UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Sequence specific termination of transcription of RNA polymerase III

Permalink
https://escholarship.org/uc/item/7p377328

Author
De Gennaro, Louis James

Publication Date
1980

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

TABLE OF CONTENTS

	PAGE
LIST OF FIGURES	ii
LIST OF TABLES	v
ABSTRACT OF THE DISSERTATION	vi
Introduction	1
References	6
Sequence Specific Termination of Transcription by RNA Polymerase III	8
References	37
Appendix 1. Ribosomal RNA Genes of *Saccharomyces cerevisiae* I. Physical map of the repeating unit and location of the regions codings for 5s, 5.8s, 18s and 25s ribosomal RNAs	39
References	45
Appendix 2. Ribosomal RNA Genes of *Saccharomyces cerevisiae* II. Physical map and nucleotide sequence of the 5s ribosomal RNA gene and adjacent intergenic regions	47
References	55
Appendix 3. Nucleotide sequence of the yeast 5s ribosomal RNA gene and adjacent putative control regions	57
References	59

LIST OF FIGURES

PAGE

Sequence Specific Termination of Transcription by RNA Polymerase III

Figure 1. Gel electrophoresis analysis of RNA synthesized from the 2.5 kbp yeast DNA fragment by yeast RNA polymerases 16

Figure 2. Agarose gel electrophoresis of the 2.5 kb fragment digested with HaeIII, HincII and HindIII and localization of sequences complementary to polymerase III transcript II. 20

Figure 3. Gel electrophoresis of polymerase III transcription products utilizing α-^{32}P-UTP, γ-^{32}P ATP, CTP or GTP 23

Figure 4. RNAse T1 oligonucleotides from α-^{32}P-labeled transcripts I and II synthesizes by RNA polymerase III with the 2.5 kb fragment as template 25

Figure 5. RNAse T1 oligonucleotides from transcripts I and II displayed by homochromotography on DEAE-cellulose plates. Homomixture b was used (14) 27

Figure 6. Nucleotide sequence of polymerase III transcripts I and II, and DNA template sequence. Major RNAse T1 oligonucleotides (Tables 2 and 3) are indicated by underlining 32

Ribosomal RNA Genes of *Saccharomyces cerevisiae*
I. Physical map of the repeating unit and location of the regions coding for 5s, 5.8s, 18s, and 25s ribosomal RNAs

Figure 1. Agarose gel electrophoresis of restriction endonuclease digested yeast and plasmid DNA 40

Figure 2. Agarose gel electrophoresis of restriction endonuclease digested *Xma* I-cleaved yeast rDNA repeat containing plasmid pBD4 and plasmid DNA vector pSF 2124. 41

PAGE

Figure 3.	Autoradiogram of an incomplete EcoRI digestion of purified Kpn I-digested yeast rDNA fragments	42
Figure 4.	Restriction map of yeast rDNA repeating unit	43
Figure 5.	Agarose gel electrophoresis of restriction endonuclease digests of purified Kpn I-generated rDNA fragments	43
Figure 6.	Hybridization of ^{125}I-labeled 5s, 5.8s, 18s, and 25s rRNA to EcoRI fragments of the yeast rDNA repeat	43
Figure 7.	Autoradiogram of the electrophoretic separation of ^{32}P-labeled oligonucleotides obtained by base specific cleavages of the portion of the 375,000 EcoRI fragments containing the 5.8s rRNA coding sequence	44
Figure 8.	Nucleotide sequence of 5' end and adjacent regions of 5.8s rRNA gene	44

Ribosomal RNA Genes of Saccharomyces cerevisiae
II. Physical map and nucleotide sequence of the 5s ribosomal RNA gene and adjacent intergenic regions

Figure 1.	Restriction endonuclease cleavage map of the yeast 2.5-kb DNA fragment containing the 5s rRNA gene	49
Figure 2.	Acrylamide gel electrophoresis of total restriction endonuclease digests of the yeast 2.5-kb DNA fragment	50
Figure 3.	Autoradiograph of partial endonuclease cleavage products of terminally labeled yeast 2.5-kb fragment	50
Figure 4.	Acrylamide gel electrophoresis of secondary restriction endonuclease digests of restriction fragments derived from the yeast 2.5-kb DNA	50
Figure 5.	Hybridization of restriction fragments derived from the 2.5-kb DNA with ^{125}I-labeled yeast 5s rRNA	50

PAGE

Figure 6. Autoradiogram of the electrophoretic separation of ^{32}P-labeled oligonucleotides obtained by base-specific cleavage of fragments containing the 5s rRNA gene 51

Figure 7. Autoradiogram of the electrophoretic separation of ^{32}P-labeled oligonucleotides obtained by base-specific cleavages of a fragment containing the region adjacent to the 5' end of the 5s rRNA gene 52

Figure 8. Nucleotide sequence of *Saccharomyces cerevisiae* 5s ribosomal RNA gene and adjacent regions 53

Figure 9. Autoradiogram of sequence gels from the 165-bp DNA fragment located between the *Eco*RI and *Hind*III sites 54

Figure 10. Summary of partial nucleotide sequence of the regions between the 5s and 35s rRNA genes of *Saccharomyces cerevisiae* 55

Nucleotide sequence of the yeast 5s ribosomal RNA gene and adjacent putative control regions

Figure 1. Strategy for isolation and sequencing of yeast DNA fragments containing the 5s rRNA gene 57

Figure 2. Autoradiogram of the electrophoretic separation of ^{32}P-labeled oligonucleotides obtained by base-specific cleavages of fragments a and b 58

Figure 3. Nucleotide sequence of *S. cerevisiae* 5s rRNA gene and adjacent regions 58

LIST OF TABLES

PAGE

Sequence Specific Termination of Transcription by RNA Polymerase III

Table I. Hybridization analysis of RNA polymerase transcripts from 2.5 kb DNA fragment containing the yeast 5s rRNA gene 18

Table II. RNAse T1 digestion products of *in vitro* transcripts I and II 29

Table III. 3' terminal RNAse T1 products of *in vitro* transcripts 30

Ribosomal RNA genes of *Saccharomyces cerevisiae*
I. Physical map of the repeating unit and location of the regions coding for 5s, 5.8s, 18s, and 25s ribosomal RNAs

Table I. Molecular weight and length of restriction fragments from *Xma* I generated yeast DNA segment containing the coding sequences for 5s, 5.8s, 18s, and 25s rRNAs 41

Table II. Molecular weights of DNA fragments generated on restriction endonuclease digestion of *Kpn* I fragments of yeast rDNA unit 43

Ribosomal RNA genes of *Saccharomyces cerevisiae*
II. Physical map and nucleotide sequence of the 5s ribosomal RNA gene and adjacent intergenic regions

Table I Molecular weight and length of restriction endonuclease digestion fragments from the *Eco*RI-generated yeast DNA segment containing the 5s rRNA gene 48

ABSTRACT

Sequence Specific Termination of Transcription

by RNA Polymerase III

by Louis James DeGennaro

When a cloned EcoRI fragment of yeast DNA containing the 5s rRNA gene is transcribed *in vitro* by yeast RNA polymerase III, two major transcripts are synthesized. The transcripts constitute 18-20% of the total RNA synthesized, are 108 and 103 bases long respectively, and do *not* contain sequences homologous to 5s rRNA. We have investigated the initiation and termination of these transcripts at the level of nucleotide sequence. Hybridization to restriction endonuclease subfragments of the DNA template indicates that the products are transcribed from one of the termini. Analysis by two-dimenstional "fingerprinting" of T1 oligonucleotides reveals that transcription is initiated at the first base pair beyond the EcoRI single stranded end and terminates at two specific sites within a short region of the template. This region is characterized by an AT rich primary structure 32 base pairs long in which the A residues occur predominantly in the coding strand. The two transcripts are produced in equal amounts when the transcription reaction contains Mn^{++} as the divalent cation. In contrast, Mg^{++} enhances the termination at the first site such that the shorter transcript constitutes 75-80% of the RNA synthesized as these two products. Thus, purified yeast RNA polymerase III is capable of recognizing and terminating transcription at specific base sequences in the DNA template.

Introduction

Gene expression in eucaryotes requires the transfer of genetic information from deoxyribonucleic acid (DNA) to ribonucleic acid molecules (RNA), the direct templates for protein synthesis (1). This transcription of genetic sequences is mediated by DNA-dependent RNA polymerase. This enzyme catalyzes the initiation, elongation and termination of polyribonucleotide chains employing ribonucleoside triphosphates as substrates (2). In the current model for the selective transcription of gene sequences, initiation and termination of RNA transcripts is dictated by specific signals on the DNA template that are recognized and translated into biochemical events by the transcription apparatus. It is at the level of the recognition of these signals that the regulation of the synthesis of a given transcript is thought to be effected.

The complexity of the overall transcription process makes it essentially impossible to study individual steps precisely, or to ascertain the effect of various factors (alteration of assay conditions, stimulatory factors, etc.) by simply measuring the total amount of RNA formed in a given amount of time (3,4). To obtain an unambiguous result, it is necessary to follow each step of the reaction and to know the size and properties of all the transcription units on the DNA template. Therefore, a detailed study of transcription requires its reconstruction *in vitro* from purified molecular components. Such *in vitro* analysis of transcription in higher eucaryotic cells is seriously hampered by two characteristics of these systems. First, as a consequence of their genetic com-

plexity, the concentration of a particular gene sequence in transcripts from isolated DNA or chromatin is very low. Second, the concentration of the transcriptive enzymes in higher cells is relatively low (5,6), making purification of these enzymes in sufficient quantities for detailed biochemical studies difficult.

In an attempt to circumvent these problems we have chosen to develop the yeast, Saccharomyces cerevisiae as a model eucaryotic system in which to study the mechanisms of transcription. Yeast, a unicellular lower eucaryote, is a appealing organism for the study of transcription. Yeast can be obtained in large quantities and at little cost. The genetics of yeast are well understood (7), and a large number of mutants defective in RNA synthesis are available (8,9). The haploid yeast genome contains 6.5-9.2 x 10^9 daltons of DNA. This represents a genetic complexity only three times that of the bacterium, Escherichia coli. Thus, the effective concentration of a gene sequence in the yeast genome is at least 100 times greater than that of a unique gene in a vertebrate genome. This low genetic complexity has facilitated the isolation of recombinant DNA molecules containing specific yeast genes which can be employed as templates in studies of transcription (see below).

Yeast contains multiple forms of DNA-dependent RNA polymerase which are similar in structure and properties to those of higher eucaryotic cells (10-12). These enzymes, originally designated I, II, and III by Roeder and Rutter (13), can be distinguished by their subunit structure, behavior on ion-exchange chromotography, cellular location, and sensitivity to the inhibitor α-amanitin (reviewed in ref. 14). This division of

enzyme forms is reflected likewise in a division of functional roles in the cell. It has been shown that RNA polymerase I synthesizes ribosomal RNA (15-17), polymerase II synthesizes heterogeneous nuclear RNA (15), and polymerase III is involved in the synthesis of 5s ribosomal RNA and transfer RNA (17,18). Each form of RNA polymerase has been purified in large quantity from yeast, and these enzymes have been characterized at the level of general enzymatic properties (12).

As indicated above, the use of a well defined template in studies of transcriptive specificity is important to the precise definition of the processes of initiation and termination. Previous attempts to reconstruct specific transcription using complex templates have met with limited success. For example, reports from our laboratory (19) and other (20) have described some degree of selective transcription of yeast ribosomal genes by purified RNA polymerase I when total yeast DNA was the template. Investigation of specificity in these systems was limited, however, by incomplete knowledge of the sites of initiation and termination of transcription in the ribosomal genes. Therefore, in order to obtain well characterized DNA templates suitable for detailed studies of transcription, we have isolated by molecular cloning the ribosomal RNA genes of *Saccharomyces cerevisiae*, and analyzed their structural organization (21,22). This dissertation describes the isolation and characterization of these genes, and the use of one of them in a detailed study of transcription by purified yeast RNA polymerase III.

Analysis of the cloned yeast ribosomal DNA (rDNA) indicates that it is composed of tandemly repeating head-to-tail units of 5.8×10^6 daltons

(9300 base pairs). Each unit contains a single gene of each of the rRNA species: 5s, 5.8s, 18s, and 25s. The physical structure of the cloned rDNA repeating unit was determined by restriction endonuclease digestion and is described in Appendix 1. Locations of the 5s, 5.8s, 18s, and 25s rRNA coding regions within the repeating unit were determined by hybridization of each purified rRNA species to the isolated DNA restriction fragments using the procedure of Southern (23). Positions of the 5s and 5.8s rRNA coding regions have been confirmed by DNA sequence analysis (Appendix 1 and 2). The primary structure of these genes has established the polarity of transcription of the 5s rRNA and high molecular weight rRNA precursor. The polarity of the high molecular weight precursor is 5'-18s, 5.8s, 25s-3'. This analysis also demonstrated that 5s rRNA and the rRNA precursor are transcribed from DNA strands with opposite polarity, as previously suggested by the hybridization results of Aarstad and Oyen (24).

The 121 base pair 5s rRNA coding sequence is contained within a 1.6×10^6 dalton (2500 base pair) Eco RI restriction enzyme subfragment of the yeast rDNA repeating unit (Appendix 1). The primary sequence of yeast 5s rRNA has previously been determined by Miyazaki (25). A significant fraction of the molecules synthesized *in vivo* were shown to possess a 5'-terminal triphosphate. This fact, and analyses of 5s rRNA biosynthesis *in vivo* (26,27), suggest that 5s rRNA is a primary transcription product. Thus, the control regions (sites of initiation and termination of transcription) for its biosynthesis should be adjacent to the DNA sequences from which it is transcribed. Therefore, the 1.6×10^6 dalton

yeast DNA fragment should contain the structural and flanking gene sequences recognized and transcribed by RNA polymerase III *in vivo*. Based on this premise, we have utilized the 1.6×10^6 dalton DNA fragment and purified yeast RNA polymerase III, to reconstruct an *in vitro* transcription system. Our goal was to establish the basal level of specific transcription by RNA polymerase III on this template. Whereas purified RNA polymerase III exhibited no obvious transcriptive specificity of the 5s rRNA gene sequences in this system, we have defined a region of DNA which is recognized by RNA polymerase III for chain termination, and have identified the exact site at which transcription terminates.

REFERENCES

1. Watson, J.D. (1976) Molecular Biology of the Gene, W.A. Benjamin Inc., Menlo Park, California.
2. Losick, R. and Chamberlain, M. (1976) RNA Polymerase, Cold Spring Harbor Laboratory, New York.
3. Chamberlain, M.J. (1974) Annu. Rev. Biochem. 43, 721-775.
4. Chamberlain, M.J. (1976) RNA Polymerase, pp. 17-67, Cold Spring Harbor Laboratory, New York.
5. Rutter, W.J., Goldberg, M.I., and Perriard, J.C. (1974) in Biochemistry of Cell Differentiation (Paul, J., ed.) London, Butterworth: p. 267.
6. Schwartz, L.B., and Roeder, R.G. (1974) J. Biol. Chem. 249, 5898-5908.
7. Hartwell, L.H. (1970) Ann. Rev. Genet. 4, 373-396.
8. Hartwell, L.H., McLaughlin, C.S., and Warner, J.R. (1970) Mol. Gen. Genet. 109, 42-56.
9. Thonart, P. Bechet, J., Hilger, F., and Burny, A. (1976) J. Bact. 125, 25-32.
10. Roeder, R.G. (1969) Doctoral Thesis, University of Washington, Seattle.
11. Hager, G.L., Holland, M.J., and Rutter, W.J. (1976) Biochemistry 15, 1-7.
12. Valenzuela, P., Bell, G.I., Weinberg, F., and Rutter, W.J. (1978) Methods in Cell Biology 18, 1-26.
13. Roeder, R.G. and Rutter, W.J. (1969) Nature 224, 234-237.
14. Roeder, R.G. (1970) in RNA Polymerase (Losick, R. and Chamberlain, M., eds.) pp. 285-329, Cold Spring Harbor Laboratory, New York.

15. Blatti, S.P., Ingles, C.J., Lindell, T.J., Morris, P.W., Weaver, R.F., Weinberg, F., and Rutter, W.J. (1970) Cold Spring Harbor Symp. Quant. Biol. 35, 649.

16. Reeder, R.H., and Roeder, R.G. (1972) J. Mol. Biol. 67, 433-442.

17. TeKamp, P.A., Valenzuela, P., Maynard, T., Bell, G.I., and Rutter, W.J. (1979) J. Biol. Chem. 254, 955-963.

18. Weinmann, R., and Roeder, R.G. (1974) Proc. Natl. Acad. Sci. U.S.A. 71, 1790-1795.

19. Holland, M.J., Hager, G.L., and Rutter, W.J. (1977) Biochemistry 16, 16-24.

20. Van Keulen, H., and Retel, J. (1977) Eur. J. Biochem. 79, 579-588.

21. Bell, G.I., DeGennaro, L.J., Gelfand, D.H., Bishop, R.J., Valenzuela, P., and Rutter, W.J. (1977) J. Biol. Chem. 252, 8118-8125.

22. Valenzuela, P., Bell, G.I., Venegas, A., Sewall, E.T., Masiarz, F.R., DeGennaro, L.J., Weinberg, F., and Rutter, W.J. (1977) J. Biol. Chem. 252, 8126-8135.

23. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.

24. Aarstad, K., and Oyen, T.B. (1975) FEBS Letters 51, 227-231.

25. Miyazaki, M. (1974) J. Biochem. 75, 1407-1410.

26. Udem, S.A., and Warner, J.R. (1972) J. Mol. Biol. 65, 227-242.

27. Trapman, J., and Planta, R.J. (1975) Biochim. Biophys. Acta 414, 115-125.

SEQUENCE SPECIFIC TERMINATION OF TRANSCRIPTION BY RNA POLYMERASE III

The use of a well defined template in studies of transcriptive specificity is important to the precise definition of the processes of initiation and termination. We have previously attempted to reconstitute specific transcription using purified components. This approach has met with limited success in experiments utilizing purified RNA polymerase with total genomic DNA as template. For example, reports from our laboratory (1) and others (2) have described some degree of selective transcription of the yeast rRNA genes by purified RNA polymerase I when total yeast DNA was the template. Investigation of specificity in these systems was limited, however, by incomplete knowledge of the sites of initiation and termination of transcription in the ribosomal genes.

Conflicting reports exist in the literature concerning the level of transcriptive specificity exhibited by purified RNA polymerase III from higher eucaryotes. Parker et al. (3,4) have reported no selectivity toward 5s rRNA synthesis by Xenopus laevis polymerase III on total or cloned Xenopus DNA. In contrast, Ackerman and Furth (5) have recently reported a 10 to 100 fold selectivity for 5s synthesis by calf thymus RNA polymerase III over E. coli polymerase on calf thymus DNA.

We report here a detailed study of the products of transcription by purified yeast RNA polymerase III utilizing a cloned fragment of yeast DNA as template. This DNA fragment contains the structural and flanking gene sequences for the 5s species of yeast ribosomal RNA which is transcribed by RNA polymerase III in vivo. The fragment is 2.5 kilobase pairs in length and has been extensively characterized at the level of DNA base sequence (6). The fragment also contains the intergenic regions before and after the 35s ribosomal RNA precursor gene (7). Our goal has been to establish the basal level of specific transcription by RNA polymerase III on this template. Whereas purified RNA polymerase III exhibits

no obvious transcriptive specificity for the 5s rRNA gene sequences in this system, we have defined a region of DNA which is recognized by RNA polymerase III for chain termination, and have identified the exact site at which transcription terminates.

EXPERIMENTAL PROCEDURES

Purification of 2.5 kb Yeast DNA Fragment

pSF2124, a colicin El-ampicillin resistant plasmid (8) containing the 2.5 kb EcoRI fragment of yeast ribosomal DNA was maintained in a derivative of E. coli K-12, DG-75 (an F^- derivative of DG-73 (9)). The plasmid DNA was isolated as described by Meagher *et al.* (10). After EcoRI digestion the 2.5 kb yeast insert was separated from the bacterial plasmid vector by two successive sucrose gradient centrifugations (20 hr at 20°C in a SW27 rotor) through 5-20% w/v linear sucrose gradients prepared in 1M NaCl, 25mM Tris-HCl, pH 8.0, 1mM EDTA. Removal of sucrose and concentration of fragment was accomplished by ethanol precipitation (2 volumes at -20°C, overnight).

Purification of Yeast RNA Polymerase

RNA polymerase III was purified as described by Valenzuela *et al.* (11) After RNA-cellulose chromatography, fractions containing enzyme activity were pooled and dialyzed against Buffer A; 50% glycerol, 0.4M KCl, 1mM PMSF. The enzyme was stored under liquid nitrogen at a concentration of 2 mg/ml in this buffer. Final specific activity was 100-140 units/ml depending on the preparation used.

Transcription Reactions and RNA Purification

All glassware was siliconized and baked at 350°C overnight before use. A typical transcription reaction contained 50mM Tris-HCl, pH 8.0, 0.5mM ATP, CTP and GTP, 0.1mM UTP, 10 μCi of α-^{32}P-UTP, 10mM 2-mercaptoethanol, 50mM KCl, 2 μg of the 2.5 kb yeast fragment, 2 μg yeast RNA polymerase III, and 1.3mM $MnCl_2$ or 8.3mM $MgCl_2$ as indicated. Reactions,

assembled on ice, were started by the addition of RNA polymerase and were incubated at 30°C for 20 min. Preparative reactions were scaled up 10 fold. In experiments with gamma-^{32}P-labeled nucleoside triphosphates 10 μCi of the indicated isotope were utilized. Concentrations of the remaining triphosphates were as indicated for the typical reaction.

Reactions were stopped by the addition of an equal volume of stop buffer which contained 0.2M Na OAC, 10mM EDTA, 0.5% SDS, pH 5.0. Incorporation of α-^{32}P-UTP was assayed by applying an aliquot of the reaction directly onto a Whatman DE-81 filter disc which had been pre-wet with stop buffer. The filters were washed seven times with 5% Na_2HPO_4, 10mM Na-pyrophosphate, twice with water, twice with ethanol and dried. Radioactivity was measured in a liquid scintillation counter after immersing the discs in a solution of 4 g/l Omnifluor (New England Nuclear) in toluene. Reactions were then extracted with an equal volume of phenol saturated with 0.2M NaOAc, 1mM EDTA, pH 5.0. After centrifugation at 4000xg for 10 minutes RNA was precipitated from the aqueous phase by the addition of two volumes of ethanol and stored overnight at -20°C. After centrifugation at 27,000xg for 30 minutes the precipitates were resuspended in sample buffer, heated at 65°C for 60 seconds and applied to a polyacrylamide gel as described by Rubin (25). Analytical gels were 1.5 mm thick, preparative gels were 3.0 mm thick. Electrophoresis was at constant 120 volts for 5 hr. Gels were then either dried under vacuum or wrapped in Saran wrap and placed on No-Screen Medical X-ray film. RNA products were located on preparative gels by using the autoradiogram as a template. The bands were excised, crushed and extracted twice for two hr. in 10mM Tris-HCl, pH 7.5, 20mM $MgCl_2$, 0.5M NaCl. Salt removal and concentration of extracts were achieved by ethanol precipitation.

Hybridization Analysis

The 2.5kb fragment purified as described above was immobilized on 47-mm diameter nitrocellulose filters (12) from which 6.5-mm diameter filters were punched. Each filter contained 6.6-7.7 μg of 5s DNA. DNA-containing filters as well as blank filters were prehybridized for 16 hrs. at 50°C in 200 μl of 5x SSC (SSC is 0.15M NaCl, 0.015M solium citrate), 50% formamide, and purified yeast 5s RNA. After prehybridization, filters were washed twice for 10 minutes at 50°C in 5x SSC, 50% formamide, then washed twice at room temperature in 2x SSC. Hybridization was performed under the same conditions described above with 4000-20,000 cpm of *in vitro* synthesized ^{32}P-labeled RNA and ^{3}H-labeled 5s rRNA standard included. Reactions were performed in duplicate. Filters were washed as described above and then incubated for 1 hour at room temperature in 2x SSC and 20 μg/ml RNase A. Filters were washed again twice in 2x SSC, dried, and counted in toluene/omnifluor.

Restriction Endonuclease Reactions and Southern Hybridization

Restriction endonucleases were obtained and reactions were carried out as previously described (7). Restricted 2.5 kb yeast DNA was fractionated on 2% agarose gels. The separated fragments were transferred to a nitrocellulose filter according to Southern (13), and hybridized with ^{32}P-in vitro RNA (7). Approximately $2x10^{5}$ cpm of RNA was used per filter.

RNA Sequencing

Ribonuclease T1 and pancreatic ribonuclease A digestions of *in vitro* RNA, and two dimensional fractionation of oligonucleotides were performed according to Brownlee (14). An enzyme to substrate ratio of 1:20 in a

buffer containing 0.01M Tris-HCl, pH 7.4, 1mM EDTA was used for both ribonuclease T1 and ribonuclease A digestions. Thin layer homochromatography was on DEAE-cellulose plates (1:7.5 DEAE - cellulose:cellulose, Analtech, Inc., Neward, Delaware) using homomixture b.

Results

Gel Electrophoretic Analysis of Transcription Products - The EcoRI fragment was transcribed by RNA polymerases I, II and III as described in Experimental Procedures. As illustrated in figure 1, RNA synthesized from this template in the presence of manganese is polydisperse, ranging in size from 8s to less than 4s when displayed by polyacrylamide gel electrophoresis. RNA synthesized with magnesium as the divalent cation is also heterogeneous but there is a notable increase in size.

Superimposed on the polydisperse products in the reactions containing polymerase III are two discrete transcripts (I,II) approximately the size of 5s yeast rRNA. Quantitative analysis of transcription products of RNA polymerase III was performed by determining the ^{32}P incorporated into different regions of polyacrylamide gels similar to those shown in figure 1. It was found that the two discrete products (I, II) represented 18-20 percent of the total RNA synthesized in the presence of either divalent cation. Transcripts I and II occured in equivalent amounts when the transcription reaction contained manganese. However, when magnesium was used as the divalent cation, the shorter transcript (II) constituted 80 percent of the RNA found in these two discrete species. Transcripts which were larger than the discrete species I and II represented 45 percent of the total, while transcripts smaller than I and II contained approximately 30-35% of the total RNA synthesized.

Competition Hybridization Analysis of RNA Products - To test for the presence of 5s rRNA transcripts the reaction products were hybridized to the 2.5 kb fragment in the presence and absence of competing yeast 5s RNA. Hybridization analysis was performed with the discrete polymerase III transcripts (I and II), and on total RNA synthesized in the presence of manganese or magnesium (table 1). In each case, except for the transcrip-

Fig. 1. Gel electrophoresis analysis of RNA synthesized from the 2.5 kbp yeast DNA fragment by yeast RNA polymerases. RNA was synthesized under standard reaction conditions in the presence of magnesium or manganese. Electrophoresis was as described in Procedures. a,b RNA polymerase I Mn^{++}, Mg^{++}. c,d RNA polymerase II Mn^{++}, Mg^{++}. e,f RNA polymerase III Mn^{++}, Mg^{++}. In reactions containing polymerase III a total of 0.4 nmoles of UTP was incorporated in the presence of manganese; 2.5 nmoles was incorporated with magnesium. Positions of purified yeast 5.8s, 5s and 4s markers are indicated.

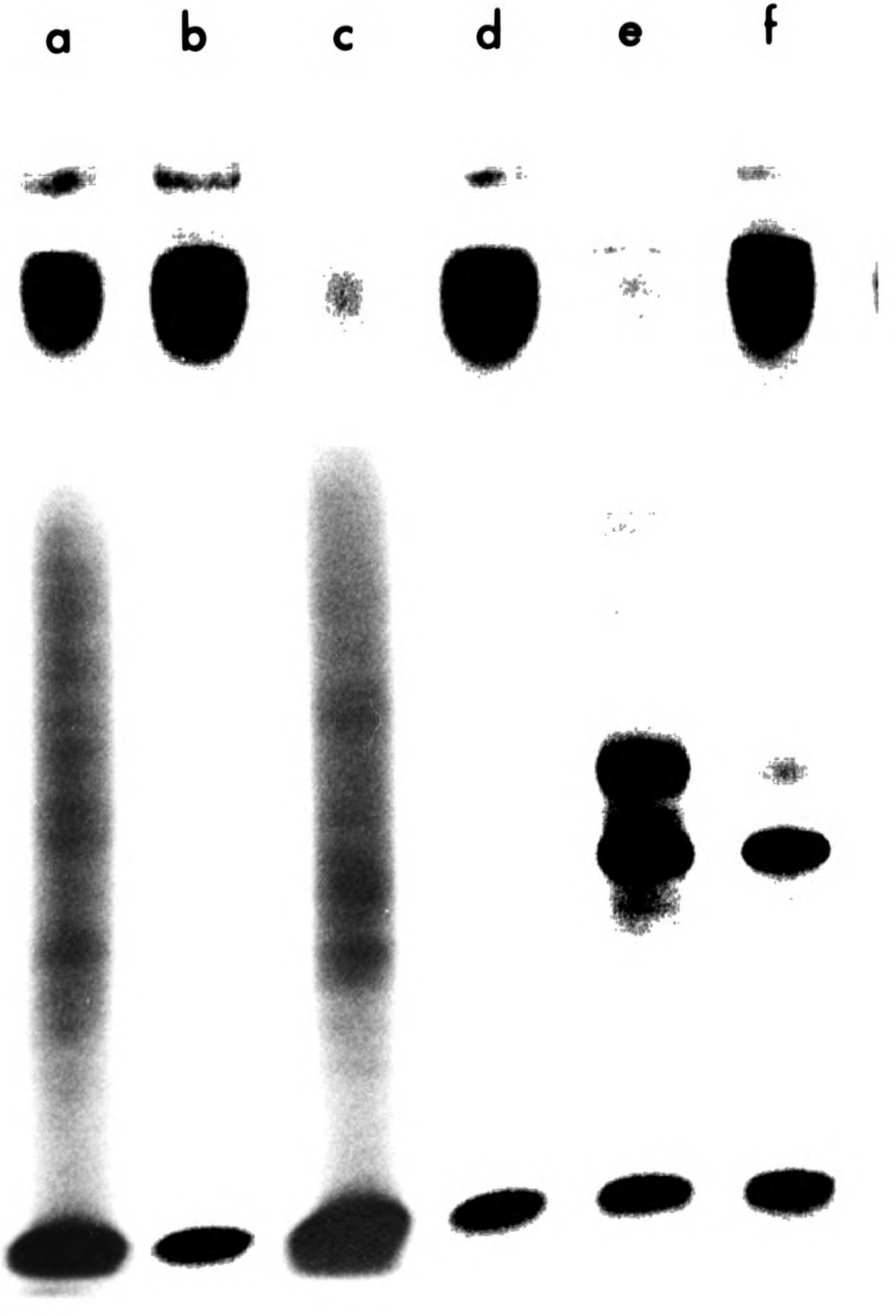

a
b
c
d
e
f

TABLE 1

Hybridization Analysis of RNA Polymerase Transcripts from 2.5 Kb DNA Fragment Containing the Yeast 5s rRNA Gene

RNA Polymerase	Divalent Cation	^{32}P Input cpm	cpm Hybridized		5s RNA Hybridization Efficiency * %	^{32}P cpm in 5 s RNA	
			-5s RNA	+5s RNA		total	% input
I, total	Mn	6,400	816	529	88.6	323	5.0
I, total	Mg	9,581	349	247	87.5	113	1.1
III, total	Mn	14,850	554	394	87.4	183	1.2
III, total	Mg	6,710	251	196	88.3	62	0.9
III, Transcript I + II	Mn	4,800	295	237	77.5	74	1.5

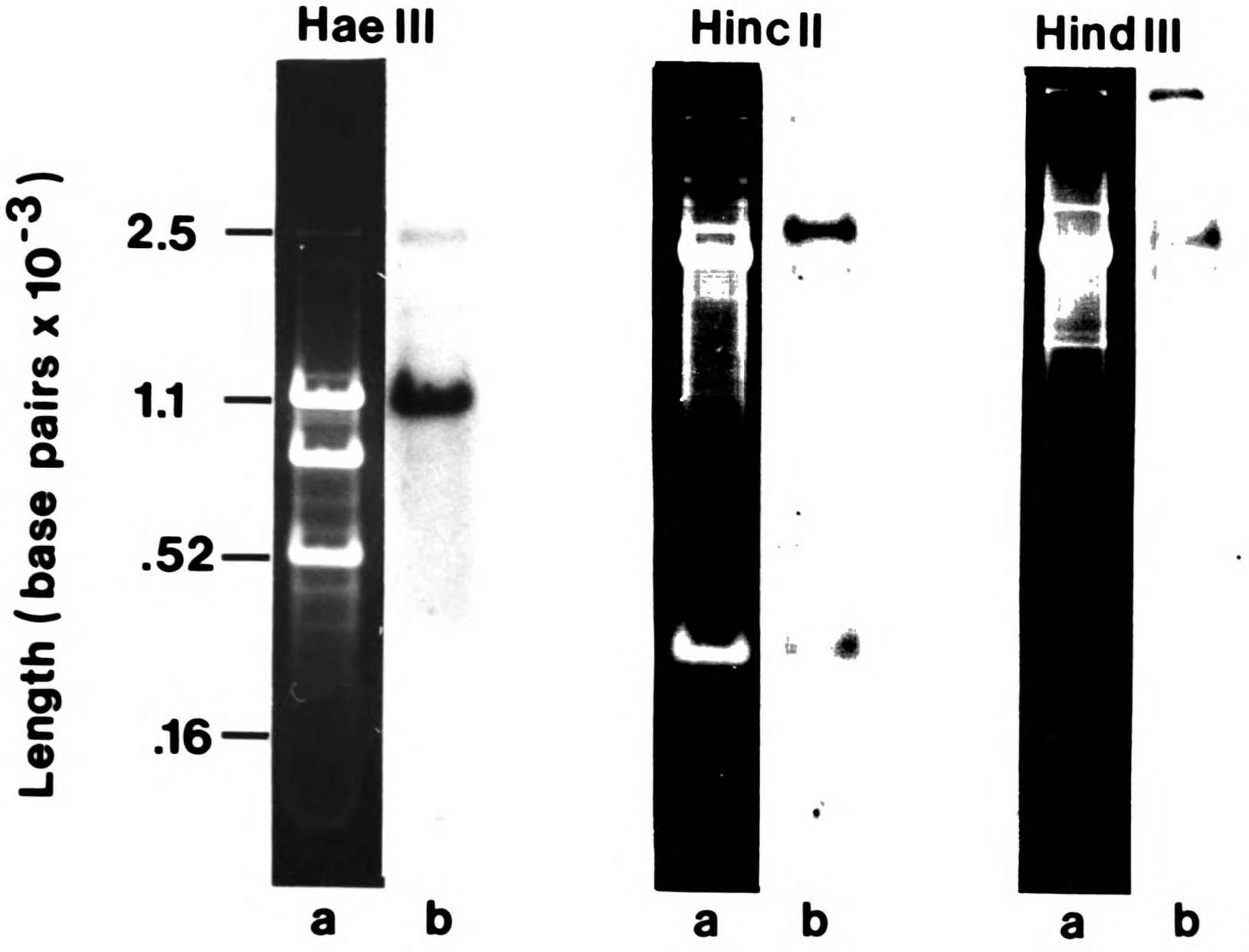
* Calculated from the efficiency of hybridization of an internal ^{3}H-labeled 5s RNA standard.

tion products of RNA polymerase I with magnesium as the cation, 5s RNA sequences comprised a smaller fraction of the input RNA than that (2.4%) expected from random transcription.

Location of Template Coding Sequence for Transcripts I and II - The location of the DNA sequences which code for transcripts I and II was determined by hybridization of these RNA products, after isolation from preparative polyacrylamide gels, to restriction endonuclease sub-fragments of the 2.5 kpb template. Template DNA was treated with the endonucleases HaeIII, HincII and HindIII. Resulting restriction fragments were separated by agarose gel electrophoresis, transferred to a nitrocellulose filter as described by Southern (13), and hybridized to transcript I or II RNA. Figure 2 depicts the hybridization results using ^{32}P-labeled transcript II. Lanes marked (a) show the restriction endonuclease subfragments of the 2.5 kbp template DNA visualized under ultraviolet light after staining with ethidium bromide. Lanes marked (b) are the respective nitrocellulose hybridization filters. Transcript II hybridizes to undigested template DNA in each case. In addition hybridization can be detected to the 1.1 kbp HaeIII subfragment, the 0.298 kbp HincII subfragment, and the 0.166 kbp HindIII subfragment.

This hybridization data locates the coding sequence for transcript II at the left end of the 2.5 kbp template. The coding sequence must lie primarily within the terminal 166 base pairs at the left end since no hybridization is detectable to the large (approximately 2.4 kbp) HindIII subfragment. Similar hybridization experiments utilizing transcript I as the probe give the same result. Thus, the coding sequences for both transcripts are located within the 166 bp HindIII subfragment.

Fig. 2. Agarose gel electrophoresis of the 2.5 kb fragment digested with HaeIII, HincII and HindIII and localization of sequences complementary to polymerase III transcript II. Fragments were separated by gel electrophoresis in 2% agarose. Columns marked (a) show the ethidium bromide staining pattern. Columns marked (b) show autoradiograms of hybridization with ^{32}P-labeled transcript II after transfer from gel to a nitrocellulose filter. (13)


Hae III
Hinc II
Hind III
Length (base pairs x 10^-3)
2.5
1.1
.52
.16
a
b
a
b
a
b

Determination of Transcript Base Sequence

Incorporation of γ-^{32}P Nucleoside Triphosphates - Nucleoside triphosphates labeled in the gamma phosphate position were utilized to determine the initiating triphosphate at the 5'-end of transcripts I and II. Figure 3 shows the polyacrylamide gel analysis of transcription products from reactions which contained γ-^{32}P-labeled ATP, CTP, or GTP, compared to the reaction products from synthesis utilizing α-^{32}P-UTP. Both transcripts I and II are specifically labeled only by γ-^{32}P cytidine triphosphate. γ-^{32}P-UTP (data not shown) was not incorporated into either transcript I or II. Transcription of both products is therefore initiated by RNA polymerase III with cytidine triphosphate.

Direct RNA Sequence Determination - Further RNA sequence information was obtained from transcripts I and II by two dimensional oligonucleotide mapping (fingerprinting) after ribonuclease T1 digestion. A comparison of the RNAse T1 fingerprints of transcripts I and II is shown in Figure 4. The simplicity and similarity of these patterns suggest that the transcripts represent discrete molecular species which are identical within the analytical range of this resolution system. Spots 1-8 were cut from the fingerprints and the relative molar yield of ^{32}P for each oligonucleotide was determined by Cerenkov counting. The oligonucleotides were then eluted from the DEAE-paper and subjected to further sequence analysis by electrophoretic fractionation after alkaline hydrolysis or digestion with RNAse A. Table 2 contains a summary of the data obtained.

In order to increase the resolution of material which appears at the second dimension origin of the fingerprints, T1 digests of transcripts I and II were fractionated by homochromatography on DEAE-cellulose plates. Figure 5 shows a comparison of the T1 oligonucleotides separated by homomixture b. Again numbered spots were eluted and analyzed after alkaline

Fig. 3. Gel electrophoresis of polymerase III transcription products utilizing α-^{32}P-UTP, γ-^{32}P ATP, CTP or GTP. Rna was synthesized under standard conditions in the presence of manganese. Gel electrophoresis was as described in Procedures. The right-most lane displays α-^{32}P-labeled yeast marker RNAs.

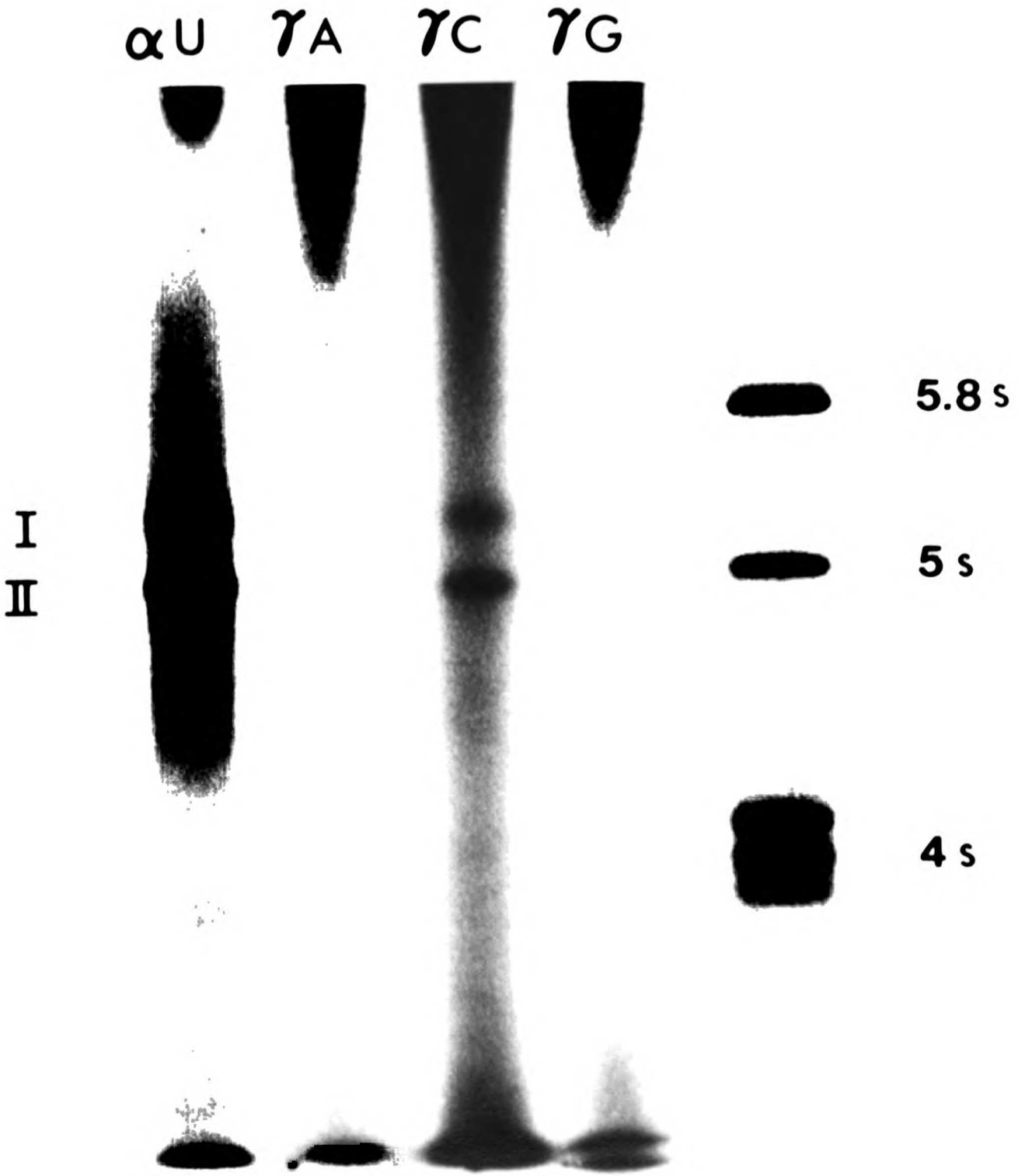

αU
γA
γC
γG
I
II
5.8 s
5 s
4 s

Fig. 4. RNAse T1 oligonucleotides from α-^{32}P-labeled transcripts I and II synthesizes by RNA polymerase III with the 2.5 kb fragment as template. Bands corresponding to transcripts I and II were excised from preparative gels as described in Procedures. The extracted RNA was digested by RNAse T1 and the resulting oligonucleotides were separated by the standard fingerprint method (14). Identity of the numbered oligonucleotides was confirmed by digestion with RNAse A or alkaline hydrolysis and subsequent electrophoresis on paper (Table 2, (14)).

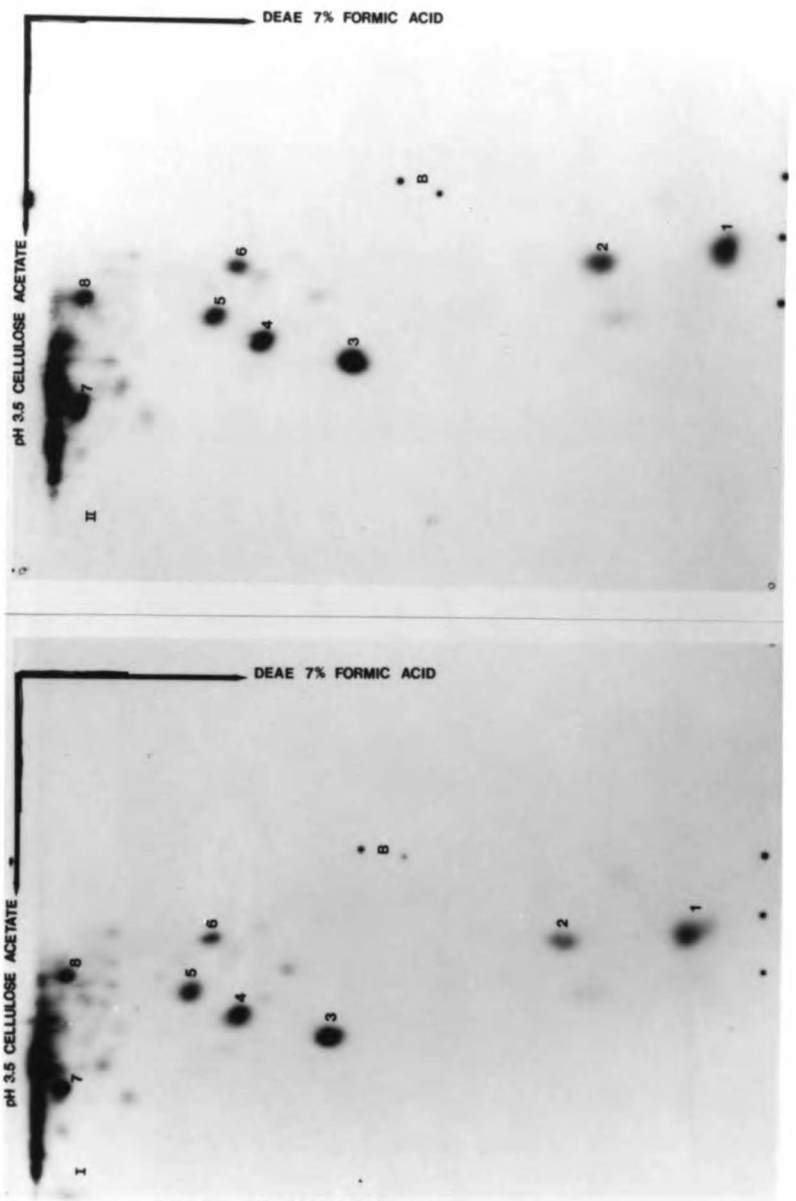

pH 3.5 CELLULOSE ACETATE
DEAE 7% FORMIC ACID
I
7
8
5
6
4
3
B
2
1
pH 3.5 CELLULOSE ACETATE
DEAE 7% FORMIC ACID
II
7
8
5
6
4
3
B
2
1

Fig. 5. RNAse T1 oligonucleotides from transcripts I and II displayed by homochromotography on DEAE-cellulose plates. Homomixture b was used (14). Numbered oligonucleotides were further analyzed by RNAse A digestion or alkaline hydrolysis after recovery from the plates (Tables 2 and 3, (14)).

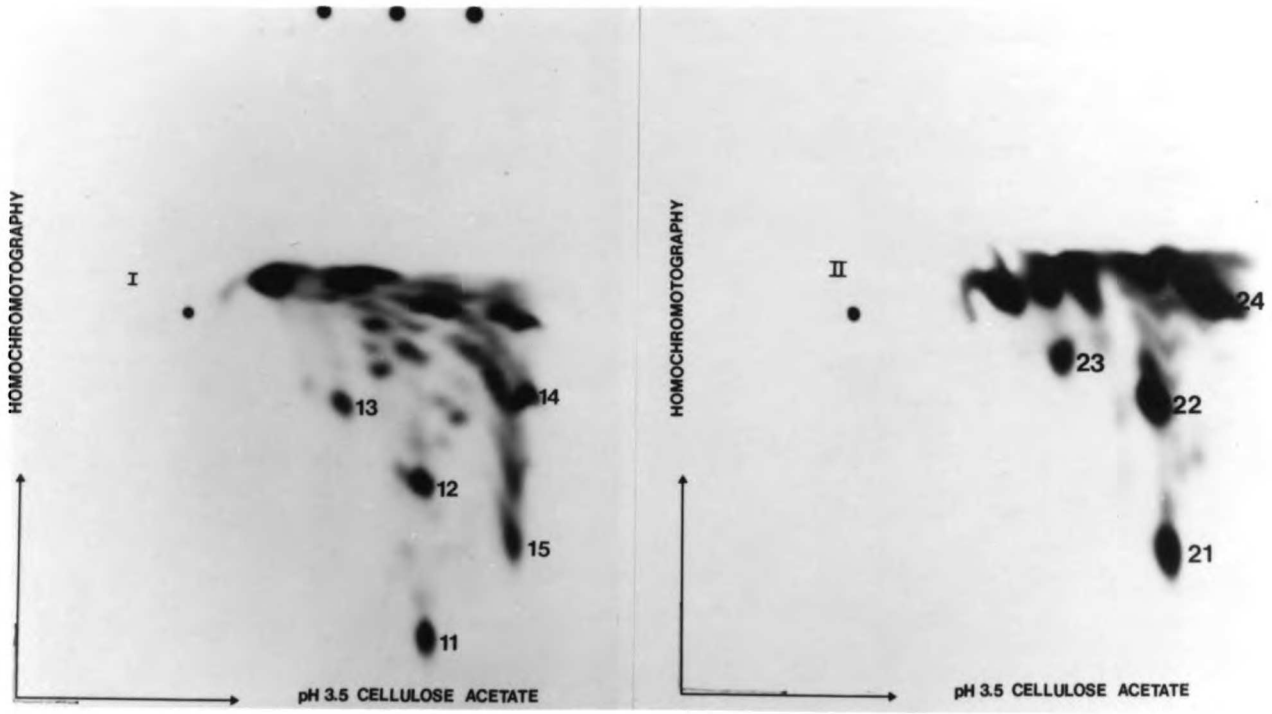

HOMOCHROMOTOGRAPHY
I
13
14
12
15
11
pH 3.5 CELLULOSE ACETATE
HOMOCHROMOTOGRAPHY
II
24
23
22
21
pH 3.5 CELLULOSE ACETATE

TABLE 2

RNAse T1 Digestion Products of In Vitro Transcripts I and II

Spot Number	T1 Oligo-nucleotide	Average Number Phosphates		Mole Ratio Pancreatic RNAse Redigestion Products			Mole Ratio Alkaline Hydrolysis Products		
		Predicted	Obtained	Predicted	Obtained		Predicted	Obtained	
1	Gp	2	2.03						
2	AGp	1	1.12						
3	UGp	3	3.04						
4	AUGp	2	2.31						
5	AUAGp	2	1.62	1	AU	0.84	1	G	1.0
				1	AG	1.0	1	A	1.25
6	AUCCGp	1	1.01						
7	UAUUGp	3	2.67	1	G	1.0	1	U	1.3
				2	AU	2.3	1	G	1.0
							1	A	1.7
8	pppCUAUGp	2	1.8	1	pppC	1.15	1	C	1.0
				1	AU	1.0	1	A	1.29
11,21	UAUAUAUCUAUUA UAAUAUACGp	9	8.8	1	C	0.89	1	U	1.06
				7	AU	6.98	7	A	6.7
				1	A2U	1.18	1	C	1.0
12,22	UACCAUUUACUAU UUGp	7	6.25	2	U	1.84	4	U	3.83
				1	AC	1.0	2	A	2.09
				2	AU	2.03	1	C	0.98
13,23	UAAAAACAUGp	2	1.8	1	G	1.0	1	G	1.0
				1	AU	1.07	1	A	1.09

TABLE 3

3' Terminal RNAse T1 Products of In Vitro Transcripts

Spot Number	Pancreatic RNAse Redigestion Products			Alkaline Hydrolysis Products			Proposed Sequence
	Obtained		Proposed	Obtained		Proposed	
Transcript I							
14	U	8.1	8	U	11.3	11	UCUUUUUAUU
	C	1.0	1	A	1.2	2	UUUUAUUU-OH
	AU	2.0	2	C	1.0	1	
Transcript II							
24	U	6.10	6	U	7.8	8	UCUUUUUAUUU
	C	0.92	1	A	1.06	1	UU-OH
	AU	2.0	2	C	1.0	1	

hydrolysis and RNAse A digestion. The data obtained is summarized in Tables 2 and 3.

Comparison of the RNA sequence data to the DNA base sequence of the left end of the 2.5 kbp fragment (shown in figure 6, also Ref. 7) establishes the complete sequence of transcripts I and II. Both transcripts initiate at the first base pair adjacent to the single stranded EccRI terminus of the left end of the template. Chain initiation occurs with cytidine triphosphate and elongation then continues until a signal for termination is reached. Table 3 shows the fingerprint analysis and proposed sequence of the 3'-terminal oligonucleotides of transcripts I and II. Both products terminate within a short AT-rich region of the template. The DNA base sequence of this region consists of a run of 31 AT base pairs broken by a single GC pair. The adenosine residues occur predominantly in the coding strand. Transcript II terminates in a run of five uridylates after transcription has proceeded 11 residues into the AT-rich region. Transcript I terminates in a run of three uridylates at the sixteenth residue into the AT-rich region.

Fig. 6. Nucleotide sequence of polymerase III transcripts I and II, and DNA template sequence. Major RNAse T1 oligonucleotides (Tables 2 and 3) are indicated by underlining. Terminal sequences for the yeast 5s rRNA and $tRNA^{Phe}$ genes are also presented. Sequences are aligned with respect to the first T of the T cluster in which termination occurs.

FIGURE 6

```
                1                   20                  40                  60
5'...AATTCTATGATCCGGGTAAAAACATGTATTGTATATATCTATTATAATATACGATGAGGATGATAGTGTG
        GATACTAGGCCCATTTTTGTACATAACATATATAGATAATATTATATGCTACTCCTACTATCACAC

pppCUAUGAUCCGGGUAAAAACAUGUAUUGUAUAUAUCUAUUAUAAUAUACGAUGAGGAUGAUAGUGUG
```

Major RNAse T1 Oligonucleotides

```
(8)  (6)   (13,23)          (11,21)

        80                  100                 120
TAGAGTGTACCATTTACTATTTGGTCTTTTTATTTTTTATTTTCTTTTTTTTTTTTTTCGTTGCAAAGAT...3'
ATCTCACATGGTAAATGATAAACCAGAAAAATAAAAAATAAAAGAAAAAAAAAAAAAAAGCAACGTTTCTA...5'

UAGAGUGUACCAUUUACUAUUUGGUCUUUUUAUUUUU-OH                    Transcript II

(12,22)

                    UCUUUUUAUUUUUUAUUU-OH                   Transcript I
```

3'-terminal sequence yeast 5s rRNA gene

```
5'...TCAGGTGCTGCAATCTTTATTTCTTTTTTTTTTTTCTAGTT...3'
3'...AGTCCACGACGTTACAAATAAAGAAAAAAAAAAAAAGATCAA...5'

5'...UGCAAUCU-OH     Yeast 5s rRNA
```

3'-terminal sequence yeast $tRNA^{Phe}$ gene

```
5'...CAGAATTCGCATTAATTTTTTTTACTTTTCATTCGTTTTGG...3'
3'...GTCTTAAGCGTAATTAAAAAAAATGAAAAGTAAGCAAAACC...5'

5'...AAUUCGCA-OH     Yeast tRNA^Phe
```

Discussion

In eucaryotes, the functional definition of DNA base sequences and their interaction with eucaryotic RNA polymerases remain unclear. In order to obtain information concerning these interactions, we have reconstituted an *in vitro* transcription system from purified, homologous components. Thus, a cloned fragment of yeast DNA which contains the 5s ribosomal RNA structural gene and the intergenic regions before and after the 35s ribosomal RNA precursor gene was used as a template for transcription by purified yeast RNA polymerase III. We report here the detailed analysis of the products of transcription in this system.

Transcription of 5s rRNA - Competition hybridization data presented in Table 1 indicate that transcription of the 2.5 kb yeast DNA fragment by purified yeast RNA polymerase I or III does not result in the selective synthesis of 5s rRNA. Indeed, 5s gene transcription by polymerase III is below that expected from random transcription. This result may reflect the initiation properties of that enzyme. Analysis of transcripts I and II conclusively shows that polymerase III is capable of initiation at a free end of the linear DNA template. The 5s rRNA coding sequence is approximately 1400 base pairs from the free end of the coding strand. Thus, preferred end initiation followed by a high probability of termination before the 5s sequences are encountered may lead to the apparent discrimination against 5s RNA synthesis. Initiation at free ends and single strand breaks (nicks) in DNA templates has often been inferred (15-17), and is now confirmed as an operational characteristic of purified yeast RNA polymerase III in the absence of additional factors.

Termination of transcription by RNA polymerase III - Purified yeast polymerase III synthesizes two predominant products as the result of specific termination at a site on the 2.5 kb fragment (fig. 1). These

transcripts represent 18-20% of the RNA synthesized, and are 108 and 103 bases long, respectively. RNAse T1 two-dimensional fingerprint analysis indicates that these products are simple and discrete molecular species (fig. 4). Southern hybridization (fig. 2) and RNA base sequence analysis (tables 2, 3; fig. 6) have been used to define the location of the coding sequence within the template, and the actual sequence of these transcripts. Thus, RNA polymerase III initiates transcription at the left end of the linear DNA template. Incorporation begins at the first base pair available and cytidine triphosphate is the initiating nucleotide. Incorporation continues until the polymerase encounters a signal for chain termination. Transcription terminates within a run of 31 AT base pairs in which the adenine residues occur primarily in the coding strand. Three deviations from this sequence pattern, at base pairs 98, 105 and 110, occur in the AT run. Termination of transcription to produce the discrete products occurs two base pairs 5' to the pattern deviations at position 105 (TA pair), and position 110, (CG pair), indicating that these pattern interruptions may help signal termination. The efficiency of the more proximal termination site is markedly dependent upon the divalent cation present (fig. 1). Whereas the ratio of product II to product I is approximately one in the presence of manganese, magnesium increases that ratio to 4:1 or greater. The locus of cation interaction which is pertinent to termination remains unclear. However, evidence from procaryotic systems suggests that the effect may be on the stability of potential stem-loop structures in the transcript or helical regions in the DNA template (18-22).

The DNA base sequence at the 3'-end of the yeast 5s rRNA gene (6) and the yeast $tRNA^{Phe}$ gene (23), both transcribed by RNA polymerase III *in vivo*, are presented in figure 6 for comparison. It can be seen that their general structure is remarkably similar to the termination site

recognized by polymerase III defined here. These similarities include the run of AT pairs and the TA and CG interruptions. Because of the location of the sequences which signal termination in this system, we have previously postulated that they may represent the termination site for the 35s ribosomal RNA precursor synthesized by polymerase I *in vivo* (7). The fact that this region is recognized by RNA polymerase III, and the previously established similarities in subunittstructure of polymerase I and III (24, 25), suggest that such regions may represent a general structure required to signal transcription termination. Finally, data presented here indicate that the functional ability to recognize sequence specific sites and terminate transcription is an inherent property of purified yeast RNA polymerase III.

REFERENCES

1. Holland, M.J., Hager, G.L., and Rutter, W.J. (1977) Biochemistry 16, 16-24.
2. Van Keulen, H., and Retel, J. (1977) Eur. J. Biochem. 79, 579-588.
3. Parker, C.S., Ng, S.Y., and Roeder, R.G. (1976) in Molecular Mechanisms in the Control of Gene Expression (Nierlich, D.P., Rutter, W.J., and Fox, C.F.), pp. 223-241, Academic Press, New York.
4. Parker, C.S., and Roeder, R.G. (1977) Proc. Natl. Acad. Sci. U.S.A. 74, 44-48.
5. Ackerman, S., and Furth, J.J. (1979) Biochemistry 18, 3243-3248.
6. Valenzuela, P., Bell, G.I., Masiarz, F.R., De Gennaro, L.J., and Rutter, W.J. (1977) Nature 267, 641-643.
7. Bell, G.I., DeGennaro, L.J., Gelfand, D.H., Bishop, R.J., Valenzuela, P., and Rutter, W.J. (1977) J. Biol. Chem. 252, 8118-8125.
8. So, M., Gill, R., and Falkow, S. (1975) Mol. Gen. Genet. 142, 239-249.
9. Polisky, B., Bishop, R.J., and Gelfand, D.H. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 3900-3904.
10. Meagher, R.B., Tait, R.C., Betlach, M., and Boyer, H.W. (1977) Cell 10, 521-536.
11. Valenzuela, P., Bell, G.I., Weinberg, F., and Rutter, W.J. (1978) Methods in Cell Biol. 18, 1-26.
12. Gillespie, D., and Spiegelman, S. (1965) J. Mol. Biol. 12, 829-842.
13. Southern, E.M. (1975) J. Mol. Biol. 98, 503-517.
14. Brownlee, G.G. (1972) in Laboratory Techniques in Biochemistry and Molecular Biology (Work, T.S., and Work, E., eds.) pp. 66-69, American Elsevier, New York.

15. Chamberlain, M. (1970) Cold Spring Harb. Symp. Quant. Biol. 35, 851-873.

16. Hecht, R.M., and Birnstiel (1972) Eur. J. Biochem. 29, 489-499.

17. Mandel, J.C., and Chambon, P. (1974) Eur. J. Biochem. 41, 367-378.

18. Yanofsky, C., and Soll, L. (1977) J. Mol. Biol. 113, 663-677.

19. Stauffer, G.V., Zurawski, G., and Yanofsky, C. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 4833-4837.

20. Adhya, S. and Gottesman, M. (1978) Ann. Rev. Biochem. 47, 967-996.

21. Rosenberg, M., Court, D., Shimatake, H., Brady, C., and Wulff, D.L. (1978) Nature 272, 414-423.

22. Kupper, H., Sekiya, T., Rosenberg, M., Egan, J., and Landy, A. Nature (1978) 272, 423-428.

23. Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R., and Rutter, W.J. (1978) Proc. Natl. Acad. Sci. U.S.A. 75, 190-194.

24. Valenzuela, P., Hager, G.L., Weinberg, F., and Rutter, W.J. (1976) Proc. Natl. Acad. Sci. U.S.A. 73, 1024-1028.

25. Roeder, R.G. (1976) in RNA Polymerase (Losick, R. and Chamberlain, M., eds.) pp. 285-329, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

26. Rubin, G.M. (1975) Methods in Cell Biology 12, 45-64.

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 252, No. 22, Issue of November 25, pp. 8118-8125, 1977
Printed in U.S.A.

Ribosomal RNA Genes of *Saccharomyces cerevisiae*

I. PHYSICAL MAP OF THE REPEATING UNIT AND LOCATION OF THE REGIONS CODING FOR 5 S, 5.8 S, 18 S, AND 25 S RIBOSOMAL RNAs*

(Received for publication, June 7, 1977)

GRAEME I. BELL, LOUIS J. DEGENNARO, DAVID H. GELFAND,‡ ROBERT J. BISHOP, PABLO VALENZUELA, AND WILLIAM J. RUTTER

From the Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143

The organization of the ribosomal DNA repeating unit from *Saccharomyces cerevisiae* has been analyzed. A cloned ribosomal DNA repeating unit has been mapped with the restriction enzymes *Xma* I, *Kpn* I, *Hin*dIII, *Xba* I, *Bgl* I + II, and *Eco*RI. The locations of the sequences which code for 5 S, 5.8 S, 18 S, and 25 S ribosomal RNAs have been determined by hybridization of the purified RNA species with restriction endonuclease generated fragments of the repeating unit. The position of the 5.8 S ribosomal DNA sequences within the repeat was also established by sequencing the DNA which codes for 83 nucleotides at the 5′ end of 5.8 S ribosomal RNA. The polarity of the 35 S ribosomal RNA precursor has been established by a combination of hybridization analysis and DNA sequence determination and is 5′-18 S, 5.8 S, 25 S-3′.

In the yeast, *Saccharomyces cerevisiae*, as in other eukaryotes, the rRNA genes are highly reiterated. Hybridization analysis suggests that there are approximately 140 cistrons for 18 S and 25 S rRNA per haploid DNA complement (1) and that at least 70% of these are located on chromosome I (2). Other studies have indicated that there are approximately equal numbers of cistrons for 5 S, 5.8 S, 18 S, and 25 S rRNA and that the 5 S rRNA cistrons are interspersed with those for the other rRNA species (3, 4). 5.8 S, 18 S, and 25 S rRNA are derived from a high molecular weight (35 S) precursor. In contrast, 5 S rRNA is synthesized independently, apparently as a primary transcript (5, 6). 35 S rRNA precursor and 5 S rRNA also appear to be transcribed from opposite DNA strands (7). These studies suggest a model for the yeast rDNA cistrons in which the 5 S and 35 S rRNA genes are present in blocks of tandemly repeating units, the basic unit of which contains a single gene for each of these RNA species.

In order to probe the structural organization of this DNA in more detail, we have cloned a yeast rDNA repeating unit as well as fragments of this unit. This paper reports a restriction endonuclease cleavage map for the yeast rDNA repeating unit, the location of the 5 S, 5.8 S, 18 S, and 25 S rRNA coding sequences within the framework of this restriction map and the polarity of the 35 S rRNA gene.

* This work was supported by a grant from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Present address, Cetus Corporation, Berkeley, Calif. 94710.

EXPERIMENTAL PROCEDURES

Purification of Saccharomyces cerevisiae DNA and RNA – Yeast DNA was isolated from commercial yeast and an auxotrophic tetraploid strain 5178 $1C^2 \times 2B^2$, obtained from L. Hartwell (University of Washington, Seattle, Wash.), as described by Holland, *et al.* (8) and Bell *et al.* (9), respectively. 18 S and 25 S rRNAs were prepared from spheroplasts of yeast strain A364A by phenol extraction as described by Udem and Warner (5). The individual rRNA species were resolved by sucrose density gradient centrifugation as described by Rubin (10). Electrophoresis in formamide-polyacrylamide gels indicated that the extent of cross-contamination of the purified RNAs was less than 5%. The low molecular weight rRNAs, 5 S and 5.8 S, were prepared by phenol extraction of whole yeast cells, strain A364A, as described by Rubin (10). 5 S and 5.8 S rRNA were separated by chromatography on Sephadex G-100 as described by Monier (11). Polyacrylamide gel electrophoresis indicated that the 5 S rRNA was apparently homogeneous, while the contamination of 5.8 S rRNA by 5 S rRNA and other low molecular weight RNA species was less than 5%.

Cloning – Yeast DNA, prepared from a commercial strain of yeast, and pSF2124 DNA, a colicin E1 ampicillin-resistant plasmid (12), were digested with *Xma* I or *Eco*RI, ligated, and then used to transform a derivative of *Escherichia coli* K-12, DG75 (an F^- derivative of DG73 (13)). Recombinant clones were selected by ampicillin resistance and inability to produce colicin E1. Yeast rDNA-containing recombinants were selected by analysis of individual colonies for those containing yeast DNA insertions with the molecular weight of the *Xma* I (*Xanthomonas malvacearum*) or *Eco*RI (*Escherichia coli* RY13) generated rDNA fragments and by hybridization to yeast rRNA. Plasmid DNA was purified by the method of Meagher *et al.* (14). All experiments were conducted in a P2 facility using precautions outlined in the "NIH Guidelines for Recombinant DNA Research."

In Vitro Labeling of DNA and RNA – DNA was labeled with [γ-^{32}P]ATP and polynucleotide kinase as described by Maxam and Gilbert (15). RNA was labeled with $Na^{125}I$ according to the procedure developed by Tereba and McCarthy (16).

Enzyme Digestions – Restriction endonuclease *Eco*RI (*Escherichia coli* RY13) was purified and assayed according to Greene *et al.* (17). *Xma* I (*Xanthamonas malvacearum*) was purified according to Roberts[1] and assayed in 6 mM $MgCl_2$, 7 mM 2-mercaptoethanol, and 6 mM Tris/HCl, pH 7.9. *Kpn* I (*Klebsiella pneumoniae* OK8), *Hin*dIII (*Haemophilus influenzae* R_d), *Bam* HI (*Bacillus amyloliquefaciens* H). and *Bgl* I + II (*Bacillus globiggi*) were obtained from New

[1] R. Roberts, unpublished procedures.

England Biolabs (Beverly, Mass.). *Kpn* I, *Hin*dIII and *Bam* HI were assayed in 0.06 M NaCl, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, and 7 mM Tris/HCl, pH 7.4. *Bgl* I + II mixture was assayed in 0.10 M NaCl, 10 mM $MgCl_2$, 7 mM 2-mercaptoethanol, and 10 mM Tris/HCl, pH 7.4. *Xba* I (*Xanthamonas badrii*), *Pst* I (*Providencia stuartii* 164), and *Alu* I (*Arthobacter luteus*), provided by W. Brown of this department, were assayed in 60 mM NaCl, 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, and 7 mM Tris/HCl, pH 7.4; 50 mM NaCl, 5 mM $MgCl_2$, and 100 mM Tris/HCl, pH 7.5; and 7 mM $MgCl_2$, 7 mM 2-mercaptoethanol, and 10 mM Tris/HCl, pH 7.6, respectively. *Sal* I (*Streptomyces albus* 6), provided by F. Bolivar of this department, was assayed in 0.15 M NaCl, 6 mM $MgCl_2$, 8 mM Tris/HCl, pH 7.6, and 50 μg/ml of bovine serum albumin. Digestions were generally performed in a final volume of 0.020 ml and contained 0.5 to 1.0 μg of DNA. After 3 h at 37°, the digestions were terminated with 10 μl of a solution of 25% glycerol, 5% sodium dodecyl sulfate, and 0.025% bromphenol blue. Samples were heated at 65° for 5 min prior to electrophoresis.

Gel Electrophoresis – Samples were loaded on agarose (SeaKem) slab gels (10 × 14 × 0.3 cm) utilizing 40 mM Tris/acetate, pH 8.1, 20 mM sodium acetate, and 2 mM disodium ethylenediaminetetraacetate (18). Electrophoresis was at 25 or 50 mA, constant current, as indicated. Gels were stained with 1 μg/ml of ethidium bromide and photographed (18). Before autoradiography, agarose gels were dried onto DE81 filter paper (19).

Hybridization – After electrophoresis, DNA was transferred to nitrocellulose filters (Millipore) as described by Southern (20). The filters were rolled and placed in a test tube (1.5 × 15 cm) containing 5 × SSC (SSC is 0.15 M NaCl, 0.015 M sodium citrate), 50% formamide, and approximately 1 to 5 × 10^6 cpm of ^{125}I-rRNA. After 24 h at 50°, the filter was removed, washed twice with 250 ml of 5 × SSC and 50% formamide at 50° for 30 min each, rinsed twice in 250 ml of 2 × SSC at room temperature for 10 min each, wrapped in Saran Wrap, and placed on No-Screen x-ray film.

Sequencing of the 5′ End of the 5.8 S rRNA Coding Region – One milligram of pBD4 DNA, which contains the yeast rDNA repeat, was digested with *Eco*RI and the resulting fragments were separated in a preparative 5% polyacrylamide slab gel as described in Valenzuela *et al.* (21). The 375,000 fragment which contains part of the 5.8 S rRNA coding sequences, was eluted from the gel and labeled at the two *Eco*RI sites with [γ-^{32}P]ATP and polynucleotide kinase as described by Maxam and Gilbert (15). The ^{32}P-labeled fragment was digested with *Alu* I which generates three fragments with approximate molecular weights of 210,000, 150,000, and 10,000. The 210,000 and 10,000 fragments were ^{32}P-labeled. The sequence of the 210,000 fragment, which contains the 5.8 S rRNA coding sequences, was determined from the ^{32}P-labeled *Eco*RI site using the chemical method of Maxam and Gilbert (15) as described by Valenzuela *et al.* (21). The ^{32}P-labeled oligonucleotides obtained by base-specific cleavages were separated by electrophoresis in a 20% acrylamide, 7 M urea slab gel.

RESULTS

Restriction Endonuclease Cleavage of Yeast DNA and Cloning of Ribosomal DNA Fragments – *Eco*RI digestion of yeast DNA and γ-DNA, a nuclear fraction which has a unique density and contains the rRNA genes, generates discrete fragments (22, 23), as shown in Fig. 1*A*, *columns 1* and *2*. Three prominent fragments with approximate molecular weights of 1.85 × 10^6, 1.60 × 10^6, and 1.30 × 10^6 are evident on digestion of yeast DNA. These fragments hybridize with 25 S, 5 S, and 18 S rRNA, respectively, and are therefore generated by digestion of the reiterated rDNA cistrons (data not shown; Refs. 22 and 25). In addition to these three high molecular weight fragments, four smaller ones with molecular weights of 375,000, 325,000, 200,000, and 150,000 are also produced during *Eco*RI digestion of the rRNA genes (23). Several of these *Eco*RI fragments have been cloned in the plasmid vector pSF2124 as shown in Fig. 1*A* and in Cramer *et al.* (26), and also in bacteriophage λ (27). Of these cloned fragments of the rDNA repeating unit, the 1.6 × 10^6 fragment which hybridizes with 5 S rRNA, has been the subject of a

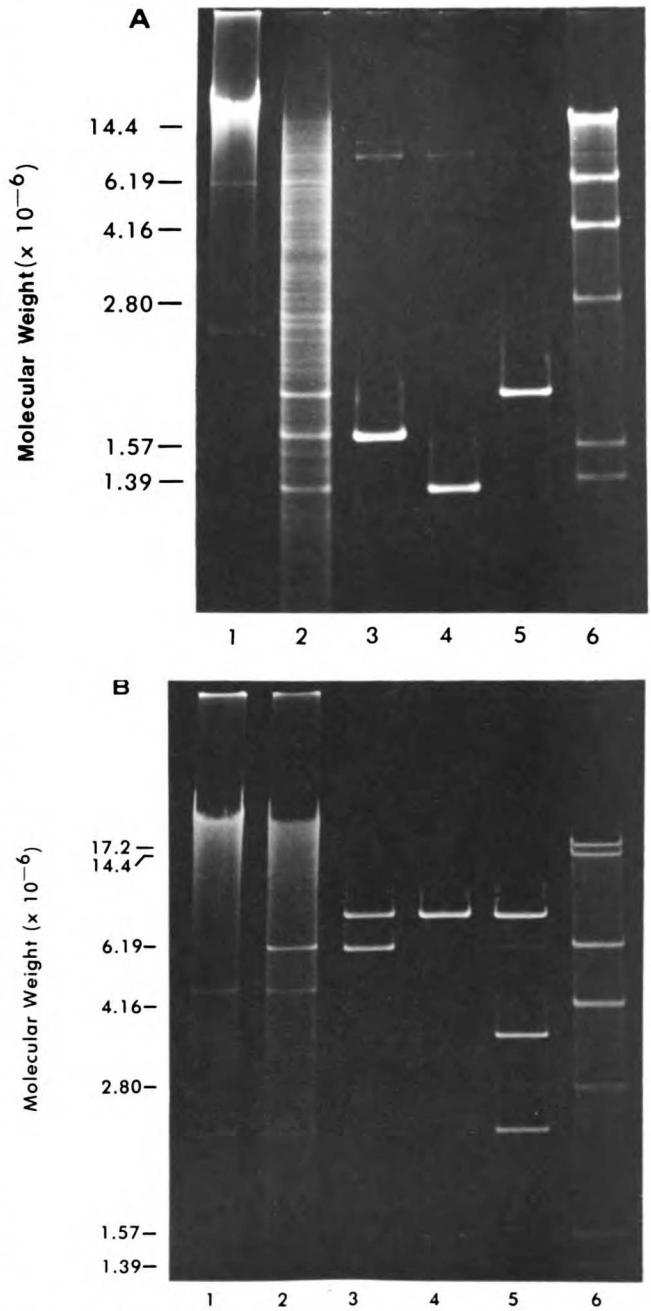


FIG. 1. Agarose gel electrophoresis of restriction endonuclease-digested yeast and plasmid DNA. *A*, *Eco*RI-digested DNA and purified cloned yeast rDNA fragments. *1*, undigested yeast DNA, strain 5178 $1C^2$ × $2B^2$; *2*, *Eco*RI-digested yeast DNA; *3*, purified, cloned 1.60 × 10^6 yeast *Eco*RI-rDNA fragment; *4*, purified, cloned 1.30 × 10^6 yeast *Eco*RI-rDNA fragment; *5*, purified, cloned 1.85 × 10^6 yeast *Eco*RI-rDNA fragment; *6*, *Hin*dIII-digested λ-DNA. Electrophoresis was in 1.0% agarose for 10 h at 25 mA. *B*, *Xma* I-digested yeast DNA and *Xma* I and *Xma* I-*Kpn* I-digested yeast rDNA repeat containing plasmid pBD4. *1*, undigested yeast DNA, strain 5178 $1C^2$ × $2B^2$; *2*, *Xma* I-digested yeast DNA; *3*, *Xma* I-digested pBD4 DNA; *4*, *Xma* I-digested pSF2124 DNA; *5*, *Xma* I-*Kpn* I-digested pBD4 DNA; *6*, *Hin*dIII-digested λ-DNA. Electrophoresis was in 0.7% agarose for 10 h at 25 mA. The discrete, low molecular weight DNA fragments present in undigested yeast DNA are presumably different allomorphs of 2-μm plasmid DNA (24). The 2-μm DNA is a substrate for *Eco*RI but apparently not for *Xma* I.

detailed primary structural analysis, the results of which are presented in Valenzuela *et al.* (21).

Cramer *et al.* (23) have reported that digestion of purified yeast rRNA genes with *Sma* I (*Serratia marcescens* Sb_b) produces a single discrete fragment. Thus, *Sma* I digests the tandemly repeated rRNA genes into monomer units. When yeast DNA was digested with an isoschizomer of *Sma* I, *Xma* I, a large fraction of the DNA appeared to remain intact as shown in Fig. 1*B*, *columns 1* and *2*. However, this enzyme did generate a fragment of approximately 5.8×10^6 which is the molecular weight predicted for the rDNA repeating unit (23). This suggested that this *Xma* I-generated fragment from yeast DNA was the rDNA repeating unit. Therefore, this DNA fragment was cloned in a bacterial plasmid. Yeast and pSF2124 DNA were digested with *Xma* I, ligated, and used to transform *Escherichia coli*. Ampicillin-resistant and colicin E1-nonproducing transformants were selected and the *Xma* I digestion pattern of the plasmid DNA from each clone was determined by agarose gel electrophoresis. A plasmid (pBD4) containing a fragment of 5.8×10^6 was obtained as shown in Fig. 1*B*, *columns 3* and *4*. Subsequent hybridization analysis indicated that this DNA-encoded yeast 5 S, 5.8 S, 18 S, and 25 S rRNAs.

Restriction Endonuclease Map of the Yeast rDNA Repeating Unit—The physical structure of the cloned yeast rDNA repeat was determined by restriction mapping. This DNA fragment was a substrate for the enzymes *Kpn* I, *Xba* I, *Bgl* I + II, *Hin*dIII, and *Eco*RI, as indicated in Fig. 1*B*, *column 5* and Fig. 2, but not for *Sal* I, *Pst* I, or *Bam* HI. The vector DNA, pSF2124, was digested by *Bgl* I + II, *Eco*RI, *Pst* I, and *Bam* HI but not *Kpn* I, *Xba I*, *Hin*dIII, or *Sal* I. The molecular weights of the fragments generated by these enzymes on digestion of the yeast rDNA repeat are summarized in Table I. The positions of the sites for these enzymes within the rDNA repeat were determined by analyzing the products of a partial digest of this DNA as described by Smith and Birnstiel (19) and by combined enzyme digests.

Xma I-digested pBD4 DNA was terminally labeled with [γ-^{32}P]ATP and polynucleotide kinase as described under "Experimental Procedures." The labeled DNA was then digested with *Kpn* I, which has a single site within the rDNA repeat, producing two fragments of 3.5 and 2.3×10^6 daltons (pSF2124 DNA is not a substrate for this enzyme). Following preparative electrophoresis in a 1% agarose gel, the radioactive *Kpn* I-generated fragments were located by autoradiography and excised using the autoradiogram as a template. The DNA was eluted from the agarose using the freeze-squeeze procedure of Thuring *et al.* (28). The purified rDNA fragments, labeled at the *Xma* I site, were used for restriction mapping analysis. The two *Kpn* I fragments were incubated with *Eco*RI for varying intervals of time and the pattern of radioactive partial digestion products analyzed by electrophoresis in a 2% agarose gel. The results are shown in Fig. 3. The molecular weights of the radioactive partial digestions were determined graphically using *Hin*dIII-*Eco*RI-digested λ-DNA standards (29). The positions of the *Eco*RI sites from the radioactive end of the fragment were inferred from the molecular weights of the labeled partial digestion products. There were three *Eco*RI sites in the 3.5×10^6 fragment at approximately 1.45×10^6, 1.8×10^6, and 2×10^6 daltons from the ^{32}P-labeled *Xma* I site and four sites in the 2.3×10^6 fragment at 150,000, 300,000, 1.6×10^6, and 2.0×10^6 daltons from the *Xma* I site. The 150,000 *Eco*RI fragment, Fig. 3, *column 1*, is in fact a doublet as indicated in Table I. It was thus possible to order the eight *Eco*RI-derived fragments within the ribosomal unit. The order of the *Eco*RI fragments within the repeat is 150,000, 150,000, 1.30×10^6, 375,000, 1.85×10^6, 200,000, 325,000, and 1.45×10^6. This is indicated schematically in Fig. 4. The assignment of the *Eco*RI fragments within the *Kpn* I fragments was confirmed by digesting the *Kpn* I-derived rDNA fragments to completion with *Eco*RI and then comparing the molecular weights of the fragments obtained with those from an *Eco*RI digest of *Xma* I-digested pBD4

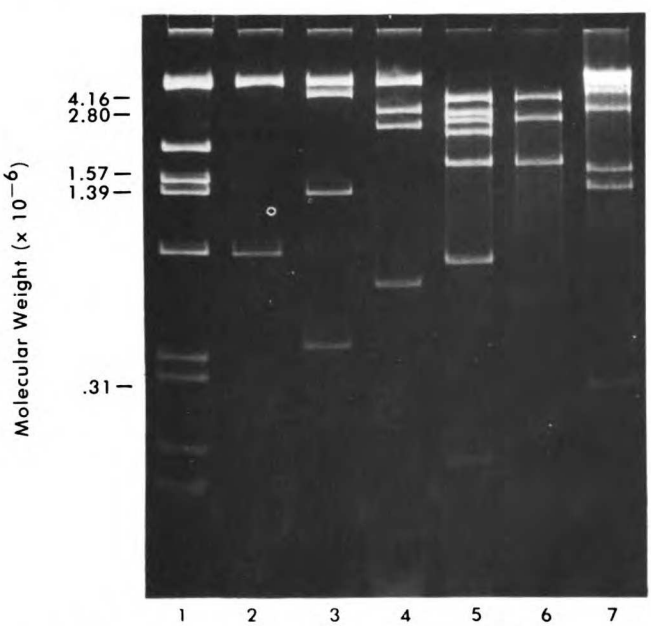


FIG. 2. Agarose gel electrophoresis of restriction endonuclease digested *Xma* I-cleaved yeast rDNA repeat containing plasmid pBD4 and plasmid DNA vector pSF2124. *1*, *Xma* I-*Eco*RI-digested pBD4 DNA; *2*, *Xma* I-*Eco*RI-digested pSF2124 DNA; *3*, *Xma* I-*Hin*dIII-digested pBD4 DNA; *4*, *Xma* I-*Xba* I-digested pBD4 DNA; *5*, *Xma* I-*Bgl* I + II-digested pBD4 DNA; *6*, *Xma* I-*Bgl* I + II-digested pSF2124 DNA; *7*, *Hin*dIII-digested λ-DNA. Electrophoresis was in 2% agarose for 4 h at 50 mA.

TABLE I

Molecular weight and length of restriction fragments from Xma I-generated yeast DNA segment containing the coding sequences for 5 S, 5.8 S, 18 S, and 25 S rRNAs

Enzyme	Fragment
1. Undigested	5.80×10^6 (9300)[a]
2. *Kpn* I	3.50×10^6 (5600); 2.30×10^6 (3700)
3. *Hin*dIII	4.05×10^6 (6500); 1.35×10^6 (2200); 4.10×10^5 (660)
4. *Xba* I	2.80×10^6 (4500); 2.35×10^6 (3800); 6.30×10^5 (1000)
5. *Bgl* I + II	2.70×10^6 (4300); 2.15×10^6 (3400); 7.85×10^5 (1300); 1.85×10^5 (300)
6. *Eco*RI	1.85×10^6 (3000); 1.45×10^6 (2300); 1.30×10^6 (2100); 3.75×10^5 (600); 3.25×10^5 (520); 2.00×10^5 (320); 1.50×10^5[b] (240); 1.50×10^5[b] (240)

[a] Numbers in parentheses indicate the approximate length, in base pairs, of the fragments.

[b] These two fragments cannot be resolved by electrophoresis in a 2% agarose gel but can be separated by electrophoresis in a 5% acrylamide gel.

DNA. The results of this analysis, shown in Fig. 5 and summarized in Table II, indicate that the 3.5 × 10^6 *Kpn* I fragment contains the 1.45 × 10^6, 325,000, and 200,000 *Eco*RI fragments as well as an additional fragment of approximately 1.55 × 10^6. The smaller 2.3 × 10^6 *Kpn* I fragment contains the 1.3 × 10^6, 375,000, and the two 150,000 *Eco*RI fragments as well as a fragment of 300,000. These results confirm that the *Kpn* I site is within the 1.85 × 10^6 *Eco*RI fragment as indicated by the analysis of partial digestion products and validate the relative positions of the *Eco*RI fragments. The positions of the three *Bgl* I + II sites were also determined by an analysis of partial *Bgl* I + II digests of the two *Kpn* I rDNA fragments (data not shown). Fig. 5 also shows complete *Bgl* I + II digestion products of the two *Kpn* I fragments. The molecular weights of these fragments are summarized in Table II. The ^{32}P-labeled terminal fragments are also indicated. Comparison of these values with those obtained on complete digestion of *Xma* I-digested pBD4 DNA supports the designated positions of the *Bgl* I + II sites determined from an analysis of partial digestions. The positions of the two *Xba* I sites were deduced from an analysis of the molecular weights of the digestion products of the two *Kpn* I fragments and the identification of the ^{32}P-labeled terminal fragments, Fig. 5 and Table II. The positions of the two *Hin*dIII sites shown in Fig. 4 were determined from an analysis of the molecular weights of the rDNA-derived fragments produced on digestion of *Xma* I-digested pBD4 with *Hin*dIII and *Kpn* I, *Xba* I, and *Bgl* I + II (data not shown). The orientation of the *Hin*dIII sites was confirmed by isolating the terminal 410,000 and 1.35 × 10^6 *Hin*dIII fragments and digesting them with *Eco*RI. As expected, digestion of the 410,000 fragment produced the two 150,000 fragments and a fragment of 110,000. There were no *Eco*RI sites in the 1.35 × 10^6 fragment. Fig. 4 summarizes the results of the restriction mapping of the yeast rDNA repeat. The order of the *Eco*RI fragments presented in Fig. 4 has also been obtained by Cramer *et al.* (26) and Petes *et al.*[2]

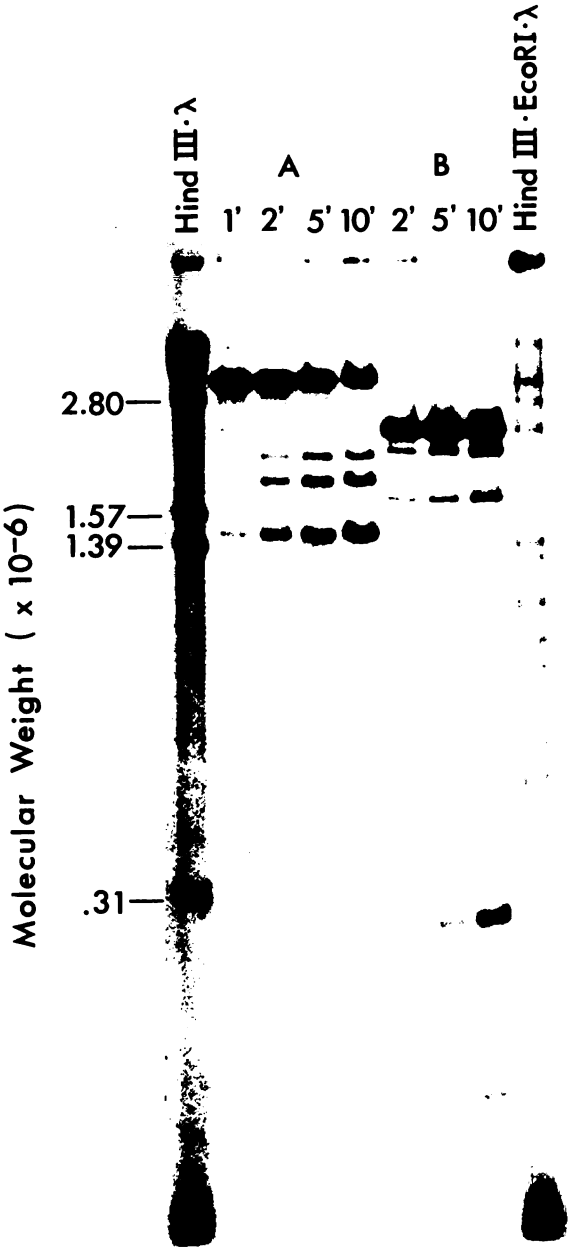


FIG. 3. Autoradiogram of an incomplete *Eco*RI digestion of purified *Kpn* I-digested yeast rDNA fragments ^{32}P-labeled at the *Xma* I site. The ^{32}P-labeled rDNA fragments were incubated with *Eco*RI for the indicated times. The partial digestion products were resolved by electrophoresis in a 2% agarose slab gel (28 × 14 × 0.3 cm) for 10 h at 50 mA. The gel was dried onto DE81 paper as described by Smith and Birnstiel (19) and placed on No-Screen x-ray film. Terminally labeled *Hin*dIII and *Hin*dIII-*Eco*RI-digested λ-DNA fragments were included as standards. The 3.5 × 10^6-dalton *Kpn* I fragment is contaminated with a small amount of the 2.3 × 10^6-dalton *Kpn* I fragment. *A*, 3.5 × 10^6 *Kpn* I-yeast rDNA fragment; *B*, 2.3 × 10^6 *Kpn* I-yeast rDNA fragment.

A comparison of the molecular weights of the three large *Eco*RI-generated fragments produced on digestion of yeast DNA or γ-DNA with those produced from cleavage of the cloned rDNA repeat suggested that the *Xma* I site was within the 1.60 × 10^6-dalton *Eco*RI fragment which hybridizes with 5 S rRNA (described previously). Thus, *Xma* I digestion produces two fragments with molecular weights of 1.45 × 10^6 and 150,000. The former fragment hybridizes with 5 S rRNA. The presence of the *Xma* I site within the 1.60 × 10^6 *Eco*RI fragment has also been established by restriction mapping and sequence analysis of this fragment (21). The correspondence of the restriction sites and fragment molecular weights within the 1.60 × 10^6 *Eco*RI rDNA fragment and the 1.45 × 10^6 and 150,000 *Eco*RI fragments derived from the cloned rDNA repeat proves that there is only a single *Xma* I site within the yeast rDNA repeat and that no low molecular weight *Xma* I-derived fragments are unaccounted for. This is consistent with the yeast rDNA repeating units as tandemly repeating "head-to-tail" units as proposed by Cramer *et al.* (23).

Localization of 5 S, 5.8 S, 18 S, and 25 S rRNA Coding Regions within the Yeast rDNA Repeating Unit and Polarity of 35 S rRNA Precursor—Having developed a restriction map for the yeast rDNA monomer, it was now possible to identify the DNA fragments which coded for each of the rRNAs by hybridization of the labeled rRNA species with isolated restriction fragments. Previous experiments using *Eco*RI-digested yeast DNA, and plasmids or λ phage containing *Eco*RI-generated fragments of yeast rDNA, demonstrated that the 1.85 × 10^6-dalton fragment hybridized with 25 S rRNA, the 1.60 × 10^6-dalton fragment with 5 S rRNA and the 1.30 × 10^6-dalton fragment with 18 S rRNA (22, 25, 27). In addition, analysis of the sequence of 5.8 S rRNA (30) indicated

[2] T. D. Petes, L. Hereford, and K. G. Skryabin, personal communication.

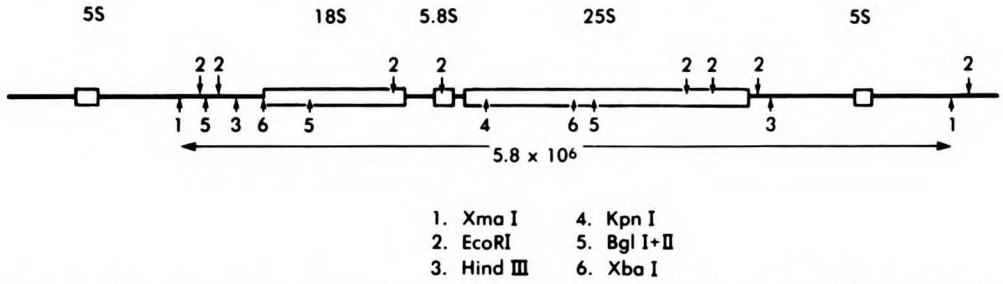


FIG. 4. Restriction map of yeast rDNA repeating unit generated with the endonucleases *Xma* I, *Eco*RI, *Hin*dIII, *Kpn* I, *Bgl* I + II, and *Xba* I. The approximate positions of the 5 S, 5.8 S, 18 S, and 25 S rRNA coding regions are also indicated. The 5′ end of 18 S rRNA appears to be between the *Xba* I and *Bgl* I + II sites. The exact positions of the 5′ and 3′ ends of 18 S and 25 S rRNAs have not been determined.

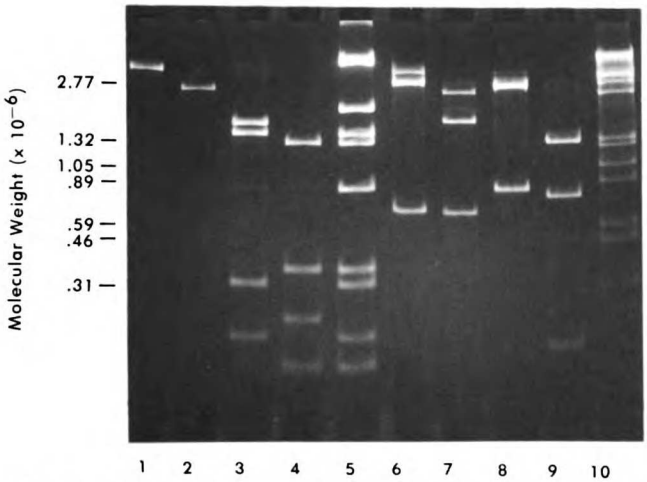


FIG. 5. Agarose gel electrophoresis of restriction endonuclease digests of purified *Kpn* I-generated rDNA fragments. *1*, 3.5 × 10^6-dalton *Kpn* I fragment, undigested; *2*, 2.3 × 10^6-dalton *Kpn* I fragment, undigested; *3*, *Eco*RI-digested 3.5 × 10^6-dalton fragment; *4*, *Eco*RI-digested 2.3 × 10^6-dalton fragment; *5*, *Xma* I-*Eco*RI-digested pBD4 DNA; *6*, *Xba* I-digested 3.5 × 10^6-dalton fragment (reaction is incomplete); *7*, *Xba* I-digested 2.3 × 10^6-dalton fragment; *8*, *Bgl* I + II-digested 3.5 × 10^6-dalton fragment; *9*, *Bgl* I + II-digested 2.3 × 10^6-dalton fragment; *10*, *Hin*dIII-*Eco*RI-digested λ-DNA. Electrophoresis was in 2% agarose for 4 h at 50 mA.

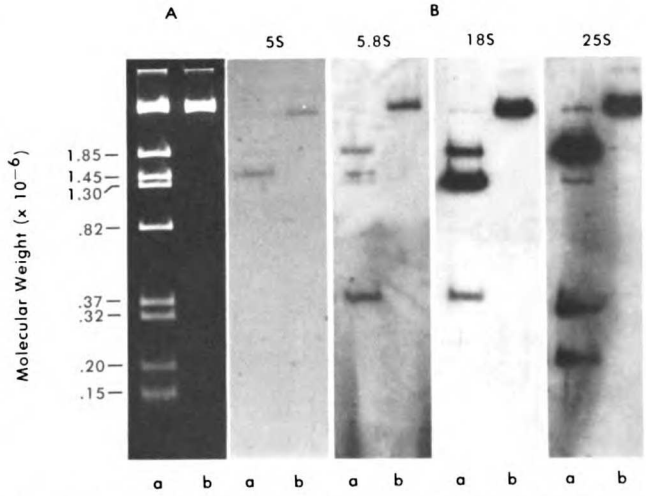


FIG. 6. Hybridization of ^{125}I-labeled 5 S, 5.8 S, 18 S, and 25 S rRNA to *Eco*RI fragments of the yeast rDNA repeat. The hybrid plasmid, pBD4, containing the yeast rDNA repeating unit was digested with *Xma* I to excise the inserted yeast DNA fragment and then with *Eco*RI. The DNA digest was analyzed in a 2% agarose gel. The separated fragments were transferred to nitrocellulose filters as described by Southern (20) and then hybridized with ^{125}I-labeled rRNA as indicated under "Experimental Procedures." *A*, photograph of the ethidium bromide-stained agarose gel. In each case, *a* denotes the *Eco*RI-digested rDNA repeat and *b* is an undigested control for hybridization. In *column a*, the largest unlabeled DNA fragment and the 820,000-dalton fragment are derived from pSF2124 DNA. Also, as described in the text, the 150,000-dalton fragment is a doublet. *B*, autoradiographs of the nitrocellulose filters after hybridization with the indicated, purified ^{125}I-labeled rRNA species.

TABLE II

Molecular weights of DNA fragments generated on restriction endonuclease digestion of Kpn I fragments of yeast rDNA unit

Enzyme	Fragment
A. Undigested	3.5 × 10^6
*Eco*RI	1.55 × 10^6; 1.45 × 10^{6a}; 325,000; 200,000
Bgl I + II	2.70 × 10^{6a}; 820,000
Xba I	2.8 × 10^{6a}; 660,000
B. Undigested	2.3 × 10^6
*Eco*RI	1.30 × 10^6; 375,000; 300,000; 150,000; 150,000[a]
Bgl I + II	1.30 × 10^6; 790,000; 185,000[a]
Xba I	1.65 × 10^6; 630,000[a]

[a] ^{32}P-labeled terminal fragment.

that the DNA coding for this RNA species would be cut by *Eco*RI. *Xma* I-digested pBD4 was incubated with *Eco*RI and the resulting fragments separated in a 2% agarose gel and then transferred to nitrocellulose filters by the technique described by Southern (20). Identical filters containing *Eco*RI-digested DNA and an undigested control were then hybridized with ^{125}I-labeled 5 S, 5.8 S, 18 S, and 25 S rRNA. The results are presented in Fig. 6. The 5 S rRNA hybridized to the 1.45 × 10^6 fragment, 5.8 S rRNA hybridized predominantly with the 1.85 × 10^6 and 375,000 fragments and to a lesser extent with the 1.45 × 10^6 and possibly the 1.30 × 10^6 fragments. The small amount of hybridization to these later two fragments probably represents hybridization of contaminating 5 S and 18 S rRNA, respectively. The 18 S rRNA hybridized to the 1.30 × 10^6 and 375,000 fragments. The hybridization to the 1.85 × 10^6 fragment is presumably due to contaminating 25 S rRNA. The 25 S rRNA hybridized predominantly with 1.85 × 10^6, 325,000, and 200,000 fragments. There was a very small amount of hybridization to the 1.30 × 10^6 fragment, probably due to contaminating 18 S rRNA. There was no observed hybridization of any of the rRNA species to the 150,000 fragments.

When the order of the *Eco*RI fragments within the rDNA repeating unit, Fig. 4, and the hybridization data are consid-

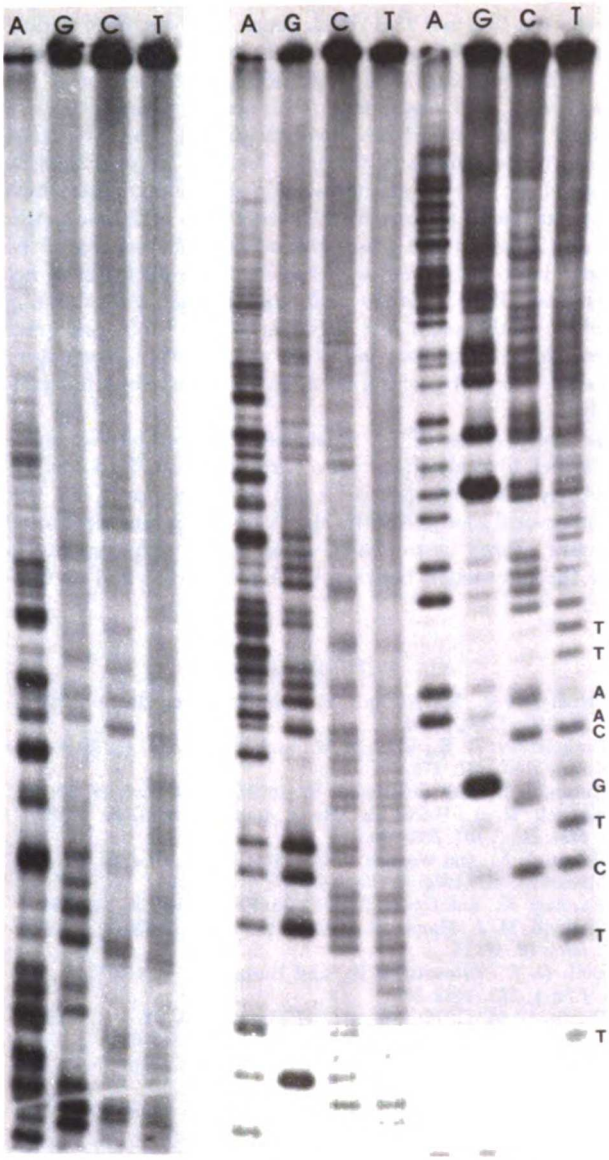


FIG. 7. Autoradiogram of the electrophoretic separation of ^{32}P-labeled oligonucleotides obtained by base-specific cleavages of the portion of the 375,000 *Eco*RI fragment containing the 5.8 S rRNA coding sequences. The columns *A*, *G*, *C*, and *T* indicate wells loaded with products from reactions partially specific for adenine, guanine, cytosine, and thymine, respectively, at 0, 12, and 24 h, from left to right. The reading of bases 3 to 12 from the ^{32}P-labeled 5′ end is illustrated. At the second base from the 5′ end, adenine is just apparent at the bottom of the gel. The first *A* from the 5′ end, determined from the specificity of *Eco*RI (31), ran off the gel. The sequence determined is complementary to the 5′ end of 5.8 S rRNA. The cytosine reactions also produced a more slowly migrating "ghost" in the autoradiogram probably due to incomplete phosphate elimination.

ered together, they are consistent with a model in which 5 S rRNA hybridizes with the 1.45 × 10^6 fragment, 5.8 S rRNA with the 375,000 and 1.85 × 10^6 fragments, 18 S rRNA with the 375,000 and 1.30 × 10^6 fragments, and 25 S rRNA with the 200,000, 325,000, and 1.85 × 10^6 fragments. The relative positions of the sequences coding for each of the rRNA species are schematically presented in Fig. 4. The position of the 5 S rRNA gene near the middle of the 1.45 × 10^6 fragment has

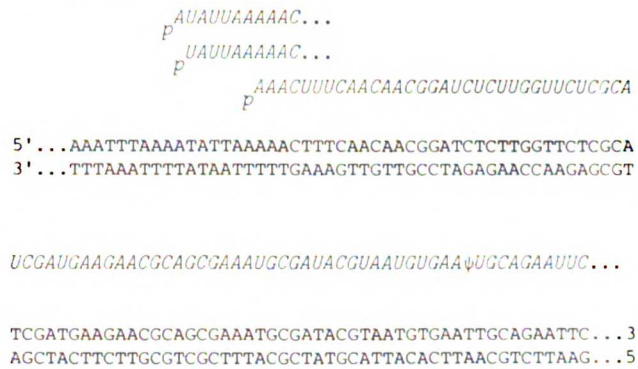


FIG. 8. Nucleotide sequence of 5′ end and adjacent regions of 5.8 S rRNA gene. The sequence of yeast 5.8 S rRNA is in *italics*. The 5′ ends of each of the RNA species are indicated (32). The major species, comprising 90% of the cellular 5.8 S rRNA, has the 5′ sequence of pAAAC.

been established by more detailed hybridization analysis and nucleotide sequence determination (21). To further confirm the presence of the 5.8 S rRNA gene within the 375,000-dalton *Eco*RI DNA fragment and also to establish the polarity of the 35 S rRNA precursor, part of this DNA fragment was sequenced. The sequence obtained, which is shown in Fig. 7 and 8, includes 83 nucleotides which code for the 5′ end of the major 5.8 S rRNA species (30) and 16 nucleotides from the adjacent transcribed spacer of the 35 S rRNA gene. Skryabin *et al.*[3] have also sequenced part of this 5.8 S rRNA gene. This DNA sequence analysis confirms the RNA sequence and hybridization analyses which predicted an *Eco*RI site within the 5.8 S rRNA gene. The polarity of the 35 S rRNA transcript, 5′-18 S, 5.8 S, 25 S-3′, is also established from this sequence analysis and the hybridization data described above. The 5.8 S rRNA sequence deduced from the DNA sequence agrees exactly with that determined by Rubin (30) for the major 5.8 S rRNA species. The DNA sequence presented here also contains sequences which code for the two minor 5.8 S rRNA species which differ only at their 5′ ends from the major RNA species (32). Thus, the presence of these three 5.8 S rRNAs is probably attributable to differences in the processing of a single molecular species of 35 S rRNA rather than to three distinct 35 S rRNA precursors. The presence of 5.8 S rDNA sequences between those coding for 18 S and 25 S rRNAs has also been observed in *Dictyostelium* and vertebrates (33, 34). The data presented in Fig. 6 indicate that 18 S and 25 S rRNAs hybridize predominantly with the 1.30 × 10^6 and 1.85 × 10^6 *Eco*RI fragments, respectively. Therefore, these fragments probably contain most of the DNA sequences coding for these two rRNA species. The localization of the domains for the 18 S and 25 S rRNA genes within the *Eco*RI fragments which hybridize with these RNAs has not been determined by more detailed hybridization or nucleotide sequence analyses. However, preliminary experiments suggest that the 5′ end of 18 S rRNA is between the *Xba* I and *Bgl* I + II site as indicated in Fig. 4. The relative sizes of the regions which code for these RNA species have been deduced from a comparison of their molecular weights (5) and the hybridization data.

[3] K. G. Skryabin, A. Maxam, and W. Gilbert, personal communication.

DISCUSSION

The yeast rDNA repeating unit consists of a single gene for 5 S rRNA and the 35 S rRNA precursor. The units are joined in a "head-to-tail" arrangement to generate tandemly repeated arrays. The data presented here and in Valenzuela *et al.* (21) indicate the positions of the coding sequences for 5 S, 5.8 S, 18 S, and 25 S rRNAs within the rDNA repeating unit. Furthermore, the data establish that the 5 S, 18 S, and 25 S rRNA coding sequences are linked as suggested by Rubin and Sulston (3) and that the 5 S rRNA gene is approximately 1.2 $\times 10^6$ daltons (1900 base pairs) and 8×10^5 daltons (1300 base pairs) from the 5′ end of the 18 S rRNA and 3′ end of the 25 S rRNA coding sequences, respectively. Of course, the 5 S rRNA gene is even nearer the ends of the 35 S rRNA precursor coding sequences.

The molecular weight of the yeast 35 S rRNA precursor is uncertain, but several estimates suggest that it is approximately 2.5 to 2.8×10^6 (5, 35). The rDNA repeating unit has a molecular weight of 5.8×10^6 which suggests that 86 to 96% of the repeating unit codes for the 35 S rRNA precursor. The 5 S rRNA coding sequence comprises 121 base pairs (6) or approximately 1% of the rDNA repeating unit. Thus, 3 to 13% of the repeat is nontranscribed DNA or has as yet undetermined function(s). This nontranscribed DNA must contain the control sequences for the synthesis of 5 S and 35 S rRNAs.

The data presented here establish that the polarity of the 35 S rRNA precursor is 5′-18 S, 5.8 S, 25 S-3′. The same polarity for the 35 S rRNA precursor has also been deduced from a kinetic analysis of rRNA synthesis in yeast (36). Thus, the polarity of the high molecular weight rRNA precursor is the same for yeast and vertebrates (37, 38). The sequence of 5 S rDNA and adjacent regions has been determined by Valenzuela *et al.* (21) and, therefore, establishes the strand of the rDNA repeat unit from which 5 S rRNA is synthesized. Their results indicate that 5 S and 35 S rRNAs are synthesized from DNA strands with opposite polarity, that is, from complementary strands. This also confirms the results of Aarstad and Oyen (7) who observed that 5 S rRNA and 18 S and 25 S rRNA hybridized to complementary strands of yeast rDNA.

The yeast, *Saccharomyces cerevisiae*, rDNA repeating unit with a molecular weight of 5.8×10^6 is approximately 50% larger than an *E. coli* rRNA precursor gene (39) but is the smallest rDNA repeating unit described for a eukaryote. It is smaller than those of *Tetrahymena*, *Physarum*, and *Dictyostelium* which have molecular weights of approximately 6.2×10^6, 19.5×10^6, and 25×10^6, respectively (40–44), and also of higher eukaryotes. For example, *Drosophila melanogaster*, *Lytechinus variegatus*, *Xenopus laevis*, and *Bos taurus* have rDNA repeat molecular weights of 7 to 11×10^6, 8×10^6, 6.8 to 10.5×10^6, and 21×10^6, respectively (45–48). The significance of this difference in size of the rDNA repeating units of eukaryotes is unknown. The yeast rDNA repeating units also seem to be homogeneous (23) like those described from *Tetrahymena* (40) and *Dictyostelium* (44), but unlike those from *Drosophila* (45), *Lytechinus variegatus* (46), and *Xenopus laevis* (47, 49). In the latter three species, chromosomal, sex-linked, and individual differences in the structure of the units have been observed.

In yeast as in bacteria, the 5 S rRNA gene is linked with the genes for the large rRNA species. In yeast, however, the 5 S and large rRNAs are synthesized independently (5) and are not part of a common precursor as they are in *E. coli* (50). The *E. coli* rRNA precursor gene also contains a tRNA gene between the 16 S and 23 S rRNA genes. We have been unable to detect any tRNA genes within the yeast rDNA repeating unit by hybridization with yeast tRNA or comparison of the known yeast tRNA sequences with the sequences we have obtained for several rDNA spacer segments. The 5 S rRNA genes are also linked with those for the large rRNAs in *Dictyostelium* (43) but are unlinked in other eukaryotes which have been examined, including *Tetrahymena* (41), *Drosophila* (51), and *Xenopus* (52). The significance of the linkage of the 5 S and large rRNA genes in some lower eukaryotes is not understood. It is possible that the linkage and opposing polarities of the yeast 5 S rRNA and 35 S rRNA precursor genes allows these two genes to share certain regulatory elements but still permits them to be transcribed independently.

The cloned yeast rDNA repeat is an appropriate template with which to study the *in vitro* transcription of 35 S and 5 S rRNAs by RNA polymerases I and III, respectively (53, 54). The structural analysis of the rDNA repeating unit presented here now allows us to design more incisive experiments to probe the interaction between RNA polymerase and its template.

REFERENCES

1. Schweizer, E., MacKechnie, C., and Halvorson, H. O. (1969) *J. Mol. Biol.* **40**, 261–277
2. Finkelstein, D. B., Blamire, J., and Marmur, J. (1972) *Nature New Biol.* **240**, 279–281
3. Rubin, G. M., and Sulston, J. E. (1973) *J. Mol. Biol.* **79**, 521–530
4. Kaback, D. B., Halvorson, H. O., and Rubin, G. M. (1976) *J. Mol. Biol.* **107**, 385–390
5. Udem, S. A., and Warner, J. R. (1972) *J. Mol. Biol.* **65**, 227–242
6. Miyazaki, M. (1974) *J. Biochem.* **75**, 1407–1410
7. Aarstad, K., and Oyen, T. B. (1975) *FEBS Lett.* **51**, 227–231
8. Holland, M. J., Hager, G. L., and Rutter, W. J. (1977) *Biochemistry* **16**, 16–24
9. Bell, G. I., Valenzuela, P., and Rutter, W. J. (1977) *J. Biol. Chem.* **252**, 3082–3091
10. Rubin, G. M. (1975) *Methods Cell Biol.* **12**, 45–64
11. Monier, R. (1971) *Proc. Nucleic Acid Res.* **2**, 618–622
12. So, M., Gill, R., and Falkow, S. (1975) *Mol. Gen. Genet.* **142**, 239–249
13. Polisky, B., Bishop, R. J., and Gelfand, D. H. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73**, 3900–3904
14. Meagher, R. B., Tait, R. C., Betlach, M., and Boyer, H. W. (1977) *Cell* **10**, 521–536
15. Maxam, A., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 560–564
16. Tereba, A., and McCarthy, B. J. (1973) *Biochemistry* **12**, 4675–4679
17. Greene, P. J., Betlach, M. C., Goodman, H. M., and Boyer, H. W. (1974) in *Methods in Molecular Biology* (Wickner, R. B., ed) Vol. 1, pp. 87–111, Marcel Dekker, New York
18. Helling, R. B., Goodman, H. M., and Boyer, H. W. (1974) *J. Virol.* **14**, 1235–1244
19. Smith, H. O., and Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387–2398
20. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517
21. Valenzuela, P., Bell, G. I., Venegas, A., Sewell, E. T., Masiarz, F. R., DeGennaro, L. J., Weinberg, F., and Rutter, W. J. (1977) *J. Biol. Chem.* **252**, 8126–8135
22. Nath, K., and Bollon, A. P. (1975) *Nature* **257**, 155–157
23. Cramer, J. H., Farrelly, F. W., and Rownd, R. H. (1976) *Mol. Gen. Genet.* **148**, 233–241
24. Livingston, D. M., and Klein, H. L. (1977) *J. Bacteriol.* **129**, 472–481
25. Nath, K., and Bollon, A. P. (1976) *Mol. Gen. Genet.* **147**, 153–168
26. Cramer, J. H., Farrelly, F. W., Barnitz, J. T., and Rownd, R. H. (1977) *Mol. Gen. Genet.* **151**, 229–244
27. Kramer, R. A., Cameron, J. R., and Davis, R. W. (1976) *Cell* **8**, 227–232

28. Thuring, R. W. J., Sanders, J. P. M., and Borst, P. (1975) *Anal. Biochem.* **66**, 213-220
29. Murray, K., and Murray, N. E. (1975) *J. Mol. Biol.* **98**, 551-564
30. Rubin, G. M. (1973) *J. Biol. Chem.* **248**, 3860-3875
31. Hedgpeth, J., Goodman, H. M., and Boyer, H. W. (1972) *Proc. Natl. Acad. Sci. U. S. A.* **69**, 3448-3452
32. Rubin, G. M. (1974) *Eur. J. Biochem.* **41**, 197-202
33. Spiers, J., and Birnstiel, M. (1974) *J. Mol. Biol.* **87**, 237-256
34. Frankel, G., Cockburn, A. F., Kindle, K. L., and Firtel, R. A. (1977) *J. Mol. Biol.* **109**, 539-558
35. Dudov, K., Dabeva, M., and Hadijilov, A. (1976) *Anal. Biochem.* **76**, 250-258
36. Trapman, J., and Planta, R. J. (1975) *Biochim. Biophys. Acta* **414**, 115-125
37. Dawid, I. B., and Wellauer, P. K. (1976) *Cell* **8**, 443-448
38. Reeder, R. H., Higashinakagawa, T., and Miller, O., Jr. (1976) *Cell* **8**, 449-454
39. Steitz, J. A., Ginsburg, D., Garwin, J. L., and Goldberg, D. A. (1976) in *Control of Ribosome Synthesis* (Kjeldgaard, N. O., and Maaloe, O., eds) pp. 178-189, Academic Press, New York
40. Gall, J. G. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71**, 3078-3081
41. Tonnesen, T., Engberg, J., and Leick, V. (1976) *Eur. J. Biochem.* **63**, 399-407
42. Molgaard, H. V., Matthews, H. R., and Bradbury, E. M. (1976) *Eur. J. Biochem.* **68**, 541-549
43. Maizels, N. (1976) *Cell* **9**, 431-438
44. Cockburn, A. F., Newkirk, M. J., and Firtel, R. A. (1976) *Cell* **9**, 605-613
45. Tartof, K. D., and Dawid, I. B. (1976) *Nature* **203**, 27-30
46. Wilson, F. E., Blin, N., and Stafford, D. W. (1976) *Chromosoma (Ber.)* **58**, 247-253
47. Wellauer, P. K., Reeder, R. H., Dawid, I. B., and Brown, D. D. (1976) *J. Mol. Biol.* **105**, 487-505
48. Blin, N., Stephenson, E. C., and Stafford, D. W. (1976) *Chromosoma (Ber.)* **58**, 41-50
49. Reeder, R. H., Brown, D. D., Wellauer, P. K., and Dawid, I. B. (1976) *J. Mol. Biol.* **105**, 507-516
50. Nomura, M. (1976) *Cell* **9**, 633-644
51. Wimber, D. E., and Steffensen, D. M. (1970) *Science* **170**, 630-641
52. Pardue, M. L., Brown, D. D., and Birnstiel, M. (1973) *Chromosoma (Ber.)* **42**, 191-203
53. Valenzuela, P., Weinberg, F., Bell, G., and Rutter, W. J. (1976) *J. Biol. Chem.* **251**, 1464-1470
54. Valenzuela, P., Hager, G. L., Weinberg, F., and Rutter, W. J. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73**, 1024-1028

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 252, No. 22, Issue of November 25, pp. 8126-8135, 1977
Printed in U.S.A.

Ribosomal RNA Genes of *Saccharomyces cerevisiae*

II. PHYSICAL MAP AND NUCLEOTIDE SEQUENCE OF THE 5 S RIBOSOMAL RNA GENE AND ADJACENT INTERGENIC REGIONS*

(Received for publication, June 7, 1977)

PABLO VALENZUELA,‡ GRAEME I. BELL, ALEJANDRO VENEGAS,‡, § ELAINE T. SEWELL, FRANK R. MASIARZ, LOUIS J. DEGENNARO, FANYELA WEINBERG, AND WILLIAM J. RUTTER

From the Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143

A DNA fragment containing the structural gene for the 5 S ribosomal RNA and intergenic regions before and after the 35 S ribosomal RNA precursor gene of *Saccharomyces cerevisiae* has been amplified in a bacterial plasmid and physically mapped by restriction endonuclease cleavage and hybridization to purified yeast 5 S ribosomal RNA.

The nucleotide sequence of the DNA fragments carrying the 5 S ribosomal RNA gene and adjacent regions has been determined. The sequence unambiguously identifies the 5 S ribosomal RNA gene, determines its polarity within the ribosomal DNA repeating unit, and reveals the structure of its promoter and termination regions. Partial DNA sequence of the regions near the beginning and end of the 35 S ribosomal RNA gene has also been determined as a preliminary step in establishing the structure of promoter and termination regions for the 35 S ribosomal RNA gene.

The 5 S, 5.8 S, 18 S, and 25 S rRNA genes of *Saccharomyces cerevisiae* are closely linked in a unit with a molecular weight of approximately 5.8×10^6 which is repeated 140 times per haploid DNA complement (1-3). The physical structure of the rDNA repeating unit has been determined and the approximate positions of each of the rRNA genes mapped (3-5). These studies localized the 5 S rRNA gene in a 2.5-kb (kilobase pairs) *Eco*RI-generated DNA fragment which hybridizes with this rRNA, but not with 5.8 S, 18 S, or 25 S rRNAs.

The 121-bp (base pairs) 5 S rRNA from *S. cerevisiae*, whose sequence has been determined by Miyazaki (6), possesses a 5'-terminal triphosphate. This and the analysis of its biosynthesis (7) suggest that this RNA is not derived from a larger precursor. Thus, the control regions for its biosynthesis should be adjacent to the DNA sequences from which it is transcribed and would, therefore, also be expected within the 2.5-kb DNA fragment. In contrast, 5.8 S, 18 S, and 25 S rRNAs are produced from a high molecular weight (35 S) precursor. The domain of the 35 S rRNA precursor is unknown, but a comparison of the amount of DNA required to code for this RNA, approximately 5×10^6 daltons (8), and the size of the yeast rDNA repeating unit, 5.8×10^6 daltons, suggests that the coding regions for the 3' and 5' ends of this molecule may also be within the 2.5-kb *Eco*RI fragment.

Therefore, we have initiated a detailed analysis of this region of the rDNA repeating unit. We have determined precisely the position of the single 5 S rRNA gene within this fragment by hybridization with 5 S rRNA and subsequently by DNA sequence determination. The sequence of the DNA flanking the 5 S rRNA gene was also determined in order to elucidate the structure of the promoter and terminator regions of a eukaryotic gene. The sequences at the ends of the 2.5-kb DNA fragment have also been examined since these may contain the signals for initiation and termination of transcription of the 35 S rRNA gene. A comparison of the DNA sequences for these intercistronic regions with those from the vicinity of the 5 S rRNA gene indicated the presence of a sequence which was similar to one at the termination region (5' end of the coding strand) of the 5 S rRNA gene. Therefore, the 2.5-kb DNA fragment appears to contain the termination region for the synthesis of the 35 S rRNA precursor. We have not been able to localize the promoter region of the 35 S rRNA precursor.

Preliminary accounts of the sequence of the 5 S rRNA gene have been reported by Valenzuela *et al.* (9) and Maxam *et al.* (10).

EXPERIMENTAL PROCEDURES

Restriction Endonucleases – Endonucleases *Hae* III (*Haemophilus aegyptius*), *Hha* I (*Haemophilus haemolyticus*), and *Hpa* II (*Haemophilus parainfluenzae*) were obtained from New England Biolabs, Beverly, Mass. *Alu* I (*Arthrobacter luteus*), *Hin*dII (*Haemophilus influenzae*), and *Hpa* I were gifts from W. Brown and H. M. Goodman of this department. *Dpn* II (*Diplococcus pneumoniae*) was a gift from F. W. Studier, Brookhaven National Laboratory, Upton, N. Y. *Taq* I (*Thermus aquaticus* YT1) was purified and assayed as described by Sato *et al.* (11). *Alu* I, *Hae* III, *Hha* I, *Hpa* I, and *Hpa* II were assayed in 6 mM $MgCl_2$, 6 mM NaCl, 7 mM 2-mercaptoethanol, and 10 mM Tris/HCl, pH 7.4. *Hin*dII was assayed in 6 mM $MgCl_2$, 60 mM NaCl, 7 mM 2-mercaptoethanol, and 10 mM Tris/HCl, pH 7.5. *Dpn* II was assayed in 6 mM $MgCl_2$, 20 mM NaCl, 7 mM 2-

* This research was supported by Grant GM 21830 from the National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ On leave from Departamento de Biologia Celular, Universidad Catolica, Santiago, Chile.

§ Recipient of a National Institutes of Health International Fellowship.

mercaptoethanol, 5 mM Tris/HCl, pH 7.5, and 50 μg/ml of bovine serum albumin. *Eco*RI (*Escherichia coli* RY13), *Xma* I (*Xanthamonas malvacearum*), *Kpn* I (*Klebsiella pneumoniae*), *Hin*dIII, *Bgl* I + II (*Bacillus globiggii*), *Xba* I (*Xanthomonas badrii*), *Pst* I (*Providencia stuartii*), *Bam* HI (*Bacillus amyloliquefaciens* H), and *Sal* I (*Streptomyces albus* G) were obtained and assayed as described in Bell *et al.* (3).

Isolation of Yeast rDNA Fragment Containing 5 S rRNA Gene – The *Eco*RI-generated yeast 2.5-kb rDNA fragment containing the 5 S rRNA gene was cloned using pSF2124 as vector as described in Bell *et al.* (3). Plasmid DNA was isolated as described by Meagher *et al.* (12). After *Eco*RI digestion, the yeast 2.5-kb insert was separated from the bacterial vector DNA by sucrose gradient centrifugation. Five to twenty per cent (w/v) sucrose gradients were prepared in 1 M NaCl, 25 mM Tris/HCl, pH 8, and 1 mM EDTA. Centrifugation was for 20 h at 20° in an SW27 rotor.

DNA Labeling – DNA fragments were dephosphorylated by incubation with bacterial alkaline phosphatase (Worthington) and then labeled at their 5′ ends using polynucleotide kinase (P-L Laboratories) and [γ-^{32}P]ATP as described by Maxam and Gilbert (13).

Gel Electrophoresis – Analytical and preparative electrophoreses of restriction fragments were performed in 5% or 8% acrylamide slab gels in 50 mM Tris/borate (pH 8.3) and 1 mM EDTA buffer (13) at constant 10 V/cm. Samples were applied in 2% sodium dodecyl sulfate, 12% glycerol, 5 mM EDTA, and 0.50% bromphenol blue. Electrophoretic DNA strand separation was performed in 8% acrylamide gels at constant 10 V/cm. The samples were applied in 0.3 M NaOH, 10% glycerol, 1 mM EDTA, 0.05% xylene cyanol, 0.05% bromphenol blue. Electrophoretic fractionation of oligonucleotides for DNA sequence was carried out as described by Maxam and Gilbert (13) in 20% acrylamide gels containing 7 M urea, 50 mM Tris/borate (pH 8.3), 1 mM EDTA at constant 20 V/cm. Agarose gel electrophoresis was performed as described in Bell *et al.* (3). DNA in gels was visualized by fluorescence after staining with 1 μg ml of ethidium bromide or by autoradiography with Kodak No-Screen medical x-ray film.

Purification of Restriction Fragments – Preparative amounts of restriction fragments were obtained by digesting 50 to 200 μg of DNA with the appropriate restriction enzyme. The products were fractionated by preparative acrylamide gel electrophoresis. After visualization, the DNA bands were excised, minced, and extracted three times with 500 mM *ammonium* acetate, 10 mM magnesium acetate, 0.1 mM EDTA, and 0.1% sodium dodecyl sulfate. Salt removal and concentration of the extracts were achieved by ethanol precipitation (3 volumes, −70°, 15 min).

DNA Sequencing – The chemical method of Maxam and Gilbert (13) was used. Preferential cleavage at guanine was achieved by methylation with 50 mM dimethylsulfate (10 min, 21°), ring opening and displacement at neutral pH, and phosphate elimination by 1 M piperidine (60 min, 90°). Cleavage at adenine was achieved by methylation with dimethylsulfate (23 min, 21°), depurination in 0.1 N HCl (120 min, 0°), and phosphate elimination by 0.1 N NaOH (30 min, 90°). Modification and cleavage at cytosine and thymine residues was achieved by treatment with 15 M hydrazine (20 min, 21°). Cleavage at only cytosine was achieved by performing the reaction in 2 M NaCl (25 min, 21°). In both cytosine and thymine reactions, phosphates were eliminated by treatment with 1 M piperidine (60 min, 90°). Aliquots of the products from the four reactions were loaded on a 20% acrylamide, 7 M urea gel at times 0, 12, and 24 h.

Hybridization – Restricted 2.5-kb DNA was fractionated on a 2% agarose gel. The separated fragments were transferred to a nitrocellulose filter (14) and hybridized with ^{125}I-labeled 5 S rRNA as described in Bell *et al.* (3).

RESULTS

Restriction Endonuclease Cleavage Map of Yeast EcoRI DNA Fragment Which Contains 5 S rRNA Gene – As discussed in the accompanying paper (3), digestion of yeast DNA with *Eco*RI produces a number of discrete fragments. One of these fragments with a molecular weight of 1.6 × 10^6 (approximately 2500 base pairs) hybridizes only with 5 S rRNA. This rDNA fragment was cloned in an *Escherichia coli* plasmid and a restriction endonuclease cleavage map was constructed as a prerequisite to localizing the 121-bp 5 S rRNA gene. This map, deduced from the experimental evidence described below, is presented in Fig. 1.

The yeast fragment was separated from the bacterial vector DNA and incubated with a number of restriction endonucleases. The results of these digestions are shown in Fig. 2 and Table I. The 2.5-kb fragment was cleaved by *Alu* I, *Hae* III, *Hha* I, *Hin*dII, *Hpa* I, *Hin*dIII, *Hpa* II, *Taq* I (not shown in Fig. 2), and *Xma* I but was not digested by *Sal* I, *Pst* I, *Bgl* I + II, *Dpn* II, *Bam* HI, *Kpn* I, and *Xba* I. As expected from previous results (3), *Hin*dIII cleaves this fragment only once, 166 bp from one end of the *Eco*RI fragment, and *Xma* I cuts it 235 bp from the other end. *Hpa* I, *Hha* I, *Hae* III, *Alu* I, *Taq* I, and *Hpa* II cleave the 2.5-kb fragment at 1, 1, 2, 5, 3, and 5 sites, respectively. The sizes of the DNA fragments produced by these enzymes are summarized in Table I. The order of the various restriction fragments was determined by an analysis of partial digestion products of the 2.5-kb fragment labeled at the end with ^{32}P (16). The DNA was labeled at both 5′ ends with [γ-^{32}P]ATP and polynucleotide kinase and then digested with *Hin*dIII. Since the smaller ^{32}P-labeled *Hin*dIII fragment could be readily identified, it was not separated from the larger fragment before the following analysis. The ^{32}P-labeled, *Hin*dIII-digested 2.5-kb fragment was incubated with *Alu* I, *Hae* III, and *Hpa* II for various times. The ^{32}P-labeled partial digestion products obtained are shown in Fig. 3. There are four *Alu* I sites within the larger, 2.4-kb, *Hin*dIII fragment at approximately 0.19, 1.41, 1.57, and 2.25 kb from the ^{32}P-labeled *Eco*RI site. The fifth *Alu* I site is within the *Hin*dIII site. There are two *Hae* III sites at 0.88 and 1.40 kb, and three *Hpa* II sites at 0.24, 1.01, and 1.72 kb from the *Eco*RI site. The smaller, 166-bp, *Hin*dIII fragment was not digested by *Alu* I or *Hae* III but was cleaved by *Hpa* II. The position of this site was subsequently determined by nucleotide sequence to be 12 bases from the *Eco*RI site (*Hin*dIII side). An additional *Hpa* II site, 2.32 kb from the *Eco*RI site (*Xma* I side) of the 2.4-kb fragment, was also detected. This site was not detected in the analysis of the partial digests because the 2.4-kb and 2.32-kb fragments were not resolved by the electrophoretic analysis employed. The position of the cleavage sites and the order of the restriction fragments was inferred from the molecular weight of the partial digestion products except

TABLE I

Molecular weight and length of restriction endonuclease digestion fragments from the EcoRI-generated yeast DNA segment containing the 5 S rRNA gene

The length of the fragments in base pairs is indicated between parentheses. Values are approximate except those obtained directly from sequence data which are italic. The fragments are arranged in the order (from left to right) in which they occur in Fig. 1.

Enzyme	Fragments
Undigested	1,600,000 (2,500)
*Hin*dIII	104,000 (*166*); 1,500,000 (2,400)
Alu I	105,000 (*168*); 71,000 (*114*); 430,000 (680); 103,000 (*165*); 770,000 (1,200); 110,000 (*174*)
Taq I	600,000 (950); 258,000 (*412*); 750,000 (1,200); 13,000 (*21*)
Hae III	710,000 (1,100); 328,000 (*524*); 560,000 (880)
Hpa II	7,500 (*12*); 113,000 (*181*); 380,000 (600); 444,000 (*710*); 490,000 (770); 148,000 (*236*)
Xma I	1,450,000 (2,300); 147,000 (*235*)
Hha I	1,350,000 (2,160); 215,000 (*344*)
*Hin*dII/*Hpa* I	1,400,000 (2,200); 186,000 (*298*)

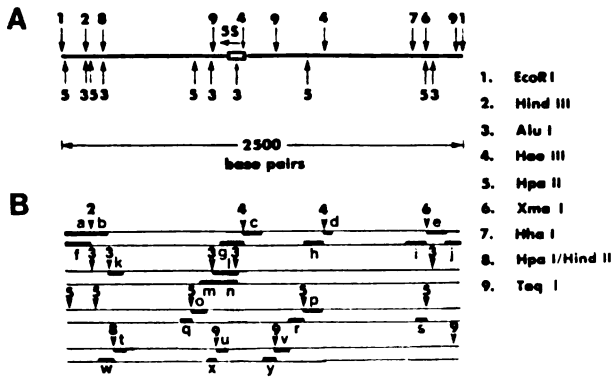


FIG. 1. Restriction endonuclease cleavage map of the yeast 2.5-kb DNA fragment containing the 5 S rRNA gene. *A*, cleavage map; *B*, scheme indicating the positions of the sequenced regions.[1] *a*, fragment *Eco*RI-2.50 (a DNA fragment is named by the enzyme(s) used to produce it and its size in kilobase pairs) was labeled with ^{32}P at both ends and then digested with *Hin*dIII. The *Eco*RI-*Hin*dIII-0.16 fragment was isolated and sequences of 78 and 80 bp (two experiments) were determined from the *Eco*RI end. Alternatively, fragment *Hin*dIII-0.16 was labeled at both ends, strand-separated, and a sequence of 101 bp from the *Eco*RI end was determined. *b*, fragment *Hin*dIII-2.40 (Table I) was labeled with ^{32}P at both ends and digested with *Hae* III. The *Hin*dIII-*Hae* III-0.94 fragment was isolated and a sequence of 130 bp was determined from the *Hin*dIII site. *c*, fragment *Hae* III-0.52 (Table I) was labeled with ^{32}P at both ends and digested with *Hpa* II. The *Hae* III-*Hpa* II-0.38 fragment was isolated and sequences of 60 and 90 bp (two experiments) were determined from the *Hae* III site. Alternatively, fragment *Hae* III-0.52 was labeled with ^{32}P at both ends and digested with *Taq* I. The *Hae* III-*Taq* I-0.30 fragment was isolated and a sequence of 105 bp was determined from the *Hae* III site. *d*, fragment *Hae* III-0.88 (Table I) was labeled with ^{32}P at both ends and digested with *Hha* I. The *Hae* III-*Hha* I-0.53 fragment was isolated and a sequence of 23 bp was determined from the *Hae* III site. *e*, fragment *Xma* I-0.24 (Table I) was labeled with ^{32}P at both ends and then strand-separated. A sequence of 144 bp was determined from the *Xma* I site. *f*, fragment *Hin*dIII-0.16 was labeled with ^{32}P at both ends and then strand-separated. A sequence of 78 bp was determined from the *Hin*dIII end. *g*, fragment *Hae* III-1.10 (Table I) was labeled with ^{32}P at both ends and digested with *Hpa* II. The *Hae* III-*Hpa* II-0.30 fragment was isolated and a sequence of 116 bp was determined from the *Hae* III end. *h*, fragment *Hae* III-0.52 was labeled with ^{32}P at both ends and digested with *Hpa* II. The *Hae* III-*Hpa* II-0.14 fragment was isolated and sequences of 60 and 80 bp (two experiments) were determined from the *Hae* III end. *i*, fragment *Xma* I-2.30 was labeled with ^{32}P at both ends and digested with *Hae* III. The *Xma* I-*Hae* III-0.64 fragment was isolated and sequences of 90 and 112 bp were determined (two experiments) from the *Xma* I end. *j*, fragment *Eco*RI-2.50 was labeled with ^{32}P at both ends and digested with *Xma* I. The resulting *Eco*RI-*Xma* I-0.24 fragment was isolated and sequences of 95 and 100 bp (two experiments) were determined from the *Eco*RI end. Alternatively, fragment *Eco*RI-*Xma* I-0.24 was labeled with ^{32}P at both ends and strand-separated. A sequence of 135 bp was determined from the *Eco*RI site. *k*, fragment *Alu* I-0.68 (Table I) was labeled with ^{32}P at both ends and digested with *Hpa* II. After isolation of the *Alu* I-*Hpa* II-0.56 fragment, a sequence of 33 bp was determined from the *Alu* I end. *l*, fragment *Alu* I-0.16 was isolated from the *Hpa* II-0.71 fragment (Table I) and labeled with ^{32}P at both ends. After strand separation, a sequence of 105 bp was determined from one of the *Alu* I sites. *m*, fragment *Alu* I-0.68 was labeled with ^{32}P at both ends and digested with *Hpa* II. After isolation of the *Alu* I-*Hpa* II-0.11 fragment, a sequence of 74 bp was determined from the *Alu* I end. *n*, fragment *Alu* I-0.16 was isolated from the *Hpa* II-0.71 fragment and labeled with ^{32}P at both ends. After strand separation, a sequence of 97 bp was determined from one of the *Alu* I sites. *o*, fragment *Hpa* II-0.71 was labeled at both ends with ^{32}P and digested with *Alu* I. After isolation of the *Hpa* II-*Alu* I-0.11 fragment, a sequence of 49 bp was determined from the *Hpa* II ends. Alternatively, fragment *Hpa* II-*Alu* I-0.11 was labeled with ^{32}P at both ends and then strand-separated. A sequence of 92 bp was determined from the *Hpa* II site. *p*, fragment *Hae* III-0.52 was digested with *Hpa* II. The *Hae* III-*Hpa* II-0.14 fragment was isolated and labeled with ^{32}P at both ends. After strand separation, a sequence of 120 bp was determined from the *Hpa* II end. *q*, fragment *Hpa* II-0.60 (Table I) was labeled with ^{32}P at both ends and digested with *Hin*dII. After isolation of the *Hpa* II-*Hin*dII-0.49 fragment, the sequence of 58 bp was determined from the *Hpa* II end. *r*, fragment *Hpa* II-0.71 was labeled with ^{32}P at both ends and digested with *Hae* III. The *Hpa* II-*Hae* III-0.38 fragment was isolated and a sequence of 80 bp was determined from the *Hpa* II site. *s*, fragment *Hpa* II-0.77 was labeled with ^{32}P at both ends and digested with *Hae* III. After isolation of the *Hpa* II-*Hae* III-0.63 fragment, a sequence of 51 bp was determined from the *Hpa* II site. *t*, fragment *Hin*dII-*Eco*RI-2.20 (Table I) was labeled with ^{32}P at both ends and digested with *Hpa* II. The *Hin*dII-*Hpa* II-0.49 fragment was isolated and a sequence of 70 bp was determined from the *Hin*dII site. *u*, fragment *Taq* I-0.41 (Table I) was labeled with ^{32}P at both ends and digested with *Alu* I. After isolation of the *Taq* I-*Alu* I-0.14 fragment, sequences of 82 and 87 bp (two experiments) were determined from the *Taq* I site. *v*, fragment *Taq* I-1.20 (Table I) was labeled with ^{32}P at both ends and digested with *Hae* III. The *Taq* I-*Hae* III-0.30 fragment was isolated and sequences of 89 and 92 bp (two experiments) were determined from the *Taq* I site. *w*, fragment *Eco*RI-*Hin*dII-0.30 (Table I) was labeled with ^{32}P at both ends and digested with *Hpa* II. After isolation of the *Hin*dII-*Hpa* II-0.10 fragment, a sequence of 48 bp was determined from the *Hin*dII end. *x*, fragment *Taq* I-0.95 (Table I) was labeled with ^{32}P at both ends and digested with *Hpa* II. The *Taq* I-*Hpa* II-0.14 fragment was isolated and sequences of 77 and 103 bp (two experiments) were determined from the *Taq* I end. *y*, fragment *Taq* I-0.41 was labeled with ^{32}P at both ends and digested with *Hae* III. After isolation of the *Taq* I-*Hae* III-0.22 fragment, a sequence of 130 bp was determined from the *Taq* I site.

[1] The primary data (films of sequence gels) will be available to interested scientists for a period of 3 years after publication of this paper.

for the *Taq* I fragments which were ordered as indicated by Maxam *et al.* (10) and later confirmed by our nucleotide sequence.

Secondary restriction endonuclease digestions were employed to confirm the arrangement of the *Alu* I, *Hae* III, and *Hpa* II fragments obtained from partial digestion experiments and also to correlate these data with those from total digestion. Fragments obtained by total digestion of the 2.5-kb DNA fragment with *Alu* I, *Hae* III, and *Hpa* II were isolated by preparative gel electrophoresis and subjected separately to the action of a second restriction endonuclease. The results of this analysis are shown in Fig. 4. The number and the molecular weight of the DNA fragments obtained on secondary digestion were in complete agreement with the physical order inferred from partial digestions, except in one instance. Secondary digestion of *Hae* III-1.1 (a fragment is named by the enzyme that is used to produce it and its size in kilobase pairs) by *Hpa* II generated three fragments of approximately 0.60, 0.32, and 0.18 kb. This revealed the additional *Hpa* II cleavage site near to the *Hin*dIII site which was not detected by the partial digestion experiments discussed above.

Localization of 5 S rRNA Cistron within the 2.5-kb Fragment—The sequence of *S. cerevisiae* 5 S rRNA (6) predicts the presence of an *Hae* III site near the beginning (3′ end of the coding strand) of the 5 S rRNA gene and an *Alu* I site 49 bp from the *Hae* III site towards the middle of the gene. Inspection of the restriction map of Fig. 1 shows that an *Alu* I and *Hae* III site are close together only once, near the center of the 2.5-kb fragment. The relative positions of these two sites predict that the 5 S rRNA gene is transcribed from right to left of Fig. 1. The presence of the 5 S rRNA gene within

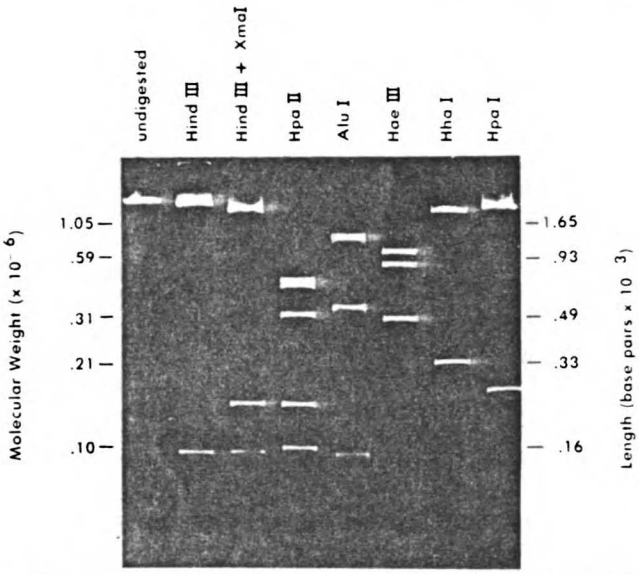


FIG. 2. Acrylamide gel electrophoresis of total restriction endonuclease digests of the yeast 2.5-kb DNA fragment containing the 5 S rRNA gene. Electrophoresis is in a 5% acrylamide slab gel as described under "Experimental Procedures." The approximate size of the DNA fragments was obtained by comparison with an *Eco*RI + *Hin*dIII digest of phage λ-DNA (15).

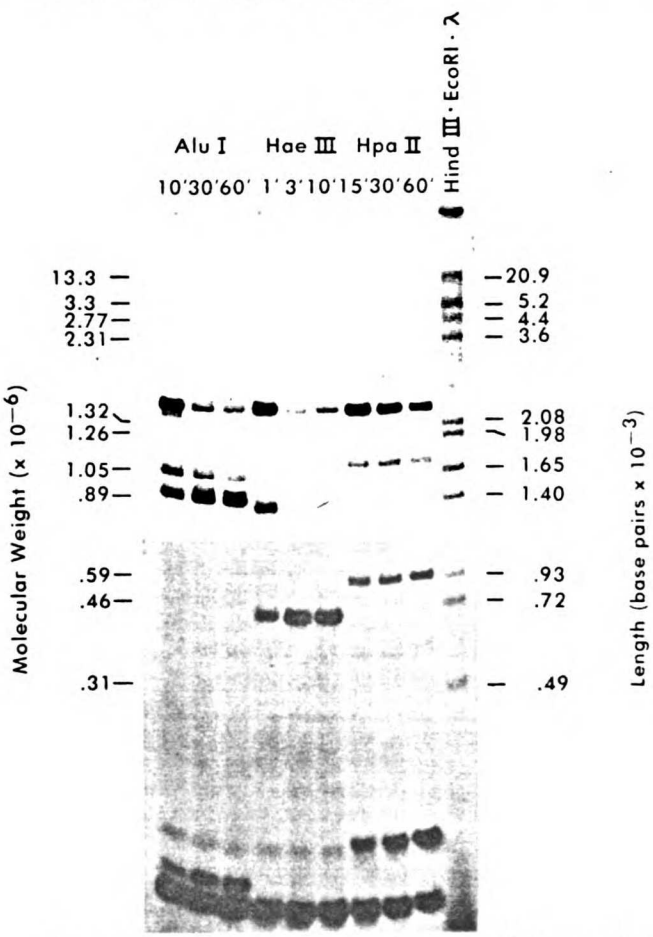


FIG. 3. Autoradiograph of partial endonuclease cleavage products of terminally labeled yeast 2.5-kb fragment. Electrophoresis is in 2% agarose. The size of the DNA fragments was determined as indicated in Fig. 2.

this region was confirmed by hybridization of 5 S rRNA to *Hae* III and *Alu* I restriction fragments. These fragments, separated by agarose gel electrophoresis, were transferred to a nitrocellulose filter as described by Southern (14) and hybridized with ^{125}I-labeled yeast 5 S rRNA. The fragments containing sequences homologous to 5 S rRNA were visualized by autoradiography, as shown in Fig. 5. The *Hae* III-1.1, *Alu* I-1.2, and *Alu* I-0.16 fragments hybridized to 5 S rRNA. The difference in hybridization detected between the *Alu* I-1.2 and *Alu* I-0.16 fragments may reflect less efficient absorption of the smaller fragment during transfer of the DNA from the agarose gel to the nitrocellulose filter as well as preferential loss of the small fragment from the filter during the hybridization. No hybridization to the *Hae* III-0.52 fragment was detected. This is explained by the inability of the 8 bp of the 5

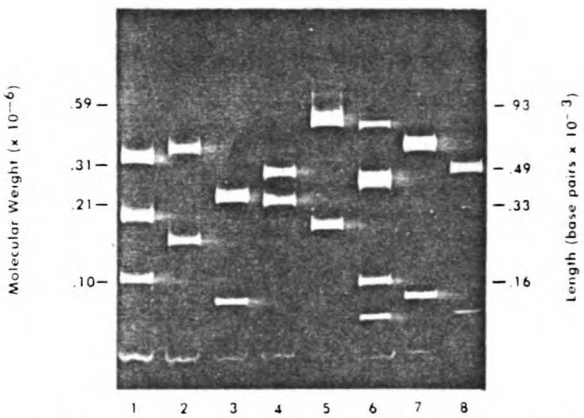


FIG. 4. Acrylamide gel electrophoresis of secondary restriction endonuclease digests of restriction fragments derived from the yeast 2.5-kb DNA. *1*, *2*, and *3*, digestion of *Hae* III-1.1, *Hae* III-0.88, and *Hae* III-0.52 fragments, respectively, with *Hpa* II; *4*, digestion of *Hae* III-0.52 with *Hha* I; *5*, digestion of *Hae* III-1.1 with *Hpa* I; *6*, digestion of *Hpa* II-0.71 with *Alu* I; *7*, digestion of *Hpa* II-0.77 with *Hae* III; *8*, digestion of *Alu* I-0.68 with *Hpa* II. The gel was 5% acrylamide. The size of the DNA fragments was determined as indicated in Fig. 2.

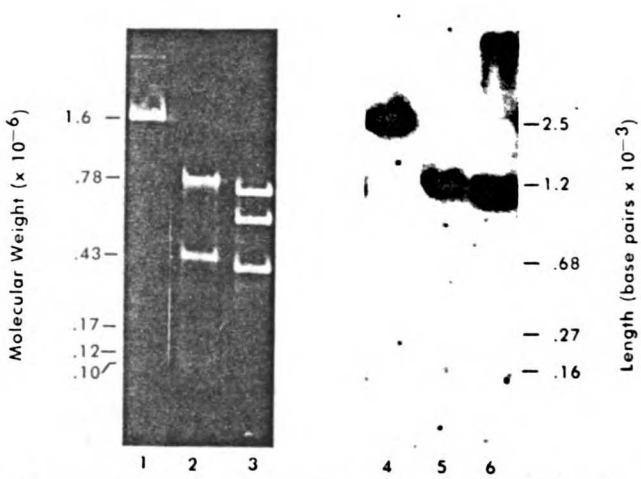


FIG. 5. Hybridization of restriction fragments derived from the 2.5-kb DNA with I^{125}-labeled yeast 5 S rRNA. *1*, *2*, and *3*, undigested, *Alu* I and *Hae* III DNA fragments, respectively, visualized by staining with ethidium bromide; *4*, *5*, and *6*, autoradiograph of the same DNA bands after hybridization with yeast ^{125}I-labeled 5 S rRNA. Electrophoresis is in 2% agarose.

S rRNA gene present in this fragment to form a stable hybrid (17).

Nucleotide Sequence of the 5 S rRNA Gene and Flanking Regions—The primary structure of most of the 5 S rRNA gene was obtained from the *Hae* III-1.1 fragment (Fig. 1, Table I). This fragment was purified by preparative gel electrophoresis, labeled at the 5′ ends with [γ-^{32}P]ATP and polynucleotide kinase, and digested with *Hpa* II. A fragment of approximately 300 bp, labeled with ^{32}P at the *Hae* III site, was isolated and sequenced by the method of Maxam and Gilbert (13). The results are shown in Fig. 6*a*.

The nucleotide sequence of the first 8 bp at the beginning of the gene and the immediately adjacent spacer region which should contain the promoter, were obtained from the *Hae* III-0.52 fragment (Fig. 1, Table I). This fragment was labeled with ^{32}P at the 5′ ends and cleaved by *Hpa* II into two approximately 140- and 380-bp segments, each labeled at their *Hae* III site. Both fragments were isolated and sequenced. The sequence gel of the 380-bp fragment which contains part of the 5 S rRNA gene and the promoter region is shown in Fig. 6*a*. Extension of the sequence of the promoter region was obtained by digesting the fragment *Hpa* II-0.71 with *Taq* I, labeling with ^{32}P, digesting with *Hae* III, and sequencing the resulting ^{32}P-labeled fragments from their *Taq* I site (Fig. 1*B*).

The primary structure of the region after the 5 S rRNA gene was obtained by sequencing the *Alu* I-0.16 fragment from this region (Fig. 1, Table I). Since there are two *Alu* I segments of this molecular weight, we first isolated *Hpa* II-0.71 containing the entire 5 S rRNA gene and treated it with *Alu* I. As expected, three fragments of approximately 0.11, 0.16, and 0.44 kb were obtained. The *Alu* I-0.16 fragment was isolated and labeled with ^{32}P. The strands were separated by gel electrophoresis and sequenced. Fig. 7 shows the gel from

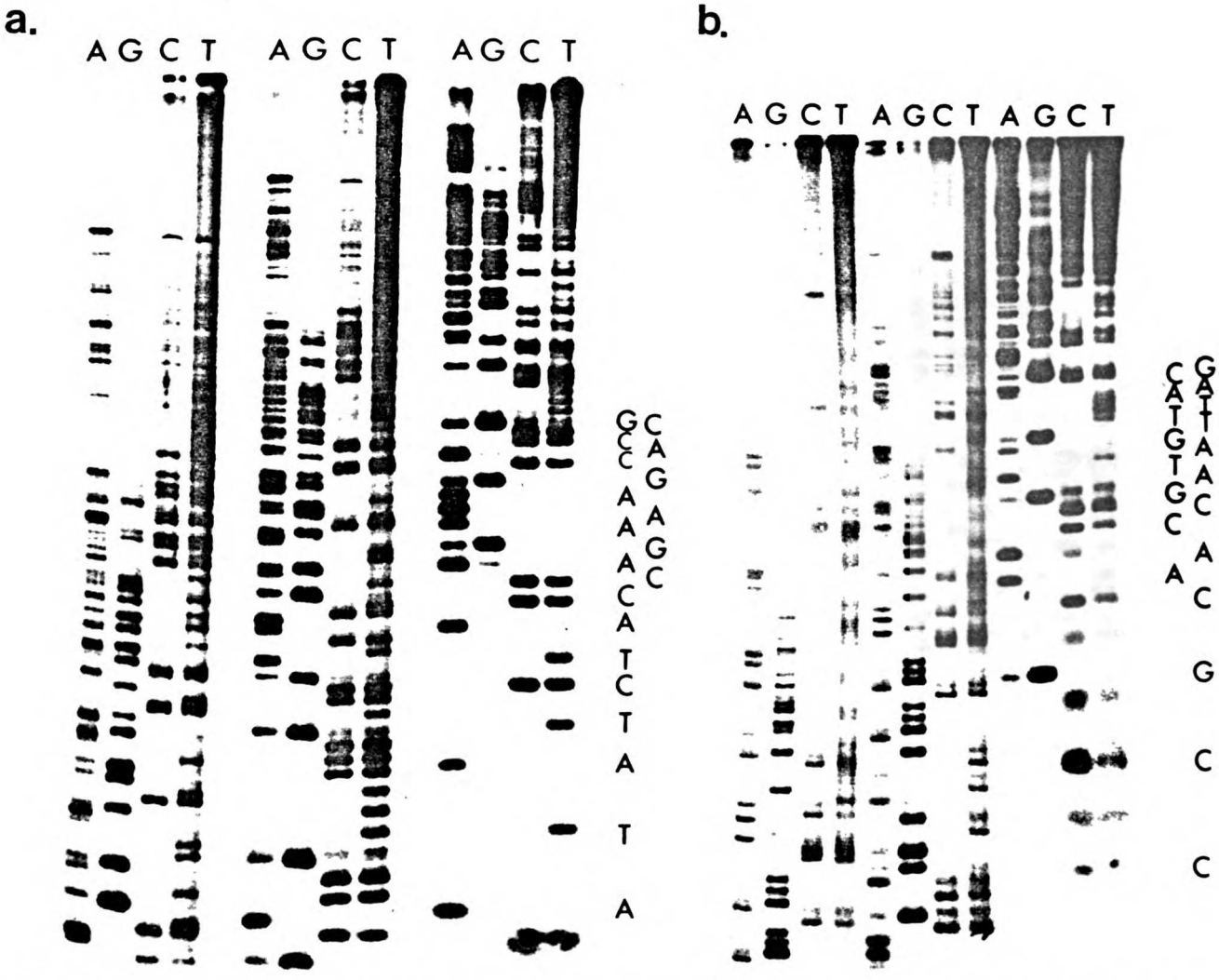


FIG. 6. Autoradiogram of the electrophoretic separation of ^{32}P-labeled oligonucleotides obtained by base-specific cleavage of fragments containing the 5 S rRNA gene. *a*, partial sequence of the 300-bp fragment obtained by *Hpa* II digestion of the fragment *Hae* III-1.1 (see Fig. 1, Table I); *b*, partial sequence of the 400-bp fragment obtained by *Hpa* II digestion of the fragment *Hae* III-0.52 (Fig. 1, Table I). The columns *A*, *G*, *C*, and *T* indicate wells loaded with products from reactions partially specific for adenine, guanine, cytosine, and thymine, respectively. The reading of the first 10 bases is illustrated at the *right* of each autoradiogram. The first two oligonucleotides of the labeled 5′ end of fragment *b* ran off the gel and were determined from the specificity of *Hae* III (18).

which the sequence of the termination region was derived.

Extension of the sequence in this region was obtained by isolating the fragments *Hpa* II-0.60 and *Hpa* II-0.71 (Fig. 1, Table I), labeling with ^{32}P at both ends, and digesting with *Alu* I. Sequence around the *Hpa* II site separating the 0.60- and 0.71-kb *Hpa* II fragments (gels not shown) defined the primary structure of a 273-bp segment at the end of the 5 S rRNA gene.

By similar strategy, we have sequenced (gels not shown) the region around the *Hae* III site separating the 0.52- and 0.88-kb *Hae* III fragments and the region around the *Hpa* II site separating the 0.71- and 0.77-kb *Hpa* II fragments (Fig. 1*B*)

The complete nucleotide sequence of the 5 S rRNA gene and adjacent regions is shown in Fig. 8.

Nucleotide Sequence of Regions at Ends of 2.5-kb DNA Fragment—The ends of the 2.5-kb fragment may contain portions of the 35 S rRNA gene and its control regions. To sequence these regions, the 2.5-kb fragment was successively digested with *Xma* I and *Hin*dIII. Three fragments of 166, 235, and 2100 bp were obtained. They were labeled at their 5′ ends with ^{32}P and digested with *Hae* III. As expected from the restriction map, Fig. 1, this enzyme cleaves the 2100-bp fragment at two sites, generating a 940-bp fragment (*Hin*dIII-*Hae* III-0.94) labeled with ^{32}P only at the *Hin*dIII site, an unlabeled 520-bp fragment (*Hae* III-0.52) and a 640-bp fragment (*Xma* I-*Hae* III-0.64) labeled with ^{32}P at the *Xma* I site. The four labeled fragments were then isolated. The *Hin*dIII-0.16 and *Xma* I-0.24 fragments were sequenced after strand separation. The sequence gels of the *Hin*dIII-0.16 strands are shown in Fig. 9. The isolated *Hin*dIII-*Hae* III-0.94 and *Xma* I-*Hae* III-0.64 fragments were also sequenced from their labeled *Hin*dIII and *Xma* I ends, respectively (gels not shown).

To sequence the region at the right of the *Hpa* I-*Hin*dII site (see Fig. 1), the *Alu* I-0.68 fragment was isolated, labeled with ^{32}P, and digested with *Hpa* II. The larger fragment (approximately 0.56 kb) was isolated and sequenced from the *Alu* I site (gel not shown). Fig. 10 summarizes the available sequence data for both ends of the 2.5-kb fragment.

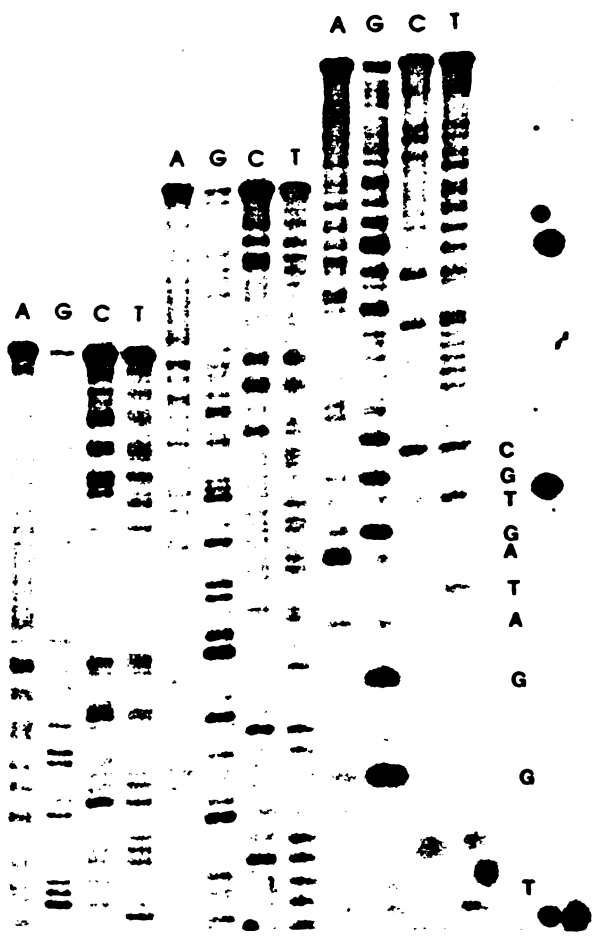


FIG. 7. Autoradiogram of the electrophoretic separation of ^{32}P-labeled oligonucleotides obtained by base-specific cleavages of a fragment containing the region adjacent to the 5′ end (coding strand) of the 5 S rRNA gene. This DNA corresponds to the electrophoretically faster moving strand from a 167-bp fragment obtained by digestion of *Hpa* II-0.71 with *Alu* I. Other conditions are the same as indicated under Fig. 6.

DISCUSSION

In bacteria, gene expression is regulated in part by the primary structure of regions adjacent to the gene (reviewed by Gilbert, Ref. 19). From an examination of the sequence of a number of prokaryotic promoters, it has been possible to deduce some general features of their structure and to identify RNA polymerase and regulatory protein binding sites. Regulatory proteins appear to bind to DNA regions which possess complete or partial 2-fold symmetry. In contrast, RNA polymerase binding sites have been more difficult to identify but they do not appear to involve palindromic sequences.

In eukaryotes, the identification of the structures involved in regulation of transcription has not been achieved. Recent advances in techniques for cloning eukaryotic DNA fragments in bacterial plasmids have facilitated the isolation of genetic segments whose structure and function can be studied in detail. We have used these techniques for studying the rRNA genes of *Saccharomyces cerevisiae*.

The 5 S rRNA is apparently an unprocessed transcript (6, 7); therefore, the control sequences for transcription initiation and termination should be adjacent to the 3′ and 5′ ends of the gene itself. By restriction endonuclease mapping and hybridization, we have located the 5 S rRNA gene near the center of a 2.5-kb *Eco*RI fragment. The nucleotide sequence of the 5 S rRNA gene reported here establishes its polarity and confirms that the 5 S gene is transcribed in opposite direction to the 35 S rRNA precursor gene (20).

The sequence of most of the 5 S RNA from *S. cerevisiae* has been reported by Miyazaki (6). However, a small region around the 35th position (from the 5′ end) remained uncertain. Our results establish the sequence of this region and confirm the remainder of Miyazaki's results.

The region before the 5 S rRNA gene must contain the promoter. Pribnow (21) has observed that several prokaryotic promoters have a related AT-rich heptanucleotide sequence centered 9 to 10 bases before the transcription initiation site. In the *lac* operon, mutations in this region that convert GC to AT pairs result in increased promoter activity (19). We detect an AT-rich sequence GTTTAAC similar to the "Pribnow heptanucleotide" centered about 8 bases prior to the 5 S gene. This AT-rich region, located about one double helical turn before the start of transcription, may facilitate local denaturation by RNA polymerase. The yeast promoter region does not possess sequences homologous with TTGA and GTTG found in prokaryotes in a region about 35 bases from the start of the gene or palindromes that have been identified as binding sites for regulatory proteins in some prokaryotic

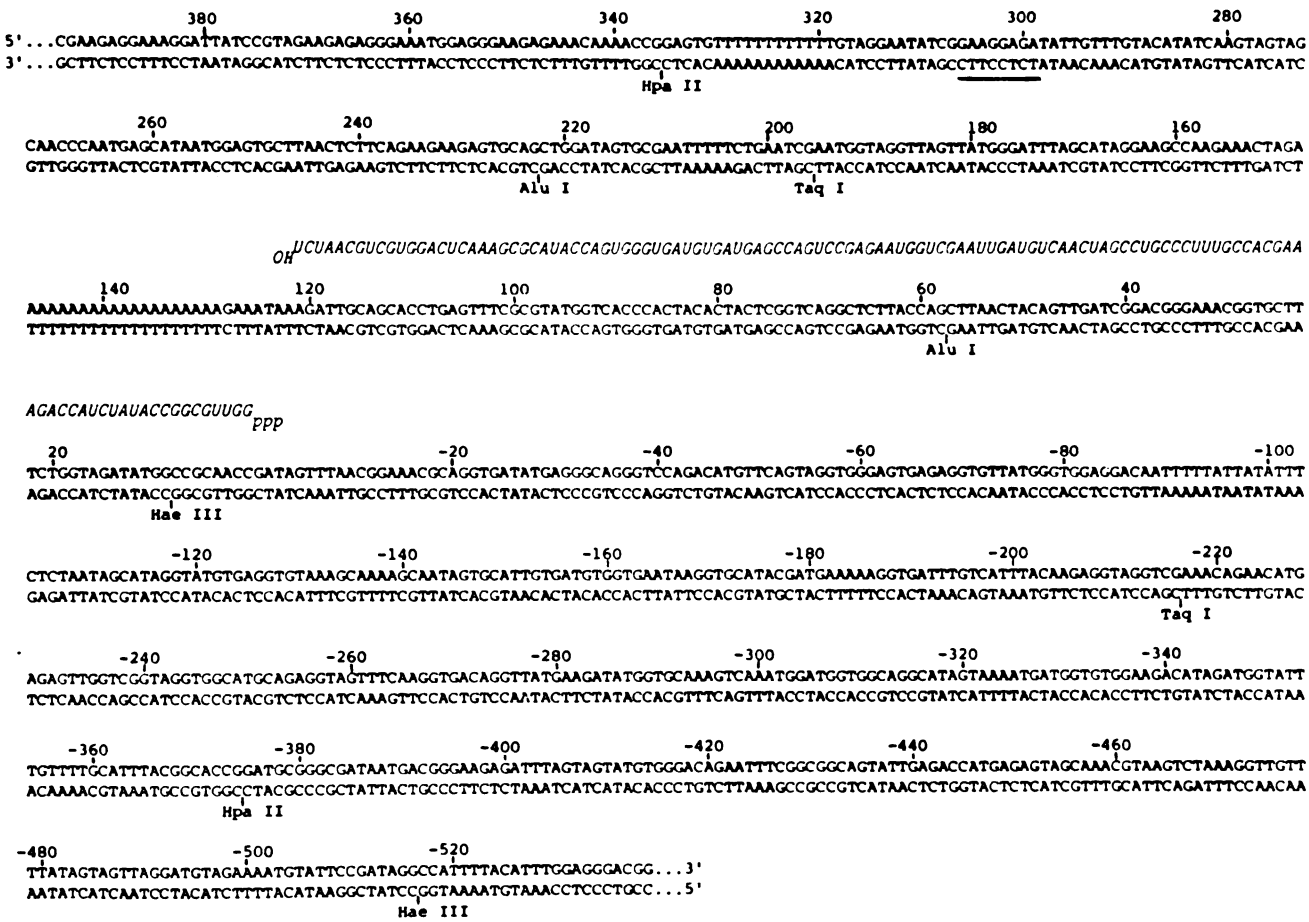


FIG. 8. Nucleotide sequence of *Saccharomyces cerevisiae* 5 S ribosomal RNA gene and adjacent regions. The sequence of the corresponding 5 S rRNA is shown in *italics*. The *underlining* indicates a region where the sequence is not fully confirmed.

promoters (21). There is no obvious relationship between the primary structures of the yeast 5 S rRNA promoter and the promoter region of *Xenopus laevis* oocyte 5 S rRNA gene.[2] The sequence relationships which determine transcription promotion in eukaryotes are currently difficult to perceive.

There is a long run of adenosines in the coding strand at the end of the 5 S rRNA gene. This unique structure could represent a termination signal or a processing sequence. We believe it is the former since no precursors of 5 S RNA have been detected in any system, and *in vitro* transcription experiments carried out in this laboratory suggest that RNA polymerase III terminates at poly(A) regions.[3] Termination regions in other systems appear to contain adenosine-rich sequences. For example, short runs of uridines have been found at the 3′ end of 4 S RNA of λ phage (22), a small RNA from phage Φ80 (23) and the attenuator region of tryptophan mRNA of *Escherichia coli* (24). In *Xenopus*, Denis and Wegnez (25) have found that a small fraction of newly formed 5 S RNA molecules are slightly longer and contain extra nucleotides at the 3′ end. These ends are uridine-rich of the type CUU, CUUU, and CUUUU. Brown and Brown (26) have demonstrated a short AT-rich region immediately after the *Xenopus laevis* 5 S rRNA gene and postulated the sequence $\frac{\text{TTTT}}{\text{AAAA}}$ as part of the termination signal.

[2] N. Fedoroff and D. D. Brown, personal communication.

[3] L. J. DeGennaro, unpublished results.

In most of the examples cited above, part of an adenosine-rich sequence is transcribed giving rise to RNAs with uridine-rich 3′ termini. In yeast, the majority of the 5 S RNA transcripts end with a single uridine. Whether a small percentage of these molecules have additional uridines is not known. These adenosine-rich regions could facilitate the detachment of the enzyme from the DNA or serve as a distinctive recognition site for other molecules which influence termination.

In contrast to the 5 S rRNA gene, it is more difficult to identify the control regions for the 35 S rRNA precursor gene. This RNA, like tRNA and several eukaryotic mRNAs, is derived from a large precursor. A comparison of the sizes of the rDNA repeating unit and the 35 S rRNA suggest that control regions of this gene may be within the 2.5-kb fragment (3). Partial nucleotide sequence of the regions between the 5 S and 35 S rRNA genes are shown in Fig. 10. At the *Hin*dIII end of the DNA, the nucleotide sequence of 355 bp has been determined. At a distance of 96 bp from the *Eco*RI cleavage site, there is a long run of 31 AT pairs interrupted once by a GC pair. This region is strikingly similar to that found at the termination of the 5 S rRNA gene, except that it is in the opposite DNA strand. The homology includes a sequence of 18 bases that is exactly the same in both regions: $\frac{\text{AAAGAAAAAA}}{\text{TTTCTTTTTT}}$ $\frac{\text{AAAAAAAA}}{\text{TTTTTTTT}}$. This justifies the hypothesis that the adenosine-rich segment near the *Hin*dIII site is part of the termination

region of the 35 S rRNA gene. Determination of the nucleotide sequence at the 3′ end of the 35 S rRNA will be necessary to test this.

At the *Xma* I end, we have determined the sequence of 349 bp as a preliminary step in our attempts to locate the 35 S rRNA promoter. The exact determination of this site must await determination of the nucleotide sequence at the 5′ end of the precursor RNA. A striking feature in this region of the DNA is a peculiar 60-bp sequence starting 6 bases at the right of the *Hha* I site. The sequence CGTATTTTCCGCT GCATAAAAGGCGA TC AG is repeated sequentially four times. Repeated units similar to this one have been described in the intergenic region between the 5 S genes in *X. laevis* oocytes (27). The biological significance of this interesting structure is yet to be determined.

Cramer *et al.* (4) have examined the topology of the DNA sequences of the yeast rDNA repeating unit by partial denaturation mapping. They observed that more highly denatured AT-rich segments occurred in the intercistronic regions. This includes the transcribed spacer segments of the 35 S rRNA precursor and the nontranscribed spacer between the 5 S and 35 S rRNA genes. The most highly denatured regions observed under their conditions occurred in the 2.5-kb *Eco*RI fragment. In particular, two regions near the *Xma* I and *Hin*dIII sites were more extensively denatured than elsewhere in the rDNA repeating unit. Our DNA sequence of these regions demonstrates their high AT character. In addition, two less prominent AT regions are located between the *Xma* I and *Hin*dIII sites (4). The DNA sequence suggests that these probably occur on both sides of the 5 S rRNA coding sequences. One region corresponds to the AT-rich sequence immediately after the 5 S rRNA gene and the other to an AT-rich region approximately 100 bp from the beginning of the 5 S rRNA gene. The present data and that of Maxam *et al.* (10) suggest that these AT-rich regions might be involved in identifying the 5 S rRNA gene. We propose that the AT-rich regions centered at the *Xma* I and *Hin*dIII sites in the 2.5-kb *Eco*RI fragment identify the limits of the 35 S rRNA gene. Wensink and Brown (28) have observed more easily denatured regions at the ends of *Xenopus laevis* 40 S rRNA precursor coding regions. The tandemly repeated 5 S rRNA genes of *X. laevis* are also separated by AT-rich spacers, a short region after the gene and a much longer region before the 5 S rRNA gene (26). The presence of AT-rich regions adjacent to the rRNA genes could well be a common feature of these genes. Whether such regions are a characteristic of all genes remains to be determined.

The AT regions found at the beginning of the gene differ from those at the terminal regions. At the beginning, most of the adenosines are in the noncoding strand, or are mixed apparently randomly with thymine residues, whereas the adenosines after the gene are present in the coding strand. The similarity in sequence and opposite orientation of the adenosine-rich strand at the *Hin*dIII site and after the 5 S rRNA gene are consistent with the suggested role for these sequences in the termination of transcription. There are two other AT-rich regions within the rDNA repeating unit that

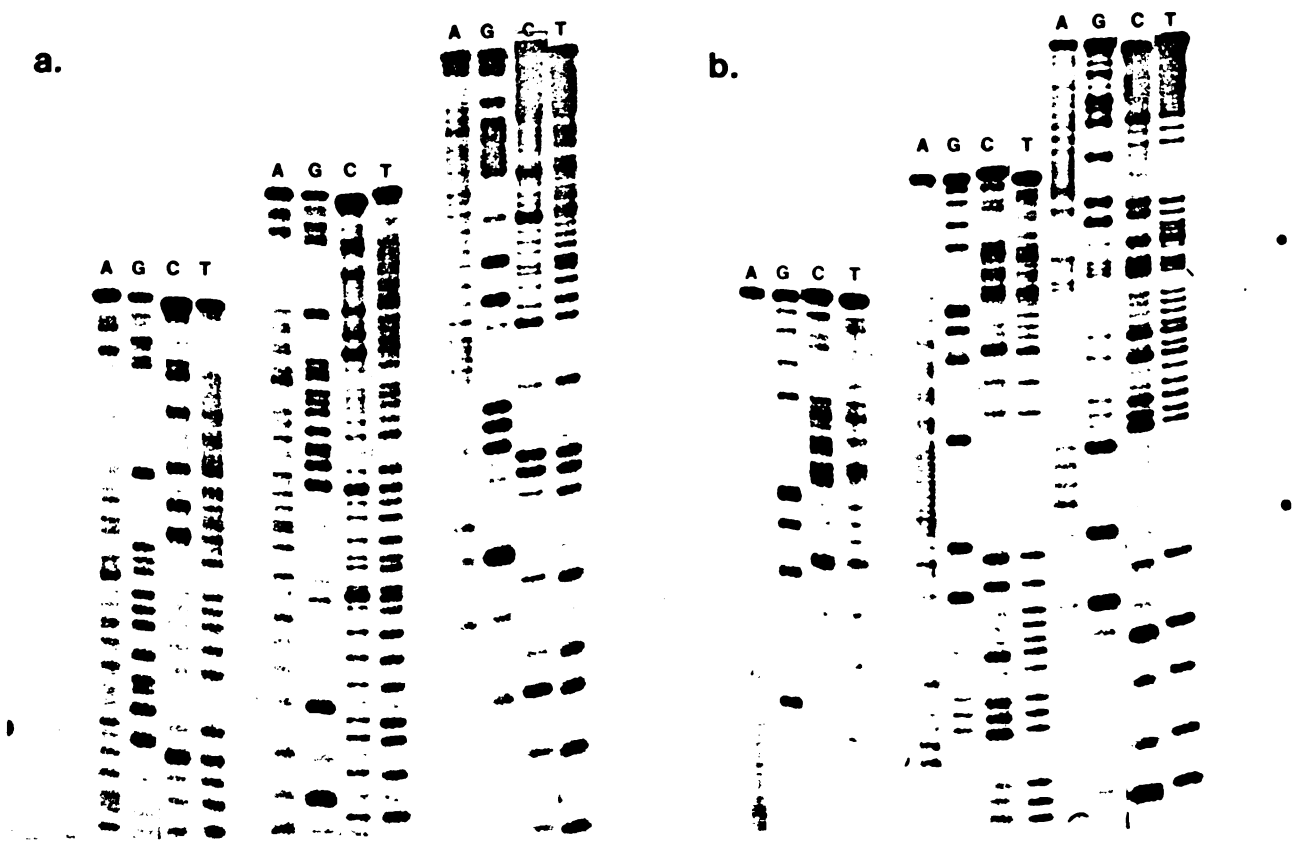


FIG. 9. Autoradiogram of sequence gels from the 165-bp DNA fragment located between the *Eco*RI and *Hin*dIII sites. *a*, electrophoretically slower moving strand; *b*, electrophoretically faster moving strand. Other conditions are as indicated under Fig. 6.

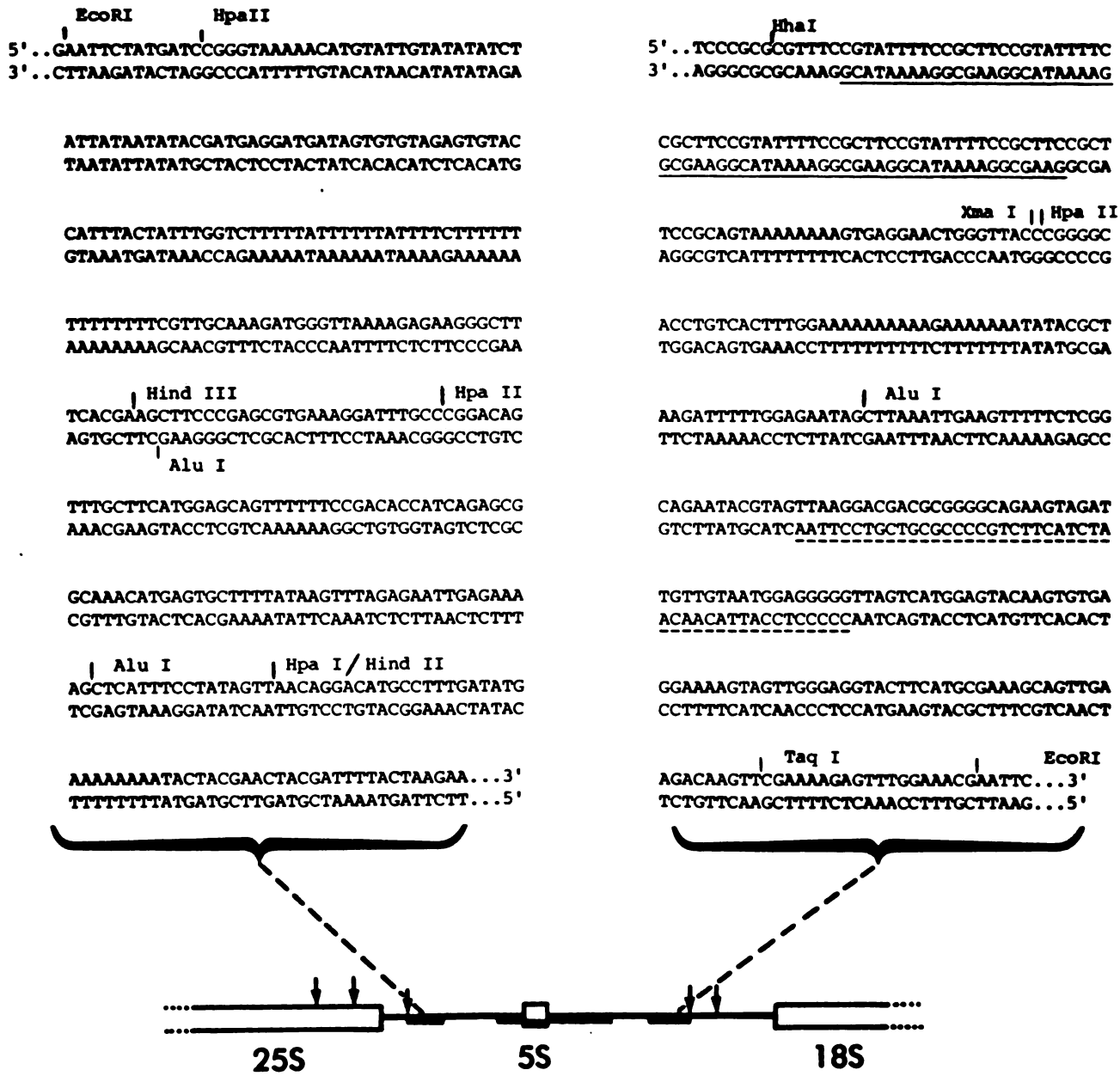


FIG. 10. Summary of partial nucleotide sequence of the regions between the 5 S and 35 S rRNA genes of *Saccharomyces cerevisiae*. The *solid underlining* indicates a 60-bp region comprised of a 15-bp sequence repeated four times. The *dashed underlining* indicates a region where the sequence is not fully confirmed. The *arrows* in the lower portion of the figure indicate the position of *Eco*RI sites in the vicinity of the 2.5-kb fragment.

have been observed by denaturation mapping (4) and sequence analysis.[4] One is before the 18 S rRNA gene and the other between the 18 S and 5.8 S rRNA genes. These AT-rich regions may play a role in RNA processing.

Our sequence studies have rendered the 2.5-kb *Eco*RI fragment an appropriate template for studying transcription. By analogy with vertebrates (29), yeast RNA polymerases I and III probably transcribe the 35 S rRNA and 5 S rRNA genes, respectively. Although these enzymes have been extensively purified and characterized (30-32), a study of their role in transcription has been limited due to the lack of a suitable assay. Using this fragment as a template, initiation and termination of 5 S rRNA and 35 S rRNA synthesis can now be directly tested by hybridization and RNA sequencing. Thus, the molecular requirements for transcriptive specificity can be directly monitored.

[4] G. I. Bell and P. Valenzuela, unpublished results.

Acknowledgments – We thank A. Maxam and W. Gilbert for providing a detailed DNA sequence protocol prior to publication.

REFERENCES

1. Schweizer, E., MacKechnie, C., and Halvorson, H. O. (1969) *J. Mol. Biol.* 40, 261-277
2. Rubin, G. M., and Sulston, J. E. (1973) *J. Mol. Biol.* 79, 521-530

3. Bell, G. I., DeGennaro, L. J., Gelfand, D. H., Bishop, R. J., Valenzuela, P., and Rutter, W. J. (1977) *J. Biol. Chem.* **252**, 8118-8125
4. Cramer, J. H., Farrelly, F. W., Barnitz, J. R., and Rownd, R. H. (1977) in *Molecular Approaches to Eukaryotic Genetic Systems* (Wilcox, G., and Abelson, J., eds) Academic Press, New York, in press
5. Kramer, R. A., Cameron, J. R., and Davis, R. W. (1976) *Cell* **8**, 227-232
6. Miyazaki, M. (1974) *J. Biochem.* **75**, 1407-1410
7. Trapman, J., and Planta, R. J. (1975) *Biochim. Biophys. Acta* **414**, 115-125
8. Udem, S. A., and Warner, J. R. (1972) *J. Mol. Biol.* **65**, 227-242
9. Valenzuela, P., Bell, G. I., Masiarz, F. R., DeGennaro, L. J., and Rutter, W. J. (1977) *Nature* **267**, 641-643
10. Maxam, A. M., Tizard, R., Skryabin, K. G., and Gilbert, W. (1977) *Nature* **267**, 643-645
11. Sato, S., Hutchinson, C. A., III, and Harris, J. I. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 542-546
12. Meagher, R. B., Tait, R. C., Betlach, M., and Boyer, H. W. (1977) *Cell* **10**, 521-536
13. Maxam, A. M., and Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 560-564
14. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503-517
15. Murray, K., and Murray, N. E. (1975) *J. Mol. Biol.* **98**, 551-564
16. Smith, H. O., and Birnstiel, M. L. (1976) *Nucleic Acids Res.* **3**, 2387-2398
17. McCarthy, B. J., and Church, R. B. (1970) *Annu. Rev. Biochem.* **39**, 131-150
18. Roberts, R. G. (1976) *CRC Crit. Rev. Biochem.* **4**, Suppl. 2, 123-163
19. Gilbert, W. (1976) in *RNA Polymerase* (Losick, R., and Chamberlin, M., eds) pp. 193-205, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
20. Aarstad, K., and Oyen, T. B. (1975) *FEBS Lett.* **51**, 227-231
21. Pribnow, D. (1975) *J. Mol. Biol.* **99**, 419-443
22. Sklar, J., Yot, P., and Weissman, S. M. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 1817-1821
23. Pieczenik, G., Barrell, B. G., and Gefter, M. L. (1972) *Arch. Biochem. Biophys.* **152**, 152-165
24. Bertrand, K., Korn, L., Lee, F., Platt, T., Squires, C. L., Squires, C., and Yanofsky, C. (1975) *Science* **189**, 22-26
25. Denis, H., and Wegnez, M. (1973) *Biochimie* **55**, 1137-1151
26. Brown, R. D., and Brown, D. D. (1976) *J. Mol. Biol.* **102**, 1-14
27. Brownlee, G. G., Cartwright, E. M., and Brown, D. D. (1974) *J. Mol. Biol.* **89**, 703-718
28. Wensink, P. C., and Brown, D. D. (1971) *J. Mol. Biol.* **60**, 235-247
29. Weinmann, R., and Roeder, R. G. (1974) *Proc. Natl. Acad. Sci. U. S. A.* **71**, 1790-1794
30. Valenzuela, P., Weinberg, F., Bell, G., and Rutter, W. J. (1976) *J. Biol. Chem.* **251**, 1464-1470
31. Valenzuela, P., Hager, G. L., Weinberg, F., and Rutter, W. J. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73**, 1024-1028
32. Bell, G. I., Valenzuela, P., and Rutter, W. J. (1977) *J. Biol. Chem.* **252**, 3082-3091

(Reprinted from Nature, Vol. 267, No. 5612, pp. 641–643, June 16, 1977)
© Macmillan Journals Ltd., 1977

Nucleotide sequence of the yeast 5S ribosomal RNA gene and adjacent putative control regions

In bacteria, the sequence of several promoters has been determined and some of the features of their interaction with RNA polymerase, repressors and other regulatory molecules are beginning to be determined[1,2]. In eukaryotes, however, the isolation and analysis of the DNA structures involved in the regulation of transcription has not yet been achieved. Recent advances in techniques for cloning eukaryotic DNA in bacterial plasmids[3] have allowed the isolation of genetic segments whose structure and function can be studied in detail. Using these methods we have isolated a DNA fragment containing the 5S ribosomal RNA gene. The 5S RNA seems to be a primary transcript in yeast[4–6]; therefore the flanking regions of the 5S gene should contain the putative promoter and termination regions. We report here the nucleotide sequence of *Saccharomyces cerevisiae* 5S DNA and these flanking regions.

Yeast DNA and the plasmid vector pSF2124 (ref. 7) were digested with *Eco*RI, ligated, and used to transform a derivative of *E. coli* K12, DG-75 (a F^- derivative of DG73 ref. 8)). Recombinant clones were selected by ampicillin resistance and inability to produce colicin E1. Recombinants containing 5S rDNA were selected by hybridisation with purified yeast 5S rRNA. A clone (pDB25) containing a 2,520 base pair yeast DNA fragment which hybridises with 5S rRNA was obtained.

The tentative position of the 5S rRNA gene within the 2,520 base pair fragment was determined by restriction mapping with endonucleases *Hae*III, *Alu*I and *Hpa*II and by hybridisation of the restriction fragments with ^{125}I-labelled yeast 5S rRNA (P.V. *et al.* in preparation). The sequence of 5S RNA of *S. cerevisiae*[4] and of the related species, *S. carlsbergensis*[5], predicts a *Hae*III site eight base pairs from the 3′ end of the gene. Therefore as indicated in Fig. 1, we isolated the two *Hae*III fragments (1,120 and 520 base pairs, respectively) which encompass the 5S rRNA gene, labelled them at their 5′ ends with γ-^{32}P-ATP and polynucleotide kinase, and digested them with *Hpa*II. Two fragments, each containing a complementary portion of the 5S RNA gene, of 300 (*a*) and 400 (*b*) base pairs were isolated by preparative gel electrophoresis and partially sequenced from their ^{32}P-labelled *Hae*III ends. DNA was sequenced by using the chemical cleavage technique developed by Maxam and Gilbert[9]. Figure 2 shows an autoradiogram of the electrophoretic separation of oligonucleotides resulting from base specific cleavages of fragments *a* and *b*. The sequence can be directly read from the autoradiogram as exemplified in the figure for the first 10 nucleotides. The experiment with fragment *a* (Fig. 2*a*) generated a sequence comprising 112 bases from the 5S rRNA gene and 61 bases from the region adjacent to the 5′ end (termination region). An *Alu*I fragment of 160 base pairs containing this region was also sequenced (P.V. unpublished). The sequence of fragment *b* (Fig. 2*b*) furnished a sequence of 68 base pairs, 8 from the 5S gene and 60 from the region adjacent to its 3′ end (initiation region). The complete sequence of the 5S RNA gene and adjacent regions is shown in Fig. 3. The sequence of other regions, comprising about 70%, of the 2,520 base pair fragment has been completed and will be reported elsewhere.

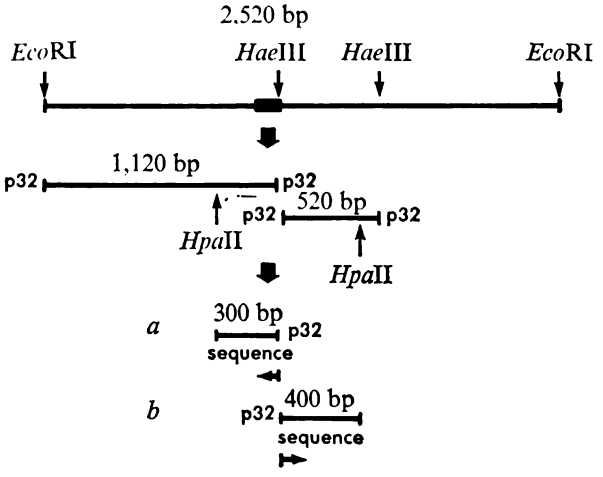


Fig. 1 Strategy for isolation and sequencing of yeast DNA fragments containing the 5S rRNA gene. The 2,520 base pair yeast DNA insert containing the 5S RNA gene was excised from the hybrid plasmid, pDB25, by digestion with *Eco*RI and preparatively separated from the bacterial vector by sucrose gradient centrifugation (5–20% sucrose, 20 h, SW-27 rotor, 20 °C). The *Hae*III fragments were prepared by digestion of 100 μg of the yeast DNA insert with *Hae*III and separated by preparative slab gel electrophoresis (5% acrylamide, 50 mM Tris-borate (*p*H 8.3), 1 mM EDTA, at constant 10 V cm^{-1}). After elution from the gel, the isolated fragments were labelled at their 5′ ends (T4 polynucleotide kinase and ^{32}P-ATP (ref. 9)) and digested with *Hpa*II. Fragments *a* and *b* were separated by preparative gel electrophoresis (as above) and sequenced from their labelled 5′ ends.

The sequence of most of the 5S RNA from *S. cerevisiae* has been reported by Miyazaki[4]. However, a small region around the 35th position (from the 5′ end) remained uncertain[4]. The sequence we have determined for the 5S RNA gene confirms Miyazaki's results and establishes the sequence of this region. Comparison with the sequence of 5S RNA from *S. carlsbergensis*[5] indicates small differences. Some may be due to species differences while others are probably due to minor errors in the proposed RNA sequence. Four of these changes occur in a region where, according to Hindley and Page[5], the sequence was uncertain.

The nucleotide sequence of the 60 base pairs adjacent to the 3′ end of the 5S gene should contain the site where RNA polymerase attaches to DNA to initiate transcription and possibly the sites for other regulatory molecules. This region may be compared with the known promoter regions of prokaryotes. Pribnow[10] has observed that several prokaryotic promoters contain an AT-rich heptanucleotide centred 9 to 10 bases before the transcription initiation site. Mutations that increase promotion in the *lac* operon are of the type GC to AT and occur in these regions or the immediate vicinity[1]. The putative promoter in yeast contains an AT-rich sequence, GTTAAAC, centred 8 bases from the start of 5S RNA transcription and bears some similarity to the 'Pribnow box'. The yeast promoter does not possess the general sequence homology (TTGA and GTTG) found

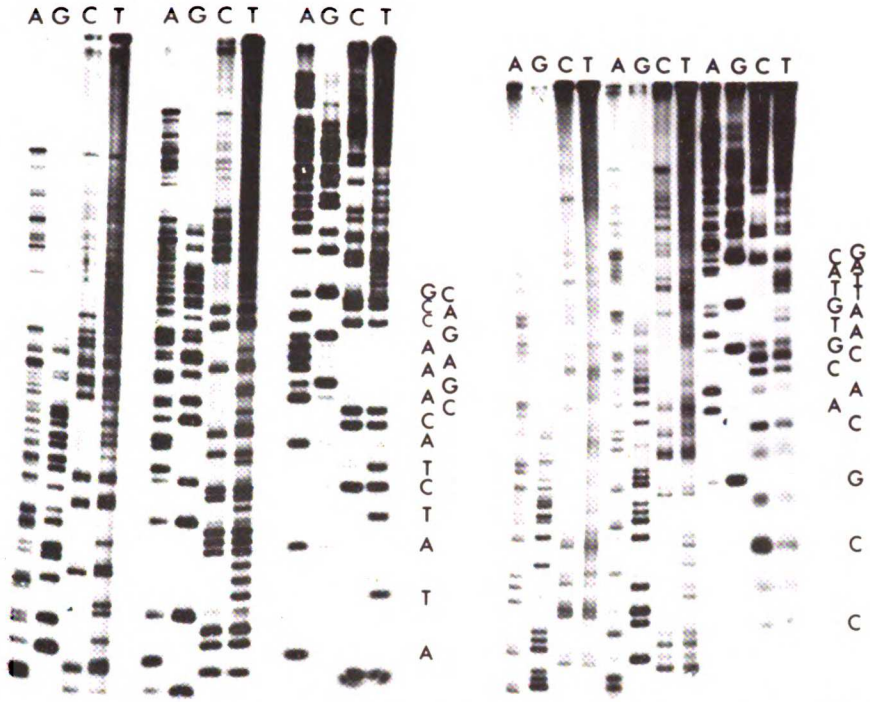


Fig. 2 Autoradiogram of the electrophoretic separation of ^{32}P-labelled oligonucleotides obtained by base-specific cleavages of fragments *a* and *b*. Both fragments (prepared as indicated in Fig. 1) were chemically degraded for sequence analysis as indicated by Maxam and Gilbert[9]. Preferential cleavage at guanine was achieved by methylation with 50 mM dimethylsulphate (10 min, 22 C), ring opening and displacement at neutral *p*H and phosphate elimination by 1 M piperidine (60 min, 90 C). Cleavage at adenine was achieved by methylation with dimethylsulphate (23 min, 22 C), depurination in 0.1 N HCl (120 min, 0 C) and phosphate elimination by 0.1 N NaOH (30 min, 90 C). Modification and cleavage at cytosine and thymine residues were obtained by treatment with 15 M hydrazine (20 min, 22 C). Cleavage only at cytosine was achieved by performing the reaction in the presence of 2 M NaCl (25 min, 22 C). In both instances, phosphates were eliminated by treatment with 1 M piperidine (60 min, 90 C). Aliquots of the products from the four reactions were loaded on a 20% acrylamide–7 M urea gel[9] at times 0, 12 and 24 h. Electrophoresis was carried out in 50 mM Tris-borate (*p*H 8.3), 1 mM EDTA at constant 20 V cm^{-1}. Both fragments were isolated and sequenced twice. The columns A, G, C and T indicate wells loaded with products from reactions partially specific for adenine, guanine, cytosine, and thymine, respectively. The reading of the first 10 bases is illustrated at the right of each autoradiogram. The first two nucleotides of the labelled 5′ end of fragment *b* ran off the gel and are determined from the specificity of *Hae*III. Fragment *a* is at the left, fragment *b* is at the right.

in prokaryotes in a region centred about 35 bases from the start of the gene, nor palindromes which in some prokaryotic promoters have been identified as binding sites for regulatory proteins[1,2]. Fedoroff and Brown (personal communication) have recently determined the structure of the promoter region for *Xenopus* oocyte 5S RNA genes. There is no obvious relationship between the structures of this region and the region we have sequenced. The fundamental structural relationships which determine transcription promotion in eukaryotes are as yet difficult to perceive.

In contrast, at the end of the 5S gene there is a striking stretch of As in the coding strand (Fig. 3). We believe this to be a termination signal for RNA polymerase. Termination in other systems may be related to A-rich regions. For example, variable runs of Us have been found at the 3′ end of 4S (ref. 11) and 6S (ref. 12) RNAs of λ phage, a small RNA from phage φ80 (ref. 13) and the attenuator region of tryptophan mRNA of *E. coli*[14]. There is no evidence for transcription of this A region in yeast; however, transcription and rapid processing has not been eliminated. In eukaryotes, Brown and Brown[15] have found a much shorter AT-rich region at the 5′ end of the *Xenopus* 5S RNA gene and postulated the sequence $\frac{\text{TTTT}}{\text{AAAA}}$ as part of a transcription termination signal.

RNA polymerase binding and *in vitro* transcription experiments will be necessary to determine the role of these sequences in selective gene transcription.

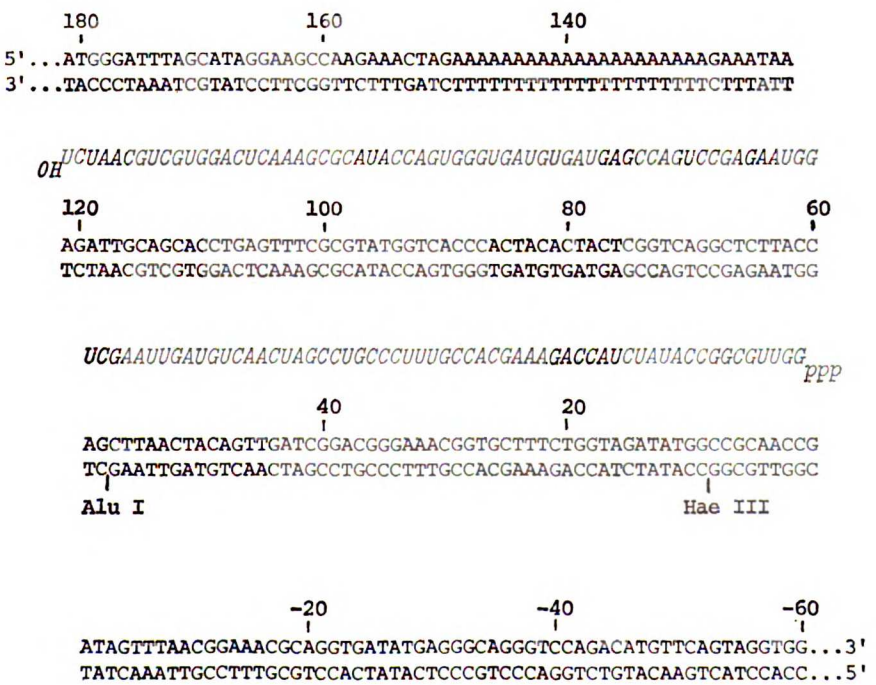


Fig. 3 Nucleotide sequence of *S. cerevisiae* 5S rRNAgene and adjacent regions. The corresponding 5S rRNA sequence is shown in italics.

This research was supported by a grant from the National Institutes of Health. We thank J. Shine for advice and A. Maxam and W. Gilbert for providing a detailed protocol of their DNA sequencing method before publication.

PABLO VALENZUELA
GRAEME I. BELL
FRANK R. MASIARZ
LOUIS J. DEGENNARO
WILLIAM J. RUTTER

Department of Biochemistry and Biophysics,
University of California,
San Francisco, California 94143

Received 24 March; accepted 19 April 1977.

1 Gilbert, W. in *RNA Polymerase* (eds Losick, R. & Chamberlin, M.) 193–205 (Cold Spring Harbor Press, 1976).
2 Dickson, R. C., Abelson, J., Barnes, W. M. & Reznikoff, W. S. *Science* **187**, 27–35 (1975).
3 Morrow, J. F., Cohen, S. N., Chang, A. C. Y., Boyer, H. W., Goodman, H. M. & Helling, R. B. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1743–1747 (1974).
4 Miyazaki, M. *J. Biochem.* **75**, 1407–1410 (1974).
5 Hindley, J. & Page, S. M. *FEBS Lett.* **26**, 157–160 (1972).
6 Trapman, J. & Planta, R. J. *Biochim. biophys. Acta* **414**, 115–125 (1975).
7 So, M., Gill, R. & Falkow, S. *Molec. gen. Genet.* **142**, 239–249 (1975).
8 Polisky, B., Bishop, R. J. & Gelfand, D. H. *Proc. natn. Acad. Sci. U.S.A.* **73**, 3900–3904 (1976).
9 Maxam, A. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
10 Pribnow, D. *Proc. natn. Acad. Sci. U.S.A.* **72**, 784–788 (1975).
11 Sklar, J., Yot, P. & Weissman, S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 1817–1821 (1975).
12 Dahlberg, J. & Blattner, F. *Fedn Proc.* (Abstr.) **32**, 664 (1973).
13 Pieczenik, G., Barrell, B. G. & Gefter, M. L. *Archs Biochem. Biophys.* **152**, 152–165 (1972).
14 Bertrand, K. *et al. Science* **189**, 22–26 (1975).
15 Brown, R. D. & Brown, D. D. *J. molec. Biol.* **102**, 1–14 (1976).