JOURNAL OF VIROLOGY, May 1987, p. 1398–1406
0022-538X/87/051398-09$02.00/0

Vol. 61, No. 5

Nucleotide Sequence and Transcript Organization of a Region of the Vaccinia Virus Genome Which Encodes a Constitutively Expressed Gene Required for DNA Replication†

NANCY A. ROSEMAN AND DENNIS E. HRUBY*

Center for Gene Research, Department of Microbiology, Oregon State University, Corvallis, Oregon 97331-3804

Received 17 November 1986/Accepted 30 January 1987

A vaccinia virus (VV) gene required for DNA replication has been mapped to the left side of the 16-kilobase (kb) VV *Hin*dIII D DNA fragment by marker rescue of a DNA$^-$ temperature-sensitive mutant, *ts*17, using cloned fragments of the viral genome. The region of VV DNA containing the *ts*17 locus (3.6 kb) was sequenced. This nucleotide sequence contains one complete open reading frame (ORF) and two incomplete ORFs reading from left to right. Analysis of this region at early times revealed that transcription from the incomplete upstream ORF terminates coincidentally with the complete ORF encoding the *ts*17 gene product, which is directly downstream. The predicted proteins encoded by this region correlate well with polypeptides mapped by in vitro translation of hybrid-selected early mRNA. The nucleotide sequences of a 1.3-kb *Bgl*II fragment derived from *ts*17 and from two *ts*17 revertants were also determined, and the nature of the *ts*17 mutation was identified. S1 nuclease protection studies were carried out to determine the 5′ and 3′ ends of the transcripts and to examine the kinetics of expression of the *ts*17 gene during viral infection. The *ts*17 transcript is present at both early and late times postinfection, indicating that this gene is constitutively expressed. Surprisingly, the transcriptional start throughout infection occurs at the proposed late regulatory element TAA, which immediately precedes the putative initiation codon ATG. Although the biological activity of the *ts*17-encoded polypeptide was not identified, it was noted that in *ts*17-infected cells, expression of a nonlinked VV immediate-early gene (thymidine kinase) was deregulated at the nonpermissive temperature. This result may indicate that the *ts*17 gene product is functionally required at an early step of the VV replicative cycle.

Vaccinia virus (VV) is a large, DNA-containing animal virus which replicates in the cytoplasm of susceptible host cells. VV is an attractive system for the study of gene regulation in higher organisms owing to the temporally coordinated expression of its approximately 200 genes (25). At the onset of infection, immediate-early viral RNAs are transcribed by the virally encoded and packaged RNA polymerase (26). A protein product of this temporal class is then required for the expression of the delayed-early genes, as their expression is blocked by protein synthesis inhibitors such as cycloheximide. Approximately 2 h postinfection (p.i.), viral DNA replication is initiated. Concomitant with DNA replication is repression of the expression of most early genes and the initiation of late gene transcription. Thus, the expression of most VV genes can be classified as immediate early, delayed early, or late. However, there is also an additional constitutive temporal class of genes which appear to be expressed at both early and late times during infection as typified by the well-studied 7.5-kilodalton gene (5) and other VV genes (41).

The precise mechanisms involved in the modulation of VV gene expression are unknown, although they are believed to operate at the transcriptional, posttranscriptional, translational, and posttranslational levels. Through the identification and analysis of a number of genes from specific temporal classes it should be possible to identify the regulatory elements involved in VV gene expression. Sequence analyses of various regions of the viral genome (2, 11, 15, 27, 28, 33, 38, 39, 40, 43, 44) have identified many open reading frames (ORFs) which correlate well with earlier transcriptional and translational mapping studies (1, 3, 4). Upstream *cis*-acting regulatory sequences have been identified for both early and late genes, but they appear to have little obvious homology to previously identified procaryotic or eucaryotic promoters (2, 10, 19, 37, 44).

In this study we have mapped a viral gene belonging to the constitutive temporal class by marker rescue of a temperature-sensitive (*ts*) DNA$^-$ mutant, *ts*17 (6, 7). The *ts*17 locus was mapped within the *Hin*dIII D fragment, and the nucleotide sequence of this region was determined from wild-type (wt) viral DNA. While this work was in progress the nucleotide sequence of the entire *Hin*dIII D fragment was reported (27; nomenclature for ORFs is in relation to the reported sequence and according to the recently proposed system for VV genes [32]). Here, we independently confirm a portion of the reported sequence and show that this region is transcriptionally active in vivo. Comparison of sequences derived from DNA isolated from *ts*17 virus, from two *ts*17 revertants (*ts*17^r), and from wt virus identified the mutation responsible for the temperature sensitivity of *ts*17. Also, analysis of the sequence in conjunction with transcriptional and translational studies of this region showed that transcription of D4, which lies directly upstream of D5 (to which *ts*17 maps), terminates coincidentally with the D5 transcript.

The transcriptional kinetics of the *ts*17 gene was examined by S1 nuclease protection studies of steady-state mRNA levels. The results showed that the *ts*17 gene is a member of the constitutive temporal class of viral genes. The putative

* Corresponding author.

† Oregon Agricultural Experiment Station Technical Paper no. 8050.

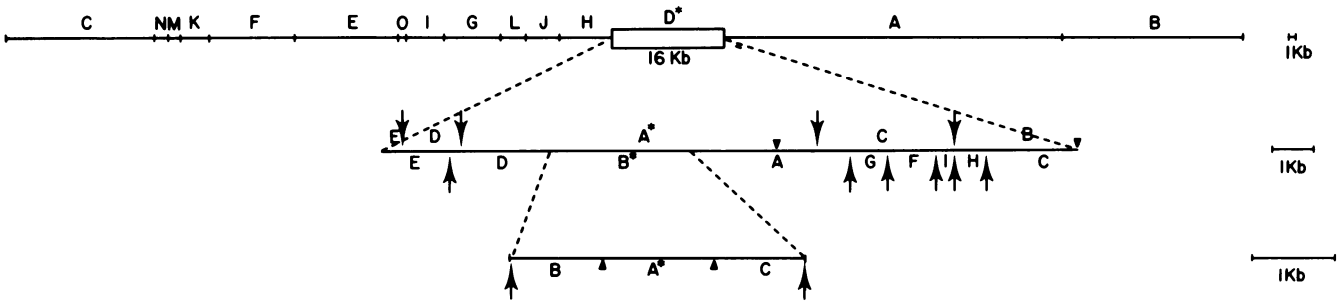


FIG. 1. Diagrammatic representation of the *Hin*dIII restriction map of the VV genome. Restriction fragments generated by digestion of the 16-kilobase-pair *Hin*dIII D fragment with *Bam*HI (↓) and *Eco*RI (↑) are shown. Restriction fragments generated by digestion of the 3.3-kb *Eco*RI B fragment by *Bgl*II (▲) are also shown. The 7.0-kb fragment designated clone pSW1 is shown (▼). Those frements used in the marker rescue study are indicated in the text and in Table 1. Fragments able to rescue *ts*17 to temperature insensitivity are indicated (*).

promoter of the *ts*17 gene was compared with those of previously reported VV genes.

MATERIALS AND METHODS

Cells and virus. BSC_{40} cells were grown in monolayer cultures of Eagle minimum essential medium (Flow Laboratories) supplemented with 10% heat-inactivated fetal calf serum, 2 mM L-glutamine, and 50 μg of gentamicin sulfate per ml. VV (strain WR) is designated as wt in these experiments. Revertants of *ts*17 were generated by repeated passages of *ts*17 virus at 40°C. *ts*17 revertants were then plaque purified at 40°C.

Marker rescue. Cloned wt VV DNA fragments were purified and sterilized by sequential extractions with phenol and ether, followed by ethanol precipitation. From 1 to 2 μg of viral DNA was suspended in HEPES (*N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid)-buffered saline and coprecipitated with salmon sperm DNA to a final concentration of 20 μg/ml using calcium phosphate (9). Confluent 60-mm plates of BSC_{40} cells were infected with *ts*17 at a multiplicity of infection of 0.1 and incubated at 31°C for 3 h. The medium was removed, the monolayer was washed with Eagle minimum essential medium, and 15% glycerol was added for 40 s. The monolayer was then washed twice with Eagle minimum essential medium, and the precipitated DNA was pipetted onto the cells. Plates were incubated at 40°C for 3 h and then washed, and minimum essential medium plus 10% fetal calf serum was added. Plates were incubated for 48 h at 40°C, the infected cells were harvested, and titers of infectious virus were determined at 31 and 40°C.

Cloning and sequencing. A clone of the *Hin*dIII D fragment of VV DNA was originally obtained from B. Moss (National Institutes of Health). Subfragments of viral DNA were cloned into pUC vectors by standard methodology (22). M13 phage vectors were used for sequencing by the dideoxynucleotide chain termination method (34). Sequences were analyzed on an IBM personal computer using Microgenie software obtained from Beckman Instruments, Inc. (29). Chemical sequencing was performed essentially as previously described (23). Enzymes and reagents were obtained from Bethesda Research Laboratories, Boehringer Mannheim, P-L Biochemicals, and Aldrich Chemical Co., Inc. Radioisotopes were obtained from New England Nuclear Corp.

Northern analysis. Two micrograms of polyadenylated viral RNA (isolated essentially as described [42]) was electrophoresed in a formaldehyde-agarose gel as described (18). After transfer to nitrocellulose, specific transcripts were detected by using the indicated nick-translated viral DNA fragments as probes.

Hybrid selection. A 20-μg sample of recombinant plasmid DNA was digested to completion with an appropriate restriction endonuclease and bound to five-3-mm-square nitrocellulose filters. Twenty micrograms of polyadenylated RNA isolated from VV-infected cells in the presence of 100 μg of cycloheximide per ml was used in hybridization reactions as described (1). Eluted mRNA was translated in vitro in rabbit reticulocyte lysates which contained 5 μCi of L-[^{35}S]methionine (New England Nuclear Corp.; 1,085 Ci/mmol) (14).

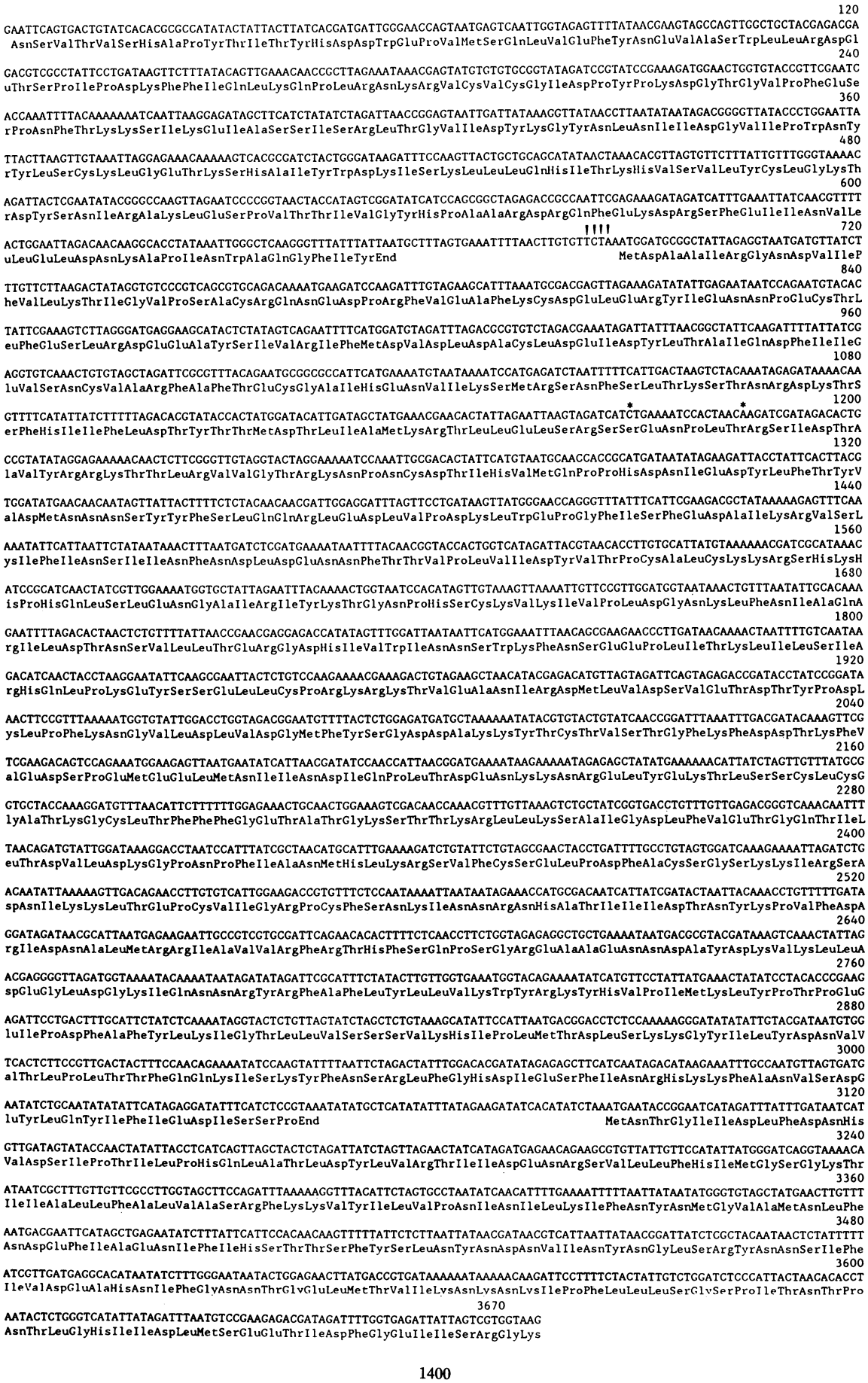
S1 nuclease mapping. 5'- or 3'-end-labeled DNA fragments were mixed with 20 μg of viral RNA in a total volume of 30 μl containing 80% formamide–40 mM PIPES [piperazine-*N*,*N'*-bis(2-ethanesulfonic acid)] (pH 6.4)–0.4 M NaCl–1 mM EDTA. The samples were heated to 90°C for 5 min and then incubated at 42°C for 4 h. Hybridization reactions were then plunged into ice, and 0.3 ml of ice-cold 0.28 M NaCl–0.05 M sodium acetate (pH 4.6)–4.5 mM $ZnSO_4$–400 U of nuclease S1 per ml was added. After 5 min on ice, the samples were incubated at 25°C for 1 h and then extracted with phenol-chloroform and ether and ethanol precipitated. Protected fragments were analyzed on sequencing gels.

Thymidine kinase assay. Duplicate sets of 60-mm dishes of Ltk$^-$ cells were infected at a multiplicity of 20 PFU per cell with wt VV or the indicated *ts* mutants. After a 30-min adsorption at 25°C, prewarmed (31 or 40°C) medium was added. One set of plates was incubated at the permissive temperature of 31°C, and the other was kept at the nonpermissive temperature of 40°C. At 2-h intervals, cytoplasmic extracts were prepared and frozen. Extracts were subsequently assayed for thymidine kinase activity (13).

TABLE 1. Titers of progeny of marker rescue by subclones of VV *Hin*dIII fragment D[a]

Viral DNA fragment	Titer at temp:	
	31°C	40°C
*Hin*dIII D	1.7×10^5	1.9×10^5
pSW1	1.2×10^3	0
*Bam*HI A	6.5×10^4	1.2×10^5
*Eco*RI B	2.6×10^5	2.8×10^5
*Bgl*II A	8.0×10^4	1.6×10^4
*Bgl*II B	8.3×10^3	0
*Bgl*II C	2.4×10^4	0
SS	4.2×10^3	0

[a] Viral DNA was coprecipitated with salmon sperm DNA (SS) and then transfected into *ts*17-infected BSC_{40} cells.


120
GAATTCAGTGACTGTATCACACGCGCCATATACTATTACTTATCACGATGATTGGGAACCAGTAATGAGTCAATTGGTAGAGTTTTATAACGAAGTAGCCAGTTGGCTGCTACGAGACGA
AsnSerValThrValSerHisAlaProTyrThrIleThrTyrHisAspAspTrpGluProValMetSerGlnLeuValGluPheTyrAsnGluValAlaSerTrpLeuLeuArgAspGl
240
GACGTCGCCTATTCCTGATAAGTTCTTTATACAGTTGAAACAACCGCTTAGAAATAAACGAGTATGTGTGTGCGGTATAGATCCGTATCCGAAAGATGGAACTGGTGTACCGTTCGAATC
uThrSerProIleProAspLysPhePheIleGlnLeuLysGlnProLeuArgAsnLysArgValCysValCysGlyIleAspProTyrProLysAspGlyThrGlyValProPheGluSe
360
ACCAAATTTTACAAAAAAATCAATTAAGGAGATAGCTTCATCTATATCTAGATTAACCGGAGTAATTGATTATAAAGGTTATAACCTTAATATAATAGACGGGGTTATACCCTGGAATTA
rProAsnPheThrLysLysSerIleLysGluIleAlaSerSerIleSerArgLeuThrGlyValIleAspTyrLysGlyTyrAsnLeuAsnIleIleAspGlyValIleProTrpAsnTy
480
TTACTTAAGTTGTAAATTAGGAGAAACAAAAAGTCACGCGATCTACTGGGATAAGATTTCCAAGTTACTGCTGCAGCATATAACTAAACACGTTAGTGTTCTTTATTGTTTGGGTAAAAC
rTyrLeuSerCysLysLeuGlyGluThrLysSerHisAlaIleTyrTrpAspLysIleSerLysLeuLeuGlnHisIleThrLysHisValSerValLeuTyrCysLeuGlyLysTh
600
AGATTACTCGAATATACGGGCCAAGTTAGAATCCCCGGTAACTACCATAGTCGGATATCATCCAGCGGCTAGAGACCGCCAATTCGAGAAAGATAGATCATTTGAAATTATCAACGTTTT
rAspTyrSerAsnIleArgAlaLysLeuGluSerProValThrThrIleValGlyTyrHisProAlaAlaArgAspArgGlnPheGluLysAspArgSerPheGluIleIleAsnValLe
720
ACTGGAATTAGACAACAAGGCACCTATAAATTGGGCTCAAGGGTTTATTTATTAATGCTTTAGTGAAATTTTAACTTGTGTTCTAAATGGATGCGGCTATTAGAGGTAATGATGTTATCT
uLeuGluLeuAspAsnLysAlaProIleAsnTrpAlaGlnGlyPheIleTyrEnd
MetAspAlaAlaIleArgGlyAsnAspValIleP
840
TTGTTCTTAAGACTATAGGTGTCCCGTCAGCGTGCAGACAAAATGAAGATCCAAGATTTGTAGAAGCATTTAAATGCGACGAGTTAGAAAGATATATTGAGAATAATCCAGAATGTACAC
heValLeuLysThrIleGlyValProSerAlaCysArgGlnAsnGluAspProArgPheValGluAlaPheLysCysAspGluLeuGluArgTyrIleGluAsnAsnProGluCysThrL
960
TATTCGAAAGTCTTAGGGATGAGGAAGCATACTCTATAGTCAGAAATTTTCATGGATGTAGATTTAGACGCGTGTCTAGACGAAATAGATTATTTAACGGCTATTCAAGATTTTATTATCG
euPheGluSerLeuArgAspGluGluAlaTyrSerIleValArgIlePheMetAspValAspLeuAspAlaCysLeuAspGluIleAspTyrLeuThrAlaIleGlnAspPheIleIleG
1080
AGGTGTCAAACTGTGTAGCTAGATTCGCGTTTACAGAATGCGGCGCCATTCATGAAAATGTAATAAAATCCATGAGATCTAATTTTTCATTGACTAAGTCTACAAATAGAGATAAAACAA
luValSerAsnCysValAlaArgPheAlaPheThrGluCysGlyAlaIleHisGluAsnValIleLysSerMetArgSerAsnPheSerLeuThrLysSerThrAsnArgAspLysThrS
1200
GTTTTCATATTATCTTTTTAGACACGTATACCACTATGGATACATTGATAGCTATGAAACGAACACTATTAGAATTAAGTAGATCATCTGAAAATCCACTAACAAGATCGATAGACACTG
erPheHisIleIlePheLeuAspThrTyrThrThrMetAspThrLeuIleAlaMetLysArgThrLeuLeuGluLeuSerArgSerSerGluAsnProLeuThrArgSerIleAspThrA
1320
CCGTATATAGGAGAAAAACAACTCTTCGGGTTGTAGGTACTAGGAAAAATCCAAATTGCGACACTATTCATGTAATGCAACCACCGCATGATAATATAGAAGATTACCTATTCACTTACG
laValTyrArgArgLysThrThrLeuArgValValGlyThrArgLysAsnProAsnCysAspThrIleHisValMetGlnProProHisAspAsnIleGluAspTyrLeuPheThrTyrV
1440
TGGATATGAACAACAATAGTTATTACTTTTCTCTACAACAACGATTGGAGGATTTAGTTCCTGATAAGTTATGGGAACCAGGGTTTATTTCATTCGAAGACGCTATAAAAAGAGTTTCAA
alAspMetAsnAsnAsnSerTyrTyrPheSerLeuGlnGlnArgLeuGluAspLeuValProAspLysLeuTrpGluProGlyPheIleSerPheGluAspAlaIleLysArgValSerL
1560
AAATATTCATTAATTCTATAATAAACTTTAATGATCTCGATGAAAATAATTTTACAACGGTACCACTGGTCATAGATTACGTAACACCTTGTGCATTATGTAAAAAACGATCGCATAAAC
ysIlePheIleAsnSerIleIleAsnPheAsnAspLeuAspGluAsnAsnPheThrThrValProLeuValIleAspTyrValThrProCysAlaLeuCysLysLysArgSerHisLysH
1680
ATCCGCATCAACTATCGTTGGAAAATGGTGCTATTAGAATTTACAAAACTGGTAATCCACATAGTTGTAAAGTTAAAATTGTTCCGTTGGATGGTAATAAACTGTTTAATATTGCACAAA
isProHisGlnLeuSerLeuGluAsnGlyAlaIleArgIleTyrLysThrGlyAsnProHisSerCysLysValLysIleValProLeuAspGlyAsnLysLeuPheAsnIleAlaGlnA
1800
GAATTTTAGACACTAACTCTGTTTTATTAACCGAACGAGGAGACCATATAGTTTGGATTAATAATTCATGGAAATTTAACAGCGAAGAACCCTTGATAACAAAACTAATTTTGTCAATAA
rgIleLeuAspThrAsnSerValLeuLeuThrGluArgGlyAspHisIleValTrpIleAsnAsnSerTrpLysPheAsnSerGluGluProLeuIleThrLysLeuIleLeuSerIleA
1920
GACATCAACTACCTAAGGAATATTCAAGCGAATTACTCTGTCCAAGAAAACGAAAGACTGTAGAAGCTAACATACGAGACATGTTAGTAGATTCAGTAGAGACCGATACCTATCCGGATA
rgHisGlnLeuProLysGluTyrSerSerGluLeuLeuCysProArgLysArgLysThrValGluAlaAsnIleArgAspMetLeuValAspSerValGluThrAspThrTyrProAspL
2040
AACTTCCGTTTAAAAATGGTGTATTGGACCTGGTAGACGGAATGTTTTACTCTGGAGATGATGCTAAAAAATATACGTGTACTGTATCAACCGGATTTAAATTTGACGATACAAAGTTCG
ysLeuProPheLysAsnGlyValLeuAspLeuValAspGlyMetPheTyrSerGlyAspAspAlaLysLysTyrThrCysThrValSerThrGlyPheLysPheAspAspThrLysPheV
2160
TCGAAGACAGTCCAGAAATGGAAGAGTTAATGAATATCATTAACGATATCCAACCATTAACGGATGAAAATAAGAAAAATAGAGAGCTATATGAAAAAACATTATCTAGTTGTTTATGCG
alGluAspSerProGluMetGluGluLeuMetAsnIleIleAsnAspIleGlnProLeuThrAspGluAsnLysLysAsnArgGluLeuTyrGluLysThrLeuSerSerCysLeuCysG
2280
GTGCTACCAAAGGATGTTTAACATTCTTTTTTGGAGAAACTGCAACTGGAAAGTCGACAACCAAACGTTTGTTAAAGTCTGCTATCGGTGACCTGTTTGTTGAGACGGGTCAAACAATTT
lyAlaThrLysGlyCysLeuThrPhePhePheGlyGluThrAlaThrGlyLysSerThrThrLysArgLeuLeuLysSerAlaIleGlyAspLeuPheValGluThrGlyGlnThrIleL
2400
TAACAGATGTATTGGATAAAGGACCTAATCCATTTATCGCTAACATGCATTTGAAAAGATCTGTATTCTGTAGCGAACTACCTGATTTTGCCTGTAGTGGATCAAAGAAAATTAGATCTG
euThrAspValLeuAspLysGlyProAsnProPheIleAlaAsnMetHisLeuLysArgSerValPheCysSerGluLeuProAspPheAlaCysSerGlySerLysLysIleArgSerA
2520
ACAATATTAAAAAGTTGACAGAACCTTGTGTCATTGGAAGACCGTGTTTCTCCAATAAAATTAATAATAGAAACCATGCGACAATCATTATCGATACTAATTACAAACCTGTTTTTGATA
spAsnIleLysLysLeuThrGluProCysValIleGlyArgProCysPheSerAsnLysIleAsnAsnArgAsnHisAlaThrIleIleIleAspThrAsnTyrLysProValPheAspA
2640
GGATAGATAACGCATTAATGAGAAGAATTGCCGTCGTGCGATTCAGAACACACTTTTCTCAACCTTCTGGTAGAGAGGCTGCTGAAAATAATGACGCGTACGATAAAGTCAAACTATTAG
rgIleAspAsnAlaLeuMetArgArgIleAlaValValArgPheArgThrHisPheSerGlnProSerGlyArgGluAlaAlaGluAsnAsnAspAlaTyrAspLysValLysLeuLeuA
2760
ACGAGGGGTTAGATGGTAAAATACAAAATAATAGATATAGATTCGCATTTCTATACTTGTTGGTGAAATGGTACAGAAAATATCATGTTCCTATTATGAAACTATATCCTACACCCGAAG
spGluGlyLeuAspGlyLysIleGlnAsnAsnArgTyrArgPheAlaPheLeuTyrLeuLeuValLysTrpTyrArgLysTyrHisValProIleMetLysLeuTyrProThrProGluG
2880
AGATTCCTGACTTTGCATTCTATCTCAAAATAGGTACTCTGTTAGTATCTAGCTCTGTAAAGCATATTCCATTAATGACGGACCTCTCCAAAAAGGGATATATATTGTACGATAATGTGG
luIleProAspPheAlaPheTyrLeuLysIleGlyThrLeuLeuValSerSerSerValLysHisIleProLeuMetThrAspLeuSerLysLysGlyTyrIleLeuTyrAspAsnValV
3000
TCACTCTTCCGTTGACTACTTTCCAACAGAAAATATCCAAGTATTTTAATTCTAGACTATTTGGACACGATATAGAGAGCTTCATCAATAGACATAAGAAATTTGCCAATGTTAGTGATG
alThrLeuProLeuThrThrPheGlnGlnLysIleSerLysTyrPheAsnSerArgLeuPheGlyHisAspIleGluSerPheIleAsnArgHisLysLysPheAlaAsnValSerAspG
3120
AATATCTGCAATATATATTCATAGAGGATATTTCATCTCCGTAAATATATGCTCATATATTTATAGAAGATATCACATATCTAAATGAATACCGGAATCATAGATTTATTTGATAATCAT
luTyrLeuGlnTyrIlePheIleGluAspIleSerSerProEnd
MetAsnThrGlyIleIleAspLeuPheAspAsnHis
3240
GTTGATAGTATACCAACTATATTACCTCATCAGTTAGCTACTCTAGATTATCTAGTTAGAACTATCATAGATGAGAACAGAAGCGTGTTATTGTTCCATATTATGGGATCAGGTAAAACA
ValAspSerIleProThrIleLeuProHisGlnLeuAlaThrLeuAspTyrLeuValArgThrIleIleAspGluAsnArgSerValLeuLeuPheHisIleMetGlySerGlyLysThr
3360
ATAATCGCTTTGTTGTTCGCCTTGGTAGCTTCCAGATTTAAAAAGGTTTACATTCTAGTGCCTAATATCAACATTTTGAAAATTTTTAATTATAATATGGGTGTAGCTATGAACTTGTTT
IleIleAlaLeuLeuPheAlaLeuValAlaSerArgPheLysLysValTyrIleLeuValProAsnIleAsnIleLeuLysIlePheAsnTyrAsnMetGlyValAlaMetAsnLeuPhe
3480
AATGACGAATTCATAGCTGAGAATATCTTTATTCATTCCACAACAAGTTTTTATTCTCTTAATTATAACGATAACGTCATTAATTATAACGGATTATCTCGCTACAATAACTCTATTTTT
AsnAspGluPheIleAlaGluAsnIlePheIleHisSerThrThrSerPheTyrSerLeuAsnTyrAsnAspAsnValIleAsnTyrAsnGlyLeuSerArgTyrAsnAsnSerIlePhe
3600
ATCGTTGATGAGGCACATAATATCTTTGGGAATAATACTGGAGAACTTATGACCGTGATAAAAAATAAAAACAAGATTCCTTTTCTACTATTGTCTGGATCTCCCATTACTAACACACCT
IleValAspGluAlaHisAsnIlePheGlyAsnAsnThrGlyGluLeuMetThrValIleLysAsnLysAsnLysIleProPheLeuLeuLeuSerGlySerProIleThrAsnThrPro
3670
AATACTCTGGGTCATATTATAGATTTAATGTCCGAAGAGACGATAGATTTTGGTGAGATTATTAGTCGTGGTAAG
AsnThrLeuGlyHisIleIleAspLeuMetSerGluGluThrIleAspPheGlyGluIleIleSerArgGlyLys

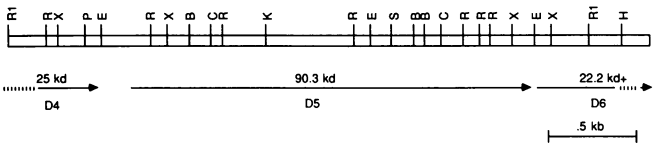


FIG. 3. Restriction map of the *Eco*RI B fragment. The complete ORF is shown by a solid line; incomplete ORFs are indicated with a broken line. Predicted molecular weights of the ORFs are indicated along with their numerical designations. R1, *Eco*RI; R, *Rsa*I; X, *Xba*I; P, *Pst*I; E, *Eco*RV; B, *Bgl*II; C, *Cla*I; K, *Kpn*I; S, *Sal*I; H, *Hin*fI.

RESULTS

Marker rescue mapping. Previously reported marker rescue mapping, using the *Hin*dIII fragments of VV, located the *ts*17 locus within the 16-kilobase (kb) *Hin*dIII D fragment (7). Subclones derived by restriction enzyme digests of wt *Hin*dIII D which were used for fine mapping of the *ts*17 locus are shown (Fig. 1). Viral DNA fragments were calcium phosphate precipitated and transfected into BSC_{40} cells which had been previously infected with *ts*17 virus. After a 48-h incubation at the nonpermissive temperature of 40°C, the infected cells were harvested, and titers of viral progeny were determined at the permissive temperature of 31°C and at the nonpermissive temperature of 40°C (Table 1). At permissive temperature, *ts*17 replication was independent of the identity of the transfected DNA fragment, whereas at nonpermissive temperature replication was dependent upon the homologous recombination of a wt viral DNA fragment which contained the appropriate sequences. Marker rescue using pSW1, a 7.0-kb fragment derived from the right-hand side of *Hin*dIII D, resulted in *ts*17 remaining temperature sensitive. However, the 8.1-kb *Bam*HI A fragment from the left-hand side of *Hin*dIII D was able to confer temperature insensitivity. The mutation was mapped further to an *Eco*RI subclone of *Bam*HI A, the 3.3-kb *Eco*RI B fragment. *Eco*RI B was subcloned into three *Bgl*II fragments, and of these the 1.3-kb *Bgl*II A fragment contained the sequences required to confer temperature insensitivity to *ts*17.

Nucleotide sequence and identification of the *ts*17 mutation. To facilitate molecular studies of the *ts*17 gene, the nucleotide sequences of the 3,309-base-pair *Eco*RI B fragment and of 300 base pairs reading rightward into the neighboring *Eco*RI A fragment were determined by the dideoxynucleotide chain termination method (Fig. 2). These sequences were determined by cloning appropriate overlapping subfragments of viral DNA into M13 vectors in both orientations so that both strands could be sequenced. By this strategy, over 80% of the sequence of both strands was determined. The sequence reported here is in complete agreement with that recently reported (27).

Analysis of ORFs found within the sequence reported here showed tandemly oriented ORFs which read from left to right (Fig. 3). In accordance with the recently proposed nomenclature for VV genes, these were designated D4, D5, and D6. D4 is incomplete in this sequence, with the G at position 1 actually being the third base of the initiation ATG codon of D4 (27). Thus, the ORF D4 has the capacity to encode a 25,055-molecular-weight polypeptide. The complete D5 ORF, which spans the site of the *ts*17 mutation, is

VV ISOLATE		
WT	AGT AGA TCA TCT GAA AAT CCA CTA ACA AGA TCG	
	Ser Arg Ser Ser Glu Asn Pro Leu Thr Arg Ser	
ts 17	AGT AGA TCA TTT* GAA AAT CCA CTA ACC* AGA TCG	
	Ser Arg Ser Phe Glu Asn Pro Leu Thr Arg Ser	
ts 17^{r}1	AGT AGA TCA TCT GAA AAT CCA CTA ACC* AGA TCG	
	Ser Arg Ser Ser Glu Asn Pro Leu Thr Arg Ser	
ts 17^{r}2	AGT AGA TCA TCT GAA AAT CCA CTA ACC* AGA TCG	
	Ser Arg Ser Ser Glu Asn Pro Leu Thr Arg Ser	

FIG. 4. Nucleotide sequence of the *Bgl*II A fragment was determined from viral DNA isolated from wt, *ts*17, and two *ts*17 revertants (*ts*17^{r}1 and *ts*17^{r}2). Shown is the sequence from nucleotide positions 1158 to 1190. The amino acid encoded is also shown. The nucleotide positions at which there is a change from wt sequences are indicated (*).

2,297 nucleotides long and could specify a polypeptide with a predicted molecular weight of 90,367. The incomplete ORF D6 encodes 22,234 daltons of what is reported to be a 68,362-dalton polypeptide (27).

To confirm the location of the *ts*17 gene and to determine the nature of the mutation giving rise to it, the 1.3-kb *Bgl*II A fragment was subcloned from *ts*17 viral DNA and the nucleotide sequence was determined. Furthermore, a number of *ts*17 revertants were isolated by repeated passages of *ts*17 virus at 40°C. The *Bgl*II A fragments from two revertants (*ts*17^{r}1 and *ts*17^{r}2) were also cloned and sequenced (Fig. 4). A comparison between the wt and *ts*17 sequences identified two closely linked mutations. The first is a transition mutation of a C to a T at nucleotide position 1168, which causes an amino acid change of a serine to a phenylalanine. The second mutation is a transversion mutation of an A to a C at nucleotide position 1184; in this case the mutation is silent, and the amino acid being encoded remains threonine. Sequences derived from the two *ts*17 revertants were identical and showed both *ts*17^{r}1 and *ts*17^{r}2 to be true revertants. The nucleotide at position 1168 had reverted back to a C, returning the codon to a serine. However, the silent mutation remained unchanged, thus indicating that *ts*17^{r}1 and *ts*17^{r}2 were *ts*17 revertants and not wt VV. Therefore, the sequence data (Fig. 4) and the genetic evidence (Table 1) demonstrate that a transition mutation at nucleotide position 1168 is responsible for the *ts* phenotype of *ts*17 virus.

Hydrophilicity-hydrophobicity analysis. The effect of the single amino acid change caused by the transition mutation on the hydropathic index (17, 35) of the predicted *ts*17 protein was determined by a comparison of the hydropathy profiles of *ts*17 and wt (Fig. 5). The amino acid change clearly caused a localized change in the hydropathy profile of *ts*17 as compared to wt. The codons which directly precede and follow the mutation reflect a higher degree of hydrophobicity. The hydropathy profile of the entire *ts*17 protein, as predicted from wt sequences, was also analyzed. The *ts*17 protein appears to have a profile typical of a globular protein, although there is a strong hydrophobic region at the carboxy terminus of the polypeptide (data not shown).

Hybrid selection and in vitro translation. To determine the tra-nslation products of the transcripts encoded for within

FIG. 2. Nucleotide sequence reading rightward beginning at the left-hand *Eco*RI site of *Eco*RI B and extending 300 base pairs rightward into the *Eco*RI A fragment. The RNA start sites determined by S1 nuclease mapping are marked with arrowheads. The sites of the *ts*17 mutations are indicated (*).

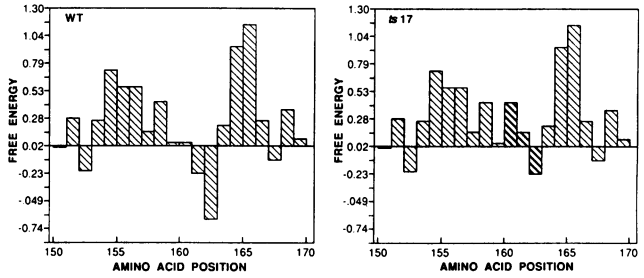


FIG. 5. Hydrophobicity profile of wt (left) and *ts*17 (right) in the region of the *ts*17 mutation. The *y* axis is the free energy of transfer of amino acid sidechains; the *x* axis is the amino acid position. The analysis shown here begins at codon 150, which corresponds with nucleotide position 1137. The *ts*17 mutation changes the codon at amino acid position 162.

*Eco*RI B, polyadenylated immediate-early viral RNA hybrid selected by *Eco*RI B was translated in vitro in rabbit reticulocyte lysates in the presence of [^{35}S]methionine. Labeled protein products were then electrophoresed on a denaturing sodium dodecyl sulfate-polyacrylamide gel and autoradiographed (Fig. 6).

The in vitro translation system programmed with mRNAs selected by the *Eco*RI B fragment yielded two major polypeptides with molecular weights of approximately 25,500 and 88,000. This correlates well with the size of polypeptides that can be predicted from the coding capacity of D4 and D5, i.e., 25,055 and 90,367 molecular weight, respectively. There does not appear to be a translation product that correlates to D6. This may be due to D6 being of the late viral gene class, and thus its transcript would not be present in the mRNA population used in this experiment.

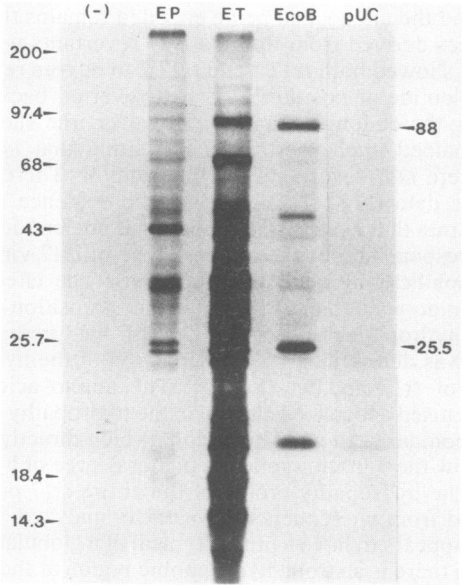


FIG. 6. Hybrid-selected cell-free translation. Polypeptides synthesized in rabbit reticulocyte lysates were separated by gel electrophoresis. Lanes: (−), no RNA added; ET, total early viral RNA; EcoB, VV early RNA hybrid-selected by *Eco*RI B; pUC, early viral RNA hybrid-selected by pUC-18. Lane EP, Polypeptides pulse-labeled at 1.5 h p.i. in wt VV-infected cells. Numbers at the left indicate molecular weights ($\times 10^3$) of protein molecular weight markers.

The two minor polypeptides of approximately 22,000 and 45,000 molecular weight are endogenous proteins of the translation system and not of viral origin. None of the *Eco*RI B-derived polypeptides comigrated with any of the major early viral proteins produced either in vivo or in vitro (Fig. 6, lanes EP and ET).

Identification of transcripts by Northern analysis. To identify mRNAs that are transcribed within *Eco*RI B, polyadenylated immediate-early viral RNA was electrophoresed under denaturing conditions on a formaldehyde-agarose gel and transferred to nitrocellulose (Fig. 7). Nick-translated *Eco*RI B DNA used as a probe revealed a major transcript of 2.7 kb and a minor transcript of 3.6 kb. To localize these transcripts more specifically, the *Bgl*II subclones of *Eco*RI B were used as probes; it was found that both transcripts span the three *Bgl*II fragments (data not shown).

To further map these transcripts, the *Eco*RI D and A fragments, which border *Eco*RI B, were used as probes. Neither *Eco*RI D nor A appeared to hybridize to the two transcripts which are coded for within *Eco*RI B. The 2.6-kb transcript which hybridized to *Eco*RI D correlated well with D1 and may encode the viral guanyl transferase (24). The 0.57- and 0.98-kb transcripts which map to *Eco*RI A can possibly be assigned to D7 and D8, respectively. The absence of a large transcript which can be assigned to D6 may again be due to this gene's being expressed late in infection.

This analysis appears to show that the 2.7- and 3.6-kb transcripts are coded for entirely within *Eco*RI B. In the case of D4, there does not appear to be a transcript of the appropriate size that would be predicted to encode the 25-kilodalton polypeptide whose presence was indicated by the sequence data and by the results of the hybrid selection in vitro translation. This result is not due to potential pleiotropic effects of the drug block (cycloheximide) used to isolate abundant early viral RNA. Only the 2.7- and 3.6-kb transcripts were hybridized by *Eco*RI B when a Northern analysis was repeated using viral RNA isolated at 2 h p.i. in the absence of drug (data not shown).

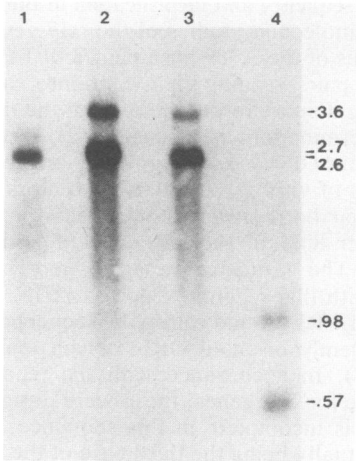


FIG. 7. Northern blot analysis. Two micrograms of polyadenylated early viral RNA was electrophoresed on a formaldehyde-agarose gel and then blotted onto nitrocellulose. Nick-translated plasmids containing fragments of VV DNA were used as probes: lane 1, *Eco*RI D; lane 2, *Eco*RI B; lane 3, *Bgl*II A; lane 4, *Eco*RI A. Numbers at right indicate calculated transcript sizes (kb) determined by coelectrophoresis of brome mosaic viral RNAs as molecular weight markers.

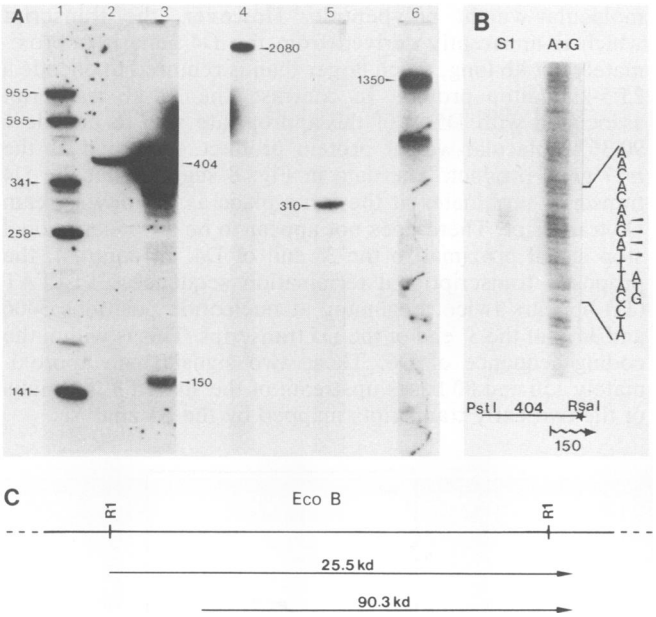


FIG. 8. (A) S1 nuclease mapping of transcripts encoded by *Eco*RI B. Labeled DNA fragments were hybridized with early RNA and treated with S1 nuclease, and protected hybrids were then analyzed on sequencing gels. Lane 1, *Sau*3A digest of pUC18; lane 2, full-length 404-nucleotide 5′-end-labeled probe; lane 3, S1 analysis using 404-nucleotide probe; lane 4, full-length 2,080-nucleotide 5′-end-labeled probe; lane 5, S1 analysis using 2,080-nucleotide probe; lane 6, S1 analysis using 3′-end-labeled 3,286-nucleotide probe (the full-length 3,286-nucleotide probe does not enter the gel system). Sizes in nucleotides of the probes and protected fragments are indicated. (B) S1 nuclease mapping of the mRNA start site. Lane S1 is the S1 nuclease-resistant fragment (wavy line) from the hybridization of a 5′ single end-labeled probe (solid line) and early mRNA. Lane A+G is a Maxam-Gilbert A+G reaction of the indicated probe. The sizes of the probes and the protected fragments are indicated in nucleotides. The sites of initiation are indicated with arrowheads within the context of the sequence. (C) Diagrammatic representation of the organization of early mRNAs transcribed in *Eco*RI B.

S1 nuclease mapping. To determine the organization of the two early transcripts encoded within *Eco*RI B, S1 nuclease protection studies were done. In these studies the indicated probe was hybridized to 20 μg of total RNA isolated from infected cells in the presence of cycloheximide. After S1 digestion, the protected fragments were electrophoresed on a denaturing polyacrylamide gel (Fig. 8A).

To map the 5′ end of the D4 transcript, a 2,080-nucleotide 5′ single end-labeled *Xba*I probe was isolated. This probe protected a fragment of approximately 310 bases, mapping the 5′ start of D4 to approximately 23 nucleotides upstream of the *Eco*RI D-B junction and the initiation ATG. The 5′ end of D5, which encodes the *ts*17 gene, was determined by utilizing a 404-nucleotide 5′ single end-labeled *Pst*I-*Rsa*I fragment as a probe. A protected fragment of approximately 150 nucleotides maps the transcriptional start of D5 about 690 nucleotides downstream of the *Eco*RI D-B junction.

To map the transcriptional stop of the two transcripts, DNA fragments which spanned the entire *Eco*RI B fragment were 3′ end labeled and used as probes. The results showed full-length protection of these probes, indicating that there is no early termination of transcription within the *Eco*RI B fragment (data not shown). Therefore, a 3,286-nucleotide *Sal*I fragment which spans the *Eco*RI B-A junction was 3′ end labeled and used as a probe. A protected fragment of 1,350 nucleotides mapped a single 3′ end approximately 200 nucleotides downstream of the *Eco*RI B-A junction. A second protected fragment (Fig. 8A, lane 6) does not have a 3′ end within *Eco*RI B, since a smaller overlapping probe showed only full-length protection. This protection must arise from a leftward-reading early transcript encoded by *Eco*RI A.

This analysis identified two early transcripts reading from left to right which are approximately 3,590 and 2,877 nucleotides in length. This is consistent with the Northern data and shows that the two transcripts terminate coincidentally (Fig. 8C).

To facilitate future studies of the *ts*17 gene and its promoter, the transcriptional start site for D5 was specifically mapped. The 404-nucleotide *Pst*I-*Rsa*I fragment was 5′ end labeled at the *Rsa*I site, gel isolated, and hybridized to 20 μg of viral RNA as described above. After S1 nuclease digestion, protected fragments were electrophoresed alongside the products of an A+G chemical sequencing reaction of the 404 nucleotide probe (Fig. 8B). Multiple transcription initiation sites were mapped at positions −2 to −5 nucleotides, upstream of the first ATG of D5. Initiation occurs within the sequence TAA, which lies directly upstream of the putative initiation codon, ATG.

Transcript kinetics. Since the sequence element TAAAT has been proposed to be a late regulatory element and has here been shown to occur in a gene the expression of which is required early, the kinetics of transcription initiation of the *ts*17 gene was determined. This was accomplished by S1 analysis of the steady-state mRNA levels over time.

The 5′ probe previously described (Fig. 8B) was hybridized to 20 μg of total RNA, which was isolated at hourly intervals throughout infection. This analysis also included RNA isolated in the presence of cycloheximide and hydroxyurea. Cycloheximide treatment allows transcription

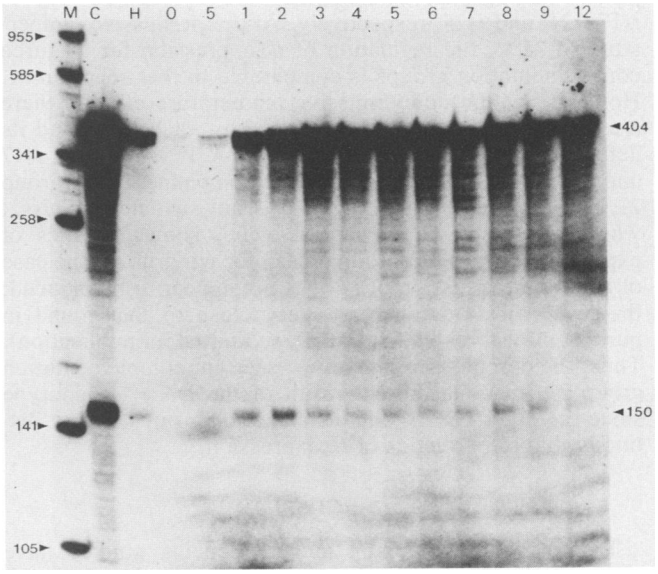


FIG. 9. Kinetics of transcription initiation. A 5′ single end-labeled probe (see Fig. 8B) was hybridized to 20 μg of total viral RNA isolated in the presence of cycloheximide (C) or hydroxyurea (H) or at the indicated times postinfection in the absence of drugs. Marker is a *Sau*3A digest of pUC18 (M). The 404-nucleotide full-length probe is shown, along with the size of the S1 nuclease-protected fragment.

of immediate-early genes only, whereas with hydroxyurea treatment, delayed-early genes are also transcribed. Both treatments block late gene expression.

After S1 nuclease digestion, a protected fragment of approximately 150 base pairs was detected (Fig. 9). This corresponds to the 5′ start located in Fig. 8A and B. The *ts*17 gene was first transcribed at 1 h p.i. The level of transcription initiation remained stable until 3 h p.i., when it appeared to decrease. The *ts*17 transcript was then present for up to 12 h p.i. Since the half-life of VV mRNA late in infection has been estimated to be approximately 1 to 2 h (13, 30), the presence of *ts*17 RNA at late times is most likely due to continued transcription.

Regulation of thymidine kinase expression by VV DNA$^-$ mutants. Of the three complementation groups which have been determined to have a DNA$^-$ phenotype, only one has been assigned a biological function, that being the DNA polymerase gene (16, 36). It remains to be shown whether the loci of the other two DNA$^-$ complementation groups encode proteins which have a biological activity that is an integral part of DNA synthesis, or rather play an essential role early in the VV replicative cycle which is not directly related to DNA metabolism. It was therefore of interest to examine the regulation of the VV thymidine kinase gene (*tk*) in DNA$^-$ mutant-infected Ltk$^-$ cells to ascertain at which point the ordered expression of this gene might be interrupted. Expression of the VV *tk* gene is regulated at the transcriptional, translational, and posttranslational levels by other VV gene products from both the early and late gene classes (12, 13).

The kinetics of *tk* expression in cells infected with wt or with *ts* mutants of the three DNA$^-$ complementation groups was determined (Fig. 10). In wt-infected cells *tk* is expressed at the onset of infection, is attenuated at 4 to 6 h p.i., and then remains stable. The level of expression is higher at 40°C than at 31°C, probably due to metabolic processes being increased at the higher temperature. The three DNA$^-$ complementation groups, 5, 21, and 24, are represented by *ts*42, *ts*17/*ts*24, and *ts*25, respectively. At the permissive temperature of 31°C, the regulation of *tk* expression for all three complementation groups is comparable to that seen for wt. However, at the nonpermissive temperature of 40°C there are distinct differences in the impact of each locus and its defective gene product on *tk* expression. In the case of *ts*17 and *ts*24, which together comprise complementation group 21, the expression of the *tk* gene is not switched off. *ts*42, which is the DNA polymerase gene, shows kinetics of expression very similar to that seen for wt virus. In the case of *ts*25, the expression of *tk* does not appear to be induced; the levels of *tk* activity are very close to that found in purified viral cores (D. E. Hruby, submitted for publication). Thus, the derepressed phenotype seen for complementation group 21 is not the direct result of the DNA$^-$ phenotype since the other DNA$^-$ complementation groups exhibit distinctly different kinetics of *tk* expression.

DISCUSSION

Transcriptional mapping and sequencing studies have shown VV genes to be organized in a tightly clustered, tandemly oriented fashion (8, 20, 21, 27, 28, 40, 45). Here, we independently confirm a portion of previously reported sequence and show that the predicted D4 and D5 ORFs are transcriptionally active. D4 is transcribed early in infection. Whether it is also transcriptionally active at later times is unknown. D4 has a predicted coding capacity for a 25,055-molecular-weight polypeptide. However, the transcript which is apparently derived from the D4 gene is approximately 3.7 kb long, much larger than is required to encode a 25.5-kilodalton protein. In contrast, the 3.0-kb transcript associated with D5 is of the appropriate size to encode a 90,367-molecular-weight protein product, identified as the *ts*17 gene product. The data in Fig. 8 suggest that the D4 transcript terminates at the same place as the downstream D5 transcript. There does not appear to be a transcriptional stop signal proximal to the 3′ end of D4. In contrast, the proposed transcriptional termination sequence TTTTTAT (31) appears twice, beginning at nucleotide positions 3406 and 3475 at the 3′ end of the D5 transcript. This is within the coding sequence of D6. These two signals map approximately 150 and 80 bases upstream of the shared 3′ terminus of the two early transcripts mapped by the S1 analysis.

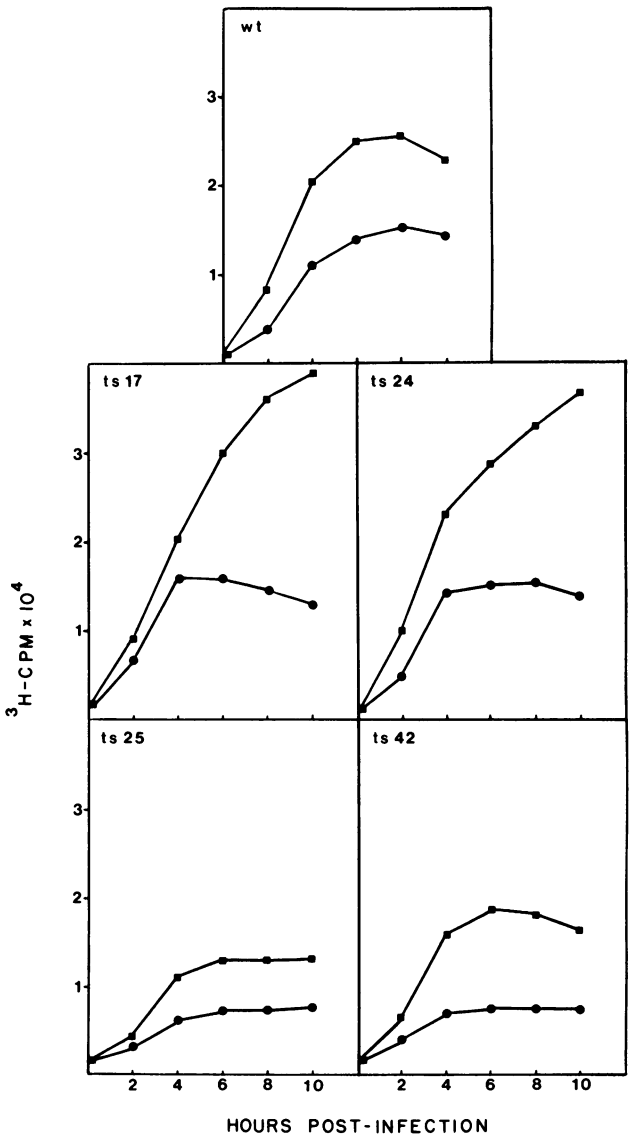


FIG. 10. Thymidine kinase activity. Thymidine kinase assays were carried out on duplicate sets of 60-mm dishes of Ltk$^-$ cells which were infected with the indicated *ts* mutant. The plates were incubated at either the permissive temperature of 31°C (●) or the nonpermissive temperature of 40°C (■) and harvested at 2-h intervals.

Marker rescue mapping and the comparison of sequences derived from wt, *ts*17, and two *ts*17 revertant viral stocks have conclusively demonstrated that D5 is the *ts*17 locus. This locus encodes a gene product which appears to be functionally required early in the viral replication cycle. Without a functional *ts*17 gene, the viral replication cycle is blocked prior to DNA replication. Yet, a kinetic analysis shows that the *ts*17 locus is transcribed continuously throughout infection, beginning at 1 h p.i., and is therefore a member of the constitutive temporal class.

The increasing amount of VV sequence data available allows for a comparative analysis of putative viral promoters and thus facilitates the identification of consensus sequences which may be regulatory elements. Transcription initiation at D5 has multiple start sites which occur between 2 and 5 nucleotides upstream of the initiation ATG. Similar findings have been reported for both early (5, 39, 43) and late (2, 11, 33, 44) VV genes. In the latter case, transcriptional initiation occurs within the sequence TAA, which is directly upstream of the initiation ATG. It has been proposed that the sequence TAAAT is essential for late promoter function (10) and may actually be part of a signal sequence for late gene expression (33). However, since the sequence TAAAT serves as the transcriptional start site for the D5 locus at both early and late times postinfection, this indicates that there must be additional factors involved with specifying the temporal expression of VV genes. This finding illustrates the flexibility that must be required to regulate the complex expression of a variety of specific temporal classes of virally encoded gene products.

The *ts*17 gene is comparable to previously sequenced VV genes in that it is 66% A-T rich, with the first 100 nucleotides upstream of the 5′ start being 71% A-T rich. A close examination of the first 100 nucleotides shows an A-T-rich octet at positions 20 and 40 nucleotides upstream of the first ATG. A-T octets at this position have been proposed as regulatory elements for early VV genes (38, 39, 43, 44). However, the octets of D5 are not homologous to the proposed consensus sequence proposed for these regulatory elements. It appears that the D5 gene has promoter elements which have been associated with both early and late consensus sequences. This makes clear the necessity for further sequence analysis in conjunction with kinetic studies for identifying temporal subsets of viral gene expression. A directed genetics approach will then be required to determine which elements are involved in regulating gene expression. Such an approach is currently being used to define the elements involved in the expression and regulation of the *ts*17 gene.

ACKNOWLEDGMENTS

We thank R. C. Condit for generously providing his entire collection of temperature-sensitive mutants to our laboratory. We thank Lisa Wilson, Walt Hodges, and Chris Franke for excellent technical support. We also thank Scott L. Weinrich for the use of time course RNAs and for helpful discussions throughout this work.

This study was supported by a Public Health Service grant from the National Institutes of Health (AI-21335) and by a grant from the American Cancer Society (MV-194). D.E.H. was supported by a Research Career Development Award (AI-00666) from the National Institutes of Health.

LITERATURE CITED

1. **Belle Isle, H., S. Venkatesan, and B. Moss.** 1981. Cell-free translation of early late mRNAs selected by hybridization to cloned DNA fragments derived from the left 14 million to 72 million daltons of the vaccinia virus genome. Virology **112**:306–317.
2. **Bertholet, C., R. Drillien, and R. Wittek.** 1985. One hundred base pairs of 5′ flanking sequence of a vaccinia virus late gene are sufficient to temporally regulate late transcription. Proc. Natl. Acad. Sci. USA **82**:2096–2100.
3. **Cabrera, C. V., M. Esteban, R. McCarron, W. T. McAllister, and J. A. Holowzak.** 1978. Vaccinia virus transcription: hybridization of mRNA to restriction fragments of vaccinia DNA. Virology **86**:102–114.
4. **Chipchase, M., F. Schwendimann, and R. Wyler.** 1980. A map of the late proteins of vaccinia virus. Virology **105**:261–264.
5. **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.
6. **Condit, R. C., and A. Motyczka.** 1981. Isolation and preliminary characterization of temperature-sensitive mutants of vaccinia virus. Virology **113**:224–241.
7. **Condit, R. C., A. Motyczka, and G. Spizz.** 1983. Isolation, characterization, and physical mapping of temperature-sensitive mutants of vaccinia virus. Virology **128**:429–443.
8. **Cooper, J. A., R. Wittek, and B. Moss.** 1981. Extension of the transcriptional and translational map of the left end of the vaccinia virus genome to 21 kilobase pairs. J. Virol. **39**:733–745.
9. **Graham, F. L., and A. J. van der Eb.** 1973. Transformation of rat cells by DNA of human adenovirus 5. J. Virol. **54**:536–539.
10. **Hanggi, M., W. Bannwarth, and H. G. Stunnenberg.** 1986. Conserved TAAAT motif in vaccinia virus late promoters: overlapping TATA box and site of transcription initiation. EMBO J. **5**:1071–1076.
11. **Hirt, P., G. Hiller, and R. Wittek.** 1986. Localization and fine structure of a vaccinia virus gene encoding an envelope antigen. J. Virol. **58**:757–764.
12. **Hruby, D. E.** 1985. Inhibition of vaccinia virus thymidine kinase by the distal products of its own metabolic pathway. Virus Res. **2**:151–156.
13. **Hruby, D. E., and L. A. Ball.** 1981. Control of expression of the vaccinia virus thymidine kinase gene. J. Virol. **40**:456–464.
14. **Hruby, D. E., and L. A. Ball.** 1982. Mapping and identification of the vaccinia virus thymidine kinase gene. J. Virol. **43**: 403–409.
15. **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.
16. **Jones, E. V., and B. Moss.** 1984. Mapping of the vaccinia virus DNA polymerase gene by marker rescue and cell-free translation of selected RNA. J. Virol. **49**:72–77.
17. **Lawson, E., A. Sadler, D. Harmatz, D. Brandau, R. Micanovic, R. MacElroy, and C. R. Middaught.** 1984. A simple experimental model for hydrophobic interactions in proteins. J. Biol. Chem. **259**:2910–2912.
18. **Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker.** 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry **16**:4743–4752.
19. **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.
20. **Mahr, A., and B. E. Roberts.** 1984. Organization of six early transcripts synthesized from a vaccinia virus *Eco*RI DNA fragment. J. Virol. **49**:497–509.
21. **Mahr, A., and B. E. Roberts.** 1984. Arrangement of late RNAs transcribed from a 7.1-kilobase *Eco*RI vaccinia DNA fragment. J. Virol. **49**:510–520.
22. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
23. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
24. **Morgan, J. R., L. K. Cohen, and B. E. Roberts.** 1984. Identification of the DNA sequences encoding the large subunit of the

mRNA-capping enzyme of vaccinia virus. J. Virol. **52**:206–214.
25. **Moss, B.** 1985. Replication of poxviruses, p. 685–703. *In* B. N. Fields (ed.), Virology. Raven Press, New York.
26. **Munyon, W., E. Paoletti, and J. T. Grace, Jr.** 1967. RNA polymerase activity in purified infectious vaccinia virus. Proc. Natl. Acad. Sci. USA **8**:2280–2287.
27. **Niles, E. G., R. C. Condit, P. Caro, K. Davidson, L. Matusick, and J. Seto.** 1986. Nucleotide sequence and genetic map of the 16-kb vaccinia virus Hind III D fragment. Virology **153**:96–112.
28. **Pluciennicak, A., E. Shroeder, C. Zettlmeissl, and R. E. Streeck.** 1985. Nucleotide sequence of a cluster of early and late genes in a conserved segment of the vaccinia virus genome. Nucleic Acids Res. **13**:985–998.
29. **Queen, C., and L. J. Korn.** 1984. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. **12**:581–599.
30. **Rice, A. P., and B. E. Roberts.** 1983. A novel mechanism of viral induced inhibition of host cell protein synthesis: vaccinia virus accelerates the labilization of mouse L-cell mRNAs. J. Virol. **47**:529–539.
31. **Rohrmann, G., L. Yuen, and B. Moss.** 1986. Transcription of vaccinia virus early genes by enzymes isolated from vaccinia virions terminates downstream of a regulatory sequence. Cell **46**:1029–1035.
32. **Rosel, J., P. L. Earl, J. Weir, and B. Moss.** 1986. Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the *Hin*dIII H genome fragment. J. Virol. **60**:436–449.
33. **Rosel, J., and B. Moss.** 1985. Transcriptional and translational mapping and nucleotide sequence analysis of a vaccinia virus gene encoding the precursor of the major core polypeptide 4b. J. Virol. **56**:830–838.
34. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
35. **Taylor, W. R.** 1986. The classification of amino acid conservation. J. Theor. Biol. **119**:219–234.
36. **Traktman, P., P. Sridhar, R. C. Condit, and B. E. Roberts.** 1984. Transcriptional mapping of the DNA polymerase gene of vaccinia virus. J. Virol. **49**:125–131.
37. **Vassef, A., M. Mars, A. Dru, A. Pluciennicak, R. Streeck, and G. Beaud.** 1985. Isolation of *cis*-acting vaccinia virus DNA fragments promoting the expression of herpes simplex virus thymidine kinase by recombinant virus. J. Virol. **55**:163–172.
38. **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.
39. **Venkatesan, S., A. Gershowitz, and B. Moss.** 1982. Complete nucleotide sequences of adjacent early vaccinia virus genes located within the inverted terminal repetition. J. Virol. **44**: 637–646.
40. **Weinrich, S. L., and D. E. Hruby.** 1986. A tandemly-oriented late gene cluster within the vaccinia virus genome. Nucleic Acids Res. **14**:3003–3016.
41. **Weinrich, S. L., and D. E. Hruby.** 1987. Noncoordinate regulation of a vaccinia virus late gene cluster. J. Virol. **61**:639–645.
42. **Weinrich, S. L., E. G. Niles, and D. E. Hruby.** 1985. Transcriptional and translational analysis of the vaccinia virus late gene L65. J. Virol. **55**:450–457.
43. **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.
44. **Weir, J. P., and B. Moss.** 1984. Regulation of expression and nucleotide sequence of a late vaccinia virus gene. J. Virol. **51**:662–669.
45. **Wittek, R., B. Richner, and G. Hiller.** 1984. Mapping of the genes coding for the two major vaccinia virus core polypeptides. Nucleic Acids Res. **12**:4835–4848.