

A Discrete Region Centered 22 Base Pairs Upstream of the Initiation Site Modulates Transcription of *Drosophila* tRNA^{Asn} Genes

ALAN K. LOFQUIST, ALONZO D. GARCIA, AND STEPHEN J. SHARP*

*Department of Microbiology and Molecular Genetics, California College of Medicine,
University of California, Irvine, California 92717*

Received 18 April 1988/Accepted 5 July 1988

We have studied the mechanism by which 5'-flanking sequences modulate the in vitro transcription of eucaryotic tRNA genes. Using deletion and linker substitution mutagenesis, we have found that the 5'-flanking sequences responsible for the different in vitro transcription levels of three *Drosophila* tRNA^{Asn} genes are contained within a discrete region centered 22 nucleotides upstream from the transcription initiation site. In conjunction with the A-box intragenic control region, this upstream transcription-modulatory region functions in the selection mechanism for the site of transcription initiation. Since the transcription-modulatory region directs the position of the start site and the actual sequence of the transcription-modulatory region determines the level of tRNA^{Asn} gene transcription, the possibility is raised that the transcription-modulatory region directs a transcription initiation event similar to open complex formation at procaryotic promoters.

In procaryotic organisms, initiation of RNA synthesis is controlled by small, conserved DNA sequences located 10 and 35 nucleotides upstream of the transcription initiation site. RNA polymerase holoenzyme recognizes and binds these regions in this first step of the pathway that leads to open-complex formation and transcriptional activation (5, 42). The three nuclear RNA polymerases of eucaryotes do not recognize promoters efficiently and, as a result, are unable to independently initiate gene-specific transcription (for a review, see reference 38). Instead, eucaryotic transcription is dependent upon promoter-specific transcription factors that recognize and bind discrete sequence elements. Current models suggest that the binding of these protein factors activates transcription by allowing each gene to be recognized by the appropriate polymerase. For RNA polymerases I and II, the known sites of transcription factor interaction are 5' of the site of transcription initiation (9, 15, 45). For RNA polymerase III-directed transcription of eucaryotic 5S RNA, tRNA and adenovirus VA RNA genes, the sites of transcription factor interaction are 3' of the transcription initiation site, within the sequences encoding the RNA (6, 16, 17, 48, 50). RNA polymerase III-directed transcription of 7SL RNA, 7SK RNA, and U6 RNA genes requires enhancerlike sequence elements for transcription factor interaction 5' of the transcription initiation site, similar to RNA polymerase II promoters (3, 7, 30, 47).

The intragenic control regions (ICRs) in tRNA genes, the A- and B-box ICRs, contain the DNA sequence determinants that direct transcription factor activities (reviewed in reference 40). For tRNA genes, the transcription factors bind to the gene to form a complex that is stable throughout multiple rounds of transcription initiation (4, 23, 36). This stable complex, comprised of tRNA gene and transcription factors, is considered to form the recognition substrate for RNA polymerase III.

Transcriptional activity of tRNA genes is particularly sensitive to mutations in the A- and B-box ICRs. All of the point mutations generated in vivo in the *Saccharomyces*

SUP4-o tRNA^{Tyr} gene which lead to loss of suppression were mapped intragenically (1). These results suggested that all the sequences necessary for tRNA gene transcription were contained within the sequence encoding the RNA. Although the ICR A- and B-box sequences are essential for tRNA gene transcription, in vitro studies with *Drosophila*, *Bombyx*, *Saccharomyces*, *Xenopus*, and mouse tRNA genes provide evidence that tRNA gene transcription is also dependent upon sequences that extend beyond the A- and B-box ICRs (11, 20, 27, 41, 49). In particular, 5'-flanking sequences can modulate the transcriptional activity of tRNA genes dramatically (13, 51). Suppressor tRNA genes have been used in *Saccharomyces* sp. to demonstrate that 5'-flanking sequences also have a major influence on tRNA gene transcription in vivo (32, 41), but the mechanism by which tRNA gene transcription is modulated by extragenic sequences has not been determined.

We previously demonstrated the differential transcription activation of three tRNA^{Asn} genes of *Drosophila melanogaster* (25), and in the present study we have used these genes to investigate the mechanism by which 5'-flanking sequences modulate the transcription of tRNA genes. Each of the three genes, *Asn6*, *Asn7*, and *Asn8*, exhibits a different relative in vitro transcription efficiency (*Asn8* > *Asn6* > *Asn7*) dictated by their respective 5'-flanking sequences and not related to the ability of each gene to form a stable transcription complex. This observation led to the conclusion that RNA polymerase III recognized the stable transcription complex of each gene equally, but that the actual ability of RNA polymerase III to initiate transcription was determined by the sequence composition of the 5'-flanking DNA. The present study has investigated which sequences in the tRNA gene 5'-flanking region are responsible for this phenomenon. Using deletion and linker substitution mutations in *Asn8*, *Asn7*, and *Asn6*, we show that tRNA gene transcription is modulated by a discrete region centered 22 nucleotides upstream from the transcription start site. Furthermore, it appears that this region is involved in the selection mechanism for the site of transcription initiation. The present study identifies a modulatory region upstream of

* Corresponding author.

the control regions directing transcription factor interaction. Our considerations, therefore, focus on the possible interactions occurring in this upstream region during transcription activation.

MATERIALS AND METHODS

Plasmid DNAs. The isolation and characterization of the three *D. melanogaster* tRNA^{Asn} genes used in this study have previously been described (25). The three tRNA^{Asn} genes, *Asn6*, *Asn7*, and *Asn8*, have identical mature tRNA coding regions with different 5'- and 3'-flanking sequences. Plasmids pAsn6/8 and pAsn7/8 contain hybrid tRNA^{Asn} genes which differ in 5'-flanking sequence only. The construction of these hybrid genes, which contain the 5'-flanking sequence of either *Asn6* (pAsn6/8) or *Asn7* (pAsn7/8), the mature tRNA^{Asn}-coding region, and the 3'-flanking sequence of *Asn8*, was also previously described (25).

Construction of 5'- and 3'-deletion mutations. Deletion mutations in the 5'-flanking sequence of *Asn8* and *Asn7* were constructed by standard *Bal* 31 exonuclease procedures (24). 5' deletions were generated from the *Hind*III site at position -49 of pAsn8. 3' deletions were generated from either the *Hpa*I site (position 37) or the *Sma*I site (position 67) of both *Asn8* and *Asn7*. After exonuclease treatment, ends were made blunt by treatment with T4 DNA polymerase. Deletion fragments were then liberated from the vector with either *Sal* I (position 167 of *Asn8*) for 5' deletions or *Hind*III (position -49 of *Asn8*; position -480 of *Asn7*) for 3' deletions and electrophoretically sized by using 1.5% low-melting-point agarose gels. DNA fragments were purified out of agarose and ligated into either the *Sma*I-*Sal*I site (5' deletions) or the *Hind*III-*Sma*I site (3' deletions) of M13mp9. The resulting deletion mutations are named for their deletion endpoints as determined by dideoxy sequencing (34).

Construction of LS mutations. The 5'- and 3'-deletion mutations were used to construct linker substitution (LS) mutations as described by Garcia et al. (19). 5'- and 3'-deletion fragments were ligated via the *Eco*RI site originally located in the polylinker of M13mp9. The resulting DNAs were inserted into the *Hind*III/*Sal*I site of pUC81, a pUC8 derivative lacking an *Eco*RI site (25). Selection of appropriate 5'- and 3'-deletion endpoints resulted in replacing 10 base pairs (bp) of wild-type *Drosophila* sequence with the 10-bp *Eco*RI linker sequence GGAATTC. For some LS constructions, only a 6-bp linker was used. In these cases, the construction containing the 10-bp linker was linearized with *Eco*RI, treated with mung bean nuclease to remove the 4-bp, single-stranded overhangs, and recircularized with T4 DNA ligase. This resulted in replacing 6 bp of wild-type *Drosophila* sequence with the 6-bp *Apa*I linker sequence GGGCCC.

Three types of LS constructions were used in this study. Constructions in which DNA both 5' and 3' of the linker sequence originated from *Asn8* are termed simply LS. Constructions in which the DNA 5' of the linker sequence is from *Asn7* and that 3' of the linker sequence is from *Asn8* are termed *Asn7* hybrid LS (7HLS). Constructions in which the DNA 5' of the linker sequence is from *Asn6* and that 3' of the linker sequence is from *Asn8* are termed *Asn6* hybrid LS (6HLS). Some of the 7HLS and all of the 6HLS mutations were constructed by exchanging the 5'-flanking sequence of *Asn7* (for 7HLS) or *Asn6* (for 6HLS) with that of *Asn8* LS constructions via the intragenic *Bgl*II site at nucleotide position 7. The nomenclature of LS mutations is as described by McKnight and Kingsbury (26) and indicates the 5'- and

3'-deletion mutations joined by either an *Eco*RI or *Apa*I linker. For example, LS-15/-4 has a 10-bp *Eco*RI linker replacing *Drosophila* DNA between nucleotide positions -15 and -4 of *Asn8*. Similarly, LS-25/-18 has a 6-bp *Apa*I linker replacing *Drosophila* DNA between positions -25 and -18.

Preparation of cell-free extract and in vitro transcription reactions. *D. melanogaster* Schneider S3 and S2 cells were grown as described previously (25). Cells were collected by centrifugation (Beckman JA-10 centrifuge, 4,000 rpm, 1 min), washed in 1 to 2% of the culture volume with insect Ringer solution (128.34 mM NaCl, 4.7 mM KCl, 1.89 mM CaCl₂), suspended in 2 volumes of buffer H (10 mM HEPES [N-2-hydroxyethylpiperazine-N'-2-ethanesulfonic acid]-KOH [pH 8.0], 10 mM KCl, 1.5 mM MgCl₂, 0.5 mM dithiothreitol) and allowed to swell on ice for 30 min. Cells were lysed by using 10 strokes of a tight-fitting Dounce homogenizer. One-tenth volume of buffer I (0.3 M HEPES-KOH [pH 8.0], 1.27 M KCl, 40 mM magnesium acetate, 50 mM dithiothreitol) was added, after which the mixture was centrifuged at 100,000 × *g* for 1 h at 4°C. The supernatant (S-100) was collected and glycerol was added to a final concentration of 20% (vol/vol). The S-100 was stored in small samples at -70°C. The final apparent concentrations of buffer and salt components in the S-100 were 27 mM HEPES-KOH (pH 8.0), 94 mM KCl, 3.75 mM MgCl₂, 3.75 mM dithiothreitol, and 20% (vol/vol) glycerol.

Transcription reactions were incubated at 24°C and contained 200 ng of covalently closed circular template DNA, 800 ng of carrier DNA (pUC8), 15 μl of cell extract (S-100), 30 mM HEPES-KOH (pH 8.0), 70 mM KCl, 5 mM MgCl₂, 4 mM dithiothreitol, 500 μM each ATP, CTP, and UTP, 100 μM [α-³²P]GTP (0.1-0.4 Ci/mmol), 6.4 mM creatine phosphate, and 3.5 U of creatine phosphokinase in a 60-μl reaction volume. The reactions were stopped after 90 min by the addition of 50 μl of a solution containing 1 mg of proteinase K per ml and 0.1% (wt/vol) sodium dodecyl sulfate in 30 mM HEPES-KOH (pH 8.0) (preincubated at 37°C, 20 min), after which the reactions were incubated at 37°C for 1 h. After phenol and chloroform extraction, transcription products were collected by ethanol precipitation, suspended in a solution containing 80% (vol/vol) formamide-0.1% (wt/vol) xylene cyanol FF-0.1% (wt/vol) bromophenol blue-2 mM EDTA, and electrophoretically resolved with an 8% (wt/vol) polyacrylamide gel (acrylamide-N,N'-methylenebisacrylamide; 20:1) containing 8.3 M urea-89 mM Tris-89 mM boric acid-2 mM EDTA. Relative transcription levels were determined by densitometer scanning of autoradiographs (LKB Ultrosan XL densitometer).

5'-End analysis of precursor tRNA transcripts. The transcription initiation site(s) for each template was determined by nearest-neighbor nucleotide analysis. For each template, four sets of transcription reactions were performed, each incorporating a different α-³²P-labeled nucleoside triphosphate. Precursor RNAs were eluted from gel slices by incubating at 37°C overnight in a solution containing 10 mM Tris hydrochloride (pH 7.5), 0.3 M NaCl, 1 mM EDTA, and 1% (wt/vol) phenol. *Escherichia coli* carrier tRNA (5 μg) was added, and after ethanol precipitation RNA was suspended in 20 mM ammonium acetate (pH 4.5) and digested to completion with 2 U of RNase T2 by incubating at 37°C for 2.5 h. The initial nucleotide(s), determined as T2-liberated 5'-triphosphate 3'-phosphate (pppNp), was analyzed by polyethyleneimine cellulose thin-layer chromatography (20-by 20-cm plates) developed with 0.75 M KH₂PO₄ (pH 3.5). The polyethyleneimine cellulose plates were autoradio-

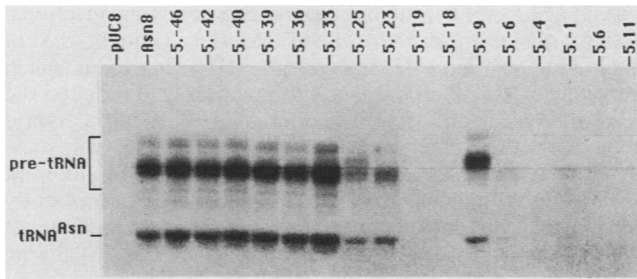


FIG. 1. 5' limit of the sequences modulating transcription. Shown is an autoradiograph of the in vitro transcription products of *Asn8* 5'-deletion mutations as electrophoretically separated on a denaturing polyacrylamide gel. Template DNAs with the deletion endpoints indicated above each lane were transcribed in *Drosophila* cell extracts as described in Materials and Methods. Precursor and mature tRNA transcripts for each template are indicated. The transcription of M13mp9 replicative form and wild-type *Asn8* served as controls.

graphed after UV visualization of marker nucleotides pppG and pppA.

The transcription initiation site(s) for each template was further confirmed by nuclease P1 digestion of the T2-digested RNAs. To remove 3'-phosphate groups, 1 U of nuclease P1 was added to samples previously digested with RNase T2 and incubated at 37°C for 2 h. The initial nucleotide(s), determined as T2/P1-liberated pppN, were chromatographed on polyethyleneimine cellulose plates with the same solvent system as above. For the T2/P1 digests, pppG, pppA, pppU, and pppC were used as marker nucleotides.

RESULTS

5' limit of the sequences modulating transcription. We previously determined that the region responsible for efficient transcription of *Asn8* (relative to *Asn6* and *Asn7*) was contained within the 50 nucleotides immediately upstream from the mature tRNA-coding region (25). To more accurately delimit the 5' end of this region, a series of 16 5'-deletion mutations, encompassing the entire 5'-flanking region of *Asn8* and extending into the gene itself, were constructed by using *Bal* 31 exonuclease as described under Materials and Methods. The 5'-deletion mutations were transcribed in extracts of *Drosophila* cells, and their transcription patterns are shown in Fig. 1. The multiple bands indicated as pre-tRNA in Fig. 1 arise from multiple transcription termination sites as determined by 5'-end analysis of pre-tRNAs (see below) and by 5'-flanking sequence exchange experiments (25). No appreciable difference in the relative transcription level (compared with that of wild-type *Asn8*) occurred when *Drosophila* DNA was deleted to nucleotide position -33 (5.-33). When *Drosophila* DNA was deleted to position -25 (5.-25), a dramatic decrease in the ability of the template to support efficient transcription was observed (35% of the level of *Asn8*). Further deletion resulted in template DNAs that either did not support transcription (5.-19, 5.-18, and 5.11) or supported a low level of transcription (5.-23, 5.-6, 5.-4, 5.-1, and 5.6 at 50, 10, 1, 13, and 2% of the level of *Asn8*, respectively). The only exception was the deletion mutation 5.-9, which, presumably due to the fortuitous conservation of critical upstream nucleotides by the substituted plasmid DNA, was transcribed at a level 70% that of wild-type *Asn8*. The 5' limit of the region required for efficient transcription of *Asn8*,

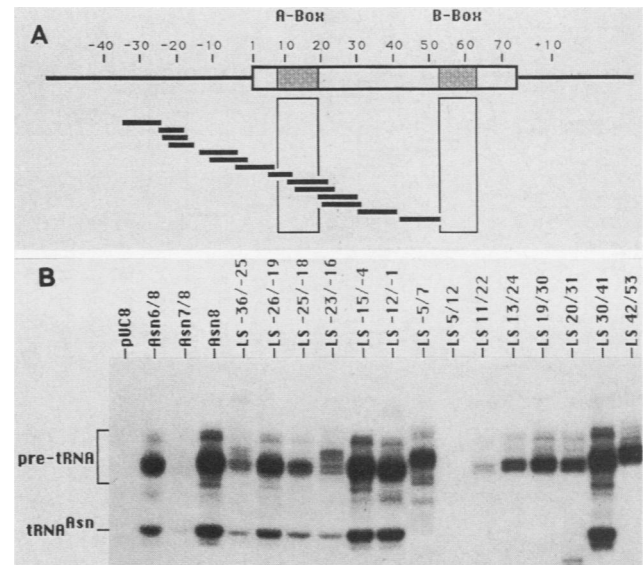


FIG. 2. Three distinct regions control tRNA gene activity in vitro. (A) Schematic representation of the *Asn8* LS mutations. The mature tRNA-coding sequence is represented by the open box, and the positions of the consensus A- and B-box ICRs are indicated. Solid bars below the gene indicate the positions of the different LS mutations. (B) Autoradiograph of the in vitro transcription products of *Asn8* LS mutations as electrophoretically separated on a denaturing polyacrylamide gel. Template DNAs with LS mutations in the positions indicated were transcribed in *Drosophila* cell extracts as described in Materials and Methods. Precursor and mature tRNA transcripts for each template are indicated. The transcription of pUC8, *Asn6/Asn8*, *Asn7/Asn8*, and *Asn8* served as controls.

therefore, occurs between nucleotide positions -33 and -25, since deletion of wild-type *Asn8* 5'-flanking sequence to position -33 had little effect on the relative transcription level, whereas all further deletions had significant effects on the in vitro transcription characteristics of this gene.

Transcription efficiency is directed by a short upstream region. To further define the region(s) responsible for transcriptional modulation of tRNA^{Asn} genes, we constructed a series of 14 linker mutations in *Asn8* extending from nucleotide position -36 in the 5'-flanking region to nucleotide position 53 in the mature tRNA-coding region. The LS mutations presented in this study were generated by ligating tDNA 5'- and 3'-deletion mutations by either an *Eco*RI or an *Apa*I linker, which resulted in exactly replacing the wild-type *Drosophila* sequence with either the 10-bp *Eco*RI linker sequence GGG AAT TCCC or the 6-bp *Apa*I linker sequence GGGCCC.

The results of LS analysis revealed the presence of a discrete region in the 5'-flanking sequence important for efficient transcription of *Asn8*. LS mutations which replaced wild-type *Asn8* 5'-flanking sequence between nucleotide positions -36 and -16 (LS-36/25, LS-26/-19, LS-25/-18, and LS-23/-16) reduced the transcriptional activity of *Asn8* significantly (26, 49, 30, and 35% of the level of the wild type, respectively) (Fig. 2; see Fig. 5A). LS mutations which replaced the *Asn8* sequence between positions -15 and -1 (LS-15/-4 and LS-12/-1) were transcribed at nearly wild-type levels (86 and 79% of the level of the wild type, respectively). The mutation LS-5/7, which disrupts the wild-type *Asn8* start site (-4A) and extends into the coding region, was transcribed at a level 52% of that observed for *Asn8*. Mutations which resulted in disrupting the A-box ICR

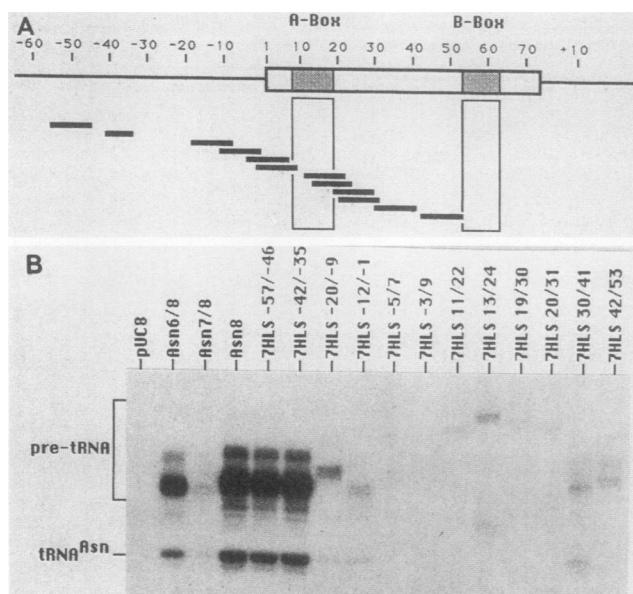


FIG. 3. Transcription efficiency is directed by a short upstream region. (A) Schematic representation of the 7HLS mutations. The mature tRNA-coding sequence is represented by the open box, and the positions of the consensus A- and B-box ICRs are indicated. Solid bars below the gene indicate the positions of the different 7HLS mutations. For these mutations, the DNA 5' of the linker sequence is from *Asn7*, and that 3' of the linker sequence is from *Asn8*. (B) Autoradiograph of the in vitro transcription products of tDNAs containing 7HLS mutations as electrophoretically separated on a denaturing polyacrylamide gel. DNAs with linker mutations in the positions indicated were transcribed in *Drosophila* cell extracts as described in Materials and Methods. Precursor and mature tRNA transcripts for each template are indicated. The transcription of pUC8, *Asn6/Asn8*, *Asn7/Asn8*, and *Asn8* served as controls.

(LS5/12, LS11/22, and LS13/24) decreased *Asn8* transcription significantly (0, 8, and 24% of wild-type level, respectively). The mutations LS19/30, LS20/31, and LS42/53, which altered the tRNA gene (tDNA) immediately adjacent to the consensus A- and B-box ICRs, were transcribed at levels 39, 36, and 47% of that of wild-type *Asn8*, respectively (consensus A- and B-box ICRs for these tRNA^{Asn} genes are located from positions 8 through 19 and 53 through 63, respectively [18]). The mutation LS30/41, which replaced DNA in the region between the two ICRs, was transcribed at approximately wild-type *Asn8* levels (113%), demonstrating that the linker sequence in itself was not deleterious to transcription activity.

We previously exchanged the 5'-flanking regions of *Asn6* and *Asn7* with that of *Asn8* to demonstrate that the differential transcription activation of *Asn6*, *Asn7*, and *Asn8* is directly attributable to their corresponding 5'-flanking sequences (25). To determine whether the upstream region responsible for efficient transcription of *Asn8* might be the same region responsible for poor transcription of *Asn7*, we constructed a series of 12 LS mutations in which DNA 5' of the linker sequence was derived from *Asn7* and DNA 3' of the linker sequence was derived from *Asn8*. These LS mutations were named *Asn7* hybrid linker substitution (7HLS) mutations to distinguish them from the *Asn8* LS mutations.

Introduction of *Asn7* 5'-flanking sequence to nucleotide position -42 (7HLS-57/-46 and 7HLS-42/-35) had little effect on the transcriptional activity of wild-type *Asn8* (85%;

Fig. 3; see Fig. 5B), indicating that the poor transcriptional activity of *Asn7* is not a result of inhibitory sequences 5' of nucleotide position -42. In contrast, 7HLS mutations which introduced *Asn7* flanking DNA to position -20 reduced the level of *Asn8* transcription to that of *Asn7/Asn8*, a hybrid gene with the same transcriptional activity as wild-type *Asn7*. (To eliminate any possible effects due to different 3'-flanking sequences, all of the templates used in this study were constructed to contain the 3'-flanking sequence of *Asn8*. For example, *Asn7/Asn8* consists of the 5'-flanking sequence of *Asn7*, the mature tRNA^{Asn}-coding region, and the 3'-flanking sequence of *Asn8*.) These results indicate that the region responsible for the poor transcription of *Asn7* is located between nucleotide positions -42 and -20. The tDNAs containing 7HLS mutations extending from nucleotide position -12 to 53 all supported a level of transcription equal to or less than that of wild-type *Asn7*. The sequences responsible for the poor transcriptional activity of *Asn7* (positions -42 to -20), therefore, are located in the same approximate region as those responsible for the efficient transcriptional activity of *Asn8* (positions -36 to -16). The combined results of deletion, LS and 7HLS analyses define the approximate boundaries for an upstream control region to be nucleotide positions -33 and -20.

The site of transcription initiation is determined by both the upstream control region and the A-box ICR. Several deletion, LS, and 7HLS mutant templates gave rise to precursor tRNA transcripts that had electrophoretic mobilities different from those produced from p*Asn6* or p*Asn7/8*, suggesting that several of the mutations caused a selection of alternative or additional sites of transcription initiation. To examine this more closely, we used nearest-neighbor nucleotide analysis to determine the transcription initiation site(s) for each template. Precursor tRNAs were isolated after gel electrophoresis and digested to completion with RNase T2 and then with nuclease P1. The initial nucleotides were determined as T2-liberated pppNp (Fig. 4) and were verified as T2-P1-liberated pppN (data not shown).

Transcription of wild-type *Asn8* initiated predominantly at nucleotide position -4A (Fig. 4). In addition, approximately 5% of the transcripts initiated at position -3G. Four of the LS mutations, LS-36/-25, LS-26/-19, LS-25/-18, and LS-23/-16, resulted in transcription initiation at position -6G in addition to the wild-type sites of initiation. Two of these mutations, LS-36/-25 and LS-23/-16, resulted in transcription initiation at position -6G as often as initiation at -4A. With the exception of the mutation LS5/12, which eliminated transcription in vitro, the remaining LS mutations resulted in tDNAs that all directed initiation at the wild-type position (-4A) (Fig. 5A). This selection occurred regardless of the nucleotide composition at the start site location. For example, in the mutation LS-12/-1 nucleotides at positions -4A and -3G were replaced with cytidine residues, whereas in the mutation LS-5/7 nucleotides at both positions were replaced with guanine residues. For each of these mutant tDNAs, transcription initiated predominantly at nucleotide position -4, indicating that RNA polymerase III is able to initiate transcription by using a pyrimidine. The only mutations in *Asn8* that resulted in the selection of alternative start sites occurred in the 5'-flanking control region.

Transcription of the hybrid gene *Asn7/Asn8* initiated predominantly at nucleotide position -3A with a low level of transcription initiation also occurring at position -1A (Fig. 4). These sites were the same as those selected for transcription of the wild-type *Asn7* gene (data not shown). Two of the 7HLS mutations, 7HLS-57/-46 and 7HLS-42/-35, re-

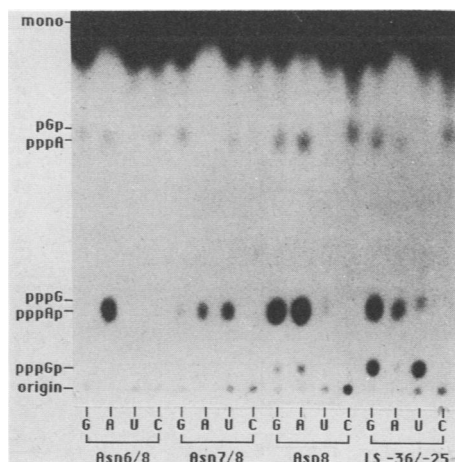


FIG. 4. Determination of the start site for transcription. The transcription initiation site for each template was determined by nearest-neighbor nucleotide analysis. For each template, four sets of transcription reactions were performed, each incorporating a different α - ^{32}P -labeled nucleoside triphosphate. Precursor RNA species (Fig. 3 and 4) were isolated and digested to completion with RNase T2, which liberates the 5'-terminal nucleoside pppNp. The 3'-phosphate of pppNp is derived from the second residue (nearest-neighbor) of the transcript RNA. In combination with the electrophoretic mobility of the transcript RNA, RNase T2 digestion of RNA which has been synthesized in the presence of α - ^{32}P -labeled nucleoside triphosphate, therefore, allows an unambiguous identification of the transcription initiation site by revealing the first two nucleotides of the transcript RNA. After digestion, the samples were chromatographed on a polyethyleneimine cellulose plate. The 5'-terminal pppNp was autoradiographed after UV visualization of marker nucleotides. Shown are representative RNase T2 digests for *Asn6/Asn8*, *Asn7/Asn8*, *Asn8*, and LS-36/-25. The α - ^{32}P -labeled nucleoside triphosphate used is indicated below each lane.

sulted in initiation at nucleotide position -4A , the normal start site for wild-type *Asn8* (Fig. 5B). The result supports the conclusion that *Asn7* sequences upstream of nucleotide position -42 have no effect on the in vitro transcription activity of this gene and, in addition, demonstrates that sequences upstream of position -42 have no effect on the selection of the transcription initiation site.

Alternative sites for transcription initiation were selected as a result of mutations 7HLS-20/-9 (-6G) and 7HLS-12/-1 (1G; the first nucleotide of the mature tRNA-coding region). The tDNA containing the mutation 7HLS-5/7 directed initiation at nucleotide position -4G ; the same position as that selected by the corresponding LS mutation in *Asn8* (LS-5/7). These results show that, as for *Asn8*, the 5'-flanking control region is involved in determining the site for transcription initiation of *Asn7*.

Whereas all 7HLS mutations extending into the coding region resulted in transcription initiation at the wild-type *Asn7* positions (-3 and -1), 7HLS mutations disrupting either the A-box ICR (7HLS11/22 and 7HLS13/24) or the tDNA immediately adjacent to the A-box ICR (7HLS19/30 and 7HLS20/31) also directed selection of an additional site of initiation at position -14G (Fig. 5B). For tDNAs containing these mutations, transcription initiation at this additional site occurred as often as initiation at the wild-type site. It is particularly striking that identical mutations in the A-box ICR of *Asn8* did not result in the selection of additional sites of transcription initiation (Fig. 2 and 5A).

The differences in start site selection observed between the LS and 7HLS A-box mutations could only have occurred

as a result of differences in the 5'-flanking sequences of *Asn8* and *Asn7*, since the sequences of the coding and 3'-flanking regions of these tDNAs were identical (Fig. 5). To determine whether A-box mutations in *Asn6* would also result in start site heterogeneity as a consequence of 5'-flanking sequence differences, six linker mutations were constructed in the hybrid gene *Asn6/Asn8* by exchanging the 5'-flanking sequence of *Asn6* with that of *Asn8* LS constructions via the intragenic *Bgl*I site at nucleotide position 7. These constructions, named 6HLS mutations to distinguish them from the LS and 7HLS mutations, resulted in tDNAs that transcribed similarly to their LS counterparts in *Asn8* (Fig. 5). Transcription of the hybrid gene *Asn6/Asn8* initiated predominantly at nucleotide position -4A (Fig. 4), the same site as that selected for wild-type *Asn6* gene transcription (data not shown). Whereas A-box mutations in *Asn7* resulted in the selection of additional sites of initiation, tDNAs with identical 6HLS mutations directed initiation of transcription only at the wild-type site (Fig. 5C). In summary, mutations in a short, upstream region resulted in selection of alternative or additional sites for transcription initiation of *Asn7* and *Asn8*; mutations in the A-box ICR resulted in additional sites of transcription initiation for *Asn7* but not for either *Asn8* or *Asn6*.

DISCUSSION

It is generally accepted that transcription of eucaryotic tRNA genes is dependent upon two intragenic sequence elements, the A- and B-box ICRs (18, 39), which function as the sites for interaction of RNA polymerase III-specific transcription factors (23). Several studies have demonstrated an additional level of complexity in that the transcriptional activity of tRNA genes can be modulated dramatically by different 5'-flanking sequences (40). This extragenic modulation is particularly interesting in view of the differential levels of transcription observed for individual copies of gene members coding for the same isoaccepting tRNA. Such members have identical, or nearly identical, mature tRNA-coding regions and, accordingly, identical A- and B-box ICRs. The majority of sequence differences in these gene families occur within the DNA flanking the gene; indeed, the differential in vitro expression of *Drosophila* tRNA^{Lys}, tRNA^{Arg}, and tRNA^{Asn}, *Bombyx* tRNA^{Ala}, *Xenopus* tRNA^{Tyr}, and mouse tRNA^{His} genes has been attributed specifically to 5'-flanking sequence differences (12, 13, 20, 25, 27, 51).

Nearly all deletion analyses performed on tRNA genes have indicated that the 5' limit for transcription lies approximately 35 bp upstream from the mature tRNA-coding region. In the present study, we have used a series of deletion and LS mutations in both wild-type and hybrid tRNA genes to examine the 5'-flanking sequences responsible for transcriptional modulation. Deletion analysis of a *Drosophila* tRNA^{Asn} gene, *Asn8*, indicated that the 5' limit of the region responsible for efficient transcription of this gene is between nucleotide positions -33 and -25 . LS mutagenesis has allowed us to define the boundaries of this region even further. The only LS mutations introduced into the 5'-flanking DNA of *Asn8* that significantly altered the transcriptional activity of the gene were those located between nucleotide positions -36 and -16 . Mutations in the 5'-flanking DNA between this region and the gene had little effect on transcription. The importance of the discrete upstream region for transcriptional modulation was further examined by experiments in which this region of *Asn8* was

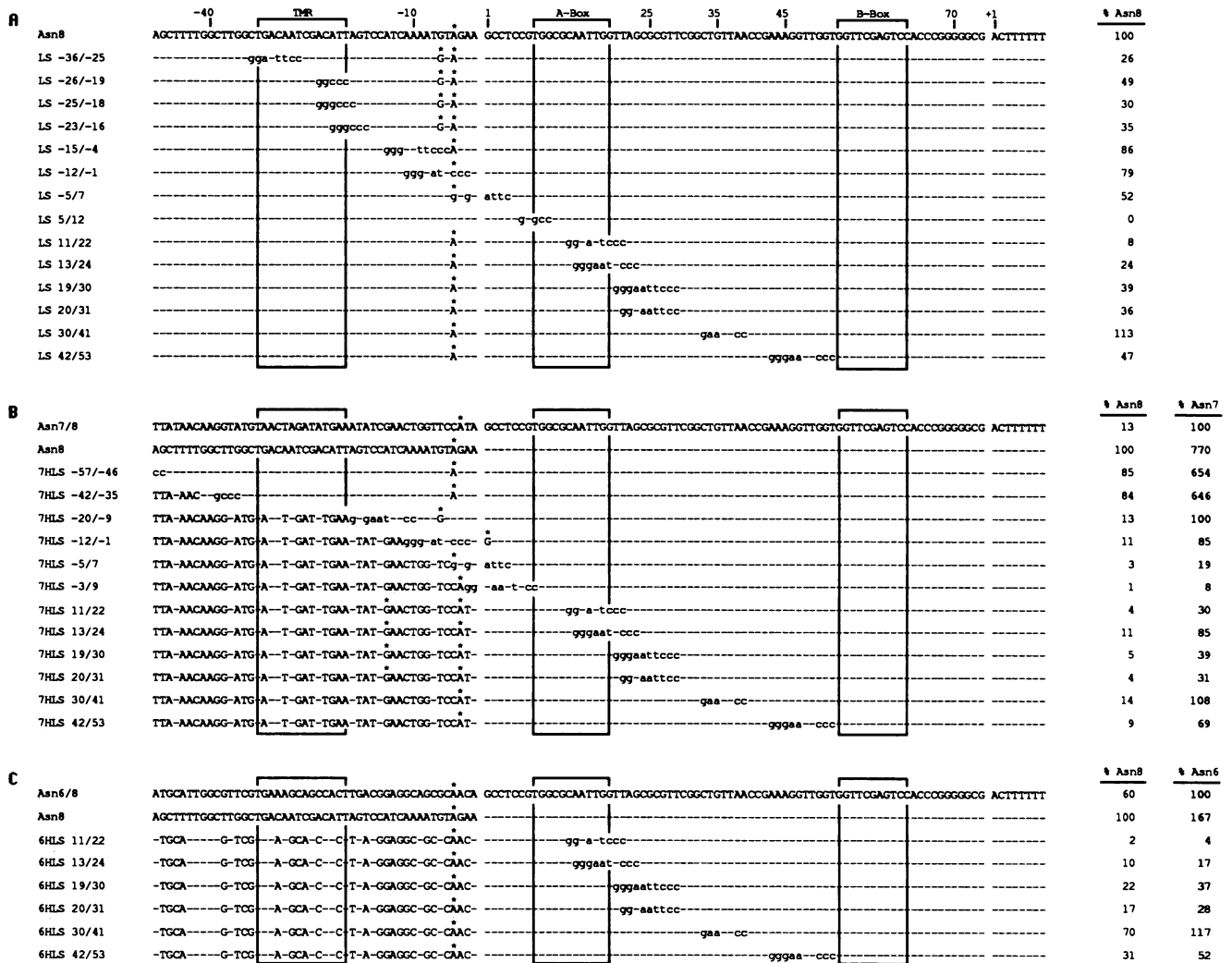


FIG. 5. Sequence comparison of wild-type, hybrid, and LS-mutation-containing tRNA^{Asn} genes. (A) The nucleotide sequences of the noncoding strand of wild-type *Asn8* and LS-containing mutations are shown. Nucleotide positions are numbered with respect to the mature tRNA-coding sequence (position 1), and the consensus A- and B-box ICRs and TMR are outlined (see text). The predominant initiating nucleotide(s) for each template is indicated by an asterisk. Mutations introduced by the linker sequence are presented in lowercase letters, and nucleotides corresponding to the wild-type sequence are represented by a dash. Relative transcription levels are presented as percentages of the transcription level of wild-type *Asn8*. (B) The nucleotide sequences of the noncoding strands of the hybrid *Asn7/Asn8*, wild-type *Asn8*, and the tDNAs containing 7HLS mutations are shown. Mutations introduced by the linker sequence are presented in lowercase letters, whereas mutations which replace the *Asn8* sequence with that of *Asn7* are presented in uppercase letters. Nucleotides corresponding to wild-type *Asn8* are represented by a dash. Relative transcription levels are presented as percentages of the transcription levels of both wild-type *Asn8* and the hybrid *Asn7/Asn8*. (C) The nucleotide sequences of the noncoding strands of the hybrid *Asn6/Asn8*, wild-type *Asn8*, and the tDNAs containing 6HLS mutations are shown. Mutations introduced by the linker sequence are presented in lowercase letters, whereas mutations which replace the *Asn8* sequence with that of *Asn6* are presented in uppercase letters. Nucleotides corresponding to wild-type *Asn8* are represented by a dash. Relative transcription levels are presented as percentages of the transcription levels of both wild-type *Asn8* and the hybrid *Asn6/Asn8*.

replaced with the corresponding region of *Asn7*. When the 5'-flanking sequence of *Asn8* between nucleotide positions -42 and -20 was replaced with that of *Asn7*, transcription decreased to a level equal to that of *Asn7*, demonstrating that, as for *Asn8*, the in vitro transcription level of *Asn7* is determined by the sequence composition of a discrete region located a short distance upstream of the gene. Taken together, the results of deletion, LS, and 7HLS analyses define the boundaries for an upstream control region to be nucleotide positions -33 and -20. To facilitate our discussion, we have called this region the transcription-modulatory region (TMR). The different TMR sequences of *Asn6*, *Asn7*, and

Asn8 are responsible for the different levels of in vitro transcription observed for these tRNA^{Asn} genes.

Although selection of the transcription initiation site is dependent upon the A-box ICR (2, 8), our results indicate that the A-box ICR is not the only sequence involved in the selection mechanism. LS mutations which disrupted the TMR of *Asn8* resulted in the selection of an additional transcription initiation site (Fig. 5). Similarly, the position of the transcription initiation site also changed when the TMR of *Asn8* was replaced with that of *Asn7*. In addition, mutations in the A-box ICR of *Asn7* resulted in the selection of an additional upstream site of transcription initiation. This was

not observed for identical mutations in the A-box ICRs of *Asn8* and *Asn6*. Recently, studies with a *Saccharomyces* tRNA_{3^{Leu}} gene showed that, under certain conditions, the yeast transcription factor τ , which recognizes both ICRs, could interact with an alternative A-box-like sequence in the 5'-flanking region, thereby directing initiation upstream of the site usually selected (2). We cannot exclude the possibility that mutations in the A-box of *Asn7* cause utilization of an alternative A-box. Nevertheless, because identical sets of mutations in *Asn8* and *Asn6* do not result in the selection of alternative start sites, use of an alternative A-box for *Asn7* transcription seems unlikely. The observation that A-box ICR mutations in *Asn7* do not display the same characteristics as identical mutations in *Asn8* and in *Asn6* suggests that the TMRs of these genes function in conjunction with the A-box ICRs in selecting the sites for transcription initiation. It is possible that the position of the TMR is determined by transcription factor interaction with the A-box ICR. Mutations in the A-box ICR of *Asn7*, therefore, could allow the utilization of an alternative upstream TMR and, as a result, allow selection of an alternative upstream start site.

Interaction of *trans*-acting transcription factors with the 5'-flanking region of some tRNA genes may serve special regulatory functions. Limited sets of 5'-flanking sequence similarities (28, 33) and two specific negative controlling sequence elements have been identified in some tRNA gene families (12, 21). In general, however, the 5'-flanking sequences of tRNA genes show little similarity or homology between members of the same or different tRNA gene families (40). The lack of observable 5'-flanking sequence similarities, although not exclusive, argues against the involvement of a sequence-specific DNA-binding protein as a general mechanism for the modulation of tRNA gene transcription. Specific transcription factor binding to 5'-flanking sequences has not been reported. Although the yeast transcription factor τ binds both tRNA gene ICRs and extends 8 to 13 nucleotides into the 5'-flanking sequence, the interaction of the factor with the 5' half of the gene is weak (2, 6, 22). Similar results have been observed for the human transcription factor TFIIC (17). The TMR defined in the present study is located upstream of the known limits for transcription factor interaction with tDNA and, as a result, may not interact directly with any known tDNA transcription factors.

As has been the case for other tRNA genes studied by deletion analysis (21, 35), a simple correlation does not exist between transcriptional activity of the tRNA^{Asn} genes and the base composition at any position upstream of the initiation site. It has been suggested that extensive DNA melting may occur in the -20 region during initiation of RNA polymerase III-directed transcription (41). Reduced transcriptional activity was observed for those LS mutations in *Asn8*, which increased the G+C content of the TMR (Fig. 5A). It appears, however, that G+C content is not the only important feature and that the modulatory mechanism must be more complex since the TMR of wild-type *Asn7* is not G+C rich and yet *Asn7* is a relatively poor template for transcription.

We previously demonstrated that the differential in vitro transcription activation of *Asn6*, *Asn7*, and *Asn8* was directly attributable to their corresponding 5'-flanking sequences and not related to the ability of each gene to sequester transcription factors (25). The present results reveal that the functional sequence determinants in the 5'-flanking sequence are within the TMR. The effect of TMR mutations on the position of the transcription initiation site

suggests a direct interaction between the TMR and RNA polymerase III, consistent with the suggestion that the structural conformation of the 5'-flanking region may be important to correctly orient RNA polymerase III to the initiation site (33, 44). The lack of strict sequence similarity in the tDNA TMR, however, combined with the mechanism of tRNA gene transcription, in which specific transcription by RNA polymerase III is dependent upon the prior binding of transcription factors, suggests that RNA polymerase III does not require specific recognition sequences for binding as, for example, is the case for procaryotic RNA polymerases. The TMR may instead act by providing the critical nucleotide(s) necessary for nucleation of strand separation, analogous to the formation of an open transcription complex by procaryotic RNA polymerases (5, 14). If this is the case, many different 5'-flanking sequences could be acceptable for RNA polymerase III activity, but the polymerase may not interact with different sequences equally, which would explain the modulatory effects observed for several wild-type tDNAs.

For genes transcribed by RNA polymerase III, the requirement for upstream signals has now become the rule rather than the exception (for a review, see reference 43). Recent studies with mouse U6, *Xenopus* U6, and human 7SK genes indicate that the 5'-flanking signals can be extreme in that these genes are subject to enhancer control and, furthermore, are seemingly not dependent upon internal control regions (3, 7, 10, 30). Although the major regulatory regions for these genes appear to be far upstream, they also demonstrate sensitivity to mutations in their -25 regions, which coincides well with the TMR defined in the tRNA^{Asn} genes. Upstream signals in this region are also essential for the in vitro transcription of adenovirus virus-associated RNA I genes (31, 44) and of 5S RNA genes from *Bombyx*, *Neurospora*, and *Drosophila* species (19, 29, 37, 46). It appears, then, that the TMR may serve a common function in RNA polymerase III-directed transcription. The one common component in the transcription of these genes is the RNA polymerase III itself. At present we do not know whether the TMR acts directly or indirectly (through another factor) on RNA polymerase III, but one clearly testable feature of this model is that the TMR acts by modulating the rate of transcription initiation.

ACKNOWLEDGMENTS

We thank Irene Smith, Chris Greer, and Bert Semler for helpful discussions and critical reading of the manuscript. In addition we thank Mark Lawson for help in the sequence verification of mutant tDNA templates.

This work was supported by Public Health Service grant GM 32486 from the National Institutes of Health. A.K.L. is supported by Public Health Service predoctoral training grant GM 07134-13 from the National Institutes of Health.

LITERATURE CITED

- Allison, D. S., S. H. Goh, and B. D. Hall. 1983. The promoter sequence of a yeast tRNA^{Val} gene. *Cell* 34:655-664.
- Baker, R. E., S. Camier, A. Sentenac, and B. D. Hall. 1987. Gene size differentially affects the binding of yeast transcription factor τ to two intragenic regions. *Proc. Natl. Acad. Sci. USA* 84: 8768-8772.
- Bark, C., P. Weller, J. Zabielski, L. Janson, and U. Petterson. 1987. A distant enhancer element is required for polymerase III transcription of a U6 RNA gene. *Nature (London)* 328:356-359.
- Bogenhagen, D. F., W. M. Wormington, and D. D. Brown. 1982. Stable transcription complexes of *Xenopus* 5S RNA genes: a means to maintain the differentiated state. *Cell* 28:413-421.

5. Buc, H., and W. R. McClure. 1985. Kinetics of open complex formation between *Escherichia coli* RNA polymerase and the *lac* UV5 promoter. Evidence for a sequential mechanism involving three steps. *Biochemistry* **24**:2712-2723.
6. Camier, S., O. Gabrielsen, R. Baker, and A. Sentenac. 1985. A split binding site for transcription factor τ on the tRNA₃^{Glu} gene. *EMBO J.* **4**:491-500.
7. Carbon, P., S. Murgu, J.-P. Ebel, A. Krol, G. Tebb, and I. W. Mattaj. 1987. A common octamer motif binding protein is involved in the transcription of U6 snRNA by RNA polymerase III and U2 snRNA by RNA polymerase II. *Cell* **51**:71-79.
8. Ciliberto, G., G. Raugel, F. Costanzo, L. Dente, and R. Cortese. 1983. Common and interchangeable elements in the promoters of genes transcribed by RNA polymerase III. *Cell* **32**:725-733.
9. Clos, J., D. Buttgerit, and I. Grummt. 1986. A purified transcription factor (TIF-1B) binds to essential sequences of the mouse rDNA promoter. *Proc. Natl. Acad. Sci. USA* **83**:604-608.
10. Das, G., D. Henning, D. Wright, and R. Reddy. 1988. Upstream regulatory elements are necessary and sufficient for transcription of a U6 RNA gene by RNA polymerase III. *EMBO J.* **7**:503-512.
11. DeFranco, D., O. Schmidt, and D. Söll. 1980. Two control regions for eukaryotic tRNA gene transcription. *Proc. Natl. Acad. Sci. USA* **77**:3365-3368.
12. DeFranco, D., S. Sharp, and D. Söll. 1982. Identification of regulatory sequences contained in the 5'-flanking region of *Drosophila* lysine tRNA₂ genes. *J. Biol. Chem.* **256**:12424-12429.
13. Dingermann, T., D. J. Burke, S. Sharp, J. Schaack, and D. Söll. 1982. The 5' flanking sequences of *Drosophila* tRNA^{Arg} genes control their *in vitro* transcription in a *Drosophila* cell extract. *J. Biol. Chem.* **257**:14738-14744.
14. Dubendorff, J. W., P. L. deHaseth, M. S. Rosendahl, and M. H. Caruthers. 1987. DNA functional groups required for formation of open complexes between *Escherichia coli* RNA polymerase and the λ P_R promoter. Identification via base analog substitutions. *J. Biol. Chem.* **262**:892-898.
15. Dynan, W. S., and R. Tjian. 1983. The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell* **35**:79-87.
16. Engelke, D. R., S.-Y. Ng, B. S. Shastri, and R. G. Roeder. 1980. Specific interaction of a purified transcription factor with an internal control region of 5S RNA genes. *Cell* **19**:717-728.
17. Fuhrman, S. A., D. R. Engelke, and E. P. Geiduschek. 1984. HeLa cell RNA polymerase III transcription factors. Functional characterization of a fraction identified by its activity in a second template rescue assay. *J. Biol. Chem.* **259**:1934-1943.
18. Galli, G., H. Hofstetter, and M. L. Birnstiel. 1981. Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. *Nature (London)* **294**:626-631.
19. Garcia, A. D., A. M. O'Connell, and S. J. Sharp. 1987. Formation of an active transcription complex in the *Drosophila melanogaster* 5S RNA gene is dependent on an upstream region. *Mol. Cell. Biol.* **7**:2046-2051.
20. Gouilloud, E., and S. G. Clarkson. 1986. A dispersed tyrosine tRNA gene from *Xenopus laevis* with high transcriptional activity *in vitro*. *J. Biol. Chem.* **261**:486-494.
21. Hipkind, R. A., and S. G. Clarkson. 1983. 5'-Flanking sequences that inhibit *in vitro* transcription of a *Xenopus laevis* tRNA gene. *Cell* **34**:881-890.
22. Huibregtse, J. M., C. F. Evans, and D. R. Engelke. 1987. Comparison of tRNA gene transcription complexes formed *in vitro* and in nuclei. *Mol. Cell. Biol.* **7**:3212-3220.
23. Lassar, A. B., P. L. Martin, and R. G. Roeder. 1983. Transcription of class III genes: formation of preinitiation complexes. *Science* **222**:740-748.
24. Lergerski, R. J., J. L. Hodnett, and H. B. Gray. 1978. Extracellular nucleases of *Pseudomonas* Bal 31. III. Use of the double-strand deoxyriboexonuclease activity as the basis of a convenient method for the mapping of fragments of DNA produced by cleavage with restriction enzymes. *Nucleic Acids Res.* **5**:1445-1464.
25. Lofquist, A., and S. Sharp. 1986. The 5'-flanking sequences of *Drosophila melanogaster* tRNA^{Asn} genes differentially arrest RNA polymerase III. *J. Biol. Chem.* **260**:14600-14606.
26. McKnight, S. L., and R. Kingsbury. 1982. Transcriptional control signals of a eukaryotic protein-coding gene. *Science* **217**:316-324.
27. Morry, M. J., and J. D. Harding. 1986. Modulation of transcriptional activity and stable complex formation by 5'-flanking regions of mouse tRNA^{His} genes. *Mol. Cell. Biol.* **6**:105-115.
28. Morton, D. G., and K. U. Sprague. 1982. Silkworm 5S RNA and alanine tRNA genes share highly conserved 5'-flanking and coding sequences. *Mol. Cell. Biol.* **2**:1524-1531.
29. Morton, D. G., and K. U. Sprague. 1984. *In vitro* transcription of a silkworm 5S RNA gene requires an upstream signal. *Proc. Natl. Acad. Sci. USA* **81**:5519-5522.
30. Murphy, S., C. Di Liegro, and M. Melli. 1987. The *in vitro* transcription of the 7SK RNA gene by RNA polymerase III is dependent only on the presence of an upstream promoter. *Cell* **51**:81-87.
31. Railey, J. F., II, and G.-J. Wu. 1988. Organization of multiple regulatory elements in the control region of the adenovirus type 2-specific VARNA1 gene: fine mapping with linker-scanning mutants. *Mol. Cell. Biol.* **8**:1147-1159.
32. Raymond, K. C., G. J. Raymond, and J. D. Johnson. 1985. *In vivo* modulation of yeast tRNA gene expression by 5'-flanking sequences. *EMBO J.* **4**:2659-2656.
33. Rubacha, A., W. Sumner III, L. Richter, and K. Beckingham. 1984. Conserved 5' flank homologies in dipteran 5S RNA genes that would function on 'A' form DNA. *Nucleic Acids Res.* **12**:8193-8207.
34. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
35. Schaack, J., S. Sharp, T. Dingermann, D. J. Burke, L. Cooley, and D. Söll. 1984. The extent of a eukaryotic tRNA gene. 5'- and 3'-flanking sequence dependence for transcription and stable complex formation. *J. Biol. Chem.* **259**:1461-1467.
36. Schaack, J., S. Sharp, T. Dingermann, and D. Söll. 1983. Transcription of eukaryotic tRNA genes *in vitro*. II. Formation of stable complexes. *J. Biol. Chem.* **258**:2447-2453.
37. Selker, E. U., E. Morzycka-Wroblewska, J. N. Stevens, and R. L. Metzberg. 1986. An upstream signal is required for *in vitro* transcription of *Neurospora* 5S RNA genes. *Mol. Gen. Genet.* **205**:189-192.
38. Sentenac, A. 1985. Eukaryotic RNA polymerases. *CRC Crit. Rev. Biochem.* **18**:31-90.
39. Sharp, S., D. DeFranco, T. Dingermann, P. Farrell, and D. Söll. 1981. Internal control regions for transcription of eukaryotic tRNA genes. *Proc. Natl. Acad. Sci. USA* **78**:6657-6661.
40. Sharp, S. J., J. Schaack, L. Cooley, D. J. Burke, and D. Söll. 1985. Structure and transcription of eukaryotic tRNA genes. *CRC Crit. Rev. Biochem.* **19**:107-144.
41. Shaw, K. J., and M. V. Olson. 1984. Effects of altered 5'-flanking sequences on the *in vivo* expression of a *Saccharomyces cerevisiae* tRNA^{Tyr} gene. *Mol. Cell. Biol.* **4**:657-665.
42. Siebenlist, U., R. B. Simpson, and W. Gilbert. 1980. *E. coli* RNA polymerase interacts homologously with two different promoters. *Cell* **20**:269-281.
43. Sollner-Webb, B. 1988. Surprises in polymerase III transcription. *Cell* **52**:153-154.
44. Thimmappaya, B., N. Jones, and T. Shenk. 1979. A mutation which alters initiation of transcription by RNA polymerase III on the Ad5 chromosome. *Cell* **18**:947-954.
45. Tower, J., V. C. Culotta, and B. Sollner-Webb. 1986. Factors and nucleotide sequences that direct ribosomal DNA transcription and their relationship to the stable transcription complex. *Mol. Cell. Biol.* **6**:3451-3462.
46. Tyler, B. M. 1987. Transcription of *Neurospora crassa* 5S rRNA genes requires a TATA box and three internal elements. *J. Mol. Biol.* **196**:801-811.
47. Ullu, E., and A. M. Weiner. 1985. Upstream sequences modulate the internal promoter of the human 7SL RNA gene. *Nature*

- (London) 318:371-374.
48. **Van Dyke, M. W., and R. G. Roeder.** 1987. Multiple proteins bind to VA RNA genes of adenovirus type 2. *Mol. Cell. Biol.* 7: 1021-1031.
 49. **Wilson, E. T., D. Larson, G. I. Hall, and K. U. Sprague.** 1985. A large region controls tRNA gene transcription. *J. Mol. Biol.* 183: 153-163.
 50. **Yoshinaga, S. K., P. A. Boulanger, and A. J. Berk.** 1987. Resolution of human transcription factor TFIIC into two functional components. *Proc. Natl. Acad. Sci. USA* 84:3585-3589.
 51. **Young, L. S., N. Takahashi, and K. U. Sprague.** 1986. Upstream sequences confer distinctive transcriptional properties on genes encoding silk gland-specific tRNA^{Ala}. *Proc. Natl. Acad. Sci. USA* 83:374-378.