JOURNAL OF VIROLOGY, Oct. 1986, p. 320–323
0022-538X/86/100320-04$02.00/0
Copyright © 1986, American Society for Microbiology
Vol. 60, No. 1

Multiple 3′ Ends of mRNA Encoding Vaccinia Virus Growth Factor Occur within a Series of Repeated Sequences Downstream of T Clusters

LEONARD YUEN AND BERNARD MOSS*

Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20892

Received 16 April 1986/Accepted 23 June 1986

Analysis of the 5′ ends of six apparently full-length cloned cDNA copies of the vaccinia virus growth factor gene suggested precise transcriptional initiation at the first purine following a run of five pyrimidines in the noncoding strand. By contrast, the 3′ ends exhibited heterogeneity and were distributed over a 46-base-pair region. In each of the six cDNAs, the nucleotide immediately preceding the retained copy of the poly(A) tail corresponded to the first or second T of a TATGT repeat. Clusters of Ts occurred upstream of the 3′ ends, but the AATAAA polyadenylation consensus sequence of higher eucaryotes was absent. Determination of the complete sequence of one cDNA revealed exact correspondence with the vaccinia virus growth factor gene, indicating the absence of internal RNA processing.

Vaccinia virus, the prototypal member of the poxvirus family, has a 185-kilobase-pair double-stranded DNA genome (for a review, see reference 18). A transcription system, packaged within the virus particle, is responsible for the formation, capping, and polyadenylation of early mRNAs. Recent studies demonstrated that the RNA polymerase has virus-encoded subunits (7, 17; E. V. Jones, C. Puckett, and B. Moss, manuscript in preparation) and recognizes virus-specific DNA signals for transcription of early genes (8, 11, 16, 19, 21). Nevertheless, at least one (7) and probably additional subunits have significant sequence homology with eucaryotic RNA polymerase polypeptides.

There are about 100 early genes dispersed through the vaccinia virus genome, several of which have been transcriptionally mapped and sequenced (7, 10, 12, 26, 27). Attention has been directed primarily to the RNA start sites and the promoter sequences present upstream. By contrast, there is less information available regarding the precise location of the 3′ ends of mRNAs, their degree of heterogeneity, and their mechanism of formation. The diversity of vaccinia virus early transcripts precludes direct isolation and biochemical characterization. cDNA cloning, however, provides a unique way of obtaining information concerning individual mRNAs within a population.

The primary purpose of the present study was to prepare a cDNA library of full-length early transcripts and characterize several that correspond to a single gene. We chose to analyze cDNA copies of mRNA that encodes the vaccinia virus growth factor (VGF). This gene is located within the inverted terminal repetition and was transcriptionally mapped and sequenced several years ago (26). Recently, it was found to be structurally (3, 6, 20) and functionally (14, 23, 24) related to epidermal and transforming growth factors of higher eucaryotes. The physical properties of the secreted protein suggest that VGF undergoes processing steps, including proteolytic cleavage and glycosylation (23). In this study, we analyzed the 5′ and 3′ ends of six full-length VGF cDNA clones and determined the complete sequence of one.

A cDNA library was prepared by oligo(dT)-primed reverse transcription of polyadenylated RNA from the cytoplasm of HeLa cells that had been infected with 30 PFU of vaccinia virus (strain WR). Cycloheximide, a protein synthesis inhibitor, was added at 100 μg/ml to amplify early transcripts and prevent progression to the late phase of the replicative cycle. Cytoplasmic RNA was isolated at 4 h postinfection, as described previously (9). Polyadenylated RNA was purified by oligo(dT)-cellulose column chromatography, and 30 μg was used as a template for oligo(dT)-primed reverse transcription and cDNA cloning (15).

If there is equal efficiency of cDNA cloning, the relative abundance of mRNA species can be estimated by screening the library with appropriate hybridization probes. When total vaccinia virus DNA was used as the probe, approximately 50% of the colonies were scored as positive. Although seemingly high, this ratio of viral to cellular mRNA is consistent with previous estimates made with saturation hybridization techniques (5). The library was also screened with probes for some previously sequenced early vaccinia virus genes. Of 2,500 colonies, 16 were positive when probed with a 250-base-pair (bp) *Acc*I-*Dde*I DNA segment containing the 3′ terminus of the VGF gene (26), whereas only 1 was positive when probed with the thymidine kinase gene. The low number of thymidine kinase cDNAs is in agreement with data suggesting that this gene has a relatively weak promoter (1, 16).

Eight of the sixteen VGF cDNA clones were chosen for further analysis. Agarose gel electrophoresis of the *Pst*I-digested plasmids indicated that six of them (clones 2-1, 2-3, 2-7, 2-8, 3-1, and 3-9) contained possible full-length inserts longer than 600 bp (Fig. 1). Previous nuclease S1 analysis indicated that the VGF transcript is about 530 nucleotides long. The additional lengths of the cDNAs were attributed to the poly(A) and synthetic homopolymer tails.

The inserts from the six potentially full-length cDNAs were excised from the plasmids and recloned in bacteriophage M13 to facilitate dideoxynucleotide sequencing. We initially thought that the sequence of the ends could be readily determined by using a universal M13 oligonucleotide primer. Reliable data were not obtained, however, because of the inefficiency with which DNA polymerase traversed the homopolymeric tails. To circumvent this problem, we sequenced the opposite strand of DNA. A synthetic 18-mer

* Corresponding author.

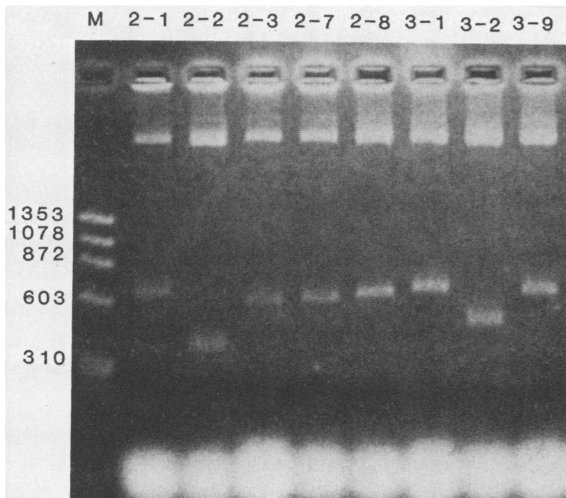

FIG. 1. Analysis of sizes of cloned VGF cDNAs. Plasmid minipreparations of eight cDNA clones (2-1, 2-2, 2-3, 2-7, 2-8, 3-1, 3-2, and 3-9) were digested with *Pst*I and analyzed by electrophoresis in a 1% agarose gel. *Hae*III-digested φX174 replicative form DNA was used as length markers. The size of marker (M) DNA fragments are given in base pairs.

oligonucleotide complementary to a region just upstream of the termination codon of the VGF gene was used as a primer for sequencing the six cDNAs. Each cDNA has a poly(A) tail which serves to mark the location of the 3′ ends of the transcripts (Fig. 2). Three of the cDNAs, 2-3, 2-7, and 3-9, have the same 3′ terminal sequence, with a T residue adjacent to the poly(A). That T is indicated by the first 3′ mark in Fig. 3. Since the next nucleotide in the viral DNA is an A, it is impossible to determine whether the RNA terminated with T or A. The T residue just 2 nucleotides downstream is adjacent to the poly(A) tract of cDNA 3-1. In that case there is no ambiguity since the next nucleotide in the viral DNA is a G (Fig. 3). The remaining two cDNAs, 2-8 and 2-1, also have T residues adjacent to the poly(A) (Fig. 3), and again there is an adjacent A residue in the viral DNA. Inspection of the 3′ noncoding sequence in Fig. 3 reveals multiple TGTAT and TATGT repeats. Remarkably, the final nucleotide of each cDNA corresponds to the first or second T of a TATGT repeat.

Analysis of the 5′ ends of the cDNAs indicated a high degree of specificity. In each case, the same A nucleotide abutted the poly(G) tail added during the cDNA cloning (data not shown). The absence of secondary structure or an unusual sequence upstream of this A suggests that it represents the 5′ end of the mRNA rather than a block to reverse transcription. Furthermore, the 5′ end of the VGF mRNA was located at this position by nuclease S1 analysis, and examination of the capped end has indicated that the 5′ terminal nucleotide is predominantly A (26). Since vaccinia virus mRNAs initiate exclusively with purines (4, 13), the presence of five consecutive pyrimidines upsteam of the 5′ end nucleotide of VGF mRNA (Fig. 3) may have contributed to the precision of transcription initiation. By contrast, previous analysis of the capped 5′ ends of another early mRNA, encoding a 7,500-dalton polypeptide, indicated microheterogeneity (25).

Although no previous studies of vaccinia virus mRNAs indicated the occurrence of splicing, the availability of cDNAs permitted a more rigorous analysis in this case. Most (six of eight) cDNA clones examined had inserts that corresponded to the predicted full-length VGF transcript. (We presume that the smaller inserts resulted from cDNAs with missing 5′ or 3′ sequences, but we have not checked this.) In addition, cDNA 2-8 was sequenced completely with synthetic oligonucleotide primers. The sequence extending from the 5′ end to the third 3′ end in Fig. 3 was identical to that previously obtained by analysis of cloned genomic DNA (26). The absence of splicing indicates that the loss of predicted N-terminal and C-terminal sequences in mature VGF (23) results from processing at the protein level.

The primary reason for performing this study was to determine the exact 3′ ends of a vaccinia virus early mRNA.

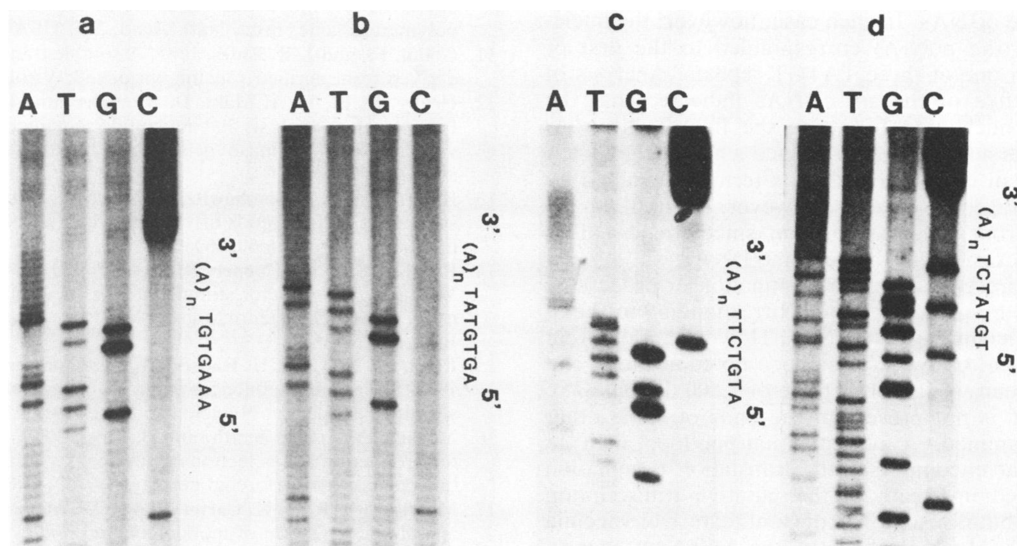

FIG. 2. Sequences of the 3′ ends of VGF cDNAs. cDNA inserts were excised from the plasmids and recloned into the *Pst*I site of M13mp19. The 3′ terminal nucleotides of six apparently full-length VGF cDNA clones were determined by dideoxynucleotide sequencing (22) with the oligonucleotide CAAGATATGGTTGTGCCATA as primer. The seven nucleotides adjacent to the poly(A) tails are shown. (a) Identical results obtained with clones 2-3, 2-7, and 3-9; (b) clone 3-1; (c) clone 2-8; (d) clone 2-1.

```
TTTTTAACAGCAAACACATTCAATATTGTATTGTTATTTTTATGTATTATTTACACAATT
                                                 5'
AACAATATATTATTAGTTTATATTACTGAATTAATAATATAAAATTCCCAATCTTGTCAT
AAACACACACTGAGAAACAGCATAAACACAAAATCCATCAAAA[ATG]TCGATGAAATATCT
GATGTTGTTGTTCGCTGCTATGATAATCAGATCATTCGCCGATAGTGGTAACGCTATCGA
AACGACATCGCCAGAAATTACAAACGCTACAACAGATATTCCAGCTATCAGATTATGCGG
TCCAGAGGGAGATGGATATTGTTTACACGGTGACTGTATCCACGCTAGAGATATTGACGG
TATGTATTGTAGATGCTCTCATGGTTATACAGGCATTAGATGTCAGCATGTAGTATTAGT
AGACTATCAACGTTCAGAAAACCCAAACACTACAACGTCATATATCCCATCTCCCGGTAT
TATGCTTGTATTAGTAGGCATTATTATTATTACGTGTTGTCTATTATCTGTTTATAGGTT
CACTCGACGAACTAAACTACCTATACAAGATATGGTTGTGCCA[TAA]TTTTTATAAATTTT
                                    3'3'   3'
TTTATGAGTATTTTTACAAAAAAAATGTATAAAGTGTATGTCTTATGTATATTTATAAAA
                     3'
ATGCTAAGTATGCGATGTATCTATGTTATTTGTATTTATCTAAACAATACCTCTACCTCT
```

FIG. 3. Sequence of VGF gene and location of the 5′ and 3′ ends of the mRNAs. The previously determined (26) sequence of the VGF gene is shown. The position of the 5′ terminal nucleotide of six cDNA clones was determined by dideoxynucleotide sequencing with the oligonucleotide AGCGAACAACATCAG. The 3′ ends were positioned from data shown in Fig. 2. The translation initiation and termination codons are boxed. Runs of T residues downstream of the stop codon are in italics. TGTAT repeats are underlined, and TATGT repeats are underscored by dots.

Although nuclease S1 analysis provides useful information, it is difficult to distinguish between true heterogeneity and technical artifacts resulting from incomplete digestion or nibbling of DNA in RNA-DNA hybrids. By contrast, sequencing of individual cDNA clones can yield a less ambiguous result provided the poly(A) tail is reverse transcribed and retained. Previously (26), we attempted to perform an analysis of this type; however, very few of the clones had poly(A) tails. Since then, improvements in cDNA cloning methods have made it possible to obtain a high percentage of full-length inserts. The data presented here indicate that the 3′ ends of VGF mRNA occur 50 to 100 bp downstream of the translational stop codon of the six cDNAs examined. The nucleotide adjacent to the copy of the poly(A) tail was identical in three cDNAs. In each case, however, the nucleotide preceding the poly(A) corresponded to the first or second T within one of three TATGT repeats. Analysis of the results for five of the six cDNAs indicated that the genomic DNA nucleotide following the final T is an A. Therefore, it was not possible to deduce whether the first A of the poly(A) tail is transcribed or is terminally added by poly(A) polymerase. In one case, however, the first A must have resulted from terminal addition since the last T is followed by a G residue in the genomic DNA.

If TATGT represents a unique termination or processing signal, it should be present at sites corresponding to the 3′ ends of other vaccinia virus mRNAs. However, a different repeated sequence, CTATTC, has been noted at the 3′ end of an mRNA specifying a polypeptide of 7,500 daltons (25), and the TATGT is not present at the ends of some other early genes examined. A potential hairpin loop with 15 paired bases that encompasses the translation termination codon was noted upstream of the putative transcription termination site but is also not a constant feature of vaccinia virus early genes. Most significantly, the AATAAA eucaryotic consensus polyadenylation sequence (2) does not precede the 3′ end of VGF mRNA. Instead, about 50 bp upstream are three clusters of T residues. Similar T clusters were found in the 3′ noncoding regions of several other early genes, and evidence that these clusters are the essential elements of the termination signal will be presented elsewhere.

We thank J. Baldick for preparation of oligonucleotides, P. Earl for assistance in computer analysis, and G. Rohrmann and N. Elango for helpful discussions.

LITERATURE CITED

1. **Bajszár, G., R. Wittek, J. P. Weir, and B. Moss.** 1983. Vaccinia virus thymidine kinase and neighboring genes: mRNAs and polypeptides of wild-type virus and putative nonsense mutants. J. Virol. **45:**62–72.
2. **Birnsteil, M. L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3′ processing: the end is in site! Cell **41:**349–359.
3. **Blomquist, M. C., L. T. Hunt, and W. C. Barker.** 1984. Vaccinia virus 19-kilodalton protein: relationship to several mammalian proteins, including two growth factors. Proc. Natl. Acad. Sci. USA **81:**7363–7367.
4. **Boone, R. F., and B. Moss.** 1977. Methylated 5′-terminal sequences of vaccinia virus mRNA species made at early and late times after infection. Virology **79:**67–80.
5. **Boone, R. F., and B. Moss.** 1978. Sequence complexity and relative abundance of vaccinia virus mRNA's synthesized in vivo and in vitro. J. Virol. **26:**554–569.
6. **Brown, J. P., D. R. Twardzik, H. Marquardt, and G. J. Todaro.** 1985. Vaccinia virus encodes a polypeptide homologous to epidermal growth factor and transforming growth factor. Nature (London) **313:**491–492.
7. **Broyles, S. S., and B. Moss.** 1986. Homology between RNA polymerases of poxviruses, prokaryotes and eukaryotes: nucleotide sequence and transcriptional analysis of vaccinia virus genes encoding 147-kDa and 22-kDa subunits. Proc. Natl. Acad. Sci. USA **83:**3141–3145.
8. **Cochran, M. A., C. Puckett, and B. Moss.** 1985. In vitro mutagenesis of the promoter region for a vaccinia virus gene: evidence for tandem early and late regulatory signals. J. Virol. **54:**30–37.
9. **Cooper, J. A., and B. Moss.** 1979. *In vitro* translation of immediate early, early and late classes of RNA from vaccinia virus-infected cells. Virology **96:**368–380.
10. **Earl, P. L., E. V. Jones, and B. Moss.** 1986. Homology between DNA polymerases of poxviruses, herpesviruses, and adenoviruses: nucleotide sequence of the vaccinia virus DNA polymerase gene. Proc. Natl. Acad. Sci. USA **83:**3659–3663.
11. **Golini, F., and J. R. Kates.** 1985. A soluble transcription system derived from purified vaccinia virions. J. Virol. **53:**205–213.
12. **Hruby, D. E., R. A. Maki, D. B. Miller, and L. A. Ball.** 1983. Fine structure analysis and nucleotide sequences of the vaccinia virus thymidine kinase gene. Proc. Natl. Acad. Sci. USA **80:**3411–3415.
13. **Keith, J. M., A. Gershowitz, and B. Moss.** 1980. Dinucleotide sequences at the 5′ ends of vaccinia virus mRNA's synthesized in vitro. J. Virol. **36:**601–605.
14. **King, C. S., J. A. Cooper, B. Moss, and D. R. Twardzik.** 1986. Vaccinia growth factor stimulates tyrosine protein kinase activity of A431 cell epidermal growth factor receptors. Mol. Cell. Biol. **6:**332–336.
15. **Land, H., M. Grez, H. Hauser, W. Lindenmaier, and G. Schutz.** 1981. 5′-Terminal sequences of eucaryotic mRNA can be cloned with high efficiency. Nucleic Acids Res. **9:**2251–2266.
16. **Mackett, M., G. L. Smith, and B. Moss.** 1984. General method for production and selection of infectious vaccinia virus recombinants expressing foreign genes. J. Virol. **49:**857–864.
17. **Morrison, D. K., J. K. Carter, and R. W. Moyer.** 1985. Isolation and characterization of monoclonal antibodies directed against two subunits of rabbit poxvirus-associated, DNA-directed RNA polymerase. J. Virol. **55:**670–680.
18. **Moss, B.** 1985. Replication of poxviruses, p. 685–703. *In* B. N. Fields, D. M. Knipe, R. M. Chanock, J. L. Melnick, B. Roizman, and R. E. Shope (ed.), Virology. Raven Press,

Publishers, New York.
19. **Puckett, C., and B. Moss.** 1983. Selective transcription of vaccinia virus genes in template dependent soluble extracts of infected cells. Cell **35:**441–448.
20. **Reisner, A. H.** 1985. Similarity between the vaccinia virus 19K early protein and epidermal growth factor. Nature (London) **313:**801–803.
21. **Rohrmann, G., and B. Moss.** 1985. Transcription of vaccinia virus early genes by a template-dependent soluble extract of purified virions. J. Virol. **56:**349–355.
22. **Sanger, F., A. R. Coulson, B. J. Barrell, A. J. Smith, and H. B. A. Roe.** 1980. Cloning in single-stranded bacteriophage as an aid to rapid DNA sequencing. J. Mol. Biol. **143:**161–178.
23. **Stroobant, P., A. P. Rice, W. J. Gullick, D. J. Cheng, I. A. Kerr, and M. D. Waterfield.** 1985. Purification and characterization of vaccinia virus growth factor. Cell **42:**383–393.
24. **Twardzik, D. R., J. P. Brown, J. E. Ranchalis, G. J. Todaro, and B. Moss.** 1985. Vaccinia virus-infected cells release a novel polypeptide functionally related to transforming and epidermal growth factors. Proc. Natl. Acad. Sci. USA **82:**5300–5304.
25. **Venkatesan, S., B. M. Baroudy, and B. Moss.** 1981. Distinctive nucleotide sequences adjacent to multiple initiation and termination sites of an early vaccinia virus gene. Cell **25:**805–813.
26. **Venkatesan, S., A. Gershowitz, and B. Moss.** 1982. Complete nucleotide sequences of two adjacent early vaccinia virus genes located within the inverted terminal repetition. J. Virol. **44:**637–646.
27. **Weir, J. P., and B. Moss.** 1983. Nucleotide sequence of the vaccinia virus thymidine kinase gene and the nature of spontaneous frameshift mutations. J. Virol. **46:**530–537.