Cyanogen Bromide Digestion of the Avian Myeloblastosis Virus pp19 Protein: Isolation of an Amino-Terminal Peptide that Binds to Viral RNA

S. P. JOHNSON,¹ M. VEIGL,² T. VANAMAN,² AND J. LEIS^{1*}

Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106,¹ and Department of Microbiology and Immunology, Duke University Medical Center, Durham, North Carolina 27710²

Received 21 June 1982/Accepted 7 October 1982

The avian myeloblastosis virus pp19 protein was separated from the other virus proteins by a rapid and simple purification procedure which yields milligram amounts of homogeneous protein. This protein was then fragmented by digestion with cyanogen bromide. When the mixture of the cyanogen bromide peptides was passed through a 60S avian myeloblastosis virus RNA-cellulose column, only one peptide bound with high affinity to the resin. The peptide migrated on a sodium dodecyl sulfate-polyacrylamide gel with an approximate molecular weight of 2,900 and will be referred to as the p3B peptide. This peptide was also isolated directly by chromatography of the cyanogen bromide-digested pp19 protein on a reverse-phase high-pressure liquid chromatography column. It was again the only cyanogen bromide peptide of the pp19 protein that bound to the RNA affinity resin. The p3B peptide is a basic peptide, as was seen by its rapid migration on acid-urea-polyacrylamide gels and its amino acid composition. A partial amino acid sequence analysis of the p3B peptide indicated that it was derived from the amino terminus of the intact protein. Although the p3B peptide bound to 60S RNA, it did not demonstrate the selective binding of native pp19 to regions of the RNA containing secondary structure.

The function of each of the major internal structural proteins, p27, pp19, p15, and pp12, of avian retroviruses has been extensively studied during the past several years (for a review, see reference 1). One of these, the pp19 polypeptide, appears to have at least two functions in the virion. First, it is a matrix protein, as evidenced by it being found disulfide-linked to the virus glycoprotein (9) and by cross-linking studies showing that the pp19 protein is in close proximity to the lipid bilayer membrane (11). Second, a small number of pp19 molecules are found associated with the viral RNA (12). Previous studies have suggested that this binding may regulate the level of *in vivo* processing of genome-length RNA (6, 7). Since the pp19 protein appears to be multifunctional, we attempted to isolate fragments generated by chemical cleavage of the protein which have distinct functional properties. In this report, we describe a novel purification procedure of the pp19 protein and the identification of a small polypeptide, produced by cyanogen bromide cleavage of pp19, that binds to viral RNA.

The avian myeloblastosis virus (AMV) pp19 protein used in these studies was purified by using a modification of the procedure outlined

by Herman (A. Herman, Ph.D. thesis, Duke University, Durham, N.C., 1975). This simple procedure yielded homogeneous viral proteins in milligram quantities without the use of gel filtration in the presence of guanidine hydrochloride (2). Sucrose gradient-purified virus (2 to 3 g), suspended in 100 ml of 0.1 M NaCl, was extracted with 210 ml of chloroform-methanol (2:1, vol/vol). The phases were separated by centrifugation. The p15 protein is found in the aqueous phase. The other viral proteins were found at the interface. The protein pellet was dried under nitrogen and extracted once with 0.1 M sodium phosphate (pH 7.0)-1% 2-mercaptoethanol-2 mM phenylmethylsulfonyl fluoride at 37°C for 2 h and then four times with the same buffer with 0.2% Triton X-100 and no phenylmethylsulfonyl fluoride at 4°C for 12 h each. During one of the extractions, 0.2% Sarkosyl was substituted for Triton X-100.

The combined supernatants, containing the solubilized proteins, were applied to a carboxymethyl-Sephadex C-50 column (26 by 2.5 cm) equilibrated with 0.1 M sodium phosphate (pH 7.0)-1 mM 2-mercaptoethanol (column buffer) at room temperature, and the column was washed with column buffer containing 0.1 M NaCl.

FIG. 1. Chromatography of AMV pp19 cyanogen bromide peptides on 60S AMV RNA-cellulose. AMV pp19 (20 to 25 μ g) was dissolved in 70% formic acid and reacted for 18 to 20 h at room temperature with cyanogen bromide (100-fold in molar excess of methionine residues). The mixture was diluted with 10 volumes of water and lyophilized twice. The soluble protein (14 mg) was applied to an AMV 60S RNA-cellulose column (1.5 by 22 cm) which had been prepared by cross-linking AMV 60S RNA (87 absorbance units at 260 nm) to 4.4 g of cellulose (Whatman CF11, washed with absolute ethanol) by irradiating for 15 min at a distance of 8 cm from two 15-W germicidal UV lamps. The irradiated cellulose resin was washed with 0.01 M sodium phosphate (pH 7.0)–0.002 M 2-mercaptoethanol–1 M NaCl to remove unbound RNA (less than 15% of the viral RNA was removed by this wash treatment). The column was equilibrated with 0.01 M sodium phosphate (pH 7.0)–0.002 M 2-mercaptoethanol at 4°C with a flow rate of 15 ml/h before the protein was loaded. The column was washed with two column volumes of this buffer and then eluted with one column volume of 1 M NaCl–0.01 M sodium phosphate (pH 7.0)–0.002 M 2-mercaptoethanol at a rate of 20 ml/h. Column fractions (1 ml) were analyzed for absorbance at 280 nm and for RNA-binding activity as measured by retention on nitrocellulose filters (6), using 32 P-labeled RSV 60S RNA (13). (A) Elution profile of the cyanogen bromide-treated pp19 protein applied to the column. (B) The 1 M NaCl buffer eluate fractions shown in panel A were combined, desalted

Under these conditions, the only viral protein retained by the column was pp12, which could be recovered in good yield by elution with column buffer containing 1 M NaCl. The pass-through fraction was dialyzed against 35 mM sodium phosphate (pH 7.0)–1 mM 2-mercaptoethanol and applied to a second carboxymethyl-Sephadex column (32 by 2.5 cm) equilibrated with dialysis buffer. The pp19 protein, which is the only component retained under these conditions, was eluted with column buffer containing 1 M NaCl. The pp19 protein was detected by monitoring absorbance at 280 nm. Its purity was greater than 95% as judged by analysis on sodium dodecyl sulfate (SDS)-polyacrylamide gels.

The purified pp19 protein was fragmented by site-specific chemical cleavage with cyanogen bromide. The amino acid composition of AMV pp19 showed that there are five methionine residues per molecule of protein (data not shown). Since the amino terminus of the pp19 protein has been reported to contain a blocked methionine (3, 10), cyanogen bromide cleavage may produce up to five polypeptide fragments.

In an attempt to resolve a cyanogen bromide peptide which retained the ability to bind to viral RNA, the cyanogen bromide-treated pp19 protein was chromatographed on an AMV 60S RNA-cellulose affinity resin as described in the legend to Fig. 1. A total of 78% of the material applied to the column was recovered in the pass-through fraction from the column, whereas 22% was bound and eluted from the RNA resin with 1 M NaCl buffer. The recovery of total protein was quantitative. Only the bound material caused the retention of RNA on nitrocellulose filters (Fig. 1). (Chromatography on a sham CF11 cellulose column indicated that these retained peptides were interacting specifically with the RNA moiety and not with the cellulose.) Polyacrylamide gel electrophoretic analyses of the proteins in these two fractions revealed that the binding fraction (Fig. 2, lane B) contained a doublet migrating with an approximate molecular weight of 10,500 and a group of unresolved Coomassie blue-staining bands migrating between 1,500 and 4,000 daltons. The nonbinding fraction contained only a doublet at the 10,500-dalton positive (Fig. 2, lane C). When this latter fraction was again chromatographed on the AMV 60S RNA column, none of the UV-

by gel filtration, lyophilized, taken up in 1.5 ml of 0.01 M sodium phosphate (pH 7.0)–0.002 M 2-mercaptoethanol, and applied to the AMV RNA-cellulose column equilibrated with the same buffer. The column was washed and monitored as described above. (C) The wash-through fractions of the column shown in panel A were chromatographed on the AMV RNA-cellulose column as described above.

FIG. 2. Gel electrophoresis analysis of the cyanogen bromide polypeptides of AMV pp19 chromatographed on an AMV 60S RNA-cellulose column. Samples containing 10 to 20 μ g of the cyanogen bromide peptides chromatographed on AMV 60S RNA-cellulose as described in legend to Fig. 1 were analyzed for size by polyacrylamide gel electrophoresis. The 15% polyacrylamide gel (acrylamide-methylenebisacrylamide, 40:1) contained 0.1% SDS, 0.1 M Tris, and 0.1 M Bicine and was electrophoresed at 25 mA. Polypeptides were detected by staining with Coomassie brilliant blue R-250. Lanes: A, protein molecular weight standards myoglobin (25,000 daltons), chymotrypsinogen (17,800 daltons), cytochrome *c* (12,500 daltons), and bacitracin (1,450 daltons); B, 15 μ g of the 1 M eluate fraction in Fig. 1A; C, 20 μ g of the wash-through fraction in Fig. 1A; D, 10 μ g of the wash-through fraction in Fig. 1B; E, 10 μ g of the 1 M eluate fraction in Fig. 1B.

absorbing material bound to the resin (Fig. 1C).

Chromatography of the bound fraction (after desalting and lyophilization) resulted in UV-absorbing material in both the unbound and the 1 M NaCl-eluted fractions (Fig. 1B). An analysis by SDS-polyacrylamide gel electrophoresis (Fig. 2) revealed that the material bound to the affinity resin contained a Coomassie blue-staining band of approximately 3,000 daltons (Fig. 2, lane E), whereas the wash-through fraction contained the remaining 10,500-dalton polypeptide as well as material which migrated with a molecular weight of approximately 3,000 (Fig. 2, lane D). The bound and unbound polypeptide fractions from this second chromatography on the affinity resin were analyzed by acid-urea gels. In this gel system, the RNA-binding fraction contained only one Coomassie blue-staining band. This band migrated much more rapidly than any peptide band seen in the nonbinding fractions. The amino acid composition of this peptide showed it to be basic, which would explain its rapid migration in the acid-urea gel (data not

shown). This RNA-binding peptide was greater than 90% pure as judged by the amount of contaminating peptides present in the stained gels and was used to determine the amino acid sequence (see below) without further purification. We refer to this RNA-binding peptide as p3B below.

We also purified the p3B peptide by reverse-phase high-pressure liquid chromatography (HPLC) on a Synchropak C-8 column. The elution pattern of the AMV pp19 cyanogen bromide peptides is shown in Fig. 3. Peaks A and B contained only the p3B peptide when analyzed by SDS- (data not shown) and acid-urea-polyacrylamide gel electrophoresis (Fig. 4, lanes A and B, respectively). The same two peaks were observed when the p3B peptide was first purified on the RNA affinity resin as described in the legend to Fig. 1 and may represent equilibrium distributions of homoserine and homoserine lactone as carboxy-terminal residues. Some of the p3B peptide was also found eluting in peak D (Fig. 3), suggesting that it associated with other peptides. The remaining peptides of the pp19

FIG. 3. HPLC of cyanogen bromide-treated pp19. Cyanogen bromide-treated AMV pp19 (2 mg) was suspended in 0.05 ml of 6 M guanidine hydrochloride. Separation of the fragments was performed by reverse-phase chromatography (8) on a Synchropak RP-P column (4.1 by 250 mm) (Synchrom, Linden, Ind.) with a DuPont model 850 liquid chromatograph. Run parameters: column and solvent temperatures, 25°C; flow rate, 0.7 ml/min; solvent A, 15% isopropanol-0.013 M trifluoroacetic acid; solvent B, isopropanol; linear gradient program, 90 min from 0 to 60% solvent B.

FIG. 4. Acid-urea-gel electrophoresis of HPLC fractions. HPLC fractions were electrophoresed on a resolving gel of 15% polyacrylamide (acrylamide-methylenebisacrylamide, 40:1) in 5% acetic acid-5 M freshly deionized urea with a running buffer of 5% acetic acid. Peptides were detected by staining with Coomassie brilliant blue G-250. Lanes A, peak A of Fig. 3 (13 μ g of protein); B, peak B of Fig. 3 (8 μ g of protein). Arrow, position of migration of the 10,500-dalton peptide.

protein were not clearly resolved from one another in this HPLC system (Fig. 3).

The HPLC fractions from peak B (Fig. 3) were chromatographed on an AMV 60S RNA-cellulose column, and any peptides which bound were eluted with 1 M NaCl. The results are shown in Fig. 5A. A total of 72% of the UV-absorbing material bound to the affinity resin. The bound and pass-through fractions were analyzed by acid-urea-polyacrylamide gel electrophoresis after being desalted. The 1 M eluate contained only the rapidly migrating band which is diagnostic for the p3B peptide. No Coomassie blue-staining bands were seen from the pass-through fraction.

A sample of the pooled HPLC peaks E through H containing all of the other peptides was similarly chromatographed on the AMV 60S RNA-cellulose resin. In this instance, less than 7% of the UV-absorbing material was bound and eluted with 1 M NaCl (Fig. 5B). An examination of these column fractions on acid-urea gels revealed no observable bands in the bound fraction, whereas the pass-through fraction appeared identical to the sample loaded on the column. These results suggest that p3B binding to RNA does not require the presence of any other portion of the pp19 molecule and that no other cyanogen bromide peptide binds RNA.

The p3B peptide in a concentration of less than 24 μ g/ml, as well as the unfractionated cyanogen bromide digest of pp19 protein, demonstrated only trace amounts of detectable RNA-binding activity as measured by the nitrocellulose filter assay (6). The loss of RNA-binding units probably reflects the inability of the cyanogen bromide peptides of pp19 to be retained on the nitrocellulose filter. This was suggested by an experiment in which a pp19 cyanogen bromide digest was passed through nitrocellulose filters. An SDS-polyacrylamide gel electrophoretic analysis revealed similar patterns for the peptides in the filtrate and those in the original protein digest (data not shown). Since the RNA binding assay assumes retention of the protein component on the filters, the decrease in RNA-binding activity results, at least in part, from peptide-bound RNA not being retained. The addition of the nonbinding polypeptides to the p3B peptide did not restore the RNA-binding units in this assay. At higher concentrations of the p3B peptide, RNA was retained on the filters. This appears to be the result of aggregation of the peptide at these concentrations, which allows it to be retained on the nitrocellulose filters.

FIG. 5. Affinity chromatography on AMV 60S RNA-cellulose of HPLC-isolated peptides. Fractions from the HPLC column were chromatographed on AMV 60S RNA-cellulose as described in the legend to Fig. 1. (A) Elution profile of fractions from peak B of Fig. 3 (0.15 mg of protein). (B) Elution profile of a sample (0.75 mg of protein) of peaks E to H of Fig. 3.

FIG. 6. Inhibition by the p3B peptide of RSV 60S RNA binding by AMV pp19 or pp12. 32 P-labeled RSV 60S RNA (200 pmol, 140 to 250 cpm/pmol) was incubated with the p3B peptide at 23°C in 0.03 ml of 10 mM Tris-hydrochloride (pH 7.5)–10 mM sodium chloride–1 mM dithioerythritol for 10 min. AMV pp19 (0.08 μg) or AMV pp12 (0.16 μg) was added to the incubation mixture, and incubation was continued for another 10 min at 23°C. Samples were filtered through nitrocellulose filters to determine the amount of RNA bound (6). Symbols: ○, pp12 protein; △, pp19 protein.

The specificity of p3B binding to 60S RNA was therefore monitored by assaying its ability to block the 60S RNA-binding sites of intact pp19 (which binds specifically to secondary structure-containing regions of 60S RNA) and of intact pp12 (which binds to single-stranded nucleic acids) (Fig. 6). At peptide concentrations at which there was negligible retention of the RNA on nitrocellulose filters, preincubation of the p3B peptide with the RNA caused a reduction in the amount of RNA bound to the filter by AMV pp19. Increasing the amount of the competing p3B peptide resulted in decreased binding of RNA to AMV pp19. (These studies were performed with a constant amount of pp19, which was less than saturating for RNA binding.) A 50% reduction in binding was obtained at a molar ratio of p3B to pp19 of approximately 78:1. This high ratio of peptide to pp19 protein required for competition suggested that the peptide bound to RNA nonselectively. Using similar procedures, we observed that increasing amounts of p3B peptide caused a decrease in the RNA bound to the filters by pp12 (Fig. 6). In this case, 50% inhibition was obtained at a molar ratio of p3B to pp12 of approximately 10:1. Since p3B competes with pp12 even more effectively than with pp19, the peptide appears to bind RNA without the specificity for double-stranded RNA which is characteristic of the pp19 protein. Thus, the p3B peptide does bind RNA. However, it does not appear to contain the entire RNA-binding site of the intact protein,

since it has lost the binding specificity characteristically displayed by pp19.

The region of the pp19 molecule contained within the p3B peptide was deduced from its chemical characteristics, a limited sequence analysis, and knowledge of the *gag* gene sequence. The results of an automated Edman degradation of peptide p3B (Table 1) allowed the assignment of residues at positions 1 to 7, 10 to 11, and 14 of the p3B peptide. These assignments agree exactly with those predicted for residues 2 to 15 of subgroup A Schmidt-Ruppin Rous sarcoma virus (RSV) pp19 (14). The additional positions 8, 9, 11, and 13, are serine, serine, cysteine, and threonine, whose phenylthiohydantoin derivatives are labile and therefore difficult to determine during an automated sequence analysis. The major sequence assigned here is also in complete agreement with that obtained from an analysis of products from *in vitro* translation of RSV PR-C 35S RNA (10), in which the initiation methionyl residue was followed by the sequence deduced for p3B in Table 1. Cyanogen bromide cleavage would have removed this methionine residue, which appears to be acetylated in the intact protein (3, 10). We therefore conclude that the p3B peptide is derived from the amino terminus of AMV pp19.

The complete amino acid sequence of the p3B peptide can be predicted from the nucleic acid sequence of the amino-terminal portion of the *gag* gene (14; D. Schwartz, personal communication). After the amino-terminal methionine, there are 28 amino acids before the next methionine is found in the sequence. This would be the p3B peptide. Its sequence has a cluster of basic amino acid residues where in one area 4 of 12 residues are basic. There is no other such cluster of basic amino acids in the predicted sequence of the pp19 protein. Such a grouping of basic residues could be involved in the RNA binding. This is consistent with the retention of only this peptide on the RNA affinity resin. However, since the p3B peptide does not retain the specific recognition of intact pp19 for double-stranded RNA, some other region of the protein chain must restrict its recognition to double-stranded RNA sites. Indeed, an interaction between the p3B peptide and the other peptides was seen during chromatography on the 60S AMV RNA-cellulose column. This is also in agreement with the finding that the apparent association constant of the cyanogen bromide-treated pp19 protein for RSV RNA was 30-fold lower than that of the intact protein and is consistent with the idea that the tertiary structure of the protein is important for specific recognition of regions of viral RNA. Studies are currently in progress to further define the RNA-binding structure of pp19 by isolating larger peptides of the pp19 protein

TABLE 1. Automated Edman degradation of AMV pp19 RNA-binding cyanogen bromide fragment^a

Cycle	Major residue ^b /amt (nmol)	Minor residue ^b /amt (nmol)	Sequence deduced	Sequence predicted ^c
1	Glu/0.8	Val/1.0	Glu/Val	Glu
2	Ala/1.7		Ala	Ala
3	Val/6.6	Ala/0.1	Val	Val
4	Ile/4.6	Val/0.1	Ile	Ile
5	Lys/2.0		Lys	Lys
6	Val/5.3		Val	Val
7	Ile/5.2	Val/0.4	Ile	Ile
8				Ser
9				Ser
10	Ala/3.6		Ala	Ala
11		Ala/0.4		Cys
12	Lys/0.5		Lys	Lys
13				Thr
14	Tyr/1.8	Pro/0.3	Tyr	Tyr

^a Automated amino acid sequence analysis was performed with an updated Beckman model 890B sequencer using a modification of the Beckman Quadrol peptide program and 0.55 M Quadrol buffer in the presence of Polybrene (Aldrich Chemical Co.) (3 mg) added to prevent peptide losses as described by Klapper et al. (5). Degradation was performed with 30 nmol of starting material. Repetitive yield (based on Val-3, Val-6, Ala-10, Tyr-14) = 92%. Initial yield (extrapolated from Val-3 assuming 92% repetitive yield) = 30%.

^b Phenylthiohydantoin amino acids, obtained from sequencer fractions after acid treatment (10 min at 80% in 1 N HCl-0.01% dithioerythritol), were identified and quantified by HPLC. Separation was achieved by a slight modification of the procedure of Hunkapiller and Hood (4) with a DuPont Zorbax-ODS column (0.46 by 25 cm) on a Hewlett-Packard 1084B liquid chromatograph. Run parameters: column and solvent temperatures, 60°C; flow rate, 1.5 ml/min; solvent A, 0.020 M sodium acetate (pH 4.8); solvent B, acetonitrile; linear gradient program, 0 to 6 min from 10 to 35% solvent B, 6 to 18 min from 35 to 45% solvent B, 18 to 20 min at 45% solvent B, 20 to 22 min at 80% solvent B, 22 to 24 min at 100% solvent B. Phenylthiohydantoin amino acids were detected at 268 nm, except for the dehydro derivatives of phenylthiohydantoin-Ser, -Thr, and -CMCys, which were monitored at 313 nm.

^c Based on the DNA sequence of the pp19 coding region of subgroup A Schmidt-Ruppin RSV determined by Swanstrom and co-workers (14).

that duplicate the RNA-binding specificity of the intact protein.

We are indebted to J. Beard, Life Sciences Inc., St. Petersburg, Fla., for providing the AMV through the National Cancer Institute Virus Cancer Program.

This research was supported by the Public Health Service grants CA 27414 (to J.L.) and CA 29467 (to T.V.) from the National Institutes of Health and by National Science Foundation grant PCM 77-00809 (to J.L.).

LITERATURE CITED

- Bolognesi, D. P., R. C. Montelaro, H. Frank, and W. Schafer. 1978. Assembly of type C oncornaviruses: a model. *Science* 199:183-186.
- Green, R. W., and D. P. Bolognesi. 1974. Isolation of proteins by gel filtration in 6 M guanidinium chloride: application to RNA tumor viruses. *Anal. Biochem.* 59:479-490.
- Herman, A. C., R. W. Green, D. P. Bolognesi, and T. C. Vanaman. 1975. Comparative chemical properties of avian oncornavirus polypeptides. *Virology* 64:339-348.
- Hunkapiller, M. W., and L. E. Hood. 1978. Direct microsequence analysis of polypeptides using an improved sequenator, a nonprotein carrier (Polybrene), and high pressure liquid chromatography. *Biochemistry* 17:2124-2133.
- Klapper, D. G., C. E. Wilde III, and D. J. Capra. 1978. Automated amino acid sequence of small peptides utilizing Polybrene. *Anal. Biochem.* 85:126-131.
- 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.
- 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.
- Mahoney, W. C., and M. A. Hermodson. 1980. Separation of large denatured peptides by reverse phase high performance liquid chromatography. *J. Biol. Chem.* 255:11199-11203.
- Montelaro, R. C., S. J. Sullivan, and D. P. Bolognesi. 1978. An analysis of type-C retrovirus polypeptides and their association in the virion. *Virology* 84:19-31.
- Palmiter, R. D., J. Gagnon, V. M. Vogt, S. Ripley, and R. N. Eisenman. 1978. The NH₂-terminal sequence of the avian oncovirus gag precursor polypeptide (Pr76^{gag}). *Virology* 91:423-433.
- Pepinsky, R. B., and V. M. Vogt. 1979. Identification of retrovirus matrix proteins by lipid-protein cross-linking. *J. Mol. Biol.* 131:819-837.
- Sen, A., and G. J. Todaro. 1977. The genome-associated specific RNA binding proteins of avian and mammalian type C virus. *Cell* 10:91-99.
- Smith, R. E., S. Nebes, and J. Leis. 1977. Production of large amounts of 35S RNA and complementary DNA from avian RNA tumor viruses. *Anal. Biochem.* 77:226-234.
- Swanstrom, R., H. E. Varmus, and J. M. Bishop. 1982. Nucleotide sequence of the 5' noncoding region and part of the gag gene of Rous sarcoma virus. *J. Virol.* 41:535-541.