JOURNAL OF VIROLOGY, Mar. 1984, p. 873–880
0022-538X/84/030873-08$02.00/0

Vol. 49, No. 3

Protein Processing Map of Poliovirus

MARK A. PALLANSCH,[1†] OLEN M. KEW,[1‡] BERT L. SEMLER,[2§] DANIEL R. OMILIANOWSKI,[1] CARL W. ANDERSON,[3] ECKARD WIMMER,[2] AND ROLAND R. RUECKERT[1*]

Biophysics Laboratory, Graduate School, and Biochemistry Department, College of Agricultural and Life Sciences, University of Wisconsin—Madison, Madison, Wisconsin 53706[1]; Department of Microbiology, School of Medicine, State University of New York at Stony Brook, Stony Brook, New York 11794[2]; and Department of Biology, Brookhaven National Laboratory, Upton, New York 11973[3]

Received 31 August 1983/Accepted 7 November 1983

Five previously unmapped proteins (5a, 7d, 8, 9b, and 10) were located on the proteolytic processing map of the polyprotein. One of the proteins, 9b, appears to be the sister fragment of a cleavage reaction (P3-9 → P3-9b + VPg). Two of the other newly mapped proteins, 8 and 10, have been identified as sister fragments of X-related proteins 3b and 5b; thus, P2-3b → P2-8 + P2-5b and P2-5b → P2-10 + P2-X. The remaining proteins, 5a and 7d, mapped in the 1b protein and appear to result from the cleavages P3-1b → P3-5a + P3-6b and P3-4b → P3-7d + P3-6b. These assignments account for over 95% of the total polioviral proteins and complete the mapping of the major processing pathways.

Synthesis of polioviral polyprotein (molecular weight, 247,000) is accompanied by extensive processing, which usually begins while still in the nascent state to form three major precursors called P1-1a, P2-3b, and P3-1b. These precursors are then cleaved into a variety of smaller products (Fig. 1A and B) that have been mapped by a variety of approaches including kinetic studies on the flow of radiolabeled amino acids through proteins synthesized in the presence and absence of the drug pactamycin (5, 23, 32, 33), tryptic analysis (13, 21, 25, 26, 37), immunoprecipitation (35), and radiochemical sequencing (14, 16, 26–28), which became feasible with the availability of the RNA sequence (14, 20, 22).

As a result of these studies, 19 of 26 electrophoretic components found in poliovirus-infected cells, after polyacrylamide gel electrophoresis of the radiolabeled proteins, have been correlated with positions in the protein processing map (Fig. 1C). We now report the identity and location of proteins representing the last three major gaps in this map (dotted lines). We also report the location of two minor proteins that appear to originate from previously unrecognized cleavages of proteins P3-1b and P3-4.

MATERIALS AND METHODS

Virus and cells. Poliovirus type 1 (Mahoney) was obtained from C. Pfau at Rensselaer Polytechnic Institute; C. Pfau had obtained the virus from S. Penman at the Massachusetts Institute of Technology. A second Mahoney strain from E. Wimmer (24) was used for the mapping of VPg and P3-1b and for radiochemical microsequencing. This second virus is the strain sequenced recently by Wimmer's group (14). Virus used for tryptic mapping was progagated in H-HeLa cells that were grown in medium B (19) containing 10% added volume of calf serum (K. C. Biologicals, Inc., Lenexa, Kans.). The infection of cells in suspension was with 50 to 100 PFU per cell by the procedures of McGregor et al. (17), except that a 30-minute virus attachment period was used. Virus used for microsequencing was propagated in S3-HeLa cells as described previously (27).

Preparation and purification of radiolabeled poliovirus VPg. Virions were radiolabeled by the addition of L-[4,5-^{3}H]lysine (5.0 mCi to 4 × 10^7 infected cells) at 2.5 h postinfection. Cell-associated virus was harvested at 6 h postinfection, the virus was purified, and the VPg RNA was phenol extracted as described previously (21). The [^{3}H]lysine VPg RNA was used directly in tryptic peptide analysis.

Preparation of radiolabeled proteins for tryptic peptide analysis. Because processing rates differ for various proteins, it was necessary to utilize three types of labeling conditions in preparing the proteins used for tryptic analysis: (i) continuous labeling for short time periods (pulse-labeling), (ii) short labeling followed by the addition of a large excess of unlabeled amino acid and continued incubation (pulse-chase labeling), and (iii) continuous labeling in the presence of $ZnCl_2$.

The pulse-labeling protocol was used to isolate proteins P1-1a, P3-1b, P1-3a, P2-3b, P3-5a, P2-5b, P2-X, P2-8, P3-9a, and P2-10. This was carried out by adding radiolabeled amino acids at 3.5 h postinfection. Incorporation was for 15 or 30 min. The cells were then lysed with sodium dodecyl sulfate as described below.

The pulse-chase labeling protocol was used to isolate proteins P3-2, P1-3c, P3-4b, P3-6a, P3-6b, P3-7d and P3-9b. Incorporation of radiolabeled amino acids was for 30 min beginning at 3 h postinfection. At this point cells were sedimented (820 × *g* for 3 min) and then suspended in the same volume of medium A (19) with leucine and lysine at 0.1 mM, a 100-fold excess over the initial concentration of the radiochemical. The cells were then incubated for an additional 90 min before harvesting by sodium dodecyl sulfate lysis.

The third procedure, labeling viral proteins in the presence of $ZnCl_2$ (6), was used to isolate rapidly cleaved precursors. This was used primarily to obtain relatively large amounts of protein P3-1b. Fetal calf serum (0.05%) and $ZnCl_2$ (1.0 mM in 0.1 mM HCl, final concentration) were added 6 min before the addition of the radiochemical at 3.5 h postinfection. Incorporation was allowed to continue for 60 min, at which time cells were harvested by sodium dodecyl sulfate lysis. The concentration of zinc (1 mM) chosen yielded the highest amount of protein P3-1b.

* Corresponding author.
† Present address: Rockefeller University, New York, NY 10021.
‡ Present address: Division of Viral Diseases, Center for Infectious Diseases, Centers for Disease Control, Atlanta, GA 30333.
§ Present address: Department of Microbiology, College of Medicine, University of California, Irvine, CA 92717.

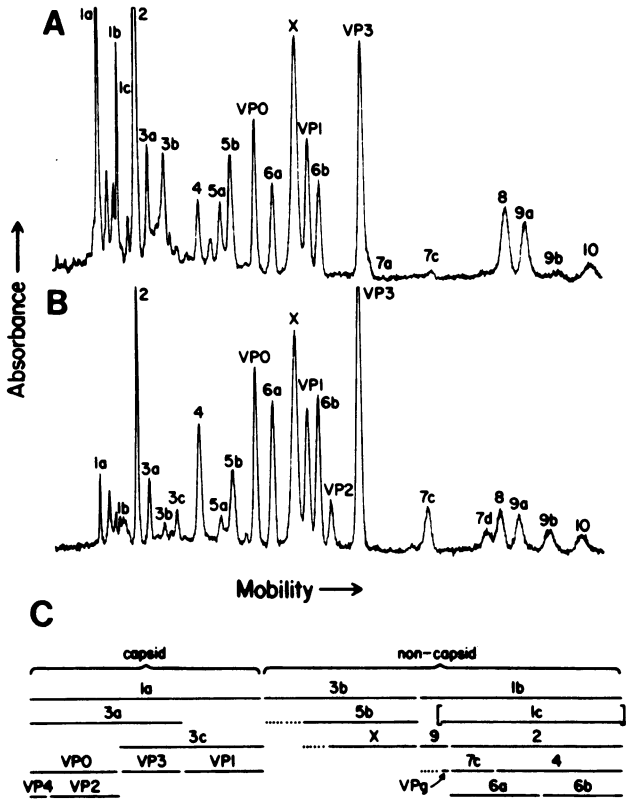

FIG. 1. Electrophoretic profiles and processing map of polioviral proteins found in lysates of virus-infected HeLa cells (A) after a 10-min labeling period with [^{35}S]methionine and (B) as in A, followed by a 90-min chase. Electrophoresis was carried out as described in the text. The processing map of the polioviral polyprotein (C) is from reference 25 modified to reflect recent mapping of VPg-containing precursor 9 (4, 26) and the tentative (in brackets) location of protein 1c. Most cleavages occur at glutamine-glycine pairs (14, 26–28).

Four radiochemicals were used to label polioviral proteins: L-[4,5-^{3}H]leucine, L-[*U*-^{14}C]leucine, L-[4,5-^{3}H]lysine, and L-[*U*-^{14}C]lysine. A typical lysate was prepared by labeling 4×10^7 cells in a 10-ml culture in amino acid-deficient medium A (19) with 500 μCi/ml for the ^{3}H-labeled compounds or 50 μCi/ml for the ^{14}C-labeled compounds.

Cell lysates were made by sedimenting the infected cells ($820 \times g$ for 5 min), suspending them at 8×10^7 cells per ml in distilled water, and adding an equal volume of solubilizing solution (2% sodium dodecyl sulfate, 1 M urea, and 0.2% mercaptoethanol). The lysate was then placed in a boiling water bath for 5 min. The preparation of samples for electrophoresis by dialysis has been described previously (19).

Preparative electrophoresis for peptide mapping. Preparative electrophoresis of cell lysates was carried out with cylindrical gels (25 by 1.2 cm) of 10% polyacrylamide in phosphate buffer as described previously (30). Five percent polyacrylamide was used for preparative electrophoresis of protein P3-1b (13, 21). After electrophoresis, gels were

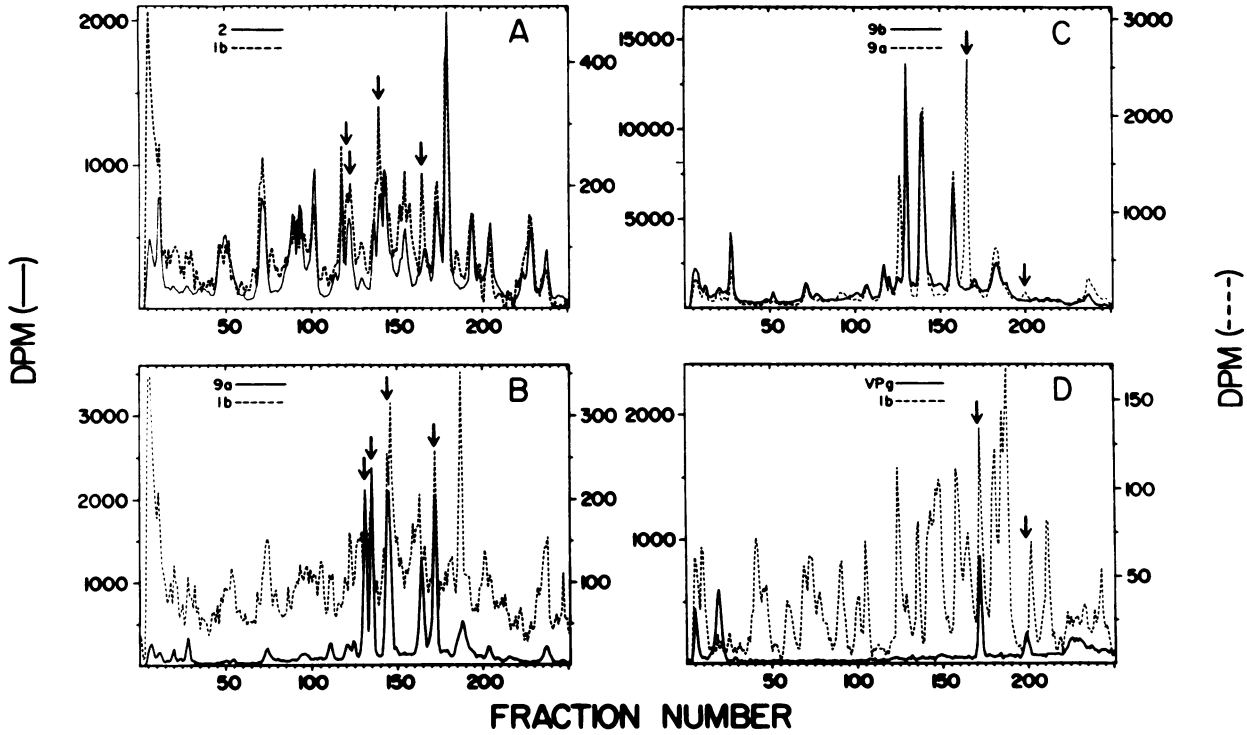

FIG. 2. Tryptic peptide analysis showing relationship of proteins 1b, 2, 9a, 9b, and VPg. Proteins were radiolabeled with [^{3}H]- or [^{14}C]lysine as described in the text. Protein 1b was radiolabeled with [^{3}H]lysine (A and B) or [^{14}C]lysine (D) in the presence of 1 mM $ZnCl_2$. Proteins 2 and 9a were labeled with [^{14}C]lysine; protein 9b was labeled with [^{3}H]lysine. Protein 1b was isolated on a 5% polyacrylamide gel; the others were labeled on a 10% gel (13). A mixture of differentially labeled proteins was reduced and alkylated before trypsin treatment; the digest was chromatographed on a column of Chromobeads type P and assayed for radioactivity as described previously (13, 31).

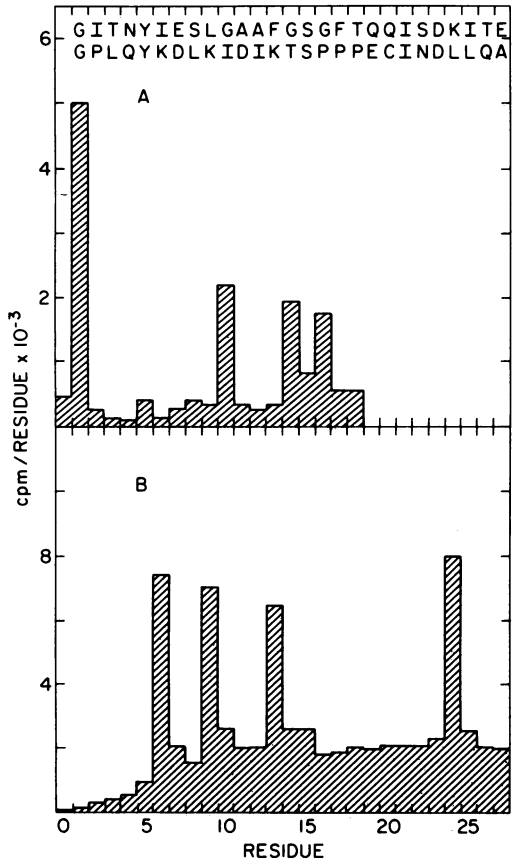

FIG. 3. Amino-terminal sequence analysis of polioviral peak 10 from electrophoretic system II. Two preparations were analyzed. One (A) was labeled with [^{3}H]glycine; the other (B) was labeled with [^{3}H]lysine. After elution from the gel the protein, about 2×10^5 cpm of lysine label or 4.1×10^4 cpm of glycine label, was subjected to automated sequential Edman degradation as described previously (3, 27). The yield of glycine in position 1 was roughly twice that found in positions 10, 14, and 16, consistent with the presence of roughly equimolar proportions of two different proteins each carrying an amino-terminal glycine residue. During the first cycle (residue 0) the coupling reagent phenylisothiocyanate was omitted to distinguish sample washout from sequential degradation in the first cycle. The total radioactivity recovered in the amino acid fraction after each cycle has been plotted as a function of the cycle (residue) number. The amino terminal sequences of P2-5b (top line) and P3-1b (second line) are given at the top of panel A (13).

fractionated in 1-mm segments with an automatic gel fractionator (10). The proteins were recovered by leaching from the gel (30), and samples of each fraction were counted for radioactivity. The resulting gel profile was used to identify proteins, and the fractions corresponding to the various proteins were pooled by centrifugation through filters as described previously (13). If any doubt existed as to a protein peak's identity, that protein was not used without further direct electrophoretic comparisons to a reference lysate. The criteria for peak identity were (i) reference to the electrophoretic profiles of Butterworth (5), (ii) kinetic behavior of the protein during pulse-chase analysis, and (iii) relative amount of protein (e.g., tracings in Fig. 1).

Tryptic peptide mapping. Proteins differentially labeled with [^{3}H]- or [^{14}C]lysine were mixed, reduced, alkylated, acid precipitated, digested with tolylsulfonyl phenylalanyl

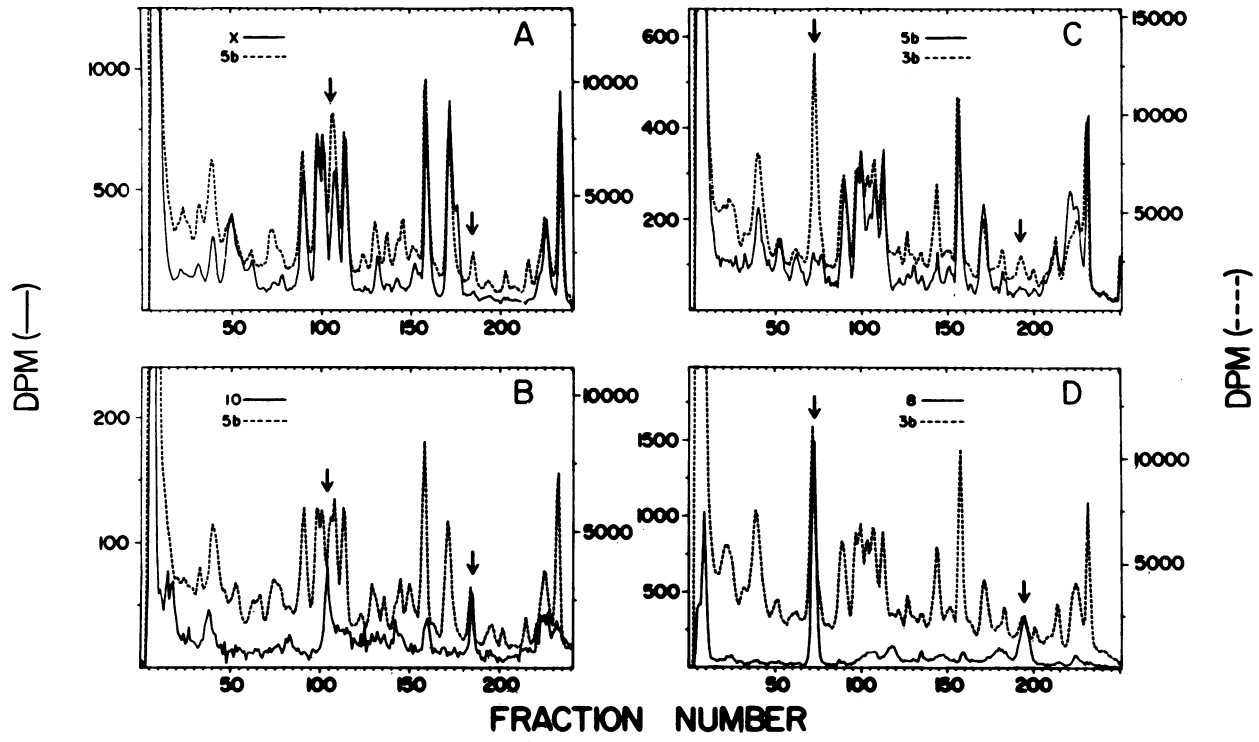

FIG. 4. Tryptic peptide analysis of leucine-labeled proteins 3b, 5b, X, 8, and 10. Arrows in panels A and B indicate differences between proteins 5b and X, which appear to be accounted for by peptides from 10. Arrows in panels C and D indicate two peptides from protein 8 that are found in 3b, but not in 5b.

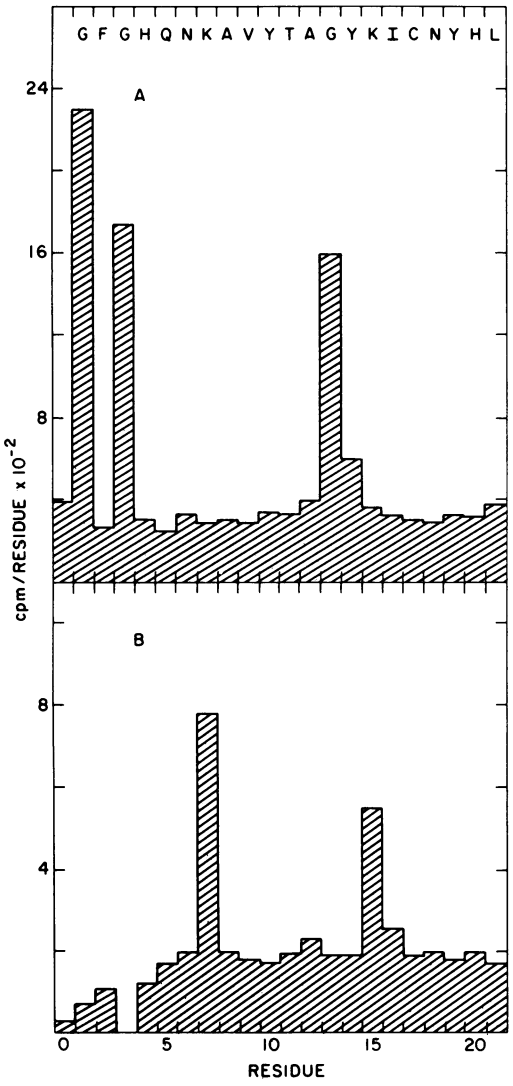

FIG. 5. Amino-terminal sequence analysis of protein 8 radiolabeled with [^{3}H]glycine (A) or [^{3}H]lysine (B). The amount of radioactivity applied to the sequencer was 10^5 cpm for glycine and 4.2×10^4 cpm for lysine. A mechanical malfunction resulted in the loss of residue 3 from the lysine-labeled protein. The predicted amino acid sequence for the amino terminus of protein P2-3b is indicated at the top of panel A.

chloromethyl ketone (TPCK)-trypsin, and analyzed by cation-exchange chromatography exactly as previously described (13, 30). If protein sample volumes exceeded 2.2 ml they were reduced in volume by lyophilization before tryptic analysis.

Preparation of radiolabeled proteins for radiochemical microsequencing. Suspension cultures of S3 HeLa cells (2.5 ml at 5×10^6 cells per ml) were infected at a multiplicity of 50 PFU per cell. Infected cell cultures were labeled from 3 to 5 h postinfection (400 μCi of [^{3}H]glycine or [^{3}H]lysine per ml) and were harvested immediately after the labeling period. Infected cell lysates were made, the lysates were run on preparative gels, and the proteins were electrophoretically eluted from gel strips as described previously (27).

Radiochemical microsequencing. Radioactively labeled proteins that had been polyacrylamide gel purified, were subjected to automated Edman degradation in a Beckman 890C sequencer as described previously (2, 27, 28).

RESULTS

Isolation and identification of polioviral proteins for tryptic peptide analysis. Two different sodium dodecyl sulfate-polyacrylamide gel systems were used in this study, reflecting the different laboratories in which the work was carried out. The Laemmli system (discontinuous Tris-glyine) was used to isolate proteins used for microchemical sequencing, whereas a modified Maizel system (continuous, phosphate buffered) was used for proteins analyzed by tryptic comparisons. The electrophoretic patterns obtained with these two systems, with one important exception to be discussed below, were generally very similar (compare Fig. 2 in reference 11 with Fig. 1 in this work). With the exception of proteins 7a and 1c, all of the numbered peaks shown in Fig. 1 were characterized by tryptic mapping by procedures described previously (13, 25). These tryptic maps confirmed all of the previously reported relationships (Fig. 1C), and the results will therefore not be presented except as they relate to the identification of previously unmapped proteins.

Characterization of proteins 9a and 9b. Earlier studies have shown that proteins 1b, 2, and 4 are related by the cleavage pathway 1b → 2 → 4 (5, 32, 33) via the pathways 2 → 7c + 4 (25, 27) and 1b → 9 + 2 (4, 26). Protein P3-9, also called preVPg-3 (4), contains at its carboxy-terminal end sequences of the genome-linked protein VPg (2, 26). One of the points clarified by the present study is the relationship of protein P3-9 to peaks 9a and 9b in Fig. 1.

As shown in Fig. 2A, the tryptic profile of protein 2 (solid line) lacks four peaks (arrows) found in protein 1b (dotted line). These four missing peaks were present in digests of peak 9a (solid line in Fig. 2B). The profiles in Fig. 2C show that protein 9b (solid line) is closely related to 9a (dotted line), but lacks two peaks found in digests of VPg (Fig. 2D).

The presence of VPg sequences in protein 9a indicates that the latter protein corresponds to protein P3-9, whereas the absence of VPg sequences in protein 9b suggests that it is the cleavage complement produced by removal of VPg; thus 9a → 9b + VPg. This conclusion was reinforced by microchemical sequencing studies. These studies also revealed that the reason protein 9b had not previously been detected in the Laemmli gel system is because it migrates in that system with band 10. Thus the pattern of glycine and lysine residues obtained from sequence analysis of band 10 could not be matched to any single stretch of amino acids predicted by the polioviral RNA sequence (Fig. 3). However the sequencing profiles of peak 10 matched perfectly the amino-terminal sequences predicted for an equimolar mixture of proteins P2-5b and P3-1b. Thus in the first 25 residues, glycine appeared at positions 1, 10, 14, and 16 exactly as predicted by the RNA sequence (14) for the amino terminus of protein P2-5b (first line at the top of Fig. 3); in the corresponding 25 residues of protein P3-1b glycine appeared only at position 1 (second line in Fig. 3). Similarly lysine occurred only once, at position 24, in the first 25 residues of protein P2-5b, but it occurred three times in protein P3-1b, at positions 6, 9, and 13. We conclude that protein peak 10 on Laemlli gels contained an equimolar mixture of two proteins. Because peak 10 is distinct from the previously identified protein 9 it must contain the protein corresponding to 9b in gel system I (Fig. 1). We show below that the other component probably corresponds to protein 10.

Mapping of protein 10. The presence in peak 10 of sequences from the amino-terminal end of protein P2-5b sug-

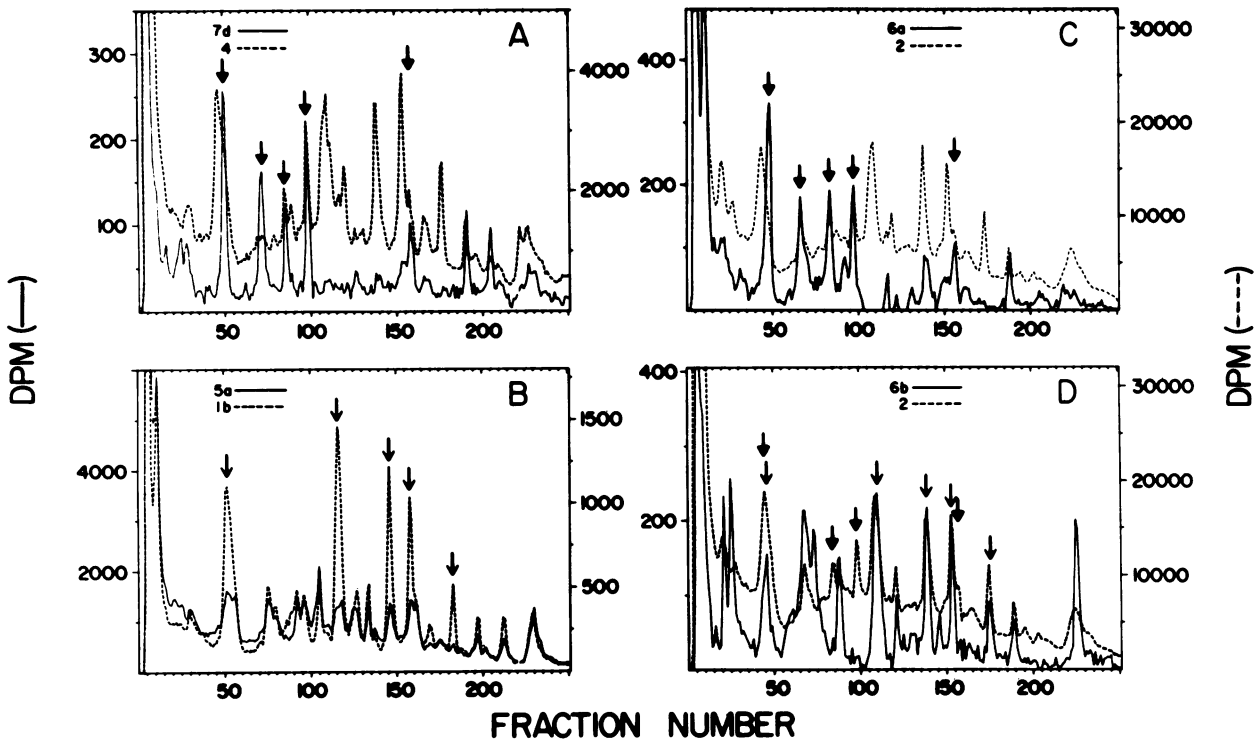

FIG. 6. Tryptic peptide analysis of proteins P3-1b, P3-2, P3-4, 5a, and 7d. Proteins P3-1b, P3-6a, P3-6b, and 7d were labeled with [^{14}C]leucine; the others, P3-2, P3-4, and 5a, were labeled with [^{3}H]leucine. Filled arrows in panels A, C, and D indicate peptides from 7d that are found in proteins P3-4 and P3-6a, but not in protein P3-6b. Open arrows in panels B and D indicate peptides of P3-6b and P3-1b that are not found in protein 5a.

gested that it corresponded to the heretofore unidentified sister fragment of protein P2-X in the cleavage P2-5b → P2-10 + P2-X. The electrophoretically determined molecular weight of protein 10 is also consistent with this hypothesis. Attempts to confirm this identity by using tryptic peptide analysis were inconclusive, but consistent with this possibility. As shown in Fig. 4A, the tryptic profile of protein P2-5b (dotted line) contained a number of minor peaks not found in protein P2-X (solid line). At least two of these missing peaks (arrows) appeared to be present in the profiles of protein 10 (FIg. 4B). Thus the cumulative evidence, based on sequence analysis, apparent molecular weight, and tryptic analysis provides consistent evidence in favor of identification of protein 10 as the sister fragment produced when protein P2-5b is cleaved to form protein P2-X.

Mapping of protein 8. As shown in Fig. 4C, protein P2-3b (dotted line) contains at least two tryptic peptides lacking in protein 5b (solid line). These two peaks, shown by arrows, are found in protein 8 (compare Fig. 4C and D), suggesting that protein 8 is the sister of protein P2-5b in the cleavage reaction P2-3b → P2-8 + P2-5b. This assignment was confirmed by microchemical sequencing with protein 8 radiolabeled with [^{3}H]glycine or [^{3}H]lysine (Fig. 5). Thus glycine occurred at positions 1, 3, and 13, whereas lysine appeared at positions 7 and 15, exactly as predicted for the unique amino-terminal sequence of protein P2-3b (Fig. 5A).

Identification of protein 7d as a cleavage product of protein P3-2. As shown in Fig. 6A, tryptic digests of protein 7d (solid line) contain five major peaks (filled arrows) that coincide with peaks in protein P3-4b (dotted line). These same five peaks also occur in protein P3-6a (Fig. 6C), but not in protein P3-6b (filled arrows, Fig. 6D). This result indicates that protein 7d is the product of a previously unrecognized cleavage of protein P3-4b or P3-6a.

Origin of protein 5a. Figure 6B shows that all five major peaks in the tryptic profile of protein 5a (solid line) are also present in the tryptic profile of protein 1b (dotted line). In addition the profile of 5a also contains a number of tryptic peaks (filled arrows) common to protein 7d (compare solid arrows in Fig. 6A and B). Finally, a number of the remaining 1b peptides, i.e., those not present in the 5a profile, coincide with peaks observed in digests from protein P3-6b (compare Fig. 6B and D, open arrows). These results indicate that peak 5a originates from cleavage of protein P3-1b; thus P3-1b → P3-5a + P3-6b. The kinetic behavior of peak 5a in a pulse-chase experiment (Fig. 1) is compatible with this hypothesis; so is its electrophoretic mobility, which is about that expected for the 49,000-dalton mass computed for this protein from the RNA sequence (Table 1).

DISCUSSION

A revised processing map of the polioviral polyprotein is illustrated in Fig. 7. We have divided the polyprotein into three regions, P1, P2, and P3 (14), allowing for easier recognition of the numerous cleavage products. For example, the proteins are designated P1-1a, P2-3b, P3-7c, etc.

All 27 proteins in this map can be accounted for by 12 cleavage sites (Table 1). Nine of these cleavages, which are probably carried out by virus-coded protein P3-7c (11), occur in glutamine-glycine (QG) pairs, of which there are a total of 13 (14). One of these sites, QG-11, is cleaved only rarely, yielding a protein, designated P3-4a (29), which is associated with the viral RNA replication complex (9). In addition to the nine known QG cleavage sites, three other

TABLE 1. Proteins of poliovirus

Protein	L434 name[a]	Mol wt[b]	NH_2 terminus or NH_2 terminal cleavage site[c]	C terminus or COOH terminal cleavage site[c]
P1-1a	1	97,247	Blocked G	YG1
P3-1b	3	84,234	QG-8	P
3b/9	2-3AB	77,000	YG-1	QG-10
P3-1c	3BCD	76,000[d]	QG-9?	P?
P3-2	3CD	72,132	QG-10	P
P1-3a	1ABC	63,786	Blocked G	QG-3
P2-3b	2	64,953	YG-1	QG-8
P1-3c	1CD	59,930	QG-2	YG-1
P3-4a		60,000	QG-11	P
P3-4b	3D	52,481	QG-12	P
X/9	2C-3AB	50,000	GQ-6	GQ-10
P3-5a[e]		48,550	QG-8	YG-2
P2-5b	2BC	48,273	QG-5	QG-8
VP0	1AB	37,352	Blocked G	QG-2
P3-6a	3C′	36,450	QG-10	YG-2
P2-X	2C	37,555	QG-6	QG-8
VP1	1D	33,521	QG-3	YG-1
P3-6b	3D′	35,700	YG-2	P?
VP2	1B	29,985	NS-2	QG-2
VP3	1C	26,410	QG-2	QG-3
P2-7a	2AB	25,500[d]	YG-1?	QG-6?
P3-7c	3C	19,669	QG-10	QG-12
P3-7d[e]		16,780	QG-12	YG-2
P2-8	2A	16,680	YG-1	QG-5
P3-9	3AB	12,100	QG-8	QG-10
P3-9b	3B	9,750	QG-8	QG-9
P2-10	2B	10,720	QG-5	QG-6
VP4	10	7,385	Blocked G	NS-2
VPg	3B	2,354	QG-9	QG-10

[a] From R. R. Rueckert and E. Wimmer, submitted for publication.

[b] The molecular weight given for each protein is that predicted from the RNA sequence (14) using the indicated amino and carboxy termini.

[c] Sites are as shown in Fig. 7. Glycine (G) is the amino terminus of the polyprotein; phenylalanine (P) is the carboxy terminus. Other abbreviations: N, asparagine; Q, glutamine; S, serine; Y, tyrosine; and ?, uncertain.

[d] Apparent molecular weight was determined by electrophoretic mobility in the continuous phosphate-buffered gel system.

[e] The molecular weight estimate of P3-7d is based on the difference in molecular weights between P3-6a and P3-7c. The molecular weight estimate of protein 5a is based on the sum of the molecular weights of proteins 6a and 9a (see the text).

cleavage sites distinct from QG cleavages have been identified. One is an asparagine-serine pair in VP0 (16), which is cleaved by an unidentified protease to form VP4 and VP2 during the virion maturation step (12). The other two cleavage sites are tyrosine-glycine (YG) pairs, one of which is located between proteins 1a and 3b (14); the other is located between 6a and 6b (11). Inhibition of viral processing in vitro by antibodies against protein 7c fail to inhibit YG cleavages (11), indicating involvement of another protease, possibly a cellular protease with a chymotrypsin-like specificity (15). For all three types of cleavages (QG, YG, asparagine-serine) proteolytic processing occurs by a single cleavage event without further trimming of the newly generated amino acid termini (2, 8).

In the interest of standardization we have adopted the nomenclature of Kitamura et al. (14). Thus the protein earlier called 9a in this report is hereafter designated protein 9. Similarly protein 4 is hereafter designated 4b to distinguish it from 4a (9).

P2 family of proteins. The identification of protein P2-8 as the cleavage complement of P2-5b and of P2-10 as the sister fragment of X (Fig. 7) completes the last two major gaps in the P2 region. These assignments do not however necessarily imply that protein P2-X is derived only through cleavage of precursors P2-3b and P2-5b, since kinetic studies (M. A. Pallansch and R. R. Rueckert, manuscript in preparation) indicate that the polyprotein can also be cleaved at sites QG-5 and QG-6 while still in the nascent state to release proteins P2-8, P2-10, and P2-X directly as primary cleavage products. Moreover, two other minor cleavage products, 3b/9 and X/9, which are unusual in that they span the P2-P3 cleavage site (Fig. 7), have recently been identified by immune precipitation and microsequence analysis (34). Protein X/9 (molecular weight, 50,000) is similar in size to protein P3-5a (Table 1), but is probably produced in substantially smaller amounts (34). It may correspond to the unidentified precursor-like protein that migrates between bands 4 and 5a (Fig. 1A).

The assignment of protein P2-7a as an uncleaved form of proteins (P2-8 and P2-10) is based upon its electrophoretically determined molecular weight, about 27,400, and is therefore still highly tentative. Attempts to confirm its identity by tryptic analysis were hindered by difficulty in completely freeing P2-7a from the large peak of VP3 with which it comigrates (Fig. 1).

Multiple cleavage pathways for P3 proteins. The complexity of the P3 region of the processing map is due to multiple cleavages, a phenomenon first observed in the homologous P3 proteins of human rhinovirus 1A (18). The same kind of complex cleavages are observed in the P3 proteins of aphthoviruses (7), but appear to be absent in the cardioviruses (5).

Radiochemical sequencing studies have shown that proteins P3-7c and P3-4b are produced by cleaving protein P3-2 at QG-12 (27), whereas proteins P3-6a and P3-6b are produced by an alternative cleavage at YG-8 (11); thus (Fig. 7) P3-2 → P3-6a + P3-6b. The identification of protein P3-7d as a fragment from the left end of protein P3-4 suggests that this protein may arise from a cleavage at YG-8 in P3-4b. If P3-7d is produced by such a cleavage, one must conclude that the cleavages of sites QG-12 and YG-8 are not mutually exclusive; in other words protein P3-4b may not be a stable end product but could be further cleaved at the YG-8 site, and thus 4b → 7d + 6b; or at site QG-12, and thus 6a → 7c + 7d. In view of the multiple potential cleavage sites in P3-2, however, the precise mapping of P3-7d must await terminal amino acid sequence analysis. Similarly it appears that the same YG-8 site can be cleaved in protein 1b to produce the cleavage pathway 1b → 5a + 6b.

Finally P3-1b may be cleaved at QG-11 to yield P3-4a (9, 29), but the corresponding N-terminal fragment of this cleavage has not been detected and may be rapidly degraded.

VPg and potential precursors. Protein P3-5a, like proteins P3-9 and P3-1b, contains VPg sequences (Fig. 7) and is therefore of potential interest as a VPg donor, which is believed to be a requirement for initiating synthesis of new polioviral RNA molecules (38). It is of interest that 5a and 1b were not immunoprecipitated from labeled poliovirus-infected HeLa cell extracts by anti-VPg serum made against a synthetic peptide corresponding to the seven carboxy-terminal amino acids of VPg (27). As previously suggested (26), when VPg sequences are not at a terminal location they may be inaccessible to certain antibody molecules due to the conformation of surrounding amino acid sequences. On the

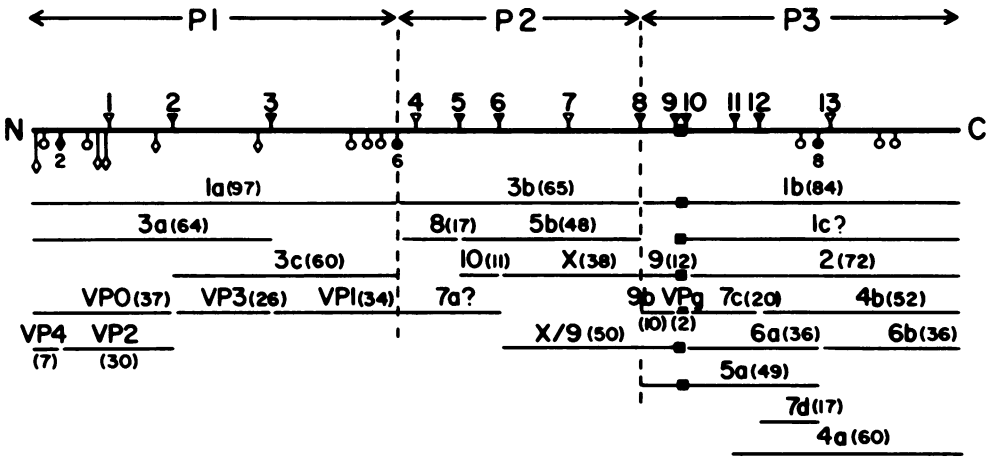

FIG. 7. Processing map of the polioviral polyprotein. The polyprotein (heavy line) is divided into three regions (P1, P2, and P3) for convenience in classifying cleavage products. Amino acid pairs (sites) known to be cleaved are indicated by filled symbols; apparently uncleaved sites are indicated by open symbols: (▽, ▼) glutamine-glycine (QG); (○, ●) tyrosine-glycine (YG); (◇, ◆) asparagine-serine (NS). The glutamine-glycine sites are all believed to be cleaved by P3-7c, a virus-encoded coded protease. The agents responsible for cleavage of sites NS-2, YG-6, and YG-8 have not yet been identified. Proteins P3-4a (9, 29), X/9 (34), and 3b/9 (data not shown; 34) are produced in only trace amounts. This implies that site QG-11 is cleaved only rarely and that site QG-8 is rarely left uncleaved. Assignments for proteins 7a and 1c are tentative (see the text). (■) VPg.

other hand, protein 1b and a protein that may correspond to 5a (pre-VPg 2) were immunoprecipitated by antisera directed against larger synthetic peptides (corresponding to all 22 or the 14 amino-terminal amino acids of VPg) (4). The increased efficacy of immunoprecipitation in the latter case is evidently due to the different and increased number of epitopes present in the longer peptides.

Another molecule of interest as a potential VPg donor is protein 1c, which has a proteolytic map similar to that of protein P3-2 (37). Thus P3-1c appears to be generated by cleavage of QG-9 just to the left of QG-10 and would contain the VPg sequences at its amino terminus. It is worth noting in this regard that the relatively minor amount of these VPg-containing molecules is no argument against their potential importance as VPg donors since the number of RNA molecules required for synthesis of the 60-subunit virions is only a minor fraction of the total number of translations.

Finally identification of protein P3-9b as the cleavage complement of VPg suggests that it is generated by the reaction, 9 → 9b + VPg. Proteins P3-9b and P2-10 comigrate in the Laemmli gel electrophoresis system, and this mixture was earlier designated P3-10 (26, 29, 34). The nomenclature used in Fig. 7 will be followed hereafter. The amount of protein P3-9b seen after a 90-min chase is substantial (Fig. 1B), yet attempts to locate VPg in infected cells have been unsuccessful (4, 14, 26) even with radioactive amino acids other than methionine, which is absent from VPg. The reason for this is still unclear.

Other minor proteins. The map illustrated in Fig. 7 represents over 95% of the total methionine incorporated into poliovirus-infected cells and therefore encompasses the major cleavage reactions. Nevertheless the map does not exhaust the complete repertoire of proteins produced in poliovirus-infected cells (1). Inspection of the electrophoretic profiles in Fig. 1 reveals several unidentified minor peaks in the region between 1a and 2 and between 3a and 5a. Additional unidentified spots have been detected with two-dimensional electrophoresis, although none is present in large amounts (36). This suggests that some of the cleavage sites mapped as inactive are in fact cleaved on rare occasions. The possible functional significance of such minor cleavages is currently unclear.

ACKNOWLEDGMENTS

This work was supported by grant MV-331 from the American Cancer Society, by Public Health Service Grants AI 15122 and CA 28146 from the National Institutes of Health, and by the U.S. Department of Energy. M.A.P. was a National Science Foundation predoctoral Fellow, O.M.K. was the recipient of a Public Health Service Postdoctoral Fellowship from the National Institute of Allergy and Infectious Diseases, and B.L.S. was supported by Public Health Service Postdoctoral Fellowship AI 05935 from the National Institutes of Health.

LITERATURE CITED

1. **Abraham, G., and P. D. Cooper.** 1975. Poliovirus polypeptides examined in more detail. J. Gen. Virol. **29**:199–213.
2. **Adler, C. J., M. Elzinga, and E. Wimmer.** 1983. The genomic-linked protein of picornavirus. VIII. Complete amino acid sequence of poliovirus VPg and carboxyterminal analysis of its precursor, P3-9. J. Gen. Virol. **64**:349–355.
3. **Anderson, C. W.** 1982. Partial sequence determination of metabolically labeled radioactive proteins and peptides, p. 147–167. *In* J. Setlow and A. Hollaender (ed.), Genetic engineering, vol. 4. Plenum Publishing Corp., New York.
4. **Baron, M. H., and D. Baltimore.** 1982. Antibodies against the chemically synthesized genome-linked protein of poliovirus react with native virus-specific proteins. Cell **28**:395–404.
5. **Butterworth, B. E.** 1973. A comparison of the virus-specific polypeptides of encephalomyocarditis virus, human rhinovirus-1A, and poliovirus. Virology **56**:439–453.
6. **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.
7. **Doel, T. R., D. V. Sangar, D. J. Rowlands, and F. Brown.** 1978. A re-appraisal of the biochemical map of foot-and-mouth disease virus RNA. J. Gen. Virol. **41**:395–404.
8. **Emini, E. A., M. Elzinga, and E. Wimmer.** 1982. Carboxy-terminal analysis of poliovirus proteins: the termination of poliovirus RNA translation and the location of unique poliovirus polyprotein cleavage sites. J. Virol. **42**:194–199.
9. **Etchison, D., and E. Ehrenfeld.** 1980. Viral polypeptides associated with the RNA replication complex in poliovirus-infected cells. Virology **107**:135–143.

10. **Gilson, W., R. Gilson, and R. R. Rueckert.** 1972. An automatic high precision acrylamide gel fractionator. Anal. Biochem. **47**:321–328.
11. **Hanecak, R., B. L. Semler, C. W. Anderson, and E. Wimmer.** 1982. Proteolytic processing of poliovirus polypeptides: antibodies to polypeptide P3-7c inhibit cleavage at glutamine-glycine pairs. Proc. Natl. Acad. Sci. U.S.A. **79**:3973–3977.
12. **Jacobson, M. F., J. Asso, and D. Baltimore.** 1970. Further evidence on the formation of poliovirus proteins. J. Mol. Biol. **49**:657–669.
13. **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 A poliovirus. J. Virol. **33**:256–263.
14. **Kitamura, N., B. L. Semler, P. G. Rothberg, G. R. Larsen, C. J. Adler, A. J. Dorner, E. A. Emini, R. Hanecak, J. J. Lee, S. van der Werf, C. W. Anderson, and E. Wimmer.** 1981. Primary structure, gene organization and polypeptide expression of poliovirus RNA. Nature (London) **291**:547–553.
15. **Korant, B. D.** 1972. Cleavage of viral precursor proteins *in vivo* and *in vitro*. J. Virol. **10**:751–759.
16. **Larsen, G. R., C. W. Anderson, A. J. Dorner, B. L. Semler, and E. Wimmer.** 1982. Cleavage sites within the poliovirus capsid protein precursors. J. Virol. **41**:340–344.
17. **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.
18. **McLean, C., T. J. Matthews, and R. R. Rueckert.** 1976. Evidence of ambiguous processing and selective degradation in the noncapsid proteins of rhinovirus 1A. J. Virol. **19**:903–914.
19. **Medappa, K. C., C. McLean, and R. R. Rueckert.** 1971. On the structure of rhinovirus 1A. Virology **44**:259–270.
20. **Nomoto, A., T. Omata, H. Toyoda, S. Kuge, H. Horie, Y. Kataoke, Y. Genba, Y. Nakano, and N. Imura.** 1982. Complete nucleotide sequence of the attenuated poliovirus Sabin 1 strain genome. Proc. Natl. Acad. Sci. U.S.A. **79**:5793–5797.
21. **Pallansch, M. A., O. M. Kew, A. C. Palmenberg, F. Golini, E. Wimmer, and R. R. Rueckert.** 1980. Picornaviral VPg sequences are contained in the replicase precursor. J. Virol. **35**:414–419.
22. **Racaniello, V. R., and D. Baltimore.** 1981. Molecular cloning of poliovirus cDNA and determination of the complete nucleotide sequence of the viral genome. Proc. Natl. Acad. Sci. U.S.A. **78**:4887–4891.
23. **Rekosh, D.** 1972. Gene order of the poliovirus capsid proteins. J. Virol. **9**:479–487.
24. **Rothberg, P. G., T. J. R. Harris, A. Nomoto, and E. Wimmer.** 1978. O^4-(5′-Uridylyl)tyrosine is the bond between the genome-linked protein and the RNA of poliovirus. Proc. Natl. Acad. Sci. U.S.A. **75**:4868–4872.
25. **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 Publishing Corp., New York.
26. **Semler, B. L., C. W. Anderson, R. Hanecak, L. F. Dorner, and E. Wimmer.** 1982. A membrane-associated precursor to poliovirus VPg identified by immunoprecipitation with antibodies directed against a synthetic heptapeptide. Cell **28**:405–412.
27. **Semler, B. L., C. W. Anderson, N. Kitamura, P. G. Rothberg, W. L. Wishart, and E. Wimmer.** 1981. Poliovirus replication proteins: RNA sequence encoding P3-1b and the sites of proteolytic processing. Proc. Natl. Acad. Sci. U.S.A. **78**:3464–3468.
28. **Semler, B. L., R. Hanecak, C. W. Anderson, and E. Wimmer.** 1981. Cleavage sites in the polypeptide precursors of poliovirus protein P2-X. Virology **114**:589–594.
29. **Semler, B. L., R. Hanecak, L. F. Dorner, C. W. Anderson, and E. Wimmer.** 1983. Poliovirus RNA synthesis *in vitro*: structural elements and antibody inhibition. Virology **126**:624–635.
30. **Shealy, D. J., and R. R. Rueckert.** 1978. Proteins of Rous-associated virus 61, an avian retrovirus: common precursor for glycoproteins gp85 and gp35 and use of pactamycin to map translational order of proteins in the *gag*, *pol*, and *env* genes. J. Virol. **26**:380–388.
31. **Shih, D. S., C. T. Shih, O. Kew, M. Pallansch, 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.
32. **Summers, D. F., and J. V. Maizel, Jr.** 1971. Determination of the gene sequence of poliovirus with pactamycin. Proc. Natl. Acad. Sci. U.S.A. **68**:2852–2856.
33. **Taber, R., D. Rekosh, and D. Baltimore.** 1971. Effect of pactamycin on synthesis of poliovirus proteins: a method for genetic mapping. J. Virol. **8**:395–401.
34. **Takegami, T., B. L. Semler, C. W. Anderson, and E. Wimmer.** 1983. Membrane fractions active in poliovirus RNA replication contain VPg precursor polypeptides. Virology **128**:33–47.
35. **Vrijsen, R., P. Brioen, and A. Boeye.** 1980. Identification of poliovirus precursor proteins by immunoprecipitation. Virology **107**:567–569.
36. **Wiegers, K. J., and R. Dernick.** 1981. Poliovirus-specific polypeptides in infected HeLa cells analyzed by isoelectric focusing and 2D-analysis. J. Gen. Virol. **52**:61–69.
37. **Wiegers, K. J., and R. Dernick.** 1981. Peptide maps of labeled poliovirus proteins after two-dimensional analysis by limited proteolysis and electrophoresis in sodium dodecyl sulfate. Electrophoresis **2**:98–103.
38. **Wimmer, E.** 1982. Genome-linked proteins of viruses. Cell **28**:199–201.