JOURNAL OF VIROLOGY, Aug. 1988, p. 3020–3026
0022-538X/88/083020-07$02.00/0

Vol. 62, No. 8

Two Nuclear Location Signals in the Influenza Virus NS1 Nonstructural Protein

DEBORAH GREENSPAN, PETER PALESE, AND MARK KRYSTAL*

Department of Microbiology, Mount Sinai School of Medicine, City University of New York, New York, New York 10029

Received 3 March 1988/Accepted 4 May 1988

The NS1 protein of influenza A virus has been shown to enter and accumulate in the nuclei of virus-infected cells independently of any other influenza viral protein. Therefore, the NS1 protein contains within its polypeptide sequence the information that codes for its nuclear localization. To define the nuclear signal of the NS1 protein, a series of recombinant simian virus 40 vectors that express deletion mutants or fusion proteins was constructed. Analysis of the proteins expressed resulted in identification of two regions of the NS1 protein which affect its cellular location. Nuclear localization signal 1 (NLS1) contains the stretch of basic amino acids Asp-Arg-Leu-Arg-Arg (codons 34 to 38). This sequence is conserved in all NS1 proteins of influenza A viruses, as well as in that of influenza B viruses. NLS2 is defined within the region between amino acids 203 and 237. This domain is present in the NS1 proteins of most influenza A virus strains. NLS1 and NLS2 contain basic amino acids and are similiar to previously defined nuclear signal sequences of other proteins.

Sequences have been characterized which can direct proteins to discrete cellular compartments, e.g., membranes, mitochondria, chloroplasts, and most recently the nucleus (7, 18, 41, 42, 44). Nuclear location signals consisting of short stretches of basic amino acids have been defined for many proteins, including the simian virus 40 (SV40) T antigen (19, 20, 26) and VP1 protein (46), the rat glucocortocoid receptor (39), nucleoplasmin (12), and adenovirus E1A protein (29). Conversely, nonbasic amino acid domains have also been implicated in the nuclear migration of certain proteins, such as the influenza virus NP protein (10) and the *Saccharomyces cerevisiae matα2* protein (16). In addition, it has been reported that some proteins, such as polyomavirus large T antigen, contain two regions within their sequences which may act as nuclear localization signals (39, 40). The NS1 and NS2 proteins of influenza virus-infected cells are nuclear proteins (3, 14, 21, 27, 30, 49). When CV-1 cells are infected with a recombinant SV40 containing the NS gene, both the NS1 and NS2 proteins are expressed and migrate into the nucleus (14, 25). Additionally, an NS1 protein which is expressed from a recombinant vaccinia virus enters and accumulates in the nuclei of infected cells independently of other influenza virus proteins (43). In this system, splicing does not occur and the NS2 protein is not detectable. It therefore seems likely that the information necessary for the nuclear location of NS1 is encoded within its sequences.

In a continuing effort to examine the functional domains of the NS1 polypeptide, we expressed altered NS1 proteins to define its nuclear signal. In this study, we identified two different domains containing nuclear localization signals. Either region alone is sufficient to direct the NS1 protein to the nucleus, but in the absence of both of these regions the protein is localized in the cytoplasm.

MATERIALS AND METHODS

Plasmids and cells. The cDNAs of the NS genes of A/PR/8/34 and A/Alaska/6/77 were cloned into the *Eco*RI site of pBR322 and are designated pAPR801 and PQ7, respectively (1, 5); plasmid pSPR1 is derived from pSP64 and pSP65 (23); pAII-SVL2 is an SV40 late region expression vector (14, 15). Plasmid pMCα19, a pBR322 derivative which contains the cDNA of the chimpanzee α-globin gene, was obtained from Stephen Liebhaber (28). All plasmids were grown in *Escherichia coli* HB101 cells. DNA manipulations, including restriction enzyme digestion, ligation of DNA fragments, and transformation of *E. coli* with plasmid DNA, were done by standard procedures (31).

Construction of NS1 deletion mutants. Synthetic oligonucleotides were prepared by using either a Biosearch SAM1 oligonucleotide synthesizer or an Applied Biosystems model 380B synthesizer. For construction of pNS-Δ2-7, which lacks codons 2 to 7 of the NS gene, plasmid pAPR801 was used. Plasmid pAPR801 was digested with *Hin*dIII (Fig. 1), the fragments were end filled by using DNA polymerase I large fragment, and *Eco*RI linkers containing an initiation codon (CATTACGAATTCGTAATG) were added. The *Eco*RI fragment containing the desired NS sequences was isolated on agarose gels with DEAE-paper (45) and cloned into the *Eco*RI site of plasmid pSPR1. Constructions pNS-Δ2-51 and pNS-Δ2-81 were generated by digestion of plasmid pQ7 with either *Xba*I or *Nco*I, respectively. *Xba*I and *Nco*I cut within codon 51 and 81, respectively, in the NS gene (Fig. 1). The fragments created by *Xba*I digestion were end filled, whereas the *Nco*I-digested DNA was treated with S1 nuclease. These DNAs were then ligated to the *Eco*RI linker described above. After *Eco*RI digestion of the DNA, the fragments containing the 3′ portion of the NS sequences were isolated and independently cloned into the *Eco*RI site of pSPR1. The fourth construct, pNS-Δ34-38, was created by digestion of pQ7 with *Eco*RI to release the NS insert. The isolated insert was digested with *Sau*3AI (Fig. 1), and fragments were recloned into *Eco*RI-digested pSPR1. This allows removal of an internal 15-base-pair *Sau*3AI fragment. The NS insert isolated from *Eco*RI-digested pQ7 was also ligated directly into pSPR1 to create pNS-Ala. pNS-Ala was then used to generate the fifth construct, pNS-Δ82-237. For this purpose, plasmid pNS-Ala was *Nco*I digested, end filled, and religated. This causes a frameshift mutation in the

* Corresponding author.

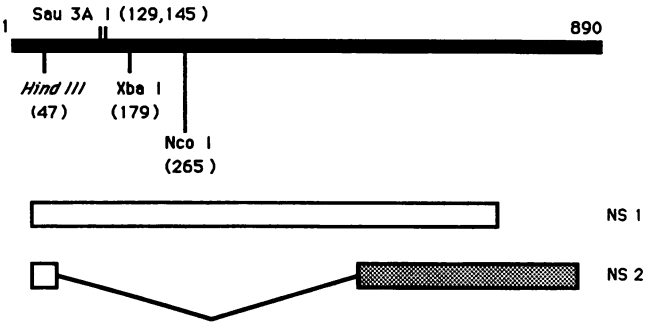

FIG. 1. Restriction enzyme sites used in construction of NS gene mutants. The restriction enzyme sites that were used for construction of plasmids expression NS1 deletion mutants and NS1–α-globin fusion proteins are shown. The solid line indicates the cDNA of the 890-nucleotide NS gene of influenza virus. The nucleotide positions of the restriction enzyme sites are in parentheses. The *Hin*dIII site is present only in pAPR801, which contains a cDNA copy of the NS gene of influenza virus A/PR/8/34; the *Xba*I site is found only in pQ7, a recombinant pBR322 containing the cDNA of the A/Alaska/6/77 NS gene. The organization of the NS1 and NS2 proteins is also shown. The NS2 protein shares the first 10 amino acids with the NS1 protein (open boxes), and the carboxyl-terminal sequences are derived from a plus one open reading frame (stippled box).

NS1-coding sequence after codon 81. All plasmid DNAs were checked for correct construction by restriction enzyme analysis and partial RNA sequencing with SP6 polymerase (38). All deleted genes were first engineered into the *Eco*RI site of pSPRI so that NS inserts could then be released by *Bam*HI digestion (23). Deletion mutants pNS-Δ2-7, pNS-Δ2-51, pNS-Δ2-81, pNS-Δ34-38, and pNS-Δ82-237 were *Bam*HI digested to liberate the cDNA inserts, and the isolated inserts were ligated into *Bam*HI-digested pA11-SVL2 to generate the recombinant SV40 vectors pSV-NS-Δ2-7, pSV-NS-Δ2-51, pSV-NS-Δ2-81, pSV-NS-Δ34-38, and pSV-NS-Δ82-237, respectively.

Construction of NS1–α-globin chimerae. pNS-Ala and pNS-Δ34-38 were digested with *Eco*RI to release the cDNA inserts. The inserts were end filled, and *Bgl*II linkers were added (CAGATCTG) (New England BioLabs, Beverly, Mass.). After digestion with *Bgl*II to remove excess linkers, the DNAs were digested with *Nco*I, which cuts the NS sequence after nucleotide 265 (codon 81). The fragments that encode the 5′ end of the NS gene were isolated. Plasmid pMCα19, containing the entire α-globin cDNA sequence, was *Nco*I and *Mbo*I digested. *Nco*I cuts within the α-globin initiation codon, whereas *Mbo*I cuts downstream of the stop codon within vector sequences. The isolated NS- and α-globin-containing fragments then were ligated with *Bam*HI-digested pA11-SVL2 to generate the recombinant SV40 vectors pSV-NSglobin-fpI(1-81) and pSV-NSglobin-fpII(1-33/39-81).

Constructs that encode NS1 proteins with carboxyl-terminal deletions. To obtain vectors that express NS proteins with COOH terminal deletions, pSV-NS-Δ34-38 was digested with *Hpa*I. This enyzme cuts in SV40 sequences downstream of the NS sequences. The DNA was then treated for various times with *BAL* 31 exonuclease, and synthetic *Bam*HI linkers (TAGCTAACTAGGATCCTAGTTAGCTA) containing stop codons in all three reading frames were ligated to the DNA. The DNA was then *Kpn*I (a single *Kpn*I site is located upstream of the NS sequences) and *Bam*HI digested, and fragments from 600 to 900 base pairs long were isolated. This mixture of DNA fragments was cloned into *Kpn*I/*Bam*HI-digested pA11-SVL2. The exact end of the coding region in each deletion mutant was determined by double-stranded DNA sequencing (9).

Expression of recombinant NS1 proteins. For transient expression, approximately 4×10^5 COS-1 cells were seeded on cover slips which were placed in a 60-mm-diameter tissue culture dish. On the next day, 5 to 10 μg of one of the recombinant SV40-derived expression vectors was transfected into cells by either the calcium phosphate precipitation procedure described by Parker and Stark (36) or that modified by Chen and Okayama (8). Recombinant SV40 stocks were also made and used in this study. These stocks were made by cotransfection of CV-1 cells with the recombinant SV40 NS DNAs and an early region deletion mutant, pSVr-Ins7, as described previously (14). The protocol requires removal of pBR322 plasmid sequences from pSV-NS constructions by digestion with *Xba*I. However, since the A/Alaska/6/77 NS gene contains an internal *Xba*I site, recombinant pA11-SVL2 vectors containing this gene were digested with *Aha*II instead of *Xba*I to eliminate the pBR322 sequences.

Immunofluorescence. Either subconfluent CV-1 monolayers infected with recombinant SV40 or transfected COS-1 cells were used for immunofluorescence. Cells were fixed 48 or 72 h postinfection or posttransfection, as described below (14). Briefly, cells were treated with 1% formaldehyde in phosphate-buffered saline (PBS; pH 7.5) for 1 h, washed with PBS, and permeabilized with 0.1% Triton X-100 in PBS for 5 min. This step was followed by methanol-acetone (1:1) treatment at −20°C for 5 min. Samples were stored at 4°C in PBS until stained. Cover slips were stained by incubation at 37°C for 30 min with a 1:100 dilution of rabbit anti-NS1 serum in PBS containing 5% normal goat serum. Cells were then washed with PBS and treated with fluorescein-conjugated goat anti-rabbit immunoglobulin G (Cappel Biomedical, Inc.) for 30 min at 37° C. Samples were washed, mounted and viewed on a Zeiss fluorescent microscope, model D-7082.

RESULTS

Expression of deletion mutants of the NS1 gene. The NS1 protein has been expressed in recombinant SV40-infected CV-1 cells and shown to localize in the nucleus (14, 25). In this follow-up study, mutants with deletions in the NS1-coding region were constructed, and the effects of these deletions on the migration of the protein were investigated (Fig. 2). First it was thought that, since both the NS1 and NS2 proteins localize to the nucleus in virus-infected cells, a nuclear localization signal may be present in the region common to both proteins. The NS1 and NS2 polypeptides are encoded by overlapping reading frames which share 10 amino acids at their amino termini but differ over the rest of their sequences (Fig. 1) (24). To obtain a deletion mutant of NS1 lacking the amino-terminal amino acids, plasmid pSV-NS-Δ2-7 was constructed as described in Materials and Methods. Plasmid pSV-NS-Δ2-7 then was used to obtain recombinant SV40 stocks. CV-1 cells were infected with these recombinant SV40 stocks and labeled with [^{35}S]methionine at 3 days postinfection. Cells lysates were analyzed by sodium dodecyl sulfate-gel electrophoresis and autoradiography (Fig. 3). Cells infected with SV-NS-Δ2-7 virus (Fig. 3, lane 3) expressed NS1 and NS2 proteins which migrated slightly faster than wild-type NS1 and NS2 proteins (Fig. 3, lanes 1 and 2). The faster migration of these proteins is explained by the deletion of amino acids 2 to 7 in the mutant

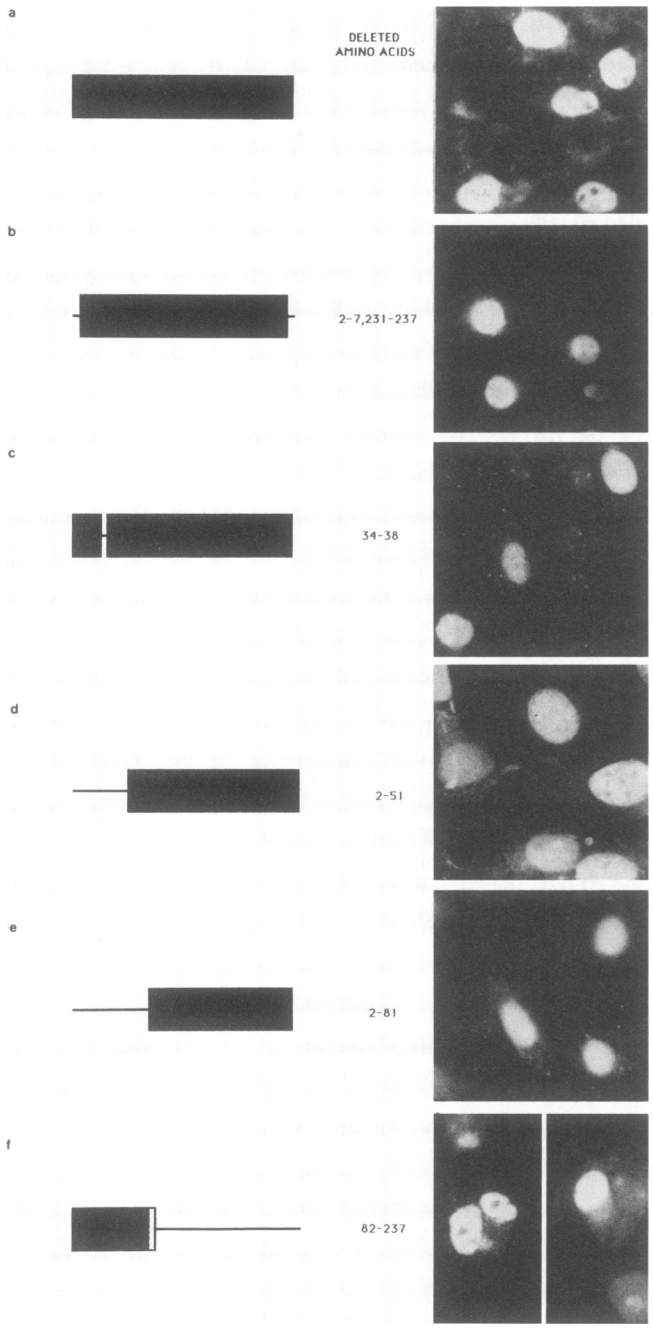

FIG. 2. Expression of deletion mutants of the NS1 gene. A schematic diagram illustrating the expressed sequence from each construct is shown. Solid boxes represent expressed proteins. Thin lines show the deleted portion of the protein, and the numbers to the side of each protein indicate deleted amino acids. Immunofluorescence data of recombinant SV40-infected CV-1 cells are shown beside each construction. Cells were fixed at 72 h postinfection and stained with polyclonal monospecific rabbit anti-NS1 serum (49). The viruses used were as follows (panels): a, SV-NS; b, SV-NS-Δ2-7; c, SV-NS-Δ34-38; d, SV-NS-Δ2-51; e, SV-NS-Δ2-81; f, SV-NS-Δ82-237. In panel f, the NS1 protein of SV-NS-Δ82-237-infected cells contains an additional 11 nonviral amino acids (open box).

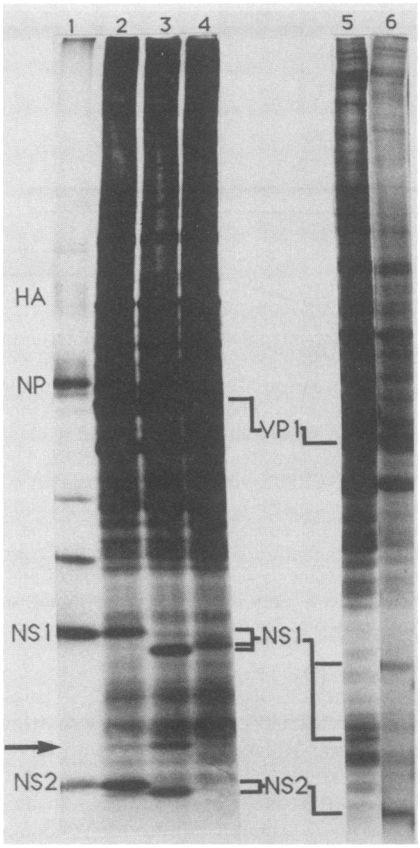

FIG. 3. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of viral proteins. CV-1 cells were infected with virus and pulse-labeled for 1 h with [^{35}S]methionine at 5 h postinfection for influenza virus or 72 h postinfection for the SV40 recombinant. Samples were electrophoresed on 7 to 14% polyacrylamide gels (36). Gels were dried and autoradiographed. The influenza virus and recombinant SV40 used for infection were as follows (lanes): 1, A/PR/8/34; 2, SV-NS; 3, SV-NS-Δ2-7; 4, SV-NS-Δ34-38; 5, SV-NS-Δ2-51; 6, SV-NS. The arrow indicates a presumed breakdown product of the NS1 protein present in SV-NS-Δ2-7-infected cells. HA, Hemagglutinin.

proteins. The truncated NS1 and NS2 proteins could be immunoprecipitated with specific polyclonal antiserum raised to bacterially synthesized NS1 and NS2 proteins (14, 49), respectively (data not shown). An additional band which migrated slightly slower than the NS2 protein was seen in Fig. 3, lane 3 (arrow). This protein band was immunoprecipitated with antiserum to NS1 and presumably represents a breakdown product of the mutant NS1 polypeptide. CV-1 cells infected with SV-NS-Δ2-7 virus were also analyzed by indirect-immunofluorescence experiments. The truncated NS1 protein, like full-length NS1 protein, was localized to the nucleus (Fig. 2, a and b). Therefore, removing amino acids 2 to 7 of the NS1 protein did not alter the cellular location of the protein. The truncated NS2 protein was also found in the nuclei of SV-NS-Δ2-7-infected cells (data not shown).

Plasmid pSV-NS-Δ34-38 has an internal deletion of 15 nucleotides in the NS1-coding sequence (Fig. 2c). This region, amino acids 34 to 38, was chosen because it is similar to the basic amino acid signal sequence of SV40 T antigen (19, 20). In addition, these amino acids are conserved among all influenza A virus NS1 proteins examined. The NS1

protein expressed in SV-Δ34-38 virus-infected cells migrated faster than wild-type NS1 and slower than the NS1 in SV-NS-Δ2-7-infected cells (Fig. 3, lanes 4, 1, and 3, respectively). The NS2 protein was not detected in several experiments, although the coding sequence of the protein should not be affected in the SV-NS-Δ34-38 construct. This observation was not examined further. When virus-infected cells were analyzed by immunofluoresecence, the NS-Δ34-38 protein was found to localize in the nuclei (Fig. 2c), presenting the same staining pattern as wild-type NS1 (Fig. 2a).

Since the amino acids deleted in NS-Δ34-38 are located in a very polar area of the protein, it was thought that a larger deletion in this domain would be necessary for an effect on the nuclear location of NS1. Construction of pSV-NS-Δ2-51 resulted in a protein lacking amino acids 2 to 51 (Fig. 2d). Lysates of CV-1 cells infected with SV-NS-Δ2-51 virus had a band that corresponds to the expected molecular weight of the truncated NS1 protein (Fig. 3, lane 5). NS2 protein was not detected in these virus-infected cells (lane 5). This was expected, since pSV-NS-Δ2-51 lacks the NS2 donor splice site (at nucleotide 57 of the wild-type NS gene). When infected cells were analyzed by immunofluorescence, the mutant NS1 protein was found to localize in the nucleus, despite the deletion of amino acids 2 to 51 (Fig. 2d).

The next experiment was designed to examine the nuclear localization of NS1 mutant proteins with even larger deletions. In fact, constructs pSV-NS-Δ2-81 and pSV-NS-Δ82-237 allowed for separate expression of the COOH-terminal and NH_2-terminal domains of the protein (Fig. 2 e and f). Virus SV-NS-Δ2-81 expresses an NS1 protein which lacks the first 81 amino acids, whereas SV-NS-Δ82-237 expresses only amino acids 1 to 81 plus an additional 11 amino acids (HGLHTCFAIHN), which are derived as a consequence of a frame shift caused by end filling of the *Nco*I site. Cells infected with these deletion mutants were analyzed by immunofluorescence (Fig. 2 e and f). In both instances, the protein was found to be localized to the nucleus. Since these recombinant viruses independently express the NS1 amino-terminal sequences (1 to 81) and the NS1 carboxyl-terminal sequences (82 to 237), it seems likely that there is at least one signal present in each mutant protein. Therefore, more than one karyophilic signal should exist in the full-length A/Alaska/6/77 NS1 protein.

Expression of NS1–α-globin chimeric proteins. The next step was to define the nuclear signal sequence present in the polypeptide region containing amino acids 1 to 81. Since introduction of deletions in the 81-amino-acid fragment could adversely affect the stability of the proteins, fusion proteins were expressed by using the chimpanzee α-globin gene. The latter protein is normally cytoplasmic but is small enough to allow for diffusion into the nucleus (10). Two constructs were generated. One, (pSV-NSglobin-fpI), contained the first 81 codons of the NS1 gene fused to the entire α-globin-coding region. The second construct (pSV-NSglobin-fpII) codes for the same chimeric protein, except that it lacks codons 34 to 38 in the NS1 region. Sequences are joined at *Nco*I sites (CCATGG) present at codon 81 (Met) of the NS1 gene and at the initiation codon of the α-globin gene. The fusion proteins were transiently expressed in COS-1 cells and localized by immunofluorescence microscopy (Fig. 4). The fusion protein expressed from pSV-NSglobin-fpI was found to localize in the nucleus (Fig. 4b), as was the NS1

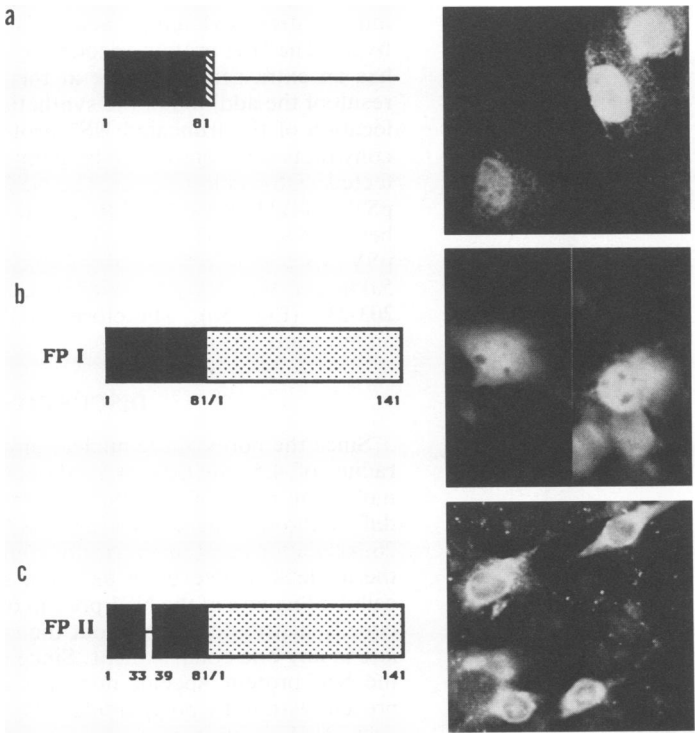

FIG. 4. Expression of NS1–α-globin chimeric proteins. Fusion proteins containing NS1 sequences upstream of α-globin sequences were constructed. (a) The truncated NS1 protein of pSV-NS-Δ82-237 contains amino acids 1 to 81 as well as an additional 11 amino acids, as described in the text. (b) Fusion protein I (FPI) is expressed from pSV-NSglobin-fpI and codes for NS1 amino acids 1 to 81 (solid box) fused with α-globin sequences (stippled box). (c) FPII is expressed from pSV-NSglobin-fpII. This protein is the same as FPI, except that it lacks the NS1 sequences encoding amino acids 34 to 38. The full-length α-globin sequences are present in FPI and FPII. COS-1 cells were transfected with each construct and used for indirect immunofluorescence.

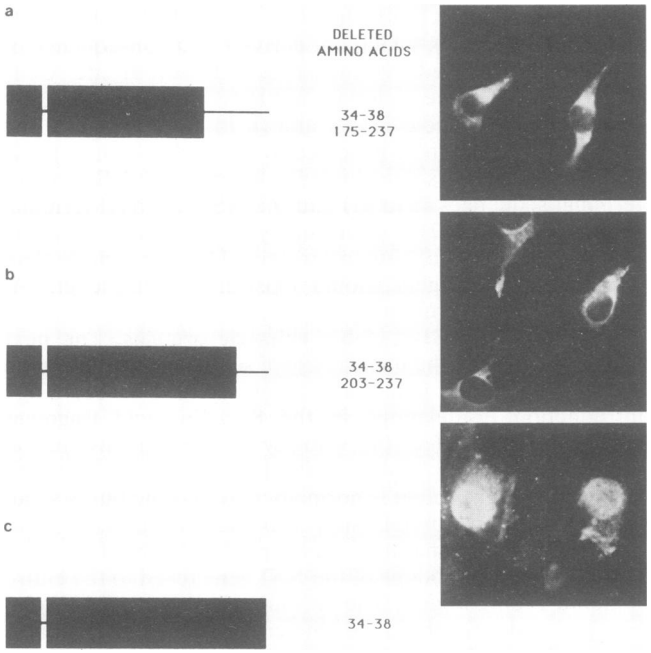

FIG. 5. Expression of NS1 proteins with carboxyl-terminal deletions. Plasmid pSV-NS-Δ34-38, containing a cDNA of an NS1 protein lacking amino acids 34 to 38, was used to generate carboxyl-terminal deletions. The NS1 proteins were expressed in a transient system in COS-1 cells. Cells were fixed 72 h posttransfection and stained with anti-NS1 serum to localize the truncated proteins. The mutant NS1 protein encoded by each construction is diagrammed, and indirect immunofluorescence is shown for cells transfected with the following: a, pSV-NS-Δ34-38/175-237; b, pSV-NS-Δ34-38/203-237; c, pSV-NS-Δ34-38.

protein expressed from SV-NS-Δ82-237 (Fig. 4a). In contrast, when the fusion protein lacking NS1 amino acids 34 to 38 was expressed in COS-1 cells, NS1-specific immunofluorescence was observed mainly in the cytoplasm (Fig. 4c). We thus conclude that removal of the amino acids Asp-Arg-Leu-Arg-Arg in the amino-terminal NS1 fragment alters the location of the fusion protein. This domain will be referred to as nuclear localization signal 1 (NLS1).

A/HT/24269/85	(H3N2)	R
A/Alaska/6/77	(H3N2)	R
A/Udorn/72	(H3N2)	
A/Berkeley/1/68	(H2N2)	
A/AA/6/60	(H2N2)	
A/Denver/1/57	(H1N1)	K
A/HT/23284/85	(H1N1)	K
A/HT/18515/84	(H1N1)	K
A/Maryland/2/80	(H1N1)	
A/USSR/90/77	(H1N1)	
A/FW/1/50	(H1N1)	
A/FM/1/47	(H1N1)	K
A/Bellamy/42	(H1N1)	
A/PR/8/34	(H1N1)	
A/WSN/33	(H1N1)	DAPFL**DRLRR**DQKSLRG
A/Swine/Iowa/15/30	(H1N1)	
A/FPV/Rostock/34	(H7N1)	
A/Turkey/Oregon/71	(H7N3)	D A K
A/Duck/Alberta/60/76	(H12N5)	D A K
B/Lee/40		DYPGQ**DRLHR**LKRKLES

FIG. 6. Comparison of NLS1 sequences of influenza virus NS1 protein. Amino acids 28 to 44 of the NS1 protein of A/WSN/33 are compared with those of other influenza A virus isolates (1, 2, 6, 22, 32). The region defined as NLS1 is in boldface. The homologous region in B/Lee/40 NS1 (4), amino acids 46 to 50, is underlined.

SV40 T antigen	PRO LYS+ LYS+ LYS+ ARG+ LYS+ MET
SV40 VP1	PRO THR LYS+ ARG+ LYS+ GLY SER
SV40 VP2/3	PRO ASN LYS+ LYS+ LYS+ ARG+ LYS+
Polyoma T antigen	PRO LYS+ LYS+ ALA ARG+GLU ASP VAL SER ARG+ LYS+ ARG+PRO ARG+
Yeast L3 protein *	PRO ARG+ LYS+ ARG+

FIG. 7. Amino acid sequence of NLS2. The sequence of the NS1 protein of influenza virus A/Alaska/6/77 is shown. Amino acids 203 to 237 are indicated in the single-letter code. The sequence 216 to 221 is enlarged and compared with previously published signal sequences of other nuclear proteins. Basic amino acids are marked with plus signs, and proline residues are underlined (19, 20, 26, 32, 40, 46, 47). *, Only a portion of the reported *S. cerevisiae* L3 protein signal is shown (32).

Determination of NLS2. After the identification of NLS1, attempts were made to identify a second nuclear localization signal in the carboxyl-terminal region. BAL 31 nuclease digestion was used to remove carboxyl-terminal coding sequences of the NS1 gene from plasmid pSV-NS-Δ34-38. NS1 proteins expressed from these deletion mutants lack NLS1, as well as a defined number of amino acids from the carboxyl terminus. pSV-NS-Δ34-38/175-237, pSV-NS-Δ34-38/203-237, and pSV-NS-Δ34-38 encode proteins which terminate after NS1 amino acids 174, 202, and 237, respectively. The NS1 protein encoded by pSV-NS-Δ34-38/203-237 has an extra serine residue at the carboxyl terminus as a result of the addition of the synthetic stop codon linker. The location of the truncated NS1 proteins expressed by these constructs was determined by immunofluorescence of transfected COS-1 cells (Fig. 5). The NS1 protein expressed from pSV-NS-Δ34-38 localized to the nucleus (Fig. 5c), as shown before (Fig. 2f). However, the NS1 protein expressed from pSV-NS-Δ34-38/175-237 was found to be cytoplasmic (Fig. 5a) as was the NS1 protein expressed from pSV-NS-Δ34-38/203-237 (Fig. 5b). Therefore, NLS2 is defined between amino acids 203 and 237.

DISCUSSION

Since the pores in the nuclear envelope have a functional radius of 4.5 nm (35), it is thought that proteins with a molecular mass of less than 67,000 daltons can passively diffuse through these nuclear pores (35). The NS1 protein is 26,000 daltons and theoretically could passively diffuse into the nucleus. However, if diffusion alone accounted for the cellular location of the NS1 protein one would expect it to be equally distributed throughout the cell rather than accumulate in any one compartment. Since this is not the case with the NS1 protein, specific nuclear signal sequences must be present within the polypeptide (11). Expression of influenza virus NS1 deletion mutants and NS1-α-globin fusion proteins have been used to identify two amino acid regions (NLS1 and NLS2) that affect the cellular location of the protein. Truncated NS1 proteins with small deletions (NS-Δ2-7 and NS-Δ34-38) or large deletions (NS-Δ2-51, NS-Δ2-81, and NS-Δ82-237) were expressed and found to localize to

the nucleus. Since mutant NS proteins NS-Δ82-237 and NS-Δ1-81 contain no overlapping sequence, at least two distinct nuclear signal sequences are present within the NS1 protein. Two nuclear signal sequences have been reported for a number of viral and cellular proteins, including the polyomavirus T antigen (40), the rat glucocorticoid receptor (39), and the *S. cerevisiae matα2* protein (16, 17).

The expression and localization of fusion proteins enabled us to define NLS1. This signal contains the amino acids Asp-Arg-Leu-Arg-Arg. We cannot eliminate the possibility that this short amino acid region comprises a portion of a larger signal disrupted by removal of these amino acids. This domain of the NS1 protein is highly charged and conserved in all influenza A virus isolates studied, including avian and swine species (Fig. 6) (1, 2, 6, 22, 33). It is also highly conserved within the influenza B virus NS1 protein (Fig. 6) (4, 48). The homology of NLS1 among the NS1 proteins of influenza A and B viruses is noteworthy, since these proteins differ by over 80% of their amino acid sequence and are antigenically distinct (49). It is tempting to hypothesize that nuclear localization of the protein during the virus life cycle is critical and that there are tight constraints upon the amino acid sequence of this region.

The expression of NS1 proteins with carboxyl-terminal deletions has allowed for definition of NLS2 to within the region encompassing amino acids 203 to 237 (Fig. 7). This domain contains a short stretch of basic amino acids at residues 216 to 221. This stretch of amino acids, Pro-Lys-Gln-Lys-Arg-Lys, is similar to other known nuclear signal sequences (Fig. 7). NLS2 is present in most influenza A virus isolates. Of 25 human, avian, and swine isolates examined, 18 were found to contain the basic amino acid stretch (216 to 221) of NLS2. Variability in the length of the NS1 proteins (6, 38) results in the absence of NLS2 in certain naturally occurring isolates. For example, the NS1 of the avian virus A/Turkey/Oregon/71 has only 124 amino acids (34) and the human isolate A/FM/1/47 has only 202 amino acids (22). Although NLS2 is not an absolute requirement for protein function of the NS1 protein, the existence of multiple nuclear signal sequences in one protein may effect the kinetics of entry into the nucleus (40). A comparative study of the kinetics of migration of the NS1 protein from different viruses has not been done. In addition, it has been postulated that a protein may enter the nucleus by either specific and selective entry or by diffusion followed by specific binding to nondiffusible nuclear components (11). It has not been determined which is the functional mechanism for the NLS1 and NLS2 domains of the NS1 protein.

A common theme of known nuclear signal sequences is the presence of a proline residue before or after a stretch of basic amino acids. It is believed that this proline residue is important in the three-dimensional structure of the signal (42). In the NS1 protein, a conserved proline is found 3 amino acids upstream from the basic NLS1 sequence. Also, within NLS2 there is a region of basic amino acids preceded by a proline. Thus, NLS1 and NLS2 are very similar to the known nuclear localization signals of SV40 T antigen (19, 20, 26), polyomavirus T antigen (40), SV40 capsid polypeptides VP1, VP2, and VP3 (13, 46, 47), and *S. cerevisiae* ribosomal protein L3 (32) (Fig. 7). Conversely, in a few proteins, nuclear signal sequences have been defined which are not made up of basic amino acids (10, 16). For example, an accumulation signal has been defined for the influenza virus nucleoprotein, NP, which contains only one basic amino acid and no prolines (10). Therefore, the NP signal is very different from NLS1 and NLS2 of the NS1 protein. It is interesting that such different signals could exist in nuclear proteins of the same virus.

ACKNOWLEDGMENTS

We thank S. Liebhaber for kindly providing plasmid pMCα-19.

This work was supported in part by Public Health Service grants AI-11823 and AI-18998 (P.P.)and grants from the A. and A. Sinsheimer Foundation, the Charles H. Revson Foundation, and the National Foundation for Infectious Disease (M.K.).

LITERATURE CITED

1. **Baez, M., R. Taussig, J. J. Zazra, J. F. Young, P. Palese, A. Reisfeld, and A. Skalka.** 1980. Complete nucleotide sequence of the influenza A/PR/8/34 virus NS gene and comparison with the NS genes of A/Udorn/72 and A/FPV/Rostock/34 strains. Nucleic Acids Res. **8**:5845–5858.
2. **Baez, M., J. J. Zazra, R. M. Elliott, J. F. Young, and P. Palese.** 1981. Nucleotide sequence of the influenza A/duck/Alberta/60/76 virus NS RNA: conservation of the NS1/NS2 overlapping gene structure in a divergent influenza virus RNA segment. Virology **113**:397–402.
3. **Briedis, D. J., G. Conti, E. A. Munn, and B. W. J. Mahy.** 1981. Migration of influenza virus-specific polypeptides from cytoplasm to nucleus of infected cells. Virology **111**:154–164.
4. **Briedis, D. J., and R. A. Lamb.** 1982. Influenza B virus genome: sequences and structural organization of RNA segment 8 and the mRNAs coding for the NS_1 and NS_2 proteins. J. Virol. **42**: 186–193.
5. **Buonagurio, D. A., M. Krystal, P. Palese, D. Deborde, and H. F. Maassab.** 1984. Analysis of an influenza A virus mutant with a deletion in the NS segment. J. Virol. **49**:418–425.
6. **Buonagurio, D. A., S. Nakada, J. D. Parvin, M. Krystal, P. Palese, and W. M. Fitch.** 1986. Evolution of human influenza A viruses over 50 years: rapid, uniform rate of change in NS gene. Science **232**:980–982.
7. **Cashmore, A., L. Szabo, M. Timko, A. Kausch, G. Van den Broeck, P. Schreier, H. Bohnert, L. Herrera-Estrella, M. Van Montagu, and J. Schell.** 1985. Import of polypeptides into chloroplasts. Bio/Technology **3**:803-808.
8. **Chen, C., and H. Okayama.** 1987. High-efficiency transformation of mammalian cells by plasmid DNA. Mol. Cell. Biol. **7**: 2745–2752.
9. **Chen, E. Y., and P. H. Seeburg.** 1985. Supercoil sequencing: a fast and simple method for sequencing plasmid DNA. DNA **4**: 165–170.
10. **Davey, J., N. J. Dimmock, and A. Colman.** 1985. Identification of the sequence responsible for the nuclear accumulation of the influenza virus nucleoprotein in xenopus oocytes. Cell **40**:667–675.
11. **DeRobertis, E. M.** 1983. Nucleocytoplasmic segregation of proteins and RNAs. Cell **32**:1021–1025.
12. **Dingwall, C., S. M. Dilworth, S. J. Black, S. E. Kearsey, L. S. Cox, and R. A. Laskey.** 1987. Nucleoplasmin cDNA sequence reveals polyglutamic acid tracts and a cluster of sequences homologous to putative nuclear localization signals. EMBO **6**: 69–74.
13. **Gharakhanian, E., J. Takahashi, and H. Kasamatsu.** 1987. The carboxyl 35 amino acids of SV40 Vp3 are essential for its nuclear accumulation. Virology **157**:440–448.
14. **Greenspan, D., M. Krystal, S. Nakada, H. Arnheiter, D. S. Lyles, and P. Palese.** 1985. Expression of influenza virus NS2 nonstructural protein in bacteria and localization of NS2 in infected eucaryotic cells. J. Virol. **54**:833–843.
15. **Gruss, P., N. Rosenthal, M. Koenig, R. W. Ellis, T. Y. Shih, E. M. Scolnick, and G. Khoury.** 1982. The expression of viral and cellular p21 ras genes using SV40 as a vector, p 13. *In* Y. Gluzman (ed.), Eukaryotic viral vectors. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Hall, M. N., L. Hereford, and I. Herskowitz.** 1984. Targeting of *E. coli* β-galactosidase to the nucleus in yeast. Cell **36**:1057–1065.
17. **Hall, M. N., and A. D. Johnson.** 1987. Homeo domain of the

yeast repressor α2 is a sequence-specific DNA-binding domain but is not sufficient for repression. Science **237**:1007–1012.
18. **Hay, R., P. Bohni, and S. Gasser.** 1984. How mitochondria import proteins. Biochim. Biophys. Acta **779**:65–87.
19. **Kalderon, D., W. D. Richardson, A. F. Markham, and A. E. Smith.** 1984. Sequence requirements for nuclear location of simian virus large T antigen. Nature (London) **311**:33–38.
20. **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.
21. **Krug, R. M., and P. R. Etkind.** 1973. Cytoplasmic and nuclear virus specific proteins in influenza virus-infected MDCK cells. Virology **56**:334–348.
22. **Krystal, M., D. A. Buonagurio, J. F. Young, and P. Palese.** 1983. Sequential mutations in the NS genes of influenza virus field strains. J. Virol. **45**:547–554.
23. **Krystal, M., R. Li, D. Lyles, G. Pavlakis, and P. Palese.** 1986. Expression of the three influenza virus polymerase proteins in a single cell allows growth complementation of viral mutants. Proc. Natl. Acad. Sci. USA **83**:2709–2713.
24. **Lamb, R. A.** 1983. The influenza virus RNA segments and their encoded proteins, p. 21–69. *In* P. Palese and D. W. Kingsbury (ed.), Genetics of influenza viruses. Springer Verlag, Vienna.
25. **Lamb, R. A., and C.-J. Lai.** 1984. Expression of unspliced NS1 mRNA, spliced NS2 and mRNA, and a spliced chimera mRNA from cloned influenza virus NS DNA in an SV40 vector. Virology **135**:139–147.
26. **Landford, 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.
27. **Lazarowitz, S. G., R. W. Compans and P. W. Choppin.** 1971. Influenza virus structural and nonstructural proteins in infected cells and their plasma membranes. Virology **46**:830–843.
28. **Liebhaber, S. A., and K. A. Begley.** 1983. Structural and evolutionary analysis of two chimpanzee β-globin mRNAs. Nucleic Acids Res. **11**:8915–8929.
29. **Lyons, R. H., B. Q. Ferguson, and M. Rosenberg.** 1987. Pentapeptide nuclear localization signal in adenovirus E1A. Mol. Cell. Biol. **7**:2451–2456.
30. **Mahy, B. W. J., T. Barrett, J. M. Briedis, J. M. Brownson, and A. J. Wolstenholme.** 1980. Influence of the host cell on influenza virus replication. Philos. Trans. R. Soc. London Biol. Sci. **288**: 349–357.
31. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
32. **Moreland, R. B., H. G. Nam, L. M. Hereford, and H. M. Fried.** 1985. Identification of a nuclear localization signal of a yeast ribosomal protein. Proc. Natl. Acad. Sci. USA **82**:6561–6565.
33. **Nakajima, K., E. Nobusawa, and S. Nakajima.** 1984. Genetic relatedness between A/swine/Iowa/15/30(H1N1) and human influenza viruses. Virology **139**:194–198.
34. **Norton, G. P., T. Tanaka, K. Tobita, S. Nakada, D. A. Buonagurio, D. Greenspan, M. Krystal, and P. Palese.** 1987. Infectious influenza A and B virus variants with long carboxyl terminal deletions in the NS1 polypeptides. Virology **156**:204–213.
35. **Paine, P. L., L. C. Moore, and S. B. Horowitz.** 1975. Nuclear envelope permeability. Nature (London) **254**:109–114.
36. **Parker, B. A., and G. R. Stark.** 1979. Regulation of simian virus 40 transcription: sensitive analysis of the RNA species present early in infections by virus or viral DNA. J. Virol. **31**:360–369.
37. **Parvin, J. D., F. I. Smith, and P. Palese.** 1986. Rapid RNA sequencing using double-stranded template DNA, SP6 polymerase, and 3′-deoxynucleotide triphosphates. DNA **5**:167–171.
38. **Parvin, J. D., J. F. Young, and P. Palese.** 1983. Nonsense mutations affecting the lengths of the NS1 nonstructural proteins of influenza A virus isolates. Virology **128**:512–517.
39. **Picard, D., and K. Yamamoto.** 1987. Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. EMBO J. **6**:3333–3340.
40. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44**:77–85.
41. **Schatz, G., and R. A. Butow.** 1983. How are proteins imported into mitochondria? Cell **37**:316–318.
42. **Smith, A. E., D. Kalderon, B. L. Roberts, W. H. Colledge, M. Edge, P. Gillett, A. Markham, E. Paucha, and W. D. Richardson.** 1985. The nuclear location signal. Proc. R. Soc. London B. Biol. Sci. **226**:43–58.
43. **Smith, G. L., J. Z. Levin, P. Palese, and B. Moss.** 1987. Synthesis and cellular location of the ten influenza virus polypeptides individually expressed by recombinant vaccinia viruses. Virology **160**:336–345.
44. **Watson, M. E. E.** 1984. Compilation of published signal sequences. Nucleic Acids Res. **12**:5145–5164.
45. **Winberg, G., and M.-L. Hammarskjold.** 1980. Isolation of DNA from agarose gels using DEAE-paper. Application to restriction site mapping of adenovirus type 16 DNA. Nucleic Acids Res. **8**: 253–264.
46. **Wychowski, C., D. Benichou, and M. Girard.** 1986. A domain of SV40 capsid polypeptide VP1 that specifies migration into the cell nucleus. EMBO J. **5**:2569–2576.
47. **Wychowski, C., D. Benichou, and M. Girard.** 1987. The intranuclear location of simian virus 40 polypeptides VP2 and VP3 depends on a specific amino acid sequence. J. Virol. **61**:3862–3869.
48. **Yamashita, M., M. Krystal, W. M. Fitch, and P. Palese.** 1988. Influenza B virus evolution: cocirculating lineages and comparison of evolutionary pattern with those of influenza A and C viruses. Virology **163**:112–122.
49. **Young, J. F., U. Desselberger, P. Palese, B. Ferguson, A. R. Shatzman, and M. Rosenberg.** 1983. Efficient expression of influenza virus NS1 nonstructural proteins in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **80**:6105–6109.