Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 3049–3053, May 1984
Biochemistry

Intervening sequence in the thymidylate synthase gene of bacteriophage T4

(prokaryotic intron/sequence of *td* gene/primary structure of thymidylate synthase/terminal overlap with *frd* gene)

FREDERICK K. CHU, GLADYS F. MALEY, FRANK MALEY, AND MARLENE BELFORT

Center for Laboratories and Research, New York State Department of Health, Empire State Plaza, Albany, NY 12201

Communicated by Seymour S. Cohen, February 6, 1984

ABSTRACT **The continuous sequence of 2.3 kilobases in a 3-kilobase DNA fragment encoding the structural gene for coliphage T4 thymidylate synthase (5,10-methylenetetrahydrofolate:dUMP *C*-methyltransferase, EC 2.1.1.45) was determined by using the M13 dideoxy chain-termination method. From the coding information within this gene and that provided by sequence analysis of selected CNBr peptides from the protein product, the primary structure of T4 thymidylate synthase was determined. The most significant finding of these studies is the presence of a 1017-base-pair interruption two-thirds of the way through the nucleotide sequence of the structural gene. The 5′- and 3′-terminal ends of this intron are demarcated by an apparent stop and start codon, respectively. The corresponding methionine preceding the second coding region of the synthase is not incorporated into the final protein product. Structural evidence confirming the presence of the intervening sequence in the phage genome was obtained by restriction and hybridization analysis. Support for the presence of the intron was also obtained at the functional level by enzyme expression studies using selected *td* gene fragments. This work also confirms the findings of Purohit and Mathews [Purohit, S. & Mathews, C. K. (1983) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* 42, 1759], which reveal that the termination codon for the dihydrofolate reductase gene and the triplet initiating thymidylate synthase overlap by a four-base stretch, A-T-G-A. The implications of this unusual gene arrangement are discussed.**

Although intervening sequences are commonly found in the genes of eukaryotes (1), no interruption in a protein coding region of a prokaryotic gene has been reported to date. Recently, however, it was shown by Kaine *et al.* (2) that two tRNAs from archaebacteria, which possess features of both eukaryotes and prokaryotes, contain small introns. It is therefore of particular interest that we have uncovered an intervening sequence in the coliphage T4 *td* gene encoding thymidylate synthase (5,10-methylenetetrahydrofolate: dUMP *C*-methyltransferase, EC 2.1.1.45).

Although this enzyme is also expressed by the *thy*A gene of *Escherichia coli*, the phage and host synthases exhibit striking dissimilarities in both structural and functional properties (3–5). In an effort to better understand the reasons for these differences as well as those underlying the differential regulation of the *td* and *thy*A genes, we undertook to determine the sequences of both genes and their protein products. The sequence of the *thy*A gene and its thymidylate synthase product was described recently (6). We now present the complete coding sequence of the *td* gene and the primary structure of its protein product. In the course of these studies it became evident that unlike the *thy*A structural gene, which exists as a continuous open reading frame encoding 264 amino acids, the *td* gene is interrupted by a 1-kilobase-pair (kb) polynucleotide segment. Thus, in contrast to what has been found for all known prokaryotic genes, the T4 phage thymidylate synthase is divided into two separate open reading frames by an intron. Evidence in support of this significant finding forms the basis of this paper.

MATERIALS AND METHODS

Bacterial Strains, Phage Strains, and Plasmids. *E. coli* strain JM103, obtained from Bethesda Research Laboratories, was used for the propagation of M13 phage (7). Strain AB2497, from B. Bachmann of the *E. coli* Genetic Stock Center, was used as the *thy*A$^-$ host for complementation experiments. The double-stranded replicative forms of M13 strains mp8 and mp9 (7) were from P-L Biochemicals. The 3-kb *Eco*RI fragment† encoding the T4 *td* gene was originally cloned from strain T4*alc*4, in which about 50% of the glucosylated hydroxymethylcytosine residues are replaced by cytosine (9). It was ligated into a phage λ vector to yield λTd (8) and kindly provided to us by N. Murray. The 3-kb *td* fragment was later subcloned into the *Eco*RI site of pBR322, pUC8 or pUC9 (10), and M13 vectors to generate pBTd (4), pUCTd, and M13Td, respectively. The 3-kb fragment was also cloned into expression plasmid pKC30 (11) by a blunt-end cloning event that destroyed the *Eco*RI sites flanking the insert to yield recombinant pKTd (4). T4 DNA for restriction hybridization experiments was extracted from strain T4*alc*7 (provided by H. Revel), in which cytosine residues are completely unmodified and the DNA is fully restriction enzyme sensitive (9).

Construction of Clones for DNA Sequence Analysis. Cloning procedures (plasmid DNA preparation, restriction, ligation, linker addition, and transformation) have been described (4–6). Two libraries of DNA fragments were generated for cloning into M13 vectors by (*i*) *Pvu* II restriction of the purified 3-kb *td* fragment and (*ii*) *Hin*dIII restriction of pBTd followed by BAL-31 digestion of the linearized DNA (6) (Fig. 1*B*). Oligonucleotide linkers (*Bam*HI, *Hin*dIII, *Eco*RI) used in cloning of blunt ends were from New England Biolabs.

DNA Sequence Analysis. The nucleotide sequence was determined by the dideoxy chain-termination method of Sanger *et al.* (12, 13) with single-stranded DNA from M13-mp8 or -mp9 clones of either *Pvu* II restriction fragments or overlapping BAL-31-generated deletions (Fig. 1*B*). [α-^{32}P]dATP (400 Ci/mmol, 10 mCi/ml; 1 Ci = 37 GBq) was purchased from Amersham. Dideoxynucleoside triphosphates and the heptadecanucleotide universal primer were from P-L Biochemicals. Nucleoside triphosphates were obtained from Boehringer Mannheim.

Amino Acid Sequence Analysis. T4 coliphage thymidylate synthase produced by pKTd was purified to homogeneity as

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: kb, kilobase pair(s); FdUMP, 5-fluorodeoxyuridylate.

†This fragment was assessed previously as 2.7 kb in length (4, 8).

described (4). The ternary complex of purified synthase containing 5-fluorodeoxy[2-^{14}C]uridylate (FdUMP) was carboxymethylated and then used for NH_2-terminal sequence analysis and conversion to peptides by reaction with CNBr (14). The peptides were separated by Bio-Gel P-60 (Bio-Rad) and HPLC and their partial amino acid sequences were determined as described (6), thus locating the coding regions for five of the six CNBr peptides within the structural gene for *td*. The sixth peptide was identified on the COOH-terminal end of the derivatized enzyme by progressive digestion with carboxypeptidase Y (6). CNBr was obtained from Pierce and chemicals for sequence analysis were from Pierce or Beckman. Fd[2-^{14}C]UMP (52 mCi/mmol) was purchased from Moravek Biochemicals (Brea, CA). Sequence analysis was performed with a Beckman 890B Sequencer.

Enzymes. The Klenow fragment of *E. coli* DNA polymerase was from Bethesda Research Laboratories and T4 DNA ligase from New England Biolabs. BAL-31 nuclease and restriction enzymes *Eco*RI, *Pst* I, *Pvu* II, *Bam*HI, *Hin*cII, *Hin*dIII, and *Hpa* I were from either of these suppliers. Carboxypeptidase Y was kindly supplied by T. Plummer, Jr.

RESULTS

Sequence Analysis of the T4 *td* Gene. The strategy for sequence determination of restriction fragments and BAL-31-generated deletion fragments of the *td* gene is depicted in Fig. 1*B*. The extent and direction of sequence determination using the dideoxy chain-termination method (12, 13) for each fragment cloned into M13 vectors mp8 or mp9 is indicated by the arrows in Fig. 1*B*. From the sequence analysis of B1, B2, and P3, it was apparent that there might be an interruption in the *td* gene, not only because of the UAA stop signal at nucleotide 769 (Fig. 2) but also because nucleotides coding for amino acids beyond residue 183 could not be found. However, this problem was resolved with BAL-31 fragments B3–B8, which revealed that the coding sequence for the COOH-terminal end of the synthase appeared after an interruption of 1017 nucleotides. As discussed below, the resumption of the synthase sequence was confirmed by sequence analysis of peptides from this region. These data make it apparent that the synthase coding region contains a 1-kb interruption two-thirds of the way from the 5′ end of the *td* structural gene. The intron sequence, determined from one strand, will be published elsewhere pending sequence verification from the complementary strand.

The Extended Functional Length of the *td* Gene. To correlate the structural organization of the gene with its function, expression of the 3-kb *Eco*RI fragment as well as that of relevant subfragments was monitored after transcriptional fusion to the p_{Lac} promoter in plasmid pUC8 or pUC9 to yield pUCTd (Fig. 1*C*). By introducing pUCTd into *thy*A$^-$ *E. coli* we were able to demonstrate that the intact 3-kb fragment (Fig. 1, C1) allowed complementation of this synthase-deficient strain to Thy$^+$. As expected, Thy$^+$ transformants were obtained only when *td* was in transcriptional alignment with p_{Lac} and not when the orientation was reversed (pUCdT), since the *td* fragment does not have an internal promoter (4). Extracts of cells with the Thy$^+$ phenotype were shown by gel electrophoresis to form a ternary complex when treated with Fd[5-^{3}H]UMP and 5,10-methylenetetrahydrofolate, which corresponded in size to the ternary complex of the T4 enzyme (4). This complex was present only in Thy$^+$ pUCTd transformants and was not formed with extracts from Thy$^-$ pUCdT cultures (data not shown). Complementation therefore provides a useful test for T4 synthase function.

Recombinants containing the *Hin*dIII–*Eco*RI subfragment depicted in Fig. 1, C2 (pUCTd-HE), were predictably Thy$^-$ since the *Hin*dIII site is within the NH_2-terminal coding sequence of the synthase [nucleotide 376, corresponding to Lys-53 (Figs. 1*A* and 2)]. Further, pUCTd-ET recombinants, containing the *Eco*RI–*Taq* I subfragment shown in Fig. 1, C3, are also Thy$^-$, regardless of insert orientation. This finding strongly suggests that the length of the *td* structural gene is greater than the coding requirements for the synthase product. Based on data from the *td* gene sequence analysis and studies with the purified protein, including amino acid composition and sequence, we have concluded that the synthase subunit comprises 286 amino acid residues. The coding sequence of *td*, if present as a continuous open reading frame starting at nucleotide 220 (Fig. 2), should extend no further

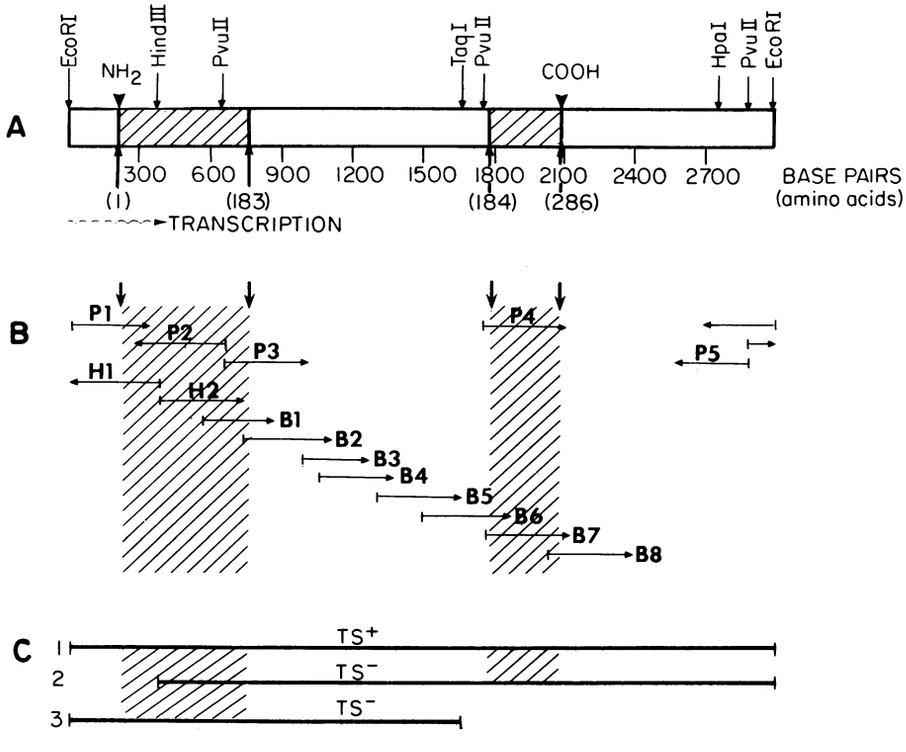

FIG. 1. (*A*) Organization of the thymidylate synthase gene of T4. Pertinent restriction enzyme sites within the 3-kb *td* fragment are shown. The numbers without parentheses indicate the scale in base pairs. The bracketed numbers represent the corresponding amino acid positions in the protein. The DNA sequence of the shaded parts of the *td* fragment containing the entire T4-thymidylate synthase coding region is presented in Fig. 2. (*B*) Sequence strategy. DNA sequence was determined by using M13mp8 or M13mp9 as a vector to generate clones derived from *Pvu* II restriction of the 3-kb *Eco*RI fragment (P series) or from *Hin*dIII restriction of pBTd (H1 and H2) followed by progressive BAL-31 digestion (B series) (6). The arrows show the direction of sequence analysis and the length of the sequence determined by the dideoxy chain-termination method of Sanger *et al.* (12, 13). (*C*) The extended functional length of the *td* gene. Fragment 1, 2, or 3 was cloned into vector pUC8 or pUC9 (10) in transcriptional alignment with the p_{Lac} promoter. Although clones containing fragment 1 complemented *thy*A$^-$ *E. coli* and produced the phage enzyme as monitored on FdUMP gels (TS$^+$) (4), fragments 2 and 3 did not encode a functional T4 thymidylate synthase (TS$^-$).

than nucleotide 1078 (Fig. 1*A*). The finding that the pUCTd-ET clones are Thy$^-$ despite the fact that their DNA encompasses the entire NH_2-terminal portion of the gene and extends all the way to residue number 1672 (Fig. 1, C3) strongly supports a physical interruption in the coding sequence of the *td* gene.

Relationship of the T4 *td* Gene to Its Protein Product. Confirmation of the amino acid sequence deduced from the nucleotide data was greatly facilitated by information derived from the partial sequence analysis of several peptides of the T4 synthase. Of particular relevance was the peptide resulting from the CNBr cleavage of Met-158, which was instru-

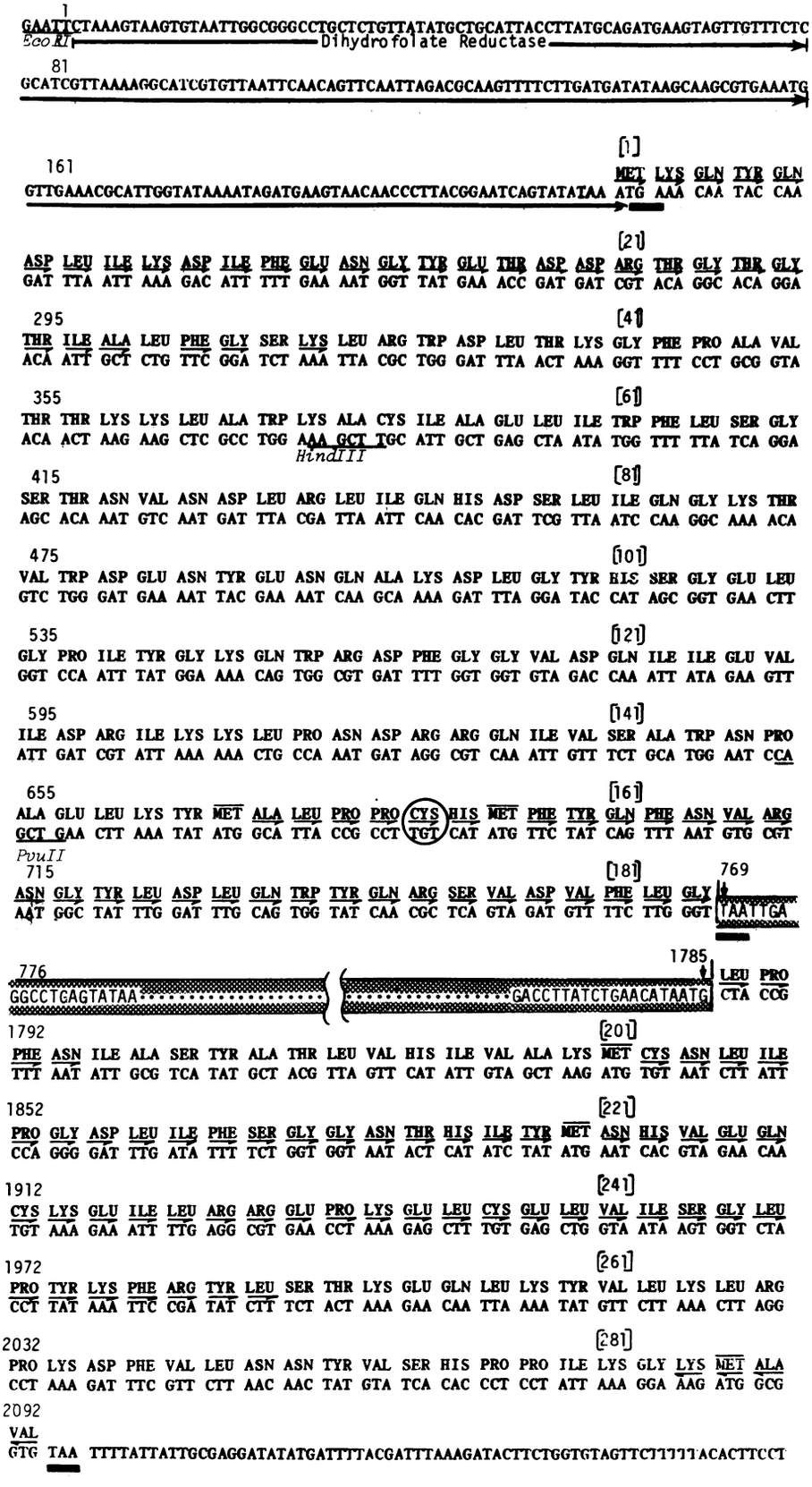

FIG. 2. DNA sequence of the T4 thymidylate synthase gene and its deduced amino acid sequence. The left-most nucleotide on each line is numbered above the sequence beginning at the 5′ end of the antisense strand. The 5′-terminal *Eco*RI site and the internal *Hin*dIII and *Pvu* II sites in the *td* structural gene are labeled under the corresponding sequences. The three stop codons are underscored. The triplet TGA (nucleotides 221–223) spanning the second to fourth nucleotides in the *td* gene is the termination codon for the T4 dihydrofolate reductase gene (15). The sequence from the *Eco*RI site to the AAA (lysine) preceding the termination codon encodes the COOH terminus of T4 dihydrofolate reductase. The amino acid sequence of the synthase is numbered in parentheses above specific amino acids beginning at the NH_2-terminal methionine residue. Those residues confirmed by the analysis of amino acid sequence are underscored by arrows in the direction of the determination. The five internal methionines of the enzyme are overscored with bars, whereas the FdUMP-binding site at Cys-156 is circled. The intron stretching from nucleotides 769 to 1785, a total of 1017 base pairs, is boxed in. Twenty nucleotides corresponding to the 5′ and 3′ ends of the intron are presented.

mental in establishing that the interruption in the open reading frame for the synthase occurred within the nucleotide domain coding for this peptide. Thus, following the stop signal at nucleotide 769 (TAA), which is preceded by the codon for Gly-183 (GGT), it was not possible to locate additional codons for this peptide until after the ATG codon at nucleotide 1783 (Fig. 2). This ATG is followed by the nucleotides coding for the Leu-Pro-Phe-Asn residues of this peptide, representing the beginning of the second open reading frame for the *td* gene. It is of interest to note that the methionine preceding the second coding segment is not included in the union between Gly-183 and Leu-184. Carboxypeptidase Y analysis of the COOH end of the enzyme and the amino acid sequence of two additional CNBr peptides indicated the absence of additional interruptions in the DNA sequence of the *td* gene (Fig. 2).

Previous data on the amino acids at the active sites of other thymidylate synthases, including those from *Lactobacillus casei*, *E. coli*, and *Saccharomyces cerevisiae* (16), all of which contain a Pro-Cys-His sequence, would appear by analogy to implicate Cys-156 as the FdUMP-binding nucleophile in this case (circled in Fig. 2). This was verified by locating Fd[2-^{14}C]UMP in the CNBr2 peptide (residues 152–157) by amino acid sequence analysis.

The sequence of 20 amino acids on the NH_2 terminus of the T4 synthase, beginning with the methionine start codon (ATG) at nucleotide 220, had been determined previously by us (4), and, as indicated in Fig. 2, it is in agreement with that deduced from the DNA sequence of the *td* gene. This information recently enabled Purohit and Mathews (15) to infer that the dihydrofolate reductase structural gene terminates within the same four bases (A-T-G-A, 220–223) associated with the initiation codon for the *td* gene. Thus, although TGA specifies a stop signal for dihydrofolate reductase, the ATG portion of this 4-nucleotide region is responsible for initiating translation of thymidylate synthase.

Verification of the Organization of the T4 *td* Region by Restriction Analysis. Because of the unusual gene arrangement within the *td* fragment (Figs. 1 and 2), it was important to show that the same configuration exists in the intact T4 genome. Thus, restriction enzyme-sensitive T4*alc*7 DNA was compared with various recombinant DNAs containing the intact *td* gene in hybridization experiments using the 3-kb fragment originally derived from T4*alc*4 (8, 9) as a probe (Fig. 3).

It was assuring to note that the *Eco*RI-restricted DNA fragments (Fig. 3*A*, left panel) derived from T4*alc*7 (lane 1), λTd (lane 2), pBTd (lane 3), M13Td (lane 4), pUCTd (lane 5), and the purified *td* fragment itself (lane 6) all contained, in addition to other bands characteristic of each DNA, a 3-kb species (band a). More definitively, the 3-kb band hybridized to the ^{32}P-labeled 3-kb *td* fragment in all cases (Fig. 3*A*, right panel, lanes 1–6). Since the *Eco*RI sites flanking the 3-kb fragment have been destroyed in constructing pKTd (4), *Eco*RI simply linearizes this plasmid by cleavage in the vector portion, accounting for the band at 9.4 kb, which also hybridized to the probe (lane 7). To further probe the internal organization of the *td* fragment, a similar analysis was conducted after mixed digestion with *Eco*RI and *Pvu* II (Fig. 3*B*). Here two fragments, b and c, comigrating at about 1.2 kb hybridized to the probe throughout lanes 1–7. These correspond to the two internal *Pvu* II fragments (see Fig. 1). Moreover, while the one terminal *Pvu* II–*Eco*RI fragment (<0.2 kb) migrated off the gel, the other appeared as a 0.65-kb band (d), which is identical in all lanes except lane 7. Again the exception is attributable to the destruction of the relevant *Eco*RI sites in pKTd.

The striking similarity of the hybridization pattern between the T4 digests and that of the purified or cloned 3-kb *td* fragment indicates that the DNA used in our cloning, amplification, and sequencing experiments is a direct derivative of the T4 genome rather than a product of an artefactual genetic rearrangement. Furthermore, this genetic configuration is the property of several T4 strains, as demonstrated in Fig. 3 by the similarity of T4*alc*7- and T4*alc*4-derived *td* fragments and as suggested by hybridization of an intron-specific fragment to wild-type T4 DNA (unpublished dot blot experiments). Moreover, enhanced thymidylate synthase levels are observed after wild-type T4 and T4*alc*7 infection as well as after transformation by pKTd, the plasmid containing the *td* fragment from T4*alc*4. These results corroborate previous findings that the purified product of the cloned gene is indistinguishable from the T4 phage-specified enzyme by immunologic, kinetic, and structural criteria. The latter include size, amino acid composition, NH_2-terminal amino acid sequence, and peptide maps (4, 18).

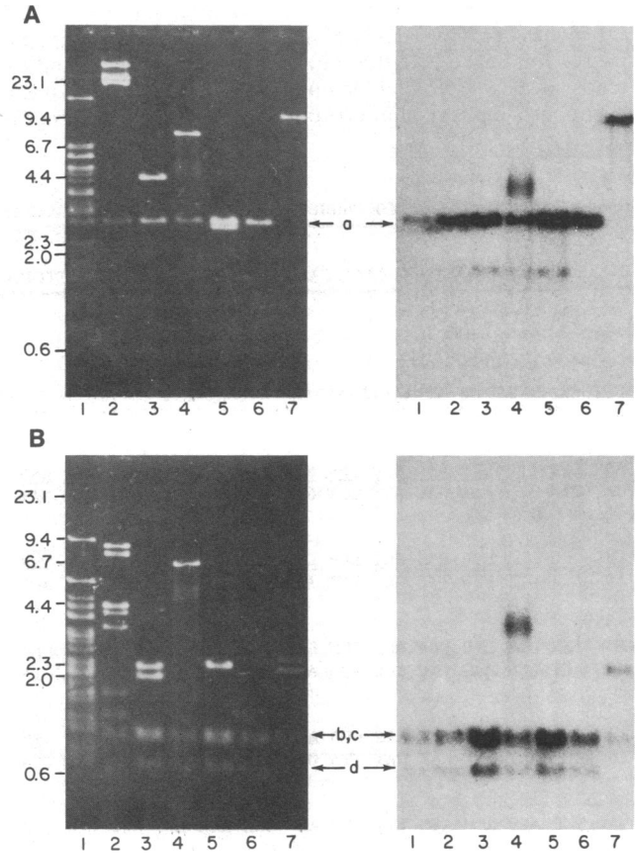

FIG. 3. Organization of DNA fragments in the *td* region of the T4 genome. By using methods described by Thomas (17), hybridization analysis with the nick-translated 3-kb *Eco*RI fragment as probe was performed on *Eco*RI-restricted (*A*) or *Eco*RI–*Pvu* II double-digested DNA (*B*). Lanes 1–7 contain the following DNAs: lanes 1, T4*alc*7; lanes 2, λTd; lanes 3, pBTd; lanes 4, M13Td; lanes 5, pUCTd; lanes 6, purified *Eco*RI *td* fragment; lanes 7, pKTd. Bands of interest are the intact 3-kb *Eco*RI fragment (a), the *Pvu* II doublet of 1.2 kb (b and c), and the 0.65-kb *Eco*RI–*Pvu* II band (d). The hybridizing band above the 3-kb insert in lanes 4 (*A* and *B*) corresponds to single-stranded DNA contaminating the M13 DNA preparation, whereas the faint hybridization below the 3-kb band (*A*, lanes 2–5) probably reflects *Eco*RI activity with reduced specificity. Sizes of molecular weight markers are shown on the left.

DISCUSSION

Although the unique sequence of the coliphage T4 *td* gene described in this report is the only known example of an intervening sequence within a gene coding for a prokaryotic protein, its presence has not been entirely unanticipated. As

indicated in a recent review (19), the T4 phage-specified enzymes RNA ligase and polynucleotide kinase have been waiting for a role to be ascribed to them. Participation in messenger splicing, following intron excision, might be the logical function for these enzymes. However, our results indicate that processing events necessary to produce mature thymidylate synthase can occur in the absence of T4 phage infection since active enzyme is produced from the cloned gene. To account for the production of active enzyme, the possible involvement of host processing enzymes such as endoribonucleases (20) and RNA ligase (21) or intron encoded functions analogous to the yeast maturases (22) must therefore be considered. An autocatalytic RNA splicing reaction similar to that described by Cech *et al.* (23) is yet another possibility.

Regardless of the mechanism involved, RNA blotting experiments, in which regional probes of the *td* fragment were hybridized to RNA extracted from T4 as well as T4*alc*7 infected cells, suggest that *td* processing indeed occurs at the RNA level (to be published elsewhere). Interestingly, the nucleotides T at the 3′ end of the first exon and G at the 3′ end of the intron (Fig. 2) correspond to the only two invariant nucleotides at the splice junctions of those eukaryotic genes whose introns appear to play an active role in their own excision (24). The relevance of this finding as well as of the UAA stop codon that signals the end of the first coding region remains to be determined. The termination codon may in fact facilitate RNA processing by freeing the message of ribosomes.

Alternatively, it might have been suggested in view of the stop and start signals at the ends of the 1-kb intron, that the transcription products of the *td* gene are translated into two separate peptide subunits, which are subsequently covalently joined to produce mature thymidylate synthase. This situation is in contrast to that encountered in eukaryotes, where start and stop signals are not evident at the 5′ and 3′ termini of the intron, which in effect signifies that the fusion of exons in these two systems probably differs. However, preliminary experiments, including pulse–chase protein labeling studies measuring precursor–product relationships of proteins encoded by the 3-kb *td* fragment, argue against such a peptide fusion model (unpublished results). It is, nevertheless, of interest to note that the methionine start codon at the NH_2 terminus of the second segment of the synthase is not included in the final product (Fig. 2). Although this triplet is probably excised as part of the synthase RNA maturation process, a role for this codon in independent translational initiation of the COOH-terminal end of the synthase from the primary transcript remains a possibility.

Another region of interest associated with the *td* gene is that coding for the NH_2-terminal end of the synthase, which has been shown recently to overlap with the COOH terminus of the *frd* gene (16). Thus, it has been found that the two genes have the four nucleotides A-T-G-A in common, providing both a termination signal (TGA) for *frd* and an initiation codon (ATG) for *td* when these genes are read out of phase by one nucleotide. One possibility that is suggested by this overlap is an evolutionary link to certain protozoa that contain dihydrofolate reductase and thymidylate synthase as a bifunctional fusion product, with a molecular weight consistent with these proteins being joined at their respective COOH- and NH_2-terminal ends (25, 26). On the other hand, this four-nucleotide overlap may possess a regulatory function, which may in fact account for our earlier observation that despite the cotranscription and proximity of the *frd* and *td* genes, thymidylate synthase production is delayed *in vivo* and dramatically retarded *in vitro* relative to dihydrofolate reductase (27). Aside from this explanation, the processing events required to allow synthesis of catalytically active thymidylate synthase from the split *td* gene no doubt account for some temporal delay between *frd* and *td* expression.

Perhaps a clue to the unusual structure of the *td* gene resides in the fact that its product plays a dual role. Thus, in addition to its catalytic role the synthase molecule forms a structural component of the phage base plate (28, 29). In this regard, the splicing event may provide a means for differentially regulating synthesis of the catalytically active and structural components of thymidylate synthase.

We thank Don Guarino, Judith Reidl, Joan Pedersen-Lane, and Karen Ehrenman for their capable technical assistance and Judy Valentino for her care in preparing this manuscript. Helpful discussions with Drs. Hanna Engelberg-Kulka and Robert Trimble are gratefully acknowledged. We thank Drs. B. Bachmann, N. Murray, and H. Revel for phage and bacterial strains. This work was supported in part by grants from the National Science Foundation (PCM-8118368, PCM-8302005) and the National Institutes of Health (GM26387).

1. Breathnach, R. & Chambon, P. (1981) *Annu. Rev. Biochem.* **50,** 349–383.
2. Kaine, B. P., Gupta, R. & Woese, C. R. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 3309–3312.
3. Maley, G. F., Maley, F. & Baugh, C. M. (1979) *J. Biol. Chem.* **254,** 7485–7487.
4. Belfort, M., Moelleken, A., Maley, G. F. & Maley, F. (1983) *J. Biol. Chem.* **258,** 2045–2051.
5. Belfort, M., Maley, G. F. & Maley, F. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 1858–1861.
6. Belfort, M., Maley, G. F., Pedersen-Lane, J. & Maley, F. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 4914–4918.
7. Messing, J. & Vieira, J. (1982) *Gene* **19,** 269–276.
8. Mileham, A. J., Revel, H. R. & Murray, N. E. (1980) *Mol. Gen. Genet.* **179,** 227–239.
9. Wilson, G. G., Tanyashin, V. I. & Murray, N. E. (1977) *Mol. Gen. Genet.* **156,** 203–214.
10. Vieira, J. & Messing, J. (1982) *Gene* **19,** 259–268.
11. Shimatake, H. & Rosenberg, M. (1981) *Nature (London)* **292,** 128–132.
12. Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463–5467.
13. Sanger, F., Coulson, A. R., Barrell, B. G., Smith, A. J. H. & Roe, B. A. (1980) *J. Mol. Biol.* **143,** 161–178.
14. Maley, G. F., Bellisario, R. L., Guarino, D. U. & Maley, F. (1979) *J. Biol. Chem.* **254,** 1288–1295.
15. Purohit, S. & Mathews, C. K. (1983) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* **42,** 1759.
16. Maley, F., Belfort, M. & Maley, G. (1984) *Adv. Enzyme Regul.* **22,** in press.
17. Thomas, P. (1983) *Methods Enzymol.* **100,** 255–266.
18. Maley, F. & Maley, G. F. (1981) in *Molecular Actions and Targets for Cancer Chemotherapeutic Agents*, eds., Sartorelli, A. C., Bertino, J. R. & Lazlo, J. J. (Academic, New York), Vol. 2, pp. 265–283.
19. Snyder, L. (1983) in *Bacteriophage T4*, eds., Mathews, C. K., Kutter, E. M., Mosig, G. & Berget, P. B. (American Society for Microbiology, Washington, DC), pp. 351–355.
20. Gegenheimer, P. & Apirion, D. (1981) *Microbiol. Rev.* **45,** 502–541.
21. Greer, C. L., Javor, B. & Abelson, J. (1983) *Cell* **33,** 899–906.
22. Lazowska, J., Jacq, C. & Slonimski, P. P. (1980) *Cell* **22,** 333–348.
23. Cech, T. R., Zaug, A. J. & Grabowski, P. J. (1981) *Cell* **27,** 487–496.
24. Cech, T. R. (1983) *Cell* **34,** 713–716.
25. Ferone, R. & Roland, S. (1980) *Proc. Natl. Acad. Sci. USA* **77,** 5802–5806.
26. Coderre, J. A., Beverley, S. M., Schimke, R. T. & Santi, D. V. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 2132–2136.
27. Trimble, R. B., Galivan, J. & Maley, F. (1972) *Proc. Natl. Acad. Sci. USA* **69,** 1659–1663.
28. Capco, G. R. & Mathews, C. K. (1973) *Arch. Biochem. Biophys.* **158,** 736–743.
29. Kozloff, L. M., Lute, M. & Crosby, L. K. (1977) *J. Virol.* **23,** 637–644.