Volume 9 Number 9 1981 Nucleic Acids Research

Nucleotide sequence of an Escherichia coli tRNA (Leu 1) operon and identification of the transcription promoter signal

Gregg Duester, Renee K.Campen and W.Michael Holmes

Medical College of Virginia, Virginia Commonwealth University, Box 678-MCV Station, Richmond, VA 23298, USA

Received 10 December 1980

ABSTRACT

Fourteen different DNA fragments containing *Escherichia coli* tRNA genes have been cloned using the vector pBR322. We report the methods of cloning, the identification of specific tRNA genes, and the presence or absence of rRNA genes on these cloned DNA fragments. In particular, one chimeric plasmid contains a 17.0 kilobase pair *Eco*RI fragment bearing tRNA(Leu 1) sequences. Using nucleotide sequence analysis we have identified a cluster of three tandem tRNA(Leu 1) genes separated by intergenic spacers of 27 and 34 base pairs, respectively. The nucleotide sequence upstream of the first gene contains a transcription promoter site. A G-C rich sequence (5'-CGCCTCC-3') found between the Pribnow box and initiation site is very similar to the corresponding sequence found in other genes which are under stringent control.

INTRODUCTION

The genome of *Escherichia coli* has been shown to contain about 60 tRNA genes (1). At least 18 of these genes have been shown to be closely associated with rRNA operons as spacer tRNA genes between 16S and 23S rRNA genes, or as distal tRNA genes lying downstream from the 5S rRNA genes (2,3,4). In addition, other tRNA genes have been located at various locations distinct from the rRNA operons, and some clustering of these genes has been observed (5). Since all tRNA and rRNA genes are subject to the coordinate control mechanism known as the stringent response (6), a comparison of the various DNA regulatory sequences of these genes may help elucidate the molecular mechanisms of this phenomenon. The seven rRNA operons have been studied in detail and many of the promoters (7,8,9) and terminators (10,11,12) have been sequenced. However, there is very limited information on regulatory sequences for tRNA operons. The promoter (13) and terminator (14) of the tRNA(Tyr 1) operon are the only well-characterized regulatory sequences for the tRNA genes unlinked to rRNA operons.

We have cloned several *E. coli* DNA fragments containing tRNA genes. These various cloned segments of the chromosome have been examined for rRNA

© IRL Press Limited, 1 Falconberg Court, London W1V 5FG, U.K.

and tRNA gene content. One particular clone was found to possess three tandem tRNA(Leu 1) genes. The nucleotide sequence of this tRNA operon was determined and evidence is presented that a transcription promoter signal lies just upstream of the first tRNA(Leu 1) gene.

MATERIALS AND METHODS

Materials

Restriction endonucleases were obtained from Bethesda Research Laboratories or New England Biolabs, Inc. Purified E. coli 5S rRNA, 16S + 23S rRNA mixture, bulk tRNA, calf intestine alkaline phosphatase, and the Klenow (A) fragment of E. coli DNA polymerase I were purchased from Boehringer Mannheim Biochemicals. Purified individual E. coli tRNA species were from Plenum Scientific Research, Inc. T-4 DNA ligase and agarose were from Bethesda Research Laboratories. Acrylamide, bis-acrylamide, and urea were from Biorad. T-4 RNA ligase and pppGp were from P-L Biochemicals. E. coli RNA polymerase holoenzyme was purchased from Enzo Biochem, Inc. Pancreatic RNAase A was from Sigma Chemical Co. and RNAase T_1 (Sankyo) was obtained from Calbiochem. [5'-^{32}P] cytidine 5',3'-diphosphate (2000-3000 Ci/mmol) was from Amersham. [γ-^{32}P] GTP (1000-3000 Ci/mmol), [α-^{32}P] ribonucleoside-5'-triphosphate mixture (400 Ci/mmol), [α-^{32}P] dATP (400 Ci/mmol) and [α-^{32}P] dCTP (400 Ci/mmol) were from ICN.

Cloning of E. coli DNA Fragments and Identification of Chimeric Plasmids Bearing tRNA Genes

High molecular weight E. coli K12 (C600) DNA was prepared as described (15). The vector pBR322 (16) was kindly provided by Dr. Patricia Shipley. One μg of E. coli DNA and 3.0 μg of pBR322 were digested to completion with EcoRI. The pBR322 was then treated with calf intestine alkaline phosphatase (to prevent self-ligation) in 0.1 M Glycine•NaOH (pH 9.5), 1.0 mM $MgCl_2$, 1.0 mM $ZnSO_2$ using 0.1 μg enzyme per 1.0 μg DNA for 30 min at 37°C. The pBR322 was then phenol extracted twice and ethanol precipitated three times. E. coli DNA and alkaline phosphatase-treated pBR322 were ligated in 50 mM Tris•HCl (pH 7.6), 10 mM $MgCl_2$, 0.1 mM ATP, 10 mM DTT, 100 μg/ml BSA using 10 units of T-4 DNA ligase at 4°C for 18 hr. Transformation of E. coli K12, MV-12 (recA$^-$) was performed as described (17). The transformed cells were spread onto Luria broth plates containing 25 μg/ml ampicillin and grown overnight at 37°C.

In order to enrich our collection of transformants for plasmids contain-

ing large molecular weight E. coli EcoRI fragments, the above procedure was repeated using EcoRI fragments fractionated by agarose gel electrophoresis. Ten μg of EcoRI digested E. coli DNA was electrophoresed at 100 V for 16 h in a horizontal 0.5% agarose gel, and all the bands above 8 kilobase pairs long were isolated as previously described (18). This DNA was ligated to pBR322 and transformed as mentioned above. In all, 1000 individual transformants were collected.

The transformants were screened for chimeric plasmids carrying tRNA genes by the following procedure. Individual transformants were grown in 1.5 ml of Luria broth overnight. Cells were pelleted in an Eppendorf microfuge for 30 sec, then resuspended with vortexing in 150 μl of 50 mM NaCl, 50 mM EDTA, 15% Sucrose, 90 mM Tris-Borate (pH 8.3), 1.0 mg/ml lysozyme, 0.1 mg/ml RNAase A. This solution was incubated for 10 min at room temperature, then 10 μl of 10% SDS was added and mixed in gently by hand (no vortexing). The solution was then incubated at 0° for 30 min, followed by a centrifugation for 15 minutes in an Eppendorf microfuge at room temperature. About 100 μl of the supernatant fraction was carefully removed with a pipet making sure absolutely none of the pellet (containing the majority of the chromosomal DNA) was disturbed. To the supernate (enriched for plasmid DNA) was added 15 μl of 50% glycerol, 1% SDS, 0.1% bromphenol blue. Fifty μl of this solution was loaded onto a 0.9% agarose gel and electrophoresed for 2 hr at 120 V using a buffer composed of 90 mM Tris-Borate (pH 8.3) and 2.5 mM EDTA. DNA fragments were transferred to nitrocellulose filters using the Southern blot method (19), and hybridized to ^{32}P-radiolabeled bulk E. coli tRNA as described in the following section. Autoradiography was used to locate the plasmids bearing tRNA genes. The transformants containing cloned tRNA genes were maintained in Luria broth with 25 μg/ml ampicillin and 25% glycerol at -20°C.

The preparation of milligram quantities of plasmid DNA was by the CsCl-ethidium bromide method of Clewell (20).

Restriction Endonuclease Digestion, Fractionation of DNA and Hybridization to Radiolabeled RNA

DNA was digested with restriction endonucleases under conditions specified by the enzyme manufacturers. Digested DNA was fractionated by electrophoresis in agarose or polyacrylamide gels using 90 mM Tris-borate pH 8.3, 2.5 mM EDTA as the buffer. Fractionated DNA was transferred to nitrocellulose filters and hybridized to ^{32}P-radiolabeled tRNA or rRNA by the method of

Southern (19) as modified by Campen et. al. (21). All tRNA and rRNA species were labeled *in vitro* at the 3' terminus as outlined by Bruce and Uhlenbeck (22), using T4 RNA ligase and [5'-^{32}P] cytidine-5',3'-diphosphate. The purity of the radiolabeled tRNA and rRNA was checked by electrophoresis on 10% acrylamide-7M urea gels. To overcome the problem of contaminating RNA species, each labeled probe was competed with unlabeled RNA species during hybridization. In experiments where bulk tRNA was used as a probe, unlabeled 16S + 23S rRNA (mixture) and 5S rRNA in 10-fold molar excess were added to the hybridization mixture. In the case of individual tRNA probes, a 10-fold molar excess of unlabeled bulk tRNA, 16S + 23S rRNA and 5S rRNA were added. For the 16S + 23S rRNA probe, unlabeled 5S rRNA and bulk tRNA were added, and for the 5S rRNA probe, unlabeled 16S + 23S rRNA and bulk tRNA were added in 10-fold molar excess.

The length in base pairs of fractionated DNA fragments was determined by comparison to standard DNA fragments of known molecular size liberated by *Hind* III or *Eco*RI cleavage of λ DNA (23), and *Alu*I or *Hinf* I cleavage of pBR322 (24).

Restriction Endonuclease Mapping

Restriction endonuclease sites were mapped by using double digestions employing various enzymes, or by the partial digestion technique of Smith and Birnstiel (25).

Purification and Labeling of DNA Fragments for DNA Sequencing

DNA fragments were purified by electro-elution from acrylamide gels as described previously (11).

Radiolabeling of DNA fragments was performed by partially filling 5' protruding sticky ends. The reaction was carried out for 3 hr at 12°C in 100 μl of 20 mM Tris-HCl (pH 7.4), 60 mM NaCl, 7.0 mM $MgCl_2$, 10 mM DTT, with 3 units of the Klenow (A) fragment of *E. coli* DNA polymerase I (26), using 10 nmoles of a single ^{32}P-labeled deoxynucleoside triphosphate per 1.0 nmole of DNA ends. For example, ^{32}P-dCTP was used to label DNA fragments obtained by *Hpa*II digestion.

DNA sequencing was performed by the chemical modification method of Maxam and Gilbert (27).

In Vitro Transcription and Fractionation of Transcripts

Transcription of DNA fragments was accomplished using RNA polymerase

holoenzyme purchased from Enzo Biochem, Inc. A standard reaction contained 50 mM Tris-HCl pH 8.0, 10 mM $MgCl_2$, 0.1 mM EDTA, 10 mM dithiothreitol, 75 mM KCl, 20% (vol/vol) glycerol, 100 μM (ATP, CTP, and UTP), and 1.0 mM GTP. A 20 μl reaction was initiated by the addition of 0.25 μg of DNA template plus 1.0 μg of RNA polymerase. For analytical purposes 5.0 μCi of radiolabeled nucleoside-5'-triphosphates was used; 50 μCi was used for preparative experiments. The sample was incubated at 37°C for 60 min, then stopped by addition of 250 μl of a solution containing 0.3 M sodium acetate, 0.1% sodium dodecyl sulfate and 25 μg carrier tRNA. One phenol extraction was performed followed by an ethanol precipitation. The pellet was resuspended in 10 μl of 80% formamide, 5 mM Tris-borate pH 8.3, 0.1 mM EDTA, 0.1% xylene cyanol, and 0.1% bromphenol blue. This sample was loaded on a 15% acrylamide-7 M urea gel and electrophoresis was carried out at 300 V for 1.5 hours. The labeled transcripts were detected by autoradiography.

RNA transcripts were isolated by elution from the acrylamide gel. The gel slice containing the transcript was minced with a razor and soaked at 37°C overnight in 400 μl of a solution containing 0.5 M KCl, 0.01 M Tris-HCl pH 7.0, 1.0 mM EDTA, and 50 μg of carrier tRNA. The solution was phenol extracted once and ethanol precipitated three times before further use.

The RNA transcripts were fingerprinted as described by Brownlee (28). PEI-cellulose chromatography was performed as described by Cashel, et. al. (29).

RESULTS

Cloning E. coli tRNA Genes

We used the restriction endonuclease *Eco*RI to cleave *E. coli* DNA for cloning since no *Eco*RI sites are present in the nucleotide sequences of any *E. coli* tRNA species examined so far (30). In addition, we had previously determined the molecular sizes of 27 *Eco*RI fragments carrying tRNA genes as determined by Southern blot hybridization analysis using ^{32}P-labeled bulk *E. coli* tRNA (21). We attempted to clone as many of these *Eco*RI fragments as possible for detailed studies. The alkaline phosphatase treatment of the *Eco*RI-cleaved pBR322 (mentioned in Materials and Methods) helped enormously in the cloning procedure since 95% of all ampicillin-positive clones were shown to possess chimeric plasmids. A total of 1000 clones carrying various fragments of the *E. coli* chromosome were screened for tRNA genes as described in Materials and Methods. In all, 20 clones were shown to contain tRNA genes, and chimeric plasmid DNA from each was purified in large quantitiy.

All were digested with EcoRI and subjected to electrophoresis in 1.0% agarose. Fragments larger than 10 kilobase pairs were also run on 0.5% agarose gels. The molecular sizes of the cloned EcoRI fragments were estimated by comparison of their mobilities with those of lambda DNA cleaved with Hind III or EcoRI. The presence of 16S + 23S rRNA and/or 5S rRNA genes on each cloned fragment was tested by Southern blotting and hybridization to a ^{32}P-labeled 16S + 23S rRNA mixture or 5S rRNA. In addition, the presence of tRNA(Leu 1) and tRNA(Thr 1) genes was determined by Southern blotting and hybridization to these ^{32}P-labeled individual tRNAs (obtained from Plenum Scientific Research, Inc.). All this data is summarized in Table 1.

Our data indicate that many cloned tRNA genes reside in or near the rRNA operons, and indeed Morgan, et. al. have shown that many tRNA genes are in

CHIMERIC PLASMID	tRNA	16S+23S rRNA	5S rRNA	kilobase pairs of EcoRI insert
1. RA76, C43	+	+	-	2.0
2. C61	+	+	-	2.1
3. RB7, D96	+	-	-	3.0
4. G4, C24	Thr 1	+	+	3.9
5. B76, D25	+	+	+	4.4
6. B49	+	-	-	6.0
7. LB35	+	-	-	7.3
8. LA37	+	+	+	8.7
9. B74, G36	+	-	-	9.5
10. LB1	+	+	+	10.5
11. LA39	+	+	-	12.0
12. D63	+	-	-	12.0
13. LB16	+	-	-	14.5
14. LA4, LA70	Leu 1	-	-	17.0

Table 1. Summary of chimeric plasmids bearing *E. coli* tRNA genes. EcoRI fragments of the *E. coli* genome were ligated to pBR322, transformed, and screened as described in Materials and Methods. In some cases the plasmids from two separate transformants contained the same cloned EcoRI fragment. The presence of various genes was determined by Southern blot hybridization of ^{32}P-labeled individual tRNAs, 16S + 23S rRNA, 5S rRNA, and bulk tRNA to EcoRI digested plasmid DNA as described in Materials and Methods.

such a location (3). In Table 1 it can be seen that we have cloned 7 unique EcoRI fragments derived from the rRNA operons. Boros, et. al. (31) recently provided EcoRI restriction endonuclease maps of all seven rRNA operons in E. coli. Using this data, three of our cloned fragments appear to be about the same size as the spacer tRNA regions in various rRNA operons (see Table 1, clones RA76, C61, B76); also, three cloned fragments (G4, LA37, LB1) can be identified as containing tRNA genes distal to rRNA operons, and one fragment (LA39) appears to contain a tRNA gene somewhere upstream of a rRNA operon. In particular, we have verified by DNA sequence analysis that G4 (pGD4) contains a tRNA(Thr 1) gene at the distal end of the rRNA operon rrnD (11). It appears by in vitro transcription studies that the tRNA(Thr 1) gene is co-transcribed with rrnD (W.M. Holmes, unpublished results). It is not known if the other two distal tRNA genes are co-transcribed with rRNA, and the upstream tRNA gene is presumably located in an operon separate from the rRNA operon it precedes.

Seven Eco RI fragments have been cloned which contain tRNA but not rRNA genes (Table 1). Our results (see next section) have shown that plasmid LA4 contains three tandem tRNA(Leu 1) genes which reside on a 17 kilobase pair EcoRI fragment. Genes for tRNA(Leu 1) have been mapped near 83 and 93 minutes (5). LA4 is most likely derived from the 93 minute region of the chromosome since the other tRNA(Leu 1) locus (83 minutes) has recently been cloned and consists of a single tRNA(Leu 1) gene along with tRNA(His) and tRNA(Pro) genes (M. Fournier, personal communication). In addition, we have shown that tRNA(His) and tRNA(Leu 1) genes may reside on the same 14.5 kilobase pair EcoRI fragment in the chromosome (21). The precise identity of the other tRNA genes awaits more detailed studies.

Since reliable assignments of particular genes to particular EcoRI fragments depended on pure RNA probes, there were some difficulties encountered. Analysis of the ^{32}P-labeled tRNA and rRNA on polyacrylamide gels indicated that minor contaminants were present (data not shown). The use of unlabeled competing RNA in the hybridization mixture was employed to hopefully overcome this problem (see Materials and Methods).

Characterization of a Plasmid Carrying Three Leucine tRNA Genes

Of particular interest to us is the plasmid LA4 which hybridizes to tRNA(Leu 1). There are two loci for tRNA(Leu 1) genes as mentioned in the previous section, and we had reason to believe that at least one locus would contain multiple copies of this gene. This is because Ilgen, et. al. (32)

had shown that an RNase P mutant of E. coli would accumulate a large tRNA precursor of 440 nucleotides in length which could be processed in vitro to yield 4.9 mature tRNA(Leu 1) species per precursor. We decided to determine by physical mapping and DNA sequence analysis whether or not plasmid LA4 contains multiple tRNA(Leu 1) genes.

Since plasmid LA4 is so large in size (21 kilobase pairs including the pBR322 vector) the tRNA(Leu 1) gene(s) were sub-cloned on a smaller plasmid. Plasmid LA4 is cleaved into 4 SalI fragments and the 3.2 kilobase pair fragment hybridizes to tRNA(Leu 1), as shown in Fig. 1 (lanes 2 and 4). By simply re-ligating the SalI fragments, a 7 kilobase pair plasmid bearing the pBR322 origin and ampicillin resistance gene was obtained as depicted in Fig. 2. Hybridization showed that tRNA(Leu 1) sequences were also retained on

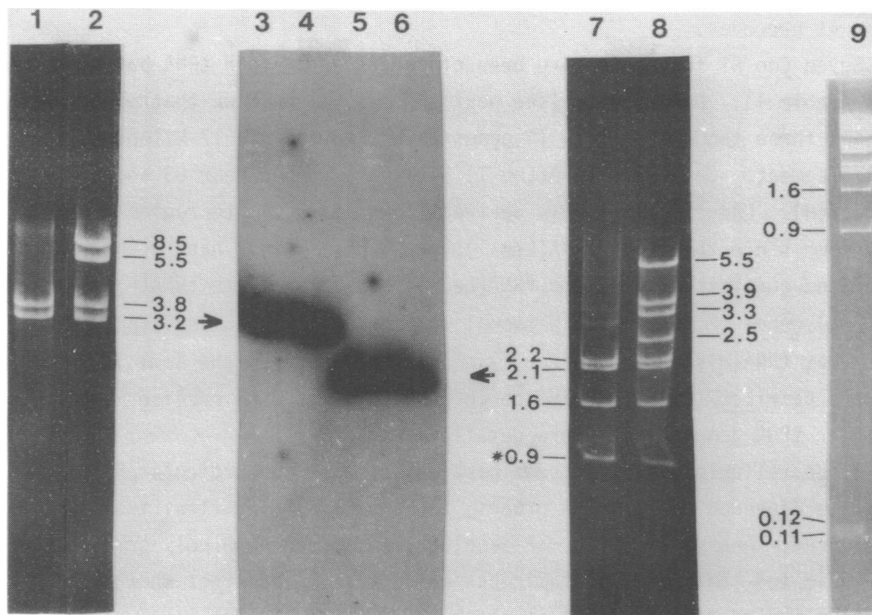

Figure 1. Restriction endonuclease digests of plasmid LA4 and pLEU1. Lanes 1 and 2 are SalI digests and lanes 7 and 8 are AccI digests of pLEU1 and LA4, respectively, run on 1% agarose gels and stained with ethidium bromide. Lanes 3, 4, 5, and 6 are the autoradiographs of the Southern blotted gels in lanes 1, 2, 7, and 8 after hybridization to ^{32}P-labeled tRNA(Leu 1). Lane 9 is an AccI digest of LA4 run on a 7% acrylamide gel. The numbers are in Kilobase pairs. The asterisk next to the 0.9 kilobase pair fragment indicates that it does hybridize to tRNA(Leu 1) very weakly, but is not visible in the autoradiographs in lanes 5 and 6.

this plasmid, designated pLEU1; i.e. the 3.2 kilobase pair SalI fragment was retained (Fig. 1, lanes 1 and 3). A SalI site located very near the pBR322 vector in LA4 made it possible to liberate a small mini-vector (3.8 kilobase pairs) which contained the pBR322 origin and ampicillin resistance gene (Fig. 2).

The tRNA(Leu 1) sequences were localized on pLEU1 using the enzyme AccI which was predicted to cut in the gene on the basis of the known sequence of tRNA(Leu 1) (33, 34). Figure 1 (lanes 5, 6, 7, 8) shows the AccI digestion patterns of LA4 and pLEU1. The lower 4 fragments of LA4 have been preserved in pLEU1. It can be seen that the 2.1 kilobase pair fragment hybridizes strongly to tRNA(Leu 1), and the 0.9 kilobase pair fragment hybridizes very weakly. Since AccI cuts 20 base pairs from the 5' end of the gene (which is 87 base pairs long), the AccI fragment containing the 5' sequences (20 base pairs) would probably hybridize much less efficiently than the AccI fragment containing the 3' sequences (67 base pairs). Therefore, the direction of transcription was tentatively assigned as proceeding from the 0.9 to the 2.1 kilobase pair fragment. The presence of multiple tRNA(Leu 1) genes was suggested by the finding of two very small AccI fragments (0.12 and 0.11 kilobase pairs) as shown in Fig. 1, lane 9. Two such small fragments would be liberated if three tRNA(Leu 1) genes lie very close together. This was confirmed by DNA sequencing as discussed in the next section.

DNA Sequence Analysis of an Operon Containing Three Tandem Leucine tRNA Genes

A more detailed restriction map of the region containing tRNA(Leu 1) was obtained (expanded map in Fig. 2), and DNA sequencing was performed by the method of Maxam and Gilbert (27). The DNA strands sequenced are shown on the bottom of Fig. 2. Figure 3 shows a representative sequencing gel which corresponds to a region upstream of the first tRNA(Leu 1) gene. The DNA sequence from 145 base pairs upstream of the first tRNA(Leu 1) gene to 231 base pairs beyond the third tRNA(Leu 1) gene is shown in Fig. 4. DNA sequencing showed in fact that three tRNA(Leu 1) genes were present. The sequence of all three genes matches the previously determined sequence of tRNA(Leu 1) (33, 34). The first and second spacer regions between the tRNA(Leu 1) genes are 27 and 34 base pairs in length, respectively. No additional tRNA genes are located in the sequence.

About 40 base pairs upstream of the first gene is a heptanucleotide sequence 5'-TAGAATG-3' (boxed in Fig. 4) which differs by only one base pair from the classic "Pribnow box" or consensus -10 sequence found in many

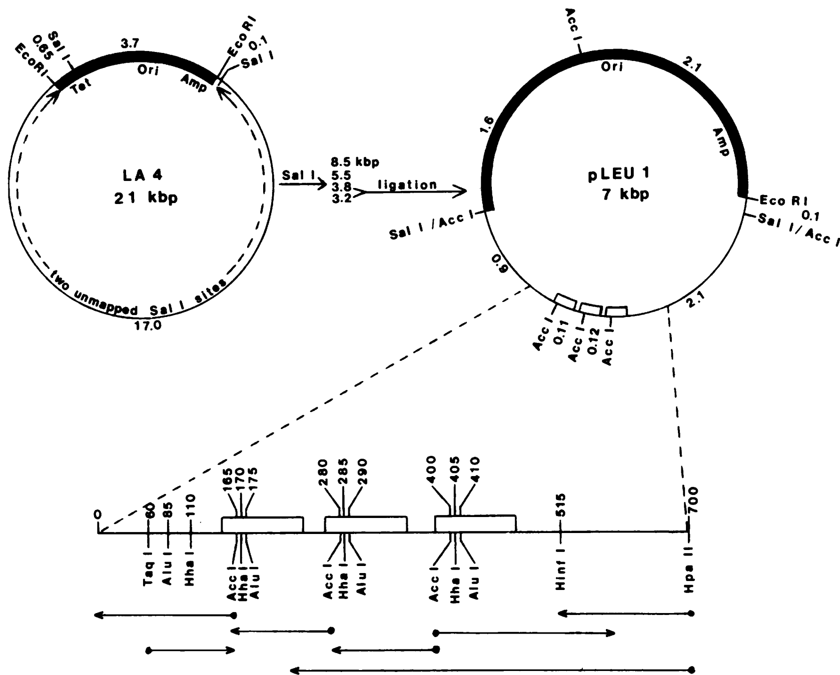

Figure 2. Restriction endonuclease map of pLEU1. The plasmid pLEU1 is derived from the two small SalI fragments in LA4. The numbers on LA4 and pLEU1 are in kilobase pairs (kbp). The numbers on the expanded DNA segment are in base pairs. The solid bars refer to pBR322 sequences, and the open bars refer to tDNA(Leu 1). The leftward-pointing arrows below the map denote regions sequenced 3' to 5' on the top strand; rightward-pointing arrows denote regions sequenced 3' to 5' on the bottom strand.

prokaryotic transcription promoter regions (35). In addition, an octanucleotide sequence 5'-TATTGACG-3' (boxed in Fig. 4) differs from the consensus -35 promoter sequence (35) by only two base pairs; also it is located about 25 base pairs upstream of the putative -10 sequence. One additional feature of the upstream sequence is the presence of an A+T rich inverted repeat (overlined in Fig. 4) just prior to the 5' end of the first tRNA(Leu 1) gene. The sequence downstream of the third tRNA(Leu 1) gene is very A+T rich; 25 of the first 40 downstream bases on the anti-sense strand are thymine residues. There is also a small inverted repeat about 90 base pairs downstream.

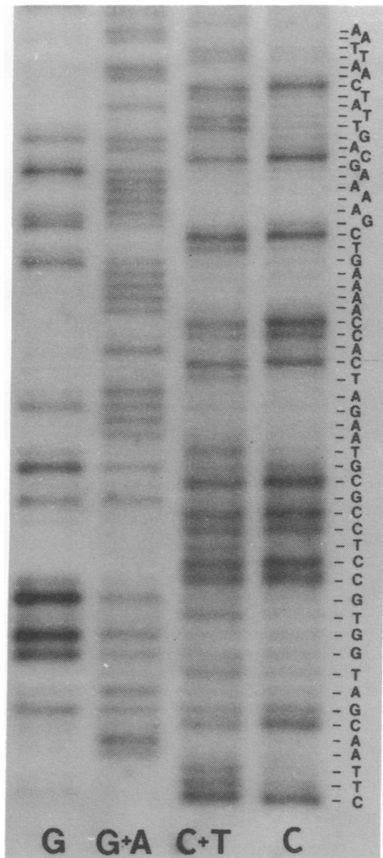

Figure 3. DNA sequencing gel. The 0.9 kilobase pair AccI fragment of pLEU1 was end-labeled with ^{32}P-dATP using the Klenow fragment of DNA polymerase, subjected to the chemical sequencing method of Maxam and Gilbert (27), then electrophoresed in an 8% acrylamide-50% urea gel. The sequence as read from top to bottom shows the region from base pair 67 to 124 on the top strand in Figure 4; this region contains the promoter of the tRNA(Leu 1) gene cluster.

In Vitro Transcription of a Taq I-Acc I Fragment

The region immediately upstream of the first tRNA(Leu 1) gene was transcribed in vitro to identify a transcription promoter signal. The 104 base pair DNA fragment spanning the region from the Taq I site (base pair 63, Fig. 4) to the Acc I site (base pair 166, Fig. 4) was isolated from pLEU 1 after electrophoresis in a 10% acrylamide gel. The optimum conditions for transcription (Materials and Methods) were found to be very similar to those

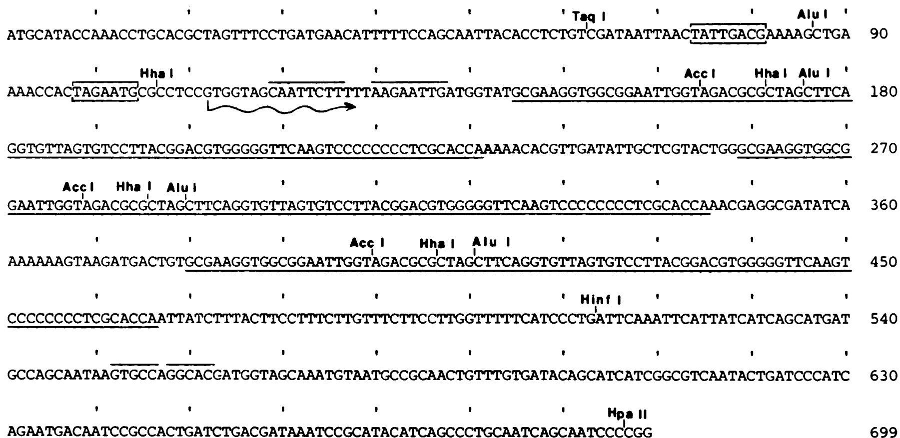

Figure 4. DNA sequence of a cluster of three tandem tRNA(Leu 1) genes. Only the anti-sense strand is shown. The three tRNA(Leu 1) sequences are underlined. A -10 promoter sequence is boxed near base pair 100 and a -35 promoter sequence is boxed near base pair 80. A region of A-T rich dyad symmetry is overlined from base pair 119 to 137. Another region of dyad symmetry is overlined from base pair 552 to 562. The initiation site for transcription is indicated by the wavy arrow.

used by Kupper, et al. for the tRNA(Tyr 1) promoter (36). Figure 5 shows the results of an _in vitro_ transcription experiment. The major RNA product was estimated to be 55 nucleotides in length. Assuming that transcription

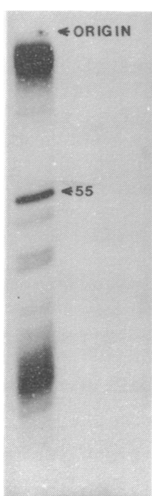

Figure 5. In vitro transcription. The _Taq_ I-_Acc_ I DNA fragment upstream of the first tRNA(Leu 1) gene was transcribed in vitro with _E. coli_ RNA polymerase holoenzyme as mentioned in Materials and Methods. An equal mixture of all four [α-^{32}P] ribonucleoside-5'-triphosphates ensured that the products were uniformly labeled. After fractionation on a 15% acrylamide-7 M urea gel, the labeled products were detected by autoradiography as shown above. The major product (marked with an arrow) was estimated to be about 55 nucleotides in length by comparison of its migration distance to that of tRNA(Thr 1) [76 nucleotides], tRNA(Leu 1) [87 nucleotides], and 5S rRNA [120 nucleotides]. The standards were detected by staining the gel in a solution of 1 μg/ml ethidium bromide and subjecting it to 260 nm ultraviolet light (data not shown).

proceeds up to (or close to) the *Acc* I end of the DNA template, the initiation event appears to have occurred just downstream of the proposed -10 sequence shown in Fig. 4. The bands below the major one may correspond to products produced by premature termination events or incomplete transcripts.

The 900 base pair *Acc* I fragment from pLEU 1 was also transcribed *in vitro* and shown to produce a major product of about 55 nucleotides in length. This may indicate that the sequences upstream of the *Taq* I site are unimportant for *in vitro* initiation.

The 55 nucleotide long transcript was isolated and Fig. 6 shows its fingerprint after RNase T-1 digestion. The identify of individual oligonucleotides was determined by secondary digestion with pancreatic RNase; this was facilitated by using a uniformly labeled transcript to ensure identification of each nucleotide. The spot labeled pppGp was reduced in yield, presumably due to breakdown to ppGp and pGp during isolation of the transcript; it was identified by co-migration with a pppGp standard on PEI-cellulose chromatography (detected by UV light). The sequences and molar yields of each oligonucleotide (indicated in Fig. 6) corresponded to the sequence of a transcript initiating at position 112 in Fig. 4 and terminating in the *Acc* I site near position 163. In particular, the initiating nucleotide was determined to be the guanosine residue at position 112 by various criteria. First pppGp was detected, thus indicating that initiation takes place at a guanosine residue. Second, there were two moles of UpGp released per transcript and these could only come from two sites: one from position 113-114 and one from position 152-153 on Fig. 4. Third, none of the oligonucleotides upstream of position 112 were seen. Also, a 55 nucleotide long transcript could be detected using γ-^{32}P-GTP, once again indicating that a guanosine is the initiating nucleotide. Therefore, the site of transcription initiation lies 8 base pairs downstream of the Pribnow box at position 112.

DISCUSSION

We have cloned 14 *Eco*RI fragments bearing *E. coli* tRNA genes, and have been able to identify which fragments also contain rRNA genes. The sizes of all the cloned *Eco*RI fragments match very closely the sizes of fragments in *Eco*RI digested *E. coli* chromosomal DNA which hybridize to bulk tRNA as described by Campen, et al. (21). In addition, the assignment of all 16S + 23S rRNA and 5S rRNA genes to cloned fragments of certain sizes matches our previously published results obtained by Southern blot analysis of

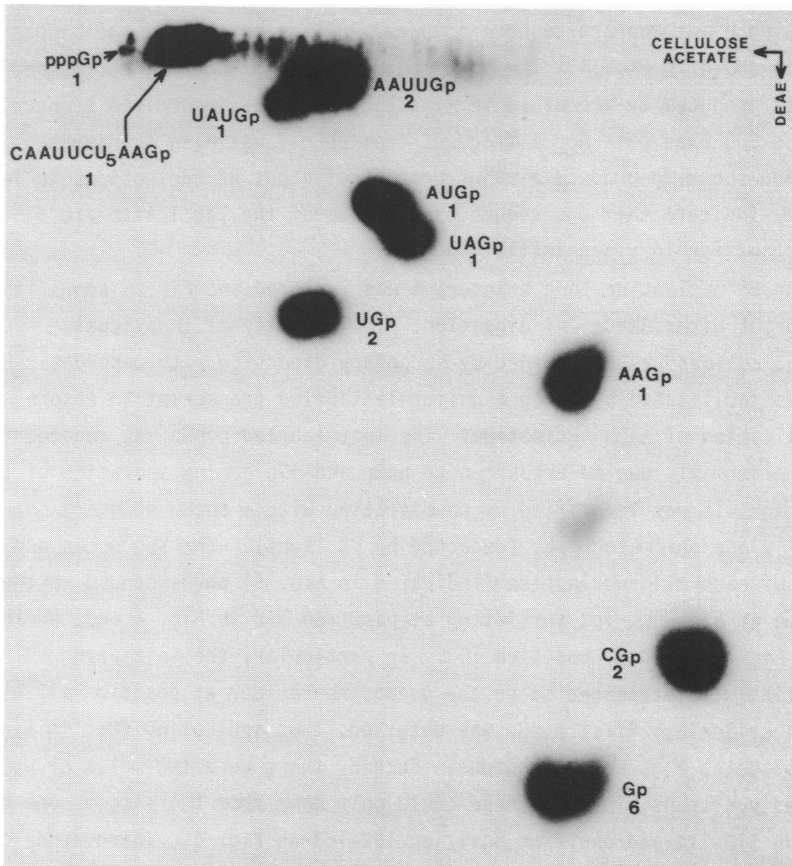

Figure 6. Fingerprint of the Taq I-Acc I transcript. The 55 nucleotide long RNA product in Fig. 5 was eluted from the gel and digested with RNase T-1. The resulting oligonucleotides were separated in the first dimension by electrophoresis on cellulose acetate paper at pH 3.5, and in the second dimension by electrophoresis on DEAE-cellulose paper at pH 1.9. The spots were cut out and radioactivity was quantitated by Cerenkov counting to help determine the molar yield of each oligonucleotide. The identity of each oligonucleotide was determined by secondary digestion with pancreatic RNase, and is indicated next to each spot along with a number for its molar yield per transcript.

chromosomal EcoRI fragments (21). However, for fragments above about 10 kilobase pairs in length accurate molecular size determinations are very difficult; size comparisons of cloned fragments with chromosomal fragments are probably reliable only within a few hundred base pairs even when side by

side in the same gel. Therefore, even though we have no evidence that major rearrangements have occurred during the cloning procedure, the large size of many of the cloned EcoRI fragments would prevent us from detecting subtle rearrangements which slightly decrease or increase their size.

In particular, a 17 kilobase pair EcoRI fragment bearing tDNA(Leu 1) has been shown to reside in the E. coli chromosome (21), but the cloned 17 kilobase pair fragment could have been slightly rearranged. This is particularly important in view of the fact that multiple tRNA(Leu 1) genes are present. In fact, in a previous report by Ilgen, et al. (32) it was estimated that there were 4.9 tRNA(Leu 1) molecules per precursor rather than the three suggested by our sequencing data. However, these studies were based on the recovery of a mature tRNA(Leu 1) product following digestion of a 440 base pair precursor with an S30 extract. The measurements could possibly have overestimated the amount of tRNA(Leu 1) released. In addition, if five 87 base pair tRNA(Leu 1) genes were fit into 440 base pairs, this would leave only five base pairs remaining (not enough for the spacer DNA). Also, since we cloned chromosomal EcoRI fragments directly into a $recA^-$ background, this may have prevented $recA^+$-dependent homologous recombination events which might lead to rearrangements. To clear up this problem, however, comparative Southern blot hybridization analysis will have to be performed on chromosomal and cloned fragments that are smaller in size than the 17 kilobase pair EcoRI fragment.

The region upstream of the first tRNA(Leu 1) gene is interesting in that it appears to contain a transcription promoter signal. Our in vitro transcription experiments have defined the site of initiation. The tRNA(Leu 1) promoter behaves similarly to the tRNA(Tyr 1) promoter with respect to in vitro transcription experiments. Maximum promoter-dependent transcription from the Taq I-Acc I tDNA(Leu 1) fragment was seen in the presence of 20% glycerol and 75 mM KCl. Optimal transcription of the tRNA(Tyr 1) promoter was observed at 20% glycerol and 50 mM KCl (36).

As mentioned in the Results section, our DNA sequence data indicates that the -10 and -35 regions upstream from the tRNA(Leu 1) transcription start site are very similar to the consensus promoter sequences for E. coli (35). In addition, the region between the Pribnow box and the start site (5'-CGCCTCC-3') is very G-C rich. This has been found to be the case for all promoters under stringent control which have been sequenced (37). Interestingly, the tRNA(Leu 1) G-C rich sequence is identical with those in the first promoters of the rRNA operons rrnD and rrnX (7), and rrnE (8). It differs from that in

the tRNA(Tyr 1) gene (5'-CGCCCC-3'; ref. 13) only by the insertion of a thymidine residue. Also, the supB-E tRNA operon containing 7 tRNA genes was recently sequenced by Nakajima, et al. (38). No in vitro transcription data was presented, but the sequence between the proposed Pribnow box and transcription start site is identical with that in the tRNA(Tyr 1) gene. Thus, all three tRNA operons have very similar structures in this region. It has been suggested by Travers (37) that a G-C rich sequence at the transcription initiation site would hinder separation of the DNA strands by RNA polymerase. However, the importance of this in the mechanism of the stringent response (or stable RNA gene expression) is still unclear.

Another feature of the DNA sequence upstream of the first tRNA(Leu 1) gene is the presence of a region of dyad symmetry as depicted in Fig. 4. The RNA transcript could possibly form a hairpin in this region consisting of an 8 base pair stem and 3 base loop. Note that 6 of the 8 base pairs are A-T and a run of 5 thymidine residues occurs in the loop. The tRNA(Tyr 1) operon possesses a region of dyad symmetry in the 5' leader region before the first gene (39). Also, all 5 sequenced rRNA promoter regions (7,8,9) have a segment of dyad symmetry just prior to the 5' end of the 16S rRNA gene; it starts 28 bases before the 16S gene and ends at the fifth base of this gene. The supB-E tRNA operon also possesses a region of dyad symmetry just prior to the first gene; it runs from position 5-22 in the published sequence (38). Since none of these regions of dyad symmetry are followed by multiple thymidine residues, they are not similar to known terminators or attenuators of transcription (35). However, they do resemble the nut sites of λ bacteriophage which are stem and loop structures (not followed by multiple thymidine residues) necessary for the action of the anti-termination factor N (40). Morgan recently reported on the possibility that rRNA operons may need a specialized mechanism which prevents premature termination of transcription (polarity) due to the absence of coupled translation (41). The regions of dyad symmetry mentioned above could possibly be the site of action of an E. coli N-like factor which performs this function for both tRNA and rRNA operons.

In Fig. 4 it can be seen that the two spacer regions between the three tRNA(Leu 1) genes are non-identical. The first spacer is 27 base pairs long and exhibits a G+C content of 41%. The second spacer is 34 base pairs long and has a G+C content of 35%. There is nothing particularly striking about these spacer regions except for their low G+C contents as compared to the surrounding tDNA(Leu 1) which is 61% G+C. In addition, the 5' leader sequence

is 33 base pairs long and is 30% G+C; the distal 50 base pairs past the third tRNA(Leu 1) are 32% G+C. The flanking DNA of other sequenced tRNA genes has also been found to be much lower in G+C content than the tRNA sequences (4, 10-12, 38, 42). If all tRNAs are processed by the same set of enzymes, including RNase P (39), then some common features of tRNA precursors should be evident. Some tRNA genes are preceded by hairpin loops as mentioned above, but many are not. In particular, the first gene of the tRNA(Leu 1) operon is preceded by a hairpin structure, but the second and third genes are not. Therefore, this structure would appear to be unimportant for producing a mature 5' terminus on tRNA(Leu 1) molecules assuming the same enzyme performs this function for all three tRNA(Leu 1) species on the precursor. From our sequencing data and that of others (4, 10-12, 38, 42) it appears, for the moment, that the only two things in common for all tRNA precursors are the folded structure of a tRNA, and flanking sequences which are much lower in G+C content than the tRNAs themselves. It is possible that this is all that is necessary for correct tRNA processing.

The region downstream of the third tRNA(Leu 1) gene has not yet been examined for the presence of a transcription termination signal. However, some analogies can be drawn between it and the downstream sequences of the tRNA(Tyr 1) operon. The tRNA(Tyr 1) distal region is comprised of a 178 base pair sequence that begins with the last 19 bases of the tRNA(Tyr 1) mature sequence and is repeated 3.14 times (14). It has been determined that a rho-dependent _in vitro_ site of termination for the tRNA(Tyr 1) operon is located near the end of the sequence 5'-CAATCAAATAT-3 found in the second repeat unit (43). We see no large repeated sequences in the distal region of the tRNA(Leu 1) operon as far out as we have sequenced, but we can locate a sequence 5'-GATTCAAATTC-3' found about 50 base pairs downstream of the last tRNA(Leu 1) gene which is similar to the rho-dependent termination site for tRNA(Tyr 1). In addition, the sequence 5'-GTAGCAAATGT-3' found 100 base pairs downstream of the last tRNA(Leu 1) gene resembles the tRNA(Tyr 1) terminator; this sequence is preceded by a small inverted repeat (see Fig. 4). Interestingly, these sites lie about 410 and 460 base pairs, respectively, downstream of the site of transcription initiation, thus correlating well with the 440 base pair tRNA(Leu 1) precursor isolated by Ilgen, et al. (32). If the precursor they isolated contained the unprocessed 5' and 3' ends, as well as only 3 instead of 5 genes, then one (or both) of these sites may be a transcription termination signal. Studies are underway to define the tRNA(Leu 1) terminator.

In order to study the regulation of the tRNA(Leu 1) operon *in vivo*, the *Taq* I-*Acc* I fragment bearing the promoter has been fused to the structural gene for galactokinase using the plasmid system of McKenney, et al. (44). This work will be published at a later date, but our studies clearly indicate that the DNA fragment contains an *in vivo* transcription promoter signal.

ACKNOWLEDGMENTS

We thank T. Platt for interesting discussions on the tRNA(Leu 1) promoter and M. Fournier for information concerning the cloning of tRNA(Leu 1) genes.

This work was supported by Grant GM-25871 from the National Institutes of Health.

REFERENCES

1. Smith, J.D. (1976). Prog. Nucleic Acid Res. and Mol. Biol. 16: 25-73.
2. Nomura, M., Morgan, E.A. and Jaskunas, S.R. (1977). Ann. Rev. Genet. 11: 297-347.
3. Morgan, E.A., Ikemura, T., Lindahl, L., Fallon, A.M. and Nomura, M. (1978). Cell 13: 335-344.
4. Rossi, J.J. and Landy, A. (1979). Cell 16: 523-534.
5. Ikemura, T. and Ozeki, H. (1977). J. Mol. Biol. 117: 419-446.
6. Gallant, J.A. (1979). Ann. Rev. Genet. 13: 393-415.
7. Young, R.A. and Steitz, J.A. (1979). Cell 17: 225-234.
8. deBoer, H.A., Gilbert, S.F., and Nomura, M. (1979). Cell 17: 201-209.
9. Csordas-Toth, E., Boros, I., and Venetianer, P. (1979). Nucleic Acids Res. 7: 2189-2197.
10. Young, R.A. (1979). J. Biol. Chem. 254: 12725-12731.
11. Duester, G.L. and Holmes, W.M. (1980). Nucleic Acids Res. 8: 3793-3807.
12. Sekiya, T., Mori, M., Takahashi, N., and Nishimura, S. (1980). Nucleic Acids Res. 8: 3809-3827.
13. Sekiya, T., Takeya, T., Contreras, R., Kupper, H., Khorana, H.G., and Landy,A. (1976). in RNA Polymerase, Losick, R. and Chamberlin, M., eds., pp. 455-472, Cold Spring Harbor Lab., Cold Spring Harbor, NY.
14. Egan, J. and Landy, A. (1978). J. Biol. Chem. 253: 3607-3622.
15. Marmur, J. and Doty, P. (1962). J. Mol. Biol. 5: 109-118.
16. Boyer, H.W., Betlach, M., Bolivar, F., Rodriquez, R.L., Heyneker, H.L., Shine, J., and Goodman, H.M. (1977). in Recombinant Molecules: Impact on Science and Society, Beers, R.F. and Bassett, E.G., eds., pp. 9-20, Raven Press, New York.
17. Brown, M.G.M., Weston, A., Saunders, J.R., and Humphreys, G.D. (1979). FEMS Microbiology Letters 5: 219-222.
18. Tabak, H.F. and Flavell, R.A. (1978). Nucleic Acids Res. 5: 2321-2332.
19. Southern, E.M. (1975). J. Mol. Biol. 98: 503-517.
20. Clewell, D.B. (1972). J. Bacteriol. 110: 667-676.
21. Campen, R.K., Duester, G.L., Holmes, W.M., and Young, J.M. (1980). J. Bacteriol. 144: 1083-1093.
22. Bruce, A.G. and Uhlenbeck, O.C. (1978). Nucleic Acids Res. 10: 3666-3677.
23. Szybalski, E.H. and Szybalski, W. (1979). Gene 7: 217-270.
24. Sutcliffe, J.G. (1979). Cold Spring Harbor Symp. Quant. Biol. 43: 77.

25. Smith, H.A. and Birnstiel, M.L. (1976). Nucleic Acids Res. 9: 2387-2398.
26. Klenow, H. and Overgaard-Hansen, K. (1970). FEBS Lett. 6: 25.
27. Maxam, A.M. and Gilbert, W. (1980). Methods Enzymol. 65: 499.
28. Brownlee, G.G. (1972). Determination of Sequences in RNA. T.S. Work and E. Work, eds., North Holland Publishing Company, Amsterdam.
29. Cashel, M., Lazzarini, R.A., and Kalbacher, B. (1969). J. Chromatography 40: 103-109.
30. Gauss, D.H., Gruter, F., and Sprinzl, M. (1979). Nucleic Acids Res. 6: r1-r19.
31. Boros, I., Kiss, A., Venetianer, P. (1979). Nucleic Acids Res. 6: 1817-1830.
32. Ilgen, C., Kirk, L., and Carbon, J. (1976). J. Biol. Chem. 251: 922-929.
33. Dube, S.K., Marcker, K.A., and Yudelevich, A. (1970). FEBS Letters 9: 168-170.
34. Blank, H.-U. and Soll, D. (1971). Biochem. Biophys. Res. Com. 43: 1192-1197.
35. Rosenberg, M. and Court, D. (1979). Ann. Rev. Genet. 13: 319-353.
36. Kupper, H., Contreras, R., Landy, A., and Khorana, G. (1975). Proc. Nat. Acad. Sci. USA. 72: 4754-4758.
37. Travers, A.A. (1980). J. Bacteriol. 141: 973-976.
38. Nakajima, N., Ozeki, H., and Shimura, Y. (1981). Cell 23: 239-249.
39. Altman, S. and Smith, J.D. (1971). Nature New Biol. 233: 35-39.
40. Rosenberg, M., Court, D., Shimatake, H., Brady, C., and Wulff, D. (1978). Nature 272: 414-422.
41. Morgan, E.A. (1980). Cell 21: 257-265.
42. Young, R.A., Mackliss, R., and Steitz, J.A. (1979). J. Biol. Chem. 254: 3264-3271.
43. Kupper, H., Sekiya, T., Rosenberg, M., Egan, J., and Landy, A. (1978). Nature 272: 423-428.
44. McKenney, K., Shimatake, H., Court, D., Schmeissner, U., Brady, C., and Rosenberg, M. (1981). in Gene Amplification and Analysis, Vol. II. J. Chirikjian and T. Papas, eds., in press, (Elsevier-North Holland).