Volume 8 Number 11 1980 Nucleic Acids Research

Nucleotide sequences at the termini of La Crosse virus RNAs

John F.Obijeski*, John McCauley and John J.Skehel

Division of Virology, National Institute for Medical Research, Mill Hill, London, NW7 1AA, UK

Received 11 April 1980

ABSTRACT

The 5' and 3'-terminal sequences of the three RNA molecules which make up the genome of La Crosse virus are reported. Fleven nucleotides at both the 5' and 3' termini of all three RNAs are conserved and complementary. In addition more extensive unique sequence complementarity is present in at least two of the three RNAs.

INTRODUCTION

The genome of La Crosse virus, a member of the Bunyaviridae (1), is made up of three RNA molecules of approximate weight 2.5, 1.6 and 0.4 x 10^6 daltons, designated L, M and S respectively (2). All three RNAs have been shown to be linear single strands of negative polarity (1,3,4). They can, however, be isolated from virus particles in circular nucleocapsid structures (2) and electron microscopic studies of other bunyavirus genomes have indicated that each RNA can under certain conditions assume a circular configuration (5,6). To investigate further the structure of the virus RNAs the nucleotide sequences at their 5' and 3' termini have been determined and are reported in this communication.

MATERIALS AND METHODS

Preparation of virus RNA. BHK-21 monolayer cell cultures were infected with La Crosse virus at an input multiplicity of 0.01-0.001 pfu/cell. After incubation at 33°C for 30 hr the virus was concentrated from clarified supernatant fluids with polyethylene glycol 6000 and NaCl and was purified by centrifugation in potassium tartrate:glycerol and sucrose gradients (7). Viral RNA was extracted from purified virus with sodium dodecyl sulphate (SDS) and phenol as previously described (2).

In vitro labelling of virus RNA. Virus RNAs were labelled at their 3'-termini with 200 µCi cytidine 3', 5'- (5'-^{32}P) bisphosphate and

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

T_4-RNA ligase by overnight incubation at 4°C (8). The 5'-termini were dephosphorylated with calf intestine alkaline phosphatase (20 units/ml) and labelled with 100-150 µCi (γ-^{32}P)-ATP and polynucleotide kinase as described (9). The labelled products were separated by electrophoresis at pH 3.5 in 2.2% polyacrylamide gels containing 7M urea and were eluted from the gel as described in (9). The eluted RNAs and carrier tRNA were precipitated with ethanol and the RNAs further purified by precipitation with cetyltrimethylammonium bromide (10).

Nucleotide sequence analysis. Partial nuclease digestion was done as described by Donis-Keller et al. (11). On occasion identification of pyrimidines was based primarily on the results of digestion with physarum RNase. These identities have not been confirmed and are, therefore, bracketed in Table 2. Chemical sequencing of 3'-labelled RNAs was by the method of Peattie (12).

Materials. γ-^{32}P-ATP (3000-5000 Ci/mmol) and cytidine 3',5'-(5'-^{32}P) bisphosphate (^{32}pCp) were obtained from the Radiochemical Centre, Amersham. ^{32}pCp was also obtained from New England Nuclear, Boston, Mass. Calf intestine alkaline phosphatase was from Boehringer-Mannheim; polynucleotide kinase and T_4-RNA ligase from PL Biochemicals; pancreatic RNase from Worthington; and T_1 and U_2 nucleases from Calbiochem. The nuclease preparation from physarum polycephalum (19) was given by Helen Donis-Keller of the Department of Biochemistry and Molecular Biology, Harvard University.

RESULTS

Terminal labelling of La Crosse virion RNAs. 5' and 3'-labelled RNAs were obtained by using γ-^{32}P-ATP and polynucleotide kinase and 5'-^{32}P pCp and T_4 RNA ligase respectively. It was noticed that with both procedures the efficiency of labelling decreased with increasing RNA molecular weight (Figure 1). In particular labelling of the large RNA (L) at the 5' terminus was so inefficient that the nucleotide sequence could not be determined. In addition to the three previously reported La Crosse virion RNAs (2), L, M and S, another labelled RNA species, designated X, of apparent moelcular weight 0.6 X 10^6 was obtained in different amounts in different virus RNA preparations. Terminal sequence analyses of this RNA species were also made.

Nucleotide sequence analysis of 3' and 5'-labelled La Crosse virus RNAs. The 3' terminal sequences of the RNAs were determined using both chemical

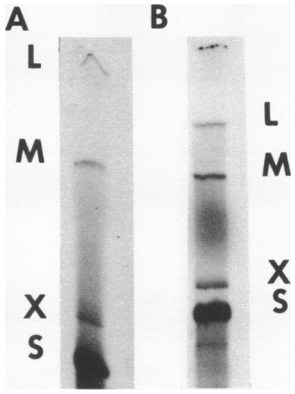

Figure 1. Autoradiograms of 5'-terminal (A) and 3'-terminal (B) ^{32}P-labelled La Crosse virus RNAs.

Table 1. 3'-terminal sequences of La Cross virus RNAs

```
                    10                  20                  30
L   U C A U C A C A U G A G G A U A G A U G U U U U G A A U G U
    C U U U U A A G U C A G U A U A G U G U U A U A U A C G U A
    U U A C C U G A U A G U U C U C C A U A G U U G U U A A G A
    A C C G A U C

                    10                  20                  30
M   U C A U C A C A U G A U G G U U C A U A U C U A U U G C A A
    A C U U A U A A U U U C A A A A C U U A G U U U C G G U U U
    C U A C U A A A C A U A U A A C C A C G A U U A A U G U C A
    A C G U C G A C G U U C G G G U C A C A U A G U U U C A C A
    A A G G U U C U A C

                    10                  20                  30
S   U C A U C A C A U G A G G U G A A C U U A U G A A A C U U U
    U A U U U A A C A A C A A C U G A C A A A A A A U G G A U U
    C C C C U U U A A U A G U U C U C A C A C U A C A G C C U A
    A A C C A C A A A A U A C U A C A G C G U A G U U G U C C A
    C G U U U A C C U A A A C U A G G A C U A C G U C C A U A U
    A C C U G A A G A C A C A A U U U U U A C G U C U U A G U G
    A G U U G G A A C G A C G U C A A U C U A G A A G A A G G A
    G U A C G
```

and partial nuclease digestion procedures and are shown in Table 1. The first eleven nucleotides of all the RNAs were identical; thereafter the sequences were unique for each RNA.

The results in Table 2 were obtained by partial nuclease digestion of 5'-labelled S and M RNAs. The first eleven residues were identical for both and the sequence of these nucleotides was apparently complementary to that of the eleven conserved residues at the 3' termini. In addition the 5'-terminal sequences of the S and M RNAs and the homologous 3'-terminal sequences were extensively complementary for the first fifty to sixty residues. The 5'-terminal sequence of the X-RNA was similar to that of the M RNA but additional nucleotides were also detected. As a consequence the relationship between the X RNA and the other virus RNAs is not clear and experiments on this point are in progress.

There are two regions near the 5' terminus of the SRNA for which ambiguous sequence results were obtained. The first between nucleotides 20 and 24 was apparent following heat denaturation of the RNA to be digested which resulted in nonspecific cleavage at these positions (Figure 2A and 2B). The second was between nucleotides 65 and 90 where nuclease digestion of non-denatured SRNA was not obtained (Figure 2C and 2D). However, following heat denatureation this region of the molecule could be sequenced and was shown to be pyrimidine-rich (Table 2).

DISCUSSION

The information collected in this study has shown that eleven nucleotides at both the 5' and 3' termini of all three La Crosse virus RNAs are conserved and are probably complementary. Similar observations have been made regarding the conservation of terminal sequences amongst the segmented genomes of other viruses such as brome mosaic virus, influenza virus,

Table 2. 5' terminal sequences of La Crosse virus RNAs*

	10	20	30
M	A G U A G U G U G(C)U U C C A A G U A U A A A A U A A U G U		
	U(U)G C A A A A C A A A U U U U C G(C)U G U U G U C A G U U		
	C		
	10	20	30
	A G U A G U G U G(C)U(C)C A(C)U G A A U A C A U U U U A U U		
	U U A U U U A U G A(C)(C)U G A U U U A(U)A G U U U U U U G C		
	U G U(U)(C)(C)(C)U U C C U C C C U C(C)C A U U U A G(C)U G(C)U		
	A		

* Residues not unambiguously identified are bracketed.

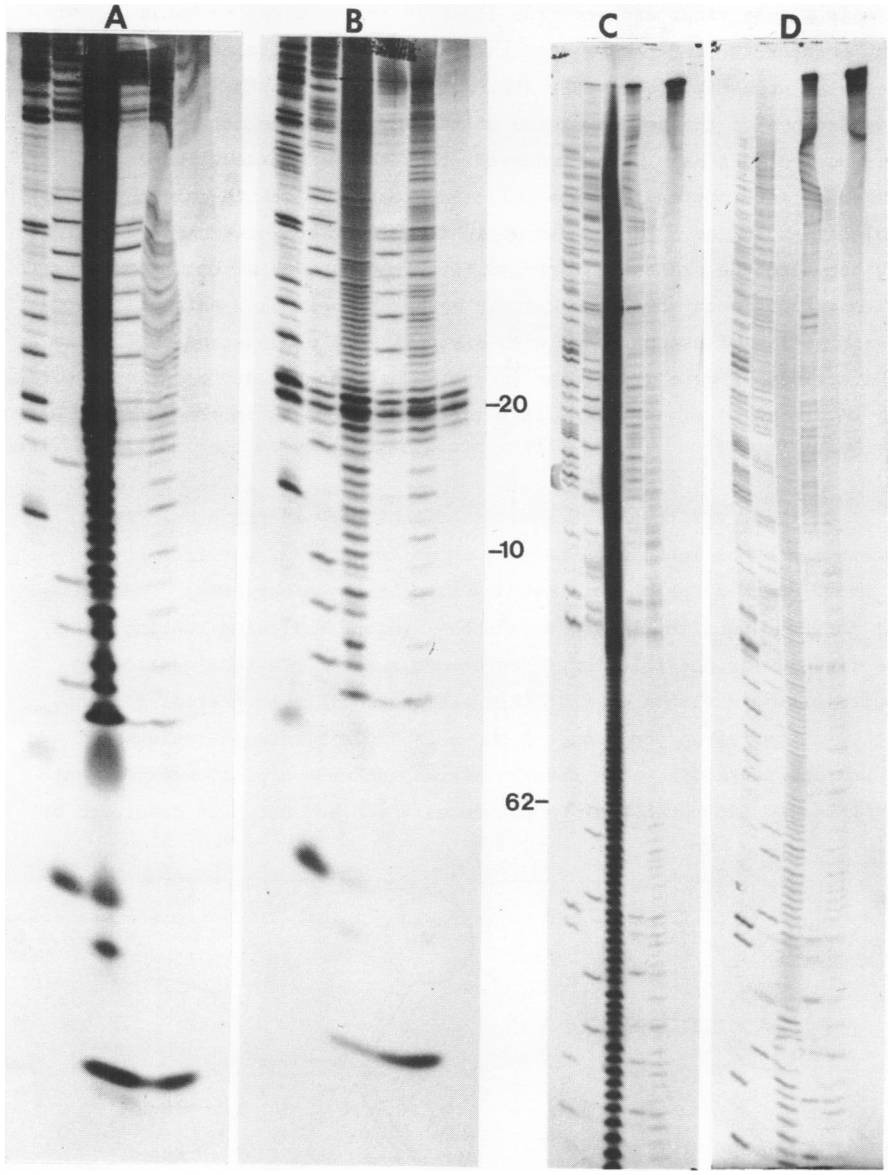

Figure 2. Autoradiograms of partial nuclease digestion products derived from 5'-labelled La Crosse S RNA. Electrophoresis was for 16 hr at 5 volts/cm in gels containing 20% acrylamide (A and B) or 24 hr at 10 volts/cm in 10% gels (C and D). The RNA used in B and D was heated at 100° for 2 min before digestion. From left to right in each section the oligonucleotides were obtained after digestion with pancreatic RNase; T_1-RNase; alkaline hydrolysis; U_2-RNase; Physarum RNase; and no nuclease control. The numbers indicate the distance in nucleotides from the 5' terminus.

alfalfa mosaic virus and reovirus (13,9,14,15). Others have also shown complementarity of 3' and 5'-terminal sequences of viral RNAs such as vesicular stomatitis virus RNA (16) and Sendai virus RNA (17) and the significance of the sequences and of the possible interaction of the oligonucleotides have been discussed. For the La Crosse virus RNAs a number of interactions are possible since not only are the eleven nucleotides at the 5' and 3' termini of the RNAs conserved and complementary but also the preceding forty to fifty nucleotides at both termini which are different for each RNA may be extensively base paired. The possibilities of establishing such distinct heterologous and homologous interactions (Figure 3) and the interplay between the different structures may be important during virus infection, for example, in the assembly of the virus genome and in the differential expression of the individual virus RNAs.

In the S RNA there are regions of ambiguity in the sequences determined by nuclease digestion at about twenty residues from both the 3' and 5' termini which were not observed in the other RNAs. The reason for the non-specific cleavages in these regions following heating is unknown. It is possible that they become susceptible to digestion by contaminating nuclease or that they are susceptible to thermal fission. Whatever the reason the sequence shown in Table 2 between residues 20 and 24 remains tentative. The results of the nuclease digestion experiments of SRNA also indicated that in molecules which had not been denatured by

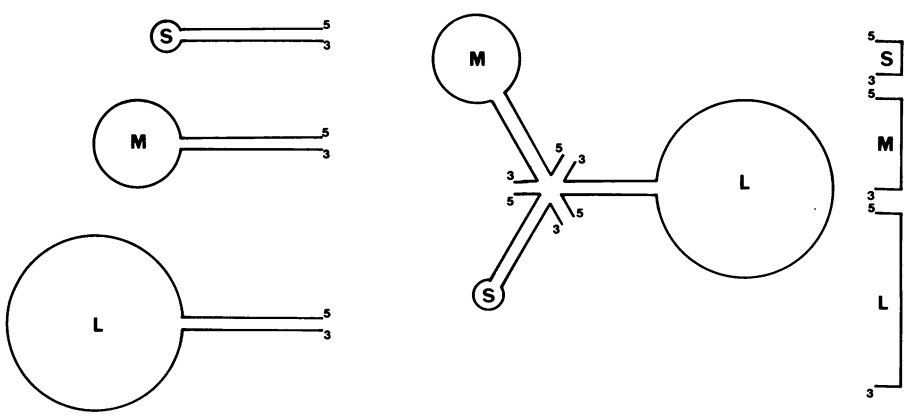

Figure 3. Examples of possible homologous and heterologous interactions of La Crosse virus RNAs.

Table 3. Amino acid sequences deduced from 3'-terminal sequences of La Crosse virus RNAs.

```
                    10
L   M H N G L S R G I N N S W L

                    10                  20
M   M I C I L V L I T V A A A S P V Y Q S V S K M

                    10                  20                  30
S   M S D L V F Y D V A S T G A N G F D P D A G I W T S V L K M
    Q N H S T L L Q L D L L P H A

 and                10                  20
    M D F C V K N A E S L N L A A V R S S S S C
```

heating the nucleotides between residues 65 and 90 from the 5'-terminus were insusceptible to digestion. The reason for this is also unknown particularly since the generation of 5'-terminal fragments longer than 90 residues was observed (Figure 2C). Since the nucleotide composition of this region is primarily pyrimidine the RNA may be simply resistant to digestion under the conditions used but this is contraindicated by the observations that heating renders the nucleotides susceptible to digestion. It is more likely that nucleotides 65 to 90 are involved in interactions with other regions of the molecule which are destabilized following nuclease digestion at nucleotides beyond residue 90.

Finally, from the 3' terminal sequences presented in Table 1 the complements of possible initiation codons were detected at residues 82-84, 101-103 and 150-152 for S RNA, 62-64 for M RNA and 55-57 and 62-64 for L RNA. In the S RNA the second of these was followed by a termination codon, at residues 164-166, and the other two by at least 131 and 63 bases respectively in which no termination codons were observed.
The first initiation codon in the L RNA is followed by at least 42 nucleotides; the second is followed by a termination codon at residues 95-97. The amino acid sequences deduced from these nucleotide sequence data are shown in Table 3. In the M RNA, which is proposed to code for the virus glycoproteins, the first initiation codon is followed by a sequence coding for 20 amino acids which are predominantly hydrophobic. These may represent a signal sequence involved in associating the nascent polypeptide chain with the endoplasmic reticulum (18).

ACKNOWLEDGEMENTS

We thank H. Donis-Keller for the gift of Physarum RNase, D. Peattie for communicating her sequencing procedure before publication and B.P. Holloway and D. Stevens for excellent assistance.

*Permanent address: Viral Exanthems Branch, Virology Division, Bureau of Laboratories, Center for Disease Control, U.S. Public Health Service, Department of Health, Education and Welfare, Atlanta, Georgia 30333, USA.

REFERENCES

1. Obijeski, J.F. and Murphy, F.A. (1977) J. gen. Virol., 37, 1-14.
2. Obijeski, J.F., Bishop, D.H.L., Palmer, E.L. and Murphy, F.A. (1976) J. Virol., 20, 664-675.
3. Bishop, D.H.L. and Shope, R.E. (1979) Comprehensive Virology, 14, 1-156.
4. Dahlberg, J.E., Obijeski, J.F. and Korb, J. (1977) J. Virol., 22, 203-209.
5. Hewlett, M.J., Pettersson, R.F. and Baltimore, D. (1977) J. Virol., 21, 1085-1093.
6. Samso, A., Bouloy, M. and Hannoun, C. (1975) C.R. Acad. Sci. Ser. D., 280, 779-782.
7. Obijeski, J.F., Bishop, D.H.L., Murphy, F.A. and Palmer, E.L. (1976) J. Virol., 19, 985-997.
8. England, T.E. and Uhlenbeck, O.G. (1978) Nature, 275, 560-561.
9. Skehel, J.J.and Hay, A.J. (1978) Nuc. Ac. Res., 5, 1207-1219.
10. Ralph, R.K. and Bellamy, A.R. (1964) Biochim. Biophys. Acta., 87, 9-16.
11. Donis-Keller, H., Maxam, A. and Gilbert, W. (1977) Nuc. Ac. Res., 4, 2527-2538.
12. Peattie, D.A. (1979) Proc. Nat. Acad. Sci., 76, 1760-1764.
13. Dasgupta, R. and Kaesberg, P. (1977) Proc. Nat. Acad. Sci., 74, 4900-4904.
14. Gunn, M.R. and Symons, R.H. (1980) Febs Letts., 109, 145-150.
15. Li, J.K.K., Keene, J.D., Scheible, P. and Joklik, W.K. (1980) in Animal Virus Genetics 9th ICN-UCLA symposium. Fields, B. and Jaenisch, R., eds. Alan R. Liss Inc., New York.
16. Keene, J.D., Schubert, M. and Lazzarini, R.A. (1979) J. Virol., 32, 167-174.
17. Kolakofsky, D., Personal communication.
18. Blobel, G. and Dobberstein, B. (1975) J. Cell. Biol., 67, 835-851.
19. Rommelaere, J., Donis-Keller, H. and Hopkins, N. (1979) Cell, 16, 43-50.