Proc. Natl. Acad. Sci. USA
Vol. 92, pp. 5431–5435, June 1995
Biochemistry

Translational termination efficiency in mammals is influenced by the base following the stop codon

(release factor/5′ deiodinase/selenocysteine)

KIM K. MCCAUGHAN*†, CHRIS M. BROWN*†‡, MARK E. DALPHIN*, MARLA J. BERRY§, AND WARREN P. TATE*¶

*Department of Biochemistry and Center for Gene Research, University of Otago, Dunedin, New Zealand; and §Thyroid Division, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115

Communicated by Douglas S. Coombs, University of Otago, Dunedin, New Zealand, February 27, 1995

ABSTRACT **The base following stop codons in mammalian genes is strongly biased, suggesting that it might be important for the termination event. This proposal has been tested experimentally both *in vivo* by using the human type I iodothyronine deiodinase mRNA and the recoding event at the internal UGA codon and *in vitro* by measuring the ability of each of the 12 possible 4-base stop signals to direct the eukaryotic polypeptide release factor to release a model peptide, formylmethionine, from the ribosome. The internal UGA in the deiodinase mRNA is used as a codon for incorporation of selenocysteine into the protein. Changing the base following this UGA codon affected the ratio of termination to selenocysteine incorporation *in vivo* at this codon: 1:3 (C or U) and 3:1 (A or G). These UGAN sequences have the same order of efficiency of termination as was found with the *in vitro* termination assay (4th base: A ≈ G >> C ≈ U). The efficiency of *in vitro* termination varied in the same manner over a 70-fold range for the UAAN series and over an 8-fold range for the UGAN and UAGN series. There is a correlation between the strength of the signals and how frequently they occur at natural termination sites. Together these data suggest that the base following the stop codon influences translational termination efficiency as part of a larger termination signal in the expression of mammalian genes.**

In prokaryotes there has been a strong suspicion that the translational termination signal might involve more than just the stop codon or at least be profoundly influenced by its context (reviewed in ref. 1). The efficiency of suppression of nonsense codons varies with the nucleotides surrounding the stop codon. For example, the first nucleotide following the codon (2), the second nucleotide following (3), the third nucleotide following, and the preceding codon (1, 4) all modulate suppression. Less information is available for mammalian contexts, but a study of suppression at UAG codons in human cells demonstrated that the 3′ nucleotide influenced the efficiency of suppression (C > G > U = A) in a different order from that of *Escherichia coli* (A > G > C > U) (5). Translational readthrough at a stop codon in the gene encoding the heat shock transcription factor from *Saccharomyces cerevisiae* was shown to occur when the stop codon was followed by A or C but not when it was followed by G (6). High readthrough of stop codons occurs at a particular context where C is the 4th base in the yeast *Ste6p* gene. This may explain the observation that premature termination mutations in a conserved region of the homologous human cystic fibrosis transmembrane conductance regulator are less severe (7). Some genes have evolved termination contexts that favor significant readthrough of the signal—for example, the signal for translational readthrough of a UGA codon in Sindbis virus RNA involves predominantly a 4th-base C (8), but evidence has been obtained that a 5th-base U and 6th-base purine might also be important (9). These effects could be on aminoacyl-tRNA selection, on release factor selection at the stop codon, or on both.

Statistical analyses of mammalian gene sequences have shown that the context of initiation codons is extremely biased (10), and this context profoundly affects the initiation step (11), but codon usage biases for other sense codons are determined by the regional G+C content of the genome (12). What then is the case for termination? A small initial study (13), and more recent larger studies (14–16), of eukaryotic stop codon contexts suggested that purines were the most common 3′ nucleotide following the stop codon. The biases were not as dramatic as those for *E. coli* genes, in particular highly expressed genes, where there are strong preferences for both codon (UAA) and 3′ nucleotide (U) (17, 18). Our analysis of stop codon contexts in genes from a wide variety of eukaryotic species and gene families had indicated that certain stop codons and nucleotides immediately following the stop codons are statistically overrepresented (particularly in yeast genes, *Drosophila* genes, and highly expressed genes of all species) (14). We have proposed a model in which the single polypeptide release factor decodes the 12 possible tetranucleotide termination sequences as signals with differing efficiencies (18). Indeed, the eukaryotic *in vitro* termination assay, established by Beaudet and Caskey (19), does not function to direct the release of a model peptide with the trinucleotide stop codon alone. It requires a longer oligoribonucleotide. As concluded recently by Tuite and Stansfield (20), our understanding of translation termination in eukaryotes is still quite rudimentary.

Recently, alternative coding events such as frameshifting and selenocysteine incorporation have been discovered at some relatively rare 4-base stop signals (21). In *E. coli*, for example, UGAC occurs at the frameshifting site in release factor 2 (RF-2) mRNA (22) and also at the site for selenocysteine incorporation in formate dehydrogenase (23). During RF-2 translation, the stop signal competes relatively poorly with the alternative recoding event (18, 24). Furthermore, changing the codon and the 4th base of the stop signals, within the same context of the RF-2 frameshift site, changes the rate of factor selection by the ribosome exactly as predicted by statistical analysis (25, 26). There is a 50-fold range for the rate of selection of RF-1 or RF-2 at the 4-base signals *in vivo*. UAAU, the signal most commonly found in highly expressed genes, is most rapidly decoded as stop, while UGAC, the signal found at the recoding sites, is decoded most slowly (26).

Here we have tested the hypothesis in mammalian systems that 12 different 4-base stop signals in the same context would

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: RF, prokaryotic release factor; eRF, eukaryotic release factor; DI, deiodinase.

†K.K.M. and C.M.B. contributed equally to this paper.

‡Present address: Plant Pathology Department, 351 Bessey Hall, Iowa State University, Ames, IA 50011.

¶To whom reprint requests should be addressed.

be decoded with different efficiencies, dependent on the 4th base. We have correlated the experimental findings with the abundance of these signals in mammalian genes.

MATERIALS AND METHODS

Statistical Analysis of the Context of Mammalian Termination Codons. The sequences around the termination codons of 5208 mammalian genes (human, rat, mouse, cow, pig, and rabbit) were extracted from GenBank release 80 and aligned; the nonrandomness (χ^2 analysis) observed at each position was determined as described for other species (14). The contexts used and summary statistics are available from the authors. The information in the feature table of GenBank entries was used to extract the contexts, as described in ref. 27. Each "CDS" (coding sequence) or "mat_peptide" (mature peptide) described in the feature table was interpreted by using feature locations, qualifiers, and join specifications. Entries were rejected if (*i*) they were duplicates in the termination region; (*ii*) they had no stop codon; (*iii*) the stop codon was not preceded by a valid open reading frame (that is, the sequence is not consistent with that specified in the feature table); or (*iv*) the open reading frame was shorter than 100 bases. For valid coding regions, the stop codon, the next base, and the following 100 bases of noncoding sequence were analyzed. If the feature table described another following coding sequence, the flanking sequence was truncated to include only noncoding sequences. Then the frequency of occurrence of the 12 tetranucleotides in any "frame" of these noncoding regions was calculated.

***In Vitro* Termination Assays.** Assays were performed essentially as described (28). The factor and ribosomes were from rabbit reticulocytes. A typical reaction mixture was incubated for 30 min at 24°C and contained within 50 μl 2 pmol of f[^{3}H]Met-tRNA·AUG·ribosome substrate; 0–1 nmol of the tetranucleotides of the UAAN, UGAN, or UAGN series or the pentanucleotides of the UGAAN or UGACN series; 1–2 μl of partially purified eukaryotic release factor (eRF); and 1 mM GTP in a buffer containing 20 mM Tris·HCl (pH 7.4), 60 mM KCl, and 11 mM $MgCl_2$. To test whether the tetranucleotide preparations (Macromolecular Resources, Fort Collins, CO) were equivalent, all 12 were also tested in the prokaryotic assay *in vitro*. This *in vitro* assay is not sensitive to the identity of the 4th base (ref. 29; K.K.M. and W.P.T., unpublished data). They all supported termination with equal specific activities in this system. The mammalian experiments were repeated both with the same and with a second preparation of the tetranucleotides with completely consistent results. Backgrounds of nonspecific f[^{3}H]Met release (0.1–0.2 pmol) were subtracted in each case, and generally the ratio specific release of f[^{3}H]Met by eRF/nonspecific release was ≥3:1.

The formation of eRF-ribosomal complexes with the ^{32}P-labeled pUAAN and ^{32}P-labeled pUAGN series of tetramers was essentially as described (23) except that the tetramers were ^{32}P-end-labeled using T4 polynucleotide kinase and [γ-^{32}P]ATP. A typical reaction mixture was incubated for 20 min at 4°C and contained within 50 μl 50 pmol of ribosomes, 5–100 pmol of the labeled oligonucleotides (10,000 cpm/pmol), eRF (5–10 μl), and 0.1 mM guanosine 5′-[β,γ-methylene]triphosphate in a buffer containing 50 mM Tris·HCl (pH 7.2), 100 mM NH_4Cl, 20 mM $Mg(OAc)_2$, and 20% (vol/vol) ethanol. Backgrounds were generally ≈0.15 pmol for 5 pmol of added labeled oligonucleotide whether eRF or the ribosome was omitted. They were similar for most radiolabeled preparations of oligonucleotide and, depending on the specific oligonucleotide, the amount of complex formed with eRF was at least 2-fold over background. It should be noted that pure tryptophanyl-tRNA synthetase, a contaminant in purified eRF preparations, has no activity in these assays.

The Synthesis of 5′ Deiodinase (DI) *in Vivo*. Constructs containing the rat 5′ DI cDNA with UAAC or UGAN at codon position 126 in the CDM-8 vector (GenBank accession no. X57999) were transfected into human embryonic kidney 293 cells (10 μg per 60-mm dish) by $CaPO_4$ coprecipitation as described (30). Cells were harvested 2 days after transfection and sonicated, and the proteins in the extract were analyzed. Competition between the eRF and the selenocysteyl-tRNA for decoding of the internal UGA of the mRNA expressed from the transfected 5′ DI gene gives either a termination product (14 kDa) or a full-length readthrough product (28 kDa) (31). The readthrough is enhanced by a selenocysteine insertion element (32). The products were detected immunologically by using an antibody to the N terminus of the DI protein after separation of the cell extracts on a 12.5% denaturing polyacrylamide gel and blotting to an Immobilon membrane (33).

RESULTS

The Relative Frequency of Occurrence of 4-Base Stop Signals in Mammalian Genes. A data base was compiled of the sequences around the stop codons of 5208 genes from six mammalian species (human, mouse, rat, cow, pig, and rabbit), which had >100 sequences available. A frequency table detailing the incidence of each of the four nucleotides in each position around the stop codon was constructed. We then determined the bias at each position. This analysis, shown in Fig. 1, demonstrated a highly significant bias in the position immediately following the stop codon as had previously been seen and reported for the prokaryote *E. coli* and the eukaryote *S. cerevisiae* (14, 17). There is also nonrandomness in the 5th base and, indeed, in the first 8 positions following the stop codon and 3 nucleotides preceding the codon ($P < 0.001$), but the most striking bias is in the base immediately following the codon (the 4th base). The nonrandomness in the 5th base and beyond was not seen in the analysis of the *E. coli* genes where, of the bases following the stop codon, the 4th base alone showed the striking bias (17).

We have previously proposed a model in which a stop signal comprised 4 bases (the termination codon plus the following base) as a result of these earlier theoretical analyses (14, 17). Based on such a model, the frequency of occurrence in mammalian genes of all 12 possible signals was determined. A nonspecific bias due to mono-, di-, tri-, or tetranucleotide nonrandomness was controlled for by analysis of frequencies in the adjacent regions of the genome.

The occurrence of the 4-base signals in the total mammalian gene group generally reflected the frequency of these sequences in the noncoding regions. However, the three signals

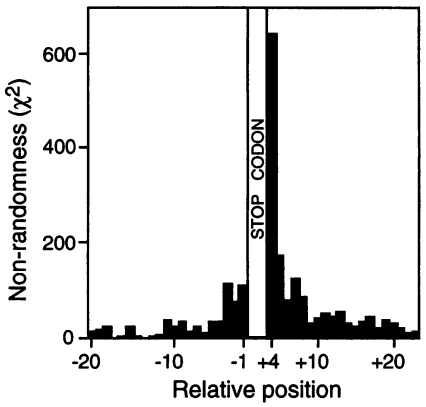

FIG. 1. Statistical analysis of the context of mammalian termination codons. Sequences around the termination codons of 5208 mammalian genes (human, rat, mouse, cow, pig, and rabbit) were analyzed as described.

with G as the 4th base were statistically more abundant than expected from the noncoding data, whereas those with U are less abundant than expected ($P < 0.001$). The A and C series are as expected, except for UAAA, which is underrepresented ($P < 0.001$). In *E. coli*, highly expressed genes have been defined from their codon adaptation index (34), and in these genes two stop signals (on the 4-base model), UAAU and UAAG, are predominantly used (26). While definition of a set of highly expressed genes in mammals is somewhat more arbitrary, the frequency with which the particular 4-base signals occurred in such highly expressed genes as globins, histones, actins, immunoglobulins, and albumins was also quite different from that found in noncoding regions of this class of genes or from that found in all mammalian genes. For example, certain signals (UAAG, UAGG) are overrepresented ($P < 0.01$), and certain signals are not used at all (UGAC, UGAU, UAGC, UAGU) (data not shown). Individual analyses on the subsets of the human and rabbit genes (the species used for the *in vivo* and *in vitro* experimental studies described below) gave similar results.

Collectively, these data might reflect a particular usage of a subset of signals in the special circumstance of high expression, when a selective pressure may be strong enough to overcome mutational biases. In the case of prokaryotes, termination optimization may be a prime factor in the selection; however, it is more likely that the selective pressure on the sequences at the end of the mammalian genes is multifactorial and reflects pressure on more than just the 4th base.

The Base Following the Termination Codon Influences Decoding Efficiency *in Vitro*. The concept that the termination signal might consist of 4 bases, as suggested by the statistical analyses (Fig. 1), was tested experimentally. All of the 12 possible 4-base stop signals were synthesized and used in termination assays to direct the release of a tRNA-linked model peptide (fMet) on a ribosome complex in the presence of eRF (Fig. 2*A*; ref. 28). This assay has been used for the past 25 years to determine much of what we know today about mammalian termination.

There was significant variation in the ability of the tetranucleotides to promote release of a product encompassing a 70-fold variation with the UAAN series and ≈8-fold with the UGAN and the UAGN series. This is illustrated with UGAN in Fig. 2*B*. UGAA (squares) was slightly more effective than UGAG (diamonds), and these two were much more effective than UGAC (circles) and UGAU (triangles). Although there were minor variations in this pattern within the UAAN and UAGN series, they mirrored these results in general terms with the influence of the 4th base A ≈ G >> C ≈ U (Fig. 2*C Left*). Some stop codon dependence can also be seen in these assays—for example, UGAA ≈ UAAA > UAGA, as observed with this series in an earlier study (19). However, this dependence was much less pronounced than the 4th-base effect.

The termination activities of the 12 signals are compared with their usage in highly expressed mammalian genes in Fig. 2*C*, ranked in the order of their functional effectiveness. In broad terms, the signals with a purine in the 4th position are more common and more active as a termination signal than those with a pyrimidine. Within each of these two groups there is some variation between functional activity and the frequency of occurrence. Evolutionary factors other than selection for efficiency in termination—for example, suppression by a particular codon-specific tRNA—are presumably responsible for this variation within each class.

A stoichiometric assay can measure the apparent codon recognition step at equilibrium—namely, the formation of a ribosome·stop signal·eRF complex—in the absence of a tRNA-linked model peptide and without involvement of the peptidyl transferase center (see Fig. 2*A*). It was of interest whether the 4th-base influence could be observed under these equilibrium conditions. Radiolabeled UAAN tetranucleotides were incorporated into the termination complex with varying efficiencies, with the 4th-base effect being in the order A > G > C > U, consistent with the data for the catalytic assay (Fig. 2*B*). In the UAGN series, the order was C ≈ A ≈ G > U. Another step of termination or a kinetic effect must relegate the UAGC to a relatively poor signal in the overall termination assay. The third series, UGAN, could not be assessed because the eRF alone binds these tetranucleotides in the absence of ribosomes but without influence of the 4th base (K.K.M. and W.P.T., unpublished data).

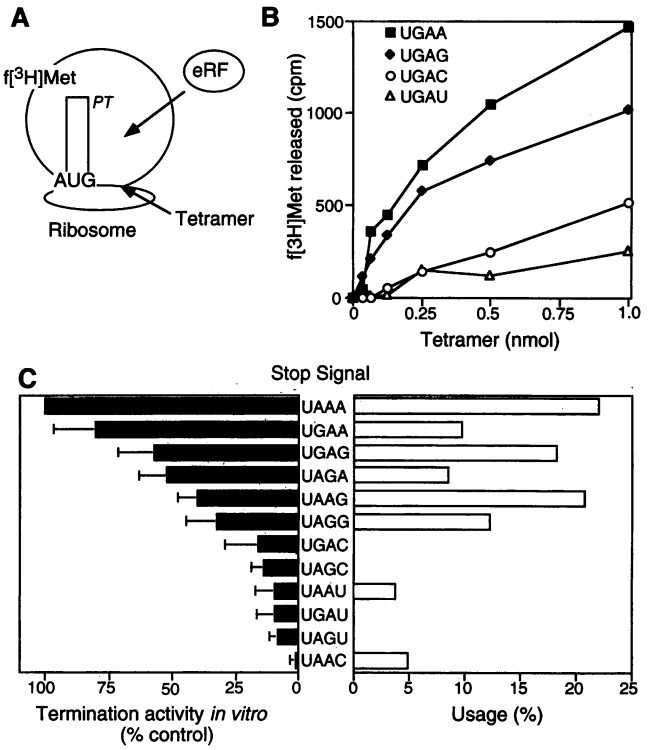

FIG. 2. (*A*) Schematic of *in vitro* termination assay (19, 28). *PT*, peptidyl transferase. The model peptide is provided by the f[^{3}H]Met on its tRNA bound at the ribosomal P site with cognate codon. The tetramer containing the stop codon, in a complex with the eRF at the A site, directs the release. (*B*) Release of the model peptide f[^{3}H]Met by the eRF with the UGAN series of signals. (*C*) Comparison of termination activity *in vitro* (*Left*) with the usage of these stop signals for termination in highly expressed mammalian genes (*Right*). The 12 stop signals (center) are ranked in order of termination activity *in vitro*. Termination assays were performed in duplicate with a range of concentrations of tetramer (as in *B*). Values for each concentration were divided by the highest values (UAAA) and expressed as percentage of the UAAA activity and then averaged for each series. Error bars represent SD.

Since the statistical analysis (Fig. 1) had indicated nonrandom occurrence of nucleotides in the 5th base, we have begun a comprehensive study on the influence of this base and beyond as to the strength of the termination signal both *in vitro* (as in Fig. 2) and *in vivo* (as in Fig. 3). For example, the strongest and weakest tetranucleotides of the UGAN series (UGAA and UGAC) were used to create two series, UGAAN and UGACN. Preliminary data indicate the specificity of the 5th base has little effect on the strong UGAA signal but that there is a pyrimidine/purine split in how the 5th base can affect the weak UGAC signal, with the 5th-base A/G increasing the signal's effectiveness much more than U/C. In contrast to what we have found to date in *E. coli*, these preliminary data already suggest the influence of the nucleotides 3′ to the stop codon extends beyond the 4th base in mammalian termination.

The 4th Base Influences Termination Signal Decoding *in Vivo*. Are the effects observed with these *in vitro* assays

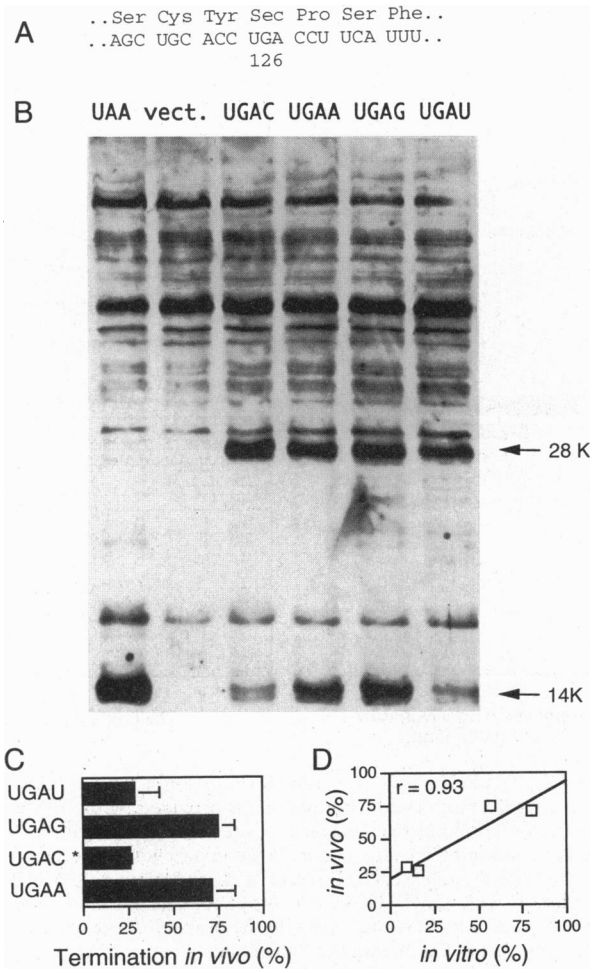

FIG. 3. Termination *in vivo* at the internal UGA sense codon of the 5′ DI mRNA. (*A*) Natural context of the UGA sense codon. (*B*) *In vivo* expression of the DI gene. Constructs containing the rat 5′ DI were transfected in human kidney cells, and after 2 days culture extracts were prepared and separated by SDS/PAGE. Proteins were blotted onto a membrane for detection of the specific 5′ DI termination and complete products (arrows) using antibodies against the N-terminal region. Sequence at codon position 126 is shown above, with lane 2 being the proteins from transfection by vector alone. (*C*) Amount of termination product (14 kDa) as a percentage of the total of the two products (28 + 14 kDa). Natural sequence (UGAC) is indicated by asterisk. Values are averages from three separate experiments; error bars represent SD. (*D*) Correlation between *in vitro* data (from Fig. 2*C*) and *in vivo* data. r, Linear correlation coefficient.

mimicked *in vivo*? An ideal system to investigate this is the incorporation of selenocysteine into type I iodothyronine deiodinase, 5′ DI. Berry *et al.* (31) have shown that the mRNA for this protein contains an internal UGA sense codon, at which selenocysteine is incorporated (32). The incorporation of selenocysteine by a special tRNA is promoted by a selenocysteine insertion element in the 3′ untranslated region (32). The context of the mRNA is shown in Fig. 3*A*, where the UGA at codon position 126 is followed by the base cytosine (UGAC was demonstrated in the *in vitro* experiments above to be relatively poor at directing the release of model peptide having <20% of the activity of the best tetranucleotide UAAA). Potentially, there is competition between termination and selenocysteine incorporation mechanisms at this internal UGA.

Constructs containing the four possible bases in the position following codon 126 were produced by site-directed mutagenesis, and the stop codon at 126 was also changed from UGA to UAA to preclude selenocysteine incorporation. Two products may be synthesized, a 14-kDa termination product or the 28-kDa mature protein if selenocysteine is incorporated (Fig. 3*B*). These products were detected by an antibody to the N terminus of the DI protein after expression in human cells in culture (33). As shown in Fig. 3*B*, only the termination product is detected as expected if UAA is the stop codon at the site, since there can be no selenocysteine incorporation, and neither product is detected from transfections of vector alone. When the base following the internal UGA was changed from a C to one of the other three bases so that the competitiveness of the two events as a function of this 4th base could be assessed, UGAU, like the naturally occurring UGAC, allowed termination in ≈1 in 4 ribosomal passages (26%), thus allowing selenocysteine incorporation at most passes. However, when this base was changed to a purine, the competing mechanism, termination, occurred in ≈3 of 4 ribosomal passages (Fig. 3*C*). An excellent correlation ($r = 0.93$) between the *in vitro* results with this UGAN series and the *in vivo* study is shown in Fig. 3*D*. Both studies showed that a purine in the 4th base of the stop signal was significantly more effective for termination than a pyrimidine.

DISCUSSION

The implication of these studies is that the base following the stop codon can have considerable influence on the mammalian termination signal.

The *in vivo* study with the 5′ DI gene reflects a competition between two events at the UGA codon—selenocysteyl-tRNA selection and eRF selection. No information is available on whether the bases adjacent to the UGA codon affect selenocysteine-tRNA selection specifically. We must assume that here we are measuring a net effect of the 4th base on the two processes. A 4th-base purine/pyrimidine split is observed, with termination favored when the 4th base is a purine and selenocysteyl-tRNA selection favored when the 4th base is a pyrimidine.

The *in vitro* studies utilize a model system where termination can be studied in isolation from suppression and where the codon plus the 4th base can be studied in isolation from the surrounding context influences. The same 4th-base pyrimidine/purine split observed in translation of the 5′ DI mRNA *in vivo* was found in these experiments. Moreover, the oligonucleotides UAAN and UAGN containing the other two stop codons also supported termination with varying efficiencies with the 4th-base purines contributing to a strong signal and the 4th-base pyrimidines contributing to relatively weak signals. These data suggest that where the surrounding context is constant (as is the case of the 5′ DI genes when only the 4th base is changed) the 4th base can influence the strength of the termination signal through its ability to be decoded by the eRF.

The results imply that mammalian genes contain a set of stop signals that are decoded with varying efficiencies. Highly expressed genes presumably require a signal that will be decoded rapidly. It is perhaps not surprising then that such genes tend to have the most efficient signals: UAAR, UAGR, UGAR (R = purine). The majority of genes, which have relatively low expression, could accommodate a diversity of stop signals without compromising translational efficiency, and the nature of such signals will have been largely determined by genetic drift. Normally, only when relatively poor signals are challenged by the presence of a strong recoding signal (for example, a selenocysteine element) or by a mutant suppressor tRNA do their inherent weaknesses become apparent. However, in a reticulocyte lysate UGAC, a poor stop signal by our analysis, is apparently a sufficient context to permit up to 10% readthrough by natural tRNAs (8, 35).

In the highly expressed genes of many species, but not mammals, it appears that a strong bias in sense codons reflects an evolutionary selection pressure for optimal codons. These

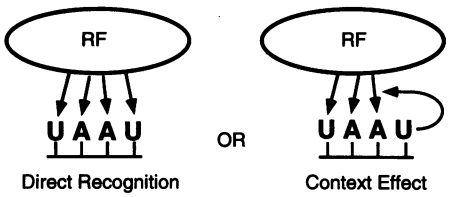

FIG. 4. Possible models for influence of the 4th base on termination.

codons are best recognized by the most abundant tRNAs and are translated faster than others. However, in lower expression genes the pattern of codon usage will be determined by mutational bias (12). In this case, the codon bias will reflect the A+T or G+C richness of the genome or subsections of it. For stop signals and their contexts, the bias could provide optimal termination to avoid competing mechanisms, such as suppression or frameshifting. Recent studies of UAGN suppression by a mutated tRNA in human cells showed that the 4th-base effect on suppression was A ≈ U < G < C (5). UAGA is suppressed poorly, whereas UAGC is well suppressed, consistent with our analysis; however, the order of UAGU and UAGG was reversed (predict A ≈ G < U ≈ C), perhaps indicating a tRNA-dependent preference for G. In *E. coli*, stop codons followed by G and A are the most efficiently suppressed (36).

This study provides direct experimental evidence that mammalian termination may be modulated by nucleotides additional to those of the stop codon. The mechanism as to how the 4th base influences the decoding of the stop signal is yet to be determined. Two models are possible; we favor a mechanism involving direct recognition of this base by the eRF, but it may be indirect, as a result of a context effect on the conformation of the codon and thereby its interaction with the eRF (Fig. 4). Recognition of the signal by eRF need not be restricted to 3 (or 4) bases and the factor may make contact with a wider context. Indeed, the fact that a 5th-base purine can restore the specific activity of a weak 4-base signal for an eRF-mediated event, whereas it has little influence on the strong 4-base signal, would be consistent with this idea. Alternatively, the 4th base may interact with the small subunit rRNA and provide a substrate of varying affinity for the eRF. We believe the 4-base concept may be applicable to most organisms. Using an *in vivo* competition assay in *E. coli* between frameshifting and termination at the frameshift site of the RF-2 gene, we have found that the 4th base influences the RF selection rate over a wide range (26). While the signal strength hierarchy is different from that reported here for the mammalian genes, it is consistent with that predicted from statistical analysis of the prokaryotic genes (26). The generality of the concept of a 4-base stop signal remains to be determined.

We thank S. Mannering for help in manuscript preparation and P. Stockwell for help with the computers used. M.J.B. is supported by National Institutes of Health Grant DK 47320; W.P.T. is an International Scholar of the Howard Hughes Medical Institute and is supported by grants from the Human Frontier Science Program, the New Zealand Health Research Council, and the New Zealand Lotteries Board.

1. Buckingham, R. H. (1990) *Experientia* **46,** 1126–1133.
2. Bossi, L. (1983) *J. Mol. Biol.* **164,** 73–87.
3. Stormo, G. D., Schneider, T. D. & Gold, L. (1986) *Nucleic Acids Res.* **14,** 6661–6679.
4. Buckingham, R. H., Sörensen, P., Pagel, F. T., Hijazi, K. A., Mims, B. H., Brechemier-Baey, D. & Murgola, E. J. (1990) *Biochim. Biophys. Acta* **1050,** 259–262.
5. Martin, R., Phillips-Jones, M. K., Watson, F. J. & Hill, L. S. (1993) *Biochem. Soc. Trans.* **21,** 846–851.
6. Kopczynski, J. B., Raff, A. C. & Bonner, J. J. (1992) *Mol. Gen. Genet.* **234,** 369–378.
7. Fearon, K., McClendon, V., Bonetti, B. & Bedwell, D. M. (1994) *J. Biol. Chem.* **269,** 17802–17808.
8. Li, G. & Rice, C. M. (1993) *J. Virol.* **67,** 5062–5067.
9. Farabaugh, P. J. (1993) *Cell* **74,** 591–596.
10. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8148.
11. Kozak, M. (1992) *J. Cell Biol.* **115,** 887–903.
12. Sharp, P. M., Stenico, M., Peden, J. F. & Lloyd, A. T. (1993) *Biochem. Soc. Trans.* **21,** 835–841.
13. Kohli, J. & Grosjean, H. (1981) *Mol. Gen. Genet.* **182,** 430–439.
14. Brown, C. M., Stockwell, P. A., Trotman, C. N. A. & Tate, W. P. (1990) *Nucleic Acids Res.* **18,** 6339–6345.
15. Cavener, D. R. & Ray, S. C. (1991) *Nucleic Acids Res.* **19,** 3185–3192.
16. Sharp, P. M., Burgess, C. J., Cowe, E., Lloyd, A. T. & Mitchell, K. J. (1992) in *Transfer RNA in Protein Synthesis*, eds. Hatfield, D. L., Lee, B. J. & Pirtle, R. M. (CRC, Boca Raton, FL), pp. 397–425.
17. Brown, C. M., Stockwell, P. A., Trotman, C. N. A. & Tate, W. P. (1990) *Nucleic Acids Res.* **18,** 2079–2086.
18. Tate, W. P. & Brown, C. M. (1992) *Biochemistry* **31,** 2443–2450.
19. Beaudet, A. L. & Caskey, C. T. (1971) *Proc. Natl. Acad. Sci. USA* **68,** 619–624.
20. Tuite, M. F. & Stansfield, I. (1994) *Mol. Biol. Rep.* **19,** 171–181.
21. Gesteland, R. F., Weiss, R. B. & Atkins, J. F. (1992) *Science* **257,** 1640–1641.
22. Craigen, W. J., Cook, R. G., Tate, W. P. & Caskey, C. T. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 3616–3620.
23. Zinoni, F., Birkmann, A., Stadtman, T. C. & Bock, A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 4650–4654.
24. Donly, B. C., Edgar, C. D., Adamski, F. M. & Tate, W. P. (1990) *Nucleic Acids Res.* **18,** 6517–6522.
25. Pedersen, W. T. & Curran, J. F. (1991) *J. Mol. Biol.* **219,** 231–241.
26. Poole, E. S., Brown, C. M. & Tate, W. P. (1995) *EMBO J.* **14,** 151–158.
27. Brown, C. M., Dalphin, M. E., Stockwell, P. A. & Tate, W. P. (1993) *Nucleic Acids Res.* **21,** 3119–3123.
28. Tate, W. P. & Caskey, C. T. (1990) in *Ribosomes and Protein Synthesis: A Practical Approach*, ed. Spedding, G. (Oxford Univ. Press, Oxford), pp. 81–100.
29. Ganoza, M. C., Buckingham, K., Hader, P. & Neilson, T. (1984) *J. Biol. Chem.* **259,** 14101–14104.
30. Brent, G. A., Harney, J. W., Chen, Y., Warne, R. L., Moore, D. D. & Larsen, P. R. (1989) *Mol. Endocrinol.* **3,** 1996–2004.
31. Berry, M. J., Banu, L., Chen, Y., Mandel, S. J., Kieffer, J. D., Harney, J. W. & Larsen, P. R. (1991) *Nature (London)* **353,** 273–276.
32. Berry, M. J., Banu, L., Harney, J. W. & Larsen, P. R. (1993) *EMBO J.* **12,** 3315–3322.
33. Berry, M. J., Harney, J. W., Ohama, T. & Hatfield, D. L. (1994) *Nucleic Acids Res.* **22,** 3753–3759.
34. Sharp, P. M. & Li, W. (1987) *Nucleic Acids Res.* **15,** 1281–1295.
35. Martin, R. (1994) *Nucleic Acids Res.* **22,** 15–19.
36. Yarus, M. & Curran, J. (1992) in *Transfer RNA in Protein Synthesis*, eds. Hatfield, D. L., Lee, B. J. & Pirtle, R. M. (CRC, Boca Raton, FL), pp. 319–365.