Volume 235, number 1,2, 1–15 FEB 06102 August 1988

Review Letter

Stop making sense

or

Regulation at the level of termination in eukaryotic protein synthesis

Rosaura P.C. Valle and Marie-Dominique Morch

**Institut Jacques Monod, 2 Place Jussieu, 75251 Paris Cedex 05, France*

Received 13 May 1988

An increasing number of examples of translational regulation at the level of termination has been recently reported in eukaryotes. This paper reviews our present knowledge on this topic and proposes an understanding of these regulations by relating the study of viral gene expression to a comprehensive view of the mechanisms and components of the translational process.

Viral gene expression; Nonsense suppression; Frameshift; suppressor tRNA; Protein synthesis termination; Translational accuracy

1. INTRODUCTION

Termination is the last step of mRNA translation and has never been as extensively studied as the initiation or elongation steps. However, termination has revealed itself to be the scene of unexpected events. A termination codon can play different characters in different scenarios: it can meet its usual partner, the release factor, and thus promote proper release of the terminated peptide chain, but it can also be recognized by an 'unfair' tRNA which will turn it into a sense codon, or it can even simply be ignored by a ribosome which either shifts reading frame before reaching it or slides past it. These various scenarios are just as many means of regulating gene expression in prokaryotes as well as in eukaryotes. A general view of the control of termination in eukaryotes is presented here: the concept of termination is defined first, then the different scenarios are described in some detail and, finally, an attempt is made to understand regulation itself, i.e. its significance and the factors that modulate the efficiency of termination.

Correspondence address: R.P.C. Valle, Institut Jacques Monod, 2 Place Jussieu, 75251 Paris Cedex 05, France

* The Institut Jacques Monod is an 'Institut Mixte, CNRS-Université Paris 7'

Abbreviations: AIDS, acquired immunodeficiency syndrome; AlMV, alfalfa mosaic virus; BLV, bovine leukemia virus; BNYVV, beet necrotic yellow vein virus; bp, base pair; CaMV, cauliflower mosaic virus; CarMV, carnation mottle virus; FeLV, feline leukemia virus; HIV-1, human immunodeficiency virus type 1; HTLV I, human T-cell leukemia virus type I; HTLV II, human T-cell leukemia virus type II; IBV, infectious bronchitis virus; nt, nucleotide; LTSV, lucerne transient streak virus; Mo-MuLV, Moloney murine leukemia virus; MMTV, mouse mammary tumor virus; MPMV, Mason Pfizer monkey virus; ORF, open reading frame; P-Ser, phosphoserine; Ra-MuLV, Rauscher murine leukemia virus; RF, release factor; RSV, Rous sarcoma virus; Se-Cys, selenocysteine; TMV, tobacco mosaic virus; TRV, tobacco rattle virus; TYMV, turnip yellow mosaic virus; SBWMV, soil-borne wheat mosaic virus

2. THE CONCEPT OF TERMINATION

The termination step can be defined as the active

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/88/$3.50

process involved on the ribosome for the release of a polypeptide chain from the tRNA carrying the last amino acid incorporated in this polypeptide, a process that occurs when the ribosome encounters one of the so-called termination or nonsense codons. The denomination 'termination' codon suggests that such codons specify the end of the elongation of the polypeptide chain; the denomination 'nonsense' codon suggests that these codons cannot be recognized by any tRNA. The so-called termination or nonsense codons are the three codons UAA (ochre), UAG (amber) and UGA (opal) that are found at the end of coding regions on mRNAs. However, this classical definition was jeopardized by the discovery that UGA specifies Trp in mitochondria [1] and by the existence of variations from the standard genetic code in many organisms.

Setting aside the variations concerning the meanings of several sense codons encountered in the mitochondrial DNA from many animal cells, yeast and filamentous fungi [2], it is noteworthy that a different version of the 'universal' genetic code is observed in the nucleus of some lower eukaryotes such as the ciliates. *Tetrahymena thermophila, Stylonichia lemnae* and *Paramecium primaurelia* or *tetraurelia*, all use UGA as sole termination codon. UAA interrupts the ORF of the histone H3 [3] and α-tubulin [4] genes in *T. thermophila* and *S. lemnae* respectively, whereas both UAA and UAG interrupt the ORF of the variable surface antigen [5,6] in the Paramecia; in all cases UAA and UAG encode Gln codons. A Gln-tRNA and its gene have been isolated from *T. thermophila* with an anticodon sequence complementary to the UAA codon [7]. This Gln-tRNA$_{UmUA}$ recognizes in vitro both the UAA and the UAG codons [8]. Although a Gln-tRNA with a 5′-CUA-3′ anticodon, complementary to the UAG codon has also been found [8], none of the available gene sequences to date show internal UAG codons in *T. thermophila*. In addition to UAA and possibly UAG, the standard CAA and CAG codons are used to specify Gln but they are decoded by a Gln-tRNA$_{UmUG}$ [8].

The finding that different versions of the code are encountered in the nucleus of ciliates and in the mitochondria of various other eukaryotic organisms affects the concept of the 'universality' of the genetic code and makes it necessary to relate the mechanism of termination to a given organism.

All the cases of regulation that are discussed here are taken from organisms that use the three standard UAA, UAG and UGA termination codons. In addition it should be stressed that only natural situations and no mutant systems are considered.

3. THE SCENARIOS

3.1. *'Bona fide' termination*

The sequence of events at peptide chain termination in eukaryotes (reviews [9,10]) is still sketchy. When a ribosome reaches a termination codon, a peptidyl-tRNA sits at the 'P site', and the 'A site' is free. The unique eukaryotic RF initially isolated from rabbit reticulocytes clearly recognizes any of the three termination codons, and binds together with GTP to the ribosome. Hydrolysis of the peptidyl-tRNA ensues, catalyzed by the ribosomal peptidyltransferase. The release of the completed peptide chain is followed by GTP hydrolysis that promotes liberation of the RF from the ribosome, release of the deacylated tRNA and finally dissociation of the ribosome from the mRNA.

Insights into the molecular mechanism of peptide chain termination should come from cloning and further characterization of the eukaryotic RF from different organisms. Comparing the sequence of the mammalian RF with the prokaryotic and ciliate counterparts would help in pinpointing regions of the RF involved in recognition of the termination codons and in GTP binding. Indeed, prokaryotes use three different RFs for recognition of UAA and UAG, UAA and UGA, and GTP respectively (in addition to GTP recognition, RF3 stimulates binding of RF1 and RF2 to the ribosome-termination codon complex). Thus, it could be speculated that domains related to these three different activities should be found in the unique mammalian RF. On the other hand, the putative ciliate RF, which should only recognize the UGA codon, would have a single UGA recognition domain in addition to a GTP-binding domain.

In addition, studies of the antinomic process, namely nonsense suppression (i.e. reading of a nonsense codon leading to suppression of the termination process), are also a tool to investigate further the mechanism of termination. Since

nonsense suppressor tRNAs can outcompete RF [11] for their common target, parameters that influence the activity of the former should be taken into consideration to understand the activity of the latter. In particular, the demonstration that certain codon contexts around the termination codon favour nonsense suppression (see section 4.2.1) suggests that other contexts can be favourable for termination [12]. Finally, the recent finding that an *Escherichia coli* nonsense suppressor mutant is mutated in the 16 S ribosomal RNA [13] opens the door to new experiments aimed at defining the sites in the ribosome involved in the interaction with the RF.

3.2. *tRNA recognition of termination codons*

Among tRNAs that recognize termination codons, tRNAs isolated from non-mutagenized cells are referred to as 'natural' suppressor tRNAs as opposed to 'genetic' suppressor tRNAs, extensively studied in prokaryotes and yeast [14,15]. Two types of natural suppressor tRNAs should be distinguished [16]. First, normal major cytoplasmic tRNAs that misread termination codons: these tRNAs are involved in suppression of nonsense codons in eukaryotic viral mRNAs, thereby acting as essential mediators of regulation of gene expression at the translational level [17]. Second, minor aminoacyl-tRNA species harbouring an anticodon complementary to a termination codon: these tRNAs are possibly involved in the insertion of modified amino acids into proteins in response to specific termination codon sites.

3.2.1. Nonsense suppression via misreading

The first group of natural suppressor tRNAs has been investigated in a study of gene expression in tobacco mosaic virus (TMV), a plant virus with an RNA genome of '+' polarity. Translation of the genomic RNA of this virus gives rise to two overlapping proteins initiated at the same AUG codon but differing in their C-terminal region. It was first demonstrated in vitro that the longer 183 kDa protein results from the readthrough of a leaky UAG termination codon located at the end of the gene of the shorter 126 kDa protein [18]. This was later confirmed by determining the nucleotide sequence of the viral RNA genome [19]. The observation that both proteins are also produced in vivo in tobacco protoplasts [20] led to the search for the tRNA species responsible for this natural suppression. Two closely related Tyr-tRNAs, with a 5′-GΨA-3′ anticodon, that specifically stimulate the readthrough of the TMV UAG codon in vitro, were isolated from uninfected tobacco leaves [21]. A Tyr-tRNA isolated from *Drosophila melanogaster*, also bearing a 5′-GΨA-3′ anticodon, had already been shown to act as natural UAG suppressor during in vitro translation of TMV RNA [22]. Interestingly, the Tyr isoacceptor isolated from either system and bearing a 5′-QΨA-3′ anticodon (Q is a hypermodified G base) is inactive in amber suppression. This observation suggests that the presence of the Q modification in the wobble position of the anticodon restricts recognition of the tRNA to the cognate UAC and UAU Tyr codons whereas the presence of an unmodified G permits the additional recognition of UAG. A correlation between suppression activity and undermodification in the wobble position was also observed in wheat germ [23] and lupin tRNAs [24].

Misreading of a UAG nonsense codon by a Tyr 5′-GΨA-3′ anticodon implies an unconventional G:G interaction at the first position of the anticodon. It has been proposed [23] that the G of the anticodon could adopt a '*syn*' conformation that would allow it to hydrogen bond with the G of the codon. On the contrary, the presence of a bulky side chain which projects out of the QΨA anticodon loop [25], would prevent by steric hindrance the Q nucleoside from adopting such a conformation. From a structural point of view this model would support the molecular basis for suppression of the TMV amber codon. However, recent crystal structure analyses of synthetic oligomers show that non-Watson-Crick pairings occur in B-DNA and that they do not involve the unusual '*syn-anti*' but rather an '*anti-anti*' base conformation [26,27]. Thus minor base conformations do not always need to be invoked to explain unusual base-pair interactions (see also section 4.2.2).

A search for natural suppressor tRNAs active in TMV suppression has also been made in mammalian systems. Surprisingly, neither of the two major Tyr isoacceptor tRNAs isolated from calf liver displays UAG suppressor activity: both possess a Q nucleoside in the anticodon, thus confirming again the role of the level of base modifica-

tion on the codon reading pattern of tRNAs (review [28]). Yet, amber suppressor tRNAs have been purified from calf liver: they are Leu-tRNAs harbouring a 5′-CAA-3′ and a 5′-CAG-3′ anticodon, respectively. Mispairing (with regard to conventional rules) occurs at the second position of the anticodon in both species and also at the third position in the latter species [29]. Further discussion on the molecular mechanisms of non-cognate codon recognition can be found in a recent review [30].

A novel natural UAG suppressor tRNA has recently been isolated from mouse cells; it is a Gln-tRNA with a 5′-UmUG-3′ anticodon. Here, recognition involves mispairing in the third position of the anticodon. This tRNA which exists as a minor species in normal mouse cells, is markedly increased in Mo-MuLV infected cells [31]. This observation supports the hypothesis that Gln-$tRNA_{UmUG}$ is actually involved in the readthrough occurring during Mo-MuLV gene expression in vivo: indeed it was shown by protein analysis that a Gln residue is inserted in response to the UAG codon located at the end of the *gag* gene [32].

The ability of the natural suppressor tRNAs described above to recognize termination codons and the very likely involvement of some of them in regulation of viral gene expression are based on in vitro experiments. A clear-cut demonstration that unusual codon-anticodon pairing can also occur in vivo comes from recent results obtained in yeast by transformation of mutant strains with a multicopy plasmid carrying a normal tRNA gene. Overexpression of a Gln-tRNA with a 5′-UUG-3′ anticodon or of a Gln-tRNA with a 5′-CUG-3′ anticodon (the anticodon sequences being deduced from the corresponding tRNA genes) can partially complement ochre or amber mutations, respectively [33,34]. In both cases in vivo suppression involves a G:U mispairing in the third position of the anticodon, as it does in the case of the mouse Gln-$tRNA_{UmUG}$. It should be recalled that in ciliates Gln-$tRNA_{UmUG}$ is incapable of reading either UAG or UAA codons (see above). In conclusion, the nature of the suppressor tRNA may vary with the system studied. This indicates that different anticodons may combine with a termination codon, provided that the unorthodox interaction is properly stabilized, for example by the pattern of modified bases in the whole structure of the suppressor tRNA. The effect of the overall primary sequence and conformation of the tRNA on the stabilization of codon-anticodon interaction may be so strong that, as in the case of the 5′-UmUG-3′ anticodon, an anticodon deprived of suppressor activity in a given tRNA (Gln-$tRNA_{UmUG}$ of ciliates) could acquire it when included in another tRNA body (Gln-$tRNA_{UmUG}$ of mouse).

3.2.2. Nonsense suppression via specific recognition

The tRNAs described above are able at times to misread nonsense codons, while maintaining their original specificity and function in protein elongation (this latter point is demonstrated in the case of yeast [34] and is presumably also true in the other cases). Thus, they differ from authentic suppressor tRNAs that specifically read termination codons. Such nonsense suppressors, initially isolated from bovine liver, have now been discovered in human, avian and *Xenopus laevis* tissues (review [16]). They correspond to two minor Ser isoacceptors with anticodons 5′-CmCA-3′ and 5′-NCA-3′ (N is an unknown modification of U), the latter being complementary to the UGA termination codon. Both of them only read UGA codons, as demonstrated by ribosomal binding assays and by in vitro translation of β-globin mRNA. In addition, these Ser-tRNAs can be phosphorylated in the presence of a specific kinase, yielding P-Ser-tRNA. These functional peculiarities may be correlated to the unusual structural features of these tRNAs ([16] and references therein). Sequencing of the gene encoding the opal suppressor Ser-$tRNA_{NCA}$ from human, rabbit, chicken and *Xenopus* has confirmed that the genome of higher eukaryotes encodes authentic suppressor tRNAs.

The function of eukaryotic opal suppressors is still an open question. Direct involvement in protein synthesis is dependent on the presence of a UGA codon interrupting an 'open' reading frame. Stewart and Sharp [35] have discussed in the following terms the possibility that the opal suppressor tRNA may serve to insert P-Ser into proteins; examining the codon usage for P-Ser residues in a P-Ser-rich protein such as casein (by comparison of the cDNA and protein sequences), they found that all P-Ser are inserted in response to standard Ser codons. However, of the four

AGU codons used to specify P-Ser in bovine α_{S1} casein, two occur in the sequence AGUGA. Therefore, the authors suggest the interesting possibility that in these two cases the opal suppressor tRNAs would act as frameshifting tRNAs inserting P-Ser in response to an out-of-frame UGA codon overlapping a standard Ser codon [35]. This hypothesis would eliminate the uncomfortable necessity of a UGA codon interrupting an open reading frame, but it assumes that a second frameshift event has to take place downstream of the P-Ser site of insertion in order to recover the original reading frame. Two recent findings shed new light on potential roles of the opal suppressor tRNAs. Mammalian glutathione peroxidase [36,37] and *E. coli* formate dehydrogenase [38] both contain a UGA codon interrupting the ORF and encoding the Se-Cys found in the active site of both enzymes. Se-Cys is cotranslationally incorporated into proteins [38,39] and it has been shown that the Se-Cys backbone in the glutathione peroxidase arises from Ser [40]. A Se-Cys-specific tRNA has been identified in rat liver [41] but its primary sequence remains to be determined to verify its identity with the opal suppressor tRNAs described above. Meanwhile, one could test whether or not the opal suppressor tRNAs are able to direct Se-Cys insertion at least into mouse glutathione peroxidase. In *E. coli* a gene whose product is involved in the incorporation of selenium into selenoproteins has just been sequenced [42]. This gene encodes a Ser-tRNA with an anticodon complementary to the UGA termination codon and bears unusual structural features. This finding defines the Se-Cys incorporation pathway: the UGA termination codon is recognized by a natural suppressor tRNA aminoacylated with Ser; a specific enzyme most likely modifies the Ser residue bound to the tRNA into Se-Cys which would then be incorporated into proteins. In conclusion, the UGA codon generally specifies chain termination but it can also be recognized, within the same cell and in some mRNAs only, as encoding Se-Cys. This raises the intriguing question of how the cell distinguishes between these two functions (see section 4).

3.3. *Frameshift*

Skipping a termination codon may occur during elongation when a ribosome shifts from one reading frame to another (+1 or −1 frameshift) at a position in the mRNA upstream from that termination codon. Several natural frameshifts have been described in prokaryotes and eukaryotes and recently reviewed ([43], table 2). In eukaryotes, frameshift is required for the expression of retroviruses and related systems [43] and for the expression of a coronavirus [44]. In some cases as discussed below, the frameshifted products have been characterized by protein sequencing and the site of frameshift localized.

3.3.1. Retroviral systems

All retroviruses express their *pol* gene by synthesizing a *gag-pol* polyprotein that is subsequently cleaved to mature components by a virus-encoded protease. On the basis of their genetic organization and with regards to the expression of the *pol* (reverse transcriptase) and the *pro* (protease) genes, four classes may be distinguished among retroviruses. They are depicted in fig.1.

Representatives of the first class are Mo-MuLV and other murine leukemia and sarcoma viruses. The *gag* and *pol* genes are in the same reading frame and expression of the latter occurs via suppression of the *gag* gene termination codon. The region spanning the end of the *gag* gene and the beginning of the *pol* gene (encompassing the amber termination codon), encodes the viral protease responsible for cleavage of the *gag* polyprotein precursor into the mature core proteins [45]. The protease from Mo-MuLV [32] has been sequenced establishing the suppression event and revealing the presence of a Gln inserted in response to the nonsense codon (see section 3.2).

The prototype of the second class is RSV in which the *pol* ORF is −1 with respect to that of *gag*. The N-terminus of the mature reverse transcriptase maps downstream of the *gag* amber termination codon, itself at the end of the overlap between the two ORFs. The protease cleaving the *gag* precursor is encoded by the *gag* ORF and lies at the C-terminus of the *gag* polyprotein [46]. Despite early statements envisioning a splicing event that would piece together on the same reading frame the *gag* and the *pol* genes with removal of the termination codon [46], RSV is one of the first examples of frameshift in eukaryotes. Strong evidence in favour of a frameshift comes from experiments using in vitro transcripts ob-

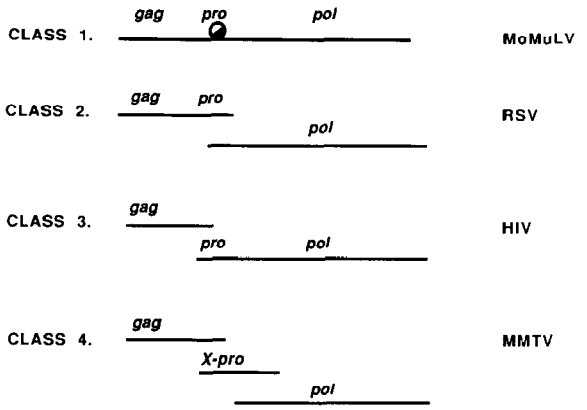

Fig.1. Schematic representation of the relative arrangements of *gag, pro* and *pol* genes in retroviruses. Four classes are distinguished on the basis of the type of expression of the protease and the reverse transcriptase genes. Regions encoding *gag, pro, 'X'* and *pol* are indicated. All overlapping reading frames are −1 compared to the upstream ORF. Coding regions are not drawn to scale, since ORF and overlaps vary in size among viruses of the same class; only the relative positions of each ORF are highlighted. The prototype virus of each class is indicated. ◕ = suppressible termination codon.

tained from cloned RSV cDNA [47]. These transcripts begin upstream of the *gag* initiation site and terminate within the *pol* reading frame far downstream from the *gag* termination codon. Upon in vitro translation in a reticulocyte lysate, the *gag-pol* RNA transcript produces the expected *gag* polyprotein and a larger polypeptide that can be immunoprecipitated with anti-reverse transcriptase and anti-*gag* serum. The size of the larger polypeptide is in agreement with that of a *gag-pol* fusion protein synthesized from the cloned RSV fragment indicating that these sequences can promote frameshift in vitro.

To the third class belong viruses such as the AIDS virus: HIV-1 [48–50] and Visna virus [51] in which the protease is encoded by the 5′-end of the *pol* ORF which is out of frame with the *gag* ORF. Therefore, it is likely that as in RSV, a single frameshift event is required for the synthesis of the *gag-pol* precursor. Indeed, it has been very recently demonstrated that a −1 frameshift occurs in HIV-1 *gag-pol* expression by in vitro translation of constructs carrying the presumptive frameshifting region. The amino acid sequence of the fusion protein derived from these constructs has made it possible to localize the site of frameshift ([52] and see section 4.2.1.).

Finally, in the fourth class fall viruses in which separate reading frames are used for the synthesis of *gag, pro* and *pol*. For example, sequence analysis of the MMTV genome [53,54] has revealed three overlapping ORFs in the 5′-part of the genome. The presence in the derived amino acid sequence of protease and reverse transcriptase domains common to other retroviruses suggests assignment of each ORF to *gag, X-pro* (*'X'* is a protein domain of unknown function) and *pol*, respectively. Three nested *gag*-related polyproteins can be detected in vitro and in vivo that are consistent in size with the stepwise extended translation of the *gag* ORF. Similar experiments to those designed for RSV, i.e. transcription and translation of variously truncated cDNA fragments encompassing one or two overlaps, produce the expected *gag-X-pro* or *gag-X-pro-pol* fused polyproteins. Clearly, the presumed single and double −1 frameshifts have taken place in vitro [54]. More recently, complete amino acid sequencing of a minor p30 protein found in MMTV virions has shown that this protein is encoded by the *gag-X-pro* overlap. Alignment of the amino acid sequence of this viral protein with the nucleotide sequence of the MMTV genome, provides a paramount demonstration of the frameshift event in vivo and highlights the site of frameshift [55]. It is interesting to note that in this class of viruses the region where frameshift occurs corresponds to a protein domain (p30 = *'X'*) different from that of the protease.

A similar genetic organization to MMTV is shared by BLV [56,57] where the protease has been sequenced and shown to be fully encoded by the *X-pro* ORF [58]. Direct amino acid analysis of the *gag-X-pro* fusion protein at the junction between the two reading frames should definitely prove the occurrence of frameshifting. Analyses of their genome sequence suggest that HTLV I [59] and II [60] and MPMV, a type D retrovirus [61] also belong to this fourth class of retroviruses.

Accumulation of sequences might lead to the creation of more than four classes based on the strategies for *pro* and *pol* gene expression. Defining a class requires, in addition to the nucleotide sequence, the complete amino acid sequence of the protein bridging two out of frame ORFs. In this regard, the case of FeLV is instructive. The nucleotide sequence of one strain (subgroup B)

suggests a strategy of expression of the third class [62] whereas the partial sequence of the protease of another strain (strain AB) seems to indicate a first class type of gene expression [63]. This discrepancy may of course rely on the fact that the sequenced cDNA clone of FeLV (subgroup B) could correspond to a defective RNA [63]. However, if one assumes that the nucleotide sequence of strain AB is identical to that of subgroup B, one should envisage the combination of a readthrough and a frameshift event to account for FeLV protease synthesis, and thus the creation of a fifth class of gene expression among retroviruses.

3.3.2. Retroviral-related systems

Eukaryotic transposons show structural and functional similarities to retroviral proviruses [64]. Recently, it has been shown that yeast Ty transposons (Ty 1–15, Ty 1–17 and Ty 912) also share with retroviruses a common mechanism of gene expression: frameshift appears to be required for the expression of the *tyb* ORF through the synthesis of a *tya-tyb* fusion protein [65–67]. The *tyb* ORF overlaps the *tya* ORF by 38 nt in a +1 reading frame. Protein *tya* exhibits structural features characteristic of a DNA-binding protein and protein *tyb* presents homologies to several reverse transcriptases. Careful S_1 mapping analyses, able to detect splicing of fragments as short as 5 bp, have definitely excluded the possibility that RNA processing is responsible for the synthesis of the *tya-tyb* fusion protein. On the contrary, *tya-tyb*-α_2 interferon [65,67] or *tya-tyb*-lac z [65] fusion proteins are expressed exclusively when the reporter genes are in frame with the *tyb* ORF, leaving frameshift as the only possible mechanism of expression. Deletion studies have brought down to 31 nt the region of overlap in which frameshifting occurs, and within these 31 nt there is an 11 nt block that is completely conserved between different classes of transposons [67]. The translational strategy used in transposons for the *tya-tyb* fusion protein is remarkably similar to that of the retroviral *gag-pol* polyprotein. Since the *tya-tyb* fusion protein synthesized in vivo is accompanied by a shorter product, presumably a cleavage product [65,66], it is reasonable to hypothesize that the *tya-tyb* moiety also encodes a protease as in retroviruses.

On the basis of sequence homology with retroviral polymerases, other candidates for gene expression via frameshifting are the 17.6 *copia*-like transposable element of *D. melanogaster* [68], hepadnaviruses and CaMV where the ORF encoding the putative reverse transcriptase overlaps the ORF encoding antigen or coat proteins, respectively (review [69]). This genetic organization is once again reminiscent of the *gag-pol* ORFs of retroviruses. However, attempts to detect a CaMV fusion protein resulting from the suspected frameshift have so far failed. Recent observations indicate that an insertion in the CaMV genome that destroys the overlap between the coat protein and polymerase genes is viable; this suggests that the fusion of the translation products of these two genes is not obligatory [70].

Another non-retroviral-like case should be mentioned: the −1 frameshift occurring in a coronavirus, IBV, for the synthesis of the F1-F2 fusion protein, presumably encoding the putative viral polymerase [44].

3.3.3. Efficiency and possible mechanisms of frameshift

In retroviral systems frameshift occurs with relatively high efficiencies ranging from 5 to 30%. For example, the *gag-pol* fusion protein in RSV is synthesized at 5% efficiency in vitro and in infected cells. The *gag-X-pro* and *X-pro-pol* fusion proteins of MMTV are synthesized in vitro at 23 and 8% efficiency respectively, and these figures are consistent with the levels observed in vivo. Frameshift efficiency in vitro is 11% in the case of HIV-1 and it is estimated to be as high as 25–30% for the F1-F2 fusion protein of IBV.

Although the site of frameshifting has been localized in at least two cases [52,55] the exact mechanism of frameshifting is still unclear. In principle, frameshifting could involve decoding of a triplet in another reading frame (implying skipping of one or two but not three nt), non-triplet decoding, non-triplet translocation or some other translational quirk [43,71,72]. The data reported above are in favour of another model: a 'slippage model' [52,55]. The frameshifting tRNA would initially bind to its cognate codon in a standard triplet interaction at the ribosomal A site. Before the peptide bond is formed, a −1 slippage would bring the aminoacyl-tRNA to mispair with the triplet in the −1 reading frame, generating the

frameshift. Such a model is based on the assumption that the interaction between the peptidyl-tRNA and its cognate codon is in fact wobbly, so that the third base of the codon at the P site can be read twice, once by its cognate tRNA and once by the frameshifting tRNA. Features that may promote this series of unusual events are discussed below (see section 4.2.1), at the level of both the primary and secondary structure of the mRNA.

Contrary to the work on nonsense suppressor tRNAs, isolation of frameshifting tRNAs has not been conducted in eukaryotic systems. Hopefully, the recent availability of natural mRNAs requiring a frameshift event for gene expression will help in the detection of such frameshifting tRNA species.

Finally, in addition to frameshifting, another mechanism of skipping termination codons should be mentioned, although it has so far only been demonstrated in prokaryotes: ribosomal hops of up to 6 nt have been shown to occur on overlapping or non-overlapping homologous codons surrounding a termination codon [73]. More intriguing is the fact that in gene 60 of the *E. coli* bacteriophage T4, a 50 nt long intracistronic region that contains 6 termination codons is apparently skipped upon translation of the corresponding mRNA [74].

4. UNDERSTANDING REGULATION

The scenarios described above, which all directly or indirectly involve a termination codon, are means of regulating gene expression at the translational level. In the following section an understanding of the meanings of such regulations is proposed and the parameters that influence these regulations and that are ultimately responsible for the modulation of the different happenings at the termination codons are discussed.

4.1. *The rationales of regulation*

Translating a termination codon as a sense codon or skipping it by frameshift results in the synthesis of elongated proteins. These translational strategies appear to be designed for the synthesis, from one mRNA species and one single

Table 1

List of animal and plant viral mRNAs with a suppressible termination codon

Virus	Readthrough		Nucleotide sequence	Termination codon	Size of proteins	Nature of proteins	Amino acid sequence	Ref.
	In vitro	In vivo						
Molony murine leukemia virus (Mo-MuLV)	+	+	+	UAG	$Pr76^{gag}/Pr180^{gag\text{-}pol}$	s/ns	+	32,45
Rauscher murine leukemia virus (Ra-MuLV)	+	+	ND	UAG	$Pr80^{gag}/Pr200^{gag\text{-}pol}$	s/ns	–	75
Sindbis virus	+	+	+	UGA	p230/p270	ns/ns	–	76,77
Middleburg virus	+	ND	+	UGA	p230/p270	ns/ns	–	76
Tobacco mosaic virus (TMV)	+	+	+	UAG	126 kDa/183 kDa	ns/ns	–	18–23
Carnation mottle virus (CarMV)	+	+	+	UAG	p30/p77/p100	ns/ns	–	78,79
Beet necrotic yellow vein virus (BNYVV)	+	ND	+	UAG	21 kDa/75 kDa	s/ns	–	80,81
Turnip yellow mosaic virus (TYMV)	+	ND	+	UAG	206 kDa/221 kDa	ns/ns	–	82
Tobacco rattle virus (TRV)	+	+	+	UGA	140 kDa/170 kDa	ns/ns	–	83
Lucerne transient streak virus (LTSV)	+	ND	ND	ND	p78/p100	ns/ns	–	84
Soil-borne wheat mosaic virus (SBWMV)	+	ND	ND	ND	19.7 kDa/28 kDa/ 90 kDa	s/ns/ns	–	85

ND, not determined; ns, nonstructural protein; s, structural protein

initiation event, of two (or more) related proteins differing in their C-terminal region and thus capable of different functions. Moreover, the two or more proteins are produced in different amounts depending on the efficiency of suppression or frameshift. These strategies are essentially encountered in animal and plant viruses. This is not surprising because, due to the relatively small size of their genome, viruses possess an extremely compact genetic organization. Indeed, expression of proteins via nonsense suppression and frameshift fulfills two requirements: maximum exploitation of nucleotide sequence and regulation of protein synthesis. For example, a retrovirus needs both the *gag* and *pol* proteins, but whereas the former is required in large amounts, as the precursor of the structural core proteins, the latter is only required in catalytic amounts; thus, a regulated suppression or frameshift mechanism ensures the production of an adequate ratio between the two precursor proteins. Tables 1 and 2 list the cases of regulation through nonsense suppression and frameshifting discussed in this paper; the termination codon involved and the nature of the proteins concerned are indicated as well as whether the nucleotide and amino acid sequences have been determined.

As discussed above, another rationale of the regulated recognition of termination codons is the possible insertion at specific sites in the polypeptide chain of modified amino acids such as Se-Cys by natural opal suppressor tRNAs (see section 3.2.2). It has recently been speculated, at a time when the exact nature of the Se-Cys inserting tRNA of *E. coli* was not known, that such an insertion could reflect an adaptation to changes in cellular physiology [87].

4.2. *The determinants of regulation*

The question to be addressed now is how 'natural errors' such as nonsense suppression and frameshifting can become regulatory devices. Two aspects have to be considered. On the one hand, the site of regulation: certain UAG codons are suppressible, others are not; certain UGA codons dictate the insertion of a modified amino acid, others do not; some codons are frameshift-prone, others are not. On the other hand, the efficiency of

Table 2

List of viruses and retroviral-related systems with frameshifting sites

System	Sequence		Test in vitro	Proteins concerned	Frameshift type	Ref.
	Nucleotide	Amino acid				
Rous sarcoma virus (RSV)	+	–	+	*gag-pro/pol*	–1	46,47
AIDS virus (HIV)	+	+	+	*gag/pro-pol*	–1	48–50,52
Visna virus	+	–	–	*gag/pro-pol*	–1	51
Mouse mammary tumor virus (MMTV)	+	+	+	*gag/X-pro*	–1	53–55
	+	–	+	*X-pro/pol*	–1	
Bovine leukemia virus (BLV)	+	+	–	*gag/X-pro*	–1	56–58
	+	–	–	*X-pro/pol*	–1	
Human T-cell leukemia virus (HTLVI)	+	–	–	*gag/X-pro/pol*	–1	59
Human T-cell leukemia virus (HTLVII)	+	–	–	*gag/X-pro/pol*	–1	60
Mason Pfizer monkey virus (MPMV)	+	–	–	*gag/X-pro/pol*	–1	61
Feline leukemia virus (FeLV)	+	+	–	*gag/pol*	[a]	62,63
Ty transposon (1–15)	+	–	+	*tya/tyb*	+1	65,67
Ty transposon (1–17)	+	–	+	*tya/tyb*	+1	67
Ty transposon (912)	+	–	+	*tya/tyb*	+1	66
17.6 copia-like element	+	–	–	ORF1/ORF2	+1	68
Avian infectious bronchitis virus (IBV)	+	–	+	F1/F2	–1	44
Hepadnaviruses	+	–	–	*'gag'/'pol'*	–1	69
Cauliflower mosaic virus (CaMV)	+	–	–	ORFV/ORFVI	–1	69
Alfalfa mosaic virus (AlMV)	+	–	+	35 kDa/54 kDa	[a]	86

[a] The mechanism remains to be clarified

regulation: which parameters modulate the level of frameshift or nonsense suppression to obtain the required amount of frameshift product. However, the aim here is not to discern among the factors those determining the specificity of regulation from those modulating the efficiency of regulation but rather to show that these two aspects are interrelated.

4.2.1. Codon context: primary sequences and secondary structures

To explain why some termination codons are suppressible and others not, and also to explain the variable efficiencies of suppression of different termination codons, a major parameter that is generally invoked is 'codon context'. This concept originated from studies on nonsense and missense suppression in prokaryotes (i.e. suppression of mutations leading to the appearance of a termination codon in an ORF or to the replacement of a 'correct' codon by an 'incorrect' one). It appears that translation of a given codon is influenced by the surrounding nucleotides. In particular, the presence of a purine 3′ of an amber codon stimulates nonsense suppression in vivo [88–90], suggesting that signals for termination of translation could include sequences downstream of the nonsense codon. This hypothesis is supported by the analysis of the sequences surrounding natural termination codons. The compilation by Köhli and Grosjean [12] of the termination sequences of many eukaryotic and prokaryotic mRNAs reveals a significant bias in the usage of bases at the 3′-side of termination codons. All tight termination codons in prokaryotes are followed by a U, and all leaky UGA codons are followed by an A [91]. Unfortunately, in eukaryotes no extensive comparison of the contexts of leaky with non-leaky termination codons that includes the cases of viral regulations examined to date has been made. One further observation may still give weight to the hypothesis of codon context in eukaryotes if not exactly to the conclusion drawn by the analysis of the mRNA sequences mentioned above. In three plant RNA viruses (TMV, BNYVV and TYMV) using natural suppression as a translational strategy, the UAG codon is flanked on either side by a CAA (Gln) codon [19,81,92]. This may be more than mere coincidence.

If contexts favourable for suppression (and thus unfavourable for termination) do exist, one may also envisage extra signals that could protect cellular termination codons from recognition by endogenous suppressor tRNAs. In this regard, the compilation of termination sequences in eukaryotic cellular mRNAs is deceiving: the frequency of two termination codons in tandem is very low, 60% of the mRNAs having a short ORF of more than 10 triplets downstream of their terminator [12]. However, it still remains that the injection of yeast genetic suppressor tRNAs into *Xenopus* oocytes has no dramatic consequences, leading only to a very limited number of changes (about 10%) in the two-dimensional pattern of the oocyte proteins [93]. Similarly, mammalian cell lines have been obtained that constitutively express low levels of genetically engineered suppressor tRNAs indicating that cell viability is not affected by the presence of tRNAs recognizing termination codons ([94] and references therein). These observations all agree with the existence of protective codon contexts in cellular mRNAs, although the identity of such contexts remains undetermined. With the accumulation of viral and non-viral eukaryotic mRNA sequences, it may be possible to discern some pattern among translation termination sequences from which a model for recognition of termination of protein synthesis by RF could emerge. A termination signal might be composed of two domains: a relatively constant region, the nonsense codon, and a variable region, the downstream sequence affecting efficiency of termination [90]. To this speculation is probably related the fact that so far only amber and opal codons have been found to be naturally suppressible. As suggested by Geller and Rich [95], termination codons may well not be synonymous but the translational machinery may distinguish UAA as absolute termination signal from UAG and UGA as facultative termination codons.

Codon context is not limited to the vicinity of a termination codon. Rather, it can be enlarged to more distant environmental sequences as observed in the case of frameshift. Inspection of the various shifty overlaps indicates that two types of sequences, located upstream of the termination codon of the ORF of reference, can be found in the different systems where frameshift is operational [54]. The first conserved sequence is U.UUA block (where the triplet is located in the ORF of

reference) which is found at the *gag/pol* junction in RSV, at the *gag/pro* junction in HIV-1, at the *X-pro/pol* junction of MMTV, BLV, HTLV I and II and at the F1/F2 junction of IBV. Interestingly, the conserved U.UUA block is also found in the intergenic region of RNA3 of AlMV where frameshift has been hypothesized to explain the synthesis of an elongated viral translation product observed in vitro and in vivo [86]. The other conserved block, A.AAA.AAC, is found at the *gag/X-pro* overlaps of MMTV, BLV and HTLV I and II (see table 2 for references). These homopolymeric sequences may favour slippage of ribosomes: frameshifting might occur within such sequences or downstream. In the case of HIV-1 it has been demonstrated that the site of frameshifting lies within the conserved U.UUA block. In MMTV the site is still debatable: it lies either at the end of the conserved A.AAA.AAC block or just beyond. The two conserved sequences seem to be so far specific for -1 frameshift since they are not found in the *tya-tyb* overlap of Ty transposons. Other conserved sequences [67] might specify $+1$ frameshifts.

Constructs containing the homopolymeric sequences alone do not induce frameshift, as suggested by experiments in which the short putative frameshift window was inserted upstream of an out-of-frame reporter gene [44,52,54]. One cannot exclude the possibility that downstream sequences of the reporter gene exert a negative effect on frameshifting. A simpler explanation is that all the information required for frameshifting does not merely reside in the short conserved stretches of nucleotides but rather that additional sequences and/or structures bordering the overlap are also involved in promoting efficient frameshift.

The search for a potential secondary structure of the mRNA in the vicinity of the frameshift site has been rewarding. Convincing stem and loop structures can be drawn downstream of the *gag/pol* junction of HIV-1 [52] and of the F1/F2 junction of IBV [44]. They are also present downstream of the *gag/X-pro* junction of BLV, MMTV, HTLV I and II as well as downstream of the *X-pro/pol* junction of MMTV [54]. The size of the loops ranges from 18 nt (MMTV *gag/X-pro*) to 81 nt (MMTV *X-pro/pol*) and in all cases, except IBV, the termination codon of the first ORF lies upstream of the hairpin. These predicted secondary structures may provide important contributions to an efficient frameshift by stalling ribosomes along the mRNA [41,51] and possibly masking the termination codon from the RF thereby increasing the chance for a frameshift event.

4.2.2. A kinetic view of regulation

The determinants of regulation of termination discussed so far are contained within nucleotide sequence elements. This is not sufficient to give insights into the molecular interactions required for regulation. A comprehensive picture should consider protein synthesis as the outcome of competing interactions involving the aminoacyl-tRNA pool, the ribosomal binding sites and the translational factors. Such a dynamic view provides possible interpretations of the effects of the regulatory sequences presented above and leads to the introduction of new parameters of regulation, namely competition between tRNAs or between tRNA and RF, tRNA availability in the cell and numerous other metabolic factors and physiological conditions that may affect such competitions.

The kinetic theory of accuracy elaborated by Hopfield [96] and by Ninio [97] has been used by these authors to describe among other mechanisms the elongation step of protein synthesis in prokaryotes (reviews [98,99]). Its predictions have been confirmed in many cases by experimental observations (recent reviews [100,101]). The main features of the theory and its consequences can apply to termination in both prokaryotes and eukaryotes. Although all the experimental support mainly comes from prokaryotic systems, we shall transpose here the theory to the eukaryotic situation, assuming that the mechanisms of elongation and termination are, if not identical, at least very similar in both systems.

Let us consider a ribosome translating a given codon in an mRNA. In principle, all the aminoacyl-tRNAs present in the cell in the form of ternary complexes with elongation factor 1 (EF1: the counterpart of the prokaryotic EF-Tu factor) and GTP are candidates for interacting with the codon at the ribosomal A site. The minimal mechanism of elongation involves two steps. In the first one, called 'substrate selection' [100] or 'initial recognition' [101], one of the ternary com-

plexes binds to the ribosomal A site and either dissociates or undergoes GTP hydrolysis. In the second step that follows GTP hydrolysis and is called 'proofreading', the aminoacyl-tRNA either is incorporated into the nascent protein chain or falls off the ribosome. Peptide bond formation requires that the EF1-GDP complex has left the ribosome [102]. Thus, three situations are encountered at a given codon: (i) ternary complexes that weakly interact with the ribosome dissociate before GTP hydrolysis; (ii) ternary complexes that stick better to the ribosome, pass the GTP hydrolysis screen but may still dissociate from the ribosome faster than does the EF1-GDP complex; (iii) finally, ternary complexes that stay on the ribosome long enough for GTP hydrolysis and dissociation of the EF1-GDP complex to occur, thus allowing peptide bond formation. In a highly accurate system, all non-cognate ternary complexes are rejected at either the first or second step and only cognate ternary complexes lead to peptide bond formation.

This model of accuracy provides a good basis for understanding the parameters of inaccuracy and in particular the regulation of the suppression of termination codons. Indeed, suppression can be regarded as the competition for a nonsense codon between the RF and all the ternary complexes among which one is acting as a suppressor. It follows first that crucial determinants of regulation of suppression are the availability and relative amount of tRNA species, natural suppressor tRNAs and RF. Any imbalance in the aminoacyl-tRNA population may favour misreading and nonsense suppression. This is illustrated in the case of Mo-MuLV (see section 3.2.1). One may also suggest that the imbalance concerns the concentration of RF by some sequestration mechanism occurring upon viral infection. Similarly, compartmentalization of natural suppressor tRNAs may control their utilization for specific purposes.

Returning to the minimal model of peptide chain elongation, it is also clear that the efficiency of suppression is determined by the value and possible variations of the 'sticking time' [99] of the suppressor ternary complex to the ribosome relative to the average time of GTP hydrolysis and when the second step is allowed to proceed relative to the average time required for the dissociation of the EF1-GDP complex from the ribosome. Any factor increasing the sticking time of a natural suppressor tRNA on its 'non-cognate' termination codon will increase its probability of participating in peptide chain elongation. Undermodification of bases at the level of the anticodon (see section 3.2.1) as well as codon context are examples of such factors. At least two hypotheses have been proposed to explain the involvement of the bases adjacent to a termination codon in increasing suppression efficiency: the 'swollen codon' hypothesis and the 'tRNA-tRNA interaction' hypothesis [80,103]. The swollen codon hypothesis suggests a 4 bp or even a 5 bp codon-anticodon interaction; however, this hypothesis appears rather unlikely when considering the conformation of the tRNA loop at the ribosomal A site [104]. The second hypothesis suggests that tRNA-tRNA interaction can occur between the peptidyl-tRNA at the P site and the acylated suppressor tRNA at the A site. If favourable, this interaction would counterbalance the weak non-cognate interaction of the natural suppressor tRNA with the termination codon, thereby increasing the sticking time of this tRNA and allowing peptide bond formation. In this respect, the tertiary structure of the suppressor tRNA and its pattern of base modifications, especially outside of the anticodon, play a decisive role.

One of the most interesting predictions of the kinetic theory is that factors which affect the rate of GTP hydrolysis or of dissociation of the EF1-GDP complex from the ribosome in a non-specific manner (i.e. independently of the nature of the ternary complex bound to the ribosome) may have a specific effect on the probability of a given ternary complex participating in peptide chain elongation. Indeed, for a given system and in a given environment, the average time of GTP hydrolysis and of EF1-GDP complex dissociation from the ribosome is very likely constant [101]. Physiological or metabolic variations in the cell may affect these constants. Consequently, the ratios of the sticking times of all ternary complexes to these modified constants are changed. For some ternary complexes this may transform the interaction with the ribosome [99,101]. For example, let us imagine a metabolic variation leading to more rapid GTP hydrolysis on the ribosome (modification of a GTPase activity, variations in the GTP or

GDP concentrations...). The average time of GTP hydrolysis is thus shortened and becomes closer to the sticking time of some weakly interacting ternary complexes. Such complexes that would have otherwise been rejected can now pass the 'initial recognition' step. Similarly, changes that accelerate the dissociation of the EF1-GDP complex from the ribosome (mutation or alteration in EF1 or in the ribosome...) will increase the probability of incorporating non-cognate amino acids into the peptide chain. Thus, both GTP metabolism and EF1 recycling appear to play a key role in accuracy and therefore in suppression [98–101,105,106].

A great variety of other factors have been described that influence the efficiency of suppression. High Mg^{2+} concentrations as well as polyamines stimulate the suppression of UAG and of UGA codons [18,107,108]. Readthrough is also more efficient at low than at high temperatures [109]. However, these effects are more difficult to interpret precisely, since ions and temperature likely affect both non-specific kinetic parameters (GTP hydrolysis and dissociation of the EF1-GDP complex) and the respective sticking times of each ternary complex and RF entering the competition.

Not only suppression of termination codons but also frameshift can be explained by the kinetic theory. Recently, a kinetic model of ribosome translocation has been proposed [110] and the increasing number of natural frameshifts observed in eukaryotic mRNAs should provide a good experimental basis for developing this model.

5. CONCLUSION

The study of termination leads to the shattering of some of the early, established rules of protein biosynthesis. The universality of the genetic code is denied by the finding that nonsense codons in one organism are sense codons in another organism. A nonsense codon can no longer be defined as a codon not recognized by any tRNA, since naturally occurring tRNAs exist that possess an anticodon complementary to a termination codon. The dogma of the non-ambiguity of the genetic code is also broken: a codon may have two or more meanings. The UAG codon of TMV, for instance, means either termination, Tyr when translated in the presence of tobacco tRNAs, Leu in the presence of calf liver tRNAs, or Gln in the presence of mouse liver tRNAs. Rules for codon-anticodon interactions are upset in nonsense suppression where unorthodox G:G or A:A pairings are involved. The translated reading frame is no longer reliably determined by the initiation codon, since frameshifts may occur to skip undesirable termination codons and synthesize complete translation products.

The translational regulations of gene expression discussed in this paper take advantage of and amplify defects or rather properties inherent to the translation process and apparatus. Indeed, since the early genetic studies on prokaryotic translation, it has become apparent that protein biosynthesis is the outcome of elementary processes of low fidelity and that evolution has developed numerous devices for achieving high accuracy in gene expression [111]. Hence, frameshifting and nonsense suppression were first observed in the selection of phenotypic reversion of frameshift and nonsense mutations and they presumably occur at low levels in all cellular translations. Of course if these inherent 'errors' of translation were above background level, accuracy of translation would be definitely at stake in any living organism. It seems rather that evolution has optimized the efficiency of these 'errors' at specific sites, turning them into biologically significant sites of regulation.

Acknowledgements: We wish to thank Anne-Lise Haenni for encouragement and critical reading of the manuscript, and Jacques Ninio for stimulating discussions. We are grateful to François Chapeville for his constant interest in our work. R.V. acknowledges the 'Société de Secours des Amis des Sciences, Paris' for a fellowship. This work was partly supported by a grant from the 'Ligue Nationale Française contre le Cancer'.

REFERENCES

[1] Barrell, B.G., Bankier, A.T. and Drouin, J. (1979) Nature 282, 189–194.
[2] Attardi, G. (1985) Int. Rev. Cytol. 93, 93–145.
[3] Horowitz, S. and Gorovsky, M.A. (1985) Proc. Natl. Acad. Sci. USA 82, 2452–2455.
[4] Helftenbein, E. (1985) Nucleic Acids Res. 13, 415–433.
[5] Caron, F. and Meyer, E. (1985) Nature 314, 185–188.
[6] Preer, J.R. jr, Preer, L.B., Rudman, B.M. and Barnett, A.J. (1985) Nature 314, 188–190.
[7] Kuchino, Y., Hanyu, N., Tashiro, F. and Nishimura, S. (1985) Proc. Natl. Acad. Sci. USA 82, 4758–4762.
[8] Hanyu, N., Kuchino, Y., Nishimura, S. and Beier, H. (1986) EMBO J. 5, 1307–1311.

[9] Steege, D.A. and Söll, D.G. (1979) in: Biological Regulation and Development (Goldberger, R.F. ed.) Gene Expression, vol.1, pp.433–485, Plenum, New York.
[10] Caskey, C.T. (1980) Trends Biochem. Sci. 5, 234–237.
[11] Beaudet, A.L. and Caskey, C.T. (1970) Nature 227, 38–40.
[12] Köhli, J. and Grosjean, H. (1981) Mol. Gen. Genet. 182, 430–439.
[13] Murgola, E.J., Hijazi, K.A., Göringer, H.U. and Dahlberg, A.E. (1988) Proc. Natl. Acad. Sci. USA, in press.
[14] Celis, D.G. and Smith, J.D. (1979) Nonsense Mutations and tRNA Suppressors, Academic Press, New York.
[15] Söll, D.G., Abelson, J. and Schimmel, P.R. (1980) Transfer RNA: Biological Aspects, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
[16] Hatfield, D.L. (1985) Trends Biochem. Sci. 10, 201–204.
[17] Morch, M.D., Valle, R.P.C. and Haenni, A.L. (1987) in: The Molecular Basis of Viral Replication (Perez Bercoff, R. ed.) pp.113–151, Plenum, New York.
[18] Pelham, H.R.B. (1978) Nature 272, 469–471.
[19] Goelet, P., Lomonossoff, G.P., Butler, P.J.G., Akam, M.E., Gait, M.J. and Karn, J. (1982) Proc. Natl. Acad. Sci. USA 79, 5818–5822.
[20] Beier, H., Mundry, K.W. and Issinger, O.-G. (1980) Intervirology 14, 292–299.
[21] Beier, H., Barciszewska, M., Krupp, G., Mitnacht, R. and Gross, H.J. (1984) EMBO J. 3, 351–356.
[22] Bienz, M. and Kübli, E. (1981) Nature 294, 188–190.
[23] Beier, H., Barciszewska, M. and Sickinger, H.-D. (1984) EMBO J. 3, 1091–1096.
[24] Barciszewski, J., Barciszewska, M., Suter, B. and Kübli, E. (1985) Plant Sci. 40, 193–196.
[25] Yokohama, S., Miyazawa, T., Iitaka, Y., Yamaizumi, Z., Kasai, H. and Nishimura, S. (1979) Nature 282, 107–109.
[26] Hunter, W.N., Brown, T., Anand, N.N. and Kennard, O. (1986) Nature 320, 552–555.
[27] Privé, G.G., Heinemann, H., Chandrasegaran, S., Kan, L.S., Kopka, M.L. and Dickerson, R.E. (1987) Science 238, 498–504.
[28] Björk, G.R., Ericson, J.U., Gustafsson, C.E.D., Hagervall, T.G., Jönsson, Y.H. and Wikström, P.M. (1987) Annu. Rev. Biochem. 56, 263–287.
[29] Valle, R.P.C., Morch, M.-D. and Haenni, A.-L. (1987) EMBO J. 6, 3049–3055.
[30] Buckingham, R.H. and Grosjean, H. (1987) in: Accuracy in Molecular Processes: its Control and Relevance to Living Systems (Kirkwood, T.B.L. et al. eds) pp.83–126, Chapman and Hall, London.
[31] Kuchino, Y., Beier, H., Akita, N. and Nishimura, S. (1987) Proc. Natl. Acad. Sci. USA 84, 2668–2672.
[32] Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) Proc. Natl. Acad. Sci. USA 82, 1618–1622.
[33] Pure, G.A., Robinson, G.W., Naumovski, L. and Friedberg, E.C. (1985) J. Mol. Biol. 183, 31–42.
[34] Weiss, W.A. and Friedberg, E.C. (1986) J. Mol. Biol. 192, 725–735.
[35] Stewart, T.S. and Sharp, S. (1984) Methods Enzymol. 106, 157–161.
[36] Chambers, I., Frampton, J., Godfarb, P., Affara, N., McBain, W. and Harrison, P.R. (1986) EMBO J. 5, 1221–1227.
[37] Ishida, K., Morino, T., Takagi, K. and Sukenaga, Y. (1987) Nucleic Acids Res. 15, 10051.
[38] Zinoni, F., Birkmann, A., Leinfelder, W. and Böck, A. (1987) Proc. Natl. Acad. Sci. USA 84, 3156–3160.
[39] Hawkes, W.C. and Tappel, A.L. (1983) Biochim. Biophys. Acta 739, 225–234.
[40] Sunde, R.A. and Evenson, J.K. (1987) J. Biol. Chem. 262, 933–937.
[41] Hawkes, W.C., Lyons, D.E. and Tappel, A.L. (1982) Biochim. Biophys. Acta 699, 183–191.
[42] Leinfelder, W., Zehelin, E., Mandrand-Berthelot, M.-A. and Böck,A. (1988) Nature 331, 723–725.
[43] Craigen, W.J. and Caskey, C.T. (1987) Cell 50, 1–2.
[44] Brierly, I., Boursnell, M.E.G., Binns, M.M., Bilimoria, B., Blok, V.C., Brown, T.D.K. and Inglis, S.C. (1987) EMBO J. 6, 3779–3785.
[45] Shinnick, T.M., Lerner, R.A. and Sutcliffe, J.G. (1981) Nature 293, 543–548.
[46] Schwartz, D.E., Tizard, R. and Gilbert, W. (1983) Cell 32, 853–869.
[47] Jacks, T. and Varmus, H.E. (1985) Science 230, 1237–1242.
[48] Wain-Hobson, S., Sonigo, P., Danos, O., Cole, S. and Alizon, M. (1985) Cell 40, 9–17.
[49] Ratner, L., Haseltine, W., Patarca, R., Livak, K.J., Stacich, B., Josephs, S.F., Doran, E.R., Rafalski, J.A., Whitehorn, E.A., Baumeister, K., Ivanoff, L., Petteway, S.R. jr, Pearson, M.L., Lautenberger, J.A., Papas, T.S., Ghrayeb, J., Chang, N.T., Gallo, R.C. and Wong-Staal, F. (1985) Nature 313, 277–284.
[50] Sanchez-Pescador, R., Power, M.D., Barr, P.J., Steimer, K.S., Stempien, M.M., Brown-Shimer, S.L., Gee, W.W., Renard, A., Randolph, A., Levy, J.A., Dina, D. and Luciw, P.A. (1985) Science 227, 484–492.
[51] Sonigo, P., Alizon, M., Staskus, K., Klatzmann, D., Cole, S., Danos, O., Retzel, E., Tiollais, P., Haase, A. and Wain-Hobson, S. (1986) Cell 42, 369–382.
[52] Jacks, T., Power, M.D., Masiarz, F.R., Luciw, P.A., Barr, P.J. and Varmus, H.E. (1988) Nature 331, 280–283.
[53] Moore, R., Dixon, M., Smith, R., Peters, G. and Dickson, G. (1987) J. Virol. 61, 480–490.
[54] Jacks, T., Townsley, K., Varmus, H.E. and Majors, J. (1987) Proc. Natl. Acad. Sci. USA 84, 4298–4302.
[55] Hizi, A., Henderson, L.E., Copeland, T.D., Sowder, R.C., Hixson, C.V. and Oroszlan, S. (1987) Proc. Natl. Acad. Sci. USA 84, 7041–7045.
[56] Rice, N.R., Stephens, R.M., Burny, A. and Gilden, R.V. (1985) Virology 142, 357–377.
[57] Sagata, N., Yasunaga, T., Tsuzuku-Kawamura, J., Ohishi, K., Ogawa, Y. and Ikawa, Y. (1985) Proc. Natl. Acad. Sci. USA 82, 677–681.
[58] Yoshinaka, Y., Katoh, I., Copeland, T.D., Smythers, G.W. and Oroszlan, S. (1986) J. Virol. 57, 826–832.
[59] Seiki, M., Hattori, S., Hirayama, Y. and Yoshida, M. (1983) Proc. Natl. Acad. Sci. USA 80, 3618–3622.

[60] Shimotohno, K., Takahashi, Y., Shimizu, N., Gojobori, T., Golde, D.W., Chen, I.S.Y., Miwa, M. and Sugimura, T. (1985) Proc. Natl. Acad. Sci. USA 82, 3101–3105.
[61] Sonigo, P., Barker, C., Hunter, E. and Wain-Hobson, S. (1986) Cell 45, 375–385.
[62] Laprevotte, I., Hampe, A., Sherr, C.J. and Galibert, F. (1984) J. Virol. 50, 884–894.
[63] Yoshinaka, Y., Katoh, I., Copeland, T.D. and Oroszlan, S. (1985) J. Virol. 55, 870–873.
[64] Varmus, H.E. (1983) in: Mobile Genetic Elements (Shapiro, J.A. ed.) pp.411–503, Academic Press, New York.
[65] Mellor, J., Fulton, S.M., Dobson, M.J., Wilson, W., Kingsman, S.M. and Kingsman, A.J. (1985) Nature 313, 243–246.
[66] Clare, J. and Farabaugh, P. (1985) Proc. Natl. Acad. Sci. USA 82, 2829–2833.
[67] Wilson, W., Malim, M.H., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Nucleic Acids Res. 14, 7001–7016.
[68] Saigo, K., Kugimiya, W., Matsuo, Y., Inouye, S., Yoshioka, K. and Yuki, S. (1984) Nature 312, 659–661.
[69] Hull, R. and Covey, S.N. (1986) J. Gen. Virol. 67, 1751–1758.
[70] Penswick, J., Hübler, R. and Hohn, T. (1988) J. Virol. 62, 1460–1463.
[71] Roth, J.R. (1981) Cell 24, 601–602.
[72] Curran, J.F. and Yarus, M. (1987) Science 238, 1545–1550.
[73] Weiss, R.B., Dunn, D.M., Atkins, J.F. and Gesteland, R.F. (1988) Cold Spring Harbour Symposium on Quantitative Biology LII, 685–693.
[74] Huang, W.M., Ao, S.-Z., Casjens, S., Orlandi, R., Zeikus, R., Weiss, R., Winge, D. and Fang, M. (1988) Science 239, 1005–1011.
[75] Murphy, E.C. jr, Wills, N. and Arlinghaus, R.B. (1980) J. Virol. 34, 464–473.
[76] Strauss, E.G., Rice, C.M. and Strauss, J.H. (1983) Proc. Natl. Acad. Sci. USA 80, 5271–5275.
[77] Strauss, E.G., Rice, C.M. and Strauss, J.H. (1984) Virology 133, 92–110.
[78] Harbison, S.-A., Davies, J.W. and Wilson, T.M.A. (1985) J. Gen. Virol. 66, 2597–2604.
[79] Guilley, J., Carrington, J.C., Balazs, E., Jonard, G., Richards, K. and Morris, T.J. (1985) Nucleic Acids Res. 13, 6663–6677.
[80] Ziegler, V., Richards, K., Guilley, H., Jonard, G. and Putz, C. (1985) J. Gen. Virol. 66, 2079–2087.
[81] Bouzoubaa, S., Ziegler, V., Beck, D., Guilley, H., Richards, K. and Jonard, G. (1986) J. Gen. Virol. 67, 1689–1700.
[82] Morch, M.-D., Drugeon, G. and Benicourt, C. (1982) Virology 119, 193–198; Morch, M.-D., Boyer, J.C. and Haenni, A.L. (1988) Nucleic Acids Res. 16, in press.
[83] Pelham, H.R.B. (1979) Virology 97, 256–265.
[84] Morris-Krsinich, B.A.M. and Forster, R.L.S. (1983) Virology 128, 176–185.
[85] Hsu, Y.H. and Brakke, M.K. (1985) Virology 143, 272–279.
[86] Joshi, S., Neeleman, L., Pleij, C.W.A., Haenni, A.-L., Chapeville, F., Bosch, L. and Van Vloten-Doting, L. (1984) Virology 139, 231–242.
[87] Engelberg-Kulka, H. and Schulaker-Schwarz, R. (1988) J. Theor. Biol. 131, 477–485.
[88] Bossi, L. and Roth, J.R. (1980) Nature 286, 123–127.
[89] Miller, J.H. and Albertini, A.M. (1983) J. Mol. Biol. 164, 59–71.
[90] Bossi, L. (1983) J. Mol. Biol. 164, 73–87.
[91] Engelberg-Kulka, H. (1981) Nucleic Acids Res. 9, 983–991.
[92] Haenni, A.-L., Morch, M.-D., Drugeon, G., Valle, R.P.C., Joshi, R.L. and Denial, T.M. (1987) in: Positive Strand RNA Viruses (Brinton, M.A. and Rueckert, R.R. eds) pp.149–158, A.R. Liss, New York.
[93] Bienz, M., Kübli, E., Kohli, J., De Henau, S., Huez, G., Marbaix, G. and Grosjean, H. (1981) Nucleic Acids Res. 9, 3835–3850.
[94] Sedivy, J.M., Capone, J.P., RajBhandary, U.L. and Sharp, P.A. (1987) Cell 50, 379–389.
[95] Geller, A.I. and Rich, A. (1980) Nature 283, 41–46.
[96] Hopfield, J.J. (1974) Proc. Natl. Acad. Sci. USA 71, 4135–4139.
[97] Ninio, J. (1975) Biochimie 57, 587–595.
[98] Ninio, J. (1986) in: Accuracy in Molecular Processes: its Control and Relevance to Living Systems (Kirkwood, T.B.L. et al. eds) pp.291–328, Chapman and Hill, London.
[99] Ninio, J. (1988) Cold Spring Harbor Symposium on Quantitative Biology LII, 639–646.
[100] Kurland, C.G. (1987) Trends Biochem. Sci. 12, 169–171.
[101] Thompson, R.C. (1988) Trends Biochem. Sci. 13, 91–93.
[102] Wolf, H., Chinali, G. and Parmeggiani, A. (1971) Eur. J. Biochem. 75, 67–75.
[103] Murgola, E.J., Pagel, F.T. and Hijazi, K.A. (1984) J. Mol. Biol. 175, 19–27.
[104] Fuller, W. and Hodgson, A. (1967) Nature 215, 817–820.
[105] Vijgenboom, E., Vink, T., Kraal, B. and Bosch, L. (1985) EMBO J. 4, 1049–1059.
[106] Hughes, D., Atkins, J.F. and Thompson, S. (1987) EMBO J. 6, 4235–4239.
[107] Morch, M.-D. and Benicourt, C. (1980) Eur. J. Biochem. 105, 445–451.
[108] Hryniewicz, M.M. and Von der Haar, R.A. (1983) Mol. Gen. Genet. 190, 336–343.
[109] Kurkinen, M. (1981) FEBS Lett. 124, 79–83.
[110] Kurland, C.G. (1987) Trends Biochem. Sci. 12, 210–212.
[111] Kurland, C.G. (1987) Trends Biochem. Sci. 12, 126–128.