JOURNAL OF VIROLOGY, May 1996, p. 3313–3318 Vol. 70, No. 5
0022-538X/96/$04.00+0

Analysis of Rous Sarcoma Virus Gag Proteins by Mass Spectrometry Indicates Trimming by Host Exopeptidase

R. BLAKE PEPINSKY,[1]* IOANNIS A. PAPAYANNOPOULOS,[1] STEPHEN CAMPBELL,[2] AND VOLKER M. VOGT[2]

Department of Protein Chemistry, Biogen, Inc., Cambridge, Massachusetts 02142,[1] and Section of Biochemistry and Molecular and Cell Biology, Cornell University, Ithaca, New York 14853[2]

Received 8 November 1995/Accepted 13 February 1996

We have used electrospray ionization-mass spectrometry to investigate Gag protein structure and processing in Rous sarcoma virus, the prototype of the avian sarcoma and leukemia viruses. Molecular masses determined for the mature virion proteins MA, CA, NC, and PR agree closely with those predicted by currently accepted models for their structures. However, the data for p10 imply that only about 10% of the product has the predicted mass while the remainder is missing the C-terminal methionine residue. Molecular masses also were obtained for products generated by PR cleavage in vitro of a Gag precursor polyprotein expressed in *Escherichia coli*. The data confirm the predicted Gag cleavage sites for PR. Thus, carboxypeptidase activity appears to be responsible for generating the des-Met form of p10. The same activity may account for the small amount of the mature des-Met CA, as previously reported. Analysis of cleavage products generated in vitro also serves to define the PR processing site separating the p2a and p2b peptides, Asn-164–Cys-165. In conjunction with published characterizations of these two peptides processed from the segment of Gag between MA and p10, these data suggest trimming of p2b by an aminopeptidase. Finally, the molecular masses determined for the MA-related species p19f, p23, and p35 now accurately define the structures of these proteins.

The internal structural proteins of retroviruses are generated through proteolytic processing of a single precursor polyprotein encoded by the *gag* gene. Cleavage of the Gag polyprotein leads to a morphological change in the virus particle known as maturation and is essential for infectivity. In avian sarcoma and leukemia viruses (ASLV), the Gag precursor is processed into five major proteins, from the N to C terminus: MA (matrix), p10, CA (capsid), NC (nucleocapsid), and PR (protease), in addition to several small peptides between MA and p10 (19) and between CA and NC (1, 5) (Fig. 1A). With the exception of PR and NC, which have been characterized by classical protein sequencing methods (16, 24), information on the primary structures of the ASLV Gag proteins is based largely on N- and C-terminal analyses and on comparison of these data with the DNA sequences (25). Recently, we investigated the structure of CA in the mature virus through peptide mapping and mass spectrometry (MS) (20). These studies revealed that ASLV CA is not a homogeneous species as previously inferred but rather exists in three forms that differ by several amino acid residues at the C terminus, because of differential processing. Recombinant viruses lacking one or more of these processing sites were noninfectious, suggesting that they are critical for maturation. The CA protein of human immunodeficiency virus type 1 shows processing events at its C terminus remarkably similar to those in ASLV (7, 8, 15, 23, 27).

The ability to rapidly access the CA structure by electrospray ionization (ESI)-MS suggested that this strategy could be readily applied to the other Gag proteins. Here we have expanded these studies to test the fidelity of the predictions of the primary structures of mature ASLV Gag proteins and to analyze the cleavage products generated by cleavage of recombinant Gag precursors with PR in vitro. The ESI-MS data provide a detailed view of the Gag protein structure and processing that is unattainable by other current methods and demonstrates the power of this strategy for probing structure. We infer from the results that at least one and probably other non-PR proteolytic activities modify Gag cleavage products to generate the forms found in the mature virus.

Analysis of proteins in virions. The proteins from the Prague C strain of Rous sarcoma virus (PrC RSV) were isolated by reversed-phase high-performance liquid chromatography (HPLC) on a C_4 column, which led to resolution of all the major species (Fig. 1C). Proteins from each of the peaks were identified by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The five major Gag proteins, which together account for about 80% of the total virion protein, are readily apparent. The other prominent peak between p10 and PR contained the Env proteins SU (gp85) and TM (gp37), which together account for about 20% of the total virion protein. Each of the Gag protein peaks was subjected to ESI-MS analysis and produced a mass spectrum that was readily interpretable on the basis of the known PrC RSV Gag sequence (25) (Table 1). In each case a component with the expected mass was observed.

ESI-MS data for MA showed two primary species with masses of 16,574 and 16,654 Da. The 16,574-Da ion has the exact mass predicted for MA based on termination at Tyr-155 and on acetylation of the initiating methionine (predicted mass of 16,574.1 Da with acetylation versus a mass of 16,532.1 Da without acetylation), thus confirming the N- and C-terminal predictions that were made by biochemical methods (17, 28). The difference in mass, 80 Da, reflects the contribution of a single phosphate. These observations are consistent with the known migration of MA as a doublet of bands on SDS-PAGE, with the upper band being phosphorylated but not the lower (11, 22), and rule out phosphorylation at more than one site on a single MA molecule, as was previously suggested to occur (14). The mass spectrum (Fig. 2B) also revealed a series of

* Corresponding author. Mailing address: Department of Protein Chemistry, Biogen, Inc., 14 Cambridge Center, Cambridge, MA 02142. Phone: (617) 679-3310. Fax: (617) 679-2616.

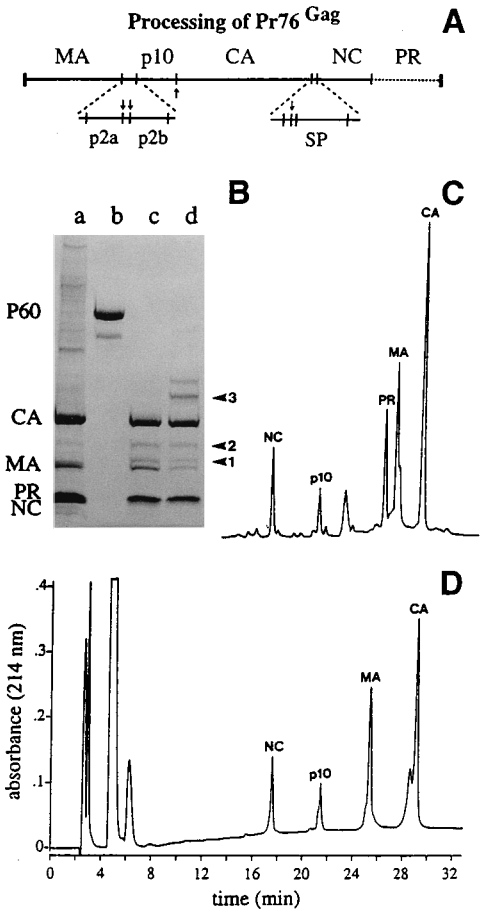

FIG. 1. Analysis of P60 Gag cleavage products. (A) Diagram of Gag Pr76 with sections containing closely spaced cleavage sites expanded. Cleavage sites are indicated with hatch marks. The P60 Gag polyprotein purified from *E. coli* is derived from a DNA construct in which a termination codon was introduced in place of the first amino acid residue of PR and therefore contains all of the Pr76 sequence except PR. Small arrows indicate discrepancies in PR-processing sites that we investigated by ESI-MS. Those at the p10-CA junction and in SP reflect sites requiring carboxypeptidase activity. The two sites between p2a and p2b reflect a series of processing steps that remain to be definitively elucidated. (B) SDS-PAGE analysis of in vitro cleavage products. P60 in 50 mM sodium acetate (pH 6.0)–600 mM NaCl was treated with AMV PR for 16 h at 37°C and analyzed for the extent of cleavage by SDS-PAGE. Gels were stained with Coomassie blue. Lanes: a, PrC RSV virus (10 μg); b, P60 (5 μg); c, P60 + PR (100:1, 10 μg); d, P60 + PR (1,000:1, 10 μg). Arrowheads mark the positions of MA-p2a (1), MA-p2a-p2b (2), and MA-p2a-p2b-10 (3). The unmarked band above MA-p2a-p2b-p10 in lane d corresponds to the CA-SP-NC partial cleavage product. (C and D) HPLC analysis of viral proteins. Aliquots of PrC RSV or P60-PR cleavage products (1,000:1 ratio) containing approximately 100 μg of total protein were dissolved in 150 μl of 6 M guanidine HCl–12.5 mM HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid; pH 8.0)–5 mM dithiothreitol and incubated at 37°C for 1 h. The samples were subjected to reversed-phase HPLC on a C_4 column (Vydac catalog no. 214TP104 column dimensions, 0.46 cm [inside diameter] by 25 cm) at ambient temperature. Bound components were eluted with a 30-min 0 to 70% gradient of acetonitrile in 0.1% trifluoroacetic acid at a flow rate of 1.4 ml/min. The column effluent was monitored at 280 nm, and 0.5-min fractions were collected. Aliquots of protein-containing fractions (50 μl) were dried in a Speed Vac concentrator, dissolved in electrophoresis sample buffer, and analyzed by SDS-PAGE. The identity of the samples is indicated in the figure. The peak eluting at 29 min as a shoulder to the CA peak contains the 35-kDa CA-SP-NC partial cleavage product. The remainder of the samples were concentrated 10-fold and subjected to ESI-MS. (C) PrC RSV; (D) in vitro cleavage products.

TABLE 1. Analysis of PrC RSV Gag proteins by ESI-MS

Protein	Position[a]	Mass (Da)	
		Calculated	Measured in virus
MA[b]	1–155	16,574.1	16,654, 16,574
p10	178–238	6,044.6	6,044
CA[c]	240–479	25,815.9	25,815, 25,683, 25,543
NC[d]	489–577	9,492.6	9,492
PR	578–701	13,609.9	13,609
MA-p19f	1–123, 1–129	13,176.3, 13,776.6	13,857, 13,777, 13,258, 13,177

[a] Positions refer to the sequence of the Gag precursor polyprotein (25).
[b] The calculated mass of MA assumes that Met-1 is acetylated (17). The 16,654-Da species corresponds to the phosphorylated form of MA.
[c] CA variants are due to processing at its C terminus (20). The 25,815-, 25,683-, and 25,543-Da species account for 36, 9, and 55% of total CA, respectively.
[d] The original sequence data for PrC RSV revealed several amino acid variations for NC and p10 because sequencing was performed on cDNAs obtained from a mixture of viruses (25). The molecular masses observed here indicate the presence of the Lys-561-to-Thr variant for NC and the Ser-204-to-Pro variant for p10. The calculated masses for NC and p10 are based on these variations. Calculated and observed masses for recombinant NC and p10 reflect the presence of Lys at position 561 and Ser at 204 (Table 2).

minor species at 13,777, 13,857, 13,258, and 13,177 Da. These products appear to be phosphorylated-unphosphorylated pairs of p19-related forms that correspond by mass to residues 1 to 129 and 1 to 123 in the MA sequence. These cleavage sites occur within a stretch of charged amino acids ending at Asp-123 and Gly-129, respectively, and appear not to conform to the consensus cleavage site sequences for PR, which has a preference for exposed neutral or hydrophobic amino acids. The fragments probably represent the truncated form of MA, called p19f, that was first recognized by Shealy and coworkers (26) but had not been characterized. We speculate that these species are produced by a protease other than PR that is incorporated into virions. A third pair of fragments with masses of 16,119 and 16,199 Da, which correspond to phosphorylated and unphosphorylated forms of residues 1 to 151, was also observed. This cleavage site, after Gly-151, bears resemblance to the sequence at the junction between p2b and p10 and seems likely to be recognized as an alternative processing site by PR. Finally, a small peak at 16,556 Da (18 Da smaller than intact MA), which probably was generated by dephosphorylation of MA, was observed. The difference in mass is consistent with the formation of dehydroalanine from phosphoserine, which is frequently observed as a by-product of this reaction and is 18 Da smaller than serine. Whether the dehydroalanine form of MA reflects a natural product or one generated during storage and purification of virus and viral proteins is unknown.

Inferences about the structure of p10 were previously based on N-terminal sequencing, amino acid analysis, and predictions made from DNA sequencing (2, 25). Since the N terminus of CA, the adjoining protein on the Gag precursor, is Pro-240, the C terminus of p10 was assumed to be Met-239 (9, 21). However, studies with p10 labeled with radioactive methionine, which were designed to test this assumption, revealed that at best only a small fraction of p10 ended in methionine (9, 18, 21). Since radioactive labeling allowed conclusions only about the presence or absence of methionine, it was unclear whether p10 was missing only this residue or additional upstream residues. The latter possibility would suggest an additional PR cleavage site in close proximity. The ESI-MS data for p10 (Fig. 3A) revealed that mature p10 ends at Ala-238. The observed mass of 6,044 Da agrees exactly with the predicted mass of p10 once it has been adjusted for loss of the C-terminal methionine (6,044.6 Da). A minor peak with a mass of 6,176 Da also was observed and accounted for about 10% of the p10 on the basis of peak height. This peak presumably represents the fraction of p10 ending in methionine, which has a predicted mass of 6,175.8 Da. Since retroviral proteases have no known exopep-

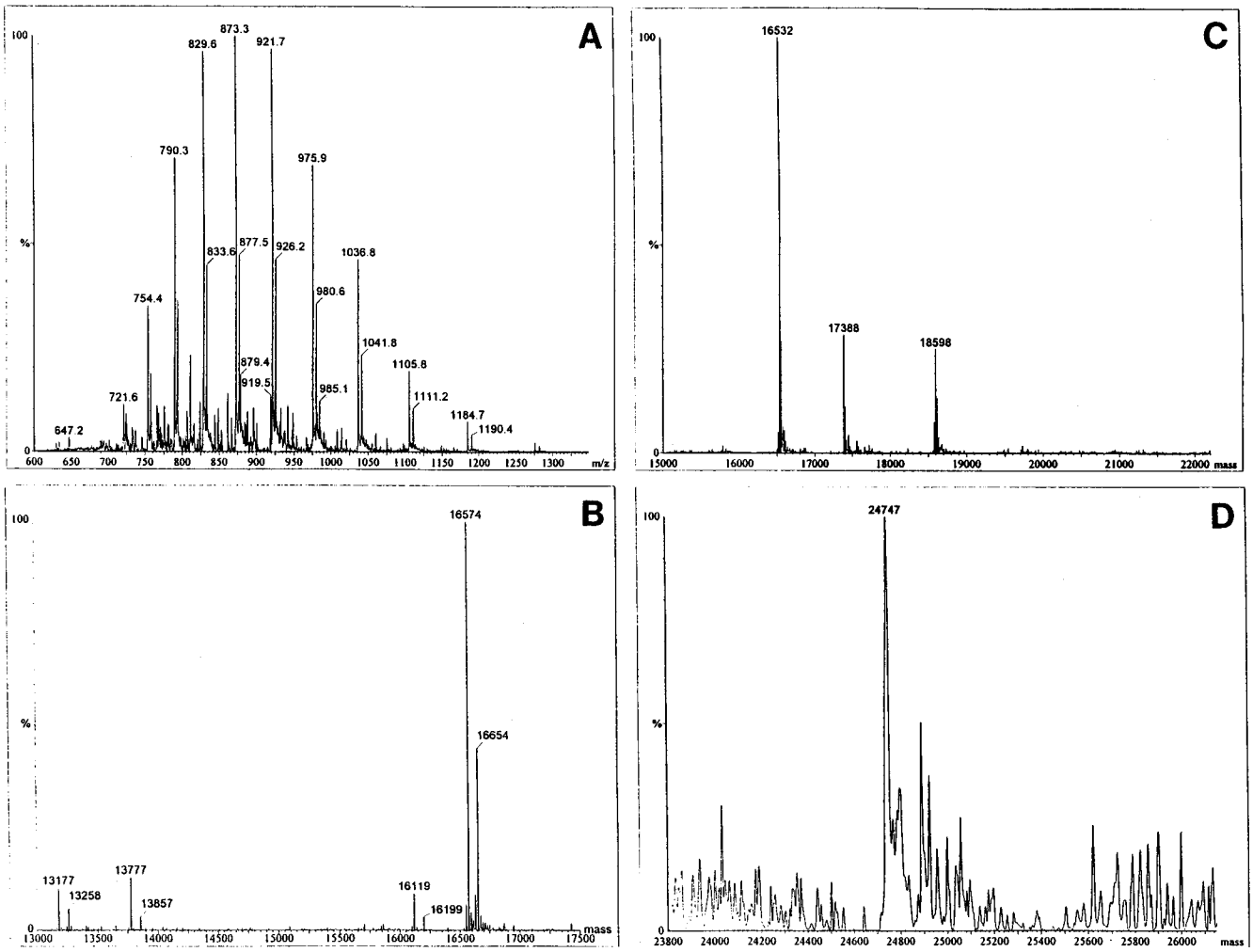

FIG. 2. Analysis of MA and its related products by ESI-MS. The MA-related proteins from PrC RSV and in vitro cleavage of P60 were analyzed by ESI-MS on a VG Quattro II triple quadrupole mass spectrometer equipped with an electrospray ion source. A volume of 20 μl of HPLC-purified MA in 50% water–50% acetonitrile (with 0.1% trifluoroacetic acid) was directly infused into the ion source at a rate of 5 μl/min. Scans were acquired throughout the sample infusion. All ESI-MS data were acquired and stored in profile mode and were processed by using the VG MassLynx data system. (A) Mass/charge data that were collected for MA from virus; (B) Molecular mass spectrum for MA generated from the data shown in panel A. Peak assignments that were generated by the data system are shown. The molecular mass spectra for in vitro-cleaved MA after digestion with PR at PR:P60 ratios of 1:100 (C) and 1:1,000 (D) are shown.

tidase activity, we infer that PR cleaves the Gag precursor at the Met-239–Pro-240 site and that the methionine residue then is removed by a carboxypeptidase. As discussed below, this hypothesis is supported by the results of Gag protein cleavage in vitro, which show formation of only the p10 species ending in Met-239. An additional component in the p10 preparation that we were unable to identify had a mass of 5,955 Da.

We described previously the characterization and origin of three different CA components in ASLV (20). The species with a mass of 25,815 Da, representing about a third of CA, has exactly the theoretical mass, 25,815.9 Da, for CA with the C terminus being Met-479, as originally inferred by carboxypeptidase analysis (2). The species with a mass of 25,543 Da, representing the majority of CA, has its C terminus at Ala-476, with a predicted mass of 25,542.6 Da. This species is the product of alternative processing by PR. The minor species with a mass of 25,683 Da has its C terminus at Ala-478, with a predicted mass of 25,684.6 Da. This species thus is analogous to p10, both of which lack a C-terminal Met residue.

The molecular mass of 9492 Da for NC agrees exactly with the predicted mass for one of three known PrC RSV NC variants (Table 1) previously identified by DNA sequencing (25) and indicates that the virus characterized contains threonine at position 561. No attempt was made to verify the presence of this residue, however. According to two earlier reports, some NC molecules in ASLV are phosphorylated (6, 11). Later reports also described a phosphorylated NC species (12, 13). Therefore we checked the masses of the minor peaks in the chromatogram for evidence of phosphate. No such evidence was found. The discrepancy in part might be explained by the confusion of NC and the MA fragments discussed above, which comigrate in some SDS-PAGE systems. This possibility would be consistent with the recent finding that in an RSV mutant engineered to lack the single MA phosphorylation site, the virions show no phosphorylated proteins whatsoever (16a). Alternatively, it is possible either that NC phosphorylation is highly transient and the phosphate was removed before analysis or that phosphorylation occurs on such a small fraction of the NC molecules that it was not detected.

The predicted and observed masses for PR agreed exactly,

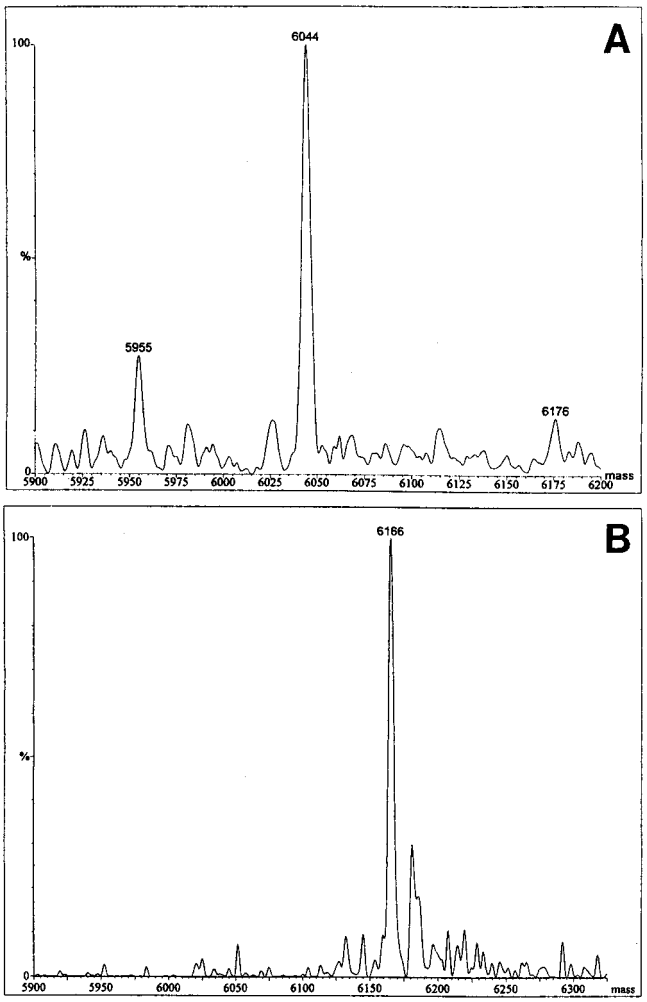

FIG. 3. Analysis of p10 by ESI-MS. The p10 products from PrC RSV and in vitro cleavage of P60 were analyzed by ESI-MS. The molecular spectra for p10 from virus (A) and from in vitro cleavage of P60 (B) are shown (mass assignments were generated by the data system).

13,609.9 and 13,609 Da, respectively (Table 1). In addition, the mass spectrum for PR contained a minor component with mass of 13,294 Da. We infer that this form is a proteolytic product of PR that is missing the tripeptide Leu-Ala-Met at the N terminus (predicted mass of 13,294.5 Da). The truncated form accounts for 7% of the PR by peak height. The observation of an alternative processing site near the junction between NC and PR is consistent with analyses of a mutant virus with a defective cleavage site between NC and PR (24a).

Analysis of proteins generated by cleavage in vitro. Processing of the Gag polyprotein precursor was tested in vitro with a purified recombinant protein expressed in *E. coli* as a substrate. This protein, here referred to as P60, consists of the Gag polypeptide beginning at the natural initiating methionine residue and terminating at the last residue of NC, Ser-577. The sequence of P60 thus is identical to that of Gag except that it is missing the PR domain (Fig. 1A). The purification of P60 consisted of solubilization of inclusion bodies at pH 12, neutralization, ammonium sulfate precipitation, and chromatography on phosphocellulose, yielding a protein that was over 95% pure and soluble at neutral pH and 0.5 M NaCl to greater than 1 mg/ml (4a). Some in vitro cleavage experiments also were performed with a shorter Gag protein, CA-SP-NC, purified as described previously (4). Purified PR obtained from avian myeloblastosis virus (AMV) by extraction with chloroform-methanol (10) was used as the enzyme. Cleavages were performed at pH 5.5 or pH 6.8 at enzyme-to-substrate ratios of 1:100, 1:1,000, and 1:2,000 and for 3 or 18 h at 37°C. The SDS-PAGE profiles from pH 5.5, 18-h digests of P60 treated with PR at ratios of 1:100 and 1:1,000 appeared similar, with both treatments promoting the formation of CA, MA, and NC-like products (Fig. 1B, lanes c and d). However, at the lower protease concentration, some larger polypeptides were apparent and the yield of MA was reduced, indicating incomplete digestion. The products were separated on the same reversed-phase C_4 column described above (Fig. 1D for the 1:1,000 sample; data for the 1:100 sample are not shown). At the characteristic elution point for MA, bands with apparent masses of 17, 19, 23, and 35 kDa on SDS-PAGE were observed. The ratios of these MA-containing forms differed considerably at the two protease concentrations, indicating that the larger polypeptides were incompletely digested forms of MA. The elution profiles of NC and CA were the same as those for the mature proteins (Fig. 1C), while p10 eluted 1 min later and MA eluted 3 min earlier than the corresponding virion proteins.

Selected HPLC samples were analyzed by ESI-MS. Results from these analyses are presented in the individual mass spectra indicated below and are summarized in Table 2. Each of the products was readily identified by its molecular mass. Several conclusions are apparent from these studies. For NC, processing led to the predicted NC cleavage product. Similarly, processing yielded the predicted p10 product ending in Met-239. For CA, at both of the PR concentrations tested, processing was incomplete, yielding CA-SP, i.e., CA with the spacer peptide attached to its C terminus. None of the mature CA species found in virions was detected. In parallel experiments we also investigated in vitro processing of the shorter Gag product CA-SP-NC. This species also was converted to CA-SP under standard digestion conditions. However, it could be converted to mature CA after prolonged treatment with higher doses of protease. Table 2 shows the results from a study in which 150 μg of CA-SP-NC was treated with 3 μg of PR for 2 h or with 9 μg of PR for 18 h. The mature CA forms with masses of 25,815 Da and 25,543 Da, resulting from cleavage at Ala-476 and Met-479, were detected only at the higher protease con-

TABLE 2. Analysis of in vitro-cleaved Gag products by ESI-MS

Protein	Position	Mass (Da)	
		Calculated	Measured (in vitro cleaved)
P60 cleavage data			
MA[a]	1–155	16,532.1	16,532
p10	178–239	6,165.7	6,166
NC	489–577	9,464.6	9,461
CA-SP	240–488	26,756.8	26,756
MA-p2a	1–164	17,389.1	17,388
MA-p2a-p2b (p23)	1–177	18,599.4	18,598
MA-p2a-p2b-p10 (p35)	1–239	24,749.2	24,747
P60	1–577	60,934.9	60,938
CA-SP-NC cleavage data			
CA (low protease concn)	240–488	26,756.8	26,754
CA (high protease concn)	240–479	25,815.9	25,815, 25,543
CA-SP-NC	240–577	36,203.0	36,200

[a] The observed mass for MA implies that the cleavage product starts with a free methionine at position 1. The calculated mass is based on this conclusion.

centration and longer incubation time. There was no evidence of the CA moiety at Ala-478 under any of the cleavage conditions tested.

For MA, as expected, all the protein generated by in vitro cleavage was present in a nonacetylated form with no evidence for phosphorylation. The slight difference in the HPLC profiles of natural and recombinant MA presumably reflects the effect of acetylation. Molecular masses for the other MA-containing cleavage products indicate cleavages at Asn-164, Gly-177, and Met-239. These species thus correspond to MA-p2a, MA-p2a-p2b, and MA-p2a-p2b-p10, respectively (Table 2). We infer that MA-p2a-p2b and MA-p2a-p2b-p10 account for the p23 and p35 species that are often detected in virus and after digestion of the Gag precursor with PR (28, 29, 30). The molecular mass of MA-p2a indicates that Asn-164–Cys-165 is a specific PR cleavage site and supports our previous hypothesis that p2a ends with Asn-164 (19). In this earlier study we had been unable to account for the Cys-165–Ala-166 junction peptide that is absent from both p2a and p2b in the mature virus.

To further investigate processing of p2 we tested the ability of PR to cleave the synthetic peptide Ile-Gly-Cys-Asn-Cys-Ala-Thr-Ala-Ser-Ala, which spans the p2a-p2b cleavage site. Twenty micrograms of the peptide was incubated at 37°C for 40 h with 4 μg of AMV PR in 100 μl of 50 mM sodium acetate (pH 5.5)–900 mM NaCl. The reaction mixture was desalted on a Superdex peptide HR10/30 column (Pharmacia), and the peptide-containing fractions were analyzed by ESI-MS (data not shown). The peptide proved to be an excellent substrate for PR, being cleaved at 10 times the rate of the CA-SP peptides that we had tested previously (20). As expected, the major products in this reaction were the tetrapeptide and hexapeptide predicted by cleavage after Asn-164. However, in different experiments, between 30 and 50% of the peptide was cleaved instead between the residues corresponding to Ala-166 and Thr-167. Thus, the peptide contains two potential cleavage sites for PR, separated by two residues. Since the alternative product was not detected in the analysis of products of in vitro cleavage of the recombinant P60 protein, we infer that structural constraints in the protein prevent access of PR to one of the sites.

Together these data suggest the model that in vivo PR first cleaves at the Asn-164–Cys-165 site to generate the p2a previously characterized in virions (19) and a p2b-related peptide with Cys-Ala attached at its N terminus. According to this model, Cys-Ala is then removed by an aminopeptidase to generate the mature p2b species. PR presumably is incapable of removing the dipeptide since it has no exopeptidase or dipeptidase activity. While there are no published data directly bearing on the possibility that RSV particles carry host proteases, an aminopeptidase activity has been reported in murine leukemia virus, providing a precedent for this possibility in the avian system (31). An alternative model to account for the results is that cleavage in the maturing virion takes place at either of two sites, after Asn-164 or after Ala-166, as suggested by the results of cleavage of the peptide described above. This would lead to two different p2a species and two different p2b species. This model would then require that for unknown reasons in the analysis of these peptides reported previously (19) only one of each of the pairs of peptides was recovered. In any case, the fact that similar yields of the p2 peptides were obtained from PrC RSV and from AMV indicates that processing in this region is the same in diverse ASLV.

Many studies on the cleavage of synthetic peptides by retroviral proteases have been reported (for example, see reference 3 for references on ASLV). Our evaluation of the p2a-p2b junction peptide adds to the list of ASLV Gag cleavage sites tested in peptide. In its ability to be processed at either of two adjoining sites, separated by two amino acid residues, the sequence represented in this peptide is analogous to that at the C terminus of CA. Perhaps two closely spaced sites reflects evolutionary pressure to ensure cleavage in a particular region.

In the analysis of peptide substrates as models for processing sites in the Gag precursor, the p2b-p10 junction peptide in ASLV was found to act as a competitive inhibitor of PR, binding tightly to the enzyme but being hydrolyzed slowly (3). These results led the investigators to speculate that this cleavage site might play a role in the regulation of cleavage in vivo and, in particular, in the slow maturation of MA. On the basis of these observations, one might predict a build-up of the p2b-p10 species in our experiments with Gag protein processed in vitro. In fact, we found no evidence for such a cleavage product, even at a low PR concentration. Discordance between the cleavage of a sequence in a peptide and its cleavage in a protein is not surprising, however, since the folding of the protein may either prevent or enable access of PR to a site.

The discrepancy in the molecular masses of p10 generated in vitro and p10 found in virions is highly surprising, although it is consistent with a previous analysis of this protein (18). The 6,166 Da molecular mass for in vitro cleaved p10 revealed that Met-239–Pro-240 is the PR cleavage site between p10 and CA. The fact that the predominant form of p10 in virions is missing the predicted C-terminal methionine residue suggests strongly that a carboxypeptidase is incorporated into virions. It is provocative that the minor component of CA ending at Ala-478 also is missing a C-terminal methionine residue, compared with the more common component ending at Met-479, and that both have an Ala residue preceding Met. Thus we speculate that both proteins are generated by the same carboxypeptidase activity. The fact that both AMV and RSV, which were derived from very different cell types, showed similar amounts of the des-methionine forms of p10 and CA suggests that incorporation of this putative carboxypeptidase is a common feature in ASLV.

The ESI-MS data presented here refine the currently accepted structural predications for ASLV viral Gag proteins and provide a glimpse of the structures of various intermediate and minor Gag proteins that have been observed in virions or in in vitro processing studies but whose structures were unknown. The data reveal unexpected features of proteolytic processing that should stimulate further studies. The accuracy of modern mass spectrometers has revolutionized the ability to evaluate protein structure. Current systems can detect an error of less than 1 Da for a protein of 150 amino acids and are useful for analyzing the masses of proteins up to 200 kDa. The ability to rapidly test structure by MS will simplify the means by which critical size data are obtained, allowing for more highly controlled structure-function analysis.

REFERENCES

1. **Bennett, R. P., S. Rhee, R. C. Craven, E. Hunter, and J. W. Wills.** 1991. Amino acids encoded downstream of *gag* are not required by Rous sarcoma virus protease during Gag-mediated assembly. J. Virol. **65:**272–280.
2. **Bhown, A. S., J. C. Bennett, and E. Hunter.** 1980. Alignment of the peptides derived from acid-catalyzed cleavage of an aspartylprolyl bond in the major internal structural polypeptide of avian retroviruses. J. Biol. Chem. **255:** 6962–6965.
3. **Cameron, C. E., B. Grinde, J. Jentoft, J. Leis, I. T. Weber, T. Copeland, and A. Wlodawer.** 1992. Mechanism of inhibition of the retroviral protease by a Rous sarcoma virus peptide substrate representing the cleavage site between the gag p2 and p10 proteins. J. Biol. Chem. **267:**23735–23741.
4. **Campbell, S., and V. Vogt.** 1995. Self-assembly in vitro of purified CA-NC proteins from Rous sarcoma virus and human immunodeficiency virus type 1. J. Virol. **69:**6487–6497.

4a. **Campbell, S., and V. Vogt.** Unpublished data.

5. **Craven, R. C., A. E. Leure-duPree, C. R. Erdie, C. B. Wilson, and J. W. Wills.** 1993. Necessity of the spacer peptide between CA and NC in the Rous sarcoma virus Gag protein. J. Virol. **67:**6246–6252.

6. **Erikson, E., J. Brugge, and R. L. Erikson.** 1977. Phosphorylated and nonphosphorylated forms of avian sarcoma virus polypeptide p19. Virology **80:** 177–185.
7. **Henderson, L. E., M. A. Bowers, R. C. Sowder, S. A. Serabyn, D. G. Johnson, J. W. Bess, 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 processing, and complete amino acid sequences. J. Virol. **66:**1856–1865.
8. **Henderson, L. E., T. D. Copeland, R. C. Sowder, A. M. Schultz, and S. Oroszlan.** 1988. Analysis of proteins and peptides from sucrose banded HTLV-III, p. 135–147. *In* D. Bolognesi (ed.), Human retroviruses, cancer and AIDS: approaches to prevention and therapy. Alan R. Liss Inc., New York.
9. **Hunter, E., J. C. Bennett, A. Bhown, R. B. Pepinsky, and V. M. Vogt.** 1983. Amino-terminal amino acid sequence of p10, the fifth major *gag* polypeptide of avian sarcoma and leukemia viruses. J. Virol. **45:**885–888.
10. **Johnson, S. P., M. Veigl, T. Vanaman, and J. Leis.** 1983. Cyanogen bromide digestion of the avian myeloblastosis virus pp19 protein: isolation of an amino terminal peptide that binds to viral RNA. J. Virol. **45:**876–881.
11. **Lai, M. M. C.** 1976. Phosphoproteins of Rous sarcoma viruses. Virology **74:** 287–301.
12. **Leis, J., and J. Jentoft.** 1983. Characteristics and regulation of interaction of avian retrovirus pp12 protein with viral RNA. J. Virol. **48:**361–369.
13. **Leis, J., S. Johnson, L. S. Collins, and J. A. Traugh.** 1984. Effects of phosphorylation of avian retrovirus nucleocapsid protein pp12 on binding of viral RNA. J. Biol. Chem. **259:**7726–7732.
14. **Leis, J., N. Phillips, X. Fu, P. T. Tuazon, and J. A. Traugh.** 1989. Phosphorylation of avian retrovirus matrix protein by Ca^{2+}/phospholipid-dependent protein kinase. Eur. J. Biochem. **179:**415–422.
15. **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.
16. **Misono, K. S., F. S. Sharief, and J. P. Leis.** 1980. Primary structure of Rous sarcoma virus RNA-associated polypeptide pp12. Fed. Proc. Exp. Biol. Abstr. **60:**1611.

16a. **Nelle, T., and J. Wills.** Submitted for publication.

17. **Palmiter, R. D., J. Gagnon, V. M. Vogt, S. Ripley, and R. Eisenman.** 1978. The NH2-terminal sequence of an avian oncovirus gag precursor polyprotein (Pr76gag). Virology **91:**423–433.
18. **Pepinsky, R. B.** 1983. Ph.D. thesis. Cornell University, Ithaca, N.Y.
19. **Pepinsky, R. B., R. J. Mattaliano, and V. M. Vogt.** 1986. Structure and processing of the p2 region of avian sarcoma and leukemia virus *gag* precursor polyproteins. J. Virol. **58:**50–58.
20. **Pepinsky, R. B., I. A. Papayannopoulos, E. P. Chow, N. K. Krishna, R. C. Craven, and V. M. Vogt.** 1995. Differential proteolytic processing leads to multiple forms of the CA protein in avian sarcoma and leukemia viruses. J. Virol. **69:**6430–6438.
21. **Pepinsky, R. B., and V. M. Vogt.** 1983. Purification and properties of a fifth major viral Gag protein from avian sarcoma and leukemia viruses. J. Virol. **45:**648–658.
22. **Pepinsky, R. B., and V. M. Vogt.** 1984. Fine-structure analyses of lipid-protein and protein-protein interactions of gag protein p19 of the avian sarcoma and leukemia viruses by cyanogen bromide mapping. J. Virol. **52:** 145–153.
23. **Pettit, S. C., M. D. Moody, R. S. Wehbie, A. H. Kaplan, P. V. Nantermet, C. A. Klein, and R. Swanstrom.** 1994. The p2 domain of human immunodeficiency virus type 1 *gag* regulates sequential proteolytic processing and is required to produce fully infectious virions. J. Virol. **68:**8017–8027.
24. **Sauer, R. T., D. W. Allen, and H. D. Niall.** 1981. Amino acid sequence of p15 from avian myeloblastosis virus complex. Biochemistry **20:**3784–3791.

24a. **Schatz, G., I. Pichova, and V. Vogt.** Unpublished data.

25. **Schwartz, D., R. Tizard, and W. Gilbert.** 1983. Nucleotide sequence of Rous sarcoma virus. Cell **32:**853–869.
26. **Shealy, D. J., A. G. Mosser, and R. R. Reukert.** 1980. Novel p19-related protein in Rous-associated virus type 61: implications for avian *gag* gene order. J. Virol. **34:**431–437.
27. **Tritch, R. J., Y. E. Cheng, F. H. Yin, and S. Erickson-Viitanen.** 1991. Mutagenesis of protease cleavage sites in the human immunodeficiency virus type 1 Gag polyprotein. J. Virol. **65:**922–930.
28. **Vogt, V. M., R. B. Pepinsky, and L. E. Southard.** 1985. The primary structure of p19 species of avian sarcoma and leukemia viruses. J. Virol. **56:**31–39.
29. **Wills, J. W., C. E. Cameron, C. B. Wilson, Y. Xiang, R. P. Bennett, and J. Leis.** 1994. An assembly domain of the Rous sarcoma virus Gag protein required late in budding. J. Virol. **68:**6605–6618.
30. **Wills, J. W., R. C. Craven, and J. A. Achacoso.** 1989. Creation and expression of myristylated forms of Rous sarcoma virus Gag proteins in mammalian cells. J. Virol. **63:**4431–4343.
31. **Yoshinaka, Y., C. Kirk, and R. B. Luftig.** 1980. Aminopeptidase activity associated with purified murine leukemia viruses. J. Gen. Virol. **46:**363–372.