JOURNAL OF VIROLOGY, Jan. 1988, p. 349–353
0022-538X/88/010349-05$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 62, No. 1

Avian Sarcoma-Leukosis Virus *pol*-endo Proteins Expressed Independently in Mammalian Cells Accumulate in the Nucleus but Can Be Directed to Other Cellular Compartments

CLAIRE MORRIS-VASIOS,[1] JAREMA P. KOCHAN,[2] AND ANNA MARIE SKALKA[1]†*

Department of Molecular Oncology, Roche Institute of Molecular Biology,[1] and Department of Molecular Genetics, Hoffmann-La Roche Inc.,[2] Roche Research Center, Nutley, New Jersey 07110

Received 29 June 1987/Accepted 14 September 1987

Eucaryotic expression vectors have been used to study transient expression of the avian sarcoma-leukosis retrovirus *pol*-endo protein in COS cells. The constructs encode proteins with N termini identical to that of authentic viral pp32 endonuclease with the exception of a single *met* residue encoded by the initiator AUG. The C termini correspond to unprocessed viral *pol* protein, authentic processed pp32, or a derivative which includes eight amino acids from the unprocessed portion. All three proteins localize to the nucleus. However, when the *pol*-endo domain is fused to a secretory signal peptide, the protein is found in culture medium and appears also to localize in the Golgi bodies and the cell membrane. These and derivative vectors will make it possible to assess the consequence of retroviral *pol* gene expression in eucaryotic cells.

Retroviral genomes encode a highly conserved domain located in the C-terminal portion of the *pol* gene product (4, 11) whose function is required for the integration of proviral DNA into a host cell chromosome (9, 10, 16, 17). Biochemical studies with avian sarcoma-leukosis virus (ASLV) virion-derived proteins have identified a DNA endonuclease activity encoded in this domain, which we refer to as the *pol*-endo protein (for a review, see reference 18). Efforts to evaluate the role of this activity in integrative recombination are hampered to some extent by the fact that (i) the *pol*-endo domain is synthesized as part of a polyprotein precursor that is only processed during or shortly after virus budding and (ii) the mature protein is sequestered in the core of infecting virions. Thus, during a normal infection the activity does not exist free in a cell, and complementation and other genetic and biochemical studies are difficult if not impossible to perform. To circumvent this problem, we have constructed expression vectors that produce proteins which correspond to the mature viral *pol* products, free of other virion components. In this report we describe results of transient expression of the *pol*-endo in monkey kidney (COS) cells. This system was chosen for initial studies because the presence of the simian virus 40 origin allows efficient replication in the T-antigen-expressing COS cells. The resultant high DNA copy number leads to high expression of genes cloned in the vector. The transient assay allows us to avoid potential problems which could result from constitutive expression of the cloned genes in stably transfected cell lines.

The *pol*-endo domain used in these studies was obtained from the ASLV Schmidt-Ruppin B strain. In ASLV, the *pol*-endo domain is contained in a number of differentially processed proteins. It is found at the C terminus of the β chain of the viral reverse transcriptase, in a phosphorylated 32-kilodalton protein, pp32, and in a minor virion component, p36 (1). The latter protein is identical in sequence to pp32, with the exception of an additional 37 amino acids at its C terminus that are encoded in the viral genome and are most likely removed during final processing of both β and pp32.

Figure 1A shows the structures of the pp32-related clones which were made for these studies. We took advantage of a previously assembled *Escherichia coli* expression clone, pFA3-RT36 (1), in which an ATG initiation codon had been introduced 5′ to the authentic N-terminal codon of pp32. This protein terminates at the functional TAA codon present in the viral sequence and includes the C-terminal 37-amino-acid (4-kilodalton) segment described above. The *E. coli* eucaryotic shuttle vector has been described previously. The derivative used (pBC12BI) contained the rat preproinsulin II gene (5). Cleavage with *Cla*I and *Sma*I removes most of the coding region; a 5′ intron and 3′ polyadenylation signal remain with the vector. To construct pBC12-p36, an *Aha*II-*Xho*I fragment containing the ASLV *pol*-endo domain was isolated from pFA3-RT36 (1) and inserted into pBC12BI. The *Xho*I end was filled in with DNA polymerase Klenow fragment before digestion with *Aha*II so that it would ligate to the blunt *Sma*I end in the vector (14). The ATG of *pol*-endo was repositioned next to the nontranslated preproinsulin sequences by oligonucleotide-directed deletion mutagenesis (8). A 30-base oligonucleotide complementary to a sequence of 15 nucleotides in pBC12BI (immediately upstream of the ATG codon) and 15 nucleotides in the endonuclease gene beginning 89 nucleotides downstream of the ATG codon was used to create the deletion. This is illustrated underneath the vector in Fig. 1A. Clones with the desired deletions were selected from transformed *E. coli* by colony hybridization with the same oligonucleotide end labeled with ^{32}P. A similar strategy was used for cloning and alignment in pBC12-32 and pBC12-32+. The endo domain in p32+ was from plasmid pLD6T (R. A. Katz and A. M. Skalka, J. Virol., in press) and is identical to that in pBC12-p36, with the exception of three single-nucleotide substitutions generated by site-directed mutagenesis (15), one of which introduces a new termination site after the eighth codon downstream from the mature end of viral pp32. The endo domain of pBC12-p32 was from pRC23-p32 (R. Terry, D. Soltis, M. Katzman, J. Leis, and A. M. Skalka,

* Corresponding author.
† Present address: Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Ave., Philadelphia, PA 19111.

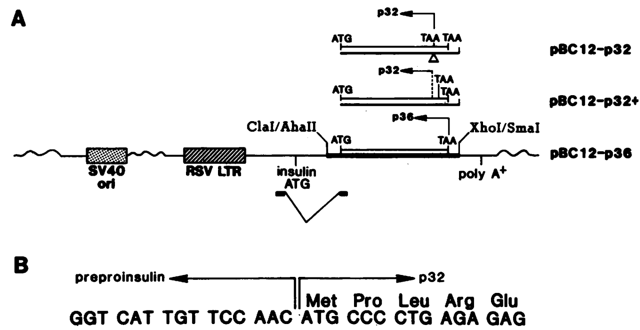

FIG. 1. Construction of plasmid expression clones. (A) Map showing the important features of the eucaryotic ASLV p36 expression vector pBC12-p36. The maps above pBC12-p36 show the relevant differences in the additional expression vectors pBC12-p32+ and pBC12-p32. (B) Nucleotide sequence aligned by oligonucleotide-directed deletion mutagenesis as illustrated below the diagram in panel A.

manuscript in preparation), a derivative of pFA3-RT36 which contains an insertion of a 20-base-pair segment at a *Ban*II site near the end of the pp32 coding region. The insert includes codons corresponding to the mature terminus of pp32 plus a functional termination signal. The newly introduced termination signals in pBC12-p32 and pBC12-p32$^+$ were shown to be functional in *E. coli* and animal cells (Katz and Skalka, in press; Terry et al., in preparation). All three clones should produce proteins that are identical to the authentic viral protein in their pp32 domains. Comparison of the three clones allow us to evaluate the significance of the 4-kilodalton C-terminal domain of *pol*.

The expression clones were introduced into COS cells by transfection (5, 6). After 2 days, cell extracts were analyzed by immunoprecipitation with goat polyclonal avian myeloblastosis virus reverse transcriptase antiserum and subsequent immunoblotting with affinity-purified peptide-specific polyclonal rabbit antibodies directed against a peptide whose sequence corresponds to the C-terminal end of pp32 (1). Significant levels of *pol*-endo proteins were observed in cells transfected with pBC12-p36 and pBC12-p32$^+$, and the proteins migrated as expected for their length differences (Fig. 2A). Very low levels of p32 were expressed in pBC12-p32-transfected cells; a band corresponding to p32 was visible only after prolonged exposure of the blot (Fig. 2B). Separate analyses to be reported elsewhere (R. A. Katz, M. Kotler, and A. M. Skalka, manuscript in preparation) indicate that the low production of protein from pBC12-p32 reflects a difference in p32 mRNA processing caused by the insertion used to introduce the functional termination codon. With the long exposure used in Fig. 2B, some antiserum-reactive protein which migrated as p32 was also detected in the pBC12-p36-transfected cells. It seems probable that this represents cleavage by a cellular protease which recognizes sites near or at the viral processing site in p36. A similar processing was previously observed in *E. coli* (1). As expected, the pp32-specific antiserum did not recognize any proteins in mock-transfected COS cells (Fig. 2A, lane C).

The *pol*-endo proteins were localized by indirect immunofluorescence in COS cells transfected with each of the three expression plasmids. Figure 3 shows the results of analyses made 2 days posttransfection. In all cases, nuclear staining was observed, and the number of positive nuclei (ca. 10%) corresponded to the expected transfection efficiency. Thus, the nonreactive cells in each field serve as an internal control. COS cells that were mock transfected did not react with the antibody. In cells synthesizing p36 and p32$^+$, the proteins were easily visualized and appeared to accumulate around the nucleoli but to be excluded from them. This pattern is similar to that observed for condensed perinucleolar chromatin (2). The significance of this observation remains to be explored. As expected, the amount of p32 visualized in the nuclei of cells transfected with pBC12-p32 was much lower than in those transfected with pBC12-p36 or pBC12-p32$^+$. Since all the *pol*-endo proteins exhibited similar cellular localization, we conclude that the 4-kilodalton C-terminal domain of p36 does not affect this property.

Localization of the p36 and p32+ *pol*-endo protein products was confirmed by fractionation of disrupted cells. Transfected cells were lysed 2 days posttransfection and separated into nuclear and cytoplasmic components. Proteins were resolved by gel electrophoresis, and endo-related products were detected by immunoblotting (Fig. 4). Both p36 and p32+ were readily detected in whole cell extracts (Fig. 4, lane W). When the *pol*-endo protein was separated into nuclear and cytoplasmic fractions, it was apparent that most of it was associated with the nuclei. Approximately half of the nuclear-associated protein was tightly bound, as evidenced by the fact that it was not readily released by salt extraction (Fig. 4; compare lanes N1 and N2).

In a normal infection, the endonuclease domain is synthesized in the cytoplasm as part of a *gag-pol* precursor and accumulates at budding sites underneath the cell membrane. This suggests that the nuclear localization signals present in the mature endonuclease protein are masked in the precursor or superceded by other localization sequences. As a control for the nuclear localization and to test whether the responsible sequences could be superceded by another sig-

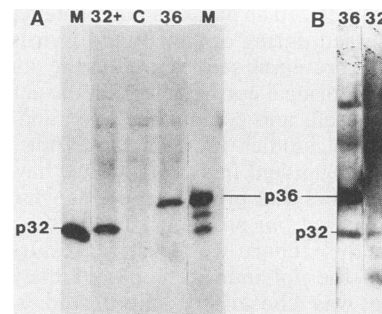

FIG. 2. Detection of *pol*-related proteins in extracts from transfected cells. Transfection and preparation of cell lysates for immunoprecipitation were performed as described by Cullen et al. (5, 6). Samples of each cleared lysate were treated with 5 μl of goat polyclonal antiserum directed against purified avian myeloblastosis virus reverse transcriptase (765-168; obtained from the National Cancer Institute Office of Program Resources and Logistics) for 45 min on ice. Immune complexes were formed with addition of *Staphylococcus aureus* protein A (5). After washing and dissociation of the complex, proteins were run on 9% sodium dodecyl sulfate–polyacrylamide gels (12) at 45 mA for approximately 1.5 h. The separated proteins were then transferred electrophoretically to nitrocellulose filters (0.2-μm pore size; Schleicher & Schuell Co.) at 0.8 mA/cm^2 for 3 to 4 h with a Sartorius semidry blotting apparatus. Filters were reacted with an affinity purified rabbit antiserum directed against a peptide extending from amino acids 840 to 854 in the pp32 domain (1). Lanes: M, bacterial p32 protein (left lane) or bacterial p36 protein (right lane) (Terry et al., in preparation) used as markers; C, untransfected control cell lysates; 36, 32+, and 32, extracts from cells transfected with pBC12-p36, pBC12-p32+, and p32, respectively.

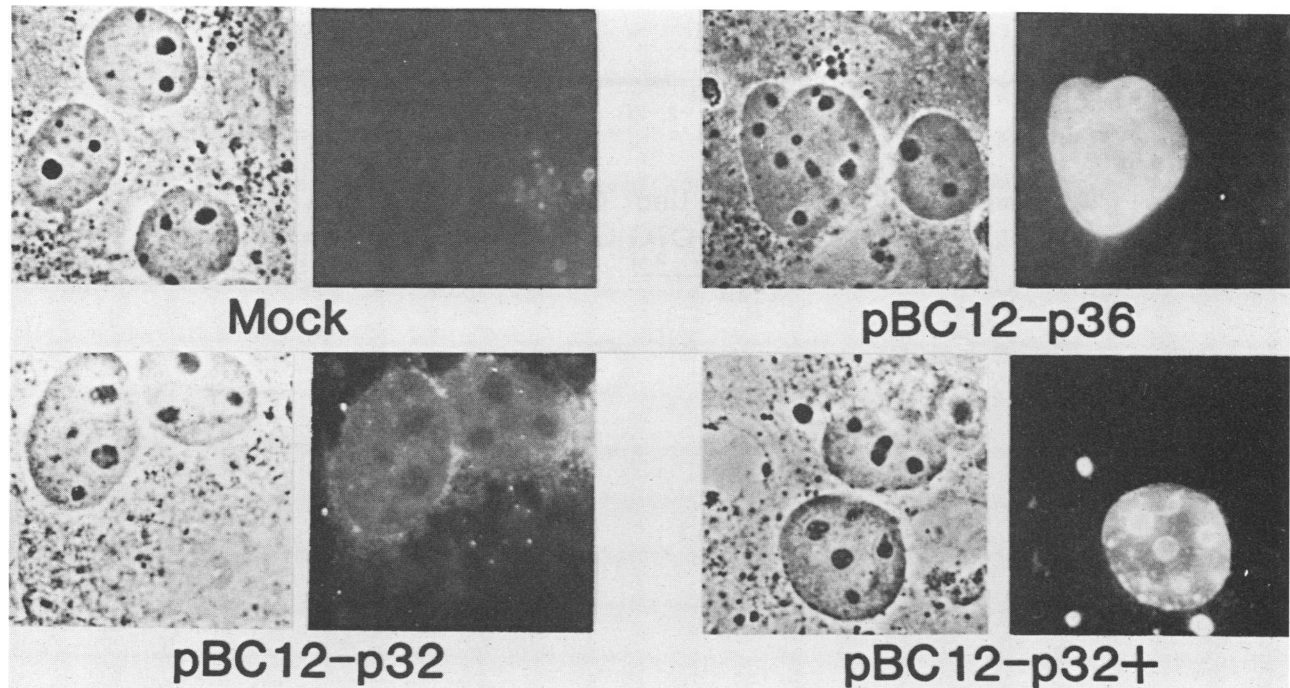

FIG. 3. Indirect immunofluorescence. Cells were grown and transfected on glass cover slips in 35-mm tissue culture dishes and fixed in situ for indirect immunofluorescence 48 to 60 h posttransfection as described by Curran et al. (5). After permeabilization of cells, cover slips were exposed to a 1:100 dilution of rabbit anti-pp32 peptide antibody (described in the legend to Fig. 2), washed, and exposed to a 1:100 dilution of a rhodamine conjugated goat anti-rabbit immunoglobulin G (Cappel Laboratories) for 30 min at room temperature. The cover slips were washed, mounted in 50% glycerol containing 0.25 M *n*-propyl gallate, observed on a Leitz diavert fluorescence microscope (40× oil objective), and photographed with Kodak Tri-X film.

nal, we constructed a fourth expression clone, pBC12-SP36, in which the p36 coding sequences were fused to the secretory signal sequence of an interleukin-2 receptor gene (Fig. 5A). If, as expected, processing occurs by the signal peptidase of the cell, our newly directed mature protein should contain the first two amino acids of the processed interleukin-2 receptor, fused to a glutamine residue, followed by the seventh amino acid of mature viral pp32 and the remainder of the p36 protein. COS cells transfected with this plasmid were metabolically labeled with [^{35}S]methionine for 1 h. The cells were then collected and lysed. *pol*-endo proteins in both the cell lysate and the culture supernatant were concentrated by immunoprecipitation and then analyzed on a polyacrylamide gel (Fig. 5B); a band corresponding to p36 is seen both in cells and in supernatants of the culture transfected with pBC12-SP36 (lanes T) but not in the mock-transfected culture (lanes M). We estimate from the relative exposures required for visualization of the bands that approximately one-fourth of the newly synthesized p36 protein was secreted into the medium during the labeling period. We note that some of the protein in both cell and supernatant fractions migrated more slowly than p36. Since there are three potential N-glycosylation sites in the p36-coding region, it seems possible that this represents a fraction that has been modified during passage through the endoplasmic reticulum. Indirect immunofluorescence analy-

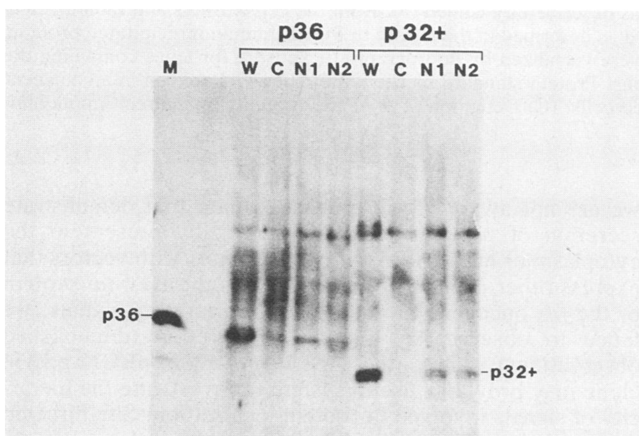

FIG. 4. Fractionation of cell extracts. Transfections were as described in the legend to Fig. 2. Separation of nuclei and cytoplasm was essentially as described by Challberg and Kelly (3), except that 2.0 μg of leupeptin per ml, 5.0 μg of aproteinin per ml, and 0.1 mM phenylmethylsulfonyl fluoride were added to the hypotonic buffer. Cytoplasmic supernatant fractions were adjusted to RIPA buffer conditions. Nuclei were suspended in 10 ml of 0.5% Nonidet P-40–0.4 M NaCl and incubated for 30 min on ice to release nuclear proteins (N1). After centrifugation at 80,000 × *g* for 30 min, the pellets were suspended in 100 μl of a buffer consisting of 0.5% sodium dodecyl sulfate, 50 mM Tris (pH 8.0), and 1 mM dithiothreitol, heated for 5 min at 100°C to release the more tightly associated nuclear proteins (N2), and spun at 80,000 × *g* for 30 min. The supernatants from each extraction were adjusted to RIPA conditions (5) (without sodium dodecyl sulfate) and clarified by centrifugation at 30,000 rpm for 30 min in a 70.1 Ti rotor in a Beckman L8-80M ultracentrifuge. All samples were immunoprecipitated and prepared for blotting as described in the legend to Figure 2. Lanes: W, whole cell extract; C, cytoplasmic material; N1, nuclear proteins eluted with 0.4 M NaCl; N2, nuclear material eluted by boiling samples.

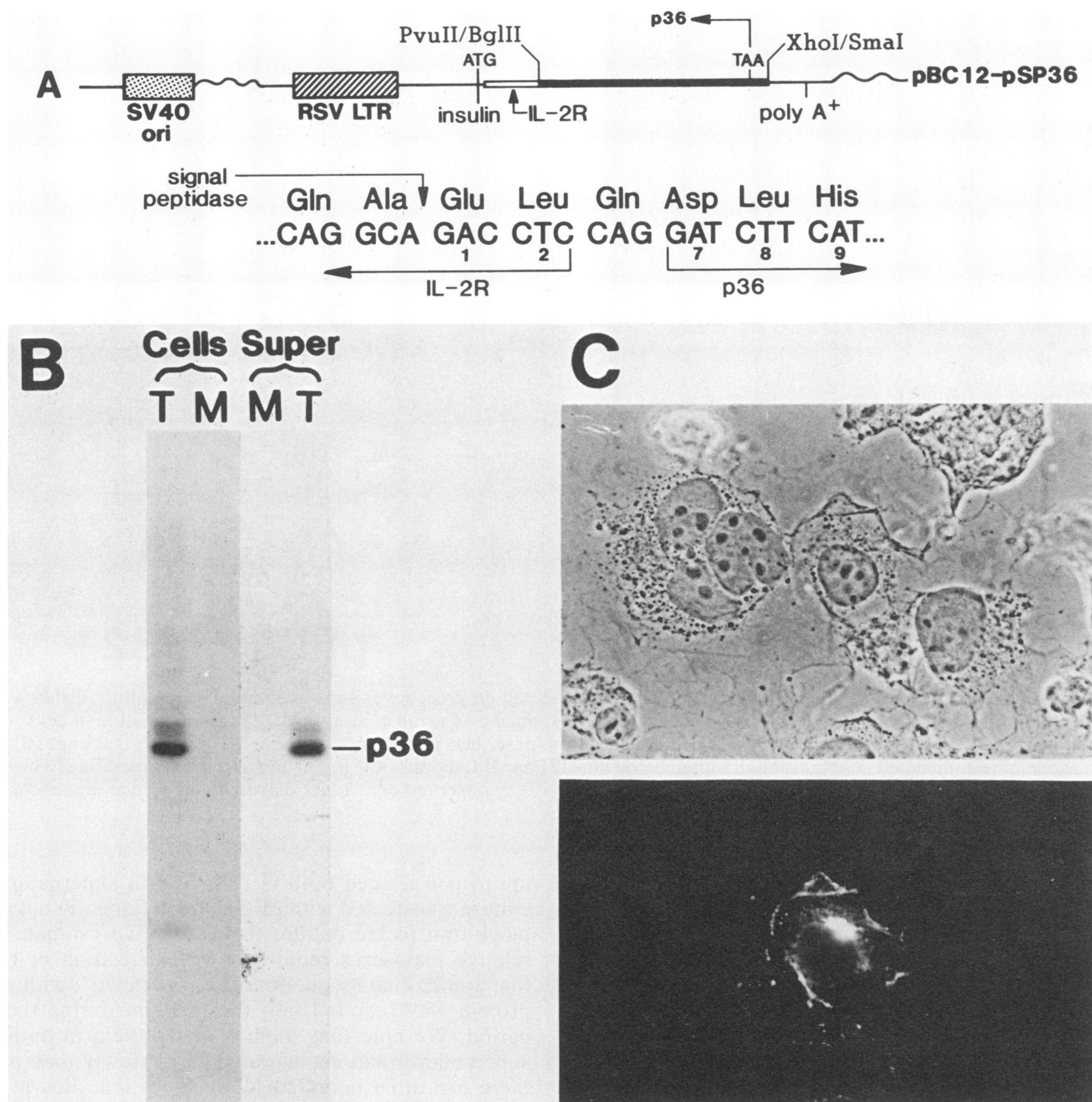

FIG. 5. Expression of a p36 *pol*-endo protein containing an N-terminal secretory signal. (A) Structure of the expression clone. The basic vector backbone is the same as in Fig. 1 (5). A linker adapter molecule was inserted into the derivative so that a unique *Pvu*II site would be produced following the second Gln codon shown below the diagram. The plasmid pFA3-RT36 was digested with *Bgl*II and *Xho*I, and the ends were filled in with DNA polymerase (Klenow fragment) (14). The blunt ended *Bgl*II-*Xho*I fragment was ligated to a *Pvu*II-digested vector, and clones containing the correctly oriented insert were identified by restriction analysis. The sequence below the line shows the predicted N terminus of the processed p36 protein expressed from this vector. (B) Immunoprecipitation of [^{35}S]methionine-labeled *pol*-endo protein. Conditions for transfection, radiolabeling, and extract preparation were as described by Cullen (5). Both the cell extracts and radiolabeling medium were treated with a 1:200 dilution of goat antiserum and processed as described in the legend to Fig. 2. Immunoprecipitated proteins were fractioned on a 10% polyacrylamide gel, and the labeled proteins were visualized by fluorography. Exposure for lanes containing the supernatant fractions was 4 times longer than that for the cellular fractions. Protein standard markers (not shown) were run in an adjacent lane. Lanes: T, cells transfected with pBC12-pSP36; M, mock-transfected cells. (C) Detection of *pol*-endo proteins by indirect immunofluorescence. Conditions were as described in the legend to Fig. 3.

ses (Fig. 5C) demonstrate that the cell-associated p36 protein is found primarily at the cell membrane and in a perinuclear region presumed to be the Golgi apparatus.

Our results show that the mature ASLV *pol*-endo proteins can be expressed independently in mammalian cells and that, as such, they are localized to the nucleus. However, when appropriate signal sequences are added, the protein can be directed to different cellular compartments and even secreted. The ability to redirect proteins with a signal peptide was first reported by Lingappa et al. (13). However, we are not aware of any published data that demonstrate secretion of proteins which are normally present in the cytoplasm or nucleus. In our own attempts with vectors that express other nuclear proteins (such as the HIV *tat* protein or the *fos* oncogene protein) fused to secretory signals, we failed to observe any significant secretion (unpublished observations). Thus, the ASLV *pol*-endo pBC12-pSB36 clone may provide a useful system to investigate the hierarchy of signals involved in protein localization. Site-directed mutagenesis studies with the clones without the secretory

signal should allow us to identify the sequences required for the nuclear localization and to determine whether they are distinct from the yet-to-be-identified DNA-binding domains. It is possible that the mature, processed *pol*-endo protein may play a role in the translocation of retroviral DNA integration intermediates into the nucleus during a normal infection. The data included in this report show that the 4-kilodalton C-terminal domain encoded in *pol* precursors does not affect the localization of the *pol*-endo proteins.

We have not determined whether the proteins produced in the transfected COS cells have enzymatic activity. However, this seems most likely, at least for those without the secretory signal, since similar constructs expressed in bacteria possess Mn^{2+}-dependent endonuclease activity that is similar to that of authentic viral pp32 (Terry et al., in preparation). The nuclear localization suggests that these proteins will be available in the correct cellular compartment to evaluate their role in integrative recombination.

We thank Bryan Cullen for the gift of the expression vectors, Laura Pettine for excellent technical assistance with the signal peptide *pol*-endo protein experiments, and Tom Curran and Lidia Sambucetti for help and advice with indirect immunofluorescence and cell fractionation procedures.

LITERATURE CITED

1. **Alexander, F., J. Leis, D. A. Soltis, R. M. Crowl, W. Danho, M. S. Poonian, Y.-C. E. Pan, and A. M. Skalka.** 1987. Proteolytic processing of avian sarcoma and leukosis viruses *pol-endo* recombinant proteins reveals another *pol* gene domain. J. Virol. **61**:534–542.
2. **Busch, H., and K. Smetana.** 1970. The nucleolus. Academic Press, Inc., New York.
3. **Challberg, M. D., and T. J. Kelly.** 1979. Adenovirus DNA replication *in vitro*. Proc. Natl. Acad. Sci. USA **76**:655–659.
4. **Chiu, I.-M., R. Callahan, S. R. Tronick, J. Schlom, and S. A. Aaronson.** 1984. Major *pol* gene progenitors in the evolution of oncoviruses. Science **223**:364–370.
5. **Cullen, B. R.** 1987. Use of eukaryotic expression technology in the functional analysis of cloned genes. Methods Enzymol. **152**:684–764.
6. **Cullen, B. R., P. T. Lomedico, and G. Ju.** 1984. Transcriptional interference in avian-retroviruses—implications for the promoter insertion model of leukaemogenesis. Nature (London) **307**:241–245.
7. **Curran, T., A. D. Miller, L. Zokas, and I. M. Verma.** 1984. Viral and cellular *fos* protein: a comparative analysis. Cell **36**: 259–268.
8. **DeChiara, T. M., F. Erlitz, and S. J. Tarnowski.** 1986. Procedures for *in vitro* DNA mutagenesis of human leukocyte interferon sequences. Methods Enzymol. **119**:403–415.
9. **Donehower, L. A., and H. E. Varmus.** 1984. A mutant murine leukemia virus with a single missense codon in *pol* is defective in a function affecting integration. Proc. Natl. Acad. Sci. USA **81**:6461–6465.
10. **Hippenmeyer, P. J., and D. P. Grandgenett.** 1984. Requirement of the avian retrovirus pp32 DNA binding protein domain for replication. Virology **137**:358–370.
11. **Johnson, M. S., M. A. McClure, D. F. Feng, J. Gray, and R. F. Doolittle.** 1986. Computer analysis of retroviral *pol* genes: assignment of enzymatic functions to specific sequences and homologies with nonviral enzymes. Proc. Natl. Acad. Sci. USA **83**:7648–7652.
12. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
13. **Lingappa, V. R., J. Chaidez, C. S. Yost, and J. Hedgpeth.** 1984. Determinants for protein localization: β-lactamase signal sequence directs globin across microsomal membranes. Proc. Natl. Acad. Sci. USA **81**:456–460.
14. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
15. **Morinaga, Y., T. Franceschini, S. Inouye, and M. Inouye.** 1984. Improvement of oligonucleotide directed site specific mutagenesis using double stranded plasmid DNA. Biotechnology **2**:636–639.
16. **Panganiban, A. T., and H. M. Temin.** 1984. The retrovirus *pol* gene encodes a product required for DNA integration: identification of a retrovirus *int* locus. Proc. Natl. Acad. Sci. USA **81**:7885–7889.
17. **Schwartzberg, P. J., J. Colicelli, M. L. Gordon, and S. P. Goff.** 1984. Mutations on the *gag* gene of Moloney murine leukemia virus: effects on production of virions and reverse transcriptase. J. Virol. **49**:918–924.
18. **Skalka, A. M., and J. Leis.** 1984. Retroviral DNA integration. BioEssays **1**:206–210.