JOURNAL OF VIROLOGY, Mar. 1990, p. 1241–1249
0022-538X/90/031241-09$02.00/0

Vol. 64, No. 3

Identification of Precursor to Cytomegalovirus Capsid Assembly Protein and Evidence that Processing Results in Loss of Its Carboxy-Terminal End

WADE GIBSON,* ALICE I. MARCY,† JAMES C. COMOLLI,‡ AND JOANNA LEE§

Virology Laboratories, Department of Pharmacology and Molecular Sciences, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205

Received 2 October 1989/Accepted 30 November 1989

The 37-kilodalton (kDa) assembly protein of cytomegalovirus (strain Colburn) B capsids is shown to have a 40-kDa precursor. Pulse-chase radiolabeling experiments revealed that conversion of the precursor to the product was slow, requiring over 6 h for completion, and correlated with movement from the cytoplasmic to the nuclear fraction of Nonidet P-40-disrupted cells. Of these two proteins, only the 40-kDa precursor was synthesized in vitro from infected-cell RNA, consistent with its being the primary translation product. Amino acid sequence data obtained from CNBr-treated, high-performance liquid chromatography-purified assembly protein indicated that precursor translation begins at the first of two closely spaced potential initiation sites and that precursor maturation involves the loss of at least 32 amino acids from its carboxy-terminal end. It is also shown by immunological cross-reactivity and peptide similarity that three low-abundance B-capsid proteins (i.e., the 45-kilodalton [45K], 39K, and 38K proteins) are closely related to the assembly protein; the nature of this relatedness is discussed.

The capsid in mature virions of cytomegalovirus (CMV) has a protein composition different from that of the predominant intranuclear capsid form that is thought to be its precursor (9, 10, 17). Both contain as their most abundant constituents the major capsid protein (MCP), the minor capsid protein (mCP), and the smallest capsid protein. The principal intranuclear capsid form (i.e., the B capsid), however, is distinguished by having an additional abundant protein species referred to as the assembly protein (9, 17). As summarized elsewhere (9, 17), the assembly protein has a direct counterpart in herpes simplex virus (HSV) (i.e., VP22a [also called NCP-3, p40, and ICP35e]) and appears to represent a species common to the herpesvirus group. General characteristics shared by the CMV and HSV assembly proteins are their sizes of 35 to 40 kilodaltons (kDa), slightly positive net charge, unusual staining properties, distinction as the only abundant phosphorylated capsid protein, and absence from the mature virion.

Although the function of the assembly protein is not known, evidence that this protein is associated with capsids lacking DNA (16, 19, 22–24, 26) has fostered the idea that its role may be similar to that of the bacteriophage scaffolding protein, which facilitates empty-prohead assembly (18). The identification of an HSV temperature-sensitive mutant (*ts*1201) that produces intranuclear capsids but is unable to fully process the assembly protein precursor and fails to package DNA at the restrictive temperature suggests a link between precursor processing and DNA packaging (24). However, CMV capsids lacking DNA yet having a processed assembly protein demonstrate that processing is not necessarily accompanied by DNA encapsidation (16, 17). Further, pulse-chase radiolabeling experiments done in conjunction with immunoprecipitations using monoclonal antibodies to p40 or ICP35 showed that the HSV assembly protein counterpart has multiple related forms (2, 24), indicating that its processing involves several steps.

The work reported here identifies the CMV assembly protein precursor and provides a basis for studying its processing, comparing its synthesis and modification with those of counterpart proteins from other herpesviruses, and studying the structure and expression of the CMV assembly protein gene.

MATERIALS AND METHODS

Cells, virus, and fractionation procedures. Human foreskin fibroblasts were prepared, cultured, and infected with CMV Colburn (simian strain) as described previously (9). Cells were separated into nuclear and cytoplasmic fractions by using 0.5% Nonidet P-40 (NP-40) in phosphate buffer (40 mM sodium phosphate–150 mM NaCl, pH 7.4) containing 1 mM phenylmethylsulfonyl fluoride as described previously (9). Capsids were recovered from the NP-40 nuclear fraction of infected cells by sonication followed by rate-velocity sedimentation in linear gradients of sucrose (15 to 50% in phosphate buffer) and were concentrated by pelleting, all as described elsewhere (22).

RNA isolation, in vitro translation, and SDS-PAGE. Total-cell RNA was prepared from CMV Colburn-infected fibroblasts approximately 70 h postinfection by disrupting the cells in guanidine thiocyanate (4) and recovering the RNA by pelleting it through a CsCl cushion (14). In vitro translations were done with a rabbit reticulocyte lysate preparation (Amersham Corp., Arlington Heights, Ill.) and [^{35}S]methionine, according to manufacturer recommendations. Protein separations were done by (i) size (sodium dodecyl sulfate-polyacrylamide gel electrophoresis [SDS-PAGE]) (9, 20), (ii) electrofocusing followed by size (charge/size two-dimen-

* Corresponding author.
† Present address: Department of Biophysical Chemistry, Merck Sharp & Dohme Research Laboratories, Rahway, NJ 07065.
‡ Present address: Biochemistry, Molecular, Cellular, and Developmental Biology Program, College of Arts and Sciences, Harvard University, Cambridge, MA 02138.
§ Present address: Department of Toxicology, The Johns Hopkins University School of Public Health and Hygiene, Baltimore, MD 21205.

sional [2-D] PAGE) (10), or (iii) nonreducing SDS-PAGE followed by reducing SDS-PAGE (nonreducing/reducing 2-D SDS-PAGE) (see below). Radiolabeled proteins in gels were detected by fluorography (1, 3, 21); densitometric measurements from fluorograms were made as described before (9).

Peptide comparisons. Peptide comparisons following partial proteolysis were done essentially as described by Cleveland et al. (5); cleavage products were resolved with a 16% polyacrylamide gel cross-linked with diallyltartardiamide (17).

2-D separations of tryptic peptides were done as follows. Proteins were separated by SDS-PAGE, and the resulting gel was soaked for 30 min in water containing a mixed bed resin (AG 501-X8; Bio-Rad Laboratories, Richmond, Calif.) to reduce the amount of SDS and other ions. The gel was then soaked in 1 M sodium salicylate for 1 h (3) and dried at 60°C, and protein bands were located by alignment of positioning dots (Glow Juice; IBI, New Haven, Conn.) following fluorography. Bands of interest were excised from the gel, rehydrated, soaked extensively in water to eliminate sodium salicylate, equilibrated in 50 mM sodium bicarbonate, and pulverized by forcing them through a stainless steel screen (CX-60; Small Parts, Inc., Miami, Fla.) at the bottom of a syringe barrel. The remaining gel "paste" was removed from the screen and syringe tip by rinsing with a small volume of 50 mM sodium bicarbonate. Tolylsulfonyl phenylalanyl chloromethyl ketone-trypsin (Worthington/Du Pont, Wilmington, Del.; 10 to 20 μg/ml of gel slurry) was added, the suspension was incubated with gentle mixing for 12 to 18 h at room temperature, and then 10 μg of fresh trypsin was added and incubation was continued for 2 h. Gel fragments were removed by gently forcing the slurry through two disks of prefilter paper (AP25 042 00; Millipore Corp., Bedford, Mass.); if necessary, the gel fragments were soaked again in 50 mM sodium bicarbonate and similarly filtered to more completely elute the peptides. Clarified peptide-containing solutions were lyophilized, suspended in water, and relyophilized; then they were suspended in peptide electrophoresis buffer (pH 1.9) and concentrated again by lyophilization; finally they were resuspended in 10 μl of peptide electrophoresis buffer and spotted onto the lower left corner of a microcrystalline cellulose thin-layer plate (20 by 20 cm) (no. 5757; E. Merck AG, Darmstadt, Federal Republic of Germany). The plate was immediately dampened by covering it with a piece of blotter paper (23 by 23 cm) prewetted in electrophoresis buffer and having a 1.5-cm hole corresponding in position to the sample origin on the thin-layer plate. Electrophoresis toward the cathode was at 1 kV for 20 min. The plates were then dried and subjected to ascending chromatography in a large chamber (model 500; Shandon Southern Instruments, Inc., Sewickley, Pa.; four plates per chamber) containing butanol-pyridine-acetic acid-water (75:50:15:60) until the solvent front was 1 to 2 cm from the top of the plates.

Radiolabeled peptides were detected in gels by using diphenyloxazole in dimethylsulfoxide (1, 21) or on thin-layer plates by using En^{3}Hance spray (Du Pont, NEN Research Products, Boston, Mass.) according to manufacturer instructions.

Antibody production. CMV Colburn B-capsid proteins were subjected to SDS-PAGE in 10% polyacrylamide gels and visualized by staining with Coomassie brilliant blue (7); the proteins were excised from the gel, completely destained in 10% acetic acid–25% 2-propanol, equilibrated against water for 1 h, and then stored at −80°C until used. In preparation for immunization, gel fragments containing the assembly protein were pulverized by forcing them through a fine screen (CX-60; Small Parts, Inc.), and combined with an equal volume of Freund complete adjuvant (GIBCO Laboratories, Grand Island, NY). This material was injected subcutaneously and intramuscularly into a female New Zealand White rabbit. Booster injections of antigen in Freund incomplete adjuvant were given 5, 11, and 17 weeks later. Approximately 50 μg of assembly protein was used for the primary injection; later injections contained approximately 20 μg each. Two weeks after the final injection, blood was obtained from the ear vein of the rabbit and allowed to clot for 1 h at room temperature. The clot contracted overnight at 4°C and was removed. The resulting serum preparation was clarified by low-speed centrifugation (1,500 × *g* for 5 min at 4°C) followed by high-speed centrifugation (12,000 × *g* for 3 min at 4°C) and stored at −80°C until used.

Immunoassays. Proteins synthesized in rabbit reticulocyte lysates containing [^{35}S]methionine or [^{35}S]cysteine were immunoprecipitated as follows. Ten microliters of lysate was combined with 5 μl of rabbit antiserum to the CMV Colburn assembly protein, and phenylmethylsulfonyl fluoride was added to a final concentration of 1 mM. The reaction mixture was incubated for 60 min at room temperature and then overnight at 4°C with continuous gentle agitation. After incubation, the reaction mixtures were vortexed, transferred into new tubes, combined with a slurry of protein A-Sepharose beads (5 mg per reaction), and incubated for 4 h at 4°C with continuous gentle mixing. Bead-bound antigen-antibody complexes were recovered by low-speed centrifugation, and the beads were rinsed four times in 1 h with 0.1% NP-40 in phosphate buffer. The beads were then rinsed in phosphate buffer without detergent, collected by low-speed centrifugation, and combined with an equal volume of twice-concentrated SDS-PAGE sample buffer (containing or lacking β-mercaptoethanol, as indicated below).

Immunoassay of proteins after electrotransfer to nitrocellulose (27) was done as described before (8, 11). Pronase was included to promote transfer of high-molecular-weight proteins; rabbit antiserum to the CMV Colburn assembly protein (see above) was used at a dilution of 1:100; antibodies were detected by using ^{125}I-protein A followed by fluorography.

Protein sequencing. B capsids were disrupted by heating them at 80°C for 3 min in 50 mM Tris, pH 7.0, containing 0.1% SDS and 0.1% β-mercaptoethanol, and proteins were separated by high-performance liquid chromatography (HPLC). A C-4 column (214TP54; Vydac, Hesperia, Calif.) was developed with a reverse-phase gradient of acetonitrile in 0.1% trifluoroacetic acid (mobile phase, 0 to 80%) and 0.1% trifluoroacetic acid (aqueous phase). At a flow rate of 1 ml/min, the assembly protein eluted at approximately 30 min and at an acetonitrile concentration of approximately 45%. Similar results were obtained when the capsids were subjected to HPLC after being disrupted in 6 M guanidine and 1% β-mercaptoethanol. The assembly protein-containing fraction was lyophilized, and the protein was suspended in acetonitrile-water-acetic acid (30:60:10), spotted onto a Polybrene-treated glass fiber filter, and subjected to gas-phase sequencing (model 470A; Applied Biosystems, Inc., Foster City, Calif.). When no amino acids were observed after four cycles of sequencing, the disk was removed and treated with 50 μl of CNBr (70 mg/ml of 70% formic acid; treatment at room temperature overnight in the dark). After removal of the CNBr by lyophilization, the treated protein was again subjected to gas-phase amino acid sequence analysis.

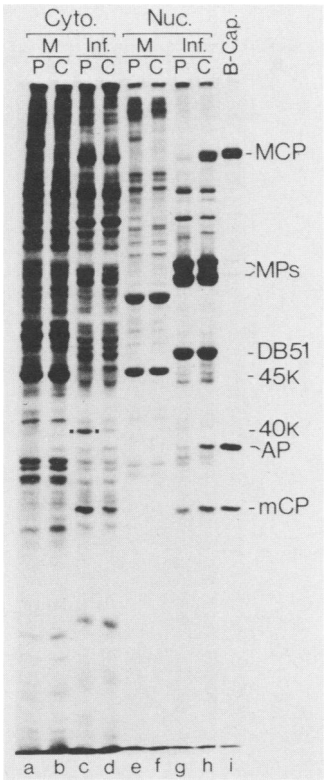

FIG. 1. Identification of possible precursor to CMV Colburn assembly protein. Cells either infected with virus (Inf.) or mock infected (M) were pulse-radiolabeled for 2 h (P) or pulse-labeled and chased for 5 h (C), separated into cytoplasmic (Cyto.) and nuclear (Nuc.) fractions, and subjected to SDS-PAGE followed by fluorography, as described in the text. [^{35}S]methionine-labeled CMV Colburn B-capsid proteins were included as markers (B-Cap.). Proteins are designated as explained elsewhere (9, 17): MCP, 145-kDa MCP; MPs, 69- and 66-kDa upper and lower matrix proteins; DB51, 51-kDa DNA-binding protein; 45K, 45-kDa B-capsid protein; 40K, 40-kDa putative precursor to assembly protein; AP, 37-kDa assembly protein; mCP, 34-kDa mCP. Dots adjacent to lane c indicate the position of the 40K protein.

RESULTS

Pulse-chase radiolabeling identifies apparent precursor to assembly protein. Cultures were labeled for 20 min or 2 h with [^{35}S]methionine (200 μCi/ml in normal medium) 3 days after infection, and then they were either processed immediately for SDS-PAGE (pulse) or rinsed to remove the radiolabel and incubated in fresh medium for an additional 30 min to 24 h (chase) (see legends to Fig. 1 and 3 for specific times). All cell preparations were separated into NP-40 cytoplasmic and nuclear fractions and analyzed by SDS-PAGE (Fig. 1). As noted before (9), the assembly protein became prominent only after a chase interval and was localized to the nuclear fraction (Fig. 1, lane h). An apparent precursor form of the 37-kDa assembly protein was also detected. This slightly larger protein (i.e., the 40,000-molecular-weight [40K] protein) was present primarily in the cytoplasmic fraction of 2-h pulse-labeled cells (Fig. 1, lane c) and had approximately the same radiolabeling intensity as the 5-h chase-labeled assembly protein. 2-D separation (charge/size PAGE) (Fig. 2) showed that this 40-kDa apparent precursor (Fig. 2, pAP) of the assembly protein is slightly more basic in net charge than the assembly protein.

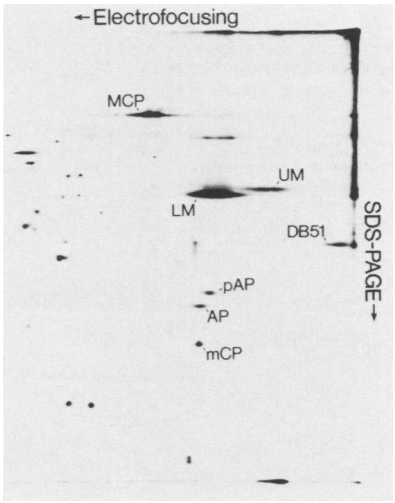

FIG. 2. 40K protein is slightly more basic than assembly protein. Equal volumes of the NP-40 cytoplasmic fraction containing pulse-labeled 40K protein (i.e., the material in Fig. 1, lane c) and the NP-40 nuclear fraction containing labeled assembly protein (i.e., material in Fig. 1, lane h) were combined and subjected to electrofocusing (basic proteins on right, acidic proteins on left) followed by SDS-PAGE and fluorography as described in Materials and Methods. Protein abbreviations: UM and LM, upper and lower matrix proteins, respectively; pAP, 40K protein (putative precursor to assembly protein); others are described in the legend to Fig. 1.

Further evidence for a kinetic relationship between the 40K protein and the assembly protein came from a second pulse-chase experiment in which shorter pulse (20 min) and chase intervals were used (Fig. 3). The results showed that as the 40K protein slowly disappeared from the cytoplasmic fraction over a 24-h period, the amount of assembly protein in the nuclear fraction slowly increased, with an initial lag period of 1 to 2 h. After 6 h, 80% of the radioactivity in the pulse-labeled precursor band was present in the nuclear assembly protein band. This calculation assumes the loss of two methionines from the carboxy terminus of the precursor during processing, as discussed below. Compartmentalization of these two proteins was strong.

40K protein is synthesized in vitro. RNA from Colburn-infected cells was translated in vitro to determine whether either the 40K protein or the assembly protein could be demonstrated to be a primary translation product. When the resulting proteins were separated by SDS-PAGE, a band was seen at the position of the 40K protein, but nothing was seen at the position of the assembly protein (Fig. 4A). To verify that the pulse-labeled intracellular 40K protein and the in vitro synthesized 40K protein were related to each other and to the assembly protein, a peptide comparison was done. [^{35}S]cysteine labeling followed by cyanogen bromide cleavage was considered, but it proved unfeasible because of the absence of cysteine in the assembly protein (Fig. 4B, AP). Parenthetically, all of the other capsid proteins (i.e., MCP, 45K, mCP, 28K, and the smallest capsid protein [Fig. 4B, SCP]) and tegument proteins (e.g., the high-molecular-weight protein, basic phosphoprotein, and the upper and lower matrix proteins) were found to contain cysteine; the 38K and 39K proteins were not detected by cysteine labeling, possibly because of their small amounts (9, 10).

As an alternative to cyanogen bromide, V-8 protease was used as described by Cleveland et al. (5) to partially digest the proteins and compare their peptide patterns. The results

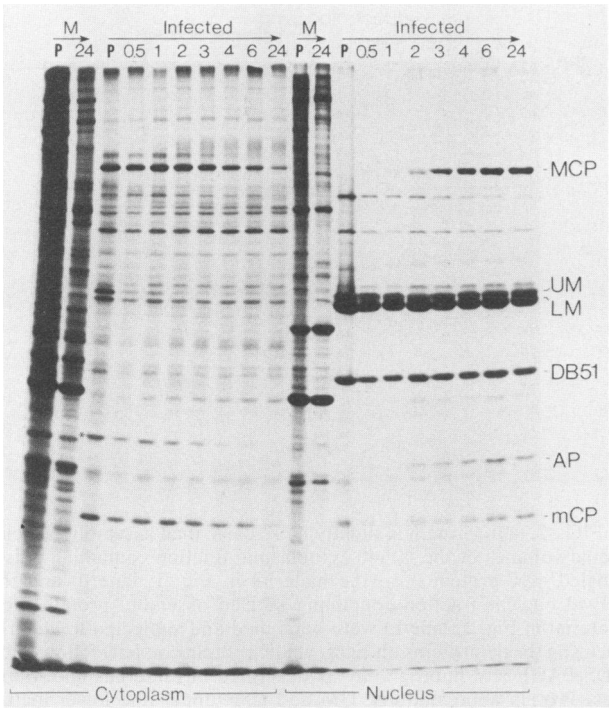

FIG. 3. Pulse-chase radiolabeling of 40K and assembly proteins in CMV Colburn-infected cells. Cells infected with CMV Colburn (Infected) or mock infected (M) were pulse-labeled for 20 min (P) or pulse-labeled and chased for 0.5, 1, 2, 3, 4, 6, or 24 h, separated into cytoplasmic (Cytoplasm) and nuclear (Nucleus) fractions, and subjected to SDS-PAGE followed by fluorography, as described in the text. Protein abbreviations are described in the legends to Fig. 1 and 2; the position of the 40K protein is indicated by an asterisk to the left of the lane containing the pulse-labeled, infected-cell cytoplasmic fraction.

of this experiment (Fig. 4C) demonstrated that intracellular pulse-labeled 40K protein, the 40K protein synthesized in vitro, and the assembly protein gave rise to the same [^{35}S]methionine peptide pattern. The absence of additional or altered peptides in the two 40K protein patterns, compared with that of the 37-kDa assembly protein, suggests (i) that the V-8 cleavage site which generates the 33-kDa peptide from all three proteins is cleaved with high efficiency and (ii) that the cleavage product that would be expected to differ between the assembly protein and the 40K proteins (i.e., 4 kDa versus 7 kDa, respectively), was too small to have been resolved in this gel.

Other B-capsid and in vitro-synthesized proteins immunologically cross-reactive with the assembly protein. Other related forms of the assembly protein were identified by using an antiserum to SDS-PAGE-purified assembly protein, prepared as described in Materials and Methods. The antiserum was specific for the assembly protein, as demonstrated by immunoblotting (Fig. 5, lanes b and c), but it also reacted with three much less abundant B-capsid proteins (lane c) referred to as the 45K, 39K, and 38K proteins (17). The relatedness of these minor species to the assembly protein was confirmed by the similarity of their tryptic peptide patterns following 2-D separations (Fig. 6).

Immunoprecipitation assays to identify immunologically related proteins synthesized in vitro were also done with this antiserum. Because cysteine is present in the 45K protein (Fig. 4B), as well as in the deduced amino acid sequence of

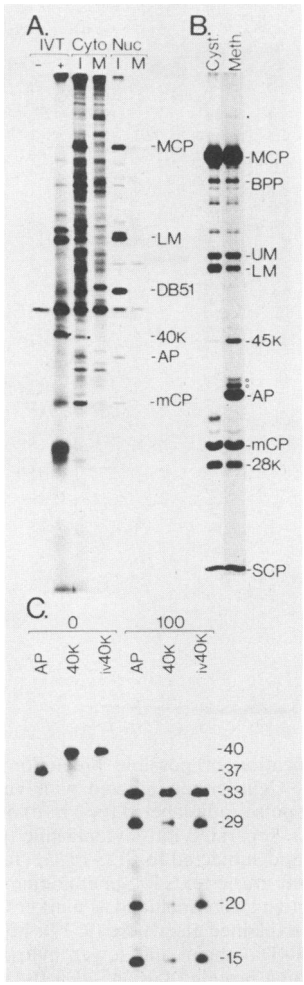

FIG. 4. Comparisons of the 40K and assembly proteins by in vitro synthesis, [^{35}S]cysteine labeling, and partial proteolysis. (A) RNA from Colburn-infected cells was prepared and translated in a rabbit reticulocyte lysate (IVT+) as described in Materials and Methods. A control lysate lacking added RNA was also tested (IVT−). The 40K protein was identified in the NP-40 cytoplasmic fraction (Cyto) of CMV Colburn-infected cells (I) pulse-labeled with [^{35}S]methionine, and the assembly protein (AP) was identified in the corresponding pulsed and chased NP-40 nuclear fraction (Nuc). Parallel fractions from mock-infected cells (M) were also analyzed. Protein abbreviations are described in the legends to Fig. 1 and 2. (B) B capsids were recovered from the NP-40 nuclear fraction of 5-day-infected cells labeled with either [^{35}S]cysteine (200 μCi/ml of normal medium) or [^{35}S]methionine (200 μCi/ml of normal medium). Shown is a fluorogram of the radiolabeled proteins after SDS-PAGE. Protein abbreviations: BPP, 119-kDa basic phosphoprotein; 28K, 28-kDa B-capsid protein; SCP, 10-kDa smallest capsid protein; others are described in the legends to Fig. 1 and 2. Open circles above AP indicate the positions of the 38-kDa and 39-kDa proteins; the dot in the right margin above MCP indicates the position of the high-molecular-weight protein (17). (C) [^{35}S]methionine-labeled assembly protein (AP) from B capsids, 40K protein (40K) from the NP-40 cytoplasmic fraction of pulse-labeled infected cells, and 40K protein from infected-cell RNA translated in vitro (iv40K) were subjected to partial proteolysis during SDS-PAGE as described by Cleveland et al. (5). V-8 protease (100 μg) was added to the three lanes on the right; no protease was added to the three lanes on the left. Protein and fragment sizes (10^3 kDa) are indicated in the right margin.

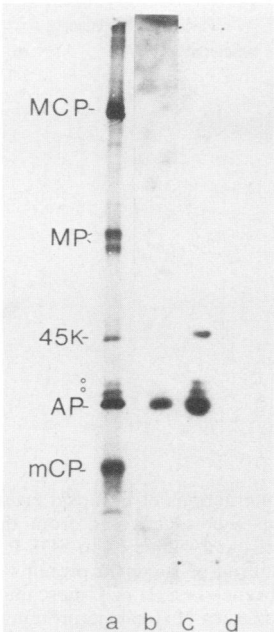

FIG. 5. Specificity of rabbit antiserum to CMV Colburn assembly protein. Unlabeled and [^{35}S]methionine-labeled B capsids were recovered from the NP-40 nuclear fraction of infected cells, subjected to SDS-PAGE and electrotransferred to nitrocellulose, and probed with antiserum to the assembly protein or with a preimmune serum from the same animal. Shown here are the following samples: lane a, [^{35}S]methionine-labeled B-capsid proteins revealed by a long autoradiographic exposure prior to immunoassay; lane b, a comparatively short fluorographic exposure of lane a after Western immunoassay with antiserum to the assembly protein and [^{125}I]protein A used to reveal only immunoreactive bands (note absence of prominent [^{35}S]methionine-labeled bands, i.e., MCP, MP, and mCP); lane c, nonlabeled B-capsid proteins (approximately three times the amount present in lane a) after immunoassay and fluorographic exposure together with lane b (note detection of 45K and 39- and 38-kDa proteins [the last two indicated by open circles in left margin]); lane d, a replicate of lane c reacted with the preimmune serum before fluorographic exposure together with lanes b and c.

the assembly protein precursor (see Fig. 10), both [^{35}S]cysteine and [^{35}S]methionine were tested as labels. The translation products were immunoprecipitated with the anti-assembly protein serum, as described in Materials and Methods, and separated by SDS-PAGE. Fluorograms prepared from the resulting gel showed that over 90% of the radioactivity immunoprecipitated from the [^{35}S]methionine-labeled preparations was present in the 40K band (Fig. 7, lane d). Prolonged exposure revealed three additional minor bands (i.e., less than 1% of the intensity of the 40K band) at approximately 94, 85, and 63 kDa, respectively. The 40K protein was also detected in the [^{35}S]cysteine-labeled preparation (Fig. 7, lane c; the band is faint and may not show in photo reproductions), but the heavy and light chains of the antibody molecules in the immunoprecipitates were labeled too (e.g., asterisks in Fig. 7, lanes c and e), indicating disulfide formation between cysteine residues in these proteins and free [^{35}S]cysteine in the lysate. Nevertheless, whether the observed 40K protein labeling was due to biosynthetic incorporation, nonmetabolic disulfide formation, or both, it indicates the presence of cysteine.

Disulfide interactions of assembly protein and related forms. Confirming the predicted cysteine in the assembly protein precursor was important to the argument that posttranslational processing of the precursor removed its carboxy end (discussed below). Because the [^{35}S]cysteine labeling results described above were ambiguous, the ability of the precursor to form intermolecular disulfide cross-links was tested as an alternative assay for the presence of cysteine. It was first shown by using B capsids that the 45K protein, which contains cysteine (Fig. 4B) and is closely related to the assembly protein (Fig. 5, lane d, Fig. 6), disappeared from the pattern under nonreducing conditions, and a 95-kDa band of about the same intensity appeared (Fig. 8A). The identity of this 95-kDa band as a dimer of the 45K protein was established by 2-D (nonreducing/reducing) SDS-PAGE separation (Fig. 8B). This experiment also showed that the assembly protein did not form intermolecular disulfide linkages (i.e., it was present only on the diagonal), as expected from its lack of cysteine, and that the mCP (i.e., 34-kDa) and 28K proteins were present both as monomers and in disulfide-linked complexes with M_rs which suggest the following compositions: (i) mCP-28K heterodimers (i.e., 60 to 62 kDa), (ii) pentameric forms of the mCP (i.e., 160 to 170 kDa), and (iii) large aggregates unable to enter the first-dimension gel and containing MCP in addition to the mCP and the 28K protein.

When in vitro-synthesized 40K protein (Fig. 9, iv40K) was mixed with B-capsid marker proteins and similarly analyzed, most remained monomeric and appeared on the diagonal;

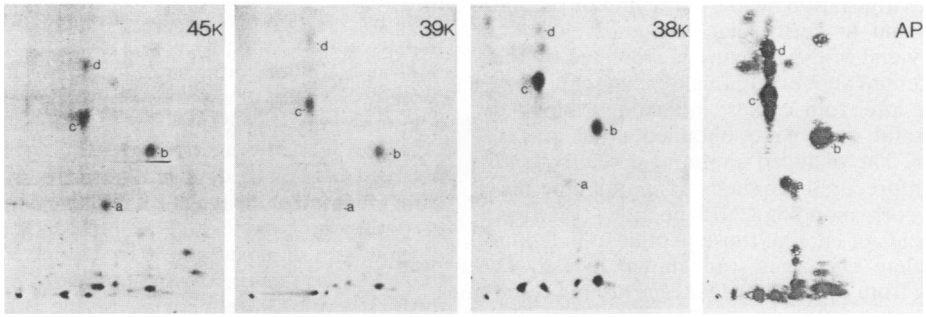

FIG. 6. Tryptic peptide comparison of assembly protein with immunologically related 45-, 39-, and 38-kDa B-capsid proteins. [^{35}S]methionine-labeled B-capsid proteins were separated by SDS-PAGE and processed for tryptic peptide comparison as described in Materials and Methods. The peptides were separated first by electrophoresis (left to right) at pH 1.9 and then by chromatography (bottom to top). The sample origin appears as the lower leftmost spot. Fluorographic exposures were made after the plates were sprayed with En^{3}Hance. Letters denote spots that appeared to be common to all four patterns. The assembly protein (AP) was present in much greater amounts than the other proteins (i.e., 20- to 90-fold) and yielded a correspondingly darker pattern.

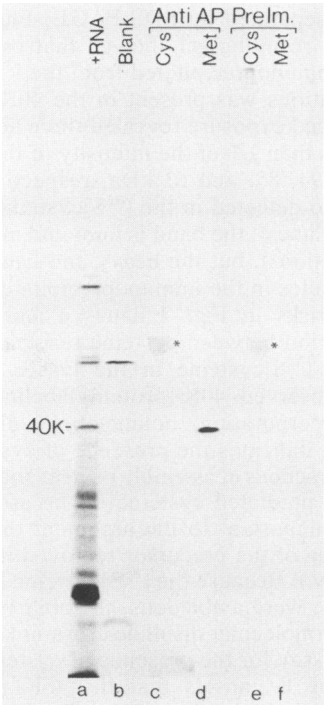

FIG. 7. Antiserum to assembly protein immunoprecipitates in vitro-synthesized 40K protein. RNA was prepared from CMV Colburn-infected cells and translated in rabbit reticulocyte lysates containing [^{35}S]methionine (Met) or [^{35}S]cysteine (Cys); the resulting proteins were subjected to immunoprecipitation with either antiserum to the assembly protein (Anti AP) or a preimmune serum from the same animal (PreIm.), all as described in Materials and Methods. Shown here is a fluorogram of the resulting proteins following SDS-PAGE. The position of the 40K protein is indicated in lane a, which contains [^{35}S]methionine translation products before immunoprecipitation (+RNA). A [^{35}S]methionine-labeled lysate with no added RNA (Blank) is shown in lane b. Asterisks indicate positions of immunoglobulin G heavy chain as determined from the Coomassie brilliant blue staining pattern.

however, some formed homodimers (Fig. 9; indicated by an open circle below the 45K dimer), confirming the presence of cysteine.

Amino acid sequencing. Approximately 200 pmol of HPLC-purified assembly protein was subjected to amino acid sequence analysis in order to substantiate the deduced amino acid sequence of the assembly protein precursor (25), to determine whether translation begins at the first or second potential start site, and to verify that precursor cleavage removes the carboxy end of the protein, as indicated by the cysteine results presented above. Although the intact protein yielded no sequence after four cycles, indicating a blocked amino terminus, useful data were obtained after in situ cleavage with CNBr. The deduced amino acid sequence of the assembly protein precursor contains 11 methionine residues, predicting 11 corresponding CNBr peptides (Fig. 10). The first 10 amino acids of each of these peptides, including products of incomplete cleavages, are shown in Fig. 11 (extensions resulting from incomplete cleavage are indicated by lowercase letters in peptides 1, 4, 5, and 6). With the exceptions of peptides 4 and 11, each of these was predicted to give rise to amino acids in one or more cycles that were unique to it (i.e., "indicator residues" that were not expected at that cycle in the other peptides), and which thereby established the presence of that peptide in the mixture.

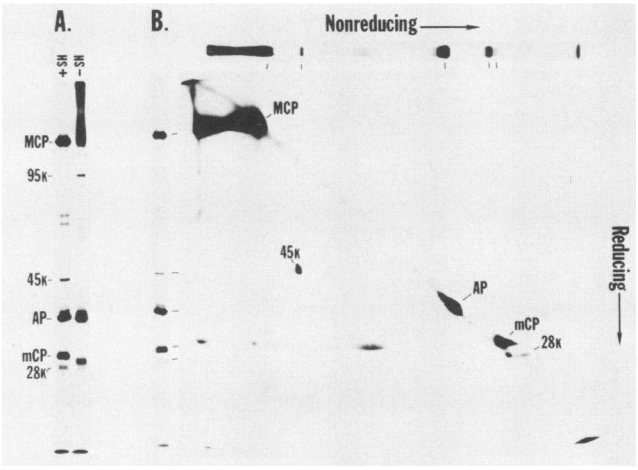

FIG. 8. Disulfide interactions of B-capsid proteins. [^{35}S]methionine-labeled B capsids were recovered from the NP-40 nuclear fraction of infected cells and subjected to SDS-PAGE. (A) Fluorogram showing the patterns of B-capsid proteins solubilized in the presence (+SH) or absence (−SH) of β-mercaptoethanol. (B) Fluorogram showing the pattern of B-capsid proteins after 2-D separation, first in the absence of β-mercaptoethanol (Nonreducing) and then in its presence (Reducing). Single-dimension separations of the same preparation under similar conditions are shown for reference above and to the left of the 2-D pattern.

The data obtained from this experiment are presented in Fig. 11 and can be summarized as follows. First, the presence of all predicted indicator residues (underlined and boldface in Fig. 11) except cycle 9 Arg in peptide 9 (signal too weak to confirm) support the deduced amino acid sequence (Fig. 10) and establish the presence of peptides 2, 3, 5, 6, 7, 8, and 9 in the mature assembly protein. Amino acids that would have been diagnostic for incompletely cleaved peptides were not detected (e.g., Met, Ser, Val, Thr, and Pro in peptide 1; Met, Arg, Asp, and Pro in peptide 4; Met in peptide 5; and Trp in peptide 6). Second, the absence of all predicted indicator amino acids for peptide 10 (i.e., Val, Asp, Leu, Asn, Arg, Leu, Phe, and Val) establishes the absence of this peptide from the mixture and strengthens the argument that maturational processing removes the carboxy end of the assembly protein precursor. Third, the histidine indicator for peptide 1 was not detected, consistent with this peptide representing the amino-terminal end of the protein and containing the blocked amino acid that prevented sequencing of the intact protein. Conversely, all four indicator residues for peptide 2 were detected, thereby establishing that its amino terminus is not blocked and that it therefore does not contain the amino-terminal end of the assembly protein. Finally, the presence of methionine and aspartic acid in the first cycle (not shown in Fig. 11, but 169 and 558 pmol, respectively, relative to 323 pmol for two prolines) is not readily explained. Although methionine could originate from incomplete cleavage by CNBr at the Met-Met junction of peptides 5 and 6 and aspartic acid could result from deamidation of asparagine 1 in peptide 3, it is also possible that these residues reflect unanticipated modifications during the first cycle of sequencing or the presence of contaminants.

DISCUSSION

This report presents evidence that the 37-kDa B-capsid assembly protein of strain Colburn CMV is derived from a

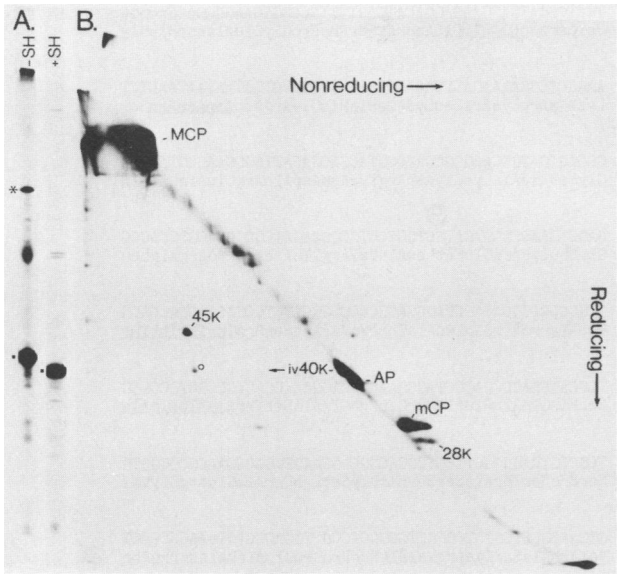

FIG. 9. 40K protein synthesized in vitro forms disulfide-linked homodimers. RNA from infected cells was translated in vitro, and the resulting [^{35}S]methionine-labeled proteins were combined with [^{35}S]methionine-labeled B capsids and subjected to SDS-PAGE. (A) Fluorogram showing the patterns of the mixture solubilized in the presence (+SH) or absence (−SH) of β-mercaptoethanol. ■, monomeric AP; *, dimeric 45K. (B) Fluorogram showing the pattern of the protein mixture after 2-D separation, first in the absence of β-mercaptoethanol (Nonreducing) and then in its presence (Reducing). Protein abbreviations: iv40K, 40-kDa protein synthesized in vitro; others are described in the legends to Fig. 1 and 4. The open circle is just to the right and above the position of the dimeric form of iv40K.

40-kDa precursor by processing that includes eliminating a carboxy-terminal portion of the precursor. Pulse-chase radiolabeling experiments showed that these two proteins exhibited a typical precursor-product relationship (Fig. 3). Essentially no product (i.e., assembly protein) was labeled during the pulse, little precursor (i.e., 40K protein) remained following a chase interval of 6 h or longer, and the final amount of product approximated the starting amount of precursor. Also consistent with a precursor-product relationship between these two proteins was the finding that only the 40K protein was synthesized in vitro (Fig. 7). The relatedness of the 40K proteins synthesized in vivo and in vitro to the assembly protein was established by peptide comparison and immunoprecipitation (Fig. 6 and 7). These results indicate that the 40K protein is the primary translation product of the assembly protein mRNA but do not exclude the possibility that it represents a comparatively stable intermediate derived from a primary product that is very short lived, both in vivo and in vitro.

Modification of the precursor to assembly protein occurred slowly, with a 1- to 2-h lag before the appearance of readily detectable product, and required over 6 h for maximal conversion (Fig. 3). It also correlated with the appearance of the assembly protein in the nuclear fraction and with the movement of MCP and mCP into the nuclear fraction following a 1- to 2-h lag. The reason for this lag in precursor cleavage and its correlation with nuclear partitioning is unknown, but one plausible explanation is that cleavage accompanies or catalyzes the incorporation of NP-40-soluble capsid proteins into comparatively insoluble structures (e.g., procapsid) which then remain in the NP-40 nuclear pellet fraction. In this connection, the observation of capsids presumably containing noncleaved p40 (i.e., the counterpart of assembly protein precursor) in nuclei of cells infected at the restrictive temperature with HSV type 1 mutant *ts*1201 (24) suggests that precursor cleavage follows (perhaps to stabilize) some earlier organizational step.

The estimated 3-kDa difference in size between the assembly protein and its precursor is due to removal of the carboxy-terminal end of the precursor. Three lines of evidence support this conclusion. First, the cysteine residue predicted near the carboxy end of the precursor (Fig. 10) (25) and evidenced here by intermolecular disulfide cross-linking (Fig. 9) is not present in the assembly protein (Fig. 4B and 8). Second, partial sequence analysis of purified assembly protein showed that although a peptide beginning 4 residues from the predicted amino terminus of the precursor was present in the mature protein, a peptide beginning 16 residues from the predicted carboxy terminus of the precursor was absent (i.e., CNBr peptides 2 and 10, respectively; Fig. 10 and 11). Third antipeptide sera to the amino-terminal 13 residues of the deduced precursor (i.e., anti-AP-N1) reacted with the assembly protein as well as the precursor, but a similar antiserum to the carboxy-terminal 22 residues of the precursor (i.e., anti-AP-C1) reacted only with the precursor (J. Comolli, J. Slemmons, P. Schenk, and W. Gibson, unpublished findings). Although the site of cleavage has not been determined, the 3.8-kDa estimated size of the minimal cleavage peptide (i.e., 32 carboxy-terminal residues through cysteine) is close to the 3-kDa observed difference in M_r between the precursor and the assembly protein and suggests that cleavage may occur just before the cysteine residue. We are investigating the possibility that an eight-amino-acid sequence around the cysteine (i.e., -Gly-Val-Val-Asn-Ala-Ser-Cys-Arg-) contains or determines the cleavage site. This sequence is absolutely conserved between the assembly proteins of simian and human strains of CMV (i.e., Colburn and AD169, respectively) and partially conserved (i.e., -Val-Asn-Ala-Ser-) at a corresponding position in the deduced amino acid sequence of the HSV type 1 assembly protein (i.e., p40/ICP35e) precursor.

Resolving the question of whether assembly protein translation begins at the first or second potential start site (i.e., nucleotide 96 or 108, respectively) (Fig. 10) was complicated by the finding that the amino terminus of the HPLC-purified protein was blocked. However, by cleaving the protein with CNBr and then subjecting the resulting mixture of peptides to simultaneous sequencing (Fig. 11), it was determined that translation most likely proceeds from the first methionine. This conclusion was based on the detection of CNBr peptide 2 (i.e., the presence of indicator Thr, Pro, Ser, and Thr in cycles 5, 6, 9, and 10, respectively) but not peptide 1 (i.e., the absence of indicator His in cycle 2), a result which would be expected if peptide 1 contained the blocked amino terminus of the protein and peptide 2 were internal and opened to sequencing by CNBr cleavage. It is noted that the same results could have been obtained if translation began at the second methionine, but only if that initiation methionine was not removed cotranslationally as usual and subsequently became blocked.

It was also demonstrated that three low-abundance B-capsid proteins, referred to as the 45K, 39K, and 38K proteins (17), are closely related to the assembly protein (Fig. 5 and 6). On the basis of results presented here and data from unpublished nucleotide sequencing, peptide comparisons, and experiments with antipeptide sera (unpublished

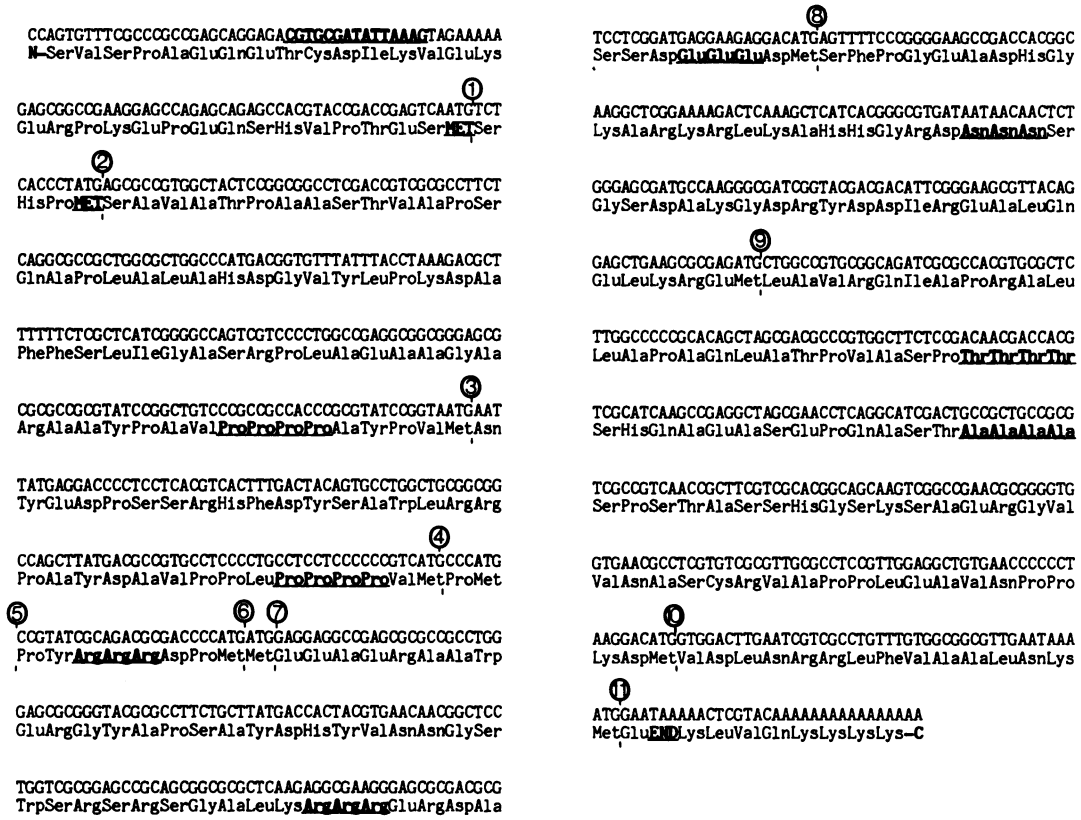

FIG. 10. Deduced amino acid sequence of the CMV Colburn assembly protein precursor containing 11 methionine residues and corresponding CNBr peptides. Shown here for orientation are the previously determined nucleotide and amino acid sequences for the assembly protein (25) with the predicted CNBr cleavage sites numbered sequentially from the amino-terminal end of the protein.

data), the simplest interpretation of the relationship of these proteins to each other and to the assembly protein is as follows. The 39K protein has the same amino- and carboxy-terminal ends as the precursor (i.e., it reacted with anti-peptide sera specific for the amino- and carboxy-terminal portions of the precursor [anti-AP-N1 and anti-AP-C1, respectively]); it comigrated during SDS-PAGE with the assembly protein precursor, whose size was previously estimated to be 40 kDa (25); and it is thought to be a capsid-incorporated assembly protein precursor. The 38K protein has the same amino terminus as the precursor (i.e., it reacted with anti-AP-N1); it lacks the extreme carboxy-terminal portion of the precursor (i.e., it did not react with anti-AP-C1); it has an M_r between those of the 40K precursor and the 37-kDa assembly protein; and it is therefore thought to represent a capsid-incorporated, partially processed form of the precursor (i.e., it retains some, but not all, of the precursor carboxy domain absent from the assembly protein). The 45K protein has the same carboxy terminus as the assembly protein and appears to contain the entire amino acid sequence of the assembly protein but differs from the assembly protein by having a peptide extension at its amino-terminal end that appears to arise from translation beginning at a potential start site 345 nucleotides upstream of and in frame with the assembly protein-coding sequence.

	Cycle Number									
Met	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
1	Ser	[redacted]	Pro	*met*	*ser*	*ala*	*val*	*ala*	*thr*	*pro*
2	Ser	Ala	Val	Ala	**Thr**	**Pro**	Ala	Ala	**Ser**	**Thr**
3	**Asn**	Tyr	**Glu**	**Asp**	**Pro**	**Ser**	**Ser**	**Arg**	**His**	**Phe**
4	Pro	*met*	*pro*	*tyr*	*arg*	*arg*	*arg*	*asp*	*pro*	*met*
5	Pro	Tyr	**Arg**	Arg	Arg	**Asp**	**Pro**	*met*	*met*	*glu*
6	**Met**	*glu*	*glu*	*ala*	*glu*	*arg*	*ala*	*ala*	*trp*	*glu*
7	Glu	Glu	**Ala**	**Glu**	Arg	Ala	Ala	Trp	**Glu**	**Arg**
8	Ser	**Phe**	Pro	**Gly**	Glu	Ala	**Asp**	**His**	**Gly**	**Lys**
9	**Leu**	Ala	Val	Arg	**Gln**	**Ile**	Ala	**Pro**	**Arg**	Ala
10	[redacted]				Arg	[redacted]				Ala
11	Glu									

FIG. 11. Summary of partial amino acid sequence data obtained from purified assembly protein.HPLC-purified assembly protein was subjected to 4 cycles of sequencing without results, cleaved in situ with CNBr, and subjected to another 10 cycles of sequencing, as described in the text. Shown here are the first 10 residues of the 11 predicted CNBr peptides indicated in Fig. 10. Amino acids in bold underlined type are diagnostic (indicators) for the presence of the corresponding CNBr peptides; the only such diagnostic residues that were not detected are highlighted by shading. Amino acids in lowercase italic type (see peptides 1, 3, 4, and 5) would have been detected if CNBr cleavage had not been complete; none was detected.

The recognition of a precursor-product relationship between the 40K protein and the assembly protein further underscores the similarity between this CMV capsid protein and its HSV counterpart VP22a (12, 13; now referred to as p40 [15, 24] or ICP35e [2]), which is also derived from a larger precursor (2, 24). Furthermore, our recent finding that the CMV assembly protein arises from the 3′ end of a much larger gene (i.e., 1,770 bp; L. McNally and W. Gibson, unpublished results) indicates additional similarities in the genetic organization of these counterpart herpesvirus pro-

teins and raises new questions about their transcription, translation, and maturation.

ACKNOWLEDGMENTS

We thank P. Shenbagamurthi of the Medical School Protein/Peptide Facility for help with protein sequencing and W. Lindenmaier and D. McGeoch for sharing their respective human CMV and HSV type 1 nucleotide sequences with us. The experiment presented in Fig. 3 was done by Amina Woods as part of a research rotation project.

A.I.M. and J.L. were predoctoral fellows in the Biochemistry, Cellular, and Molecular Biology Training Program and the Pharmacology and Molecular Sciences Training Program, respectively. J.C.C. was an undergraduate Biology major at the Johns Hopkins University and was supported in part by funds from a work-study program. This work was aided by Public Health Service grants AI 13718 and AI 22711 from the National Institute of Allergy and Infectious Diseases.

LITERATURE CITED

1. **Bonner, W. M., and R. A. Laskey.** 1974. A film detection method for tritium-labeled proteins and nucleic acids in polyacrylamide gels. Eur. J. Biochem. **46**:83–88.
2. **Braun, D. K., B. Roizman, and L. Pereira.** 1984. Characterization of post-translational products of herpes simplex virus gene 35 proteins binding to the surfaces of full capsids but not empty capsids. J. Virol. **49**:142–153.
3. **Chamberlain, J. P.** 1979. Fluorographic detection of radioactivity in polyacrylamide gels with the water-soluble fluor, sodium salicylate. Anal. Biochem. **98**:132–135.
4. **Chirgwin, J. M., A. E. Przybyla, R. J. MacDonald, and W. J. Rutter.** 1979. Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry **18**:5294–5299.
5. **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. **232**:1102–1106.
6. **Cohen, G. H., M. Ponce de Leon, H. Diggelmann, W. C. Lawrence, S. K. Vernon, and R. J. Eisenberg.** 1980. Structural analysis of the capsid polypeptides of herpes simplex virus types 1 and 2. J. Virol. **34**:521–531.
7. **Fairbanks, G., T. L. Steck, and D. F. H. Wallach.** 1971. Electrophoretic analysis of the major polypeptides of the human erythrocyte membrane. Biochemistry **10**:2606–2617.
8. **Gibson, W.** 1981. Protease-facilitated transfer of high-molecular-weight proteins during electrotransfer to nitrocellulose. Anal. Biochem. **118**:1–3.
9. **Gibson, W.** 1981. Structural and nonstructural proteins of strain Colburn cytomegalovirus. Virology **111**:516–537.
10. **Gibson, W.** 1983. Protein counterparts of human and simian cytomegaloviruses. Virology **128**:391–406.
11. **Gibson, W., and D. M. Benko.** 1988. Protease-facilitated protein transfer, p. 93–99. *In* O. J. Bjerrum and N. H. H. Heegaard (ed.), CRC handbook of immunoblotting of proteins. CRC Press, Inc., Boca Raton, Fla.
12. **Gibson, W., and B. Roizman.** 1972. Proteins specified by herpes simplex virus. VIII. Characterization and composition of multiple capsid forms of subtypes 1 and 2. J. Virol. **10**:1044–1052.
13. **Gibson, W., and B. Roizman.** 1974. Proteins specified by herpes simplex virus. X. Staining and radiolabeling properties of B-capsid and virion proteins in polyacrylamide gels. J. Virol. **13**:155–165.
14. **Glisin, V., R. Crkvenjakov, and C. Byus.** 1974. Ribonucleic acid isolated by cesium chloride centrifugation. Biochemistry **13**:2633–2637.
15. **Heilman, C. J., Jr., M. Zweig, J. R. Stevenson, and B. Hamper.** 1979. Isolation of a nucleocapsid polypeptide of herpes simplex virus types 1 and 2 possessing immunologically type-specific and cross-reactive determinants. J. Virol. **29**:34–42.
16. **Irmiere, A., and W. Gibson.** 1983. Isolation and characterization of a noninfectious virion-like particle released from cells infected with human strains of cytomegalovirus. Virology **130**:118–133.
17. **Irmiere, A., and W. Gibson.** 1985. Isolation of human cytomegalovirus intranuclear capsids, characterization of their protein constituents, and demonstration that the B-capsid assembly protein is also abundant in noninfectious enveloped particles. J. Virol. **56**:277–283.
18. **King, J., and S. Casjens.** 1974. Catalytic head assembling protein in virus morphogenesis. Nature (London) **251**:112–119.
19. **Ladin, B. F., S. Ihara, H. Hampl, and T. Ben-Porat.** 1982. Pathway of assembly of herpesvirus capsids: an analysis using DNA+ temperature-sensitive mutants of pseudorabies virus. Virology **116**:544–561.
20. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
21. **Laskey, R. A., and A. D. Mills.** 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. Eur. J. Biochem. **56**:335–341.
22. **Lee, J. Y., A. Irmiere, and W. Gibson.** 1988. Primate cytomegalovirus assembly: evidence that DNA packaging occurs subsequent to B-capsid assembly. Virology **167**:87–96.
23. **O'Callaghan, D. J., and C. C. Randall.** 1976. Molecular anatomy of herpesviruses; recent studies. Prog. Med. Virol. **22**:152–210.
24. **Preston, V. G., J. A. V. Coates, and F. J. Rixon.** 1983. Identification and characterization of a herpes simplex virus gene product required for encapsidation of virus DNA. J. Virol. **45**:1056–1064.
25. **Robson, L., and W. Gibson.** 1989. Primate cytomegalovirus assembly protein: genome location and nucleotide sequence. J. Virol. **63**:669–676.
26. **Sherman, G., and S. L. Bachenheimer.** 1988. Characterization of intranuclear capsids made by ts morphogenic mutants of HSV-1. Virology **163**:471–480.
27. **Towbin, H., T. Staehelin, and J. Gordon.** 1979. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA **76**:4350–4354.