Synthetic deoxyoligonucleotides as general probes for chloroplast transfer RNA genes

Jac A. Nickoloff and Richard B. Hallick

Department of Chemistry, University of Colorado, Boulder, CO 80309, USA

Received 22 July 1982; Revised and Accepted 19 November 1982

ABSTRACT

The utility of chemically synthesized deoxyoligonucleotides as hybridization probes for the detection of tRNA genes has been examined. Chloroplast tRNA genes were chosen for this study. Deoxyoligonucleotides complementary to highly conserved regions of chloroplast tRNA genes of both higher plants and *Euglena gracilis* were chemically synthesized. These synthetic probes have been used to detect tRNA genes by Southern hybridizations to restriction fragments of chloroplast DNAs. This new method of tRNA gene mapping and the oligonucleotides synthesized may be of general application to many chloroplast genomes. This is illustrated by the detection of known and unknown tRNA genes of *Euglena gracilis* and spinach, and unknown tRNA genes of maize and cucumber chloroplast DNAs. The precise locus and polarity of the *Euglena gracilis* chloroplast tRNA^{Phe} gene has been determined. We also describe experiments which relate to the effects of the time of hybridization, the stringency of washing, and of base pair mismatches on the hybridization signal.

INTRODUCTION

Synthetic deoxyoligonucleotides have been useful for the detection of specific mRNAs (1,2,3,4,5), cDNAs (6), and for screening large numbers of bacterial colonies for specific DNA sequences (5). The use of synthetic oligonucleotides to study tRNA genes has not previously been reported. We decided to synthesize several deoxyoligonucleotides to be used for the detection of tRNA genes. *Euglena gracilis* chloroplast (ct) DNA was chosen for this study. Several transfer RNA coding loci of this ct genome have been described (7,8,9), and the nucleotide sequences for 17 different tRNA genes and one tRNA are known (10, 11, 12, 13, 14, 15). From a comparison of the primary structures of known chloroplast tRNAs and tRNA genes, we detected several different highly conserved nucleotide sequences in the DHU stem and loop and the T ψ C stem and loop (16,17). It was reasoned that oligonucleotides complementary to these regions could be used in the detection of many ct tRNA genes, not only from *Euglena* but also from both higher and lower plants, and possibly from non-chloroplast sources such as *E. coli*. The greatest homologies

between all tRNAs occur in the DHU stem and loop. This region is highly conserved between isoaccepting and non-isoaccepting tRNAs both within and between species. In contrast, except for invariant bases, the T ψ C stem and loop is generally conserved only among isoaccepting tRNAs, either within or between species. Therefore an oligonucleotide which is complementary to the DHU stem and loop of a tRNA may be useful in detecting tRNA genes in many organisms, although the genes detected with a given probe may not be isoaccepting species. Probes complementary to the T ψ C stem and loop will be most useful in detecting isoacceptors, although the possibility of detecting non-isoacceptors cannot be ruled out. The usefulness of these oligonucleotides can be further enhanced by designing them to be complementary to the template strand of the gene, allowing their use to probe membrane filter blots of either DNA or RNA.

Although the sequence of the *Euglena gracilis* var *bacillaris* tRNA^{Phe} gene has been known since the report of Chang et al. (10), the gene locus has not previously been reported. Based on the results of tRNA^{Phe}-DNA hybridization experiments, Schwartzbach et al. (18) concluded that there is a single tRNA^{Phe} gene in *Euglena gracilis* var *bacillaris* chloroplast DNA, and none in nuclear DNA. We decided to use synthetic oligonucleotides complementary to the known tRNA^{Phe} sequence to precisely map the tRNA^{Phe} gene(s) of *Euglena gracilis* Z chloroplast DNA. Based on a *Dde* I restriction endonuclease cleavage site internal to the tRNA^{Phe}, to one of the synthetic oligonucleotides used as a hybridization probe, and to the tRNA^{Phe} coding locus of *Euglena gracilis* Z chloroplast DNA, we have been able to confirm that a single tRNA^{Phe} gene is present, and to define the precise location and polarity of the gene.

El-Gewely et al. (19) have independently mapped the tRNA^{Phe} gene locus to a 660 bp *Pvu* II-*Bal* I fragment of *Euglena gracilis* var *bacillaris* chloroplast DNA. This locus coincides with the tRNA^{Phe} locus we report for *Euglena gracilis* strain Z chloroplast DNA. El-Gewely et al. (19) did not determine the precise gene locus, the number of genes, or the gene polarity.

In this paper we describe the synthesis of tetradecanucleotides, the characterization of some of the hybridization properties of synthetic tRNA probes to membrane bound DNA, the results of Southern hybridizations to both known and unknown tRNA genes on total ct DNA and cloned restriction fragments from *Euglena* and higher plant ct DNAs, and experiments which define the precise locus and polarity of the *Euglena* ct tRNA^{Phe} gene.

MATERIALS

DNA Synthesis

Deoxynucleosides, benzoyl chloride, and isobutyryl chloride were purchased from Sigma Chemical Co. Trityl chloride, dimethylaminopyridine, 1*H*-tetrazole, and ZnBr₂ were purchased from Aldrich Chemical Co., Inc. 3'-Tri-tylthymidine covalently bound to a silica support by the 5'-hydroxyl group and (CH₃)₂NPClOCH₃ were gifts from Dr. Robert A. Goldman. All other chemicals were of reagent grade.

Chloroplast DNA

Euglena gracilis Klebs, strain Z Pringsheim ct DNA was prepared as previously described (20,21). Spinach ct DNA was provided by Charles Passavant. Maize and cucumber ct DNAs were a gift from Dr. Emil M. Orozco, Jr. A recombinant plasmid containing the spinach ct DNA *Pst* 5 fragment was generously provided by Drs. Jeff Palmer and Wm. Thompson (22).

Plasmid DNA

Recombinant plasmid DNAs were prepared by either a rapid alkaline method (23), or by CsCl-ethidium bromide gradient centrifugation (24,25). The plasmid pEZC2 was isolated from a library of recombinant plasmids generated by ligation of *Hind*III isolated *Euglena gracilis* Z chloroplast DNA and *Hind*III digested pBR322 plasmid DNA.

Enzymes and Radiochemicals

All restriction endonucleases were purchased from New England Biolabs. T4 polynucleotide kinase was purchased from New England Nuclear Co. or New England Biolabs. [γ -³²P]ATP (1000-3000 Ci/mmol) was purchased from New England Nuclear Co.

METHODS

DNA Synthesis

Amino and 5'-hydroxyl groups of deoxynucleotides were blocked by the method of Khorana et al. (26). Tetradecanucleotides were synthesized by a modification of the method of Matteucci and Caruthers (27). Instead of using a machine, all operations were done manually in a 15 ml scintered glass funnel. Condensations were carried out under a dry N₂ atmosphere.

Purification of Tetradecanucleotides

Following removal of protecting groups and cleavage of the DNA-silica linkage, the tetradecanucleotides were ethanol precipitated and washed once with 80% ethanol, then dried *in vacuo*. 50 μ g of each crude DNA mixture was loaded onto a well (0.2 cm X 0.65 cm) of a 20% acrylamide/7 M urea gel (12.5 cm X 14.5 cm X 0.2 cm) with 0.025% xylene cyanol (XC) and 0.025% bromophenol

blue (BPB) dyes. Gels were run at 400 volts for 2-3 h. Under these conditions tetradecanucleotides run nearly between the two dyes. DNA was visualized by UV shadowing. The highest MW bands were excised and the DNA extracted by the crush and soak method of Maxam and Gilbert (28). Electrophoretically purified tetradecanucleotides were ethanol precipitated and washed with 80% ethanol.

5'-³²P Radiolabeling of Tetradecanucleotides

The polynucleotide kinase procedure used was basically that of Maxam and Gilbert (28). 50 pmoles of DNA was incubated with 50 pmoles of [γ -³²P]ATP and 4 units (NEN) of T4 polynucleotide kinase at 37°C for 15 min. Then four additional units of enzyme were added and the reaction was incubated at 37°C for an additional 15 min. The reaction mixture was ethanol precipitated and washed with 80% ethanol, then dried *in vacuo*, and brought up in 0.4 ml H₂O. Typically, the labeled DNA could be used for hybridization experiments without further purification. Alternatively, unincorporated [γ -³²P]ATP as well as ³²P_i could be removed using a 20 ml Sephadex G-25-150 column.

Endonuclease Digestion of Plasmid and ct DNAs

All digestions were carried out according to the suppliers' recommendations. Following digestion, samples were ethanol precipitated and washed twice with 80% ethanol, dried *in vacuo*, and redissolved in 10 μ l of 5% glycerol, 0.025% each of XC and BPB dyes, 18 mM Tris-OH, 18 mM boric acid, 0.5 mM Na₂EDTA.

Transfer of DNA to Nitrocellulose

The procedure of Southern (29) as previously described (30) was used.

Hybridization of ³²P-tetradecanucleotides

The procedure of Wallace et al. (5) was followed. Filters were washed with 6 X SSC at 0°C twice for 1 h and once for 24 h, except as noted. Filters were exposed to Kodak XRP film at -70°C with an intensifying screen for 8 h to 7 days.

RESULTS

Design of Synthetic Oligonucleotides

We have synthesized five deoxytetradecanucleotides. They are listed in Table I. Probes 1-4 are complementary to the DHU stem and loop of a number of tRNAs. Probe 5 is complementary to the T ψ C stem and loop of tRNA^{Phe} of *Euglena*, bean, spinach and maize ct DNAs. Probes 1-3 were designed to hybridize to overlapping *Alu* I and *Dde* I restriction endonuclease sites. They all contain the sequence 5'-CTGAGCT-3'. When a restriction fragment is digested with either of these enzymes, any hybridization signal to Probes 1-3 should

Table I. Chemically Synthesized Tetradecanucleotide tRNA Gene Probes

<u>Probes Synthesized</u>		<u>Complementary tRNAs^a</u>	<u>Positive Hybridization to ct DNA Restriction Fragments</u>
5'	3'		
1.	d(CTACCAACTGAGCT)	tRNA ^{Ala} (<i>Euglena</i> , Maize, Tobacco), tRNA ^{Val, Gly} (<i>Euglena</i>), tRNA ^{Phe} (<i>Euglena</i> , Maize, Spinach)	<i>Eco</i> A, <i>Eco</i> G, <i>Eco</i> J', <i>Eco</i> P, <i>Eco</i> V, <i>Hind</i> 3, <i>Hind</i> 6, <i>Hind</i> 21, <i>Hind</i> 27, 28, 29
2.	d(TAACCAACTGAGCT)	tRNA ^{Ile} (<i>Euglena</i>)	<i>Eco</i> G, <i>Eco</i> J', <i>Eco</i> P, <i>Eco</i> V
3.	d(TCTACCACTGAGCT)	tRNA ^{Asn} (<i>Euglena</i>), tRNA ^{Ile} (Tobacco, Maize)	<i>Eco</i> G, <i>Eco</i> Q
4.	d(GTGCTACCAATTT)	tRNA ^{Leu} 1,2,3 (Bean) tRNA ^{Leu} 3 (Spinach)	None
5.	d(AGATTGAACTGGT)	tRNA ^{Phe} (<i>Euglena</i> , Bean, Spinach)	<i>Eco</i> A, <i>Hind</i> 21

^aBased on Sprinzl and Gauss (16, 17).

be reduced or eliminated. Therefore by comparing both restriction endonuclease mapping data and corresponding Southern hybridization data, the exact location of a tRNA gene can be deduced. In the case where two different probes are complementary to different regions of the same tRNA gene, the gene polarity can also be determined.

The previously described sequence of *Euglena gracilis* chloroplast tRNA^{Phe} (10) is shown in Fig. 1 in the cloverleaf form. Probe 1 is complementary to positions 9-23 of tRNA^{Phe} (Fig. 1). Since this oligonucleotide

Fig. 1. The structure of the *Euglena gracilis* chloroplast tRNA^{Phe} (10), and the region of complementary oligonucleotide probes. Probe 1 is complementary to the 14 bp sequence in the DHU stem and loop (left). Probe 5 is complementary to the 14 bp sequence in the TΨC stem and loop (right).

contains sequences complementary to the single *Dde* I and *Alu* I restriction endonuclease cleavage sites in the tRNA^{Phe} gene, chloroplast or plasmid DNAs digested with either *Dde* I or *Alu* I would not be expected to hybridize this probe. This is a negative control for tRNA^{Phe} mapping. Also, this oligonucleotide is complementary to the known tRNA^{Ala}, tRNA^{Val}, and tRNA^{Gly} genes of *Euglena gracilis* Z chloroplast DNA, as well as at least two other unidentified genes. In addition to the tRNA^{Phe} gene, these genes are all expected to hybridize Probe 1 in membrane blots of total chloroplast DNA restriction fragments. This is a positive control for tRNA^{Phe} gene mapping. Probe 5 is complementary to positions 49-62 of tRNA^{Phe} (Fig. 1). Of the eighteen known *Euglena* tRNA structures, this sequence is unique to tRNA^{Phe}. As will be shown below, there is only a single locus on the *Euglena* chloroplast DNA that has coincident hybridization of Probes 1 and 5. Probes 1 and 5 are also targeted at the 5' and 3' halves of the tRNA^{Phe} gene, respectively. This was exploited in determining the tRNA^{Phe} gene polarity.

Characterization of Synthetic Oligonucleotides

Following preparative polyacrylamide gel electrophoresis of a crude mixture of synthetic DNA, oligonucleotides can be visualized by UV shadowing. The highest molecular weight DNA band is excised. A typical synthesis yields 20-50 nmoles of purified tetradecanucleotides from 250 nmoles of silica-linked nucleoside. Following 5'-³²P labeling of the gel purified tetradecanucleotides, unincorporated [γ -³²P]ATP can be separated from ³²P-tetradecanucleotides using Sephadex chromatography. Typically, $>4.0 \times 10^7$ cpm in <50 pmoles is obtained after gel filtration chromatography. This is a minimum specific activity, since the loss of DNA during kination and column chromatography cannot be accurately determined.

Determination of Hybridization Conditions

The effect of hybridization reaction time and washing conditions on signal strength and background was investigated. Eight identical *Eco* RI digests of plasmid pPG689 DNA, which contains the *Euglena* ct DNA *Eco* RI fragment *Eco* J' (9) cloned in the plasmid vector pMB9, were electrophoresed on a 1% agarose gel. Following transfer to nitrocellulose, lanes were cut apart, and the DNA blots were hybridized with Probe 1 under various conditions. The results are shown in Fig. 2. Although hybridization is evident after a 30 min reaction, the signal increases continuously from 30 min to 6 h, and nearly doubles in strength from 6 h to 20 h. The effect of different washing procedures on signal strength and background is illustrated in lanes 7, 8, and 9 of Fig. 2. The strongest signal is achieved with the least washes,

Fig. 2. Effects of hybridization time and washing conditions on hybridization signal. Eight identical samples of electrophoretically separated *Eco* RI digested plasmid pPG689 DNA (lane 1) were transferred to nitrocellulose filters and hybridized as follows: lane 2, 0.5 h; lane 3, 1 h; lane 4, 2 h; lane 5, 4 h; lane 6, 6 h; lanes 7, 8 and 9, 20 h. Following hybridization, filters were washed as follows: lanes 2, 3, 4, 5, 6 and 9, twice for 1 h each, plus once for 20 h; lane 7, twice for 1 h each; lane 8, three times for 1 h each. All washes in 6 X SSC at 0°C.

although the background is also highest. No detectable background is present when an overnight wash is included although the signal strength is reduced. Unless otherwise noted, the conditions chosen for hybridization were 20 h at 37°C followed by 3 washes at 0°C with 6 X SSC. Oligonucleotide concentration

Fig. 3. Effect of base mismatching on hybridization signal. Lane 1 and 3: *Hae* III restriction fragments of the *Euglena* ct DNA *Eco* RI fragment *Eco* P. The DNA was transferred to nitrocellulose and hybridized with either Probe 1 (lane 2) or Probe 2 (lane 4). The 520 bp fragment contains the tRNA^{Ala} gene complementary to Probe 1. The 230 bp fragment contains the portion of the tRNA^{Ile} gene complementary to Probe 2. The signal from the 230 bp fragment in lane 2 is barely visible in the original autoradiograph. The mismatch hybridization in lane 4 is example 2 in Table II.

Table II. Effects of Base Pair Mismatches on the tRNA Gene: Oligonucleotide Probe Hybridization Signal Strength

	Probe ^a	tRNA Genes ^b	Mismatch Hybridization ^c	Signal ^d
1.	1	tRNA ^{Ile}	5'-AGCTCAGTTGGT <i>TA</i> TCGAGTCAACCA <i>TC</i> -5'	+/-
2.	2	tRNA ^{Ala}	5'-AGCTCAGTTGGT <i>AG</i> TCGAGTCAACCA <i>AT</i> -5'	+
3.	1	tRNA ^{Gly, Thr}	5'-AGCTCAGT·GGTAG TCGAGTCAACCA <i>TC</i> -5'	+
4.	2	tRNA ^{Gly, Thr}	5'-AGCTCAGT·GGT·A TCGAGTCAACCA <i>AT</i> -5'	+/-
5.	3	tRNA ^{Ala}	5'-AGCTCAGT <i>T</i> GGTAGA TCGAGTCA·CCATCT-5'	+/-
6.	1	tRNA ^{Phe}	5'-AGCTCAGTGGTAG TCGAGTCAACCA <i>TC</i> -5'	+
7.	3	tRNA ^{Thr}	5'-A <i>A</i> CTCAGTGGATAGA TCGAGTCACC·ATCT	+/-
8.	1	tRNA ^{Met}	5'-AGCTCAG·GGTAG TCGAGTCAACCA <i>TC</i> -5'	-
9.	1	tRNA ^{Trp}	5'-AGTTC <i>A</i> ATTGGTAG TCGAGTCAACCA <i>TC</i> -5'	-
10.	5	tRNA ^{MetE}	5'-ACAGGTTCAAATCT TGGTCAAGTTT <i>AGA</i> -5'	-
11.	5	tRNA ^{Ile}	5'-GCTAGTTCAAGTCT TGGTCAAGTTT <i>AGA</i> -5'	-
12.	5	tRNA ^{Thr}	5'-GTCGGTTCGAATCC TGGTCAAGT <i>T</i> AGA-5'	-
13.	5	tRNA ^{Tyr}	5'-GCTGGTTCGAATCC TGGTCAAGT <i>T</i> AGA-5'	-
14.	5	tRNA ^{Val}	5'-AGCGGTTCGAATCC TGGTCAAGT <i>T</i> AGA-5'	-
15.	5	tRNA ^{Leu}	5'-GAGAGTTCGAGTCT TGGTCAAGT <i>T</i> AGA-5'	-
16.	5	tRNA ^{Ala}	5'-AGCGGTTCGAGTCC TGGTCAAGT <i>T</i> AGA-5'	-
17.	3	tRNA ^{MetE}	5'-AGCTCAGAGGTAGA TCGAGTCACCATCT-5'	-
18.	3	tRNA ^{Trp}	5'-AGTTC <i>A</i> ATTGGTAGA TCGAGTCA·CCATCT-5'	-
19.	3	tRNA ^{His}	5'-AGCCAAGTGGTA <i>AG</i> TCGAGTCACCATCT-5'	-
20.	2	tRNA ^{MetE}	5'-AGCTCAG·GGTAG TCGAGTCAACCA <i>AT</i> -5'	-
21.	2	tRNA ^{Trp}	5'-AGTTC <i>A</i> ATTGGTAG TCGAGTCAACCA <i>AT</i> -5'	-

Table II. Effects of Base Pair Mismatches on the tRNA Gene: Oligonucleotide Probe Hybridization Signal Strength (continued)

22.	2	tRNA ^{His}	5'-AGCCAAGT·GGTAA TCGAGTCAACCAAT-5'	-
23.	4	tRNA ^{Leu}	5'-AAATTGGTAGACGC TTTAACCATCTGTG-5'	-
24.	4	tRNA ^{Trp}	5'-CAATTGGTAGAACG TTTAACCATCTGTG-5'	-
25.	2	tRNA ^{Phe}	5'-AGCTCAGCTGGAGA TCGAGTCAACCAAT-5'.	-
^a Tetradecanucleotide probes are identified as in Table I.				
^b All examples are for <i>Euglena gracilis</i> ct tRNA genes except 6, 7 and 25, which are spinach ct DNA tRNA genes.				
^c The tRNA sequence is on top, the oligonucleotide probe is on the bottom. Mismatches are denoted by italics and/or gaps, denoted by symbols (·). In examples 6 and 7 additional unknown tRNA genes may be present and contributing to the hybridization signals.				
^d (+) = detectable signal, but weaker than signals from totally matching structures. (+/-) = barely detectable signal. (-) = signal absent. For an illustration of a weak signal, see Fig. 3.				

was 1 ng/ml.

Effect of Base Mismatches on Hybridization

Probes 1 and 2, which differ by only two bases out of 14, are complementary to the tRNA^{Ala} and tRNA^{Ile} genes, respectively, in the 16 S-23 S rRNA spacer regions of the *Euglena* ct DNA ribosomal RNA operons. These probes and genes were chosen for an initial analysis of the effects of base mismatches on hybridization. The *Euglena* ct DNA *Eco* RI restriction fragment *Eco* P containing both the tRNA^{Ala} and tRNA^{Ile} genes (31) was digested with *Hae* III. *Hae* III cleaves between the DHU regions of the two genes where the probes hybridize. Products were electrophoresed on a 1% agarose gel, transferred to nitrocellulose, and hybridized with either Probe 1 or Probe 2. The results are shown in Fig. 3. Both probes gave strong signals to the fragment containing the totally matching gene sequence. Probe 1 hybridized only weakly to the tRNA^{Ile} gene, despite 12 of 14 consecutive, homologous bases. The remaining two bases give T-T and C-A mismatches. Probe 2 hybridized detectably to the tRNA^{Ala} gene, but this signal is still weaker than that from the totally matching Probe 2. This hybrid also contains 12 of 14 consecutive homologous bases, ending in A-A and G-T mismatches.

Many tRNAs from *Euglena* and higher plant ct DNAs have been completely

Fig. 4. Hybridizations of Probe 1 to *Euglena* ct DNA and cloned *Euglena* ct DNA restriction fragments. Lanes 1-4 are *Eco* RI digests of the following DNAs: lane 1, *Euglena* ct DNA; lane 2, pPG14 plasmid DNA (13), containing an *Eco* G/pMB9 recombinant; lane 3, pPG689 plasmid DNA (33), containing an *Eco* J'/pMB9 recombinant; lane 4, pPG50 plasmid DNA (31), containing an *Eco* P/pMB9 recombinant. Lane 5 is a *Bam* HI, *Sal* I, *Eco* RI digest of pPG76 plasmid DNA (14), containing a *Bam*-*Sal* 9/pBR322 recombinant. Each lane has a 1% agarose gel stained with ethidium bromide and the corresponding autoradiogram of the nitrocellulose filters.

sequenced. For *Euglena* chloroplast DNA, 17 tRNA gene sequences and one tRNA sequence are known. We have used these genes as both positive and negative controls for evaluating the specificity of hybridization with synthetic deoxyoligonucleotides. The ability of each synthetic oligonucleotide to hybridize to non-homologous, but closely related DNA sequences in each of these 17 known genes was evaluated. In addition, hybridization to closely related sequences of spinach chloroplast tRNA genes was monitored. A comprehensive summary of the effect of base pair mismatches on hybridization signals is given in Table II. There are several noteworthy results. First, from examples 3 and 5 we see that single base bulges in either the probe or the target sequence reduce, but do not eliminate, the hybridization signal. Second, an example of the destabilizing effect of multiple G-T mismatching is shown in example 11. All three mismatches in this structure are G-T mismatches and no signal is evident. Other examples of the destabilizing effect of G-T pairs are 12, 13 and 14. The presence of a terminal C-A mismatch in

Fig. 5. Hybridizations of Probe 2 to *Euglena* ct DNA and cloned *Euglena* ct DNA restriction fragments. The DNA samples in all lanes are identical to those shown in Fig. 4. Note that signals in lanes 2, 4 and 5 are due to base mismatch structures (see text and Table II).

these hybrids cannot by itself be responsible for the loss of signal, since a terminal C-A mismatch occurs in example 1 which has a weak signal. Third, the effect of a mismatch near the middle of a duplex is shown in example 17. In this position, even a single base pair mismatch eliminates the hybridization signal. These data are in agreement with the conclusions of Wallace et al. (5), that a single base pair mismatch three bases from one end destabilizes duplexes enough to eliminate hybridization signals. It should be noted that examples 6 and 7 may not be the best structures possible. The spinach 7.6 md *Pst* I fragment which was used in these experiments contains as many as ten tRNA genes (32), several of which have not been sequenced. The structures presented are the most stable based on known tRNA gene sequences in this DNA.

Hybridization to Known and Unknown *Euglena* and Higher Plant Chloroplast tRNA Genes

Euglena chloroplast DNA has at least 8-9 loci where tRNA genes are clustered (9). Four of these loci, containing 17 different genes have been sequenced (10,11,12,13,14,15). The results of hybridizations of synthetic probes to both total *Euglena* ct DNA and to cloned fragments of ct DNA, which

Fig. 6. Hybridizations of Probe 3 to cloned *Euglena* ct DNA restriction fragments. All lanes are *Eco* RI digested DNAs. The DNAs are as follows: lane 1, pPG689 plasmid DNA (33); lane 2, pPG50 plasmid DNA (31); lane 3, pEZC300 plasmid DNA (15), containing an *Eco* Q/pBR325 recombinant. The signal in lane 2 is due to a base mismatch structure (see Table II).

have been transferred to nitrocellulose are shown in Figs. 4, 5, and 6. In every case, previously sequenced tRNA genes give the expected hybridization signals. Genes not complementary to probes do not. Besides known tRNA genes giving signals, two fragments which contain known tRNA loci (9), but in which the identity of the genes has not been determined, also give signals with one or more deoxyoligonucleotides. For example, both *Eco* J' and *Eco* Q were identified as containing at least one tRNA locus (9). They gave signals when hybridized to Probe 1 and Probe 3, respectively. These probes were used to further fine map the tRNA gene loci within these fragments. *Eco* Q has now been sequenced (15). Of the five tRNA genes found, two (tRNA^{Thr}, tRNA^{Gly}) were complementary to Probe 3. There are two or three known tRNA loci of *Euglena* ct DNA not giving signals with any of the probes used in this study. These are in *Eco* I, one of two loci in *Eco* A, and possibly one in *Eco* B.

A detailed description of spinach chloroplast tRNA coding loci has been reported by Weil et al. (32). Several of our synthetic oligonucleotides are complementary to spinach chloroplast tRNAs of known sequence (Table I). The results of hybridizing Probe 5 to total spinach ct DNA digested with *Pst*

Fig. 7. Hybridizations of Probe 5 to spinach ct DNA and a cloned spinach *Pst* I fragment. Lane 1, spinach ct DNA digested with *Pst* I; lane 2, spinach ct DNA restriction fragment *Pst* 5 cloned in pBR322 (22), digested with *Kpn* I and *Sal* I. Each lane has a 1% agarose gel stained with ethidium bromide (left) and the corresponding autoradiogram from the hybridization (right).

I, and to a *Kpn* I/*Sal* I digest of the 7.6 md *Pst* I fragment designated *Pst* 5 cloned in pBR322 are shown in Fig. 7. There is a signal from the 7.6 md *Pst* I fragment, which encodes tRNA^{Phe}. Probe 5 is complementary to spinach tRNA^{Phe}. In addition, the probe is hybridizing to the 8.4 md *Pst* I fragment. The latter contains several tRNA genes (32). When Probe 5 is hybridized to the cloned 7.6 md *Pst* I fragment digested with *Kpn* I and *Sal* I, two signals are evident. One is due to the tRNA^{Phe} gene in the 2.0 *Kpn* I/*Sal* I fragment. The other is due to the 4.0 kbp *Kpn* I fragment which is known to contain both tRNA^{Met2} and tRNA^{Val} genes. However, the known tRNA^{Met2} and tRNA^{Val} gene sequences both show mismatching when the probe sequence is aligned. It is likely that a third tRNA gene complementary to Probe 5 is present in the 4.0 md *Kpn* I fragment. A restriction endonuclease map of the 7.6 md *Pst* I fragment and the results of hybridizations with Probe 5 are summarized in Fig. 8. In preliminary work with cucumber ct DNA and maize ct DNA positive hybridizations with Probe 1 have been obtained (data not shown).

Fig. 8. Restriction endonuclease map of spinach ct DNA fragment *Pst* 5, (a). Summary of hybridization results with Probe 5, (b). The *tRNA^{Phe}* gene is located on the 2.0 kbp *Kpn* I/*Sal* I fragment. A second region complementary to Probe 5 is located on the 3.9-4.0 kbp *Kpn* I fragment. Map adapted from ref. 32.

Fig. 9. Locating the *Euglena* chloroplast *tRNA^{Phe}* gene on a *Hind*III restriction fragment with membrane filter hybridization. Lane 1, (left) *Hind*III digestion product of *Euglena* ct DNA electrophoretically separated on a 1% agarose gel stained with ethidium bromide, and photographed under UV illumination; and the corresponding autoradiogram (right) of the membrane filter hybridized with probe 1. Lane 2, same as lane 1, except membrane filter hybridized with Probe 5. *Hind*III fragments containing known tRNAs complementary to Probe 1 and/or Probe 5 are listed on the right.

Fig. 10. Location and polarity of *Euglena* ct tRNA^{Phe} gene on the *Euglena* ct genome. (a) Restriction endonuclease map of *Eco* A, with the location of *Hind*21 (sizes in kbp). The map is complete for all enzyme sites in *Eco* A except *Hind*III and *Hae* III. (b) Restriction endonuclease map of *Hind*21 (sizes in bp). An unmapped *Taq* I site is present between the two *Taq* I sites denoted "300 + 80" for the sizes of the unmapped *Taq* I fragments. (c) Restriction endonuclease map of the internal 230 bp *Taq* I fragment of *Hind*21. (d) Hybridization boundaries with Probe 1. (e) Hybridization boundaries with Probe 5. (f) Exact location and polarity of *Euglena* ct tRNA^{Phe} gene relative to the single *Dde* I site in *Hind*21.

Locating the tRNA^{Phe} Gene on *Euglena* ct DNA

The results of hybridizations of Probes 1 and 5 to membrane filter blots of total *Euglena* ct DNA digested with *Hind*III are shown in Fig. 9. Probe 1 hybridizes to the following four DNAs: (1) The 6.0 kbp *Hind*3, which contains both tRNA^{Gly} and tRNA^{Val} genes; (2) the 4.2 kbp *Hind*6 which contains at least two tRNA genes that have not been identified; (3) the 2.2 kbp triplet *Hind*27, *Hind*28, *Hind*29 which contains the repeated tRNA^{Ala} genes within the spacer regions of the 16 S-23 S rRNA genes; and (4) the 2.5 kbp *Hind*21. Probe 5 hybridized only to *Hind*21. *Hind*21 must be the fragment containing the tRNA^{Phe} gene. In a separate experiment, Probe 5 hybridized to the 22 kbp *Eco* RI fragment *Eco* A of *Euglena* ct DNA (data not shown).

*Hind*III fragments from *Euglena* ct DNA were shotgun cloned into pBR322. Several clones containing *Hind*21 sized inserts were isolated and screened for hybridization with Probe 5. A plasmid DNA designated pEZC 2 gave a positive hybridization signal (data not shown). This cloned recombinant DNA was used in all subsequent experiments. A restriction endonuclease map of the chloroplast DNA insert in pEZC2 was determined (Fig. 10b). The chloroplast DNA

Fig. 11. Location and polarity of *Euglena* ct tRNA^{Phe} gene using membrane filter hybridization. Each pair of lanes has a photograph of electrophoretically separated restriction endonuclease fragments of either plasmid pEZC2 DNA, which contains the *Euglena* ct DNA *Hind*III fragment *Hind*21 cloned into pBR322, or purified *Hind*21. Also shown is the corresponding autoradiogram of the membrane filter following hybridization. Lane 1, *Hind*III/*Taq* I digest of pEZC2; lane 2, *Hind*III/*Bst*NI digest of pEZC2; lane 3, *Hae* III digest of *Hind*21; lanes 4 and 5, *Dde* I digest of *Hind*21. Lanes 1-4 hybridized with Probe 1, Lane 5 hybridized with Probe 5. Samples for 1, 2, 4 and 5 were electrophoresed in a 1% agarose gel, and for lane 3, a 2% agarose gel. Some pBR322 contamination of the purified *Hind*21 is evident in lanes 3, 4 and 5. Sizes in bp. The fragment locations for *Hind*21 digestion products indicated in the margins can be deduced from Figure 10b.

insert of pEZC2 was found to be a 2.5 kbp *Hind*III fragment which comigrates with authentic *Hind*21 (data not shown). *Hind*21 is cleaved internally at a single loci by *Pvu* II and *Bal* I to yield a DNA fragment of 0.5 kbp. There is only a single region in *Euglena* ct DNA, located within *Eco* RI fragment *Eco* A, where *Pvu* II and *Bal* I sites are in such close proximity (Fig. 10a) (11,12). This region corresponds to a previously described tRNA coding region in *Euglena gracilis* Z (12) and var *bacillaris* (13), and to the tRNA^{Phe} locus identified by El-Gewely et al. (7). The 2.5 kbp *Hind*21 was found to contain a single *Dde* I cleavage site, which is thus the presumptive tRNA^{Phe} gene locus.

The locus and polarity of the tRNA^{Phe} gene were confirmed by hybridization of either Probe 1 or 5 to membrane filter blots of pEZC2 DNA restriction

fragments (Fig. 11). Probe 1 hybridized specifically to a 230 bp *Taq* I fragment (Fig. 11, lane 1), and more precisely to a region defined by *Bst*NI and *Hae* III (*Bal* I) sites (Fig. 11, lanes 2 and 3) separated by approximately 120 bp. Digestion of *Hind*21 with *Dde* I, which cuts between these *Bst*NI and *Hae* III cleavage sites, results in no Probe 1 hybridization (Fig. 11, lane 4). Therefore the *Dde* I cleavage site in *Hind*21 is within the DHU stem and loop coding region near the 5'-end of the tRNA^{Phe} gene. Probe 5 hybridizes only to the 460 bp *Dde* I fragment of *Hind*21, and a *Dde* I partial digestion product (Fig. 11, lane 5; Fig. 10e). Therefore the tRNA^{Phe} gene polarity could be assigned (Fig. 10f). If we assume that the tRNA^{Phe} gene is not interrupted by intervening sequences, the gene can be located with absolute precision with respect to the *Dde* I site in *Hind*21.

DISCUSSION

The experiments presented are an illustration of the usefulness of synthetic deoxyoligonucleotides as hybridization probes for tRNA genes. Chloroplast tRNA genes were chosen as models for this study because a large number of tRNA genes are present in a relatively small DNA. The current approach could be used for tRNA genes from any genome. All of the probes synthesized for this study are complementary to the template strand of the gene and therefore can also be used as RNA transcript probes. Probes 1 and 3 have been successfully used to detect *Euglena* ct tRNAs (data not shown). The usefulness of Probes 1-3 is enhanced by the presence of restriction endonuclease sites within their sequences, allowing tRNA genes to be mapped precisely on a restriction fragment.

Synthetic deoxyoligonucleotides complementary to tRNA genes have also been useful in studying the effects of mismatches on hybridization. Since tRNAs have very closely related primary structures, only a few deoxyoligonucleotides need to be synthesized in order to study many different types of mismatching effects. As more tRNA sequences become known, the effective usefulness of a particular probe in mismatch studies increases. In addition to studies which are designed to detect the presence or absence of signals, direct blotting of DNA onto nitrocellulose could be employed in studying the effects on thermal stability of different types of base pair mismatches.

We report here the first use of synthetic deoxyoligonucleotides to locate a specific tRNA gene precisely on a chloroplast genome, and to the best of our knowledge, any genome. As illustrated in this paper, they can be used for precise mapping of a gene with respect to a restriction endonuclease

site if two conditions are met. The probe must contain a restriction endonuclease recognition sequence, and the particular restriction endonuclease site must be mapped. The polarity of the gene can be determined if two probes are synthesized and hybridization signals occur on different restriction fragments or if two probes are designed with a restriction site recognition sequence in one of the probe sequences. The techniques described are generally applicable to many systems for which tRNA sequences are known. The two probes used to locate the *Euglena* tRNA^{Phe} gene were designed to be complementary to both bean and maize chloroplast tRNA^{Phe} genes. Probe 5 is also complementary to spinach chloroplast tRNA^{Phe} (Table I).

Recently El-Gewely et al. (19) located the *Euglena gracilis* var *bacillaris* tRNA^{Phe} gene on a 660 bp *Pvu* II/*Bal* I fragment using purified, *in vitro* ³²P-labeled tRNA^{Phe} as a probe. We can confirm their findings and further define the locus to the exact site of the gene if it is assumed there are no intervening sequences within the tRNA^{Phe} gene. In addition, we have defined the polarity of the gene. This is the first example of a specific *Euglena* ct tRNA gene that has been mapped for both the Z-strain and the *bacillaris* strain. Even though these strains have some differences in their physical maps (33), this tRNA^{Phe} coding locus seems to be highly conserved.

Besides use as hybridization probes, these deoxyoligonucleotides may be used as primers for the "dideoxy" DNA sequencing method of Sanger (34). In *Euglena* all known tRNA genes occur in clusters (9, 14, 15, 31). Therefore these primers might be used to prime sequencing 5'- to the hybridizing gene. Such a sequence should contain either the 5'-proximal tRNA gene(s), or the promoter for the tRNA gene cluster. The locus of the tRNA^{Phe} gene is within a region previously defined among nine major *Euglena* ct tRNA coding loci (8,9). It seems possible that the tRNA^{Phe} gene might be part of a yet uncharacterized tRNA gene cluster. It is noteworthy that both Probes 1 and 5 can be used to prime sequencing within the tRNA^{Phe} gene and 5'-flanking sequences. These same two probes should be of general use in the future to define the locus of many higher plant ct tRNA^{Phe} genes (Table I).

The advantages of using synthetic oligonucleotides as probes for tRNA genes are many: (1) Probes are of defined sequence and relatively easy to isolate and purify in large amounts. (2) Probes are single stranded and are synthesized without a 5'-phosphate so they are easy to end label with 5'-³²P to very high specific activities (>10⁶ cpm/pmole). Also they do not need to be denatured before hybridizations and do not self anneal. (3) Probes typically give very strong signals with little or no background. (4) Oligo-

nucleotides can be used as primers for initial sequencing using dideoxy methods to screen for tRNA gene clusters. (5) They can be used as probes of RNAs to detect both mature tRNAs and possibly unprocessed primary tRNA transcripts. (6) Relatively few oligonucleotides need to be synthesized to locate many tRNA genes from many sources.

The set of probes prepared for the present study should have broad, general utility for the study of many genomes. We have demonstrated their use with both *Euglena* and spinach chloroplast DNAs. We anticipate that many, and in some cases all, of the tRNA probes will be complementary to tRNA genes on chloroplast DNA from almost any plant or algae source. It seems likely that many of the probes could be applied to the study of procaryote genomes in general.

ACKNOWLEDGEMENTS

This work was supported by Grant GM28463 from the National Institutes of Health. We thank Marvin Caruthers, Robert Goldman and Eric Fisher for assistance in deoxyoligonucleotide synthesis; Jesse Dodd for preparing *Euglena* ct DNA and plasmid DNA; Margaret Hollingsworth and Jerry Karabin for plasmid DNA; Chuck Passavant for spinach ct DNA; Scott E. Bingham and H. Marshall Matthews for the construction and initial characterization of pE2C2; Emil Orozco and Nam-Hai Chua for gifts of maize and cucumber ct DNAs.

REFERENCES

1. Tullis, R. H., Gutierrez, R. and Rubin, H. (1980) Biochem. Biophys. Res. Comm. 93, 941-947.
2. Mevarech, M., Noyes, B. E. and Agarwal, K. L. (1979) J. Biol. Chem. 254, 7472-7475.
3. Noyes, B. E., Mevarech, M., Stein, R. and Agarwal, K. L. (1979) Proc. Natl. Acad. Sci. USA 76, 1770-1774.
4. Agarwal, K. L., Brunstedt, J. and Noyes, B. E. (1981) J. Biol. Chem. 256, 1023-1028.
5. Wallace, R. B., Johnson, M. J., Hirose, T., Miyake, J., Kawashima, E. H. and Itakura, K. (1981) Nucl. Acids Res. 9, 879-894.
6. Sood, A. U., Pereira, D. and Weissman, S. M. (1981) Proc. Natl. Acad. Sci. USA 78, 616-620.
7. Hallick, R. B., Rushlow, K. E., Orozco, E. M., Jr., Stiegler, G. L. and Gray, P. W. (1979) ICN-UCLA Symp. Mol. Cell. Biol. 15, 127-141.
8. El-Gewely, M. R., Lomax, M. I., Lau, E. T., Helling, R. B., Farmerie, W. and Barnett, W. E. (1981) Mol. Gen. Genet. 181, 296-305.
9. Orozco, E. M., Jr. and Hallick, R. B. (1982) J. Biol. Chem. 257, 3258-3264.
10. Chang, S. H., Brum, C. K., Silverklang, M., RajBhandary, U. L., Hecker, L. I. and Barnett, W. E. (1976) Cell 9, 717-723.
11. Orozco, E. M. Jr., Rushlow, K. E., Dodd, J. R., Hallick, R. B. (1980) J. Biol. Chem. 255, 10997-11003.

12. Graf, L., Kössel, H. and Stutz, E. (1980) *Nature* 286, 908-910.
13. Orozco, E. M., Jr. and Hallick, R. B. (1982) *J. Biol. Chem.* 257, 3265-3275.
14. Hollingsworth, M. J. and Hallick, R. B. (1982) *J. Biol. Chem.*, in press.
15. Karabin, J. and Hallick, R. B., submitted for publication.
16. Sprinzl, M. and Gauss, D. H. (1982) *Nucleic Acids Res.* 10, r1-r55.
17. Sprinzl, M. and Gauss, D. H. (1982) *Nucleic Acids Res.* 10, r57-r81.
18. Schwartzbach, S. D., Hecker, L. I. and Barnett, W. E. (1976) *Proc. Natl. Acad. Sci. USA* 73, 1984-1988.
19. El-Gewely, M. R., Helling, R. B., Farmerie, W. and Barnett, W. E. (1982) *Gene* 17, 337-339.
20. Gray, P. W. and Hallick, R. B. (1977) *Biochemistry* 16, 1665-1671.
21. Hallick, R. B., Richards, O. C. and Gray, P. W. in *Methods in Chloroplast Molecular Biology*, Edelman, M., Hallick, R. B., and Chua, N.-H. Eds., Elsevier/North Holland, Amsterdam, in press.
22. Palmer, J. D. and Thompson, W. F. (1981) *Gene* 15, 21-26.
23. Birnboim, M. C. and Doly, J. (1979) *Nucleic Acids Res.* 7, 1513-1523.
24. Guerry, P., LeBlance, D. J. and Falkow, S. (1973) *J. Bacteriol.* 116, 1064-1066.
25. Cohen, S. N. and Miller, C. A. (1970) *J. Mol. Biol.* 50, 671-687.
26. Schaller, H., Weimann, G., Lerch, B. and Khorana, H. G. (1963) *J. Amer. Chem. Soc.* 85, 3821-3827.
27. Matteucci, M. D. and Caruthers, M. H. (1981) *J. Amer. Chem. Soc.* 103, 3185-3191.
28. Maxam, A. M. and Gilbert, W. (1980) *Meth. Enzymol.* 65, 499-560.
29. Southern, E. M. (1975) *J. Mol. Biol.* 98, 503-517.
30. Rushlow, K. E., Orozco, E. M., Jr., Lipper, C. and Hallick, R. B. (1980) *J. Biol. Chem.* 255, 3786-3792.
31. Orozco, E. M., Jr., Gray, P. W. and Hallick, R. B. (1980) *J. Biol. Chem.* 255, 10991-10996.
32. Driesel, A. J., Crouse, E. J., Gordon, U., Bohnert, H. J., Herrmann, R. G., Steinmetz, A., Mubumbila, M., Keller, M., Burkard, G. and Weil, J. H. (1979) *Gene* 6, 285-306.
33. Hallick, R. B. in *The Biology of Euglena*, Buetow, D., Ed., Vol IV, Academic Press, New York, in press.
34. Sanger, F., Nicklen, S. and Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.