Transcription of the murine gammaherpesvirus 68 ORF73 from promoters in the viral terminal repeats

Heather M. Coleman, Stacey Efstathiou and Philip G. Stevenson

Gammaherpesviruses persist as latent episomes in a dynamic lymphocyte pool. The regulated production of an episome maintenance protein is therefore crucial to their survival. The transcription initiation site of the murine gammaherpesvirus 68 episome maintenance protein, ORF73, was mapped to the viral terminal repeats, more than 10 kb distant from the open reading frame (ORF) itself. A 5’ non-coding exon in the terminal repeats was spliced to the right end of the viral unique sequence, and then across ORFs 75a, 75b, 75c and 74 to ORF73. The right-hand portion of a single repeat unit was sufficient for constitutive promoter activity. The unique left end of the viral genome further enhanced ORF73 transcription. This, together with the large size of the predominant ORF73 mRNA, suggested that transcription initiates in distal repeat units and then splices between repeats to generate an extensive 5’ untranslated region. A second promoter in the left-hand portion of the proximal terminal repeat unit generated a transcript which overlapped that of ORF73, but failed to splice to the ORF73 coding exon and so transcribed ORF75a. In distal repeat copies, however, transcription from this promoter would enter the next repeat unit to become an ORF73 mRNA. There was a third promoter just upstream of ORF73 itself. These data indicate that ORF73 transcription is highly complex, and support the idea that the terminal repeats of gamma-2-herpesviruses constitute a vital component of episomal persistence.

INTRODUCTION

Gammaherpesviruses drive lymphocyte proliferation during acute infection and persist in an intermittently replicating memory lymphocyte pool (Babcock et al., 1998). They therefore have a fundamental requirement for episome maintenance: the replication of latent viral episomes and their segregation between daughter cells at mitosis. Viral episome maintenance proteins constitute a potential immune target (Lee et al., 2004; Tellam et al., 2004), so their expression has to be tightly regulated. The Epstein–Barr virus (EBV) episome maintenance protein, EBNA-1, (Yates et al., 1985) is co-expressed with other EBNAs from the Wp and Cp promoters in latency III (Rogers et al., 1990; Sample & Kieff, 1990), and independently from the Qp promoter in latency I (Nonkwelo et al., 1996; Schaefer et al., 1995; Tsai et al., 1995). EBNA-1 autoregulates the Qp promoter (Sample et al., 1992) and its own translation (Yin et al., 2003). The Qp promoter is also cell cycle-dependent (Davenport & Pagano, 1999). Thus although EBNA-1 is expressed in all latently infected cells in vitro (Rowe et al., 1987), it is often absent from latently infected cells in vivo (Thorley-Lawson, 2001).

How episome maintenance functions in vivo is germane to both combating gammaherpesvirus disease and developing episomal vectors for gene delivery. However, such questions are difficult to address with human gammaherpesviruses due to their strict species tropisms. Consequently, the murine gammaherpesvirus 68 (MHV-68) has become an important experimental tool for understanding in vivo gammaherpesvirus gene functions. Like EBV, MHV-68 persists in memory B cells (Flano et al., 2002; Marques et al., 2003; Willer & Speck, 2003). The episome maintenance proteins of MHV-68 and the Kaposi’s sarcoma-associated herpesvirus (KSHV) are encoded by ORF73 (Ballestas et al., 1999; Rainbow et al., 1997). The MHV-68 ORF73 is transcribed in latency both in vitro and in vivo (Marques et al., 2003; Martinez-Guzman et al., 2003; Virgin et al., 1999), and ORF73-deficient MHV-68 shows a profound latency deficit in vivo (Fowler et al., 2003; Moorman et al., 2003). The task now with MHV-68 is to define the molecular details that underlie in vivo ORF73 function. For this we must identify how ORF73 production is controlled and how this control relates to its function.

Latent KSHV in B cell tumour lines transcribes ORF73, ORF72 and ORF71 from a common promoter just upstream of ORF73 (Dittmer et al., 1998; Rainbow et al., 1997; Sarid et al., 1999; Talbot et al., 1999). In contrast, MHV-68 has no ORF71, its ORF16 homologue is sited between ORF72 and ORF73 (Virgin et al., 1997), and there are consensus polyadenylation signals just 3’ of ORF73. In vivo analysis
indicates that ORF73 transcription is largely distinct from that of ORF72 (Virgin et al., 1999). Rightwards of ORF73, MHV-68 lacks a K14 homologue, may not have a K15 homologue, and has three copies of ORF75 rather than one. Otherwise, MHV-68 appears fairly similar to KSHV.

Surprisingly, analyses of MHV-68 latent gene transcription have identified loci rightwards of ORF73 that are not transcribed in KSHV latency, including ORF74 in latently infected peritoneal exudate cells (Virgin et al., 1999) and ORF74, ORF75a and ORF75c in latently infected S11E cells (Martinez-Guzman et al., 2003). The S11E cell line contains at least one integrated viral genome (Husain et al., 1999) and may therefore show abnormal patterns of transcription, but ORF74, ORF75a, ORF75b and ORF75c transcripts in 3T3 cells are resistant to inhibition by cycloheximide (Martinez-Guzman et al., 2003), consistent with latency-associated promoter activity. These results have raised the possibility that MHV-68 latency gene expression is fundamentally different to that of KSHV. However, the splicing patterns of MHV-68 latency transcripts remain largely undefined, and without knowing the structure of each gene it is difficult to interpret definitively the detection of small mRNA segments. In order to understand more about MHV-68 latency, we have mapped the 5' ends of the major ORF73 transcripts and identified the functional cores of the associated promoters.

**METHODS**

**Cells and viruses.** 3T3 cells were maintained in Dulbecco’s modified Eagle’s medium supplemented with penicillin (100 U ml\(^{-1}\)), streptomycin (100 µg ml\(^{-1}\)), l-glutamine (2 mM) and 10% fetal bovine serum (PAA laboratories). MHV-68 was grown and titrated in BHK-21 cells (Bridgeinan et al., 2001).

**Promoter constructs and CAT ELISA.** The MHV-68 terminal repeats were cloned from the MHV-68 BAC with BglII (117840) and PstI, which cuts at the left end of the BAC cassette, into the BglII and EcoRV sites of pSV40-ZEO2 (Invirogen). This plasmid was then digested with PstI and ligated back to itself, leaving one intact repeat unit. This was subcloned as a PstI/Xhol fragment into the Xhol/SalI sites of pCAT-enhancer (Promega), inserting the CAT gene in ORF73 exon 2 with a single upstream terminal repeat unit (118210–119450). We generated a 5' truncation of this construct by digestion with PstI (118700) and HincIII, which cuts in pCAT-enhancer, blunting with T4 DNA polymerase (New England Biolabs) and self-religation of the plasmid.

CAT was joined to ORF73 exon 1 by first PCR amplifying genomic DNA, co-ordinates 118616–118719, including Xhol and HincIII sites in the respective downstream and upstream primers, and ligating this into the Xhol/HincIII sites of pCAT-enhancer (Promega). A 1-2 kb PstI genomic clone (Estasthiou et al., 1990), comprising the 5' 750 bp of one terminal repeat unit and the 3' 500 bp of the next, was then cloned into this plasmid at the 118700 PstI site, upstream of CAT. This construct was truncated at the 5' end by digestion with HincIII and either PstI (118700), SacI (119078), BsaI (119087), NotI (119105), AscI (119151), SfiI (119348), XcmI (119359) or BbvCI (118616, cuts in the upstream terminal repeat copy). We used MspAI (118720) to truncate at the 3' end of the SacI clone and PstI (118700) plus Xhol to truncate at the 3' end of the BbvCI clone. We also used BAL-31 exonuclease to digest from the 3' end of the promoter construct after linearization with HindIII. The cut ends were blunted with T4 DNA polymerase and ligated together. These constructs are summarized in Fig. 6(a).

The promoter upstream of ORF75a (genomic co-ordinates 118117–118598) was amplified by PCR, including an Xhol site in each primer, and cloned in either orientation into the Xhol site of pCAT-enhancer. Genomic co-ordinates 104871–105430, just upstream of ORF73 were amplified by PCR with a PstI restriction site in each primer, and cloned in either orientation into the PstI sites of pCAT-enhancer. All constructs were checked by DNA sequence analysis. Promoter activity was tested by transfection (3 µg plasmid per 10\(^5\) 3T3 cells) using Eugene-6 (Roche Diagnostics), followed 40–48 h later by CAT ELISA (Roche Diagnostics). In some experiments, the cells were co-transfected with either pCDNA3-ORF50 (Stevenson et al., 2000), pCDNA3-ORF73 (genomic co-ordinates 104868–103924, PCR-cloned into the EcoRI/Xhol sites of pCDNA3) or empty pCDNA3 vector (Invitrogen). All transfection experiments were performed at least three times. In at least one replicate experiment, equal transfection efficiency was checked by co-transfection of a GFP expression plasmid. For significant negative results we tested at least two independent plasmid preparations.

**Rapid amplification of cDNA ends (RACE).** RNA was extracted from MHV-68-infected 3T3 cells (5 p.f.u. per cell, 18 h) by using RNazol B (Tel-Test). For 5' RACE of ORF73 (Roche Diagnostics), total RNA was reverse transcribed from a primer corresponding to genomic co-ordinates 104638–104659. A 5' poly(A) tail was added with terminal transferase, and the product was amplified by PCR, using a 5' poly(A)-specific primer and a 3' ORF73 primer (genomic co-ordinates 104699–104717). The products were gel-purified (QIAquick gel extraction kit; Qiagen), sequenced directly, and T/A-cloned into pGEM-T-easy. For 3' RACE of ORF73, we used a poly(A)-specific RACE kit primer for reverse transcription, which incorporated a 3' tail for subsequent PCR amplification. RACE products were amplified by using a tail-specific primer and sequence-specific primers, matching genomic co-ordinates 104503–104483, 104383–104363, 104263–104243 or 104050–104030. For 5' RACE of ORF75a (genomic co-ordinates 117904–114032), the reverse transcription primer matched 117720–117740 and the PCR amplification primer matched 117741–117760. PCR products were cloned into pSP73 by using the Clal site in the 5' poly(A) tail-specific primer and a PstI site at 117755.

For RT-PCR based confirmation of the RACE-mapped ORF73 transcript, we paired a 104699–104717 ORF73 primer with upstream primers corresponding to genomic co-ordinates 117904–117920, 118100–118117, 118182–118199 or 118616–118632. A positive control for these primers was a BglII/PstI genomic fragment (117840–118700) cloned into the NruI/PstI sites of pCDNA3-ORF73, i.e. an 860 bp genomic fragment covering the right-hand end of the viral unique sequence, 32 bp upstream of ORF73.

**Northern blotting.** RNA was purified from uninfected or MHV-68-infected (5 p.f.u. per cell, 18 h) 3T3 cells using RNAzol-B, electrophoresed (20 µg per lane) on a 0.8% formaldehyde agarose gel and blotted overnight unto uncharged nylon membranes (Roche Diagnostics). Probe templates for β-actin and ORF73 (a BsaI/Bsu36I genomic fragment, co-ordinates 103126–104829) were random-prime labelled (Qbiogene) with \(^{32}P\)-dCTP (AP Biotech). Blots were hybridized with probe overnight at 45 °C in 50% formamide, 5× Denhardt’s solution, 6× SSC, 0.1% SDS, 100 µg sonicated salmon sperm DNA ml\(^{-1}\), followed by washing (0.1× SSC, 0.1% SDS, 65 °C) and exposure to X-ray film (Coleson et al., 2003).

**RNase protection assay.** We generated four probes. A Syl genomic fragment (104714–105433) was cloned into the EcoRV site of pSP72 (Promega), linearized with XcmI (105088), and transcribed...
with T7 RNA polymerase (Ambion). A BglII/PstI genomic fragment (117840–118700) was cloned into the BglII/PstI sites of pSP72, linearized with BseYl (118292) and transcribed with T7 RNA polymerase. An ORF73 RACE clone (genomic co-ordinates 104699–104871, 118055–118160, 118605–118658) in pGEM-T-easy was linearized with SalI and transcribed with T7 RNA polymerase to give a 514 nt probe (60 nt of plasmid sequence, 121 nt of poly(A) tail added during RACE, and 333 nt of MHV-68 exons). An SfoI/ PstI genomic fragment (118552–118700) cloned into pSP72 was linearized with XhoI, 28 bp downstream of the PstI cloning site, and transcribed with T7 RNA polymerase. Each anti-sense riboprobe was labelled with 32P-UTP (Amersham Biosciences) and hybridized to mRNA from MHV-68-infected 3T3 cells (5 p.f.u. per cell, 6 h) by using the Direct Protect lysate RPA kit (Ambion). Following RNase A/RNase T1 treatment, protected fragments were purified, separated on 8 M urea/6 % polyacrylamide gels, and exposed to X-ray film.

RESULTS

RACE mapping of the ORF73 transcription start site

A major difficulty with analysing MHV-68 gene expression is that the virus does not transform primary cells (Dutia et al., 1999; Stevenson & Doherty, 1999). MHV-68 also infects lymphocyte cell lines rather inefficiently and with a high rate of spontaneous reactivation (Sunil-Chandra et al., 1993). Lymphocyte tumour lines are unlikely to show normal viral gene expression. However, although MHV-68 persists mainly in memory B cells (Flano et al., 2002; Marques et al., 2003; Willer & Speck, 2003), it also establishes latency in epithelial cells (Stewart et al., 1998). Latency transcripts such as M2 and ORF73 are detectable in infected fibroblasts (Marques et al., 2003), and infection is far from uniformly lytic (Martinez-Guzman et al., 2003; May et al., 2004). We therefore used MHV-68-infected 3T3 cells as a basis for 5’ RACE-mapping of the ORF73 transcription start site (Fig. 1a).

DNA sequence analysis of the major 360 bp product and of three clones derived from it identified a transcription start site at genomic co-ordinates 118640–118658, within the viral terminal repeats (118210–119450) (Fig. 1b, c). The start site at genomic co-ordinates 118640–118658, within DNA sequence analysis of the major 360 bp product and transcription start site (Fig. 1a).

May et al. established latency in epithelial cells (Stewart et al., 1998). Thus, in vivo latent ORF73 transcripts also initiated in the viral terminal repeats.

RT-PCR based confirmation of the RACE-mapped ORF73 mRNA

We confirmed the presence of the spliced ORF73 mRNA in different virus-infected cell populations by RT-PCR. A 3’ primer in the ORF73 coding sequence amplified a product from MHV-68-infected 3T3 cells with a 5’ primer sited in ORF73 exon 1 or exon 2, but not with a 5’ primer in either intron (Fig. 2a). Cycloheximide treatment (Fig. 2b) reduced the abundance of a thymidine kinase mRNA but not that of the spliced ORF73 mRNA, indicating that the latter had immediate-early kinetics. The spliced mRNA was also detectable in cDNA derived from lungs, spleens and purified germinal centre B cells of MHV-68-infected mice (Fig. 2c). Germinal centre B cells appear to support almost entirely latent infection (Stevensons et al., 2002). Thus, in vivo latent ORF73 transcripts also initiated in the viral terminal repeats.

RNase protection analysis of the ORF73 mRNA

We used RNase protection assays (Fig. 3) to provide a reverse transcriptase-independent and PCR-independent confirmation of the ORF73 RACE product. We first hybridized mRNA from MHV-68-infected 3T3 cells with a 32P-labelled anti-sense probe derived from a StyI/XmnI genomic clone (104714–105088) (Fig. 3b). The 160 nt protected fragment was consistent with use of the splice acceptor site at genomic co-ordinate 104871 (158 nt predicted fragment). The larger protected fragment corresponded to the full-length of the MHV-68 sequence in the probe (374 nt) and implied an additional, unspliced ORF73 mRNA – undetected by RACE – starting at least as far upstream as genomic co-ordinate 105080. Notably, there are no AUG codons between 105098 and the start of the ORF73 coding region.

We used a BglII/BseYI genomic clone (co-ordinates 117840–118272) to cover the ORF73 exon 2 (Fig. 3c). However, background probe protection obscured the position of the predicted 106 nt band. There was a 370 nt protected fragment, consistent with an mRNA starting upstream of the ORF73 exon 2 splice acceptor site (approximate genomic co-ordinate 118210) and continuing across its splice donor site to the BglII site at 117840. This mRNA would therefore encode ORF75a rather than ORF73. The
370 nt band was not diminished in abundance by cycloheximide treatment of the infected cells.

A probe derived from the longest ORF73 RACE clone (333 nt of MHV-68 sequence: 104699–104871, 118055–118160, 118605–118658) was fully protected (Fig. 3d), confirming the existence of the RACE-mapped mRNA in infected cells. There was also a 170 nt band, consistent with 3' probe protection by an unspliced ORF73 mRNA (Fig. 3b). A protected 5' end would have required an ORF75a rather than ORF73 mRNA and would have protected a 321 nt fragment of the Bgl II/Bse YI probe, of which we saw no sign (Fig. 3c).

We used an Sfo I/Pst I genomic clone (118552–118700) to look for evidence of mRNAs initiating upstream of genomic co-ordinate 118640 (Fig. 3e). Protected fragments were seen below 50 nt even with unhybridized probe, presumably because this G/C-rich RNA has considerable secondary structure (see also Fig. 8c). It was therefore not
possible to identify or to rule out a 40–50 nt band corresponding to exon 1 of the RACE clones. The 150 nt fragment corresponded to the full-length probe (148 nt of MHV-68 sequence). This probably reflected protection by G/C-rich, terminal repeat DNA, since cycloheximide treatment of the infected cells, which would prevent viral DNA replication, reduced its abundance. In contrast, a 90–95 nt protected fragment was more abundant in cycloheximide-treated cells, indicating protection by a latency transcript. This was presumably a 5' probe fragment, from the splice donor site at genomic co-ordinate 118605 to the PstI site at genomic co-ordinate 118700. 3' Probe protection would

Fig. 2. PCR-based confirmation of the RACE-mapped ORF73 mRNA. (a) A 3' primer in ORF73 exon 3 (genomic co-ordinates 104699–104717) was paired with 5' primers A, B, C or D as shown. As positive controls, we used either the ORF73 RACE product (primer D) or a PstI/BglII genomic fragment (117840–118700) joined to the ORF73 coding sequence (primers A, B and C), effectively a plasmid-engineered deletion of most of intron 2. Thus the positive control for primer B was 478 bp, while the predicted size for the RACE-mapped mRNA was 234 bp. (b) The spliced ORF73 mRNA was more readily detected in 3T3 cells after a 6 h infection in the presence of cycloheximide than after overnight infection without cycloheximide. The opposite was true of thymidine kinase mRNA. (c) PCR confirmed the presence of the spliced ORF73 mRNA in cDNA samples derived from infected mice. Latently infected spleen cells and CD19+PNAhi germinal centre B cells, flow cytometrically sorted (>95% pure) from spleen cells, were harvested 14 days after intranasal infection and pooled from five mice. Lungs were harvested from individual mice at 5 days post-infection – samples from three mice are shown. Amplification was with the 3' ORF73 primer and the 5' primer D, as in (b).
have required an mRNA that missed the 118605 splice donor site, and was therefore inconsistent with the RACE product sequencing and with direct RT-PCR (primer C, Fig. 2).

Protection of the 118605–118700 fragment suggested that ORF73 transcription initiated at least as far upstream as 118700. However, it was also possible that an mRNA initiating 3' or 5' of 118700 in an upstream repeat unit acquired the 118605–118700 segment from downstream repeat units. Significantly, a splice acceptor site at genomic co-ordinate 118695 was identified in two cDNA library clones derived from S11E cells (Husain et al., 1999). These transcribed a 91 nt exon, and used the 118605 splice donor, but were then fused to an immunoglobulin mRNA, indicating integration of the viral genome. Our data suggested that these cDNA clones were transcribed from an ORF73 promoter.

With both a splice acceptor and a splice donor in each terminal repeat, there was the potential for the 5' untranslated region of an ORF73 mRNA to accumulate multiple copies of the 118605–118695 exon. RACE would not have detected these, since internal PCR priming would have preferentially amplified from the most 3' terminal repeat copy. Northern blotting (Fig. 4) showed a predominant ORF73 mRNA at least 4000 nt in size, consistent with published data (Virgin et al., 1999). There was also a 2000 nt mRNA, consistent with an unspliced ORF73 mRNA initiating around genomic co-ordinate 105100 (see below). The 4000 nt mRNA was too large for this, and so was likely to be the RACE-mapped transcript.

With a poly(A) tail of approximately 300 nt, the RACE-mapped lower limit of ORF73 transcription initiation (118640) would generate an mRNA of approximately 1500 nt. Therefore some 2500 nt of the 4000 nt ORF73 mRNA was likely to be 5' untranslated region, corresponding to 25–30 copies of the 91 nt terminal repeat exon. There are at least 10 terminal repeat copies at each end of linear MHV-68 genomes (Efstathiou et al., 1990), implying that the predominant ORF73 mRNA start site was at least as far upstream as 118700 in an upstream repeat unit.

---

**Fig. 3.** RNase protection analysis of the ORF73 mRNA in MHV-68-infected cells. (a) 32P-labelled RNA probes were designed to cover the start of ORF73 (StyI probe), ORF73 exon 2 (BglII/BseYI probe), the RACE mapped mRNA itself and ORF73 exon 1 in the terminal repeats (SfoI probe). (b) RNA from uninfected (nil) or MHV-68-infected (6 h, m.o.i. = 5) 3T3 cells (inf) was hybridized with the StyI riboprobe and digested with RNase. Unless stated otherwise, the infected cells were treated with cycloheximide to confine the analysis to immediate-early MHV-68 mRNAs. Protected fragments are indicated by arrowheads. The smaller protected fragment corresponded to exon 3 (173 nt). The larger protected fragment (approx. 375 nt) matched the full-length of MHV-68 sequence in the probe, and implied the existence of an unspliced ORF73 mRNA starting at least as far 5' as genomic co-ordinate 105088. (c) As in (b), RNA samples were used to protect a labelled riboprobe, this time derived from a BglII/BseYI genomic fragment that spanned ORF73 exon 2. The smaller protected bands present in both uninfected and infected samples probably reflected probe secondary structure. They obscured the expected 106 nt band corresponding to ORF73 exon 2. The specific protected fragment of approximately 370 nt (arrowhead) indicated an mRNA initiating upstream of the exon 2 splice acceptor site and running beyond the exon 2 splice donor site. This mRNA therefore overlapped with ORF73 exon 2, but used a distinct promoter and transcribed ORF75a. (d) With a riboprobe derived from an ORF73 RACE clone, there was a protected fragment of approximately 330 nt that corresponded to protection of the full-length of spliced MHV-68 sequence in the probe (333 nt). Thus, RNase protection confirmed the existence of the RACE-mapped mRNA. The smaller protected fragment corresponded to protection of the 3' end of the probe up to the exon 3 splice acceptor site (173 nt), consistent with the 375 nt fragment protected by the StyI probe in (b). (e) With the SfoI probe, RNA samples were from uninfected cells (nil), cells infected for 6 h in the presence of cycloheximide (inf) in the presence of cycloheximide (inf) or cells infected overnight without cycloheximide (O/N). The smallest protected fragments (<50 nt) were also seen with RNase-treated, unhybridized probe (probe). Why they were more abundant in the cycloheximide-treated, infected cells is not clear. The larger fragments (approx. 160 nt and 90 nt) were specific for infected cells. Only the 90 nt fragment was prominent when infected cells were treated with cycloheximide. This fragment a known 118697–118605 terminal repeat exon (Husain et al., 1999).

**Fig. 4.** Northern blot analysis of the ORF73 mRNA. RNA from uninfected (UI) or MHV-68-infected (inf) 3T3 cells was denatured, electrophoresed and probed for either ORF73 or β-actin. The predominant ORF73 mRNA was approximately 4200 nt. The ORF itself is 944 bp. β-Actin mRNA (approx. 2300 nt) is less abundant in infected cells because of host shut-off (Stevenson et al., 2000; Rochford et al., 2001). The 5000 nt and 2000 nt bands seen with the β-actin probe probably represent some background hybridization to rRNA.
least 20 copies in circular episomes. The ORF73 mRNA therefore appeared to initiate close to the 3’ end of the viral unique sequence and acquire exons from most of the terminal repeat copies. RNase protection assays clearly had a limited capacity to resolve complex splicing across the G/C-rich terminal repeats. We therefore tested whether there was an initiation site within the terminal repeats, or whether transcription had to start in the unique left end of the viral genome, by analysing the MHV-68 terminal repeats for promoter activity.

**Overlapping ORF73 and ORF75a transcripts**

We first cloned a single repeat unit and tested its capacity to drive CAT expression in transfected 3T3 cells. Initially, we retained ORF73 exon 1 and intron 1 and fused CAT to ORF73 exon 2 at genomic coordinate 118117 (Fig. 5a, 1 x TR). However, even a negative control plasmid, extending only 40 bp upstream of the RACE-mapped ORF73 transcription start site, retained some promoter activity (Fig. 5b, exon 2/PstI clone).

RNase protection assays (Fig. 3c) had indicated that 5’ portions of ORF73 and ORF75a mRNAs could overlap. The most likely explanation, therefore, for the transcriptional activity of the exon 2/PstI clone was that it incorporated an ORF75a promoter. We tested this possibility by RACE mapping the ORF75a transcription start site in MHV-68-infected 3T3 cells and the CAT transcription start site in 3T3 cells transfected with the exon 2/PstI plasmid (Fig. 5c). Infected cells gave a predominant RACE band for ORF75a (114029–117904) of approximately 500 bp, corresponding to transcription initiation near genomic co-ordinate 118220 (Fig. 5a, RACE start site 2). DNA sequence of bulk RACE products identified 5’ mRNA ends at 117950 (RACE start site 1) and 118110. Cloned RACE products gave start sites

---

**Fig. 5.** Overlap of ORF73 and ORF75a mRNAs. (a) Schematic diagram of the key features of the extreme right-hand end of the viral unique sequence and the adjacent terminal repeat copy. The RACE-mapped ORF73 mRNA splices out of the proximal terminal repeat, collects a short exon from the unique right end of the viral genome and then splices to ORF73 itself. The promoter constructs used in (b) and (d) are shown, as well as the RACE-mapped ORF75a transcription start sites corresponding to panel (c). (b) A single terminal repeat (1 x TR) was sufficient to drive the transcription of CAT fused to ORF73 exon 2. However, the same promoter still had activity when truncated at a PstI site only 50 bp upstream of the RACE-mapped ORF73 transcription start site. (c) ORF75a-specific RACE using mRNA from MHV-68-infected 3T3 cells gave a product of approximately 500 bp (arrowhead). DNA sequence analysis of this product identified the two transcription start sites shown. The sub-100 bp band probably represents primer dimers. It does not correspond to an ORF75a mRNA, since the 3’ PCR primer was 163 bp 3’ of the start of the ORF. (d) CAT assays confirmed the presence of a directional promoter in ORF73 intron 1.
at 117981, 117926, 117929 and 118221. Transfected cells gave a transcription start site at 118216. Therefore, there appeared to be two major ORF75a transcripts of very similar size in infected cells: one initiated within ORF73 intron 2; the other used the same transcription start site as the exon 2/PstI CAT clone.

We confirmed that genomic co-ordinates 118117–118598, essentially ORF73 intron 1, contained a constitutively active promoter (Fig. 5d). Crucially, transcription from this promoter initiated in the terminal repeats (ORF75a, RACE start site 2). The copy in the proximal repeat unit transcribed ORF75a: all ORF73 RACE clones were spliced to genomic co-ordinate 118605; PCR with a primer at the 3’ end of intron 1 gave no ORF73-specific product (Fig. 2b); and RNase protection assays indicated an ORF75a transcript, but not an ORF73 transcript, starting close to genomic co-ordinate 118210 (Fig. 3c). However, transcripts from distal copies of this ‘ORF75a’ promoter would run into the next terminal repeat rather than the unique right-hand end of the genome, and would therefore join the 118605–118695 exon to become ORF73 mRNAs. Interestingly, a cDNA clone derived from S11E cells has a terminal repeat exon matching genomic co-ordinates 119373–119195, followed by the 118605–118695 exon (Husain et al., 1999). This fitted nicely with transcription from the promoter in the left-hand portion of an upstream terminal repeat copy.

A major ORF73 promoter in the MHV-68 terminal repeats

Although transcription starting from the left-hand end of an upstream repeat unit could have generated the ORF73 RACE product, the greater activity of the 1 × TR clone than the exon 2/PstI clone in CAT assays (Fig. 5b) suggested additional promoter elements in the terminal repeats. To look at this further, we generated MHV-68 genomic clones with a 3’ end at 118616, thereby fusing CAT to ORF73 exon 1, upstream of the ‘ORF75a’ promoter (Fig. 6a). Extending this region to genomic co-ordinate 118700 gave no CAT transcription. Further 5’ extensions up to the end of the terminal repeat unit gave increasing CAT transcription; adding portions of the next terminal repeat unit gave little additional activity (Fig. 6b). Thus the promoter driving transcription from ORF75a RACE start site 2 had, in this setting, less impact on ORF73 transcription than a promoter in the right-hand portion of the terminal repeat, upstream of the 118700 PstI site. Deleting genomic co-ordinates 118700–119105 abolished promoter activity (Fig. 6b, clone BbvCl-PstI/NotI), as did reversing the orientation of the region upstream of 118700 (Fig. 6c, clone PstI-TR-rev). The lack of transcription from the upstream ‘ORF75a’ promoter in the BbvCl-PstI/NotI construct probably reflected a loss of the 118695 splice acceptor site.

The activity of the SacII genomic clone (5’ limit 119078) (Fig. 6b, d) argued against transcription initiation much upstream of 118800. We used 3’ deletions to try to identify a more precise transcription start site. A 3’ deletion up to 118720 abolished promoter activity (Fig. 6d, compare clones SacII-MspAI1 and SacII). A deletion up to 118705 reduced promoter activity, but did not abrogate it entirely (Fig. 6d, clone BbvCl-PstI/XbaI). Both 3’ deletions also destroyed the 118695 acceptor site. Thus transcription initiation probably occurred just 5’ of the 118700 PstI site, suggesting a TATAA box function for the striking, A/T-rich region at genomic co-ordinates 118756–118778 in the otherwise G/C-rich repeat unit (Fig. 6a, see also Fig. 1b).

Unbiased transcription initiation in the terminal repeats would vary the 5’ untranslated region by 2500 nt, with about 30 different, discrete lengths. There was some evidence for a variable ORF73 mRNA length (Fig. 4), but the predominance of a 4000 nt band implied that initiation was biased, with distal repeat units transcribing ORF73 more actively than proximal ones. A possible explanation was that the unique left end of the viral genome promoted transcription from its adjacent repeat unit, that furthest away from ORF73. Consistent with this idea, a 1.8 kb PstI genomic clone, which contained genomic co-ordinates 118700–119450 of the right-most terminal repeat copy plus 1225 bp of the left end of the genome, gave very high transcriptional activity (Fig. 6e). A 3’ deletion of this genomic clone up to its SacII site at genomic co-ordinate 118720 abolished its promoter activity (data not shown), indicating that transcription remained dependent on the promoter in the right-hand portion of the terminal repeat. Thus, the predominant 4000 nt ORF73 mRNA appeared to reflect that the unique left end of the viral genome enhanced ORF73 transcription from the A/T-rich region of the adjacent terminal repeat copy. Overall, CAT assays provided strong supporting evidence for ORF73 transcription from the viral terminal repeats.

A third ORF73 promoter proximal to the ORF

Despite all RACE clones starting in the terminal repeats, RNase protection assays (Fig. 3b, d) had suggested that the MHV-68 ORF73 was also transcribed from a promoter overlapping the 5’ end of ORF74 (Fig. 7a). RT-PCR (Fig. 7b) confirmed the existence of an unspliced ORF73 transcript, corresponding to the protection of the StyI probe. An unspliced transcript across the 5’ end of ORF73 was also present in cycloheximide-treated, MHV-68-infected cells (Fig. 7c), consistent with the RNase protection data. CAT assays (Fig. 7d) established that the 3’ end of the RACE-mapped ORF73 intron 2 (104872–105430) contained a constitutive promoter, albeit a weaker one than those in the terminal repeats. Thus, at least three separate promoters could transcribe ORF73 with immediate-early kinetics. The 104871–105430 region in reverse orientation also gave CAT transcription, possibly reflecting an ORF74 promoter.

The influence of ORF50 and ORF73 expression on ORF73 promoter activity

A notable feature of the KSHV ORF73 promoter is its upregulation by ORF73 gene expression (Jeong et al., 2004;
Fig. 6. Identification of an ORF73 promoter in right-hand portion of the terminal repeat unit. (a) Schematic diagram of the promoter constructs used. The unique left end of the viral genome contains four of eight viral tRNA homologues. The BAL-31-digested clone that extended 5’ to genomic co-ordinate 118243, just within the upstream terminal repeat copy. (b) Each promoter construct was tested for promoter activity by linking it to CAT, transfecting it into 3T3 cells, and measuring CAT production by ELISA. A single terminal repeat copy was sufficient for strong promoter activity. (c) Reversing the terminal repeat abolished CAT activity. The Tth1111-TR clone is the 1 × TR construct in Fig. 5(b). (d) The 3’ promoter deletions suggested that the transcription start site was just upstream of genomic co-ordinate 118700, consistent with a TATAA-box-like function for the nearby A/T-rich region in the terminal repeat. (e) A genomic clone extending from the likely transcription start site (118700) into the unique left end of the viral genome gave very strong promoter activity.
Renne et al., 2001). In contrast, co-transfected ORF73 down-regulated the MHV-68 ORF73 promoters (Fig. 7e), although to a fairly minor degree. The MHV-68 ORF50 lytic transactivator, which strongly upregulates the MK3 promoter (Coleman et al., 2003), downregulated the promoter just upstream of ORF73, consistent with this promoter being active in latency rather than lytic infection. ORF50 had little effect on the major terminal repeat promoter.
**DISCUSSION**

Regulated episome maintenance is crucial to gammaherpesvirus persistence, as it allows latent virus to respond to the physiological state of the cell while minimizing the presentation of foreign antigens. We identified two promoters in the MHV-68 terminal repeats that were capable of transcribing ORF73, and one promoter upstream of the ORF itself. MHV-68 will no doubt prove just as complex as EBV, which has at least four promoters for EBNA-1. Our data represent a first step in elucidating how ORF73 regulation in MHV-68 relates to its function.

In conjunction with previously published cDNA clones (Husain et al., 1999), our data explain some of the structure of the MHV-68 terminal repeats (Fig. 8). The downstream transcription initiation site lay within 27 bp that are duplicated at the ends of each repeat unit (see the annotation of GenBank accession number af105037). Nineteen basepairs of this 27 bp are duplicated at the extreme right end of the viral unique sequence, thereby mimicking the junction between repeat units. This arrangement may be important for promoter function, although the significance of generating an immediate-early ORF75a promoter in the proximal repeat unit is unclear. Both the terminal repeat promoters lay mainly within introns, separated by a 91 nt 5′ untranslated ORF73 exon. The considerable secondary structure predicted for this exon (Fig. 8c) and supported by the self-protection of the SfoI RNase protection probe (Fig. 3e) may, particularly in tandem copies, contribute to the control of mRNA splicing, nuclear export or translation. Thus an involvement in splicing would explain why the downstream promoter in the proximal repeat unit gave an unspliced ORF75a mRNA, whereas the upstream promoter gave a spliced ORF73 mRNA; there is precedent from the 5′ untranslated regions of other latency transcripts for internal ribosome entry site function (Bieleski & Talbot, 2001; Coleman et al., 2003; Grundhoff & Ganem, 2001; Isaksson et al., 2003; Low et al., 2001); and translational autoregulation is an important immune evasion mechanism of EBNA-1 (Yin et al., 2003).

Transcription from the MHV-68 terminal repeats was evident in infected lungs, spleens and germinal centre B cells, as well as in fibroblasts. In fibroblasts, ORF73 transcription appeared to initiate close to the left end of the viral genome, perhaps explaining the latency deficit seen with some left-end viral mutants (Adler et al., 2001; Clambey et al., 2000). One function of viral tRNA transcription (Bowden et al., 1997) might therefore be to keep the left end of the genome open for ORF73 transcription. The KSHV ORF73 gene product binds to the viral terminal repeats (Ballestas & Kaye, 2001; Garber et al., 2002). In MHV-68, this would provide an opportunity for transcriptional autoregulation. We saw little sign of this in our promoter constructs, but binding to multiple repeat units may be required for a large effect.

---

**Fig. 8.** Structure of the MHV-68 terminal repeat unit. (a) Promoters are shown as shaded blocks. Transcription from the upstream promoter generates ORF73 exon 1. This can then collect additional exon 1 copies before it joins exon 2. Transcription from the downstream promoter generates the Husain et al. exon, which joins the mRNA from the upstream promoter in the next repeat copy at a splice acceptor site just within exon 1. Again this mRNA could accumulate multiple exon 1 copies from adjacent repeat units before reaching exon 2. (b) Each terminal repeat unit has a 27 bp direct repeat at its ends. A 19 bp portion of this repeat is duplicated at the extreme right end of the unique sequence. The transcription start site is in the upstream copy of each pair. (c) Analysis of ORF73 exon 1 with MFOLD (Mathews et al., 1999; Zuker, 2003) suggested that this RNA would have extensive secondary structure. A 5′ untranslated region with multiple ORF73 exon 1 copies would have tandem repeats of this structure, separated by 11 nt linkers.
The large ORF73 intron 2 highlighted a general problem with tracking MHV-68 latency transcripts by RT-PCR. This is necessary to study low-abundance viral transcripts, but the data can be difficult to interpret without transcript mapping. For example, ORF74, which is spanned by ORF73 intron 2 but encoded on the opposite strand, has been identified as an immediate-early transcript (Martinez-Guzman et al., 2003) and as a latency transcript (Virgin et al., 1999) by RT-PCR, but as a late-lytic transcript by strand-specific RNase protection assays (Rochford et al., 2001). Unspliced, nuclear ORF73 mRNAs are a potential source of RT-PCR signal for ORF74, as well as for ORF75a, ORF75b, ORF75c. Also, ORF75a could be transcribed from the proximal repeat unit with immediate-early kinetics, but in the absence of ORF57 (Malik et al., 2004) the unspliced mRNA may not be exported from the nucleus. One way to tackle such ambiguity would be to restrict RT-PCR to cytoplasmic extracts rather than whole-cell lysates.

The promoter just upstream of ORF73 was not a major one in infected fibroblasts, but may be more active in other cell types. It may also play a role in the lytic/latent switch of viral gene expression, since it was strongly downregulated by ORF50; there is clearly the potential for reciprocal inhibition between ORF73 and ORF50 (Lan et al., 2004). Defining the roles of the three different ORF73 promoter in host colonization will probably require in vivo analysis, since normal gene regulation is unlikely to be reproduced in immortalized tumour cells. The present data provide a basis for such a molecular description, and ultimately for understanding how gammaherpesvirus episome maintenance works in vivo.

ACKNOWLEDGEMENTS

Sofia Marques kindly provided some of the cDNA samples used in this study. Janet May generated some of the initial terminal repeat constructs. The Medical Research Council (UK) provided support through grants G108/462, G9901295 and G8000903. P.G.S. is an MRC/Academy of Medical Sciences Clinician Scientist.

REFERENCES


