Overlapping divergent transcripts mapping to the HindIII F region of the Autographa californica nuclear polyhedrosis virus

Steven E. Hardin† and Robert F. Weaver*

Department of Biochemistry, University of Kansas, Lawrence, Kansas 66045, U.S.A.

We have mapped a set of transcripts that cross the XhoI site in the HindIII F fragment of the genome of the Autographa californica nuclear polyhedrosis virus. These transcripts overlap at their 5' ends by about 550 bases and run in opposite directions. We have tentatively identified two open reading frames corresponding to the leftward transcript. The rightward transcript is present from 8 h post-infection (p.i.) to 24 h p.i.; the leftward transcripts are present from 2 h p.i. to 24 h p.i. The early (2 h) transcript is about 2-1 kb in size and its 5' end maps about 504 bp to the right of the XhoI site. Beginning at about 8 h p.i. a new transcription start site is used, about 80 bp downstream (to the left) of the first. The late (rightward) transcript is about 1.2 kb in size; its 5' end seems to be heterogeneous and maps about 44 to 60 bp to the left of the XhoI site. Late in infection transcription proceeds in both directions at the same time through the overlapping region of the DNA encoding these transcripts.

The nuclear polyhedrosis virus of the alfalfa looper, Autographa californica, (AcNPV) replicates in the nuclei of infected lepidopteran insect cells. It is a large double-stranded DNA virus with a closed circular genome of 128 kbp. Early transcription is accomplished by use of the host RNA polymerase II (Fuchs et al., 1983). Later in infection, a novel virus-induced RNA polymerase is responsible for most viral transcription (Fuchs et al., 1983; Grula et al., 1981). During the infection cycle there is a very strict temporal regulation of gene expression (Gordon & Carstens, 1984; Kelly & Lescott, 1981; Miller et al., 1983).

Overlapping transcripts appear to be a common feature of AcNPV; they have now been found in at least five different regions of the genome (Guarino & Summers, 1986; Lübbert & Doerfler, 1984; Kuzio et al., 1984; Rankin et al., 1986; Friesen & Miller, 1985; Oellig et al., 1987). One of the best studied examples was discovered by Friesen & Miller (1985). This is a nested set of rightward transcripts beginning in the HindIII K region and terminating in the HindIII Q region. The 3' ends of these transcripts are all the same, but the cap sites are progressively more leftward (upstream) as the infection progresses. All of these transcripts, including the earliest (α1), encode only one, 35K protein. Two additional transcripts (α3 and α4) run in the opposite direction through the same region and encode a 95K protein (Friesen & Miller, 1987). The controls involved in this complex transcription system are not yet clear.

We present here evidence for divergent early and late transcripts in the HindIII F fragment of AcNPV that overlap at their 5' ends by over 500 bases.

As an approach to the investigation of transcripts of the region upstream of the polyhedrin gene, we prepared Northern blots of RNAs harvested from infected cells at 2, 6, 12 and 24 h post-infection (p.i.), and probed them with a nick-translated plasmid containing the HindIII F fragment. Fig. 1(a) shows that two major transcripts were detected early in the infection, with sizes of approximately 2-1 kb and 1-2 kb. It appears that late in infection an additional smaller transcript of about 0.8 kb is detected. Since the polyhedrin gene is partially contained within the HindIII F fragment and has a transcript size of approximately 1.2 kb, the 1-2 kb band detected here at 12 and 24 h p.i. probably includes polyhedrin transcripts. However, the 1.2 kb band was detected as early as 2 and 6 h p.i.; at these early times, this band must include another transcript of approximately the same size as the polyhedrin mRNA.

To locate the 5' termini of the major leftward transcripts that cross the XhoI site in the HindIII F region, we performed S1 nuclease mapping. Fig. 2 shows a partial map of the HindIII F/V region, along with the overlapping transcripts mapped in this work, and the polyhedrin transcript. The unique XhoI site in the HindIII F fragment was end-labelled, and the right-hand XhoI–HindIII fragment was used as a probe for S1
Fig. 1. (a) Northern blot of transcripts from the HindIII F region. RNAs from cells 2, 6, 12 and 24 h p.i., as indicated at top, were electrophoresed on 1% agarose/methylmercuric hydroxide gels (Bailey & Davidson, 1976) and Northern-blotted as described by Maniatis et al. (1982). The blot was hybridized to a 32P-labelled HindIII F DNA probe and autoradiographed. In all panels, the mobilities of RNA markers (kb) are shown at the right, and the sizes (kb) of major RNA species are given at the left. (b) Northern blot of leftward transcripts. RNAs from mock-infected cells (mi) and cells 2 to 24 h p.i., as indicated at the top, were electrophoresed and Northern-blotted as described for (a), and hybridized to a 32P-labelled rightward transcript of the XhoI-SalI fragment made with T7 RNA polymerase in vitro. (c) Northern blot of rightward transcripts. RNAs from mock-infected cells (mi) and from cells at the times p.i. indicated at the top were electrophoresed and Northern-blotted as described for (a). The blot was hybridized to a 32P-labelled leftward transcript of the XhoI–SalI fragment, made with T3 RNA polymerase in vitro.

Fig. 2. Summary of transcript mapping. The arrows represent major transcripts mapped in the right-hand EcoRI–HindIII region of HindIII F. The rightmost arrow represents the polyhedrin transcript. The double vertical line at the right end of the leftward arrow represents the early and major late 5' ends of the leftward transcripts.

mapping. The results are shown in Fig. 3(a). RNAs from cells infected for 2 to 24 h were used in this experiment, and the 5' terminus of an early transcript was detected approximately 504 bases upstream from the XhoI site. The transcript was present as early as 2 h p.i. and its steady-state concentration increased up to 8 h p.i. At 8 h p.i. two new transcription start sites were detected. One of these lies about 90 bp upstream, the other about 80 bp downstream, from the original cap site. The transcript corresponding to the upstream start site is short-lived and is no longer detected by 12 h p.i. The transcript corresponding to the new downstream start site, however, continues to accumulate until at least 24 h p.i. Therefore, this DNA region appears to be transcribed during the late phase of infection, taking advantage of a new, downstream start site.

To make sure of the location and orientation of the 2-1 kb transcript detected with the whole HindIII F probe, a Northern blot was performed with a shorter, strand-specific RNA probe, generated by transcribing the XhoI–SalI region in the rightward direction. This method was used to probe Northern blots of RNA harvested from cells infected for 2 to 24 h. The results are shown in Fig. 1(b). A transcript approximately 2-1 kb in size was detected. This transcript was present as early as 2 h p.i. and increased in abundance between 8 and 12 h p.i.; it remained detectable at 24 h p.i. These results match those of the 5' end S1 mapping experiment discussed above. Thus, it appears that the transcripts whose 5' ends map about 500 bp to the right of the XhoI site are represented by the 2-1 kb band detected by Northern blotting. Allowing 200 bases for a poly(A) tail, the coding region for the early transcript is about 1-9 kb. This positions the 3' end of this transcript about 1-5 kb to the left of the XhoI site.
To study rightward transcripts crossing the \textit{XhoI} site, the left-hand \textit{XhoI–HindIII} fragment was 5' end-labelled at the \textit{XhoI} site and used for S1 mapping. Fig. 3(b) shows the results of S1 mapping of RNA isolated from cells infected for 2 to 24 h. A set of signals is first detected at 8 h p.i. and becomes more prominent at least until 18 h p.i. The transcription start site represented by these bands appears to be heterogeneous and maps to positions from 44 to 60 bp to the left of the \textit{XhoI} site.

Primer extension analysis was conducted to verify the 5' ends of these transcripts. The 234 bp \textit{XhoI–SphI} fragment (see Fig. 2) was 3' end-labelled at the \textit{XhoI} site and used as the primer for this experiment. As Fig. 3(c) shows, an extended product was detected with RNA from cells 8 to 24 h p.i. The product is heterogeneous, suggesting heterogeneous 5' start sites. The primer was extended 44 to 60 bases, positioning the \textit{XhoI} site 44 to 60 bp downstream from the 5' ends of the transcripts. All of these results corroborate those obtained by S1 mapping.

To measure the size of the rightward transcript, an \textit{in vivo} leftward transcript of the \textit{XhoI–SalI} fragment (Fig. 2) was used to probe Northern blots of RNA prepared at 2 to 24 h p.i. (Fig. 1c). A 1-2 kb transcript was detected as early as 8 h p.i. This result (allowing for 200 bases of poly(A) tail), positions the 3' end of the late transcripts about 1 kb to the right of the \textit{XhoI} site, near the leftmost \textit{SalI} site.

Fig. 2 summarizes the mapping data on the rightward and leftward transcripts crossing the \textit{XhoI} site. These transcripts run in opposite directions and overlap by about 550 bases. It appears that late in infection transcription is proceeding in two directions at the same time, through the same DNA region.

To test this hypothesis, nuclear run-on assays were performed. Nuclei were harvested from infected cells at 4, 8 and 12 h p.i. Transcription in the isolated nuclei yielded radioactive RNAs that were then hybridized to blots of leftward and rightward transcripts of the \textit{XhoI–SalI} fragment. T7 transcripts are strand-specific probes for the early and late leftward transcripts, and T3
transcripts are strand-specific probes for the late, rightward transcripts. As a negative control, T7 and T3 were prepared, mixed, and blotted also. Probes did not hybridize to a significant degree to either T3 or T7 transcripts when nuclei harvested at 4 h p.i. were used (compare with the level of hybridization to the late, rightward transcript). However, with RNA synthesized in vitro, two open reading frames corresponding to the leftward transcript, but none corresponding to the rightward transcript. The latter transcript may have a control function.

The data presented here suggest that a switch in transcription pattern around the XhoI site in HindIII F occurs around 8 h p.i. Leftward transcription occurs from two new cap sites, and transcripts from one of these persist during the remainder of the infection cycle. This suggests tandem promoters that are used sequentially. Also, at this time the late, rightward transcription begins.

Other genes of AcNPV have been shown to acquire new transcription start sites as infection progresses. The p39 gene has a tight cluster of three early start sites but later in infection an additional site is detected approximately 20 bp upstream (Guarino & Summers, 1986). The set of overlapping transcripts reported by Friesen & Miller (1985) that run from the HindIII K to HindIII Q region, have a number of new 5' cap sites that are systematically used as infection progresses. Further-
more, these workers report another set of overlapping transcripts in the same region that run in the opposite direction (Friesen & Miller, 1987). Therefore, this area contains overlapping transcripts with a number of different cap sites. Each new site probably represents a new promoter.

The transcripts reported here are similar to those described by Friesen & Miller (1987) in that they overlap and run in opposite directions. The expression of each is temporally controlled, one beginning early and the other late. As with the transcripts in HindIII K, the gene represented by the early transcript continues to be expressed throughout the entire infection cycle by taking advantage of new transcription start sites (and probably new promoters) during the late phase.

This virus may use this promoter switching mechanism to guarantee expression of a gene throughout the infection cycle. This is probably necessitated by the fact that two different RNA polymerases are active early and late in infection: the host RNA polymerase II transcribes the early genes, whereas a virus-induced RNA polymerase transcribes the late genes (Fuchs et al., 1983). This latter polymerase apparently recognizes its own class of promoters (Possee & Howard, 1987; Matsuura et al., 1987; Weyer & Possee, 1988; Rankin et al., 1988; Qin et al., 1989). Therefore, the presence of both types of promoters in certain genes ensures their expression both early and late.

We thank Judy Bevan for expert technical assistance. This work was supported by grant no. MV259 from the American Cancer Society and by grant no. GM22127 from the National Institutes of Health.

References


(Received 12 April 1989; Accepted 12 September 1989)