridine acetate buffer (200 ml glacial acetic acid and 5 ml pyridine in 1 l, pH 2.8) before electrophoresis at 1,000 V for 30 min in the same pyridine acetate buffer. Following electrophoresis, the TLC plate was dried using a heat gun, and peptide spots were detected by phosphorimage analysis. The fractional yield of the peptides and free [^{35}S]Met in each reaction at different times were quantified and fit using Kaleidagraph (Synergy Software) to the single exponential equation: $y = Y_{max} (1-\exp (-k_{obs}t))$, where Y_{max} is the maximum fraction of peptide formed and k_{obs} is observed rate constant.

Preparation of Initiation and Elongation Factors

Initiation factors eIF1, eIF1A and eIF5 were expressed in *E. coli* BL21 CodonPlus (DE3)-RIL cells (Agilent Technologies) and purified using the IMPACT Protein Purification System (New England Biolabs) as described previously [231]. Initiation factor eIF5B was purified from yeast strain using Glutathione Sepharose 4B affinity chromatography as described

previously [232] with minor modification. Briefly, yeast strain H1511 harboring the expression vector pEG-KT-eIF5B$^{397\text{-}1002}$ [232,239] was grown in S-raffinose medium (0.145% yeast nitrogen base, 0.5% ammonium sulfate, and 2% raffinose plus required supplements) until $A_{600} = 0.5$, and GST-eIF5B$^{397\text{-}1002}$ expression was induced by adding galactose to final 2% (v/v) and incubating the culture with shaking at 30°C for 14 h. Cells were harvested and the cell pellet was suspended in an equivalent volume of Lysis Buffer A (1X phosphate buffered saline [PBS] solution containing Complete Protease Inhibitor cocktail [EDTA-free, Roche], 0.5 mM 4-[2-Aminoethyl] benzenesulfonyl fluoride hydrochloride [AEBSF], 5 μg/ml pepstatin). Cells were broken by adding 50% (v/v) glass beads to the cell suspension and then mixing vigorously on a vortex for 5 min at 4°C. Following removal of the glass beads and unbroken cells by centrifugation at 1,900 x *g* for 10 min, the extract was clarified by centrifugation at 27,000 x *g* for 30 min, and then mixed with 1 ml of a 50% slurry of Glutathione Sepharose 4B (GE Healthcare) at 4°C for 2 h. The resin was then washed extensively with 20-fold excess volume of 1× PBS buffer, and eIF5B was eluted by adding 40 U/ml thrombin in 1× PBS buffer and incubating at room temperature for 2 h and then overnight at 4°C. The supernatant containing released eIF5B was dialyzed against Storage Buffer (20 mM HEPES-KOH [pH 7.5], 100 mM potassium acetate, 2.5 mM magnesium acetate, 2 mM DTT, and 10% glycerol) and stored at -80°C.

Elongation factor eEF1A was purified from yeast strain YRP840 [238]. Cells were grown in 2 l YPD to $A_{600} = 3.0$, harvested and broken in 50 ml Lysis Buffer B (60 mM Tris–HCl [pH 7.5], 50 mM KCl, 5 mM $MgCl_2$, 0.1 M EDTA [pH 8.0], 10% glycerol, 1 mM DTT, 0.2 M AEBSF) using glass beads as described above. After removal of unbroken cells by centrifugation at 7,600 x *g* for 10 min, the supernatant was clarified by centrifugation at 150,000 x *g* for 3 h, and then mixed gently with 10 ml DE52 resin (Whatman, pre-equilibrated with Lysis Buffer B) for 1 h at

4°C. The unbound fraction containing eEF1A was isolated by pouring the mixture into a column and collecting the elute, which was then applied to a HiTrap CM Sepharose column (GE Healthcare), and eEF1A was eluted with a linear gradient to 300 mM KCl. Fractions containing eEF1A were identified by SDS-PAGE, pooled, dialyzed against Storage Buffer and store at -80°C.

Poly-histidine tagged versions of elongation factors eEF2 and eEF3 were purified from yeast strains TKY675 and TKY702, respectively, using published protocols [233,234] with some modifications. Cells were grown in 2.5 l YPD to A_{600} = 1.5, harvested, and then suspended in Lysis Buffer C (50 mM Tris-HCl [pH 8.0], 300 mM KCl, 1 mM 2-mercaptoethanol, 1 mM AEBSF, 10 mM imidazole, and 1X Complete protease inhibitor EDTA-free). After the cells were broken with glass beads as described above, the lysate was cleared of unbroken cells by centrifugation at 17,000 x *g* for 30 min, clarified by centrifugation at 180,000 x *g* for 80 min, and then gently mixed with 1 ml Ni-NTA resin for 2 h at 4°C. The resin was then washed with 5 vol of the Lysis Buffer C containing 20 mM imidazole, and the His-tagged proteins were eluted in Buffer C containing 250 mM imidazole. The eluted proteins were mixed with a 6-fold excess of Dilution Buffer (20 mM Tris-HCl [pH 7.5], 2 mM DTT and 10% [v/v] glycerol), loaded on a HiTrapQ HP column (GE Healthcare), and then eluted by a linear KCl gradient to 1 M. Fractions containing eEF2 or eEF3 were identified by SDS-PAGE, pooled, dialyzed against Storage Buffer, and then stored at -80°C.

Preparation of eIF5A

The polycistronic expression system developed by Song Tan [240] was used to produce hypusinated eIF5A in *E. coli* (Figure S4A). First, an N-terminally His_6-tagged version of the

eIF5A open reading frame was cloned between the *Nde*I and *Bam*HI sites of the vector pET3a Trm, and then moved as an *Xba*I-*Bam*HI fragment to the expression vector pST39 generating the plasmid pC4181. The *Dys1* and *Lia1* open reading frames were cloned into pET3a Trm using *Nde*I-*Hin*dIII and *Nde*I-*Mlu*I sites, respectively, and then sequentially transferred to pST39 using *Eco*RI-*Hin*dIII and *Bsp*EI-*Mlu*I sites, respectively, to make the plasmid pC4183.

To purify eIF5A, *E. coli* strain BL21 CodonPlus (DE3)-RIL (Agilent) was transformed with pC4181 (for unmodified eIF5A) or pC4183 (for hypusinated eIF5A) and cells were grown in 1 l LB medium containing 100 μg/ml ampicillin at 37°C to A_{600} = 0.5. Then, 0.5 mM IPTG was added and the culture was incubated at 25°C for 14 h. Following harvesting, the cell pellet was suspended in 40 ml Lysis Buffer D (50 mM Tris-HCl [pH 7.5], 300 mM KCl, 10 mM imidazole, 0.5 mM AEBSF) and cells were broken by sonication using a microtip (5 cycles of 30 sec pulse followed by 30 sec cooling at 4°C). The cell lysate was cleared by centrifugation at 27,000 x *g* for 30 min and then mixed gently with 1 ml Ni-NTA resin (Qiagen) at 4°C for 2 h. The resin was transferred to a disposable column (Qiagen), washed sequentially with 10 ml Lysis Buffer D and then 10 ml Lysis Buffer D containing 20 mM imidazole, and then protein was eluted in 10 ml Lysis Buffer D containing 0.5 M imidazole. The elute was diluted five times with Dilution Buffer (20 mM Tris-HCl [pH 7.5], 10% glycerol, 2 mM DTT), loaded on a HiTrapQ FPLC column, and the bound proteins were eluted in a 100 mM to 1M KCl gradient in Dilution Buffer. Fractions containing eIF5A (eluting near 0.3 M KCl) were identified by SDS-PAGE, pooled, and then dialyzed against 30 mM Hepes-KOH (pH 7.5), 150 mM potassium chloride, 2 mM DTT and 10% glycerol.

Hypusine modification of eIF5A was analyzed by ElectroSpray-Ionization Quadrupole-Time-of-Flight Mass Spectrometry (ESI QTOF MS, Agilent). As shown in Figure S4B, analysis

of unmodified eIF5A purified from *E. coli* showed a single major peak with a calculated molecular weight of 17,805.94, consistent with the predicted mass of eIF5A with the first methionine removed. Analysis of the eIF5A purified from *E. coli* co-expressing Dys1 and Lia1 revealed a major peak with molecular weight of 17892.79. The 87 Da increase in mass is consistent with the mass of the hypusine modification and indicates that the majority of eIF5A was hypusinated.

Cys mutants of eIF5A used for hydroxyl radical cleavage studies were purified from yeast (Table 1), according to Saini et al. [204] with minor modification. Cells were grown in 2.5 l YPD to A_{600} = ~3.0, harvested, and then the cell pellet was washed in 1X TBS (50 mM Tris-HCl [pH 7.5], 150 mM NaCl), suspended in 4 pellet vol 1X TBS containing 1X Complete Protease Inhibitor and 1mM AEBSF, mixed with 50% vol glass beads, and broken by vigorous mixing on a vortex for 5 min at 4°C. Lysates were cleared by centrifugation at 27,000 x *g* for 30 min, and then gently mixed for 2 h at 4°C with 1 ml anti-flag-M2 affinity gel (Sigma, pre-equilibrated with 1X TBS). The resin was transferred to a 1 ml disposable column, washed with 10 ml 1X TBS, and bound proteins were eluted in 1 ml 1X TBS containing 200 µg/ml FLAG peptide (Sigma). Eluted proteins were dialyzed against eIF5A Storage Buffer (30 mM HEPES-KOH [pH 7.5], 150 mM KCl, 2 mM DTT, 10% glycerol), and stored at -80°C. Fe(II)-BABE derivatization of purified eIF5A was performed as described previously [231].

Preparation of mRNA, tRNA, and Ribosomes

Unstructured model mRNAs based on the template: 5′-GGAA$(UC)_7$U-*peptide-coding-sequence*-$(CU)_{10}$C-3′ were prepared by T7 *in vitro* transcription or purchased (Integrated DNA

Technologies) and used for preparation of elongation complexes. The following codons were used to encode: Met(AUG), Pro(CCA), Phe(UUC), and Stop (UAA).

The UGG isoacceptor of $tRNA^{Pro}$ was purified from bulk *S. cerevisiae* tRNA (Roche) using the biotinylated oligo 5'-CCAAAGCGAG AATCATACCA CTAGAC-3' (BioTEG) as follows [198]: 400 µl streptavidin beads (Pierce) were washed three times with 400 µl 10 mM Tris-HCl (pH 7.5), bound to 8 nmol biotinylated oligonucleotide at 25°C for 30 min, and then washed twice with 10 mM Tris-HCl (pH 7.5). The beads were pelleted by centrifugation for 30 sec at 1000 x *g* after each wash to remove the supernatant. Bulk tRNA (180 nmol in 300 µl) was mixed with an equal volume of 2 M TMA buffer (20 mM Tris-HCl [pH 7.5], 1.8 M tetramethylammonium chloride, 0.2 mM EDTA), incubated with the streptavidin beads at 65°C for 10 min to denature the tRNA, then the mixture was slowly cooled to 25°C over ~10 min to allow annealing. The beads were then washed eight times with 400 µl 10 mM Tris-HCl (pH 7.5) to remove unbound tRNA. The $tRNA_{UGG}$ was melted off the beads by heating to 65°C for 5 min and then eluted by centrifugation of the beads in a new tube pre-loaded with 2 µl 1M magnesium acetate. After repeating this melting and elution process, the two eluted fractions were combined and precipitated with ethanol. The tRNA was then resuspended in 50 µl water and quantified by measuring the A_{260}. The purified $tRNA^{Pro}$ was further treated with CCA-adding enzyme to increase the proline charging efficiency. For the CCA-adding reaction, 20 µM $tRNA^{Pro}$ was mixed with 10 mM ATP, 10 mM CTP, and 6 µM nucleotidyl transferase (CCA-adding enzyme) in 1X CCA Buffer (50 mM Tris-HCl [pH 7.5], 20 mM magnesium chloride, 0.5 mM DTT), and then incubated at 37°C for 30 min. For aminoacylation of $tRNA^{Pro}$, 5 µM $tRNA^{Pro}$ was mixed with 2 mM ATP-Mg^{2+}, 0.3 mM proline, and 2.5% reaction vol ProRS in 1X Reaction Buffer (40 mM Tris-HCl [pH 7.6], 10 mM magnesium acetate, 1 mM DTT), and then incubated at 30°C for

30 min. Typical charging efficiency of $tRNA^{Pro}$ with yeast ProRS was ~20%, as determined using [^{3}H]proline (Perkin-Elmer). The *S. cerevisiae* His_6-tagged ProRS was purified as described [241].

$tRNA^{Phe}$ (Chemical Block) and $tRNA^{Lys}$ (tRNA Probes, College Station, TX) were aminoacylated using a post-ribosomal supernatant (S100) as the source of PheRS and LysRS. The S100 was prepared by growing strain BY4741 in 10 l YPD to A_{600}=1.0 and then lysing cells with glass beads as described above in 10 mM potassium phosphate buffer (pH 7.2) containing EDTA-free Complete Protease Inhibitor cocktail (Roche). Lysates were cleared by centrifugation at 27,000 x *g* for 30 min, and the S100 was obtained following centrifugation at 100,000 x *g* for 3 h. The S100 was mixed with DE52 resin for 1 h at 4°C, then the mixture was poured into a column, washed with 5 column vol 10 mM potassium phosphate buffer (pH 7.2), and bound proteins were eluted with 250 mM potassium phosphate (pH 6.5). Following addition of glycerol (5%) and DTT (2 mM) to the eluate, aliquots were stored at -80°C.

Ribosomal subunits were prepared from the yeast strain YRP840 (also known as YAS2488) as described previously [235].

Polysome Analysis

Yeast cultures were either treated with 50 µg/ml CHX for 5 min before collection or left untreated, transferred to a 500-ml centrifuge bottle containing shaved ice, pelleted, and washed with 10 ml Buffer P (20 mM Tris-HCl [pH 7.5], 50 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, Complete Protease Inhibitor cocktail [Roche]). In all subsequent steps, the CHX-treated cells were treated with 50 µg/ml CHX. Cell pellets were suspended in 300–500 µl Buffer P, mixed with an equal volume of glass beads, and then cells were broken by 5 cycles of vigorous mixing

on a vortex for 1 min followed by 1 min on ice. Following clarification, ten A_{260} units of the whole cell extracts (WCEs) were layered on 4.5–45% sucrose gradients prepared in 20 mM Tris-HCl [pH 7.5], 50 mM KCl, 10 mM $MgCl_2$ and 1 mM DTT, and then subjected to centrifugation in a Beckman SW41 rotor for 2.5 h at 260,000 x g. Gradients were fractionated while monitoring absorbance at A_{254}.

Toeprinting Assay

Toeprinting was performed as described [230] with minor variations. Initiation complexes were prepared as described above except that 0.4 µM Met-$tRNA_i^{Met}$ was used in place of 4 nM [^{35}S]Met-$tRNA_i^{Met}$, and the mRNAs coded for MPPPPP (5'-GGAA$[UC]_7$UAAAA-AUGCCACCACCACCACCAUAA$[UC]_{22}$GUUAAUAAGCAAAAUUCAUUAUAACC-3') or MFFFFF (5'-GGAA$[UC]_7$UAAAAAUGUUCUUCUUCUUCUUCUAA$[UC]_{22}$ GUUAAUAAGCAAAAUUCAUUAUAACC-3'). Initiation complexes were reacted with elongation factors and tRNA as described in the peptide formation assay in a final vol of 15 µl and incubated at 26°C for 5 min. Following addition of 0.3 µl 1 M $MgCl_2$, two pmol ^{32}P-labeled toeprinting primer (5′-GGTTATAATGAATTTTGCTTATTAAC-3′) [96] and 0.1 µl SUPERase-In (Ambion) were added to each reaction mixture. Samples were diluted with 3 vol Primer Extension Mix (10 mM Tris-HCl [pH 7.5], 10 mM $MgCl_2$, 60 mM NH_4Cl, 6 mM 2-mercaptoethanol, 360 µM each dNTP and 0.5 U/µl AMV Reverse Transcriptase [Roche]) and then incubated for 30 min at 30°C to allow primer extension. Products were diluted with 2 vol formamide/EDTA loading buffer (1x TBE, 17 mM EDTA [pH 8.0], 90% formamide, 0.001% bromophenol blue, 0.001% xylene cyanol) and then resolved on a sequencing gel (8% polyacrylamide, 7 M Urea).

Directed Hydroxyl Radical Cleavage Analysis

80S initiation complexes were assembled in 1X Recon Buffer A as schemed in Figure 6B, and then Fe(II)-BABE-modified eIF5A was added to the reaction. Following incubation on ice for 5 min, hydroxyl radicals were generated by the Fenton reaction as described in Shin et al. [231]. Final concentrations for each component in the reaction were: 0.4 μM eIF2, 0.4 μM Met-$tRNA_i^{Met}$, 1 μM eIF1, 1 μM eIF1A, 1 μM mRNA, 0.2 μM 40S, 0.8 μM eIF5, 0.4 μM eIF5B, 0.2 μM 60S and 0.2 μM Fe(II)-BABE-modified eIF5A. To analyze cleavage of $tRNA_i^{Met}$, 0.1 μM Met-[^{32}P]$tRNA_i^{Met}$ was used in place of unlabeled Met-$tRNA_i^{Met}$ and the concentration of 40S and 60S subunits was increased to 0.4 μM. Primer extension analysis of rRNA cleavage sites were performed as described previously [231] using the previously described seven primers for 18S rRNA [231] and 13 primers for 25S rRNA:

25-1, 5′-CAGACAACAAAGGCTTAATCTC-3′;

25-2, 5′-CTGACTTAGAGGCGTTCAGCC-3′;

25-3, 5′-CTGCCACAAGCCAGTTATCCC-3′;

25-4, 5′-AGCTCCGCTTCATTGAATAAG-3′;

25-5, 5′-TCATAGTTACTCCCGCCGTTTACC-3′;

25-6, 5′-CCAAGCAGTCCACAAGCACGC-3′;

25-7, 5′-GTGATAAGCTGTTAAGAAGAA-3′;

25-8, 5′-GCTACTACCACCAAGATCTGC-3′;

25-9, 5′-GCACCTTAACTCTACGTTCGG-3′;

25-10, 5′-TATACCCAAATTCGACGATCG-3′;

25-11, 5′-GCGGCATATAACCATTATGCC-3′;

25-12, 5′-TTCCATCACTGTACTTGTTCG-3′;

25-13, 5′-AGGAACATAGACAAGGAACGG-3′.

References

(1) Ban, N.; Nissen, P.; Hansen, J.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 905–920.

(2) Woese, C. R. *RNA N. Y. N* **2001**, *7*, 1055–1067.

(3) Schmeing, T. M.; Ramakrishnan, V. *Nature* **2009**, *461*, 1234–1242.

(4) Voorhees, R. M.; Ramakrishnan, V. *Annu. Rev. Biochem.* **2013**, *82*, 203–236.

(5) Watson, J. D.; Baker, T. A.; Bell, S. P.; Gann, A.; Levine, M.; Losick, Ri. *Molecular Biology of the Gene*; 7th ed.; Pearson ; Cold Spring Harbour Laboratory Press: Boston [u.a.]; Cold Spring Harbour, NY, 2013.

(6) Boelens, R.; Gualerzi, C. O. *Curr. Protein Pept. Sci.* **2002**, *3*, 107–119.

(7) Simonetti, A.; Marzi, S.; Jenner, L.; Myasnikov, A.; Romby, P.; Yusupova, G.; Klaholz, B. P.; Yusupov, M. *Cell. Mol. Life Sci. CMLS* **2009**, *66*, 423–436.

(8) Carter, A. P.; Clemons, W. M., Jr; Brodersen, D. E.; Morgan-Warren, R. J.; Hartsch, T.; Wimberly, B. T.; Ramakrishnan, V. *Science* **2001**, *291*, 498–501.

(9) Marzi, S.; Myasnikov, A. G.; Serganov, A.; Ehresmann, C.; Romby, P.; Yusupov, M.; Klaholz, B. P. *Cell* **2007**, *130*, 1019–1031.

(10) Shine, J.; Dalgarno, L. *Proc. Natl. Acad. Sci. U. S. A.* **1974**, *71*, 1342–1346.

(11) Diaconu, M.; Kothe, U.; Schlünzen, F.; Fischer, N.; Harms, J. M.; Tonevitsky, A. G.; Stark, H.; Rodnina, M. V.; Wahl, M. C. *Cell* **2005**, *121*, 991–1004.

(12) Ogle, J. M.; Brodersen, D. E.; Clemons, W. M., Jr; Tarry, M. J.; Carter, A. P.; Ramakrishnan, V. *Science* **2001**, *292*, 897–902.

(13) Battle, D. J.; Doudna, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 11676–11681.

(14) Gromadski, K. B.; Rodnina, M. V. *Mol. Cell* **2004**, *13*, 191–200.

(15) Ogle, J. M.; Murphy, F. V.; Tarry, M. J.; Ramakrishnan, V. *Cell* **2002**, *111*, 721–732.

(16) Pape, T.; Wintermeyer, W.; Rodnina, M. *EMBO J.* **1999**, *18*, 3800–3807.

(17) Moazed, D.; Noller, H. F. *Nature* **1989**, *342*, 142–148.

(18) Schmeing, T. M.; Voorhees, R. M.; Kelley, A. C.; Gao, Y.-G.; Murphy, F. V., 4th; Weir, J. R.; Ramakrishnan, V. *Science* **2009**, *326*, 688–694.

(19) Valle, M.; Sengupta, J.; Swami, N. K.; Grassucci, R. A.; Burkhardt, N.; Nierhaus, K. H.; Agrawal, R. K.; Frank, J. *EMBO J.* **2002**, *21*, 3557–3567.

(20) Daviter, T.; Wieden, H.-J.; Rodnina, M. V. *J. Mol. Biol.* **2003**, *332*, 689–699.

(21) Voorhees, R. M.; Schmeing, T. M.; Kelley, A. C.; Ramakrishnan, V. *Science* **2010**, *330*, 835–838.

(22) Berchtold, H.; Reshetnikova, L.; Reiser, C. O.; Schirmer, N. K.; Sprinzl, M.; Hilgenfeld, R. *Nature* **1993**, *365*, 126–132.

(23) Nissen, P.; Kjeldgaard, M.; Thirup, S.; Polekhina, G.; Reshetnikova, L.; Clark, B. F.; Nyborg, J. *Science* **1995**, *270*, 1464–1472.

(24) Valle, M.; Zavialov, A.; Li, W.; Stagg, S. M.; Sengupta, J.; Nielsen, R. C.; Nissen, P.; Harvey, S. C.; Ehrenberg, M.; Frank, J. *Nat. Struct. Biol.* **2003**, *10*, 899–906.

(25) Thompson, R. C.; Stone, P. J. *Proc. Natl. Acad. Sci. U. S. A.* **1977**, *74*, 198–202.

(26) Kim, D. F.; Green, R. *Mol. Cell* **1999**, *4*, 859–864.

(27) Schmeing, M. T.; Huang, K. S.; Strobel, S. A.; Steitz, T. A. *Nature* **2005**, *438*, 520–524.

(28) Nissen, P.; Hansen, J.; Ban, N.; Moore, P. B.; Steitz, T. A. *Science* **2000**, *289*, 920–930.

(29) Voorhees, R. M.; Weixlbaumer, A.; Loakes, D.; Kelley, A. C.; Ramakrishnan, V. *Nat. Struct. Mol. Biol.* **2009**, *16*, 528–533.

(30) Sievers, A.; Beringer, M.; Rodnina, M. V.; Wolfenden, R. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 7897–7901.

(31) Carrasco, N.; Hiller, D. A.; Strobel, S. A. *Biochemistry (Mosc.)* **2011**, *50*, 10491–10498.

(32) Hiller, D. A.; Singh, V.; Zhong, M.; Strobel, S. A. *Nature* **2011**, *476*, 236–239.

(33) Kuhlenkoetter, S.; Wintermeyer, W.; Rodnina, M. V. *Nature* **2011**, *476*, 351–354.

(34) Dorner, S.; Panuschka, C.; Schmid, W.; Barta, A. *Nucleic Acids Res.* **2003**, *31*, 6536–6542.

(35) Zaher, H. S.; Shaw, J. J.; Strobel, S. A.; Green, R. *EMBO J.* **2011**, *30*, 2445–2453.

(36) Agirrezabala, X.; Lei, J.; Brunelle, J. L.; Ortiz-Meoz, R. F.; Green, R.; Frank, J. *Mol. Cell* **2008**, *32*, 190–197.

(37) Guo, Z.; Noller, H. F. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 20391–20394.

(38) Dorner, S.; Brunelle, J. L.; Sharma, D.; Green, R. *Nat. Struct. Mol. Biol.* **2006**, *13*, 234–241.

(39) Frank, J.; Agrawal, R. K. *Nature* **2000**, *406*, 318–322.

(40) Spiegel, P. C.; Ermolenko, D. N.; Noller, H. F. *RNA N. Y. N* **2007**, *13*, 1473–1482.

(41) Zhang, W.; Dunkle, J. A.; Cate, J. H. D. *Science* **2009**, *325*, 1014–1017.

(42) Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H. D. *Science* **2005**, *310*, 827–834.

(43) Rodnina, M. V.; Savelsbergh, A.; Katunin, V. I.; Wintermeyer, W. *Nature* **1997**, *385*, 37–41.

(44) Zhou, J.; Lancaster, L.; Donohue, J. P.; Noller, H. F. *Science* **2013**, *340*, 1236086.

(45) Rodnina, M. V. *Science* **2013**, *340*, 1534–1535.

(46) Lu, J.; Kobertz, W. R.; Deutsch, C. *J. Mol. Biol.* **2007**, *371*, 1378–1391.

(47) Bhushan, S.; Gartmann, M.; Halic, M.; Armache, J.-P.; Jarasch, A.; Mielke, T.; Berninghausen, O.; Wilson, D. N.; Beckmann, R. *Nat. Struct. Mol. Biol.* **2010**, *17*, 313–317.

(48) Wilson, D. N.; Beckmann, R. *Curr. Opin. Struct. Biol.* **2011**, *21*, 274–282.

(49) Voss, N. R.; Gerstein, M.; Steitz, T. A.; Moore, P. B. *J. Mol. Biol.* **2006**, *360*, 893–906.

(50) Ito, K.; Chiba, S.; Pogliano, K. *Biochem. Biophys. Res. Commun.* **2010**, *393*, 1–5.

(51) Brenner, S.; Stretton, A. O.; Kaplan, S. *Nature* **1965**, *206*, 994–998.

(52) Brenner, S.; Barnett, L.; Katz, E. R.; Crick, F. H. *Nature* **1967**, *213*, 449–450.

(53) Korostelev, A. A. *RNA* **2011**, *17*, 1409–1421.

(54) Youngman, E. M.; McDonald, M. E.; Green, R. *Annu. Rev. Microbiol.* **2008**, *62*, 353–373.

(55) Nakamura, Y.; Ito, K. *Wiley Interdiscip. Rev. RNA* **2011**, *2*, 647–668.

(56) Klaholz, B. P.; Pape, T.; Zavialov, A. V.; Myasnikov, A. G.; Orlova, E. V.; Vestergaard, B.; Ehrenberg, M.; van Heel, M. *Nature* **2003**, *421*, 90–94.

(57) Nakamura, Y.; Ito, K.; Ehrenberg, M. *Cell* **2000**, *101*, 349–352.

(58) Scarlett, D.-J. G.; McCaughan, K. K.; Wilson, D. N.; Tate, W. P. *J. Biol. Chem.* **2003**, *278*, 15095–15104.

(59) Frolova, L. Y.; Tsivkovskii, R. Y.; Sivolobova, G. F.; Oparina, N. Y.; Serpinsky, O. I.; Blinov, V. M.; Tatkov, S. I.; Kisselev, L. L. *RNA N. Y. N* **1999**, *5*, 1014–1020.

(60) Shaw, J. J.; Green, R. *Mol. Cell* **2007**, *28*, 458–467.

(61) Trobro, S.; Aqvist, J. *Biochemistry (Mosc.)* **2009**, *48*, 11296–11303.

(62) Zavialov, A. V.; Buckingham, R. H.; Ehrenberg, M. *Cell* **2001**, *107*, 115–124.

(63) McDonald, M. E.; Green, R. *RNA N. Y. N* **2012**, *18*, 605–609.

(64) Kaji, A.; Kiel, M. C.; Hirokawa, G.; Muto, A. R.; Inokuchi, Y.; Kaji, H. *Cold Spring Harb. Symp. Quant. Biol.* **2001**, *66*, 515–530.

(65) Peske, F.; Rodnina, M. V.; Wintermeyer, W. *Mol. Cell* **2005**, *18*, 403–412.

(66) Petry, S.; Weixlbaumer, A.; Ramakrishnan, V. *Curr. Opin. Struct. Biol.* **2008**, *18*, 70–77.

(67) Pech, M.; Nierhaus, K. H. *Mol. Microbiol.* **2012**, *86*, 6–9.

(68) Janssen, B. D.; Hayes, C. S. *Adv. Protein Chem. Struct. Biol.* **2012**, *86*, 151–191.

(69) Moore, S. D.; Sauer, R. T. *Annu. Rev. Biochem.* **2007**, *76*, 101–124.

(70) Karzai, A. W.; Susskind, M. M.; Sauer, R. T. *EMBO J.* **1999**, *18*, 3793–3799.

(71) Gutmann, S.; Haebel, P. W.; Metzinger, L.; Sutter, M.; Felden, B.; Ban, N. *Nature* **2003**, *424*, 699–703.

(72) Neubauer, C.; Gillet, R.; Kelley, A. C.; Ramakrishnan, V. *Science* **2012**, *335*, 1366–1369.

(73) Sundermeier, T. R.; Dulebohn, D. P.; Cho, H. J.; Karzai, A. W. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 2316–2321.

(74) Hayes, C. S.; Sauer, R. T. *Mol. Cell* **2003**, *12*, 903–911.

(75) Bessho, Y.; Shibata, R.; Sekine, S.; Murayama, K.; Higashijima, K.; Hori-Takemoto, C.; Shirouzu, M.; Kuramitsu, S.; Yokoyama, S. *Proc. Natl. Acad. Sci.* **2007**, *104*, 8293–8298.

(76) Miller, M. R.; Buskirk, A. R. *RNA N. Y. N* **2014**, *20*, 228–235.

(77) Weis, F.; Bron, P.; Giudice, E.; Rolland, J.-P.; Thomas, D.; Felden, B.; Gillet, R. *EMBO J.* **2010**, *29*, 3810–3818.

(78) Ivanova, N.; Pavlov, M. Y.; Ehrenberg, M. *J. Mol. Biol.* **2005**, *350*, 897–905.

(79) Tu, G. F.; Reid, G. E.; Zhang, J. G.; Moritz, R. L.; Simpson, R. J. *J. Biol. Chem.* **1995**, *270*, 9322–9326.

(80) Flynn, J. M.; Levchenko, I.; Seidel, M.; Wickner, S. H.; Sauer, R. T.; Baker, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 10584–10589.

(81) Keiler, K. C.; Waller, P. R.; Sauer, R. T. *Science* **1996**, *271*, 990–993.

(82) Chadani, Y.; Ono, K.; Ozawa, S.-I.; Takahashi, Y.; Takai, K.; Nanamiya, H.; Tozawa, Y.; Kutsukake, K.; Abo, T. *Mol. Microbiol.* **2010**, *78*, 796–808.

(83) Chadani, Y.; Ono, K.; Kutsukake, K.; Abo, T. *Mol. Microbiol.* **2011**, *80*, 772–785.

(84) Chadani, Y.; Ito, K.; Kutsukake, K.; Abo, T. *Mol. Microbiol.* **2012**, *86*, 37–50.

(85) Handa, Y.; Inaho, N.; Nameki, N. *Nucleic Acids Res.* **2011**, *39*, 1739–1748.

(86) Kogure, H.; Handa, Y.; Nagata, M.; Kanai, N.; Güntert, P.; Kubota, K.; Nameki, N. *Nucleic Acids Res.* **2013**.

(87) Ito, K.; Chiba, S. *Annu. Rev. Biochem.* **2013**, *82*, 171–202.

(88) Chiba, S.; Lamsa, A.; Pogliano, K. *EMBO J.* **2009**, *28*, 3461–3475.

(89) Saller, M. J.; Fusetti, F.; Driessen, A. J. M. *J. Bacteriol.* **2009**, *191*, 6749–6757.

(90) Kol, S.; Majczak, W.; Heerlien, R.; van der Berg, J. P.; Nouwen, N.; Driessen, A. J. M. *J. Mol. Biol.* **2009**, *390*, 893–901.

(91) Luo, Z.; Sachs, M. S. *J. Bacteriol.* **1996**, *178*, 2172–2177.

(92) Wei, J.; Wu, C.; Sachs, M. S. *Mol. Cell. Biol.* **2012**, *32*, 2396–2406.

(93) Spevak, C. C.; Ivanov, I. P.; Sachs, M. S. *J. Biol. Chem.* **2010**, *285*, 40933–40942.

(94) Cruz-Vera, L. R.; Rajagopal, S.; Squires, C.; Yanofsky, C. *Mol. Cell* **2005**, *19*, 333–343.

(95) Nakatogawa, H.; Ito, K. *Cell* **2002**, *108*, 629–636.

(96) Vazquez-Laslop, N.; Thum, C.; Mankin, A. S. *Mol. Cell* **2008**, *30*, 190–202.

(97) Ramu, H.; Vázquez-Laslop, N.; Klepacki, D.; Dai, Q.; Piccirilli, J.; Micura, R.; Mankin, A. S. *Mol. Cell* **2011**, *41*, 321–330.

(98) Tanner, D. R.; Cariello, D. A.; Woolstenhulme, C. J.; Broadbent, M. A.; Buskirk, A. R. *J. Biol. Chem.* **2009**, *284*, 34809–34818.

(99) Gong, F.; Yanofsky, C. *Science* **2002**, *297*, 1864–1867.

(100) Hayes, C. S.; Bose, B.; Sauer, R. T. *J. Biol. Chem.* **2002**, *277*, 33825–33832.

(101) Hayes, C. S.; Bose, B.; Sauer, R. T. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 3440–3445.

(102) Oliver, D.; Norman, J.; Sarker, S. *J. Bacteriol.* **1998**, *180*, 5240–5242.

(103) Driessen, A. J. *Trends Microbiol.* **2001**, *9*, 193–196.

(104) Vrontou, E.; Economou, A. *Biochim. Biophys. Acta* **2004**, *1694*, 67–80.

(105) Mori, H.; Ito, K. *Trends Microbiol.* **2001**, *9*, 494–500.

(106) Nakatogawa, H.; Ito, K. *Mol. Cell* **2001**, *7*, 185–192.

(107) Butkus, M. E.; Prundeanu, L. B.; Oliver, D. B. *J. Bacteriol.* **2003**, *185*, 6719–6722.

(108) Yap, M.-N.; Bernstein, H. D. *Mol. Cell* **2009**, *34*, 201–211.

(109) Bhushan, S.; Hoffmann, T.; Seidelt, B.; Frauenfeld, J.; Mielke, T.; Berninghausen, O.; Wilson, D. N.; Beckmann, R. *PLoS Biol* **2011**, *9*, e1000581.

(110) Woolhead, C. A.; Johnson, A. E.; Bernstein, H. D. *Mol. Cell* **2006**, *22*, 587–598.

(111) Lawrence, M. G.; Lindahl, L.; Zengel, J. M. *J. Bacteriol.* **2008**, *190*, 5862–5869.

(112) Fulle, S.; Gohlke, H. *J. Mol. Biol.* **2009**, *387*, 502–517.

(113) Gumbart, J.; Schreiner, E.; Wilson, D. N.; Beckmann, R.; Schulten, K. *Biophys. J.* **2012**, *103*, 331–341.

(114) Vazquez-Laslop, N.; Ramu, H.; Klepacki, D.; Kannan, K.; Mankin, A. S. *EMBO J.* **2010**, *29*, 3108–3117.

(115) Seidelt, B.; Innis, C. A.; Wilson, D. N.; Gartmann, M.; Armache, J.-P.; Villa, E.; Trabuco, L. G.; Becker, T.; Mielke, T.; Schulten, K.; Steitz, T. A.; Beckmann, R. *Science* **2009**, *326*, 1412–1415.

(116) Muto, H.; Nakatogawa, H.; Ito, K. *Mol. Cell* **2006**, *22*, 545–552.

(117) Muto, H.; Ito, K. *Biochem. Biophys. Res. Commun.* **2008**, *366*, 1043–1047.

(118) Wohlgemuth, I.; Brenner, S.; Beringer, M.; Rodnina, M. V. *J. Biol. Chem.* **2008**, *283*, 32229–32235.

(119) Brock, T. D.; Brock, M. L. *Biochim. Biophys. Acta* **1959**, *33*, 274–275.

(120) Taubman, S. B.; So, A. G.; Young, F. E.; Davie, E. W.; Corcoran, J. W. *Antimicrob. Agents Chemother.* **1963**, *161*, 395–401.

(121) Tenson, T.; Lovmar, M.; Ehrenberg, M. *J. Mol. Biol.* **2003**, *330*, 1005–1014.

(122) Weisblum, B. *Antimicrob. Agents Chemother.* **1995**, *39*, 797–805.

(123) Gupta, P.; Sothiselvam, S.; Vázquez-Laslop, N.; Mankin, A. S. *Nat. Commun.* **2013**, *4*, 1984.

(124) Gryczan, T. J.; Grandi, G.; Hahn, J.; Grandi, R.; Dubnau, D. *Nucleic Acids Res.* **1980**, *8*, 6081–6097.

(125) Horinouchi, S.; Weisblum, B. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, *77*, 7079–7083.

(126) Ramu, H.; Mankin, A.; Vazquez-Laslop, N. *Mol. Microbiol.* **2009**, *71*, 811–824.

(127) Schlünzen, F.; Zarivach, R.; Harms, J.; Bashan, A.; Tocilj, A.; Albrecht, R.; Yonath, A.; Franceschi, F. *Nature* **2001**, *413*, 814–821.

(128) Tu, D.; Blaha, G.; Moore, P. B.; Steitz, T. A. *Cell* **2005**, *121*, 257–270.

(129) Vazquez-Laslop, N.; Klepacki, D.; Mulhearn, D. C.; Ramu, H.; Krasnykh, O.; Franzblau, S.; Mankin, A. S. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 10496–10501.

(130) Kazarinoff, M. N.; Snell, E. E. *J. Biol. Chem.* **1977**, *252*, 7598–7602.

(131) Deeley, M. C.; Yanofsky, C. *J. Bacteriol.* **1981**, *147*, 787–796.

(132) Kamath, A. V.; Yanofsky, C. *J. Bacteriol.* **1997**, *179*, 1780–1786.

(133) Konan, K. V.; Yanofsky, C. *J. Bacteriol.* **2000**, *182*, 3981–3988.

(134) Gong, F.; Ito, K.; Nakamura, Y.; Yanofsky, C. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 8997–9001.

(135) Cruz-Vera, L. R.; Yanofsky, C. *J. Bacteriol.* **2008**, *190*, 4791–4797.

(136) Cruz-Vera, L. R.; Gong, M.; Yanofsky, C. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 3598–3603.

(137) Cruz-Vera, L. R.; New, A.; Squires, C.; Yanofsky, C. *J. Bacteriol.* **2007**, *189*, 3140–3146.

(138) Martínez, A. K.; Gordon, E.; Sengupta, A.; Shirole, N.; Klepacki, D.; Martinez-Garriga, B.; Brown, L. M.; Benedik, M. J.; Yanofsky, C.; Mankin, A. S.; Vazquez-Laslop, N.; Sachs, M. S.; Cruz-Vera, L. R. *Nucleic Acids Res.* **2014**, *42*, 1245–1256.

(139) Cruz-Vera, L. R.; Yang, R.; Yanofsky, C. *J. Bacteriol.* **2009**, *191*, 7001–7006.

(140) Martínez, A. K.; Shirole, N. H.; Murakami, S.; Benedik, M. J.; Sachs, M. S.; Cruz-Vera, L. R. *Nucleic Acids Res.* **2012**, *40*, 2247–2257.

(141) Trabuco, L. G.; Harrison, C. B.; Schreiner, E.; Schulten, K. *Struct. Lond. Engl. 1993* **2010**, *18*, 627–637.

(142) Schmeing, T. M.; Huang, K. S.; Kitchen, D. E.; Strobel, S. A.; Steitz, T. A. *Mol. Cell* **2005**, *20*, 437–448.

(143) Weixlbaumer, A.; Jin, H.; Neubauer, C.; Voorhees, R. M.; Petry, S.; Kelley, A. C.; Ramakrishnan, V. *Science* **2008**, *322*, 953–956.

(144) Lovett, P. S.; Rogers, E. J. *Microbiol. Rev.* **1996**, *60*, 366–385.

(145) Tenson, T.; Ehrenberg, M. *Cell* **2002**, *108*, 591–594.

(146) Morris, D. R.; Geballe, A. P. *Mol. Cell. Biol.* **2000**, *20*, 8635–8642.

(147) Kimchi-Sarfaty, C.; Oh, J. M.; Kim, I.-W.; Sauna, Z. E.; Calcagno, A. M.; Ambudkar, S. V.; Gottesman, M. M. *Science* **2007**, *315*, 525–528.

(148) Thanaraj, T. A.; Argos, P. *Protein Sci. Publ. Protein Soc.* **1996**, *5*, 1594–1612.

(149) Tsai, C.-J.; Sauna, Z. E.; Kimchi-Sarfaty, C.; Ambudkar, S. V.; Gottesman, M. M.; Nussinov, R. *J. Mol. Biol.* **2008**, *383*, 281–291.

(150) Gong, F.; Yanofsky, C. *J. Biol. Chem.* **2001**, *276*, 1974–1983.

(151) Berisio, R.; Schluenzen, F.; Harms, J.; Bashan, A.; Auerbach, T.; Baram, D.; Yonath, A. *Nat. Struct. Biol.* **2003**, *10*, 366–370.

(152) Gabashvili, I. S.; Gregory, S. T.; Valle, M.; Grassucci, R.; Worbs, M.; Wahl, M. C.; Dahlberg, A. E.; Frank, J. *Mol. Cell* **2001**, *8*, 181–188.

(153) Mitra, K.; Schaffitzel, C.; Fabiola, F.; Chapman, M. S.; Ban, N.; Frank, J. *Mol. Cell* **2006**, *22*, 533–543.

(154) Garza-Sánchez, F.; Janssen, B. D.; Hayes, C. S. *J. Biol. Chem.* **2006**, *281*, 34258–34268.

(155) Tanner, D. R.; Dewey, J. D.; Miller, M. R.; Buskirk, A. R. *J. Biol. Chem.* **2006**, *281*, 10561–10566.

(156) Sunohara, T.; Jojima, K.; Tagami, H.; Inada, T.; Aiba, H. *J. Biol. Chem.* **2004**, *279*, 15368–15375.

(157) Nurizzo, D.; Shewry, S. C.; Perlin, M. H.; Brown, S. A.; Dholakia, J. N.; Fuchs, R. L.; Deva, T.; Baker, E. N.; Smith, C. A. *J. Mol. Biol.* **2003**, *327*, 491–506.

(158) Li, X.; Hirano, R.; Tagami, H.; Aiba, H. *RNA N. Y. N* **2006**, *12*, 248–255.

(159) Roche, E. D.; Sauer, R. T. *EMBO J.* **1999**, *18*, 4579–4589.

(160) Dong, H.; Nilsson, L.; Kurland, C. G. *J. Mol. Biol.* **1996**, *260*, 649–663.

(161) Craigen, W. J.; Caskey, C. T. *Nature* **1986**, *322*, 273–275.

(162) Poole, E. S.; Brown, C. M.; Tate, W. P. *EMBO J.* **1995**, *14*, 151–158.

(163) Gurvich, O. L.; Baranov, P. V.; Zhou, J.; Hammer, A. W.; Gesteland, R. F.; Atkins, J. F. *EMBO J.* **2003**, *22*, 5941–5950.

(164) Mottagui-Tabar, S.; Björnsson, A.; Isaksson, L. A. *EMBO J.* **1994**, *13*, 249–257.

(165) Pavlov, M. Y.; Watts, R. E.; Tan, Z.; Cornish, V. W.; Ehrenberg, M.; Forster, A. C. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 50–54.

(166) Doronina, V. A.; Wu, C.; de Felipe, P.; Sachs, M. S.; Ryan, M. D.; Brown, J. D. *Mol. Cell. Biol.* **2008**, *28*, 4227–4239.

(167) Beringer, M. *RNA N. Y. N* **2008**, *14*, 795–801.

(168) Collier, J.; Bohn, C.; Bouloc, P. *J. Biol. Chem.* **2004**, *279*, 54193–54201.

(169) Hentzen, D.; Mandel, P.; Garel, J. P. *Biochim. Biophys. Acta* **1972**, *281*, 228–232.

(170) Janzen, D. M.; Frolova, L.; Geballe, A. P. *Mol. Cell. Biol.* **2002**, *22*, 8562–8570.

(171) Miller, J. H. In *Cold Springs Harbor Laboratory Press*; Cold Springs Harbor: NY, 1972.

(172) Woolstenhulme, C. J.; Parajuli, S.; Healey, D. W.; Valverde, D. P.; Petersen, E. N.; Starosta, A. L.; Guydosh, N. R.; Johnson, W. E.; Wilson, D. N.; Buskirk, A. R. *Proc. Natl. Acad. Sci.* **2013**, *110*, E878–E887.

(173) Yanagitani, K.; Kimata, Y.; Kadokura, H.; Kohno, K. *Science* **2011**, *331*, 586–589.

(174) Komar, A. A. *Trends Biochem. Sci.* **2009**, *34*, 16–24.

(175) Zhang, G.; Hubalewska, M.; Ignatova, Z. *Nat. Struct. Mol. Biol.* **2009**, *16*, 274–280.

(176) Dove, S. L.; Joung, J. K.; Hochschild, A. *Nature* **1997**, *386*, 627–630.

(177) Joung, J. K.; Ramm, E. I.; Pabo, C. O. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 7382–7387.

(178) Levchenko, I.; Seidel, M.; Sauer, R. T.; Baker, T. A. *Science* **2000**, *289*, 2354–2356.

(179) Wah, D. A.; Levchenko, I.; Rieckhof, G. E.; Bolon, D. N.; Baker, T. A.; Sauer, R. T. *Mol. Cell* **2003**, *12*, 355–363.

(180) Sunohara, T.; Abo, T.; Inada, T.; Aiba, H. *RNA N. Y. N* **2002**, *8*, 1416–1427.

(181) Roche, E. D.; Sauer, R. T. *J. Biol. Chem.* **2001**, *276*, 28509–28515.

(182) Cochella, L.; Brunelle, J. L.; Green, R. *Nat. Struct. Mol. Biol.* **2007**, *14*, 30–36.

(183) Youngman, E. M.; Brunelle, J. L.; Kochaniak, A. B.; Green, R. *Cell* **2004**, *117*, 589–599.

(184) Doerfel, L. K.; Wohlgemuth, I.; Kothe, C.; Peske, F.; Urlaub, H.; Rodnina, M. V. *Science* **2013**, *339*, 85–88.

(185) Ude, S.; Lassak, J.; Starosta, A. L.; Kraxenberger, T.; Wilson, D. N.; Jung, K. *Science* **2013**, *339*, 82–85.

(186) Glick, B. R.; Ganoza, M. C. *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 4257–4260.

(187) Glick, B. R.; Chládek, S.; Ganoza, M. C. *Eur. J. Biochem. FEBS* **1979**, *97*, 23–28.

(188) Hanawa-Suetsugu, K.; Sekine, S.; Sakai, H.; Hori-Takemoto, C.; Terada, T.; Unzai, S.; Tame, J. R. H.; Kuramitsu, S.; Shirouzu, M.; Yokoyama, S. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 9595–9600.

(189) Blaha, G.; Stanley, R. E.; Steitz, T. A. *Science* **2009**, *325*, 966–970.

(190) Navarre, W. W.; Zou, S. B.; Roy, H.; Xie, J. L.; Savchenko, A.; Singer, A.; Edvokimova, E.; Prost, L. R.; Kumar, R.; Ibba, M.; Fang, F. C. *Mol. Cell* **2010**, *39*, 209–221.

(191) Yanagisawa, T.; Sumida, T.; Ishii, R.; Takemoto, C.; Yokoyama, S. *Nat. Struct. Mol. Biol.* **2010**, *17*, 1136–1143.

(192) Peil, L.; Starosta, A. L.; Virumäe, K.; Atkinson, G. C.; Tenson, T.; Remme, J.; Wilson, D. N. *Nat. Chem. Biol.* **2012**, *8*, 695–697.

(193) Li, G.-W.; Oh, E.; Weissman, J. S. *Nature* **2012**, *484*, 538–541.

(194) Ingolia, N. T.; Lareau, L. F.; Weissman, J. S. *Cell* **2011**, *147*, 789–802.

(195) Brunelle, J. L.; Youngman, E. M.; Sharma, D.; Green, R. *RNA N. Y. N* **2006**, *12*, 33–39.

(196) Shimizu, Y.; Inoue, A.; Tomari, Y.; Suzuki, T.; Yokogawa, T.; Nishikawa, K.; Ueda, T. *Nat. Biotechnol.* **2001**, *19*, 751–755.

(197) Blanchard, S. C.; Kim, H. D.; Gonzalez, R. L., Jr; Puglisi, J. D.; Chu, S. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12893–12898.

(198) Yokogawa, T.; Kitamura, Y.; Nakamura, D.; Ohno, S.; Nishikawa, K. *Nucleic Acids Res.* **2010**, *38*, e89.

(199) Walker, S. E.; Fredrick, K. *Methods San Diego Calif* **2008**, *44*, 81–86.

(200) Gutierrez, E.; Shin, B.-S.; Woolstenhulme, C. J.; Kim, J.-R.; Saini, P.; Buskirk, A. R.; Dever, T. E. *Mol. Cell* **2013**, *51*, 35–45.

(201) Benne, R.; Hershey, J. W. *J. Biol. Chem.* **1978**, *253*, 3078–3087.

(202) Kemper, W. M.; Berry, K. W.; Merrick, W. C. *J. Biol. Chem.* **1976**, *251*, 5551–5557.

(203) Schreier, M. H.; Erni, B.; Staehelin, T. *J. Mol. Biol.* **1977**, *116*, 727–753.

(204) Saini, P.; Eyler, D. E.; Green, R.; Dever, T. E. *Nature* **2009**, *459*, 118–121.

(205) Gregio, A. P. B.; Cano, V. P. S.; Avaca, J. S.; Valentini, S. R.; Zanelli, C. F. *Biochem. Biophys. Res. Commun.* **2009**, *380*, 785–790.

(206) Scuoppo, C.; Miething, C.; Lindqvist, L.; Reyes, J.; Ruse, C.; Appelmann, I.; Yoon, S.; Krasnitz, A.; Teruya-Feldstein, J.; Pappin, D.; Pelletier, J.; Lowe, S. W. *Nature* **2012**, *487*, 244–248.

(207) Silvera, D.; Formenti, S. C.; Schneider, R. J. *Nat. Rev. Cancer* **2010**, *10*, 254–266.

(208) Park, M. H. *J. Biol. Chem.* **1989**, *264*, 18531–18535.

(209) Park, M. H.; Wolff, E. C.; Smit-McBride, Z.; Hershey, J. W.; Folk, J. E. *J. Biol. Chem.* **1991**, *266*, 7988–7994.

(210) Roy, H.; Zou, S. B.; Bullwinkle, T. J.; Wolfe, B. S.; Gilreath, M. S.; Forsyth, C. J.; Navarre, W. W.; Ibba, M. *Nat. Chem. Biol.* **2011**, *7*, 667–669.

(211) Park, J.-H.; Johansson, H. E.; Aoki, H.; Huang, B. X.; Kim, H.-Y.; Ganoza, M. C.; Park, M. H. *J. Biol. Chem.* **2012**, *287*, 2579–2590.

(212) Ganoza, M. C.; Aoki, H. *Biol. Chem.* **2000**, *381*, 553–559.

(213) Zuk, D.; Jacobson, A. *EMBO J.* **1998**, *17*, 2914–2925.

(214) Letzring, D. P.; Dean, K. M.; Grayhack, E. J. *RNA N. Y. N* **2010**, *16*, 2516–2528.

(215) Gelperin, D. M.; White, M. A.; Wilkinson, M. L.; Kon, Y.; Kung, L. A.; Wise, K. J.; Lopez-Hoyo, N.; Jiang, L.; Piccirillo, S.; Yu, H.; Gerstein, M.; Dumont, M. E.; Phizicky, E. M.; Snyder, M.; Grayhack, E. J. *Genes Dev.* **2005**, *19*, 2816–2826.

(216) Eyler, D. E.; Green, R. *RNA N. Y. N* **2011**, *17*, 925–932.

(217) Ben-Shem, A.; Jenner, L.; Yusupova, G.; Yusupov, M. *Science* **2010**, *330*, 1203–1209.

(218) Basavappa, R.; Sigler, P. B. *EMBO J.* **1991**, *10*, 3105–3111.

(219) Anand, M.; Balar, B.; Ulloque, R.; Gross, S. R.; Kinzy, T. G. *J. Biol. Chem.* **2006**, *281*, 32318–32326.

(220) Triana-Alonso, F. J.; Chakraburtty, K.; Nierhaus, K. H. *J. Biol. Chem.* **1995**, *270*, 20473–20478.

(221) Pech, M.; Karim, Z.; Yamamoto, H.; Kitakawa, M.; Qin, Y.; Nierhaus, K. H. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 3199–3203.

(222) Driscoll, D. M.; Copeland, P. R. *Annu. Rev. Nutr.* **2003**, *23*, 17–40.

(223) Park, M. H.; Nishimura, K.; Zanelli, C. F.; Valentini, S. R. *Amino Acids* **2010**, *38*, 491–500.

(224) Frigieri, M. C.; Thompson, G. M.; Pandolfi, J. R.; Zanelli, C. F.; Valentini, S. R. *Genet. Mol. Res. GMR* **2007**, *6*, 152–165.

(225) Valentini, S. R.; Casolari, J. M.; Oliveira, C. C.; Silver, P. A.; McBride, A. E. *Genetics* **2002**, *160*, 393–405.

(226) Bullwinkle, T. J.; Zou, S. B.; Rajkovic, A.; Hersch, S. J.; Elgamal, S.; Robinson, N.; Smil, D.; Bolshan, Y.; Navarre, W. W.; Ibba, M. *J. Biol. Chem.* **2013**, *288*, 4416–4423.

(227) Ramachandran, G. N.; Mitra, A. K. *J. Mol. Biol.* **1976**, *107*, 85–92.

(228) Ingolia, N. T.; Ghaemmaghami, S.; Newman, J. R. S.; Weissman, J. S. *Science* **2009**, *324*, 218–223.

(229) Acker, M. G.; Kolitz, S. E.; Mitchell, S. F.; Nanda, J. S.; Lorsch, J. R. *Methods Enzymol.* **2007**, *430*, 111–145.

(230) Mitchell, S. F.; Walker, S. E.; Algire, M. A.; Park, E.-H.; Hinnebusch, A. G.; Lorsch, J. R. *Mol. Cell* **2010**, *39*, 950–962.

(231) Shin, B.-S.; Kim, J.-R.; Walker, S. E.; Dong, J.; Lorsch, J. R.; Dever, T. E. *Nat. Struct. Mol. Biol.* **2011**, *18*, 1227–1234.

(232) Shin, B.-S.; Dever, T. E. *Methods Enzymol.* **2007**, *429*, 185–201.

(233) Andersen, C. F.; Anand, M.; Boesen, T.; Van, L. B.; Kinzy, T. G.; Andersen, G. R. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 1304–1307.

(234) Ortiz, P. A.; Ulloque, R.; Kihara, G. K.; Zheng, H.; Kinzy, T. G. *J. Biol. Chem.* **2006**, *281*, 32639–32648.

(235) Shin, B.-S.; Acker, M. G.; Maag, D.; Kim, J.-R.; Lorsch, J. R.; Dever, T. E. *Mol. Cell. Biol.* **2007**, *27*, 1677–1685.

(236) Foiani, M.; Cigan, A. M.; Paddon, C. J.; Harashima, S.; Hinnebusch, A. G. *Mol. Cell. Biol.* **1991**, *11*, 3203–3216.

(237) Jørgensen, R.; Carr-Schmid, A.; Ortiz, P. A.; Kinzy, T. G.; Andersen, G. R. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, *58*, 712–715.

(238) Hatfield, L.; Beelman, C. A.; Stevens, A.; Parker, R. *Mol. Cell. Biol.* **1996**, *16*, 5830–5838.

(239) Mitchell, D. A.; Marshall, T. K.; Deschenes, R. J. *Yeast Chichester Engl.* **1993**, *9*, 715–722.

(240) Tan, S. *Protein Expr. Purif.* **2001**, *21*, 224–234.

(241) SternJohn, J.; Hati, S.; Siliciano, P. G.; Musier-Forsyth, K. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 2127–2132.