A Re-appraisal of the Biochemical Map of Foot-and-mouth Disease Virus RNA

By T. R. DOEL, D. V. SANGAR, D. J. ROWLANDS and F. BROWN
Animal Virus Research Institute, Pirbright, Surrey, U.K.

(Accepted 13 June 1978)

SUMMARY
The proteins induced by infection of BHK 21 cells with foot-and-mouth disease virus have been compared by tryptic peptide analysis. The results indicate that there are three primary products 5'-P88, P52, P100-3'. The polypeptide P56, which we considered previously to be a primary product, is derived from the region of the genome that codes for P100. The results indicate that there are alternative cleavage pathways of P100, the polypeptide coded for by the 3' end of the genome.

INTRODUCTION
The genetic information of the picornaviruses is encoded in a single stranded RNA of mol. wt. about 2.8 × 10^6. It has been demonstrated in several laboratories that only one major initiation site for protein synthesis is expressed in vivo. The genome is therefore translated entirely and without interruption. This total translation product or 'polyprotein' has, with one exception (Kiehn & Holland, 1970), only been detected when further processing was inhibited by incubating the infected cells in the presence of amino acid analogues or inhibitors of proteolytic enzymes. Under normal conditions the first cleavages occur on the nascent polypeptide chain and produce the proteins first detected during a pulse label; these proteins have been termed the primary products. Three primary products have been detected for several picornaviruses. In every case, the product derived from the 5' end of the RNA has a mol. wt. of approx. 90 to 100 × 10^3 and is the precursor of the structural polypeptides of the virus particle. A polypeptide of similar size is translated from the 3' end of the RNA while the middle region of the genome gives rise to the third primary product with a mol. wt. of 30 to 50 × 10^3 (see review by Rekosh, 1977, for references).

In a previous paper we provided evidence for the presence in foot-and-mouth disease virus (FMDV) infected cells of four primary products which we designated 5'-P88, P52, P56, P100-3', the number representing the mol. wt. in thousands (Sangar et al. 1977). The major difference between FMDV and the other picornaviruses was the presence of an additional primary product with a mol. wt. of about 50 × 10^3. Pactamycin mapping showed clearly the relative positions of P100, P88 and P52 but difficulty was found in determining the relative positions of P100 and P56. However, several experiments indicated that P56 was nearer to the 5' end of the genome. The four putative primary products were found in approximately equimolar amounts, providing evidence that they arose from distinct regions of the genome.

Because of the apparent difference between the number of primary products in FMDV infected cells compared with the number in cells infected with other picornaviruses, we have now examined in more detail the primary products of FMDV and their major cleavage
products by tryptic peptide mapping. We showed previously that the P56 region of the polyacrylamide gels used to separate the induced proteins contained two polypeptides, P56a and P56b, which we considered from pacytamycin mapping and kinetic data to be alternative forms of the same polypeptide. By omitting urea from the gels we have been able to resolve a third protein migrating between those we originally designated P56a and P56b. We have designated this third protein P56b, our original P56b becoming P56c.

Tryptic peptide mapping of the primary products has shown that P100, P88 and P52 are different from each other. However, the maps produced from P56a and P56c together gave a pattern similar to that from P100, providing strong evidence that P56a and P56c are derived from the part of the genome that codes for P100.

METHODS

Virus. FMDV of serotype A (A 6I, subtype 10) was used in all experiments. When 3H-leucine labelled virus was required, BHK 21 cell monolayers were infected at a m.o.i. of 10. After 30 min the inoculum was replaced with leucine-free Eagle's medium and after a further 60 min, 3H-leucine (5 µCi/ml) was added. The medium was separated from the harvest when the cells had left the glass and the virus particles purified by the method described by Harris et al. (1977) for swine vesicular disease virus.

Labelling of virus induced polypeptides. Monolayers of BHK cells in 4 oz bottles were infected at 37 °C at a m.o.i. of 100. Virus was allowed to attach and absorb for 30 min after which the monolayers were washed in either methionine- or leucine-free Eagle's medium. The cells were incubated for a further 90 min before they were pulsed with either 200 µCi 35S-methionine or 50 µCi 14C-leucine for 5 min. 3H-labelled induced proteins were prepared by pulsing infected cells for 30 min with 500 µCi of either 3H-methionine or 3H-leucine at 120 min post-infection in the presence of 2.5 x 10⁻⁴ M-zinc acetate.

Separation and isolation of the induced polypeptides. The labelled cells were lysed, after removing the supernatant fluid, by adding 1 ml of disrupting solution (0.125 M-tris, pH 6.8, 2% SDS, 2% mercaptoethanol, 10% glycerol, 0.5 M-urea) directly to the cells. The solution produced was boiled for 15 min to reduce the viscosity and then analysed either on 7.5% polyacrylamide slab gels, for the 35S and 14C preparations, or cylindrical gels for the 3H preparations, using the gel system described by Laemmli (1970). Slab gels were fixed in 20% methanol, 9% acetic acid for 30 min after which they were washed in distilled H2O for 30 min. The gels were then dried on wettable cellophane using an apparatus similar to that described by Fairbanks et al. (1965). The dried gels were left in contact with Kodirex film for appropriate times; the developed film was then placed over the gel and the positions of the polypeptides marked by piercing the gel through the autoradiograph with a pin. The strips of gel containing the polypeptides were cut out and eluted by boiling for 2 min in 5 ml 0.125 M-tris, pH 6.8, 0.5% SDS, 2 mM-phenyl methyl sulphonyl fluoride followed by a further 18 h at 37 °C. Cylindrical gels were fractionated, without fixing, into 1 mm slices and the proteins eluted at 37 °C for 18 h with 1 ml of 0.1 M-tris, pH 8.0, 1% SDS, 1% 2-mercaptopethanol, 2mM-phenyl methyl sulphonyl fluoride. The purity of the eluted proteins was checked by precipitating about 10% of the sample with 67% acetone at −20 °C in the presence of 20 µg bovine plasma albumin as carrier and analysing the precipitate on a 7.5% polyacrylamide slab gel. The polypeptide was visualized either by autoradiography or fluorography.

Preparation of tryptic peptides. Mixtures of the 35S-methionine and 3H-methionine or
Biochemical map of FMDV, a re-appraisal

14C-leucine and 3H-leucine labelled proteins to be compared were carboxymethylated and then digested with trypsin as described by Doel & Brown (1978).

Peptide mapping. Two methods were used:

Ion exchange chromatography. The peptides were separated on a 30 x 0.9 cm Chromosorb P resin column using a linear gradient of pyridine-acetate buffer, from 0.04 M-pyridine, pH 2.45, to 2.0 M-pyridine, pH 4.9, maintained at 54 °C (Burroughs et al. 1978). In the experiments with 14C- and 3H-leucine labelled peptides, fractions of 2.4 ml were collected and the data were analysed using a Fortran computer program. The data are shown as disintegrations/min.

High voltage paper electrophoresis. Mixtures of 35S- and 3H-methionine tryptic peptides were electrophoresed at 3 kV for 2 h on Whatman 3 MM paper using a pyridine-acetate buffer, pH 6.5. Usually, six tryptic digests were electrophoresed in parallel lanes, 60 cm long and 1 cm wide, on a single sheet of paper. Each lane was cut into 1 cm sections and these were counted in a Packard scintillation counter.

RESULTS

Separation and isolation of the virus induced proteins

In an earlier publication we deduced a scheme, shown in Fig. 1, for the translation and processing of virus-specified proteins found in FMDV infected BHK cells. The evidence for this mapping of the genome products was based on pulse-chase experiments using (a) the drug pactamycin to inhibit initiation of protein synthesis and (b) iodoacetamide to inhibit secondary proteolytic cleavages of the primary gene products.

To test the validity of this map we have isolated the major induced proteins and compared their tryptic peptides. An autoradiograph of a typical slab gel separation of the proteins in virus infected cells labelled for 5 min with 35S-methionine is shown in Fig. 2. Four proteins were seen routinely in the 50 to 60 x 10^3 mol. wt. region of the gel and these are designated, in decreasing order of apparent size P56a, P56b, P56c and P52. Polypeptide P56c is equivalent to P56b in our previous paper (Sangar et al. 1977) and P56b is a newly resolved protein. Attempts to improve the separation of these proteins by increasing the time of electrophoresis or by altering the gel concentration were unsuccessful. Although sufficiently pure preparations of P56a, P56c and P52, as judged by PAGE of the isolated proteins, could be obtained for peptide analysis, P56b was too heavily contaminated with P56a to merit analysis.

To prepare 3H-methionine or -leucine labelled polypeptides, infected cells were pulsed for 30 min in the presence of 2.5 x 10^{-4} M-zinc acetate. This compound greatly inhibits the secondary cleavages of the primary products, resulting in a build up of these polypeptides. The samples were electrophoresed in cylindrical gels which were then sliced into 1 mm fractions. This method allowed the isolation of P88, P100 and P122, the latter being found only in the presence of inhibitors of proteolytic enzymes. Polypeptides P56a, P56b, P56c and P52 could not be resolved by this method.

Tryptic peptide analysis of the primary products

The ion exchange chromatogram of a mixture of the tryptic peptides of 14C-leucine labelled P100 and 3H-leucine labelled P88 is shown in Fig. 3(a). It is clear, both from the positions at which the peptides elute and from the distribution of radioactivity in the peaks that P100 and P88 are not related. This result was confirmed by high voltage electrophoretic separation of a mixture of the tryptic peptides of 35S-methionine labelled P100
Fig. 1. Scheme proposed by Sangar et al. (1977) for the cleavage pathways of FMDV induced polypeptides.

Fig. 2. Autoradiograph of a section of a preparative PAGE of the polypeptides induced in FMDV-infected BHK cells. The numbers represent the mol. wt. $\times 10^{-3}$ of the polypeptides. The faint band between P34 and the front contains the structural polypeptides; that between P52 and P38 is found consistently but not in quantities sufficient for peptide mapping. The band between P72 and P56a is found only rarely and has not been characterized.
Fig. 3. Ion exchange chromatograms of the tryptic peptides of (a) $^{14}$C-leucine P100 (—) and
$^3$H-leucine P88 (——); (b) $^3$H-methionine P100 (—) and $^{35}$S-methionine P52 (—); (c) $^3$H-
methionine P100 (—) and $^{35}$S-methionine P56a (—); (d) $^3$H-methionine P100 (—) and $^{35}$S-
methionine P56c (—).
Fig. 4. High voltage electropherograms of the tryptic peptides of (a) $^{35}$S-methionine P100 (---) and $^3$H-methionine P88 (----); (b) $^{35}$S-methionine P52 (-----) and $^{35}$S-methionine P34 (------). The P52 and P34 separations were made in separate lanes and have been superimposed to allow easier comparison. The arrow at fraction 37 indicates the origin.
Biochemical map of FMDV, a re-appraisal

Fig. 5. Ion exchange chromatograms of the tryptic peptides of (a) $^{14}$C-leucine P88 (---) and $^3$H-leucine virus structural polypeptides (---); (b) $^{35}$S-methionine P100 (---) and $^3$H-methionine P72 (---).

and $^3$H-methionine labelled P88 (Fig. 4a). In isolation, the peptide mapping data do not unequivocally eliminate the possibility of overlap between the sequences of P100 and P88. However, the present result is in complete agreement with our previous work which showed that P100 and P88 were distinct primary products (Sangar et al. 1977). The chromatogram of a mixture of the tryptic peptides of $^{14}$C-leucine labelled virion proteins and $^3$H-leucine labelled P88 confirms our earlier conclusion that P88 is the precursor of the structural proteins (Fig. 5a).

The ion exchange chromatogram of a mixture of the tryptic peptides of $^{35}$S-methionine labelled P52, the third putative primary product, and $^3$H-methionine labelled P100 shows that the majority of the P100 peptides are not seen in P52 (Fig. 3b). Polypeptide P52 is exceptional in that its tryptic peptides lack a major peak at or near fraction 27 but instead there is a considerable number of radioactive counts in a unique peak at fraction 35. The peak at fraction 27 is common to most proteins that we have mapped by this method and we consider it to be a high mol. wt. partially cleaved peptide. However, the number of radioactive counts in a single peak at or about fraction 35 is, in our experience, unique to P52. This protein is also exceptional in that most of the $^{35}$S-methionine labelled tryptic peptides are either neutral or acidic. This is best illustrated in Fig. 4(b) which shows that
most of the radioactivity is either at or to the acidic side of the origin in high voltage paper electropherograms.

The close similarity of P52 and P34 is also seen in Fig. 4(b), confirming our earlier results (Sangar et al. 1977) which indicated that P52 was the precursor of P34. The peptides from P34 were electrophoresed in a lane parallel to the P52 peptides but are shown as superimposed profiles for convenience. Because we have found that reproducible comparisons can be made between tryptic digests electrophoresed in parallel lanes it is valid to compare the positions of the P52 peptides in Fig. 4(b) with those of P88 and P100 in Fig. 4(a). This comparison shows that P52 and P88 are unrelated.

As explained above, the P56 region of the gel resolved into three proteins P56a, P56b and P56c but only P56a and P56c were extracted in a form sufficiently pure for analysis. An ion exchange chromatogram of a mixture of the tryptic peptides of 35S-methionine labelled P56a and 3H-methionine labelled P100 is shown in Fig. 3(c). Most of the peptides are common to both proteins, the major feature being the different relative amounts of certain peptides in P100 and P56a. We have confirmed this observation with leucine labelled peptides separated by the same method. A similar situation is found with P56c (Fig. 3d) which was originally regarded as a primary product and was thought to give rise to P56a by a post-translational modification. It is interesting that some of the peptides that are prominent in P56a are greatly reduced in P56c (e.g. fraction 13) and vice versa (e.g. fraction 83). These results suggest (1) that P56a and P56c are contaminated with each other when they are extracted from gels, although they appear to be pure when monitored by PAGE and (2) that they are derived from different regions of the part of the genome that codes for P100. The results do not permit us to eliminate the possibility that there is some overlap in the sequences of the two polypeptides.

P56a and P56c were also shown to be related to P100 by using either the technique of partial proteolytic digestion (Cleveland et al. 1977) or cleavage with cyanogen bromide (D. V. Sangar & D. J. Rowlands, unpublished results) followed by PAGE. In contrast the 35S-methionine containing tryptic peptides of P56a or P56c did not show any similarities to those obtained from P88 when compared by high voltage paper electrophoresis (T. R. Doel, unpublished data).

At present we cannot speculate usefully as to whether the combination of P56a and P56c (total mol. wt. $112 \times 10^3$) is equivalent to P100 alone (with the discrepancy in mol. wt. being due to the anomalies frequently observed in mol. wt. determination by PAGE) or whether they are derived from P122, a protein normally seen only in the presence of inhibitors of proteolytic enzymes (Sangar et al. 1977). Polypeptide P122 has been found to be closely related to P100 by tryptic peptide analysis (T. R. Doel, unpublished data).

In our earlier paper (Sangar et al. 1977) it was shown by pactamycin mapping that P100 gave rise to P72 via a series of cleavages. Tryptic peptide analysis of the two proteins supports this conclusion (Fig. 5b).

**DISCUSSION**

The work described in this paper was done to test the validity of the biochemical map for FMDV that we had constructed mainly from the results of pulse-chase experiments and the use of pactamycin (Sangar et al. 1977). We concluded that four primary products were produced and not the three found in cells infected with other picornaviruses because, in the presence of proteolytic cleavage inhibitors, four polypeptides were detected in approximately equimolar amounts.

In agreement with our previous studies we have now shown, by tryptic peptide analysis,
that the primary products P100, P88 and P52 are derived from different regions of the FMDV genome. We have also confirmed that (1) P100 is related to P72; (2) P88 is the precursor of the virus structural proteins and (3) P52 and P34 are closely related.

A major difference from our previous work was found, however, when P100 and P56 were analysed. P56 is a mixture of at least three polypeptides, P56a, b and c. We had found previously that P56c was unstable during a chase and that the radioactivity lost from this protein was equal to the increase in radioactivity associated with P56a. Because P100 appeared to be converted quantitatively to P72 and could not therefore represent a precursor to P56a we had concluded that P56a was derived from P56c by a post-translational modification. It should be emphasized that in our previous work only two polypeptides were detected in the P56 region and these were designated P56a and P56b. Polypeptide P56a has now been resolved into two peaks which we have designated P56a and P56c.

Tryptic peptide analysis of P56a and P56c suggested that these polypeptides were different from each other but contained only those peptides found in P100. We have not been able to isolate P56b sufficiently free from P56a for it to merit analysis. However, in contrast to P56a and P56c, polypeptide P56b is not detected in infected cells which have been treated with pactamycin for 5 min before pulsing for 5 min with $^{35}$S-methionine, suggesting that it originates from the 5' end of the genome.

The relationships of the polypeptides derived from P100 and the cleavages involved appear to be very complex. It seems that P100 can be cleaved in two ways. One results in the formation of P72 via a series of intermediates P80, P78 and P76. The second results in the formation of P56a and P56c, the cleavages being very rapid and possibly on the nascent polypeptide. In this case P56a and P56c could be considered to be alternative primary products to P100 and this concept is supported by the insensitivity of the cleavages producing P56a and P56c to inhibitors which prevent all other secondary cleavages.

The mol. wt. of P100 is not large enough to accommodate two P56 proteins, unless P56a and P56c share common sequences. In addition we regularly detect a protein of mol. wt. $122 \times 10^3$ following treatment of infected cells with zinc acetate or the proteolytic enzyme inhibitor 'Trasylol'. Furthermore, P122 is essentially identical to P100 by tryptic
peptide analysis, which could mean that PI22 is the rarely observed precursor of P56a and P56c. Alternatively, the discrepancy in mol. wt. may be another example of the well known errors inherent in the determination of mol. wt. by PAGE. The amount of P56a increases during a chase but we do not know whether this increase arises from P100 or P72. The alternative cleavages exhibited by P100 may be a mechanism for making maximum use of limited genetic material.

The probable interrelationship between the induced proteins of FMDV are shown in Fig. 6. This 'biochemical map' is very similar to that published for EMC virus in that both possess three major primary products, the precursor of the structural proteins is coded for by the 5' end of the genome and the polypeptide coded for by the 3' and gives rise to alternative cleavage products.

As mentioned above, the strongest evidence for the postulate that there were four primary products was the observation that P100, P88, P56, and P52 were present in equimolar amounts and all were found in conditions which inhibited secondary cleavages. Because P56a and P56c are not related and there is an additional polypeptide P56b in this region, the apparent molar ratio of the P56 region is considerably lower than we considered previously. We have calculated the molar ratio of P100:P56a:P88:P52 to be approximately 1:3:1:1. It is not known at present whether these ratios indicate that the 3' region of the FMDV genome is translated more frequently than the remainder of the genome or whether the polypeptides translated from the 5' region are preferentially degraded.

It is worth noting that alternative cleavage pathways for the primary product coded for by the 3' end of the genome were also observed by McLean et al. (1976) in cells infected with a human rhinovirus, although in this system one of the pathways greatly predominated. It seems possible that this mechanism could be a common method for making maximum use of the available genetic information.

Note added. We have recently detected in infected cells a polypeptide with a mol. wt. of $20 \times 10^3$ which has the characteristics of a primary product. Preliminary mapping experiments using pactamycin suggest that this protein is located towards the 5' end of the genome.

REFERENCES


(Received 12 April 1978)