Volume 7 Number 6 1979 Nucleic Acids Research

New procedure for the production of influenza virus-specific double-stranded DNA's

Andrew J.Caton and James S.Robertson

Division of Virology, Department of Pathology, University of Cambridge, Laboratories Block, Addenbrooke's Hospital, Hills Road, Cambridge, UK

Received 11 September 1979

ABSTRACT

A novel technique is described for the production of pure, full-length influenza virus ds DNA's corresponding to each segment of the influenza virus genome, and suitable for molecular cloning and restriction endonuclease mapping. The method involves the synthesis of DNA complementary to both virion (negative strand) and messenger (positive strand) RNA, gel purification and annealing. By avoiding the use of SI nuclease, which often removes the terminal regions of DNA duplexes, the method allows transcription of the total sequence information of influenza virion and messenger RNA's into a ds DNA form.

INTRODUCTION

For sequence analysis of RNA from a variety of sources, the production of corresponding double-stranded DNA's (ds DNA's) is of great value, since this allows analysis of the primary structure either by restriction endonuclease mapping, or by total nucleotide sequence determination by established rapid DNA sequencing techniques after molecular cloning and amplification in bacteria. In recent years many reports have appeared describing methods by which ds DNA's can be produced from an RNA molecule, most of these involving β-globin or ovalbumin mRNA as the RNA source (1-5). Typically, the RNA is used as a template for the synthesis of a single-stranded DNA copy using the enzyme reverse transcriptase. It has generally been found that the 3' end of these cDNA products fold back upon themselves, forming a hairpin loop structure which allows the 3' end to act as a primer for the synthesis of a second strand of DNA complementary to the first using reverse transcriptase or polymerase I. The resulting ds DNA molecule is covalently closed at one end and can be converted to an open duplex by the use of SI nuclease. The production of influenza virus ds DNA's by this method, using influenza virus genome RNA polyadenylated *in vitro* as the initial template, has been described (6). One disadvantage of this method is that the use of SI nuclease

© Information Retrieval Limited 1 Falconberg Court London W1V 5FG England

results in the loss of a portion of the terminal sequences of the ds DNA.

The influenza virus genome is comprised of eight negative-stranded RNA segments (7), each of which is transcribed *in vivo* into corresponding discrete mRNA species (8). These mRNA species are not exact transcripts of the virion RNA segments but contain additional host derived sequences of approximately 15 nucleotides in length at the 5' end (9), and at the 3' end lack 20-30 nucleotides complementary to the 5' end of virion RNA (10). This paper describes the synthesis of single-stranded DNA copies of both classes of RNA and their subsequent hybridisation to produce ds DNA's which represent the structure and properties of the genome segments of influenza virus, and can be used for sequence analysis of both the influenza virion RNA segments and their *in vivo* mRNA transcripts.

MATERIALS AND METHODS

Materials

ATP:RNA adenyltransferase was purified as described by Sippel (11) from *E. coli* MRE 600 cells (MRE, Porton Down). Reverse transcriptase purified from avian myeloblastosis virus (Lot No. G-179) was supplied by the Office of Program Resources and Logistics, Virology Cancer Program, Viral Oncology, Division of Cancer Cause and Prevention, National Cancer Institute, Bethesda, Maryland. Calf intestinal phosphatase and polynucleotide kinase were obtained from Boehringer, restriction endonuclease Hind III from Miles Research Ltd., SI nuclease from Sigma, and Pronase from Calbiochem.

The primers $p(dT)_{12-18}$ and $p(dT)_8{}^{dA}$ were obtained from P-L Biochemicals, and cycloheximide, bovine serum albumin and phage λ DNA were obtained from Sigma. $[\alpha\text{-}^{32}P]$ dTTP (>350 Ci/mmol), $[^{3}H]$ dTTP (15-30 Ci/mmol), and $[\gamma\text{-}^{32}P]$ ATP (>2000 Ci/mmol) were all obtained from the Radiochemical Centre, Amersham.

Purification and polyadenylation of influenza virion RNA

Fowl plague virus (Rostock strain) was grown in fertile hen eggs and purified by sucrose gradient velocity centrifugation as described by Inglis *et al.* (12). Purified virions were disrupted in 0.1M tris-HCl, pH 7.5, 0.05 M NaCl, 0.01M EDTA, 0.5% SDS and digested with 0.5 mg/ml Pronase for 1 h at 37°C. Total RNA was extracted with phenol:chloroform (1:1) and the aqueous layer precipitated twice with ethanol. The purified virus RNA was resuspended in H_2O to a final concentration of 0.5 mg/ml, as determined by UV absorbance (1.0 $OD_{260} \simeq 40$ μg/ml).

Virus RNA was polyadenylated at the 3' end using ATP:RNA adenyltrans-

ferase (11,13). The reaction mixture for RNA polyadenylation contained 50 mM tris-HCl pH 7.9, 10 mM $MgCl_2$, 2.5 mM $MnCl_2$, 0.2M NaCl, 0.5 mg/ml bovine serum albumin, 0.1 mM ATP, 200 µg/ml virus RNA and $^1/10$ volume of enzyme. After incubation at 37° for 10 min, the reaction was terminated by extraction with an equal volume of phenol:chloroform (1:1). The RNA was precipitated twice with ethanol, resuspended in H_2O to a final concentration of 0.5 mg/ml, and stored at -70°C.

Isolation of influenza virus-infected cell mRNA

Primary chick embryo fibroblasts were infected with influenza virus, harvested and the cytoplasmic RNA extracted as described by Inglis *et al.* (14), except that the cells were treated with cycloheximide (100 µg/ml) from 2.5 h post infection (15) and harvested 4.5 h post infection. Poly(A)-containing RNA (mRNA) was purified by oligo(dT)-cellulose chromatography essentially according to Glass *et al.* (16), except that the RNA binding buffer contained 0.4 M LiCl instead of 0.4M NaCl, and the RNA bound to the column was washed with 0.1M LiCl, 10 mM tris-HCl pH 7.4, 1 mM EDTA, and 0.1% SDS before elution. Following elution, the mRNA was precipitated twice with ethanol, dissolved in H_2O to a final concentration of 0.5 mg/ml, and stored at -70°C.

Synthesis of cDNA

Single-stranded, complementary DNA was synthesised in a reaction mixture containing 50 mM tris-HCl pH 8.3, 10 mM $MgCl_2$, 20 mM 2-mercaptoethanol, 0.4 mM dATP, dCTP, dGTP, 50 µM dTTP (α-^{32}P or ^3H -labelled), 100 µg/ml primer, 60 µg/ml RNA (mRNA or polyadenylated virion RNA) and 150 units/ml of reverse transcriptase. Incubation was at 46° for 1 hour. The primer used was different for the two classes of RNA; for mRNA, the reaction was primed with $p(dT)_{12-18}$, but for the polyadenylated virion RNA $p(dT)_8dA$ was used. The latter will hybridise specifically to the 5' end of the poly(A) tail since the 3' terminal nucleotide of influenza virus genome RNA is known to be U (17). Following incubation, the solution was made to 0.01M EDTA, 0.1M NaOH, incubated at 60°C for 30 min, neutralised with HCl, 5 µg tRNA carrier added and the DNA recovered by ethanol precipitation.

Gel electrophoresis

RNA and single-stranded DNA were electrophoresed for 15 h at 300V on 2.5% polyacrylamide gels (40 x 20 x 0.15 cm) containing 90 mM tris-borate, 2.5 mM EDTA and 7 M urea (pH 8.3). The wet gel was exposed to Kodak X-Omat H film at -70°C with an intensifying screen.

Double-stranded DNA's were analysed by electrophoresis for 15 h at 150V on 1.5% flat-bed agarose gels (24 x 18 x 0.5 cm) containing 40 mM tris, 8 mM Na acetate, and 0.4 mM EDTA, pH 7.9. Agarose gels were dried down under vacuum before exposing to Kodak X-Omat H film.

Production of ds DNA

Individual cDNA's were eluted from gel slices by soaking for 16 h in 0.5 M NH_4 acetate, 5 mM Mg acetate, 1% SDS, and the DNA recovered by ethanol precipitation with 40 µg tRNA carrier (18). Corresponding cDNA's were dissolved in 10 mM tris-HCl pH 7.6, and prior to mixing, boiled for 30 seconds and cooled on ice. They were then mixed together in a volume which ensured a final cDNA concentration of greater than 1 µg/ml, and an equal volume of 20 mM tris HCl pH 7.6, 2 M NaCl added. The mixture was incubated at 70°C for 16 h in a sealed glass capillary and the ds DNA recovered by ethanol precipitation.

End-labelling of RNA and preparation of DNA markers

The 5' termini of RNA were dephosphorylated by incubation at 37° for 30 min with calf intestinal phosphatase purified as described by Efstratiadis *et al.* (19). The phosphatase was removed by phenol extraction, and following ethanol precipitation, the RNA was 5' end-labelled in a mixture containing polynucleotide kinase and $[\gamma\text{-}^{32}P]$ ATP and incubated at 37° for 30 min (17).

For the production of DNA markers, phage λ DNA was digested using Hind III under the conditions recommended by the suppliers. The restricted DNA was 5' end labelled using polynucleotide kinase and $[\gamma\text{-}^{32}P]$ ATP as described by Miller *et al.* (20). The molecular weights of the restricted fragments are from Murray & Murray (21).

SI nuclease treatment

SI nuclease treatment was carried out by incubation at 37°C for 30 min in reaction mixtures containing 30 mM Na acetate pH 4.6, 50 mM NaCl, 1 mM $ZnSO_4$, DNA and SI nuclease as indicated.

RESULTS

Polyadenylation of influenza virion RNA

Influenza virion RNA was polyadenylated *in vitro* at the 3' end using the enzyme ATP:RNA adenyltransferase (11,13) prior to labelling of the 5' end, and electrophoresed in parallel with 5'-labelled virion RNA in a polyacrylamide-urea gel (Fig.1, tracks a & b). The eight RNA segments which range in

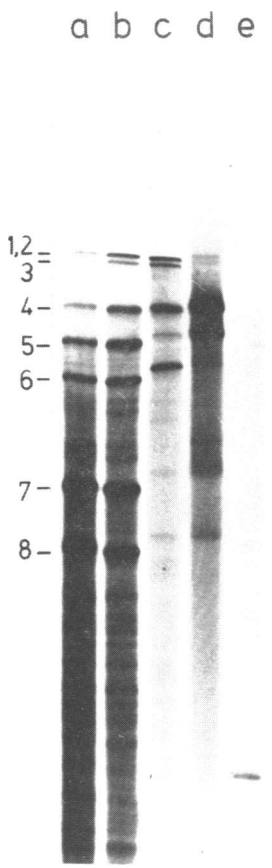

Fig. 1. Polyacrylamide gel electrophoresis of influenza virus RNA and influenza virus-specified cDNA products.

(a) 5' end-labelled virion RNA polyadenylated *in vitro*.

(b) 5' end-labelled virion RNA.

(c) internally ^{32}P-labelled cDNA synthesised from 1 µg virion RNA polyadenylated *in vitro*.

(d) internally ^{32}P-labelled cDNA synthesised from 1 µg mRNA.

(e) internally ^{32}P-labelled cDNA synthesised from 1 µg non-polyadenylated virion RNA.

size from 900 to 2400 nucleotides in length can be identified, with segments 1 and 2 comigrating in this gel system (7,22). Polyadenylation resulted in a slight decreased mobility of each segment, and degradation of the virion RNA was negligible. Evidence from the polyadenylation of virion RNA with $[^{3}H]$ ATP as substrate indicated that this shift in mobility was due to the addition of adenylate residues, and that under these conditions approximately 400 pmol of AMP are added per µg virion RNA (data not shown).

Reverse transcription of influenza virion and messenger RNA's

Influenza virion RNA polyadenylated *in vitro*, and mRNA extracted from influenza virus-infected cells were used as templates for cDNA synthesis by reverse transcriptase, following which the RNA was destroyed by hydrolysis with NaOH. Analysis of the products by gel electrophoresis indicated 8

major species which migrated more slowly than their corresponding templates (Fig.1, tracks b, c & d). The similarity between the patterns obtained on electrophoresis of the cDNA products produced from both classes of RNA and the electrophoretic pattern of virion RNA indicated full length transcription of all 8 RNA segments as has been observed elsewhere (22).

The background of minor cDNA products was presumably produced due to premature termination of reverse transcription and transcription of fragments of the templates. In addition, the pattern of cDNA synthesised in response to virion RNA indicated that not all templates were transcribed with the same efficiency, since although the individual *in vitro* polyadenylated virion RNA segments were present in equimolar quantities (Fig.1, track a, and ref.7), the corresponding cDNA's were not. This differential transcription was not related to the degree of polyadenylation or the amount of reverse transcriptase used in the reaction (data not shown), and could represent uneven priming or the different secondary structure of individual RNA segments. DNA copies of influenza virus mRNA were not synthesised in equimolar amounts, since the cDNA's corresponding to segments 4 & 5 were overrepresented, and the cDNA corresponding to segment 6 was underrepresented relative to other segments. This uneven representation may be due in part to the relative transcriptional ability of each mRNA template as shown above with the virion RNA template, and in part to the relative abundance of each message. It has been well established by a variety of techniques, including cell free translation (23,24), hybridisation studies (25), and the direct electrophoretic analysis of labelled mRNA's following hybridisation with unlabelled virion RNA (8), that there is an unequal representation of influenza virus mRNA's in the infected cell, and this overall pattern of representation of each mRNA is reflected in the pattern of the cDNA products shown here. It is also clear from the pattern of the cDNA synthesised that the majority of the mRNA in the infected cell is influenza virus-specified.

In the absence of polyadenylation, influenza virion RNA does not direct the synthesis of corresponding full-length cDNA's (Fig.1, track e). A prominent low molecular weight product and some additional minor products can be seen which occur either by hybridisation of the primer to internal sites in the template, or by self-priming of the reaction due to the presence of hairpin loops at the 3' end of the virion RNA. This clearly shows that polyadenylation and priming are necessary steps for the production of full length cDNA products.

The production of influenza virus ds DNA's

For the production of influenza virus ds DNA, the two classes of influenza virus cDNA, i.e. cDNA to virion RNA and cDNA to mRNA, were annealed together, and on analysis by electrophoresis in a non-denaturing agarose gel, bands were observed which corresponded in molecular weight to that of influenza virus ds DNA (Fig.2, tracks a & e). There was, however, an exceedingly high background caused by the hybridisation of incomplete cDNA transcripts. Treatment with SI nuclease reduced but did not remove this background by degrading the single-stranded regions of these partial hybrids. Indeed, it was observed that this treatment led to a reduction in size of the ds DNA's, and that the extent of nucleotide loss from the ds DNA's increased with the amount of nuclease used (Fig.2, tracks b, c & d).

To overcome the problem of contamination with partial hybrids and the sequence loss from SI nuclease treatment, selection for full length transcripts of individual virion RNA segments and messenger RNA's was made at the level of single-stranded cDNA synthesis. Labelled cDNA was synthesised from virion RNA, and non-labelled cDNA from messenger RNA, and the individual cDNA segments excised and eluted from a polyacrylamide-urea gel as in Fig.1.

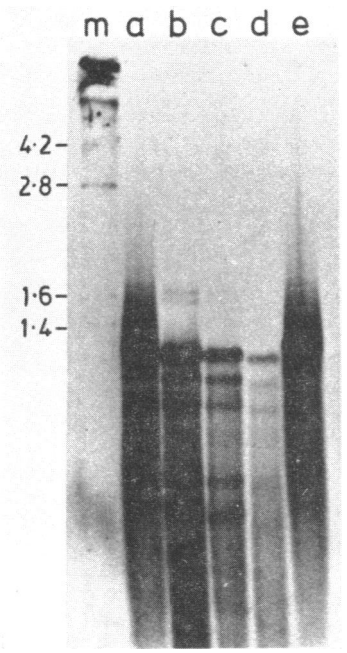

Fig. 2. Hybridisation of unfractionated cDNA synthesised from mRNA and virion RNA polyadenylated *in vitro*. Internally ^{32}P-labelled cDNA synthesised from virion RNA polyadenylated *in vitro* was annealed together with a 2-4 fold excess of unlabelled cDNA synthesised from mRNA and samples analysed on a non-denaturing agarose gel (a&e). Further samples, estimated to contain approximately 2 µg of total nucleic acid, were treated with 0.2(b), 2(c) and 20(d) units of SI nuclease. Molecular weights (x 10^{-6}) are derived from the Hind-IIIλ markers (m).

Analysis of the individual labelled cDNA segments on a non-denaturing agarose gel indicated only a trace of contamination by incomplete segments. Also, some of the cDNA segments (tracks 5s, 7s & 8s) migrate as doublets in this non-denaturing gel system. When individual labelled cDNA segments produced from virion RNA were annealed together with an excess of the corresponding non-labelled cDNA segments derived from mRNA, the resulting ds DNA's showed a considerably lower mobility than the non-annealed segments (Fig.3, tracks labelled d), and the overall pattern of the individual ds DNA's showed the characteristic pattern obtained on electrophoresis of influenza virion RNA and both classes of single-stranded cDNA products (Fig.1).

It is clear from comparison with the mobility of the molecular weight markers that these species represent complete ds DNA copies of each of the influenza virus genome segments. Sleigh *et al.* (22) recently determined the chain length of influenza virus genome segments 1 & 2 to be 2390 nucleotides; thus ds DNA's corresponding to these segments will have a molecular weight of 1.6×10^6 daltons. Figure 3 (track 1/2, d) shows that these ds DNA's comigrate exactly with the 1.6×10^6 MW marker. When equimolar amounts of corresponding cDNA segments are annealed together, the entire sequence information is converted to a ds DNA form (data not shown) giving rise to individual ds DNA's with no detectable contamination by non-annealed sequences, partial duplexes, or ds DNA's derived from any other segment.

DISCUSSION

This paper describes a new method for the production of influenza virus ds DNA's which makes use of the fact that the virion RNA segments are of opposite polarity to the virus mRNA segments. By annealing together the single-stranded DNA copies of virion RNA and messenger RNA made using reverse transcriptase, we have been able to produce individual ds DNA's corresponding to each of the influenza virus genome segments. These individual ds DNA's are free of contamination by ds DNA's derived from any other segment and of sufficient chemical purity to act as substrates for restriction endonucleases (data not shown). This approach has two main advantages over a previously described method for the production of influenza virus ds DNA's (6). First, the process of duplex formation following polyadenylation of the virion RNA requires only one enzymatic step, does not involve treatment with SI nuclease, and thus allows the incorporation of the full length of the cDNA into the DNA duplex. Second, by incorporating the sequence informa-

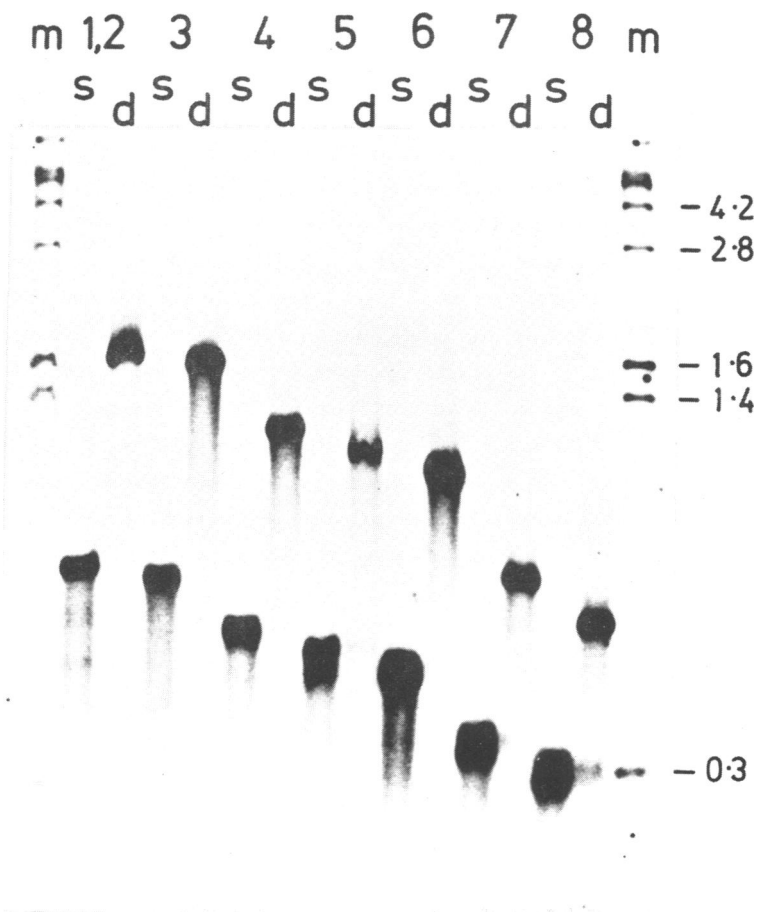

Fig. 3. Production of individual influenza virus ds DNA's. ^{32}P-labelled cDNA synthesised from virion RNA polyadenylated *in vitro* was fractionated on a polyacrylamide-urea gel and the individual species extracted. These single-stranded cDNA's were further analysed on a non-denaturing agarose gel (tracks labelled (s)). In addition, double-stranded DNA's were produced by hybridisation of each individual cDNA with a 2-4 fold excess of the corresponding non-labelled cDNA synthesised from messenger RNA (tracks labelled (d)). Molecular weights (x 10^{-6}) are derived from the Hind III-λ markers (m). Numbers refer to the segment from which the cDNA's are derived.

tion of both virion RNA and mRNA into the DNA duplex, properties of both classes of RNA can be studied directly by DNA sequence analysis.

The use of SI nuclease results in the loss of terminal sequence information and it is apparent from Fig.2 that increasing the ratio of SI nuclease

to ds DNA increases the extent to which the DNA duplexes are degraded. Our approach eliminates the necessity for the use of SI nuclease and allows selection of full length single-stranded cDNA's, ensuring that the total sequence information of both the virion RNA and mRNA are present in the final ds DNA. Although it is apparent in Fig.1 that some incomplete cDNA transcripts are produced, the similarity between the electrophoretic pattern of the virion RNA segments and the major cDNA products transcribed from both mRNA and virion RNA polyadenylated *in vitro* strongly suggests that these products represent full length copies of the templates, and are not the result of premature termination of transcription of one of the larger templates. Recent reports on the reverse transcription of influenza virion RNA polyadenylated *in vitro* also show this characteristic migrational pattern of the cDNA products (6,22), and in other systems, high yields of full length transcripts have been obtained using similar conditions to those used here (26-29). This series of observations indicates that the major cDNA products shown in Fig.1 represent full length transcripts of the influenza virion and messenger RNA templates.

The presence of the sequence information of both virion RNA and messenger RNA in the DNA duplex is particularly important in the study of influenza virus, since these two classes of RNA are not exact complements of each other. When influenza virus mRNA is synthesised, modifications occur at both the 3' and 5' end. In the first instance, influenza virus mRNA is known to lack 20-30 nucleotides complementary to the 5' end of virion RNA (10), and direct analysis of these ds DNA's will allow determination of the precise location at which termination of *in vivo* mRNA transcription occurs. Secondly, influenza virus mRNA contains additional, host-derived sequences of approximately 15 nucleotides in length at the 5' end (9), and these will be rendered accessible either to direct DNA sequence analysis or to total sequence determination following molecular cloning. On repair of these ds DNA's following insertion into molecular cloning vectors by 3' end tailing, no influenza virus-specific sequences will be lost. Clones isolated from ds DNA's produced using this approach will contain the total sequence of influenza virion RNA and the additional host derived 5' mRNA primer sequence. Thus, in addition to allowing the simple production of influenza virus ds DNA's suitable for molecular cloning, the method also provides a technique for the isolation and sequence analysis of the putative host-derived 5' mRNA primer (9).

ACKNOWLEDGEMENTS

This work was supported by a grant from the Medical Research Council. A.J.C. is the recipient of a S.R.C. CASE award. We thank Dr. B.W.J. Mahy for helpful discussion and encouragement.

REFERENCES

1. Higuchi, R., Paddock, G.V., Wall, R. and Salser, W. (1976). *Proc. Natl. Acad. Sci. U.S.A.* 73, 3146-3150.
2. Efstratiadis, A., Kafatos, F.C., Maxam, A.M. and Maniatis, T. (1976). *Cell* 7, 279-288.
3. Rougeon, F. and Mach, B. (1976). *Proc. Natl. Acad. Sci. U.S.A.* 73, 3418-3422.
4. Monahan, J.J., McReynolds, L.A. and O'Malley, B.W. (1976). *J. Biol. Chem.* 251, 7355-7362.
5. Wilson, J.T., Forget, B.G., Wilson, L.B. and Weissman, S.M. (1977). *Science* 196, 200-202.
6. Emtage, J.S., Catlin, G.H. and Carey, N.H. (1979). *Nucl. Acids Res.* 6, 1221-1240.
7. McGeoch, D., Fellner, P. and Newton, C. (1976). *Proc. Natl. Acad. Sci. U.S.A.* 73, 3045-3049.
8. Hay, A.J., Lomniczi, B., Bellamy, A.R. and Skehel, J.J. (1977). *Virology* 83, 337-355.
9. Plotch, S.J., Bouloy, M. and Krug, R.M. (1979). *Proc. Natl. Acad. Sci. U.S.A.* 76, 1618-1623.
10. Hay, A.J., Abraham, G., Skehel, J.J., Smith, J.C. and Fellner, P. (1977). *Nucl. Acids Res.* 4, 4197-4209.
11. Sippel, A.E. (1973). *Eur. J. Biochem.* 37, 31-40.
12. Inglis, S.C., Carroll, A.R., Lamb, R.A. and Mahy, B.W.J. (1976). *Virology* 74, 489-503.
13. Sano, H. and Feix, G. (1976). *Eur. J. Biochem.* 71, 577-583.
14. Inglis, S.C., McGeoch, D. and Mahy, B.W.J. (1977). *Virology* 78, 522-536.
15. Etkind, P.R. and Krug, R.M. (1975). *J. Virol.* 16, 1464-1475.
16. Glass, S.E., McGeoch, D. and Barry, R.D. (1975). *J. Virol.* 16, 1435-1443.
17. Lewandowski, L.J., Content, J. and Leppla, S.H. (1971). *J. Virol.* 8, 701-707.
18. Simoncsits, A., Brownlee, G.G., Brown, R.S., Rubin, J.R. and Guilley, H. (1977). *Nature* 269, 833-836.
19. Efstratiadis, A., Vournakis, J.N., Donis-Keller, H., Chaconas, G., Dougall, D.K. and Kafatos, F.C. (1977). *Nucl. Acids Res.* 4, 4165-4174.
20. Miller, J.R., Cartwright, E.M., Brownlee, G.G., Fedoroff, N.V. and Brown, D.D. (1978). *Cell* 13, 717-725.
21. Murray, K. and Murray, N.E. (1975). *J. Mol. Biol.* 98, 551-564.
22. Sleigh, M.J., Both, G.W. and Brownlee, G.G. (1979). *Nucl. Acids Res.* 4, 1309-1321.
23. Inglis, S.C., Conti, G. and Mahy, B.W.J. (1978). In *"Negative Strand Viruses and the Host Cell"*, (B.W.J. Mahy and R.D. Barry, eds.), pp.239-248. Academic Press, New York/London.
24. Inglis, S.C. and Mahy, B.W.J. (1979). *Virology* 95, 154-164.
25. Barrett, T., Wolstenholme, A.J. and Mahy, B.W.J. (1979). *Virology* 98, 211-225.
26. Efstratiadis, A., Maniatis, T., Kafatos, F.C., Jeffrey, A. and Vournakis, J.N. (1975). *Cell* 4, 367-378.

27. Chang, J.C., Temple, G.F., Poon, R., Neumann, K.H. and Kan, Y.W. (1977). *Proc. Natl. Acad. Sci. U.S.A.* 74, 5145-5149.
28. Friedman, E.Y. and Rosbash, M. (1977). *Nucl. Acids Res.* 4, 3455-3471.
29. McReynolds, L.A., Catterall, J.F. and O'Malley, B.W. (1977). *Gene* 2, 217-231.