Transcriptional analyses of baculovirus polyhedrin and foreign gene expression relative to baculovirus p10 mRNA levels

Mi-Kyung Min and David H. L. Bishop*

NERC Institute of Virology and Environmental Microbiology, Mansfield Road, Oxford OX1 3SR, U.K.

Comparisons have been made between the p10 and polyhedrin mRNA levels recovered from Spodoptera frugiperda cells infected with Autographa californica nuclear polyhedrosis virus (AcNPV). In molar terms and from 18 h post-infection (p.i.), the polyhedrin mRNA species increased to levels one and a half times to twice as high as the p10 levels. The influence of the polyhedrin leader sequence on the expression of a foreign gene under the control of the polyhedrin promoter was investigated using a series of four recombinant baculoviruses expressing the lymphocytic choriomeningitis virus nucleocapsid (N) protein gene. The different recombinants varied in the length and composition of the upstream polyhedrin mRNA leader sequence. The recombinant containing the full-length polyhedrin leader sequence gave levels of N mRNA comparable to those of AcNPV polyhedrin mRNA. These levels were either equal to (12 h p.i.) or higher (18 to 42 h p.i.) than the p10 levels at corresponding times. Three other recombinants, with different lengths of leader sequence, accumulated significantly lower quantities of N mRNA in comparison to the p10 mRNA levels. However the mRNA levels for the three recombinants were similar (20 to 50% of the p10 level) and did not correspond to their N protein expression levels. By comparing the mRNA and protein levels, it is concluded that the sequence between -8 to +1 of the AcNPV polyhedrin translation-initiating ATG has an important function for mRNA transcription (or accumulation), while the sequences between -32 to -8 affect the overall translation efficiencies.

The baculovirus Autographa californica nuclear polyhedrosis virus (AcNPV) which is pathogenic for certain lepidopteran insects, has a double-stranded, covalently closed, circular DNA genome of some 130 kbp (Adang & Miller, 1982). The infection course of the virus is separated into early and late phases with certain viral genes expressed at high levels only during the latter phase. Infectious enveloped nucleocapsids (virions) are produced from the beginning of the late phase. They bud from the surface plasma membrane of infected cells. These virions are responsible for spreading virus infection from cell to cell (in tissue culture or within an infected host). During the 'very late' phase, two virus-induced proteins are synthesized in large amounts (representing up to 50% of the stainable protein in cell or larval extracts). These are the polyhedrin and p10 proteins. They are involved in occluding viral nucleocapsids into polyhedron-shaped occlusion bodies in the nucleus of infected cells. The inclusion bodies protect the viral nucleocapsids in the environment and between hosts.

Recently, baculoviruses have become intensely studied because of their usefulness as expression vectors (Smith et al., 1983a). The AcNPV polyhedrin and p10 promoters have been used to drive expression of foreign genes originating from prokaryotes or eukaryotes (Luckow & Summers, 1988; Matsuura et al., 1986, 1987; Weyer & Possee, 1988, 1989). Although high level expression of a foreign gene is frequently obtained, for certain genes the expression is low. The reason is unknown; it could be due to a variety of causes such as mRNA transcription rates, stability of the mRNA species, or translation efficiencies.

It has been shown previously that the length of the normal 5' non-coding leader sequence of the baculovirus polyhedrin mRNA influences the level to which a foreign gene such as the nucleocapsid (N) protein of lymphocytic choriomeningitis virus (LCMV; Matsuura et al., 1987) is expressed. In that report, protein expression was investigated for four LCMV N recombinant baculoviruses which differed in the amount and sequence of the normal polyhedrin mRNA leader. These and similar data for a variety of other foreign genes have shown that the level of expression is related to the extent that the polyhedrin leader sequence is conserved, although there are also other factors that affect expression levels (see Takehara et al., 1990). In order to determine the molecular basis of these observations we
have now analysed the accumulated mRNA levels of the four recombinants between 12 and 42 h post-infection (p.i.) using Northern blotting techniques. We have also quantified the relative accumulation (in molar terms) of polyhedrin and p10 mRNA in wild-type AcNPV infections. In order to reduce inaccuracies due to differences between or within experiments, we devised a normalization procedure using the mRNA of the AcNPV p10 gene as the basis for comparison, and chimeric sequences to provide gene probes of similar specific activities. Kinetic graphs of p10: polyhedrin and p10: LCMV N mRNA synthesis were derived in this way.

Two probes, pLCMNp10 and pPOLYp10, were constructed. The pLCMNp10 probe contained an LCMV N oligomer and an AcNPV p10 oligomer, and pPOLYp10 probe contained an AcNPV polyhedrin oligomer and the same p10 oligomer (Fig. 1). The probes were constructed so that the content of AT residues was as nearly equivalent as possible for all three gene sequences. The 222 nucleotide LCMV N oligomer had a base composition of A, 63; G, 64; C, 41; T, 54; the 224 nucleotide p10 oligomer a composition of A, 66; G, 46; C, 55; T, 57; and the 226 nucleotide polyhedrin oligomer a composition of A, 68; G, 49; C, 55; T, 54.

The RP5N, RP6N, Ac373N and YON recombinants used in these studies were those previously designated RP5YN12, RP6YN2, S373YN1 and YM1YN1 respectively (Matsuura et al., 1987). All four viruses contain the LCMV S sequences from the LCMV BamHI site (LCMV residues 9 to 14; Romanowski et al., 1985), through the ATG codon of the N gene (LCMV residues 61 to 63) and the complete N coding region (LCMV residues 61 to 1734; Romanowski & Bishop, 1985). RP5N, RP6N and Ac373N also contain the LCMV S sequences up to residue 3356 (Romanowski et al., 1985).
Fig. 2. Northern blot analysis of mRNA species. *S. frugiperda* cells were infected with (a) AcNPV (i), RP5N (ii), RP6N (iii), Ac373N (iv) or YON (v), or (b) RP6N at an m.o.i. of 10 p.f.u/cell and incubated at 28°C in a spinner flask. Total cellular RNA was extracted at 6 h intervals until 42 h p.i. (lanes 1 to 6 respectively; for RP6N the first time point was at 15 h p.i.). The poly(A)-containing RNA species were recovered by a small volume batch technique using oligo(dT)-cellulose and using two elutions with Tris–EDTA buffer to recover the mRNA (Manley et al., 1979). Re-extraction of the material bound to the cellulose did not elute any more mRNA as shown by the appropriate Northern analyses (i.e. <5% of the eluted species was mRNA). Analysis of the unbound material did not identify additional RNA that hybridized to the probes (<5% of the eluted species), indicating that the genes were efficiently transcribed into mRNA species containing poly(A) and that degradation of mRNA was not a significant factor in the analyses.

Preliminary analyses established the conditions to obtain reproducible RNA transfer and hybridization and to ascertain that all the RNA was transferred (none was detected on re-blotting), that there was no RNA that passed through the filters (using double membranes etc.), and that further hybridization (increased time or amount of probe etc.) did not give higher signals.

Riboprobes were initially considered for the hybridization analyses in view of the high specific activities that can be obtained. However it was found that the riboprobes transcribed by T7 polymerase from pLCMNp10 or pPOLYbp10 contained a mixture of product sizes. Since the p10 oligomer was placed 3' to the LCMV (or polyhedrin) sequence in each probe, it was considered inadvisable to use probes that may have variable representations of the two gene sequences. Because of this, nick-translated probes were employed. In a preliminary study, analysis of mRNA preparations recovered from YON-infected cells using Northern blotting procedures showed that at a multiplicity of 10, the peak of p10 and N mRNA accumulation was between 24 and 36 h p.i., although mRNA was identified up to 60 h p.i. A significant proportion (>20%) of the infected cells began to lyse by 48 h p.i., accounting for the reduced quantities of RNA recovered (although this was not reflected by the analyses of accumulated proteins, presumably since the proteins were cell-associated and less readily degraded). Consequently the mRNA analyses were restricted to times between 12 and 42 h p.i.

In time course studies with YON, Ac373N, RP6N, RP5N or AcNPV, Northern blot analyses using the pLCMNp10 and pPOLYbp10 probes clearly identified the major bands of mRNA representing the p10, LCMV N, or polyhedrin mRNA species (Fig. 2a). The identity of each band was confirmed using individual plasmid- and strand-specific probes representing each gene sequence (data not shown). Only one major band was identified for either the p10 or polyhedrin genes. When exposed for longer periods, minor bands larger than the polyhedrin (or p10) mRNA species were identified using gene-specific probes; however, they constituted <5% of

as well as the carboxy-proximal polyhedrin coding region (AcNPV residues 177 to 738, Hooft van Idderkinge et al., 1983) and the AcNPV 3' flanking sequences (i.e. 0-33 kbp to the polyhedrin transcription termination site). The YON recombinant lacks LCMV 1774 to 3356, as well as the carboxy-proximal polyhedrin coding sequences and 13 residues of the region immediately 3' of the polyhedrin gene (Matsuura et al., 1987).

Virus-infected cells were lysed using SDS and poly(A)-containing mRNA was prepared. The presence of ribosomal RNA inhibited the identification and resolution of the mRNA species when high amounts of RNA were loaded on the gels (i.e. for the early time points). For all samples thereafter, the mRNA species were isolated by a small volume batch technique using oligo(dT)-cellulose and using two elutions with Tris–EDTA buffer to recover the mRNA (Manley et al., 1979). Re-extraction of the material bound to the cellulose did not elute any more mRNA as shown by the appropriate Northern analyses (i.e. <5% of the eluted species was mRNA). Analysis of the unbound material did not identify additional RNA that hybridized to the probes (<5% of the eluted species), indicating that the genes were efficiently transcribed into mRNA species containing poly(A) and that degradation of mRNA was not a significant factor in the analyses.

Preliminary analyses established the conditions to obtain reproducible RNA transfer and hybridization and to ascertain that all the RNA was transferred (none was detected on re-blotting), that there was no RNA that passed through the filters (using double membranes etc.), and that further hybridization (increased time or amount of probe etc.) did not give higher signals.

Riboprobes were initially considered for the hybridization analyses in view of the high specific activities that can be obtained. However it was found that the riboprobes transcribed by T7 polymerase from pLCMNp10 or pPOLYbp10 contained a mixture of product sizes. Since the p10 oligomer was placed 3' to the LCMV (or polyhedrin) sequence in each probe, it was considered inadvisable to use probes that may have variable representations of the two gene sequences. Because of this, nick-translated probes were employed. In a preliminary study, analysis of mRNA preparations recovered from YON-infected cells using Northern blotting procedures showed that at a multiplicity of 10, the peak of p10 and N mRNA accumulation was between 24 and 36 h p.i., although mRNA was identified up to 60 h p.i. A significant proportion (>20%) of the infected cells began to lyse by 48 h p.i., accounting for the reduced quantities of RNA recovered (although this was not reflected by the analyses of accumulated proteins, presumably since the proteins were cell-associated and less readily degraded). Consequently the mRNA analyses were restricted to times between 12 and 42 h p.i.

In time course studies with YON, Ac373N, RP6N, RP5N or AcNPV, Northern blot analyses using the pLCMNp10 and pPOLYbp10 probes clearly identified the major bands of mRNA representing the p10, LCMV N, or polyhedrin mRNA species (Fig. 2a). The identity of each band was confirmed using individual plasmid- and strand-specific probes representing each gene sequence (data not shown). Only one major band was identified for either the p10 or polyhedrin genes. When exposed for longer periods, minor bands larger than the polyhedrin (or p10) mRNA species were identified using gene-specific probes; however, they constituted <5% of
the amount in the principal band and consequently were not taken into consideration in the subsequent data analyses. In the case of the Ac373N and RP6N virus-infected cells, two LCMV N-specific mRNA species were consistently detected (Fig. 2a; labelled N-short transcript and N-long transcript). The short transcript predominated. For RP5N, the N-long transcript could only be detected after extended exposure of the blots. Discounting any contribution of the poly(A) sequence, the expected sizes of the AcNPV mRNA species are (p10) 0.6 kb and (polyhedrin) 1.2 kb (Smith et al., 1983b). The expected size of the LCMV N mRNA species from Ac373N, RP6N and RP5N terminating at the polyhedrin transcription termination site is 4-3 kb, since these constructs contain the complete 3.4 kbp LCMV S coding region (Matsuura et al., 1987) as well as the 0.9 kbp coding and non-coding polyhedrin sequence 3' to the insert. The sizes of the p10 and N mRNA species were determined (Fig. 2b) using marker RNA standards (BRL, 0.24 to 9.5 kb) run in parallel and stained with ethidium bromide to identify their positions. The location of the p10 mRNA agreed with the 0.6 kb value. However, by comparison with the markers it was concluded that the N-long transcript was about 3 kb and the N-short transcript 1.8 kb. In view of this and in relation to the known sequence of the incorporated LCMV sequence, it was concluded that the N-short transcript represents transcription termination at or near the LCMV N GPC intergenic sequence (located 1.8 kb into the S coding sequence), whereas the N-long transcript represents termination at a site within the GPC region of LCMV S cDNA (Romanowski et al., 1985). For YON, the mRNA size was estimated to be 2-1 kb. Since the intergenic sequence, GPC, and residual polyhedrin coding sequence 3' to the insert were absent (i.e. the remaining non-coding 3' sequence to the polyhedrin transcription termination site was 0.3 kbp; see Matsuura et al., 1987), it was concluded that the 2.1 kb YON N mRNA terminated at the normal polyhedrin termination site (i.e. 1.8 kb of N and 0.3 kb of the 3' AcNPV sequence). No analyses were undertaken to confirm these conclusions.

The data shown in Fig. 2(a) are representative of three separate infections undertaken using the same experimental protocol. For RP6N, 15 h time points were analysed instead of 12 h. Longer blotting periods or longer hybridization regimes gave the same results when duplicate samples were employed.

The oligonucleotide probes were designed so that each probe (i.e. LCMV N plus p10, or polyhedrin plus p10) would have similar specific activities for each gene sequence it contained. It is reasonable to assume, therefore, that the amount of label which hybridized to the respective mRNA species in each Northern analysis corresponded to the relative molar proportions of these mRNAs in the extracts. As mentioned, longer blotting or hybridization regimes did not affect the results. Kinetic graphs of the amounts of polyhedrin of LCMV N mRNA synthesis relative to the AcNPV p10 mRNA levels were produced for each recombinant virus (Fig. 3). The results shown represent the average of three independent analyses. Although there was some variation between the analyses (see Fig. 3), it was observed that the ratio of polyhedrin to p10 mRNA increased from 18 h p.i. In molar terms; the amounts of polyhedrin mRNA were between one and a half times to twice as high as the p10 mRNA by 24 to 30 h p.i. YON accumulated the highest amount of LCMV N mRNA, the level of which, at all times was greater than the p10 mRNA level. By contrast, the accumulation of LCMV N mRNA by the Ac373N, RP6N and RP5N recombinants corresponded to some 20 to 50% of the p10 mRNA levels at any time p.i. Surprisingly, in view of the major differences in N protein expression (Fig. 4), there were no substantial differences in N mRNA accumulation relative to the p10 mRNA levels between these three viruses.

In this study, we have obtained data that confirm that the complete 5' non-coding sequence of the polyhedrin gene is required for high level expression of the LCMV N gene. Analyses of the N protein and accumulated mRNA species agree with this conclusion. The data from the Northern analyses demonstrate that the p10, N and polyhedrin mRNA amounts peaked between 24 and 36 h p.i. (Fig. 2b). When the same protocol was employed, little variation was observed in the relative levels of the mRNA species between experiments (see the ranges indicated in Fig. 3). Other protocols (e.g.
Fig. 4. Expression of LCMV N protein by recombinant baculoviruses. *S. frugiperda* cells were infected with the recombinants YON (lane 3), Ac373N (lane 4), RP6N (lane 5) or RP5N (lane 6) at an m.o.i. of 10 p.f.u./cell and incubation at 28 °C for 48 h. Lysates of the infected or uninfected (lane 2) insect cells were electrophoresed in a 10 to 30% gradient polyacrylamide Laemmli gel and stained with Coomassie brilliant blue. The positions of the LCMV N and AcNPV gp64 and p10 proteins are indicated. Lane 1, molecular weight markers.

Different m.o.i. were not examined and studies at the later times of infection (> 42 h p.i.) were not undertaken owing to cell death and lysis in the infected cell population. The sizes of the p10 (0.6 kb), polyhedrin (1.2 kb) and N (2.1 kb) mRNA species recovered from YON or AcNPV infections were those expected if the mRNAs terminated at the usual AcNPV transcription termination sites. For Ac373N, RP6N and RP5N, the sizes of the N-specific mRNAs were smaller than the expected 4.3 kb size (3-4 kb of LCMV plus 0.9 kb of polyhedrin coding and non-coding sequences; see Matsuura et al., 1987). The N-short transcript predominated (1.8 kb). Based on the known sequence of the LCMV S cDNA incorporated into the recombinants, it is likely that the mRNA terminated in the vicinity of the LCMV S intergenic region, a region which has the capability to form a hairpin of 21 bp of inverted complementary sequence (LCMV S residues 1748 to 1793; Romanowski & Bishop, 1985). [This region is involved in LCMV mRNA transcription termination by the LCMV RNA polymerase (D. H. L. Bishop & V. C Emery, unpublished data.)] If this hairpin causes termination in the AcNPV system, then the polymerase must recognize that particular sequence (product or template DNA) and stop mRNA synthesis, following which the mRNA is processed and polyadenylated. It will be of interest to determine precisely where the N-short transcript terminates. The N-long transcript has an estimated size of 3 kb, i.e. it is also shorter than the expected 4-3 kb product. In fact no 4-3 kb transcript was identified even when the blots were overexposed. Based on its size, the N-long transcript probably terminates within the GPC coding region of these recombinants. Heterologous transcription termination signals have previously been shown to be effective for polyhedrin expression vectors (Possee & Howard, 1987). Again, direct analyses of the 3' end sequences of the N-long mRNA species are required to confirm the site of polyadenylation. Although the possibility that cryptic promoters may exist in one or other strand of the LCMV sequence cannot be totally discounted, strand-specific riboprobes complementary to mRNA identified the same sizes and relative amounts of RNA species as the nick-translated probes (data not shown). This and the fact that expression levels were the same for the RP6N recombinant and an identical recombinant from which the superfluous 3' (GPC) sequences had been removed (Matsuura et al., 1987) argue against the possibility that the identified shorter mRNA species represent either negative-sense RNA species, or mRNA species initiated downstream of the known polyhedrin promoter.

The analyses of the relative amounts of the N and p10 mRNAs assume that the lack of polyhedrin had no effect on p10 synthesis. This has not been directly investigated. Protein analyses (Fig. 4) did not detect any major differences between the expression levels of p10 and gp64 for the various recombinants in comparison to their levels in AcNPV-infected cells (not shown). However, it will be of interest to determine whether in the absence of the polyhedrin gene and its promoter, the p10 gene transcription and expression rates are altered relative to other viral genes (i.e. due to lack of competition from the polyhedrin site). On the basis of the above assumptions, the data may be interpreted to indicate that for foreign gene expression the complete polyhedrin leader sequence is required to obtain mRNA transcription levels comparable to those of polyhedrin mRNA in normal AcNPV infections. Whether changes within the −8 to +1 sequence (deletions or substitutions) affect transcription is not known, but can be examined by the types of mRNA and protein analyses reported here.

Most interesting is the evidence that the N mRNA levels relative to p10 levels for the recombinants Ac373N, RP6N and RP5N were essentially similar and some four- to six-fold lower than the YON levels. By comparison, and as assessed from gel scans, the N protein expression levels were YON > Ac373N > RP6N > RP5N, with up to four times more N produced by
YON than Ac373N, which in turn gave three times more N than RP6N which gave over four times more N than RP5N (data not shown; see Matsuura et al., 1987). These results suggest that the sequences between -32 and -8 have some role in the overall expression levels from the available mRNA species (e.g. a translational effect). Of particular interest is the observation that Ac373N made available mRNA species (e.g. a translational effect). Of results suggest that the sequences between -32 and -8 nucleotide and the BamHI linker (CGGATCC) in pAc373 have some effect on the translational efficiency of the recombinant gene. Notwithstanding this, the recombinant with the complete leader sequence gave both higher N protein and mRNA expression levels than did pAc373. Although it is possible that the longer 3' non-coding sequences in the mRNAs of the Ac373N, RP6N and RP5N recombinants may affect the levels of mRNA transcription (relative to YON), the expression of N from an RP6-derived recombinant from which the same 3' sequences were removed was unaltered by comparison with RP6N (Matsuura et al., 1987).

We thank Dr Robert D. Possee for providing pAcE1-P. The study was supported in part by contract DAMD17-87-C-7069.

References


(Received 27 November 1990; Accepted 1 July 1991)