Volume 12 Number 24 1984 Nucleic Acids Research

Transfer RNA splicing in *Saccharomyces cerevisiae:* defining the substrates

Richard C.Ogden, Ming-Chou Lee+ and Gayle Knapp+*

Agouron Institute, 505 Coast Boulevard South, La Jolla, CA 92037, and +Department of Microbiology, University of Alabama, Birmingham, AL 35294, USA

Received 14 September 1984; Revised and Accepted 14 November 1984

ABSTRACT

The primary sequences of all the tRNA precursors which contain intervening sequences and which accumulate in the Saccharomyces cerevisiae rnal mutant are presented. A combination of DNA and RNA sequence analysis has led to elucidation of the primary sequence of four hitherto uncharacterized precursors. The location of the intervening sequence has in all cases been unambiguously determined by analysis of the intermediates in the splicing reaction. Secondary structures based upon the tRNA cloverleaf are shown for all the tRNA precursors and discussed with respect to common recognition by the yeast splicing endonuclease.

INTRODUCTION

Transfer RNA splicing in the yeast Saccharomyces cerevisiae has been well described in recent years. The initial report by Goodman et al. (1) showed that three different $tRNA^{Tyr}$ genes are not colinear with the mature RNA product but contain an intervening sequence of 14 base pairs (bp) interrupting the anticodon loop. The sequence of an ochre suppressor allele (SUP4) of one of the $tRNA^{Tyr}$ genes was shown to contain the same intervening sequence, thus providing good evidence that removal of the intervening sequence was a necessary step in the biosynthesis of the mature suppressor tRNA. Subsequently, three $tRNA^{Phe}$ genes were shown to contain intervening sequences of 18 or 19 bp (2) bearing no obvious homology to the $tRNA^{Tyr}$ intervening sequence other than in their position with respect to the mature RNA sequence. The discovery that tRNA precursors accumulate in the yeast rnal mutant (3) prompted a detailed investigation into the identity of these precursors and an analysis by RNA fingerprinting of the precursor-specific sequences (4). The identities of individual precursors were confirmed in some cases by hybridization to a collection of E. coli recombinant plasmid clones containing known tRNA genes (5). In this way, five known tRNAs, including $tRNA^{Tyr}$ and $tRNA^{Phe}$, were shown to accumulate as precursors in the rnal mutant. By fingerprint analysis, it was shown that the precursors to $tRNA^{Tyr}$ and $tRNA^{Phe}$

differed from the mature sequences by the presence of additional oligonucleotides consistent with the presence of the intervening sequence. A wild type yeast extract was shown to be capable of precisely removing the intervening sequence from the precursor to generate the mature tRNA sequence (4,6).

Studies on the mechanism of the yeast tRNA splicing reaction initially described the existence and sequence of linear intermediates, the two tRNA halves and the intervening sequence, generated in crude extract in the absence of ATP and the chemical nature of the termini of these fragments as isolated from crude extracts, a hydroxyl group at the 5' end and a phosphatase sensitive group at the 3' end (7,8). The positional isomer (2' or 3') was not revealed by the method of analysis but the mobility of the 3' terminal oligonucleotide and the phosphatase sensitivity suggested a phosphomonoester. Furthermore, removal of this group from the 5' tRNA half prevented ligation to the 3' half. More recently, a reevaluation of the mechanism of the reaction using partially purified splicing endonuclease and ligase revealed that 2',3'-cyclic phosphodiester termini are initially generated by the endonuclease (9) and that, prior to ligation, the 5' terminus of the 3' tRNA half is phosphorylated and activated by addition of AMP (10) in a fashion much resembling the mechanism of T4 RNA ligase (11). Ligation occurs either with or subsequent to opening of the cyclic phosphodiester to the 2' position to generate a 3',5'-phosphodiester, 2'-phosphate structure at the junction. The subsequent removal of the 2' phosphate has not been reported in vitro using the purified activities but can be implied from RNA analysis of products of splicing in crude extracts reported below.

The sequences of three additional tRNA precursors whose identities were known from colony filter hybridization have been reported, the data deriving from both DNA and RNA sequence analysis (12,13). The present paper completes the characterization of the primary and putative secondary structures of the remaining uncharacterized substrates and catalogs the splicing endonuclease substrates that accumulate in the rnal mutant. Although we know that RNA splicing is not an obligatory step in the biosynthesis of all yeast tRNAs, we do not know whether the precursors which are presented here are the only substrates for tRNA splicing in yeast. At the present time, they define the range of substrates for this particular step in yeast tRNA biosynthesis.

MATERIAL AND METHODS

Preparation of ^{32}P-labelled tRNA for sequencing.

Uniformly ^{32}P-labelled tRNA precursors were prepared and isolated as previously described (14) with modifications in the gel system which are presented below and in Results. The 5' and 3' half tRNA fragments and introns were prepared from uniformly ^{32}P-labelled pre-tRNAs as previously described (8). Terminally labelled pre-tRNAs were generated using yeast CCA·tRNA nucleotidyl transferase and unlabelled RNA which was prepared from the rnal mutant of Saccharomyces cerevisiae, strain ts136, as described elsewhere (15). The labelled RNAs were separated in a two-dimensional polyacrylamide gel electrophoresis system. This two-dimensional system consisted of a first dimension 10% (w/v) polyacrylamide (30:1, acrylamide:bis-acrylamide) gel containing 4 M urea, 1 mM EDTA and 90 mM Tris·borate (pH 8.3) and a second dimension 20% (w/v) polyacrylamide (30:1, acrylamide:bis-acrylamide) gel containing 7 M urea, 1 mM EDTA and 90 mM Tris·borate (pH 8.3). Each intron-containing tRNA precursor was identified by colony filter hybridization to matrices of clones selected from the Hind III collection described by Beckmann et al. (5) and an Eco RI genomic library (unpublished data).

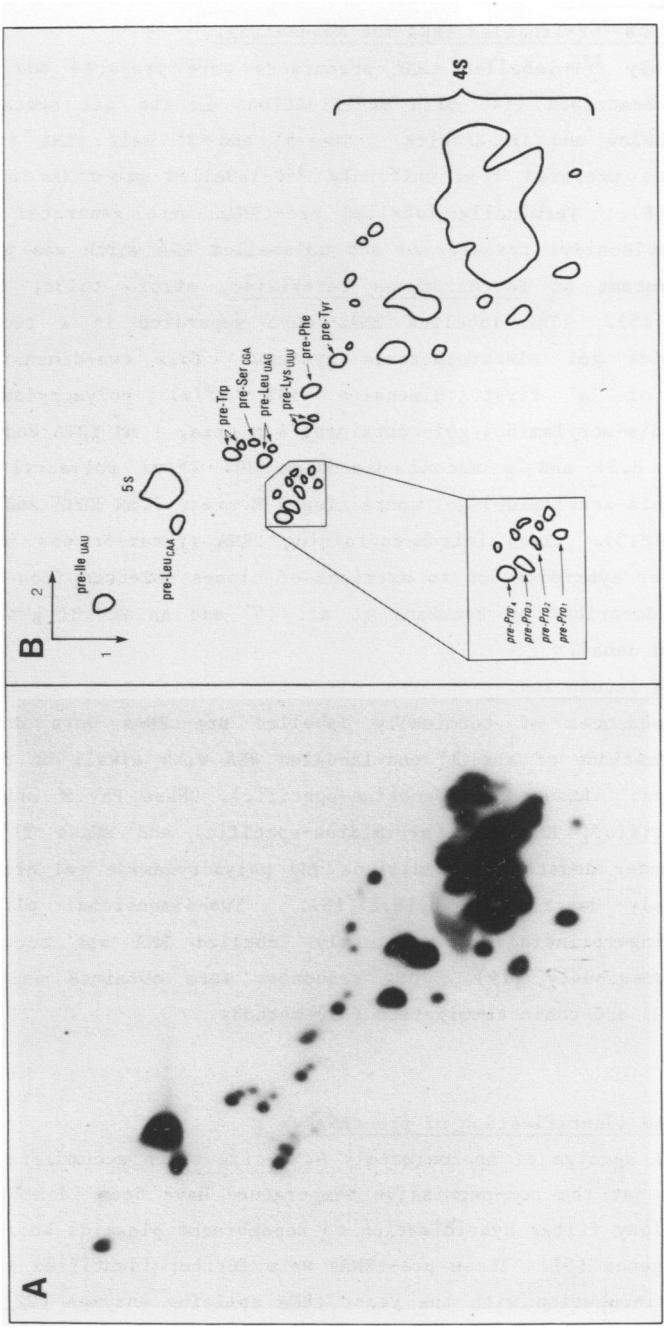
Nucleic acid sequencing.

The sequences of terminally labelled pre-tRNAs were determined by partial digestions of the 3' end-labelled RNA with alkali or base-specific ribonucleases: RNase U2 (adenosine-specific), RNase Phy M (adenosine- and uridine-specific), RNase A (pyrimidine-specific) and RNase T1 (guanosine-specific) under denaturing conditions and polyacrylamide gel electrophoresis as previously described (15,16,17,18). Two-dimensional oligonucleotide mapping (fingerprinting) of uniformly labelled RNA was accomplished as described previously (19). DNA sequences were obtained using both the chemical (20) and chain termination (21) methods.

RESULTS

Isolation and identification of pre-tRNAs.

The RNA species of approximately 4.5S size which accumulate in the rnal yeast mutant at the non-permissive temperature have been identified as pre-tRNAs by colony filter hybridization to recombinant plasmids known to contain yeast tRNA genes (5). These pre-tRNAs were further identified as containing introns by incubation with the yeast tRNA splicing enzymes (22). Figure 1 summarizes the precursor identities including their former designations in

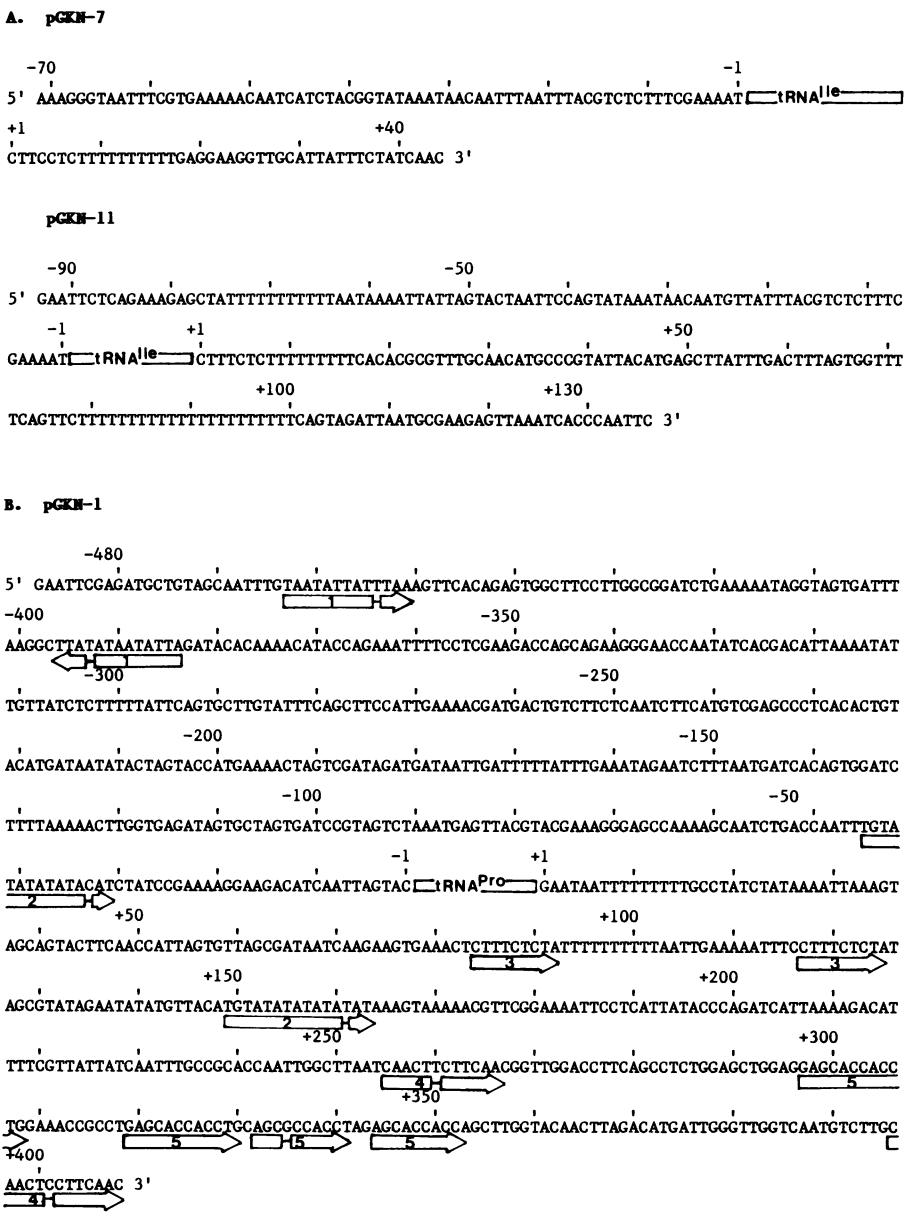

A
B
2
1
pre-Ile UAU
5S
pre-Leu CAA
pre-Trp
pre-Ser CGA
pre-Leu UAG
pre-Lys UUU
pre-Phe
pre-Tyr
4S
pre-Pro4
pre-Pro3
pre-Pro2
pre-Pro1

previously published work (4,22). From Figure 1 it is obvious that the original 11/12 designation underestimated the numbers of precursors to $tRNA^{Pro}_{UGG}$. Some of the eleven RNA species which hybridize to pGKN-1 and pGKN-2 (the hybridization is not sensitive to minor sequence variations in the intervening sequence) may be attributable to partial 3' terminal maturation. Those RNAs which are on a 45° diagonal toward faster mobility in both dimensions have not been detected on duplicate gels separating pre-tRNAs that have been 3' terminally labelled using nucleotidyl transferase (i.e., only intact CCA_{OH} species are observed). There are five RNA species in this region of a two-dimensional separation of 3' end-labelled pre-tRNAs which hybridize to pGKN-1 which is known (by sequence data) to contain only the $tRNA^{Pro}_{UGG}$ gene. The other pre-tRNAs which have been identified and sequenced in the work presented here are pre-$tRNA^{Ile}_{UAU}$ and pre-$tRNA^{Leu}_{UAG}$. A gene encoding $tRNA^{Lys}_{UUU}$ has been sequenced previously and shown to contain an intervening sequence (23). We have confirmed that the precursor to this tRNA accumulates in the rnal mutant and have unambiguously assigned the location of the intervening sequence. It should be noted that the relative mobilities of pre-$tRNA^{Lys}_{UUU}$ and pre-$tRNA^{Phe}$ differ between Figure 1 and previously published data (22). The salient difference between the two gel systems is the anticipated amount of denaturation (7 M versus 4 M urea, respectively). Furthermore, the increased urea concentration decreases the resolution between pre-$tRNA^{Leu}_{CAA}$ and 5S rRNA and increases the resolution of at least four precursors to $tRNA^{Pro}_{UGG}$.

DNA sequence analysis.

Extensive fingerprint analysis of the precursors, intermediates and the spliced mature-sized tRNA revealed that two of the unidentified precursor RNAs, designated 19 and 11/12, had not previously been reported as mature yeast tRNAs. Their primary structure has been established by sequencing cloned genes. In the case of the two genes for precursor 19, plasmid pGKN-11

Figure 1. Separation of tRNA precursors. Two-dimensional polyacrylamide gel electrophoretic separation is shown. Uniformly ^{32}P-labelled RNA was isolated (as previously described (14)) at the non-permissive temperature from a temperature-sensitive diploid yeast, M304, that is homozygous at the rnal locus. (A) Autoradiography. (B) Schematic representation and identification of the RNA species. The two dimensions indicated at the top left of the schematic are: first, 10% polyacrylamide (30:1 acrylamide:bis-acrylamide), 4 M urea; second, 20% polyacrylamide (30:1 acrylamide:bis-acrylamide), 7 M urea. Both dimensions contained 90 mM Tris·borate, pH 8.3, 1 mM EDTA. Correspondence between the identified precursors and their designations in previous work (22) are as follows: pre-Ile_{UAU}=precursor 19; pre-Leu_{CAA}=pre-$tRNA^{Leu}_{3}$; pre-Ser_{CGA}=pre-$tRNA^{Ser}_{UCG}$; pre-Leu_{UAG}=precursor 13; pre-Pro4=precursor 12; pre-Pro3=precursor 11; pre-Lys_{UUU}=precursor 9a.


A. pGKN-7
-70
-1
5' AAAGGGTAATTTCGTGAAAAACAATCATCTACGGTATAAATAACAATTTAATTTACGTCTCTTTCGAAAAT tRNAIle
+1
+40
CTTCCTCTTTTTTTTTTGAGGAAGGTTGCATTATTTCTATCAAC 3'
pGKN-11
-90
-50
5' GAATTCTCAGAAAGAGCTATTTTTTTTTTAATAAAATTATTAGTACTAATTCCAGTATAAATAACAATGTTATTTACGTCTCTTTC
-1
+1
+50
GAAAAT tRNAIle CTTTCTCTTTTTTTTTCACACGCGTTTGCAACATGCCCGTATTACATGAGCTTATTTGACTTTAGTGGTTT
+100
+130
TCAGTTCTTTTTTTTTTTTTTTTTTTTTTCAGTAGATTAATGCGAAGAGTTAAATCACCCAATTC 3'
B. pGKN-1
-480
5' GAATTCGAGATGCTGTAGCAATTTGTAATATTATTTAAAGTTCACAGAGTGGCTTCCTTGGCGGATCTGAAAAATAGGTAGTGATTT
1
-400
-350
AAGGCTTATATAATATTAGATACACAAAACATACCAGAAATTTTCCTCGAAGACCAGCAGAAGGGAACCAATATCACGACATTAAAATAT
-300
-250
TGTTATCTCTTTTTATTCAGTGCTTGTATTTCAGCTTCCATTGAAAACGATGACTGTCTTCTCAATCTTCATGTCGAGCCCTCACACTGT
-200
-150
ACATGATAATATACTAGTACCATGAAAACTAGTCGATAGATGATAATTGATTTTTATTTGAAATAGAATCTTTAATGATCACAGTGGATC
-100
-50
TTTTAAAAACTTGGTGAGATAGTGCTAGTGATCCGTAGTCTAAATGAGTTACGTACGAAAGGGAGCCAAAAGCAATCTGACCAATTTGTA
-1
+1
TATATATACATCTATCCGAAAAGGAAGACATCAATTAGTAC tRNAPro GAATAATTTTTTTTTGCCTATCTATAAAATTAAAGT
2
+50
+100
AGCAGTACTTCAACCATTAGTGTTAGCGATAATCAAGAAGTGAAACTCTTTCTCTATTTTTTTTTTAATTGAAAAATTTCCTTTCTCTAT
3
3
+150
+200
AGCGTATAGAATATATGTTACATGTATATATATATATAAAGTAAAAACGTTCGGAAAATTCCTCATTATACCCAGATCATTAAAAGACAT
2
+250
+300
TTTCGTTATTATCAATTTGCCGCACCAATTGGCTTAATCAACTTCTTCAACGGTTGGACCTTCAGCCTCTGGAGCTGGAGGAGCACCACC
4
5
+350
TGGAAACCGCCTGAGCACCACCTGCAGCGCCACCTAGAGCACCACCAGCTTGGTACAACTTAGACATGATTGGGTTGGTCAATGTCTTGC
5
5
5
+400
AACTCCTTCAAC 3'
4
pGKN-2
-20
-1
+1
5' TTTATTTAGACACCAACAGATACAT tRNAPro ATTATAATTTTTTCACTTTTTTGTTCTTTGCAAGAACGCGCGGTAC
+50
+100
GCAGTTATGAGATGATGTAGGCAATCTCGAGAATTGAAACTTCTGCATTACCAT '3

contains a 0.73 kilobase (kb) Eco RI fragment inserted in pBR322 and plasmid pGKN-7 contains two Eco RI fragments in pBR322 (5 kb and 2 kb), the larger of which contains the gene as shown by hybridization (24). The two genes and their flanking regions were sequenced by the chemical method (20) according to the strategy outlined in Figure 2. The sequences of the non-transcribed strand flanking regions are shown in Figure 2A and the sequences of the genes are presented in Figure 4 as RNAs. Fingerprint analysis of precursor 19 (Ogden and Harrell, unpublished data; 9) is consistent with the sequence between positions 1 and 133 (the terminal CCA_{OH} not being encoded) and confirms the position of the intervening sequence (see Fig. 4). The sequence of the mature tRNA can be folded into a conventional cloverleaf secondary structure, indicating that it decodes AUA and corresponds to $tRNA^{Ile}_{UAU}$. The two genes show a single difference, within the mature portion of the coding sequence at position 119, which is also observed in the RNA fingerprint analysis of precursor 19 (data not shown). This demonstrates that both genes are transcribed in vivo in the rnal mutant.

Two clones containing tRNA genes which hybridize with precursor 11/12 and a lower abundance RNA species with distinct mobility on two-dimensional polyacrylamide gel systems (pre-Prol) were similarly subjected to extensive restriction mapping and sequence analysis. Plasmid pGKN-1 contains a 1.1 kb Eco RI fragment of yeast DNA inserted in pBR322 and plasmid pGKN-2 contains a 4.1 kb Eco RI fragment in pBR322. The chemical sequencing data for pGKN-1 was additionally confirmed by enzymatic sequencing (chain-termination method (21)) of the Eco RI fragment cloned in M13mp9. The sequences of the non-transcribed strand flanking regions are shown in Figure 2B. The sequences of the $tRNA^{Pro}_{UGG}$ genes are presented in Figure 4 as RNAs. Fingerprint analyses of precursors in the 11/12 region (data not shown) were consistent with the DNA sequences obtained (the terminal CCA_{OH} not being encoded) and confirms the position of the intervening sequence as assigned in Figure 4.

Figure 2. Gene Flanking sequences of yeast $tRNA^{Ile}_{UAU}$ and $tRNA^{Pro}_{UGG}$. The sequences of flanking regions of 4 yeast tRNA genes are shown: (A) $tRNA^{Ile-1}_{UAU}$ and $tRNA^{Ile-2}_{UAU}$ and (B) $tRNA^{Pro}_{UGG}$. Only the non-transcribed strand is shown for simplicity although both strands were sequenced. Flanking sequences are numbered from the gene using a minus or plus symbol for 5' or 3' flanking sequences, respectively. Restriction sites used for sequencing are located as follows: pGKN-7--Hae III at 84 and Xho I at 125; pGKN-11--Eco RI at -93, Xho I at 125; pGKN-1--Eco RI at -488, Hinf I at -156 and 22, Sph I at 65 (confirmatory data obtained using Rsa I at +359, +41, -4 and -79 and Taq I at -355, -238 and -199); pGKN-2--Sph I at 65, Hinf I at 22 and Xho I at +72. Numbers without minus or plus symbols refer to nucleotide positions in the gene as shown in Figure 4. Arrows below the sequence mark repetitive sequences and can be matched via the associated number.

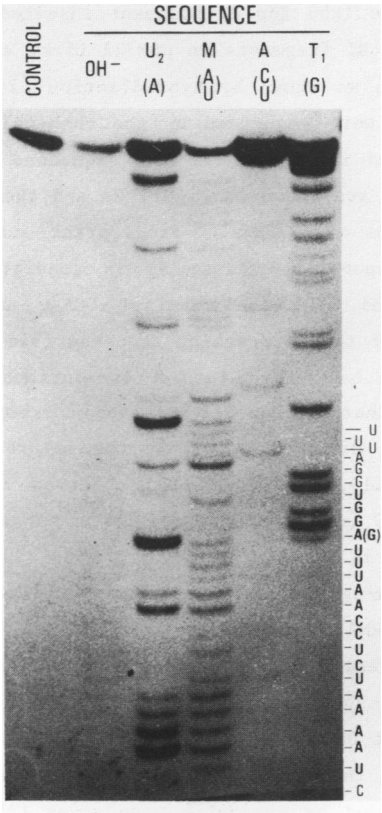


Figure 3. Enzymatic sequencing of pre-tRNA$^{Leu}_{UAG}$. Pre-tRNA$^{Leu}_{UAG}$, labelled at the 3' end was subjected to limited hydrolysis by a variety of reagents according to standard procedures (15,16,17). Following separation of the products by gel electrophoresis, the sequence was read directly from the autoradiograph. The portion of the autoradiograph shown corresponds to the intervening sequence which is presented in bold type. A brief portion of the mature sequence is also shown. The single sequence heterogeneity detected within the intervening sequence is indicated. The nucleases used in sequencing are as follows: U_2=RNase U_2 (1 x 10^{-1} u/uRNA); M=RNase PhyM (1 u/ug RNA); A=RNase A (1 x 10^{-6} u/ug RNA); T1=RNase T1 (1 x 10^{-2} u/ug RNA). Base specificities are indicated in brackets, OH^- represents hydrolysis with alkali (non-specific). The control is untreated end-labelled RNA.

RNA sequence analysis.

The two remaining precursors, formerly designated 9a and 13, have been shown by RNA fingerprint analysis (19) and end-labelled RNA sequencing (16,17) to correspond to precursors to tRNA$^{Lys}_{UUU}$ (tRNA$^{Lys}_{2}$) (25) and tRNA$^{Leu}_{UAG}$ (26), respectively. The precursor fingerprints do show additional oligonucleotides and the absence of oligonucleotides corresponding to the

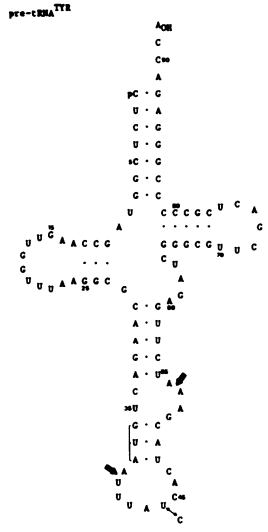
mature anticodon loop sequences (data not shown). The DNA sequence of a tRNA gene corresponding to $tRNA^{Lys}_{UUU}$ has been published (16). Our data confirm the published sequence. The correct position of the intervening sequence within the mature tRNA was deduced directly from RNA fingerprint data (not shown).

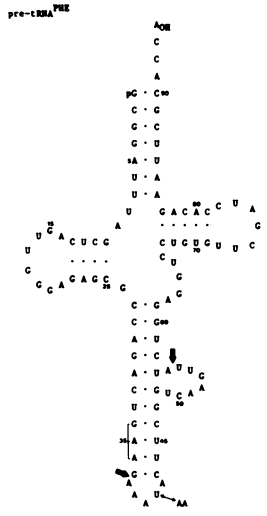
We chose to determine the primary sequence of the intervening sequence of pre-$tRNA^{Leu}_{UAG}$ by enzymatic sequencing of 3' terminally labelled precursor (Fig. 3) as previously described (15,16,17). The nucleotide sequence of the tRNA portion (not shown) of the precursor is in agreement with the published sequence (26). The sequence derived for the intervening sequence in the precursor to $tRNA^{Leu}_{UAG}$ is supported by the identity of additional oligonucleotides in RNase A fingerprint analyses (data not shown).

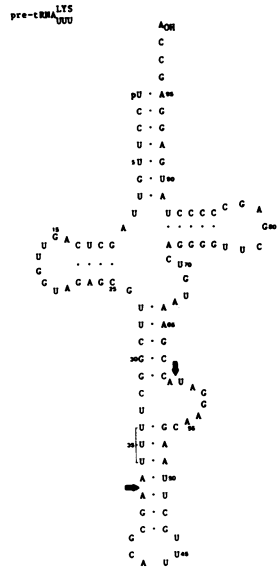
The fingerprint analysis of the spliced $tRNA^{Leu}_{UAG}$ was additionally informative. The ligated junction occurs following a guanosine. The 3' adjacent RNase T1 oligonucleotide (CUCUG) is present in molar yield. This indicates that the splice junction produced in this experiment is completelv sensitive to RNase T1. This would not be the case if the 2' position at the junction were blocked by a phosphomonoester as is observed when partially purified splicing ligase is used to produce mature tRNA (10). Similar RNase T1 sensitivity at the splice junction (although not to the same extent) has been previously reported (7). The most reasonable explanation for the different results between the crude and partially purified splicing activities is that the crude fraction (4) used in these experiments contained an activity which removed the 2'-phosphate group—a step presumed to be obligatory in the yeast tRNA splicing pathway as it is currently envisaged (9,10). This result is also found for the other mature-sized tRNAs produced in crude yeast extracts although the proof, relying as it does on chromatographic and electrophoretic mobility of the junction species, is less compelling than in this particular case.

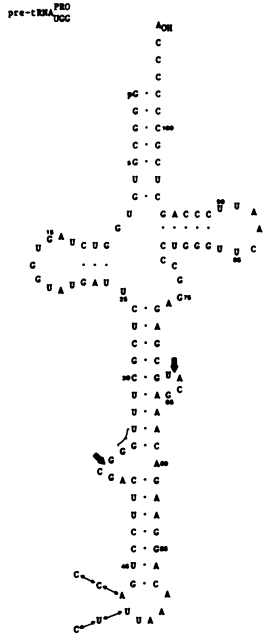
DISCUSSION

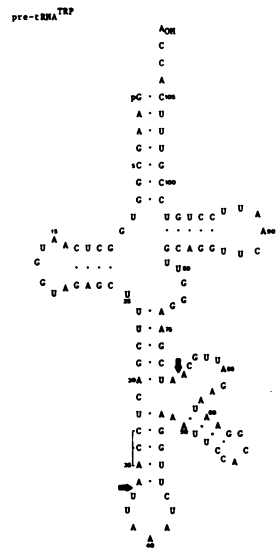
This paper completes an analysis of the primary sequence of the known yeast tRNA splicing substrates. A combination of DNA and RNA sequencing has revealed that tRNAs for eight amino acids accumulate at the nonpermissive temperature in the yeast rnal mutant as precursors containing intervening sequences. We have not sequenced every gene for each tRNA and certainly additional minor variants, masked by low abundance or by the method of RNA analysis and the general insensitivity of molar yield calculations, will be uncovered in due course. This point is particularly well illustrated in the

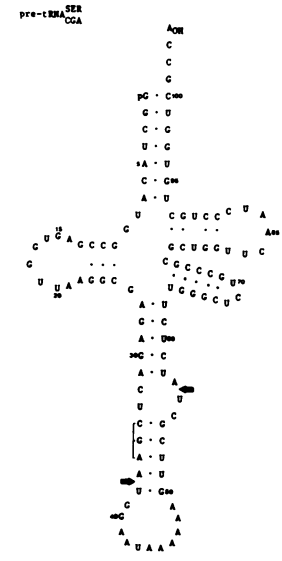

pre-tRNA^{TYR}


pre-tRNA^{PHE}


pre-tRNA$^{LYS}_{UUU}$


pre-tRNA$^{PRO}_{UGG}$


pre-tRNA^{TRP}


pre-tRNA$^{SER}_{CGA}$

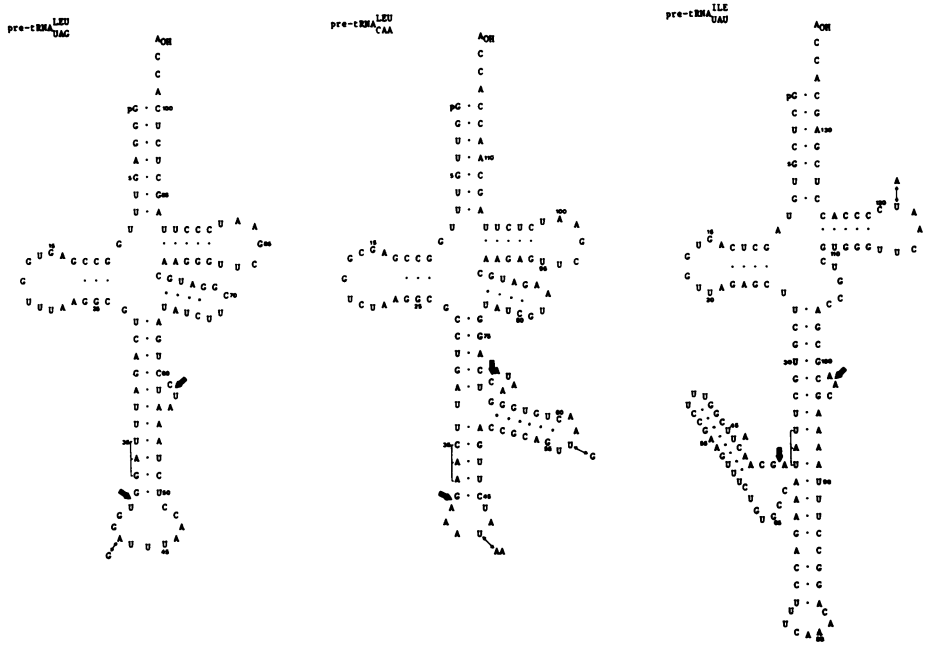


Figure 4. Sequences of yeast tRNA precursors containing intervening sequences. Sequences of tRNA precursors that accumulate in the rnal yeast mutant at the non-permissive temperature are summarized. Heavy arrows mark the positions of the 5' and 3' intron ends. The anticodon triplet is indicated adjacent to the sequence by square brackets. The variations in sequence that have been observed for each pre-tRNA family are indicated by the double-headed arrows. The sequence shown for pre-tRNA^{Ile} is that containing tRNA$^{Ile-1}_{UAU}$ found in pGKN-7; the variation (for tRNA$^{Ile-2}_{UAU}$) is found in pGKN-11. In cases where there are more than two sequence variations, the variants are shown in layers. The sequence variations for pre-tRNA$^{Pro}_{UGG}$ include two sequences obtained from sequencing pGKN-1 and pGKN-2 (inner and outer layers, respectively) and one sequence from another gene sequenced by C. Cummins (middle layer; manuscript in preparation). An additional intron sequence variation in pre-tRNA$^{Pro}_{UGG}$ has been reported elsewhere (15). References for sequence information and intron position assignments were obtained as follows: pre-tRNA^{Tyr} (1,4); pre-tRNA^{Phe} (2,4); pre-tRNA$^{Lys}_{UUU}$ (23; this paper); pre-tRNA$^{Ser}_{CGA}$ (12); pre-tRNA^{Trp} (13); pre-tRNA$^{Pro}_{UGG}$ (41; this paper); pre-tRNA$^{Leu}_{CAA}$ (13); pre-tRNA$^{Leu}_{UAG}$ (this paper, 26) and pre-tRNA$^{Ile}_{UAU}$ (this paper).

case of precursors to tRNA$^{Pro}_{UGG}$ (see Fig. 1). The data indicate a probable minimum of 5 precursors to tRNA$^{Pro}_{UGG}$ of which the sequences of three variants are presented in Figure 5. Sequence analysis of two 3' end-labelled RNA species from the same region of the two-dimensional polyacrylamide gel revealed two additional intervening sequence variants (Lee and Knapp, unpublished data; 15). For clarity in Figure 4, these have not been shown

TABLE 1

POSITIONS OF VARIATION IN INTERVENING SEQUENCES

INTRON:	region 1 ANTICODON LOOP COMPLEMENT				region 2
INTRON IN	TYR	PHE	LEU_{UAG}	LEU_{CAA}	PRO_{UGG}
TOTAL LENGTH	14	18,19	19	32,33	28,31
POSITION OF ANTICODON LOOP COMPLEMENT[a]	9-11	6-11	12-17	7-11	25-28
POSITION(S) OF SEQUENCE VARIATION:					
REGION 1	5[b]	4[b]	4[b]	4[b]	multiple[c]
REGION 2	-	-	-	19[b]	multiple[d]

[a]The numbers (counting starts with the first nucleotide of intron) and sequence refer to the one given in Fig. 5.
[b]Shown in Fig. 5.
[c]The $tRNA^{Pro}$ genes present a more complex picture. Three variants, representing one or two base substitutions, can be observed in Fig. 5. Two other sequences, pre-pro1 and pre-pro4, are more varied. Pre-Pro1 has four changes: deletion of #8 and #16 and base substitutions at #21 and 22 (A to G). Pre-pro4 has eight changes: deletion of #1-3 and 6-8, insertion of A after #10, and base change at #12 (U to A).
[d]Insertion of two or one pyrimidines after #30 in pre-pro1 and pre-pro4, respectively.

but they contain several small sequence changes which are summarized in Table 1.

Classifying the precursor tRNAs according to the translational role of their mature counterparts, variability within the classes can be found within the anticodon (Leu_{CAA} and Leu_{UAG}), within other portions of the mature tRNA (Ile) and most commonly within the intervening sequence (Tyr, Phe, Leu, Pro). It is also known that isoacceptor tRNAs (e.g., for Ser (12,27) and Pro (28)) can differ in whether they are transcribed from genes containing intervening sequences or not.

The two $tRNA^{Ile}_{UAU}$ sequences present an interesting case. Southern blot hybridization of pre-$tRNA^{Ile}$ to restriction digests of total yeast DNA indicate two copies of the gene (data not shown). The sequences of both are presented in Figures 2A and 4. In this instance, although the anticodons are the same, the two tRNAs differ by a single base in the TΨC loop. The

intervening sequences are identical and there is considerable conservation of the 5' flanking sequence to position -38. This latter observation is notable in that most yeast tRNA genes that have been sequenced seldom exhibit homology in 5' or 3' flanking regions. Both genes are observed to be equally active in vivo in the rnal mutant (as judged from molar ratios of the appropriate RNase T1 oligonucleotide (9)). Presumably the base change in what has been defined as the distal promoter element in other yeast tRNA genes (29) does not impair expression in this case.

The presence of an intervening sequence within a tRNA gene does not correlate with either high or low copy number or with the level of usage of the mature tRNA in decoding known yeast genes (reviewed in 30). For example, the gene encoding $tRNA^{Ser}_{CGA}$ is single copy and the codon UCG, which can only be read by this tRNA in yeast, is very infrequently used. Nevertheless, this represents an instance where cell viability rests heavily on a faithful splicing reaction. Similarly, the two genes for $tRNA^{Leu}_{UAU}$ contain intervening sequences and the tRNAs decode rarely used AUA codons. On the other hand, $tRNA^{Leu}_{CAA}$ and $tRNA^{Pro}_{UGG}$ genes are multicopy and the tRNAs are frequently used. Intervening sequences appear to be distributed throughout yeast tRNA genes without regard for tRNA function. It is true that in our analysis so far, a given anticodon is either always associated with an intervening sequence or not, however the intervening sequences may show heterogeneity in terms of length and/or sequence.

Figure 4 summarizes the sequences of precursors of the nine tRNAs that accumulate at the nonpermissive temperature in the rnal mutant. Not all variants are shown in the figure to avoid ambiguity. Additional sequences are referred to in the legend or in Table 1. The sequences are presented in secondary structures which maintain the four helical stems of the tRNA and maximize favorable free energies (hand estimated using the guidelines of Tinoco and coworkers (31)) of the base pairing in the anticodon loop-intervening sequence regions. In two precursors (Ile and Pro) alternate structures with similar free energies are possible. In those cases we chose the structure which conformed with structure-probing data (15). In all nine precursors, a portion of the anticodon loop including part or all of the anticodon may form base pairs with a complementary sequence in the intervening sequence. Additional secondary structure is possible in the cases of precursors which have larger introns (i.e., Ile, Leu_{CAA}, Trp and Pro). The sites of cleavage by the splicing endonuclease are in single-stranded regions except for the 5' site in pre-$tRNA^{Lys}_{UUU}$ which is shown base-

paired in the structure having the most favorable estimated free energy. It is probable that at least some of these common structural features will prove to be recognition points for the yeast splicing endonuclease. The partially purified endonuclease has been shown to cleave all nine precursors equally well (9) and the significance of these structural features is under investigation. Table 1 presents a summary of sequence variations among intervening sequences found in these pre-tRNAs. The one region of the intervening sequence that is invariant is the region which contains the sequence which is complementary to the anticodon loop.

The function of intervening sequences in yeast tRNA genes is not immediately apparent from an inspection of their occurrence. Early experiments were directed towards exploring a transcriptional role for the intervening sequences in this subset of genes (32) and, with the subsequent description of the split internal promoter for eukaryotic tRNA genes (reviewed in 32), further investigations along these lines were pursued (33). A powerful and different approach investigated the effects of precisely deleting the intervening sequence from a suppressor allele of a $tRNA^{Tyr}$ gene (SUP6) on suppressor function in vivo (35,36). It was found that suppressor function was impaired if the gene containing the deletion was present in single or low copy (as opposed to its presence on a multicopy 2 micron plasmid derived vector when suppression was not distinguishable from that conferred by the intact suppressor gene). Furthermore, a difference in modification in the anticodon was clearly shown between in vivo transcripts from the normal versus the deletion-containing suppressor gene. This result suggests that at least in this system ($tRNA^{Tyr}$) the pseudouridylate synthetase responsible for introducing the modification in the anticodon recognizes the precursor and not the mature sequence.

The relationship between intervening sequences in tRNA genes of S. cerevisiae and of other organisms is another mystery. Heterologous splicing studies, where precursors from one system (most frequently, S. cerevisiae, S. pombe or Drosophila) are processed in extracts from another (most frequently X. laevis) generally in a coupled transcription/processing system (15,36), suggest some common features for the precursors. This does not always appear to be displayed at the secondary structure level upon inspection (37,38) but no experiments along the lines of those done in S. cerevisiae (15,39) have been pursued. Possible secondary structure in the intervening sequence-anticodon loop region is not always found in pre-tRNAs from other organisms, suggesting that the tRNA splicing endonucleases from

various organisms recognize preferentially different pre-tRNA structural features. Conversely, the conservation of position of the intervening sequence in nuclear tRNA genes among organisms suggests a common basis for structural recognition by the tRNA splicing ligases. This is indeed observed, however, the nature of the reaction catalyzed as well as the specificity of the reaction for tRNA is variable (reviewed in (40) and references therein). It is not known whether these alternate mechanisms are involved in homologous tRNA splicing systems. S. cerevisiae remains the best studied system with regard to tRNA splicing. It is anticipated that the scope of the reaction, as represented by the substrates defined in this work, will yield to further biochemical and genetic analysis.

ACKNOWLEDGEMENTS

We wish to thank John Abelson for his support and our colleagues, past and present, in his laboratory for their interest in this work. Beverly Harrell and Joe Sandmeyer helped significantly with much of the RNA analysis; Stephanie Bragg, Yvonne Pfanenstiel and Yvette Cook provided excellent technical assistance. This work was supported by National Institutes of Health Grant GM 30356 (R.C.O.) and American Cancer Society Grant MV-138 and a University of Alabama in Birmingham Graduate School Faculty Research Grant (G.K.).

* To whom correspondence should be addressed.

REFERENCES

1. Goodman,H.M., Olson,M.V. and Hall,B.D. (1977) Proc. Natl. Acad. Sci. USA 74, 5453-5457.
2. Valenzuela,P., Venegas,A., Weinberg,F., Bishop,R. and Rutter,W.J. (1978) Proc. Natl. Acad. Sci. USA 75, 190-194.
3. Hopper,A.K., Banks,F. and Evangelidis,V. (1978) Cell 14, 211-219.
4. Knapp,G., Beckmann,J.S., Johnson,P.F., Fuhrman,S.A. and Abelson,J. (1978) Cell 14, 221-236.
5. Beckmann,J.S., Johnson,P.F., Knapp,G., Sakano,H., Fuhrman,S.A., Ogden, R.C. and Abelson,J. (1978) in Nonsense Mutations and tRNA Suppressors, Celis,J.E. and Smith,J.D. Eds., pp. 207-234, Academic Press, London.
6. O'Farrell,P.Z., Cordell,B., Valenzuela,P., Rutter,W.J. and Goodman,H.M. (1978) Nature 274, 438-445.
7. Peebles,C.L., Ogden,R.C., Knapp,G. and Abelson,J. (1979) Cell 18, 27-35.
8. Knapp,G., Ogden,R.C., Peebles,C.L. and Abelson,J. (1979) Cell 18, 37-45.
9. Peebles,C.L., Gegenheimer,P. and Abelson,J. (1983) Cell 32, 525-536.
10. Greer,C.L., Peebles,C.L., Gegenheimer,P. and Abelson,J. (1983) Cell 32, 537-546.
11. Uhlenbeck,O.C. and Gumport,R.I. (1982) in The Enzymes, Boyer,P.D. Ed., Vol. 15, pp. 31-58, Academic Press, New York.
12. Etcheverry,T., Colby,D. and Guthrie,C. (1979) Cell 18, 11-26.

13. Kang,H.S., Ogden, R.C. and Abelson,J. (1980) in Mobilization and Reassembly of Genetic Information, Scott,W.A., Werner,R., Joseph,D.R., and Shultz,J. Eds., pp. 317-334, Academic Press, New York.
14. Ogden,R.C., Beckmann,J.S., Abelson,J., Kang,H.S., Soll,D. and Schmidt,O. (1979) Cell 17, 399-406.
15. Lee, M.-C. and Knapp,G. (1984) Submitted to J. Biol. Chem.
16. Donis-Keller,H., Maxam,A.M. and Gilbert,W. (1977) Nucleic Acids Res. 4, 2527-2531.
17. Vournakis,J.N., Celantano,J., Finn,M., Lockard,R.E., Mitra,T., Pavlakis, G., Troutt,A., van den Berg,M. and Wurst,R.M. (1981) in Gene Amplification and Analysis, Chirikjian,J.G. and Papas,T.S. Eds., Vol. 2, pp. 267-298, Elsevier-North Holland, New York.
18. Maxam,A.M. and Gilbert,W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
19. Brownlee,G.G. (1972) in Laboratory Techniques in Biochemistry and Molecular Biology, Work,T.S. and Work,E. Eds., pp. 67-99, American Elsevier Publ. Co. Inc., New York.
20. Maxam,A.M. and Gilbert,W. (1980) Methods Enzymol. 65, 499-560.
21. Sanger,F., Nicklen,S. and Coulson,A.R. (1977) Proc. Natl. Acad. Sci. USA 74, 5463-5467.
22. Ogden,R.C., Knapp,G., Peebles,C.L., Kang,H.S., Beckmann,J.S., Johnson, P.F., Fuhrman,S.A. and Abelson,J. (1980) in Transfer RNA: Biological Aspects, Soll,D., Abelson,J. and Schimmel,P.R. Eds., pp. 173-190, Cold Spring Harbor Laboratory.
23. Del Rey,F.J., Donahue,T.F. and Fink,G.R. (1982) Proc. Natl. Acad. Sci. USA 79, 4138-4142.
24. Southern,E.M. (1979) Methods Enzymol. 65, 58-109.
25. Smith,C.J., Teh,H.-S., Ley,A.N. and D'Obrenan,P. (1973) J. Biol. Chem. 248, 4475-4485.
26. Randerath,E., Gupta,R.C., Chia,L.S.Y., Chang,S.H. and Randerath,K. (1979) Eur. J. Biochem. 93, 79-94.
27. Broach,J.R., Friedman,L. and Sherman,F. (1981) J. Mol. Biol. 150, 375-387.
28. Cummins,C.M., Donahue,T.F. and Culbertson,M.T. (1982) Proc. Natl. Acad. Sci. USA 79, 3565-3569.
29. Koski,R.A., Allison,D.S., Worthington,M. and Hall,B.D. (1982) Nucleic Acids Res. 10, 8127-8143.
30. Guthrie,C. and Abelson,J. (1982) in Molecular Biology of the Yeast Saccharomyces: Metabolism and Gene Expression, Strathern,J.N., Jones,E.W. and Broach,J.R. Eds., pp. 487-528, Cold Spring Harbor Laboratory.
31. Tinoco,I. Jr., Borer,P.N., Dengler,B., Levine,M.D., Uhlenbeck,O.C., Crothers,D.M. and Gralla,J. (1973) Nature New Biol. 246, 40-41.
32. Johnson,J.D., Ogden, R.C., Johnson,P.F., Abelson,J., Dembeck,P. and Itakura,K. (1980) Proc. Natl. Acad. Sci. USA 77, 2564-2568.
33. Hall,B.D., Clarkson,S.G. and Tocchini-Valentini,G.P. (1982) Cell 29, 3-5.
34. Raymond, G.J. and Johnson,J.D. (1983) Nucleic Acids Res. 11, 5969-5988.
35. Wallace,R.B., Johnson,P.F., Tanaka,S., Schold,M., Itakura,K. and Abelson,J. (1980) Science 209, 1396-1400.
36. Johnson,P.F. and Abelson,J. (1983) Nature 302, 681-687.
37. Gamulin,V., Mao,J., Appel,B., Sumner-Smith,M., Yamao,F. and Soll,D. (1983) Nucleic Acids Res. 11, 8537-8546.
38. Muller,F. and Clarkson,S.G. (1980) Cell 19, 345-353.
39. Swerdlow,H. and Guthrie,C. (1984) J. Biol. Chem. 259, 5197-5207.
40. Greer, C., Javor,B. and Abelson, J. (1983) Cell 33, 899-906.
41. Cummins,C.M., Culbertson,M.R. and Knapp,G. (1984) Submitted to Mol. Cell. Biol.