Cell, Vol. 101, 349–352, May 12, 2000, Copyright ©2000 by Cell Press

Mimicry Grasps Reality in Translation Termination

Minireview

Yoshikazu Nakamura,*‡ Koichi Ito,* and Måns Ehrenberg†
*Department of Tumor Biology
The Institute of Medical Science
The University of Tokyo
Minato-ku, Tokyo 108-8639
Japan
†Department of Cell and Molecular Biology
Uppsala University
S-75124 Uppsala
Sweden

The termination of protein synthesis takes place on the ribosomes as a response to a stop, rather than a sense, codon in the "decoding" site (A site). Polypeptide release factors (RFs) play an essential role in this process. Although the termination process and the RF activity were discovered in the late 1960s (Goldstein and Caskey, 1970, and references therein), much of the mechanisms remained obscure. After three decades of investigation, we now know that two classes of RFs are required for translation termination in prokaryotes and eukaryotes: a class I factor, codon-specific RFs (RF1 for UAG/UAA and RF2 for UGA/UAA in prokaryotes; eRF1 in eukaryotes), and a class II factor, nonspecific RFs (RF3 in prokaryotes; eRF3 in eukaryotes) that bind GTP or GDP and stimulate class I RF activity (reviewed by Nakamura et al., 1996). In contrast, eukaryotic eRF1s normally recognize all three stop codons (reviewed by Buckingham et al., 1997). This review will focus on the mechanisms of stop codon recognition and translation termination by RFs from the perspective of molecular mimicry between translation factors and tRNA. It is surprising that it took four decades since the discovery of the genetic code to figure out the basic mechanisms behind the deciphering of its 64 codons.

Peptide "Anticodon"

tRNA-like properties have been suggested for RFs (reviewed by Tate et al., 1996) and the idea of an RF-tRNA molecular mimicry led to the proposal that there exists a peptide anticodon in RFs for reading stop codons (Nakamura et al., 1996, and references therein). This prediction was confirmed recently by the discovery of a peptide determinant in RFs equivalent to the anticodon of tRNA (Ito et al., 2000). Genetic selection combined with biochemical studies showed that the tripeptides Pro-Ala-Thr in RF1 and Ser-Pro-Phe in RF2 determine RF identity and that the first and third amino acids independently discriminate between the second and third purine bases, respectively (Figure 1). Thus, at the first position, Pro is restrictive to A (RF1), while Ser is permissive to both A and G (RF2). At the third position, Thr is permissive to A and G (RF1), while Phe is restrictive to A (RF2). These two discrimination switches operate separately since the Pro-Pro-Phe variant recognizes only UAA while the Ser-Pro-Thr variant recognizes three stop codons and UGG as well. They are referred to as a tripeptide "anticodon" that deciphers stop codons in mRNA.

Pauling's Paradox and the Release Factor Solution

In addition to the recognition of two stop codons each, RF1 and RF2 must discriminate against the 61 sense codons of the genetic code. This negative function may well be the most difficult design feature of termination of protein synthesis. Linus Pauling argued that the ability of proteins to discriminate between similar substrates is quite limited. If true this would lead to the apparent paradox that protein synthesis is much more accurate in amino acid selection than the physical chemistry seems to allow. Pauling's "paradox" was finally resolved when it was realized that the accuracy of enzymatic selections can be enhanced by energy-driven, multiple step "proofreading" of substrates (Kurland et al., 1996, and references therein). The mimicry hypothesis suggests that class II RFs share structural and functional properties with elongation factor EF-Tu (eEF-1α) or EF-G (Nakamura et al., 1996, and references therein). At face value mimicry would imply that class II RFs enhance the speed of peptide release since EF-Tu accelerates the rate of peptidyl transfer to aminoacyl-tRNA by orders of magnitude. It would also imply that the very tight discrimination against premature termination at sense codons observed in bacteria depends on proofreading driven by GTP hydrolysis on RF3 (eRF3) since hydrolysis of GTP on EF-Tu drives proofreading of aminoacyl-tRNAs (Kurland et al., 1996, and references therein). However, surprising results from very recent experiments on the peptide release reaction suggest that proofreading is not the only way to escape Pauling's paradox.

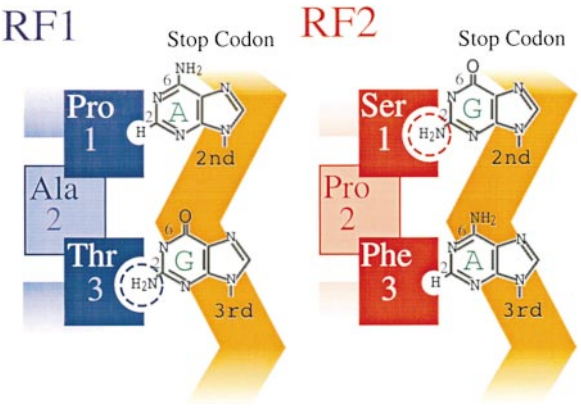

Figure 1. The Tripeptide Anticodon of Bacterial Release Factors

The first and third amino acids discriminate the second and third purine bases. The C-2 amino group of G is a primary target for discrimination by Pro and Phe, while Thr and Ser permit both C-2 amino group and proton of purine (dotted circles suggest a contribution of hydrogen bonding).

‡To whom correspondence should be addressed (e-mail: nak@ims.u-tokyo.ac.jp).

Table 1. The Accuracy of Release Factors for Reading Triplets: The Order of Magnitude of Fold Decrease in k_{cat}/K_M

Codon Change	RF1		RF2	
	−RF3	+RF3	−RF3	+RF3
[1st position]				
A	>6	>6	>5	>5
G	>6	5	5	5
C	5	4 ∼ 5	4	3 ∼ 4
[2nd position]				
G	4 ∼ 5	4		
C	4	3	3	3
U	4	3	4	4
[3rd position]				
G			3	3
C	4	3	4	4
U	3	2	4	4

A high value for the fold decrease in k_{cat}/K_M means that the accuracy is high. One nucleotide of cognate stop codons (U·A·A/G for RF1 and U·A/G·A for RF2) is changed to noncognate triplets.

Freistroffer et al. (2000) examined how well RF1 and RF2 discriminate similar mRNA sequences containing either of the three stop codons or a sense codon differing from any of the stop signals by a single mutation in the in vitro tetrapeptide synthesis. The first surprise was that RF3 increases the rate of termination at sense codons, but leaves the rate of termination at stop codons unchanged (Table 1), contrary to the idea that energy-driven proofreading enhances the accuracy of termination of protein synthesis. The second surprise was the enormous precision by which RF1 can discriminate between a U (cognate) and a C (noncognate) in the first position of codons. The efficiency of termination at CAA and CAG by RF1 is reduced a little less than a million-fold while the corresponding error by RF2 at CAA and CGA codons is about 20 times larger. This means that protein–RNA interactions can be so precise that Pauling's paradox doesn't arise. The third surprise was that processivity losses in vivo caused by RF (Kurland et al., 1996, and references therein) seem to be dominated by two hot spots for false termination events: UGG (Trp) codon for RF2 and UGU (Cys) for RF1 (Freistroffer et al., 2000).

Accurate Chemical Interaction between Peptide Anticodon and Stop Codon

The anticodon tripeptide defines the identity of RF1 and RF2 and explains how they avoid terminating at Trp codons (UGG). RF1 reads the RF2 stop codon UGA more than 100,000 times less efficiently and UGG about 50,000 times less efficiently than it reads UAA/G (Table 1). RF2 is sloppier and reads UGG as well as the RF1 codon UAG with about 2,000 times less efficiency than its own codons. The exclusive recognition of A is accomplished by the bulky hydrophobic amino acids Pro and Phe, neither of which can contribute to hydrogen bonding, and Pro in RF1 excludes reading of G in the second codon position more efficiently than Phe in RF2 excludes reading of G in the third position. The main mechanism of discrimination is steric exclusion of the C-2 amino group of G by either of the two bulky amino acids, while "wobble" recognition of both purines by the two hydrophilic amino acids is likely to involve hydrogen bonding (Figure 1).

One relevant example of precise recognition of specific RNA sequences by protein is the discriminator peptides identified in methionyl- and isoleucyl-tRNA synthetases that selectively recognize tRNAs (Auld and Schimmel, 1995, and references therein). When, for example, the native CAU anticodon of $tRNA^{Met}$ is changed to UAU, the efficiency of aminoacylation by methionyl-tRNA synthetase is reduced 10^4-fold, and when the anticodon becomes GAU instead the synthetase activity drops about 10^5-fold. Although the accuracy in tRNA recognition by this protein is impressive, the precision by which RF1 discriminates between a cognate U and a noncognate C or purines in the first position of a stop codon is between one and two orders of magnitude higher (Table 1). It is likely that a very careful design of the A site to hold both protein and mRNA (Yoshizawa et al., 1999) in sterically well defined positions is one factor behind this remarkable selectivity of RF1.

Mimicry Suggests that the Accuracy of RFs Can Be Altered in Two Principally Different Ways

The accuracy of reading of stop codons can be reduced by mutations in a region designated domain C of RF, about 40 amino acids apart from the tripeptide anticodon. Single Glu-to-Lys changes in domain C confer omnipotent decoding activity on RF2 and allow bacterial cells to grow with these single RF variants instead of two RFs (reviewed by Nakamura and Ito, 1998). Since these variants also catalyze protein termination at sense codons, they appear to reduce accuracy in a general fashion that charge-flip changes may strengthen the binding of RFs or influence the docking position to ribosomes in a codon-independent way. When the affinity between an enzyme and its cognate and noncognate substrates is increased in a nonspecific way, the accuracy of the reaction will be reduced (Kurland et al., 1996). This suggests that although the charge-flip variant omnipotence may appear similar to the "Ser-Pro-Thr" tripeptide anticodon, the mechanisms may be completely different: the former may be a loss of specificity due to a nonspecific increase in binding between factor and ribosome while the latter is a genuine loss of intrinsic selectivity of codon–anticodon interactions.

Recycling

After release of polypeptides by RFs, the ribosome is in complex with deacylated tRNA (in P site), a class I RF (in A site) and mRNA. In bacteria, the first step in ribosomal recycling back to initiation is dissociation of RF1 or RF2 from the ribosome, which is accelerated by RF3 in a GTP-dependent manner (Karimi et al., 1999, and references therein). Efficient release of RF1 and RF2 by RF3 is particularly important at strong stop signals (Crawford et al., 1999, and references therein). After GTP hydrolysis and dissociation of RF3 from the ribosome it contains mRNA, deacylated tRNA in the P site, and an empty A site. Studies by Kaji and colleagues in the 1970s have shown that another factor, RRF, together with EF-G, is required for dissociation of the bacterial posttermination ribosomal complex (Selmer et al., 1999, and references therein), but the mechanism behind this event has remained unclear.

Recently, Liljas and colleagues solved the 2.55 Å crystal structure of *Thermotoga maritima* RRF and found that it superimposes almost perfectly with tRNA except that the amino acid binding 3′ end is missing (Selmer

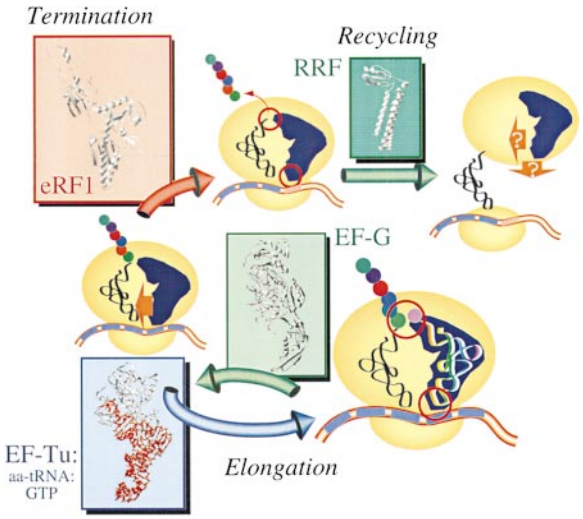

Figure 2. Crystal Structures of Translation Factors that Mimic tRNA and Their Working Steps during Protein Synthesis

Thermus thermophilus RRF (Protein Data Bank accession code 1EH1), yeast Phe-tRNA^{Phe}, EF-Tu:GDPNP:Phe-tRNA^{Phe} (1TTT), EF-G:GDP (1DAR), and human eRF1 (1DT9). Arrows and circles mean the target or the site of action.

et al., 1999, see Figure 2). They speculate that RRF interacts with the posttermination ribosome complex mimicking a peptidyl-tRNA in the A site thereby luring EF-G to translocate deacylated tRNA from the P to the E site from where it rapidly dissociates. However, this model is inconsistent with biochemical experiments by Karimi et al. (1999) which show, first, that RRF and EF-G split the ribosome into subunits in a reaction that requires GTP hydrolysis and, second, that initiation factor IF3 is required for the removal of deacylated tRNA from the P site of the 30S particle. Therefore, the mechanistic significance of a tRNA mimic by RRF remains to be verified.

Structure and Function of eRF1

The recently published crystal structure of human eRF1 to 2.8 Å by Barford and colleagues (Song et al., 2000) revealed that the overall shape and dimensions of eRF1 resemble a tRNA with domains 1, 2, and 3 of eRF1 corresponding, respectively, to an anticodon stem, an aminoacyl acceptor stem, and a T stem (see Figure 2). This domain assignment relies on the assumption that the universal GGQ motif (Frolova et al., 2000, and references therein) located at the tip of domain 2 is a structural counterpart of the tRNA aminoacyl group on the CCA-3′ acceptor stem and that a codon-specific discrimination defect can be created in domain 1 (Song et al., 2000). In this view the eRF1 anticodon is located at or near the tip of domain 1 in the N-terminal region. Extending the bacterial tripeptide "anticodon" analogy to eRF1s, it is speculated that a "Thr-Ala-Ser" tripeptide adjacent to the helical hairpin might play the role of an omnipotent discriminator tripeptide. This gets some support from the strong conservation of Thr-Ala-Ser in eukaryotes, with *Tetrahymena* as the exception with Lys-Ala-Ser in this position. Interestingly, *Tetrahymena* has UGA as sole stop codon with UAA and UAG reassigned to glutamine codons, which is consistent with its exceptional tripeptide. A simple omnipotent discriminator tripeptide of *E. coli* type could not account for the exclusive recognition of all three stop codons, since it would recognize UGG as well (Ito et al., 2000). It remains to be seen whether the predicted Thr-Ala-Ser tripeptide in eRF1 can permit any two purines in second and third codon positions except double Gs. It is worth mentioning that functional assessment of each domain of eRF1 still remains to be verified because the mutant phenotypes of the GGQ motif or the disabled codon recognition, which are behind the prediction of domain function, can be generated by several topologically distinct alterations in bacterial RF1 and RF2 (Nakamura and Ito, 1998, and references therein).

What Is the Primary Role of eRF3?

eRF3 is essential for cell growth and forms a stable complex with eRF1 off the ribosome (Eurwilaichitr et al., 1999, and references therein), while RF3 is dispensable and does not bind to RF1 or RF2 off the ribosome (reviewed by Buckingham et al., 1997). Therefore, eRF3 seems to play a role distinct from RF3, which only removes class I factors from the ribosome, and to more resemble an "EF-Tu-like" protein that brings class I proteins to the A site of the ribosome. Deletion analyses have shown that residues 281–415 of human eRF1 are necessary and sufficient for its interaction with eRF3 (Frolova et al., 2000, and references therein), and this is a region of eRF1 that corresponds exactly to the core secondary structure of its domain 3 (Song et al., 2000). At the same time, eRF1 derivatives lacking the C-terminal region necessary for strong binding to eRF3 remain active as RFs in vivo (Eurwilaichitr et al., 1999, and references therein) and in vitro (Frolova et al., 2000, and references therein). Therefore, eRF3 must have an unknown essential function in addition to what is suggested by a simple "eRF3-EF-Tu mimicry." This is further emphasized by the fact that the C-terminal one-third domain of eRF3, without the G domain, is sufficient for binding to eRF1 (Frolova et al., 2000). This is in sharp contrast with other translational G proteins (e.g., EF-Tu, eEF1) whose substrate binding is controlled by their G domains. Therefore, the significance of the G domain of eRF3 remains obscure, but also the main role in eukaryotic termination of translation of this essential factor is still unknown. It is likely that one function of eRF3 is to recycle class I RFs, in analogy with RF3. However, eRF3 may also enhance accuracy of termination by proofreading. This may be necessary since eRF1s normally recognize all three stop codons, possibly making discrimination against sense codons more difficult. One may also ask whether eRF3 has another novel function(s) associated with its prion-like N-terminal extension (reviewed by Tuite, 2000).

A Clue to Protein–tRNA Mimicry in the Translational Apparatus

Molecular mimicry between protein and RNA was first suggested by Nyborg, Clark, and colleagues when they saw that the crystal structure of the EF-Tu:GTP:aminoacyl-tRNA ternary complex has high structural similarity with EF-G, e.g., so that domain IV of EF-G mimics the anticodon stem of tRNA (Nissen et al., 1995) (Figure 2). Tethered radical footprinting and cryo-electron microscopy analysis of the EF-G:ribosome complex reveals A site occupation by domain IV of EF-G and proximity of

the tip of its domain IV to the 30S decoding site (Wilson and Noller, 1998, and references therein). Thus, the structural mimicry of domain IV, inferred from the crystallographic comparison, extends to its position in or near the tRNA binding region of the ribosome.

Although inspired by these data, the RF-tRNA mimicry hypothesis was inferred from both genetic and biochemical studies (reviewed by Nakamura et al., 1996; Tate et al., 1996). The main point was not structural but to use mimicry to direct an experimental search for a functional peptide anticodon in RFs. This successfully unraveled the tripeptide anticodon determinant, which is completely conserved in class I RFs of prokaryotes. It is all the same that the crystal structure of human eRF1 significantly resembles a tRNA molecule (Figure 2) even if it may turn out that it uses a different protein architecture than prokaryotic RFs. Selmer et al. (1999) have argued that the mimicry analogy may hold also for the recycling step of protein synthesis from the crystallographic data (see Figure 2).

A mimic of the shape of a tRNA works as an entrance pass to sit in the cockpit (A site) in a ribosome "machine." However, the action once sitting there is diverse for the different translation factors as schematically presented in Figure 2: pushing peptidyl-tRNA, deciphering stop codons, triggering hydrolysis, and splitting off the 50S subunit. Nature must have evolved this "art" of molecular mimicry using different protein architectures (see ribbon diagrams in Figure 2) for the diverse actions, still keeping a similar shape to fulfill the requirement of the ribosome.

Selected Reading

Auld, D.S., and Schimmel, P. (1995). Science *267*, 1994–1996.

Buckingham, R.H., Grentzmann, G., and Kisselev, L. (1997). Mol. Microbiol. *24*, 449–456.

Crawford, D.-J.G., Ito, K., Nakamura, Y., and Tate, W.P. (1999). EMBO J. *18*, 727–732.

Eurwilaichitr, L., Graves, F.M., Stansfield, I., and Tuite, M. (1999). Mol. Microbiol. *32*, 485–496.

Freistroffer, D.V., Kwiatkowski, M., Buckingham, R.H., and Ehrenberg, M. (2000). Proc. Natl. Acad. Sci. USA *97*, 2046–2051.

Frolova, L.Y., Merkulova, T.I., and Kisselev, L.L. (2000). RNA *6*, 381–390.

Goldstein, J.L., and Caskey, C.T. (1970). Proc. Natl. Acad. Sci. USA *67*, 537–543.

Ito, K., Uno, M., and Nakamura, Y. (2000). Nature *403*, 680–684.

Karimi, R., Pavlov, M.Y., Buckingham, R.H., and Ehrenberg, M. (1999). Mol. Cell. *3*, 601–609.

Kurland, C.G., Hughes, D., and Ehrenberg, M. (1996). In Escherichia coli and Salmonella, E.C. Neidhardt, ed. (Washington, D.C.: ASM Press), pp. 979–1004.

Nakamura, Y., and Ito, K. (1998). Genes Cells *3*, 263–278.

Nakamura, Y., Ito, K., and Isaksson, L.A. (1996). Cell *87*, 147–150.

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

Selmer, M., Al-Karadaghi, S., Hirokawa, G., Kaji, A., and Liljas, A. (1999). Science *286*, 2349–2352.

Song, H., Mugnier, P., Das, A.K., Webb, H.M., Evans, D.R., Tuite, M.F., Hemmings, B.A., and Barford, D. (2000). Cell *100*, 311–321.

Tate, W.P., Poole, E.S., and Mannering, S.A. (1996). Prog. Nucleic Acids Res. *52*, 293–335.

Tuite, M.F. (2000). Cell *100*, 289–292.

Wilson, K.S., and Noller, H.F. (1998). Cell *92*, 131–139.

Yoshizawa, S., Fourmy, D., and Puglisi, J.D. (1999). Science *285*, 1722–1725.