2018-06-25

Mechanism of premature translation termination on a sense codon

Egor Svidritskiy
University of Massachusetts Medical School

Et al.

JBC Papers in Press. Published on June 25, 2018 as Manuscript AW118.003232
The latest version is at http://www.jbc.org/cgi/doi/10.1074/jbc.AW118.003232

Mechanism of premature translation termination on a sense codon

Egor Svidritskiy[1], Gabriel Demo[1] and Andrei A. Korostelev[1,*]

[1]RNA Therapeutics Institute, University of Massachusetts Medical School, 368 Plantation St., Worcester, MA 01605, USA.

*Corresponding author. Correspondence: andrei.korostelev@umassmed.edu

Abstract

Accurate translation termination by release factors (RFs) is critical for the integrity of cellular proteomes. Premature termination on sense codons, for example, results in truncated proteins, whose accumulation could be detrimental to the cell. Nevertheless, some sense codons are prone to triggering premature termination, but the structural basis for this is unclear. To investigate premature termination, we determined a cryo-EM structure of the Escherichia coli 70S ribosome bound with RF1 in response to a UAU (Tyr) sense codon. The structure reveals that RF1 recognizes a UAU codon similarly to a UAG stop codon, suggesting that sense codons induce premature termination because they structurally mimic a stop codon. Hydrophobic interaction between the nucleobase of U3 (the third position of the UAU codon) and conserved Ile 196 in RF1 is important for misreading the UAU codon. Analyses of RNA binding in ribonucleoprotein complexes or by amino acids reveal that Ile–U packing is a frequent protein–RNA binding motif with key functional implications. We discuss parallels with eukaryotic translation termination by the release factor eRF1.

Translation termination defines the lengths of all cellular proteins. Three stop codons—UAA, UAG and UGA—signal the end of the mRNA open reading frame (ORF). A stop codon in the ribosomal A site is recognized by a bifunctional protein called a release factor (RF), which (i) recognizes stop codons and discriminates against sense codons, and (ii) catalyzes peptidyl-tRNA hydrolysis, releasing the peptide from the ribosome. Bacteria express two release factors: RF1 recognizes UAA/UAG codons and RF2 recognizes UAA/UGA codons. A single eukaryotic release factor eRF1 recognizes all three stop codons.

Accurate termination on stop codons is crucial for the cell and organism. Premature termination, either on sense codons or on premature stop codons arising from nonsense mutations, would result in accumulation of truncated proteins with compromised or toxic activities. Many genetic diseases are caused by premature stop codons (1,2), highlighting the deleterious effects of premature termination. In an intact ORF (i.e., no premature stop codon), however, near-stop codons may cause premature termination (3). Sense codons differing from the stop codon in the third-nucleotide (wobble) position are most promiscuous (4). A study measuring the activity of bacterial release factors on near-stop codons identified the UAU sense codon as a “hot spot” for RF1 (4). We recently confirmed these findings (5), showing that RF1 can catalyze release on a UAU sense codon almost as efficiently as on a UAA codon when RF1 is in excess (Fig. 1A). Affinity of release factors to some sense codons does not lead to predominant premature termination *in vivo* due to efficient decoding of sense codons by aminoacyl-tRNAs, but is likely responsible for background levels of premature termination (3).

The structural mechanism of stop-codon recognition was elucidated by high-resolution crystal structures (6-9) and further investigated by molecular-dynamics

simulations (10). Release factors strongly discriminate against purines in the first position, so that no product can be detected when the ribosome encounters sense codons with A1 or G1 (i.e., A or G in position 1 of the A-site codon; Fig. 1A) (4,11,12). The structures showed that the Watson-Crick side of the first nucleotide faces the backbone of an α-helix of RF1 or RF2, so the strict requirement for pyrimidine in the first position is explained by base recognition by a rigid structural element of release factor. The second and third positions are limited to purines, and interact with side chains of RF1 or RF2 (reviewed in refs (13-17)). The third nucleotide is also sandwiched between the conserved 16S rRNA nucleotide G530 and a side chain of release factor: Ile 196 of RF1 or Arg 218 of RF2 (*E. coli* numbering, unless noted otherwise). The Hoogsteen side of the third nucleotide (at atoms N6/O6 and N7) is stabilized by a conserved threonine (Thr 194 in RF1 or Thr 216 in RF2). In the eukaryotic 80S termination complex formed with eRF1, the third position of the stop codon similarly packs between a purine and Ile 62 (*Homo sapiens*) of the essential NIKS motif of eRF1 (18), and interacts with Thr 58 (19,20).

In this work, we asked how a UAU sense codon is misread by RF1 as a stop codon, and why U in the third position is preferred over C, rendering UAU a more efficient "mis-terminator" than UAC. We determined a cryo-EM structure of the bacterial 70S ribosome complex that helps answer these questions. Our analyses suggest that the packing of U3 against Ile 196 of RF1 is critical for the preference of U over C. We find that the Ile-U packing is a prevalent motif in the ribosome and other protein-RNA structures, consistent with the role of Ile and other non-aromatic hydrophobic side chains in specific recognition of uridine.

Mechanism of termination on the UAU sense codon

We determined a cryo-EM structure of the *E. coli* 70S•RF1 complex formed in response to the UAU codon in the A site at 3.7 Å average resolution (Fig. 1B and Table 1), with local resolution achieving ~3 Å in ribosomal functional centers (Fig. 1C), allowing near-atomic interpretation (Figs. 1D and 1E). The overall conformations of release factor and the UAU codon are similar to those in canonical termination complexes formed with stop codons. We note one main difference between our structure and previous high-resolution 70S•RF1 structures from *Thermus thermophilus* and heterologous systems. In our *E. coli* structure, domain 1 of RF1 binds both the large-subunit ribosomal protein L11 and 23S ribosomal RNA of the L11 stalk, consistent with observations at lower resolutions for *E. coli* 70S•RF1 complexes formed on a stop codon (21). Specifically, hydrophobic Ile 29 of RF1 docks at the proline-rich region of L11 (aa Pro 21 to Pro 25) and Phe 35 binds near A1095 of 23S rRNA, similar to that seen for RF2 (6,8,22,23). In other high-resolution structures, however, domain 1 of RF1 is either unresolved (11,24-26) or binds near the L11 stalk without contacting it (7,9,23). In our structure, domain 1 is not as well resolved as other domains of RF1 (Fig. 1C), consistent with the idea that it is dynamic. The different positions of domain 1 in different structures highlight that interactions between domain 1 and the ribosome are dynamic and are likely important at early

stages of RF1 binding or during RF1 dissociation, in keeping with functional interaction between RF1 and L11 (27-29).

Density in the peptidyl-transferase and decoding centers (Fig. 1D and 1E) showed interactions similar to those found in canonical termination complexes with RF1 (7,9). Consistent with the catalytic activity of RF1 on the UAU codon, the catalytic ^{233}GGQ^{235} loop is positioned next to the terminal nucleotide A76 of the P-site tRNA and is stabilized by interactions with A2602, which is essential for termination (Fig. 1D) (30-32). In the decoding center, the first two nucleotides of the UAU codon interact with the codon-recognition residues of RF1 similarly to stop-codon nucleotides. The sense-codon U3 residue is sandwiched between Ile 194 of RF1 and G530 of 16S rRNA. Its carbonyl oxygen O4 faces the hydrogen-bond donors OH of Thr 196 and NH of Gln 185 (Fig. 1E and 2A), similar to the carbonyl oxygen of G3 in the UAG stop codon (Fig. 2B) (9). However, whereas Thr 196 can form a hydrogen bond with one of two possible hydrogen-bond acceptors within G3 (at N7 and O6), the threonine can interact with a single acceptor within U3 of the UAU sense codon. Moreover, the sense codon U3-G530 pyrimidine-purine stacking is expected to be less stable than the stop codon G3-G530 purine-purine stacking (33-36). Thus, the interactions between RF1 and UAU sense codon resemble those between RF1 and UAG stop codon, and suggest why RF1 recognizes UAU less efficiently than UAG (Fig. 1A).

Unlike a UAU codon, a UAC sense codon inefficiently triggers peptide release, similar to sense codons that differ from stop codon at their second position, e.g., UCA or UGG (4). This might appear surprising, since the hydrogen-bonding valences of the amino group N4 of cytosine, placed similarly to the carbonyl oxygen O4 of uracil, could be satisfied by interactions with Thr196 (its OH group becoming the H-bond acceptor) and the amide group of Gln 185 (9) or with an ordered water molecule (10). In fact, just as U3 in UAU mimics a G3 in UAG stop codon, a C3 in UAC could mimic A3 in UAA stop codon. We propose that difference in hydrophobicity between the cytidine and uridine make termination on UAC less efficient than on UAU. Since C is substantially less hydrophobic than U (37-39), packing of C between Ile 169 and G530 would be less energetically favorable than the packing of U. Base-stacking energy of cytosine on guanosine is similar to that of uracil on guanosine (33-36), suggesting that the major discrimination between U and C in position 3 results from favorable hydrophobic packing of Ile 169 on U rather than on the less hydrophobic C.

In summary, our structure shows that recognition of the sense codon UAU by RF1 is similar to that of stop codon UAG. This suggests that other hot-spot sense codons likely undergo similar conformational rearrangements and are recognized by release factors similarly to stop codons, with which they have partial stereochemical resemblance. Structural analysis points at Ile-U packing being critical for making UAU a hot spot for mis-termination by RF1.

Ile-U interactions in protein-RNA complexes

The role of Ile-U packing in mis-termination by RF1 prompted us to investigate whether Ile-U is a common interaction employed in protein-RNA recognition. Nucleotide stacking is the major stabilizing interaction in secondary and tertiary structures of nucleic acids. In protein-RNA complexes, the energy of unstacking a nucleotide from its stacking partner(s) is usually compensated by interaction between the aromatic base of the nucleotide and protein side chain(s). The best-characterized interactions include stacking of nucleotides on aromatic side chains and on positively charged side chains ((40) and references therein). Because stacking interactions involve a hydrophobic energy contribution (non-polar, non-electrostatic, solvent entropy) (34,41,42), nucleotides also stack on aliphatic hydrophobic side chains. Isoleucine is the most hydrophobic side chain, according to many hydrophobicity scales (43-46), and the uracil and adenine nucleobases are more hydrophobic than cytosine and guanine (37-39). Thus, uridine and adenosine are more likely than cytidine and guanosine to interact with isoleucine and other aliphatic side chains. Indeed, in computational simulations, U was the only nucleotide with a favorable free energy of binding to Ile (i.e., negative $\Delta\Delta G$) in methanol, which is thought to represent the environment for nucleic acid-protein interfaces more accurately than water (47).

To test whether packing of isoleucine on uracil is among preferred protein-RNA interactions, we calculated the number of stacking interactions between RNA nucleotides and aliphatic, aromatic, or charged protein side chains in high-resolution crystal structures of ribosomes, including the 2.4-Å resolution *E.coli* 70S ribosome (48), 2.5-Å resolution *T. thermophilus* 70S ribosome (49) and 3.0-Å resolution *S. cerevisiae* 80S ribosome (50). Collectively, these structures provide a large pool of protein-RNA interactions comprising ~29,000 amino acids and ~19,000 nucleotides. We normalized each type of stacking to the number of nucleotides of each type (i.e., number of amino acids per thousand nucleotides).

These data show that Ile does indeed prefer to pack on uridine (Fig. 3). Ile-A packing is also well represented, whereas Ile packed the least on less hydrophobic C and G nucleotides. As expected, most of packing interactions of nucleotides occur with the positively charged Arg. Aliphatic, aromatic, and positively charged side chains pack more frequently on U or A (collectively > 50%) than on the less hydrophobic C or G. This preference is notable for Ile (72%), Pro (85%), Phe (70%), and Tyr (65%). Negatively charged Asp and Glu are the least represented among protein side chains packing on aromatic bases, as expected. Interestingly, however, 70S rescue complexes formed with ArfA and RF2 on truncated mRNAs, employ the packing of Glu 30 of ArfA on G530 of 16S rRNA (51-55), suggesting functional roles for this underrepresented group of interactions. The carboxyl group of Glu 30 is stabilized by interactions with the side chain of Arg 213 of RF2 and with 2′-OH of G530.

The specific affinity between U and Ile is also emphasized by amino acid-RNA affinity selection experiments, in which isoleucine selectively bound to UAU-containing RNA motifs (56,57). This specificity plays key functional roles in protein-RNA complexes. For example, isoleucine-U interaction is

critical for specific and efficient recognition of bovine immunodeficiency virus transactivation response element (TAR) by the Tat protein (58). Crystal structure revealed that the side chain of Ile 79 of Tat packs on the aromatic ring of U10 and stabilizes the U10–A13:U24 base triplet (59). In *Drosophila*, Sxl regulates alternative splicing by specific recognition of a U-rich sequence in pre-mRNA, involving a uridine sandwiched between two isoleucines (Ile-U-Ile) (60). In archaeal RNase P, protein Rpp38 provides Ile 63 to stabilize bulged U19 of the enzyme's RNA (61,62).

Implications for eukaryotic termination

Structural and biochemical work has yielded detailed insights into the mechanism of eukaryotic termination, but the mechanism of premature termination on sense codons in eukaryotes remains poorly understood. The UGG codon binds eRF1 (63-65), but termination activity was not detected *in vitro* (66). The sequences and structures of codon-recognition domains of eukaryotic and bacterial release factors are very different. Perhaps not surprisingly, therefore, structural recognition of stop codons by eukaryotic eRF1 also differs from that by bacterial release factors. For example, eRF1 recognizes the U-turn-like geometry of the stop codon (19,20,67) and a nucleotide downstream of the stop codon (68). However, recognition of the third nucleotide by eRF1 is remarkably similar to that by RF1. In the structures of 80S-eRF1 complexes (19,20,67), the third nucleotide is sandwiched between a purine (second base of stop codon) and Ile 62 of the universally conserved NIKS motif of eRF1 (Fig. 2C). Furthermore, the O6 atom of the guanosine of the UAG stop codon (19) hydrogen bonds with Thr 58. The similar structural environment of the third nucleotide in the eukaryotic and bacterial complexes suggests that UAU might also be a hot spot for termination in eukaryotes. Premature termination on UAU, however, is likely less pronounced than on bacterial ribosomes, due to stringent codon discrimination facilitated by GTPase eRF3 (69,70).

MATERIALS AND METHODS

Preparation of the 70S•mRNA(UAU)•tRNA^{fMet}•RF1 complex

C-terminally His-tagged *E. coli* RF1 was purified as described (5). 70S ribosomes were prepared from *E. coli* (MRE600) as described (5), and stored in the ribosome-storage buffer (20 mM Tris-HCl (pH 7.0), 100 mM NH_4Cl, 12.5 mM $MgCl_2$, 0.5 mM EDTA, 6 mM βME) at -80°C. Ribosomal 30S and 50S subunits were purified using sucrose gradient (10-35%) in a ribosome-dissociation buffer (20 mM Tris-HCl (pH 7.0), 300 mM NH_4Cl, 1.5 mM $MgCl_2$, 0.5mM EDTA, 6 mM βME). The fractions containing 30S and 50S subunits were collected separately, concentrated and stored in the ribosome-storage buffer at -80°C. *E. coli* tRNA^{fMet} was purchased from Chemical Block. RNA, containing the Shine-Dalgarno sequence and a linker to place the AUG codon in the P site and the UAU codon in the A site (GGC AAG GAG GUA AAA AUG UAU AAAAAA) was synthesized by IDT.

The 70S•mRNA•tRNA^{fMet}•RF1 complex was prepared by reconstitution *in vitro*. 1 μM 30S subunit (all concentrations are specified for the final solution) were pre-activated at 42°C for 5 minutes in the ribosome-reconstitution buffer (20 mM Tris-

HCl (pH 7.0), 100 mM NH_4Cl, 20.5 mM $MgCl_2$, 0.5 mM EDTA, 6 mM βME). After pre-activation, 0.9 μM 50S subunit with 12 μM mRNA and 5 μM $tRNA^{fMet}$ were added to the 30S solution and incubated for 15 minutes at 37°C. Equal volume of 40 μM RF1 was then added resulting in the following final concentrations: ~0.45 μM 70S, 6 μM mRNA, 2.5 μM $tRNA^{fMet}$ and 20 μM RF1. The solution was incubated for 15 minutes at 37°C and applied on cryo-EM grids at room temperature.

Cryo-EM and image processing

Holey-carbon grids (C-flat 1.2-1.3, Protochips) were glow discharged with 20 mA with negative polarity for 30 seconds in a PELCO easiGlow glow discharge unit. 1.5 μl of the 70S•mRNA•$tRNA^{fMet}$•RF1 complex was applied to the grids. The grids were blotted for 3.5 s at blotting power 8 at 4°C and ~95% humidity and plunged into liquid ethane, using an FEI Vitrobot MK4. The grids were stored in liquid nitrogen.

A dataset of 1,065,147 particles was collected as follows. 3,963 movies were collected using SerialEM (71) on a Talos Arctica (FEI) microscope operating at 200 kV equipped with a K2 Summit camera system (Gatan) with -0.7 to -1.7 μm defocus. Each exposure was acquired with continuous frame streaming with the exposure length of 80 frames per movie yielding a total dose of 37.7 $e^-/Å^2$. The nominal magnification was 22,000 and the corrected super-resolution pixel size at the specimen level was 0.944 Å. The frames for each movie were processed using IMOD (72). The movies were motion-corrected and frame averages were calculated using frames 3 to 42 within each movie, using alignframes (IMOD), after multiplication with the corresponding gain reference. cisTEM (73) was used to determine defocus values for each resulting frame average and for particle picking. The stack and FREALIGN parameter file were assembled in via cisTEM with the binning of 1x, 3x and 6x (box size of 480 for a non-binned stack).

Data processing was performed essentially as described (74). FrealignX v9.11 in FrealignX mode was used for all steps of refinement and reconstruction (75). The 6x-binned image stack (1,065,147 particles) was initially aligned to a ribosome reference (PDB 5J4D: (11)) without RF1 and E-tRNA, using 3 cycles of mode 3 (global search) alignment, including data in the resolution range from 30 Å to 300 Å. Subsequently, the 6x binned stack was refined using mode 1 (refine) in the resolution ranges (sequentially): 30-300, 24-300, 18-300 and 15-300 Å (3 cycles for each range). Using the 3x binned image stack, the particles were successively aligned in mode 1 (refine) by gradually increasing the high resolution limit to 12, 10, 9, 8 and 7 Å (3 cycles for each resolution limit). In the last step, the unbinned (full-resolution) image stack was used to successively align particles against the common reference using mode 1 (refine; 3 cycles) at the resolution limit of 6 Å. 3D density reconstruction was obtained using 60% of particles with highest scores. The map contained density for the P- and E-site tRNAs, mRNA and RF1. The resolution of the resulting reconstruction was ~3.7 Å (Fourier Shell Correlation (FSC) = 0.143); local resolution for the codon-recognition domain of RF1 in the decoding center and the catalytic domain in the peptidyl-transferase center achieves ~3 Å resolution, allowing near-atomic-resolution

interpretation of nucleotide and side-chain interactions (Figs. 1C-E). Additional classification into 16 classes yielded 3 classes (87% of all particles) with the occupancy of RF1 similar to that in the initial map. The initial reconstruction was B-factor sharpened using B-factors of -150, -200 and -225 Å^2 in bfactor.exe (part of the FrealignX distribution) and used for model building and structure refinements. B-factor of -100 Å^2 was also used to visualize lower-resolution details. FSC curve was calculated by FrealignX for even and odd particle half-sets (Fig. 1B). Blocres was used to assess local resolution of the unfiltered and unmasked volume using a box size of 56 pixels, step size of 10 pixels, and resolution criterion of FSC value at 0.143 (76).

Model building and refinement

Recently reported cryo-EM structure of *E. coli* 70S•ArfA•RF2 complex (55), excluding ArfA and RF2, was used as a starting model for structure refinement. The initial model of *E. coli* RF1 (domains 2-4) was extracted from the crystal structure of the 70S•RF1 complex (11) and domain 1 was obtained by homology modeling from *Thermus thermophilus* RF1 (9) using SWISS-PROT (77). Initial protein and ribosome domain fitting into cryo-EM maps was performed using Chimera (78), followed by manual modeling using PyMOL (79). The linker between domain 1 and domain 2 (aa 99-105) was not defined in the cryo-EM map and was not modeled.

The structural model was refined by real-space simulated-annealing refinement using atomic electron scattering factors in RSRef (80,81) as described (82). Secondary-structure restraints, comprising hydrogen-bonding restraints for ribosomal proteins and base-pairing (distance and co-planarity) restraints for RNA nucleotides, were implemented in CNS format (83). Refinement parameters in RSRef, such as the relative weighting of stereochemical restraints and experimental energy term, were optimized to produce the stereochemically optimal models that closely agree with the corresponding maps. Refinement was performed using B-factor sharpened maps: all-atom refinement (-200 Å^2), then local refinement of P-tRNA, mRNA, RF1 and neighboring residues (-225 Å^2) at starting annealing temperature 1000 K. In the final stages, the structures were refined using phenix.real_space_refine (84) against a B-sharpened map (-150 Å^2) at 300 K, followed by a round of refinement in RSRef applying harmonic restraints to preserve protein backbone geometry (-150 Å^2 and 300 K). The refined structural model closely agrees with the corresponding maps, as indicated by low real-space R-factor of ~0.19 (RSRef) and high correlation coefficient of 0.86 (PHENIX, CC around atoms). The resulting models have good stereochemical parameters, characterized by low deviation from ideal bond lengths and angles, low number of protein-backbone and rotamer outliers, as shown in Table 1.

Structure superpositions and comparisons were performed in PyMOL.

Structural analyses of protein side chain interactions with RNA nucleotides

The following ribosome structures were downloaded from RCSB (www.rcsb.org) for the analyses of amino-acid-nucleotide interactions: *E. coli* 70S ribosome (PDB 4YBB (48)), *T. thermophilus* 70S ribosome

(PDB 4Y4P: (49) and *S. cerevisiae* 80S ribosome (PDB 4V88: (50). PyMOL was used to calculate the number of side chains packed on RNA nucleotides (#aa per 1000 nucleotides, Fig. 3). The following distance cutoff criterion was used: at least one side-chain non-hydrogen atom (i.e. any atom excluding the backbone atoms) within 3.7 Å from the following carbon atoms of the aromatic base of a nucleotide: C2, C4 or C5 (U or C), C4, C5 or C6 (A) and C2, C4, C5 or C6 (G). Selected amino acids were confirmed by visual inspection of the PDB structures in PyMOL, supporting the stringency of the selection criterion.

Structure Accession Codes

The cryo-EM map and PDB coordinates have been deposited in EMDB and the Protein Data Bank with accession codes EMD-7970 and 6DNC.

Acknowledgements

We thank Chen Xu for help with screening cryo-EM grids and for assistance with data collection at the UMass Medical School cryo-EM facility; Anna B. Loveland for assistance with data processing; Darryl Conte Jr. for helpful comments on the manuscript. This study was supported by NIH Grant R01 GM107465.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

Author contributions

E.S. assembled the ribosome complex, collected and processed cryo-EM data; G.D. purified ribosome subunits and RF1, and assisted with data processing; A.A.K modeled and refined the structure, and wrote the manuscript. All authors finalized the manuscript.

References

1. Frischmeyer, P. A., and Dietz, H. C. (1999) Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet* **8**, 1893-1900
2. Keeling, K. M., and Bedwell, D. M. (2011) Suppression of nonsense mutations as a therapeutic approach to treat genetic diseases. *Wiley Interdiscip Rev RNA* **2**, 837-852
3. Jorgensen, F., Adamski, F. M., Tate, W. P., and Kurland, C. G. (1993) Release factor-dependent false stops are infrequent in Escherichia coli. *J Mol Biol* **230**, 41-50
4. Freistroffer, D. V., Kwiatkowski, M., Buckingham, R. H., and Ehrenberg, M. (2000) The accuracy of codon recognition by polypeptide release factors. *Proc Natl Acad Sci U S A* **97**, 2046-2051
5. Svidritskiy, E., and Korostelev, A. A. (2018) Conformational control of translation termination on the 70S ribosome. *Structure* **26**, 821-828
6. Weixlbaumer, A., Jin, H., Neubauer, C., Voorhees, R. M., Petry, S., Kelley, A. C., and Ramakrishnan, V. (2008) Insights into translational termination from the structure of RF2 bound to the ribosome. *Science* **322**, 953-956
7. Laurberg, M., Asahara, H., Korostelev, A., Zhu, J., Trakhanov, S., and Noller, H. F. (2008) Structural basis for translation termination on the 70S ribosome. *Nature* **454**, 852-857
8. Korostelev, A., Asahara, H., Lancaster, L., Laurberg, M., Hirschi, A., Zhu, J.,

Trakhanov, S., Scott, W. G., and Noller, H. F. (2008) Crystal structure of a translation termination complex formed with release factor RF2. *Proc Natl Acad Sci U S A* **105**, 19684-19689
9. Korostelev, A., Zhu, J., Asahara, H., and Noller, H. F. (2010) Recognition of the amber UAG stop codon by release factor RF1. *EMBO J* **29**, 2577-2585
10. Sund, J., Ander, M., and Aqvist, J. (2010) Principles of stop-codon reading on the ribosome. *Nature* **465**, 947-950
11. Svidritskiy, E., Madireddy, R., and Korostelev, A. A. (2016) Structural Basis for Translation Termination on a Pseudouridylated Stop Codon. *J Mol Biol* **428**, 2228-2236
12. Korkmaz, G., and Sanyal, S. (2017) R213I mutation in release factor 2 (RF2) is one step forward for engineering an omnipotent release factor in bacteria Escherichia coli. *J Biol Chem* **292**, 15134-15142
13. Korostelev, A. A. (2011) Structural aspects of translation termination on the ribosome. *RNA* **17**, 1409-1421
14. Dunkle, J. A., and Cate, J. H. (2010) Ribosome structure and dynamics during translocation and termination. *Annu Rev Biophys* **39**, 227-244
15. Zaher, H. S., and Green, R. (2009) Fidelity at the molecular level: lessons from protein synthesis. *Cell* **136**, 746-762
16. Nakamura, Y., and Ito, K. (2011) tRNA mimicry in translation termination and beyond. *Wiley Interdiscip Rev RNA* **2**, 647-668
17. Ramakrishnan, V. (2011) Structural studies on decoding, termination and translocation in the bacterial ribosome. in *Ribosomes: Structure, Function, and Dynamics* (Rodnina, M. V., Wintermeyer, W., and Green, R. eds.), Springer Vienna, Vienna. pp 19-30
18. Frolova, L., Seit-Nebi, A., and Kisselev, L. (2002) Highly conserved NIKS tetrapeptide is functionally essential in eukaryotic translation termination factor eRF1. *RNA* **8**, 129-136
19. Brown, A., Shao, S., Murray, J., Hegde, R. S., and Ramakrishnan, V. (2015) Structural basis for stop codon recognition in eukaryotes. *Nature* **524**, 493-496
20. Matheisl, S., Berninghausen, O., Becker, T., and Beckmann, R. (2015) Structure of a human translation termination complex. *Nucleic Acids Res* **43**, 8615-8626
21. Rawat, U., Gao, H., Zavialov, A., Gursky, R., Ehrenberg, M., and Frank, J. (2006) Interactions of the release factor RF1 with the ribosome as revealed by cryo-EM. *J Mol Biol* **357**, 1144-1153
22. Jin, H., Kelley, A. C., Loakes, D., and Ramakrishnan, V. (2010) Structure of the 70S ribosome bound to release factor 2 and a substrate analog provides insights into catalysis of peptide release. *Proc Natl Acad Sci U S A* **107**, 8593-8598
23. Petry, S., Brodersen, D. E., Murphy, F. V. t., Dunham, C. M., Selmer, M., Tarry, M. J., Kelley, A. C., and Ramakrishnan, V. (2005) Crystal structures of the ribosome in complex with release factors RF1 and RF2 bound to a cognate stop codon. *Cell* **123**, 1255-1266
24. Svidritskiy, E., and Korostelev, A. A. (2018) Mechanism of Inhibition of Translation Termination by Blasticidin S. *J Mol Biol* **430**, 591-593
25. Pierson, W. E., Hoffer, E. D., Keedy, H. E., Simms, C. L., Dunham, C. M., and Zaher, H. S. (2016) Uniformity of Peptide Release Is Maintained by Methylation of Release Factors. *Cell Rep* **17**, 11-18
26. Florin, T., Maracci, C., Graf, M., Karki, P., Klepacki, D., Berninghausen, O., Beckmann, R., Vazquez-Laslop, N., Wilson, D. N., Rodnina, M. V., and Mankin, A. S. (2017) An antimicrobial peptide that inhibits translation by trapping release factors on the

ribosome. *Nat Struct Mol Biol* **24**, 752-757
27. Van Dyke, N., and Murgola, E. J. (2003) Site of functional interaction of release factor 1 with the ribosome. *J Mol Biol* **330**, 9-13
28. Tate, W. P., Dognin, M. J., Noah, M., Stoffler-Meilicke, M., and Stoffler, G. (1984) The NH2-terminal domain of Escherichia coli ribosomal protein L11. Its three-dimensional location and its role in the binding of release factors 1 and 2. *J Biol Chem* **259**, 7317-7324
29. Bouakaz, L., Bouakaz, E., Murgola, E. J., Ehrenberg, M., and Sanyal, S. (2006) The role of ribosomal protein L11 in class I release factor-mediated translation termination and translational accuracy. *J Biol Chem* **281**, 4548-4556
30. Polacek, N., Gomez, M. J., Ito, K., Xiong, L., Nakamura, Y., and Mankin, A. (2003) The critical role of the universally conserved A2602 of 23S ribosomal RNA in the release of the nascent peptide during translation termination. *Mol Cell* **11**, 103-112
31. Youngman, E. M., Brunelle, J. L., Kochaniak, A. B., and Green, R. (2004) The active site of the ribosome is composed of two layers of conserved nucleotides with distinct roles in peptide bond formation and peptide release. *Cell* **117**, 589-599
32. Amort, M., Wotzel, B., Bakowska-Zywicka, K., Erlacher, M. D., Micura, R., and Polacek, N. (2007) An intact ribose moiety at A2602 of 23S rRNA is key to trigger peptidyl-tRNA hydrolysis during translation termination. *Nucleic Acids Res* **35**, 5130-5140
33. Jafilan, S., Klein, L., Hyun, C., and Florian, J. (2012) Intramolecular base stacking of dinucleoside monophosphate anions in aqueous solution. *J Phys Chem B* **116**, 3613-3618
34. Friedman, R. A., and Honig, B. (1995) A free energy analysis of nucleic acid base stacking in aqueous solution. *Biophys J* **69**, 1528-1535
35. Brown, R. F., Andrews, C. T., and Elcock, A. H. (2015) Stacking Free Energies of All DNA and RNA Nucleoside Pairs and Dinucleoside-Monophosphates Computed Using Recently Revised AMBER Parameters and Compared with Experiment. *J Chem Theory Comput* **11**, 2315-2328
36. Florián, J., Šponer, J., and Warshel, A. (1999) Thermodynamic Parameters for Stacking and Hydrogen Bonding of Nucleic Acid Bases in Aqueous Solution: Ab Initio/Langevin Dipoles Study. *J. Phys. Chem.* **103**, 884–892
37. Tsukada, K., Ueda, S., and Okada, R. (1984) Preparation of Liposome-Encapsulated Anti-tumor Drugs ; Relationship between Lipophilicity of Drugs and in Vitro Drug Release. *Chem. Pharm. Bull* **32**, 1929–1935
38. Cullis, P. M., and Wolfenden, R. (1981) Affinities of nucleic acid bases for solvent water. *Biochemistry* **20**, 3024-3028
39. Shih, P., Pedersen, L. G., Gibbs, P. R., and Wolfenden, R. (1998) Hydrophobicities of the nucleic acid bases: distribution coefficients from water to cyclohexane. *J Mol Biol* **280**, 421-430
40. Wilson, K. A., Holland, D. J., and Wetmore, S. D. (2016) Topology of RNA-protein nucleobase-amino acid pi-pi interactions and comparison to analogous DNA-protein pi-pi contacts. *RNA* **22**, 696-708
41. Norberg, J., and Nilsson, L. (1998) Solvent influence on base stacking. *Biophys J* **74**, 394-402
42. Luo, R., Gilson, H. S., Potter, M. J., and Gilson, M. K. (2001) The physical basis of nucleic acid base stacking in water. *Biophys J* **80**, 140-148
43. Zhu, C., Gao, Y., Li, H., Meng, S., Li, L., Francisco, J. S., and Zeng, X. C. (2016) Characterizing hydrophobicity of amino

acid side chains in a protein environment via measuring contact angle of a water nanodroplet on planar peptide network. *Proc Natl Acad Sci U S A* **113**, 12946-12951
44. Woese, C. R., Dugre, D. H., Saxinger, W. C., and Dugre, S. A. (1966) The molecular basis for the genetic code. *Proc Natl Acad Sci U S A* **55**, 966-974
45. Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydropathic character of a protein. *J Mol Biol* **157**, 105-132
46. Hessa, T., Kim, H., Bihlmaier, K., Lundin, C., Boekel, J., Andersson, H., Nilsson, I., White, S. H., and von Heijne, G. (2005) Recognition of transmembrane helices by the endoplasmic reticulum translocon. *Nature* **433**, 377-381
47. de Ruiter, A., and Zagrovic, B. (2015) Absolute binding-free energies between standard RNA/DNA nucleobases and amino-acid sidechain analogs in different environments. *Nucleic Acids Res* **43**, 708-718
48. Noeske, J., Wasserman, M. R., Terry, D. S., Altman, R. B., Blanchard, S. C., and Cate, J. H. (2015) High-resolution structure of the Escherichia coli ribosome. *Nat Struct Mol Biol* **22**, 336-341
49. Polikanov, Y. S., Melnikov, S. V., Soll, D., and Steitz, T. A. (2015) Structural insights into the role of rRNA modifications in protein synthesis and ribosome assembly. *Nat Struct Mol Biol* **22**, 342-344
50. Ben-Shem, A., Garreau de Loubresse, N., Melnikov, S., Jenner, L., Yusupova, G., and Yusupov, M. (2011) The structure of the eukaryotic ribosome at 3.0 A resolution. *Science* **334**, 1524-1529
51. James, N. R., Brown, A., Gordiyenko, Y., and Ramakrishnan, V. (2016) Translational termination without a stop codon. *Science* **354**, 1437-1440
52. Zeng, F., Chen, Y., Remis, J., Shekhar, M., Phillips, J. C., Tajkhorshid, E., and Jin, H. (2017) Structural basis of co-translational quality control by ArfA and RF2 bound to ribosome. *Nature* **541**, 554-557
53. Ma, C., Kurita, D., Li, N., Chen, Y., Himeno, H., and Gao, N. (2017) Mechanistic insights into the alternative translation termination by ArfA and RF2. *Nature* **541**, 550-553
54. Huter, P., Muller, C., Beckert, B., Arenz, S., Berninghausen, O., Beckmann, R., and Wilson, D. N. (2017) Structural basis for ArfA-RF2-mediated translation termination on mRNAs lacking stop codons. *Nature* **541**, 546-549
55. Demo, G., Svidritskiy, E., Madireddy, R., Diaz-Avalos, R., Grant, T., Grigorieff, N., Sousa, D., and Korostelev, A. A. (2017) Mechanism of ribosome rescue by ArfA and RF2. *Elife* **6**
56. Lozupone, C., Changayil, S., Majerfeld, I., and Yarus, M. (2003) Selection of the simplest RNA that binds isoleucine. *RNA* **9**, 1315-1322
57. Majerfeld, I., and Yarus, M. (1998) Isoleucine:RNA sites with associated coding sequences. *RNA* **4**, 471-478
58. Chen, L., and Frankel, A. D. (1995) A peptide interaction in the major groove of RNA resembles protein interactions in the minor groove of DNA. *Proc Natl Acad Sci U S A* **92**, 5077-5081
59. Moras, D., and Poterszman, A. (1996) Getting into the major groove. Protein-RNA interactions. *Curr Biol* **6**, 530-532
60. Handa, N., Nureki, O., Kurimoto, K., Kim, I., Sakamoto, H., Shimura, Y., Muto, Y., and Yokoyama, S. (1999) Structural basis for recognition of the tra mRNA precursor by the Sex-lethal protein. *Nature* **398**, 579-585
61. Oshima, K., Gao, X., Hayashi, S., Ueda, T., Nakashima, T., and Kimura, M. (2018) Crystal structures of the archaeal RNase P protein Rpp38 in complex with RNA fragments containing a K-turn

motif. *Acta Crystallogr F Struct Biol Commun* **74**, 57-64
62. Oshima, K., Kakiuchi, Y., Tanaka, Y., Ueda, T., Nakashima, T., Kimura, M., and Yao, M. (2016) Structural basis for recognition of a kink-turn motif by an archaeal homologue of human RNase P protein Rpp38. *Biochem Biophys Res Commun* **474**, 541-546
63. Chavatte, L., Seit-Nebi, A., Dubovaya, V., and Favre, A. (2002) The invariant uridine of stop codons contacts the conserved NIKSR loop of human eRF1 in the ribosome. *EMBO J* **21**, 5302-5311
64. Kryuchkova, P., Grishin, A., Eliseev, B., Karyagina, A., Frolova, L., and Alkalaeva, E. (2013) Two-step model of stop codon recognition by eukaryotic release factor eRF1. *Nucleic Acids Res* **41**, 4573-4586
65. Chavatte, L., Frolova, L., Laugaa, P., Kisselev, L., and Favre, A. (2003) Stop codons and UGG promote efficient binding of the polypeptide release factor eRF1 to the ribosomal A site. *J Mol Biol* **331**, 745-758
66. Frolova, L., Le Goff, X., Rasmussen, H. H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A. L., Celis, J. E., Philippe, M., and et al. (1994) A highly conserved eukaryotic protein family possessing properties of polypeptide chain release factor. *Nature* **372**, 701-703
67. Shao, S., Murray, J., Brown, A., Taunton, J., Ramakrishnan, V., and Hegde, R. S. (2016) Decoding Mammalian Ribosome-mRNA States by Translational GTPase Complexes. *Cell* **167**, 1229-1240 e1215
68. McCaughan, K. K., Brown, C. M., Dalphin, M. E., Berry, M. J., and Tate, W. P. (1995) Translational termination efficiency in mammals is influenced by the base following the stop codon. *Proc Natl Acad Sci U S A* **92**, 5431-5435
69. Salas-Marco, J., and Bedwell, D. M. (2004) GTP hydrolysis by eRF3 facilitates stop codon decoding during eukaryotic translation termination. *Mol Cell Biol* **24**, 7769-7778
70. Frolova, L., Le Goff, X., Zhouravleva, G., Davydova, E., Philippe, M., and Kisselev, L. (1996) Eukaryotic polypeptide chain release factor eRF3 is an eRF1- and ribosome-dependent guanosine triphosphatase. *RNA* **2**, 334-341
71. Mastronarde, D. N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51
72. Kremer, J. R., Mastronarde, D. N., and McIntosh, J. R. (1996) Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**, 71-76
73. Grant, T., Rohou, A., and Grigorieff, N. (2018) cisTEM, user-friendly software for single-particle image processing. *Elife* **7**
74. Loveland, A. B., Demo, G., Grigorieff, N., and Korostelev, A. A. (2017) Ensemble cryo-EM elucidates the mechanism of translation fidelity. *Nature* **546**, 113-117
75. Grigorieff, N. (2016) Frealign: An Exploratory Tool for Single-Particle Cryo-EM. *Methods Enzymol* **579**, 191-226
76. Cardone, G., Heymann, J. B., and Steven, A. C. (2013) One number does not fit all: mapping local variations in resolution in cryo-EM reconstructions. *J Struct Biol* **184**, 226-236
77. Bairoch, A., Boeckmann, B., Ferro, S., and Gasteiger, E. (2004) Swiss-Prot: juggling between evolution and stability. *Brief Bioinform* **5**, 39-55
78. Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-1612
79. DeLano, W. L. (2002) The PyMOL Molecular Graphics System. in *The PyMOL Molecular Graphics System*, The

PyMOL Molecular Graphics System Ed., DeLano Scientific, Palo Alto, CA, USA

80. Chapman, M. S. (1995) Restrained real-space macromolecular atomic refinement using a new resolution-dependent electron density function. *Acta Crystallogr. A* **51**, 69-80

81. Korostelev, A., Bertram, R., and Chapman, M. S. (2002) Simulated-annealing real-space refinement as a tool in model building. *Acta Crystallogr D Biol Crystallogr* **58**, 761-767

82. Svidritskiy, E., Brilot, A. F., Koh, C. S., Grigorieff, N., and Korostelev, A. A. (2014) Structures of Yeast 80S Ribosome-tRNA Complexes in the Rotated and Nonrotated Conformations. *Structure* **22**, 1210-1218

83. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54**, 905-921

84. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* **66**, 213-221

Table 1

Cryo-EM data collection and refinement statistics

Data collection	
EM equipment	FEI Talos Arctica
Voltage (kV)	200
Detector	Gatan K2 Summit
Software	SerialEM
Pixel size (Å)	0.944
Electron dose ($e^-/Å^2$)	37.7 (used 19)
Defocus range (μm)	-0.7 to -1.7
Data set size (# particles)	1,065,147
Reconstruction	
Software	cisTEM and Frealign v9.10-9.11
Number of particles used	639,088
Final resolution (Å, FSC=0.143)	3.7
Model composition	
Non-hydrogen atoms	152439
Protein residues	6520
RNA bases	4728
Refinement	
Software	RSRef and Phenix
Correlation Coeff	0.86
R-factor	0.186
Ramachandran-plot statistics (%)	
Favored (overall)	86.8
Allowed (overall)	12.4
Outlier (overall)	0.9
Rotamer outliers (%)	0.02
C-beta deviations	0
R.m.s. deviations	
Bond length (Å)	0.005
Bond angle (°)	0.84

Figure legends

Figure 1. (A) Michaelis-Menten curves for *in vitro* formyl-methionine release from fMet-tRNA by RF1 on the stop codon UAA and sense codons UAU, UGG and AAA (adopted from refs (5,11). (B) Cryo-EM structure of the 70S•RF1 complex formed on the UAU sense codon. Fourier shell correlation (FSC) curve is shown for the 70S•RF1 cryo-EM map (lower left). (C) Local resolution of RF1 in the cryo-EM map, determined using Blocres (76). RF1 is oriented similarly to the view shown in panel B. The map was sharpened by applying a B-factor of −100 Å^2 and is shown at 2σ, colored using a resolution scale ranging from 2.8 Å to 5.3 Å (left). (D) Cryo-EM map (mesh) in the peptidyl-transferase center (PTC). (E) Cryo-EM map (mesh) in the decoding center (DC). In structural models, the large 50S ribosomal subunit is shown in cyan, small 30S subunit in yellow, RF1 in green, mRNA in dark blue, P-site tRNA in orange and E-site tRNA in pink. Domains of RF1 are labeled in panels B and C.

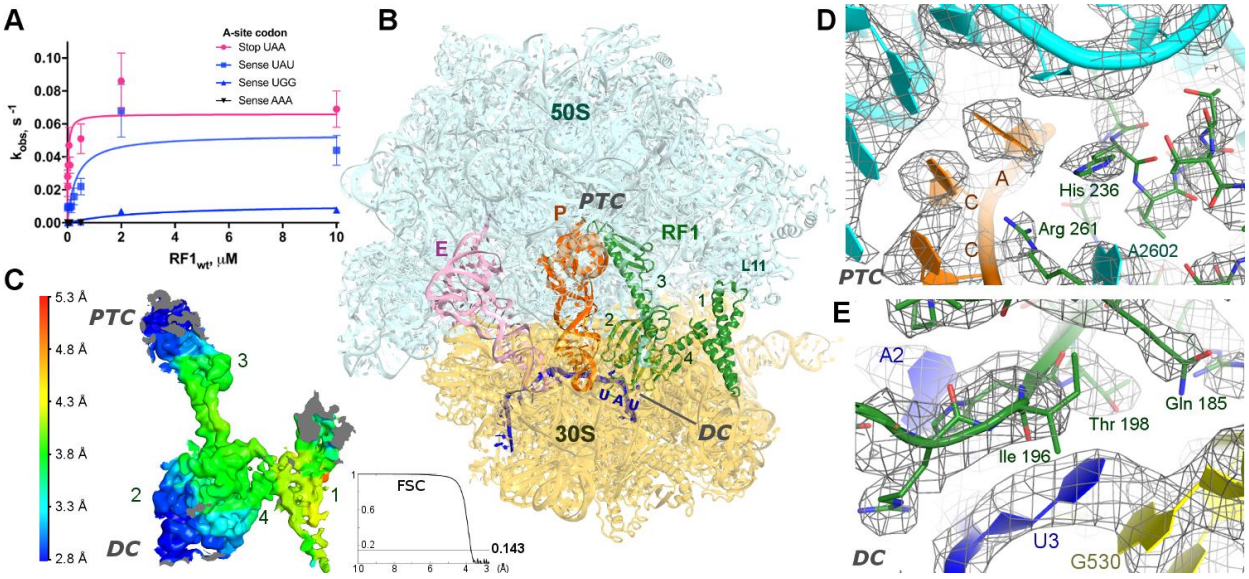


Figure 2. Recognition of the third nucleotide of the UAU and UAG codons by RF1 and eRF1. (A) Interactions of *Escherichia coli* RF1 with U3 of the UAU sense codon in the 70S ribosome (this work) (B) Interactions of *Thermus thermophilus* RF1 with G3 of the UAG stop codon in the 70S ribosome (PDB 4V7P; (9)). Thr 198 forms one of two possible hydrogen bonds with G3 (shown with the dashed and dotted lines). (C) Interactions of *Oryctolagus cuniculus* eRF1 with G3 of the UAG stop codon in the 80S ribosome (PDB 3JAH; (19)).

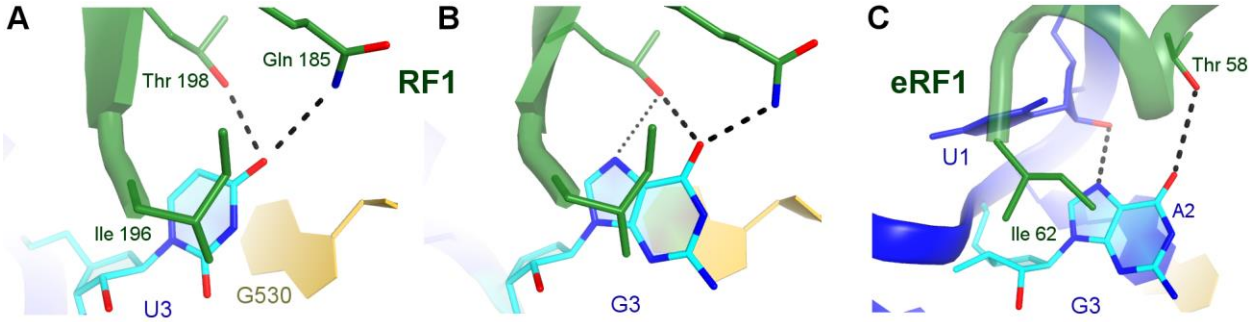
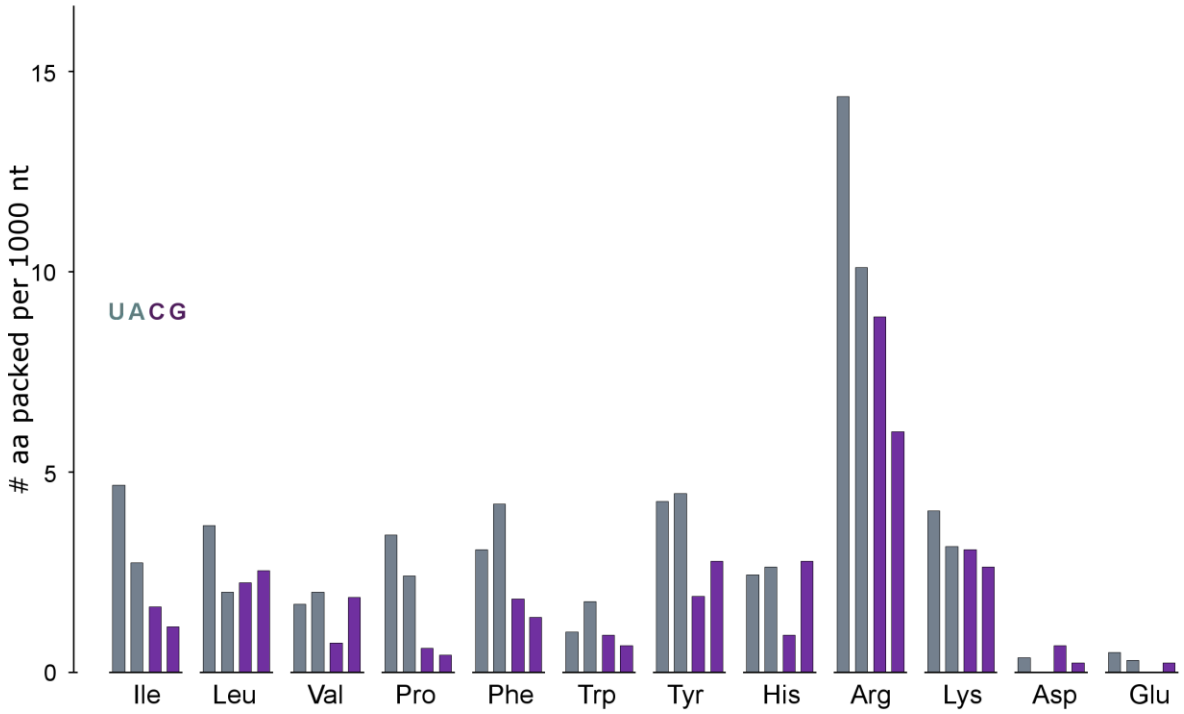


Figure 3. Distribution of amino acid side chains packing on the four types of RNA nucleotide bases (the more hydrophobic bases U and A are shown in gray, C and G – in purple) in three high-resolution ribosome structures from *E. coli* (PDB 4YBB: (48)), *T. thermophilus* (PDB 4Y4P: (49) and *S. cerevisiae* (PDB 4V88: (50).



Mechanism of premature translation termination on a sense codon

Egor Svidritskiy, Garbriel Demo and Andrei Korostelev

J. Biol. Chem. published online June 25, 2018
