JOURNAL OF VIROLOGY, Aug. 1980, p. 414-419
0022-538X/80/08-0414/06$02.00/0 Vol. 35, No. 2

Picornaviral VPg Sequences Are Contained in the Replicase Precursor

MARK A. PALLANSCH,[1*] OLEN M. KEW,[1†] ANN C. PALMENBERG,[1] FRED GOLINI,[2‡] ECKARD WIMMER,[2] AND ROLAND R. RUECKERT[1]

Biophysics Laboratory of the Graduate School and Department of Biochemistry of the College of Agricultural and Life Sciences, University of Wisconsin, Madison, Wisconsin 53706,[1] and Department of Microbiology, State University of New York at Stony Brook, Stony Brook, New York 11794[2]

It has previously been shown that the RNA replicase of encephalomyocarditis virus contains two virus-coded proteins, D and E, which are produced in two successive proteolytic steps: (i) C → D + ?; and (ii) D → p22 + E. It is here shown (i) that virus protein H (molecular weight, 12,000) is the previously unidentified product of the first step and (ii) that VPg, a protein linked covalently to the virion RNA, yields two tryptic peptides found in protein C but not in protein D. The results suggest that VPg is derived by cleavage of protein C and that protein H may be an intermediate. Preliminary experiments with VPg sequences in poliovial noncapsid protein 1b, the counterpart of encephalomyocarditis viral protein C, were inconclusive.

In all picornaviruses examined there is a small protein (VPg) covalently attached to the 5'-terminal residue of the genome RNA (8, 10, 13, 21). Circumstantial evidence suggests that VPg is encoded by the viral genome (8, 13); however, its function is presently unknown. Because VPg has been found linked to nascent RNA strands of the replicative intermediate (7, 16), it has been proposed that it is involved in RNA replication, possibly at the initiation step (7, 13, 16, 18). An unexpected result, however, was the observation that free VPg cannot be detected in extracts of poliovirus-infected cells (Golini and Wimmer, unpublished data). This led to the suggestion that VPg is generated by proteolytic cleavage from a larger precursor protein during initiation of RNA synthesis (16, 25), possibly from noncapsid viral protein 2 (NCVP 2; reference 6). (Nomenclature for polioviral protein is that of Summers et al. [24] as modified by Butterworth [1], and that for encephalomyocarditis [EMC] viral proteins is that of Butterworth et al. [2].) Alternatively, it has been suggested that VPg may represent the product of a less frequently translated second cistron (11). There is, however, no direct evidence supporting either of these proposals.

We now report evidence, based on tryptic mapping, that VPg is indeed a virus-coded protein and that it is derived from protein C of EMC virus, possibly via an intermediate protein, H.

† Present address: Division of Virology, Center for Disease Control, Atlanta, GA 30333.

‡ Present address: Department of Cellular Biology, Scripps Clinic and Research Foundation, La Jolla, CA 92037.

MATERIALS AND METHODS

Virus and cells. The EMC virus and the poliovirus type 1 (Mahoney) used have been described previously (9, 19). Virus was propagated in H-HeLa cells which were grown in medium B (15) containing 10% calf serum (K. C. Biologicals, Inc., Lenexa, Kans.) except where otherwise noted.

Preparation of radiolabeled EMC VPg. Infection of cells in suspension culture followed the procedure of McGregor et al. (14), with the exception of a 30-min virus attachment period. Virions were labeled by addition of either L-[4,5-^{3}H]lysine (5.0 mCi to 4 × 10^{7} cells) or L-[4,5-^{3}H]leucine (10.0 mCi to 4 × 10^{8} cells) at 2.5 h postinfection. Incubation (37°C) in the presence of label continued until 6 h postinfection, at which time the infected cell culture was subjected to three freeze-thaw cycles. The released virus was purified by pelleting, followed by banding in an isopycnic cesium chloride gradient as previously described (15). The virus was removed from cesium chloride by pelleting through a sucrose cushion. The RNA was extracted from virions and precipitated as described previously (22). The VPg-containing RNA was digested for 60 min at 37°C with a mixture of pancreatic RNase A (10 μg/ml), RNase T_1 (10 μg/ml), and RNase T_2 (5 U/ml) in 0.05 M sodium acetate (pH 4.5). VPg was dissolved in 0.01% sodium dodecyl sulfate by heating for 5 min at 100°C.

Preparation of radiolabeled poliovirus VPg. Poliovirus VPg labeled with L-[4,5-^{3}H]lysine was prepared as described previously (19).

Synthesis of protein in cell-free extracts. In vitro EMC RNA-dependent protein synthesis was carried out exactly as previously described (22, 23). Translation (90-μl reaction mixture) was in the presence of L-[4,5-^{3}H]leucine (24 μCi) for 60 min, at which time 5 μl of pancreatic RNase A (1 mg/ml) and 10 μl of cycloheximide (1 mg/ml) were added. After an additional 60 min of incubation, the sample was acetone

precipitated as previously described (17), dissolved in electrophoresis sample buffer, and dialyzed as previously described (2).

Synthesis of protein in infected cells. Lysates were prepared from 2×10^7 cells infected at a multiplicity of 100 PFU/cell. To obtain proteins D or NCVP 2, cells were exposed at 240 min postinfection to either L-[4,5-^{3}H]lysine or L-[4,5-^{3}H]leucine for 15 min. To obtain proteins C or NCVP 1b, cells were infected in the presence of 0.05% fetal calf serum and treated with 0.91 mM (EMC virus) or 0.70 mM (poliovirus) $ZnCl_2$ at 204 min postinfection as described (3). Either L-[*U*-^{14}C]lysine or L-[*U*-^{14}C]leucine was added from 210 min to 310 min postinfection. In both cases, cells were lysed at the end of the labeling period without washing, using hot sodium dodecyl sulfate, and the samples were dialyzed (2).

Preparative electrophoresis. Preparation of samples for electrophoresis has been described (15). Preparative electrophoresis for infected cell lysates to isolate proteins C, D, NCVP 1b, or NCVP 2 was the same as previously described (22) except that the gel composition was 4.9% acrylamide, 0.2% *N,N'*-methylenebisacrylamide, 0.1% sodium dodecyl sulfate, 0.1 M sodium phosphate at pH 7.2, 0.1% *N,N,N',N'*-tetramethylenediamine, 0.5 M urea, and ammonium persulfate at a final concentration of 0.05%. Preparative electrophoresis of in vitro reaction mixtures to isolate protein H was performed on 10% acrylamide gels containing 8 M urea (12).

Tryptic mapping. Proteins were mixed, reduced, alkylated, acid precipitated, digested with trypsin, and analyzed by cation-exchange chromatography exactly as previously described (12).

RESULTS

Evidence that VPg sequences of EMC virus reside in protein C. Evidence that VPg sequences are contained in protein C was obtained by digesting a mixture of the proteins which had been differentially labeled by incorporation of radiolabeled leucine. To obtain labeled VPg, purified RNA, extracted from virions grown in the presence of [^{3}H]leucine, was digested for 1 h with a mixture of RNases using conditions which had previously been shown to release the covalently attached protein of poliovirus in the form VPg-pUp (7, 13). Protein C was isolated by preparative electrophoresis of lysates from cells which had been exposed to [^{14}C]leucine from 3.5 to 4.5 h postinfection. Zinc chloride (0.91 mM) was included during the incorporation period to enhance the yield of EMC viral protein C (3), which is otherwise present in only small amounts. The mixture of VPg and protein C was alkylated, digested with trypsin, and chromatographed on a cation-exchange column to resolve the peptides (Fig. 1A).

VPg yielded a single leucine-containing peak which emerged at fraction 11 (Fig. 1A, arrow) just behind the flow-through peak (fraction 5). The digest from protein C also exhibited a peak in this position which was, however, somewhat broader on the right shoulder. That this broader peak is a compound one is shown by comparing the tryptic profile of protein C with that of its cleavage product (2), protein D (Fig. 1B). This profile shows that the tryptic peak from VPg is missing from protein D. There remains, however, a smaller peak in D which accounts for the noncoincidence of the peaks from protein C and VPg (seen in Fig. 1A, arrow). The difference between the profiles of proteins C and D was, however, not completely accounted for by the tryptic peptide of VPg. Thus one sees in the profile of protein C noncoincidences at fractions 27, 62, 93, and 98.

As seen in Fig. 1C, the tryptic profile of protein H accounts for all of these differences between C and D, including the tryptic peak emerging with that in the VPg profile (arrow). These data indicate that H is the previously unidentified cleavage complement of D (i.e., C → H + D). The data also strongly suggest that VPg is part of the H sequences contained in protein C but not in protein D.

The latter conclusion is greatly strengthened by a second experiment modified by substituting radiolabeled lysine for radiolabeled leucine (Fig. 2A). The lysine-labeled peak coincides with one found in protein C but not in protein D (Fig. 2B), and this peak is different from the leucine-labeled peak in Fig. 1. Thus VPg exhibits two tryptic peaks, one containing lysine but no leucine and a second that contains leucine but no lysine, which coincide with tryptic peaks from protein C but not from protein D.

VPg in polioviral proteins. The polioviral homolog of protein C is NCVP 1b, which is cleaved to NCVP 2 plus a still unidentified protein. Based upon an amino acid analysis of VPg, the RNA sequence coding for VPg has been determined (N. Kitamura, C. Adler, J. Martinko, S. G. Nathenson, and E. Wimmer, submitted for publication). This RNA sequence is located approximately 2,000 nucleotides from the 3′ end of the poliovirus genome, that is, within a segment of RNA coding for NCVP 1b. Moreover, the amino acid sequence of VPg, as deduced from the RNA sequence, predicts that digestion of VPg with trypsin should yield three peptides. These three peptides have indeed been observed (Fig. 3A).

To determine whether the VPg sequences of poliovirus could be demonstrated in NCVP 1b, the proteins, differentially labeled with lysine, were compared as before. In this case polioviral VPg exhibited three tryptic peaks, a, b, and c (Fig. 3A). None coincided exactly with the major peaks in NCVP 1b; moreover, peak a from VPg appeared to be entirely absent from NCVP 1b.

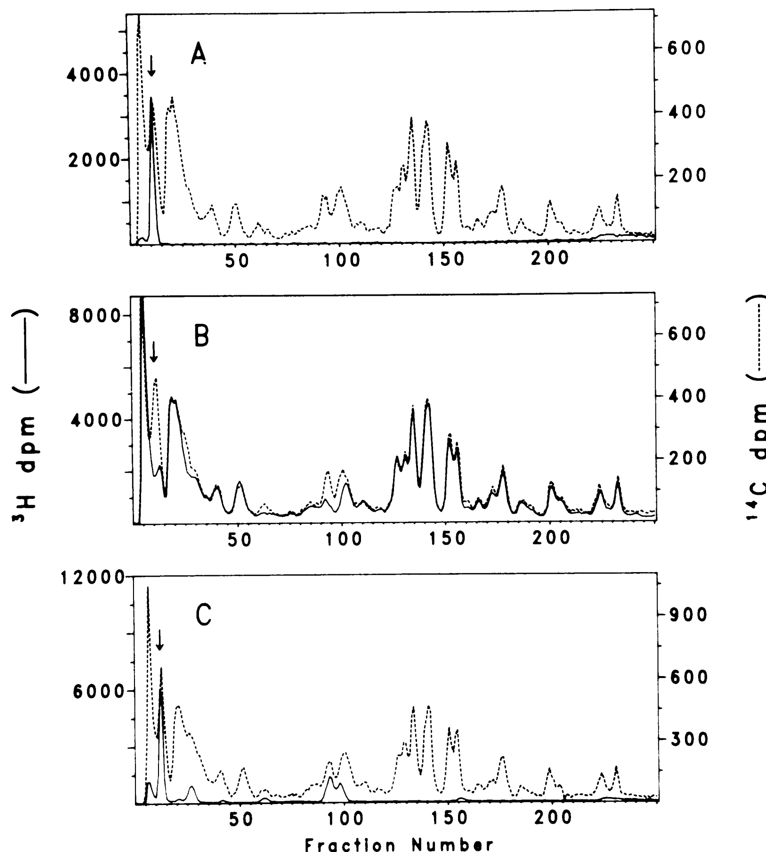

FIG. 1. *Tryptic profiles of leucine-labeled EMC proteins C, D, H, and VPg. Differentially labeled proteins were mixed, digested with trypsin, chromatographed on a cation-exchange column, and assayed for radioactivity as described in the text. (A) ^{3}H-labeled VPg (——) versus ^{14}C-labeled protein C (-----). (B) ^{3}H-labeled protein D (——) versus ^{14}C-labeled protein C (-----). (C) ^{3}H-labeled protein H (——) versus ^{14}C-labeled protein C (-----). Protein H was isolated from a cell-free translation product (see the text). Scales were adjusted such that the ratio of their heights equaled the specific activity of the proteins (disintegrations per minute [dpm] per dalton).*

These results do not exclude the possibility, however, that VPg peaks b and c are members of compound tryptic peaks from NCVP 1b, and that VPg peak a represents the new tryptic peptide produced when NCVP 1b is cleaved by the viral processing enzyme. Alternatively, peak a could be a tryptic peptide linked to -pUp. This nucleotidyl peptide would have no counterpart in NCVP 1b.

The possibility of compound peaks in NCVP 1b was further examined by comparing the tryptic profiles of NCVP 1b and its cleavage product NCVP 2 (Fig. 3B). The results obtained in this experiment were inconclusive. This was due in part to inability to resolve the VPg peptides definitively within the large number of peptides found in NCVP 1b and 2. A second difficulty was the presence of an unrelated protein, detected by analytical electrophoresis in slab gels, in our purified preparations of NCVP 1b. This contaminant was also manifested by tryptic profiles which became significantly simpler when NCVP 1b was enriched by treating infected cells with zinc (results not shown). When VPg and NCVP 2 were compared directly, two of the three VPg peptides were clearly resolved from any in NCVP 2, suggesting the absence of VPg sequences in this protein (not shown). Other experiments with lysine-labeled VPg excluded the presence of VPg sequences in polioviral capsid precursor NCVP 1a, in NCVP 5b, which is related to stable protein X (20), and in the small unmapped proteins 9a, 9b, and 10. These experiments were repeated using leucine as the radiolabel, but the results again failed to establish conclusive evidence for VPg sequences in 1b.

DISCUSSION

The data presented here indicate (i) that protein H is the previously unidentified product

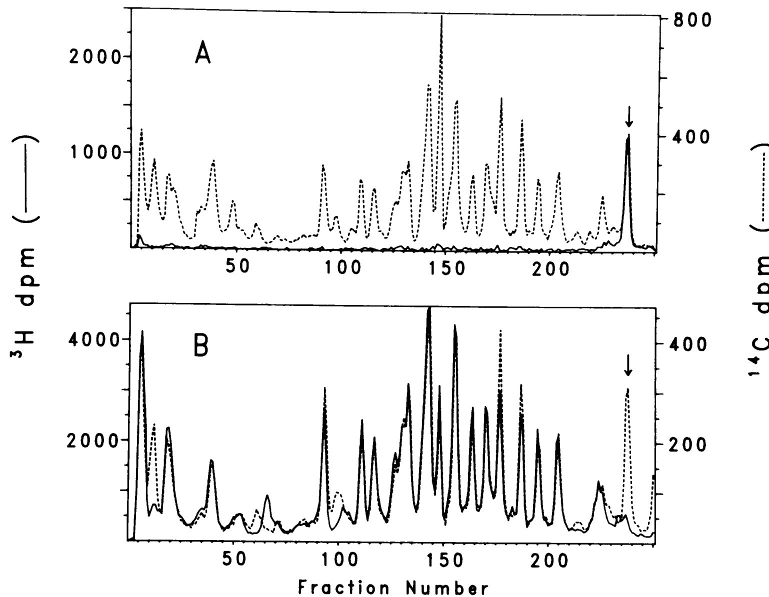

FIG. 2. *Tryptic profiles of lysine-labeled EMC viral proteins C, D, and VPg. (A) ^{3}H-labeled VPg (——) versus ^{14}C-labeled protein C (-----). (B) ^{3}H-labeled protein D (——) versus ^{14}C-labeled protein C (-----).*

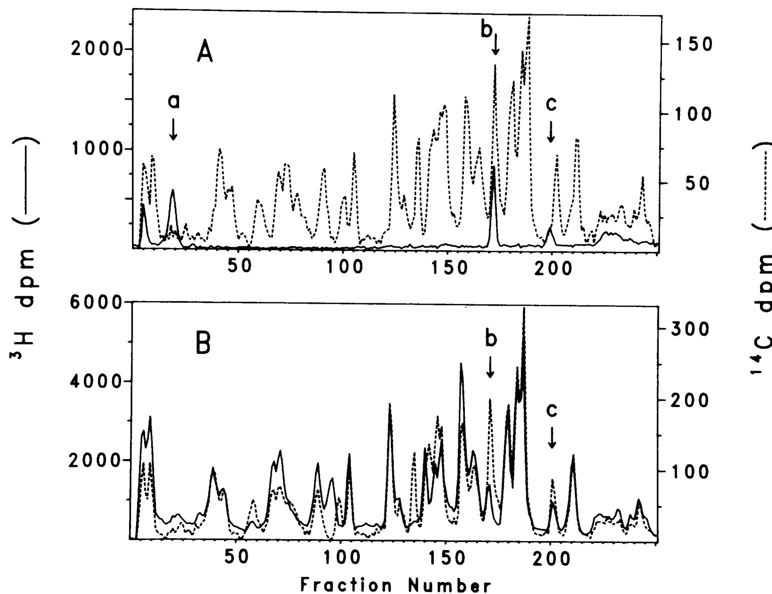

FIG. 3. *Tryptic profiles of lysine-labeled poliovirus proteins NCVP 1b, NCVP 2, and VPg. (A) ^{3}H-labeled VPg (——) versus ^{14}C-labeled NCVP 1b (-----). (B) ^{3}H-labeled NCVP 2 (——) versus ^{14}C-labeled NCVP 1b (-----).*

obtained when EMC viral protein C is cleaved to form protein D, and (ii) that VPg is part of protein H. These experiments are consistent with models of picornavirus replication proposed previously (7, 13, 16, 18, 25) which suggest involvement of VPg in initiation of RNA synthesis. They are also consistent with the proposal (17) that the polymerase precursor C donates VPg to the nascent RNA strand during initiation of synthesis. The finding of VPg sequences in protein H suggests that generation of VPg from protein C involves more than one proteolytic

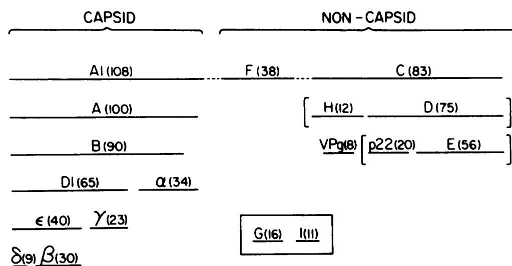

FIG. 4. *Processing map of EMC viral proteins revised to include proteins H and VPg. Brackets enclose proteins whose order is still unclear. Box encloses unmapped proteins. Electrophoretically determined molecular masses (parentheses) are given in kilodaltons (17). The electrophoretically determined molecular mass of VPg is 8,000 to 10,000 daltons (8); however, sequencing studies on polioviral-VPg suggest a mass less than 3,000.*

step. It remains to be clarified, however, whether H is an intermediate in the pathway of VPg addition to RNA.

The new data have been incorporated into the proteolytic processing map of EMC virus (Fig. 4). The position of protein H on the amino-terminal side of protein D is based on pactamycin mapping (4); however, no conclusion can be drawn from the tryptic data on the position of VPg within protein H.

If EMC and poliovirus follow the same protein cleavage pattern, then the presence of VPg in protein H could explain why polio VPg is not found free in the infected cell (25). Because it is known that the number of translations of the viral genome significantly exceeds the number of new viral RNA strands made in an infection (reviewed in N. Kitamura, C. Adler, and E. Wimmer, Ann. N. Y. Acad. Sci., in press, and reference 5), there should be a significant excess of VPg sequences. If the sequences contained in proteins C and H are only released when RNA synthesis occurs, free VPg may not be detectable in infected cells. In other words, protein H may represent the excess synthesis of VPg, although its role in RNA replication is presently unclear.

The major cleavage steps in the processing of replicase-related proteins of poliovirus and EMC virus are analogous (1b $\rightarrow$ 2 + ? is analogous to C $\rightarrow$ D + H; reference 20). Accordingly, we anticipate that VPg of poliovirus is contained in NCVP 1b and not in NCVP 2. The results presented here are compatible with this model, but inconclusive.

ACKNOWLEDGMENTS

We thank Daniel Omilianowski for excellent technical assistance in running the tryptic peptide column and Jane Hubertz for assistance in the preparation of VPg.

This work was supported by grant MV-33I from the American Cancer Society and by Public Health Service grants CA-08662, CA-16879, and AI-15122 from the National Institutes of Health. M.A.P. is a National Science Foundation Predoctoral Fellow.

LITERATURE CITED

1. **Butterworth, B. E.** 1973. A comparison of the virus-specific polypeptides of encephalomyocarditis virus, human rhinovirus-1A, and poliovirus. Virology **56**:439–453.
2. **Butterworth, B. E., L. Hall, C. M. Stoltzfus, and R. R. Rueckert.** 1971. Virus-specific proteins synthesized in encephalomyocarditis virus-infected HeLa cells. Proc. Natl. Acad. Sci. U.S.A. **68**:3083–3087.
3. **Butterworth, B. E., and B. D. Korant.** 1974. Characterization of the large picornaviral polypeptides produced in the presence of zinc ion. J. Virol. **14**:282–291.
4. **Butterworth, B. E., and R. R. Rueckert.** 1972. Gene order of encephalomyocarditis virus as determined by studies with pactamycin. J. Virol. **9**:823–828.
5. **Cooper, P. D., A. Steinor-Pryor, and P. J. Wright.** 1973. A proposed regulator for poliovirus: the equestron. Intervirology **1**:1–10.
6. **Dasgupta, A., M. H. Baron, and D. Baltimore.** 1979. Poliovirus replicase: a soluble enzyme able to initiate copying of poliovirus RNA. Proc. Natl. Acad. Sci. U.S.A. **76**:2679–2683.
7. **Flanegan, J. B., R. F. Pettersson, V. Ambros, M. J. Hewlett, and D. Baltimore.** 1977. Covalent linkage of a protein to a defined nucleotide sequence at the 5'-terminus of virion and replicative intermediate RNAs of poliovirus. Proc. Natl. Acad. Sci. U.S.A. **74**:961–965.
8. **Golini, F., A. Nomoto, and E. Wimmer.** 1978. The genome-linked protein of picornaviruses. IV. Difference in the VPg's of encephalomyocarditis virus and poliovirus as evidence that the genome-linked proteins are virus-coded. Virology **89**:112–118.
9. **Hall, L., and R. R. Rueckert.** 1971. Infection of mouse fibroblasts by cardioviruses: premature uncoating and its prevention by elevated pH and magnesium chloride. Virology **43**:152–165.
10. **Hruby, D. E., and W. K. Roberts.** 1978. Encephalomyocarditis virus RNA. III. Presence of a genome-associated protein. J. Virol. **25**:413–415.
11. **Humphries, S., F. Knauert, and E. Ehrenfeld.** 1979. Capsid protein precursor is one of two initiated products of translation of poliovirus RNA in vitro. J. Virol. **30**: 481–488.
12. **Kew, O. M., M. A. Pallansch, D. R. Omilianowski, and R. R. Rueckert.** 1980. Changes in three of the four coat proteins of oral polio vaccine strain derived from type 1 poliovirus. J. Virol. **33**:256–263.
13. **Lee, Y. F., A. Nomoto, B. M. Detjen, and E. Wimmer.** 1977. A protein covalently linked to poliovirus genome RNA. Proc. Natl. Acad. Sci. U.S.A. **74**:59–63.
14. **McGregor, S., L. Hall, and R. R. Rueckert.** 1975. Evidence for the existence of protomers in the assembly of encephalomyocarditis virus. J. Virol. **15**:1107–1120.
15. **Medappa, K. C., C. McLean, and R. R. Rueckert.** 1971. On the structure of rhinovirus 1A. Virology **44**:259–270.
16. **Nomoto, A., B. Detjen, R. Pozzatti, and E. Wimmer.** 1977. The location of the polio genome protein in viral RNAs and its implication for RNA synthesis. Nature (London) **268**:208–213.
17. **Palmenberg, A. C., M. A. Pallansch, and R. R. Rueckert.** 1979. Protease required for processing of picornaviral coat protein resides in the viral replicase gene. J. Virol. **32**:770–778.
18. **Pettersson, R. F., J. B. Flanegan, J. K. Rose, and D. Baltimore.** 1977. 5'-Terminal nucleotide sequences of poliovirus polyribosomal RNA and virion RNA are identical. Nature (London) **268**:270–272.
19. **Rothberg, P. G., T. J. R. Harris, A. Nomoto, and E.**

Wimmer. 1978. Genome-linked protein of picornaviruses. 5. O_4-(5'-uridylyl) tyrosine is bond between genome-linked protein and RNA of poliovirus. Proc. Natl. Acad. Sci. U.S.A. **75**:4868–4872.

20. **Rueckert, R. R., T. J. Matthews, O. M. Kew, M. Pallansch, C. McLean, and D. Omilianowski.** 1979. Synthesis and processing of picornaviral polyprotein, p. 113-125. *In* R. Perez-Bercoff (ed.), The molecular biology of picornaviruses. Plenum Press, New York.
21. **Sangar, D. V., D. J. Rowlands, T. J. R. Harris, and F. Brown.** 1977. Protein covalently linked to foot-and-mouth disease virus RNA. Nature (London) **268**:648–650.
22. **Shih, D. S., C. T. Shih, O. Kew, M. Pallansch, R. R. Rueckert, and P. Kaesberg.** 1978. Cell-free synthesis and processing of the proteins of poliovirus. Proc. Natl. Acad. Sci. U.S.A. **75**:5807–5811.
23. **Shih, D. S., C. T. Shih, D. Zimmern, R. R. Rueckert, and P. Kaesberg.** 1979. Translation of encephalomyocarditis virus RNA in reticulocyte lysates: kinetic analysis of the formation of virion proteins and a protein required for processing. J. Virol. **30**:472–480.
24. **Summers, D. F., J. V. Maizel, Jr., and J. E. Darnell, Jr.** 1965. Evidence for virus-specific noncapsid proteins in poliovirus-infected HeLa cells. Proc. Natl. Acad. Sci. U.S.A. **54**:505–513.
25. **Wimmer. E.** 1979. The genome-linked protein of picornaviruses: discovery, properties and possible functions, p. 175-190. *In* R. Perez-Bercoff (ed.), The molecular biology of picornaviruses. Plenum Press, New York.