Biochemistry

 Volume 31, Number 9 March 10, 1992

Perspectives in Biochemistry

Translational Termination: "Stop" for Protein Synthesis or "Pause" for Regulation of Gene Expression†

Warren P. Tate* and Chris M. Brown

Department of Biochemistry and Centre for Gene Research, University of Otago, Dunedin, New Zealand

Received December 18, 1991

The termination of protein synthesis has traditionally referred to the polypeptide release factor mediated release of the completed polypeptide from the ribosome. This occurs in response to the appearance of a stop codon in the decoding site. Understanding this last phase of protein synthesis has been slower than that of the other phases, partly because it has been difficult to establish appropriate in vitro systems to study the event. There had also been a perception that termination was of less profound significance to the cell. It is now clear that the efficiency of reading particular stop signals can be very important in the regulation of certain cellular events. Translational termination can now be thought of in two ways: either as a complete stop in protein synthesis mediated by the polypeptide chain release factor (RF), the general and fundamental event resulting in release of the protein product, or as a pause at these stop signals, the signals being used as a "give way" (or "yield") for more specialized purposes. Translational stopping involves an intimate relationship between the ribosome, the mRNA, and the polypeptide release factors (Figure 1; Tate et al., 1990; Craigen et al., 1990) but the pause may allow a diverse array of possible competing events at the signal, such as specific amino acid incorporation (Stadtman 1991; Bock et al., 1991), read through [as reviewed in Valle and Morch (1988), Hatfield and Oroszlan (1990), and Eggertsson and Soll (1988)], or frame shifting (Atkins et al., 1990; Jacks, 1990).

What Is a Stop Signal? Although the stop signal in the mRNA is an important component of the termination event, identication of three codons specifying stop, UAA, UAG, and UGA, seemed to resolve the mRNA's contribution and resulted in little attention being given to this aspect for some time. Evidence accumulated, however, that these codons do not always signal stop, and researchers discovered a class of tRNAs, the suppressor tRNAs, that in many organisms recognize the codons as sense and compete with the normal termination mechanism [reviewed in Eggertsson and Soll (1988) and Hatfield et al. (1990)]. This suppression occurs with different efficiency depending on the context, leading to early suggestions that the stop signal was longer than the triplet codon (Salser, 1969), but only recently has this idea been reexamined. Similarly, with the discovery of alternative genetic codes involving the code words for stop, there has been renewed interest in how these stop signals have evolved.

(*i*) *Evolution of the Stop Signal.* What was the early stop signal, how did it evolve into the stop signal used today, and what is the nature of these efficient stop signals? An interesting theory of "stop codon takeover" has been developed by Lehman and Jukes (1988), who suggest that all codons were originally nonsense until tRNAs evolved to recognize them. The protein release factors would have evolved to recognize those nonsense codons remaining unassigned to amino acids. The alternative genetic codes provide compelling evidence for this idea, for example, the UGA stop codon being taken over by tryptophan, cysteine, or selenocysteine and the UAA and UAG stop codons by glutamine in some alternative genetic systems. The assignment of a universal stop codon to an amino acid code word or stop in a specific organism or cell organelle might have been influenced by genomic AT or GC pressure, as seen in the replacement of UGG tryptophan codons with UGA (Osawa et al., 1990). This would have to be accompanied by the apppearance of a tRNA with the appropriate anticodon and subsequently the loss of the factor recognizing the UGA as stop. Indeed, rat mitochondria, which use UGA for tryptophan, have lost a polypeptide release factor that can recognize UGA as stop but retained one which recognizes UAA and UAG (Lee et al., 1987). Within genetic systems using the universal code the use of UGA in some genes as a code word for selenocysteine may be an example of "stop codon takeover in progress" (Bock et al., 1991; Stadtman, 1991). In the most extreme case it occurs in 10 places within a single

† The work of the authors described in this review has been supported by the Health Research Council of New Zealand and the Alexander von Humboldt Foundation, Germany.

0006-2960/92/0431-2443$03.00/0

Table I: Relative Occurrence of Stop Signals in Genes of Several Organisms

S. cerevisiae			human			*E. coli*		
	occurrence (%)			occurrence (%)			occurrence (%)	
signal	total	high expression	signal	total	high expression	signal	total	high expression
(i) preferred[a]								
UAAG	18	50	UGAG[b]	20	24	UAAU	28	56
UAAA	20	31	UAAG[b]	8	16	UAAG	15	32
			UAAA	10	15			
			UAGG[b]	7	13			
(ii) intermediate								
UAAU	15	9	UGAA	14	7	UGAU	13	7
UAGA	13	3	UGAC	8	5	UAAA	13	1
UGAA	6	6	UAAU[b]	7	3	UAAC	10	3
UAGU	8		UAAC	6	5	UGAA	7	1
UGAU	6		UGAU	6	3			
			UAGA	6	9			
			UAGC[c]	6	1			
(iii) rare								
UAAC	2		UAGU[c]	3	1	UGAG	4	
UGAC	2					UGAC	4	
UAGC	2					UAGU	3	
UGAG	2					UAGG	2	
UAGG	2					UAGC	2	
						UAGA	1	

[a] Used in >10% of highly expressed genes. [b] Most suppressible by aminoglycosides. [c] Least suppressible by aminoglycosides.

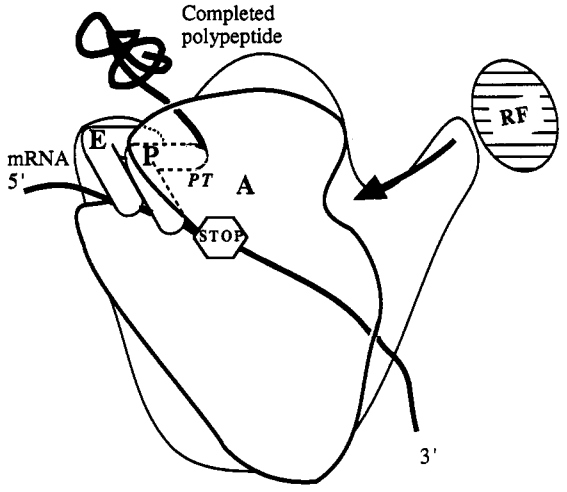

FIGURE 1: Translational stopping. The release factor (RF) binds the ribosome in response to the appearance of a stop signal in the mRNA at the decoding site. The completed polypeptide is cleaved from the peptidyl site (P site) bound tRNA by the peptidyl transferase (PT). Abbreviations: E, exit site bound tRNA; P, peptidyl site bound tRNA; A, aminoacyl site.

gene (Hill et al., 1991). The terms "stop codon" and "nonsense codon" were once considered synonymous, but it is now important to distinguish the use of these terms. Stop codons are actively recognized by a release factor whereas a nonsense codon has no active recognition process. An example of a true nonsense codon is CGG in AT-rich mycoplasma. Although no tRNA exists to decode this codon, it signals stop poorly and causes a stop in translation without termination, presumably because no factor has evolved to recognize it (Oba et al., 1991). In GC-rich *Micrococcus luteus* codons ending in A may also be true nonsense codons (Kano et al., 1991).

(*ii*) *Efficient Stop Signals.* The active process to terminate protein synthesis that is found in organisms today involves an interaction between the mRNA code word and a specific factor on the ribosome. In the evolution of such a protein-mediated response to a stop codon, there was no necessity to restrict the signal to the nonsense triplet. Indeed, there is a rapidly increasing array of exceptions to the rule that a triplet stop codon signals stop with very high efficiency, renewing ideas that an efficient stop signal might be more than just the codon.

(*iii*) *The Signal in Prokaryotes.* Analysis of the region around the stop codons of many genes from prokaryotes revealed local nonrandomness (see Figure 2), particularly within and in the nucleotide following the triplet (Brown et al., 1990a). In *Escherichia coli* there is a preference for the UAA stop codon (Sharp & Bulmer, 1988) and a strong preference for U or G and against A or C immediately following. This is accentuated in highly expressed genes which are terminated mainly by UAAU (56%) of UAAG (32%) (Table I). The bias suggests that, for the efficient termination of protein synthesis in *E. coli*, the stop signal is a tetranucleotide and that the polypeptide chain release factors recognize this extended signal. Indeed, there is a heirarchy of termination signals whose efficiency is influenced by the fourth base position. As predicted by this hypothesis, it has been shown that the rate of release factor-1 (RF-1) selection at UAGN termination sites varies 3-fold, with UAGU being the best and UAGA the poorest (Pedersen & Curran, 1991). The translational pause would be shortest at those signals which select the RF more quickly and longest at poor stop signals.

(*iv*) *The Signal in Eukaryotes.* It had been known for some time that a minimum of a tetranucleotide was required for factor-mediated termination to be measured in an in vitro assay, but this was thought to be a requirement of the ribosome to bind the codon rather than to relate to its recognition (Caskey, 1980). Analysis of the region around the stop signal in genes from a wide variety of eukaryotes gave strong suggestive evidence that the termination signal is also a tetranucleotide (Brown et al., 1990b). For example, the preferred signal for highly expressed genes in *Saccharamyces cerevisiae* (Table I) and *Drosophila melanogaster* is UAAG (Brown et al., 1990b). These preferences are most evident in organisms in which rapid protein synthesis is an advantage but are not as obvious in mammals or plants (for example, in the human genes as shown in Table I). It has recently been shown that termination is a relatively slow step involving a translational pause in rabbit reticulocyte lysates (Wolin & Walter, 1988). Comparative studies of the effects of aminoglycoside suppression in human cultured cells and *E. coli* have shown that failure to terminate is context dependent in human cells and that these context effects differ between *E. coli* and human

cells (Martin et al., 1989), consistent with our analysis (as indicated in Table I). Generally, the signals UAA(A/G) and UGAG are preferred in eukaryotes where the fourth nucleotide is affected by the GC % of the organism or DNA region (Angenon et al., 1990; Brown et al., 1990b), with a strong bias against the CG dinucleotide immediately following the stop codon (Cavener & Ray, 1991).

Stop versus Pause. Poor stop signals may be advantageous in some circumstances. The importance of having a heirarchy of efficiencies for stop signals is that the cell may be able to use these when alternative events to stopping protein synthesis are important. In most of the cases these competing events are enhanced by other factors such as the secondary or tertiary structure of the mRNA. Presumably there is a complex kinetic balance between termination and suppression or frame shifting at these sites [reviewed in Parker (1989)].

(*i*) *Alternative Events at Pause Sites.* It is of interest that UGAC is rarely used at a natural termination site in *E. coli*, and this sequence is found in two well-characterized failures to terminate in this organism. In the formate dehydrogenase gene UGA(C) is recognized by a specific tRNA to incorporate selenocysteine and stop is excluded (Bock et al., 1991). Similarly, an in-phase stop codon [UGA(C)] signals stop in only 50–70% of ribosomal passes at a site within the coding region of the release factor-2 (RF-2) gene (Craigen & Caskey, 1986; Donly et al., 1990). On the remainder of passes it is circumvented by a high-frequency frame-shifting event. The stop signal can be made more efficient to give almost 100% termination by increasing the level of the factor that recognizes it (Donly et al., 1990), despite there being other context elements surrounding the site which favor frame shifting. The significance of this to the cell seems to be in autoregulation of the RF-2 concentration, as high levels would act to reduce RF-2 synthesis whereas low levels would permit it. Perhaps the most bizarre alternative ribosomal event involving a stop codon yet reported is the bypassing of 50 nucleotides by an elongating ribosome in phage T4 DNA topoisomerase mRNA. Here the in-phase UAGC, also very rare at natural termination sites, is a part of an elaborate mechanism to achieve the bypass (Weiss et al., 1990b).

In these examples, there are other "stimulators" promoting the event competing with the "stop mechanism" at the site. It is not clear whether a poor stop codon has evolved to increase the efficiency of the alternative event or whether such events have evolved around sites of poor stop codons. Perhaps then subtle differences in stop signal strength would become unimportant over time, while the need for the stop codon itself remained significant. An analysis of RF-2 genes across a wide range of species would test whether the internal stop codon is always accompanied by upstream elements that enhance frame shifting or whether one or other of the elements can exist within the gene on its own.

Other examples where the stop codon contributes to the balanced production of important protein products can also be found. Not surprising, many of the examples are found in viruses where such mechanisms allow high economy in the use of nucleic acid sequence to code for proteins. Stop codons occur in frame immediately after the shifty sequences between the *gag* and *pol* or *pro* and *pol* genes in several retroviruses (class II, IV), whereas in the case of another class of retroviruses (class I) it is a read-through event of a leaky UAG codon at the end of the *gag* gene that allows balanced production of gag and the gag–pol fusion protein (Jacks, 1990; Hatfield & Oroszlan, 1990). In the frame-shifting retroviruses stem loops or pseudoknots enhance frame shifting, and recently

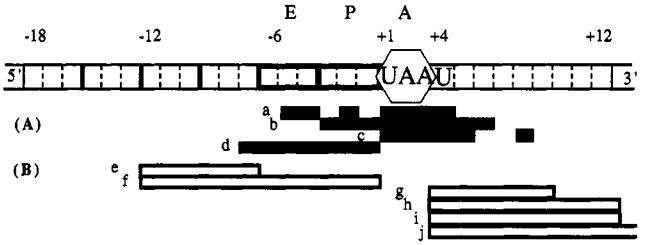

FIGURE 2: Regions around the stop codon that may affect its function. The mRNA is positioned with the stop codon in the A site. The dotted lines separate nucleotides and the solid lines codons. (A) General features (filled bars): sequences (a) around natural stop codons (*E. coli*), (b) affecting read through (*E. coli*), (c) around natural stop codons (*S. cerevisiae*), and (d) promoting frame shifting (many systems). (B) Features found in specific genes (open bars): sequences (e) involved in a Shine–Dalgarno-like interaction (*E. coli*), (f) essential for heat shock response (reticulocyte lysate), (g) affecting reinitiation (*S. cerevisiae*), (h) purine-rich affecting read through during retrovirus gene expression (reticulocyte lysate), (i) affecting read through in tobacco mosaic virus infected plants, and (j) forming stem loops and pseudoknots (often very distant from the stop signal) affecting frame shifting and read through during retroviral gene expression (mammals) or decoding of UGA as selenocysteine (*E. coli*, human). See text for references.

Atkins has obtained evidence for a pseudoknot promoting read through of the Moloney murine leukemia virus gag stop codon (Wills et al., 1991). It may be that in this case a purine-rich sequence following the stop codon is more important than the nature of the stop codon itself, since in this example the UAG can be replaced by either of the other two stop codons UAA or UGA (but not by CAG) to give productive viral infection (Honigman et al., 1991; Jacks, 1990).

Commonly the expression of open reading frames in plant RNA viruses occurs through read through of the leaky UAG stop codons which separate two protein coding regions. For example, in tobacco mosaic virus, a UAGC (a rare tetranucleotide in our analysis) allows 2–5% read through, whereas G, A, or U in the fourth position lowers this 20–50-fold (Skuzeski et al., 1991). Skuzeski et al. have also shown that the presence of this C and the following five nucleotides influences the efficiency of leakiness in vitro. It might be expected that these poor signals would be decoded relatively slowly, enhancing the possibility of read-through expression, but it must be noted again that the stop signal appears to be only one of several elements of mRNA structure influencing the event.

(*ii*) *Other Elements of Stop Signal Context.* What other context effects in the mRNA can influence how a stop codon functions and influence the termination of protein synthesis? Upstream elements have also been shown to be important in several cases. For example, under heat shock conditions the last four codons (encoding Glu, Glu, Val, Asp) in the heat shock protein family (hsp70) are critical for translation in reticulocyte lysates and may affect termination (Denisenko & Yarchuk, 1989). Other examples in which frame shifting is enhanced include the Shine–Dalgarno sequence upstream of the in-frame stop codon in the RF-2 gene and homopolymeric sequences prior to stop codons (Atkins et al., 1990; Weiss et al., 1987, 1990a). Examples of the influence of nucleotides downstream from the stop codon on the efficiency of the stop signal have already been discussed; however, one further specific example is found in the translational control of the GCN4 gene of *S. cerevisiae*. In this case 10 bases downstream of a stop codon are important for a strong translational barrier (Miller & Hinnebusch, 1989, 1990).

A summary of the potential influences on a stop codon which have been described in the previous sections is shown in Figure

2. This figure illustrates that the region of the mRNA which is in contact with the ribosome, that is, 18 nucleotides upstream and the 12 nucleotides downstream from the beginning of the termination codon, can potentially influence the functioning of the stop signal.

Stop Signal Decoding—Its Interactions with the Ribosome and Release Factor. (i) The Decoding Site. Recent work has focused on the role of the *E. coli* 16S rRNA in termination and possible RNA/RNA decoding. The 16S rRNA of the *E. coli* ribosome has been cross-linked to a UGA trinucleotide in a termination complex with release factor (Lang et al., 1989). A more defined approach utilized small designed mRNAs containing a thio-U nucleotide as part of the stop signal placed in the ribosomal A site. This thio residue can act as a zero-length cross-linking reagent to identify parts of the ribosome or release factors in close contact with the thio-UAA. With this approach a single cross-linked residue in the 16S rRNA, A1408, was detected (Tate et al., 1990a). The same residue can also be protected from chemical attack by an A site bound tRNA (Noller, 1991). This region of the rRNA has been placed in the cleft of the small ribosomal subunit and within that defined as the decoding site (Brimacombe, 1988; Stern et al., 1989). It has been suggested that recognition of the stop signal may occur as a base-paired structure with the rRNA, and it is therefore of interest that residue 1408 forms part of a single-stranded region of sequence Um^{5}CA1408, complementary in three positions with UGA and in two positions with UAA and UAG. Interaction of the m^{5}C with the middle A of UAA and UAG is possible however, since m^{5}C can base pair with A in the wobble position during the decoding of UGA as selenocysteine by a specific tRNA (Hatfield et al., 1990).

This 16S rRNA sequence is highly conserved in most ribosomal RNAs except those of a few mitochondria (which do not use UGA as stop), and it seems to be a key region in the fidelity of decoding (Noller, 1991). It is also very close to the supposed antibiotic binding sites of hygromycin and neomycin which inhibit all termination reactions (Tate et al., 1990b). Ribosomal proteins S18 and S21 were also identified as being close to the stop signal by cross-linking from the thio-U in this site (Tate et al., 1990a), and these proteins have also been located adjacent to the cleft in the subunit which defines the decoding site (Moore & Capel, 1988).

(ii) Other Regions of the Small Subunit. An alternative model has been suggested for the decoding of UGA through base pairing with tandem UCA triplets (1192–1204) in helix 34 of the 16S rRNA (Brimacombe's numbering; Brimacombe, 1988). This model was derived from the observation of Murgola and colleagues that an *E. coli* mutant which suppressed some UGA stop codons had a deletion of a single base, C1054 (Murgola et al., 1988). Furthermore, changing the nucleotide at this position to any other nucleotide also resulted in elevated suppression of UGA (Hanfler et al., 1990). The tandem UCA triplets are located near this nucleotide in this same helix. Goringer and colleagues have altered the triplets by site-directed mutagenesis and found that the double mutant is lethal while a single mutant has specifically enhanced UGA suppression (Goringer et al., 1991). Phylogenetically these UCA triplets are well conserved in the rRNA. At least one is found in all eubacteria and eukaryotes, including mycoplasma (contrary to earlier reports) which do not use UGA for stop, although most archaebacteria have neither despite using UGA as stop (Neefs et al., 1991; Olsen et al., 1991).

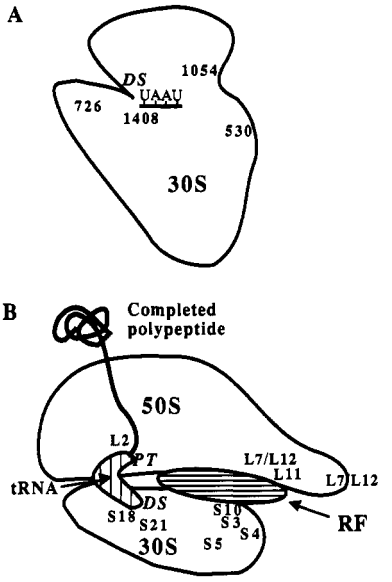

FIGURE 3: (A) Approximate location of regions of the 16S rRNA implicated in translational termination, derived from current models of the *E. coli* 30S subunit. The numbers represent nucleotides in the 16S rRNA numbered from the 5′ terminus. Abbreviations: DS, decoding site; UAAU, a stop signal shown approximately to scale. (B) Schematic view of the RF interaction with the *E. coli* ribosome (top view). The elongated RF (horizontal stripes) is shown penetrating deep into the cavern between the subunits. The approximate locations of proteins on the large 50S subunit (L2, two dimers of L7/L12 and L11) and small 30S subunit (S3, S4, S5, S10, S18, S21) implicated in RF function are shown. For clarity the mRNA is not shown. Symbols: P site bound tRNA, vertical stripes; DS, decoding site; PT, peptidyltransferase.

A mutation at position C726 in the rRNA (distant from helix 34 on the platform side of the small subunit) also causes increased UGA suppression (Prescott & Goringer, 1990), while a mutation in yeast mitochondrial rRNA of the small subunit at a position corresponding to the G517 in the "530 loop" of the *E. coli* rRNA causes increased UAA suppression (Shen & Fox, 1989). Models involving base pairing for the decoding of UAA and UAG have not been tested. Their approximate locations in the current three-dimensional model of the 16S rRNA with respect to the decoding site, and hence the stop signal, are shown in Figure 3A.

Interpretation of genetic studies that use stop codon read through as a negative test for termination has been hampered by the multiplicity of "context effects" that act on stop codons. For example, efficiency of UAG read through varies over a 500-fold range in plant virus derived sequences (Skuzeski et al., 1991) or 3-fold depending on just the 3′ nucleotide in *E. coli* (Pederson & Curran, 1991). These effects may be due to either the termination or the competing suppression mechanism or both, and it is hazardous to compare the effects of changes in a RF-1/UAG suppressor competition at a UAG context with those in a RF-2/UGA suppressor competition at a UGA context. Very recently, for example, Prestcott et al. (1991) have reported that the C1054 deletion in 16S rRNA discussed above, specific for UGA suppression in the contexts originally tested, can suppress the other termination codons as well. More carefully defined systems are becoming available which are helping to separate the various components contributing to the net suppression observed (Martin et al., 1988; Pederson & Curran, 1991; Buckingham et al., 1990).

Do these findings of other regions distant from the decoding site involved in termination imply that there are two phases to codon recognition by the release factor, perhaps involving a major conformational change in the ribosome and site change

for the mRNA? This may involve either a prereading of the stop signal at one of these sites, perhaps out of frame (Buckingham et al., 1987; Tate et al., 1983a), or alternatively a second step following the decoding site event?

Perhaps not, since recent cross-linking studies aimed at tracing the path of the mRNA through the ribosome have suggested that these regions of the rRNA, which are placed somewhat apart on the subunit on current three-dimensional models of the ribosome, might in fact be in closer proximity (Brimacombe, personal communication). A codon in a designed mRNA positioned in the "A" decoding site of a 30S·mRNA·$tRNA_f^{Met}$ initiation complex cross-linked from a position equivalent to the third nucleotide of the stop signal (+3) to a nucleotide around 1050 in "helix 34" of the rRNA (Dontsova et al., 1991). This suggests that a UGA stop codon in this site in the mRNA might be able to make direct contact with helix 34. Moreover, a separate study using similar technology has suggested that the "1400" rRNA region at the decoding site, and the "530 loop" currently placed on the opposite side of the subunit, may indeed be in close proximity (Rinke-Appel et al., 1991). More detailed studies will help to clarify the relative locations of these structures.

(*iii*) *Interaction of the Stop Signal with RF.* The first direct evidence for physical recognition of the codon by the release factor was provided using thio-U-containing RNA with its stop codon in the A site; factor/codon cross-linked products were isolated (Tate et al., 1990a). How does this correlate with other studies defining the binding site of the release factors on the ribosome? Paradoxically the major ribosomal binding site of RF-2 has been placed on the side *opposite* the decoding site at the ribosomal interface of the 70S bacterial ribosome by immunoelectron microscopy, although it was not clear from these studies how far the factor penetrated into the cavern between the subunits and hence across to the decoding site (Kastner et al., 1990). Indeed, when the factor was cross-linked to the isolated small subunit in a functional complex, there were epitopes of the factor on both sides of the subunit, implying that the factor extended across the subunit (Kastner and Tate, unpublished results).

A number of other bacterial mutants with reduced termination activity have been characterized as being in specific ribosomal proteins (Tate et al., 1990b). In most cases these mutations are within the binding sites of the factors on the ribosome, and although not tested, their effects could be explained by a perturbation of this interaction. It is well established, however, that L7/L12 and L11 affect the ribosomal binding of RF-1 and RF-2 specifically (Tate et al., 1986, 1990b). A number of mutations have been found in specific positions in the genes for the release factors (Chang et al., 1990; Elliott & Wang, 1991; Kawakami et al., 1988a; Wu et al., 1990). While these are scattered throughout the genes, two have been identified in RF-1 (amino acids 180 and 182) in a small isolated region of homology between the yeast mitochondrial RF (Pel, personal communication) and RF-1. In these cases the relationship between the factors, ribosome, and mRNA has been perturbed in an unidentified manner. However, what is clear is that there is a selective advantage for other events if the partially disabled factors are unable to decode the stop signals with normal efficiency. For example, in two cases RF-2 mutations increased transcription termination at the *trp* operon attenuator, possibly as a result of a longer pause at the UGA stop signal of the leader peptide and thereby a greater chance for the formation of a transcription-terminating structure in the RNA (Roesser et al., 1989). Eventually it may be possible to derive information on important structural or functional domains from altered or hybrid factors.

What Happens after Stop Signal Recognition by the Release Factor? The ultimate consequence of the recognition of the stop signal in the mRNA by the release factor is the hydrolysis of the ester bond between the tRNA and the C-terminal amino acid and the release of the completed polypeptide from the ribosome. This involves the peptidyltransferase center on the large subunit, but how the codon recognition event on the small subunit communicates with this center to change it from a peptide bond forming mode to the hydrolysis mode remains obscure.

Where is the release factor in relation to the peptidyltransferase center and how might it affect its activity? The mapping of RF-2 onto the 70S ribosome by immunoelectron microscopy (Kastner et al., 1990) suggested the factor extended well into the interface between the two subunits and could make contact with the region of the large subunit known to include the peptidyltransferase center. Moreover, when 50S subunits containing RF-2 were derived from these 70S cross-linked complexes, under conditions which dissociated the ribosome into its subunits, it could be seen that the factor extended from the base of the L7/L12 stalk well across the subunit. The ribosomal binding domains of the two factors overlap but are not identical. A number of studies had identified L7/L12, L11, and L16 of the large subunit and S3, S4, S5, and S10 of the small subunit as part of these binding domains (Tate et al., 1990b), with the C-terminal regions of the body dimer of L7/L12 being the critical determinants for these proteins (Tate et al., 1990c). L16 is believed to be essential for maintaining the conformation of the peptidyltransferase center (Teraoka & Nierhaus, 1978), and its importance to RF function implies that the factor requires a conformationally intact site for peptidyl-tRNA hydrolysis. RF-2 binds with higher affinity to ribosomes lacking L11. In this case the peptidyltransferase function is less dependent on L16. This suggests that the RF might interact directly with the peptidyltransferase domain to influence its conformation (Tate et al., 1983b). Moreover, L2, mapped within the center (and still a putative candidate for a peptidyltransferase enzyme), can be cross-linked to RF-2 (Stoffler et al., 1982; Lang et al., 1989). A summary of the ribosomal determinants, their spacial orientation within the two subunits of the 70S ribosome (presented as a view from the top of the ribosome), and their relationships with the release factor are shown in Figure 3B.

There is growing evidence that rRNA depleted of protein can catalyze the peptidyltransferase reaction albeit with reduced efficiency (Noller, 1991), and therefore the role of the ribosomal proteins or the release factors may be to hold the rRNA in optimum conformations to achieve maximum functional activity of either peptide formation or polypeptidyl-tRNA hydrolysis. Indeed, if the protoribosome were composed entirely of RNA, while an RNA world was evolving to an RNA–protein world, and there was no specific stop mechanism at that time, then the evolution of ribosomal proteins and release factors and their interactions at the active site might have been a secondary development to increase the efficiency not only of protein synthesis but also for its termination. Such secondary events would still not preclude the involvement of functional groups on these ribosomal proteins and factors to increase the efficiency of catalysis. This question still remains unresolved. A model representing how the ribosome, release factor, and the mRNA might interact and function in these events in *E. coli* is presented in Figure 4. It includes the major new features of direct RF/codon recog-

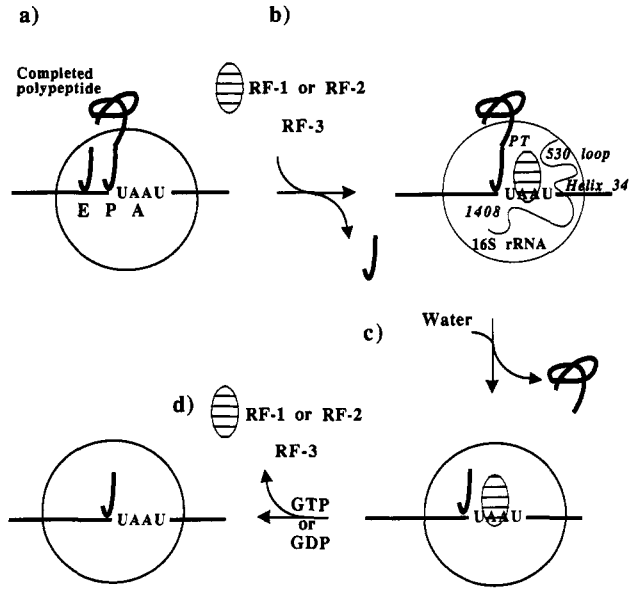

FIGURE 4: Proposed scheme for translational termination in *E. coli*. (a) Following the last round of elongation, the stop codon is located in the A site (A), the peptidyl-tRNA in the P site (P), and a deacylated tRNA in the E site (E). (b) The presence of a tetranucleotide stop signal in the decoding site (e.g., UAAU) creates a high-affinity binding site for the decoding factor (RF-1 or RF-2). This factor binds to the ribosome, penetrating into the cleft between the subunits and interacting directly with parts of the stop signal (5′ U), the 16S rRNA (nucleotide 1408 and possibly helix 34 and/or the 530 loop), and proteins of both subunits (see also Figure 3). RF-1 or RF-2 binding is enhanced by a complex with the accessory factor RF-3. The E site bound tRNA would be ejected, assuming a conformational change. (c) The RF also causes a conformational change in the peptidyltransferase (PT), and this catalyzes hydrolysis of the peptidyl-tRNA bond, allowing release of the completed polypeptide. (d) The RF is ejected. This would be followed by dissociation of the complex, or reinitiation (not shown).

nition, the involvement of the rRNA, and a four-nucleotide stop signal. In addition, the third RF, RF-3, which has received little attention since the 1970s, is included as a factor involved in the binding and dissociation cycle of the decoding factors, RF-1 and RF-2 (Goldstin & Caskey, 1970). In a series of ribosome reconstitution experiments, however, no ribosomal requirements could be found for RF-3 (apart from those already defined for RF-1 and RF-2), implying that RF-3 has a site of interaction only with the other factors (Tate, unpublished results). Although there is no information yet on how the ribosomal E site is involved in termination, we have included it in the model. The deacylated tRNA is shown as leaving the ribosome once the A-site codon is recognized by the release factor, consistent with the principles of the allosteric three-site model of Nierhaus (1990).

Structure of the Release Factors. The independence of the partial reactions of protein synthesis in which release factors 1 and 2 participate, that is, stop codon recognition, ribosome binding and peptidyl-tRNA hydrolysis, and, possibly, interaction with a third release factor (RF-3 in *E. coli*), indicates that they contain distinct functional domains and suggests that semiautonomous structural domains might exist within the proteins. The most highly conserved regions may reflect structural and functional elements essential for steps of the termination reaction common to both factors.

The amino acid sequences of *E. coli* RF-1 and RF-2 have been deduced from the sequences of their genes, and while they show only 30% identity, many of the substitutions are conservative (Craigen et al., 1985). The sequence identity is clustered particularly in the middle third of the molecules. Recently, Pel and colleagues (personal communication) have isolated the putative gene for yeast mitochondrial RF (the RF-1 homologue), and the derived amino acid sequence also shows similar clustering. We have also observed that antibodies directed against these conserved clusters in the *E. coli* factors recognize the equivalent factors from *Bacillus subtilis*. These regions may have been functionally conserved for activity. The N-terminal two-thirds of the release factor molecules are quite resistant to proteolytic digestion, but a protease-sensitive region is found adjacent to it. This susceptible sequence is highly conserved, and it has been proposed that such an unstructured domain may extend from the main ribosomal binding site of the factors across to the decoding site on the small subunit (see Figure 3) and be involved in codon recognition (Moffat et al., 1991). In the absence of a three-dimensional structure we are developing a battery of region-specific antibodies to probe for these structural elements.

The first report of a sequence of a eukaryotic release factor, the rabbit eRF, derived from the sequence of a cDNA clone surprisingly revealed no homology to the bacterial RF's but instead with bacterial tryptophanyl-tRNA synthetases (Lee et al., 1990) and near identity with mammalian tryptophanyl-tRNA synthetases (Rubin et al., 1991; Garret et al., 1991). The eRF protein has biophysical characteristics very similar to those of the synthetase. This raised the intriguing possibility that the two functions might be carried out by a single protein. However, eRF activity could not be demonstrated in extracts of a γ interferon stimulated human fibroblast, which exhibited highly elevated Trp-tRNA synthetase protein and activity (Rubin et al., 1991). The relationship between the eRF and the Trp-tRNA synthetase is currently under study in several laboratories.

Although the isolation of a yeast eRF has not been reported, Tuite and co-workers have developed a simple quantitative in vivo assay for measuring the efficiency of translation of the three termination codons in *S. cerevisiae* (Firoozan et al., 1991), and such an assay may facilitate characterization of the factor from this organism. Possible candidates have included the product of the allosuppressor gene, *SAL4/SUP1* (Tuite et al., 1990), and the psi+ determinant (Tuite et al., 1987).

Regulation of Release Factor Synthesis and Usage. Attempts to express active RF-2 in large amounts in a bacterial cell from plasmids or to clone the gene in high copy number plasmids have been largely unsuccessful, suggesting the factor is toxic in unregulated amounts, possibly due to an effect on selenoprotein synthesis at UGA sense codons (Nakamura et al., 1990). Furthermore, the specific activity of the expressed product is proportional to the level of expression, suggesting that another specific deactivation or activation mechanism may exist to control the activity of the factor posttranslationally (Adamski, Donly, and Tate, unpublished results). A number of observations demonstrate that RF-2 regulates its own synthesis. The RF-2 mRNA contains an in-frame UGA at codon 26 which can either be read as stop or be circumvented by a +1 frame shift (Craigen et al., 1985). A Shine–Dalgarno interaction six bases upstream from the stop favors the frame-shifting event (Weiss et al., 1988), whereas exogenous RF-2 has been shown to decrease frame shifting in vitro, thereby regulating its own synthesis (Craigen & Caskey, 1986; Donly et al., 1990). An altered RF-2, which can bind to the ribosome but not release a product, has been shown to abolish termination at the site in vivo (Donly et al., 1990), and a reduced level of activity in an RF-2 mutant also increased the frequency of the competing frame-shift event (Nakamura et

al., 1990). The gene for the bacterial RF-2, *prfB*, is within an operon encoding a lysyl-tRNA synthetase (*herC*) (Kawakami et al., 1988b).

No such regulatory mechanism has been demonstrated for the other factor, RF-1, but Elliott and co-workers have made the intriguing suggestion that RF-1 synthesis might also be autoregulated by the sensing of RF-1 activity at the poor stop signal, UAGC, in the mRNA for hemA (Elliott, 1989; Verlamp & Chelm, 1989). The gene for bacterial RF-1, *prfA*, has been mapped in an operon following from the *hemA* gene (encoding the first commited step in heme biosynthesis).

In *E. coli* both RF-1 and RF-2 can decode UAA whereas RF-1 is specific for the rare stop codon, UAG, and RF-2 for the more frequent UGA. Could the two release factors be utilized differentially to decode UAA stop signals depending on its context, particularly when the signal may in fact be composed of four bases? RF-1 has been shown to be favored at UAA codons which are efficiently suppressed, whereas RF-2 is preferred at poorly suppressed signals or at those signals used by genes highly expressed at high growth rates (Martin et al., 1988). RF-2 may be the preferred factor at high growth rates (Kurland, 1991), and therefore there may need to be fine-tuning mechanisms (such as the frame shifting/termination regulation) to adjust its synthesis to changes in protein synthesis rates. On the other hand, to prevent premature chain termination by the factors at slowly decoded sense codons or UGA sense codons (Parker, 1989), it may be necessary to maintain their concentrations within narrow limits.

Summary. The focus on alternative events to translational stop involving a stop codon has enlivened interest in the termination event of protein synthesis. A number of new predictions have been made from this recent flurry of new information on different aspects of peptide chain termination. It should be an exciting time as these undergo rigorous testing and a better understanding of this fundamental cellular mechanism gradually unfolds. Not only has an alternative function for stop signals in gene regulation been implicated but efforts have been renewed to redefine the nature of the stop signal. The picture unfolding is suggestive that the stop codon may be central to previously unsuspected subtle cellular regulation mechanisms, although whether the examples elucidated to date represent the enigmas or the first examples of more common events remains a critical unanswered question.

An intimate relationship between the ribosome, the mRNA, and the polypeptide chain release factors is intrinsic to the decoding of stop signals and the release of the completed polypeptide in translational termination. The focus on the rRNA as a dynamic structure possibly having a "ribozyme function", and the advances in models of the structure of these RNAs, means the events of translational termination can be visualized in three dimensions. As a result, a better appreciation of the mechanisms of factor-mediated stop signal recognition and polypeptidyl-tRNA hydrolysis events is developing. Complementary advances in our knowledge of release factor structural and functional domains must also occur, and studies are underway to provide this.

Acknowledgments

We thank colleagues who supplied preprints and Mervyn G. Smith, Mark E. Dalphin, and John G. Moffat for their critical comments on the manuscript.

References

Angenon, G., Vanmontagu, M., & Depicker, A. (1990) *FEBS Lett. 271*, 144–146.

Atkins, J. F., Weiss, R. B., & Gesteland, R. F. (1990) *Cell 62*, 413–423.

Bock, A., Forchhammer, K., Heider, J., Leinfelder, W., Sawers, G., Veprek, B., & Zinoni, F. (1991) *Mol. Microbiol. 5*, 515–520.

Brimacombe, R. (1988) *Biochemistry 27*, 4207–4213.

Brown, C. M., Stockwell, P. A., Trotman, C. N. A., & Tate, W. P. (1990a) *Nucleic Acids Res. 18*, 2079–2086.

Brown, C. M., Stockwell, P. A., Trotman, C. N. A., & Tate, W. P. (1990b) *Nucleic Acids Res. 18*, 6339–6345.

Buckingham, K., Chung, D. G., Neilson, T., & Ganoza, M. C. (1987) *Biochim. Biophys. Acta 909*, 92–98.

Buckingham, R. H., Murgola, E. J., Sorensen, P., Pagel, F. T., Hijazi, K. A., Mims, B. H., Figueroa, N., Brechemier-Baey, D., & Coppin-Raynal, E. (1990) in *The Ribosome—Structure, Function, and Evolution* (Dahlberg, A., Hill, W. E., Garrett, R. A., Moore, P. B., Schlessinger, D., & Warner, J. R., Eds.) pp 541–545, American Society for Microbiology, Washington, DC.

Caskey, C. T. (1980) *Trends Biochem. Sci.*, 234–237.

Cavener, D. R., & Ray, S. C. (1991) *Nucleic Acids Res. 19*, 3185–3192.

Chang, Z., Inokuchi, H., & Ozeki, H. (1990) *Jpn. J. Genet. 65*, 71–81.

Craigen, W. J., & Caskey, T. (1986) *Nature 322*, 273–275.

Craigen, W. J., Cook, R. G., Tate, W. P., & Caskey, C. T. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 3616–3620.

Craigen, W. J., Lee, C. C., & Caskey, C. T. (1990) *Mol. Microbiol. 4*, 861–865.

Denisenko, O. N., & Yarchuk, O. B. (1989) *FEBS Lett. 247*, 251–254.

Donly, B. C., Edgar, C. D., Adamski, F. M., & Tate, W. P. (1990) *Nucleic Acids Res. 18*, 6517–6522.

Dontsova, O., Kopylov, A., & Brimacombe, R. (1991) *EMBO J. 10*, 2613–2620.

Eggertsson, G., & Soll, D. (1988) *Microbiol. Rev. 52*, 354–374.

Elliott, T. (1989) *J. Bacteriol. 171*, 3948–3960.

Elliott, T., & Wang, X. H. (1991) *J. Bacteriol. 173*, 4144–4154.

Firoozan, M., Grant, C. M., Duarte, J. A. B., & Tuite, M. F. (1991) *Yeast 7*, 173–183.

Garret, M., Pajot, B., Trezeguet, V., Labouesse, J., Merle, M., Gandar, J.-C., Benedetto, J.-P., Sallafranque, M.-L., Alterio, J., Gueguen, M., Sarger, C., Labouesse, B., & Bonnet, J. (1991) *Biochemistry 30*, 7809–7817.

Goldstein, J. L., & Caskey, C. T. (1970) *Proc. Natl. Acad. Sci. U.S.A. 67*, 537–543.

Goringer, H. U., Hijazi, K. A., Murgola, E. J., & Dahlberg, A. E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 6603–6607.

Hanfler, A., Kleuvers, B., & Goringer, H. U. (1990) *Nucleic Acids Res. 18*, 5625–5632.

Hatfield, D., & Oroszlan, S. (1990) *Trends Biochem. Sci. 15*, 186–190.

Hatfield, D. L., Smith, D., Lee, B. J., Worland, P. J., & Oroszlan, S. (1990) *Crit. Rev. Biochem. Mol. Biol. 25*, 71–96.

Hill, K. E., Lloyd, R. S., Yang, J. G., Reed, R., & Burk, R. F. (1991) *J. Biol. Chem. 266*, 10050–10053.

Honigman, A., Wolf, D., Yaish, S., Falk, H., & Panet, A. (1991) *Virology 183*, 313–319.

Jacks, T. (1990) *Curr. Top. Microbiol. Immunol. 157*, 93–124.
Kano, A., Andachi, Y., Ohama, T., & Osawa, S. (1991) *J. Mol. Biol. 221*, 387–401.
Kastner, B., Trotman, C., & Tate, W. P. (1990) *J. Mol. Biol. 212*, 241–245.
Kawakami, K., Inada, T., & Nakamura, Y. (1988a) *J. Bacteriol. 170*, 5378–5381.
Kawakami, K., Jonsson, Y. H., Bjork, G. R., Ikeda, H., & Nakamura, Y. (1988b) *Proc. Natl. Acad. Sci. U.S.A. 85*, 5620–5624.
Kurland, C. G. (1991) *FEBS Lett. 285*, 165–169.
Lang, A., Friemert, C., & Gassen, H. G. (1989) *Eur. J. Biochem. 180*, 547–554.
Lee, C. C., Timms, K. M., Trotman, C. N., & Tate, W. P. (1987) *J. Biol. Chem. 262*, 3548–3552.
Lee, C. C., Craigen, W. J., Muzny, D. M., Harlow, E., & Caskey, C. T. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 3508–3512.
Lehman, N., & Jukes, T. H. (1988) *J. Theor. Biol. 135*, 203–214.
Martin, R., Weiner, M., & Gallant, J. (1988) *J. Bacteriol. 170*, 4714–4717.
Martin, R., Mogg, A. E., Heywood, L. A., Nitschke, L., & Burke, J. F. (1989) *Mol. Gen. Genet. 217*, 411–418.
Miller, P. F., & Hinnebusch, A. G. (1989) *Genes Dev. 3*, 1217–1225.
Miller, P. F., & Hinnebusch, A. G. (1990) *Biochim. Biophys. Acta 1050*, 151–154.
Moffat, J. G., Timms, K. M., Trotman, C. N. A., & Tate, W. P. (1991) *Biochimie 73*, 1113–1120.
Moore, P. B., & Capel, M. S. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 349–367.
Murgola, E. J., Hijazi, K. A., Goringer, H. U., & Dahlberg, A. E. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 4162–4165.
Nakamura, Y., Kawakami, K., & Mikuni, O. (1990) in *Post-Transcriptional Control of Gene Expression* (McCarthy, J. E. G., & Tuite, M. F., Eds.) pp 455–464, Springer-Verlag, Berlin and Heidelberg.
Neefs, J. M., Vandepeer, Y., Derijk, P., Goris, A., & Dewachter, R. (1991) *Nucleic Acids Res. 19*, 1987–2015.
Nierhaus, K. H. (1990) *Biochemistry 29*, 4997–5008.
Noller, H. F. (1991) *Annu. Rev. Biochem. 60*, 191–227.
Oba, T., Andachi, Y., Muto, A., & Osawa, S. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 921–925.
Olsen, G. J., Larsen, N., & Woese, C. R. (1991) *Nucleic Acids Res. 19*, 2017–2021.
Osawa, S., Muto, A., Jukes, T. H., & Ohama, T. (1990) *Proc. Royal Soc. London, B 241*, 19–28.
Parker, J. (1989) *Microbiol. Rev. 53*, 273–298.
Pedersen, W. T., & Curran, J. F. (1991) *J. Mol. Biol. 219*, 231–241.
Prescott, C. D., & Goringer, H. U. (1990) *Nucleic Acids Res. 18*, 5381–5386.
Rinke-Appel, J., Junke, N., Stade, K., & Brimacombe, R. (1991) *EMBO J. 10*, 2195–2202.
Roesser, J. R., Nakamura, Y., & Yanofsky, C. (1989) *J. Biol. Chem. 264*, 12284–12288.
Rubin, B. Y., Anderson, S. L., Xing, L., Powell, R. J., & Tate, W. P. (1991) *J. Biol. Chem.* (in press).
Salser, W. (1969) *Mol. Gen. Genet. 105*, 125–130.
Sharp, P. M., & Bulmer, M. (1988) *Gene 63*, 141–145.
Shen, Z., & Fox, T. D. (1989) *Nucleic Acids Res. 17*, 4535–4539.
Skuzeski, J. M., Nichols, L. M., Gesteland, R. F., & Atkins, J. F. (1991) *J. Mol. Biol. 218*, 365–373.
Stadtman, T. C. (1991) *J. Biol. Chem. 266*, 16257–16260.
Stern, S., Powers, T., Changchien, L.-M., & Noller, H. F. (1989) *Science 244*, 783–790.
Stoffler, G., Tate, W. P., & Caskey, C. T. (1982) *J. Biol. Chem. 257*, 4203–4206.
Tate, W. P., Hornig, H., & Luhrmann, R. (1983a) *J. Biol. Chem. 258*, 10360–10365.
Tate, W. P., Schulze, H., & Nierhaus, K. H. (1983b) *J. Biol. Chem. 258*, 12810–12815.
Tate, W. P., McCaughan, K. K., Ward, C. D., Sumpter, V. G., Trotman, C. N. A., Stoffler, M. M., Maly, P., & Brimacombe, R. (1986) *J. Biol. Chem. 261*, 2289–2293.
Tate, W., Greuer, B., & Brimacombe, R. (1990a) *Nucleic Acids Res. 18*, 6537–6544.
Tate, W. P., Brown, C. M., & Kastner, B. (1990b) in *The Ribosome—Structure, Function, and Evolution* (Dahlberg, A., Hill, W. E., Garrett, R. A., Moore, P. B., Schlessinger, D., & Warner, J. R., Eds.) pp 393–401, American Society for Microbiology, Washington, DC.
Tate, W. P., Kastner, B., Edgar, C. D., McCaughan, K. K., Timms, K. M., Trotman, C. N. A., Stoffler-Meilicke, M., Stoffler, G., Nag, B., & Traut, R. R. (1990c) *Eur. J. Biochem. 187*, 543–548.
Teraoka, H., & Nierhaus, K. H. (1978) *FEBS Lett. 88*, 223–226.
Tuite, M. F., Cox, B. S., & McLaughlin, C. S. (1987) *FEBS Lett. 225*, 205–208.
Tuite, M. F., Akhmaloka, Firoozan, M., Duarte, J. A. B., & Grant, C. M. (1990) in *Post-Transcriptional Control of Gene Expression* (McCarthy, J. E. G., & Tuite, M. F., Eds.) pp 611–622, Springer-Velag, Berlin and Heidelberg.
Valle, R. P., & Morch, M. D. (1988) *FEBS Lett. 235*, 1–15.
Verlamp, E., & Chelm, B. (1989) *J. Bacteriol. 171*, 4728–4735.
Weiss, R. B., Dunn, D. M., Atkins, J. F., & Gesteland, R. F. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 687–693.
Weiss, R. B., Dunn, D. M., Dahlberg, A. E., Atkins, J. F., & Gesteland, R. F. (1988) *EMBO J. 7*, 1503–1507.
Weiss, R. B., Dunn, D. M., Atkins, J. F., & Gesteland, R. F. (1990a) *Prog. Nucleic Acid Res. Mol. Biol. 39*, 159–183.
Weiss, R. B., Huang, W. M., & Dunn, D. M. (1990b) *Cell 62*, 117–126.
Wills, N. M., Gesteland, R. F., & Atkins, J. F. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 6991–6995.
Wolin, S. L., & Walter, P. (1988) *EMBO J. 7*, 3559–3569.
Wu, E. D., Inokuchi, H., & Ozeki, H. (1990) *Jpn. J. Genet. 65*, 115–119.
Yano, R., & Yura, T. (1989) *J. Bacteriol. 171*, 1712–1717.