Volume 9 Number 8 1981 Nucleic Acids Research

Sequence of DNA complementary to a small RNA segment of influenza virus A/NT/60/68

Bernard A.Moss* and George G.Brownlee

School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, UK

Received 4 March 1981

ABSTRACT

A small RNA segment from the influenza virus strain A/NT/60/68 (H3N2) was converted to cDNA and then to double-stranded DNA using synthetic oligodeoxynucleotide primers. The double-stranded form was cloned into the bacteriophage M13mp7. Clones yielding single-strand recombinant templates in opposite orientation were sequenced by the Sanger dideoxynucleotide chain termination technique. The small viral RNA was 422 nucleotides long and the evidence indicated that it was formed by internal deletion of segment 3. It also contained sequences homologous to segment 1.

INTRODUCTION

Influenza A virus normally has a genome composed of eight unique negative-stranded RNA segments ranging in size from about 900 to 2,400 nucleotides [1]. In addition to these eight components some influenza viruses contain extra RNA molecules which are mostly shorter than 1000 nucleotides (reviewed in 2). Such small RNA segments are apparently associated with defective von Magnus particles of influenza virus and interfere with viral replication [2-4]. Current evidence reveals an inverse relationship between the three viral polymerase genes segments 1-3 and the appearance of these small RNA molecules [2,5-8]. Moreover, it has been shown by oligonucleotide mapping and partial 3'- and 5'- terminal sequence analyses that the small RNA molecules are derived from these polymerase genes by internal deletion [2,6-11], although the mechanism is uncertain.

Previous work with the Hong Kong strain A/NT/60/68 consistently showed two extra RNA components which were designated 9 and 10 [1]. Their electrophoretic mobilities suggested lengths of about 400 to 500 nucleotides with sizes similar to the bands 9 and 10 observed by Skehel and Hay [11] for fowl plague virus. In A/NT/60/68, band 10 was the more prominent RNA component. We have now sequenced this small RNA segment by the strategy of preparing double-stranded cDNA copies with synthetic primers [12], cloning

into the bacteriophage M13mp7 [13,14] and sequencing by the Sanger dideoxy-nucleotide chain-termination method [15].

MATERIALS AND METHODS

Preparation of cDNA

The influenza virus, A/NT/60/68 (H3N2), a field strain antigenically equivalent to the first 1968 Hong Kong isolate [16], was propagated in embryonated hen eggs and purified as described previously [17]. The RNA was extracted [18] and individual segments separated on 2.8% polyacrylamide gels [1]. Band 10 RNA was isolated [18] and sequenced using a synthetic deoxydodecanucleotide primer [20,21]. Alternatively, single-stranded cDNA was prepared from total RNA with reverse transcriptase and dodecanucleotide prior to fractionation [12]. Band 10 cDNA was isolated, then back-copied for 1 hour at 23°C with the Klenow fragment of DNA polymerase I (Boehringer) and a synthetic deoxy-tridecanucleotide primer [12] to form double-stranded DNA. This dsDNA was then phosphorylated with T4 polynucleotide kinase (Boehringer) [12].

Cloning of double-stranded DNA

The cloning vector was M13mp7, a derivative of the DNA phage M13mp2 [13,14]. The replicative form (RF) was cut at the HindII site and phosphatase treated before blunt-end ligation of the phosphorylated dsDNA at 23°C for 20 hours with T4 DNA ligase [12]. The ligated material was used in transformations of E.coli strain JM101 [13,23] and white recombinant plaques harbouring phage with band 10 dsDNA inserts were grown and processed [23].

Nucleotide sequence analysis

The portion of the single stranded M13 template representing the inserted DNA was sequenced by the dideoxynucleotide chain termination method [15]. A chemically synthesised 17-long deoxynucleotide primer (M.J. Gait, unpublished and [24]) which hybridizes next to the insertion site, was used in conjunction with DNA polymerase I (Klenow fragment) to sequence through the band 10 DNA insert [12,23-25]. Two clones (C1 and C3) in opposite orientation provided the evidence for the 422 nucleotide sequence. C3, which gave the positive sense copy, started with the dodecamer primer sequence and, apart from only two difficult regions, the sequencing gels could be read clearly for 366 bases. C1 (negative sense) was used to complete the sequence and resolve the two uncertainties. This clone began with the

tridecamer primer d(AGTAGAAACAAGG) and could be read quite satisfactorily for 275 bases, which gave an overlap of 221 bases containing all the uncertain residues needing to be checked. The non-overlapping regions did not give any ambiguities.

RESULTS

Preliminary studies on band 10 of A/NT/60/68 using direct RNA sequencing methods [36] showed the presence of the 13-long sequence common to the eight standard virion RNA segments at its 5' terminus. Furthermore band 10 contained the 12-long common sequence at its 3' terminus as the specific 3' terminal oligonucleotide, d(AGCAAAAGCAGG), was able to prime the synthesis of cDNA in a dideoxy sequencing reaction with reverse transcriptase [19] (results not shown).

Fig 1 shows the DNA sequence deduced for the positive strand of band 10 derived by sequence analysis of two full length cDNA clones inserted in opposite orientations in the vector M13mp7 (see methods). A comparison of fig 1 with the limited sequence available for segments 1-3 of fowl plague virus shows that the two end regions of band 10 are derived from segment 3. Specifically there are only 6 differences covering residues 1-56 and 3 differences in residues 350-422 (fig 1), this being the full extent of the information available for segment 3 of fowl plague virus. The difference at residue 4 (fig 1) may not be real as this residue (A) in A/NT/60/68 was derived directly from the dodecamer primer and was not independently checked by direct RNA sequencing. The sequence conservation at the 3' end is extensive, suggesting this region is particularly important, possibly in viral replication.

Fig 2 shows that band 10 is also related to a region of segment 1 of fowl plague. Specifically residues 111-138 of band 10 and residues 21 to 48 of segment 1 are homologous with only 4 mismatches. This homology, by itself, might be considered only marginally significant but it is extended by 40 residues in further studies (Fields, Winter & Brownlee, unpublished) of segment 1 of influenza A/PR/8/34.

The three potential amino acid reading frames for band 10 (fig 1) show that no phase is open covering a greater part of the length of the molecule. This is also true for the negative strand. Therefore, if a protein is made from this mRNA, at least two separate splicing events would have to occur to give an open translation frame covering most of the template strand. While there are some potential splice consensus sequences (e.g.

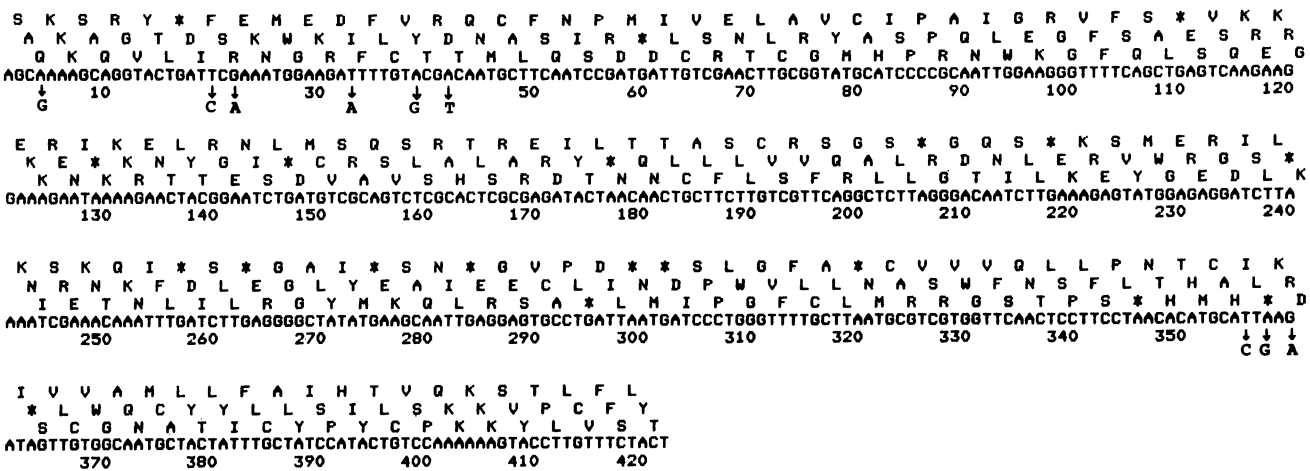

Figure 1. The nucleotide sequence of band 10 of A/NT/60/68 written as a DNA sequence in the mRNA sense. The predicted amino acid sequences for the three reading frames are depicted by the one letter amino acid code and * marks terminators. Differences between the 5'- and 3'- terminal regions of this sequence and the first 56 residues and last 72 residues of segment 3 of fowl plague are indicated by arrows.

```
                                111         120         130
                                 .           .           .
A/NT/60/68  Band 10              AGTCAAGAAGGAAAGAATAAAAGAACTA
                                  x     x x    x
Fowl Plague Segment 1            ATTCAATATGGAGAGAATAAAAGAACTA
                                 .           .           .
                                21          30          40
```

Figure 2. A comparison of a region of band 10 of A/NT/60/68 and segment 1 of fowl plague virus. x marks mismatches.

around residues 77 and 107) there do not appear to be enough [37] to yield any protein of moderate size. Therefore, the likely consequence of band 10 in A/NT/60/68 would be the synthesis of a 28-long peptide, ME . . . FS* which is very unlikely to be functional.

DISCUSSION

The complete sequence of band 10 of A/NT/60/68 derived here and its comparison with segment 3 of fowl plague clearly shows that its two ends arose from segment 3 by a process of internal deletion. This confirms work of others based on much more limited sequence analysis [2] showing that the defective interfering influenza RNA species retain the two common ends characteristic of all influenza segments. Quite unexpectedly we found that an internal region of band 10 appears to derive from a different gene, viz. segment 1. We deduce therefore that band 10 is either a "mosaic" derived from both segments 1 and 3 or that segments 1 and 3 show extensive nucleotide sequence homology to one another, thus invalidating our conclusion that an internal region of band 10 is uniquely homologous to segment 1. Clearly to resolve these possibilities, and the mechanism of origin of band 10, it will be necessary to sequence both segments 1 and 3 from the homologous viral strain A/NT/60/68. In advance of this knowledge we might speculate that internal deletions could arise by abnormal splicing of the influenza genes either by the normal host enzyme [37] or by an influenza coded protein splicing activity [38] during viral replication. A knowledge of the sequence at the exact splice point could resolve which of these mechanisms is operating.

It is likely that similar internal deletions are responsible for the production of defective interfering RNAs of positive strand viruses such as Semliki Forest virus [29,30], Sindbis virus [31] and poliovirus [32]. Thus an understanding of the phenomenon in influenza might be applicable to these viruses as well. By contrast, a different mechanism seems to account for

the generation of defective interfering RNA in Sendai virus and Vesicular Stomatitis virus [32-35].

ACKNOWLEDGEMENTS

We are grateful to Stan Fields and David Bentley for T4 DNA ligase, to Andrew Caton for advice and to Barbara Paxman for typing the manuscript. GGB gratefully acknowledges financial support from the M.R.C. and a gift of reverse transcriptase from Dr Joseph Beard.

* Permanent address: C.S.I.R.O. Molecular and Cellular Biology Unit, P.O. Box 184, North Ryde, N.S.W., Australia, 2113.

REFERENCES

1. Sleigh, M.J., Both, G.W. and Brownlee, G.G. (1979) Nucleic Acids Res. 6, 1309-1321.
2. Nayak, D.P. (1980) Ann. Rev. Microbiol. 34, 619-644.
3. Von Magnus, P. (1954) Adv. Virus Res. 2, 59-78.
4. Janda, J.M. and Nayak, D.P. (1979) J. Virol. 32, 697-702.
5. Crumpton, W.M., Dimmock, N.J., Minor, P.D. and Avery, R.J. (1978) Virology 90, 370-373.
6. Nakajima, K., Ueda, M. and Sugiura, A. (1979) J. Virol. 29, 1142-1148.
7. Ueda, M., Nakajima, K. and Sugiura, A. (1980) J. Virol. 34, 1-8.
8. Janda, J.M., Davis, A.R., Nayak, D.P. and De, B.K. (1979) Virology 95, 48-58.
9. Davis, A.R. and Nayak, D.P. (1979) Proc. Natl. Acad. Sci. USA 76, 3092-3096.
10. Davis, A.R., Hiti, A.L. and Nayak, D.P. (1980) Proc. Natl. Acad. Sci. USA. 77, 215-219.
11. Skehel, J.J. and Hay, A.J. (1978) Nucl. Acids Res. 5, 1207-1219.
12. Winter, G., Fields, S., Gait, M.J. and Brownlee, G.G. (1981) Nucleic Acids Res. 9, 237-245.
13. Messing, J. (1979) Recombinant DNA Technical Bulletin 2, 43-48.
14. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
15. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA. 74, 5463-5467.
16. Fazekas de St. Groth, S. (1969) Bull. Wld. Hlth. Org. 41, 651-657.
17. Moss, B.A. and Underwood, P.A. (1978) in "The influenza virus haemagglutinin" eds. W.G. Laver, H. Bachmayer and R. Weil, Topics in Infectious Diseases, Vol. 3, pp. 145-166. Springer-Verlag, Wien New York.
18. Both, G.W. and Air, G.M. (1979) Eur. J. Biochem. 96, 363-372.
19. Hamlyn, P.H., Brownlee, G.G., Cheng, C.C., Gait, M.J. and Milstein, C. (1978) Cell 15, 1067-1075.
20. Air, G.M. (1980) in "Structure and variation in influenza virus" eds. W.G. Laver and G.M. Air, Developments in Cell Biology, Vol. 5, pp. 135-146, Elsevier/North-Holland, Amsterdam.
21. Cummings, I. and Salser, W. (1980) as in [20] pp. 147-155.
22. Maxam, A. and Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA. 74; 560-564.
23. Winter, G. and Fields, S. (1980) Nucleic Acids Res. 8, 1965-1974.

24. Anderson, S., Gait, M.J., Mayol, L. and Young, I.G. (1980) Nucleic Acids Res. 8, 1731-1743.
25. Heidecker, G., Messing, J. and Gronenborn, B. (1980) Gene 10, 69-73.
26. Air, G.M. (1979) Virology 97, 468-472.
27. Robertson, J.S. (1979) Nucleic Acids Res. 6, 3745-3757.
28. Desselberger, U., Racaniello, V.R., Zazra, J.J. and Palese, P. (1980) Gene 8, 315-328.
29. Kennedy, S.I.T. (1976) J. Mol. Biol. 108, 491-511.
30. Stark, C. and Kennedy, S.I.T. (1978) Virology 89, 285-299.
31. Dohner, D., Monroe, S., Weiss, B. and Schlesinger, S. (1979) J. Virol. 29, 794-798.
32. Leppert, M., Kort, L. and Kolakofsky, D. (1977) Cell 12, 539-552.
33. Schubert, M., Keene, J.D., Lazzarini, R.A. and Emerson, S.U. (1978) Cell 15, 103-112.
34. Keene, J.D., Schubert, M., Lazzarini, R.A. and Rosenberg, M. (1978) Proc. Natl. Acad. Sci. USA. 75, 3225-3229.
35. Semler, B.L., Perrault, J., Abelson, J. and Holland, J.J. (1978) Proc. Natl. Acad. Sci. USA. 75, 4704-4708.
36. Simoncsits, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilley, H. (1977) Nature 269, 833-836.
37. Catterall, J.F., O'Malley, B.W., Robertson, M.A., Staden, R., Tanaka, Y. and Brownlee, G.G. (1978) Nature, 257, 510-513.
38. Robertson, H.D., Dickson, E., Plotch, S.J. and Krug, R.M. (1980) Nucleic Acids Res. 8, 925-942.