JOURNAL OF VIROLOGY, May 1983, p. 530–537
0022-538X/83/050530-08$02.00/0

Vol. 46, No. 2

Nucleotide Sequence of the Vaccinia Virus Thymidine Kinase Gene and the Nature of Spontaneous Frameshift Mutations

JERRY P. WEIR AND BERNARD MOSS*

Laboratory of Biology of Viruses, National Institute of Allergy and Infectious Diseases, Bethesda, Maryland 20205

Received 5 January 1983/Accepted 15 February 1983

Nucleotide sequencing of a 1,300-base-pair vaccinia virus DNA segment previously shown to contain a thymidine kinase (TK) gene revealed an uninterrupted reading frame of 177 codons capable of producing a polypeptide with a molecular weight of 20,102. Mapping of the TK mRNA by primer extension indicated a unique 5′ end that precedes the initiation codon by only six nucleotides. Multiple 3′ ends within a 10-nucleotide region, about 30 nucleotides beyond the termination codon, were located by nuclease digestion of DNA-RNA hybrids, and the length of the TK transcript, exclusive of polyadenylate, was estimated to be approximately 570 nucleotides. The region preceding the TK mRNA start site is extremely A+T rich and has sequence homologies with three other early genes. Genetic information is so compressed in this region of the DNA that the putative transcriptional regulatory sequence of the TK gene overlaps the coding sequence of a late gene. Only nine nucleotides separate the termination codon of the late gene from the initiation codon of the TK gene. Downstream, 66 nucleotides separate the TK termination codon from the apparent initiation codon of another early gene. The nature of three independent TK^- mutants was revealed by nucleotide sequencing. Each has a nucleotide reiteration leading to a +1 frameshift and a nonsense codon downstream. The location of one frameshift mutation provided evidence that the first ATG is used for initiation of protein synthesis.

Poxviruses are large DNA viruses that replicate within the cytoplasm of infected cells, encode enzymes for RNA and DNA synthesis, and regulate the expression of about 200 genes in a temporally coordinated fashion (19). The ability to readily select thymidine kinase-positive (TK^+) and TK^- mutants makes this gene particularly suitable for detailed analysis (7). Marker rescue experiments demonstrated that the TK gene is contained within the 5,000-base-pair (bp) *Hin*dIII-J fragment of the 187,000-bp linear, double-stranded DNA genome (27). This location was confirmed by synthesis of the active enzyme in cell-free extracts programmed with mRNA selected by hybridization to immobilized DNA (11, 27). Further studies indicated that the TK gene is bisected by a unique *Eco*RI site within the *Hin*dIII-J fragment (1) and that the insertion of exogenous DNA at that site leads to the formation of TK^- recombinant virus (16). TK mRNA is about 590 nucleotides long and codes for a 19,000 (19K)-molecular-weight polypeptide (1, 11). A minor transcript of 2,380 nucleotides contains the same 5′ end but extends through the next gene downstream (1). The absence of the 19K polypeptide from the cell-free translation products of hybridization-selected mRNAs from several TK^- vaccinia virus mutants provided an independent identification of the TK polypeptide (1). The small size of the vaccinia virus TK distinguishes it from the corresponding polypeptides of procaryotes (20), eucaryotes (8), and herpesvirus (10).

We now report the nucleotide sequence of the vaccinia virus TK gene, the precise location of the mRNA start site, and the occurrence of A+T-rich putative transcriptional regulatory signals within the coding sequence of a late gene immediately upstream. In addition, nucleotide sequences of three spontaneous TK^- mutants in each case revealed a frameshift due to a single nucleotide reiteration.

MATERIALS AND METHODS

Recombinant plasmids and phages. Plasmids derived from pBR322, containing either the *Hin*dIII-J fragment or *Hin*dIII-J subfragments of wild-type vaccinia virus strain WR have been described previously (1, 2). Similar procedures were used to clone the *Hin*dIII-J fragments of spontaneous TK^- mutants of vaccinia virus (1). Various restriction endonuclease sites were used to clone small subfragments of *Hin*dIII-J into the single-stranded DNA phage cloning vectors M13mp8

and M13mp9 (18). For this purpose, DNA fragments were purified from agarose gels by glass extraction (25) or by electrophoresis onto DEAE-paper (28).

DNA sequencing. The dideoxynucleotide chain termination method (21) was used with a recombinant single-stranded phage M13 DNA template and a 15-base single-stranded DNA primer.

Mapping the ends of mRNA by primer extension and single-strand nuclease digestion. Cytoplasmic RNA, obtained by Dounce homogenization of HeLa cells infected 4 h earlier with 15 PFU of vaccinia virus in the presence of 100 μg of cycloheximide per ml, was purified by CsCl centrifugation as described previously (6). Primer extension was carried out by a modification of previous methods (3, 9). Approximately 5 to 10 ng of an *Rsa*I-*Eco*RI restriction fragment was denatured by heating for 5 min at 100°C in 5 μl of 80% formamide–10 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid), pH 6.4]–0.4 M NaCl. The denatured primer was incubated with 40 μg of RNA in the same buffer for 45 min at 68°C. After ethanol precipitation, the material was dissolved in 5 μl of 50 mM NaCl–34 mM Tris-hydrochloride (pH 8.3)–6 mM $MgCl_2$–5 mM dithiothreitol, and 1 μl of [α-^{32}P]dATP (800 Ci/mmol, 10 μCi/μl) was added. Of this mixture, 9 μl was mixed with 6 μl of 0.125 mM dGTP-dCTP-dTTP–1.75 mM Tris-hydrochloride (pH 7.5)–1.75 mM $MgCl_2$–12.5 mM NaCl. After 15 min at 37°C, 3 μl of 0.5 mM dATP was added, and the incubation was continued for 15 min more.

The 3′ end of the TK mRNA was mapped by single-strand nuclease digestion of RNA-DNA hybrids. An *Eco*RI-*Hpa*II restriction fragment was labeled with [α-^{32}P]dATP, using the Klenow fragment of DNA polymerase and dTTP to fill in the recessed *Eco*RI site. The end-labeled fragment was mixed with RNA in 10 ml of 90% formamide–40 mM PIPES (pH 6.4)–0.4 M NaCl and heated to 90°C for 5 min and then at 42°C for 3 h. Unhybridized DNA was digested with mung bean nuclease in 0.03 M sodium acetate (pH 4.5)–0.25 M NaCl–1 mM $ZnSO_4$–5% glycerol at room temperature for 60 min. After ethanol precipitation, RNA-DNA hybrids were dissolved in electrophoresis buffer containing 80% formamide and heated at 100°C for 5 min.

Materials. Restriction endonucleases were purchased from Bethesda Research Laboratories or New England Biolabs. Mung bean nuclease and dideoxynucleotides were from P-L Biochemicals, Inc. Avian myeloblastosis virus reverse transcriptase and DNA polymerase were supplied by Boehringer Mannheim Corp., and radioactive nucleotides were from Amersham Corp.

Computer analysis. Routine management of nucleotide sequences was performed on an IBM 370 computer, using a program described by Queen and Korn (14). An addition to this program (15) was used to search for homologies among different sequences. Codon frequences were tabulated on a Dec 10 computer, using the SEQ program (5). An addition to this program, SEQ·TRANS (G. Evans, personal communication) was used to print the nucleotide and derived amino acid sequences in a compact form.

RESULTS

Nucleotide sequence of the TK gene. A restriction endonuclease map of the left portion of the *Hin*dIII-J fragment and the location of the TK gene are shown in Fig. 1A. Arrows underneath the map summarize the strategy used for cloning DNA fragments in M13 phage and sequencing them by the dideoxynucleotide chain termination method. Both DNA strands and overlapping fragments were sequenced to minimize errors. The nucleotide sequence of the 1,300-bp fragment is shown in Fig. 1B.

Location of the ends of the TK mRNA. A primer extension method was used to precisely locate the 5′ end of the TK mRNA. Early RNA from the cytoplasm of cells infected with vaccinia virus in the presence of cycloheximide, an inhibitor of protein synthesis, was hybridized to a 220-bp *Rsa*I-*Eco*RI fragment and extended with reverse transcriptase. The product of this reaction was analyzed by polyacrylamide gel electrophoresis alongside a sequence ladder prepared by extension of the same primer hybridized to single-stranded genomic DNA. The cDNA extension product migrated as a single band alongside a C residue (Fig. 2) corresponding to the indicated G residue of the noncoding DNA strand shown in Fig. 1. This result was consistent with previous single-strand S1 nuclease protection experiments which placed the 5′ end about 285 bp to the left of the *Eco*RI site (1). Similar S1 nuclease experiments placed the 3′ end about 305 bp to the right of the *Eco*RI site (1). This was confirmed by electrophoresis of mung bean nuclease-digested RNA-DNA hybrids next to a sequence ladder. Multiple bands between nucleotides 1,060 and 1,070 were obtained (Fig. 1).

Derived amino acid sequence. Inspection of the nucleotide sequence indicated that the first possible translational initiation codon occurs six nucleotides after the start of the TK message and is followed by a 176-codon open reading frame (Fig. 1). The second ATG occurs in phase 12 codons further downstream; however, only the first is flanked by nucleotides characteristic of those found near functional eucaryotic initiation sites (13). Other ATG sequences are located too far downstream to code for a polypeptide of the requisite size and are in reading frames with numerous stop codons. Additional data indicating translation of the first AUG are presented below. Although a six-base untranslated leader sequence for the TK mRNA is quite short, a similar size has been found for another vaccinia virus mRNA as well (24). The 177-codon open reading frame ends with TAA in an A+T-rich region about 30 nucleotides before the 3′ end of the TK message (Fig. 1). A molecular weight of 20,102 for the TK polypeptide, calculated from the deduced amino acid sequence, corresponds closely to the value of 19K determined by polyacrylamide gel electrophoresis of the TK poly-

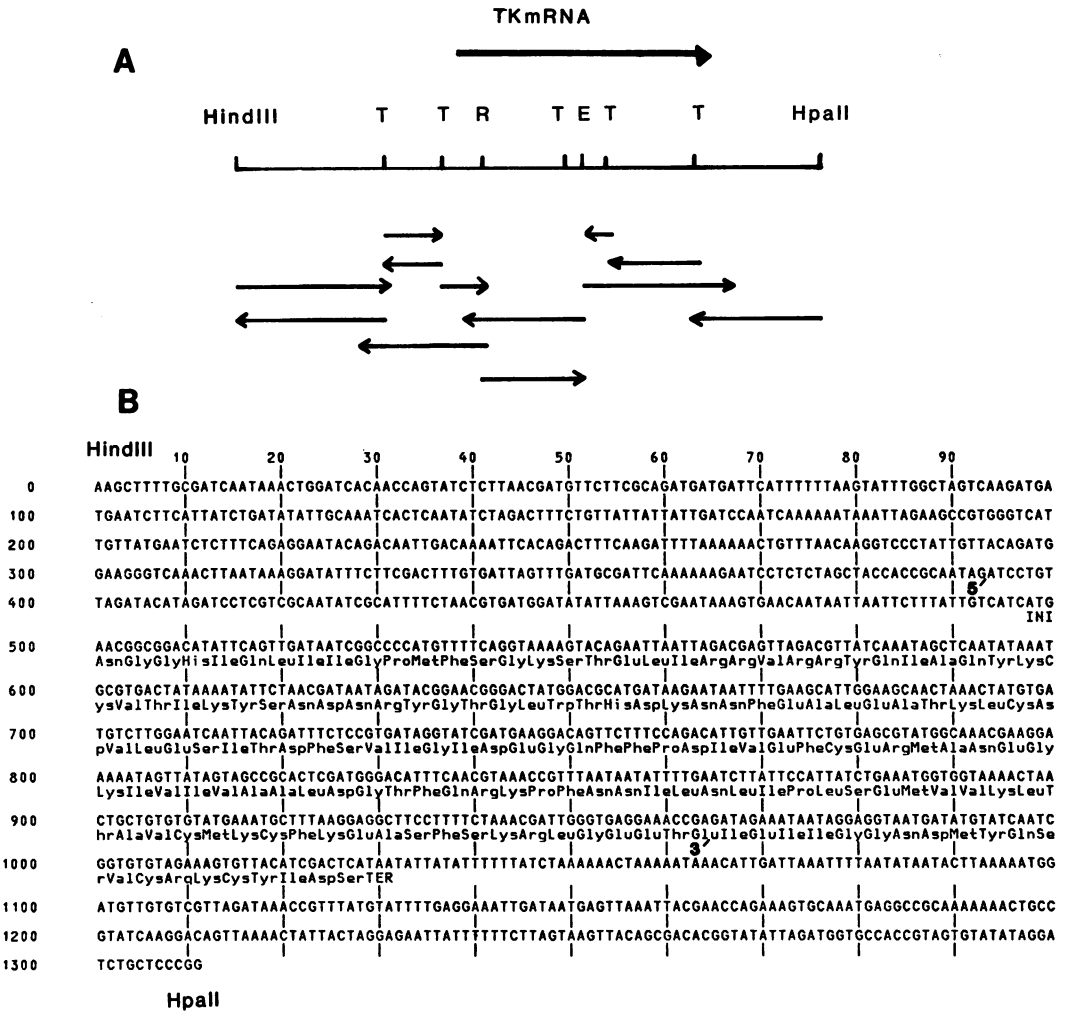

FIG. 1. Nucleotide sequence of the vaccinia virus TK gene and flanking regions. A 1,300-bp DNA segment of the *Hin*dIII-J fragment, extending from the left *Hin*dIII site to the *Hpa*II site, was digested with several restriction endonucleases and subcloned in the single-stranded DNA phages M13mp8 and M13mp9. (A) The dideoxynucleotide chain termination method was used for sequencing the segments indicated. The arrows show the direction and extent of sequence determination from each restriction site. Restriction enzyme cleavage sites are indicated as follows: T, *Taq*I; R, *Rsa*I; E, *Eco*RI. (B) The nucleotide sequence obtained, including the locations of *Hin*dIII and *Hpa*II sites, the 5′ and 3′ ends of the TK mRNA, and the derived amino acid sequence of the TK polypeptide. INI, Initial methionine residue; TER, termination codon.

peptide synthesized in reticulocyte cell-free extracts (1, 11).

Inspection of the nucleotide sequence preceding the start of the TK message revealed it to be extremely A+T rich. The 44-bp sequence immediately upstream of the TK gene is 84% A+T (Fig. 1). Sequences with similar base compositions precede three other early vaccinia virus genes examined (23, 24), and detailed comparisons are presented below. Remarkably, the A+T-rich putative regulatory sequence preceding the TK gene lies within the coding region of an adjacent gene. An examination of the nucleotide sequence (Fig. 1) revealed an open translational reading frame extending from the *Hin*dIII site to within 12 nucleotides of the start of the TK message. Previous mapping studies (1) and additional unpublished work indicate that this open reading frame represents the distal half of a late gene encoding a polypeptide of 28K and that the corresponding late mRNA overlaps the TK gene.

Nucleotide sequences of TK$^-$ mutants. One useful characteristic of the TK gene is the ability

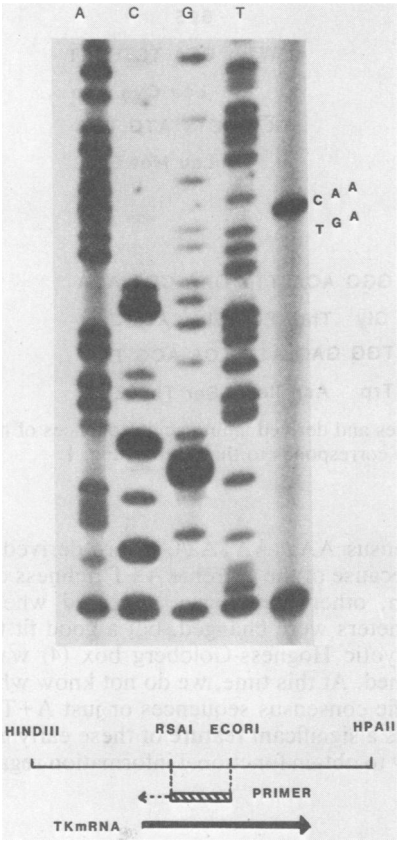

FIG. 2. Mapping of the 5' ends of TK mRNA by primer extension. Purified cytoplasmic RNA from cells infected with vaccinia virus in the presence of cycloheximide was hybridized to a 220-bp *Rsa*I-*Eco*RI fragment. The primer was extended with avian myeloblastosis virus reverse transcriptase in the presence of [α-^{32}P]dATP and unlabeled dGTP, dCTP, and dTTP. The same *Rsa*I-*Eco*RI fragment was used as a primer for dideoxynucleotide sequencing of a genomic *Hin*dIII-*Eco*RI fragment cloned in M13mp8. An autoradiograph of the polyacrylamide gel is shown. The extreme right lane contains the extended primer corresponding to the indicated C residue.

to readily select spontaneous TK$^-$ mutants by plaque formation in the presence of 5-bromodeoxyuridine (7). Three such mutants have been found to make normal amounts of TK mRNA that does not direct synthesis of a 19K TK polypeptide in reticulocyte extracts (1). Preliminary experiments indicated that mutants TK$^-$ 1 and TK$^-$ 20 were rescued by the left *Hin*dIII-*Eco*RI segment of the *Hin*dIII-J fragment of wild-type vaccinia virus DNA, whereas TK$^-$ 16 was rescued by the right *Eco*RI- *Hin*dIII fragment. To locate precisely the site of mutation, the *Hin*dIII-J fragment of each mutant was first cloned in a plasmid vector, and the TK gene was sequenced as described for the wild-type DNA (Fig. 1). In each case, a mutation consisting of the addition of a single nucleotide identical to one preceding it was found (Fig. 3). Frameshifts caused by this nucleotide reiteration are shown in Fig. 4.

DISCUSSION

The nucleotide sequence of the vaccinia virus TK gene revealed an open reading frame of 177

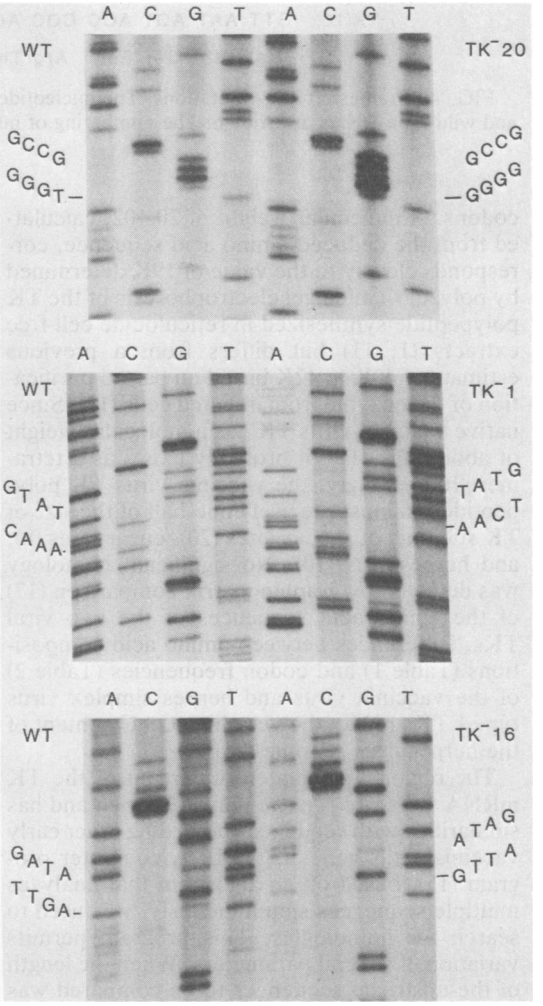

FIG. 3. Nucleotide sequencing TK genes of TK$^-$ mutants. The nucleotide sequences of cloned DNA fragments of wild-type virus (WT), TK$^-$ 20, TK$^-$ 1, and TK$^-$ 16 were determined as described in the legend to Fig. 1. Autoradiographs of segments containing the sites of mutation are shown. By reference to Fig. 1, the template strand of DNA from mutants TK$^-$ 20, TK$^-$ 1, and TK$^-$ 16 have an extra G between nucleotides 533 and 534, an extra T between nucleotides 694 and 695, and an extra A between nucleotides 810 and 811, respectively.

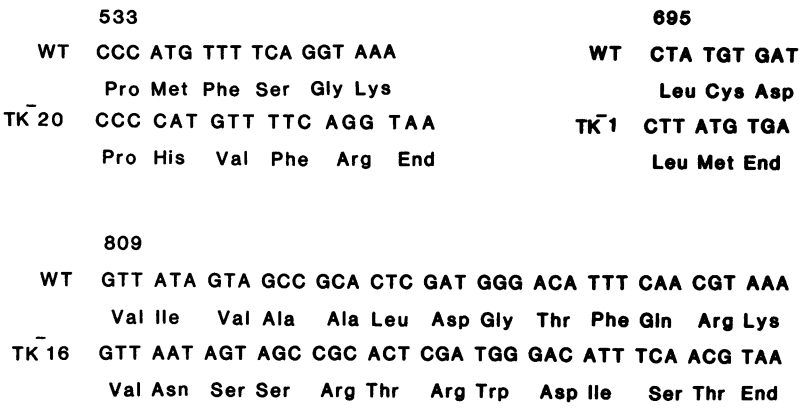

FIG. 4. Frameshift TK⁻ mutations. The nucleotide sequences and derived amino acid sequences of mutant and wild-type DNAs are shown. The numbering of nucleotides corresponds to that used in Fig. 1.

codons. A molecular weight of 20,102, calculated from the deduced amino acid sequence, corresponds closely to the value of 19K determined by polyacrylamide gel electrophoresis of the TK polypeptide synthesized in reticulocyte cell-free extracts (1, 11) but differs from a previous estimate of 40K to 42K based on partial purification of the enzyme from infected cells (12). Since native vaccinia virus TK has a molecular weight of about 80K (12), it probably exists as a tetramer. Interestingly, the vaccinia virus TK polypeptide is almost precisely one-half of the size of TK subunits of procaryotes (20), eucaryotes (8), and herpesvirus (10). No significant homology was detected by graphic matrix comparison (17) of the amino acid sequences of the two viral TKs. Differences between amino acid compositions (Table 1) and codon frequencies (Table 2) of the vaccinia virus and herpes simplex virus type 1 TKs reflect the very high G+C content of the herpesvirus genome (26).

The region immediately upstream of the TK mRNA start site is extremely A+T rich and has similarities with sequences preceding other early vaccinia virus genes (23, 24). A computer program (15), based on an algorithm that analyzes multiple sequences simultaneously, was used to search for homologies. This program permits variation of several parameters. When the length of the upstream sequences to be compared was set to 100 nucleotides, the length of the homology to 6 nucleotides, the number of mismatches between homologous sequences to 1 or less, and the maximum difference in the relative position of homologous sequences to 1, a hexanucleotide consensus sequence was found near position −40, and three overlapping hexanucleotides were found near position −13. By allowing additional A-T substitutions, an octanucleotide consensus TATAATa/tA and a nanonucleotide consensus AAa/tAATAa/tA were derived (Fig. 5). Because of the extreme A+T richness of this region, other homologies appeared when the parameters were changed, but a good fit to the eucaryotic Hogness-Goldberg box (4) was not obtained. At this time, we do not know whether specific consensus sequences or just A+T richness is a significant feature of these early genes. Partly to obtain functional information regarding

TABLE 1. Derived amino acid compositions of vaccinia virus and herpesvirus TKs

Amino acid	No. (%) in TK of:	
	Vaccinia virus	Herpes virus[a]
Ala	8 (4.5)	46 (12.2)
Arg	9 (5.1)	28 (7.4)
Asn	10 (5.6)	8 (2.1)
Asp	9 (5.1)	19 (5.0)
Cys	7 (3.9)	4 (1.1)
Gln	6 (3.4)	19 (5.0)
Glu	14 (7.9)	14 (3.7)
Gly	13 (7.3)	30 (8.0)
His	2 (1.1)	12 (3.2)
Ile	18 (10.1)	16 (4.2)
Leu	12 (6.7)	41 (10.9)
Lys	12 (6.7)	4 (1.1)
Met	6 (3.4)	13 (3.4)
Phe	10 (5.6)	8 (2.1)
Pro	4 (2.2)	30 (8.0)
Ser	10 (5.6)	17 (4.5)
Thr	9 (5.1)	28 (7.4)
Trp	1 (0.6)	5 (1.3)
Tyr	6 (3.4)	12 (3.2)
Val	11 (6.2)	22 (5.8)
Total	177	376

[a] Data from reference 26.

TABLE 2. Codon frequency of vaccinia virus and herpesvirus TKs

Virus TK	Codon-amino acid	No. (%)	Codon-amino acid	No. (%)	Codon-amino acid	No. (%)	Codon-amino acid	No. (%)
Vaccinia	TTT-Phe	7 (3.9)	TCT-Ser	3 (1.7)	TAT-Tyr	4 (2.2)	TGT-Cys	5 (2.8)
	TTC-Phe	3 (1.7)	TCC-Ser	2 (1.1)	TAC-Tyr	2 (1.1)	TGC-Cys	2 (1.1)
	TTA-Leu	2 (1.1)	TCA-Ser	3 (1.7)	TAA-	1 (0.6)	TGA-	0 (0.0)
	TTG-Leu	5 (2.8)	TCG-Ser	1 (0.6)	TAG-	0 (0.0)	TGG-Trp	1 (0.6)
	CTT-Leu	1 (0.6)	CCT-Pro	0 (0.0)	CAT-His	2 (1.1)	CGT-Arg	3 (1.7)
	CTC-Leu	1 (0.6)	CCC-Pro	1 (0.6)	CAC-His	0 (0.0)	CGC-Arg	0 (0.0)
	CTA-Leu	3 (1.7)	CCA-Pro	2 (1.1)	CAA-Gln	4 (2.2)	CGA-Arg	2 (1.1)
	CTG-Leu	0 (0.0)	CCG-Pro	1 (0.6)	CAG-Gln	2 (1.1)	CGG-Arg	0 (0.0)
	ATT-Ile	6 (3.4)	ACT-Thr	3 (1.7)	AAT-Asn	7 (3.9)	AGT-Ser	1 (0.6)
	ATC-Ile	3 (1.7)	ACC-Thr	1 (0.6)	AAC-Asn	3 (1.7)	AGC-Ser	0 (0.0)
	ATA-Ile	9 (5.1)	ACA-Thr	3 (1.7)	AAA-Lys	9 (5.1)	AGA-Arg	4 (2.2)
	ATG-Met	6 (3.4)	ACG-Thr	2 (1.1)	AAG-Lys	3 (1.7)	AGG-Arg	0 (0.0)
	GTT-Val	3 (1.7)	GCT-Ala	3 (1.7)	GAT-Asp	7 (3.9)	GGT-Gly	4 (2.2)
	GTC-Val	1 (0.6)	GCC-Ala	1 (0.6)	GAC-Asp	2 (1.1)	GGC-Gly	2 (1.1)
	GTA-Val	2 (1.1)	GCA-Ala	4 (2.2)	GAA-Glu	10 (5.6)	GGA-Gly	6 (3.4)
	GTG-Val	5 (2.8)	GCG-Ala	0 (0.0)	GAG-Glu	4 (2.2)	GGG-Gly	1 (0.6)
Herpes	TTT-Phe	5 (1.3)	TCT-Ser	2 (0.5)	TAT-Tyr	4 (1.1)	TGT-Cys	0 (0.0)
	TTC-Phe	3 (0.8)	TCC-Ser	3 (0.8)	TAC-Tyr	8 (2.1)	TGC-Cys	4 (1.1)
	TTA-Leu	1 (0.3)	TCA-Ser	1 (0.3)	TAA-	0 (0.0)	TGA-	1 (0.3)
	TTG-Leu	6 (1.6)	TCG-Ser	6 (1.6)	TAG-	0 (0.0)	TGG-Trp	5 (1.3)
	CTT-Leu	6 (1.6)	CCT-Pro	4 (1.1)	CAT-His	6 (1.6)	CGT-Arg	3 (0.8)
	CTC-Leu	6 (1.6)	CCC-Pro	15 (4.0)	CAC-His	6 (1.6)	CGC-Arg	14 (3.7)
	CTA-Leu	1 (0.3)	CCA-Pro	2 (0.5)	CAA-Gln	6 (1.6)	CGA-Arg	3 (0.8)
	CTG-Leu	21 (5.6)	CCG-Pro	9 (2.4)	CAG-Gln	13 (3.4)	CGG-Arg	7 (1.9)
	ATT-Ile	1 (0.3)	ACT-Thr	1 (0.3)	AAT-Asn	1 (0.3)	AGT-Ser	1 (0.3)
	ATC-Ile	11 (2.9)	ACC-Thr	13 (3.4)	AAC-Asn	7 (1.9)	AGC-Ser	4 (1.1)
	ATA-Ile	4 (1.1)	ACA-Thr	5 (1.3)	AAA-Lys	4 (1.1)	AGA-Arg	1 (0.3)
	ATG-Met	13 (3.4)	ACG-Thr	9 (2.4)	AAG-Lys	0 (0.0)	AGG-Arg	0 (0.0)
	GTT-Val	3 (0.8)	GCT-Ala	7 (1.9)	GAT-Asp	3 (0.8)	GGT-Gly	4 (1.1)
	GTC-Val	7 (1.9)	GCC-Ala	27 (7.2)	GAC-Asp	16 (4.2)	GGC-Gly	11 (2.9)
	GTA-Val	2 (0.5)	GCA-Ala	0 (0.0)	GAA-Glu	2 (0.5)	GGA-Gly	1 (0.3)
	GTG-Val	10 (2.7)	GCG-Ala	12 (3.2)	GAG-Glu	12 (3.2)	GGG-Gly	14 (3.7)

SEQUENCES UPSTREAM OF FOUR EARLY VACCINIA VIRUS GENES

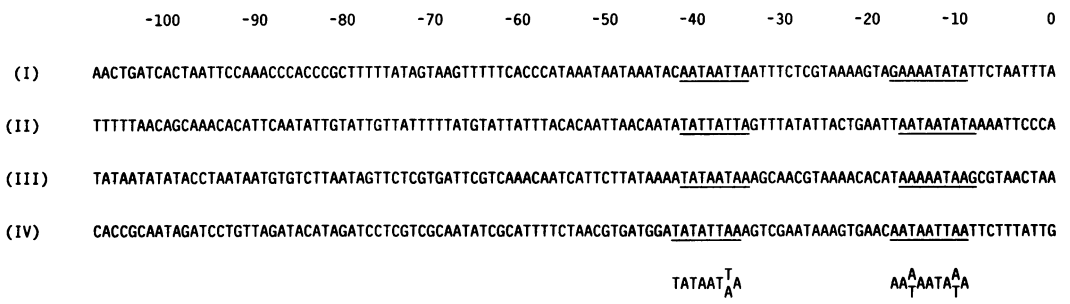

FIG. 5. Comparison of the nucleotide sequences preceding the RNA start sites of several early genes. Nucleotide sequences preceding the 7.5K gene (23) (I), the 19K gene (24) (II), the 42K gene (24) (III), and the TK gene (IV) are indicated. Consensus sequences indicated at the bottom were determined as described in the text.

transcriptional regulation, chimeric genes containing vaccinia virus DNA and foreign protein-coding sequences have been inserted into the vaccinia virus genome (16). A 275-bp DNA fragment, containing an RNA start site and upstream sequences of a gene encoding an early 7.5K polypeptide, served to promote expression of herpesvirus TK (16). Expression of procaryotic chloramphenical acetyltransferase and herpesvirus TK was obtained when the coding sequences of these genes were inserted four nucleotides downstream of the TK mRNA start site (M. Mackett, personal communication).

Protein-coding sequences are very closely packed in the region of the vaccinia virus genome described here. Only nine nucleotides separate the termination codon of a late gene from the TK initiation codon. Whether transcription of the late gene plays any role in modulating TK expression is not known. Downstream, 66 nucleotides separate the TK termination codon from the initiation codon of another early gene.

Primer extension mapping suggested that the TK mRNA has a unique 5′ end. By contrast, other early mRNAs examined appeared to have multiple 5′ ends that map close together. As determined with S1 nuclease (1) and mung bean nuclease, multiple 3′ ends of the TK mRNA mapped within a 10-bp segment of an extremely A+T-rich region. Whether this composition is related to termination or to the start of another early mRNA that maps just downstream (1) cannot be ascertained at this time. The 570- to 580-nucleotide length of the TK mRNA, determined by 5′ and 3′-end mapping, is similar to that determined by other methods (1). Nucleotide sequence data also is consistent with the absence of introns, a conclusion previously deduced from S1 nuclease mapping (1).

Previously, we found that three spontaneous TK$^-$ mutants had nonsense phenotypes (1). The nature of their mutations was revealed by nucleotide sequencing. In each case, an extra nucleotide identical to one preceding it was found. This extra nucleotide results in a frameshift and the occurrence of a termination codon downstream (Fig. 4). Perhaps a tendency to occasionally reiterate a previous nucleotide is characteristic of the vaccinia virus DNA polymerase. Such a mechanism for spontaneous frameshift mutations was inferred by Streisinger et al. (22) from studies with phage T4. They suggested that insertions would most likely occur in regions of repeating bases, as in fact occurred after four C residues in vaccinia virus TK$^-$ 20. Significantly, the location of the frameshift between the first and second ATG triplets of TK$^-$ 20 provides evidence that the first ATG initiates translation.

The availability of the nucleotide sequence for the vaccinia virus TK gene and flanking sequences, coupled with the powerful selection procedure available, should make in situ mutagenesis a feasible way of learning more about the novel regulatory mechanisms of vaccinia virus.

ACKNOWLEDGMENTS

We thank Alan Gershowitz for providing cloned DNA fragments of the TK$^-$ mutants and for assistance with sequence determinations. J. Owens kindly helped with graphic matrix analysis of sequence homologies.

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. **Belle Isle, H., S. Venkatesan, and B. Moss.** 1981. Cell-free translation of early and late mRNAs selected by hybridization to cloned DNA fragments from the left 14 million to 72 million daltons of the vaccinia virus genome. Virology **112**:306–317.

3. **Bina-Stein, M., J. A. Thompson, N. P. Salzman, and M. Thoren.** 1979. Rapid sequence determination of late simian virus 40 16S mRNA leader by using inhibitors of reverse transcriptase. Proc. Natl. Acad. Sci. U.S.A. **76**:731–735.
4. **Breathnach, R., and P. Chambon.** 1981. Organization and expression of eukaryotic split genes coding for proteins. Annu. Rev. Biochem. **50**:349–383.
5. **Brutlag, D. L., J. Clayton, P. Friedland, and L. H. Kedes.** 1982. SEQ: a nucleotide sequence analysis and recombination system. Nucleic Acids Res. **10**:279–295.
6. **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.
7. **Dubbs, D. R., and S. Kit.** 1964. Isolation and properties of vaccinia mutants deficient in thymidine kinase-inducing activity. Virology **22**:214–225.
8. **Ellims, P. H., and M. B. Van der Weyden.** 1980. Human liver thymidine kinase. Purification and some properties of the enzyme. J. Biol. Chem. **255**:11290–11295.
9. **Ghosh, P. K., V. B. Reddy, J. Swinscoe, P. Lebowitz, and S. M. Weissman.** 1978. Heterogeneity and 5'-terminal structures of the late RNAs of simian virus 40. J. Mol. Biol. **126**:813–846.
10. **Honess, R. W., and D. H. Watson.** 1974. Herpes simplex virus-specific polypeptides studied by polyacrylamide gel electrophoresis of immune precipitates. J. Gen. Virol. **22**:171–185.
11. **Hruby, D. E., and L. A. Ball.** 1982. Mapping and identification of vaccinia virus thymidine kinase gene. J. Virol. **43**:403–409.
12. **Kit, S., G. N. Jorgensen, A. Liav, and V. Zaslavsky.** 1977. Purification of vaccinia virus-induced thymidine kinase activity from [^{35}S]methionine-labeled cells. Virology **77**:661–676.
13. **Kozak, M.** 1981. Possible role of flanking nucleotides in recognition of the AUG initiator codon by eukaryotic ribosomes. Nucleic Acids Res. **9**:5233–5252.
14. **Queen, C. L., and L. J. Korn.** 1980. Computer analysis of nucleic acids and proteins. Methods Enzymol. **65**:595–609.
15. **Queen, C. L., M. N. Wegman, and L. J. Korn.** 1982. Improvements to a program for DNA analysis: a procedure to find homologies among many sequences. Nucleic Acids Res. **10**:449–456.
16. **Mackett, M., G. L. Smith, and B. Moss.** 1982. Vaccinia virus: a selectable eukaryotic cloning and expression vector. Proc. Natl. Acad. Sci. U.S.A. **79**:7415–7419.
17. **Maizel, J. V., Jr., and R. P. Lenk.** 1981. Enhanced graphic matrix analysis of nucleic acids and protein sequences. Proc. Natl. Acad. Sci. U.S.A. **78**:7665–7669.
18. **Messing, J., and J. Viera.** 1982. A new pair of M13 vectors for selecting either strand of double digest restriction fragments. Gene **19**:269–276.
19. **Moss, B.** 1974. Reproduction of poxviruses. Compr. Virol. **3**:405–473.
20. **Rhode, W., and A. G. Lezius.** 1971. The purification of thymidine kinase from Escherichia coli by affinity chromatography. Hoppe Seyler's Z. Physiol. Chem. **32**:1507–1516.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U.S.A. **74**:5463–5467.
22. **Streisinger, G., Y. Okada, J. Emrich, J. Newton, A. Tsugita, E. Terzaghi, and M. Inouye.** 1966. Frame shift mutations and the genetic code. Cold Spring Harbor Symp. Quant. Biol. **31**:77–84.
23. **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 **125**:805–813.
24. **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.
25. **Vogelstein, B., and D. Gillespie.** 1979. Preparative and analytical purification of DNA from agarose. Proc. Natl. Acad. Sci. U.S.A. **76**:615–619.
26. **Wagner, M. J., J. A. Sharp, and W. C. Summers.** 1981. Nucleotide sequence of the thymidine kinase gene of herpes simplex virus type 1. Proc. Natl. Acad. Sci. U.S.A. **78**:1441–1445.
27. **Weir, J. P., G. Bajszar, and B. Moss.** 1982. Mapping of the vaccinia virus thymidine kinase gene by marker rescue and by cell-free translation of selected mRNA. Proc. Natl. Acad. Sci. U.S.A. **79**:1210–1214.
28. **Winberg, G., and M.-L. Hammerskjold.** 1980. Isolation of DNA from agarose gels using DEAE-paper. Application to restriction site mapping of adenovirus type 16 DNA. Nucleic Acids Res. **8**:253–264.