Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

6 Armstrong, C. M., Bezanilla, F. M. and Horowicz, P. (1972) *Biochem. Biophys. Acta* 267, 605–608
7 Brum, G., Rios, E. and Stefani, E. (1988) *J. Physiol.* 398, 441–473
8 Baylor, S. and Hollingworth, S. (1988) *J. Physiol. (London)* 403, 151–192
9 Volpe, P., Krause, K-H., Hashimoto, S., Zorzato, F., Pozzan, T., Melcolesi, J. and Lew, D. P. (1988) *Proc. Natl Acad. Sci. USA* 85, 1091–1095
10 Benedetti, A., Fulceri, R., Romani, A. and Comporti, M. (1988) *J. Biol. Chem.* 263, 3466–3473
11 Melzer, W., Rios, E. and Schneider, M. F. (1986) *J. Physiol.* 372, 261–292
12 Volpe, P., Salviati, G., DiVirgilio, F. and Pozzan, T. (1985) *Nature* 316, 347–349
13 Vergara, J., Tsien, R. Y. and Delay, M. (1985) *Proc. Natl Acad. Sci. USA* 82, 6352–6356
14 Donaldson, S. K., Goldberg, N. D., Walseth, T. F. and Huetmann, D. A. (1988) *Proc. Natl Acad. Sci. USA* 85, 5749–5753
15 Walker, J. W., Somlyo, A. V., Goldman, Y. E., Somlyo, A. P. and Trentham, S. R. (1987) *Nature* 327, 249–252
16 Fabiato, A. (1986) *Biophys. J.* 49, 190a
17 Carrasco, M. A., Magenzo, K., Jaimovich, E. and Hidalgo, C. (1988) *Arch. Biochem. Biophys.* 262, 360–366
18 Varsanyi, M., Messer, M. and Brandt, N. R. (1989) *Eur. J. Biochem.* 179, 473–479
19 Schneider, M. F. and Chandler, W. K. (1973) *Nature* 242, 244–246
20 Meissner, G., Darling, E. and Eveleth, J. (1986) *Biochemistry* 25, 236–244
21 Chandler, W. K., Rakowski, R. F. and Schneider, M. F. (1976) *J. Physiol.* 254, 285–316
22 Cadwell, J. J. S. and Caswell, A. H. (1982) *J. Cell Biol.* 93, 543–550
23 Kawamoto, R. M., Brunschwig, J-P., Kim, K. C. and Caswell, A. H. (1986) *J. Cell Biol.* 103, 1405–1414
24 Inui, M., Saito, A. and Fleischer, S. (1987) *J. Biol. Chem.* 262, 1740–1747
25 Imagawa, T., Smith, J. S., Coronado, R. and Campbell, K. P. (1987) *J. Biol. Chem.* 2621, 16636–16643
26 Lai, A. F., Erickson, H. P., Rousseau, E., Liu, Q. Y. and Meissner, G. (1988) *Nature* 331, 315–319
27 Hui, C. S., Milton, R. L. and Eisenberg, R. S. (1984) *Proc. Natl Acad. Sci. USA* 81, 2582–2585
28 Rios, E. and Brum, G. (1987) *Nature* 325, 717–720
29 Lamb, G. D. (1986) *J. Physiol.* 376, 85–100
30 Beam, K. G. and Knudson, C. M. (1988) *J. Gen. Physiol.* 91, 799–815
31 Block, B. A., Imagawa, T., Campbell, K. P. and Franzini-Armstrong, C. (1988) *J. Cell Biol.* 107, 2587–2600
32 Chadwick, C. C., Inui, M. and Fleischer, S. (1988) *J. Biol. Chem.* 262, 10872–10877

Talking Point

How 'hidden' reading frames are expressed

Roberto Cattaneo

Secondary reading frames, 'hidden' under other reading frames, are used for coordinated expression of proteins in several eukaryotic viruses. In some genes, ribosomal frameshifting and initiation or reinitiation of protein synthesis on internal AUG codons are translational mechanisms allowing access to such 'hidden' reading frames. In others, secondary reading frames are translated from alternatively spliced or edited mRNAs.

Eukaryotic mRNAs are generally monocistronic, but for several viral transcripts[1–8] as well as for one cellular transcript[9], it has been shown that proteins encoded in different reading frames are expressed. Since eukaryotic ribosomes typically start protein synthesis at the first AUG codon only[4], it is not immediately evident how secondary reading frames overlapping with or following the first reading frame can be expressed.

Ribosomal frameshifting

A situation present in many viral eukaryotic transcripts is illustrated in Fig. 1: the first open reading frame (top, stippled box) overlaps with a second one (top, black box). Thus, if translation begins at the first AUG codon, only one protein will be expressed (bottom left, stippled box). Ribosomal frameshifting, subsequent to initiation on the first AUG codon, may enable expression of the overlapping reading frame, creating a fusion protein (bottom right, fused stippled and black boxes). Since only a fraction of the ribosomes change frame at a frameshift signal, fusion proteins are produced in addition to, rather than instead of, the 'normal' protein.

Frameshifting in the −1 frame is used by most retroviruses to access the reverse transcriptase reading frame[5,6], hidden in the gag mRNA, and by coronaviruses to express their RNA replicase[10], encoded by two different reading frames. Until now, no clear-cut case of ribosomal frameshifting has been observed in cellular genes of higher eukaryotes, and it seems that cellular eukaryotic genes avoid sequences on which frameshifting could occur[6]. In contrast, the yeast retroviral-like element Ty shifts to the +1 frame to express its reverse transcriptase[7], and the ribosomes of *E. coli* can slip, shift, step backward and hop forward (−2, −1, +1, +2, +5 and +6 frameshifts)[11].

Internal initiation

The use of an internal AUG codon for initiation of translation is an alternative to ribosomal frameshifting, and does not result in production of fusion proteins (Fig. 1, bottom center, black box). Several conditions may allow the use of an internal AUG codon. (1) An internal AUG is sometimes used in cases where the first AUG occurs in an unfavourable context for translation initiation[4]. (2) Termination of protein synthesis at a stop codon may lead to reinitiation at a nearby AUG in another frame[9,12]. (3) A 'ribosome landing pad' may direct the ribosome to an internal position in the mRNA, as described for the uncapped genomic RNA of picornaviruses[13]. (4) In a capped mRNA of a paramyxovirus, some ribosomes pass from the cap directly to an initiation codon far downstream[14].

It is interesting to note that the cousins of retroviruses and the yeast Ty elements, the hepatitis B-like viruses (HBVs), also have the reverse transcriptase reading frame hidden under an overlapping reading frame initiating upstream. However, in contrast to the retroviruses, HBVs avoid ribosomal frameshifting and instead use an internal AUG codon to express reverse

R. Cattaneo is at the Institut für Molekularbiologie I, Universität Zürich, Hönggerberg, 8093 Zürich, Switzerland.

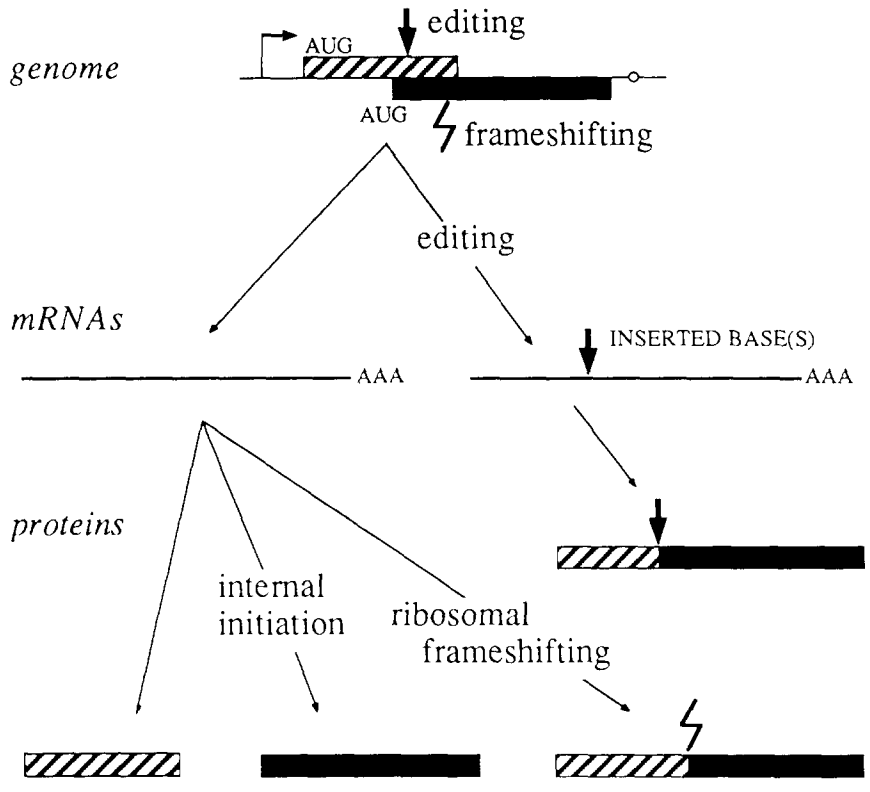

Fig. 1. Three mechanisms allowing expression of 'hidden' reading frames.

transcriptase[8,15]. Even more intriguingly, ribosomes can use ACG or AUA for initiation on the HBV reverse transcriptase reading frame, albeit at considerably reduced efficiency[8]. This is reminiscent of a trick used by the adeno-associated virus (a parvovirus) and Sendai virus (a paramyxovirus) to initiate protein synthesis with different efficiency on ACG or AUG codons in the same mRNA[16,17].

Portable ribosomal frameshifting signals[6], as well as portable ribosome landing pads[13], have been defined and used to study the expression of foreign proteins[18]. These tools can now be used experimentally to co-ordinate expression of specific proteins.

Viral RNA editing: reading frame switch

In other viral and cellular eukaryotic genes, secondary reading frames are translated from alternatively spliced or edited mRNAs. As alternative splicing as an expression mechanism has been reviewed recently[19], I will concentrate on RNA editing. The three different forms of RNA editing (Table I) are all characterized by addition or by substitution of nucleotides during or after transcription. The viral form of RNA editing is a paradigm of viral thriftiness. Paramyxoviruses appear to edit one of their mRNAs by a mechanism similar to the one they use for polyadenylation, i.e. 'stuttering' of the polymerase. This enzyme is believed to stutter at the end of each gene on short stretches of template U residues to polyadenylate viral mRNAs. In 20–50% of the phosphoprotein mRNAs of SV5 (Ref. 20) and measles virus[21], nontemplated G residues are inserted at a precise position, probably by means of intragenic polymerase stuttering on a stretch of templated Cs. The protein products of the viral type of RNA editing are fusion proteins (Fig. 1, right). As with ribosomal frameshifting (which also produces fusion proteins), viral RNA editing does not act on all template molecules, and allows coordinated expression of two types of proteins.

Cellular RNA editing: new stop codon or alteration of the reading frame

Cellular types of RNA editing seem to be mechanistically unrelated to the viral type. In the mammalian apolipoprotein gene, RNA editing is characterized by mutation of a single nucleotide (Table I). This editing process allows tissue-specific expression of proteins differing in length at the carboxyl terminus[22,23]. Proteins with carboxyl termini of different lengths can also result from naturally occurring suppression of stop codons, a phenomenon well known in prokaryotic[24] and eukaryotic[25] systems.

A perplexing type of RNA editing has been described to occur in the mitochondria of trypanosomes. For several genes of these organelles, faithfully transcribed mRNAs do not have AUG codons in the protein-coding frame near the 5′ end of the mRNA. Faithful transcripts of these genes are detected, if at all, only in certain developmental stages, whereas in other stages mRNAs with U residues inserted at multiple positions have been observed[26,27]. Further extraordinary observations have recently been made: in the transcript of one trypanosomal gene, the coding region is extensively altered by insertion of one or more U residues at 121 sites and by deletion of Us at seven sites[28]. Moreover, mRNAs coding for very similar proteins are expressed in three species of trypanosomes which carry remarkably different genes in their mitochondrial genomes[29]. These two last observations have been challenged by Maizels and Weiner[30], who have postulated the existence of an RNA template or of a nuclear gene coding for these mRNAs.

'Hidden' reading frames in cellular genes?

Ribosomal frameshifting, internal initiation of protein synthesis, and RNA editing are frequently used in viral systems to allow the expression of 'hidden' reading frames. In contrast, in cellular genes overlapping reading frames are present only rarely, and to my knowledge none has been shown to code for an expressed protein. In cellular genes of higher eukaryotes, coordi-

Table I. Three types of RNA editing

Gene and organism	Type of modification	Base	Site(s), frequency and other characteristics	Consequences for protein expression
Phosphoprotein gene of paramyxoviruses	Insertion	G	Single, 20–50%	Additional protein (reading frame switch)
Apolipoprotein gene of mammals	Point mutation	C to U	Single, up to 100%, tissue-specific	Shorter protein (new stop codon)
Mitochondrial genes of trypanosomes	Insertion and deletion	U	Multiple, up to 100%, developmentally regulated	Limited or extensive alteration of reading frame

nated expression of multiple proteins from a single transcription unit is obtained generally by alternative splicing[19] or by initiation on different in-frame AUGs[4,31]. In the first case, the reading frames do not generally overlap, and in the second, only one reading frame is used. A notable exception to this rule might be the gene coding for a transcription factor, OCT2 (Ref. 32). In this mRNA, a second reading frame, encoding 278 amino acids, largely overlaps with the main reading frame. It should, however, be mentioned that the OCT2 coding region is very GC-rich and so stop codons are expected only at a very low frequency. I predict that cellular genes, because of the minimal constraints imposed on their size, will only rarely express proteins from overlapping reading frames. The exceptions might be found in genes coding for proteins which have to be regulated in a strictly coordinated fashion, either during development or in a specific tissue. However, 'hidden' reading frames may be rare, but the corresponding proteins may well have important biological functions[33].

Acknowledgements

I thank Hans Weber for discussions and linguistic assistance.

References

1 Bos, J. L., Polder, L. J., Bernards, R., Schrier, P. I., van den Ellen, P. J., van der Eb, A. J. and Ormondt, H. (1981) *Cell* 27, 121–131
2 Shaw, M. W., Choppin, P. W. and Lamb, R. A. (1983) *Proc. Natl Acad. Sci. USA* 80, 4879–4883
3 Giorgi, C., Blumberg, B. M. and Kolakofsky, D. (1983) *Cell* 35, 829–836
4 Kozak, M. (1986) *Cell* 47, 481–483
5 Jacks, T. and Varmus, H. E. (1985) *Science* 230, 1237–1242
6 Jacks, T., Madhani, H. D., Masiarz, F. R. and Varmus, H. E. (1988) *Cell* 55, 447–458
7 Wilson, W., Malim, M. H., Mellor, J., Kingsman, A. J. and Kingsman, S. M. (1986) *Nucleic Acids Res.* 14, 7001–7016
8 Schlicht, H-J., Radziwill, G. and Schaller, H. (1989) *Cell* 56, 85–92
9 Thomas, K. R. and Capecchi, M. R. (1986) *Nature* 324, 34–38
10 Brierley, I., Boursnell, M. E. G., Binns, M. M., Bilimoria, B., Block, V. C., Brown, T. D. K. and Inglis, S. C. (1987) *EMBO J.* 6, 3779–3785
11 Weiss, R. B., Dunn, D. M., Atkins, J. F. and Gesteland, R. F. (1987) *CSH Symp. Quant. Biol.* 52, 687–693
12 Liu C-C., Simonsen, C. C. and Levinson, A. D. (1984) *Nature* 309, 82–85
13 Pelletier, J. and Sonenberg, N. (1988) *Nature* 334, 320–325
14 Curran, J. A. and Kolakofsky, D. (1988) *EMBO J.* 7, 2869–2874
15 Chang, L. J., Pryciak, P., Ganem, D. and Varmus, H. (1989) *Nature* 337, 364–368
16 Becerra, S.P., Rose, J. A., Hardy, M., Baroyoy, B. M. and Anderson, C. W. (1985) *Proc. Natl Acad. Sci. USA* 82, 7919–7923
17 Curran, J. A. and Kolakofsky, D. (1988) *EMBO J.* 7, 245–251
18 Edery, I., Petryshyn, R. and Sonenberg, N. (1989) *Cell* 56, 303–312
19 Breitbart, R. E., Andreadis, A. and Nadal-Ginard, H. (1987) *Annu. Rev. Biochem.* 56, 467–495
20 Thomas, S. M., Lamb, R. A. and Paterson, R. G. (1988) *Cell* 54, 891–902
21 Cattaneo, R., Kaelin, K., Baczko, K. and Billeter, M. A. (1989) *Cell* 56, 759–764
22 Powell, L. M., Wallis, S. C., Pease, R. J., Edwards, Y. H., Knott, T. J. and Scott, J. (1987) *Cell* 50, 831–840
23 Chen, S-H., Habib, G., Yang, C. Y., Gu, Z. N., Lee, B. R., Weng, S. A., Silberman, S. R., Cai, S. J., Deslypere, J. P., Rossenev, M., Gotto, A. M., Li, W. H. and Chan, L. (1987) *Science* 238, 363–366
24 Weiner, A. M. and Weber, K. (1971) *Nature (London), New Biol.* 234, 206–209
25 Pelham, H. R. B. (1978) *Nature* 272, 469–471
26 Benne, R., van den Burg, J., Brakenhoff, J. P. J., Sloof, P., van Boom, J. H. and Tromp, M. C. (1986) *Cell* 46, 819–826
27 Feagin, J. E., Jasmer, D. P. and Stuart, K. (1987) *Cell* 49, 337–345
28 Feagin, J. E., Abraham, J. M. and Stuart, K. (1988) *Cell* 53, 413–422
29 Shaw, J. M., Feagin, J. E., Stuart, K. and Simpson, L. (1988) *Cell* 53, 401–411
30 Maizels, N. and Weiner, A. (1988) *Nature* 334, 469–470
31 Kozak, M. (1988) *J. Cell. Biol.* 107, 1–7
32 Clerc, R. G., Corcoran, L. M., LeBonite, J. H., Baltimore, D. and Sharp, P. A. (1988) *Genes Dev.* 2, 1570–1581
33 Berkhout, B. and van Duin, J. (1985) *Nucleic Acids Res.* 13, 6955–6968