Polyadenylation and reverse transcription of influenza viral RNA

J.S. Emtage, G.H. Catlin and N.H. Carey

Searle Research Laboratories, Lane End Road, High Wycombe, Bucks., HP12 4HL, UK

Received 21 February 1979

ABSTRACT

The polyadenylation of Fowl Plague Viral RNA and of Influenza A/Victoria Viral RNA using E.coli poly(A) polymerase and the subsequent reverse transcription of the polyadenylated species is reported. We have shown that all 8 genome fragments are adenylated and that an average of 25-30 adenylic acid residues per molecule is sufficient for maximal transcription with reverse transcriptase. The cDNA product is 95% sensitive to Sl-nuclease and hybridisation analysis against viral RNA reveals it to be a faithful copy of the RNA. Amongst the transcription products are long, discrete copies of genes 1-8, the lengths of which are comparable with those of the vRNA determined by electrophoresis on formamide acrylamide gels. These single-stranded cDNAs have been further transcribed to form double-stranded products with hair-pin structures at one end. Analysis of this material on native acrylamide gels revealed some DNA bands corresponding to the predicted sizes for genes 4-8.

INTRODUCTION

The Influenza A Viruses are members of the Orthomyxovirus group (1) and have a segmented genome consisting of eight unique single-stranded RNA molecules of molecular weights between $2 \times 10^5$ and $10^6$. These RNAs lack infectivity and must be replicated during infection to produce both viral mRNA and templates for further vRNA synthesis (2). The mRNAs for Influenza Virus proteins have been extensively studied and shown to produce virus-specified polypeptides in cell-free protein synthesising systems (3). Further, the genomes for a number of Influenza A Viruses have now been mapped and the genes for most of the virus-specified proteins identified (4, 5, 6).

Thus, the nature of the genome provides a unique system for studying the relationship between structure and function for eight individual genes and it is already becoming clear that there are regions of similarity, not only among genes of a given strain of virus, but between strains. Skehel and Hay (7) have recently shown the first thirteen nucleotides of the 5'-end of all eight genome RNAs of FPV and of the antigenically unrelated X-31 Influenza
Virus to be the same. As well, they show that the 5' termini of the RNA transcripts produced in vitro are similar for the first twelve nucleotides.

To study the nucleotide structure of the genome further, we have investigated the synthesis of DNA complementary to the vRNA. In this publication, we describe the polyadenylation and transcription of vRNA from two strains of Influenza A Virus, properties of the resulting cDNA and synthetic genes, and make a reassessment of the sizes of the vRNAs themselves.

MATERIALS AND METHODS

Virus

The Rostock strain of Fowl Plague Virus (FPV) was grown in 10-day-old embryonated hens' eggs at 37°C for 18-30 hr. The allantoic fluid was collected and centrifuged at 2,000 rpm for 10 min to remove red cells and other cellular debris. Virus was then harvested from the resulting supernatant by centrifugation at 18,000 rpm for 90 min (IEC B-60, 6 x 250 ml angle rotor). Virus pellets were resuspended in NTE (100 mM NaCl, 10 mM Tris-HCl pH 7.5, 1 mM EDTA) by gentle homogenisation, spun at 5,000 rpm for 4 min (Sorvall) to remove aggregated yolk protein, and the supernatant removed. The virus was finally purified on a 15%-60% w/v linear sucrose gradient in NTE by banding to equilibrium at 24,000 rpm for 3 hr at 4°C. The virus band, appearing about 1/3 from the bottom of the tube, was removed by suction, diluted to twice its volume with NTE, and pelleted at 25,000 rpm for 40 min at 4°C.

Influenza A/Victoria/75 was supplied by Evans Medical Co., Liverpool.

RNA extraction

a) FPV:

The pellet of virus was resuspended in 5 ml NTE, SDS added to 0.2% and the mixture extracted with phenol and chloroform as described (8). The final aqueous phase was made to 200 mM with NaCl and RNA precipitated by the addition of 2.5 volumes of ethanol. Ethanol precipitates were pelleted (11K x 20 min, Sorvall), washed with ethanol, dried in vacuo and dissolved in TE (10 mM Tris-HCl pH 7.5, 1 mM EDTA). RNA was quantitated assuming that 1 mg/ml = 20 A\text{260}^\nu.

b) A/Victoria:

To 10 ml virus suspension (at 4 mg protein/ml) was added an equal volume of 0.4% SDS, 20 mM EDTA. The mixture was extracted with 20 ml of phenol and 20 ml of chloroform. After phase separation, the organic layer was re-extracted with 5 ml NTE. The two aqueous phases (plus interfaces) were pooled and
extracted twice with equal volumes of chloroform. RNA was finally precipitated by the addition of 2.5 volumes of ethanol.

Ethanol precipitates were pelleted, dissolved in 1 ml TE and spun (10K x 10 min, Sorvall) to remove insoluble material. Viral RNA was re-precipitated by the addition of 3 volumes of 4 M NaCl. After standing at -20°C for 16 hr, the RNA was again pelleted, washed with 70% ethanol, dried and dissolved in TE.

**Enzymes**

Reverse transcriptase was isolated as described (9) from AMV-infected plasma supplied through the office of Program Resources and Logistics, Viral Cancer Program, National Cancer Institute, Bethesda, Maryland, U.S.A. Poly(A) polymerase was isolated (10) from late-log _E. coli_ B cells (Miles Laboratories). Sl-nuclease was prepared by the method of Vogt (11).

**Polyadenylation**

Viral RNA was polyadenylated at 37°C for the indicated times. The reaction mixture contained 50 mM Tris-HCl pH 8, 10 mM MgCl₂, 1 mM MnCl₂, 50 μg/ml purified ovalbumin, 0.1 mM ³H ATP (5 μCi/ml unless otherwise stated), 200 mM NaCl, 50 μg/ml viral RNA and 1/10 volume of poly(A) polymerase. Incubation was terminated by the addition of SDS to 0.2% and EDTA to 20 mM. The extent of adenylation was determined by TCA precipitation of a 5 μl aliquot of the incubation. The remainder was extracted with phenol and chloroform as described above and the final aqueous phase precipitated by the addition of an equal volume of 4 M NH₄Ac and then 2.5 volumes of ethanol.

**Synthesis of cDNA**

DNA complementary to the polyadenylated viral RNA was synthesised in a mixture containing 50 mM Tris-HCl pH 8.3, 0.2 mM dGTP, 0.2 mM dATP, 0.2 mM TTP, 50 μM dCTP (³²P or ³H as indicated), 5 mM DTT, 0.01% Triton X-100, 50 μg/ml actinomycin D, 40 mM KCl, 10 units/ml rat liver ribonuclease inhibitor, 5 μg/ml oligo(dT)₁₂₋₁₈, 20 μg/ml adenylated RNA and 60 units/ml reverse transcriptase. After incubation at 37°C for 90 min, the mixture was made 0.2% with SDS, 20 mM with EDTA and a 10 μl aliquot removed for TCA precipitation. To the remainder was added an equal volume of 4 M NH₄Ac and the cDNA precipitated by the addition of 2.5 volumes of ethanol. cDNA was recovered by centrifugation, dissolved in 100 μl H₂O and then made to 100 mM with NaOH. After incubation at 70°C for 20 min, the sample was neutralised with acetic acid and chromatographed on a 50 x 0.7 cm column of Sephadex G-150 equilibrated in 50 mM NaCl/0.1% SDS. DNA eluting in the void volume was pooled and concentrated by ethanol precipitation.
Synthesis of double-stranded cDNA

Double-stranded cDNA was synthesised from \(^3\)H-labelled single-stranded cDNA in an incubation containing 50 mM Tris pH 8.3, 20 mM DTT, 10 mM MgCl\(_2\), 0.4 mM dGTP, 0.4 mM dATP, 0.4 mM TTP, 0.1 mM dCTP (the concentration of radioactivity is given in the Figure Legends), 2.5-10 \(\mu\)g/ml cDNA and 400 units/ml reverse transcriptase. Incubation was at 45°C for the indicated periods. The reaction mixture was then made 20 mM with EDTA, 0.2% with SDS, extracted with phenol and chloroform, and placed on a Sephadex G-50 column (20 x 0.7 cm) equilibrated in 50 mM NaCl and 0.1% SDS. The excluded fractions were pooled.

Gel electrophoresis

RNA samples were electrophoresed on 3.7% or 4% acrylamide disc gels in 98% formamide prepared as described (12). After electrophoresis, gels were soaked for 2-3 hr in 30% glycerol/5% acetic acid and then scanned in a Joyce-Loebl u.v. gel scanner.

Samples of viral RNA dissolved in 7 M urea were also analysed on 3% acrylamide gels containing 7 M urea. The electrophoretic procedure was as described (2), except that both the gels and upper reservoir buffer contained 7 M urea. The RNA bands were stained for 1 hr in 0.02% methylene blue in citrate-EDTA and then gel destained in water.

Native DNA samples were examined on 1.4% agarose gels in Tris-acetate-EDTA buffer (13). For single-stranded DNA molecules, a 1.4% agarose slab gel (20 x 15 x 0.5 cm) in 30 mM NaOH, 2 mM EDTA was used (14). These agarose gels were dried (15) before autoradiography using Kodak XH-1 film and intensifying screens.

S1 nuclease treatment

The products of the reverse transcriptase reactions, reannealed double-stranded cDNA and RNA/cDNA hybrids were treated with S1 nuclease purified from α-amylase (11). The reaction mixtures (100-200 \(\mu\)l) contained 150 mM NaCl, 25 mM sodium acetate pH 4.6, 1 mM ZnSO\(_4\), DNA and 2.5-5 units of S1 nuclease and incubation was at 37°C for 30 min. Percent hybridisation was determined by TCA precipitation while samples for electrophoresis were phenol extracted and ethanol precipitated.

Hybridisation

Samples of viral RNA and cDNA for hybridisation or double-stranded cDNA for reassociation were dissolved in 10-20 \(\mu\)l of 300 mM NaCl, 10 mM Heps pH 6.8, 0.5 mM EDTA, 0.1% SDS, sealed in siliconised glass capillaries, denatured for 3 min at 100°C and incubated at 68°C for the indicated times.
Double-strand formation was assayed using Sl-nuclease as described above.

**Preparation of $^{32}$P-labelled RNA fragments**

Viral RNA was fractionated on 3% polyacrylamide gels in 7 M urea as described above, the bands excised and RNA eluted as described (16). The RNA was fragmented by digestion in 0.1 M sodium hydroxide for 1 hr in ice then neutralised and precipitated with ethanol. The fragments were terminally labelled by incubation at 37°C for 1 hr in a mixture (10 ul) containing 10 uCi $\gamma$$^{32}$P-polynucleotide kinase. After incubation, the labelled fragments were chromatographed on a 0.7 x 20 cm column of Sephadex G-50 (M).

**Nitrocellulose filter methodology**

cDNA was electrophoresed on alkaline agarose gels as described above and then transferred to nitrocellulose sheets as described (17). Filters were hybridised to labelled probes in 50% formamide, 5 x SSC at 37°C for 16 hr.

**RESULTS**

**Analysis of the virus genome**

The eight single-stranded RNA molecules comprising the genome of Influenza A Viruses are readily resolved on polyacrylamide or agarose gels in the presence of urea. The chain lengths of the RNA species have been estimated, in a number of cases, using ribosomal RNAs as standards, to range from $10^5$ to $1.2 \times 10^6$ daltons (see for example 18, 19).

We have analysed the virion RNA from both Fowl Plague and A/Victoria Virus on two denaturing gel systems: 7 M urea, pH 3.5 and 98% formamide. Fig. 1a compares the mobility of A/Victoria RNA with that of E.coli rRNA and 18S chick rRNA on acrylamide gels in the presence of urea. The viral RNA separates into eight species, although Band 1 tended to be more diffuse than the others. By comparing the mobilities of the RNA species with the mobilities of the ribosomal RNAs, we have estimated their chain lengths; these are shown in Table 1 and, generally, agree with published data (18, 19).

Electrophoresis of virion RNA on formamide disc gels is shown in Fig. 1b. Although this system failed to resolve genes 1-3, it is clear that relative to E.coli rRNA, the six large species migrate more rapidly. This is reflected in the lower chain lengths estimated from these formamide gels (Table 1). This difference in migration on the two gel systems is discussed below.

**Characterisation of the polyadenylation reaction**

The poly(A) polymerase showed an absolute requirement for either Mg$^{++}$ or Mn$^{++}$ ions. Maximal activity was found with a combination of 10 mM MgCl$_2$
Figure 1. Separation of vRNA segments of A/Victoria by polyacrylamide gel electrophoresis.

a) RNA was separated on a 3.0% polyacrylamide slab gel in 7 M urea (20 cm) at 80 V for 16 hr and visualised by staining with methylene blue. a = A/Victoria vRNA, b = E.coli 16S and 23S ribosomal RNA, c = hen oviduct 18S and 28S ribosomal RNA. The viral RNA genes are numbered 1-8.

b) RNA samples in 98% formamide were heated at 68°C for 4 min before electrophoresis on a 3.7% polyacrylamide disc gel (10 cm) in 98% formamide at 150 V for 6 hr. Gels were scanned at 265 nm. The positions of E.coli 16S and 23S ribosomal RNA and of oviduct 18S and 28S ribosomal RNA on parallel gels are shown.

and 1 mM MnCl₂. In agreement with the results of Sippel (10), the enzyme also had a specific requirement for high NaCl concentrations for maximal activity; this was obtained at 260 mM NaCl with a sharp decrease in activity at higher ionic strength (data not shown).

For the poly(A) polymerase to be useful in further studies involving transcription, it was important that the enzyme was free of nucleases. Consequently, the poly(A) polymerase was assayed for ribonuclease contamination using E.coli rRNA as a substrate by examining its migration on denaturing formamide gels after incubation. A 5 min incubation at 37°C produced little degradation of the RNA relative to the control incubated in the absence of enzyme (data not shown but see Fig. 2).
### Table 1

**RNAs of FPV and A/Victoria Virus**

<table>
<thead>
<tr>
<th>Gene</th>
<th>FPV (nucleotides)</th>
<th>A/Victoria (nucleotides)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7 M Urea</td>
<td>Formamide</td>
</tr>
<tr>
<td>1</td>
<td>3500</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>3000</td>
<td>2660-2720</td>
</tr>
<tr>
<td>3</td>
<td>2950</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>2450</td>
<td>2060</td>
</tr>
<tr>
<td>5</td>
<td>2000</td>
<td>1800</td>
</tr>
<tr>
<td>6</td>
<td>1720</td>
<td>1560</td>
</tr>
<tr>
<td>7</td>
<td>1080</td>
<td>980</td>
</tr>
<tr>
<td>8</td>
<td>870</td>
<td>830</td>
</tr>
</tbody>
</table>

*Data taken from (19)*

Viral RNA (5-10 µg) was electrophoresed either on slab gels containing 7 M urea in citrate/EDTA buffer and then stained with methylene blue (Fig. 1a) or on disc gels containing 98% formamide and then scanned at 265 nm (Fig. 1b). RNA chain lengths were deduced from plots of migration vs log (chain length) using E. coli rRNA (1680 and 3240 nucleotides) and chick 18S rRNA (2100) as standards.

**All eight viral genes are polyadenylated**

Although the poly(A) polymerase preparation was completely primer dependent, it was important to demonstrate that all eight genome segments were capable of being polyadenylated and to show a physical attachment of the poly(A) region to them. A standard assay mixture containing 10 µg FPV-RNA was incubated at 37°C for 3 min and then analysed on 4% acrylamide gels in 98% formamide. Fig. 2 shows that after this short incubation period, the incorporated radioactivity coincided exactly with the optical density profile obtained by scanning the gel. We calculated that in this experiment, an average of 40 adenylate residues was added per molecule. These poly(A) tails are evidently not long enough to produce the drastic change in migration and sedimentation properties observed for MS-2 RNA (20).

To achieve maximal transcription of the polyadenylated RNA using oligo-(dT)12-18 as a primer, it is essential that (a) the majority of the RNA segments are adenylated and (b) the length of the poly(A) tail is enough to
Figure 2. Electrophoresis of FPV-RNA after polyadenylation. 10 μg FPV-RNA was adenylated at 37°C for 3 min in the presence of 20 μCi/ml 3H-ATP. After phenol extraction and ethanol precipitation, vRNA was dissolved in 98% formamide, heated at 68°C for 2 min. After rapid cooling, the RNA was electrophoresed on a 4% polyacrylamide gel in 98% formamide. The gel was scanned at 265 nm with a Joyce-Loebl ultraviolet scanner ( ) and then cut into 100 x 1 mm slices. Slices were incubated at 37°C overnight with 0.5 ml of 95% NCS and then counted for radioactivity ( ).

promote transcription. Thus we incubated 2.5 μg aliquots of FPV-RNA with increasing quantities of poly(A) polymerase at 37°C for 2½ min. The polyadenylated RNA was recovered, added to a transcription mixture and incubated with oligo(dT) and reverse transcriptase at 37°C for 90 min. The results of this experiment are shown in Figs. 3a and 3b. The extent of adenylation was proportional to the enzyme concentration over the range examined (up to 50 A's per molecule). Transcription, however, began to plateau after approximately 30 A's per molecule had been added*, indicating that a poly(A) tail of sufficient length for efficient priming had been produced.

The results of Fig. 3 apply to the bulk of the RNA and no conclusion can be drawn as to the relative efficiency of transcription of individual genes. To examine this, we adenylated the A/Victoria RNA for 1 or 2 min (the latter time was sufficient for maximal transcription), transcribed it and electrophoresed the DNA products on a denaturing gel (Fig. 4). Although total incorporation after 1 min was only 50% of that after 2 min, it is clear that transcription was biased in favour of Band C (gene 4) at the early time. After 2 min adenylation, however, the DNA bands were of more regular intensity.

*The number of adenylate residues added per molecule was calculated as the ratio umoles AMP incorporated/umoles RNA present. In the case of FPV-RNA which has RNA species varying in molecular weight of 0.3–1.2 x 10^6 daltons, an average molecular weight of 0.7 x 10^6 daltons was used in the calculation.
Figure 3. Effect of poly(A) polymerase concentration on adenylation and reverse transcription of FPV-RNA. Five standard adenylation reactions, 50 µl each and containing 1 µCi $^3$H-ATP, were incubated at 37°C for 24 min with increasing quantities of poly(A) polymerase. Incubations were terminated by the addition of SDS to 0.2% and EDTA to 20 mM and 5 µl aliquots were removed for TCA precipitation (a). The remainder of each incubation was extracted with phenol and chloroform, precipitated with ethanol and then added to a 100 µl reverse transcription incubation containing $^{32}$P-dCTP (30 µCi/ml) as the radioactive deoxynucleotide triphosphate. After 90 min at 37°C, 10 µl aliquots were removed from each reaction for TCA precipitation (b).

Figure 4. cDNA production: the effect of length of polyadenylation. Standard adenylation reactions, 50 µl each, containing 2.5 µg A/Victoria RNA and 5 µl poly(A) polymerase, were incubated at 37°C for 1 and 2 min. Reactions were terminated, extracted with phenol and chloroform, precipitated with ethanol and transcribed as described in Fig. 3. cDNA was synthesised using $^{32}$P-dCTP (50 μM, 100 μCi/ml) as radioactive precursor. After 90 min at 37°C, 10 µl aliquots were removed for TCA precipitation. TCA precipitable radioactivity was 9950 cpm/10 µl and 22000 cpm/10 µl for the 1 and 2 min incubations respectively. Samples of each were electrophoresed on a 1.4% agarose slab gel (20 cm) in 30 mM NaOH at 40 V for 16 hr. After electrophoresis, the gel was dried and autoradiographed. a = $^{32}$P-Hind III fragments of PM-2 DNA, b = 4775 cpm of 1 min adenylation, c = 5400 cpm of 2 min adenylation.
Analysis of the transcription products

cDNA was recovered after NaOH digestion as described in the Methods section. This material was >96% sensitive to Sl-nuclease, indicating that it indeed is single-stranded, and became >90% resistant to Sl-nuclease after annealing with an excess of viral RNA (Table 2). We examined the lengths of the cDNA transcripts by electrophoresis on 1.4% agarose gels run in 30 mM NaOH to ensure completely denaturing conditions. Fig. 5 shows an autoradiograph of cDNA synthesised from polyadenylated FPV and A/Victoria RNA and the Hind III fragments of bacteriophage PM-2 DNA. The sizes of these latter fragments were determined on both agarose and acrylamide gels using the Hpa I and Hae III fragments of ØX174 as references and are given in the legend to Table 3.

Both cDNA preparations resolved into 7 bands, A-G, the lengths of which are given in Table 3. On the basis of their estimated sizes and relative mobilities, we have tentatively assigned them as follows: Band A is comprised mainly of genes 1 and 2, Band B contains mainly gene 3 transcripts, Band C is equivalent to gene 4, Band D to gene 5, Band E to gene 6, Band F to gene 7 and Band G to gene 8. We provide further evidence for these assignments below. Assuming the cDNA bands are equivalent to the RNA segments of the virus, we calculate a genome size of 14,600 nucleotides (4.8 x 10^6 daltons) for A/Victoria and 14,820 nucleotides (4.9 x 10^6 daltons) for Fowl Plague Virus.

The cDNA bands represent copies of individual genes.

Although we have shown that all the viral RNA species are polyadenylated and that the pattern of cDNA produced reflects that of the starting RNA, we cannot exclude the possibility that the smaller cDNA bands are produced by

Table 2

<table>
<thead>
<tr>
<th>Sample</th>
<th>% Sl resistant</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. FPV-cDNA</td>
<td>2.6</td>
</tr>
<tr>
<td>2. denatured cDNA</td>
<td>1.4</td>
</tr>
<tr>
<td>3. denatured cDNA, hybridised</td>
<td>7.4</td>
</tr>
<tr>
<td>4. cDNA + 1 µg FPV-RNA, denatured and hybridised</td>
<td>95.7</td>
</tr>
</tbody>
</table>
Figure 5. Comparison of cDNA synthesised from adenylated FPV-RNA and A/Victoria RNA. Viral RNAs were adenylated for 3 min, reverse transcribed and analysed on a 1.4% agarose gel (20 cm) in 30 mM NaOH. a = FPV cDNA; b = 32P-Hind III fragments of PM-2 DNA; c = A/Victoria cDNA.

Table 3

<table>
<thead>
<tr>
<th>Band</th>
<th>FPV-cDNA</th>
<th>A/Victoria cDNA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(lengths in nucleotides*)</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>2550</td>
<td>2500</td>
</tr>
<tr>
<td>B</td>
<td>2430</td>
<td>2400</td>
</tr>
<tr>
<td>C</td>
<td>1940</td>
<td>1920</td>
</tr>
<tr>
<td>D</td>
<td>1730</td>
<td>1680</td>
</tr>
<tr>
<td>E</td>
<td>1550</td>
<td>1580</td>
</tr>
<tr>
<td>F</td>
<td>1100</td>
<td>1100</td>
</tr>
<tr>
<td>G</td>
<td>970</td>
<td>930</td>
</tr>
</tbody>
</table>

*These values take no account of the length of the poly(T) tail on the cDNA.

cDNAs and 32P-labelled Hind III fragments of PM-2 were electrophoresed on alkaline agarose gels and the gels dried and autoradiographed. Lengths of cDNA bands were deduced from plots of migration vs log (chain length) of the PM-2 fragments. The lengths of these, estimated from 8X174 Hae III and Hpa I fragments (36), were 5400, 2350, 1050, 475, 450, 270 and 110 nucleotides.
"strong stop" signals in some of the larger RNA species. To demonstrate a relationship between individual RNA and DNA bands, we have fractionated the genome of A/Victoria, isolated the RNA bands corresponding to genes 4 and 7 and labelled them using polynucleotide kinase and $\gamma^{32}$P-ATP. These labelled fragments were then used as probes and hybridised to a "Southern-blot" of cold A/Victoria cDNA. The results, shown in Figure 6, clearly demonstrate that the gene 4 probe hybridises specifically to Band C while the gene 7 probe hybridises to Band F. We conclude, therefore, that the cDNA bands do indeed represent individual genes. Whether they are complete or lack the sequence corresponding to the 5'-terminal end of the virion RNA as the mRNA does (21) remains to be determined.

Hybridisation studies

We next attempted to characterise the cDNA to establish whether it represented a true complementary copy of the viral RNA; although the cDNA bands

---

Figure 6. Hybridisation of fractionated virion RNA to cDNA on Millipore filters. A/Victoria RNA was polyadenylated, $^3$H- and $^{32}$P-cDNA prepared and electrophoresed on a 1.4% agarose gel in 30 mM NaOH. After electrophoresis, the gel was neutralised and the cDNA transferred to a nitrocellulose filter. $a = 20 \text{ ng } ^3$H-cDNA; $b = 32^P$-cDNA, $10^3 \text{ cpm}; c = 20 \text{ ng } ^3$H-cDNA. The filters were cut into strips and hybridised to $^{32}$P-labelled RNA probes prepared from genes 4 and 7. The strips were wet with a minimum volume of hybridisation solution (50% formamide, 5 x SSC and $10^5 \text{ cpm/ml of } ^32P$-RNA), covered with liquid paraffin and left at 37°C overnight. Following hybridisation, the filter strips were washed four times with 50% formamide, 5 x SSC, once with 2 x SSC, once with 2 x SSC containing 10 μg/ml pancreatic RNAase and finally twice with 2 x SSC. Dried strips were then autoradiographed. $a = $ RNA probe from gene 4; $c = $ RNA probe from gene 7.
were large enough to contain near complete transcripts of Influenza RNA, its length might have been partly the result of anomalous synthesis of a homopolymer (22, 23) or "slippage" during transcription so that part of the mRNA was transcribed more than once.

Figure 7 shows the melting profile of vRNA/cDNA hybrids and the kinetics of hybridisation of cDNA to an excess of vRNA. The sharp melting transition at 88°C and the lack of significant melting below 80°C indicates that very little base pair mis-matching occurred. Similarly, the cDNA hybridisation reaction occurred within a 100-fold range of vRNA Cot values with a Cot\textsuperscript{1} of 4.8 x 10^{-3} mol. sec. lit^{-1} (at 0.3 M NaCl) and obtained a value of greater than 90% hybridisation.

Synthesis of double-stranded DNA

To produce synthetic genes from the single-stranded cDNA, we utilised the self-priming ability of the cDNA (24). Total \(^3\)H-cDNA was purified free of vRNA as described and used as template in a second reverse transcriptase reaction, in the presence of \(^32\)P-dCTP but without actinomycin D. As seen in Fig. 8, incorporation of radioactivity into acid insoluble material proceeded for 3-4 hr. During this time, DNA synthesis amounted to 40-50% of the input cDNA. In agreement with this figure, 40-50% of the \(^3\)H counts were rendered SI-resistant while 100% of the \(^32\)P counts were SI-resistant.

As shown for globin (24) and for ovalbumin (25), the second DNA strand

![Figure 7. Hybridisation kinetics of FPV cDNA to FPV RNA. \(^{32}\)P-labelled FPV cDNA (2.5 ng, 4000 cpm) was hybridised at 68°C for 2 min to 2 hr with an excess of vRNA (30 ng) in a final volume of 20 \(\mu\)l. The percentage of cDNA hybridised was determined using SI nuclease as described in Materials and Methods. The insert shows the melting profile of the cDNA/RNA hybrid. The hybrid was absorbed to a 0.5 ml hydroxyapatite column at 65°C in 0.14 M Na-phosphate pH 7. The temperature of the column was raised as indicated and single-stranded material eluted with 2.5 ml 0.14 M phosphate pH 7. The eluted material was TCA precipitated and counted.](image-url)
Figure 8. Kinetics of second-strand synthesis. 

$^{3}H$-labelled FPV cDNA (5 µg/ml) was incubated at 45°C with reverse transcriptase (200 units/ml) as described in Materials and Methods. 5 µl aliquots were removed at the indicated times and TCA precipitable radioactivity estimated. 

- - - $^{32}P$ cpm; - - - - , $^{3}H$ cpm.

is covalently bound to the cDNA by a "hairpin" structure that is susceptible to S1-nuclease. Such a structure implies that after denaturation, the molecules should "snap-back" immediately; after S1-nuclease treatment, however, reannealing of the strands should follow normal kinetics. The results of such an experiment are shown in Fig. 9. Double-stranded DNA, labelled with $^{32}P$ in the second strand and either treated or non-treated with S1-nuclease, was denatured (100°C for 3 mins) and then reannealed at 68°C for the indicated periods. The majority of the DNA with no prior S1-treatment was in a double-stranded form, even at the very early times (estimated Cot of less than $10^{-6}$ mol. sec. lit$^{-1}$), while that with S1-treatment reannealed with a Cot$^{1}$ of approximately $10^{-3}$ mol. sec. lit$^{-1}$.

Figure 9. "Hairpin" structures are cut by S1-nuclease to produce "open" genes. Double-stranded FPV DNA, labelled with $^{32}P$ in the second strand, was produced as described in Materials and Methods. A portion was treated with S1-nuclease and then reannealed after denaturation while a second sample was denatured and reannealed directly. ( - - - - ), reannealing after S1-nuclease digestion; ( - - - - ), reannealing without S1-nuclease treatment. Each incubation contained approximately 0.5 ng DNA/10 µl.
Thus the combination of reverse transcriptase and S1-nuclease is sufficient to produce synthetic viral genes. However, when double-stranded DNA was synthesised from unfractionated cDNA, the product, although containing some DNA bands equivalent to complete gene copies, consisted mainly of smaller molecules (results not shown). In an attempt to increase the yield of long gene copies, the single-stranded cDNA was fractionated prior to the second reverse transcriptase reaction. Thus, cDNA from A/Victoria vRNA was fractionated on an NaOH gel as in Figs. 4 and 5 and the bands corresponding to genes 4, 5, 6, 7 and 8 eluted from gel slices. Following reverse transcription and S1-nuclease digestion, samples were analysed on a neutral agarose gel. The results, shown in Fig. 10, indicate that the products in each case consist mainly of double-stranded molecules approximately equal in size to the cDNA molecules used in the incubation.

**DISCUSSION**

The Influenza Virus particle contains on its surface two immunologically and morphologically separate antigens, the hemagglutinin and the neuraminidase. It is the antigenic variability of these proteins that accounts for the

![Figure 10. Agarose gel electrophoresis of synthetic genes for A/Victoria 4-8. cDNA, specific activity $10^6$ cpm/μg, was synthesised from 100 μg polyadenylated A/Victoria RNA and electrophoresed on a 1.4% agarose gel in NaOH. Bands corresponding to genes 4, 5, 6, 7 and 8 (see Fig. 5) were recovered and double-stranded DNA, specific activity $10^6$ cpm/μg, produced. Following incubation with S1-nuclease, aliquots of each incubation were analysed on a native 1.4% agarose gel (30 cm). The gel was dried and autoradiographed. a) gene 4 DNA, 1850 bp; b) gene 5 DNA, 1620 bp; c) gene 6 DNA, 1500 bp; d) gene 7 DNA, 1050 bp; e) gene 8 DNA, 880 bp. DNA lengths, in base-pairs (bp), were calculated from photographs of the ethidium stained gel using the Hind III fragments of PM-2 DNA as standards.](image-url)
success of Influenza A Virus as a human pathogen. Two types of antigenic change are known to occur; the first, termed antigenic "shift", arises when viruses with major antigenic changes emerge, and is associated with pandemics. In between pandemics, frequent minor but progressive changes occur in the surface antigens, termed antigenic "drift" - these comprise the second type (for a review see (26)). The ability to produce synthetic DNA genes from mRNA with reverse transcriptase and the rapid development of genetic engineering and DNA sequencing methods now allow us to determine how the primary structure of the surface antigens changes in human pandemic Influenza Virus strains and to examine the regulatory sequences in the relevant genes.

Our approach to transcribing the viral genome was to adenylate the RNA with poly(A) polymerase and then use oligo(dT) as primer. The enzyme from E. coli has been described (10, 20, 23) and, as shown in Fig. 2, can be purified free of ribonuclease. Optimal adenylation occurred very quickly and, assuming that all RNA molecules are adenylated, maximal transcription occurred when ~30 A's per molecule were added (Fig. 3). Interestingly, from Fig. 4, it appears that gene 4 of A/Victoria is adenylated more rapidly than the others, possibly reflecting a more open structure near the 3' end of the molecule.

The differences observed in the sizes of viral RNA in urea gels (Table 1) and the cDNA bands (Table 3) prompted us to look at the RNA on other denaturing systems. The determination of RNA chain length by gel electrophoresis is fraught with difficulties stemming from the fact that RNA molecules generally do not have hydrodynamically equivalent conformations in aqueous solution. Several procedures have been developed to ensure denaturation; among these are buffers containing 6 M urea at pH 3.5, 5 mM methyl mercury or 98% formamide. Lehrach et al. (27) have recently examined a number of such systems and observed that, although urea gels at acid pH gave good separation of RNA molecules, they were unreliable for molecular weight determinations because differences in structure and protonation resulted in anomalous migration. Moreover, the work of Boedtker (28) and Rejinders et al. (29) suggests that 6 M urea at neutral pH, the method generally used in determining the size of Influenza Viral RNAs, does not fully denature most RNA molecules at room temperature. Electrophoresis in 98% formamide, especially after the samples are heated at 68°C for 4 min in formamide, should ensure more complete denaturation. Under these conditions, the sizes of the viral RNA (Table 1) and the cDNA products are in better agreement, leading to the conclusion that the cDNA bands probably represent complete or near-complete copies of
the viral genome and that the sizes of the Influenza genes may have been over-estimated in previous work. Indeed, Desselberger and Palese (30) have recently published revised molecular weights for Influenza A and B Viruses based on denaturation in 1 M glyoxal at 50°C. Their findings agree well with ours; thus the genome size is estimated as 4.9 x 10^6 daltons compared to the values of 4.88 x 10^6 for A/Victoria RNA (Table 1) and 4.8 x 10^6 and 4.9 x 10^6 for the cDNAs of A/Victoria and FPV respectively (Table 3).

There is also the question of whether the sizes of the RNA in formamide and the single-strand cDNA contain sufficient information to encode the observed viral proteins. Consider, for example, the gene for the hemagglutinin. Inglis et al. (3) and Palese and Schulman (4) have shown by biochemical and genetic methods that the hemagglutinins of FPV and A/Hong Kong, an H3/N2 strain like A/Victoria, are produced by gene 4 RNA. The sizes of these RNAs in formamide are estimated as 2060 and 1950 nucleotides for FPV and A/Victoria respectively while the cDNA sizes were 1940 and 1920 nucleotides respectively. If completely translated, these will produce a protein of 640-690 amino acids (70,000-75,000 daltons). Although the Influenza hemagglutinin from a number of strains has been examined and found to range from 75,000-85,000 daltons, a large proportion of this is carbohydrate. White (31) has estimated that HA1 comprises about 60% of the viral carbohydrate and Ward and Dopheide (32) have shown that glycosylated proteins exhibit anomalously low electrophoretic mobility in SDS-polyacrylamide gels. Similar results have been reported (33) for the major Avian Myeloblastosis Virus glycoprotein. When this effect of carbohydrate content on mobility is accounted for, the apo-proteins HA1 and HA2 have molecular weights of 35,500 and 28,300 daltons respectively (32); this total of 63,800 daltons (580 amino acids) is thus easily contained in both RNA and cDNA molecules of the size reported here (Tables 1 and 3). We have also examined the RNA/protein relationships for the other genes and, as seen in Table 4, have found good agreement between them.

The results presented here also indicate that the cDNA is a faithful copy of the vRNA. Thus greater than 95% is protected from S1-nuclease by the template RNA (Table 1) and the kinetics of hybrid formation and melting are as expected (Fig. 7). We have found no evidence of "slippage" occurring at the oligo(dT) primer.

In conclusion, we have outlined a procedure for obtaining double-stranded DNA copies of Influenza RNA genes. Although initially covalently joined, the connecting linkage is susceptible to S1-nuclease (Fig. 9), yielding DNA (Fig.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Virus</th>
<th>( M_W ) (daltons)</th>
<th><em>aa</em></th>
<th>Minimum number of nucleotides required in mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>P1–P3</td>
<td>A/FPV</td>
<td>81–94,000</td>
<td>730–850</td>
<td>2190–2560</td>
</tr>
<tr>
<td>HA</td>
<td>A/Memphis</td>
<td>35,500</td>
<td>320</td>
<td>1730</td>
</tr>
<tr>
<td></td>
<td>(A/Hong Kong variant)</td>
<td>28,300</td>
<td>257</td>
<td></td>
</tr>
<tr>
<td>NP</td>
<td>FPV</td>
<td>53–60,000</td>
<td>480–540</td>
<td>1440–1620</td>
</tr>
<tr>
<td>NA</td>
<td>FPV</td>
<td>45,000</td>
<td>410</td>
<td>1230</td>
</tr>
<tr>
<td>NA</td>
<td>X-7</td>
<td>58–66,000</td>
<td>530–600</td>
<td>1590–1800</td>
</tr>
<tr>
<td>M</td>
<td>FPV</td>
<td>25,000</td>
<td>230</td>
<td>690</td>
</tr>
<tr>
<td>NS</td>
<td>FPV</td>
<td>23,000</td>
<td>210</td>
<td>630</td>
</tr>
</tbody>
</table>

* Based on normal amino acid composition and an average molecular weight of 110 per amino acid.

† HA1 has a molecular weight of 47,000 of which 35,500 is protein (32).

§ Molecular weights are taken from (31), except those for the hemagglutinin, which are from (32).

10) that can be cloned in plasmids using synthetic linkers (34) or "tailing" procedures (35). Finally, we have revised the sizes of the viral genes. The new estimates, based on migration of the RNA in 98% formamide gels and the cDNA on alkaline gels, indicate that the viral genes probably contain few non-essential nucleotides and thus make it unlikely that the genes contain untranslated inserts within the structural sequence.

**Acknowledgements**

We thank Hilary Watkins and Parmjit Jat for help with virus and enzyme preparations, and Dr. A.J. Hale for provision of excellent research facilities. We are grateful to Anne Henry for typing the manuscript and John Hobbs for drafting graphical figures.
REFERENCES
