JOURNAL OF VIROLOGY, Feb. 1995, p. 801–808
0022-538X/95/$04.00+0

Vol. 69, No. 2

Human Foamy Virus Bel1 Transactivator Contains a Bipartite Nuclear Localization Determinant Which Is Sensitive to Protein Context and Triple Multimerization Domains

JUN CHANG,[1] KI JEONG LEE,[1] KYUNG LIB JANG,[1]† EUN KYEONG LEE,[1]‡ GWAN HYUK BAEK,[1] AND YOUNG CHUL SUNG[1,2]*

Department of Life Science[1] and Center for Biofunctional Molecules,[2] Pohang University of Science and Technology, Pohang 790-784, Republic of Korea

Received 26 July 1994/Accepted 25 October 1994

The Bel1 protein of human foamy virus is a 300-amino-acid nuclear regulatory protein which transactivates the gene expression directed by the homologous long terminal repeat and the human immunodeficiency virus type 1 long terminal repeat. While previous reports suggested that the single basic domain of Bel1 from residues 211 to 222 and/or 209 to 226 is necessary and sufficient for efficient nuclear localization (L. K. Venkatesh, C. Yang, P. A. Theodorakis, and G. Chinnandurai, J. Virol. 67:161–169, 1993; F. He, J. D. Sun, E. D. Garrett, and B. R. Cullen, J. Virol. 67:1896–1904, 1993), our recent data showed that another basic domain, from amino acid residues 199 to 200, is also required for nuclear localization of Bel1 (C. W. Lee, C. Jun, K. J. Lee, and Y. C. Sung, J. Virol. 68:2708–2719, 1994). To clarify this discrepancy, we constructed various *bel1-lacZ* chimeric constructs and several linker insertion mutants and determined their subcellular localization. When the region of Bel1 containing basic domains was placed at an internal site of the *lacZ* gene, the nuclear localization signal (NLS) of Bel1 consisted of two discontinuous basic regions separated by an intervening sequence. Moreover, insertion of specific amino acids between two basic regions disrupted the activity of the Bel1 NLS. On the other hand, Bel1 residues 199 and 200 were not required to direct the Bel1–β-galactosidase chimeric protein to the nucleus when the Bel1 NLS was appended to the amino terminus of β-galactosidase. These results indicate that the function of the Bel1 NLS is sensitive to the protein context within which the sequence is present. In addition, we demonstrated that the Bel1 protein forms a multimeric complex in the nuclei of mammalian cells by using a sensitive in vivo protein-protein interaction assay. Mutational analyses revealed that the regions which mediate multimer formation map to three domains of Bel1, i.e., residues 1 to 31, 42 to 82, and 82 to 111. Furthermore, our results show that the region of Bel1 from residues 202 to 226 prevents Bel1 from forming a multimeric complex.

The *bel1* gene of human foamy virus (HFV) encodes a 300-amino-acid regulatory protein termed Bel1, which is a potent transcriptional activator required for transcription from the homologous and human immunodeficiency virus type 1 long terminal repeat promoters (18a, 18b, 25, 27, 35). The Bel1 protein has been shown to be essential for virus replication in vitro (28) and is localized to the nuclei of cells (18b, 28). Previous studies predict that the Bel1 protein is composed of discrete, function-specific modules (13, 26, 39a, 39b). The carboxy terminus of Bel1 represents the domain responsible for the autonomous 30-amino-acid transcriptional activation domain (26, 39b) and the augmenting domain for HFV long terminal repeat-directed transactivation (39b). The central region of Bel1 contains the promoter-binding (13) and/or regulatory domain which controls the transcriptional activation domain (26).

Most nuclear proteins contain specific sequences that facilitate their transport into the nucleus through the nuclear membrane. The nuclear localization signals (NLSs) can be classified into three categories: prototypic NLSs consisting of short stretches of basic amino acids, such as the simian virus 40 large T antigen (12, 17a, 17b, 23, 24, 38); relatively rare motifs with few basic residues, such as the influenza virus nucleoprotein NLS (5); and bipartite NLSs consisting of two clusters of basic residues separated by 10 to 12 amino acids, commonly including proline residues, such as those of *Xenopus laevis* nucleoplasmin (7, 8, 36) and N1 (20). Many NLSs show little homology to other sequences except in that they possess many basic residues (6). On the basis of the simian virus 40 large T antigen paradigm, two basic clusters within Bel1 located at positions 193 to 200 and 214 to 223 were suggested as a putative NLS (11). Our recent mutational analyses of the Bel1 protein have revealed that missense mutations altering either one of the two basic segments (amino acids 199 to 200 and 213 to 223) result in predominantly cytoplasmic accumulation of Bel1 (26), suggesting that the Bel1 NLS consists of two basic regions separated by a 12-amino-acid spacer (a bipartite NLS). However, other mutational studies have reported that the Bel1 NLS is localized to the region defined by amino acid residues 209 to 226 (13) and/or 211 to 225 (39b) as a single basic cluster.

To identify and characterize the NLS sequence, most studies have used chimeric constructs consisting of various portions of a nuclear protein and a cytosolic protein. In this study, to define the precise region required for nuclear targeting of Bel1, various portions of Bel1 were placed at an internal region or at the amino terminus of β-galactosidase (β-gal) and tested for the ability to locate a large cytosolic β-gal polypeptide to the nucleus. Our results show that the Bel1 NLS consists of two essential basic amino acid domains which function in an inter-

* Corresponding author. Phone: 82-562-279-2294. Fax: 82-562-279-2199.
† Present address: Department of Microbiology, Pusan National University, Pusan, Republic of Korea.
‡ Present address: Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT 06269.

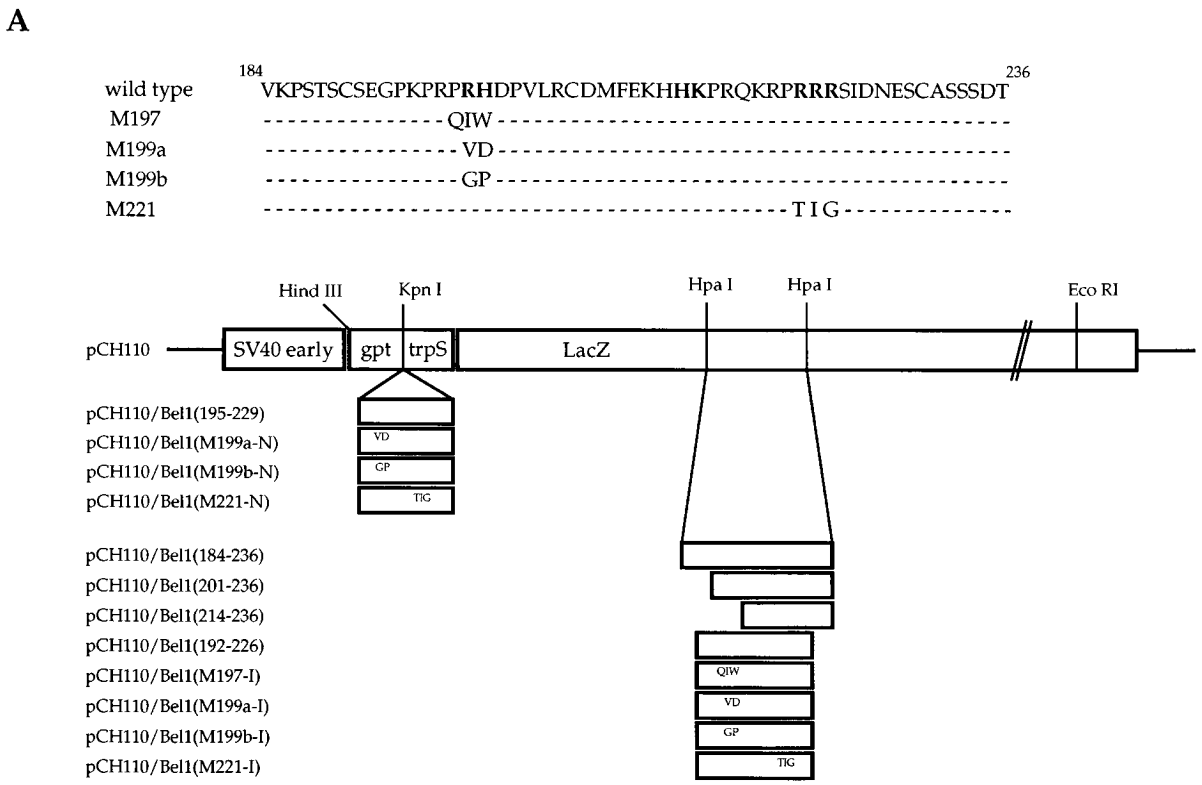

FIG. 1. (A) Schematic diagram of chimeric Bel1–β-gal expression vectors. The simian virus 40 early promoter of pCH110 directs high-level expression of a chimeric protein composed of 40 residues of *Escherichia coli gpt* and 20 residues derived from *trpS* fused to the *lacZ*-encoded β-gal protein. A series of internal Bel1–β-gal fusion vectors were constructed by replacing various regions of the Bel1 and mutant derivatives with the 624-bp *Hpa*I-*Hpa*I DNA fragment of pCH110. The amino-terminal chimera plasmids were constructed by inserting PCR-amplified DNA fragments of Bel1 into the *Kpn*I site of pCH110. The amino acid sequences of Bel1 and mutant proteins encoded by these chimeric expression vectors are presented at the top in the single-letter code, and residues important for Bel1 nuclear targeting, identified by point mutational analysis, are in boldface. Amino acid residues of Bel1 in the chimeric constructs are in parentheses. *Hpa*I recognition sites were removed by insertion of *bel1* DNA fragments. (B) Immunofluorescence images of wild-type β-gal and internal Bel1–β-gal fusion derivatives. (C) Immunofluorescence images of amino-terminal Bel1–β-gal chimeras.

dependent manner, as does that of *Xenopus* nucleoplasmin. However, the carboxy-terminal single basic domain of Bel1 is sufficient to confer the nucleus localization property on the cytosolic β-gal protein only when the basic domain is attached to the amino terminus of β-gal, indicating that the ability of the Bel1 NLS is location dependent.

Many eukaryotic transcription factors take the form of dimers or higher-order multimers as an absolute requirement for biological activity (2a, 4, 15, 22, 29, 40–42). Structural domains or motifs mediating specific interactions have been characterized, such as the leucine zipper, helix-loop-helix, and helix-span-helix proteins (16, 21, 22). In addition, several proteins, such as Myc, E12, and AP4, possess more than one dimerization domain (19, 30). To investigate whether the Bel1 protein functions as a monomer or an oligomer, we used an in vivo assay of protein-protein interaction (10). Our results show that the Bel1 protein forms a multimeric complex in the nuclei of mammalian cells, and domains for this specific multimerization were mapped to three regions i.e., residues 1 to 31, 42 to 82, and 82 to 111. Furthermore, we found that the region of Bel1 from residues 202 to 226 inhibits multimer formation.

MATERIALS AND METHODS

Construction of various Bel1–β-gal chimeric plasmids and linker insertion mutants. Plasmid pRc/CMVBel1, containing the entire *bel1* open reading frame, was described previously (26). Plasmid pCH110 (Pharmacia LKB) contains the full-length prokaryotic *lacZ* gene under control of the simian virus 40 early promoter. To construct pCH110/Bel1(184-236), the A nucleotide at position 9988 (all sequence positions refer to the full-length proviral HFV DNA; 28) of the *bel1* gene in pRc/CMVBel1 was changed to C by site-directed mutagenesis to create an *Hpa*I site. The plasmid DNA was then digested with *Hpa*I and *Nco*I and the end was made flush with T4 DNA polymerase. The 150-bp DNA fragment was isolated by electroelution from a 4% polyacrylamide gel and was replaced with the *Hpa*I fragment (624 bp) of vector pCH110 (Fig. 1A). pCH110/Bel1(201-236) and pCH110/Bel1(214-236) were constructed by the same procedure, except that the inserts were the *Hin*cII-*Nco*I fragment (105 bp) of pRc/CMVBel1-M199a (26) and the *Hin*cII-*Nco*I fragment (66 bp) of pRc/CMVBel1-M213 (26), respectively. To construct pCH110/Bel1(192-226), the Bel1 sequences from amino acids 192 to 226 were amplified with two primers (5′-TTCTAC CAGTTGCCTCGAGGGTCCA-3′ and 5′-AAGCACATGACTCGAGATCGA TGGATC-3′; [the *Xho*I recognition sequence is underlined]) and then digested with *Xho*I. The resulting DNA fragment was cloned in frame into the *Hpa*I site of pCH110 (Fig. 1A). Several derivatives of pCH110/Bel1(192-226), except pCH110/Bel1(M197-I), were also generated by the same procedure from the pRc/CMVBel1 derivatives (26; Fig. 1A). pCH110/Bel1(M197-I) was constructed by the same method following site-directed mutagenesis to change residues P-197, R-198, and P-199 to Q, I, and W, respectively, with primer 5′-GACAG GATCGTGCCCAAGATCTGGTTTTGGACCTT-3′. Suitable constructs were identified by both DNA sequencing and Western blotting (immunoblotting) with a monoclonal anti-β-gal antibody (Boehringer Mannheim Biochemicals). To construct amino-terminal chimeric plasmids, the region of Bel1 from residues 195 to 229 and its missense mutant derivatives (26) were amplified by PCR with primers 5′-AGTTGCTCAGAGGTACCAAAACCAAGACCT-3′ and 5′-TGT CACTACTGGGTACCCATGACTCATTATC-3′ (the *Kpn*I recognition sequence is underlined) and then digested with *Kpn*I. The resulting fragments were cloned into the unique *Kpn*I site of pCH110 (Fig. 1A). A synthetic oligonucleotide (5′-TGATGCTTTGTTAACATGTCACAG-3′ [the *Hpa*I recognition sequence is underlined]) complementary to nucleotides 10048 to 10071 of the *bel1* gene was used to create a novel *Hpa*I site at nucleotide 10057 to generate pRc/CMVBel1-M, resulting in the conversion of F-209 and D-210 to L and T, respectively. The 12-mer phosphorylated linkers (*Sal*I, *Eco*RI, *Hin*dIII, and *Xba*I

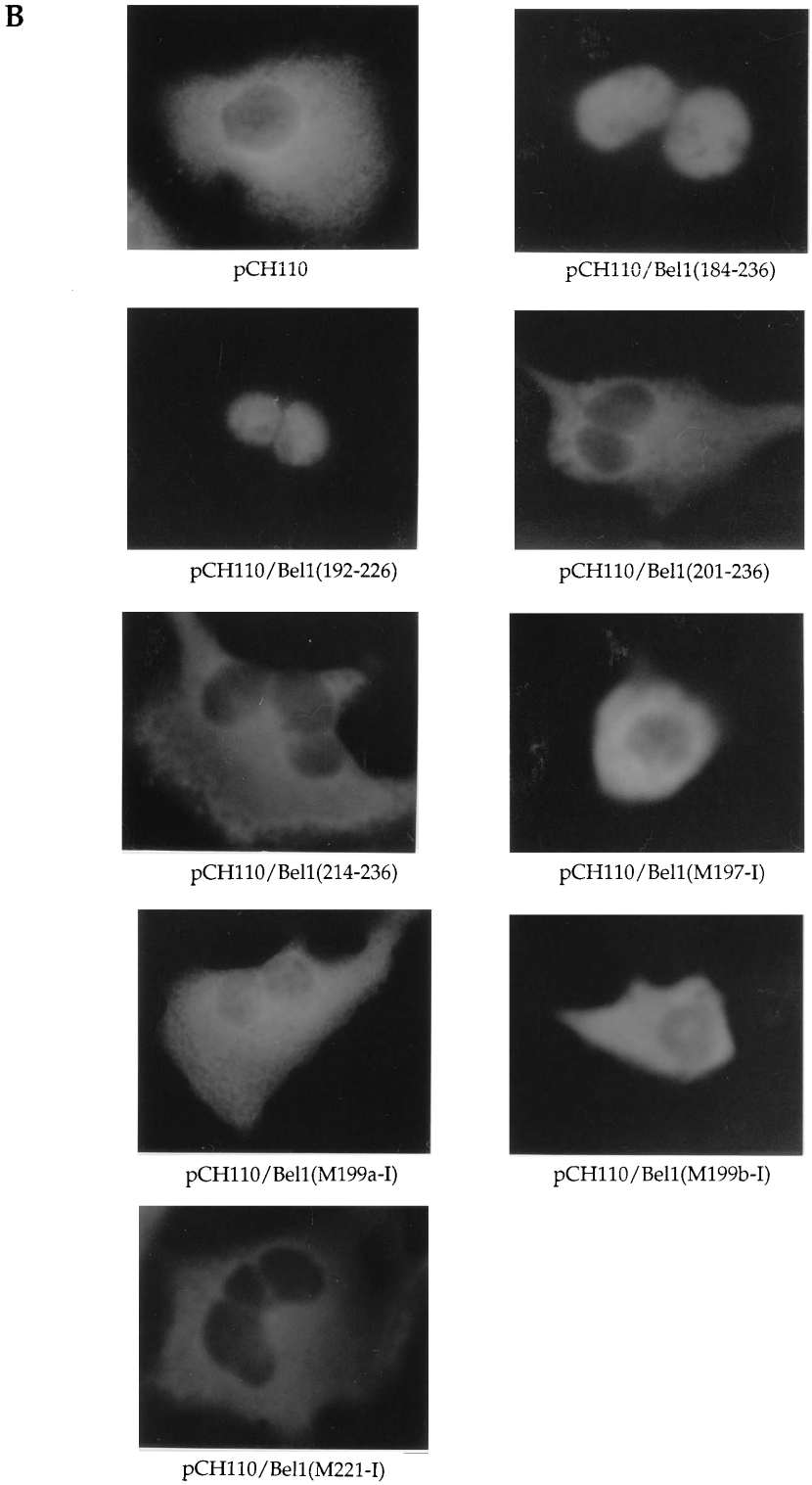

FIG. 1—*Continued.*

linkers; New England Biolabs) were inserted into the *Hpa*I site of pRc/CMV Bel1-M to generate pRc/CMVBel1-M-SalI, pRc/CMVBel1-M-HindIII, pRc/CMVBel1-M-EcoRI, and pRc/CMVBel1-M-XbaI, respectively. The correct in-frame sequences of these plasmids were confirmed by DNA sequencing and immunoblotting with rabbit anti-Bel1 serum (26).

Construction of Gal4-Bel1 fusion plasmids. Gal4-Bel1 derivatives were generated by inserting various *bel1* DNA fragments obtained by appropriate restriction enzyme digestions and PCR into the *Sma*I site of pSG424 (9). Gal4-Bel1(1-260), Gal4-Bel1(1-226), Gal4-Bel1(1-82), and Gal4-Bel1(82-150) were previously described (26). Gal4-Bel1(1-201), Gal4-Bel1(1-196), Gal4-Bel1(1-180), Gal4-Bel1(1-173), Gal4-Bel1(1-150), Gal4-Bel1(1-123), Gal4-Bel1(1-73), Gal4-Bel1(1-54), Gal4-Bel1(1-43), and Gal4-Bel1(1-31) were constructed by inserting the

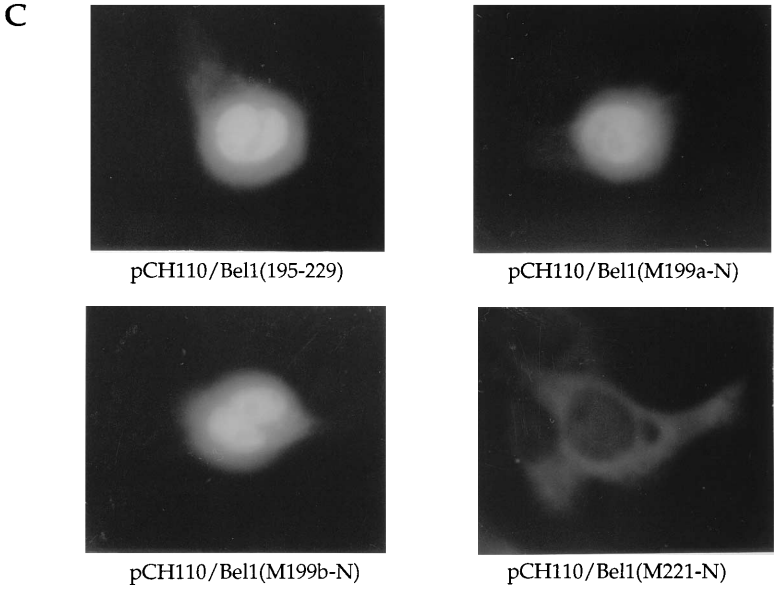

FIG. 1—*Continued.*

PCR-amplified DNA fragments into the *Sma*I site of pSG424. An *Sma*I site was generated in a 5′ primer to facilitate in-frame fusion of the construction (5′-AG ATTGTATCCCGGGTTCCTACGAAAA-3′ [the *Sma*I recognition sequence is underlined]). Several 3′ primers were designed for generation of various carboxy terminus deletion mutants (201, 5′-ACAGCGAAGGACCATATGGT GCCTAGG-3′; 196, 5′-CAGGTTTTGGTTCATGATCCGTGCTAG-3′; 180, 5′-TTAACAAGAGATCTGTTTCTTTGAT-3′; 173, 5′-AATAAACACAGCA GATCCAAC-3′; 150, 5′-GGGTATTCTACCCGTCGACTTCAACCTTA-3′; 123, 5′-TTCCCAAGAAACTGTAAAAGG-3′; 73, 5′-AGGATGTTTGTCCT CAATTTC-3′; 54, 5′-TTTGGTATATCGTCTGGGGCG-3′; 43, 5′-TTCAGGT TCCTCAGCAATAGT-3′; 31, 5′-CCTCTAGAGCATTTTCAGGGCCAAC-3′). Gal4-Bel1(1-18) was obtained by partial digestion of Gal4-Bel1(1-54) with *Pst*I and then religation. Gal4-Bel1(42-82) and Gal4-Bel1(82-120) were constructed by inserting a *Kpn*I-*Bam*HI fragment and an *Apa*I-*Bam*HI fragment from pRc/CMVBel1-M40 and pRc/CMVBel1-M120 (26) into the *Sma*I site of pSG424, respectively. Gal4-Bel1(112-150) and Gal4-Bel1(150-260) were also generated by inserting the PCR-amplified DNA fragments into the *Sma*I site of pSG424 (112sense, 5′-AAGGATCCCTTTTACAGTTTCTTG-3′; 150sense, 5′-CTGAA TTCGGAATTTGGGTAAAAA-3′; 260antisense, 5′-AAGTAGCCCTGATAG TAGCGGTCC-3′). pSbel1-S and Bel1(1-260)VP16 were previously described (26). pMC1, encoding herpes simplex virus type 1 VP16 amino acids 1 to 490, was previously described (1).

Cell culture, transfection, and chloramphenicol acetyltransferase (CAT) assays. COS-7 and BHK-21 cells were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum. For indirect immunofluorescence, approximately 2×10^5 cells on a glass coverslip were transfected with the indicated plasmid DNA by the DEAE-dextran method (34). Cells (10^6) were plated on a 100-mm-diameter dish at 24 h before transfection and transfected with 2 μg of each of the reporter and activator plasmids. The cells were harvested and assayed for CAT activity at 48 h after transfection. The difference in transfection efficiency was normalized by using a second reporter plasmid, pGL2 (Promega), containing the luciferase gene. Luciferase activity was measured with a luciferase assay kit (Promega) in accordance with the supplier's recommendation. CAT enzyme reactions were carried out for 2 h of incubation. When the level of acetylation of chloramphenicol was more than 90%, the lysates were diluted for the reaction and the CAT values were corrected by the dilution factor. All of the CAT assay data reported are from points in the linear range of the assay.

Immunofluorescence. For indirect immunofluorescence, transfected cells were fixed and permeabilized at 48 h after transfection with 95% methanol in phosphate-buffered saline at −20°C for 10 min. Cells were then reacted with a 1:100 dilution of a mouse anti-β-gal monoclonal antibody (Boehringer Mannheim Biochemicals), followed by a 1:80 dilution of anti-mouse immunoglobulin–fluorescein (Boehringer Mannheim Biochemicals). When pRc/CMVBel1 derivatives were transfected, a 1:80 dilution of polyclonal anti-Bel1 rabbit serum (26) and a 1:80 dilution of anti-rabbit immunoglobulin G–fluorescein isothiocyanate (GIBCO BRL) were used as the primary and secondary antibodies, respectively. The cells were photographed on a Carl Zeiss microscope equipped for fluorescent illumination at a magnification of ×400 with Kodak Gold 400 film.

RESULTS

The Bel1 NLS is a bipartite motif and is sensitive to the protein context. Previous reports have identified the Bel1 NLS by several methods, such as deletion (39b), missense mutation (13, 26), and Bel1–β-gal chimeric constructs (13). However, there is a conspicuous discrepancy in the results. In one case, missense mutations introduced into Bel1 residues 199 and 200 (R-199 and H-200→VD or GP) abolished nuclear targeting of Bel1 (26), while in another case, an M197 missense mutant (R-197, P-198, and R-199→QIW) and a Bel1 deletion mutant (Δ194-200) retained nuclear targeting activity (13, 39b). To determine the precise peptide sequence within Bel1 that is sufficient to function independently as an NLS, expression vectors containing several subregions of Bel1 fused to β-gal were constructed and their ability to target a large cytoplasmic β-gal protein to the nucleus was tested (Fig. 1A). We placed a portion of Bel1 in an internal region of β-gal since the Bel1 NLS, in its normal context, is located not at the amino- or carboxy-terminal end but in an internal region. The subcellular localization of these chimeric proteins was examined by indirect immunofluorescence staining of COS-7 cells transfected with each expression plasmid (Fig. 1B). pCH110 expressing wild-type β-gal was detected predominantly in the cytoplasm with little nuclear accumulation (Fig. 1B). In contrast, chimeric proteins containing the region of Bel1 from either residues 184 to 236 or 192 to 226 were detected predominantly in the nucleus. However, chimeric proteins of pCH110/Bel1(201-236) and pCH110/Bel1(214-236) containing the single basic domain of Bel1 displayed a mostly cytoplasmic staining indistinguishable from that of pCH110 (Fig. 1B). Correct expression of the fusion proteins was verified by Western blot analysis with an anti-β-gal monoclonal antibody (data not shown). These results suggest that the region of Bel1 from residues 192 to 226 including two basic segments is required for a large cytoplasmic protein to be localized to the nucleus.

Since some deletion mutants could change the structure significantly enough to mask the NLS motif required for intra- or intermolecular interactions, we generated various substitution mutant derivatives of pCH110/Bel1(192-226) (Fig. 1A).

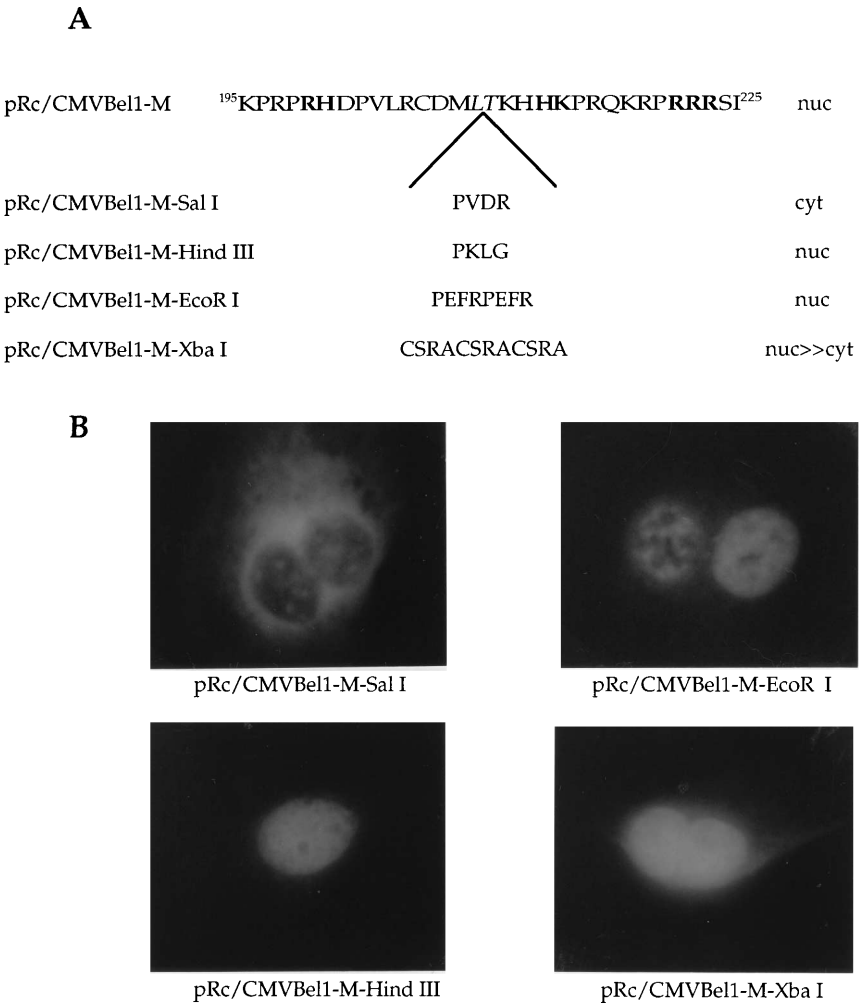

FIG. 2. Structure and immunofluorescence of mutant Bel1 proteins with linker insertions. (A) Sequences of Bel1 from residues 195 to 225 and pRc/CMVBel1-M linker insertion mutants. Phosphorylated synthetic linkers (12 bp) containing restriction endonuclease cleavage sites were inserted in frame into pRc/CMVBel1-M. The subcellular locations of the Bel1 proteins are given on the right. nuc, nucleus; cyt, cytoplasm. (B) Immunofluorescence images of the mutants indicated.

pCH110/Bel1(M197-I), pCH110/Bel1(M199a-I), and pCH110/Bel1(M199b-I), which have missense mutations in the first basic cluster of the Bel1 NLS, showed severely impaired ability to direct chimeric proteins to the nucleus (Fig. 1B). Interestingly, pCH110/Bel1(M197-I) has the same mutation as M197, which was previously reported to be localized in the nucleus of transfected COS cells (13). We think that the inconsistency resulted from the use of different constructs and conditions in which the segment of Bel1 was fused to β-gal. As expected, pCH110/Bel1(M221-I) was localized exclusively in the cytoplasm (Fig. 1B). These results further confirmed that two discontinuous basic regions of Bel1, from residues 199 to 223, are absolutely required to direct cytoplasmic β-gal to the nucleus.

However, previous reports showed that the single basic amino acid cluster of Bel1 from residues 209 to 226 and/or 211 to 225 was both necessary and sufficient for nuclear localization of the Bel1 protein (13, 39b). Previous studies investigated the activity of the Bel1 NLS by placing the potential NLS regions at the amino terminus of β-gal. To determine whether the discrepancy between our results and other previous reports is due to the difference in the position of the NLS within the *lacZ* gene, we fused the region of Bel1 from residues 195 to 229 and its derivatives to the amino terminus of β-gal (Fig. 1A) and then determined the subcellular localization of these Bel1–β-gal chimeric proteins (Fig. 1C). As expected, pCH110/Bel1(195-229), containing the wild-type Bel1 sequence, but not pCH110/Bel1(M221-N), can direct β-gal to the nuclei of transfected cells. In contrast, pCH110/Bel1(M199a-N) and pCH110/Bel1(M199b-N) are localized preferentially in the nuclei of transfected cells, which is consistent with earlier reports (13). These results suggest that the second basic cluster of the Bel1 NLS itself is sufficient to function as a minimal nuclear localization signal only when it is located at the amino terminus of β-gal, indicating that the ability of the Bel1 NLS is dependent on the protein context within which it is present.

Effects of linker insertion between the two basic clusters. Sequence comparison of the Bel1 NLS with other bipartite NLSs suggests that the NLS of the Bel1 protein is composed of a bipartite basic amino acid motif (Table 1). It was reported that the correct positioning of the two basic amino acid motifs could be essential for efficient nuclear targeting, whereas the precise sequence of the spacer region might not be important (6, 32, 36). To investigate whether the activity of the Bel1 NLS is affected by alteration of the length of the spacer between the two basic clusters, we introduced an *Hpa*I site between two basic domains by site-directed mutagenesis to construct pRc/

TABLE 1. Alignment of nuclear targeting sequences in Bel1 with the bipartite motifs of other nuclear proteins

Protein	Species of origin	Position[a]	Sequence
Nucleoplasmin	X	155	**KR**paatKKagqa**KKKK**l
N1/N2	X	534	**KR**Kteeesplkd**K**da**KK**
c-FOS	H, M, R		**RR**eRnKmaaaKc**R**n**RRR**
c-JUN	H, M		**KR**mRnRiaasKc**RKRK**l
C/EBP	R	275	**KK**svdKnsneyRv**RR**e**R**
C/EBP	R	288	**RR**eRnniavRKs**R**d**K**a**K**
GCN4	Y	230	**KR**aRnteaaRRs**R**a**RK**l
Bel1	H	199	**RH**dpvlRcdmfeKh**HK**p**R**q**KR**p**RRR**

[a] Each number refers to the position of the first amino acid shown in the primary sequence of the protein. The single-letter amino acid code is used, basic amino acids are in uppercase, and all other amino acids are in lowercase. X, xenopus; H, human; M, mouse; R, rat; Y, yeast (*Saccharomyces cerevisiae*).

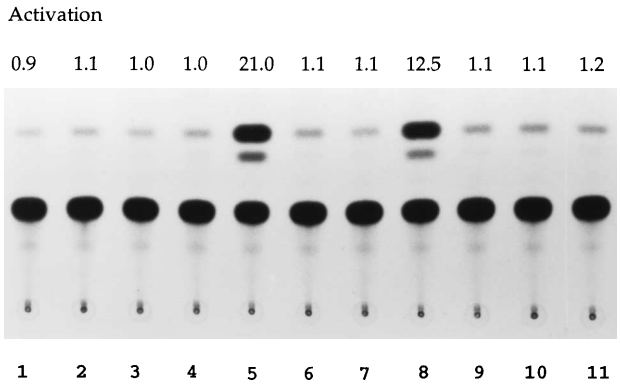

FIG. 3. The Bel1 protein forms a multimeric complex in nuclei of mammalian cells. Plasmids pSG424, pSbel1-S, pMC1, and G5E1bCAT are described in the text. G5E1bCAT and various effector plasmids were cotransfected into BHK-21 cells, and CAT activity was measured at 48 h after transfection as described in Materials and Methods. Lanes: 1, pSG424; 2, Gal4-Bel1(1-260); 3, Bel1(1-260)VP16; 4, pSG424 and Bel1(1-260)VP16; 5, Gal4-Bel1(1-260) and Bel1(1-260)VP16; 6, pSbel1-S; 7, pSG424 and pSbel1-S; 8, Gal4-Bel1(1-260) and pSbel1-S; 9, pMC1; 10, pMC1 and pSG424; 11, pMC1 and Bel1(1-260)VP16.

CMVBel1-M and then inserted several synthetic linkers into the *Hpa*I site (Fig. 2A). As expected, pRc/CMVBel1-M was concentrated in the nuclei of transfected cells. Insertion of the amino acid sequence PVDR (5′-CCGGTCGACCGG-3′; *Sal*I linker) caused cytoplasmic accumulation of Bel1 (Fig. 2B). In contrast, insertion of single or double repeats of other linkers (PKLG [pRc/CMVBel1-M-HindIII] and PEFRPEFR [pRc/CMVBel1-M-EcoRI]) to give additional spacer lengths of four and eight amino acids had no effect on the nuclear targeting of Bel1. When the spacer length was increased by insertion of 12 amino acids (CSRACSRACSRA [pRc/CMVBel1-M-XbaI]), predominant nuclear staining with slight cytoplasmic accumulation was observed (Fig. 2B). These results indicate that the spacer region of the Bel1 NLS is moderately tolerant of insertion mutations, as is nucleoplasmin (36), but not when their separation is extended by a certain specific amino acid sequence, such as PVDR.

The Bel1 protein forms multimeric complexes in the nuclei of mammalian cells. Recently, it has been reported that several transactivators of human retroviruses, such as the Tat and Rev proteins of human immunodeficiency virus type 1 and the Rex protein of human T-cell leukemia virus type 1, form multimeric complexes in the nuclei of eukaryotic cells (2a, 2b, 42). To investigate whether the Bel1 protein functions as a monomer or an oligomer, like human immunodeficiency virus type 1 Tat and Rev, an in vivo protein-protein interaction assay was performed. We constructed a Gal4-Bel1(1-260) chimeric plasmid that contains the amino-terminal 147 residues of Gal4 to which Bel1 residues 1 to 260 were fused and fusion plasmid Bel1(1-260)VP16, which expresses a fusion protein consisting of the Bel1 sequence from residues 1 to 260 and the herpes simplex virus type 1 VP16 transcriptional activation domain (amino acids 423 to 490). The various expression vectors were transfected into BHK-21 cells, either alone or in combination but always in the presence of CAT reporter plasmid G5E1bCAT (3). Effective protein-protein interaction by Bel1 is predicted to spatially juxtapose the Gal4 DNA-binding domain and VP16 transcriptional activation domain encoded by the different plasmids, which in turn leads to transcriptional activation of the G5E1bCAT reporter gene, which contains Gal4-binding sites in the promoter. As shown in Fig. 3, expression of pSG424 containing a DNA-binding domain (amino acids 1 to 147) of Gal4, Gal4-Bel1(1-260), Bel1(1-260)VP16, pMC1, or pSbel1-S containing the entire *bel1* open reading frame alone produced no significant activation of G5E1bCAT. However, coexpression of Gal4-Bel1(1-260) with Bel1(1-260)VP16 or pSbel1-S resulted in stimulation of CAT gene expression by about 21- and 12-fold, respectively (Fig. 3, lanes 5 and 8), while cotransfection of pSG424 and Bel1(1-260)VP16 or pSbel1-S did not increase G5E1bCAT-derived gene expression (Fig. 3, lanes 4 and 7). The significant increases in CAT gene expression appear to reflect a specific Bel1-Bel1 protein interaction in vivo. The higher level of CAT activity obtained with Bel1(1-260)VP16 than with pSbel1-S may be due to the difference in the strength of the transcriptional activation domain between VP16 and Bel1. To examine whether the Bel1-Bel1 interaction is specific, we cotransfected plasmid pMC1, expressing the entire domain of VP16, with Gal4-Bel1(1-260). Cotransfection of pMC1 and Gal4-Bel1(1-260) failed to produce significant stimulation of the G5E1b promoter (Fig. 3, lane 11), indicating that the Bel1 protein did not interact with VP16 in vivo. These results suggest that the Bel1 protein forms a multimeric complex in the nuclei of mammalian cells.

Multiple domains are involved in the oligomerization of Bel1. To identify the sequences required for multimerization of Bel1, we generated a batch of Gal4-Bel1 chimeric proteins by introducing a series of deletions toward the amino-terminal side starting at residue 260 by using appropriate restriction enzymes or PCR and tested their ability to form multimers by assaying transient cotransfection into BHK-21 cells (Fig. 4). Deletion of residues 260 to 227 from Bel1 produced a slight increase in the efficiency of multimerization. Interestingly, further deletion of residues 226 to 202 from Bel1 produced an increase in CAT activity of about fourfold. This suggests that the region from residues 202 to 226 contains a negative regulatory region which inhibits multimer formation with the Bel1 protein. This result is partially consistent with our recent finding that Bel1 contains a negative regulatory region from residues 153 to 226 which down-regulates the transactivation domain of Bel1 (26) if multimerization is an absolute requirement of transcriptional activation. Further successive deletions from residue 201 to residues 197, 181, 174, 151, 124, and 83 did not show any significant effect on the multimerization of Bel1, indicating that the region of Bel1 from residues 1 to 82 still contains sufficient information to form a multimeric complex with Bel1. Deletion of Bel1 sequences from this point to residues 74, 55, 44, and 32 resulted in slightly progressive reductions in the efficiency of multimerization which may have been

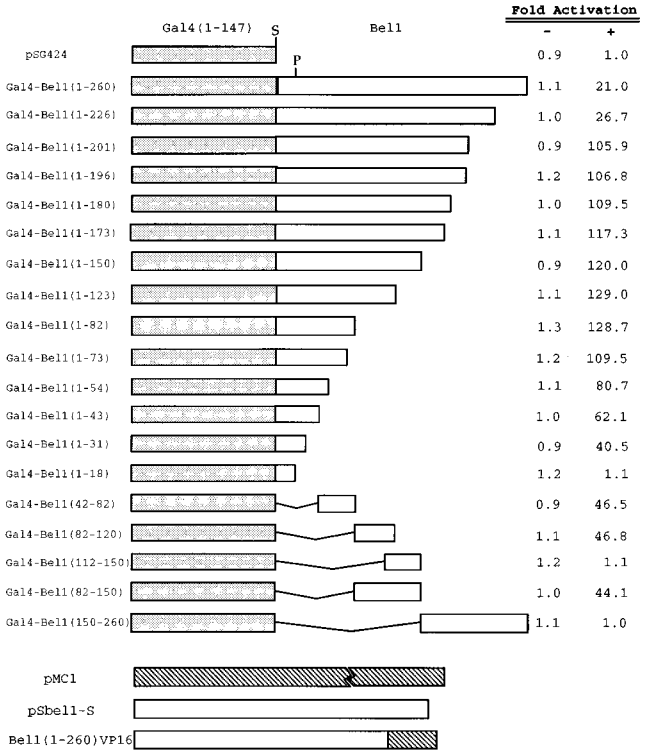

FIG. 4. Identification of regions affecting the multimerization of Bel1. A series of amino- and carboxy-terminal deletion mutant Bel1 proteins were generated by using the appropriate restriction enzymes and PCR and then were fused in frame downstream of the Gal4 DNA-binding domain of pSG424. The ability of each Gal4-Bel1 fusion protein to multimerize with the Bel1-VP16 fusion protein expressed by the Bel1(1-260)VP16 plasmid was determined by assaying transient cotransfection into BHK-21 cells. pRc/CMV was used as a negative control (−) instead of Bel1(1-260)VP16 (+). The Gal4 DNA-binding domain, a part of Bel1, and part or all of VP16 are represented by the stippled, open, and hatched boxes, respectively. S, *Sma*I; P, *Pst*I.

due to removal of the region required for optimal multimerization. In contrast, an additional deletion of residues 31 to 19 from Bel1 resulted in complete loss of in vivo multimerization activity. Therefore, the serial carboxy-terminal deletion analysis demonstrates that multimer formation with Bel1 was mediated by the region from residues 1 to 82 and down-regulated by the region from residues 202 to 226.

Further deletion analysis revealed that Gal4-Bel1(42-82) has CAT activity comparable to that of Gal4-Bel1(1-31). This suggests that the regions of Bel1 from residues 42 to 82 and 1 to 31 are required for effective multimerization. Surprisingly, Gal4-Bel1(82-120) and Gal4-Bel1(82-150), but not Gal4-Bel1 (112-150), produced an increase in CAT activity of about 45-fold, indicating that the region of Bel1 from residues 82 to 111 also can mediate the protein-protein interaction in vivo. In contrast, Gal4-Bel1(150-260) did not show any detectable CAT activity, suggesting that the region of Bel1 from residues 150 to 260 is not required for multimer formation. The expression levels of Gal4-Bel1 derivatives were confirmed by Western blot analysis with an anti-Gal4 rabbit antibody which reacts to the DNA-binding domain of Gal4. The Western blot analysis result did not show any significant difference in the protein levels of Gal4-Bel1 derivatives (data not shown). Taken together, the available data show that the Bel1 protein contains three regions required for effective multimerization of Bel1 (residues 1 to 31, 42 to 82, and 82 to 111) and one region (residues 202 to 226) required for down-regulation of Bel1 multimerization.

DISCUSSION

In this report, we have demonstrated that the nuclear targeting signal of Bel1 is bipartite, comprising two interdependent clusters of basic amino acids separated by an intervening spacer. In addition, our results showed that the second basic cluster is sufficient to target the Bel1–β-gal chimeric polypeptide to the nucleus when it is located at the amino terminus of *lacZ* but not when it is in an internal region. When the basic cluster is located at the amino terminus of the chimeric protein, it may be free enough to interact with factors involved in nuclear transport. Presumably, an amino acid sequence as basic as the NLS would normally tend to be exposed at the hydrophilic surface of a protein. The second basic cluster thus appears to be sufficient to direct Bel1–β-gal chimeric proteins to the nucleus. These results indicate that the activity of the Bel1 NLS is dependent on the protein context within which it is present. It is likely that both basic clusters located in an internal region of a chimeric protein are required to act directly in some steps involved in nuclear transport, or one of the two basic regions may be indirectly involved in providing the proper structural conformation for presenting the "true" signal. It was reported that the capacity of the NLS to direct cytoplasmic proteins to the nucleus depends on the chimeric protein context (33, 37). Therefore, the conclusion that the Bel1 NLS is bipartite appears to be more justified since it is located in the internal region in the natural context of Bel1.

A sequence comparison identified similar compositions of a bipartite nucleoplasmin-like motif between Bel1 and a number of nuclear proteins (Table 1). The typical bipartite NLS consists of 2 basic residues followed by a spacer of 10 other residues and then a second cluster in which 3 of 5 residues are basic. Insertional mutagenesis of Bel1 showed that the spacer length can be increased to 14, 18, or 22 amino acids without significantly affecting nuclear targeting, indicating that there is no strict requirement for spacer length. However, we observed that the activity of the Bel1 NLS was abolished when a specific amino acid sequence, PVDR, was inserted into the spacer region. Similar results were reported in an earlier study in which the spacer length of the nucleoplasmin NLS was altered without disrupting the nuclear targeting activity, yet targeting was abolished by insertion of only one copy of the sequence QPWL (36). It is likely that inappropriate folding of the nuclear targeting sequence by insertion of a specific foreign sequence into the spacer region impairs the ability of the protein to interact with relevant cellular factors required for nuclear transport. Thus, efficient nuclear targeting presumably requires correct positioning of the two basic elements relative to each other.

We have used an in vivo assay of protein-protein interaction (10) to assess the potential of the Bel1 protein to form specific multimers in the nuclei of mammalian cells. The results presented in this report demonstrate that Bel1 can indeed form a multimeric complex, and the regions mediating the multimerization were mapped to three subdomains (residues 1 to 31, 42 to 82, and 82 to 111). Furthermore, the region from residues 202 to 226 negatively regulates multimer formation. It has been reported that several proteins, such as Myc, E12, and AP4, possess more than one dimerization domain and these multiple dimerization domains regulate dimer specificity (14, 21, 30). The presence of multiple domains in Bel1 for oligomerization could be expected to give specificity of multimer formation. In fact, the region of Bel1 from residues 1 to 88 is shared with Bet, a cytoplasmic protein expressed at a very high level in HFV-infected cells; Beo; and Bel3 (31). These facts indicate that the region of Bel1 from residues 82 to 111 may play a specific role

in multimeric complex formation with the Bel1 protein. This is supported by the fact that Gal4-Bel1(82-150), which contains both the positive regulatory domain and the multimerization subdomain of Bel1, effectively inhibits the native function of Bel1 as determined by an in vivo competition assay, whereas Gal4-Bel1(1-82) cannot (26). While the amino-terminal 55 residues were shown to be dispensable for activation (39a), the 82-to-120 region is included in the essential effector region (26, 39b). Although the data reported here demonstrate that HFV Bel1 can form specific multimers in the nuclei of mammalian cells, we do not know whether in vivo multimerization involves a bridging cellular factor. Also, it remains to be determined whether oligomerization of Bel1 is important for the transactivation function.

ACKNOWLEDGMENTS

We are grateful to Y. S. Lee for providing the microscope and technical advice.

This work was supported by grant BM-9411-23 from the Center for Biofunctional Molecules.

REFERENCES

1. **Ace, C. I., M. A. Dalrymple, F. H. Ramsay, V. G. Preston, and C. M. Preston.** 1988. Mutational analysis of the herpes simplex virus type 1 trans-inducing factor Vmw65. J. Gen. Virol. **69:**2595–2605.
2. **Bogerd, H. P., R. A. Fridell, W. S. Blair, and B. R. Cullen.** 1993. Genetic evidence that the Tat proteins of human immunodeficiency virus types 1 and 2 can multimerize in the eukaryotic cell nucleus. J. Virol. **67:**5030–5034.

2a. **Bogerd, H. P., and W. C. Greene.** 1993. Dominant negative mutants of human T-cell leukemia virus type 1 Rex and human immunodeficiency virus type 1 Rev fail to multimerize in vivo. J. Virol. **67:**2496–2502.

3. **Cathy, Q. L., Y. Yun, J. P. Hoeffler, and J. F. Habener.** 1990. Cyclic-AMP-responsive transcriptional activation of CREB-327 involves interdependent phosphorylated subdomains. EMBO J. **9:**4455–4465.
4. **Chen, J., C. Panagiotidis, and S. Silverstein.** 1992. Multimerization of ICP0, a herpes simplex virus immediate-early protein. J. Virol. **66:**5598–5602.
5. **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.
6. **Dingwall, C., and R. A. Laskey.** 1991. Nuclear targeting sequences—a consensus? Trends Biochem. Sci. **16:**478–481.
7. **Dingwall, C., J. Robbins, and S. M. Dilworth.** 1989. Characterization of the nuclear location sequence of *Xenopus* nucleoplasmin. J. Cell Sci. **11**(Suppl.)**:** 243–248.
8. **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 SV 40 large T antigen. J. Cell Biol. **107:**841–849.
9. **Fields, S., and S. K. Jang.** 1990. Presence of a potent transcription activating sequence in the p53 protein. Science **249:**1046–1049.
10. **Fields, S., and O. K. Song.** 1989. A novel genetic system to detect protein-protein interaction. Nature (London) **340:**245–246.
11. **Flügel, R. M.** 1991. Spumaviruses: a group of complex retroviruses. J. Acquired Immune Defic. Syndr. **4:**739–750.
12. **Goldfarb, D. S., J. Gariepy, G. Schoolnick, and R. D. Kornberg.** 1986. Synthetic peptides as nuclear localization signals. Nature (London) **322:**641–644.
13. **He, F., J. D. Sun, E. D. Garrett, and B. R. Cullen.** 1993. Functional organization of the Bel1 transactivator of human foamy virus. J. Virol. **67:**1896–1904.
14. **Herkowitz, I.** 1987. Functional inactivation of genes by dominant negative mutations. Nature (London) **329:**219–222.
15. **Hu, Y.-F., B. Lusher, A. Admon, N. Mermod, and R. Tjian.** 1990. Transcription factor AP-4 dimerization domains that regulate dimer specificity. Genes Dev. **4:**1741–1752.
16. **Jones, N.** 1990. Transcriptional regulation by dimerization: two sides to an incestuous relationship. Cell **61:**9–11.

17a. **Kalderon, D., W. D. Richardson, A. T. Markham, and A. E. Smith.** 1984. Sequence requirements for nuclear location of simian virus 40 large T antigen. Nature (London) **311:**33–38.

17b. **Kalderon, D., W. D. Richardson, A. T. Markham, and A. E. Smith.** 1984. A short amino acid sequence able to specify nuclear location. Cell **39:**499–509.

18a. **Keller, A., E. D. Garrett, and B. R. Cullen.** 1992. The Bel1 protein of human foamy virus activates human immunodeficiency virus type 1 gene expression via a novel DNA target site. J. Virol. **6:**3946–3949.

18b. **Keller, A., K. M. Partin, M. Löchelt, H. Bannert, R. M. Flügel, and B. R. Cullen.** 1991. Characterization of the transcriptional *trans* activator of human foamy retrovirus. J. Virol. **65:**2589–2594.

19. **Kerppola, T. K., and T. Curran.** 1991. Transcription factor interactions; basics on zippers. Curr. Opin. Struct. Biol. **1:**71–79.
20. **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.
21. **Lamb, P., and S. L. McKnight.** 1991. Diversity and specificity in transcriptional regulation: the benefits of heterotypic dimerization. Trends Biochem. Sci. **16:**417–426.
22. **Landschulz, W. H., P. F. Johnson, and S. L. McKnight.** 1988. The leucine zipper: a hypothetical structure common to a new class of DNA binding proteins. Science **240:**1759–1764.
23. **Lanford, R. E., and J. S. Butel.** 1984. Construction and characterization of an SV 40 mutant defective in nuclear transport of T antigen. Cell **37:**801–813.
24. **Lanford, R. E., P. Kanda, and R. C. Kennedy.** 1986. Induction of nuclear transport with a synthetic peptide homologous to the SV 40 T antigen transport signal. Cell **46:**575–582.
25. **Lee, A. H., K. J. Lee, S. Kim, and Y. C. Sung.** 1992. Transactivation of human immunodeficiency virus type 1 long terminal repeat-directed gene expression by the human foamy virus *bel1* protein requires a specific DNA sequence. J. Virol. **66:**3236–3240.
26. **Lee, C. W., C. Jun, K. J. Lee, and Y. C. Sung.** 1994. The Bel1 protein of human foamy virus contains one positive and two negative control regions which regulate a distinct activation domain of 30 amino acids. J. Virol. **68:**2708–2719.
27. **Lee, K. J., A. H. Lee, and Y. C. Sung.** 1993. Multiple positive and negative *cis*-acting elements that mediate transactivation by Bel1 in the long terminal repeat of human foamy virus. J. Virol. **67:**2317–2326.
28. **Löchelt, M., H. Zentgraf, and R. M. Flügel.** 1991. Construction of an infectious DNA clone of the full-length human spumaretrovirus genome and mutagenesis of the *bel1* gene. Virology **184:**43–54.
29. **McIntyre, M. C., M. G. Frattini, S. R. Grossman, and L. A. Laimins.** 1993. Human papillomavirus type 18 E7 protein requires intact Cys-X-X-Cys motifs for zinc binding, dimerization, and transformation but not for Rb binding. J. Virol. **67:**3142–3150.
30. **Mordacq, J. C., and D. I. H. Linzer.** 1989. Co-localization of elements required for phorbol ester stimulation and glucocorticoid repression of proliferin gene expression. Genes Dev. **3:**760–769.
31. **Muranyi, W., and R. M. Flügel.** 1991. Analysis of splicing patterns of human spumaretrovirus by polymerase chain reaction reveals complex RNA structures. J. Virol. **65:**727–735.
32. **Nath, S. T., and D. P. Nayak.** 1990. Function of two discrete regions is required for nuclear localization of polymerase basic protein 1 of A/WSN/33 influenza virus (H1N1). Mol. Cell. Biol. **10:**4139–4145.
33. **Nelson, M., and P. Silver.** 1989. Context affects nuclear protein localization in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **9:**384–389.
34. **Queen, C., and D. Baltimore.** 1983. Immunoglobulin gene transcription is activated by downstream sequence elements. Cell **33:**741–748.
35. **Rethwilm, A., K. Mori, B. Maurer, and V. ter Meulen.** 1990. Transacting transcriptional activation of human spumaretrovirus LTR in infected cells. Virology **175:**568–571.
36. **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.
37. **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.
38. **Smith, A. E., D. Kalderon, B. L. Roberts, W. D. Colledge, M. Edge, P. Gillett, A. F. Markham, E. Paucha, and W. D. Richardson.** 1985. The nuclear localization signal. Proc. Trans. R. Soc. Lond. B **226:**43–58.

39a. **Venkatesh, L. K., and G. Chinnadurai.** 1993. The carboxy-terminal transcription enhancement region of the human spumaretrovirus transactivator contains discrete determinants of the activator function. J. Virol. **67:**3868–3876.

39b. **Venkatesh, L. K., C. Yang, P. A. Theodorakis, and G. Chinnandurai.** 1993. Functional dissection of the human spumaretrovirus transactivator identifies distinct classes of dominant-negative mutants. J. Virol. **67:**161–169.

40. **Williams, T., and R. Tjian.** 1991. Characterization of a dimerization motif in AP-2 and its function in heterologous DNA-binding proteins. Science **251:** 1067–1070.
41. **Xia, Y.-P., and M. M. C. Lai.** 1992. Oligomerization of hepatitis delta antigen is required for both the *trans*-activating and *trans*-dominant inhibitory activities of the delta antigen. J. Virol. **66:**6641–6648.
42. **Zapp, M. L., T. J. Hope, T. G. Parslow, and M. R. Green.** 1991. Oligomerization and RNA binding domains of the type 1 human immunodeficiency virus Rev protein: a dual function for an arginine-rich binding motif. Proc. Natl. Acad. Sci. USA **88:**7734–7738.