Volume 8 Number 17 1980 Nucleic Acids Research

Sequence of the distal $tRNA_1^{Asp}$ gene and the transcription termination signal in the Escherichia coli ribosomal RNA operon rrnF(or G)

Takao Sekiya, Midori Mori, Nozomi Takahashi and Susumu Nishimura

Biology Division, National Cancer Center Research Institute, Tsukiji 5-1-1, Chuo-ku, Tokyo, Japan

Received 21 May 1980

ABSTRACT

Several DNA fragments carrying tRNA genes have been cloned from *Eco*RI endonuclease digests of *Escherichia coli* DNA. Using cloned DNA, the sequence of the region around the distal gene for $tRNA_1^{Asp}$(*F*(or *G*)) in the *E. coli* ribosomal RNA operon [*rrnF*(or *G*)] has been determined. In the distal portion of *rrnF*(or *G*), the genes for 23S, 5S rRNA and $tRNA_1^{Asp}$(*F*(or *G*)) are located in that order and separated by intergenic spacers of 93 and 52 base pairs, respectively. A possible hairpin structure, with its center between the 22nd and 23rd base pair downstream from the 3'-end of the $tRNA_1^{Asp}$(*F*(or *G*)) gene, followed by a sequence of eight thymidine residues was identified as the transcription termination signal for *rrnF*(or *G*). The termination is rho-independent, at least *in vitro*, and occurs within the region of the contiguous thymidine residues. A possible promoter for a protein gene is present about 50 base pairs downstream from the *rrnF*(or *G*) terminator.

INTRODUCTION

The primary sequences of more than 30 *Escherichia coli* tRNA species have been determined (1) and the genes for many tRNAs have been mapped on the *E. coli* chromosome (2 and H.Ozeki personal communication). The amount of tRNA molecules in *E. coli* cells is known to differ significantly in different species (2). Since the amount of tRNA molecules in the cells does not correspond to the number of genes in some tRNA species, as for example in $tRNA_1^{Tyr}$ and $tRNA_2^{Tyr}$ (3-5), the transcription signals, promoters and terminators, must play an important role in tRNA gene expression. Processing of tRNA precursors may also be involved in regulation of the amount of tRNA molecules. Determinations of the DNA sequences of many tRNA genes, including their promoters, terminators and intergenic spacers, should provide information for understanding the detailed mechanism of tRNA gene expression in *E. coli*.

The genes for tRNA so far analyzed can be divided into two groups. In group 1, one or more tRNA genes constitute a tRNA operon. Promoter-dependent transcription of the operon produces a precursor RNA that can be processed to the mature tRNA(s). In this group, the structure of the gene for $tRNA_1^{Tyr}$,

including the signals for gene expression, has been most extensively characterized (6-12). The genes for tRNA in group 2 belong to the seven ribosomal RNA operons (*rrn*) in *E. coli* as genes for spacer or distal tRNAs (13,14). The DNA sequences of the regions around the genes for spacer tRNAs, $tRNA_1^{Ile}$ (15, 16) $tRNA_{1B}^{Ala}$ (16) and $tRNA_2^{Glu}$ (17), have been determined. Furthermore, recently Young has sequenced the distal end region of *rrnC* containing the genes for both $tRNA_1^{Asp}$(*C*) and $tRNA^{Trp}$ (18).

We have cloned and isolated several *E. coli* DNA fragments carrying tRNA genes. Using one of the cloned DNAs carrying tRNA genes of group 2, the DNA sequence of the region around the distal $tRNA_1^{Asp}$(*F*(or *G*)) gene in *rrnF*(or *G*) was elucidated. Here we report evidence that the gene for $tRNA_1^{Asp}$(*F*(or *G*)) follows the 5S rRNA gene in *rrnF*(or *G*)) and that the transcription termination signal for the operon is located shortly after the tRNA gene. Furthermore, a sequence that seems to be a promoter for an adjacent protein gene in the region downstream from the *rrn* operon is described.

MATERIALS AND METHODS

Materials. Restriction endonucleases were purchased from Bøehringer Mannheim Biochemicals or Bethesda Research Laboratories. T_4-DNA ligase (19), RNA polymerase holoenzyme (20) and rho-factor (21) were prepared according to published procedures. RNase T_1 was from Sankyo Co. T_4-Polynucleotide kinase was from Boehringer Mannheim Biochemicals and bacterial alkaline phosphatase (BAPF) was from Worthington. [α-^{32}P]Ribonucleoside-5'-triphosphates were prepared as described by Reeve and Huang (22) using ribonucleoside [5'-^{32}P]monophosphates obtained by nuclease P1 (Yamasa Shoyu Co.) digestion of total *E. coli* RNA labeled *in vivo*. The preparation had a specific activity of 1 Ci/mmol. Agarose for gels was Seakem brand, and acrylamide and bisacrylamide were from Eastman.

Restriction Endonuclease Digestions and Fractionation of DNA. Digestions of DNA with restriction endonucleases were performed under the conditions specified by the enzyme suppliers. After digestion, DNA fragments were fractionated by electrophoresis on 1 % agarose or 5 to 10 % polyacrylamide gels. Agarose gels were used for separation of DNA fragments bigger than 2000 bp ; electrophoresis being carried out at 15 mA for 16 hrs in 90 mM Tris-borate (pH 8.3)-4 mM EDTA. DNA fragments smaller than 2000 bp and 500 bp were analyzed by electrophoresis on 5 % and 10 % polyacrylamide gels, respectively ; the ratio of acrylamide to bisacrylamide was 29 to 1 and electrophoresis was performed in 45 mM Tris-borate (pH 8.3)-2 mM EDTA.

Recovery of DNA Fragments from Gels. The bands of DNA in gels were located by staining with ethidium bromide (1 μg/ml) or by autoradiography. DNA was recovered from agarose gels by the method of Tabak and Flavell (23). DNA in bands excised from acrylamide gels was eluted by the method of Maxam and Gilbert (24).

Cloning of E. coli DNA Fragments and Detection of Plasmids Carrying tRNA Genes. *E. coli* C600 DNA (500 μg) was digested with 500 units of restriction endonuclease *Eco*RI in 500 μl of reaction mixture at 37° for 16 hrs.

After digestion, the mixture was applied to a wide well (200 x 2 x 5 mm) of a preparative 1 % agarose gel (250 x 170 x 5 mm) and the DNA fragments were fractionated by electrophoresis. A portion of the gel (20 x 170 x 5 mm) was excised and subjected to hybridization as described by Southern (25). When hybridization was performed using total *E. coli* tRNA labeled with [^{32}P] phosphate *in vivo* as a probe, thirteen distinct radioactive bands were observed on the autoradiogram. DNA fragments were eluted from the regions corresponding to these bands. Samples of 2 μg of these fractionated DNA fragments were ligated to the *Eco*RI site of RSF2124 (26) (4 μg) in a solution (100 μl) of 50 mM Tris-HCl (pH 7.5)-5 mM $MgCl_2$-10 mM DTT-100 μM ATP using 40 units of T_4-DNA ligase at 4° for 16 hrs. Transformation of *E. coli* C600 (rK^-mK^-) by the ligated DNA was performed as described by Enea *et al.* (27).

Ampicillin resistant and colicin E1 non-producing transformants were collected. Each transformant was cultured in 10 ml of Luria broth and cleared lysate was prepared by chloramphenicol amplification as described by Clewell and Helinski (28). Proteins and RNA were removed from the lysate by extraction with an equal volume of 1 x SSC saturated phenol and by treatment with RNase A (20 μg/ml) at 37° for 1 hr. Then the crude plasmids were precipitated by addition of 2.5 volumes of ethanol, dissolved in 50 μl of 10 mM Tris-HCl (pH 7.5)-0.1 mM EDTA and subjected to electrophoresis on 1 % agarose gels. The region of the gel containing plasmids was excised and subjected to Southern hybridization (25) using [^{32}P] labeled total *E. coli* tRNA as a probe. The plasmids carrying the sequence complimentary to tRNA were transferred to *E. coli* HB101 (*hsm*$^-$,*hrs*$^-$,*rec*A$^-$,*gal*$^-$,*pro*$^-$,*str*R) and maintained in Luria broth containing 50 % glycerol.

Labeling of DNA fragments and DNA Sequencing. The phosphate groups at the 5'-ends of DNA fragments generated by restriction endonucleases were removed as described previously (15). The fragments were labeled with [^{32}P] phosphate at the 5'-ends as described by Maxam and Gilbert (24), except that

the T_4-polynucleotide kinase reaction was carried out at pH 7.5. DNA fragments were sequenced using the chemical modification procedure of Maxam and Gilbert (24).

In vitro transcription of the cloned DNA. Transcription and processing of transcripts were performed as described previously (29). [α-^{32}P]UTP and [α-^{32}P]CTP were used as labeled ribonucleoside-5'-triphosphates. Reactions were stopped by heating at 90° for 3 min. Denatured proteins, if present, were removed by centrifugation, and then the solution was mixed with one-third volume of 50 % glycerol-0.1 % xylene cyanol-0.1 % bromphenol blue and subjected to two-dimensional electrophoresis using 10 % polyacrylamide gel in the first dimension and 20 % polyacrylamide gel in the second dimension. Fingerprinting of the RNA products was performed as described previously (29).

RESULTS

Plasmids carrying the genes for tRNA. No sites for restriction endonuclease *Eco*RI are present in the known nucleotide sequences of *E. coli* tRNA species. Therefore, by using *Eco*RI endonuclease, most of the tRNA genes should be cleaved from the chromosome without damaging the structural genes. At least 13 distinct DNA bands complementary to tRNA or 5S rRNA were obtained when *Eco*RI fragments were separated by electrophoresis on 1 % agarose gel and then subjected to the Southern hybridization procedure using total *E. coli* tRNA labeled with [^{32}P]phosphate *in vivo* as a probe. Unfractionated *Eco*RI fragments (Table 1, Exp.1) and these fractionated as described above (Table 1, Exp.2) were cloned using plasmid RSF2124 (26). The nine DNA fragments so far obtained are shown in Table 1. Among the DNA fragments thus cloned, only two DNAs contained the tRNA gene of group 1. All the other cloned fragments in Table 1 contained group 2 genes, which belong to *rrn* operons as genes for spacer or distal tRNAs. Assignments of these DNAs carrying the genes for tRNA or 5S rRNA were performed by comparison with the known chain lengths of *Eco*RI fragments generated from *rrn* operons (13,14,32). The DNA sequence of the region around the $tRNA_1^{Ile}$ gene in pTS228 was reported previously (15).

When pMM10-19 was digested with *Eco*RI, the fragment of 2.7 kb was generated. Since this fragment carried the gene for $tRNA_1^{Asp}$ (*F*(or *G*)) (Table 1), it was designated as $tDNA_1^{Asp}$. In *E. coli*, the genes for $tRNA_1^{Asp}$ thus far known are located in the *rrnC* and the *rrn* (Group I) (*rrnF*(or *G*), see the legend to Table 1) operons as distal genes (13,14). The *rrnF*(or *G*) operon contains only the $tRNA_1^{Asp}$(*F*(or *G*)) gene (13), while *rrnC* has the genes for both $tRNA_1^{Asp}$(*C*) and $tRNA^{Trp}$ (14). The cloned $tDNA_1^{Asp}$ was compared with the

Recombinant DNA		Size of DNA cloned (kb)	tRNA or 5S RNA Hybridized	Remarks on tRNA genes
Experiment 1	Experiment 2			
pYK134		6	X	tRNA operon
	pMM1,2-23	10	Asn	tRNA operon
pTS228	pMM13-6	2	Ala1,Ile1	spacer tRNAs in *rrnA*(85') or *D*(71')
pYK206	pMM8-23	4	Ala1,Ile1,5S	spacer tRNAs in *rrnF*(or *G*)
pYK50	pMM13-18	2	Glu2	spacer tRNA in *rrnB*(88'), *C*(83'), *E*(89') or *rrnG*(or *F*)
pYK137		8	Asp1,Trp	distal tRNAs in *rrnC*(83')
	pMM10-19	3	Asp1	distal tRNA in *rrnF*(or *G*)
pYK24		9	5S	*rrnA*(85')
	pMM11-3	2	5S	*rrnC*(83')

Table 1. *Eco*RI fragments of *E. coli* DNA carrying the genes for tRNA or 5S rRNA. In experiment 1, total *Eco*RI digests of *E. coli* DNA were ligated to RSF2124 (for pYK Strains) or pMB9 [for pTS228 (15)]. In experiment 2, *Eco*RI fragments obtained from *E. coli* DNA were first fractionated by electrophoresis on 1 % agarose gel. The DNA fragments carrying tRNA genes were located by blotting hybridization as described by Southern (25) and eluted from the gel. The DNA fragments thus fractionated were subjected to molecular cloning using the RSF2124 plasmid (26) as a vector. The tRNA species of the gene in the cloned DNA was identified by hybridization with [^{32}P]tRNA fractionated by two dimensional polyacrylamide gel electrophoresis (30). The chain lengths of the cloned DNAs were estimated by comparison of their mobilities on agarose gel electrophoresis with those of λDNA treated with both *Eco*RI and *Hin*dIII (31). The *rrn* (Group I) operon reported by Morgan *et al.* (13), corresponds to one of two *rrn* operons, *rrnF* and *rrnG* (14,32). The restriction maps for *rrnF*(or *G*) and *rrnG*(or *F*) have been reported by Boros *et al.* (32) and Bram *et al.* (33). Presence of an *Eco*RI site in the distal region suggests that *rrnF*(or *G*) may correspond to the *rrn* (Group I) operon. According to the terminology of Boros *et al.* (32), *rrnF*(or *G*) is used instead of *rrn* (Group I) in this paper.

plasmid pLC7-21, which is known to carry *rrnF*(or *G*) (13). As shown in Fig.1a, digestion of pLC7-21 with *Eco*RI generated fragment B (2.7 kb), with a chain length corresponding to that of $tDNA_1^{Asp}$, besides the plasmid vector and fragment A of 4.3 kb. The cloned fragment $tDNA_1^{Asp}$ and fragment B were further cleaved with *Hae*III, *Hin*dII and a mixture of *Hae*III and *Hin*dII, and the digests were analyzed by 10 % polyacrylamide gel electrophoresis. The results shown in Fig. 1b clearly indicate that $tDNA_1^{Asp}$ and fragment B were identical. It is therefore concluded that $tDNA_1^{Asp}$ is a distal portion of the *rrnF*(or *G*) operon as indicated in Fig. 2. Fig. 2 shows the expected structure of *rrnF*(or *G*)

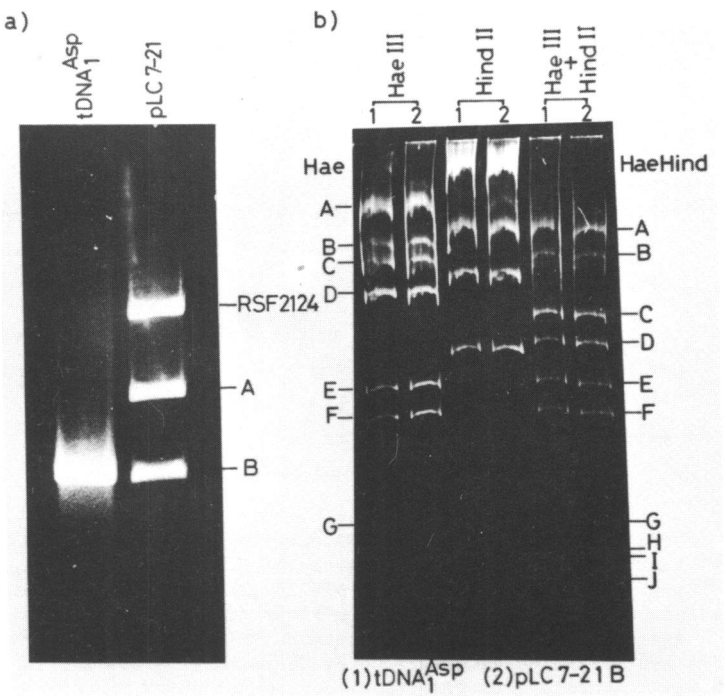

Fig. 1. Comparison of $tDNA_1^{Asp}$ with fragment B obtained by *Eco*RI digestion of pLC7-21. Two µg of pLC7-21 was digested with 5 units of *Eco*RI. The digests obtained and the cloned $tDNA_1^{Asp}$ fragment were subjected to electrophoresis side by side on 1 % agarose gel (a). Fragment B (1 µg) obtained in (a), which has the same mobility as $tDNA_1^{Asp}$, and $tDNA_1^{Asp}$ (1 µg) were digested with 2 units of *Hae*III, *Hin*dII and a mixture of *Hae*III and *Hin*dII. The sub-fragments obtained from the two DNAs were compared by electrophoresis on 10 % polyacrylamide gel (b).

and also includes the results described below.

Since the gene for $tRNA_1^{Asp}$(*F*(or *G*)) is located in the distal portion of *rrnF*(or *G*) (13), $tDNA_1^{Asp}$ must contain the region of transcription termination of the operon. Being interested in the structure of the transcription termination signal, we determined the DNA sequence of the region around the gene for $tRNA_1^{Asp}$(*F*(or *G*)) using $tDNA_1^{Asp}$. We also sequenced a terminal portion of fragment A, where the genes for 23S and 5S rRNA are located.

Restriction map of the distal region of rrnF(or G). For determination of its terminus adjacent to $tDNA_1^{Asp}$, the *Eco*RI fragment A of 4.3 kb obtained from pLC7-21 (Fig.1a and Fig.2) was labeled with [^{32}P]phosphate at the 5'-ends

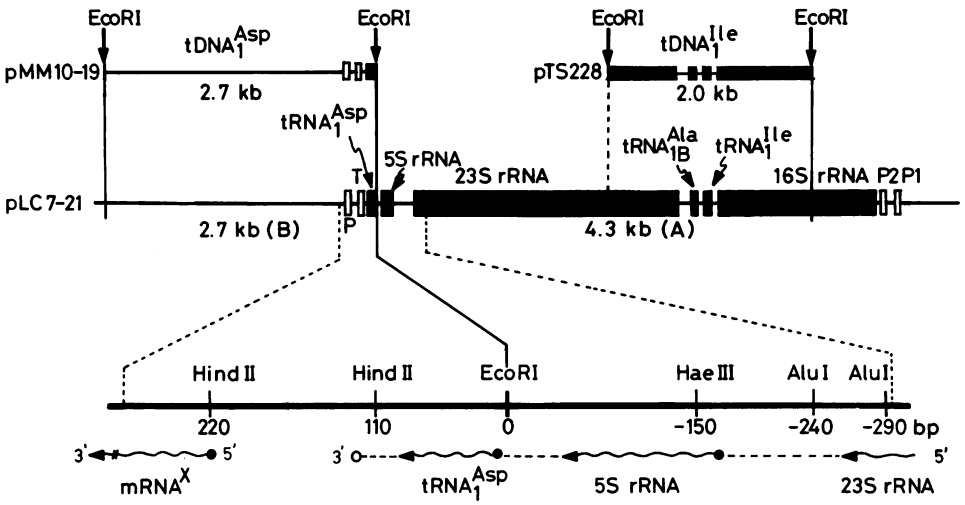

Fig. 2. Structure of the *rrnF*(or *G*) operon. Top, two cloned DNA fragments, $tDNA_1^{Ile}$ carried by pTS228 (15) and $tDNA_1^{Asp}$ in pMM 10-19 are shown. Middle, the region of *rrnF*(or *G*) carried by pLC7-21 (13) is illustrated with the positions of the constituent genes. The positions of the promoters for *rrnF* (or *G*) (p1 and p2) are deduced by analogy to those for *rrnE*, *rrnA* (34), *rrnD*, *rrnX* (35) and *rrnB* (36). The positions of the transcription termination signal for *rrnF*(or *G*) and the promoter for the assumed gene adjacent to *rrnF* (or *G*) are also illustrated. Bottom, a restriction map of the region joining fragment A (4.3 kb) and $tDNA_1^{Asp}$ (*i.e.* fragment B, 2.7 kb) is shown.

using T_4-polynucleotide kinase and [γ-^{32}P]ATP. The labeled fragment A was digested with *Alu*I and the digests were compared by electrophoresis on 10 % polyacrylamide gel with those obtained from $tDNA_1^{Ile}$, which carried the 3'-end region of the 16S rRNA gene, the genes for $tRNA_1^{Ile}$-$tRNA_{1B}^{Ala}$ and the 5'-end region of the 23S rRNA gene (15 and Fig.2). Among the 19 DNA bands obtained from fragment A, a subfragment of 240 bp (*Alu* D ; subfragments were designated alphabetically according to their chain lengths) was radioactive and not present in the digests of $tDNA_1^{Ile}$. Therefore, the labeled end of *Alu* D must be adjacent to $tDNA_1^{Asp}$. Partial *Alu* I digestion of the [5'-^{32}P]*Hin*fI fragment of 380 bp obtained similarly from the same labeled end of fragment A revealed that a small *Alu* fragment of 47 bp (*Alu* R) was adjacent to *Alu* D. The results are illustrated at the bottom of Fig. 2(0 to -290 bp).

The cleavage sites of $tDNA_1^{Asp}$ with several restriction endonucleases are summarized in Fig. 3. Digestion of $tDNA_1^{Asp}$ with *Hae*III generated seven fragments (*Hae* A to *Hae* G in Figs.1 and 3a). When the digests were separated by

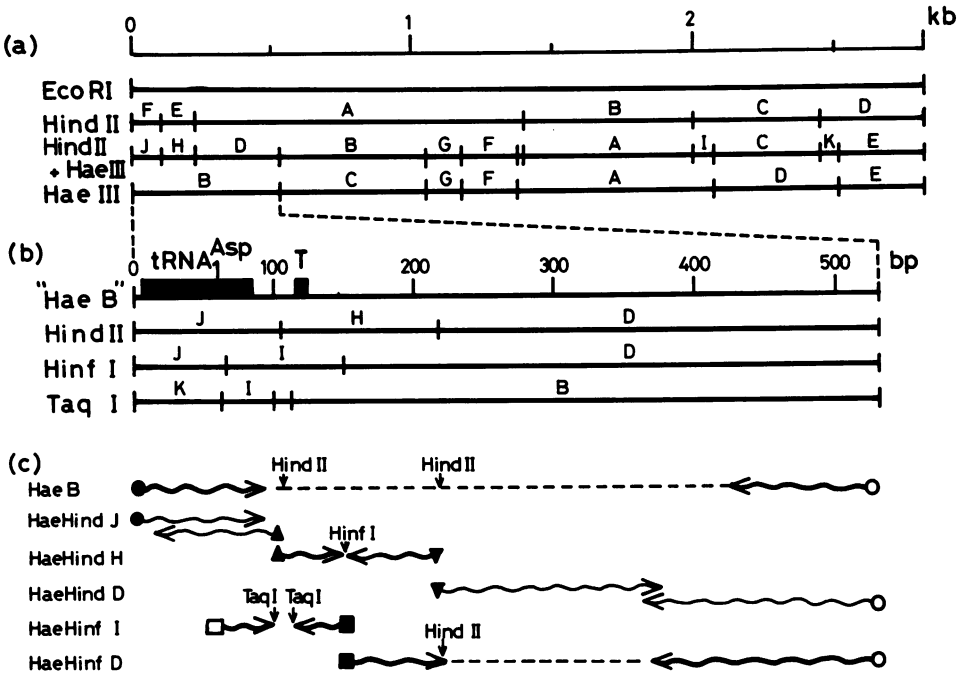

Fig. 3. Restriction map of the cloned *E. coli* DNA fragment (tDNA$_1^{Asp}$). (a) Cleavage sites on the DNA fragment (tDNA$_1^{Asp}$) were mapped for *Hin*dII, *Hae*III and a mixture of *Hin*dII and *Hae*III. The fragment B obtained by *Hae*III digestion (*Hae* B) was shown to carry the gene for tRNA$_1^{Asp}$(*F*(or *G*)). (b) Cleavage sites on the *Hae* B fragment were mapped for *Hin*dII, *Hin*fI and *Taq*I. The position of the gene for tRNA$_1^{Asp}$(*F*(or *G*)) and its transcription termination signal (T) are also indicated. (c) The DNA fragments used for nucleotide sequence determination are illustrated. All the double stranded DNAs were first labeled with [^{32}P]phosphate at their 5'-termini using T4-polynucleotide kinase and [γ-^{32}P]ATP. The labeled DNAs were then subjected to either strand separation or further digestion with appropriate restriction endonucleases.

electrophoresis on 2 % agarose gel and then subjected to blotting hybridization by the method of Southern (25) using total *E. coli* tRNA labeled with [^{32}P]phosphate *in vivo* as a probe, only the fragment *Hae* B could be detected as a labeled DNA on the autoradiogram (data not shown). Fig. 3b shows a cleavage map of fragment *Hae* B thus identified to carry the tRNA$_1^{Asp}$(*F*(or *G*)) gene, and also the position of the gene.

DNA sequencing. The DNA sequence of the region containing the 3'-end of the 23S rRNA gene and the gene for 5S rRNA was elucidated using the *Alu* D and *Alu* R fragments obtained from *Eco*RI fragment A of pLC7-21. The separated

single strands were examined by the method of Maxam and Gilbert (24). The region around the gene for $tRNA_1^{Asp}$ (*F*(or *G*)) was sequenced using the fragments generated from $tDNA_1^{Asp}$ by digestion with both *Hae*III and *Hin*dII (*HaeHin*d D, H and J, see Fig.1b and 3b) or with both *Hae*III and *Hin*dII(*HaeHin*f I, J and D, Fig.3b). The method for DNA sequencing of $tDNA_1^{Asp}$ is illustrated in Fig. 3c. The following fragments were obtained ; i) two labeled terminal fragments produced by *Hin*dII digestion of [5'-^{32}P] *Hae* B, ii) separated strands of the [5'-^{32}P] labeled *HaeHin*d J fragment, iii) two labeled fragments obtained from [5'-^{32}P] *HaeHin*d H by *Hin*fI digestion, iv) two terminal fragments generated by *Taq*I digestion of [5'-^{32}P] *HaeHin*f I and v) two labeled fragments obtained from [5'-^{32}P] *HaeHin*f D by *Hin*dII digestion. Sequence determination of these fragments was also performed according to the procedure of Maxam and Gilbert (24). The DNA sequences thus determined for *Alu* D and R of fragment A (positions 1-292) and for the *HaeHin*d J-H region and a portion of *HaeHin*d D (positions 293-600) are shown in Fig. 4. A remarkable structure of two-fold rotational symmetry followed by eight T-A base pairs (positions 384-417) is present in the region beyond the gene for $tRNA_1^{Asp}$ (*F*(or *G*)) (positions 301-377) and this seems to be a transcription termination signal.

In vitro transcription of the EcoRI fragment, $tDNA_1^{Asp}$. To identify the presumed terminator, $tDNA_1^{Asp}$ was transcribed with *E. coli* RNA polymerase and the products were analysed. As shown in Fig. 5a, six transcripts of distinct sizes were separated from other heterogeneous products by two-dimensional electrophoresis on polyacrylamide gel. Of these six transcripts, transcript 2, containing about 120 nucleotides, was the main product. A significant amount of transcript 5, which contained less than 80 nucleotides, but a few more than $tRNA_1^{Asp}$, was also produced. As shown in Fig. 5b, the sizes of the major product 2 and of other transcripts were not affected by addition of the rho-factor.

The fingerprints of RNase T_1-oligonucleotides obtained from the primary transcripts 1 to 6 revealed that all these transcripts except transcript 1 contained the nucleotide sequence of $tRNA_1^{Asp}$. The region of $tDNA_1^{Asp}$ giving transcript 1 is not identified yet. The fingerprint obtained from the major transcript 2 is shown in Fig. 6a. In fact, the primary transcripts could be processed to $tRNA_1^{Asp}$ as shown in Fig. 5c (spot 3) and Fig. 6b when the transcription reaction mixture was treated with the S30 fraction of *E. coli* extracts. Fingerprint analysis of the major transcript 2, shown in Fig. 6a, also revealed that besides the T_1-oligonucleotides from the sequence of $tRNA_1^{Asp}$, those expected from the region beyond the 3'-end of tRNA were

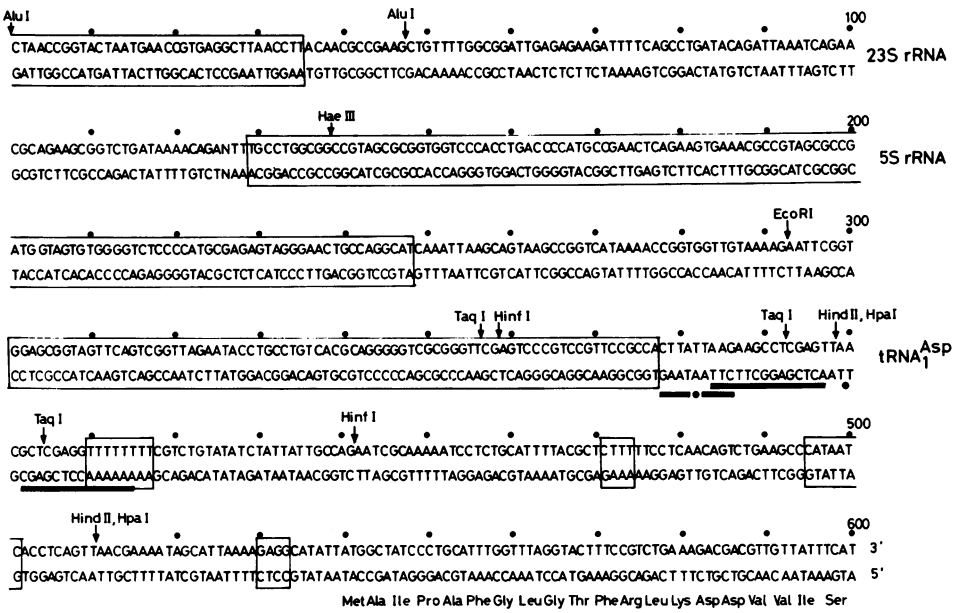

Fig. 4. DNA sequence in the region around the distal gene for $tRNA_1^{Asp}$ (F(or G)) in the *rrnF*(or *G*) operon. The DNA sequence from position 1 to the *Eco*RI site (Position 292) is from the fragment A obtained from pLC7-21 (Figs.1 and 2). The nucleotide sequence from the *Eco*RI site (position 292) to position 600 is from one end of $tDNA_1^{Asp}$. The genes for 23S and 5S rRNA and $tRNA_1^{Asp}$(F(or G)) are indicated by the solid boxes. Regions of two-fold rotational symmetry are underlined and dots indicate centers of symmetry. The region of eight thymidine residues, 32 base pairs downstream from the CCA-end of the tRNA sequence, is also shown by a box. The boxes at positions 471-474, 495-501 and 530-533 are expected signals for RNA polymerase recognition, binding and ribosome binding, respectively. The assumed sequence of amino acids beginning at the ATG codon (position 540) is shown under the DNA sequence.

present. The presence of uridine-rich oligonucleotides (p1, p2 and p3 in Fig. 6a) strongly suggests that the transcription terminated within the region of eight uridine residues. Transcript 3 which is few nucleotides shorter than transcript 2 contains all RNase T_1-oligonucleotides found in transcript 2 except p1 and p3. It is therefore concluded that the major transcript 2 of about 120 nucleotides and also minor transcript 3 are promoter-independent products of RNA polymerase that starts transcription of $tDNA_1^{Asp}$ from the *Eco*R1 terminus and terminates at the proposed signal. Transcript 5 does not contain precursor specific oligonucleotides (p1, p2, p3 and p5) except some amount of p4. Therefore, it might be a product of

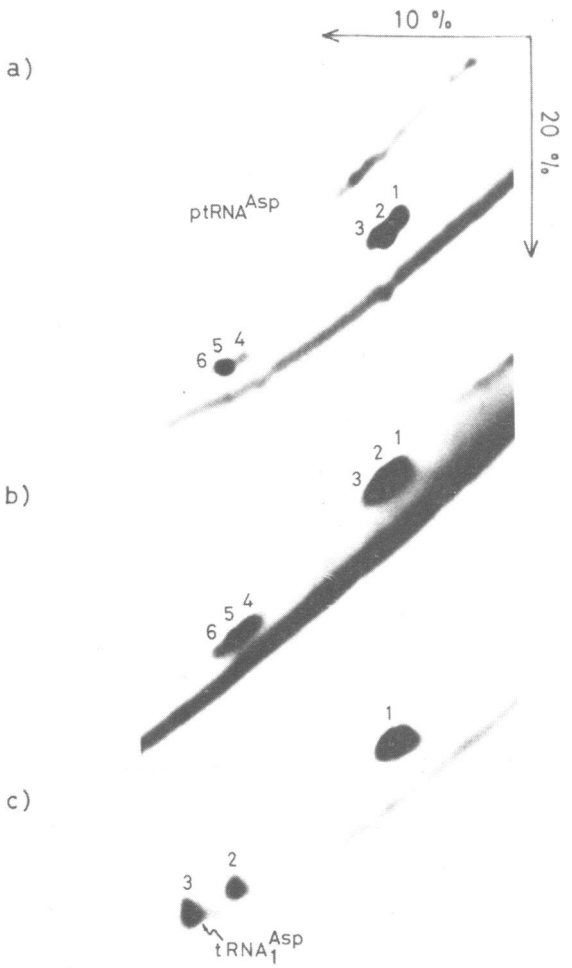

Fig. 5. *In vitro* transcription of $tDNA_1^{Asp}$. *In vitro* Transcription was performed as described under Materials and Methods. The products of the transcription reaction were subjected to two-dimensional electrophoresis on 10 % polyacrylamide gel in the first dimension and 20 % polyacrylamide gel in the second dimension. a) and b), Transcripts from $tDNA_1^{Asp}$ in the absence and presence of rho-factor, respectively; c), the primary transcripts processed with the S30 fraction of an *E. coli* cell extract. Spot 3 contains $tRNA_1^{Asp}$. Since the processing was not complete, spots 1 and 2 corresponding to the major primary transcripts 2 and 5, respectively, are also present.

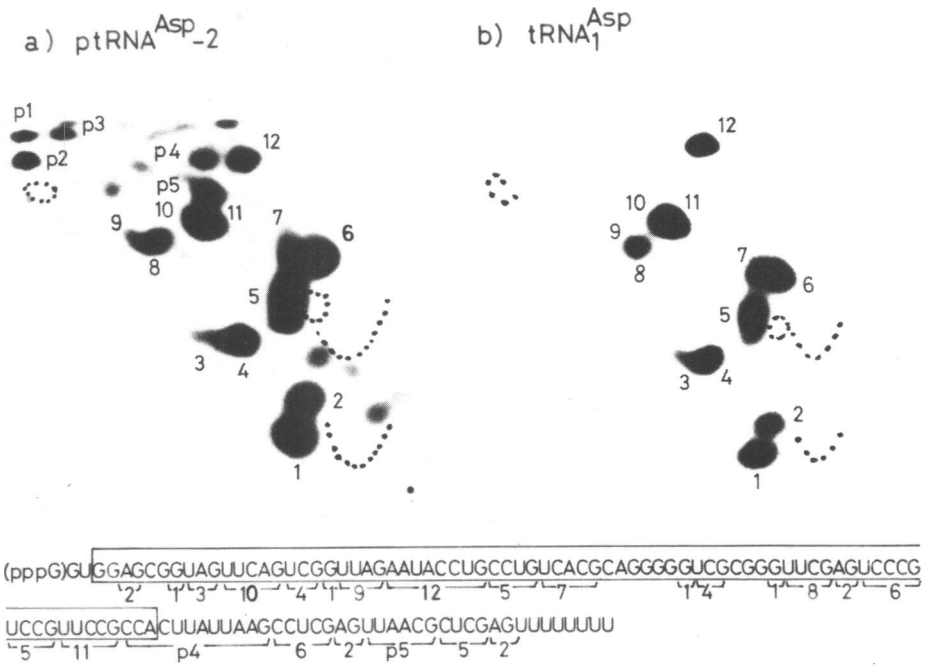

Fig. 6. Fingerprints of the products obtained by RNase T_1 digestion of the primary transcript and of the processed product. The primary transcript (ptRNA^{Asp}-2 in Fig. 5a), and the processed product (spot 3 in Fig. 5c) were prepared using [α-^{32}P]UTP and [α-^{32}P]CTP as labeled ribonucleoside triphosphates and were digested with RNase T_1. The fingerprints for the primary and processed transcripts are shown in (a) and (b), respectively. The oligonucleotides expected from the primary transcript and from the processed transcript are shown systematically under the fingerprints. The sequence corresponding to tRNA$_1^{Asp}$(*F*(or *G*)) is shown by a box.

RNA polymerase which terminates transcription somehow in the region few nucleotides beyond the 3'-end of the tRNA gene.

DISCUSSION

By molecular cloning of restriction endonuclease *Eco*R1 fragments of *E. coli* DNA, we were able to obtain almost all of tRNA genes belonging to the *rrn* operons, but only two genes composing the tRNA operons have been cloned thus far. Four genes for tRNA$_2^{Glu}$, three for tRNA$_1^{Ile}$ and tRNA$_{1B}^{Ala}$, two for tRNA$_1^{Asp}$ and one for tRNA^{Trp} and tRNA^{Thr} are present in seven *E. coli rrn* operons (13,14,17). The copies of these tRNA genes are not very different

from those of other genes composing the tRNA operons. Therefore, it is very strange that the genes belonging to the tRNA operons have been cloned less efficiently than those in the *rrn* operons. The reason for this is unknown.

Structure of the distal region of rrnF(or G). Using one of the DNA fragments thus cloned, we determined the nucleotide sequence of the gene for $tRNA_1^{Asp}$(*F*(or *G*)) and the surrounding region. The genes for $tRNA_1^{Asp}$ so far known are located in *rrn* operons as distal genes. Morgan *et al.* reported that the *rrnF*(or *G*) operon contains only the $tRNA_1^{Asp}$(*F*(or *G*)) gene while *rrnC* contains the genes for both $tRNA_1^{Asp}$(*C*) and $tRNA^{Trp}$ (13,14,17). Since we have also cloned the DNA of 8 kb containing the genes for both $tRNA_1^{Asp}$(*C*) and $tRNA^{Trp}$ (Table 1), the DNA fragment of 2.7 kb carrying only the $tRNA_1^{Asp}$ gene that we analyzed in this work seemed to correspond to a portion of the *rrn F*(or *G*) operon. This possibility was confirmed by analysing the DNA that was cloned independently and known to contain the *rrnF*(or *G*) operon (13).

The sequence of 35 base pairs determined for the 3'-end of the 23S rRNA gene in *rrnF*(or *G*) (positions 1-35 in Fig.4) corresponds exactly to those recently reported for *rrnX* and *rrnD* (33) and *rrnB* (37). The spacer between the 23S and 5S rRNA genes in *rrnF*(or *G*) is 93 base pairs (positions 46-128), being one base pair longer than that of *rrnX* and *rrnD* reported by Bram *et al.* (33). The nucleotide sequence of the intergenic spacer is identical to that of *rrnX* and *rrnD* except for following positions (Fig.4). The C-G pair at position 48 in the *Alu*I recognition site is A-T in *rrnX* and G-C in *rrnD*. No T-A pair at position 62 is present in *rrnX* and *D*. This residue makes it possible to add an A-U pair in the stem region of the secondary structure proposed for the flanking regions of the 23S rRNA precursor (33). The A-T pair at position 118 is identical to that in *rrnD*, but is replaced to G-C in *rrnX*. The nucleotide sequence of the corresponding region of *rrnB* has also been sequenced (J. Brosius, personal communication) and is identical to that of *rrn F*(or *G*), except for the base pair at position 62 discussed above.

The DNA seuqence of the 5S rRNA gene in *rrnF*(or *G*) is identical to those sequenced for *rrnC* (18) and *rrnB* (J. Brosius, personal communication) except that C-G at position 140 is A-T in *rrnB*. The first 27 base pairs of 5S rRNA genes in *rrnX* and *rrnD* have also been sequenced (33) and differ from that of *rrnF*(or *G*) at the heterogeneous position described above.

The spacer DNA between the genes for 5S rRNA and $tRNA_1^{Asp}$(*F*(or *G*)) is 52 base pairs (positions 248-300) and its nucleotide sequence is identical to that determined for *rrnC* carrying genes for both $tRNA_1^{Asp}$(*C*) and $tRNA^{Trp}$ at the end of the operon (18). In this region, it is noteworthy that the hepta-

nucleotide sequence GTAAAA is present 10 nucleotides upstream from the 5'-end of the $tRNA_1^{Asp}$(*F*(or *G*)) gene. Although their biological meanings are unknown, exactly the same sequences are also observed in the corresponding regions of the $tRNA_1^{Tyr}$ gene and the gene for the spacer $tRNA_1^{Ile}$ in *rrn* operons (15). The DNA sequence of the $tRNA_1^{Asp}$(*F*(or *G*)) gene (positions 310-377) corresponds exactly to that for the tRNA, determined previously by the RNA sequencing method (38).

Following the $tRNA_1^{Asp}$(*F*(or *G*)) gene, a small structure of two-fold rotational symmetry is present (positions 378-386). The center of symmetry is located at the 5th base pair from the 3'-end of the gene. The regions beyond the 3'-end in several *E. coli* tRNA genes have been sequenced. Some of them carry a similar structure of dyad symmetry immediately after the end of the gene (15-17,29,39,40). The biological significance of this structure, if any, is not clear, but it has been suggested that the structure found in the $tRNA_1^{Tyr}$ gene may be involved in processing the 3'-end of the tRNA (29).

Transcription termination. Another structure of two-fold rotational symmetry (positions 384-415) is remarkable in the region adjacent to the first small symmetry described above. The center of symmetry is located between the 22nd and 23rd base pair downstream from the 3'-end of the $tRNA_1^{Asp}$ (*F*(or *G*)) gene. Fig. 7a shows the possible secondary structure of the $tRNA_1^{Asp}$ (*F*(or *G*)) gene and the surrounding region in an RNA form. A hairpin structure composed of eight base pairs in the stem and four nucleotides in the loop is followed by eight contiguous U residues. Several transcription termination signals for *E. coli* RNA polymerase have been sequenced (for review, see ref. 41). Rho-independent signals are known to have the common features in their primary sequences that the region of a stem-loop structure in the RNA is near the 3'-end and termination occurs within a run of uridine residues (41). On the other hand, in rho-dependent terminations, a common sequence of CAATCA proximal to the termination point seems to be recognized (10,42,43). The stem-loop structure followed by U residues shown in Fig. 7a shares the common features of those in known rho-independent terminators. The assumed terminator was evident when the *Eco*RI fragment, $tDNA_1^{Asp}$, was transcribed *in vitro* with *E. coli* RNA polymerase. Two of the promoter-independent transcripts starting from the *Eco*RI end of $tDNA_1^{Asp}$ terminated within the contiguous U residues.

Since the DNA spacer of 52 base pairs between the genes for 5S rRNA and $tRNA_1^{Asp}$(*F*(or *G*) seems to be too short to contain both the transcription

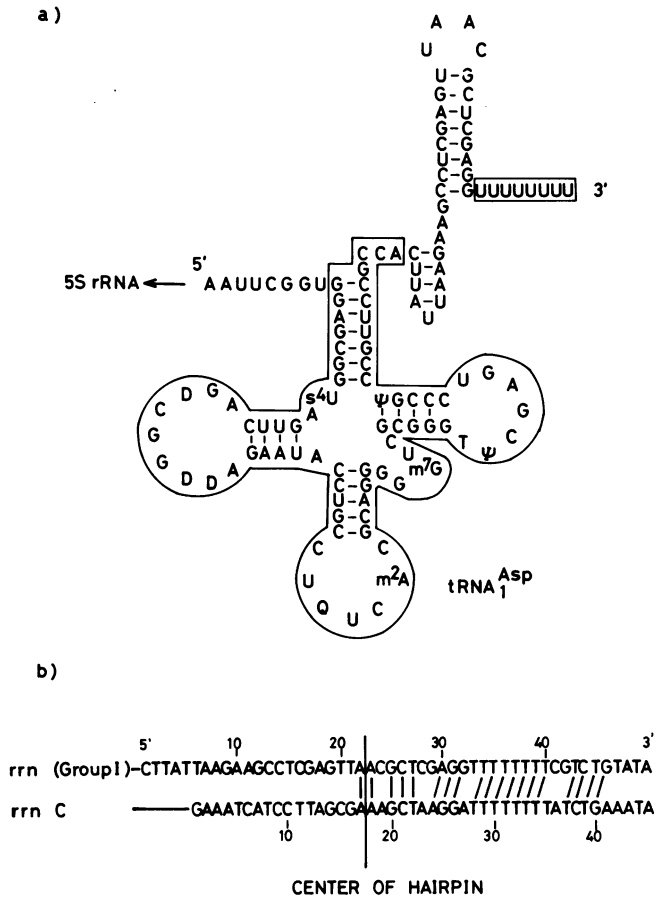

Fig. 7. A possible secondary structure of the precursor for distal $tRNA_1^{Asp}$ from *rrnF*(or *G*) and comparison of the *rrn* terminators. (a), The nucleotide sequence of the top strand of the region around the gene for $tRNA_1^{Asp}$(*F*(or *G*)) in Fig. 4 is shown in RNA form. The sequence corresponding to $tRNA_1^{Asp}$, shown in the fully modified form, and the possible transcription termination site composed of eight uridine residues are enclosed by continuous lines. (b), The DNA sequences of the transcription termination signals in *rrnF*(or *G*) and *rrnC* (15) in the antisense strands are compared. The vertical line shows the centers of the hairpin structures. Sequence similarities are indicated by short lines. The nucleotides are numbered from the adjacent nucleotide to the end of the distal gene.

termination signal for the 5S rRNA gene and the promoter for the $tRNA_1^{Asp}$ (*F*(or *G*)) gene and does not contain any suggestive sequence for these transcriptional signals, the gene for $tRNA_1^{Asp}$(*F*(or *G*)) must be co-transcribed with

the rRNA genes present upstream, as proposed by Ikemura and Nomura (44). We therefore conclude that the terminator found in the region beyond the 3'-end of the gene for $tRNA_1^{Asp}$(*F*(or *G*)) must be the signal for transcription termination of the *rrnF*(or *G*) operon.

Recently, Young determined the DNA sequence of the *rrnC* operon in the region containing the distal genes for both $tRNA_1^{Asp}$(*C*) and $tRNA^{Trp}$ in this order and the accompanying rho-independent transcription termination signal (18). The secondary structure of the *rrnC* terminator is very similar to that for *rrnF*(or *G*). It is composed of a hairpin structure of similar size to that of *rrnF*(or *G*), followed by eight thymidine residues. When the nucleotide sequences of these two *rrn* terminators are compared, the great similarities shown in Fig. 7b are seen in the region downstream from the center of the hairpin structures. Recently, the region of transcription termination in *rrn*B has been sequenced by Brosius *et al.* (J. Brosius, personal communication). A secondary structure similar to that shown in Fig. 7a is present as a possible terminator. Although the sequence similarities observed between the *rrnF*(or *G*) and the *rrnC* terminator are not seen in *rrnB*, the sequence of $\begin{matrix}G\\A\end{matrix}$TCTG in the region immediately beyond the series of thymidine residues is common to the three *rrn* terminators so far sequenced. Interestingly, Young has pointed out the fact that the rho-independent termination of λ6S RNA (45, 46) occurs immediately before the sequence ATCTG (R.A. Young, personal communication). We also noticed that the similar sequence TTCTG, ATCTG, and ATCTA are present immediately after the recognition signal of CAATCA in the rho-dependent terminators of λtR1 (42), the $tRNA_1^{Tyr}$ operon (10) and IS2 (I) (43), respectively. Since a conserved short sequence is present in a number of terminators, its biological function, if any, may be related to terminator recognition by RNA polymerase.

Possible protein gene adjacent to rrnF(or G). As shown in Fig. 4, in the region downstream from the *rrnF*(or *G*) terminator, the first ATG sequence appears at position 540. For the following reasons, we propose the presence of a gene for an unknown protein (X), the translation of which starts at the ATG codon. i) The sequence GAGG, present six base pairs upstream from the initiation codon, corresponds to the recognition site for ribosome binding proposed by Shine and Dalgarno (47). ii) The heptanucleotide sequence CATAATC at positions 495-501 is very similar to that of the RNA polymerase binding site (CATAATG) in the λoop promoter (48) and therefore may correspond to a Pribnow box (49). iii) The sequence CTTT is present in the region about 25 base pairs upstream from the center of the Pribnow box. In the promoters

sequenced thus far for *E. coli* operons, those for $tRNA_1^{Tyr}$ (6), *lac* (50), *gal* (51), *ala* (52) and β (53) have the sequence CTTT in a corresponding region, and it believed to be the recognition site for RNA polymerase (54). These features of the DNA sequence strongly suggest that a promoter for a transcription unit might be present about 50 bp downstream from the *rrnF*(or *G*) terminator. If the assumed promoter is functional, transcription may start at the A or G residue (position 507 or 508), 9 or 10 base pairs downstream from the center of the Pribnow box. We are now examining whether this proposed gene really exists in cloned DNA. If it does, the function of the gene product and its relationshop to the *rrn* operon should be of interest.

ACKNOWLEDGEMENTS

We are grateful to Dr. M. Nomura for pLC7-21, to Dr. R.A. Young for pointing out the common sequences in *rrn* terminators and for useful discussion, and to Drs. J.A. Steitz, R.A. Young, J. Brosius and M. Nomura for showing us their unpublished results. We also thank Miss. Y. Kanemaru for help in DNA cloning. This work was partly supported by Grants-in-Aid for Cancer Research and Scientific Research from the Ministry of Education, Science and Culture of Japan.

REFERENCES

1. Sprinzl, M., Grueter, F., Spelzhaus, A. and Gauss, D.H. (1980) *Nucleic Acids Res. 8*, rl-r22.
2. Ikemura, T. and Ozeki, H. (1977) *J. Mol. Biol. 117*, 419-446.
3. Nishimura, S., Harada, F., Narushima, U. and Seno, T. (1967) *Biochim. Biophys. Acta 142*,133-148.
4. Goodman, H.M., Abelson, J., Landy, A., Brenner, S. and Smith, J.D. (1968) *Nature 217*, 1019-1024.
5. Russell, R.L., Abelson, J.H., Landy, A., Gefter, M., Brenner, S. and Smith, J.D. (1970) *J. Mol. Biol. 47*, 1-13.
6. Sekiya, T. and Khorana, H.G. (1974) *Proc. Natl. Acad. Sci., USA 71*, 2978-2982.
7. Sekiya, T., van Ormondt, H. and Khorana, H.G. (1975) *J. Biol. Chem. 250*, 1087-1098.
8. Sekiya, T., Gait, M.J., Noris, K., Rammamoorthy, B. and Khorana, H.G. (1976) *J. Biol. Chem. 251*, 4481-4489.
9. Loewen, P.C., Sekiya, T. and Khorana, H.G. (1974) *J. Biol. Chem. 249*, 217-226.
10. Küper, H., Sekiya, T., Rosenberg, M., Egan, J. and Landy, A. (1978) *nature 272*, 423-428.
11. Egan, J. and Landy, A. (1978) *J. Biol. Chem. 253*, 3607-3622.
12. Berman, M.L. and Landy, A. (1979) *Proc. Natl. Acad. Sci., USA 76*, 4304-4307.
13. Morgan, E.A., Ikemura, T. and Nomura, M. (1977) *Proc. Natl. Acad. Sci., USA 74*, 2710-2714.

14. Morgan, E.A., Ikemura, T., Lindahl, L., Fallon, A.M. and Nomura, M. (1978) *Cell 13*, 335-344.
15. Sekiya, T. and Nishimura, S. (1979) *Nucleic Acids Res. 6*, 575-592.
16. Young, R.A., Macklis, R. and Steitz, J.A. (1979) *J. Biol. Chem. 254*, 3264-3271.
17. Morgan, E.A., Ikemura, T., Post, L. and Nomura, M. (1979) In *Transfer RNA* (Abelson, J., Schimmel, P. and Söll, D., eds) Cold Spring Harbor Laboratory, Cold Spring Harbor, in press.
18. Young, R.A. (1979) *J. Biol. Chem. 254*, 12725-12731.
19. Panet, A., Van de Sande, J.H., Loewen, P.C., Khorana, H.G., Raae, A.J., Lillehaug, J.R. and Kleppe, K. (1973) *Biochem. 12*, 5045-5050.
20. Burgess, R.R. and Jendrisak, J.J. (1975) *Biochem. 14*, 4634-4638.
21. Roberts, J.W. (1969) *Nature 224*, 1168-1174.
22. Reeve, A.E. and Huang, R.C. (1979) *Nucleic Acid Res. 6*, 81-90.
23. Tabak, H.F. and Flavell, R.A. (1978) *Nucleic Acids Res. 5*, 2321-2332.
24. Maxam, A. and Gilbert, W. (1977) *Proc. Natl. Acad. Sci., USA 74*, 560-564.
25. Southern, E.M. (1975) *J. Mol. Biol. 98*, 503-517.
26. So, M., Gill, R. and Flakow, S. (1975) *Molec. Gen. Genet. 142*, 239-249.
27. Enea, V., Vovis, G.F. and Zinder, N.D. (1975) *J. Mol. Biol. 96*, 495-509.
28. Clewell, D.B. and Helinski, D.R. (1970) *Biochem. 9*, 4428-4439.
29. Sekiya, T., Contreras, R., Takeya, T. and Khorana, H.G. (1979) *J. Biol. Chem. 254*, 5802-5816.
30. Ikemura, T. and Dahlberg, J.E. (1973) *J. Biol. Chem. 248*, 5024-5032.
31. Thomas, M. and Davis, R.W. (1975) *J. Mol. Biol. 91*, 315-328.
32. Boros, I., Kiss, A. and Venetianer, P. (1979) *Nucleic Acids Res. 6*, 1817-1830.
33. Bram, R.J., Young, R.A. and Steitz, J.A. (1980) *Cell 19*, 393-401.
34. De Boer, H.A., Gilbert, S. and Nomura, M. (1979) *Cell 17*, 201-209.
35. Young, R.A. and Steitz, J.A. (1979) *cell 17*, 225-234.
36. Csordás-Tóth, É., Boros, I. and Venetianer, P. (1979) *Nucleic Acid Res. 7*, 2189-2197.
37. Brosius, J., Dull, T.J. and Noller, H.F. (1980) *Proc. Natl. Acad. Sci., USA 77*, 201-204.
38. Harada, F., Yamaizumi, K. and Nishimura, S. (1972) *Biochem. Biophys. Res. Comm. 49*, 1605-1609.
39. Sekiya, T., Contreras, R., Küpper, H., Landy, A. and Khorana, H.G. (1976) *J. Biol. Chem. 25*, 5124-5140.
40. Rossi, J.J. and Landy, A. (1979) *Cell 16*, 523-534.
41. Adhya, S. and Gottesman, M. (1978) *Ann. Rev. Biochem. 47*, 967-996.
42. Rosenberg, M., Court, D., Shimatake, H., Brady, C. and Wulff, D.L. (1978) *Nature 272*, 414-423.
43. Ghosal, D., Sommer, H. and Saedler, H. (1979) *Nucleic Acids Res. 6*, 1111-1122.
44. Ikemura, T. and Nomura, M. (1977) *Cell 11*, 779-793.
45. Lebowitz, P., Weissman, S.M. and Radding, C.M. (1971) *J. Biol. Chem. 246*, 5120-5139.
46. Rosenberg, M., DeCrombrugghe, B. and Musso, R. (1976) *Proc. Natl. Acad. Sci., USA 73*, 717-721.
47. Shine, J. and Dalgarno, L. (1974) *Proc. Natl. Acad. Sci., USA 71*, 1342-1346.
48. Schere, G., Hobom, G. and Kössel, H. (1977) *Nature 265*, 117-121.
49. Pribnow, D. (1975) *J. Mol. Biol. 99*, 419-443.
50. Dickson, R.C., Abelson, J., Barnes, W.M. and Reznikoff, W.S. (1975) *Science 187*, 27-35.
51. Musso, R., DiLauro, R., Rosenberg, M. and DeCrombrugghe, B. (1977) *Proc. Natl. Acad. Sci., USA 74*, 106-110.

52. Greenfield, L., Boon, T. and Wilcox, G. (1978) *Proc. Natl. Acad. Sci., USA 75*, 4724-4728.
53. Post, L.E., Strycharz, G.D., Nomura, M., Lewis, H. and Dennis, P.P. (1979) *Proc. Natl. Acad. Sci., USA 76*, 1697-1701.
54. Gilbert, W. (1976) In *RNA polymerase* (Losick, R. and Chamberlin, M. eds.) Cold Spring Harbor Laboratory, Cold Spring Harbor, p.193-225.