

## 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 inclusion 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'-TAAATG-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 most 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'-TAAAT-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.

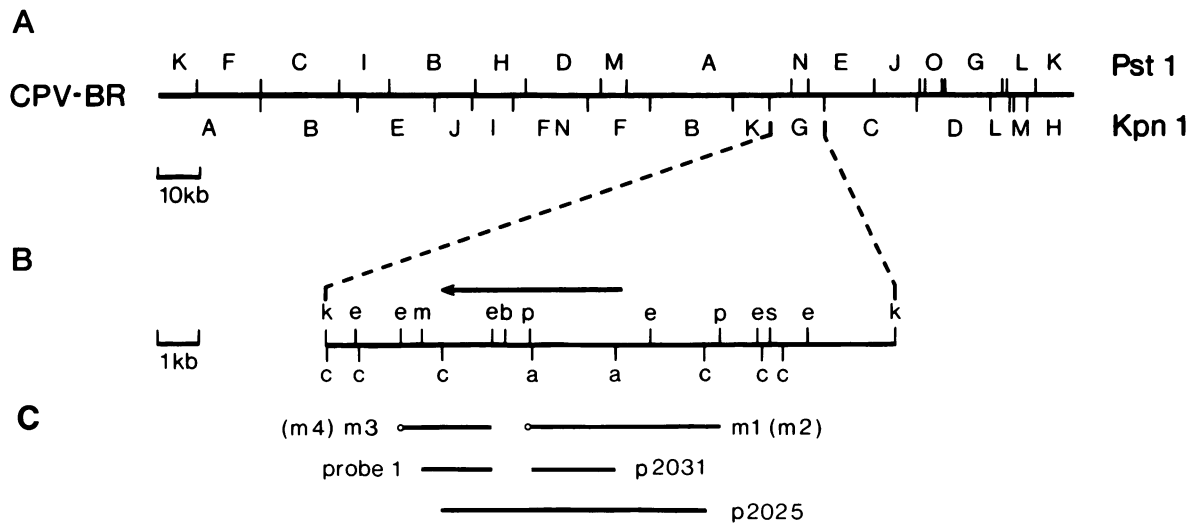
This paper describes the characterization of the gene (designated the 160K gene) that encodes the 160-kd ATI protein. It also describes the unusual structure of this gene's mRNAs.

### 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 *Pst*I and *Kpn*I 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 *Kpn*I G fragment. Subsequent hybridization-selection experiments showed that the complementary region was within a 6.4-kb *Cla*I fragment derived from the *Kpn*I G fragment (Figure 1B). Furthermore, the transfection of plasmid p2025 (containing this *Cla*I 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 *Cla*I 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.*,



**Fig. 1.** The position of the 160K gene in the CPV genome. (A) *Pst*I and *Kpn*I restriction maps of the genome of the Brighton Red strain of CPV (Archard *et al.*, 1984; Pickup *et al.*, 1984). (B) Restriction map of the *Kpn*I G fragment. The arrow represents the direction of transcription of mRNA of the 160K gene; it also represents the position and length of the transcribed region of the 160K gene. Restriction sites are abbreviated as follows: a, *Acc*I; b, *Bam*HI; c, *Cl*aI; e, *Eco*RI; k, *Kpn*I; m, *Msp*I; p, *Pst*I; and s, *Sal*I. The only *Acc*I and *Msp*I sites shown are those used to obtain the DNA fragment corresponding to Probe 1 and also the *Acc*I fragment contained in plasmid p2031. (C) Nature and orientation of the cloned DNAs used to determine both the position of the 160K gene and the structure of its transcripts. Sub-fragments of the *Kpn*I G fragment were inserted into phage M13 vectors and pUC19 plasmid vectors. m1 and m3 correspond to single-stranded DNA subfragments of *Kpn*I G contained in DNAs of M13 vector phages; the open circles indicate the 5'-end orientation of the single-stranded DNAs. The designations of corresponding single-stranded DNA subfragments of opposite orientation are indicated in parentheses. Probe 1 represents the 1.7 kb *Msp*I–*Eco*RI fragment used in the nuclease S1 analysis of the 3'-end of the 160K gene transcripts. Plasmid p2031 contains a 2.0-kb *Acc*I fragment located within the coding region of the 160K gene. Plasmid p2025 contains a 6.4-kb *Cl*aI fragment that spans the coding region of the 160K gene.

1986; Patel *et al.*, 1986; J.Esposito, D.Patel and D.Pickup, unpublished data).

#### Mapping of the transcribed region of the 160K gene

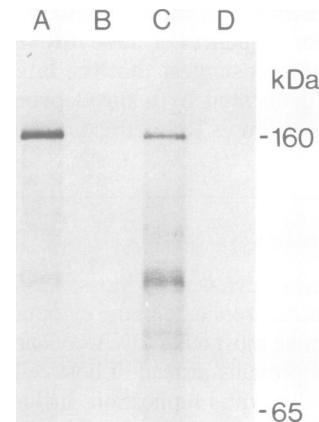
The direction of transcription of the 160K gene was determined by hybridization selection of mRNA with single-stranded DNAs derived from phage M13 recombinants m1, m2, m3 and m4 (Figure 1C). Only the DNAs of clones m1 and m3 were complementary to mRNA of the 160K gene. This demonstrated that the direction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

Nuclease 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 *Eco*RI 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.

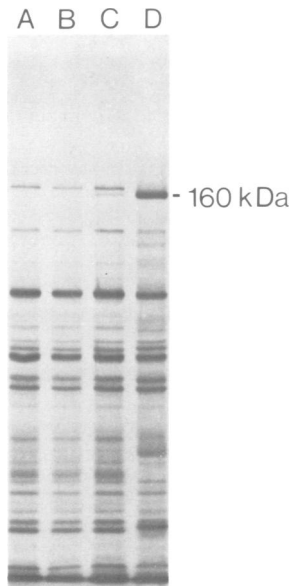
#### Nucleotide sequence analysis

Figure 6 shows the nucleotide sequence of the coding strand of the DNA at the beginning of the transcribed region of the 160K

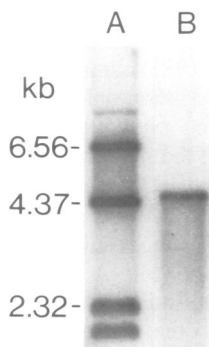


**Fig. 2.** Hybridization-selection of mRNA encoding the 160-kd protein of CPV A-type inclusions. Various cloned fragments of the CPV DNA were used in the hybridization-selection of mRNA synthesized late in cells infected with CPV. The selected RNAs were translated *in vitro* in a rabbit reticulocyte lysate containing [<sup>35</sup>S]methionine. The proteins in the lysate were then immunoprecipitated with antigen-purified antibody against the 160-kd ATI protein. The proteins were electrophoresed in an 8% polyacrylamide gel, and the labeled proteins were visualized by autoradiography. The autoradiogram shows the immunoprecipitated *in vitro* translation products of unselected RNA (lane A), and RNA selected with the *Kpn*I K fragment (lane B), the *Kpn*I G fragment (lane C), and the *Pst*I J fragment (lane D). The translation products of unselected RNA were co-electrophoresed to provide standards for the estimation of molecular masses of the products of the selected mRNAs.

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

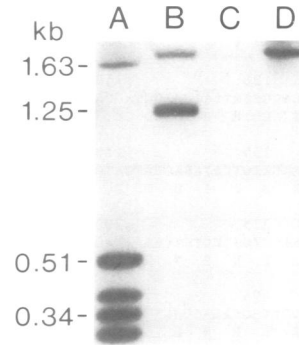


**Fig. 3.** Transient expression of the 160K gene. Preconfluent monolayers of human 143 cells were infected with vaccinia virus at a multiplicity of 10 p.f.u./cell. Two hours after infection the cells were transfected with plasmid DNA. Eighteen hours after infection the cells were incubated in the presence of 50  $\mu$ Ci [ $^{35}$ S]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. [ $^{35}$ S]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) *Hind*III fragments of phage lambda DNA, which were used as size markers; and (B) the RNAs detected by hybridization to a probe derived from an *Acc*I fragment present within the transcribed region of the 160K gene.

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. 5.** Nuclease S1 protection analysis of the extent of complementarity between the 160K gene and the 3'-end of its mRNA. A 1.7-kb *Msp*I-*Eco*RI fragment (probe 1, Figure 1C) was [ $^{32}$ P]-labeled at the 3'-end generated by the *Eco*RI cleavage. This hybridization probe was annealed to RNA extracted from 143 cells 24 h after infection with CPV (10 p.f.u./cell). After hybridization, the residual single-stranded nucleic acids were digested with nuclease S1. The nuclease S1-resistant hybrids were resolved by electrophoresis through a 1% neutral agarose gel. They were then transferred to a nylon membrane, and the [ $^{32}$ P]-labeled fragments were visualized by autoradiography. The autoradiogram shows: (A) [ $^{32}$ P]-end-labeled *Hin*fI restriction fragments of pBR322 DNA; (B) [ $^{32}$ P]-labeled probe DNA protected from digestion by hybridization with RNA from CPV-infected cells; (C) the products of nuclease S1 digestion of probe 1 DNA after hybridization with calf liver tRNA; and (D) the intact probe 1 DNA. The size of the protected fragment (1.25 kb) was estimated from its electrophoretic mobility relative to the mobilities of the *Hin*fI fragments of pBR322 DNA.

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 dideoxynucleotide 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.

```

      15          30          45          60          75          90          105          120
GAATTCATGCTATAAAGACTCATCAACGATAGTACTGATGCTCAACATATACATTTTGGATTTAGAAAATATGGTAATAATAGACAATGAATGCGCTAATATTGAGTCAAGTGCATGAAAATG
I B A I R L I N D S T D A Q H I H F G F R N M V I I D N E C A N I Q S S A E N

      135          150          165          180          195          210          225          240
CAACTGATACAGGACATCATCAAGATAGCAAAAATAATCGAAGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
A T D T G H H Q D S K I N I E V E D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D D
      255          270          285          300          315          330          345          360
ATGATGATGATGATGTCATAGACGATGATGATTATAATCCAAAAACCCACTCCGATACCGGACCCCTACCCCTAGACCCCGTTCACAGACATGAATATCATAAAGAGCCGAAAAGTCTCTC
D D D D D D V I D D D D D Y N P K P T P I P E P H P R P P F F P R H E Y H K R P K V L

      375          390          405          420          435          450          465          480
CTGTAGAAGAACCTGATCCTCTCAAAAAGACGGCGATCGTATAAGACTTGATAATCATATATAAACACAGTTGGATCATAATCTTAATTCGATCGGACACTTGTGTGATAGACGAC
P V E E P D P V K K D A D R I R L D N H I L N T L D H N L N S I G H Y C C D T A

      495          510          525          540          555          570          585          600
CAGTTGATAGTTAGAACATCACATTGAAAACATTGGGACAATATGCGACTAATACTAGCAAGAAAGATAAAATATGCAAAACATTACTGTTCCCATGGCCATTACTCTCCATCCACATG
A V D R L E E H R I E T L G Q Y A V I L A R K I N M Q T L L F P W P L P T V H P H

      615          630          645          660          675          690          705          720
CGATAGATGGTACTATTCCGGCCACATGGGAGATCTACGATCTTATAATACACGATTGTAAGTTTGAATAAAAATTTTTTTATAATAAAATGGAGGTCCAGCAACCTTATTGAAAAT
A I D G S I P P H G R S T I L M E V T N L I E K

      735          750          765          780          795          810          825          840
GTACCAAGCACTCCAAAAGATTTCCGCCACTGAGGTAGAAAACTATGGAAACGATGACTTGAAGTTCTGAATCAGGCTCTCTCAAGAAAAACAGAAATGTAATTCGTAATATTCTTCGTGATA
C T K H S K D F A T E V E K L W N D E L S S E S G L S R K T R N V I R N I L R D

      855          870          885          900          915          930          945          960
TCACTAAGTCACTAACTACAGATAAGAAATCAAAAGTGTTCGCTATACTAGAACGTTCCAGCATTAAACCGGAGCAGATTAAGATGTATATAAACTATTTTTAATAATGGTGTGATG
I T K S L T T D K K S K C F R I L E R S T I N G E Q I K D V Y K T I F N N G V D

      975          990          1005          1020
TGGAGCTCAGAATCAACACTACAGAAACTGTTCTATTACAGTTATGACTTATGCTGC
V E S R I N T T G K Y V L F T V M T Y A

```

Fig. 6. Nucleotide sequence of the regions flanking the initiation codon of the 160K gene. Shown is the 1021 nucleotide sequence of an *EcoRI*-*Fnu* 4H1 segment of the *PstI* N fragment of CPV DNA. The *EcoRI* site (nucleotides 1–6) is 693 nucleotides upstream of the initiation codon of the 160K gene (see Figure 1B). Upstream of the 160K gene is an open reading-frame (nucleotides 1–644) capable of encoding 214 amino acids, 45 of which would be aspartic acid residues. The unusual sequence between nucleotides 171 and 254 consists of 28 direct repeats of the triplet GAT.

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, [<sup>32</sup>P] 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 dideoxynucleotide 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.

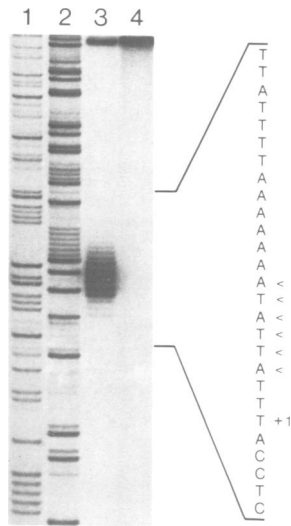
The results of this nuclease S1 protection experiment are shown in Figure 10. RNA extracted from CPV-infected cells protected 18–33 nucleotides at the 5'-end of the oligonucleotide (Figure 10, lane 1), whereas RNA extracted from uninfected cells failed to provide significant protection of the 5'-end 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 trimolecular 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

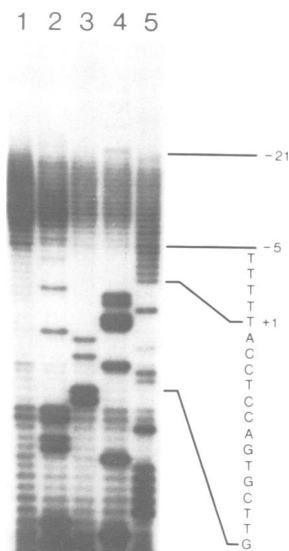
The most remarkable feature of the mRNAs encoding the 160-kd ATI protein is the structure of the 5'-ends of these molecules. Each mRNA contains a 5'-terminal poly(A) sequence immediately upstream of the initiation codon. However, the template strand of the 160K gene does not contain poly(dT) sequences 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.



**Fig. 7.** Complementarity between the 160K gene and its mRNA extends only a few nucleotides upstream of the initiation codon. A *Bgl*III–*Hin*I fragment (nucleotides 630–786, Figure 6) was 5'-end labeled at the *Hin*I cleavage site, and then hybridized with RNA extracted from CPV-infected cells. Residual single-stranded nucleic acids were digested with nuclease S1, and the S1-resistant nucleic acids were resolved by electrophoresis in an 8% polyacrylamide gel containing 7 M urea. The products of dideoxynucleotide chain-termination sequence reactions were coelectrophoresed to provide size markers. Components of the sequence reactions included an oligonucleotide primer (5'-ATTCAGAACTCAAC-3'), m2 single-stranded DNA (template), [ $\alpha$ - $^{35}$ S]dATP, and either dideoxy TTP or dideoxy ATP. Similarly-sized fragments of primer-containing DNA and end-labeled DNA contained the same nucleotide sequence, because the nucleotide sequence of the primer was identical to that at the 5'-end of the hybridization probe. Primer-containing DNAs, which were not phosphorylated at their 5'-ends, migrated through the gel as if they were the equivalent of half of one nucleotide longer than probe DNAs of the same sequence. After electrophoresis the labeled DNAs were visualized by autoradiography of the dried gel. The autoradiogram shows: the products of sequence reactions containing dideoxy TTP (lane 1); the products of sequence reactions containing dideoxy ATP (lane 2); nuclease-S1 resistant DNA (lane 3); and the hybridization probe (lane 4). The sequence of the template DNA strand is shown; +1 corresponds to the nucleotide complementary to the first nucleotide in the gene's initiation codon. The arrowheads indicate the 3'-ends of the protected DNAs.

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'-TAAATG-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'-TAAAT-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



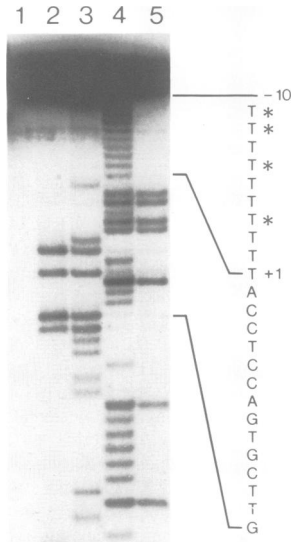
**Fig. 8.** Primer-extension analyses of the 5'-ends of the mRNAs of the 160K gene. [ $^{32}$ P] 5'-end labeled oligonucleotide (5'-CTACCTCAGTGGCGA-3') primers were hybridized with RNA extracted from CPV-infected cells. The oligonucleotides (complementary to sequences between 49 and 63 nucleotides downstream of the initiation codon) were extended by AMV reverse transcriptase in reactions containing dNTPs either without dideoxy NTPs (lane 1), or with dideoxy ATP (lane 2), dideoxy GTP (lane 3), dideoxy CTP (lane 4) and dideoxy TTP (lane 5). The primer-extended products were resolved in an 8% polyacrylamide gel containing 7 M urea, and then visualized by autoradiography. The nucleotide sequence derived from the sequence reactions is shown; nucleotide +1 is complementary to the first nucleotide of the initiation codon. The major 3'-ends of the primer-extended products are between 5 and 21 nucleotides upstream of the initiation codon. Minor products longer than these are derived from the extension of primers hybridized to the overlapping transcripts of the gene upstream of the 160K gene.

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 RNAs 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 ( $m^7$ G(5')pppN<sub>1</sub>pN<sub>2</sub>) of late VV RNAs. They found that on average  $88 \pm 12\%$  of methylated N<sub>1</sub> residues were either 2'-*O*-methyladenosine ( $A^m$ ) or  $N^6,2'$ -*O*-dimethyladenosine ( $m^6A^m$ ), and  $73 \pm 12\%$  of methylated N<sub>2</sub> residues were either  $A^m$  or  $m^6A^m$ . The results of Whitkop *et al.*'s (1982) analysis of RNase T<sub>1</sub>-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). However, given the methods used, 5'-terminal poly(A), containing on average only 10–15 residues, would have been more difficult to detect than 3'-terminal poly(A), which contains on average 100 residues (Nevins and Joklik, 1975).

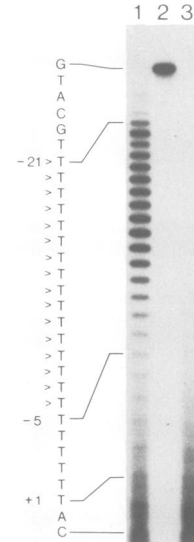
Currently there is little information on the 5'-terminal structures of mRNA of individual late VV genes. The structures of most have been inferred from the results of nuclease S1 protection experiments. When attempted, it proved to be difficult to confirm the results of nuclease S1 protection experiments 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 <sup>32</sup>P 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.

upstream of the poly(A) sequences. The results of Schwer *et al.* (1987) indicated that the mRNAs of both the 11K gene and the 4b gene contained 5'-terminal poly(A) sequences (of ~35 residues) fused to the coding region of each gene. Therefore, with respect to the poly(A)-initiation codon sequence, there is considerable similarity between the VV 11K mRNA, the VV 4b mRNA and the CPV 160K mRNA. In each the nucleotide sequence of the mRNA and the corresponding DNA diverge at the site of the three A residues in the conserved DNA sequence 5'-TAAATG-3'. However, there is no evidence that the 160K



**Fig. 10.** The mRNAs of the 160K gene contain poly(A) leader sequences. Thirty-nine base oligonucleotides were synthesized whose sequence (5'-CGTGACCTCCATTTTTTTTTTTTTTTTTTTTTTTTGCATG-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 <sup>32</sup>P 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 S1. The nuclease S1 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 S1 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 S1 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 nuclease S1 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 nuclease S1 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'-ATTTAC-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 Bertholet *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 coronaviruses (Spaan *et al.*, 1983; Baric *et al.*, 1987) 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'-ATTTAC-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 virus

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.

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-ml cushion of 5.7 M CsCl (Gliisin *et al.*, 1974). Polyadenylated RNA was purified by oligo(dT)-cellulose chromatography (Aviv and Leder, 1972).

### Identification of the gene encoding the 160-kd ATI 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 Caruthers (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 (Boedtker, 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 [<sup>32</sup>P]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 <sup>32</sup>P-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 <sup>32</sup>P-labeled at their 3'-ends according to the method of Drouin (1980). The denatured probes were hybridized with 25 µg of RNA



in 30  $\mu$ l 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  $\mu$ l of S1 buffer [0.25 M NaCl, 30 mM NaOAc (pH 4.5), 1 mM ZnSO<sub>4</sub> and 5% (v/v) glycerol] at 20°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 <sup>32</sup>P-labeled fragments which were electrophoresed in agarose gels were visualized by autoradiography after being transferred to nylon membranes (Southern, 1975). The <sup>32</sup>P-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 as a hybridization probe. The 5'-end of a 39-mer oligonucleotide (5'-CGTGACCTCCATTTTTTTTTTTTTTTTTTTTTTTCATG-3') was labeled with <sup>32</sup>P. This probe was hybridized with 20  $\mu$ g of RNA in 20  $\mu$ l of a buffer containing 0.3 M NaCl, 10 mM Tris-HCl (pH 8.0), 5 mM EDTA and 50% formamide at 60°C for 16 h. Residual single-stranded nucleic acids were digested with nuclease S1 (75 units/ml, at 37°C for 30 min) in 200  $\mu$ l of S1 buffer. The products were electrophoresed in a 20% polyacrylamide gel containing 7 M urea, and the labeled fragments were visualized by autoradiography.

#### Primer-extension analyses

Primer extension analyses (Ghosh *et al.*, 1978) including nucleotide sequence determination (Hamlyn *et al.*, 1978) were used to characterize the 5'-ends of the transcripts of the 160K gene. Oligonucleotide primers (5'-CTACCTCA-GTGGCGA-3') complementary to the region between 49 and 63 nucleotides downstream of the initiation codon of the mRNA were labeled with [<sup>32</sup>P] at their 5'-ends. The primers were annealed to CsCl-purified RNA as described by Lane *et al.* (1985). They were then extended under the following conditions: 50  $\mu$ M dNTPs, 75 mM KCl, 50 mM Tris-HCl (pH 8.3), 10 mM MgCl<sub>2</sub>, 10 mM DTT and 400 units/ml reverse transcriptase from avian myeloblastosis virus (Boehringer Mannheim Biochemicals) at 42°C for 30 min. In order to determine the nucleotide sequence of the RNA these conditions were modified by the addition of either 10  $\mu$ M dideoxy ATP, 10  $\mu$ M dideoxy GTP, 10  $\mu$ M dideoxy CTP or 10  $\mu$ M dideoxy TTP. The reactions containing dideoxynucleotides were completed by the addition of dNTPs to 200  $\mu$ M, and incubation at 42°C for 30 min. Products of these reactions were electrophoresed in 7 M urea-polyacrylamide gels, and then visualized by autoradiography.

The sequence at the 3'-ends of the primer-extended products was determined as follows. The 5'-end labeled products of primer-extension reactions were resolved by PAGE. Products of specific lengths were eluted from the gel and then their nucleotide sequences were determined by the chemical degradation method of Maxam and Gilbert (1980).

#### Acknowledgements

We thank Carol Ray for expert technical assistance, and Elaine Veazey for help in preparation of the manuscript. This work was supported by grant R01 AI232886 from the National Institutes of Health. D.D.P. was supported by Medical Scientist Training Program No. 5T32 GM0717.

#### 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.
- Baric, R.S., Shieh, C.-K., Stohman, 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 Sharp, P.A. (1977) *Cell*, **12**, 721–732.
- Bertholet, C., Drillien, R. and Wittek, R. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 2096–2100.
- Bertholet, C., Stocco, P., Van Meir, E. and Wittek, R. (1986) *EMBO J.*, **5**, 1951–1957.
- Bertholet, C., Van Meir, E., ten Heggeler-Bordier, B. and Wittek, R. (1987) *Cell*, **50**, 153–162.
- Biggin, M.D., Gibson, T.J. and Hong, G.F. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3963–3965.
- Boedtke, H. (1971) *Biochim. Biophys. Acta*, **240**, 448–453.
- Boone, R.F. and Moss, B. (1977) *Virology*, **79**, 67–80.
- Cochran, M.A., Mackett, M. and Moss, B. (1985a) *Proc. Natl. Acad. Sci. USA*, **82**, 19–23.
- Cochran, M.A., Puckett, C. and Moss, B. (1985b) *J. Virol.*, **54**, 30–37.
- Cooper, J.A., Wittek, R. and Moss, B. (1981) *J. Virol.*, **37**, 284–294.
- Cox, R.A. (1968) *Methods Enzymol.*, **12B**, 120–129.
- Drouin, J. (1980) *J. Mol. Biol.*, **140**, 15–34.
- Esposito, J.J. and Knight, J.C. (1985) *Virology*, **143**, 230–251.
- Fritz, H.-J., Belagaje, R., Brown, E.L., Fritz, R.H., Jones, R.A., Lees, R.G. and Khorana, H.G. (1978) *Biochemistry*, **17**, 1257–1267.
- Ghosh, P.K., Reddy, V.B., Swinscoe, J., Lebowitz, P. and Weissman, S.M. (1978) *J. Mol. Biol.*, **126**, 813–846.
- Glisin, V., Crkvenjakov, R. and Byus, C. (1974) *Biochemistry*, **13**, 2633–2637.
- Hamlyn, P.H., Brownlee, G.G., Cheng, C.-C., Gait, M.J. and Milstein, C. (1978) *Cell*, **15**, 1067–1075.
- Hänggi, M., Bannwarth, W. and Stunnenberg, H.G. (1986) *EMBO J.*, **5**, 1071–1076.
- Jackson, R.J. and Hunt, T. (1983) *Methods Enzymol.*, **96**, 50–74.
- Kassavetis, G.A., Zentner, P.G. and Geiduschek, E.P. (1986) *J. Biol. Chem.*, **261**, 14256–14265.
- Kessler, S.W. (1975) *J. Immunol.*, **115**, 1617–1624.
- Kitamoto, N., Tanimoto, S., Hiroi, K., Tanaka, T., Miyamoto, H., Wakamiya, N., Ikuta, K., Ueda, S. and Kato, S. (1986) *Arch. Virol.*, **89**, 15–28.
- Kozak, M. (1981) *Nucleic Acids Res.*, **9**, 5233–5252.
- Kozak, M. (1984) *Nature*, **308**, 241–246.
- Kozak, M. (1986) *Cell*, **44**, 283–292.
- Lane, D.J., Pace, B., Olsen, G.J., Stahl, D.A., Sogin, M.L. and Pace, N.R. (1985) *Proc. Natl. Acad. Sci. USA*, **82**, 6955–6959.
- Lopata, M.A., Sollner-Webb, B. and Cleveland, D.W. (1985) *Mol. Cell. Biol.*, **5**, 2842–2846.
- Mackett, M. and Archard, L.C. (1979) *J. Gen. Virol.*, **45**, 683–701.
- Mahr, A. and Roberts, B.E. (1984) *J. Virol.*, **49**, 510–520.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Maxam, A.M. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA*, **74**, 560–564.
- Maxam, A. and Gilbert, W. (1980) *Methods Enzymol.*, **65**, 499–560.
- Moss, B. (1985) In Fields, B.N. (ed.), *Virology*. Raven Press, New York, pp. 685–703.
- Nevins, J.R. and Joklik, W.K. (1975) *Virology*, **63**, 1–14.
- Niles, E.G., Condit, R.C., Caro, P., Davidson, K., Matusick, L. and Seto, J. (1986) *Virology*, **153**, 96–112.
- Norlander, J., Kempe, T. and Messing, J. (1983) *Gene*, **26**, 101–106.
- Oda, K. and Joklik, W.K. (1967) *J. Mol. Biol.*, **27**, 395–419.
- Patel, D.D., Pickup, D.J. and Joklik, W.K. (1986) *Virology*, **149**, 174–189.
- Pickup, D.J., Ink, B.S., Parsons, B.L., Hu, W. and Joklik, W.K. (1984) *Proc. Natl. Acad. Sci. USA*, **81**, 6817–6821.
- Plotch, S.J., Bouloy, M., Ulmanen, I. and Krug, R.M. (1981) *Cell*, **23**, 847–858.
- Plucieniczak, A., Schroeder, E., Zettlmeissl, G. and Streeck, R.E. (1985) *Nucleic Acids Res.*, **13**, 985–998.
- Puckett, C. and Moss, B. (1983) *Cell*, **35**, 441–448.
- Ricciardi, R.P., Miller, J.S. and Roberts, B.E. (1979) *Proc. Natl. Acad. Sci. USA*, **76**, 4927–4931.
- Rice, A.P. and Roberts, B.E. (1983) *J. Virol.*, **47**, 529–539.
- Rosel, J. and Moss, B. (1985) *J. Virol.*, **56**, 830–838.
- Rosel, J.L., Earl, P.L., Weir, J.P. and Moss, B. (1986) *J. Virol.*, **60**, 436–449.
- Roseman, N.A. and Hruby, D.E. (1987) *J. Virol.*, **61**, 1398–1406.
- Sanger, F., Coulson, A.R., Barrel, B.G., Smith, A.J.H. and Roe, B.A. (1980) *J. Mol. Biol.*, **143**, 161–178.
- Schwer, B., Visca, P., Vos, J.C. and Stunnenberg, H.G. (1987) *Cell*, **50**, 163–169.
- Sheldon, R., Kates, J., Kelley, D.E. and Perry, R.P. (1972) *Biochemistry*, **11**, 3829–3834.
- Shenk, T.E., Rhodes, C., Rigby, P.W.J. and Berg, P. (1975) *Proc. Natl. Acad. Sci. USA*, **72**, 989–993.
- Southern, E.M. (1975) *J. Mol. Biol.*, **98**, 503–517.
- Spaan, W., Delius, H., Skinner, M., Armstrong, J., Rottier, P., Smeekens, S., Van der Zeust, B.A.M. and Siddell, S.G. (1983) *EMBO J.*, **2**, 1839–1844.
- Vogt, V.M. (1973) *Eur. J. Biochem.*, **33**, 192–200.
- Weaver, R.F. and Weissman, C. (1979) *Nucleic Acids Res.*, **7**, 1175–1193.
- Weinrich, S.L. and Hruby, D.E. (1986) *Nucleic Acids Res.*, **14**, 3003–3016.
- Weir, J.P. and Moss, B. (1984) *J. Virol.*, **51**, 662–669.
- Weir, J.P. and Moss, B. (1987) *J. Virol.*, **61**, 75–80.
- Whitkop, C., Lipinskas, B.R., Mercer, S., Panicali, D. and Paoletti, E. (1982) *J. Virol.*, **42**, 734–741.

Received on July 10, 1987; revised on August 20, 1987