Mapping of the Gene Coding for the Major Late Structural Polypeptide on the Frog Virus 3 Genome

By T. N. THAM, J. M. MESNARD,† L. TONDRE AND A. M. AUBERTIN*

Groupe de Recherches de l'INSERM (U 74) sur la Pathogénie des Infections Virales et Laboratoire de Virologie de la Faculté de Médecine de l'Université Louis Pasteur, 3 rue Koeberlé, 67000 Strasbourg, France

(Accepted 14 October 1985)

SUMMARY

The gene encoding the major capsid polypeptide (MCP 48) of frog virus 3 (FV 3) has been mapped on the viral DNA. Late FV 3 messenger RNA, hybrid-selected by the SalI-F fragment or a subset of these sequences, BamHI-L and -W fragments, directed the synthesis in vitro of a 48,000 mol. wt. (48K) polypeptide. This product was recognized by monospecific antibodies raised against the major capsid polypeptide. The RNA complementary to these DNA sequences was about 1350 nucleotides in size. This transcript, encoding MCP 48, was precisely located; S1 nuclease analysis indicated that its 5' end mapped at 1250 nucleotides to the right and its 3' end at 160 nucleotides to the left of the BamHI site at the junction between the BamHI-W and -L fragments.

INTRODUCTION

Frog virus 3 (FV 3), a virus belonging to the family Iridoviridae (Matthews, 1982), has a linear, double-stranded DNA genome approximately 107 kilobase pairs in length (Lee & Willis, 1983). The genome is highly methylated (Willis & Granoff, 1980) and organized into molecules which are circularly permuted and terminally redundant (Goorha & Murti, 1982; Murti et al., 1982). The initial round of DNA synthesis takes place in the host nucleus (Goorha et al., 1978; Martin et al., 1984); after expression of the late genes, replication occurs in the cytoplasm (Martin et al., 1984) where the concatemers are processed and encapsidated into mature particles (Goorha, 1982). Many of these characteristics are unique for a DNA virus multiplying in eukaryotic cells. Another peculiar feature is that, although host RNA polymerase II is involved in its transcription (Goorha, 1981), the purified viral DNA is not infectious (Willis et al., 1979). During the course of infection the genes are sequentially expressed and coordinately regulated (Willis et al., 1977; Elliott & Kelly, 1980; Elliott et al., 1980) but the elements allowing the expression of the different classes of genes to be controlled remain unknown. Only one gene coding for an immediate early RNA has been located and sequenced (Willis et al., 1984). As a prelude to identifying the structural sequences involved in this regulation, the mapping of a FV 3 late gene encoding a 48,000 mol. wt. (48K) polypeptide is described here. This protein is additionally interesting insofar as it is one of the few FV 3 polypeptides to which biological functions can be ascribed. Polypeptide 48K is the major capsid protein of the morphological subunits (Tripier-Darcy et al., 1982) and antibodies directed against it neutralize the infectivity of the virion in a complement-dependent fashion (Aubertin et al., 1981). In solution, this protein interacts with cells producing, in the presence of polycations, a stimulation in the influx of Na+ and K+ ions (Descamps, 1980).

METHODS

Cells and virus. FV 3 was produced and purified as described (Aubertin et al., 1980) and the viral structural proteins were solubilized as previously reported (Aubertin et al., 1973, 1977). Chinese hamster ovary cells (CHO,

† Present address: Laboratoire de Virologie, Institut de Biologie Moléculaire et Cellulaire du C.N.R.S., 15 rue Descartes, 67084 Strasbourg, France.
proline auxotroph) were grown in alpha medium supplemented with 10% newborn calf serum. They were infected at a multiplicity of 40 p.f.u./cell.

**Bacterial strains and plasmids.** The plasmid pUC8 (Vieira & Messing, 1982) was used to subclone the BamHI fragments of the FV 3 SalI-F fragment originally cloned in pAT153. The plasmid and the recombinants were propagated in *Escherichia coli* strain JM103 (Messing et al., 1981). Growth of bacteria, purification of plasmids and all recombinant DNA methodology were as summarized by Maniatis et al. (1982).

**RNA production, hybridization selection and in vitro translation.** Total cellular RNA was isolated at 8 h after infection (late viral RNAs) and was purified as previously described (Glisin et al., 1974). RNA was selected by hybridization to recombinant plasmid DNA carrying FV 3 restriction fragments as described (Belle Isle et al., 1981) except that hybridizations were performed at 47 °C. The selected RNA was translated in the presence of [35S]methionine in a message-dependent rabbit reticulocyte lysate (purchased from Amersham) according to the conditions described by the manufacturer. The products were analysed by SDS–PAGE as described (Aubertin et al., 1980).

**Immunoprecipitation.** A monospecific antibody directed against the major structural polypeptide (48K) was used. Isolation of the antigen by two-dimensional immunoelectrophoresis and the production of antibodies have been described (Aubertin et al., 1981). The labelled proteins were first diluted in 150 μl buffer A (0-5% NP40, 150 mM-NaCl, 5 mM-EDTA, 25 mM-Tris–HCl pH 7-4) containing 15 μl antisemur or control serum. In the case of polypeptides made *in vitro*, 10 μl of the translation lysate was diluted with 125 μl of unlabelled soluble proteins obtained by ultrasonic dissociation of infected cells suspended in buffer A (106 cells/ml) followed by an elimination of the insoluble material by centrifugation (100000 g, 1 h). Proteins were immunocomplexed by the addition of 15 μl antiserum. After incubation at room temperature for 3 h, 15 μl *Staphylococcus aureus* cell suspension (100 mg/ml) was added to the immunocomplex bound to the cells was collected by centrifugation and washed three times with 1% sodium deoxycholate, 0.5% NP40, 500 mM-NaCl, 1 mM-EDTA, 20 mM-Tris–HCl pH 7-4 and once with 10 mM-NaCl, 10 mM-Tris–HCl pH 7-4. Finally, the pellet was resuspended in 2% 2-mercaptoethanol, 10% glycerol, 0-003% bromophenol blue, 62.5 mM-Tris–HCl pH 6-8 and heated at 100°C for 2 min to dissociate the immunoprecipitate. The supernatant, recovered after centrifugation, was analysed by SDS–PAGE.

**Nuclease S1 analysis.** Nuclease S1 analysis (Berk & Sharp, 1977) was carried out as previously described (Willis et al., 1984). One-hundred ng of suitably labelled probe was mixed with 25 μg of the total cell RNA. After hybridization and nuclease S1 treatment, S1 nuclease-resistant products were sized by electrophoresis on alkaline agarose gels.

**Northern hybridizations.** Samples of total late RNA were glyoxalated and electrophoresed through 1.4% agarose gels as described (McMaster & Carmichael, 1977). Gel-fractionated RNA was blotted onto nitrocellulose membranes (Thomas, 1980) and was probed with cloned restriction fragments labelled by nick translation (Rigby et al., 1977).

**RESULTS**

**Analysis of the specificity of the major capsid polypeptide antiserum**

To ascertain the specificity of the immune antiserum directed against the 48K protein, the major capsid polypeptide (MCP 48), 35S-labelled structural proteins solubilized from the virions were incubated with antibodies as described in Methods. The immunocomplex was isolated by binding to *S. aureus* cells, then re-solubilized and analysed by SDS–PAGE. MCP 48 was the only prominent polypeptide precipitated from the viral protein solution (Fig. 1, lane 4). As a control, incubation with an antiserum raised against the structural polypeptide 44K led to the precipitation of the 44K protein (Fig. 1, lane 3).

**Recognition of a translation product from infected cell mRNAs by MCP 48 antiserum**

A combination of hybridization selection, *in vitro* translation and immunoprecipitation was used to map the gene coding for the MCP 48.

In a series of experiments to be described elsewhere, total RNA from FV 3-infected cells isolated at 8 h after infection (late RNA) was purified by hybridization to cloned SalI and HindIII restriction fragments immobilized onto nitrocellulose membranes. Fragments were chosen so that all the unique sequences of the genome were included in the analysis. Hybrid-selected RNA was translated *in vitro* and the products were characterized by SDS–PAGE. Only late RNA selected by a single recombinant DNA containing the SalI-F fragment directed the synthesis of a 48K polypeptide (Fig. 2). The location of this SalI-F fragment relative to the map of the XbaI sites (Lee & Willis, 1983) on the genome was obtained by double digestion and hybridization of 32P-labelled SalI-F recombinant to a Southern blot of an XbaI digest of the
Mapping of FV 3 capsid protein gene

Fig. 1. Characterization of the major capsid polypeptide 48K (MCP 48) antiserum. The $^{35}$S-labelled solubilized virus structural proteins (lane 2) were immunoprecipitated by the structural 44K polypeptide antiserum (lane 3) or by the MCP 48 antiserum (lane 4) and analysed by SDS-PAGE. The $^{35}$S-labelled FV 3 structural proteins (lane 1) were used as size markers (mol. wt. \times 10^{-3}).

FV 3 DNA. SalI-F partially overlaps XbaI-J and -F fragments, 900 nucleotides of the left side being contained in XbaI-J.

For finer mapping of the gene encoding this 48K protein, cleavage sites generated by BamHI were located. For this, the SalI-F fragment excised from the recombinant was hydrolysed with BamHI and the fragments obtained were ligated to BamHI- or to BamHI- and SalI-digested plasmid pUC8 and cloned. These cloned fragments were used to generate the BamHI map (Fig. 2b) and to select messenger RNA as described above. BamHI-W and BamHI-L hybrid-selected the mRNA encoding the 48K polypeptide as defined by in vitro translation (Fig. 2). The molecule of approximately 46K is an endogenous in vitro product. This band was found also in the sample not containing mRNA (not shown). Its intensity varied with the batch of lysate as also did the intensity of polypeptides smaller than 48K. The minor species were not detected for example in the experiment with SalI-F. The 48K polypeptide synthesized in vitro was immunoprecipitated by the MCP 48 antiserum but did not react with the normal serum (Fig. 3). The gene encoding MCP 48 therefore spans the BamHI site between fragments W and L.

Mapping of the transcript by nuclease S1 analysis

To map the gene more precisely and also to determine the direction of transcription of the mRNA, nuclease S1 analysis of RNA-DNA duplexes was performed.

BamHI-W and -L fragments labelled at the 3' or 5' ends were used as hybridization probes.
T. N. THAM AND OTHERS

The end-labelled probes were hybridized to late RNA isolated from FV 3-infected cells. The residual single-stranded DNA was hydrolysed with nuclease S1 and the resistant DNA was separated by electrophoresis on an alkaline 1.5% agarose gel. With the BamHI-W fragment labelled at its 5' termini, no nuclease-resistant products were detected (Fig. 4b). The 3' end-labelled BamHI-W fragment was protected by late RNA to produce a single fragment which was 160 nucleotides in size (Fig. 4c). As the RNA identified by hybridization to the 3' end-labelled probe was the only one to be detected with the BamHI-W fragment, it most likely codes for the MCP 48 and its transcription ends in the BamHI-W fragment.
Mapping of FV 3 capsid protein gene

Fig. 3. Immunoprecipitation of in vitro translation products of RNA selected by the SalI-F sequences. Late RNAs purified by hybridization to individual BamHI-N, -W, -L or -S fragments (a to d respectively) were translated in vitro and the polypeptides were analysed by SDS-PAGE either directly (lane 1) or after immunoprecipitation by MCP 48 antiserum (lane 2) or with normal serum (lane 3). Lane 4 contains the polypeptides synthesized in the reticulocyte lysate without the addition of RNA. 35S-labelled FV 3 structural proteins were used as size markers (lane 5, mol. wt. × 10^-3).

This was further borne out by the results obtained with the BamHI-L fragment. As shown in Fig. 4(b, c), two nuclease-resistant fragments were generated after hybridization of late RNA with the BamHI-L fragment, one with the 3’ end-labelled fragment and another with the 5’ end-labelled one. The 1150-nucleotide band obtained with the 3’ end-labelled probe was produced by an RNA of approximately 1400 nucleotides whose transcription begins in the BamHI-S fragment and ends in the BamHI-L fragment (J. M. Mesnard et al., unpublished results). The 5’ end of one late RNA which extends 1250 nucleotides into the BamHI-L fragment was detected with the 5’ end-labelled probe.

To confirm the orientation of the transcript, the BamHI-L fragment was cleaved with HincII (Fig. 4d) and the resulting fragment of 2000 base pairs was labelled at the BamHI 5’ end. After hybridization with late RNA and digestion with nuclease S1, no resistant DNA fragments were observed (results not shown). This means that the end of this transcript mapped at 1250 nucleotides to the right of the BamHI site at the junction between the BamHI-W and -L fragments, the message being transcribed from the leftward reading strand. The size of the transcript was determined by Northern blot analysis. Total late RNAs were, after glyoxylation, fractionated by electrophoresis in an agarose gel, transferred to a nitrocellulose membrane and visualized by hybridization with nick-translated BamHI-W or -L fragments (Fig. 4a). As only one abundant transcript of about 1350 nucleotides was common to these two fragments, we concluded that this RNA encoded MCP 48; its transcription began in the BamHI-L fragment and ended in the BamHI-W fragment, as shown in the diagram (Fig. 4).
Fig. 4. Fluorograph of electrophoretically separated nuclease S1-resistant DNA fragments. Mapping of the 5' termini and 3' termini of late RNA by hybridization to BamHI-L and -W DNA fragments labelled at the 5' position (b) or at the 3' end (c) was followed by nuclease S1 treatment. End-labelled BamHI-W and -L fragments (lanes 1 and 3) and resistant DNA fragments (lanes 2 and 4) were resolved by electrophoresis on a 1.5% alkaline agarose gel. (a) The late RNA detected by Northern blotting with BamHI-W (lane 1) or BamHI-L (lane 2) fragments labelled by nick translation. The sizes of the fragments (kb) were determined by comparison with the mobilities of end-labelled λ HindIII fragments (scale on the left). The map position and transcription direction of the late RNA encoding MCP 48 is indicated in (d).

DISCUSSION

The identification of the regions coding for the different classes of proteins is essential for unravelling the elements controlling the temporal expression of FV 3 genes. The MCP 48 gene was chosen for several reasons which simplify the analysis involving hybrid selection of mRNAs and in vitro translation. Firstly, the relative molar amount of this polypeptide synthesized late in infection was more than tenfold higher than that of any other polypeptide (Martin et al., 1984; Willis et al., 1977) and one RNA that could potentially code for the major capsid protein was found to be the most abundant message (Willis et al., 1977). Secondly, this polypeptide is not phosphorylated (Aubertin et al., 1980), phosphorylation being the sole post-translational modification found among FV 3 proteins (Elliott & Kelly, 1980). Thus, a comparison of in vitro and in vivo translation products with structural polypeptides on the basis of their electrophoretic mobility was possible. Thirdly, among in vitro products, MCP 48 can be unambiguously characterized by means of immunoprecipitation with monospecific antibodies. This polypeptide is predominant in the particle and represents 90% of the proteins solubilized from the virions (Fig. 1, lane 4). Purification of MCP 48 in quantities sufficient to produce a monospecific immune serum was possible, as had seemed likely. This approach involving immunological characterization was necessary as late messages are poorly translated in cell-free protein-synthesizing extracts (Raghow & Granoff, 1983). The abundance of the late proteins compared to the early proteins produced in vitro differed significantly from what was found in infected cells and this was an obstacle to the mapping of some polypeptides.
Mapping of FV 3 capsid protein gene

By combining hybridization selection, in vitro translation and immunoprecipitation, we were able to map the gene encoding the MCP 48 on the BamHI-L and -W fragments. As shown in Fig. 3, one polypeptide of 48K is recognized by the monospecific antibodies directed against the major capsid protein (48K). Polypeptides of sizes smaller than 48K share antigenic determinants with this polypeptide as they were precipitated by the antibody directed against MCP 48. The 44K polypeptide is not related to the 44K structural protein as determined by a similar experiment carried out with the 44K antiserum; the gene encoding this protein has been located in the fragment SalI-C (unpublished data). Thus, the smaller polypeptides may result from a premature termination of translation of the 48K mRNA. Although partial degradation of mRNA could also lead to similar translation patterns, this hypothesis can be excluded; incomplete products should be observed for each message, particularly for the mRNA encoding polypeptide 70K which is even larger. Moreover, translation patterns were identical for mRNA selected by both BamHI-L and -W fragments: if truncated RNAs were present, they would hybridize poorly if at all to the BamHI-W fragment which contains only 160 nucleotides complementary to the 3' end of the MCP 48 mRNA. An alternative explanation is the involvement of separate AUG codons in the initiation of translation as has been shown for herpes simplex virus thymidine kinase mRNA (Marsden et al., 1983).

Finer localization of the gene was accomplished by nuclease S1 analysis. Transcription was found to be initiated within the BamHI-L fragment at 1250 nucleotides to the right of the BamHI site at the junction between the BamHI-W and -L fragments and terminated within BamHI-W. The sum of the sequences transcribed from BamHI-L and -W (1410 nucleotides) is close to the size determined for the putative message [FV 3 RNA lacks poly(A) tracts (Willis et al., 1977)], an observation which implies that the transcript is not extensively spliced. This information is sufficient to encode a polypeptide of 50K.

In addition, our data indicate that another gene, transcribed from the same DNA strand and expressed at early and late times after infection, is located very close to, if not overlapping, the MCP 48 gene or its promoter. Sequencing this region will help in clarifying the organization of the transcripts and allow a comparison of the putative promoter sequence of this late gene, the first to be mapped, with other FV 3 genes.

We are grateful to Professor A. Kirn for constant interest and encouragement during this study.

REFERENCES


(Received 5 August 1985)