FEBS Letters 514 (2002) 30–33 FEBS 25818

Minireview

A tripeptide discriminator for stop codon recognition

Yoshikazu Nakamura*, Koichi Ito

Department of Basic Medical Sciences, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8630, Japan

Received 14 November 2001; revised 20 November 2001; accepted 21 November 2001

First published online 1 February 2002

Edited by Lev Kisselev

Abstract Only recently has it been established that a tripeptide in the bacterial release factors (RFs), RF1 and RF2, is responsible for the stop codon recognition. This functional mimic of the anticodon of tRNA is referred to as a tripeptide 'anticodon' or a tripeptide discriminator. Here we review the experimental background and process leading to this discovery, and strengthen functional evidence for the tripeptide determinant for deciphering stop codons in mRNAs in prokaryotes.

Key words: Translation termination; Stop codon; Tripeptide anticodon; Polypeptide release factor; tRNA mimic

1. Introduction

Since the deciphering of the genetic code in the early 1960s, it has remained uncertain how the three stop codons, UAG, UGA and UAA, are recognized and how the ribosome terminates translation. The first step of translational termination involves the recognition of a stop codon in mRNA by polypeptide release factors (RFs), which then induce hydrolysis of the ester bond linking the nascent polypeptide to tRNA on the ribosome [1]. Prokaryotes have generally two codon-specific factors with overlapping specificity, RF1 (for UAG/UAA) and RF2 (for UGA/UAA), while eukaryotes have only one factor, eRF1, which normally recognizes all three stop codons [2]. By virtue of their functions, RFs have long been thought to mimic tRNA [3,4]. The fact that RF1 and RF2 exhibit codon specificity led many researchers to speculate that they interact directly with their codons. However, evidence has been lacking for such direct contact or direct deciphering of stop codons by protein RFs. Only recently a functional mimic of the anticodon of tRNA has been discovered in RF1 and RF2, which we refer to as a tripeptide 'anticodon' or a tripeptide discriminator [1,5]. In this article, we aim to review and strengthen evidence for the discriminator function of the peptide element and gain further insight into the mechanism.

2. Peptide swapping

A clue to the problem was realized when we cloned and sequenced the structural gene for RF3 [6], which is known to bind guanine nucleotides and stimulate RF1 and RF2 activity. Its primary protein sequence resembled the N-terminal part of elongation factor G (EF-G), rather than the elongation factor Tu. This led us to speculate that RF1 and RF2 might be equivalent to the C-terminal part of EF-G given that the RFs evolved from a progenitor of EF-G [7,8]. This initial idea, though speculative, urged us to compare the primary sequences of RFs from different organisms and other translation factors. Thus, we became aware of universally conserved seven-domain structures in the primary sequences, domains A through G, in prokaryotic RFs [9]. Although the predicted domain structure does not necessarily reflect the tertiary structure, the successful partitioning of RF1 and RF2 sequences into seven common conservative regions provided us with a theoretical means by which domain function could be investigated by swapping domains between RF1 and RF2.

The seven domains were swapped combinatorially between RF1 (UAG-specific) and RF2 (UGA-specific) to screen for active RF hybrids that would display altered codon specificity in vivo. For this purpose, common restriction sites were introduced into a clone at, or near, the sequences encoding the domain junctions and, thus, these base changes did not affect activity [5]. A combinatory set of 128 RF hybrids was examined for the ability to complement RF1 knockout (*prfA*::Km^R) and RF2 knockout (*prfB*::Cm^R) alleles. For initial screening, phenotypic in vivo assessment of RF1 and RF2 activity was more reliable than in vitro analysis. This was because the specific activity of RF1 greatly differs from – is much higher than – that of RF2, and Glu-to-Lys (charge-flip) changes at specific positions in RFs are capable of triggering polypeptide release at all three stop codons [10,11], which would interfere with proper assessment of amino acid changes responsible for the altered stop codon selectivity.

Fig. 1 shows hybrid RF proteins that are capable of restoring viability to either the RF1 knockout strain or the RF2 knockout strain. These hybrid proteins cannot reverse the lethality of the RF1/RF2 doubly defective strain, showing that each hybrid RF possesses RF1-specific or RF2-specific complementation activity in vivo and thus does not exert omnipotent activity. Next, based on the in vivo profiling, each hybrid RF protein was purified and the activity in vitro was

*Corresponding author. Present address: Department of Basic Medical Sciences, Institute of Medical Science, University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan. Fax: (81)-3-5449 5415. *E-mail address:* nak@ims.u-tokyo.ac.jp (Y. Nakamura).

Abbreviations: RF, release factor; PAT, Pro–Ala–Thr; SPF, Ser–Pro–Phe

PII: S0014-5793(02)02330-X

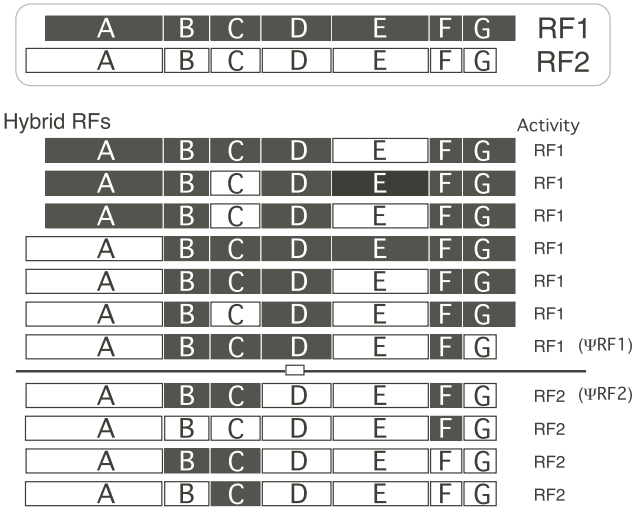

Fig. 1. Active RF hybrids. Seven conserved domains A through G were swapped between bacterial RF1 and RF2 sequences. The activity of the hybrid proteins was monitored by the in vivo complementation test of RF1 null and RF2 null strains as well as by the in vitro fMet release assay. A pair of RF hybrids, ΨRF1 and ΨRF2, exerted RF1- and RF2-specific complementation activities due to the respective domain D inserts. A hybrid RF construct, designated ΨRF, which is identical to ΨRF1 and ΨRF2 except for domain D [5], was used as a test RF backbone to monitor specificity of transplanted peptides and an amino acid swap made in the discriminator tripeptide. Closed and open boxes represent RF1 and RF2 segments, respectively. Discriminator–peptide swapping was conducted within the domain D subregion shown by the small open box.

monitored by f[^{3}H]Met release from ribosomes complexed with f[^{3}H]Met–tRNA$_{fMet}$ and a 9-mer mini-messenger RNA (5′-UUC AUG-3′, followed by stop or test triplets). The in vitro data was consistent with the in vivo results; namely, RF1-complementing hybrids respond to UAA and UAG, while RF2-complementing hybrids respond to UAA and UGA. These results firmly establish that the codon-responsive profiling shown in Fig. 1 represents the nature of stop codon selectivity by RF1 and RF2.

From Fig. 1, we could make several important conclusions. First, these hybrid RFs were scored in pairwise combinations in which only one domain differed. The criterion used to assign the discriminator domain was its ability to switch, without exception, the complementation activity exclusively and efficiently between RF1 and RF2. We found that a unique set of functional RF hybrids, designated ΨRF1 and ΨRF2, containing a single distinct domain, domain D, showed RF1- and RF2-specific in vivo complementation and in vitro peptide release activities. Any active RF hybrids other than these two switched the selectivity upon substituting domain D between RF1 and RF2. Hence, we conclude that domain D contains the peptide determinant, or discriminator, for deciphering stop codons. Second, we also conclude that the hybrid ΨRF sequence can play a vital role in the assessment of the discriminator activity upon swapping the peptide 'anticodon' segment between RF1 and RF2. Third, and importantly, swapping of any domain(s) other than domain D did not switch selectivity between RF1 and RF2. For example, domain A swapping never changes the response to stop codons if domain D is unchanged (see Fig. 1). This means that any regions other than domain D do not encode a functional discriminator activity for stop codons in mRNA. Fourth, swapping of domain D in the native RF1 and RF2 sequences simply inactivated the RF activity (data not shown). This suggests that even if RF1 and RF2 share a progenitor, both factors established their own structural signatures, not only in domain D but also in other regions, to achieve a selective recognition of stop codons during evolution.

3. Amino acid shuffling

By swapping peptides within domain D, the discriminator region was further defined to the central 15 amino acid segment, in which five amino acids are different and perfectly conserved in RF1s and RF2s of many species [5]. The conserved RF1- and RF2-specific motifs in this region were 'Q--PAT-----I' and 'V--SPF-----R', respectively. Transfer of a 24-mer segment containing these 15 amino acids from RF1 to ΨRF2 switched the specificity from RF2 to RF1. Hence, subsequent experiments used variants of this 15-mer segment as discriminator cassettes, in which five amino acids were shuffled between RF1 and RF2. Using designed primers, hybrid ΨRF1 and ΨRF2 proteins containing 256 (i.e. $4\times2\times2\times4\times4$) combinatory sets of conservative (and variant) five amino acids were constructed, transformed into the RF1 knockout and RF2 knockout strains, and viable colonies were selected as described [5]. Fig. 2 represents amino acid requirements for each position that confer RF1 or RF2 activity on hybrid ΨRF in vivo. All five positions preferred the authentic amino acids to the shuffled amino acids, highlighting the amino acid requirements for the relevant RF activity under the structural constraint of ΨRF1 and ΨRF2

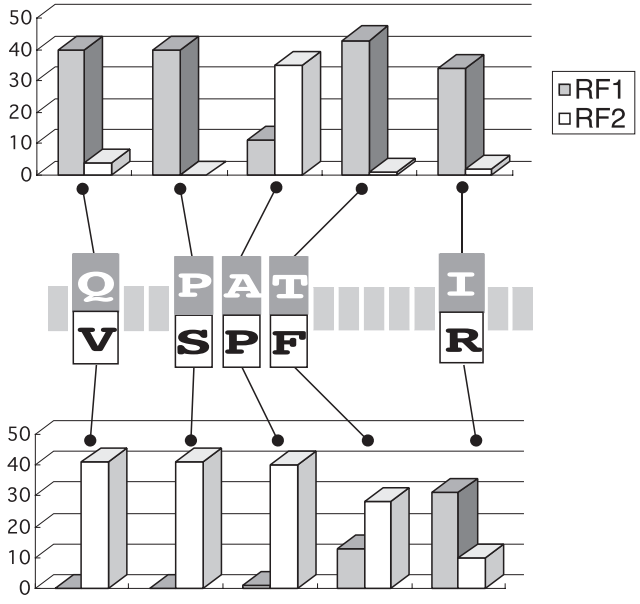

Fig. 2. Shuffling of conserved amino acids between RF1 and RF2 sequences at the assigned discriminator region. The indicated five residues are differentially conserved in RF1s and RF2s of many species. Each position was partially randomized using primer sequences: position Q/V contained Q, V, E and L; position P/S contained P and S; position A/P contained A and P; position T/F contained T, F, I and S; position I/R contained I, R, L and S. Hybrid ΨRF1 and ΨRF2 proteins containing all shuffled combinations (i.e. 256 cases) were synthesized and transformed into RF1 null and RF2 null strains. The number of winning amino acids at each position through each mutant selection and the specificity are shown. Closed and open bars represent RF1- and RF2-type hybrids, respectively.

sequences. In spite of the conservative residues at five positions, we became aware through site-directed mutagenesis that substitutions of amino acids other than PAT (Pro–Ala–Thr) and SPF (Ser–Pro–Phe) in this segment did not affect the selectivity of codons (data not shown). The conserved residues, 'Q/V' and 'I/R', were indirectly involved in respective RF activity, but the requirement for RF2's ability to discriminate was relatively more restrictive than that for RF1, preferring 'V' to 'Q', for instance (data not shown). Hence, we assume that tripeptides PAT and SPF play a crucial role in deciphering stop codons in RF1 and RF2. Nevertheless, the screening of requirements for the discriminator as shown in Fig. 2 indicated that 'A' of PAT and 'F' of SPF were apparently less stringent because of a relaxed amino acid requirement for the former position and an occurrence of omnipotent RF variant by the latter change (see below).

4. Selection of amino acids for functional discriminators

The systematic selection of the discriminator motifs was performed using designed tripeptide collections in which the second position was fixed to Pro, which functions in both RF1 and RF2 (see Fig. 2), while the first or third position was fixed to either amino acid of RF1 and RF2 in otherwise random sequences: XPT (pool-1), XPF (pool-2), PPX (pool-3) and SPX (pool-4), where P = Pro, S = Ser, T = Thr, F = Phe, and X = 20 amino acids. These collections were transformed into the RF1 knockout and RF2 knockout strains, and viable colonies were selected. Winners from these pools that reversed the lethality of RF1 or RF2 null strain are summarized in Fig. 3 as a spatial presentation. Several important conclusions are drawn here. Winners from pool-1 that rescued the RF1 null contained Ser or Pro at the first position, while the only winner in the RF2 null was SPT (omnipotent; see below). When the third position was fixed as Phe (pool-2), the sole winner was SPF, which appeared in the RF2 null: no winners appeared in the RF1 null (see Fig. 3).

Winners from pool-3 that rescued the RF1 null strain contained non-charged amino acids at the third position to give rise to PPL, PPS, PPT, PPQ and PPN, which function as RF1, and none rescued the RF2 null strain. Winners from pool-4 that rescued RF1 and RF2 null strains shared amino acids at the third position to give rise to SPL, SPI, SPM, SPT, SPH, SPF and SPR, most of which are omnipotent for three stop codons except for Phe (in SPF), which was restrictive to the RF2 null strain (see Fig. 3).

The above genetic selection revealed the tripeptide amino acid species, including both conserved and non-conserved residues, required for discriminator activity. The codon specificity of these variant RFs was examined using the in vitro release assay. Of the tripeptides selected using either RF1 null or RF2 null conditions, SPT, SPL, SPR, SPI, SPM and SPH variants actually rescued both the RF1 null and the RF2 null. Indeed, these variants catalyzed polypeptide release at all three stop codons (omnipotent) and were able to complement the RF1/RF2 double knockout strain [5]. On the other hand, variants such as SPF, PPT and PPN rescued exclusively either the RF1 null or the RF2 null. These variants catalyzed polypeptide release either at UAA/UAG or at UAA/UGA, and were only able to complement either defect in RF1 or RF2, but not the doubly defective strain.

Based on these patterns and codon specificity, we concluded that the tripeptides PAT in RF1 and SPF in RF2 determine RF specificity and that the first and third amino acids independently discriminate the second and third purine bases, respectively [5]. 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). The amino acid profiling shown in Fig. 3 further demonstrates that the principle rule of base discrimination by the discriminator tripeptide is that amino acids used for restricted recognition of A are bulky hydrophobic residues, such as Pro and Phe, whereas those for relaxed recognition of A and G are small hydrophilic residues, such as Ser and Thr. The latter relaxed recognition of A and G may be phenotypically equivalent to wobble pairing between the mRNA triplet and tRNA anticodon at the third position. What is unique to RF is to allow wobble not only at the third position of the stop codon but also at the second position. These two discrimination switches operate separately since the PPF variant recognizes only UAA while the SPT variant recognizes the three stop codons and also UGG [5].

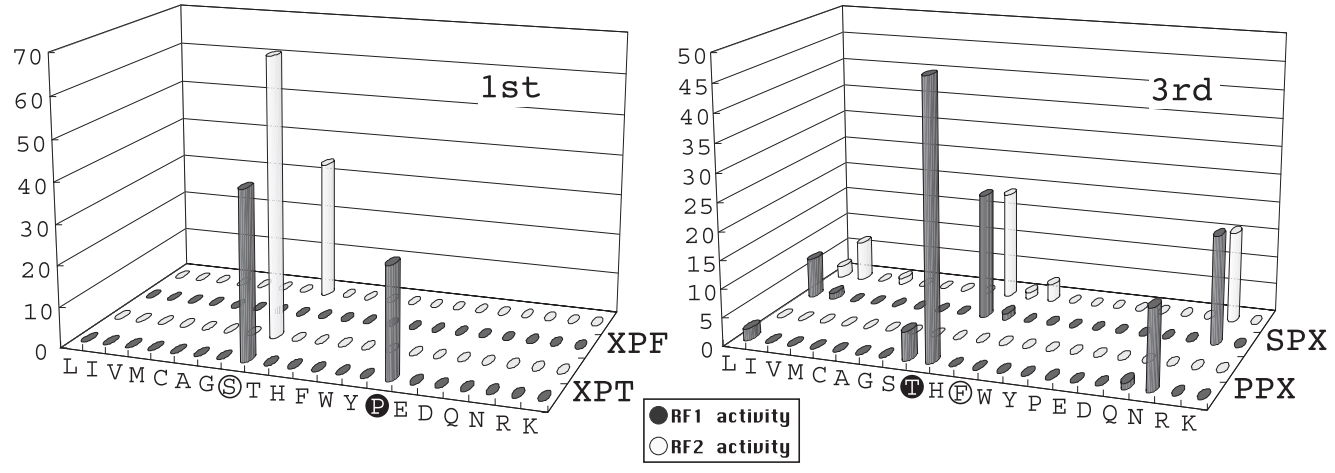

Fig. 3. Selection of discriminator tripeptides of peptide 'anticodon'. ΨRF variants carrying designed (random) sequence pools of discriminator tripeptides were transformed into RF1 null and RF2 null strains, and active variants were selected as described previously [5]. Winners were examined for their sequence and activity to complement RF1 and/or RF2 null mutations. The number of winning amino acids at the first (left panel) and third (right panel) positions and the specificity are shown here as a spatial presentation. Amino acids are presented by single letter codes and X represents 20 amino acids.

These results led us to conclude that the defined tripeptide represents the discriminator element that 'functionally' deciphers stop codons in mRNA; hence referred to as a tripeptide 'anticodon' or a tripeptide discriminator.

5. Chemical basis of base discrimination

Two purine bases, A and G, differ in two side groups – C-2 amino (G) or C-2 proton (A) and C-6 carbonyl (G) or C-6 amino (A). Hence, either or both of them may be potent targets for discrimination by the RF tripeptides. To test these possibilities, we substituted inosine (I) for G in the stop codon and analyzed polypeptide release in vitro. Both UAI and UIA were recognized by RF1 and RF2, showing that the I substitution removed the primary discrimination target. On the other hand, UIG and UGI were still recognized selectively by RF1 and RF2, respectively [5]. Therefore, it is very likely that the tripeptide discriminates primarily the C-2 amino group of G at both the second and third bases. Nevertheless, the potential contribution, if any, of the C-6 amino group of A remains to be examined because of the apparently lower release activity of A-to-I variants at the discriminator position. In addition, a minor part of the discrimination may come from the C-6 group of the purines, since substituting 2-amino purine for guanine, or inosine for adenine impaired termination somewhat. Although the molecular or structural basis of these interactions is unknown, the highly specific 'functional' interplay of amino acid species in the discriminator tripeptides and the C-2 side group of purine bases might be consistent with the idea that the tripeptide discriminator interacts directly with the stop codons.

6. Topological evidence for A-site docking of a tripeptide discriminator

Functional sites of interaction between RF1 and the ribosome have been assigned by directed hydroxyl radical probing [12]. Interestingly, the site-directed radical cleavages tethered at positions 187 and 192 of RF1 in 16S rRNA are very similar to the corresponding positions on EF-G, when bound to the ribosome in the post-translocational state. In particular, Fe(II) tethered to EF-G at the tip of its anticodon-mimicking domain IV results in identical rRNA cleavages in the head and platform of the small subunit surrounding the decoding site [13]. RF1 positions 187 and 192 flank the tripeptide 'anticodon' (188–190). Thus, the anticodon mimic inferred from the functional study is within or near the tRNA-binding region of the ribosome, and thus provides strong support for A-site docking of the tripeptide discriminator.

More evidence that the RFs are in close contact with the translational stop signal comes from site-directed crosslinks formed from the first position (+1) of the stop codon (where the +1 U was replaced by photoactive 4-thio-U). A crosslink from the first position of the stop codon was identified in domain D near to or within the peptide 'anticodon' region [14]. A second crosslink was found at positions 130–140 in the RF2 sequence [14,15]. This region is located in domain C adjacent to the tripeptide discriminator. Since specific amino acid substitutions in domain C are known to alter the decoding capacity of the tripeptide discriminator [10,11], one could speculate that these residues interact functionally with the tripeptide. Hence, this provides us with another support for the spatial localization of these positions, including the tripeptide, near the decoding pocket of the 30S subunit.

7. Conclusions

The work reviewed here has established that the PAT and SPF tripeptides in the bacterial RFs, RF1 and RF2, respectively, serve as the functional elements to decipher stop codons in mRNA. Since these two tripeptides are able to discriminate the second and third purine bases in stop codons, they are referred to as tripeptide 'anticodons' or tripeptide discriminators. We should emphasize here that these tripeptides are 'functionally' defined. Therefore, it is of particular interest whether or not the structure of bacterial RFs, once known, can solve the mechanism of the discriminating function of the tripeptides. Unless the structure could easily account for the mechanism due to an altered structural motif (or a structural non-resemblance) between the tripeptide region and the anticodon arm of tRNA, we believe that the mechanism of stop codon recognition by RFs should be more complex than what we speculate from a simple tRNA mimic. Nevertheless, the true answer should be able to explain the discriminating action of the tripeptides PAT and SPF for deciphering stop codons in bacterial RFs RF1 and RF2.

Acknowledgements: We thank Colin Crist for reading the manuscript. This work was supported in part by grants from the Ministry of Education, Sports, Culture, Science and Technology of Japan, the Human Frontier Science Program, the Basic Research for Innovation Biosciences Program of the Bio-oriented Technology Research Advancement Institution (BRAIN), and the Mitsubishi Foundation.

References

[1] Nakamura, Y., Ito, K. and Ehrenberg, M. (2000) Cell 101, 349–352.
[2] Kisselev, L.L. and Buckingham, R.H. (2000) Trends Biochem. Sci. 25, 561–567.
[3] Nakamura, Y., Ito, K. and Isaksson, L.A. (1996) Cell 87, 147–150.
[4] Tate, W.P., Poole, E.S. and Mannering, S.A. (1996) Prog. Nucleic Acids Res. 52, 293–335.
[5] Ito, K., Uno, M. and Nakamura, Y. (2000) Nature 403, 680–684.
[6] Mikuni, O., Ito, K., Moffat, J., Matsumura, K., McCaughan, K., Nobukuni, T., Tate, W. and Nakamura, Y. (1994) Proc. Natl. Acad. Sci. USA 91, 5798–5802.
[7] Nakamura, Y. and Ito, K. (1998) Genes Cells 3, 265–278.
[8] Nakamura, Y., Kawazu, Y., Uno, M., Yoshimura, K. and Ito, K. (2000) in: The Ribosome: Structure, Function, Antibiotics and Cellular Interactions (Garrett, R.A., Douthwaite, S.R., Liljas, A., Matheson, A.T., Moore, P.B. and Noller, H.F., Eds.), pp. 519–562, The American Society for Microbiology, Washington, DC.
[9] Ito, K., Ebihara, K., Uno, M. and Nakamura, Y. (1996) Proc. Natl. Acad. Sci. USA 93, 5443–5448.
[10] Ito, K., Uno, M. and Nakamura, Y. (1998) Proc. Natl. Acad. Sci. USA 95, 8165–8169.
[11] Uno, M., Ito, K. and Nakamura, Y. (2002) Proc. Natl. Acad. Sci. USA, in press.
[12] Wilson, K., Ito, K., Noller, H. and Nakamura, Y. (2000) Nat. Struct. Biol. 7, 866–870.
[13] Wilson, K.S. and Noller, H.F. (1998) Cell 92, 131–139.
[14] Poole, E. and Tate, W. (2000) Biochim. Biophys. Acta 1493, 1–11.
[15] Wilson, D.N., Dalphin, M.E., Pel, H.J., Major, L.L., Mansell, J.B. and Tate, W.P. (2000) in: The Ribosome: Structure, Function, Antibiotics and Cellular Interactions (Garrett, R.A., Douthwaite, S.R., Liljas, A., Matheson, A.T., Moore, P.B. and Noller, H.F., Eds.), pp. 495–508, The American Society for Microbiology, Washington, DC.