Volume 10 Number 6 1982 Nucleic Acids Research

New rapid methods for DNA sequencing based on exonuclease III digestion followed by repair synthesis

Li-He Guo and Ray Wu

Section of Biochemistry, Molecular and Cell Biology, Cornell University, Ithaca, NY 14853, USA

Received 2 November 1981; Revised and Accepted 16 February 1982

ABSTRACT

We describe improve enzymatic methods for sequencing DNA. They are based on partial digestion of duplex DNA with exonuclease III to produce DNA molecules with 3' ends shortened to varying lengths, followed by repair synthesis to extend and label the 3' ends. After asymmetrical cleavage of the DNA with a restriction enzyme, the labeled products are separated by gel electrophoresis and the sequence read from the autoradiogram. The entire procedure, beginning with unrestricted DNA and followed through gel electrophoresis, takes only one day for sequencing both strands of the DNA molecule. These methods are especially suitable for sequencing DNA cloned in plasmid vectors, and they greatly extend the usefulness of the dideoxynucleotide chain termination method of Sanger et al. (Proc. Natl. Acad. Sci. USA 74, 5463, 1977). Using these methods we have determined the sequence of a 410 base pair fragment which includes the yeast SUP3 tyrosine tRNA gene.

INTRODUCTION

The development of rapid methods for sequencing DNA has led to recent important advances in understanding gene structure and function. These procedures fall into two categories: chemical and enzymatic.

The method developed by Maxam and Gilbert (1) uses chemical reactions to cleave terminally labeled DNA at specific nucleotides producing fragments of different lengths.

Sanger and co-workers (2,3) have developed two enzymatic procedures which depend on primed synthesis with DNA polymerase to generate fragments of varying lengths. One problem with these latter methods is their requirement for single-stranded DNA as a template, as well as a specific primer for the synthesis. Several procedures have been developed to simplify the isolation of template and primer DNAs. One involves cloning DNA fragments into the double-stranded replicative form of M13 bacteriophage and using single-stranded DNA from the phage as a template (4,5,6). Another procedure uses Escherichia coli exonuclease III (7) to digest duplex DNA, producing

0305-1048/82/1006-2065$2.00/0

partially single-stranded DNA template suitable for hybridization to an added primer (8-11). In no report was the exonuclease III used to generate the primer for sequence analysis. Seif et al. (12) have developed a method similar to Sanger's "plus and minus" method (2), but which obviates the necessity of isolating either primer or single-stranded template DNAs. However, it requires terminal labeling and gel fractionation of DNA fragments prior to sequence analysis.

These methods for DNA sequencing rely on electrophoresis on polyacrylamide gels to separate fragments which share a common end but vary in length (1,2,3,11). However, they also all require at least one additional preparative gel electrophoresis step to isolate either the primer (2-5,10,11), or the fragment to be sequenced (1,12).

In this paper we describe two closely-related enzymatic methods which are simpler and more rapid than those mentioned above. They start with uncut plasmid, need no added primer, and require only a single gel electrophoresis step to obtain the DNA sequence. These methods are made possible by the observation that under controlled digestion of DNA at 23°C, exonuclease III removes synchronously approximately 10 nucleotides/minute from 3' termini (10).

EXPERIMENTAL DESIGN

These methods include the following steps (Fig. 1):

(a) The digestion of the DNA (e.g., a plasmid carrying a cloned DNA to be sequenced) with a restriction enzyme which cuts at only one site to give a linear molecule. The enzyme may give either 3' recessed ends or blunt ends. 3' Protruding ends cannot be efficiently cut by exonuclease III (Yang and Wu, unpublished). If there is a need to remove 3' protruding ends, 1-2 units of DNA polymerase (Klenow fragment) are added to the restriction enzyme reaction mixture and incubated at 37° for 15-30 min after completion of the restriction enzyme digestion.

(b) The controlled digestion of duplex DNA with exonuclease III to produce shortened 3' ends, and long 5'-ended single strands. For Method I, aliquots are removed from the reaction mixture at regular intervals and pooled, yielding fragments of every length over a broad range. For Method II, the reaction is stopped after a fixed time, yielding DNA with long 5'-ended single strands.

(c) The incorporation of labeled nucleotides into the 3' termini of the exonuclease III-digested DNA. These molecules are ideal template-primer

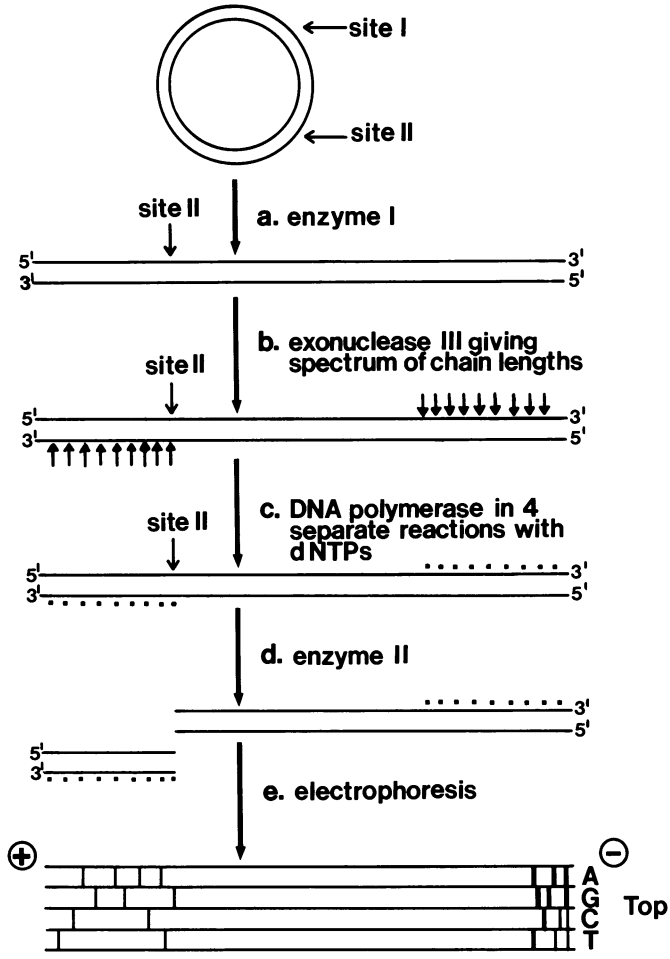


Figure 1. The principle of the exonuclease III-repair synthesis Method I for sequencing DNA. A fragment of DNA to be sequenced is cloned into a plasmid (represented by a double circle) between restriction sites I and II. In this example, the plasmid is first digested with restriction enzyme I to give a linear DNA. The digestions by exonuclease III (step b) gave a family of DNA molecules with 3' ends shortened to different extents. The closely spaced arrows indicate the end point of digestion of different molecules of DNA. The digested DNA in step (b) is distributed into 4 tubes, and the 3' ends of the DNA are labeled according to Method I (step c) using a different [α-^{32}P] dNTP in each of the 4 tubes. The squares (▪) after steps c and d indicate the positions of terminal labeling of this population of chains. After digestion with a second restriction enzyme (step d), the DNA fragments are fractionated on a denaturing polyacrylamide gel (step e). The shorter fragments are well separated on the lower part of the gel (left-hand side of gel), and the longer fragments are retained near the top of the gel (right-hand side of gel). Four lanes represent DNA fragments terminated with each of the 4 different dNTPs.

systems in which the shortened strands with 3' ends serve as the primer. These may be used in two different ways for the two different labeling methods (Table 1). Method I uses the 4 different [α-^{32}P] deoxynucleoside triphosphates in 4 separate reactions (13). The only fragments labeled in each reaction will be those in which the polymerase can add the specific [α-^{32}P] dNTP present, where the first base on the single-stranded part of the template is complementary. Thus, the only bands on the autoradiogram in any lane will be those ending in that particular base. Method II uses Sanger's chain termination method (3) to generate labeled fragments of specific lengths. The 4 dideoxy NTPs (ddNTPs) are used separately in each of 4 reactions, together with all 4 dNTPs including one [α-^{32}P] dNTP.

(d) The digestion of labeled DNA molecules asymmetrically with a second restriction enzyme to produce two families of labeled fragments of very different lengths.

(e) Separation of the members of these 2 families of DNA by high

Table I
Composition of the repair synthesis mixtures

Method	I				II			
Reaction	A	G	C	T	A	G	C	T
ExoIII-DNA mix (µl)	4	4	4	4	3	3	3	3
[α-^{32}P]dATP[1] (µl)	1				2	2	2	2
[α-^{32}P]dGTP[1] (µl)		1						
[α-^{32}P]dCTP[1] (µl)			1					
[α-^{32}P]dTTP[1] (µl)				1				
ddA-dN mix[2] (µl)					1			
ddG-dN mix[3] (µl)						1		
ddC-dN mix[4] (µl)							1	
ddT-dN mix[5] (µl)								1

[1]The specific activity of [α-^{32}P]dNTP is 410 Ci/mmole (1 mCi/ml). The [α-^{32}P]dNTP is dried down in a tube before the addition of the exonuclease III digested DNA to this tube. After the addition of nucleotides, one µl of DNA polymerase (Klenow fragment, 0.2-0.4 units/µl) is added to each tube.

[2]ddA-dN mix contains 250 µM dGTP, 250 µM dCTP, 250 µM dTTP, 2.5 µM dATP, and 0.6 mM ddATP

[3]ddG-dN mix: 250 µM dCTP, 250 µM dTTP, 20 µM dGTP, and 1 mM ddGTP

[4]ddC-dN mix: 250 µM dGTP, 250 µM dTTP, 15 µM dCTP, and 1 mM ddCTP

[5] ddT-dN mix: 250 µM dGTP, 250 µM dCTP, 20 µM dTTP, and 2 mM ddTTP

resolution electrophoresis on a denaturing polyacrylamide gel.

MATERIALS AND METHODS

Materials

E. coli exonuclease III was from Bethesda Research Laboratories, Inc. *E. coli* DNA polymerase I (Klenow fragment) was from New England Biolabs, Inc. The restriction enzymes were from BRL or New England Biolabs.

The [α-^{32}P] dNTPs (410 Ci/mmol, 1 mCi/ml) were from Amersham Corporation. The dNTPs and ddNTPs were from P-L Biochemicals, Inc.

Plasmid pYT2 was constructed by ligating a 410 base pair *EcoRI-Sau3AI* fragment from plasmid pWJ3 (supplied by Rod Rothstein) containing the yeast SUP3 tyrosine tRNA gene (15) to the large *EcoRI-BamHI* fragment of plasmid pBR322. Plasmid DNA was isolated and purified by a modification of the method of Birnboim and Doly (16). It is important to check whether the DNA is free of contaminating DNase by incubating the DNA at 37° in a restriction enzyme buffer without the addition of enzyme. The intactness of DNA is then examined by agarose gel electrophoresis on a 3" x 4" glass plate.

Methods

(a) *Digestion of DNA with a restriction enzyme*. Five µg of plasmid DNA (for pYT2, about 1.7 pmole) were added to a 1.5 ml Eppendorf tube, which contained 1.5 µl of a 10 fold concentration (10 X) repair synthesis buffer (500 mM Tris-HCl, pH 7.6, 100 mM $MgCl_2$, 100 mM dithiothreitol, 500 mM KCl) and 5-15 units of a restriction enzyme. The volume was adjusted to 15 µl with distilled water and the mixture incubated at 37°C for 15 to 60 minutes.

(b) *Digestion of DNA with exonuclease III*. To the reaction mixture from step (a) 38 µl of H_2O, 6 µl of 10 X exonuclease III buffer (660 mM Tris-HCl, pH 8.0, 770 mM NaCl, 50 mM $MgCl_2$, 100 mM dithiothreitol) and 1 to 4 µl of exonuclease III (25 units/µl) were added (see Table II). Digestion was at room temperature (23°C) for 10 to 150 min. For Method I, aliquots were removed from the reaction mixture at 2 minute intervals during the appropriate digestion period and transferred into another tube containing 15 µl of ice cold stop solution (0.25 M EDTA, 1.5 M sodium acetate) and 75 µl of phenol saturated with 1 M Tris-HCl, pH 8.0. For Method II, only a single incubation period is needed. The aqueous-phenol mixture was vortexed and centrifuged at 12,000 x g (Eppendorf centrifuge) for 2 min. The aqueous phase was transferred to another tube and extracted with 100 µl of chloroform-isoamyl alcohol (24:1). After centrifugation for 1 min, the aqueous phase was removed to another tube and the DNA precipitated by adding

200 µl of ethanol. The tube was chilled for 5 min in dry ice, and centrifuged at 0°C for 5 min. The supernatant was removed and 300 µl of ethanol was added to rinse the DNA pellet. The tube was centrifuged 3 min and the supernatant removed. The DNA pellet was dried under vacuum for 5 min.

(c) Labeling of partially-digested DNA. The [α-^{32}P] dNTPs as shown in Table 1 were first dried down in each tube under vacuum.

For Method I, the pellet from step (b) was resuspended in 2 µl of 10 X repair synthesis buffer and 16 µl of H_2O. Aliquots of 4 µl were pipetted into each tube containing one of the 4 [α-^{32}P] dNTPs, and 0.3 unit of DNA polymerase (Klenow fragment) was added. The reactions were incubated for 15 min at room temperature.

For Method II, the DNA pellet was resuspended in 6 µl of 10 X repair synthesis buffer and 54 µl of H_2O. Aliquots of 3 µl were added to 4 tubes containing the dried [α-^{32}P] dATP. To these tubes were added 1 µl of the appropriate ddNTP-dNTP mix (Table 1) and 0.3 units DNA polymerase (Klenow fragment). Synthesis was at 37°C for 15 min to add labeled nucleotide and to extend the DNA chain up to several hundred nucleotides in length. A 1 µl chase solution (0.5 mM unlabeled dATP) was added to each tube, and synthesis allowed to continue an additional 15 min at 37°C.

All reactions were stopped by heating at 70°C for 10 min.

(d) Digestion of DNA with a second restriction enzyme. After cooling the samples from step (c), the second restriction enzyme (2 units) was added to each tube. The tubes were incubated at 37°C for 15 min. The reactions were stopped by the addition of 1.5 µl of 1 N NaOH, 100 mM EDTA, and 5 µl of 10 M urea containing 0.3% bromophenol blue and xylene cyanol. Alternatively, the DNA sample is precipitated by the addition of 1 µl tRNA (10 µg/µl), 2 µl of 0.25 M EDTA -1.5 M NaoAC and 25 µl of ethanol; after centrifugation, the DNA is suspended in 3 µl of 0.2 N NaOH -10 mM EDTA and 3 µl of 10 M urea containing dyes.

(e) Gel electrophoresis. The samples (1-2 µl) were loaded on an 8% polyacrylamide gel (may vary between 4% and 20% depending on the size of DNA fragments to be fractionated), 40 cm x 34 cm x 0.06 cm in dimension, 8 M urea, and 50 mM Tris-borate. Electrophoresis was at 20 mA constant current for 3-8 hours with 1 to 3 loadings. After electrophoresis, the gel was either exposed to x-ray film directly or transferred to a sheet of Whatman 3 MM paper and dried (to give sharper bands) at 90°C under vacuum for 30 min before exposing to x-ray film. In some cases an intensifying screen was

used and the film exposed at -70°C.

RESULTS

Exonuclease III digestion of DNA. In order to determine the rate of digestion, linearized plasmid DNA was subjected to exonuclease III digestion and samples taken at several points in the digestion. These samples were labeled using Method I, but adding two ^{32}P nucleotides to the 3' ends of the shortened strands. This was followed by adding a second restriction enzyme, which digests the DNA asymmetrically, and samples were loaded on an 8% sequencing gel, along with a size marker. Figure 2(a) shows that approximately 10 nucleotides/minute are removed at 23°C. The rate does decrease during longer digestions, but this can be compensated for by using higher levels of exonuclease III (Figure 2b). Table 2 shows the number of units of enzyme/ pmole of DNA required for different extents of digestion. It was found that sampling every two minutes during the appropriate digestion period produced a DNA population which included all lengths within the desired range.

DNA Sequencing. Plasmid pYT2 was used to test the new methods. Figure 3(a) shows part of the strategy used to sequence the 410 base pair EcoRI-Sal3AI fragment of pYT2 from a point within the tRNA gene. The autoradiogram shown in Figure 3(b) contains sequences between EcoRI and SmaI site on one strand of DNA (nucleotides L98-L200) and between BstNIand SmaI site on the other strand of DNA (nucleotides U69-U-18) in Figure 4. Results obtained using both Methods I and II are shown in parallel. Both methods gave clear patterns and the reading of the sequence is relatively easy. In Method I, when the same nucleotide is repeated, frequently the band corresponding to the shortest fragment is weak or absent. Because of uniform spacing, reading of the sequence is possible. All but about 30 nucleotides around the SmaI site were determined in this experiment. These nucleotides plus others have been determined by using a strategy described below.

Figure 5 shows the nucleotide sequence around and across the SmaI site obtained by cutting the DNA at a point outside of the region of interest. The pYT2 DNA was digested first with SalI (see Fig. 3a) and then with exonuclease III for 70 min. This was followed by digestion with EcoRI and the sequence of the lower strand (nucleotides L10 to L155) in Figure 4 was obtained by using Method II. The sequence of the complementary strand was obtained (data not shown) by cutting the DNA first with EcoRI. After exonuclease III digestion and labeling by Method II, it was cut with BstNI.

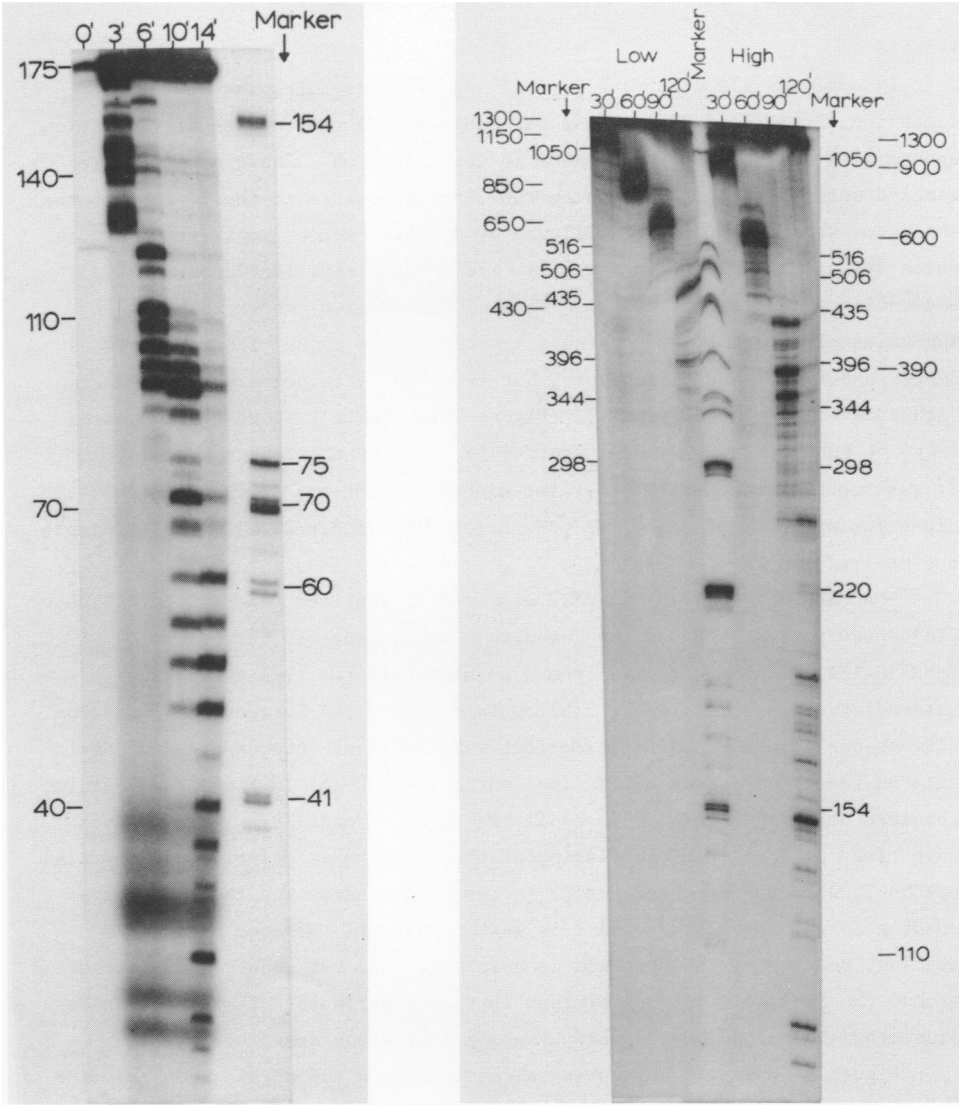


Figure 2. The rate of digestion of duplex DNA by exonuclease III (ExoIII). (a) pYT2 DNA linearized by SmaI was digested with 15 units ExoIII/pmole DNA and aliquots taken at the times indicated. These samples were labeled by Method I, but two labeled dNTPs were used (dATP and dTTP), cut with a second enzyme (EcoRI) and loaded on a sequencing gel with HinfI digested pYT2 as a size marker. The initial length of the DNA fragment was 175 base pairs. (b) A longer digestion of a linearized DNA. "Low" is 15 units/pmole DNA and "high" is 45 units/pmole. The second enzyme was AvaI and the initial length of the DNA fragment was 1300 base pairs.

Table II
Digestion of DNA with exonuclease III

To remove approximately 10 nucleotides* per minute from each 3' end of DNA, the concentration of E. coli exonuclease III (ExoIII) needed is as follows:

Number of nucleotides removed from each end of DNA	Incubation (min)	ExoIII(u/pmole DNA)
100-250	10-25	15-20
250-500	25-50	20-25
500-750	50-75	25-30
750-1000	75-100	30-35
1000-1500	100-150	35-45

A unit of exonuclease III is the amount of enzyme that liberates 1 nmole of mononucleotides from a sonicated DNA substrate in 30 minutes at 37°C. BRL ExoIII, lot 2429, was used.

*The salt concentration (sodium and potassium ions) in ExoIII digestion solution influences appreciably the rate of ExoIII digestion of DNA. If the final salt concentration including that in enzyme preparation is 90 mM, ExoIII can remove approximately 10 nucleotides per minute from each 3' end of DNA. If the salt concentration is lower than 90 mM, the rate of ExoIII digestion is more rapid, but the DNA products are less desirable primers for Method II. If the salt concentration is 105 mM or 125 mM, the rate is about 7.5 or 5 nucleotides per minute, respectively from each 3' end of DNA.

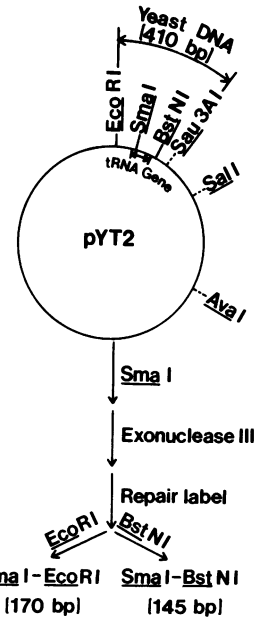
Thus, by selecting two different restriction enzymes for the first digestion, the sequence of both strands of the cloned DNA was established. With two or three loadings of samples for gel electrophoresis, longer sequence information was obtained (data not shown). In fact, the entire 410 base apir sequence can be determined by making the first cut of the DNA outside the tRNA gene (by not using the SmaI site within the gene), and by method II alone.

The complete sequence of the 410 base pair EcoRI-Sau3AI fragment of pYT2 is shown in Figure 4. The entire sequence was unambiguously determined using the methods presented in this paper. This sequence includes the yeast SUP3 tyrosine tRNA gene analyzed from both directions as detailed in Figures 3 and 5. A portion of this yeast sequence was identical to the sequence in tWJ3, which has been determined by R. Rothstein (personal communication). Although the coding sequence of SUP3 tyrosine tRNA gene is identical to that of SUP4 tyrosine tRNA gene (15), the 5' and 3' non-coding sequence is different except that the first 11 base pairs of the 5' non- coding sequence in front of both tyrosine tRNA genes are rich in A/T pairs.

These methods depend upon the availability of specific restriction sites that occur once or twice in the plasmid. These critical sites may be as far as 1300 base pairs from the region of interest as indicated from the following experiment. Figure 6 gives the nucleotide sequence near the AvaI site in pYT2 which is 1300 base pairs away from the site of the initial cut at the SmaI site. The DNA was digested with exonuclease III for 130 min. The sequence between base pairs 1237 and 1341 of the pBR322 (numbering from the EcoRI site) part of pYT2 was determined. This sequence agrees with the published sequence of Sutcliffe (17).

DISCUSSION

E. coli exonuclease III is a 3'→5' exonuclease with little base specificity. Using this enzyme, linear duplex DNA can be degraded almost synchronously from the 3' ends of both strands. The rate of digestion can be controlled conveniently, and averages about 10 nucleotides/min at each end at 23°C. If as in Method I aliquots of the reaction are removed at regular intervals during the digestion and then pooled, the molecules will have 3' ends of every length within a range depending on the length of the digestion. If the entire reaction is stopped at once as in Method II, the fragments will be of approximately uniform length. The DNA thus digested



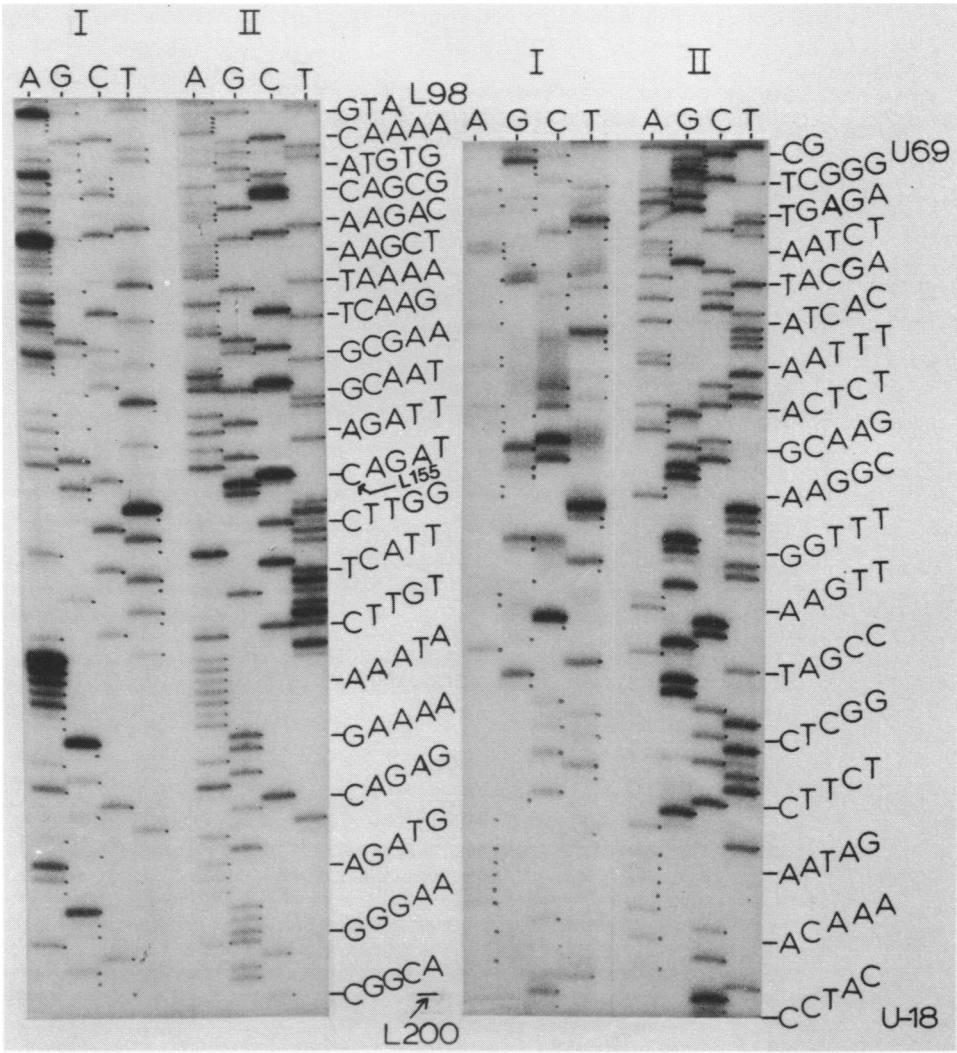


Figure 3. Sequencing the SmaI-EcoRI and SmaI-BstNI fragments of pYT2. (a) Schematic representation of the sequencing strategy used. pYT2 (410 base pairs yeast DNA cloned between EcoRI and BamHI sites of pBR322) was first cut with SmaI and then digested for 16 min with exonuclease III. The DNA was divided and labeled using Methods I and II, and then each portion was divided and digested separately with EcoRI and BstNI. (b) Autoradiogram of the sequencing gel. The SmaI-EcoRI fragment is at the left and shows nucleotides 98 to 200 on lower strands in Fig. 4. The SmaI-BstNI fragment is at the right and shows nucleotides 69 to -18 of upper strand in Fig. 4. Fragments were labeled by Method I (I) and Method II (II).

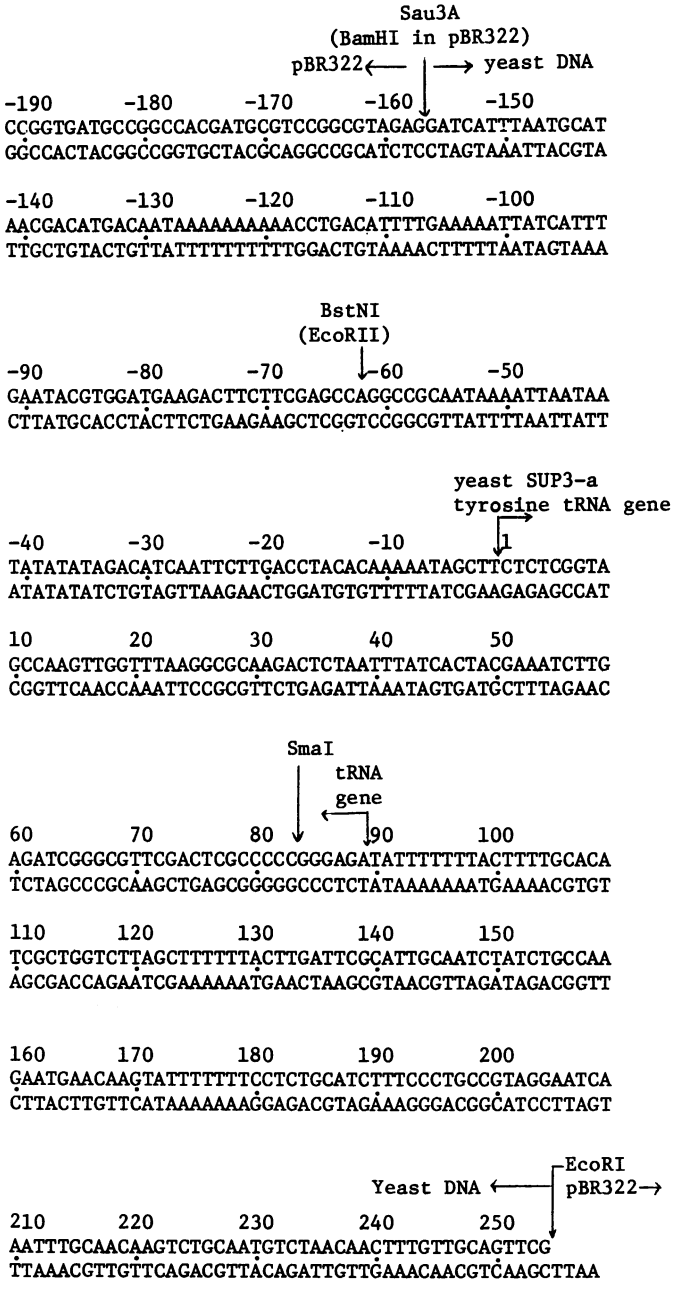


Figure 4. Sequence of the yeast SUP3-a tyrosine tRNA gene and its 5' and 3' flanking sequence in pYT2.

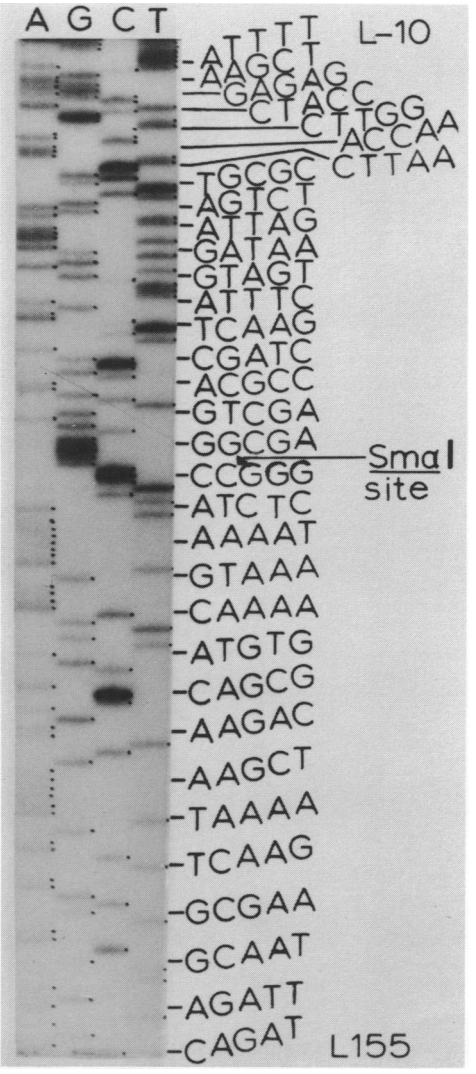


Figure 5. Autoradiogram of a sequencing gel. pYT2 DNA was cut with SalI followed by a 70 min exonuclease III digestion. The DNA was labeled using Method II, then digested with EcoRI. The SmaI site can be seen at the middle of the figure. The sequence below the SmaI site is identical to that shown at the left of Fig. 3(b), for nucleotides L155-L98, and above the site it is complementary to the sequence shown at the right of Fig. 3(b).

can be used for sequencing by the two methods presented here. It is easy to determine the rate of degradation as shown in Figure 2, and when first applying the method to a specific fragment, it may be desirable to do so.

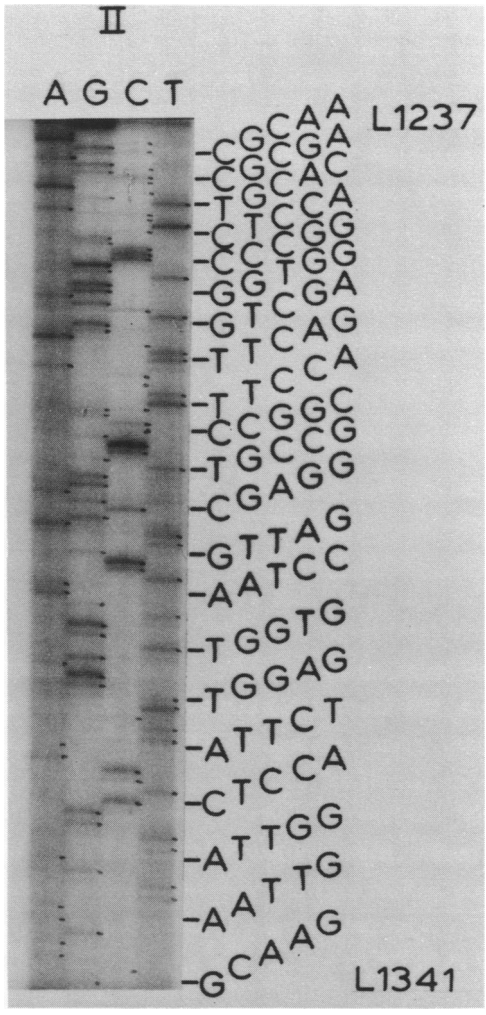


Figure 6. Autoradiogram of a sequencing gel. pYT2 was cut with SmaI and digested with exonuclease III for 130 min to remove about 1300 nucleotides from each 3' end. Method II was used to label the DNA. The DNA was then digested with AvaI to give labeled fragments about 100 to 250 nucleotides long. The sequence shown is from 1237 to 1341 in pYT2 numbering from the EcoRI site.

This is the case because while it is true that the rate of digestion of DNA by exonuclease III is not greatly affected by base composition, there is some sequence dependence. Furthermore, it should be noted that the digestion is more efficient with blunt-ended DNA than with cohesive-ended DNA.

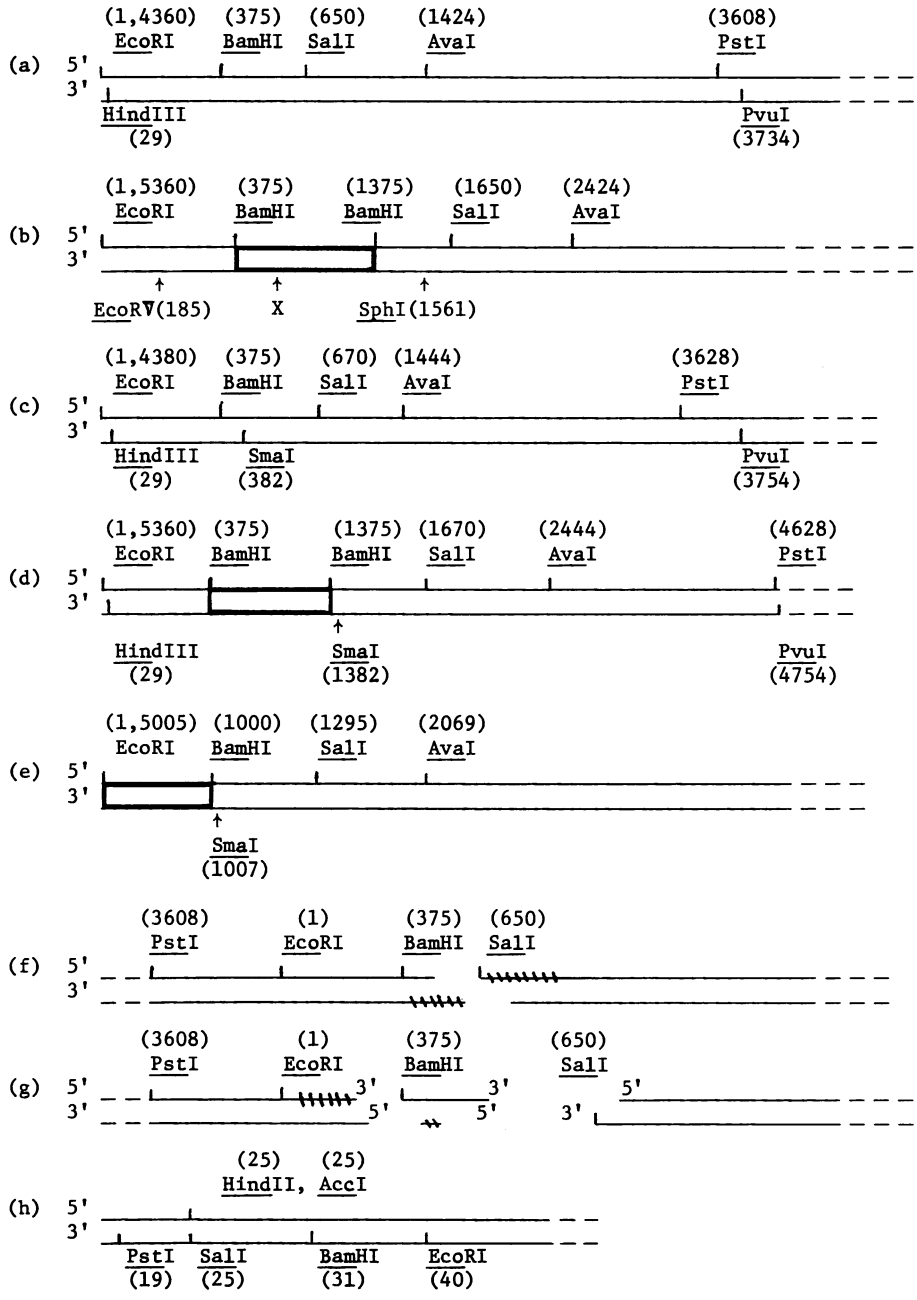
It is not necessary to strictly control the exonuclease III digestion. In fact, for reading close to the site of the second restriction enzyme, the digestion should be carried out until some molecules have reached and others have passed the second site (see Fig. 1, step b).

Both the E. coli and the T_4 DNA polymerase also possess 3'→5' exonuclease activity. Potentially, either enzyme may serve both as an exonuclease (without adding dNTPs) and for repair synthesis (after adding dNTPs). If it is possible to use either one for sequence analysis according to our scheme shown on Fig. 1, then it would require only one enzyme instead of exonuclease III plus E. coli polymerase. However, our preliminary studies showed that E. coli DNA polymerase is not active enough as a 3'→5' exonuclease to remove several hundreds of nucleotides required by our method. T_4 DNA polymerase is sufficiently active, but repair synthesis tends to stop at sequence specific locations along the template, which gives very strong bands in certain regions and practically no bands in other regions on a sequencing gel. Thus, T_4 polymerase cannot be used for DNA sequence analysis by replacing exonuclease III plus E. coli polymerase, but it can be used as an exonuclease in some case (Method II) instead of exonuclease III.

For Method I, the shortened strands are of every length over a chosen range. To label specific fragments in each of 4 separate reactions, a different [α-^{32}P] dNTP is used in each case with no unlabeled dNTPs. The polymerase will only add that nucleotide if the base residue on the template strand is complementary. It is also possible that one nucleotide at the 3' end is removed by the exonuclease activity of the polymerase and a labeled nucleotide of the same type subsequently incorporated (18). Thus the labeled strands in any one of the 4 reactions all end with the same nucleotide.

For Method II, the shortened 3' ends are labeled and then the chain terminated by using a dideoxynucleoside monophosphate, as in the technique of Sanger et al. (3).

Both methods label the linearized plasmid at both ends. The DNA is then digested with a second restriction enzyme which cuts the plasmid asymmetrically to provide the common 5' end to the labeled strands. This produces two families of labeled fragments, one family usually at least several hundred base pairs longer than the other. By proper choice of the second restriction enzyme, there will be no overlap in the sizes of the fragments from each end, and so the two sets of labeled DNA fragments can be


(a)
(1,4360) EcoRI
(375) BamHI
(650) SalI
(1424) AvaI
(3608) PstI
HindIII (29)
PvuI (3734)
(b)
(1,5360) EcoRI
(375) BamHI
(1375) BamHI
(1650) SalI
(2424) AvaI
EcoRV(185)
X
SphI(1561)
(c)
(1,4380) EcoRI
(375) BamHI
(670) SalI
(1444) AvaI
(3628) PstI
HindIII (29)
SmaI (382)
PvuI (3754)
(d)
(1,5360) EcoRI
(375) BamHI
(1375) BamHI
(1670) SalI
(2444) AvaI
(4628) PstI
HindIII (29)
SmaI (1382)
PvuI (4754)
(e)
(1,5005) EcoRI
(1000) BamHI
(1295) SalI
(2069) AvaI
SmaI (1007)
(f)
(3608) PstI
(1) EcoRI
(375) BamHI
(650) SalI
(g)
(3608) PstI
(1) EcoRI
(375) BamHI
(650) SalI
(h)
(25) HindII, (25) AccI
PstI (19)
SalI (25)
BamHI (31)
EcoRI (40)

fractionated on a single sequencing gel.

One advantage of these methods is their versatility. In most cases, suitable restriction sites are available for the first and second digestion. Depending on the location of restriction sites in and around the fragment to be sequenced, several options are available. For example, to sequence a 1000 base pair BamHI fragment (boxed area of Figure 7b) cloned at the BamHI site of pBR322 (19), if there is a site (X) such as SmaI within the fragment which does not occur elsewhere in the plasmid, this may be used for the first digestion. After exonuclease III treatment and labeling, the sequence on both sides of the cut at X may be determined after digesting the DNA with BamHI if X is located asymmetrically (e.g., around 200 to 400 base pairs from one end) within the cloned BamHI fragment. If X is located in the center of this fragment, EcoRV and SphI can be used for 2 separate second digestions to determine the sequence independently. The site chosen for the second digestion may either be unique (e.g., the EcoRV site) or there may be several sites in the plasmid, provided that the other sites are sufficiently distant that no other labeled fragments fall into the size range of the fragments to be sequenced. This principle applies also for restriction site X chosen for the first digestion if there are two or three X sites in the cloned DNA and cloning vehicle. The sequence around X can be obtained by using the EcoRI site for the first digestion and SphI for the second. The sequence of the complementary strand can be obtained by using SalI first and EcoRV second.

We have modified plasmid pBR322 by introducing a SmaI site 7 base pairs downstream from the BamHI site (20) and named it pLL10 (Fig. 7c). A DNA fragment cloned into the BamHI site of pLL10 (see Fig. 7d) can be readily analyzed by our exonuclease III methods by making the first restriction

Figure 7. Physical map of several plasmids. The circular map of each plasmid is represented as a linear molecule for simplicity of comparison. (a) pBR322, linearized at the EcoRI site, showing certain unique restriction sites. Numbers in parentheses indicate nucleotide number from the EcoRI site. (b) A 1000-base pair-long DNA fragment (boxed area), with BamHI cohesive ends, cloned into the BamHI site of pBR322. (c) pLL10, linearized at the EcoRI site. (d) A 1000-base pair-long DNA fragment, with BamHI cohesive ends, cloned into the BamHI site of pLL10. (e) A 1000- base pair-long DNA fragment, with a BamHI and a EcoRI cohesive end, cloned into pLL10. (f) pBR322, linearized at the PstI site, and cut at the SalI site. The hatched areas near the SalI site represent the portions digested by λ exonuclease. (g) pBR322 shown in (f) is digested by BamHI. The hatched areas represent the portions digested by exonuclease III. (h) A partial map of pUR222.

enzyme cut at this SmaI site. A first cut with PvuI can be used to analyze an insert at the PstI site, and a first cut with HindIII or ClaI for an insert at the EcoRI site. A DNA fragment cloned in between the EcoRI and BamHI sites by replacing the 375-long fragment of pLL10 (Fig. 7e) can be cut at either site before exonuclease III digestion, and the sequence of both strands of the cloned fragments can be determined from the same plasmid. These examples show how the exonuclease III method for sequence analysis can be initiated at a variety of restriction sites either within the fragment (Figs. 3b and 7b) or outside of the segment of interest (Figs. 4, 7a, 7c, 7d, 7e). In contrast, sequence analysis using the M13 system requires that sequence analysis be initiated from only one end near the primer (5,6), and for sequencing both strands of DNA two M13 phages that carry the inserted DNA in opposite orientations are required.

Either Method I or II can be used alone for DNA sequencing. Each has its strengths and weaknesses. Using Method I, there may be some variation in the intensity of the bands. This is especially true where there is a run of several residues of the same base because synthesis goes beyond a single base (2). The last base in a run gives the most intense signal with successive identical bases giving progressively weaker signals. For long runs of identical nucleotides, the last member that corresponds to the lowest band may even be undetectable (Fig. 3b, Method I), which is a disadvantage of this method. This problem can be minimized by using more DNA or less dNTP so that the molar ratio of DNA to [α-^{32}P] dNTP is greater than 0.17. Another disadvantage of Method I is that it requires four labeled dNTPs which are expensive to buy, although only a small amount of each is needed (e.g., 1 μC for each analysis). The strength of Method I is that no extraneous bands are observed after gel electrophoresis.

The intensity of the bands for Method II is much greater, and there is little or no variation in intensity within a run of the same base. Occasionally, however, one may find extraneous bands in the sequence, as also observed by Sanger et al. (3). To minimize this problem, the samples may be heated to 70°C for 5 min after the resuspension following exonuclease III digestion and before the dNTPs, ddNTPs and enzyme are added. In general, Method II is recommended since it requires only one labeled dNTP and the digestion with exonuclease III can be carried out with a single incubation period. Any ambiguity in sequence analysis of one strand of DNA can often be resolved by determining the complementary sequence.

In cases where one end of the DNA is 3' recessed and the other is 3'

protruding (e.g., EcoRI and PstI digested pBR322 DNA), exonuclease III digests the latter inefficiently (Yang and Wu, unpublished observation). Thus, only the 3' recessed end is extensively shortened by exonuclease III digestion. The advantage is that after incorporation of labeled nucleotides by DNA polymerase (Fig. 1, Step c) and digestion by a second enzyme there is only one family of labeled fragments. Thus, the flexibility of the exonuclease III methods is further extended in these cases. It is also possible to use λ exonuclease (21) to shorten one of the 5' ends of a DNA fragment so that the end becomes 3' protruding and no longer susceptible to exonuclease III. This can be done, for example, by digesting pBR322 DNA at the SalI site (Fig. 7a) and then by λ exonuclease (in 50 mM Tris-HCl, pH 9.5, 2 mM $MgCl_2$, 3 mM dithiothreitol) at 23°C for 10 min. to remove about 50 nucleotides from each 5' end to give a molecule similar to that depicted in Fig. 7f. The DNA molecules are next digested with BamHI, and then with exonuclease III to give DNA molecules as shown in Fig. 7g. It may be noted that only the 3' end of the upper strand at the original BamHI site junction is appreciably shortened by exonuclease III digestion. Thus, sequence analysis of the DNA fragment to the left of the BamHI site can be easily determined after digestion by a second restriction enzyme without interference by a second family of labeled fragments.

In summary, the exonuclease III methods for sequencing DNA described here offer several advantages. They are very simple and reproducible and they do not require time-consuming steps of isolating DNA fragments or primers thus permiting determination of a sequence with one day's work.

While this manuscript was in preparation, RÜther et al. (22) reported the construction of a multipurpose plasmid, pUR222, to speed up cloning of DNA and sequence analysis by an improved chemical method. The polylinker plasmid pUR222 is an excellent cloning vehicle, since DNA can be cloned into any one of the six unique restriction enzyme sites (Fig. 7h), and the selection of recombinant plasmids is simple. DNA cloned into this plasmid can be analyzed by using an improved chemical method which does not involve isolation of the labeled fragments. However because RÜther's method includes 4 out of 5 steps of our exonuclease III methods plus all the steps involved in the chemical method, it lacks the ease and simplicity of the present method. We believe that the use of a plasmid like pUR222 or pLL10 (or pBR322) as a vector for cloning DNA and our exonuclease III methods for sequencing constitutes the simplest system for rapid cloning and sequence analysis of DNA.

ACKNOWLEDGEMENTS

This work was supported by the research grants GM 27365 and GM 29179 from the National Institutes of Health, U.S. Public Health Service. We thank Robert Yang for valuable help and discussion.

REFERENCES

1. Maxam, A. and Gilbert, W. (1980) Methods in Enzymol. 65, 499-560.
2. Sanger, F. and Coulson, A.R. (1975) J. Mol. Biol. 94, 441-448.
3. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 5463-5467.
4. Gronenborn, B. and Messing, J. (1978) Nature (London) 272, 275-277.
5. Schreier, P.H. and Cortese, R. (1979) J. Mol. Biol. 129, 169-172.
6. Messing, J., Crea, R. and Seeburg, P.H. (1981) Nucleic Acids Res. 9, 309-321.
7. Richardson, C.C., Lehman, I.R. and Kornberg, A. (1964) J. Biol. Chem. 239, 251-258.
8. Wu, R., Tu, D.C.-P. and Padamanabhan, R. (1973) Biochem. Biophys. Res. Commun. 55, 1092-1099.
9. Smith, A.J.R. (1979) Nucleic Acids Res. 6, 831-848.
10. Wu, R., Ruben, G., Siegel, B., Jay, E., Spielman, P. and Tu, C.D. (1976) Biochem. 15, 734-740.
11. Zain, B.S. and Roberts, R.J. (1979) J. Mol. Biol. 131, 341-352.
12. Seif, I., Khoury, G. and Dhar, R. (1980) Nucleic Acids Res. 8, 2225-2240.
13. Wu, R. and Taylor, E. (1971) J. Mol. Biol. 57, 491-511.
14. Goodman, H.M., Olson, M.V. and Hall, B.D. (1977) Proc. Natl. Acad. Sci. USA 74, 5453-5457.
15. Rothstein, R.J., personal communication.
16. Birnboim, H.C. and Doly, J. (1979) Nucleic Acids Res. 7, 1513-1523.
17. Sutcliffe, J.G. (1978) Cold Spring Harbor Symposium 43, 77-90.
18. Donelson, J.E. and Wu, R. (1972) J. Biol. Chem. 247, 4654-4660.
19. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.W., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) Gene 2, 95-113.
20. Rothstein, R.J., Lau, L.F., Bahl, C.P., Narang, S.A. and Wu, R. (1979) Methods in Enzymol. 68, 98-109.
21. Little, J.W., Lehman, I.R. and Kaiser, A.D. (1967) J. Biol. Chem. 242, 672-
22. Rüther, U., Koenen, M., Otto, K. and Müller-Hill, B. (1981) Nucleic Acids Res. 9, 4087-4098.