TRANSCRIPTION AND PROCESSING OF TRANSFER RNA IN *S. CEREVISIAE*

Thesis by

Vicente Mendoza Reyes

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1988

(Submitted December 21, 1987)

ACKNOWLEDGMENTS

There are several persons who have made my stay at Caltech worth all those five, long years. First and foremost, I would like to thank my Ph.D. committee chairman, Professor Norman Davidson, for his sound advice, concern, and, most important of all, encouragement, without which I might not have lasted as long as I did. I also thank John Abelson for the general direction of this thesis and for financial support. I also would like to acknowledge several members of our laboratory who have substantially contributed to the overall development of this work: particularly, I would like to thank Calvin Ho, Kyle Tanner, and Reinhard Rauhut for valuable scientific discussions; Shawn Westaway, Jennifer Normanly and Usha Vijayraghavan for important technical suggestions; and Gloria McFarland and Soo-Chen Cheng for helping me recall my enzyme kinetics and radiochemical calculations, respectively.

I am also lucky to have a number of small, but very close circles of friends inside and outside Pasadena with whom I share many fond memories. I thank my friends from U. C. Irvine, Rosie Panganiban, Ces Veto, Edith Dizon and friends for occassional parties and get-togethers; Tony Reyes, Chi Goh, Purit, Jonnelle and Claude, for the Sunday afternoon bridge games and birthday celebrations; and Danny Casimiro, for introducing me to his nice relatives and family friends in the Norwalk area. I also thank very much my uncle and aunt, Mr. and Mrs. Rene and Glo Mendoza of Los Angeles, for providing me with a home away from home. I also thank and express my deep appreciation for my voice teacher, Miss Margaret Keane, who not only provided me with inspiration through music, but also moral support—both very important ingredients which kept me going. Through her and our always-enjoyable Saturday evening workshops, I met several good friends, namely, Mr. and Mrs. Manny and Libby Fernando, Hong Chuan, Lilian and David, Laurine, Naomi, Harold, and the rest of Margaret's nice students.

I also would like to specially thank the staff of the Church Biology Office, especially Ms. Renée Thorf, for excellent word processing.

Finally, I would like to share the success of this work with my family: my parents, Luis and Tita Reyes, and my brothers, Ponchit and Ed.

ABSTRACT

A collection of studies on tRNA transcription and splicing in the yeast *S. cerevisiae* are presented. These studies employ a combination of recombinant DNA, oligonucleotide-directed mutagenesis and *in vitro* synthetic technologies applied using the reverse genetics approach. Chapter I introduces the reader to these topics. In Chapter II, an attempt to solve the puzzle of tandem tRNA gene transcription in yeast is described. This tandem, the *S. cerevisae* $tRNA^{Arg}$-$tRNA^{Asp}$ gene pair, is transcribed solely by use of the upstream gene promoter signals, giving rise to a dimeric precursor which is processed into two mature tRNA molecules, much like prokaryotic systems. A collection of specific point and deletion mutations were constructed to answer the question of why the downstream $tRNA^{Asp}$ gene is apparently inactive. Our results show that it is so only in this configuration; the $tRNA^{Arg}$ and spacer sequences, which constitute its upstream flanking sequences in this arrangement, seem inhibitory to the independent activity of the $tRNA^{Asp}$ gene. Taken together, these results emphasize the importance of flanking regions in eukaryotic tRNA gene transcription. In Chapter III, the construction and characterization of a heterologous system for the *in vitro* synthesis of pre-tRNA are presented. A strong bacteriophage T7 promoter was fused to a *S. cerevisiae* pre-$tRNA^{Phe}$ gene. We show that pre-$tRNA^{Phe}$ is synthesized efficiently from this system by the cognate T7 RNA polymerase, and that this RNA has the correct sequence, mature terminii, and is spliced efficiently and accurately by our *in vitro* splicing system, which consists of highly purified tRNA splicing endonuclease and ligase enzymes. In Chapter IV, an extensive investigation on tRNA splicing substrate specificity is described. A collection of 15 carefully-designed mutant pre-$tRNA^{Phe}$ genes were constructed, and then transcribed and analyzed as above.

We find that the endonuclease recognizes two highly conserved, surface residues in pre-tRNA^{Phe}-U8 and C56-and probably contacts these bases during the splicing reaction. We also find that splice site selection by this enzyme is a function of the length of the anticodon stem, and thus proceeds by a simple distance measurement mechanism. We have evidence that this measuring process commences in the thoracic region of the pre-tRNA, where the endonuclease probably binds. Finally, we demonstrate that the highly conserved purine residue 3' proximal to the anticodon may be important for cleavage at the nearby 5' splice site.

TABLE OF CONTENTS

Chapter I

INTRODUCTION

Transcription of Transfer RNA in S. cerevisiae

The *S. cerevisiae* haploid genome contains approximately 360 tRNA genes and the great majority of these occur as highly dispersed and independent transcriptional units (Guthrie and Abelson, 1982). This lack of clustering has been shown by UV light mapping techniques (Feldman, 1977), as well as gene cloning experiments (Beckmann *et al.*, 1977a). In the latter, out of 4,000 different yeast genomic clones, only 187 hybridized to tRNA-specific probes. This implied that there is roughly one tRNA gene per 2.0 kb of yeast DNA, a highly dispersed arrangement (Beckmann *et al.*, 1977b; Beckmann *et al.*, 1979). However, tRNA gene dispersal is not a general feature among eukaryotic organisms. In *Drosophila* and *Xenopus*, most tRNA genes are tightly linked, whereas in silkworm and nematode systems, these genes are dispersed, as in yeast (Johnson, 1983). Only three cases of tight clustering of tRNA genes are known in yeast systems. In *S. cerevisiae*, there is a $tRNA^{Arg}$-$tRNA^{Asp}$ gene tandem separated by 10 bp (Schmidt *et al.*,1980), and in *S. pombe*, there is a $tRNA^{Ser}$-$tRNA^{Met}$ gene tandem 7 bp apart (Mao *et al.*, 1980). The *S. cerevisiae* mitochondrial DNA, which codes for approximately 20 tRNA genes, contains a $tRNA^{Arg}$-$tRNA^{Ser}$ gene tandem separated only by 3 bp (Abelson, 1980). In the first two cases, it is known that transcription from the gene tandem is controlled solely by the first gene (Kjellin-Sträby *et al.*, 1984; Mao *et al.*, 1980), and that the primary RNA product is a dimeric precursor which is known to undergo a sequence of processing reactions to produce the two mature tRNAs (Mao *et al.*, 1980; Pearson *et al.*, 1985; Engelke *et al.*, 1985). Through precise deletion of the $tRNA^{Arg}$ gene and spacer sequences in the $tRNA^{Arg}$-$tRNA^{Asp}$ gene pair, it has been shown that the downstream $tRNA^{Asp}$ gene can be activated (Reyes *et al.*, 1986). This is in contrast to a set of identical *C. elegans* $tRNA^{Pro}$ genes artificially placed end to end in mono-, di-, and trimeric arrangements, and where in each case, each individual gene was independently active in transcription (Ciliberto *et al.*, 1984).

Eukaryotic tRNA genes are transcribed by RNA polymerase III (pol III) (for a review, see Ciliberto *et al.*, 1983a). Sequences necessary for promotion of transcription of, as well as binding of transcription factors to, tRNA genes, reside within the gene coding sequence as two noncontiguous, 10 bp-blocks called the A- and B-blocks, or 5' and 3' ICRs, respectively (Stewart *et al.*, 1985; Baker and Hall, 1984; DeFranco *et al.*, 1980; Sharp *et al.*, 1982; Folk and Hofstetter, 1983). By analysis of a series of 5' and 3' deletion mutant tRNA genes (Hofstetter *et al.*, 1981; Sharp *et al.*, 1981 Allison *et al.*, 1983), the borders of these internal promoter sequences have been mapped to positions 8-25 and 50-60, respectively (exact position depends on the particular gene). Thus the A-block maps to the upper half of the D-arm, and the B-block to the TψC loop region (Hall *et al.*, 1982; Newman *et al.*, 1983, Ciliberto *et al.*, 1983a; Clarkson, 1983). These two sequence blocks can tolerate reasonably wide variations in separation, with an optimum of 35-45 bp; above and below this range, transcription efficiency tends to fall off (Ciliberto *et al.*, 1982a; Hofstetter *et al.*, 1981). Increase in distance is particularly tolerated since approximately a tenth of *S. cerevisiae* tRNA genes contain introns which further separate the two sequence blocks; these introns range in size from 14-60 bp, yet these genes are known to be transcriptionally active. In some cases, a third short region of consensus sequence AGGTC and mapping to the extra-loop of the tRNA, is necessary for optimal transcriptional efficiency (Allison *et al.*, 1983; Willis *et al.*, 1984). The A- and B-blocks appear to act as independent control elements in spite of the overall regularities of the tRNA molecule; for example, a hybrid tRNA gene, constructed by fusing the 5' half of a $tRNA^{Pro}$ gene with the 3' half of a $tRNA^{Leu}$ gene, both from *C. elegans*, exhibited normal transcriptional phenotype *in vitro* (Ciliberto *et al.*, 1982b; Galli *et al.*, 1981). The above general scenario resembles that of the 5S RNA gene, which is also transcribed by pol III (Birkenmeier *et al.*, 1978). In this gene, the

promoter elements are known to be composed of three closely-spaced intragenic sequences: an A-block homolog, a short intermediate element, and a third block designated C, which is thought to be the main recognition element by the well-known 5S-specific transcription factor, TFIIIA (Pieler *et al.*, 1987). Additionally, two hybrid pol III genes, one consisting of the 5' half of a *C. elegans* $tRNA^{Pro}$ gene and the 3' half of a *X. borealis* 5S RNA gene, and the other, the converse arrangement, are both efficiently transcribed by pol III in *X. laevis* oocytes (Ciliberto *et al.*, 1983b). Finally, some (e.g., *E. coli* $tRNA^{Asp}$ and $tRNA^{Trp}$ genes), but not all (e.g., *E. coli* $tRNA^{Tyr}$ gene), bacterial tRNA genes can be transcribed by the *X. laevis* system, which suggests that the forces that maintain tRNA structure and function in bacterial tRNA genes are by themselves sufficient to preserve ICR sequences that can be recognized by eukaryotic pol III (Folk *et al.*, 1982).

Although tRNA genes possess intragenic promoter elements, 5' and 3' flanking sequences greatly influence their *in vitro* and *in vivo* expression (Raymond *et al.*, 1985; Allison and Hall, 1985). In the *Drosophila* $tRNA_2^{Arg}$ gene, sequences up to 60 bp upstream and 35 bp downstream of the coding sequence are required for optimal levels of transcription (Sharp *et al.*, 1983; Schaack *et al.*, 1984). In the $tRNA_2^{Lys}$ gene, on the other hand, an inhibitory 11-base sequence in the 5' flanking region, GGCAGTTTTTG, was detected by successive resection of flanking sequences and analysis of transcriptional activity (DeFranco *et al.*, 1981). Similarly, two sequence blocks, each consiting of alternating purines and pyrimidines, were shown to repress transcription of a *Xenopus* $tRNA_i^{Met}$ gene (Hipskind *et al.*, 1983). This last result is particularly interesting since such a sequence has been previously implicated in the formation of Z-DNA structures (Kolata, 1983). However, the existence of such inhibitory 5'-flanking oligonucleotides are the exception rather than the rule, and could simply play

regulatory roles in their natural setting rather than inhibition *per se*. A great majority of tRNA genes thus require the normal 5' and 3' flanking regions for full transcriptional efficiency *in vitro* (Raymond and Johnson, 1983; Johnson and Raymond, 1984) and optimal function *in vivo* (Shaw and Olson, 1984). In the *in vivo* case, particularly, suppression efficiency correlated directly with the amount of wild-type sequences that flank the $tRNA^{Tyr}$ SUP4-o gene (Shaw and Olson, 1984; Allison and Hall, 1985). Upstream sequences have also been shown to play a regulatory role in tissue-specific expression of a silkworm $tRNA^{Ala}$ gene (Sprague *et al.*, 1980; Sprague *et al.*, 1987). In this system, the constitutive $tRNA^{Ala}$ gene (C-ala) is transcribed in all cell types, but during rapid silk production, a second $tRNA^{Ala}$ gene (SG-ala) is produced at greatly elevated levels by the silkgland. Yet, these two $tRNA^{Ala}$ genes differ only by a C to T transition in the anticodon stem (Young *et al.*, 1986). Construction of $tRNA^{Ala}$ gene chimeras from C-ala and SG-ala by assembling upstream, coding and downstream sequences in different combinations, has shown that the 5' flanking sequence is the element responsible for the difference in expression of these two genes (*Ibid.*).

A 34-bp yeast repetitive element called *sigma* is also found in close association with certain tRNA genes in the yeast genome (Del Rey *et al.*, 1982). These elements are found 16-18 bp upstream of some, but not all, *S. cerevisiae* tRNA genes. Although there is presently no evidence that *sigma* can in fact transpose, an interesting possibility if it does is that it recognizes tRNA-like sequences or structures in the genome for transposition.

In all known cases, transcription of tRNA genes initiates at a purine residue in the immediate 5' flanking region of the gene (exact position depends on the particular gene), and terminates within the first stretch of four or more T residues in the 3' flanking region (Schaack *et al.*, 1984; for a review, see Sharp *et al.*, 1986). As a result, primary tRNA transcripts always contain extra sequences at

their extremities (and some contain introns in the middle of the tRNA sequence) which have to be enzymatically deleted during the course of tRNA biosynthesis. In the case of the $tRNA_3^{Leu}$ gene family, transcription initiation has been associated with the second purine residue in the 15-mer consensus upstream sequence, TTTCAACAAATAAGT (Raymond *et al.*, 1985).

Systematic study of tRNA gene transcription was made possible by the development of several *in vitro* pol III transcription systems derived from crude cell extracts that could faithfully and accurately transcribe cloned tRNA genes. Such systems are now available from *Xenopus* (Ciliberto *et al.*, 1982a), *S. cerevisiae* (Weil *et al.*, 1979; Koski *et al.*, 1982; Klekamp and Weil, 1982), HeLa cells (Strandring *et al.*, 1981), *Drosophila* (Dingermann *et al.*, 1981), *C. elegans* (Honda *et al.*, 1986), silkworm (Sprague *et al.*, 1980), and human cells (Zasloff *et al.*, 1982c). Characterization of such systems by chromatographic fractionation into several different activities is currently the subject of intense investigation.

A hallmark of eukaryotic cells is the presence of three different RNA polymerases, termed I, II and III, classed according to the type of genes they transcribe (Sklar *et al.*, 1976; Sentenac, 1985). Pol I transcribes rRNA genes, pol II, protein-coding genes, and pol III, tRNA, 5S RNA, low-molecular weight viral RNAs, and a number of other genes. These three RNA polymerases differ in their chromatographic properties (Sentenac and Hall, 1982) and α-amanitin sensitivities, with pol II being the most sensitive, and pol III, most resistant, to the drug (Hager *et al.*, 1976). Each is a high-molecular weight, multisubunit complex made up of approximately ten polypeptides, some of which are shared between two, and some by all three RNA polymerases (Sentenac and Hall, 1982). In *S. cerevisiae*, pol I is 600 kdal, pol II, 540 kdal, and pol III, 390 kdal, and all require Zn^{++} as cofactor (*Ibid.*). Previously, the involvement of DNA topoisomerase II in pol III transcription was implicated, and the anti-DNA topoisomerase antibiotics,

novobiocin and coumermycin was shown to inhibit pol III transcription as well. More recently, however, it has been shown that pol III transcription inhibition by these two antibiotics does not involve DNA topoisomerase II (Gottesfeld, 1986).

Purified RNA polymerase III by itself is unable to transcribe tRNA or 5S RNA genes; certain DNA-binding proteins called transcription factors (TFs) are required as well for precise transcription of these genes *in vitro* (Shastry *et al.*, 1982). By chromatographic fractionation of crude cell extracts which can accurately transcribe class III genes, at leat three pol III transcription factors are presently recognized, and are termed TFIIIA, B, and C (Sentenac and Hall, 1982; Shastry *et al.*, 1982). TFIIIA is a 38 kdal protein and is a 5S RNA gene-specific initiation factor known to bind to both 5S RNA and DNA (Pieler *et al.*, 1987). Binding to the former gives rise to a 7S ribonucleoprotein complex and is presumed to have storage function. From its amino acid sequence, TFIIIA has been predicted to contain nine loop-like domains or "fingers", each stablized by a zinc atom; this is now popularly known as the "finger motif", and which might be a general feature among eukaryotic regulatory DNA-binding proteins (Vincent, 1986). TFIIIB and C bind to both tRNA and 5S RNA genes (Lassar *et al.*, 1983). At 300 kdal, TFIIIC is a large protein and has a very high affinity for the B-block promoter element as shown by DNAse I protection experiments and transcription-competition assays. TFIIIC, and another factor called τ, purified from another laboratory, are presumed to be one and the same (Sylvie *et al.*, 1985). This factor is known to protect not only the B-block, but almost the entire gene, from DNAse I (Baker *et al.*, 1986; Stillman and Geiduschek, 1984). From methylation studies, several contact points which cluster within the B-block are now known. Mutations in both the A- and B-blocks, as well as in the 5' and 3' flanking regions, are known to diminish its affinity with the template, as shown by transcription-competition experiments. By direct electron microscopic examination, this factor

has been visualized; it is a large, structurally isometric particle and apparently causes a drastic structural disruption of the tRNA gene on binding, whereupon the gene is almost bent into a U (Stillman *et al.*, 1985a). This observation is quite interesting since it offers an explanation whereby the two tRNA promoter blocks could be brought closer together during the transcription reaction.

Binding of TFIIIC to the tRNA gene is a well-studied process. Stability of binding is diminished at high ionic concentrations, and favored in the presence of nonpolar solvents like DMSO; for example, 10% DMSO can offset the effects of 50-60 mM added NaCl (Stillman *et al.*, 1984). At elevated temperatures, binding is also stabilized, suggesting that melting of the template is favorable (Stillman *et al.*, 1985b). In the presence of single-stranded DNA, footprint analyses show that binding to the A-block and neighborhood diminishes, whereas the B-block footprint is not affected; however, the overall stability of the binding reaction is diminished. Through selective proteolysis, this factor was shown to consist of two domains, τ_B and τ_A. τ_B has a very high affinity for the B-block, while τ_A interacts with the A-block less tightly; in fact, τ can bind to the B-block even after the removal of its τ_A domain (Marzouki *et al.*, 1986). The ability of the intact protein to bind to two noncontiguous regions of tRNA genes which are variably spaced (i.e., IVS-containing versus intronless tRNA genes) has led to the hypothesis that τ_A and τ_B might be connected by a flexible arm or hinge structure, which itself is accessible to proteolysis (*Ibid.*).

The third pol III transcription factor, TFIIIB, is the least characterized of the three factors. It is a 60 kdal protein and thought to interact with the DNA template primarily through protein-protein interactions with factors already bound to the gene (Weil and Klekamp, 1986). Successive binding of these factors to the tRNA gene results in the formation of protein-DNA preinitiation complexes of varying stabilities. It is currently thought that TFIIIC, through its B-domain,

initially binds rapidly and reversibly to the B-block, producing a metastable complex. The bound factor then slowly isomerizes and its A-domain then interacts with the A-block to produce a stable complex (Stillman *et al.*, 1985b). TFIIIB then associates with this complex, probably mainly through interaction with TFIIIC to form a stable preinitiation complex, which is then ready for transcription by pol III in the presence of the four NTPs (Lassar *et al.*, 1983). Thus, the actual template for pol III transcription is not a naked tRNA gene, but the preinitiation complex consisting of the gene plus bound factors; this has been kinetically and biochemically substantiated (Lassar *et al.*, 1983; Schaack *et al.*, 1983).

Mutations within the coding as well as flanking regions of the tRNA gene affect a whole spectrum of steps in tRNA biosynthesis, from transcription factor binding to intron excision. Most studies in this area have employed *in vitro* transcription systems from *Xenopus* and HeLa cell extracts. Point mutations which affect transcription efficiency and factor binding stability cluster within the two ICRs and the extra loop promoter element, but wild-type sequences in the flanking regions also exert influence to a large degree. In one experiment, a $tRNA_3^{Leu}$ gene with an insertional mutation in the IVS and truncated in that region and therefore lacks the B-block, is still transcriptionally active, implying that the B-block is dispensable for template activity in this system (Carrara *et al.*, 1981). In the same, but intact, gene, a three-base mutation in the A-block that disrupts the complementarity within the D-stem does not abolish transcription, but severely diminishes 5'-end maturation and IVS excision (Mattoccia *et al.*, 1983). In a study involving a $tRNA^{Pro}$ gene from *C. elegans*, a direct correlation was observed between primary transcript processibility and stability in the extract (Traboni *et al.*, 1984). In a more recent finding, inhibition of transcription by base mismatches within, but not outside, the ICRs, has been interpreted as an

indication of the involvement of repeated DNA denaturation and renaturation during the transcription process, much like prokaryotic RNA polymerase systems (Sullivan and Folk, 1987). More recently, it has been found that a $tRNA^{Glu}$ gene with a deleted A-block is inactive in dilute KCl solutions, but is transcribed if KOAc is instead present, suggesting that transcriptional requirement for the A-block depends on the ionic environment of the gene (Gabrielsen and Øyen, 1987).

Processing of Transfer RNA in *S. cerevisiae*

RNA processing is the collection of enzymatic reactions that transform the primary RNA transcript of a gene into a mature, biologically functional molecule (Abelson, 1979). RNA splicing is the subcategory of RNA processing that refers to the removal of introns from the RNA transcript of split genes. Three classes of RNA splicing are generally recognized: class I RNA splicing involves nuclear-coded mRNA (Busch *et al.*, 1982; Flint, 1974; Sharp *et al.*, 1984; Green, 1986; Padgett *et al.*, 1986); class II, rRNA and some fungal mitochondrial RNA (Sharp, 1985; Cech and Bass, 1986); and class III, tRNA. Two crucial questions always arise from the general problem of RNA splicing (Abelson *et al.*, 1986), namely: (a) how are the splice junctions brought into correct alignment to ensure correct cleavage and ligation; and (b) what are the biochemical mechanisms of the cleavage and ligation reactions?

In the case of yeast tRNA genes, it is known that in all cases, transcription initiation and termination occur at the 5' and 3' flanking sequences, respectively, and therefore the primary transcript has short extra sequences at either end which are enzymatically deleted during the tRNA maturation pathway. Further, in approximately a tenth of the 360 tRNA genes in *S. cerevisiae*, there is an intervening sequence near the middle of the gene—specifically, one base 3' to the anticodon (Kang *et al.*, 1979; Kang *et al.*, 1980; Ogden *et al.*, 1984; Valenzuela

et al., 1980)—which must be removed at the RNA level by a set of excision and ligation reactions known as tRNA splicing (Knapp *et al.*, 1979; Peebles *et al.*, 1979; Abelson *et al.*, 1981; Knapp *et al.*, 1978; Ogden *et al.*, 1979; Ogden *et al.*, 1980). While these processes take place, certain bases and ribose residues in the tRNA are covalently modified by addition or alteration of certain functional groups, in a series of steps collectively designated as base modification reactions (Kohli, 1983).

The study of tRNA splicing in *S. cerevisiae* has been facilitated by the discovery of a temperature-sensitive mutant strain, designated *ts*136, which accumulates pre-tRNA-sized molecules, as well as pre-mRNA and pre-rRNA in the nucleus at the nonpermissive temperature. This mutation defines the *rna*1 genetic locus and the defect is generally thought to be in the process of RNA transport from the nucleus to the cytoplasm (Hopper *et al.*, 1978). A related mutation, *los*1, results in the accumulation of only pre-tRNA-sized molecules (Hurt *et al.*, 1987). The identity and structure of each pre-tRNA that accumulate in *ts*136 is now known; the group is a defined subset of tRNAs in the cell, which possess mature 5' and 3'-ends but still contain the IVS (Hopper, 1984; Deutscher, 1984). End-mature precursors to Tyr, Phe, Lys, Pro, Trp, Ser, Leu (UAG and CAA) and Ile tRNAs accumulate; in the cases of Tyr, Phe, Pro, and both Leu tRNAs, there are slight sequence heterogeneities in the IVS (Ogden *et al.*, 1984). The defect in *ts*136 probably does not involve the splicing enzymes directly because extracts prepared from this strain efficiently splices pre-tRNAs at the permissive temperature. Analyses of these *in vivo* precursors have revealed a wealth of information. We now know that IVS from different tRNA isoaccepting families are extremely heterogeneous in size (13-60 bases) and sequence; the only constant feature they possess is their location (one base 3' to the anticodon), base composition (A+U-rich, with ratio A+U:G+C of approximately 2:1), and possession

of a sequence complementary to the anticodon and its immediate vicinity, suggesting that these sequences probably form a stem structure in the pre-tRNA There are no conserved sequences at the splice junctions. In contrast, *S. pombe* pre-tRNA introns do not contain sequences which are complementary to the anticodon (Gamulin *et al.*, 1983).

tRNA splicing is a two-stage reaction mediated by two distinct and separable enzymes: an endonuclease to locate and cleave the 5' and 3' splice sites, producing a free IVS and two annealed tRNA halves; and a ligase, to join the two tRNA halves at the nicked junction, producing an intact tRNA (Peebles *et al.*, 1979; Peebles *et al.*, 1983; Greer *et al.*, 1983b). A detailed mechanism for tRNA splicing in *S. cerevisiae* has been proposed (Ogden *et al.*, 1981; Greer *et al.*, 1983b). First, the endonuclease cleaves the pre-tRNA at the 5' and 3' splice sites to produce a 5' half tRNA with a 2',3'-cyclic phosphate terminus, a 3' half tRNA with a 5' OH terminus, and a free IVS with both of the above termini; the 5' and 3' halves are presumed to remain annealed. Then, the cylic phosphodiesterase activity of ligase cleaves the 2',3'-cyclic phosphate terminus of the 5' half tRNA to produce a 2' phosphate terminus, and its polynucleotide kinase activity phosphorylates the 5' OH terminus of the 3' half tRNA by transfer of the γ-phosphate of an ATP cofactor, producing a 5' phosphate end. The ligase protein then itself becomes adenylylated, with the release of a pyrophosphate corresponding in the γ and β phosphates of an ATP cofactor. The adenylyl residue, consisting of the α-phosphate and adenosine, is then transferred to the 5' phosphate terminus of the 3' half, producing an activated 3' half tRNA with a 5'-5'-phosphoanhydride bond. This activated species is then ligated to the 5' half accompanied by the release of the previously added adenylyl residue and formation of a mature tRNA possessing a 2' phosphate at the splice junction, and whose phosphate at the 3',5'-phosphodiester bond there was derived from the γ-

position of an ATP cofactor. Finally, a 2' phosphatase activity in the cell, believed to be not associated with ligase, deletes the 2' phosphate at the splice junction.

Processing of tRNA has also been characterized in mammalian systems. In HeLa cells, for example, transcription and processing can be readily uncoupled at low Mg^{++} concentrations, in which tRNA genes are transcribed but not processed; at relatively high Mg^{++} concentrations, both reactions take place (Laski *et al.*, 1983). Cleavage at the splice sites also produces 5' OH and 3' phosphate termini (possibly 2',3'-cyclic phosphate as in yeast), but the major difference from yeast systems is that the 3' phosphate is retained at the splice junction after ligation; in yeast, this phosphate comes from the γ-position of the ATP cofactor (*Ibid.*). Further, some mammalian tRNAs are first matured at the 5' end (e.g., mouse $tRNA^{Gly}$), while others are end-matured in the reverse order (e.g., mouse $tRNA^{His}$); thus, there probably exist two distinct end-maturation pathways in these systems (Rooney and Harding, 1986).

In *X. laevis*, both the 5'- and 3'-end maturation enzymes have been extensively purified and characterized and have been designated as 5'- and 3'-pre-tRNAses, respectively (Castaño *et al.*, 1985; Castaño *et al.*, 1986). The 3'-pre-tRNAse is a single polypeptide of 98 kdal and requires a 5' end-mature, 3' trailer-containing pre-tRNA as substrate; reaction produces an intact 3' trailer sequence and the end-mature tRNA (Castaño, *et al.*, 1985). In *Drosophila*, a similar scenario is evident, wherein 3' end-maturation is endonucleolytic and occurs after 5'-end maturation (Frendewey *et al.*, 1985; Willis *et al.*, 1986). In contrast, the *S. cerevisiae* 3' pre-tRNAse appears to be exonucleolytic (Willis *et al.*, 1986; Engelke *et al.*, 1985). The *X. laevis* 5'-pre-tRNAse has a very complex structure, being made up of at least 14 polypeptides with molecular weights ranging from 20-32 kdal, and visible under the electron microscope as a cylinder made up of

four identical rings stacked on top of the other (Castaño *et al.*, 1986). It has no RNA component, in contrast with the prokaryotic RNAse P which requires an RNA fragment, termed M1 RNA, for activity. The *S. pombe* RNAse P has also been purified, and it was found to be associated not just with one, but with two, RNA species, designated K1 and K2 (Kline *et al.*, 1981; Krupp *et al.*, 1986). However, neither RNA possesses any sequence homology to their *E. coli* counterpart.

The *S. cerevisiae* tRNA splicing endonuclease and ligase have both been extensively studied in this laboratory. The endonuclease has the properties of an integral membrane protein and has eluded complete purification to date. It is particulate even at high ionic concentrations, but can be solubilized in the presence of nonionic detergents such as Triton X-100 (cited in Clark and Abelson, 1987). Preliminary calculations indicate there are 200 molecules of this enzyme per cell, and at the present stage of purification, activity copurifies with three bands of molecular weights 52.5, 40.7, and 29.5 kdal (P. Green, pers. comm.). In contrast, the tRNA splicing ligase has been purified to homogeneity, the protein sequenced at its amino terminus, and this sequence used to generate a collection of synthetic oligodeoxynucleotides as a mixed probe to screen a yeast genomic library for the cloning of the corresponding gene (Phizicky *et al.*, 1986). Complete sequence of the gene shows that it is a single, basic polypeptide of 827 amino acids of molecular weight 95.4 kdal, without any significant homology to any known protein of similar activity (Westaway *et al.*, 1987). There are roughly 400 molecules of ligase per cell (Phizicky *et al.*, 1986). The flanking sequence has also revealed a second open reading frame, termed ORF2, which is essential for cell viability (as shown by gene disruption experiments) and which may code for an unknown tRNA maturation activity. It is located unusually close (125 bp) to the ligase gene and may share the same promoter signals with it; and although

transcribed divergently from that of ligase, the ORF2 transcript is of the same abundance as the ligase transcript (Westaway *et al.*, 1987). It is generally believed that the endonuclease and ligase recognize common structural features in the pre-tRNA, but there is evidence that each has its own unique recognition elements, as indicated by recent findings that there are certain mutations affecting only excision but not ligation, and some which affect excision and ligation differentially (Greer *et al.*, 1987).

Wheat germ has a ligase activity that can join tRNA halves using a similar mechanism as that of the yeast tRNA ligase, although 3' terminal phosphates seem not to be tightly required (Gegenheimer *et al.*, 1983; Schwartz *et al.*, 1983). *E. coli* also has a ligase activity, but the product has a 2',5'-phosphodiester bond at the ligated junction, derived from the 2',3'-cyclic phosphate at the 3' terminus (Greer *et al.*, 1983a). Since *E. coli* genes do not possess any introns, the precise biological role of this RNA ligase activity remains obscure.

In *X. laevis*, the tRNA splicing endonuclease has been extensively purified and characterized. Results involving a yeast $tRNA_3^{Leu}$ gene containing mutations that disrupt the anticodon-intron base-pairing show that binding and cleavage of the pre-tRNA by the endonuclease are two distinct and separate steps (Baldi *et al.*, 1986). Recent findings from this laboratory indicate that this is also true for the *S. cerevisiae* tRNA splicing endonuclease (our unpublished results).

Development of synthetic systems that involve *in vitro* production of pre-tRNA from heterologous promoters that allow mutation of any portion of the pre-tRNA has also greatly facilitated the study of tRNA processing (Reyes and Abelson, 1987b; Uhlenbeck, 1986; Lowary *et al.*, 1986). For example, by construction of predesigned pre-$tRNA^{Phe}$ variants, we have been able to identify two probable endonuclease contact points in the pre-$tRNA^{Phe}$ (U8 and C56), as well formulate a splice site selection model which involve distance measurement as opposed to sequence recognition at the splice sites (Reyes and Abelson, 1987a).

Although the endonuclease and ligase are two distinct enzymes that can act separately and independently *in vitro*, there are biochemical and kinetic evidences which suggest that they act concertedly *in vivo*, and that a tRNA splicing complex or spliceosome does form in the cell (Greer, 1986). Also, binding studies indicate that ligase, even in the absence of endonuclease, binds end-mature, intron-containing pre-tRNA more tightly than any other type of pre-tRNA species (Greer, C. L., unpublished results). This observation, and the fact that ligase is nucleoplasmic while the endonuclease seems nuclear membrane-bound and possibly in close association with a nuclear pore, has prompted the formulation of a tRNA shuttle model, wherein ligase initially interacts with end-mature, intron-containing pre-tRNA to form a pre-splicing complex, which in turn interacts with the membrane-bound endonuclease, where splicing takes place concertedly and is perhaps coupled to tRNA translocation into the cytoplasm (Greer, 1986).

A hallmark of a fully-mature, functional tRNA is the presence of modified base and ribose moieties at certain, mostly constant, positions in the molecule (Björk, 1984). Base modifications are present in both prokaryotic and eukaryotic tRNAs, although the latter tend to have more of them. There are base modifications which are found exclusively in either class (e.g., yW and i^6A specific for eukaryotes; mnm^5s^2U specific for eubacteria; and m_2^2Gm specific for archaebacteria), and there are those which are found in both (e.g., m^5U and Ψ) (Björk, 1986). Some positions in the tRNA are always modified (e.g., position 54, usually with m^5U, and position 53, usually with Ψ), and certain portions of the molecule tend to have more base modifications (e.g., anticodon region and the D- and TΨc-loops) than others (e.g., acceptor stem). Some base modifications take place in the nucleus (e.g., D, Ψ, m^5C) while others take place in the cytoplasm (e.g., a poorly-characterized modification of G14 in $tRNA^{Tyr}$) (Melton *et al.*, 1980). The most complex nature and variety of base modifications are usually

found at the wobble (position 34, the first base in the anticodon) and hypermodified (position 37, that immediately following the anticodon) positions. In the biosynthesis of these complex base modifications, usually more than one base-modifying enzyme is involved, in contrast to the simple ones wherein only one is involved. For example, the Q (queuosine) base at position 37 of mammalian $tRNA^{Lys}$, the most complex base modification known, contains a 7-deazaguanine ring and a cyclopentendiol side chain as its distinctive structural feature; it is synthesized by a preliminary abstraction of the guanine base at this position by a transglycosylase to leave an apurinic site, subsequent insertion of queuine (q), and then further addition of side-chains, probably by another enzymatic activity to give queuosine (Nishimura, 1980). For at least one base modification enzyme, the wobble 2'-O-methylase, it has been shown that recognition elements lie in the mature domain of the pre-tRNA, as in the case of the tRNA processing enzymes (Droogmans *et al.*, 1986). Several mutations are known in a variety of organisms that lead to undermodification of tRNA (e.g., *trm*A for ribothymidine deficiency, and *his*T, pseudouridine), (Kohli, 1983) and certain neoplastic disorders are known to be associated with tRNA undermodification as well (Nishimura, 1980). Although the structural role of tRNA base modifications remains obscure (Kim, 1979) and certain strains defective in base modifications are viable, it is now known that certain base modifications, especially in the anticodon region, are necessary for optimal biological function of tRNA *in vivo*, specially codon recognition and ribosomal interaction (Bjork, 1986; Kohli, 1983) during the translational event.

The spatial and temporal order of events involved in tRNA maturation has been extensively studied in *X. laevis* oocytes through subcellular microinjection of either pre-tRNA or cloned tRNA genes, and it is generally believed that a similar scenario is true for most other eukaryotes as well (Melton *et al.*, 1980; Nishikura

and DeRobertis, 1981). All processing activities involving size alterations of the transcript—end-maturation, 3' CCA addition, IVS removal and exon ligation—are always detected in the nucleoplasm and never in the nuclear envelope or cytoplasm (DeRobertis *et al.*, 1981). This is slightly at odds with results involving yeast, wherein the endonuclease is believed to be an integral nuclear membrane protein; however, this might simply reflect the evolutionary distance between *X. laevis* and *S. cerevisiae*. As a rule, the IVS and the 5' and 3' extensions in the primary transcript are never base-modified. However, base modifications occur as early as the primary transcript, which is usually found to be completely modified at the TΨC loop (e.g., Ψ, m^1A and m^5U). The end-mature, IVS-containing pre-tRNA usually contains most of the base modifications in the D-loop (e.g., D, m^2G) and extra arms. This precursor is then spliced, and the remaining base modifications, mostly in the anticodon region (e.g., Q, i^6A), are added.

The tRNA splicing ligase has been localized, by electron microscopic and immune fluorescence techniques, in the inner periphery of the nucleus, and appears to be distributed in a distinct nucleoplasmic staining pattern, indicating close association of this enzyme with some nucleoplasmic substructure (Clark and Abelson, 1987).

Transport of tRNA from the nucleus to the cytoplasm, studied in detail in *X. laevis* oocyte systems, is one of the last steps in the biosynthetic pathway of this molecule. In this system, tRNA transport has a K_M of 10^{-7} M and a V_{max} of $2x10^9$ molecules/min/nucleus (Tobian *et al.*, 1985). Moreover, several point mutations, notably in the TψC- and D-stem/loop regions, affect transport of tRNA; for example, residue G_{57} is critical, suggesting that the transport machinery recognizes primary sequence in the tRNA (Zasloff *et al.*, 1982a). Further, all mutants impaired in transport are likewise impaired in processing, and no mutants have so far been detected that were transport-defective but were

normally processed. Hence, it seems likely that the transport machinery recognizes overall tertiary conformation of the tRNA as well. In *X. laevis*, tRNA transport is believed to be carrier-mediated as opposed to free diffusion through the nuclear pores, since the process is saturable by tRNA (Zasloff, 1983). Further, there is mutual competitive inhibition by two different tRNAs, suggesting that all tRNA molecules in the cell utilize the same transport apparatus.

Crucial to the study of tRNA processing is knowledge of the higher-order structure of the pre-tRNA, which contains the essential recognition elements of the various processing enzymes (see Guthrie, 1980). For this problem, we turn to the well-known three-dimensional structure of yeast $tRNA^{Phe}$. There is plenty of evidence, including results of chemical (Swerdlow and Guthrie, 1984) and enzymatic (Lee and Knapp, 1985) structure-probing analyses as well as direct NMR studies on pre-$tRNA^{Phe}$, which suggest that this molecule closely resembles $tRNA^{Phe}$ in structure, the only discrepancy lying on the anticodon stem/loop region, where the intron is located (C. Hall, unpublished results). Thus all the known structural features of $tRNA^{Phe}$ may be assumed to exist in pre-$tRNA^{Phe}$, except those in that region.

Crystallization of $tRNA^{Phe}$ was an absolute prerequisite for a rigorous physical study of its three-dimensional structure. This was initially achieved by inclusion of spermine, giving rise to orthorhombic and monoclinic crystal forms of the molecule (Rich and RajBhandary, 1976; Sussman and Kim, 1976; Kim ***et al.***, 1974). This crucial development allowed rigorous X-ray crystallographic analyses of the tRNA, and a vast wealth of structural information was quickly unearthed. The two-dimensional secondary cloverleaf structure proposed by Holley ***et al.*** (1965) was found to be correct. In three dimensions, however, the molecule has an overall L-shaped configuration, with the acceptor and TΨC stems forming a continuous helix and constituting one leg of the L, and the D- and anticodon stems

similarly making up the other leg (Quigley and Rich, 1976; Cedergren *et al.*, 1981). This L-shaped molecule is 75Å from tip to tip. The tertiary configuration of the molecule is stabilized by extensive base stacking (e.g., all but 5 - D_{16}, D_{17}, G_{20}, V_{47}, and A_{70} - of the 76 bases in the molecule and stacked) and by a multitude of tertiary interactions consisting of 10 base-base, 13 base-backbone, and 8 backbone-backbone interactions (Quigley and Rich, 1976; Klug *et al.*, 1974; Sundaralingam, 1977). Most of the double helical stems are of the 11-fold type RNA-A helix, with a very deep major groove and a shallow minor groove. Sharp turns in the loop regions, designated U- or uridine-turns, are stabilized by pi interactions involving uridine rings. Arch conformations—the bulging out of the polynucleotide chain which has the effect of excluding the arched-out residues from stacking—are stabilized by interactions involving a neighboring ribosyl 2' OH. It is clear from the general picture that there is extensive involvement of the ribosyl 2'OH, a feature that distinguishes RNA from DNA, and which may be the reason why nature has bestowed structural roles upon RNA, but not upon DNA. It should be commented, however, that certain rare species of tRNA do not obey this general structural motif; for example, a human mitochondrial $tRNA^{Ser}$ completely lacks the D-arm (De Bruijn *et al.*, 1980).

More recently, yeast $tRNA^{Asp}$ has been crystallized and subjected to the same X-ray crystallographic analyses done on yeast $tRNA^{Phe}$. Although a number of differences were found, a similar L-shaped tRNA molecule was immediately evident (Romby *et al.*, 1987). A striking difference, however, stems from the fact that $tRNA^{Asp}$ has an anticodon sequence of GUC, and hence can dimerize by anticodon-anticodon pairing between two identical molecules. This association is believed to mimic codon-anticodon interactions in the ribosome during translation, and by telestability triggers a long-distance conformational change in the molecule manifested as a departure in structure at the TψC loop (particularly C_{56}) from that of $tRNA^{Phe}$.

With the availability of several cloned IVS-containing tRNA genes and powerful site-directed mutagenesis techniques, the problem of tRNA processing has been dealt with using the reverse genetics approach: precise alteration of the tRNA precursor, and analysis of the splicing products of these altered precursors. The *S. cerevisiae* $tRNA_3^{Leu}$ gene, for example, contains a 32-bp IVS with an *Hpa*I site in the middle; insertion of a 21-bp synthetic DNA fragment corresponding to the *E. coli lac* operon at this site did not affect transcription and processing, suggesting that intact IVS is not required in these reactions (Johnson *et al.*, 1980). In another study, a transversion involving the highly conserved purine residue at the border of the 5' splice site was found to have no effect on the accuracy of IVS excision, supporting previous observations that splicing of tRNA does not depend on splice junction primary sequence *per se*, but on the overall tertiary structure of the tRNA domain of the precursor (Colby *et al.*, 1981). Indeed, analysis of a collection of single point mutations of $tRNA^{Tyr}$ revealed that a great majority of those which affect splicing map to the mature portion of the pre-tRNA (Nishikura *et al.*, 1982). Residue G_{57}, on the other hand, has been implicated in 3'-end maturation of the precursor, and transport of the mature tRNA across the nuclear membrane (Zasloff *et al.*, 1982a; Zasloff *et al.*, 1982b). Integrity of the double helical structure of the D-stem seems required for 5'-end maturation and IVS excision (Baldi *et al.*, 1983; Gandini-Attardi *et al.*, 1985; Mattoccia *et al.*, 1983). In contrast, end-maturation is slowed down if the 5' and 3' extensions are base-paired to form an extension of the acceptor stem, although the sequence of these terminal extensions *per se* do not affect either the accuracy or the efficiency of end-maturation (Castagnoli *et al.*, 1982). These *in vitro* results probably reflect the real situation *in vivo*, since a high correlation between *in vitro* and *in vivo* results have been consistently found (Strobel and Abelson, 1986).

Perhaps the best way to analyze tRNA function *in vivo* is through the use of tRNA suppressors, the most common type of which contain base changes in the anticodon that enable them to insert amino acids in response to the termination codons UAG (amber), UGA (opal), and UAA (ochre). Structure and function of tRNA suppressors are best characterized in the prokaryotes *E. coli* and *S. thypimurium*, and an almost complete collection of cloned tRNA suppressor genes are now available in these organisms (Sherman, 1982). In *S. cerevisiae*, however, tRNA nonsense suppressors inserting only tyrosine, serine and leucine are known; it is not clear why glutamine, lysine, glutamic acid, and tryptophan-inserting nonsense suppressors have not been found, when these tRNAs possess anticodons which can be easily altered to suppress nonsense codons (*Ibid.*). Like *S. cerevisiae*, *S. pombe* has a well-characterized tRNA suppressor system, and the first eukaryotic opal (UGA) suppressor tRNA was first isolated and characterized in this organism (Kohli *et al.*, 1980). A wide variety of cloned nonsense mutants of the iso-1-cytochrome C gene, termed *cycl* mutants, are available and provide an unequivocal method in *S. cerevisiae* for determining which nonsense codon is recognized and which amino acid is inserted by a tRNA suppressor, and is a widely-used system in the study of tRNA suppression (Sherman *et al.*,, 1979). Through total *de novo* synthesis, *S. cerevisiae* tRNA suppressor genes have been synthesized, and efforts to construct a similar tRNA suppressor gene library as in *E. coli* are underway (Masson *et al.*, 1987).

The widespread use of *S. cerevisiae*—which will be referred to as "yeast" from here on—as the experimental organism in studies on eukaryotic gene expression stems not only from the fact that it is unicellular and has a small and simple genome (only about three times that of *E. coli*), but also because it has a well-characterized genetics and is amenable to genetic manipulation, allowing identification and characterization of many mutations (Hopper, 1984; Warner,

1987). Its haploid genome size of 10^{10} daltons or 1.7×10^{4} kilo bp, is extremely small compared to those of other eukaryotes (Petes, 1980). Protocols are available for transformation of yeast by cloned DNA, and by the gene transplacement technology, a resident gene in the yeast chromosome may be precisely replaced by a cloned, site-specific mutant copy, thus allowing systematic analysis of the expression of such genes *in vivo* (Hopper *et al.*, 1981). Yeast is stable as haploids, allowing easy isolation of mutants, and as diploids, allowing quick determination of gene dominance and their classification into complementation groups (Petes, 1980). Yeast can be induced to sporulate and as such are easily micromanipulated; this feature has been exploited in tetrad analysis, wherein the segregation pattern of a certain gene or phenotype is used to quickly determine whether the phenotype is controlled by one or more genes, whether the gene is chromosomal or episomal, or whether two genes are linked (Mortimer and Schild, 1982). Further, several types of cloning vectors are available which afford a wide variety of genetic manipulations. YIp (yeast integrating plasmids) vectors are stably maintained in the cell by homologous integration into the chromosome, and is the most useful for genetic analyses; YEp (E = episomal) vectors contain a control segment of the yeast 2μ plasmid, transform at very high frequency, and are maintained at high copy numbers per cell; YRp (R = replicating) vectors contain a yeast ARS sequence, and can either be maintained episomally or integrated chromosomally; and YCp (C = centromeric) vectors contain yeast chromosomal centromeric sequences, and thus replicate and segregate as small, independent chromosomes of stable copy number of unity (Botstein and Davis, 1982). There are also several well-characterized and powerful methods of mapping genes in the yeast genome which involve meiotic (tetrad, random spore, or trisomic analyses) or mitotic (crossing-over or chromosome loss) techniques (Mortimer and Schild, 1982). A large set of tester

yeast strains are available at the Yeast Genetics Stock Center in Berkeley, California, that can be used in genetic crosses involving these gene mapping procedures (*Ibid.*). The fact that several yeast genes (excluding those with introns) can function efficiently in *E. coli* have been exploited in the cloning of many yeast genes by complementation of bacterial mutations, a vast collection of strains of which are available (Olson, 1982). Through yeast genetics and recombinant DNA techniques, therefore, it is now formally possible to isolate and clone the DNA corresponding to any yeast gene for which there exists a scorable phenotype for mutations in that gene (Roman, 1982).

Processing of tRNA may at present be the subject of much intense research, but the central and most fundamental biological significance of a tRNA molecule is its adaptor function during protein synthesis (for a review, see Schimmel and Söll, 1979). In this process, the tRNA molecule makes several homotropic (e.g., between nucleic acids) and heterotropic (e.g., between proteins and nucleic acids) interactions with a wide variety of macromolecules in the cell (Kisselev, 1985). Since much more is known about tRNA structure and function, both at the DNA and RNA levels, than any other component of the translational machinery, much of the information about these interactions concerns which portion of the tRNA interacts with which components, rather than *vice versa*. For instance, all elongator tRNAs contain the sequence TΨCG at the T-loop, and it has been suggested that this sequence interacts with the complementary sequence CGAAC in 5S RNA during translation in the ribosome (Erdman *et al.*, 1973). The invariant U_8 residue has been implicated in the formation of a transient covalent Michael adduct with the corresponding synthetase during aminoacylation (Koontz and Schimmel, 1979; Starzyk *et al.*, 1982), and there is strong evidence that the anticodon is a specific recognition element for the synthetase during the aminoacylation reaction (Kisselev, 1985). For instance, binding of complementary

oligonucleotides to the anticodon inhibits charging of the tRNA; also, mutation of the anticodon alters the specificity and efficiency of charging. Besides its major role in protein synthesis, tRNAs have been shown to participate in several nonconventional cellular functions; for example, it has been shown to be a component of the ubiquitin-dependent proteolysis system in the cell (Ferber and Ciechanover, 1986; Ciechanover *et al.*, 1985). More recently, it was found that a chloroplast $tRNA^{Glu}$ is a precursor to chlorophyll biosynthesis (Schön *et al.*, 1987).

Study of tRNA processing is only peripheral to the more profound problem of intron origin and function. Introns (intragenic regions) were first discovered a decade ago during routine electron microscopic examination of RNA-DNA hybrids wherein looped out regions in the DNA suggested noncolinearity between gene and message (Crick, 1979). Since that time, it has been well-documented that introns is a widespread feature among eukaryotic genes (with some examples in archaebacterial systems) coding for tRNA, rRNA and mRNA, and organellar, as well as chromosomal genes (*Ibid.*; Greer and Abelson, 1984). Introns break the continuity of genes into fragments called exons (expressed regions). Intron function in yeast tRNA genes have been directly addressed by site-directed deletion of the intron and introduction of the resulting intronless gene into the yeast cell by transformation. Thus far, no biological function has been ascribed to tRNA introns, except as possible recognition elements for a couple of anticodon base modification enzymes (Wallace *et al.*, 1980; Johnson and Abelson, 1983; Strobel and Abelson, 1982). Initially, four different possible routes were considered whereby these split genes are expressed (Gilbert, 1978): i) DNA could rearrange in somatic cells to get rid of introns in their germ line counterparts and then be transcribed as such; ii) RNA polymerase skips introns and transcribes only the exons; iii) each exon could be transcribed separately and the resulting transcripts subsequently ligated; and (iv) the split gene is transcribed in its

entirety including the introns, and then these introns are excised and the resulting exons ligated, all at the level of RNA. It has been shown that some immunoglobulin genes (e.g., mouse κ or λ light chain genes) indeed rearrange as in (i), but it is now known that the great majority of eukaryotic split genes are expressed as in (iv).

Sequence comparison of a large number of intron-containing, protein-coding genes reveal consensus sequences at the 5'- and 3'-splice sites: $AG^{\downarrow}GURAG$ and $Y\text{-}YYY\text{-}CAG^{\downarrow}$, respectively (Sharp, 1981). These signals seem to act independently, since a chimeric gene containing the 5' splice site of SV40 early mRNA and the 3' splice site from the mouse β-globin gene is spliced correctly (cited in Sharp, 1981).

A basic question regarding intron origin is whether they are primordial features that are slowly being lost with time, or recent insertions being slowly added into previously unsplit genes (Lewin, 1982; Cornish-Bowden, 1985). Previous findings that introns are usually located between functional and/or structural domains in proteins favor the first hypothesis, since this would indicate that introns function as recombinational hot spots during exon-shuffling for the production of novel genes. However, exceptions are known wherein introns are found in the middle of protein functional domains, as in the actin and myosin introns (cited in Lewin, 1982). However, by an intricate analysis of α-carbon distance maps of these types of proteins, it was found that such domains are further subdivided into subregions termed modules, and that introns are always located between modules. This observation obviously strengthens the first hypothesis.

REFERENCES

Abelson, J. (1979). RNA processing and the intervening sequence problem. **Ann. Revs. Biochem.** *48*, 1035-1069.

Abelson, J. (1980). The organization of tRNA genes. In: *Transfer RNA: Biological Aspects*, Söll, D., Abelson J. and Schimmel, P. R., eds., pp. 211-220, Cold Spring Harbor Laboratory, New York.

Abelson, J. N., Brody, E. N., Cheng, S.-C. *et al.* (1986). RNA splicing in yeast. *Chemica Scripta 26B* 127-137.

Abelson, J., Knapp, G., Peebles, C. L., Ogden, R. C., Johnson, P. F. and Johnson, J. D. (1981). tRNA splicing in yeast. In: *Molecular and Cellular Aspects of Microbial Evolution.* Carlilie, M. J., Collins J. F. and Moseley, B. E. B., eds., pp. 151-173, Soc. for Gen. Microbiol. Ltd., Symp #32, Cambridge University Press.

Allison, D. S., Goh, S. H. and Hall, B. D. (1983). The promoter sequence of a yeast $tRNA^{Tyr}$ gene. **Cell** *34*, 655-664.

Allison, D. S. and Hall, B. D. (1985). Effects of alterations in the 3' flanking sequence on *in vivo* and *in vitro* expression of the yeast SUP4-o $tRNA^{Tyr}$ gene. **EMBO J.** *4*, 2657-2664.

Baker, R. E., Gabrielsen, O., and Hall, B. D. (1986). Effects of $tRNA^{Tyr}$ point mutations on the binding of yeast RNA polymerase III transcription factor C. **J. Biol. Chem.** *261*, 5275-5282.

Baker, R. E. and Hall, B. D. (1984). Structural features of yeast tRNA genes which affect transcription factor binding. **EMBO J.** *3*, 2793-2800.

Baldi, M. I., Mattoccia, E., Ciafre, S., Gandini-Attardi, D. and Tocchini-Valentini, G. P. (1986). Binding and cleavage of pre-tRNA by the *Xenopus* splicing endonuclease: Two separable steps of the intron excision reaction. **Cell** *47*, 965-971.

Baldi, M. I., Mattoccia, E. and Tocchini-Valentini, G. P. (1983). Role of RNA structure in splicing: Excision of the intervening sequence in yeast $tRNA_3^{leu}$ is dependent on the formation of a D-stem. **Cell** *35*, 109-115.

Beckmann, J. S., Johnson, P. F. and Abelson, J. (1977a). Cloning of yeast transfer RNA genes in *Escherichia coli*. **Science** *196*, 205-208.

Beckmann, J. S., Johnson, P. F., Abelson, J. and Fuhrman, S. A. (1977b). Isolation and characterization of *Escherichia coli* clones containing genes for the stable yeast RNA species. In: *Eukaryotic Genetic Systems: ICN-UCLA Symposia on Molecular and Cellular Biology*, Vol. 8, pp. 213-226, Academic Press, Inc., New York.

Beckmann, J. S., Johnson, P. F., Knapp, G., Sakano, H., Fuhrman, S. A., Ogden, R. C. and Abelson, J. (1979). Organization and expression of yeast tRNA genes. In: *Nonsense Mutations and tRNA Suppressors*, Celis, J. E. and Smith, J. D., eds., pp. 207-234, Academic Press, Inc., New York.

Birkenmeier, E. H., Brown, D. D. and Jordan, E. (1978). A nuclear extract of *Xenopus laevis* oocytes that accurately transcribes 5S RNA genes. **Cell** *15*, 1077-1086.

Björk, G. R. (1984). Modified nucleosides in RNA: Their formation and function. In: *Processing of RNA*, Apirion, D., ed., pp. 291-330, CRC Press, Inc., Florida.

Björk, G. R. (1986). Transfer RNA modification in different organisms. **Chemica Scripta** *26B*, 91-95.

Botstein, D. and Davis, R. W. (1982). Principles and practice of recombinant DNA research with yeast. In: *The Molecular Biology of the Yeast Saccharomyces cerevisiae: Metabolism and Gene Expression*, Strathern, J. N., Jones, E. W. and Broach, J. R., eds., pp. 607-636, Cold Spring Harbor Laboratory, New York.

Busch, H., Reddy, R., Rothblum, L. and Choi, Y. C. (1982). SnRNAs, SnRNPs, and RNA processing. **Ann Revs. Biochem.** *51*, 617-654.

Carrara, G., DiSegni, G., Otsuka, A., Tocchini-Valentini, G. P. (1981). Deletion of the 3' half of the yeast $tRNA_3^{Leu}$ gene does not abolish promoter function *in vitro*. **Cell** *27*, 371-379.

Castagnoli, L., Ciliberto, G. and Cortese, R. (1982). Processing of eukaryotic tRNA precursors: Secondary structure of the precursor-specific sequences affects the rate, but not the accuracy of processing reactions. **Nucl. Acids Res.** *10*, 4135-4145.

Castaño, J. G., Ornberg, R., Koster, J. G., Tobian, J. A. and Zasloff, M. (1986). Eukaryotic pre-tRNAse 5' processing nuclease: Copurification with a complex cylindrical particle. **Cell** *46*, 377-387.

Castaño, J. G., Tobian, J. A. and Zasloff, M. (1985). Purification and characterization of an endonuclease from *Xenopus laevis* ovaries which accurately processes the 3' terminus of human pre-$tRNA_i^{Met}$ (3' pre-tRNAse). **J. Biol. Chem.** *260*, 9002-9008.

Cech, T. R. and Bass, B. L. (1986). Biological catalysis by RNA. **Ann. Revs. Biochem.** *55*, 599-629.

Cedergren, R. J., Sankoff, D., LaRue, B. and Grosjean, H. (1981). The evolving tRNA molecule. **CRC Crit. Rev. Biochem.** *11*, 35-104.

Ciechanover, A., Wolin, S., Steitz, J. A. and Lodish, H. F. (1985). Transfer RNA is an essential component of the ubiquitin and ATP-dependent proteolytic system. **Proc. Natl. Acad. Sci. USA** *82*, 1341-1345.

Ciliberto, G., Castagnoli, L., and Cortese, R. (1983a). Transcription by RNA polymerase III. **Curr. Top. Dev. Biol.** *18*, 59-88.

Ciliberto, G., Castagnoli, L., Melton, D. A. and Cortese, R. (1982a). Promoter of a eukaryotic $tRNA^{Pro}$ gene is composed of three noncontiguous regions. **Proc. Natl. Acad. Sci. USA** *79*, 1195-1199.

Ciliberto, G., Raugei, G., Constanzo, F., Dente, L. and Cortese, R. (1983b). Common and interchangeable elements in the promoters of genes transcribed by RNA polymerase III. **Cell** *32*, 725-733.

Ciliberto, G., Traboni, C. and Cortese, R. (1982b). Relationship between the two components of the split promoter by eukaryotic tRNA genes. **Proc. Natl. Acad. Sci. USA** *79*, 1921-1925.

Ciliberto, G., Traboni, C. and Cortese, R. (1984). Transcription of multimeric tRNA genes. **Nucl. Acids Res.** *12*, 1277-1285.

Clark, M. W. and Abelson, J. (1987). Yeast tRNA ligase has a primary location at the inner membrane of the nuclear envelope. **J. Cell. Biol.,** in press.

Clarkson, S. G. (1983). Transfer RNA genes. In: *Eukaryotic Genes: Their Structure, Activity and Regulation*, Maclean, N., Gregory, S. P. and Flavell, R. A., eds., pp. 239-261, Butterworths Pub. Co., Ltd., Boston.

Colby D., Leboy, P.S., and Guthrie, C. (1981). Yeast tRNA precursor mutated at a splice junction is correctly processed in vivo. **Proc. Natl. Acad. Sci. USA,** *78*, 415-419.

Cornish-Bowden, A. (1985). Are introns structural elements or evolutionary debris? **Nature** (London) *313*, 434-435.

Crick, F. (1979). Split genes and RNA splicing. **Science** *204*, 264-271.

De Bruijn, M. H. L., Schreier, P. H., Eperon, I. C., and Barrell, B. G. (1980). A mammalian serine transfer RNA lacking the "dihydrouridine" loop and stem. **Nucl. Acids. Res.** *8*, 5213-5222.

DeFranco, D., Schmidt, O. and Söll, D. (1980). Two control regions for eukaryotic tRNA gene transcription. **Proc. Natl. Acad. Sci. USA** *77*, 3365-3368.

DeFranco, D., Sharp, S. and Söll, D. (1981). Identification of regulatory sequences contained in the 5' flanking region of *Drosophila* lysine $tRNA_2$ genes. **J. Biol. Chem.** *256*, 12424-12429.

Del Rey, F. J., Donahue, T. F. and Fink, G. R. (1982). *Sigma*, a repetitive element found adjacent to tRNA genes of yeast. **Proc. Natl. Acad. Sci. USA** *79*, 4138-4142.

DeRobertis, E. M., Black, P. and Nishikura, K. (1981). Intranuclear location of the tRNA splicing enzymes. **Cell** *23*, 89-93.

Deutscher, M. P. (1984). Processing of tRNA in prokaryotes and eukaryotes. **CRC Crit. Rev. Biochem.** *17*, 45-71.

Dingermann, T., Sharp, S., Appel, B., DeFranco, D., Mount, S., Heiermann, R., Pongs, O. and Söll, D. (1981). Transcription of cloned tRNA and 5S RNA genes in a *Drosophila* cell extract. **Nucl. Acids Res.** *9*, 3907-3918.

Droogmans, L., Haumont, E., de Henau, S. and Grosjean, H. (1986). Enzymatic 2'-O-methylation of the wobble nucleoside of eukaryotic $tRNA^{Phe}$: Specificity depends on structural elements outside the anticodon loop. **EMBO J.** *5*, 1105-1109.

Engelke, D. R., Gegenheimer P. and Abelson, J. (1985). Nucleolytic processing of a $tRNA^{Arg}$-$tRNA^{Asp}$ dimeric precursor by a homologous component from *S. cerevisiae*. **J. Biol. Chem.** *260*, 1271-1279.

Erdman, V. A., Sprinzl, M. and Pongs, O. (1973). The involvement of 5S RNA in the binding of tRNA to ribosomes. ***Biochem. Biophys. Res. Commun.*** ***54,*** 942-948.

Feldman, H. (1977). A comparison of transcriptional linkage of tRNA cistrons in yeast and ***E. coli*** by the ultraviolet light technique. **Nucl. Acids Res.** *4*, 2831-2840.

Ferber, S. and Ciechanover, A. (1986). Transfer RNA is required for conjugation of ubiquitin to selective substrates of the ubiquitin-and ATP-dependent proteolytic system. **J. Biol. Chem.** *261*, 3128-3134.

Flint, S. J. (1984). Processing of mRNA precursors in eukaryotic cells. In: *Processing of RNA*, D. Apirion, ed., pp. 151-179, CRC Press, Inc., Boca Raton, Florida.

Folk, W. R. and Hofstetter, H. (1983). A detailed mutational analysis of the eukaryotic $tRNA_i^{Met}$ gene promoter. **Cell** *33*, 585-593.

Folk, W. R., Hofstetter, H. and Birnstiel, M. L. (1982). Some bacterial tRNA genes are transcribed by eukaryotic RNA polymerase III. **Nucl. Acids Res.** *10*, 7153-7162.

Frendewey, D., Dingermann, T., Cooley, L. and Söll, D. (1985). Processing of tRNA Precursors in *Drosophila*: Processing of the 3' end involves an endonucleolytic cleavage and occurs after 5' end maturation. **J. Biol. Chem.** *280*, 449-454.

Gabrielsen, O. S., and Øyen, T. B. (1987). The requirement for the A-block promoter element in tRNA gene transcription in vitro depends on the ionic environment. **Nucl. Acids Res.** *15*, 5699-5713.

Galli, G., Hofstetter, H., and Birnstiel, M. L. (1981). Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. **Nature** (London) *294*, 626-631.

Gamulin, V., Mao, J., Appel, B., Sumner-Smith, M., Yamao, F., and Söll, D. (1983). Six ***Schizosaccharomyces pombe*** tRNA genes including a gene for a $tRNA^{Lys}$ with an intervening sequence which cannot base-pair with the anticodon. **Nucl. Acids. Res.** *11*, 8537-8546.

Gandini-Attardi, D., Margarit, I. and Tocchini-Valentini, G. P. (1985). Structural alterations in mutant precursors of the yeast $tRNA_3^{leu}$ gene which behave as defective substrates for a highly purified splicing endonuclease. **EMBO J.** *4*, 3289-3297.

Gegenheimer, P., Gabius, H. J., Peebles, C. L. and Abelson, J. (1983). An RNA ligase from wheat germ which participates in transfer RNA splicing *in vitro*. **J. Biol. Chem.** *258*, 8365-8373.

Gilbert, W. (1978). Why genes in pieces? **Nature** (London) *271*, 501.

Gottesfeld, J. (1986). Novobiocin inhibits RNA polymerase III transcription *in vitro* by a mechanism distinct from DNA topoisomerase II. **Nucl. Acids Res.** *14*, 2075-2088.

Green, M. R. (1986). Pre-mRNA splicing. **Ann. Rev. Genet.** *20*, 671-708.

Greer, C. L. (1986). Assembly of a tRNA splicing complex: Evidence for concerted excision and joining steps in splicing *in vitro*. **Mol. Cell. Biol.** *6*, 635-644.

Greer, C. L., and Abelson, J. (1984). RNA splicing: Rearrangement of RNA sequences in the expression of split genes. **Trends Biochem. Sci.** *9*, 139-141.

Greer, C. L., Javor, B. and Abelson, J. (1983a). RNA ligase in bacteria: Formation of a 2,'5'-linkage by an *E. coli* extract. **Cell** *33*, 899-906.

Greer, C. L., Peebles, C. L., Gegenheimer, P., and Abelson, J. (1983b). Mechanism of action of a yeast RNA ligase in tRNA splicing. **Cell** *32*, 537-546.

Greer, C. L., Söll, D., and Willis, I. (1987). Substrate recognition and identification of splice sites by the tRNA-splicing endonuclease and ligase from *Saccharomyces cerevisiae*. **Mol. Cell. Biol.** *7*, 76-84.

Guthrie, C. (1980). Folding up a transfer RNA molecule is not simple. **Quart. Rev. Biol.** *55*, 335-352.

Guthrie, C. and Abelson, J. (1982). Organization and expression of tRNA genes in *Saccharomyces cerevisiae*. In: ***Molecular Biology of the Yeast Saccharomyces cerevisae: Metabolism and Gene Expression.*** Strathern, J. N. Jones, E. W. and Broach, J. R., eds., pp. 191-208, Cold Spring Harbor Laboratory, New York.

Hager, G., Holland, M., Valenzuela, P., Weinberg, F. and Rutter, W. J. (1976). RNA polymerases and transcriptive specificity in *Saccharomyces cerevisiae*. In: *RNA Polymerase*, Losick, R. and Chamberlin, M., eds., pp. 745-762, Cold Spring Harbor Laboratory, New York.

Hall, B. D., Clarkson, S. G. and Tocchini-Valentini, G. P. (1982). Transcription initiation of eukaryotic trasnfer RNA genes. **Cell** *29*, 3-5.

Hipskind, R. A. and Clarkson, S. G. (1983). 5' flanking sequences that inhibit *in vitro* transcription of a *Xenopus laevis* tRNA gene. **Cell** *34*, 881-890.

Hofstetter, H., Kressmann, A. and Birnstiel, M. (1981). A split promoter for a eukaryotic tRNA gene. **Cell** *24*, 573-585.

Holley, R. W., Apgar, J., Everett, G. A., Madison, J. T., Marquisee, J. T., Mervil, S. H., Penswick, J. R. and Zamir, A. (1965). Structure of a ribonucleic acid. ***Science* 147**, 1462-1465.

Honda, B. M., Devlin, R. H., Nelson, D. W. and Khosla, M. (1986). Transcription of class III genes in cell-free extracts from the nematode *Caenorhabditis elegans*. **Nucl. Acids Res.** *14*, 869-881.

Hopper, A. K. (1984). Genetic and biochemical studies of RNA processing in yeast. In: ***Processing of RNA***. D. Apirion, ed., pp. 91-117, CRC Press, Inc. Boca Raton, Florida.

Hopper, A. K., Banks, F. and Evangelidis, V. (1978). Yeast mutant which accumulates precursor tRNAs. **Cell** *14*, 211-219.

Hopper, A. K., Nolan, S. L., Kurjan, J. and Furukawa, A. H. (1981). Genetic and biochemical approaches to studying *in vivo* intermediates in tRNA biosynthesis. In: ***Molecular Genetics in Yeast***, Alfred Benzon Symp. 16, von Wettstein, D., Früs, J., Brandt, M. K. and Stenderup, A., eds., pp. 302-325, Copenhagen, Denmark.

Hurt, D. J., Wang, S. S., Lin, Y. H. and Hopper, A. K. (1987). Cloning and characterization of *LOS1*, a Saccharomyces cerevisae gene that affects tRNA splicing. **Mol. Cell Biol.** *7*, 1208-1216.

Johnson, J. D., Ogden, R., Johnson, P., Abelson, J., Dembeck, P. and Itakura, K. (1980). Transcription and processing of a yeast tRNA gene containing a modified intervening sequence. **Proc. Natl. Acad. Sci. USA** *77*, 2564-2568.

Johnson, J. D., and Raymond, G. J. (1984). Three regions of a yeast $tRNA_3^{Leu}$ gene promote RNA polymerase III transcription. **J. Biol. Chem.** *259*, 5990-5994.

Johnson, P. F. (1983) Transfer RNA gene expression in *S. cerevisiae*: Studies on gene organization and intervening sequence function. Ph.D. Thesis, University of California, San Diego, California.

Johnson, P. F. and Abelson, J. (1983). The yeast $tRNA^{Tyr}$ gene intron is essential for correct modification of its tRNA product. **Nature** (London) *302*, 681-687.

Kang, H. S., Ogden, R. C. and Abelson, J. (1980). Two yeast tRNA genes containing intervening sequences. In: ***Mobilization and Reassembly of Genetic Information***, Scott, W. A., Werner, R., Joseph, D. R. and Schultz, J., eds., pp. 317-334, Academic Press, Inc., New York.

Kang, H. S., Ogden, R. C., Knapp, G., Peebles, C. L. and Abelson, J. (1979). Structure of yeast tRNA precursors containing intervening sequences. In: ***Eukaryotic Gene Regulation***, Axel, R., Maniatis, T. and Fox, F., eds., pp. 69-84, Academic Press, Inc., New York.

Kim, S. H. (1979). Crystal structure of yeast $tRNA^{Phe}$ and general structural features of other tRNAs. In: ***Transfer RNA: Structure, Properties and Recognition***, Schimmel, P. R., Söll, D., and Abelson, J., eds., pp. 83-100, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Kim, S. H., Sussman, J. L., Suddath, F. L., Quigley, G. J., McPherson, A., Wang A. H. J., Seeman, N. C. and Rich, A. (1974). The general structure of transfer RNA molecules. **Proc. Natl. Acad. Sci. USA** 71, 4970-4974.

Kisselev, L. L. (1985). The role of the anticodon in recognition of tRNA by aminoacyl-tRNA synthetases. **Prog. Nucl. Acids Res. Mol. Biol.** *32*, 237-266.

Kjellin-Sträby, K. S., Engelke, D. R. and Abelson, J. (1984). Homologous *in vitro* transcription of linear DNA fragments containing the $tRNA^{Arg}$-$tRNA^{Asp}$ gene pair from *Saccharomyces cerevisiae*. **DNA** *3*, 167-171.

Klekamp, M. S. and Weil, P. A. (1982). Specific transcription of homologous class III genes in yeast-soluble cell-free extracts. **J. Biol. Chem.** *257*, 8432-8441.

Kline, L., Nishikawa, S. and Söll, D. (1981). Partial Purification of RNAse P from Schizosaccharomyces pombe. **J. Biol. Chem.** *256*, 5058-5063.

Klug, A., Ladner, J., and Roberts, J. D. (1974). The structural geometry of coordinated base changes in transfer RNA. **J. Mol. Biol.** *89*, 511-516.

Knapp, G. Beckmann, J. S., Johnson, P. F., Fuhrman, S. A. and Abelson, J. (1978). Transcription and processing of intervening sequences in yeast tRNA genes. **Cell** *14*, 221-236.

Knapp, G., Ogden, R. C., Peebles, C. L., and Abelson, J. (1979). Splicing of yeast tRNA precursors: Structure of the reaction intermediates. **Cell** *18*, 37-45.

Kolata, G., (1983). Z-DNA moves toward "real biology". **Science,** *222*, 495-496.

Kohli, J. (1983). The genetics of modified nucleosides in transfer RNA: Known mutations leading to modification defects and their influence on transfer RNA function. In: ***The Modified Nucleosides of Transfer RNA II***, pp. 1-10, Alan R. Liss, Inc., New York.

Kohli, J., Mao, J. and Söll, D. (1980). Transfer RNA and nonsense suppression in ***Schizosaccharomyces pombe.*** In: ***Genetics and Evolution of RNA***

Polymerase, tRNA, and Ribosomes, Osawa, S., Ozeki, H. Uchida, H. and Yura, T., eds., pp. 233-250, Tokyo University Press, Tokyo, Japan.

Koski, R. A., Allison, D. S., Worthington, M. and Hall, B. D. (1982). An *in vitro* RNA polymerase III system from *S. cerevisiae*: Effects of deletions and point mutations upon SUP4 gene transcription. **Nucl. Acids Res.** *10*, 8127-8143.

Koontz, S. W. and Schimmel, P. R. (1979). Aminoacryl synthetase-catalyzed cleavage of the glycosidic bond of 5-halogenated uridines. **J. Biol. Chem.** *254*, 12277-12280.

Krupp, G., Cherayil, B., Frendewey, D., Nishikawa, S. and Söll, D. (1986). Two RNA species copurify with RNAse P from the fission yeast *Schizosaccharomyces pombe*. **EMBO J.** *5*, 1697-1703.

Laski, F. A., Fire, A. Z., RajBhandary, U. L. and Sharp, P. A. (1983). Characterization of tRNA precursor splicing in mammalian extracts. **J. Biol. Chem.** *258* 11974-11980.

Lassar, A. B., Martin, P. L. and Roeder, R. G. (1983). Transcription of class III genes: Formation of preinitiation complexes. **Science** *222*, 740-748.

Lee, M. C. and Knapp, G. (1985). Transfer RNA splicing in *Saccharomyces cerevisiae*: Secondary and tertiary structures of the substrates. **J. Biol. Chem.** *260*, 3108-3115.

Lewin, R. (1982). On the origin of introns. **Science** *217*, 921-922.

Lowary, P., Sampson, J., Milligan, J., Groebe, D. and Uhlenbeck, O. (1986). A better way to make RNA for physical studies, (submitted for publication).

Mao, J., Schmidt, O. and Söll, D. (1980). Dimeric transfer RNA precursors in *S. pombe*. **Cell** *21*, 509-516.

Marzouki, N., Camier, S., Ruet, A., Moenne, A. and Sentenac, A. (1986). Selective proteolysis defines two DNA binding domains in yeast transcription factor τ. **Nature** (London) *323*, 176-178.

Masson, J. M., Meuris, P., Grunstein, M., Abelson, J. and Miller, J. (1987). Expression of a set of synthetic suppressor $tRNA^{Phe}$ genes in *Saccharomyces cerevisiae*. **Proc. Natl. Acad. Sci. USA,** in press.

Mattoccia, E., Baldi, M. I., Pande, G., Ogden, R. and Tocchini-Valentini, G. P. (1983). Mutation in the A-block of the yeast $tRNA_3^{Leu}$ gene that allows transcription, but abolishes splicing and 5'-end maturation. **Cell** *32*, 67-76.

Melton, D. A., DeRobertis, E. M., and Cortese, R. (1980). Order and intracellular location of the events involved in the maturation of a spliced tRNA. **Nature** *284*, 143-148.

Mortimer, R. K. and Schild, D. (1982). Genetic mapping in *Saccharomyces cerevisiae*. In: *Molecular Biolog of the Yeast Saccharomyces cerevisiae: Life Cycle and Inheritance*, Strathern, J. N., Jones, E. W. and Broach, J. R., eds., pp. 11-26, Cold Spring Harbor Laboratory, New York.

Nevins, J. R. (1984). Poly(A) in eukaryotic mRNA. In: *Processing of RNA*, D. Apirion, ed., pp. 133-150, CRC Press, Inc., Boca Raton, Florida.

Newman, A. J., Ogden, R. C. and Abelson, J. (1983). tRNA transcription in yeast: Effects of specified base substitutions in the intragenic promoter. **Cell** *35*, 117-125.

Nishikura, K. and DeRobertis, E. M. (1981). RNA processing in microinjected *Xenopus* oocytes: Sequential addition of base modifications in a spliced transfer RNA. **J. Mol. Biol.** *145*, 405-420.

Nishikura, K., Kurjan, J., Hall, B. D., DeRobertis, E. M. (1982). Genetic analysis of the processing of a spliced tRNA. **EMBO J.** *1*, 263-268.

Nishimura, S. (1980). Structure and biosynthesis of hypermodified nucleosides in tRNA. In: *Genetics and Evolution of RNA Polymerase, tRNA and Ribosomes*, Osawa, S., Ozeki, H., Uchida H., and Yura, T., eds., pp. 251-265, Tokyo University Press, Tokyo, Japan.

Ogden, R. C., Beckmann, J. S., Abelson, J., Kang, H. S., Söll, D. and Schmidt, O. (1979). *In vitro* transcription and processing of a yeast tRNA gene containing an intervening sequence. **Cell** *17*, 399-406.

Ogden, R. C., Knapp, G., Peebles, C. L., Johnson, J., and Abelson, J. (1981). The mechanism of tRNA splicing. **Trends Biochem. Sci.** *6*, 154-158.

Ogden, R. C., Knapp, G., Peebles, C. L., Kang, H. S., Beckmann, J. S., Johnson, P. F., Fuhrman, S. A., and Abelson, J. N. (1980). Enzymatic removal of intervening sequences in the synthesis of yeast tRNAs. In: *Transfer RNA: Biological Aspects*, Söll, D., Abelson, J., and Schimmel, P. R., eds. pp. 173-190, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Ogden, R., Lee, M. C., Knapp, G. (1984). Transfer RNA splicing in *S. cerevisiae*: Defining the substrates. **Nucl. Acids Res.** *12*, 9367-9382.

Olson, M. V. (1982). Applications of molecular cloning to *Saccharomyces*. In: **Genetic Engineering,** Williamson, R., ed., Vol. 3, pp. 57-88, Academic Press, Inc., New York.

Padgett, R. A., Grabowski, P. J., Konarska, M. M., Seiter, S. and Sharp, P. A. (1986). Splicing of messenger RNA precursors. **Ann. Revs. Biochem.** *55*, 1119-1150.

Pearson, D., Willis, I., Hottinger, H., Bell, J., Kumar, A., Leupold, U. and Söll, D. (1985). Mutations preventing expression of *sup*3 $tRNA^{Ser}$ nonsense suppressors of *S. pombe*. **Mol. Cell. Biol.** *5*, 808-815.

Peebles, C. L., Gegenheimer, D., and Abelson, J. (1983). Precise excision of intervening sequences from precursor tRNAs by a membrane-associated yeast endonuclease. **Cell** *32*, 525-536.

Peebles, C. L., Ogden, R. C., Knapp, G., and Abelson, J. (1979) Splicing of yeast tRNA precursors: A two-stage reaction. **Cell** *18*, 27-35.

Petes, T. D. (1980). Molecular genetics of yeast. **Ann. Revs. Biochem.** *49*, 845-876.

Phizicky, E. M., Schwartz, R. C. and Abelson, J. (1986) *S. cerevisiae* tRNA ligase: Purification of the protein and isolation of the structural gene. **J. Biol. Chem.** *261*, 2978-2986.

Pieler, T., Hamm, J. and Roeder, R. G. (1987). The 5S gene internal control region is composed of three distinct sequence elements, organized as two functional domains with variable spacing. **Cell** *48*, 91-100.

Quigley, G. J. and Rich, A. (1976). Structural domains of transfer RNA molecules. **Science** *194*, 796-806.

Raymond, G. J. and Johnson, J. D. (1983). The role of non-coding DNA sequences in transcription and processing of a yeast tRNA. **Nucl. Acids Res.** *11*, 5969-5988.

Raymond, K. C., Raymond, G. J. and Johnson, J. D. (1985). *In vivo* modulation of yeast tRNA gene expression by 5' flanking sequences. **EMBO J.** *4*, 2249-2256.

Reyes, V. M. and Abelson, J. (1987a). Substrate recognition and splice site determination in yeast tRNA splicing. **Cell**, submitted for publication.

Reyes, V. M., and Abelson, J. (1987b). A synthetic substrate for tRNA splicing. **Anal. Biochem.** *166*, 90-106.

Reyes, V. M., Newman, A., and Abelson, J. (1986). Mutational analysis of the coordinate expression of the yeast $tRNA^{Arg}$-$tRNA^{Asp}$ gene tandem. **Mol. Cell. Biol.** *6*, 2436-2442.

Rich, A., and RajBhandary, U. L. (1976). Transfer RNA: Molecular structure, sequence and properties. **Ann. Revs. Biochem.** *45*, 805-860.

Roman, H. (1982). Development of yeast as an experimental organism. In: ***Molecular Biology of the Yeast Sacharomyces cerevisiae: Life Cycle and***

Inheritance, pp. 1-9, Strathern, J. N., Jones, E. W. and Broach, J. R., eds., Cold Spring Harbor Laboratory, New York.

Romby, P., Moras, D., Dumas, P., Ebel, J. P. and Giege, R. (1987). Comparison of the tertiary structure of yeast $tRNA^{Asp}$ and $tRNA^{Phe}$ in solution: Chemical modification study of the bases. **J. Mol. Biol.** *195*, 193-204.

Rooney, R. J. and Harding, J. D. (1986). Processing of mammalian tRNA transcripts *in vitro*: Different pre-tRNAs are processed along alternative pathways that contain a common rate-limiting step. **Nucl. Acids Res.** *14*, 4849-4864.

Schaack, J., Sharp, S., Dingermann, T. and Söll, D. (1983). Transcription of eukaryotic tRNA genes *in vitro*. II. Formation of stable complexes. **J. Biol. Chem.** *258*, 2447-2453.

Schaack, J., Sharp, S., Dingermann, T., Burke, D. J., Cooley, L. and Söll, D. (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.

Schimmel, P. R. and Söll, D. (1979). Aminoacyl-tRNA synthetases: General features and recognition of transfer RNAs. **Ann. Revs. Biochem.** *48*, 601-648.

Schmidt, O., Mao, J., Ogden, R., Beckmann, J., Sakano, H., Abelson, J. and Söll, D. (1980). Dimeric tRNA precursors in yeast. **Nature** (London) *287*, 750-752.

Schön, A., Krupp, G., Gough, S., Berry-Lowe, S., Kannangara, C. G. and Söll, D. (1987). The RNA required in the first step of chlorophyll biosynthesis is a chloroplast glutamate tRNA. **Nature** (London) *322*, 281-284.

Schwartz, R. C., Greer, C. L., Gegenheimer, P. and Abelson, J. (1983). Enzymatic mechanism of an RNA ligase from wheat germ. **J. Biol. Chem.** *258*, 8374-8383.

Sentenac, A. (1985). Eukaryotic RNA polymerases. **CRC Crit. Rev. Biochem.** *18*, 31-90.

Sentenac, A. and Hall, B. D. (1982). Yeast nuclear RNA polymerases. In: ***Molecular Biology of the Yeast Saccharomyces cerevisiae: Metabolism and Gene Expression.*** Strathern, J. N. Jones, E. W. and Broach, J. R., eds., pp. 561-606, Cold Spring Harbor Laboratory, New York.

Sharp, P. A. (1981). Speculations on RNA splicing. **Cell** *23*, 643-646.

Sharp, P. A. (1985). On the origin of RNA splicing and introns. **Cell** *42*, 397-400.

Sharp, P. A., Hardy, S. and Padgett, R. A. (1984). Splicing of messenger RNA precursors. In: *Transfer and Expression of Eukaryotic Genes*, Ginsberg, H. S. and Vogel, H. J., eds., pp. 183-202, Academic Press, Inc., New York.

Sharp, S., Schaack, J., Cooley, L., Burke, D. J. and Söll, D. (1986). Structure and transcription of eukaryotic tRNA genes. **CRC Crit. Rev. Biochem.** *19*, 107-144.

Sharp, S., DeFranco, D., Dingermann, T., Farrell, D. and Söll, D. (1981). Internal control regions for transcription of eukaryotic tRNA genes. **Proc. Natl. Acad. Sci. USA** *78*, 6657-6661.

Sharp, S., Dingermann, T., Schaack, J., DeFranco, D. and Söll, D. (1983). Transcription of eukaryotic tRNA genes *in vitro*: I. Analysis of control regions using a competitive assay. **J. Biol. Chem.** *258*, 2440-2446.

Sharp, S., Dingermann, T., and Söll, D. (1982). The minimum intragenic sequences required for promotion of eukaryotic tRNA gene transcription. **Nucl. Acids. Res.** *10*, 5393-5406.

Shastry, B. S., Ng, S. Y. and Roeder, R. G. (1982). Multiple factors involved in the transcription of class III genes in *Xenopus laevis*. **J. Biol. Chem.** *257*, 12979-12986.

Shaw, K. J. and Olson, M. V. (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.

Sherman, F. (1982). Suppression in the yeast *Saccharomyces cerevisiae*. In: *The Molecular Biology of the Yeast Saccharomyces cerevisiae: Metabolism and Gene Expression*, Strathern, J. N., Jones, E. W. and Broach, J. R., eds., pp. 463-486, Cold Spring Harbor Laboratory, New York.

Sherman, F., Ono, B. and Stewart, J. W. (1979). Use of the iso-1-cytochrome C system for investigating nonsense mutants and suppressors in yeast. In: *Nonsense Mutations and tRNA Suppressors*, Celis, J. E. and Smith, J. D., eds., pp. 133-153, Academic Press.

Silvie, C., Gabrielsen, O., Baker, R. and Sentenac, A. (1985). A split binding site for transcription factor τ on the $tRNA_3^{Glu}$ gene. **EMBO J.** *4*, 491-500.

Sklar, V. E. F., Yamamoto, M., and Roeder, R. G. (1976). Molecular structures of Eukaryotic class III RNA polymerases. In: *RNA polymerase*, Losick, R. and Chamberlin, M., eds., pp. 803-817, Cold Spring Harbor Laboratory, New York.

Sprague, K. U., Larson, D. and Morton, D. (1980). 5' Flanking sequence signals are required for activity of silkworm alanine tRNA genes in homologous *in vitro* transcription systems. **Cell** *22*, 171-178.

Sprague, K. U., Ottonello, S., Rivier, D. H. and Young, L. S. (1987). Control of tRNA gene transcription. In: *RNA Polymerase and the Regulation of Transcription*, W. Reznikoff, S. *et al.*, eds., pp. 195-207, Elsevier Science Publishing Co., Inc., New York.

Standring, D. N., Venegas, A. and Rutter, W. J. (1981). Yeast $tRNA_3^{Leu}$ gene transcribed and spliced in a HeLa cell. **Proc. Natl. Acad. Sci. USA** *78*, 5963-5967.

Starzyk, R. M., Koontz, S. W. and Schimmel, P. R. (1982). A covalent adduct between the uracil ring and the active site of an aminoacyl tRNA synthetase. **Nature** (London) *298*, 136-140.

Stewart, T. S., Söll, D. and Sharp, S. (1985). Point mutations in the 5' ICR and anticodon region of a *Drosophila* $tRNA^{Arg}$ gene decrease *in vitro* transcription. **Nucl. Acids. Res.** *13*, 435-447.

Stillman, D. J., Better, M., Geiduschek, E. P. (1985a). Electron microscopic examination of the binding of a large RNA polymerase III transcription factor to a tRNA gene. **J. Mol. Biol.** *185*, 451-455.

Stillman, D. J., Caspers, P. and Geiduschek, E. P. (1985b). Effects of temperature and single-stranded DNA on the interaction of an RNA polymerase III transcription factor with a tRNA gene. **Cell** *40*, 311-317.

Stillman, D. J. and Geiduschek, E. P. (1984). Differential binding of a *S. cerevisiae* RNA polymerase III transcription factor to two promoter segments of a tRNA gene. **EMBO J.** *3*, 847-853.

Stillman, D. J., Sivertsen, A. L., Zentner, P. G. and Geiduschek, E. P. (1984). Correlations between transcription of a yeast tRNA gene and transcription factor-DNA interactions. **J. Biol. Chem.** *259*, 7955-7962.

Strobel, M. C., and Abelson, J. (1986). Effect of intron mutations on processing and function of *Saccharomyces cerevisiae* SUP53 tRNA *in vitro* and *in vivo*. **Mol. Cell. Biol.** *6*, 2663-2673.

Struhl, K. (1986). Yeast promoters. In: *Maximizing Gene Expression*, Reznikoff, W. and Gold, L., eds., pp. 35-78, Butterworths Pub. Co., Ltd., Boston.

Sullivan, M. A. and Folk, W. R. (1987). Transcription of a eukaryotic $tRNA_i^{Met}$ and 5S RNA genes by RNA polymerase III is blocked by base mismatches in the intragenic control regions. **Nucl. Acids Res.** *15*, 2059-2068.

Sundaralingam, M. (1977). Non-Watson-Crick base pairs in ribonucleic acids. **Intl. J. Quant. Chem: Quant. Biol. Symp.** *4*, 11-23.

Sussman, J. L. and Kim, S. H. (1976). Three-dimensional structure of a transfer RNA in two crystal forms. **Science** *192*, 853-858.

Swerdlow, H. and Guthrie, C. (1984). Structure of intron-containing tRNA precursors. **J. Biol. Chem.** *259*, 5197-5207.

Tobian, J. A., Drinkard, L. and Zasloff, M. (1985). tRNA nuclear transport: Defining the critical regions of human $tRNA_i^{Met}$ by point mutagenesis. **Cell** *43*, 415-422.

Traboni, C., Ciliberto, G. and Cortese, R. (1984). Mutations in the B-box of the promoter of a eukaryotic $tRNA^{Pro}$ gene affects rate of transcription, processing, and stability of transcripts. **Cell** *36*, 179-187.

Uhlenbeck, O. C. (1986). Structure and function of RNA. **Chemica Scripta** *26B*, 97-101.

Valenzuela, P., O'Farrell, P. Z., Cordell, B., Maynard, T., Goodman, H. M. and Rutter, W. J. (1980). Yeast tRNA precursors: Structure and removal of intervening sequences by an excision-ligation activity. In: *Transfer RNA: Biological Aspects*, Söll, D., Abelson, J. N. and Schimel, R. R. eds., pp. 191-208, Cold Spring Harbor Laboratory, New York.

Vincent, A. (1986). TFIIIA and homologous genes: The "finger" proteins. **Nucl. Acids Res.** *14*, 4385-4391.

Wallace, R. B., Johnson, P. F., Tanaka, S., Schöld, M., Itakura, K. and Abelson, J. (1980). Directed deletion of a yeast transfer RNA intervening sequence. **Science** *209*, 1396-1400.

Warner, J. (1987). Applying genetics to the splicing problem. **Genes Dev.**, *1*, 1-3.

Weil, P. A. and Klekamp, M. S. (1986). Partial purification and characterization of *Saccharomyces cerevisiae* transcription factor IIIB. **J. Biol. Chem.** *261*, 2819-2827.

Weil, P. A., Segall, J., Harris, B., Ng, S. Y. and Roeder, R. G. (1979). Faithful transcription of eukaryotic genes by RNA polymerase III in systems reconstituted with purified DNA templates. **J. Biol. Chem.** *254*, 6163-6173.

Westaway, S. K., Phizicky, E. M. and Abelson, J. (1987). Sequence of a yeast tRNA splicing gene. **Mol. Cell. Biol.,** in press.

Willis, I., Frendewey, D., Nichols, M., Hottinger-Werlen, A., Schaack, J. and Söll, D. (1986). A single base change in the intron of a serine tRNA affects the rate of RNAse P cleavage *in vitro* and suppressor activity *in vivo* in *Saccharamyces cerevisiae*. **J. Biol. Chem.** *261*, 5878-5885.

Willis, I., Hottinger, H., Pearson, D., Chisolm, V., Leupold, U. and Söll, D. (1984). Mutations affecting excision of the intron from a eukaryotic dimeric tRNA precursor. **EMBO J.** *3*, 1573-1580.

Young, L. S., Takahashi, N. and Sprague, K. U. (1986). Upstream sequences confer distinctive transcriptional properties on genes encoding silkgland-specific $tRNA^{Ala}$. **Proc. Natl. Acad. Sci. USA** *83*, 374-378.

Zasloff, M. (1983). tRNA transport from the nucleus in a eukaryotic cell: Carrier-mediated translocation process. **Proc. Natl. Acad. Sci. USA** *80*, 6436-6440.

Zasloff, M., Rosenberg, M. and Santos, T. (1982a). Impaired nuclear transport of a human variant $tRNA_i^{Met}$. **Nature** (London) *300*, 81-84.

Zasloff, M., Santos, T. and Hamer, D. (1982b). tRNA precursor transcribed from a mutant human gene inserted into a SV40 vector is processed incorrectly. **Nature** (London) *295*, 533-535.

Zasloff, M., Santos, T., Romeo, P. and Rosenberg, M. (1982c). Transcription and precursor processing of normal and mutant human $tRNA_i^{Met}$ genes in a homologous cell-free system. **J. Biol. Chem.** *257*, 7857-7863.

Chapter II

MUTATIONAL ANALYSIS OF THE COORDINATE EXPRESSION OF THE YEAST $tRNA^{Arg}$-$tRNA^{Asp}$ GENE TANDEM

[The text of this chapter appeared in: Reyes, V. M., Newman, A. J. and Abelson, J. (1986) **Mol. Cell. Biol.** *6*, 2436-2442.]

MOLECULAR AND CELLULAR BIOLOGY, July 1986, p. 2436–2442
0270-7306/86/072436-07$02.00/0

Vol. 6, No. 7

Mutational Analysis of the Coordinate Expression of the Yeast tRNA^{Arg}-tRNA^{Asp} Gene Tandem

VICENTE M. REYES, ANDREW NEWMAN,† AND JOHN ABELSON*

Divisions of Biology and Chemistry, California Institute of Technology, Pasadena, California 91125

Received 9 December 1985/Accepted 31 March 1986

tRNA genes occur in the yeast genome as highly dispersed and independent transcriptional units. The 5′-tRNA^{Arg}-tRNA^{Asp}-3′ gene tandem, separated by a 10-base-pair spacer sequence, thus represents a rare case of tight clustering. Previous in vitro studies did not reveal any primary transcript from the tRNA^{Asp} gene, but rather a dimeric precursor containing both gene sequences plus spacer, which undergoes a series of maturation steps. This seems anomalous since the tRNA^{Asp} gene contains the sequences necessary for its own transcription. We found that site-directed mutation of the highly conserved C at position 56 to a G in the tRNA^{Arg} gene suppresses all transcription and does not activate the tRNA^{Asp} gene. Precise deletion of the entire tRNA^{Arg} gene gives a similar result. Rescue of tRNA^{Asp} gene transcription is effected either by the precise deletion of both the tRNA^{Arg} gene and spacer or by the precise deletion of this gene with concomitant introduction of an artificial RNA polymerase III start site in the spacer. This artificial start site is ineffective if the tRNA^{Arg} gene is present upstream.

Eucaryotic tRNA genes are transcribed by RNA polymerase III (Pol III) (for a review, see reference 2). The recognition elements that specify transcription of tRNA genes have been shown to reside within the structural gene, although it has been shown that sequences upstream and downstream of the gene can play a role (4, 5, 15, 25). There are two internal control regions (ICRs) of highly conserved primary sequence which define the intragenic promoter of tRNA genes (3, 8, 10, 24): the 5′ ICR, or A-block, which maps to the dihydrouridine loop, and the 3′ ICR, or B-block, which maps to the TψCG loop. It would seem to follow that all of the tRNA genes should be transcribed as monocistronic transcripts since all tRNA genes contain these conserved sequences. This is generally the case. In *Saccharomyces cerevisiae*, for example, there are about 350 tRNA genes, and very few of these are clustered (9). There are, however, several sets of tRNA^{Arg}-tRNA^{Asp} pairs in *S. cerevisiae* (22), and in *Schizosaccharomyces pombe* there is a tRNA^{Ser}-tRNA^{Met} pair (17). In both cases, the two tRNAs are transcribed together to give a dimeric tRNA precursor. Thus, only one of the tRNA genes is normally recognized by the transcription apparatus, and this gene serves as the promoter for the other. To determine which of the genes serves as promoter, we carried out a transcriptional study of the tRNA^{Arg}-tRNA^{Asp} pair using an in vitro system derived from a nuclear extract of *S. cerevisiae* (6, C. S. Parker and J. Topol, unpublished data). We found that linear DNA fragments resulting from restriction endonuclease digestion of the plasmid DNA carrying the tRNA^{Arg}-tRNA^{Asp} genes are excellent templates for RNA synthesis. When restriction sites in the downstream tRNA^{Asp} gene are cleaved, the intact tRNA^{Arg} gene is transcribed, but when the tRNA^{Arg} gene is cleaved, there is no transcription of the intact tRNA^{Asp} gene (12). This experiment showed that the tRNA^{Arg} gene serves as the promoter for the production of the dimeric transcript and further that even when the tRNA^{Arg} gene is inactivated by cleavage, the tRNA^{Asp} gene is incapable of promoting its own transcription.

The question remains: why does the tRNA^{Asp} gene fail to direct its own transcription? We considered three possibilities. (i) The tRNA^{Asp} ICR sequences are intrinsically weak promoter sequences. This hypothesis seems invalidated by the fact that not all tRNA^{Asp} genes are found as tRNA^{Arg}-tRNA^{Asp} pairs. In fact, a monomeric tRNA^{Asp} gene has been isolated which has a sequence identical to that found in the tRNA^{Arg}-tRNA^{Asp} gene tandem (7). Nonetheless, this hypothesis seemed worth testing since it had not been proved that the isolated tRNA^{Asp} gene is transcriptionally active in vivo.

(ii) The binding of a transcription factor to the upstream tRNA^{Arg} gene physically prevents recognition and binding of the transcriptional apparatus to the downstream tRNA^{Asp} gene. It is known that a transcription factor binds specifically and tightly to sequences in the B-block (13). This factor interacts strongly with about 20 base pairs (bp) in the B-block and less tightly to sequences upstream of that. It seems possible that binding of this factor to the upstream tRNA^{Arg} gene could sterically prevent initiation of transcription of the tRNA^{Asp} gene 10 to 15 bp downstream from the binding domain.

(iii) The sequences in the 10-bp intergenic spacer or in the tRNA^{Arg} gene or both may be inhibitory. It is known that certain 5′-flanking sequences to tRNA genes can inhibit transcription (4, 5), although the mechanism of inhibition is not clear.

We used the technique of oligonucleotide-directed mutagenesis (29) to construct precise alterations of the tRNA^{Arg}-tRNA^{Asp} gene pair. These mutants have been used as templates to direct the synthesis of tRNA in a homologous Pol III system (6; Parker and Topol, unpublished data) to determine what alterations allow for independent expression of the tRNA^{Asp} gene.

MATERIALS AND METHODS

pJB19f DNA (1) containing the tRNA^{Arg}-tRNA^{Asp} gene pair and pYSUP6 DNA (11) used as a control template for in vitro transcription experiments have been previously described. All radioisotopes were obtained from Amersham

* Corresponding author.
† Present address: MRC Laboratory of Molecular Biology, Cambridge, CB2 2QH, England.

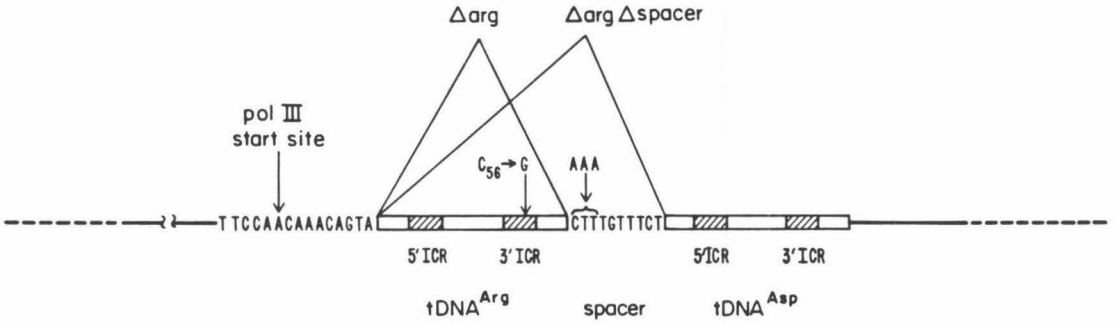


FIG. 1. Organization of the $tRNA^{Arg}$-$tRNA^{Asp}$ gene system. The $tRNA^{Arg}$ and $tRNA^{Asp}$ genes are shown as hollow boxes, with their A- and B-blocks as hatched regions. Neither gene contains introns. The various mutations are indicated. Each gene is 72 bp long, and the spacer is 10 bp. Yeast flanking regions, shown as thick black lines, are 119 bp upstream of the gene tandem and 45 bp downstream of it. Altogether, the yeast insert is 318 bp cloned into the unique *Sma*I site of the M13mp8 phage vector, which is indicated as broken lines.

Corp. (Arlington Heights, Ill.). Restriction endonucleases were from Boehringer Mannheim Biochemicals (Indianapolis, Ind.), while DNA polymerase I, large fragment (Klenow fragment), T4 DNA ligase, and T4 polynucleotide kinase were from either Bethesda Research Laboratories, Inc. (Gaithersburg, Md.) or New England BioLabs, Inc. (Beverly, Mass.). Mini- and large-scale preparations of double-stranded, replicative-form (RF) DNA were done by the alkaline lysis procedure of Maniatis et al. (16). Large-scale preparation of single-stranded bacteriophage DNA was done by the method of Yamamoto et al. (28). For DNA sequencing, single-stranded phage DNA was recovered from a clarified culture supernatant by polyethylene gycol-NaCl precipitation, followed by phenol extraction. Bacterial transformation was done by the method of Maniatis et al. (16). Polyacrylamide gel electrophoresis for DNA sequencing and transcription reactions were done in 8 or 14% thin polyacrylamide gels containing 7.8 M urea. Dideoxy and chemical sequencing of DNA were done by methods described previously (18, 21). Chemical DNA syntheses either were done manually via phosphotriester chemistry (19) or were made on an Applied Biosystems 380A DNA synthesizer (deoxyoligonucleotides were kindly provided by S. H. Horvath). Yeast nuclear extract used for in vitro transcription experiments was prepared by the method of C. Parker and J. Topol (unpublished data), as described by Engelke et al. (6). Oligonucleotide-directed mutagenesis and in vitro Pol III transcription were performed as previously described (20). DNase I protection analyses (footprinting) were done by the method of Schmitz and Galas (23) as previously described (20), except that radioactive labeling of the DNA fragment was done using [α-^{32}P]dCTP and Klenow fragments instead of [γ-^{32}P]ATP and T4 polynucleotide kinase.

RESULTS AND DISCUSSION

The organization of the $tRNA^{Arg}$-$tRNA^{Asp}$ gene system is shown in Fig. 1. This gene system had been cloned as a *Hin*dIII yeast fragment of about 3.8 kilobase pairs (2.5×10^6 daltons) in the plasmid pBR313 to give the recombinant plasmid pJB19f (1). A 318-bp fragment containing the tandem gene pair was produced from pJB19f by cleaving it with *Hae*III and *Hpa*I restriction endonucleases. This fragment was cloned into the unique *Sma*I site in the phage vector M13mp8 to obtain a source of single-stranded DNA for oligonucleotide-directed mutagenesis. The orientation of the insert in the phage vector is such that the sequence of the single-stranded DNA is homologous with the sequences of $tRNA^{Arg}$ and $tRNA^{Asp}$. We refer to this recombinant phage as M13ArgAsp. The double-stranded RF DNA isolated from M13ArgAsp-infected cells served as the template for Pol III transcription of the $tRNA^{Arg}$-$tRNA^{Asp}$ tandem gene pair. The products of transcription are shown in Fig. 2 (lane 1). The products and extent of transcription from this template were identical to those obtained with pJB19f plasmid DNA (lane 15), so the 318-bp fragment cloned into M13mp8 contains all of the sequences required for synthesis and processing of the two tRNAs.

We have previously characterized the transcription and processing products of the $tRNA^{Arg}$-$tRNA^{Asp}$ gene system in some detail (6). The largest product, band A, is the de novo transcript consisting of a 5′ leader, $tRNA^{Arg}$, the 10-bp intergenic spacer, $tRNA^{Asp}$, and a variable number of 3′ trailer residues including the terminal uridines which specify termination of transcription by Pol III. The de novo transcript is processed in this extract to give the mature $tRNA^{Arg}$ (band D) and $tRNA^{Asp}$ (band C). Band B is the principal intermediate in processing and consists of $tRNA^{Arg}$ flanked by the 5′ leader and spacer sequences. In this extract, the spacer sequence at the 3′ end of the principal intermediate is removed by an exonuclease.

Mutation in $tRNA^{Arg}$ 3′ ICR. In the introduction, we suggested that binding of a transcription factor to the $tRNA^{Arg}$ ICRs could prevent initiation of transcription at the $tRNA^{Asp}$ gene. We have shown with a yeast $tRNA_3^{Leu}$ gene that mutation of the conserved C-56 to G in the TψCG sequence of the 3′ ICR results in a 20-fold decrease in the template activity of the $tRNA_3^{Leu}$ gene in a yeast nuclear extract (20). Previously, it had been shown that an equivalent change in the $tRNA^{Tyr}$ *SUP4* gene results in loss of suppression in vivo and of template activity in vitro (14). The decrease in template activity is correlated with a complete loss of transcription factor binding to the 3′ ICR of the $tRNA_3^{Leu}$ gene as judged by the footprinting assay. We therefore decided to construct a C-56 to G mutation in the $tRNA^{Arg}$ 3′ ICR to test the possibility that this change would unmask the $tRNA^{Asp}$ promoter (mutant designated M13C56G). The mutation was constructed by oligonucleotide-directed mutagenesis (29). The C-56 to G change destroys a *Taq*I restriction site, and this could readily be seen in restriction digests of the RF DNA (data not shown). In addition to this evidence, the C-56 to G change was confirmed by dideoxy sequencing with a synthetic oligonucleotide primer specific for the 3′-flanking region of the $tRNA^{Asp}$

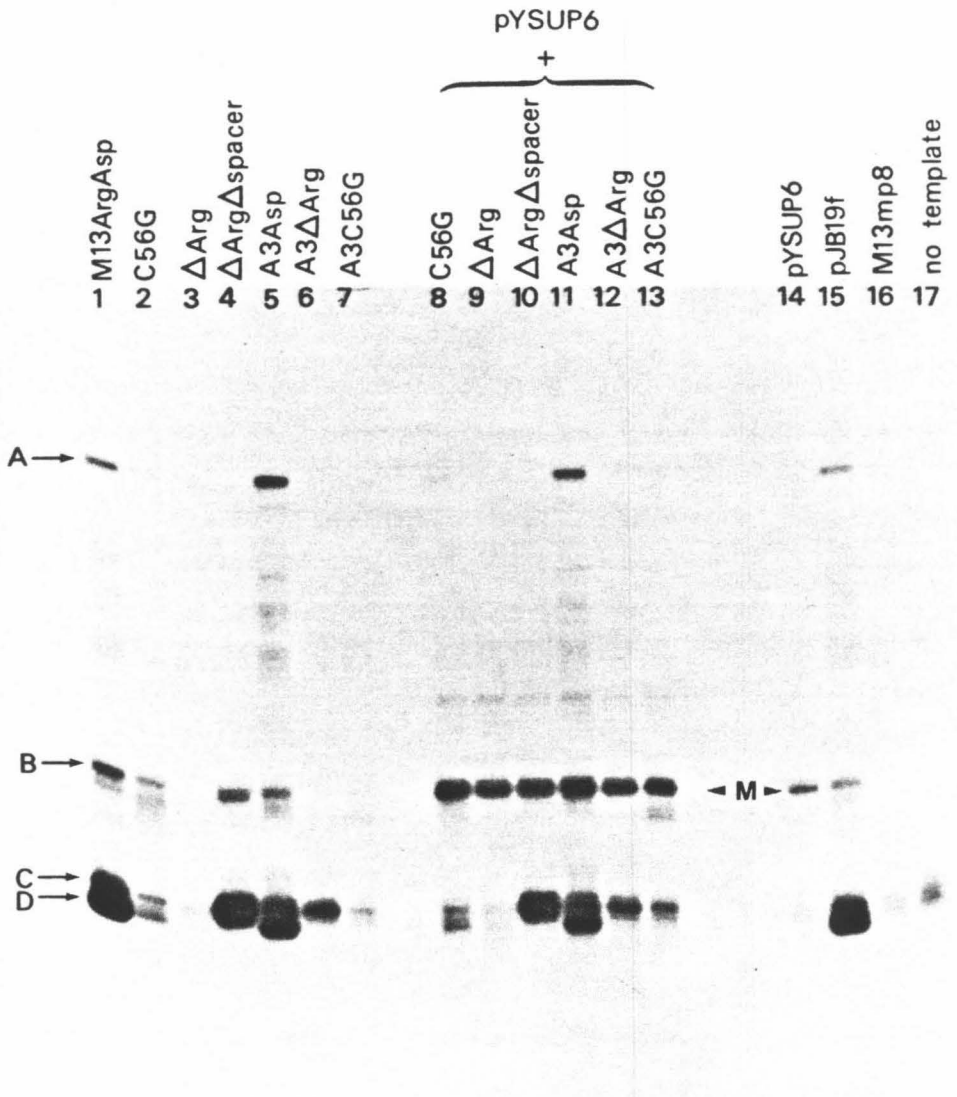


FIG. 2. In vitro transcriptional analysis of the wild-type $tRNA^{Arg}$-$tRNA^{Asp}$ gene system and its various mutant variants. Detailed protocol for the transcription assay is in Materials and methods. The DNA templates are indicated at the top of each lane. Lanes 8 to 13 are mixed-template transcription controls, such that to each corresponding template in lanes 1 to 7 has been added an equal amount of pYSUP6 DNA. The *SUP6* transcript has a size distinct from that of either $tRNA^{Arg}$ or $tRNA^{Asp}$. Band A is the primary dimeric transcript containing the two tRNAs joined together by the spacer sequence, including a 5′ leader and a 3′ trailer sequence. Band B is a major processing intermediate which contains $tRNA^{Arg}$ flanked by the 5′ leader sequence and the spacer sequence. Bands C and D are the mature $tRNA^{Asp}$ and $tRNA^{Arg}$ molecules, respectively. Band M is the pYSUP6 transcript.

gene. The sequences of all mutations described below were confirmed by the same procedure.

When M13C56G RF DNA was used as the template in the nuclear extract, it was found that the C-56 to G mutation results in decrease in template efficiency of about 20-fold when compared with that of M13ArgAsp (Fig. 2, lane 2). This is similar to the effect of the same mutation in the $tRNA_3^{Leu}$ gene. To exclude the possibility that there is an inhibitory substance in the M13C56G reaction mixture, DNA mixing experiments were performed in which an equal amount of pYSUP6 DNA containing the $tRNA^{Tyr}$ *SUP6* gene (which contains a 14-bp intron) was added to the reaction mixture. Transcription of this DNA produced unspliced products which were distinct in size from the mature $tRNA^{Arg}$-$tRNA^{Asp}$ transcripts (Fig. 2, lane 14). This control was performed with each of the mutants, and in no case was the transcription of the pYSUP6 gene inhibited (Fig. 2, lanes 8 to 13).

The C-56 to G change is believed to inactivate the $tRNA^{Arg}$ gene transcription by preventing transcription factor binding to its 3′ ICR. This has been demonstrated directly in the $tRNA^{Arg}$-$tRNA^{Asp}$ gene system by doing DNase I protection assays (23) on the wild-type M13ArgAsp DNA and the C-56 to G mutant (Fig. 3). In the wild-type M13ArgAsp DNA, a footprint (absence of bands) was readily visible in and around the B-block of the $tRNA^{Arg}$

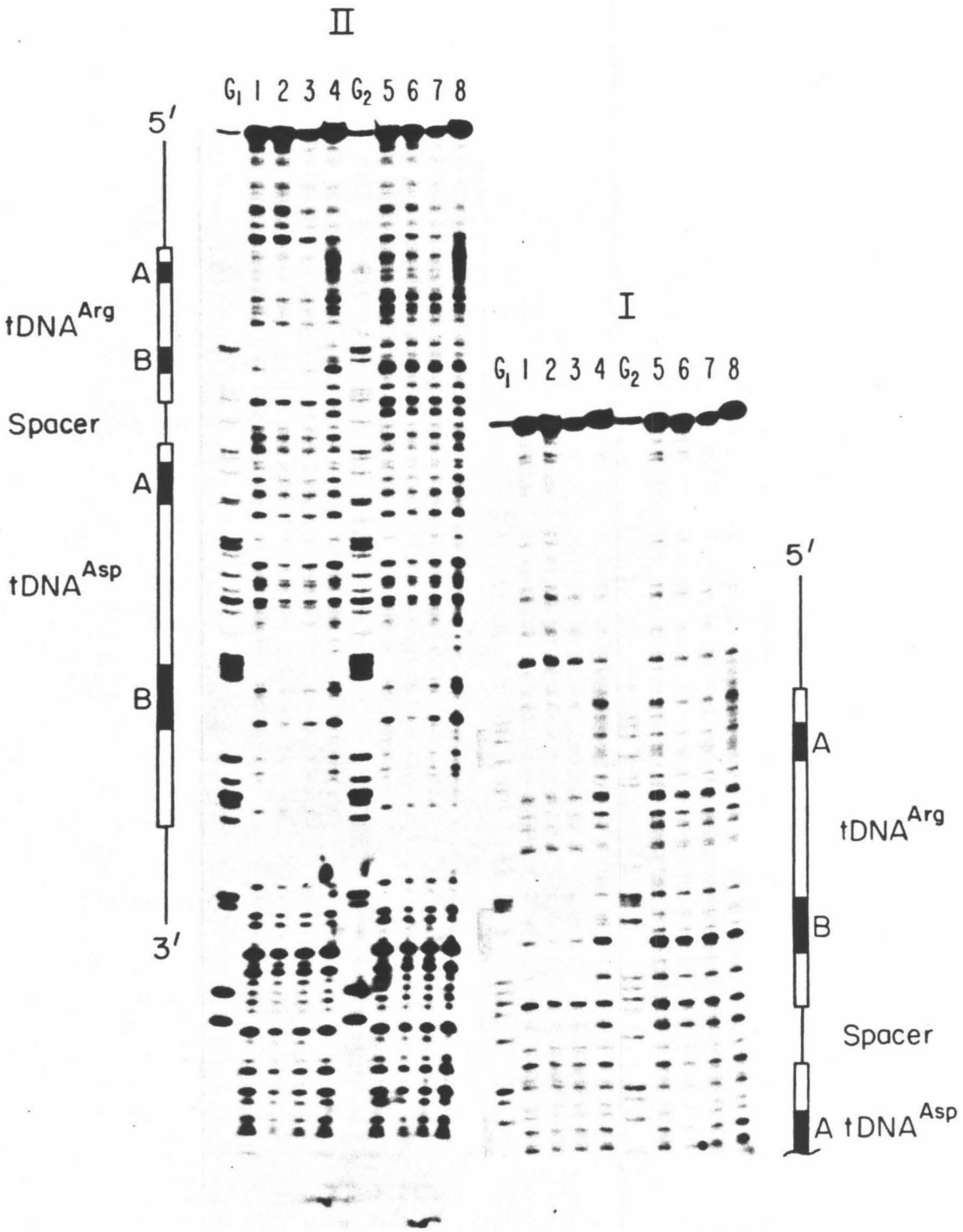


FIG. 3. DNase I protection analysis of M13C56G. A 333-bp *Eco*RI-*Acc*I restriction fragment labeled at the 3′ end (*Acc*I end) and carrying the $tRNA^{Arg}$-$tRNA^{Asp}$ gene tandem was subjected to the footprint assay of Schmitz and Galas (23) (see Materials and methods). Two gel loadings were done to visualize more clearly the transcription factor-binding activity of the $tRNA^{Arg}$ gene. The lanes headed by I constitute the first loading; those headed by II constitute the second loading. Lanes G_1 and G_2 are the Maxam-Gilbert G lanes of M13ArgAsp and M13C56G, respectively. Lanes 1, 2, and 3 are the plus-factor lanes for M13ArgAsp, and lanes 5, 6, and 7 are the corresponding ones for M13C56G. In lanes 1 and 5, the end-labeled fragment was preincubated with 1 μl of transcription factor solution per reaction before DNase I treatment; in lanes 2 and 6, 2 μl; and in lanes 3 and 7, 3 μl. Lanes 4 and 8 are the minus-factor control lanes for M13ArgAsp and M13C56G, respectively, in which the transcription factor preincubation step was omitted. The positions of the $tRNA^{Arg}$ and $tRNA^{Asp}$ genes, together with their A- and B-blocks, are indicated by the adjoining diagrams.

gene, as well as around its A-block, but no such footprints were visible in the $tRNA^{Arg}$ gene of the C-56 to G mutant. There is some protection around the B-block sequence of the $tRNA^{Asp}$ gene of both DNAs, and a little around the A-block sequence. These results immediately eliminate the steric hindrance hypothesis described in the Introduction. The C-56 to G mutation in the B-block has this same effect on both the $tRNA_3^{Leu}$ and $tRNA^{Tyr}$ genes (14, 20). It therefore appears that prevention of transcription factor binding to the 3′ ICR of the $tRNA^{Arg}$ gene does not unmask the $tRNA^{Asp}$ gene promoter, but it remains possible that the 5′ ICR or other sequences in the $tRNA^{Arg}$ gene prevent independent $tRNA^{Asp}$ gene transcription.

Precise deletion of the $tRNA^{Arg}$ gene and the spacer sequence. The technique of oligonucleotide-directed mutagenesis can be used to produce precise deletions as well as point mutations (27). To provide a more stringent test of the effects of the $tRNA^{Arg}$ gene on $tRNA^{Asp}$ expression, we constructed a precise deletion of the $tRNA^{Arg}$ gene. This deletion joined the 5′-flanking sequences of the $tRNA^{Arg}$ gene directly to the 10-bp spacer region (Fig. 1). RF DNA (designated M13ΔArg) carrying this mutation was inactive in transcription (Fig. 2, lane 3). No detectable $tRNA^{Asp}$ was produced. The transcription products were identical to those seen when the control M13mp8 RF DNA was used as the template (Fig. 2, lane 16). Since a wild-type 5′-flanking region and both $tRNA^{Asp}$ gene ICRs were present in this inactive arrangement, we considered that sequences in the spacer could be having an inhibitory effect. To test this hypothesis, we constructed a precise deletion of the spacer region, together with the entire $tRNA^{Arg}$ gene. This deletion joined the 5′-flanking sequence of the $tRNA^{Arg}$ gene directly to the $tRNA^{Asp}$ gene. The RF DNA (designated M13ΔArgΔspacer) carrying this mutation was transcriptionally active, and the product was $tRNA^{Asp}$ (Fig. 2, lane 4). This was confirmed by elution of the tRNA product from the gel, digestion with T_1 RNase, and separation and characterization of the digestion products by two-dimensional thin-layer chromatography (26) (data not shown). A precursor band was also seen in the M13ΔArgΔspacer transcription, and although we did not characterize this product further, its size, about 10 bases larger than $tRNA^{Asp}$, is consistent with the notion that the start site used for wild-type dimer transcription is also utilized for the transcription of the $tRNA^{Asp}$ monomeric gene. We conclude from this experiment that the $tRNA^{Asp}$ promoter is, as expected, transcriptionally competent. Its activity, however, is masked when it is preceded by the 10-bp spacer sequence. This result emphasizes the importance of 5′-flanking sequences in tRNA gene transcription by Pol III. We considered two possible explanations for the inhibitory effects of the spacer sequence on $tRNA^{Asp}$ transcription. First, there is the ever-present "poison sequence" hypothesis. The spacer sequence could possibly assume a configuration which prevents initiation by Pol III or recognition by the transcription factors. Alternatively, it is possible that the sequence of the spacer is inappropriate for initiation of transcription by Pol III. Pol III transcripts in yeasts generally begin with A. Many yeast tRNA genes contain the sequence $(Py)_3CAACAAA$ in the 5′-flanking region (6). Initiation occurs at an A residue about 10 bases upstream from the mature 5′ end of the tRNA. The spacer sequence does not contain any A residue. We elected to test the possibility that the introduction of a string of A residues into the spacer region in M13ΔArg can activate transcription of the $tRNA^{Asp}$ gene.

Introduction of three A residues into the spacer sequence. By oligonucleotide-directed mutagenesis of M13ΔArg, the spacer sequence CTTTGTTTCT was altered to AAATGT TTCT. The RF DNA (designated M13A3ΔArg) was transcribed, and the product was $tRNA^{Asp}$ (Fig. 2, lane 6). Thus, the negative effect of the spacer can be overcome by introduction of A residues 8 bp upstream from the 5′ end of the $tRNA^{Asp}$ gene.

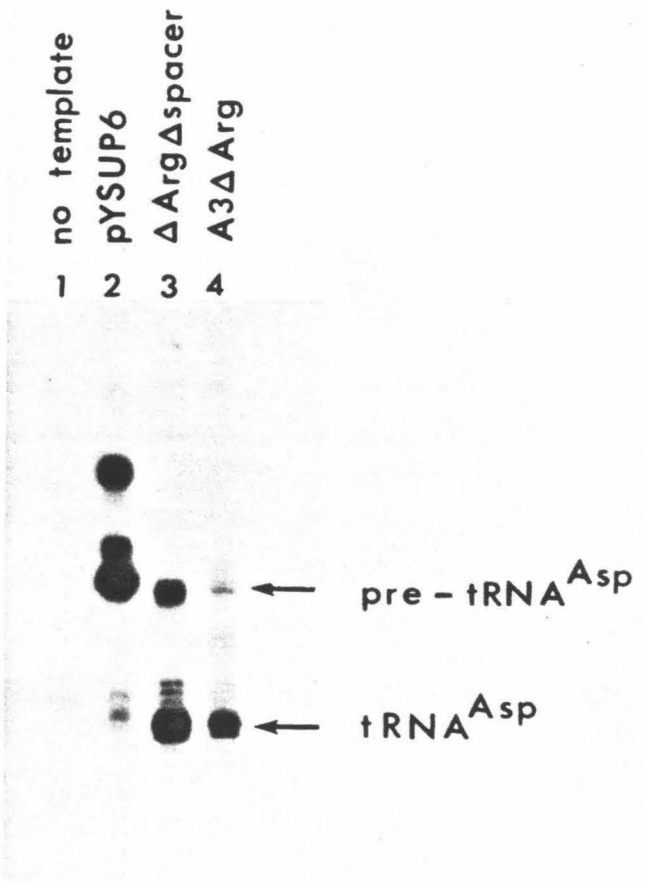


FIG. 4. Evidence of utilization of the artificially introduced Pol III start site in the M13A3ΔArg double mutant. Transcription of the M13ΔArgΔspacer and M13A3ΔArg mutants were run in parallel. Overexposure of the gel (5 days at −70°C) reveals that the precursors for $tRNA^{Asp}$ produced from both mutants have the same electrophoretic mobility and are most probably the same size (lanes 3 and 4). It is known that transcription starts at the A residue at position −10 in the wild-type M13ArgAsp (6). On the basis of the relative sizes of precursor and mature transcript, we assume that this is the same start site utilized for precursor production in the M13ΔArgΔspacer. Since the artificially introduced Pol III start site is at position −10 relative to the $tRNA^{Asp}$ gene in M13A3ΔArg, we conclude that it is indeed the one being utilized.

The precursor for M13A3ΔArg transcription was visible upon overexposure of the gel (5 days at −70°C) and was about 8 to 10 bp larger than the mature product ($tRNA^{Asp}$) as judged by their relative electrophoretic mobilities (Fig. 4). It also comigrated exactly with the precursor tRNA for M13ΔArgΔspacer transcription. Although we did not characterize this product further, these results strongly suggest that the artificially introduced start site is functional in M13A3ΔArg.

The artificial Pol III start site was also introduced into the

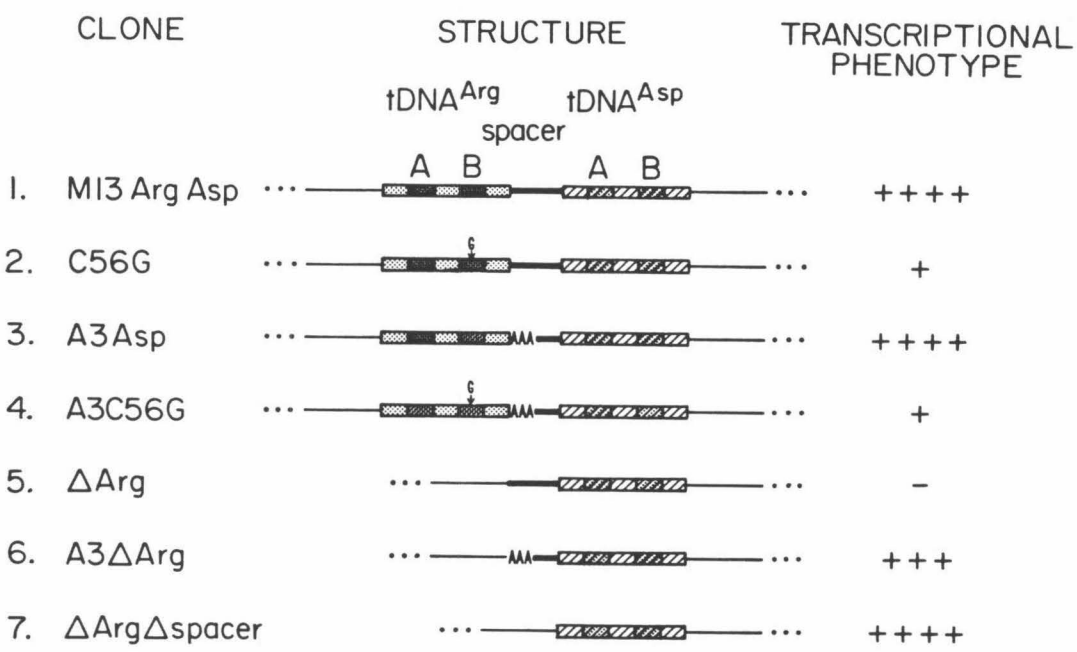


FIG. 5. Summary of results. The structures of the various mutants under study are indicated. The introduction of a G residue in the B-box of the $tRNA^{Arg}$ gene is indicated. The artificially introduced Pol III start site is indicated as AAA. Levels of transcription in vitro are indicated as plus signs, with wild-type level designated as ++++ and no transcription being −. These designations are all in reference to Fig. 2.

intergenic spacer of the wild-type gene tandem (mutant designated A3Asp) and into the C-56 to G mutant (designated A3C56G). Transcription of the wild-type gene was not affected by this alteration of the spacer sequence (Fig. 2, lane 5). Since the spacer sequence inhibited transcription of $tRNA^{Asp}$ in the ΔArg construct, it is possible that it is the spacer which inhibits the transcription of $tRNA^{Asp}$ in the C56G construct. If so, introduction of A3 into the intergenic spacer of C56G should activate $tRNA^{Asp}$ transcription. This, however, was not the case (Fig. 2, lane 7). Apparently, the inhibition of the $tRNA^{Asp}$ gene promoters in the tandem construct is complex.

The results of this study are summarized in Fig. 5. Independent transcription of the $tRNA^{Asp}$ gene was achieved by precise deletion of the $tRNA^{Arg}$ gene together with the 10-bp spacer. Deletion of the $tRNA^{Arg}$ gene sequence by itself was not sufficient to activate transcription of the $tRNA^{Asp}$ gene, suggesting an inhibitory effect of the spacer sequence. This inhibition was partially overcome by the introduction of three A residues at the 5′ edge of the intergenic spacer. The failure, however, of this sequence to activate transcription of $tRNA^{Asp}$ in the $tRNA^{Arg}$ C-56 to G mutant suggests that the $tRNA^{Arg}$ sequences, as well as the spacer region, inhibit the utilization of the $tRNA^{Asp}$ gene promoters in the tandem by Pol III. Therefore, it must be the sequence of the $tRNA^{Arg}$ gene together with the spacer sequence which suppresses an otherwise competent $tRNA^{Asp}$ intragenic promoter.

ACKNOWLEDGMENTS

We thank Peter Johnson, David Engelke, and Marjorie Strobel for their help and for useful discussions.

This work was supported by a grant from the American Cancer Society (NP302D) and Public Health Service grant GM30356 from the National Institutes of Health.

LITERATURE CITED

1. **Beckmann, J. S., P. F. Johnson, and J. N. Abelson.** 1977. Cloning of yeast transfer RNA genes in *Escherichia coli*. Science **196**:205–208.
2. **Ciliberto, G., L. Castagnoli, and R. Cortese.** 1983. Transcription by RNA polymerase III. Curr. Top. Dev. Biol. **18**:59–88.
3. **Ciliberto, G., L. Castagnoli, D. A. Melton, and R. Cortese.** 1982. Promoter of a eukaryotic $tRNA^{Pro}$ gene is composed of three noncontiguous regions. Proc. Natl. Acad. Sci. USA **79**:1195–1199.
4. **DeFranco, D., S. Sharp, and D. Soll.** 1981. Identification of regulatory sequences contained in the 5′ flanking region of *Drosophila* lysine $tRNA_2$ genes. J. Biol. Chem. **256**:12424–12429.
5. **Dingermann, T., D. S. Burke, S. Sharp, J. Schaack, and D. Soll.** 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.
6. **Engelke, D. R., P. Gegenheimer, and J. Abelson.** 1985. Nucleolytic processing of a $tRNA^{Arg}$-$tRNA^{Asp}$ dimeric precursor by a homologous component from *Saccharomyces cerevisiae*. J. Biol. Chem. **20**:1271–1279.
7. **Gafner, J., E. M. De Robertis, and P. Philippsen.** 1983. Delta sequences in the 5′ noncoding region of yeast tRNA genes. EMBO J. **2**:583–591.
8. **Galli, G., H. Hofstetter, and M. L. Birnstiel.** 1981. Two conserved sequence blocks within eucaryotic tRNA genes are major promoter elements. Nature (London) **394**:626–631.
9. **Guthrie, C., and J. Abelson.** 1982. Organization and expression of tRNA genes in *Saccharomyces cerevisiae*, p. 487–528. *In* J. N. Strathern, E. W. Jones, and J. R. Broach (ed.), Molecular biology of the yeast *Saccharomyces*: metabolism and gene expression. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
10. **Hofstetter, H., A. Kressmann, and M. L. Birnstiel.** 1981. A split promoter for a eucaryotic tRNA gene. Cell **24**:573–585.
11. **Johnson, P. F., and J. Abelson.** 1983. The yeast $tRNA^{Tyr}$ gene intron is essential for correct modification of its tRNA product. Nature (London) **302**:681–687.

12. **Kjellin-Straby, K., D. R. Engelke, and J. Abelson.** 1984. Homologous in vitro transcription of linear DNA fragments containing the $tRNA^{Arg}$-$tRNA^{Asp}$ gene pair from *Saccharomyces cerevisiae*. DNA **3**:167–171.
13. **Klemenz, R., D. J. Stillman, and E. P. Geiduschek.** 1982. Specific interactions of *Saccharomyces cerevisiae* proteins with a promoter region of eukaryotic tRNA genes. Proc. Natl. Acad. Sci. USA **79**:6191–6195.
14. **Koski, R. A., D. S. Allison, M. Worthington, and B. D. Hall.** 1982. An in vitro RNA polymerase III system from *S. cerevisiae*: effects of deletions and point mutations upon SUP4 gene transcription. Nucleic Acids Res. **10**:8127–8143.
15. **Larson, D., J. Bradford-Wilcox, L. S. Young, and K. U. Sprague.** 1983. A short 5′ flanking region containing conserved sequences is required for silkworm alanine tRNA gene activity. Proc. Natl. Acad. Sci. USA **80**:3416–3420.
16. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
17. **Mao, J., O. Schmidt, and D. Soll.** 1980. Dimeric transfer RNA precursors in *S. pombe*. Cell **21**:509–516.
18. **Maxam, A. M., and W. Gilbert.** 1977. A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA **74**:560–564.
19. **Miyoshi, K., T. Huang, and K. Itakura.** 1980. Solid-phase synthesis of polynucleotides. III. Synthesis of polynucleotides with defined sequences by the block-coupling phosphotriester method. Nucleic Acids Res. **8**:5491–5505.
20. **Newman, A. J., R. C. Ogden, and J. Abelson.** 1983. tRNA gene transcription in yeast: effects of specified base substitutions in the intragenic promoter. Cell **35**:117–125.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Schmidt, O., J. Mao, R. Ogden, J. Beckmann, H. Sakano, J. Abelson, and D. Soll.** 1980. Dimeric tRNA precursors in yeast. Nature (London) **287**:750–752.
23. **Schmitz, A., and D. J. Galas.** 1983. The study of protein-DNA binding specificity: DNAse footprinting, p. 305–347. *In* S. M. Weissman (ed.), Methods of DNA and RNA sequencing. Praeger Press, New York.
24. **Sharp, S., D. DeFranco, T. Dingermann, P. Farrell, and D. Soll.** 1981. Internal control regions for transcription of eucaryotic tRNA genes. Proc. Natl. Acad. Sci. USA **78**:6657–6661.
25. **Sprague, K. U., D. Larson, and D. Morton.** 1980. 5′ flanking sequence signals are required for activity of silkworm alanine tRNA genes in homologous in vitro transcription systems. Cell **22**:171–178.
26. **Volckaert, G., and W. Fiers.** 1977. Micro thin layer techniques for rapid sequence analysis of ^{32}P-labeled RNA: double digestion and pancreatic ribonuclease analyses. Anal. Biochem. **83**:228–239.
27. **Wallace, R. B., P. F. Johnson, S. Tanaka, M. Schöld, K. Itakura, and J. Abelson.** 1980. Directed deletion of a yeast transfer RNA intervening sequence. Science **209**:1396–1400.
28. **Yamamoto, K. R., B. M. Alberts, R. Benzinger, L. Lawhorne, and G. Treiber.** 1970. Rapid bacteriophage sedimentation in the presence of polyethylene glycol and its application to large scale virus purification. Virology **40**:734–744.
29. **Zoller, M. J., and M. Smith.** 1982. Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. Nucleic Acids Res. **10**:6487–6500.

Chapter III

A SYNTHETIC SUBSTRATE FOR tRNA SPLICING

[The text of this chapter appeared in: Reyes, V. M. and Abelson, J. (1987). **Anal. Biochem.** *166*, 90-106.]

ANALYTICAL BIOCHEMISTRY **166,** 90–106 (1987)

A Synthetic Substrate for tRNA Splicing[1]

VICENTE M. REYES* AND JOHN ABELSON†

*Divisions of *Chemistry and †Biology, California Institute of Technology, Pasadena, California 91125*

Received May 4, 1987

A novel method for the synthesis of precursor tRNA as substrate for *in vitro* splicing is reported. A construct consisting of the *Saccharomyces cerevisiae* pre-tRNA^{Phe} gene under the control of a bacteriophage T7 promoter was assembled from a set of synthetic oligonucleotides and cloned into an M13 vector. By the use of T7 RNA polymerase, *Bst*NI-runoff transcripts were produced. The resulting pre-tRNA was shown to possess mature termini and was accurately spliced by highly purified yeast tRNA-splicing endonuclease and ligase. Using this synthetic pre-tRNA, the kinetic parameters of the tRNA-splicing endonuclease were also determined. Use of this system provides several advantages for the study of tRNA-splicing mechanisms. Mutant tRNA precursors can be readily synthesized. It is also possible to synthesize large quantities of pre-tRNA for structural studies.

KEY WORDS: yeast tRNA splicing; T7 RNA polymerase; synthetic pre-tRNA; *S. cerevisiae* pre-tRNA^{Phe}.

The *Saccharomyces cerevisiae* haploid genome contains approximately 360 tRNA genes, and about a tenth of these contain intervening sequences (IVSs)[2] that interrupt the gene (1). In all cases, the start of transcription by RNA polymerase III (pol III) is a few bases upstream of the 5′ end of the gene and the termination is several bases downstream of its 3′ end (2–4). Therefore, the initial product of transcription is an RNA molecule which contains the tRNA sequence, the IVS, and short 5′ and 3′ extensions. Before it can be functional as a tRNA, this primary transcript must undergo a complex series of maturation steps (for a review, see (5)) which include (a) the removal of the 5′ and 3′ extensions and addition of the trinucleotide CCA to the resulting 3′ end to produce mature termini; (b) the excision of the intron and joining of the resulting half-tRNAs; and (c) the introduction of certain base modifications. Each of these steps is accomplished by a unique enzyme or set of enzymes, which act in specific sequential order.

Previous studies on tRNA biosynthesis suggest that, among the IVS-containing tRNAs, splicing is one of the last steps in the maturation process (6). The primary transcript has to be end-matured and certain base modifications added before the two splicing enzymes—an endonuclease to excise the IVS and a ligase to join the resulting half-tRNAs —act. There is as yet no evidence, at least in yeast, for the removal of IVS prior to 5′ and 3′ end maturation (7). The splicing substrate is therefore a pre-tRNA molecule possessing mature ends and certain base modifications, but still interrupted by an IVS.

Previously, it has also been shown that mutations affecting the maturation steps map to the mature domain (corresponding to the two tRNA halves) of the tRNA gene (8–10). End maturation is the step most sensitive to these mutations, although splicing and base modifications may also be blocked (10). Thus, there is appreciable polarity in

[1] This paper is dedicated to the memory of Professor Nathan O. Kaplan.

[2] Abbreviations used: IVS, intervening sequence; pol III, RNA polymerase III; BSA, bovine serum albumin; SDS, sodium dodecyl sulfate; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; DTT, dithiothreitol; RF, replicative form; YNE, yeast nuclear extract.

0003-2697/87 $3.00

these mutations. A mutation which blocks an early step may affect later ones; for example, a block in end maturation could produce an inappropriate splicing substrate.

We are interested in studying mutations in IVS-containing tRNA genes which particularly affect the splicing reaction, apart from effects on the other maturation steps. Our general approach is to construct predesigned mutations in the pre-tRNA gene, synthesize a splicing substrate carrying this mutation, react this pre-tRNA with a yeast tRNA-splicing endonuclease (11) and ligase (12) *in vitro,* and analyze the products by gel electrophoresis. However, due to the complexities of tRNA biosynthesis discussed above, a problem that naturally arises from this approach is that the mutation in question might block or adversely affect the biosynthetic steps prior to the splicing event, and indeed in the past it had been difficult and sometimes even impossible to synthesize tRNA precursors containing certain mutations. For example, it is of great interest to specifically alter the highly conserved residues (e.g., those in the D- and TψC-loops) or those otherwise engaged in tertiary base interactions in the molecule and then to examine the structural and functional effects of these changes. However, these residues largely map to the internal promoter elements of the tRNA gene and hence any alteration is likely to block transcription by pol III (13,14); or if the RNA is made at all, the resulting primary transcript may be rendered unrecognizable to RNase P for correct processing of its 5′ terminus (15).

In this paper, we report a new method for the synthesis of precursor tRNA. A synthetic gene which allows the *in vitro* synthesis of yeast pre-tRNA^{Phe} containing a 19-base IVS has been constructed (Fig. 1). In this construct, the pre-tRNA^{Phe} sequence is linked directly to a bacteriophage T7 promoter. The sequence is terminated with a *Bst*NI restriction endonuclease site. Run-off transcription

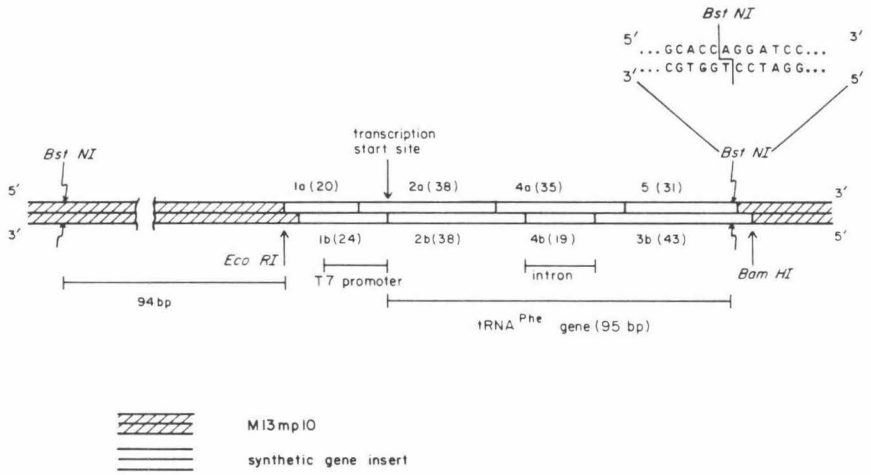


FIG. 1. The structure of the T7 promoter/pre-tRNA^{Phe} gene system. The entire construct is cloned into the *Eco*RI/*Bam*HI site of M13mp10 RF and is 124 base pairs long. The upper strand is homologous to the tRNA^{Phe} sequences. The eight synthetic oligonucleotides (**1a, 1b, 2a, 2b, 4a, 4b, 5,** and **3b**) together with their corresponding sizes (in number of bases) in parentheses are indicated. The region of the 3′ terminus of the pre-tRNA^{Phe} gene (*Bst*NI cleavage site) is enlarged (note the single T-overhang in the coding strand). The immediate upstream *Bst*NI cleavage site in M13Phe is 94 base pairs away from the insert, thus giving a *Bst*NI fragment of 215 base pairs as template for T7 RNA polymerase (recombinant plasmid M13Phe is 7349 base pairs long and is cleaved eight times by *Bst*NI to generate eight fragments of sizes 3975, 1809, 952, 215, 139, 127, 74, and 57 base pairs).

by T7 RNA polymerase (16,17) of *Bst*NI-cleaved DNA gives rise to a precursor which has mature termini and an IVS. Site-specific mutations can be introduced at will anywhere in the pre-tRNA gene and the corresponding mutant pre-tRNA can be tested as a splicing substrate, thus affording a systematic genetic study of tRNA splicing. Large amounts of transcript can be synthesized, allowing structural analysis of the pre-tRNA molecule. Yeast $tRNA^{Phe}$ was chosen for this purpose because more is known about its structure than that of any other tRNA molecule (18,19).

MATERIALS AND METHODS

Materials. T7 RNA polymerase was from United States Biochemical Corp. (Cleveland, OH) and *Bst*NI restriction endonuclease was from New England Biolabs (Beverly, MA). All radioactive ribonucleoside triphosphates were from Amersham Corp. (Arlington Heights, IL); RNasin was from Promega Biotech (Madison, WI). Double-stranded M13mp10 recombinant plasmid as transcriptional template was prepared by the method of Maniatis *et al.* (20), except that the CsCl-ethidium bromide gradient ultracentrifugation step was done twice; i.e., the plasmid band was removed from the first gradient and loaded onto a second, new gradient and centrifuged again. All preparative transcription gels were done in 8% polyacrylamide/4.0 M urea gels, 0.3 mm thick, 12 × 18 cm. All splicing assay gels were done in 10% polyacrylamide/7.8 M urea gels, 0.3 mm thick, 38 × 19 cm. Two-dimensional oligonucleotide mapping of RNase T1 or A digests of radioactively labeled RNA has been described previously (21). Secondary digestion analysis of oligonucleotides is described elsewhere (22). Base analysis by the method of Saneyoshi *et al.* was done as described (23).

Synthesis of pre-$tRNA^{Phe}$ by T7 RNA polymerase. Twenty microliters of supercoiled, double-stranded DNA template at 1.0 μg/μl is mixed with 10 μl of 10× transcription buffer (0.5 M NaCl/0.1 M Tris, pH 7.5/0.5 M $MgCl_2$/0.01 M DTT), 65 μl of dH_2O, and 5.0 μl of *Bst*NI (10 u/μl), for a total volume of 100 μl. This mixture is covered with paraffin oil and incubated at 65°C for 1.0 h. The aqueous phase is then transferred to another tube and the following reagents were added directly: 30.0 μl of NTP mix (1 mM UTP/5 mM ATP, CTP, and GTP); 35 μl of water; 5.0 μl of [α-^{32}P]UTP (3000 Ci/mmol); 5.0 μl of RNasin (40 u/μl); and 25 μl of T7 RNA polymerase diluted to 12.5 u/μl with polymerase diluent (1.0 μg/μl BSA/2.5 mM β-mercaptoethanol/4 mM Tris, pH 8). Incubation is resumed at 37°C for 1 h. Transcription may also be done using plasmid DNA that has been previously cleaved with *Bst*NI in large scale. This DNA is stored at a final concentration of 1.0 μg/μl in TE (10 mM Tris–HCl, pH 8.0/1 mM EDTA). In this case, 20 μl of the *Bst*NI-cut plasmid at 1.0 μg/μl is mixed with 4 μl of 10× transcription buffer, 24 μl of NTP mix, 4 μl of [α-^{32}P]UTP (3000 Ci/mmol), 20 μl of T7 RNA polymerase diluted to 12.5 u/μl, 2 μl of RNasin (40 u/μl), 2 μl of 4% Triton X-100, and enough water to bring the volume to 80 μl, and the reaction is incubated at 37°C as before. The reaction is then terminated by addition of 50–80 μl of carrier-containing dilution buffer (15 mM EDTA/0.05 μg/μl RNA or glycogen carrier/0.5% Sarkosyl) and by extraction with phenol:chloroform (1:1). The aqueous phase is ethanol precipitated, washed, and dried. The final pellet is dissolved in 20–25 μl of formamide loading buffer (95% formamide/0.5% each of xylene cyanol and bromphenol blue), and is loaded onto a transcription gel and run at 500–600 V. Under these conditions, the pre-tRNA (95 bases) comigrates with the XC dye; when XC migrates halfway down the gel, electrophoresis is stopped and the RNA band is visualized by autoradiography and sliced. If there is significant background smearing, a second electrophoresis is done as follows. The rest of the gel is lifted from the plate, leaving behind the RNA-

containing slice. A 20% polyacrylamide/4 M urea gel is then formed around the slice and the electrophoresis is resumed (in the same direction as before) until the XC dye reaches the bottom of the gel. This second electrophoresis in 20% polyacrylamide separates most of the background smear from the RNA. The RNA is then visualized by autoradiography and sliced out, and the RNA is isolated either by overnight diffusion in 0.4 ml of extraction buffer (0.5 M NH_4 acetate/0.05 M Mg acetate/1 mM EDTA/0.1% SDS) or by electroelution as described (Electroelution Manual, International Biotechnologies, Inc., New Haven, CT). This finally yields a total of approximately 2–5 $\times$ 10^6 cpm of transcript with a specific activity of approximately 8–9 $\times$ 10^4 dpm/pmol, or approximately 2–4 μg of pre-tRNA (pre-tRNA^{Phe} is 95 bases, $pppA_{26}C_{21}G_{26}U_{22}$, and has a molecular mass of 32.5 kDa). The RNA is precipitated using glycogen rather than RNA as carrier (carrier RNA inhibits splicing, M. Strobel, personal communication), and is dissolved in appropriate volume of TE (10 mM Tris, pH 7.3/1 mMEDTA) to give exactly 1000 cpm/μl, or approximately 10^{-3} μg/μl, or 35 nM pre-tRNA^{Phe}.

In vitro splicing assay with partially purified endonuclease and ligase. The following are mixed together (enzymes always added last): 2 μl of 1000 cpm/μl pre-tRNA^{Phe}; 5 μl of 2x splicing buffer (40 mM HEPES, pH 7.5/5 mM spermidine/10 mM $MgCl_2$/0.2 mM DTT/0.8% Triton X-100/4 mM ATP/20% glycerol); 1 μl of 1:25 dilution of endonuclease; 1 μl of 1:18 to 1:25 dilution of ligase; and enough water to give a total of 10 μl. The endonuclease has been purified approximately 200-fold over a Triton X-100 extract of crude yeast membranes (P. Green, personal communication); the ligase is a heptylagarose pool (Fraction IV) purified approximately 300- to 400-fold (E. Phizicky, personal communication). Endonuclease diluent is 20 mM Tris, pH 8.0/0.5 mM EDTA/1 mM DTT/20 mM $(NH_4)_2SO_4$/0.5% Triton X-100/20% glycerol; while ligase diluent is 20 mM sodium phosphate, pH 7.5/1 mM EDTA/1 mM β-mercaptoethanol/200 mM NaCl/10% glycerol. RNasin does not inhibit the splicing reaction and may be added at 0.5 μl/10 μl reaction (final concentration of 2 units/μl) if nuclease degradation is suspected. In "buffers-only" controls, diluents without enzymes are added; in "endonuclease-only" controls, ligase diluent without ligase is added. The final concentration of pre-tRNA in each 10-μl reaction is approximately 7–10 nM. Incubation is for 20 min at 30°C. Reactions are terminated by addition of 50–80 μl dilution buffer, extraction with phenol:chloroform, and ethanol precipitation. Electrophoresis is performed at 1500–1800 V for 1.5 h or until the bromphenol blue dye is approximately 2–3 in. from the bottom of the gel.

Determination of endonuclease kinetic parameters. Fifty-microliter endonuclease reaction mixtures containing different concentrations of pre-tRNA were prepared. The precursor concentrations were varied from 3 to 110 nM at a specific activity of 8–9 $\times$ 10^4 dpm/pmol. Ten-microliter aliquots were removed from each reaction mixture (30°C) at 0, 3, 6, and 9 min into formamide loading buffer, and the entire sample was loaded onto a 10% polyacrylamide gel/7.8 M urea, 1.5 mm thick, 12 $\times$ 18 cm. After electrophoresis, the gel was autoradiographed, bands (combined exons or introns) were excised, and the Cerenkov radiation was quantitated. Under these conditions the accumulation of products was linear with time.

RESULTS

Synthesis and splicing of the synthetic precursor tRNA. The yeast pre-tRNA^{Phe} gene (Fig. 1) linked to a bacteriophage T7 promoter (24) was constructed by self-assembly of eight synthetic oligodeoxynucleotides, as shown in Table 1. This construct was then cloned into the *Eco*RI/*Bam*HI site of phage M13mp10 replicative form (RF) DNA. M13mp10 RF was chosen as vector so as to

TABLE 1

OLIGONUCLEOTIDES FOR T7 PROMOTER/PRE-tRNA^{Phe} GENE CONSTRUCTION

```
                    (P)  ↓
1a        AATTGCTGCAGTAATACGAC
1b             CGACGTCATTATGCTGAGTGATAT

2a  TCACTATAGCGGATTTAGCTCAGTTGGGAGAGCGCCAG
2b               CGCCTAAATCGAGTCAACCCTCTCGCGGTCTGACTTCT

4a  ACTGAAGAAAAAACTTCGGTCAAGTTATCTGGAGG
4b               TTTTTGAAGCCAGTTCAAT

                                    (E)
5                TCCTGTGTTCGATCCACAGAATTCGCACCAG
3b  AGACCTCCAGGACACAAGCTAGGTGTCTTAAGCGTGGTCCTAG
```

Note. The eight oligonucleotides **1a, 1b, 2a, 2b, 4a, 4b, 5,** and **3b** are shown as base-paired duplexes. The upper strands (anticoding) are written 5′ to 3′; the lower strands (coding) are written 3′ to 5′. They self-assemble as shown in Fig. 1. The entire construct is cloned into the *Eco*RI/*Bam*HI site of M13mp10; the *Eco*RI site is not preserved but is converted to an *Eco*R* site (there is an *Eco*RI site far downstream in the gene [E]). A *Pst*I site (P) right next to the *Eco*RI site may be used together with the preserved *Bam*HI site for subcloning purposes. Start of transcription is right where oligonucleotide **2b** begins. The beginning of the 17-base pair T7 promoter is within the **1a**/**1b** duplex (↓). The *Bst*NI cleavage site right upstream from the *Bam*HI terminus is indicated by two arrowheads.

have a convenient source of single-stranded DNA for subsequent site-directed mutagenesis using synthetic oligonucleotides (25). This recombinant plasmid is designated M13Phe. The identity of the insert was proven by dideoxy sequencing (26) using as primer a 20-base oligonucleotide whose 3′ end is 8 bases downstream of the pre-tRNA^{Phe} gene (data not shown). In this construct transcription by T7 RNA polymerase starts exactly at the 5′ end of the pre-tRNA^{Phe} gene and terminates at the *Bst*NI site at the 3′ end. To synthesize pre-tRNA^{Phe}, the plasmid is first incubated with *Bst*NI restriction endonuclease which cleaves precisely at the 3′ end of the gene to give rise to a coding strand ending with 3′-GGT-5′, thus ensuring that the transcript terminates with 5′-CCA-3′. The restricted plasmid is then incubated with T7 RNA polymerase, the four nucleoside triphosphates, and the appropriate radioactive label. The T7 promoter/pre-tRNA^{Phe} gene construct was recloned into the *Pst*I/*Bam*HI site of the vector pUC13, to give a recombinant plasmid designated pUC13Phe. Transcription of M13Phe and pUC13Phe results in identical pre-tRNA^{Phe}, but it is easier to obtain larger quantities of pUC13 DNA.

Figure 2A shows the transcription products of pUC13Phe DNA produced under different conditions. In each case, the major band is a 95-base RNA, which corresponds to the pre-tRNA^{Phe} molecule (Fig. 2B). A single transcription reaction using 20 μg of DNA template and 250 units of T7 RNA polymerase in a total transcription volume of 80–100 μl yields approximately 2–5 $\times$ 10^6 cpm of precursor with a specific radioactivity of 8–9 $\times$ 10^4 dpm/pmol, which corresponds to approximately 2–4 μg of pre-tRNA. To prove that this transcript is indeed pre-tRNA^{Phe}, transcription was done using either [α-^{32}P]ATP or [α-^{32}P]CTP as label, and the transcripts were subjected to two-dimensional oligonucleotide mapping, using either RNase A or RNase T1. As seen in Figs. 3A–D, all labeled fragments are exactly those predicted from the pre-tRNA^{Phe} gene se-

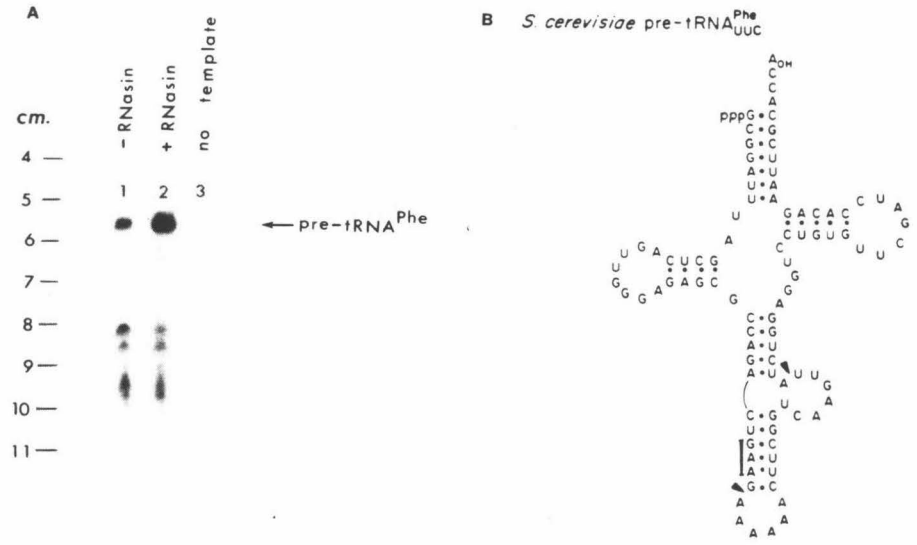
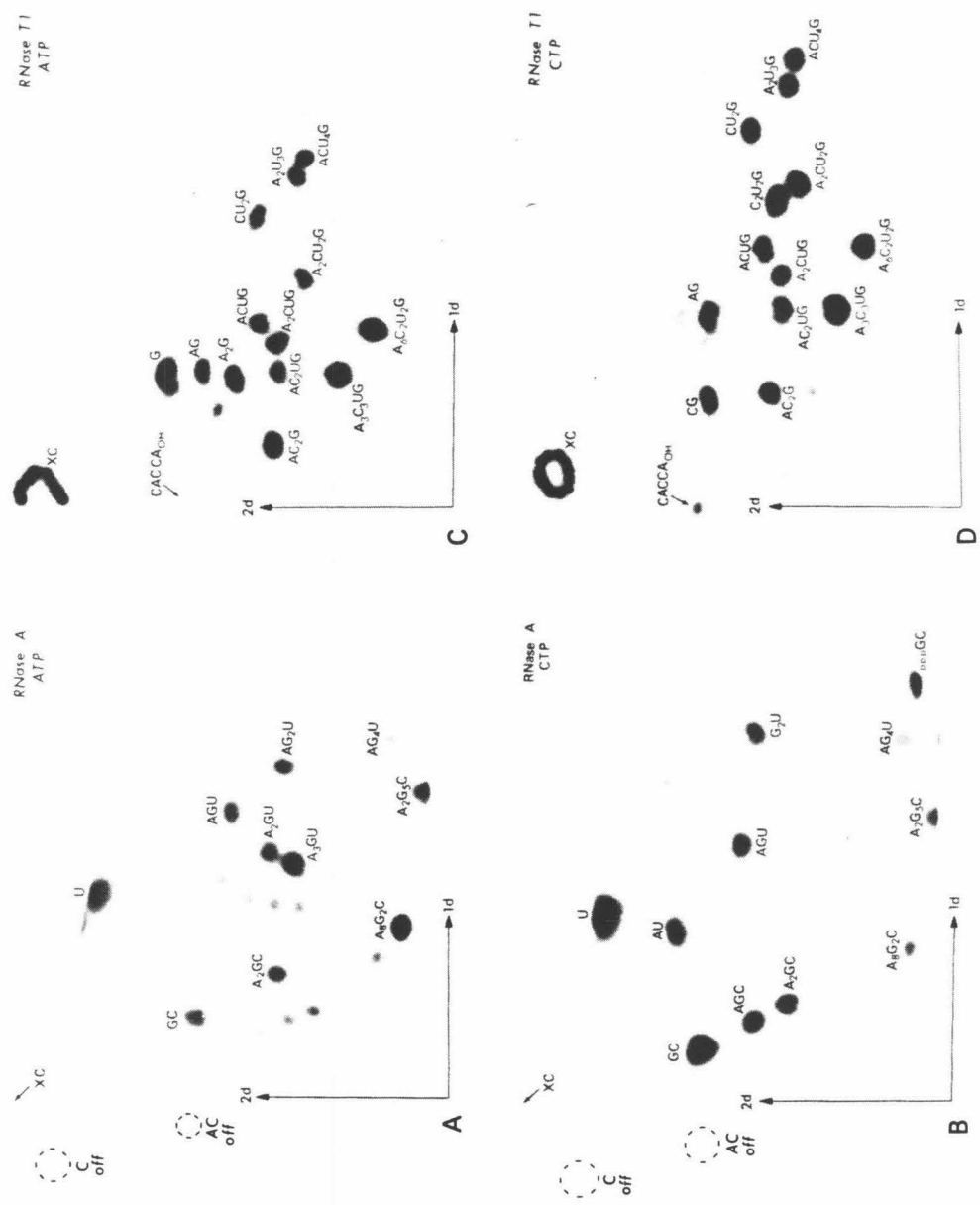


FIG. 2. (A) T7 transcription products of the T7 promoter/pre-tRNA^{Phe} gene system. pUC13Phe is cleaved with *Bst*NI and transcribed with T7 RNA polymerase as described under Materials and Methods. The major band (arrow) is pre-tRNA^{Phe}; the identities of the minor bands below the precursor have not been determined. Transcript yield is enhanced by RNasin approximately five- to eightfold (by liquid scintillation counting of RNA-containing gel slices). The numbers on the left indicate distances in centimeters from the origin. (B) The cloverleaf structure of the synthetic pre-tRNA^{Phe}. The anticodon (GAA) is indicated. The splice sites are indicated by arrowheads. *S. cerevisiae* has two types of pre-tRNA^{Phe} molecules, differing only in the length of the IVS. The other type has an 18-base IVS; in that variant, the fourth and fifth bases (both A's in the present variant) are replaced by a single U residue (27).

quence. Table 2 shows the sequence of pre-tRNA^{Phe} as well as an enumeration of the labeled oligonucleotide fragments generated from it by digestion with either RNase A or T1, as well as those of its splicing products. Recently, end-labeled pre-tRNA^{Phe} has been directly sequenced using the gel method (28) (K. Tanner, personal communication). These results show that the pre-tRNA^{Phe} gene is being transcribed faithfully.

To show that this synthetic pre-tRNA possesses mature termini, its 5′ and 3′ ends were mapped as follows. The 5′ nucleotide of pre-tRNA^{Phe} is a G and the next is a C. To prove that transcription actually starts at this G, transcription was done using [α-^{32}P]CTP as label, and the transcript was digested with RNase A. The oligonucleotide presumed to be pppGCp (Fig. 3B) was isolated and redigested with RNase T1 to produce labeled pppGp. Using two-dimensional base analysis (23) and secondary oligonucleotide mapping (22), this nucleotide was shown to comigrate with authentic pppGp (data not shown).

To show that the pre-tRNA terminates correctly with CCA_{OH}, transcription was done using either [α-^{32}P]CTP or [α-^{32}P]ATP as label. The transcript was digested with RNase T1 and an oligonucleotide with mobility expected for $CACCA_{OH}$ (Figs. 3C and 3D) was isolated, redigested with RNase A, and subjected to secondary thin-layer chromatography (22). Upon autoradiography, the spots corresponding to ACp (Fig. 4I) and to Cp (Fig. 4II) were visible with C and A labels, respectively, consistent with $CACCA_{OH}$, but not directly establishing the 3′ end. To determine the precise 3′ terminus, unlabeled transcript was prepared, and the purified RNA product was labeled at the 3′


RNase A
ATP
RNase T1
ATP
RNase A
CTP
RNase T1
CTP
XC
C off
AC off
pppGp off
2d
1d
A
B
C
D

end with [5′-^{32}P]cytidine-3,5′-biphosphate by the action of T4 RNA ligase. The products were then separated by gel electrophoresis, and the major band (corresponding to pre-tRNA^{Phe}-pCp) was isolated, digested simultaneously with RNases A, T1, and T2, and subjected to two-dimensional base analysis. By nearest-neighbor transfer, only the 3′ terminal residue in the initial transcript should be visible upon autoradiography. However, all four nucleotides were labeled, indicating heterogeneity at the 3′ end (data not shown). This pre-tRNA population cannot be resolved using the usual gel electrophoresis, because only a single band is apparent. Thus, to gain greater resolution, use was made of the fact that shorter fragments are resolved much better than longer ones, so the same end-labeled transcript was digested with RNase T1 and the products were separated on a thin, 20% polyacrylamide gel. Upon autoradiography, two equally intense major bands (A and B) were observed, differing by one nucleotide (B longer than A), plus a third, very faint band (C), apparently one nucleotide longer than B (Fig. 4III). Each band was isolated and subjected to two-dimensional base analysis. In band A, only Ap was visible (Fig. 4IV), implying that this transcript corresponds to the correctly terminated 3′ end ($CACCA_{OH}$). In Band B, mostly Gp was labeled (Fig. 4V) suggesting that in this RNA population, a G residue is preferentially added by T7 RNA polymerase after the DNA-encoded CCA terminus ($CACCA-G_{OH}$). In Band C, all four spots are visible (Fig. 4 VI), suggesting that this population of RNA ends with -CCAGN, where N is a random nucleotide added by T7 RNA polymerase.

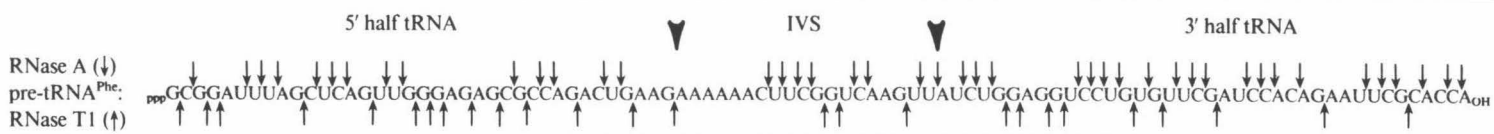
To determine whether this synthetic precursor is a substrate for the yeast tRNA splicing system, transcription was done using [α-^{32}P]UTP as label, and the precursor band was eluted and subjected to *in vitro* splicing by incubation with partially purified yeast tRNA splicing endonuclease and ligase. The products of splicing under a variety of conditions are shown in Fig. 5. Apparently, the pre-tRNA is accurately cleaved by the endonuclease to produce two tRNA^{Phe} halves (5′ half-tRNA is 37 bases; 3′ half-tRNA is 39 bases), a presumptive 2/3 pre-tRNA, and the IVS (19 bases). Upon inclusion of ligase and ATP, the two halves are joined to form mature tRNA^{Phe} (76 bases). Since the transcript is devoid of base modifications, this result implies that base modifications are not required for correct folding of the pre-tRNA or for recognition by the splicing endonuclese and ligase. Reannealing of the precursor by heating to 80°C followed by slow cooling to 30°C did not alter its substrate properties.

To show that cleavage by endonuclease and joining by ligase are both accurate, [α-^{32}P]ATP-labeled pre-tRNA^{Phe} was subjected to the *in vitro* splicing reaction, and bands presumed to correspond to the mature tRNA, 5′ and 3′ half-tRNAs, IVS, and 2/3 pre-tRNA were isolated and fingerprinted using RNase T1. The results are shown in Figs. 6A, B, C, D, and E, respectively. As can be seen from Table 2, the fragments visible in the 5′ half, 3′ half, and IVS fingerprints cor-

FIG. 3. Two-dimensional oligonucleotide mapping of RNase A (A and B) or T1 (C and D) digestion products of pre-tRNA^{Phe} labeled either with [α-^{32}P]ATP (A and C) or with [α-^{32}P]CTP (B and D). First dimension (1d) is electrophoresis at pH 3.5 in cellulose acetate strips; second dimension (2d) is thin-layer homochromatography on polyethyleneimine plates. Each labeled fragment is identified (see also Table 2). RNases A (cleaves after Us and Cs) and T1 (cleaves after Gs) both cleave internally to leave a 5′-hydroxyl and a 3′-monophosphate to generate fragments of the form $5'_{HO}AB \ldots Np\ 3'$ (N = U or C for RNase A; G for RNase T1), whose migration behavior in this system is known (15). (A, B, and D) Spots labeled "off" indicate that the corresponding labeled oligonucleotides (indicated) missed being transferred to the second dimension. (C) The labeled fragment $CACCA_{OH}$ has only been partially transferred to the second dimension.

TABLE 2

SEQUENCE OF PRE-tRNA^{Phe} AND LABELED OLIGONUCLEOTIDE FRAGMENTS ARISING FROM RNase A OR T1 DIGESTION



RNA species	Precursor tRNA^{Phe}				Mature tRNA^{Phe}	5′ half	3′ half	IVS
Ribonuclease	A	A	T1	T1	T1	T1	T1	T1
α-^{32}P-label	ATP	CTP	ATP	CTP	ATP	ATP	ATP	ATP
	U (×2)	U (×4)	G (×3)	$_{ppp}$G	G (×3)	G (×2)	G	UU P
	C (×5)	C (×2)	AG	AG	AG	AG	UUCG	UCAAG
	AC	GC (×2)	AAG	CG	CCAG	AAG P	CACCA$_{OH}$	AAAAAACUUCG
	GC	$_{ppp}$GC	CCAG	CCAG	UUCG	CCAG	AAUUCG	
	GAU	AC (×2)	UUCG	ACUG	ACUG	ACUG	AUCCACAG	
	GGAU	AU	ACUG	UUCG	CUCAG	CUCAG		
	AGAC	GAU	CUCAG	CUCAG	CACCA$_{OH}$	AUUUAG		
	AAGU	AGC	UCAAG	UCAAG	AUUUAG			
	AGAAU	GGU	CACCA$_{OH}$	UCCUG	AAUUCG			
	GGAGGU	AGAC	AUUUAG	CACCA$_{OH}$	AUCCACAG			
	GGGAGAGC	GGAGGU	AAUUCG	AUUUAG	AAG ^{P}AUCUG			
	GAAGAAAAAAC	GGGAGAGC	UUAUCUG	AAUUCG				
		GAAGAAAAAC	AUCCACAG	UUAUCUG				
			AAAAAACUUCG	AUCCACAG				
				AAAAAACUUCG				

Note. All oligonucleotides begin with a 5′-hydroxyl and terminate with a 3′-monophosphate, unless otherwise indicated. Unlabeled fragments are not indicated in the table. Splice sites are indicated by arrowheads. RNase A (arrows pointing down) and T_1 (arrows pointing up) cleavages are indicated in the sequence, as well as exon 1, IVS, and exon 2 regions of the pre-tRNA. If an oligonucleotide is generated n times, this is indicated as ($\times n$). The last oligonucleotide in the mature tRNA column contains an RNase T1-resistant G residue (see Text). The 5′ half-tRNA and IVS (and 3′ half-tRNA) used in the fingerprint assays (Fig. 6) were prepared by reacting the pre-tRNA with endonuclease alone. Hence their terminal residues most probably still possess a 2′,3′-cyclic phosphate, indicated as P, which is characteristic of endonuclease action at a 3′ terminus (ligase has a 2′,3′-cyclic phosphate opening activity) (12).

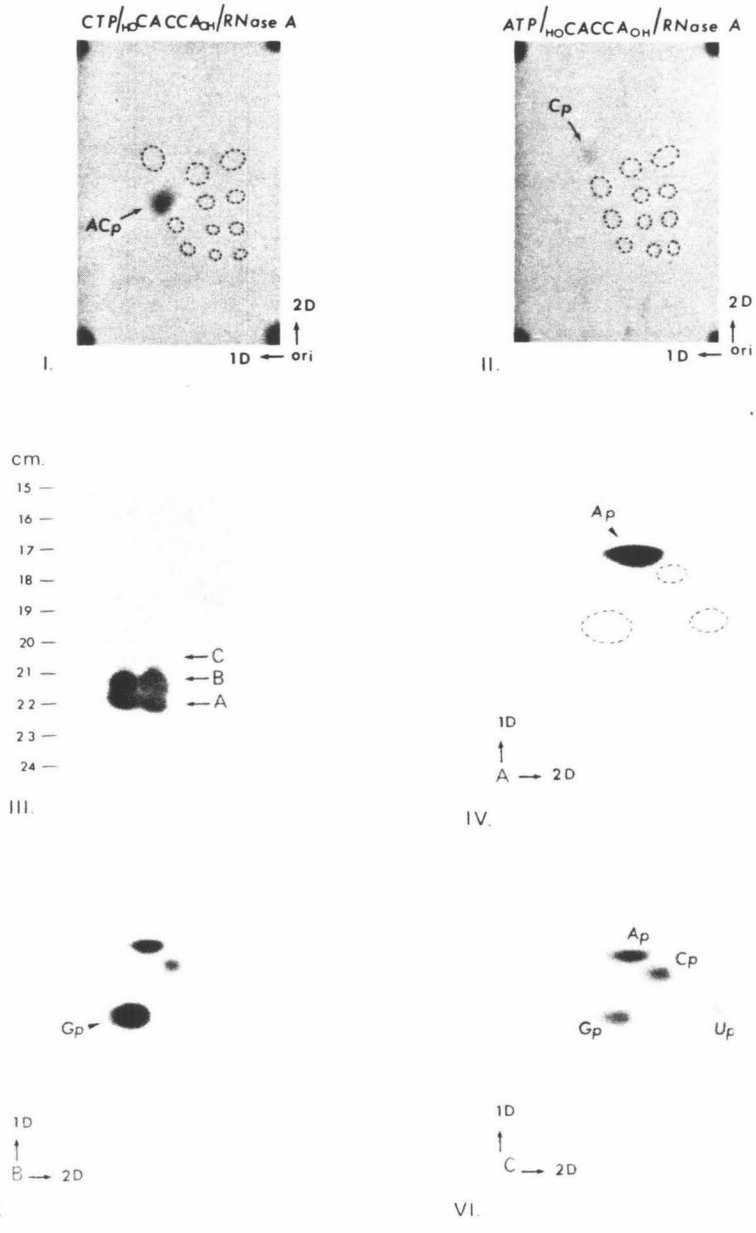


FIG. 4. Mapping of the 3′-terminus of the $tRNA^{Phe}$ precursor. (I and II) Two-dimensional TLC of RNase A digestion products of the 3′-terminal RNase T1-oligonucleotide, $CACCA_{OH}$, labeled either with [α-^{32}P]CTP or -ATP, respectively. (III) The three populations of 3′-terminal fragments, A, B, and C, on a thin polyacrylamide gel; the numbers on the left refer to distance from the origin in centimeters. (IV, V, and VI) two-dimensional base analyses of the three populations, A, B and C, respectively. The origins (point of sample application) are indicated.

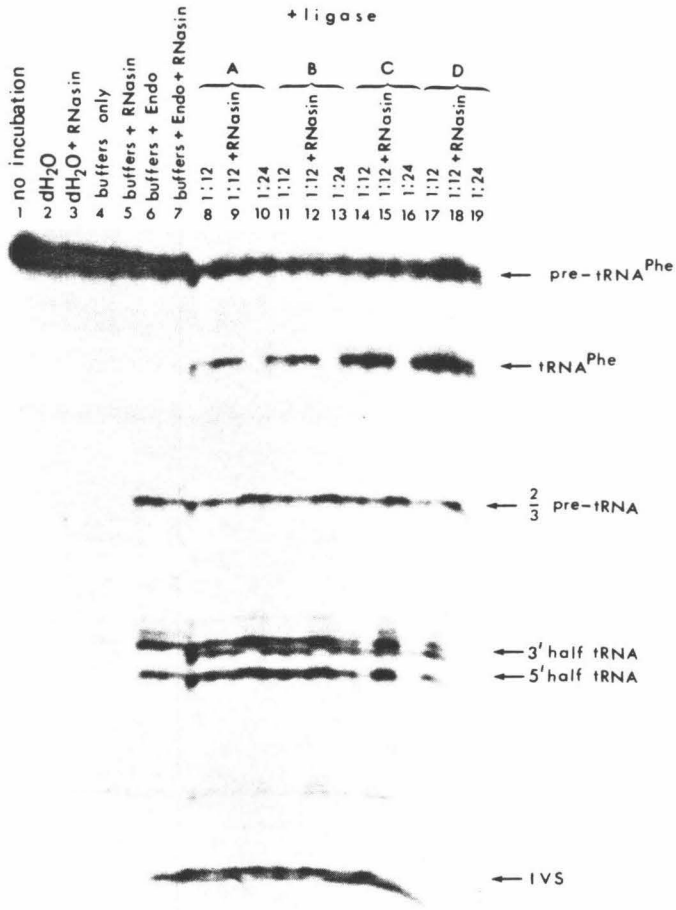


FIG. 5. Splicing assay of the synthetic pre-tRNA^{Phe} under a variety of conditions. All lanes (1–19) contain the same amount of pre-tRNA. The initial concentration of pre-tRNA in each 10-μl splicing reaction is approximately 7–10 nM and the concentration of endonuclease is estimated to be 10^{-11} to 10^{-10} M. For each reaction, the precursor has been incubated for 20 min at 30°C with the reagent indicated above each lane. The tubes in lanes 8–10, 11–13, 14–16, and 17–19 each contain a different ligase preparation (ligase A, B, C, and D), each in increasing dilution (braces); the endonuclease preparation and dilution (1:25) are the same in all lanes. This gel reveals that the activity of the ligase preparations increase in the order A < B < C < D; also, RNasin does not interfere with the splicing reaction. Overexposure of this same gel (data not shown) shows that RNasin decreases nonspecific nuclease degradation during the splicing reaction (lanes 3, 5, 7, 9, 12, 15, and 18).

respond exactly to those predicted from the gene sequence if the endonuclease cleaved exactly at the predicted splice sites. In Fig. 6E, the fingerprint of the presumptive 2/3 pre-tRNA molecule reveals that it is composed of 5′ half-specific (Fig. 6B) and IVS-specific (Fig. 6D) fragments, indicating that the endonuclease cuts sequentially and its first cleavage is at the 3′ splice site. Precursors containing a single cut at the 3′ splice site are

not dead-end products since the 2/3 molecule is observed to disappear upon longer incubation of inclusion of more enzyme. The 5′ half-tRNA ends with a G; hence, if the cleavage and ligation reactions both proceeded accurately, then the splice junction in the mature tRNA should be preceded immediately by this G. In Fig. 6A the new spot (arrow) is the putative splice junction oligonucleotide, AAGAUCUGp. This oligonucleotide contains an internal G residue which is RNase T1-resistant by virtue of the 2′-phosphate at the splice junction resulting from ligase action (12). To ascertain the nature of this oligonucleotide, it was isolated, digested with nuclease P1, and subjected to two-dimensional base analysis. Upon autoradiography, two spots were visible as expected. The first comigrated with pA_{OH}, but the second, although comigrating with pG_{OH} in the first dimension, migrated slower in the second dimension (data not shown). This is the mobility expected for the nuclease P1-resistant dinucleotide, $pG^{p}_{p}A_{OH}$. We therefore conclude that this synthetic precursor tRNA is both cleaved and joined accurately by the splicing endonuclease and ligase, respectively.

Determination of apparent kinetic parameters of the yeast splicing endonuclease. Using this synthetic pre-tRNA^{Phe}, the kinetics of cleavage by the endonuclease was studied. Cleavage reaction rates were measured at various concentrations from 3 to 110 nM pre-tRNA^{Phe}, of known and uniform specific radioactivity ($8–9 \times 10^4$ dpm/pmol). At 30°C, cleavage is linear with time for at least 20 min and dependent on substrate concentration in this range. A typical experiment is shown in Fig. 7A.

The velocity of the cleavage reaction can be represented by either the rate of intron or the combined exon appearance. Thus for each substrate concentration, the rate of endonuclease cleavage was measured by plotting the amount of intron or combined exons appearing as a function of time. The reaction velocity at a particular substrate concentration was taken to be the slope of the regression line through four points (0, 3, 6, and 9 min) in a time course. When reaction rates, either as combined exon appearance or as intron appearance, are plotted versus substrate concentrations, a typical hyperbolic curve is seen (Fig. 7B). Lineweaver–Burk, Eadie–Hofstee, and Hanes–Woolf analyses of these data give comparable results: The apparent K_M is 25–30 nM, while the apparent V_{max} is 10^{-3}–10^{-2} pmol/min.

DISCUSSION

The pre-tRNA expressed from the T7 promoter differs from precursor synthesized *in vivo* or by yeast nuclear extract (YNE) in two respects. First, it is totally devoid of base modifications. The fact that it serves as an excellent substrate for the yeast tRNA splicing enzymes suggests that base modifications are not necessary for recognition by these enzymes. (Because the exact concentration of *in vivo*-labeled tRNA precursors cannot be easily determined, we have not compared the substrate properties of *in vivo* precursors to those of the synthetic precursor described here.) Second, the *in vitro* transcript starts with a 5′-triphosphate, instead of the 5′-monophosphate found in the 5′ end-matured precursor. This problem has recently been circumvented by including a three-fold excess of 5′-GMP over GTP in the transcription reaction, resulting in the production of transcripts almost entirely starting with a 5′-monophosphate (J. Sampson and O. Uhlenbeck, personal communication; our unpublished results).

The *Bst*NI-restricted template ends with a single 5′ T-overhang in the coding strand (see Fig. 1). The transcript is heterogeneous at the 3′ end and consists of two major subpopulations, one of which possesses the correct, mature 3′ end CCA_{OH} terminus, and the other an extra G residue, i.e., $-CCAG_{OH}$. A third, minor subpopulation possesses an ad-

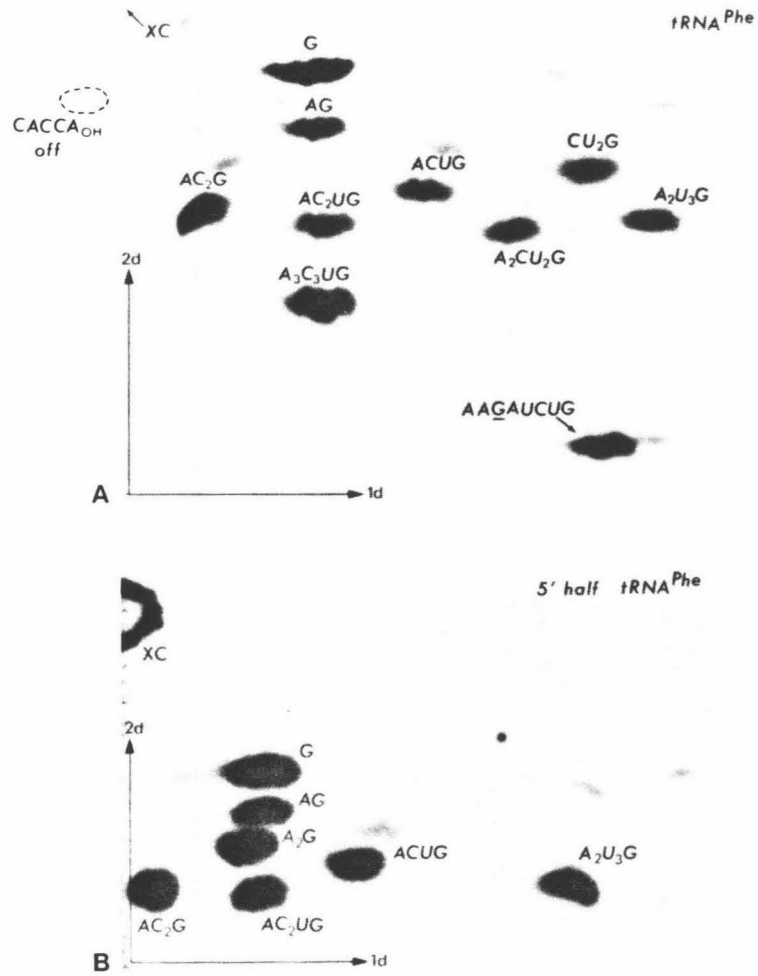


FIG. 6. Two-dimensional oligonucleotide mapping of RNase T1 digestion products of mature $tRNA^{Phe}$ (A), 5′ half-$tRNA^{Phe}$ (B), 3′ half-$tRNA^{Phe}$ (C), IVS (D), and $\frac{2}{3}$ pre-$tRNA^{Phe}$ (E) generated from the reaction products of pre-$tRNA^{Phe}$ labeled with [α-^{32}P]ATP. Each labeled fragment is identified (see also Table 2). (A) The spot indicated by an arrow is the splice junction oligonucleotide, AAGAUCUGp, identified by virtue of its anomalous mobility (see text). The extra spots appearing near the upper left-hand corner of the IVS fingerprint (D) have not been further analyzed, but are presumed to be degradation products of the poly(A)-containing, IVS-specific oligonucleotide, $A_6C_2U_2G$ (see Table 2). These spots are consistently observed in the same relative position whenever this oligonucleotide is present (i.e., pre-tRNA, $\frac{2}{3}$-pre-tRNA, and IVS fingerprints).

ditional random residue after the G, i.e., -$CCAGN_{OH}$. This incorporation of extra nucleotides at the 3′ end after the template overhang-encoded nucleotide is also observed by J. Sampson and O. Uhlenbeck (personal communication) in the purification of $tRNA^{Phe}$ synthesized by T7 RNA polymerase. Apparently, heterogeneity with

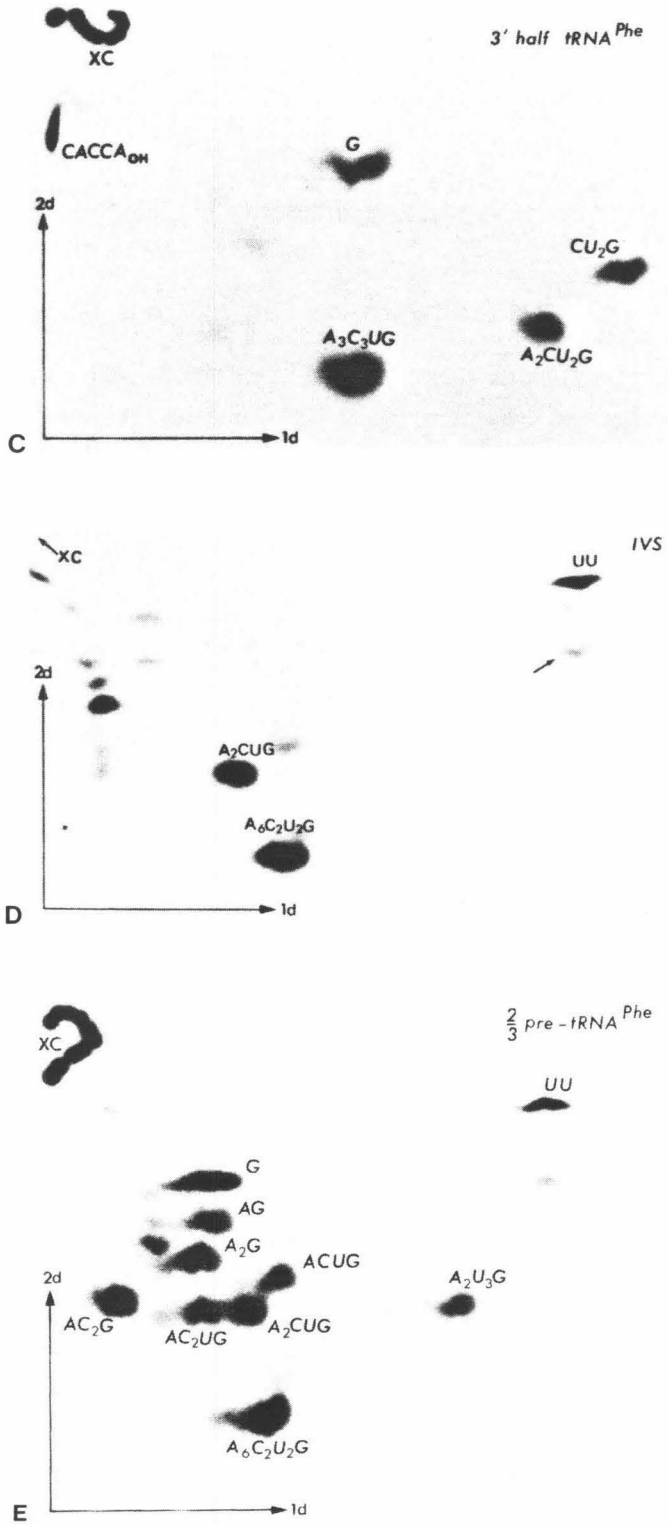
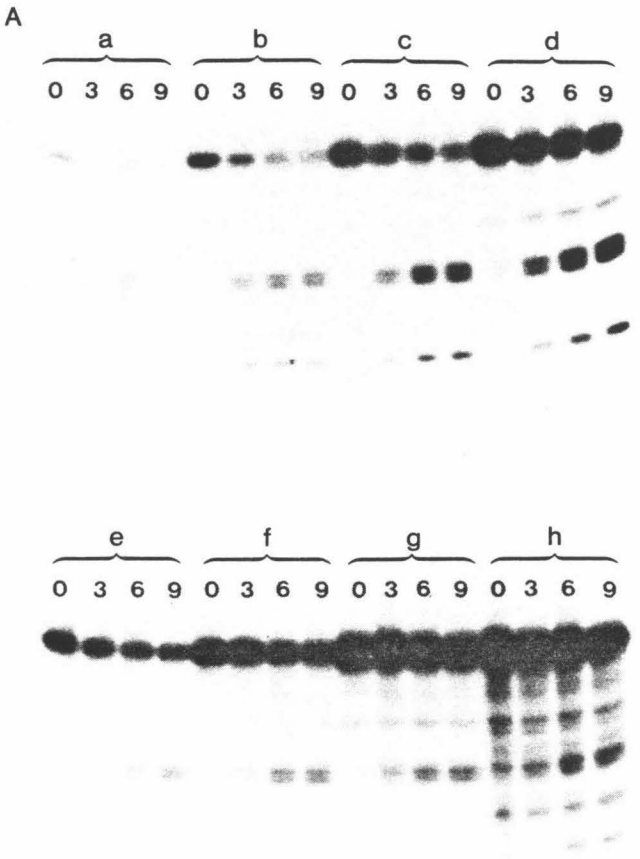
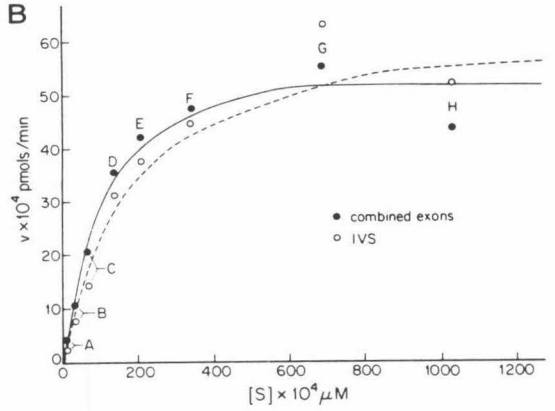


FIG. 6—*Continued.*


A
a
b
c
d
0 3 6 9 0 3 6 9 0 3 6 9 0 3 6 9
e
f
g
h
0 3 6 9 0 3 6 9 0 3 6 9 0 3 6 9


B
v x 10⁴ pmols/min
[S] x 10⁴ μM
combined exons
IVS
A
B
C
D
E
F
G
H
0
10
20
30
40
50
60
200
400
600
800
1000
1200

respect to the 3′ terminus is not a serious problem, since the +1 and the +2 subpopulations are cleaved accurately by the endonuclease, and the corresponding +1 and +2 3′ halves are joined to the 5′ half-tRNA by the ligase. We observe this heterogeneity as a triplet band corresponding to the 3′ half-molecule (c.f. Fig. 5). All of these molecules are converted to product if ligase is added.

Previously, pre-tRNA splicing substrates were either isolated from *in vivo* ^{32}P-labeled RNA synthesized at the nonpermissive temperature in an *rna*1 mutant yeast strain (*ts*136) (29) or prepared using a crude YNE containing RNA polymerase III, necessary transcription factors, end maturation, and base modification enzymes (3), under conditions not optimal for splicing. The *in vivo* method precludes design and construction of site-specific mutations of the pre-tRNA gene, is laborious, and involves working with high levels of radioactivity (50 mCi per preparation). In the YNE method, on the other hand, transcription and end-processing take place simultaneously in the test tube. In this system, mutations which affect either transcription or end maturation cannot be tested for their effects on splicing. Using the T7 polymerase method, however, the effects of any mutation in the precursor upon IVS removal can be studied, and we are now embarked on an extensive investigation of the structural requirements for splicing substrates. In addition, the present technology allows the synthesis of precursors with precisely determined specific activities.

Finally, this method allows large quantities of the pre-tRNA to be synthesized for direct structural studies. Recently, 15 mg of the pure precursor tRNA^{Phe} has been made in a large-scale transcription of pUC13Phe DNA by C. Hall of Brandeis University. NMR studies on this RNA have been initiated (C. Hall, personal communication). The stage is now set for a detailed study of the structural basis for substrate recognition by the yeast tRNA-splicing endonuclease and ligase.

ACKNOWLEDGMENTS

We thank Marjorie Fidler for her assistance during the early phases of this work; Phillip Green and Eric Phizicky for providing the partially purified endonuclease and ligase, respectively; Gloria McFarland and Calvin Ho for helpful discussions; Ren-Jang Lin for technical advice; J. Sampson and Professor O. Uhlenbeck for the synthetic oligonucleotides and for communicating results prior to publication; and Cathy Elkins for preparing the manuscript. This work was supported by grants from the American Cancer Society and the NIH.

REFERENCES

1. Guthrie, C., and Abelson, J. (1982) *in* The Molecular Biology of the Yeast *Saccharomyces cerevisiae:* Metabolism and Gene Expression (Strathern, J. N., Jones, E. W., and Broach, J. R., Eds.), pp. 487–528, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
2. Raymond, K. C., Raymond, G. J., and Johnson, J. D. (1985) *EMBO J.* **4**, 2249–2256.

FIG. 7. Typical kinetics experiment. (A) Each 10-μl aliquot (described under Materials and Methods) is loaded onto a single lane. Hence, each of the eight different substrate concentrations (a–h) gives rise to four different lanes corresponding to the four different time points (0, 3, 6, and 9 min). The combined exon bands and the intron band in each lane were sliced out and the radioactivity in each sample was determined. This allowed two determinations of the velocity of the reaction for each substrate concentration. The identity of the band that appears between the 5′ half-tRNA and the IVS at very high substrate concentrations (h) has not been determined. The first gel (a–d) was exposed with intensifying screen; the second (e–h), without intensifying screen. (B) A hyperbolic Michaelis–Menten Plot of velocity v versus substrate concentration [S]. Solid and open circles refer to reaction rates in terms of combined exon or intron appearance rate, respectively. The substrate concentraitons ([S]) versus reaction rate (v) values were analyzed using Lineweaver–Burk ($1/v = K_M/V_{max}(1/[S]) + 1/V_{max}$), Eadie–Hofstee ($v = -K_M(v/[S]) + V_{max}$), and Hanes–Woolf ($[S]/v = [S]/V_{max} + K_M/V_{max}$) plots.

3. Engelke, D. R., Gegenheimer, P., and Abelson, J. (1985) *J. Biol. Chem.* **260,** 1271–1279.
4. Allison, D. S., and Hall, B. D. (1985) *EMBO J.* **4,** 2657–2664.
5. Sharp, S. J., Schaack, J., Cooley, L., Burke, D. J., and Söll, D. (1986) *CRC Crit. Rev. Biochem.* **19,** 107–144.
6. Melton, D. A., DeRobertis, E. M., and Cortese, R. (1980) *Nature (London)* **284,** 143–148.
7. Hopper, A. K. (1984) *in* Processing of RNA (Apirion, Ed.), pp. 91–117, CRC Press, Boca Raton, FL.
8. Pearson, D., Willis, I., Hottinger, H., Bell, J., Kumar, A., Leupold, U., and Söll, D. (1985) *Mol. Cell Biol.* **5,** 808–815.
9. Willis, I., Hottinger, H., Pearson, D., Chisolm, V., Leupold, U., and Söll, D. (1984) *EMBO J.* **3,** 1573–1580.
10. Nishikura, K., Kurjan, J., Hall, B. D., and DeRobertis, E. M. (1982) *EMBO J.* **1,** 263–268.
11. Peebles, C. L., Gegenheimer, P., and Abelson, J. (1983) *Cell* **32,** 525–536.
12. Greer, C. L., Peebles, C. L., Gegenheimer, P., and Abelson, J. (1983) *Cell* **32,** 537–546.
13. Hofstetter, H., Kressmann, A., and Birnstiel, M. L. (1981) *Cell* **24,** 573–585.
14. Newman, A. J., Ogden, R. C., and Abelson, J. (1983) *Cell* **35,** 117–125.
15. Mattoccia, E., Baldi, M. I., Pande, G., Ogden, R., and Tocchini-Valentini, G. P. (1983) *Cell* **32,** 67–76.
16. Tabor, S., and Richardson, C. C. (1985) *Proc. Natl. Acad. Sci. USA* **82,** 1074–1078.
17. Davanloo, P., Rosenberg, A. H., Dunn, J. J., and Studier, F. W. (1984) *Proc. Natl. Acad. Sci. USA* **81,** 2035–2039.
18. Ladner, J. E., Jack, A., Robertis, J. D., Brown, R. S., Rhodes, D., Clark, B. F. C., and Klug, A. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4414–4418.
19. Quigley, G. J., Wang, A. H. J., Seeman, N. C., Suddath, F. L., Rich, A., Sussman, J. L., and Kim, S. H. (1975) *Proc. Natl. Acad. Sci. USA* **72,** 4866–4870.
20. Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982) *in* Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
21. Volckaert, G., Min Jou, W., and Fiers, W. (1976) *Anal. Biochem.* **72,** 433–446.
22. Volckaert, G., and Fiers, W. (1977) *Anal. Biochem.* **83,** 228–239.
23. Saneyoshi, M., Ohashi, Z., Harada, F., and Nishimura, S. (1972) *Biochim. Biophys. Acta* **262,** 1–10.
24. Dunn, J. J., and Studier, W. (1983) *J. Mol. Biol.* **166,** 477–535.
25. Zoller, M. J. and Smith, M. (1982) *Nucleic Acids Res.* **10,** 6487–6500.
26. Sanger, F., Nicklen, S., and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
27. Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R., and Rutter, W. J. (1978) *Proc. Natl. Acad. Sci. USA* **75,** 190–194.
28. Donis-Keller, H. (1980) *Nucleic Acids Res.* **8,** 3133–3142.
29. Hopper, A. K., Banks, F., and Evangelidis, V. (1978) *Cell* **14,** 211–219.

Chapter IV

SUBSTRATE RECOGNITION AND SPLICE SITE SELECTION BY THE *S. CEREVISIAE* tRNA SPLICING ENDONUCLEASE

[A condensed version of the text of this chapter has been submitted for publication in **Cell.**]

Summary

We have fused a S. cerevisiae pre-tRNA^{Phe} gene to a bacteriophage T7 promoter and shown that the pre-tRNA transcribed from this heterologous construct is an efficient substrate for our in vitro splicing system. This method has allowed the precise alteration of any portion of the pre-tRNA for the extensive and systematic study of tRNA splicing substrate specificity. A total of 15 carefully designed pre-tRNA^{Phe} variants were constructed and subjected to in vitro splicing analyses. Our results suggest that: (i) residues U8 and C56 are recognized by the endonuclease, and are probably contact points during the splicing event; (ii) the endonuclease selects the 5' and 3' splice sites in the pre-tRNA by a simple distance measurement mechanism; and (iii) the highly conserved purine residue 3' proximal to the anticodon may be important for cleavage at the 5' splice site. The possible borders of the endonuclease-bound region in the pre-tRNA are also discussed.

Introduction

tRNA splicing in S. cerevisiae is mediated by two distinct enzymes: an endonuclease to locate and cleave the 5' and 3' splice sites, producing a free IVS and two annealed tRNA halves; and a ligase, to join the two tRNA halves at the nicked junction, producing an intact tRNA (Peebles et al., 1979; Peebles et al., 1983; Greer et al., 1983). The ligase has been purified to homogeneity and the gene cloned (Phizicky et al., 1986) and sequenced (Westaway et al., 1987) and through electron microscopic analyses, its cellular location has been determined to be in the nuclear periphery (Clark and Abelson, 1987). The endonuclease, being an integral membrane protein, has eluded homogeneous purification to date, but has been purified to a very high degree (P. Green, personal communication). The basic molecular mechanism of the tRNA splicing reaction, as well as the nature of the intermediates involved, has been worked out to fine detail: the endonuclease cleaves at the 5' and 3' splice sites of the pre-tRNA to generate 3' and 5' termini possessing 2'3'-cyclic phosphate and 5' OH, respectively; and ligase, possessing at least three distinct activities, joins the 3' terminus of the 5' half tRNA to the 5' terminus of the 3' half tRNA in a series of cyclic phosphatase, phosphorylation and adenylylation reactions, with the participation of two molecules of ATP (Knapp et al., 1979; Greer et al., 1983).

Despite the fact that tRNA splicing is more fully characterized than either class I (nuclear-coded mRNA) or class II (rRNA and some fungal mitochondrial RNA) RNA splicing, the problems of substrate recognition and its interaction with the tRNA splicing enzymes have not been thoroughly addressed. In class I splicing, conserved sequences at the splice junctions and elsewhere in the mRNA molecule are known to interact with certain components of the splicing machinery and mutations at these sites impair the reaction (Mount, 1982; Mount et al., 1983;

Langford and Gallwitz, 1983); in contrast, no such consensus splice junction sequences exist in pre-tRNAs (Greer and Abelson, 1984; Ogden, et al., 1981; Ogden et al., 1984). tRNA IVSs are extremely heterogeneous in size (ranging from 13 to 60 bases) and sequence; the only constant feature they possess is that they are always located one base 3' to the anticodon, and they are A + U-rich (A + U to G + C ratio of approx. 2:1) (Ogden et al., 1980). Further, pre-tRNA splice junction mutations do not affect the accuracy of IVS excision, or processing in general (Colby et al., 1981). In class II splicing, the IVS takes an active role during the splicing event - its several blocks of conserved primary sequence allow it to assume a unique secondary and tertiary conformation for proper alignment of the splice sites prior to the self-catalyzed reaction (Kruger et al., 1982; Davies et al., 1982; Michel and Dujon, 1983); in contrast, tRNA IVSs are generally thought to play a passive role in their own removal. This is evidenced by the rarity of intron mutations that affect splicing (Nishikura et al., 1982), and the fact that no biological function has yet been ascribed to tRNA introns, except as probable recognition elements for a couple of anticodon base modification enzymes (Johnson and Abelson, 1983; Strobel and Abelson, 1986).

To probe the enzyme-substrate interaction during the tRNA splicing event, there has been a need for a system that allows one to construct precise alterations anywhere in the pre-tRNA and analyze the effects of such alterations on splicing. Our in vitro splicing system consisting of highly purified endonuclease and ligase enzymes provides a convenient and accurate way of analyzing splicing phenotypes, but an ideal pre-tRNA synthetic system has generally been lacking. The main deterrent is the inherent complexity of the tRNA biosynthetic pathway itself. tRNA gene promoters reside in the coding sequence as two noncontiguous 10-base pair blocks (Galli et al., 1981; Sharp et al., 1981) and in all cases, transcription

initiates and terminates in the 5' and 3' flanking regions, respectively (Raymond et al., 1985; Allison and Hall, 1985). As a result, the primary transcript from these genes contain 5' and 3' extensions which have to be enzymatically deleted and the trinucleotide CCA added to the trimmed 3' end before the IVS can be excised; meanwhile, certain base modifications take place during each of these steps (Melton et al., 1980; Nishikura et al., 1982). The enzymes which carry out these early processing steps recognize elements of conserved sequence and structure in the pre-tRNA, and it is generally believed that the splicing enzymes interact with these elements as well (Traboni et al., 1984; Greer et al., 1987). Thus, mutations which affect any of these earlier steps cannot be tested for their effects on splicing (Newman et al., 1983; Mattoccia et al., 1983).

We have recently overcome the above problem. A S. cerevisiae pre-tRNA^{Phe} gene linked to a bacteriophage T7 promoter was synthesized de novo and cloned into an M13 vector (Reyes and Abelson, 1987). We have shown that the synthetic pre-tRNA transcribed from this system is an efficient substrate for our in vitro splicing system (Ibid.). Yeast tRNA^{Phe} was chosen for this purpose because more is known about its structure than that of any other tRNA molecule (Quigley and Rich, 1976; Sussman and Kim, 1976; Rich and RajBhandary, 1976). We have taken advantage of the power and versatility of this method by synthesizing a total of 15 carefully designed pre-tRNA^{Phe} variants to probe the mode of interaction of this molecule with the splicing machinery.

In this report, we show evidence that the endonuclease recognizes at least two of the totally conserved bases in tRNAs, U8 and C56, as well as the highly conserved purine residue immediately preceding the 5' splice junction. From our results also emerges a model for splice site selection by the endonuclease: the enzyme selects the 5' and 3' cleavage sites in the pre-tRNA by measurement of

fixed distances from some reference point in the molecule. Finally, we have attempted to delineate this endonuclease-bound region of the pre-tRNA molecule during the splicing event.

Results and Discussion

Bacteriophage T7 Promoter/Pre-tRNA^{Phe} Gene System Allows Synthesis of Any Predesigned Mutant Pre-tRNA

Our system consists of the S. cerevisiae pre-tRNA^{Phe} gene containing a 19-base IVS, linked to a bacteriophage T7 promoter (Figure 1A); the system was constructed by self-assembly of a set of complementary synthetic deoxyoligonucleotides and cloned in M13 mp10 (see Tables IA and IC). It was designed such that the start of transcription by T7 RNA polymerase is at the 5' end of the gene, and that a BstNI site coincides with its 3' end. To synthesize pre-tRNA, the plasmid is cleaved with BstNI and then transcribed with T7 RNA polymerase. We have shown that the pre-tRNA transcribed from this system has correct sequence and mature 5' and 3' termini and is cleaved and ligated accurately by the endonuclease and ligase, respectively (Reyes and Abelson, 1987). This pre-tRNA differs from natural pre-tRNAs in that it has no base modifications and it starts with a 5' triphosphate, being a primary transcript from the bacteriophage T7 promoter. Natural pre-tRNAs start with a 5' monophosphate due to 5'-end maturation by RNAse P (Kline et al., 1981; Engelke et al., 1985). Previous studies have shown that base modifications have no defined role in the stabilization of the tertiary structure of tRNAs (Kim, 1979), but that the presence of a 5' triphosphate may mimic a 5' leader sequence and lower splicing efficiency due to steric hindrance (Greer et al., 1987). This pre-tRNA^{Phe} starts with a G. Thus, to resolve the above problem,

transcription was done in the presence of a 3-fold excess of 5' GMP over GTP (plus the other 3 NTPs). 5' GMP can take part only in the initiation of transcription, and since it considerably exceeds GTP, the majority of the transcripts should start with 5' monophosphate. The second base in this pre-tRNA is a C; thus, to detect the identity of the 5' nucleotide, α-^{32}P-CTP was used as transcription label, and the resulting transcript subjected to complete hydrolysis and Nishi base analysis. We found that almost 100% of the transcripts start with 5' monophosphate (Figure 1B). To compare splicing efficiencies, pre-tRNAs transcribed with or without 5' GMP was subjected to a standard splicing assay. We find that cleavage and ligation are both approximately 5-fold more efficient for the transcript that starts with a 5' monophosphate than for the one with a 5' triphosphate (Figure 1C). Hence, all subsequent transcriptions, of wild type or mutant pre-tRNA genes, were done in the presence of 5' GMP.

Fusion of the pre-tRNA^{Phe} gene with the heterologous T7 promoter has allowed us to design any change anywhere in the gene and synthesize the corresponding mutant pre-tRNA. In this study, the altered pre-tRNAs were constructed either by oligonucleotide-directed mutagenesis (Zoller and Smith, 1982) or total gene synthesis (see Tables IA, B and C). In designing these changes, we let ourselves be guided by the well-known three-dimensional structure of tRNA^{Phe} (Quigley and Rich, 1976; Sussman and Kim, 1976). Previous work on chemical and enzymatic structure probing analyses of pre-tRNA^{Phe} (Lee and Knapp, 1985; Swerdlow and Guthrie, 1984), as well as recent data on direct NMR structural analysis of this molecule (C. Hall, personal communication), strongly suggest that its mature domain closely resembles tRNA^{Phe}. Although pre-tRNA^{Phe} has not yet been subjected to x-ray crystallographic analyses and exact atomic coordinates are not yet known, we assume that all nine tertiary base-base interactions in tRNA^{Phe}

also exist in pre-tRNA^{Phe}. In this study, we have made single base substitutions involving the G19-C56 and G15-C48 tertiary base pairs, and the residue U8. These 5 residues map to the internal promoter elements of the gene. Residues U8, G15, and G19 map to the A-block; C56 maps to the B-block; and C48 to the extra-loop promoter element (Willis et al., 1984; Traboni et al., 1984). Thus, synthesis of precursors altered at these sites may not be possible using conventional pol III-derived systems (Engelke et al., 1985; C. Parker and J. Topol, unpublished data); in particular, the C56 and G19 alterations are both well-documented transcription-down mutations (Newman et al., 1983; Reyes et al., 1986). Using our T7 promoter/pre-tRNA^{Phe} gene system, however, these mutant pre-tRNAs are readily synthesized.

We have made single-base as well as double, compensatory mutations involving these residues. Generated were four single-base substitution mutants and two double mutants of pre-tRNA^{Phe}. The single mutants are G15A, C48T, G19C and C56G, where residues G15, C48, G19, and C56 were changed to A, T, C and G, respectively. The double mutants are G15A-C48T and G19C-C56G. The former contains both G15 to A and C48 to T alterations, while the latter contains both G19 to C and C56 to G changes. We also changed the totally conserved U8 to a C. Lastly, a truncated pre-tRNA^{Phe} was constructed by cleaving the plasmid with *Ava*II restriction endonuclease instead of BstNI prior to transcription; this cleavage site maps to the extra arm of the pre-tRNA, thus the result is a molecule lacking the entire TΨC arm and the 3' strand of the acceptor stem. For a summary of mutations that were constructed, refer to Fig. 2C. Each of these pre-tRNA^{Phe} gene mutants were transcribed and each precursor subjected to standard in vitro splicing assay (Reyes and Abelson, 1987). In all splicing assays, pre-tRNA species were incubated in the presence of endonuclease alone to detect cleavage

efficiencies, or with endonuclease plus ligase to detect ligation efficiencies. Further, two control experiments were always done: pre-tRNA is gel electrophoresed without any prior incubation to detect spontaneous breakdown, or after incubation with buffers only (see Experimental Procedures) to detect any degradation due to nonspecific nucleases in the buffers. In all cases, no such untoward cleavages occurred.

Base Substitution Mutants of Pre-tRNA^{Phe} Define Probable Contact Points with the Endonuclease

G19 and C56 are both surface residues of pre-tRNA^{Phe} and form a tertiary base pair situated on the outer corner of the molecule (Fig. 2, panels A and B). Changing G19 to C does not affect cleavage (Figure 3, Panel III, lane 2), but changing C56 to G abolishes it (lane 3). In the double mutant, G19C-C56G, the presence of both changes is expected to restore the tertiary base pairing between these two residues, but this mutant is not cleaved either (lane 4). We do not have any information as to the existence of tertiary base pairing between residues 19 and 56 in the G19C, C56G and G19C-C56G mutant precursors. However, in the wild-type, G19 is engaged in a Watson-Crick base pairing interaction with C56 (Quigley and Rich, 1976; Sussman and Kim, 1976). Thus, it is most reasonable to assume that the bases in these 2 positions can engage in tertiary pairing only if it is of Watson-Crick type. From this argument it follows that tertiary base pairing is allowed in the G19C-C56G double mutant, but not in the G19C nor C56G single mutants. This means that splicing phenotypes correlate with the identity of the base at position 56, and not with the existence of tertiary base pairing between residues 19 and 56. The inability of the compensating mutations to rescue cleavage in the double mutant suggests that there is a tight requirement by the endonuclease for a C

residue at position 56. This result, together with the fact that C56 is totally conserved among (non-initiator) tRNAs and is located on the surface of the molecule, strongly suggests that this residue is a recognition element for the endonuclease and is probably a contact point during the splicing event. In contrast, there seems to be no requirement for a unique residue at position 19, and it seems unlikely that the endonuclease contacts this residue.

It is also formally possible that the cleavage defect in the C56G and G19C-C56G mutants is due to a drastic alteration in tertiary structure brought about by the C56 to G alteration and not due to the change in base identity per se at this position (It also follows from this argument that the G19 to C change has no structural effects). However, the C56G precursor is not an absolute defective, but can be cleaved using a purer endonuclease preparation at higher activities, and at certain [Mg^{++}] optima, although in all cases it is severely defective compared to wild-type pre-$tRNA^{Phe}$ (data not shown). This means that the C56G precursor can still interact with the enzyme (although apparently at lower affinity), and thus the structure is probably not drastically changed. This is consistent with the notion that the pre-$tRNA^{Phe}$ contacts the endonuclease at several points, and in the C56G precursor, only one of these contact points was lost.

The G15-C48 tertiary base pair consists of two semi-conserved residues, both located in the interior of the thoracic portion of the molecule (Fig. 2, Panels A and B). Residue 15 is always a purine, and residue 48, a pyrimidine. This lack of total conservation as well as the interior location of this base pair, makes it unlikely that these two residues are recognition elements by the endonuclease. In fact, it has previously been shown that both G15 and C48 are unreactive toward chemical modification (cited in Rich and RajBhandary, 1976). Changing G15 to A, or C48 to U both diminish splicing by a small but detectable amount (Figure 3,

panel III, lanes 5 and 6). Incorporation of both changes in the same molecule to restore base pairing between the two residues rescues splicing to wild-type levels (lane 7). These results suggest that there are no requirements for unique residues at positions 15 and 48 and G15 and C48 are probably not recognized by the endonuclease. A15/C48 and G15/U48 configurations are both apparently well-tolerated; however, G15/C48 and A15/U48 seem more favorable.

Structural integrity of the thoracic region of the pre-tRNA is largely maintained by four tertiary interactions: three triads (U8-A14-A21, A9-A23-U12, and C13-G22-G46) and the G15-C48 tertiary pair (Klug et al., 1974). The slight diminution of cleavage efficiency of the G15A and C48T single mutants may be attributed to the loss of tertiary base pairing between residues 15 and 48, but this is apparently well-tolerated, probably due to the existence of the three triads. In the double mutant, the change from G15-C48 to A15-U48 tertiary base pairing is expected to result in a lateral displacement of this base pair by approximately 2Å, as well as some alteration in H-bonding (cited in Rich and RajBhandary, 1976). These changes are apparently well-tolerated in pre-tRNA^{Phe}, and do not affect endonuclease action.

U8 is another invariant residue in pre-tRNA^{Phe}, but unlike C56, it is on the inner, not outer, corner of the molecule. Alteration of U8 to C inhibits splicing (Figure 3, panel III, lane 8). An endonuclease binding assay involving competition between radioactively labeled U8C and increasing amounts of unlabeled wild-type precursor indicates that this mutant binds the enzyme weakly compared to wild type (data not shown). U8 also forms a tertiary base pair with A14, and although we have not made any compensating mutations at position 14, we think that the cleavage defect in the U8C mutant is due to the base change per se and not to a drastic structural alteration of the molecule resulting from this base change. First,

the U8C mutant, like the C56G, is not an absolute defective, and can be cleaved using higher activities of the endonuclease at certain [Mg^{++}] optima (data not shown), implying that interaction with the endonuclease is still possible. Second, the same U8 to C change was made in S. cerevisiae pre-$tRNA_3^{Leu}$, and this mutant is also severely defective in cleavage by purified X. laevis tRNA splicing endonuclease. Enzymatic structure probing of this mutant pre-tRNA shows that it is indistinguishable from wild-type (G. Tocchini-Valentini, pers. comm.). These results strongly suggest that U8, like C56, is another recognition element for the emdonuclease, and is probably among the points of contact during the splicing event.

A more direct and precise way to detect any structural alteration brought about by the C56 to G and U8 to C mutations is to subject these two mutant precursors to a similar NMR structural analysis done with wild-type pre-$tRNA^{Phe}$ (C. Hall, personal communication). A result indicating identical or similar tertiary structure to the wild-type precursor would constitute stronger argument for the hypothesis that these two residues are endonuclease contact points during the splicing reaction.

The *Ava*II-truncated precursor is totally resistant to the endonuclease (Fig. 3, Panel III, lane 12). In this variant, several recognition elements of the enzyme may have been lost as a result of the missing TΨC arm and surrounding regions. Thus, this variant may not interact with the endonuclease at all. We have not performed any computer analyses to determine the precise secondary structure of this precursor, and the possibility remains as well that this RNA folds in an entirely novel way.

Table 2 summarizes all the splicing results. It can be observed that there is very high correlation between cleavage and ligation efficiencies for each precursor

studied (compare also corresponding lanes in panels III and IV, Fig. 3); specifically, we did not find any mutation that impairs ligation but not cleavage (except for poly U precursor; see last section).

Endonuclease Cleavage Sites Are a Function of the Length of the Anticodon, but not Acceptor, Stem of the pre-tRNA

All nine S. cerevisiae pre-tRNAs contain anticodon stems of length 5 base pairs, and the 5' and 3' splice sites are always 11 and 6 bases away respectively, from the first (i.e., topmost) base pair of the stem (Ogden et al., 1984). This constancy in distance of the splice sites from the central portion of the molecule has prompted us to investigate the possibility that splice site selection in tRNA splicing is determined by distances from a fixed point in the thorax of the molecule. This seems reasonable since tRNA IVSs are extremely heterogeneous and do not possess any consensus splice junction sequences. This hypothesis was tested by making insertions into various parts of the pre-tRNA; the structures and designations of these mutant pre-tRNAs are shown in Figures 2C and 2E.

The G▽ variant, having a single G insertion in the anticodon stem is not cleaved (Fig. 3, panel III, lane 9). In the GC▽ variant, a C residue was introduced in the other strand to base-pair with this G residue, and cleavage was rescued; however, the major IVS is now two bases longer than wild type IVS (Figure 3, panel III, lane 10). These results suggest that the double-helical structure of the anticodon stem may be important for endonuclease recognition. However, the fact that the IVS increased in length by two bases indicates that additional factors are at play. Fingerprint analysis of the three IVSs produced from the GC▽ mutant indicates that the major IVS (topmost band) is generated by cleavages one base upstream and one base downstream of the normal 5' and 3' splice sites, respectively (compare Panels A and B, Fig. 5). The second IVS is generated by a cut at the

normal 5' splice site and at the new 3' splice site above (compare panels A and C, Fig. 5); and the third IVS, by cuts at the normal 5' and 3' splice sites (compare panels A and D, Figure 5). The upstream shift of the 5' splice site in the longest GC▽ IVS is also clear from the faster migration of the A_6GC fragment relative to A_6C in the first dimension of an RNase A fingerprint (compare Panels E and F, Figure 5). The enzyme has preserved the distance of the two splice sites from the central portion of the pre-tRNA ; the increase in distance resulting from the increase in the length of the anticodon stem due to the base-pair insertion did not affect the exon lengths, but instead was absorbed by the IVS. These results therefore suggest a cleavage mechanism for the endonuclease: it selects the splice sites by measuring fixed distances from its binding site, as opposed to recognition of short sequences at the splice junctions.

Splicing of the GC▽ mutant precursor generates two shorter IVSs in addition to the major IVS, which was generated by cuts at novel 5' and 3' splice sites. The novel 5' splice site in this mutant precursor is in a double-stranded region, while novel 3' splice site is right next to the double-helical anticodon stem. Neither extra IVS was produced by cuts at the novel 5' splice site, while one (middle IVS) was generated by a cut at the novel 3' splice site. Thus, cleavage at a double-stranded region seems unfavorable, and the endonuclease appears to prefer to cleave at sites as well-removed as possible from a double helical stem.

In order to further test the hypothesis that the endonuclease selects the 5' and 3' splice sites by measurement of fixed distances from the central region of the pre-tRNA molecule, the anticodon stem was prolonged by 2 base pairs by insertion of U•A and G•C base pairs (see Figure 2, panel E); this mutant is designated $(^{UA}_{GC})$▽. Splicing of this precursor shows that the IVS is now four bases longer than wild type (Figure 4, panel A, lanes 10 and 12). Fingerprint analysis of this IVS show

that the 5' and 3' splice sites have been shifted two bases upstream and downstream of the wild-type 5' and 3' splice sites, respectively (compare Panels A and G, Figure 5). Two shorter IVSs were also produced as in GC▽, but in considerably diminished amounts. Thus, this result further strengthens the above distance measurement hypothesis for splice site selection.

In the ΔU mutant, deletion of the invariant U33 residue preceding the anticodon did not affect cleavage and ligation, but the IVS is one base shorter and is unique (Fig. 3, panel III, lane 11). Fingerprint analysis reveals that the 5' splice site has been shifted one base downstream, but the 3' splice site is unaffected (Fig. 5, Panel H). This result is consistent with the measuring mechanism of splice site selection described above. Further, it suggests that selection of the 5' and 3' splice sites are two distinct and independent steps. In this variant, the single deleted residue is from the left hand strand of the IVS-anticodon stem and the shift, one base downwards, is only at the 5' splice site; the 3' splice site is not affected. Thus, it seems probable that the endonuclease may have 2 distinct catalytic centers, one for each splice site.

Both alterations in G▽ and ΔU are expected to produce a kink in the anticodon and IVS-anticodon stems, respectively, disrupting the respective double-helical structures there. However, ΔU cleaves normally, whereas G▽ is inactive (Figure 3, panel III, lanes 9 and 11). Apparently, disruption of helicity is tolerated in the IVS-anticodon stem, but not in the anticodon stem. NMR structural analysis of wild type pre-tRNA^{Phe} reveals that the IVS-anticodon stem region has a very low Tm and probably has an unstable secondary structure (C. Hall, personal communication). Thus, helicity of the IVS-anticodon stem is probably transient in the first place.

An a priori requirement for distance measurement is the existence of a fixed reference point from which distances can be precisely and reproducibly measured. The thoracic region of the pre-tRNA contains the greatest concentration of the totally- and semi-conserved residues as well as tertiary base pairs in the molecule (Rich and RajBhandary, 1976); therefore, logic dictates that such a reference point - and hence the endonuclease-bound domain - must be in this region. To test this hypothesis, a base pair was inserted into the portion of the molecule above this region - the acceptor stem. This mutant is designated AU▽ (see Figure 2, panel E).

Splicing of the AU▽ mutant reveals that the 5' and 3' halves are each lengthened by one base due to the bp insertion, and the IVS did not change in size (Figure 4A, lane 9). Thus, cleavages were at the wild type splice sites. This means that the endonuclease binds below the AU insertion rather than above it since if it bound above it, the IVS would have increased in length by two bases, while the two exons would have been unaffected. Thus, the measurement process commences at a point below the AU insertion. But what is below the AU insertion is the thoracic portion of the molecule. This is an entirely reasonable result, since this region of the molecule has the most concentration of invariant residues most of which are those involved in tertiary interactions. The reference point for the measurement mechanism is thus most probably located there. For a summary of novel cleavage sites, please see Fig. 6A.

The measuring activity is not likely to be a property of some exogenous factor/s present in the enzyme preparation, but is rather an inherent property of the endonuclease itself. This is supported by the observation (data not shown) that use of a relatively impure enzyme preparation (Peebles et al., 1983) gave identical results as that of a highly purified preparation (P. Green, personal

communication). The distance measurement mechanism for splice site selection is depicted by the cartoon in Fig. 6B.

Highly Conserved Purine Residue 3' Proximal to the Anticodon May Be Essential for Cleavage at the 5' Splice Site

This synthetic wild-type pre-tRNA^{Phe} has been synthesized in large-scale, purified by HPLC, and subjected to NMR structural analysis. Preliminary results indicate that the mature domain of pre-tRNA^{Phe} closely resembles the well-known three-dimensional structure of tRNA^{Phe} (Quigley and Rich, 1976; Sussman and Kim, 1976), the difference lying only in the anticodon-IVS region. Although possessing internal primary sequence complementarity, this region seems to have very low melting temperature and its secondary structure is probably unstable (C. Hall, pers. comm.). To provide a stronger argument in support of this observation, we constructed a pre-tRNA which has completely no potential to form a secondary structure in this region. In this pre-tRNA variant, the entire portion below the anticodon stem is a poly U sequence; a couple of G residues were included in this sequence to facilitate subsequent intron sequence analysis by RNAse T1 fingerprinting (Figure 2, panel D). This mutant pre-tRNA is designated poly U. Wild-type splice sites are those that preserve the exon lengths, and give rise to a 14-base IVS. Splicing of poly U pre-tRNA shows that only a 3' half molecule and a 2/3 intermediate are produced (Figure 4B, lane 7). We have confirmed this through fingerprint analyses (data not shown). Thus, cleavage is only at the 3' splice site. Moreover, no mature tRNA is produced in the presence of ligase (Figure 4B, lane 10).

The 2/3 intermediate produced also migrates as a triplet. As will be demonstrated later, this is due to a length heterogeneity at the poly U sequence

resulting from imprecise transcription of the poly(dA) template strand by T7 RNA polymerase.

All nine S. cerevisiae pre-tRNAs contain a purine residue (guanines in 4 of the 9 cases) immediately preceding the 5' splice site (Ogden et al., 1984). In pre-$tRNA^{Phe}$, this residue is a G. We were interested in the possibility that the poly U sequence, which has replaced the 5' splice site region (but not the 3' splice site region in the poly U precursor), inhibits cleavage at that site. Alternatively, loss of the highly conserved purine (G) residue preceding the 5' splice site renders this site unrecognizable to the endonuclease. The G residue 3 bases downstream of this site apparently could not function as such (Figure 2, panel D).

Thus, we reintroduced a G residue at the appropriate position in the poly U pre-tRNA; this mutant is designated poly U(G) (Figure 2, panel D). Splicing of this precursor shows that cleavage at the 5' splice site has been rescued, and all splicing intermediates - 5' half, 3' half, IVS and 2/3 molecule - are now produced (Figure 4B, lane 9); if ligase is included, mature tRNA appears (Figure 4B, lane 12). Fingerprint analysis of the mixture of all three poly U(G) introns shows all the expected oligonucleotides from RNAse T1 digestion of this RNA fragment (Fig. 5, Panel I). These results suggest that the highly conserved purine residue immediately preceding the 5' splice site is essential for cleavage at this site. The intron produced (as well as the 2/3 intermediate), however, migrates as a triplet, IVS-1, -2, and -3. To determine the nature of each band comprising this triplet, each was isolated, digested with RNAse T1, and the products separated by gel electrophoresis. The results are shown in Fig. 6 and Table 2. Lanes I, II, III, and IV were run in parallel to provide oligonucleotide length markers. Lanes V, VI and VII show clearly that the 5' and 3' terminal oligonucleotides, U_2G and UU>P, respectively, are identical in each intron, but the internal fragments are different.

The longest IVS contains 8 consecutive uridine residues (U_8G), corresponding to precise transcription of the 8 consecutive adenine residues in the gene; the second IVS, 7 uridine residues (U_7G); and the third IVS, six (U_6G). Thus, the triplet nature of the intron is not due to imprecise cleavage by the endonuclease at the 3' splice site, rather to imprecise transcription of the poly U stretch by T7 polymerase, resulting in the "skipping" of one, and sometimes two, adenine residues in the DNA template. These results also support the idea that intron primary sequence does not influence cleavage or splice site selection by the endonuclease.

The cleavage rescue in poly U(G) is also partial, since splicing efficiency is less than wild type (compare lanes 8 and 9, and 11 and 12, Fig. 4B). This reduced efficiency could be the effect of the novel environment brought about by the poly U sequence. This result nonetheless indicates that secondary structure in the anticodon-IVS stem-loop region is not necessary for splicing, and clearly supports the previous NMR structural finding that this region has low Tm and might not possess a stable secondary structure.

Finally, except for the poly U precursor whose cleavage products - a 2/3 molecule and a 3' half tRNA - are not substrates for ligase, all the variant precursors which were cleaved can also be ligated with equivalent efficiency (see Table 3). This high correlation between cleavage and ligation efficiencies demonstrates that the endonuclease and ligase have a lot of recognition elements in common. However, it does not rule out the possibility that these enzymes each have unique recognition elements of their own. In the S. pombe pre-*sup*3-e, replacement of U33 with a C does not affect cleavage, but ligation is severely diminished (Greer et al., 1987). Thus, in that system, U33 is probably recognized by the ligase, but not by the endonuclease. In the present study, however, the ΔU mutant lacks this residue, but both cleavage and ligation are normal although the 5'

splice site has been shifted one base downstream. The simplest explanation for this apparent disparity is that this may reflect an inherent difference between substrate properties of class I (with small extra arms) and class II (with large extra arms) pre tRNAs, to which pre-tRNA^{Phe} and pre-*sup*3-e belong, respectively.

Conclusion

We have presented evidence that the invariant tRNA residues U8 and C56 are probably among its contact points with the endonuclease during the splicing reaction, and that the highly conserved purine residue immediately upstream of the 5' splice site may be essential for cleavage at that site. We have also demonstrated that the endonuclease selects its cleavage sites in the pre-tRNA by a mechanism that maintains the distances of these sites from a point within the central portion of the molecule, where the enzyme probably binds (see Fig. 7A and 7B).

The ultimate test for the validity of these preliminary conclusions is to crystallize highly purified pre-tRNA with highly-purified endonuclease, and subject this cocrystal to a rigorous x-ray crystallographic analysis. This has already been possible for yeast tRNA^{Asp} and its corresponding synthetase (D. Moras, unpublished results). In our laboratory, the endonuclease has proven to be very difficult to purify, but we are now quickly nearing that goal. That should allow us to clone the gene for this enzyme and produce it in massive amounts and in pure form. The present bacteriophage T7 promoter/pre-tRNA^{Phe} gene system, on the other hand, is uniquely suited for the production of large quantities of pure pre-tRNA. The precise mode of interaction of pre-tRNA^{Phe} with the endonuclease can thus be studied to the finest detail.

Experimental Procedures

Materials

All oligodeoxynucleotides were synthesized on an Applied Biosystems 380A DNA synthesizer and purified on a 20% polyacrylamide 7.8M urea gel. All other reagents, enzymes and equipment were as described previously (Reyes and Abelson, 1987).

Gene Construction and RNA Synthesis

Gene constructions were done either by total synthesis (see Table IA and IC) or by site-directed mutagenesis of appropriate single-stranded DNA templates by synthetic oligonucleotides (see Table IA and IB). The latter method has been described previously (Zoller and Smith, 1982). In total gene synthesis, each gel-pure deoxyoligonucleotide was first 5' phosphorylated with ATP by the action of PNK. 1 to 3 μg of each of the 5' phosphorylated oligonucleotide was mixed together and NaCl added to a final concentration of 50 mM and total volume of 100 μl. Exact amount, in μg, of each oligonucleotide was calculated by dividing the length (in number of bases) of each oligonucleotide with that of the shortest oligonucleotide. The shortest oligonucleotide will then automatically be added at 1.0 μg. This solution was covered with paraffin oil, heated to 90°C. for 5 min and slowly cooled to 25°C. 10X ligation buffer and 100 ng of *Eco*RI/*Bam*HI-cleaved M13mp10 RF DNA was then added directly and the mixture incubated at 37°C for 30 min, then cooled on ice to let the sticky ends anneal. Then 50 units of T4 DNA ligase was added and the solution incubated at 15°C overnight. Competent JM101 E. coli cells were then transformed as described (Maniatis et al., 1982), and the resulting plaques screened by hybridization with the appropriate radioactively labeled oligonucleotide. The gene system emerging from this construction consists of a

pre-tRNA^{Phe} gene linked downstream to a bacteriophage T7 promoter (Figure 1A). In all cases, the identity of the wild-type or mutant gene insert was determined by dideoxy sequencing (Sanger et al., 1977) using oligonucleotide ϕ (Table IA) as sequencing primer. Synthesis of the pre-tRNA is done by first cleaving the plasmid with *Bst*NI, and incubating this cleaved DNA with T7 RNA polymerase, as described (Reyes and Abelson, 1987).

In Vitro Splicing Assays

This procedure involves incubation of 7-10 nM of pre-tRNA at 30°C for 30 min. with a highly purified preparation of endonuclease and/or ligase, as described (Ibid.). The above pre-tRNA concentration corresponds to approximately 1/3 to 1/2 the endonuclease K_M, which has been determined previously (Ibid.).

RNA Sequence Analysis

RNA sequence analyses were either Nishi base analyses or fingerprint analyses. The former has been described (Saneyoshi et al., 1972). Fingerprint analysis was done by first incubating approx. 2000 cpm of radioactively labeled RNA with either RNase A or T1 in TE (10/1) in the presence of 25 to 30 µg RNA carrier at a total volume of approximately 8 µl, under paraffin oil, at 42°C for 1 to 2 hrs. Meanwhile, cellulose acetate strips (Schleicher and Schuell, CA 2500, 30 x 550 mm) were soaked in electrophoresis buffer (5 M urea/2 mM EDTA/5% HOAc, pH = 3.5 with pyridine) for 30 min. The digested RNA is then spotted centrally on the strip at a point 10 cm from one end; spotting is done 1 µl at a time to allow drying between each application. However, the rest of the area of the strip is prevented from drying by applying 3 mm Whatman paper strips, soaked in electrophoresis buffer, as wicks. The RNA is then electrophoresed on the strip in pyridine acetate at pH 3.5

using a high-voltage electrophoresis apparatus (Savant Instruments, Inc., Hicksville, N.Y., Model EJB) at 2500 to 3000 volts for 1 to 2 hrs. The electrophoresed RNA fragments were then transferred onto dried 20 x 20 cm PEI plates (Macherey-Nagel, Postfach, Germany), which have been prewashed with water and formic acid (pH = 2.2) each for 30 min, as follows: the cellulose acetate strip was laid along its length on a glass stirring rod and 20 x 2 cm Whatman filter paper strips soaked in water placed along each edge as wicks. In most cases, the portion from the BpB dye and downstream was transferred. The PEI plates, fastened securely onto 20 x 20 cm glass plates by double-stick tapes, was made to contact the strip lengthwise on a line 2 cm from one edge of the plate. A weight was placed on top of the plate and the set-up left undisturbed for approximately 30 min to allow transfer of RNA by capillary action. The plates were then soaked in 90% ethanol, dried, and TLC performed as follows: Two pieces of 20 x 4 cm Whatman Filter paper were attached (touching the PEI surface) to the end of the plate opposite the one where RNA was transferred. This is then soaked in dH_2O, patted dry to remove excess water, and then placed in a glass chamber (l x w x h = 28 x 8 x 21 cm) containing 100 mls of C10 (for RNAse T1 fragments) or C15 (for RNase A fragments) homomix. Glass cover is placed such that the wicks protrude to outside, and the chamber placed in a 65°C oven for 3 to 4 hrs. The plates were then washed in 90% ethanol, dried, and autoradiographed. Spots corresponding to labeled internal fragments in the RNA were then identified according to their known mobility in this system. 5'- and 3'-terminal fragments are expected to migrate anomalously due to extra or missing phosphate residues, respectively.

Acknowledgements

We thank Phillip Green and Eric Phizicky for providing the highly purified preparations of endonuclease and ligase, respectively. We also thank Suzanna Horvath of the Caltech Microchemical Facility for providing the synthetic oligonucleotides, and Jennifer Normanly for the total gene synthesis protocol. We acknowledge the numerous technical suggestions from J. Sampson and L. Behlen of the O. Uhlenbeck laboratory at the University of Colorado; and valuable discussions with Professor O. Uhlenbeck, Calvin Ho, and Kyle Tanner. This work was supported by grants from the American Cancer Society and the NIH.

References

Allison, D. S. and Hall, B. D. (1985). Effects of alterations in the 3' flanking sequence on in vivo and in vitro expression of the yeast SUP4-o $tRNA^{Tyr}$ gene. EMBO J. *4*, 2657-2664.

Clark, M. W. and Abelson, J. (1987). The subnuclear localization of tRNA ligase in yeast. J. Cell Biol. *105*, 1515-1526.

Colby D., Leboy, P.S., and Guthrie, C. (1981). Yeast tRNA precursor mutated at a splice function is correctly processed in vivo. Proc. Natl. Acad. Sci. USA, *78*, 415-419.

Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., and Scazzocchio, C. (1982). Making ends meet: a model for RNA splicing in fungal mitochondria. Nature (London) *300*, 719-724.

Engelke, D. R., Gegenheimer, D., and Abelson, J. (1985). Nucleolytic processing of a $tRNA^{Phe}$ - $tRNA^{Asp}$ dimeric precursor by a homologous component from Saccharomyces cerevisiae. J. Biol. Chem. *20*, 1271-1279.

Galli, G., Hofstetter, H., and Birnstiel, M. (1981). Two conserved sequence blocks within eukaryotic tRNA genes are major promoter elements. Nature (London) *394*, 626-631.

Greer, C. L., and Abelson, J. (1984). RNA splicing: Rearrangement of RNA sequences in the expression of split genes. Trends Biochem. Sci. *9*, 139-141.

Greer, C. L., Peebles, C. L., Gegenheimer, P., and Abelson, J. (1983). Mechanism of action of a yeast RNA ligase in tRNA splicing. Cell *32*, 537-546.

Greer, C. L., Söll, D., and Willis, I. (1987). Substrate recognition and identification of splice sites by the tRNA-splicing endonuclease and ligase from Saccharomyces cerevisiae. Molec. Cell. Biol. *7*, 76-84.

Johnson, P. F. and Abelson, J. (1983). The yeast $tRnA^{Tyr}$ gene intron is essential for correct modification of its tRNA product. Nature (London) *302*, 681-687.

Kim, S. H. (1979). Crystal structure of yeast $tRNA^{Phe}$ and general structural features of other tRNAs. In: Transfer RNA: Structure, Properties and Recognition; Schimmel, P. R., Söll, D., and Abelson, J., eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.) pp. 83-100.

Kline, L., Nishikawa, S. and Söll, D. (1981). Partial Purification of RNAse P from Schizosaccharomyces pombe. J. Cell Biol. *256*, 5058-5063.

Klug, A., Ladner, J., and Roberts, J. D., (1974). The structural geometry of coordinated base changes in transfer RNA. J. Mol. Biol. *89*, 511-516.

Knapp, G., Ogden, R. C., Peebles, C. L., and Abelson, J. (1979). Splicing of yeast tRNA precursors: Structure of the reaction intermediates. Cell *18*, 37-45.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell *31*, 147-157.

Langford, C. J. and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcript. Cell *33*, 519-527.

Lee, M. C. and Knapp, G. (1985). Transfer RNA splicing in Saccharomyces cerevisiae: Secondary and tertiary structures of the substrates. J. Biol. Chem. *260*, 3108-3115.

Maniatis, T., Fritsch, E. F., and Sambrook, J. (1982). Molecular cloning: A laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.

Mattoccia, E., Baldi, M. I., Pande, G., Ogden, R., and Tocchini-Valentini, G. P. (1983). Mutation in the A-block of the yeast $tRNA_3^{leu}$ that allows transcription but abolishes splicing and 5'-end maturation. Cell *32*, 67-76.

Melton, D. A., DeRobertis, E. M., and Cortese, R. (1980). Order and intracellular location of the events involved in the maturation of a spliced tRNA. Nature *284*, 143-148.

Michel, F. and Dujon, B. (1983). Conservation of RNA secondary structures in two intron families including mitochondrial-, chloroplast-, and nuclear-encoded members. EMBO J. *2*, 33-38.

Mount, S. M. (1982). A catalog of splice junction sequences. Nucl. Acids Res. *10*, 459-472.

Mount, S. M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J. (1983). The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell *33*, 509-518.

Newman, A. J., Ogden, R. C., and Abelson, J. (1983). tRNA gene transcription in yeast: effects of specified base substitutions in the intragenic promoter. Cell *35*, 117-125.

Nishikura, K., Kurjan, J., Hall, B. D., DeRobertis, E. M. (1982). Genetic analysis of the processing of a spliced tRNA. EMBO J. *1*, 263-268.

Ogden, R. C., Knapp, G., Peebles, C. L., Johnson, J., and Abelson, J. (1981). The mechanism of tRNA splicing. Trends Biochem. Sci. *6*, 154-158.

Ogden, R. C., Knapp, G., Peebles, C. L., Kang, H. S., Beckmann, J. S., Johnson, P. F., Fuhrman, S. A., and Abelson, J. N. (1980). Enzymatic removal of intervening sequences in the synthesis of yeast tRNAs. In Transfer RNA: Biological Aspects, Söll, D., Abelson, J., and Schimmel, P. R., eds. (Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.), pp. 173-190.

Ogden, R., Lee, M. C., Knapp, G. (1984). Transfer RNA splicing in *S. cerevisiae*: Defining the substrates. Nucl. Acids Res. *12*, 9367-9382.

Peebles, C. L., Gegenheimer, D., and Abelson, J. (1983). Precise excision of intervening sequences from precursor tRNAs by a membrane-associated yeast endonuclease. Cell *32*, 525-536.

Peebles, C. L., Ogden, R. C., Knapp, G., and Abelson, J. (1979) Splicing of yeast tRNA precursors: A two-stage reaction. Cell *18*, 27-35.

Phizicky, E. M., Schwartz, R. C. and Abelson, J. (1986) *S. cerevisiae* tRNA ligase: Purification of the protein and isolation of the structural gene. J. Biol. Chem. *261*, 2978-2986.

Quigley, G. J. and Rich, A. (1976). Structural domains of transfer RNA molecules. Science *194*, 796-806.

Raymond, K. C., Raymond, G. J., and Johnson, J. D. (1985). In vivo modulation of yeast tRNA gene expression by 5'-flanking sequences. EMBO J. *4*, 2649-2656.

Reyes, V. M., Newman, A., and Abelson, J. (1986). A mutational analysis of the coordinate expression of the yeast $tRNA^{Arg}$-$tRNA^{Asp}$ gene tandem. Molec. Cell. Biol. *6*, 2436-2442.

Reyes, V. M., and Abelson, J. (1987). A synthetic substrate for tRNA splicing. Anal. Biochem. *166*, 90-106.

Rich, A., and RajBhandary, U. L. (1976). Transfer RNA: Molecular structure, sequence and properties. Ann. Revs. Biochem. *45*, 805-860.

Saneyoshi, M., Ohashi, Z., Harada, F., and Nishimura, S. (1972). Isolation and characterization of 2-methyladenosine from Escherichia coli $tRNA_2^{Glu}$, $tRNA_1^{Asp}$, $tRNA_1^{His}$, and $tRNA^{Arg}$. Biochim. Biophys. Acta *262*, 1-10.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463-5467.

Sharp, S., DeFranco, D., Dingermann, T., Farrell, D., and Söll, D. (1981). Internal control regions for transcription of eukaryotic tRNA genes. Proc. Natl. Acad. Sci. USA *78*, 6657-6661.

Strobel, M. C., and Abelson, J. (1986). Effect of intron mutations on processing and function of Saccharomyces cerevisiae SUP53 tRNA in vitro and in vivo. Molec. Cell. Biol. *6*, 2663-2673.

Sussman, J. L. and Kim, S. H. (1976). Three-dimensional structure of a transfer RNA in two crystal forms. Science *192*, 853-858.

Swerdlow, H. and Guthrie, C. (1984). Structure of intron-containing tRNA precursors. J. Biol. Chem. *259*, 5197-5207.

Traboni, C., Ciliberto, G. and Cortese, R. (1984). Mutations in the B-box of the promoter of a eukaryotic $tRNA^{Pro}$ gene affects rate of transcription, processing and stability of transcripts. Cell *36*, 179-187.

Westaway, S. K., Phizicky, E. M. and Abelson, J. (1987). Structure and function of the yeast tRNA ligase gene. J. Biol. Chem., in press.

Willis, I., Hottinger, H., Pearson, D., Chisolm, V., Leupold, U. and Söll, D. (1984). Mutations affecting excision of the intron from a eukaryotic dimeric tRNA precursor. EMBO J. *3*, 1573-1580.

Zoller, M. J. and Smith, M. (1982). Oligonucleotide-directed mutagenesis using M13-derived vectors: an efficient and general procedure for the production of point mutations in any fragment of DNA. Nucl. Acids. Res. *10*, 6487-6500.

Figure Legends

Figure 1.

(A) Schematic Diagram of the pre-tRNA^{Phe} Synthetic System.

The pre-tRNA^{Phe} gene is linked to a bacteriophage T7 promoter and transcription starts precisely at the 5' end of the gene. The *Bst*NI site at the 3' end of the gene is magnified; note the 3'-GGT-5' terminal sequence produced by BstNI cleavage in the coding strand, ensuring that the transcript ends with 5'-CCA-3'. The deoxyoligonucleotides (1a, b, 2a, etc.) refer to those shown in Table 1A.

(B) Two-dimensional Nishi Base Analysis of RNA to Determine the Identity of Its 5' Nucleotide.

The second nucleotide of the transcript is a C; by nearest-neighbor transfer, the first base can only be detected if the label is α-^{32}P - CTP. In the lower right panel, no 5' GMP is included during transcription and all of the 5' base (and hence the transcript) contain a 5' triphosphate (pppGp). In the lower left, 5' GMP was included in the transcription, and almost all the transcripts start with 5' monophosphate (pGp). The top 2 panels are controls wherein the transcription label is α-^{32}P-UTP; the first uridine residue is the sixth base in this transcript, hence the 5' nucleotide cannot be detected in either case.

(C) Comparison of Splicing Efficiency of pre-tRNA^{Phe} Containing a 5' Monophosphate (5' P) with That Containing a 5' Triphosphate (5' PPP).

The plus signs in the headings depict relative amount of indicated component included in the incubation. Splicing products are identified: p = pre-tRNA; m = tRNA; 2/3 = 2/3 intermediate; 5' h and 3' h = 5' and 3' halves, respectively; i = IVS. We have previously shown (Reyes and Abelson, 1987) that the 2/3 intermediate consists of the 5' half and IVS; thus, endonuclease cleavage is ordered, and the first cut is at the 3' splice site. Further, the triplet nature of the band corresponding to

the 3' half is due to the heterogeneity of the 3' end of the transcript: a G and a random residue N are usually added by T7 RNA polymerase after the encoded-CCA 3' terminus.

Figure 2.

(A) The three-dimensional structure of $tRNA^{Phe}$ showing the locations of some of the important residues (numbers).

(B) Two-dimensional cloverleaf structure of $tRNA^{Phe}$, showing its 9 tertiary base interactions.

(C) Cloverleaf structure of the present pre-$tRNA^{Phe}$ molecule, showing the base substitutions, insertions and deletions constructed in this study, together with their respective designations (boxed entries); solid arrowheads indicate wild-type splice sites.

(D) Cloverleaf structure of the poly U precursor showing the expected splice sites (solid arrowheads), and the site (small arrow) where the G residue has been reintroduced in the poly U(G) precursor.

(E) Cloverleaf structure of wild-type pre-$tRNA^{Phe}$ showing the locations of the base-pair insertions in the AU▽, GC▽ and $\binom{UA}{GC}$ ▽ mutants.

Figure 3.

Splicing Assay of the First Eleven pre-$tRNA^{Phe}$ Mutants. Identities are indicated above each lane. Panels I (unincubated precursors) and II (precursors incubated in buffers only) are controls done to detect any spontaneous or nonspecific nuclease-mediated degradations, respectively. Panels III (incubation with endonuclease plus buffers) and IV (incubation with endonuclease and ligase, plus buffers) show the products of cleavage and ligation, respectively.

Figure 4.

(A) Splicing Assay of the Base-Pair Insertion Mutants of pre-$tRNA^{Phe}$.

Identities are written above each lane; panel U = unincubated precursors; B =

precursors incubated with buffers only; E = precursors incubated with endonuclease; and E + L = precursors incubated with endonuclease and ligase. The various splicing products are indicated, as in Figure 1C legend. Note specifically the increase in length of the IVS in steps of two as a base pair is inserted into the anticodon stem (lanes 10, 11, 12; and 14, 15, 16).

(B) Splicing Assay of the Poly U and Poly U(G) Precursors.

Designations U, B, E and E + L refer to incubation conditions as described in figure legend 4A. Wild-type IVS has 19 bases, while poly U and poly U(G) IVSs both have 14 bases; thus, their respective 2/3 intermediates and IVSs (poly U is cleaved only at the 3' splice site and does not release an IVS) differ in length, and hence, mobility, (lanes 8, 9, and 10). Both poly U and poly U(G) precursors are apparently cleaved by the endonuclease at the 3' splice site in a staggered fashion as evidenced by their respective triplet 2/3 intermediate bands, and the triplet IVS band of poly U(G). Further, ligase has a 2',3'-cyclic phosphatase activity, and the splitting of the wild-type and poly U(G) IVS bands in the presence of ligase is most probably due to this activity (compare lanes 8 and 9 with lane 11).

Figure 5.

(A-I) Fingerprint Analyses of Introns from Some pre-tRNA^{Phe} Mutants.

The parent precursor of the IVS is indicated (first entry). "GC▽ top," "GC▽ mid," and "GC▽ low" refer to the three intron bands generated in the splicing of the GC▽ mutant (lane 10, Figure 3, panels III and IV; also Figure 4A, lanes 11 and 15). ATP or UTP indicates that the RNA was labeled using α-^{32}P-ATP or -UTP, respectively. T1 or A indicates digestion with either RNAse T1 or A, respectively. The first (horizontal) and second (vertical) dimensions of each fingerprint is indicated. Each spot is identified; minor, unindentified spots are degradation products.

Figure 6.

One-dimensional 20% polyacrylamide gel electrophoresis of RNAse T1-digestion fragments of the three poly U(G) introns (lanes V, VI and VII). RNAase T1-digestion products of wild-type (lane I), poly U (lane II) poly U(G) (lane III), and 5' half poly U(G) (lane IV), was run in parallel to provide length markers. Each band is identified in Table 2. Numbers on the left-hand side denote oligonucleotide lengths in number of bases. Fragment identities printed on the right-hand side refer to those on lanes V, VI, and VII only.

Figure 7.

(A) Summary of Endonuclease Cleavage Sites in the Various pre-tRNAs.

Wild-type splice sites are indicated by closed arrowheads; novel splice site in the mutants, by open arrowheads. In each precursor (boxed entry) splice sites are indicated below it, e.g., in the AU▽ precursor, the 3' splice site is 1 and the 5' splice site is 4.

(B) A Cartoon Illustrating the Endonuclease Distance Measurement Mechanism for Selecting the 5' and 3' Splice Sites in the pre-tRNA.

The ruler is meant to emphasize that the measurement process is precise, and that these distances are fixed. The bracket refers to the putative endonuclease-bound region in the pre-tRNA where measurement is initiated; the question mark signifies the present uncertainty as to the boundaries and exact location of this region. Scissors were meant to signify cutting without regard to sequences, but only distances.

Table 1. Construction of Wild-Type and Variant pre-tRNA^{Phe} Gene Systems

A. Sequences of Synthetic Oligodeoxynucleotides Employed

	Name	Length (number of bases)	Sequence (5' → 3')
1.	1a	20	AAT TGC TGC AGT AAT ACG AC
2.	2a	38	TCA CTA TAG CGG ATT TAG CTC AGT TGG GAG AGC GCC AG
3.	4a	35	ACT GAA GAA AAA ACT TCG GTC AAG TTA TCT GGA GG
4.	5	31	TCC TGT GTT CGA TCC ACA GAA TTC GCA CCA G
5.	1b	24	TAT AGT GAG TCG TAT TAC TGC AGC
6.	2b	38	TCT TCA GTC TGG CGC TCT CCC AAC TGA GCT AAA TCC GC
7.	4b	19	TAA CTT GAC CGA AGT TTT T
8.	3b	43	GAT CCT GGT GCG AAT TCT GTG GAT CGA ACA CAG GAC CTC CAG A
9.	19	19	GGC GCT CTC GCA ACT GAG C
10.	56	19	CTG TGG ATC CAA CAC AGG A
11.	15	19	CTC TCC CAA TTG AGC TAA A
12.	48	25	TCT GTG GAT CGA ACA CAG AAC CTC C
13.	8	20	CAA CTG AGC TGA ATC CGC TA
14.	gi	21	TCT TCA GTC TCG GCG CTC TCC
15.	ci	21	CAG GAC CTC CGA GAT AAC TTG
16.	du	22	GTT TTT TCT TCG TCT GGC GCT C
17.	7a'	39	TGA CTG AAG AAA AAA CTT CGG TCA AGT TAT CAC TGG AGG
18.	8b'	40	TCT TCA GTC ACT GGC GCT CTC CCA ACT GAG CTA AAT CCG C
19.	9b'	45	GAT CCT GGT GCG AAT TCT GTG GAT CGA ACA CAG GAC CTC CAG TGA
20.	8a	39	TCA CTA TAG CGG AAT TTA GCT CAG TTG GGA GAG CGC CAG
21.	6	32	TCC TGT GTT CGA TCC ACA GAA TTT CGC ACC AG
22.	10b	39	TCT TCA GTC TGG CGC TCT CCC AAC TGA GCT AAA TTC CGC
23.	11b	44	GAT CCT GGT GCG AAA TTC TGT GGA TCG AAC ACA GGA CCT CCA GA
24.	6a	30	ACT TTT TTT GTT TTT TTT GTT ATC TGG AGG
25.	7b'	38	ACA AAA GTC TGG CGC TCT CCC AAC TGA GCT AAA TCC GC
26.	6b	14	TAA CAA AAA AAA CA
27.	9a	30	ACT TTT GTT GTT TTT TTT GTT ATC TGG AGG
28.	ϕ	20	GGC TGC AGG TCG ACT CTA GA

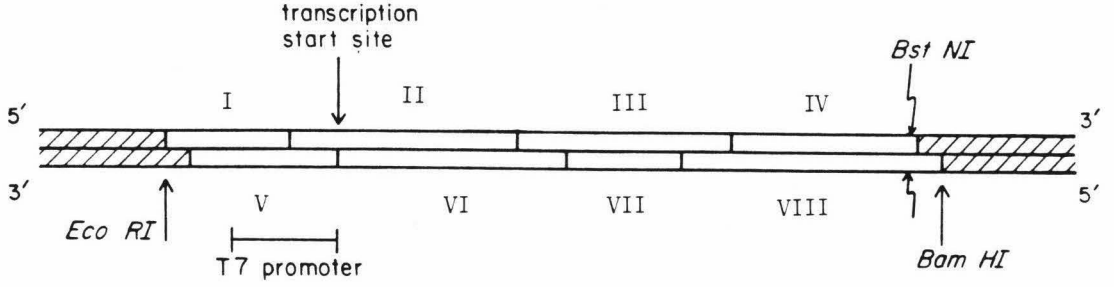
Table 1 (continuation)

B. Construction by Oligonucleotide-Directed Mutagenesis

	Pre-tRNA Gene	Single-stranded DNA Template	Mutagenic Oligodeoxynucleotide Primer
1.	G19C	WT	19
2.	C56G	WT	56
3.	G19C-C56G	C56G	19
4.	G15A	WT	15
5.	C48T	WT	48
6.	G15A-C48T	G15A	48
7.	U8C	WT	8
8.	GV	WT	gi
9.	GCV	GV	ci
10.	ΔU	WT	du
11.	AvaII	n/a[a]	n/a

[a]This pre-tRNA variant was synthesized by cleaving wild-type pre-tRNA^{Phe} plasmid DNA with AvaII restriction endonuclease (instead of BstNI) and then transcribing this truncated wild-type pre-tRNA^{Phe} gene with T7 RNA polymerase.

C. Construction by Total Gene Synthesis [a]



		I	II	III	IV	V	VI	VII	VIII
1.	WT	1a	2a	4a	5	1b	2b	4b	3b
2.	poly U	1a	2a	6a	5	1b	7b	6b	3b
3.	poly U(G)	1a	2a	9a	5	1b	7b′	6b	3b
4.	AUV	1a	8a	4a	6	1b	10b	4b	11b
5.	$(^{UA}_{GC})$V	1a	2a	7a′	5	1b	8b′	4b	9b′

[a]Entries found in the table refer to the oligonucleotides (see Part A, this table) in the various positions (I, II, III, etc.) indicated in the diagram above it. Hatched regions in the diagram represent M13mp10 sequences.

Table 2. Labeled RNAse T1 digestion products of α-^{32}P-UTP-labeled wild-type pre-tRNA^{Phe} (a-i), poly U pre-tRNA (j-s), poly U(G) pre-tRNA (t-bb), 5' half poly U(G) pre-tRNA (cc-ff), and the three poly U(G) introns, IVS-1 (gg-ii), IVS-2 (jj-ll), and IVS-3 (mm-oo).

I.	a	$A_6C_2U_2G$		l[2]	A_3C_3UG		v[3]	ACU_4G (x2)		ff	U_2G
	b	A_3C_3UG		m[3]	ACU_4G		w	A_2U_3G, A_2CU_2G	V.	gg	U_8G
	c	ACU_4G		n	A_2U_3G, A_2CU_2G		x	AC_2UG, C_2U_2G		hh[4]	UU>P
	d	A_2U_3G, A_2CU_2G		o	AC_2UG, C_2U_2G		y	CU_2G		ii	U_2G
	e,f[1]	AC_2UG, A_2CUG, C_2U_2G		p	CU_2G		z	U_2G (X2)	VI.	jj	U_7G
		ACUG, CU_2G		q	U_2G		aa	UG		kk[4]	UU>P
	g	U_2G		r	UG		bb	G		ll	U_2G
	h	UG		s	G	IV.	cc	ACU_4G>P	VII.	mm	U_6G
	i	G(X2)	III.	t	U_8G		dd	A_2U_3G		nn[4]	UU>P
II.	j	ACU_7G		u[2]	A_3C_3UG		ee	AC_2UG		oo	U_2G
	k	U_8G									

Letter designations of the various fragments are from Fig. 6, where the length of each fragment is also indicated.

(1)The heterogeneity of the fragments comprising these two populations might have given rise to the extra band between those for n=5 and n=4.

(2)These bands probably contain small amounts of U_7G resulting from T7 RNA polymerase "slipping", as described in the text.

(3)Similarly, these bands probably contain small amounts of U_6G.

(4)The identities of these bands as UU P have been verified by RNAse T1 fingerprinting (data not shown).

Table 3. Summary of In Vitro Splicing Results

	pre-tRNA[a]	Molecular Formula A	C	G	U	Relative Cleavage Efficiency[b]	Relative Ligation Efficiency[b]	Length (number of bases) Precursor[c]	Mature tRNA	2/3 pre-tRNA[d]	5' Half	3' Half	IVS
I.	WT	26	21	26	22	+++++	+++++	95	76	56	37	39	19
II.	U8C	26	22	26	21	-	-	95	n/a	n/a	n/a	n/a	n/a
III.	1. G15A	27	21	25	22	+++	++++	95	76	56	37	39	19
	2. C48T	26	20	26	23	+++	+++	95	76	56	37	39	19
	3. G15A-C48T	27	20	25	23	+++++	+++++	95	76	56	37	39	19
IV.	1. G19C	26	22	25	22	++++	+++++	95	76	56	37	39	19
	2. C56G	26	20	27	22	-	-	95	n/a	n/a	n/a	n/a	n/a
	3. G19C-C56G	26	21	26	22	-	-	95	n/a	n/a	n/a	n/a	n/a
V.	1. AU$_\nabla$	27	21	26	23	++++	++++	97	78	57	38	40	19
	2. ΔU	26	21	26	21	++++	++++	94	76	55	37	39	18
	3. G$_\nabla$	26	21	27	22	-	-	96	n/a	n/a	n/a	n/a	n/a
	4. GC$_\nabla$	26	22	27	22	+++	++	97	76	58	37	39	21
	5. $\binom{UA}{GC}_\nabla$	27	22	27	23	+++	++	99	76	60	37	39	23
VI.	1. poly U	16	18	23	33	+++[e]	-	90	n/a	51[e]	n/a	39	n/a
	2. poly U(G)	16	18	24	32	+++	++	90	76	51	37	39	14
VII.	AvaII	19	11	20	14	-	-	64	n/a	n/a	n/a	n/a	n/a

[a]These precursors have been grouped into seven families according to the mutations they possess, as shown.

[b]Relative cleavage and ligation efficiencies were determined by liquid scintillation counting of gel slices containing the combined exons and mature tRNA bands, respectively.

[c]These numbers refer to the pre-tRNA population terminating in 3'CCA_{OH}; the minor subpopulations ending in 3'CCAG and 3'CCAGN (N = any nucleotide) are not considered.

[d]The 2/3 pre-tRNA molecule is composed of the 5' half and IVS, as the endonuclease cleaves first at the 3' splice site; this has been proven for wild type precursor by RNA sequence analysis , and assumed for all these variant pre-tRNAs.

[e]The endonuclease cleaves this precursor only at the 3' splice site giving a 2/3 molecule and the 3' half; cleavage is staggered, giving rise to a triplet band corresponding to the 2/3 molecule; the value of 51 for the length of the 2/3 molecule refers to cleavage at the wild type 3' splice site; the other two cleavages are each tentatively assigned one base upstream and downstream of this cleavage site.

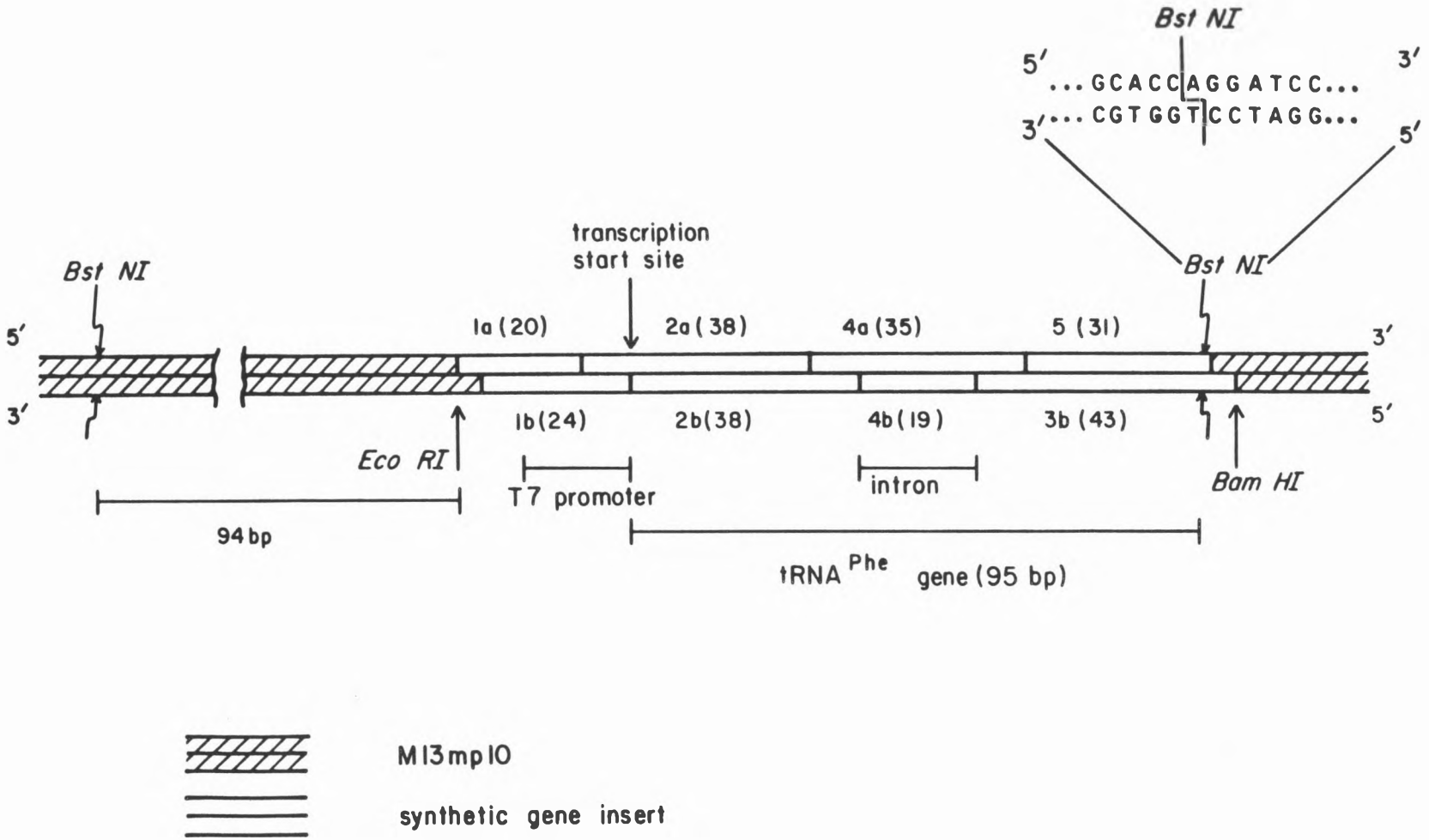

Bst NI
5′
3′
...GCACCAGGATCC...
3′...CGTGGTCCTAGG...
5′
Bst NI
transcription
start site
Bst NI
5′
3′
1a (20)
2a (38)
4a (35)
5 (31)
3′
5′
1b (24)
2b (38)
4b (19)
3b (43)
Eco RI
T7 promoter
intron
Bam HI
94 bp
tRNA Phe gene (95 bp)
M13mp10
synthetic gene insert

Figure 1A

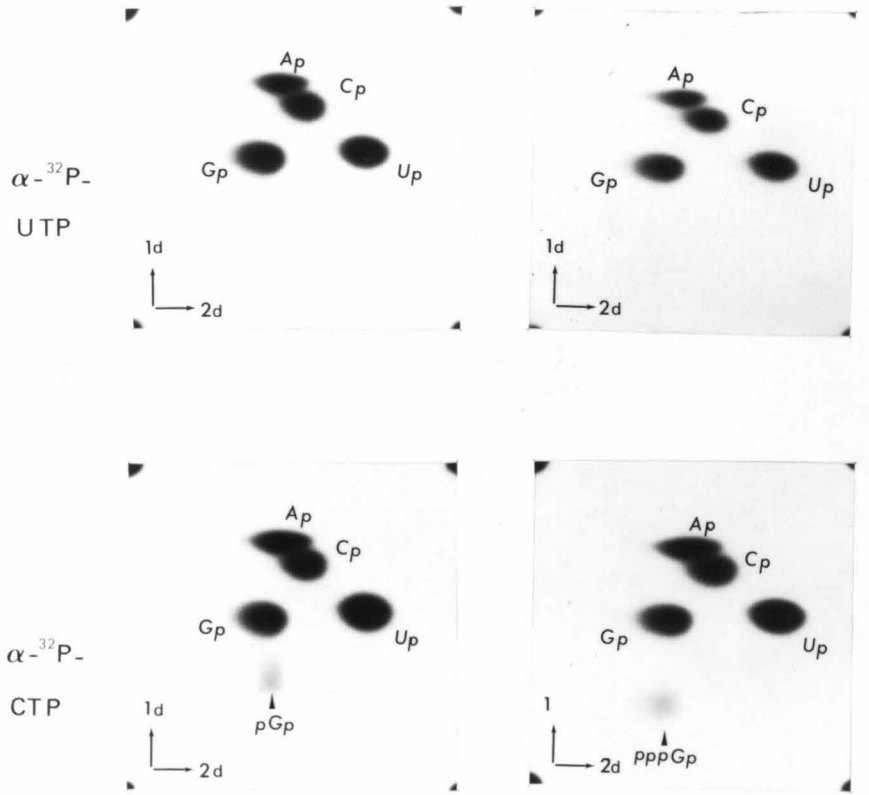


Figure 1B

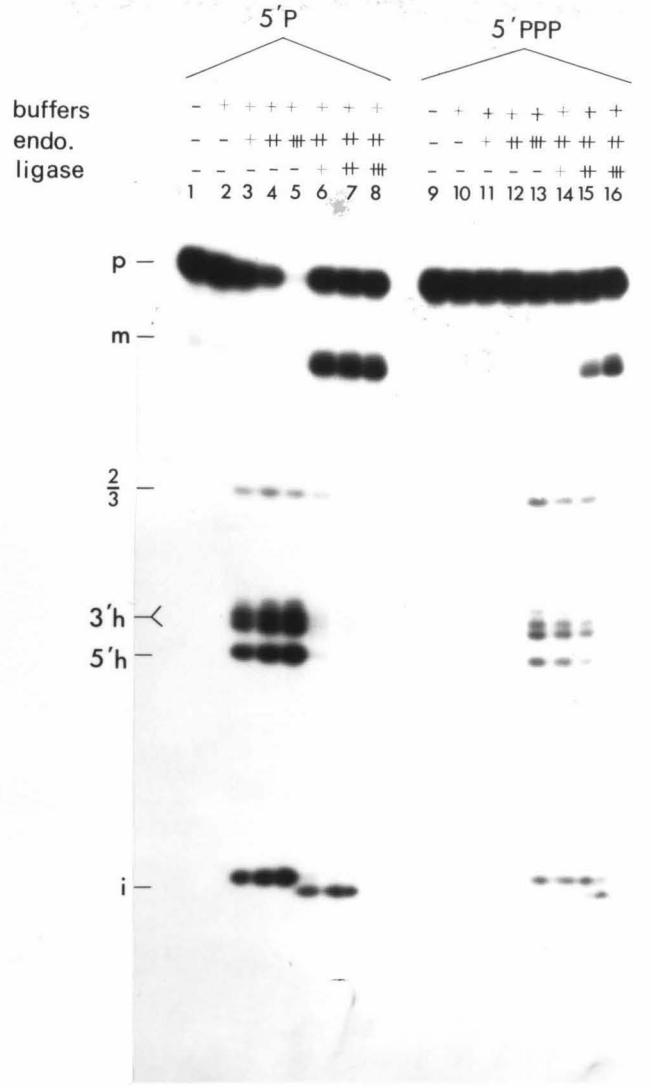


Figure 1C

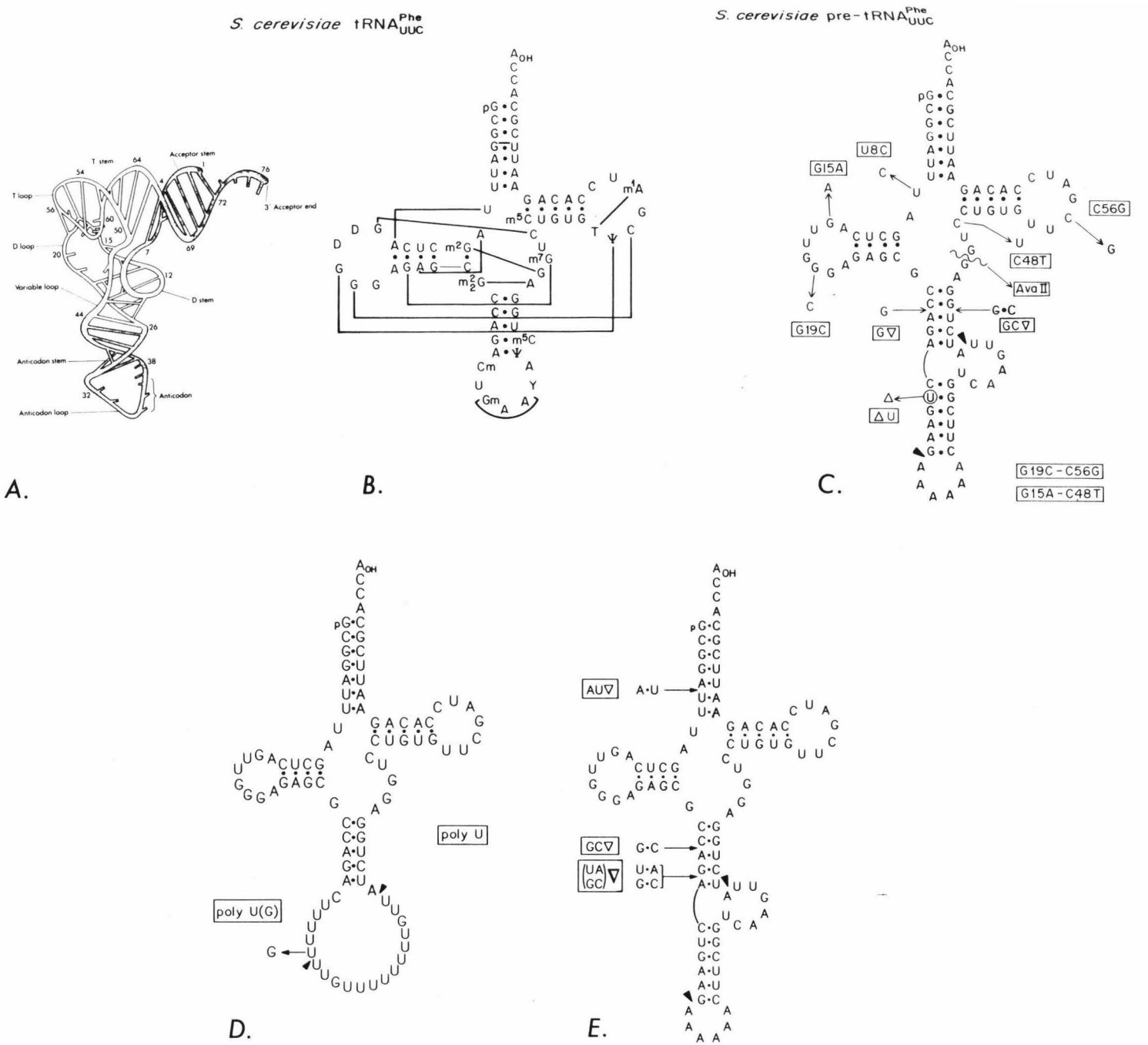
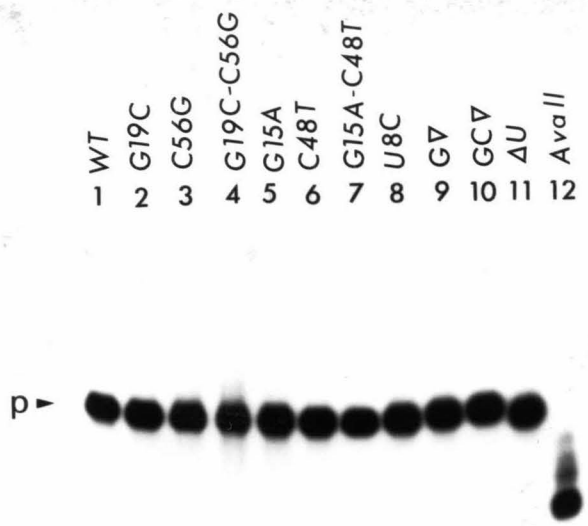
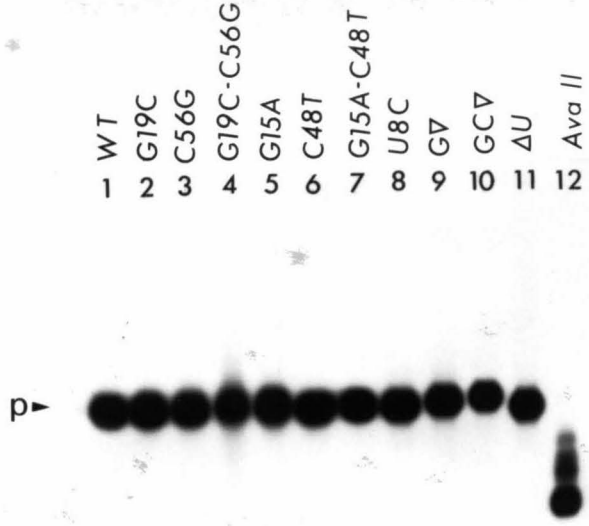

S. cerevisiae tRNA$^{Phe}_{UUC}$
S. cerevisiae pre-tRNA$^{Phe}_{UUC}$
T stem
Acceptor stem
T loop
D loop
Variable loop
D stem
Anticodon stem
Anticodon
Anticodon loop
3' Acceptor end
A.
B.
C.
U8C
G15A
C56G
C48T
AvaII
G19C
GC∇
ΔU
G19C - C56G
G15A - C48T
poly U
poly U(G)
AU∇
GC∇
D.
E.

Figure 2

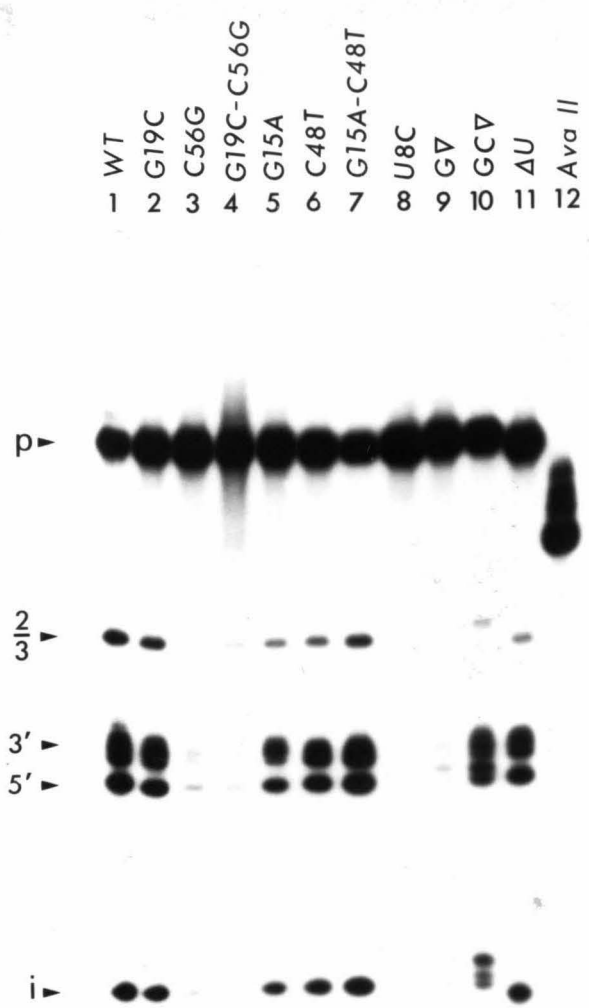


I. unincubated precursors

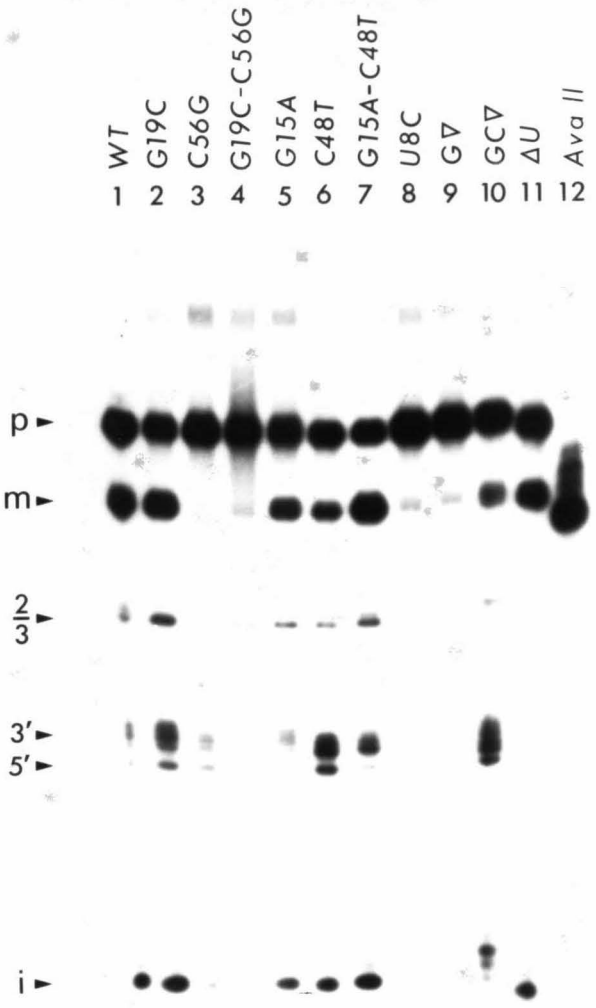


II. buffers only

Figure 3



III. endonuclease only



IV. endonuclease + ligase

Figure 3

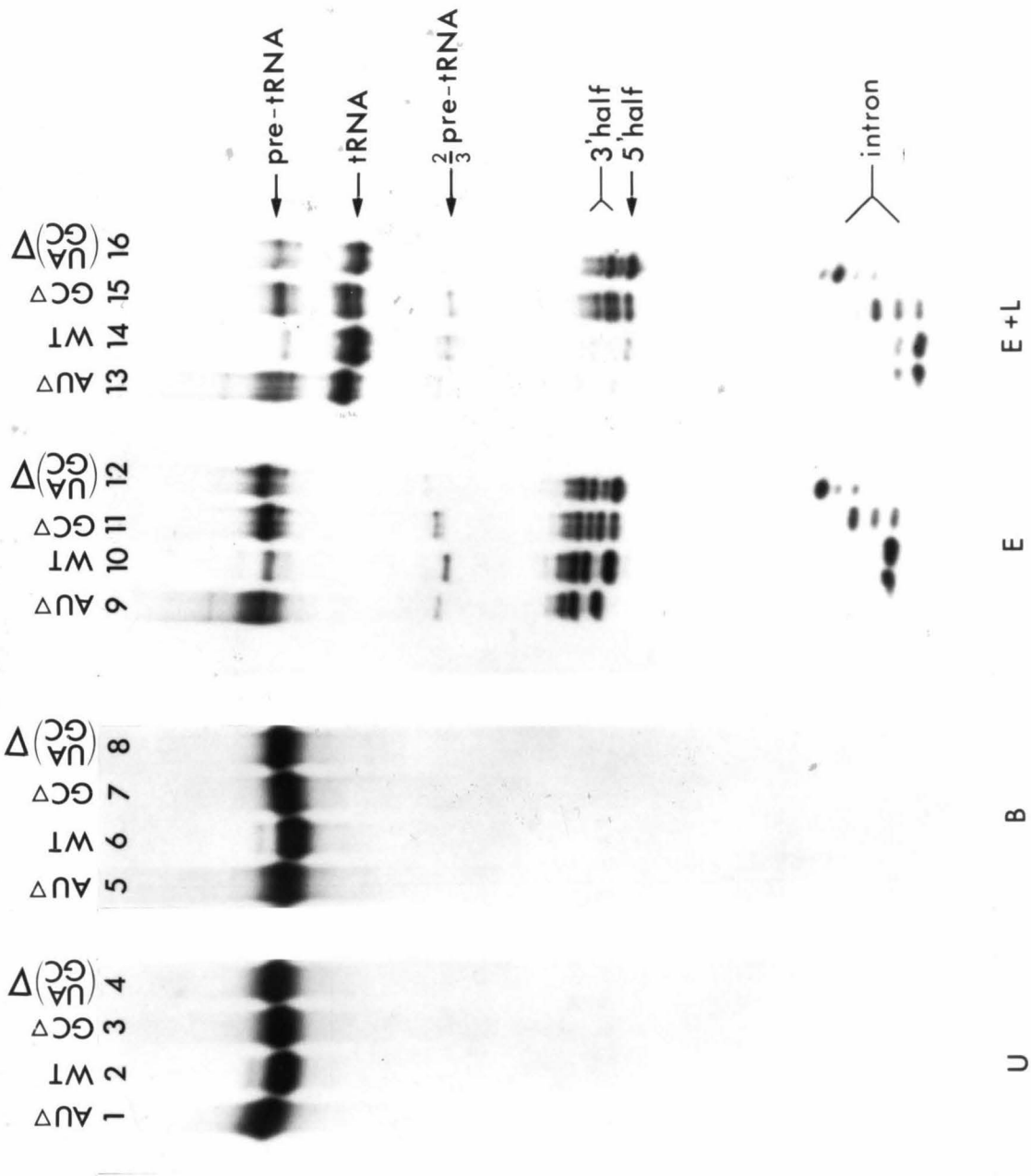

pre-tRNA
tRNA
⅔ pre-tRNA
3'half
5'half
intron
Δ(GUA GC)
ΔGC
WT
ΔU
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
U B E E+L

Figure 4A

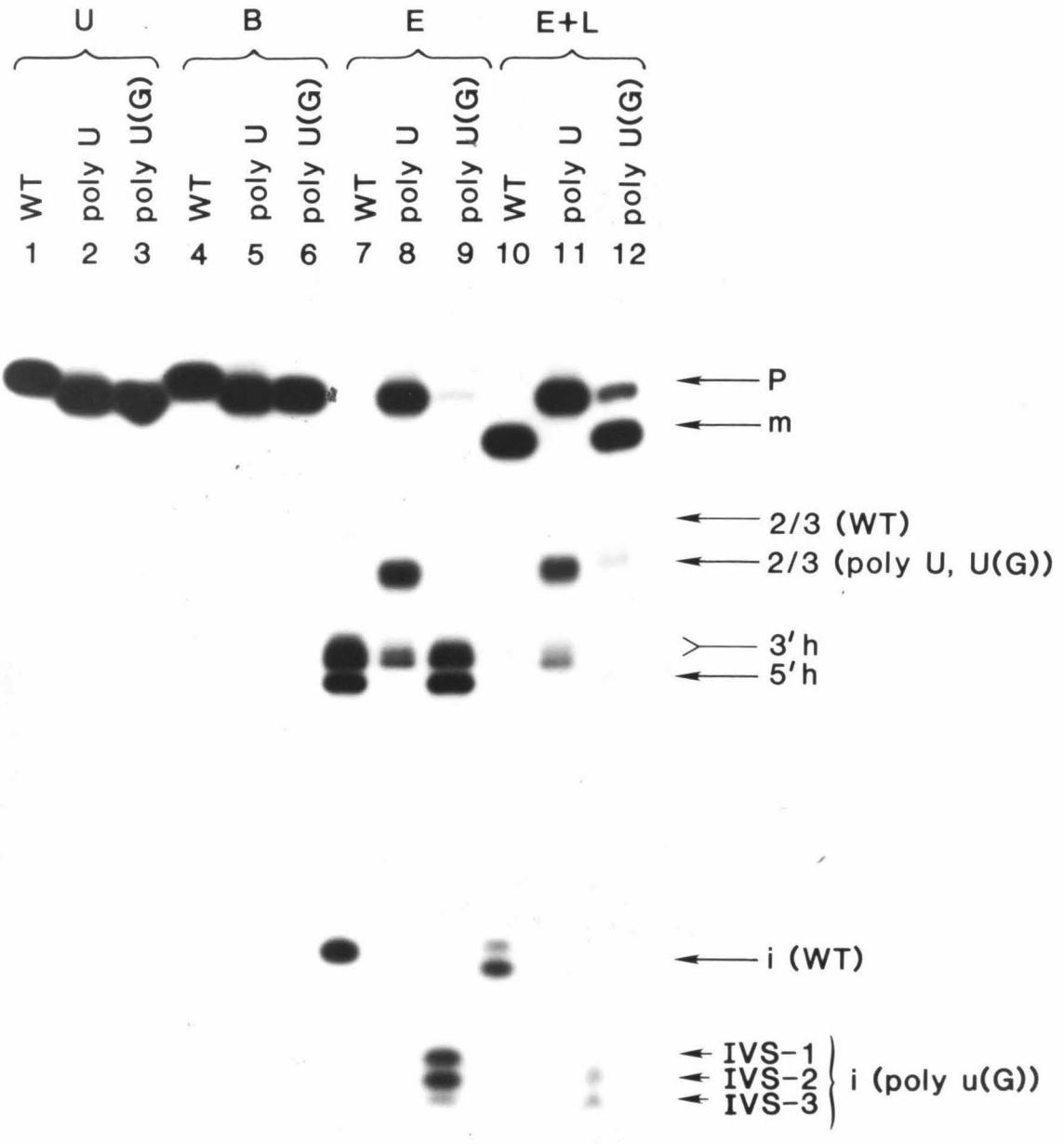
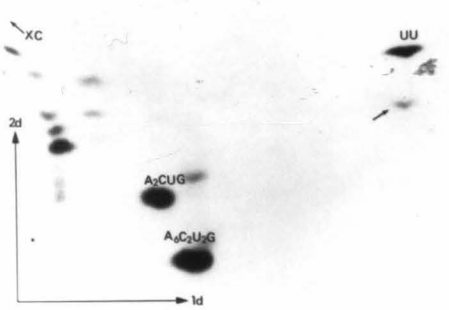
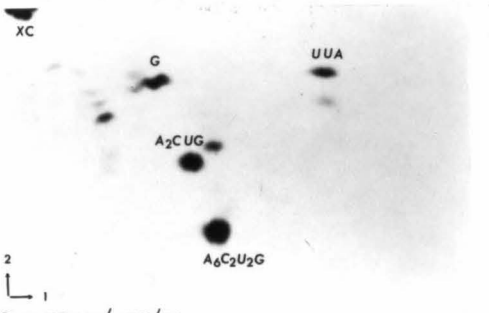

U
B
E
E+L
WT
poly U
poly U(G)
1 2 3 4 5 6 7 8 9 10 11 12
P
m
2/3 (WT)
2/3 (poly U, U(G))
3′h
5′h
i (WT)
IVS-1
IVS-2
IVS-3
i (poly u(G))

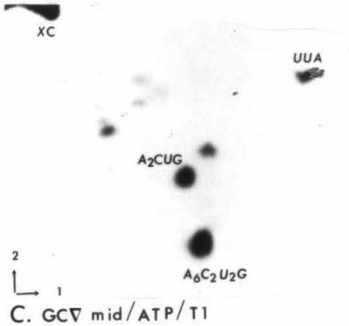
Figure 4B



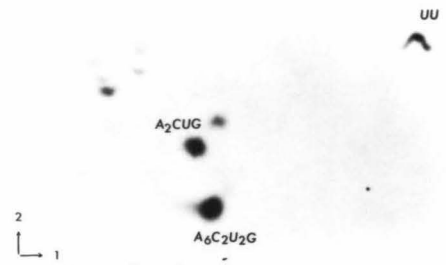
A. WT/ATP/T1



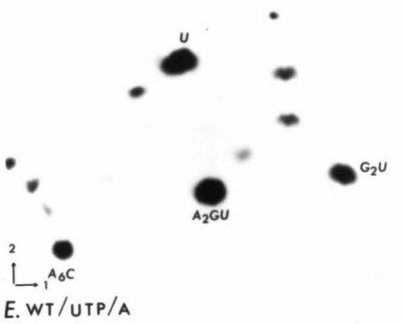
B. GC∇ top/ATP/T1



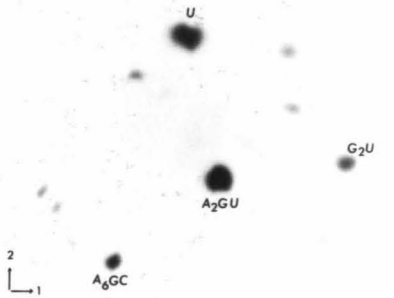
C. GC∇ mid/ATP/T1



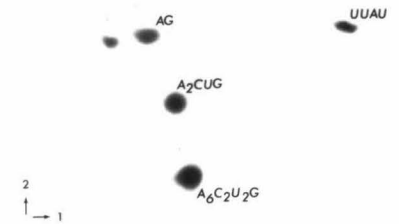
D. GC∇ low/ATP/T1



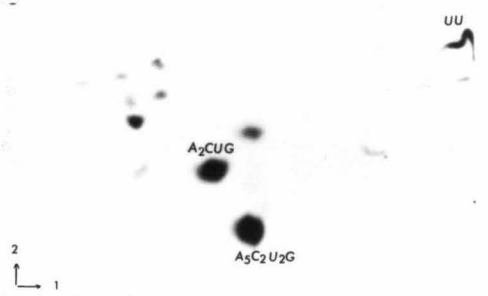
E. WT/UTP/A



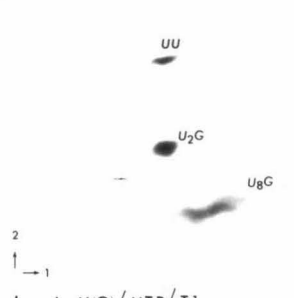
F. GC∇ top/UTP/A



G. $\binom{UA}{GC}\nabla$/ATP/T1



H. ΔU/ATP/T1



I. poly U(G)/UTP/T1

Figure 5

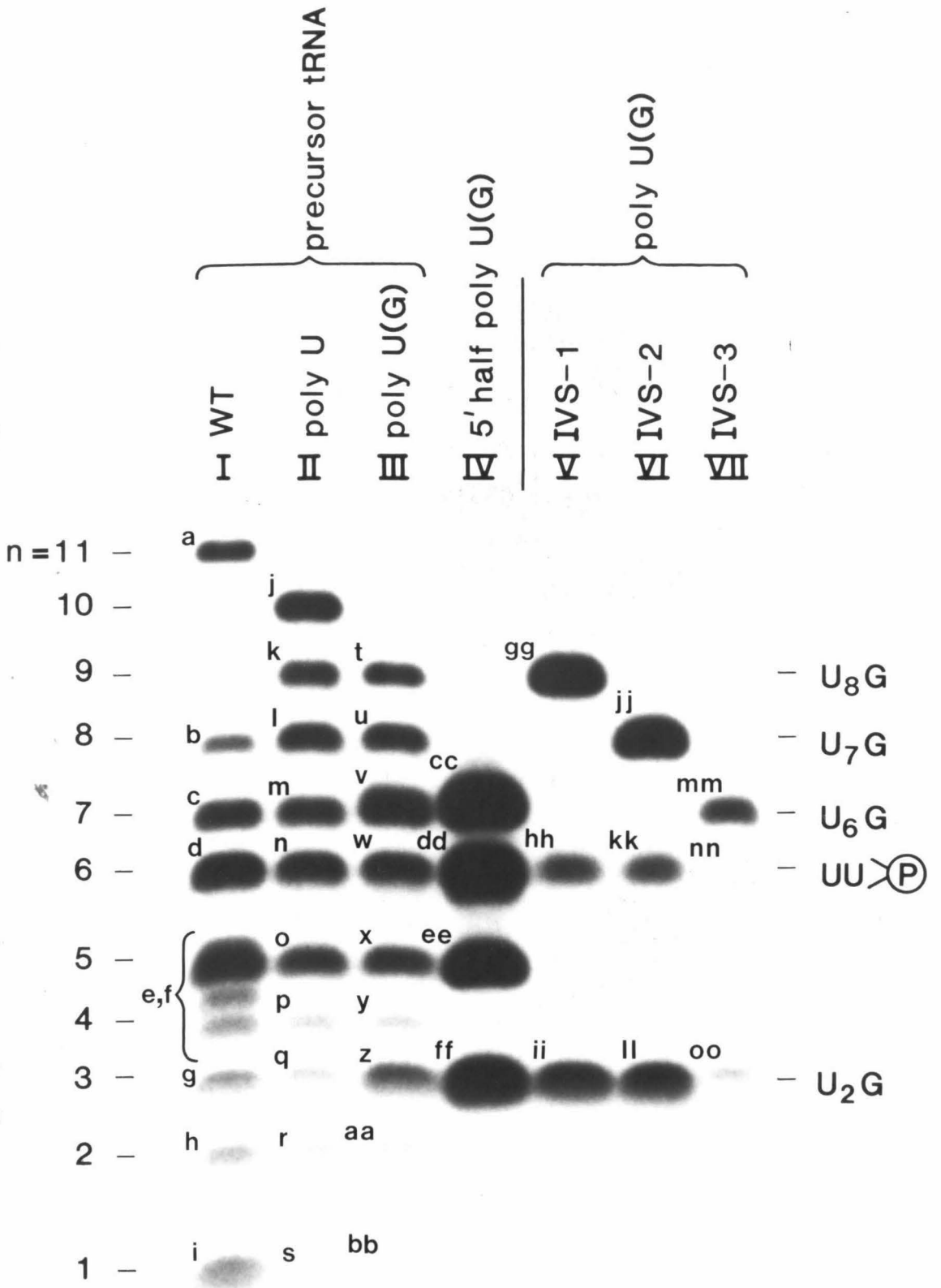
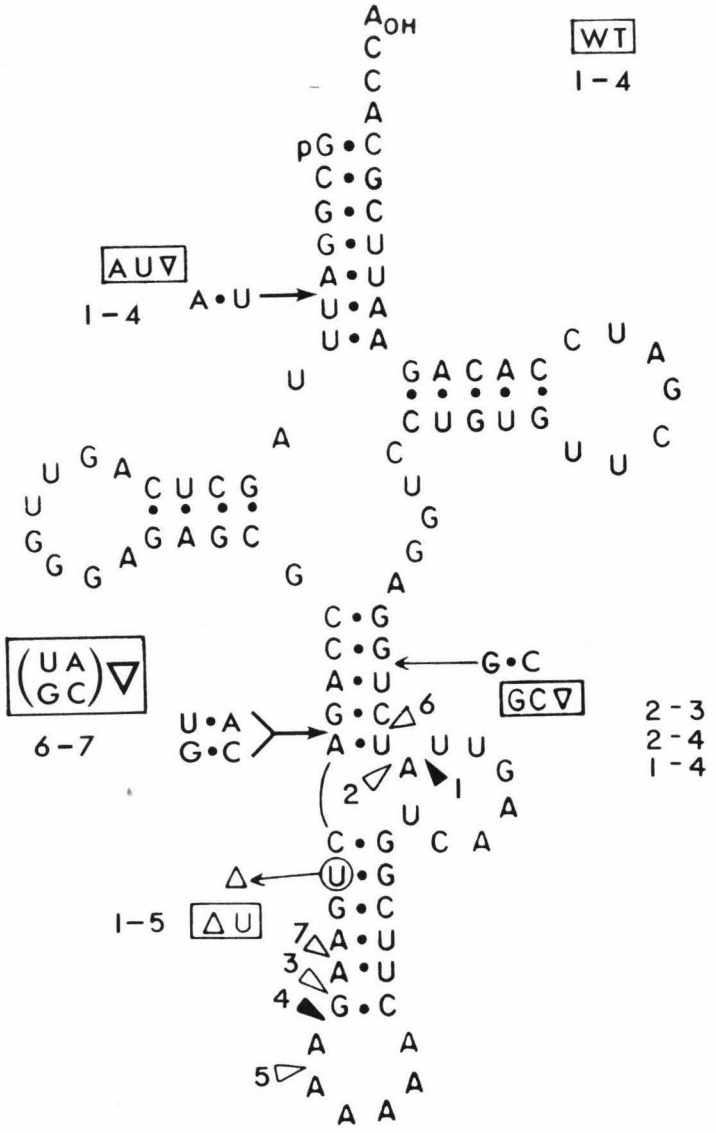
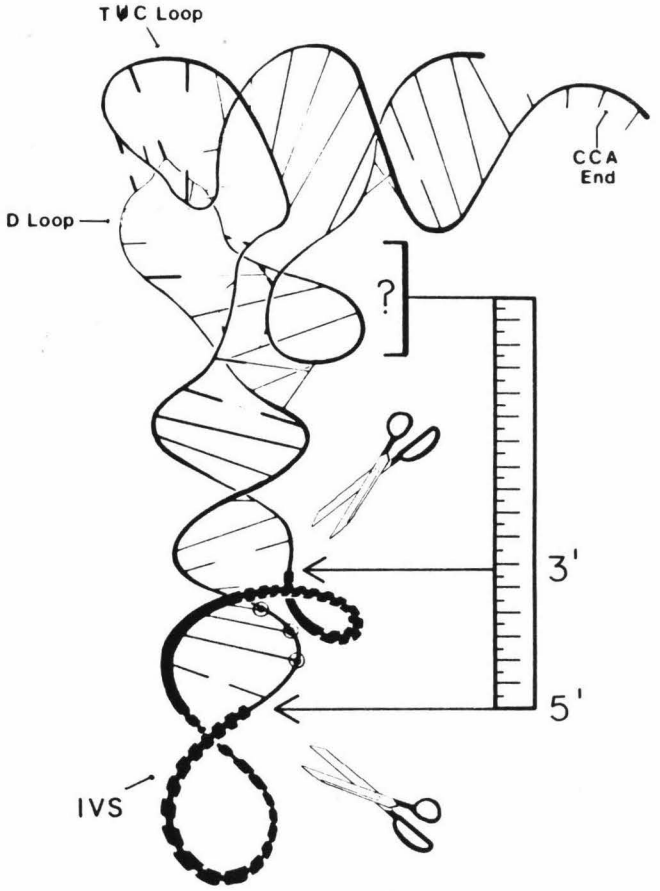

precursor tRNA
poly U(G)
I WT
II poly U
III poly U(G)
IV 5' half poly U(G)
V IVS-1
VI IVS-2
VII IVS-3
n = 11
10
9
8
7
6
5
4
3
2
1
a
j
k
t
gg
jj
l
u
b
cc
c
m
v
mm
d
n
w
dd
hh
kk
nn
o
x
ee
e,f
p
y
q
z
ff
ii
ll
oo
g
h
r
aa
i
s
bb
U8G
U7G
U6G
UU
U2G

Figure 6


WT
1-4
AU∇
1-4
A•U
(UA/GC)∇
6-7
U•A
G•C
GC∇
2-3
2-4
1-4
1-5
ΔU

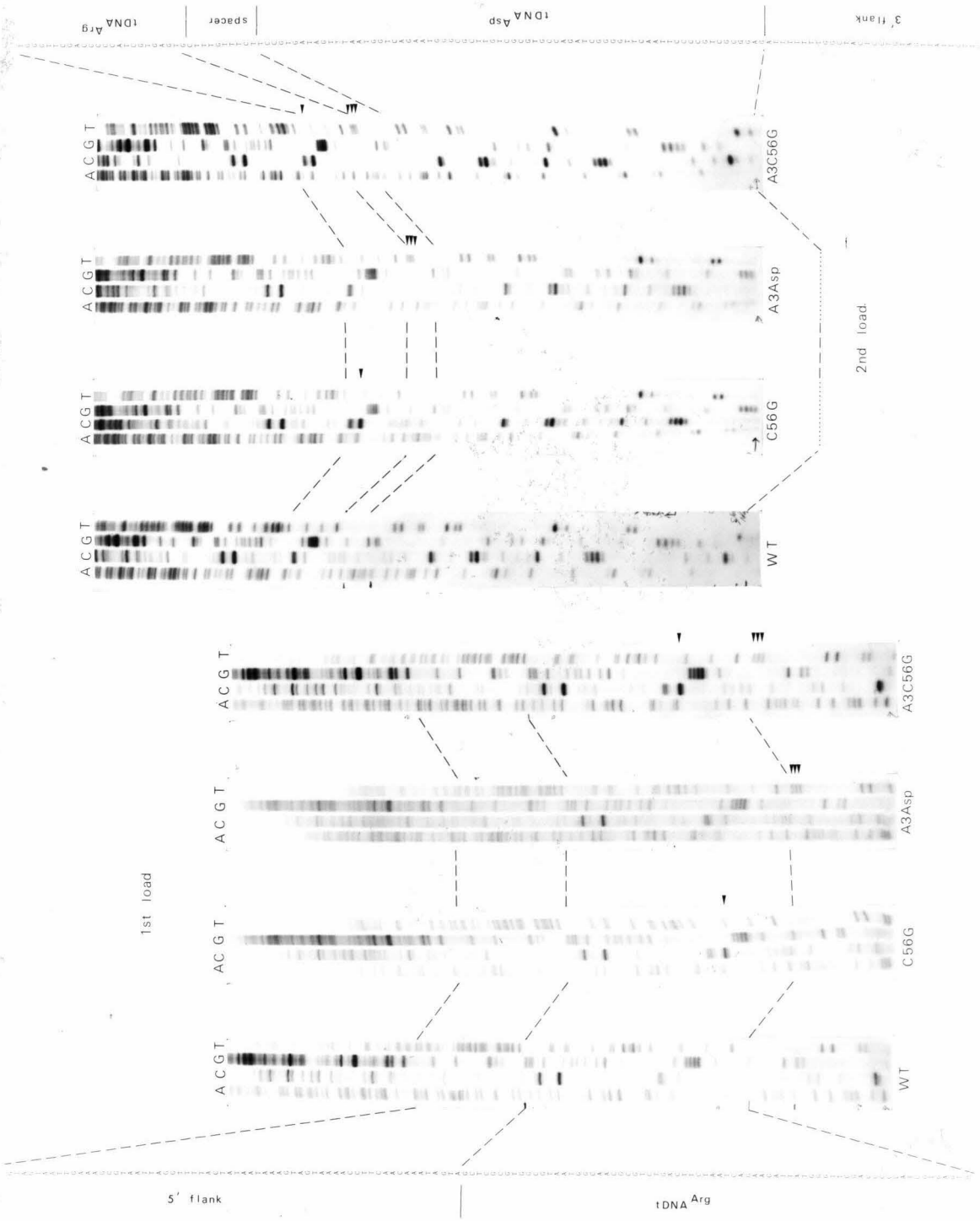
A.

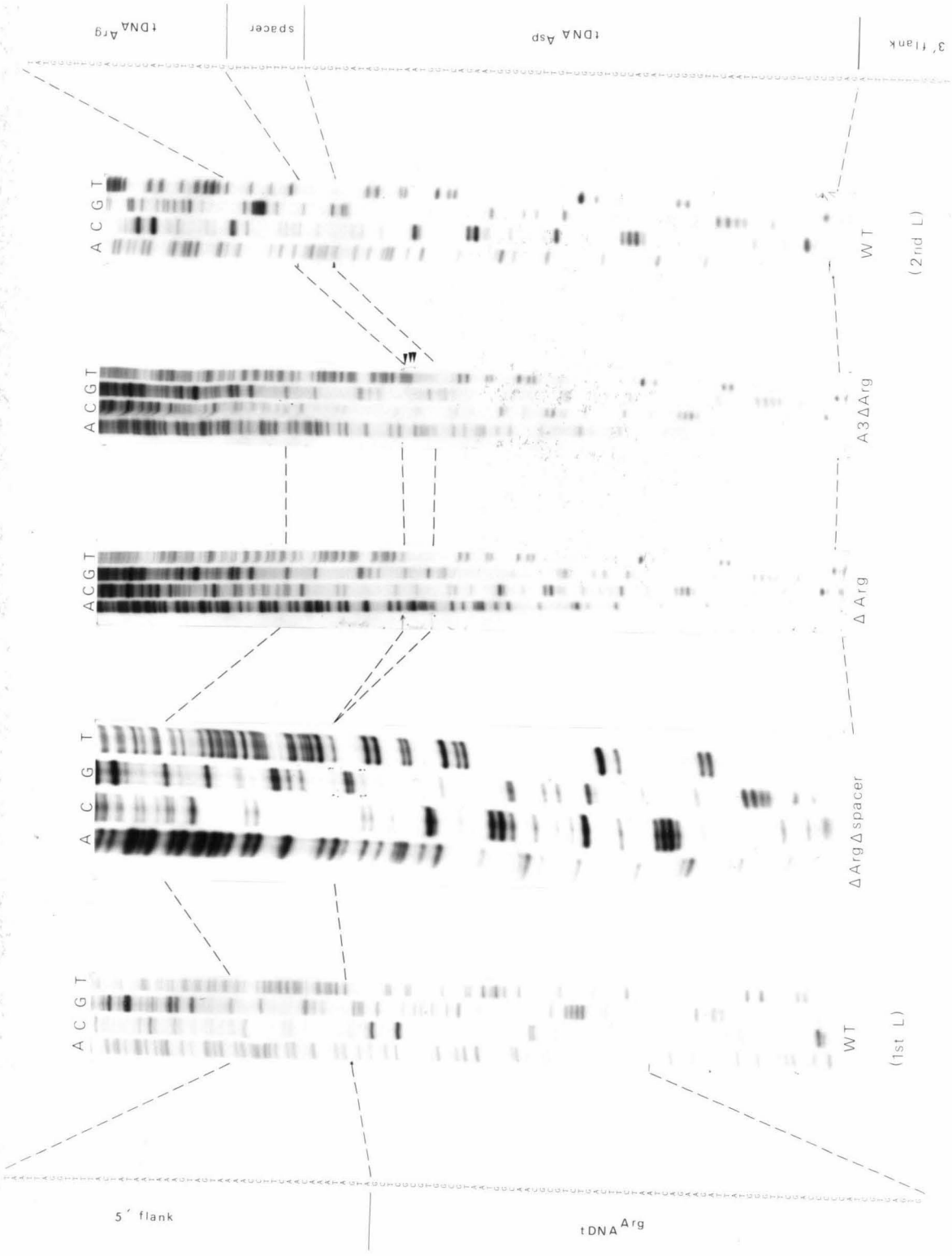

TψC Loop
CCA End
D Loop
?
3'
5'
IVS

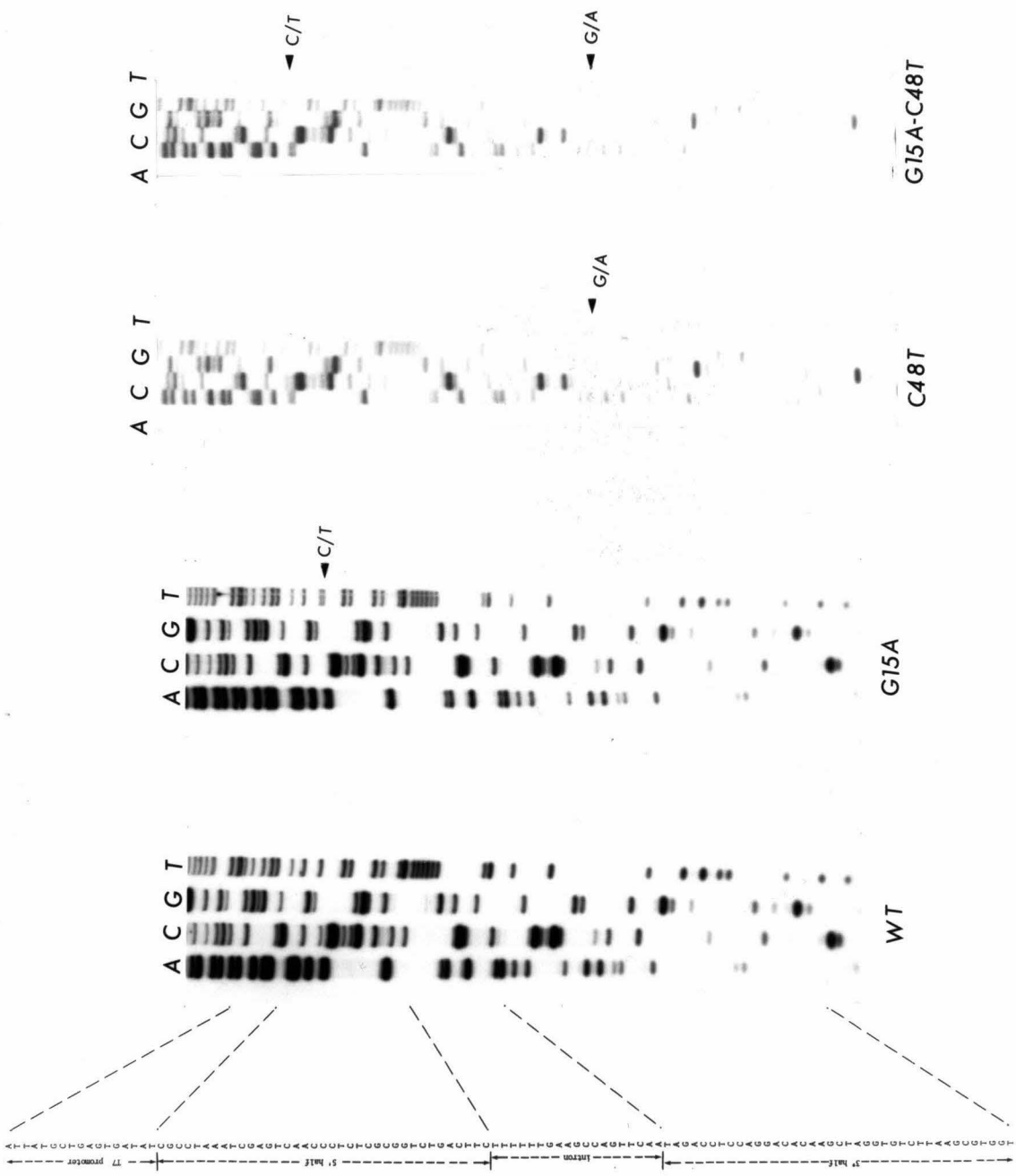
B.

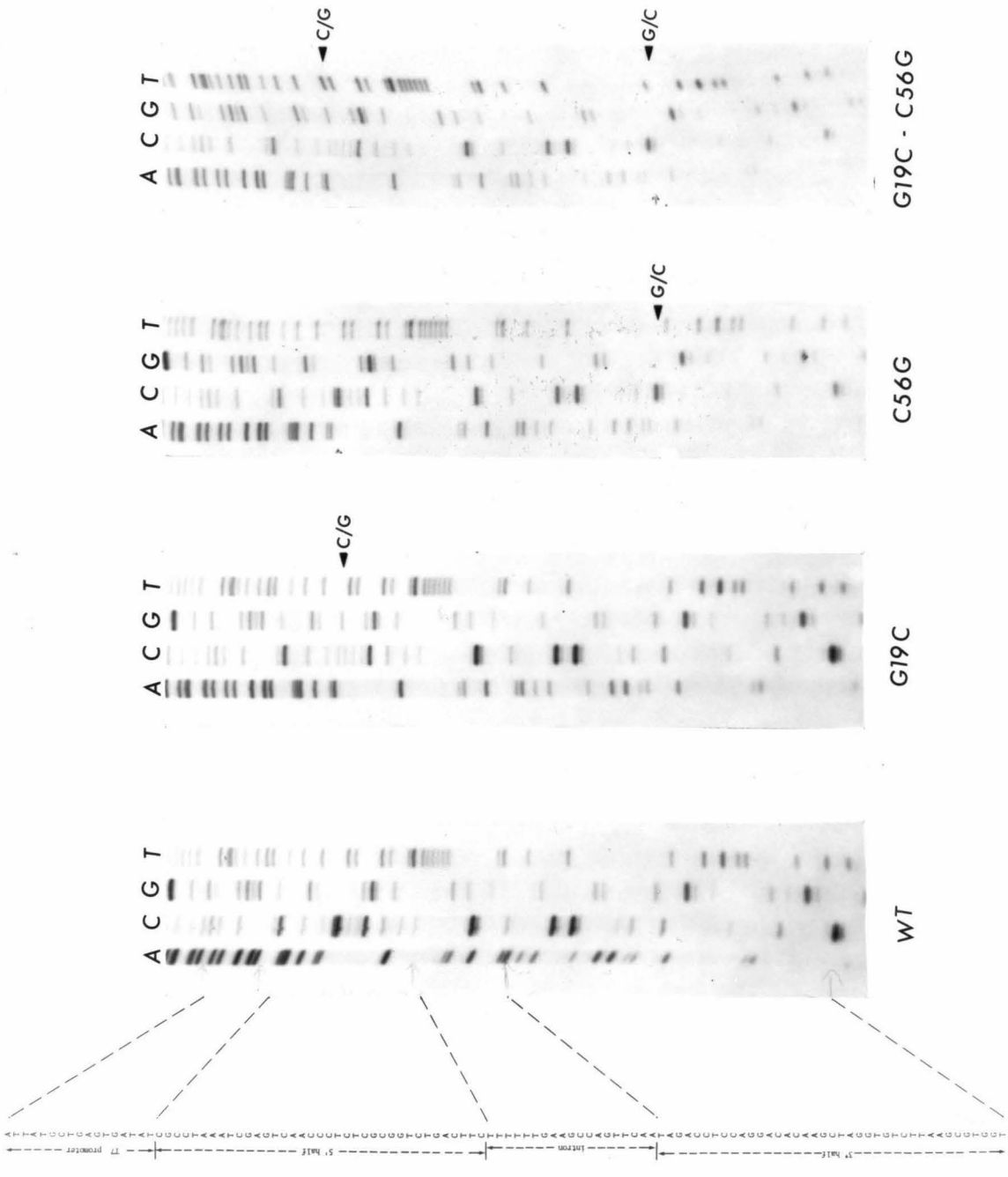
Figure 7

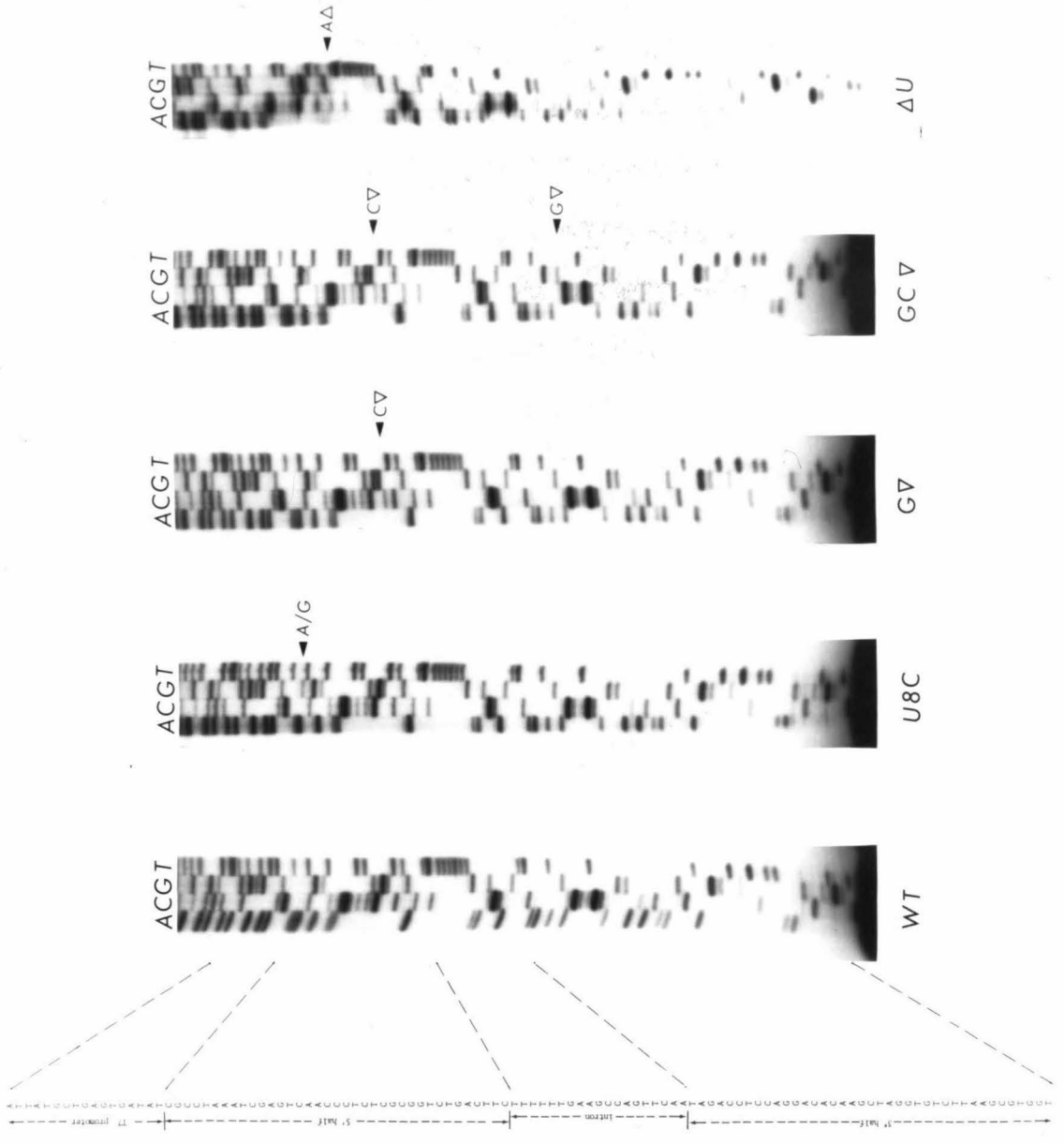
APPENDICES

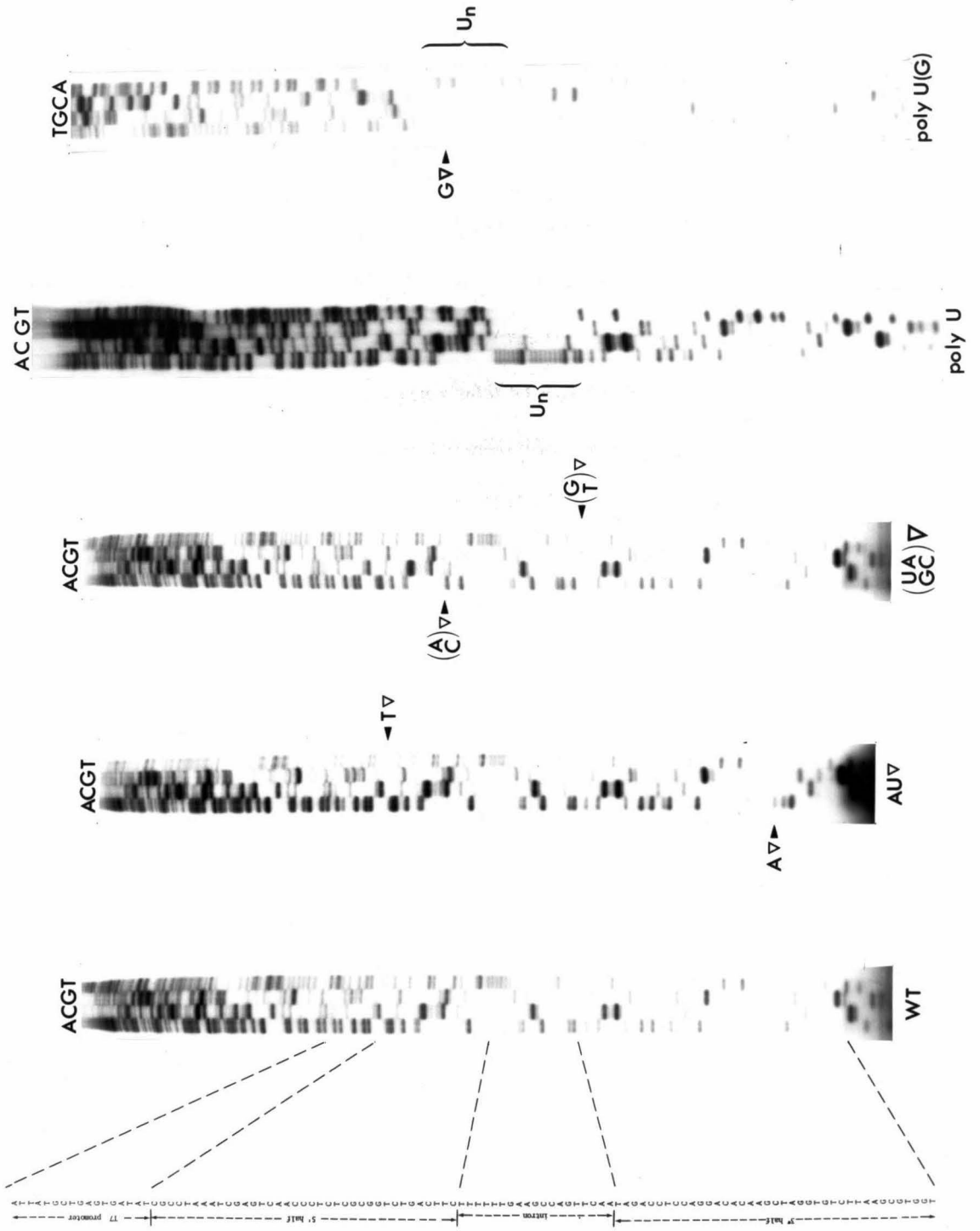

tDNA Arg
spacer
tDNA Asp
3' flank
A C G T
A3C56G
A3Asp
C56G
WT
2nd load
1st load
5' flank
tDNA Arg


tDNA Arg
spacer
tDNA Asp
3' flank
A C G T
WT
(2nd L)
A3ΔArg
ΔArg
ΔArgΔspacer
(1st L)
5' flank
tDNA Arg


C/T
G/A
A C G T
G15A-C48T
G/A
A C G T
C48T
C/T
A C G T
G15A
A C G T
WT
T7 promoter
5' half
intron
3' half


C/G
G/C
A C G T
G19C - C56G
G/C
A C G T
C56G
C/G
A C G T
G19C
A C G T
WT
T7 promoter
5' half
intron
3' half


ACGT
ΔU
GC∇
G∇
U8C
WT
A/G
T7 promoter
5' half
intron
3' half


TGCA
Un
G
poly U(G)
ACGT
Un
poly U
(G/T)
ACGT
(A/C)
(UA/GC)
T
ACGT
A
AU
ACGT
WT
17 promoter
5' half
intron
3' half

Research Proposition I

Pre-tRNA Binding and Cleavage by the *S. cerevisiae* tRNA Splicing Endonuclease

Summary

An investigation on pre-tRNA primary sequence requirement for binding to the *S. cerevisiae* tRNA splicing endonuclease is proposed. A uniform concentration of radioactively labeled wild-type pre-RNA^{Phe} is incubated in the presence of a constant but limiting amount of the endonuclease and with unlabeled mutant competitor pre-tRNA at a range of concentrations. Labeled splicing products were separated by gel electrophoresis, quantitated, and the yields compared with that of a control experiment consisting of an analogous competition between labeled and unlabeled wild-type pre-tRNA. Inhibition of cleavage of the labeled pre-tRNA is taken as a measure of binding affinity of the competitor precursor to the endonuclease relative to that of the wild-type precursor. Our preliminary results show that the U8C mutant pre-tRNA^{Phe} binds the endonuclease very poorly, whereas a poly(U)-intron variant of the same pre-tRNA binds the enzyme with at least the same affinity as does wild-type pre-tRNA. The implication of these results on endonuclease-pre-tRNA interactions are discussed.

Introduction

Splicing of tRNA in *S. cerevisiae* is a two-stage reaction mediated by two distinct and separable enzymes: an endonuclease to excise the IVS, and a ligase to join the resulting half-tRNA moelcules to produce an intact, mature tRNA. In *X. laevis*, the reaction follows a similar course, and the endonuclease has been purified and extensively characterized. For instance, it has been demonstrated that binding and cleavage of pre-tRNA by the *X. laevis* endonuclease are two distinct and separable steps; certain mutant pre-tRNAs are available which can bind to the enzyme but cannot be cleaved by it (Gandini-Attardi *et al.*, 1985). Endonuclease-pre-tRNA complex in this system has been detected by means of a native gel retardation assay.

Is this phenomenon of separable binding and cleavage steps by the endonuclease a general feature among eukaryotic tRNA splicing systems? In this laboratory, a large collection of site-specific mutant *S. cerevisiae* pre-tRNA^{Phe} gene clones is available (Reyes and Abelson, 1987b). With the availability of a quick and accurate endonuclease binding assay, this problem can be readily addressed. Greer *et al.* (1987) have briefly described an *in vitro* binding assay for the endonuclease which is essentially a standard *in vitro* splicing assay done in the absence of ligase, modified to include only a limiting amount of the endonuclease (exact amount determined in a previous set of experiments), and done in the presence of an unlabeled, competitor pre-tRNA whose binding affinity to the enzyme is being sought. In this short report, we demonstrate the feasibility of this assay in studying binding affinities of various pre-tRNAs, and show some preliminary results which suggest that binding and cleavage are distinct and separable steps in *S. cerevisiae*, as they are in *X. laevis*.

Experimental Procedures

The procedure is basically the standard *in vitro* splicing assay described previously (Reyes and Abelson, 1987a), but done in the absence of ligase (i.e., ligase diluent is added instead of ligase) and in the presence of unlabeled competitor pre-tRNA at a range of concentrations, typically ½, 1, 3, 5, and 7 times that of the radiolabeled wild-type pre-tRNA^{Phe} which, in contrast, is present at a uniform concentration. It is also important that endonuclease concentration is limiting (concentration determined previously). A control endonuclease binding assay consists of competition between labeled and unlabeled wild-type pre-tRNA^{Phe}. All subsequent competition experiments involving unlabeled mutant pre-tRNAs are compared against this control.

Preliminary Results and Discussion

The structures of the various pre-tRNA species used in this study is shown in Fig. 1. Inclusion of unlabeled wild-type competitor pre-tRNA at a gradually increasing concentration with respect to the labeled wild-type pre-tRNA results in a parallel gradual decrease in cleavage products of the latter (Fig. 2, panel A, lanes 3 to 7). At unlabeled competitor concentration 5x that of labeled precursor, inhibition is approximately 66% (panel 6); and at 7.5x competitor concentration, inhibition is almost complete (lane 7). In Fig. 2, panel B, the unlabeled competitor pre-tRNA has a U8 to C mutation (Fig. 1, panel A), and it is evident that inhibition is not detectable even at 2.5x unlabeled competitor concentration (lane 5). At 7.5x unlabeled competitor concentration, inhibition is just approximately 50% (lane 7); note again that in panel A, inhibition at this stage is complete. The same results can be seen in Fig. 2, panel C, lanes 4 to 6, where the unlabeled competitor is also the U8C mutant, present at concentrations ½x, 1.5x and 5x that of the labeled wild-type pre-tRNA. We thus conclude from these results that the U8C precursor binds

very poorly to the endonuclease; affinity of this mutant precursor to the enzyme is at most half that of wild-type pre-tRNA^{Phe}.

In Fig. 2, panel C, lanes 7 to 9, competition between wild-type and poly(U)-IVS pre-tRNA^{Phe} is shown. The structure of the poly(U)-IVS variant is shown in Fig. 1, panel B. As small as $\frac{1}{2}$x competitor concentration results in a readily detectable diminution of cleavage (lane 7). At 5x competitor concentration, inhibition is almost complete (lane 9). We conclude from these results that the poly(U)-IVS mutant binds the endonuclease very tightly, and may even have slightly greater affinity for the enzyme than does the wild-type pre-tRNA^{Phe} (compare lane 7, panel A, with lane 9, panel C).

The finding that the U8C mutant binds the endonuclease poorly and that the poly(U)-IVS mutant binds it very efficiently is not surprising considering our previous splicing results: the U8C mutant is not cleaved by the endonuclease, whereas the poly(U)-IVS precursor is cleaved efficiently, albeit only at the 3' splice site (Reyes and Abelson, 1987b). But what seems ironic from these results is the fact that the U8C mutant has only a single-base substitution, whereas in the poly(U)-IVS precursor, a whole block of 20 bases in the molecule was replaced by an entirely novel sequence, poly(U). But what must be remembered is that the alteration in the poly(U)-IVS precursor, although extensive, is at the IVS-anticodon loop region, and that in the U8C mutant is in a totally conserved position in the mature domain of the molecule. Thus, this result emphasizes the importance of the mature portion of the pre-tRNA for the recognition and binding by the various processing enzymes, and to the general belief that the IVS-anticodon region is essentially passive and unimportant for these processes.

In summary, we conclude that the above endonuclease binding assay is not only feasible, but is a quick and accurate method as well for determining

endonuclease binding affinities of various forms of pre-tRNA^{Phe} molecules, and may be used with the rest of the variant pre-tRNAs in a large collection in this laboratory (Reyes and Abelson, 1987b) to allow classification of these altered precursors as either binding-defective or cleavage-defective pre-tRNAs.

References

Gandini-Attardi, D., Margarit, I., and Tocchini-Valentini, G. P. (1985) Structural alterations in mutant precursors of the yeast $tRNA_3^{Leu}$ gene which behave as defective substrates for a highly purified splicing endonuclease. **EMBO J.** *4*, 3289-3297.

Greer, C. L., Söll, D., and Willis, I. (1987) Substrate recognition and identification of splice sites by the tRNA splicing endonuclease and ligase from *Saccharomyces cerevisiae*. **Mol. Cell. Biol.** *7*, 76-84.

Reyes, V. M. and Abelson, J. (1987a) A synthetic substrate for tRNA splicing. **Anal. Biochem.** *166*, 90-106.

Reyes, V. M. and Abelson, J. (1987b) Substrate recognition and splice site selection by the *S. cerevisiae* tRNA splicing endonuclease. **Cell,** submitted for publication.

Figure Legends

Figure 1. Secondary cloverleaf structures of wild-type and U8C mutant pre-tRNA^{Phe} (panel A) and the poly(U)-IVS pre-tRNA (panel B).

Figure 2.

Panel A. Competition between labeled an unlabeled wild-type pre-tRNA^{Phe} for the endonuclease. Group I are the controls: lane 1 is unincubated wild-type precursor; lane 2, wild-type precursor incubated with buffers only; and lane 3, wild-type precursor with endonuclease, in the absence of competitor. Group II are the competition experiments: lanes 4, 5, 6 and 7: presence at concentrations of 1x, 2.5x, 5x and 7.5x, respectively, of unlabeled wild-type competitor with respect to labeled pre-tRNA. Bands a, b, c, d and e are the precursor, 2/3 molecule, 3' half, 5' half and IVS, respectively.

Panel B. Competition between labeled wild-type pre-tRNA^{Phe} and unlabeled U8C mutant for the endonuclease. This experiment was done exactly as in panel A, except that the unlabeled competitor in group II is the U8C mutant pre-tRNA^{Phe}.

Panel C. Competition between labeled wild-type pre-tRNA^{Phe} with U8C pre-tRNA (Group II) and poly(U)-IVS pre-tRNA (Group III). Group I are the controls and are exactly the same as the corresponding lanes in panels A and B. Group II are the competition experiments between labeled wild-type precursor and unlabeled U8C precursor: lanes 4, 5 and 6: presence of $\frac{1}{2}$x, 1.5x, and 5x, respectively, of unlabeled U8C mutant competitor. Group III are the competition experiments between labeled wild-type precursor and the poly(U)-IVS precursor: concentrations of competitor is the same as the corresponding lanes in Group II, except that the poly(U)-IVS competitor is labeled; this is permissible in this case because the wild-

type pre-tRNA^{Phe} and poly(U)-IVS precursors have dissimilar sizes and migrate differently; they and their splicing products (except 3' half) are separately detectable as well for the same reason. Band identities are as follows: a = pre-tRNA^{Phe}; b = poly(U)-IVS precursor; c, d, and i = most probably degradation products of the poly(U)-IVS precursor; e and f = 2/3 molecule of pre-tRNA^{Phe} and poly(U)-IVS precursor, respectively; g = wild-type and poly(U)-IVS 3' half; h and i = wild-type 5' half and IVS, respectively.

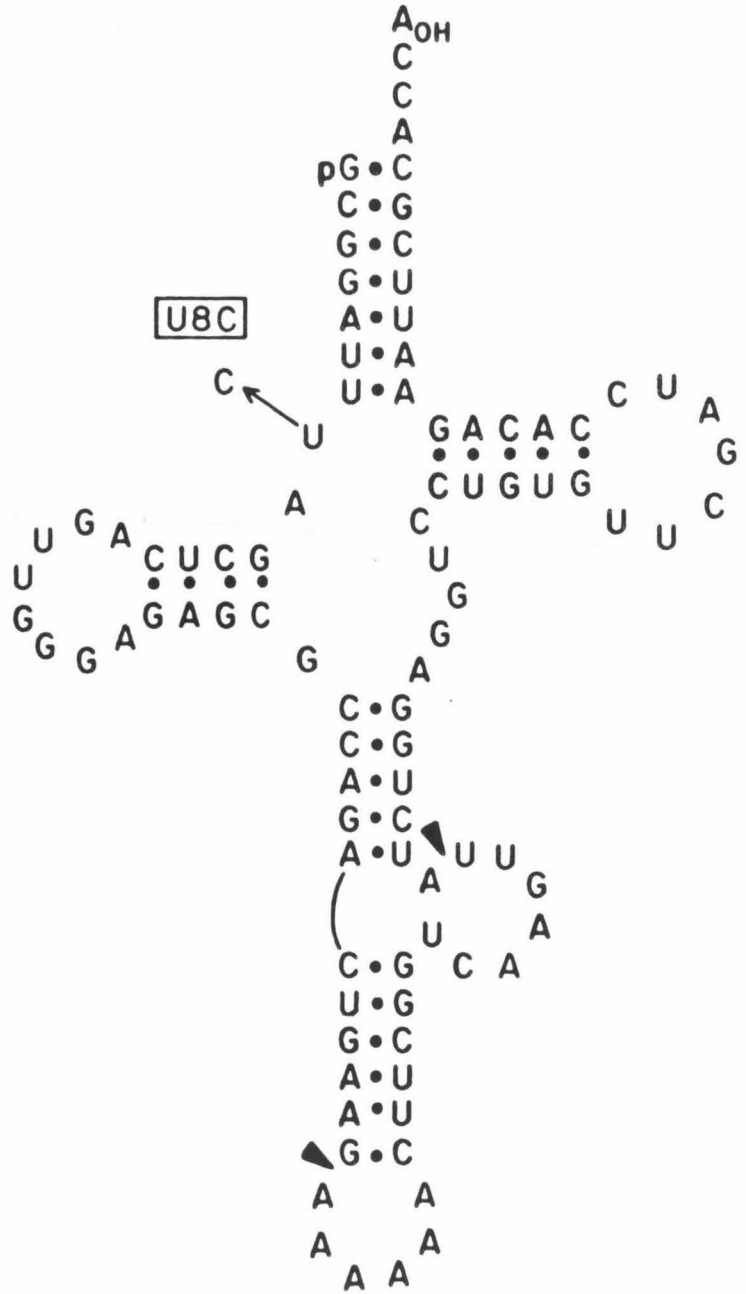


Figure 1A

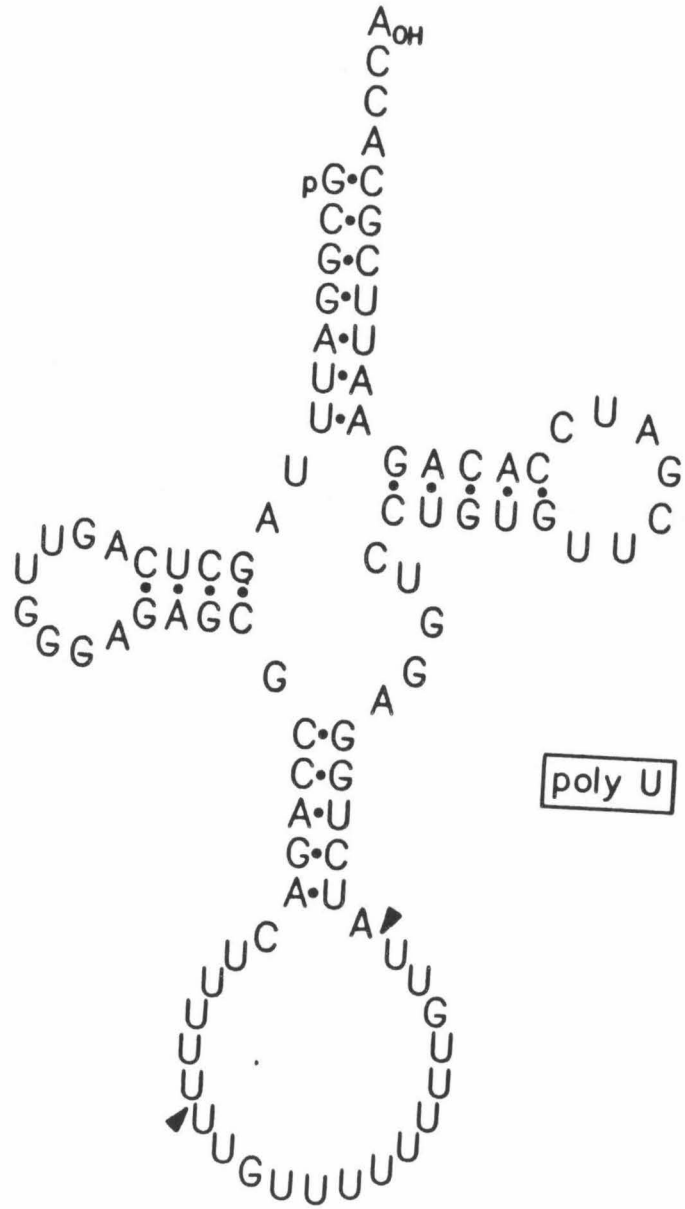


Figure 1B

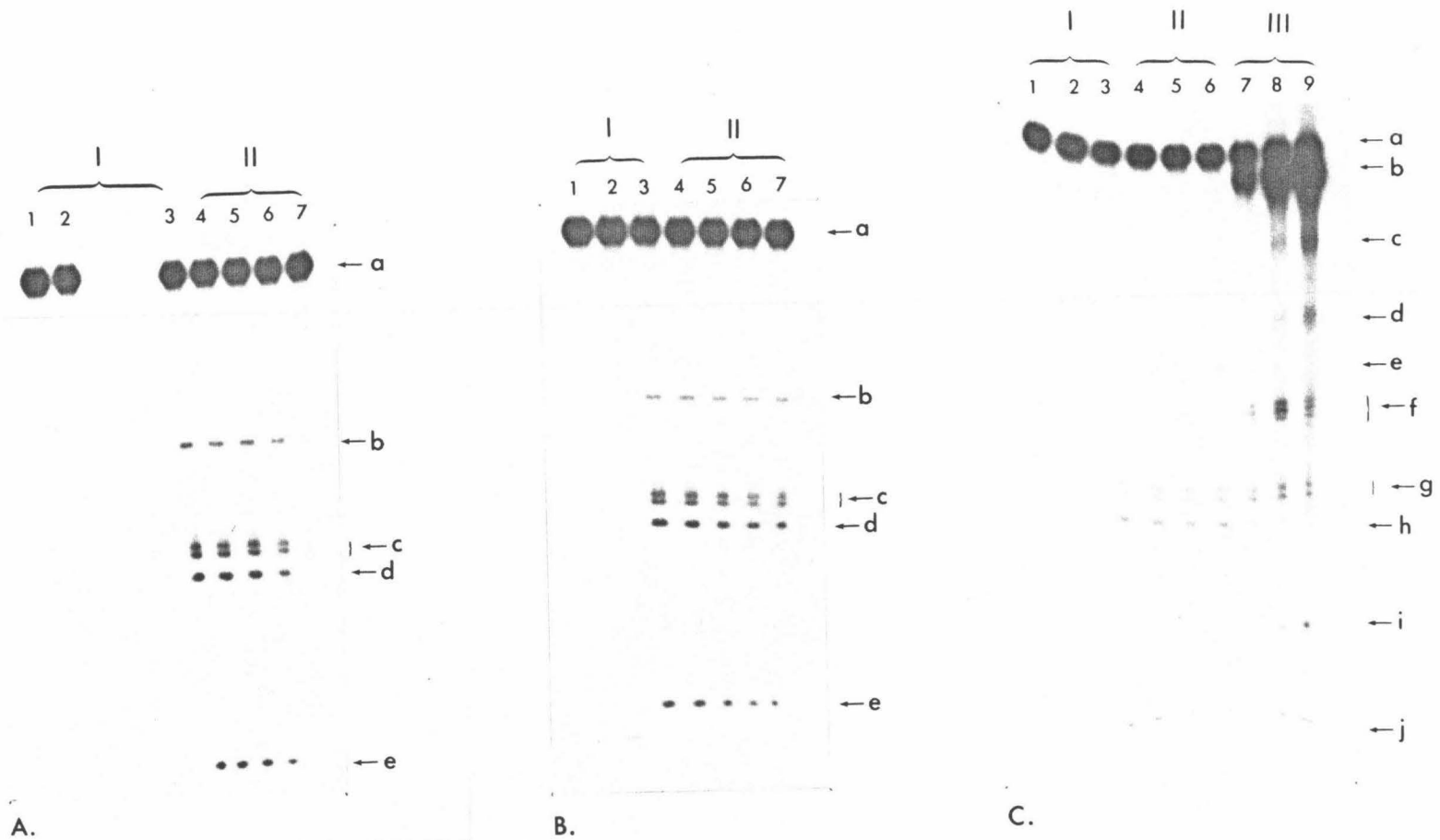


Figure 2

Research Proposition II

Approximate Structure-Probing of RNA Molecules By Plumbous Ion-Mediated RNA Cleavage

Summary

A quick method to approximately compare tertiary structures of RNA molecules is presented. Use is made of the known property of Pb^{++} ions to specifically cleave RNA molecules on the basis of higher-order structure as opposed to primary sequence. This method is applied to a *S. cerevisiae* pre-$tRNA^{Phe}$ and two of its variants which are known to be defective in splicing. A control experiment involving Pb^{++}-induced cleavage of mature $tRNA^{Phe}$, a well-characterized reaction, is also run in parallel. Our preliminary results suggest that two single-base substitution mutants of pre-$tRNA^{Phe}$, U8C and C56G, are cleaved with the same efficiency and specificity as the wild-type precursor, and that these three tRNA species are cleaved at the same site as mature $tRNA^{Phe}$, suggesting that all four RNA molecules have approximately the same tertiary conformation. The general applicability of this technique for probing RNA structures is discussed.

Introduction

The main function of DNA and RNA in biological systems is to code for genetic information. However, one fundamental difference between these two nucleic acids is the structural role bestowed as well by nature upon RNA, but not upon DNA. There are a multitude of examples which demonstrate that the structure of RNA--its precise three-dimensional molecular contour--is crucial for its biological function. Thus, structure-probing analyses are of prime importance in studies on structure and function of RNA molecules. For example, the structure of both mature and precursor $tRNA^{Phe}$ of *S. cerevisiae* have been successfully analyzed using a battery of chemical, enzymatic, and physical techniques (for a review, see Rich and RajBhandary, 1976).

Plumbous ions have been known for some time to depolymerize RNA (Werner *et al.*, 1976; Brown *et al.*, 1985). In the case of tRNAs, for example, it was demonstrated that cleavage by this ion is usually specific between U16 and U17 and between U17 and G18 positions. Although some cleavages are unique for certain tRNA species (e.g., $tRNA^{Asp}$ is cleaved between residues 35 and 36, but not $tRNA^{Phe}$), it is generally believed that cleavage of RNA by Pb^{++} ions is determined by the overall 3-dimensional conformation of the RNA molecule, and not by any specific primary sequence it may contain. In the light of the above information, we propose to investigate the possibility of employing Pb^{++}-induced cleavage as an approximate, but general, probe for higher-order structure of RNA molecules. As a preliminary test, we turn to two members of a large collection of site-specific mutant pre-$tRNA^{Phe}$ molecules available in this laboratory (Reyes and Abelson, 1987). Our preliminary results show that a C56-to-G and a U8-to-C mutant of pre-$tRNA^{Phe}$ are cleaved with the same specificity and efficiency as does wild-type precursor and mature $tRNA^{Phe}$, suggesting that these four tRNA species all have similar tertiary structures.

Experimental Procedures

The reaction conditions for Pb^{++}-mediated RNA cleavage consists of 2 μM tRNA species; 15 mM $MgCl_2$; 1.5 mM spermine or spermidine; and 15 mM MOPS (morpholino propane-sulfonic acid) buffer, pH = 7.0. Pb^{++} is added as $Pb(OAc)_2$ at a range of concentrations, usually from 1 to 8 mM. Incubation is at room temperature for 15 to 20 min. A 4x buffer containing 60 mM each of MOPS buffer, pH = 7.0, and $MgCl_2$, and 6 mM of spermine (or spermidine) is usually convenient to use. The reaction is stopped by adding gel loading dye containing 15 mM EDTA and 8.0 M urea. The tRNA species may be reannealed before the reaction by heating to 65°C for 5 min., and then slowly cooling to room temperature; however, this is not absolutely necessary. Cleavage products are separated by gel electrophoresis in 10% polyacrylamide/8.0 M urea gel. $Pb(OAc)_2$ is dissolved in distilled water and stored in small aliquots as 0.5 to 1.0 M solutions; each aliquot is discarded after every use to avoid concentration changes due to reaction of Pb^{++} with atmospheric CO_2, which would result in the precipitation of $PbCO_3$. The MOPS buffer may be prepared as a 0.1 M stock solution in H_2O, adjusted to pH = 7.0 with NaOH.

Preliminary Results and Discussion

Cleavage of wild-type, U8C, and C56G pre-$tRNA^{Phe}$ and of mature $tRNA^{Phe}$ (structures shown in Fig. 1) generates a large and a small fragment. The large fragment from the three pre-tRNAs appears to be of the same size and is just slightly larger than uncleaved mature tRNA (compare band b with band c, Figure 2). The large fragment generated from mature $tRNA^{Phe}$ (band d) is shorter than the corresponding fragment generated from the three pre-tRNAs. Strikingly, however, the small fragment generated from all four tRNA species (the 3

precursors and mature tRNA) are all of the same size (band e; band f is probably the salt front). This result strongly suggests that the site of cleavage is occurring at the same position relative to a terminus of the RNA, presumably 5', in each case.

At first glance, cleavage appears more efficient for wild-type precursor and mature $tRNA^{Phe}$ than for the two mutants, U8C and C56G (compare intensities of band e for all three groups I, II and III). However, by liquid scintillation counting (data not shown) of the uncleaved RNA bands (bands a or b) and the small fragment (bands e), it was found that the ratio of uncleaved to small fragment is the same in all four cases. Although the same amount of total input cpm is present in all lanes in Figure 2, there is apparently more nonspecific background cpm (reasons are unclear, but may be due to labeled polyphosphates, etc.) present in the U8C and C56G pre-tRNA lanes (lanes 2, 3, 6, 7, 10 and 11), hence less total tRNA species is actually being added. Further, as the concentration of Pb^{++} ions is increased from 4 mM in Group I to 8 mM in group III, the cleavage efficiency (ratio of species a or b to species c) and accuracy (size of fragment in band e) remains essentially the same.

Taken together, these results suggest that accuracy (exact site) and efficiency (cpm ratio of uncleaved to small fragment) of cleavage of these four tRNA species by Pb^{++} is indistinguishable, and we conclude that they have very similar three-dimensional configurations. These results agree very well with NMR structural data which indicate that pre-$tRNA^{Phe}$ and $tRNA^{Phe}$ have the same tertiary structure, the difference lying only in the IVS-anticodon loop region (C. Hall, pers. comm.). Further, although the U8C and C56G mutant precursors have not yet been directly analyzed structurally, these results indicate that they have the same tertiary structure as the wild-type precursor. Since both mutant precursors only have a single base change involving the U8-A14 and C56-G19

tertiary base pairs, respectively, it is not surprising that the tertiary structures have not been drastically altered since tRNAs have nine tertiary base pairs and apparently, the other 8 such pairs suffice to stabilize and preserve the tertiary structure, not to mention the stability resulting from the extensive base stacking interactions in the molecule as well.

Incidentally, both U8C and C56G mutant precursors are cleavage-defective (Reyes and Abelson, 1987). Thus, in the light of the present structural results, it may be concluded that the splicing defect of both mutants is due to primary sequence alteration *per se*, and not to higher-order structural changes (i.e., tertiary conformation).

In conclusion, we would like to suggest that the above Pb^{++}-induced cleavage of RNA is useful as a general method of approximate structure-probing of RNA molecules, provided there is a standard RNA species to which all variants can be compared, corresponding to mature $tRNA^{Phe}$ in the present case. Although this procedure still needs further optimization to eliminate the extensive degradation, (see Fig. 2) for example by inclusion of RNasin in the incubation buffer, (which has been shown not to affect even splicing of pre-tRNAs), we think it will be a useful procedure for the study of RNA structure and function.

References

Brown, R. S., Dewan, J. C., and Klug, A. (1985) Crystallographic and biochemical investigation of the lead(II)-catalyzed hydrolysis of yeast phenylalanine tRNA. **Biochem.** *24*, 4785-4801.

Reyes, V. M. and Abelson, J. (1987) Substrate recognition and splice site selection by the *S. cerevisiae* tRNA splicing endonuclease. **Cell,** submitted for publication.

Rich, A. and RajBhandary, U. L. (1976) Transfer RNA: Molecular structure, sequence, and properties. **Ann. Revs. Biochem.** *45*, 805-860.

Werner, C., Krebs, B., Keith, G., and Dirheimer, G. (1976) Cleavages of pure tRNAs by plumbous ions. **Biochem. Biophys. Acta** *432*, 161-175.

Figure Legends

Figure 1. Secondary cloverleaf structures of wild type, U8C and C56G pre-$tRNA^{Phe}$ molecules.

Figure 2. Plumbous ion-mediated cleavage of wild-type pre-$tRNA^{Phe}$ (lane 1, 5 and 9), and its U8C (lanes 2, 6 and 10) and C56G (lanes 3, 7 and 11) variants, and of mature $tRNA^{Phe}$ (lanes 4, 8 and 12). Groups I, II and III represent increasing concentrations of Pb^{++} in the incubation mixture: Group I, 4 mM Pb^{++}; Group II, 6 mM Pb^{++}; and Group III, 8 mM Pb^{++}. Band a are uncleaved wild-type, U8C or C56G precursors, which are of the same size; band b is the large fragment generated by Pb^{++}-induced cleavage of these three precursors; band c is uncleaved mature $tRNA^{Phe}$; band d is the large fragment generated from c and band e represents the small fragment , all of the same size, generated by Pb^{++} cleavage from all four tRNA species. Band f is the putative salt front. Position indicated as "o" is the origin of electrophoresis.

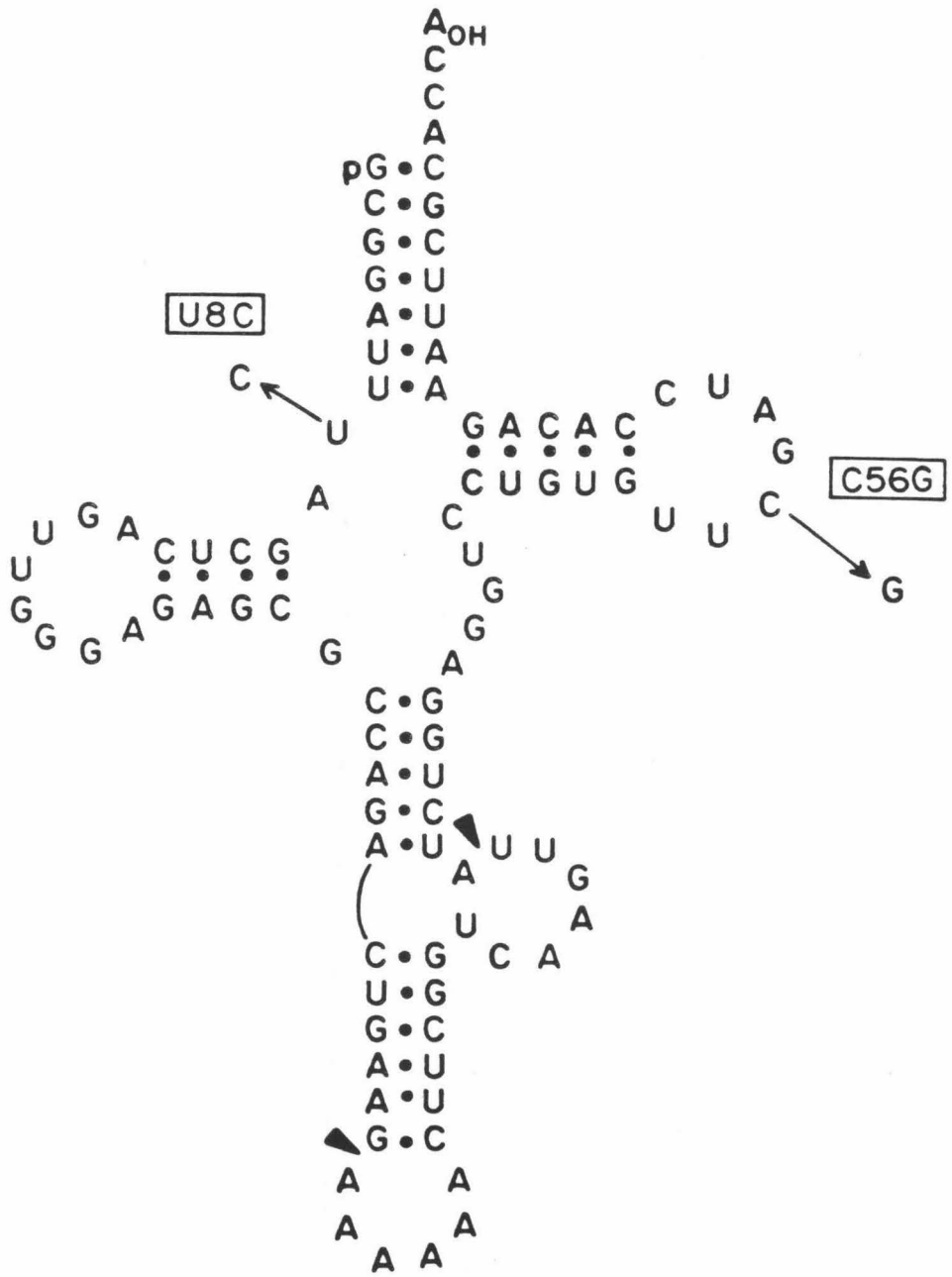


Figure 1

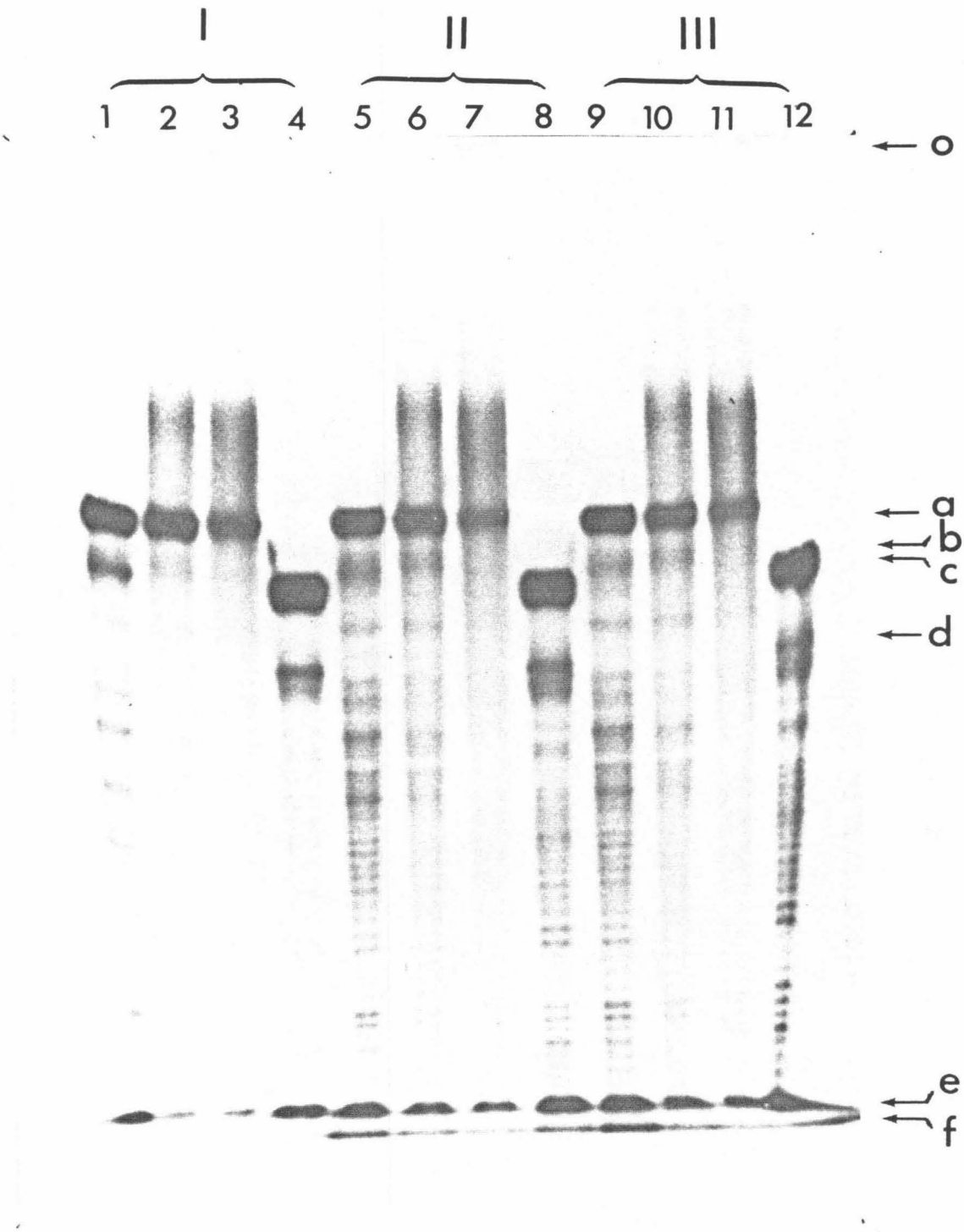


Figure 2

Research Proposition III

Two-Dimensional Oligoribonucleotide Mapping of RNA Fragments Generated by RNAse U_2

Summary

A two-dimensional oligonucleotide mapping of RNA fragments generated by digestion with RNAse U_2 is presented. These RNA fragments, except the 3' terminal one, end in A. Our preliminary results show that the two-dimensional map is similar, but not identical, to one generated by RNAse T_1 digestion fragments. The slight differences between the two maps are described. The usefulness of this mapping technique in routine RNA sequence analyses is also discussed.

Introduction

RNA sequence analysis is essential to all investigations pertaining to structure, function and interaction of RNA. Presently, there exist two general methods of RNA sequence analysis. Both methods require prior digestion of a labeled RNA fragment with base-specific chemical or enzymatic reagents, but one involves separation of uniformly labeled digestion fragments in two dimensions producing a pattern of spots upon autoradiography called a map, or fingerprint (Volckaert *et al.*, 1976; Volckaert and Fiers, 1977), while the other involves separation of end-labeled fragments in one dimension by gel electrophoresis, producing a sequencing ladder equivalent to those generated by the Maxam-Gilbert DNA sequencing method (Donis-Keller, 1980). In this preliminary investigation, we focus on the first method and propose a two-dimensional map for a new family of RNA fragments. Previously, the labeled RNA to be analyzed is digested either with RNAse T_1 or RNAse A, and families of fragments ending either with G or with U and C (except the 3' terminal fragments) are generated. In this report, we investigate the possibility of using RNAse U_2 to generate fragments ending in A, and determine the two-dimensional map of this family of oligoribonucleotides. We also discuss the possibility of employing this procedure as a general method of RNA sequence analysis.

Experimental Procedures

The complete procedure is identical to previously published protocols for two-dimensional RNA oligonucleotide mapping (Volckaert *et al.*, 1976; Volckaert and Fiers, 1977; Reyes and Abelson, 1987) except that the enzyme used is RNAse U_2 instead of RNAse T_1 or RNAse A. The reading chart for an RNAse U_2 fingerprint is presented in Fig. 2.

Preliminary Results and Discussion

Two-dimensional fingerprint analysis of RNA fragments generated from α-^{32}P-UTP-labeled pre-tRNA^{Phe} by RNAse U_2 reveals that a map similar, but not identical, to one expected for an RNAse T_1 fingerprint is obtained. (Fig. 1, panel A). In the latter, an (n + 1)-mer family of fragments migrate to form a triangle pointing down, with C_nG, A_nG, and U_nG forming the three corners (Fig. 2A). In an RNAse U_2 family, a similar triangle is obtained, only this time the corners of the triangle correspond to C_nA, G_nA and U_nA, respectively (fragments migrating within the triangle are altered accordingly). G migrates slightly faster than A in the first dimension, so the lower vertex of the triangle is now nearer U_nA than C_nA (Fig. 2, panel B), unlike in an RNAse T_1 fingerprint where the corresponding corner is nearer C_nG than U_nG (Fig. 2, panel A).

Comparison of panel A with panel B of Fig. 1 shows that the fingerprint is somewhat cleaner if the digestion with RNAse U_2 is carried out at the lower pH of 3.8, suggesting that this enzyme has a greater specificity for cleavage after A residues at this pH.

Although the present method still needs further optimization as to digestion conditions to achieve greater specificity as well as refinement of the 2-dimensional map presented in Fig. 2, we deem it useful as a general method of RNA sequence analysis. It is particularly useful in mapping terminal or internal fragments that cannot possibly be uniquely labeled if RNAse T_1 or RNAse A are used, and thus would otherwise necessitate tedious secondary digestion analyses.

References

Donis-Keller, H. (1980) Phy M: An RNase activity specific for U and A residues useful in RNA sequence analysis. **Nucl. Acids Res.** *8*, 3133-3142.

Reyes, V. M. and Abelson, J. (1987) Substrate recognition and splice site selection by the *S. cerevisiae* tRNA splicing endonuclease. **Cell,** submitted for publication.

Volckaert, G., and Fiers, W. (1977) Micro thin-layer techniques for rapid sequence analysis of ^{32}P-labeled RNA: Double digestion and pancreatic ribonuclease analyses. **Anal. Biochem.** *83*, 228-239.

Volckaert, G., Min Jou, W., and Fiers, W. (1976) Analysis of ^{32}P-labeled bacteriophage MS2 RNA by a mini-fingerprinting procedure. **Anal. Biochem.** *72*, 433-446.

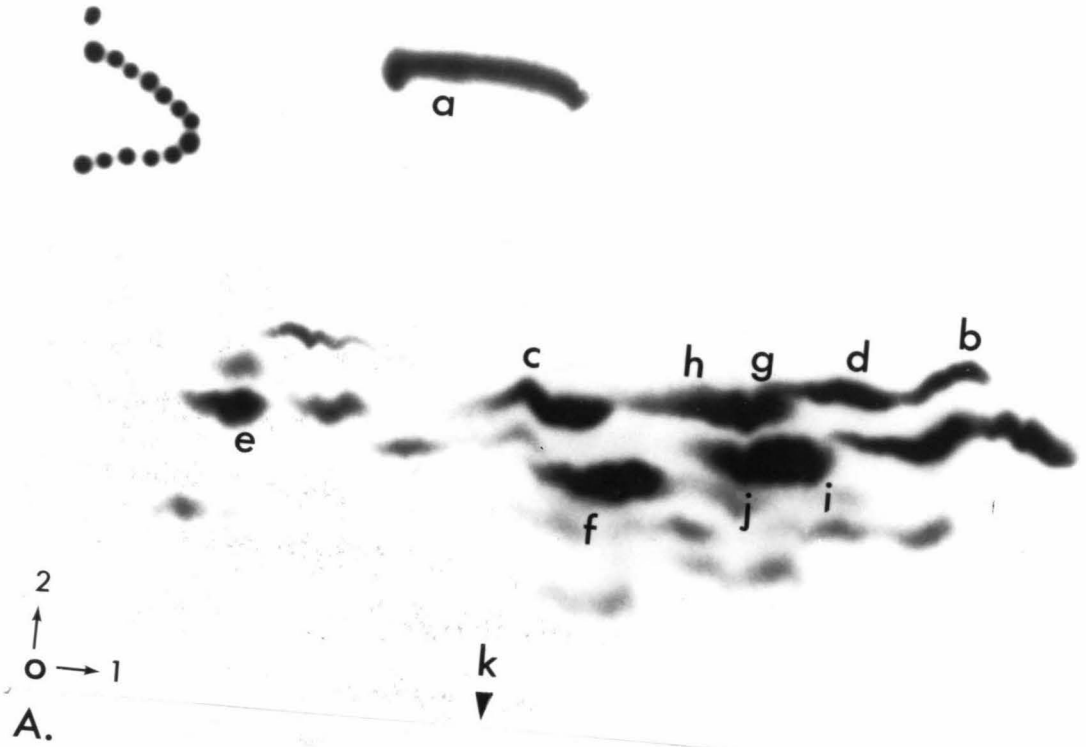
Figure Legends

Figure 1, Panel A. Two dimensional oligonucleotide mapping of RNAse U_2 fragments of α-^{32}P-UTP-labeled wild-type pre-tRNA^{Phe}. Digestion with RNAse U_2 was done in plain TE (10 mM Tris, pH 7.3/1 mM EDTA). The origin (o) and the first and second dimensions (electrophoresis at pH = 3.5 and thin layer chromatography in polyethyleneimine plates, respectively) are identified. The RNAse U_2-generated fragments are identified by letters as follows: a = A; b = U_3A; c = CGUA; d = GU_2A; e = pCG_3A (5' terminus); f = C_2GUA; g = CG_2U_2A; h = C_2GU_2A; i = G_4U_2A; j = $C_3G_2U_3A$; and k = $C_3G_5U_5A$. Fragment k, containing five guanine and five uridine residues, probably streaked very badly all over the bottom part of the plate and is not visible (uridine and guanine residues induce streaking in the first and second dimensions, respectively). The dotted region on the upper left hand corner is the xylene cyanol marker dye. Unidentified spots are all presumed to be either degradative or partial digestion products.

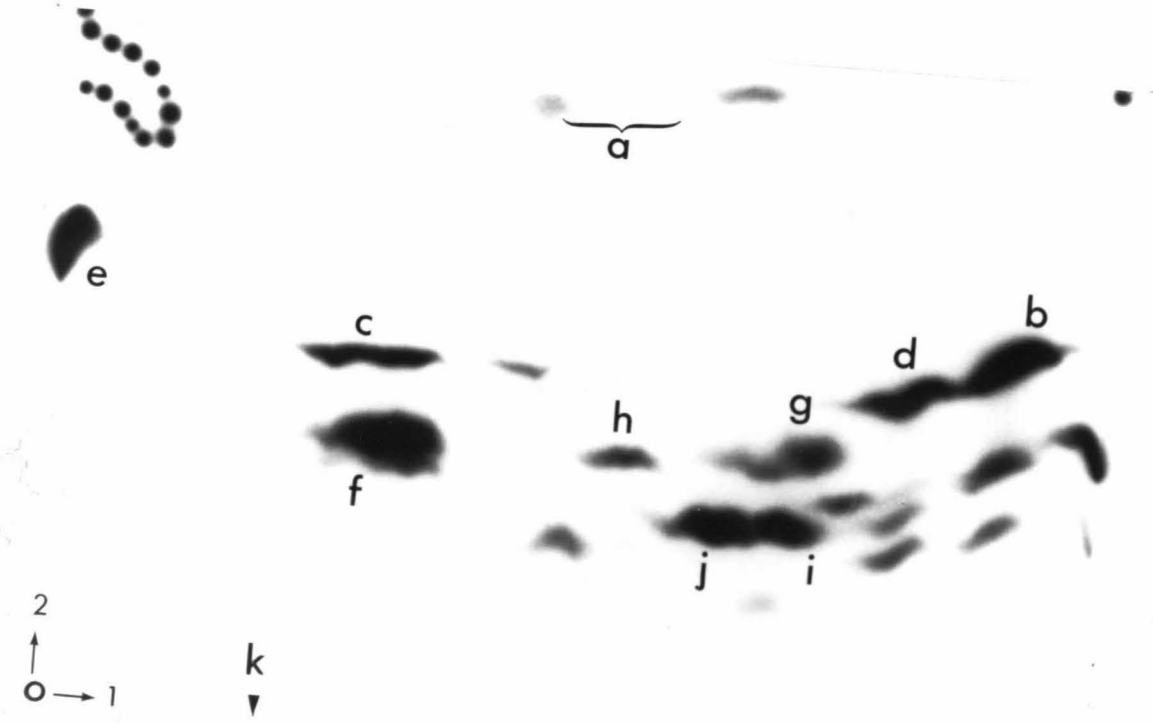
Panel B. This experiment is exactly the same as that in panel A, except that digestion with RNAse U_2 was done in TE solution buffered at pH 3.8, and the first dimension was not run very far as in panel A.

Figure 2. Oligoribonucleotide map of RNAse T_1-generated (panel A) and RNAse U_2-generated (panel B) RNA fragments. These two maps both apply to (n+1)-mers; shorter oligonucleotides migrate faster, and longer oligonucleotides, slower, and would therefore form a corresponding triangle above and below, respectively, the ones shown here.

Figure 1

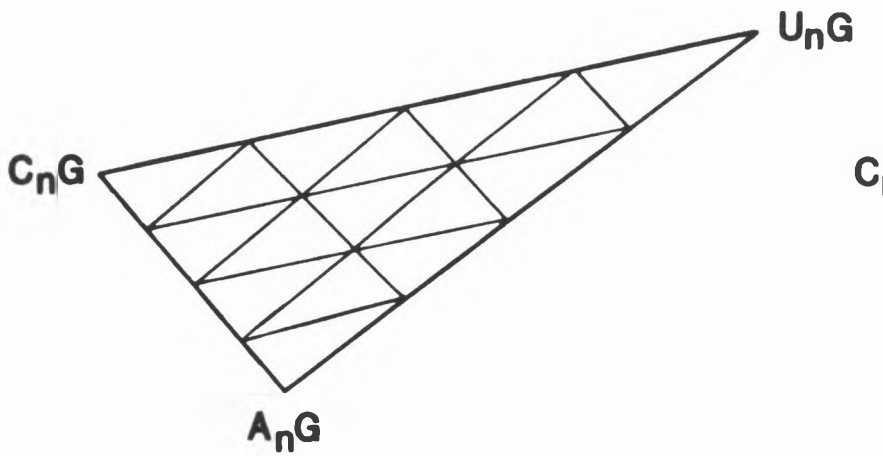


A.



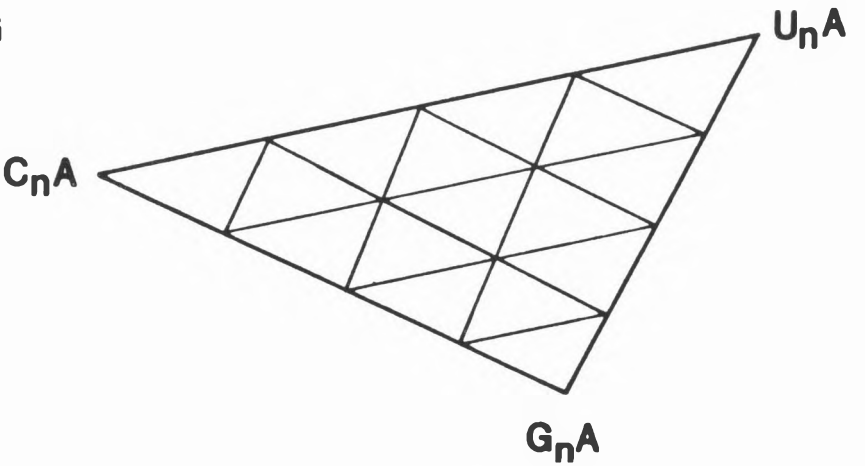
B.

Figure 1



RNase T_1

Figure 2A



RNase U_2

Figure 2B

Research Proposition IV

Testing for Gene Positional Effects on the Transcription of the *S. cerevisiae* $tRNA^{Arg}$ - $tRNA^{Asp}$ Gene Tandem

Summary

A method to switch the positions of the $tRNA^{Arg}$ and $tRNA^{Asp}$ genes in the *S. cerevisiae* $tRNA^{Arg/Asp}$ gene tandem is presented. By oligonucleotide-directed mutagenesis, *Sca*I and *Eco*RV recognition sequences will be introduced into the 5' end of the $tRNA^{Arg}$ gene and the 3' end of the $tRNA^{Asp}$ gene, respectively, in the plasmid M13ArgAsp (Reyes *et al.*, 1986). These two restriction endonucleases were chosen because they do not cleave in M13mp8 and the introduction of their recognition sequences cause the least perturbation of the flanking sequences. Cleavage of this modified M13ArgAsp plasmid with *Sca*I and *Eco*RV will precisely excise the $tRNA^{Arg/Asp}$ gene pair (and spacer sequence). Self-assembly of 13 synthetic oligodeoxynucleotides (Table I and Fig. 1) generates the $tRNA^{Asp/Arg}$ gene tandem, and cloning of this synthetic DNA fragment into the *Sca*I/*Eco*RV-cleaved M13ArgAsp plasmid gives the desired plasmid, M13AspArg.

Introduction

In *S. cerevisiae*, tRNA genes are generally highly dispersed in the genome and each is transcribed as a monomeric transcriptional unit. A striking exception is the $tRNA^{Arg}$ - $tRNA^{Asp}$ gene tandem separated only by 10 bp (Schmidt *et al.*, 1980). This gene tandem is known to be cotranscribed to produce a dimeric precursor, which is then converted into the two mature tRNAs by RNA processing (Engelke *et al.*, 1985), a scenario reminiscent of prokaryotic systems. It has been shown that it is the $tRNA^{Arg}$ gene ICRs that direct this cotranscription, and that the $tRNA^{Asp}$ gene ICRs are inactive in this arrangement (Kjellin-Sträby *et al.*, 1984). By precise deletion of the $tRNA^{Arg}$ gene and spacer sequences, the $tRNA^{Asp}$ gene has been fused to the 5' flanking region of the $tRNA^{Arg}$ gene, and in this construct, the $tRNA^{Asp}$ gene can be transcribed (Reyes *et al.*, 1986). Thus, the $tRNA^{Asp}$ gene ICRs are capable of directing transcription, and that the inactivity of the $tRNA^{Asp}$ gene in the tandem arrangement is merely due to an apparent "inappropriate" 5' flanking region, consisting of the $tRNA^{Arg}$ gene and spacer sequences.

Can the $tRNA^{Arg}$ gene be transcribed independently if the $tRNA^{Asp}$ gene and spacer sequences are put in front of it? Or will a tRNA gene in the immediate downstream flanking region of a second tRNA gene be automatically turned off, and instead be automatically cotranscribed with the upstream gene, regardless of the identities of the genes, and hence of relative internal promoter strengths? In this report, we describe a quick procedure by which the $tRNA^{Arg}$ and $tRNA^{Asp}$ genes in the yeast $tRNA^{Arg/Asp}$ gene tandem can be precisely switched, and thus enable us to directly address the above questions.

Proposed Experiments and Discussion

Briefly, two unique restriction sites will be introduced into the plasmid M13ArgAsp by oligonucleotide-directed mutagenesis. One of the sites will be precisely at the 5' end of the $tRNA^{Arg}$ gene, and the other, precisely at the 3' end of the $tRNA^{Asp}$ gene. Cleavage of this modified M13ArgAsp plasmid with the two restriction endonucleases will precisely excise the $tRNA^{Arg/Asp}$ gene tandem (including the spacer sequence between them), leaving behind the natural 5' and 3' flanking sequences. Then, a set of self-complementary synthetic oligodeoxynucleotides will be cloned into this gap to generate the inverted gene tandem, $tRNA^{Asp/Arg}$.

In the M13ArgAsp plasmid, the *S. cerevisiae* $tRNA^{Arg/Asp}$ gene tandem plus both flanking regions (119 bp of 5' and 45 bp of 3' flanking sequence) is cloned into the *Sma*I site of M13mp8 (Reyes *et al.*, 1986). The 5' end of the $tRNA^{Arg}$ gene consists of the sequence (5')ACAAATAGTA↓GCTCGCGTGG(3'), and the 3' end of the $tRNA^{Asp}$ gene is (5')CGTCGCGGAG↓ATTTTTTTGG(3'). The aim is to convert each of these sequences into the recognition sequence of a restriction endonuclease which does not cleave in M13mp8, so that this conversion will give rise to two unique restriction sites in M13ArgAsp. The gene tandem can thus be precisely excised without disturbing the two flanking sequences. There are at least 34 restriction endonucleases which do not cleave M13mp8 and inspection of the recognition sequence of each reveals that *Sca*I, which cleaves in (5')AGT↓ACT(3'), is closest in sequence to the 5' end of the $tRNA^{Arg}$ gene, and *Eco*RV, cleaving at (5')GAT↓ATC(3'), is closest to that of the 3' end of the $tRNA^{Asp}$ gene. By oligonucleotide-directed mutagenesis, the two 20-base oligonucleotides, S and E (Table I) can be used to create the *Sca*I and *Eco*RV sites mentioned above. Cleavage of this modified M13ArgAsp plasmid with *Sca*I and *Eco*RV releases the $tRNA^{Arg/Asp}$ gene pair (plus spacer sequence) and leaves behind the 5' and 3'

flanking sequences with the vector sequences. Self-assembly of 13 synthetic oligonucleotides (Table I) and cloning of the resulting blunt-ended synthetic DNA fragment into the *Sca*I/*Eco*RV-linearized M13ArgAsp plasmid gives the desired plasmid, M13AspArg (Fig. 1). M13AspArg is identical to M13ArgAsp in all respects except that the two tRNA genes have been switched in position, and that there is a GTA → AGT change at positions -3 to -1 in the 5' flank, and a T → C change at position +3 in the 3' flank. These minor changes in the flanking regions are due to the introduction of the *Sca*I and *Eco*RV sites earlier. It is unlikely that these changes would cause drastic transcriptional effects on the resulting gene tandem since the first change is just a simple permutation of the three nucleotides, and the second, a single transition. If there are any template effects at all, we think they would be negligible.

Finally, to determine the transcriptional phenotype of this altered gene tandem, a crude yeast nuclear extract (YNE) described previously (Reyes *et al.*, 1986) and which contains pol III and necessary transcription factors, will be employed.

References

Engelke, D. R., Gegenheimer, P., and Abelson, J. (1985) Nucleolytic processing of a $tRNA^{Arg}$-$tRNA^{Asp}$ dimeric precursor by a homologous component from *S. cerevisiae*. **J. Biol. Chem.** *260*, 1271-1279.

Kjellin-Sträby, K. S., Engelke, D. R., and Abelson, J. (1984) Homologous *in vitro* transcription of linear DNA fragments containing the $tRNA^{Arg}$-$tRNA^{Asp}$ gene pair from *Saccharomyces cerevisiae*. **DNA** *3*, 167-171.

Reyes, V. M., Newman, A. J. and Abelson, J. (1986) Mutational analysis of the coordinate expression of the yeast $tRNA^{Arg}$-$tRNA^{Asp}$ gene tandem. **Mol. Cell. Biol.** *6*, 2436-2442.

Schmidt, O., Mao, J., Ogden, R., Beckmann, J., Sakano, H., Abelson, J. and Söll, D. (1980) Dimeric tRNA precursors in yeast. **Nature (London)** *287*, 750-752.

Figure and Table Legends

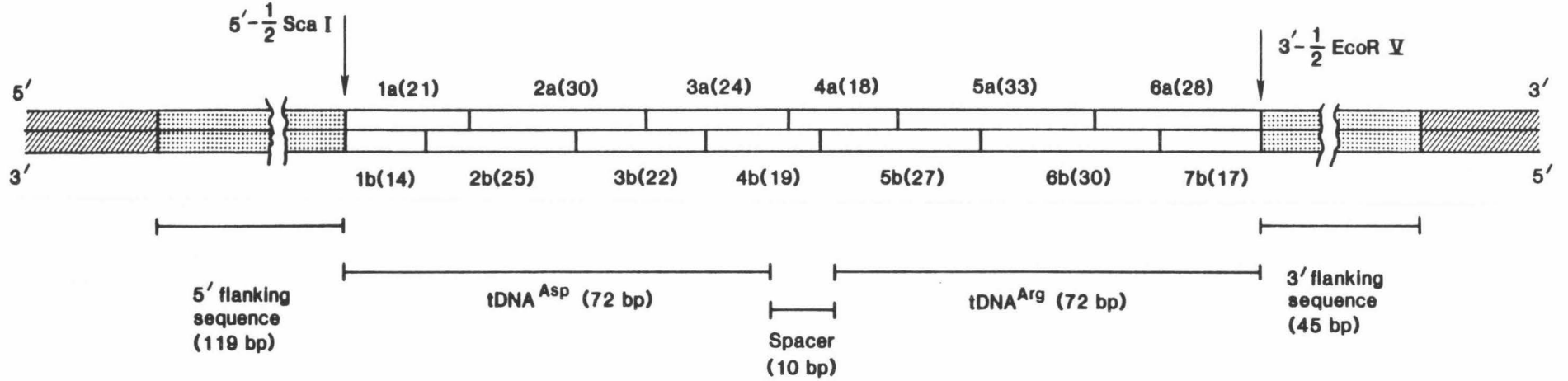
Table I. Oligonucleotide Sequences. Oligonucleotides *S* and *E* will be used as mutagen in the site-specific introduction of *Sca*I and *Eco*RV recognition sequences at the 5' end of the $tRNA^{Arg}$ and the 3' end of the $tRNA^{Asp}$ genes, respectively, in the parental plasmid, M13ArgAsp. Oligonucleotides 1a - 6a and 1b - 7b are the thirteen synthetic, self-complementary oligonucleotides whose self-assembly generates the switched gene tandem, $tRNA^{Asp/Arg}$. The structure of the cloned, self-assembled arrangement is shown in Fig. 1.

Figure 1. Structure of the M13AspArg Plasmid. The self-assembled synthetic oligodeoxynucleotides are indicated above and below each strand (for sequences, refer to Table I). Numbers in parentheses indicate lengths of the oligonucleotides as number of bases. Natural yeast flanking sequences are represented by the stippled blocks. Only the 5' half of the *Sca*I and the 3' half of the *Eco*RV recognition sequences were retained in M13AspArg, as indicated.

Table I

Oligonucleotide Sequences

Name	Length (No. of bases)	Sequence (5' → 3')
S	20	CCA CGC GAG TAC TTA TTT GT
E	20	CCA AAA AGA TAT CCG CGA CG
1a	21	TCC GTG ATA GTT TAA TGG TCA
2a	30	GAA TGG GCG CTT GTC GCG TGC CAG ATC GGG
3a	24	GTT CAA TTC CCC GTC GCG GAG CTT
4a	18	TGT TTC TGC TCG CGT GGC
5a	33	GTA ATG GCA ACG CGT CTG ACT TCT AAT CAG AAG
6a	28	ATT ATG GGT TCG ACC CCC ATC GTG AGT G
7b	17	CAC TCA CGA TGG GGG TC
6b	30	GAA CCC ATA ATC TTC TGA TTA GAA GTC AGA
5b	27	CGC GTT GCC ATT ACG CCA CGC GAG CAG
4b	19	AAA CAA AGC TCC GCG ACG G
3b	22	GGA ATT GAA CCC CGA TCT GGC A
2b	25	CGC GAC AAG CGC CCA TTC TGA CCA T
1b	14	TAA ACT ATC ACG GA


5′-½ Sca I
3′-½ EcoR V
5′
3′
1a(21)
2a(30)
3a(24)
4a(18)
5a(33)
6a(28)
1b(14)
2b(25)
3b(22)
4b(19)
5b(27)
6b(30)
7b(17)
5′ flanking sequence (119 bp)
tDNA Asp (72 bp)
Spacer (10 bp)
tDNA Arg (72 bp)
3′ flanking sequence (45 bp)

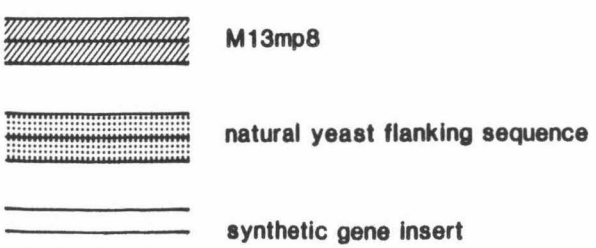

M13mp8
natural yeast flanking sequence
synthetic gene insert

Figure 1

Research Proposition V

Role of a Proline-Rich Polypeptide in Germinating Soybean

Summary

1A10 is one of five mRNAs that accumulate in soybean during germination. Sequence analysis of the cDNA copy of 1A10 reveals an open reading frame corresponding to a short polypeptide that is 40% proline. Since proline residues are potential hydroxylation sites in plant proteins, and germinating seeds require large quantities of glycoproteins for rapid primary cell wall synthesis, we would like to investigate the possibility that the 1A10 polypeptide is a glycoprotein precursor wherein the hydroxylated prolines are the major glycosylation sites. To address this problem, it is proposed that a synthetic polypeptide corresponding to 1A10 be prepared and used as substrate for hydroxylation and glycosylation activities in a crude soybean cell extract in the presence of H^3-arabinose or H^3-mannose. Long-term studies involving 1A10 and its polypeptide product are also discussed.

Introduction

It is generally believed that glycoproteins are the morphogenic substrates in multicellular organisms. In plants, particularly, cell type is determined by the type and composition of carbohydrates that deposit in the form of a cell wall.

Soybean is a useful and convenient organism for investigations on plant cell morphogenesis because of its rapid growth rate and well-characterized germination growth stage. Furthermore, a soybean genomic library is now available for the screening of genes for which appropriate probes are available (A. Marcus, personal communication).

In the laboratory of Dr. Abraham Marcus at the Fox Chase Cancer Center in Philadelphia, five mRNAs that substantially increase in soybean between 16 and 31 hr after germination are being studied. cDNAs corresponding to each of these five mRNAs are available. It is believed that none of these five mRNAs are simply growth maintenance-related, since four of them are absent from the 0.5 cm tip of 31-hr seedlings, and the fifth mRNA, although present in the seedling tip, is not found in soybean cells growing in culture. It is more likely that these mRNAs code for proteins involved in the various stages of maturation of the seedling primary cell wall (A. Marcus, personal communication).

To address the above hypothesis, attention was focused on one of the above five mRNAs, a 0.9 kb mRNA designated 1A10. In the 31-hr seedlings, 90% of the transcripts is a 1.2 kb mRNA, and 5% is 3 kb. In growing soybean cells in culture and in the seedling roots, the same mRNAs are present in a similar ratio, but the total quantity is greatly elevated. In the stem of the hypocotyl stage after 4 days, the total RNA is present at a reduced level, and 50% of it is a new transcript, the 0.9 kb 1A10 mRNA. By 5 days, the 1A10 level rises to 90% of total mRNA. If germination is done in the presence of abscidic acid, a hormone that inhibits seed

germination, the 31-hr "seedlings" contain only the 1A10 transcript, but at a greatly elevated level. Sequence analysis of the cDNA copy of 1A10 reveals the presence of two open reading frames (ORF) in a 365 nucleotide sequence. One of these ORFs code for a protein which possesses 40% proline, 20% lysine, and 17% tyrosine.

Proline residues in a number of polypeptides are usually hydroxylated (Alberts *et al.*, 1983). We would like to investigate the possibility that the abundant proline residues in the 1A10 polypeptide are hydroxylated, and the resulting hydroxyl residues linked to sugar residues, for example arabinose or mannose, to form a glycoprotein required for primary cell wall formation.

Proposed Experiments and Discussion

As a source of hydroxylation and sugar-attachment enzymatic activities, a crude extract will be prepared from cultured soybean cells using standard protocols (Montreuil *et al.*, 1956; Scopes, 1982). Extract preparation from plant tissues, however, is quite different from ordinary procedures. First, presence of the tough, cellulosic cell wall must be taken into account. Second, only a small fraction of the volume of plant tissue is intracellular; the most part is largely either vacuolar or intercellular space. This means that upon tissue disruption, much liquid is released, therefore only a small volume of extractant liquid is necessary. Finally, there is a great tendency for acidification and oxidation of certain extract components. For example, most plant tissues contain phenolic compounds which are easily oxidized (either by atmospheric O_2 or by endogenous oxidases in the extract) to form dark pigments. These pigments tend to covalently attach themselves to proteins, including the enzymes, which might then be inactivated. Inclusion of thiol compounds such as β-mercaptoethanol, inhibit such oxidation; addition of polyvinyl pyrrolydone (PVP), on the other hand, excludes the phenolic compounds by adsorbing them.

A synthetic peptide corresponding to the 1A10 polypeptide will also be made. Synthesis of peptides of up to 100 amino acid residues is now possible (Caltech Microchemical Facility, personal communication). This synthetic peptide will be incubated with the soybean cell extract in the presence of various tritiated sugars. Incubation in the presence of H^3-arabinose will be attempted first since hydroxyproline-arabinose linkages are usually common among plant glycoproteins, particularly in soybean (A. Marcus, personal communication). The putative labeled glycoprotein will then be extracted from the cell debris, and then analyzed by gel electrophoresis. Two control experiments, the first done in the presence of unlabeled sugar, and the second in the absence of sugar, will be performed, and gel electrophoresed in parallel with the experimental reactions. Polypeptide bands will be visualized by a combination of autoradiography and silver staining.

Long-term studies involving 1A10 and its protein product would include cloning of the gene using the soybean genomic library, and the 1A10 cDNA as probe. Sequencing of the gene will be performed, for it might reveal the presence of an intron, whose splicing, in turn, might be involved in the regulation of its expression, a problem that is of prime significance in the morphogenesis of any plant.

References

Alberts, B., Bray, D., Lewis, J., Raff, M., Roberts, K., and Watson, J. D (1983). Molecular Biology of the Cell, Garland Publishing, Inc., New York, pp. 1099-1146.

Marcus, Abraham, Institute for Cancer Research, Fox Chase Cancer Center, 7701 Burholme Avenue, Philadelphia, Pennsylvania, personal communications.

Montreuil, J., Bouquelet, S., Debray, H., Fournet, B., Spit, G., and Strecker, G. (1986) Glycoproteins. In: Carbohydrate Analysis, A Practical Approach, Chaplin, M. F. and Kennedy, J. F. (eds.), IRL Press, Washington, D.C., pp. 143-204.

Scopes, R. (1982) Protein Purification, Principles and Practice, Springer-Verlag, New York.