of k_3. As $1/[S]$ gets large, eq 9 takes the following form in the asymptote:

$$1/k_{obsd} = [(k_2 + k_3)/k_1k_3](1/[S]) + k_2/[k_3(k_2 + k_3)] \quad (11)$$

Note that the Y-axis intercept of this asymptotic solution is nonzero for nonzero k_2. In the limiting case of $k_2 \gg k_3$, eq 10 and 11 become equal and the observed plot is linear. If $k_2 \ll k_3$, the scheme becomes one of consecutive first-order irreversible reactions and the relative amplitudes of the two exponentials depends on the relative values of $k_1[S]$ and k_3. Whichever rate is smaller will have the larger amplitude.

In general, double-reciprocal plots like Figure 1 have three measurable parameters: the Y-axis intercept (inter), the initial slope (inits), and the asymptotic slope (asymps). These can be combined to give the three rate constants as the initial point for a nonlinear least-squares analysis, viz.

$$k_1 = \frac{1}{(\text{asymps} - \text{inits})}$$

$$k_2 = \frac{1}{\text{inter}[(\text{asymps}/\text{inits}) - 1]}$$

$$k_3 = \frac{1}{\text{inter}}$$

Depending upon the degree of curvature in the double-reciprocal plot, the asymptotic slope may be severely underestimated if a simple graphical procedure is used. Conversely, if $k_2 \gg k_3$, the plot will appear to be linear and the asymptotic slope can be determined readily. In other words, the rapid equilibrium assumption will be valid. In cases of pronounced curvature, as seen in Figure 1, it seems that much better results are obtained by fitting all of the data at once rather than attempting to determine the parameters graphically.

References

Bevington, P. R. (1969) *Data Reduction and Error Analysis for the Physical Sciences*, McGraw-Hill, New York.

Chiba, K., Peterson, L. A., Castagnoli, K. P., Trevor, A. J., & Castagnoli, N., Jr. (1985) *Drug Metab. Dispos. 13*, 343–347.

Hiromi, K. (1979) *Kinetics of Fast Enzyme Reactions*, Halsted Press, New York.

Husain, M., Edmondson, D. E., & Singer, T. P. (1982) *Biochemistry 21*, 595–600.

Marquardt, D. W. (1963) *J. Soc. Ind. Appl. Math. 11*, 431–441.

Pearce, L. B., & Roth, J. A. (1985) *Biochemistry 24*, 1821–1826.

Ramsay, R. R., Husain, M., & Steenkamp, D. J. (1987) *Biochem. J. 241*, 883–892.

Salach, J. I. (1979) *Arch. Biochem. Biophys. 192*, 128–137.

Salach, J. I., Singer, T. P., Castagnoli, N., Jr., & Trevor, A. (1984) *Biochem. Biophys. Res. Commun. 125*, 831–835.

Singer, T. P., Salach, J. I., & Crabtree, D. (1985) *Biochem. Biophys. Res. Commun. 127*, 707–712.

Singer, T. P., Salach, J. I., Castagnoli, N., Jr., & Trevor, A. (1986a) *Biochem. J. 235*, 785–789.

Singer, T. P., Salach, J. I., Castagnoli, N., Jr., & Trevor, A. (1986b) in *MPTP: A Neurotoxin Producing a Parkinsonian Syndrome* (Markey, S. P., Castagnoli, N., Jr., Trevor, A., & Kopin, I. J., Eds.) pp 235–251, Academic, New York.

Strickland, S., Palmer, G., & Massey, V. (1975) *J. Biol. Chem. 250*, 4048–4052.

Mechanism and Requirements of in Vitro RNA Splicing of the Primary Transcript from the T4 Bacteriophage Thymidylate Synthase Gene†

Frederick K. Chu,* Gladys F. Maley, and Frank Maley

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

Received November 17, 1986; Revised Manuscript Received January 7, 1987

ABSTRACT: The splicing of a procaryotic precursor RNA transcribed from the T4 phage thymidylate synthase (*td*) gene with SP6 RNA polymerase was investigated in vitro. The intron excision–cyclization reaction increased progressively to 60 °C. Exon ligation, though barely detectable at the lower temperatures, was greatly enhanced at 60 °C. Both reactions required Mg^{2+}. The addition of guanosine to the 5′ end of an intron–exon II intermediate via a 3′,5′-phosphodiester bond was essential for the ligation of exon I to exon II. The added guanosine and the first intron-encoded uridine are subsequently lost as a dinucleotide from the 5′ end during cyclization of the linear form of the excised intron RNA. Exon ligation is intramolecular and occurs more readily in the nascent RNA molecule (cotranscriptionally) than in the finished transcript (posttranscriptionally). These data and the identification of various structural elements (P, Q, R, S, E, E′) in the *td* intron that are found typically in eucaryotic class I introns firmly establish the *td* intron as the first example of class I intron of procaryotic origin.

The presence of introns in eucaryotic genes has been amply documented. The expression of these genes involves RNA splicing reactions that are either small nuclear ribonucleoprotein particle (snRNP) dependent (Lerner et al., 1980; Mount et al., 1983; Greer et al., 1983; Peebles et al., 1983; Pikielny et al., 1983; Padgett et al., 1983; Kramer et al., 1984) or self-splicing (Cech et al., 1981; Kruger et al., 1983; Zaug et al., 1983; Tabak et al., 1984; Garriga & Lambowitz, 1984). Examples of self-splicing RNAs include the precursors of *Tetrahymena thermophila* nuclear rRNA, yeast and fungal mitochondrial rRNA and mRNA, and some chloroplast

† This investigation was supported in part by Public Health Grant GM26387 from the National Institutes of General Medical Science, U.S. Public Health Service, Department of Health and Human Services, and by National Science Foundation Grant PCM8402005.

0006-2960/87/0426-3050$01.50/0

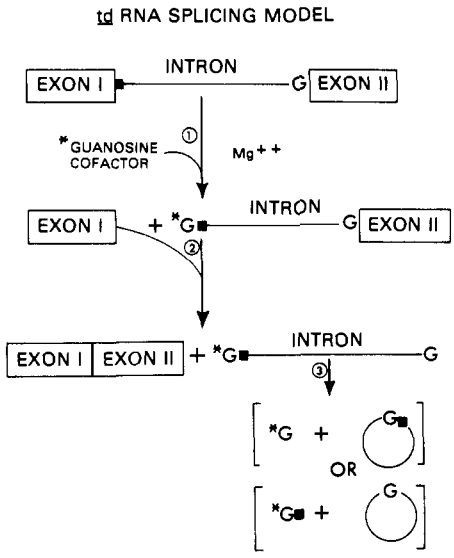


FIGURE 1: RNA splicing model for the *td* primary transcript. The scheme shown is similar to that proposed for the *T. thermophila* nuclear large rRNA precursor (Cech, 1983). Three transesterification reactions (1–3) are involved that do not require participation of proteins or an external energy source. The solid square represents one or more consecutive bases that may be cleaved off together with the noncoded G during intron cyclization.

tRNAs. These self-splicing introns, collectively known as class I introns, are defined by a set of conserved oligonucleotide sequences (P or A, Q or B, R or 9L, S or 2, E or 9R′, E′ or 9R) (Michel et al., 1982; Michel & Dujon, 1983; Waring et al., 1983; Cech et al., 1983; Waring & Davies, 1984). RNA splicing in this class is characterized by a series of three transesterification reactions in vitro in the absence of proteins (Cech, 1983), schematically depicted in Figure 1. The first transesterification involves a guanosine added to the 5′ end of the intron, releasing the 5′ exon; the second involves the ligation of the 5′ and 3′ exons, releasing the intron as a linear molecule; the third joins the 3′ nucleotide (invariably a G) of the intron to a nucleotide near the 5′ end of the intron to form a circular intron molecule, releasing an oligonucleotide containing the added guanosine.

We have described the presence of a 1017 base pair intervening sequence in the thymidylate synthase (*td*) gene of the T4 bacteriophage (Chu et al., 1984). It was shown that expression of this gene involves RNA splicing (Chu et al., 1985; Belfort et al., 1985) and, furthermore, that splicing in vitro occurs in an autocatalytic manner (Chu et al., 1985). Our recent determination of the intron nucleotide sequence in the *td* gene confirms its identity as a member of the class I self-splicing introns (Chu et al., 1986). The *td* intron not only contains the conserved structural elements typical of class I introns but also possesses a 735 nucleotide open reading frame characteristic of many yeast and fungal mitochondrial introns.

In this work, we have investigated the properties and requirements for splicing of the *td* primary transcript. We report here that, mechanistically, the overall splicing reaction and the individual reactions of intron excision–cyclization and exon ligation proceed in a manner reminiscent of class I self-splicing RNAs in eucaryotes.

Experimental Procedures

Plasmids, Phage, and Bacterial Strains. The 2.85-kb[1] *Eco*RI fragment containing the *td* gene provided by Noreen Murray was cloned from phage strain T4 *alc*4 into a phage λ vector (Mileham et al., 1980). This *td* fragment was subcloned into pSP64 plasmid (Promega Biotec) to generate pSP64td (Chu et al., 1986). The 1.6-kb *Hin*dIII–*Eco*RV subfragment containing truncated exon I, entire intron, and truncated exon II was also cloned into pSP64 to generate pSP64-1.6 (Chu et al., 1986). Both pSP64 derivatives contain gene inserts in transcriptional alignment with the SP6 promoter. *Escherichia coli* strain JM83 was used for the propagation of SP6 recombinant plasmids.

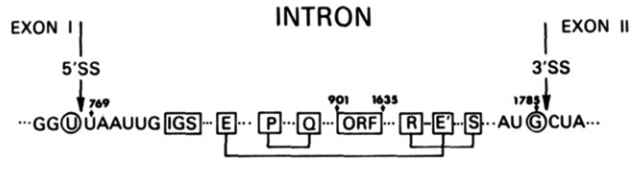
Enzymes and Chemicals. Restriction enzymes *Eco*RI, *Eco*RV, and *Hin*dIII were obtained from New England Biolabs while *Hpa*I was from Bethesda Research Laboratories. SP6 RNA polymerase and transcription size markers were from Promega Biotec. Nucleoside triphoshates were obtained from Boehringer Mannheim. The other nucleotides were dGTP (ultrapure), pd(GC), and pd(GT)$_2$ from Pharmacia, GMP from P-L Biochemicals, and guanosine and guanine from Mann Research Laboratories. D-Ribose was from Nutritional Biochemicals Corp. [γ-^{32}P]ATP (3000 Ci/mmol) for end labeling the synthetic oligodeoxynucleotides was from New England Nuclear Corp. [α-^{32}P]ATP and [α-^{32}P]GTP (each at 400 Ci/mmol) were from Amersham.

Preparation and Linearization of DNA Templates. Recombinant plasmids pSP64td and pSP64-1.6 amplified in *E. coli* JM83 were purified by banding in CsCl–ethidium bromide gradients (Clewell & Helinski, 1969). Plasmid DNA at 0.5 mg/mL was treated with restriction enzyme at a substrate to enzyme ratio of 1 μg to 5 units for 2–4 h at 37 °C, heated at 65 °C for 10 min, and then precipitated with 2 volumes of ethanol in the presence of 0.2 M NaCl at −20 °C. The washed DNA precipiate was dissolved in sterile water to a final concentration of 0.5–1 mg/mL and stored at −20 °C to be used as template for in vitro transcription.

In Vitro Transcription and Isolation of RNA Products. Transcription of linearized SP6 recombinant plasmid DNA with SP6 RNA polymerase was as described previously (Chu et al., 1985). Where radioactive RNA was desired, [α-^{32}P]-ATP was added at a concentration 5 μCi/5 μL of reaction. After transcription, the DNA template was digested with DNase and the synthesized RNA was isolated free of proteins, ribonucleotides, and salts by passage through a Nensorb 20 cartridge (Du Pont). Elution of bound RNA from the cartridge was effected with 20% ethanol followed by precipitation with 2 volumes of prechilled ethanol in the presence of 0.2 M NaCl and 0.1 mg/mL of yeast tRNA as carrier at −20 °C. The RNA precipitate was washed twice with 70% ethanol, dissolved in a volume of sterile water equal to that of the original transcription mixture, and stored at −70 °C until use.

In Vitro RNA Splicing Reaction and Gel Electrophoretic Analysis. Two splicing buffers were used. The first was used initially (in Figure 3) and contained all the components of the transcription reaction (minus SP6 RNA polymerase) as reported previously (Chu et al., 1986). The other was used throughout this work (except in Figure 3) and contained 40 mM tris(hydroxymethyl)aminomethane hydrochloride (pH 7.5), 10 mM $MgCl_2$, and 100 μM GTP (or GMP). Where [α-^{32}P]GTP was used, 25 μCi of the nucleotide (400 Ci/mmol) was added in place of unlabeled GTP to 5 μL of splicing mixture. The final chemical concentration of [α-^{32}P]GTP was 2.4 μM. The spliced RNA products were electrophoresed in a 0.75 mm thick slab gel containing 5% polyacrylamide and 8 M urea at room temperature for 4 h at 10 mA/slab. Processing of the gel for autoradiography was as described previously (Chu et al., 1986). For analysis of oligonucleotides

[1] Abbreviations: kb, kilobase; bp, base pair; IGS, internal guide sequence; ORF, open reading frame; G, guanylate; U, uridylate; A, adenylate; C, cytidylate; T, thymidylate.



		T4 td	CONSENSUS
IGS	(775)	AGGCCUGAG	NONCONSERVED
E	(807)	UCUAUC	NNNC
P	(814)	AAACGGGGAA	AAGUGCUGGA G A
Q	(838)	AAUCC CGUGC	AAU CA GCA GG
R	(1636)	UAACGAC UAUCC	UCAGAGACUACA
E'	(1679)	GAUAGA	G NNN
S	(1705)	GAGAUAUAGUCU	AAGAUAUAGUCC

FIGURE 2: Proposed sequence determinants of self-splicing in the T4 phage *td* intron. The order of the conserved and nonconserved sequence determinants indicated in boxes in the upper panel (U 5′ to intron and G in 3′ end of intron, IGS, E, P, Q, R, E′, and S sequences) is typical of most class I self-splicing eucaryotic introns. However, the distances in between the sequences do vary widely. The sequences proposed to base pair with each other (E with E′, P with Q, and R with S) are joined by lines. The internal guide sequence (IGS) near the 5′ end of the intron is postulated to align the exons for splicing by base pairing with sequences in the 5′ and 3′ splice sites (Davies et al., 1982). The intron open reading frame (ORF) is situated between the Q and R sequences, similar to the chloroplast ribosomal intron ORF in *Chlamydomonas reinhardii* (Rochaix et al., 1985). The table in the lower panel compares these sequences found in the *td* intron with the corresponding consensus sequences in eucaryotes. The nucleotides potentially involved in base pairing are underscored. The number in parentheses to the left of each T4 *td* intron sequence determinant indicates the position of the first nucleotide in each case. Numbering is from the 5′ end of the original *Eco*RI *td* containing 2.85-kb DNA fragment derived from the *T4 alc*4 phage.

or mononucleotides, a 20% polyacrylamide–urea gel was used.

Dot Hybridization Analysis of RNA. Isolated RNA (0.1 pmol in 1 μL) was spotted on Hybond nylon membrane (Amersham). The air-dried membrane containing the RNA dots was wrapped in Saran wrap and irradiated with ultraviolet light for 5 min to fix the RNA onto the membrane. The hybridization buffer contained 5 × SSPE buffer [0.05 M sodium phosphate (pH 7.7), 0.9 M NaCl, 0.005 M disodium ethylenediaminetetraacetate], 50% deionized formamide, 10% dextran sulfate, and 0.01% each of bovine serum albumin, Ficoll, and poly(vinylpyrrolidine). Prehybridization for 3 h (in the presence of 0.1 mg/mL of sonicated salmon sperm DNA) and hybridization for 16 h (in the presence of ^{32}P-labeled synthetic oligodeoxynucleotide probe) were performed at 48 °C with gentle shaking. After hybridization, the membrane was washed at room temperature first with 5 × SSPE buffer (3 times) and then with 1 × SSPE buffer (3 times) before autoradiography. The washing procedure removes unhybridized and partially hybridized oligodeoxynucleotides.

RESULTS

Structural Determinants for Self-Splicing of td Primary Transcript. We have demonstrated previously the in vitro self-splicing capacity of the primary transcript from the intron-containing *td* gene encoding thymidylate synthase of the T4 phage (Chu et al., 1985). Furthermore, the *td* intron contains the four conserved internal sequences (P, Q, R, S) as well as the internal guide sequence (IGS) characteristic of eucaryotic class I introns (Chu et al., 1986). As the conserved sequences were initially identified on the basis of nucleotide sequence alone, they were tentative at best, as pointed out to us by Michel and Dujon (personal communication). The method of comparative secondary structure analysis (Michel et al., 1982; Michel & Dujon, 1983; Michel & Cummings, 1985) should facilitate a more accurate identification of these conserved sequences as well as of the nonconserved ones known

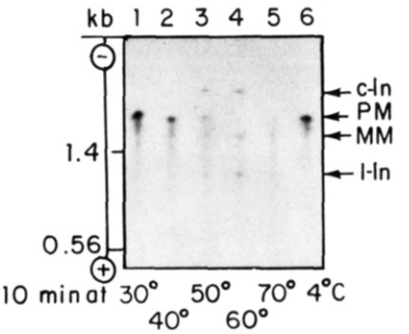


FIGURE 3: Temperature dependence of splicing of isolated *td* precursor RNA. ^{32}P-Labeled *td* 2.7-kb precursor RNA was synthesized at 28 °C with SP6 RNA polymerase and then isolated with a Nensorb 20 cartridge. Aliquots of the isolated RNA (50000 cpm) were incubated for 10 min in 5 μL of transcription buffer at 30 (lane 1), 40 (lane 2), 50 (lane 3), 60 (lane 4), and 70 °C (lane 5). Unprocessed RNA shown in lane 6 was incubated at 4 °C. The resulting RNA products were analyzed in 5% polyacrylamide–8 M urea slab gel electrophoresis and visualized by autoradiography. Abbreviations: PM, 2.7-kb *td* precursor RNA; MM, 1.7-kb mature mRNA; c-In and l-In, circular and linear forms of intron RNA, respectively. The RNA size markers (1.4 and 0.56 kb) were transcripts of pSP-λ marker DNA (Promega Biotec).

as E (9R′) and E′ (9R). The *td* intron sequence was thus reanalyzed with this method. Figure 2 shows the relative positions and structures of the conserved and nonconserved sequence elements (E, P, Q, R, E′, S) and the IGS in the 1017 base long intron. The probable base pairings are between P and Q, R and S, and E and E′. In our previous work, the elements E and E′ were not identified and the elements P and Q were wrongly assigned (Chu et al., 1986). The secondary structure predicted by the comparative method for the *td* intron RNA (Michel & Dujon, 1986) shares numerous features in common with those in the *Chlamydomonas* chloroplast ribosomal intron RNA (Rochaix et al., 1985).

Effect of Temperature on Self-Splicing of td Primary Transcript. Our working model for the processing of the *td* precursor RNA as depicted in Figure 1 shows many characteristics reminiscent of RNA splicing in the *T. thermophila* nuclear large rRNA precursor (Cech et al., 1981). The determinants of the self-splicing property should reside in the RNA sequence, which in turn determines the higher order of RNA structure essential for splicing in the absence of protein factors in vitro. Since the secondary and tertiary structures of RNA folding are temperature dependent, splicing should be coordinately sensitive to temperature. Figure 3 shows the effect of temperature on the extent of self-splicing of isolated SP6 transcript (2.7 kb) of the *td* gene. The spliced products (MM for ligated exon I–exon II RNA; c-In and l-In for circular and linear intron RNAs, respectively) were identified by hybridization analysis with synthetic oligodeoxynucleotide probes specific for various regions of the *td* gene. Although the yield of linear and circular forms of the 1-kb intron RNA increased progressively with temperature (lanes 1–4), the formation of the 1.7-kb ligated exon I–exon II RNA (MM) occurred sharply at 60 °C (lane 4). It should be noted that the 2.7-kb precursor RNA (PM) was completely processed at this temperature to c-In, MM, and l-In. At 70 °C, the yield of all the products of splicing diminished (lane 5), probably due to melting of secondary structures essential for efficient RNA cleavage and ligation reactions.

Cofactor Requirements for Self-Splicing of td Primary Transcript Involving Intron Excision and Exon Ligation. To examine the requirements for splicing of the *td* precursor RNA, various components originally present in the SP6 RNA polymerase in vitro transcription system were omitted from

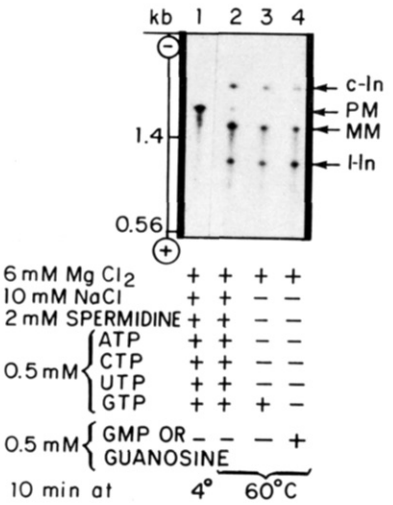


FIGURE 4: Chemical requirements of splicing of *td* precursor RNA. Isolated ^{32}P-labeled *td* precursor RNA was incubated for 10 min at 60 °C in transcription buffer with various omissions or substitutions as indicated in the lower panel. The resulting products were electrophoretically analyzed in 5% polyacrylamide–8 M urea slab gel and visualized by autoradiography. The abbreviations and size markers are the same as described in the Figure 3 legend.

the splicing reaction, with [α-^{32}P]ATP-labeled 2.7-kb SP6 *td* transcript as the starting RNA substrate. As shown in Figure 4, the only absolute splicing requirements were Mg^{2+} ion and a guanosine cofactor (lanes 3 and 4). As indicated, GTP could be replaced by guanosine, GMP, or GDP (data not shown) without affecting the yield of spliced products, but not by guanine or D-ribose (data not shown). Other ribonucleotides (ATP, CTP, UTP) were inactive in promoting splicing (data not shown), as was dGTP (Figure 5B, lane 9). In order to determine the minimum concentration of Mg^{2+} and GTP required for splicing, incubation at 60 °C was shortened to 2 min instead of the usual 10 min. RNA splicing in the presence of 100 μM GTP reached a maximum at 3 mM Mg^{2+} (Figure 5A, lane 5) and was not affected on increasing Mg^{2+} to 20 mM (lane 8). For GTP (or GMP or guanosine), splicing occurred at a minimum of 0.2 μM in the presence of 10 mM $MgCl_2$ (Figure 5B, lane 4). However, increasing GTP to 1 μM doubled the yield of spliced products (Figure 5B, lane 6).

A comparison of the amount and nature of the spliced RNA products made in the absence of either Mg^{2+} or GTP reveals a striking difference between the two conditions. Omission of Mg^{2+} from the splicing reaction halted the production of both excised intron and ligated exon RNAs (Figure 5A, lane 2). In contrast, omission of GTP abolished the formation of ligated exon RNA but still permitted the formation of both linear and circular forms of intron RNA (Figure 5B, lane 2), although to a lesser extent. Quantitation of the radioactive RNA in the intron bands showed a 50% reduction in the circular form and a 70% reduction in the linear form.

Mechanics of the Guanosine Cofactor Requirement. To investigate the mechanics of the guanosine requirement in splicing of *td* precursor RNA, unlabeled SP6 *td* transcript (2.7 kb) was isolated and its interaction with [α-^{32}P]GTP was examined. In the presence of 10 mM Mg^{2+} and at 60 °C for only 2 min, [α-^{32}P]GTP was rapidly incorporated into two species of RNA of 1 and 1.9 kb, corresponding in size to the linear intron RNA and intron–exon II intermediate, respectively (Figure 6, lane 2). The point of GTP addition was implied from the following results to be at the 5′ end of the intron segment via a 3′,5′-phosphodiester bond. First, the 2.7- and 1.7-kb RNAs, corresponding to the precursor RNA (PM) and ligated exon RNA (MM), respectively, were not labeled

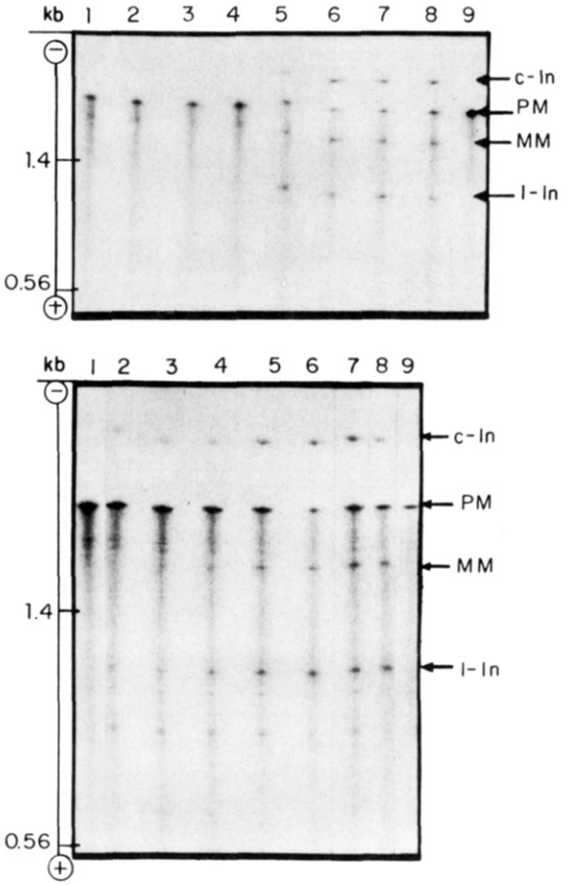


FIGURE 5: Determination of splicing requirements. Incubation was carried out for 2 min instead of 10 min at 60 °C. (A) Magnesium chloride concentrations of 0 (lane 2), 1 (lane 3), 2 (lane 4), 3 (lane 5), 5 (lane 6), 10 (lane 7), and 20 mM (lane 8) were tested in the *td* splicing reaction containing 100 μM GTP. Unspliced controls are shown in lane 1 (PM incubated at 4 °C in 10 mM $MgCl_2$ and 100 μM GTP) and in lane 9 (PM incubated at 60 °C in the absence of $MgCl_2$ and GTP). (B) Concentrations of 0 (lane 2), 0.05 (lane 3), 0.2 (lane 4), 0.5 (lane 5), and 1 μM GTP (lane 6), as well as 1 μM GMP (lane 7), 1 μM guanosine (lane 8), and 1 μM dGTP (lane 9), were tested in the *td* splicing reaction containing 10 mM $MgCl_2$. Unspliced control is shown in lane 1 (PM incubated at 4 °C in 10 mM $MgCl_2$ and 1 μM GTP). See the Figure 3 legend for abbreviations and size markers.

with [α-^{32}P]GTP (Figure 6, lanes 2–5). Second, treatment of the ^{32}P-labeled 1-kb linear intron (l-In) and 1.9-kb intron–exon II intermediate with bacterial alkaline phosphatase completely removed the label (data not shown). Incubation for longer than 10 min in $MgCl_2$ at 60 °C led to the disappearance of the ^{32}P-labeled 1.9-kb RNA species (lane 5), probably due to its conversion to the 1-kb linear intron species by cleavage at the junction between intron and exon II (3′ splice site). The circular intron RNA (c-In) did not contain the noncoded [α-^{32}P]GTP (lanes 2–5) although the formation of c-In species was evident when ^{32}P uniformly labeled PM was used as the splicing substrate (lane 1). These data suggest that the guanosine cofactor, which was originally added to the 5′ end of the intron, was released during intron cyclization. Omission of Mg^{2+} resulted in no incorporation of GTP into any RNA species (lane 6). Furthermore, if [α-^{32}P]GTP was replaced with [α-^{32}P]ATP, none of the splicing RNA intermediates was labeled (lane 7).

Intron Cyclization under Splicing Conditions Results in Release of Dinucleotide GU from the 5′ End. Autocyclization of the linear intron molecule in vitro during processing of the *T. thermophila* nuclear large rRNA precursor led to the release of a 15-mer oligonucleotide containing the added gua-

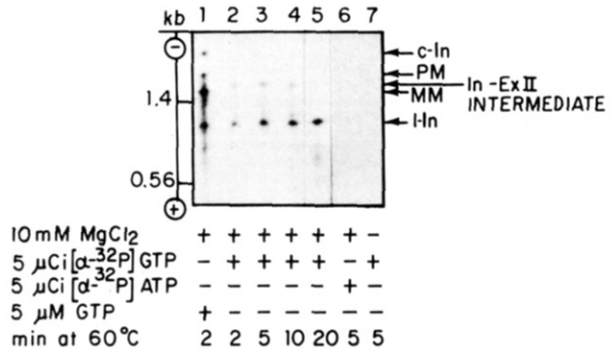


FIGURE 6: Guanosine addition during splicing of *td* precursor RNA. Isolated unlabeled *td* 2.7-kb precursor RNA (0.5 pmol) was incubated with 25 μCi of [α-^{32}P]GTP (400 Ci/mmol) in the presence of 10 mM $MgCl_2$ at 60 °C in a volume of 10 μL. At times indicated in the lower panel (lanes 2–5), 2-μL aliquots containing 5 μCi of radioactivity were chilled on ice and then electrophoretically analyzed in 5% polyacrylamide–8 M urea slab gel. The negative control lanes are lane 6 (GTP was replaced by ATP) and lane 7 (0 mM $MgCl_2$). Lane 1 shows the spliced products from ^{32}P-labeled *td* precursor RNA. The In–Ex II (intron–exon II) RNA intermediate has an expected size of 1.9 kb. The RNA size markers are the same as in Figure 3.

nosine from the 5′ end of the excised intron (Zaug et al., 1983). Although characterization of total RNA extracted from induced cells harboring the cloned *td* gene by primer extension dideoxy sequencing suggested the intron cyclization junction to be missing a GU dinucleotide originally present in the 5′ end of the linear intron molecule, this product was not observed or characterized (Ehrenman et al., 1986). In an effort to identify the cleavage product of intron cyclization, we added [α-^{32}P]GTP to the 5′ end of the linear intron which was derived from the SP6 2.7-kb primary transcript of the *td* gene by incubating at 60 °C for 3 min in the presence of 10 mM $MgCl_2$. The ^{32}P-labeled products were isolated free of unreacted [α-^{32}P]GTP by passage through Nensorb (see Experimental Procedures). Intron cyclization was effected by incubating the 5′ end labeled linear intron at 60 °C for 2–10 min in the presence of 10 mM $MgCl_2$. The final products of incubation were electrophoretically analyzed in a 5% gel (for detection of the radioactive 1-kb linear intron) and in a 20% gel (for detection of any released radioactive short oligonucleotides). As shown in Figure 7, the linear 1-kb intron RNA, which was 5′ end labeled with [α-^{32}P]GTP (shown in lane 1), readily loses its ^{32}P label on incubation at 60 °C in the presence of $MgCl_2$ in the cyclization step (shown in lanes 2 and 3). Analysis of these same products in a 20% gel (lanes 4–7) reveals a time-dependent appearance of a [α-^{32}P]GTP-containing dinucleotide (lanes 5–7). This dinucleotide comigrates with a dinucleotide marker [pd(GC)] at a slightly slower rate than [α-^{32}P]GTP (lane 4) but faster than a tetranucleotide marker [$pd(GT)_2$]. The faster moving ^{32}P-labeled band in lanes 5–7 may be residual [α-^{32}P]GTP not removed by Nensorb or from degradation of the released dinucleotide.

To demonstrate that the second nucleotide in the dinucleotide was a U (the first intron-encoded residue), the SP6 *td* transcript was synthesized in the presence of [α-^{32}P]UTP and then isolated free of unpolymerized [α-^{32}P]UTP. The [α-^{32}P]UTP uniformly labeled *td* transcript was subjected to incubation at 60 °C from 0 to 10 min in the presence of 100 μM GTP and 10 mM $MgCl_2$, and the UTP-labeled products were electrophoretically analyzed in a 20% gel for oligonucleotide products. As shown in Figure 7, a U-containing dinucleotide was generated during intron cyclization (lanes 8–10). The specificity of the G-U nucleotide content of the observed dinucleotide was confirmed by the failure of an [α-^{32}P]ATP uniformly labeled *td* transcript to generate a ^{32}P-labeled dinucleotide under identical incubation conditions (lane 11).

td Exon Ligation Is Primarily Intramolecular. It was of interest to determine the mode of ligation of exon I to exon II RNA. The products of splicing from a mixture of two SP6 *td* primary transcripts of different sizes were analyzed by 5% polyacrylamide–urea gel electrophoresis. One primary transcript was the 2.7-kb species, containing intact exon I (0.82-kb), intron (1-kb), and exon II (0.85-kb) segments, while the other was 1.6 kb in size, containing truncated exon I (0.4 kb), intact intron (1 kb), and truncated exon II (0.2 kb) (see Figure 8, lower panel). Cis or intramolecular splicing from this mixture should produce only 1.7- (exon I–exon II from the 2.7-kb PM) and 0.6-kb (truncated exon I–truncated exon II from the 1.6-kb PM) ligated exon RNAs whereas trans or intermolecular splicing should produce four ligated exon RNAs: the 1.7- and 0.6-kb species as well as a 1.04- (exon I–truncated exon II) and a 1.25-kb (truncated exon I–exon II) species. As shown in the upper panel of Figure 8, only the 1.7- and 0.6-kb ligated exon RNAs were produced from the primary transcript mixture (lane 3). The data demonstrate a predominantly intramolecular mode of exon ligation during splicing of *td* precursor RNA.

Cotranscriptional Splicing Is More Efficient Than Post-transcriptional Splicing. As indicated in Figure 3, the temperature dependence of the *td* precursor RNA splicing reactions, particularly for exon ligation, is inefficient below 50 °C. However, our previous work showed that transcription of the *td* gene in vitro with SP6 RNA polymerase at 40 °C led to the production of not only the 2.7-kb primary transcript but also the spliced products (Chu et al., 1985). To explain this apparent discrepancy, we compared splicing of the *td* precursor

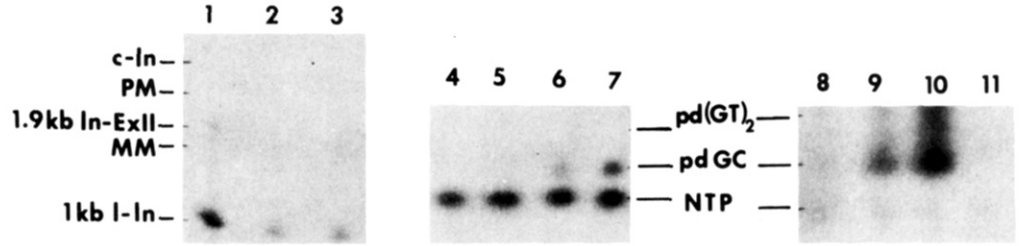


FIGURE 7: Dinucleotide cleavage during *td* intron cyclization. A two-step incubation was carried out as follows. In the first step, addition of [α-^{32}P]GTP (3000 Ci/mmol) to isolated unlabeled *td* 2.7-kb precursor RNA was effected as described in the legend to Figure 6. ^{32}P-Labeled RNA was reisolated by passage through Nensorb and subjected to the second step of incubation in 10 mM $MgCl_2$ at 60 °C from 0 to 10 min. The resulting products were electrophoretically analyzed in 5% gel (lanes 1–3) or 20% gel (lanes 4–11). In a parallel experiment, GTP (100 μM) was incubated with isolated *td* radioactive precursor RNA [uniformly labeled with [α-^{32}P]UTP (lanes 8–10) or [α-^{32}P]ATP (lane 11)] at 60 °C from 0 to 10 min in the presence of 10 mM $MgCl_2$, and the resulting products were similarly analyzed. The gel lanes are 1, 5, and 8, unincubated; 2, 6, and 9, 2 min at 60 °C; 3, 7, 10, and 11, 10 min at 60 °C; and 4, [α-^{32}P]GTP control. The dinucleotide and tetranucleotide standards for the 20% gel were pd(GC) (=dGC) and $pd(GT)_2$ (=dGTGT), respectively. See Figure 6 legend for explanation of abbreviations in lanes 1–3.

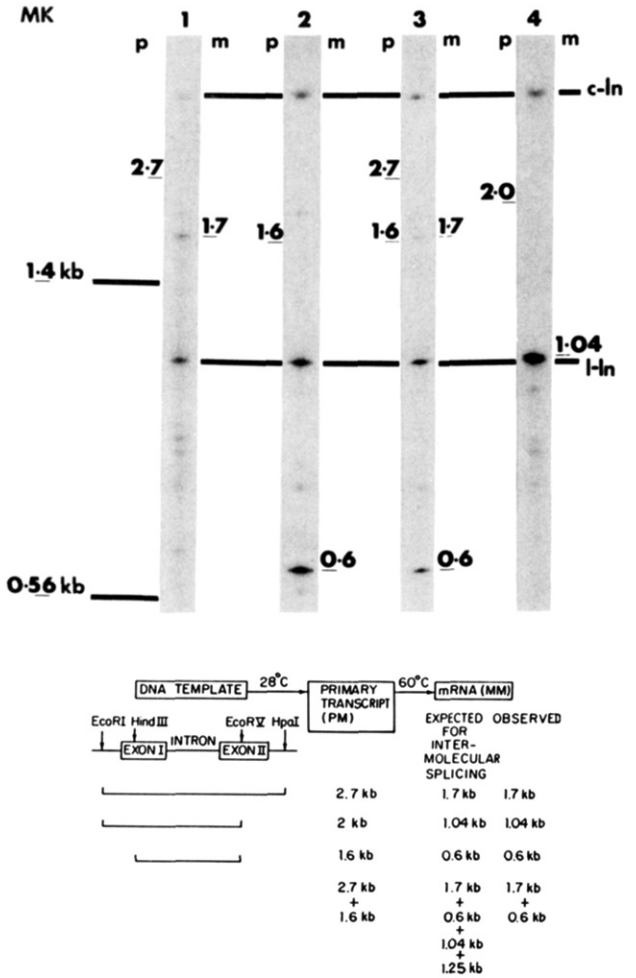


FIGURE 8: Ligation of *td* exons is intramolecular. Splicing was carried out at 60 °C for 10 min with the following ^{32}P-labeled primary transcripts as starting substrates: 2.7-kb PM containing exon I–intron–exon II (lane 1), 1.6 kb containing truncated exon I–intron-truncated exon II (lane 2), 2.7 kb + 1.6 kb (lane 3), and 2 kb containing exon I–intron-truncated exon II (lane 4). These are indicated under "p" to the left of each gel lane. Under "m" to the right of each lane are the observed spliced exon RNAs. The size markers are the same as in Figure 3. The scheme in the lower panel shows the segment of *td* gene giving rise to each primary transcript (PM) and the corresponding expected and observed sizes for ligated exon RNAs (MM).

RNA during transcription at 40 °C and after transcription. The latter condition was simulated by first transcribing the *td* gene at 28 °C, which produced only the 2.7-kb primary transcript (Chu et al., 1986), then by destroying the DNA template by treatment with deoxyribonuclease, and followed by shifting the nontranscribing mixture to higher temperatures. As shown in Figure 9, incubation of the finished 2.7-kb *td* primary transcript (shown in lane 1) at 40 °C resulted in some production of both linear and circular 1-kb intron RNAs but little or no 1.7-kb MM RNA (lane 2). In contrast, when transcription was carried out at 40 °C, some production of the 1.7-kb MM RNA was detected (lane 4). In either case, shifting to 60 °C enhanced the splicing reaction greatly (lanes 3 and 5). The identity of the 1.7-kb ligated exon RNA and the observed enhancement of its formation during transcription were confirmed by RNA dot hybridization analysis with *td* sequence specific synthetic oligodeoxynucleotide probes (Figure 10).

Discussion

We have reexamined the T4 phage thymidylate synthase (*td*) intron not only with respect to its nucleotide sequence,

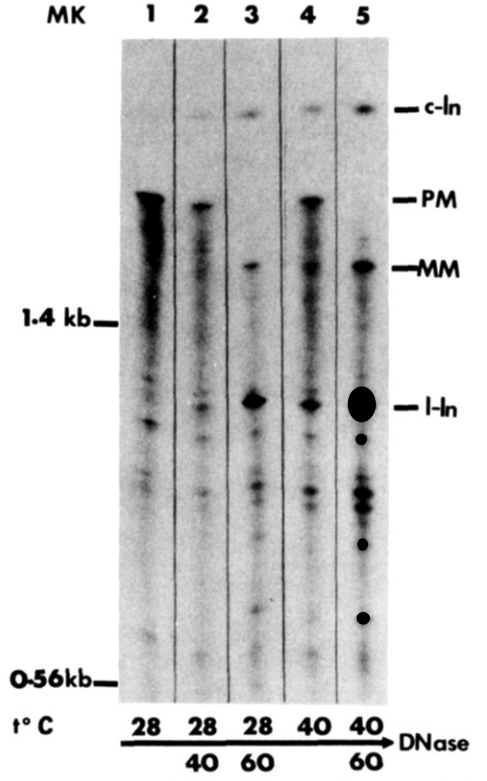


FIGURE 9: Comparison of RNA products of cotranscriptional and posttranscriptional *td* splicing. Transcription of *Hpa*I-linearized pSP64td DNA with SP6 RNA polymerase was carried out for 60 min at either 28 (lanes 1–3) or 40 °C (lanes 4 and 5) in the presence of [α-^{32}P]ATP. At the end of transcription, the DNA template was digested with DNase (RNase free) for 20 min at room temperature (22 °C). The resulting nontranscribing mixtures were incubated for an additional 10 min at 4 (lanes 1 and 4), 40 (lane 2), and 60 °C (lanes 3 and 5) and then electrophoretically analyzed in 5% polyacrylamide–8 M urea slab gel. The abbreviations and size markers are the same as in Figure 3.

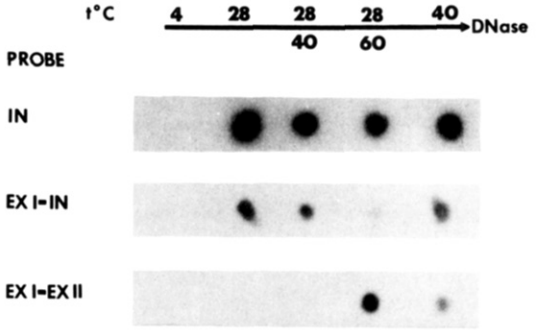


FIGURE 10: Hybridization analysis of RNA products of co- and posttranscriptional splicing. Experimental conditions were essentially identical with those in Figure 9 except that unlabeled RNA was made. Following transcription, RNA was isolated by the Nensorb procedure and subjected to dot hybridization analysis with *td* region specific ^{32}P-labeled synthetic oligodeoxynucleotide probes as described under Experimental Procedures. The transcription and splicing temperatures are indicated at the top of the figure. The hybridization probes used were IN [intron 22-mer containing the sequence 5′-(TGTCTAC-TAGAGAGGTTCCCCG)-3′ 49 bases downstream of the 5′ splice site], ExI–IN [exon I–intron junction 23-mer containing 5′-(GGCCTCAATT AACCCAAGAAAAC)-3′], and ExI–ExII [exon I–exon II spliced junction 24-mer containing 5′-(ATTAAACGGTAG ACCCAAGAAAAC)-3′].

but more importantly, in terms of possible secondary structures relative to those of other eucaryote class I introns. By so doing, we have identified in the *td* intron all the known structural conserved elements (P, Q, R, S) as well as nonconserved ones (IGS, E, E′) which have been implicated as determinants of

the self-splicing property of class I introns (Figure 1). The previously suggested P and Q sequences (Chu et al., 1986) were found to fall within a 631 base pair intron deletion that did not affect *td* RNA splicing either in vivo (Ehrenman et al., 1986) or in vitro (unpublished results). This finding appeared to contradict the suggestion that P and Q sequences base pair to form a secondary structure essential for RNA self-splicing (Waring & Davies, 1984). However, the corrected location of the P and Q sequences in this work reveals that they not only lie outside and upstream of the intron deletion but that they also have the potential to base pair to form a stable secondary structure very similar to the "c" helix in eucaryote class I introns, particularly that postulated for *Chlamydomonas reinhardii* chloroplast ribosomal intron (Rochaix et al., 1985). The presently identified E and E′ sequences would form the equivalent of the "b" helix in *Chlamydomonas*. However, at this juncture, the role of these structural elements in the self-splicing of *td* precursor RNA is purely conjecture. Oligonucleotide-directed mutagenesis of these structural elements in the *td* intron and determination of the effects of these mutations on splicing in vitro should yield valuable information on the importance of these sequences in *td* splicing. Such an approach has been successfully applied in elucidating the role of R and S sequences in the splicing of *T. thermophila* nuclear large rRNA precursor (Burke et al., 1986).

The dramatic effect of temperature on self-splicing of the *td* precursor RNA (Figure 3), particularly for exon ligation, reinforces the idea that the determinants of self-splicing reside only in the RNA nucleotide sequence, similar to that described for the *T. thermophila* nuclear large rRNA (Cech et al., 1983; Waring & Davies, 1984). RNA folding is profoundly influenced by Mg^{2+}, which, in the present case, is believed to stabilize structures essential for both intron excision and exon ligation (Figures 4 and 5A). The effect of temperature on splicing suggests that, below 60 °C, the *td* precursor RNA assumes structures unfavorable for exon ligation. At 60 °C, these structures may be altered to facilitate efficient splicing. Pertinent to this idea is the observation that splicing of the *td* precursor RNA seems to be more efficient during transcription (Figures 9 and 10). Apparently, the nascent or unfinished RNA molecule is better able to effect structures in the intron and in its flanking regions, which are favorable for splicing. Conversely, we can postulate that perhaps the RNA segment downstream of the 3′ splice site contains sequences that may base pair with intron elements, thus preventing the formation of structural determinants of splicing at low temperature.

There is little doubt that the *td* precursor RNA self-splices in vitro and that its intron contains all of the known class I intron elements. It is therefore not surprising to find that in vitro splicing of this RNA has identical requirements (Mg^{2+} and guanosine) (Figure 4) and invokes a mechanism (Figures 6–8) analogous to that of its eucaryote counterparts. In the case of the T4 phage *td* transcript, intron cyclization occurs when the 3′-hydroxyl group of the 3′ terminal guanosine residue attacks the phosphodiester linkage between the first (U) and second (A) residue in the 5′ end of the intron. An interesting finding is that although guanosine is absolutely required for *td* exon ligation, its omission does not prevent intron excision and cyclization in vitro (Figure 5B). This phenomenon has been very recently reported for the *Tetrahymena* rRNA precursor (Inoue et al., 1986), where it was shown that, in the absence of guanosine, the first cleavage event occurs at the 3′ splice site instead of at the 5′ splice site, giving rise to the exon I–intron intermediate which is unable to undergo exon ligation. Interestingly, the guanosine residue at the 3′ end of the intron in this intermediate can attack the adenosine at the 5′ end of the *Tetrahymena* intron, causing a complete intron circle to be formed. This may be the scenario for the guanosine-independent excision of the *td* intron. To confirm that this is the case, determination of nucleotide sequence across the cyclization site in the circular intron RNA formed in the presence and absence of guanosine will be necessary.

It should be pointed out that the covalent addition of guanosine to the 5′ end of the intron RNA segment during the first transesterification reaction (depicted in Figure 1) has been adapted by others to probe for total self-splicing introns in the T4 phage and by us for similar introns in several procaryotic systems (Chu et al., unpublished results).[2] Thus, a 625 base pair intervening sequence has been discovered in the *nrd*B gene encoding the T4 phage ribonucleotide reductase B2 subunit, and the nucleotide sequence bears remarkable resemblance to the *td* gene in the exon I–intron and intron–exon II boundaries (Sjoberg et al., 1986). The self-splicing nature of this intron has been subsequently demonstrated in vitro by Gott et al. (1986).

Does self-splicing occur in vivo? If the reaction occurs in vitro, it can certainly do so in vivo. However, it is likely that in vivo there are protein factors serving as "scaffolds" to stabilize RNA structure for self-splicing. In this respect, we have observed *td* precursor RNA splicing to yield ligated exon RNA during transcription in the absence of *E. coli* proteins only at 40 °C or higher (Chu et al., 1985) and posttranscription at a high temperature of 60 °C (this work). This is in contrast to the synthesis of thymidylate synthase enzyme from the *td* gene in vitro at 25 °C in a coupled transcription translation system derived from an *E. coli* S-30 fraction (unpublished observation) and in intact cells (West et al., 1986). The much lower temperature for the in vitro RNA splicing dependent expression of the *td* gene in the presence of *E. coli* proteins suggests that the *td* precursor RNA splicing reaction may be enhanced by some *E. coli* proteins, which in some way facilitate the correct folding of the RNA for splicing.

In conclusion, we have provided unequivocal evidence that the *td* gene, the first example of an intron-containing, procaryotic protein encoding gene, is transcribed to a precursor RNA whose processing requirements are identical with those of eucaryote class I self-splicing introns. The origin of these self-splicing introns is a question interesting to speculate on but difficult to answer. One possibility is that these introns may have evolved early in evolution in procaryotes but most were disposed of during the course of evolution as a means of increasing the efficiency of expression of specific genes and at the same time streamlining genetic material. A few were retained (e.g., in the T4 phage). The other possibility is that introns occurred late in evolution in eucaryotes from which the T4 bacteriophage or its ancestor acquired its self-splicing intron.

ACKNOWLEDGMENTS

We are indebted to Francois Michel and Bernard Dujon of CNRS, Gif-sur-Yvette, France, and Richard Waring of the University of Medicine and Dentistry of New Jersey for their expert advice on the secondary structure model of the *td* intron. We thank Debbie West for her helpful discussion, Joanne Martinez for her excellent technical assistance, and Judy

[2] Multiple self-splicing intron RNAs were detected in all T-even phage infected cells, including a 0.6-kb species (approximate size of the *nrd*B-derived intron) present only in T4 phage infected cells.

Valentino for her care in typing the manuscript.

REFERENCES

Belfort, M., Pedersen-Lane, J., West, D., Ehrenman, K., Maley, G., Chu, F., & Maley, F. (1985) *Cell (Cambridge, Mass.) 41*, 375–382.

Burke, J. M., Irvine, K. D., Kaneko, K. J., Kerker, B. J., Oettgen, A. B., Tierney, W. M., Williamson, C. L., Zaug, A. J., & Cech, T. R. (1986) *Cell (Cambridge, Mass.) 45*, 167–176.

Cech, T. R. (1983) *Cell (Cambridge, Mass.) 34*, 713–716.

Cech, T. R., Zaug, A. J., & Grabowski, P. J. (1981) *Cell (Cambridge, Mass.) 27*, 487–496.

Cech, T. R., Tanner, N. K., Tinoco, I., Jr., Weir, B. R., Zuker, M., & Perlman, P. S. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 3903–3907.

Chu, F. K., Maley, G. F., Maley, F., & Belfort, M. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 3049–3053.

Chu, F. K., Maley, G. F., Belfort, M., & Maley, F. (1985) *J. Biol. Chem. 260*, 10680–10688.

Chu, F. K., Maley, G. F., West, D. K., Belfort, M., & Maley, F. (1986) *Cell (Cambridge, Mass.) 45*, 157–166.

Clewell, D. B., & Helinski, D. R. (1969) *Proc. Natl. Acad. Sci. U.S.A. 62*, 1159–1166.

Davies, R. W., Waring, R. B., Ray, J. A., Brown, T. A., & Scazzocchio, C. (1982) *Nature (London) 300*, 719–724.

Ehrenman, K., Pedersen-Lane, J., West, D., Herman, R., Maley, F., & Belfort, M. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 5875–5879.

Garriga, G., & Lambowitz, A. M. (1984) *Cell (Cambridge, Mass.) 39*, 631–641.

Gott, J. M., Shub, D. A., & Belfort, M. (1986) *Cell (Cambridge, Mass.) 47*, 81–87.

Greer, C. L., Peebles, C. L., Gegenheimer, P., & Abelson, J. (1983) *Cell (Cambridge, Mass.) 32*, 537–546.

Inoue, T., Sullivan, F. X., & Cech, T. R. (1986) *J. Mol. Biol. 189*, 143–165.

Kramer, A., Keller, W., Apel, B., & Luhrmann, R. (1984) *Cell (Cambridge, Mass.) 38*, 299–307.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E., & Cech, T. R. (1982) *Cell (Cambridge, Mass.) 31*, 147–157.

Lerner, M. R., Boyle, J. A., Mount, S. M., Wolin, S. L., & Steitz, J. (1980) *Nature (London) 283*, 220–224.

Michel, F., & Dujon, B. (1983) *EMBO J. 2*, 33–38.

Michel, F., & Cummings, D. J. (1985) *Curr. Genet. 10*, 69–79.

Michel, F., & Dujon, B. (1986) *Cell (Cambridge, Mass.) 46*, 323.

Michel, F., Jacquier, A., & Dujon, B. (1982) *Biochimie 64*, 867–881.

Mount, S. M., Petterson, I., Hinterberger, M., Karmas, A., & Steitz, J. A. (1983) *Cell (Cambridge, Mass.) 33*, 509–518.

Mileham, A. J., Revel, H. R., & Murray, N. E. (1980) *MGG, Mol. Gen. Genet. 179*, 227–239.

Padgett, R. A., Mount, S. M., Steitz, J. A., & Sharp, P. A. (1983) *Cell (Cambridge, Mass.) 35*, 101–107.

Peebles, C. L., Gegenheimer, P., & Abelson, J. (1983) *Cell (Cambridge, Mass.) 32*, 525–536.

Pikielny, C. W., Teem, J. L, & Rosbash, M. (1983) *Cell (Cambridge, Mass.) 34*, 395–403.

Rochaix, J. D., Rahire, M., & Michel, F. (1985) *Nucleic Acids Res. 13*, 975–984.

Sjoberg, B.-M., Hahne, S., Mathews, C. Z., Mathews, C. K., Rand, K. N., & Gait, M. J. (1986) *EMBO J. 5*, 2031–2036.

Tabak, H. F., Van der Horst, G., Osinga, K. A., & Arnberg, A. C. (1984) *Cell (Cambridge, Mass.) 39*, 623–629.

Waring, R. B., & Davies, R. W. (1984) *Gene 28*, 277–291.

Waring, R. B., Scazzocchio, C., Brown, T. A., & Davies, R. W. (1983) *J. Mol. Biol. 167*, 595–605.

West, D. K., Belfort, M., Maley, G. F., & Maley, F. (1986) *J. Biol. Chem. 261*, 13446–13450.

Zaug, A. J., Grabowski, P. J., & Cech, T. R. (1983) *Nature (London) 301*, 578–583.