Minireview

Cytoplasmic Transcription System Encoded by Vaccinia Virus

Bernard Moss, Byung-Yoon Ahn, Bernard Amegadzie, Paul D. Gershon, and James G. Keck

From the Laboratory of Viral Diseases, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892

The poxviruses comprise the only known family of DNA viruses that propagate entirely within the cytoplasm of eukaryotic cells and that encode most, if not all, of the specific enzymes and factors needed for transcription and replication (reviewed in Refs. 1 and 2). The “life cycle” of vaccinia virus, the prototypal member of the family, is surprisingly complex with at least three temporally regulated classes of gene products (Fig. 1). The 200,000-base pair (bp) 1 linear double-stranded DNA genome and the components of the early transcription system are packaged within the core of the infectious virus particle so that mRNA synthesis begins immediately after entry into the cytoplasm (3, 4). The subsequent synthesis of DNA polymerase and other viral early proteins leads to DNA replication, which is followed by the intermediate and late phases of gene expression (5–8). The considerable progress that has been made in identifying the cis- and trans-acting components of the regulatory cascade is described in this review.

Early Transcription

Methods of Study—Vaccinia virus transcription can be studied conveniently in vivo or in vitro. Structurally and functionally, the RNA species made in vitro by permeabilized purified virus particles resemble those made early in infection (9–11). ATP hydrolysis is required for initiation (12) and elongation (13) of transcription. The transcripts accumulate transiently in the virus core, and high concentrations of ATP with a hydrolyzable β-γ bond are required for extrusion of the RNA (14). The average time for synthesis and release of mRNAs is about 2.2 min (15).

Highly active template-dependent transcription systems were obtained by disruption of the virions and removal of insoluble structural proteins and nucleic acids (16, 17). The ability to faithfully initiate and terminate transcription of added early gene templates is retained in these extracts (18) which also serve as a source of individual enzymes and factors involved in the synthesis and processing of mRNA.

The ability to ligate natural or mutated transcriptional regulatory signals to reporter genes and integrate them into the genome of infectious virus particles (19) has provided a reliable way of correlating structure with function in vivo.

mRNA—The mRNAs produced in vitro by virus cores or in infected cells are capped (10, 20), polyadenylated (9), and of discrete length with no indication of splicing. Some early mRNAs also contain short 5’ poly(A) leaders (21, 22) that are synthesized by an RNA editing process that will be discussed in conjunction with late mRNAs.

Promoters—The promoters for several early genes were found to extend only about 30 bp upstream of the RNA start sites. Within one early promoter, the effects of all single nucleotide substitutions were measured by expression of the β-galactosidase reporter gene (Fig. 2) and confirmed by in vivo and in vitro transcription (23). On the basis of these results, the promoter was divided into three regions relative to the RNA start site at +1: a 15-bp A-rich critical region (−13 to −28) in which many single nucleotide substitutions have a major effect, separated by 11 bp of a less critical T-rich sequence from a 7-bp region within which initiation at a purine occurs. Within the critical region, A residues are essential at certain locations and optimal at some others. A G residue is needed at −21 and Ts are important at −22 or −23. Similar to the TATA box of higher eukaryotic RNA polymerase II promoters, the critical region specifies the distance to the transcription initiation site downstream. Most natural promoters do not have optimal nucleotides in all positions, and variability in promoter strength may provide a way of regulating gene expression. When the promoter sequences of a large number of early genes were lined up, however, the predominant nucleotide at each position corresponded closely to the optimal one determined by mutagenesis.

Termination Signal—The DNA sequence corresponding to the eukaryotic processing signal AAUAAA (24), which appears about 20 nucleotides before the polyadenylation site of eukaryotic mRNAs, is not present near the 3’ ends of poxvirus genes. Instead, the ends of vaccinia virus early mRNAs occur 20–50 bp downstream of the sequence TTTTTNT, in which N can be any nucleotide including T (25). Termination, rather than RNA processing, is believed to occur for kinetic reasons and because of the absence of specific endonuclease activity (18). Incorporation of halogenated UTP derivatives into nascent transcripts prevented termination in a template-dependent system, suggesting that the signal is actually recognized in the RNA as UUUUUNU (15a). The same derivatives also blocked RNA extrusion from cores and led to the accumulation therein of long transcripts (15b).

As predicted, TTTTTTNT occurs near the ends of most early genes, but in its absence mRNAs with ends that are co-terminal with those of downstream mRNAs may be formed. In vivo studies suggest that termination is less than 100% efficient (26), and the occasional occurrence of TTTTTTNT sequences within early genes may therefore down-regulate but not entirely prevent expression.

Virus Core-associated Enzymes

DNA-dependent RNA Polymerase—The α-aminitin-resistant RNA polymerase from vaccinia virions has an apparent molecular mass of nearly 500 kDa with two large and many small subunits (27), thus resembling its eukaryotic counterpart. Nevertheless, all of the subunits appear to be virus-encoded (28). The genes for the two large (rpo147 and rpo132) and five of the small (rpo35, rpo30, rpo22, rpo19, and rpo18) polypeptides (22, 29, 30–32) have been identified (Table I). The large subunits are homologous to the corresponding subunits of eukaryotes and have significant but less sequence similarity to those of Escherichia coli. The rpo147 and rpo132
VETF has DNA-dependent ATPase activity that might account for the ATP hydrolysis requirement for initiation of RNA synthesis (38). The viral genes encoding the 70- and 82-kDa subunits of VETF have been identified (Table II), and both are expressed late in infection (39, 40). The protein sequence predicted for the smaller of the two subunits contains a motif associated with ATP binding or ATPase activity consistent with the properties of VETF; the larger polypeptide contains variations of the canonical zinc finger and leucine zipper motifs (39).

**Capping and Methylating Enzymes**—Capping and methylating of vaccinia virus mRNA is accomplished by a multifunctional enzyme complex composed of 97- and 33-kDa subunits that catalyze the removal of the terminal or γ-phosphate from an RNA chain, the transfer of the GMP moiety of GTP to the now diphosphate-ended RNA, and finally the transfer of a methyl group from S-adenosylmethionine to the 7-position of the added guanosine which stabilizes the final product against reversal by pyrophosphate ion (41, 42). The 97-kDa subunit reacts with GTP to form a covalent lysine-GMP intermediate (43, 44). Since short oligoribonucleotides can serve as cap acceptors, capping might occur shortly after transcription initiation, consistent with the finding of capping enzyme associated with the transcription complex (45). The genes encoding both subunits of capping enzyme have been identified (46, 47) and expressed as a functional heterodimer in *Escherichia coli* (48–50). The large subunit alone has RNA triphosphatase and guanylyltransferase but not methyltransferase activities.

A 38-kDa RNA (nucleoside-2′-O-)methyltransferase catalyzes the transfer of a methyl group from S-adenosylmethionine to the ribose of the first encoded nucleotide, which is separated by a triphosphate bridge from the terminal 7-methylguanosine (51). Uncapped polynucleotides or even capped ones lacking the 7-methyl group are not methyl acceptors, indicating that 2′-O-methylation is the final step in the formation of the cap structure m^7^G(5')pppNm- (52). The gene encoding this enzyme has not been identified nor has the role of ribose methylation been determined.

### Table I

| Gene  | Homolog   | Sequence motif | Ref. |
|-------|-----------|----------------|------|
| rpo147| RPB1 (yeast) | Zinc finger | 29   |
| rpo132| RPB2 (yeast) | Zinc finger, leucine zipper | 30, 31 |
| rpo35 | SI1 (mouse) | Zinc finger | 32   |
| rpo30 | Leucine zipper | 29   |
| rpo19 | Leucine zipper, acidic | N terminus | 22   |
| rpo18 |            |                |      |

* B. Amegadzie, unpublished data.
* B-Y. Ahn, unpublished data.

subunits both have zinc finger motifs, but the rpo147 subunit lacks the heptapeptide C-terminal repeats of the related RNA polymerase II subunit. Interestingly, the rpo30 subunit appears to be a homolog of eukaryotic transcription elongation factor SI1 (32). All of the vaccinia virus RNA polymerase subunit genes have early promoters and some have late promoters in addition. A role for the host RNA polymerase, or some of its subunits, in expression of poxvirus genes has been suggested (33–35).

**Early Transcription Factor**—Purified vaccinia virus RNA polymerase lacks the ability to transcribe double-stranded DNA templates unless complemented with an early transcription factor called VETF that was isolated from virion extracts (36). DNA affinity-purified VETF is a heterodimeric protein that works in conjunction with all early promoters tested. The DNA binding of VETF is specific and dependent on the critical region of the early promoter, being abrogated by single nucleotide substitutions that decrease transcription (37).

![Diagram of vaccinia virus gene expression](image-url)

**Fig. 1.** Temporal regulation of vaccinia virus gene expression.

![Graph of β-galactosidase activity](image-url)

**Fig. 2.** Effect of single nucleotide substitutions on early promoter activity. Recombinant vaccinia viruses that contain the indicated natural promoter sequences were used to infect cells in the presence of cytosine arabinoside, an inhibitor of DNA replication. The activities obtained with the mutated promoters are given relative to the 100% value obtained with the natural promoter. Data are from Ref. 23.

### Table II

| Factor | Sequence motif | Ref. |
|--------|----------------|------|
| VETF (82 kDa) | Leucine zipper, zinc finger | 39   |
| VETF (70 kDa) | Helicase | 39, 40 |
| VLTA (30 kDa) | Zinc finger | 69   |
| VLTA (26 kDa) | Zinc finger | 69   |
| VLTA (17 kDa) | Zinc finger | 69   |

* VLTA, vaccinia virus late transactivator.
polypeptides associated with purified poly(A) polymerase have been identified. Additional Enzymes—Other enzymes that may have roles in transcription have been purified from vaccinia virus cores. These include nucleoside-triphosphate phosphohydrolase I, a virus-encoded DNA-dependent ATPase of M, 61,000 (57–59). A conditionally lethal temperature-sensitive nucleoside-triphosphate phosphohydrolase I mutant is defective in intermediate and late gene expression suggesting that the enzyme may be involved in transcription (60, 61).

Nucleoside-triphosphate phosphohydrolase II, a M, 68,000 protein that is immunologically distinct from nucleoside-triphosphate phosphohydrolase I, is stimulated by a wider range of nucleic acids than nucleoside-triphosphate phosphohydrolase I and hydrolyzes all four nucleoside triphosphates rather than just ATP (62). The gene encoding nucleoside-triphosphate phosphohydrolase II has not yet been identified nor is there information regarding its function.

The vaccinia virus-encoded DNA topoisomerase is able to relax both positively and negatively supercoiled DNA and has the properties of a cellular type I enzyme (63, 64). Although failed attempts to insertional inactivate the topoisomerase gene suggest that it is essential (65), the enzyme is not required for transcription of linear DNA templates in vitro (36). The topoisomerase might be required with more topologically constrained natural templates in vivo or have an entirely unrelated function. Additional core-associated enzymes (reviewed in Ref. 2) include a protein kinase and deoxyribonuclease/ligase.

Intermediate Transcription

Although the existence of at least two temporal classes of postreplicative genes was predicted by analyzing the time course of polypeptide synthesis in vaccinia virus-infected cells (7, 8, 66), direct evidence for a class of intermediate genes, distinct from late, was only recently obtained (67). Whereas viral DNA replication is required for expression of true late genes regardless of whether they are genomic or present in transfected plasmids (68), transfected copies of intermediate genes are expressed in the presence of inhibitors of DNA replication (67). The latter data suggest that transacting factors required for intermediate expression are present prior to replication and hence are early viral and/or cellular proteins. There are now five known examples of intermediate genes (67, 69), but a promoter consensus sequence has not been determined.

Late Transcription

mRNA—Late mRNAs have several unusual features. Transcription termination signals for late transcription have not been recognized, and the 3′ ends of late mRNAs are heterogeneous in length (70, 71). Moreover, both strands of DNA are transcribed, isolated late RNAs can be annealed with other late RNAs or with early RNAs to form ribonuclease-resistant hybrids (72, 73). Whether anti-sense RNA has a role in vivo is unknown, however. The discrete 3′ end of one major cowpox virus transcript provides an exception to the rule of length heterogeneity of late RNA (74).

A capped 5′ poly(A) tract of approximately 35 nucleotides is a characteristic feature of late mRNAs (74–77) whereas a shorter one is present in a minority of early mRNAs (21, 22). The poly(A) leader is not encoded as such within the genome and probably arises by RNA polymerase slippage when initiating within the highly conserved TAAAT sequence, as discussed below. Some speculations regarding the role of the capped 5′ poly(A) leader include a binding site for initiation factors and the 40 S ribosomal subunit, which would then scan, unimpeded by anti-sense RNA, to the first AUG (usually located immediately after the poly(A) leader) where ribosome assembly and translation occur.

Promoters—The late promoter may be considered in terms of three regions: an upstream sequence of about 20 bp with some consecutive T or A residues, separated by a region of about 6 bp from a highly conserved TAAAT element within which transcription initiates (78). Mutations within the A triplet of TAAAT drastically decreased transcription (78–80), and substitution of the flanking T residues also had a negative effect but to a degree that depended inversely on the strength of the promoter as determined by upstream sequences (78). Immediately downstream of the TAAAT, a G is optimal for promoter activity and an A is second best. Single nucleotide substitutions within the 6 bp upstream and 3 bp downstream of TAAAT had relatively modest effects on promoter strength (78). The region upstream of −7 is essential for late promoter function and usually contains runs of A or T residues with the latter having a much greater activating effect. A very strong synthetic promoter was constructed with runs of 15 or 20 T residues (78). Regardless of promoter strength, however, initiation always occurred within the A triplet.

The 5′ poly(A) leader was diminished in length when the T residues of TAAAT were mutated, and the leader was absent or limited to a few nucleotides when any of the three A residues themselves were mutated (78). Shortening of poly(A) also was noted when mutations further downstream were made (81). These data suggested that the poly(A) leader is formed by a mechanism involving backward slippage of the RNA polymerase.

In Vitro Transcription System—Extracts prepared from cells at late times after infection are capable of transcribing vaccinia virus late genes, and the resulting RNAs contain 5′ poly(A) leaders (82, 83). After passage of the cytoplasmic extract through a phosphocellulose column and stepwise elution, three fractions were obtained which together but not separately could transcribe late promoter templates (64). One of the factors called VLTF-1 was partially purified. It seems likely that the RNA polymerases used for early and late transcription are similar or at least share subunits since conditionally lethal mutations in the large subunit and in small subunit include decreased expression of late genes (85, 86).

Transactivators of Late Gene Expression—A novel transfection approach was used to identify intermediate genes needed for late transcription (69). Earlier studies had indicated that intermediate and late promoter controlled reporter genes were expressed from transfected plasmids provided the cells were infected with vaccinia virus (67, 68). Significantly, however, only the intermediate promoters were active when viral DNA replication was inhibited. This replication block to expression of a late promoter-controlled reporter gene on a plasmid could be bypassed by also transfecting naked virion DNA (69). The role of the transfected virion DNA was investigated by substituting a library of cloned vaccinia virus DNA fragments. It was found, by repeated subcloning and transfection, that three intermediate class genes encoding polypeptides of 17, 26, and 30 kDa were both necessary and sufficient for expression of a transfected late gene in the presence of an inhibitor of DNA replication. Further studies revealed that the 30-kDa protein corresponds to VLTF-1 and is thus a transcription factor.

There are several possible reasons why the DNA in the

---

5 P. D. Gershon, B.-Y. Ahn, M. Garfield, and B. Moss, manuscript in preparation.

6 C. Wright, J. Keck, and B. Moss, manuscript in preparation.
VIRUS CORE

ENTER 

LATE mRNA

FIG. 3. Transcriptional regulatory cycle. ETF, early transcription factor; ITF, intermediate transcription factor(s); LTF, late transcription factor(s).

infecting virus particle only serves as a template for early gene expression whereas replicated or transfected DNA serves as a template for intermediate and late gene expression. Initially, the virion proteins may sequester the input DNA from newly synthesized enzymes and factors required for the expression of later classes of genes. Any input DNA that becomes more accessible to exogenous proteins may rapidly associate with the replication apparatus. Alternatively, expression of the intermediate and late classes of genes might be prevented by the presence of specific repressor proteins on the packaged DNA.

Regulatory Cycle

An oversimplified diagram of the transcriptional regulatory cycle is presented in Fig. 3. Poxvirus gene expression begins with the entry of the virus into the cytoplasm of the cell. The virus core, with its associated enzymes including RNA polymerase, VETF, guanylyltransferase, methyltransferases, and poly(A) polymerase, is programmed to express early genes and produce functional mRNAs. The translation products of these early mRNAs include RNA polymerase subunits, DNA polymerase, and putative factors for specific transcription of intermediate genes. The parental DNA does not serve as a template for intermediate gene expression, perhaps because its location in the core or commitment to DNA replication restricts access by a newly synthesized transcription complex or because specific repressors of intermediate genes. With the occurrence of DNA replication, the cycle enters its second phase and templates are available for transcription of the three intermediate genes identified as transactivators of late gene expression. In the third phase of the cycle, the late genes encoding virion enzymes, early transcription factor, and structural proteins are expressed. Finally, the progeny virus particles are assembled and some are released ready to begin the cycle again.

REFERENCES

1. Moss, B. (1989) in Virology (Fields, B. N., Knipe, D. M., Chanock, R. M., Hirsch, M. S., Melnick, J., Monath, T. P., and Roizman, B., eds.) pp. 2079-2112, Raven Press, New York.
2. Moss, B. (1990) Annu. Rev. Biochem. 59, 661-688.
3. Munyon, W. H., and Kat, S. (1986) Virology 150, 303-306.
4. Kates, J. R., and McAulay, B. (1987) Proc. Natl. Acad. Sci. U. S. A. 87, 3142-320.
5. Oda, K., and Joklik, W. K. (1977) J. Mol. Biol. 77, 305-419.
6. salmon, N. P., and Sebring, E. (1967) J. Virol. 1, 16-23.
7. Moss, B., and Salmon, N. P. (1989) J. Virol. 2, 1016-1027.
8. Pennington, T. H. (1974) J. Gen. Virol. 25, 443-444.
9. Kates, J. R., and Beeson, J. (1970) J. Mol. Biol. 50, 19-23.
10. Wei, C. M., and Moss, B. (1975) Proc. Natl. Acad. Sci. U. S. A. 72, 318-322.
11. Cooper, J. A., and Moss, B. (1978) Virology 88, 149-165.
12. Gershovitz, A., Boone, R. F., and Moss, B. (1978) J. Virol. 27, 399-408.
13. Shuman, E., Spenczyk, E., and Hurwitz, J. (1966) J. Biol. Chem. 235, 5364-5460.
14. Veomett, G. E., and Kates, J. R. (1973) ICN-UCLA Symp. Mol. Cell. Biol. 127-142.
15. Shuman, S., and Moss, B. (1988) J. Biol. Chem. 263, 6220-6225.