Title : Adenovirus serotype 5 L4-22K and L4-33K proteins have distinct functions in regulating late gene expression

Susan J. Morris and Keith N. Leppard*

Department of Biological Sciences, University of Warwick, Coventry, CV4 7AL, U.K.

* Author for correspondence: Keith N. Leppard

Phone: +44 24 7652 3579

Fax: +44 24 7652 3701

E-mail: Keith.Leppard@warwick.ac.uk

Short running title: adenovirus L4 22K upregulates late genes

Abstract word count: 243

Main text word count: 6601

Number of figures and tables: 7

Key words: adenovirus, L4 proteins, late gene expression, early-late switch

Abstract

Adenoviruses express up to 20 distinct mRNAs from five major late transcription unit (MLTU) regions L1-L5 by differential splicing and polyadenylation of the primary transcript. MLTU expression is regulated at transcriptional and post-transcriptional levels. L4-33K protein acts as a splicing factor to upregulate several MLTU splice acceptor sites as the late phase progresses. The L4 region also expresses a 22K protein that is sequence-related to L4-33K. L4-22K is shown here also to have an important role in regulating the pattern of MLTU gene expression. An adenovirus genome containing a stop codon in the L4-22K ORF expressed low levels of both structural and non-structural late proteins compared to a wt genome; a decrease in intermediate proteins, IVa2 and IX, was also observed. However, early protein synthesis and replication were unaffected by the absence of L4-22K. Intermediate and late protein expression was restored to wt levels by L4-22K expressed *in trans* but not by L4-33K. Increased MLTU promoter activity, resulting from stabilisation of the transcriptional activator IVa2 by L4-22K, made a small contribution to this restoration of late gene expression. However, the principal effect of L4-22K was on the processing of MLTU RNA into specific cytoplasmic mRNA. L4-22K selectively increased expression of penton mRNA and protein whereas splicing to create penton mRNA is known not to be increased by L4-33K. These results indicate that L4-22K plays a key role in the early-late switch in MLTU expression, additional to and distinct from the role of L4-33K.

Introduction

Human adenovirus serotype 5 (Ad5) is considered as a candidate delivery vector for gene therapy and vaccination. However, despite already achieving use in a clinical setting, some key questions about the basic biology of Ad5 remain. One of these unsolved questions is how control is achieved of late protein expression from the major late promoter (MLP), which directs the production of an array of structural proteins from the major late transcription unit (MLTU) regions L1 – L5 via differential splicing and polyadenylation of the primary transcript (Fig. 1A, reviewed in (1)). Understanding late gene expression is important for further Ad vector development since Ad $E1A^-$ vectors that retain late genes show poor persistence of transgene expression due to immune responses to residual late gene expression products (27, 41-44).

Full expression from the Ad5 MLTU is the end result of a temporal pattern of regulated gene expression that has several phases (reviewed in (35)). The earliest event is expression of the E1A transcriptional activator proteins. These then upregulate transcription of the other early viral genes E1B, E2, E3 and E4. Despite its name, the MLP is active at a low level during this early phase, producing the abundant i-leader protein and L1-52/55K (38). Accumulation of E2A-DNA binding protein (DBP) and E2B- precursor terminal protein (pTP) & DNA polymerase (Pol) allows viral DNA replication to begin. Around the onset of DNA replication, MLTU expression expands to include L4 proteins (22). The intermediate gene products, IVa2 and IX are also expressed at this time. IVa2 expression is regulated by a cellular repressor that is titrated out upon genome replication (19). IX and IVa2 are involved in upregulating transcriptional activity of the MLP during the early to late phase transition of MLTU expression (25, 26, 32, 40). However, it is only as viral DNA replication progresses that the MLP is activated to produce maximal expression of the late structural genes from all of MLTU regions.

The temporal pattern of MLTU expression involves regulation at both the transcriptional and post-transcriptional levels. Early in infection, transcription does not proceed beyond the L3 polyadenylation site, and the L1 polyadenylation site is preferentially used leading to the accumulation of L1-52/55K protein (30). After the initiation of viral DNA replication, MLP activity increases and transcription extends through L4 and L5 regions, allowing a full complement of structural proteins to be produced. Alongside these transcriptional changes, RNA processing is regulated to give altered patterns of mRNA. For example, in the L1 region the proximal 52/55K splice acceptor site is used almost exclusively at early times but during the late phase usage shifts to the distal IIIa acceptor site, due to the production of an infection-specific splicing factor (1). This has now been shown to be L4-33K (39).

L4 is now also known to encode a 22 kDa protein (L4-22K) (31), as predicted from bioinformatic analysis (11), that shares its N-terminus with L4-33K but has a unique C-terminal domain (Fig. 1B). L4-22K has a role in genome packaging via its binding, in association with IVa2, to the A2 sequence within the packaging signal (31, 45, 46). L4-22K has also been shown to bind, with IVa2, to the downstream elements (DE1, DE2) of the MLP (31), which have been shown previously to bind IVa2 as both a homodimer (DEF-B) (25, 40) and in complex (DEF-A) with an unidentified viral protein (25, 28) suggesting that L4-22K, in conjunction with IVa2, may play a role in activating the MLP (31). However, DEF-A has also been shown to contain L4-33K, and IVa2 plus L4-33K activated a MLP reporter (2). Thus it is uncertain what the relative contributions of these two L4 proteins are to regulation of the MLP. The sequence-relatedness between L4-22K and L4-33K also raises the possibility that they might have overlapping functions. We therefore sought to analyse the impact of L4-22K on the expression of Ad5 late proteins. Our data show that L4-22K is required for the transition to late phase MLTU expression and acts primarily at the level of RNA

processing. However, the action of L4-22K is different from that of the splice factor L4-33K and the two proteins are not functionally redundant.

Materials and Methods

Plasmids. pTG3602 (pWT), containing wild type Ad5 genome, and pTG3602-L4-22K$^-$ (pL4-22K$^-$), containing a premature stop codon within the C-terminal unique portion of the L4-22K ORF that results in an additional A*vr*II restriction recognition site (Fig. 1B), have been described previously (8, 31). Expression plasmids for C-terminal FLAG-tagged L4-22K and L4-33K (pCMV-22KFLAG and pCMV-33KFLAG) were constructed by insertion of X*ba*I/P*ac*I fragments, amplified respectively from viral genomic DNA or viral mRNA using primers containing the restriction enzyme recognition sites (primer sequences available on request), into pCMV-FLAG (12). pBiL1-3NheI, pCMV-100KFLAG, and pCMV-22/33KFLAG, which expresses both L4-22K (no epitope tag) and L4-33KFLAG (12), pCMV-IX (5) and pMEPCMV-IVa2 (4) have been described previously. pMLP+DEluc was generated by amplifying the MLP including DE1 and DE2 (Ad5 5939-6174) using primers containing restriction recognition sites for *Kpn*I (5' primer) and *Nhe*I (3' primer) and cloning into pGL3-Basic luciferase reporter plasmid (Promega). pMLP-DEluc, containing the MLP without the downstream elements (Ad5 5939-6133), was generated in a similar way. pcDNA3.1HisLacZ (Invitrogen) was used as a transfection control.

Cells and Viruses. 293 cells were maintained in Dulbecco's MEM (DMEM) supplemented with 5% fetal bovine serum (FBS). 293-IVa2 cells constitutively express IVa2 from an EBV replicon plasmid (4). 293TetOFF cells (Clontech) were maintained in DMEM supplemented with 10% Tet-free FBS and 100µg/ml geneticin (G418 sulphate, Melford Laboratories). An L4-22K complementing, stable

cell line (293-L4-22K) was generated by transfecting 293 cells with pCMV-22KFLAG and selecting clones in DMEM supplemented with 10% FBS and 500µg/ml geneticin. L4-22KFLAG expression was confirmed by immunoblotting and immunofluorescence using a mouse mAb to FLAG (M2, Sigma).

Wild-type virus was Ad5 strain 300 (wt300) (21). To generate an L4-22K$^-$ virus, 293-L4-22K cells were seeded into a 12 well plate and transfected with L4-22K$^-$ genome, excised from pTG3602-L4-22K$^-$ by P*ac*I digestion, using Lipofectamine2000 (Invitrogen). When 60-80% cpe was observed cells were harvested, virus released by three freeze/thaw cycles and cellular debris removed by centrifugation. This L4-22K$^-$ virus stock was passaged 5 times to ensure removal of transfected genomic DNA. To confirm the presence of the mutated sequence in the P5 stock, virus particle DNA was isolated (16) and used as template for a PCR to amplify the Ad5 L4 region 26018-27086bp. PCR products were then digested with A*vr*II, which is diagnostic for the 22K$^-$ mutant (31). Virus titres were determined by fluorescent focus assay. Briefly, 293 cells were infected with a 10-fold dilution series of L4-22K$^-$ virus or wt300, fixed at 20h.p.i. with 10% formalin in phosphate-buffered saline (PBS) and permeablized with 0.5% NP40 in PBS before staining with a mouse mAb to DBP (B6-8, (34)) and goat anti-mouse Alexafluor488 antibody (Invitrogen). Virus titres (FFU/ml) were determined by the number of DBP positive cells at a given dilution (33). For wt300, 1 FFU was equivalent to 7 pfu.

Transfection and Immunoblotting. All transient transfections were carried out in 12 well plates at a density of $7x10^5$ cells/well using either Lipofectamine2000 (Invitrogen) or TransLT (Cambridge Bioscience) at a ratio of 3µl / µg DNA following the manufacturers' protocols. Transfections used 1µg L4-22K$^-$ or wt genome, excised from pL4-22K$^-$ or pWT respectively by *Pac*I digestion, and

500ng various expression plasmids or empty vector (pCMV-FLAG). Experiments using the pBi plasmid system were carried out as previously described (12). Transfected cells were harvested 48h post-transfection directly into sample loading buffer (2% (w/v) sodium dodecyl sulphate [SDS], 50mM dithiothreitol [DTT], 10 % (v/v) glycerol, 25mM Tris pH 6.8 and 0.01 % bromophenol blue) and from 10 - 33% cell lysate volume resolved through either 10% or 15% SDS-polyacrylamide gels as appropriate. Proteins were transferred to ECL nitrocellulose membrane (GE Healthcare) and western blot analysis carried out as described (23). Proteins were detected using the following antibodies: FLAG, M2 mouse mAb (Sigma) at 1:10,000; Ad5 late proteins, AbJLB1 rabbit polyclonal serum at 1:10,000 (12); fibre, RI 89&99 rabbit polyclonal serum 1:5000 (obtained from V. Mautner, University of Birmingham, U. K.); E2A-DBP, mouse mAb B6-8 at 1:10,000; L1-52/55K, rabbit polyclonal serum raised to glutathione-S-transferase (GST) - L1-52/55K fusion protein at 1: 10,000, (3); L4-100K, mouse mAb at 1:10,000 (obtained from W.C. Russell, University of St. Andrews, (36)); IVa2, rabbit polyclonal serum raised to GST-IVa2 at 1:10,000, (4); IX, rabbit polyclonal serum at 1:5000, (5). Secondary antibodies were goat-anti-mouse-horseradish peroxidise (HRP; Sigma) at 1:5000 and goat-anti-rabbit-HRP (Santa Cruz) at 1:100,000. HRP signal was detected using ECL-Advance western blot detection kit (GE Healthcare) according to the manufacturer's instructions.

RT-PCR. Cells were scraped into PBS, pelleted by centrifugation, resuspended in isotonic buffer (150 mM NaCl, 10 mM Tris.HCl pH 7.6, 1.5 mM $MgCl_2$) containing 0.6% v/v NP40 and incubated on ice for 10 min. before pelleting the nuclei by centrifugation at 16000 x g for 3 min. The cytoplasmic fraction was reserved and the nuclei resuspended in lysis solution (isotonic buffer containing 1% NP40 and 1% sodium deoxycholate). RNA was isolated from both the nuclear and

cytoplasmic fraction using TRI-reagent (Sigma) as described (29). 2µg RNA and 100U Superscript reverse transcriptase (RT; Invitrogen) were used with specific 3' primers to generate cDNA in a 25µl volume reaction using manufacturer's protocol. Equivalent RT^- negative reactions omitted Superscript. 2.5µl RT^+ or RT^- reaction were added to a PCR (1x PCR buffer, 0.5mM each dNTP, 0.5µM 5' and 3' primer, 3mM $MgCl_2$ and 1-2.5U *Taq* polymerase (Fermentas) in a total volume of 20µl). Primers used were complementary to Ad5 tripartite leader 1 sequence 6049-6069 and Ad5 L2-penton 14889-14863, Ad5 L2-V 17445-17415 or Ad5 L3-hexon 19254-19228, or to sequences flanking the Ad5 L3 poly(A) and E2A poly(A) sites 22034-22061 and 22749-22722, or to human β-actin cDNA positions 1430-1455 and 2383-2368. Reactions were incubated at 94°C for 5 min followed by 30 cycles of 94°C, 1 min, 65°C, 1 min and 72°C, 1 min. PCR products were resolved through 1% agarose alongside size markers (1kb Generuler, Fermentas).

Southern Blot Analysis. 293 cells were infected with 2 x 10^{-3} FFU/cell L4-22K$^-$ virus or wt300, and harvested 3 or 20h. p.i. Cells were processed either to purify packaged viral DNA, using spermine precipitation to remove unpackaged material (16), or for total low molecular weight DNA by HIRT extraction (18). DNA was digested with *Hind*III, resolved through 0.7% agarose and transferred to Hybond-N membrane (GE Healthcare) and viral DNA detected using an L1 probe (10,589-11,565bp from pGem4-L1 (37)) labelled using the AlkPhos direct labelling kit (GE Healthcare).

Luciferase Assay. 293 and 293-IVa2 cells were transfected with 500ng reporter plasmid (pMLP+DEluc or pMLP-DEluc) and 100ng pcDNA3.1HisLacZ, together with 500ng pCMV-22KFLAG and/or 1µg L4-22K$^-$ genome. All transfections were equalised for DNA content by the addition of either salmon sperm DNA or pCMV-FLAG as appropriate. Cells were lysed 24h post-

transfection in 1x passive lysis buffer and a luciferase assay carried out according to the manufacturer's instructions (Promega). Cell lysates were also analysed for β-galactosidase activity using 2-nitrophenyl-β-D-galactopyranoside (ONPG) as substrate. The ratio of firefly luciferase to β-galactosidase was calculated to correct for differences in transfection efficiency. Data are shown as fold induction where luciferase expression from the reporter plasmids is set as 1. Statistical analysis was performed using Student's T-Test.

Results

L4-22K is required for efficient late gene expression in 293 cells. Our previous studies of the early to late transition in MLTU expression utilised plasmid pBiL1-3NheI, in which MLTU genomic sequence encoding the tripartite leader (TPL) and regions L1 – L3 was placed under the control of a heterologous promoter. Full late pattern expression could be induced from this plasmid by CMV expression vectors encoding L4-33K and -100K proteins (12). The "33K" plasmid utilised was also capable of L4-22K expression, although L4-22K had not been proven to exist at that time. Subsequently, L4-33K was shown by Tormanen et al. to be a splicing factor capable of activating certain MLTU splice sites (39). To determine whether L4-22K made any contribution to the MLTU activation observed in this system, plasmids expressing either L4-22K or L4-33K with C-terminal FLAG tags were constructed. When co-transfected with pBiL1-3NheI and an L4-100K vector to permit efficient hexon protein accumulation, L4-22KFLAG increased levels of several late proteins independently of L4-33K (Fig. 1C). These effects were qualitatively distinct, with L4-22K increasing penton expression more than, and hexon expression less than did L4-33K. These data suggested that L4-22K made a contribution to late protein expression from L1-L3, although it could not be excluded

198 that this effect was due to expression of a truncated L4-33K-related protein from the L4-22K plasmid
199 clone.
200 To investigate fully the role of L4-22K in regulating MLTU expression, given the prior
201 evidence that L4-22K and/or L4-33K might act on the MLP, we focused further studies on a whole
202 genome system, in which the MLTU was controlled by the native MLP within linear full-length
203 genome excised from pTG3602 (8). This also provided the possibility of IVa2 production that was
204 not present in the pBi plasmid system. The role and mechanism of action of L4-22K in late gene
205 expression from this genome was assessed by comparison with an L4-22K$^-$ genome that retained the
206 ability to express full-length L4-33K (Fig. 1B, (31)), kindly provided by P. Hearing.
207 293 cells transfected with L4-22K$^-$ genome showed a decrease in late structural protein
208 production (hexon, penton, IIIa, V and fibre) compared to a wt genome (Fig. 2A & B); while the
209 absolute level of this defect varied between experiments, as exemplified here, it was always
210 substantial. In addition, expression of the non-structural late proteins L4-100K and L1-52/55K from
211 L4-22K$^-$ genome was reduced to virtually undetectable levels (Fig. 2B). The level of E2A-DBP
212 detected was at least equivalent for L4-22K$^-$ and the wt genome, and possibly somewhat increased in
213 the mutant, demonstrating that the defect in late protein expression was not due to differences in
214 transfection efficiency and that the absence of L4-22K does not cause a global downregulation of
215 viral protein production. Expression of all late proteins from the L4-22K$^-$ genome was restored to
216 wild-type levels when cells were co-transfected with an expression plasmid for L4-22KFLAG (Fig.
217 2A), confirming that the reduced late protein expression seen from L4-22K$^-$ genome was due solely
218 to the lack of L4-22K expression. L4-22K has been shown previously to be involved in the
219 packaging of Ad genomes into particles (31), so at 48 h post-transfection it was possible that the
220 observed differences in late gene expression resulted from reduced levels of second-round infection

in L4-22K$^-$ genome-transfected cultures consequent upon a failure to package genomes efficiently. However, this possibility was discounted as similar results were observed at 20 h post-transfection, a time point when no progeny virus would have been able to reach the late phase of a secondary infection (data not shown). The late protein defect of L4-22K$^-$ is kinetic rather than absolute as, when analysis was delayed to 60 h post-transfection, L4-22K$^-$ genome produced similar levels of late proteins to those seen at 20 h post-transfection for either wt genome or L4-22K$^-$ genome co-transfected with L4-22KFLAG expression plasmid (data not shown). Thus L4-22K upregulates late gene expression in 293 cells.

L4-22K acts via a different mechanism to L4-33K. The stop codon inserted into the L4-22K ORF in pL4-22K$^-$ is located within the intron sequence of L4-33K ORF and it was possible that this mutation might affect the expression of L4-33K, hence causing some part of the observed defect in late protein production. Although the defect was complemented by L4-22KFLAG, this plasmid also has the capacity to express a truncated L4-33K protein consisting of the N-terminal 129 amino acids that might contribute to this complementation. The truncated protein would lack the C-terminal RS motifs that have been show to be essential for this activity (39), making this unlikely. However, to exclude the possibility, L4-22K$^-$ was co-transfected with an expression plasmid for L4-33KFLAG. Expressing L4-33KFLAG *in trans* did not restore expression of either the structural proteins or L4-100K (Fig. 2B), although a small increase in L1-52/55K was observed. Late protein expression from L1 through to L5 was only restored when L4-22K was added *in trans* either from pCMV-22KFLAG or from a plasmid, pCMV-22/33KFLAG, capable of expressing both L4-22K and L4-33KFLAG from alternatively spliced transcripts (Fig. 2B). During complementation by pCMV-22K/33KFLAG, expression of the structural proteins and L4-100K was greater than that seen with L4-22K$^-$ genome

alone but was reduced compared to when L4-22KFLAG was expressed from pCMV-22KFLAG. This may reflect preferential expression of the spliced L4-33KFLAG transcript from this plasmid and consequently a lower expression level of L4-22K protein than was achieved with pCMV-22KFLAG. These data show that the defect in late gene expression is not due to any unexpected aberrations in L4-33K expression occurring due to the mutation in the intronic sequence of L4-33K ORF. Furthermore, the evidence clearly shows the mechanism of action of L4-22K is independent of L4-33K splicing function and thus indicates a direct role for L4-22K in the activation of the late phase of infection.

Reduced late gene expression in the absence of L4-22K is not due to a defect in replication. Full expression from the Ad5 MLTU is the end result of a temporal pattern of regulated gene expression. It is only once viral DNA replication has begun that the MLP is activated to produce maximal expression of the late structural genes from all MLTU regions. Thus a defect in replication could account for the reduction in late protein production observed from L4-22K$^-$ genome. To allow replication studies to be carried out without the concern of transfection efficiencies or the confusion of large quantities of input DNA, L4-22K$^-$ virus was isolated by transfection of L4-22K$^-$ genome into a stable L4-22K complementing cell line, 293-L4-22K. Although this cell line only expressed a low level of L4-22KFLAG, it was sufficient to permit the growth of virus to low titre.

The production of L4-22K$^-$ virus was confirmed by *Avr*II digest of PCR products amplified from DNA isolated from L4-22K$^-$ virus particles (Fig. 3A). Using particle DNA ensured that only newly synthesised and packaged DNA was acting as a template in the PCR and not any remaining transfected genome. Incubation of the L4-22K$^-$ virus product with A*vr*II resulted in the virtually complete cleavage to bands 532 and 536bp in size, indicating presence of the site diagnostic of the

L4-22K$^-$ mutation (31). The small amount of uncut DNA is likely to represent incomplete digestion. If any wild-type DNA was present, it would rapidly overgrow the slow-growing mutant but no increase in the proportion of this uncut species was seen upon further passage; the presence of the inserted stop codon was also confirmed by sequencing (data not shown). The L4-22K$^-$ virus had the packaging defect expected from previous work (Fig. 3B) (31). At 20 h. p.i. viral DNA was not detected for L4-22K$^-$ unless complemented by prior expression of L4-22KFLAG; when complemented, amounts of L4-22K$^-$ DNA were similar to wt virus. No viral DNA was detected at 3 h. p.i. indicating that viral DNA detected at 20 h. p.i. was due to newly synthesised, packaged genomic DNA and not virus input. These data confirm the L4-22K deficient nature of the L4-22K$^-$ virus generated in this study both genetically and biologically.

To investigate the growth properties of L4-22K$^-$, the late protein expression defect observed with transfected genome was first confirmed. Non-complementing cells infected with L4-22K$^-$ virus behaved identically to cells transfected with L4-22K$^-$ genome in showing a substantial impairment of structural protein expression (data not shown). Genome replication was then assessed (Fig. 3C). At 3h. p.i. a low level of viral DNA was detected; this time point is too early for the synthesis of progeny viral DNA and thus shows the amount of input viral DNA. The increased level of viral DNA detected at 20 h. p.i. was similar for uncomplemented and complemented L4-22K$^-$ virus, showing that the absence of L4-22K has no effect on virus DNA replication. This result agrees with that reported previously in which replication of the L4-22K$^-$ genome was investigated in HeLa cells (31). The block in MLTU gene expression must therefore occur after DNA replication.

A reduction in intermediate proteins IVa2 and IX is not fully responsible for the reduction in late protein expression from L4-22K$^-$. The two intermediate proteins, IVa2 and IX upregulate the

MLP activity during the transition between the early and late phases of infection (25, 26, 32, 40). IVa2 acts via binding, either as a homodimer (DEF-B) or heterodimer (DEF-A), to sequences DE1 and DE2 within the MLP (20, 28). Furthermore, the pattern of differential protein expression observed here in the absence of L4-22K is similar to that reported previously for MLP DE1 and DE2 mutants (32) suggesting that L4-22K may be acting via IVa2, perhaps as a component of DEF-A since enhanced binding of IVa2 to DE1,2 in the presence of 22K has been reported (31).

Expression of IVa2 and IX from L4-22K$^-$ genome alone was barely detectable, whereas levels of both proteins were substantially increased when L4-22KFLAG was co-expressed (Fig. 4A). To determine whether this reduction in intermediate proteins was responsible for the limited expression of late proteins observed for L4-22K$^-$ genome, these two proteins were expressed *in trans*. Co-transfection of L4-22K$^-$ genome with IVa2 expression plasmid increased expression of hexon, penton and the smaller isoform of protein V, but failed to increase IIIa, the upper isoform of V or pVI (Fig. 4B). In contrast, co-transfection with IX expression plasmid failed to increase the level of late gene expression significantly. When IX and IVa2 were expressed together, the level of all late proteins tested increased a little as compared to either protein alone. However, this late protein expression was very much lower than that seen when the L4-22K$^-$ genome was directly complemented by L4-22KFLAG (Fig. 4B).

L4-22K clearly increases the level of both IVa2 and IX proteins produced from transfected genome (Fig. 4A). In the case of IVa2 this may reflect a stabilisation of the protein rather than a transcription or RNA processing event since L4-22KFLAG also increases the level of IVa2 protein expressed from cDNA under the control of the heterologous CMV promoter (Fig. 4C); this effect was specific to L4-22K as neither L4-33KFLAG nor L4-100KFLAG had any effect (Fig. 4C). This observation suggests that the increase in IVa2 seen when L4-22K$^-$ genome is complemented is not

mediated through IVa2's interaction with L1-52/55K, its known partner in viral DNA packaging (15), expression of which also increases upon complementation of L4-22K$^-$ (Fig. 2B). The IVa2 sequence within pMEPCMV-IVa2 does not contain the binding site for the cellular repressor that blocks IVa2 transcription prior to replication (9, 24) and thus L4-22K cannot be acting by relieving this repression. In addition, no effect of L4-22K on expression of other proteins from the CMV promoter was observed (data not shown). Therefore, we hypothesise that L4-22K acts to stabilise IVa2 protein.

It is possible that lack of IVa2 stabilisation by L4-22K is the reason why IVa2 expressed *in trans* can only partially complement the late protein expression defect of L4-22K$^-$ genome. However, this seems unlikely as the levels of IVa2 achieved were similar to those seen when L4-22K$^-$ was fully complemented by L4-22KFLAG. Furthermore, similar results were obtained when a cell line constitutively expressing high levels of IVa2 was used (Fig. 5C), excluding the possibility that L4-22K somehow affects IVa2 transfection efficiency. Alternatively, the requirement for L4-22K may reflect a direct contribution of this protein to the DEF-A transcription factor that is required for maximal activation of the MLP (31).

L4-22K increases MLP activity in a DE-dependent manner. To determine if L4-22K directly regulated transcription from the MLP, and whether this was DE-dependent, luciferase expression from MLP reporters either containing or lacking DE1,2 was assessed in the presence of the L4-22K$^-$ genome or complemented genome (Fig. 5A). The presence of the genome ensured all viral proteins, other than L4-22K, that might be required for transcriptional activation were present in the system. Luciferase expression was increased 45-60 fold by the presence of the L4-22K$^-$ genome, regardless of the presence or absence of the DE elements. This significant increase occurred despite the fact that the L4-22K$^-$ genome displays a severe reduction in IVa2, IX and late protein synthesis. Expression of

L4-22KFLAG further increased the activity of pMLP+DEluc by two-fold compared to L4-22K$^-$ genome alone ($p<0.005$), whereas the activity of pMLP-DEluc was not significantly altered ($p>0.1$), demonstrating that the increase was DE-dependent (Fig. 5A). In contrast, L4-22KFLAG expressed in the absence of viral genome had no effect on the activity of MLP containing the DEs (Fig. 5B), suggesting that another viral protein was required in association with L4-22K.

A likely candidate protein was IVa2, which was known to act via the DEs (40) and had been shown to be stabilised in the presence of L4-22K (Fig. 4). To test whether DE-dependent MLP activation by L4-22K was indirectly due to increased IVa2 levels or due to a direct interaction of L4-22K and IVa2 at the DEs, MLP reporter activity was analysed in a cell line constitutively expressing high levels of IVa2 (293-IVa2) (4). Activity of the DE-containing MLP in 293-IVa2 cells was increased a modest 1.2 fold compared to its activity in 293 cells. When L4-22KFLAG was co-expressed in these cells, the increase was 1.6 fold compared to 293 cells, a 1.3 fold increase compared to 293-IVa2 cells (Fig. 5B). Therefore, both L4-22K and IVa2 contribute to MLP activation. However, the small increase in MLP activity observed when L4-22KFLAG is expressed in the presence of high levels of IVa2 does not explain the large differences in the expression of late proteins, notably penton and IIIa, from complemented genome in the same cell line (Fig. 5C). This suggests that L4-22K may have a role in late gene expression that is post-transcriptional, in addition to its effects, with or via IVa2, at the MLP. This conclusion is supported by the observed increase in L1-3 gene expression from a heterologous promoter in the presence of L4-22KFLAG, when no IVa2 was present (Fig. 1C).

L4-22K acts at the level of late gene mRNA production/stability. To determine whether L4-22K was exerting its post-transcriptional effect at the level of mRNA or protein production, the amounts

of penton, V and hexon mRNA produced by L4-22K$^-$ and wt virus were assessed. Penton and V mRNA production from L4-22K$^-$ virus was barely detectable, compared to readily detectable expression from wt virus (Fig. 6); this defect was fully complemented when L4-22KFLAG was expressed *in trans* (Fig. 6). In contrast, the level of hexon mRNA produced by L4-22K$^-$ virus was similar to that from wt and not further increased when complemented by L4-22K. PCR products were not from DNA contamination of the template as equivalent reactions containing no reverse transcriptase were negative (Fig. 6). The quantity and quality of RNA used was confirmed by the equivalent amplification of β-actin mRNA (Fig. 6). Therefore, a major component of the reduction in late gene expression observed for L4-22K$^-$ is at the level of late mRNA production and/or stability, and the action of L4-22K is selective for specific mRNAs.

L4-22K does not increase the level of un-processed late mRNA. To further clarify the role of L4-22K in transcription and RNA processing, the level of unspliced nuclear MLTU RNA was determined by RT-PCR amplification of a region spanning the L3 polyadenylation signal sequence (Ad5 22034-22722); this region is only present in RNA unprocessed at the L3 polyadenylation site. Amounts of unprocessed L3 RNA in wt and L4-22K$^-$ virus infections were comparable, and complementation of L4-22K$^-$ virus with L4-22KFLAG had no effect (Fig. 7 left upper panel). The level of unprocessed E2A RNA, tested as a control since its protein product was unaffected by L4-22K, was equal for all samples as expected (Fig. 7 right upper panel) and control reactions lacking reverse transcriptase were negative (Fig 7 right lower panel), indicating that the PCR products were generated from RNA template. These data provide further evidence against a major role for L4-22K in transcriptional activation of the MLP, and focus attention instead on MLTU mRNA processing as its principal site of action.

Discussion

Ad5 late protein synthesis is regulated at the level of transcription, mRNA processing and translation. Early in infection, MLTU expression is limited to L1-52/55K. However, after DNA replication begins, expression extends across MLTU regions L1 through to L5 resulting in over 20 distinct mRNAs and thus proteins. The L4 region is expressed immediately after DNA replication (22) and two of its three non-structural protein products have been described as factors involved in the regulation of late protein synthesis. L4-33K is required for late mRNA splicing (12, 39) whereas L4-100K is responsible for selective translation of late mRNAs (17) and for the stabilisation and assembly of hexon trimers (6). The results presented here demonstrate that the third L4 non-structural protein, L4-22K, is also involved in the regulation of late protein expression and appears to have distinct functions in transcription and RNA processing.

A L4-22K deficient genome exhibited a reduction in both structural and non-structural late protein expression from MLTU regions L1 through to L5 (fibre, hexon, penton, IIIa, V, VI, L4-100K and L1-52/55K) compared to a wt genome, and this defect was fully complemented by exogenous L4-22K expression. The first impact of L4-22K deficiency was in the late phase as early gene expression, demonstrated by E2A-DBP levels, was not affected and newly synthesised DNA levels from L4-22K$^-$ virus-infected cells were comparable to those when L4-22K$^-$ was complemented with exogenous L4-22K. L4-22K has been shown previously to be required for viral genome packaging (31), and this same defect was also observed in this study. However, failure to package genome efficiently could not be the cause of the observed reduced late gene expression since the defect was seen at times post-infection when any secondary infections could not have reached the late phase of

gene expression. Thus, defective late gene expression is a primary consequence of L4-22K deficiency.

L4-33K has been shown previously to upregulate late gene expression post-transcriptionally (12, 39), so it was possible that lack of L4-22K affected late protein production via a failure to express L4-33K correctly. No antibodies specific for L4-33K were available to test this directly. However, L4-33K expressed *in trans*, in constrast to L4-22K, had almost no impact on the defective late gene expression from L4-22K$^-$ genome. Therefore the reduction in late gene expression from this mutant genome is a direct effect of the absence of L4-22K and a not a consequence of either unanticipated aberrant expression of L4-33K due to the L4-22K$^-$ mutation that lies within the L4-33K intron or reduction in L4-33K levels that might result from the lack of L4-22K effects on the MLP. Moreover, the action of L4-22K must be independent of the demonstrated effect of L4-33K on late mRNA splicing.

When high level transcription from L1-L3 regions was rendered independent of Ad5 IVa2 and other viral factors by the use of a heterologous promoter, complementation by L4-22K upregulated penton levels disproportionately to hexon, the reverse of the effect of L4-33K. These observations fit with the work of Tormanen et al. who, when showing that L4-33K was the previously demonstrated Ad late splicing factor (39), found that L4-33K had no effect on penton mRNA splicing *in vitro*. These authors classified splice sites by length of polypyrimidine tract and showed that splicing stimulation by L4-33K correlated inversely with tract length. Our findings therefore raise the possibility that L4-22K might also be selective for a particular class of splice site. If so, it does not correspond to polypyrimidine tract strength; penton and V mRNA levels were both strongly L4-22K – responsive in our study even though their 3’ splice sites have very different polypyrimidine tracts and V mRNA splicing was reported to be strongly stimulated by L4-33K in

contrast to the lack of effect on penton (39). Hexon mRNA accumulation did not require L4-22K in our study despite hexon protein production being strongly dependent on L4-22K. This discordance is explained by the dependence of hexon protein accumulation on the presence of L4-100K (12), which was itself found to be strongly dependent on L4-22K for its expression. Taken together, these data suggest that L4-33K and L4-22K each act post-transcriptionally to support the development of the full pattern of late viral gene expression in different ways.

As well as the late proteins, the two intermediate proteins, IVa2 and IX were reduced in the absence of L4-22K. These two proteins are only expressed after viral DNA replication has begun (10, 19), and both have been implicated in activation of the MLP. The data presented here show that L4-22K increases IVa2 protein levels, independent of the IVa2 promoter and known regulatory sequences. IVa2 has been shown previously to bind to L4-22K during Ad DNA packaging, and L4-22K also promotes the binding of IVa2 to the MLP DEs (31). Thus, we suggest that L4-22K increases IVa2 levels by binding and stabilising it, and via this action L4-22K achieves a modest activation of the MLP that is dependent on the DEs. However, this effect is minor in comparison with the overall effect of L4-22K on late gene expression.

Previous studies have sought to address the function of L4 proteins through directed mutation of the L4 reading frames. A virus with a stop codon at position 20 of the shared L4-33K/22K reading frame (v33K.1) and expected to lack both proteins was viable, with normal early gene expression and viral DNA replication but a substantial defect in late protein synthesis (13), a phenotype that is similar to that found here for L4-22K$^-$. However, a virus with its L4-33K reading frame truncated by stop codons and expected to lack the C-terminal 47 residues proved impossible to isolate, suggesting it was a lethal mutation (14). This finding suggests that v33K.1 must be leaky for functional L4-33K expression and hence also for L4-22K, although since we have not been able to complement the Δ47

mutant genome (kindly provided by S. J. Flint) in a cell line expressing L4-33K (data not shown), it is also possible that the severe Δ47 phenotype includes pleiotropic effects not directly attributable to a lack of L4-33K. In this case, v33K.1 may represent the true null phenotype for both L4-22K and L4-33K. The L4-22K$^-$ virus that was isolated here, using the mutated plasmid genome constructed by Ostapchuk and colleagues (31), was clearly viable and showed defects in late gene expression and genome packaging. It retained the restriction site difference diagnostic of the original mutation and its phenotype was identical to that of the transfected genome from which it was derived. Thus, although it cannot be excluded that its viability results from second-site mutation(s) that compensate for aspects of the L4-22K$^-$ phenotype, this seems unlikely. Because of the shared L4-33K/22K exon, the L4-22K$^-$ mutation truncates the reading frame only after 113 residues. It thus has the potential to express a substantive protein, albeit containing only 8 residues of the 22K-unique sequence, which might have residual functions distinct from those of L4-33K. Although no shorter protein immunoreactive with anti-33K serum was detected from this genome (31), a contribution of such a protein to the viability of L4-22K$^-$ virus cannot be excluded. However, our data, taken together with previous studies, suggest that Ad5 can grow, albeit with low efficiency, without L4-22K and that impaired late gene expression and particle formation is the null phenotype for this protein.

Our results on late gene expression in the absence of L4-22K differ from those previously reported using the L4-22K$^-$ genome. Ostapchuk and colleagues reported no differences from wt in replication, DBP, L1-52/55K, hexon or penton protein levels (31). In this study we also detected no differences in replication or DBP levels, but detected a clear reduction in the levels of L1-52/55K, penton and to a lesser extent hexon. We believe the previous data do in fact show some differences in L1-52/55K and penton levels but to a considerably lesser degree than our data indicate. Different cell types, 293 and HeLa, were used for the two studies but this does not appear to explain this

difference since we consistently observed the same L4-22K$^-$ late gene expression defect in infected HeLa cells as seen in 293 cells by either genome transfection or virus infection (data not shown). Possibly, either culture conditions or HeLa cell strain differences affect the expression of host cell factors with which L4-22K needs to interact to upregulate late gene expression and hence modulate the severity of the phenotype observed. Alternatively, differences between the two studies in the effective time post-infection / transfection that assays were conducted may provide an explanation, since the defect in late gene expression that we observed diminished as the time to assay was extended.

The MLP is activated by several factors, including the binding of DEF-A and DEF-B to the downstream elements DE1 and DE2. The importance of these factors is most apparent in the absence of activation via upstream promoter elements (32). DEF-A is a heterodimer of IVa2 and another protein, and there is published evidence for this protein being either L4-33K or L4-22K (2, 31). Our data do not bear directly on this issue, since no studies of protein interactions with the DEs were performed. However, MLP activity was shown to be only modestly stimulated by L4-22K, in cooperation with IVa2, suggesting that L4-22K is not the unknown component of DEF-A. This does not exclude that a complex of IVa2 and L4-22K can bind to the DEs, as previously reported (31); indeed, our data on IVa2 stabilisation suggest that increased IVa2 complex formation on the DEs when L4-22K is present should be expected.

In contrast to the modest effects on MLP of IVa2 and L4-22K, the presence of the full viral genome *in trans* with the MLP reporter caused a 60-fold increase in activity. This suggests that something other than IVa2 or L4-22K is required, either another factor expressed from the genome or the presence of replicating genome itself. One possible factor is E2A-DBP, which has previously been shown to increase MLP activity substantially (7) and is expressed from L4-22K$^-$. Another

potential activating factor coming from L4-22K$^-$ is E1A. However, 293 cells already express E1A, and we have observed only a slight increase in luciferase expression from another E1A-responsive reporter in these cells when E1A is further over-expressed, Nonetheless, it is possible that the population of E1A isoforms expressed in 293 cells is functionally altered during viral infection and that the activation of the MLP in this context is a reflection of this. Further experiments are therefore required to determine the basis of MLP activation by Ad genome.

This study, together with previous reports, shows that the L4 region of the MLTU provides two crucial regulators of the temporal transitions in MLTU RNA processing that are observed over the course of Ad infection. L4-22K and L4-33K have complementary effects that together provide the normal pattern of mRNA production during the late phase of infection. L4-33K was shown previously to be a splicing factor that activates splicing to specific late mRNA 3' splice sites. However, this activity did not extend to all splice sites whose activity is seen to increase during the early-late transition in MLTU expression. It is possible that L4-22K is also a splicing factor, but with a distinct specificity from L4-33K. Alternatively, it may act less directly, perhaps via changes in host proteins. Further experiments are needed to address these questions.

Acknowledgements

We gratefully acknowledge the following colleagues for their generous gifts of reagents used in this study: P. Hearing, Stony Brook University, N.Y.; S. J. Flint, Princeton University, N.J.; W. C. Russell, University of St. Andrews, U.K.; V Mautner, University of Birmingham, U.K. This work was supported by the Biotechnology and Biological Sciences Research Council (grant number BB/E014550/1).

References

1. **Akusjarvi, G.** 2008. Temporal regulation of adenovirus major late alternative RNA splicing. Frontiers in Bioscience **13:**5006-15.
2. **Ali, H., G. LeRoy, G. Bridge, and S. J. Flint.** 2007. The adenovirus L4 33-kilodalton protein binds to intragenic sequences of the major late promoter required for late phase-specific stimulation of transcription. J Virol **81:**1327-38.
3. **Arslanoglu, A.** 1999. Studies of adenovirus L1 gene aimed at L1 gene deficiencies for use in gene therapy vectors. PhD Thesis, University of Warwick, U.K.
4. **Brey, S.** 1999. Construction of IVa2-deficient adenovirus using complementing cell lines. PhD Thesis, University of Warwick, U.K.
5. **Caravokyri, C., and K. N. Leppard.** 1995. Constitutive episomal expression of polypeptide IX (pIX) in a 293-based cell line complements the deficiency of pIX mutant adenovirus type 5. J Virol **69:**6627-33.
6. **Cepko, C. L., and P. A. Sharp.** 1982. Assembly of adenovirus major capsid protein is mediated by a nonvirion protein. Cell **31:**407-15.
7. **Chang, L. S., and T. Shenk.** 1990. The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters. J Virol **64:**2103-9.
8. **Chartier, C., E. Degryse, M. Gantzer, A. Dieterle, A. Pavirani, and M. Mehtali.** 1996. Efficient generation of recombinant adenovirus vectors by homologous recombination in Escherichia coli. J Virol **70:**4805-10.
9. **Chen, H., R. Vinnakota, and S. J. Flint.** 1994. Intragenic activating and repressing elements control transcription from the adenovirus IVa2 initiator. Mol Cell Biol **14:**676-85.

10. **Crossland, L. D., and H. J. Raskas.** 1983. Identification of adenovirus genes that require template replication for expression. J Virol **46:**737-48.

11. **Davison, A. J., M. Benko, and B. Harrach.** 2003. Genetic content and evolution of adenoviruses. J Gen Virol **84:**2895-2908.

12. **Farley, D. C., J. L. Brown, and K. N. Leppard.** 2004. Activation of the early-late switch in adenovirus type 5 major late transcription unit expression by L4 gene products. J Virol **78:**1782-91.

13. **Fessler, S. P., and C. S. H. Young.** 1999. The role of the L4 33K gene in adenovirus infection. Virology **263:**507-516.

14. **Finnen, R. L., J. F. Biddle, and J. Flint.** 2001. Truncation of the human adenovirus type 5 L4 33-kDa protein: evidence for an essential role of the carboxy-terminus in the viral infectious cycle. Virology **289:**388-99.

15. **Gustin, K. E., P. Lutz, and M. J. Imperiale.** 1996. Interaction of the adenovirus L1 52/55-kilodalton protein with the IVa2 gene product during infection. J Virol **70:**6463-7.

16. **Hardy, S., M. Kitamura, T. Harris-Stansil, Y. Dai, and M. L. Phipps.** 1997. Construction of adenovirus vectors through Cre-lox recombination. J Virol **71:**1842-9.

17. **Hayes, B. W., G. C. Telling, M. M. Myat, J. F. Williams, and S. J. Flint.** 1990. The adenovirus L4 100-kilodalton protein is necessary for efficient translation of viral late mRNA species. J Virol **64:**2732-42.

18. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J Mol Biol **26:**365-9.

19. **Iftode, C., and S. J. Flint.** 2004. Viral DNA synthesis-dependent titration of a cellular repressor activates transcription of the human adenovirus type 2 IVa2 gene. Proc Natl Acad Sci U S A **101:**17831-6.

20. **Jansen-Durr, P., G. Mondesert, and C. Kedinger.** 1989. Replication-dependent activation of the adenovirus major late promoter is mediated by the increased binding of a transcription factor to sequences in the first intron. J Virol **63:**5124-32.

21. **Jones, N., and T. Shenk.** 1978. Isolation of deletion and substitution mutants of adenovirus type 5. Cell **13:**181-8.

22. **Larsson, S., C. Svensson, and G. Akusjarvi.** 1992. Control of adenovirus major late gene expression at multiple levels. J Mol Biol **225:**287-98.

23. **Lethbridge, K. J., G. E. Scott, and K. N. Leppard.** 2003. Nuclear matrix localization and SUMO-1 modification of adenovirus type 5 E1b 55K protein are controlled by E4 Orf6 protein. J Gen Virol **84:**259-68.

24. **Lin, H. J., and S. J. Flint.** 2000. Identification of a cellular repressor of transcription of the adenoviral late IVa(2) gene that is unaltered in activity in infected cells. Virology **277:**397-410.

25. **Lutz, P., and C. Kedinger.** 1996. Properties of the adenovirus IVa2 gene product, an effector of late-phase-dependent activation of the major late promoter. J Virol **70:**1396-405.

26. **Lutz, P., M. Rosa-Calatrava, and C. Kedinger.** 1997. The product of the adenovirus intermediate gene IX is a transcriptional activator. J Virol **71:**5102-9.

27. **Michou, A. I., L. Santoro, M. Christ, V. Julliard, A. Pavirani, and M. Mehtali.** 1997. Adenovirus-mediated gene transfer: influence of transgene, mouse strain and type of immune response on persistence of transgene expression. Gene Ther **4:**473-82.

28. **Mondesert, G., C. Tribouley, and C. Kedinger.** 1992. Identification of a novel downstream binding protein implicated in late-phase-specific activation of the adenovirus major late promotor. Nucleic Acids Res **20:**3881-9.

29. **Morris, S. J., H. Smith, and C. Sweet.** 2002. Exploitation of the Herpes simplex virus translocating protein VP22 to carry influenza virus proteins into cells for studies of apoptosis: direct confirmation that neuraminidase induces apoptosis and indications that other proteins may have a role. Arch Virol **147:**961-79.

30. **Nevins, J. R., and M. C. Wilson.** 1981. Regulation of adenovirus-2 gene expression at the level of transcriptional termination and RNA processing. Nature **290:**113-8.

31. **Ostapchuk, P., M. E. Anderson, S. Chandrasekhar, and P. Hearing.** 2006. The L4 22-kilodalton protein plays a role in packaging of the adenovirus genome. J Virol **80:**6973-6981.

32. **Pardo-Mateos, A., and C. S. Young.** 2004. Adenovirus IVa2 protein plays an important role in transcription from the major late promoter in vivo. Virology **327:**50-9.

33. **Philipson, L.** 1961. Adenovirus assay by the fluorescent cell-counting procedure. Virology **15:**263-8.

34. **Reich, N. C., P. Sarnow, E. Duprey, and A. J. Levine.** 1983. Monoclonal antibodies which recognize native and denatured forms of the adenovirus DNA-binding protein. Virology **128:**480-4.

35. **Russell, W. C.** 2000. Update on adenovirus and its vectors. J Gen Virol **81:**2573-2604.

36. **Russell, W. C., G. Patel, B. Precious, I. Sharp, and P. S. Gardner.** 1981. Monoclonal antibodies against adenovirus type 5: preparation and preliminary characterization. J Gen Virol **56:**393-408.

37. **Soloway, P. D., and T. Shenk.** 1990. The adenovirus type 5 i-leader open reading frame functions in cis to reduce the half-life of L1 mRNAs. J Virol **64:**551-8.

38. **Symington, J. S., L. A. Lucher, K. H. Brackmann, A. Virtanen, U. Pettersson, and M. Green.** 1986. Biosynthesis of adenovirus type 2 i-leader protein. J Virol **57:**848-56.

39. **Tormanen, H., E. Backstrom, A. Carlsson, and G. Akusjarvi.** 2006. L4-33K, an adenovirus-encoded alternative RNA splicing factor. J Biol Chem **281:**36510-7.

40. **Tribouley, C., P. Lutz, A. Staub, and C. Kedinger.** 1994. The product of the adenovirus intermediate gene IVa2 is a transcriptional activator of the major late promoter. J Virol **68:**4450-7.

41. **Yang, J., T. Tsukamoto, N. Popnikolov, R. C. Guzman, X. Chen, J. H. Yang, and S. Nandi.** 1995. Adenoviral-mediated gene transfer into primary human and mouse mammary epithelial cells in vitro and in vivo. Cancer Lett **98:**9-17.

42. **Yang, Y., H. C. Ertl, and J. M. Wilson.** 1994. MHC class I-restricted cytotoxic T lymphocytes to viral antigens destroy hepatocytes in mice infected with E1-deleted recombinant adenoviruses. Immunity **1:**433-42.

43. **Yang, Y., F. A. Nunes, K. Berencsi, E. E. Furth, E. Gonczol, and J. M. Wilson.** 1994. Cellular immunity to viral antigens limits E1-deleted adenoviruses for gene therapy. Proc Natl Acad Sci U S A **91:**4407-11.

44. **Zabner, J., L. A. Couture, R. J. Gregory, S. M. Graham, A. E. Smith, and M. J. Welsh.** 1993. Adenovirus-mediated gene transfer transiently corrects the chloride transport defect in nasal epithelia of patients with cystic fibrosis. Cell **75:**207-16.

45. **Zhang, W., and M. J. Imperiale.** 2000. Interaction of the adenovirus IVa2 protein with viral packaging sequences. J Virol **74:**2687-2693.

46. **Zhang, W., J. A. Low, J. B. Christensen, and M. J. Imperiale.** 2001. Role for the adenovirus IVa2 protein in packaging of viral DNA. J Virol **75:**10446-10454.

Figure legends

Figure 1. (A) Schematic diagram of the major late transcription unit showing the alternatively spliced mRNAs and encoded proteins. (B) Organisation of the L4-22K and L4-33K ORF. The position of the L4-22K truncation mutation in L4-22K$^-$ is indicated (31). (C) 293TETOFF cells, which constitutively express the tetracycline transactivator protein were transfected with pBiL1-3NheI (which expresses the MLTU L1-3 region under the control of a tetracycline-regulated promoter), pCMV100KFLAG and either empty vector or the L4-22KFLAG or L4-33KFLAG expression plasmids. Cell lysates were harvested 48h post-transfection, separated on 10% polyacrylamide gel and late proteins, as indicated on the right, detected by western blot analysis using AbJLB1 polyclonal anti-late protein serum. The positions to which proteins of known molecular mass migrated are shown on the left (kDa).

Figure 2. Late protein expression from L4-22K$^-$ genome. Cells were mock transfected or transfected with L4-22K$^-$ genome and either empty vector or various L4 expression plasmids as indicated. Ad5 wt genome was used as a control. Cell lysates were separated on 10% polyacrylamide gels and expressed proteins, as indicated at the right of each panel, detected by western blot analysis using either AbJLB1 polyclonal anti-late protein antiserum (top panels, A & B), RI polyclonal anti-late protein antiserum to detect fibre (B) or monospecific / monoclonal antibodies as indicated. In top panel B, all lanes shown derive from the same exposure of a single blot with irrelevant lanes excised for clarity of presentation. Note that no FLAG-tagged 22K is detected from transfection of the single

L4-22/33K plasmid as the tag sequence is only attached to the 33K reading frame. However, the expected expression of 22K from this plasmid can be inferred from its effect on late protein expression. The positions to which proteins of known molecular mass migrated are shown on the left (kDa).

Figure 3. L4-22K has no effect on viral DNA replication. (A) Confirmation of L4-22K$^-$ mutant status. L4-22K$^-$ virus was generated and passaged five times in a stable L4-22K complementing cell line, 293-L4-22K. Ad5 26018-27086 bp region was amplified by PCR from DNA isolated from either L4 22K$^-$ or wt virus particles. PCR products were either undigested or digested with *Avr*II as indicated. Lane M, 1kb ladder, sizes marked on the left (kbp). (B) L4-22K effects on packaging. 293 cells were either mock transfected or transfected with L4-22KFLAG and 24 h. later infected with 2 x 10^{-3} FFU/cell L4-22K$^-$ virus or Ad5 wt. Packaged DNA isolated from infected cells at either 3 h. p.i. or 20 h. p.i. was digested with *Hind*III and analysed by southern blotting using an L1 probe. (C) L4-22K effects on replication. As B. except DNA was obtained by HIRT extraction of cell cultures at 3 h. or 20 h. 50ng Ad5wt viral DNA was used as a positive control (B & C).

Figure 4. Restoring expression levels of the intermediate proteins IVa2 and IX does not fully complement L4-22K$^-$. (A&B) 293 cells were mock transfected, or transfected with L4-22K$^-$ genome and either empty vector or various expression plasmids as indicated. (C) 293 cells were mock transfected, or transfected with the IVa2 expression plasmid with either empty vector or various L4-protein expression vectors as indicated. Cell lysates were separated on 10% or 15% polyacrylamide gels and expressed proteins detected by western blot analysis using anti-IVa2, anti-

IX or AbJLB1 anti-late protein antisera. The positions to which proteins of known molecular mass migrated are shown on the left (kDa).

Figure 5. L4-22K increases transcription activity of IVa2 at the MLP. (A) 293 cells were transfected with pMLP+DEluc (black bars) or pMLP-DEluc (grey bars), alone or with L4-22K$^-$ genome and various expression plasmids as indicated. (B) 293 cells (grey bars) or 293-IVa2 (black bars) cells were transfected with pMLP+DEluc (black bars), alone or with L4-22K$^-$ genome and various expression plasmids as indicated. Firefly luciferase expression was detected using a luciferase assay, corrected for transfection efficiency using β-galactosidase expression from an independent control plasmid, and expressed as fold induction where expression from the reporter plasmids alone was set as 1. Error bars show standard deviations where n = 3. (C) 293 or 293-IVa2 cells were transfected with L4-22K$^-$ genome and either empty vector or L4-22KFLAG expression plasmid. Cell lysates were separated on 10% polyacrylamide gels and expressed proteins detected by western blot analysis using AbJLB1 anti-late protein or anti-IVa2 antisera. The positions to which proteins of known molecular mass migrated are shown on the left (kDa).

Figure 6. L4-22K increases the level of late mRNA. RT-PCR amplifications were performed on total cytoplasmic RNA from 293 cells that were either mock transfected or transfected with L4-22KFLAG and 24 h. later infected with $2x10^{-3}$ FFU/cell L4-22K$^-$ virus or Ad5 wt as indicated. Late mRNAs were detected by RT-PCR using a 5' primer specific for tripartite leader 1 and 3' primers directed to L2 penton (947bp product), L2 V (1138bp) or L3 hexon (662bp) 3'exons. Equivalent control reactions were performed lacking reverse transcriptase (RT$^-$); amplification of β-actin mRNA (511bp product) served as a positive control. For V mRNA, both RT-PCR and RT$^-$ results derive

699 from the same experiment and are presented at the same exposure. The positions to which DNA size
700 markers migrated are indicated (kbp).
701
702 **Figure 7. L4-22K acts at the level of RNA processing.** 293 cells were either mock transfected or
703 transfected with L4-22KFLAG expression vector. 24h post-transfection, cells were mock-infected, or
704 infected with 2 x 10^{-3} FFU/cell L4-22K^{-} virus or Ad5 wt and nuclear RNA isolated 20h p.i. Amounts
705 of L3 RNA unprocessed at the L3 polyA site (upper left panel), and of E2A RNA unprocessed at its
706 polyA site (upper right panel), were determined by RT-PCR using specific primers. β-actin mRNA
707 served as a positive control (lower left panel). To confirm that amplified products derived from
708 RNA template, the same amount of RNA was added to reactions where the reverse transcriptase had
709 been omitted (lower right panel). The positions to which DNA size markers migrated are indicated
710 (kbp).

Figure 1

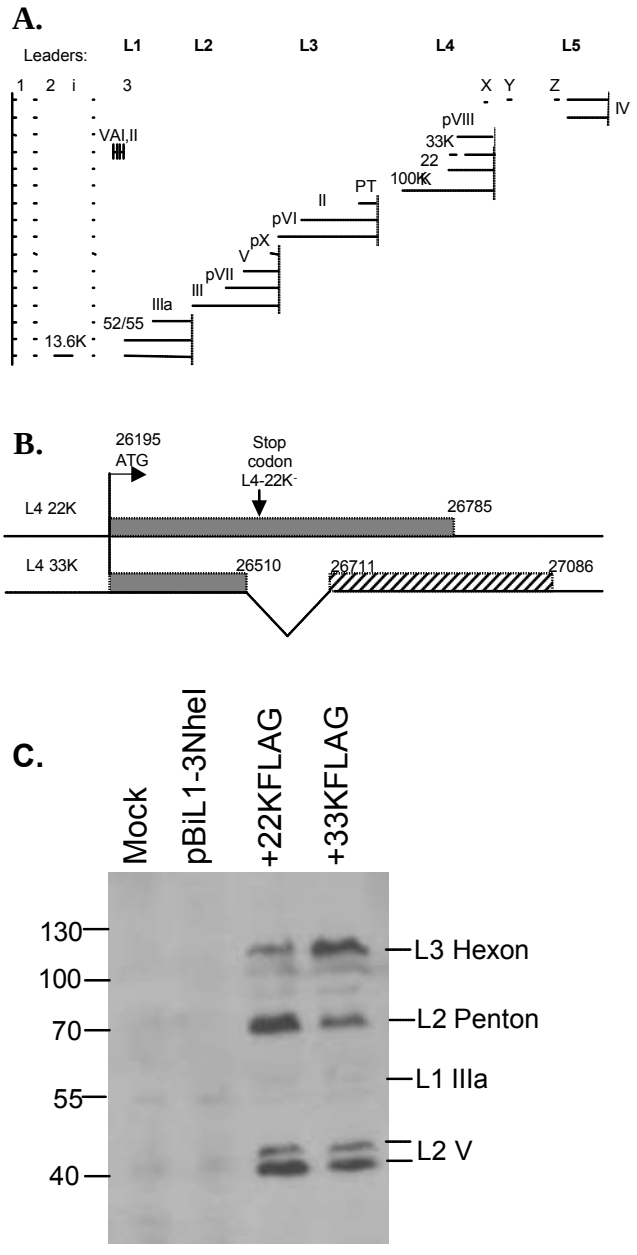

A.
Leaders:
L1
L2
L3
L4
L5
1
2
i
3
X
Y
Z
IV
VAI,II
pVIII
33K
22
100K
PT
II
pVI
pX
V
pVII
III
IIIa
52/55
13.6K
B.
26195
ATG
Stop
codon
L4-22K
L4 22K
26785
L4 33K
26510
26711
27086
C.
Mock
pBiL1-3Nhel
+22KFLAG
+33KFLAG
130
100
70
55
40
L3 Hexon
L2 Penton
L1 IIIa
L2 V

Figure 2

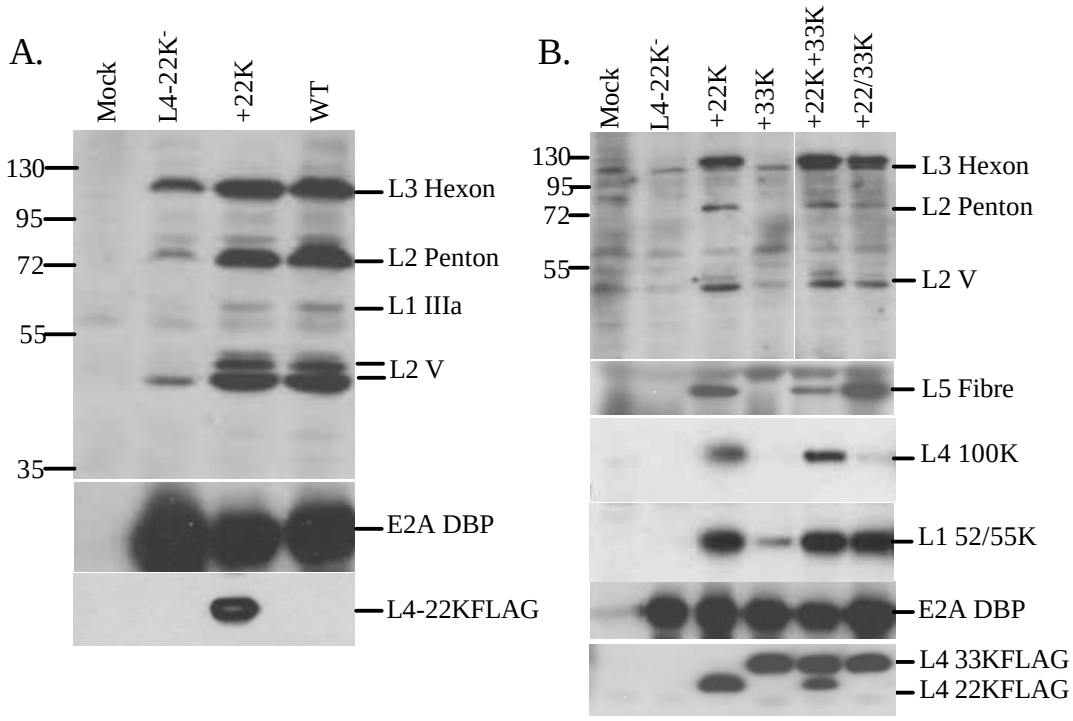

A.
Mock
L4-22K-
+22K
WT
130
95
72
55
35
L3 Hexon
L2 Penton
L1 IIIa
L2 V
E2A DBP
L4-22KFLAG
B.
Mock
L4-22K-
+22K
+33K
+22K+33K
+22/33K
130
95
72
55
L3 Hexon
L2 Penton
L2 V
L5 Fibre
L4 100K
L1 52/55K
E2A DBP
L4 33KFLAG
L4 22KFLAG

Figure 3

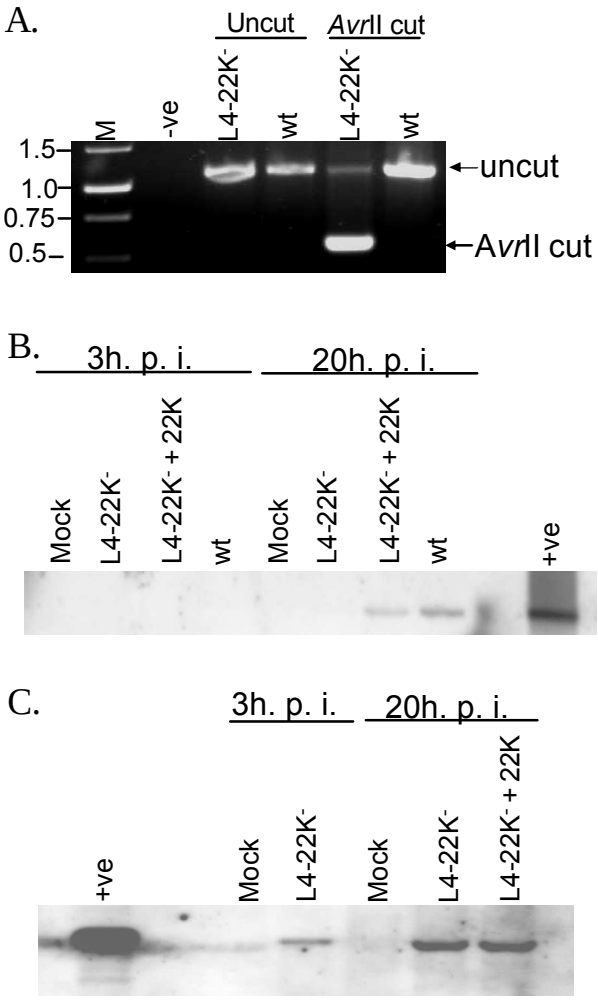

A.
Uncut
AvrII cut
M
-ve
L4-22K-
wt
L4-22K-
wt
1.5
1.0
0.75
0.5
uncut
AvrII cut
B.
3h. p. i.
20h. p. i.
Mock
L4-22K-
L4-22K- + 22K
wt
Mock
L4-22K-
L4-22K- + 22K
wt
+ve
C.
3h. p. i.
20h. p. i.
+ve
Mock
L4-22K-
Mock
L4-22K-
L4-22K- + 22K

Figure 4

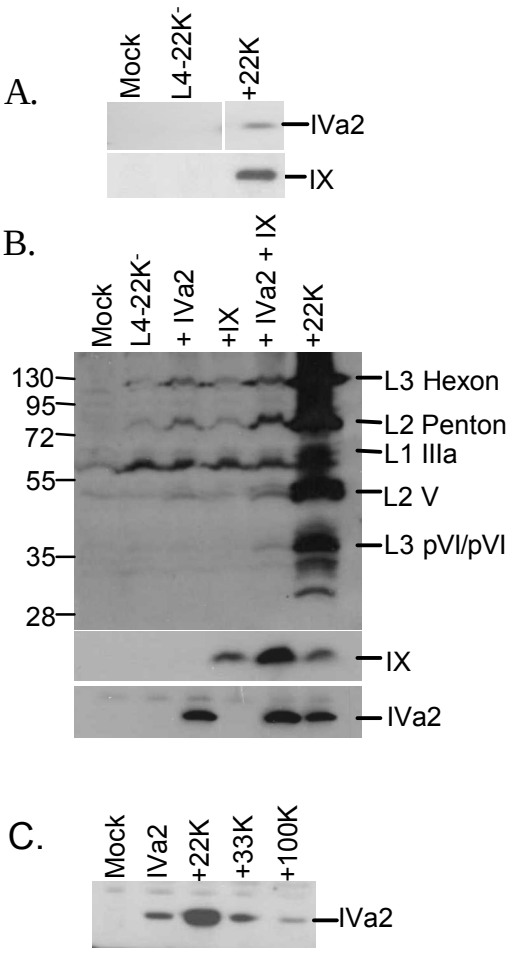
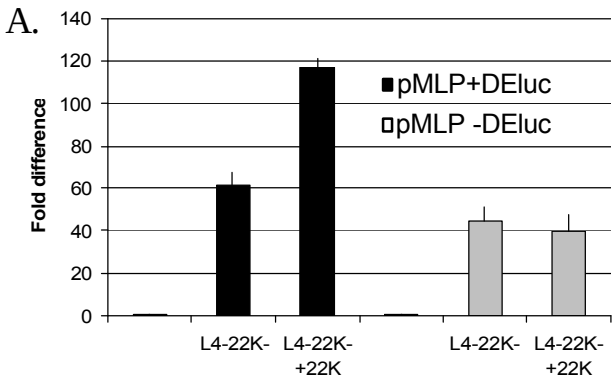
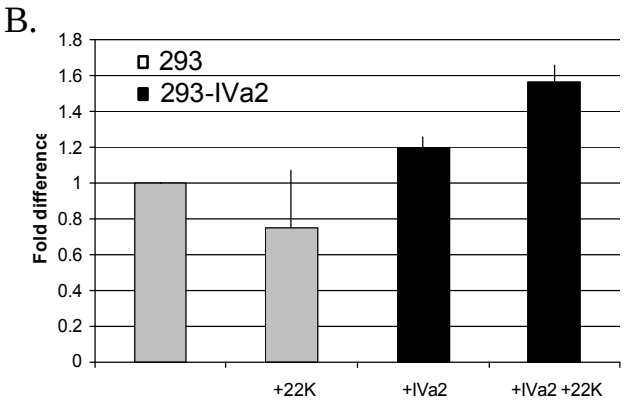

A.
Mock
L4-22K-
+22K
IVa2
IX
B.
Mock
L4-22K-
+ IVa2
+IX
+ IVa2 + IX
+22K
130
95
72
55
35
28
L3 Hexon
L2 Penton
L1 IIIa
L2 V
L3 pVI/pVI
IX
IVa2
C.
Mock
IVa2
+22K
+33K
+100K
IVa2

Figure 5


A.
140
120
100
80
60
40
20
0
Fold difference
pMLP+DEluc
pMLP -DEluc
L4-22K-
L4-22K-
+22K
L4-22K-
L4-22K-
+22K


B.
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Fold difference
293
293-IVa2
+22K
+IVa2
+IVa2 +22K

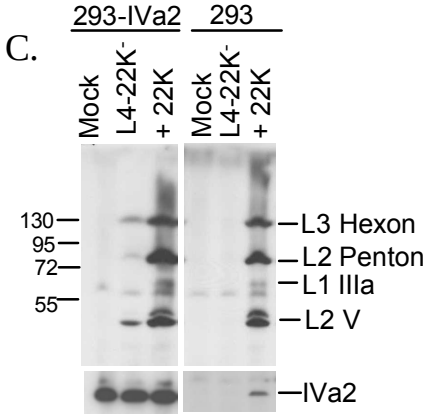

C.
293-IVa2
293
Mock
L4-22K-
+ 22K
Mock
L4-22K-
+ 22K
130
95
72
55
L3 Hexon
L2 Penton
L1 IIIa
L2 V
IVa2

Figure 6

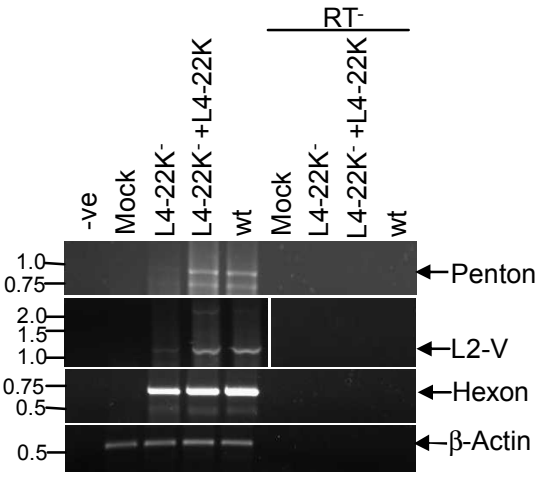
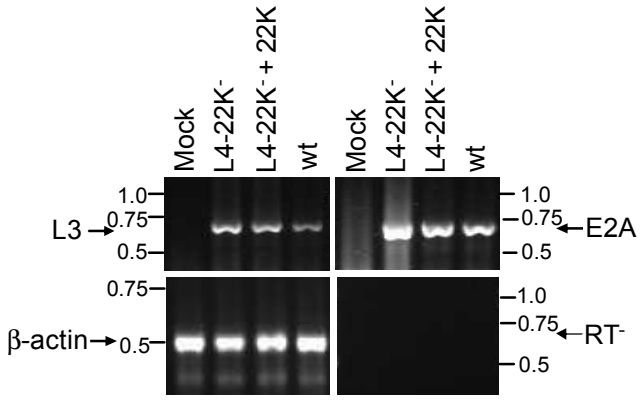

RT-
-ve
Mock
L4-22K-
L4-22K-+L4-22K
wt
Mock
L4-22K-
L4-22K-+L4-22K
wt
1.0
0.75
Penton
2.0
1.5
1.0
L2-V
0.75
0.5
Hexon
0.5
β-Actin

Figure 7


Mock
L4-22K-
L4-22K- + 22K
wt
Mock
L4-22K-
L4-22K- + 22K
wt
1.0
0.75
0.5
L3
E2A
0.75
0.5
β-actin
1.0
0.75
0.5
RT-