MOLECULAR AND CELLULAR BIOLOGY, Aug. 1990, p. 4139–4145
0270-7306/90/084139-07$02.00/0

Vol. 10, No. 8

Function of Two Discrete Regions Is Required for Nuclear Localization of Polymerase Basic Protein 1 of A/WSN/33 Influenza Virus (H1 N1)

SUREND T. NATH† AND DEBI P. NAYAK*

Department of Microbiology and Immunology, Jonsson Comprehensive Cancer Center, University of California—Los Angeles School of Medicine, Los Angeles, California 90024-1747

Received 28 November 1989/Accepted 7 May 1990

Polymerase basic protein 1 (PB1) of influenza virus (A/WSN/33), when expressed from cloned cDNA in the absence of other viral proteins, accumulates in the nucleus. We have examined the location and nature of the nuclear localization signal of PB1 by using deletion mutants and chimeric constructions with chicken muscle pyruvate kinase, a cytoplasmic protein. Our studies showed some novel features of the nuclear localization signal of PB1. The signal was present internally within residues 180 to 252 of PB1. Moreover, unlike most nuclear localization signals, it was not a single stretch of contiguous amino acids. Instead, it possessed two discontinuous regions separated by an intervening sequence which could be deleted without affecting its nuclear localization property. On the other hand, deletion of either of the two signal regions rendered the protein cytoplasmic, indicating that the function of both regions is required for nuclear localization and that one region alone is not sufficient. Both of these signal regions contained short stretches of basic residues. Possible ways by which this novel bipartite signal can function in nuclear localization are discussed.

In eucaryotic cells, the nucleus is separated from the cytoplasm by a nuclear membrane, which is punctuated by the presence of nuclear pores. These nuclear pores have an average diameter of 9 nm (range, 7 to 11 nm [8, 34]) and function as the gateways for the nuclear-cytoplasmic exchange of macromolecules. Therefore, while smaller proteins (M_r >60,000) may passively diffuse through the nuclear pores, larger proteins or protein complexes are actively transported into the nucleus (6. 7, 10, 43). The proteins that are selectively transported into the nucleus have been shown to possess the nuclear localization signal(s). The function of these signals in nuclear localization has been clearly demonstrated for a number of nuclear proteins (10, 43). For example, when these signals are deleted, the proteins become cytoplasmic, whereas when these signals are attached to a cytoplasmic protein, the marker protein is transported into the nucleus (5, 43). For the majority of the nuclear proteins studied to date, the nuclear signal appears to be a short stretch of seven to nine contiguous basic amino acids, which directs the protein into the nucleus.

Among the well-characterized nuclear signals, the one derived from the simian virus 40 (SV40) T antigen has been studied in great detail. It consists of a proline residue followed by a stretch of basic residues (PKKKRKV [19, 24]). This short sequence of seven amino acids, when attached to cytoplasmic proteins, can translocate the protein efficiently into the nucleus (19). However, this sequence motif shows little homology to other sequences implicated in nuclear localization except that many of them also possess basic residues. Moreover, similar stretches of basic residues have also been found in many nonnuclear cytoplasmic proteins (43). Furthermore, in a number of proteins, e.g., influenza virus NP protein (6) and *MAT*α2 protein of *Saccharomyces cerevisiae* (17, 18), nonbasic residues have also been implicated in nuclear localization. In addition, some proteins, such as polyoma virus T antigen (39, 40), human T-lymphotropic virus (HTLV) p27^x (41), and influenza virus NS1 protein (16), possess multiple nuclear signals which can function independently.

Among the RNA viruses, influenza virus, a segmented negative-strand RNA virus, is unique in that both the replication and the transcription of the viral RNA occur within the nucleus of the infected cell (12). Furthermore, the influenza virus transcriptase consists of three polymerase proteins (PB1, PB2, and PA), which function as a heterotrimer complex. The heterotrimer complex plays a multifunctional role in the transcription and replication of the viral RNA and must be transported into the nucleus of the infected cell (22). Earlier studies have shown that each of the three polymerase proteins as well as NS1, NS2, and NP of influenza virus, when expressed individually from the cloned cDNA in the absence of other viral proteins, accumulates in the cell nucleus (1, 6, 16, 22, 23, 43), and therefore, each of these proteins probably has the nuclear localization signal(s).

Although the nuclear localization signals of influenza virus NP (6) and NS1 (12) have been identified, little is known about the nature of the signals of the three polymerase proteins and the mechanism of their nuclear entry. We have therefore undertaken systematic studies to define the function of different domains within the individual polymerase proteins, including the mechanism of their nuclear localization.

In the present study, our data show that the influenza virus polymerase protein PB1 possesses two discrete domains, both of which are required for nuclear localization; one domain alone is not sufficient.

MATERIALS AND METHODS

Virus, cDNA clones, and antiserum. Strain A/WSN/33 influenza virus was grown as described previously (33). The isolation and analysis of the cDNA clone of WSN PB1 have been reported (42). The cDNA of the PB1 gene of A/WSN/33 was cloned into the *Hin*dIII site of pSP64. Recombinant

* Corresponding author.
† Present address: Diagnostic Products Corporation, 5700 W. 96th St., Los Angeles, CA 90045.

vaccinia virus VTF7.3 and plasmid pTF7.5 were obtained from B. Moss (14). Monospecific antiserum against PB1 was obtained as described before (1). The cDNA of chicken muscle type M1 pyruvate kinase (PK) and the rabbit antiserum against PK were provided by A. Smith (19). All plasmids were grown in *Escherichia coli* SCS-1 cells. All DNA manipulations, including restriction digestion, linker addition, ligation of DNA fragments, and transformation of *E. coli* with plasmid DNA, were carried out by standard procedures (28). All expression experiments were carried out in CV-1 cells (1).

Construction of PB1 deletion mutants. For the construction of PB1Δ667, the plasmid PB1/SP64 was digested with *Bam*HI. The large fragment (4,973 base pairs [bp]) was isolated, end-filled by using DNA polymerase, and then digested with *Hin*dIII to obtain a 2,026-bp PB1 fragment possessing a *Hin*dIII site at its 5′ end and a blunt end at its 3′ end. The plasmid pGem-3 was digested with *Pst*I, blunt-ended with mung bean nuclease, and then digested with *Hin*dIII. The vector was dephosphorylated and ligated to the PB1 DNA fragment. This recombinant plasmid encodes amino acids (aa) 1 to 667 of PB1.

For the construction of PB1Δ711, the plasmid PB1/SP64 was digested with *Hin*dIII to release the 2,341-bp PB1 DNA. This DNA fragment was then digested with *Eco*RV to generate a 2,154-bp PB1 DNA having a *Hin*dIII site at its 5′ end and a blunt end at its 3′ end. Plasmid pGem-3 was digested with *Eco*RI, and the ends were filled by DNA polymerase I and then digested with *Hin*dIII. This vector was then ligated to the PB1 fragment to generate a deletion mutant having aa 1 to 711 of PB1. Three (GRL) and five (IPVSL) amino acids at the COOH end of PB1Δ667 and PB1Δ711, respectively, came from the restriction linkers. The *Hin*dIII-*Ava*I fragment of PB1Δ667 and *Hin*dIII-*Hha*I fragment of PB1Δ711 containing the PB1 region were blunt-ended and then ligated to the blunt-ended *Bam*HI site of the plasmid pTF7.5 (14).

Construction of PB1-PK chimera. PB1-PK chimeras were constructed by using the convenient restriction sites within the PB1 cDNA (42). Three single restriction sites for *Acc*I (nucleotide [nt] 778), *Eco*RI (nt 1495), and *Bgl*II (nt 2258) of PB1 were used to construct PB1(1-252)-PK, PB1(254-490)-PK, and PB1(492-746)-PK, respectively. *Eco*RI linkers were used to construct chimeras so that the PB1 sequences were always at the amino terminus joined to the PK sequences (aa 17 to 529 of PK) in phase at the COOH terminus. Subsequently, the PB1(1-252)-PK chimera was used for making additional deletion mutants from appropriate restriction sites. For example, *Bal*I (nt 452), *Nco*I (nt 117), *Hph*I (nt 621), and *Sau*3AI (nt 564) were used to construct PB1(1-143)-PK, PB1(1-33)-PK, PB1(1-203)-PK, and PB1(182-252)-PK, respectively. Chimeras which were constructed from pGem-4 with the gene under the T7 promoter were used directly for transfection. The pGem-3 chimeras were digested by *Hin*dIII and *Bam*HI to release the PB1-PK chimeric fragment, which was blunt-ended and then ligated to the blunt-ended *Bam*HI site of pTF7.5 (14). In chimeras which required an ATG initiation codon, an *Nco*I linker sequence was ligated in phase to the 5′ end of the PB1 fragment. Figure 1 lists the PB1 constructs and chimeras used in this study. Two to three amino acid residues were often acquired from the restriction linkers at the junction of PB1 and PK. For site-specific deletions (Table 1), the PB1 cDNA fragment (bp 1 to 778) was first inserted into the *Hin*dIII-*Eco*RI site of M13mp19. Oligonucleotides ranging from 30 to 36 bases long were used to create specific deletions of 7 to 20 amino acids with an Amersham in vitro mutagenesis system. All deletions were checked by sequence analysis.

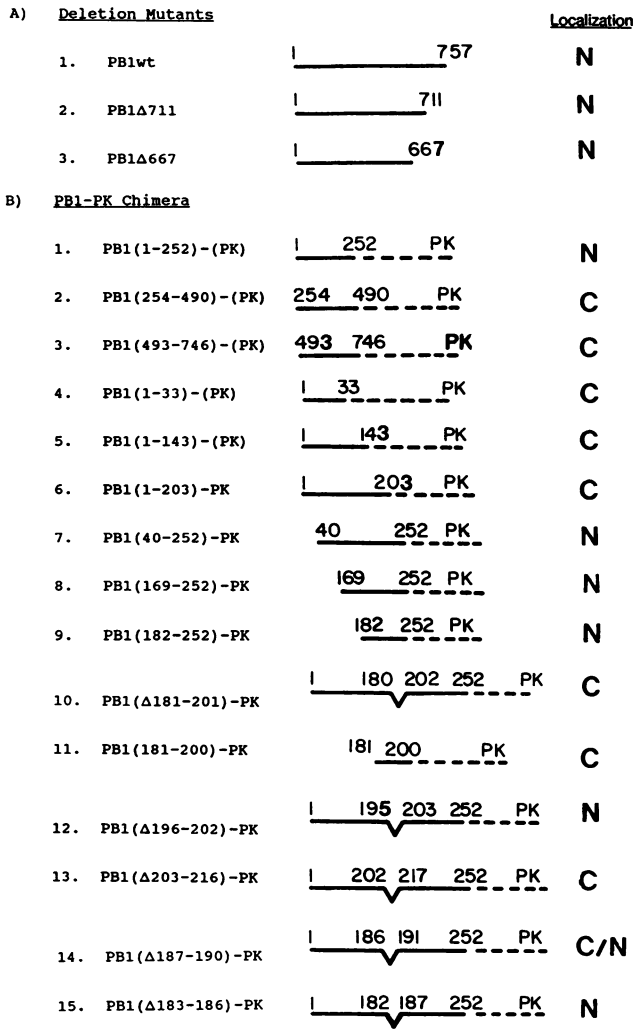

FIG. 1. Schematic presentation of PB1 deletion mutants and PB1-PK chimeras. Solid line, PB1; broken line, PK; N, nucleus; C, cytoplasm; wt, wild type.

Expression of PB1-PK chimeras and PB1 deletion mutants. For transient expression in CV-1 cells, the procedure described by Fuerst et al. (14) was used. Briefly, approximately 4×10^5 CV-1 cells were seeded on cover slips placed in a 30-mm-diameter tissue culture dish. On the following day, the cells were infected with recombinant vaccinia virus VTF7.3 (2 PFU per cell). One hour after infection, 5 μg of the recombinant expression plasmid containing the desired PB1 construct was introduced by lipofection-mediated transfection (12). At 16 h posttransfection, cells were examined for expression of the chimeric proteins by indirect immunofluorescence.

Immunofluorescence. Cells were fixed with methanol-acetone (1:1) at −20°C and used for intracellular immunofluorescence (1). Usually, PB1 antibody (1:20) for the deletion mutants and PK antibody (1:20) for the PK chimeras was added to the cover slips containing fixed cells and incubated in a humidified chamber for 30 min at 37°C. The antibody-treated cells were then washed with phosphate-buffered saline and layered with a 1:50 dilution of fluorescein isothiocyanate-conjugated goat anti-rabbit immunoglobulin G (Cap-

TABLE 1. Analysis of the region involved in nuclear localization of influenza virus PB1

Construct	PB1 sequence[a]	Localization[b]
	1 180 190 200 210 217 252	
PB1 (1–252) PK	MXXX EITTHFQ**RKRR**VRDNM [TKKMVTQ] RTIG**KRKQR**LNKRSY XXXV VNS–PK	N
PB1 (182–252) PK	MG TTHFQ**RKRR**VRDNM TKKMVTQ RTIG**KRKQR**LNKRSY XXXV VNS–PK	N
PB1 (Δ181–201) PK	MXXX E Q RTIG**KRKQR**LNKRSY XXXV VNS–PK	C
PB1 (Δ187–190) PK	MXXX EITTHFQ VRDNM TKKMVTQ RTIG**KRKQR**LNKRSY XXXV VNS–PK	C/N
PB1 (Δ183–186) PK	MXXX EIT **RKRR**VRDNM TKKMVTQ RTIG**KRKQR**LNKRSY XXXV VNS–PK	N
PB1 (181–200) PK	M ITTHFQ**RKRR**VRDNM TKKMV NS–PK	C
PB1 (Δ196–202) PK	MXXX EITTHFQ**RKRR**VRDNM RTIG**KRKQR**LNKRSY XXXV VNS–PK	N
PB1 (Δ203–216) PK	MXXX EITTHFQ**RKRR**VRDNM TKKMVTQ Y XXXV VNS–PK	C

[a] Amino acids added during cDNA construction are underlined. Basic residues involved in nuclear localization are in boldface letters. An intervening sequence that can be deleted without affecting nuclear localization is boxed. X, Any amino acid residue.
[b] N, Nucleus; C, cytoplasm.

pel Laboratories) for 30 min at 37°C and examined with a Nikon photomicroscope by epifluorescence.

RESULTS

Nuclear localization of signal PB1 is present in the NH_2-terminal third of the PB1 sequence. Earlier studies have shown that the influenza virus PB1, when expressed from a cloned cDNA in CV-1 cells, is transported into the nucleus (1, 44). Furthermore, it was predicted from computer analysis that a short region, Pro-668-Lys-Arg-Asn-Arg-Ser, near COOH terminus of PB1, which possesses partial homology to the nuclear localization signal of the SV40 large T antigen, might function as a possible nuclear localization signal of PB1 (1). To test the function of this region in the nuclear transport of PB1, two deletion mutants were made and expressed in CV-1 cells. PB1Δ711 lacks sequences encoding 47 aa at the COOH end but retains the putative nuclear localization signal, whereas PB1Δ667 lacks sequences for the last 89 aa (aa 668 to 757), including the putative signal sequence. When PB1Δ667 and PB1Δ711 were expressed in CV-1 cells, both proteins, PB1Δ667 (Fig. 2C) and PB1Δ711 (Fig. 2B), migrated into the nucleus, indicating that neither the Pro-668-Lys-Arg-Asn-Arg sequence nor any other sequence in the last 89 aa of the COOH terminus of PB1 was required for transport of PB1 into the nucleus. We therefore made three chimeric constructs in which either the NH_2-terminal third, the middle third, or the COOH-terminal third of the PB1 cDNA was joined in phase to the PK cDNA encoding aa 17 to 519 of PK (Fig. 1). The chimeric proteins encoded by these constructs possessed over 750 aa and therefore are unlikely to be transported by passive diffusion into the nucleus. When expressed in CV-1 cells, the PB1 (254-490)-PK and PB1(493-746)-PK proteins were present only in the cytoplasm (Fig. 2E and F), whereas the chimeric protein PB1(1-252)-PK was found in the nucleus (Fig. 2D). These results indicate that the nuclear localization signal of PB1 is present within the NH_2-terminal 252 residues of PB1 and is absent in aa 254 to 757. In addition, these results show that the PB1 signal can transport a heterologous cytoplasmic protein into the nucleus.

Identification of the nuclear localization signal in PB1. To define the nuclear localization signal of PB1 more precisely, a number of deletion mutations in the NH_2 terminus of PB1 were made and expressed in CV-1 cells as a chimeric protein with the chicken PK. The results showed that PB1(1-33)-PK and PB1(1-143)-PK were cytoplasmic proteins (Fig. 3A and B), whereas PB1(40-252)-PK, PB1(169-252)-PK, and PB1 (182-252)-PK were transported into the nucleus (Fig. 3D, E, and F). These results indicate that the nuclear localization signal was present within residues 182 to 252 of PB1.

To further define the location of the signal, internal deletions were made within these regions. In one such deletion construct, 21 aa (aa 181 to 201) were removed while aa 1 to 180 and 202 to 252 of PB1 were retained. This protein, PB1(Δ181-201)-PK, when expressed in CV-1 cells, was cytoplasmic (Fig. 4A), suggesting that residues 181 to 201 might possess the nuclear localization signal. Two constructs were then made: in one [PB1(1-203)-PK], the first 203 aa were joined in phase to PK; in another [PB1(181-200)-PK], only 20 aa (aa 181 to 200) plus a methionine at the NH_2 terminus were fused to PK. However, both of these proteins, when expressed in CV-1 cells, were cytoplasmic (Fig. 3C and 4B). These results showed that although the removal of 21 aa (aa 181 to 201) caused disruption of the nuclear signal (Fig. 4A), these residues alone [as in PB1(181-200)-PK] or along with the rest of the PB1 NH_2 terminus [as in PB1(1-203)-PK] were not sufficient for transporting the cytoplasmic PK into the nucleus.

These results suggested two possibilities: (i) the residues around aa 202 were an integral part of a single nuclear signal and the deletion of the amino acids in this region [as in PB1(Δ181-201)-PK, PB1(1-203)-PK, or PB1(181-200)-PK] caused disruption of the nuclear signal and consequently of the nuclear localization of the chimeric proteins; (ii) alternatively, aa 202 and the adjacent region did not constitute a nuclear signal, but two discrete signals were present in the flanking regions of residue 202. These two signals need to function cooperatively, and neither signal alone is sufficient for nuclear localization. Consequently, the removal of either of these signals would disrupt nuclear localization.

To distinguish between these two possibilities, a construct was made in which seven amino acids (residues 196 to 202) were deleted. This protein, PB1(Δ196-202)-PK, was present in the nucleus (Fig. 4C), demonstrating that residues 196 to 202 were not a part of the nuclear localization signal and that a single continuous sequence was not responsible for the nuclear localization of PB1. These results also indicated that two of more discrete regions were functioning cooperatively for nuclear localization.

To further define the location of these signals, smaller deletions were made in which fewer amino acids were removed from each of these two regions of PB1. Two

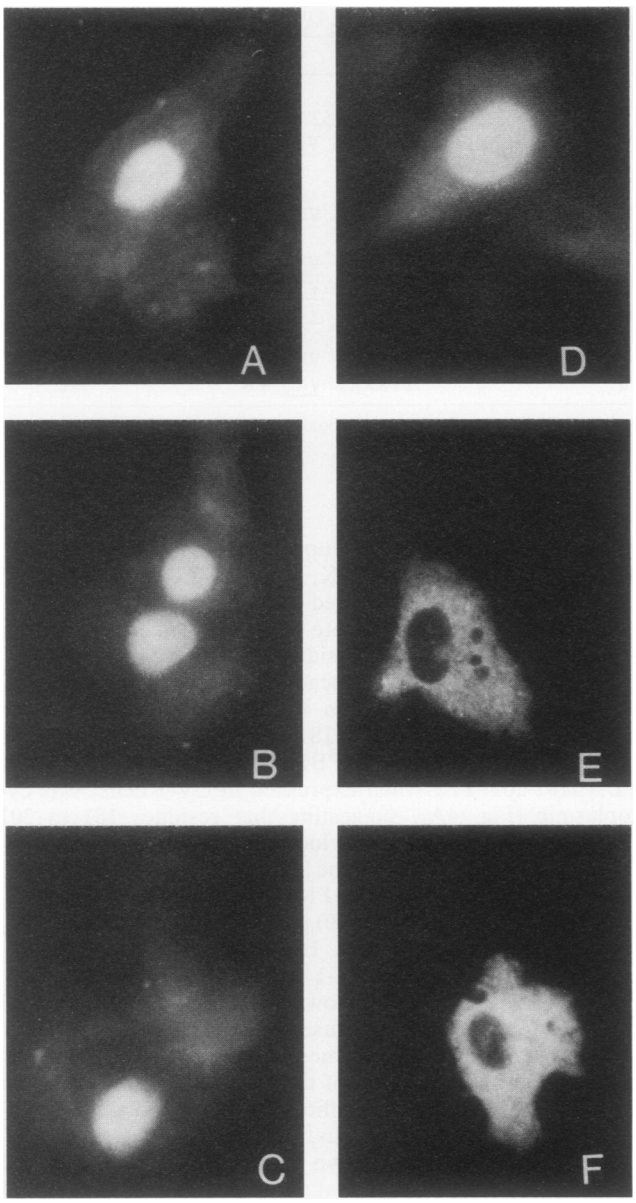

FIG. 2. Localization of PB1 wild type, deletion mutants, and chimeric PB1-PK proteins in CV-1 cells. CV-1 cells were infected with recombinant vaccinia virus VTF7.3 and transfected with a plasmid construct as described in Materials and Methods. At 16 h posttransfection, cells were stained by indirect immunofluorescence with specific antiserum against either PB1 (A, B, C) or PK (D, E, F). Approximately 10% of the cells exhibited specific immunofluorescence. Although the degree of expression as demonstrated by the intensity of fluorescence did vary, all cells transfected with a given construct exhibited a similar phenotype. Kodak Ektachrome 400 film was used with an exposure time of 5 to 15 s. (A) PB1 wild type; (B) PB1Δ711; (C) PB1Δ667; (D) PB1(1-252)-PK; (E) PB1(254-490)-PK; (F) PB1(493-746)-PK.

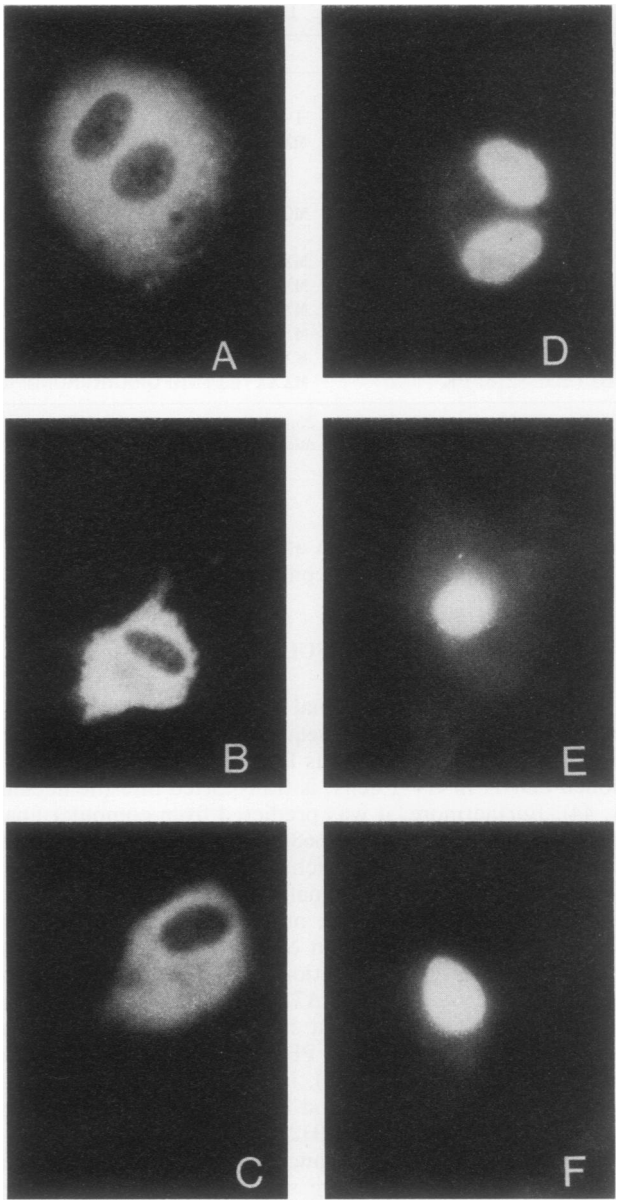

FIG. 3. Localization of chimeric PB1-PK proteins in CV-1 cells. The conditions for transfection and immunofluorescence were same as described in the legend to Fig. 2. (A) PB1(1-33)-PK; (B) PB1(1-143)-PK; (C) PB1(1-203)-PK; (D) PB1(40-252)-PK; (E) PB1(169-252)-PK; (F) PB1(182-252)-PK.

constructs were made on the left-hand flanking region. In PB1(Δ187-190)-PK, four basic amino acids (RKRR) were deleted, whereas PB1(Δ183-186)-PK lacked four adjacent residues (THFQ). One construct [PB1(Δ203-216)] was made on the right-hand flanking region in which 14 residues (RTIGKRKQRLNKRS), including a set of basic residues (KRKQR), were deleted. Upon expression, PB1(Δ187-190)-PK was found in both the cytoplasm and the nucleus (Fig. 4D). Since the deletion of four basic amino acids only partially disrupted the nuclear localization whereas the deletion of 20 amino acids (aa 181 to 200) made the protein entirely cytoplasmic, these results indicate that the four basic residues (RKRR) constitute a part of a larger nuclear signal present within the 20 aa of this region (aa 181 to 200). Another construct, PB1(Δ183-186)-PK, was present primarily in the nucleus (Fig. 4E), suggesting that these residues are not a part of the nuclear signal. PB1(Δ203-216)-PK was present only in the cytoplasm (Fig. 4F), demonstrating the presence of another signal sequence in this region, which also contained a set of charged residues (KRKQR). These data indicate that two discontinuous signal sequences are present and that both of them are required for nuclear localization.

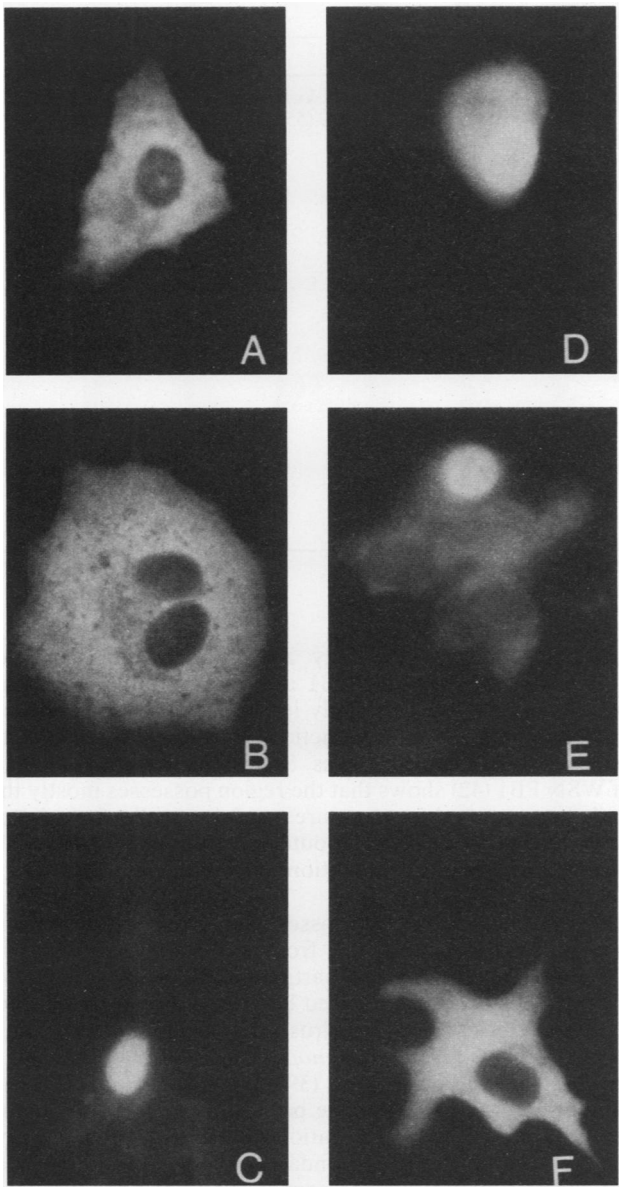

FIG. 4. Localization of chimeric PB1-PK proteins after finer deletions. The conditions for transfection and immunofluorescence were same as described in the legend to Fig. 2. (A) PB1(Δ181-201)-PK; (B) PB1(181-200)-PK; (C) PB1(Δ196-202)-PK; (D) PB1(Δ187-190)-PK; (E) PB1(Δ183-186)-PK; (F) PB1(Δ203-216)-PK.

DISCUSSION

The majority of the nuclear localization signals studied to date consist of a single stretch of seven to nine amino acids possessing a cluster of basic amino acids along with one or more proline residues (43) (Table 2). However, the nuclear localization signal of PB1 possesses some novel features. For example, it is not a single stretch of contiguous residues but appears to be a bipartite signal consisting of two discontinuous regions separated by an intervening sequence. Our study shows that both regions are required for nuclear localization and that either region alone is not sufficient. Although both regions possess a cluster of positive charges (Table 2), they lack the proline residue often present in nuclear signals or the proposed consensus sequence for nuclear transport (K-R/K-X-R/K [5]). Also, the nonbasic nuclear accumulation sequence found in smaller nuclear proteins, such as influenza virus NP and yeast MATα2 (6, 17, 18), is absent in both of these two regions (Table 2). The left-hand domain of the PB1 nuclear signal appears to be present within the nine residues (residues 187 to 195, Table 1), since the removal of residues 183 to 186 or 196 to 202 did not affect nuclear transport. Furthermore, the left-hand boundary of this part of the signal may be Arg-187, since it is the first of the four charged residues (RKRR) present in this region and removal of these four charged residues also partially disrupted nuclear localization (Fig. 4D). However, since the nuclear transport was only partially disrupted, it would suggest that the left-hand domain of the signal extends beyond the four charged residues but ends before residue 196. The right-hand subdomain of the signal is contained within the 14 residues (residues 203 to 216) which also contain a set of charged residues (KRKQR, Table 2).

Although further work is needed to define the precise boundaries of the intervening sequence and the two subdomains as well as the role of the individual amino acids in nuclear localization of PB1, the data presented here demonstrate that two spatially distinct sequences are required for nuclear localization. However, it should be noted that since these experiments were done with chimeric proteins which, although it is unlikely, may have different requirements for nuclear localization, the function of the signal sequences needs to be verified with the native PB1 protein. Also, the minimal sequence requirement of a bipartite signal to translocate a heterologous cytoplasmic protein into the nucleus remains to be determined.

Even though the nuclear localization signals in many proteins appear to be single continuous stretches of amino acids, recent studies have suggested that some nuclear proteins may require the function of two discontinuous regions either for nuclear localization or for localization with the specified intranuclear components. For example, it has been suggested that two signals may be required for the nuclear transport of *Xenopus* nucleoplasmin (4, 9, 11), although recent studies have demonstrated that a single stretch of amino acids of *Xenopus* nucleoplasmin can translocate a cytoplasmic protein (PK) into the nucleus (11). However, this signal is more complex and larger than that of the SV40 T antigen (11). Similarly, the additive function of multiple signals appears to be more efficient in the nuclear translocation of polyomavirus T antigen (39). However, unlike the influenza virus PB1, each of these signals can partially transport the polyomavirus large T antigen into the nucleus. In HTLV $p27^x$ proteins, two adjacent signals have been found around its NH_2 terminus (41). Each signal can independently transport the protein into the nucleus, but both of them are required for the nucleolar localization of $p27^x$.

Recent studies with deletion and mutation analyses have shown that at least two other proteins, *Xenopus* protein N1 (21) and adenovirus DNA-binding protein (31), possess bipartite signals for nuclear localization. *Xenopus* protein N1 requires two signal sequences (VRKKRKT and AK KSKQE) separated by a 10-aa spacer (21). Similarly, it has been shown that the adenovirus DNA-binding protein requires two signals (aa 42 to 46 and aa 84 to 89, Table 2) for nuclear transport and that deletion of either of these sequences disrupts nuclear localization (31). However, deletion analysis of the spacer region was not carried out for either *Xenopus* N1 or adenovirus DNA-binding protein. Like the nuclear localization signal of influenza virus PB1, the signal regions of both *Xenopus* N1 and adenovirus DNA-binding protein also possess clusters of basic amino

TABLE 2. Comparison of nuclear localization signal sequences[a]

Protein (reference)	Sequence (aa nos.)
Influenza virus PB1 (42)	(181)-I T T H F Q **R K R R** V R D N M[T K K M V T Q]R T I G **K R K Q R** L N K R S Y-(217)
Xenopus protein N1 (21)	(530)-L V **R K K R K** T E E E S P L K D K D A K K S K **Q E**-(554)
Adenovirus DNA-binding protein (31)	(42)-**P P K K R** X X X X **P K K K K K**-(84)
SV40 T antigen (19)	P **K K K R K** V E
SV40 VP1 (45)	P T **K R K** G S
SV40 VP2/3 (15, 46)	P N **K K K R K**
Yeast *Mat*α2 (17)	M N K I P I K
Polyomavirus large T antigen (40)	(i) V S **R K R P R** P (ii) P P **K K A R** E D
Influenza virus NP (6)	A A F E D L R V I S
Xenopus nucleoplasmin (4, 9, 11)	A V **K R** P A A T **K K** A G Q A **K K K K** L D
Influenza virus NS1 (16)	(i) Q **R L R R** (ii) P **K K K R K**
HTLV p27[x] (41)	(i) M P K T **R R R P R R** (ii) S Q **R K R** P P T P
Human A/C[b] lamin (26)	S V T **K K R K** L E
Adenovirus E1A (27)	L S V **K R P R** C N
Consensus sequence (5)	**K-R/K-X-R/K**

[a] See Table 1, footnote *a*.
[b] A/C, A lamin and C lamin.

acids. However, the spacer region appears to be variable in both size and content of amino acids. Although the precise boundaries of the signal regions and spacer regions have yet to be determined, these studies show that some nuclear proteins do possess multiple discontinuous signals required either for entry into the nucleus or for binding to specific components within the nucleus.

The mechanism by which two distinct regions can function in nuclear translocation is not clear. One possibility is that one region functions as the nuclear transport signal and the other is the nuclear accumulation (or retention) signal (43), but this is unlikely for the influenza virus PB1. If this were the case, one of the chimeric proteins possessing the nuclear transport signal but lacking the nuclear retention signal would be present both in the nucleus and in the cytoplasm since large proteins, once translocated into the nucleus, are unlikely to passively diffuse out into the cytoplasm efficiently. The second possibility is that these signals have an additive function in nuclear transport, each being rather inefficient alone. Previous studies have shown that the efficiency of nuclear transport may be dependent on the multiple signals present in the proteins (25). A number of proteins, including influenza virus NS1 (16), HTLV p27 (41), rat glycocorticoid receptor (37), and polyomavirus T antigen (39), have two nuclear signals, each of which, however, can function independently of the other, albeit inefficiently, in transporting the protein into the nucleus. However, for the PB1 signal, this possibility also appears unlikely, because even at 16 to 20 h posttransfection, the chimeric proteins containing only one signal sequence were present only in the cytoplasm and not in the nucleus (Fig. 4A, B, and F). The third possibility, that two discrete sequences function cooperatively for nuclear transport, is the most plausible explanation of the data presented in this article. Both regions may act directly in some of the steps involved in nuclear transport, or one of the regions may be indirectly involved in providing the proper structural conformation for presenting the "true" signal. Therefore, the deletion of one region may alter the conformation of the protein, thereby masking or inactivating a more distant signal. Further work is needed to define the specific role of each region in the steps involved in the nuclear transport of PB1.

The sequences of PB1 from influenza A and B viruses possess extensive homology and conservation. The sequence in this region of PB1 also exhibits homology and conserved charges, especially in the basic amino acid residues, and could therefore function as a nuclear signal in both A (42) and B (20) PB1 proteins. Secondary structure analysis of WSN PB1 (42) shows that the region possesses mostly the undefined or random structures which usually form loops and become exposed to the outer surface, thus making this region available for interaction with other proteins, e.g., nuclear pore proteins.

In summary, WSN PB1 possesses a novel type of nuclear localization signal different from the prototype found in SV40 T antigen. It is a bipartite signal consisting of two independent regions separated by an intervening sequence. Recent studies with adenovirus DNA-binding protein (31), *Xenopus* protein N1 (21), *Xenopus* nucleoplasmin (4, 9, 11), and polyomavirus T antigen (39) also support the idea that discontinuous signals may be present in other proteins and required for nuclear localization. Further analysis will be required to define the boundaries of each of the signal regions and the intervening sequence and to delineate their specific roles in transporting proteins into the nucleus.

ACKNOWLEDGMENT

This work was supported by Public Health Service grants R01-AI-12749 and R01-AI-16348 from the National Institutes of Health.

LITERATURE CITED

1. **Akkina, R. K., T. M. Chambers, D. R. Londo, and D. P. Nayak.** 1987. Intracellular localization of the viral polymerase proteins in cells infected with influenza virus and cells expressing PB1 protein from cloned cDNA. J. Virol. **61**:2217–2224.
2. **Ariga, H., H. Klein, A. Levine, and M. Horowitz.** 1980. A cleavage product of adenoviral DNA-binding protein is active in DNA replication in vitro. Virology **101**:307–310.
3. **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.
4. **Burglin, T. R., and E. M. De Robertis.** 1987. The nuclear migration signal of *Xenopus laevis* nucleoplasmin. EMBO J. **6**:2617–2625.
5. **Chelsky, D., R. Ralph, and G. Jonak.** 1989. Sequence require-

ments for synthetic peptide-mediated translocation to the nucleus. Mol. Cell. Biol. 9:2487–2492.
6. **Davey, J., N. J. Dimmock, and A. Coleman.** 1985. Identification of the sequence responsible for nuclear accumulation of the influenza virus nucleoprotein in *Xenopus* oocytes. Cell **40:** 667–675.
7. **De Robertis, E. M.** 1983. Nucleoplasmic segregation of proteins and RNAs. Cell **32:**1021–1025.
8. **Dingwall, C.** 1985. Accumulation of proteins in the nucleus. Trends Biol. Sci. **85:**64–66.
9. **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 J. **6:**69–74.
10. **Dingwall, C., and R. A. Laskey.** 1986. Protein import into the cell nucleus. Annu. Rev. Cell Biol. **2:**367–390.
11. **Dingwall, C., J. Robbins, S. M. Dilworth, B. Roberts, and W. D. Richardson.** 1988. The nucleoplasmin nuclear location sequence is larger and more complex than that of SV40 large T antigen. J. Cell Biol. **107:**841–849.
12. **Flegner, P. L., T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northop, G. M. Ringold, and M. Danielsen.** 1987. Lipofection: a highly efficient, lipid mediated DNA-transfection procedure. Proc. Natl. Acad. Sci. USA **84:**7413–7417.
13. **Franke, W. W., U. Scheer, G. Krohne, and E. D. Jarasch.** 1981. The nuclear envelope and the architecture of the nuclear periphery. J. Cell Biol. **91:**(3, pt. 2):39s–50s.
14. **Fuerst, T. R., A. L. Earl, and B. Moss.** 1987. Use of a hybrid vaccinia virus-T7 RNA polymerase system for expression of target genes. Mol. Cell. Biol. **7:**2538–2544.
15. **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.
16. **Greenspan, D., P. Palese, and M. Krystal.** 1988. Two nuclear location signals in the influenza virus NS1 nonstructural protein. J. Virol. **62:**3020–3026.
17. **Hall, M. N., L. Hereford, and I. Herskowitz.** 1984. Targeting of *E. coli* β-galactosidase to the nucleus in yeast. Cell **36:**1057–1065.
18. **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.
19. **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.
20. **Kemdirim, S., J. Palefsky, and D. J. Briedis.** 1986. Influenza B virus PB1 protein: nucleotide sequence of genome RNA predicts a high degree of structural homology with the corresponding influenza A virus polymerase protein. Virology **152:**126–135.
21. **Kleinschmidt, J. A., and A. Seiter.** 1988. Identification of domains involved in nuclear uptake and histone binding of protein N1 of *Xenopus laevis*. EMBO J. **7:**1605–1614.
22. **Krug, R. M., F. V. Alonso-Caplen, I. Julkunen, and M. G. Katze.** 1989. Expression and replication of influenza virus genome, p. 89–252. *In* R. M. Krug (ed.), Influenza viruses. Plenum Publishing Corp., New York.
23. **Lamb, R. A., and C.-J. Lai.** 1984. Expression of unspliced NS1 mRNA, spliced NS2 mRNA, and a spliced chimera RNA from cloned influenza virus NS DNA in an SV40 vector. Virology **135:**139–147.
24. **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.
25. **Landford, R. E., P. Kanda, and R. C. Kennedy.** 1986. Induction of nuclear transport with a synthetic peptide homologous to the SV40 T antigen transport signal. Cell **46:**575–582.
26. **Loeinger, L., and F. McKeon.** 1988. Mutations in the nuclear lamin proteins resulting in their aberrant assembly. EMBO J. **7:**2301–2309.
27. **Lyons, R. H., B. Q. Ferguson, and M. Rosenberg.** 1987. Pentapeptide nuclear localization signal in adenovirus E1A. Mol. Cell. Biol. **7:**2451–2456.
28. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Moreland, R. B., G. L. Langevin, R. H. Singer, R. L. Garcea, and L. M. Hereford.** 1987. Amino acid sequences that determine the nuclear localization of yeast histone H2B. Mol. Cell. Biol. **7:**4048–4057.
30. **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.
31. **Morin, N., C. Delsert, and D. F. Klessig.** 1989. Nuclear localization of adenovirus DNA-binding protein: requirements of two signals and complementation during viral infection. Mol. Cell. Biol. **9:**4372–4380.
32. **Muirhead, H., D. A. Clayden, D. Barford, C. G. Lorimer, L. A. Forthergill-Gilmore, E. Shiltz, and W. Schmitt.** 1986. The structure of cat muscle pyruvate kinase. EMBO J. **5:**475–481.
33. **Nayak, D. P., K. Tobita, J. M. Janda, A. R. Davis, and B. K. De.** 1978. Homologous interference mediated by defective interfering influenza virus derived from a temperature-sensitive mutant of influenza virus. J. Virol. **28:**375–386.
34. **Newmeyer, D. D., and D. J. Forbes.** 1988. Nuclear import can be separated into distinct steps in vitro: nuclear pore binding and translocation. Cell **52:**641–653.
35. **Newmeyer, D. D., J. M. Lucocq, T. R. Bürglin, and E. M. De Robertis.** 1986. Assembly in vitro of nuclei active in nuclear protein transport: ATP is required for nucleoplasmin accumulation. EMBO J. **5:**501–510.
36. **Paine, P. L., L. C. Moore, and S. B. Horowitz.** 1975. Nuclear envelope permeability. Nature (London) **254:**109–114.
37. **Picard, D., and K. Yamamoto.** 1987. Two signals mediate hormone-dependent nuclear localization of the glucocorticoid receptor. EMBO J. **6:**3333–3340.
38. **Richardson, W. D., A. D. Mills, S. M. Dilworth, R. A. Laskey, and C. Dingwall.** 1988. Nuclear protein migration involves two steps: rapid binding at the nuclear envelope followed by slower translocation through nuclear pores. Cell **52:**655–664.
39. **Richardson, W. D., B. L. Roberts, and A. E. Smith.** 1986. Nuclear location signals in polyoma virus large-T. Cell **44:** 75–85.
40. **Roberts, B. L., W. D. Richardson, and A. E. Smith.** 1987. The effect of protein context on nuclear location of signal function. Cell **50:**465–475.
41. **Siomi, H., H. Shida, S. H. Nam, T. Nosaka, M. Maki, and M. Hatanaka.** 1988. Sequence requirements for nuclear localization of human T cell leukemia virus type I pX protein which regulates viral RNA processing. Cell **55:**197–209.
42. **Sivasubramanian, N., and D. P. Nayak.** 1982. Sequence analysis of the polymerase 1 gene and secondary structure prediction of polymerase 1 protein of human influenza virus A/WSN/33. J. Virol. **44:**321–329.
43. **Smith, A. E., D. Kalderon, B. L. Roberts, W. H. Colledge, M. Edge, P. Gillet, A. Markham, E. Paucha, and W. D. Richardson.** 1985. The nuclear location signal. Proc. R. Soc. London Biol. Sci. **226:**43–58.
44. **Smith, G. L., J. Z. Levin, P. Palese, and B. Moss.** 1987. Synthesis and cellular location of the ten influenza polypeptides individually expressed by recombinant vaccinia viruses. Virology **160:**336–345.
45. **Wychowski, C., D. Benichou, and M. Girard.** 1986. Domain of SV-40 capsid polypeptide VP1 that specifies migration into the cell nucleus. EMBO J. **5:**2569–2576.
46. **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.
47. **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 *E. coli*. Proc. Natl. Acad. Sci. USA **80:**6105–6109.