Volume 4 Number 12 December 1977 Nucleic Acids Research

The nucleotide sequence at the termini of adenovirus type 5 DNA

P.H. Steenbergh, J. Maat*, H. van Ormondt*, and J.S. Sussenbach

Laboratory for Physiological Chemistry, State University, Utrecht and *Laboratory for Physiological Chemistry, State University, Leiden, Netherlands

Received 27 October 1977

ABSTRACT

The sequences of the first 194 base pairs at both termini of adenovirus type 5 (Ad5) DNA have been determined, using the chemical degradation technique developed by Maxam and Gilbert (Proc. Nat. Acad. Sci. USA 74 (1977), pp. 560-564).
The nucleotide sequences 1-75 were confirmed by analysis of labeled RNA transcribed from the terminal *Hha*I fragments *in vitro*. The sequence data show that Ad5 DNA has a perfect inverted terminal repetition of 103 base pairs long.

INTRODUCTION

Human adenoviruses contain linear double-stranded DNA molecules with a molecular weight of 20-25 million[1,2]. It has been shown that in the virion a protein molecule is covalently linked to both 5'-terminal nucleotides of adenovirus DNA and that a circular DNA-protein complex can be isolated from virions[3,4]. The mechanism of replication of this DNA molecule has been the subject of extensive investigations (for a review see Levine et al.[5]).

Recently, experiments have been performed in various laboratories on the location of initiation and termination sites of adenovirus DNA replication[6-9]. Combination of all available data[6-12] suggests that adenovirus DNA replication starts at either one of the two molecular ends, proceeds displacing one of the parental strands, and terminates at the other end of the molecule. The displaced strand is replicated in the opposite direction after completion of the displacement synthesis. Recent studies on the initiation sites of replication have indicated that initiation occurs within the terminal 75 base pairs[13]. However, the mechanism of initiation is still completely obscure.

A potential structural basis for the initiation of adenovirus DNA replication at both molecular ends is provided by the identical nucleotide sequences at both termini of the type

```
3'-A B C D ------- D'C'B'A'-5'
5'-A'B'C'D'------- D C B A -3'
```

The presence of this inverted terminal repetition was suggested by the observation, that denaturation of native adenovirus DNA followed by renaturation at low concentrations leads to the formation of single-stranded circles[14,15]. The length of the terminal repetition in Ad2 DNA was estimated to be between 100 and 140 nucleotide pairs by restriction enzyme analysis[16]. This, however, does not prove that the terminal repetition is a perfect one. Both molecular ends of Ad2 as well as of Ad5 DNA were shown to terminate with the sequence pCpC- - - -pGpApTpG$_{OH}^{3'}$, indicating that the inverted terminal repetition extends to the very ends of the molecule[17].

The present report concerns the sequence analysis of the terminal 194 nucleotides at both ends of the Ad5 genome, with the aim of determining the exact length of the inverted terminal redundancy as well as the degree to which this repetition is perfect. The possible relation between the deduced sequences and the mechanism of DNA replication will be discussed.

MATERIALS AND METHODS

Materials

Ad5 DNA was isolated from purified virions as described previously[18]. Restriction endonucleases *Hpa*I and II, *Sma*I and *Hsu*I were purified according to Sharp et al.[19], Greene and Mulder[20], and Roberts and Myers [21], respectively. *Hph*I was a generous gift of Dr. R.J. Roberts, and *Hha*I and *Hae*III were purchased from Biolabs (Boston, Mass.). Bacterial alkaline phosphatase (BAP F) was purchased from Worthington Biochemicals (Freehold, N.J.). T_4-polynucleotide kinase was purified according to the method described by Richardson[22]. Fraction VI was again eluted over a DEAE-cellulose column, concentrated on a small phosphocellulose column and stored until use at -20° C in 50% glycerol. T_1 RNAase was obtained from Sankyo Co. (Tokyo). Pancreatic RNAase and spleen phosphodiesterase were from Boehringer (Mannheim). *E. coli* DNA dependent RNA polymerase, free from ATPase, DNAase and RNAase activities, was a generous gift of Drs. C. van Kreyl and H. Bos. Carrier tRNA and yeast RNA used in homomixtures were purchased from BDH (Poole). (γ-^{32}P) ATP (specific activity 2000-4000 Ci/mmole) and (α-^{32}P) NTP's (specific activity 250 Ci/mmole) were obtained from The Radiochemical Centre (Amersham). Unlabeled ribonucleoside triphosphates were purchased from Merck (Darmstadt). Dimethylsulfate (99% pure, gold label) was obtained from Aldrich-Europe (Beerse), hydrazine (95% pure) from Eastman Chemicals (Rochester, N.Y.) and piperidine (99% pure) from Merck (Darmstadt). Acrylamide and bisacrylamide (ultra pure) were from BDH

(Poole) or Serva (Heidelberg). Urea (ultra pure) was purchased from Schwarz-Mann (Orangeburg, N.Y.). Cellulose acetate strips for electrophoresis were obtained from Schleicher and Schüll (Dassel). Polyethyleneimine thin layer plates (20 x 20 cm) for homochromatography from Macherey-Nagel Co. (Düren) were washed according to the procedure described by Southern and Mitchell[23].

Isolation of restriction fragments

Large restriction enzyme fragments (>2%), derived from the termini of Ad5 DNA, were isolated by introducing the material of the 2.5% polyacrylamide slabgel, on which they were separated, into a glasswool-plugged 10 ml disposable pipette. A dialysis membrane was attached to the end of the pipette. Membrane, pipette and reservoirs of the apparatus in which the pipette was fitted contained 20 mM Tris-HCl pH 7.8, 2.5 mM sodium acetate, 0.5 mM EDTA. The fragments were collected in the membrane by electrophoresis[24]. After 16 hrs at 100 V the current was reversed for 30 min. The contents of the dialysis membrane was combined with an equal volume of buffer in which the emptied membrane had been heated at 60° C for 30 min. After concentration with n-butanol the fragments were purified further by elution over a Sephadex G-100 column in 2 mM Tris-HCl pH 8.0, 0.05 mM EDTA. Small restriction enzyme fragments were recovered, usually from 5% polyacrylamide gels, by the gel elution procedure described by Maxam and Gilbert[25].

Dephosphorylation

Five to ten pmoles of one of the larger terminal restriction enzyme fragments, i.e. the *Hsu*I fragment I or the *Hpa*I fragment E, were digested with a second restriction enzyme. The resulting mixture of fragments was eluted over a Sephadex G-100 column in 50 mM Tris-HCl pH 8.0 to remove the phosphate usually present in restriction enzyme preparations. The eluted fragments were dephosphorylated by the addition of alkaline phosphatase (final concentration 4 U/ml) and incubation for 1 hr at 37° C. The phosphatase was removed by three extractions with an equal volume of freshly destilled phenol saturated with the Tris-buffer. After concentration with n-butanol the fragments were again eluted over a Sephadex G-100 column in the same Tris-buffer and precipitated with ethanol after the addition of sodium acetate to 0.3 M and standing at -70° C for 15 min. The precipitate was washed once with ice-cold 75% ethanol.

Labeling of 5'-ends

The 5'-ends of the dephosphorylated fragments were labeled with (γ-^{32}P) ATP and polynucleotide kinase as described by Maxam and Gilbert[25]. The precipitate was dissolved in 42.5 µl 5 mM Tris-HCl, pH 9.5, 10 mM EDTA, 0.1 mM spermidine, heated at 100° C for 3 min and chilled in ice-water. After addition of 5 µl 0.5 M Tris-HCl, pH 9.5, 0.1 M $MgCl_2$, 50 mM dithiotreitol, the mixture was transferred to the dried (γ-^{32}P) ATP. The final ATP concentration was 2 µM. Several units of polynucleotide kinase were added and phosphorylation was allowed to take place for 30 min at 37° C. The reaction was stopped by the addition of 4 µl 0.5 M EDTA, followed by heating for 2 min at 100° C. The mixture was eluted over a Sephadex G-100 column in 2 mM Tris-HCl pH 8.0, 0.05 mM EDTA, to separate the phosphorylated fragments from unreacted (γ-^{32}P) ATP. The fractions containing the DNA were concentrated with n-butanol, $MgCl_2$ was added to 3 mM and the fragments were renatured by incubation at 60° C for 2 hrs.

The original 5'-ends of Ad5 DNA will not be labeled by this procedure[26]. Terminal fragments therefore will be labeled at the 5'-ends resulting from the restriction enzyme digestion only. Non-terminal fragments will be labeled at both 5'-ends. Before separating the labeled fragments on a 5% polyacrylamide slabgel, a restriction enzyme digestion was in some cases included at this point to cut a non-terminal fragment of interest.
Fragments from which the sequence was to be derived were cut from the gel after autoradiography for 30 min, isolated after addition of 50 µg carrier tRNA, and divided into four equal portions.

DNA-sequence analysis

The chemical degradation reactions were carried out exactly as described by Gilbert and Maxam[25]. Specific cleavage at guanine residues was achieved by methylation with dimethylsulfate, at adenine residues by ring-opening with alkali, at cytosine and thymine residues by hydrazinolysis and at cytosine residues alone by hydrazinolysis in 2 M NaCl, in all cases followed by incubation with piperidine. The reaction products were fractionated on 12 and 20% polyacrylamide slabgels (40 x 20 x 0.15 cm), containing 90 resp. 50 mM Tris-borate pH 8.3, 1 mM EDTA, 7 M urea, at 600-1000 V. Usually two or three runs were performed with each set of samples. In short runs the bromphenol-blue tracking-dye moved about 30 cm, in long runs the xylene cyanol F.F.-marker moved about 30 cm. The same Tris-borate buffer without urea was used

as reservoir-buffer. Samples were loaded in 50% formamide directly after heating at 100° C for 3 min. After electrophoresis the DNA-bands were immobilized within the gel by precipitation with 10% acetic acid. Autoradiography was for 3-14 days.

Transcription-reaction

Incubation mixtures were essentially the same as described by Blackburn[27]: 25 mM Tris-HCl pH 7.8, 8 mM KCl, 10 mM $MgCl_2$, 5 mM dithiotreitol, 0.5 mM EDTA, 0.8 mM potassium phosphate, three unlabeled ribonucleoside triphosphates at a concentration of 0.5 mM each and one (α-^{32}P) ribonucleoside triphosphate at a concentration of 0.05 mM.
0.1 µg of the double-stranded restriction enzyme fragment was denatured by heating at 100° C for 3 min, cooled rapidly in ethanol-dry ice and added to the reaction mixture. RNA polymerase was added to bring the molar template to polymerase ratio to 1:3. The final volume was 25 µl. The transcription reaction was carried out for 3 hrs at 37° C and was stopped by the addition of 25 µl 20 mM EDTA, containing 2 mg/ml tRNA. This mixture was extracted with an equal volume of freshly destilled phenol, saturated with 10 mM Tris-HCl pH 7.5, 0.1 M NaCl, 1 mM EDTA, and eluted over a Sephadex G-50 column in the latter buffer to remove the low molecular weight materials. The RNA eluting in the void volume was precipitated with ethanol and dissolved in a small volume of water.

T_1 RNAase and pancreatic RNAase fingerprinting

The labeled transcripts were digested with either T_1 RNAase or RNAase A for 30 min at 37° C at an enzyme to carrier tRNA ratio of 1:5 or 1:10, respectively. The digestion products were separated on cellulose acetate strips and polyethyleneimine thin layer plates by the standard two-dimensional system described by Brownlee and Sanger [28], in the modification of Volckaert et al.[29]. The homochromatography was performed at 60° C with 3% RNA homomix, which was hydrolysed for 30 min at pH 12.8 at room temperature. The labeled oligonucleotides were located by autoradiography and eluted as described by Volckaert et al.[29].

Analysis of oligonucleotides

The oligonucleotides from the T_1 RNAase fingerprints were digested with pancreatic RNAase or hydrolysed with alkali. The products were characterized

by electrophoresis on DEAE paper at pH 3.5 or on Whatman 540 paper at pH 3.5, respectively, according to Barrell[30]. The oligonucleotides from pancreatic RNAase fingerprints were digested with T_1 RNAase or hydrolysed in alkali. The products were characterized in the same way. All products were located by autoradiography. The relative amounts of the products were determined by cutting the spots from the paper and measuring the Cerenkov radiation in a liquid scintillation counter.

RESULTS

Restriction enzyme maps of the termini of Ad5 DNA

Right-hand terminus

The restriction enzymes *Hsu*I (*Hind*III) and *Sma*I produce terminal fragments of 3% (*Hsu*I-I) and 2% (*Sma*I-L) of genome length, respectively[7,31]. Subfragments of *Hsu*I-I produced by various restriction enzymes were mapped according to the method described by Smith and Birnstiel[32]. The *Hsu*I-I fragment was labeled at both 3'-ends and cut with *Sma*I, which results in two fragments, each labeled at one 3'-end. The labeling procedure with exonuclease III and Klenow-polymerase will be described[33]. The labeled fragments were isolated from a 3% polyacrylamide gel and digested with different restriction enzymes under such conditions that all possible partial digestion products were generated. The products were separated by electrophoresis on a 5% polyacrylamide slabgel. The lengths of the terminal 3'-labeled subfragment and of the partial digestion products containing this terminal subfragment were calculated from the autoradiogram. From these data the restriction enzyme maps could directly be constructed. The positions of the restriction enzyme sites relevant to the experiments described in the next section are shown in Fig. 1.

Left-hand terminus

The restriction enzymes *Hpa*I and *Sma*I produce terminal fragments of 4% (*Hpa*I-E) and 3% (*Sma*I-K), respectively[34,31]. Detailed restriction enzyme maps of the *Hpa*I-E fragment with *Hpa*II, *Hae*III, *Alu*I, *Taq*I, and *Hin*f will be described[33]. The *Hha*I and *Hph*I recognition sites within the *Hpa*I-E fragment were mapped by a combination of data obtained from partial digests, according to the method described for the right-hand terminus, from cross-digestions of the *Hha*I and *Hph*I fragments with restriction enzymes of which the map was already known, and from sequence analysis results available at that time.

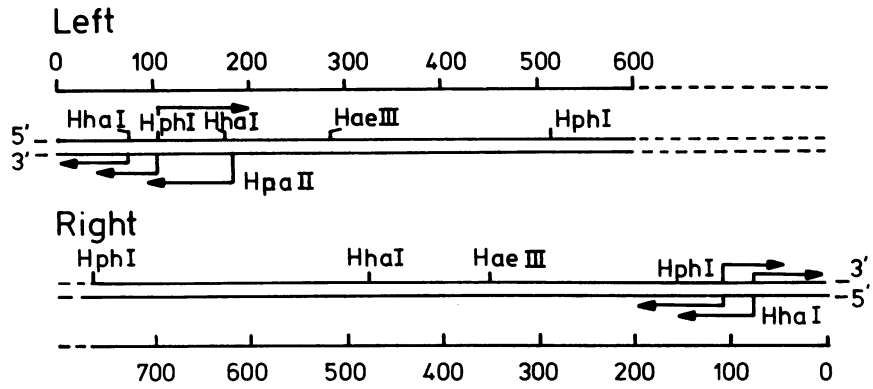

Figure 1. Location of restriction endonuclease recognition sites at the termini of Ad5 DNA. The nucleotide sequences derived by the chemical modification technique are indicated by arrows starting from the labeled 5'-terminal nucleotide.

These maps were confirmed by the complete nucleotide sequence of the *Hpa*I-E fragment[35]. The sites mentioned in the following section are indicated in Fig. 1.

DNA sequence analysis

Right-hand terminus

The 3% terminal *Hsu*I-I fragment was digested with *Hha*I and the resulting subfragments were labeled at their 5'-ends with kinase. The 75 base pairs fragment derived from the end of the Ad5 genome will be labeled at the 5'-end of the *Hha* cleavage site only, since the original terminus of Ad5 DNA is insensitive to the action of phosphatase[26]. The adjacent *Hha*I fragment, 400 base pairs long, will be labeled at both 5'-ends. In order to make the latter fragment directly available for sequence analysis the whole mixture of fragments was digested with *Hae*III before electrophoretic separation. The 75 base pairs terminal *Hha*I fragment and the 275 base pairs *Hha*I/*Hae*III fragment were cut from the gel and used for sequence analysis. From these experiments the sequence from nucleotide 70 towards the terminus and from 80 to 140 could be derived. To establish the sequence between nucleotides 70 and 80 and from 140 onwards, the *Hsu*I-I fragment was digested with *Hph*I. Again the subfragments were labeled, digested with *Hae*III and separated. The 105 base pairs terminal *Hph*I fragment and the 245 base pairs *Hph*I/*Hae*III fragment were analysed. Representative gels for the analysis of the first 100 nucleotides are shown in Fig. 2.

The sequence from 1-194 determined from these experiments is shown in Fig. 3.

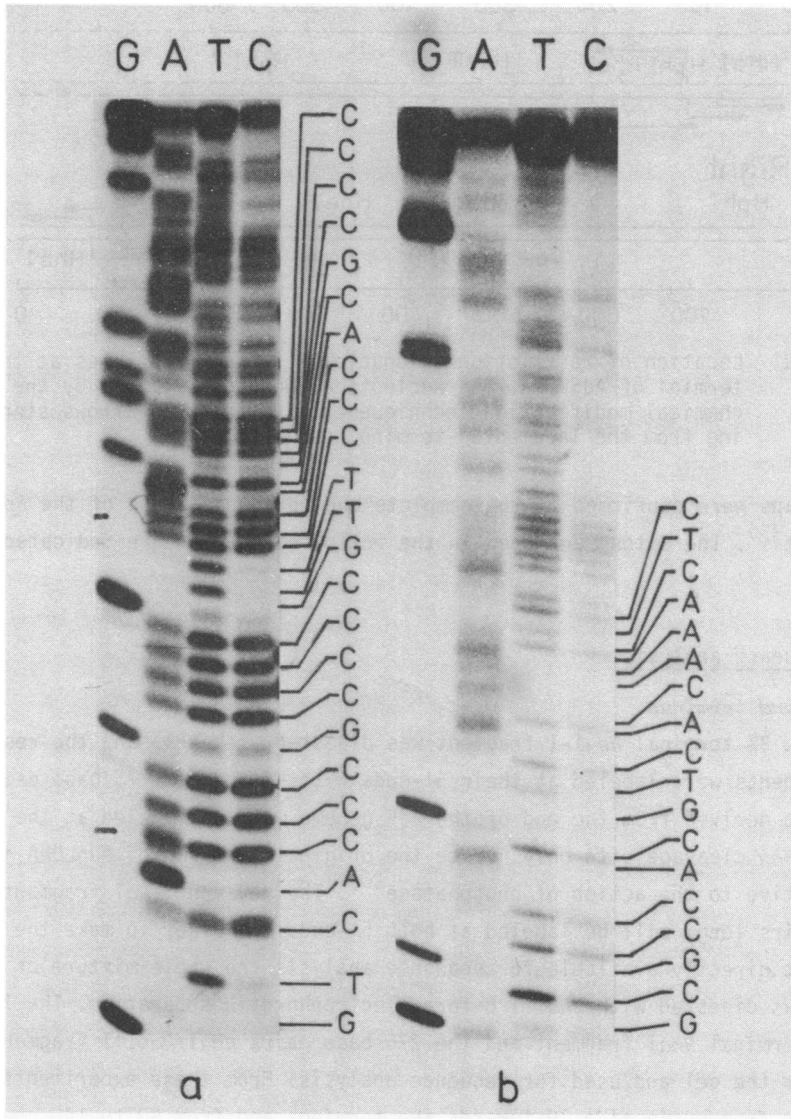

Figure 2.
Sequencing gels of the right-hand terminal *Hph*I fragment (a and b, short and long run, respectively) and *Hha*I fragment (c, short run, and d and e, two longer runs).

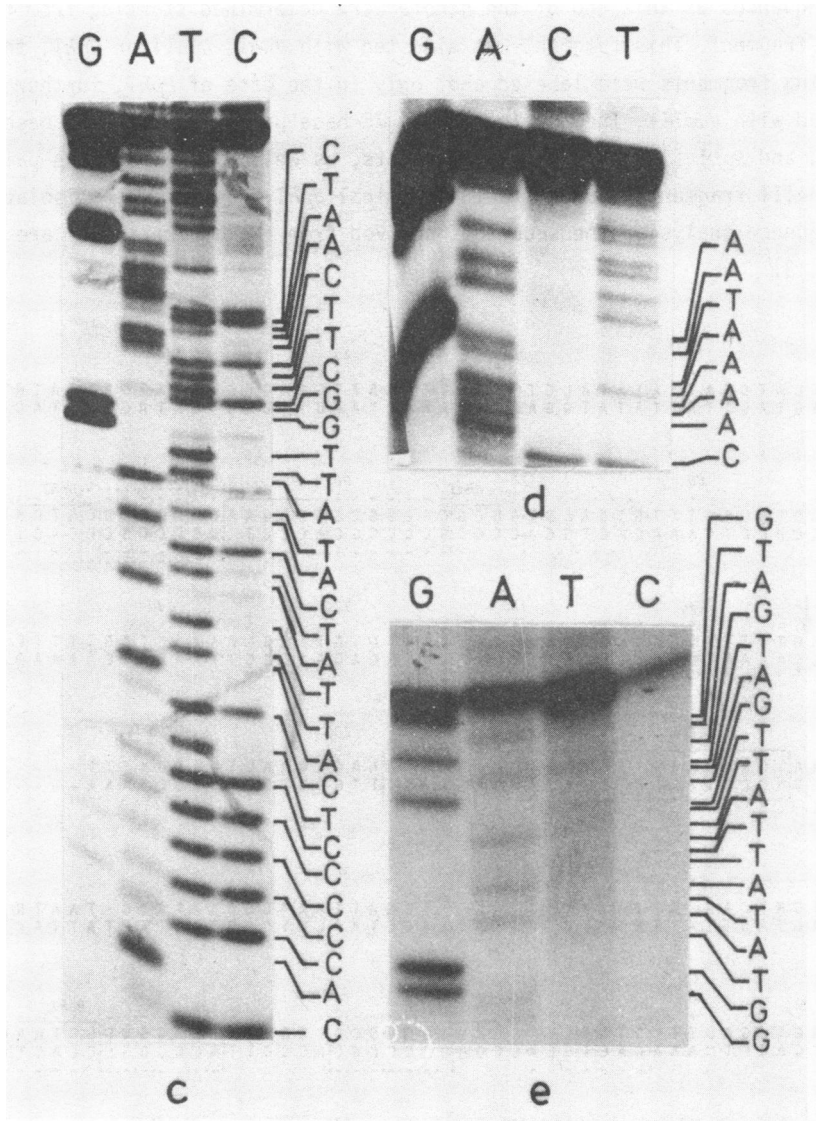

Figure 2.
(Continued)

Left-hand terminus

Sequences at this end of the genome were determined starting from the *Hpa*I-E fragment. This fragment was digested with *Hha*I, *Hpa*II or *Hph*I, the resulting fragments were labeled and, only in the case of *Hph*I, further digested with *Hae*III. The terminal *Hha*I (75 base pairs), *Hpa*II (190 base pairs), and *Hph*I (105 base pairs) fragments, as well as the 180 base pairs *Hph*I/*Hae*III fragment adjacent to the terminal *Hph*I fragment, were isolated for sequence analysis. The sequences derived from these experiments are shown in Fig. 3.

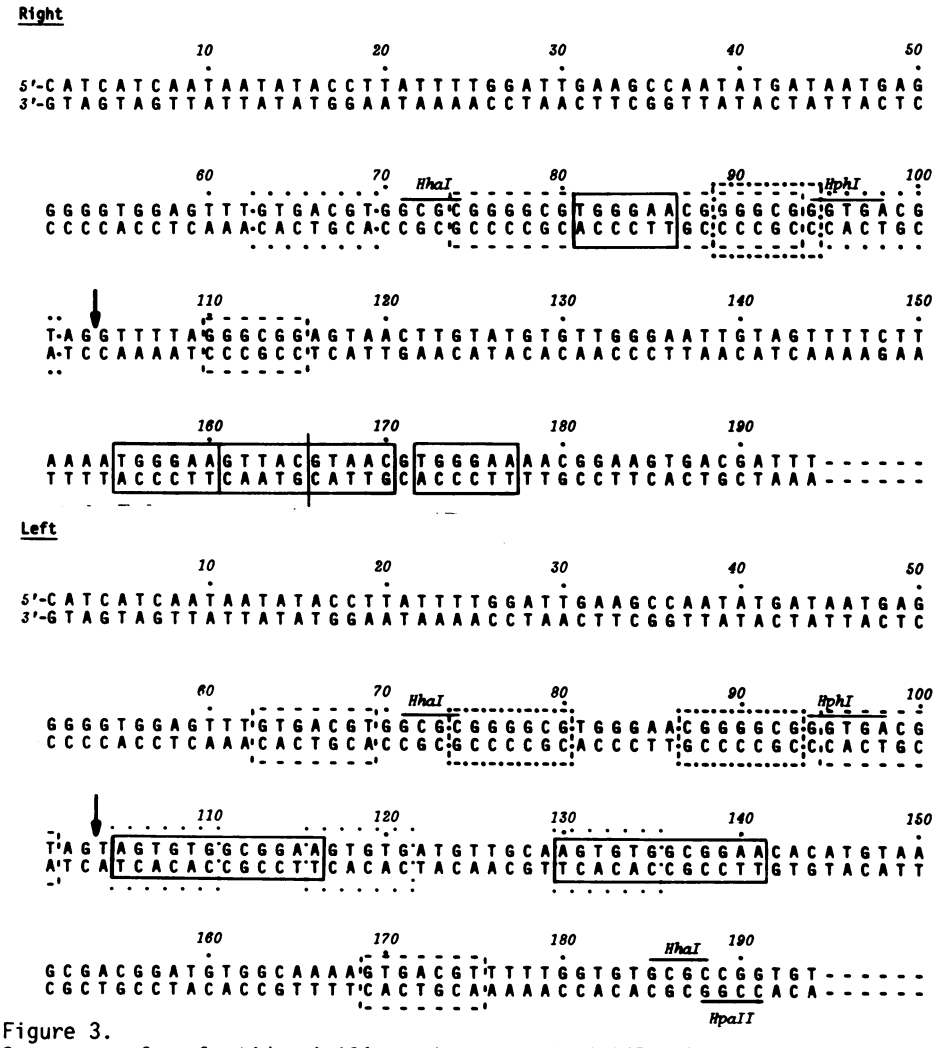

Figure 3.
Sequences of nucleotides 1-194 at the termini of Ad5 DNA. The arrows indicate the end of the inverted terminal repetition.

RNA sequence analysis

The terminal 75 base pairs *Hha*I fragments from both ends of the Ad5 DNA were isolated. The DNA was first digested with *Sma*I, the terminal fragments K (3%) and L (2%) were isolated, and these were digested with *Hha*I. The products were visualized after electrophoretic separation under UV light by ethidium bromide staining. The L fragment was cut into three fragments (400, 160 and 75 base pairs in length) while six fragments were produced from the K fragment (550, 190, 120, 115, 75 and 35 base pairs).

Transcripts were made from both 75 base pairs terminal fragments. The kinetics of RNA synthesis by *E. coli* RNA polymerase using a 71 base pairs denatured restriction enzyme fragment in the same system have been described[36]. Analysis of the RNA on urea gels[36] showed that only very little specific transcription products are formed. They are visible in the autoradiogram (Fig. 4) as distinct bands superimposed on a smear, indicating that transcription is initiated and/or terminated with little preference for any specific site on the fragment.

T_1 RNAase and pancreatic RNAase fingerprints of the transcripts, labeled separately with each of the four ribonucleoside triphosphates were made (Fig. 5) and all products were analysed. Fingerprints and oligonucleotide analysis results from transcripts synthesized on the two terminal *Hha*I fragments were completely identical. In Tables 1 and 2 only the analysis results for the longer T_1 RNAase and RNAase·A products that were reproducibly present

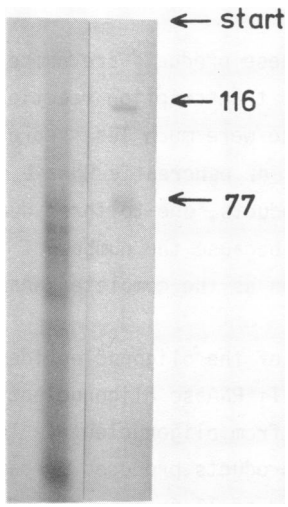

Figure 4.

Urea-gel analysis of the RNA transcribed from one of the terminal *Hha*I fragments. ^{32}P-labeled 5S RNA from *Bacillus licheniformis* (116 nucleotides in length) and tRNA from yeast (mean length 77 nucleotides) were used as length markers.

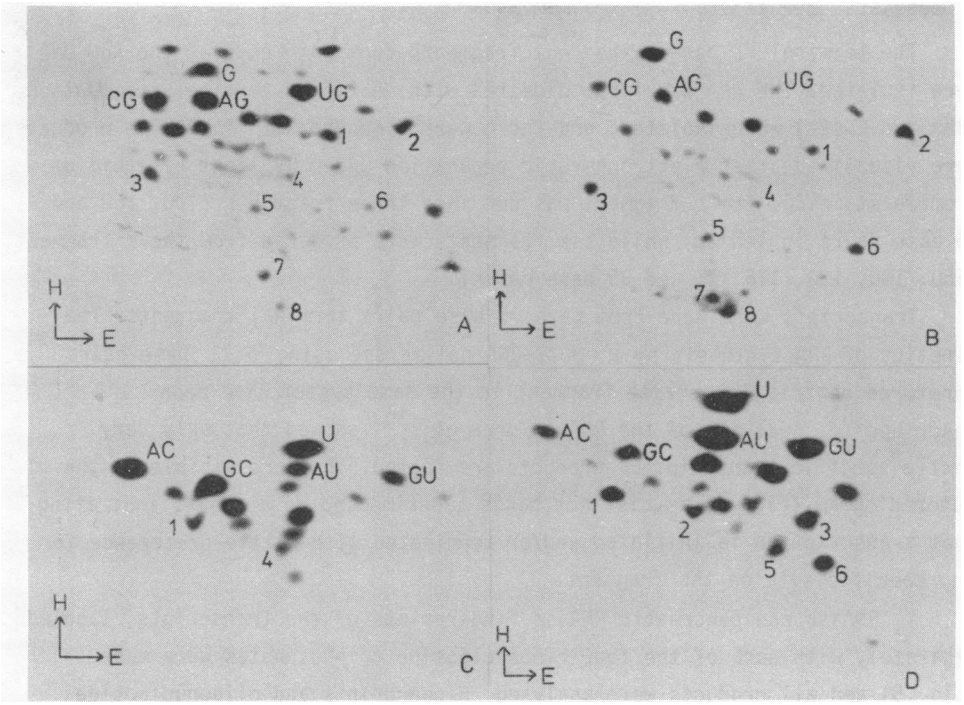

Figure 5. T_1 RNAase fingerprints of transcripts labeled with (α-^{32}P) GTP (A) and (α-^{32}P) UTP (B) and pancreatic RNAase fingerprints of transcripts labeled with (α-^{32}P) CTP (C) and (α-^{32}P) UTP (D). Electrophoresis (E) was from left to right and homochromatography (H) from the bottom upwards.

in all four different fingerprints are listed. These products are numbered in Fig. 5. As a consequence of the randomness of the transcription reaction many minor spots are present in the fingerprints; these were much less reproducible and after analysis turned out to be incomplete T_1 or pancreatic RNAase oligonucleotides. The analysis results for shorter products, one to three nucleotides in length, could easily be misinterpreted, because the number of incomplete products having the same overall composition as the complete RNAase oligonucleotides is evidently much higher here.

From the data in Table 1 and 2 the sequence of the oligonucleotides could be derived directly, except for the longer T_1 RNAase oligonucleotides T6, T7 and T8. The correct order of the products from oligonucleotide T6 was determined by two-dimensional separation of the products produced by partial spleen phosphodiesterase digestion from the 5'-end of T6 from a transcript

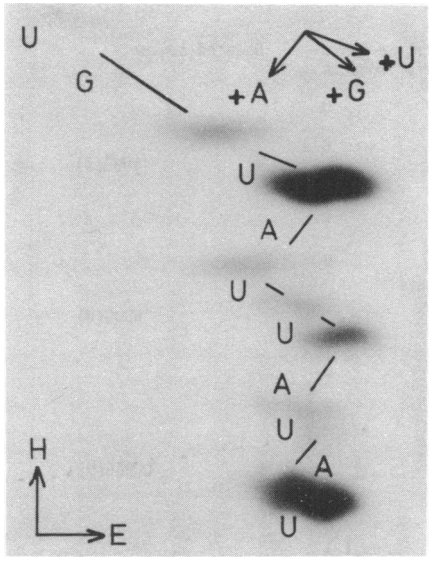

Figure 6.

Two dimensional separation of the products generated by partial digestion of T6 from a transcript labeled with (α-^{32}P) GTP with spleen phosphodiesterase. The sequence UAUAUUAUUG corresponds to nucleotides 16-7 in the DNA sequence.

labeled with (α-^{32}P) GTP, in which case only the 3'-terminal nucleotide is labeled (Fig. 6). From the data in Table 1 no unique sequence can be concluded for oligonucleotides T7 and T8. All data are in agreement with the sequences given in the last column of the table, which are derived from nucleotides 34-17 and 66-35 in the DNA sequence. They are also confirmed in part by sequences of oligonucleotides complementary to T7 and T8, i.e. P3 and P4 confirm T7, while T2, T4, T5 and P6 confirm T8, and by analysis of some incomplete T_1 RNAase oligonucleotides. Fig. 7 shows the location of the oligonucleotides from Table 1 and 2 in the sequence.

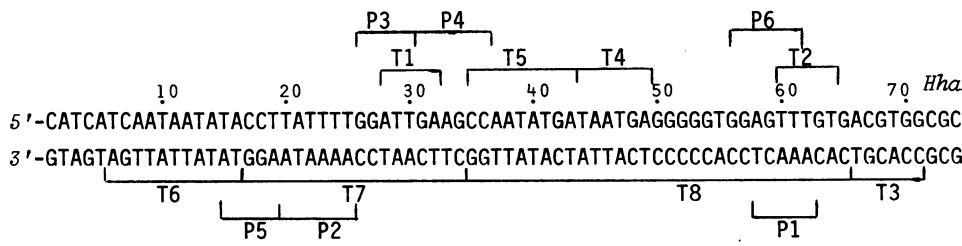

Figure 7. The DNA sequence of the terminal *Hha*I fragments, showing the location of the nucleotides which correspond to the T_1 and pancreatic RNAase products in Tables 1 and 2.

Table 1 Oligonucleotides produced by T_1 RNAase digestion

Oligo-nucleotide	Label	Pancreatic-RNAase products	Alkaline hydro-lysis products	Deduced sequence
T1	G	U	U	
	A	G	G	AUUG(A)
	U	AU	A,U 1 1	
	C	-	-	
T2	G	U	U	
	A	-	-	UUUG(U)
	U	U,G 2 1	U,G 2 1	
	C	-	-	
T3	G	AC	C	
	A	C	C	CCACG(U)
	U	G	G	
	C	C,AC 1 1	A,C 1 1	
T4	G	AAU	U	
	A	G,AU,AAU 1 1 1	G,A,U 1 1 1	AUAAUG(A)
	U	AU,AAU 1 1	A	
	C	-	-	
T5	G	AU	U	
	A	C,G,AAU 1 1 2	G,A,U,C 1 1 1 1	CCAAUAUG(A)
	U	AU,AAU 1 1	A	
	C	C	C	
T6	G	U	U	
	A	G,U,AU 1 2 1	G,U 1 3	UAUAUUAUUG(A)
	U	AU	A,U 3 2	
	C	-	-	
T7	G	AAG	A,G 1 1	
	A	C,AAG,AAU,AAAAU 2 1 1 4	A,U,C 5 1 2	CUUCAAUCCAAAAUAAG(G)
	U	U,C,AAU,AAAAU 1 1 1 1	A,U,C 2 1 1	
	C	U,C,AAU 1 1 1	U,C 2 1	
T8	G	G,U 1 1	G,U 1 1	
	A	U,C,AU,AC,AAAC 1 4 1 1 2	A,U,C 2 2 5	UCACAAACUCCACCCCC-
	U	C,AU,AAAC 1 6 1	A,U,C 4 2 2	-UCAUUAUCAUAUUG(G)
	C	U,C,AU,AC,AAAC 3 4 1 3 1	A,U,C 3 4 5	

Table 2 Oligonucleotides produced by pancreatic RNAase digestion

Oligo-nucleotide	Label	T_1 RNAase products	Alkaline hydrolysis products	Deduced sequence
P1	G	-	-	
	A	AAAC	A	AAAC(U)
	U	AAAC	C	
	C	AAAC	A	
P2	G	-	-	
	A	AAAAU	A,U 3 1	AAAAU(A)
	U	AAAAU	A	
	C	-	-	
P3	G	G	G	
	A	G	G	GGAU(U)
	U	AU	A,U 1 1	
	C	-	-	
P4	G	AAG	A	
	A	G,AAG 1 1	G,A 1 1	GAAGC(C)
	U	-	-	
	C	C,AAG 1 1	G,C 1 1	
P5	G	AAG	A,G 1 1	
	A	U,AAG 1 1	A,U 1 1	AAGGU(A)
	U	G	G	
	C	-	-	
P6	G	AG,G 1 1	A,G 1 1	
	A	G	G	GGAGU(U)
	U	U,AG 1 1	G,U 1 1	
	C	-	-	

Tables 1 and 2.
Analysis results for the longer T_1 RNAase and pancreatic RNAase oligonucleotides. When more than one product was formed, the relative amounts are given.

DISCUSSION

DNA sequence analysis

From the sequencing gels of the right-hand terminal *Hha*I and *Hph*I fragments of Ad5 DNA the sequence of the first 100 nucleotides of the l-strand

could be determined. Analysis of the adjacent *Hha*I and *Hph*I fragments resulted in the sequence of nucleotides 80 to 194 in the r-strand. The sequence of the first 185 nucleotides from the r-strand at the left-hand terminus was established by analysis of the terminal *Hha*I, *Hph*I and *Hpa*II fragments. By analysing the second *Hph*I fragment the sequence of nucleotides 110 to 185 in the l-strand was confirmed and extended to nucleotide 194, just beyond the *Hpa*II site.

RNA sequence analysis

The sequences derived by the analysis of the longer T_1 RNAase and pancreatic RNAase oligonucleotides confirm the sequences of nucleotides 6 to 72 in the terminal *Hha*I fragments (Fig. 7). No oligonucleotides, complete or incomplete, were detected with a sequence corresponding to the first 25 nucleotides of the DNA sequence at the 5'-ends of the Ad5 DNA. Variations in polymerase to template ratio, salt or ribonucleoside triphosphate concentrations, temperature and addition of synthetic primers to the transcription reaction-mixture gave the same results. The reason for this is not clear, but it might be caused by the fact, that the RNA-polymerase preferentially initiates transcription with the incorporation of a guanosine residue; transcription in that case would start at nucleotide 26, which is confirmed by the presence of P3 in the fingerprints. A second unique oligonucleotide not present in the fingerprints predicted by nucleotides 48 to 56 in the DNA sequence is the pancreatic RNAase product GAGGGGGU(G). Probably this oligonucleotide was specifically lost by aggregation, which is a well known phenomenon for such guanosine-rich sequences, especially when the G-residues form a consecutive run[29,37].

Inverted terminal repetition

From the sequence analysis results Ad5 DNA is shown to contain an inverted terminal repetition of 103 base pairs, i.e. the first 103 base pairs at the right-hand terminus of the Ad5 genome are completely identical to the first 103 base pairs at the left-hand terminus, while the adjacent sequences up to nucleotide 194 do not show any significant homology. A striking feature of the terminal repetition is the asymmetric distribution of AT and GC base pairs. The first 50 base pairs are 72% AT and 28% GC while the next 50 are 26% AT and 74% GC.

In Fig. 3 repeated and palindromic sequences longer than 5 nucleotides have been marked. The longest repetition is located at the left-hand terminus, where the sequence between 105 and 116 is repeated from 130 to 141. The only

palindromic sequence is located between nucleotides 161 and 170 at the right-hand terminus.

Comparison with the sequence at the termini of Ad2 DNA

Recently the nucleotide sequence of the first 134 nucleotides at the right-hand terminus and the first 156 nucleotides at the left-hand terminus of Ad2 DNA have been determined[38]. These sequences are nearly completely identical to the Ad5 DNA sequences. At the left-hand terminus the adenosine residue in position 154 in Ad5 DNA has been replaced by a cytidine residue in Ad2 DNA. This difference accounts for the fact, that Ad2 DNA, containing the sequence 5'-AGCGCCGG-3', is cut at this point by the restriction endonucleases *Hpa*II, *Hha*I, and *Hae*II, while Ad5 DNA with the sequence 5'-AGCGACGG-3' is not. Two other differences occur between nucleotides 120 and 130 at the right-hand terminus, viz. the adenosine residue in position 125 and the guanosine residue in position 130 in Ad5 DNA are reported to be cytidine and adenosine residues in Ad2 DNA, respectively. The only difference between Ad2 DNA and Ad5 DNA within the inverted terminal repetition consists of the deletion in Ad2 DNA of the A/T base pair in position 8 of the Ad5 sequence.

The 3'-terminal nucleotide may constitute another difference between the two serotypes. The T4 polymerase exchange reaction, used in earlier experiments to determine the terminal nucleotide sequence in Ad2 and Ad5 DNA, does not discriminate between the structures

$$\begin{matrix} \text{pCpC . . . . pGpApTpG}_{\text{OH}}^{3'} \\ \text{GpGp. . . . .CpTpApCp}_{5'} \end{matrix} \quad \text{and} \quad \begin{matrix} \text{pCpC . . . . pGpApT}_{\text{OH}}^{3'} \\ \text{GpGp. . . . .CpTpApCp}_{5'} \end{matrix}$$ [17].

The present data show, that Ad5 DNA does contain the 3'-terminal guanosine residue. In Ad2 DNA, however, the 3'-terminal nucleotide has been reported to be a thymidine residue. Both DNA's contain the 5'-terminal cytosine residue, which is probably covalently linked to the circularizing protein.

The mechanism of initiation of DNA replication

A model for the initiation of replication of eukaryotic chromosomes has been proposed by Cavalier-Smith[39]. An analogous mechanism, based on self-priming, i.e. the formation of hairpin-loop structures at the termini of the duplex, has been suggested for adenovirus DNA replication[40]. The sequences at the termini of Ad5 DNA presented in this paper, however, do not support this mechanism. Inverted repeats, which would allow the formation of such structures, are not present at the termini. Possibly, nucleotide sequences located beyond nucleotide 194 are involved in the formation of such structures, and

palindromic sequences have been reported to be present about 180 nucleotides from each terminus of Ad2 DNA[41]. On the other hand, the length of the region involved in the formation of such hairpin-loops was estimated to be less than 50-20 nucleotides by *in vivo* labeling experiments[42]. If the termini of the adenovirus DNA are covalently linked in some stage of the replication process, the inverted repetition would form a long palindromic sequence, possibly functioning as recognition signal for one or more proteins involved in the initiation of DNA replication. Alternatively, the protein covalently linked to the 5'-ends of adenovirus DNA[3] itself may serve as a primer in the initiation event.

Transcription of the terminal sequences

The terminal nucleotide sequences described in this paper are probably not transcribed into stabile mRNA. Hybridization studies, in which early as well as late mRNA were annealed to restriction enzyme fragments of adenovirus DNA, have indicated, that regions at both ends of the genome, amounting to 1-2% of the total DNA, are not complementary to cytoplasmic RNA[43].
The sequences certainly do not encode protein products, because stopcodons are present in all possible reading frames, 13 in the right and 12 in the left terminal 194 nucleotides, 8 of which are located within the inverted terminal repetition.

ACKNOWLEDGEMENTS

The authors thank dr. R.J. Roberts for a gift of *Hph*I, drs. C. van Kreyl and H. van Keulen for gifts of RNA polymerase, dr. H. Raué for a gift of ^{32}P-labeled 5S tRNA from *B. licheniformis* and yeast tRNA, and mrs. H.S.C. Lupker-Wille, miss A.M. Pisters, and mr. R. Schiphof for the purification of enzymes. We are grateful to drs. A. Maxam and W. Gilbert for providing us with detailed information on their DNA sequence analysis method before publication, and drs. J.R. Arrand and R.J. Roberts for communicating their sequencing results prior to publication. We thank dr. H.S. Jansz for critical reading of the manuscript. This investigation was supported in part by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organisation for the Advancement of Pure Research (ZWO).

REFERENCES

1. Green,M., Pina,M., Kimes,R., Wensink,P.C., Machattie,L.A., and Thomas,C.A. (1967) Proc.Nat.Acad.Sci. USA *57*, 1302-1309.

2. van der Eb,A.J., van Kesteren,L.W., and van Bruggen,E.E.J. (1969) Biochim. Biophys.Acta *182*, 530-541.
3. Rekosh,D.M.K., Russel,W.C., Bellett,A.J.D., and Robinson,A.J. (1977) Cell *11*, 283-295.
4. Robinson,A.J., and Bellett,A.J.D. (1975) Cold Spring Harbor Symp.Quant. Biol. *39*, 523-531.
5. Levine,A.J., van der Vliet,P.C., and Sussenbach,J.S. (1976) Curr.Topics Microbiol.Immunol. *73*, 67-124.
6. Horwitz,M.S. (1976) J.Virol. *18*, 307-315.
7. Sussenbach,J.S., and Kuyk,M.G. (1977) Virology *77*, 149-157.
8. Weingärtner,B., Winnacker,E.L., Tolun,A., and Petterson,U. (1976) Cell *9*, 259-268.
9. Ariga,H., and Shimojo,H. (1977) Virology *78*, 415-424.
10. Lavelle,G., Patch,C., Khoury,G., and Rose,J. (1975) J.Virol. *16*, 775-782
11. Flint,S.J., Berget,S.M., and Sharp,P.A. (1976) Cell *9*, 559-571.
12. Sussenbach,J.S., Tolun,A., and Petterson,U. (1976) J. Virol. *20*, 532-534.
13. Sussenbach,J.S., and Kuyk,M.G. Virology, in press.
14. Garon,C.F., Berry,K.W., and Rose,J.A. (1972) Proc.Nat.Acad.Sci. USA *69*, 2391-2395.
15. Wolfson,J., and Dressler,D. (1972) Proc.Nat.Acad.Sci. USA *69*, 3054-3057.
16. Roberts,R.J., Arrand,J.R., and Keller,W. (1974) Proc.Nat.Acad.Sci. USA *71*, 3829-3833.
17. Steenbergh,P.H., Sussenbach,J.S., Roberts,R.J., and Jansz,H.S. (1975) J. Virol. *15*, 268-272.
18. Sussenbach,J.S. (1971) Virology *46*, 969-972.
19. Sharp,P.A., Sugden,B., and Sambrook,J. (1973) Biochem. *12*, 3055-3063.
20. Greene,R., and Mulder,C. Unpublished observations.
21. Roberts,R.J., and Myers,P.A. Unpublished observations.
22. Richardson,C.C. (1965) Proc.Nat.Acad.Sci. USA *54*, 158-165.
23. Southern,E.M., and Mitchell,A.R. (1971) Biochem.J. *123*, 613-617.
24. Galibert,F., Sedat,J., and Ziff,E. (1974) J.Mol.Biol. *87*, 377-407.
25. Maxam,A.M., and Gilbert,W. (1977) Proc.Nat.Acad.Sci. USA *74*, 560-564.
26. Carusi,E.A. (1977) Virology *76*, 380-394.
27. Blackburn,E.H. (1975) J.Mol.Biol. *93*, 367-374.
28. Brownlee,G.G., and Sanger,F. (1969) Eur.J.Biochem. *11*, 395-399.
29. Volckaert,G., Min Jou,W., and Fiers,W. (1976) Anal.Biochem. *72*, 433-446.
30. Barrell,B.G. (1971) in Procedures in Nucleic Acid Research (Cantoni,G.L., and Davies,D.R. eds.), vol. 2 pp 751-779, Harper & Row, New York.
31. Sambrook,J., Williams,J. Sharp,P.A., and Grodzicker,T. (1975) J.Mol.Biol. *97*, 369-390.
32. Smith,H.O., and Birnstiel,M.L. (1976) Nucl.Acid Res. *3*, 2387-2398.
33. Maat,J., and Lupker-Wille,H.S.C. Manuscript in preparation.
34. Mulder,C., Arrand,J.R., Delius,H., Keller,W., Petterson,U., Roberts,R.J., and Sharp,P.A. (1974) Cold Spring Harbor Symp.Quant.Biol. *39*, 397-400.
35. Van Ormondt,H., Maat,J., de Waard,A. and van der Eb,A.J., manuscript in preparation.
36. Van Mansfeld,A.D.M., Vereijken,J.M., and Jansz,H.S. (1976) Nucl.Acid Res. *3*, 2827-2844.
37. Ralph,R.K., Connors,W.J., and Khorana,H.G. (1962) J.Amer.Chem.Soc. *84*, 2265-2270.
38. Arrand,J.R., and Roberts,R.J. Submitted for publication.
39. Cavalier-Smith,T. (1974) Nature *250*, 467-470.
40. Wu,M., Roberts,R.J., and Davidson,N. (1977) J.Virol. *21*, 766-777.
41. Padmanabhan,R., Padmanabhan,R., and Green,M. (1976) Biochem.Biophys.Res. Comm. *69*, 860-867.
42. Sussenbach,J.S., and Kuyk,M.G. Submitted for publication.
43. Sharp,P.A., Gallimore,P.H., and Flint,S.J. Cold Spring Harbor Symp.Quant. Biol. *39*, 457-474.