Cell, Vol. 61, 801–809, June 1, 1990, Copyright © 1990 by Cell Press

Role of DNA Replication in Vaccinia Virus Gene Expression: A Naked Template Is Required for Transcription of Three Late *Trans*-Activator Genes

James G. Keck, Carl J. Baldick, Jr., and Bernard Moss
Laboratory of Viral Diseases
National Institute of Allergy and Infectious Diseases
National Institutes of Health
Bethesda, Maryland 20892

Summary

The DNA replication requirement for vaccinia virus late gene expression was mimicked by transfecting a late promoter–controlled reporter gene into infected cells in the presence of a DNA synthesis inhibitor. This late promoter activation block was overcome by cotransfecting either naked linear vaccinia virion DNA or three cloned viral genes encoding *trans*-activator polypeptides of 17, 26, and 30 kd. These newly identified *trans*-activator genes were independently transcribed only from replicated or transfected DNA. These data suggest a regulatory cascade in which the parental viral genome serves as a template for the RNA polymerase and early promoter–specific transcription factors that are packaged in the infectious particle; the newly replicated DNA is accessible to sequentially synthesized intermediate promoter– and late promoter–specific *trans*-activators.

Introduction

DNA replication plays a central role in the regulation of gene expression in prokaryotes and eukaryotes. Because of their rapid and precisely programmed life cycle, DNA viruses provide advantageous systems for studying such phenomena. Experiments with adenoviruses (Crossland and Raskas, 1983; Gaynor and Berk, 1983; Thomas and Mathews, 1980) and herpesviruses (Mavromara-Nazos and Roizman, 1987) suggested that regulation occurs at the transcriptional level and requires a still undefined *cis*-acting modification of the template associated with DNA replication. The replication fork and its associated proteins may act as a mobile enhancer for expression of bacteriophage T4 late genes (Herendeen et al., 1989).

The stringent division of poxvirus genes into pre- and postreplicative classes was recognized more than 20 years ago, but the regulatory mechanisms underlying this division have remained elusive (for review, see Moss, 1990). Vaccinia virus, the prototypic member of the poxvirus family, contains a double-stranded linear DNA genome that encodes between 150 and 200 genes, of which approximately half are expressed only after DNA replication. Several features of poxviruses lend themselves to combined biochemical and genetic investigations. Importantly, they encode most if not all of the enzymes and factors needed for their replication in the cytoplasmic compartment of the cell. A complete transcription system is packaged within the infectious virus particle and is activated upon entry into the cytoplasm (Kates and McAuslan, 1967; Munyon and Kit, 1966). Therefore, neither protein nor DNA synthesis is required for transcription of early genes. In the presence of inhibitors of DNA replication, the life cycle of vaccinia virus is arrested; early mRNA synthesis continues, but late mRNA is not made (Boone and Moss, 1978; Kaverin et al., 1975; Oda and Joklik, 1967; Paoletti and Grady, 1977). When the inhibitor is removed, two successive classes of postreplicative proteins, called intermediate and late, are detected (Moss and Salzman, 1968; Pennington, 1974; Vos and Stunnenberg, 1988).

Detailed biochemical studies have provided evidence for differences in the promoter sequences of early and late class genes (Davison and Moss, 1989a, 1989b). A factor specific for early promoter transcription has been purified and characterized as a two subunit protein with DNA-dependent ATPase activity (Broyles and Moss, 1988; Broyles et al., 1988). The factors necessary for intermediate gene transcription have not been identified but must be present in the infected cell prior to DNA replication (Vos and Stunnenberg, 1988). Factors for late transcripton appear after replication (Wright and Moss, 1987; Schwer and Stunnenberg, 1988) and have been partially purified (Wright and Moss, 1989). Vos and Stunnenberg (1988) suggested that intermediate genes may be *trans*-activators of late gene transcription and proposed a cascade model. However, until now, no evidence for a connection between expression of intermediate and late genes has been reported.

In the first part of the present study, transfection procedures were used to determine whether the replication requirement for late gene expression can be circumvented. In the second part, we identified three intermediate genes that are only transcribed from replicated (or transfected) DNA and that are necessary and sufficient to overcome a replication block to late promoter expression.

Results

Transfected Virion DNA Activates Late Gene Expression in the Absence of DNA Replication

A transient expression assay in which cells are infected with vaccinia virus and transfected with a plasmid containing a reporter gene controlled by a vaccinia virus late promoter was described by Cochran et al. (1985). It was observed that the expression of the reporter gene was prevented by addition of cytosine arabinoside (AraC), an inhibitor of DNA replication. We considered that by transfecting infected cells with appropriate forms of DNA, we might overcome the AraC block and thereby learn how late gene expression and DNA synthesis are coupled.

First, we improved the sensitivity of the transient assay by employing plasmid 11Xβ (p11Xβ), which has the strong and well-characterized late promoter of the vaccinia virus 11K structural protein gene (Bertholet et al., 1985) driving expression of β-galactosidase. Our general protocol con-

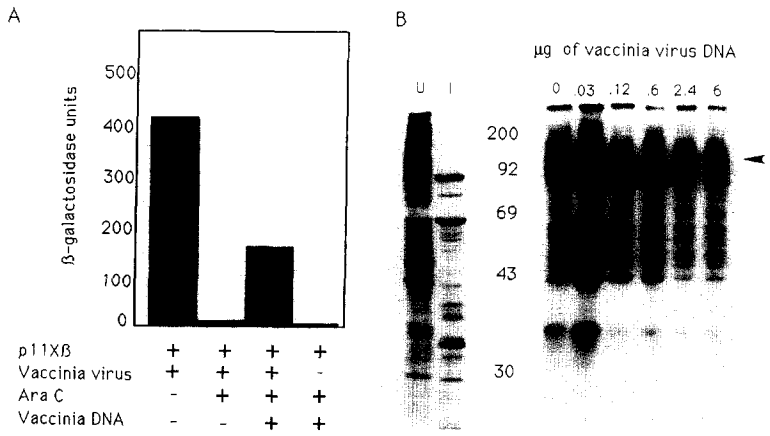

Figure 1. Transfected Vaccinia Virion DNA *Trans*-Activates a Late Promoter–Regulated Reporter Gene in the Absence of DNA Replication

(A) Vaccinia virus– or mock-infected 293 cells (1×10^6) were treated where indicated (+) with 40 μg/ml AraC, starting 30 min after infection, and transfected with 1.125 μg of calcium phosphate–precipitated p11Xβ, a pUC19-derived plasmid containing the vaccinia virus 11K late promoter regulating expression of the *lacZ* reporter gene, and 0.6 μg of either vaccinia virus (+) or salmon sperm (–) DNA. At 18 hr postinfection, the cells were lysed and β-galactosidase activity was measured.

(B) Vaccinia virus–infected cells were transfected with 0–6 μg of vaccina virus DNA and 1.125 μg of p11Xβ in the presence of AraC. At 4 hr postinfection, the regular medium was replaced with methionine-free medium still containing AraC. The cells were labeled with ^{35}S-methionine (30 μCi/ml) from 16 to 16.5 hr and lysed, and the proteins were analyzed by electrophophoresis on 10.5% polyacryamide gels. An autoradiogaph is shown; the arrowhead points to the β-galactosidase protein. The U and I lanes contain proteins labeled in the absence of AraC from uninfected and infected cells, respectively. Molecular weight markers in kd are indicated.

sisted of infecting human 293 cells with vaccinia virus at a multiplicity of 10 plaque-forming units per cell, adding AraC after 30 min, transfecting with p11Xβ and additional DNA as indicated, and assaying cell lysates for β-galactosidase after 16–18 hr. The AraC prevented detectable viral DNA and late protein synthesis as measured by DNA dot blot hybridization and polyacrylamide gel electrophoresis of extracts of [^{35}S]methionine-labeled cells, respectively. Moreover, AraC abrogated β-galactosidase expression by more than 99% and prevented detectable transcription of the latter gene (shown later).

To overcome the AraC block, we first tried transfecting naked DNA that had been purified from virus particles together with p11Xβ. β-galactosidase activity was stimulated 40-fold over controls in which salmon sperm DNA was transfected with p11Xβ (Figure 1A). With 1.25 μg of p11Xβ, maximal β-galactosidase activity was obtained with 2.4 μg of transfected virion DNA (data not shown). As expected, the transfected virion DNA supplemented but did not replace the requirement for virus infection (Figure 1A).

The relative amount of β-galactosidase synthesized in the presence of AraC was evaluated by infecting cells with vaccinia virus, transfecting with 1.25 μg of p11Xβ and 0–6 μg of virion DNA, and metabolically labeling with [^{35}S]methionine from 16 to 16.5 hr. Autoradiographs of polyacrylamide gels revealed that the labeled β-galactosidase band (indicated by an arrowhead) was clearly present when the AraC-treated cells were transfected with 0.6 μg of virion DNA and became prominent with 2.4 μg; otherwise, the early pattern of protein synthesis appeared unchanged by transfection (Figure 1B). For comparison, the uninfected and late viral patterns are shown in a parallel gel containing proteins from cells untreated with AraC. The inability to detect individual late viral proteins expressed from the transfected virion DNA in the presence of AraC could reflect its low (1:40 or less) molar ratio to reporter plasmid or some structural difference allowing only the plasmid to serve as a template. To discriminate between these possibilities, DNA was extracted from a recombinant vaccinia virus that had the late promoter–controlled reporter gene integrated into the vaccinia genome. With high (50 μg) amounts of sheared recombinant DNA, β-galactosidase activity was readily detected in the presence of AraC (data not shown).

In summary, these data demonstrate that the stringent replication requirement for late promoter activation can be overcome by providing naked virion DNA in *trans* or *cis*. Further experiments were designed to determine the role of the transfected virion DNA.

Cloned DNA Segments Can Substitute for Virion DNA

To determine whether molecularly cloned DNA can substitute for virion DNA in rescuing expression of late pro-

Table 1. *Trans*-Activation of Late Promoter Expression

Cosmids						
19	3	21	10	A	6	β-Gal Units
–	–	–	–	–	–	0.6
+	+	+	+	+	+	50.1
–	+	+	+	+	+	24.9
+	–	+	+	+	+	32.6
+	+	–	+	+	+	5.1
+	+	+	–	+	+	51.6
+	+	+	+	–	+	17.1
+	+	+	+	+	–	27.7
–	–	–	–	–	–	4.9
+	–	+	–	–	–	7.1
–	+	+	–	–	–	3.7
–	–	+	+	–	–	5.1
–	–	+	–	+	–	28.1
–	–	+	–	–	+	1.7

Vaccinia virus–infected 293 cells were incubated with AraC and transfected with p11Xβ and cosmids. The vaccinia DNA segments within the cosmids are listed in order from left to right on the virus genome. Salmon sperm DNA was used in controls. β-galactosidase (β-gal) activities of cell lysates were measured after 16–18 hr.

```
        10         20         30         40         50         60
TCTAGATACTCTACCATTTCTGCCGCATACAATAACTTGTTAGATAAAATCAGGGTTATC

        70         80         90        100        110        120
AAAGTGTTTAGCGTGGCTAGAATAGTGGGCTTGCATGTATTAAAGAATGCGGTAGTATGA

       130        140        150        160        170        180
GTAAACCGTTTTAACGAATTATATAGTCTCCAGAAATCTGTGGCGTTACATACATGAGCC

       190        200        210        220        230        240
GAATGACATCGAAGATTGTCCAATATTTTTAATAGCTGCTCTTTGTCCATTATTTCTATA

       250        260        270        280        290        300
TTTGACTCGCAACAATTGTAGATACCATTAATCACCGATTCCTTTTTCGATGCCGGACAA

       310        320        330        340        350        360
TAGCACAATTGTTTAGCTTTGGACTCTATGTATTCAGAATTAATAGATATATCTCTTAAT

       370        380        390        400        410        420
ACAGATTGCACTATACATTTTGAAACTATGTCAAAAATTGTAGAACGACGCTGTTCTGCA

       430        440        450      ▼▼460        470        480
GCCATTTAACTTTAAATAATTTACAAAAATTTAAAATGAGCATCCGTATAAAAATCGATA
                                    MetSerIleArgIleLysIleAspL

       490        500        510        520        530        540
AACTGCGCCAAATTGTGGCATATTTTTCAGAGTTCAGTGAAGAAGTATCTATAAATGTAG
ysLeuArgGlnIleValAlaTyrPheSerGluPheSerGluGluValSerIleAsnValA

       550        560        570        580        590        600
ACTCGACGGATGAGTTAATGTATATTTTTGCCGCCTTGGGCGGATCTGTAAACATTTGGG
spSerThrAspGluLeuMetTyrIlePheAlaAlaLeuGlyGlySerValAsnIleTrpA

       610        620        630        640        650        660
CCATTATACCTCTCAGTGCATCAGTGTTCTACCGCGGAGCCGAAAATATTGTGTTTAATC
laIleIleProLeuSerAlaSerValPheTyrArgGlyAlaGluAsnIleValPheAsnL

       670        680        690        700        710        720
TTCCTGTGTCCAAGGTAAAATCGTGTTTGTGTAGTTTTCACAATGATGCCATCATAGATA
euProValSerLysValLysSerCysLeuCysSerPheHisAsnAspAlaIleIleAspI

       730        740        750        760        770        780
TAGAACCTGATCTGGAAAATAATCTAGTAAAACTTTCTAGTTATCATGTAGTAAGTGTCG
leGluProAspLeuGluAsnAsnLeuValLysLeuSerSerTyrHisValValSerValA

       790        800        810        820        830        840
ATTGTAATAAGGAACTGATGCCTATTAGGACAGATACTACTATTTGTCTAAGTATAGATC
spCysAsnLysGluLeuMetProIleArgThrAspThrThrIleCysLeuSerIleAspG

       850        860        870        880        890        900
AAAAGAAATCTTACGTGTTTAATTTTCACAAGTATGAAGAAAAATGTTGTGGTAGAACCG
lnLysLysSerTyrValPheAsnPheHisLysTyrGluGluLysCysCysGlyArgThrV

       910        920        930        940        950        960
TCATTCATTTAGAATGGTTGTTGGGCTTTATCAAGTGTATTAGTCAGCATCAGCATTTGG
alIleHisLeuGluTrpLeuLeuGlyPheIleLysCysIleSerGlnHisGlnHisLeuA

       970        980        990       1000       1010       1020
CTATTATGTTTAAAGATGACAATATTATTATGAAGACTCCTGGTAATACTGATGCATTTT
laIleMetPheLysAspAspAsnIleIleMetLysThrProGlyAsnThrAspAlaPheS

      1030       1040       1050       1060       1070       1080
CCAGGGAATATTCTATGACTGAATGTTCTCAAGAACTACAAAAGTTTTCTTTCAAAATAG
erArgGluTyrSerMetThrGluCysSerGlnGluLeuGlnLysPheSerPheLysIleA

      1090       1100       1110       1120       1130       1140
CTATCTCGTCTCTCAACAAACTACGAGGATTCAAAAAGAGAGTCAATGTTTTTGAAACTA
laIleSerSerLeuAsnLysLeuArgGlyPheLysLysArgValAsnValPheGluThrA

      1150       1160       1170       1180       1190       1200
GAATCGTAATGGATAATGACGATAACATTTTAGGAATGTTGTTTTCGGATAGAGTTCAAT
rgIleValMetAspAsnAspAspAsnIleLeuGlyMetLeuPheSerAspArgValGlnS

      1210       1220       1230       1240       1250       1260
CCTTTAAGATCAACATCTTTATGGCGTTTTTAGATTAATACTTTCAATGAGATAAATATG
erPheLysIleAsnIlePheMetAlaPheLeuAsp

      1270       1280       1290       1300       1310       1320
GGTGGCGGAGTAAGTGTTGAGCTCCCTAAACGGGATCCGCCTCCGGGAGTACCCACTGAT

      1330       1340       1350       1360       1370
GAGATGTTATTAAACGTGGATAAAATGCATGACGTGATAGCTCCCGCTAAGCTT
```

Figure 2. Sequence of the GK1 ORF and Flanking Regions

The DNA sequence between the XbaI and HindIII sites encompassing the GK1 ORF and the predicted polypeptide sequence of the latter are shown. The 5′ ends of the GK1 transcripts, determined by nuclease S1 protection, are indicated by arrowheads.

moters in AraC-treated infected cells, we employed a library of six overlapping cosmids that together contain the entire vaccinia virus genome except for the terminal sequences (Jones et al., 1987). The degree of overlap of the cosmid sequences was sufficient to ensure that uninterrupted copies of all genes were present in the library, but this also meant that many genes were contained in two different cosmids. When infected cells were transfected with all six cosmids and p11Xβ, β-galactosidase activity was significantly increased over controls transfected with salmon sperm DNA and p11Xβ (Table 1). No single cosmid was effective by itself (data not shown), however, suggesting that more than one gene might be required for *trans*-activation. Results obtained with mixtures of five cosmids indicated that omission of cosmid 21 resulted in the lowest β-galactosidase activity (Table 1). To analyze the cosmid requirement further, sets of two cosmids, each including cosmid 21, were tested in the assay. The combination of cosmids 21 and A was found to give the greatest stimulation of β-galactosidase activity (Table 1). These results indicated that transfection of infected cells with at least two nonoverlapping segments of the vaccinia genome was required for *trans*-activation of late promoter expression in the absence of DNA replication.

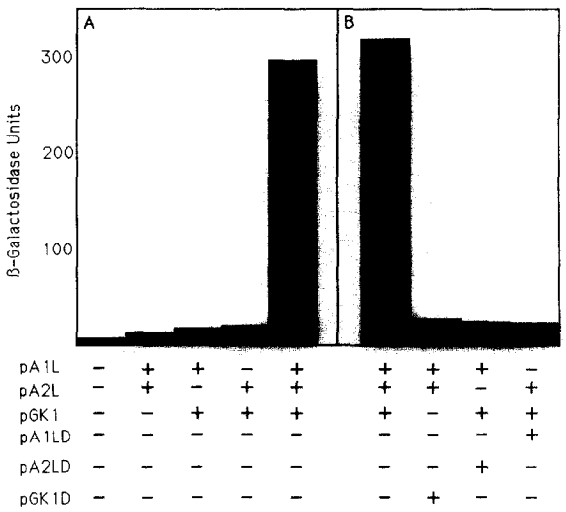

Figure 3. Three ORFs Are Required for *Trans*-Activation of a Late Promoter

Infected cells were treated with AraC and transfected with p11Xβ and plasmids pA1L, pA2L, or pGK1 or with the corresponding plasmids pA1LD, pA2LD, or pGK1D containing point mutations as indicated. The cells were lysed at 18 hr, and β-galactosidase activity was measured.

Identification of Three Late *Trans*-Activator Genes

From this point on, our goal was to obtain a minimal set of DNA segments necessary and sufficient to overcome the AraC block to expression of p11Xβ. To locate them, we transfected the vaccinia virus–infected cells with successively smaller plasmids containing subfragments of cosmids 21 and A as detailed in Experimental Procedures. In this manner, we identified three open reading frames (ORFs) required for *trans*-activation. Two of them, A1L and A2L (previously called ORF2 and ORF1, respectively; Weinrich and Hruby, 1986), are located adjacent to one another within a previously sequenced region at the extreme left end of the HindIII A fragment. A1L and A2L are predicted to encode polypeptides of 16,908 and 26,291 daltons, respectively. The third ORF, from a previously unsequenced region of the HindIII G fragment, was provisionally named GK1 and is predicted to encode a protein of 260 amino acids with a molecular weight of 29,914 (Figure 2). In cells infected with vaccinia virus and treated with AraC, cotransfection of all three ORFs with p11Xβ was required for significant β-galactosidase activity (Figure 3A).

To determine whether the intact ORFs were required for *trans*-activation, point mutations were made by filling in or digesting staggered ends produced by restriction endonucleases so as to shift the reading frames. DNA sequencing confirmed that the mutated A1L and A2L ORFs had two nucleotide insertions at codons 30 and 10, respectively. In the case of the GK1 ORF, sequencing revealed that a 3-base rather than the expected 4-base deletion had occurred. This mutation removed the arginine at amino acid 60 without altering the remaining downstream reading frame. The ability of each mutated plasmid to substitute for the original one was tested by transfection assays. In each case, the mutation destroyed the activity when assayed with authentic copies of the other two ORFs (Figure 3B), strongly suggesting that the translation products of all three ORFs are required for *trans*-activation and that the arginine at codon 60 of GK1 is essential.

The A1L, A2L, and GK1 ORFs Are Transcribed Only from Replicated or Transfected Templates

If the role of the transfected plasmids is to provide templates for transcription of the A1L, A2L, and GK1 ORFs, then it follows that they are normally not expressed until after DNA replication. Previous studies demonstrated that RNAs transcribed from A1L and A2L are present late in infection, but earlier times were not examined nor was the effect of inhibitors of DNA replication reported (Miner et al., 1988; Weinrich and Hruby, 1986). No data at all were available with regard to the time of transcription of GK1. To obtain this information, total RNA was isolated at 0, 2, 3.5, and 7 hr after vaccinia virus infection and hybridized to 5′ end ^{32}P-labeled single-stranded DNA probes specific for the 5′ ends of the putative *trans*-activator genes. As controls, the same RNA preparations also were hybridized to probes for the well-characterized early VGF (Yuen and Moss, 1986) and late 11K (Bertholet et al., 1985) genes, respectively. Following nuclease S1 digestion, the protected bands were resolved by polyacrylamide gel electrophoresis and visualized by autoradiography (Figure 4). The early VGF RNA was present at 2 hr after infection, increased at 3.5 hr, but was largely gone by 7 hr as revealed by the intensity of the 205 nucleotide protected band. In contrast, the 120 nucleotide protected band indicative of the 5′ end of the 11K late transcript was detected at 7 hr. Similarly, bands of 220, 232, and 135 nucleotides, indicative of the A1L, A2L, and GK1 transcripts, were present in appreciable amounts starting at 7 hr, consistent with

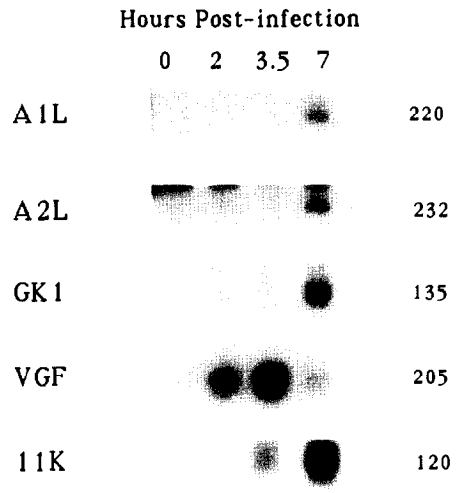

Figure 4. Presence of A1L, A2L, and GK1 Transcripts at Late Times after Infection

Total RNA was purified from HeLa suspension cells at 0, 2, 3.5, or 7 hr after vaccinia virus infection in the absence of inhibitors. The RNA was hybridized to 5′ ^{32}P-labeled single-stranded DNA probes specific for A1L, A2L, GK1, VGF, and 11K transcripts and then digested with nuclease S1. The protected DNA fragments were resolved by polyacrylamide gel electrophoresis alongside DNA markers. The relevant portions of the autoradiograms are shown, and the deduced lengths of the protected DNA segments are indicated on the right. At late times, there was an additional band apparently resulting from full probe protection by upstream readthrough RNAs (data not shown).

their synthesis after DNA replication. The sizes of the protected A1L and A2L DNA fragments indicated that the 5′ ends of the RNAs mapped just upstream of their respective ORFs in agreement with previous results (Weinrich and Hruby, 1986). A sequence ladder was used to locate precisely the 5′ end of the GK1 transcript (data not shown) to a location only a few nucleotides upstream of its ORF as well.

Next, we demonstrated that AraC prevented transcription of viral genomic A1L, A2L, and GK1 but not of the corresponding ORFs from transfected plasmids. High specific activity complementary RNA probes were used for hybridization to increase the sensitivity of RNA detection. Nuclease S1 digestion was carried out and the protected probes were resolved by polyacrylamide gel electrophoresis and autoradiography. As a control, we demonstrated that AraC had no effect on expression of the early VGF gene but completely abrogated transcription of the late 11K gene in virus-infected cells (Figure 5A). In the absence of AraC, correctly initiated A1L, A2L, and GK1 transcripts were detected (Figure 5B, right lanes). The higher of the two bands, seen with A1L and A2L, apparently results from readthrough by upstream RNA (vaccinia virus late RNAs are known to be long and heterogeneous, and it is typical to obtain complete probe protection). When AraC was present, transcription of genomic A1L, A2L, and GK1 was not detected (Figure 5B, left lanes). Significantly, however, the corresponding genes were transcribed from individually transfected pA1L, pA2L, and pGK1 in the presence of AraC (Figure 5B, middle lanes). In the case of pA2L, we differentiated the genome- and plasmid-derived transcripts by including a segment of plasmid sequence in the complementary RNA probe, which only hybridized to the latter transcript.

Specificity of *Trans*-Activation

Thus far, our attention had been directed toward the *trans*-activating genes rather than toward the late promoter target in p11Xβ. In particular, we needed to determine whether *trans*-activation occurred at the transcriptional level and whether the known sequence requirement of late promoters was retained. RNA probes complementary to the 5′ end of p11Xβ transcripts were hybridized to RNA from vaccinia virus–infected cells transfected with p11Xβ in the presence and absence of AraC. Formation of a nuclease S1–resistant band of approximately the size expected, when compared with DNA markers, occurred only when DNA replication was permitted (Figure 5C, upper panel). However, p11Xβ was transcribed despite the presence of AraC when pA1L, pA2L, and pGK1 also were transfected (Figure 5C, lower panel). Moreover, the major band detected under these conditions corresponded in size to the one formed in the absence of AraC. These data are consistent with transcriptional activation of the reporter gene, although an effect on mRNA stability was not ruled out.

If the late promoter was *trans*-activated by the normal transcriptional mechanism despite the absence of DNA replication in our transient assay, then the known stringent late promoter sequence specificity should be retained. To test this prediction, we used a set of four plasmids containing β-galactosidase controlled by a late promoter either with the conserved TAAAT sequence or with single inactivating nucleotide substitutions (Davison and Moss, 1989b). When the middle nucleotide was the authentic A, plasmids A1L, A2L, and GK1 stimulated activity 20-fold whereas relatively little or no *trans*-activation occurred when any other nucleotide was in this position (Table 2). This effect was similar to that observed with genome-integrated copies of the same reporter genes when assayed at late times of a normal infection (Davison and Moss, 1989b). Thus, *trans*-activation by A1L, A2L, and GK1 appears to be mediated by the normal late transcription mechanisms.

Plasmids A1L, A2L, and GK1 also stimulated β-galactosidase expression under control of additional late promoters, suggesting that they encode general late promoter *trans*-activators (data not shown).

Discussion

Coupling of Gene Regulation to Replication

In principle, DNA synthesis could regulate gene expression in at least three different but not necessarily exclusive ways. The first category involves template alterations associated with DNA replication such as replication forks, single strandedness, nicks, and concatemer junctions that are absent from nonreplicating molecules. Only a cru-

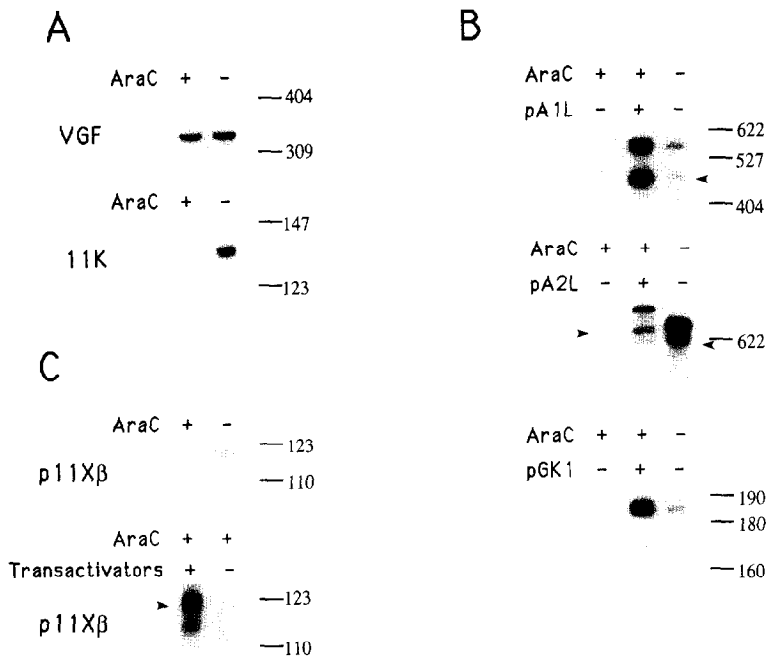

Figure 5. Effect of AraC on Transcription of Genomic and Transfected DNA

(A) RNA was purified from 293 cells infected in the presence or absence of AraC and hybridized to uniformly labeled VGF and 11K gene complementary RNA probes. After nuclease S1 digestion, the protected probes were resolved by polyacrylamide gel electrophoresis. The relevant portions of the autoradiogram and the positions of DNA markers are shown.

(B) RNA was analyzed as in (A), except that one set of infected and AraC-treated cells was transfected with either pA1L, pA2L, or pGK1, and uniformly labeled complementary RNA probes specific for these genes were employed. Arrowheads point to the predicted size products generated by RNAs that initiate just upstream of the ORFs.

(C) RNA was analyzed as in (A), except that all infected cells were transfected with p11Xβ and one set also was transfected with the three *trans*-activator plasmids pA1L, pA2L, and pGK1. The uniformly labeled complementary RNA probe was specific for the 5′ end of the reporter gene. The arrowhead points to the nuclease S1–protected band generated by RNA that initiates at the expected position.

cial *trans*-activator gene would have to be expressed from an altered region of the template, and the product of this gene could set off a cascade that leads to activation of late promoters. The second category, which may be difficult to distinguish from the previous one, involves protein complexes with a dual role in replication and transcription. In the third category, replication simply provides additional DNA, perhaps to titrate out an inhibitory binding protein or to provide an accessible template for newly synthesized transcription factors. We believe that accessibility of the template is important for the regulation of vaccinia virus late gene expression and that changes in primary DNA structure or replicating complexes need not be invoked.

Validity of the Experimental Approach

Before interpreting the results further, it is important to consider the validity of our experimental approach with regard to a natural infection. The basic protocol consisted of infecting cells with vaccinia virus and then transfecting them with a plasmid containing a reporter gene controlled by a natural and well-characterized late promoter. Expression not only was dependent upon both the virus infection and the late promoter but was blocked at the transcriptional level by inhibitors of DNA replication. In addition, single base mutations that abrogated activity of the promoter when recombined back into the virus genome did the same in the transfection assay. Thus, in every sense, this experimental system mimicked late promoter regulation during natural infection.

Table 2. Effect of Late Promoter Mutations on *Trans*-Activation

	β-Gal Units	
	A1L + A2L + GK1	
Late Promoter Mutations	−	+
TATAAAT	7.0	149.7
TATAGAT	6.9	8.7
TATATAT	6.0	18.9
TATACAT	11.6	9.9

Vaccinia virus–infected and AraC-treated cells were transfected with a plasmid containing a strong late promoter or a promoter with one of the above underlined mutations, attached to the β-galactosidase (β-gal) gene and with either pA1L, pA2L, and pGK1 (+) or salmon sperm DNA (−). At 18 hr postinfection cell lysates were assayed for β-gal activity.

Role of DNA Replication Is to Provide Unsequestered DNA Templates

A key finding was that transfected, deproteinized linear DNA that had been purified from virus particles was able to overcome the DNA replication requirement for expression of a late promoter–controlled reporter gene. The transfected viral genomic DNA worked in *cis* when it contained an integrated copy of the reporter gene and in *trans* when the reporter gene was in a cotransfected plasmid. Thus, the very same DNA that was nonfunctional with regard to activation of late promoters when its entry was via an infectious particle was functional when it was introduced as naked DNA. Additional experiments (data not shown) demonstrated that the reporter gene also had to be in replicated or transfected DNA, since it was not efficiently *trans*-activated when present in unreplicated genomic DNA that was introduced by infection at 0 time or by superinfection at late times in the presence of AraC.

The specificity of *trans*-activation was demonstrated by achieving the same result with two nonoverlapping cosmids containing vaccinia DNA and ultimately by three separate plasmids, each of which contained a single ORF. Moreover, transfection of plasmids containing the three

ORFs A1L, A2L, and GK1 were necessary and sufficient to activate the late promoter. We have not ruled out the possibilities, however, that some other ORFs could substitute for A1L, A2L, or GK1 or enhance their expression. Although the experiments described here were carried out in 293 cells with AraC as the inhibitor, similar *trans*-activation results were obtained using CV-1 cells and hydroxyurea (data not shown).

A1L, A2L, and GK1 Are Transcriptional *Trans*-Activators

Several lines of evidence indicate that the three ORFs encode *trans*-activators of late gene transcription. First of all, point mutations that altered the reading frame or deleted a single codon in each case led to loss of activity. Second, these ORFs are naturally transcribed only after DNA replication but are transcribed from the transfected plasmids even in the presence of AraC. Third, in the presence of AraC, transcription of the reporter gene depended upon transfection of the three ORFs. Although we have no data regarding the mechanism of *trans*-activation, it may not be coincidental that preliminary biochemical studies also point to a minimum of three late transcription factors (Wright and Moss, 1989).

The three ORFs have the capacity to encode proteins of 17, 26, and 30 kd. Computer searches revealed no significant relationship to any known viral or cellular proteins. However, the N-terminal sequence 1-MNLRLCSGCRHNGIVSEQGYEYCIFCESVF-30 of the 26 kd protein contains a potential zinc finger of the C4 type that is also present as a single copy in the *trans*-activation region of the adenovirus E1A protein (Moran and Mathews, 1987; Evans and Hollenberg, 1988; Culp et al., 1988; Pusztai et al., 1989). The sequence 51-KCWFCNQDLVFKPISIETFKGGEVGYFCSKICRNS-85 within the 17 kd protein resembles a zinc finger except that it has a 3 amino acid space between the second pair of cysteines. The sequence $IX_6LX_6LX_6$, resembling an atypically short leucine zipper (Vinson et al., 1989), occurs between amino acids 86 and 100 in the 30 kd protein (Figure 2). Further experiments are needed to determine whether these motifs have any functional correlates.

Comparison of A1L, A2L, and GK1 with Previously Described Intermediate Class Genes

The principle criterion used by Vos and Stunnenberg (1988) to classify I3 and I8 as intermediate rather than as late class postreplicative genes was that transfected copies were transcribed in the absence of viral DNA replication. Based on this, A1L, A2L, and GK1 also belong to the intermediate class. In the case of A2L, the transfected plasmid retained about 26 bp of DNA upstream of the RNA start site, indicating a promoter region no longer than that of vaccinia early promoters (Davison and Moss, 1989a). A comparison of the promoter regions of the five intermediate genes indicated some similarities to one another, particularly with regard to the sequence TNAAAT near the RNA start sites of all except I8, which has TANAAT, reminiscent of the highly conserved TAAAT of late promoters (Rosel et al., 1986; Hanngi et al., 1986). In addition, conserved A and T residues are present at specific upstream locations. Detailed mutational analyses will be necessary, however, to determine the significance of these sequences. The functions of I3 and I8 are unknown, and there is no evidence that either is involved in regulation of gene expression. However, because I3 has a compound promoter with both early and intermediate RNA start sites (Vos and Stunnenberg, 1988), it would be made from the infecting virus even in the presence of AraC and hence might not appear to stimulate late promoter expression in our assay even if it were a *trans*-activator.

A Regulatory Cascade

The data reported here provide evidence for the regulation of vaccinia virus gene expression by a cascade mechanism. The infectious virus particle contains the early transcription factor VETF as well as the multisubunit RNA polymerase and is programmed to synthesize early mRNA. The RNA polymerase and factors needed to express the A1L, A2L, and GK1 intermediate genes are present in infected cells prior to DNA replication and are likely to be early viral proteins, although some components might be released from the virion or derived from the host. The failure of the parental DNA to serve as a template for transcription of A1L, A2L, and GK1 suggests that it is sequestered either by residual virion components that deny access to newly synthesized transcription complexes (but not DNA polymerase) or by specific repressors that bind to intermediate promoters. Following replication, the progeny DNA serves as a template for transcription of A1L, A2L, and GK1, the products of which then *trans*-activate transcription of late genes. Because recent data indicate that the genes encoding both subunits of VETF have late promoters (P. Gershon and B. M., submitted), expression of A1L, A2L, and GK1 leads to the synthesis of the early transcription factor, which is then packaged into virus particles. In this cyclical model, early transcription factors are made at late times in the previous infection, intermediate transcription factors are made early in infection, and late transcription factors are made at intermediate times. DNA replication, by providing naked templates for intermediate and late gene expression, regulates the entire process.

Experimental Procedures

Viruses and Cells

Human 293 and CV-1 cells were grown as monolayer cultures in minimal essential media (MEM) with 5% fetal calf serum. Infections were usually carried out with vaccinia virus strain WR. In some experiments recombinant vaccinia virus vSC8 (Chakrabarti et al., 1985), which contains the Escherichia coli *lacZ* gene under control of the vaccinia virus 11K gene late promoter, was used. Cells were infected with a multiplicity of 10 plaque-forming units of virus- or mock-infected cells for 30 min at 37°C. Unadsorbed virus was removed by aspiration and the cell monolayers were covered with normal medium or medium containing 40 μg/ml of AraC to block DNA replication.

Transfection

Cells (1×10^6) were transfected with 3–40 μg of calcium phosphate–precipitated DNA (Cochran et al., 1985). The precipitate was formed by mixing the DNA in 250 mM $CaCl_2$ with 280 mM NaCl, 40 mM HEPES (pH 7.05), 2 mM Na_2HPO_4. Four hours after transfection, the liquid was removed and replaced with fresh normal medium or medium containing AraC. After 16–18 hr, the monolayers were washed

in phosphate-buffered saline (pH 7.0) and freeze-thawed lysates were tested for β-galactosidase activity (Miller, 1972). The activity was standardized to 1 mg of lysate protein. All transfections were carried out in duplicate and the average value was used.

Mapping the *Trans*-Activator Genes within Cosmids 21 and A

Vaccinia virus–infected cells were treated with AraC and transfected with p11Xβ, cosmids A and 19, and individual plasmids containing either the HindIII I, G, L, J, H, or D fragments of the vaccinia virus genome (Belle Isle et al., 1981). Cosmid 19 was not essential but was included because in some early experiments it appeared to increase activity. Significant β-galactosidase activity was obtained only when the HindIII G fragment was transfected. A similar experiment was carried out except that cosmids 21 and 19 were used with individual plasmids containing HindIII–BamHI or BamHI–BamHI subfragments of cosmid A. Of the eight plasmids tested, only the one with the leftmost HindIII–BamHI fragment (pA34) stimulated β-galactosidase. To consolidate these findings, we demonstrated that plasmids HindIII G and A34, together and without any other DNA in addition to p11Xβ, stimulated activity. Sequence information was used to identify further the DNA segments in plasmids A34 and HindIII G that are required to overcome the replication block for late promoter expression. The DNA in pA34 was sequenced previously (Weinrich and Hruby, 1986) and contains four complete ORFs: A1L, A2L, A3L, and A4L. To determine whether any of these correlated with *trans*-activating activity, a variety of subclones containing complete individual ORFs were made. Cotransfections of individual plasmids containing ORFs A1L and A2L were necessary to replace pA34.

We turned next to a dissection of the HindIII G fragment. Because this 10,000 bp segment of the vaccinia virus genome had not yet been sequenced, several restriction endonucleases were tested for their ability to cleave the DNA. Plasmids recircularized after complete or partial restriction endonuclease digestions were cotransfected with plasmids A34 and 11Xβ and tested for their ability to *trans*-activate p11Xβ. Examination of the cloned active fragments suggested that the right end of the HindIII G is necessary for *trans*-activation of late genes. We sequenced the vaccinia DNA contained in the smallest active plasmid and found only one complete ORF, provisionally named GK1.

DNA Sequencing

DNA sequencing was performed on double-stranded plasmids using the dideoxynucleotide procedure described in the Sequenase Kit (United States Biochemical). Specific oligonucleotide primers were used to sequence both strands of two individual XbaI clones containing the GK1 ORF.

Nuclease S1 Analysis of RNA

RNA was extracted from uninfected or infected cells at indicated times and purified using the guanidine isothiocyanate–CsCl method (Maniatis et al., 1982). Complementary DNA probes were synthesized by the unidirectional polymerase chain reaction. Primers were end labeled using [γ-^{32}P]ATP and T4 polynucleotide kinase. Each gene was cut at sites up- and downstream of the primer binding site. A 643 base 11K gene probe was prepared from a 650 bp HindIII–BamHI fragment; a 730 base VGF early RNA probe from a 880 bp HindIII–AccI fragment; a 569 base GK1 probe from a 630 bp KspI–XbaI fragment; a 310 base A1L probe from a 555 bp HindIII–BglII fragment; and a 230 base A2L probe from a 640 bp BglII–XbaI fragment. The polymerase chain reaction was for 35 cycles and the products were purified on 6% polyacrylamide gels. Probes, ~50,000 cpm, were added to 25 μg of total RNA in a final volume of 30 μl of 40 mM PIPES (pH 6.4), 1 mM EDTA (pH 8.0), 0.4 M NaCl, 80% formamide and heated at 100°C for 2 min prior to an overnight incubation at 42°C. The RNA–DNA hybrids were digested with 500 U of S1 nuclease at room temperature for 60 min. The primer used to make the single-stranded probe was also used to generate the sequencing ladder for mapping the 5′ end of the GK1 mRNA.

Internally labeled [α-^{32}P]UTP complementary RNA probes were also used for nuclease S1 analysis of RNA made from infected and transfected cells. For these experiments, RNA present at 18 hr after infection was extracted by the RNAzol (Cinna/Biotecx) method. The following fragments were cloned into pGEM-3Z or pGEM-4Z (Promega): A1L, −111 (BglII) to +453 (HindIII); A2L, −24 (XbaI) to +617 (BglII); GK1, −448 (XbaI) to +181 (KspI); 11Xβ, +407 (AatII) to +113 (PvuII); VGF, +328 (HincII) to +309 (AccI); 11K, −550 (BamHI) to +120 (HindIII). Plasmids were purified and in vitro T7 transcripts were made as recommended by Promega. Total RNA (10 μg) was hybridized to the appropriate probe overnight at 42°C and digested with 500 U of S1 nuclease at room temperature for 1 hr. Protected fragments were then analyzed on 6% or 8% polyacrylamide sequencing gels.

Mutagenesis Procedures

The GK1 ORF was mutated by cutting with restriction enzyme KspI, removing the two nucleotides overhanging both strands with T4 DNA polymerase, and religating the plasmid. The A1L and A2L ORFs were partially digested with AccI, the two nucleotide staggered ends were filled in with T4 DNA polymerase, and the plasmid was recircularized.

Acknowledgments

We thank M. Merchlinsky for providing many of the reagents and for helpful discussions and M. Challberg, P. Gershon, B. Ahn, and other members of the Laboratory of Viral Diseases for suggestions and criticisms. J. G. K. is a National Research Council Associate.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 5, 1990; revised March 13, 1990.

References

Belle Isle, H., Venkatesan, S., and Moss, B. (1981). Cell-free translation of early and 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.

Bertholet, C., Drillien, R., and Wittek, R. (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.

Boone, R. F., and Moss, B. (1978). Sequence complexity and relative abundance of vaccinia virus mRNA's synthesized *in vivo* and *in vitro*. J. Virol. *26*, 554–569.

Broyles, S. S., and Moss, B. (1988). DNA-dependent ATPase activity associated with vaccinia virus early transcription factor. J. Biol. Chem. *263*, 10761–10765.

Broyles, S. S., Yuen, L., Shuman, S., and Moss, B. (1988). Purification of a factor required for transcription of vaccinia virus early genes. J. Biol. Chem. *263*, 10754–10760.

Chakrabarti, S., Brechling, K., and Moss, B. (1985). Vaccinia virus expression vector: coexpression of β-galactosidase provides visual screening of recombinant virus plaques. Mol. Cell. Biol. *5*, 3403–3409.

Cochran, M. A., Mackett, M., and Moss, B. (1985). Eukaryotic transient expression system dependent on transcription factors and regulatory DNA sequences of vaccinia virus. Proc. Natl. Acad. Sci. USA *82*, 19–23.

Crossland, L. D., and Raskas, H. J. (1983). Identification of adenovirus genes that require template replication for expression. J. Virol. *46*, 737–748.

Culp, J. S., Webster, L. C., Friedman, D. J., Smith, C. L., Huang, W.-J., Wu, F. Y.-H., Rosenberg, M., and Ricciardi, R. P. (1988). The 289-amino acid E1a protein of adenovirus binds zinc in a region that is important for *trans*-activation. Proc. Natl. Acad. Sci. USA *85*, 6450–6454.

Davison, A. J., and Moss, B. (1989a). The structure of vaccinia virus early promoters. J. Mol. Biol. *210*, 749–769.

Davison, A. J., and Moss, B. (1989b). The structure of vaccinia virus late promoters. J. Mol. Biol. *210*, 771–784.

Evans, R. M., and Hollenberg, S. M. (1988). Zinc fingers: gilt by association. Cell *52*, 1–3.

Gaynor, R. B., and Berk, A. J. (1983). *Cis*-acting induction of adenovirus transcription. Cell *33*, 683–693.

Hanngi, M., Bannwarth, W., and Stunnenberg, H. G. (1986). Con-

served TAAAT motif in vaccinia virus late promoters: overlapping TATA box and site of transcription initiation. EMBO J. *5*, 1071–1076.

Herendeen, D. R., Kassavetis, G. A., Barry, J., Alberts, B. M., and Geiduschek, E. P. (1989). Enhancement of bacteriophage T4 late transcription by components of the T4 DNA replication apparatus. Science *245*, 952–958.

Jones, E. V., Puckett, C., and Moss, B. (1987). DNA-dependent RNA polymerase subunits encoded within the vaccinia virus genome. J. Virol. *61*, 1765–1771.

Kates, J. R., and McAuslan, B. (1967). Messenger RNA synthesis by a "coated" viral genome. Proc. Natl. Acad. Sci. USA *57*, 314–320.

Kaverin, N. V., Varich, N. L., Surgay, V. V., and Chernos, V. I. (1975). A quantitative estimation of poxvirus genome fraction transcribed as early and late mRNA. Virology *65*, 112–119.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Mavromara-Nazos, P., and Roizman, B. (1987). Activation of herpes simplex virus 1 γ_2 genes by viral DNA replication. Virology *161*, 593–598.

Miller, J. H. (1972). Experiments in Molecular Genetics (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory), pp. 352–355.

Miner, J. N., Weinrich, S. L., and Hruby, D. E. (1988). Molecular dissection of *cis*-acting regulatory elements from 5′-proximal regions of a vaccinia virus late gene cluster. J. Virol. *62*, 297–304.

Moran, E., and Matthews, M. B. (1987). Multiple functional domains in the adenovirus E1A gene. Cell *48*, 177–178.

Moss, B. (1990). Regulation of vaccinia virus transcription. Annu. Rev. Biochem., in press.

Moss, B., and Salzman, N. P. (1968). Sequential protein synthesis following vaccinia virus infection. J. Virol. *2*, 1016–1027.

Munyon, W. H., and Kit, S. (1966). Induction of cytoplasmic ribonucleic acid synthesis in vaccinia-infected LM cells during inhibition of protein synthesis. Virology *29*, 303–306.

Nevins, J. R. (1989). Mechanisms of viral-mediated *trans*-activation of transcription. Adv. Virus Res. *37*, 35–83.

Oda, K., and Joklik, W. K. (1967). Hybridization and sedimentation studies on "early" and "late" vaccinia messenger RNA. J. Mol. Biol. *27*, 395–419.

Paoletti, E., and Grady, L. J. (1977). Transcriptional complexity of vaccinia virus in vivo and in vitro. J. Virol. *23*, 608–615.

Pennington, T. H. (1974). Vaccinia virus polypeptide synthesis: sequential appearance and stability of pre- and post-replicative polypeptides. J. Gen. Virol. *25*, 433–444.

Pusztai, R., Loewenstein, P. M., and Green, M. (1989). Mutational analysis of autonomously functioning *trans*-activating peptides encoded by adenovirus E1A: evidence for two subdomains. J. Virol. *63*, 3516–3518.

Rosel, J. L., Earl, P. L., Weir, J. P., and Moss, B. (1986). Conserved TAAATG sequence at the transcriptional and translational initiation sites of vaccinia virus late genes deduced by structural and functional analysis of the HindIII H genome fragment. J. Virol. *60*, 436–439.

Schmitt, J. F. C., and Stunnenberg, H. G. (1988). Sequence and transcriptional analysis of the vaccinia virus HindIII I fragment. J. Virol. *62*, 1889–1897.

Schwer, B., and Stunnenberg, H. G. (1988). Vaccinia virus late transcripts generated in vitro have a poly(A) head. EMBO J. *7*, 1183–1190.

Thomas, G. P., and Mathews, M. B. (1980). DNA replication and the early to late transition in adenovirus infection. Cell *22*, 523–533.

Vinson, C. R., Sigler, P. B., and Mcknight, S. L. (1989). Scissors-grip model for DNA recognition by a family of leucine zipper proteins. Science *246*, 911–916.

Vos, J. C., and Stunnenberg, H. G. (1988). Derepression of a novel class of vaccinia virus genes upon DNA replication. EMBO J. *7*, 3487–3492.

Weinrich, S. L., and Hruby, D. E. (1986). A tandemly-oriented late gene cluster within the vaccinia virus genome. Nucl. Acids Res. *14*, 3003–3016.

Wright, C. F., and Moss, B. (1987). In vitro synthesis of vaccinia virus late mRNA containing a 5′ poly(A) leader sequence. Proc. Natl. Acad. Sci. USA *84*, 8883–8887.

Wright, C. F., and Moss, B. (1989). Identification of factors specific for transcription of the late class of vaccinia virus genes. J. Virol. *63*, 4224–4233.

Yuen, L., and Moss, B. (1986). Multiple 3′ ends of mRNA encoding vaccinia virus growth factor occur within a series of repeated sequences downstream of T clusters. J. Virol. *60*, 320–323.