More surprises in translation: Initiation without the initiator tRNA

Uttam L. RajBhandary*

Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139

One of the most widespread beliefs in molecular biology is that protein synthesis is initiated with methionine or formylmethionine in all organisms, by using AUG as the initiation codon and a special methionine tRNA called the initiator tRNA (1, 2). Eubacteria, mitochondria, and chloroplasts initiate protein synthesis with formylmethionine, whereas archaeobacteria and eukaryotic cytoplasm initiate with methionine. In eubacteria, codons such as GUG and UUG, which are related to AUG by a single base change, are occasionally used for initiation. However, these codons are read by the same initiator tRNA and are, therefore, translated as formylmethionine. In eukaryotic systems, AUG is almost exclusively the codon used for initiation. In the rare cases where ACG, CUG, AUU, and AGG are used, protein synthesis is still thought to be initiated with methionine (3–5).

The first surprising exception to the universally accepted notion that protein synthesis is always initiated with methionine or formylmethionine was provided by the demonstration that protein synthesis in *Escherichia coli* could be initiated with codons other than AUG and amino acids other than methionine (6, 7). This work involved the use of anticodon sequence mutants of the *E. coli* initiator tRNA (Fig. 1). The anticodon mutation allows the mutant initiator tRNA to initiate protein synthesis by using a codon complementary to the new anticodon. Because the anticodon is, in many cases, an important

determinant for recognition of the tRNA by aminoacyl-tRNA synthetases, the mutant tRNAs are often aminoacylated with different amino acids. Thus, protein synthesis could even be initiated with UAG, a termination codon, and with formylglutamine (6). Recent studies have extended some of these findings also to mammalian cells (8). Surprising and unexpected as these results were, translation of the mutant mRNAs still required an initiator tRNA, although in a mutated form, for initiation of protein synthesis.

A paper by Sasaki and Nakashima now provides another surprising exception (9). It shows that translation of an insect viral RNA is initiated with glutamine and without the participation of the initiator methionine tRNA. A special RNA structure, including a pseudoknot involving nucleotides immediately upstream of the glutamine codon, is shown to be necessary for this exceptional mechanism of translation initiation. Coming on the heels of a paper by Shastri and coworkers (10) on the antigen presentation by Class I MHC molecules of a cryptic peptide, whose synthesis is likely initiated with leucine by using CUG as the initiation codon, these two papers together provide strong evidence that translation initiation in insect cells and in mammalian cells can sometimes occur with amino acids other than methionine and without the obligatory participation of the initiator methionine tRNA. Working with another insect virus, Sarnow and coworkers have independently come to a similar conclusion recently (J. E. Wilson, M. J. Powell, S. E. Hoover, and P. Sarnow, personal communication).

Sasaki and Nakashima (9) were studying the translation of the capsid protein gene of the plautia stali intestinal virus (PSIV). This RNA virus is classified as a picorna-like insect virus. Its genome organization (Fig. 2) is very similar to that of three other insect viruses that have been sequenced recently (11–14). These insect viral RNAs contain two long ORFs, which are thought to be translated independently. This is different from other eukaryotic mRNAs, which are mostly mono-

cistronic. The upstream ORF codes for the nonstructural viral proteins such as RNA helicase, a cysteine protease and an RNA replicase, whereas the downstream ORF codes for the structural proteins. The capsid protein is synthesized as a polyprotein precursor, with the capsid protein as the N-terminal protein of the precursor. This polyprotein is then cleaved, presumably, by the virally encoded protease, as in poliovirus (15), to produce the structural proteins of the virus. The capsid protein contains glutamine at its N terminus. However, sequence analysis of the RNA showed that there was no AUG codon preceding the CAA glutamine codon (11). Although ORF1 and ORF2 in the PSIV RNA are in frame and separated by the termination codon of ORF1, it is unlikely that they are translated together. For example, introduction of a UAA termination codon three codons upstream of the CAA codon had no effect on synthesis of the capsid protein *in vitro* (16), suggesting that initiation, at least *in vitro*, was occurring at the CAA codon or one or two codons upstream of the CAA codon.

Translation of the PSIV RNA involves “cap-independent” internal initiation, which is different from the “cap-dependent” initiation used on most eukaryotic cellular and viral mRNAs. Translation initiation involves the assembly of the 40S and the 60S ribosomal subunits, mRNA, initiator methionyl-tRNA (Met-tRNA), GTP, and at least nine different initiation factors to form the 80S initiation complex (1, 17–19). Eukaryotic mRNAs contain a cap structure at the 5′-end and a poly(A) sequence at the 3′-end (20). The first step of translation initiation on these mRNAs (Fig. 3) is the formation of a specific ternary complex between the initiation factor eIF2, GTP, and the initiator Met-

Fig. 1. Codon-anticodon pairing between wild-type mRNA and wild-type *E. coli* initiator tRNA^{fMet} and mutant mRNA (AUG→UAG) and mutant initiator tRNA (CAU→CUA).

See companion article at www.pnas.org/cgi/doi/10.1073/pnas.010426997

*To whom reprint requests should be addressed. E-mail: bhandary@mit.edu.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.040579197. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.040579197

Fig. 2. Genome organization of PSIV RNA (adapted from ref. 9).

tRNA. This ternary complex is thought to bind to the 40S ribosomal subunit to form a 43S complex, which in turn binds to the capped 5'-end of the mRNA. The ribosome then moves along the mRNA in a process called scanning until it reaches the AUG codon (21) to form the 48S initiation complex. The joining of the 60S subunit to this complex requires the prior action of the initiation factor eIF5 (22). The scanning model of eukaryotic translation initiation proposed by Kozak is supported by a substantial amount of experimental evidence. The first AUG codon is almost always the codon selected for initiation. When the first AUG is not part of a long ORF, it is often used for translational regulation of the mRNA (23). Genetic and biochemical experiments have shown that the initiator Met-tRNA and the initiation factors eIF1, eIF2, and eIF5 play important roles in helping the scanning ribosome locate the initiation codon on the mRNA (3, 24, 25).

Viral mRNAs such as PSIV lack the cap structure at the 5'-end. In these and some cellular mRNAs, there is an alternative pathway of translation initiation (called internal initiation) in which ribosomes bind directly to the initiation codon or

Fig. 3. Schematic diagram of "cap-dependent" 80S initiation complex formation (adapted from ref. 17).

Fig. 4. Schematic diagram of the capsid-coding region of PSIV RNA highlighting the CAA initiation codon (red) and the base pairing leading to formation of part of a pseudoknot structure (adapted from ref. 9).

bind internally in the mRNA upstream of the initiation codon and move along the mRNA until an AUG initiation codon is reached (15, 26–28). The first examples of internal initiation came from studies on poliovirus and encephalomyocarditis virus RNAs (26, 27). Besides lacking a 5'-terminal cap structure, these mRNAs have 5'-noncoding regions (5'NCRs) that are several hundred nucleotides long and that allow the ribosome to bind internally to the mRNA. These 5'-NCRs upstream of the coding sequences are highly structured and are known as internal ribosome entry site (IRES). Many of these IRES sequences also contain a widely occurring RNA structural motif called a pseudoknot (29). The classical pseudoknot is a structure formed from RNA stem-loop sequences in which the sequences in the loop form base pairs with single-stranded regions elsewhere in the RNA (Fig. 4). Pseudoknots are found in many different types of RNAs and play important roles in their structure and function (30–32).

The PSIV RNA has a pseudoknot involving nucleotides that are immediately upstream of the CAA codon used for translation initiation of the capsid protein gene (Fig. 4). Sasaki and Nakashima show that this pseudoknot structure is critical for the IRES-dependent translation of the capsid protein *in vitro* (9). Mutations that disrupt the base pairing between the AAG of the AAGUG and the CUU of the CACUU sequences highlighted in Fig. 4 abolish translation of the capsid protein mRNA, whereas compensatory mutations that regenerate the base pairing restore translation of the capsid protein. Remarkably, one of these compensatory mutations introduces UAG, a termination codon, immediately upstream of the CAA codon, and yet the mRNA is translated well. This result indicates strongly that the CAA codon is used to initiate protein synthesis without the involvement of the initiator Met-tRNA. Further evidence for this is derived from translation of mRNA

constructs corresponding to the capsid coding sequences, which do not specify any internal methionine residues. This mRNA is translated to yield a protein that does not contain methionine. Most interestingly, the IRES-dependent translation of a mutant transcript in which the CAA codon is immediately preceded by an AUG codon also yields a protein lacking methionine. In contrast, when the same coding sequence is translated by using the scanning mechanism by removal of the IRES from the transcript, translation now starts from the AUG, and the protein contains methionine.

Sasaki and Nakashima's work provides a striking example of the important role of the mRNA structure in facilitating translation initiation without the canonical initiator methionine tRNA. It also raises the question of whether there is something special about the CAA codon and the glutamine tRNA used for this unusual process of translation initiation. It is likely, however, that the nature of the codon and the tRNA used for initiation is less important than the need for a specific mRNA structure immediately upstream of the codon. First, three other insect viruses, including Drosophila C virus (DCV), Rhopalosiphum padi virus (RhPV), and Himetobi P virus (HiPV) have genome organizations very similar to that of PSIV, and they also contain similar pseudoknot structures immediately upstream of GCU or GCA codons (12–14). These codons specify alanine, and alanine is known to be the N-terminal amino acid of the capsid protein in at least DCV and HiPV (12, 14). Thus, it is quite likely that these viruses initiate protein synthesis with alanine and without the involvement of the normal initiator methionine tRNA. Such a possibility was mentioned by Johnson and Christian for DCV (12). Second, Sasaki *et al.* showed that the translational defect of a mRNA construct carrying insertion of a C residue between the CUU and CAA sequences in Fig. 4 could be

restored by deletion of a G residue downstream (16). Although the translational defect was, at the time, interpreted as being caused by a shift in the reading frame after initiation from CUU, it is equally possible that protein synthesis initiated from CCA, a proline codon, and terminated downstream.

Do some mammalian viruses use a similarly unusual mechanism of translation initiation? Translation of hepatitis C virus (HCV) RNA may have some features in common with PSIV and the other three insect viral RNAs. HCV RNA is translated in an IRES-dependent manner and has a pseudoknot structure 11–12 nucleotides upstream of the AUG codon that is important for initiation. HCV RNA requires fewer of the translation initiation factors, and ribosomes can bind directly to the HCV RNA in the absence of all initiation factors (33). Although HCV RNA uses AUG as the initiation codon, mutants in which the AUG codon has been changed to AUU or CUG are at least 80–90% as active in initiation (34). The retention of such a high level of translational activity in the AUU and CUG mutants is quite unexpected. Although this has been interpreted as being because of the strong contribution of coding sequences immediately downstream of the AUG toward IRES activity, it is possible that the HCV IRES also has the structural features and the potential to initiate protein synthesis without the obligatory involvement of the initiator methionine tRNA. This possibility can be tested with the mutants available (33, 34).

How do ribosomes initiate protein synthesis without the initiator Met-tRNA? Ribosomes can translate RNA sequences without the need for an initiation codon, as demonstrated in experiments leading to the elucidation of the genetic code (35, 36). Also, the work of Kaji and his colleagues on ribosome recycling factor (RRF) from *E. coli* may be quite relevant. RRF, an essential protein in *E. coli* (37), is involved in dissociation of the 70S ribosome after ribosomes have reached a termination codon (38, 39). Interestingly, inactivation of RRF *in vivo* leads to random translation of RNA sequences downstream of the termination codon (38). Presumably the 70S ribosomes bound to the mRNA are in an “elongation mode” and translate any codon downstream. In eukaryotes, there is no evidence for the presence of RRF in the cytoplasm, although mitochondria and chloroplasts are thought to contain RRF (38). With PSIV and the other insect viruses mentioned above, it is possible that the structure of the IRES element including the pseudoknot positions the ribosome on the mRNA in such a way that the ribosome is in an “elongation mode” and uses the next codon in line to begin translation. It might even be speculated that 80S ribosomes bind directly to the capsid-coding region of PSIV or that the binding of the 40S ribosomal subunit is followed by the joining of the 60S subunit without the requirement for an initiator tRNA. Much of the machinery for translation initiation is used for the assembly of the 48S and the 80S initiation complexes. Therefore, it is possible that assembly of such a complex on the PSIV ORF2 could

require fewer of the initiation factors, as for assembly of the HCV RNA initiation complex (33). For example, it might be predicted that eIF2, eIF4E, and eIF5 may not be necessary. Other questions that may be asked are: (i) are the glutamine tRNA and the CAA codon in the 80S ribosome-PSIV ORF2 initiation complex at the P site or at the A site? (ii) how is the complex assembled? These and many other questions will, undoubtedly, be addressed in future work.

In a recent review, Dever (40) has mentioned a common misconception among many scientists that all of the important questions in protein synthesis were solved in the 1960s and 1970s. Dever's article highlighting exciting new developments in translation and translational regulation (40) and recent reports of ribosome crystal structure from several laboratories should put to rest such a misconception. Translation remains an area of interest to many. As in other areas of biology, there is no predicting where the next advance or the next surprise will come from. As we absorb these new findings (refs. 9, 10; P. Sarnow, personal communication) and marvel at the seemingly endless diversity in biology, it is important to note that the fundamentals of the normal translation machinery and its regulation are still not understood and deserve continued investigation.

I thank Annmarie McInnis for cheerfulness and care in the preparation of this commentary. Work in our laboratory is supported by grants R37GM17151 from National Institutes of Health and DAAD19-99-1-0300 from the United States Army Research Office.

- Kozak, M. (1983) *Microbiol. Rev.* **47**, 1–45.
- RajBhandary, U. L. & Chow, C. M. (1995) in *tRNA: Structure, Biosynthesis, and Function*, eds. Söll, D. & RajBhandary, U. L. (Am. Soc. Microbiol., Washington, DC), pp. 511–528.
- Cigan, A. M., Feng, L. & Donahue, T. F. (1988) *Science* **242**, 93–97.
- Peabody, D. S. (1989) *J. Biol. Chem.* **264**, 5031–5035.
- Becerra, S. P., Rose, J. A., Hardy, M., Baroudy, B. M. & Anderson, C. W. (1985) *Proc. Natl. Acad. Sci. USA* **76**, 7919–7923.
- Varshney, U. & RajBhandary, U. L. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1586–1590.
- Chattopadhyay, R., Pelka, H. & Schulman, L. H. (1990) *Biochemistry* **29**, 4263–4268.
- Drabkin, H. J. & RajBhandary, U. L. (1998) *Mol. Cell. Biol.* **18**, 5140–5147.
- Sasaki, J. & Nakashima, N. (January 31, 2000) *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.010426997. <http://www.pnas.org/cgi/doi/10.1073/pnas.010426997>
- Malarkannan, S., Horng, T., Shih, P. P., Schwab, S. & Shastri, N. (1999) *Immunity* **10**, 681–690.
- Sasaki, J., Nakashima, N., Saito, H. & Noda, H. (1998) *Virology* **244**, 50–58.
- Johnson, K. N. & Christian, P. D. (1998) *J. Gen. Virol.* **79**, 191–203.
- Moon, J. S., Domier, L. L., Mccoppin, N. K., D'Arcy, C. J. & Jin, H. (1998) *Virology* **243**, 54–65.
- Nakashima, N., Sasaki, J. & Toriyawa, S. (1999) *Arch. Virol.* **14**, 1–8.
- Ehrenfeld, E. in *Translational Control*, eds. Hershey, J. W. B., Mathews, M. B. & Sonenberg, N. (Cold Spring Harbor Lab. Press, New York), pp. 549–573.
- Sasaki, J. & Nakashima, N. (1999) *J. Virol.* **73**, 1219–1226.
- Merrick, W. C. & Hershey, J. W. B. (1996) in *Translational Control*, eds. Hershey, J. W. B., Mathews, M. B. & Sonenberg, N. (Cold Spring Harbor Lab. Press, New York), pp. 31–70.
- Gingras, A.-C., Raught, B. & Sonenberg, N. (1999) *Annu. Rev. Biochem.* **68**, 913–963.
- Sachs, A. B., Sarnow, P. & Hentze, M. W. (1997) *Cell* **89**, 831–838.
- Both, G. W., Furuichi, Y., Muthukrishnan, S. & Shatkin, A. J. (1975) *Cell* **6**, 185–195.
- Kozak, M. (1989) *J. Cell Biol.* **108**, 229–241.
- Chakrabarti, A. & Maitra, U. (1991) *J. Biol. Chem.* **266**, 14039–14045.
- Hinnebusch, A. G. (1997) *J. Biol. Chem.* **272**, 21661–21664.
- Huang, H.-k., Yoon, H., Hannig, E. M. & Donahue, T. F. (1997) *Genes Dev.* **11**, 2396–2413.
- Pestova, T. V., Borukhov, S. I. & Hellen, C. T. (1998) *Nature (London)* **394**, 854–859.
- Jang, S. K., Krausslich, H. G., Nicklin, M. J. H., Duke, G. M., Palmenberg, A. C. & Wimmer, E. (1988) *J. Virol.* **62**, 2636–2643.
- Pelletier, J. & Sonenberg, N. (1988) *Nature (London)* **334**, 320–325.
- Macejak, D. G. & Sarnow, P. (1991) *Nature (London)* **350**, 90–94.
- Pleij, C. W. A., Rietveld, K. & Bosch, L. (1985) *Nucleic Acids Res.* **13**, 1717–1731.
- Jacks, T., Madhani, H. D., Masiarz, F. R. & Varmus, H. E. (1988) *Cell* **55**, 447–458.
- Schimmel, P. R. (1989) *Cell* **58**, 9–12.
- Kim, Y.-G., Maas, S., O'Neill, A. & Rich, A. (1999) *Proc. Natl. Acad. Sci. USA* **96**, 14234–14239.
- Pestova, T. V., Shatsky, I. N., Fletcher, S. P., Jackson, R. J. & Hellen, C. U. T. (1998) *Genes Dev.* **12**, 67–83.
- Reynolds, J. E., Kaminski, A., Kettinen, H. J., Grace, K., Clarke, B. E., Carroll, A. R., Rowlands, D. J. & Jackson, R. J. (1995) *EMBO J.* **14**, 6010–6020.
- Nirenberg, M. W. & Matthaei, J. H. (1961) *Proc. Natl. Acad. Sci. USA* **47**, 1588–1602.
- Nishimura, S., Jones, D. S. & Khorana, H. G. (1965) *J. Mol. Biol.* **13**, 302–324.
- Janosi, L., Shimizu, I. & Kaji, A. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 4249–4253.
- Janosi, L., Mottagui-Tabar, S., Isaksson, L. A., Sekine, Y., Ohtsubo, E., Zhang, S., Goon, S., Nelken, S., Shuda, M. & Kaji, A. (1998) *EMBO J.* **17**, 1141–1151.
- Karimi, R., Pavlov, M. Y., Buckingham, R. H. & Ehrenberg, M. (1999) *Mol. Cell* **3**, 601–609.
- Dever, T. E. (1999) *Trends Biochem. Sci.* **24**, 398–403.