A Putative tRNA^{Trp} Gene Cloned from *Dictyostelium discoideum*: Its Nucleotide Sequence and Association with Repetitive Deoxyribonucleic Acid†

Dennis M. Peffley and Mitchell L. Sogin*

ABSTRACT: Using a total tRNA population labeled with ^{32}P, we have cloned a number of tRNA genes from *Dictyostelium discoideum*. A partial sequence of a cloned 1250-base-pair DNA insert, pDT-513, revealed the occurrence of a putative tRNA^{Trp} gene. In addition to the cloverleaf secondary structure, the tRNA^{Trp} gene contained all of the invariant and semiinvariant residues found in most tRNA sequences and has a 13-base-pair intron which is located one base removed from the 3′ residue of the anticodon. The genomic distribution of the tRNA gene and its flanking sequences was examined via Southern annealing experiments. The structural gene is represented on at least six *Eco*RI fragments in the *D. discoideum* genome. Sequences flanking the 5′ terminus of the cloned gene are repeated many times in the genome while the sequence flanking the 3′ terminus of the pDT-513 DNA insert structural tRNA gene is present only once in the genome.

The detailed structural analysis of "related gene families" in eucaryotes is of central importance to the understanding of gene expression and regulation. Through the use of recombinant DNA technology and DNA sequencing protocols, a picture for the organization of the "related family" of transfer RNA genes is beginning to emerge. The total number of tRNA genes per haploid genome ranges from 360 in yeast (Schweizer et al., 1969) to as many as 8000 in *Xenopus* (Birnstiel et al., 1972; Clarkson et al., 1973). Transfer RNA genes may be randomly dispersed throughout the genome or organized into clusters. A dispersed organization has been demonstrated for the tRNA^{Tyr} genes in yeast (Olson et al., 1977). Clusters of identical tRNA genes have been reported for the tRNA^{Val} and tRNA^{Lys} genes in *Drosophila* (Delaney et al., 1976; Grigliatti et al., 1974). Clusters of three different tRNA genes have been reported in a 3100-base-pair DNA repeat unit from *Xenopus* (Muller & Clarkson, 1980) and in a cloned DNA fragment from *Drosophila* (Schmidt et al., 1978). The clustering of tRNA genes might play a role in coordinate tRNA biosynthesis; however, multimeric tRNA precursors similar to those found in bacterial systems have not yet been identified. An alternative mechanism for the coordinate synthesis of multiple tRNAs might involve the association of similar or identical regulatory sequences with several tRNA genes. Data supporting this notion are very limited. With the exception of two tRNA^{Phe} genes in yeast (Valenzuela et al., 1978), homologous repetitive sequences which flank a series of transcriptionally related tRNA genes and which might serve in a regulatory capacity have not been reported.

We have initiated a study of tRNA gene organization in the slime mold *Dictyostelium discoideum*. The relatively small size of its genome (2% of that of a typical mammalian genome) greatly simplifies the screening procedure required to collect a variety of tRNA genes. Furthermore, the organization of the *D. discoideum* genome resembles that of higher eucaryotes; 15% of the DNA is represented in short, middle-reiterated sequences which are interspersed among single-copy DNA sequences (Firtel & Kindle, 1975). In this paper, we report the initial characterization of a presumptive tRNA^{Trp} gene from *D. discoideum*. The structural gene is repeated on at least six different genomic DNA fragments bounded by *Eco*RI restriction endonuclease sites. Sequences adjacent to the 5′ terminus of the tRNA gene are repeated many times in the genome. The DNA sequence flanking the 3′ terminus of the gene is represented once in the genome. Finally, as in the case of a tRNA^{Tyr} gene in *Xenopus laevis* (Muller & Clarkson, 1980) and in several yeast tRNA genes (Valenzuela et al., 1978; Goodman et al., 1977; Venegas et al., 1979; Etcheverry et al., 1979), the *D. discoideum* tRNA^{Trp} gene contains an intervening sequence in the anticodon loop.

Experimental Procedures

The restriction endonucleases *Hin*fI, *Alu*I, *Hae*III, *Mbo*I, *Mbo*II, and *Taq*I were purchased from New England BioLabs. *Eco*RI endonuclease was isolated as described by Greene et al. (1974), omitting the DEAE-cellulose step. Polynucleotide kinase and RNA ligase were isolated from bacteriophage T4 am E4314 infected *Escherichia coli* as described by Uhlenbeck & Cameron (1977) and Walker et al. (1975), respectively. Carrier-free [γ-^{32}P]ATP was prepared by using the procedures of Reeve and Huang (Johnson & Walseth, 1978). HTP[1] and low gel temperature agarose were purchased from Bio-Rad Laboratories.

Cloning. *D. discoideum* nuclear DNA was prepared as previously described and inserted into the *Eco*RI site of the plasmid pMB9 (Sogin & Olsen, 1980). Recombinant plasmids carrying the tetracycline resistance gene were used to transform the tetracycline-sensitive host *E. coli* HB101. Clones were grown on tetracycline containing L agar and screened for tRNA genes by using a slight modification of the colony hybridization technique (Grunstein & Hogness, 1975). Transfer RNA coding sequences were located by incubating the colony hybridization filters with an in vitro ^{32}P-labeled tRNA probe in 5 × SSC,[1] 50% deionized formamide, and 0.5% NaDodSO$_4$ at 42 °C for 16 h. The filters were washed

† From the Department of Molecular and Cellular Biology, National Jewish Hospital and Research Center, Denver, Colorado 80206. *Received December 16, 1980.* This investigation was supported by a grant from the National Institutes of Health (GM23464).

[1] Abbreviations used: HTP, hydroxylapatite–Bio-Gel HTP; SSC, 0.15 M NaCl and 0.015 M sodium citrate; NaDodSO$_4$, sodium dodecyl sulfate; Bis, *N*,*N*′-methylenebis(acrylamide); E buffer, 40 mM Tris/HOAc, pH 7.0, and 1 mM EDTA; Hepes, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PO$_4$ buffer, 50 mM sodium phosphate, pH 6.9, and 1 mM EDTA; Tris, tris(hydroxymethyl)aminomethane; EDTA, ethylenediaminetetraacetic acid.

0006-2960/81/0420-4015$01.25/0

30 min in hybridization buffer at 42 °C, which was followed by two 30-min washes in 2 × SSC at room temperature. Colonies capable of hybridizing with the ^{32}P-labeled tRNA probe were identified by autoradiography using intensifying screens at −70 °C. All work was carried out under P2-EK1 containment conditions specified by the NIH Guidelines for Recombinant DNA Research.

Preparation of the tRNA Hybridization Probe. Total cytoplasmic RNAs were isolated as previously described (Sogin & Olsen, 1980) and were fractionated by electrophoresis through 8% polyacrylamide–NaDodSO$_4$[1] gels (Bis[1]/acrylamide ratio 1:20) in E buffer.[1] The region of the gels containing 4S-sized RNA was located by the "UV shadowing" technique (Hassur & Whitlock, 1974). Subsequent to elution from the gel and concentration by ethanol precipitation, the 4S-size RNA molecules were labeled at their 3′ termini with cytidine [5′-^{32}P]bisphosphate by using RNA ligase (Stahl et al., 1979). The intermolecular reaction catalyzed by RNA ligase was carried out in a mixture (100 μL) containing 50 mM Hepes/NaOH,[1] pH 8.3, 10 mM $MgCl_2$, 5 mM dithiothreitol, 1.2 nmol of ATP, 0.3 nmol of cytidine [5′-^{32}P]bisphosphate, 0.3 nmol of tRNA, and 25 units of RNA ligase. After incubation for 1.5 h at 37 °C, the reaction was terminated by addition of NaDodSO$_4$ to 0.5%, NaOAc to 0.2 M, and EDTA to 20 mM. Subsequent to ethanol precipitation, residual cytidine [5′-^{32}P]bisphosphate was removed by passage through a Sephadex G-50 column equilibrated in 50 mM Tris-HCl, pH 7.2, 0.3 M NaOAc, and 1 mM EDTA. After concentration by ethanol precipitation, the specific radioactivity of the tRNA probe was found to be greater than 10^7 cpm/μg of RNA.

Gel Electrophoresis and Preparation of DNA Fragments. DNA fragments greater than 150 base pairs in length were isolated by fractionation in 0.75% or 1.5% low gel temperature agarose slab gels built in 1 × PO$_4$ buffer.[1] Subsequent to electrophoresis in 0.5 × PO$_4$ buffer, the region of the gel (typically 5 × 3 × 20 mm) containing the DNA fragment, as identified by "UV shadowing" or autoradiography, was excised and placed in a vial containing 20 mL of 20 mM sodium phosphate, pH 6.9, buffer. The gel was melted by heating to 65 °C and cooled to 42 °C prior to loading an HTP column equilibrated in 20 mM sodium phosphate, pH 6.9, buffer at 42 °C. Subsequent to washing with 20 mM phosphate buffer, the DNA was eluted from the HTP with 0.6 M sodium phosphate, pH 6.9, and desalted on a Sephadex G-50 column equilibrated with 50 mM Tris-HCl, pH 7.2, 0.3 M NaOAc, and 1 mM EDTA buffer. Fractions containing the DNA were pooled and concentrated by ethanol precipitation. DNA fragments smaller than 150 nucleotides were resolved in either 5% or 8% acrylamide gels which were built and run in 50 mM Tris–borate, pH 8.2, and 1 mM EDTA electrophoresis buffer. DNA fragments in the gel were located by "UV shadowing" or autoradiography, recovered by elution into electrophoresis buffer, and then concentrated by ethanol precipitation.

RNA/DNA Hybridization and DNA/DNA Annealing. Prior to immobilization on cellulose nitrate strips (Schleicher & Schüll BA85), *D. discoideum* genomic DNA *Eco*RI limit digests were fractionated in 0.75% standard agarose gels which were built in 1 × PO$_4$ and electrophoresed in 0.5 × PO$_4$ buffer. Southern transfers were prepared as previously described (Southern, 1975). RNA hybridizations to the immobilized DNA were carried out at 37 °C in 5 × SSC, 50% deionized formamide, and 0.5% NaDodSO$_4$ containing 50 μg of *Bacillus subtilis* RNA per mL of hybridization solution to reduce backgrounds. Subsequent to hybridization for 24 h, the Southern strips were washed for 30 min in hybridization buffer and then twice for 30 min in 2 × SSC at room temperature. Prior to DNA annealing experiments, the Southern transfer strips were soaked in a solution containing 6 × SSC, 0.05 M piperazine-*N*,*N*′-bis(2-ethanesulfonic acid), pH 7.8, 100 mM potassium phosphate, pH 7.9, and 1 × Denhardt's solution (Denhardt, 1966) for 12 h at 65 °C. Annealing was carried out at 65 °C in a buffer containing 5 × SSC, 100 mM potassium phosphate, pH 7.0, 1 × Denhardt's solution, and 1 μg/mL sonicated calf thymus DNA. After incubation for 18 h, the strips were washed twice for 20 min in 4 × SSC, 1 × Denhardts, and 0.5% NaDodSO$_4$ at 65 °C followed by two 30-min washes in 2 × SSC, 1 × Denhardts, and 0.5% NaDodSO$_4$ at 65 °C, and finally two 30-min washes in 2 × SSC at room temperature. Radioactivity in the filters or Southern strips was located by autoradiography using intensifying screens at −70 °C.

DNA Sequencing. Restriction fragments labeled at their 5′ termini with [γ-^{32}P]ATP and polynucleotide kinase were either separated into single strands on polyacrylamide gels or restricted with a second enzyme and refractionated on agarose or polyacrylamide gels. The resulting DNA fragments labeled at one end were sequenced according to the chemical procedures of Maxam & Gilbert (1980). The hydrazine reaction in the presence or absence of 1 M NaCl was used for site-specific cleavage at dCyd and dCyd plus dThd. The dimethyl sulfate reaction was used for dGuo specific chemistries. Alkali treatment at 90 °C was used for the dAdo > dCyd cleavages. The products of the sequencing reactions were resolved on 25%, 20%, or 8% gels (Air et al., 1976).

Results

A total "end-labeled", 4S RNA population was used as a probe in the colony hybridization procedure to identify recombinant pMB9 clones harboring *D. discoideum* tRNA genes. Initial attempts to screen recombinant clones by using an in vivo ^{32}P-labeled tRNA probe were unsuccessful due to the low specific radioactivities of the tRNA. As an alternative, a 4S RNA probe was prepared by ligating cytidine [5′-^{32}P]-bisphosphate to the 3′ termini of a gel-purified *D. discoideum* tRNA population by using RNA ligase from T4 infected *E. coli*. This procedure circumvents the difficulty inherent with the in vitro labeling of recessed, duplex 5′ termini using [γ-^{32}P]ATP and polynucleotide kinase. The probe was judged to be representative of a normal tRNA population in that a two-dimensional gel electrophoretic characterization of the probe was essentially identical with that of a uniformly ^{32}P-labeled tRNA population isolated from vegetative cells. Using this probe, we have detected transfer RNA coding sequences in 2% of our clones. Each of these recombinant clones has been shown to contain at least one *Eco*RI fragment capable of strongly hybridizing to the in vitro ^{32}P-labeled tRNA probe in Southern transfer experiments. One of the clones, pDT-513, was selected for further analysis.

The pDT-513 clone contains a *D. discoideum* DNA insert which is 1250 base pairs long. A restriction map for the insert (Figure 1) was derived via agarose and acrylamide gel analyses of single and double restriction endonuclease digestions as well as single enzyme digests of the DNA insert labeled in vitro with [γ-^{32}P]ATP at the *Eco*RI termini. The location of the putative tRNA structural gene in the restriction map was initially determined by a primary structure analysis. Initially, the sequence of one strand was determined for nucleotide positions −725 through −453 and −134 through +525 (nucleotide position +1 is coincident with the 5′-terminal nu-

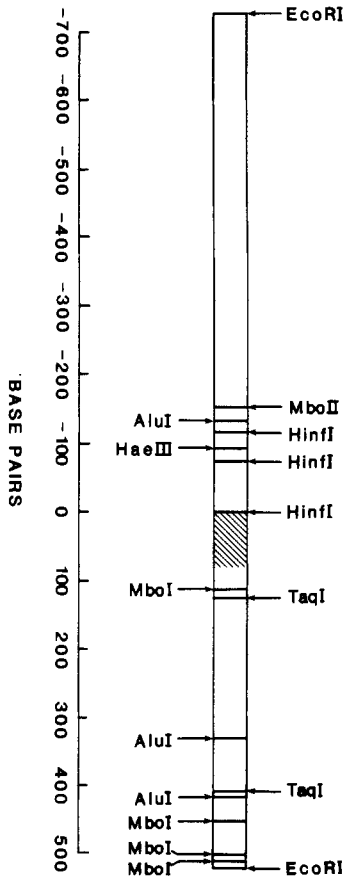

FIGURE 1: Restriction endonuclease map of the 1250-base pair pDT-513 passenger DNA. The pDT-513 DNA insert was challenged with 20 restriction endonucleases. The cleavage sites of the six enzymes capable of digesting the DNA fragment were determined by agarose or polyacrylamide gel analyses of single and double digests of the pDT-513 DNA insert. The identity of the terminal restriction fragments was verified by analyzing digests of passenger molecules labeled with ^{32}P at their *Eco*RI termini. *Tac*I, *Hin*fI, or *Taq*I digests of pBR322 plasmid DNA were used for molecular weight markers to determine the size of the restriction fragments. The positions of the cleavage sites were subsequently confirmed by DNA sequence analysis. The location of the putative tRNA^{Trp} gene was determined by a computer-assisted secondary structure analysis of the DNA sequence and is indicated by the crosshatching. Position +1 is coincident with the 5′-terminal nucleotide of the nontranscribed strand of the tRNA gene.

cleotide of the noncoding strand of the tRNA gene). A computer-assisted secondary structure analysis revealed the existence of a single tRNA-like gene in the DNA insert. Location of other tRNA genes in the unsequenced 329-base-pair region bounded by positions −454 and −133 is unlikely in that the DNA bounded by residues −725 through −154 is completely resistant to cleavage by a battery of restriction endonucleases, indicating an AT-rich stretch of DNA (Figure 1). Direct sequence analysis confirmed the AT-rich DNA composition between positions −725 and −453.[2] It is significant that the restriction endonucleases *Taq*I (specificity d-pT-C-G-A) and *Hin*fI (specificity d-pG-A-N-T-C) failed to cleave the DNA bounded by positions −725 through −154. Most tRNA genes contain the sequence d-pT-C-G-A-N-T-C in the "common arm" loop. Finally, a Southern hybridization analysis of a double digest of the DNA insert with *Alu*I and *Mbo*II failed to display homology between the in vitro ^{32}P-labeled tRNA probe and the 570-base-pair fragment defined by the *Eco*RI terminus at position −725 and the *Mbo*II site at position −154.[2]

Subsequent to identifying the location of a putative tRNA gene, both strands of the region coding for the tRNA gene in the cloned DNA insert were sequenced. Fragments labeled at only one 5′ terminus were prepared by digesting the pDT-513 DNA insert with the restriction endonuclease *Hin*fI, labeling the 5′ termini with [γ-^{32}P]ATP using polynucleotide kinase, and redigesting with *Alu*I. The resulting "end-labeled" fragment bounded by the *Hin*fI site at position +1 and the *Alu*I site at +330 was isolated by agarose gel electrophoresis and subjected to sequence analysis.[2] The opposite strand was sequenced in a similar fashion by digesting the DNA insert with *Taq*I, labeling the 5′ termini as above, and redigesting with *Alu*I. The fragment defined by the *Taq*I site at position +124 and the *Alu*I site at −134 was isolated in an 8% acrylamide gel and sequenced by using the chemical procedures of Maxam & Gilbert (1980). Sequencing both strands was required because methylated dCyd residues are not modified in the hydrazine reaction and as a result may be misinterpreted as dAdo residues. Figure 2 shows the sequence of the tRNA gene as well as the DNA regions which flank the 5′ and 3′ termini of the structural gene. The probable secondary structure of the theoretical transcript of nucleotides 1–87 is shown in Figure 3 (see paragraph at end of paper regarding supplementary material). Three lines of evidence support the presumptive identification of this DNA sequence as a tRNA gene; they are the observation that under stringent hybridization conditions the DNA insert hybridizes with an RNA sequence in our tRNA probe, the "cloverleaf" secondary structure which is characteristic of all tRNA molecules, and the presence of nucleotide residues which are common to all tRNA molecules (Gauss et al., 1979). These primary and secondary structural characteristics are found if the molecule is drawn with an intron in the anticodon loop as shown in Figure 3. The encircled residues indicate invariant or semi-invariant residues. A 3′-terminal C-C-A$_{OH}$ sequence common to mature tRNA molecules is not coded by the genome and is presumably added posttranscriptionally, as in the case of other eucaryotic tRNA sequences (McClain, 1977).

Preliminary RNA/DNA hybridization data indicate that there are approximately 1000 tRNA genes in the *D. discoideum* genome.[3] Three organizational patterns are possible for the reiterated tRNA genes, i.e., "homoclustering" of isoaccepting species, "heteroclustering" of different tRNA sequences, and random or dispersed distributions for individual tRNA genes. A possible clustering of the tRNA^{Trp} with additional tRNA genes separated by spacers longer than 500 base pairs cannot be dismissed; however, the existence of only one tRNA-like sequence in the pDT-513 DNA supports the notion that this putative tRNA^{Trp} gene is not closely linked by very short spacer sequences with other tRNA genes in the *D. discoideum* genome. Further information regarding the organization of tRNA^{Trp} genes was obtained by annealing the ^{32}P-labeled pDT-513 clone to Southern transfers of *Eco*RI-digested *D. discoideum* DNA. Lanes 1 and 3 in Figure 4 are ethidium bromide staining patterns of the pDT-513 DNA insert and *Eco*RI-digested *D. discoideum* DNA. Intense staining bands in lane 3 represent reiterated sequences in the genome bounded by *Eco*RI restriction sites. Lane 2 of Figure 4 is a Southern analysis in which ^{32}P-labeled pDT-513 DNA was annealed to *Eco*RI-digested *D. discoideum* DNA. The unexpected result demonstrates homology between at least a portion of the pDT-513 DNA insert and a large number of *Eco*RI fragments in the *D. discoideum* genome. This ob-

[2] D. M. Peffley and M. L. Sogin, unpublished data.

[3] G. J. Olsen and M. L. Sogin, unpublished data.

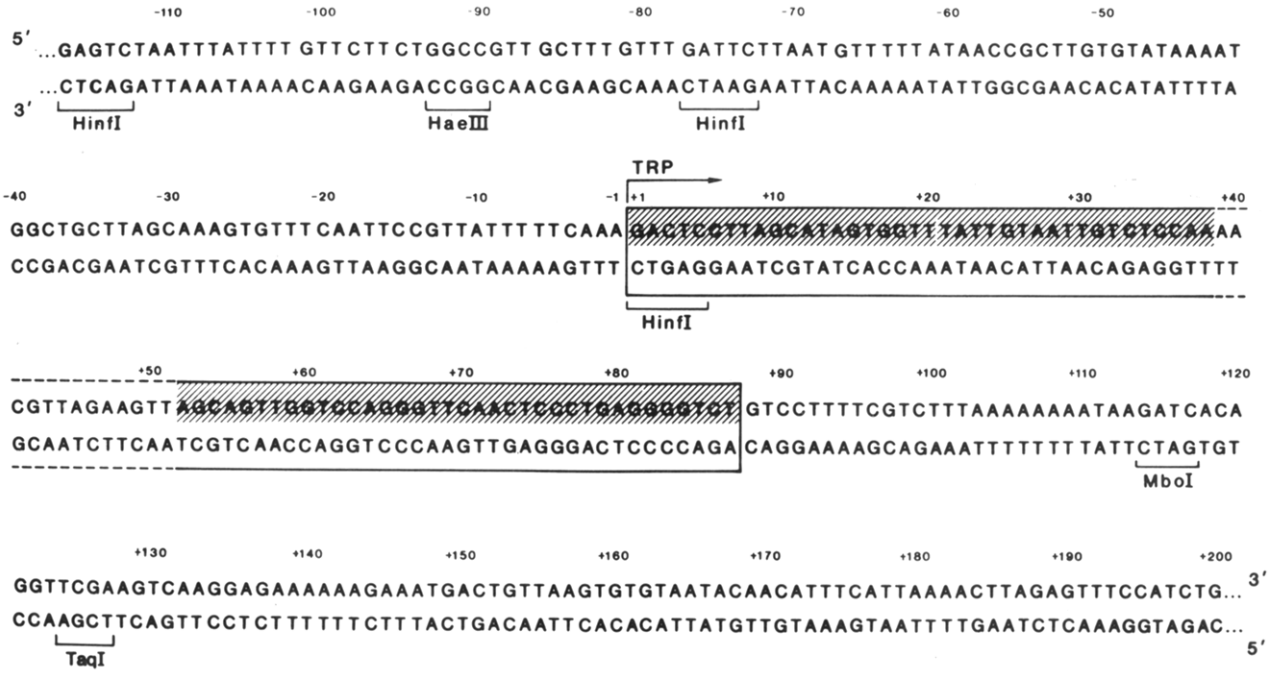

FIGURE 2: Nucleotide sequence of the putative $tRNA^{Trp}$ structural gene and its flanking sequences. The tRNA gene is enclosed in solid line boxes with crosshatched areas representing the noncoding strands of the tRNA gene. The intervening sequence is enclosed by dashed lines. Position number +1 is coincident with the 5′-terminal nucleotide of the nontranscribed strand of the tRNA gene.

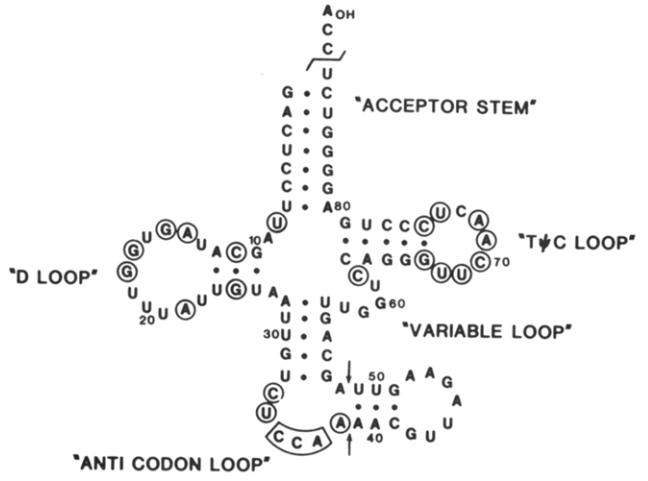

FIGURE 3: Secondary structure model of the theoretical transcript of a *Dictyostelium discoideum* $tRNA^{Trp}$ gene. The transcript of the $tRNA^{Trp}$ gene defined by residues +1–87 in the pDT-513 DNA insert is drawn in a "cloverleaf" arrangement typical of transfer RNA molecules. The invariant and semiinvariant residues are encircled, and the anticodon is enclosed within solid lines. The arrows in the anticodon loop indicate the proposed "splicing" sites for excising the intron. The 3′-terminal C-C-A_{OH} is not encoded by the gene and is presumably added posttranscriptionally.

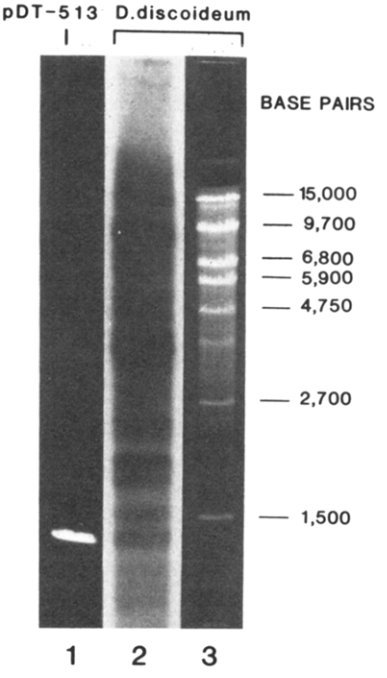

FIGURE 4: Size of *Eco*RI fragments of the *Dictyostelium discoideum* genome which are complementary to the pDT-513 DNA insert. A Southern transfer of *Eco*RI-digested *Dictyostelium discoideum* DNA was prepared and annealed with intact pDT-513 passenger labeled in vitro with [γ-^{32}P]ATP at the *Eco*RI termini. Lane 1 is an ethidium bromide stain of the pDT-513 DNA insert. Lane 3 shows an ethidium bromide stain of the *D. discoideum Eco*RI digest. Intensely staining bands represented repeated DNA sequences in the *D. discoideum* genome. Lane 2 is an autoradiogram of the ^{32}P-labeled pDT-513 probe annealed to a Southern transfer of *Eco*RI-digested *D. discoideum* genomic DNA.

servation is consistent with the presence of many different tRNA genes in the cloned pDT-513 DNA insert, a random distribution of many $tRNA^{Trp}$ sequences in the genome, a highly conserved (i.e., cross-hybridizable) primary structure for several *D. discoideum* tRNAs, or the presence of a repetitive DNA sequence within the pDT-513 DNA insert. As discussed above, the possibility that many different tRNA genes reside in the pDT-513 clone giving rise to a complex annealing pattern to the Southern transfer is ruled out by the restriction endonuclease and DNA sequencing analyses. The other possibilities were investigated by using small ^{32}P-end-labeled restriction fragments from the pDT-513 clone to probe Southern transfers of *D. discoideum* genomic DNA digested with *Eco*RI endonuclease. These probes include the 330-nucleotide fragment (positions +1 to +330) containing the $tRNA^{Trp}$ gene as well as fragments proximal to the 5′ and 3′ termini of the 330-nucleotide fragment. Lanes 1, 3, and 7 in Figure 5 are Southern images of intact pDT-513 DNA annealed to immobilized *D. discoideum Eco*RI digested DNA. The Southern image shown in lane 2 of Figure 5A for the

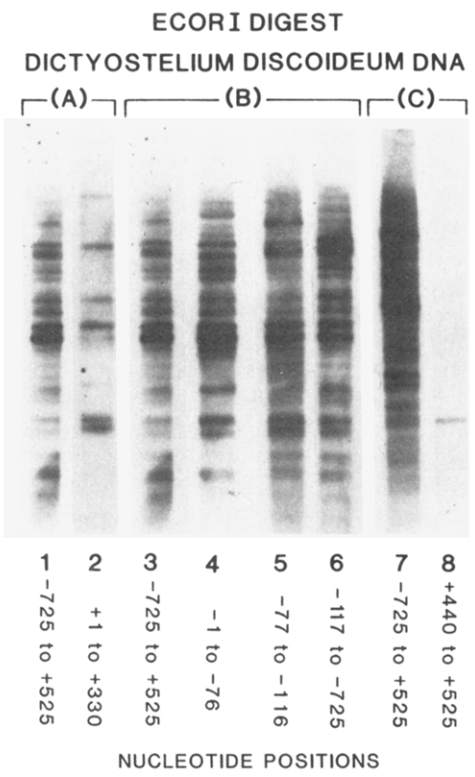

FIGURE 5: Size of *Eco*RI fragments in the *Dictyostelium discoideum* genome which are complementary to restriction fragments isolated from pDT-513. Restriction fragments coding for the tRNA^{Trp} gene (A), sequences flanking the 5′ end of the structural tRNA gene (B), and the sequence flanking the 3′ terminus of the tRNA gene (C) were labeled at their 5′ termini and isolated as described in the text. Autoradiograms show the annealing patterns of the probes to Southern transfers of *Eco*RI-digested *D. discoideum* DNA. The nucleotide position refers to the restriction map and the sequence displayed in Figures 1 and 2, respectively. Lanes 1, 3, and 7 in panels A, B, and C are control experiments which display annealing of the intact pDT-513 probe to *Eco*RI-digested *D. discoideum* DNA. Lane 2 in panel A was probed with the restriction fragment (positions 1–330) containing the tRNA^{Trp} structural gene. Lanes 4, 5, and 6 in panel B were probed with the 76-base-pair *Hin*fI fragment (positions −1 to −76), the 40-base-pair *Hin*fI fragment (positions −77 to −116), and the 609-base-pair fragment defined by the *Eco*RI site at −725 and the *Hin*fI site at −117, respectively. Lane 8 in panel C was probed with the sequence defined by the *Eco*RI site at +525 and the *Alu*I site at +440.

330-nucleotide fragment containing the tRNA structural gene demonstrates strong homology with six *Eco*RI bands and, to a lesser degree of homology, with two additional bands, as indicated by a significantly weaker autoradiographic signal. Direct sequence analysis of the "end-labeled" fragment did not reveal contamination with other DNA sequences which might result in the probe annealing to several *Eco*RI-generated bands. Hybridization of the tRNA^{Trp} DNA to only six *Eco*RI genomic DNA bands is not consistent with all *D. discoideum* tRNA genes containing stringently conserved sequences. Furthermore, this result rules out the possibility of a random distribution throughout the genome of highly reiterated tRNA^{Trp} genes.

Restriction fragments flanking the tRNA gene were analyzed in a similar fashion. The 76-nucleotide fragment (positions −1 to −76) adjacent to the 5′ terminus of the adjacent tRNA gene, the 40-nucleotide fragment (positions −77 to −116), and the terminal 609-base-pair fragment (positions −117 to −725) were labeled at their 5′ termini with [γ-^{32}P]ATP and annealed to a Southern transfer of the *Eco*RI-digested *D. discoideum* genomic DNA. These fragments displayed strong homology with a large number of *Eco*RI-generated DNA fragments (Figure 5B, lanes 4–6). Subsequent to strand separation, the 76- and 40-base-pair fragments were shown, by direct sequence analysis, to be uncontaminated with other DNA fragments which might have given rise to the annealing with many *Eco*RI fragments. Similarly the 609-base-pair fragment labeled at only the −725 position was partially sequenced and found to be uncontaminated with other DNA fragments. Finally, the fragment located adjacent to the 3′ terminus of the tRNA gene including the *Eco*RI site at +525 was labeled in vitro with [γ-^{32}P]ATP and polynucleotide kinase. It annealed to a single *Eco*RI band which comigrates on gels with the intact pDT-513 DNA insert (Figure 5C, lane 8).

Several artifacts, including contaminated fragment preparations or incomplete digestion of the genomic DNA used in preparing the Southern transfer strip, might result in the multiple-band annealing patterns observed with the short DNA fragments. However, the direct sequence analysis of the 5′-proximal probes demonstrated their purity. Incomplete digestion of the *D. discoideum* DNA with *Eco*RI endonuclease cannot be responsible for the multiple band annealing pattern because the sequences beyond the 3′ terminus of the tRNA^{Trp} gene anneal with a single band and the limited annealing pattern of the tRNA gene itself, i.e., six bands, would not have been observed. These results are consistent with the following interpretations: the putative tRNA^{Trp} gene is present on a minimum of six *Eco*RI restriction fragments in the *D. discoideum* genome, a repetitive DNA sequence exceeding 100 base pairs in length is adjacent to the 5′ terminus of the tRNA gene, and the sequence adjacent to the 3′ terminus of the gene apparently is present only once in the genome.

Discussion

The sequence analysis of pDT-513 DNA has demonstrated the presence of a putative tRNA^{Trp} gene. This identification is based on the ability of the cloned insert to hybridize with an in vitro ^{32}P-labeled tRNA probe, its content of a nucleotide sequence which can be folded to resemble a tRNA "cloverleaf" pattern containing a tryptophan anticodon, and the presence of all of the invariant and semiinvariant nucleotide residues characteristic of tRNA sequences. The putative tRNA gene does not contain the 3′-terminal C-C-A_{OH}, which is typically added posttranscriptionally to eucaryotic tRNA gene transcripts. The presumptive tRNA gene contains an intron 13 bases long, which is located one base removed from the 3′ residue of the anticodon. No detectable homology is observed between this intron and those reported in yeast tRNA sequences. Unlike intervening sequences reported in yeast tRNA anticodon loops, the *D. discoideum* tRNA^{Trp} intron does not contain a sequence complementary to the anticodon. Introns in tRNA anticodon loops have only been reported in several yeast tRNA precursors (Goodman et al., 1977; Etcheverry et al., 1979). It is likely, however, that introns in tRNA precursors for other eucaryotes will be found. Direct sequence analysis of a cloned 3100-base-pair fragment from *Xenopus* revealed the presence of a putative tRNA^{Trp} gene which contains a different 13-base intron in the anticodon loop (Schmidt et al., 1978). Furthermore, when the yeast tRNA^{Phe} gene, which contains an intron in the anticodon loop, is transcribed in vitro by using a *Xenopus* germinal vesicle extract, the intervening sequence is removed from the primary transcript by a processing activity in the enzyme preparation (Silverman et al., 1979). The presence of the excision activity supports the notion that at least in *Xenopus* there may be intervening sequences in tRNA anticodon loops of tRNA precursors.

The inferred tRNA structure is 61% homologous with bovine $tRNA^{Trp}$ and 54% homologous with yeast $tRNA^{Trp}$ (Gauss et al., 1979). The homologies are most evident in the anticodon loop and the TΨC loop of yeast and in the acceptor stem of bovine $tRNA^{Trp}$. The size of the variable loop is identical in all three organisms. A major difference is in the presence of 12 unpaired bases in the *D. discoideum* $tRNA^{Trp}$ "D" loop rather than the 7 or 8 found in most tryptophanyl tRNAs. The homology levels of 55–60% with other eucaryotic tryptophanyl tRNAs are consistent with the findings of very low ribosomal RNA homologies between *D. discoideum* 17S rRNA and yeast 17S rRNA or L cell 18S rRNA.[4] Southern analysis of the 330-base-pair region containing the putative tRNA gene reveals strong homology with at least six *Eco*RI-generated *D. discoideum* fragments. (A minimum of 80% homology between complementary sequences is required for duplex formations in 5 × SSC at 65 °C.) This result suggests the presence of at least six $tRNA^{Trp}$ genes in the genome; however, the possibility of multiple $tRNA^{Trp}$-like genes in five of the six *Eco*RI generated fragments cannot be ruled out. The ability of the region of the pDT-513 clone which contains the putative $tRNA^{Trp}$ gene to hybridize 4S-size RNA under stringent conditions and to anneal strongly with several *Eco*RI bands in the *D. discoideum* genome supports the interpretation that this sequence represents a functional tRNA gene; it is unlikely that six nonfunctional tRNA-like genes would be conserved in the genome. However, in the absence of an RNA sequence for the cytoplasmic tRNA or its precursor, it is not possible to unequivocally identify the pDT-513 DNA as a functional $tRNA^{Trp}$ gene; this is the case for any cloned, multiple-copy gene. Finally, we cannot rule out the existence of "pseudogenes", as have been found for *Xenopus* 5S rRNA genes (Jacq et al., 1977).

The annealing of pDT-513 DNA to a large number of DNA fragments generated by *Eco*RI in the *D. discoideum* genome demonstrates the presence of reiterated DNA sequences in the cloned insert. Six additional cloned *D. discoideum* DNA sequences, capable of hybridizing the bulk ^{32}P-labeled tRNA probe, revealed nearly identical multiple-band annealing patterns with *Eco*RI-digested *D. discoideum* genomic DNA. Furthermore, the general annealing pattern for all of the clones resembled the complex distribution of radioactivity obtained when a total in vitro ^{32}P-labeled tRNA probe was hybridized to similar Southern transfers of *D. discoideum Eco*RI digested genomic DNA.[2] There are three potentially interesting explanations for these observations. The first is that tRNA genes are arranged in heteroclusters in the *D. discoideum* genome. The assemblage of tRNA genes may vary from heterocluster to heterocluster. This could give rise to multiple-band annealing patterns for each of the cloned tRNA coding sequences. However, in at least one of these clones, pDT-513, the direct DNA structural analysis does not support this model. A computer-assisted structural analysis revealed the presence of only the $tRNA^{Trp}$ gene. A second possibility is that many *D. discoideum* tRNA species share extensive, common sequences. This might result in significant annealing overlap between different tRNA genes. This interpretation is unlikely, because the 330-base-pair region of the pDT-513 DNA insert containing the $tRNA^{Trp}$ gene anneals to only about six *Eco*RI-generated *D. discoideum* DNA fragments. Under these conditions, the invariant and semiinvariant sequences common to all tRNA species would not be sufficient for duplex formation. The final and most interesting possibility is that common sequences occur adjacent to many tRNA genes. Annealing a ^{32}P-labeled recombinant plasmid containing a tRNA gene and its adjacent sequences to an *Eco*RI digest of the total genome would therefore give rise to an annealing pattern similar to that observed with a hybridization probe comprised of the bulk ^{32}P-labeled tRNA population.

The multiple annealing pattern exhibited by the short DNA fragments which lie immediately adjacent to the 5′ terminus of the tRNA gene in pDT-513 demonstrates the association of reiterated DNA sequences with at least one tRNA gene. The total length of the reiterated sequences exceeds 100 base pairs and is not simply an extremely AT-rich region of DNA similar to that reported at the 5′ termini of the *D. discoideum* actin genes (Firtel et al., 1979). (It is conceivable that random AT-rich sequences might give rise to multiple annealing patterns due to homology with other AT-rich regions of the genome.) The association of specific classes of reiterated DNA sequences with tRNA genes is supported by comparing Southern images of *Eco*RI-digested *D. discoideum* DNA annealed to either the reiterated DNA fragments flanking the 5′ terminus of the tRNA gene or the 330-base-pair fragment containing the structural tRNA gene. Included in the spectrum of fragments which annealed to the reiterated probes are bands which display mobilities identical with those which hybridize the putative $tRNA^{Trp}$ structural gene. This would suggest that a highly similar reiterated DNA sequence is linked with each of the *D. discoideum* $tRNA^{Trp}$ genes. However, it will be necessary to clone and isolate each of the $tRNA^{Trp}$ genes from *D. discoideum* and explore the character of sequences flanking those structural genes to confirm these speculations. The complex annealing patterns of the other tRNA clones suggest the association of a reiterated DNA sequence or series of reiterated DNA sequences with many *D. discoideum* tRNA genes. Once again this hypothesis can be tested by identifying other examples of cloned DNA sequences which contain tRNA genes and specific classes of reiterated DNA sequences.

The interspersion of reiterated sequences among single-copy or structural genes present in low copy number in the genome has been reported for a number of organisms. These repeated elements may play a role in the coordinate control of genes which are widely dispersed in the genome, for example, by serving as targets for regulatory molecules (Britten & Davidson, 1979). Thus, the transcriptional control of several or hundreds of genes may be effected by a single species of regulatory molecule. The economy of coordinately controlling gene expression by a limited number of regulatory elements is self-evident. The reiterated sequences adjacent to tRNA genes may play such a role by regulating the transcription of these genes or RNA processing of tRNA precursors. DeFranco et al. (1980) have demonstrated that sequences which flank the 5′ termini of $tRNA^{Lys}$ genes isolated from *Drosophila* attenuate their in vitro transcription efficiencies. In contrast, the 5S rRNA genes which are also transcribed by polymerase III have a control region which is within the RNA coding sequence (Sakonju et al., 1980; Bogenhagen et al., 1980). However, this does not rule out the possibility of 5′ flanking sequences attenuating in vitro transcription. To date, there is no evidence from in vitro studies that the reiterated sequences flanking the putative $tRNA^{Trp}$ gene in *D. discoideum* play a role in transcriptional control.

Supplementary Material Available

Documentation for the sequence shown in Figure 3 (6 pages). Ordering information is given on any current masthead page.

[4] M. L. Sogin and C. R. Woese, unpublished data.

References

Air, G. M., Sanger, F., & Coulson, A. R. (1976) *J. Mol. Biol. 108*, 519–533.

Birnstiel, M. L., Sells, B. H., & Purdom, I. F. (1972) *J. Mol. Biol. 63*, 21–39.

Bogenhagen, D. F., Sakonju, S., & Brown, D. D. (1980) *Cell (Cambridge, Mass.) 19*, 27–35.

Britten, R. J., & Davidson, E. H. (1979) *Science (Washington, D.C.) 204*, 1052–1059.

Clarkson, S. G., Birnstiel, M. L., & Serra, V. (1973) *J. Mol. Biol. 79*, 391–410.

DeFranco, D., Schmidt, O., & Söll, D. (1980) *Proc. Natl. Acad. Sci. U.S.A. 77*, 3365–3368.

Delaney, A., Dunn, R., Grigliatti, T. A., Tener, G. M., Kaufman, T. C., & Suzuki, D. T. (1976) *Fed. Proc., Fed. Am. Soc. Exp. Biol. 35*, 1976.

Denhardt, D. T. (1966) *Biochem. Biophys. Res. Commun. 23*, 641–646.

Etcheverry, T., Colby, D., & Guthrie, C. (1979) *Cell (Cambridge, Mass.) 18*, 11–26.

Firtel, R. A., & Kindle, K. (1975) *Cell (Cambridge, Mass.) 5*, 401–411.

Firtel, R. A., Timm, R., Kimmel, A. R., & KcKeown, M. (1979) *Proc. Natl. Acad. Sci. U.S.A. 76*, 6206–6210.

Gauss, D. H., Gruter, F., & Sprinzl, M. (1979) *Nucleic Acids Res. 6*, R1–R19.

Goodman, M. M., Olson, M. V., & Hall, B. D. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5453–5457.

Greene, P. J., Betlach, M. C., & Bozer, H. W. (1974) in *Methods in Molecular Biology* (Wickner, R. B., Ed.) pp 87–111, Marcel Dekker, New York.

Grigliatti, T. A., White, B. N., Tener, G. M., Kaufmann, T. C., & Suzuki, D. T. (1974) *Proc. Natl. Acad. Sci. U.S.A. 71*, 3527–3531.

Grunstein, M., & Hogness, D. S. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 3961–3965.

Hassur, S. M., & Whitlock, H. W., Jr. (1974) *Anal. Biochem. 59*, 162–164.

Jacq, C., Miller, J. R., & Brownlee, G. G. (1977) *Cell (Cambridge, Mass.) 12*, 109–120.

Johnson, R. A., & Walseth, T. F. (1978) in *Advances in Cyclic Nucleotide Research* (Brooker, G., Greengard, P., & Robinson, G. A., Eds.) pp 135–167, Raven Press, New York.

Maxam, A. M., & Gilbert, N. (1980) *Methods Enzomol. 65*, 499–560.

McClain, W. H. (1977) *Acc. Chem. Res. 10*, 418–425.

Muller, E., & Clarkson, S. G. (1980) *Cell (Cambridge, Mass.) 19*, 345–353.

Olson, M. V., Montgomery, D. L., Hopper, A. K., Page, G. S., Horodyski, F., & Hall, B. D. (1977) *Nature (London) 267*, 639–641.

Sakonju, S., Bogenhagen, D. F., & Brown, D. D. (1980) *Cell (Cambridge, Mass.) 19*, 13–25.

Schmidt, T., Egg, A. H., & Kubli, E. (1978) *Mol. Gen. Genet. 164*, 249–254.

Schweizer, E., MacKechnie, C., & Halvorson, H. O. (1969) *J. Mol. Biol. 40*, 261–277.

Silverman, S., Schmidt, O., Söll, D., & Hovermann, B. (1979) *J. Biol. Chem. 254*, 10290–10294.

Sogin, M. L., & Olsen, G. J. (1980) *Gene 8*, 231–238.

Southern, E. M. (1975) *J. Mol. Biol. 98*, 503–517.

Stahl, D. A., Walker, T. A., Meyhack, B., & Pace, N. R. (1979) *Cell (Cambridge, Mass.) 18*, 1133–1143.

Uhlenbeck, O. C., & Cameron, V. (1977) *Nucleic Acids Res. 4*, 85–98.

Valenzuela, P., Venegas, A., Weinberg, F., Bishop, R., & Rutter, W. J. (1978) *Proc. Natl. Acad. Sci. U.S.A. 75*, 190–194.

Venegas, A., Quivoga, M., Zaldivan, J., Rutter, W. S., & Valenzuela, P. (1979) *J. Biol. Chem. 254*, 12306–12309.

Walker, G. C., Uhlenbeck, O. C., Bedows, E., & Gumport, R. I. (1975) *Proc. Natl. Acad. Sci. U.S.A. 72*, 122–126.