
Bioinformatic, structural, and functional analyses support release factor-like MTRF1 as a protein able to decode nonstandard stop codons beginning with adenine in vertebrate mitochondria

DAVID J. YOUNG, CHRISTINA D. EDGAR, JENNIFER MURPHY, JOHANNES FREDEBOHM, ELIZABETH S. POOLE, and WARREN P. TATE

Department of Biochemistry, University of Otago, Dunedin 9054, New Zealand

ABSTRACT

Vertebrate mitochondria use stop codons UAA and UAG decoded by the release factor (RF) MTRF1L and two reassigned arginine codons, AGA and AGG. A second highly conserved RF-like factor, MTRF1, which evolved from a gene duplication of an ancestral mitochondrial RF1 and not a RF2, is a good candidate for recognizing the nonstandard codons. MTRF1 differs from other RFs by having insertions in the two external loops important for stop codon recognition (tip of helix $\alpha 5$ and recognition loop) and by having key substitutions that are involved in stop codon interactions in eubacterial RF/ribosome structures. These changes may allow recognition of the larger purine base in the first position of AGA/G and, uniquely for RFs, only of G at position 2. In contrast, residues that support A and G recognition in the third position in RF1 are conserved as would be required for recognition of AGA and AGG. Since an assay with vertebrate mitochondrial ribosomes has not been established, we modified *Escherichia coli* RF1 at the helix $\alpha 5$ and recognition loop regions to mimic MTRF1. There was loss of peptidyl-tRNA hydrolysis activity with standard stop codons beginning with U (e.g., UAG), but a gain of activity with codons beginning with A (AAG in particular). A lower level of activity with AGA could be enhanced by solvent modification. These observations imply that MTRF1 has the characteristics to recognize A as the first base of a stop codon as would be required to decode the nonstandard codons AGA and AGG.

Keywords: vertebrate mitochondria; release factors; translation; codon recognition; MTRF1; MTRF1L

INTRODUCTION

Translation termination for most genetic systems requires recognition of one of three standard stop codons (UAG, UAA, and UGA) at the ribosomal decoding center followed by hydrolysis of the nascent polypeptide chain from the peptidyl-tRNA at the peptidyl-transferase center (for reviews, see Poole et al. 2007; Petry et al. 2008; Youngman et al. 2008). Stop codons are recognized by one of two decoding release factors (RFs) in eubacteria. RF1 restricts base recognition in positions 1 and 2 to a single base (U and A, respectively) but accommodates both purine bases at position 3, and provides a model for mitochondrial factor

recognition of the standard stop codons UAA and UAG. In contrast, RF2 restricts recognition to a single base in positions 1 and 3 (U and A, respectively) but accommodates both purine bases in position 2 (Scolnick et al. 1968).

A tripeptide motif critical for codon recognition in eubacteria, PXT in RF1 and SPF in RF2, was proposed as the key discriminator for recognition of the bases at positions 2 and 3 of the stop codon (Ito et al. 2000). The motif was later found to be within an exposed loop (sometimes referred to as the “anticodon” loop, which we now call the recognition loop) at the surface of the RFs, consistent with that proposal (Vestergaard et al. 2001; Shin et al. 2004; Graille et al. 2005; Zoldak et al. 2007). New structural studies now show these motifs are important for interaction with the first two nucleotides rather than the last base of the stop codon (Korostelev et al. 2008; Laurberg et al. 2008; Weixlbaumer et al. 2008). Moreover, other residues of the loop also form interactions with the codon.

Reprint requests to: Warren P. Tate, Department of Biochemistry, University of Otago, P.O. Box 56, Dunedin 9054, New Zealand; e-mail: warren.tate@stonebow.otago.ac.nz; fax: 0064-3-4797866.

Article published online ahead of print. Article and publication date are at <http://www.rnajournal.org/cgi/doi/10.1261/rna.1970310>.

Eubacterial RF1 and RF2 bound with a stop codon on the ribosome showed that a second structural element, the tip of helix $\alpha 5$, contacted the first base of the stop codon (Petry et al. 2005; Korostelev et al. 2008; Laurberg et al. 2008; Weixlbaumer et al. 2008). This structure packs against the Watson–Crick edge of the uridine in the first position of the stop codon, revealing the reason why discrimination at this position is remarkably accurate (Freistroffer et al. 2000). Bioinformatic analysis has also identified differentially conserved positions determining specificity in both the recognition loop and helix $\alpha 5$ region in eubacterial RF1 and RF2 (Oparina et al. 2005). These residues could also contribute to the functional specificities of RF1 and RF2.

Mitochondria contain their own protein synthesis machinery for the production of polypeptides encoded by mitochondrial DNA (Sprengli et al. 2004; Rorbach et al. 2007). For example, in mammals mtDNA encodes 13 proteins and the required complement of mt-tRNAs and mt-rRNAs, but protein synthesis relies on importing ribosomal proteins and translational factors synthesized in the eukaryotic cytosol. Mitochondria often use variations to the standard genetic code (Osawa et al. 1992; Knight et al. 2001; Santos et al. 2004; Soll and RajBhandary 2006; Lekomstev 2007), with the most common deviation being the reassignment of UGA to a codon specifying tryptophan. The genomes of organisms that have reassigned UGA to tryptophan require only a single gene encoding a mitochondrial RF (MRF1) to decode the two remaining standard stop codons, UAA and UAG. *Saccharomyces cerevisiae* MRF1 exhibits a high sequence similarity to eubacterial RF1 (Pel et al. 1992b), consistent with mitochondria originating from an endosymbiotic α -proteobacterial ancestor, and a PXT motif, indicating that it is indeed a eubacterial RF1 homolog. As expected, recombinant *S. cerevisiae* MRF1 recognized UAA and UAG in vitro, demonstrating activity on both *S. cerevisiae* mitochondrial and *Escherichia coli* ribosomes (Askarian-Amiri et al. 2000). An equivalent vertebrate RF purified from rat mitochondria also recognized UAA and UAG (Lee et al. 1987).

In vertebrate mitochondria, apart from those of rodents, not only is UGA reassigned to tryptophan but also the arginine codons AGA and AGG are reassigned as stop codons complementing UAA and UAG (Barrell et al. 1979; Anderson et al. 1981). This implies the need for a second decoding RF with altered specificity. Two mitochondrial RF genes have since been identified in the human genome, *MTRF1* (*mtRF1/HMRF1*) and *MTRF1L* (*mtRF1a/HMRF1L*) (Zhang and Sprengli 1998; Soleimanpour-Lichaei et al. 2007; Nozaki et al. 2008), and the gene products of both have been shown to localize to mitochondria (Soleimanpour-Lichaei et al. 2007; Nozaki et al. 2008). In vitro release assays have shown that recombinant human MTRF1L containing a PXT tripeptide motif similar

to other RF1-type factors is the classical RF and the homolog to that identified in rat liver (Lee et al. 1987), recognizing the standard stop codons and not AGA or AGG (Soleimanpour-Lichaei et al. 2007; Nozaki et al. 2008). Vertebrate MTRF1 is an enigma and appears to be a novel type of RF. It was inferred to be the missing mitochondrial RF needed for recognizing the reassigned stop codons AGA and AGG (Soleimanpour-Lichaei et al. 2007), but in vitro human recombinant MTRF1 had no release activity on bacterial ribosomes with any codons (Soleimanpour-Lichaei et al. 2007; Nozaki et al. 2008).

We have carried out a detailed bioinformatic analysis of the vertebrate mitochondrial RFs and detailed interpretation of the recent high-resolution structures of eubacterial RF termination complexes, and coupled this analysis with experiments aimed at mimicking MTRF1 activity to show that MTRF1 has the characteristics in its external loops to recognize the reassigned stop codons AGA and AGG.

RESULTS AND DISCUSSION

How do MTRF1L and MTRF1 differ?

To gain an understanding of the differences between the classical RF1-type factor MTRF1L and the unassigned MTRF1, we performed alignments of vertebrate MTRF1L and MTRF1 sequences. These identified a number of inaccurate annotations in predicted MTRF1 and MTRF1L protein sequences obtained from the NCBI databases (see Supplemental Information).

Genome sequence data for vertebrates that are well annotated contain both distinct mitochondrial RF genes, suggesting that each is essential. One of the groups is the classical mitochondrial MTRF1, while the second group encompasses the novel human mtRF1/HMRF1 and is annotated in the NCBI database as MTRF1. There are rare examples with one gene type only; for example, in the pig (*Sus scrofa*) a *MTRF1L* to decode the standard stop codons has not been identified, although it is likely to be essential as in yeast (Pel et al. 1992a). The recently published platypus (*Ornithorhynchus anatinus*) genome (Warren et al. 2008) has no identified *MTRF1*, although AGG is used once for stop in platypus mitochondria (Janke et al. 1996).

Aligning the two vertebrate mitochondrial RF types showed that there was a high degree of sequence homology between the predicted vertebrate MTRF1L and MTRF1 proteins, particularly in the regions that form domains 2, 3, and 4 as defined in the eubacterial RFs (Fig. 1). This implies that MTRF1 is highly likely to be functionally active on ribosomes. Two significant regions of sequence difference between MTRF1L and MTRF1 corresponded to the two structural elements identified as critical for stop codon recognition in the eubacterial RFs, the tip of helix $\alpha 5$ and the recognition loop (see Fig. 1).

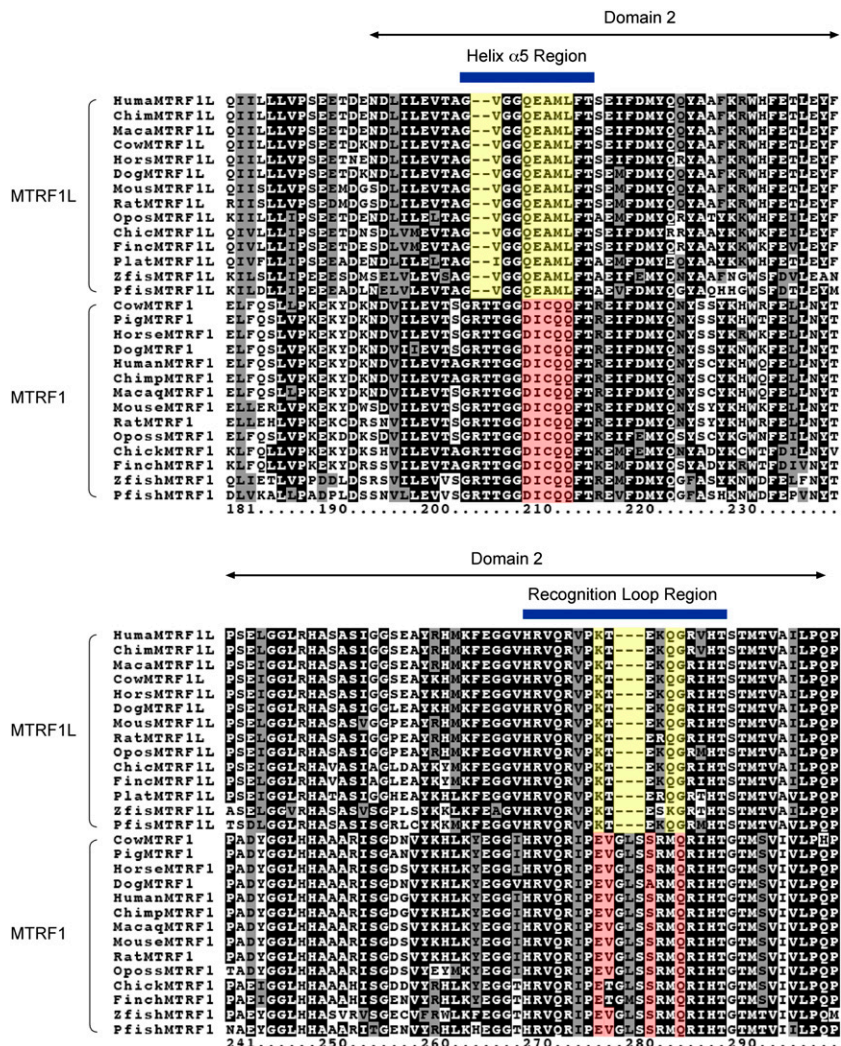


FIGURE 1. Sequence alignment of vertebrate MTRF1L and MTRF1 proteins. Sequence alignment of known vertebrate MTRF1L and MTRF1 proteins showing the region equivalent to domain 2 of the eubacterial RFs. Identical amino acids are white (black background), similar amino acids are black (gray background), and different residues are black (white background). The blue lines show the location of the helix $\alpha 5$ and the recognition loop regions. Residues within these regions that differ between MTRF1L and MTRF1 are highlighted in yellow for MTRF1L and red for MTRF1.

Did vertebrate MTRF1 evolve by gene duplication from the ancestral MRF1 or from an ancestral eubacterial RF2-like gene?

Aligning eubacterial RF1 and RF2 and vertebrate MTRF1L and MTRF1 protein sequences (see Supplemental Fig. 1) showed that vertebrate MTRF1 and MTRF1L have higher sequence similarity to eubacterial RF1 than to eubacterial RF2 sequences. A phylogenetic analysis on a selection of eubacterial RF1, nonvertebrate MRF1, vertebrate MTRF1L, and vertebrate MTRF1 proteins (Fig. 2) showed that the MTRF1 proteins branch from within the MRF1 and MTRF1L proteins of the Deuterostomes, suggesting that the MTRF1L paralog evolved during early vertebrate evolution.

The high level of sequence similarity between MTRF1 and other RF1-type factors, together with the phylogenetic analysis, suggests that vertebrate MTRF1 arose from a gene duplication of an ancestral MRF1 gene in an early member of the vertebrate lineage rather than from an ancestral RF2 gene (mitochondrial or eubacterial) of the type found today in plant mitochondria that still use UGA for stop. Intriguingly, vertebrate genomes are scattered with MRF pseudogenes encoding fragments of mitochondrial RFs. Our hypothesis is that one of the duplicated MTRF genes would have acquired new codon recognition specificities by subsequent mutations to the helix $\alpha 5$ and recognition loop regions, losing recognition of standard stop codons but allowing it to recognize AGA and AGG. The other gene would have retained its specificity for UAA and UAG.

The tip of helix $\alpha 5$ of MTRF1 has unique features

The helix $\alpha 5$ tip in RFs comprises the loop between strand $\beta 1$ and helix $\alpha 5$, and the start of helix $\alpha 5$ (Fig. 3A). It is highly conserved between eubacterial RF1 and RF2. The loop between strand $\beta 1$ and helix $\alpha 5$ contains three highly conserved glycines (Fig. 3A, cyan, right panel): one at the start of the loop (G118 in *E. coli* RF1; G135 in *E. coli* RF2) and two at the end (Fig. 3A, G120, lower panel; 121 in *E. coli* RF1; G137, 138 in *E. coli* RF2) that facilitate a tight turn at the start of helix $\alpha 5$.

High-resolution structures of both eubacterial RFs in termination complexes with the ribosome showed that the stop codon first base forms a hydrogen bond (N3 of U1) with the backbone carbonyl group of G120/G137 of this region, and the O4 of U1 forms a hydrogen bond with the backbone amide of E123/E140 in the $\alpha 5$ helix (shown in Fig. 3A, lower panel; Korostelev et al. 2008; Laurberg et al. 2008; Weixlbaumer et al. 2008). These hydrogen bonds are only possible with U in the first codon position and discriminate against a purine at this site. E123/E140 is highly conserved in both RF1 and RF2 on the first turn of helix $\alpha 5$ (see Fig. 3A, right panel). The RF1 specificity-determining positions L126 and F127, while on the second turn of this helix, are on the same helix face as E123. The loop between strand $\beta 1$ and helix $\alpha 5$ of the vertebrate MTRF1 proteins is larger than that of vertebrate MTRF1L,

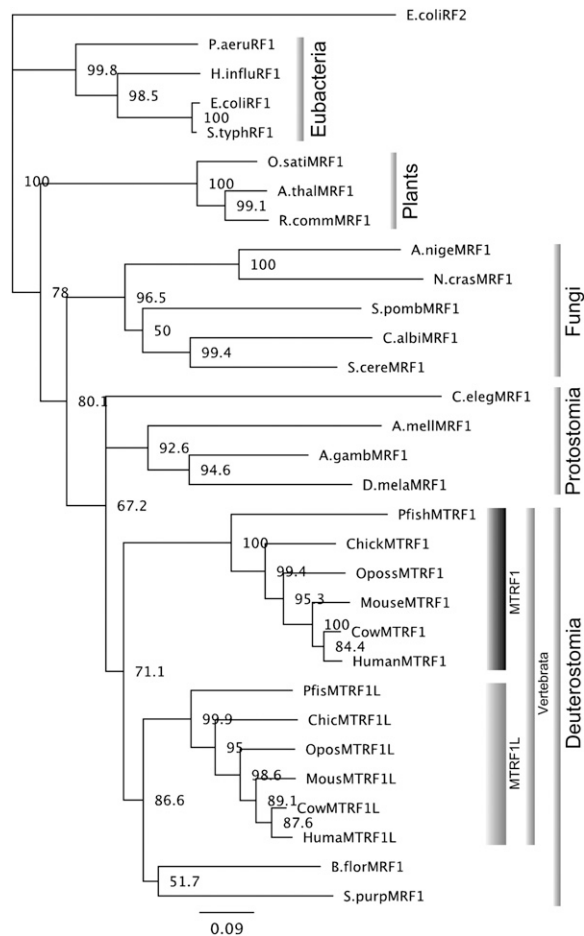


FIGURE 2. A phylogenetic tree of eubacterial RF1, RF2, nonvertebrate MRF1 and vertebrate MTRF1L and MTRF1 proteins. *E. coli* RF2 was used as an out group; 1000 neighbor joining trees were generated to produce bootstrap values. Major taxonomic divisions are shown.

eubacterial RF1, and nonvertebrate MRF1 proteins by a two-amino acid insertion, RT (see Figs. 1, 3B). The insertion is most likely tolerated structurally because it is within the external loop and does not appear to disrupt any secondary structure (see Fig. 3A, right panel). MTRF1 also has critical amino acid substitutions at the start of helix $\alpha 5$ (see Figs. 1, 3B). For example, the key glutamic acid (E123) of eubacterial RF1 (forming the hydrogen bond with O4 of U1 of the stop codon) and E141 of hMTRF1L (Fig. 3B) is not conserved in hMTRF1 but is substituted with isoleucine (I199). Additionally, the RF1 specificity determining position found at L126 of eubacterial RF1 and at L144 in hMTRF1L is substituted with Q in hMTRF1 (Fig. 3B). This strongly suggests that while the vertebrate mitochondrial MTRF1L proteins recognize the invariant U in the standard stop codons in a similar manner to that of eubacterial RF1 and RF2, MTRF1 almost certainly has this recognition disturbed.

Given these changes, could vertebrate MTRF1 proteins recognize a two-ringed purine base at the first position of the stop codon? We suggest that the two-amino acid

insertion in the loop between strand $\beta 1$ and helix $\alpha 5$ repositions the glycine residues in the loop equivalent to G120/G121 (G196/G197 in human MTRF1), allowing a larger purine base to be accommodated as would be required to recognize AGA or AGG. The other sequence differences in the vertebrate MTRF1 proteins may also allow recognition of A at the first position and contribute to the unique requirement for the RF decoding AGA and AGG to recognize only G at the second position.

The recognition loop of MTRF1 also has unique features

The recognition loop of vertebrate MTRF1 proteins is also different from other RF1-type RFs, both in size and in key conserved amino acids. Critically, while the MTRF1L loop retains a PXT tripeptide motif found in eubacterial RF1 and nonvertebrate MRF1 proteins (Figs. 1, 3C; Soleimanpour-Lichaei et al. 2007; Nozaki et al. 2008), the loop of MTRF1 instead has an expanded hexapeptide PEVGLS motif. The conserved T of the classical RF1-specific PXT motif is substituted mainly as valine (V266 in hMTRF1). This makes it unlikely that the same contacts observed for the RF1 structure would be maintained for MTRF1.

The similarities between the recognition loops of vertebrate MTRF1L proteins and eubacterial RF1 and nonvertebrate MRF1 proteins extend beyond the PXT tripeptide motif. The residues important for stop codon recognition are conserved or have conservative substitutions. For example, the specificity determining position E191 (E209 in hMTRF1L), I196 (V214 in hMTRF1L), and V184 and Q185 in strand $\beta 4$ (V202/Q203 in hMTRF1L) (see Fig. 3C). In contrast, the recognition loop of vertebrate MTRF1 proteins is three amino acids longer (shown as a GLS insertion; see Figs. 1, 3C). This insertion can be tolerated structurally because it is external and does not disrupt any secondary structure. We suggest that the insertion (as for the insertion in the $\alpha 5$ helix region) may also be important for accommodating the larger purine base at the first position of the stop codon, and critical for recognition of G only in the second position. This is a unique requirement of MTRF1. The sequences of the two loops (MTRF1 versus MTRF1L) are also substantially different; for example, there is no obvious position equivalent to the RF1/RF2 E191/D208 specificity determining position in the MTRF1 loop.

All stop codons used in vertebrate mitochondria (UAA/UAG and AGA/AGG) allow either A or G at the third position. The recognition loop residues important for determining this relaxed specificity for eubacterial RF1 are Q185 and I196 (Fig. 3A, lower panel). Significantly, in vertebrate MTRF1 these residues have been retained (Q261 and I275) (Fig. 3C). This is strongly suggestive that vertebrate MTRF1 proteins can recognize the third base of the stop codon in a similar manner to classical RF1-type

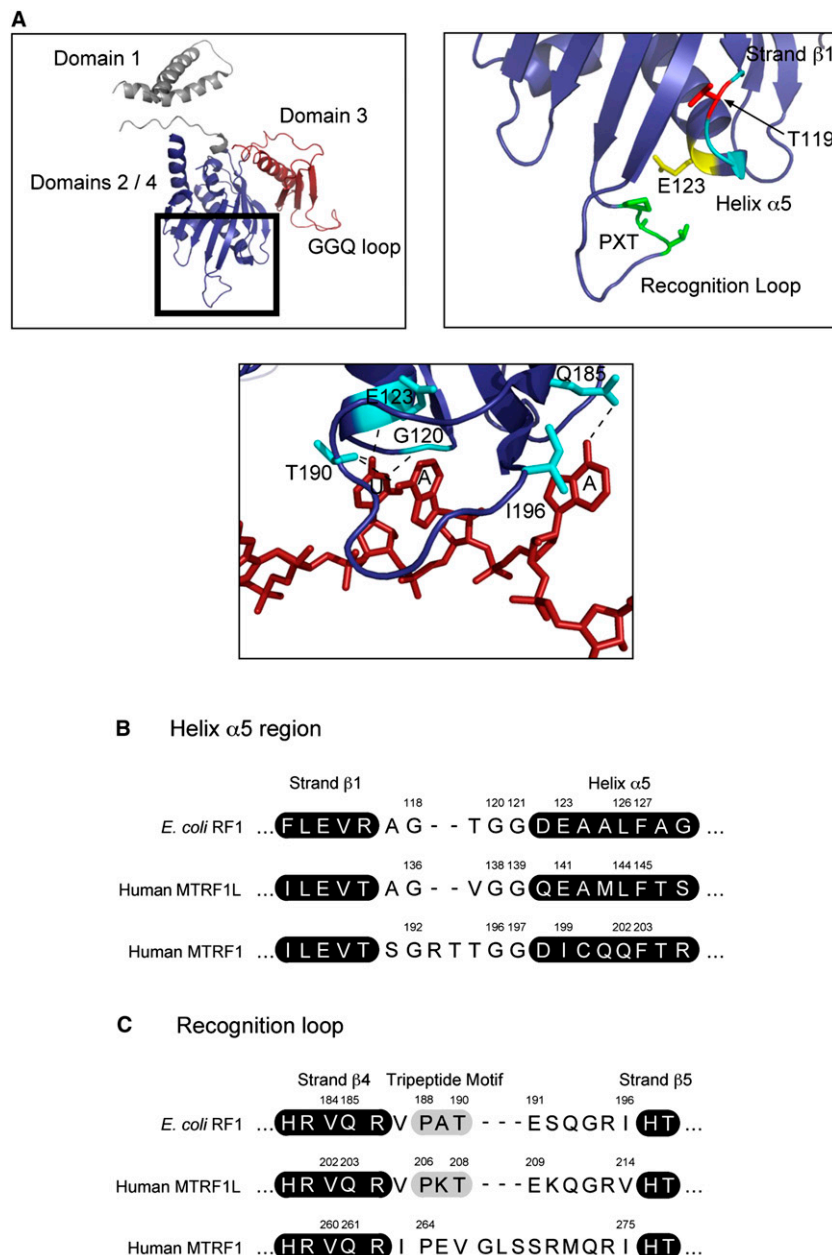


FIGURE 3. The helix $\alpha 5$ and recognition loop regions of the eubacterial and vertebrate mitochondrial RF1 proteins. (A) The helix $\alpha 5$ region of *E. coli* RF1 is based on the *E. coli* RF1 structure (Protein Data Bank [PDB] file 2B3T) (Graille et al. 2005). (Left) The location of the helix $\alpha 5$ and recognition loop regions are boxed on the solution structure of RF1. (Right) The boxed region expanded. Shown are the three highly conserved glycine residues in the loop between strand $\beta 1$ and helix $\alpha 5$ (cyan), the location of the highly conserved E123 (glutamic acid, yellow) residue that forms interactions with the first base of the stop codon, residue T119 (red) at the site of the two amino acid insertion in vertebrate MTRF1 proteins, and the RF1-specific PXT tripeptide motif (green). (Lower panel) The interactions of RF residues with UAA based on the *Thermus thermophilus* RF1 structure (PDB file 3D5A) (Laurberg et al. 2008) bound to the ribosome. T190 of the PXT motif is shown along with recognition loop residues I196 and adjacent Q185, as well as G120 of helix $\alpha 5$. (B) The helix $\alpha 5$ regions of *E. coli* and human mitochondrial RF1 proteins. Strand $\beta 1$ and helix $\alpha 5$ are shown (black). Key conserved residues and specificity determining positions are numbered. (C) The recognition loops of *E. coli* and human mitochondrial RF1 proteins. Strands $\beta 4$ and $\beta 5$ are shown (black). The PXT tripeptide motifs of *E. coli* RF1 and human MTRF1L are highlighted (gray). Key conserved residues and specificity determining positions are numbered. Note that the ClustalW sequence alignment program has determined the position of the insertion (the position of the insertion in the alignment has been used as a boundary for the PEVGLS motif).

proteins, in contrast to the recognition of the first and second bases.

Can mimicking the MTRF1 helix $\alpha 5$ region and the recognition loop in *E. coli* RF1 change codon specificity?

Recombinant human MTRF1 has no activity on bacterial ribosomes, and as yet, developing a successful vertebrate mitochondrial in vitro termination assay has not proven possible. As an alternative strategy to get indirect functional data, we constructed three *E. coli* RF1 variants containing either the mitochondrial $\alpha 5$ region (mit $\alpha 5$ /wt rl), or the mitochondrial recognition loop (wt $\alpha 5$ /mit rl), or both regions (mit $\alpha 5$ /mit rl) (Fig. 4A). These variants were tested both without and with specific codons for codon-dependent peptidyl-tRNA hydrolysis activity in vitro (Fig. 4B; Tate and Caskey 1990). Exchanging the $\alpha 5$ helix tip alone (mit $\alpha 5$ /wt rl) created an inactive factor. In contrast, exchanging the recognition loop alone (wt $\alpha 5$ /mit rl) retained a significant ($P \leq 0.01$) amount of release activity with UAA. The increased size of the anticodon loop with a wt $\alpha 5$ clearly did not fully prevent recognition of the first position U but significantly decreased recognition efficiency. There was also a small stimulation with UAA and a hint of activity stimulation by the nonstandard codon AGA with the double-loop variant (mit $\alpha 5$ /mit rl). This was explored further using a range of sense codons beginning with A (Fig. 4C). AAG showed significant stimulation ($P \leq 0.05$), but all codons tested beginning with A caused a small stimulation of release activity with the double mit $\alpha 5$ /mit rl variant that was not seen with native RF1. This double-loop variant also triggered somewhat more peptidyl-tRNA hydrolysis activity independent of any codon compared with RF1. Interestingly, there was little stimulation with UAG compared with AAG (differing only with U to A in the first position) (Fig. 4D). UGA, which is not normally recognized by RF1, only has a small gain in recognition in the double-loop variant, lower

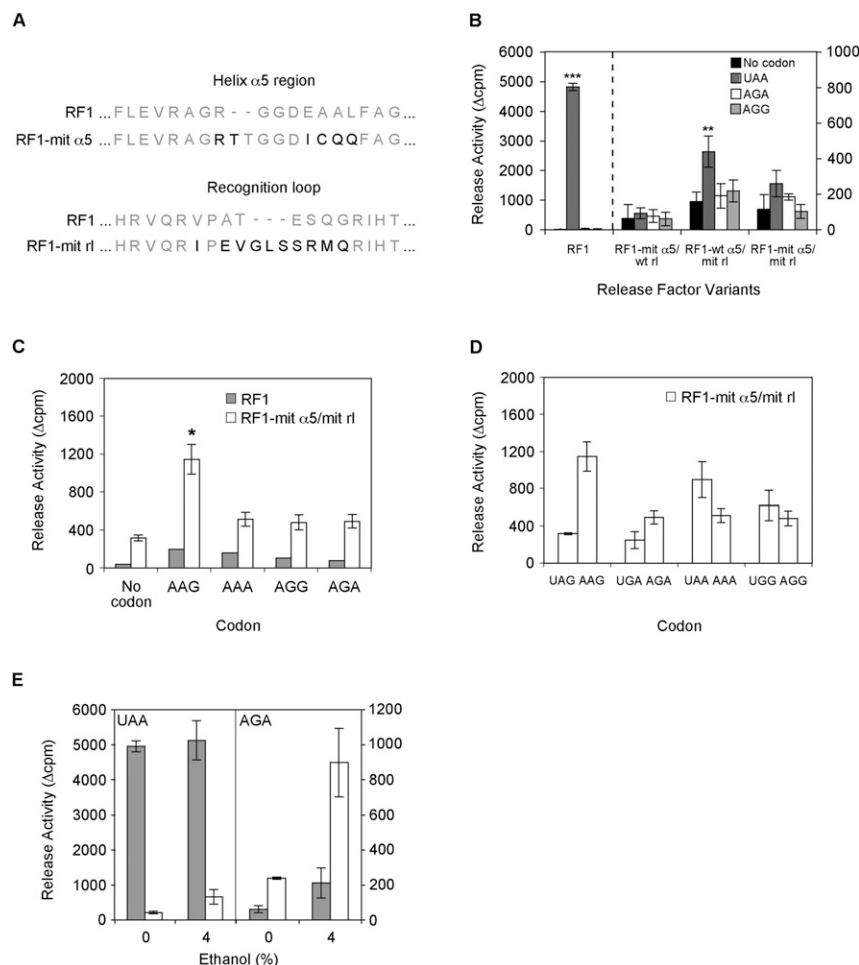


FIGURE 4. A comparison of the activities of the RF1 variants containing the hMTRF1 recognition loop and/or the helix $\alpha 5$ region at standard and nonstandard stop codons. (A) The sequence differences between the classical RF1-type factors (represented by eubacterial RF1) and MTRF1 in the helix $\alpha 5$ and recognition loop regions. (B) The codon-dependent peptidyl-tRNA hydrolysis activities of bacterial RF1 and RF1 containing MTRF1 helix $\alpha 5$ and/or the recognition loop regions with no codon and UAA, AGA, and AGG codons. The RF variants were assayed in triplicate with each experiment repeated twice to ensure consistent results between assays. (C,D) A comparison of the codon-dependent peptidyl-tRNA hydrolysis activities of the RF1-mit $\alpha 5$ /mit rl variant with a range of codons. The RF variant was assayed in duplicate. (E) Ethanol stimulation of UAA and AGA codon-dependent peptidyl-tRNA hydrolysis activities with native RF1 and the RF1-mit $\alpha 5$ /mit rl variant containing MTRF1 recognition determinants in the helix $\alpha 5$ and recognition loop regions. Release activity represents Δ cpm of [3 H]fMet released. Error bars, SD of the mean (B–D); SEM (E). Statistical significances, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ (Student’s two-tailed t -test), were determined by comparison to the activity of each factor with no codon.

than with AGA. In contrast, UAA still retains $\sim 10\%$ of its original activity (see Fig. 4B), but there is also gain of activity with AAA (Fig. 4D). UGG is excluded from recognition by RF1 as a codon with consecutive second and third position Gs not tolerated. Interestingly, the double-loop variant gains a small amount of activity with both UGG and AGG. A factor that recognizes AGG as “stop” must have the ability to recognize G in both the second and third positions, and so low-level release activity stimulation of the mit $\alpha 5$ /mit rl variant with UGG in this artificial chimera is not surprising.

When partial in vitro termination assays were developed (for description, see Tate and Caskey 1990) ethanol at 10%–20% (v/v) was used to separate the peptidyl-tRNA hydrolysis step from codon recognition. At lower concentrations, however, ethanol can enhance the codon-dependent activity of eubacterial RFs by stabilizing the termination complex in vitro. We compared water or ethanol (4%) as the solvents with native RF1 or the double-loop mit $\alpha 5$ /mit rl variant (Fig. 4E) and with UAA and AGA to find out whether peptidyl-tRNA hydrolysis activity could be enhanced. As expected, with UAA and ethanol RF1 activity was high, but the double variant had no activity above that induced without codon. In contrast, with AGA the variant showed an enhanced response in ethanol, whereas RF1 did not respond significantly to this codon above the activity seen in its absence. These results reinforce the conclusion that the helix $\alpha 5$ and recognition loops of MTRF1 can accommodate A in the first position of the stop codon, whereas the classical factor cannot.

Could the function of MTRF1 have evolved to recognize stop codons beginning with A and with G in the second position?

The vertebrate MTRF1 is deduced from our phylogenetic analysis to have evolved relatively late and acquired a new function on the 55S mitoribosome with no prior functional history on eubacterial ribosomes. The factor has clearly retained the key structural domains of a functional RF and has high conservation relationships with eubacterial RFs, nonvertebrate MRFs, other vertebrate RFs, and MTRF1L (see Fig. 1). This provides compelling evidence that MTRF1 has retained a ribosomal function. The fact that the two key elements, the helix $\alpha 5$ tip and the recognition loop that we now know are intimately involved in codon recognition, have been remodeled (see Fig. 3A) leads to an attractive hypothesis that the new factor has escaped the constraints of the classical RFs and their URR (excluding UGG) restrictions and has acquired new specificities. Of interest to the evolution of two factors from one in vertebrate mitochondria, is the recent report of a second elongation

factor, EF-G2mt, from mammalian mitochondria that has evolved as an exclusive recycling factor, with the other elongation factor, EF-G1, specifically catalyzing translocation, in contrast to the eubacterial ancestor that carries out both functions (Tsuboi et al. 2009).

The structures of the eubacterial termination complexes have identified which residues of the bacterial factors RF1 and RF2 are involved in interactions with stop codons. Even with loops of the same size, these two factors utilize their sequence diversity to create unique patterns of interactions with individual bases. Some interactions are in common, for example, from the helix $\alpha 5$ tip for first base recognition (Fig. 5A). Other interactions are different. In RF1 it is the third amino acid T190 in the PAT tripeptide motif of the recognition loop that recognizes the second base of the stop codon (see Figs. 3A, bottom panel, and

5B), but in RF2 it is the first amino acid S in the SPF motif (Fig. 5B). Third base recognition uses equivalent parts of the loop. Key residues in RF1 loops and their geometries have been modified in MTRF1, resulting in increased size that disrupts the recognition properties for the first two bases of the stop codon (Fig. 5B). In contrast, third base recognition determinants in the recognition loop are retained in MTRF1 with the same sequence (HRVQR) as in eubacterial RF1. These patterns could be consistent with a switch in specificities between RF1 and MTRF1 from URR to AGR. However, without a definitive *in vivo* or *in vitro* assay based on vertebrate mitochondrial components, it is not yet possible to generate directly the experimental evidence in support of the contention that MTRF1 is the factor that recognizes AGA and AGG on vertebrate mitochondrial ribosomes. Nevertheless, the fact that MTRF1 codon recognition loops function as part of *E. coli* RF1 on bacterial ribosomes has provided clear evidence for a loss of recognition of U in the first position of the stop codon, and suggests a preferred recognition of AXX over UXX, shown best with AAG versus UAG. Ethanol, a known stimulant of codon-directed events, could enhance recognition of “A” in AGA but did not improve the recognition of the “U” in UAA by the MTRF1 codon recognition loops. This is indirect support for the hypothesis that MTRF1 is the mitochondrial RF responsible for recognizing the reassigned stop codons, AGA and AGG. Unequivocal proof for this recognition or another mitochondrial ribosomal function awaits a system that will respond to native MTRF1.

While this paper was under revision, a brief study reported that human mitochondria could avoid AGA and AGG “hungry” codons by frameshifting -1 to result in a UAG stop codon where standard termination of protein synthesis could occur (Temperley et al. 2010). If this mechanism were generally applicable, then a specific factor for recognizing the nonstandard stop codons would not be necessary. This still leaves open the question as to the function of the highly conserved RF-like protein MTRF1. Intriguingly, many vertebrate species have the putative stop codon AGA at the end of the mitochondrial *cytb* gene with an A immediately before the AGA (or AGG). For this gene a simple -1 frameshift to a stop would not be possible. Similarly, the nonstandard termination codons AGA or AGG at the end of the *COI* gene is preceded by A or T or G among different species and the *ND6* gene also has a variable sequence preceding AGA. MTRF1 seems the likely factor that mediates termination under these circumstances.

MATERIALS AND METHODS

Bioinformatic analysis

Protein sequences were identified from the NCBI databases, downloaded in FASTA format, and aligned using ClustalW (BioManager by ANGIS) (Thompson et al. 1994). Aligned

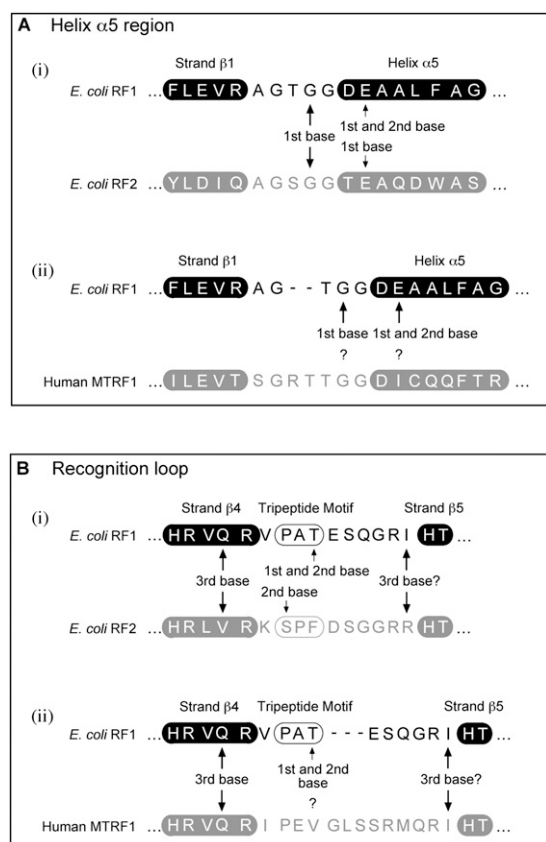


FIGURE 5. The regions and residues of RFs that interact with stop codons. (A) The helix $\alpha 5$ region. (i) A comparison between *E. coli* RF1 and RF2 showing first and second base recognition positions. (ii) A comparison of *E. coli* RF1 and MTRF1 showing the equivalent residues of MTRF1 at the sites of first and second base recognition. (B) The recognition loop. (i) A comparison of the residues of *E. coli* RF1 and RF2 important for first, second, and third base recognition. (ii) A comparison of *E. coli* RF1 with MTRF1 showing the equivalent positions in MTRF1 at sites of recognition of the individual bases. The tripeptide motifs are circled. Residues interacting directly with the stop codon and important for determining the different stop codon specificities of RF1 and RF2 are indicated.

sequences were colored using the program BoxShade to identify identical, similar, and different residues (BioManager by ANGIS).

Phylogenetic comparisons

Neighbor joining trees were generated using Geneious (v4.6) software (Drummond et al. 2009). *E. coli* RF1 or RF2 were used as outgroups; 1000 trees were generated to produce bootstrap values. Protein parsimony trees were calculated using the Protpars program in PHYLIP (BioManager by ANGIS) (Felsenstein 1989). Protein maximum likelihood trees were calculated using PROTML (BioManager by ANGIS) (Adachi and Hasegawa 1996).

Bacterial strains

All cloning was carried out in the *E. coli* strain DH5 α (*supE44* Δ *lacU169* [ϕ 80*lacZ* Δ M15] *hsdR17* *recA1* *endA1* *gyrA96* *thi-1* *relA1*) (Hanahan 1983). Expression of the RF protein variants from the pET3a vectors utilized the *E. coli* strain BL21(DE3) pLysS (*hsdS* *gal* [λ *clts* 857 *ind1* *Sam7* *nin5* *lacUV5*-T7 gene 1] pLysS [*cam*^R]) (Studier et al. 1990). The BL21(DE3) pLysS strain was maintained in 30 μ g/mL chloramphenicol due to the presence of the pLysS plasmid.

Construction of expression vectors

Standard cloning and PCR mutagenesis protocols were used to construct and purify all vector sequences. Construction of the RF1 expression vectors pETRF1 and pTGRF1 has been described previously (Wilson et al. 2000; Mansell et al. 2001). RF1 variants were made by introducing mutations into the *prfA* gene in both the pETRF1 and pTGRF1 vectors using a two-step PCR strategy (Sarkar and Sommer 1990) as described previously (Scarlett et al. 2003). The sequences of the PCR primers used are listed in the Supplemental Table. All reactions were carried out using Platinum Taq High Fidelity (Invitrogen Corporation). The first step used a mutagenic primer and a reverse cloning primer (either pET-T7r or pTGseqr dependent on the template vector) to generate a 3' fragment of RF1 containing the desired mutation. In the second step, a forward primer (pET-T7f or pTGseqf) and 180–360 ng of purified PCR1 product were used to amplify the full-length RF1 gene to generate the variant. This process was repeated for each of the individual site-directed mutations. The RF2 expression plasmids pET21aRF2* T246A and pTGRF2* T246A were constructed in a similar manner. The pET21aRF2* T246A and pTGRF2* T246A plasmids contain the *prfB* gene with a T246A mutation, which alleviates the decrease in specific activity observed when RF2 is overexpressed (Uno et al. 1996; Wilson et al. 2000).

RF protein expression

For expression of the RF proteins, individual freshly transformed colonies were inoculated into LB medium supplemented with appropriate antibiotics (100 μ g/mL ampicillin and 30 μ g/mL chloramphenicol). The overnight culture was then used to inoculate fresh selective medium (1% [v/v]) and grown at 37°C with shaking to an OD₆₀₀ of 0.4–0.6. Protein expression was induced using 1 mM (25 μ g/mL) isopropyl- β -D-thiogalactopyranoside. The cultures were then grown for a further 3 h before the cell

pellets were harvested and frozen at -80°C in preparation for protein purification.

RF purification

Crude protein preparations were obtained from the frozen pellets essentially as described previously (Tate and Caskey 1990; Wilson et al. 2000). The proteins were purified to homogeneity on a ResourceQ anion exchange column (Amersham Pharmacia Biotech) using the AKTA purifier chromatography system (Amersham Pharmacia Biotech) according to the manufacturer's instructions. The concentration of dialyzed protein fractions was measured with a NanoDrop 1000 Spectrophotometer (Nanodrop Technologies). The extinction coefficients (21,430 M⁻¹cm⁻¹RF1, 43,890 M⁻¹cm⁻¹RF2) calculated using the program ProtParam available on the ExPASy server, and molecular weights (40,517.3 Da RF1; 41,250.7 Da RF2) were used to calculate the concentrations of RF1 and RF2.

In vitro peptidyl-tRNA hydrolysis assay

MRE600 ribosomes, codons and f[³H]Met-tRNA^{Met} used in the assay were prepared according to Tate and Caskey (1990). The assay measures the release of a model peptide, f[³H]Met, from the P-site tRNA (Caskey et al. 1971) and was performed as described previously (Wilson et al. 2000). The assay substrate (f[³H]Met-tRNA^{Met}-AUG-ribosome complex) was formed by incubating 5 pmol of 70S ribosomes with 2.5 pmol of f[³H]Met-tRNA^{Met}, and 250 pmol AUG in 1 \times stage I buffer [20 mM Tris-HCl at pH 7.4, 10 mM Mg(OAc)₂, 150 mM NH₄Cl] at 30°C for 20 min. This complex was then incubated with 20 pmol RF and 250 pmol of stop codon in 1 \times RF buffer [50 mM Tris-HCl at pH 7.2, 30 mM Mg(OAc)₂, 75 mM NH₄Cl] for 15 min at 24°C. The reaction was stopped by the addition of 0.1 M HCl, and released f[³H]Met was extracted into 1 mL of ethyl acetate by vortexing for 15 sec. The phases were separated by centrifugation (13,000g for 1 min), and 750 μ L of the organic phase was mixed with 3 mL of Optiphase HiSafe 2 scintillation cocktail (Perkin Elmer Life and Analytical Sciences, USA) and analyzed in a scintillation counter. The background was subtracted by using data from an assay excluding the RF.

SUPPLEMENTAL MATERIAL

Supplemental material can be found at <http://www.rnajournal.org>.

ACKNOWLEDGMENTS

This work was supported by grants from the Marsden Fund of New Zealand (grant no. UOO210 to W.P.T. and E.S.P, and grant no. UOO0701 to W.P.T.). We thank Peter Stockwell and Craig Marshall for help with the phylogenetic analyses and Sue Cutfield for helpful discussions.

Received October 21, 2009; accepted March 1, 2010.

REFERENCES

Adachi J, Hasegawa M. 1996. Molphy version 2.3. Programs for molecular phylogenetics based on maximum likelihood. In

- Computer Science monographs No. 28*. Institute of Statistical Mathematics, Tokyo, Japan.
- Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, et al. 1981. Sequence and organization of the human mitochondrial genome. *Nature* **290**: 457–465.
- Askarian-Amiri ME, Pel HJ, Guevremont D, McCaughan KK, Poole ES, Sumpter VG, Tate WP. 2000. Functional characterization of yeast mitochondrial release factor 1. *J Biol Chem* **275**: 17241–17248.
- Barrell BG, Bankier AT, Drouin J. 1979. A different genetic code in human mitochondria. *Nature* **282**: 189–194.
- Caskey CT, Beaudet AL, Scolnick EM, Rosman M. 1971. Hydrolysis of fMet-tRNA by peptidyl transferase. *Proc Natl Acad Sci* **68**: 3163–3167.
- Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A. 2009. Geneious v4.6. <http://www.geneious.com/>.
- Felsenstein J. 1989. PHYLIP—phylogeny inference package (version 3.2). *Cladistics* **5**: 164–166.
- Freistroffer DV, Kwiatkowski M, Buckingham RH, Ehrenberg M. 2000. The accuracy of codon recognition by polypeptide release factors. *Proc Natl Acad Sci* **97**: 2046–2051.
- Graille M, Heurgue-Hamard V, Champ S, Mora L, Scrima N, Ulryck N, van Tilbeurgh H, Buckingham RH. 2005. Molecular basis for bacterial class I release factor methylation by PrmC. *Mol Cell* **20**: 917–927.
- Hanahan D. 1983. Studies on transformation of *Escherichia coli* with plasmids. *J Mol Biol* **166**: 557–580.
- Ito K, Uno M, Nakamura Y. 2000. A tripeptide “anticodon” deciphers stop codons in messenger RNA. *Nature* **403**: 680–684.
- Janke A, Gemmell NJ, Feldmaier-Fuchs G, von Haeseler A, Paabo S. 1996. The mitochondrial genome of a monotreme—the platypus (*Ornithorhynchus anatinus*). *J Mol Evol* **42**: 153–159.
- Knight RD, Freeland SJ, Landweber LF. 2001. Rewiring the keyboard: Evolvability of the genetic code. *Nat Rev Genet* **2**: 49–58.
- Korostelev A, Asahara H, Lancaster L, Laurberg M, Hirschi A, Zhu J, Trakhanov S, Scott WG, Noller HF. 2008. Crystal structure of a translation termination complex formed with release factor RF2. *Proc Natl Acad Sci* **105**: 19684–19689.
- Laurberg M, Asahara H, Korostelev A, Zhu J, Trakhanov S, Noller HF. 2008. Structural basis for translation termination on the 70S ribosome. *Nature* **454**: 852–857.
- Lee CC, Timms KM, Trotman CN, Tate WP. 1987. Isolation of a rat mitochondrial release factor. Accommodation of the changed genetic code for termination. *J Biol Chem* **262**: 3548–3552.
- Lekomstev SA. 2007. Non-standard genetic codes and translation termination. *Mol Biol (Mosk)* **41**: 964–972.
- Mansell JB, Guevremont D, Poole ES, Tate WP. 2001. A dynamic competition between release factor 2 and the tRNA(Sec) decoding UGA at the recoding site of *Escherichia coli* formate dehydrogenase H. *EMBO J* **20**: 7284–7293.
- Nozaki Y, Matsunaga N, Ishizawa T, Ueda T, Takeuchi N. 2008. HMRFL is a human mitochondrial translation release factor involved in the decoding of the termination codons UAA and UAG. *Genes Cells* **13**: 429–438.
- Oparina NJ, Kalinina OV, Gelfand MS, Kisselev LL. 2005. Common and specific amino acid residues in the prokaryotic polypeptide release factors RF1 and RF2: Possible functional implications. *Nucleic Acids Res* **33**: 5226–5234.
- Osawa S, Jukes TH, Watanabe K, Muto A. 1992. Recent evidence for evolution of the genetic code. *Microbiol Rev* **56**: 229–264.
- Pel HJ, Rep M, Grivell LA. 1992a. Sequence comparison of new prokaryotic and mitochondrial members of the polypeptide chain release factor family predicts a five-domain model for release factor structure. *Nucleic Acids Res* **20**: 4423–4428.
- Pel HJ, Maat C, Rep M, Grivell LA. 1992b. The yeast nuclear gene MRF1 encodes a mitochondrial peptide chain release factor and cures several mitochondrial RNA splicing defects. *Nucleic Acids Res* **20**: 6339–6346.
- Petry S, Brodersen DE, Murphy FVt, Dunham CM, Selmer M, Tarry MJ, Kelley AC, 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.
- Petry S, Weixlbaumer A, Ramakrishnan V. 2008. The termination of translation. *Curr Opin Struct Biol* **18**: 70–77.
- Poole ES, Young DJ, Askarian-Amiri ME, Scarlett DJ, Tate WP. 2007. Accommodating the bacterial decoding release factor as an alien protein among the RNAs at the active site of the ribosome. *Cell Res* **17**: 591–607.
- Rorbach J, Soleimanpour-Lichaei R, Lightowlers RN, Chrzanowska-Lightowlers ZM. 2007. How do mammalian mitochondria synthesize proteins? *Biochem Soc Trans* **35**: 1290–1291.
- Santos MA, Moura G, Massey SE, Tuite MF. 2004. Driving change: The evolution of alternative genetic codes. *Trends Genet* **20**: 95–102.
- Sarkar G, Sommer SS. 1990. The “megaprimer” method of site-directed mutagenesis. *Biotechniques* **8**: 404–407.
- Scarlett DJ, McCaughan KK, Wilson DN, Tate WP. 2003. Mapping functionally important motifs SPF and GGQ of the decoding release factor RF2 to the *Escherichia coli* ribosome by hydroxyl radical footprinting. Implications for macromolecular mimicry and structural changes in RF2. *J Biol Chem* **278**: 15095–15104.
- Scolnick E, Tompkins R, Caskey T, Nirenberg M. 1968. Release factors differing in specificity for terminator codons. *Proc Natl Acad Sci* **61**: 768–774.
- Shin DH, Brandsen J, Jancarik J, Yokota H, Kim R, Kim SH. 2004. Structural analyses of peptide release factor 1 from *Thermotoga maritima* reveal domain flexibility required for its interaction with the ribosome. *J Mol Biol* **341**: 227–239.
- Soleimanpour-Lichaei HR, Kuhl I, Gaisne M, Passos JF, Wydro M, Rorbach J, Temperley R, Bonnefoy N, Tate W, Lightowlers R, et al. 2007. mtRF1a is a human mitochondrial translation release factor decoding the major termination codons UAA and UAG. *Mol Cell* **27**: 745–757.
- Soll D, RajBhandary UL. 2006. The genetic code—thawing the ‘frozen accident.’ *J Biosci* **31**: 459–463.
- Spremlin LL, Coursey A, Navratil T, Hunter SE. 2004. Initiation and elongation factors in mammalian mitochondrial protein biosynthesis. *Prog Nucleic Acid Res Mol Biol* **77**: 211–261.
- Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol* **185**: 60–89.
- Tate WP, Caskey CT. 1990. Termination of protein synthesis. In *Ribosomes and protein synthesis: A practical approach* (ed. G Spedding), pp. 81–100. Oxford University Press, Oxford, UK.
- Temperley R, Richter R, Dennerlein S, Lightowlers RN, Czarnowska-Lightowlers ZM. 2010. Hungry codons promote frameshifting in human mitochondrial ribosomes. *Science* **327**: 301.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Tsuboi M, Morita H, Nozaki Y, Akama K, Ueda T, Ito K, Nierhaus K. 2009. EF-G2mt is an exclusive recycling factor in mammalian mitochondrial protein synthesis. *Mol Cell* **35**: 502–510.
- Uno M, Ito K, Nakamura Y. 1996. Functional specificity of amino acid at position 246 in the tRNA mimicry domain of bacterial release factor 2. *Biochimie* **78**: 935–943.
- Vestergaard B, Van LB, Andersen GR, Nyborg J, Buckingham RH, Kjeldgaard M. 2001. Bacterial polypeptide release factor RF2 is structurally distinct from eukaryotic eRF1. *Mol Cell* **8**: 1375–1382.
- Warren WC, Hillier LW, Marshall Graves JA, Birney E, Ponting CP, Grutzner F, Below K, Miller W, Clarke L, Chinwalla AT, et al. 2008. Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **453**: 175–183.
- Weixlbaumer A, Jin H, Neubauer C, Voorhees RM, Petry S, Kelley AC, Ramakrishnan V. 2008. Insights into translational

- termination from the structure of RF2 bound to the ribosome. *Science* **322**: 953–956.
- Wilson DN, Guevremont D, Tate WP. 2000. The ribosomal binding and peptidyl-tRNA hydrolysis functions of *Escherichia coli* release factor 2 are linked through residue 246. *RNA* **6**: 1704–1713.
- Youngman EM, McDonald ME, Green R. 2008. Peptide release on the ribosome: Mechanism and implications for translational control. *Annu Rev Microbiol* **62**: 353–373.
- Zhang Y, Spremulli LL. 1998. Identification and cloning of human mitochondrial translational release factor 1 and the ribosome recycling factor. *Biochim Biophys Acta* **1443**: 245–250.
- Zoldak G, Redecke L, Svergun DI, Konarev PV, Voertler CS, Dobbek H, Sedlak E, Sprinzl M. 2007. Release factors 2 from *Escherichia coli* and *Thermus thermophilus*: Structural, spectroscopic and microcalorimetric studies. *Nucleic Acids Res* **35**: 1343–1353.