JOURNAL OF VIROLOGY, Apr. 1982, p. 208–219
0022-538X/82/040208-12$02.00/0

Vol. 42, No. 1

Structural Analysis of Virion Proteins of the Avian Coronavirus Infectious Bronchitis Virus

DAVID F. STERN,[1,2]* LOYD BURGESS,[1] AND BARTHOLOMEW M. SEFTON[2]

Department of Biology, University of California, San Diego, La Jolla, California 92093,[1] and Tumor Virology Laboratory, The Salk Institute, San Diego, California 92138[2]

Received 13 August 1981/Accepted 25 November 1981

We have found six major polypeptides in virions of the avian coronavirus infectious bronchitis virus grown in tissue culture: four glycoproteins, GP84, GP36, GP31, and GP28, and two non-glycosylated proteins, P51 and P23. In addition, we detected three minor species: two glycoproteins, GP90 and GP59, and one non-glycosylated protein, P14. Two-dimensional tryptic peptide mapping showed that GP36, GP31, GP28, and P23 comprise a group of closely related proteins which we have designated the "P23 family," but that the other proteins are distinct. Analysis by partial proteolytic digestion of the P23 family, labeled biosynthetically with [^{35}S]methionine, and P23, labeled with [^{35}S]formyl-methionine by in vitro translation of RNA from infected cells, revealed that the proteins of the P23 family differ in their amino-terminal domains. Similar analysis of GP31 and GP36 labeled with [^{3}H]mannose showed that the partial proteolytic fragments unique to these proteins were glycosylated. This suggests that differences in glycosylation in the amino-terminal domains contributes to the marked polymorphism of the P23 family. The results are discussed with respect to possible models for synthesis of the virion proteins.

The coronaviruses are large enveloped viruses with positive-stranded RNA genomes. The genome of the coronaviruses is the largest of any RNA virus characterized, with estimates of its size ranging from 6×10^6 to 9×10^6 daltons (20). These viruses are of interest because of the diverse diseases caused by members of this family and the unique mode of gene expression.

The positive-stranded animal viruses characterized to date employ two different strategies for the expression of their genetic information. The picornavirus proteins are synthesized by translation of a single mRNA identical in sequence to the viral genome (11). In contrast, the alphaviruses employ two mRNAs, the genomic RNA which is translated to produce the nonstructural proteins, and a single subgenomic mRNA, corresponding to a 3'-terminal portion of the genome, which directs the synthesis of virion structural proteins. The translated regions of these mRNAs do not overlap; translation of the genome terminates at a point upstream from that corresponding to the 5' end of the subgenomic mRNA, and the smaller mRNA is translated from an initiation site latent in the genome (3, 27). The production of two mRNAs permits independent transcriptional regulation of the synthesis of the nonstructural and structural polyproteins.

Coronavirus infection results in the synthesis of at least five subgenomic RNA species (13, 28, 30). These RNAs are likely to be functional mRNAs because they are polyadenylated, are present in polysomes, and can be translated in vitro to yield viral proteins (21, 25; see below). Ribonuclease T_1 oligonucleotide fingerprint analysis of these RNAs demonstrates that they form a nested set (13, 30). Comparison of the oligonucleotide composition of the subgenomic mRNAs of the avian coronavirus infectious bronchitis virus (IBV) with a T_1 oligonucleotide order obtained for the IBV genome revealed that each of the five major RNAs is colinear with the 3' end of the genome (31). In this sense, the transcriptional pattern of the coronaviruses resembles that of the alphaviruses.

We wish to understand eventually how the genetic information in these overlapping mRNAs is expressed. As a first step we have characterized the IBV virion proteins because they are likely to be the polypeptides encoded by these mRNAs. There has been some disagreement as to the polypeptide content of coronavirus particles. Virions of the murine coronaviruses contain a large glycoprotein with an estimated molecular weight of approximately 90,000, a smaller glycoprotein or group of glycoproteins with molecular weights of 20,000 to 30,000, and a phosphoprotein with a molecular weight of 50,000 to 60,000 (24, 32, 33, 35). Analysis of the structure of viral particles has shown that most of the large glycoprotein is

external to the lipid bilayer and forms the characteristic coronavirus peplomers, that the smaller glycoprotein is a membrane protein which is largely but not completely protected from protease digestion in the intact virion, and that the phosphoprotein and associated genomic RNA comprise the helical nucleocapsid of the virus (32–35).

Whereas a rather similar polypeptide composition of IBV has been reported (6, 9, 16), others have identified as many as sixteen proteins in IBV virions (2, 8). The variation among these studies may be attributed to the different strains of virus employed, the different cellular and embryonic hosts used to grow the virus, the lack of uniformity among procedures for virus purification and electrophoretic analysis, and the difficulty in distinguishing minor viral proteins from contaminating host-encoded proteins.

We demonstrate here that virions of the Beaudette strain of IBV contain six major proteins, GP84, P51, GP36, GP31, GP28, and P23, and three minor proteins, GP90, GP59, and P14, when grown in tissue culture. GP36, GP31, GP28, and P23 appear to be closely related to one another. Further analysis suggests that these proteins differ in their amino-terminal domains and that it is the amino-terminal portions which are glycosylated in GP36 and GP31. The remaining proteins which were mapped appeared unrelated to one another. We discuss these results with respect to those obtained with other coronaviruses and possible models for translation of the coronavirus mRNAs.

MATERIALS AND METHODS

Virus and cells. The Beaudette strain (strain 42) of IBV was propagated in primary chicken embryo kidney (CEK) cells as described previously (30), except that the cells were seeded in Dulbecco modified Eagle medium (DMEM) containing 10% horse serum. Cells were incubated at 37 or 38.5°C.

Virus growth and purification. CEK cells in 160-mm culture dishes (Falcon Plastics) were washed once with Tris-buffered saline and infected with IBV at a multiplicity of approximately 10 PFU/cell. After incubation for 90 min, the inoculum was replaced with 12 ml of DMEM containing 2% calf serum (for nonlabeled virions) or with 11 ml of one of the labeling media described below. At 4 h postinfection (defined relative to the end of the adsorption period), radioactive label was added in 1.0 ml of labeling medium, giving a final volume of 12 ml. Cultures were labeled with [^{35}S]methionine (1,200 Ci/mmol; Amersham Corp.) at a concentration of 0.08 mCi/ml in methionine-free DMEM. Cells were labeled with [^{35}S]cysteine (1,150 Ci/mmol; Amersham) at a concentration of 0.08 mCi/ml in cysteine-free DMEM. Cells were labeled with ^{3}H-mixed amino acids (Amersham) at a concentration of 0.08 mCi/ml in amino acid-free DMEM supplemented with glutamine. Cultures were labeled with [2-^{3}H]mannose (14.5 Ci/mmol; New England Nuclear Corp.) at a concentration of 0.13 mCi/ml or [6-^{3}H]glucosamine (New England Nuclear) at a concentration of 0.13 mCi/ml in minimal essential medium supplemented with 5 mM pyruvate and with nonessential amino acids. All labeling media contained 2% calf serum dialyzed against phosphate-buffered saline. At 15 h postinfection the medium from the infected cultures was harvested and clarified. Clarification and all subsequent steps were performed at 4°C. Virions were sedimented through 6.5 ml of 20% (wt/vol) sucrose in TNE (50 mM Tris [pH 7.4], 100 mM NaCl, 1 mM EDTA) onto a cushion of 5.0 ml of 55% (wt/wt) sucrose-TNE by centrifugation in a SW27 rotor at 75,000 × *g* for 3 h. The virus was located visually, aspirated, diluted with TNE, and centrifuged to equilibrium in 16-ml linear 20%–55% sucrose-TNE gradients at 75,000 × *g* for 18 h in an SW27 rotor. The virus was then diluted and layered on linear 10 to 50% (vol/vol) Renografin (Renografin-76; E. R. Squibb and Sons, Inc.)–TNE gradients which were centrifuged to equilibrium under the same conditions. Finally the virus was diluted with TNE, and virions were pelleted by sedimentation at 65,000 × *g* for 3 h in a 30 rotor. Pellets were stored at −70°C until used.

SDS-polyacrylamide gel electrophoresis. Virions were suspended in sample buffer (5 mM phosphate buffer [pH 7.0], 2% sodium dodecyl sulfate [SDS], 10% mercaptoethanol, 100 mM dithiothreitol, 10% glycerol, bromophenol blue), boiled for 30 s, and analyzed by electrophoresis on discontinuous 15% acrylamide–0.09% bisacrylamide gels as described before (23). Analytical gels were 14 cm long by 1 mm thick. For peptide mapping GP90 was isolated from a gel 40 cm long by 1.5 mm thick. All other proteins were isolated from gels 14 cm long by 2 cm thick.

Molecular weight determination. The molecular weights of virion proteins were estimated by comparison of their electrophoretic mobilities to those of the following standards (and molecular weights): horse heart cytochrome *c* (12,400); soybean trypsin inhibitor (21,500); chicken ovalbumin (45,000); bovine serum albumin (68,000); human transferrin (82,000); rabbit muscle phosphorylase A (90,000); and *Escherichia coli* RNA polymerase (155,000 and 165,000).

Molar ratios and relative glycosylation of virion proteins. The fluorograph of virions labeled with ^{3}H-mixed amino acids (Fig. 1, lane b) was scanned with a Zeineh soft laser scanning densitometer (LKB Instruments, Inc.), and peak areas were computed with a Hewlett-Packard digitizer. The relative molarities of the proteins were calculated and normalized to a value of 1.0 for the abundance of GP31 as follows

$$\text{Relative abundance}_x = \frac{\text{area of peak}_x}{\text{molecular weight}_x} \div \frac{\text{area of peak}_{GP31}}{\text{molecular weight}_{GP31}}$$

The fluorographs of virion proteins labeled with [^{3}H]mannose and [^{3}H]glucosamine (Fig. 1, lanes a and c) were scanned and quantified as above. The mannose and glucosamine contents of the proteins were calculated per polypeptide chain by correcting for the relative abundance of each of the proteins and the values were normalized to a figure of 1.0 for the mannose and glucosamine contents of GP31 as follows

$$\text{Relative content}_x = \frac{\text{area of peak}_x}{\text{relative abundance}_x} \div \frac{\text{area of peak}_{GP31}}{\text{relative abundance}_{GP31}}$$

Tryptic peptide mapping. Virions were labeled biosynthetically with [^{35}S]methionine or [^{35}S]cysteine and purified as described above. For maps of methionine-labeled peptides, virions were purified and pooled from three to five radiolabeled, infected cultures. Virions purified from two labeled and six unlabeled cultures were pooled for maps of [^{35}S]cysteine-labeled proteins. The virion proteins were resolved on 2-mm-thick 15% acrylamide–0.09% bisacrylamide gels as described above and eluted by electrophoresis as previously described (36) up to the first lyophilization step. Bovine immunoglobulin G carrier (10 μg) was then added, and the proteins were precipitated with trichloroacetic acid, oxidized with performic acid, and digested with tolylsulfonyl phenylalanyl chloromethyl ketone-trypsin (Worthington Diagnostics) as described before (1). GP84 was purified by a different procedure. Virion proteins were fractionated on a preparative gel 14 cm long by 2.0 mm thick, and GP84 was eluted by homogenization of the excised gel slice (1). The protein was then dissolved in electrophoresis sample buffer and applied to a 14-cm by 1.0-mm gel. Electrophoresis of this gel was allowed to continue for 4 h after the bromophenol blue dye had reached the bottom reservoir. GP84 was eluted from the gel by homogenization and prepared for peptide mapping as described above. Two-dimensional peptide mapping on 0.1-mm thin-layer cellulose plates (E. M. Reagents) by electrophoresis at pH 4.7 in the first dimension, and ascending chromatography in the second dimension was performed as described previously (1). The plates were prepared for fluorography (5) and exposed to preflashed film (Kodax X-Omat R) at −70°C.

Partial proteolytic analysis of virion proteins. Virion proteins were reduced and carboxymethylated as follows. The purified virions were resuspended in 1.0 M Tris (pH 8.8)–2% 2-mercaptoethanol–2% SDS and boiled for 1 min. Iodoacetamide (1 M, BDH Biochemicals) was added to yield a final concentration of 0.33 M, and the suspension was incubated at 20°C for 30 min. The proteins were precipitated with trichloroacetic acid, and the pellets were washed twice with ethanol and once with ethanol-ether (1:1, vol/vol) and suspended in sample buffer. The proteins were resolved in a 1-mm 15% preparative acrylamide gel which was dried without fixation. Polypeptides were detected by autoradiography (^{3}H-labeled bands were located according to the positions of ^{35}S-labeled proteins in adjacent lanes) and mapped as described by Cleveland et al. (7), except that EDTA was omitted from the gel solutions, and the buffers for enzyme dilution and rehydration of the gel slices contained 1 mM 2-mercaptoethanol. The gel slices were inserted into wells of a 20% acrylamide–0.08% bisacrylamide discontinuous SDS-polyacrylamide gel prepared as described previously (23). Proteolytic digestion was for 30 min at room temperature.

Preparation of cytoplasmic RNA for in vitro translation. Six CEK cultures in 160-mm dishes were washed with Tris-buffered saline and infected with IBV at a multiplicity of approximately 10 PFU/cell. After the 90-min adsorption period, the inoculum of one plate was replaced with 20 ml of phosphate-free DMEM containing 1 μg of actinomycin D (Calbiochem-Boehring Corp.) per ml and 2% heat-inactivated calf serum dialyzed against 0.9% NaCl. After 1 h ^{32}P (ICN Pharmaceuticals; 285 Ci/mg) was added to this culture to a final concentration of 0.05 mCi/ml. The remaining cultures were incubated in DMEM containing 2% heat-inactivated calf serum. At 8 h postinfection the cells were washed three times in cold phosphate-buffered saline and once with cold TNE, scraped into cold TNE, and pelleted at 4°C (labeled and unlabeled cultures were worked up separately and in parallel). After resuspension of the pellets in TNE, three volumes of TNE containing 2% Triton N-101 (Sigma Chemical Co.) was added, and the cells were dissolved by repeated inversion for 1 min at room temperature and centrifuged at 4°C to remove remaining cells and nuclei. The supernatants were transferred to tubes containing SDS, the pellets were reextracted with 2% Triton N-101-TNE, and the resulting supernatants were pooled with the supernatants obtained from the first extraction. The final concentration of SDS was 2%. The supernatants were extracted with an equal volume of phenol-chloroform (1:1, vol/vol; saturated with TNE), the organic phases were reextracted with TNE, and the combined aqueous phases were then extracted once more with phenol-chloroform and precipitated with ethanol.

Nucleic acid recovered from the labeled and unlabeled cultures was pooled, and polyadenylated RNA was selected by two cycles of chromatography on oligodeoxythymidylic acid-cellulose (Collaborative Research, Inc., type 3) as previously described (7), except that the RNA was not boiled before chromatography. The RNA was then precipitated twice with ethanol, suspended, lyophilized, and dissolved in water to a final concentration of 1 mg/ml.

A portion of the RNA was fractionated by preparative polyacrylamide gel electrophoresis in a 2% acrylamide–0.1% bisacrylamide gel as described previously (30), except that the chamber buffer contained 0.2% SDS. The IBV intracellular RNAs were located by autoradiography of the wet gel and excised. RNA was eluted from the gel slices by homogenization and prepared for in vitro translation as described previously (10).

In vitro translation. Samples (0.5 μl) of RNA were translated in a messenger-dependent reticulocyte lysate (19) as described before (1). [^{35}S]formyl methionyl $tRNA_f^{met}$ (4.5×10^6 cpm/μl) was prepared essentially as described by Stanley (29), except that it was treated with $CuSO_4$ to hydrolyze methionyl $tRNA_f^{met}$ (22). It was a gift from B. Adkins and J. Cooper (The Salk Institute). In vitro translation was performed as above, except that 1 μl of [^{35}S]formyl methionyl $tRNA_f^{met}$ was added instead of [^{35}S]methionine, and the reaction mixture was supplemented with 0.4 mM methionine and 0.04 mM dithiothreitol.

RESULTS

Identification of IBV structural proteins. All experiments reported here employed IBV grown in CEK cells in tissue culture. Virions were purified by centrifugation onto a sucrose cush-

ion, sedimentation to equilibrium in sucrose, and sedimentation to equilibrium in Renografin. Virions pelleted from the Renografin band were boiled in the presence of SDS, dithiothreitol, and 2-mercaptoethanol, and the proteins were analyzed by SDS-polyacrylamide gel electrophoresis on 15% low-bisacrylamide gels. Six major and three minor polypeptides were found in virions labeled with ^{3}H-mixed amino acids (Fig. 1, lane b). The molecular weights of these proteins were estimated by comparison of their electrophoretic mobilities with those of several marker proteins, and the bands in Fig. 1 are labeled accordingly. P51, P23, and P14 were apparently not glycosylated because they were not labeled with either [^{3}H]mannose or [^{3}H]glucosamine (Fig. 1, lanes a and c, respectively). The remaining proteins, GP90, GP84, GP59, GP36, GP31, and GP28, were labeled with one or both ^{3}H-sugars and have therefore been designated GP.

The molar ratios of virion proteins were determined from a densitometric scan of the fluorograph in Fig. 1, lane b, which shows virion proteins labeled with an amino acid mixture (Table 1). There is some variability in the proportion of the less prominent proteins among different preparations, so these numbers are necessarily approximate. The major structural proteins were GP31 and P51. Relatively small amounts of GP90, GP59, and P14 were present. The extent of labeling by [^{3}H]mannose and [^{3}H]glucosamine was also calculated (Table 1). The mannose content of the five glycoproteins was approximately equal. In contrast, there was significant variation in glucosamine content. Particularly striking was the lack of labeling of GP28 by [^{3}H]glucosamine (Fig. 1, lane c). Again, because of the variability in proportions of these proteins, these numbers must be considered approximate.

TABLE 1. Molar ratios and relative glycosylation of IBV virion proteins

Protein	Relative amt of polypeptide[a]	Relative content of mannose and glucosamine[b]	
		Mannose	Glucosamine
GP90	0.001	2.2	1.5
GP84	0.01	0.8	0.5
GP59	0.006	ND[c]	ND
P51	0.5		
GP36	0.07	1.0	5.0
GP31	1.0	1.0	1.0
GP28	0.04	1.0	0
P23	0.06		
P14	0.01		

[a] The fluorograph of virion proteins labeled with ^{3}H-mixed amino acids (Fig. 1, lane b) was scanned, and the relative molarity of each protein was calculated from the peak areas as described in the text. The data have been normalized such that the relative abundance of GP31 is 1.0.

[b] The fluorograph of virion proteins labeled with [^{3}H]mannose and [^{3}H]glucosamine (Fig. 1, lanes a and c) was scanned, and peak areas were computed. The mannose and glucosamine contents per polypeptide chain were calculated and normalized as described above, except that the values were corrected for the relative abundance of each polypeptide.

[c] ND, Not determined.

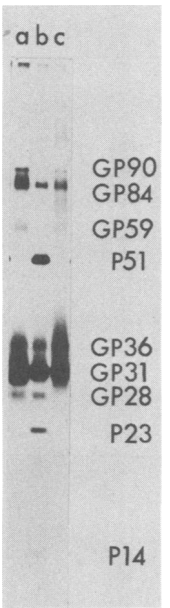

FIG. 1. Virion proteins of IBV grown in CEK cells. Virions labeled biosynthetically with [^{3}H]mannose, ^{3}H-mixed amino acids, or [^{3}H]glucosamine were purified and analyzed by SDS-polyacrylamide gel electrophoresis as described in the text. Approximately 50,000 cpm of radioactive virus was analyzed in each lane, and proteins were visualized by fluorography with preflashed film (Kodak X-Omat R) for 8 days at −70°C. Lanes: a, [^{3}H]mannose-labeled IBV virions; b, ^{3}H-mixed amino acid-labeled virions; lane c, [^{3}H]glucosamine-labeled virions.

Tryptic peptide analysis of IBV proteins. To determine whether the proteins represented unique gene products or were related to one another, we prepared tryptic peptide maps of the structural proteins labeled with methionine and cysteine (Fig. 2, 3, and 4). It was difficult to obtain preparations of GP90 and GP84 free of cross-contamination when the proteins were isolated from 14-cm preparative gels. To circumvent this problem we either used two consecutive gel purification steps, with the second gel run for a longer time than usual (GP84, Fig. 2), or fractionated the virion proteins on a single 40-cm gel (GP90, Fig. 2).

It was evident from maps of methionine-containing peptides that GP90 and GP84 were struc-

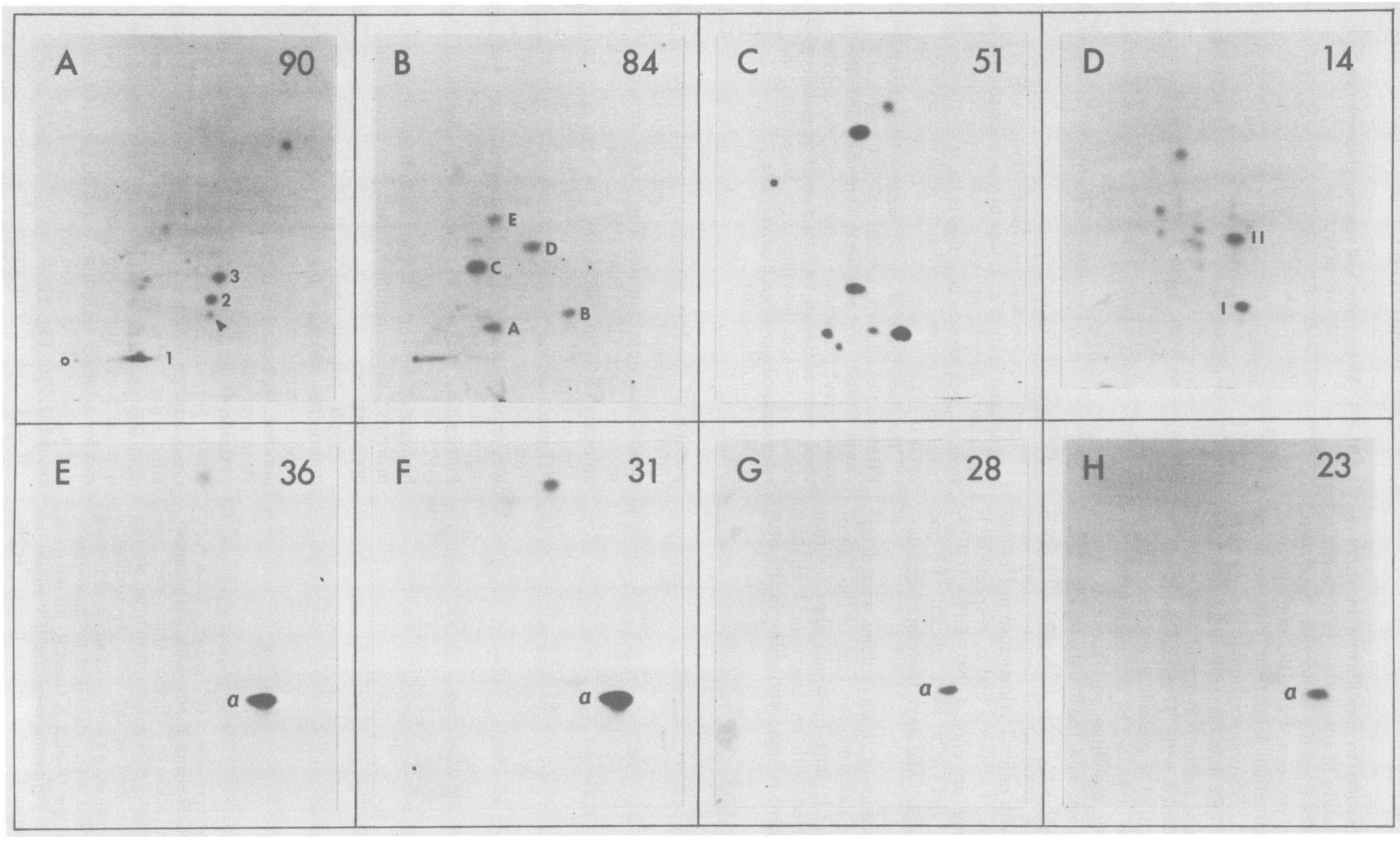

FIG. 2. Two-dimensional tryptic peptide maps of IBV virion proteins labeled with [^{35}S]methionine. Tryptic digests of [^{35}S]methionine-labeled IBV virion proteins were prepared as described in the text. In this figure and subsequent figures peptides were resolved on thin-layer plates by electrophoresis in the first dimension (from left to right) and ascending chromatography in the second dimension (bottom to top). The plates were dipped in 2-methyl naphthalene containing 0.4% diphenyloxazole and exposed at −70°C. Sample activity ranged from 200 cpm (GP90) to 50,000 cpm (P51), and exposure ranged from 10 days to 49 days with preflashed film. "O" indicates the origin in panel A. The other symbols designate specific peptides and are discussed in the text. Panels: A, GP90; B, GP84; C, P51; D, P14; E, GP36; F, GP31; G, GP28; and H, P23.

turally distinct (Fig. 2A and B). GP90 contained two major tryptic peptides (designated 2 and 3) and a streak (designated 1); GP84 yielded five major tryptic peptides, designated A, B, C, D, and E. Mixture of peptides from the two proteins confirmed that none of the peptides derived from them comigrated (Fig. 3A). Preliminary data suggest that the minor glycoprotein GP59 may be related to GP84.

The map of P51, labeled with [^{35}S]methionine (Fig. 2C), contained eight peptides, none of which was present in the maps of the other proteins. Thus P51 appeared unrelated to the other virion proteins.

The maps of methionine-labeled GP36, GP31, GP28, and P23 (Fig. 2E, F, G, and H) each contained one major peptide, designated α. Analysis of mixtures showed that these peptides had the same mobility (Fig. 3 E and F). The ovoid shape of peptide α in these maps was due to overloading of the thin-layer plate. When smaller amounts of material were analyzed it was possible to resolve this peptide as a small round spot (Fig. 3D).

Because of the simplicity of the maps of methionine-containing peptides obtained from these proteins, we prepared maps of the cysteine-containing tryptic peptides as well. These maps contained four major peptides each (Fig. 4A, B, C, and D) and were almost identical. A mixture of peptides from GP31 and P23 (Fig. 4E) showed that all of these peptides comigrated. We concluded that GP36, GP31, GP28, and P23 constitute a family of closely related proteins, which we will refer to as the "P23 family."

The map of methionine-labeled P14 (Fig. 2D) contained two major tryptic peptides (I and II). Peptide I had a mobility similar to that of peptide α, but a mixture of peptides from P14 and GP31 (Fig. 3C) showed that peptides I and α do not have identical mobilities. P14 therefore appears to be unrelated to other virion proteins.

One minor peptide in GP90 (see arrow in Fig. 2A) had the same mobility as peptide α of the P23 family (compare Fig. 2A, Fig. 3D, and the mixture in Fig. 3C). This peptide was also detected in some preparations of GP84. It was not, however, visible in the preparation of GP84 mapped in Fig. 2. The amount of this peptide was highly variable in preparations of GP90 and GP84, which suggested that the peptide was not derived from either of the two proteins. A more likely explanation is that some preparations of GP90 and GP84 are contaminated with aggregates of one of the P23 family proteins (see below).

Sequence relationships within the P23 family. We wished to localize the structural differences between proteins of the P23 family. Partial proteolytic digestion products of the proteins were therefore analyzed by the method of Cleveland et al. (7). This analysis was initially complicated by the tendency of proteins of the P23 family to aggregate, a phenomenon first reported by Sturman (32). The mapping procedure involves gel purification of a protein followed by proteolytic digestion and electrophoresis in a second gel. The paucity of reducing agent in the digestion buffer apparently promotes the formation of multimers. To minimize aggregation it was necessary to reduce and alkylate the proteins prior to gel fractionation. This treatment lessened, but did not completely eliminate, aggregation; discrete multimers of gel-purified GP31 are visible in Fig. 5 and Fig. 6, section I, lane 31. Carboxymethylation did not significantly alter the gel profile obtained for the structural proteins (data not shown).

Figure 5 shows a partial proteolytic map of GP31. The undigested protein is in lane a, and lanes b, c, and d show the results of digestion with increasing amounts of the protease from *Staphylococcus aureus*, strain V8. With 10 ng of protease (Fig. 5, lane b) one major fragment was observed. This fragment was slightly smaller than the undigested protein and has been designated fragment V12. With higher amounts of protease (lanes c and d) smaller fragments, designated V11, V3, V2, and V1, appeared. Digestion with 10 and 50 ng of papain produced five fragments, which we have designated P12, P11, P3, P2, and P1 (Fig. 5, lanes e and f). The presence of the V12 fragment in lane a was due to the leakage of the protease from the adjacent lane. This fragment was not visible in undigested GP31 in similar mapping experiments where the protein was well separated from lanes containing protease (e.g., Fig. 6, section I, lane 31).

In Figure 6 the proteolysis products of GP31 are compared to those derived from GP36, GP28, and P23. The figure is divided into three sections, each containing all four proteins of the P23 family. Section I shows the untreated proteins, section II shows the results of proteolysis with 50 ng of *S. aureus* protease, and section III shows the results of proteolysis with 10 ng of papain. The first four lanes in each section show the profiles obtained for the GP36, GP31, GP28, and P23 from virions. Bands smaller than the undigested GP31 and P23 may have resulted from spontaneous proteolysis. P23 did not migrate as a doublet in other experiments. The smaller *S. aureus* protease and papain fragments (V1, V2, V3, P1, P2, and P3) were of equal size in all four proteins. This demonstrated the homology of the four proteins which was apparent from the peptide maps. The large *S. aureus* protease fragments V11 and V12 and the large papain fragments P11 and P12 varied in size among the proteins in much the same way as the

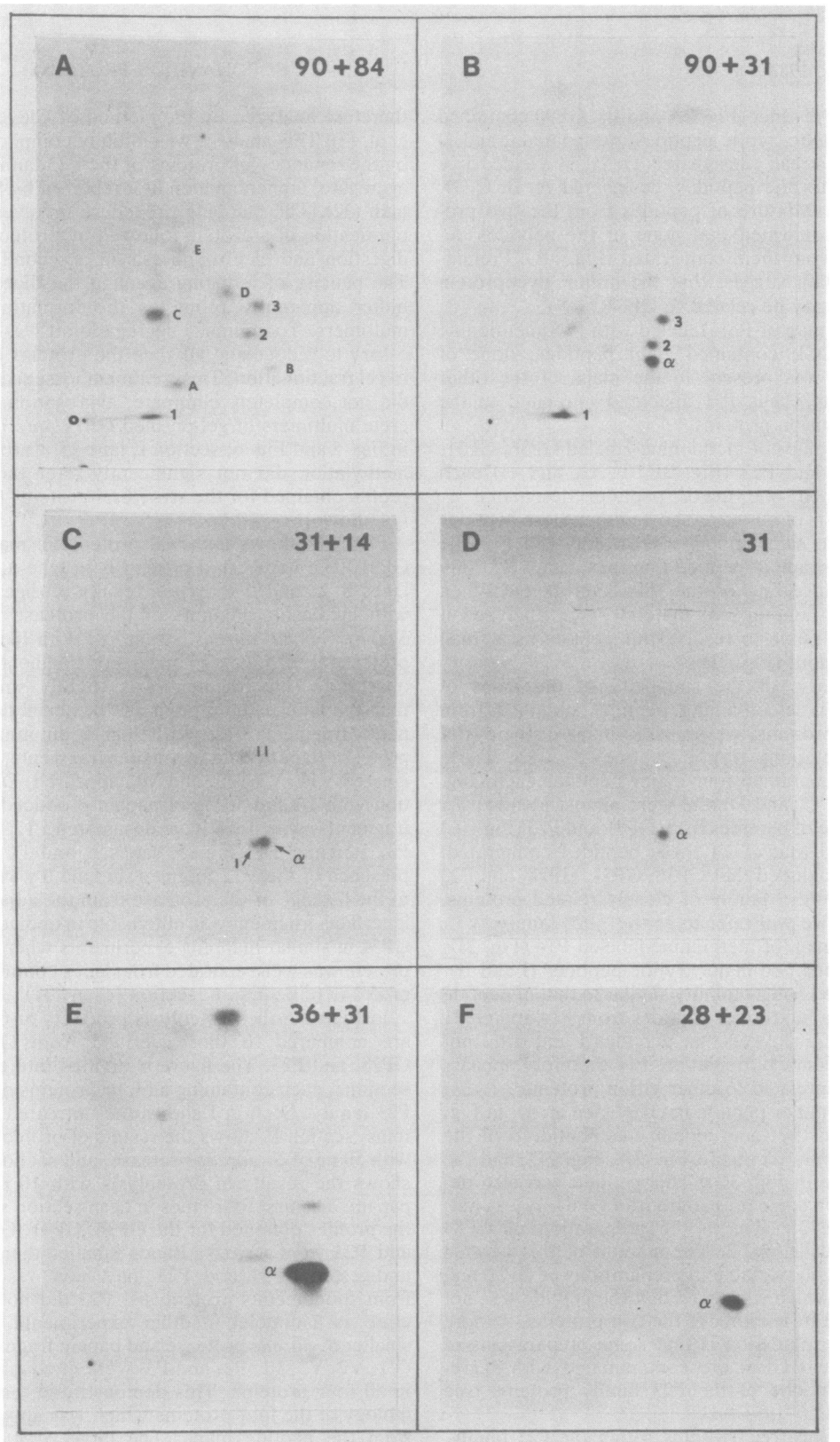

FIG. 3. Maps of mixtures of methionine-containing tryptic peptides of IBV virion proteins. Maps of most of the proteins analyzed in this figure are also shown in Fig. 2. However, GP31 in panels B, C, and D was from a different preparation than the other maps of this protein. Mixtures of peptides obtained from the virion proteins were analyzed in two dimensions. Exposures were for 21 or 28 days. "O" indicates the origin in panel A. The remaining symbols are discussed in the text. Panels: A, GP90 plus GP84; B, GP90 plus GP31; C, GP31 plus P14; D, GP31; E, GP36 plus GP31; and F, GP28 plus GP23.

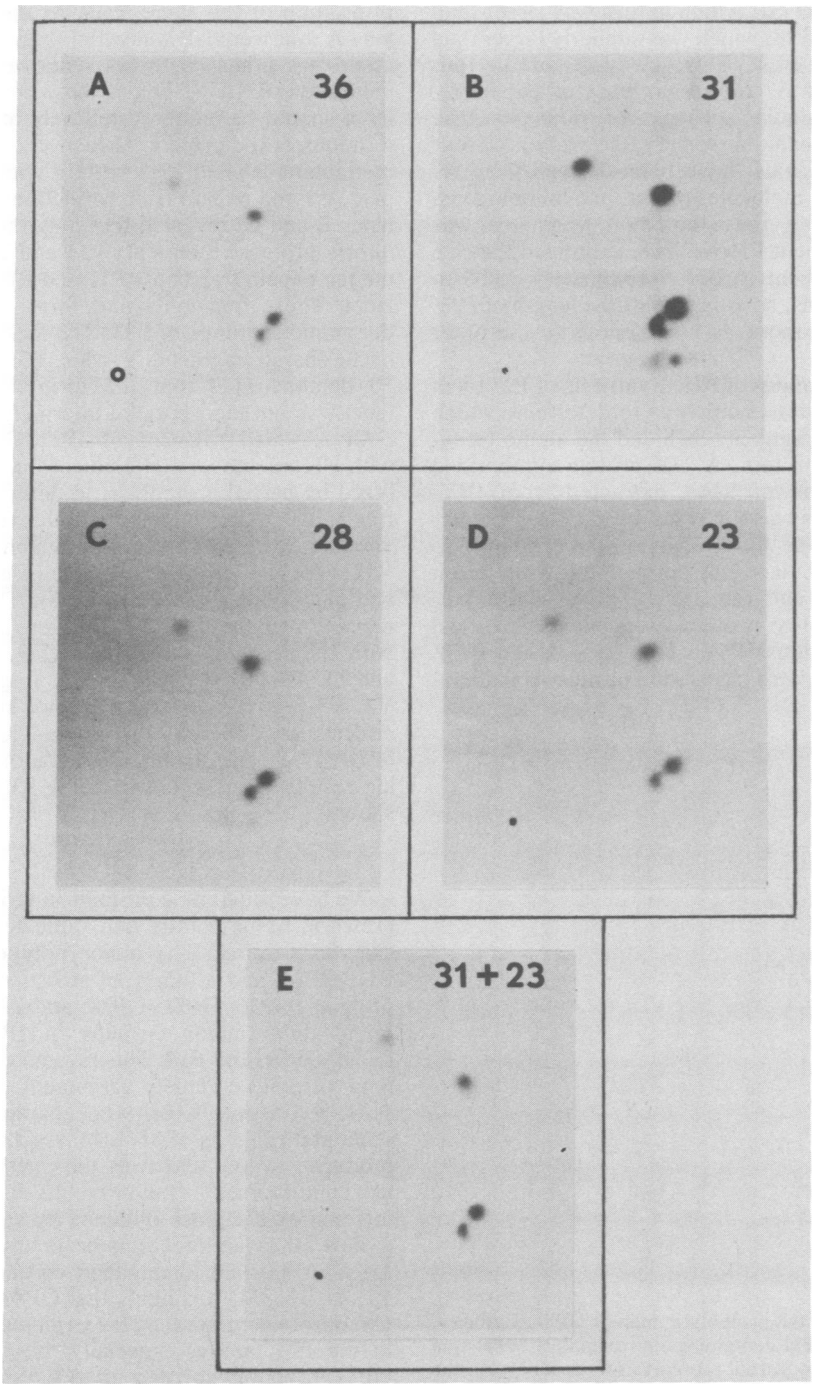

FIG. 4. Two-dimensional tryptic peptide maps of IBV virion proteins labeled with [^{35}S]cysteine. Virions were labeled with [^{35}S]cysteine, and tryptic digests were prepared for two-dimensional mapping as described in the text; 10,000 to 40,000 cpm was applied to each plate, and fluorography was for 5 days. "O" designates the origin in panel A. Panels: A, GP36; B, GP31, C, GP28; D, P23; and E, GP31 plus P23.

intact proteins vary. For example, the V12 fragment from GP36 was slightly larger than the V12 fragment of GP31, which in turn was larger than the V12 of GP28, which was similarly larger than the V12 of P23. Thus, cleavage of all four proteins with *S. aureus* protease or papain resulted in a similar spectrum of fragments. The smaller fragments were the same size for all four proteins and must have been derived from regions of the molecules which are homologous. The largest of these common fragments was papain fragment P3. We have estimated the size of this fragment to be approximately 9,000 to 10,000 daltons, so at least half the length of P23, and perhaps more, is homologous to the other proteins.

Amino terminus of P23. Synthesis of P23 by in vitro translation enabled us to determine which proteolytic fragments included the amino terminus of that protein. A 23K protein synthesized by translation of RNA extracted from IBV-infected cells was analyzed by partial proteolysis (Fig. 6, lane A). The undigested protein (Fig. 6, section I, lane A) comigrated with virion protein P23 and the proteolytic cleavage patterns of the two proteins were almost identical (Fig. 6, sections II and III, lanes A and P23). Thus, this in vitro translation product is indistinguishable from P23. The 23K in vitro translation products in lanes B and C were prepared by translation of the same RNA as the protein in lane A, but were labeled with [^{35}S]formyl methionine by synthesis in the presence of [^{35}S]formyl methionyl $tRNA_f^{met}$. Under these conditions the label should be incorporated only at the amino terminus of the protein. Digestion of the formyl methionine-labeled 23K protein with *S. aureus* protease and papain (Fig. 6, sections II and III, lanes B and C) revealed that only the unique *S. aureus* protease fragments V11 and V12 and the unique papain fragments P11 and P12 contained label. These fragments therefore must include the amino terminus of P23. The lack of labeling of the shared fragments V1, V2, V3, P1, P2, and P3 demonstrated that the formyl methionine label was not incorporated internally.

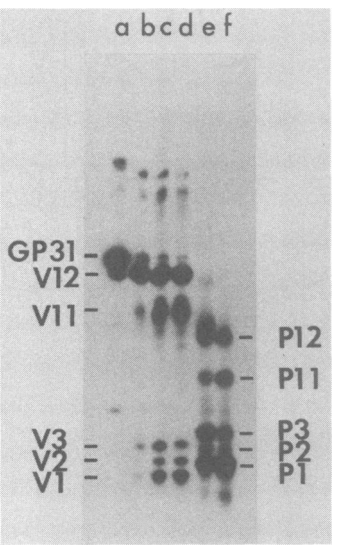

FIG. 5. Partial proteolytic map of GP31. Carboxymethylated GP31 containing approximately 3,000 cpm was mapped by partial proteolysis as described in the text. The proteolytic cleavage fragments are designated as discussed in the text. The fluorograph was exposed for 60 days. Lanes: a, no protease; b, GP31 digested with 10 ng of *S. aureus* V8 protease (Miles Laboratories); c, 50 ng of *S. aureus* protease; d, 100 ng of *S. aureus* protease; e, 10 ng of papain (Sigma); and f, 50 ng of papain.

Glycosylated fragments of GP31. GP31 labeled with [^{3}H]mannose or [^{3}H]glucosamine was analyzed by partial proteolysis to determine which fragments contain oligosaccharides (Fig. 7). The variable *S. aureus* protease fragments V11 and V12 were labeled with mannose (Fig. 7, lane 1) and glucosamine (Fig. 7, lane 3). The variable papain fragments P11 and P12 were also labeled with the ^{3}H-monosaccharides (Fig. 7, lanes 4 and 6). However, the common fragments V1, V2, V3, P1, P2, and P3 were not labeled with either sugar. Therefore, the variable fragments of GP31 bear the oligosaccharide moieties. Similar results were obtained for GP36 (data not shown).

DISCUSSION

We have found nine proteins in virions of IBV grown in tissue culture and subjected to extensive purification. The major polypeptides are P51, GP84, and a family of structurally related proteins, GP36, GP31, GP28, and P23. We also found three minor proteins in IBV virions, GP90, GP59, and P14. This protein composition is in virtually complete agreement with the results of Cavanagh (6), who characterized the structural proteins of M41 (Massachusetts) IBV produced in ovo and with the results of Lomniczi and Morser (14). Our results also resemble to some extent those obtained by several other groups for avian infectious bronchitis virus (12, 16, 17, 34a) and mammalian coronaviruses (4, 15, 20, 32, 35). It is likely that GP84 is the large glycoprotein comprising the peplomers. Proteins in the P23 family, especially the predominant GP31, probably correspond to the smaller major glycoprotein(s) described by other groups.

Several observations suggest that GP84, P51, GP36, GP31, GP28, and P23 are virus-induced or virus-encoded proteins: (i) infection of CEK cells with IBV results in de novo synthesis of proteins comigrating with GP84, P51, G31, and

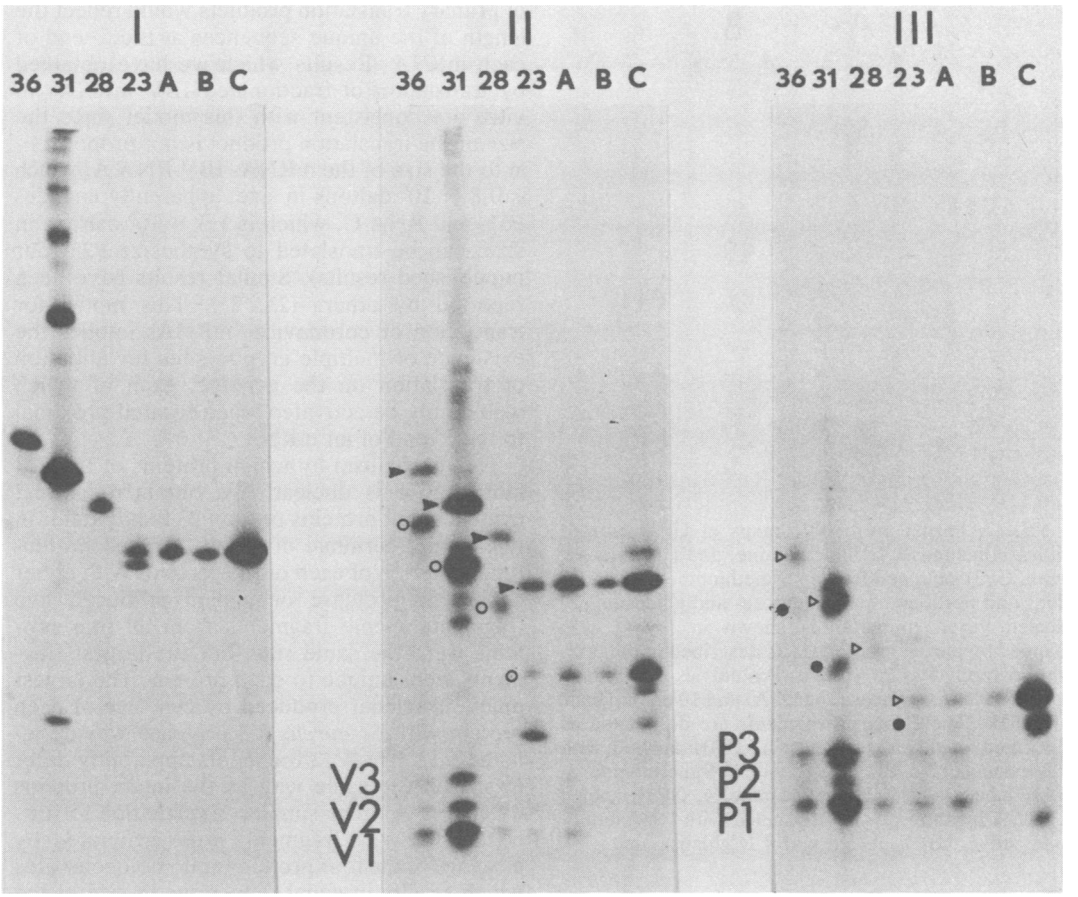

FIG. 6. Partial proteolytic maps of the P23 family and the 23K in vitro translation product. A 23K protein was synthesized by in vitro translation of RNA isolated from IBV-infected cells as described in the text. Gel-purified RNA C was translated for lanes A and B, and unfractionated RNA was translated for lane C. The 23K protein was labeled by translation in the presence of [^{35}S]methionine (lanes A) or [^{35}S]formyl methionyl $tRNA_f^{met}$ (lanes B and C). [^{35}S]methionine-labeled virion proteins and the in vitro translation products were carboxymethylated and mapped by partial proteolysis as described in the text. Fluorography was for 90 days. Sections: I, undigested proteins; II, products of digestion with 50 ng of *S. aureus* protease; and III, products of digestion with 10 ng of papain. Lanes 36, 31, 28, and 23 contain virion proteins GP36, GP31, GP28, and P23. Lanes A contain the 23K in vitro translation product synthesized by translation of RNA C and labeled with [^{35}S]methionine; lanes B contain the 23K product synthesized from RNA C and labeled with [^{35}S]formyl methionine; and lanes C contain the 23K product synthesized from unfractionated RNA and labeled with [^{35}S]formyl methionine. The unique *S. aureus* protease and papain fragments are marked as follows: ▶, V12; ○, V11; ▷, P12; ●, P11.

P23; (ii) proteins comigrating with GP84, P51, GP31, GP28, and P23 can be immunoprecipitated from infected, but not mock-infected, cell lysates with rabbit antiserum raised against purified IBV virions; and (iii) P51 and P23 can be translated in vitro from cytoplasmic RNA extracted from IBV-infected, but not mock-infected, cells.

We do not yet know whether the minor proteins are virus-specific or host cell contaminants. It is clear from our tryptic peptide maps that GP90 and P14 are unrelated to the remaining structural proteins. Cavanagh (6) concluded that GP94, which may be analogous to our GP90, is a virus-specific protein, but that P14 was host specific. Interestingly, Siddell et al. (26) have detected a 14K protein synthesized in cells infected with the murine coronavirus JHM virus which was not found in JHM virus virions.

Two-dimensional tryptic peptide mapping and one-dimensional mapping by limited proteolysis demonstrated that GP36, GP31, GP28, and P23 constitute a group of related proteins which we have designated the P23 family. It is probable that such a family of proteins is a feature of other coronaviruses, since other groups have reported a heterogeneous series of proteins in this size range (6, 21, 26). For example, Siddell et al. (26)

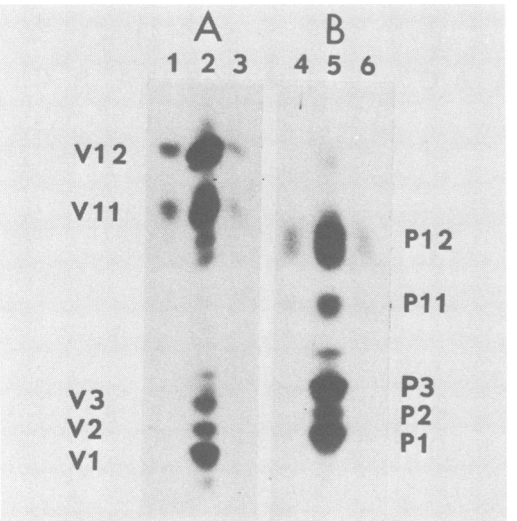

FIG. 7. Partial proteolytic maps of GP31 labeled with [^{3}H]mannose, [^{35}S]methionine, and [^{3}H]glucosamine. GP31 obtained from [^{35}S]methionine-labeled virions and portions of the mannose- and glucosamine-labeled virion preparations shown in Fig. 1 was mapped by partial proteolysis as described in the text. Fluorography was for 42 days. Proteolysis was with 50 ng of *S. aureus* protease (panel A) and 10 ng of papain (panel B). The cleavage fragments are designated as described in the text. Lanes: 1, GP31 labeled with [^{3}H]mannose; 2, GP31 labeled with [^{35}S]methionine; 3, GP31 labeled with [^{3}H]glucosamine; 4, GP31 labeled with [^{3}H]mannose; 5, GP31 labeled with [^{35}S]methionine; and 6, GP31 labeled with [^{3}H]glucosamine.

identified a 23K non-glycosylated protein in the murine coronavirus JHM as well as a series of slightly larger glycoproteins, and Cheley and Anderson (6a) have shown that two such mouse hepatitis virus proteins are structurally related.

Our results imply that as few as three and as many as nine viral genes may encode the IBV structural proteins. Since there are at least five subgenomic mRNAs it is possible that each virion protein is encoded by a different mRNA.

The IBV mRNAs comprise a 3′-coterminal nested set. Each mRNA contains all the sequences of the smaller RNA species and additional "unique" sequences at its 5′ end (30, 31). Since the mRNAs overlap, it is conceivable that the IBV proteins would be synthesized by translation of the same region in different mRNAs. However, we have shown that the IBV proteins, excepting the P23 family, are likely to be structurally distinct, so any overlapping translation in the same reading frame must only occur over short regions. The simplest hypothesis is that each mRNA is translated only over the unique portion at its 5′ end, as occurs with the two alphavirus mRNAs. A prediction of this non-overlapping translational model is that the sizes of primary translation products would reflect the length of the unique sequences at the 5′ end of each mRNA. Results which we have obtained by translation of fractionated IBV mRNAs in vitro are consistent with this model since the size of the translation product is not proportional to the size of the mRNA. IBV RNA A, which is 0.8×10^6 daltons in size, apparently encodes P51, and RNA C, which is 1.3×10^6 daltons in size, can be translated to synthesize P23 (our unpublished results). Similar results have been reported by others (21, 25). This model for translation of coronavirus mRNAs implies the existence of multiple cryptic sites for initiation of translation on the genome, each of which would only be activated when located proximal to the 5′ end of an mRNA.

The mechanism by which proteins of the P23 family arise is unclear. We obtained indirect evidence that proteins of the P23 family differ in their amino-terminal domains. Limited proteolytic digestion of each of the proteins with either *S. aureus* protease or papain produced five fragments. Some fragments from all four proteins were the same size, but the largest fragments were unique to each protein. The largest unique fragment produced by cleavage of each protein with *S. aureus* V8 protease was designated V12. The V12 fragments apparently differ in size in the same way as the intact proteins differ (Fig. 6). The simplest explanation for this is that the V12 fragments were generated by cleavage of the four proteins at homologous sites which results in equal reductions in size of the proteins. We showed that the V12 fragments derived from P23 labeled with [^{35}S]formyl-methionine retain the label. Thus, the P23 V12 fragment contains the amino terminus of P23. If all of the V12 fragments are in fact generated by cleavage at the same site, then each should also contain the amino terminus of the proteins from which it is derived. Thus, it is likely that the proteins of the P23 family differ in their amino terminal domains.

The P23 family could arise from synthesis of four primary translation products, by differential processing of a single translation product, or through some combination of both schemes. Some of the differences among the P23 family proteins must be ascribed to glycosylation since P23 is not glycosylated and the remaining proteins are. It is possible that the four proteins are variants of P23 which are glycosylated to different extents and contain identical primary amino acid sequences. This model is consistent with the fact that P23 can be synthesized in vitro and the finding that the unique partial proteolytic fragments of GP36 and GP31 are glycosylated.

We are now attempting to distinguish between these possibilities with pulse-chase experiments

and by the generation of non-glycosylated forms of GP36, GP31, and GP23 by using tunicamycin and endoglycosidase H. We are also translating the fractionated viral RNAs in vitro. This should reveal the unprocessed form of each of the viral polypeptides. In addition, it will permit determination of the coding assignments for the viral mRNAs, which, since they have been mapped to the genome, will reveal the genetic map of the virus.

ACKNOWLEDGMENTS

We thank Tony Hunter for discussions and Jon Cooper for advice about the manuscript.

D.F.S. and L.B. were supported by Public Health Service training grants GM 07313 and CA 09345 from the National Institutes of Health. This work was supported by Public Health Service grants CA 14195 and CA 17289 from the National Institutes of Health.

LITERATURE CITED

1. **Beemon, K., and T. Hunter.** 1978. Characterization of Rous sarcoma virus *src* gene products synthesized in vitro. J. Virol. **28**:551–566.
2. **Bingham, R. W.** 1975. The polypeptide composition of avian infectious bronchitis virus. Arch. Virol. **49**:207–216.
3. **Bonatti, S., N. Sonenberg, A. J. Shatkin, and R. Cancedda.** 1980. Restricted initiation of protein synthesis on the potentially polycistronic Sindbis virus 42S RNA. J. Biol. Chem. **255**:11473–11477.
4. **Bond, C. W., J. L. Leibowitz, and J. A. Robb.** 1979. Pathogenic murine coronaviruses. II. Characterization of virus-specific proteins of murine coronaviruses, JHMV and A59V. Virology **94**:371–384.
5. **Bonner, W. M., and J. D. Stedman.** 1978. Efficient fluorography of ^{3}H and ^{14}C on thin layers. Anal. Biochem. **89**:247–256.
6. **Cavanagh, D.** 1981. Structural polypeptides of coronavirus IBV. J. Gen. Virol. **53**:93–103.

6a. **Cheley, S., and R. Anderson.** 1981. Cellular synthesis and modification of murine hepatitis virus polypeptides. J. Gen. Virol. **54**:301–311.

7. **Cleveland, D. W., S. G. Fischer, M. W. Kirschner, and U. K. Laemmli.** 1977. Peptide mapping by limited proteolysis in sodium dodecyl sulfate and analysis by gel electrophoresis. J. Biol. Chem. **252**:1102–1106.
8. **Collins, M. S., and D. J. Alexander.** 1980. Avian infectious bronchitis virus structural polypeptides: effect of different conditions of disruption and comparison of different strains and isolates. Arch. Virol. **63**:239–251.
9. **Davies, H. A., R. R. Dourmashkin, and M. R. MacNaughton.** 1981. Ribonucleoprotein of avian infectious bronchitis virus. J. Gen. Virol. **53**:67–74.
10. **Hunter, T., and J. I. Garrels.** 1977. Characterization of the mRNAs for α-, β-, and γ-actin. Cell **12**:767–781.
11. **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.
12. **Lanser, J. A., and C. R. Howard.** 1980. The polypeptides of infectious bronchitis virus (IBV-41 strain). J. Gen. Virol. **46**:49–361.
13. **Leibowitz, J. L., K. C. Wilhelmsen, and C. W. Bond.** 1981. The virus-specific intracellular RNA species of two murine coronaviruses: MHV-A59 and MHV-JHM. Virology **114**:39–51.
14. **Lomniczi, B., and J. Morser.** 1981. Polypeptides of infectious bronchitis virus. I. Polypeptides of the virion. J. Gen. Virol. **55**:155–164.
15. **MacNaughton, M. R.** 1980. The polypeptides of human and mouse coronaviruses. Arch. Virol. **63**:75–80.
16. **MacNaughton, M. R., and J. H. Madge.** 1977. The polypeptide composition of avian infectious bronchitis virus particles. Arch. Virol. **55**:47–54.
17. **Nagy, E., and B. Lomniczi.** 1979. Polypeptide patterns of infectious bronchitis virus serotypes fall into two categories. Arch. Virol. **61**:341–345.
18. **Papkoff, J., T. Hunter, and K. Beemon.** 1980. In vitro translation of virion RNA from Moloney murine sarcoma virus. Virology **101**:91–103.
19. **Pelham, H. R. B., and R. J. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
20. **Robb, J. A., and C. W. Bond.** 1979. Coronaviridae, p. 193–247. *In* H. Fraenkel-Conrat and R. R. Wagner (ed.), Comprehensive virology, vol. 14. Plenum Publishing Corp., New York.
21. **Rottier, P. J. M., W. J. Spaan, M. C. Horzinek, and B. A. M. van der Zeijst.** 1981. Translation of three mouse hepatitis virus strain A59 subgenomic RNAs in xenopus laevis oocytes. J. Virol. **38**:20–26.
22. **Schofield, P., and P. C. Zamecnik.** 1968. Cupric ion catalysis in hydrolysis of amino acyl-tRNA. Biochim. Biophys. Acta **155**:410–416.
23. **Sefton, B. M., K. Beemon, and T. Hunter.** 1978. Comparison of the expression of the *src* gene in vitro and in vivo. J. Virol. **28**:957–971.
24. **Siddell, S. G., A. Barthel, and V. Ter Meulen.** 1981. Coronavirus JHM: a virion-associated protein kinase. J. Gen. Virol. **52**:235–243.
25. **Siddell, S., H. Wege, A. Barthel, and V. Ter Meulen.** 1980. Coronavirus JHM: cell-free synthesis of structural protein p60. J. Virol. **33**:10–17.
26. **Siddell, S., H. Wege, A. Barthel, and V. Ter Meulen.** 1981. Coronavirus JHM: Intracellular protein synthesis. J. Gen. Virol. **53**:145–155.
27. **Simmons, D. T., and J. H. Strauss.** 1974. Translation of Sindbis virus 25S RNA and 49S RNA in lysates of rabbit reticulocytes. J. Mol. Biol. **86**:397–409.
28. **Spaan, W. J. M., P. J. M. Rottier, M. C. Horzinek, and B. A. M. van der Zeijst.** 1981. Isolation and identification of virus-specific mRNAs in cells infected with mouse hepatitis virus (MHV-A59). Virology **108**:424–434.
29. **Stanley, W. M., Jr.** 1972. Preparation and analysis of L-[^{35}S]-methionine labeled transfer ribonucleic acids from rabbit liver. Anal. Biochem. **48**:202–216.
30. **Stern, D. F., and S. I. T. Kennedy.** 1980. The coronavirus multiplication strategy. I. Identification and characterization of virus-specific RNA. J. Virol. **34**:665–674.
31. **Stern, D. F., and S. I. T. Kennedy.** 1980. Coronavirus multiplication strategy. II. Mapping the avian infectious bronchitis virus intracellular RNA species to the genome. J. Virol. **36**:440–449.
32. **Sturman, L. S.** 1977. Characterization of a coronavirus. I. Structural proteins: effects of preparative conditions on the migration of protein in polyacrylamide gels. Virology **77**:637–649.
33. **Sturman, L. S., and K. V. Holmes.** 1977. Characterization of a coronavirus. II. Glycoproteins of the viral envelope: tryptic peptide analysis. Virology **77**:650–660.
34. **Sturman, L. S., K. V. Holmes, and J. Behnke.** 1980. Isolation of coronavirus envelope glycoproteins and interaction with the viral nucleocapsid. J. Virol. **33**:449–462.

34a. **Wadey, C. N., and E. G. Westaway.** 1981. Structural proteins and glycoproteins of infectious bronchitis virus particles labeled during growth in chick embryo cells. Intervirology **15**:19–27.

35. **Wege, H., H. Wege, K. Nagashima, and V. Ter Meulen.** 1979. Structural polypeptides of the murine coronavirus JHM. J. Gen. Virol. **42**:37–47.
36. **Welch, W. J., B. M. Sefton, and F. S. Esch.** 1981. Amino terminal sequence analysis of alphavirus polypeptides. J. Virol. **38**:968–972.