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

R. BLAKE PEPINSKY[†] AND VOLKER M. VOGT*

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, New York 14853

Received 3 April 1984/Accepted 21 June 1984

In avian sarcoma and leukemia viruses, the *gag* protein p19 functions structurally as a matrix protein, connecting internal components with the viral envelope. We have used a combination of in situ cross-linking and peptide mapping to localize within p19 the regions responsible for two major interactions in this complex, p19 with lipid and p19 with p19. Lipid-protein cross-links were localized near the amino terminus within the first 35 amino acids of the polypeptide. Homotypic protein-protein disulfide bridges were found to originate from near the carboxy terminus of p19, from cysteine residues at amino acids 111 and 153. These results suggest that p19 is divided into domains with distinct functions. The peptide maps constructed for p19, and for the related proteins p23 in avian sarcoma and leukemia viruses and p19 β in recombinant avian sarcoma viruses, should serve as useful tools for other types of studies involving these proteins.

The major class of proteins in avian sarcoma and leukemia viruses (ASLVs) consists of the five internal structural proteins p27, p19, p15, p12, and p10. They account for 80% of the total viral protein and form the internal structure that is detected in thin sections by electron microscopy. During assembly these polypeptides are processed proteolytically from a single large precursor polypeptide that is encoded by the *gag* gene (9). In mature ASLVs, p19 is associated with the viral envelope (20, 24); p10 is within the space between the membrane and core (25); p27 forms the core shell (2, 34); and p12 is within the core associated with viral RNA in a ribonucleoprotein complex (2, 8). The location of p15 is unknown. The observed homotypic associations of p19, p27, and p12 in virus particles (23) are consistent with structural models in which a virus is constructed of concentric shells of protein (3). In virus particles, p19-p19 associations are stabilized covalently by intermolecular disulfide bridges (23) (see below). Since p19-p19 cross-linked oligomers also are detected as prominent species in lipid-protein cross-linking studies (24), both lipid and protein interactions of p19 represent a single population of molecules.

In addition to acting as a matrix protein, p19 interacts with genomic RNA. A small fraction of the ca. 2,000 p19 molecules is reported to be associated with genomic RNA in virions (30). In vitro, purified p19 binds with high affinity to its genomic RNA but not to other types of RNA (16). This binding appears to be specific for double-stranded regions of RNA (16). These observations have led to models in which p19 regulates diverse activities such as splicing (17), RNA packaging (7), and translation (7). Thus p19 is potentially a key component in gene expression and assembly of the virus. In this study, we have used a combination of cross-linking and peptide mapping analysis (22) to localize within the polypeptide backbone of p19 the sites of lipid-protein and protein-protein interactions. The fine-structure analyses presented suggest that the different functions reside within different domains of the protein.

MATERIALS AND METHODS

Cells and viruses. Chicken embryo fibroblasts infected with avian myeloblastosis virus (AMV) or with the Prague C strain of Rous sarcoma virus (PrC) were grown as monolayers at 39°C in Dulbecco modified Eagle medium containing antibiotics (penicillin, 50 U/ml; streptomycin, 50 μ g/ml) and 10% fetal calf serum. Quail cell lines transformed with avian sarcoma viruses (SE21Q1B, PrE21-21) were grown under similar conditions, except that the growth medium was supplemented with 2% heat-inactivated chicken serum, 1% dimethyl sulfoxide, and twice the normal dosage of vitamins. The SE21Q1b $\times$ Rous-associated virus type 1 (RAV-1) virus was obtained from M. Linial. The ev-1 virus was from fibroblasts of a chicken embryo spontaneously turned on for ev-1 (obtained from R. Eisenman). These cells were infected with PrC and thus were producing a mixture of virus particles with distinguishable p19 forms. Procedures for isolation of viruses on isopycnic sucrose gradients were described previously (24). Preparation of radioactive virus was also as described previously (24), except that during labelings one of the following formulations was substituted for complete medium. For tritium and ¹⁴C-labeled amino acids, virus-producing cells were labeled for 5 h in a modified Eagle minimum essential medium prepared without the amino acid used for labeling. For ³²P_i, virus-producing cells also were labeled to minimal medium, in which the phosphate and serum concentrations were reduced to 0.05 mM and 2.5%, respectively. For preparing virus of high specific activity, cells were pulse-labeled with [³⁵S]methionine or [³⁵S]cysteine. In the latter case, monolayers (100-mm dishes) were washed twice with phosphate-buffered saline and then incubated at 39°C for 30 min with periodic shaking in 1.5 ml of Earle balanced salt solution containing 500 μ Ci of the radioactive amino acid per ml (ca. 600 Ci/mmol). Then complete medium was added, and virus collections were performed as previously described.

SDS-PAGE. Viral proteins were resolved by electrophoresis in sodium dodecyl sulfate (SDS)-polyacrylamide gels (15% acrylamide, 0.18% methylene bisacrylamide) by the procedure of Laemmli (15). Stacking gels contained 7.6%

* Corresponding author.

[†] Present address: Biogen, Cambridge, MA 02142.

acrylamide and 0.21% methylene bisacrylamide. Before electrophoresis, samples were disrupted in electrophoresis sample buffer (2% SDS, 0.05 M Tris-hydrochloride [pH 6.8], 12.5% glycerol, 1.5% 2-mercaptoethanol). Unlabeled proteins were visualized by staining with Coomassie brilliant blue. Radioactive bands were visualized by fluorography (5). To locate bands to be excised and subsequently treated with cyanogen bromide (CNBr), one of several techniques that do not involve fixation was used. For ^{35}S - and ^{32}P -labeled proteins, radioactive bands were visualized by autoradiography of the wet gel at 4°C. For nonradioactive proteins and for tritiated proteins, bands were visualized either by precipitation in ice-cold 0.5 M KCl plus 0.1% 2-mercaptoethanol (12) or by rapid surface staining of parallel tracks with Coomassie blue (6). The latter procedure was used in instances in which the precipitation procedure was not sensitive enough.

CNBr mapping. Details of the CNBr mapping technique have been described previously (22). Briefly, several gel slices containing viral protein were cut orthogonal to the protein band and either stored at -20°C or directly incubated with CNBr for 1 h at 23°C in 0.1 N HCl plus 0.2% 2-mercaptoethanol. Gel slices were then washed, and the cleavage products were subjected to SDS-urea polyacrylamide gel electrophoresis (PAGE) (22, 35). Cleavage profiles were visualized by silver staining (21) for unlabeled protein or by fluorography (5) for radioactive bands. Unless noted otherwise, cleavage of p19 and its related proteins was performed with 10 mg of CNBr per ml.

Preparation of purified p19. AMV from leukemic chicken

plasma (10 mg of protein per ml) was incubated at 0°C for 30 min in buffer (50 mM Tris-hydrochloride [pH 7.5], 100 mM NaCl, 1 mM EDTA, 10 mM dithiothreitol [DTT]) containing 1% Triton X-100. The sample was diluted 10-fold with ice-cold 20 mM Tris-hydrochloride (pH 7.5) and subjected to phosphocellulose chromatography as described previously (25) in the presence of 1 mM phenylmethylsulfonyl fluoride and 1 mM DTT. Approximately 70% of the loaded protein was retarded by this procedure. The resin was then washed with 5 column volumes of column buffer (20 mM Tris-hydrochloride [pH 7.5], 0.1% Triton X-100, 1 mM DTT) and then with 2 volumes of this buffer containing 100 mM NaCl, and finally p19 was eluted with a salt step containing 350 mM NaCl. p19 and its related proteins were the only viral structural proteins recovered by this procedure (Fig. 1). In instances in which it was necessary to separate phosphorylated and unphosphorylated forms of p19 and p23, i.e., for two-dimensional CNBr mapping, preparations were subjected to DEAE-cellulose chromatography (DE52 cellulose from Whatman Ltd). We have observed that in buffer (20 mM Tris-hydrochloride [pH 7.5], 1 mM EDTA, 0.1% Triton X-100) containing 5 mM DTT, only phosphorylated p19 binds to the resin, whereas in the absence of DTT both forms of p19 are retained. Phosphorylated p19 can be eluted with buffer containing 100 mM NaCl.

RESULTS

Construction of a peptide map for p19. The protein component of ASLVs is made up predominantly of structural proteins that are derived from the *gag* and *env* genes. After

FIG. 1. Partial peptide mapping analysis of *gag* proteins of ASLVs. Preparations of ASLVs were disrupted in electrophoresis sample buffer, and proteins were subjected to SDS-PAGE. Profiles were visualized directly by fluorography (a, d, and e) or by staining with Coomassie blue (b, c, and f). In lanes g to m, individual proteins were further analyzed by CNBr mapping (22). Gel slices containing the specified proteins were incubated with 10 mg of CNBr per ml and washed, and cleavage fragments were electrophoresed into SDS-urea gels. Cleavage products were visualized by fluorography (g to j) or by silver staining (k to m). Lanes: a, [^3H]tryptophan-labeled AMV; b, unlabeled AMV (25 μg); c, p19-enriched fraction from AMV after phosphocellulose chromatography; d, [^{35}S]cysteine-labeled PRE21-21; e, p19 β -enriched fraction from [^{35}S]cysteine-labeled PRE21-21 after phosphocellulose chromatography; f, unlabeled SE21Q1b (25 μg); g, PrC-p27; h, PrC-p19; i, PrC-p23; j, PrC-p12; k, AMV-p23; l, AMV-p19; m, AMV-p19f.

SDS-PAGE, these products form a characteristic pattern of bands (Fig. 1a). The five *gag* proteins (p27, p19, p15, p12, p10) and two *env* proteins (gp85, gp37) are denoted at the left of the panel. When gel slices containing individual ASLV *gag* proteins are treated with CNBr and the cleavage products are resolved by SDS-urea PAGE, unique cleavage profiles are observed for each of the *gag* proteins. Lanes g to j (Fig. 1) display profiles for p27, p19, p15, and p12, respectively. The differences in the patterns presented illustrate the usefulness of the CNBr profiles for identification and characterization of the individual proteins. Here p10 is not shown since it is not cleaved by CNBr.

The cleavage profile of p19 consists of six fragments. The relative intensities of these products vary as a function of CNBr concentration (Fig. 2b to d) and as a function of reaction time (data not shown), but no additional cleavage products are detected. In the experiment in Fig. 2, after a 1-h incubation with 3 mg of CNBr per ml (lane b), ca. 20% of the p19 molecules were cleaved at least once by the reagent, whereas at 10 mg of CNBr per ml (lane c), more than 90% of the molecules were cleaved. This range spans the useful CNBr concentrations for detecting all six cleavage products simultaneously. At 30 mg of CNBr per ml (lane d), all of the p19 was cleaved at least once. In our analysis of p19, the appearance of the large cleavage products as a double image results from the slight electrophoretic difference between the phosphorylated and unphosphorylated forms of p19 (Fig. 1a).

FIG. 2. Localization of phospholipid cross-link to a 4-Kd fragment at the amino terminus of p19. Samples containing unlabeled AMV (10 μ g), [14 C]ethanolamine-labeled AMV that was cross-linked with 0.25 mg of dimethyl suberimidate per ml for 2 h at 23°C (3×10^5 dpm), or [32 P]phosphate-labeled PrC (1×10^6 dpm) were disrupted in electrophoresis sample buffer and subjected to SDS-PAGE. Gel slices containing p19 regions were excised with a razor blade, incubated with CNBr for 1 h at 23°C, and washed, and cleavage fragments were electrophoresed into SDS-urea gels. Profiles were visualized by silver staining (a to d) or by fluorography (e to h). Lanes: a to d, unlabeled p19; e and f, [32 P]phosphate-labeled p19; g and h, cross-linked p19-[14 C]phosphatidylethanolamine complexes. Lanes: a, no CNBr; b, e, and g, 3 mg of CNBr per ml; c, f, and h, 10 mg of CNBr per ml; d, 30 mg of CNBr per ml. Fragmentary compositions indicated at the left are based on the mapping analysis presented in Fig. 3.

To determine the fragment compositions of the six products derived from p19, we applied a two-dimensional mapping analysis described in detail previously for the *gag* protein of Gazdar murine sarcoma virus (22). The mapping technique is based on nearest-neighbor analysis of fragments that are released by CNBr treatment of partial cleavage products in gel slices. As shown by the silver-stained gel in Fig. 3A and by the diagrammatic representation of this gel in Fig. 3C (filled circles), the mapping technique distinguishes complete cleavage fragments (e and g) from partial cleavage fragments (b, c, d, and f) by their resistance to further cleavage in a subsequent treatment with CNBr. In this instance, however, the two-dimensional profile alone is not sufficient for determining fragment order, since two missing cleavage products generate ambiguities in the analysis. The fact that two separate fragments are lost is clearly illustrated by the cleavage profiles derived from spots c and d. Both partials release only fragment e. The size differences between e and d (2.3 kilodaltons [Kd]) on the one hand and between e and c (3.5 Kd) on the other suggest that the missing fragments may be simply too small to be detected by SDS-urea PAGE.

One method to distinguish these small fragments without visualizing them directly is to modify one of them and then determine the effect of this change on the cleavage profile. The expectation is that the modification should alter only the one complete fragment and the corresponding partials that contain this fragment. This type of analysis by difference has been documented previously (22). Rather than chemically modifying p19, we relied on analysis of the p19-related protein p23 that is found in virus particles. p23 is a minor cleavage product of the *gag* precursor (31). It contains all of the methionine tryptic peptides of p19, including the peptide derived from the N-terminus (unpublished data; R. Eisenman, personal communication) and thus presumably differs from p19 by a stretch of amino acids at its carboxy terminus. This region of the *gag* sequence does not contain any methionine residues (29), and thus the modification (i.e., elongation at the C-terminus) should alter the cleavage profile only for those products that contain the C-terminal fragment of p19.

As shown in Fig. 1k, the cleavage profile of p23 contains five fragments that are identical to the ones obtained from p19 (lane 1), two new fragments, and one missing fragment. To determine the composition of these fragments, two-dimensional CNBr mapping also was performed on p23. In this analysis (Fig. 3B and D), the cleavage profiles for the five fragments common to p19 and p23 (b, d, e, f, and g) are identical to ones obtained for the corresponding fragments of p19, whereas profiles of the new fragments specific to p23 (denoted with prime marks) are unique to p23. The two p23-specific fragments, c' and b', which are related to each other, presumably contain the modified C-terminus. By similar reasoning, spot c of p19 must be a partial that contains the C-terminus of this protein. Since spots b and d are not affected by the modification, we infer that they contain the other missing cleavage fragment.

Together the two-dimensional analyses of cleavage fragments from p19 and from p23 allowed us to construct peptide maps (Fig. 4). In these and in all subsequent analyses, we adopt a simple nomenclature in which fragment numbers correspond to the relative positions of polypeptides from the amino terminus. Since p19 generates four complete cleavage fragments, fragment 1 represents the polypeptide that contains the amino terminus of p19 and fragment 4 represents the fragment containing the carboxy terminus. The deriva-

FIG. 3. Two-dimensional electrophoretic analysis of CNBr fragments derived from p19 and from p23. Gel slices containing 3 μ g of p19 or p23 from AMV were treated with 8 mg of CNBr per ml for 1 h at 23°C and washed, and cleavage fragments were subjected to SDS-urea PAGE. The entire gel strips were then treated with 15 mg of CNBr per ml for 1 h at 23°C and washed, and fragments were electrophoresed into SDS-urea slab gels. A and B display silver stains of two-dimensional profiles for p19 and p23, respectively. C (p19) and D (p23) are schematic representations of the same data, in which the missing fragments (denoted by open circles) are drawn at positions roughly consistent with their molecular weights. Tracks of cleavage products are marked by a letter code. At the left of the figure are fragmentary compositions of cleavage products based on nearest-neighbor analysis.

tion of the cleavage fragments is presented with respect to the actual positions of methionines in the predicted amino acid sequence of p19 (29) (Fig. 4A). The composition of the fragments seen in one-dimensional cleavage profiles is summarized (Fig. 4B). Below we illustrate by two methods the consistency in this analysis.

First, peptide order can be derived directly from the ladder of cleavage fragments that are specific to p23, since each of these products has a fixed carboxy terminus. When analyzed sequentially, the difference between each pair of partials yields a fragment family that is derived from the amino terminus of the larger partial. Thus the difference between c' (4' + 3) and b' (4' + 3 + 2) defines fragment 2-containing products, and the difference between b' and a' (intact p23, 4' + 3 + 2 + 1) defines fragment 1-containing products. As observed with p19, the C-terminal fragment of p23 is not itself detected by silver staining. However, using [35 S]cysteine labeling of p23, which results in a very large apparent enrichment for fragment 4' and its partials, we have detected radioactive spots near the top of the gel that are not

visible by staining or with other radioactive amino acids. We infer these to correspond to fragment 4', since they are released only from cleavage fragments specific to p23 (data not shown). The nature of the aberrant electrophoretic mobility was not determined. The peptide compositions of other cleavage fragments are consistent with the assignment of the p23-specific fragments. In these profiles, spot d releases fragments 2 and 3, and spot b releases fragments 1, 2, and 3. Finally, the analysis predicts that g is the amino-terminal fragment, since it is detected only in the largest of the p23-specific products, p23 itself, implying that it is the most distal fragment to fragment 4'. Supporting this identification are the observations that it is released from p19f (Fig. 1m), a protein fragment shown previously to be derived from the amino-terminal half of p19 (32), and that it is released from spot g in Fig. 3A and B. The latter fact, which implies that the two small fragments are next to each other, is supported by the predicted amino acid sequence of p19 (29).

Alternatively, cleavage fragments can be mapped by nearest-neighbor analysis. To clarify the logic underlying this

FIG. 4. Schematic summary of CNBr cleavage data. The distribution of methionines in p19 based on nucleic acid sequencing is summarized in A. The five methionines yield only four fragments that are detected by the gel analysis; 1 (4.3 Kd), 2 (2.3 Kd), 3 (8.4 Kd), and 4 (3.5 Kd). The apparent molecular weights of fragments 2 and 4 are based on size differences of partials since the complete cleavage products are not observed. Similar data are also shown for p23. The single p23-specific fragment is denoted 4'. In B these data are summarized with respect to the electrophoretic displays of cleavage fragments. Lanes: a, schematic representation of p19 profile; b, p19 cleavage profile; c, p23 cleavage profile; d, schematic representation of p23 profile. The faint spot between fragment 2 + 3 and fragment 1 + 2 + 3 in lane c, which can also be seen in Fig. 1, results from a 3 + 4' partial with an intrachain disulfide bond that forms in the gel (see equivalent spot above the diagonal in Fig. 3B).

analysis, in the schematic profiles shown in Fig. 3C (p19) and D (p23), we have added back the two missing fragments (denoted by open circles) at positions where they might be expected to migrate based on their relative sizes. Thus in Fig. 3D, the three partials f (1 + 2), d (2 + 3), and c' (3 + 4) that release two cleavage fragments automatically define a peptide order of 1-2-3-4. The cleavage profiles of the two partials, b (1 + 2 + 3) and b' (2 + 3 + 4), that release three cleavage fragments are consistent with this assignment. In Fig. 3C, as described above, spots f, d, and c automatically define peptide order. However, the two partials that release three fragments comigrate, thus generating an ambiguity, which contributed to our original inability to order fragments from the profile of p19 alone.

Although the order of the two small amino-terminal fragments (1 and 2) is obvious from the analysis, the final location of the stretch of 11 amino acids (between methionine residues at positions 29 and 40) that separates these fragments is uncertain. To address this question, we characterized the peptide compositions of fragments 1 and 1 + 2 (Fig. 4) by analyzing a series of metabolically labeled p19 molecules, each containing radioactive label that is distributed differently in the subset of cleavage fragments. The expected distribution of five amino acids tested, glycine, cysteine, leucine, arginine, and tryptophan, with respect to the five complete CNBr peptides of p19 is tabulated in Table 1. The letter δ corresponds to the stretch of 11 amino acids separating fragments 1- δ and 2.

When gel slices containing the various radiochemically labeled p19 species were treated with CNBr and the cleavage products were resolved by SDS-urea PAGE, we obtained the profiles shown in Fig. 5. Lane a, [^{14}C]glycine-labeled p19, is representative of uniformly labeled protein, since all fragments contain glycine. The expected fragmentary composition of the spots is indicated at the left. Lanes b through e contain labeled p19 species that correspond to the other radioactive amino acids tested. In these profiles, cysteine (lane b) was used to monitor 1- δ ; leucine (lane c) was used to monitor δ , although it labels fragment 2 as well; and tryptophan (lane d) and arginine (lane e) were used to monitor fragment 2. The facts that leucine labels both of the visible low-molecular-weight bands whereas tryptophan and arginine label only the larger of these imply that the δ region is contained in both of these two bands. One possible explanation for the δ region remaining with fragment 1 is that the methionine that separates these fragments is not very reactive to CNBr under the conditions used. Similar variations in reactivity were observed in previous studies (22).

p19-related proteins from recombinant avian sarcoma virus. In addition to p19 from AMV and from PrC, we have analyzed by CNBr mapping the corresponding proteins from three recombinant viruses (18, 25, 31), SE21Q1b, PRE21-21, and SE21Q1b $\times$ RAV-1, and from an endogenous virus derived from locus ev-1. In all four of these preparations, p19 is replaced by a related protein termed p19 β , which is similar in function to p19 (24, 31) but appears 4,000 daltons smaller by SDS-PAGE (cf. Fig. 1a and f). We have shown elsewhere that the actual difference between p19 and p19 β is not size but rather a charge difference due to a single amino acid substitution (V. M. Vogt, R. B. Pepinsky, and L. Southard, manuscript in preparation). All the viruses with p19 β also contain a protein termed p19 α , which corresponds to p23.

One of the predicted amino acid differences between p19 and PrC and p19 β in the ev-1 virus, a methionine at amino acid 144 in p19 β instead of a threonine found in p19 (Vogt et

TABLE 1. Predicted distribution of amino acids based on sequence data

Amino acids ^a	Distribution with respect to:				
	1- δ	δ	2	3	4
Total	29	11	23	75	16
Glycine (a)	2	1	1	11	1
Cysteine (b)	2	0	0	1	1
Leucine (c)	0	5	2	7	0
Tryptophan (d)	0	0	1	2	0
Arginine (e)	0	0	1	5	0

^a The letters within parentheses correspond to the lanes in Fig. 5.

FIG. 5. Partial peptide mapping analysis of metabolically labeled p19 molecules. Preparations of AMV that were metabolically labeled with radioactive amino acids were subjected to SDS-PAGE. The p19 regions were visualized by precipitation with ice-cold KCl and excised with a razor blade. Gel slices were incubated with 10 mg of CNBr per ml for 1 h at 23°C and washed, and cleavage fragments were electrophoresed into SDS-urea gels. Cleavage products were visualized by fluorography, and the cleavage profiles obtained are shown. The fragmentary compositions denoted at the left are based on the nomenclature used throughout the text. Lanes: a, [^{14}C]glycine-p19; b, [^{35}S]cysteine-p19; c, [^3H]leucine-p19; d, [^3H]tryptophan-p19; e, [^3H]arginine-p19. The predicted distribution of amino acids based on sequence data is shown in Table 1 (29).

al., manuscript in preparation), should produce a new CNBr cleavage site near the C-terminus of the protein, splitting fragment 4 into two smaller peptides. Cleavage at this site should yield an additional family of fragments that contain the seven amino acids from the amino-terminal half of fragment 4, a region we refer to as ϵ . Cleavage profiles of p19 β from SE21Q1b and ev-1 viruses support this prediction (Fig. 6c and d). Both profiles contain eight cleavage fragments, six that are characteristic of p19 and two new fragments, denoted by arrows, which correspond in size to ones predicted by the ev-1 DNA sequence. The smaller of these fragments, located between fragments 3 and 3 + 2, presumably contains fragments 3 and ϵ , whereas the larger product, between fragments 1 + 2 + 3 and 1 + 2 + 3 + 4, probably contains fragments 1, 2, 3, and ϵ . A third new fragment, which is resolved on two-dimensional gels (SDS-PAGE followed by SDS-urea PAGE), comigrates with fragments 1 + 2 + 3 and 2 + 3 + 4 by a one-dimensional analysis (data not shown). This product presumably corresponds to 2 + 3 + ϵ . Thus the data presented for p19 β further confirm our alignment of cleavage fragments with the known DNA sequence. The p19 β from PRE21-21 virus yielded a cleavage profile like the two described above (data not shown).

A third cleavage pattern was obtained for p19 β from SE21Q1b $\times$ RAV-1, an infectious virus that is a recombinant between RAV-1 and the non-infectious packaging mutant SE21Q1b (M. Linial, personal communication). This profile differs from that of other p19 β proteins studied (Fig. 6c and d) by the absence of two cleavage products, fragments 1 and 2 + 3, marked by arrows in lane e. Since generation of both

of these products relies on cleavage at the methionine between fragments 1 and 2, we infer that this amino acid has been altered. One interpretation of this alteration is that this part of the p19 sequence is derived from RAV-1 and that RAV-1 does not have a methionine residue at position 40. DNA sequencing of RAV-1 should clarify this point.

CNBr mapping of phosphorylated p19. With a detailed understanding of the makeup of each of the observed cleavage fragments of p19 (Fig. 4), one can use the one-dimensional gel profile as a map for localizing sites of modifications within the polypeptide backbone. Alterations in p19 can be mapped rapidly by comparing the cleavage profile of the modified protein with the characteristic profile of p19. A simple application of this technique, the localization of the site of phosphorylation within p19 (10), is shown in Fig. 2e and f. When gel slices containing [^{32}P]phosphate-labeled p19 from PrC were subjected to CNBr mapping, only the four cleavage fragments that contain fragment 3 were detected. Thus we infer that the site of phosphorylation is between amino acid residues 63 and 138, within fragment 3. This inference is consistent with the observation noted previously, that fragment 3 also is the source of the electrophoretic variation between phosphorylated and unphosphorylated p19. Similar results were obtained for ^{32}P -labeled p23 (data not shown). p19f, a naturally occurring protein derived from the amino-terminal half of p19 (32), also yields a double

FIG. 6. Peptide mapping analysis of p19-related proteins in ASLV. Proteins from unlabeled AMV and SE21Q1b, [^{35}S]cysteine-labeled PrC and ev-1, and [^{14}C]glycine-labeled SE21Q1b $\times$ RAV-1 were subjected to SDS-PAGE. The appropriate proteins were excised and treated with 10 mg of CNBr per ml, and cleavage products were analyzed by SDS-urea PAGE. Unlabeled spots were visualized by silver staining. Radioactive spots were visualized by fluorography. Lanes: a, p19 from AMV; b, p19 from PrC; c and h, p19 from SE21Q1b; d, p19 from ev-1; e, p19 from SE21Q1b $\times$ RAV-1. In f and g, cleavage profiles of the two disulfide-linked p19 dimers of [^{35}S]cysteine-labeled PRE21-21 are presented. To obtain these profiles, from a first-dimensional SDS-gel a longitudinal gel strip was cut containing a portion of each of the two dimer bands. The strip was treated with CNBr and then placed on top of the second SDS-urea gel. The distance between f and g thus represents the separation of the two bands in the original gel. The spot visible just to the right of lane g, between fragments 1 + 2 and 3, is due to spillover from the adjacent part of the gel.

profile (Fig. 1m) and thus contains the phosphorylated amino acid. Since the proteolytic cleavage that generates p19f from p19 must occur approximately in the middle of fragment 3, we infer that the site of phosphorylation is in the N-terminal portion of this fragment.

Localization of the site in p19 associated with membrane. In ASLVs we have shown by lipid-protein cross-linking that most of the p19 is associated with the viral membrane (24). This result was obtained by cross-linking ^{14}C -phosphatidylethanolamine-labeled AMV with the bifunctional agent dimethyl suberimidate and then identifying lipid-protein complexes from fluorographs of gels after SDS-PAGE. When gel slices containing radioactive lipid-p19 complexes were subjected to the CNBr mapping analysis, we obtained the patterns presented in Fig. 2g and h. Of the six fragments that are generated from p19, only three are labeled with lipid, fragments 1, 1 + 2, and 1 + 2 + 3. This pattern indicates an interaction with membrane that is specific for fragment 1, the amino-terminal fragment of p19. The specificity of this association with membrane can be inferred from the fact that no radioactivity is detected in fragments 3, 3 + 4, and 2 + 3. As noted previously (22), the slight decrease in electrophoretic mobility of the cross-linked fragments results from the added lipid moiety. In addition to p19, the corresponding protein from recombinant viruses, p19 β , as well as the larger forms p23 and p19 α , also become cross-linked to lipid by lipid-protein cross-linking (24); however, we have not subjected these products to the mapping analysis.

Localization of intermolecular disulfide bridges within p19-p19 dimers. For ASLVs we observed previously that most of the p19 is associated in homotypic complexes that are stabilized in situ by disulfide bridges (23). To better characterize these complexes, we repeated the studies with the following modifications. First, complexes were resolved on longer gels. Second, since p19 is rich in cysteine, the apparent concentration of p19 was enriched by metabolic labeling of the virus with [^{35}S]cysteine. Third, based on our observation that complexes derived from p19 β are better resolved than those of p19, we analyzed several recombinant avian sarcoma viruses.

When [^{35}S]cysteine-labeled PRE21-21 was subjected to two-dimensional diagonal gel electrophoresis (23), we obtained the profiles shown in Fig. 7. Panel A (Fig. 7) shows the analysis of virus on two consecutive 20-cm gels, the first dimension in the absence of reducing agent and the second dimension after a brief incubation with 2-mercaptoethanol. The majority of the p19 β molecules (ca. 80% based on densitometry) is associated in disulfide-linked multimeric complexes, which also contain p19 α and p19f species. Other dimers observed, gp85-gp37 and p12-p12, have been described previously (23). Panel B (Fig. 7) shows the dimer-trimer region from a similar gel in which the protein complexes were first resolved on longer 40-cm gels. On these gels, at least four species of p19 dimers are apparent: p19 α -p19 α , p19 α -p19 β , p19 β -p19 β , and p19 β -p19f. The dumbbell appearance of released products under reducing conditions, which is particularly apparent for the p19 β -p19 β dimer, suggests that each cross-linked dimer exists in two forms that exhibit different mobilities in SDS-gels. This phenomenon is further characterized below. In the expanded two-dimensional profile, a fifth minor complex is also observed (p19 α -p19f); however, it is not obvious without a spot-to-spot comparison of the profiles presented here with a profile obtained for PrC under the same conditions (data not shown). In PRE21-21 this form is found between the two p19 β -p19 β dimers. The presence of all possible combina-

FIG. 7. Identification of disulfide-linked complexes in PRE21-21 by two-dimensional SDS-PAGE. [^{35}S]cysteine-labeled PRE21-21 was treated with 50 mM iodoacetamide for 30 min at room temperature in the dark, then dissolved directly in SDS-gel sample buffer, and subjected to diagonal gel electrophoresis as described previously (23). First dimension, no reducing agent; second dimension, 3% 2-mercaptoethanol. Radioactive spots were visualized by fluorography. A, Gels were 20 by 20 cm; B, the p19 dimer-trimer region from a gel measuring 40 by 20 cm. At the left of A is a strip from a one-dimensional gel profile of [^{35}S]cysteine-labeled virus.

tions of cross-linked products of p19 α , p19 β , and p19f, at intensities that roughly reflect their relative concentrations in virions, suggests that they are localized together. This inference was drawn previously for p19 α and p23 from lipid-protein cross-linking studies (24).

In the analysis of p19 β -containing viruses, it was fortuitous that SDS-gels resolved cross-linked dimeric complexes into two components, thus allowing the structure of each to be analyzed separately. When a gel strip cut orthogonal to the two dimer bands from a one-dimensional gel was subjected to CNBr mapping (Fig. 6f and g), each band yielded a unique cleavage profile that differed from the profile for uncross-linked p19 β . For dimer I (lane f), the only difference between the cleavage profiles of the disulfide-linked dimer and p19 β itself (lane c, d, or h) is a single missing fragment (denoted by arrow), which corresponds to fragment 4 + 3

(Fig. 6). No other cleavage products are affected. Since this is the only cleavage product that is unique to fragment 4, we conclude that the disulfide linkage in dimer I is homotypic, coupling the single cysteine at residue 154 in fragment 4 of one p19 molecule to the same position on a second p19 molecule. Similarly, we infer that dimer II is a homotypic dimer of fragment 3, formed at the single cysteine at residue 112. In this profile (Fig. 6g), the entire collection of fragment 3-containing products is absent, whereas fragments 1 and 1 + 2 are not affected. In this type of analysis, any heterotypic dimers (e.g., 3-4 dimers) would yield a collection of complete and partial cleavage fragments that includes all those seen from p19 itself. Thus, clearly there is not evidence of heterotypic p19 dimers. CNBr cleavage of dimers should also yield disulfide-banded fragments not seen in monomeric p19, namely 3-3 and 4-4 (and partials built upon these). We have not been able to see such predicted fragments. The former is too large to be resolved in this gel system, whereas the latter presumably does not enter the gel for the same reason that fragment 4 itself does not enter the gel. In summary, as in the case of Pr65, the murine retrovirus *gag* precursor (22), CNBr mapping of naturally occurring disulfide-bonded p19 leads to the conclusion that like domains of the protein interact.

DISCUSSION

We have investigated the lipid-protein and protein-protein interactions of the ASLV *gag* protein p19 by a combination of cross-linking and CNBr peptide analysis. The data presented imply that the site of interaction with membrane is within the amino-terminal quarter of p19 and that homotypic protein-protein interactions occur within the carboxy-terminal quarter of p19. Two distinguishable p19-p19 dimers are formed by disulfide linkages at amino acid 112 on adjoining molecules and at amino acid 154 on adjoining molecules, since these are the only cysteines in the two reactive fragments (29). We infer that the observed p19 oligomers and p19-lipid complexes all are drawn from the same population, since oligomers were shown previously to become cross-linked to lipid (24). The specificity of the observed interactions together with the abundance of the species analyzed suggest that they represent major structural features of ASLVs. This conclusion is further supported by observations from murine retroviruses (24; Y. Yoshinaka, I. Katoh, and R. B. Luftig, *Virology*, in press) and from mouse mammary tumor virus (4, 27, 37), in which the corresponding *gag* proteins exhibit similar properties.

The nature of the association of avian p19 with membrane differs from that of its murine retrovirus counterpart p15. Whereas murine p15 is transmembranous (1) and is presumably anchored into the membrane via covalently bound lipid molecules (13, 28), p19 exhibits neither of these properties (19, 28). These findings are consistent with the observations that purified p15 can be reconstituted into liposomes prepared with phosphatidylcholine (1), whereas p19 cannot (unpublished data). They indicate that some other component may be necessary for the association of p19 with membrane. Montelaro et al. (20) suggested that p19 may be anchored to the viral envelope by viral glycoproteins. Although some contact between p19 and gp37 may occur, this interaction cannot play an obligatory role in virus assembly since viruses containing large *env* deletions, such as the Bryan high titer strain of RSV, bud from the plasma membrane properly (9). Another possibility is that a cellular component is required. This possibility is suggested by studies of RSV-infected mammalian cell lines, in which the

gag precursor polypeptide is synthesized as a stable protein (11, 36) yet no virus particles are formed (11). Infectious virus can be recovered from mammalian cell lines by fusion with chicken cells (33). Low levels of infectious virus also can be rescued by fusing the infected mammalian cells with red cell ghosts containing total AMV proteins (11), suggesting that components necessary for assembly may be carried by virions.

A third possibility is that p19 associates with the membrane through electrostatic interactions with charged phospholipids, as has been inferred for the matrix, or M, protein of vesicular stomatitis virus (38, 39). M protein appears to interact with liposomes *in vitro* only if they contain a high concentration of negatively charged phospholipids such as phosphatidylserine. Lipid-protein cross-linking studies (24, 40) show that there are amino groups in M protein closely associated with the membrane. It remains to be seen whether these amino groups form an essential part of the lipid binding site. Although similar reconstitutions with liposomes have not been performed with p19, two observations suggest that p19 and M protein may be similar in this regard. First, the region in p19 that we have shown to contain the lipid-protein cross-links is highly charged, containing six lysyl and three glutamyl residues. At least one of the positively charged amino groups must be located in close apposition to the membrane. Second, RSV particles are enriched for phosphatidylserine, containing twice the concentration found in the plasma membrane from which the virus buds (26).

In virions a small fraction of p19 molecules is reported to interact with genomic RNA (30). Specific p19-RNA interactions have been identified *in situ* by UV cross-linking (30) and *in vitro* by nitrocellulose filter binding assays (16). Although high-affinity RNA binding proteins are common to all retroviruses, p19 is unique among these proteins in that it also is the *gag* protein that is associated with lipid. In *in vitro* binding studies (14) in which purified CNBr cleavage fragments of p19 were used, only the amino-terminal fragment exhibited a binding activity. However, this fragment did not retain specificity for viral RNA. Using published UV-cross-linking procedures (7, 30), we were unable to obtain sufficient quantities of p19-RNA complexes for CNBr analysis. Thus a more complete analysis will be required before a full picture of the functional domains within p19 is obtained.

The ability to localize interactions within the polypeptide backbone of p19 relies on a peptide mapping technique that we developed (22), in which the fragmentary compositions of cleavage fragments are determined by nearest-neighbor analysis. This approach is a powerful extension of the limited proteolysis method (6), in that the peptide maps serve as a tool for mapping interactions or modifications within a protein. In addition to the structural studies described, the CNBr map of p19 has been used for studying the primary structure of p19 (Vogt et al., manuscript in preparation) and for epitope assignment of monoclonal antibodies (unpublished data). The peptide maps constructed for p19 and its related proteins should serve as useful tools for future studies.

ACKNOWLEDGMENTS

This work was supported by Public Health Service grant CA-20081 from the National Cancer Institute.

We thank Maxine Linial for the quail cell lines and the gift of the radioactive recombinant virus SE21Q1b $\times$ RAV-1, Robert Eisenman for the ev-1-producing chicken cells, and Carol Miernicki Steeg for critical reading of the manuscript.

LITERATURE CITED

1. Barbacid, M., and S. A. Aaronson. 1978. Membrane properties of the *gag* gene-coded p15 protein of mouse type-C RNA tumor viruses. *J. Biol. Chem.* **253**:1408-1414.
2. Bolognesi, D. P., R. Luftig, and J. H. Shaper. 1973. Localization of RNA tumor virus polypeptides. I. Isolation of further virus substructures. *Virology* **56**:549-564.
3. Bolognesi, D. P., R. C. Montelaro, H. Frank, and W. Schäfer. 1978. Assembly of type C oncornaviruses: a model. *Science* **199**:183-186.
4. Cardiff, R. D., M. J. Puentes, L. J. T. Young, G. H. Smith, Y. A. Teramoto, B. W. Altrock, and T. S. Pratt. 1978. Serological and biochemical characterization of the mouse mammary tumor virus with localization of p10. *Virology* **86**:398-412.
5. Chamberlain, J. P. 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. *Anal. Biochem.* **98**:132-135.
6. Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli. 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. *J. Biol. Chem.* **252**:1102-1106.
7. Darlix, J.-L., and P. F. Spahr. 1982. Binding sites of viral protein p19 onto Rous sarcoma virus RNA and possible controls of viral functions. *J. Mol. Biol.* **160**:147-161.
8. Davis, N. L., and R. R. Rueckert. 1972. Properties of a ribonucleoprotein particle isolated from Nonidet P-40-treated Rous sarcoma virus. *J. Virol.* **10**:1010-1020.
9. Eisenman, R. N., and V. M. Vogt. 1978. The biosynthesis of oncovirus proteins. *Biochem. Biophys. Acta* **473**:187-239.
10. Erikson, E., J. S. Brugge, and R. L. Erikson. 1977. Phosphorylated and nonphosphorylated forms of avian sarcoma virus polypeptide p19. *Virology* **80**:177-185.
11. Frixen, U., K. Von der Helm, R. Friis, W. Wille, and K. Willecke. 1982. Fusion infection of Rous sarcoma virus proteins into Rous sarcoma virus-transformed non-producing hamster cells causes release of infectious virus. *J. Gen. Virol.* **61**:83-86.
12. Hager, D. A., and R. R. Burgess. 1980. Elution of proteins from sodium dodecyl sulfate-polyacrylamide gels, removal of sodium dodecyl sulfate, and renaturation of enzymatic activity: results with sigma subunit of *Escherichia coli* RNA polymerase, wheat germ DNA topoisomerase, and other enzymes. *Anal. Biochem.* **109**:76-86.
13. 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. U.S.A.* **80**:339-343.
14. 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.
15. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680-685.
16. Leis, J. P., J. McGinnis, and R. W. Green. 1978. Rous sarcoma virus p19 binds to specific double-stranded regions of viral RNA: effect of p19 on cleavage of viral RNA by RNase III. *Virology* **84**:87-98.
17. Leis, J. P., P. Scheible, and R. E. Smith. 1980. Correlation of RNA binding affinity of avian oncornavirus p19 proteins with the extent of processing of virus genome RNA in cells. *J. Virol.* **35**:722-731.
18. Linial, M., E. Medeiros, and W. S. Hayward. 1978. An avian oncovirus mutant (SE21Qb) deficient in genomic RNA: biological and biochemical characterization. *Cell* **15**:1371-1381.
19. Montelaro, R. C., and R. R. Rueckert. 1975. On the use of chloramine-T to iodinate specifically the surface proteins of intact enveloped viruses. *J. Gen. Virol.* **29**:127-131.
20. Montelaro, R. C., S. J. Sullivan, and D. P. Bolognesi. 1978. An analysis of type-C retrovirus polypeptides and their associations in the virion. *Virology* **84**:19-31.
21. Oakely, B. R., D. R. Kirsch, and N. R. Morris. 1980. A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. *Anal. Biochem.* **105**:361-363.
22. Pepinsky, R. B. 1983. Localization of lipid-protein and protein-protein interactions within the murine retrovirus *gag* precursor by a novel peptide mapping technique. *J. Biol. Chem.* **258**:11229-11235.
23. Pepinsky, R. B., D. Cappiello, C. Wilkowski, and V. M. Vogt. 1980. Chemical cross-linking of proteins in avian sarcoma and leukemia viruses. *Virology* **102**:205-210.
24. Pepinsky, R. B., and V. M. Vogt. 1979. Identification of retrovirus matrix proteins by lipid-protein cross-linking. *J. Mol. Biol.* **131**:819-837.
25. Pepinsky, R. B., and V. M. Vogt. 1983. Purification and properties of a fifth major *gag* protein from avian sarcoma and leukemia viruses. *J. Virol.* **45**:648-658.
26. Pessin, J. E., and M. Glaser. 1983. Budding of Rous sarcoma virus and vesicular stomatitis virus from localized lipid regions in the plasma membrane of chicken embryo fibroblasts. *J. Biol. Chem.* **258**:9044-9050.
27. Racevskis, J., and N. H. Sarkar. 1980. Murine mammary tumor virus structural protein interactions: formation of oligomeric complexes with cleavable cross-linking agents. *J. Virol.* **35**:937-948.
28. Schultz, A. M., and S. Oroszlan. 1983. In vivo modification of retroviral *gag* gene-encoded polypeptides by myristic acid. *J. Virol.* **46**:355-361.
29. Schwartz, D., R. Tizard, and W. Gilbert. 1983. Nucleotide sequence of Rous sarcoma virus. *Cell* **32**:853-869.
30. Sen, A., and G. J. Todaro. 1977. The genome-associated specific RNA binding proteins of avian and mammalian type C viruses. *Cell* **10**:91-99.
31. Shaikh, R., M. Linial, S. Brown, A. Sen, and R. Eisenman. 1979. Recombinant avian oncoviruses. II. Alterations in the *gag* proteins and evidence for intragenic recombination. *Virology* **92**:463-481.
32. Shealy, D. J., A. G. Mosser, and R. R. Rueckert. 1980. Novel p19-related protein in Rous-associated virus type 61: implications for avian *gag* gene order. *J. Virol.* **34**:431-437.
33. Steimer, K. S., and D. Boettiger. 1977. Complementation rescue of Rous sarcoma virus from transformed mammalian cells by polyethylene glycol-mediated cell fusion. *J. Virol.* **23**:133-141.
34. Stromberg, K., N. E. Hurley, N. L. Davis, R. R. Rueckert, and E. Fleissner. 1974. Structural studies of avian myeloblastosis virus: comparison of polypeptides in virion and core component by dodecyl sulfate-polyacrylamide gel electrophoresis. *J. Virol.* **13**:513-528.
35. Swank, R. T., and K. D. Munkres. 1971. Molecular weight analysis of oligopeptides by electrophoresis in polyacrylamide gel with sodium dodecyl sulfate. *Anal. Biochem.* **39**:462-477.
36. Vogt, V. M., D. A. Bruckenstein, and A. P. Bell. 1982. Avian sarcoma virus *gag* precursor polypeptide is not processed in mammalian cells. *J. Virol.* **44**:725-730.
37. Weiss, R., N. Teich, H. Varmus, and J. Coffin (ed). 1982. Molecular biology of tumor viruses: RNA tumor viruses. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
38. Wiener, J. R., R. Pal, Y. Barenholz, and R. R. Wagner. 1983. Influence of the peripheral matrix protein of vesicular stomatitis virus on the membrane dynamics of mixed phospholipid vesicles: fluorescence studies. *Biochemistry* **22**:2162-2170.
39. Zakowski, J. J., W. A. Petri, and R. R. Wagner. 1981. Role of matrix protein in assembling the membrane of vesicular stomatitis virus: reconstitution of matrix protein with negatively charged phospholipids. *Biochemistry* **20**:3902-3907.
40. Zakowski, J. J., and R. R. Wagner. 1980. Localization of membrane-associated proteins in vesicular stomatitis virus by use of hydrophobic membrane probes and cross-linking reagents. *J. Virol.* **36**:93-102.