

---

**A method for sequencing restriction fragments with dideoxynucleoside triphosphates**

---

Jan Maat\* and Andrew J.H.Smith+

---

**\*Department of Physiological Chemistry, Leiden University, Leiden, The Netherlands, and  
+MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH,UK**

---

Received 26 October 1978

---

ABSTRACT

A rapid enzymatic approach is described for the sequence analysis of a 5' terminally labelled restriction fragment. It involves limited nicking of the strands of the molecule throughout the sequence by pancreatic DNAase I. The 3' hydroxyl groups exposed by each nick are then used to prime chain extension by DNA polymerase I in four separate reactions. Each reaction uses one of the four chain terminating dideoxynucleoside triphosphates (ddNTPSs), together with the four deoxynucleoside triphosphates (dNTPs). In a single reaction all the 3' ends are terminated in positions of the same base, which is different for each of the four reactions. When the products of these reactions are resolved by gel electrophoresis according to size, a sequence can be deduced from the pattern of radioactive bands. Sequences can be determined onwards from 10-20 residues from the 5' labelled end. The length of sequence which can be determined is only limited by the resolution of the gel.

INTRODUCTION

A number of rapid methods have been developed for DNA sequence analysis. In this paper a new method is described which, like the chemical degradation procedure of Maxam and Gilbert [1], can be used to sequence 5' terminally labelled restriction fragments. This is an enzymatic approach, which relies on the ability of DNA polymerase I to use the chain terminating 2',3'-dideoxynucleoside triphosphates (ddNTPs) as substrates in a nick-translation reaction [2] of the labelled fragment.

Sanger *et al.* [3] have described a method of rapid DNA sequencing. It involves the synthesis of a radioactive complementary copy of a single stranded target sequence with DNA polymerase, using the directly adjacent annealed strand of a restriction fragment as a primer. In four separate reactions, which each contain a different ddNTP and, in addition, the four deoxynucleoside triphosphates (dNTPs) (one or more of which are  $\alpha$ -<sup>32</sup>P-labelled), there is a partial incorporation of the terminator onto the radioactive elongation products. In a single reaction a size range of labelled

oligonucleotides is produced, all with a common 5' end, but with the 3' end terminating in positions of the same base throughout the sequence. Parallel fractionation by denaturing polyacrylamide gel electrophoresis of the products of the four separate reactions, which each contain in turn one of the four ddNTPs, resolves the oligonucleotides of different sizes. These reveal in order the positions of each base, allowing a sequence to be determined.

The drawback of this method is that it requires a pure single stranded template. In the approach described here, use is made of the ddNTPs to sequence double stranded DNA in the form of restriction fragments, labelled at only one 5' end.

In the protocol a 5' <sup>32</sup>P-labelled fragment is incubated in the presence of DNA polymerase I and pancreatic DNAase I in four separate reactions. Each reaction contains all four dNTPs and, in turn, one different ddNTP. The pancreatic DNAase I introduces nicks, the 3' hydroxyls of which serve as priming points for DNA polymerase I. Chain extension then proceeds from the 3' end of every nick, leading to a base specific chain termination in each reaction, as in the method of Sanger *et al.* [3]. The products of the reactions are fractionated on denaturing polyacrylamide slab gels, resolving the oligonucleotides with the common labelled 5' end.

Although the nicking and subsequent chain extension occurs on both strands, since only one is labelled a pattern of radioactive bands is produced for only that strand, allowing its sequence to be deduced. Nicking of the fragment by DNAase I does not occur at every residue along its length but the fact that both deoxy- and dideoxy- NTPs are present in each reaction ensures that chain extension continues through several residues of the same base from the site of the nick. Every residue in the sequence should therefore give rise to a band.

### EXPERIMENTAL PROCEDURES

#### Materials

Bovine pancreatic DNAase I, grade I, Boehringer Mannheim (FRG); solution of 1 mg/ml in distilled water, stored at -20°C.

DNA polymerase I, grade I, Boehringer Mannheim (FRG); 5 units/μl.  
ddNTPs and dNTPs, P.-L. Biochemicals (Milwaukee, Wis., USA); dissolved to a concentration of 10 mM in TE buffer (1 mM Tris.HCl, pH 7.6, 0.05 mM EDTA) and stored at -20°C.

T4 polynucleotide kinase was made according to a modification [4] of the

procedure of Richardson [5].

$\gamma$ - $^{32}$ P-ATP, 3000 Ci/mole, the Radiochemical Centre (Amersham, UK).

DNA molecules, 5' terminally labelled, were prepared from adenovirus type 5 DNA which was a gift of B.M.M. Dekker.

All other materials used are as described in references 3, 4 and 6.

#### 5' terminal labelling of restriction fragments

2-5 pmoles of restriction fragments are labelled with T4 polynucleotide kinase and  $\gamma$ - $^{32}$ P-ATP as described in reference 4. In order to obtain DNA molecules with only one  $^{32}$ P-labelled 5' terminus, the fragments are further digested with a restriction enzyme of different specificity, which cleaves at one or more sites within the fragment. The resulting mixture of labelled fragments is then fractionated by electrophoresis on polyacrylamide slab gels (400 x 180 x 1.5 cm) in 90 mM Tris-borate pH 8.3, 1 mM EDTA. The radioactive fragments are recovered from the gel with the elution procedure described by Maxam and Gilbert [1] and precipitated with ethanol without carrier tRNA. The precipitate is taken up in 10  $\mu$ l water.

#### Sequencing procedure

To 2-5 pmoles of labelled fragment in 10  $\mu$ l is added 1  $\mu$ l of pancreatic DNAase I at a concentration of 0.1  $\mu$ g/ml. In four capillaries 2.5  $\mu$ l of this mixture is added to 2.5  $\mu$ l of each of the four base specific mixtures and 0.5  $\mu$ l of DNA polymerase I (= 2.5 units). The composition of the base specific mixtures is given in Table 1.

Table 1. Composition of the base specific mixtures			
A	G	C	T
ddATP 1 mM	ddGTP 1 mM	ddCTP 1 mM	ddTTP 1 mM
dATP 2 $\mu$ M	40 $\mu$ M	40 $\mu$ M	40 $\mu$ M
dGTP 40 $\mu$ M	2 $\mu$ M	40 $\mu$ M	40 $\mu$ M
dCTP 40 $\mu$ M	40 $\mu$ M	2 $\mu$ M	40 $\mu$ M
dTTP 40 $\mu$ M	40 $\mu$ M	40 $\mu$ M	2 $\mu$ M

Each mix also contains: 10 mM Tris.HCl pH 7.6, 10 mM MgCl<sub>2</sub>, 10 mM dithiothreitol, 100 mM NaCl.

After incubation for 30 min at room temperature, the reaction is terminated by adding 15  $\mu$ l of formamide, containing 10 mM EDTA, 0.05% XCFE and 0.05% BFB.

### Gel electrophoresis

5  $\mu$ l of each sample in formamide dye is then heat denatured at 100°C for 3 min and loaded in a 5 mm wide well of an 8% polyacrylamide slab gel (400 x 180 x 0.35 mm) containing 7 M urea. Gel electrophoresis is carried out according to Sanger and Coulson [6]. Enough material is left to allow several loadings on the gel at different times after the electrophoresis has been started. In this way expansion of the band pattern in different regions of the sequence is obtained. On an 8% gel, for example, up to 300 nucleotides from the 5' end can be resolved.

The gels are wrapped in Saranwrap and autoradiographed at -20°C on pre-flashed Sakura QH no-screen medical X-ray film. In some cases enhanced autoradiographic detection is obtained with Ilford fast tungstate 30 x 40 cm intensifying screens; exposure is at -70°C.

### RESULTS

Figures 1 and 3 show examples of sequences from part of the HindIII fragment E of adenovirus type 5 DNA obtained using the procedure described in this paper. The HindIII fragment E (between map positions 8.0 and 17.0) was 5' labelled and cleaved with HinfI to yield labelled fragments of 630 and 440 nucleotides long. The sequence deduced in Figure 1 using the 440 nucleotide long fragment starts at position 2816, 10 nucleotides away from the 5' labelled end, and reads through to position 3035 (counting from the left terminus of adenovirus type 5 DNA in the l-strand). Figure 2 shows the result of the Maxam and Gilbert procedure applied to the same fragment. The sequences obtained from Figures 1 and 2 are in agreement. Figure 3 shows a part of the sequence of the r-strand of the 630 nucleotide long fragment. The label resides at the HindIII site at map position 17.0. The sequence can be deduced 17 nucleotides from the 5' end to a point 200 nucleotides away. It also has been shown to be in full agreement with the sequence deduced employing the Maxam and Gilbert method on the same fragments (results not shown).

### DISCUSSION

The nick-translation sequencing procedure described here has been shown to be reliable, rapid and simple. The reaction makes use of the intact DNA polymerase I enzyme which is necessary for carrying out nick-translation; the 5'-3' exonuclease activity of the enzyme hydrolyses the 5' end of the nick as the 3' end is simultaneously extended by the 5'-3' polymerase activity. Experience has shown that there is no significant loss of labelled material

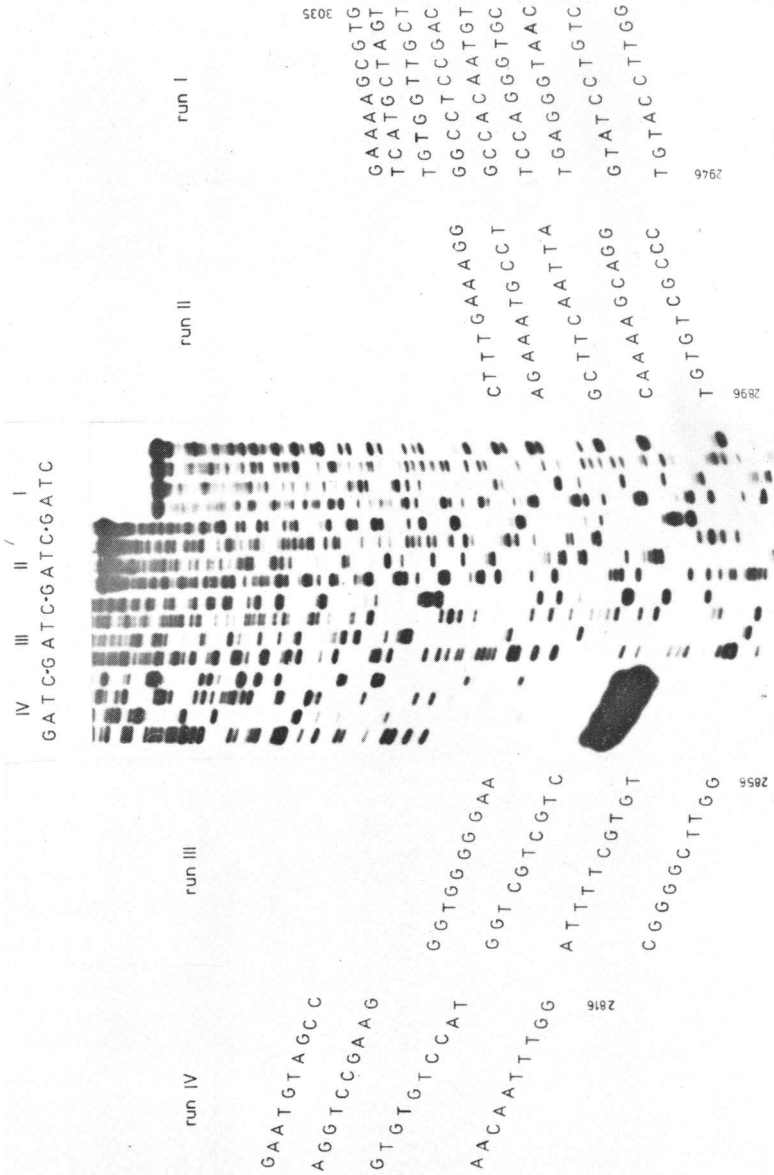


Figure 1. Result of the procedure described in this paper applied to a 440 nucleotide long fragment from a HindIII digest of the 5' end labelled HindIII-E fragment of adenovirus type 5 DNA. Samples from each base specific reaction mixture were loaded every 2 h. Electrophoresis was carried out at a constant current of 30 mA. Faint bands in the original autoradiograph are not visible in the photograph shown here.

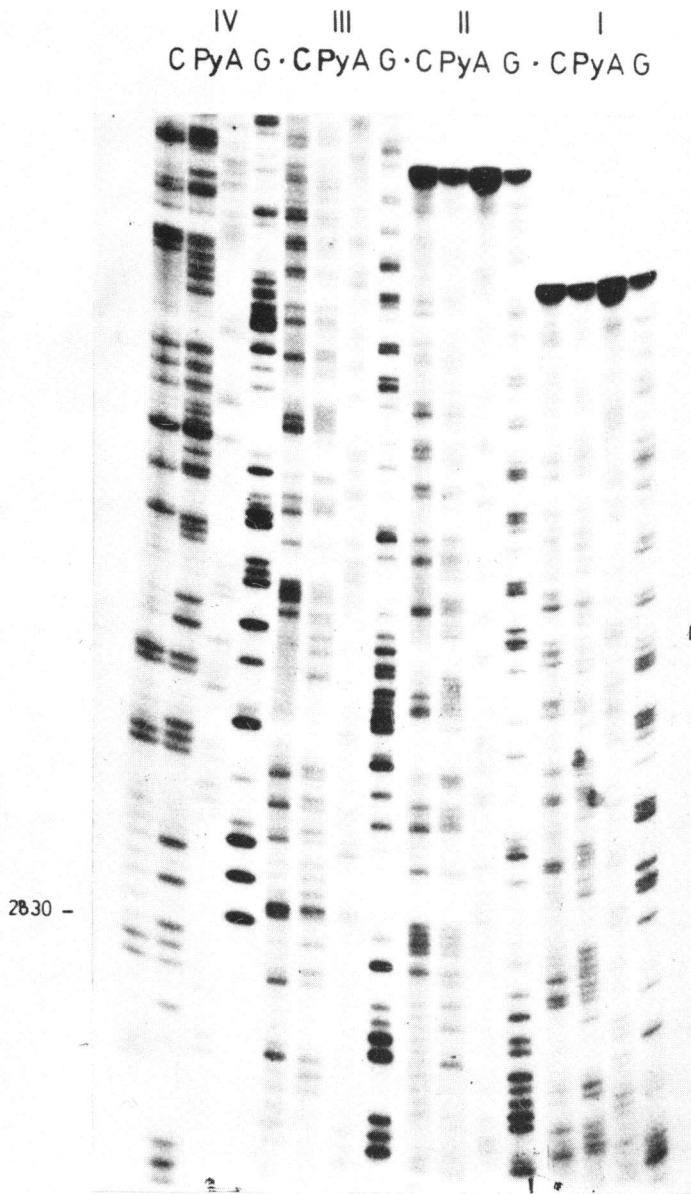


Figure 2. Result of the procedure of Maxam and Gilbert [1] applied to the same fragment as in Figure 1. Samples from each base specific chemical degradation were loaded in formamide and electrophoresed as described in the legend to Figure 1.

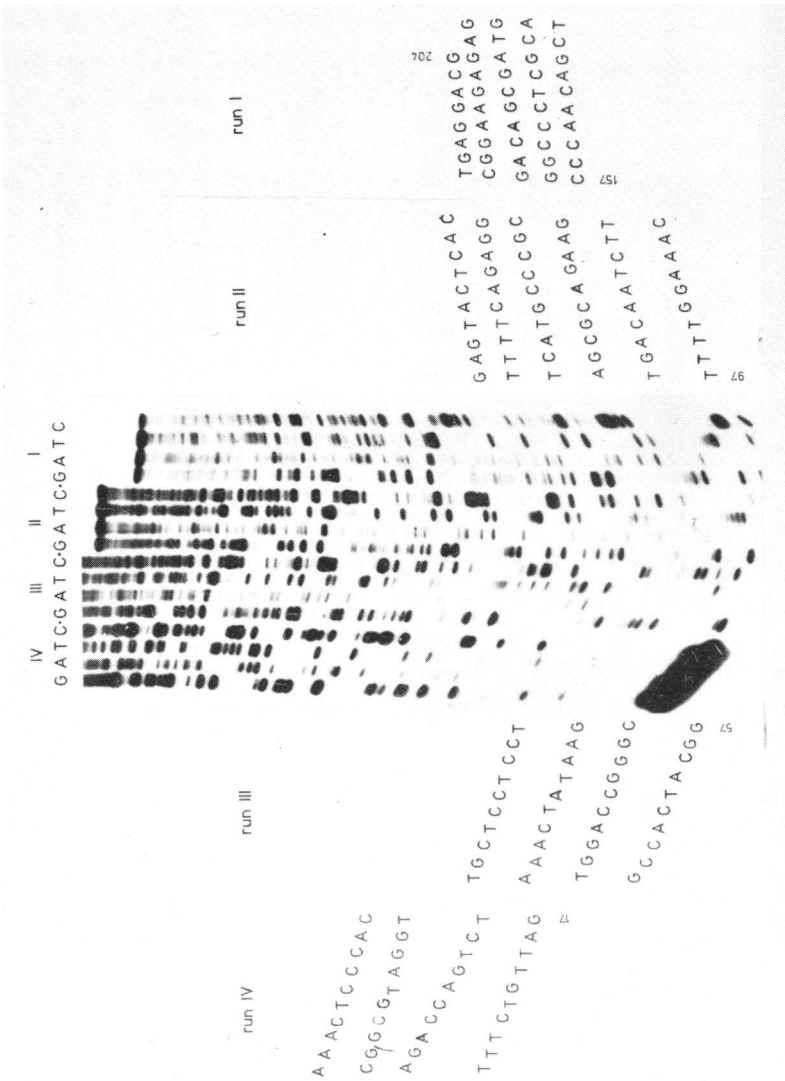


Figure 3. Result of the procedure described in this paper applied to a 630 nucleotide long fragment from a HinfI digest of the 5' end labelled HindIII-E fragment of adenovirus type 5 DNA. Electrophoresis was carried out as described in the legend to Figure 1. Faint bands in the original autoradiograph are not visible in the photograph shown here.

due to hydrolysis of the labelled 5' end by the 5'-3' exonuclease activity. Preliminary experiments indicate that the Klenow subfragment of DNA polymerase I may also be suitable for this procedure. It may somehow be capable of displacing the 5' end of the nick as polymerisation proceeds from the 3' end.

The sequencing approach can be applied to DNA fragments of any length. The sequence that can be deduced is only limited by the resolving power of the gel electrophoresis (which is at present 300-400 nucleotides long). The conditions described in the protocol are, on the whole, suitable for application to fragments of 100-500 nucleotides long. For longer fragments the DNAase I concentration may need to be increased in order to obtain most of the labelled DNA within a size range of oligonucleotides below 300-400 nucleotides long.

It is apparent from the results obtained using the nick-translation procedure that there are a number of peculiarities of the method, some of which are evident in Figures 1 and 3.

- (a) Every residue gives rise to a band apart from the first 10-20 residues from the 5' labelled end. In some cases a sequence can be deduced directly after the position of the strong band which runs across all four tracks and appears to have a size corresponding to an oligonucleotide of chain length of 8-10 residues. No bands are seen corresponding to oligonucleotides of smaller chain length than this. Generally bands corresponding to the first few residues of the sequence are very weak in intensity. In other cases not even weak bands are seen directly after the position of this strong artefact band and sometimes sequences cannot be deduced until some 10 nucleotides from this position. We do not completely understand the reasons for these phenomena.
- (b) Throughout the sequence band intensity is very variable. This particularly happens in a run of the same residue, where it decreases in a 5' to 3' direction. This is thought to be due in part to the fact that nicking of the fragment does not occur randomly throughout the sequence and is especially infrequent in runs of the same residue. Chain extension through a run in most cases therefore takes place from the 3' end of a nick lying adjacent 5' to the run. Partial incorporation of the chain terminators at each residue in a run results in an exponential fall-off (in a 5' to 3' direction) of the numbers of oligo-



nucleotides terminated at those residues.

- (c) Although not seen in Figures 1 and 3, DNA polymerase I sometimes has difficulties in copying certain sequences of the template strand. This leads to pile-ups of bands lying across all tracks of the gel. There seems to be no consistent feature of the sequences at these points.

In spite of the disadvantages mentioned above and the obvious limitation of the method - that it can only be applied to 5' terminally labelled fragments (as opposed to 5' or 3' end labelled fragments, as is the case with the "Maxam-Gilbert" procedure) - we feel that this procedure may have use as a general method of rapid DNA sequencing and have some advantages over existing methods. It has the advantage over the "classical" chain terminator method that it can be applied directly to double stranded DNA. In comparison with the chemical cleavage method of Maxam and Gilbert, the manipulations involved are less laborious and more rapid, and the bands produced in thin gels are sharper, enabling longer sequences to be read.

#### ACKNOWLEDGEMENTS

The authors thank Drs. F. Sanger, A. de Waard and H. van Ormondt for advice and help in the course of their work.

The research reported in this manuscript was supported in part by the Medical Research Council and in part by the Netherlands Foundation for Chemical Research (S.O.N.) with financial aid from the Netherlands Organisation for Advancement of Pure Research (Z.W.O.)

A.J.H. Smith was supported by a Research Studentship from the Medical Research Council. J. Maat was supported by the Koningin Wilhemmina Fonds.

#### REFERENCES

1. Maxam, A.M. and Gilbert, W. (1977) Proc. Nat. Acad. Sci. USA 74, 560-564.
2. Kelly, R.B., Cozzarelli, N.R., Deutscher, M.P., Lehman, I.R. and Kornberg, A. (1970) J. Biol. Chem 245, 39-45.
3. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Nat. Acad. Sci. USA 74, 5463-5467.
4. van Ormondt, H., Maat, J., de Waard, A. and van der Eb, A.J. (1978) Gene, in press.
5. Richardson, C.C. (1972) Proc. Nucleic Acid Res. 2, 815-828.
6. Sanger, F. and Coulson, A.R. (1978) FEBS Letters 87, 107-110.