Volume 6 Number 7 1979 Nucleic Acids Research

Sequence and location of the poly C tract in aphtho- and cardiovirus RNA

D.N.Black, Pauline Stephenson, D.J.Rowlands and F.Brown

Animal Virus Research Institute, Pirbright, Surrey, UK

Received 2 April 1979

ABSTRACT

The poly C tract in the RNA of the aphtho- and cardio viruses has been examined in several isolates of foot-and-mouth disease virus (FMDV) and encephalomyocarditis (EMC) virus. The length of the tract is variable, containing 100 to 170 bases in the FMDV isolates and 80 to 250 bases in the EMC virus isolates. Each poly C tract contains c. 10% A and U residues, located at the 5' end, i.e. most of the tract is a continuous run of C residues. The position of the tract on the genome was the same in each of the FMDV isolates, about 400 bases from the 5' end, whereas in the EMC virus isolates it was about 150 bases from the 5' end.

INTRODUCTION

The family Picornaviridae is composed of more than 200 members, which have been divided into 4 genera: entero-, cardio-, rhino- and aphthoviruses (1). All the members examined contain one molecule of a single-stranded RNA, molecular weight c. 2.5 x 10^6. The RNAs of the aphtho- and cardioviruses differ from those of the entero- and rhinoviruses in containing a long tract (100 to 250 residues) which is resistant to RNAse T_1 (2, 3). This tract comprises c. 90% C residues and hence has been termed poly C. It is located towards the 5' end of the RNA in EMC and Mengo viruses (4, 5) and FMDV (6). However its precise location has been determined for only one virus strain, FMDV, type A, where it was found to be 400-500 bases from the 5' end (7). In a study to determine whether the poly C tract has a specific function in the replication of the virus RNA, it seemed necessary to determine first some general features regarding its sequence and precise location on the genome of several isolates of the two genera. The present study has been directed towards obtaining answers to the following three questions: 1. Is the variability in length of the poly C tract found in FMDV isolates (3) also present in different isolates of EMC virus? 2. Where are the A and U residues

located in the poly C tract of different isolates? 3. Is the poly C tract at the same position on the genome of the different viruses?

MATERIALS AND METHODS

Viruses

Nine virus isolates were examined. Their origin is given in Table 1. Unlabelled virus was grown in monolayers containing 10^8 BHK21 cells in the presence of Eagle's medium at 37°C, using a multiplicity of infection of 0.01. Virus labelled with ^{3}H-cytidine was grown in the same medium in the presence of 200-400μCi ^{3}H-cytidine (sp. act. 27Ci/mmole). Virus labelled with ^{32}P was grown in the presence of 150μCi/ml carrier free ^{32}P orthophosphate in phosphate free Earle's saline after infecting at a multiplicity of 100.

Preparation of virus RNA

The harvested viruses were purified by the method described by Brown and Cartwright (8) and the RNA extracted with phenol:chloroform (2:1) in the

TABLE 1. Species of origin of virus isolates.

Virus	Species of origin	Source
Aphtho		
FMDV-A61	Bovine	World Reference Library, A.V.R.I.
-SAT_1-7*	Bovine	Dr. G.N. Mowat, A.V.R.I.
-SAT_1-82*	Bovine	
Cardio		
EMC-R^3	Chimpanzee	Dr. R.R. Rueckert, Madison, Wisconsin,USA
-DW	Pig	Dr. D. Watt, Brisbane, Australia.
-V251d	Water rat	Dr. I. Littlejohns, Glenfield, N.S.W., Australia.
-V297d	Pig	
-RS 3	Red squirrel	Dr. A.D. Vizoso, Unit of Invertebrate Virology, Oxford, U.K.
-GS 8	Grey squirrel	

* SAT_1-7 and SAT_1-82 were derived from cattle epithelium by 7 and 82 passages respectively in baby hamster kidney 21 cells.

presence of 0.5% SDS. The aqueous layer was then precipitated with 2 vol. ethanol at -20°C overnight. The precipitate was washed twice with ethanol, dissolved in 200mM NaCl, 10mM tris, 1mM EDTA, 0.2% SDS (NTE/SDS) and centrifuged at 20°C in a 10ml 5 to 25% (w/v) sucrose gradient in NTE/SDS for 16hr. at 60,000g. The RNA fractions sedimenting at 37S were combined and precipitated with 2 vol. ethanol at -20°C overnight.

Digestion with RNAse T_1

Purified RNA was dissolved in 10mM tris, 1mM EDTA, pH 7.4 to a final concentration of 10-15µg RNA/ml and incubated at 37°C for 1hr. with RNAse T_1 (Sankyo, Japan) at a ratio of 100 units of enzyme/100µg RNA. The digestion was stopped by adding 0.5ml of 200mM NaCl, 10mM tris, 1mM EDTA, 1% sarcosyl, pH 7.5 (NTE/sarc). The enzyme resistant fragment was separated from the smaller hydrolysis products on a Sephadex G-100 column equilibrated with NTE/sarc and precipitated with 2 vol. ethanol at -20°C overnight.

Estimation of size of poly C tract

^{32}P-poly C tracts prepared from each of the virus RNAs were electrophoresed on 20% acrylamide, 7M urea slab gels in 50mM tris-borate, 1mM EDTA, pH 8.3 at 1000V, 35mA for 24 hr. at 60-80°C. Internal standards of the poly C tracts of EMC virus-R^3 and FMDV-A61 were used.

^{32}P labelling of the 5' end of the poly C tract

Preparations of poly C were dissolved in kinase buffer (50mM glycine - NaOH, 10mM $MgCl_2$, 5mM dithiothreitol, 100mM spermidine, 25% glycerol, pH 9.5) and added to 300-500µCi γ^{32}P-ATP which had been evaporated to dryness in a siliconised Eppendorf tube. Polynucleotide kinase(1unit/10µl reaction mixture) was then added and the mixture incubated at 37°C for 30 min. The reaction was stopped with 0.1mM EDTA (3µl/10µl reaction mixture), an equal vol. of formamide-dye mixture (50ml deionised formamide, 10ml 100mM EDTA, 0.02% bromophenol blue, 0.02% xylene cyanol FF) (9) added, the mixture heated at 100°C for 2 min and then immediately cooled on ice. The "kinased" poly C tract was separated from the excess γ^{32}P ATP by electrophoresis on a 20% acrylamide 7M urea gel. This separation step incidentally demonstrated the homogeneity of the "kinased" poly C tract. The poly C tract, located by autoradiography using an exposure of 1 to 5 min, was eluted from the gel with 500mM NH_4 acetate, 100mM EDTA, 0.5% SDS, pH 7.0 at 37°C for 2 hr., using the technique described by Simoncsits et al. (9) and precipitated with 2 vol. ethanol at -20°C.

Sequencing of the poly C tracts

The sequence was determined by the rapid gel sequencing technique described by Simoncsits et al. (9), with modifications subsequently devised by the same authors. All the RNA samples contained 1µg t-RNA as carrier.

C and U analysis

1µl RNA, 7µl H_2O, 1µl buffer (200mM tris, 20mM EDTA, pH 7.5) and 1µl pancreatic RNAse (0.1µg/ml) were kept at 0°C and 3µl aliquots were removed at 3, 10 and 20 min, pooled and stored at -70°C until analysed.

C, U and A analysis

0.5µl RNA, 7µl H_2O, 1µl buffer (100mM Na acetate, 10mM EDTA, pH 4.5) and 1µl Physarum RNAse 1 (7 units/ml) were kept at room temperature. 3µl aliquots were removed at 3, 10 and 30 min, pooled and stored at -70°C until analysed.

A analysis

1µl RNA, 8µl buffer (8.75M urea, 20mM Na acetate, 2mM EDTA, pH 4.5) and 1µl RNAse U_2 (1 unit/ml) were keptfor 2.5 min at 50°C. The samples were then stored at -70°C until analysed.

All samples were mixed with an equal volume of formamide dye and boiled for 2 min before PAGE analysis.

Preparation and sizing of the S fragment by RNAse H treatment

5µg RNA, labelled with ^{3}H-cytidine, 1µg oligo dG_{8-12}, 1µl RNAse H (500 units/ml) in 60µl buffer (50mM tris, 1mM EDTA, 10mM $MgCl_2$, 10mM DTT, pH 7.9) were incubated at 30°C for 40 min. The reaction was terminated by the addition of 10µl 175mM EDTA, 2.5% SDS. The resultant S fragments were sized on a 10% acrylamide gel (10).

RESULTS

Size of the poly C tracts

The observation that the length of the poly C tract was different in different strains of FMDV (3) led us to determine whether the same variability occurred in different isolates of EMC virus. Fig. 1 shows the PAGE analysis on a 7M urea denaturing gel of the poly C tract from 6 isolates of EMC virus.

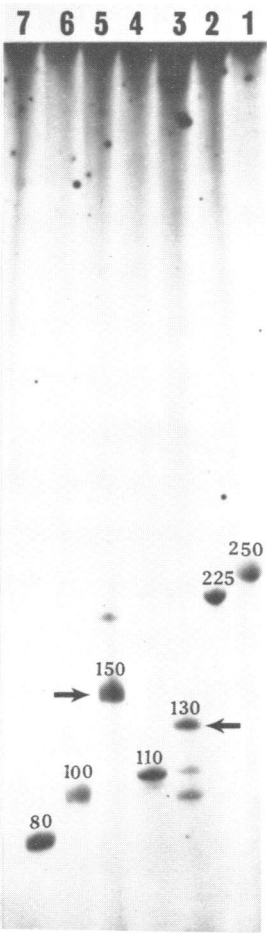

Fig.1. Analysis of the size of the poly C tracts of the RNAs of 7 isolates of FMDV and EMC virus. The RNAse T_1 digests of the ^{32}P-RNAs were electrophoresed on 20% polyacrylamide slab gels. Lane 1, EMC-DW; lane 2, EMC-251d; lane 3, EMC-GS 8; lane 4, EMC-RS 3; lane 5, FMDV-A61; lane 6, EMC-R^3; lane 7, EMC-297d. In lanes 3 and 5, the poly C tracts are indicated by an arrow. The other bands in these lanes were not C rich. The number of bases in the poly C tracts are indicated on the individual lanes.

The size ranges from 80 to 250 bases compared with 100 to 170 bases for the 3 isolates of FMDV examined in this study (3).

Sequence of the bases in the poly C tracts

Previous analyses (3) of the poly C tracts from the R^3 isolate of EMC virus and 4 isolates of FMDV had shown that they contained c. 90% C residues with the remainder A and U. The method of preparation, RNAse T_1 digestion, ensured that only one G residue could be present.

The tracts were uniquely labelled at the 5' end with ^{32}P and the sequence determined by the method of Simoncsits et al. (9). The results for FMDV, A61 and 2 EMC viruses, DW and GS 8 are shown in Fig. 2 and 3 respectively. The

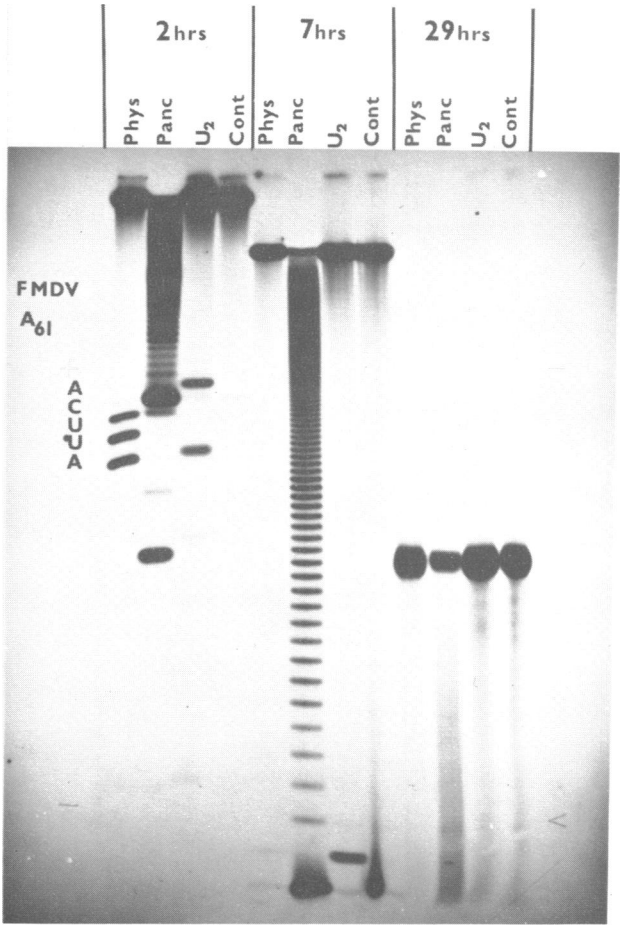

Fig. 2. Resolution by size on a 20% polyacrylamide slab gel of products obtained by controlled RNAse Physarum I, pancreatic and U_2 digestion of the poly C tract of FMDV-A61 RNA, labelled at the 5' end with ^{32}P.

most striking feature to emerge was that the C residues are present as a continuous tract from the 3' end with the A and U residues grouped at the 5' end. The sequences of these poly C tracts and those from additional isolates of FMDV and EMC virus are given in Table 2.

Location of the poly C tract on the genome

Rowlands et al (7) showed that in FMDV, type A61, the poly C tract was

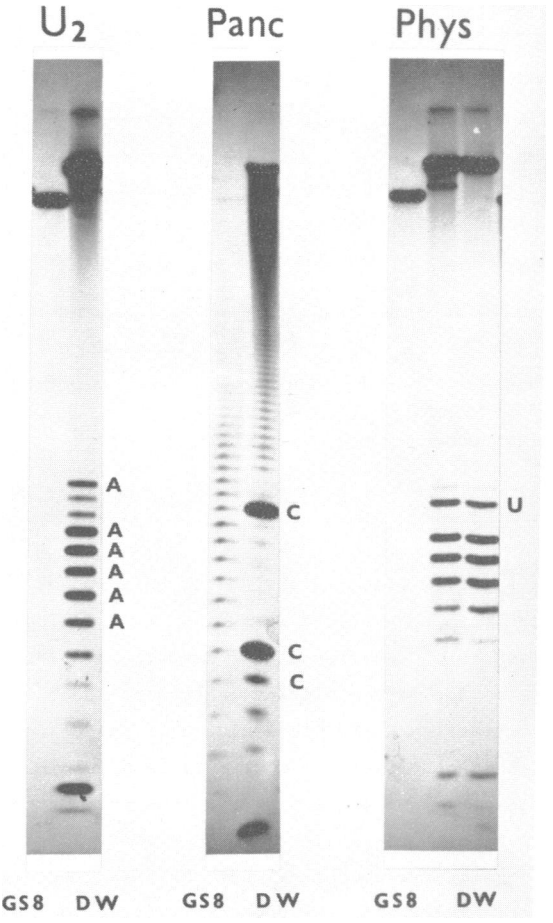

Fig. 3. Resolution by size on a 20% polyacrylamide slab gel of products obtained by controlled RNAse U_2, pancreatic and Physarum I digestion of the poly C tract from the RNAs of EMC viruses GS 8 and DW.

located 400-500 bases from the 5' end of the RNA. The size of this short section (the so-called S fragment) was determined by cutting the RNA with RNAse H at the poly C tract following its hybridisation to oligo dG. Application of this method to the isolates being examined in this study showed that the S fragments from the FMDV isolates were similar in size (400-500 bases) but were considerably longer than those from the EMC virus isolates (150 bases).

TABLE 2. Base sequence of poly C tracts from aphtho- and cardiovirus isolates.

Aphtho	5'	3'
FMDV-A61	CCAUUCAC ————	CG
-SAT_1-7	CCCUC ————	CG
-SAT_1-82	CACCUC ————	CG
Cardio		
EMC-DW	CAAAAACUAC ————	CG
-GS 8	CUUUUC$^{a}_{u}$CUAC ————	CG
-V297d	CUUCU$^{u}_{a}$CAC ————	CG

DISCUSSION

The presence of a poly C tract in the RNA of only two of the four genera of the Picornaviridae raises questions regarding its function, particularly since most of the structural features of the members of all the genera are essentially identical. Another perplexing feature is the variability in length of the poly C tract in different isolates of one of the genera, the aphthoviruses (3). This observation prompted us to determine whether a similar variability exists in members of the second poly C containing genus, the cardioviruses. The evidence presented in this study shows that the poly C tract in EMC virus strains isolated from many different sources is also highly variable in length. The reason for this variability, which is of the same order of magnitude in both FMDV and EMC viruses, is not known. However, it is noteworthy that the length of the tract in the attenuated strain of FMDV, SAT_1 - 82 is only about one-half of that in the virulent strain, SAT_1-7 from which it was derived (11). In unpublished work Miss K.J.H. Robson from this laboratory has also found a similar difference between the length of the poly C tract in a virulent and an attenuated strain of FMDV, type A.

Before a role can be proposed for the poly C tract it seemed necessary to determine the location of the A and U residues which comprise about 10% of the tract. The following structures seemed possible: 1. block repeats of $(C_{18}AU)_nG$ where the variation in n would generate the different lengths of poly C observed; 2. random distribution of the A and U residues in the tract; 3. specific grouping of the A and U residues, possibly giving some function

to those regions. The results show categorically that the A and U residues in the three viruses from each genus are located at the 5' end of the tract. A variety of codon functions for the 5' end of the tract can be proposed, depending on the reading frame that is used. However, the continuous run of C residues can only code for polyproline and there has been no demonstration of such a polypeptide or even one with enriched amounts of proline, in cells infected with FMDV (D.V. Sangar, unpublished observation). This provides good evidence that the poly C tract is not translated. The absence of a terminator codon at the 5' end of the poly C tract implies that the bases at that end also are not in a message area. The large fragment of RNA to the 3' side of the poly C tract in FMDV, type A61 translates as efficiently and completely *in vitro* in a rabbit reticulocyte system (12) as the complete RNA (13). All this information suggests that the regions of the RNA at the 5' end, including the poly C tract, are not translated, either *in vivo* or *in vitro*.

The tract is located *c.* 400 bases from the 5' end in all the FMDV RNAs examined, irrespective of the length of the poly C, but is only *c.* 150 bases from the 5' end in the EMC virus RNAs. The constant distance of the poly C tract from the 5' end of the genome in each genus strongly suggests that this region has a function in the replication of the viruses. The differences in the lengths of the poly C tracts of individual viruses and their sequence and precise position on the genome of the two genera may help to throw light on their function in the replication of the viruses.

ACKNOWLEDGEMENTS

We wish to thank Dr. Tim Harris for his help in estimating the size of the S fragments and Dr. George Brownlee for the *Physarum* enzyme. We are also grateful to the people mentioned in Table 1 for the gifts of viruses.

REFERENCES

1 Cooper, P.D. et al. (1978). Intervirology 10, 165-180.
2 Porter, A., Carey, N. and Fellner, P. (1974). Nature 248, 675-678.
3 Brown, F., Newman, J., Stott, J., Porter, A., Frisby, D., Newton, C., Carey, N. and Fellner, P. (1974). Nature 251, 342-344.
4 Chumakov, K.M. and Agol, V.I. (1976). Biochem. and Biophys. Res. Comm. 71, 551-557.
5 Perez-Bercoff, R., Gander, M. and Preisig, E. (1978). Virol. 85, 378-386.
6 Harris, T.J.R. and Brown, F. (1976). J. gen. Virol. 33, 493-501.
7 Rowlands, D.J., Harris, T.J.R. and Brown, F. (1978). J. Virol. 26, 335-343.
8 Brown, F. and Cartwright, B. (1963). Nature 199, 1168-1170.

9 Simoncsits, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilley, H. (1976). Nature 269, 833-836.
10 Harris, T.J.R., Dunn J.J. and Wimmer, E. (1978). Nucleic Acids Research 5, 4039-4054.
11 Harris, T.J.R. and Brown, F. (1976). J. gen. Virol. 34, 87-105.
12 Pelham, H.B. and Jackson, R.J. (1976). Eur. J. Biochem. 67, 247-256.
13 Sangar, D.V., Black, D.N. and Harris, T.J.R., personal communication.