VIROLOGY 185, 432–436 (1991)

Antitermination of Vaccinia Virus Early Transcription: Possible Role of RNA Secondary Structure

YAN LUO AND STEWART SHUMAN

Program in Molecular Biology, Sloan-Kettering Institute, New York, New York 10021

Received June 10, 1991; accepted July 18, 1991

Transcription of vaccinia early genes by the viral RNA polymerase terminates downstream of a signal sequence TTTTTNT in the nontemplate DNA strand. Signal recognition occurs at the level of the sequence UUUUUNU in nascent RNA and depends on a virus-encoded termination factor (VTF). The presence of TTTTTNT elements within protein encoding regions of some early genes requires that these 5′ proximal signals be ignored in order to achieve early expression of the full-sized proteins. In the case of the A18R gene, which contains a proximal terminator that is not utilized *in vivo* (Pacha *et al., J. Virol.* 64, 3853–3863 (1990)), the TTTTTNT sequence can be folded into a potential hairpin structure such that UUUUUNU would be part of a duplex stem in the nascent RNA. We find that the A18R putative hairpin is unable to promote factor-dependent termination in a purified *in vitro* transcription system. Sequence manipulations that abrogate the potential to form an RNA hairpin restore the activity of the TTTTTNT motif. The *in vitro* studies suggest that antitermination at the proximal site of the A18R gene may be mediated by secondary structure in the nascent RNA, and that early termination involves recognition by VTF and/or RNA polymerase of the UUUUUNU sequence in single-stranded form. © 1991 Academic Press, Inc.

Recent discoveries have focused attention on transcription termination as an important regulatory step in eukaryotic gene expression (reviewed in Refs. *1, 2*). Synthesis of mRNA can be modulated by premature termination of transcription (or elongation block) within promoter proximal regions, a process loosely comparable to prokaryotic attenuation. Examples include the proto-oncogenes c-myc (*3, 4*), c-fos (*5*), c-myb (*6*), and L-myc (*7*), the Drosophila hsp70 gene (*8*), the murine and human adenine deaminase genes (*9, 10*), and the *Xenopus laevis* α-tubulin gene (*11*). Termination control has also been implicated in the expression strategies of many animal viruses, including adenovirus (*12*), polyoma (*13*), SV40 (*14*), minute virus of mice (*15*), human immunodeficiency virus (*16*), and vaccinia virus (*17*).

The molecular details of mRNA transcription termination by a eukaryotic RNA polymerase are most clearly appreciated in the case of vaccinia virus. Vaccinia, which replicates exclusively in the cytoplasm, encapsidates within the virion all the components required for transcription of the approximately 100 early viral genes (*17*). These include a virus-encoded multisubunit RNA polymerase with structural and functional homology to cellular RNA polymerase II, a virus-encoded mRNA capping enzyme, and a virus-encoded poly(A) polymerase (*17*). What makes vaccinia especially well suited to studies of termination is the fact that the 3′ ends of primary transcripts of early genes arise via bona fide termination rather than endonucleolytic cleavage. Moreover, accurate termination can be recapitulated *in vitro* with soluble transcription systems derived from virion extracts (*18*). Termination requires a simple *cis*-acting sequence TTTTTNT in the nontemplate DNA strand that is sufficient to induce termination *in vivo* and *in vitro* at heterogeneous sites 20–50 nucleotides downstream of the signal (*19*). Vaccinia RNA polymerase cannot terminate in response to the TTTTTNT signal, but requires a separate *trans*-acting termination factor (VTF) that is identical to the vaccinia-encoded mRNA capping enzyme (*20, 21*). Base analog substitution effects indicate that the signal for factor-dependent termination is actually the sequence UUUUUNU in the nascent RNA (*21, 22*).

Termination of transcription is subject to strict temporal control during the vaccinia life cycle. The transition from early to late gene expression (contingent upon DNA replication) entails not only a switch in promoter recognition by vaccinia RNA polymerase, but also an abrogation of the early mode of 3′ end formation. Indeed, for the majority of late transcription units, termination occurs nonspecifically. Thus, while early messages are of discrete size, late transcripts are severalfold longer and considerably more heterogeneous (*23*). That the early termination mechanism is overridden at late times is suggested by the distribution of TTTTTNT sequences within viral transcription units. The TTTTTNT motif is found upstream of the 3′ ends of all early transcripts whose termini have been determined experimentally and is conspicuously absent from the coding regions of most early genes. In contrast, TTTTTNT sequences are encountered often (and

0042-6822/91 $3.00

in multiple copies) within the coding regions of most late viral genes. The basis of the failure to terminate late is obscure, but of considerable interest, particularly because the early termination factor (capping enzyme/VTF) is present at late times and is clearly accessible to late mRNAs, as evinced by the fact that late transcripts are efficiently capped (*24, 25*).

Several observations suggest that termination may be subject to regulation even at early times of infection. First, in some early genes, e.g., thymidine kinase, the efficiency of termination at the TTTTTNT element 3′ of the coding region is incomplete; in this event, polymerase reads through to a distal terminator found at the 3′ end of a downstream early gene (*26*). This results in two distinctly sized mRNAs encoding the same protein, but differing in their 3′ termini. Whether there are any consequences for differential stability or translation of such mRNAs, including possible utilization of polycistronic transcripts, is unclear. Second, although the termination signal tends to be excluded from proximal positions within early genes, several exceptions do occur in which a single TTTTTNT is present either within the coding region or within a 5′ nontranslated leader sequence (*27*). These include early genes E9L (encoding DNA polymerase (*28*)), A24R (*rpo*132, encoding the second largest subunit of RNA polymerase (*29, 30*)), J3R (encoding the small subunit of poly(A) polymerase (*31, 32*), E1L (encoding the large subunit of poly(A) polymerase (*32*)), G5R (*33*), A18R (*34*), and I8R (*35*). In the case of the I8R gene (which encodes an essential 78 kDa virion protein (*36*)), two early transcripts are detected *in vivo*. One is a "full-length" RNA (capable of encoding the entire open reading frame) that terminates in response to a TTTTTT element downstream of the translation stop codon; the other is an "attenuated" transcript whose 3′ end maps within the I8R ORF, just downstream of an internal TTTTTGT sequence (*35*). The attenuated early transcript, which can encode only the N-terminal 70% of the protein, predominates over the full-length RNA, indicating that termination at an intragenic site can be fairly efficient *in vivo* (*35*). The I8R gene is also transcribed late (from a tandem late promoter element) when, absent termination, all mRNAs would encode the full-sized polypeptide (*35*). Thus, for this gene, it is quite possible that the level (and type) of gene product at early times is regulated specifically by premature termination. Other examples of termination at TTTTTNT sequences within protein encoding regions of early transcription units have been noted using recombinant vaccinia viruses that express heterologous genes (*37*).

Do the other early genes containing proximal TTTTTNT sequences also produce attenuated transcripts? Pacha *et al.* have examined early transcripts of the A18R gene that contains a TTTTTT sequence within the coding region and failed to detect (by Northern analysis and S1 probe protection) any 3′ end that would result from termination near this sequence (*34*). How then are some early terminators "ignored" *in vivo* (e.g., in A18R) while others (I8R) are utilized? Pacha *et al.* point out that the internal TTTTTT signal in the A18R gene is found within a sequence that can be folded into a hairpin structure, as shown in Fig. 1A. This potential hairpin contains a 5 nucleotide loop and a 16-bp stem (within which the termination signal would be completely base-paired). They suggest that formation of the hairpin structure might prevent termination (*34*).

We have approached this question using a reconstituted early transcription system amenable to analysis of factor-dependent termination (*21*). A DNA oligonucleotide containing the A18R putative hairpin sequence (Fig. 1A) was inserted downstream of a synthetic early transcription unit in the plasmid pSB24 (shown in Fig. 2) to generate pYL7. Plasmid DNAs were linearized with *Nde*I and used as templates for vaccinia RNA polymerase with or without addition of VTF. The parent plasmid pSB24 directed the synthesis of a 630 nucleotide transcript resulting from initiation at the early promoter and synthesis of a runoff RNA (Fig. 3, lane 1). Because pSB24 contained no TTTTTNT motif between the promoter and the *Nde*I restriction site, no terminated RNAs were evident in reactions containing VTF. pYL7, containing the 50 nucleotide hairpin insert, programmed the synthesis of a correspondingly longer 680 nucleotide readthrough RNA (Fig. 3, lane 4). Inclusion of VTF in the reaction failed to induce the appearance of a terminated transcript (Fig. 3, lane 5). A strikingly different result was obtained using the template pYL5 that contained the "half-hairpin" insert shown in Fig. 1B. Deletion of the proximal portion of the stem in the putative hairpin restored the ability of the TTTTTT motif to promote the synthesis of terminated transcripts (with concomitant diminution of readthrough synthesis) in a VTF-dependent fashion (Fig. 3, lanes 2 and 3). The terminated transcripts were of two distinct size classes, consistent with heterogeneity in the site of 3′ end formation (*18*). The efficiency of termination by the half-hairpin construct was comparable to that of other active terminators that were tested previously in this *in vitro* transcription system (*21*). Thus, we have recapitulated *in vitro* with purified enzymes the *in vivo* observation that the internal A18R hairpin element does not act in *cis* as a factor-dependent terminator. This effect is not dependent on the homologous A18R promoter, it does not require RNA or DNA sequences other than the hairpin, and it ap-

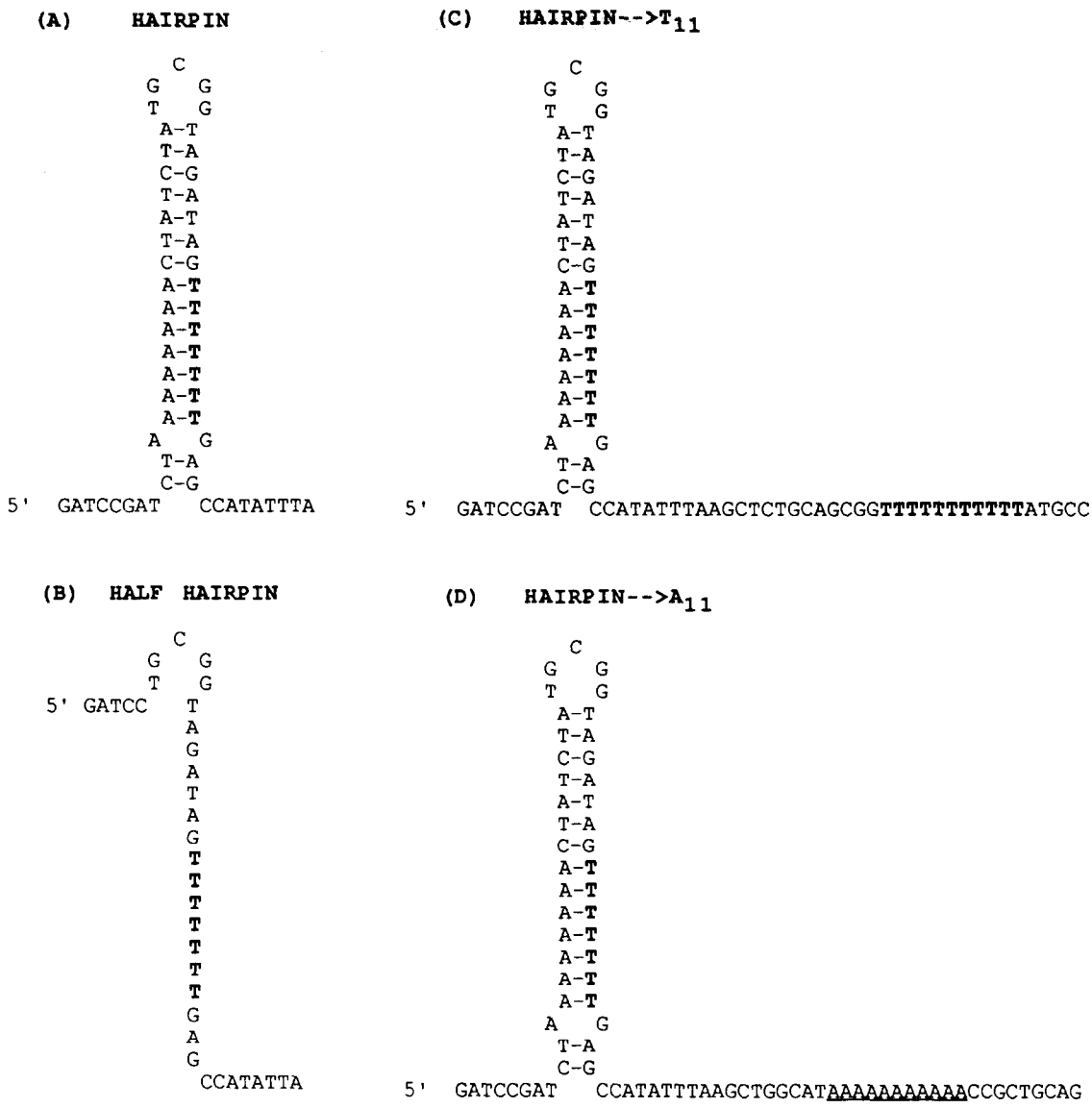

FIG. 1. Structure of A18R proximal termination signal and derivatives thereof. The nucleotide sequence of the vaccinia A18R gene in the vicinity of the unutilized proximal terminator is shown in (A). The sequence (corresponding to the nontemplate DNA strand) is displayed in the form of the putative hairpin structure suggested by Pacha *et al.* (*34*). The TTTTTT signal is displayed in bold type. Complementary synthetic DNA oligonucleotides corresponding exactly to this sequence (and containing single-strand extensions to facilitate molecular cloning) were hybridized and inserted between the *Bam*H1 and *Hin*dIII restriction sites of plasmid pSB24 (Fig. 2) to generate pYL7. The potential for secondary structure was eliminated by deletion of the entire proximal stem sequence while leaving intact the distal segment containing the terminator motif, as shown in (B). Complementary synthetic oligonucleotides corresponding exactly to the half hairpin sequence (and containing single-strand extensions) were hybridized and inserted between the *Bam*H1 and *Hin*dIII restriction sites of plasmid pSB24 to generate pYL5. The terminator elements shown in (C) and (D) were constructed by insertion into the *Hin*dIII site of pYL7 a 29-bp oligonucleotide containing a run of 11 consecutive T:A base pairs. Plasmids of both orientation were obtained: pYL11-2 containing a T_{11} tract in the nontemplate strand (C), and pYL11-3 with a A_{11} sequence (D). The structures of all the recombinant plasmids was verified by dideoxy sequencing.

pears to involve no antiterminator protein acting in *trans.*

The finding that antitermination requires the proximal stem region of the putative hairpin structure can be interpreted in two simple ways: (i) the duplex context masks the recognition of the TTTTTT signal by VTF and/or RNA polymerase, or (ii) the hairpin structure per se has an inhibitory effect on termination. The latter case is relevant to models of termination in which VTF might be obliged to translocate through the entire length of the nascent transcript in order to dissociate the ternary elongation complex. In this scenario, the hairpin may arrest VTF "scanning" based on inability of the protein to disrupt the putative RNA hairpin (e.g.,

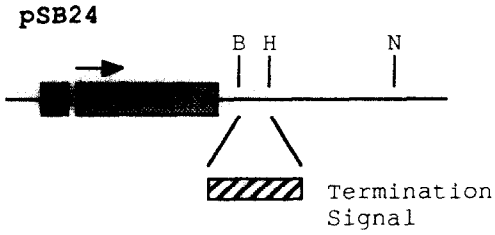

FIG. 2. Transcription template for assay of termination signal activity. Plasmid pSB24 contains a 382 nucleotide G-less cassette (shaded box) downstream of a synthetic vaccinia early promoter (black box, see Ref. 21 for promoter sequence). The direction of transcription is indicated by the arrow. DNA oligonucleotides containing TTTTTNT termination signals were inserted between restriction sites for endonucleases *Bam*H1 (B) and *Hin*dIII (H). All plasmids were linearized at the downstream *Nde*I site (N) before use in *in vitro* transcription reactions.

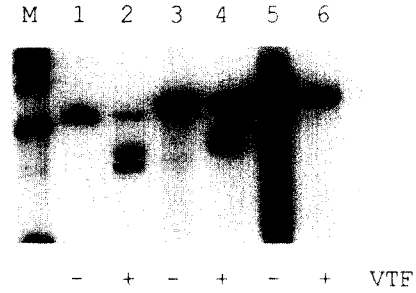

FIG. 4. The A18 hairpin does not inhibit factor-dependent termination in *cis*. Transcription reactions were performed as described in the legend to Fig. 3. *Nde*I-cut DNA templates were included as follows: lanes 1 and 2, pYL-5; lanes 3 and 4, pYL-11-2; lanes 5 and 6, pYL-11-3. An autoradiogram of the dried gel is shown. The inclusion of VTF/capping enzyme in the reactions is indicated by a + below the lanes. Lane M contains radiolabeled *Hae*III restriction fragments of ϕX 174 RF DNA.

through hypthetical helicase activity). In order to test this possibility, we inserted a new terminator (containing a T_{11} sequence in the nontemplate strand) downstream of the hairpin element to create the "hairpin-T_{11}" plasmid pYL11-2 (the insert of which is shown in Fig. 1C). The control "hairpin-A_{11}" plasmid pYL11-3 contained an oppositely oriented oligonucleotide insert with run of 11 consecutive A residues in the nontemplate strand (Fig. 1D). *In vitro* transcription analysis was performed as before using plasmid templates linearized by *Nde*I. Insertion of a T_{11} sequence distal to the hairpin was sufficient to override the antiterminator effect of the putative stem-loop structure (Fig. 4, lanes 3 and 4). The synthesis of terminated RNAs from this template was VTF-dependent; the efficiency of termination was similar to that of the half-hairpin template (Fig. 4, compare lanes 2 and 4). Factor-dependent termination was not observed in transcription reactions programmed by the hairpin-A_{11} template. The sizes of the terminated transcripts were longer for the hairpin-T_{11} construct than for the half-hairpin template because of the larger size of the DNA insert in the pYL11-2. These data indicate that a potential hairpin structure is not antagonistic per se to VTF-dependent termination. Furthermore, we can exclude a model involving antitermination due to a duplex roadblock to VTF movement; clearly, the termination factor can recognize and act upon a terminator downstream of the potential intramolecular duplex (Fig. 4). Our experiments do not address the possibility that looping of the nacent RNA might circumvent the hairpin and allow interaction with a downstream signal.

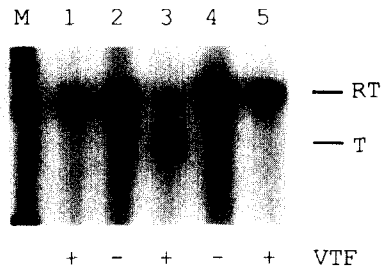

FIG. 3. Termination signal activity of A18 hairpin and half-hairpin elements. Transcription reaction mixures contained 20 m*M* Tris–HCl, pH 8.0, 2 m*M* dithiothreitol, 6 m*M* $MgCl_2$, 1 m*M* ATP, 1 m*M* CTP, 1 m*M* GTP, 100 μM [α^{32}P]UTP, RNA polymerase (1.25 units, phosphocellulose fraction (*21*)), VTF/capping enzyme (100 fmol, phosphocellulose fraction; inclusion indicated by + below the lanes) and 300 ng *Nde*I-cut DNA template as follows: lane 1, pSB24; lanes 2 and 3, pYL-5; lanes 4 and 5, pYL-7. Incubation was for 30 min at 30°. Transcription reaction products were processed as described (*21*) and analyzed by electrophoresis through 4% polyacrylamide gels under denaturing conditions. An autoradiogram of the dried gel is shown. Transcripts corresponding to read-through RNA (RT) and terminated RNA (T) species are indicated on the right. Contained in lane M are radiolabeled *Hae*III restriction fragments of ϕX 174 RF DNA.

The *in vitro* experiments suggest (i) that antitermination at the proximal site of the A18R gene may be mediated by secondary structure in the nascent RNA, and (ii) that early termination involves recognition of the UUUUUNU sequence in single-stranded form. (The existence of the antiterminating stem-loop is inferred from the sequence and remains to be demonstrated directly.) Whether other early genes containing internal terminators are regulated similarly remains unclear. Pacha *et al.* note that proximal signals within other early transcription units can also be folded into secondary structures (*34*). Should RNA analyses confirm that these terminators are also ignored *in vivo* rather than utilized, then the *in vitro* approach described here should simplify identification of common sequence or structural motifs involved in vaccinia antitermination.

In several respects, the VTF-dependent vaccinia termination system resembles ρ-dependent termination by *Escherichia coli* DNA-dependent RNA polymerase. Similarities include: (i) recognition of a *cis*-acting termination signal in the nascent RNA, (ii) 3′ end formation at heterogenous sites downstream of the signal itself, and (iii) recognition of the RNA signal in a context largely devoid of secondary structure. RNA elements involved in ρ-mediated termination (*rut* sites) are more complex and lack the strict conservation of primary sequence seen with the vaccinia motif. In general, *rut* sites contain a high proportion of C resudues and are not amenable to intramolecular base-pairing (*38*). Increasing the potential for base-pairing within *rut* sites (e.g., by base analog substitution or via hybrid formation between *rut* sequences and complementary DNA oligonucleotides) tends to abrogate termination (*39, 40*). ρ action in termination is contingent on NTP hydrolysis by an intrinsic RNA-dependent NTPase activity (*41*). The RNA cofactor requirements for NTPase reflect known characteristics of *rut* sites, e.g., the most potent effectors of NTPase are C-rich RNAs devoid of secondary structure (*42*). Does VTF-dependent termination also require NTP hydrolysis, and, if so, how might this requirement relate to earlier observations (*43*) that vaccinia capping enzyme/VTF posseses an intrinsic purine nucleoside triphosphate phosphohydrolase activity? Furthermore, does VTF, like ρ (*44*), possess RNA:DNA helicase activity? Methods to study VTF-dependent termination within a single round of transcription have been developed recently (*21*) and should be applicable to these outstanding mechanistic questions.

ACKNOWLEDGMENTS

This work was supported by Grants GM 42498 from the National Institutes of Health, and JFRA 274 from the American Cancer Society. S.S. is the recipient of a Pew Scholarship in the Biomedical Sciences.

REFERENCES

1. SPENCER, C. A., and GROUDINE, M., *Oncogene* **5,** 777–785 (1990).
2. KERPOLLA, T. K., and KANE, C. M., *FASEB J.,* in press (1991).
3. BENTLEY, D. M., and GROUDINE, M., *Cell* **53,** 245–256 (1988).
4. WRIGHT, S., and BISHOP, J. M., *Proc. Natl. Acad. Sci. USA* **86,** 505–509 (1989).
5. MECHTI, N., PIECHACZYK, M., BLANCHARD, J., JEANTEUR, P., and LEBLEU, B., *Mol. Cell. Biol.* **11,** 2832–2841 (1991).
6. BENDER, T. P., THOMPSON, C. B., and KUEHL, W. M., *Science* **237,** 1473–1476 (1987).
7. KRYSTAL, G., BIRRER, M., WAY, J., NAU, M., SAUSVILLE, E., THOMPSON, C., MINNA, J., and BATTEY, J., *Mol. Cell. Biol.* **8,** 3373–3381 (1988).
8. ROUGVIE, A. E., and LIS, J. T., *Cell* **54,** 795–804.
9. CHEN, Z., HARLESS, M. L., WRIGHT, D. A., and KELLEMS, R. E., *Mol. Cell. Biol.* **10,** 4555–4564 (1990).
10. RAMAMURTHY, V., MAA, M., HARLESS, M. L., WRIGHT, D. A., and KELLEMS, R. E., *Mol. Cell. Biol.* **10,** 1484–1491 (1990).
11. MIDDLETON, K. M., and MORGAN, G. T., *Mol. Cell. Biol.* **10,** 727–735 (1990).
12. WIEST, D. K., and HAWLEY, D. K., *Mol. Cell. Biol.* **10,** 5782–5795 (1990).
13. SKARNES, W. C., TESSIER, D. C., and ACHESON, N. H., *J. Mol. Biol.* **203,** 153–171 (1988).
14. HAY, N., SKOLNIK-DAVID, H., and ALONI, Y., *Cell* **29,** 183–193 (1982).
15. RESNEKOV, O., and ALONI, Y., *Proc. Natl. Acad. Sci. USA* **86,** 12–16 (1989).
16. KAO, S-Y., CALMAN, A. F., LUCIW, P. A., and PETERLIN, B. M., *Nature* **330,** 489–493 (1987).
17. MOSS, B., *Annu. Rev. Biochem.* **59,** 6661–688 (1990).
18. ROHRMANN, G., YUEN, L., and MOSS, B., *Cell* **46,** 1029–1035 (1986).
19. YUEN, L., and MOSS, B., *Proc. Natl. Acad. Sci. USA* **84,** 6417–6421 (1987).
20. SHUMAN, S., BROYLES, S. S., and MOSS, B. J., *Biol. Chem.* **262,** 12372–12380 (1987).
21. LUO, Y., HAGLER, J., and SHUMAN, S., *J. Biol. Chem.* **266,** 13,303–13,310 (1991).
22. SHUMAN, S., and MOSS, B., *J. Biol. Chem.* **263,** 6220–6225 (1988).
23. MAHR, A., and ROBERTS, B. E., *J. Virol.* **49,** 510–520 (1984).
24. BOONE, R. F., ENSINGER, M. J., and MOSS, B., *J. Virol.* **21,** 475–483 (1977).
25. BOONE, R. F., and MOSS, B., *Virology* **79,** 67–80 (1977).
26. BAJSZAR, G., WITTEK, R., WEIR, J. P., and MOSS, B., *J. Virol.* **45,** 62–72 (1983).
27. GOEBEL, S. J., JOHNSON, G. P., PERKUS, M. E., DAVIS, S. W., WINSLOW, J. P., and PAOLETTI, E., *Virology* **179,** 247–266 (1990).
28. EARL, P. L., JONES, E. V., and MOSS, B., *Proc. Natl. Acad. Sci. USA* **83,** 3659–36663 (1986).
29. PATEL, D. D., and PICKUP, D. J., *J. Virol.* **63,** 1076–1086 (1989).
30. AMEGADZIE, B. Y., HOLMES, M. H., COLE, N. B., JONES, E. V., EARL, P. L., and MOSS, B., *Virology* **180,** 88–98 (1991).
31. PLUCIENNICZAK, A., SCHROEDER, E., ZETTLMEISSL, G., and STREECK, R. E., *Nucleic Acids Res.* **13,** 985–998 (1985).
32. GERSHON, P. D., and MOSS, B. (personal communication).
33. MEIS, R. J., and CONDIT, R. C., *Virology* **182,** 442–454 (1991).
34. PACHA, R. F., MEIS, R. J., and CONDIT, R. C., *J. Virol.* **64,** 3853–3863 (1990).
35. FATHI, Z., and CONDIT, R. C., *Virology* **181,** 258–272 (1991).
36. FATHI, Z., and CONDIT, R. C., *Virology* **181,** 273–276 (1991).
37. EARL, P. L., HUGIN, A. W., and MOSS, B., *J. Virol.* **64,** 2448–2451 (1990).
38. FAUS, I., and RICHARDSON, J. P., *J. Mol. Biol.* **212,** 53–66 (1990).
39. ADHYA, S., SARKAR, P., VALENZUELA, D., and MAITRA, U., *Proc. Natl. Acad. Sci. USA* **76,** 1613–1617 (1979).
40. CHEN, C-Y., GALLUPPI, G. R., and RICHARDSON, J. P., *Cell* **46,** 1023–1028 (1986).
41. SHARP, J. A., and PLATT, T., *J. Biol. Chem.* **259,** 2268–2273 (1984).
42. LOWERY, C., and RICHARDSON, J. P., *J. Biol. Chem.* **252,** 1381–1385 (1977).
43. SHUMAN, S., SURKS, M., FURNEAUX, H., and HURWITZ, J., *J. Biol. Chem.* **255,** 11588–11598 (1980).
44. BRENNAN, C. A., STEINMETZ, E. J., SPEAR, P., and PLATT, T., *J. Biol. Chem.* **265,** 5440–5447 (1990).