Vol. 140, No. 1, 1986 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
October 15, 1986 Pages 288-293

THE SELECTION OF THE FIRST AUG AS THE INITIATOR OF EUCARYOTIC mRNAs TRANSLATION IS FAVORED BY A 5'-TERMINAL CAP GROUP AND A PURINE IN THE -3' POSITION

Elzbieta Szczesna-Skorupa, David A. Mead, and Byron Kemper

Department of Physiology and Biophysics, and College of Medicine at Champaign-Urbana, University of Illinois, Urbana, IL 61801

Received June 18, 1986

SUMMARY: We have constructed a series of plasmids containing variations in the DNA that codes for the translation initiation region of preproparathyroid hormone. Gel electrophoretic analyses of the character and extent of synthesis of various modified preproparathyroid hormone like proteins derived from the mRNAs transcribed from these plasmids reveal that the presence of a 5'-terminal cap group on mRNA's facilitates recognition of the most 5'-terminal AUG sequence on a mRNA and that AUG sequences within the consensus sequence PuXXAUGPu are favored sites for the initiation of mRNA translation.

Several mechanisms have been implicated in the initiation of eucaryotic mRNA translation. First, the presence of a 5'-terminal cap group enhances the initiation of eucaryotic mRNA translation (1,2). Secondly, most initiations of eucaryotic mRNA translations occur on the first AUG from the 5'-terminus (3). This led to an initial theory that translation of these mRNAs occurs at the first 5'-terminal AUG codon (4). However, initiation of translation of eucaryotic mRNAs sometimes occurs at AUG sequences other than the most 5'-terminal AUG (3). This has prompted a theory that the sequence context of the initiator AUG is important. Thus, an AUG within the sequence PuXXAUGPu is preferred as the initiator codon over a more 5' proximal AUG in an unfavorable sequence context (3,5,6). These mechanisms have been incorporated into a generalized "scanning mechanism" in which ribosomes initially bind to a 5'-terminal cap group and then "scan down" from the 5'-terminus of the mRNA to stop at the first AUG or first PuXXAUGPu sequence (7). Much of this reasoning is based upon sequence analyses of eucaryotic mRNAs and few experimental studies of the translation of specific mRNAs

0006-291X/86 $1.50

with specific nucleotide sequences have been designed to test the relative importance of these mechanisms. This paper therefore reports the translation of a series of modified mRNAs containing multiple AUGs near the 5'terminus of the mRNAs. Gel electrophoretic analyses of proteins derived from translation of these mRNAs suggest that the 5'terminal cap as well as the sequence context of the AUG influences the selection of an initiator AUG.

MATERIALS AND METHODS

Enzymes and Chemicals

Restriction endonucleases were obtained from Bethesda Research Laboratories, New England Biolabs, or Promega Biotec. T4 DNA ligase, RNasin, and SP6 RNA polymerase were the products of Promega Biotec. E. coli DNA polymerase (Klenow fragment) was from Boehringer Mannheim. T4 DNA polymerase, nucleotides and m^7GpppG were from PL Biochemicals. The Nco I linker, CCCATGGG, was purchased from New England Biolabs.

Plasmid constructions.

Plasmid pPNTH, containing an insert of a modified preProPTH cDNA was used as the parent plasmid for further construction (8). Introduction of the additional ATG triplets into different positions, in phase with the preproPNTH coding region was obtained by inserting the Nco I linker directly into a restriction site. Plasmids p23Ctr and p23C were constructed by inserting an Nco I linker into a Hinc II site of pPNTH. The protein encoded by p23C contains 23 amino acids added to the amino terminus of pPNTH. p23Ctr lost a G residue at the site of the Nco I linker insertion during the cloning procedure, which resulted in the upstream terminated tripeptide. The protein encoded by p23G also contains an additional 23 amino acids. To construct p23G, the Sca I to Nco I fragment to the 5' of the coding region of p23C was replaced with the analogous Sca I to Sal I fragment of pPNTH. The Sal I and Nco I ends were repaired by incubation with the Klenow fragment of E. coli DNA polymerase I before ligation.

In vitro transcription and translation.

Conditions for transcription were as described (8). For production of capped RNA 500 uM m^7GpppG was added and the GTP concentration was decreased to 50 uM (9).

Translations in the wheat germ and reticulocyte lysate cell-free systems were carried out as described (10) except that optimal potassium concentrations of 70 mM and 120 mM were used for uncapped and capped RNA's, respectively. The ratio of initiation at the two AUG's was independent of the potassium concentration.

RESULTS

We have constructed a series of plasmids containing modified versions of a cDNA for "stretched" preproparathyroid hormone (8). Linearization of these plasmids at a restriction site downstream from the cDNA insert followed by *in vitro* run off transcription of the linearized plasmids in the presence and absence of the cap primer

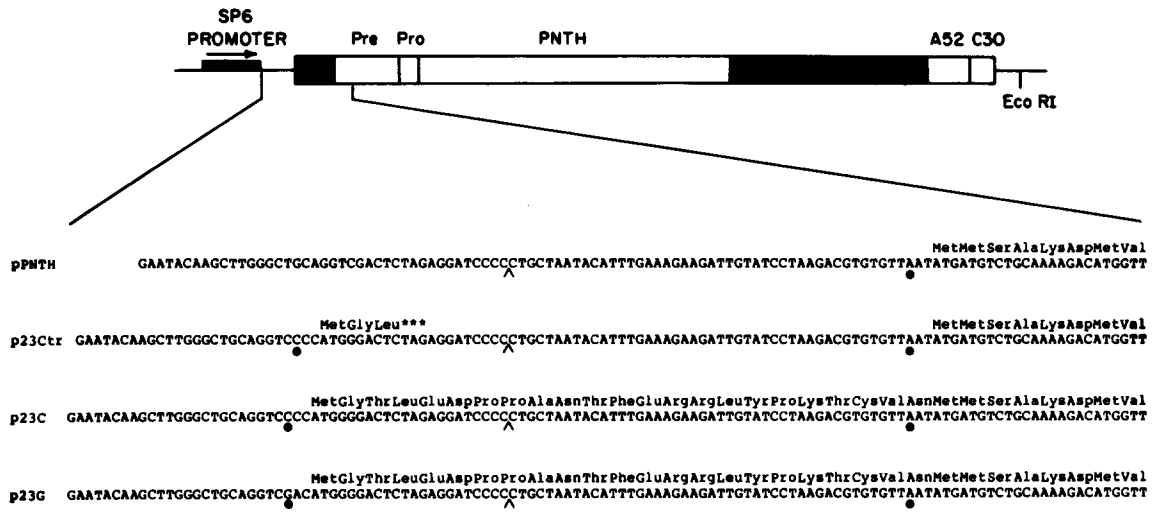

Fig.1. Alterations at the 5' end of the preProPNTH cDNA. The nucleotide sequence of the entire 5' untranslated region and a portion of the translated region of the altered proteins is shown below a diagram of the cDNA insert of plasmid pPNTH. The cDNA insert contains 52 deoxyadenylate acid residues (A52) and 30 deoxycytidylic acid residues (C30) at the 3' terminus. The filled boxes represent the 5' and 3' noncoding regions, and the open boxes represent regions coding for preProPNTH. The positions of the -3 nucleotides from the natural or new translation initiation codons are identified by the dots. Nucleotides to the left of the arrows are from the remainder of the multiple cloning site of the parent plasmid pPNTH, while nucleotides to the right of the arrows originate from PTH cDNA.

m^7GpppG (9) yields a series of 5'-terminal capped or uncapped mRNAs containing variations in their ribonucleotide sequences. Figure 1 shows the DNA sequences of these plasmids analogous to the mRNAs in question along with the expected translation products.

Translation of the varied mRNAs indicated in Figure 1 in 2 different eukaryotic *in vitro* systems and analysis by gel electrophoresis of the radiolabeled protein product(s) (Figure 2) allow a direct *in vitro* assay for the relative dominance as the initiator of the various AUGs in the 5' region. One can then examine if the presence of a 5'-terminal cap group enhances the relative use of an engineered first 5'-terminal AUG compared to that of the natural PuXXAUGPu translation start of preproparathyroid hormone. As can be seen from a comparison of the quantity of the largest protein product derived from the translation of the mRNAs from the p23G and p23C

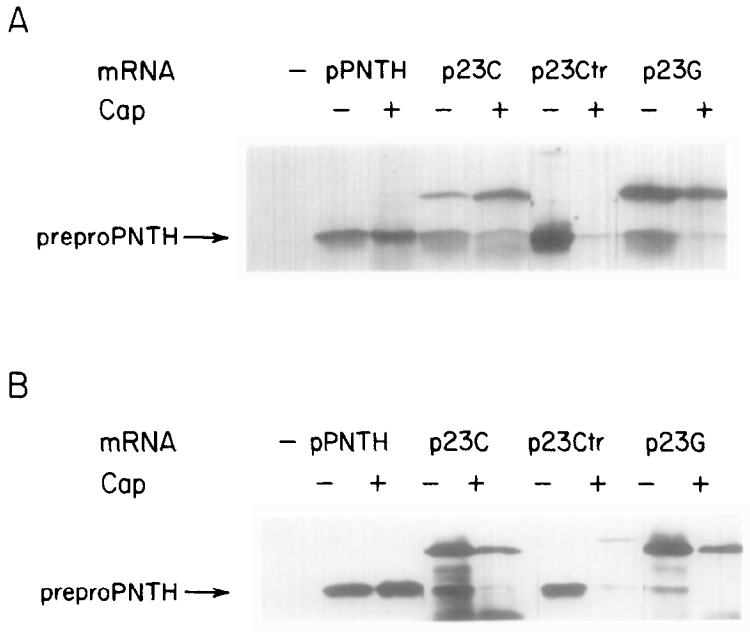

Fig.2. Analysis by a gel electrophoresis of ^{35}S-methionine labelled cell-free translational products. One hundred ng of *in vitro* transcribed capped or uncapped poly A containing RNA's were translated in reticulocyte (panel A) or wheat germ (panel B) cell-free systems. The first lanes in A and B are reactions without mRNA added and are indicated by the dash. Translation reactions were analyzed by polyacrylamide gel electrophoresis and fluorography (10). Fluorography was 16 hr except for p23Ctr. To visualize the weak signal for the capped p23Ctr RNA, fluorography was 30 hr.

plasmids containing and lacking cap groups, the presence of a 5'-terminal cap favors the first AUG sequence as an initiator codon over the following PuXXAUGPu natural start of preproparathyroid hormone synthesis. This point is strengthened by data from translation of the capped and uncapped transcripts from the p23Ctr plasmid. These transcripts contain initial AUG sequence preceeding the natural PuXXAUGPu translation initiation site. An in phase terminator codon dictates that such early initiation of translation yields only a short peptide, undetectable in our gel electrophoretic assay. After termination minimal reinitiation of translation at subsequent initiator codons occurs in cell-free systems (11). As can be seen, the presence of a 5'-terminal cap on the p23Ctr transcript decreases use of the sub sequent natural PuXXAUGPu translation start, presumably due to in-

creased use of the earlier AUG codon as an initiator codon in the capped mRNA.

One can also examine if an AUG codon with a purine rather than a pyrimidine in the -3 position, i.e. PuXXAUGPu versus PyXXAUGPu, favors use of a more 5'-terminal AUG as the initiator codon over the natural initiator codon of preproparathyroid hormone. A comparison of the + cap mRNA from the p23G and p23C plasmids addresses this point. As can be seen, the mRNAs from transcription of the p23G plasmid generally yield more efficient and more specific initiation of translation at the more 5'-terminal PuXXAUGPu site than that found with mRNA from p23C plasmid which contain a more 5'-terminal PyXXAUGPu sequence.

DISCUSSION

These results demonstrate the power of the combination of *in vitro* mutagenesis and an *in vitro* assay to study the functional significance of RNA structure. Disadvantages of the *in vitro* assay are also apparent. Specifically, both the rabbit reticulocyte and wheat germ systems often produce low levels of additional proteins besides the proteins predicted from the mRNAs transcribed from our plasmid constructs. This is most noticeable in the wheat germ translations of the capless p23C mRNA and the capped p23Ctr mRNA that yield unpredictable additional small proteins and a large protein, respectively. Proteins smaller than anticipated can arise from premature termination or cleavages of correctly initiated proteins. The abnormally large proteins probably arise from mRNA formed as a result of incomplete linearization of the p23Ctr plasmid prior to its transcription. However, in spite of these potential anomalies, the data clearly show that the presence of a 5'-terminal cap group on a mRNA enhances use of the first available AUG sequence in from the 5' end of a mRNA and that AUG sequences within a PuXXAUGPu consensus sequence are favored, when available, for the initiation of translation of mRNAs by these *in vitro* eucaryotic translation systems.

ACKNOWLEDGMENTS

We thank Dr. John Clark for critically reviewing the manuscript. This work was supported by NIH grant AM18866.

REFERENCES

1. Shatkin, A. J., (1976) Cell 9, 645-653.
2. Filipowicz, W., (1978) FEBS Lett. 96, 1-11.
3. Kozak, M., (1983) Microbiol. Rev. 47, 1-45.
4. Kozak, M., (1978) Cell 15, 1109-1123.
5. Kozak, M., (1984) Nature 308, 241-246.
6. Kozak, M., (1984) Nucleic Acids Res. 12, 857-872.
7. Kozak, M., (1984) Nucleic Acids Res. 12, 2872-2892.
8. Mead, D., Szczesna-Skorupa, E., and Kemper, B., (1985) Nucleic Acids Res. 13, 1103-1118.
9. Pelletier, J., and Sonenberg, N., (1985) Cell 40, 515-526.
10. Dorner, A. J., and Kemper, B., (1978) Biochemistry 17, 5550-5555.
11. Kozak, M., (1986) Cell 4, 283-292.