24

Dynamics of the Genetic Code

John F. Atkins
Howard Hughes Medical Institute and Department of Human Genetics
University of Utah
Salt Lake City, Utah 84112-5330;
Department of Genetics
Trinity College, Dublin, Ireland

August Böck
Lehrstuhl für Mikrobiologie der Universität München
80638 München, Germany

Senya Matsufuji
Department of Biochemistry II
Jikei University School of Medicine
Tokyo 105-8461, Japan

Raymond F. Gesteland
Howard Hughes Medical Institute and Department of Human Genetics
University of Utah
Salt Lake City, Utah 84112-5330

Shape is crucial for catalysis. In the hypothetical RNA World, the replicative RNAs, which constituted the hereditary information, also functioned as the "shapes" for catalysis. There was no need for decoding. When decoding originated, presumably discrimination between alternate coding possibilities was initially weak, but once one mode became predominant, there would have been selection to lock it in with increasing efficiency. Nontriplet translocation and nonstandard meaning of code words presumably generally approached a minimum compatible with speed and energy use optimization. In the present day, nonstandard decoding alternatives generally just contribute to a low level of translational errors, of which frameshifting errors (Atkins et al. 1972; Kurland 1992) are a grave type. However, some unknown proportion of genes in probably all organisms has special sites where efficient decoding alternatives are programmed into the mRNA. The group of mechanisms involved in redirection of decoding is called "recoding" (Gesteland et al. 1992). Either the evolution of the ability to perform recoding was coincident with evolution of the ability of the decoding apparatus to perform standard decoding, or it is a later so-

phistication. Of course, the answer is unknown, but the conservation of the required mRNA signals presented below indicates that recoding has been part of the decoding repertoire for at least several hundred million years and must therefore be favored by evolution. This is clearly distinct from the error rate that is a trade-off between energy expenditure, speed, and accuracy. In the latter case, evolution has optimized the balance as a whole. With recoding, it is the specific, nonstandard decoding events themselves that have led to selective advantage. In addition to the evolutionary questions posed by recoding, the mRNA signals involved reveal heretofore unsuspected roles of RNA sequences and structures in modern organisms.

Signals for recoding are carried in the mRNA sequence and include specification of a site in the coding sequence and stimulatory elements ranging from simple sequences on either side of the site to far removed and complex stem-loop structures. Little is known about the involvement of other cellular components, although in one case (insertion of selenocysteine in prokaryotes), a special protein factor is known to be essential.

Three classes of recoding are recognized. The meaning of specific code words can be redefined. The reading frame can be altered by ribosomes that switch from one overlapping reading frame to another, and blocks of nucleotides can be bypassed with, or without, a change in reading frame (for review, see Gesteland and Atkins 1996).

Recoding is in competition with standard decoding. For instance, the programmed ribosomal frameshifting that occurs two-thirds of the way through the *Escherichia coli dnaX* coding sequence is 50% efficient. Half the ribosomes frameshift and soon terminate at a stop codon in the new frame, whereas the other half traverse the shift site with standard decoding and synthesize a product with an extra domain. The two products are present in a 1:1 ratio in DNA polymerase III (Blinkova and Walker 1990; Flower and McHenry 1990; Tsuchihashi and Kornberg 1990; Kelman and O'Donnell 1995). The signals in other cases of recoding give different set efficiencies, sometimes for viral packaging purposes (Wickner 1989; Atkins et al. 1990; Levin et al. 1993), or the process can be regulated. The dynamic nature of recoding contrasts with the "hard-wired" derivatives of the universal code found in specialized niches such as mitochondria, where the meaning of a code word is altered wherever it occurs. In recoding, the rule change is dependent on mRNA context.

REDEFINITION

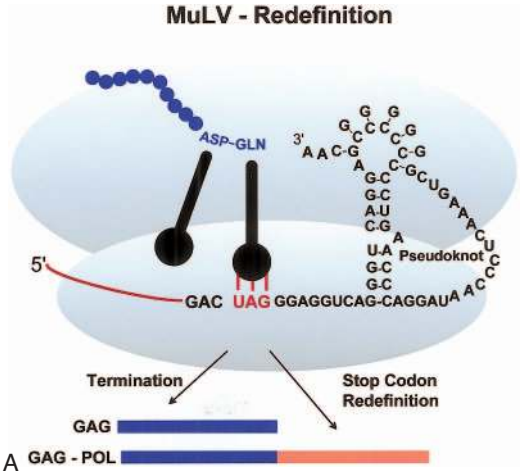
Redefinition of stop codons by recoding involves a variety of stimulatory signals. The first case of redefinition studied, the low-efficiency, but es-

sential, readthrough of the UGA at the end of the phage Qβ coat protein gene (Weiner and Weber 1973; Hofstetter et al. 1974), is probably influenced only by the identity of a few flanking nucleotides. Tate and colleagues have shown that the efficiency of release factor 2 function at UGA is influenced by the identity of the flanking 3′ nucleotide and, to a lesser extent, the following two nucleotides (Poole et al. 1998). The identity of corresponding Qβ nucleotides is expected to give somewhat inefficient termination. The identity of the last couple of amino acids encoded prior to the stop codon also influences termination efficiency (Mottagui-Tabar and Isaksson 1997). At the end of the Qβ coat protein gene, there appears to be a trade-off between amino acids that give efficient termination and 3′ nucleotides that give moderately efficient readthrough, so that modest levels of readthrough result.

In tobacco mosaic virus, the recoding signal is the identity of six nucleotides 3′ of the redefined UAG "stop" codon (Skuzeski et al. 1991; Zerfass and Beier 1992; Stahl et al. 1995). The counterpart for synthesis of the Gag-Pol precursor of murine leukemia virus is a nearby 3′ pseudoknot (Fig. 1A) (ten Dam et al. 1990; Wills et al. 1991, 1994; Felsenstein and Goff 1992; Feng et al. 1992). Although not yet studied, a pseudoknot may also be important for the putative *gag-pro* analog readthrough of the *Dictyostelium* retrotransposon, Skipper (Leng et al. 1998). For UAG stop codon redefinition in decoding barley yellow dwarf virus, both close and distant signals are important (Brown et al. 1996). Given the widespread occurrence of redefinition in plant viruses, understanding the mechanism involved is especially interesting and worthwhile. For *Drosophila* trachea branching controller, *hdc*, high-level readthrough of UAA is required, but the location of the stimulatory signals is unknown (Steneberg et al. 1998). For *Drosophila* kelch, the recoding signal is also unknown, but the efficiency of redefinition is regulated (Robinson and Cooley 1997). For these cases of redefinition, the amino acid inserted is either known, or highly likely to be one of the standard 20 amino acids having their own unique codon(s). However, the 21st (excluding formyl methionine) directly encoded amino acid, selenocysteine, is encoded only by UGA, a stop codon in the standard code.

Selenocysteine

Selenocysteine was recognized as a constituent of special proteins in 1976 (Cone et al. 1976). Experiments in two different biological systems in 1986 showed that this nonstandard amino acid is inserted cotranslationally, directed by an in-frame UGA codon in the mRNA (Chambers et al. 1986; Zinoni et al. 1986). Thus, in a single mRNA, UGA can have two



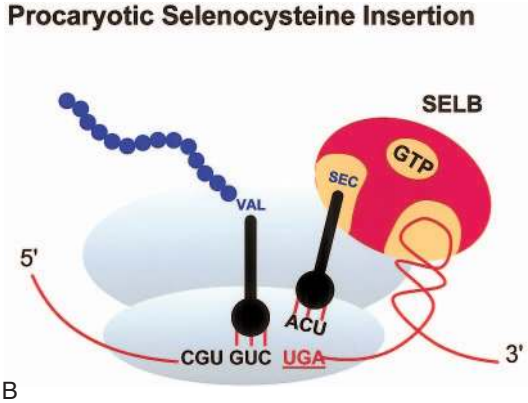


Figure 1 (*A*) Redefinition of the murine leukemia virus *gag* terminator required for synthesis of the Gag-Pol precursor. The size of loop 1 and stem 2 shown have been determined in recent work by N.M. Wills et al. (unpubl.). (*B*) Model of the ribosomal complex decoding UGA with selenocysteine. The model shows the quaternary complex between SelB, GTP, selenocysteyl-tRNA, and the SECIS element of the mRNA. Contact between SelB in the quaternary complex and the ribosome triggers GTP hydrolysis resulting in the release of selenocysteyl-tRNA (Hüttenhofer and Böck 1998b).

contrasting meanings, stop or selenocysteine insertion. This is in contrast to the use of UGA as a codon for tryptophan in *Mycoplasma* and mitochondria, where the new meaning is hard-wired and the stop function is lost. UGA thus is a "chameleon" among codons, which raises intriguing questions about whether the ancestral UGA encoded stop or an amino acid and how the evolution of the change of meaning occurred.

Distribution of Selenoproteins

Selenocysteine-containing proteins occur in all three lines of descent, but not in all organisms. Our own recent screening revealed that among twenty gram-positive and gram-negative bacteria, only five synthesized selenoproteins (S. Schorling et al., unpubl.). This finding of a nonubiquitous distribution is supported by the recent results of whole genome sequence analysis, which reveals a similar frequency. We were also unable to demonstrate the occurrence of selenoproteins in several plant cell cultures (Neuhierl and Böck 1996). Intriguingly, *Mycoplasma* species in which UGA codes only for tryptophan do not contain selenoproteins (Himmelreich et al. 1997). The numbers of selenoproteins synthesized are different among organisms. They can range from just two in the case of *Haemophilus influenzae* (Wilting et al. 1998) to three for *E. coli* (Baron and Böck 1995) to seven for *Methanococcus jannaschi* (Bult et al. 1996; Wilting et al. 1997) to an estimated more than thirty for mammals (Behne et al. 1996).

S and Se in Metabolism

Sulfur and selenium occur in the biosphere at a ratio between 10^3–10^5 to 1 and, with the exception of one major branch point leading to the specific biosynthesis of selenocysteine, they share the same metabolic paths. Free selenocysteine formed via the cysteine biosynthetic enzymes (Müller et al. 1997) can be aminoacylated onto tRNAs by cysteyl-tRNA synthetase and incorporated into any cysteine position of proteins (Müller et al. 1994). Low-molecular-weight selenocysteine is also the precursor for selenomethionine (Sliwkowski and Stadtman 1986). For our discussion, one has to keep in mind, therefore, that any cysteine in a protein is "contaminated" by selenocysteine at a ratio determined by the relative abundance of the two elements and the biochemical S/Se discrimination capacity of the respective organism. The major branch point mentioned above separates the fate of the two elements by a high-affinity metabolic route targeted to the efficient synthesis of selenocysteine under low trace element concentrations.

Selenocysteine Biosynthesis

The biosynthesis of selenocysteine in bacteria differs from that leading to cysteine as it takes place in a tRNA-bound state (Leinfelder et al. 1990; Forchhammer and Böck 1991). A specific tRNA ($tRNA^{Sec}$) is charged with L-serine by seryl-tRNA synthetase, and the seryl moiety is converted into the selenocysteyl residue by selenocysteine synthase with selenomonophosphate (SeP) as selenium donor. SeP itself is the reaction

product of selenophosphate synthetase (Leinfelder et al. 1990; Ehrenreich et al. 1992; Veres et al. 1992). Genes coding for $tRNA^{Sec}$ have been identified in many organisms within Bacteria, Archaea, and Eukarya, and their products share a number of characteristics differentiating them from ordinary elongator tRNAs. In addition to the UCA anticodon, complementary to UGA, they display sequence and architectural deviations from the consensus of classic elongator tRNAs (Baron et al. 1993; Sturchler et al. 1993). They have a 6-bp D-stem with a 4-bp loop, an extended aminoacyl-acceptor-T-stem axis of 13 bp and a large extra arm which makes these tRNAs the largest ones known. These sequence deviations have a role in maintaining novel tertiary interactions (Baron et al. 1993; Sturchler et al. 1993). $tRNA^{Sec}$, therefore, may constitute a different evolutionary line of elongator tRNAs. The structural differences are the basis of the additional functions which $tRNA^{Sec}$ has compared to elongator tRNAs, namely, serving as an adapter for the biosynthesis of selenocysteine by selenocysteine synthase, binding in a selenocysteyl-specific manner to a specialized elongation factor (see below), and precluding binding to elongation factor Tu (Baron and Böck 1991).

Biosynthesis of selenocysteine resembles that of glutamine and asparagine, which also take place starting from a precursor in the tRNA-bound state (Ibba et al. 1997). In this context, it is intriguing that archaeal genomes sequenced thus far do not contain recognizable genes for essential cysteine biosynthetic enzymes or for a cysteyl-tRNA synthetase. Thus, either the sequences of these enzymes in Archaea are highly diverged, or cysteine biosynthesis could occur from a specialized seryl-tRNA by analogy with that of selenocysteyl-tRNA (Bult et al. 1996; Ibba et al. 1997; Smith et al. 1997). If so, this would add another example of aminoacyl transformations and fill an important gap in our knowledge of the connections between evolution of the genetic code and amino acid biosynthesis (Wong 1975; Di Giulio 1997).

Elements Involved in Decoding UGA as Selenocysteine

Discrimination between UGA as the selenocysteine-specific codon and UGA as a stop codon is by an mRNA secondary/tertiary structure (the SECIS element) which, in bacteria, is located at the immediate 3′ side of UGA, i.e., within the reading frame. Swapping of the SECIS within bacteria is restricted by specific interaction of SECIS with SelB (see below) and the ribosome (Tormay and Böck 1997; Wilting et al. 1998). In Eukarya (Berry et al. 1991, 1993; Hill et al. 1993; Kollmus et al. 1996; Walczak et al. 1998) and Archaea (Wilting et al. 1997), the SECIS motif is positioned outside the reading frame in the 3′untranslated region and acts at

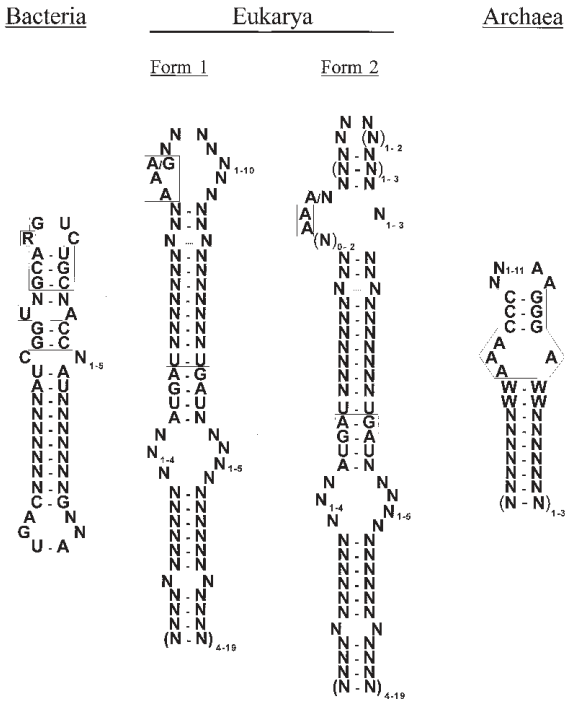


Figure 2 Consensus features of SECIS elements from Bacteria, Archaea, and Eukarya as adopted from Hüttenhofer et al. (1996), Low and Berry (1996), Wilting et al. (1997), Walczak et al. (1998), and M. Berry (pers. comm.). Consensus nucleotides are boxed and semi-conservative ones are printed in bold.

a distance of up to 2,700 bases. The SECIS motifs in bacteria are variable in sequence, based on the constraint of the reading frame, whereas those from mammalia and Archaea possess highly conserved motifs (Fig. 2). Mammalian SECIS elements are predicted to fall into two classes, which differ in the secondary structure around the conserved adenosines (Low and Berry 1996). Selenoprotein P has one of each class in its 3′ UTR, but no functional difference between the two is known (M. Berry, pers. comm.). The mammalian SECIS element can be transplanted and functions to cause selenocysteine insertion at UGA in a heterologous gene (Shen et al. 1993). Although the analysis is at an early stage, there appears to be remarkable conservation between the SECIS element(s) in lower metazoans and mammals. *Caenorhabditis elegans* thioredoxin reductase is almost certainly a selenoprotein very similar to its mammalian counterpart. The 3′ UTR of its mRNA shows a putative SECIS element that has only a single, nonsignificant, substitution from the highly conserved ver-

tebrate-like SECIS (M. Berry, pers. comm.). This SECIS, however, is very different from those identified in Archaea.

SelB, first discovered in *E. coli* (Forchhammer et al. 1989), is a specialized translation factor that interacts with guanine nucleotides, selenocysteyl-tRNA, and the SECIS element in mRNA forming a quaternary complex (Fig. 1B) (Heider et al. 1992; Hüttenhofer et al. 1996). Within this complex, SelB attains a structure suitable for interaction with the ribosome; as a consequence, GTP is hydrolyzed and selenocysteyl-tRNA is released in the proximity of the ribosomal A site (Hüttenhofer and Böck 1998a,b). Despite considerable effort, a homolog to SelB has not yet been identified in Eukarya and Archaea.

The domains of translation factor SelB involved in selenocysteyl-tRNA binding and in mRNA binding can be separated. A 17-kD carboxy-terminal domain of SelB is responsible for binding to the SECIS element (Kromayer et al. 1996) and retains this property when separated from the rest of the protein. The amino-terminal part—separated by a linker domain—has considerable sequence similarity with EF-Tu, and can bind selenocysteyl-tRNA in vivo and in vitro. SelB thus is an elongation factor homologous to EF-Tu that is tethered to the mRNA by its carboxy-terminal extension.

An intriguing consequence of the mechanism used for the "localized" decoding of the UGA codon is that the substrate—selenocysteyl-tRNA—is bound to the translation factor together with the mRNA. It will be very interesting to see the spatial relationship of codon and anticodon within this complex. Is there an interaction between the two nucleic acids before the ribosome even arrives at the UGA?

Evolution of Selenocysteine Insertion

Selective Advantage of Selenocysteine

Selenoproteins that have identified functions are enzymes with selenocysteine in their active site. Natural variants containing a cysteine in this position have been identified for many of these enzymes, showing that selenocysteine per se, in most of the selenoproteins, does not possess an essential role. Mutational change of the selenocysteine to a cysteine also gives variants that are active (Axley et al. 1991; Berry et al. 1992) but have decreased overall catalytic efficiency by a factor of 300 to 400, mostly due to a reduction of the reaction velocity (Axley et al. 1991). It is clear from these studies that although selenocysteine confers a considerable catalytic advantage, it can be replaced by a cysteine in most enzymes.

Did UGA Previously Code for Cysteine?

The discovery of UGA encoding selenocysteine raises the question whether UGA was originally a "sense codon" specifying selenocysteine incorporation or whether this is a "new" development selected to expand the genetic code (Leinfelder et al. 1988). Osawa (1995) and Jukes (1990) in their "codon capture" hypothesis present a detailed pro and con discussion of the two alternatives, and they point out that the evolution of UGA from a selenocysteine to a stop codon may be very difficult to achieve. They assume that the UGN family box originally encoded both cysteine and selenocysteine, pairing with the anticodon UCA. After duplication of the tRNA gene, one of the siblings mutated to GCA, pairing with UGY codons (the present cysteine codons). The other one (UCA) was "captured" by the newly evolving amino acid tryptophan, and then changed to CCA that only pairs with UGG (the present tryptophan codon). The only remaining function of UGA then was coding for selenocysteine. In this scheme, UGA as a stop codon then appeared by mutation of UAA (Jukes 1990).

Several of these arguments are in accord with recent biochemical facts. First, it was shown that cysteine and selenocysteine are equally well accepted by cysteinyl-tRNA synthetase and incorporated into protein (Stadtman et al. 1989; Müller et al. 1994, 1997). The previous failure to detect unspecific selenocysteine incorporation might have been due to its chemical instability or to the fact that cystathionine-β-synthase has a higher affinity for selenocysteine than cysteinyl-tRNA-synthetase, relative to the substrate cysteine. The UGY codons thus can be considered to indiscriminately code for cysteine plus its selenium analog. Second, when the UGA codon in the fdhF mRNA was changed to a UGY codon, selenocysteine was still incorporated, although to a reduced extent, since selenocysteyl-tRNA bound to SelB was competing with cysteyl-tRNA.EF-Tu (Baron et al. 1989). This indicates that the UGY codon can pair with UCA of the selenocysteyl-tRNA and also leads to the conclusion that the switch from UGN to UGY and UGA would not have been detrimental, since it interchanged chemically very similar amino acids.

The assumption that the UGN codon family, and therefore UGA, originally coded for cysteine plus selenocysteine also supports speculations on why selenocysteine is incorporated at only a few specific sites and was not maintained at "neutral" positions. It is assumed that the switch to specific selenocysteine insertion with the exclusion of cysteine insertion was a continuous, step by step, optimization process, which allowed the development of all components of the insertion machinery, SECIS, $tRNA^{Sec}$, SelB, and the biosynthetic path.

A Possible Scenario

As discussed previously, selenocysteine might have been incorporated indiscriminately with cysteine, encoded by the UGN codon family and a cysteine-specific tRNA (UCA). Indiscriminate incorporation at certain positions may have conferred to that gene product a higher reactivity and thereby a selective advantage. After duplication of the anticodon and separation of the UGY (cysteine) and UGR families, UGA was maintained for the readout of cysteine and selenocysteine and UGG was "captured" by the new amino acid tryptophan (Jukes 1990). The selective advantage forced the development of the selenium biosynthesis and insertion machinery for these special positions. UGA at other positions may have been counterselected by the high reactivity of selenol residues leading to trapped folding intermediates or by oxidative inactivation due to the appearance of oxygen in the atmosphere (Leinfelder et al. 1988).

The existence of the SECIS element and of the specialized translation factor might not have been crucial at this stage, since there was no need for discrimination against chain termination. On the other hand, later development of the SECIS motif designated the special UGA as sense, specific for selenocysteine, and forced its maintenance. "Unprotected" UGA could disappear or gain a new function, e.g., termination. It is noteworthy that the SECIS elements of Bacteria, Archaea, and Eukarya bear no structural similarity, which supports the possibility of convergent evolution. It is also an open question whether the SECIS element within the coding sequence is original and the 3′ SECIS derived from it. An argument in favor of this view is that the influence of an mRNA structure on the recoding process may be easier mechanistically if the codon is in the vicinity. However, SECIS elements within coding regions are under sequence constraint and may not be suitable to direct the insertion of more than one selenocysteine residue, whereas 3′ SECIS elements are not under such constraints and have the capacity for multiple insertions (Low and Berry 1996; Wilting et al. 1997).

The human selP gene has 10 UGA codons, and recoding is promoted by two tandem SECIS elements in the 3′ UTR (Hill et al. 1993). It is not at all clear how 3′ SECIS elements can promote multiple insertions, especially in view of the apparent inefficiency of even single insertions (see below). It is difficult to imagine that a mechanism analogous to the prokaryotic example could suffice; the 3′ element would need to cycle aminoacyl tRNA from the element to each UGA as ribosomes progressed down the message.

An alternative model is that the 3′ element interacts with the 5′ end of the mRNA analogous to the well-known communication of 3′ and 5′ ends

of eukaryotic mRNAs. Through this interaction, perhaps the initiating ribosome is modified so that for its transit of the mRNA, it reads each UGA codon as selenocysteine. In this case, the 3′ element would provide a ribosome switch rather than a tRNA delivery system (Gesteland and Atkins 1996; Kollmus et al. 1996). Even if triggering the switch for recoding is inefficient, perhaps once the ribosomes are programmed for selenocysteine incorporation, the efficiency at each subsequent site would be high, and the overall efficiency of multiple and single selenocysteine incorporations may not be very different (Kollmus et al. 1996). However, recent experiments caution against a simple version of this model and can more readily be explained by the information being delivered to ribosomes at individual UGA codons. In deiodinase mRNA, there is a single selenocysteine-encoding internal UGA codon and a single SECIS element in the 3′ UTR. Increasing the number of SECIS elements had no effect on the efficiency of selenocysteine incorporation, whereas it did when the number of UGA codons was artificially increased (S.C. Low et al., pers. comm.).

The mechanism of mammalian selenocysteine incorporation is unresolved, particularly as to how the distinction is made between termination and redefinition. Two additional results may be relevant. A UGA that is less than 55–110 nucleotides from a SECIS element functions as a terminator (Martin et al. 1996; Gu et al. 1997), and at least the second UGA in SelP mRNA may sometimes function efficiently as a terminator, since foreshortened forms of SelP protein are found (Himeno et al. 1996).

Efficiency

A recent study in *E. coli* shows that the efficiency of selenocysteine insertion is low. The normal decoding of UGA by SelB-GTP-selenocysteyl-tRNA was only 2% efficient compared to decoding of UCA by EF-Tu.GTP.selenocysteyl-tRNA (with a suitably mutant tRNA; S. Suppmann et al., unpubl.). The efficiency seems limited by nonsaturating amounts of charged $tRNA^{Sec}$ and by the kinetics of the formation and resolution of SelB quaternary complex itself (S. Suppmann et al., unpubl.). This low efficiency in *E. coli* is similar to the 1–3% levels (Berry et al. 1992; Kollmus et al. 1996) measured for selenocysteine insertion in mammalia (these efficiencies are based on transient transfection experiments, but see Martin et al. 1996).

Evolution of Selenocysteine Biosynthesis

The mode of synthesis of selenocysteine is in accord with an hypothesis for the coevolution of the genetic code and amino acid biosynthesis

(Wong 1975, 1988). UGA belongs to the serine/cysteine codon family, and both cysteine and selenocysteine are synthesized from a serine precursor. Whereas cysteine is synthesized in the low-molecular state and charged to tRNA by a specific enzyme in Bacteria and Eukarya, biosynthesis of selenocysteine takes place in the tRNA-bound state. This is similar to the biosynthesis of glutaminyl-tRNA or asparaginyl-tRNA from the glutamyl or aspartyl precursors (Ibba et al. 1997) and may reflect coevolution as postulated by Wong (1975, 1988). On the other hand, additional forces may also have been involved in necessitating tRNA-bound biosynthesis: (1) Free selenocysteine is highly toxic; (2) the development of aminoacyl-tRNA synthetases with specific recognition of selenocysteine and cysteine may be difficult to reach in view of the known lack of discrimination by cysteyl-tRNA synthetase (Müller et al. 1994). With respect to our model presented above, selenocysteine biosynthesis had to switch from a co-synthesis via the cysteine biosynthetic path to the formation in the tRNA-bound state. This had to occur early in the sequence of events described, possibly after the split of the primordial UGN codon family into UGY and UGR.

Phylogeny of Sel Gene Products

The fact that selenoproteins occur in all three lines of descent and that selenocysteine is encoded in all cases by UGA supports, but by no means proves, an early evolutionary origin. A considerable number of sel genes have been cloned and sequenced in the past years allowing (with all reservations) some conclusions about relationships.

Selenocysteine Synthase. Selenocysteine synthase is a pyridoxal-phosphate-dependent enzyme. Alignment of the known sequences shows that the enzyme belongs to the α/γ-superfamily of PLP-dependent enzymes and that it has diverged very early from the γ-family. It is intriguing that the closest relatives of selenocysteine synthase are enzymes from sulfur metabolism, namely cystathionine-γ-lyase, O-acetylhomoserine sulfhydrylase, cystathionine-γ-synthase, and cystathionine-β-lyase (Tormay et al. 1998). Thus, selenocysteine synthase may have diverged early, possibly from some enzyme of sulfur metabolism.

Translation Factor SelB. A dendrogram of the known SelB sequences (Hilgenfeld et al. 1996) revealed that the part of the SelB protein that is homologous to EF-Tu displays a greater similarity in different organisms than it does to the EF-Tu sequence from the same organism. This also holds for the relationship with IF-2. One can conclude that SelB belongs to an individual class of translation factors that separated very early from

other factors involved in protein synthesis (Hilgenfeld et al. 1996). Unfortunately, sequences of SelB homologs from archaeal or eukaryal species are not yet available for comparison.

tRNA^{Sec}. The predicted tRNA^{Sec} secondary and tertiary structures are much more conserved than the primary structure. Alignment of the sequences shows that the sequence relationships are parallel to those deduced from the 16S rRNA structures of the same organisms (Tormay et al. 1994; Baron and Böck 1995), although the small size of the molecule does not allow statistically significant conclusions

Generality of Redefinition

Sense codons can be redefined to function as start codons. GUG, UUG, and AUU specify valine, leucine, and isoleucine, respectively, when at internal positions of a coding region, but when they function as an initiator they specify methionine (or formyl methionine in *E. coli*). In *E. coli* and its phages, this redefinition requires an appropriately positioned, preceding Shine-Dalgarno sequence. The process is again dynamic; for instance, in the transposon IS911, one particular AUU acts sometimes as an initiator and sometimes as an internal sense codon (Polard et al. 1991), and in the RNA phage fr, a particular UUG behaves similarly (Adhin and van Duin 1990).

The fact that the meaning of specific codons can be redefined by mRNA context raises the important possibility of specific alteration of the meaning of one internal sense codon to another. Conventional protein chemistry could easily miss such events if their efficiency was below 10%.

An intriguing question is whether a redefinition strategy is used for the insertion of additional amino acids beyond the encoded 21, or perhaps could be experimentally exploited for the targeted insertion of normally nonencoded amino acids.

These examples of redefinition of codon meaning all use triplet translocation, the standard mechanism of mRNA readout. In contrast, the next type of recoding to be considered involves altering linear readout andx thus changing the reading frame.

REDIRECTION OF LINEAR READOUT

Frameshifting: Once-only Codon Anticodon Pairing Versus Dissociation and Re-pairing

The issue of entering and maintaining the desired reading frame must have been a significant one for the early translation apparatus. Triplet RNA:RNA interactions are inherently unstable even when the stabilizing

topology of an anticodon loop is involved (for review, see Grosjean and Chantrenne 1980). This instability is important. The potential for cognate tRNA to dissociate at initial pairing at the A-site allows near-cognate tRNAs to dissociate and to be preferentially discarded (noncognates are less of an issue) (for review, see Yarus and Smith 1995). Dissociation at the P-site is essential for some types of programmed frameshifting and perhaps one way of dealing with translation errors (Menninger 1977; for review, see Heurgué-Hamard et al. 1996). The weak triplet RNA:RNA interaction is stabilized by events at the ribosome in an active way (for review, see Yarus and Smith 1995), which themselves favor discrimination. tRNA design is integral to this process. However, the instability of triplet RNA:RNA interactions, which is advantageous now, must have posed a problem for early decoding in the absence of the stabilizing role of a sophisticated ribosome and associated factors. If more codon–anticodon bases were paired in early decoding, there is a problem in comprehending how decoding could have evolved to triplet codon–anticodon pairing without wiping out the fruits of previously selected codons. One proposed scheme for early decoding (Crick et al. 1976) was that at any one time, five codon–anticodon bases were paired, but because of a ratcheting of the tRNA (Woese 1970), only triplet "decoding" was involved. An alternative, which would also not involve whole-scale scrambling of previous information, is that six codon–anticodon bases were initially involved in pairing. If this were so, a transition to triplet pairing would just result in interspersed amino acids. Another alternative is that decoding was triplet from the start, but that stacking interactions with protoribosomal RNA stabilized the pairing (Noller et al. 1986). Whatever the explanation for early decoding, it is highly likely that modern protein synthesis involves tRNA interactions with ribosomal components that stabilize codon:anticodon pairing, and presumably these have a major role, direct or indirect (Lodmell and Dahlberg 1997), in mediating framing. Even though pairing is stabilized, if a tRNA anticodon dissociates from pairing with its cognate codon within the ribosome and quickly repairs with the same codon, this would have been undetected in the experiments performed to date. If this happens, one might imagine that the function of some ribosomal component would be to minimize this dissociation. On the basis of what has been found in genetic studies with the large ribosomal protein L9 (Herbst et al. 1994; Adamski et al. 1996; C. Johnston, unpubl.), it is a candidate for having such a function in *E. coli.*

The discrimination at the initial selection of tRNAs at the ribosomal A site is impressive. However, this selectivity can get overwhelmed, with serious consequences for frame maintenance, if the balance of competing

tRNAs is upset, especially with a small minority of tRNAs (Atkins et al. 1979; Gallant and Foley 1980; Gallant and Lindsley 1993). An imbalance can lead to acceptance of a noncognate, or near-cognate, tRNA for pairing of its anticodon with the codon. This can be important for frameshifting, even if on occasion the effect of framing is not manifest until the tRNA enters the P site (for review, see Farabaugh 1996). However, in many cases efficient programmed frameshifting follows after selection of the cognate tRNA. The tRNAs that mediate frameshifting in response to an imbalance, as well as in response to other signals, are not special in terms of their anticodon loop size; they have the same size as virtually all tRNAs. In general, this means that their anticodon size is also the standard three bases. Whether the shift tRNAs, in some cases, are special in terms of their base modifications or other features (Hatfield et al. 1989; Atkins and Gesteland 1995; Brierley et al. 1997) is currently being investigated. Following initial studies with model systems (Weiss et al. 1987), it has been found that most cases of programmed frameshifting involve tRNA dissociation and triplet re-pairing in a new frame. Because of this, weak initial pairing contributes to dissociation and so to frameshifting (Tsuchihashi and Brown 1992; Curran 1993).

A key ingredient for efficient programmed frameshifting is having an overlapping codon available for re-pairing by the P-site tRNA. For +1 frameshifting (quadruplet translocation) this is achieved by having the first base of the next zero-frame codon temporarily unoccupied. This means having an empty A site. As initially found in model systems (for review, see Gallant and Lindsley 1993), and later with yeast Ty programmed frameshifting (for review, see Farabaugh 1996), having the zero-frame A-site codon as a rare codon stimulates +1 frameshifting. The tRNAs for rare codons are themselves sparse. When one of these tRNAs is specified by the codon immediately 3′ of a shift codon, the level of aminoacylation of the tRNA becomes critical for the level of frameshifting. Hence, shortage of the amino acid used to charge such a sparse tRNA can be revealed via programmed +1 frameshifting (Kawakami et al. 1993). Amino acid starvation must have been a problem in ancient times, and it is common today for many bacteria. It will be interesting to see if the expected attendant frameshifting has regulatory significance. Various movable elements, including the yeast Ty elements, become more mobile when "hard times" are encountered. It may be of selective advantage for their hosts and consequently themselves if they transpose out of important genes whose inactivation causes hard times, or if by inserting elsewhere they provide a promoter that activates expression of desirable, but heretofore silent, genes. Whatever the reason, synthesis of the transposase in Ty

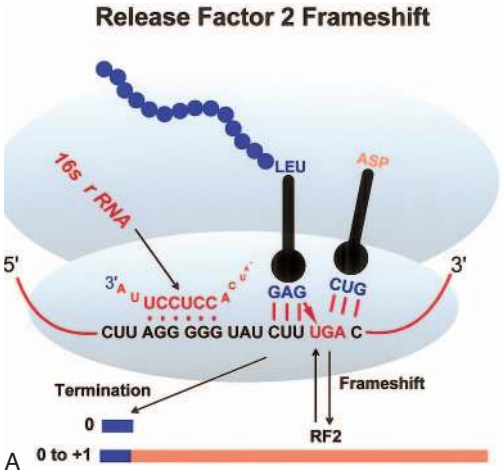
elements requires programmed +1 frameshifting that is responsive to the level of a particular aminoacylated, sparse tRNA. In addition, the yeast gene est3 (ever shorter telomeres 3), whose product is required for telomerase, has similar programmed frameshifting (Morris and Lundblad 1997). In contrast, the telomere-specific retrotransposon used for telomere maintenance in *Drosophila* apparently uses –1 frameshifting, and it is in the middle of its *gag* counterpart in distinction to the location of frameshifting in retrovirus decoding (Danilevskaya et al. 1994, 1998).

Another way to reduce competition in the A site is for the first base of the zero-frame codon 3′ adjacent to the shift codon to be part of an efficient stop codon (Poole et al. 1998). This is illustrated by the programmed frameshifting required for *E. coli* release factor 2 expression, where codons 25 and 26 are C<u>UU U</u>GA (Craigen et al. 1985). $tRNA^{Leu}$ pairs initially with the zero-frame CUU and some of the time dissociates and re-pairs with the overlapping UUU (underlined) to cause a shift to the +1 frame that encodes the rest of the release factor (Fig. 3A) (Weiss et al. 1987). Since termination at UGA is specifically mediated by release factor 2, low amounts of release factor 2 permit a greater chance of re-pairing in the +1 frame. Subsequent triplet reading of the +1 frame leads to synthesis of functional release factor 2. The converse is also true, giving an autoregulatory circuit (Craigen and Caskey 1986).

However, stops may not only stimulate shifting frame by diminishing competition from an incoming tRNA. Model systems have shown that they stimulate –1 frameshifting where a tRNA re-pairs with a triplet that overlaps the previous upstream codon (Weiss et al. 1987; 1990b; Horsfield et al. 1995). This is the basis for the programmed –1 frameshifting in decoding potato virus M that involves a single-shift tRNA (Gramstat et al. 1994), but the mechanism is unclear.

A very different way to make the last nucleotide of the previous zero-frame codon available for tRNA to re-pair with the overlapping –1 codon is for its corresponding tRNA (in the P site) to also shift –1; i.e., for tandem A-site and P-site –1 shifting. This was first discovered for retroviral, programmed –1 frameshifting (Jacks et al. 1988), and is a common type of programmed –1 frameshifting. Since both tRNAs re-pair, the characteristic shift sequence for this type of frameshifting is of the general form X-XXY-YYZ. In infectious bronchitis virus it is U-UUA-AAC (Brierley et al. 1992) and in *dnaX* it is A AAA AAG (see Fig. 3B). There have been several suggestions as to the details of re-pairing with respect to ribosomal A and P sites (Jacks et al. 1988; Weiss et al. 1989; Yelverton et al. 1994; Atkins and Gesteland 1995). However, not all programmed –1 frameshifts are tandem shifts.

Not all programmed frameshifting involves dissociation and re-pairing in an overlapping frame. In yeast Ty3 and mammalian antizyme +1 frameshifting, the evidence points to "once-only" pairing so that the first base of the next zero-frame codon is somehow unavailable for pairing with an incoming tRNA (Fig. 4A) (Farabaugh et al. 1993; Matsufuji et al. 1995). This means that finding potential frameshift sequences by looking for overlapping cognate codons will miss some examples.



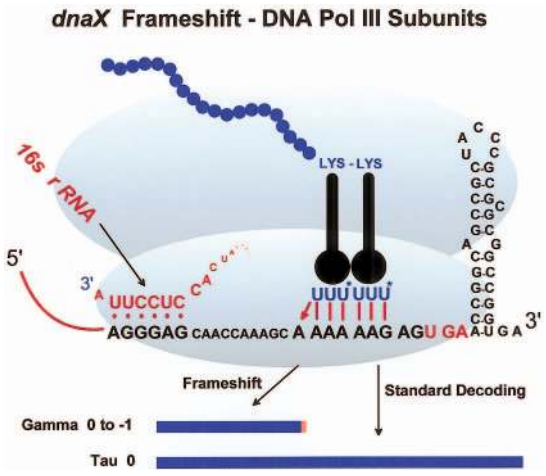


Figure 3 (*A*) The obligatory regulated +1 frameshifting required for synthesis of *E. coli* polypeptide chain release factor 2. (*B*) The tandem codon –1 frameshift required for synthesis of the γ subunit of *E. coli* DNA polymerase III.

FRAMESHIFTING: STIMULATORY SIGNALS

mRNA stimulatory signals are critical for efficient programmed frameshifting of either the "dissociation—re-pairing" or the "once-only pairing" types. In many cases, the mRNA signal(s) is 3′ of the shift site. A relatively simple stem-loop 3′ of the shift site is responsible for stimu-

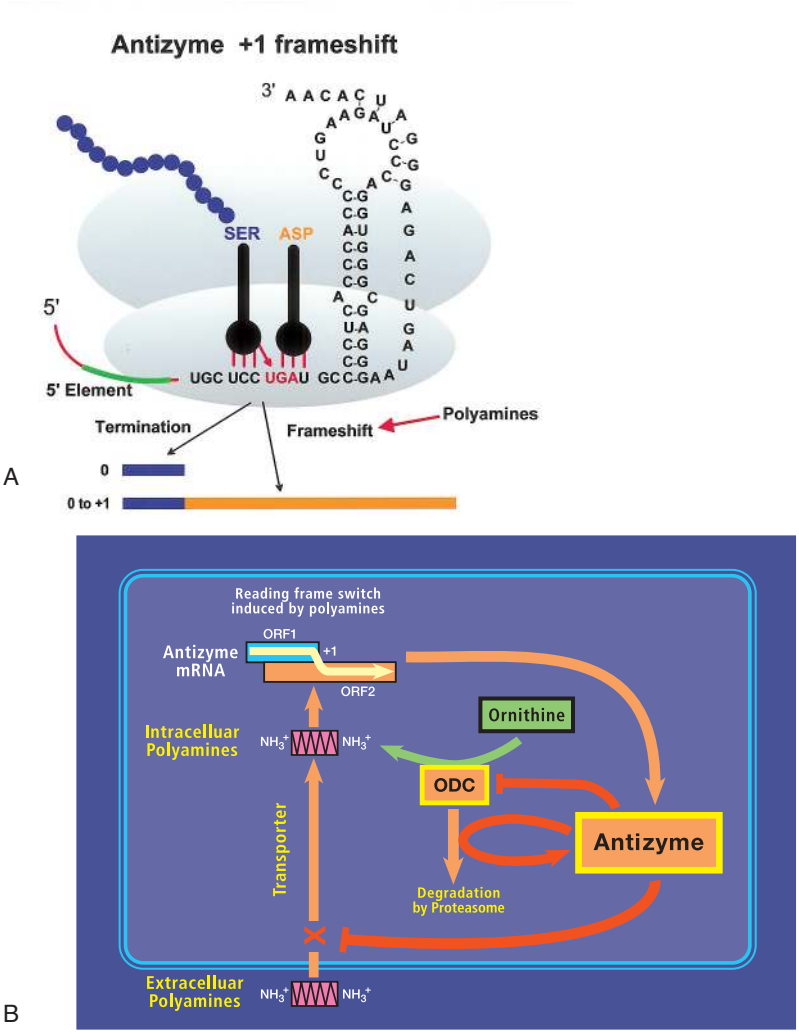


Figure 4 (*A*) The "once-only" pairing +1 frameshifting required for synthesis of human antizyme 1. (*B*) A key component of the autoregulatory circuit for polyamine homeostasis is the modulation by polyamines of the frameshifting required for antizyme synthesis.

lating the programmed frameshifting in decoding HIV *gag-pol* (Parkin et al. 1992; Bidou et al. 1997) and *E. coli dnaX* (Fig. 3B) (Larsen et al. 1997), whereas a complicated stem-loop is utilized in the bacterial transposable element, IS911 (Polard et al. 1991). In the case of *dnaX*, the frameshift efficiency is directly proportional to the predicted stability of the stem-loop structure.

The stimulatory mRNA structure is often a pseudoknot located within 8 bases 3′ of the shift site (Brierley et al. 1989; ten Dam et al. 1990). Examples are found in Coronaviruses (Brierley et al. 1991), retroviruses, the double-stranded RNA virus L-A of yeast (Dinman and Wickner 1992; Tu et al. 1992), and mammalian antizyme (Matsufuji et al. 1995, Fig. 4). With mouse mammary tumor virus *gag-pro* frameshifting, a wedge base at the junction of the two pseudoknot stems that keeps them from stacking coaxially was shown to be important for stimulation (Chen et al. 1996). However, there is controversy as to the generality of this conformation (Du et al. 1997; Sung and Kang 1998). A number of these structures are likely to interact directly with the oncoming ribosome to influence frameshifting, but it is possible that some extraribosomal factors are involved in other cases.

Tantalizingly, distant sequences in the 3′ UTR are important for programmed frameshifting in decoding barley yellow dwarf luteovirus (Miller et al. 1997), and phage T7 gene 10 (Condron et al. 1991). Phage T7 RNA polymerase transcribes faster than *E. coli* ribosomes translate and faster than *E. coli* RNA polymerase transcribes, so that distant 3′ sequences will be transcribed before ribosomes reach the shift site. The possible importance of coupling of replication and translation in some RNA animal viruses also needs investigation (Lewis and Matsui 1996).

In the few cases tested (Tu et al. 1992; Somogyi et al. 1993), pseudoknots cause pausing of ribosomes that may be necessary but not sufficient for recoding. Insertion of a stem-loop with at least equal predicted stability to a pseudoknot does not lead to frameshifting, even though it still causes pausing, albeit less effectively (Somogyi et al. 1993).

Stimulatory signals 5′ of shift sites are also found. In one case, a direct interaction with ribosomal RNA of the translocating ribosome has been characterized. Three bases 5′ of the shift site in the *E. coli* release factor 2 coding sequence, there is a Shine-Dalgarno sequence which is important for the +1 frameshifting by ribosomes that initiated 25 codons upstream of the shift site (Fig. 3A) (Weiss et al. 1987; Curran and Yarus 1988). The anti-Shine-Dalgarno sequence near the 3′ end of 16S rRNA in translocating ribosomes pairs with its mRNA complement and stimulates frameshifting (Weiss et al. 1988). Similar mRNA:16S rRNA pairing is

important for the programmed –1 frameshifting in *E. coli dnaX* decoding, but here the Shine-Dalgarno sequence is 10 bases 5′ of the shift site (Fig. 3B) (Larsen et al. 1994). Spacing of the Shine-Dalgarno sequence influences directionality of the shift at the slippery sequence. Perhaps tension in the short region of 16S rRNA between the anti-Shine-Dalgarno sequence and the part of the 16S rRNA at the decoding site upsets the framing mechanism. Since Shine-Dalgarno interactions between mRNA and rRNA were discovered for initiation before they were found to be utilized by translocating ribosomes, one tends unconsciously to think that they first arose for initiation, but of course we don't know which came first.

Although Shine-Dalgarno interactions are not used for initiation by eukaryotic ribosomes, it is much too soon to write off the possibility that, at least for translocating ribosomes and programmed frameshifting, some type of mRNA–rRNA interaction may be involved. One place to start looking is at the 5′ signal for mammalian antizyme programmed +1 frameshifting (Fig. 4A) (Matsufuji et al. 1995; Ivanov et al. 1998a,c; S. Matsufuji, in prep.). It is also possible that interactions of ribosomal RNA in translating ribosomes with mRNA sequences are not confined just to mRNA sequences 5′ of the shift site. One candidate for such an interaction is the sequence 3′ of the Ty3 shift site (Farabaugh et al. 1993).

Ancient Programmed Frameshifting

Most of the known or suspected cases of programmed frameshifting and codon redefinition, other than selenocysteine, are in viruses or transposable elements. Frameshifting is rampant in the expression of plant virus genes and probably also for bacterial insertion sequences of the IS3 family (Chandler and Fayet 1993; Ohtsubo and Sekine 1996), where approximately 60 cases are suspected (O. Fayet, pers. comm.). It is also found in the expression of quite a number of animal viruses, especially retroviruses, and also their retrotransposon counterparts. Inferring the evolutionary relationships of the recoding involved in these cases is at an early stage. Our comments on this topic, other than selenocysteine discussed above, will be restricted to the programmed frameshifting used in the expression of two nonmobile chromosomal genes. The first example is the autoregulatory frameshifting involved in decoding the bacterial gene for release factor 2 which, as described above, mediates termination at UGA (Fig. 3A).

The early evidence that the release factor 2 programmed frameshifting signals were highly similar among divergent bacteria came from a sequence comparison of the shift signals from *Bacillus subtilis* and *E. coli.* The 12 nucleotides known to be important for the autoregulatory frame-

shifting are identical (Pel et al. 1992). A recent analysis of the sequences from 20 bacteria, several of them even more distant than *B. subtilis* is from *E. coli*, has led to the inference that this frameshift mechanism was present in the common ancient ancestor of a large group of divergent bacteria but was subsequently lost in three independent lineages (Persson and Atkins 1998).

The second case is the +1 frameshifting in decoding antizyme genes. The protein antizyme governs the intracellular level of polyamines by negatively impacting the intracellular synthesis, and extracellular uptake, of polyamines. It binds to, and inactivates, ornithine decarboxylase, which catalyzes the first step of the synthesis of polyamines and also inhibits the polyamine transporter (Fig. 4B). As discovered by Matsufuji and colleagues (for review, see Gesteland et al. 1992), the programmed frameshifting required for the synthesis of antizyme is in turn regulated by polyamines, thus completing an autoregulatory circuit. Following on from the original identification of a gene in rats (Miyazaki et al. 1992), a gene for antizyme has been detected in other mammals (Tewari et al. 1994; Kankare et al. 1997; Nilsson et al. 1997), in fowl (Drozdowski et al. 1998), in zebra fish (T. Saito et al.; I.P. Ivanov et al., both unpublished), in *Xenopus* (Ichiba et al. 1995), in *Drosophila melanogaster* (Ivanov et al. 1998c), in *Schizosaccharomyces pombe* yeast, and in *C. elegans* (I.P. Ivanov, unpubl.). When a cassette containing the mammalian antizyme shift site and recoding signal is introduced into the budding yeast, *Saccharomyces cerevisiae*, high levels of frameshifting to the +1 frame occur at the shift site. However, the product has an extra amino acid as the ribosomes shift –2 instead of +1 and the utilization of the recoding signals is very different from what it is in mammals (Matsufuji et al. 1996). In contrast, the same mammalian shift cassette directs mammalian-like +1 shifting in the fission yeast, *S. pombe* (Ivanov et al. 1998b).

Recently, a second antizyme gene has been identified in mammals, and its product, antizyme 2, is distinct from the previously known mammalian antizyme 1 (Ivanov et al. 1998a). Two antizymes are now also known in zebra fish (termed Short and Long to avoid implying correspondence with the respective mammalian antizymes 1 and 2) (T. Saito et al., unpubl.). Despite substantial divergence of overall nucleotide sequence, the UGA stop codon of ORF1—the first nucleotide of which is part of the shift site (Rom and Kahana 1994; Matsufuji et al. 1995)—and 16 out of 18 nucleotides immediately 5′ of it are identical from *Drosophila* antizyme mRNA to mammalian antizymes 1 and 2 mRNAs (Ivanov et al. 1998a,c). This sequence includes much of the 5′ element discussed above, which acts in an unknown manner to stimulate frameshifting (S. Matsufuji,

unpubl.). The sequences of the stems of the stimulatory pseudoknot, 3′ of the shift site, are highly conserved between mammalian antizymes 1 and 2, but the loop sequences have diverged (Ivanov et al. 1998a). A flanking 3′ pseudoknot is not apparent in *Drosophila* by sequence inspection, but there is some sequence conservation with its mammalian counterparts in this region. It seems safe to discount convergent evolution in the case of antizyme and deduce that the shift signals have been used for efficient regulated frameshifting for hundreds of millions of years. As suggested by A.E. Dahlberg (pers. comm.), perhaps polyamines played a crucial role with primordial ribosomal RNA and, subsequently, ribosomal proteins displaced some of these roles. This raises the question of whether the sensing of polyamine levels by modern ribosomes is an evolutionary remnant.

SUBVERSION OF CONTIGUITY

Bypassing

As described above, codon–anticodon dissociation can lead to the anticodon re-pairing to an overlapping triplet resulting in frameshifting. However, the re-pairing can be elsewhere on the mRNA leading to bypassing of mRNA sequences. This was initially discovered with low efficiency (ca. 1%) to nearby sequences in special model systems (Weiss et al. 1987; O'Connor et al. 1989). However, with phage T4 gene 60 decoding, bypassing of 50 bases occurs with an efficiency of 50% from a so-called "take-off" codon to a "landing site" (Fig. 5A) (Huang et al. 1988; Weiss et al. 1990a; Maldonado and Herr 1998). The mechanism of this bypass involves 70S ribosome complexes, with peptidyl tRNA scanning the gap region to find the landing site (F. Adamski et al., unpubl.). Part of the nascent peptide, still within the ribosome, is important for this bypassing (Weiss et al. 1990a). The nascent peptide is cross-linkable to 50S subunit components (Choi and Brimacombe 1998) and appears flexible, perhaps partly folded, in an exit tunnel in that subunit. However, it is also cross-linkable to the 30S subunit, close to the decoding site. At least some of its role in bypassing may be mediated by direct contacts with the decoding area of the 30S subunit or with the tRNA–mRNA complex (Choi et al. 1998). In addition to the nascent peptide, a short stem-loop within the coding gap is important for bypassing. However, without these two special features exhibited by gene 60, efficient bypassing can occur over shorter distances if the codon following the take-off site is a rare codon and its cognate aminoacylated tRNA is limiting (J. Gallant and D. Lindsley, pers. comm.). The above-described translational bypassing is quite distinct

from the shunting of 40S ribosomal subunits to another site within the 5 untranslated regions of cauliflower mosaic virus and adenovirus mRNAs. Here the intervening sequence is not traversed; rather, specific structures appear to pass the ribosomal subunit from one site to the other (Fütterer et al. 1993; Yueh and Schneider 1996; Hemmings-Mieszczak et al. 1997).

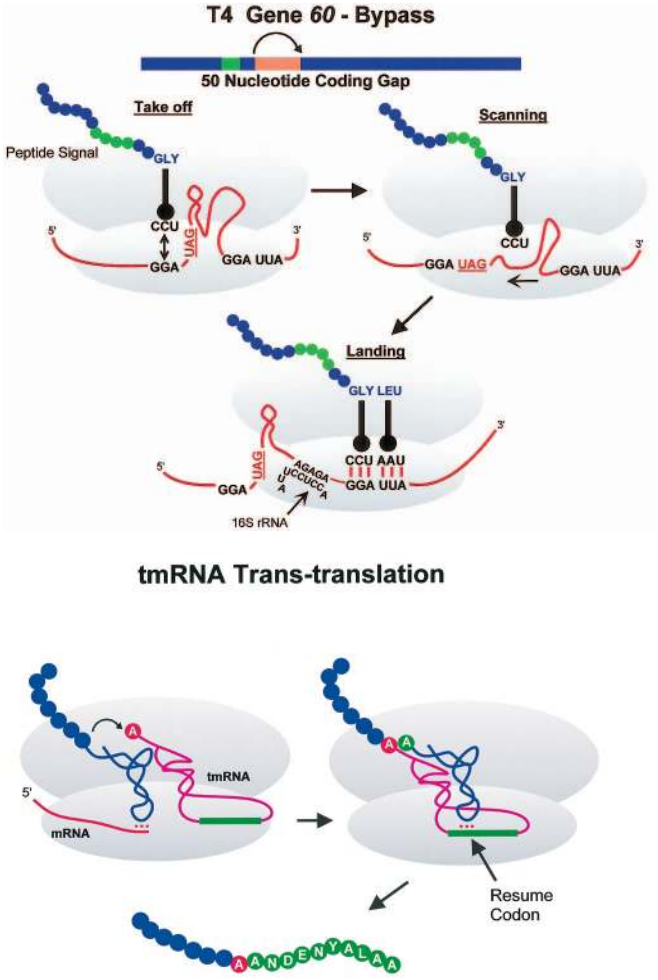


Figure 5 (*A*) The recoding signal for the 50% efficient translational bypass of 50 nucleotides in decoding phage T4 gene *60*. (*B*) Rescue of ribosomes stalled at the end of bacterial mRNAs lacking a terminator and degradation of the aberrant protein product, utilizes tmRNA which functions both as a tRNA and mRNA.

Clearly, if bypassing without special signals were rampant, decoding would be chaotic. Nonetheless, the current translational mechanism is capable of carrying out noncontiguous decoding of a message.

AN INTRAMOLECULAR RIBOSOME?

The primary (modern) function of mRNA is to be a linear tape, feeding through the ribosome readout machine. This linear property is, of course, a direct reflection of the information style of its DNA origins. However, as we have seen, mRNA sequences have additional roles in mRNA decoding. Do these observations provide an insight into primitive decoding early in evolution?

It is difficult to imagine how the early decoding apparatus could assemble amino acid chains according to a nucleic acid code. A major difficulty must have been maintaining high enough local concentrations of the reactants to drive reactions. However, the number of diffusible reactants might have been minimized by combined functions within a multifunctional mRNA molecule. Any such scheme would require evolution of complex RNA molecules, which seems counterintuitive. The trade-off between dealing with many reacting molecules and constructing a complex molecule may have been tilted toward the latter by evolution of RNA ligation activity.

The modern translation apparatus assembles the amino acid chain by sequential passage of aminoacyl tRNAs through A and P sites on the ribosome, as dictated by codon sequence in the mRNA. The success of in vitro protein synthesis experiments makes us think of the decoding apparatus operating in a soluble soup with a diffusible flow of substrate tRNAs into the ribosomal A site, where incorrect molecules are rejected until the correct one is identified. However, there is ample evidence that substrates and factors are not just free-floating, but rather are harbored in a ribosome megacomplex (Stapulionis et al. 1997; Kruse et al. 1998). Within this complex, aminoacyl-tRNAs are tested at the A site, but the volume available for free diffusion must be greatly reduced, aiding the reaction rate. The discharged tRNAs are likely recycled within the complex by resident acylating enzyme. The implication is that small molecules—ATP and amino acids—flow into the complex and the newly synthesized polypeptide chain emerges.

Could the megacomplex of current ribosomes be suggestive of an earlier strategy to deal with the substrate concentration problem in early evolution of a translation system? The ultimate megacomplex would have many functions and substrates in one molecule to maximize the number

of unimolecular reactions. The multiple capabilities of RNA are intriguing for thinking about a primitive, unimolecular decoding complex in the absence of proteins.

First, tmRNA combines mRNA and tRNA functions in one molecule (Fig. 5B) (Tu et al. 1995; Keiler et al. 1996; Himeno et al. 1997), hence its name "tmRNA" (Jentsch 1996; Atkins and Gesteland 1996). This 363-nucleotide RNA has 3′ and 5′ ends that come together to form a partial tRNA-like structure that can be charged with alanine (Komine et al. 1994; Ushida et al. 1994; Felden et al. 1998). It can access a ribosomal A site that has no tRNA and no mRNA codon, such as at the end of an mRNA with no stop codon. The alanine is donated to the growing peptide chain in the P site just as if tmRNA was an ordinary tRNA. Then, remarkably, the ribosome reads out 10 codons from an internal part of tmRNA, adding an 11-amino-acid tag to the growing chain, targeting it for degradation (Keiler et al. 1996). tmRNA has a very complex shape with at least 4 pseudoknots (Williams and Bartel 1996; Felden et al. 1997) that must provide the conditions necessary for the tRNA ends to be in the P site and the internal coding sequence to be in the mRNA track for the A site. Although this interesting mechanism employs modern-day ribosomes, it does encourage thoughts about one molecule having multiple functions involved in decoding.

Second, peptide-bond catalysis by RNA is clearly possible (Zhang and Cech 1997), and it seems likely that the peptidyl transfer function of modern ribosomes is affected by ribosomal RNA (Nitta et al. 1998; see Chapter 8). We could imagine that at early times the catalytic center for peptide-bond formation and the informational sequence (mRNA) could be in one RNA molecule; each mRNA would need to have its own, resident catalytic center.

Third, we know that folded RNA structures within an mRNA can participate in bringing an appropriate substrate aminoacyl tRNA into the ribosome for recognition of its codon, as exemplified by selenocysteine insertion, in *E. coli.* In this case, a downstream stem-loop structure (within the coding sequence) tethers the $tRNA^{Sec}$ (via a special EF-Tu protein) in order to deliver the tRNA to the waiting UGA codon in the ribosomal A site. Could we imagine that the primitive mRNA catalytic center molecule suggested above might also have an amino acid delivery system?

There are intriguing suggestions that the amino acid acceptor branch of tRNA (acceptor arm plus TψC arm) originated independently from the anticodon branch (anticodon arm plus DHU arm) (see Chapters 3 and 8) and a reason for the origin of the acceptor branch independent of its role in protein synthesis has been proposed (see Chapter 3). Perhaps the anti-

codon branch of tRNAs originated in folded structures in internal regions of primordial mRNAs that folded back and paired with "codons" in the same mRNA. One proto-anticodon branch would need to be capable of forming for each of a limited number of amino acids. How could pairing of the proto-anticodon trigger delivery of an amino acid on a proto-acceptor branch that is not contiguous with the proto-anticodon branch? Two alternatives can be considered. One is that a stereochemical, folded RNA pocket containing an amino acid is delivered by a structurally contiguous proto-anticodon branch due to an association between the two. If so, the pocket could hold the now-positioned amino acid at the catalytic site until the next amino acid is delivered. The amino acid-specific fold could then bind another amino acid, ready for delivery when its codon was required again. Thus, a series of "fingers" with bound amino acids could play back on coding sequences within one molecule to decode a part of the RNA sequence. In this scenario, activation of the amino acid takes place at the catalytic center. Alternatively, if a 3′ end is involved in primordial aminoacylation, it could act repetitively to deliver amino acids to the catalytic site. In this scenario, the amino acid could not be held at the catalytic site by the delivery system and might be held by the catalytic site until the 3′ end delivered the subsequent amino acid. Perhaps pairing of a particular anticodon branch with its codon influences, by way of tertiary interactions, the identity of the amino acid aminoacylated to the 3′ end. (For a discussion of self-aminoacylation, see Chapter 7.) In either case, as each amino acid reaches the catalytic center, a peptide bond needs to form with the growing chain. If amino acids are delivered by folded internal pockets, then subsequent evolution of 3′ end aminoacylation is a big step. However, synthesis of proteins by the internal delivery system would provide a different milieu for the subsequent but parallel development of the 3′ end delivery system.

By this imaginary scheme a single RNA molecule is mRNA, peptide bond catalyst, and "tRNA" that acts as an amino acid collection and delivery system. Many interactions would be intramolecular; the diffusion-limited reactions would be the amino acids finding their binding pockets. Another possibility, maybe a step further in evolution, might be a two-component system, with the mRNA molecule separate from a "primitive ribosome" that had the peptide bond catalyst, the tRNAs, and the delivery system in one molecule. This "ribosome" could then act on a variety of mRNAs, much like the modern ribosome megacomplex. These scenarios are admittedly farfetched and do not deal with a number of crucial issues. However, they may be illustrative of ways to think about multifunctional RNA "mega" molecules.

PERSPECTIVE

With a more complete understanding of decoding and recoding, it may be possible to consider engineering organisms with an expanded repertoire of coding capacities to include nontraditional amino acids. The challenges are clearly formidable, but there is already an impressive start by specific manipulation of an aminoacyl tRNA synthetase (Liu et al. 1997).

Early decoding likely yielded many products from a single coding sequence, because of randomness in the mechanism. Presumably, the differing specificities of these products gave some molecules with activities that provided a survival advantage. As sophistication of the decoding system evolved, the repertoire of products from a single coding sequence must have become more limited, eventually reaching the current coding rules. Although these rules, in general, result in one protein product per mRNA, recoding examples tell us that there has been coevolution of specific mechanisms to produce more than one product. The big unknown is how many coding sequences have themselves evolved to take advantage of the diversity of expression offered by recoding. Might there be many messages where 5% of the ribosomes bypass the terminator, reading codons in what is normally considered to be the 3′UTR? How commonly does decoding of mRNAs involve frameshifting akin to mammalian antizyme? Could it be that, in some mRNAs, ribosomes bypass codons by scanning from one codon to the next cognate one, as seen in T4 gene 60? Each of these acts would result in a different protein with perhaps a new or additional function. Although current methods for protein analysis are very powerful, rarely would protein variants at the level of 5% be discovered, except by a fortuitous observation or by serious digging. This also holds true for posttranslational modifications. Again, we know a great deal about some proteins, but we have little appreciation for the overall picture. A serious attack on the "proteome" is needed to begin to understand the full diversity of the products of genes.

ACKNOWLEDGMENTS

We thank Marla Berry for characteristically helpful comments and sharing unpublished data. J.F.A. is supported by National Institutes of Health grant RO1-GM48152 and Japan Health Science Foundation grant K-1005; A.B. by Deutsche Forschungsgemeinschaft and Fonds der Chemischen Industrie; S.M. by Grants-in-Aid from the Ministry of Education, Science and Culture in Japan, and the Human Frontier Sciences Program; and R.F.G. is an investigator of the Howard Hughes Medical Institute.

REFERENCES

Adamski F.M., Atkins J.F., and Gesteland R.F. 1996. Ribosomal protein L9 interactions with 23 S rRNA: The use of a translational bypass assay to study the effect of amino acid substitutions. *J. Mol. Biol.* **261:** 357–371.

Adhin, M.R. and van Duin, J. 1990. Scanning model for translational reinitiation in eubacteria. *J. Mol. Biol.* **213:** 811–818.

Atkins J.F. and Gesteland R.F. 1995. Discontinuous triplet decoding with or without repairing by peptidyl tRNA. In *tRNA: Structure, biosynthesis and function* (ed. D. Söll and U.L. RajBhandary) pp. 471–490. ASM Press, Washington, D.C.

———. 1996. A case for *trans* translation. *Nature* **379:** 769–770.

Atkins J.F., Elseviers D., and Gorini L. 1972. Low activity of β-galactosidase in frameshift mutants of *Escherichia coli. Proc. Natl. Acad. Sci.* **69:** 1192–1195.

Atkins J.F., Weiss R.B., and Gesteland R.F. 1990. Ribosome gymnastics—Degree of difficulty 9.5, style 10.0. *Cell* **62:** 413–423.

Atkins J.F., Gesteland R.F., Reid B.R., and Anderson C.W. 1979. Normal tRNAs promote ribosomal frameshifting. *Cell* **18:** 1119–1131.

Axley M.J, Böck A., and Stadtman T.C. 1991. Catalytic properties of an *Escherichia coli* formate dehydrogenase mutant in which sulfur replaces selenium. *Proc. Natl. Acad. Sci.* **88:** 8450–8454.

Baron C. and Böck A. 1991. The length of the aminoacyl acceptor stem of the selenocysteine-specific $tRNA^{Sec}$ of *Escherichia coli* is the determinant for binding to elongation factor SELB or Tu. *J. Biol. Chem.* **266:** 20375–20379.

———. 1995. The selenocysteine-inserting tRNA species: Structure and function. In *tRNA: Structure, biosynthesis and function* (ed. D. Söll and U.L. RajBhandary), pp. 529–544. ASM Press, Washington, D.C.

Baron C., Heider J., and Böck A. 1989. Mutagenesis of selC, the gene for the selenocysteine-inserting tRNA species in *E. coli*: Effects on *in vivo* function. *Nucleic Acids Res.* **18:** 6761–6766.

Baron C., Westhof E., Böck A., and Giege R. 1993. Solution structure of selenocysteine-inserting $tRNA^{Sec}$ from *Escherichia coli.* Comparison with canonical $tRNA^{Sec}$. *J. Mol. Biol.* **231:** 274–292.

Behne D., Kyriakopoeulos A., Weiss-Nowak C., Kalcklösch M., Westphal C., and Gessner H. 1996. Newly found selenium-containing proteins in the tissues of the rat. *Biol. Trace Elem. Res.* **55:** 99–110.

Berry M.J., Banu L., Harney J.W., and Larsen P.R. 1993. Functional characterization of the eukaryotic SECIS elements which direct selenocysteine insertion at UGA codons. *EMBO J.* **12:** 3315–3322.

Berry M.J., Harney J.W., Ohama T., and Hatfield D.L. 1994. Selenocysteine insertion or termination: Factors affecting UGA codon fate and complementary anticodon: codon mutations. *Nucleic Acids Res.* **22:** 3753–3759.

Berry M.J., Mai A.L., Kieffer J.D., Harney J.W., and Larsen P.R. 1992. Substitution of cysteine for selenocysteine in type I iodothyronine deiodinase reduces the catalytic efficiency of the protein but enhances its translation. *Endocrinology* **131:** 1848–1852.

Berry M.J., Banu L., Chen Y., Mandel S.J., Kieffer J.D., Harney J.W., and Larsen P.R. 1991. Recognition of UGA as a selenocysteine codon in type I deiodinase requires sequences in the 3′ untranslated region. *Nature* **353:** 273–276.

Bidou L., Stahl G., Grima B., Liu H., Cassan M., and Rousset J.-P. 1997. *In vivo* HIV-1 frameshifting efficiency is directly related to the stability of the stem-loop stimulatory signal. *RNA* **3:** 1153–1158.

Blinkowa A.L. and Walker J.R. 1990. Programmed ribosomal frameshifting generates the *Escherichia coli* DNA polymerase III γ subunit from within the τ subunit reading frame. *Nucleic Acids Res.* **18:** 1725–1729.

Brierley I., Digard P., and Inglis S.C. 1989. Characterization of an efficient Coronavirus frameshifting signal: Requirement for an RNA pseudoknot. *Cell* **57:** 537–547.

Brierley I., Jenner A.J., and Inglis S.C. 1992. Mutational analysis of the "slippery-sequence" component of a Coronavirus ribosomal frameshifting signal. *J. Mol. Biol.* **227:** 463–479.

Brierley I., Meredith M.R., Bloys A.J., and Hagervall T.G. 1997. Expression of a Coronavirus ribosomal frameshift signal in *Escherichia coli*: Influence of tRNA anticodon modification on frameshifting. *J. Mol. Biol.* **271:** 1–14.

Brierley I., Rolley N.J., Jenner A.J., and Inglis S.J. 1991. Mutational analysis of the RNA pseudoknot component of a Coronavirus frameshifting signal. *J. Mol. Biol.* **220:** 889–902.

Brown C.M., Dinesh-Kumar S.P., and Miller A. 1996. Local and distant sequences are required for efficient read-through of the barley yellow virus-PAV coat protein gene stop codon. *J. Virol.* **70:** 5884–5892.

Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D et al. 1996. Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii. Science* **273:** 1058–1072.

Chambers I., Frampton J., Goldfarb P, Affara N., McBain W., and Harrison P.R. 1986. The structure of the mouse glutathione peroxidase gene: The selenocysteine in the active site is encoded by the "termination" codon TGA. *EMBO J.* **5:** 1221–1227.

Chandler M. and Fayet O. 1993. Translational frameshifting in the control of transposition in bacteria. *Mol. Microbiol.* **7:** 497–503.

Chen X., Kang H., Shen L.X., Chamorro M., Varmus H.E., and Tinoco I, Jr. 1996. A characteristic bent conformation of RNA pseudoknots promotes –1 frameshifting during translation of retroviral RNA. *J. Mol. Biol.* **260:** 479–483.

Choi K.M. and Brimacombe R. 1998. The path of the growing peptide chain through the 23S rRNA in the 50S ribosomal subunit; a comparative cross-linking study with three different peptide families. *Nucleic Acids Res.* **26:** 887–895.

Choi K.M., Atkins J.M., Gesteland R.F., and Brimacombe R. 1998. Flexibility of the nascent polypeptide chain within the ribosome. Contacts from the peptide N-terminus to a specific region of the 30S subunit. *Eur. J. Biochem.* **255:** 409–413.

Condron B.G., Gesteland R.F., and Atkins J.F. 1991. An analysis of sequences stimulating frameshifting in the decoding of gene 10 of bacteriophage T7. *Nucleic Acids Res.* **19:** 5607–5612.

Cone J.E., del Rio M., Davis J.N., and Stadtman T.C. 1976. Chemical characterization of the selenoprotein component of clostridial glycine reductase: Identification of selenocysteine as the organoselenium moiety. *Proc. Natl. Acad. Sci.* **73:** 2659–2663.

Craigen W.J. and Caskey C.T. 1986. Expression of peptide chain release factor 2 requires high-efficiency frameshift. *Nature* **322:** 273–275.

Craigen W.J., Cook R.G., Tate W.P., and Caskey C.T. 1985. Bacterial peptide chain release factors: Conserved primary structure and possible frameshift regulation of release factor 2. *Proc. Natl. Acad. Sci.* **82:** 3616–3620.

Crick F.H.C., Brenner S., Klug A., and Pieczenik G. 1976. A speculation on the origin of protein synthesis. *Origins Life* **7:** 389–397.

Curran J.F. 1993. Analysis of effects of tRNA: Message stability on frameshift frequency at the *Escherichia coli* RF2 programmed frameshift site. *Nucleic Acids Res.* **21:** 1837–1843.

Curran J.F. and Yarus M. 1988. Use of tRNA suppressors to probe regulation of *Escherichia coli* release factor 2. *J. Mol. Biol.* **203:** 75–83.

Danilvskaya O., Slot F., Pavlova M., and Pardue M.-L. 1994. Structure of the *Drosophila* HeT-A transposon: A retrotransposon-like element forming telomeres. *Chromosoma* **103:** 215–224.

Danilevskaya O.N., Tan C., Wong J., Alibhai M., and Pardue M.-L. 1998. Unusual features of the *Drosophila melanogaster* telomere transposable element HeT-A are conserved in *Drosophila yakuba* telomere elements. *Proc. Natl. Acad. Sci.* **95:** 3770–3775.

Di Guilio M. 1997. The origin of the genetic code. *Trends Biochem. Sci.* **22:** 49.

Dinman J.D. and Wickner R.B. 1992. Ribosomal frameshifting efficiency and *gag/gag-pol* ratio are critical for yeast M_1 double-stranded RNA virus propagation. *J. Virol.* **66:** 3669–3676.

Drozdowski B., Gong T.W.-L., and Lomax M.I. 1998. The chicken cDNA for ornithine decarboxylase antizyme. *Biochim. Biophys. Acta* **1396:** 21–26.

Du Z., Holland J.A., Hansen M.R., Giedroc D.P., and Hoffman D.W. 1997. Base-pairings within the RNA pseudoknot associated with the simian retrovirus-1 *gag-pro* frameshift site. *J. Mol. Biol.* **270:** 464–470.

Ehrenreich A., Forchhammer K., Tormay P., Vepreck B., and Böck A. 1992. Selenoprotein synthesis in *E. coli.* Purification and characterization of the enzyme catalyzing selenium activation. *Eur. J. Biochem.* **206:** 767–773.

Farabaugh P.J. 1996. Programmed translational frameshifting. *Microbiol. Rev.* **60:** 103–134.

Farabaugh P.J., Zhao H., and Vimaladithan A. 1993. A novel programmed frameshift expresses the *POL3* gene of retrotransposon Ty3 of yeast: Frameshifting without tRNA slippage. *Cell* **74:** 93–103.

Felden B., Himeno H., Muto A., McCutcheon J.P., Atkins J.F., and Gesteland R.F. 1997. Probing the structure of the *Escherichia coli* 10Sa (tmRNA). *RNA* **3:** 89–104.

Felden B., Hanawa K., Atkins J.F., Himeno H., Muto A., Gesteland R.F., McCloskey J.A., and Crain P.F. 1998. Presence and location of modified nucleotides in *E. coli* tmRNA: Structural mimicry with tRNA acceptor branches. *EMBO J.* **17:** 3188–3196.

Felsenstein K.M. and Goff S.P. 1992. Mutational analysis of the *gag-pol* junction of Moloney murine leukemia virus: Requirements for expression of the *gag-pol* fusion protein. *J. Virol.* **66:** 6601–6608.

Feng Y., Yuan H., Rein A., and Levin J.G. 1992. Bipartite signal for read-through suppression in murine leukemia virus mRNA: An eight nucleotide purine-rich sequence immediately downstream of the *gag* termination codon followed by an RNA pseudoknot. *J. Virol.* **66:** 5127–5132.

Flower A.M. and McHenry C.S. 1990. The γ subunit of DNA polymerase III holoenzyme of *Escherichia coli* is produced by ribosomal frameshifting. *Proc. Natl. Acad. Sci.* **87:** 3713–3717.

Forchhammer K. and Böck A. 1991. Selenocysteine synthase from *Escherichia coli.* Analysis of the reaction sequence. *J. Biol. Chem.* **266:** 6324–6328.

Forchhammer K., Leinfelder W., and Böck A. 1989. Identification of a novel translation factor necessary for the incorporation of selenocysteine into proteins. *Nature* **342:** 453–456.

Fütterer J., Kiss-László Z., and Hohn T. 1993. Nonlinear ribosome migration on cauliflower mosaic virus 35S RNA. *Cell* **73:** 789–802.

Gallant J. and Foley D. 1980. On the causes and prevention of mistranslation. In *Ribosomes: Structure, function and genetics* (ed. G. Chambliss et al.), pp. 615–638. University Park Press, Baltimore.

Gallant J. and Lindsley D. 1993. Ribosome frameshifting at hungry codons: Sequence rules, directional specificity and possible relationship to mobile element behaviour. *Biochem. Soc. Trans.* **21:** 817–821.

Gesteland R.F. and Atkins J.F. 1996. Recoding: Dynamic reprogramming of translation. *Annu. Rev. Biochem.* **65:** 741–768.

Gesteland R.F., Weiss R.B., and Atkins J.F. 1992. Recoding: Reprogrammed genetic decoding. *Science* **257:** 1640–1641.

Gramstat A. Prüfer D., and Rohde W. 1994. The nucleic acid-binding zinc finger protein of potato virus M is translated by internal initiation as well as by ribosomal frameshifting involving a shifty stop codon and a novel mechanism of P-site slippage. *Nucleic Acids Res.* **22:** 3911–3917.

Grosjean H. and Chantrenne H. 1980. On codon-anticodon interactions. *Mol. Biol. Biochem. Biophys.* **32:** 347–367.

Gu Q.-P., Beilstein M.A., Vendeland S.C., Lugade A., Ream W., and Whanger P.D. 1997. Conserved features of selenocysteine insertion sequence (SECIS) elements in selenoprotein W cDNAs from five species. *Gene* **193:** 187–196.

Hatfield D., Feng Y.X., Lee, B.J., Rein A., Levin J.G., and Oroszlan S. 1989. Chromatographic analysis of the aminoacyl tRNAs which are required for translation of codons at and around the ribosomal frameshift sites of HIV, HTLV-1, and BLV. *Virology* **173:** 736–742.

Heider J., Baron C., and Böck A. 1992. Coding from a distance: Dissection of the mRNA determinants required for the incorporation of selenocysteine into proteins. *EMBO J.* **11:** 3759–3766.

Hemmings-Mieszczak M., Steger G., and Hohn T. 1997. Alternative structures of the cauliflower mosaic virus 35S RNA leader: Implications for viral expression and replication. *J. Mol. Biol.* **267:** 1075–1088.

Herbst K.L., Nichols L.M., Gesteland R.F., and Weiss R.B. 1994. A mutation in ribosomal protein L9 affects ribosomal hopping during translation of gene 60 from bacteriophage T4. *Proc. Natl. Acad. Sci.* **91:** 12525–12529.

Heurgué-Hamard V., Mora L., Guarneros G., and Buckingham R.H. 1996. The growth defect in *Escherichia coli* deficient in peptidyl-tRNA hydrolyase is due to starvation for Lys-tRNA^{Lys} . *EMBO J.* **15:** 2826–2833.

Hilgenfeld R., Wilting R., and Böck A. 1996. Structural model for the selenocysteine-specific elongation factor SelB. *Biochimie* **78:** 971–978.

Hill K.E., Lloyd R.S., and Burk R.F. 1993. Conserved nucleotide sequences in the open reading frame and 3′ untranslated region of selenoprotein P mRNA. *Proc. Natl. Acad. Sci.* **90:** 537–541.

Himeno S., Chittum H.S., and Burk R.F. 1996. Isoforms of selenoprotein P in rat plasma. *J. Biol. Chem.* **271:** 15769–15775.

Himeno H., Sato M., Tadaki T., Fukushima M., Ushida C., and Muto A. 1997. *In vitro trans* translation mediated by alanine-charged 10Sa RNA. *J. Mol. Biol.* **268:** 803–808.

Himmelreich R., Plagens H., Hilbert H., Reiner B., and Herrmann R. 1997. Comparative analysis of the genomes of the bacteria *Mycoplasma pneumoniae* and *Mycoplasma genitalium. Nucleic Acids Res.* **25:** 701–712.

Hofstetter H., Monstein H.-J., and Weissman C. 1974. The readthrough protein A_1 is essential for the formation of viable Qβ particles. *Biochim. Biophys. Acta* **374:** 238–251.

Horsfield J.A., Wilson D.N., Mannering S.A., Adamski F.M., and Tate W.P. 1995. Prokaryotic ribosomes recode the HIV-1 *gag-pol*–1 frameshift sequence by an E/P site post-translocation simultaneous slippage mechanism. *Nucleic Acids Res.* **23:** 1487–1494.

Huang W.M., Ao S.-Z., Casjens S., Orlandi R., Zeikus R., Weiss R., Winge D., and Fang M. 1988. A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60. *Science* **239:** 1005–1012.

Hüttenhofer A. and Böck A. 1998a. RNA structures involved in selenoprotein synthesis. In *RNA structure and function* (ed. R. Simons and M. Grunberg-Manago), pp. 603–639. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

———. 1998b. Selenocysteine-inserting RNA elements mediate GTP hydrolysis by elongation factor SelB. *Biochemistry* **37:** 885–890.

Hüttenhofer A., Heider J., and Böck A. 1996. Interaction of the *Escherichia coli* fdhF mRNA hairpin promoting selenocysteine incorporation with the ribosome. *Nucleic Acids Res.* **20:** 3903–3910.

Ibba M., Curnow A.W., and Söll D. 1997. Aminoacyl-tRNA synthesis: Divergent routes to a common goal. *Trends Biochem. Sci.* **22:** 39–42.

Ichiba T., Matsufuji S., Miyazaki Y., and Hayashi S. 1995. Nucleotide sequence of ornithine decarboxylase antizyme cDNA from *Xenopus laevis. Biochim. Biophys. Acta* **1262:** 83–86.

Ivanov I.P., Gesteland R.F., and Atkins J.F. 1998a. A second mammalian antizyme: Conservation of programmed ribosomal frameshifting. *Genomics* (in press).

Ivanov I.P., Gesteland R.F., Matsufuji S., and Atkins J.F. 1998b. Programmed frameshifting in the synthesis of mammalian antizyme is +1 in mammals, predominantly +1 in fission yeast but –2 in budding yeast. *RNA* (in press).

Ivanov I.P., Simin K., Letsou, A., Atkins J.F., and Gesteland R.F. 1998c. The *Drosophila* gene for antizyme requires ribosomal frameshifting for expression and contains an intronic gene for snRNP Sm D3 on the opposite strand. *Mol. Cell. Biol.* **18:** 1553–1561.

Jacks T., Madhani T.H.D., Masiarz F.R., and Varmus H.E. 1988. Signals for ribosomal frameshifting in the Rous sarcoma virus *gag-pol* region. *Cell* **55:** 447–458.

Jentsch S. 1996. When proteins receive deadly messages at birth. *Science* **271:** 955–956.

Jukes T.H. 1990. Genetic code 1990. Outlook. *Experientia* **46:** 1149–1157.

Kankare K., Uusi-Oukari M., and Jänne O.A. 1997. Structure, organization and expression of the mouse ornithine decarboxylase antizyme gene. *Biochem. J.* **324:** 807–813.

Kawakami K., Pande S., Faiola B., Moore D.P., Boeke J.D., Farabaugh P.J., Strathern J.N., Nakamura Y., and Garfinkel D.J. 1993. A rare tRNA-Arg(CCU) that regulates Ty1 element ribosomal frameshifting is essential for Ty1 retrotransposition in *Saccharomyces cerevisiae. Genetics* **135:** 309–320.

Keiler K.C., Waller P.R.H., and Sauer R.T. 1996. Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* **271:** 990–993.

Kelman Z. and O'Donnell M. 1995. DNA polymerase III: Structure and function of a chromosomal replicating machine. *Annu. Rev. Biochem.* **64:** 171–200.

Kollmus H., Flohe L., and McCarthy J.E.G. 1996. Analysis of eukaryotic mRNA structure directing cotranslational incorporation of selenocysteine. *Nucleic Acids Res.* **24:** 1195–1201.

Komine Y., Kitabatake M., Yokogawa T., Nishikawa K., and Inokuchi H. 1994. A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli. Proc. Natl. Acad. Sci.* **91:** 9223–9227.

Kromayer M., Wilting R., Tormay P., and Böck A. 1996. Domain structure of the prokaryotic selenocysteine-specific elongation factor SELB. *J. Mol. Biol.* **262:** 413–420.

Kruse C., Grünweller A., Willkomm D.K., Pfeiffer T., Hartmann R.K., and Müller P.K. 1998. tRNA is entrapped in similar, but distinct, nuclear and cytoplasmic ribonucleoprotein complexes, both of which contain vigilin and elongation factor 1α. *Biochem. J.* **329:** 615–621.

Kurland C.G. 1992. Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* **26:** 29–50.

Larsen B., Gesteland R.F., and Atkins J.F. 1997. Structural probing and mutagenic analysis of the stem-loop required for *E. coli dnaX* ribosomal frameshifting: Programmed efficiency of 50%. *J. Mol. Biol.* **271:** 47–60.

Larsen B., Wills N.M., Gesteland R.F., and Atkins J.F. 1994. rRNA-mRNA base pairing stimulates a programmed –1 ribosomal frameshift. *J. Bacteriol.* **176:** 6842–6851.

Leinfelder W., Zehelein E., Mandrand-Berthelot M.A., and Böck A. 1988. Gene for a novel tRNA species that accepts L-serine and cotranslationally inserts selenocysteine. *Nature* **331:** 723–725.

Leinfelder W., Forchhammer K., Veprek B., Zehelein E., and Böck A. 1990. In vitro synthesis of selenocysteyl-tRNA$_{UCA}$ from seryl-tRNA$_{UCA}$: Involvement and characterization of the selD gene product. *Proc. Natl. Acad. Sci.* **87:** 543–547.

Leng P., Klatte D.H., Schumann G., Boeke J.D., and Steck T.L. 1998. Skipper, an LTR retrotransposon of *Dictyostelium. Nucleic Acids Res.* **26:** 2008–2015.

Levin M.E., Hendrix R.W., and Casjens S.R. 1993. A programmed translational frameshift is required for the synthesis of a bacteriophage λ tail assembly protein. *J. Mol. Biol.* **234:** 124–139.

Lewis T.L. and Matsui S.M. 1996. Astrovirus ribosomal frameshifting in an infection-transfection transient expression system. *J. Virol.* **70:** 2869–2875.

Liu D.R., Magliery T.J., Pastrnak M., and Schultz P.G. 1997. Engineering a tRNA and aminoacyl-tRNA synthetase for the site-specific incorporation of unnatural amino acids into proteins *in vivo. Proc. Natl. Acad. Sci.* **94:** 10092–10097.

Lodmell J.S. and Dahlberg A.E. 1997. A conformational switch in *Escherichia coli* 16S ribosomal RNA during decoding of messenger RNA. *Science* **277:** 1262–1267.

Low S.C. and Berry M.J. 1996. Knowing when not to stop: Selenocysteine incorporation in eukaryotes. *Trends Biochem. Sci.* **21:** 203–208.

Maldonado R. and Herr A.J. 1998. The efficiency of T4 gene 60 translational bypassing. *J. Bacteriol.* **180:** 1822–1830.

Martin G.W., Harney J.W., and Berry M.J. 1996. Selenocysteine incorporation in eukaryotes: Insights into mechanism and efficiency from sequence, structure, and spacing proximity studies of the type I deiodinase SECIS element. *RNA* **2:** 171–182.

Matsufuji S., Matsufuji T., Wills N.M., Gesteland R.F., and Atkins J.F. 1996. Reading two bases twice: Mammalian antizyme frameshifting in yeast. *EMBO J.* **15:** 1360–1370.

Matsufuji S., Matsufuji T., Miyazaki Y., Atkins J.F., Gesteland R.F., and Hayashi S. 1995. Autoregulatory frameshifting in decoding mammalian ornithine decarboxylase antizyme. *Cell* **80:** 51–60.

Menninger J.R. 1977. Ribosome editing and the error catastrophe hypothesis of cellular aging. *Mech. Ageing Dev.* **6:** 131–142.

Miller W.A., Brown C.M., and Wang S. 1997. New punctuation for the genetic code: Luteovirus gene expression. *Semin. Virol.* **8:** 3–13.

Miyazaki Y., Matsufuji S., and Hayashi S. 1992. Cloning and characterization of a rat gene encoding ornithine decarboxylase antizyme. *Gene* **113:** 191–197.

Morris D.K. and Lundblad V. 1997. Programmed translational frameshifting in a gene required for yeast telomere replication. *Curr. Biol.* **7:** 969–976.

Mottagui-Tabar S., and Isaksson L.A. 1997. Only the last amino acids in the nascent peptide influence translation termination in *Escherichia coli* genes. *FEBS Lett.* **414:** 165–170.

Müller S., Heider J., and Böck A. 1997. The path of unspecific incorporation of selenium in *Escherichia coli. Arch. Microbiol.* **168:** 421–427.

Müller S., Senn H., Gsell B., Vetter W., Baron C., and Böck A. 1994. The formation of diselenide bridges in proteins by incorporation of selenocysteine residues: Biosynthesis and characterization of (Se2)-thioredoxin. *Biochemistry* **33:** 3404–3412.

Neuhierl B. and Böck A. 1996. On the mechanism of selenium tolerance in selenium-accumulating plants. Purification and properties of a specific selenocysteine methytransferase from cultured cells of *Astragalus bisulcatus. Eur. J. Biochem.* **239:** 235–238.

Nilsson J., Koskiniemi S., Persson K., Grahn B., and Holm I. 1997. Polyamines regulate both transcription and translation of the gene encoding ornithine decarboxylase antizyme in mouse. *Eur. J. Biochem.* **250:** 223–231.

Nitta I., Ueda T., and Watanabe K. 1998. Possible involvement of *Escherichia coli* 23S ribosomal RNA in peptide bond formation. *RNA* **4:** 257–267.

Noller H.F., Asire M., Barta A., Douthwaite S., Goldstein T., Gutell R.R., Moazed D., Normanly J., Prince J.B., Stern S., Triman K., Turner S., van Stolk B., Wheaton V., Weiser B., and Woese C.R. 1986. Studies on the structure and function of ribosomal RNA. In *Structure, function and genetics of ribosomes* (ed. B. Hardesty and G. Kramer), pp. 143–163. Springer-Verlag, New York.

O'Connor M., Gesteland R.F., and Atkins J.F. 1989. tRNA hopping: Enhancement by an expanded anticodon. *EMBO J.* **8:** 4315–4323.

Ohtsubo E. and Sekine Y. 1996. Bacterial insertion sequences. *Curr. Top. Microbiol. Immunol.* **204:** 1–26.

Osawa S. 1995. *Evolution of the genetic code.* Oxford University Press, New York.

Parkin N.T., Chamorro M., and Varmus H.E. 1992. Human immunodeficiency virus type 1 *gag-pol* frameshifting is dependent on downstream mRNA secondary structure: Demonstration by expression *in vivo. J. Virol.* **66:** 5147–5151.

Pel H.J., Rep M., and Grivell L.A. 1992. 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.

Persson B.C. and Atkins J.F. 1998. Does the disparate occurrence of the autoregulatory programmed frameshifting in decoding the release factor 2 gene reflect an ancient origin with loss in independent lineages? *J. Bacteriol.* **180:** 3462–3466.

Polard P., Prère M.F., Chandler M., and Fayet O. 1991. Programmed translational frameshifting and initiation at an AUU codon in gene expression of bacterial insertion sequence IS911. *J. Mol. Biol.* **222:** 465–477.

Poole E.S., Major L.L., Mannering S.A., and Tate W.P. 1998. Translation termination in *Escherichia coli*: Three bases following the stop codon crosslink to release factor 2 and affect the decoding efficiency of UGA-containing signals. *Nucleic Acids Res.* **26:** 954–960.

Robinson D.N. and Cooley L. 1997. Examination of the function of two kelch proteins generated by stop codon suppression. *Development* **124:** 1405–1417.

Rom E. and Kahana C. 1994. Polyamines regulate the expression of ornithine decarboxylase antizyme *in vitro* by inducing ribosomal frame-shifting (correction, p. 9195). *Proc. Natl. Acad. Sci.* **91:** 3959–3963.

Shen Q., Chu F.F., and Newburger P.E. 1993. Sequences in the 3′-untranslated region of the human cellular glutathione peroxidase gene are necessary and sufficient for selenocysteine incorporation at the UGA codon. *J. Biol. Chem.* **268:** 11463–11469.

Skuzeski J.M., Nichols L.M., Gesteland R.F., and Atkins J.F. 1991. The signal for a leaky UAG stop codon in several plant viruses includes the two downstream codons. *J. Mol. Biol.* **218:** 365–373.

Sliwkowski M.X. and Stadtman T.C. 1985. Incorporation and distribution of selenium into thiolase from *Clostridium kluyveri. J. Biol. Chem.* **260:** 3140–3144.

Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H., Dubois J., Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K., Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D., Spadafora R., Vicaire R., Wang Y., Wierzbowski J., Gibson R., Jiwani N., Caruso A., Bush D., Reeve J.N. et al. 1997. Complete genome sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional analysis and comparative genomics. *J. Bacteriol.* **179:** 7135–7155.

Somogyi P., Jenner A.J., Brierley I., and Inglis S.C. 1993. Ribosomal pausing during translation of an RNA pseudoknot. *Mol. Cell. Biol.* **13:** 6931–6940.

Stadtman T.C., Davis J.N., Zehelin E., and Böck A. 1989. Biochemical and genetic analysis of *Salmonella typhimurium* and *Escherichia coli* mutants defective in specific incorporation of selenium into formate dehydrogenase and tRNAs. *Biofactors* **2:** 35–44.

Stahl G., Bidou L., Rousset J.-P., and Cassan M. 1995. Versatile vectors to study recoding: Conservation of rules between yeast and mammalian cells. *Nucleic Acids Res.* **23:** 1557–1560.

Stapulionis R., Kolli S., and Deutscher M.P. 1997. Efficient mammalian protein synthesis requires an intact F-actin system. *J. Biol. Chem.* **272:** 24980–24986.

Steneberg P., Englund C., Kronhamn J., Weaver T.A., and Samakovlis C. 1998. Translational readthrough in the *hdc* mRNA generates a novel branching inhibitor in the *Drosophila* trachea. *Genes Dev.* **12:** 956–967.

Sturchler C., Westhof E., Carbon P., and Krol A. 1993. Unique secondary and tertiary structural features of the eukaryotic selenocysteine-tRNA^{Sec}. *Nucleic Acids Res.* **21:** 1073–1079.

Sung D. and Kang H. 1998. Mutational analysis of the RNA pseudoknot involved in efficient ribosomal frameshifting in simian retrovirus-1. *Nucleic Acids Res.* **26:** 1369–1372.

ten Dam E.B., Pleij C.W.A., and Bosch L. 1990. RNA pseudoknots; translational frameshifting and readthrough on viral RNAs. *Virus Genes* **4:** 121–136.

Tewari D.S., Qian Y., Thornton R.D., Pieringer J., Taub R., Mochan E., and Tewari M. 1994. Molecular cloning and sequencing of a human cDNA encoding ornithine decarboxylase antizyme. *Biochim. Biophys. Acta* **1209:** 293–295.

Tormay P. and Böck A. 1997. Barriers to heterologous expression of a selenoprotein gene in bacteria. *J. Bacteriol.* **179:** 576–582.

Tormay P., Wilting R., Heider J., and Böck A. 1994. Genes coding for the selenocysteine-inserting tRNA species from *Desulfomicrobium baculatum* and *Clostridium thermoaceticum*: Structural and evolutionary implications. *J. Bacteriol.* **176:** 1268–1274.

Tormay P., Wilting R., Lottspeich F., Mehta P. Christen P., and Böck A. 1998. Bacterial selenocysteine synthase: Enzyme properties. *Eur. J. Biochem.* **254:** 655–661.

Tsuchihashi Z. and Brown, P.O. 1992. Sequence requirements for efficient translational frameshifting in the *Escherichia coli dnaX* gene and the role of an unstable interaction between $tRNA^{Lys}$ and an AAG lysine codon. *Genes Dev.* **6:** 511–519.

Tsuchihashi Z. and Kornberg A. 1990. Translational frameshifting generates the γ subunit of DNA polymerase III holoenzyme. *Proc. Natl. Acad. Sci.* **87:** 2516–2520.

Tu C., Tzeng T.-H., and Bruenn J.A. 1992. Ribosomal movement impeded at a pseudoknot required for frameshifting. *Proc. Natl. Acad. Sci.* **89:** 8636–8640.

Tu G.F., Reid G.E., Zhang J.G., Moritz R.L., and Simpson R.J. 1995. C-terminal extension of truncated recombinant proteins in *Escherichia coli* with 10Sa RNA decapeptide. *J. Biol. Chem.* **270:** 9322–9326.

Ushida C., Himeno H., Watanabe T., and Muto A. 1994. tRNA-like structures in 10Sa RNAs of *Mycoplasma capricolum* and *Bacillus subtilis. Nucleic Acids Res.* **22:** 3392–3396.

Veres Z., Tsai L., Scholz T.D., Polotino M., Balaban R.S., and Stadtman T.C. 1992. Synthesis of 5-methylaminomethyl-2–selenouridine in tRNAs: 31P NMR studies show the labile selenium donor synthesized by the selD gene product contains selenium bonded to phosphorus. *Proc. Natl. Acad. Sci.* **89:** 2975–2979.

Walczak R., Carbon P., and Krol A. 1998. An essential non-Watson-Crick base pair motif in 3′UTR to mediate selenoprotein translation. *RNA* **4:** 74–84.

Weiner A.M. and Weber K. 1973. A single UGA functions as a natural termination signal in the coliphage Qβ coat protein cistron. *J. Mol. Biol.* **80:** 837–855.

Weiss R.B., Huang W.M., and Dunn D.M. 1990a. A nascent peptide is required for ribosomal bypass of the coding gap in bacteriophage T4 gene 60. *Cell* **62:** 117–126.

Weiss R.B., Dunn D.M., Atkins J.F., and Gesteland R.F. 1987. Slippery run, shifty stops, backward steps, and forward hops: –2, –1, +1, +2, +5 and +6 ribosomal frameshifting. *Cold Spring Harbor Symp. Quant. Biol.* **52:** 687–693.

———. 1990b. Ribosomal frameshifting from –2 to +50 nucleotides. *Prog. Nucleic Acid Res. Mol. Biol.* **39:** 159–183.

Weiss R.B., Dunn D.M., Dahlberg A.E., Atkins J.F., and Gesteland R.F. 1988. Reading frame switch caused by base-pair formation between the 3′ end of 16S rRNA and the mRNA during elongation of protein synthesis in *Escherichia coli. EMBO J.* **7:** 1503–1507.

Weiss R.B., Dunn D.M., Shuh M., Atkins J.F., and Gesteland R.F. 1989. *E. coli* ribosomes re-phase on retroviral frameshift signals at rates ranging from 2 to 50 percent. *New Biol.* **1:** 159–169.

Wickner R.B. 1989. Yeast virology. *FASEB J.* **3:** 2257–2265.

Williams K.P. and Bartel D.P. 1996. Phylogenetic analysis of tmRNA secondary structures. *RNA* **2:** 1306–1310.

Wills N.M., Gesteland R.F., and Atkins J.F. 1991. Evidence that a downstream pseudoknot is required for translational read-through of the Moloney murine leukemia virus *gag* stop codon. *Proc. Natl. Acad. Sci.* **88:** 6991–6995.

———. 1994. Pseudoknot-dependent read-through of retroviral gag termination codons: Importance of sequences in the spacer and loop 2. *EMBO J.* **13:** 4137–4144.

Wilting R., Vamvakidou K., and Böck A. 1998. Functional expression in *Escherichia coli* of the *Haemophilus influenzae* gene coding for selenocysteine-containing selenophosphate synthetase. *Arch. Microbiol.* **169:** 71–75.

Wilting R., Schorling S., Persson B.C., and Böck A. 1997. Selenoprotein synthesis in archaea: Identification of an mRNA element of *Methanococcus jannaschii* probably directing selenocysteine insertion. *J. Mol. Biol.* **266:** 637–641.

Woese C.R. 1970. Molecular mechanics of translocation: A reciprocating ratchet mechanism. *Nature* **226:** 817–820.

Wong J.T.-F. 1975. A co-evolution theory of the genetic code. *Proc. Natl. Acad. Sci.* **72:** 1909–1912.

———. 1988. Evolution of the genetic code. *Microbiol. Sci.* **5:** 174–181.

Yarus M. and Smith D. 1995. tRNA on the ribosome. In *tRNA: Structure, biosynthesis and function* (ed. D. Söll and U.L. RajBhandary), pp. 443–469. ASM Press, Washington, D.C.

Yelverton E., Lindsley D., Yamauchi P., and Gallant J. 1994. The function of a ribosomal frameshifting signal from human immunodeficiency virus-1 in *Escherichia coli. Mol. Microbiol.* **11:** 303–313.

Yueh A. and Schneider R.J. 1996. Selective translation initiation by ribosome jumping in adenovirus-infected and heat-shocked cells. *Genes Dev.* **10:** 1557–1567.

Zerfass K. and Beier H. 1992. Pseudouridine in the anticodon GψA of plant cytoplasmic $tRNA^{Tyr}$ is required for UAG and UAA suppression in the TMV-specific context. *Nucleic Acids Res.* **20:** 5911–5918.

Zhang B. and Cech T.R. 1997. Peptide bond formation by in vitro selected ribozymes. *Nature* **390:** 96–100.

Zinoni F., Birkmann A., Stadtman T.C., and Böck A. 1986. Nucleotide sequence and expression of the selenocysteine-containing polypeptide of formate-dehydrogenase (formate-hydrogen-lyase-linked) from *Escherichia coli. Proc. Natl. Acad. Sci.* **83:** 4560–4564.