LEADER POLYPEPTIDES ENCODED IN THE 5′-REGION OF THE ENCEPHALOMYOCARDITIS VIRUS GENOME

Yu. A. KAZACHKOV, T. V. CHERNOVSKAYA, E. Yu. SIYANOVA, Yu. V. SVITKIN, T. Yu. UGAROVA and V. I. AGOL

Institute of Poliomyelitis and Viral Encephalitides, USSR Academy of Medical Sciences, Moscow Region 142782 and *Moscow State University, Moscow 117234, USSR*

Received 26 February 1982

1. Introduction

The synthesis of encephalomyocarditis (EMC) virus proteins, like proteins of other picornaviruses, is initiated at a site located in the 5′-adjacent region of the viral genome; the product of this synthesis is represented by a giant polypeptide chain, polyprotein, which contains all the amino acid sequences present in the structural and non-structural viral proteins; the final products of translation, 'mature' proteins, are generated by sequential proteolytic cleavages of the polyprotein molecule starting while the latter is in the nascent form [1]. Using a variety of techniques, cleavage maps of the EMC virus polyprotein were constructed, which showed the order of different polypeptide chains relative to the N-terminus of the polyprotein molecule [1]. Most of the details of these maps are now generally accepted, but some points remain unclear or disputable. One of these important points concerns the uppermost N-terminal region of the polyprotein: which particular polypeptide is the first to be translated? The usually considered candidate is the precursor of capsid proteins, polypeptide A, which is assumed to share the N-terminal amino acid sequence with its cleavage products, polypeptides B and D1. This view was supported by the finding that, upon in vitro translation of EMC virus RNA, formyl-[^{35}S]Met-tRNA$_f^{Met}$ could donate a labelled formylmethionyl residue to the polypeptides (by inference, to their N-terminal position) that comigrated with polypeptides A, B and possibly, D1 [2]. On the other hand, earlier data obtained with the same labelled substrate but using less efficient cell-free translation systems suggested that formyl-methionine might be incorporated in peptides other than those corresponding to the capsid proteins [3–5].

We have attempted to solve the controversy and show here that:

(i) The initiator amino acid sequences, as defined by the presence of formyl-methionyl residue, are contained in a precursor polypeptide, preA;
(ii) Upon a short treatment of preA with a preparation of EMC virus-specific protease [6], these initiator sequences are cleaved off and emerge in two low-M_r closely-related polypeptides, p14 and p12;
(iii) After the removal of p14 (p12), the remaining moiety of preA corresponds to the precursor of capsid proteins, polypeptide A;
(iv) preA, p14 and p12 appear to share an N-terminal amino acid sequence.

These data suggest that all 3 latter polypeptides are initiated at a common single site on the viral genome and that p14 (p12) should occupy the N-terminal position on the EMC virus polyprotein cleavage map followed by polypeptide A, a precursor of the capsid proteins.

2. Methods

2.1. *Labelling viral polypeptides in vivo and in vitro*

Suspensions of ascites carcinoma Krebs-II cells were prepared and infected with EMC virus as in [7]. [^{35}S]Methionine (50 μCi/ml; Amersham) was added at 3.5 h postinfection, and after a 60 min additional incubation the cells were lysed in a buffer containing SDS and urea [8].

Address reprint requests to V. I. A.

Published by Elsevier Biomedical Press
00145793/82/0000–0000/$02.75

Translation of EMC virus RNA in a micrococcal nuclease-treated extracts of Krebs-II cells was done as in [8,9]. Preparation of formyl-[^{35}S]Met-tRNA$_f^{Met}$ was described in [10].

2.2. *Electrophoresis in polyacrylamide gels*

Electrophoretic separation of labelled polypeptides was performed using SDS-containing gradient (8–20%) polyacrylamide slabs and a discontinuous buffer system [11] followed by autoradiography.

2.3. *Isolation and fingerprinting of tryptic [^{35}S]methionine-containing peptides*

Regions containing labelled viral polypeptides were excised from polyacrylamide slabs, the excised gels were washed with methanol and H_2O, digested with TPCK-treated trypsin (300 μg/ml; 40 U/mg; Serva) for 16 h at 37°C, and peptides were eluted from the gels as in [12]. The eluates were lyophilized twice, dissolved in a mixture containing 15% acetic acid and 5% formic acid, applied onto 10 × 10 cm thin-layer cellulose plates (Merck) and electrophoresed for 30 min at 500 V. The second-dimension ascending chromatography was carried out in a buffer containing butanol:pyridine:acetic acid:H_2O (32.5:25:5:20, by vol.) with 7% (w/v) 2,5-diphenyloxazole [13]. The dried plates were fluorographed at −70°C.

2.4. *Isolation and analysis of tryptic formyl-[^{35}S]-methionine-containing peptides*

Products of in vitro translation of EMC virus RNA in the presence of formyl-[^{35}S]Met-tRNA$_f^{Met}$ (either unfractionated or eluted from the gel as individual polypeptides, preA, p14 and p12) were oxidized with performic acid, digested with TPCK-treated trypsin, and analyzed by high-voltage paper electrophoresis at pH 10.2 [14].

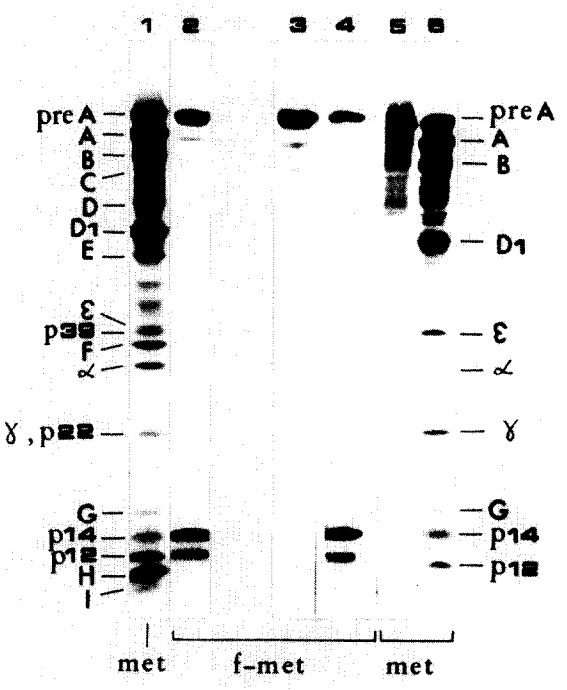

Fig.1. Electrophoretic patterns of the products of in vitro translation of EMC virus RNA. [^{35}S]Methionine (lanes 1,5,6) or formyl-[^{35}S]Met-tRNA$_f^{Met}$ (lanes 2–4) were used as labelled substrates. Samples 1,2 were incubated for 150 min at 30°C. Samples 3–6 were incubated for 30 min at 30°C, then the synthesis was stopped by the addition of cycloheximide to 250 μg/ml final conc.; simultaneously a cytoplasmic extract (S_{30} fraction) from EMC virus-infected cells was added (0.1 vol. of the sample) to samples 4,6. After the addition of cycloheximide, samples 3–6 were incubated for 5 min at 37°C.

3. Results and discussion

Prolonged translation of EMC virus RNA in the Krebs-II-derived cell-free system using [^{35}S]methionine as labelled substrate results in accumulation of a nearly full complement of virus-specific polypeptides found in the virus-infected cells as well as of some additional species (fig.1, lane 1; cf. [8,9]). On the other hand, if formyl-[^{35}S]Met-tRNA$_f^{Met}$ is used as the labelled substrate, only 3 major radioactive species can be revealed (fig.1, lane 2); these are polypeptides preA, p14, and p12 (of ~112 000, 14 000 and 12 000 M_r, respectively). Neither of the 3 is present in any significant amount in the virus-infected cells (cf. [15]). Trace formyl-methionine radioactivity may also be incorporated in vitro into polypeptides comigrating with proteins A and (sometimes) B. Thus, the labelling of high-M_r polypeptides with formyl-methionine in our system is virtually similar to that observed in a reticulocyte-derived one [2], but in the latter no incorporation of this residue in polypeptides analogous to p14 and p12 was observed.

The presence of formyl-methionine in the 3 in vitro products might be explained by their independent initiation. Alternatively, the smaller species might represent cleavage products of the larger molecule. The following experiment proved that the latter possibility was correct. Cell-free samples programmed with EMC

virus RNA were incubated with either [^{35}S]methionine or formyl-[^{35}S]Met-tRNA$_f^{Met}$ for a relatively short time interval to allow significant synthesis of only precursors of capsid proteins (fig.1, lanes 3,5). Virus-specific protease, which is encoded farther toward the 3′-end of the genome [9,16], could not be formed in appreciable amounts during this period, and this was shown by the absence of any methionine-containing mature capsid proteins. Significantly, p14 and p12 were also absent, regardless of the substrate used. Further translation was stopped by cycloheximide, and an extract from the virus-infected cells was added as a source of viral protease. After an incubation as short as 5 min, prominent formyl-methionine- and methionine-labelled bands in the p14 and p12 regions appeared (fig.1, lanes 4,6) concomitantly with the partial loss of radioactivity from the high-M_r precursor. As expected, protease also promoted the appearance of methionine-labelled mature capsid proteins (fig.1, lane 6).

The conclusion that p14 and p12 are derived from preA was further supported by the fact that [^{35}S]-methionine-labelled tryptic peptides of the two former species appeared to be (closely related to one another) subsets of tryptic peptides of the latter (fig.2). The patterns of non-p14 (p12) tryptic peptides of preA, on the one hand, and of polypeptide A synthesized either in vitro or in vivo, on the other, were nearly identical to each other.

These data establish that in the preA molecule (and by inference, in the EMC virus polyprotein molecule), the amino acid sequence of polypeptide A, a precursor of capsid proteins, is preceded by sequences corresponding to non-structural polypeptides p14 and p12.

This conclusion seems to contradict the apparent labelling of polypeptide A with formyl-methionine observed here (fig.1, lane 2,3) and in [2]. The controversy may easily be explained, however, if we assume that the labelled polypeptide comigrating with A is not an authentic A but rather an unfinished preA (A and unfinished preA lack a N-terminal and a C-terminal portion of preA, respectively).

The presence of coding sequences that precede the first capsid protein gene does not appear to be a common feature of all picornavirus genomes. While leader polypeptides similar to those described here were also reported for foot-and-mouth disease virus [17], they have been neither found among the product of in vitro translation nor deduced from the primary structure of the genome [18,19] in the case of poliovirus.

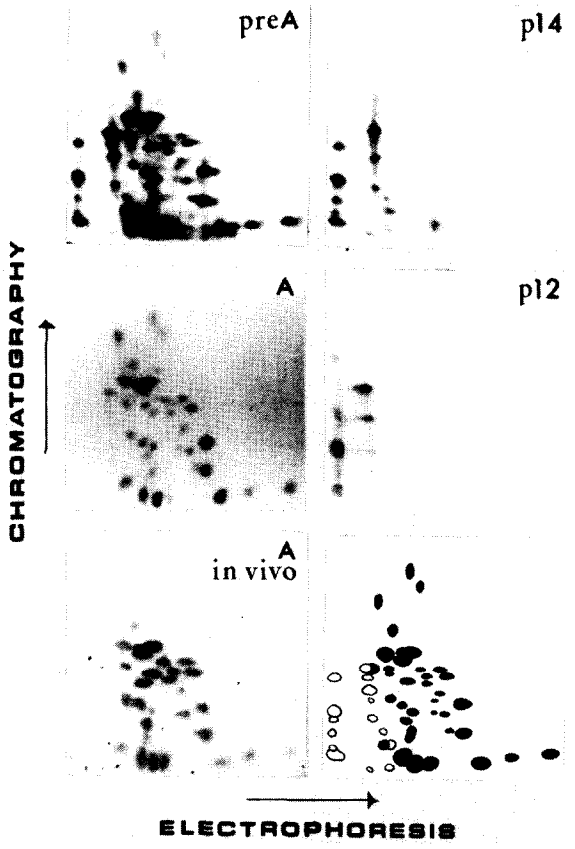

Fig.2. Tryptic peptide maps of the in vitro synthesized polypeptides preA, A, p14 and p12 as well as of the in vivo synthesized polypeptide A. On the scheme, white spots represent peptides derived from p14 and p12, black spots correspond to the non-p14(p12) peptides of preA.

The function, if any, of the picornaviral leader polypeptides remains unknown, but our impression is that they are in many respects similar to signal polypeptides of some membrane-reacting proteins [20]. Judging by the failure of the infected cells to accumulate appreciable amounts of the leader polypeptides, they are unstable, rapidly turning-over products. This inference is also supported by some in vitro observations [15].

While elucidating the gross relative location of the coding sequences in the 5′-terminal region of the EMC virus genome, the above findings pose another important problem: What is the difference between the closely related polypeptides p14 and p12, both of which could be labelled with formyl-methionine? If the smaller polypeptide is devoid of a N-terminal portion of the larger one, this would signify the existence

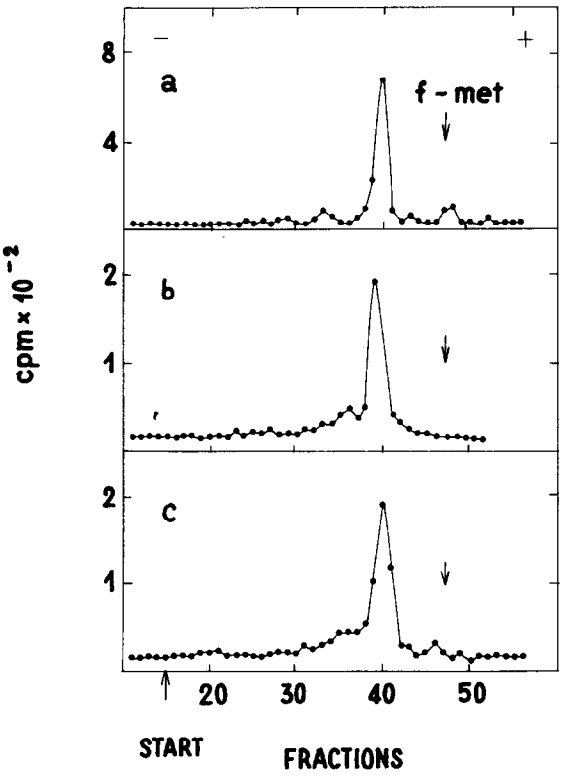

Fig.3. Electrophoretic patterns of the formyl-methionine-labelled tryptic peptides derived from the products of in vitro translation of EMC virus RNA. Paper electrophoresis was carried out at pH 10.2 at 2.2 kV for 120 min. Peptides derived from the total in vitro product (a), from p14 (b) and from p12 (c). The position of formyl-methionine is indicated by arrows.

of two different sites for initiation of translation. If, on the other hand, the difference between the two polypeptides resides in their C-termini, alternative cleavage sites should be invoked.

The formyl-methionine-labelled tryptic peptides of the preA-containing total in vitro product, p14 and p12 appear to be represented by the same single major species when electrophoresed at pH 10.2 (fig.3). This suggests that the N-terminal amino acid sequences of all 3 proteins are identical. However, trypsinolysis of formyl-methionine-containing proteins, regardless of their nature, sometimes gives rise to more than one major labelled products. Several lines of evidence suggested that these additional products were, most likely, of artifactual origin (unpublished).

Thus, a single major initiation site appears to be utilized for the synthesis of both leader polypeptides, p14 and p12, as well as of preA. Some circumstantial evidence suggests that p14 is not a precursor to p12; rather, both polypeptides are produced by alternative cleavages of preA. If this suggestion is correct, the questions concerning the heterogeneity of the N-terminal sequence of the capsid protein precursor, polypeptide A, and the possible correction of this heterogeneity, arise. Further experiments are required to answer these questions.

References

[1] Rueckert, R. R. (1976) in: Comprehensive Virology (Fraenkel-Conrat, H. and Wagner, R. eds) vol. 6, pp. 131–213, Plenum, New York.
[2] Pelham, H. R. B. (1978) Eur. J. Biochem. 85, 457–462.
[3] Smith, A. E., Marcker, K. A. and Mathews, M. B. (1970) Nature 225, 184–187.
[4] Öberg, B. F. and Shatkin, A. J. (1972) Proc. Natl. Acad. Sci. USA 69, 3589–3593.
[5] Smith, A. E. (1973) Eur. J. Biochem. 33, 301–313.
[6] Gorbalenya, A. E., Svitkin, Yu. V., Kazachkov, Yu. A. and Agol, V. I. (1979) FEBS Lett. 108, 1–5.
[7] Ginevskaya, V. A., Skarlat, I. V., Kalinina, N. O. and Agol, V. I. (1972) Arch. ges. Virusforsch. 39, 98–107.
[8] Svitkin, Yu. V. and Agol, V. I. (1978) FEBS Lett. 87, 7–11.
[9] Svitkin, Yu. V., Gorbalenya, A. E., Kazachkov, Yu. A. and Agol, V. I. (1979) FEBS Lett. 108, 6–9.
[10] Svitkin, Yu. V., Ugarova, T. Yu., Chernovskaya, T. V., Lyapustin, V. I., Lashkevich, V. A. and Agol, V. I. (1981) Virology 110, 26–34.
[11] Laemmli, U. K. (1970) Nature 227, 680–685.
[12] Morrison, T. C. and Lodish, H. F. (1975) J. Biol. Chem. 250, 6955–6962.
[13] Gautch, J. W., Lerner, R., Howard, D., Teramoto, Y. A. and Schlom, J. (1978) J. Virol. 27, 688–699.
[14] Celma, M. L. and Ehrenfeld, E. (1975) J. Mol. Biol. 98, 761–780.
[15] Agol, V. I., Chumakov, K. M., Dmitrieva, T. M. and Svitkin, Yu. V. (1980) in: Biology Reviews (Skulachev, V. P. ed) vol. 1, pp. 319–370, Soviet Sci. Rev. Harwood Acad. Publ. GmbH, Chur.
[16] Palmenberg, A. C., Pallansch, M. A. and Rueckert, R. R. (1979) J. Virol. 32, 770–778.
[17] Sangar, D. V., Black, D. N., Rowlands, D. J., Harris, T. J. R. and Brown, F. (1980) J. Virol. 33, 59–68.
[18] Kitamura, N., Semler, B. L., Rothberg, P. G., Larsen, G. R., Adler, C. J., Dorner, A. J., Emini, E. A., Hanecak, R., Lee, J. J., Van der Werf, S., Anderson, C. W. and Wimmer, E. (1981) Nature 291, 547–553.
[19] Racaniello, V. R. and Baltimore, D. (1981) Proc. Natl. Acad. Sci. USA 78, 4887–4891.
[20] Blobel, G., Walter, P., Chang, C. N., Goldman, B. M., Erickson, A. H. and Lingappa, V. P. (1979) Symp. Soc. Exp. Biol. 33, pp. 9–36, University Press, Cambridge.