JOURNAL OF VIROLOGY, Nov. 1994, p. 7620–7627
0022-538X/94/$04.00+0

Vol. 68, No. 11

Amino Acid Sequence Analysis of the Proteolytic Cleavage Products of the Bovine Immunodeficiency Virus Gag Precursor Polypeptide

GREGORY J. TOBIN,[1] RAYMOND C. SOWDER II,[2] DANIELE FABRIS,[3] MARIE Y. HU,[1] JANE K. BATTLES,[1] CATHERINE FENSELAU,[3] LOUIS E. HENDERSON,[2] AND MATTHEW A. GONDA[1]*

Laboratory of Cell and Molecular Structure[1] and AIDS Vaccine Program,[2] Program Resources, Inc./DynCorp, National Cancer Institute-Frederick Cancer Research and Development Center, Frederick, Maryland 21702-1201, and Department of Chemistry and Biochemistry, University of Maryland, Baltimore County, Baltimore, Maryland 21228[3]

Received 15 February 1994/Accepted 25 July 1994

Bovine immunodeficiency virus Gag proteins were purified from virions, and their amino acid sequences and molecular masses were determined. The matrix, capsid, and nucleocapsid (MA, CA, and NC, respectively) and three smaller proteins (p2L, p3, and p2) were found to have molecular masses of 14.6, 24.6, and 7.3 and 2.5, 2.7, and 1.9 kDa, respectively. The order of these six proteins in the Gag precursor, Pr53^{gag}, is NH_2-MA-p2L-CA-p3-NC-p2-COOH. In contrast to other retroviral MA proteins, the bovine immunodeficiency virus MA retains its N-terminal methionine and is not modified by fatty acids. In addition, the bovine immunodeficiency virus NC migrates as a 13-kDa protein in denaturing gel electrophoresis; however, its molecular mass was determined to be 7.3 kDa.

Bovine immunodeficiency virus (BIV) is a nonacute pathogenic retrovirus and a member of the lentivirus genus. BIV shares biologic, genetic, and pathologic characteristics with lentiviruses that cause chronic inflammatory diseases and immunodeficiencies (5, 16, 43; reviewed in references 13 to 15 and 17). Similarities between BIV and the human and simian immunodeficiency viruses (HIV and SIV, respectively) extend to ultrastructure, genome organization, catalytic functions of *pol* gene products, and immunological cross-reactivity of the Gag proteins (1, 4, 11, 16, 26, 35). Sequencing of functional molecular clones of BIV has shown that its genome includes the obligate *gag*, *pol*, and *env* retroviral structural genes as well as six putative nonstructural accessory genes (*tat*, *rev*, *vif*, *vpw*, *vpy*, and *tmx*) (4, 11, 13, 29, 32–34).

The BIV *gag* gene encodes a 53-kDa precursor protein, Pr53^{gag} (1, 35), that participates in the formation of immature virions and assembles into virus-like particles when produced in the baculovirus-insect cell expression system (35). The processing of lentivirus Gag precursors by an aspartyl-type viral protease (PR) leads to the formation of five to six major proteolytic cleavage products (7, 19, 20, 23). Immunologic analyses of proteins from BIV virions with monospecific anti-BIV and -HIV type 1 (HIV-1) sera have resolved three major Gag cleavage products, the matrix, capsid, and nucleocapsid proteins (MA [p16^{MA}], CA [p26^{CA}], and NC [p13^{NC}], respectively) (1, 35). A definition of the N and C termini of the three known Gag proteins and the exact number of Gag proteins contained within the BIV Gag precursor, however, have not been determined in previous studies (1, 35).

In the present report, the mature Gag proteins from BIV virions obtained from molecularly cloned BIV R29-127 were purified by reversed-phase, high-pressure liquid chromatography (rp-HPLC) and identified by immunoblotting with monospecific sera. N-terminal sequencing, amino acid composition, and mass spectrometry (MS) were used to calculate the molecular weight and to determine N and C termini of each isolated BIV Gag protein. These data precisely identified the proteolytic cleavage sites in and thus the subunit organization of Pr53^{gag}. In addition to the requisite MA, CA, and NC proteins, three small intragenic Gag cleavage fragments were identified, as was the lack of a fatty acid modification at the N terminus of the MA. The molecular masses of the BIV MA, CA, and NC proteins were determined by MS and are in agreement with those deduced from amino acid analyses. A comparison of these data with those derived from the study of HIV-1 and other retroviruses is discussed.

Isolation, amino acid analyses, and MS of BIV Gag proteins. Virus was harvested from Cf2Th cultures transfected with the BIV R29-127 molecular clone, isolated by continuous-flow sucrose density gradient centrifugation, and further concentrated by ultracentrifugation (2, 42). Virions were disrupted and reduced by the addition of solid guanidine hydrochloride and 2-mercaptoethanol in the presence of 3.7 μM zinc acetate to protect the putative zinc-binding NC protein. The virus preparation was sonicated and equilibrated to pH 2 with trifluoroacetic acid. Materials were loaded onto μBondapack C_{18} rp-HPLC supports (25 by 100 mm) in 0.05% (vol/vol) trifluoroacetic acid at pH 2 and eluted with a gradient of increasing acetonitrile concentrations in 0.05% trifluoroacetic acid, as previously described (20). Eluted proteins and peptides were detected by determining the UV A_{206}, A_{280}, and A_{294}; collected; and lyophilized. The spectra at 206 nm are shown in Fig. 1A.

Peak fractions of the rp-HPLC separation were analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and visualized by Coomassie staining (Fig. 1B) and immunoblot analysis (Fig. 1C). The predominant species in fractions 173 through 219 were identified as p16^{MA} and p26^{CA} (data not shown) but were not sufficiently resolved for subsequent analyses. Therefore, fractions 173 to 186, 187 to 199, and 201 to 219 were combined into pools I, II, and III, respectively, for additional rp-HPLC purifications (Fig. 1A, insets). Rechromatography of pooled fractions I, II, and III was performed over 3 h using 38 to 52% acetonitrile gradients at flow rates of 5 ml/min (Fig. 1A, insets). N-terminal amino acid sequence and compositional analyses were performed, as

* Corresponding author. Phone: (301) 846-1290. Fax: (301) 846-5427. Electronic mail address: gonda@lcms-1.ncifcrf.gov.

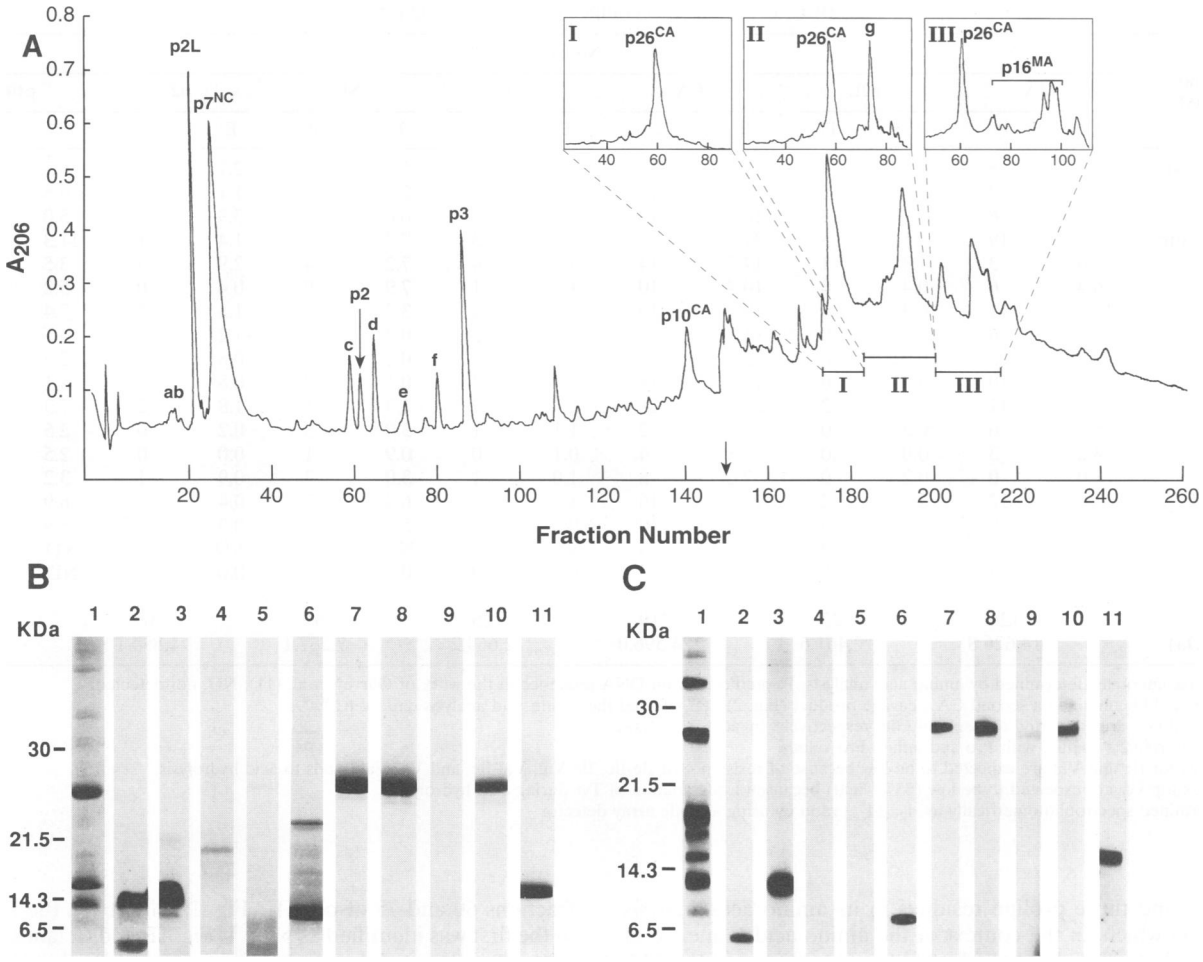

FIG. 1. Isolation and identification of BIV proteins. (A) Chromatogram of rp-HPLC elution profile of proteins from disrupted BIV virions detected by A_{206} versus fraction number. The large profile depicts the initial chromatogram. The inset panels show the absorbance profiles of rechromatographed pooled fractions I, II, and III. Peaks corresponding to p2L, $p7^{NC}$, p2, p3, $p10^{CA}$, $p26^{CA}$, and $p16^{MA}$ as well as cleavage fragments of p2L (a and b), p2 (c and d), and p3 (e and f) and a reagent peak (g) are shown. The arrow at fraction 150 indicates the time at which elution pump and chart recorder speeds were reduced from 10 to 2.6 ml/min and 2 to 0.5 mm/min, respectively, in the initial chromatograph. (B) Coomassie-stained electropherogram of the rp-HPLC peak fractions labeled in panel A obtained with a 6 to 18% linear gradient in SDS-PAGE. Lanes: 1, disrupted BIV virions used as markers for rp-HPLC-isolated proteins; 2 to 6, concentrated samples of p2L, $p7^{NC}$, p2, p3, and $p10^{CA}$, respectively, from the initial purification profile; 7, 8, and 10, $p26^{CA}$ peaks in insets I, II, and III, respectively; 9 and 11, peak g and $p16^{MA}$ from insets II and III, respectively. Molecular mass markers are indicated to the left. (C) Immunoblot analysis of rp-HPLC peak samples shown in panel A and analyzed in panel B. Samples were electrophoresed in 16% denaturing polyacrylamide gels, transferred to membranes, and cut into 0.5-cm strips. Membrane strips were reacted with a panel of eight previously described BIV-specific antisera (1) and arranged in the same lane order as shown in panel B. Results with BIV-infected cow F19 (lane 1), MA-CA intergenic peptide (lane 2), NC*term* peptide (lane 3), recombinant Pr53 (lanes 4, 5, and 9), CA peptide (lanes 6 to 8 and 10), and recombinant MA (lane 11) antisera are shown. Molecular mass markers are indicated to the left.

previously described (20), to define the N- and C-terminal boundaries of isolated Gag proteins. By comparing results of these analyses with the translation predicted from the DNA sequence (11), the probable C-terminal residues of each protein were determined. The amino acid compositions and N-terminal sequence analyses are summarized in Table 1 and Fig. 2, respectively. Two major anomalies observed during the course of this investigation, which could not be resolved by SDS-PAGE or amino acid analyses, required analysis by electrospray ionization MS (ESI-MS): (i) the HPLC elution profile of MA and CA suggested that these proteins contain multiple species, and (ii) NC migrated more slowly in SDS-PAGE than expected for a protein of the molecular weight that the amino acid sequence suggested.

(i) MA proteins. The majority of the BIV $p16^{MA}$ eluted as a series of peaks in fractions 201 to 219 in the initial acetonitrile gradient (data not shown). Although the second separation performed with pool III resolved several absorbance peaks (Fig. 1A, inset III), immunoblot analysis indicated the presence of only MA and CA (Fig. 1C). The predominant proteins contained in the later fractions 70 to 100 were MA species that appeared to migrate at the same molecular weight as determined by SDS-PAGE and immunoblot analysis (fraction 98 of pool III, Fig. 1B and C, lanes 11; fractions 70 to 97 of pool III, data not shown).

Forty-three residues of the MA protein were determined by N-terminal sequence analysis (Fig. 2). The MA-protein sequence was further characterized by the presence of six

TABLE 1. Amino acid composition of BIV Gag proteins

Amino acid(s)	No. of residues[a] in:													
	MA		p2L		CA		p3		NC		p2		$p10^{CA}$[b]	
	E	P	E	P	E	P	E	P	E	P	E	P	E	P
Asp + Asn[c]	8.4	8	3.8	4	19.2	19	0.4	0	5.8	6	2.1	2	6.9	7
Thr	7.3	7	5.0	5	14.9	15	1.1	1	2.2	2	1.1	1	4.5	5
Ser	6.5	6	3.1	3	10.4	10	1.2	1	6.0	6	3.9	4	5.0	4
Glu + Gln[c]	18.4	19	4.5	5	31.9	33	3.1	3	7.7	8	1.4	1	11.3	14
Pro[d]	3.6	3	1.3	1	14.5	14	4.2	4	7.2	4	2.9	3	3.5	8
Gly	6.4	6	0.4	0	10.4	10	1.2	1	7.9	9	0.4	0	5.2	5
Ala	10.8	11	0.3	0	18.4	19	3.9	4	3.2	3	1.3	1	7.4	10
Val[e]	6.1	6	0.1	0	11.3	12	2.0	2	0.2	0	1.2	1	4.5	4
Met	2.9	3	0.0	0	5.2	5	1.9	2	0.1	0	0.8	1	2.5	4
Ile[e]	8.9	10	0.0	0	12.2	14	1.9	2	0.1	0	0.8	1	4.1	5
Leu	11.2	11	2.1	2	21.1	21	2.1	2	1.3	1	1.8	2	7.3	8
Tyr[f]	5.7	6	0.2	0	2.6	2	1.1	1	3.0	3	0.2	0	2.5	2
Phe	3.2	3	0.0	0	4.8	4	0.1	0	0.9	1	0.0	0	2.5	2
His	1.0	0	0.2	0	7.6	8	1.0	1	3.0	3	0.9	1	3.2	4
Lys	11.0	12	2.1	2	15.9	16	1.2	1	6.8	7	0.4	0	6.9	7
Arg	8.4	9	0.3	0	9.9	10	0.2	0	5.8	6	0.3	0	3.9	3
Cys	ND	3	ND	0	ND	3	ND	0	ND	7	ND	0	ND	2
Trp[g]	2.8	3	0.0	0	4.1	4	0.0	0	0.0	0	0.0	0	ND	0
Total		126		22		219		25		66		18		94
Mass (Da)		14,626.8		2,461.6		24,596.0		2,664.2		7,287.1		1,894.1		10,381.9

[a] E, experimentally determined by amino acid analysis; P, predicted from DNA sequence in the work of Garvey et al. (11); ND, not detected.
[b] Fraction 140 contained a second CA cleavage product (Fig. 2) that affected the amino acid analysis data of $p10^{CA}$.
[c] Asn and Gln are converted to Asp and Glu, respectively, by acid hydrolysis.
[d] Cys by-products coelute with Pro and inflate Pro values.
[e] Values for Ile and Val are expected to be low because of resistance of Ile-Ile, Ile-Val, Val-Ile, and Val-Val bonds to acid hydrolysis.
[f] Values for Tyr are expected to be low (85% yield) because of degradation of Tyr during acid hydrolysis.
[g] Determined spectrophotometrically as A_{294}/A_{206} ratio by using a diode array detector.

tyrosine and three proline residues in its amino acid analysis (Table 1) which, in the context of the amino acid sequence of p2L (see below), helped to identify the C terminus of the MA protein. Thus, MA consists of 126 amino acid residues with a calculated molecular mass of 14,626.8 Da. The accessibility of the MA protein to Edman degradation and the retention of the initial methionine residue in the MA-protein sequence indicated that the N terminus was not blocked by fatty acid modification.

The profile of the chromatogram of pool III (Fig. 1A, inset III) suggested the presence of heterogenous MA species, which was confirmed by ESI-MS (Fig. 3A). The major MA peak that eluted in fractions 95 and 96 contained MA species of 14,628 and 14,708 Da (Fig. 3A and Table 2) in almost equal amounts. Proteins with masses of 14,814 and 14,628 Da were characterized in fractions 76 to 77 and 91 to 92, respectively. The experimental molecular masses observed for one major and one minor MA species (14,628 Da) were in good agreement with the mass of 14,626.8 Da predicted from the DNA sequence and the proteolytic cleavage sites determined here. Moreover, these data provided further evidence that the majority of the MA was not modified by fatty acid acylation.

(ii) CA proteins. In the initial rp-HPLC separation, the BIV $p26^{CA}$ eluted in several peaks from fractions 173 to 219 (Fig. 1A; data not shown). These fractions were pooled as indicated above and rechromatographed to improve the separation (Fig. 1A, inset I). As with MA, there appeared to be heterogenous species of CA, as multiple CA-containing peaks were identified in the rechromatographs. The major peak of pool I (fraction 60, inset I, Fig. 1A) contained a protein with a mobility of 26 kDa in SDS-PAGE, which was identified as $p26^{CA}$ in immunoblots (Fig. 1B and C, lanes 7). Of the two major peaks at fractions 60 and 75 of pool II (Fig. 1A, inset II), the material in the first was identified as $p26^{CA}$ (Fig. 1B and C, lanes 8). The second peak (g) did not contain any material that could be identified by Coomassie staining or immunoblotting (Fig. 1B and C, lanes 9). The amino acid composition and N-terminal sequence analysis of this fraction indicated that no polypeptide was present; therefore, it most likely contained a medium component such as phenol red. The first major peak of pool III (fraction 60) contained $p26^{CA}$, as indicated by SDS-PAGE and immunoblot analyses (Fig. 1B and C, lanes 10). $p26^{CA}$ was also detected in the later fractions (70 to 100) of pool III, most likely as a result of its poor solubility in the acidic liquid phase.

Amino acid sequence analysis of $p26^{CA}$ confirmed its 52 N-terminal residues, beginning at Pro-149 of $Pr53^{gag}$ (Fig. 2). The presence of four phenylalanines and 21 leucines, deduced by compositional analyses, defined the C terminus as Phe-367 of $Pr53^{gag}$ (Table 1). The $p26^{CA}$ contains 219 residues and has a calculated molecular mass of 24,596.0 Da. ESI-MS analysis demonstrated the presence of a CA species of 24,610 Da (Fig. 3B and Table 2) as the major chromatographic peak of each pool, labeled $p26^{CA}$ in the insets of Fig. 1A. Several less abundant forms of $p26^{CA}$ also were found in samples analyzed by ESI-MS from different pools, confirming the heterogeneity observed in the rp-HPLC profile. A species of 24,653 Da was present in pool I; its mass difference suggests the possibility of posttranslational acetylation of the major form. Species of 24,682 and 24,757 Da were observed in pools II and III.

A small percentage of the CA protein undergoes secondary proteolysis to generate additional peptides; one, a 10-kDa C-terminal peptide, has been immunologically identified previously (1). This 10-kDa protein, $p10^{CA}$, which eluted in fraction 140 of the acetonitrile gradient (Fig. 1A), was identi-

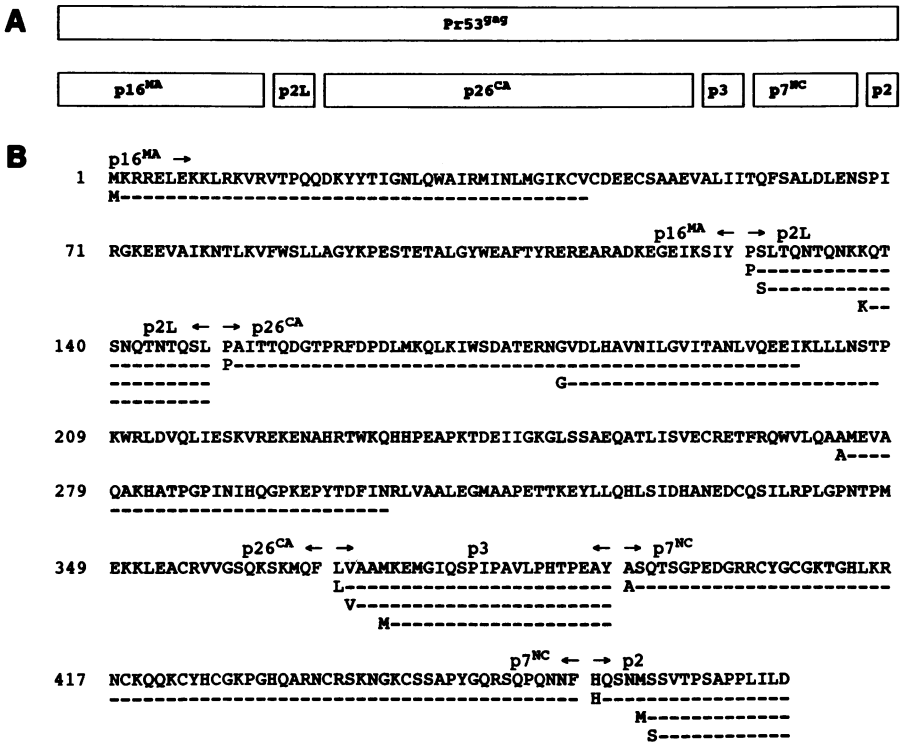

FIG. 2. Diagram of the proteolytic cleavage products of BIV Pr53^{gag} and amino acid sequence. (A) Proteolytic cleavage products of BIV Pr53^{gag}. Proteins shown are p16^{MA}, p2L, p26^{CA}, p3, p7^{NC}, and p2, as described in the text. (B) Amino acid sequence of BIV Pr53^{gag} as predicted from the DNA sequence of BIV R29-127 (11). N-terminal residues determined by N-terminal sequence analyses are indicated by single-letter abbreviations below the predicted sequence; other determined residues are indicated by dashes. Arrows indicate proteolytic cleavage sites between mature Gag proteins. Two additional cleavage products each of p2L, p3, and p2 were found to have N-terminal residues of Ser-128 and Lys-137, Val-369 and Met-372, and Met-463 and Ser-464 of Pr53^{gag}, respectively. The additional cleavage products are indicated below the major Gag subunits and are found in peaks a, b, e, f, c, and d, respectively, of Fig. 1A.

fied in SDS-PAGE and immunoblots using a peptide antiserum specific to the C terminus of CA (Fig. 1B and C, lanes 6). Edman degradation of p10^{CA} identified the Ala-274 of Pr53^{gag} as its N terminus (Fig. 2). Because of the detection of at least two phenylalanines by compositional analysis (Table 1), p10^{CA} most likely shares its C terminus with p26^{CA} (Fig. 2). Thus, p10^{CA} contains the 94 C-terminal residues of p26^{CA} and has a calculated mass of 10,381.9 Da. Although not immunologically identified, a second minor CA cleavage product with an apparent molecular mass of 11 kDa also was found in fraction 140. Sequence analysis identified its N-terminal residue as Gly-179 of Pr53^{gag} (Fig. 2). Its mobility in SDS-PAGE suggests that its C terminus was generated by proteolytic cleavage between Ala-273 and Ala-274. This CA peptide is predicted to have a mass of 10,804.3 Da and thus represents an internal CA fragment of 95 residues. Both the 10- and 11-kDa CA fragments were present in a molar ratio of approximately 2% with respect to the intact p26^{CA}. A putative peptide containing the amino-terminal 30 residues with a predicted mass of 3,445.8 Da was not identified in the rp-HPLC separation. The functions, if any, of the secondary CA cleavage products are not known.

(iii) NC proteins. A 13-kDa BIV protein previously was identified with HIV and BIV NC-specific antisera in immunoblots and radioimmunoprecipitations (1). The BIV NC eluted in fraction 25 of the HPLC gradient (Fig. 1A) and migrated as a 13-kDa protein in gel electrophoresis (Fig. 1B and C, lanes 3).

All 66 residues of the NC protein were confirmed by amino acid sequence analysis (Fig. 2). The presence of three alanine, one phenylalanine, three tyrosine, and three histidine residues characterizes the termini of the NC protein (Table 1). The N and C termini were Ala-393 and Phe-458 of Pr53^{gag}, respectively, and the BIV NC protein, p13^{NC}, has a calculated mass of 7,287.1 Da.

ESI-MS was of particular value in resolving the discrepancy between the molecular weights of the NC protein predicted by gel mobility and determined by amino acid composition. The ESI-MS of HPLC fractions 25 to 27 (Fig. 1A) characterizes an NC species of 7,283 Da (Fig. 3C and Table 2), which is in good agreement with the calculated mass of 7,287.1 Da. Minor NC proteins with masses of 7,475, 7,538, and 7,122 Da also were detected (Fig. 3C). The 7,283-, 7,475-, and 7,538-Da forms of NC were present in molar fractions of approximately 43, 18, and 37% of the total NC found. No evidence for dimeric forms of NC protein was found.

(iv) Small Pr53^{gag} cleavage products. Three small Gag-related peptides also were isolated during rp-HPLC separation of the viral proteins and were sequenced in their entirety. Fraction 21 (Fig. 1A) contained a peptide, designated p2L, which had a relative migration of 3 kDa in SDS-PAGE and was reactive to an antiserum directed to a region between the MA and CA proteins (Fig. 1B and C, lanes 2). Fraction 21 (Fig. 1A) also contained p13^{NC} that appeared to coelute with p2L (Fig. 1B, lane 2). N-terminal sequence analysis indicated that the N and C termini of p2L in fraction 21 were Pro-127 and Leu-148

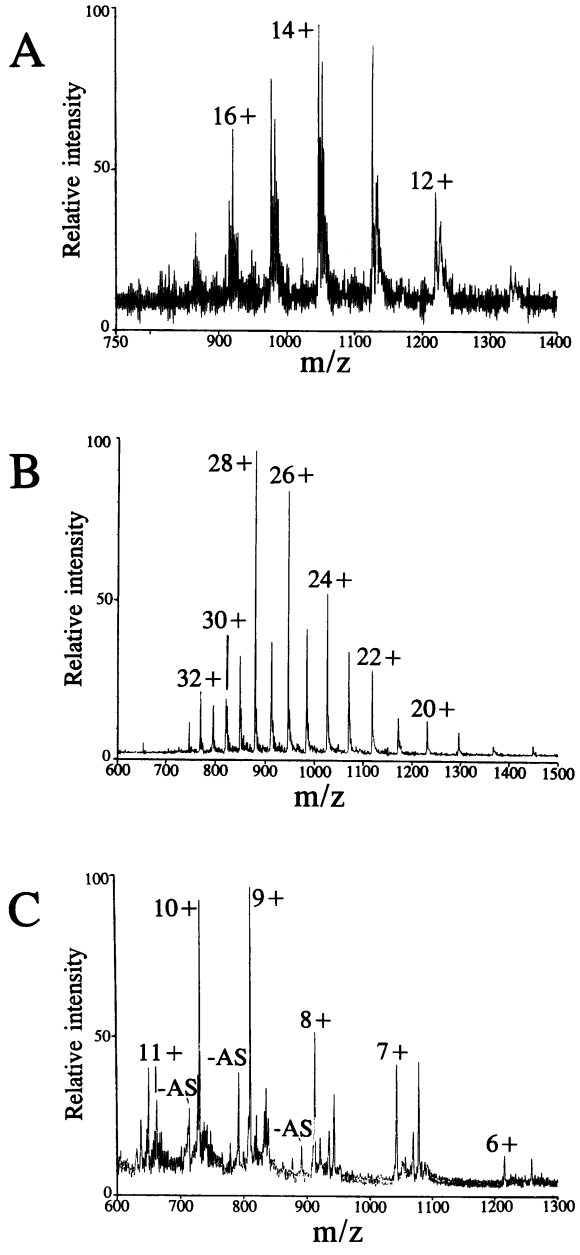

FIG. 3. Electrospray mass spectra of BIV Gag proteins isolated from virions by rp-HPLC (Fig. 1A). rp-HPLC-purified fractions dissolved in water-methanol-glacial acetic acid (48.5:48.5:3, vol/vol) were introduced by a syringe infusion pump at a flow rate of 10 μl/min into a Vestec electrospray source fitted to a Hewlett-Packard quadrupole mass spectrometer. Data were analyzed by standard deconvolution methodology (9, 20, 30). (A) MA protein contained in fractions 95 to 96 of inset III; (B) CA protein contained in fractions 58 to 61 of inset I; (C) NC protein contained in fractions 25 to 27 of initial separation. The ordinate indicates signal intensity relative to the maximum intensity in the spectrum; the abscissa indicates mass/charge ratios. Peaks are labeled with charge state values, and residue losses are indicated in panel C by single-letter abbreviations.

of the Gag precursor, respectively. Therefore, p2L has a calculated molecular mass of 2,461.6 Da. The presence of only one proline and two leucine residues confirmed the MA-p2L and the p2L-CA cleavage sites (Table 2). Two minor cleavage products of p2L that contained 7 and 8% of the total p2L were found in peaks a and b (Fig. 1A), respectively. The N-terminal proline was absent from the p2L in peak a, and the 10 N-terminal residues were missing from the p2L in peak b.

A second small peptide, p2, identified by amino acid analysis as the C-terminal protein of the Gag precursor, eluted in fraction 61 (Fig. 1A) but could not be resolved in SDS-PAGE or immunoblot analyses, probably because of its small size or solubility characteristics (Figs. 1B and C, lanes 4). Amino acid sequence analysis indicated that the N-terminal residue of p2 was His-459 of $Pr53^{gag}$, and its molecular mass is 1,894.1 Da. Fractions 58 and 63 contained two fragments of p2, denoted c and d (Fig. 1A). Peaks c and d were missing four and five N-terminal residues (Fig. 2) and contained 46 and 31% of the total p2 material, respectively.

A third peptide, p3, eluted in fraction 85 (Fig. 1A); it appeared as a diffuse band by SDS-PAGE but could not be visualized in immunoblot analysis (Fig. 1B and C, lanes 5). Amino acid compositional and sequence analyses confirmed the N and C termini of p3 (Table 1) as Leu-368 and Tyr-392 of $Pr53^{gag}$, respectively (Fig. 2). p3 has a calculated molecular mass of 2,664.2 Da. Peaks e and f, isolated from fractions 71 and 80 of the rp-HPLC separation, contained 20 and 14% of the total p3 material (Fig. 1A) and were missing one and six N-terminal residues, respectively (Fig. 2).

DNA sequence determination of MA gene. To determine whether the heterogeneity of MA observed by ESI-MS could have been caused by genetic mutations in the *gag* gene, we derived multiple clones of the MA-coding region. Extrachromosomal DNA preparations (24) were made from BIV R29-127-infected Cf2Th cultures at a passage immediately after the production of large-scale virus preparations for amino acid analyses and ESI-MS. Fragments overlapping the MA-coding region of the *gag* gene were amplified by PCR using BIV deoxyoligonucleotide primers 5′-AGAAGACTCCGGACAG GT-3′ and 5′-CCAGATCTTAAGCTGCTTCATGAGGT-3′ derived from the BIV R29-127 nucleotide sequence (11) and Vent polymerase (New England Biolabs). The DNA fragments were cloned into the *Sma*I site of pBS(II)SK+ (Stratagene). Twelve independent clones were sequenced by the method of Sanger et al. (38); only one clone was found to contain a transition (G→A at nucleotide 851). This resulted in an alanine-to-threonine replacement and increased the molecular mass by 30 Da. Thus, heterogeneity of the MA protein could have resulted, at least in part, from genetic mutations.

Phosphorylation of BIV Gag proteins. To determine if some of the heterogenity observed for Gag proteins was due to posttranslational modifications, phosphorylation of BIV proteins was studied by immunoprecipitation of $^{32}P_i$-labeled BIV-infected cell and virion lysates. BIV R29-127-infected BLAC-20 cell cultures were radiolabeled with [^{35}S]cysteine and [^{35}S]methionine (as a control for reactivity of sera in immunoprecipitation) or $^{32}P_i$, as previously described (34). Intracellular viral proteins were released by detergent lysis; virions were concentrated by ultracentrifugation of culture supernatants through 20% (wt/wt) sucrose and detergent lysed (1, 34). Radiolabeled samples were immunoprecipitated with a panel of BIV protein-specific antisera made to both structural and nonstructural proteins (1, 34), separated by SDS-PAGE, and visualized by autoradiography. No ^{32}P-labeled Gag proteins were observed after immunoprecipitation, whereas these antisera did recognize Gag proteins from ^{35}S-labeled cell

TABLE 2. Comparison of BIV Gag protein molecular masses determined by denaturing gel electrophoresis, amino acid analysis, and MS

Protein	Gel mobility[a]	No. of residues[b]	Mol wt[c]	
			Predicted	Derived ± SD
MA	p16	126	14,626.8	14,628 ± 6 14,708 ± 7 14,814 ± 12
CA	p26	219	24,596.0	24,610 ± 14 24,653 ± 15 24,682 ± 20 24,757 ± 28
NC	p13	66	7,287.1	7,283 ± 5 7,475 ± 3 7,538 ± 2

[a] Relative mobility of Gag proteins in SDS–16% PAGE.
[b] Determined by amino acid sequence.
[c] Predicted, calculated from amino acid sequence data and predicted from translation of published DNA sequence (11); Derived, calculated by ESI-MS (the most abundant mass form is listed first).

lysates (data not shown). In contrast, immunoprecipitation with Tat- and Rev-specific antisera resulted in the recovery of $^{32}P_i$-radiolabeled Tat and Rev proteins, indicating the effective incorporation of radiolabeled P_i into specific viral proteins (data not shown and reference 34).

In conclusion, several studies have determined the amino acid sequences of lentivirus Gag proteins purified from mature virions and deduced the subunit organization of the Gag precursors (7, 19, 20, 23). In addition to BIV MA, CA, and NC functional proteins, three small Gag proteins (p2L, p3, and p2) were characterized in the present study, which indicated that the processing of the BIV Gag precursor is more complex than was revealed from the previous detailed immunological studies (1, 35). Comparison of the empirically derived amino acid sequences with those predicted from translation of the BIV DNA sequence (11) demonstrates that the order of the six major proteins in the BIV Gag precursor is NH_2-MA-p2L-CA-p3-NC-p2-COOH.

Most retroviral Gag precursor N termini are posttranslationally modified by fatty acids (e.g., myristic acid), which is believed to play a role in localizing the Gag precursor to the plasma membrane during virus assembly (7, 19–21, 31, 36, 37, 39, 40). Mutated *gag* genes, which result in Gag precursors that are not myristylated, are not capable of forming virions (18, 36, 40). Although we previously demonstrated that the BIV Gag precursor or MA protein is not labeled in vivo with [^{14}C]myristic acid (1, 35), we speculated that an alternative fatty acid modification, such as acetylation, may be present. Since the N terminus of the BIV Gag precursor is not blocked to Edman degradation, retains the initiator methionine, and thus is not modified by fatty acids, it is unique among those encoded by other retroviruses studied. Interestingly, Mason-Pfizer monkey virus MA mutants that are myristylated but contain small internal deletions near the N-terminal region fail to assemble (37). Therefore, while myristylation may be important in some systems, an alternative mechanism must exist for the localization of the BIV Gag precursor in the plasma membrane and assembly of the virus particle.

A C-terminal protein is also present in BIV; however, in contrast to the MA, CA, and NC proteins, the C-terminal proteins of lentiviruses from different species are quite variable in size and range from 2 to 9 kDa. Several studies suggest that the HIV-1 C-terminal protein (p6) has a role in virus assembly (12, 28); however, deletion of C-terminal residues still results in particle formation in the baculovirus-insect cell expression system (41). Thus, the role of p6 in the HIV-1 Gag precursor remains controversial. Alignment of the C-terminal domains of several lentiviruses demonstrates that BIV p2 contains a conserved PS/TAPP motif (data not shown). The presence of this conserved domain suggests a common, albeit unknown, function. Although the amino acid sequences of ovine and caprine Gag proteins have not been determined, the proximity of the PS/TAPP motif to the C termini of their Gag precursors suggests that their putative C-terminal Gag proteins are similar in size to the BIV p2. The only lentivirus that lacks the PS/TAPP domain is equine infectious anemia virus, which is also the only lentivirus that lacks a *vif* gene (27, 32). Vif has been implicated in the increased infectivity of the lentivirus virion (3, 6). Although the mechanism of this phenomenon is unknown, it could be conjectured that Vif interacts with other viral proteins, such as the C terminus of the Gag precursor, during the budding process.

BIV is unique among the lentiviruses in having an intragenic peptide (p2L) between the MA and CA proteins. Such a peptide has been described for other retrovirus genera (22, 25). DNA sequence data indicate that new field isolates of BIV contain p2L and, despite p2L sequence heterogeneity, residues in the cleavage sites are well conserved (10). In lentiviruses, a small peptide resides between the CA and NC domains; 5-, 17-, 14-, and 9-residue peptides are removed during processing of the CA and NC proteins of equine infectious anemia virus (23), SIV (19), HIV-1 (20), and feline immunodeficiency virus (7), respectively. This peptide (p3) in BIV is slightly larger, 25 residues. A small peptide is also found between the NC and C-terminal proteins in HIV-1 and SIV (19, 20); in contrast, BIV, equine infectious anemia virus, and feline immunodeficiency virus lack this peptide (7, 23). Whether the small intragenic peptides are simply spacer peptides used in processing the Gag precursor by the viral PR or are necessary for interaction with other Gag precursor domains during virus assembly is unknown.

The BIV NC protein has been referred to as $p13^{NC}$ on the basis of its mobility in denaturing gels (1, 35). The original predicted molecular mass (13 kDa) based on the amino acid translation for the BIV NC protein did not take into account the cleavage of the CA-NC intragenic p3 or the C-terminal p2 (11). The MS and the amino acid sequence data generated in the present study (Table 2) are in good agreement and indicate that the mass of the NC protein is 7.3 kDa. Its slow electrophoretic mobility in denaturing gel electrophoresis is probably a result of a high pI of 10.47.

When the PR cleavage sites for the BIV Gag precursor are compared with those of HIV-1 (20), only the BIV MA-p2L (Tyr-Pro) and HIV-1 MA-CA (Tyr-Pro) sites appear to be conserved. Interestingly, preliminary experiments using recombinant HIV-1 and BIV PRs indicate that recombinant BIV and HIV-1 Gag precursors can be cleaved by either viral PR to yield some of the expected mature products found in the virion (41). When the BIV PR sequence was fitted to the three-dimensional structure of the HIV-1 PR (44), the active sites of the two PR dimers were superimposable with the exception of two residues, despite limited sequence similarity (8). Although the mechanism by which viral PRs recognize and cleave Gag precursor is not completely understood, the cleavage sites are believed to be determined largely by the primary amino acid sequence and conformation of precursor polyproteins in the immature particle (19). Further investigation into the processing of the recombinant BIV Gag precursor by heterologous

viral PRs should identify the cleavage sites used in vitro and provide information on the mechanism of viral PR specificity. Moreover, knowledge of viral PR cleavage sites in the Gag precursor will facilitate the molecular cloning and expression of individual BIV Gag proteins for further study.

We thank J. Bess for production of virus stocks, J. Dunkle for PCR cloning and DNA sequencing, L. Rasmussen for radioimmunoprecipitations, S. Rivard for photographic assistance, and G. Serig for help in preparation of the manuscript. We also thank the Biomedical Supercomputing Center, PRI/DynCorp, NCI-FCRDC, for computer hardware and software support for DNA and protein sequence analysis.

REFERENCES

1. **Battles, J. K., M. Y. Hu, L. Rasmussen, G. J. Tobin, and M. A. Gonda.** 1992. Immunological characterization of the *gag* gene products of bovine immunodeficiency virus. J. Virol. **66:**6868–6877.
2. **Benton, C. V., H. M. Hodge, and D. L. Fine.** 1978. Comparative large-scale propagation of retroviruses from Old World (Mason-Pfizer monkey virus) and New World (squirrel monkey virus) primates. In Vitro (Rockville) **14:**192–199.
3. **Blanc, D., C. Patience, T. F. Schulz, R. Weiss, and B. Spire.** 1993. Transcomplementation of VIF− HIV-1 mutants in CEM cells suggests that VIF affects late steps of the viral life cycle. Virology **193:**186–192.
4. **Braun, M. J., S. Lahn, A. L. Boyd, T. A. Kost, K. Nagashima, and M. A. Gonda.** 1988. Molecular cloning of biologically active proviruses of bovine immunodeficiency-like virus. Virology **167:**515–523.
5. **Carpenter, S., L. D. Miller, S. Alexandersen, C. A. Whetstone, M. J. Van Der Maaten, B. Viuff, Y. Wannemuehler, J. M. Miller, and J. A. Roth.** 1992. Characterization of early pathogenic effects after experimental infection of calves with bovine immunodeficiency-like virus. J. Virol. **66:**1074–1083.
6. **Cullen, B. R., and W. C. Greene.** 1990. Functions of the auxiliary gene products of the human immunodeficiency virus type 1. Virology **178:**1–5.
7. **Elder, J. H., M. Schnölzer, C. S. Hasselkus-Light, M. Henson, D. A. Lerner, T. R. Phillips, P. C. Wagaman, and S. B. H. Kent.** 1993. Identification of proteolytic processing sites within the Gag and Pol polyproteins of feline immunodeficiency virus. J. Virol. **67:**1869–1876.
8. **Erickson, J. (PRI/DynCorp, National Cancer Institute-Frederick Cancer Research and Development Center).** 1993. Personal communication.
9. **Fenselau, C.** 1991. Beyond gene sequencing: analysis of protein structure with mass spectrometry. Annu. Rev. Biophys. Biophys. Chem. **20:**205–220.
10. **Garvey, K. J., et al.** Unpublished data.
11. **Garvey, K. J., M. S. Oberste, J. E. Elser, M. J. Braun, and M. A. Gonda.** 1990. Nucleotide sequence and genome organization of biologically active proviruses of the bovine immunodeficiency-like virus. Virology **175:**391–409.
12. **Gelderblom, H. R., M. Özel, and G. Pauli.** 1989. Morphogenesis and morphology of HIV, structure-function relations. Arch. Virol. **106:**1–13.
13. **Gonda, M. A.** 1992. Bovine immunodeficiency virus. AIDS **6:**759–776.
14. **Gonda, M. A.** 1994. The lentiviruses of cattle, p. 83–109. *In* J. A. Levy (ed.), The Retroviridae. Plenum, New York.
15. **Gonda, M. A.** 1994. Molecular biology and virus-host interactions of lentiviruses. Ann. N. Y. Acad. Sci. **724:**22–42.
16. **Gonda, M. A., M. J. Braun, S. G. Carter, T. A. Kost, J. W. Bess, Jr., L. O. Arthur, and M. J. Van Der Maaten.** 1987. Characterization and molecular cloning of a bovine lentivirus related to human immunodeficiency virus. Nature (London) **330:**388–391.
17. **Gonda, M. A., D. G. Luther, S. E. Fong, and G. J. Tobin.** 1994. Bovine immunodeficiency virus: molecular biology and virus-host interactions. Virus Res. **32:**155–181.
18. **Göttlinger, H. G., J. G. Sodroski, and W. A. Haseltine.** 1989. Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc. Natl. Acad. Sci. USA **86:**5781–5785.
19. **Henderson, L. E., R. E. Benveniste, R. Sowder, T. D. Copeland, A. M. Schultz, and S. Oroszlan.** 1988. Molecular characterization of *gag* proteins from simian immunodeficiency virus (SIV_{Mne}). J. Virol. **62:**2587–2595.
20. **Henderson, L. E., M. A. Bowers, R. C. Sowder II, S. A. Serabyn, D. G. Johnson, J. W. Bess, Jr., L. O. Arthur, D. K. Bryant, and C. Fenselau.** 1992. Gag proteins of the highly replicative MN strain of human immunodeficiency virus type 1: posttranslational modifications, proteolytic processings, and complete amino acid sequences. J. Virol. **66:**1856–1865.
21. **Henderson, L. E., H. C. Krutzsch, and S. Oroszlan.** 1983. Myristyl amino-terminal acylation of murine retrovirus proteins: an unusual post-translational protein modification. Proc. Natl. Acad. Sci. USA **80:**339–343.
22. **Henderson, L. E., R. Sowder, T. D. Copeland, G. Smythers, and S. Oroszlan.** 1984. Quantitative separation of murine leukemia virus proteins by reversed-phase high-pressure liquid chromatography reveals newly described *gag* and *env* cleavage products. J. Virol. **52:**492–500.
23. **Henderson, L. E., R. C. Sowder, G. W. Smythers, and S. Oroszlan.** 1987. Chemical and immunological characterizations of equine infectious anemia virus *gag*-encoded proteins. J. Virol. **61:**1116–1124.
24. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–369.
25. **Hizi, A., L. E. Henderson, T. D. Copeland, R. C. Sowder, H. C. Krutzsch, and S. Oroszlan.** 1989. Analysis of *gag* proteins from mouse mammary tumor virus. J. Virol. **63:**2543–2549.
26. **Jacobs, R. M., H. E. Smith, B. Gregory, V. E. O. Valli, and C. E. Whetstone.** 1992. Detection of multiple retroviral infections in cattle and cross-reactivity of bovine immunodeficiency-like virus and human immunodeficiency virus type 1 proteins using bovine and human sera in a Western blot assay. Can. J. Vet. Res. **56:**353–359.
27. **Kawakami, T., L. Sherman, J. Dahlberg, A. Gazit, A. Yaniv, S. R. Tronick, and S. A. Aaronson.** 1987. Nucleotide sequence analysis of equine infectious anemia virus proviral DNA. Virology **158:**300–312.
28. **Lever, A., H. Gottlinger, W. Haseltine, and J. Sodroski.** 1989. Identification of a sequence required for efficient packaging of human immunodeficiency virus type 1 RNA into virions. J. Virol. **63:**4085–4087.
29. **Liu, Z. Q., D. Sheridan, and C. Wood.** 1992. Identification and characterization of the bovine immunodeficiency-like virus *tat* gene. J. Virol. **66:**5137–5140.
30. **Mann, M., C. K. Meng, and J. B. Fenn.** 1989. Interpreting mass spectra of multiply charged ions. Anal. Chem. **61:**1702–1708.
31. **Mervis, R. J., N. Ahmad, E. P. Lillehoj, M. G. Raum, F. H. R. Salazar, H. W. Chan, and S. Venkatesan.** 1988. The *gag* gene products of human immunodeficiency virus type 1: alignment within the *gag* open reading frame, identification of posttranslational modifications, and evidence for alternative *gag* precursors. J. Virol. **62:**3993–4002.
32. **Oberste, M. S., and M. A. Gonda.** 1992. Conservation of amino-acid sequence motifs in lentivirus Vif proteins. Virus Genes **6:**95–102.
33. **Oberste, M. S., J. D. Greenwood, and M. A. Gonda.** 1991. Analysis of the transcription pattern and mapping of the putative *rev* and *env* splice junctions of bovine immunodeficiency-like virus. J. Virol. **65:**3932–3937.
34. **Oberste, M. S., J. C. Williamson, J. D. Greenwood, K. Nagashima, T. D. Copeland, and M. A. Gonda.** 1993. Characterization of bovine immunodeficiency virus *rev* cDNAs and identification and subcellular localization of the Rev protein. J. Virol. **67:**6395–6405.
35. **Rasmussen, L., J. K. Battles, W. H. Ennis, K. Nagashima, and M. A. Gonda.** 1990. Characterization of virus-like particles produced by a recombinant baculovirus containing the *gag* gene of the bovine immunodeficiency-like virus. Virology **178:**435–451.
36. **Rein, A., M. R. McClure, N. R. Rice, R. B. Luftig, and A. M. Schultz.** 1986. Myristylation site in $Pr65^{gag}$ is essential for virus particle formation by Moloney murine leukemia virus. Proc. Natl. Acad. Sci. USA **83:**7246–7250.
37. **Rhee, S. S., and E. Hunter.** 1990. Structural role of the matrix

protein of type D retroviruses in Gag polyprotein stability and capsid assembly. J. Virol. **64**:4383–4389.
38. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
39. **Schultz, A. M., L. E. Henderson, and S. Oroszlan.** 1988. Fatty acylation of proteins. Annu. Rev. Cell Biol. **4**:611–647.
40. **Schultz, A. M., and A. Rein.** 1989. Unmyristylated Moloney murine leukemia virus $Pr65^{gag}$ is excluded from virus assembly and maturation events. J. Virol. **63**:2370–2373.
41. **Tobin, G. J., et al.** Unpublished data.
42. **Toplin, I., and P. Sottong.** 1972. Large-volume purification of tumor viruses by use of zonal centrifuges. Appl. Microbiol. **23**: 1010–1014.
43. **Van Der Maaten, M. J., A. D. Boothe, and C. L. Seger.** 1972. Isolation of a virus from cattle with persistent lymphocytosis. J. Natl. Cancer Inst. **49**:1649–1657.
44. **Wlodawer, A., M. Miller, M. Jaskólski, B. K. Sathyanarayana, E. Baldwin, I. T. Weber, L. M. Selk, L. Clawson, J. Schneider, and S. B. H. Kent.** 1989. Conserved folding in retroviral proteases: crystal structure of a synthetic HIV-1 protease. Science **245**:616–621.