Available online at www.sciencedirect.com
SCIENCE DIRECT®
Virology 319 (2004) 1–11
VIROLOGY
www.elsevier.com/locate/yviro

Post-transcription cleavage generates the 3′ end of F17R transcripts in vaccinia virus

Susan M. D'Costa,[a] James B. Antczak,[b] David J. Pickup,[b] and Richard C. Condit[a,*]

[a] *Department of Molecular Genetics and Microbiology, University of Florida, Gainesville, FL 32610-0266, USA*
[b] *Department of Molecular Genetics and Microbiology, Duke University Medical Center, Durham, NC 27710, USA*

Received 28 July 2003; returned to author for revision 16 September 2003; accepted 25 September 2003

Abstract

Most vaccinia virus intermediate and late mRNAs possess 3′ ends that are extremely heterogeneous in sequence. However, late mRNAs encoding the cowpox A-type inclusion protein (ATI), the second largest subunit of the RNA polymerase, and the late telomeric transcripts possess homogeneous 3′ ends. In the case of the ATI mRNA, it has been shown that the homogeneous 3′ end is generated by a post-transcriptional endoribonucleolytic cleavage event. We have determined that the F17R gene also produces homogeneous transcripts generated by a post-transcriptional cleavage event. Mapping of in vivo mRNA shows that the major 3′ end of the F17R transcript maps 1262 nt downstream of the F17R translational start site. In vitro transcripts spanning the in vivo 3′ end are cleaved in an in vitro reaction using extracts from virus infected cells, and the site of cleavage is the same both in vivo and in vitro. Cleavage is not observed using extract from cells infected in the presence of hydroxyurea; therefore, the cleavage factor is either virus-coded or virus-induced during the post-replicative phase of virus replication. The *cis*-acting sequence responsible for cleavage is orientation specific and the factor responsible for cleavage activity has biochemical properties similar to the factor required for cleavage of ATI transcripts. Partially purified cleavage factor generates cleavage products of expected size when either the ATI or F17R substrates are used in vitro, strongly suggesting that cleavage of both transcripts is mediated by the same factor.

Keywords: Vaccinia; Post-transcription site-specific cleavage; Endoribonucleolytic cleavage; F17R

Introduction

Vaccinia virus, the prototypical orthopoxvirus, contains a linear double-stranded genome of 192-kb that encodes approximately 200 genes (Moss, 2001). Poxviruses are unique among DNA viruses in that their entire replication cycle occurs in the cytoplasm of the infected cell. Because of this cytoplasmic site of infection, vaccinia encodes the majority of the enzymes required for viral DNA replication and transcription. Poxvirus transcription is temporally regulated. Early genes are transcribed before DNA replication, whereas intermediate and late genes are transcribed after DNA replication. At each stage, transcription initiation requires *trans*-acting factors that are synthesized in the previous stage. Post-initiation events, viz., elongation and termination of all three transcriptional classes, are also regulated. Thus, vaccinia serves as a good model to study the fundamental aspects of transcription.

Regulation of pre- and post-replicative RNA 3′ end formation occurs by different mechanisms. At early times during infection before DNA replication, viral transcripts are relatively homogeneous in sequence both at their 5′and 3′ ends. This homogeneity at the 3′ end of early transcripts results from a sequence-specific transcription termination event controlled by a highly specific *cis*-acting sequence 5′ UUUUUNU3′ in the nascent RNA (Rohrmann et al., 1986; Shuman and Moss, 1988; Yuen and Moss, 1987) and several early termination factors (Christen et al., 1998; Deng and Shuman, 1998; Mohamed and Niles, 2000; Shuman et al., 1987). Upon transcription of this signal, the polymerase terminates 30–50 nt downstream. The known virus-coded termination factors include the heterodimeric capping enzyme (VTF) (Shuman et al., 1987), NPH-I (Christen et al.,

* Corresponding author. Department of Molecular Genetics and Microbiology, University of Florida School of Medicine, PO Box 100266, Gainesville, FL 32610-0266. Fax: +1-352-392-3133.
E-mail address: condit@mgm.ufl.edu (R.C. Condit).

0042-6822/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.virol.2003.09.041

1998), and RAP94 (Mohamed and Niles, 2000). By contrast, at late times post-infection following the onset of DNA replication, both intermediate and late transcripts are extremely heterogeneous at their 3′ ends (Mahr and Roberts, 1984). The *cis*-acting signal that is operative at early times is not active even when present in late transcripts, and to date, no *cis*-acting sequence has been identified that might be involved in signaling the termination of intermediate and late transcripts. If such a signal does exist, it must be highly degenerate or very simple, and inefficient. However, it has been shown both genetically and biochemically that the products of vaccinia genes A18R, G2R, and J3R control 3′ end formation by influencing intermediate and late transcription elongation and termination (Black and Condit, 1996; Lackner and Condit, 2000; Latner et al., 2000; Xiang et al., 1998, 2000). The 56-kDa A18 protein is a transcript release factor required for dissociation of nascent RNA from a ternary transcription complex in vitro (Lackner and Condit, 2000). In addition, viruses with mutations in A18R produce abnormally long post-replicative transcripts in vivo (Xiang et al., 1998). In contrast to the effect of A18, the 26-kDa G2 and 39-kDa J3 proteins have been shown to be positive elongation factors of post-replicative transcription (Black and Condit, 1996; Condit et al., 1996b; Latner et al., 2000, 2002; Xiang et al., 2000). Mutants in either of these proteins produce post-replicative transcripts that are shorter than normal at their 3′ ends, resulting in reduced quantities of large proteins at late times post infection (Black and Condit, 1996; Xiang et al., 2000). In summary, although intermediate and late transcripts are heterogeneous at their 3′ ends, suggesting a lack of sequence-specific termination, 3′ end formation of post-replicative transcripts is nevertheless clearly influenced by viral-encoded factors.

Not all late transcripts are heterogeneous at their 3′ ends. Pickup et al. have identified at least three exceptions: transcripts from the gene encoding the major protein component of the A-type inclusion (ATI) bodies in cowpox and vaccinia (Amegadzie et al., 1992; Patel and Pickup, 1987; Patel et al., 1986, 1988), telomeric transcripts from cowpox, raccoonpox, and vaccinia virus (Parsons and Pickup, 1990), and transcripts from the gene encoding the second largest subunit of RNA polymerase (rpo132) in vaccinia virus (Patel and Pickup, 1989; Patel et al., 1988). In the case of the ATI transcripts, 3′ end homogeneity is a result of post-transcriptional site-specific endoribonucleolytic cleavage of the nascent RNA (Antczak et al., 1992). The 3′ ends of the *ati* transcripts map within a 345-bp *Alu*I–*Xba*I fragment (AX element) contained within the coding region of the convergently transcribed *rpo132* gene (Patel and Pickup, 1987). Further analysis of homogeneous transcripts of the *ati* gene revealed the presence of a 43-nt *cis*-acting element that directs the cleavage reaction (Howard et al., 1999). Interestingly, most of the *cis*-acting element is downstream of the cleavage site. The *cis*-acting sequence is composed of at least two noncontiguous regions with a spacer region of 10 bases whose composition may be changed without interfering with RNA 3′ end formation. Portions of the two noncontiguous regions are complementary to each other, suggesting the possibility that the two regions form a specific structure that participates in 3′ end formation. The factor responsible for cleavage of nascent ATI transcripts appears to be virus-induced or coded at late times post

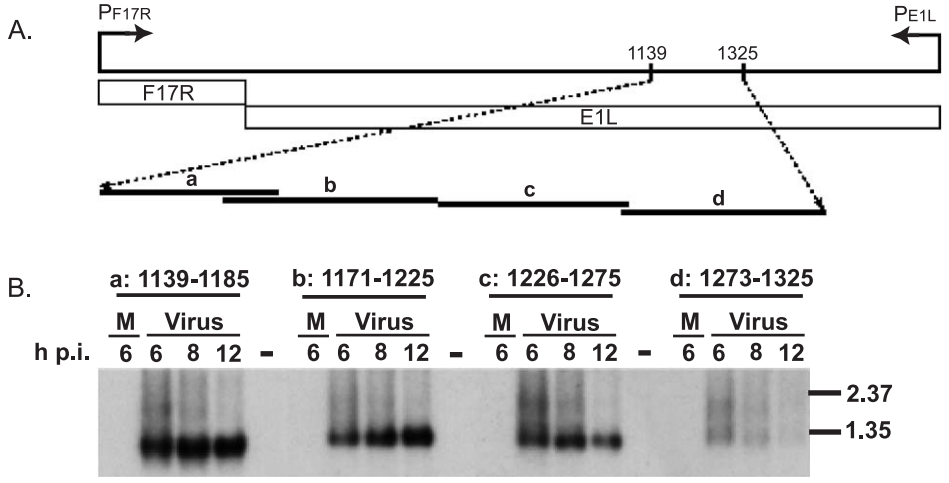

Fig. 1. Crude mapping of the 3′ end of the homogeneous F17R transcripts. (A) A diagram showing the map positions of oligonucleotides used to determine the 3′ end of F17R transcripts. The solid horizontal line represents the F17R and the E1L genes in their entirety. Arrows on either side of the line represent the F17R promoter (P_{F17}) and the E1L promoter (P_{E1L}) and the orientation of these genes on the genome. The F17R and E1L ORFs are denoted as open boxes. The numbers represent the sequence position in nucleotides downstream of the F17 translation start site. The region between 1139 and 1325 nt has been expanded to show positions of four tandem oligonucleotides, a, b, c, and d, used to probe Northern blots containing RNA from mock-infected or virus-infected BSC40 cells. (B) Northern analyses of total viral RNA using tandem oligonucleotide probes. RNA was extracted from mock-infected and wt virus-infected cells at times shown, transferred to nylon membranes, and hybridized with ^{32}P labeled oligonucleotide probes a, b, c, or d. An autoradiogram is shown. The sequence positions of the four probes are indicated above the autoradiogram. M = RNA from mock-infected cells, Virus = RNA from virus-infected cells. Times after infection are shown above each lane. Molecular weight standards are indicated on the right of the autoradiogram in kilobase.

infection. The mechanism of 3′ end formation for the other exceptions is not known at this time.

We have observed that yet another late gene, F17R, yields homogeneous transcripts (1.4 kb) late during the virus infection cycle. The object of this study was to identify the 3′ ends of these transcripts and to characterize the mechanism of 3′ end formation in F17R transcripts. We demonstrate that F17R transcripts have a well-defined 3′ end 1262 nt downstream of the F17R translation start site. In addition, these ends are formed by post-transcriptional endoribonucleolytic cleavage of the nascent message. We have also characterized and optimized the conditions of F17R cleavage. Our results suggest that the same factor is probably responsible for cleavage of both the ATI and F17R transcripts.

Results

Crude mapping of the 3′ end of F17 transcripts

The vaccinia virus F17R gene is a late gene which encodes an 11-kDa phosphoprotein present in the virion core (Wittek et al., 1984). Homogeneous transcripts (1.4 kb) of the F17 gene have been previously observed in our laboratory in Northern blots using oligonucleotide probes that span the downstream region of the F17R transcript (Xiang et al., 2000). In these experiments, we localized the 3′ end of the wt F17R transcript to a region between 1160 and 1510 nt downstream of the F17 translational start site. To obtain a more refined map position for the 3′ end of F17R transcripts, Northern blot analysis was carried out using antisense DNA oligonucleotides (approximately 50 mers) that span the entire region between 1139 and 1536 nt downstream of the F17R translation start site (Fig. 1). Homogeneous F17R transcripts were clearly detected with probes a–c. A weak, largely heterogeneous signal was detected with probe d mainly at 6 h p.i. (Fig. 1). No signal was detected with other probes spanning the regions downstream of d (data not shown). Therefore, we concluded that the 3′ ends of the F17R transcript map between 1226 and 1275 nt downstream of the F17R translation start site. We suspect that the faint signal detected with probe d represents a background of heterogeneous F17R transcripts similar to those produced by most late genes.

Nuclease S1 mapping to determine the exact 3′ end of F17R transcripts

Nuclease S1 mapping was used to precisely map the F17R 3′ end within 1226–1275 nt downstream of the F17R translation start site. Total RNA was extracted at late times postinfection from virus-infected cells and hybridized to a 3′-end-labeled DNA oligonucleotide. The RNA–DNA hybrids were then treated with nuclease S1, denatured, and fractionated on a polyacrylamide-urea gel (Fig. 2). Lanes 1–4 correspond to F17R sequence (G, A, T, C) in this region of the viral genome. Lane 5 contains the S1-treated products showing about nine products of varying intensity. This apparent 3′ end heterogeneity was observed with varying S1 nuclease concentrations and hybridization temperatures, strongly suggesting that there are in fact multiple 3′ ends within this region. An identical result was obtained using poly(A) containing infected cell RNA (not shown). The band with the highest intensity was 1262 nt downstream of the F17R translation start site and was therefore designated as the major 3′ end. Larger protected products as well as undigested probe were also observed in

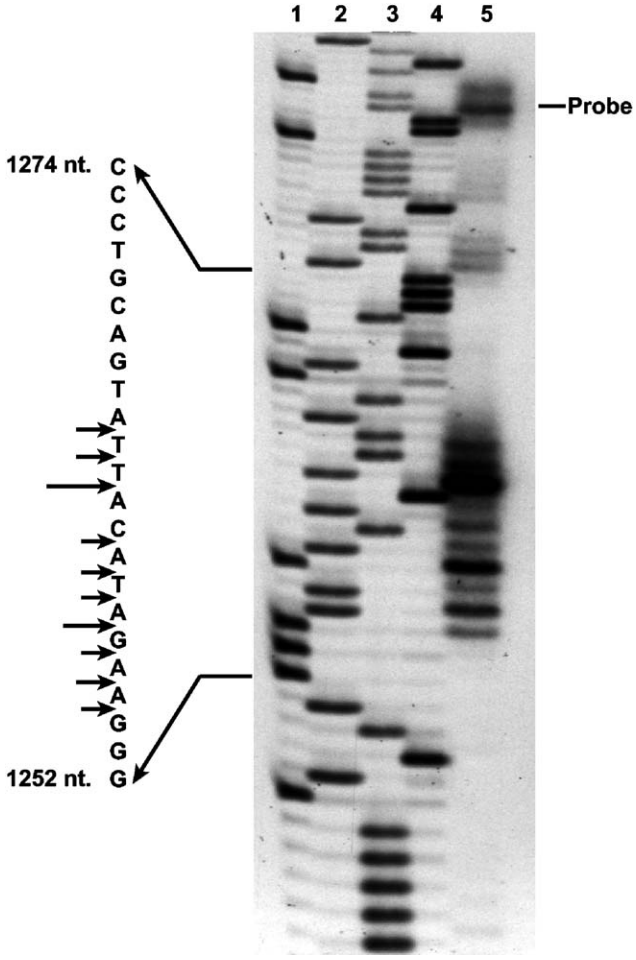

Fig. 2. Nuclease S1 protection analysis of the 3′ ends of F17R RNAs. Total RNA harvested from virus-infected cells at 12 h p.i. was incubated with a ^{32}P-3′-end-labeled DNA oligonucleotide containing 1208–1287 nt downstream of the F17R translation start site. The RNA–DNA hybrids were subjected to S1 nuclease digestion, denatured, and the nuclease-resistant products were resolved on a 10% polyacrylamide-urea gel. An autoradiogram is shown. Lanes 1–4 correspond to the F17R sequence (G, A, T, C) in this region. The 5′ end of the sequencing primer corresponds exactly with the 3' end of the labeled probe used in the S1 analysis. Lane 5 corresponds to the S1 products. The full-length undigested probe is indicated. A portion of the sequence containing the S1 products has been expanded on the left with arrows indicating the position of the S1 products. The length of the arrowheads corresponds to the intensity of the S1 products.

lane 5. Consistent with the background observed with probe d in Fig. 1, we suspect that these signals correspond to a background of heterogeneous F17R transcripts similar to those produced by most late genes.

The 3′ ends of F17R transcripts are generated by site-specific cleavage

The generation of 3′ ends in F17R transcripts could be as a result of transcription termination or post-transcriptional endoribonucleolytic cleavage of nascent transcripts. The 3′ end of the ATI transcript is formed by post-transcriptional endoribonucleolytic cleavage (Antczak et al., 1992). To determine if the same mechanism was responsible for the homogeneity of F17R transcripts, we used an adaptation of the in vitro cleavage assay reported earlier (Antczak et al., 1992) and described in detail in Materials and methods (Fig. 3). RNA containing the F17R in vivo 3′ end was synthesized in vitro as a 430-nt transcription product from clone pGF17stp DNA using SP6 polymerase (Fig. 3A; Substrate 1). The 430-nt transcript was incubated for various times at 30 °C with extracts from cells infected or mock-infected in the presence or absence of hydroxyurea, a known DNA synthesis inhibitor (Fig. 3B). When the 430-nt transcript was treated with an extract from cells infected in the absence of drug, two RNA products of approximately 190 and 240 nt were observed. Similar products were not evident with extracts from cells infected in the presence of hydroxyurea, mock-infected, and in the absence of any extract. These results show that the 3′ ends of F17R transcripts are produced by post-transcriptional cleavage of nascent RNA, and that the factor responsible for this activity is either virus-coded or virus-induced late during infection. Cleavage of ATI transcripts was only observed with extracts from cells infected in the absence of AraC, suggesting that the factor responsible for post-transcriptional cleavage of nascent ATI transcripts is also virus-induced or virus-coded late during infection (data not shown). Extracts from purified virus particles did not cleave either the F17R or the ATI transcripts (data not shown).

Products of 190 and 240 nt could result from a cleavage event of Substrate 1 at one of two sites referred to here as an upstream or downstream cleavage event. An upstream cleavage event would coincide with the in vivo cleavage site. A downstream cleavage event would coincide with cleavage 50 nt downstream from the in vivo cleavage site. To determine if the 3′ end observed in vitro was the same as that observed in vivo, we used a second RNA, 410 nt long, transcribed in vitro using T7 polymerase from clone pGF17sense609 DNA digested with *Acc*I (Substrate 2). If

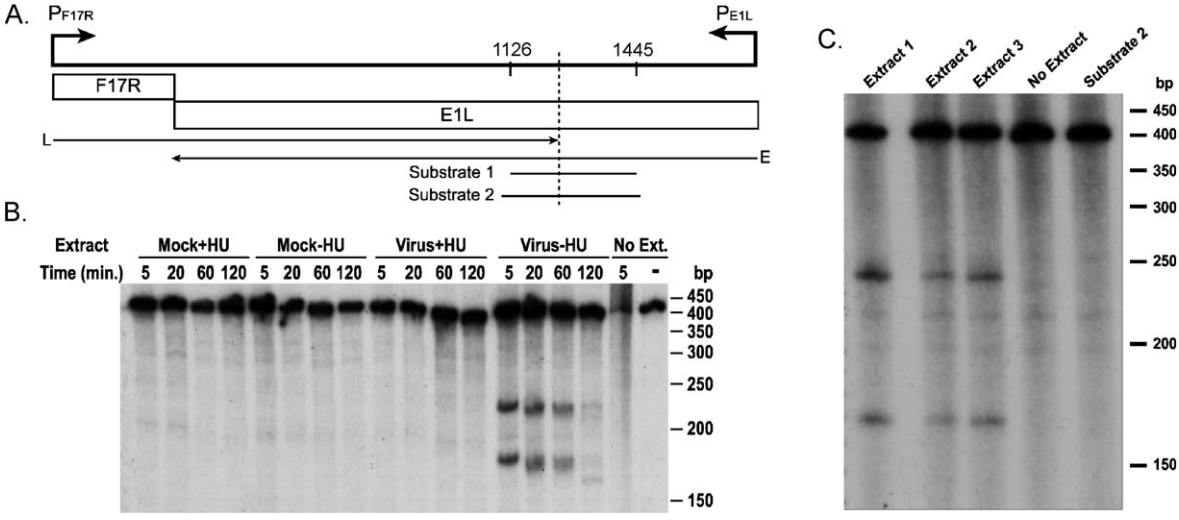

Fig. 3. 3′ ends in F17R transcripts are generated by post-transcriptional endoribonucleolytic cleavage. (A) A diagram showing the map positions of Substrates 1 and 2 used in the in vitro cleavage assay. The solid horizontal line represents the F17R and the E1L genes in their entirety. Arrows on either side of the line represent the F17R promoter (P_{F17}) and the E1L promoter (P_{E1L}) and the orientation of these genes on the genome. The F17R and E1L ORFs are denoted as open boxes. The numbers represent the sequence position in nucleotides downstream of the F17R translation start site. The orientation of the F17R and E1L transcripts in vivo are shown as solid arrows. L = late F17R transcript, E = early E1L transcript. The vertical dotted line indicates the in vivo cleavage site for the F17L transcript. Substrates 1 and 2 indicate the in vitro transcripts used in the cleavage assay. The vaccinia sequences present in Substrates 1 and 2 are denoted by solid horizontal lines. Substrate 1 also contains 49 nt of vector sequence at the 5′ end and Substrate 2 contains 9 nt of vector sequence at the 5′ end not shown in this cartoon. (B) Autoradiograph of an in vitro cleavage assay using the sense transcript of Substrate 1. ^{32}P-labeled sense transcript of Substrate 1 (430 nt) was incubated at 30 °C with extracts from cells infected (Virus + HU, Virus − HU) or mock-infected (Mock + HU, Mock − HU) in the presence or absence of hydroxyurea (HU) for various times. Times of incubation in minutes are shown above each lane. A corresponding amount of Substrate 1 was also incubated in the absence of extract (No Extract) or loaded directly on the gel (−). RNA was purified from the reactions and fractionated on a 6% polyacrylamide-urea gel. Molecular weight markers are shown on the right in base pairs. (C) Autoradiograph of an in vitro cleavage assay using the sense transcript of Substrate 2. ^{32}P-labeled sense transcript of Substrate 2 (410 nt) was incubated at 30 °C for 60 min with three independently prepared infected-cell extracts (Extract 1, Extract 2, and Extract 3), in the absence of extract (No Extract), or loaded directly on the gel (Substrate 2). RNA was purified from these samples and fractionated on a 6% polyacrylamide-urea gel. Molecular weight markers are shown on the right in base pairs.

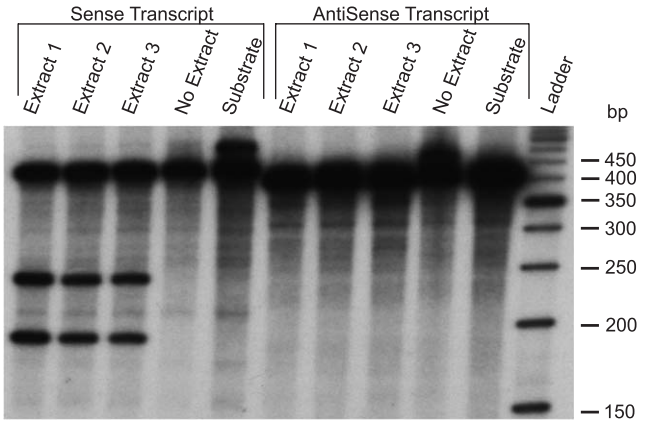

Fig. 4. The *cis*-acting sequence required for 3′ end formation in F17R is strand-specific. ^{32}P-labeled sense (430 nt) or antisense (402 nt) transcripts of Substrate 1 were incubated at 30 °C for 60 min with three independently prepared infected-cell extracts (Extract 1, Extract 2, and Extract 3), in the absence of extract (No Extract), or loaded directly on the gel (Substrate). RNA was recovered from the samples and fractionated on a 6% polyacrylamide-urea gel. Molecular weight markers are indicated on the right in base pairs.

the cleavage site was the same as in vivo, the expected products would be 164 and 246 nt long. If, however, a downstream cleavage site was used, the expected products would be 200 and 210 nt long. The transcript of Substrate 2 was incubated with three independent preparations of infected-cell extracts. Fractionation of this RNA on polyacrylamide gels revealed the presence of two RNA products approximately 160 and 250 nt long in addition to the full-length substrate (Fig. 3C). These results confirm that the 3′ ends of F17R transcripts are synthesized as a result of post-transcriptional cleavage of the nascent RNA and that the site of cleavage is the same both in vivo and in vitro.

The cis-acting sequence required for site-specific cleavage is strand specific

The generation of 3′ ends in F17R transcripts must require a *cis*-acting recognition signal for site-specific cleavage. A 43-nt *cis*-acting element has been characterized at the 3′ end of ATI transcripts to be responsible for site-specific cleavage (Howard et al., 1999). Most of this sequence was observed to be downstream of the cleavage site. The reverse complement of this sequence was inactive in cleavage. To determine if the *cis*-acting sequence in F17R transcripts has similar characteristics as the ATI *cis*-acting sequence, we incubated either sense or antisense transcripts from Substrate 1 with three independent preparations of infected-cell extracts (Fig. 4). Expected cleavage products of 190 and 240

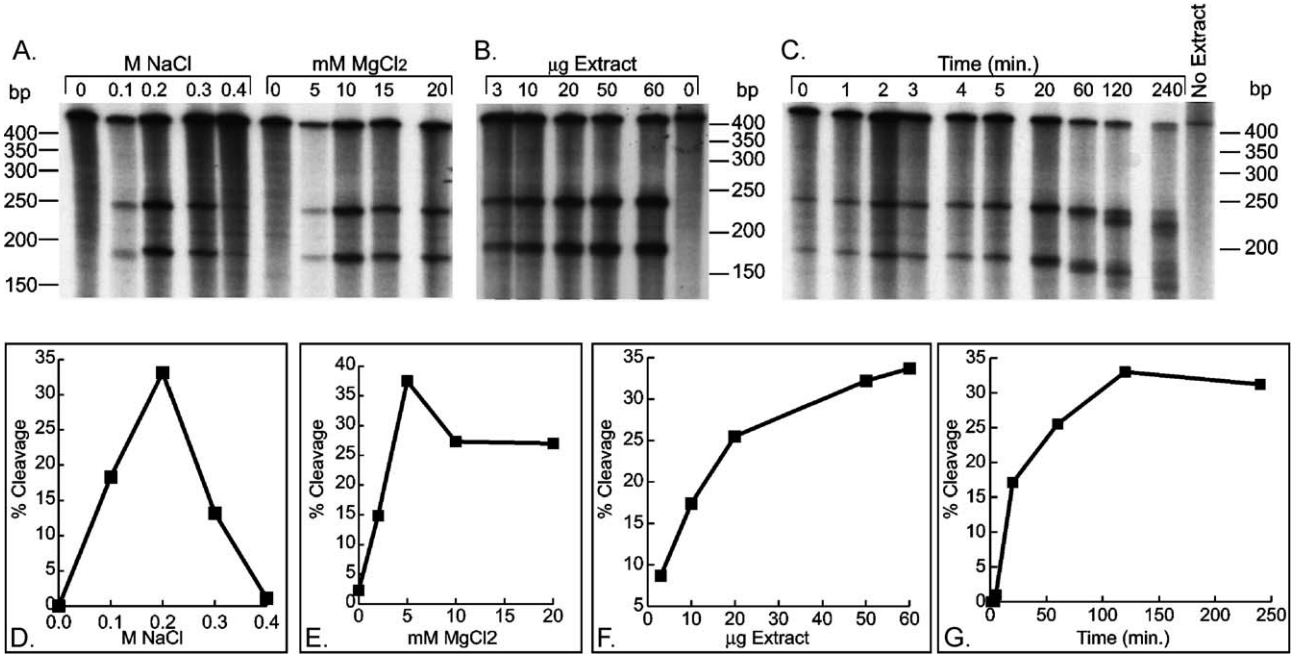

Fig. 5. Biochemical characterization of cleavage activity. (A) Effect of salt and Mg^{2+} concentration on cleavage activity. The ^{32}P-labeled sense transcript (430 nt) of Substrate 1 was incubated with 20 µg of infected-cell extract at 30 °C for 60 min in the absence or with increasing concentrations of NaCl or $MgCl_2$. RNA was recovered from these samples and fractionated on a 6% polyacrylamide-urea gel. An autoradiograph is shown. The molar concentrations (M NaCl, mM $MgCl_2$) are indicated above each lane. Molecular weights are indicated to the left in base pairs. (B) Autoradiograph showing the effect of extract concentration on cleavage activity. The ^{32}P-labeled sense transcript of Substrate 1 was incubated at 30 °C for 60 min with increasing extract concentrations or in the absence of extract (0). Extract concentration is indicated above each lane in µg. Molecular weight markers are indicated on the right in base pairs. (C) Time course of cleavage activity. The ^{32}P-labeled sense transcript of Substrate 1 was incubated at 30 °C with 20 µg of infected-cell extract for various times or in the absence of extract (No Extract). Time in minutes is indicated above each lane. Molecular weight markers are indicated to the right in base pairs. Graphs D–G are computed with data from A–C. Percentage of cleavage over background is plotted versus M NaCl (D), mM $MgCl_2$ (E), µg extract (F), or time in minutes (G).

nt were obtained using the 430-nt sense transcript of Substrate 1. The expected cleavage products with the 402-nt antisense transcript would be 152 and 250 nt if cleavage were not strand-specific. However, no cleavage products were detected when the antisense transcript was used. These results show that the *cis*-acting sequence required for cleavage is indeed strand specific.

Biochemical characterization of the cleavage activity

The data described above suggest that the factor(s) responsible for post-transcriptional cleavage of nascent F17R transcripts are virus-coded or virus-induced at late times during infection. To determine the biochemical characteristics of this activity, the effect of NaCl, $MgCl_2$, and extract concentration was titrated, and a time course was performed (Figs. 5A–C). Graphs of percentage of cleavage over background were plotted using these data (Figs. 5D–G). An obvious requirement of NaCl for cleavage activity was noted, with an optimal concentration of 200 mM (33% cleavage activity, Figs. 5A and D). The optimal salt concentration using any given extract preparation was slightly different and ranged between 0.2 and 0.3 M NaCl (data not shown). Higher concentrations of salt inhibited cleavage activity. Other salts (KCl, NH_4OAc, NH_4Cl) were also tested to determine the ideal salt for maximal activity (not shown). While cleavage activity was detected in the presence of all these salts, the activity was maximal using NaCl. The cleavage reaction also required Mg^{2+}, with an optimal concentration between 5 and 10 mM (37% cleavage activity, Figs. 5A and E). Cleavage activity increased linearly with increasing concentrations of extract up to 20 μg (26% cleavage, Figs. 5B and F). Cleavage activity also increased with time up to 120 min (33% cleavage, Figs. 5C and G). Cleavage products were observed as early as 1 min. At 120 and 240 min, two additional cleavage products were observed which were slightly smaller compared to each of the original two products. We suggest these additional products are the result of a nonspecific nuclease activity. Cleavage activity is not stimulated by added ATP and not inhibited by added AMP-PNP and therefore is not ATP dependent (not shown).

The factor responsible for cleavage of the F17R and ATI transcripts may be the same

The biochemical properties (effect of NaCl concentration, Mg^{2+} dependency, and the lack of requirement for ATP) of the factor cleaving the F17R transcripts are the same as those of the factor shown to cleave the ATI RNAs (data not shown). To determine if the factor responsible for endoribonucleolytic cleavage of F17R and ATI transcripts may be the same, we incubated either the ATI transcript (Antczak et al., 1992) or the F17R transcript from Substrate 1 with fractions from three steps of a purification protocol currently under development in our laboratory (Fig. 6). The fractions tested include a crude extract prepared from infected HeLa cells (Extract), a 400 mM NaCl DEAE

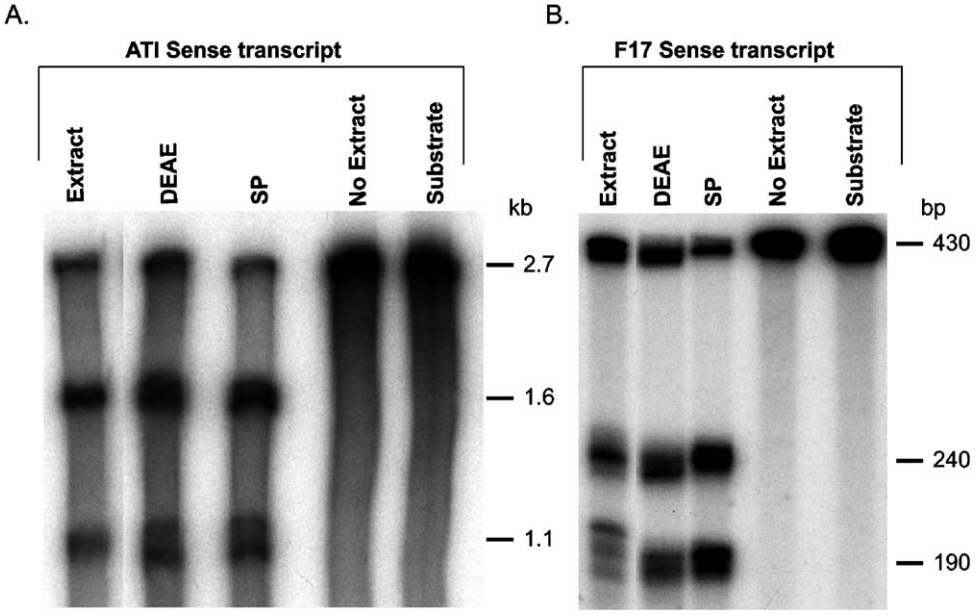

Fig. 6. The factor responsible for cleavage of F17R and ATI transcripts may be the same. (A) Autoradiograph of an in vitro cleavage assay using the ATI sense transcript. ^{32}P-labeled sense transcript of p2098 (2.7 kb) was incubated at 30 °C for 60 min with crude extract prepared from wt-infected HeLa S3 cells (Extract), DEAE flowthrough fraction (DEAE), an eluate from an SP sepharose column (SP), in the absence of extract (No Extract), or loaded directly on the gel (Substrate). RNA was purified from these samples and fractionated on a 6% polyacrylamide-urea gel. Molecular weights of each individual band are denoted to the right of the autoradiograph in kilobase. (B) Autoradiograph of an in vitro cleavage assay using the F17 sense transcript from Substrate 1. The reaction conditions and the extracts or fractions used to assay cleavage are the same as in A. Molecular weights of each individual band are denoted to the right of the autoradiograph in base pairs.

cellulose flowthrough fraction (DEAE), and pooled active fractions from an SP sepharose column eluting between 300 and 750 mM NaCl (SP). The expected cleavage products of 1.6 and 1.1 kb for the 2.7-kb ATI transcript (Antczak et al., 1992) were observed with the crude infected HeLa cell extract as well as with partially purified preparations of this extract (Fig. 6A). These same fractions also cleaved the F17 sense transcript (Fig. 6B). No cleavage products were observed for either substrate in the absence of any extract. This result strongly suggests that the factor responsible for cleavage of F17 transcripts is possibly the same as that for ATI transcripts.

Discussion

F17R provides yet another example of a late vaccinia gene that produces homogeneous transcripts. Our results demonstrate that F17R transcripts have a defined 3′ end, 1262 nt downstream of the F17R translational start site, generated by site-specific endoribonucleolytic cleavage of nascent mRNA. Homogeneous (1.4 kb) F17R transcripts were observed in Northern analyses using poly(A) RNA (not shown) as well as with total RNA, and S1 nuclease protection assays with poly(A) RNA (not shown) showed results similar to those observed with total RNA, demonstrating polyadenylation of the cleaved F17R RNA 3′ end. The cleavage factor is either virus-induced or virus-coded at late times during infection and requires divalent cation and salt for activity. The *cis*-acting sequence responsible for cleavage is strand-specific and the factor responsible for cleavage activity has similar biochemical properties as the factor required for cleavage of ATI transcripts. Partially purified cleavage factor generated cleavage products of expected size when either the ATI or F17R substrates were used in vitro, strongly suggesting that cleavage of both transcripts is mediated by the same factor.

The *cis*-acting sequence identified in ATI transcripts is 43 nt long and composed of at least two noncontiguous regions separated by a spacer of 10 bases (Howard et al., 1999). The sequence composition of the spacer can be changed without affecting 3′ end formation. Most of the *cis*-acting sequence is downstream from the cleavage site. Portions of the noncontiguous regions of the ATI *cis*-acting sequence are complementary to each other, suggesting the formation of a secondary structure that is necessary for cleavage. Using the RNA fold program (Zuker et al., 1991) and partial ribonuclease digestion of a 260-base RNA transcript within the AX element containing the 3′ end of ATI, Howard et al. (1999) identified a possible secondary structure for the ATI substrate. However, the computer-assisted folding was not completely supported by the partial RNase digestion, suggesting that the two sub-elements are not directly associated with each other but perhaps require the presence of other factors involved in RNA 3′ end formation. A comparison of the 43 nt *cis*-acting sequence identified in ATI with sequence surrounding the cleavage site in F17R did not show any obvious sequence homology. Furthermore, computer-assisted folding of the 3′ end of F17R does not yield secondary structures for the cleavage sequence that are obviously homologous to secondary structures proposed for the ATI cleavage sequence. We assume that there must exist some sequence or structure in common between the F17R and ATI transcripts which directs the site-specific RNA cleavage; however, our inability so far to identify any homology between the F17R and ATI cleavage sites suggests that these common features are extremely subtle.

It is unclear whether the cleavage reaction is confined to just the ATI and F17R transcripts, or to the few late transcripts that possess homogeneous 3′ ends, or whether it might be more generally involved in post-replicative RNA 3′ end formation. The majority of post-replicative vaccinia transcripts are extremely 3′ end heterogeneous in sequence (Mahr and Roberts, 1984). However, at least four late poxvirus genes produce homogeneous transcripts, including the ATI transcripts (Amegadzie et al., 1992; Patel and Pickup, 1987; Patel et al., 1986, 1988), the telomeric transcripts (Parsons and Pickup, 1990), transcripts from the *rpo132* gene (Patel and Pickup, 1989; Patel et al., 1988), and transcripts from the F17R gene. Clearly, the exceptional homogeneity found for late transcripts of ATI (Antczak et al., 1992) and F17R is attributable to post-transcriptional site-specific cleavage of the nascent RNA. It seems likely that the cleavage factor described here is also responsible for the RNA 3′ end maturation of the other characterized homogeneous transcripts. While the heterogeneity of most other late vaccinia transcripts suggests a different mechanism of 3′ end formation compared to F17R and ATI transcripts, it is also possible that post-transcriptional endoribonucleolytic cleavage of nascent mRNAs represents the mechanism by which all late mRNAs are produced. Specifically, the *cis*-acting signal responsible for cleavage may be highly degenerate and inefficient resulting in heterogeneous 3′ ends for most transcripts. It also remains to be determined whether termination of nascent RNA from the ternary transcription complex, site-specific cleavage of the nascent transcript, and polyadenylation are linked as is the case in eukaryotic systems.

A precedent exists for a requirement of endoribonucleolytic cleavage in the formation of Pol II transcripts in eukaryotes. In eukaryotes, two distinct cleavage events are required for the formation of mature mRNA (Dye and Proudfoot, 2001; reviewed in Proudfoot et al., 2002). An initial cleavage event, termed co-transcriptional cleavage (CoTC), occurs at the region of Pol II termination downstream of the poly(A)-recognition signal. CoTC is followed by a second cleavage event [termed here "poly(A) cleavage" for convenience] at the poly(A) site, followed by polyadenylation and ultimately resulting in formation of mature mRNA and release of the polymerase from the DNA template. The initial CoTC event requires that the polymerase traverse past the poly(A) signal. The polymerase

then slows down or pauses in an AT-rich region downstream of the poly(A) site and a primary CoTC event occurs at this AT-rich region. The efficiency of the AT-rich elements appears to differ in different genes; for example, a more diffuse AT-rich signal exists in human ε-globin transcripts as compared to human β-globin transcripts (Dye and Proudfoot, 2001). Following CoTC, the 5′ proximal cleavage product remains associated with the C-terminal domain (CTD) of Pol II. The RNA polymerase remains associated with the transcriptional complex and possibly traverses further downstream with the 5′ proximal cleavage product still associated with the CTD. Formation of the mature 3′ end of eukaryotic mRNA, release of the transcript from the polymerase, and release of the polymerase from the transcription complex involves a second cleavage event, the poly(A) cleavage. Two *cis*-acting sequences on the primary transcript and many *trans*-acting proteins are required for poly(A) cleavage. The *cis*-acting sequences are (1) the poly(A)-recognition signal, AAUAAA, 20–30 nt upstream of the poly(A) cleavage site and (2) a GU-rich region immediately downstream of the poly(A) cleavage site. These sequences are recognized by two *trans*-acting factors: (1) CPSF, the cleavage and polyadenylation specificity factor and (2) CstF, the cleavage stimulation factor. These factors then associate with two additional cleavage factors, CF1 and CF2, and cleave the nascent RNA. The 5′ proximal cleaved product is immediately polyadenylated by the poly(A) polymerase, PAP. In addition to these processing factors, the phosphorylated CTD of Pol II is also required for 3′ end processing (reviewed in Proudfoot, 2000; Proudfoot and O'Sullivan, 2002; Proudfoot et al., 2002). Interaction of the 5′ proximal cleavage product with the CTD and 3′ end processing of this transcript at the poly(A) cleavage site may have a role in mediating release of the polymerase, the final step of termination. In summary, although Pol II transcription termination determines the 3′ end of the primary transcript, the 3′ end of the mature transcript is formed by RNA processing at the poly(A) cleavage site upstream of the termination site. Both CoTC and poly(A) cleavage are ultimately required for transcription termination.

Whether 3′ end formation in post-replicative vaccinia transcripts involves a similar elaborate and complex array of protein–protein and protein–RNA interactions is yet to be determined. It is tempting to postulate that a reaction similar to the eukaryotic CoTC also occurs in all vaccinia late genes. The heterogeneity of most late vaccinia transcripts could then be attributed to the presence of diffuse cleavage elements in the termination region. However, it seems unlikely that polyadenylation and 3′ end formation in vaccinia late transcripts are linked, given that the sequence requirements for the vaccinia poly(A) polymerase to polyadenylate 3′ ends of mRNAs are more degenerate and less specific (Gershon, 1998) than its eukaryotic counterpart. In addition, homologs of CPSF, and CstF proteins have not been identified in vaccinia. Ultimately, determination of the role of the vaccinia ATI/F17R cleavage factor in viral mRNA metabolism will require identification of the gene(s) encoding the cleavage factor, and study of mutations in the cleavage factor gene(s). Experiments dedicated to this goal are in progress.

Materials and methods

Eukaryotic cells, viruses, and bacterial hosts

BSC40 cells, A549 cells, wild-type vaccinia strain WR, and the conditions for their growth and infection have been described previously (Condit and Motyczka, 1981; Condit et al., 1983, 1996a). *Escherichia coli* TB1 cells were used in transformations of pGF17stp, pGF17sense609, and p2098 plasmids.

Plasmids

pGF17stp contains vaccinia DNA from 1147 to 1466 nt downstream of the F17R (gene VACV-WR 056) ATG start site. This region was PCR amplified from the vaccinia WR genome using forward primer 5′ CGGGATCCAGTACTAGAATATCTATATCGCCG3′ and reverse primer 5′ CGGAATTCCTACTTTTCTAAACAGAC3′, digested with *Eco*RI and *Bam*HI, and cloned into the corresponding sites of pGEM-3zf(+) such that transcription from the T7 promoter would yield RNA which is antisense to the F17R transcript although transcription from the SP6 promoter would yield sense transcripts in vitro. pGF17sense609 contains vaccinia DNA from 1111 to 1696 nt downstream of the F17R start site. This region was also PCR amplified from the vaccinia genome using forward primer 5′ CGGAATTCTCTAGAATTAGTCTGAAGAA3′ and reverse primer 5′ CGGGATCCAGTACTAGATTATAGAAACCTATTTA3′, digested with *Eco*RI and *Bam*HI, and cloned into pGEM-3zf(+) such that transcripts sense to the F17R transcripts could be synthesized from the T7 promoter. Plasmid p2098, which contains the 3′ end of ATI transcript, has been previously described (Antczak et al., 1992).

In vitro transcription

pGF17stp was linearized with *Eco*RI to generate a 430-nt substrate in the sense orientation relative to F17R transcripts or with *Xba*I to generate a 402-nt substrate antisense to F17R transcripts. pGF17sense609 was linearized with *Acc*I to synthesize a 410-nt substrate in the sense orientation relative to F17R transcripts. DNA from the p2098 clone was linearized with *Nco*I to synthesize a 2.7-kb substrate in the sense orientation relative to ATI transcripts. To synthesize RNA substrates, 1 μg of linearized DNA template was incubated for 60 min at room temperature in a 20-μl reaction using a MAXIScript SP6/T7 kit (Ambion) with either T7 (antisense transcript from pGF17stp, sense transcript from

pGF17sense609, or sense transcript from p2098) or SP6 RNA polymerase (sense transcript from pGF17stp). The DNA template was then removed from the reaction by digestion with 1 unit of RNase-free DNase (Ambion) at 37 °C for 15 min. The RNA substrate was then recovered by phenol–chloroform extraction and concentrated by ethanol precipitation in the presence of 2 M NH_4-acetate. The specific activity of the sense and antisense transcripts from pGF17stp and the sense transcript from pGF17sense609 was approximately 5×10^7 cpm/pmol RNA. The specific activity of the ATI transcript from p2098 was approximately 4×10^8 cpm/pmol RNA.

Northern analyses

Confluent 100-mm dishes of BSC40 cells were either mock-infected or infected with wt vaccinia virus at an moi of 15 and incubated at 37 °C for 6, 8, and 12 h p.i. At the various times postinfection, total RNA was extracted from cells using the RNeasy kit (Qiagen). Two micrograms of each sample was then fractionated on a 1.2% agarose gel containing 2.2 M formaldehyde. Northern blotting and transfer were done using established procedures (Ausebel et al., 1994). The transferred RNA was cross-linked onto GeneScreen membrane (NEN Life Sciences) with a *StrataLinker 2400* (Stratagene). Tandem oligonucleotides to be used as probes were synthesized by *Genemed Syn* (CA). The positions of these nucleotides in the WR genome (GenBank accession, AY243312) are 44893-44840 (probe a), 44926-44862 (probe b), 44976-44927 (probe c), 45020-44974 (probe d), 45076-45020, 45126-45073, 45186-45126, and 45238-45185. The probes were 5′-end-labeled with [γ-^{32}P] ATP using T_4 kinase (New England Biolabs). Unincorporated nucleotides were removed using a 1-ml G25 Sephadex column. Hybridization was carried out at 42 °C using the protocol described by Ausebel et al. (1994) with some modifications. The prehybridization and hybridization solutions contained 6 × SSC, 0.1% SDS, 10 × Denhardt's [50 × Denhardt's is 1% (w/v) Ficoll 400, 1% (w/v) BSA (Sigma, Fraction V), and 1% (w/v) polyvinylpyrrolidone], and 100 μg single-stranded salmon sperm DNA. After hybridization, the membrane was washed twice in 5 × SSC, 0.1% SDS at 50 °C for 20 min each, blotted dry, and exposed to X-ray film.

Preparation of infected cell extracts

Confluent 100-mm dishes of A549 cells were either mock-infected or infected with wt vaccinia virus at an moi of 15 and incubated at 37 °C for 16–18 h in the presence or absence of 10 mM hydroxyurea. Extracts were prepared as described previously (Condit et al., 1996a). Briefly, infected cell monolayers were permeabilized with lysolecithin, harvested, treated with micrococcal nuclease, clarified by centrifugation, and stored in aliquots at − 70 °C. Total protein concentration was determined using the Bradford protein assay (Bio-Rad). Partially purified infected extracts used in Fig. 6 were from a purification procedure, which is currently in progress in our laboratory, the details of which will be published later. Briefly, suspension HeLa cells were infected with vaccinia virus for 16 h, and the cells were harvested by centrifugation, lysed by dounce homogenization, and separated into nuclear and cytoplasmic fractions by centrifugation. Most of the cleavage activity was associated with the nuclear fraction. A high salt, sonicated, clarified extract of the nuclear fraction, called "Extract" in Fig. 6, was used for further purification. This extract was passed over a column of DEAE cellulose at 400 mM NaCl, and the flowthrough from this column that contains the cleavage activity is called "DEAE" in Fig. 6. The DEAE fraction was applied to an SP sepharose column and eluted with a NaCl gradient. Active fractions eluting between 300 and 750 mM NaCl are called "SP" in Fig. 6.

RNA cleavage assays

^{32}P-labeled RNA substrates prepared by in vitro transcription were used in the RNA cleavage assays at 1000 or 10 000 cpm per reaction. Assays were typically carried out at 30 °C for 60 min in the presence of 20 μg infected cell extract, 40 mM Tris–HCl, pH 7.5, 300 mM NaCl, and 7.5 mM $MgCl_2$ in a final volume of 25 μl. Samples were then treated with 125 μl of "PK mix" (114 mM Tris–HCl, pH 7.5, 14 mM EDTA, 171 mM NaCl, 1.14% SDS, 40 μg glycogen, 230 μg proteinase K) at 37 °C for at least 15 min, followed by a single extraction with 150 μl phenol–chloroform. RNA was precipitated by addition of 40 μl 10 M NH_4-acetate and 500 μl of ethanol, incubated at − 20 °C for 10 min, and centrifuged at maximum speed (16 000 × *g*) for 30 min. Pellets were washed once with 70% ethanol, dried, and resuspended in 10 μl formamide loading buffer. Samples were denatured at 90–95 °C for 3 min and loaded on a 6% polyacrylamide gel containing 8 M Urea (Sequagel, National Diagnostics). Gels were fixed, dried, and visualized by autoradiography. DNA ladders (50 bp, Gibco BRL) were end-labeled with ^{32}P [ATP] using polynucleotide kinase and used as molecular weight markers. Cleavage activity was calculated as the ratio of the sum of the intensities of the two cleavage products to the sum of the intensities of the uncleaved substrate and the products of cleavage. This number was then corrected for background, multiplied by 100, and denoted as percentage of cleavage.

Nuclease S1 assays

BSC40 cells were infected with wt vaccinia virus at an moi of 15 and RNA was extracted at 12 h p.i. as for northerns. The DNA oligonucleotide, corresponding to 44998–44931 nt in the WR genome, used to detect the 3′ ends of the F17R RNA was labeled at the 3′ end with [α-^{32}P] (cordycepin-5′-triphosphate) (NEN) using terminal deoxynucleotidyl transferase (Promega). Briefly, 10 pmol of primer were incubated with 125 μCi [α-^{32}P] (cordycepin-

5′-triphosphate) (5000 Ci/mmol) and 30 units of terminal transferase at 37 °C for 30 min. The reaction was stopped by heating at 70 °C for 10 min and the radiolabeled probe was purified by phenol–chloroform extraction, and precipitated in the presence of 2 M NH_4-acetate and 20 μg glycogen with 2.5 volumes of ethanol. Pellets were washed with 70% ethanol, dried, and resuspended in 20 μl DEPC-treated water. The denatured probe (10^6 cpm) was hybridized with 10 μg RNA in 30 μl buffer containing 1 M NaCl, 0.15 M HEPES, pH 7.5, and 0.3 mM EDTA, pH 8, heated to 75 °C for 10 min, and then incubated at 42 °C overnight. RNA–DNA hybrids were digested with 75 units of nuclease S1 (Promega) in 300 μl of S1 buffer (0.28 M NaCl, 50 mM NaOAc, pH 4.5, 4.5 mM $ZnSO_4$) and 6 μg single-stranded salmon sperm DNA at 37 °C for 30 min. At the end of the incubation, the reaction was stopped by adding 3 μl of 0.5 M EDTA, 20 μg of glycogen was added as carrier and the nuclease-resistant products were precipitated with 700 μl ethanol. Pellets were washed with 70% ethanol, dried, and resuspended in 10 μl 0.1 M NaOH. Three microliters of nuclease resistant products were combined with an equal volume of formamide loading dye, heated at 90 °C for 3 min, and electrophoresed on a 10% polyacrylamide gel containing 8 M Urea alongside a DNA sequencing ladder. The gels were fixed, dried, and visualized by autoradiography. The DNA ladder used to determine the 3′ end of the protected fragment after nuclease S1 digestion was sequenced using the Sequenase v2.0 kit (USB Corporation). The 5′ end of the primer (5′ TGGCGTCTAAGATATT-CTTCC3′) used for sequencing was designed to be complementary to the 3′ end of the labeled oligonucleotide probe.

Acknowledgments

This work was supported by Grant # RO1 AI 18094 from the National Institutes of Health.

References

Amegadzie, B.Y., Sisler, J.R., Moss, B., 1992. Frame-shift mutations within the vaccinia virus A-type inclusion protein gene. Virology 186, 777–782.

Antczak, J.B., Patel, D.D., Ray, C.A., Ink, B.S., Pickup, D.J., 1992. Site-specific RNA cleavage generates the 3′ end of a poxvirus late mRNA. Proc. Natl. Acad. Sci. U.S.A. 89, 12033–12037.

Ausebel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., Struhl, K., 1994. Current protocols in molecular biology. Wiley, New York.

Black, E.P., Condit, R.C., 1996. Phenotypic characterization of mutants in vaccinia virus gene G2R, a putative transcription elongation factor. J. Virol. 70, 47–54.

Christen, L.M., Sanders, M., Wiler, C., Niles, E.G., 1998. Vaccinia virus nucleoside triphosphate phosphohydrolase I is an essential viral early gene transcription termination factor. Virology 245, 360–371.

Condit, R.C., Motyczka, A., 1981. Isolation and preliminary characterization of temperature-sensitive mutants of vaccinia virus. Virology 113, 224–241.

Condit, R.C., Motyczka, A., Spizz, G., 1983. Isolation, characterization, and physical mapping of temperature-sensitive mutants of vaccinia virus. Virology 128, 429–443.

Condit, R.C., Lewis, J.I., Quinn, M., Christen, L.M., Niles, E.G., 1996a. Use of lysolecithin-permeabilized infected-cell extracts to investigate the in vitro biochemical phenotypes of poxvirus ts mutations altered in viral transcription activity. Virology 218, 169–180.

Condit, R.C., Xiang, Y., Lewis, J.I., 1996b. Mutation of vaccinia virus gene G2R causes suppression of gene A18R ts mutants: implications for control of transcription. Virology 220, 10–19.

Deng, L., Shuman, S., 1998. Vaccinia NPH-I, a DExH-box ATPase, is the energy coupling factor for mRNA transcription termination. Genes Dev. 12, 538–546.

Dye, M.J., Proudfoot, N.J., 2001. Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II. Cell 105, 669–681.

Gershon, P.D., 1998. mRNA 3′ end formation by vaccinia virus: mechanism of action of a heterodimeric poly (A) polymerase. Semin. Virol. 8, 343–350.

Howard, S.T., Ray, C.A., Patel, D.D., Antczak, J.B., Pickup, D.J., 1999. A 43-nucleotide RNA *cis*-acting element governs the site-specific formation of the 3′ end of a poxvirus late mRNA. Virology 255, 190–204.

Lackner, C.A., Condit, R.C., 2000. Vaccinia virus gene A18R DNA helicase is a transcript release factor. J. Biol. Chem. 275, 1485–1494.

Latner, D.R., Xiang, Y., Lewis, J.I., Condit, J., Condit, R.C., 2000. The vaccinia virus bifunctional gene J3 (nucleoside-2′-O-)-methyltransferase and poly(A) polymerase stimulatory factor is implicated as a positive transcription elongation factor by two genetic approaches. Virology 269, 345–355.

Latner, D.R., Thompson, J.M., Gershon, P.D., Storrs, C., Condit, R.C., 2002. The positive transcription elongation factor activity of the vaccinia virus J3 protein is independent from its (nucleoside-2′-O-) methyltransferase and poly(A) polymerase stimulatory functions. Virology 301, 64–80.

Mahr, A., Roberts, B.E., 1984. Arrangement of late RNAs transcribed from a 7.1-kilobase *Eco*RI vaccinia virus DNA fragment. J. Virol. 49, 510–520.

Mohamed, M.R., Niles, E.G., 2000. Interaction between nucleoside triphosphate phosphohydrolase I and the H4L subunit of the viral RNA polymerase is required for vaccinia virus early gene transcript release. J. Biol. Chem. 275, 25798–25804.

Moss, B., 2001. Poxviridae: the viruses and their replication. In: Knipe, D.M., Howley, P.M. (Eds.), Fields Virology. Lippincott Williams & Wilkins, Philadelphia, pp. 2849–2884.

Parsons, B.L., Pickup, D.J., 1990. Transcription of orthopoxvirus telomeres at late times during infection. Virology 175, 69–80.

Patel, D.D., Pickup, D.J., 1987. Messenger RNAs of a strongly-expressed late gene of cowpox virus contain 5′-terminal poly(A) sequences. EMBO J. 6, 3787–3794.

Patel, D.D., Pickup, D.J., 1989. The second-largest subunit of the poxvirus RNA polymerase is similar to the corresponding subunits of procaryotic and eucaryotic RNA polymerases. J. Virol. 63, 1076–1086.

Patel, D.D., Pickup, D.J., Joklik, W.K., 1986. Isolation of cowpox virus A-type inclusions and characterization of their major protein component. Virology 149, 174–189.

Patel, D.D., Ray, C.A., Drucker, R.P., Pickup, D.J., 1988. A poxvirus-derived vector that directs high levels of expression of cloned genes in mammalian cells. Proc. Natl. Acad. Sci. U.S.A. 85, 9431–9435.

Proudfoot, N., 2000. Connecting transcription to messenger RNA processing. Trends Biochem. Sci. 25, 290–293.

Proudfoot, N., O'Sullivan, J., 2002. Polyadenylation: a tail of two complexes. Curr. Biol. 12, R855–R857.

Proudfoot, N.J., Furger, A., Dye, M.J., 2002. Integrating mRNA processing with transcription. Cell 108, 501–512.

Rohrmann, G., Yuen, L., Moss, B., 1986. Transcription of vaccinia virus

early genes by enzymes isolated from vaccinia virions terminates downstream of a regulatory sequence. Cell 46, 1029–1035.

Shuman, S., Moss, B., 1988. Factor-dependent transcription termination by vaccinia virus RNA polymerase. Evidence that the *cis*-acting termination signal is in nascent RNA. J. Biol. Chem. 263, 6220–6225.

Shuman, S., Broyles, S.S., Moss, B., 1987. Purification and characterization of a transcription termination factor from vaccinia virions. J. Biol. Chem. 262, 12372–12380.

Wittek, R., Hanggi, M., Hiller, G., 1984. Mapping of a gene coding for a major late structural polypeptide on the vaccinia virus genome. J. Virol. 49, 371–378.

Xiang, Y., Simpson, D.A., Spiegel, J., Zhou, A., Silverman, R.H., Condit, R.C., 1998. The vaccinia virus A18R DNA helicase is a postreplicative negative transcription elongation factor. J. Virol. 72, 7012–7023.

Xiang, Y., Latner, D.R., Niles, E.G., Condit, R.C., 2000. Transcription elongation activity of the vaccinia virus J3 protein in vivo is independent of poly(A) polymerase stimulation. Virology 269, 356–369.

Yuen, L., Moss, B., 1987. Oligonucleotide sequence signaling transcriptional termination of vaccinia virus early genes. Proc. Natl. Acad. Sci. U.S.A. 84, 6417–6421.

Zuker, M., Jaeger, J.A., Turner, D.H., 1991. A comparison of optimal and suboptimal RNA secondary structures predicted by free energy minimization with structures determined by phylogenetic comparison. Nucleic Acids Res. 19, 2707–2714.