JOURNAL OF VIROLOGY, Nov. 1975, p. 1248-1253

Vol. 16, No. 5
Printed in U.S.A.

N-Terminal Amino Acid Sequences in the Major Capsid Proteins of Foot-and-Mouth Disease Virus Types A, O, and C

H. D. MATHEKA[1] AND H. L. BACHRACH*

Plum Island Animal Disease Center, Agricultural Research Service, Greenport, New York 11944

Received for publication 9 June 1975

Sequences of amino acids at the N-termini of virus proteins VP_1, VP_2, and VP_3 were determined for foot-and-mouth disease virus types A_{12} strain 119, O_1 Brugge and C_3 Resende. In the polyacrylamide gel electrophoresis system used to purify the proteins, VP_3 migrated faster than VP_1 or VP_2; and in the virion, VP_3 could be cleaved by trypsin into VP_{3a} and VP_{3b}. The N-terminal amino acids for each of the virus types were glycine in VP_1, aspartic acid in VP_2, and threonine in VP_3. No divergences in sequence across the virus types were indicated until at least the fourth position in VP_1, the second in VP_2, and the third in VP_3. For virus types A_{12}, O_1 and C_3, the sequences were, respectively: for VP_1 (Gly-ile-phe,pro,val---), (Gly,ile,phe---) and (Gly-ile-phe,ala---); for VP_2 (Asp,X,met---), (Asp---) and Asp-leu---); and for VP_3 (Thr-thr-ala-thr---), (Thr-thr-ser---) and (Thr-thr---). Unresolved mixtures of VP_{3a} and VP_{3b}, from either type A_{12} or O_1 viruses, appeared to have the N-terminal amino acids threonine, which is presumed to be the same threonine as in uncleaved VP_3 and serine, which is generated by the tryptic cleavage.

Foot-and-mouth disease virus (FMDV) is an acid-sensitive picornavirus that contains four capsid proteins, three of which comprise approximately 91% of the total protein. In a previous paper, Bachrach et al. (2) isolated each of the three major proteins of type A_{12} virus by preparative polyacrylamide gel electrophoresis (PAGE) and determined their C-terminal sequences by the use of carboxypeptidase A.

Recently, we examined the N-terminal amino acid residues of the three major proteins of type A_{12} virus and found them to be similar to those reported for type O_1 Kaufbeuren by Adam and Strohmaier (1). This paper reports our findings on the N-termini of the three major proteins of FMDV types A_{12} strain 119, O_1 Brugge and C_3 Resende, as well as on short sequences of amino acids at the N-termini of these proteins.

MATERIALS AND METHODS

Virus and virus proteins. FMDV types A_{12} strain 119, O_1 Brugge and C_3 Resende were grown in rolling-bottle cultures of a baby hamster kidney cell line passage 21, clone 13 (15), concentrated by precipitation with 6% polyethylene glycol (16) and purified by CsCl density-gradient centrifugation (3). Virus proteins (VP_1, VP_2, and VP_3) were prepared directly from the viruses by disc PAGE according to Maizel (11) in hollow cylindrical gels (2.7 cm outer diameter by 1.1 cm inner diameter; 2-cm long 4% stacking gel; and a 15-cm long 10% resolving gel; bisacrylamide-acrylamide, 1:37 [wt/wt]) as described previously (2), except that 8 M urea (Schwarz/Mann, ultra pure) was present in both the sample preparation buffer and the gels, and the sodium dodecyl sulfate (SDS) used was the Sequanal grade of the Pierce Chemical Co. Approximately 5 mg of purified virus in 0.7 to 1.5 ml of 0.2 N KCl-0.05 M potassium phosphate buffer, pH 7.5, mixed with an equal volume of sample preparation buffer (0.06 M Tris-hydrochloride buffer, pH 7.1, containing 10% glycerine, 2% SDS, 5% β-mercaptoethanol [ME], 8 M urea, and phenol red) was heated to 100 C for 4 min and applied to a gel. Three gels were used for type A_{12} virus and one each for types O_1 and C_3. Electrophoresis at a constant current of 8 mA per gel for approximately 20 h effected sufficient separation (Fig. 1) to permit the bands to be cut out and eluted electrophoretically into dialysis bags containing PAGE buffer (0.05 M Tris-0.4 M glycine (pH 8.1)-0.1% SDS). The eluted proteins were dialyzed against water containing 0.1% SDS and 0.01% ME adjusted to pH 9 with sodium carbonate (pH 9-SDS-ME), lyophylized and resuspended in pH 9-SDS-ME solution. The purity and concentration of the proteins were determined (Bachrach, unpublished data) by electrophoresing 25-μl aliquots on analytical 12.5% gels (0.6 cm by 10 cm), staining the band(s) with Coomassie brilliant blue, scanning them at 600 nm in a recording spectrophotometer, and converting the peak areas to micrograms of protein by standard curves established for each protein with the Folin phenol reagent (10). Recoveries of purified proteins

[1] Permanent address: Federal Research Institute for Animal Virus Diseases, Tübingen, West Germany.

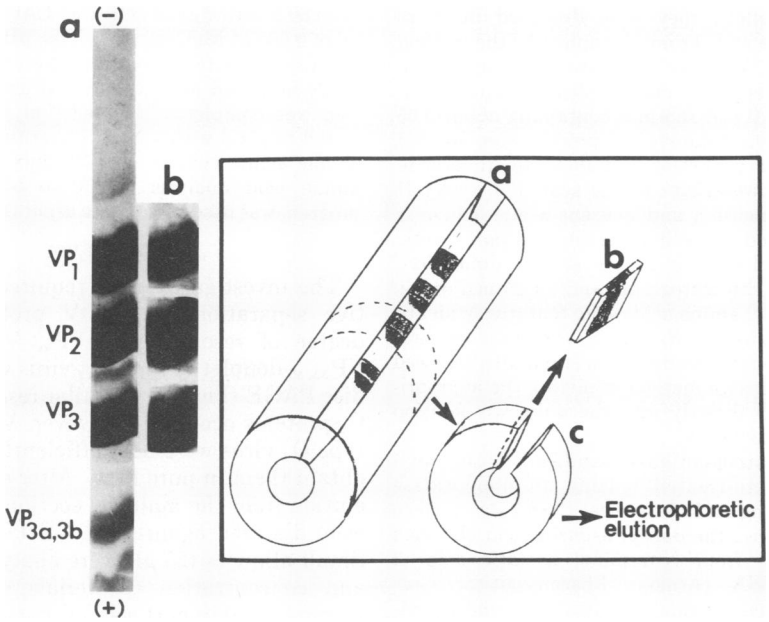

FIG. 1. *Preparative disc polyacrylamide gel electrophoresis (10% gel with 8 M urea and SDS) showing separation of type A_{12} proteins VP_1, VP_2, VP_3, and $VP_{3a,3b}$ doublet. Long narrow strip (a) is cut from preparative gel and stained. It is replaced as a guide for excising protein-containing sectors (c). The accuracy of cutting of sectors (c), before electrophoretic elution of proteins, is monitored by removing and staining the thin wafers (b).*

from the viruses used for their preparation ranged from 66 to 100%. Concentrations of stock solutions of type A_{12} proteins ranged from 2.1 to 3.1 mg/ml and from 0.60 to 0.84 mg/6 ml for type C_3 proteins. For type O_1 virus, only VP_3 was isolated in a pure form (1.5 mg/6 ml); $VP_{1,2}$ was a mixture (0.84 mg/6 ml), as was $VP_{3a,3b}$ (0.18 mg/6 ml), which formed spontaneously in the type O_1 virus from a portion of VP_3. $VP_{3a,3b}$ (0.42 mg/13 ml) was prepared also from type A_{12} virus after the virus was first treated with tosylamido-2-phenyl ethyl chloromethyl ketone trypsin (200:1 wt/wt) at 37 C for 20 min.

Sequence analysis with dansyl chloride. The N-terminal amino acids of each virus protein and of a series of shortened proteins derived from it by stepwise Edman degradations (6) were determined by dansylation and chromatography on polyamide sheets. Approximately 2, 3, 4, 5, and 6 nM of each protein were committed to 1, 2, 3, 4, and 5 Edman cycles, respectively. The protein solutions were lyophylized in 12-ml conical tubes, treated with 50 to 100 μl of a phenylisothiocyanate-pyridine-water mixture (3:57:40 vol/vol/vol) at 50 C for 45 min to form phenylthiocarbamyl proteins, and dried in a vacuum desiccator at 60 C for 30 min. The phenylthiocarbamyl proteins were cyclized and cleaved at 60 C for 10 min in 50 to 100 μl of anhydrous trifluoracetic acid followed by drying for 20 min. This complete Edman cycle was repeated four more times; representative shortened proteins were set aside at the end of each cycle. After the final cycle, all of the accumulated shortened protein samples were extracted three times with aqueous ethylacetate (tinted with methyl green to detect the phase separations) to remove by-products. The extracted samples and 10 nM samples of VP_1, VP_2, VP_3, and $VP_{3a,3b}$ in pH 9-SDS-ME were lyophylized and dansylated by the method of Gros and Labouesse as described by Niederwieser (12), except that the proteins were first denatured in SDS before the addition of 8 M urea. (In preliminary tests, this denaturing procedure gave better yields of N-terminal amino acids from myoglobin, bovine serum albumin [BSA] and lysozyme than when either SDS or urea or sodium carbonate [5] was used alone.) The virus proteins were dissolved in 0.5 M sodium phosphate buffer at pH 8.4; because the "Edman" proteins did not have SDS, the phosphate buffer for them contained 1% SDS. Solid urea (100 to 250 mg) was then added to make each protein solution approximately 8 M in urea. After the proteins were boiled at 100 C for 2 min and cooled, 50 to 100 μl of dimethyl formamide and 25 to 50 μl of 0.2 M dansylchloride in acetonitrile were added for incubation at 37 C for 1 h. The solutions were clarified at 600 g, layered onto Sephadex G25 columns (13 by 0.6 cm; 3.7-ml bed volume), and developed with sodium carbonate, pH 9. After the first 2 ml of eluate was voided, approximately 1 ml of dansylated protein (orange-fluorescence under UV light) was collected before the development of dansylic acid (bluish), dansyl amide, urea, SDS, and other low-molecular-weight contaminants. The dansyl proteins were lyophilized and portions were hydrolyzed in 6.1 N HCl at 105 C for only 5 h to prevent destruction of dansyl serine and dansyl proline; other portions were hydrolyzed for 24 h. After the dansyl amino acids were dried twice in vacuo

over NaOH pellets, they were dissolved in 10 to 50 μl of pyridine and chromatographed by the method of Woods and Wang as described by Gray (5) on polyamide sheets (5 by 5 cm) (Schleicher and Schüll). Approximately 0.1-μl samples containing about 0.05 nmol of hydrolyzed dansylated protein were spotted from fine capillary pipettes. The chromatograms were developed with water-90% formic acid (100:3 vol/vol) in the first dimension and benzene-acetic acid (9:1 vol/vol) followed by ethyl acetate-methanol-acetic acid (20:1:1 vol/vol/vol) in the second dimension. The developed chromatograms were examined under UV light. Dansyl amino acids were tentatively identified by comparison of R_f values with those in chromatograms prepared from known standards, followed by the more positive method of mixing the appropriate standard(s) directly into the unknown solution before chromatography.

Leucine aminopeptidase studies. Diisopropylfluoro-phosphate-treated leucine aminopeptidase (LAP) (Worthington Biochemical Corp.) was used to ascertain whether the dansyl aspartic acid obtained from VP_2 was from N-terminal aspartic acid or asparagine. BSA (Armour Pharmaceutical Co.), poly-L-asparagine (Sigma Chemical Co.) and leucylglycylglycine (LGG) (Mann Research Labs., Inc.) were used as controls. The BSA and LGG were each used in the presence and absence of 6 M urea. The enzyme was activated before assay with 10^{-3} M $MnCl_2$ in 0.02 M Tris buffer at pH 8.5 for 1 h at 40 C (9). Three nanomoles of LGG or 20 to 30 nmol of VP_2, poly-L-asparagine or BSA were incubated at 37 C for 24 to 60 h with 7 μg of activated LAP in a total of 0.63 ml of 0.005 M $MgCl_2$-0.02 M Tris, pH 8.5. When LGG was used as a test substrate, the enzyme was minimally inhibited by 2, 4, and 6 M urea. After incubation was completed, the reaction mixture was adjusted to pH 3 with 1 N HCl, clarified, and analyzed on the acidic column of a Phoenix model K 8000B amino acid analyzer. A 30 to 50 C temperature program was used to separate asparagine from serine.

RESULTS

The investigation first required the preparative separation of FMDV proteins. A high degree of resolution of VP_1, VP_2, VP_3, and $VP_{3a,3b}$ doublet of type A_{12} virus was effected by disc PAGE (Fig. 1). A similar resolution of type C_3 proteins occurred. However, VP_1 and VP_2 of type O_1 virus were not sufficiently separated to obtain them in pure form. After electrophoretic elution from the main gel sections, the proteins were dialyzed against pH 9-SDS-ME solution. Small aliquots (25 μl) were analyzed for purity and concentration (Materials and Methods) before the main part was submitted to N-terminal amino acid and sequence determinations.

Chromatogram tracings of the N-terminal dansyl amino acids of proteins from FMDV types A_{12}, O_1, and C_3 are shown in Fig. 2. The N-terminal amino acids of VP_1 and VP_2 appeared to be glycine and aspartic acid in both

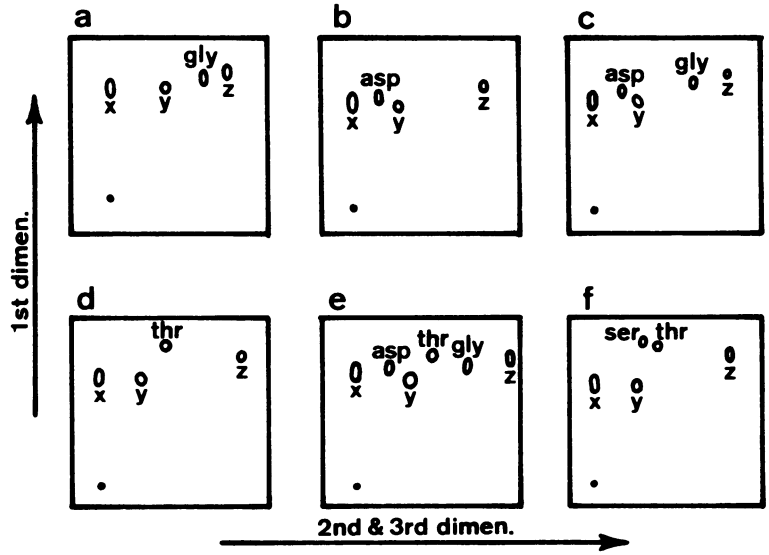

FIG. 2. *Chromatograms on polyamide sheets of N-terminal dansyl amino acids of FMDV proteins. (Dansylic acid [x], O-dansyl tyrosine [y] and dansylamide [z] spots were traced to serve as markers.) (a) VP_1 of type A_{12} virus. (Identical result for type C_3.) (b) VP_2 of type C_3 virus. (Identical result for type A_{12}.) (c) $VP_{1,2}$ of type O_1 virus. (d) VP_3 of type O_1 virus. (Identical results for types A_{12} and C_3.) (e) Total unresolved proteins of type A_{12} virus. (When these proteins were first maleylated, no N-terminal dansyl amino acids were found.) (f) $VP_{3a,3b}$ of type O_1 virus. (Identical result for type A_{12}.)*

virus types A_{12} and C_3 (Fig. 2a, b), and threonine was found to be N-terminal in VP_3 of all three viruses (Fig. 2d). In accord with the findings for virus types A_{12} and C_3, the N-terminal glycine and aspartic acid in the $VP_{1,2}$ mixture from type O_1 virus (Fig. 2c) very likely are also derived from VP_1 and VP_2, respectively. Also, Adam and Strohmaier (1) have reported that these two amino acids are N-terminal in VP_1 and VP_2 of virus type O_1 Kaufbeuren. The N-terminal amino acids, glycine, aspartic acid, and threonine, were found also in the total unresolved proteins of type A_{12} virus (Fig. 2e); none were found, however, when the proteins were first blocked with maleyl groups. The $VP_{3a,3b}$ doublet from both FMDV types A_{12} and O_1 yielded N-terminal threonine and serine (Fig. 2f). This indicates that one of the cleavage products of VP_3 in the virus retains the N-terminal region of VP_3, whereas the other one, which is generated from tryptic cleavage near the midpoint of VP_3, has N-terminal serine.

Determination of N-terminal amino acids by dansylation does not, by itself, differentiate between aspartic acid and asparagine because during hydrolysis of the dansylated protein in 6.1 N HCl, dansyl asparagine would be converted to dansyl aspartic acid. Evidence that the N-terminal residue in VP_2 is most probably aspartic acid and not asparagine was obtained by LAP treatment of VP_2 of type A_{12} virus, selected synthetic peptides and BSA (Table 1). The enzyme, which was active even in 6 M urea, released leucine and glycine from LGG and asparagine from poly-L-asparagine. It had no effect, however, on VP_2 nor on BSA, which has been found by King and Spencer (7) to have N-terminal aspartic acid. The resistance of N-terminal aspartic acid to LAP has been described (9). The results from the LAP reactions and from dansylation indicate quite clearly that the N-terminus of VP_2 is aspartic acid and not asparagine.

Sequences of amino acids in FMDV proteins determined by Edman degradation and dansylation are shown in Table 2. The similarities in the N-terminal regions of the virus proteins of virus types A_{12}, O_1, and C_3 apparently extend beyond the N-terminus at least for VP_1 and VP_3. VP_1 of each virus type appears to have the common sequence (Gly-ile-phe---) and VP_3 (Thr-thr---) before any divergence. Results for VP_2 were equivocal in residues beyond the N-terminal aspartic acid.

DISCUSSION

Virus types A_{12} strain 119, O_1 Brugge, and C_3 Resende all appear to have the same N-termini, glycine for VP_1, aspartic acid for VP_2, and threonine for VP_3; moreover, the similarity among virus types appears to extend further into the protein chains, at least for VP_1 (Gly-ile-phe---) and VP_3 (Thr-thr---). In addition, the present investigation provides evidence that the dansyl aspartic acid found for VP_2 of FMDV type A_{12} strain 119 is actually derived from N-terminal aspartic acid and not from asparagine. Also, tryptic cleavage of VP_3 in virus type A_{12} or spontaneous cleavage (presumably also by trypsin) in O_1 virus yields two fragments in each virus, VP_{3a} and VP_{3b}, one having apparently the N-terminal threonine of VP_3 and the other having N-terminal serine generated by the cleavage.

The N-terminal amino acids in the major

TABLE 1. *Reaction of leucine aminopeptidase with type A_{12} VP_2, synthetic peptides, and BSA*

Substrate[a]	Amino acids released
VP_2	None
LGG[b]	Leucine and glycine
VP_2 + LGG	Leucine and glycine
BSA[b,c]	None
Poly-L-asparagine	Asparagine

[a] Incubated with enzyme at 37 C for 24 h or longer; VP_2 was used with SDS and with SDS removed by precipitation with 0.2 M KCl or by the method of Weber and Kuter (17).

[b] Same result when the substrates were first denatured in 6 M urea.

[c] BSA has N-terminal aspartic acid (7).

TABLE 2. *N-terminal amino acid sequences[a] determined for proteins from FMDV types A_{12}, O_1, and C_3*

Virus type	VP_1	VP_2	VP_3
A_{12} strain 119	(Gly-ile-phe,pro,val---)	(Asp,X[b],met---)	(Thr-thr-ala-thr---)
O_1 Brugge[c]	(Gly, ile, phe---)	(Asp---)	(Thr-thr-ser---)
C_3 Resende	(Gly-ile-phe,ala---)	(Asp-leu---)	(Thr-thr---)

[a] Sequences linked by hyphens are more certain than those linked by commas.

[b] Unidentified.

[c] Results for the $VP_{1,2}$ mixture of type O_1 virus were assigned to VP_1 and VP_2 by analogy with sequences determined for type A_{12} and C_3 viruses.

capsid proteins of FMDV type O_1 Kaufbeuren were determined by Adam and Strohmaier (1) as being glycine for VP_1, asparagine or aspartic acid for VP_2, and threonine for VP_3. Other investigators have reported different results for other FMDVs. Laporte (8) found leucine, isoleucine, and threonine in a heat-stable mutant of type O virus; and Matheka and Dietzschold (Abstr. Proc. 2nd Int. Cong. Virol., 1972, p. 148, S. Karger, Basel) reported glycine, asparagine, and arginine in type A_2 Spain. It was suggested (1) that these chemical differences, if confirmed, might be related to type and subtype serological differences in FMDVs.

The present investigation does not support the suggestion of Adam and Strohmaier (1), because the N-termini in VP_1, VP_2, and VP_3 are glycine, aspartic acid, and threonine, respectively, in all of the FMDV types A_{12} strain 119, O_1 Brugge, and C_3 Resende. Moreover, almost identical results have been obtained for the capsid proteins of several other picornaviruses. The N-terminal amino acids in the unfractionated proteins of poliovirus strain Mahoney have been reported to be glycine, aspartic acid, and serine (Burrell and Cooper [4]). Likewise, N-terminal glycine, aspartic acid (or asparagine), and serine have been detected in the unfractionated capsid proteins of encephalomyocarditis virus (D. Rekosh, personal communication); analyses on the fractionated proteins indicate that these amino acids are N-terminal in the α-, β-, and γ-chains, respectively (D. Omilianowski and R. Rueckert, personal communication). The unfractionated capsid proteins of bovine enterovirus serotype VG-5-27 appear to have the same N-terminal amino acids as encephalomyocarditis virus (P. Carthew, personal communication), and analyses of procapsids and capsids suggest that both VP_1 and VP_4 of bovine enterovirus have N-terminal glycine.

A function for the N-terminal amino acids in several picornaviruses is suggested by their similarities. Specific N-terminal amino acids and more probably specific N-terminal sequences (e.g., [gly-ile-phe---] in VP_1 of FMDVs) may be required to maintain a uniformity of conformation in capsid-precursor polypeptides which permit primarily only those post-translational cleavages that lead to stable capsids and virions. The presence of threonine in FMDVs and serine in poliovirus, bovine enterovirus and encephalomycarditis virus would be accommodated by a point mutation in their RNAs of a single base (C↔G) in either of two triplet codons or of A↔U in four codons. The small difference in structure between threonine and serine might not be expected to interfere markedly with conformation or cleavage of precursor polypeptides or stability of virus capsids.

In accord with the findings for picornaviruses, the N-terminal regions are remarkably similar among the gs 30,000-dalton proteins of five murine, one feline, and RD-114 leukemia viruses (13, 14). All the leukemia gs antigens have the N-terminal sequence (pro-leu-arg---), and the first 24 amino acids of murine gs antigens show divergence only at residue 4. Within the first 15 amino acids of the gs antigens of feline and RD-114 viruses, divergences in sequence occur from each other and from the murine sequence only at residues 4, 6, 8, 9, and 10. Thus, the N-terminal regions in leukemia viruses gs proteins may also function, as proposed for picornavirus proteins, in controlling the conformation and post-translational cleavages of precursor polypeptides.

ACKNOWLEDGMENTS

We gratefully acknowledge the excellent technical assistance of Karl Axelson and James Dean.

ADDENDUM IN PROOF

The concept is proposed (H. L. Bachrach, submitted for publication) that post-translational cleavages of precursor proteins occur in external β-bends at protease-sensitive bonds to the left of the helix-breaker amino acids. This conclusion is based on the correlative observation that the N-terminal amino acids glycine, aspartic acid (or asparagine), serine, and threonine in picornavirus capsid proteins and proline in p30 of RNA tumor viruses are the helix-breaker amino acids which direct β-foldings in proteins.

LITERATURE CITED

1. Adam, K-H., and K. Strohmaier. 1974. Isolation of the coat proteins of foot-and-mouth disease virus and analysis of the composition and N-terminal end groups. Biophys. Biochem. Res. Commun. **61**:185-192.
2. Bachrach, H. L., J. B. Swaney, and G. F. Vande Woude. 1973. Isolation of the structural polypeptides of foot-and-mouth disease virus and analysis of their C-terminal sequences. Virology **52**:520-528.
3. Bachrach, H. L., R. Trautman, and S. S. Breese, Jr. 1964. Chemical and physical properties of virtually pure foot-and-mouth disease virus. Am. J. Vet. Res. **25**:333-342.
4. Burrell, C. J., and P. D. Cooper. 1973. N-terminal aspartate, glycine and serine in poliovirus protein. J. Gen. Virol. **21**:443-451.
5. Gray, W. R. 1972. End-group analysis using dansyl chloride, p. 121-138. *In* C. H. W. Hirs and S. N. Timasheff (ed.), Methods in enzymology, vol. XXV. Academic Press Inc., New York.
6. Gray, W. R. 1972. Sequence analysis with dansyl chloride, p. 333-344. *In* C. H. W. Hirs and S. N. Timasheff (ed.), Methods in enzymology, vol. XXV. Academic Press Inc., New York.
7. King, T. P., and M. Spencer. 1972. Amino acid sequence of the carboxyl terminal cyanogen bromide peptides of

bovine plasma albumin. Arch. Biochem. Biophys. **153**:627-640.
8. Laporte, J. 1969. The structure of foot-and-mouth disease virus protein. J. Gen. Virol. **4**:631-634.
9. Light, A. 1972. Leucine aminopeptidase in sequence determination of peptides, p. 253-262. *In* C. H. W. Hirs and S. N. Timasheff (ed.), Methods in enzymology, vol. XXV. Academic Press Inc., New York.
10. Lowry, O. H., N. J. Rosebrough, A. L. Farr, and R. J. Randall. 1951. Protein measurement with the Folin phenol reagent. J. Biol. Chem. **193**:265-275.
11. Maizel, J. V., Jr. 1971. Polyacrylamide gel electrophoresis of viral proteins, p. 179-246. *In* K. Maramorosch and H. Koprowski (ed.), Methods in virology, vol. 5. Academic Press Inc., New York.
12. Niederwieser, A. 1972. Thin layer chromatography of amino acids and derivatives, p. 60-99. *In* C. H. W. Hirs and S. N. Timasheff (ed.), Methods in enzymology, vol. XXV. Academic Press Inc., New York.
13. Oroszlan, S., T. Copeland, M. R. Summers, and R. V. Gilden. 1973. Feline leukemia and RD-114 virus group-specific proteins: comparison of amino terminal sequence. Science **181**:454-455.
14. Oroszlan, S., M. R. Summers, C. Foreman, and R. V. Gilden. 1974. Murine type-C virus group-specific antigens: interstrain immunochemical, biophysical, and amino acid sequence differences. J. Virol. **14**:1559-1574.
15. Polatnick, J., and H. L. Bachrach. 1964. Production and purification of milligram amounts of foot-and-mouth disease virus from baby hamster kidney cell cultures. Appl. Microbiol. **12**:368-373.
16. Wagner, G. G., J. L. Card, and K. M. Cowan. 1970. Immunochemical studies of foot-and-mouth disease. VII. Characterization of foot-and-mouth disease virus concentrated by polyethylene glycol precipitation. Arch. Gesamte Virusforsch. **30**:343-352.
17. Weber, K., and D. Kuter. 1971. Reversible denaturation of enzymes by sodium dodecyl sulfate. J. Biol. Chem. **246**:4504-4509.