Messenger RNAs of a strongly-expressed late gene of cowpox virus contain 5'-terminal poly(A) sequences

Dhavalkumar D.Patel and David J.Pickup
Department of Microbiology and Immunology, Duke University Medical Center, Durham, NC 27710, USA

Communicated by J.J.Skehel

We have identified and characterized one of the most strongly-expressed genes of cowpox virus (CPV). This is the gene encoding the major protein component of the A-type inclusions bodies produced by this virus. This gene (designated the 160K gene) is transcribed late during the infection. Analyses of its mRNAs showed that these late RNAs, unlike all other characterized late mRNAs of poxviruses, are uniform in length. However, the most remarkable feature of the mRNAs of the 160K gene is the structure of their 5' termini. Most of these mRNAs have 5'-terminal poly(A) sequences containing 5-21 residues. Furthermore, these 5'-terminal poly(A) sequences are not complementary to the corresponding region of the template strand of the viral DNA. Instead, the nucleotide sequences of the mRNA and the viral DNA diverge at the site of the three As in the sequence 5'-TAAAATG-3' containing the gene's initiation codon. Consequently, the poly(A) provides the leader sequences of these mRNAs. These unusual 5'-terminal structures suggest that the late mRNAs of poxvirus genes are generated by a novel process.

Key words: orthopoxviruses/RNA structure/transcription/vaccinia virus

Introduction

The orthopoxviruses (reviewed by Moss, 1985) are large, DNA-containing viruses that replicate in the cytoplasm of mammalian cells. They are unlike other DNA-containing viruses in that they employ viral proteins instead of host-cell proteins in many processes involved in virus replication, including the transcription of viral genes.

The transcription of viral genes is a tightly-regulated process. There is a temporal regulation of transcription of each gene (Oda and Joklik, 1967). Genes that are transcribed before the onset of viral DNA replication are designated early genes and those that are transcribed after the onset of viral DNA replication are designated late genes. The efficiency of the transcription of each gene is also regulated.

Currently, the mechanisms involved in the regulation of viral transcription are poorly understood. It is clear that the transcription of viral genes requires viral proteins and also specific regulatory elements within the viral DNA (Puckett and Moss, 1983; Cochran et al., 1985a). The regulatory elements appear to comprise short sequences immediately upstream of the transcribed portion of each gene (Bertholet et al., 1985, 1986; Cochran et al., 1985b; Hänggi et al., 1986; Weir and Moss, 1987). These regulatory elements share little similarity with transcriptional promoter elements of eukaryotes. And, apart from AT-rich regions, they share little similarity with each other. The most highly conserved sequence in these elements is the sequence 5'-TAAATG-3', which is present at or near the putative transcriptional initiation site of all characterized late genes of vaccinia virus (Weir and Moss, 1984; Bertholet et al., 1985; Plucienniczak et al., 1985; Rosel and Moss, 1985; Hänggi et al., 1986; Niles et al., 1986; Rosel et al., 1986; Weinrich and Hruby, 1986; Roseman and Hruby, 1987). The 5'-TAAATG-3' component of this conserved sequence appears to be essential for the transcription of the late 11K gene of vaccinia virus (Hänggi et al., 1986), but its exact role is unknown. The nature of the various cis-acting elements that control the efficiency of transcription, the site of transcriptional initiation and the temporal class of the gene are also unknown.

Recently, we have identified a gene that appears to be one of the most strongly-expressed of all of the genes of the orthopoxviruses. This is the gene encoding the 160-kd protein that is the major component of the A-type inclusions (ATIs) produced in cells infected with cowpox virus (Patel et al., 1986). Late during infection its product, the 160-kd protein, may accumulate to form up to 4% of the total protein in the cell. Studies on this gene should allow us to identify the features that direct high levels of expression of late genes.

This paper describes the characterization of the gene (designated the 160K gene) that encodes the 160-kd ATI protein. It also describes the unusual structure of this gene's mRNAs.

Results

Identification of the position of the 160K gene within the CPV genome

The mRNAs of CPV were fractionated by hybridization-selection with various cloned PstI and KpnI fragments of CPV DNA (Figure 1A). Immunoprecipitation of the translation products of these mRNAs (Figure 2) showed that the mRNAs encoding the 160-kd ATI proteins were complementary to sequences contained within the KpnI G fragment. Subsequent hybridization-selection experiments showed that the complementary region was within a 6.4-kb Clal fragment derived from the KpnI G fragment (Figure 1B). Furthermore, the transfection of plasmid p2025 (containing this Clal fragment) into cells infected with vaccinia virus induced the synthesis of a 160-kd protein (Figure 3, lane C). These results indicated that the 160-kd ATI protein was encoded by a gene contained within the 6.4-kb Clal fragment. An open reading frame of about 4.3 kb would be required to encode a gene product of this size.

The 160K gene is contained in a region of the DNA that is highly conserved among the genomes of the orthopoxviruses (Mackett and Archard, 1979; Esposito and Knight, 1986). This location is consistent with the presence of counterparts to the 160K gene in the genomes of other orthopoxviruses. These cognate genes encode antigenically-similar proteins whose apparent molecular masses (94 kd, vaccinia virus; 155 kd, raccoonpox virus; 92 kd, monkeypox virus; and 130 kd, ectromelia virus) are characteristic of the virus type (Kitamoto et al.,...
Mapping of the transcribed region of the 160K gene

The direction of transcription of the 160K gene was determined by hybridization selection of mRNA with single-stranded DNAs derived from phage M13 recombinants m1, m2, m3 and m4 (Figure 1C). Only the DNAs of clones m1 and m3 were complementary to mRNA of the 160K gene. This demonstrated that the direction of transcription of this gene was towards the central portion of the viral DNA.

Northern blot analyses were used to determine the size of transcripts of the 160K gene. A 2.0-kb AccI fragment contained in plasmid p2031 (Figure 1C) was used as the hybridization probe, because its map coordinates (2.2 kb from each end of the 6.4-kb ClaI fragment) were within the predicted 4.3-kb coding region of the 160K gene. Hybridization of this probe to polyadenylated RNA extracted from cells 24 h after CPV infection demonstrated that a late RNA of defined length (4.5 kb) was transcribed from this region (Figure 4). This 4.5-kb RNA was not detected in RNA extracted from cells in the early stages of CPV infection (data not shown).

Nucleoside S1 protection experiments were used to map the extent of complementarity between the 4.5-kb RNA and the viral DNA. Figure 5 shows that complementarity between the DNA and the 3'-end of the RNA extended to a position about 1.25 kb to the left of the 3'-end labeled EcoRI site shown in Figure 1. Complementarity between the DNA and the 5'-end of the RNA extended to a position 4.4 kb to the right of the end-point of complementarity to the 3'-end of the RNA (data not shown). The position of the transcribed region of the 160K gene is shown in Figure 1B.

Nucleotide sequence analysis

Figure 6 shows the nucleotide sequence of the coding strand of the DNA at the beginning of the transcribed region of the 160K gene. The sequence contains two long open reading-frames (nucleotides 1–644 and nucleotides 693–1021) arranged in tandem. The open reading-frame beginning at nucleotide 693 is contained within the 4.4-kb transcribed region of the 160K gene. The ATG triplet at nucleotide 693 is the only in-frame ATG within 300 bp of the beginning of the transcribed region; it is
also the only ATG capable of acting as the initiation codon of an open reading-frame of the size (4.3 kb) required to encode a 160-kd polypeptide. The 160-kd ATI protein appears to be the primary translation product of the 160K mRNA (Figure 2, lane A), therefore these results suggest that the ATG at nucleotide 693 is the translational initiation codon of the 160K gene. The other open reading-frame terminates only 46 bp upstream of the initiation codon of the 160K gene. This short non-coding region includes a 21-bp sequence containing only A and T residues. The

Fig. 3. Transient expression of the 160K gene. Preconfluent monolayers of human 143 cells were infected with vaccinia virus at a multiplicity of 10 p.f.u./cell. Two hours after infection the cells were transfected with plasmid DNA. Eighteen hours after infection the cells were incubated in the presence of 50 μCi [35S]methionine (1100 Ci/mmol). The cells were harvested 2 h later. Their proteins were solubilized and then resolved by electrophoresis in an 8% polyacrylamide gel. [35S]methionine-labeled proteins were visualized by autoradiography. The lanes show the proteins labeled in: (A) cells that were mock-transfected; (B) cells transfected with DNA of the vector pUC19; (C) cells transfected with DNA of plasmid p2025; and (D) cells that were infected with cowpox virus instead of vaccinia virus. Cells transfected with p2025 (C), and cells infected with cowpox virus (D), contain the 160-kd protein.

Fig. 4. Transcripts of the 160K gene are ~4.5 kb long. Polyadenylated RNA was electrophoresed in a 1% agarose gel containing 2.2 M formaldehyde. The resolved RNAs were transferred to a nylon membrane, and transcripts of the 160K gene were identified by DNA−RNA hybridization. The autoradiogram shows: (A) HindIII fragments of phage lambda DNA, which were used as size markers; and (B) the RNAs detected by hybridization to a probe derived from an Accl fragment present within the transcribed region of the 160K gene.

open reading-frame upstream of the 160K gene contains an unusual sequence; this sequence (nucleotides 171–254) consists of 28 repeats of the triplet GAT. Its position suggests that at least one of its functions is to encode 28 aspartic acid residues. However, its exact functions remain to be determined.

Structure of the 5'-end of the mRNA

The extent of complementarity between the 5'-end of the mRNA of the 160K gene and the viral DNA was determined by high-resolution nuclease S1 protection experiments. In these experiments, reaction conditions similar to those described by Shenk et al. (1975) were used to minimize the removal of base-paired residues at the ends of hybrid duplexes. The results, shown in Figure 7, indicated that the major end-points of complementarity were between three and eight nucleotides upstream of the initiation codon.

In contrast, primer extension analyses (Figure 8, lane 1) indicated that most of the 160K gene's mRNAs contained 5–21 nucleotides upstream of the initiation codon. This suggested that the 5'-ends of the mRNAs contained sequences that were not complementary to the nucleotide sequence of the template strand of the viral DNA.

The nucleotide sequence of the oligonucleotide-primed cDNA was determined by the dideoxyribonucleotide chain-termination method (Figure 8, lanes 2–5). The sequence of the cDNA corresponds to that of the viral DNA except at nucleotide –3 (3 nucleotides upstream of the first nucleotide of the initiation codon) where the cDNA appears to contain a T residue instead of an A residue. The autoradiogram also contains several faint bands; these correspond to an A residue at the –3 position, and various residues 20–60 nucleotides upstream of the initiation codon. The nucleotide sequence determined from these faint bands is identical to that of the viral DNA. This suggests that these minor products were derived from less abundant transcripts that extended from upstream genes into the coding region of the 160K gene.
Northern blot analyses confirmed the existence of such overlapping transcripts (data not shown).

The nucleotide sequence in the region between −5 and −21 was uninterpretable because of the heterogeneity in the length of the mRNAs. Therefore the following method was used to determine the sequence of the cDNAs extended to this region. First, [32P] 5′-end labeled cDNAs that extended to −12 nucleotides upstream of the initiation codon were purified. Then their sequence was determined by the chemical degradation method of Maxam and Gilbert (1980). This analysis (Figure 9) showed that the cDNAs were homogeneous in sequence; more importantly, it showed that the nine residues upstream of the initiation codon were all T residues, whereas the corresponding region of the template strand in the viral DNA contains A residues, not T residues, at positions −3, −6, −8 and −9. This result confirmed that obtained by the deoxyribonucleotide chain-termination method of sequence analysis (Figure 8).

The nuclease S1 protection experiment was modified to demonstrate that the 3′-end oligo (dT) sequences of the cDNAs were derived from poly(A) sequences upstream of the initiation codon of the mRNA. Instead of using viral DNA as the hybridization probe, a chemically synthesized 39-mer oligonucleotide was used. The oligonucleotide was constructed to be complementary to both the putative poly(A) region (containing up to 22 adenylate residues) immediately upstream of the initiation codon, and the first four codons of the mRNA. Five residues (3′-GTACG-5′) were included at the 3′-end of the oligonucleotide to form a non-complementary 3′-tail; this provided a means of assessing the specificity of the nuclease S1 for unprotected regions of the probe.

The results of this nuclease S1 protection experiment are shown in Figure 10. RNA extracted from CPV-infected cells protected 18−33 nucleotides at the 5′-end of the oligonucleotide (Figure 10, lane 1), whereas RNA extracted from uninfected cells failed to provide significant protection of the 5′-end of the oligonucleotide probe (Figure 10, lane 3). This indicated that the protection of the poly(dT) residues by the RNA of CPV-infected cells was due to 5′-terminal poly(A)-containing mRNA of the 160K gene. The poly(dT) portion of the oligonucleotide was also unprotected in control experiments in which the RNA of uninfected cells was supplemented with single-stranded DNA copies of the viral DNA complementary to at least 12 nucleotides at the 5′-end of the oligonucleotide (data not shown). This indicated that the protection of the poly(dT) by the CPV RNA was not a consequence of the possible formation of trinucleotide hybrids (Lopata et al., 1985) containing the 5′-end of the mRNA of the 160K gene, the 3′-terminal poly(A) tail of an mRNA and the poly(dT)-containing oligonucleotide. In addition, the incomplete protection of the short poly(dT) region is inconsistent with the degree of protection that relatively long 3′-terminal poly(A) tails of mRNAs should have been able to provide.

Figure 11 summarizes the results of the analyses of the structure of the 5′-end of the mRNA. It also shows the correlation of these results. The end-points of the primer extensions correspond almost exactly to the end-points of complementarity between the mRNA and the poly(dT)-containing oligonucleotide. The end-points of complementarity between the mRNA and the viral DNA correspond exactly to the region within which the nucleotide sequence of the mRNA begins to differ from that of the viral DNA. All these data indicate that 5′-terminal poly(A) sequences are present immediately upstream of the initiation codons of the mRNAs of the 160K gene.

Discussion

The most remarkable feature of the mRNAs encoding the 160-kd ATI protein is the structure of the 5′-ends of these molecules. Each mRNA contains a 5′-terminal poly(A) sequence immediately upstream of the initiation codon. However, the template strand of the 160K gene does not contain poly(dT) sequences immediately upstream of the initiation codon. Therefore, the 5′-terminal poly(A) sequences are not the products of straightforward transcription of DNA sequences contiguous with the coding region of the gene.
Most of the 5'-terminal poly(A) sequences appear to contain between 5 and 21 residues. The close correspondence between the 5'-ends as determined by the nuclease S1 protection experiments (using the oligonucleotide probe), and the 5'-ends as determined by primer extension analyses, suggests that these are the actual ends of the mRNA. However, the possibility that these ends are the products of either a susceptibility of the RNA to limited nuclease digestion or the premature termination of reverse transcription cannot be ruled out.

The nucleotide sequence of the 160K gene and the nucleotide sequence at the 5'-end of its mRNA diverge at the site of three A residues in the sequence 5'-TAATG-3' containing the gene's initiation codon. This sequence element is conserved at the apparent transcriptional initiation sites of all characterized late genes of vaccinia virus (Weir and Moss, 1984; Bertholet et al., 1985; Cochran et al., 1985b; Rosel and Moss, 1985; Rosel et al., 1986; Weinrich and Hruby, 1986; Roseman and Hruby, 1987). Moreover, part of the conserved 5'-TAATG-3' sequence together with 15-30 nucleotides upstream appear to be required to effect the transcription of these late genes (Cochran et al., 1985b; Bertholet et al., 1986; Hänggi et al., 1986; Weir and Moss, 1987). Therefore, the mechanism that produces the 5'-poly(A)-containing 160K mRNAs may be similar to that producing the mRNAs of most late genes of vaccinia virus.

The structures of the late mRNAs of vaccinia virus (VV) have not been fully characterized, but some of the available data are consistent with existence of 5'-terminal poly(A)-containing mRNAs. Boone and Moss (1977) examined the methylated 5'-terminal sequences (m5G(5')ppNpN2) of late VV RNAs. They found that on average 88 ± 12% of methylated N1 residues were either 2'-O-methyladenosine (A^m) or N6,2'-O-dimethyladenosine (m6A^m), and 73 ± 12% of methylated N2 residues were either A^m or m6A^m. The results of Whitkop et al.'s (1982) analysis of RNase T1-resistant 5'-terminal oligonucleotides of VV RNAs are also consistent with the presence of heterogeneous lengths of poly(A) at the 5'-ends of late mRNAs. Conversely, the results of Sheldon et al. (1972) demonstrating that most of the poly(A) of VV mRNAs are at the RNA's 3'-terminus, appear to be inconsistent with the presence of 5'-terminal poly(A).

Currently there is little information on the 5'-terminal structures of mRNA of individual late VV genes. The structures of most have been inferred from the results of nuclease S1 protection experiments. When attempted, it proved to be difficult to confirm the results of nuclease S1 protection experiments by primer-extension analyses (Rosel and Moss, 1985; Rosel et al., 1986; Weinrich and Hruby, 1986). This difficulty was attributed
to the presence of overlapping and complementary transcripts. Most likely this was not a problem in the analysis of the 5'-ends of the 160K mRNAs because of the abundance of these mRNAs late in the infection. Recently, Bertholet et al. (1987) and Schwer et al. (1987) used additional methods to characterize the 5'-end structures of mRNAs of two strongly-expressed late VV genes. The results of Bertholet et al. (1987) indicated that each mRNA of the 11K gene contained a leader sequence that was not complementary to sequences upstream of the coding region of the gene. These leaders contained poly(A) sequences fused to the initiation codon of the gene. Most leader sequences also contained additional nucleotide sequences (up to several kb long)

---

**Fig. 9.** Nucleotide sequence analysis of cDNA copies of the 5'-end region of the 160K gene's mRNA. The cDNA copies were made by primer extension reactions using the 32P 5'-end-labeled primer described in the legend to Figure 8. The cDNA products were fractionated by electrophoresis in a 10% polyacrylamide gel containing 7 M urea. cDNAs 73–77 nucleotides long (containing 10–14 nucleotides complementary to sequences upstream of the initiation codon) were eluted from the gel, and their nucleotide sequences were determined by base-specific cleavage reactions according to the method of Maxam and Gilbert (1980). The products of these reactions were resolved by electrophoresis in an 8% polyacrylamide gel containing 7 M urea. The autoradiogram of the gel shows (lane 1) the purified cDNA fragments, and (lanes 2–5) the products of base-specific cleavage reactions at G, G+A, C+T and C residues. The nucleotide sequence of the cDNA is shown, where nucleotide +1 is complementary to the first residue of the initiation codon of the mRNA. The nucleotide sequence of the cDNA is identical to that of the template strand of the viral DNA except where indicated by asterisks.

**Fig. 10.** The mRNAs of the 160K gene contain poly(A) leader sequences. Thirty-nine base oligonucleotides were synthesized whose sequence (5'-CGTGACCCACTTTTTTTTTTTTTTTTTTGTG-3') was constructed to be complementary to leader sequences of up to 22 adenosine residues adjacent to the first four codons of the 160K gene's mRNA. The oligonucleotides were 32P 5'-end labeled and hybridized either with RNA from CPV-infected cells or with RNA from uninfected cells. Residual single-stranded nucleic acids were digested with nuclease SI. The nucleic mRNAs resistant products were electrophoresed through a 20% polyacrylamide/7 M urea gel, and labeled DNAs were visualized by autoradiography. The autoradiogram shows: lane 1, portions of the oligonucleotide protected from nuclease SI digestion by hybridization with RNA extracted from CPV-infected cells; lane 2, the oligonucleotide hybridization probe; lane 3, portions of the oligonucleotide protected from nuclease SI digestion by hybridization with RNA extracted from uninfected cells. The nucleotide sequence of the 3'-end of the oligonucleotide is shown where nucleotide +1 is complementary to the first nucleotide of the initiation codon of the 160K gene's mRNA. The arrowheads indicate the major end-points of the complementarity between the oligonucleotides and the RNAs.

**Fig. 11.** Comparison of the results of nucleate SI protection and primer extension analyses of the structure of the 5'-ends of the 160K gene's mRNA. The nucleotide sequence of the DNA template strand is shown, the numbering corresponds to that of its complementary strand as shown in Figure 6. Nucleotides corresponding to the initiation codon of the mRNA are underlined. The arrows above the viral DNA sequence correspond to the major endpoints of complementarity between this template strand of viral DNA and the mRNA (shown below) as determined by nucleate SI analysis. The sequence of the mRNA was derived from nucleotide sequence analysis of cDNA copies of the mRNA. The arrows above the predicted sequence of the mRNA correspond to the major 5'-ends of the mRNA as determined by primer extension analysis. The arrows above the sequence of the oligonucleotide correspond to the major endpoints of complementarity between this DNA and the mRNA. The length of each arrow corresponds to the relative abundance of fragments having endpoints at that site.
mRNAs contain any sequences upstream of the poly(A)-initiation codon sequence. Northern blot analyses (Figure 4) showed that most of the 160K mRNAs are only slightly longer than the transcribed region of the 160K gene. This additional length (~ 100 bp) is most likely attributable to both 5'- and 3'-poly(A) sequences. Conceivably, the long leaders of the 11K mRNAs and the shorter leaders of the 160K mRNAs are generated by different mechanisms. Alternatively, one or other of these RNA structures might represent either a different stage or a different end-product of a similar pathway of synthesis.

The nature of the mechanism that generates the 5'-poly(A)-containing 160K mRNAs has not yet been determined. Possible mechanisms include the following: (i) an RNA processing event; (ii) a novel process in which a poly(A)-containing RNA primes transcription of the late gene; and (iii) the repetitive transcription of three T residues in the conserved sequence 3'-ATTTCG-5' (template strand) prior to elongation of the transcript through the coding region. The first possibility, an RNA processing event, would explain the structure of the RNAs described by Berthalet et al. (1987). However, currently there is little evidence that 5'-terminal poly(A) is attached to the 160K mRNA in this manner. The second possibility, a mechanism involving RNA-primed transcription, might have limited similarity to the primed-transcription processes of the coronavirus (Spaan et al., 1983; Baric et al., 1983) or the influenza viruses (Plotch et al., 1981). The third possibility, the repetitive transcription of T residues prior to elongation, appears to occur during transcription from certain late promoters of the bacteriophage T4. The T4-modified RNA polymerase of _E. coli_ may undergo abortive transcriptional initiation. When this occurs at three T residues, multiple rounds of (i) slippage of the _pppApApA_ product and (ii) subsequent elongation of the product may then generate transcripts containing 5'-poly(A) sequences (Kassavetis et al., 1986). Such 5'-terminal poly(A) sequences are extensions of the A triplet encoded at the transcriptional start-site. As noted above, all characterized late genes contain a sequence similar if not identical to the sequence 5'-ATTTCG-3' (template strand) at their putative transcriptional initiation sites. Therefore a similar mechanism might generate 5'-terminal poly(A) in late poxvirus mRNAs.

Whatever the mechanism of 5'-polyadenylation, the transient expression assay (Figure 3, lane C) shows that the 160K gene is transcribed from transfected plasmid DNA. This implies that a late promoter element exists in this DNA. However, this element has not yet been shown to be within the cloned fragment of poxvirus DNA.

The role of the 5'-terminal poly(A) sequences is unclear. One role could be to enhance the translational efficiency of the mRNA. Kozak (1981, 1984, 1986) demonstrated that the sequences flanking the initiation codon affect the efficiency of translational initiation. The presence of a purine residue three nucleotides upstream of the initiation codon exerts a dominant positive effect. Therefore, the 5'-poly(A) sequence should enhance the translation of the 160K mRNA by providing a purine three nucleotides upstream of the initiation codon. The influence of the additional residues in the poly(A) leader remains to be determined.

Another possible role of the 5'-terminal poly(A) might be to increase the stability of the late mRNA. Rice and Roberts (1983) showed that the degradation of cellular mRNAs is induced by infection with vaccinia virus. Perhaps the late viral mRNAs are protected from this rapid virus-induced degradation by the 5'-terminal poly(A) sequences.

Lastly, the 160K mRNAs have one other unusual structural feature. Each has a 3'-end that corresponds to a single position just downstream of the open reading-frame. In contrast, all previously described late mRNAs of _VV_ genes have heterogeneous 3'-ends (Cooper et al., 1981; Mahr and Roberts, 1984; Weir and Moss, 1984). It is not yet clear how these late transcripts of the 160K gene acquire their defined 3'-ends.

Further studies on the 160K gene and its mRNAs should provide additional insights into the mechanisms regulating the expression of poxvirus genes.

### Materials and methods

**Cells and viruses**

Strain 143 human osteosarcoma cells were grown in monolayer culture in growth medium consisting of MEME medium (GIBCO) supplemented with 5% fetal calf serum. The Brighton Red strain of cowpox virus and the Western Reserve strain of vaccinia virus were used in this study.

**Purification of RNA**

Total cellular RNA was extracted according to the method described by Cox (1968). Cells were washed with phosphate buffered saline, and lysed by the addition of 6 M guanidinium HCl in 0.1 M NaOAc pH 5.0 (at 4°C). To shear the DNA, the lysate was passed through a 22-gauge syringe needle four times. One-half volume of absolute ethanol was added to precipitate the RNA (2 h at -20°C), which was then collected by centrifugation, washed with 70% ethanol, and resuspended in water treated with diethyl pyrocarbonate. To obtain the viral transcripts synthesized in the late stages of infection, RNA was extracted from cells 24 h after they had been infected with CPV at 10 p.f.u./cell. Unless otherwise indicated, all viral RNAs referred to in this paper were extracted from cells in this late stage of infection.

**Northern blot analyses**

RNA that was used in the analysis of the 5'-terminus of 160K mRNA was purified further by centrifugation (38 000 r.p.m./12 h/SW 41) through a 2.5 M CsCl cushion of 5.7 M guanidinium HCl (Gisin et al., 1974). Polyadenylated RNA was purified by oligo(dT)-cellulose chromatography (Aviv and Leder, 1972).

**Identification of the gene encoding the 160-kd AT1 protein**

Various cloned fragments of CPV DNA (Pickup et al., 1984) were used in the hybridization-selection of viral mRNAs. Phage M13 vectors mp18 and mp19 (Norlander et al., 1983) were used to obtain single strands of the cloned DNAs. Hybridization-selection of mRNAs and their in vitro translation were done by standard methods (Ricciardi et al., 1979; Jackson and Hunt, 1983). Antigen-purified antibody against the 160-kd protein of the CPV ATIs (Patel et al., 1986) was used according to the method of Kessler (1975) to immunoprecipitate the proteins in the _in vitro_ translation mixture. The immunoprecipitated proteins were resolved by polyacrylamide gel electrophoresis (PAGE), and then visualized by autoradiography as previously described by Patel et al. (1986). The method of Cochran et al. (1985a) was used to achieve the transient expression of cloned copies of the 160K gene.

**Characterization of the DNA containing the 160K gene**

Restriction endonuclease cleavage sites were mapped by standard methods. The methods of Sanger et al. (1980) and Biggin et al. (1983) were used to determine the sequence of both strands of the 1021-bp EcoRI-Fnu 4H1 fragment containing the initiation codon of the 160K gene. Oligonucleotides used for primers and hybridization probes were produced on an Applied Biosystems 380A DNA synthesizer by the phosphoramidite method of Beaucage and Carruthers (1981). The oligonucleotides were purified by reverse-phase HPLC (Fritz et al., 1978). Hybridization analysis of DNA immobilized on nylon membranes was performed by standard methods (Southern, 1975; Maniatis et al., 1982).

**Northern blot analysis of viral RNA**

Purified, polyadenylated RNA was electrophoresed in formaldehyde-containing agarose gels (Boedter, 1971), and then transferred to nylon membranes by the method of Southern (1975). DNA fragments derived from the coding region of the 160K gene were labeled with [32P]dATP by nick-translation and used as hybridization probes for transcripts of the 160K gene. Hybridization conditions were as described by Maniatis et al. (1982). Hybridized DNAs were visualized by autoradiography.

**Nuclease S1 protection experiments**

The extent of complementarity between the transcripts of the 160K gene and the viral DNA was determined by nuclease S1 protection procedures as described by Berk and Sharp (1977). The viral DNA fragments used to detect the 5'-ends of the 160K gene transcripts were 32P-labeled at their 5'-ends according to the method of Weaver and Weissman (1979). DNA fragments used to detect the 3'-ends of the mRNAs were 32P-labeled at their 3'-ends according to the method of Drouin (1980). The denatured probes were hybridized with 25 μg of RNA.
in 30 μl of a buffer containing 0.4 M NaCl, 40 mM Pipes (pH 6.4), 0.1 mM EDTA, and 80% formamide, at 37°C for 16 h. Residual single-stranded nucleic acids were digested with nuclease S1 (40 units/ml) in 300 μl of S1 buffer [0.25 M NaCl, 30 mM NaOAc (pH 4.5), 1 mM ZnSO4 and 5% (v/v) glycerol] at 37°C for 60 min (Vogt, 1973). The ethanol-precipitated products were then electrophoresed in either 1% neutral agarose gels or 7 M urea—polyacrylamide gels (Maxam and Gilbert, 1977). The 32P-labeled fragments which were electrophoresed in agarose gels were visualized by autoradiography after being transferred to nylon membranes (Southern, 1975). The 32P-labeled fragments in acrylamide gels were visualized by autoradiography of the dried gels.

The nature of the 5’-end of the transcript of the 160K gene was characterized by nuclease S1 protection experiments using a chemically-synthesized oligonucleotide (5’-CTGACCTCAGTTTGTTCGT-3’) as a hybridization probe. The 5’-end of a 39-mer oligonucleotide was labeled and its 5’-end was 32P labeled by PAGE.

**References**

Archard,L.C., Mackett,M., Barnes,D.E. and Dumbell,K.R. (1984) J. Gen. Virol., 65, 875–886.

Aviv,H. and Leder,P. (1972) Proc. Natl. Acad. Sci. USA, 69, 1408–1412.

Bacic,R.S., Shieh,K.K., Stohlman,S.A. and Lai,M.M.C. (1987) Virology, 156, 342–354.

Beaucage,S.L. and Caruthers,M.H. (1981) Tetrahedron Lett., 22, 1859–1862.

Berk,A.J. and Kozak,M. (1985) J. Biol. Chem., 260, 1617–1624.

Berg,King,B., Sauer,R.T. and Novick,P. (1975) Proc. Natl. Acad. Sci. USA, 72, 633–637.

Birnstiel,M.L., Setlow,R.D., Erlich,H.A. and Smith,J.M. (1985) Nucleic Acids Res., 13, 3909–3911.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.

Bouloy,M., Ulmanen,I. and Krug,R.M. (1981) Cell, 23, 847–858.