Proc. Nat. Acad. Sci. USA
Vol. 68, No. 12, pp. 3083–3087, December 1971

Virus-Specific Proteins Synthesized in Encephalomyocarditis Virus-Infected HeLa Cells

(gel electrophoresis/tryptic mapping/cyanogen bromide mapping)

BYRON E. BUTTERWORTH, LINDA HALL*, C. MARTIN STOLTZFUS†, AND ROLAND R. RUECKERT

Biophysics Laboratory and Department of Biochemistry, University of Wisconsin, Madison, Wis. 53706

Communicated by Henry Lardy, September 29, 1971

ABSTRACT **The *in vivo* synthesis of encephalomyocarditis-specific proteins was studied by labeling the viral proteins with radioactive amino acids under conditions where host-protein synthesis was almost completely inhibited. To assure recovery of all proteins, intact cells were lysed in hot 1% sodium dodecyl sulfate. These lysates were analyzed by quantitative high-resolution electrophoresis on sodium dodecyl sulfate–polyacrylamide gels. This technique allowed the detection and estimation of the molecular weight of 15 virus-specific polypeptides: *A*, 100,000; *B*, 90,000; *C*, 84,000; *D*, 75,000, *D1*, 65,000; *E*, 56,000; ϵ, 40,000; *F*, 38,000; α, 34,000; β, 30,000; γ, 23,000; *G*, 16,000; *H*, 12,000; *I*, 11,000; and δ, 9,000. Pulse–chase experiments, in conjunction with cyanogen bromide and tryptic mapping of the isolated polypeptides, indicate that at least three primary gene products (A,F,C), with a cumulative weight of about 220,000, are generated during translation of the RNA genome. Chains *A* and *C* then undergo post-translational cleavages, while *F* remains uncleaved. The proteins generated by the cleavage of *A* include all of the capsid chains (α, β, γ, δ, ϵ). Those generated by the cleavage of *C* include *D* and *E*. The chains α, β, γ, δ, *E*, *F*, *G*, *H*, *I*, with a cumulative molecular weight of about 230,000, are stable and are produced in about equimolar amounts. A model for the synthesis of, and a cleavage sequence that accounts for, all of the viral polypeptides is proposed.**

As the smallest RNA-containing animal viruses, the picornaviruses (1) constitute good models with which to explore problems related to macromolecular synthesis and function in animal cells. Evidence indicates that the virions of all four subgroups (2), the *enteroviruses* (e.g., polio-, coxsackie-, echoviruses), the *cardioviruses* [e.g., mouse-Elberfeld (ME), encephalomyocarditis (EMC), mengoviruses], the *rhinoviruses* (e.g., common-cold viruses) and the *foot-and-mouth disease viruses* are similar in size and structure (3).

Poliovirus-infected cells contain at least 14 electrophoretically distinct polypeptides; many of these polypeptides arise from the cleavage of large precursors into smaller chains (4–7). Six of these chains have been identified as capsid-related: VP1, VP2, VP3, and VP4, from mature virions; VP0, an uncleaved precursor of VP2 and VP4; and NCVP1, a very large precursor of all four capsid proteins (8).

Much is known about the size and structure of the cardioviruses (9). Here, we present the results of a thorough search for the proteins synthesized by EMC virus. A determination of the kinetics of synthesis and cleavage of the polypeptides found, in addition to their characterization by tryptic- and CNBr-mapping allowed the formulation of a model for the biosynthesis of EMC-viral proteins.

Abbreviations: EMC virus, encephalomyocarditis virus; ME virus, mouse-Elberfeld virus; SDS, sodium dodecyl sulfate.

* Present address: Department of Zoology, University of British Columbia, Vancouver 8, Canada.

† Present address: Roche Institute of Molecular Biology, Nutley, N.J. 07110.

MATERIALS AND METHODS

Cells, Virus, and Plaque Assay. HeLa cells (strain W), EMC virus and plaque assay were described (10). The cells were propagated in suspension in Medium F containing 6% bovine serum.

Media and Buffers. Medium F is Eagle's minimal medium (11) in Earle's saline (12), lacking calcium and magnesium salts, and supplemented with 0.1 mM (each) of glycine and serine and 0.1% Pluronic F68. Medium AH is medium A (10) modified by the addition of *N*-2-hydroxyethylpiperazine-*N'*-2-ethanesulfonic acid (HEPES) buffer, to a final concentration of 25 mM. Medium AL is medium AH lacking amino acids. Buffer TSA is 0.25 M sodium chloride–5 mM Tris (pH 7.4)–0.01% bovine serum albumin. PBSA was described (10).

Infection. HeLa cells were harvested when the density reached 3–4 $\times$ 10^5 cells/ml; 8 $\times$ 10^7 cells were suspended in 2 ml of medium AL at room temperature. This suspension was then infected with 200 plaque-forming units (PFU) per cell of EMC virus; attachment was allowed to proceed for 15 min after infection at room temperature. 18 ml of warm medium AL containing 5.5 μg/ml of actinomycin D was then added to the suspension. With the cells at a concentration of 4 $\times$ 10^6 cells/ml, the infection was allowed to proceed at 37°C in a constant-temperature, shaken water bath.

Preparation of Cell Lysates. After two washes with ice-cold Earle's saline, the cells were suspended to about 2 $\times$ 10^7 cells/ml in distilled water; an equal volume of solubilizing solution (1 M urea–2% sodium dodecyl sulfate (SDS)–0.2% 2-mercaptoethanol) was added. The resulting viscous solution was heated for 5 min in a boiling-water bath to dissociate the proteins into their constituent polypeptide chains. The lysate was dialyzed overnight at 8°C against a large volume of dialysis buffer [0.01 M sodium phosphate (pH 7.2)–0.5 M urea–0.1% SDS–0.1% 2-mercaptoethanol–1 mM sodium azide] (4), and stored frozen at −70°C.

Electrophoresis. Gels were prepared and electrophoresed as described (13). The gel columns were 0.6 $\times$ 20–29 cm and

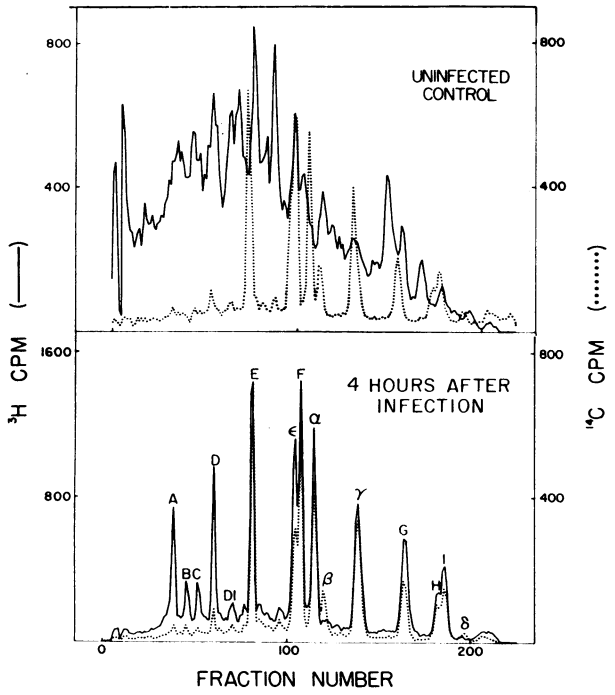

FIG. 1. Electrophoretic profiles of proteins synthesized in HeLa cells after infection with EMC virus (——). At 1 hr intervals after infection (1,2, and 3 hr after infection profiles not shown) 3-ml samples were withdrawn from the suspension of infected cells and diluted into 0.9 ml of medium AL containing 91 μCi/ml of the [^{3}H]aminoacid mixture. Labeling was allowed to proceed for 0.5 hr at 37°C, with occasional shaking. The incorporation was stopped by chilling in an ice bath. Whole-cell extracts were then made and analyzed by gel electrophoresis (as described in *Methods*). Each sample was coelectrophoresed with a ^{14}C-labeled reference lysate (· · ·) containing "stable" EMC-specific proteins (see *text*). This lysate was prepared from an infected suspension after it was labeled with 5 μCi/ml of [^{14}C]aminoacid mixture from 3.5–5 hr after infection. The absence of appreciable label at the top of the gel is good evidence that few of the protein chains are aggregated. The anode in this and subsequent figures is to the *right*.

contained 10% polyacrylamide–0.3% (v/v) ethylenediacrylate–0.5 M urea–0.1% SDS–0.1 M sodium phosphate buffer (pH 7.2)–0.1 M sodium 3-mercaptopropionate. Samples to be analyzed by electrophoresis were made 1% in SDS and 2-mercaptoethanol, 20% in sucrose, and 0.01% in bromphenol blue. From 20 to 200 μl was then layered on the gel under the electrophoresis buffer. The high viscosity of the lysate, associated with release of cellular DNA, did not hinder electrophoretic analysis; up to 10^6 solubilized cells could be applied without overloading the gels.

Gel Fractionation and Measurement of Radioactivity. Fractionation of gels has been described (13). 1-mm segments of crushed gel in 0.3 ml of water were counted either directly in 10 ml of solvent B-10, or were hydrolyzed and counted in solvent tT21 (13). Because of fluorescence problems encountered with solvent tT21 at ambient temperatures, the B-10 systems was adopted. Although the recovery of radioactivity from the gels with tT21 was about 95%, while the simpler B-10 solvent gave only 75% recovery, a comparison of duplicate gels by these two methods gave comparable patterns with all peaks present in the same relative proportions. The samples were counted on a Beckman LS-233 liquid scintillation counter equipped with a punched paper tape output. The raw data from the counter was analyzed with a Univac 1108 computer. A program was written to correct ^{3}H and ^{14}C counts for background and spillover, calculate the running area under the peaks, and compute the molecular weights of the proteins. The electropherogram was then generated with the system's CalComp Plotter.

Isotopes. The reconstituted protein hydrolysate containing 16 [^{3}H]aminoacids was obtained from Schwarz/Mann. The L-[^{14}C]amino-acid mixture containing 15 amino acids was purchased from New England Nuclear.

RESULTS

Unmasking Viral Protein Synthesis. Identification of all virus-specific polypeptides and the determination of their kinetics of synthesis requires a representative sample of all viral proteins in the cell. Analysis of cytoplasmic extracts prepared from cells homogenized after swelling in hypotonic

TABLE 1. *Size of viral polypeptides and molar ratios of end-product chains*

Peak	Apparent molecular weight*	Kinetic behavior†	% of Total viral cpm‡	Molar ratio§
A	100,000	u	0.5 (0.3)	
B	90,000	u	0.5 (0.2)	
C	84,000	u	0.5 (0.3)	
D	75,000	u	1.3 (0.4)	
D1	65,000	u	0.7 (0.3)	
E	56,000	s	18.3 (1.0)	0.72
ε	40,000	u	9.0 (0.6)	0.50¶
F	38,000	s	17.2 (1.6)	1.00
α	34,000	s	15.4 (0.3)	1.00
β	30,000	s	4.3 (0.5)	0.32¶
γ	23,000	s	13.2 (0.4)	1.27
G	16,000	s	7.7 (0.4)	1.06
H	12,000	s	3.8 (0.3)	0.70
I	11,000	s	6.6 (0.4)	1.32
δ	9,000	s	1.1 (0.2)	0.27¶

* The apparent molecular weight of each component was calculated from its mobility on SDS–gels relative to that of six standards: phosphorylase a (94,000), bovine serum albumin (68,000), ovalbumin (43,000), carboxypeptidase A (34,600), lysozyme (14,300), and cytochrome *c* (11,700). The markers were labeled by methylation (20) with $(^{14}CH_3)_2SO_4$. A straight line was fitted by regression analysis to a plot of log of molecular weight against mobility (21).

† s = stable (no loss of counts during a chase); u = unstable.

‡ Calculated from the average of six independent gels, run on three different cell extracts, in which the cleavages had gone to completion. The profiles were similar to that in Fig. 2*B*. The distribution of ^{14}C was calculated relative to the total radioactivity recovered from all of the viral peaks. The average deviations are indicated in parentheses.

§ The mass of protein was assumed to be proportional to cpm. The molar ratios were calculated by dividing the mass ratio (column 4) of each chain by its apparent molecular weight, and normalizing with respect to the alpha chain.

¶ Were the precursor chain ε completely cleaved, the molar ratios of its products, β and δ, would be 0.5 + 0.32 = 0.82 and 0.5 + 0.27 = 0.77, respectively.

buffer proved unsatisfactory in our hands. Not only were most of the labeled proteins found in the nuclear pellet, but the electrophoretic profiles of the nuclear and cytoplasmic fractions were quantitatively different. To circumvent this difficulty, the intact cells were solubilized in hot 1% SDS.

EMC virus plus actinomycin D rapidly inhibited host protein synthesis and allowed the preferential labeling of the viral proteins (14). The kinetics of this inhibition was shown by examination of the products of protein synthesis throughout the infection (Fig. 1); there was a transition from a complex pattern characteristic of the host cell (uninfected control) to a simpler pattern characteristic of virus infection (4 hr). 4 hr after infection corresponds to about the mid-log phase of the viral growth (infectivity) curve.

Classification of Virus-Specific Proteins. Of the 15 peaks observed 4 hr after infection (Fig. 1), five ($\alpha,\beta,\gamma,\delta,\epsilon$) were identified as capsid polypeptides by coelectrophoresis with homologous chains from purified ^{14}C-labeled virus (not shown). These and the remaining peaks, designated *A* through *I*, were further characterized with respect to size and kinetic behavior (Table 1).

In pulse–chase experiments, most of the label was found in large polypeptides within 10 min of synthesis (Fig. 2*A*). These precursor polypeptides were then cleaved to generate

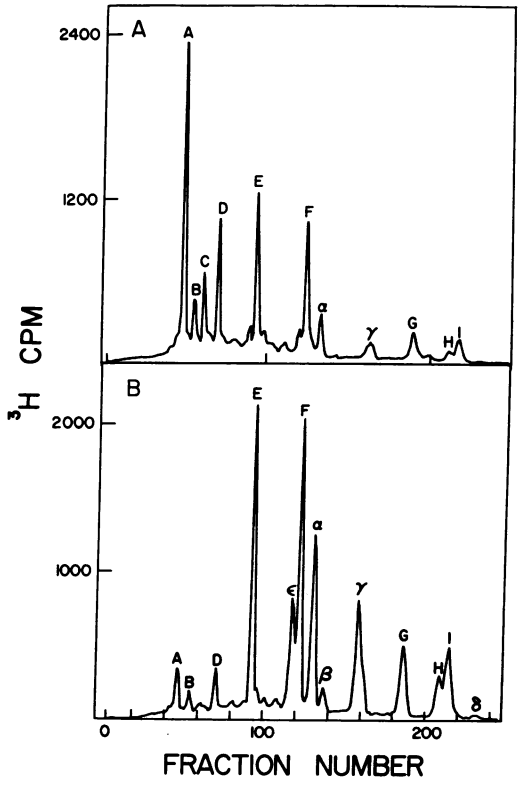

FIG. 2. Flow of radioactivity in a pulse–chase experiment. At 3.7 hr after infection, 8×10^7 cells were sedimented and resuspended in 3 ml of warm medium AL. To initiate labeling, 200 μCi of [^{3}H]aminoacid mixture in 2 ml of warm medium AL was added. (*A*) 10 min after the addition of label, a 0.6-ml sample was removed and lysed in solubilizing solution. (*B*) The remaining suspension was incubated 15 min longer. The cells were then sedimented and resuspended in warm medium AH containing no radioactive amino acids and incubated for an additional hour (chased) at 37°C before lysing.

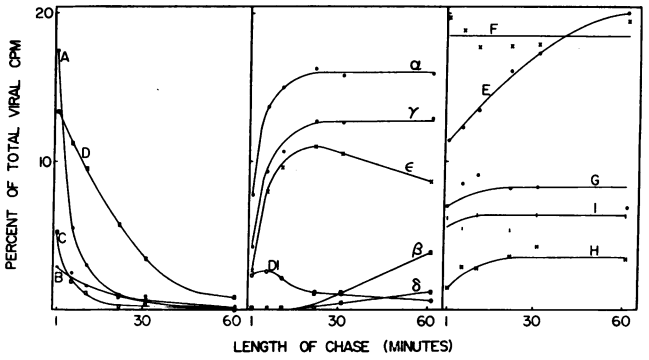

FIG. 3. Pulse–chase kinetics of EMC-virus proteins. 4 hr after infection, cells were pulsed for 10 min with a [^{3}H]amino acid mixture and chased as described in Fig. 2. Aliquots were removed at the indicated intervals, chilled, washed, lysed, dialyzed, and electrophoresed. The distribution of label was calculated as percent cpm recovered in each peak, relative to the total cpm recovered in all peaks.

the smaller stable chains. This cleavage is evidenced by the flow of label from the large to the small chains during a 1-hr chase with unlabeled amino acids (Fig. 2*B*).

More detailed kinetic analysis of such a pulse–chase experiment, in which the distribution of radioactivity was measured at intervals during the chase period (Fig. 3), illustrates the transient (precursor) nature of *A*, *B*, *C*, *D*, *D1*, and ϵ, and the flow of label into components α, β, γ, δ, *E*, *G*, *H*, and *I*. The rapid rate of disappearance of label from the *A* chain resembles the appearance rate of α, γ, and ϵ. The correspondingly slower decay-rate of *D* and slow appearance of *E* suggests that *E* is a product of *D*. In this, and other, ex-

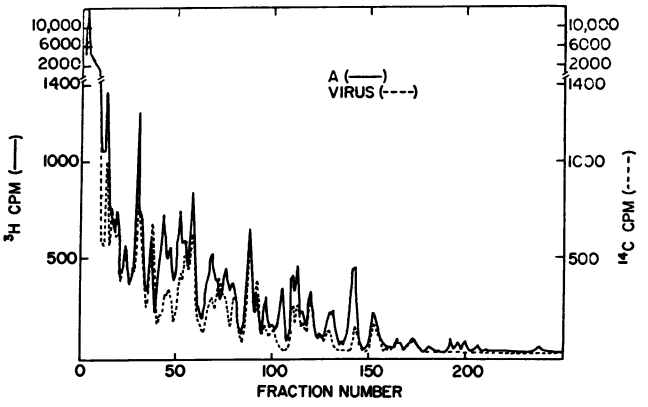

FIG. 4. Chromatographic comparison of tryptic peptides from "*A*" protein (——) with those from purified EMC virions (– – –). About 1.2×10^5 cpm of isolated, lyophilized *A* protein labeled with [^{3}H]aminoacids, mixed with 0.15 ml of EMC virions labeled with [^{14}C]aminoacids (1.0×10^5 cpm) in TSA, was heated for 1 min at 100°C. with 0.35 ml H_2O and 12 mg of *N,N'*-diacetylcystamine. After adjustment to about pH 8 with 8 mg of NH_4HCO_3, the protein was digested with TPCK-treated trypsin (treated with TPCK to remove chymotryptic activity, Worthington) (1 mg/ml) for 5 hr at 37°C. The lyophilized digest was dissolved in 1 ml of 0.2 M pyridinium acetate buffer (pH 3.1), and chromatographed on a Dowex 50 x-8 column (22). Samples of about 2 ml (each) were evaporated to dryness, dissolved in 0.2 ml of water, and counted in 10 ml of scintillation solvent B-10 (13).

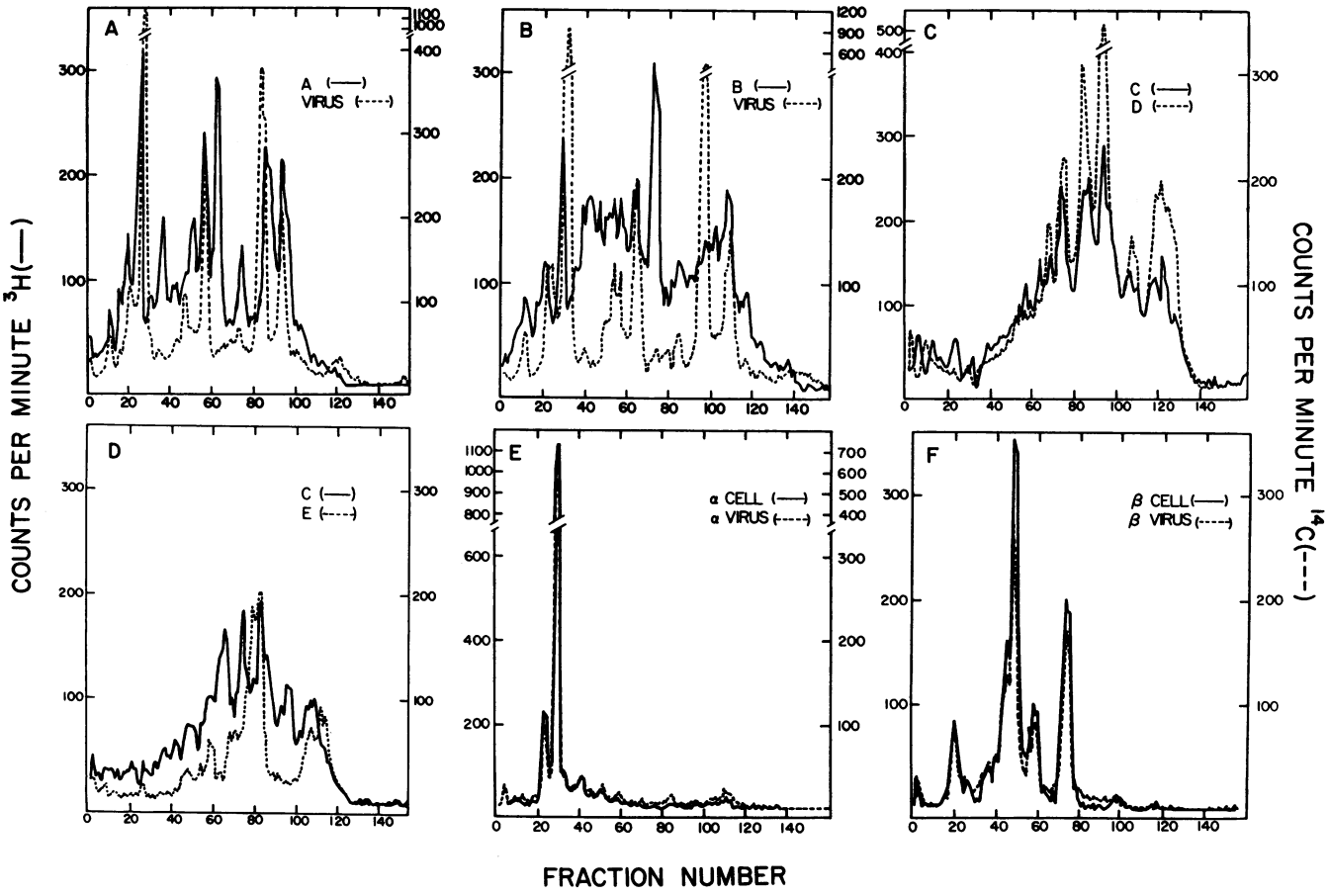

FIG. 5. Comparison CNBr fragmentation profiles of EMC-virus-specific polypeptides from cell lysates labeled with [^{3}H]amino-acids (——) and from purified virions labeled with [^{14}C] aminoacid (---). Individual chains were prepared from a cell extract labeled at 4 hr as described in Fig. 3, and each chain was isolated from polyacrylamide gels. These were compared to either whole virus or the homologous chain from EMC virions by coelectrophoresis of their CNBr digests. (*A*), chain *A* against EMC-virions; (*B*), Chain B against EMC virions; (*C*), chain *C* against chain *D* (in this case, each ^{3}H-labeled chain was electrophoresed separately in parallel gels, and the profiles were superimposed for comparison); (*D*), chain *C* against chain *E*; (*E*), α chain; (*F*), β chain.

Each mixture (about 10,000–15,000 cpm of each label) was digested 9 hr at room temperature with 2% CNBr in 70% formic acid, then lyophilized and subjected to electrophoresis on 0.6 × 15 cm SDS–polyacrylamide gels (23) for 13 hr at 4 mA per gel.

periments *F* was the only chain that consistently did not change in amount during the chase period and is, therefore, assumed to be a stable primary product.

The amount of ^{3}H in the ε peak rose to a maximum within 15–20 min, then fell at the same rate as the β and δ chains were generated. This behavior is consistent with earlier suggestions (7, 15) that the β and δ chains are generated by cleavage of the ε chain.

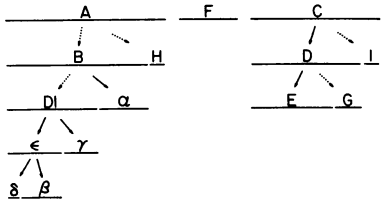

FIG. 6. Model for the biosynthesis of EMC-virus proteins. Three primary gene products A, F, and C are generated during translation of the RNA genome. Chain *F* is not further cleaved; the capsid precursor, *A*, is cleaved four times and precursor chain *C* is cleaved twice. *Dotted arrows* indicate tentative assignments. The order of the capsid polypeptides within *A* appears to be δ-β-γ-α, starting with δ at the amino-terminal end of the sequence. The order of synthesis of the primary products appears to be A-F-C.

To further relate these chains, each was isolated by extraction from polyacrylamide gels, and was analyzed by comparison of tryptic or CNBr digests with similar digests of capsid protein (16). Tryptic digests of protein *A* and EMC-capsid protein revealed many overlaps (Fig 4), thereby supporting the identification of *A* as a capsid precursor.

Comparative analysis of CNBr digests from each peak gave the following results. The electrophoretic profiles of digests from *A*, *B*, and the capsid chains were similar to each other (Fig. 5*A* and *B*). The profiles from *C*, *D*, and *E* were related, as a family distinct from the capsid proteins (Fig. 5*C* and *D*). The profile from protein *F* (not shown) showed little correlation with the profiles from *A* or *C*.

Comparative studies on the CNBr-fragmentation patterns of the ε chain revealed similarities to patterns from the β and δ chains (not shown), thus further supporting the hypothesis (7, 15) that the β and δ polypeptides are produced by the cleavage of the ε chain.

The patterns of digests from the α, β, and γ chains isolated from cell lysates coincided with those of the homologous chains

isolated from purified virions (Fig. 5*E* and *F*; γ not shown). The correspondence of these profiles from chains isolated from the two sources not only confirms their identities, but also demonstrates that the α, β, and γ peaks from the extract are relatively free of unrelated, comigrating polypeptides.

Analysis of the smaller chains *G*, *H*, and *I* showed that all have distinctive fragmentation profiles (not shown). However, the limited number of fragments produced and the apparent inability of CNBr to cleave quantitatively (16) precluded further sequence-relatedness assignments.

Discussion

The similarities of the CNBr and tryptic digests of several of the proteins allows them to be placed into the related families: (*i*) *A*, *B*, and the capsid polypeptides; (*ii*) *C*, *D*, and *E*. The kinetic experiments distinguish precursor from product chains and their results are consistent with the following classifications: capsid precursor chains (A,B,ϵ); stable capsid chains ($\alpha,\beta,\gamma,\delta$); stable primary product (F); noncapsid precursor chains (C,D); stable noncapsid chains (E,G,H,I).

Integrating these results, and assuming that the largest precursor polypeptide in each group is the primary product of protein synthesis, we propose that the EMC-virus genome is initially translated to generate three gene products (*A*,*F*, and *C*), two of which (*A*,*C*) are further cleaved into smaller chains (Fig. 6). The assignments of *D1*, *G*, *H*, and *I* are based primarily on attempts to balance differences in molecular weights.

Progressive labeling experiments, which determine the initial time and rate of labeling of the polypeptides, and studies with the protein synthesis inhibitor pactamycin, suggest that the primary products are synthesized in the order 5′ → 3′ A,F,C and that the capsid polypeptides are synthesized in the order δ, β, γ, α (Butterworth and Rueckert, in preparation). An analogous sequence has been independently proposed for the synthesis of polioviral proteins (17). Summers and Maizel also find the polioviral capsid protein to be nearest the 5′ end of the viral RNA (personal communication). Preliminary studies on the order of CNBr fragments in the parent chains of the closely related ME virus confirm that the δ-chain is derived from the amino-terminal region of the ϵ-chain (16).

If a value of 2.6×10^6 is assumed for the molecular weight of EMC-virus RNA (3, 18), the virus has a theoretical coding capacity of about 270,000. The sum of the molecular weights both of the proposed primary products (A,F,C), with a cumulative weight of 222,000, and of the stable end-products ($\alpha,\beta,\gamma,\delta,E,F,G,H,I$), with a cumulative weight of 229,000, are slightly less than the theoretical value. It remains to be determined whether this small difference is due to incorrect values for the molecular weight of the viral RNA, for the size or number of polypeptides produced, or if a part of the RNA genome has noncoding functions.

It has been suggested that the picornaviral messenger RNA possesses a single initiation site and is, therefore, completely translated by every ribosome (5, 19). The consequence of this would be the equimolar production of all stable viral polypeptides (if the complete cleavage of intermediates is assumed). If this theory is correct, the finding of about equimolar ratios of the stable EMC-virus proteins (Table 1) argues that all of the end products have been resolved and identified.

We thank Prof. G. Craven and L. Shih for assistance with tryptic mapping, and G. Radloff and P. VandeVoorde for excellent technical assistance. B. E. B. and L. H. were supported by NIH predoctoral training grant 5-T01-GM-00236. C. M. S. was supported by NSF predoctoral training grant 1-F1-GM-40, 785-01. R. R. R. is a United States Public Health Service Career Development Awardee. This work was supported by Grants VC-26A from the American Cancer Society and CA-08662 from the United States Public Health Service.

1. "International Enterovirus Study Group," *Virology*, **19**, 114 (1963).
2. Fenner, F., in *The Biology of Animal Viruses* (Academic Press, New York, N.Y., 1968), pp. 20–21.
3. Rueckert, R. R. in *Comparative Virology*, ed. K. Maramorousch and E. Kurstak (Academic Press, New York, N.Y., 1971), pp. 255–306.
4. Summers, D. F., J. V. Maizel, Jr., and J. E. Darnell, Jr., *Proc. Nat. Acad. Sci. USA*, **54**, 505 (1965); Summers, D. F., and J. V. Maizel, Jr., *Proc. Nat. Acad. Sci. USA*, **59**, 966 (1968).
5. Jacobson, M. F., and D. Baltimore, *Proc. Nat. Acad. Sci. USA*, **61**, 77 (1968).
6. Holland, J. J., and E. D. Kiehn, *Proc. Nat. Acad. Sci. USA*, **60**, 1015 (1968).
7. Jacobson, M. F., and D. Baltimore, *J. Mol. Biol.*, **33**, 369 (1968).
8. Jacobson, M. F., J. Asso, and D. Baltimore, *J. Mol. Biol.*, **49**, 657 (1970); Cooper, P. C., D. F. Summers, and J. V. Maizel, *Virology*, **41**, 408 (1970).
9. Scraba, D. G., C. M. Kay, and J. S. Colter, *J. Mol. Biol.*, **26**, 67 (1967); Dunker, A. K., and R. R. Rueckert, *J. Mol. Biol.*, **58**, 217 (1971).
10. Hall, L., and R. R. Rueckert, *Virology*, **43**, 152 (1971).
11. Eagle, H., *Science*, **130**, 432 (1959).
12. Earle, W. R., *J. Nat. Cancer Inst.*, **4**, 165 (1943).
13. Medappa, K. C., C. McLean, and R. R. Rueckert, *Virology*, **44**, 259 (1971).
14. Hall, L., Ph.D. thesis, University of Wisconsin, Madison, 1970.
15. Rueckert, R. R., A. K. Dunker, and C. M. Stoltzfus, *Proc. Nat. Acad. Sci. USA*, **62**, 912 (1969).
16. Stoltzfus, C. M., Ph.D. thesis, University of Wisconsin, Madison, 1971.
17. Taber, R., D. Rekosh, and D. Baltimore, *J. Virol.*, **8**, 395 (1971).
18. Burness, A. T. H., *J. Gen. Virol.*, **6**, 373 (1970).
19. Roumiantzeff, M., D. F. Summers, and J. V. Maizel, Jr., *Virology*, **44**, 249 (1971).
20. Kiehn, E. D., and J. J. Holland, *J. Virol.*, **5**, 358 (1970).
21. Shapiro, A. L., E. Vinuela, and J. V. Maizel, Jr., *Biochem. Biophys. Res. Commun.*, **28**, 815 (1967); Weber, J., and M. Osborn, *J. Biol. Chem.*, **244**, 4406 (1969); Dunker, A. K., and R. R. Rueckert, *J. Biol. Chem.*, **244**, 5074 (1969).
22. Craven, G. R., P. Voynow, J. S. Hardy, and C. G. Kurland, *Biochemistry*, **8**, 2906 (1969).
23. Swank, R. T., and K. D. Munkres, *Anal. Biochem.*, **39**, 462 (1971).