JOURNAL OF BACTERIOLOGY, Sept. 1987, p. 4368–4375
0021-9193/87/094368-08$02.00/0

Vol. 169, No. 9

# Interrupted Thymidylate Synthase Gene of Bacteriophages T2 and T6 and Other Potential Self-Splicing Introns in the T-Even Bacteriophages

FREDERICK K. CHU,* FRANK MALEY, JOANNE MARTINEZ, AND GLADYS F. MALEY

*Wadsworth Center for Laboratories and Research, New York State Department of Health, Albany, New York 12201*

Received 2 March 1987/Accepted 25 June 1987

**Southern hybridization analyses of procaryotic DNA from *Escherichia coli*, λ bacteriophage, and T1 to T7 phages were carried out. The hybridization probes used consisted of DNA restriction fragments derived from the T4 phage intron-containing thymidylate synthase gene (*td*) and short synthetic oligodeoxynucleotides defining specific exon and intron regions of the gene. It was shown that intact as well as restricted DNA from the T-even phages hybridized not only to both T4 phage *td* intron- and exon-specific probes but also to probes defining the *td* 5′ (exon I-intron) and 3′ (intron-exon II) presplice junctions. These data strongly suggest that, analogous to the T4 phage, only the T2 and T6 phages among the procaryotes tested contain interrupted *td* genes. The *td* intervening sequence in each phage is roughly 1 kilobase pair (kb) in size and interrupts the *td* gene at a site analogous to that in the T4 phage. This was confirmed by data from Northern (RNA) hybridization analysis of *td*-specific in vitro transcripts of these phage DNAs. [α-$^{32}$P]GTP in vitro labeling of total RNA from T4 phage-infected cells produced five species of labeled RNAs that were 1, 0.9, 0.83, 0.75, and 0.6 kb in size. Only the 1-, 0.9-, and 0.75-kb species were labeled in RNA from T2- or T6-infected cells. The commonly present 1-kb RNA is the excised *td* intron, which exists in both linear and circular forms in the respective T-even-phage-infected cells, while the 0.6-kb RNA unique to T4 may be the excised intron derived from the ribonucleotide reductase small subunit gene (*nrdB*) of the phage. The remaining labeled RNA species are likely candidates for other self-splicing introns.**

Many protein-encoding genes in eucaryotes are interrupted by one or more intervening sequences which are transcribed but not translated. Usually such noncoding sequences (introns) are excised posttranscriptionally, and the resulting coding segments (exons) are ligated via the process of RNA splicing (1). Although this process has been shown to be prevalent in eucaryotes, its possible occurrence in procaryotes was not considered until the discovery of procaryotic introns, first in two tRNA genes in archaebacteria (15) and subsequently in the thymidylate synthase gene (*td*) of the coliphage T4 (9). More recently, a second intron was found which is located in the T4 bacteriophage gene encoding ribonucleotide reductase small B2 subunit (*nrdB*) (26). Specifically in the case of the *td* gene, the formation of mature mRNA has been shown to be preceded by intron excision and exon ligation at the level of RNA in both in vitro (8) and in vivo (5) systems. Subsequent determination of the nucleotide sequence in the *td* intron has revealed that it possesses many structural features typical of eucaryotic class I introns (6, 22, 32). These include a uridine 5′ to the intron, a guanosine in the 3′ end of the intron and nonconserved (E and E′) and conserved (P, Q, R, and S) sequences (10, 11). Furthermore, the *td* primary transcript has been shown to undergo self-splicing in vitro (8, 10, 11) similar to that found for the *Tetrahymena thermophila* nuclear large rRNA (7, 35) and for many yeast and fungal mitochondrial pre-RNAs (13, 27). An internal guide sequence consisting of nine consecutive nucleotides located just downstream of the 5′ splice site has been designated for the *td* intron (11). It is believed that this guide sequence, present in the intron segment of self-splicing pre-RNA, brings the 5′ and 3′ splice sites into proximity via specific base pairing for precise exon ligation (33).

The discovery of a 1,017-base-pair (bp) intron in the T4 phage *td* gene prompted us to screen for the presence of other procaryotic introns, particularly in the T-odd and the other T-even bacteriophages. In this report, we present evidence that T4 *td* intronlike sequences are present in the genomes of the T2 and T6 phages, also interrupting the respective *td* genes. Like the T4 *td* intron, the intron in T2 and T6 was approximately 1 kilobase (kb) in size and was excised at the RNA level as a linear RNA molecule which subsequently cyclized. We also demonstrate that [α-$^{32}$P]GTP, which has been shown to be covalently joined to the 5′ end of the T4 phage *td* intron RNA during self-splicing of the *td* pre-RNAs in vitro (10) and to label RNA extracted from T4 phage-infected cells (14), also labeled several RNA species from T2- and T6-infected cells, suggesting the presence of multiple self-splicing introns in the T-even phages. Interestingly, the heteroduplex analysis of T2, T4, and T6 phage DNAs by Kim and Davidson (16) revealed the presence of a deletion loop at approximately 137 map units (*nrdB* region) and the absence of a loop at 142 map units (*td* region) for both T2-T4 and T4-T6 pairs, implicating (i) an additional sequence for T4 in the *nrdB* region and (ii) highly homologous sequences for all T-even phages in the *td* region, respectively.

## MATERIALS AND METHODS

**Bacterial strains, phage strains, and plasmids.** *Escherichia coli* B was used as the host for infection with bacteriophage λ, λtd, T1, T2, T3, T4, T5, T6, or T7. Phages were purified by CsCl step gradient centrifugation (12). λtd was originally provided by Noreen Murray of Edinburgh University (23). Recombinant plasmids pSP64td and pBTAH were the cor-

* Corresponding author.

TABLE 1. Synthetic oligodeoxynucleotides

| Designation | Definition | Size (mer) | Nucleotide position[a] |
|---|---|---|---|
| 1 | Exon I[b] | 24 | 745 to 768 |
| 2 | Exon II[b] | 20 | 1839 to 1858 |
| 3 | Intron[c] | 22 | 817 to 838 |
| 4 | Intron (5' end)[d] | 21 | 774 to 794 |
| 5 | intron (3' end)[d] | 21 | 1765 to 1785 |
| 6 | exon I-intron[c] | 23 | 757 to 779 |
| 7 | intron-exon II[c] | 25 | 1774 to 1798 |

[a] See reference 9 for numbering of nucleotides in the 2,847-bp *Eco*RI fragment. Exon I extends from nucleotide 200 to 768, the intron extends from 769 to 1785, and exon II extends from 1786 to 2094.
[b] Obtained from Pharmacia, Piscataway, N.J.
[c] Obtained from Systec, Minneapolis, Minn.
[d] Provided by James Freisheim, University of Toledo, Toledo, Ohio.

responding pSP64 and pBR322 plasmids containing the thymidylate synthase gene (*td*) from T4 phage (8) and the corresponding gene (*thyA*) from *E. coli* (3), respectively. DNA was extracted from various sources with phenol followed by exhaustive dialysis.

**Media, enzymes, and oligodeoxynucleotides.** TBYE was used for the growth of *E. coli* strains and contained 1% tryptone (Difco Laboratories), 0.5% NaCl, and 0.5% yeast extract. Ampicillin (100 μg/ml) was included in the growth medium for propagating plasmid-carrying *E. coli* strains. The restriction enzymes *Eco*RI, *Hpa*I, and *Taq*I were obtained from Bethesda Research Laboratories. Other restriction enzymes were from New England BioLabs. SP6 RNA polymerase for in vitro RNA synthesis was purchased from Promega Biotec, and *E. coli* RNA polymerase was from Sigma Chemical Co.

The synthetic oligodeoxynucleotides used as probes in hybridization are listed in Table 1, and their relative positions in the *td* gene are indicated in Fig. 1. The oligodeoxynucleotides were 5' end labeled with [γ-$^{32}$P]ATP (3,000 Ci/mmol; New England Nuclear Corp.) with polynucleotide kinase (Boehringer Mannheim Biochemicals). Purified *td*-derived DNA restriction fragments were labeled with [α-$^{32}$P]dTTP (400 Ci/mmol; Amersham Corp.) by using a Worthington nick translation kit (Cooper Biomedical, Inc.).

**Conditions for hybridization.** *Taq*I-restricted DNA fragments (0.1 to 0.5 μg) were electrophoresed in a 1% agarose slab gel in TBE buffer (100 mM Tris hydrochloride [pH 8.3], 100 mM boric acid, 2 mM disodium EDTA) and then blotted onto a nitrocellulose filter (Schleicher & Schuell, Inc.) by the procedure of Thomas (28). Unrestricted DNA (5 ng) was spotted onto a nitrocellulose filter in 1- to 2-μl volumes. The DNA-containing filters were baked at 80°C under vacuum in an oven for 2 h and stored in sealed plastic bags until use. Prehybridization was carried out in plastic bags at 48°C for at least 4 h in the presence of 0.1 mg of sonicated salmon sperm DNA per ml. Hybridization (at 48°C for 16 h) was initiated by the addition of radioactively labeled DNA fragment or oligodeoxynucleotide probe to a final specific activity of $10^5$ cpm/ml of hybridization mixture. The buffer used in both prehybridization and hybridization (31) contained 90 mM Tris hydrochloride (pH 7.5), 900 mM NaCl, 6 mM disodium EDTA, 0.1% sodium dodecyl sulfate, 10% dextran sulfate, and 0.01% each bovine serum albumin, Ficoll (Pharmacia Fine Chemicals), and polyvinylpyrrolidone. After hybridization, the filters were exhaustively washed in buffer containing 90 mM trisodium citrate and 900 mM NaCl (six to eight changes) for 6 h, either less stringently at 22°C (for DNA

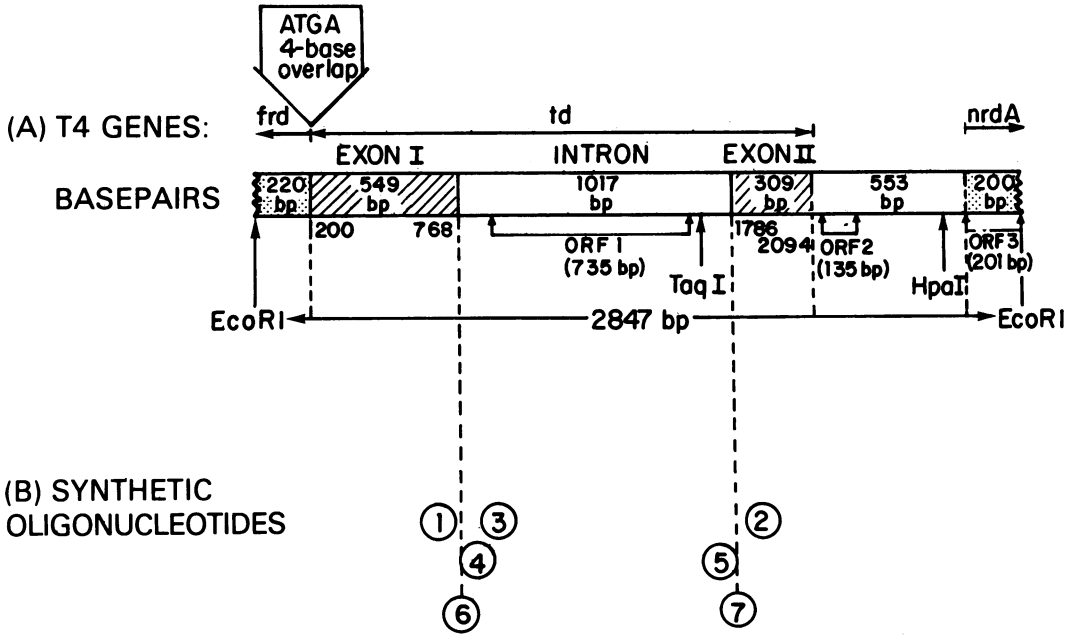


FIG. 1. (A) Genomic organization of the *Eco*RI fragment containing the thymidylate synthase gene (*td*) of T4 bacteriophage and pertinent restriction sites. The fragment is 2,847 bp in size. The numbers indicate the scale in base pairs. The *Eco*RI, *Taq*I, and *Hpa*I restriction enzyme sites are shown. The *Taq*I site is 36 bp downstream of open reading frame 1 (ORF1), while the *Hpa*I site is 354 bp downstream of ORF2. The *td* gene overlaps with the *frd* gene (encoding T4 dihydrofolate reductase) to its left by a 4-bp stretch (ATGA) (9, 24). To its right is the *nrdA* gene (encoding T4 ribonucleotide reductase B1 subunit) located at the end of the *td* fragment (truncated ORF3) (F. K. Chu, G. F. Maley, A.-M. Wang, and F. Maley, Gene, in press). ORFs 1, 2, and 3 are preceded by ribosome-binding sites. ORF1 was previously identified and located in the *td* intron (11). (B) Location of synthetic oligodeoxynucleotide hybridization probes in the *td* fragment (see Table 1 for the sequences of probes 1 to 7).

TABLE 2. DNA dot hybridization

| *td* DNA fragment probe[a] | Hybridization[b] with the following DNA source: | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | λ | λtd[c] | td | T1 | T2 | T3 | T4 | T5 | T6 | T7 | *E. coli* | pBTAH[d] |
| Exon I-intron-exon II[e] | − | + | + | − | + | − | + | − | + | − | − | − |
| Exon I[f] | − | + | + | − | + | − | + | − | + | − | − | − |
| Intron[g] | − | + | + | − | + | − | + | − | + | − | − | − |
| Exon II[h] | − | + | + | − | + | − | + | − | + | − | − | − |

[a] Purified *td* DNA restriction fragments were labeled with $^{32}P$ by nick translation.
[b] −, Nonhybridizing; +, hybridizing.
[c] λ recombinant containing the thymidylate synthase gene (*td*) of T4 phage.
[d] pBR322 plasmid containing the thymidylate synthase gene (*thyA*) of *E. coli*.
[e] 1.6-kb *Hin*dIII-*Eco*RV fragment (nucleotides 379 to 1985).
[f] 0.28-kb *Hin*dIII-*Pvu*II fragment (nucleotides 379 to 655).
[g] 0.9-kb *Acc*I-*Pvu*II fragment (nucleotides 834 to 1737).
[h] 0.87-kb *Pvu*II-*Hpa*I fragment (nucleotides 1737 to 2609).

fragment probes) or more stringently at 52°C (for oligodeoxynucleotide probes), blotted dry, and then subjected to autoradiography on Kodak X-Omat R film.

**Preparation of total RNA from phage-infected cells.** A 1-ml sample of an overnight culture of *E. coli* B was inoculated into 100 ml of TBYE medium for growth at 34°C. At an optical density at 650 nm of 0.3, the culture was infected with a T phage (T2, T4, T5, or T6) at a multiplicity of infection of 7 to 10. At intervals (5, 10, and 20 min), 25 ml of the phage-infected culture was quickly chilled for 10 s in an ethanol-dry ice bath and immediately spun at 10,000 rpm for 10 min in a Sorvall SS-34 rotor. The pelleted cells were suspended in 0.2 ml of TE buffer (10 mM Tris hydrochloride [pH 8], 1 mM disodium EDTA, 5 mM dithiothreitol). Lysozyme was added to a final concentration of 25 μg/ml. The cell suspension was then frozen and thawed twice to effect cell breakage. The resulting cell lysate was treated with RQ1 RNase-free DNase (Promega Biotec) for 1 h at 4°C and then extracted with an equal volume of redistilled phenol saturated with TE buffer for 5 min at room temperature. The upper aqueous phase was phenol extracted a second time for 5 min at room temperature. Total RNA was precipitated with 2 volumes of prechilled absolute alcohol in the presence of 0.2 M NaCl at −20°C for 16 h. Precipitated RNA (0.5 to 1 mg from a 25-ml cell culture) was collected by centrifugation at 4°C, washed twice with 70% alcohol, dissolved in 25 μl of sterile $H_2O$, and stored at −20°C until use.

**Covalent addition of [α-$^{32}$P]GTP to total RNA and gel analysis.** Total RNA (5 μg) from uninfected or T phage-infected cells at 10 min postinfection was incubated with 5 μCi of [α-$^{32}$P]GTP (400 Ci/mmol; Amersham) in 5 μl of buffer containing 40 mM Tris hydrochloride (pH 7.5) and 10 mM $MgCl_2$ at 60°C for 5 min. After incubation, the mixture was chilled on ice and mixed with 15 μl of gel loading buffer containing 0.1 M Tris base, 0.1 M boric acid, 2 mM disodium EDTA, 0.08% each xylene cyanol and bromphenol blue, and 95% deionized formamide. Gel electrophoretic analysis of the labeled products was carried out in a 5% polyacrylamide–8 M urea slab gel (0.75 mm thick) in TBE buffer. The gel was dried onto filter paper backing (Bio-Rad Laboratories) and subjected to autoradiography on Kodak X-Omat R film. RNA for hybridization was first resolved by electrophoresis in a 1.5-mm-thick slab gel, followed by electroblotting to Hybond-N nylon membrane (Amersham). Northern (RNA) hybridization analysis was carried out as described elsewhere (11).

**Other methods.** Transcription with SP6 RNA polymerase in vitro (20) was conducted as described elsewhere (8), and the resulting RNA product was collected by ethanol precipitation after isolation by passage through a Nensorb 20 cartridge (DuPont Co.) (10). Transcription with *E. coli* RNA polymerase was carried out as described by Trimble and Maley (29), but at 30°C instead of 18°C. For both transcription systems [α-$^{32}$P]ATP (400 Ci/mmol; Amersham) was used to uniformly label the transcripts. DNA concentration was determined by measuring the $A_{260}$.

## RESULTS AND DISCUSSION

**Presence of T4 phage *td* exon and intron nucleotide sequences in T2 and T6 phages.** Intron-containing procaryotic genes were first found in archaebacteria, but these genes code for tRNA molecules and the introns are quite small (15). The T4 phage *td* gene is the first instance of an intron-containing procaryotic gene that codes for a protein (9) (Fig. 1A). Recently, the *nrdB* gene coding for the ribonucleotide reductase small subunit (B2) of the T4 phage was found to be interrupted also (26), and its RNA transcript was subsequently shown to undergo self-splicing in vitro (14). To detect potential introns in other procaryotes, we utilized as probes DNA restriction fragments derived from the T4 *td* gene, as well as short synthetic oligodeoxynucleotides defining specific nucleotide sequences in the *td* gene (Fig. 1B and Table 1). These DNAs were radioactively labeled and used as probes to screen for similar sequences in DNA extracted from *E. coli*, λ phage, and coliphages T1 to T7. Table 2 summarizes the hybridization results with T4 *td*-specific DNA restriction fragment probes containing either exon I, exon II, or intron segments. The exon-specific probes from T4 phage hybridized to T2 and T6 DNA, which is not surprising as antibody to T2 thymidylate synthase has been shown to cross-react with the T4 enzyme (4). The T5 phage, the only T-odd phage which encodes its own thymidylate synthase (2), does not contain *td*-specific exon sequences, indicating that the overall T5 synthase primary structure may be significantly different from that of the thymidylate synthase in T-even phages. More interestingly, the T4 *td* intron probe from T4 phage hybridized to both T2 and T6 DNA, suggesting the presence of T4-like intron sequences in the latter phages. Although these data do not specify the location of these sequences in T2 and T6, they demonstrate a lack of significant sequence homology between the T4 phage *td* gene fragments and the genomes of phage λ, T1, T3, T5, or T7 or that of *E. coli*. In the last case, the cloned *E. coli* thymidylate synthase gene (*thyA*) in pBTAH (3) also failed to hybridize to the *td* intron probe (Table 2). These results were confirmed by data from hybridization analyses, under high-stringency conditions, of the

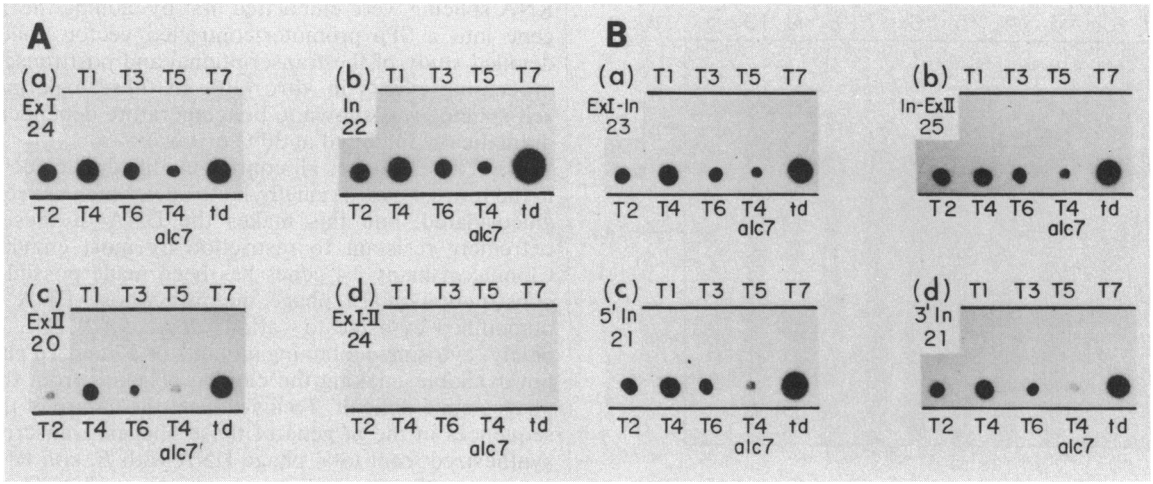


FIG. 2. Hybridization analysis of T phage DNAs with *td* region-specific synthetic oligodeoxynucleotides. The oligodeoxynucleotides were labeled with $^{32}P$ at the 5′ end. T4*alc7* is a T4 mutant whose DNA contains cytosine in place of the normally present hydroxymethylcytosine residues in wild-type T-even phage DNAs. (A) Analysis with *td* exon (Ex)- or intron (In)-specific oligodeoxynucleotide probes. (B) Analysis with pre-splice-junction-specific oligodeoxynucleotide probes. See Fig. 1 for the location of synthetic probes in the *td* gene.

various DNAs with oligodeoxynucleotide probes whose sequences (Table 1) are unique to either T4 *td* exon I (Fig. 2A, panel a), intron (panel b), or exon II (panel c). Since the DNA fragment probes (Table 2) were derived from the T4*alc7* phage, a T4 mutant which contains cytosine in place of 5-hydroxymethylcytosine (19), DNA from the mutant was included as a control.

To determine the location of the T4 *td* intronlike sequences, particularly in relation to the thymidylate synthase exons in T2 and T6 DNA, we probed the respective DNAs with synthetic oligodeoxynucleotides defining the ends of the intron as well as the flanking 5′ and 3′ splice sites of the T4 *td* gene. Conditions for washing the hybridized filters were stringent (52°C instead of 22°C), such that only complete hybrids were stable enough to be detected. For example, an oligodeoxynucleotide (24-mer) defining the contiguous T4 *td* exon I-exon II spliced junction failed to hybridize to T4 DNA containing the interrupted *td* gene or to T2 or T6 DNA (Fig. 2A, panel d). The 5′ (Fig. 2B, panel c) and 3′ (Fig. 2B, panel d) end sequences of the T4 *td* intron were present in both T2 and T6 genomes. Furthermore, the exon-intron boundary sequences defining the T4 *td* exon I-intron (Fig. 2B, panel a) and intron-exon II (Fig. 2B, panel b) splice sites were also detected in both T2 and T6, suggesting that the *td* intronlike sequence in these phages may interrupt their respective *td* genes in the same relative position as that in T4. To confirm that this is indeed the case, we carried out hybridization analyses of *Taq*I-restricted DNA fragments of T2, T4, and T6 with T4 *td* exon I-specific (24-mer) and intron-specific (22-mer) oligodeoxynucleotides as probes. *Taq*I was used for restriction because it cuts DNA containing glucosylated hydroxymethylcytosine residues. Based on the known nucleotide sequences of the T4 phage *frd* (encoding dihydrofolate reductase) and *td* genes, which were previously shown to overlap by a 4-bp stretch (antisense ATGA) (9, 24), a 1.7-kb DNA fragment comprising the 3′ portion of *frd*, *td* exon I, and a 907-bp stretch of *td* intron should be generated from T4 DNA by *Taq*I (Fig. 3B). The results of hybridization analysis of the *Taq*I-restricted T2, T4, T6, and T4*alc7* DNA fragments after electrophoresis on 1% agarose slab gel and blotting are presented in Fig. 3A. As expected for the positive-control DNAs, a 1.7-kb DNA band in both the T4 and T4*alc7* digests hybridized to T4 *td*-specific exon I and intron oligodeoxynucleotide probes (Fig. 3a and b, respectively). The identity of the 3′ end of this fragment containing the truncated *td* intron was confirmed by its failure to hybridize to oligodeoxynucleotide probes defining either the 3′ end of the T4 *td* intron or the 5′ end of exon II (Fig. 3c). Significantly, both the T2 and T6 DNA digests contained a 1.7-kb species exhibiting a T4-specific pattern of hybridization to the oligodeoxynucleotide probes (Fig. 3), implying the presence of a very similar nucleotide sequence among the T-even phages in this region of the genome containing the *frd* and *td* genes.

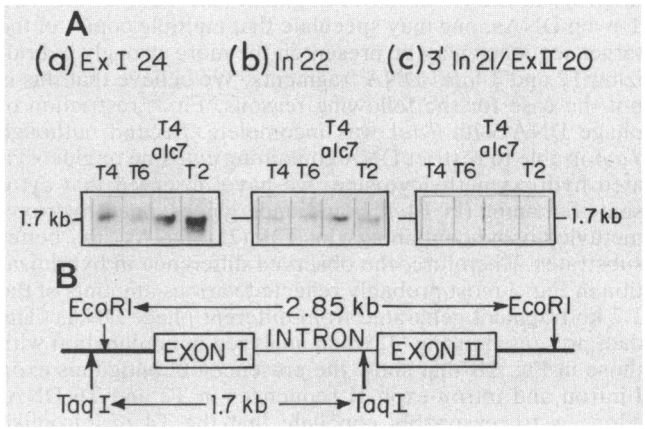


FIG. 3. Blot hybridization analysis of *Taq*I-restricted DNA from T-even phages. (A) DNA fragments generated by digestion with *Taq*I were electophoresed on an agarose gel, blotted onto a nitrocellulose filter, and subjected to hybridization analysis with $^{32}P$-labeled synthetic oligodeoxynucleotides. Ex, Exon; In, intron. See Table 1 for the nucleotide sequence of the probes. (B) *Taq*I restriction sites in the 2.85-kb *Eco*RI *td* fragment, showing the generation of a 1.7-kb fragment with *Taq*I. The second *Taq*I site is based on the T4 *td* gene sequence (11). The first *Taq*I site to the left of the first *Eco*RI site is determined from the nucleotide sequence of the T4 *frd* gene encoding dihydrofolate reductase (24).

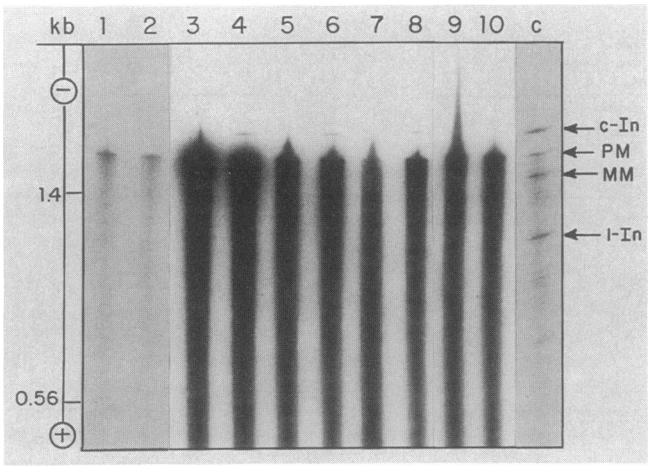


FIG. 4. In vitro synthesis and processing of transcripts of *E. coli* and phage DNA. Transcription of DNA with *E. coli* RNA polymerase was carried out at 30°C in the presence of [$\alpha$-$^{32}$P]ATP in a volume of 10 $\mu$l. After 30 min, RNase-free DNase was added to terminate transcription. After 10 min at 22°C, one half of the treated reaction was chilled on ice and the other half was shifted to 60°C for 10 min. The resulting RNA products were electrophoretically analyzed in a 5% polyacrylamide-8 M urea slab gel as described in Materials and Methods. The transcripts were transcribed from *E. coli* B DNA (lanes 1 and 2), T2 DNA (lanes 3 and 4), T4 DNA (lanes 5 and 6), T6 DNA (lanes 7 and 8), T5 DNA (lanes 9 and 10), and spliced RNA products from *Hpa*I-linearized pSP64td DNA (lane c). The odd-numbered lanes show RNA synthesized at 30°C; the even-numbered lanes show RNA synthesized at 30°C and processed at 60°C. The RNA size markers (1.4 and 0.56 kb) were transcripts of pSP-$\lambda$ marker DNA (Promega Biotec). The gel positions of the T4 *td* 1-kb circular intron RNA (c-In), 2.7-kb primary transcript (PM), 1.7-kb spliced transcript (MM), and 1-kb linear intron RNA (1-In) are indicated.

On the basis of the observed variation in hybridizing intensity of the 1.7-kb DNA fragment derived from different T-even DNAs, one may speculate that multiple copies of the target sequence may be present in the more strongly hybridizing T2 and T4*alc*7 DNA fragments. We believe that this is not the case for the following reasons. First, restriction of phage DNA with *Taq*I was incomplete. Second, although *Taq*I is able to restrict DNA containing cytosine or glucosylated hydroxymethylcytosine, we have observed that cytosine-containing (in T4*alc*7) and underglucosylated hydroxymethylcytosine-containing (in T2) (25) DNAs are better substrates. Therefore, the observed difference in hybridization in Fig. 3 most probably reflected various amounts of the 1.7-kb fragment generated from different phage DNAs. The data presented in Fig. 3, when analyzed in conjunction with those in Fig. 2B that show the presence of contiguous exon I-intron and intron-exon II sequences in T2 and T6 DNA, allow us to reasonably conclude that the T4 *td* intronlike segment in these phages interrupts their respective thymidylate synthase structural genes in precisely the same location as that in T4. Further confirmation was afforded by the experiment described below.

**Linear and circular forms of *td* intron RNA in T2 and T6 phage-specific transcripts.** T4 phage *td* gene expression invokes autocatalytic processing of the *td* primary transcript involving intron excision-cyclization and exon ligation (8, 11), which are characteristic of class I introns (32). The mechanism and requirements (guanosine and $Mg^{2+}$) of *td* RNA splicing were elucidated first by cloning the T4 phage gene into a SP6 promoter-controlled vector, followed by detailed study of the transcriptional and posttranscriptional processing events in vitro (10). In these experiments, *td* self-splicing was shown to be temperature dependent, being dramatically inhibited at 30°C or lower.

The T-even phages all contain modified cytosine residues in the form of hydroxymethylcytosines which are sometimes glucosylated, and this makes the DNA of these phages extremely resistant to restriction by most enzymes (18). Cloning of many T4 genes has been made possible by the construction of T4 phage mutants whose DNA contains unmodified cytosine to various degrees (19, 34). Unfortunately, cytosine-containing mutants of T2 and T6 phages are not available, making the cloning of genes from these two phages quite difficult. To investigate the nature of the intron sequences in the *td* gene of these phages, transcripts were synthesized from total phage DNA with *E. coli* RNA polymerase at the nonsplicing temperature of 30°C and then induced to splice at 60°C. As the *td* gene product is expressed early during infection by T-even phages, it is reasonable to expect that *E. coli* RNA polymerase would transcribe the *td* genes in vitro. Figure 4 shows an electrophoretic analysis of in vitro-transcribed products in a 5% polyacrylamide slab gel. In all cases, discrete RNA bands were not evident under nonsplicing conditions (lanes 1, 3, 5, 7, and 9). However, a slowly migrating RNA species (c-In) was generated after incubation at 60°C of RNA transcripts synthesized from T2, T4, and T6 DNAs (lanes 4, 6, and 8, respectively). This band, which comigrated with the circular form of the T4 phage *td* intron RNA formed from the cloned gene in vitro (lane c), was missing from *E. coli* and T5 phage transcripts (lanes 2 and 10, respectively).

Northern hybridization analysis of the T2, T4, and T6 RNAs after incubation at 60°C (Fig. 4, lanes 4, 6, and 8, respectively) with a T4 phage *td* intron-specific oligodeoxynucleotide 21-mer probe (Table 1) revealed two discrete hybridizing bands in each case; one band comigrated with the T4 phage *td* circular intron RNA (c-In migrating aberrantly at approximately 5 kb) and the other comigrating with the 1-kb linear form (Fig. 5, lanes 2, 3, and 4 for T2, T4, and T6, respectively). The linear form, though present, could not be resolved from other T2, T4, or T6 transcripts of similar size (Fig. 4, lanes 4, 6, and 8). Consistent with the notion that *E. coli* and T5 phage may not contain T4 phage *td* intronlike sequences was the absence of hybridizing RNA species from their corresponding in vitro transcripts (Fig. 5, lanes 1 and 5, respectively). However, it cannot be unequivocally ruled out that the failure to detect T4 *td* intronlike sequences in T5 phage and in *E. coli* might reflect a very low level of transcripts containing these sequences, particularly in the case of *E. coli*.

A similar hybridization analysis was performed on RNA made in T-even-phage-infected cells. Total RNA extracted from T2, T4, or T6 phage-infected cells at 10 min postinfection showed a similar pattern of RNA hybridization to the intron-specific 21-mer probe (data not shown), eliminating the possibility of in vitro transcription artifacts. Therefore, the data from both in vitro and in vivo systems lead to the conclusion that each T-even phage contains a 1-kb intron present in comparable location within the respective *td* gene. Furthermore, hybridization analyses suggest that the respective *td* intron shares a high degree of sequence homology and a similar mechanism of self-excision and cyclization. Finally, *td* intron RNA excision-cyclization observed at 60°C in vitro (Fig. 4 and 5) corroborates our previous finding that

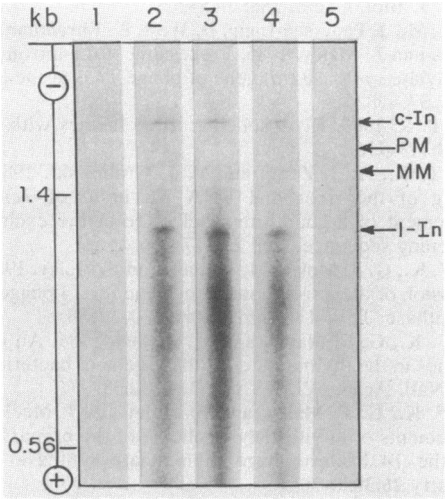


FIG. 5. Northern hybridization analysis of in vitro transcripts processed at 60°C. Unlabeled versions of the transcripts shown in Fig. 4, lanes 2, 4, 6, 8, and 10, were subjected to 5% gel electrophoresis on a 1.5-mm-thick slab gel, electroblotted to a nylon membrane, and then analyzed for an intron-specific sequence by hybridization to a $^{32}$P-labeled oligodeoxynucleotide 21-mer complementary to a region in the 5′ end of the T4 phage *td* intron (nucleotides 774 to 794 in Table 1). Hybridized RNA bands were visualized by autoradiography. The gel lanes show RNA from uninfected *E. coli* B cells (lane 1) and from *E. coli* B infected with phage T2 (lane 2), T4 (lane 3), T6 (lane 4), and T5 (lane 5). The RNA size markers and abbreviations are as described in the legend to Fig. 4.

the *td* RNA self-splicing reaction in vitro is highly temperature dependent (10).

**Other candidates for self-splicing introns in the T-even phages.** The mechanism of RNA self-splicing was first deciphered for *T. thermophila* for the processing of its nuclear large rRNA precursor (7, 17, 35). The first of three transesterification reactions (6) occurs at the 5′ splice site and requires the participation of a guanosine cofactor (GTP, GDP, GMP, or guanosine) which is covalently joined to the 5′ end of the intron, resulting in the release of exon I RNA and formation of a guanosine-intron-exon II intermediate (Fig. 6A). A second transesterification reaction at the 3′ splice site results in the ligation of exon I to exon II RNA and the production of a linear guanosine-intron RNA. Cyclization of the guanosine-intron linear RNA via another transesterification results in the loss of an oligonucleotide containing the added guanosine from the 5′ end of the RNA (reaction not shown in Fig. 6A). The covalent association of the noncoded guanosine with the 5′ end of released linear intron RNA has been observed for all known class I self-splicing introns (7, 13, 17, 21, 30). We therefore decided to utilize this reaction to screen for potential self-splicing introns in the T-even phages. Multiple $^{32}$P-labeled RNA species (Fig. 6B, arrows) from T-even-phage-infected *E. coli* cells were observed on incubation of total cellular RNA with [α-$^{32}$P]GTP under self-splicing conditions (40 mM Tris hydrochloride [pH 7.5], 10 mM $MgCl_2$), five from T4-infected cells (Fig. 6B, lane 5), and three each from T2- (lane 3) and T6 (lane 4) -infected cells. However, no tagging of uninfected *E. coli* RNA (lane 2) or T5-infected *E. coli* RNA (not shown) by GTP was observed, implying but not proving the absence of self-splicing introns in the host bacterium and T5 phage, respectively. The specificity of guanosine labeling was shown by the inability of [α-$^{32}$P]ATP to label these RNAs (lane 1, showing T4 RNA). In addition, circular *td* intron RNA did not retain the added labeled GTP which was released along with the 5′ uridine as a dinucleotide during in vitro cyclization (10). All T-even-phage-infected RNAs contained a 1-kb labeled species (band 1) which comigrated with the labeled intron RNA derived from an SP6 *td* transcript in vitro (Fig. 6B, lane 6). In addition, all of the RNAs contained a 0.9-kb (band 2) and a 0.75-kb (band 4) species, suggesting that, other than the *td* gene, two genes common among the T-even phages may contain class I intervening sequences. The T4 phage-infected RNA contained two additional guanosine-labeled RNAs (0.83 and 0.6 kb, represented by bands 3 and 5, respectively) which are absent in the T2- and T6-infected cells. The 0.6-kb species (band 5) is probably the intron RNA excised from the transcript of the *nrdB* gene encoding the T4 phage ribonucleotide reductase small subunit. This gene has been sequenced recently and was found to possess an intron of 625 bp containing an open reading frame of 291 bp (26), making it the second example of an

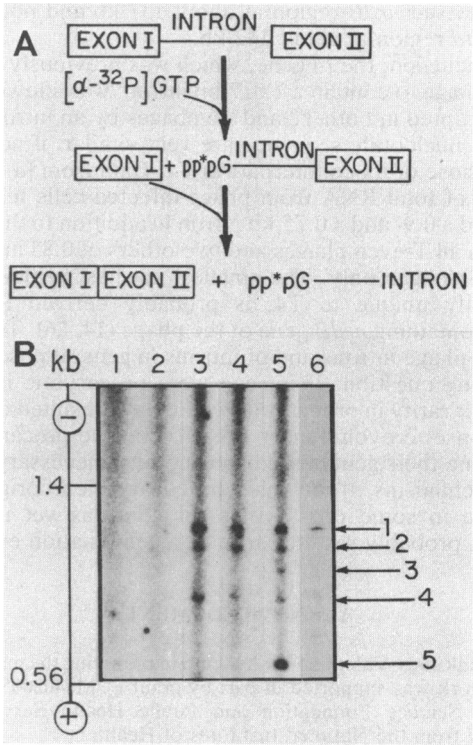


FIG. 6. (A) Reaction scheme showing the generation of [α-$^{32}$P]GTP-labeled linear intron RNA during processing of T4 *td* precursor RNA. This reaction was first reported for the self-splicing of nuclear large rRNA precursor in *T. thermophila* (7) and subsequently was shown to occur with transcripts from the T4 phage *td* gene (11). The guanosine-intron-exon II intermediate is very unstable and is readily converted to guanosine-intron via cleavage at the intron-exon II junction (10). (B) In vitro labeling of total RNA extracted from uninfected and T-even-phage-infected *E. coli* B with [α-$^{32}$P]GTP. The conditions for covalent addition of GTP to RNA in vitro are described in Materials and Methods. Lanes: 1, T4-infected RNA plus [α-$^{32}$P]ATP; 2, uninfected RNA plus [α-$^{32}$P]GTP; 3, T2-infected RNA plus [α-$^{32}$P]GTP; 4, T6-infected RNA plus [α-$^{32}$P]GTP; 5, T4-infected RNA plus [α-$^{32}$P]GTP; 6, in vitro SP6 *td* transcript plus [α-$^{32}$P]GTP. See the legend to Fig. 4 for a description of RNA size markers. The numbers and arrows to the right identify the [α-$^{32}$P]GTP-labeled RNA species according to size.

intron-containing procaryotic protein-encoding gene. In vitro self-splicing activity of the *nrdB* gene transcript was recently demonstrated by Gott et al. (14), who showed that autocatalytic addition of [$\alpha$-$^{32}$P]GTP to total T4 phage-infected RNA occurred with a major 0.6-kb RNA species corresponding to the size of the *nrdB* intron. As both T2 and T6 transcripts (Fig. 6B, lanes 3 and 4, respectively) lacked the 0.6-kb RNA species, it would be reasonable to assume that the *nrdB* gene in T2 and T6 may not be interrupted, at least not by an intron of the self-splicing variety. In this respect, an oligodeoxynucleotide (20-mer) defining the 5′ portion of the *nrdB* intron in the T4 phage failed to hybridize to T2 and T6 phage DNAs (unpublished observation). This is in contrast to the case of the *td* gene, which is interrupted by fairly homologous intervening sequences in all T-even phages (9; this work). An extremely pleasant surprise is that our biochemical data appear to corroborate the data of Kim and Davidson on heteroduplex analysis of T2, T4, and T6 phage DNA (16). Although their heteroduplex maps, constructed in 1974, are somewhat out of date due largely to revision to include newly found T4 phage genes, heteroduplex maps of both T2-T4 and T6-T4 showed a deletion loop in the revised *nrdB* region at about 137 kb and none in the revised *td* region at about 142 kb.

In conclusion, the *td* gene, which was previously found in the T4 phage to contain a 1,017-bp intron, was shown here to be interrupted in both T2 and T6 phages by an intron whose size and nucleotide sequence are very similar, if not identical, to those of its counterpart in T4. Data from [$\alpha$-$^{32}$P]GTP labeling of total RNA from phage-infected cells tentatively identified a 0.9- and a 0.75-kb intron in addition to the 1-kb *td* intron in all T-even phages and two others of 0.83 and 0.6 kb in the T4 phage only. The smallest species, of 0.6 kb, and apparently unique to T4, is probably derived from the intron-containing *nrdB* gene of the phage (14, 26). The origin of these phage introns and of introns in general poses a very interesting question. If introns were procaryotic in origin, then their rarity in procaryotes could have resulted from loss in response to evolutionary pressure on the procaryotes to streamline their genomes, disposing of unnecessary regulatory mechanisms. If introns were eucaryotic in origin, their presence in some procaryotes for some as yet unknown function probably resulted from a recombination event.

## ACKNOWLEDGMENTS

We thank Judy Valentino for her care in preparing the manuscript.

This work was supported in part by grant PCM8402005 from the National Science Foundation and Public Health Service grant GM26387 from the National Institutes of Health.

## LITERATURE CITED

1. **Abelson, J.** 1979. RNA processing and the intervening sequence problem. Annu. Rev. Biochem. **48**:1035–1067.
2. **Barner, H. D., and S. S. Cohen.** 1959. Virus-induced acquisition of metabolic function. IV. Thymidylate synthetase in thymine-requiring *Escherichia coli* infected by T2 and T5 bacteriophages. J. Biol. Chem. **234**:2987–2991.
3. **Belfort, M., G. F. Maley, J. Pedersen-Lane, and F. Maley.** 1983. Primary structure of the *Escherichia coli* thyA gene and its thymidylate synthase product. Proc. Natl. Acad. Sci. USA **80**: 4914–4918.
4. **Belfort, M., A. Moelleken, G. F. Maley, and F. Maley.** 1983. Purification and properties of the T4 phage thymidylate synthase produced by the cloned gene in an amplification vector. J. Biol. Chem. **258**:2045–2051.
5. **Belfort, M., J. Pedersen-Lane, D. West, K. Ehrenman, G. Maley, F. Chu, and F. Maley.** 1985. Processing of the intron-containing thymidylate synthase (*td*) gene of phage T4 is at the RNA level. Cell **41**:375–382.
6. **Cech, T. R.** 1983. RNA splicing: three themes with variations. Cell **34**:713–716.
7. **Cech, T. R., A. J. Zaug, and P. J. Grabowski.** 1981. In vitro splicing of the ribosomal RNA precursor of *Tetrahymena*: involvement of a guanosine nucleotide in the excision of the intervening sequence. Cell **27**:487–496.
8. **Chu, F. K., G. F. Maley, M. Belfort, and F. Maley.** 1985. In vitro expression of the intron-containing gene for T4 phage thymidylate synthase. J. Biol. Chem. **260**:10680–10688.
9. **Chu, F. K., G. F. Maley, and M. Belfort.** 1984. An intervening sequence in the thymidylate synthase gene of bacteriophage T4. Proc. Natl. Acad. Sci. USA **81**:3049–3053.
10. **Chu, F. K., G. F. Maley, and F. Maley.** 1987. Mechanism and requirements of in vitro RNA splicing of the primary transcript from the T4 bacteriophage thymidylate synthase gene. Biochemistry **26**:3050–3057.
11. **Chu, F. K., G. F. Maley, D. K. West, M. Belfort, and F. Maley.** 1986. Characterization of the intron in the phage T4 thymidylate synthase gene and evidence for its self-excision from the primary transcript. Cell **45**:157–166.
12. **Clewell, D. B., and D. R. Helinski.** 1969. Supercoiled circular DNA-protein complex in *Escherichia coli*: purification and induced conversion to an open circular DNA form. Proc. Natl. Acad. Sci. USA **62**:1159–1166.
13. **Garriga, G., and A. M. Lambowitz.** 1984. RNA splicing in *Neurospora* mitochondria: self-splicing of a mitochondrial intron in vitro. Cell **39**:631–641.
14. **Gott, J. M., D. A. Shub, and M. Belfort.** 1986. Multiple self-splicing introns in bacteriophage T4: evidence from autocatalytic GTP labeling of RNA in vitro. Cell **47**:81–87.
15. **Kaine, B. P., R. Gupta, and C. R. Woese.** 1983. Putative introns in tRNA genes of prokaryotes. Proc. Natl. Acad. Sci. USA **80**: 3309–3312.
16. **Kim, J.-S., and N. Davidson.** 1974. Electron microscope heteroduplex study of sequence relations of T2, T4, and T6 bacteriophage DNAs. Virology **57**:93–111.
17. **Kruger, K., P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech.** 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of *Tetrahymena*. Cell **31**:147–157.
18. **Kutter, E., and W. Rüger.** 1983. Map of the T4 genome and its transcription control sites, p. 277–290. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (eds.), Bacteriophage T4. American Society for Microbiology, Washington, D.C.
19. **Kutter, E., and L. Snyder.** 1983. Production of cytosine-containing T4 phage, p. 56–57. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (eds.), Bacteriophage T4. American Society for Microbiology, Washington, D.C.
20. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
21. **Michel, F., and D. J. Cummings.** 1985. Analysis of class I introns in a mitochondrial plasmid associated with senescence of *Podospora anserina* reveals extraordinary resemblance to the *Tetrahymena* ribosomal intron. Curr. Genet. **10**:69–79.
22. **Michel, F., A. Jacquier, and B. Dujon.** 1982. Comparison of fungal mitochondrial introns reveals extensive homologies in RNA secondary structure. Biochimie **64**:867–881.
23. **Mileham, A. J., H. R. Revel, and N. E. Murray.** 1980. Molecular cloning of the T4 genome: organization and expression of the *frd*-DNA ligase region. Mol. Gen. Genet. **179**:227–239.
24. **Purohit, S., and C. K. Mathews.** 1984. Nucleotide sequence reveals overlap between T4 phage genes encoding dihydrofolate reductase and thymidylate synthase. J. Biol. Chem. **259**:6261–6266.
25. **Revel, H. R.** 1983. DNA modification: glucosylation, p. 156–165.

*In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. B. Berget (ed.), Bacteriophage T4. American Society for Microbiology, Washington D.C.

26. **Sjoberg, B.-M., S. Hahne, C. Z. Mathews, C. K. Mathews, K. N. Rand, and M. J. Gait.** 1986. The bacteriophage T4 gene for the small subunit of ribonucleotide reductase contains an intron. EMBO J. **5**:2031–2036.
27. **Tabak, H. F., G. Van der Horst, K. A. Osinga, and A. C. Arnberg.** 1984. Splicing of large ribosomal precursor RNA and processing of intron RNA in yeast mitochondria. Cell **39**:623–629.
28. **Thomas, P. S.** 1983. Hybridization of denatured RNA transferred or dotted to nitrocellulose paper. Methods Enzymol. **100**: 255–266.
29. **Trimble, R. B., and F. Maley.** 1973. In vitro synthesis of deoxynucleotide kinase, dihydrofolate reductase and deoxycytidylate hydroxymethylase from RNA transcripts of T2 phage DNA. Biochem. Biophys. Res. Commun. **52**:1026–1033.
30. **Van der Horst, G., and H. F. Tabak.** 1985. Self-splicing of yeast mitochondrial ribosomal and messenger RNA precursors. Cell **40**:759–766.
31. **Wallace, R. B., M. J. Johnson, T. Hirose, T. Miyake, E. H. Kawashima, and K. Itakura.** 1981. The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequence to rabbit β-globin DNA. Nucleic Acids Res. **9**:879–894.
32. **Waring, R. B., and R. W. Davies.** 1984. Assessment of a model for intron RNA secondary structure relevant to RNA self-splicing: a review. Gene **28**:277-291.
33. **Waring, R. B., R. W. Davies, C. Scazzocchio, and T. A. Brown.** 1982. Internal structure of a mitochondrial intron of *Aspergillus nidulans*. Proc. Natl. Acad. Sci. USA **79**:6332–6336.
34. **Wilson, G. G., V. I. Tanyashin, and N. E. Murray.** 1977. Molecular cloning of fragments of bacteriophage T4 DNA. Mol. Gen. Genet. **156**:203–214.
35. **Zaug, A. J., P. J. Grabowski, and T. R. Cech.** 1983. Autocatalytic cyclization of an excised intervening sequence RNA is a cleavage-ligation reaction. Nature (London) **301**:578–583.