JOURNAL OF VIROLOGY, Feb. 1995, p. 935–947
0022-538X/95/$04.00+0

Vol. 69, No. 2

Identification of Nuclear and Nucleolar Localization Signals in the Herpes Simplex Virus Regulatory Protein ICP27

WENDY E. MEARS, VIVIAN LAM, AND STEPHEN A. RICE*

Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7

Received 9 July 1994/Accepted 31 October 1994

Previous work has shown that the herpes simplex virus type 1 (HSV-1) regulatory protein ICP27 localizes to the cell nucleus and that certain mutant ICP27 polypeptides localize preferentially in nucleoli. To map the signals in ICP27 which mediate its nuclear localization, we identified the portions of ICP27 which can direct a cytoplasmic protein, pyruvate kinase (PK), to nuclei. Our results demonstrate that ICP27 contains multiple nuclear localization signals (NLSs) that function with differing efficiencies. First, ICP27 possesses a strong NLS, mapping to residues 110 to 137, which bears similarity to the bipartite NLSs found in *Xenopus laevis* nucleoplasmin and other proteins. Second, ICP27 possesses one or more weak NLSs which map to a carboxyl-terminal portion of the protein between residues 140 and 512. Our PK-targeting experiments also demonstrate that ICP27 contains a relatively short sequence, mapping to residues 110 to 152, that can function as a nucleolar localization signal (NuLS). This signal includes ICP27's strong NLS as well as 15 contiguous residues which consist entirely of arginine and glycine. This latter sequence is very similar to an RGG box, a putative RNA-binding motif found in a number of cellular proteins which are involved in nuclear RNA processing. To confirm the results of the PK-targeting experiments, we mutated the ICP27 gene by deleting sequences encoding either the strong NLS or the RGG box. Deletion of the strong NLS (residues 109 to 138) resulted in an ICP27 molecule that was only partially defective for nuclear localization, while deletion of the RGG box (residues 139 to 153) resulted in a molecule that was nuclear localized but excluded from nucleoli. Recombinant HSV-1s bearing either of these deletions were unable to replicate efficiently in Vero cells, suggesting that ICP27's strong NLS and RGG box carry out important in vivo functions.

Herpes simplex virus type 1 (HSV-1) is a widespread and pathogenic human alphaherpesvirus. In infected individuals, HSV-1 undergoes two quite different interactions with cells: a productive (or lytic) infection, which can occur in multiple cell types, and a latent infection, which occurs in neurons. Although neuronal latency is not well understood, much is known about the productive infection, which can be studied in cultured cells in vitro. Productive infection consists of a highly ordered program of viral gene expression, DNA replication, and virion assembly and culminates in the formation of infectious viral progeny and cell death (reviewed in reference 50). Currently, there is much interest in the regulatory mechanisms which coordinate the events of productive infection. In particular, several of the HSV-1 immediate-early (IE) proteins, the first viral proteins to be synthesized in the infected cell, have been shown to have important regulatory roles and are thus the subjects of ongoing studies.

One such IE protein is ICP27 (ICP = infected cell protein). Genetic studies have demonstrated that this ~63-kDa, 512-residue phosphoprotein is required for the transition between the delayed-early (DE) and late (L) phases of productive infection and is absolutely essential for the formation of infectious progeny (31, 42, 43, 45, 52). On the basis of phenotypes of engineered viral mutants, ICP27 appears to carry out two distinct functions that contribute to the DE to L transition (42, 43, 45). First, although it is not absolutely required for viral DNA replication, ICP27 increases the level of replicated DNA approximately 10-fold. Second, ICP27 mediates a switch in viral gene expression that results in the activation of L genes and the concomitant repression of DE and IE genes. This latter function of ICP27 is genetically separable from its ability to stimulate DNA replication (42, 43).

The molecular mechanisms by which ICP27 carries out its lytic functions have not yet been defined. However, several recent studies suggest that ICP27 may work, at least in part, through posttranscriptional gene regulatory mechanisms. Compelling evidence for this possibility comes from cotransfection experiments, in which ICP27 can either activate or repress the expression of cotransfected reporter genes in uninfected cells (4, 13, 41, 55, 63). These effects appear to be independent of the promoter sequences which drive expression of the reporter genes (8, 53). Rather, regulatory sequences downstream of the protein-coding regions appear to mediate the effects of ICP27. Specifically, activation of reporter genes by ICP27 has been correlated with the presence of particular polyadenylation [poly(A)] signals (8, 53), while repression has been correlated with the presence of introns (53). Several additional studies suggest that ICP27 may regulate genes at a posttranscriptional level. First, ICP27 is required during infection for the induction of late-processing factor, a pre-mRNA processing activity that selectively increases the efficiency of polyadenylation at an HSV L gene poly(A) site (34). Second, HSV-1 causes the redistribution of cellular small nuclear ribonucleoprotein (RNP) antigens in the infected cell nucleus, and ICP27 is both required and sufficient for this effect (38). Last, studies with an HSV-1 ICP27 temperature-sensitive mutant have provided some evidence that ICP27 can regulate certain HSV-1 genes by posttranscriptional mechanisms during the productive infection (59).

ICP27 is localized to the cell nucleus during infection, as determined by both cell fractionation (1) and immunofluorescence (25) studies. ICP27 is also localized to the nucleus when

* Corresponding author. Mailing address: Department of Biochemistry, University of Alberta, 474 Medical Sciences Building, Edmonton, Alberta, Canada T6G 2H7. Phone: (403) 492-2717. Fax: (403) 492-0886. Electronic mail address: steve_rice@darwin.biochem.ualberta.ca.

it is expressed in uninfected cells (46). These studies suggest that, like most other nuclear proteins, ICP27 possesses one or more nuclear localization signals (NLSs). In addition, some evidence suggests that ICP27 interacts with cell nucleoli. First, although wild-type (WT) ICP27 protein is not preferentially localized to nucleoli, it does show significant accumulation in these structures (25, 42, 46). Second, certain mutant ICP27 polypeptides show altered nucleolar localization, as demonstrated by immunofluorescence analysis of cells expressing ICP27 nonsense mutants (42, 46). The *n*263R protein, which consists of the N-terminal half of ICP27 (263 of 512 residues), is preferentially localized in nucleoli, while the *n*406R protein, which consists of the N-terminal ~80% of ICP27 (406 residues), is localized to the nucleus but is largely excluded from nucleoli.

The objective of this study was to identify the polypeptide determinants within ICP27 which influence its ability to localize to, and within, the nucleus. We find that ICP27 contains multiple signals, both strong and weak, which mediate its nuclear localization. In addition, ICP27 possesses a relatively short sequence which can function as a nucleolar targeting signal. Interestingly, a portion of this sequence bears similarity to a putative RNA-binding motif found in a number of cellular proteins which are involved in nuclear RNA processing.

MATERIALS AND METHODS

Cells, viruses, and infections. All infections and transfections were carried out in Vero (African green monkey kidney) cells or in V27 cells (42), a derivative of Vero cells which contain an integrated copy of the ICP27 gene. The cells were propagated in Dulbecco modified Eagle medium plus 10% heat-inactivated fetal bovine serum (GIBCO). The viruses used for infections were the WT strain KOS1.1 and the ICP27 mutants *d*27-1 (42), *d*3-4, and *d*4-5 (both described below). Infections were carried out at a multiplicity of infection of 10 PFU per cell.

Construction of plasmids. To begin these studies, we constructed a vector, ppAPKX, which could be used to engineer ICP27-pyruvate kinase (PK) fusion genes. This was accomplished in several cloning steps. First, a functional poly(A) signal was cloned in pUC19. This was done by isolating a 100-bp *Xba*I-*Sal*I fragment containing the ICP27 gene poly(A) site from plasmid ptkCAT27 (44) and cloning it into the *Kpn*I site of pUC19, using *Kpn*I linkers. Next, a 1.8-kb *Bgl*II fragment containing the chicken PK-coding region (codons 17 to 529) was isolated from plasmid PK10b (15) and cloned into the *Bam*HI site of pUC19-polyA. This generated the plasmid pUC19-polyAPK. Finally, the *Xba*I site upstream of the PK-coding region in pUC19-polyAPK was changed to an *Xho*I site. To do this, pUC19-polyAPK was digested with *Xba*I, the 3′ recessed ends were filled in by using the large fragment of *Escherichia coli* DNA polymerase I, and an 8-bp *Xho*I linker was inserted. This plasmid was designated ppAPKX.

ppAPKX was used to engineer constructs in which N-terminal segments of the ICP27-coding region were fused to the PK-coding region. To do this, we used the ICP27 plasmid mutants pM1 through pM16, which contain engineered *Xho*I restriction sites in the ICP27 gene (43). The *Xho*I sites are spaced at numerous sites throughout the ICP27-coding region and occur in a common reading frame. To construct hybrid genes, pM mutants were doubly digested with *Hin*dIII and *Xho*I. The small *Hin*dIII-*Xho*I restriction fragments were then cloned into ppAPKX that had been digested with *Hin*dIII and *Xho*I. Ligation of the *Xho*I ends results in the in-frame fusion of the ICP27 and PK-coding regions. The constructs were named according to the pM plasmid which was used in each construction; e.g., pM4PK was derived using pM4. pMSPK was constructed by taking advantage of the naturally occurring *Sal*I site in the ICP27 gene. To construct pMSPK, the *Hin*dIII-*Sal*I fragment from pM27 (43) was cloned into *Hin*dIII-*Xho*I-digested ppAPKX.

To engineer PK fusion genes containing C-terminal segments of ICP27, we modified pM1PK by converting its *Xho*I site to a *Cla*I site. To do this, pM1PK was linearized with *Xho*I, the 3′ recessed ends were filled in, and an 8-bp *Cla*I linker was inserted. The resulting plasmid was designated pMCPK. Next, pMCPK was linearized with *Bst*XI, which cuts near the 3′ end of the PK open reading frame. The DNA ends were made blunt by using T4 DNA polymerase, and an 8-bp *Xho*I linker was inserted. The resulting plasmid was designated pMCPKX. To construct hybrid genes, members of the pM1–M16 series were doubly digested with *Eco*RI and *Xho*I. The small *Eco*RI-*Xho*I restriction fragments were then cloned into pMCPKX via its *Eco*RI and *Xho*I sites. Ligation of the *Xho*I ends results in the in-frame fusion of ICP27 coding segments to codon 523 of PK. As a consequence, the C-terminal six residues of PK are replaced by C-terminal portions of ICP27. The C-terminal constructs were named according to the pM mutants used; e.g., pMPK3 was derived using pM3. To study the effect of simply deleting the C-terminal six residues of PK, we constructed pMPKΔ, which contains nonsense mutations after codon 523. To make pMPKΔ, pMCPKX was linearized with *Xho*I, the 3′ recessed ends were filled in, and a 12-bp *Nhe*I linker containing stop codons in all three reading frames (New England Biolabs) was inserted.

Plasmids encoding hybrid genes which contained internal portions of the ICP27-coding region were made using two different strategies. To make pM3-SPK, pM4-SPK, pM5-SPK, pM6-SPK, and pMS-16PK, plasmids pM3, pM4, pM5, pM6, and pM16, respectively, were digested with *Xho*I and *Sal*I, and the small *Xho*I-*Sal*I restriction fragments were cloned into the unique *Xho*I site of pM1PK. Restriction enzyme analysis was used to identify those clones which had the insert in the desired orientation. The remaining hybrid plasmids were constructed by a multiple-step procedure. In the first step, pM1PK was altered by the insertion of a *Bgl*II linker immediately downstream of its unique *Xho*I site. This plasmid was designated pMXBPK. Next, oligonucleotide-directed mutagenesis (Altered Sites system; Promega) was used to engineer *Bgl*II sites into pM2, pM3, pM4, and pM5. This resulted in the plasmid derivatives pM2-4B, pM2-5B, pM3-4B, pM3-5B, pM3-6B, and pM4-6B, which contain engineered *Bgl*II sites at ICP27 codons 138/139, 153/154, 138/139, 153/154, 173/174, and 173/174, respectively. Finally, the small *Xho*I-*Bgl*II fragments from these plasmids were cloned into pMXBPK via its *Xho*I and *Bgl*II sites. This resulted in plasmids pM2-4PK, pM2-5PK, pM3-4PK, pM3-5PK, pM3-6PK, and pM4-6PK.

Two additional plasmids were constructed. The first was pMPK3-4, which encodes a hybrid protein in which residues 110 to 137 of ICP27 are tagged to the C terminus of PK. To construct pMPK3-4, pM3-4PK was modified by the insertion of a 12-bp *Nhe*I stop codon linker at its *Bgl*II site. The resulting plasmid, pM3-4NPK, was digested with *Xho*I and *Eco*RI, and the 2.0-kb *Xho*I-*Eco*RI fragment was cloned into *Xho*I-*Eco*RI-digested pMCPKX. This generated pMPK3-4. The second plasmid construct was pMd3-4, which encodes an ICP27 molecule lacking residues 109 to 138. To engineer this construct, pM3 and pM4 were digested with *Xho*I and *Eco*RI. The 1.6-kp *Xho*I-*Eco*RI fragment from pM4 was then ligated to the 6.4-kb *Eco*RI-*Xho*I fragment of pM3, generating pMd3-4.

Construction of HSV-1 recombinants. The recombinant viruses *d*3-4 and *d*4-5, bearing engineered deletions in the ICP27 gene, were generated by a strategy involving in vivo homologous recombination as previously described (42, 45). The plasmids which served as the source of the mutant ICP27 alleles were pMd3-4 (described above) and pMd4-5. pMd4-5 was constructed in a manner analogous to pMd3-4 except that pM4 and pM5 were the parental plasmids. The *d*3-4 and *d*4-5 ICP27 alleles encode ICP27 molecules that lack residues 109 to 138 and 139 to 153, respectively. For each mutant, two independent viral isolates, designated a and b, were obtained. The genomic structures of all four viral isolates were confirmed by diagnostic Southern blotting (44).

Transfections. For transient transfections, Vero cells were plated on glass coverslips in 3.8-cm^2 wells of 12-well tissue culture plates. The next day, cells in individual wells were transfected with 1 μg of an ICP27-PK hybrid plasmid, using the calcium phosphate precipitation procedure as described previously (19, 41). Plasmid DNA (pUC19) was used as the carrier DNA. Sixteen to eighteen hours later, the cells were mock infected or infected with *d*27-1. At 6 h postinfection (hpi), the cells were fixed and processed for immunofluorescence as described below.

For stable transfections, Vero cells plated in 25-cm^2 flasks were transfected with 0.5 μg of pSV2neo, which encodes the neomycin resistance gene (61), and a fivefold molar excess of ICP27-PK hybrid plasmid (~2.5 μg). Sheared salmon sperm DNA was used as carrier DNA. Approximately 18 h after transfection, the cells were subjected to glycerol shock (14) for 2 min. The following day, the cells were trypsinized and ~10^6 cells were plated in a 75-cm^2 flask in medium containing G418 (Geneticin; 300 mg/ml [active concentration]; GIBCO). Colonies were allowed to develop over the next 2 to 3 weeks, during which time the medium was changed every 3 to 4 days. After approximately 3 weeks, the colonies (~25 to 75 per flask) were trypsinized and pooled. For immunofluorescence experiments, the cell pools were plated on coverslips as described above, infected with *d*27-1, and processed for immunofluorescence at 6 hpi.

Immunofluorescence and immunoblot analyses. Cells were fixed with 3.4% formaldehyde, permeabilized with acetone, and processed for single or double indirect immunofluorescence as described by Quinlan et al. (40). The following primary antibodies were used: H1113 (1) and H1119, mouse monoclonal antibodies (MAbs) specific for ICP27 (Goodwin Institute for Cancer Research, Plantation, Fla.); anti-PK, a polyclonal rabbit antiserum raised against chicken PK (15); and ANA-N, a human antiserum specific for a nucleolar antigen (Sigma). H1113, H1119, anti-PK, and ANA-N were diluted 1:200, 1:600, 1:50, and 1:2, respectively, in phosphate-buffered saline (PBS). The secondary antibodies used for indirect immunofluorescence were tetramethyl rhodamine isothiocyanate-conjugated goat anti-mouse immunoglobulin G (IgG) and dichlorotriazinyl amino fluorescein-conjugated donkey anti-rabbit IgG and goat anti-human IgG (Jackson ImmunoResearch Laboratories Inc., Mississauga, Ontario, Canada), diluted 1:200, 1:150, and 1:200, respectively, in PBS. The cells were visualized with a Zeiss Axioskop 20 fluorescence microscope equipped with a Plan-Neofluar 63× objective lens.

To assess the integrity of the hybrid proteins in the stably transfected cells, immunoblot analyses were carried out. Cells were grown to confluence in 25-cm^2 flasks and infected with *d*27-1. At 6 hpi, total proteins were harvested and subjected to immunoblot analysis as described previously (46). The antibodies

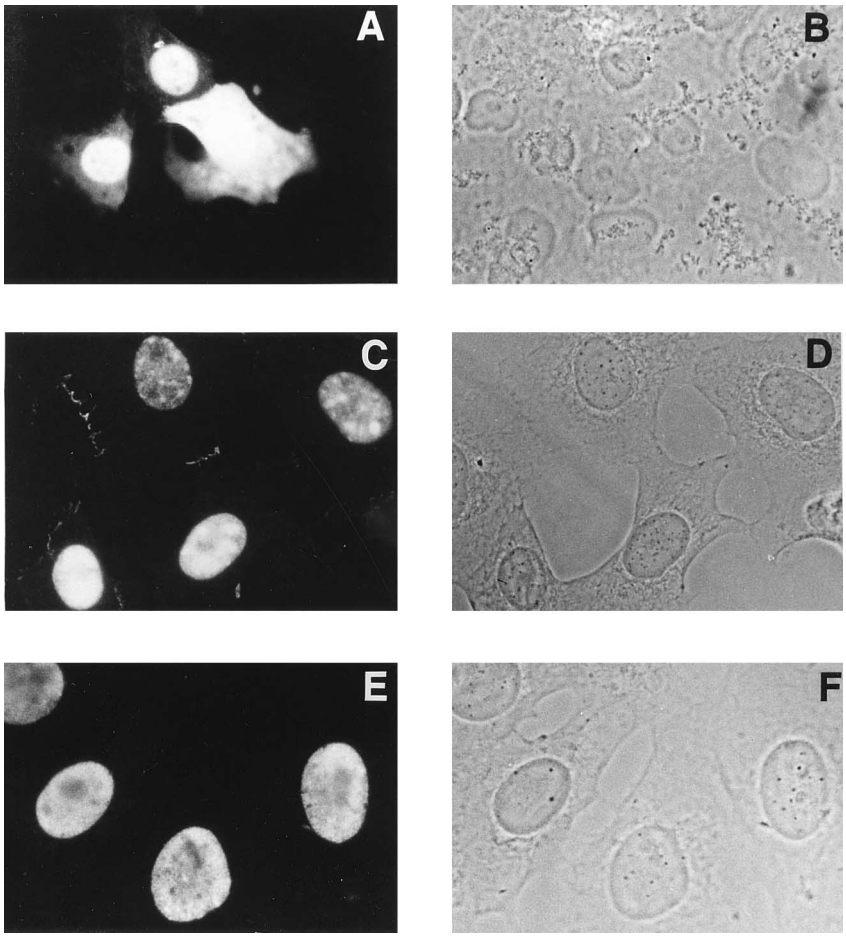

FIG. 1. ICP27 localization in transfected and infected Vero cells. (A to D) Vero cells were transiently (A and B) or stably (C and D) transfected with pM27, a plasmid encoding WT ICP27. The transfected cells were infected with *d*27-1, an HSV-1 ICP27 deletion mutant, and processed for immunofluorescence at 6 hpi. (E and F) Vero cells were infected with WT HSV-1 and processed for immunofluorescence at 5 hpi. Immunofluorescent staining was performed with H1119, an ICP27-specific MAb. Panels A, C, and E are immunofluorescence images; panels B, D, and F are the corresponding phase-contrast images.

used were H1113, H1119, and anti-PK, diluted 1:1,000, 1:1,800, and 1:1,500, respectively. Immunoreactive proteins were detected by an enhanced chemiluminescence, using commercially available reagents (ECL detection kit; Amersham).

RESULTS

A stable transfection assay for protein localization. The objective of this study was to identify the polypeptide determinants of ICP27 which mediate its nuclear and nucleolar localization. To do this, we examined which portions of ICP27 can direct a cytoplasmic protein, chicken pyruvate kinase (PK), into the cell nucleus or to nucleoli. Specifically, our experimental strategy involved the construction of hybrid ICP27-PK genes, the introduction of these genes into mammalian cells, and the localization of the expressed proteins by immunofluorescence microscopy.

We first carried out experiments to find optimal conditions for the immunofluorescent localization of proteins expressed from a transfected gene under the control of the ICP27 promoter. pM27, which contains the WT ICP27 gene (43), was used as a test plasmid. In the first set of experiments, pM27 was transfected into Vero cells, and ICP27 localization was assessed 1 day after transfection by indirect immunofluorescence. We found that the percentage of cells expressing detectable ICP27 was quite low, <0.5% of the total (data not shown). However, additional experiments showed that the problem of low expression could be overcome by infecting the cells 1 day after transfection with an HSV-1 ICP27 deletion mutant, *d*27-1 (42), and performing immunofluorescence at 6 hpi. HSV-1 infection has been previously shown to enhance the expression of transiently transfected IE promoters, through the action of the virion transactivator protein VP16 (36, 39) and perhaps other viral transactivators such as ICP0. After infection, a significant fraction (>2%) of the cells showed readily detectable expression of ICP27. However, ICP27 localization in many of the positive cells appeared aberrant. In these cells, ICP27 was distributed about equally in both the nuclear and cytoplasmic compartments (Fig. 1A and B). This pattern was in contrast to ICP27's nearly exclusive nuclear localization in WT HSV-1-infected cells (Fig. 1E and F). In addition, the fluorescent signal in many of the transfected cells was so strong that it was difficult to discern ICP27's intranuclear distribution.

The strong signals and unusual localization of ICP27 in the transiently transfected, infected cells might have resulted from nonphysiologically high expression of ICP27 due to multiple, unintegrated plasmid templates present in the cells. We therefore tested a stable transfection assay, in which the gene copy number would be expected to be lower. In this case, pM27 was cotransfected into Vero cells with pSV2neo (61), a plasmid conferring resistance to the drug G418. G418-resistant colo-

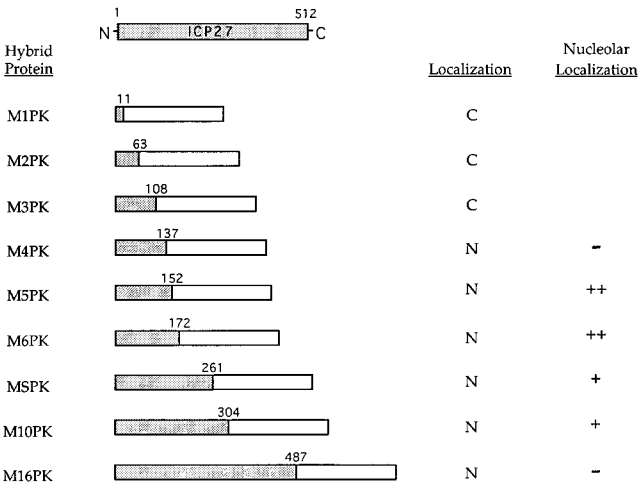

FIG. 2. Localization of PK fusion proteins containing N-terminal portions of ICP27. Plasmids which encode the PK fusion proteins illustrated were constructed. Protein sequences are represented as bars, with the stippled bars representing N-terminal ICP27 sequences and the open bars representing the PK moiety (residues 17 to 529; not to scale). The numbers above the bars denote the extent of ICP27 residues. For comparison, the WT ICP27 molecule is shown at the top. The plasmids were introduced into Vero cells by stable transfection, fusion protein expression was induced by infection with *d*27-1, and protein localization was determined at 6 hpi by immunofluorescent staining with H1119, an ICP27-specific mouse MAb, and anti-PK, a polyclonal rabbit serum. Abbreviations: C, cytoplasmic localization; N, nuclear localization; +, low level of nucleolar localization; ++, moderate nucleolar localization, with preferential staining of nucleolar peripheries; −, exclusion from nucleoli.

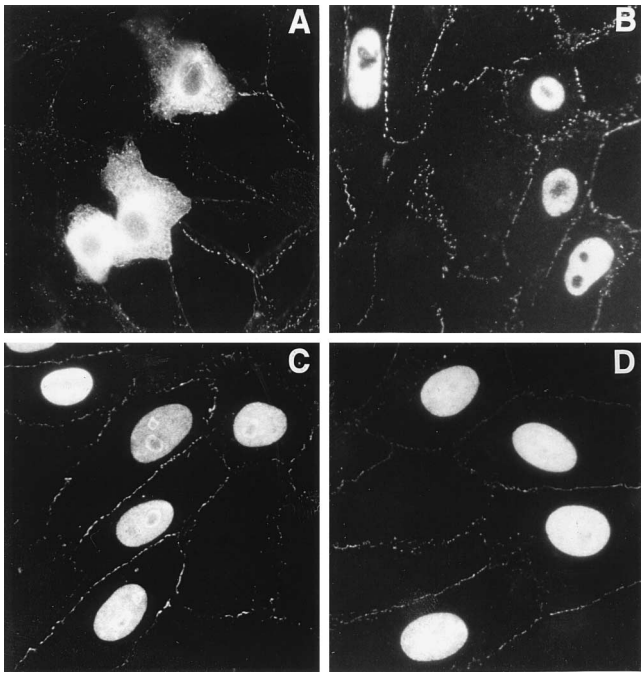

FIG. 3. Immunofluorescent staining of PK fusion proteins containing N-terminal portions of ICP27. Stably transfected Vero cells expressing various fusion proteins (M3PK [A], M16PK [B], M5PK [C], and MSPK [D]) were processed for immunofluorescence as described in the text. The photographs show immunofluorescent staining with H1119, an anti-ICP27 MAb which recognize the N-terminal 11 residues of ICP27. Note that H1119 also shows a low level of cross-reactivity with a cellular antigen which is found at or near areas of cell-cell contact, allowing the outlines of most cells to be discerned.

nies were allowed to develop over several weeks. The colonies (~100) were then pooled and tested for ICP27 expression by immunofluorescence. It was found that the pooled transfectants failed to express detectable levels of ICP27. This was not surprising, as we have previously observed that a stably transfected ICP27 gene is not constitutively expressed in Vero cells but can be induced by HSV-1 infection (42, 44). When the pooled transfectants were infected with *d*27-1, ~25% of the cells expressed ICP27 at 6 hpi (Fig. 1C and D). Significantly, ICP27 was localized nearly exclusively in nuclei and in a pattern that was very similar to that seen in the WT infection. The stable transfection assay therefore more accurately reproduces the normal physiological localization of ICP27. For this reason, we used this method in further experiments to assess the cellular localization of ICP27-PK hybrids.

Identification of an NLS in the N-terminal region of ICP27. We next constructed a set of hybrid genes in which N-terminal coding segments of the ICP27 gene were fused to the PK open reading frame (Fig. 2). The fusion genes were under the control of the ICP27 gene promoter and carried a downstream poly(A) signal also derived from the ICP27 gene. The plasmids were introduced into Vero cells by stable transfection, and the G418-resistant colonies from each transfection (ranging from ~25 to 100 colonies) were pooled. To assess the localization of the hybrid proteins, the transfectants were infected with *d*27-1 and processed for immunofluorescence at 6 hpi. Immunofluorescent staining was performed with anti-PK, a polyclonal antiserum specific for chicken PK (15), and H1119, a MAb directed against ICP27. It was found that H1119 reacted strongly with all of the hybrids, including M1PK, a protein which contains only the N-terminal 11 residues of ICP27 (Fig. 2). This result demonstrates that the epitope recognized by H1119 is contained within the ICP27's N-terminal 11 residues. This was confirmed in a separate study in which it was shown that ICP27 residues 1 to 12 could confer H1119 immunoreactivity to an unrelated poxvirus protein (65). We consistently found that immunofluorescent staining with H1119 resulted in significantly less background staining than did staining with anti-PK, although H1119 did show a low level of cross-reaction with a cellular protein localized at points of cell-cell contact. For this reason, the photographic documentation of hybrid protein localization in this study was routinely done on cells by using MAb H1119. All results, however, were confirmed by staining with the anti-PK antibody.

The cellular localizations of the N-terminal ICP27-PK hybrids are summarized in Fig. 2. Hybrid proteins possessing 11, 63, or 108 N-terminal residues of ICP27 did not show significant accumulation in the cell nucleus. An example of such cytoplasmic localization is shown for the M3PK protein in Fig. 3A. Fusion of the N-terminal 137 residues of ICP27, however, resulted in a PK hybrid that localized very efficiently to the cell nucleus. Additional constructs containing larger N-terminal portions of ICP27 were also efficiently nuclear localized. In this and all further localization experiments, control immunoblots using either anti-ICP27 or anti-PK antibodies confirmed that the expressed fusion proteins were of the approximate predicted molecular sizes and were not detectably degraded (data not shown). From these data, we conclude that the N-terminal 137 residues of ICP27 contains a strong NLS.

Although all constructs possessing 137 or more N-terminal residues of ICP27 localized to nuclei, there were some striking differences in nucleolar localization (Fig. 3B to D). The conclusions concerning nucleolar localization were reached by visually comparing the fluorescence signal with the phase-contrast image of the cells (the cell nucleoli were readily visible at

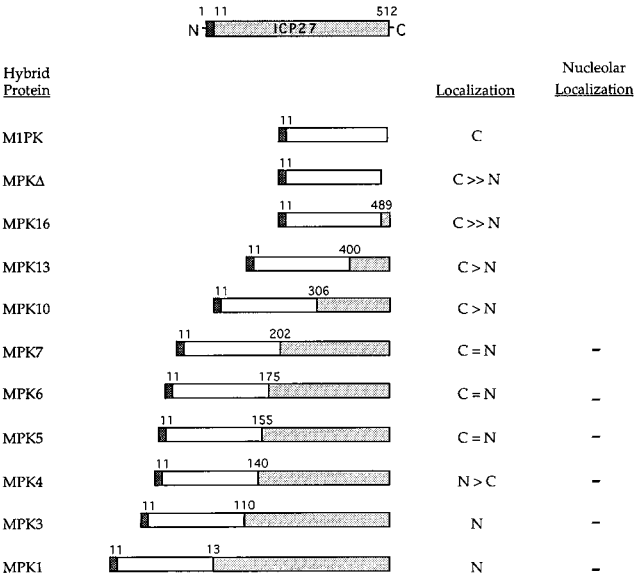

FIG. 4. Localization of PK fusion proteins containing C-terminal portions of ICP27. Plasmids which encode the ICP27-PK fusion proteins illustrated were constructed. Dark-stippled bars represent the N-terminal 11 residues of ICP27, open bars represent the PK moiety (residues 17 to 523; not to scale), and light-stippled bars represent C-terminal ICP27 sequences. The numbers above the bars denote the extent of C-terminal ICP27 residues. For comparison, the WT ICP27 is illustrated at the top. Protein localization was determined by the stable transfection assay as described in the text. Abbreviations: C, cytoplasmic localization; N, nuclear localization; −, exclusion from nucleoli.

this stage of infection). Three general patterns of nucleolar localization were observed and are summarized in Fig. 2. First, the M4PK and M16PK hybrids were largely excluded from nucleoli, as shown for M16PK in Fig. 3B. Second, the M5PK and M6PK proteins localized throughout the nucleus but showed some preferential accumulation in nucleoli, particularly around the peripheral regions. This type of staining is shown for M5PK in Fig. 3C. Finally, the MSPK and M10PK fusion proteins showed more or less even localization throughout nuclei, including significant, but not preferential, localization in nucleoli. This staining pattern is shown for MSPK in Fig. 3D.

Identification of additional weak NLSs in ICP27. The foregoing experiments demonstrated that ICP27's N-terminal 137 residues contain a strong NLS. However, if sequences C terminal to residue 137 contain additional NLSs, they would not have been identified. Therefore, we engineered additional genes in which C-terminal portions of the ICP27-coding region were fused to the C terminus of the PK gene (Fig. 4). At their N termini, all of the hybrids possessed the N-terminal 11 residues of ICP27, providing a convenient epitope tag for staining with MAb H1119. Because all of the hybrid gene constructions resulted in the replacement of PK's six C-terminal residues, we constructed an additional control plasmid, pMPKΔ, which encodes a similarly truncated PK molecule not containing attached ICP27 sequences.

The localization results for the C-terminal ICP27 fusions are summarized in Fig. 4. Truncation of PK's C terminus (MPKΔ) led to a very low but detectable amount of nuclear localization compared with M1PK, although cytoplasmic localization was still very much predominant (Fig. 5A). Fusion of ICP27 residues 489 to 512 had no further effect, but fusion of ICP27 residues 400 to 512 or 306 to 512 (MPK13 or MPK10, respectively) led to a slight increase in nuclear localization, as shown for MPK10 in Fig. 5B. Hybrid proteins containing ICP27 residues 202 to 512, 175 to 512, or 155 to 512 (MPK7, MPK6, or MPK5, respectively), on the other hand, showed significant localization in nuclei, being distributed approximately equally between the nucleus and cytoplasm. This staining pattern is shown for MPK5 in Fig. 5C. The addition of ICP27 residues 140 to 512 led to a PK hybrid (MPK4) that was predominantly nuclear, although nearly all cells still showed significant cytoplasmic localization (Fig. 5D). Complete nuclear localization was not achieved until ICP27 residues 110 to 512 or 13 to 512 had been attached to PK (MPK3 or MPK5, respectively), as shown for MPK3 in Fig. 5E. In contrast to the N-terminal ICP27-PK hybrids, all of the nuclear-localized PK hybrids in this series showed a similar nucleolar localization phenotype, being largely excluded from nucleoli.

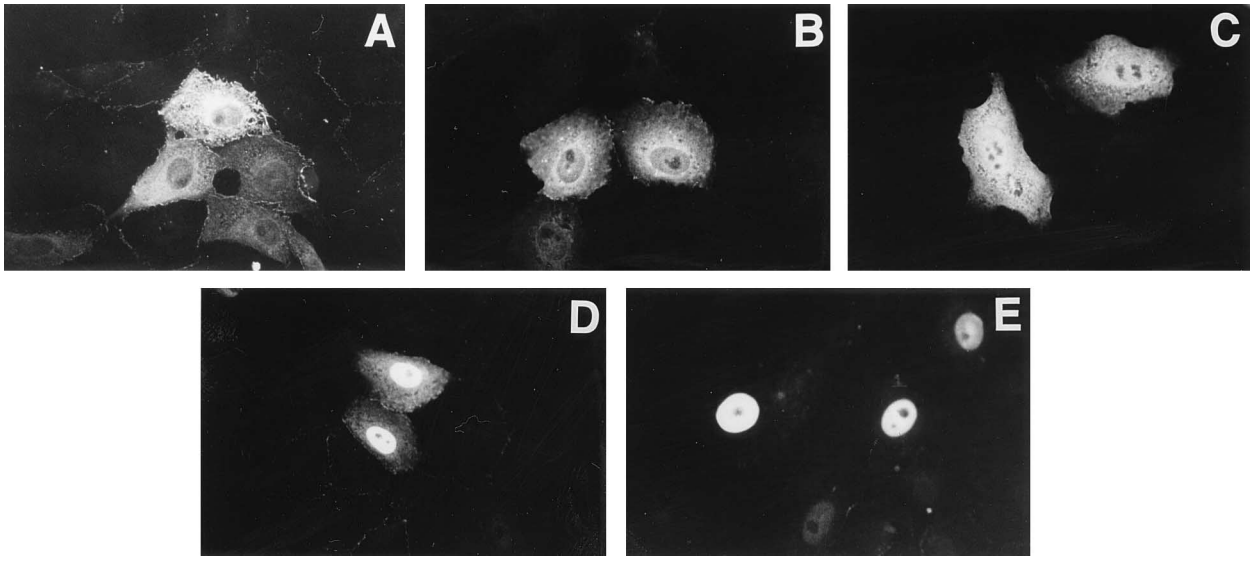

FIG. 5. Immunofluorescent staining of PK fusion proteins containing C-terminal portions of ICP27. Stably transfected Vero cells expressing various fusion proteins (MPKΔ [A], MPK10 [B], MPK5 [C], MPK4 [D], and MPK3 [E]) were processed for immunofluorescence, using MAb H1119.

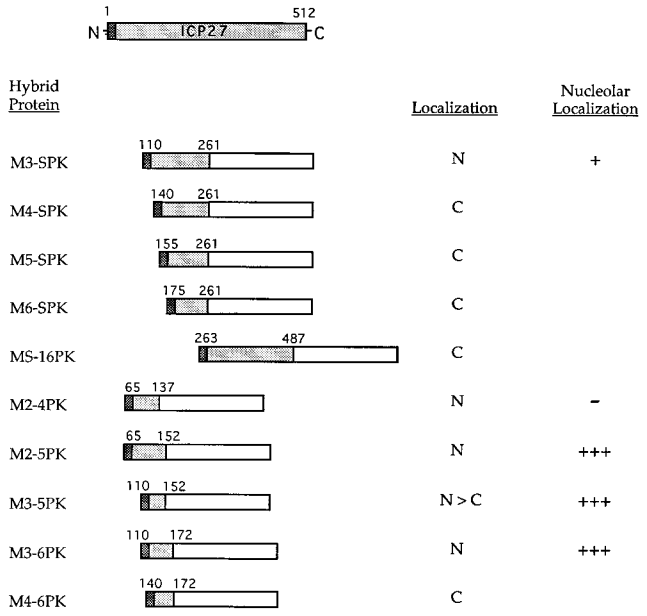

FIG. 6. Localization of PK fusion proteins containing internal portions of ICP27. Plasmids which encode the ICP27-PK fusion proteins illustrated were constructed. Dark-stippled bars represent the N-terminal 11 residues of ICP27, light-stippled bars represent internal sequences of ICP27, and open bars represent the PK moiety (residues 17 to 529; not to scale). The numbers above the bars denote the extent of internal ICP27 residues. For comparison, the WT ICP27 is shown at the top. Protein localization was determined by the stable transfection assay as described in the text. Abbreviations: C, cytoplasmic localization; N, nuclear localization; +, low level of nucleolar localization; −, exclusion from nucleoli; +++, preferential staining throughout nucleoli.

Two conclusions concerning ICP27's NLSs can be drawn from this experiment. First, the C-terminal region of ICP27, from residues 140 to 512, appears to contain several weak NLSs which function additively. However, by itself, this C-terminal portion of ICP27 is unable to confer complete nuclear localization to PK. Second, since the MPK3 hybrid localized completely to the nucleus, residues 110 to 512 contain a strong NLS.

Mapping of ICP27's strong NLS. Together, the foregoing experiments suggest that ICP27 contains a strong NLS in its N-terminal half. To more precisely map this signal, we engineered additional constructs in which internal portions of the ICP27 coding region were fused to the PK gene. The internal ICP27 sequences were linked to the H1119 epitope tag and positioned at the N terminus of the PK moiety (Fig. 6). The hybrid genes were then tested in the stable transfection assay. Hybrid proteins possessing ICP27 residues 110 to 261, 65 to 137, 65 to 152, 110 to 152, or 110 to 172 all localized efficiently to the nucleus, although M3-5PK, possessing residues 110 to 152, showed some cytoplasmic localization. None of the other constructs, however, including one that contained most of the C-terminal half of ICP27, showed significant localization in the nucleus.

The data summarized in Fig. 2, 4, and 6 demonstrate that all ICP27-PK fusion proteins which contain residues 110 to 137 of ICP27 localize efficiently to the nucleus. To test directly whether these residues define ICP27's strong NLS, we constructed two hybrid genes in which the sequence encoding these residues was fused to an epitope-tagged PK gene (Fig. 7A). One hybrid, M3-4PK, encoded a protein in which residues 110 to 137 were positioned near the N terminus of PK. The other, MPK3-4, encoded a protein in which residues 110 to 137 were tagged to the C terminus of PK. Both gene constructs were tested in the stable transfection assay, together with their appropriate control plasmids. The M3-4PK protein was cytoplasmic, indicating that residues 110 to 137 do not function as an NLS in this context. In contrast, the MPK3-4 protein showed strong, although not exclusive, nuclear localization (Fig. 7B), demonstrating that residues 110 to 137 can function as an efficient NLS in this protein context. These data, combined with the results summarized in Fig. 2, 4, and 6, demonstrate conclusively that residues 110 to 137 define a strong NLS, capable of mediating the complete nuclear localization of an unrelated protein.

Identification of a nucleolar localization sequence. The data summarized in Fig. 6 and 7A provide further information regarding the ICP27 sequences which govern nucleolar localization. Although several proteins were excluded from the nucleolus, the hybrid proteins M2-5PK, M3-5PK, and M3-6PK, which contain ICP27 residues 65 to 152, 110 to 152, and 110 to 172, respectively, all showed preferential localization in nucleoli. This staining pattern is shown for the M3-6PK hybrid in Fig. 8A and B. Double immunofluorescence labeling with a nucleolar-specific antiserum demonstrated that the observed preferential localization was indeed nucleolar (Fig. 8C).

These data demonstrate that a relatively short sequence in ICP27, from residues 110 to 152, can confer preferential nucleolar localization to PK. This sequence therefore defines a nucleolar localization signal (NuLS). Consistent with this conclusion is the fact that all of the ICP27-PK hybrids which showed some degree of nucleolar localization contained residues 110 to 152 (Fig. 2, 4, and 6). However, three proteins which did contain these residues, M16PK, MPK3, and MPK1, were excluded from nucleoli (Fig. 2 and 4). Therefore, the function of ICP27's NuLS appears dependent on protein context.

ICP27's strong NLS and NuLS function in their natural context. To see if ICP27's strong NLS functions in its natural setting, we engineered a mutant ICP27 gene, designated Md3-4, in which the sequences encoding the NLS (codons 110 to 138) were deleted. The mutant and WT genes were then tested in the stable transfection assay. As expected, WT ICP27 localized nearly exclusively in the cell nucleus (Fig. 9A). In contrast, the *d*3-4 protein was found in both the nucleus and the cytoplasm (Fig. 9B). In some cells, the mutant protein was distributed approximately equally between the nucleus and cytoplasm, while in others, nuclear localization was predominant. Therefore, deletion of ICP27's strong NLS leads to a significant defect in its nuclear accumulation. This observation demonstrates that ICP27's strong NLS functions in its natural context. However, a substantial fraction of the *d*3-4 protein was still able to enter the nucleus, indicating that other sequences in ICP27 must have NLS activity. This finding is consistent with our PK-targeting studies, wherein we found that residues 140 to 512 were able to mediate significant, but not complete, nuclear localization of PK. Together, these results indicate that ICP27 contains multiple NLSs that function with differing efficiencies.

To extend these results, we engineered two HSV-1 recombinants, *d*3-4 and *d*4-5, which had targeted disruptions in the ICP27 gene. The *d*3-4 mutant contains the *d*3-4 allele described above and thus encoded an ICP27 molecule lacking the strong NLS (residues 109 to 138). The *d*4-5 mutant, on the other hand, contains an in-frame deletion which removed codons 139 to 153. This latter alteration leaves ICP27's strong NLS intact but removes a portion of the NuLS. Both HSV-1 mutants were isolated and propagated in V27 cells (42), a derivative of Vero cells which contain a stably transfected

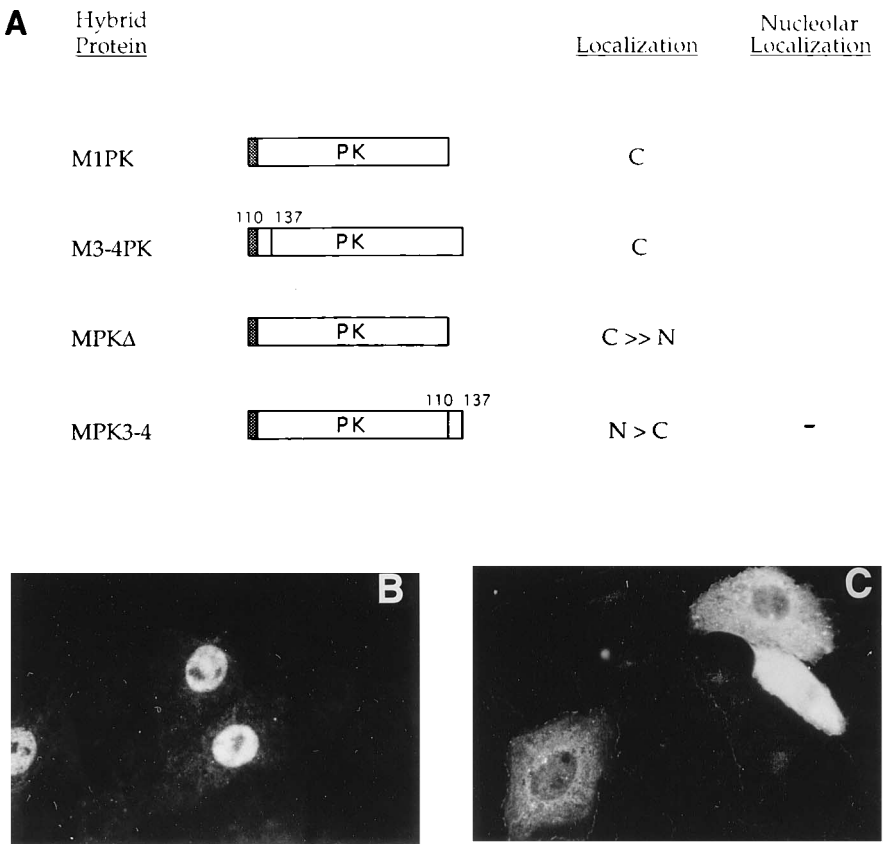

FIG. 7. Residues 110 to 137 define ICP27's strong NLS. (A) Schematic representation of ICP27-PK fusion proteins and summary of the localization results. Dark-stippled bars represent the N-terminal 11 residues of ICP27, light-stippled bars represent ICP27 residues 110 to 137, and open bars represent PK sequences. The localization of the fusion proteins was determined by the stable transfection assay as described in the text. Abbreviations: C, cytoplasmic localization; N, nuclear localization; −, exclusion from nucleoli. (B) Immunofluorescent staining of MPK3-4 with MAb H1113. (C) Immunofluorescent staining of MPKΔ with MAb H1119.

ICP27 gene and therefore complement the growth of viral ICP27 mutants. For each mutant, two independent isolates, designated a and b, were obtained and analyzed in subsequent experiments.

To see if the *d*3-4 and *d*4-5 mutants were defective for growth in Vero cells, a single-cycle growth experiment was performed. Vero cells were infected at a multiplicity of infection of 10 (using mutant virus titers determined on V27 cells), and the cultures were incubated for 24 h. As a control, the same inocula were used to infect cultures of V27 cells. The yield of infectious virus was then determined by plaque assay of the harvested cell lysates on V27 cells. The results, shown in Table 1, demonstrate that both *d*3-4 and *d*4-5 are significantly defective for growth in Vero, but not V27, cells. The *d*3-4 mutant showed a 9- to 34-fold defect in growth in Vero cells compared with its growth in V27 cells. The *d*4-5 mutant was even more deficient, replicating 58- to 84-fold less efficiently in Vero than V27 cells. In contrast, the WT parent, strain KOS1.1, showed nearly equivalent growth in Vero and V27 cells (less than a twofold difference). Therefore, the sequences deleted in *d*3-4 and *d*4-5, which encode important localization signals, are required for optimal growth of the virus in Vero cells.

We next used immunofluorescence to examine the cellular localization of the ICP27 polypeptides encoded by *d*3-4 and *d*4-5 (Fig. 10). As predicted, the *d*3-4 protein (Fig. 10C) showed a significant defect in nuclear localization compared with the WT protein (Fig. 10B), although nuclear localization was still predominant. The staining pattern was very similar to that seen earlier when the Md3-4 protein was expressed in stably transfected, *d*27-1-infected cells (Fig. 9B). The *d*4-5 ICP27 molecule, in contrast, was not deficient in its ability to localize to the cell nucleus (Fig. 10D). This finding was consistent with expectation, since this mutant ICP27 still contained the strong NLS. However, the intranuclear localization of the *d*4-5 protein differed from the WT localization in that the *d*4-5 protein was largely excluded from nucleoli (compare Fig. 10D with Fig. 10B). This was confirmed by double immunofluorescence labeling using nucleolar-specific antisera (not shown). These results demonstrate that both the strong NLS and the NuLS function during the HSV-1 infection to help mediate the nuclear and intranuclear localization of ICP27.

Mapping of an ICP27-specific MAb epitope. As demonstrated earlier, MAb H1119 maps to the N-terminal 11 residues of ICP27, since these residues alone can confer immunoreactivity (this study and reference 65). We used our collection of ICP27-PK hybrids to map the epitope for a second ICP27 MAb, H1113 (1). Key data are summarized in Table 2. The M3PK protein, possessing the N-terminal 108 residues of ICP27, did not react with H1113, but M4PK, possessing the N-terminal 137 residues, did react. MPK4, a hybrid containing ICP27 residues 140 to 512, was unable to bind to H1113, but MPK3, possessing residues 110 to 512, did bind. These data suggest that the epitope for H1113 maps to residues 110 to 137. Consistent with this hypothesis, we found that MPK3-4, which contains residues 110 to 137 tagged to the C terminus of PK,

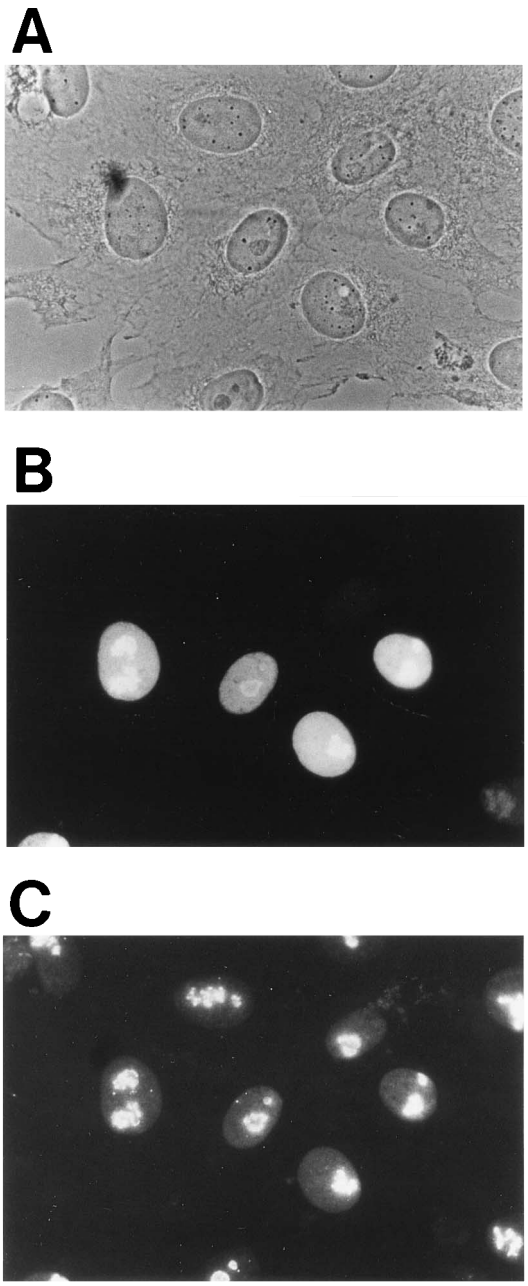

FIG. 8. Identification of a nucleolar localization signal in ICP27. Vero cells were stably transfected with pM3-6PK, which encodes a PK fusion protein containing ICP27 residues 110 to 172 (Fig. 6). The cells were infected with *d*27-1 and fixed at 6 hpi. A single field of cells processed for double-immunofluorescence microscopy is shown. (A) Phase-contrast image. (B) Immunofluorescent staining with H1119. (C) Immunofluorescent staining with ANA-N, an antiserum specific for nucleoli.

also reacted with H1113. These data do not formally exclude the possibility that residues 1 to 11 contribute to the H1113 epitope. However, we have recently engineered a mutant ICP27 mutant gene, designated M1X, which contains a stop codon inserted after codon 12 (44). This mutant gene expresses a truncated ICP27 protein, most likely by reinitiating translation at codon 50, a downstream AUG in a favorable translation context (26). Consistent with this, the M1X-encoded protein fails to react with the N-terminal specific H1119. However, it does react with H1113 (44). Together, these data demonstrate that the epitope for MAb H1113 maps to residues 110 to 137.

Residues 110 to 137 also define ICP27's strong NLS. It was therefore of interest that M3-4PK, the hybrid protein which contained the NLS sequence but failed to localize to the nucleus (Fig. 7A), was not recognized by H1113 (Table 2). Therefore, residues 110 to 137 likely have an altered conformation in this protein or are buried internally and thus are not accessible. Either possibility would explain the failure of the NLS to function in this protein context.

DISCUSSION

Identification of multiple NLSs in ICP27. Proteins enter the nucleus via the nuclear pore complexes, large proteinaceous structures which provide aqueous channels across the nuclear envelope (reviewed in reference 35). Although it is possible for small proteins to diffuse through the pore complexes, the import of large proteins (>40 to 60 kDa) is a pore-mediated, energy-dependent process (reviewed in references 16 and 56). In nearly all cases, large nuclear proteins possess one or more NLSs, which are required for import. Although there is no universal consensus for NLSs, most are short sequences which contain a high proportion of basic residues. There appear to be two major types of NLS. First, there are signals which consist of a single short basic sequence, usually less than 10 residues, which is usually flanked by a proline or glycine residue. The NLS of simian virus 40 (SV40) large T antigen, PKKKRKV, is the prototype for this type of signal (22, 23, 27). Second, there are bipartite NLSs which are composed of two clusters of basic residues separated by a 10- to 12-residue linker sequence: the first cluster contains two basic residues, while the second cluster consists of five residues, of which at least three must be basic (11). Bipartite NLSs were first characterized in the *Xenopus laevis* nucleoplasmin protein (48) and appear to be the most common type of NLS (11). In addition to these two well-characterized signals, a few other types of NLSs have been identified. For example, a relatively long (48-residue) and weakly basic sequence serves as the NLS of the human T-cell leukemia virus type I (HTLV-I) Tax protein (60).

Although the predicted sequence of ICP27 contains a relatively long arginine-rich basic region (see below) and several smaller basic regions, it does not contain sequences which are closely related to the SV40 T-antigen NLS or which fit the bipartite NLS consensus. It was thus impossible to predict, a priori, which portions of ICP27 might function as NLSs. Therefore, to map NLSs, we identified the regions of ICP27 which can direct a cytoplasmic protein, PK, into the nucleus. We found that ICP27 contains multiple signals which can mediate nuclear localization and that these signals function with differing efficiencies. First, ICP27 contains a strong NLS, mapping to residues 110 to 137, which can mediate complete nuclear localization. Second, ICP27 contains one or more weak signals, mapping to residues 140 to 512, which can mediate substantial nuclear localization but are incapable of mediating complete nuclear localization. It is relevant to point out that all of our experiments were performed in the context of HSV-1-infected cells. Thus, it is possible that other viral components contribute to the function of ICP27's NLSs and that the nuclear localization of ICP27 differs in uninfected cells.

The conclusion that residues 110 to 137 define ICP27's strong NLS comes from the PK-targeting studies which show that these residues are necessary for complete nuclear localization and that alone they can target PK into the nucleus. It is relevant to point out that although the PK fusion proteins containing residues 110 to 137 (or 110 to 152) did exhibit

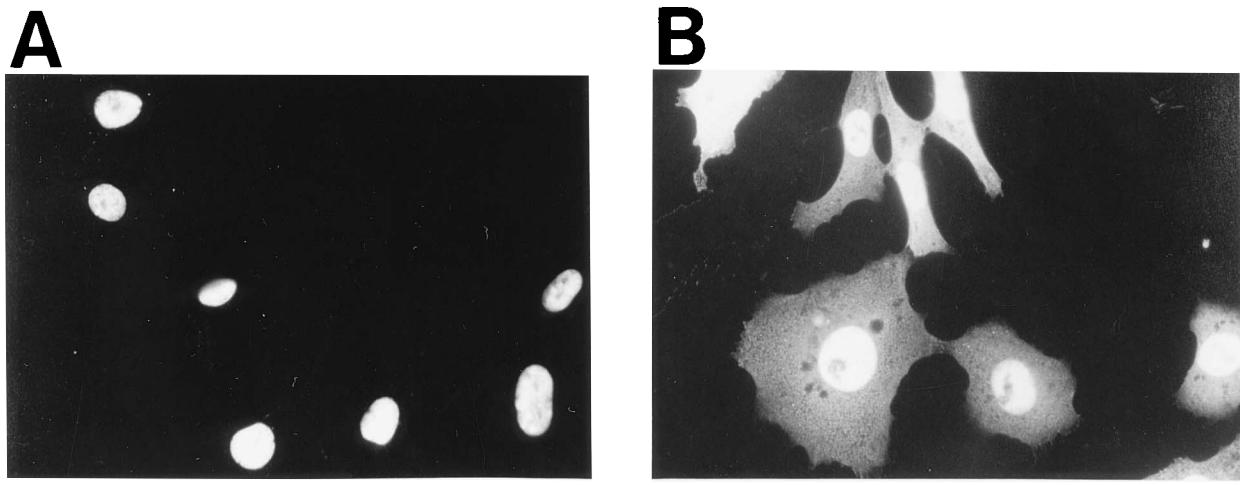

FIG. 9. Deletion of ICP27's strong NLS leads to a partial defect in nuclear localization. Vero cells were stably transfected with pM27, which encodes WT ICP27, or pMd3-4, which encodes an ICP27 molecule lacking residues 109 to 138. The cells were infected with *d*27-1 and processed for immunofluorescence at 6 hpi, using MAb H1119. (A) WT ICP27 localization from pM27. (B) Md3-4 localization.

predominant nuclear localization, some cytoplasmic localization was still evident. Thus, while residues 110 to 137 appear to define the minimal extent of the strong NLS, flanking sequences may influence its efficiency.

The results of our PK localization experiments do not formally rule out the possibility that ICP27's N-terminal 11 residues, which were included in all constructs as an epitope tag, also contribute to nuclear localization. This is very unlikely, however, for several reasons. First, the N-terminal residues did not mediate any nuclear localization on their own. In addition, an effect of the N-terminal residues on nuclear localization in our experiments would require that they interact in multiple circumstances with distal protein sequences. Finally, we have recently found that a mutant ICP27 protein which lacks its normal N terminus is still capable of efficiently localizing to the cell nucleus (44).

The predicted sequence of ICP27's strong NLS, i.e., residues 110 to 137, is ARRPSCSPERHGGKVARLQPPPTKAQPA (33, 44). Inspection of this sequence reveals a potentially significant similarity to the consensus sequence of bipartite NLSs, as exemplified by the *X. laevis* nucleoplasmin protein (Fig. 11A). ICP27's sequence, from residues 111 to 127, matches the bipartite consensus, with the exception that the second cluster of basic residues, KVARL, contains only two of five, rather than three of five, basic residues. It is noteworthy that both basic clusters, but not much of the linker sequence, is conserved in the corresponding region of the HSV-2 ICP27 (Fig. 11A). It seems likely, therefore, that ICP27's strong NLS is a member of the bipartite family. If this is so, it suggests that not all bipartite NLSs need contain three basic residues in their second basic cluster.

TABLE 1. Growth properties of HSV-1 ICP27 deletion mutants

Virus	Virus yield (PFU/cell)[a] in cell line:		Ratio, yield V27/yield Vero
	Vero	V27	
*d*3-4a	7.0	62	8.9
*d*3-4b	1.7	58	34
*d*4-5a	0.57	33	58
*d*4-5b	0.43	36	84
KOS1.1	170	104	0.61

[a] A total of 2.3×10^6 Vero cells were infected at a multiplicity of infection of 10 PFU per cell. The same inocula were used to infect parallel cultures of 2.6×10^6 V27 cells. The cultures were incubated at 37°C for 24 h and harvested. Virus yield was determined by plaque assay on V27 cells.

The function of ICP27's strong NLS is dependent on its protein context. This was demonstrated by the fact that the NLS did not function in the M3-4PK protein, in which it was positioned near the N terminus of PK. Several previous studies have shown that NLSs can be context dependent (reviewed in reference 16). For example, the NLS of SV40 large T antigen is not functional when inserted into a hydrophobic region of PK (49). The dependence of NLSs on protein context is likely based on the necessity for these sequences to fold into specific three-dimensional structures which are exposed on the protein surface. Although the three-dimensional structure of bipartite NLSs is not known, it has been proposed that the linker region loops out, juxtaposing the two clusters of basic residues. This may create a single basic surface that functions like the SV40 T-antigen class of signals (11). We found that H1113, a MAb which recognizes residues 110 to 137, did not bind to the M3-4PK protein. This observation suggests that the NLS is aberrantly folded in M3-4PK and is thus unable to interact productively with the nuclear transport machinery.

Our experiments further suggest that the C-terminal region of ICP27, from residues 140 to 512, contains multiple weak NLSs. The evidence for this comes from the analysis of PK hybrids which contain various C-terminal regions of ICP27. By increasing the extent of the attached ICP27 sequences, we were able to increase PK nuclear localization in a stepwise fashion (Fig. 4 and 5). Fusion of residues 400 to 512 led to a low level of nuclear localization, fusion of residues 202 to 512 led to ~50% nuclear localization, and fusion of residues 140 to 512 led to predominant, but incomplete, nuclear localization. The simplest explanation of these results is that residues 400 to 512, 202 to 399, and 140 to 201 each contain a weak NLS and that the effects of these NLSs are additive. In this regard, it is noteworthy that the largest effect on nuclear localization in this C-terminal fusion series resulted from the addition of residues 140 to 155 (construct MPK4). Therefore, residues 140 to 155, which include multiple arginine residues (see below), may be the most potent of ICP27's weak NLSs. However, it is possible

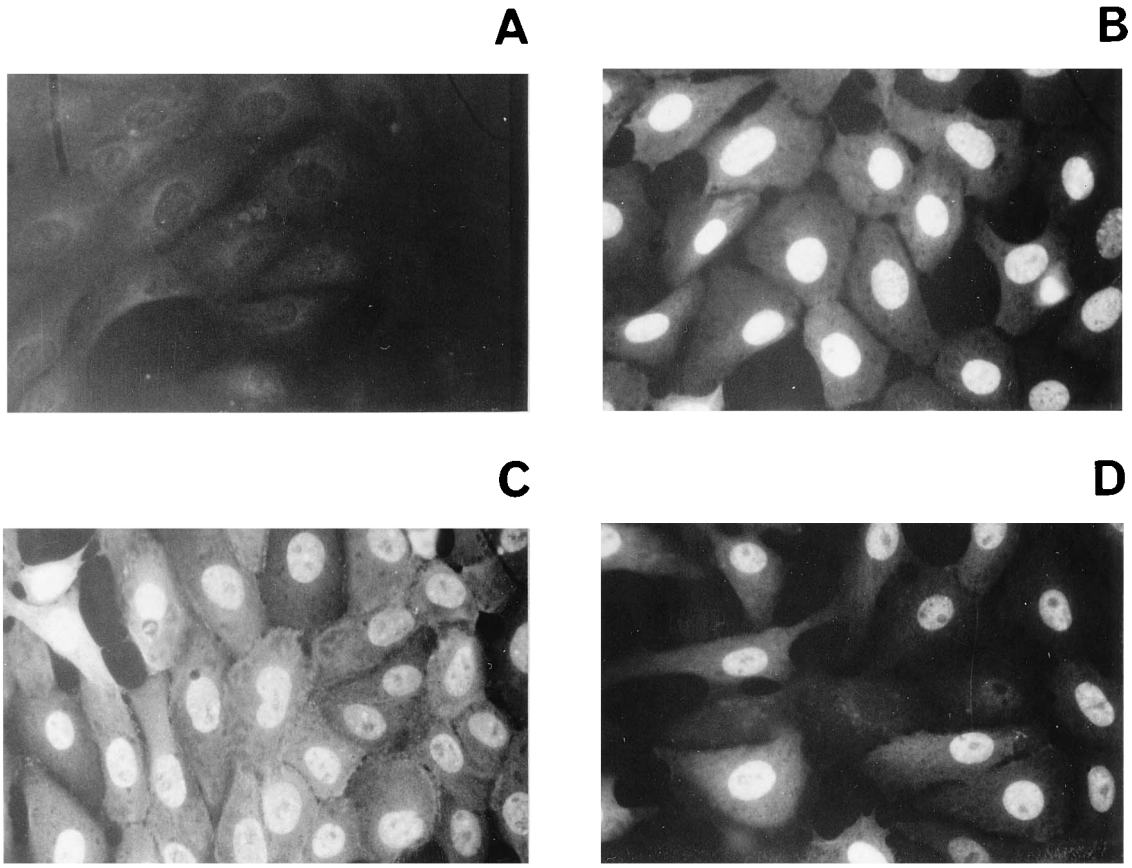

FIG. 10. Localization of mutant ICP27 molecules expressed from recombinant viruses. Vero cells were mock infected (A) or infected with WT HSV-1 (B), *d*3-4a (C), or *d*4-5a (D). At 5 hpi, the cells were fixed and processed for immunofluorescence, using MAb H1119. All four micrographs were done under the same conditions of infection, staining, exposure, and printing to allow for direct comparison.

to interpret our data in terms of a single weak NLS. In this scenario, the very C-terminal region of ICP27 could have a functional but weak NLS which is stabilized or made more efficient by adjoining ICP27 sequences. It is relevant to point out that several of the PK fusion proteins which contained the putative weak signals (e.g., M4-SPK, M5-SPK, M6-SPK, and MS-16PK) did not exhibit any nuclear localization. Therefore, protein context also appears to affect the function of ICP27's weak NLSs.

Inspection of residues 140 to 512 reveals a few basic sequences which could conceivably function as minor NLSs but no sequences which are obviously related to known NLSs. One possibility is that one or more of the weak signals are not bona fide NLSs but instead mediate an interaction with another protein which does possess an NLS. This would allow the coimport of ICP27. Such a mechanism has been proposed for the nuclear entry of several proteins, including the adenovirus DNA polymerase (68) and the *Schizosaccharomyces pombe* cdc2 kinase (5). Since our experiments were performed in the context of HSV-1-infected cells, a putative coimporter of ICP27 could be either a viral or a cellular protein.

TABLE 2. Mapping of the H1113 epitope

Protein	ICP27 residues	H1113 reactivity	Nuclear localization[a]
M3PK	1–108	−	C
M4PK	1–137	+	N
MPK4	1–11, 140–512	−	N > C
MPK3	1–11, 110–512	+	N
MPK3-4	1–11, 110–137	+	N
M3-4PK	1–11, 110–137	−	C

[a] C, cytoplasmic; N, nuclear.

To verify that ICP27's strong NLS functions in its natural context, we engineered an ICP27 gene, designated Md3-4, in which the sequences encoding the strong NLS were deleted (codons 109 to 138). The Md3-4 protein, expressed from either a stably transfected gene or an HSV-1 recombinant, showed significantly enhanced cytoplasmic localization. This finding demonstrates that the strong NLS is required for highly efficient nuclear localization. However, a substantial fraction of the Md3-4 ICP27 was still able to localize to the nucleus. This finding suggests that ICP27's weak NLSs also function in the context of the WT protein.

The presence of multiple NLSs within a single protein has been observed for several other proteins (20, 47, 64). Given the efficiency of ICP27's strong NLS, it is not clear why or if ICP27 requires secondary NLSs. One possible explanation for multiple NLSs in ICP27 is that one or more of these signals do not function in some of the cell types in which HSV-1 naturally replicates, such as neurons. Alternatively, the optimal nuclear localization of ICP27 in a given cell type may require the combined efficiency of multiple NLSs. This hypothesis is consistent with studies that show the nuclear localization of proteins can be made more efficient by increasing the number of

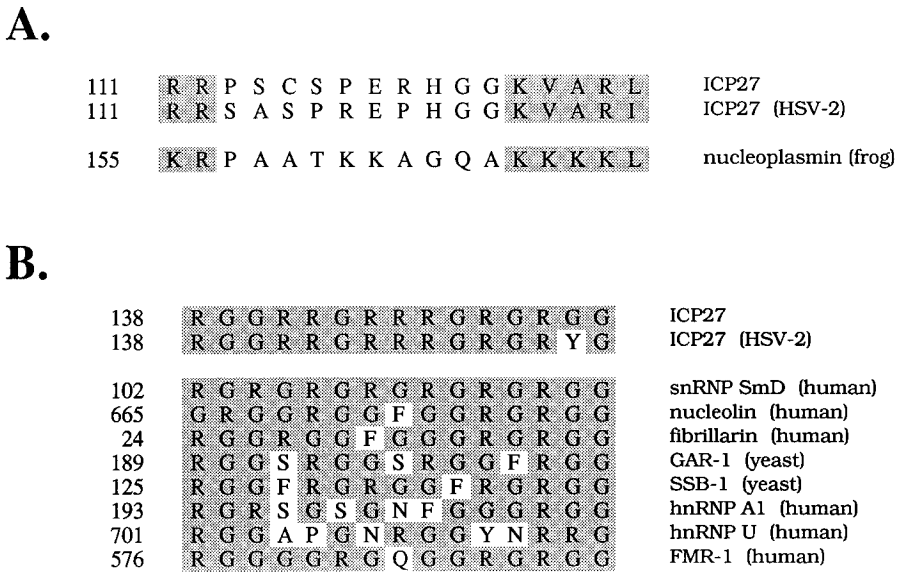

FIG. 11. Nuclear and nucleolar localization signals in ICP27. Protein sequences are shown in the single-letter code. The position in the primary sequence of the first amino acid in each line is shown at the left. (A) Similarity of ICP27's NLS to bipartite NLSs. A portion (residues 111 to 127) of ICP27's strong NLS is shown. Note that residue 119 is arginine (R) in HSV-1 strain KOS1.1 (44); this differs from the sequence of HSV-1 strain 17, in which residue 119 is a glutamine (Q) (33). The corresponding sequence from HSV-2 ICP27 (32) is shown immediately below. At the bottom, the bipartite NLS of the *X. laevis* nucleoplasmin protein (48) is shown. The two clusters of basic residues which define the bipartite NLS motif are highlighted. (B) Comparison of ICP27's RGG box region, residues 139 to 152, with the corresponding sequence from HSV-2 ICP27 and with sequences found in cellular proteins. Arginine and glycine residues are highlighted. The sequences of the cellular proteins were taken from references 2, 6, 18, 21, 24, 51, 58, and 62. snRNP, small nuclear RNP; hnRNP, heterogeneous nuclear RNP.

NLSs (12, 49; reviewed in reference 16). Our genetic analysis of the *d*3-4 mutant demonstrated that in Vero cells, the sequence encoding the strong NLS is required for optimal virus growth. Further work will be required to see if this growth deficiency results directly from the reduced level of nuclear-localized ICP27.

Identification of a nucleolar targeting signal in ICP27. Our experiments demonstrate that ICP27 possesses a relatively short sequence, mapping to residues 110 to 152, which can mediate the preferential nucleolar localization of PK. This NuLS appears to function in its natural context, since WT ICP27 shows significant, although not preferential, localization in nucleoli in both infected and uninfected cells (this study and references 25, 42, and 46). Furthermore, when a portion (residues 139 to 153) of the NuLS was deleted in the context of a recombinant virus, ICP27 was found to be excluded from nucleoli (Fig. 10D).

Similar to the function of its NLSs, the function of ICP27's NuLS is dependent on protein context. This was demonstrated by the fact that several ICP27-PK hybrid proteins which possessed the NuLS either were excluded from nucleoli (e.g., MPK1) or showed an unusual localization around the nucleolar peripheries (e.g., M5PK). It is interesting that fusion protein length appears to inversely correlate with the ability of the hybrid proteins to localize in nucleoli. Thus, it is possible that large proteins are less efficient at entering nucleoli under the conditions of our experiments. Protein context may also affect the function of the NuLS in other ways. For example, the NuLS may be nonfunctional in some contexts as a result of aberrant folding or to additional ICP27 sequences which are inhibitory to nucleolar localization.

Several studies have attempted to define protein sequences which are involved in the specific localization of proteins to nucleoli. For a few viral proteins, including HTLV-I Rex and human immunodeficiency virus Rev and Tat, relatively short NuLSs have been identified (9, 10, 57). Although a consensus sequence is not evident, two common features have been noted in these signals: the sequences are extremely rich in basic amino acids, in particular arginine residues, and they contain a functional NLS. It has been suggested that these NuLSs are modified or extended NLSs which additionally function to target proteins to nucleoli (16). However, the nucleolar localization of several cellular proteins does not appear to be mediated by such short targeting signals (28, 37, 54, 67). Rather, nucleolar localization in these cases is dependent both on an NLS and on fairly large protein regions which in some cases correspond to known functional domains. For example, the nucleolar localization of nucleolin is dependent on RNA-binding motifs which mediate its interaction with pre-rRNA (54), and the nucleolar localization of the transcription factor mUBF depends on functional DNA-binding domains and correlates with ribosomal DNA binding (28). These results have led to the proposal that the nucleolar localization of proteins is mediated not by a signal-dependent targeting mechanism but rather by stable binding of nuclear proteins to protein or nucleic acid constituents of nucleoli (28, 54, 67). The short viral NuLSs described above can be accommodated by this model if these sequences mediate a stable interaction with nucleolar components. This may be the case, since the NuLSs of human immunodeficiency virus type 1 Rev and Tat have been shown to possess RNA-binding activities (3, 7, 29, 66). Although the natural targets of Rev and Tat are believed to be specific sequences in viral pre-mRNAs, it has been suggested that related RNA sequences or structures might be present in pre-rRNA or rRNA, leading to the nucleolar accumulation of these proteins when they are expressed in cells at high levels (7, 54).

ICP27's NuLS appears similar in kind to the short basic NuLSs of Rex, Rev, and Tat in that it is composed of an NLS (residues 110 to 137) and a contiguous arginine-rich basic region (residues 138 to 152). Interestingly, the arginine-rich subdomain of ICP27's NuLS is also very rich in glycine residues

and bears similarity to a recently identified sequence motif found in a number of cellular proteins (Fig. 11B). This motif, termed the RGG box (24) or GAR (for glycine-arginine-rich) domain (18), is found in many proteins involved in nuclear RNA processing, particularly in nucleolar proteins involved in pre-rRNA processing or in heterogeneous nuclear RNPs involved in mRNA maturation (30). Most of the RGG box proteins are known or suspected to bind RNA (18, 24, 30, 58). It has been proposed that the RGG box defines a functional domain which mediates, or helps to mediate, RNA binding (24, 30). In fact, the region of the heterogeneous nuclear RNP U protein that contains the RGG box has been shown to bind to RNA (24). However, for most of the RGG box proteins, the RGG box is unlikely to be the sole determinant of RNA binding, since many also possess one or more well-characterized RNP RNA-binding motifs. It has been proposed that the RGG box of nucleolin interacts relatively nonspecifically with pre-rRNA and helps to unfold it, allowing the more specific RNP domains to gain access to their specific target sequence (17). Thus, RGG boxes may work in concert with other protein domains to mediate RNA binding.

Although there is no direct evidence that ICP27 binds to RNA, two arguments suggest that ICP27's RGG box is important in its biological functions. First, the RGG box is highly conserved in HSV-2 ICP27 (Fig. 11B), even though the overall sequence identity between the two proteins is only 67% in their N-terminal halves. Second, the *d*4-5 viral mutant, which encodes an ICP27 molecule lacking the RGG box, is unable to replicate efficiently in Vero cells (Table 1). The further phenotypic analysis of this mutant, currently in progress, should help to elucidate the biological role of the RGG box.

As reviewed earlier, studies of HSV-1 mutants suggest that ICP27 affects HSV-1 DNA replication and mRNA production, neither of which is thought to occur in nucleoli. Given additional evidence which suggests that ICP27 interacts with the cellular mRNA splicing and/or polyadenylation machinery (34, 38, 53), it is intriguing to note that a number of the cellular RGG proteins are involved in pre-mRNA maturation or processing. In fact, one of the best matches with ICP27's RGG box is a sequence from the SmD protein (51), a constituent of the spliceosomal U small nuclear RNPs (Fig. 11B). On the basis of the foregoing considerations, we hypothesize that ICP27's RGG box mediates, or helps to mediate, a specific interaction with pre-mRNA or small nuclear RNAs involved in mRNA processing. We speculate that as a secondary consequence of this RNA-binding activity, ICP27 is able to interact to some extent with rRNA, leading to the low level of nucleolar localization that is seen for the WT protein. The strong preferential nucleolar localization that is exhibited by the truncated ICP27 molecule *n*263R (42, 46) or by certain ICP27-PK hybrids (this study) might result from the deletion of specificity determinants elsewhere in the ICP27 molecule. This might relax the requirements for RNA binding by the RGG box, leading to enhanced rRNA binding and nucleolar accumulation. Currently, we are using biochemical approaches to test whether ICP27 binds to RNA and whether the RGG box is required for this putative activity.

ACKNOWLEDGMENTS

We are grateful to Min Gao for providing plasmid PK10b, to Jim Stone for helpful suggestions, and to Leslie Schiff and Mike Schultz for thoughtful reviews of the manuscript.

This research was supported by the National Cancer Institute of Canada with funds from the Canadian Cancer Society. S.A.R. is supported by a Scholarship from the Alberta Heritage Foundation for Medical Research (AHFMR). W.E.M. is supported by studentship awards from the Natural Sciences and Engineering Research Council of Canada and the AHFMR.

REFERENCES

1. **Ackermann, M., D. K. Braun, L. Pereira, and B. Roizman.** 1984. Characterization of herpes simplex virus type 1 α proteins 0, 4, and 27 with monoclonal antibodies. J. Virol. **52:**108–118.
2. **Aris, J. P., and G. Blobel.** 1991. cDNA cloning and sequencing of human fibrillarin, a conserved nucleolar protein recognized by autoimmune sera. Proc. Natl. Acad. Sci. USA **88:**931–935.
3. **Berger, J., C. Aepinus, M. Dobrovnik, B. Fleckenstein, J. Hauber, and E. Bohnlein.** 1991. Mutational analysis of functional domains in the HIV-1 rev trans-regulatory protein. Virology **183:**630–635.
4. **Block, T., and R. Jordan.** 1988. Herpes simplex virus type 1 α gene containing plasmid can inhibit expression regulated from an alpha promoter in CV-1 but not HeLa cells. Virus Res. **11:**269–279.
5. **Booher, R. N., C. E. Alfa, J. S. Hyams, and D. H. Beach.** 1989. The fission yeast cdc2/cdc13/suc1 protein kinase: regulation of catalytic activity and nuclear localization. Cell **58:**485–497.
6. **Buvoli, M., G. Biamonti, P. Tsoulfas, M. T. Bassi, A. Ghetti, S. Riva, and C. Morandi.** 1988. cDNA cloning of human hnRNP protein A1 reveals the existence of multiple mRNA isoforms. Nucleic Acids Res. **16:**3751–3770.
7. **Calnan, B. J., S. Biancalana, D. Hudson, and A. D. Frankel.** 1991. Analysis of arginine-rich peptides from the HIV tat protein reveals unusual features of RNA-protein recognition. Genes Dev. **5:**201–210.
8. **Chapman, C. J., J. D. Harris, M. A. Hardwicke, R. M. Sandri-Goldin, M. K. L. Collins, and D. S. Latchman.** 1992. Promoter-independent activation of heterologous virus gene expression by the herpes simplex virus immediate-early protein ICP27. Virology **186:**573–578.
9. **Cochrane, A. W., A. Perkins, and C. A. Rosen.** 1990. Identification of sequences important in the nucleolar localization of human immunodeficiency virus *rev*: relevance of nucleolar localization to function. J. Virol. **64:**881–885.
10. **Dang, C. V., and W. M. F. Lee.** 1989. Nuclear and nucleolar targeting sequences of c-erb-A, c-myb, N-myc, p53, HSP70, and HIV tat proteins. J. Biol. Chem. **264:**19–23.
11. **Dingwall, C., and R. A. Laskey.** 1991. Nuclear targeting sequences—a consensus? Trends Biochem. Sci. **16:**478–481.
12. **Dworetzky, S. I., R. E. Landford, and C. M. Feldherr.** 1988. The effects of variations in the number and sequence of targeting signals on nuclear uptake. J. Cell Biol. **107:**1279–1287.
13. **Everett, R. D.** 1986. The products of herpes simplex virus type 1 (HSV-1) immediate-early genes 1, 2, and 3 can activate gene expression in trans. J. Gen. Virol. **68:**2507–2513.
14. **Frost, E., and J. Williams.** 1978. Mapping temperature sensitive and host range mutants of adenovirus type 5 by marker rescue. Virology **91:**39–50.
15. **Gao, M., and D. M. Knipe.** 1992. Distal protein sequences can affect the function of a nuclear localization signal. Mol. Cell. Biol. **12:**1330–1339.
16. **Garcia-Bustos, J., J. Heitman, and M. N. Hall.** 1991. Nuclear protein localization. Biochim. Biophys. Acta **107:**83–101.
17. **Ghisolfi, L., G. Joseph, F. Amalric, and M. Erard.** 1992. The glycine-rich domain of nucleolin has an unusual supersecondary structure responsible for its RNA-helix-destabilizing properties. J. Biol. Chem. **267:**2955–2959.
18. **Girard, J.-P., H. Lehtonen, M. Caizergues-Ferrer, F. Amalric, D. Tollervey, and B. Lapeyre.** 1992. GAR1 is an essential small nucleolar RNP protein required for pre-RNA processing in yeast. EMBO J. **11:**673–682.
19. **Graham, F. L., and A. J. van der Eb.** 1973. A new technique for assay of infectivity of human adenovirus 5 DNA. Virology **52:**456–467.
20. **Hall, M. N., C. Craik, and Y. Hiraoka.** 1990. Homeodomain of yeast repressor α2 contains a nuclear localization signal. Proc. Natl. Acad. Sci. USA **87:**6954–6958.
21. **Jong, A. Y., M. W. Clark, M. Gilbert, A. Oehm, and H. L. Campbell.** 1987. *Saccharomyces cerevisiae* SSB1 protein and its relationship to nucleolar RNA-binding proteins. Mol. Cell. Biol. **7:**2947–2955.
22. **Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith.** 1984. Sequence requirements for nuclear localization of simian virus 40 large-T antigen. Nature (London) **311:**33–38.
23. **Kalderon, D., B. L. Roberts, W. D. Richardson, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear location. Cell **39:**499–509.
24. **Kiledjian, M., and G. Dreyfuss.** 1992. Primary structure and binding activity of the hnRNP U protein: binding RNA through RGG box. EMBO J. **11:**2655–2664.
25. **Knipe, D. M., D. Senechek, S. A. Rice, and J. L. Smith.** 1987. Stages in the nuclear association of the herpes simplex transcriptional activator protein ICP4. J. Virol. **61:**276–284.
26. **Kozak, M.** 1991. Structural features in eukaryotic mRNAs that modulate the initiation of translation. J. Biol. Chem. **266:**19867–19870.
27. **Lanford, R. E., and J. S. Butel.** 1984. Construction and characterization of an SV40 mutant defective in nuclear transport of T antigen. Cell **37:**801–813.
28. **Maeda, Y., K. Hisatake, T. Kondo, K.-I. Hanada, C.-Z. Song, T. Nishimura, and M. Muramatsu.** 1992. Mouse rRNA gene transcription factor mUBF requires both HMG-box 1 and an acidic tail for nucleolar accumulation: molecular analysis of the nucleolar targeting mechanism. EMBO J. **11:**3695–3704.

29. **Malim, M. H., and B. R. Cullen.** 1991. HIV-1 structural gene expression requires the binding of multiple rev monomers to the viral RRE: implications for HIV-1 latency. Cell **65:**241–248.
30. **Mattaj, I. W.** 1993. RNA recognition: a family matter? Cell **73:**837–840.
31. **McCarthy, A. M., L. McMahan, and P. A. Schaffer.** 1989. Herpes simplex virus type 1 ICP27 deletion mutants exhibit altered patterns of transcription and are DNA deficient. J. Virol. **63:**18–27.
32. **McGeoch, D. J., C. Cunningham, G. McIntyre, and A. Dolan.** 1991. Comparative sequence analysis of the long repeat regions and adjoining parts of the long unique regions in the genomes of herpes simplex viruses types 1 and 2. J. Gen. Virol. **72:**3057–3075.
33. **McGeoch, D. J., M. A. Dalrymple, A. J. Davison, A. Dolan, M. C. Frame, D. McNab, L. J. Perry, J. E. Scott, and P. Taylor.** 1988. The complete DNA sequence of the long unique region in the genome of herpes simplex virus type 1. J. Gen. Virol. **69:**1531–1574.
34. **McLauchlan, J., A. Phelan, C. Loney, R. M. Sandri-Goldin, and J. B. Clements.** 1992. Herpes simplex virus IE63 acts at the posttranscriptional level to stimulate viral mRNA 3′ processing. J. Virol. **66:**6939–6945.
35. **Newmeyer, D. D.** 1993. The nuclear pore complex and nucleocytoplasmic transport. Curr. Opin. Cell Biol. **5:**395–407.
36. **O'Hare, P., and G. S. Hayward.** 1984. Expression of recombinant genes containing herpes simplex virus delayed-early and immediate-early regulatory regions and *trans* activation by herpesvirus infection. J. Virol. **52:**522–531.
37. **Peculis, B. A., and J. G. Gall.** 1992. Localization of the nucleolar protein N038 in amphibian oocytes. J. Cell Biol. **116:**114.
38. **Phelan, A., M. Carmo-Fonseca, J. McLauchlan, A. I. Lamond, and J. B. Clements.** 1993. A herpes simplex virus type 1 immediate-early gene product, IE63, regulates small nuclear ribonucleoprotein distribution. Proc. Natl. Acad. Sci. USA **90:**9056–9060.
39. **Post, L. E., S. Mackem, and B. Roizman.** 1981. Regulation of α genes by herpes simplex virus: expression of chimeric genes produced by fusion of thymidine kinase with alpha gene promoters. Cell **24:**555–565.
40. **Quinlan, M. P., L. B. Chen, and D. M. Knipe.** 1984. The intranuclear location of a herpes simplex virus DNA binding protein is determined by the status of viral DNA replication. Cell **36:**857–868.
41. **Rice, S. A., and D. M. Knipe.** 1988. Gene-specific transactivation by herpes simplex virus type 1 ICP27. J. Virol. **62:**3814–3823.
42. **Rice, S. A., and D. M. Knipe.** 1990. Genetic evidence for two distinct transactivation functions of the herpes simplex virus α protein ICP27. J. Virol. **64:**1704–1715.
43. **Rice, S. A., and V. Lam.** 1994. Amino acid substitution mutations in the herpes simplex virus ICP27 protein define an essential gene regulation function. J. Virol. **68:**823–833.
44. **Rice, S. A., and V. Lam.** Unpublished data.
45. **Rice, S. A., V. Lam, and D. M. Knipe.** 1993. The acidic amino-terminal region of herpes simplex virus type I alpha protein ICP27 is required for an essential lytic function. J. Virol. **67:**1778–1787.
46. **Rice, S. A., L. Su, and D. M. Knipe.** 1989. Herpes simplex virus α protein ICP27 possesses separable positive and negative regulatory activities. J. Virol. **63:**3399–3407.
47. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signal in polyoma virus large-T. Cell **44:**77–85.
48. **Robbins, J., S. M. Dilworth, R. A. Laskey, and C. Dingwall.** 1991. Two interdependent basic domains in nucleoplasmin nuclear targeting sequence: identification of a class of bipartite nuclear targeting sequence. Cell **64:**615–623.
49. **Roberts, B. L., W. D. Richardson, and A. E. Smith.** 1987. The effect of protein context on nuclear location signal function. Cell **50:**465–475.
50. **Roizman, B., and A. E. Sears.** 1990. Herpes simplex viruses and their replication, p. 1795–1842. *In* B. N. Fields and D. M. Knipe (ed.), Fundamental virology. Raven Press, New York.
51. **Rokeach, L. A., J. A. Haselby, and S. O. Hoch.** 1988. Molecular cloning of a cDNA encoding the human Sm-D autoantigen. Proc. Natl. Acad. Sci. USA **85:**4832–4836.
52. **Sacks, W. R., C. C. Greene, D. P. Aschman, and P. A. Schaffer.** 1985. Herpes simplex virus type 1 ICP27 is an essential regulatory protein. J. Virol. **55:** 796–805.
53. **Sandri-Goldin, R. M., and G. E. Mendoza.** 1992. A herpesvirus regulatory protein appears to act post-transcriptionally by affecting mRNA processing. Genes Dev. **6:**848–863.
54. **Schmidt-Zachmann, M. S., and E. A. Nigg.** 1993. Protein localization to the nucleolus: a search for targeting domains in nucleolin. J. Cell Sci. **105:**799–806.
55. **Sekulovich, R. E., K. Leary, and R. M. Sandri-Goldin.** 1988. The herpes simplex virus type 1 α protein can act as a *trans*-repressor or a *trans*-activator in combination with ICP4 and ICP0. J. Virol. **62:**4510–4522.
56. **Silver, P. A.** 1991. How proteins enter the nucleus. Cell **64:**489–497.
57. **Siomi, H., H. Shida, S. H. Nam, T. Noskaka, M. Maki, and M. Hatanaka.** 1988. Sequence requirements for nucleolar localization of human T cell leukemia virus type I pX protein, which regulates viral RNA processing. Cell **55:**197–209.
58. **Siomi, H., M. C. Siomi, R. L. Nussbaum, and G. Dreyfuss.** 1993. The protein product of the fragile X gene, FMR1, has characteristics of an RNA-binding protein. Cell **74:**291–298.
59. **Smith, I. L., M. A. Hardwicke, and R. M. Sandri-Goldin.** 1992. Evidence that the herpes simplex virus immediate early protein ICP27 acts post-transcriptionally during infection to regulate gene expression. Virology **186:**74–86.
60. **Smith, M. R., and W. C. Greene.** 1992. Characterization of a novel nuclear localization signal in the HTLV-1 Tax transactivator protein. Virology **187:** 316–320.
61. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. **1:**327–341.
62. **Srivastiva, M., P. J. Fleming, H. B. Pollard, and A. L. Burns.** 1989. Cloning and sequencing of human nucleolin cDNA. FEBS Lett. **250:**99–105.
63. **Su, L., and D. M. Knipe.** 1989. Herpes simplex virus α protein ICP27 can inhibit or augment viral gene transactivation. Virology **170:**496–504.
64. **Underwood, M. R., and H. Fried.** 1990. Characterization of nuclear localizing sequences derived from yeast ribosomal protein L29. EMBO J. **9:**91–100.
65. **Upton, C. A., L. Schiff, S. A. Rice, T. Dowdeswell, X. Yang, and G. McFadden.** 1994. A novel poxvirus protein binds zinc through a ring finger motif and localizes in virus factories. J. Virol. **68:**4186–4195.
66. **Weeks, K. M., C. Ampe, S. C. Schultz, T. A. Steitz, and D. M. Crothers.** 1990. Fragments of HIV-1 tat protein specifically bind TAR RNA. Science **249:** 1281–1285.
67. **Yan, C., and T. Melese.** 1993. Multiple regions of NSR1 are sufficient for accumulation of a fusion protein within the nucleolus. J. Cell Biol. **123:**1081–1091.
68. **Zhao, L.-J., and R. Padmanabhan.** 1988. Nuclear transport of adenovirus DNA polymerase is facilitated by interaction with preterminal protein. Cell **55:**1005–1015.