



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.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

## VIRAL AND CELLULAR mRNA CAPPING: PAST AND PROSPECTS

Yasuhiro Furuichi<sup>1</sup> and Aaron J. Shatkin<sup>2</sup>

AGENE Research Institute, Kamakura, 247-0063 Japan<sup>1</sup>  
and Center for Advanced Biotechnology and Medicine<sup>2</sup>  
Piscataway, New Jersey 08854

- I. Introduction
- II. Discovery
  - A. Prologue
  - B. Stage 1
  - C. Stage 2
  - D. Stage 3
  - E. Stage 4
  - F. Stage 5
- III. Caps in Eukaryotic mRNAs
  - A. Cellular mRNAs
  - B. Viral mRNAs
- IV. Mechanisms of Synthesis
  - A. Cotranscriptional Capping: Type 1
  - B. Genes and Enzymes Directing Cotranscriptional Capping
  - C. Vesicular Stomatitis Virus mRNA Capping: Type 2
  - D. Formation of TMG Cap and Mppp Cap in snRNAs
- V. Preformed Capped 5' End Addition to Cellular and Viral mRNAs
  - A. Capped mRNA Formation by *Trans*-Splicing in Nematodes and Trypanosomes
  - B. Cap Snatching in Influenza Virus Transcription
  - C. Possibility for Capped Leader-Primed Transcription in Coronavirus mRNA Synthesis
- VI. Cap Functions
  - A. Protection of Pre-mRNA and mRNA Against 5' Exonucleolytic Attack
  - B. Splicing and Nucleocytoplasmic Transport
  - C. Translational Effects
- VII. Conclusion
- References

### I. INTRODUCTION

Cap structures of the type  $m^7\text{GpppN(m)pN(m)p}$  are present at the 5' ends of nearly all eukaryotic cellular and viral mRNAs (Fig. 3.1). A cap is added to cellular mRNA precursors and to transcripts of viruses that replicate in the nucleus during the initial phases of transcription

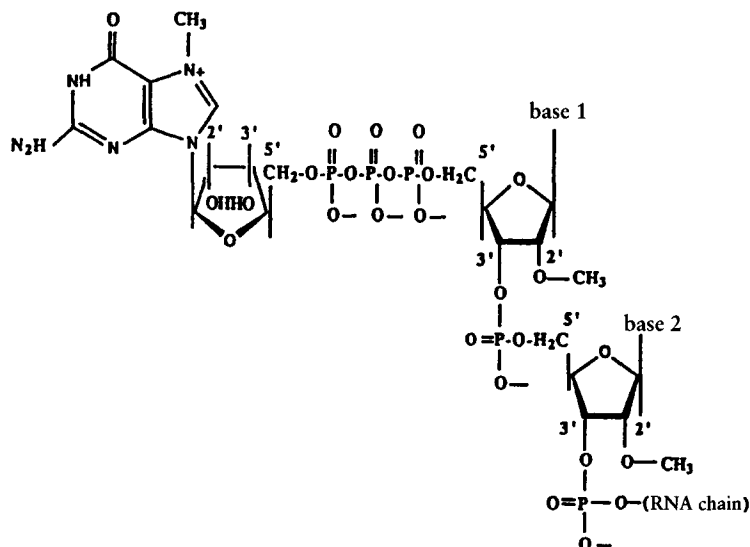


FIG 3.1. Cap structure. Cap 2 ( $m^7\text{GpppN1mpNm2p}$ ) showing the  $N^7$ -methylguanosine 5'-5' linked to the 5' end of the primary transcript.

and before other processing events including internal  $N^6\text{A}$  methylation, 3'-poly(A) addition, and exon splicing (1-5). Similar structures with multiple methyl groups on the terminal guanosine residue (i.e.,  $2,2,7 m_3 \text{GpppN}^m$ ) are also present on small nuclear (sn)RNAs U1, U2, U3, U4, and U5, which act as scaffolds for RNA splicing (6). Most, if not all, capped mRNAs have a single methyl group on the terminal G residue at the  $N^7$  position, whereas the adjacent nucleotides can be 2'-*O*-methylated to different extents, providing a basis for the following cap nomenclature:  $m^7\text{GpppN}$  (Cap 0),  $m^7\text{GpppNm}$  (Cap 1) and  $m^7\text{GpppNmpNm}$  (Cap 2). The methyl groups in Cap 0 and Cap 1 structures are added in the nucleus, and the additional 2'-*O*-methylation in Cap 2 is a cytoplasmic event. In many lower eukaryotes including yeast, mRNAs contain mainly Cap 0, and higher organisms usually have more extensively methylated caps (1-6).

Despite the variations on the methylation theme, the important biological consequences of a cap structure appear to correlate with the  $N^7$ -methyl on the 5'-terminal G and the two pyrophosphoryl bonds that connect  $m^7\text{G}$  in a 5'-5' configuration to the first nucleotide of mRNA. For example, caps increase mRNA stability by protecting against 5'→3'-exonucleolytic degradation (7-9). Unmethylated caps can also stabilize mRNA to a certain extent, but unless the 5'-terminal  $\text{GpppN}$

is converted to m<sup>7</sup>GpppN, it is capable of being hydrolyzed to (p)pN and susceptible to degradation (7).

Splicing accuracy and efficiency are both increased by the presence of the 5'-terminal m<sup>7</sup>GpppN (10, 11), possibly because of the involvement of nuclear cap-binding protein(s) (12–14) analogous to the cytoplasmic cap-binding initiation factor eIF4F that is required for eukaryotic protein synthesis (15). The 5'-terminal m<sup>7</sup>GpppN may also be involved in other critical steps in gene expression including transport from nucleus to cytoplasm. The 5' cap and 3'-poly(A) are hallmark structures of eukaryotic mRNAs and are not present in rRNA, tRNA, or most prokaryotic mRNAs. These two processing structures have been the subjects of many review articles on their biological significance. Since their discovery in the early 1970s, these 5'- and 3'-terminal elements have contributed to the progress of molecular biology research in many important ways including their utilization for distinguishing and isolating various eukaryotic mRNAs. In this article, we provide some history on the discovery of the cap and an update of research on viral and cellular mRNA capping.

## II. DISCOVERY

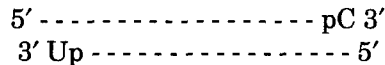
### A. Prologue

Before the identification of the cap on several viral mRNAs in 1974, it was generally believed that the 5' ends of all eukaryotic mRNAs, like those of *Escherichia coli* and bacteriophage mRNAs, are triphosphorylated, pppN forms, although no eukaryotic mRNA had been reported as 5' end-labeled by the standard, polynucleotide kinase method. Large-scale isolation of mRNA from eukaryotic cells was technically difficult, and the prevailing biochemical studies on *E. coli* and its phage had clearly shown that the mRNAs were triphosphorylated. *E. coli* mRNA contained pppA and pppG 5' ends; and all RNA phages such as MS2, Q $\beta$ , and R17 contained a 5'-triphosphorylated purine nucleoside in the genomic positive-strand RNA (and thus mRNA) extracted from purified particles. In contrast, the rRNAs and tRNAs, which are abundant in both eukaryotic and bacterial cells, had been shown to contain monophosphorylated 5' ends (16). Concerning methylated nucleotides, it was believed that such modified nucleotides were present only in tRNA and rRNA. Here, like other important biological findings made in the 1970s including 3'-polyadenylation, splicing, and reverse transcription, viral systems played a pivotal role in defining the structure of the cap and the biochemistry of cap formation.

*B. Stage 1*

Two kinds of findings that suggested the presence of methylated nucleotides in eukaryotic mRNAs were made in 1971–1972 in Japan and in the United States. During the summer of 1968, Dr. K. Miura came from Nagoya to NIH (Bethesda, MD) to begin work on reovirus genomic RNA with Shatkin. Analyses done in Bethesda and later in Mishima, Japan, in Miura's new laboratory at the National Institute of Genetics indicated the presence of 2'-*O*-methyl-Gp in reovirus dsRNA (17). In Japan, Furuichi and Miura (18,19) tried to define the 5'- and 3'-terminal sequences of dsRNA of cytoplasmic polyhedrosis virus (CPV), which contains 10 different genomic segments. Their initial interest was to understand how the 10 segments could be correctly packaged within the viral particle (an enigma not yet solved). Because their laboratory was new, they were the only researchers working on this project and without a technician. Thus, they divided the experiments in half: 3' sequencing by Furuichi and 5' sequencing by Miura. Furuichi used classic sequencing, not the currently used DNA sequencing techniques. This involved a series of tedious chemical reactions with stepwise reductive labeling of 3'-terminal nucleosides by [<sup>3</sup>H]borohydride after converting the *cis*-diol ribose moiety to two aldehyde residues by periodate-mediated oxidative ring opening.

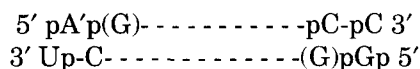
These efforts yielded the terminal sequences of



for the dsRNAs (18). Furuichi was not satisfied with the results, however, because about 30% of the <sup>3</sup>H label was incorporated into an undefined component, referred to as nonnucleosidic material (NNM), and the remaining <sup>3</sup>H counts were evenly distributed into the two terminal nucleosides C and U after digestion of labeled RNA by ribonucleases. The NNM (which later turned out to be Cap 1) was resistant to alkali, proteases, DNase, and ribonucleases but was sensitive to venom phosphodiesterase. It had a net negative charge of -4.5 at neutral pH and a size of about 5.4Å (19). Miura, on the other hand, was faced with the difficulty of trying to label the 5'-nucleotides, possibly G and A, based on the predicted end-to-end duplex structure of CPV dsRNA, by treatment with polynucleotide kinase and [ $\gamma$ -<sup>32</sup>P]ATP. This was achieved when Furuichi, who wanted to analyze the penultimate nucleotide species, prepared dsRNA in which the terminal C and U were removed by sequential oxidation (by periodate),  $\beta$ -elimination (by aniline), and dephosphorylation (by alkaline phosphatase). This RNA yielded C for

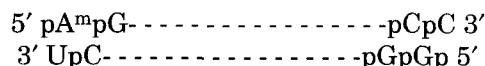
both strands in the penultimate positions to the terminal C and U, indicating that the initial 3'-terminal sequences of CPV RNA are ---pCpC and ---pCpU. Curiously, the NNM labeled by  $^3\text{H}$  in the first terminal nucleoside analysis was not obtained in the second  $^3\text{H}$ -labeling on the penultimate C. However, Miura was now satisfied because he was able to label efficiently the 5'-nucleotides pG and pA using kinase treatment and digestion with P1 nuclease to cut the phosphodiester bonds and leave the 5'-phosphate on the resulting nucleotides.

The 5' pG and pA made a perfect match with the complementary strand in the form of



However, Miura was puzzled by the slight difference in mobility of  $^{32}\text{P}$ -labeled pA compared with authentic pA in the two-dimensional paper chromatography system used for identifying the  $^{32}\text{P}$ -labeled 5'-nucleotides: the 5' pA' migrated slightly faster than the authentic marker pA. Further analyses of 5' oligonucleotides resulting from the 5'- $^{32}\text{P}$ -labeled RNA after digestion by guanine-specific RNase T<sub>1</sub>, pyrimidine-specific pancreatic ribonuclease A or nonspecific ribonuclease T<sub>2</sub> indicated that pG was the penultimate nucleotide to both 5' pA' and pG, establishing pA'pG and pGpG as the 5' sequences. More important, results showed that the 2' OH of the ribose in pA' of pA'pG was modified because the phosphodiester linkage between pA' and pG could not be digested by RNase T<sub>2</sub>, which requires a free OH for 2', 3'-cyclic phosphate formation as an RNA hydrolysis intermediate.

These were the days when chemistry was used for identifying modified bases in tRNA, and Miura readily obtained authentic 2'-O-methyl-pA to compare its chromatographic mobility with that of the 5'-[ $^{32}\text{P}$ ]A'. Both migrated in the identical position, showing clearly that A' was 2'-O-methyladenosine. This finding indicated for the first time that RNAs such as the genomic dsRNA of *Reoviridae*, CPV and reovirus, in addition to tRNA and rRNA, could contain methylated residues (17). Accordingly, the CPV genome RNA was considered to be in an end-to-end structure (20):



However, the NNM and its structural relationship to the 5' pA<sup>mp</sup>pG remained enigmatic, although it was apparent that the NNM some-

how protected the pA<sup>m</sup>pG-containing strand from the action of polynucleotide kinase and end labeling.

### C. Stage 2

Viruses in the dsRNA-containing *Reoviridae* family, including the prototype human reovirus, contain RNA polymerases, which conservatively copy the plus strands in the duplex genomic RNAs to produce viral mRNAs. Thus, one of the two strands comprising the dsRNA is the same as mRNA. Synthesis of viral mRNAs by *in vitro* transcription using purified reovirus had already been demonstrated (21), and K. Shimotohno, a graduate student who had joined Miura's group, started a study of transcription with purified CPV to examine which of the two strands had the same polarity as mRNA. The transcriptase activity of the CPV was low, about 100-fold less than reovirus, but Shimotohno succeeded in producing CPV transcripts by using a large amount of virus. He found that the CPV RNA produced *in vitro* had the same polarity as the pAmpG strand, and the 5' sequence was pppApGp (22). Indeed, the RNA could be labeled by [ $\gamma$ -<sup>32</sup>P]ATP in a 5'-triphosphorylated form, like mRNAs of *E. coli* and prokaryotic phage RNAs. Could the methyl groups in the genomic RNA be involved in the assembly of segmented genomes or transcription?

### D. Stage 3

One late night in the fall of 1973, Furuichi tested his idea that the weak CPV transcriptase activity *in vitro* might be activated by the addition of the methyl donor, *S*-adenosyl-L-methionine (AdoMet), thinking that the methylation of mRNA might be a prerequisite for natural viral mRNA synthesis in infected silkworm *Bombyx mori* cells. The result was indeed exciting, far beyond expectation, because the CPV transcriptase activity was stimulated more than 100-fold, reminiscent of similar effects seen when purified reovirus was treated with chymotrypsin (23). In addition, the resulting CPV transcripts were methylated at the 5' ends. Methylation occurred at the initiation of CPV transcription, and about two [<sup>3</sup>H]methyl groups from [<sup>3</sup>H]methyl-labeled AdoMet were incorporated per molecule of CPV RNA, one in the terminal pA and another in the NNM-like material that included the terminal pA. Furthermore, [ $\beta$ -<sup>32</sup>P]ATP could label the 5' termini of the CPV RNA, and the labeled phosphate was resistant to alkaline phosphatase, indicating that the 5' ppAm was blocked by something in

an NppAm structure in which the undefined N contained an additional methyl group. This unique "methylation-coupled transcription" was published in volume 1 of *Nucleic Acids Research* (24) in 1974, the same year in which Perry and Kelly at the Fox Chase Cancer Institute reported the presence of 2.2 residues of methylated nucleotides per 1000 residues in poly(A)-containing mRNA prepared from mouse L cells cultured in the presence of [<sup>3</sup>H-methyl]methionine (25). Here, the earlier findings of 3'-poly(A) in vaccinia virus and cellular mRNAs and a purification method for polyadenylated RNAs by oligo(dT)-cellulose greatly facilitated research on cellular mRNAs. In Rottman's laboratory at Michigan State University, m<sup>7</sup>Gp and 2'-O-methylated nucleotides as well as N<sup>6</sup>mAp were identified in Novikoff hepatoma mRNAs (26). In addition, in Dr. H. Busch's laboratory, Reddy *et al.*, (27) found the unique <sup>2,2,7</sup>m<sub>3</sub> GppAm structure in snRNAs (U1, U2, and U3) of Novikoff hepatoma cells. In Australia, Adams and Cory (28) also found several methylated nucleotides in a peculiar class of oligonucleotides, derived from <sup>32</sup>P-labeled mouse myeloma cell mRNA, which were resistant to ribonucleases and reproducibly migrated as a faint spot near the origin in two-dimensional fingerprinting analyses. After discussions of the results on cellular mRNAs and reovirus RNAs (17, 29) after a 1974 Gordon Research Conference, Rottman *et al.* (30) proposed m<sup>7</sup>GppNm as the 5' structure of eukaryotic cellular mRNAs in the newly started journal *Cell*. Although both this structure and the <sup>2,2,7</sup>m<sub>3</sub> GppAm in snRNA contained one-too-few phosphates, 1974 was the dawn of mRNA methylation and close to a complete elucidation of the cap.

#### E. Stage 4

In June of 1974, Furuichi joined Shatkin at the Roche Institute of Molecular Biology to continue the research on mRNA methylation using reovirus and CPV. A blocked 5' structure presumed to be NppG<sup>m</sup> was found in the reovirus mRNA synthesized *in vitro* in the presence of AdoMet; pG<sup>m</sup> was identified as 2'-O-methylated and the first nucleotide in the mRNA, but N remained to be identified. However, the 5' m<sup>7</sup>GppNm structure proposed by Rottman *et al.* (30) failed to account for the 5' NppAm and NppGm of CPV and reovirus mRNAs, respectively, which had a net negative charge of -2.5, because the net charge of the postulated m<sup>7</sup>GppNm was -1 to -1.5 owing to the +1 charge in m<sup>7</sup>G. Meanwhile, the N in CPV and reovirus mRNAs was found to consist of a phosphate group and a nucleoside with +1 charge,



suggesting that N might be AdoMet-p (AdoMet has a +1 charge), m<sup>7</sup>Gp, or m<sup>7</sup>Ap. An experiment that used [2-<sup>3</sup>H-methionyl]Adomet (containing an unlabeled methyl group) failed to label the mRNAs of either CPV or reovirus, showing that the N did not contain AdoMet. These results clearly limited the candidates for N to m<sup>7</sup>Gp or m<sup>7</sup>Ap, and the ultraviolet (UV)-fluorescent spot obtained with reovirus mRNA digests strongly suggested m<sup>7</sup>Gp. The final answer was obtained from an experiment in which the <sup>32</sup>P of [ $\alpha$ -<sup>32</sup>P]GTP was incorporated into the N of "NppAm" of CPV and "NppGm" of reovirus mRNAs. The results clearly indicated that the 5'-terminal structures of CPV and reovirus mRNAs were m<sup>7</sup>GpppAmp and m<sup>7</sup>GpppGmp (31,32), which fit in all the aspects with the mysterious NNM found before in CPV genomic RNA (19). At the same time, the large DNA-containing vaccinia virus that replicates in the cytoplasm also was found to produce mRNAs with 5' m<sup>7</sup>GpppA<sup>m</sup> and m<sup>7</sup>GpppG<sup>m</sup> (33,34); vaccinia also contains its own RNA polymerase, and mRNA synthesis *in vitro* facilitated the prompt identification of the blocked and methylated 5' terminus. The next question was: "Is this bizarre structure found generally in eukaryotic viral and cellular mRNAs?"

#### F. Stage 5

The poly(A)-containing mRNA of HeLa cells and nuclear precursor heterogeneous nuclear RNA (hnRNA) were examined for their 5'-terminal structures after <sup>32</sup>P-labeling and oligo (dT)-cellulose selection. The efforts made by us and in collaboration with the Darnell group at the Rockefeller University clearly showed that both the mRNA in the cytoplasm and the hnRNA in the nucleus contained the same general cap structure, m<sup>7</sup>GpppNm (35). In one of the manuscript preparation meetings, Darnell suggested that we needed a nickname to replace the complicated "blocked and methylated 5' structure" and said "Let's call it cap for short." Subsequent experiments done by us with <sup>32</sup>P pulse-labeled hnRNA suggested that caps were made at the initiation of transcription. The m<sup>7</sup>G caps were added to the products of RNA polymerase II shortly after transcriptional initiation within a narrow window of nascent oligonucleotides (+20 to +40), indicating that the process is rapid and efficient (36). Although snRNAs containing 2,2,7m<sub>3</sub>GpppNm are also products of RNA polymerase II, we have not focused on them in this review because they are not mRNAs and do not code for proteins. Excellent reviews have been written on snRNAs and their nuclear splicing functions (5,6).

## III. CAPS IN EUKARYOTIC MRNAs

## A. Cellular mRNAs

A cap structure  $m^7GpppN(m)$  consisting of 7-methylguanosine linked to the 5' end of the transcript by a 5'-5' triphosphate bridge was found in all eukaryotic cells examined (Table 3.1). Most lower eukaryotes including yeast, fungi, and amebas contain the less methylated Cap 0; higher eukaryotes including human have Cap 1 and 2 in which the first and second nucleotides of mRNA are 2'-O-methylated. Because methylation on the cap generally becomes more complex as organisms evolved, the existence of cap on mRNA and its structural complexity are among the characteristics that distinguish eukaryotes from prokaryotes and higher from lower eukaryotes, although the spliced leader sequence on trypanosome mRNAs contains several 2'-O-methylated nucleotides (37). Because capping appears to be completed at a very early stage of transcription, cotranscriptional capping, rather than posttranscriptional, is likely for cellular mRNA synthesis *in vivo* (38). Biochemically, posttranscriptional capping can occur *in vitro* on 5' triphosphorylated RNA, on oligonucleotides, or even on ppNpN dinucleotide structures by the action of purified RNA 5'-triphosphatase (RTase) and guanylyltransferase (GTase; 39-41). However, this may not occur *in vivo* because a strong 5'-exonuclease activity that hydrolyzes uncapped RNA exists in the nucleus, and nascent RNAs would be readily digested before the completion of transcription unless their 5' ends are protected by a  $m^7G$  cap (7, 8). As the action of 5'-exonuclease yields 5' monophosphorylated RNAs, which are no longer substrates for capping enzyme, capping of nascent RNAs must be completed when they are short enough to be protected by RNA polymerase II complexes.

Historically, a prototype concept "methylation-coupled transcription" was first proposed for CPV mRNA synthesis before identification of the cap structure (24), and then the potential of "post-transcriptional capping" was shown as *in vitro* systems became possible using purified vaccinia virus capping enzyme. The current concept of cotranscriptional capping is based on the direct interaction between RNA polymerase II and capping enzymes (42-45), as well as the necessity to protect uncapped transcripts from degradation by nuclear 5'-exonuclease. The concept may evolve further to linkage of 5' capping to the initiation of transcription as observed in vaccinia virus, CPV, and reovirus transcription systems. In vaccinia virus, capping enzymes are required for viral RNA synthesis (46); in CPV, transcription *in vitro* is

TABLE 3.1  
CAPS IN EUKARYOTIC CELLULAR MRNAs

Source	Types of Cap <sup>a</sup>			Other methylations
	Cap 0	Cap 1	Cap 2	
Human HeLa cells	—	N <sub>1</sub> =Pu <sup>m</sup> ,Py <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	m <sup>6</sup> A
HeLa cell histone	—	N <sub>1</sub> =A <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup> ,G <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	n.d.
Mouse myeloma	—	N <sub>1</sub> =Pu <sup>m</sup> ,Py <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	m <sup>6</sup> A
Mouse erythroid	—	N <sub>1</sub> =Pu <sup>m</sup> ,Py <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	m <sup>6</sup> A
Mouse fibroblasts	—	N <sub>1</sub> =Pu <sup>m</sup> ,Py <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	m <sup>6</sup> A
Mouse kidney	—	N <sub>1</sub> =N <sup>m</sup>	N <sub>2</sub> =N <sup>m</sup>	m <sup>6</sup> A
Rat hepatoma	—	N <sub>1</sub> =Pu <sup>m</sup> ,Py <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	m <sup>6</sup> A
Hamster kidney (BHK-21)	—	N <sub>1</sub> =Pu <sup>m</sup> ,Py <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	m <sup>6</sup> A,m <sup>5</sup> C
Monkey kidney (BSC-1)	—	N <sub>1</sub> =Pu <sup>m</sup> ,Py <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =Pu <sup>m</sup> ,Py <sup>m</sup>	m <sup>6</sup> A
Mouse immunoglobulin	—	N <sub>1</sub> =G <sup>m</sup> ,m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =A <sup>m</sup>	m <sup>6</sup> A
Human globin	—	N <sub>1</sub> =m <sup>6</sup> A <sup>m</sup> ,A <sup>m</sup>	N <sub>2</sub> =C <sup>m</sup>	n.d.
Human gp 130	—	N <sub>1</sub> =N <sup>m</sup>	n.d.	n.d.
Human WS <sup>b</sup>	—	N <sub>1</sub> =A <sup>m</sup> ,G <sup>m</sup> ,U <sup>m</sup>	n.d.	n.d.
Human RTS <sup>c</sup>	—	N <sub>1</sub> =A <sup>m</sup> ,G <sup>m</sup> ,U <sup>m</sup>	n.d.	n.d.
Mouse globin	—	N <sub>1</sub> =m <sup>6</sup> A <sup>m</sup> ,A <sup>m</sup>	N <sub>2</sub> =C <sup>m</sup>	n.d.
Rabbit globin	—	N <sub>1</sub> =m <sup>6</sup> A <sup>m</sup>	N <sub>2</sub> =C <sup>m</sup>	n.d.
Duck globin	—	N <sub>1</sub> =N <sup>m</sup>	n.d.	n.d.
Chick ovalbumin	N <sub>1</sub> =A	N <sub>1</sub> =m <sup>6</sup> A <sup>m</sup> ,A <sup>m</sup>	N <sub>2</sub> =Py <sup>m</sup>	n.d.

Trout protamine	—	$N_1=N^m$	—	n.d.
<i>Drosophila</i>	$N_1=C,A,G$	$N_1=Py^m>Pu^m$	$N_2=Pu^m, Py^m$	n.d.
<i>Bombyx mori</i> silk fibroin	—	$N_1=A^m$	$N_2=U^m$	m <sup>6</sup> A
<i>Aedes albopictus</i>	—	$N_1=Pu^m, Py^m$	$N_2=Pu^m, Py^m$	m <sup>6</sup> A
<i>C. elegans</i>	—	TMG-capped mRNA <sup>2,2,7</sup>		—
Tobacco hornworm oocyte	GpppN	—	—	n.d.
Brine shrimp	—	$N_1=Am, Gm$	n.d.	m <sup>6</sup> A
Sea urchin embryo	—	$N_1=Pu^m>Py^m$	n.d.	m <sup>6</sup> A
Slime mold	$N_1=A>G$	$N_1=A^m$ (10%)	n.d.	n.d.
<i>Neurospora</i>	$N_1=A>G$	n.d.	n.d.	n.d.
Yeast, <i>S. cerevisiae</i>	$N_1=A>G$	n.d.	n.d.	n.d.
Wheat embryo	$N_1$	n.d.	n.d.	n.d.
Maize	$N_1$	n.d.	n.d.	n.d.
Soybean seeds	$N_1$	$N_1=N^m$	n.d.	n.d.

<sup>a</sup> Cap 0: m<sup>7</sup>GpppN<sub>1</sub>pN<sub>2</sub>p-, Cap 1: m<sup>7</sup>GpppN<sub>1</sub><sup>m</sup>pN<sub>2</sub>p-, Cap 2: m<sup>7</sup>GpppN<sub>1</sub><sup>m</sup>pN<sub>2</sub>p<sup>m</sup>-. n.d., Not done; —, absent.

<sup>b</sup> WS, Human Werner syndrome gene transcript.

<sup>c</sup> RTS, Human Rothmund-Thomson syndrome gene transcript.

<sup>d</sup> TMG, Trimethylguanosine (2,2,7-m<sub>3</sub>G).

stimulated by an allosteric effect of AdoMet, which presumably reacts with the methyltransferase/RNA polymerase complex (47–49). In reovirus transcription, preformed blocked dinucleotides GpppG and GppppG can be incorporated into the 5' end of viral mRNAs (50, 51). The blocked dinucleotides of general structure GpppN (N= A,C,G and U) and GppppN exist abundantly in brine shrimp cells (52) and are also produced by the GTase of CPV and reovirus during mRNA synthesis *in vitro* (53). Since these unmethylated blocked dinucleotides (and m<sup>7</sup>GpppN) are used equally as well as ATP or GTP by RNA polymerase for initiation, albeit not for elongation, they have been used in many cases to prepare capped mRNAs of desired sequences by transcription of recombinant plasmid DNA with purified SP6 phage polymerase and vaccinia methyltransferase (MTase).

The transcriptional initiation sites of eukaryotic cellular mRNAs and most viral mRNAs are *a priori* the cap sites, based on capping enzymes necessarily using the triphosphorylated 5' ends of nascent transcripts. Recently, a method referred to as “cap site hunting” has been invented to determine the adjacent nucleotide sequence at the cap site (54,55). The method shown in Fig. 3.2 uses the structural feature of the cap to generate a library containing the cap site cDNA and permits the determination of the exact cap site of a given mRNA species. In addition, major and minor transcription initiation sites can be estimated by comparing the frequency of cap site DNA sequences cloned in an *E. coli* plasmid. Cap site analyses made recently for the transcripts of human Werner syndrome and Rothmund–Thomson syndrome genes (TATA boxless housekeeping genes regulated by the SP1 transcriptional element) have provided several important notions about the initiation site of transcription (56). First, the start site is not strict but can be located at various positions within about 50 bases downstream of the promoter. Second, the starting nucleotides are predominantly A or G, whereas C and U can also be rarely used. Third, the most frequently used start sites are not necessarily the closest geographically to the promoter but vary depending on the gene. Certainly more work is needed to understand the mechanisms behind these findings. However, this apparent lax initiation may reflect the combined effects of RNA polymerase II and capping enzymes, (i.e. the selection of efficient/inefficient transcriptional initiation site(s) by polymerase and the successful/unsuccesful capping by capping machinery). In any case, the oligocapping method and the resulting cap site cDNA should provide insights into the transcriptional initiation of various types of genes. In most TATA box-containing genes, the cap sites are located about 25 bases downstream of the TATA box (57). Dierks *et al.* (58) have examined the effects of DNA sequences preceding the rab-

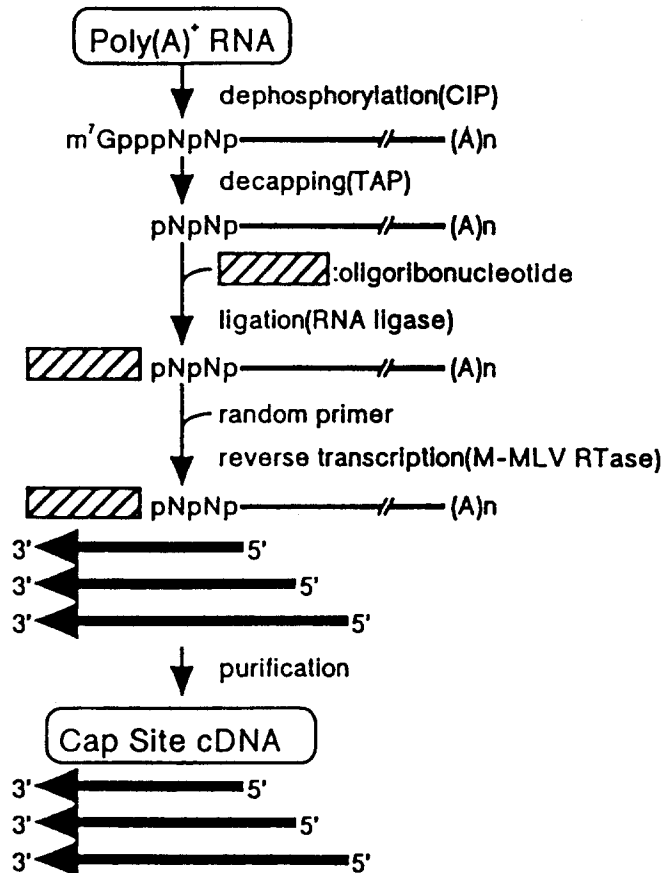


FIG 3.2. Preparation of cDNAs containing cap-site sequences by oligo-capping. CIP, Calf intestinal phosphatase; TAP, tobacco acid pyrophosphatase.

bit  $\beta$ -globin gene on the selection of correct initiation site(s) and showed that (1) in the case of the gene with only 14 base pairs of native 5' flanking sequence, a high level of  $\beta$ -globin RNA was produced, but none of the transcripts had the correct 5' end and most originated in the vector moiety; (2) with 66 base pairs of 5' flanking sequence, 5% of the cap sites were correct, and with 76 or more base pairs, 30–85% were correct. The region between 14 and 66 base pairs preceding the cap site contained the Hogness box and appeared to be essential for correct initiation of transcription. Thus, both the 5' flanking sequence and its length significantly influence the initiation/cap site of transcription.

Besides the regular m<sup>7</sup>G cap structure as represented by Cap 0, 1 or 2, there are other types of cap structures. The trimethylguanosine (TMG) cap, 2,2,7m<sub>3</sub>GpppN, was found in a fraction of the mRNAs in *Caenorhabditis elegans*, in addition to its general occurrence in eukaryotic snRNAs. The TMG cap present on the 22-nucleotide spliced leader RNA (SL RNA) exists as a small nuclear ribonucleoprotein and seems to be transferred to the pre-mRNA (59). By the same *trans*-splicing reactions, some of the *C. elegans* mRNAs containing a m<sup>7</sup>G cap may be formed posttranscriptionally by this unique *trans*-splicing reaction including translocation of a capped leader sequence from a 90–100 nucleotide precursor RNA to the pre-mRNA (60).

### B. Viral mRNAs

Viruses adopt various types of strategies for their parasitic replication and proliferation in infected cells (61). With a few exceptions, notably the picornavirus group, eukaryotic viral mRNAs contain the same cap structure as cellular mRNAs (Table 3.2), irrespective of their genomic structures and their replication strategies, probably because capped mRNAs are the functional form in host cells. Accordingly, the degree of methylation on viral mRNA caps correlates with host mRNA cap methylation, (i.e., viruses that infect higher eukaryotic cells generally contain Cap 1 and Cap 2 structures, whereas mRNAs of viruses that infect unicellular hosts and plant cells contain Cap 0). Although they may have the same cap structure, differences in the synthesis of capped mRNA exist among viruses. Smaller DNA viruses, such as adenoviruses and SV40 viruses, that use the host DNA synthetic machinery in the nucleus for their replication, also depend for their capped mRNA synthesis on the cellular system. Similarly, retroviruses such as avian sarcoma virus (ASV) and human immunodeficiency virus (HIV) also use host cell RNA polymerase II and capping enzymes for viral gene expression after their genomic RNAs are reverse-transcribed to DNAs and integrated into host cell chromosomes. Indeed, their genomic RNAs in the virus particles are capped, in the case of ASV by the presence of 5' m<sup>7</sup>GpppGmpCmp (62). On the other hand, larger DNA-containing baculoviruses that replicate in the nuclei of insect cells use both the host and their own transcription and capping systems. Baculovirus early mRNAs are synthesized by cellular RNA polymerase II, but the late and very late mRNAs are generated after DNA replication by a viral transcription complex consisting of four virus-encoded polypeptides that include RNA polymerase and capping enzymes coded for by the LEF-4 gene (63).

TABLE 3.2  
5' STRUCTURE OF VIRAL MRNAS

Viruses	Genome	Transcription	5' Structures			Other methylations
			Genome 5' end	mRNA 5' end		
<b>Mammals/Birds (RNA type)</b>						
Reo	dsRNA(±)	cytoplasm	7mGpppG <sup>m</sup> (+)/ppG(-)	7mGpppG <sup>m</sup> pC <sup>m</sup> p	n.d.	n.d.
Rous sarcoma	ssRNA(+)	nucleus	7mGpppG <sup>m</sup> pCp	n.d.	10-12 m <sup>6</sup> A	
Avian sarcoma	ssRNA(+)	nucleus	7mGpppG <sup>m</sup> pCp	n.d.	10 m <sup>6</sup> A	
Moloney murine leukemia	ssRNA(+)	nucleus	7mGpppG <sup>m</sup> p	n.d.	15-23 m <sup>6</sup> A	
Feline leukemia	ssRNA(+)	nucleus	7mGpppG <sup>m</sup> p	7mGpppG <sup>m</sup> pAp	10 m <sup>6</sup> A	
Sindbis	ssRNA(+)	cytoplasm	7mGpppA <sub>p</sub> U <sub>p</sub> Y <sub>p</sub> G <sub>p</sub> <sup>b</sup>	7mGpppApUpG	m <sup>5</sup> C	
Calci	ssRNA(+)	cytoplasm	protein combined	base unspecified	n.d.	
Dengue	ssRNA(+)	cytoplasm	7mGpppA <sup>m</sup> pNp	7mGpppm <sup>6</sup> A <sup>m</sup> p	n.d.	
VSV	ssRNA(-)	cytoplasm	(p)ppA	7mGpppm <sup>6</sup> A <sup>m</sup> pA <sup>m</sup> p	n.d.	
Influenza	ssRNA(-)	nucleus	pppA	7mGpppN <sup>m</sup> p	3 m <sup>6</sup> A	
Newcastle disease	ssRNA(-)	nucleus	n.d.	7mGpppN <sup>m</sup> p	n.d.	
Polio	ssRNA(+)	cytoplasm	protein-pUp	pUp	n.d.	
EMC	ssRNA(+)	cytoplasm	protein-pUp	pNp	n.d.	
<b>Mammals (DNA type)</b>						
Vaccinia	DNA	cytoplasm	—	7mGpppm <sup>6</sup> A <sup>m</sup> p/G <sup>m</sup> p	n.d.	
Adeno	DNA	nucleus	—	7mGpppm <sup>6</sup> A <sup>m</sup> p/A <sup>m</sup> pN <sub>2</sub> <sup>m</sup> p	m <sup>6</sup> A, m <sup>5</sup> C	
Simian virus 40	DNA	nucleus	—	7mGpppm <sup>6</sup> A <sup>m</sup> p/G <sup>m</sup> p/C <sup>m</sup> p	m <sup>6</sup> A	
Herpes simplex type 1	DNA	nucleus	—	7mGpppN <sub>1</sub> <sup>m</sup> pN <sub>2</sub> <sup>m</sup> p	m <sup>6</sup> A	
Polyoma	DNA	nucleus	—	7mGpppm <sup>6</sup> A <sup>m</sup> p/G <sup>m</sup> pN <sub>2</sub> <sup>m</sup> p	m <sup>6</sup> A	

(continued)



TABLE 3.2 (continued)

Viruses	Genome	Transcription	5' Structures		
			Genome 5' end	mRNA 5' end	Other methylations
<b>Insect</b>					
CPV	dsRNA(±)	cytoplasm	7mGpppA <sup>m</sup> p (+)/ppA (-)		n.d.
<b>Fish</b>					
Spring viremia of carp	ssRNA(-)	cytoplasm	pppA	m <sup>7</sup> GpppA <sup>m</sup> pN	n.d.
<b>Plants</b>					
Wound tumor	dsRNA(±)	m <sup>7</sup> Gppp	G <sup>m</sup> GUA(+)/AUCA(-)	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Brome mosaic	ssRNA(+)	cytoplasm	7mGpppGp	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Tobacco mosaic	ssRNA(+)	cytoplasm	7mGpppGp	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Alfalfa mosaic	ssRNA(+)	cytoplasm	7mGpppGp	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Turnip yellows	ssRNA(+)	cytoplasm	7mGpppGp/A	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Cucumber mosaic virus	ssRNA(+)	cytoplasm	7mGpppGp	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Barley stripe mosaic	ssRNA(+)	cytoplasm	7mGpppN	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Potato virus X	ssRNA(+)	cytoplasm	7mGpppGpA	m <sup>7</sup> GpppC <sup>m</sup>	n.d.
Satellite tobacco necrosis	ssRNA(+)	cytoplasm	ppApG	ppApG	n.d.
Cowpea mosaic	ssRNA(+)	cytoplasm	protein-pUp	n.d.	n.d.
<b>Mold</b>					
<i>Penicillium chrysogenum</i>	dsRNA(±)	cytoplasm	7mGpppN <sup>m</sup> p(+)		n.d.

Viruses that replicate in the cytoplasm of infected cells and are unable to use the nuclear transcription machinery have developed their own transcription and capping systems. Vaccinia and African swine fever are examples of such DNA viruses. They contain RNA polymerase and capping enzymes in the virus particle in addition to a large circular DNA genome. Studies of these viruses provided major contributions to the discovery of cap, as well as to the characterization of capping reactions by purified viral enzymes (40, 41, 64). Double-stranded RNA-containing reovirus and CPV, members of the *Reoviridae*, are other cytoplasmic viruses that contributed to the discovery of capped mRNA as described previously. It should be mentioned that the mechanism of cap formation on cellular mRNAs proved to be essentially the same as defined earlier for both vaccinia virus and *Reoviridae*, as discussed later.

The negative strand-containing RNA viruses vesicular stomatitis virus (VSV) and influenza virus also form capped viral mRNAs in infected cells but by mechanisms different from cellular capped mRNA synthesis (4). Influenza virus has negative-stranded genomic RNAs and contains in the virus particle RNA polymerase that produces positive-stranded mRNAs. Despite these characteristics of being a cytoplasmic virus, the replication of influenza virus was sensitive to actinomycin D, an inhibitor of DNA-mediated transcription by host RNA polymerase. This observation raised the question of whether the virus included a nuclear phase for replication. The Krug and Shatkin groups jointly made the surprising finding that the core of influenza virus, which migrated to the nucleus in infected cells, catalyzed capped oligonucleotide-primed viral mRNA synthesis by using as primers capped short fragments cleaved from the 5' ends of capped nascent pre-mRNA of host cells (65, 66). A virus-specific endonuclease that cleaves RNA polymerase II products 10–13 bases downstream of the 5' cap in a unique "cap-snatching mechanism" was later identified (66). Accordingly, the resulting influenza mRNAs contain randomized sequences at the 5' end derived from host pre-mRNA. Similar cap-snatching for mRNA synthesis has been found in other orthomyxoviruses (67). These findings emphasized the biological importance of 5' cap in mRNAs for their translation.

The only exceptions to capped mRNA are certain plant viruses, notably satellite tobacco necrosis virus (STNV) and viruses in the *Picornaviridae* family that includes the well-studied poliovirus and encephalomyocarditis (EMC) virus (and possibly caliciviruses) (Table 3.2). All these RNA viruses contain a single-stranded positive RNA genome, which can serve as mRNA at the early stages of infection, but genome RNAs isolated from purified virus particles do not

contain a cap. Instead, STNV RNA has a 5'-terminal ppAGUAAA-GACAGGAAACUUUACUGACUAACAUGGCAAAACAAC, which can form a perfect Watson-Crick base-paired structure between the 5'-nucleotides 1-7 and 16-22 (68). Such a stick-like secondary structure and potential hairpin loop are reminiscent of the base-paired 3' and 5' sequences of tRNA that may protect against hydrolysis by the cellular 5'→3'-exonuclease. The "capless" poliovirus genomic RNA contains instead a small protein, VPg, covalently attached to the 5' end in the form VPg-pUUAAAACAG-, via a phosphodiester bond that links pU to the OH group of tyrosine (69). Genomic RNA of EMC virus also contains 5' VPg-pU (70). The VPg seems to prime poliovirus and EMC virus RNA synthesis because there is a significant amount of abortive VPg-pU and of VPg-pU-containing short oligonucleotides in picornavirus-infected cell lysates (71). The mRNAs of poliovirus isolated from the polysomal fractions at the middle and late stage of infections, however, do not contain 5' covalently bound proteins but start with 5' pU (72). Both polio and EMC mRNAs are translated by a cap-independent mechanism in infected cells as mentioned in a later section.

#### IV. MECHANISMS OF SYNTHESIS

Subsequent to the cap discovery, the mechanism of capping was elucidated by studying the process associated with the *in vitro* synthesis of reovirus mRNA by viral cores (73). As shown in Fig. 3.3, it consists of four reactions catalyzed by distinct enzymatic activities. A very similar (type 1) mechanism of cap formation was observed for vaccinia virus and CPV (74,75). This same mechanism was confirmed for various cellular mRNAs using purified enzymes isolated from mammalian cells and yeast. Several recent studies that included gene targeting and the generation of temperature-sensitive yeast enzymes demonstrated that each of the steps involved in cap formation is essential for cell viability, as well as the replication of viruses (38,76). Besides the type 1 mechanism adopted generally by eukaryotic cells and most viruses, a different mechanism (type 2) has been found in VSV mRNA synthesis and perhaps related, negative-stranded RNA viruses (4, 64).

##### A. Cotranscriptional Capping: Type 1

Capping and subsequent methylation of most mRNAs occur by the scheme in Fig. 3.3. Nascent transcripts of RNA polymerase II are first modified by removal of the 5'- $\gamma$ -phosphate by RTase to yield a diphos-

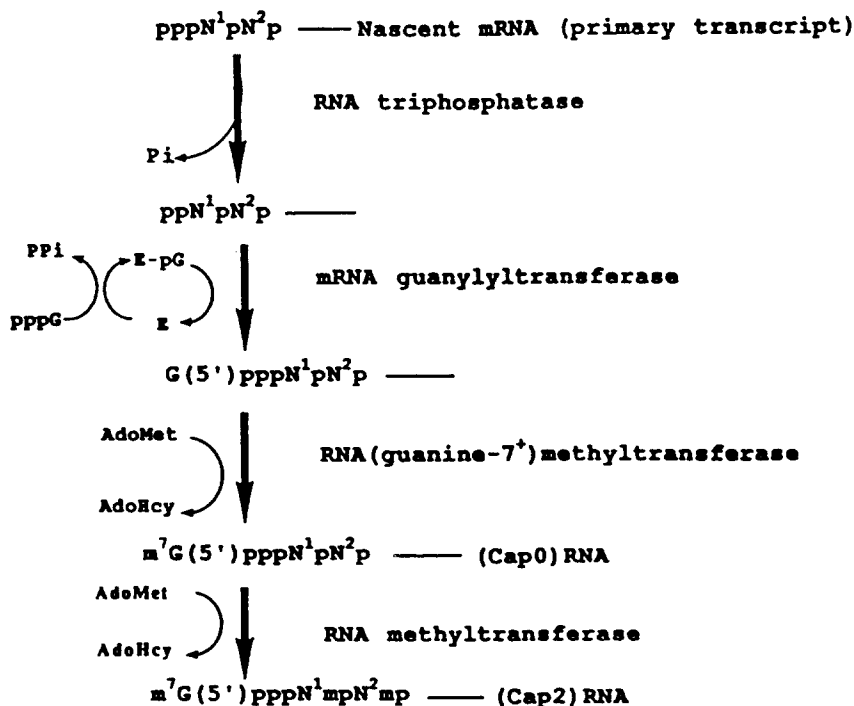


FIG 3.3. Schematic representation of the mechanism of cap formation.

phorylated end (ppNpNp), which is then converted to GpppNpNp by transfer of GMP derived from GTP by mRNA guanylyltransferase (GTase). In mammalian cells, RTase and GTase activities are on the same capping enzyme polypeptide of about 65 kDA, and the GTase located in the C-terminal region binds to the phosphorylated C-terminal domain (CTD) of the RNA polymerase II largest subunit (42–45). The finding that capping enzyme can directly and selectively bind to RNA polymerase II provided a solid molecular basis for mRNA identity and the reason why only polymerase II transcripts have a cap (77).

During the capping process GTase in the transcription complexes reacts with GTP and forms a covalent GMP–enzyme complex by phosphoamide linkage to the  $\epsilon$ -amino group of lysine in the signature KXDG sequence (77,78). Subsequent transfer to RNA to form terminal GpppN followed by methylation catalyzed by transferases using AdoMet as the methyl donor yields  $\text{m}^7\text{GpppN}(\text{m})$ . The  $\text{N}^7$ -methylation on guanine occurs cotranscriptionally by the action of RNA (gua-

nine-7-) methyltransferase (MTase) that also binds to GTase/RNA polymerase II complexes (79). A schematic representation of the interactions of capping enzyme, MTase, RNA polymerase II (pol II), and nascent RNA is shown in Fig. 3.4. The CTD of mammalian and yeast pol II largest subunit is phosphorylated on Ser residues in the tandemly repeated heptad motif YSPTSPS at the initiation/early elongation stage of transcription (80) and then recruits capping enzyme and MTase (42–45). Recently Ho and Shuman (81) found that GTase binds to Ser-2 or Ser-5 phosphorylated synthetic CTD peptides but not to unphosphorylated heptads, and that a synthetic two heptad repeat containing Ser-5-P stimulates enzyme–GMP complex formation. Transcript capping, as well as enzyme–GMP formation, by purified human GTase was similarly stimulated by a CTD 42 heptad repeat that was phosphorylated by incubation with transcription

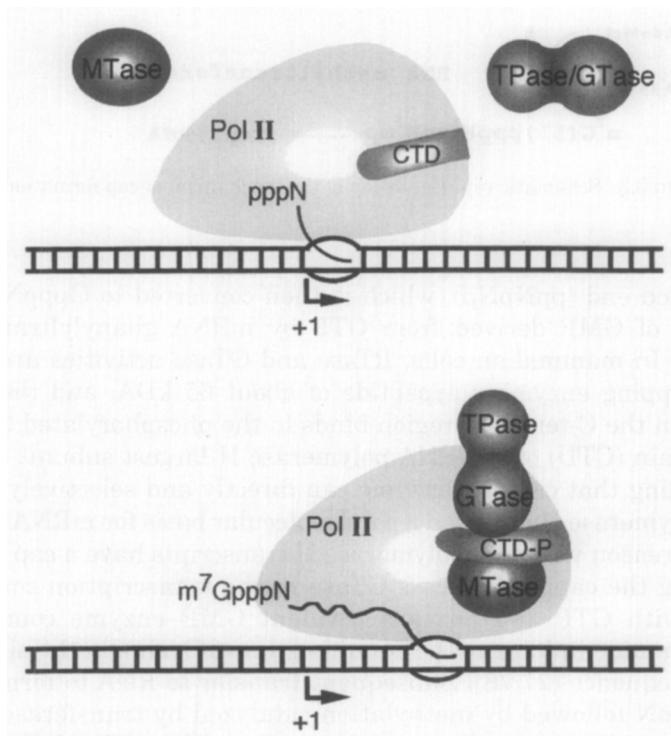


FIG 3.4. Model for recruitment of capping enzymes to the phosphorylated CTD of RNA polymerase II.

factor TFIIF (82). These results indicate that phosphorylation on Ser residues of pol II CTD regulates cotranscriptional capping.

Methylation of the first N 2'-*O*-ribose in m<sup>7</sup>GpppNpN occurs m<sup>7</sup>G cap dependently in the nucleus, whereas the second 2'-*O*-ribose methylation take places in the cytoplasm. Consistently, the 2'-*O*-ribose methyltransferase (Cap I methyltransferase) that methylates both purine and pyrimidine nucleosides in the N position of m<sup>7</sup>GpppN, and the Cap II methyltransferase involved in the second 2'-*O*-ribose methylation were found in HeLa cell nuclear and cytoplasmic fractions, respectively, and were purified by Langberg and Moss (83). However, it is unclear when the second 2'-*O*-ribose methylation take places in the cytoplasm and how or if this posttranscriptional mRNA modification is regulated. Kuge and Richter (84) reported a clue by showing that the cytoplasmic 3'-poly(A) addition (or elongation) induces ribose methylation of mRNA caps in progesterone-induced oocyte maturation in *Xenopus laevis*. Thus, posttranscriptional modifications on both the 3' and 5' ends of mRNA may be relevant to stage-specific translational stimulation of a subset of mRNAs in this system.

The detailed enzymology of mRNA capping was initially studied most extensively using purified vaccinia virus proteins. A heterodimeric 6.5S complex, consisting of 95 kDa and 31 kDa protein subunits, catalyzed the first three reactions shown in Fig. 3.3. The 95 kDa polypeptide contains both the RTase and GTase (85), whereas the 31 kDa subunit is required, but not sufficient, for N<sup>7</sup>-methyltransferase activity (78). Hagler and Shuman (86) showed that the vaccinia capping enzyme formed a binary complex with the viral RNA polymerase in solution, and the complex catalyzed viral RNA synthesis in which capping occurred only after the transcript had grown to a minimum of 27 nucleotides. Interestingly, the capping enzyme complex is also responsible for termination in vaccinia virus transcription (87), although the presence of a 5'-cap is not required for transcript termination. Vaccinia capping enzyme was also shown to be involved in the formation of a transcription initiation complex and was required for the formation of short uncapped transcripts, as well as longer capped ones, clearly indicating that viral capping machinery, but not necessarily the capping reaction itself, is coupled to the initiation of transcription (46). Vaccinia virus provides an example of a protein that affects both 5' and 3' end processing: the 5' cap-specific viral 2'-*O*-methyltransferase (VP39) was found to be a processive factor for the vaccinia-encoded poly(A) polymerase (VP55) and stimulated mRNA 3'-polyadenylation (88). VP39 catalyzed m<sup>7</sup>G-cap dependent 2'-*O*-methylation of substrates longer than m<sup>7</sup>GpppNp (89).

As described previously, CPV forms capped mRNA by virus-associated RNA polymerase and a type 1 capping mechanism like vaccinia virus and reovirus (75). A unique feature distinguishing CPV from other viral transcription systems is that CPV mRNA synthesis is greatly stimulated *in vitro* by AdoMet, the natural methyl donor for methylating caps (24). Further studies showed that the stimulatory effect of AdoMet resulted from lowering of the  $K_m$  for the initiating ATP so that the viral RNA polymerase readily forms the first dinucleotide, pppApG. The kinetic studies indicated that AdoMet and its structural analogs stimulate RNA polymerase apparently by allosteric effect(s) on the CPV transcription apparatus, perhaps by binding to the MTase associated with viral RNA polymerase/GTase complexes (48, 49). These findings prompt us to speculate that the apparatus involved in the type 1 capping may affect the initiation of cellular RNA synthesis by polymerase II via protein-protein interaction(s), as proposed previously for CPV "pretranscriptional capping" (47). A similar allosteric effect was found in influenza virus where free  $m^7GpppNm$  stimulated *in vitro* transcription but was not incorporated into mRNA (90). Recognition of cap by influenza transcription complexes and their unique cap-snatching mechanism is discussed in a later section.

### B. Genes and Enzymes Directing Cotranscriptional Capping

The genes of enzymes involved in the individual steps of capping have been cloned and characterized from various sources. In vaccinia virus, three enzymatic activities (TPase, GTase, and MTase) are contained in the 130 kDa heterodimeric complex consisting of 95 kDa and 31 kDa subunits encoded by the D1 and D12 viral genes, respectively. Biochemical and genetic analyses indicated that the domains involved in the three activities are in the 95 kDa polypeptide with RTase at the N-proximal region, GTase in the middle of the molecule and MTase at the C-proximal region (38,91–93) as schematically shown in Fig. 3.5. The 31 kDa protein associates with the MTase domain and stimulates MTase activity (94). The 130 kDa binary complex binds to viral RNA polymerase and directs capping cotranscriptionally. The vaccinia cap 2'-O-methyltransferase activity was purified by Barbosa and Moss (95) as the viral 39 kDa protein, which was later found to be a subunit of the heterodimeric poly(A) polymerase (96). VP39 binds specifically to capped RNA, and the structure of this complex has been solved by X-ray crystallography (97).

In contrast to vaccinia virus, the enzymes that catalyze yeast mRNA capping are encoded by three separate genes, all essential for cell via-

Sources	Gene	Protein Structure	Mol. Wt. (kDa)
Vaccinia virus	D1	N- TP ase GT ase MT ase -C	95
	D12	N- p31 -C	31
Yeast ( <i>S.c.</i> )	CET1	N- TP ase -C	80
	CEG1	N- GT ase -C	52
	ABD1	N- MT ase -C	50
<i>C. elegans</i>	CEL1	N- TP ase GT ase -C	66
	C25A1.f	N- MT ase -C	46
Human ( <i>HeLa</i> )	HCE1	N- TP ase GT ase -C	69
	CMT1	N- MT ase -C	55

FIG 3.5. Cloned capping enzyme genes, protein products, and activity domains.

bility or growth: RTase (80-kDa) by *ScCET1* (98), GTase (52 kDa) by *ScCEG1* (99), and MTase (50-kDa) by *ABD1* (100). *ScCEG1* and *ABD1* genes are both essential for viability, and gene disruption experiments have shown that *ScCET1* is essential for growth (101). The yeast capping enzyme exists as a heterodimeric, bifunctional complex of GTase and RTase subunits. The site on GTase that forms a covalent enzyme-GMP intermediate was shown to be lysine in the KTDG motif that is also present in GTases from vaccinia virus, Shope fibroma virus, African Swine fever virus (38) and in a homolog in *Candida albicans* (102). Similar motifs occur at the active sites of T4 RNA ligase (KEDG), mammalian DNA ligase (KYDG), and yeast tRNA ligase (KANG) (103). Mutational studies on the *CEG1* gene and its products indicated that the lysine residue (Lys-70) in the motif is conserved among capping enzymes (104). The yeast MTase was found to have a catalytic domain similar to the C-terminal segment of the vaccinia virus D1 gene product (105). In *C. albicans*, the *CaCET1* and *CaCEG1* genes coding for TPase (59kDa) and GTase (52kDa), respectively, were cloned (102, 106), and *CaCET1* was shown to rescue *ScCET1*-deficient *Saccharomyces cerevisiae*.

The first metazoan gene coding for a capping enzyme (*CEL1*) was obtained in *C. elegans* (107), and shortly after the human and mouse



species were reported (44). The *CEL1* protein consists of 574 residues, with a C-terminal domain highly related to yeast and vaccinia virus GTase. Wang *et al.* (108) also identified the capping enzyme gene of *C. elegans* based on motif sequences homologous to GTases from vaccinia virus and *Chlorella* virus. Mammalian capping enzymes and *CEL1* all contain RTase in their N-terminal domains, which include an essential cysteine in a sequence motif that is characteristic of the active site of protein tyrosine phosphatase family members. The expressed 1-236 amino acids of *CEL1* has RTase activity that removes the  $\gamma$ -phosphate group from 5'-triphosphorylated RNA, although it remains to be tested if this activity leaves  $\alpha$  and  $\beta$ -phosphate groups in the substrate RNA like the other triphosphatases involved in type 1 cap formation. Another *C. elegans* gene, C25A1.f, encodes an ABD1-like 402 residue protein, which may be the worm gene for RNA (guanine-7-) methyltransferase (79,108). The GTase genes of *Crithidia fasciculata* and *Trypanosoma brucei* have recently been characterized (109), and *Artemia salina* capping enzyme consisting of a single 73kDa polypeptide has been shown to contain both RTase and enzyme-GMP forming GTase activities. The TPase and GTase activities were shown to be separable into 20 kDa and 40 kDa polypeptide domains, respectively, by limited tryptic digestion (110), providing an early clue that metazoan capping enzymes are bifunctional, two-domain proteins (64).

Recently cDNAs for human capping enzymes (HCE) have been cloned from HeLa and adenocarcinoma cell libraries (44,111,112). HCE1 (also referred to as hCAP1 or RNGTT) is a 597-residue polypeptide with both RTase and GTase activities. Expression studies and amino acid sequence homology searches indicate that they are contained in N-terminal and C-terminal domains, respectively (111, 44,45). The human MTase was earlier shown to be a separate 55 kDa protein, and the CMT (or RNMT) cDNA encoding it has been cloned (79,113). Finally, the human capping enzyme and MTase genes were mapped to chromosome 6q16 and 18p11.22-p11.23, respectively (114). Capping enzyme cDNAs of yeast *Schizosaccharomyces pombe* (PCE1, reference 102) and mouse (MCE1, reference 44) have also been sequenced. MCE1 is closely homologous to the HCE1 and can complement the growth of haploid *S. cerevisiae* cells lacking endogenous GTase, indicating a functional conservation of eukaryotic mRNA capping apparatus from yeast to mammals (44). In addition, recombinant full-length MCE1 and its active C-terminal fragment bound to the elongating form and not to the initiating form of RNA polymerase,

indicating that the phosphorylated CTD of RNA polymerase II couples capping to transcription elongation and accounts for the selective capping of RNA polymerase II transcripts (42–45).

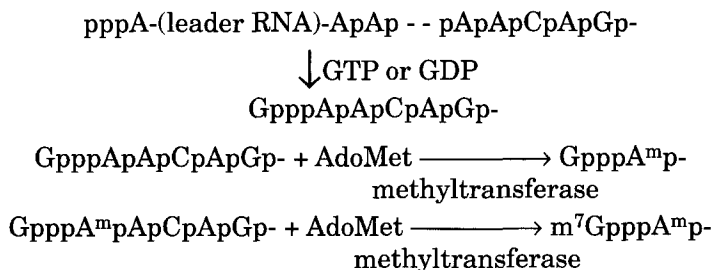
In reoviruses, the genome dsRNA segments L1, L2, and L3 were found to encode RNA polymerase, GTase/MTase, and RTase, respectively (115–118). Baculovirus *Autographa californica* nuclear polyhedrosis virus contains both RTase and GTase in the 464 residue LEF-4 gene product (119–121). Bluetongue virus (BTV), another *Reoviridae* also contains both GTase and RTase in viral polypeptide VP4 expressed in and purified from insect cells. Curiously, BTV VP4 also contained inorganic pyrophosphatase activity (122). In still another important variation on the capping theme, the nsP1 proteins of positive-stranded RNA alphaviruses, Semliki Forest and Sindbis viruses, catalyze both the GTase and MTase reactions (123–126). However, in each case in the presence of AdoMet they form an enzyme–m<sup>7</sup>GMP complex in addition to an enzyme–GMP complex, suggesting that in alphavirus mRNA capping the guanine is methylated before linkage to the mRNA molecule (125). This mechanism was later confirmed (126) and clearly provides a potential new target for developing novel antivirals.

Many if not most viral enzymes apparently evolved away from cellular and other viral capping enzymes despite similar enzymatic functions, and the active site sequences of viral GTases are probably diverged from the canonical signature motif found in cellular and poxvirus enzymes (127). The smallest GTase was identified in a *Chlorella* virus PBCV-1 polypeptide (330 amino acids) that catalyzes GMP transfer to RNA but not cap methylation or  $\gamma$ -phosphate cleavage (128). The crystal structure of monofunctional PBCV-1 GTase has been solved with GTP bound at the active site and with GMP bound covalently (129). The enzyme consists of two domains separated by a cleft, and the structure resembles T7 DNA ligase. The cleft observed in PBCV-1 GTase, however, is much deeper and narrower than in DNA ligase, which may account for the substrate specificity of capping enzyme for single-stranded RNA molecules. The enzyme structure undergoes a large conformational change from an open to a closed state after binding to manganese ions, which may facilitate the hydrolysis of bound GTP to form a stable enzyme–GMP complex (129).

### C. Vesicular Stomatitis Virus mRNA Capping: Type 2

The virus-associated RNA polymerase of the rhabdovirus vesicular stomatitis virus (VSV) synthesizes five mRNA species, all with the 5'

sequence GpppApApCpApGp (4) and quantitatively methylated to  $m^7GpppA^m pApCpApGp-$  in the presence of AdoMet. VSV contains an 11 kb single- and negative-stranded genomic RNA encoding N, P, M, G, and L proteins. The monocistronic transcripts, which are 3' polyadenylated as well as capped, are made sequentially from the N to L genes colinearly arranged on the template genomic RNA, after synthesis of a short leader RNA containing a 5'-triphosphate terminus (4,64). Although the cap is identical in VSV and cellular mRNAs, the reactions catalyzing cap formation are notably dissimilar. First, the  $\alpha$ -phosphate of the initiating ATP, and the  $\alpha$ - and  $\beta$ -phosphates of GTP are incorporated into the cap structure, suggesting that the  $\beta$ -phosphate of GTP is linked to the 5' mono- (or poly-) phosphorylated end of VSV pre-mRNA or is involved in a phosphorolysis of pre-mRNA at specific sites in nascent pre-mRNA during chain elongation. It remains to be clarified how the GDP transfer is made to the 5' sequence pAACAGUAAUC-, although it is certain that cap formation occurs co-transcriptionally. Second, unlike cellular GTases, the VSV enzyme(s) does not form a covalent enzyme-GMP intermediate, and thus exogenously added ppA-containing RNAs cannot be capped. Third, the order of the methylation reaction differs from cellular and many other viral mRNAs; here the 2'-O-ribose methylation occurs in the initial nucleotide (A in GpppA), and the guanine-7 methylation follows as shown below. Fourth, exogenously added GpppA or other blocked dinucleotides cannot be methylated, in contrast to the cap methyltransferases of vaccinia virus, reovirus and various cells, indicating that VSV RNA methylation is very tightly coupled to nascent mRNA synthesis in the following reactions:



The GTase involved in this process has not yet been identified, but VSV RNA polymerase L protein expressed in insect Sf21 (*Spodoptera frugiperda*) cells was recently shown to be associated with cellular protein synthesis elongation factor EF-1 $\alpha\beta\gamma$  and packaged in purified virus

grown in Sf21 cells (130). This is strikingly similar to phage Q $\beta$  replicase which also requires multiple host-encoded proteins as subunits. The insect cell guanylyltransferase, LEF-4 gene product was also found to be associated with the expressed VSV GTase and capable of capping uncapped transcripts *in vitro*, although the insect GTase forms a canonical guanylate-enzyme complex (131). While the VSV mechanism needs to be reconsidered, the type 2 cap formation initially found in VSV transcription may alternatively be part of a processing scheme to generate monocistronic mRNAs from an otherwise long polycistronic precursor RNA. A new model, constrained RNA structure-mediated capping, proposed by Shuman (132), may resolve the enigma that research on VSV mRNA capping has been facing for many years.

#### D. Formation of TMG cap and Mppp cap in snRNAs

Small nuclear RNAs (snRNAs) contain 5' structures,  $^{2,2,7}\text{m}_3\text{GpppN}$  (TMG cap) and Methyl-pppN (Mppp cap) (5,6). Several lines of evidence suggest that RNA polymerase II is responsible for the synthesis of TMG-capped mammalian U1, U2, and U3 snRNAs, including the finding that their synthesis is inhibited by low concentrations of  $\alpha$ -amanitin or 5,6-dichloro-1- $\beta$ -D-ribofuranosylbenzimidazole which are specific inhibitors of pol II transcription (133,134). Indeed, it was shown that the U1 primary transcripts are capped with  $\text{m}^7\text{G}$  cotranscriptionally, like most pre-mRNAs (135,136). Thus, the mechanism involved in TMG cap formation is assumed to be like type 1 cap formation on cellular mRNA. In U1 snRNA, the Cap 1 structure,  $\text{m}^7\text{GpppAmpN}$ -, is formed in the nucleus, and the RNA is then transported to the cytoplasm where additional methylations occur to complete the  $^{2,2,7}\text{m}_3\text{GpppAmpNmp}$ - (137). U1 RNA with its multiple methyl groups binds Sm antigen at a downstream AUUUUUG sequence, and the resulting ribonucleoprotein is transported back to the nucleus to play a role in mRNA splicing (138,139). However, U3 RNA is not exported to the cytoplasm but is retained in the nuclei of *Xenopus laevis* oocytes, where it undergoes 5' cap trimethylation (140). The nuclear retention and maturation of this RNA raises the possibility that other TMG-capped RNAs are also retained and modified in the nucleus. Because of multiple methylations on the capping G, it is likely that the nuclear methyltransferase that yields TMG cap is different from the monomethylating RNA-(guanine-7)methyltransferase. Alternatively, other methyl transferase(s) that cooperate with RNA-(guanine-7)methyltransferase and methylate the 2-amino group of guanine may

be included in the transcription initiation complex in a U1-U3 promoter sequence-specific manner.

Mammalian U6, 7SK, B2 RNAs, and the plant U3 RNA, which contain a Mppp cap structure, appear to be synthesized by RNA polymerase III because their synthesis is not inhibited by low concentrations of  $\alpha$ -amanitin but by tagetitoxin, a specific inhibitor of RNA polymerase III at low concentrations (141). Mppp cap formation occurs post transcriptionally and is very different from the mechanisms that have been characterized for mRNAs. U6 Mppp cap is apparently formed by enzymes unrelated to those involved in mRNA capping since the  $\gamma$ -phosphate of U6 snRNA is not removed during cap formation but is directly methylated by presumptive nuclear RNA sequence-dependent methyltransferase. A 5' proximal stem-loop structure seems to be necessary for U6 snRNA Mppp-capping. The enzyme that directs such capping reactions on the human U6 small nuclear RNA has been purified and characterized by Shimba and Reddy (142). Mppp-capped U6 RNA stays in the nucleus associated with La antigen, which may be a transcription termination factor for RNA polymerase III. Further studies are needed to understand better the biological significance of methyl-capping on snRNAs. This unusual cap structure appears to protect RNA from exonucleolytic degradation.

In addition to m<sup>7</sup>G and 2'-O-methyl groups, N<sup>6</sup>-methyladenosines (N<sup>6</sup>mA) are present in cellular and several virus mRNAs, on average one N<sup>6</sup>mA per 1000 bases (1,4,64). N<sup>6</sup>mA affects the efficiency of mRNA processing and transport, although the mechanism(s) by which this occurs is not known. The gene coding for the methyltransferase involved in methylating internal A of human mRNA has recently been identified by Bokar *et al.* (143).

#### V. PREFORMED CAPPED 5' END ADDITION TO CELLULAR AND VIRAL MRNAs

The two types of cap formation first described apparently are associated with nascent RNA synthesis in a tight cotranscriptional manner. However, there are other cases in which cells or viruses use a preformed cap to generate 5' capped mRNAs. The biological significance of these phenomena remains unclear, but one possibility is that the limited genomic capacity, for example of small viruses, is enhanced by use of a capped 5' sequence from host cell transcripts or a preformed cap-oligonucleotide for mRNA synthesis.

### A. Capped mRNA Formation by Trans-Splicing in Nematodes and Trypanosomes

One-fourth of the genes in the nematode *Caenorhabditis elegans* are in operons (i.e., polycistronic transcription units arranged with the genes only 100–400 bp apart). The resulting pre-mRNAs are processed into monocistronic mRNAs by a combination of cleavage and polyadenylation at the 3' end of the upstream mRNA and *trans*-splicing of a 22-base TMG-capped RNA (SL RNA) to the 5' end of the downstream mRNA (Fig. 3.6A) (144). The SL RNA is derived from the 5' end of a precursor RNA of approximately 100 nucleotides, called spliced-leader RNA, that is transcribed by RNA polymerase II from a 1 kb DNA repeat, which also encodes 5S ribosomal RNA (145). A subset of mRNAs in *C. elegans* and the parasitic nematode *Ascaris lumbricoides* var. *suum* both acquire SL by the *trans*-splicing mechanism shown in Fig. 3.6 (see also color insert). Correct 3' end formation of the upstream message in an operon seems to be required for specific downstream mRNA *trans*-splicing of SL RNA. By contrast, *trans*-splicing of the downstream mRNA *per se* is not necessary for correct 3' end formation of the upstream mRNA, suggesting that SL RNA mediated *trans*-splicing that adds a capped 5' oligomer is coupled mechanistically to 3' end formation in the processing of nematode polycistronic pre-mRNAs (146, 147). Similar *trans*-splicing also occurs in trypanosomes (148, 149) and *Euglena* (150). The SL RNA of *Trypanosoma brucei* and *Crithidia fasciculata* has a cap 4 structure  $m^7Gpppm_26AmAmCmUm$ , and the ribose methyl groups appear to be obligatory for SL RNA utilization in *trans*-splicing.

### B. Cap Snatching in Influenza Virus Transcription

Synthesis of influenza mRNAs involves a novel priming mechanism with initiation by host cell  $m^7GpppNm$ -containing RNA fragments derived from RNA polymerase II transcripts (66). This occurs in the nucleus of infected cells, and viral mRNA synthesis requires continuous function of RNA polymerase II. Consequently it is inhibited by  $\alpha$ -amanitin. Capped RNA primers are generated by a viral cap-dependent endonuclease that cleaves capped cellular RNAs 10–13 nucleotides from their 5' ends, preferentially at a purine residue (151–153). In this unique cap-snatching mechanism, the capped oligonucleotides are used as primers by the virion-associated RNA-dependent RNA polymerase to initiate transcription of the single-stranded negative genomic RNA (vRNA). The viral polymerase

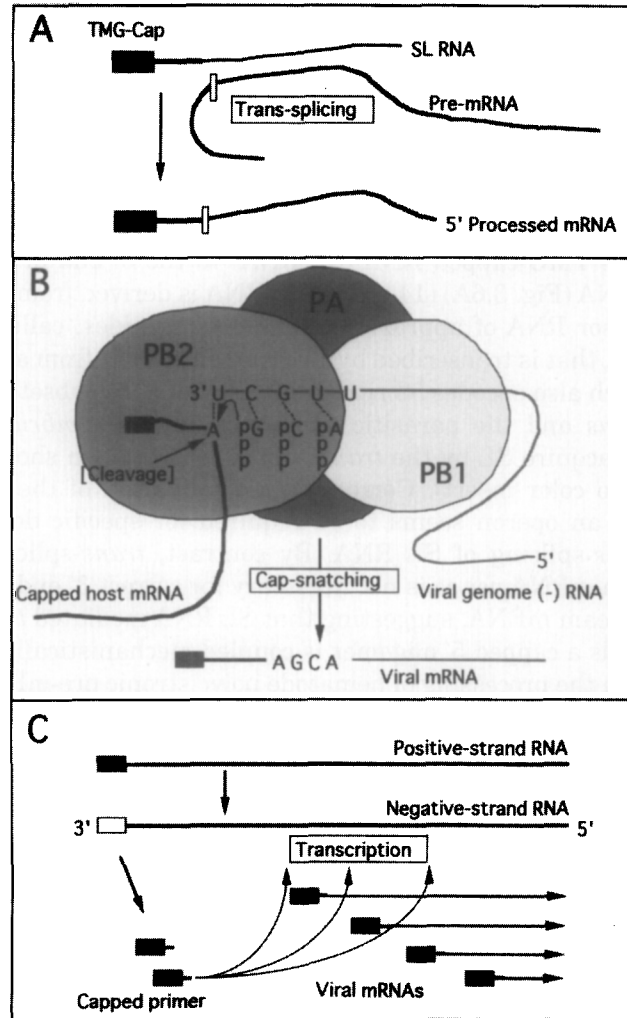


FIG 3.6. Schematic representation of the mechanism for adding preformed 5' cap sequences to cellular and viral mRNAs. (A) *Trans-splicing* in nematodes; (B) cap-snatching in influenza virus; (C) possible capped leader priming in coronavirus transcription. The horizontal boxes with red or black indicate capped oligonucleotides. See also color insert.

can efficiently bind capped RNAs shorter than 9 nucleotides but cannot utilize them or free caps as primers. During the generation of primers by endonucleolytic cleavage of capped host RNA, flu polymerase consisting of three P proteins (PB1, PB2, and PA) forms a

complex that recognizes three signals on capped RNA: cap at 5' termini, RNA sequences at the cleavage sites, and distance between these two signals (153). The PB2 protein in the polymerase complex recognizes and binds to the cap structure of host cell heterogeneous nuclear RNA and leads to endonucleolytic cleavage of the capped primer downstream of the cap (66,153). Polymerase complexes require both the 5' and 3' ends of genomic vRNA for the endonuclease activity, and the choice of cleavage site is controlled by the RNA sequence and does not depend on direct interactions with the vRNA template (154). RNA modification and electron microscopic analyses performed on native flu RNPs confirmed that the polymerase forms a complex with both conserved vRNA ends (155). The resulting capped oligonucleotides are used as primers for viral mRNA synthesis by the virion transcriptase (Fig. 3.6B). The PB1 protein, which is most likely the transcriptase, adds a G residue from GTP to the 3' end of primer and then elongates the chain. Priming does not seem to require hydrogen bonding between the capped primer fragments and the 3' ends of the vRNA template.

Initially, the *in vitro* transcription of influenza virus by virus-associated RNA polymerase was stimulated by the dinucleotides ApG and GpG, and these dinucleotides were incorporated into RNA products, indicating that influenza virus polymerase has difficulty initiating transcription without primers (66). The findings that a free cap structure such as m<sup>7</sup>GpppNm (and each of the methyl groups in the cap) can stimulate viral transcription but not be incorporated into viral mRNAs *in vitro* suggests that the cap structure stimulates viral RNA polymerase by allosteric modulation (156). The molecular mechanism(s) behind this allosteric stimulation by caps remains to be defined. However, PB1 has recently been shown to bind sequentially to 5' and 3' sequences of vRNA at different regions, which causes the polymerase complex to undergo two sequential alterations in conformation that activate binding to capped RNA followed by activation of the endonuclease (157). Perhaps, a similar allosteric conformational change occurs in PB1 in the P protein complex that stimulates RNA elongation after PB2 binds and "snatches" caps from host RNAs. The tick-borne Thogoto virus (THOV), another *Orthomyxoviridae*, also produces capped mRNA by cap-snatching, but THOV transcription is initiated preferentially by 5' m<sup>7</sup>GpppAm-containing sequences (158). A novel anti-influenza virus compound that specifically blocks the cap-dependent endonuclease of the viral transcriptase, and is inhibitory to the replication of both influenza A and B viruses, has been identified as a possible antiviral drug (159).



*C. Possibility for Capped Leader-Primed Transcription in Coronavirus mRNA Synthesis*

The human coronaviruses contain a positive-sense, single-stranded RNA genome ranging from 27.6 to 31 kb, which functions as mRNA and is also infectious. The virus replicates in the cytoplasm, and the genomic RNA is capped. It contains 7–10 functional genes, 4 or 5 encoding structural proteins. The genes are arranged in the order 5'-RNA polymerase-HE(hemagglutinin-esterase glycoprotein)-S(spike glycoprotein)-E(small membrane protein)-M(membrane glycoprotein)-N(nucleocapsid protein)-poly(A)-3'. Several other genes that code for nonstructural proteins are located between the structural protein genes. Capped mRNAs of various sizes encoding these proteins are made from negative-stranded template RNA(s) by the RNA polymerase, which apparently has associated capping enzymes. Interestingly, the individual mRNAs share a 5' capped leader sequence of approximately 60–90 nucleotides, which is derived from the 5' end of genomic RNA (160). At the mRNA start sites on the viral genomic RNA is present a short stretch homologous to the 3' end of the leader RNA. This sequence constitutes part of the signal for subgenomic mRNA transcription. Sequence comparison of viral genomic RNA and mRNA suggests that subgenomic mRNAs are derived by fusion of the 5' capped leader to the mRNA start sites on the viral genomic RNA as illustrated in Fig. 3.6C. These observations indicate that capped 5' leader sequence may be used as primer for initiation of individual coronavirus mRNAs, although possibilities remain for posttranscriptional *trans*-splicing between leader RNA and pre-mRNA, or of a polymerase jumping during minus strand synthesis to generate various subgenomic template RNAs that all contain the same leader template sequence at the 3' end (160).

## VI. CAP FUNCTIONS

Since its discovery, a variety of functions have been envisaged for the cap including (1) involvement in the initiation of mRNA biosynthesis, (2) stabilization of nuclear pre-mRNAs and cytoplasmic mRNAs, and (3) facilitation and regulation of mRNA translation. Based on a large amount of work in the past 25 years, these roles have been demonstrated as will be reviewed in the following sections. Other, related biological effects of mRNA cap raised later included (4) stimulation of pre-mRNA splicing, (5) nucleocytoplasmic transport, (6) developmental regulation in certain cells and organisms, and (7) stim-

ulation of 3'-poly(A) addition. Identifying the factors that interact with mRNA caps to mediate these important biological events should lead to new and important insights into how eukaryotic gene expression is regulated.

*A. Protection of Pre-mRNA and mRNA Against  
5' Exonucleolytic Attack*

In early biochemical studies it was shown that capped mRNAs are more stable than their uncapped counterparts (7,8). Uncapped reovirus mRNAs with 5' ppG-ends were readily digested to 5'-monophosphates when microinjected into frog oocytes, while the mRNAs with 5' m<sup>7</sup>GpppGm- or GpppG- were relatively stable. This preferential digestion of uncapped mRNA is due to the 5' exonucleases, in both cytoplasm and nucleus, that digest RNAs containing free 5' ends to 5' mononucleotides, in a 5'→3' direction, exonucleolytically and processively (7). These observations clearly indicated that cap helps to protect RNA from degradation by 5' exonucleases, and such nucleases occur in *S. cerevisiae* (161,162) and also in mammalian cells (163,164). The yeast 5' exonuclease gene *XRN1* encodes a 175 kDa protein, and its disruption restricts yeast growth (165,166). Using *XRN1* mutant cells, Lo *et al.* (167) reported a series of elegant experiments in which they placed *HIS4* under transcriptional control of the pol I ribosomal DNA promoter/enhancer region and compared the stability and translational activity of the resulting uncapped *HIS* transcripts with those of the inherent capped *HIS* mRNAs under conditions in which both types of mRNA were 3'-polyadenylated *in vivo*. The majority of capless *HIS4* mRNAs (*rhis4* mRNA) were degraded rapidly despite the *rhis4* mRNA being more actively transcribed than the wild-type capped *HIS4* gene transcripts made by pol II. Furthermore, in *XRN1* mutant yeast that lacked 5' exonuclease activity, the steady-state level of *rhis4* mRNA increased by 10-fold compared with the wild-type *HIS4* mRNA level, indicating that the cap confers stability to mRNA and 5' exonuclease scavenges uncapped mRNA.

Capped mRNA turnover likely is facilitated by decapping enzymes that cleave mRNA cap m<sup>7</sup>G-dependently and have been detected in extracts prepared from various cells (64, 168, 169). One example, the yeast *DCP1* product (Dcp1p), has been studied extensively by Parker's group (170–176). Dcp1p acts preferentially on capped RNA that contains a shortened 3'-poly(A) tail, and it can be inhibited by uncapped RNA or poly(A). Thus, a major pathway of mRNA degradation in yeast is initiated by poly(A) shortening that triggers decapping, thereby

exposing the mRNA to 5'→3'-exonucleolytic degradation. Additional strong genetic evidence was provided by Schwer *et al.* (177) who used temperature-sensitive CEG1 (capping enzyme gene)-bearing *S. cerevisiae* to examine mRNA degradation processes. They found that a shift of CEG1 mutant cells to restrictive temperature elicited a sharp reduction in the steady-state levels of multiple individual mRNAs. Also, the CEG1 mutations prevented the accumulation of mRNAs that were newly synthesized at the restrictive temperature, whereas uncapped poly(A)-containing mRNAs accumulated in cells lacking the 5'-exonuclease *Xrn1*, consistent with mRNA turnover by deadenylation-decapping-exonucleolytic degradation, where the cap plays a critical role. Similar observations that decapping and exonucleolytic digestion are preceded by a shortening of the poly(A) tail were obtained in mammalian cells (178). Besides poly(A) shortening-dependent decapping, an early, nonsense codon-mediated decapping pathway also has been demonstrated (179,180). This pathway resulted in a striking reduction in the defective mRNA in cells of patients with various genetic diseases including Werner and Rothmund–Thomson syndromes, recent findings that remain to be explored further (181,182).

### *B. Splicing and Nucleocytoplasmic Transport*

It was demonstrated early that the presence of the 5' cap structure increases both the accuracy and efficiency of pre-mRNA splicing in mammalian cell extracts (183–185) and intact *Xenopus* oocytes (186). The cap on pre-mRNA interacts with splicing components at the adjacent 5' splice site for processing the first exon and removing the first intron (185). This interaction is mediated by a nuclear cap-binding complex (CBC) consisting of two tightly associated proteins, CBP 20 and CBP 80 (187). The monomethylated 5' cap on pre-mRNA is specifically bound by CBC, which facilitates association of the cap-proximal 5' splice site with U1 snRNP that includes the TMG-capped U1 RNA. The CBC remains bound to the cap during RNA processing and plays an active role in both splicing and RNA export (188). The CBC-mediated interaction between U1 snRNP and the 5' splice site is one of the earliest steps in spliceosome assembly and is conserved in humans and yeast (189). Spliceosomes consist of U1, U2, U4/U6, and U5 snRNA particles together with a large number of non-snRNP splicing factors, most of which remain to be characterized.

The CBC/capped RNA complex also participates in 3'-polyadenylation and nucleocytoplasmic transport. As shown by Flaherty *et al.* (190), depletion of CBC from HeLa cell nuclear extract strongly

reduces the endonucleolytic cleavage step of the polyadenylation process. The overall structure of pre-mRNA is [5' cap – 5' splice site] – intron – [3' splice site – exon – 5' splice site] – intron – [3' splice site – polyadenylation site], and the end structures of exons differ depending on their locations in pre-mRNA: the first exon is associated with cap and 5' splice site, the interior exons are between 3' and 5' splice sites and the last exon is accompanied by 3' splice site and polyadenylation site (191,13). Thus, exon definition (or recognition) must be different for interior and the two terminal exons, but a shared mechanism must coordinate and regulate the entire RNA processing scheme. Perhaps factors that interact with the 3' splice site share features similar to CBC and also enhance polyadenylation. In this context, it is intriguing that both the cap and the 3' splice site similarly affect polyadenylation efficiency in HeLa cell extract (13). By contrast, genetic data obtained with *S. cerevisiae* containing temperature-sensitive alleles of *CEG1* (encoding the capping enzyme GTase subunit) indicated that cap enhances but is not essential for mRNA splicing (192). Further studies, including molecular genetic approaches, are needed to identify the structural basis for the recognition of cap/CBC, interior exons, and polyadenylation site, as well as the factor(s) involved in individual interactions. Also, the role of cap in transport has not been well characterized, even though CBC remains bound during mRNA transit from the nuclear transcription site. Several lines of evidence suggest that cap and CBC are not required for nucleocytoplasmic transport (193), although the cap stabilizes mRNA during the exit pathway. Microinjected anti-CBP20 antibodies did not alter mRNA export from *Xenopus* oocyte nuclei, whereas they completely blocked TMG-capped U snRNA export (194). Moreover, yeast strains carrying conditional mutant *CEG1* alleles, or disrupted genes encoding yCBP80 and yCBP20, did not show a defect in mRNA nuclear export under restrictive conditions (195).

### C. Translational Effects

Reovirus with its associated RNA polymerase and capping enzymes also proved extremely useful for deciphering how eukaryotic mRNAs initiate protein synthesis. Comparisons of reovirus RNAs containing various types of 5'-ends indicated that molecules with 5'-terminal m<sup>7</sup>GpppGm are preferentially translated in extracts of either plant or animal cells (196,197). Selection of capped over uncapped RNAs occurred during binding to small ribosomal subunits (198). Other observations also supported the idea that caps are recognized during

early stages of protein synthesis. For example, initiation complex formation was inhibited by cap analogs like m<sup>7</sup>Gp (199, 200), and a protein that bound to capped but not uncapped oligonucleotides was detected in salt washes of brine shrimp embryo ribosomes (201). Direct evidence was provided by the observation that the cap in reovirus mRNA comprises part of the ribosome binding site in ~48S initiation complexes as demonstrated by cap retention in the sequences protected against RNase digestion. Furthermore, the protected capped mRNA fragments rebound efficiently, whereas only low levels of ribosome rebinding were observed with the corresponding uncapped sequences that were collected from RNase-treated 80S complexes (202). These findings on the sizes, sequences, and initiation properties of 40S- vs. 80S-protected reovirus mRNAs provided insights leading to the scanning model of eukaryotic translation initiation (203,204).

The unusual 5'-5' linkage and resulting free *cis*-diol in the cap m<sup>7</sup>G ribose moiety made it possible to demonstrate that mammalian protein synthesis initiation factor preparations contain a polypeptide of molecular weight ~24 kDa that can be specifically cross-linked to the cap in oxidized mRNA (205). This cap binding protein (CBP) was purified by affinity chromatography on Sepharose-coupled m<sup>7</sup>GDP and shown to stimulate capped mRNA translation (206,207). Translation of picornavirus uncapped RNA was not increased by CBP, suggesting that these viral RNAs initiate translation by a cap-independent mechanism, as later documented (208). Although capped host mRNAs remained intact, their translation was shut off in poliovirus-infected cells and in infected cell extracts. Translation of capped mRNA could be restored by adding to the extracts a protein complex consisting of the 24 kDa CBP (later named eIF4E), initiation factor eIF4A (an ATPase/RNA helicase), and a large polypeptide (P220, later called eIF4G, 209). This heterotrimeric complex also proved to be an initiation factor required for globin mRNA translation in a reconstituted reticulocyte system and was named eIF4F (210). In the 4F complex, eIF4G was found to be a critical "bridging" polypeptide containing an N-terminal eIF4E binding region, a central domain that associates with eIF3 (and thus to 43S ribosomes), and other sequences that interact with eIF4A and poly(A)-binding (or related) protein(s) (211-214). In cells infected with poliovirus or some other picornaviruses (208), proteolytic removal of the N-terminal domain of eIF4G (and its functional homolog, eIF4GII, 215) disrupted the ribosome adapter function for capped mRNA, resulting in replacement of cap-dependent initiation by cap-independent viral protein synthesis. Thus, many diverse findings support a perhaps oversimplified model involving a sequence

of events in which the mRNA capped end is initially recognized by the 4E component of eIF4F; mRNA binding to ribosomes is facilitated by association of the 4G component of eIF4F with eIF3 on small ribosomal subunits; RNA unwinding and subunit scanning to the initiator AUG is promoted by the eIF4A RNA-dependent ATPase/helicase activity; and multiple rounds of translation are facilitated by eIF4G/poly A-binding protein interactions that increase the proximity of the 5'-capped and 3'-polyadenylated mRNA ends (216–220). As predicted for a multisubunit, key regulator of protein synthesis, eIF4F functions are apparently controlled by a myriad of reactions in addition to proteolysis of the eIF4G component exemplified by picornavirus infection. Mechanisms proposed for regulating eIF4E include activation of transcription, phosphorylation, dephosphorylation, and binding to inhibitory proteins. Many studies on the effects of altering cap binding proteins on translation and other cellular processes including growth potential have been summarized in a recent comprehensive review article (221).

## VII. CONCLUSION

It has been 25 years since the discovery of the cap on eukaryotic mRNAs. During this time much has been learned, and in this review we have attempted to give a broad overview of the subject. We apologize to the many investigators whose work could not be cited because of space limitations. In addition to elucidating the biochemical mechanisms of capping and the downstream effects of this 5'-modification on gene expression, the advent of gene cloning has made available an ever-increasing amount of information on the proteins responsible for producing caps and the functional effects of other cap-related interactions. Genetic approaches have demonstrated the lethal consequences of cap failure in yeasts, and complementation studies have shown the evolutionary functional conservation of capping from unicellular to metazoan organisms. Although deletion of capping enzyme(s) will likely result in early lethality in knockout mice, it should be possible to study phenotypes in conditional mutants, for example using tetracycline-regulated or other systems. *C. elegans* injected with gene-specific duplex RNAs (222) may provide another model system to uncover in an intact multicellular organism new and crucial cap-related events. These may be effects on RNA polymerase II transcription complexes, such as the recently described stimulatory interaction between capping enzyme and a human transcription elongation element, SPT5

(82), or the interplay of cap binding protein phosphorylation and signaling pathways that modulate translation and cell growth control (223). In parallel with genetic studies in whole organisms, continued success in defining cap-related macromolecular structures is certain to uncover exciting insights into how the cap influences the life cycle of eukaryotic cells and viruses (224–228).

#### ACKNOWLEDGMENT

We thank Ms. Midori Kawaguchi, Ms. Miki Yoshiike, Ms. Domna Hebenstreit, Dr. Saori Kitao, and Dr. Akira Shimamoto for excellent assistance in compilation of this manuscript.

#### REFERENCES

1. Shatkin, A. J. (1976). Capping of eukaryotic mRNAs. *Cell* **9**, 645–653.
2. Furuichi, Y., and Shatkin, A. J. (1989). Characterization of cap structures. *Methods Enzymol.* **180**, 164–177.
3. Furuichi, Y., and Shatkin, A. J. (1991). Capping and methylation of mRNA: In “RNA Processing: A Practical Approach,” Vol II, (S. J. Higgins and B. D. Hames, eds.). pp. 35–67, Oxford University Press, London. .
4. Banerjee, A. (1980) 5'-Terminal cap structure in eukaryotic messenger ribonucleic acids. *Microbiol. Rev.* **44**, 175–205.
5. Reddy, R., and Busch, H. (1988). Small nuclear RNAs: RNA sequences, structure, and modifications. In “Structure and Function of Major and Minor snRNPs”, (M. Birnstiel, ed.) pp. 1–37, Springer Verlag, Berlin.
6. Reddy, R., Singh, R., and Shimba, S. (1992). Methylated cap structures in eukaryotic RNAs: Structure, synthesis and functions. *Pharma. Ther.* **54**, 249–267.
7. Furuichi, Y., Lafiandra, A., and Shatkin, A. J. (1977). 5'-Terminal structure and mRNA stability. *Nature (London)* **266**, 235–239.
8. Shimotohno, K., Kodama, Y., Hashimoto, J., and Miura, K. I. (1977). Importance of 5' terminal blocking structure to stabilize mRNA in eukaryotic protein synthesis. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 2734–2738.
9. Hsu, C. L., and Stevens, A. (1993). Yeast cells lacking 5' → 3' exonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5' cap structure. *Mol. Cell. Biol.* **13**, 4826–4835.
10. Lewis, J. D., and Izaurralde, E. (1997). The role of the cap structure in RNA processing and nuclear export. *Eur. J. Biochem.* **247**, 461–469.
11. Schwer, B., and Shuman, S. (1996). Conditional inactivation of mRNA capping enzyme affects yeast pre-mRNA splicing in vivo. *RNA* **2**, 574–583.
12. McDevitt, M. A., Gilmartin, G. M., Reeves, W. H., and Nevins, J. R. (1988). Multiple factors are required for poly(A) addition to a mRNA 3' end. *Genes Dev.* **2**, 588–597.
13. Cooke, C., and Alwine, J. C. (1996). The cap and the 3' splice site similarly affect polyadenylation efficiency. *Mol. Cell. Biol.* **16**, 2579–2584.
14. Flaherty, S. M., Fortes, P., Izaurralde, E., Mattaj, I. W., and Gilmartin, G. M. (1997). Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11893–11898.
15. Sonenberg, N. (1988). Cap-binding proteins of eukaryotic mRNA: Functions in initiation and control of translation. *Prog. Nucleic Acids Res. Mol. Biol.* **35**, 173–207.

16. Kominami, R., and Muramatsu, M. (1977). Heterogeneity of 5'-termini of nucleolar 45S, 32S and 28S RNA in mouse hepatoma. *Nucleic Acids Res.* **4**, 229-240.
17. Miura, K.-I., Watanabe, K., Sugiura, M., and Shatkin, A. J. (1974). The 5'-terminal nucleotide sequence of the double-stranded RNA of human reovirus. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3979-3983.
18. Furuichi, Y., and Miura, K. (1972). The 3'-termini of the genome RNA segments of silkworm cytoplasmic polyhedrosis virus. *J. Mol. Biol.* **64**, 619-632.
19. Furuichi, Y., and Miura, K.-I. (1973). Identity of the 3'-terminal sequences in ten genome segments of silkworm cytoplasmic polyhedrosis virus. *Virology* **55**, 418-425.
20. Miura, K., Watanabe, K., and Sugiura, M. (1974). 5'-Terminal nucleotide sequence of the double-stranded RNA of silkworm cytoplasmic polyhedrosis virus. *J. Mol. Biol.* **86**, 31-48.
21. Banerjee, A. K., Lafiandra, A., and Shatkin, A. J. (1970). Transcription in vitro by reovirus-associated ribonucleic acid-dependent polymerase. *J. Virol.* **6**, 1-11.
22. Shimotohno, K., Kimura, K., and Miura, K. (1973). Transcription of double-stranded RNA in cytoplasmic polyhedrosis virus in vitro. *Virology* **53**, 283-286.
23. Shatkin, A. J., and Sipe, J. D. (1968). RNA polymerase activity in purified reoviruses. *Proc. Natl. Acad. Sci. U.S.A.* **61**, 1462-1469.
24. Furuichi, Y. (1974). Methylation-coupled transcription by virus-associated transcriptase of cytoplasmic polyhedrosis virus containing double-stranded RNA. *Nucleic Acids Res.* **1**, 809-822.
25. Perry, R. P., and Kelley, D. E. (1974). Existence of methylated messenger RNA in mouse L cells. *Cell* **1**, 37-42.
26. Desrosiers, R., Friderici, K., and Rottman, F. (1974). Identification of methylated nucleotides in messenger RNA from Novikoff hepatoma cells. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3971-3975.
27. Reddy, R., Ro-Choi, T. S., Henning, D., and Busch, H. (1974). Primary sequence of U1 nuclear ribonucleic acid of Novikoff hepatoma ascites cells. *J. Biol. Chem.* **249**, 6486-6494.
28. Adams, J. M., and Cory, S. (1975). Modified nucleotides and bizarre 5'-termini in mouse myeloma mRNA. *Nature* **255**, 28-33.
29. Shatkin, A. J. (1974). Methylated messenger RNA synthesis in vitro by purified reovirus. *Proc. Natl. Acad. Sci. U.S.A.* **71**, 3204-3207.
30. Rottman, F., Shatkin, A. J., and Perry, R. P. (1974). Sequences containing methylated nucleotides at the 5' termini of messenger RNA: possible implication for processing. *Cell* **3**, 197-199.
31. Furuichi, Y., and Miura, K.-I. (1975). A blocked structure at the 5' terminus of mRNA from cytoplasmic polyhedrosis virus. *Nature* **253**, 374-375.
32. Furuichi, Y., Morgan, M., Muthukrishnan, S., and Shatkin, A. J. (1975). Reovirus messenger RNA contains a methylated, blocked 5'-terminal structure: m7G(5')ppp(5') GmpCp-. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 362-366.
33. Wei, C. M., and Moss, B. (1975). Methylated nucleotides block 5'-terminus of vaccinia virus messenger RNA. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 318-322.
34. Urushibara, T., Furuichi, Y., Nishimura, C., and Miura, K. (1975). A modified structure at the 5'-terminus of mRNA of vaccinia virus. *FEBS Lett.* **49**, 385-389.
35. Furuichi, Y., Morgan, M., Shatkin, A. J., Jelinek, W., Salditt-Georgieff, M., and Darnell, J. E. (1975). Methylated, blocked 5' termini in HeLa cell mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1904-1908.
36. Salditt-Georgieff, M., Harpold, M., Chen-Kiang, S., and Darnell, J. E. (1980). The addition of 5'-cap structures occurs early in hnRNA synthesis and prematurely terminated molecules are capped. *Cell* **19**, 69-78.



37. Ullu, E., and Tschudi, C. (1991). Trans splicing in trypanosomes requires methylation of the 5' end of the spliced leader RNA. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10074–10078.
38. Shuman, S. (1995). Capping enzyme in eukaryotic mRNA synthesis. *Prog. Nucleic Acids Res. Mole. Biol.* **50**, 101–129.
39. Martin, S. A., and Moss, B. (1975). Modification of RNA by mRNA guanylyltransferase and mRNA-(guanine-7)-methyltransferase from vaccinia virions. *J. Biol. Chem.* **250**, 9330–9335.
40. Ensinger, M. J., Martin, S. A., Paoletti, E., and Moss, B. (1975). Modification of the 5'-terminus of mRNA by soluble guanylyl and methyl transferases from vaccinia virus. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 2525–2529.
41. Venkatesan, S., Gershowitz, A., and Moss, B. (1980). Modification of the 5' end of mRNA. Association of RNA triphosphatase with the RNA guanylyltransferase-RNA (guanine-7) methyltransferase complex from vaccinia virus. *J. Biol. Chem.* **255**, 903–908.
42. McCracken, S., Fong, N., Ronina, E., Yankulv, K., Brothers, G., Siderovski, D., Hessel, A., Foster, S., Amgen EST Program, Shuman, S., and Bentley, D. (1997). 5'-Capping enzymes are targeted to pre-mRNA by binding to the phosphorylated carboxy-terminal domain of RNA polymerase II. *Genes Dev.* **11**, 3306–3318.
43. Cho, E.-J., Takagi, T., Moore, C. R., and Buratowski, S. (1997). mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain. *Genes Dev.* **11**, 3319–3326.
44. Yue, Z., Maldonado, E., Pillutla, R., Cho, H., Reinberg, D., and Shatkin, A. J. (1997). Mammalian capping enzyme complements mutant *Saccharomyces cerevisiae* lacking mRNA guanylyltransferase and selectively binds the elongation form of polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12898–12903.
45. Ho, K. C., Sriskanda, V., McCracken, S., Bentley, D., Schwer, B., and Shuman, S. (1998). The guanylyltransferase domain of mammalian mRNA capping enzyme binds to the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* **273**, 9577–9585.
46. Harris, N., Rosales, R., and Moss, B. (1993). Transcription initiation factor activity of vaccinia virus capping enzyme is independent of mRNA guanylylation. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 2860–2864.
47. Furuichi, Y. (1978). Pre-transcriptional capping in the biosynthesis of cytoplasmic polyhedrosis virus mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 488–492.
48. Wertheimer, A. M., Chen, S. Y., Borchardt, R. T., and Furuichi, Y. (1980). S-Adenosylmethionine and its analogs. Structural features correlated with synthesis and methylation of mRNAs of cytoplasmic polyhedrosis virus. *J. Biol. Chem.* **255**, 5924–5930.
49. Furuichi, Y. (1981). Allosteric stimulatory effect of S-adenosylmethionine on the RNA polymerase in cytoplasmic polyhedrosis virus. A model for the positive control of eukaryotic transcription. *J. Biol. Chem.* **256**, 483–493.
50. Yamakawa, M., Furuichi, Y., Nakashima, K., Lafiandra, A. J., and Shatkin, A. J. (1981). Excess synthesis of viral mRNA 5-terminal oligonucleotides by reovirus transcriptase. *J. Biol. Chem.* **256**, 6507–6514.
51. Yamakawa, M., Y., Furuichi and Shatkin, A. J. (1982). Priming of reovirus transcription by GppppG and formation of CpG(5')pppp(5')GpC. *Proc. Natl. Acad. Sci. U.S.A.* **79**, 6142–6146.
52. Liu, J. J., and Melenman, A. G. (1994). Purification and properties of GTP:GTP guanylyltransferase from encysted embryos of the brine shrimp *Artemia*. *J. Biol. Chem.* **269**, 11787–11794.

53. Smith, R. E., and Furuichi, Y. (1982). A unique class of compound, guanosine-nucleotide tetraphosphate G(5')pppp(5')N, synthesized during the in vitro transcription of cytoplasmic polyhedrosis virus of *Bombyx mori*. Structural determination and mechanism of formation. *J. Biol. Chem.* **257**, 485–494.
54. Maruyama, K., and Sugano, S. (1994). Oligo-capping: A simple method to replace the cap structure of eukaryotic mRNAs with oligoribonucleotides. *Gene* **138**, 171–174.
55. Suzuki, Y., Yoshitomi-Nakagawa, K., Maruyama, K., Suyama, A., and Sugano S. (1997). Construction and characterization of a full length-enriched and a 5'-end-enriched cDNA library. *Gene*. **200**, 149–156.
56. Yamabe, Y., Shimamoto, A., Goto, M., Yokota, J., Sugawara, M., and Furuichi, Y. (1998). Spl-Mediated transcription of the Werner helicase gene is modulated by Rb and p53. *Mol. Cell. Biol.* **18**, 6191–6200.
57. Xu, L. C., Thali, M., and Schaffner, W. (1991). Upstream box/TATA box order is the major determinant of the direction of transcription. *Nucleic Acids Res.* **19**, 6699–6704.
58. Dierks, P., Van Ooyen, A., Mantei, N., and Weissmann, C. (1981). DNA sequences preceding the rabbit beta-globin gene are required for formation in mouse L cells of beta-globin RNA with the correct 5' terminus. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 1411–1415.
59. Van Doren, K., and Hirsh, D. (1988). Trans-spliced leader RNA exists as small nuclear ribonucleoprotein particles in *Caenorhabditis elegans*. *Nature* **335**, 556–559.
60. Kuersten S., Lea, K., MacMorris, M., Spieth, J., and Blumenthal, T. (1997). Relationship between 3' end formation and SL2-specific trans-splicing in polycistronic *Caenorhabditis elegans* pre-mRNA processing. *RNA* **3**, 269–278.
61. "Fields Virology." (1996). Third Edition, (B. N. Fields, D. M. Knipe, P. M. Howley et al., eds.) Lippincott-Raven Publishers, Philadelphia.
62. Furuichi, Y., Shatkin, A. J., Stavnezer, E., and Bishop, J. M. (1975). Blocked, methylated 5'-terminal sequence in avian sarcoma virus RNA. *Nature* **257**, 618–620.
63. Guarino, L. A., Jin, J., and Dong, W. (1998). Guanylyltransferase activity of the LEF-4 subunit of baculovirus RNA polymerase. *J. Virol.* **72**, 10003–10010.
64. Mizumoto, K., and Kaziro, Y. (1987). Messenger RNA capping enzymes from eukaryotic cells. *Prog. Nucleic Acids Res. Mol. Biol.* **34**, 1–28.
65. Krug, R. M. Morgan, M. A., and Shatkin, A. J. (1976). Influenza viral mRNA contains internal N6-methyladenosine and 5'-terminal 7-methylguanosine in cap structures. *J. Virol.* **20**, 45–53.
66. Krug, R. M. (1989). Expression and replication of the influenza virus genome. In "The Influenza Viruses" (R. M. Krug, ed.) pp. 89–152, Plenum Press, New York.
67. Leahy, M. B., Dessens, J. T., and Nuttall, P. A. (1997). In vitro polymerase activity of Thogoto virus: Evidence for a unique cap-snatching mechanism in a tick-borne orthomyxovirus. *J. Virol.* **71**, 8347–8351.
68. Leung, D. W., Browning, K. S., Heckman, J. E., Rajbhandary, U. L., and Clark, J. M. Jr. (1979). Nucleotide sequence of the 5' terminus of satellite tobacco necrosis virus ribonucleic acid. *Biochemistry* **18**, 1361–1366.
69. Lee, Y. F., Nomoto, A., Detjen, B. M., and Wimmer, E. (1977). A protein covalently linked to poliovirus genome RNA. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 59–63.
70. Vartapetian, A. B., Drygin, Y. F., Chumakov, K. M., and Bogdanov, A. A. (1980). The structure of the covalent linkage between proteins and RNA in encephalomyocarditis virus. *Nucleic Acids Res.* **8**, 3729–3742.

71. Vartapetian, A. B., Koonin, E. V., Agol, V. I., and Bogdanov, A. A. (1984). Encephalomyocarditis virus RNA synthesis in vitro is protein-primed. *EMBO J.* **3**, 2593–2598.
72. Nomoto, A., Detjen, B., Pozzatti, R., and Wimmer, E. (1977). The location of the polio genome protein in viral RNAs and its implication for RNA synthesis. *Nature.* **268**, 208–213.
73. Furuichi, Y., Muthukrishnan, S., Tomasz, J., and Shatkin, A. J. (1976). Mechanism of the formation of reovirus mRNA 5'-terminal blocked and methylated sequence, m7GpppGmpC. *J. Biol. Chem.* **251**, 5043–5053.
74. Moss, B., Gershowitz, A., Wei, C.-M., and Boone, R. (1976). Formation of the guanylated and methylated 5'-terminus of vaccinia virus mRNA. *Virology.* **72**, 341–351.
75. Shimotohno, K., and Miura, K.-I. (1976). The process of formation of the 5'-terminal modified structure in messenger RNA of cytoplasmic polyhedrosis virus. *FEBS Lett.* **64**, 204–208.
76. Ho, C. K., Schwer, B., and Shuman, S. (1998). Genetic, physical, and functional interactions between the triphosphatase and guanylyltransferase components of the yeast mRNA capping apparatus. *Mol. Cell. Biol.* **18**, 5189–5198.
77. Shuman, S. (1997). Commentary; Origin of mRNA identity: Capping enzymes bind to the phosphorylated C-terminal domain of RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 12758–12760.
78. Shuman, S. (1982). RNA capping by HeLa cell RNA guanylyltransferase. Characterization of a covalent protein-guanylate intermediate. *J. Biol. Chem.* **25**, 7237–7245.
79. Pillutla, R. C., Yue, Z., Maldonado, E., and Shatkin, A. J. (1998). Recombinant human mRNA cap methyltransferase binds capping enzyme/RNA polymerase II complexes. *J. Biol. Chem.* **273**, 21443–21446.
80. Dahmus, M. E. (1996). Reversible phosphorylation of the C-terminal domain of RNA polymerase II. *J. Biol. Chem.* **271**, 19009–19017.
81. Ho, C. K., and Shuman, S. (1999). Distinct roles for CTD Ser-2 and Ser-5 phosphorylation in the recruitment and allosteric activation of mammalian mRNA capping enzyme. *Mol. Cell* **3**, 405–411.
82. Wen, Y., and Shatkin, A. J. (1999). Transcription elongation factor hSPT5 stimulates mRNA capping. *Genes Dev.* **13**, 1774–1779.
83. Langberg, S. R., and Moss, B. (1981). Post-transcriptional modifications of mRNA. Purification and characterization of cap I and cap II RNA (nucleoside-2'-)-methyltransferases from HeLa cells. *J. Biol. Chem.* **256**, 10054–10060.
84. Kuge, H., and Richter, J. D. (1995). Cytoplasmic 3' poly(A) addition induces 5' cap ribose methylation: Implication for translational control of maternal mRNA. *EMBO J.* **14**, 6301–6310.
85. Shuman, S., and Hurwitz, J. (1981). Mechanism of mRNA capping by vaccinia virus guanylyltransferase: Characterization of an enzyme-guanylate intermediate. *Proc. Natl. Acad. Sci. U.S.A.* **78**, 187–191.
86. Hagler, J., and Shuman, S. (1992). A freeze-frame view of eukaryotic transcription during elongation and capping of nascent mRNA. *Science* **255**, 983–986.
87. Shuman, S., Broyles, S., and Moss, B. (1987). Purification and characterization of a transcription termination factor from vaccinia virions. *J. Biol. Chem.* **262**, 12372–12380.
88. Gershon, P. D., Shi, X., and Hodel, A. E. (1998). Evidence that the RNA methylation and poly (A) polymerase stimulatory activities of vaccinia virus protein VP39 do not impinge upon one another. *Virology* **246**, 253–265.

89. Lockless, S. W., Cheng, H-T., Hodel, A. E., Quioco, F. A., and Gershon, P. D. (1998). Recognition of capped RNA substrates by VP39, the vaccinia virus-encoded mRNA cap-specific 2'-O-methyltransferase. *Biochemistry* **37**, 8564–8574.
90. Kawakami, K., Mizumoto, K., Ishihama, A., Shinozaki-Yamaguchi, K., and Miura, K. (1985). Activation of influenza virus associated RNA polymerase by cap-1 structure (m7GpppNm). *J. Biochem. (Tokyo)* **97**, 655–661.
91. Higman, M. A., Christen, L. A., and Niles, E. G. (1994). The mRNA (guanine-7-) methyltransferase domain of the vaccinia virus mRNA capping enzyme. *J. Biol. Chem.* **269**, 14974–14981.
92. Higman, M. A., and Niles, E. G. (1994). Location of the S-adenosyl-L-methionine binding region of the vaccinia virus mRNA (guanine-7-)methyltransferase. *J. Biol. Chem.* **269**, 14982–14987.
93. Myette, J. R., and Niles, E. G. (1996). Domain structure of the vaccinia virus mRNA capping enzyme. *J. Biol. Chem.* **271**, 11936–11944.
94. Mao, X., and Shuman, S. (1994). Intrinsic RNA (guanine-7) methyltransferase activity of the vaccinia virus capping enzyme D1 subunit is stimulated by the D12 subunit. *J. Biol. Chem.* **269**, 24472–24479.
95. Barbosa, E., and Moss, B. (1978). mRNA (nucleoside-2'-)-methyltransferase from vaccinia virus. Characteristics and substrate specificity. *J. Biol. Chem.* **253**, 7698–7702.
96. Schnierle, B. S., Gershon, P. D., and Moss, B. (1992). Cap-specific mRNA (nucleoside-02'-)-methyltransferase and poly(A) polymerase stimulatory activities of vaccinia virus are mediated by a single protein. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 2897–2901.
97. Hodel, A. E., Gershon, P. D., Shi, X., and Quioco, F. A. (1996). The 1.85 Å structure of vaccinia protein VP39: A bifunctional enzyme that participates in the modification of both mRNA ends. *Cell* **85**, 247–256.
98. Tsukamoto, T., Shibagaki, Y., Imajo-Ohmi, S., Murakoshi, T., Suzuki, M., Nakamura, A., Gotoh, H., and Mizumoto, K. (1997). Isolation and characterization of the yeast mRNA capping enzyme beta subunit gene encoding RNA 5'-triphosphatase, which is essential for cell viability. *Biochem. Biophys. Res. Comm.* **239**, 116–122.
99. Shibagaki, Y., Itoh, N., Yamada, H., Nagata, S., and Mizumoto, K. (1992). mRNA capping enzyme: Isolation and characterization of the gene encoding mRNA guanylyltransferase subunit from *Saccharomyces cerevisiae*. *J. Biol. Chem.* **267**, 9521–9528.
100. Mao, X., Schwer, B., and Shuman, S. (1995). Yeast mRNA cap methyltransferase is a 50-kilodalton protein encoded by an essential gene. *Mol. Cell. Biol.* **15**, 4167–4174.
101. Mao, X., Schwer, B., and Shuman, S. (1996). Mutational analysis of the *Saccharomyces cerevisiae* ABD1 gene: Cap methyltransferase activity is essential for cell growth. *Mol. Cell. Biol.* **16**, 475–480.
102. Yamada-Okabe, T., Shimmi, O., Doi, R., Mizumoto, K., Arisawa, M., and Yamada-Okabe, H. (1996). Isolation of the mRNA-capping enzyme and ferric-reductase-related genes from *Candida albicans*. *Microbiology* **142**, 2515–2523.
103. Shuman, S., Liu, Y., and Schwer, B. (1994). Covalent catalysis in nucleotidyl transfer reactions: Essential motifs in *Saccharomyces cerevisiae* RNA capping enzyme are conserved in *Schizosaccharomyces pombe* and viral capping enzymes and among polynucleotide ligases. *Proc. Natl. Acad. Sci. U.S.A.* **91**, 12046–12050.
104. Shibagaki, Y., Gotoh, H., Kato, M., and Mizumoto, K. (1995). Localization and in vitro mutagenesis of the active site in the *Saccharomyces cerevisiae* mRNA capping enzyme. *J. Biochem. (Tokyo)* **118**, 1303–1309.

105. Wang, S. P., and Shuman, S. (1997). Structure-function analysis of the mRNA cap methyltransferase of *Saccharomyces cerevisiae*. *J. Biol. Chem.* **272**, 14683–14689.
106. Yamada-Okabe, T., Mio, T., Matsi, M., Kashima, Y., Arisawa, M., and Yamada-Okabe, H. (1998). Isolation and characterization of the *Candida albicans* gene for mRNA 5'-triphosphatase: Association of mRNA 5'-triphosphatase and mRNA 5'-guanylyltransferase activities is essential for the function of mRNA 5'-capping enzyme in vivo. *FEBS Lett.* **435**, 49–54.
107. Takagi, T., Moore, C. R., Diehn, F., and Buratowski, S. (1997). An RNA 5'-triphosphatase related to the protein tyrosine phosphatases. *Cell* **89**, 867–873.
108. Wang, S. P., Deng, L., Ho, C. K., and Shuman, S. (1997). Phylogeny of mRNA capping enzymes. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9573–9578.
109. Silva, E., Ullu, E., Kobayashi, R., and Tschudi, C. (1998). Trypanosome capping enzymes display a novel two-domain structure. *Mol. Cell. Biol.* **18**, 4612–4619.
110. Yagi, Y., Mizumoto, K., and Kaziro, Y. (1984). Limited tryptic digestion of messenger RNA capping enzyme from *Artemia salina*. Isolation of domains for guanylyltransferase and RNA 5'-triphosphatase. *J. Biol. Chem.* **259**, 4695–4698.
111. Yamada-Okabe, T., Doi, R., Shimmi, O., Arisawa, M., and Yamada-Okabe, H. (1998). Isolation and characterization of a human cDNA for mRNA 5'-capping enzyme. *Nucleic Acids Res.* **26**, 1700–1706.
112. Tsukamoto, T., Shibagaki, Y., Murakoshi, T., Suzuki, M., Nakamura, A., Gotoh, H., and Mizumoto, K. (1998). Cloning and characterization of two human cDNAs encoding the mRNA capping enzyme. *Biochem. Biophys. Res. Commun.* **243**, 101–108.
113. Tsukamoto, T., Shibagaki, Y., Nikura, Y., and Mizumoto, K. (1998). Cloning and characterization of three human cDNAs encoding mRNA (guanine-7-)-methyltransferase, an mRNA cap methylase. *Biochem. Biophys. Res. Commun.* **251**, 27–34.
114. Pillutla, R., Shimamoto, A., Furuichi, Y., and Shatkin, A. J. (1998). Human mRNA capping enzyme (RNGTT) and cap methyltransferase (RNMT) Map to 6q16 and 18p11.22-p11.23, respectively. *Genomics* **54**, 351–353.
115. Seliger, L. S., Zheng, K., and Shatkin, A. J. (1987). Complete nucleotide sequence of reovirus L2 gene and deduced amino acid sequence of viral mRNA guanylyltransferase. *J. Biol. Chem.* **262**, 16289–16293.
116. Mao, Z. X., and Joklik, W. K. (1991). Isolation and enzymatic characterization of protein lambda 2, the reovirus guanylyltransferase. *Virology* **185**, 377–386.
117. Luongo, C. L., Contreras, C. M., Farsetta, D. L., and Nibert, M. L. (1998). Binding site for S-adenosyl-L-methionine in a central region of mammalian reovirus lambda2 protein. Evidence for activities in mRNA cap methylation. *J. Biol. Chem.* **273**, 23773–23780.
118. Noble, S., and Nibert, M. L. (1997). Core protein  $\mu$ 2 is a second determinant of nucleoside triphosphatase activities by reovirus cores. *J. Virol.* **71**, 7728–7735.
119. Gross, C. H., and Shuman, S. (1998). Characterization of a baculovirus-encoded RNA 5'-triphosphatase. *J. Virol.* **72**, 7057–7063.
120. Jin, J., Dong, W., and Guarino, L. A. (1998). The LEF-4 subunit of baculovirus RNA polymerase has RNA 5'-triphosphatase and ATPase activities. *J. Virol.* **72**, 10011–10019.
121. Gross, C. H., and Shuman, S. (1998). RNA 5'-triphosphatase, nucleoside triphosphatase, and guanylyltransferase activities of baculovirus LEF-4 protein. *J. Virol.* **72**, 10020–10028.
122. Martinez-Costas, J., Sutton, G., Ramadevi, N., and Roy, P. (1998). Guanylyltransferase and RNA 5'-triphosphatase activities of the purified expressed VP4 protein of bluetongue virus. *J. Mol. Biol.* **280**, 859–866.

123. Mi, S. D., Huang, H. V., Rice, C. M., and Stiller, V. (1989). Association of the Sindbis virus methyltransferase activity with the nonstructural protein nsP1. *Virology*, **170**, 385–391.
124. Laakkonen, P., Hyvonen, M., Peranen, J., and Kääriäinen, L. (1994). Expression of Semliki Forest virus nsP1-specific methyltransferase in insect cells and in *Escherichia coli*. *J. Virol.* **74**, 7418–7425.
125. Ahola, T., and Kääriäinen, L. (1995). Reactions in alphavirus mRNA capping: Formation of a covalent complex of nonstructural protein nsP1 with 7-methyl-GMP. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 507–511.
126. Ahola, T., Laakkonen, P., Vihinen, H., and Kääriäinen, L. (1997). Critical residues of Semliki Forest virus RNA capping enzyme involved in methyltransferase and guanylyltransferase-like activities. *J. Virol.* **71**, 392–397.
127. Bisaillon, M., and Guy, L. (1997). Minireview: Viral and cellular enzymes involved in synthesis of mRNA cap structure. *Virology* **236**, 1–7.
128. Ho, C. K., Van Etten, J. L., and Shuman, S. (1996). Expression and characterization of an RNA capping enzyme encoded by *Chlorella* virus PBCV-1. *J. Virol.* **70**, 6658–6664.
129. Hakanson, K., Doherty, A. J., Shuman, S., and Wigley, D. B. (1997). X-ray crystallography reveals a large conformational change during guanyl transfer by mRNA capping enzymes. *Cell* **89**, 545–553.
130. Das, T., Mathur, M., Gupta, A. K., Janssen, G. M. C., and Banerjee, A. K. (1998). RNA polymerase of vesicular stomatitis virus specifically associates with translation elongation factor- $\alpha\beta\gamma$  for its activity. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 1449–1454.
131. Banerjee, A. K., Mathur, M., Gupta, A. K., and Datta, U. (1999). Involvement of cellular guanylyltransferase in the 5'-terminal capping of vesicular stomatitis virus mRNAs in vitro. Abstract for the International Virology Meeting, Sydney, Australia.
132. Shuman, S. (1997). Minireview: A proposed mechanism of mRNA synthesis and capping by vesicular stomatitis virus. *Virology* **227**, 1–6.
133. Ro-Choi, T. S., Raj, N. B. K., Pike, L. M., and Busch, H. (1976). Effects of  $\alpha$ -amanitin, cycloheximide, and thioacetamide on low molecular weight nuclear RNA. *Biochemistry* **15**, 3823–3828.
134. Hellung-Larsen, P., Kulamowicz, I., and Frederiksen, S. (1980). Synthesis of low molecular weight RNA components in cells with a temperature-sensitive polymerase II. *Biochem. Biophys. Acta* **609**, 201–204.
135. Eliceiri, G. L. (1980). Formation of low molecular weight RNA species in HeLa cells. *J. Cell. Physiol.* **102**, 199–207.
136. Skuzeski, J. M., Lund, E., Murphy, J. T., Steinberg, T. H., Burgess, R. R., and Dahlberg, J. E. (1984). Multiple elements upstream of the coding region are required for accumulation of human U1 RNA *in vivo*. *J. Biol. Chem.* **259**, 8345–8352.
137. Mattaj, I. W. (1986). Cap trimethylation of U snRNA is cytoplasmic and dependent on U snRNA protein binding. *Cell* **46**, 905–911.
138. Fischer, U., and Luhrmann, R. (1990). An essential signaling role for the m<sup>3</sup>G cap in the transport of U1 snRNP to the nucleus. *Science* **249**, 786–790.
139. Hamm, J., Darzynkiewicz, E., Tahara, S. M., and Mattaj, I. W. (1990). The trimethylguanosine cap structure of U1 snRNA is a component of a bipartite nuclear targeting signal. *Cell* **62**, 569–577.
140. Terns, M. P., and Dahlberg, J. E. (1994). Retention and 5' cap trimethylation of U3 snRNA in the nucleus. *Science* **264**, 959–961.

141. Steinberg, T. H., Mathews, D. E., Durbin, R. D., and Burgess, R. R. (1990). Tagetitoxin: A new inhibitor of eukaryotic transcription by RNA polymerase III. *J. Biol. Chem.* **265**, 499–505.
142. Shimba, S., and Reddy, R. (1994). Purification of human U6 small nuclear RNA capping enzyme. *J. Biol. Chem.* **269**, 12419–12423.
143. Bokar, J. A., Shambaugh, M. E., Polayes, D., Matera, A. G., and Rottman, F. M. (1997). Purification and cDNA cloning of the AdoMet-binding subunit of the human mRNA (N6-adenosine)-methyltransferase. *RNA* **3**, 1233–1247.
144. Kuersten S., Lea K., MacMorris, M., Spieth, J., and Blumenthal, T. (1997). Relationship between 3' end formation and SL2-specific trans-splicing in polycistronic *Caenorhabditis elegans* pre-mRNA processing *RNA* **3**, 269–278.
145. Van Doren, K., and Hirsh, D. (1988). Trans-spliced leader RNA exists as small nuclear ribonucleoprotein particles in *Caenorhabditis elegans*. *Nature* **335**, 556–559.
146. Blumenthal, T. (1995). Trans-splicing and polycistronic transcription in *C. elegans*. *Trends Genet.* **11**, 132–136.
147. Blumenthal, T., and Steward, K. (1997). RNA processing and gene structure. In *C. elegans* II, (D. Riddle, T. Blumenthal, B. Meyer, and J. Press eds.) pp. 117–145. Cold Spring Harbor Laboratory Press, New York.
148. Ullu, E., and Tschudi, C. (1995). Accurate modification of the trypanosome spliced leader cap structure in a homologous cell-free system. *J. Biol. Chem.* **270**, 20365–20369.
149. Lee, M. G., and Van Der Ploeg, L. H. (1997). Transcription of protein-coding genes in trypanosomes by RNA polymerase I. *Ann. Rev. Microbiol.* **51**, 463–489.
150. Breckenridge, D. G., Watanabe, Y., Greenwood, S. J., Gray, M. W., and Schnare, N. M. (1999). U1 small nuclear RNA and spliceosomal introns on *Euglena gracilis*. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 853–856.
151. Plotch, S. J., Bouloy, M., and Krug, R. M. (1979). Transfer of 5'-terminal cap of globin mRNA to influenza viral complementary RNA during transcription. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 1618–1622.
152. Plotch, S. J., Bouloy, M., Ulmanen, I., and Krug, R. M. (1981). A unique cap (m7GpppXm)-dependent influenza virion endonuclease cleaves capped RNAs to generate the primers that initiate viral RNA transcription. *Cell* **23**, 847–858.
153. Braam, J., Ulmanen, I., and Krug, R. M. (1983). Molecular model of a eucaryotic transcription complex: Functions and movements of influenza P proteins during capped RNA-primed transcription. *Cell* **34**, 609–618.
154. Hagen, M., Tiley, L., Chung, T. D., and Krystal, M. (1995). The role of template-primer interactions in cleavage and initiation by the influenza virus polymerase. *J. Gen. Virol.* **76**, 603–611.
155. Klumpp, K., Ruigrk, R. W., and Baudin, F. (1997). Role of the influenza virus polymerase and nucleoprotein in forming a functional RNP structure. *EMBO J.* **16**, 1248–1257.
156. Kawakami, K., Mizumoto, K., Ishihama, A., Shinozaki-Yamaguchi, K., and Miura, K. (1985). Activation of influenza virus-associated RNA polymerase by cap-1 structure (m7GpppNm). *J. Biochem. (Tokyo)* **97**, 655–661.
157. Li, M-L., Ramirez, B. C., and Krug, R. M. (1998). RNA-dependent activation of primer RNA production by influenza virus polymerase: Different regions of the same protein subunit constitute the two required RNA-binding sites. *EMBO J.* **17**, 5844–5852.
158. Leahy, M. B., Dessens, J. T., and Nuttall P. A. (1997). In vitro polymerase activity of Thogoto virus: Evidence for a unique cap-snatching mechanism in a tick-born orthomyxovirus *J. Virol.* **71**, 8347–8351.

159. Tomassini, J. E., Davies, M. E., Hastings, J. C., Lingham, R., Mojena, M., Raghobar, S. L., Singh, S. B., Tkacz, J. S., and Goetz, M. A. (1996). A novel antiviral agent which inhibits the endonuclease of influenza viruses. *Antimicrob. Agents Chemother.* **40**, 1189–1193.
160. Lai, M. M. C., and Cavanagh, D. (1997). The molecular biology of coronaviruses. *Adv. Virus. Res.* **48**, 1–100.
161. Hsu, C. L., and Stevens, A. (1993). Yeast cells lacking 5′-3′ exoribonuclease 1 contain mRNA species that are poly(A) deficient and partially lack the 5′ cap structure. *Mol. Cell. Biol.* **13**, 4826–4835.
162. Poole, T. L. and Stevens, A. (1997). Structural modifications of RNA influence the 5′ exoribonucleolytic hydrolysis by XRN1 and HKE1 of *Saccharomyces cerevisiae*. *Biochem. Biophys. Res. Commun.* **235**, 799–805.
163. Walther, T. N., Wittop Koning, T. H., Schimperli, D., and Muller, B. (1998). A 5′-3′ exonuclease activity involved in forming the 3′ products of histone pre-mRNA processing in vitro. *RNA* **4**, 1034–1046.
164. Sachs, A. B. (1993). Messenger RNA degradation in eukaryotes. *Cell.* **74**, 413–421.
165. Larimer, F. W., and Stevens, A. (1990). Disruption of the gene XRN1, coding for a 5′-3′ exoribonuclease, restricts yeast cell growth. *Gene* **95**, 85–90.
166. Larimer, F. W., Hsu, C. L., Maupin, M. K., and Stevens, A. (1992). Characterization of the XRN1 gene encoding a 5′→3′ exoribonuclease: Sequence data and analysis of disparate protein and mRNA levels of gene-disrupted yeast cells. *Gene* **120**, 51–57.
167. Lo, H-J., Huang, H-K., and Donahue, T. F. (1998). RNA polymerase I-promoted HIS4 expression yields uncapped, polyadenylated mRNA that is unstable and inefficiently translated in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **18**, 665–675.
168. Nuss, D. L., Furuichi, Y., Koch, G., and Shatkin, A. J. (1975). Detection in HeLa cell extracts of a 7-methyl guanosine specific enzyme activity that cleaves m7GpppNm. *Cell* **6**, 21–27.
169. Nuss, D. L., and Furuichi, Y. (1977). Characterization of the m7G(5′)pppN-pyrophosphatase activity from HeLa cells. *J. Biol. Chem.* **252**, 2815–2821.
170. Decker, C. J., and Parker, R. (1994). Mechanisms of mRNA degradation in eukaryotes. *Trends Biochem. Sci.* **19**, 336–340.
171. Caponigro, G., and Parker, R. (1995). Multiple functions for the poly (A)-binding protein in mRNA decapping and deadenylation in yeast. *Genes Dev.* **9**, 2421–2432.
172. Beelman, C. A., Stevens, A., Caponigro, G., LaGrandeur, T. E., Hatfield, L., Fortner, D. M., and Parker, R. (1996). An essential component of the decapping enzyme required for normal rates of mRNA turnover. *Nature* **382**, 642–646.
173. Hatfield, L., Beelman, C. A., Stevens, A., and Parker, R. (1996). Mutations in trans-acting factors affecting mRNA decapping in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **16**, 5830–5838.
174. LaGrandeur, T. E., and Parker, R. (1996). mRNA decapping activities and their biological roles. *Biochimie.* **78**, 1049–1055.
175. LaGrandeur, T. E., and Parker, R. (1998). Isolation and characterization of Dcp1p, the yeast mRNA decapping enzyme. *EMBO J.* **17**, 1487–1496.
176. Tharun, S., and Parker, R. (1999). Analysis of mutations in the yeast mRNA decapping enzyme. *Genetics.* **151**, 1273–1285.
177. Schwer, B., Mao, X., and Shuman, S. (1998). Accelerated mRNA decay in conditional mutants of yeast mRNA capping enzyme. *Nucleic Acids Res.* **26**, 2050–2057.
178. Couttet, P., Promont-Racine, M., Steel, D., Pictet, R., and Grange, T. (1997). Messenger RNA deadenylation precedes decapping in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 5628–5633.



179. Muhlrad, D., and Parker, R. (1994). Premature translational termination triggers mRNA decapping. *Nature* **370**, 578–581.
180. Lim, S. K., and Maquat, L. E. (1992). Human beta-globin mRNAs that harbor a nonsense codon are degraded in murine erythroid tissues to intermediates lacking regions of exon I or exons I and II that have a cap-like structure at the 5' termini. *EMBO J.* **11**, 3271–3278.
181. Yamabe, Y., Sugimoto, M., Satoh, M., Suzuki, N., Sugawara, M., Goto, M., and Furuichi, Y. (1997). Down-regulation of the defective transcripts of the Werner's syndrome gene in the cells of patients. *Biochem. Biophys. Res. Commun.* **236**, 151–154.
182. Kitao, S., Shimamoto, A., Goto, M., Miller, R. W., Smithson, W. A., Lindor, N. M., and Furuichi, Y. (2000). Identification of RecQ4 DNA helicase gene as a causative gene of Rothmund-Thomson syndrome. *Nature Genet.* **22**, 82–84.
183. Konarska, M., Padgett, R., and Sharp, P. (1984). Recognition of cap structure in splicing in vitro of mRNA precursors. *Cell* **38**, 731–736.
184. Edery, I., and Sonenberg, N. (1985). Cap-dependent RNA splicing in a HeLa nuclear extract. *Proc. Natl. Acad. Sci. U.S.A.* **82**, 7590–7594.
185. Lewis, J. D., Izaurralde, E., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1996). A nuclear cap-binding complex facilitates association of U1 snRNP with the cap-proximal 5' splice site. *Genes Dev.* **10**, 1683–1698.
186. Inoue, K., Ohno, M., Sakamoto, H., and Shimura, Y. (1989). Effect of the cap structure on pre-mRNA splicing in *Xenopus* oocyte nuclei. *Genes Dev.* **3**, 1472–1479.
187. Izaurralde, E., Lewis, J., McGuigan, C., Jankowska, M., Darzynkiewicz, E., and Mattaj, I. W. (1994). A nuclear cap binding complex involved in pre-mRNA splicing. *Cell* **78**, 657–668.
188. Visa, N., Izaurralde, E., Ferreira, J., Daneholt, B., and Mattaj, I. W. (1996). A nuclear cap-binding complex binds Balbiani ring pre-mRNA co-transcriptionally and accompanies the ribonucleoprotein particle during nuclear export. *J. Cell Biol.* **133**, 4–14.
189. Colot, H. V., Stutz, F., and Rosbash, M. (1996). The yeast splicing factor Mud13p is a commitment complex component and corresponds to CBP20, the small subunit of the nuclear cap-binding complex. *Genes Dev.* **10**, 1699–1708.
190. Flaherty, S. M., Fortes, P., Izaurralde, E., Mattaj, I. W., and Gilmartin, G. M. (1997). Participation of the nuclear cap binding complex in pre-mRNA 3' processing. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11893–11898.
191. Berget, S. M. (1995). Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**, 2411–2414.
192. Fresco, L. D., and Buratowski, S. (1996). Conditional mutants of the yeast mRNA capping enzyme show that the cap enhances, but is not required for mRNA splicing. *RNA* **2**, 584–596.
193. Jamolowski, A., Boelens, W. C., Izaurralde, E., and Mattaj, I. W. (1994). Nuclear export of different classes of RNA is mediated by specific factors. *J. Cell. Biol.* **124**, 627–635.
194. Izaurralde, E., Lewis, J., Gamberi, C., Jarmolowski, A., McGuigan, C., and Mattaj, I. W. (1995). A cap-binding protein complex mediating U snRNA export. *Nature* **376**, 709–712.
195. Izaurralde, E., and Adam, S. (1998). Transport of macromolecules between the nucleus and the cytoplasm. *RNA (review)* **4**, 351–364.
196. Both, G. W., Banerjee, A. K., and Shatkin, A. J. (1975). Methylation-dependent translation of viral messenger RNAs *in vitro*. *Proc. Natl. Acad. Sci. U.S.A.* **72**, 1189–1193.

197. Muthukrishnan, S., Both, G. W., Furuichi, Y., and Shatkin, A. J. (1975). 5'-terminal 7-methylguanosine in eukaryotic mRNA is required for translation. *Nature* **255**, 33-37.
198. Both, G. W., Furuichi, Y., Muthukrishnan, S., and Shatkin, A. J. (1975). Ribosome binding to reovirus mRNA in protein synthesis requires 5' terminal 7-methylguanosine. *Cell* **6**, 185-195.
199. Hickey, E. D., Weber, L. A., and Baglioni, C. (1976). Inhibition of initiation of protein synthesis by 7-methylguanosine-5'-monophosphate. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 19-23.
200. Adams, B. L., Morgan, M., Muthukrishnan, S., and Shatkin, A. J. (1978). The effect of "cap" analogs on reovirus mRNA binding to wheat germ ribosomes. *J. Biol. Chem.* **253**, 2589-2595.
201. Filipowicz, W., Furuichi, Y., Sierra, J. M., Muthukrishnan, S., Shatkin, A. J., and Ochoa, S. (1976). A protein binding the methylated 5'-terminal sequence, m7GpppN, of eukaryotic messenger RNA. *Proc. Natl. Acad. Sci. U.S.A.* **73**, 1559-1563.
202. Kozak, M., and Shatkin, A. J. (1976). Characterization of ribosome-protected fragments from reovirus messenger RNA. *J. Biol. Chem.* **251**, 4259-4266.
203. Kozak, M., and Shatkin, A. J. (1978). Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine. *J. Biol. Chem.* **253**, 6568-6577.
204. Kozak, M. (1978). How do eukaryotic ribosomes select initiation regions in messenger RNA? *Cell* **15**, 1109-1123.
205. Sonenberg, N., Morgan, M., Merrick, W. C., and Shatkin, A. J. (1978). A polypeptide in eukaryotic initiation factors that crosslinks specifically to the 5'-terminal cap in mRNA. *Proc. Natl. Acad. Sci. U.S.A.* **75**, 4843-4847.
206. Sonenberg, N., Rupprecht, K. M., Hecht, S. M., and Shatkin, A. J. (1979). Eukaryotic mRNA cap binding protein purification by affinity chromatography on Sepharose-coupled m7GDP. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 4345-4349.
207. Sonenberg, N., Trachsel, H., Hecht, S., and Shatkin, A. J. (1980). Differential stimulation of capped mRNA translation in vitro by cap binding protein. *Nature* **285**, 331-333.
208. Ehrenfeld, E. (1996). Initiation of translation by picornavirus RNAs. In "Translational Control" (J. W. B. Hershey, M. B. Mathews, and N. Sonenberg, eds.). pp. 549-573, Cold Spring Harbor Laboratory Press, New York.
209. Tahara, S. M., Morgan, M. A., and Shatkin, A. J. (1981). Two forms of purified m7G-cap binding protein with different effects on capped mRNA translation in extracts of uninfected and poliovirus-infected HeLa cells. *J. Biol. Chem.* **256**, 7691-7694.
210. Grifo, J. A., Tahara, S. M., Morgan, M. A., Shatkin, A. J., and Merrick, W. C. (1983). New initiation factor activity required for globin mRNA translation. *J. Biol. Chem.* **258**, 5804-5810.
211. Lamphear, B. J., Kirchweger, R., Skem, T., and Rhoads, R. E. (1995). Mapping of functional domains in eukaryotic protein synthesis initiation factor 4G (eIF4G) with picornaviral proteases. *J. Biol. Chem.* **270**, 21975-21983.
212. Mader, S., Lee, H., Pause, A., and Sonenberg, N. (1995). The translation initiation factor eIF-4E binds to a common motif shared by the translation factor, eIF-4 gamma and the translational repressors 4E-binding proteins. *Mol. Cell. Biol.* **15**, 4990-4997.
213. Hentze, M. W. (1997). eIF4G: A multipurpose ribosome adapter? *Science* **275**, 500-501.

214. Fraser, C. S., Pain, V. M., and Morley, S. J. (1999). The association of initiation factor 4F with poly(A)-binding protein is enhanced in serum-stimulated *Xenopus* kidney cells. *J. Biol. Chem.* **274**, 196–204.
215. Gradi, A., Svitkin, Y. V., Imataka, H., and Sonenberg, N. (1998). Proteolysis of human eukaryotic translation initiation factor eIF4GII, but not eIF4GI, coincides with the shutoff of host protein synthesis after poliovirus infection. *Proc. Natl. Acad. Sci. U.S.A.* **95**, 11089–11094.
216. Shatkin, A. J. (1985). mRNA cap binding proteins: Essential factors for initiating translation. *Cell* **40**, 223–224.
217. Sachs, A. B., Sarnow, P., and Hentze, M. W. (1997). Starting at the beginning, middle, and end: translation initiation in eukaryotes. *Cell* **89**, 831–838.
218. Sonenberg, N., and Gingras, A.-C. (1998). The mRNA 5' cap-binding protein eIF4E and control of cell growth. *Curr. Opin. Cell Biol.* **10**, 268–275.
219. Pain, V. M. (1996). Initiation of protein synthesis in eukaryotic cells. *Eur. J. Biochem.* **236**, 747–771.
220. Imataka, H., Gradi, A., and Sonenberg, N. (1998). A newly identified N-terminal amino acid sequence of human eIF4G binds poly(A)-binding protein and functions in poly(A)-dependent translation. *EMBO J.* **17**, 7480–7489.
221. Gingras, A.-C., Raught, B., and Sonenberg, N. (1999). eIF4 initiation factors: Effectors of mRNA recruitment to ribosomes and regulators of translation. *Annu. Rev. Biochem.* **68**, 913–963.
222. Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**, 806–811.
223. Rhoads, R. E. (1999). Signal transduction pathways that regulate eukaryotic protein synthesis. *J. Biol. Chem.* **274**, 30337–30340.
224. Doherty, A. J. (1999). Conversion of a DNA ligase into an RNA capping enzyme. *Nucleic Acids Res.* **27**, 3253–3258.
225. Ahola, T., Lampio, A., Auvinen, P., and Kääriäinen, L. (1999). Semliki Forest virus mRNA capping enzyme requires association with anionic membrane phospholipids for activity. *EMBO J.* **18**, 3164–3172.
226. Lampio, A., Ahola, T., Darzynkiewicz, E., Stepinski, J., Jankowska-Anyszka, M., Kääriäinen, L. (1999). Guanosine nucleotide analogs as inhibitors of alphavirus mRNA capping enzyme. *Antiviral Res.* **42**, 35–46.
227. Hu, G., Gershon, P. D., Hodel, A. E., and Quijcho, F. A. (1999). mRNA cap recognition: Dominant role of enhanced stacking interactions between methylated bases and protein aromatic side chains. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 7149–7154.
228. Marcotrigiano, J., Gingras, A.-C., Sonenberg, N., and Burley, S. K. (1999). Cap-dependent translation initiation in eukaryotes is regulated by a molecular mimic of eIF-4G. *Mol. Cell* **3**, 707–716.