

# Formation of a DNA Loop at the Replication Fork Generated by Bacteriophage T7 Replication Proteins\*

(Received for publication, October 1, 1997, and in revised form, December 4, 1997)

Kyusung Park<sup>‡§</sup>, Zeger Debyser<sup>||</sup>, Stanley Tabor<sup>¶\*\*</sup>, Charles C. Richardson<sup>¶\*\*</sup>,  
and Jack D. Griffith<sup>‡ ‡‡</sup>

From the <sup>‡</sup>Lineberger Comprehensive Cancer Center, University of North Carolina, Chapel Hill, North Carolina 27599-7295 and the <sup>¶</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, Massachusetts 02115

**Intermediates in the replication of circular and linear M13 double-stranded DNA by bacteriophage T7 proteins have been examined by electron microscopy. Synthesis generated double-stranded DNA molecules containing a single replication fork with a linear duplex tail. A complex presumably consisting of T7 DNA polymerase and gene 4 helicase/primase molecules was present at the fork together with a variable amount of single-stranded DNA sequestered by gene 2.5 single-stranded DNA binding protein. Analysis of the length distribution of Okazaki fragments formed at different helicase/primase concentrations was consistent with coupling of leading and lagging strand replication. Fifteen to forty percent of the templates engaged in replication have a DNA loop at the replication fork. The loops are fully double-stranded with an average length of ~1 kilobase. Labeling with biotinylated dCTP showed that the loops consist of newly synthesized DNA, and synchronization experiments using a linear template with a G-less cassette demonstrated that the loops are formed by active displacement of the lagging strand. A long standing feature of models for coupled leading/lagging strand replication has been the presence of a DNA loop at the replication fork. This study provides the first direct demonstration of such loops.**

Chromosomal replication in most complex organisms requires the coordination of events at the replication fork such that progress on the leading strand matches that of the lagging strand. One mechanism for accomplishing this coordination was proposed over twenty years ago by Alberts and colleagues (1–4) for bacteriophage T4 replication and is known as the “trombone” model. In this model the lagging strand loops back at the fork and engages a second molecule of DNA polymerase. This loop increases in size as the synthesis of each Okazaki fragment progresses, until synthesis of the Okazaki fragment abuts the 5' end of the previous fragment at which time the

loop is released and a new cycle of lagging strand synthesis begins. This model was based on mechanistic considerations and data suggesting a coupling of leading and lagging strand synthesis (reviewed in Ref. 5). It was further assumed that the proteins involved in lagging strand synthesis are recycled from the end of a completed Okazaki fragment to the next primer and that primer synthesis occurs exclusively at the fork.

The original trombone model suggested that the lagging strand polymerase remains at the fork through a direct linkage to the leading strand polymerase. Looping then allows both polymerases to synthesize in the same direction and remain in contact with the primase-helicase. Thus, both strands are synthesized at the same rate by the dimeric polymerase. In more recent variants of the trombone model, the lagging strand polymerase is allowed to replicate at a faster rate until it reaches the end of the previous fragment and then pauses while the leading strand polymerase covers the same distance (6). A diagram illustrating the looping of the lagging strand is presented in Fig. 1A. Kinetic data and the effects of reaction conditions on the size of Okazaki fragments has provided evidence for a coordination of leading and lagging strand DNA synthesis in the *Escherichia coli* (7–9) and T4 DNA replication systems (10). The presence of two DNA polymerase III cores within one holoenzyme particle provides a basis for coupling leading and lagging strand DNA synthesis in the *E. coli* system (11–13). Indeed, both DNA polymerases within the complex can function simultaneously (13).

The coupling of leading and lagging strand synthesis has been demonstrated by gel electrophoretic analyses of the amount and average length distribution of lagging strand fragments upon dilution of the polymerase or the helicase-primase. If lagging strand synthesis is carried out by a polymerase-primase complex that is not associated with the leading strand complex then dilution of either protein should result in an increase in the average length of lagging strand fragments since a greater amount of leading strand synthesis will occur between priming events. If, on the other hand, the synthesis of the two strands are coordinated by a single complex, then reducing the free polymerase or primase concentration will not affect the average length of each lagging strand fragment; each new fragment will be initiated by the same primase-polymerase complex used to synthesize the previous fragment. Thus the average length distribution of Okazaki fragments as a function of decreasing concentration of helicase-primase is often used as a measure of whether leading and lagging strand synthesis is coupled. Studies using purified bacteriophage T4 proteins (5, 10, 14), T7 proteins (15), and *E. coli* proteins (7) in which the size of Okazaki fragments was measured after dilution of replicating complexes, have generally been supportive of the model of a single replisome coordinating the synthesis of

\* Supported in part by Grant GM 31819 from the United States Public Health Service. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Laboratory of Biochemistry, NCI, National Institutes of Health, Bethesda, MD 20892.

|| Supported by Grant NP-1W from the American Cancer Society and Grant AI-06045 from the United States Public Health Service. Present address: REGA Institute for Medical Research, Minderbroedersstraat 10, B3000 Leuven, Belgium.

\*\* Supported by Grant NP-1W from the American Cancer Society and Grant AI-06045 from the United States Public Health Service.

‡‡ To whom correspondence should be addressed. Tel.: 919-966-2151; Fax: 919-966-3015.

both strands.

The T7 system offers a simple powerful tool for studies of DNA-protein interactions at the replication fork. The fundamental reactions occurring at a T7 DNA replication fork can be reconstituted using four proteins, three encoded by T7, the products of genes 2.5, 4, and 5, and one protein, thioredoxin, which is encoded by *E. coli* (16). T7 gene 5 protein is a DNA polymerase that becomes highly processive when complexed with thioredoxin (17–19). Gene 4 encodes two polypeptides of 56 and 63 kDa, in which the 56-kDa polypeptide is translated from an internal initiation codon that is in frame with the coding sequence of the 63-kDa gene 4 protein (20). Both forms of gene 4 protein have helicase activity. They bind to single-stranded (ss)<sup>1</sup> DNA and translocate 5' to 3' along the DNA using the energy of dTTP hydrolysis (21–23). Upon encountering double-stranded (ds) DNA, they unwind the DNA processively (24). The 63-residue amino-terminal domain unique to the 63-kDa gene 4 protein contains a zinc motif that enables this form of gene 4 protein to function as a primase (25–27). It synthesizes tetranucleotides at specific recognition sequences, and T7 DNA polymerase uses these primers to initiate lagging strand DNA synthesis (22, 28). The active form of gene 4 protein, like other homologous helicases, is a hexamer (29–31). Mutants of T7 gene 4 protein have been identified that have reduced ability to form hexamers, to hydrolyze TTP, and consequently to catalyze helicase activity (29, 31–34). The T7 gene 2.5 protein is a 26-kDa ssDNA binding protein that removes secondary structure in ssDNA (35, 36) and also facilitates homologous pairing reactions (37). *In vitro* studies using the T7 replication proteins have shown that the basic mechanism of leading and lagging strand DNA synthesis is fundamentally similar to that of other procaryotic and eucaryotic systems (15, 16, 38–40). In studies making use of a preformed replication fork on a duplex circular molecule, Debyser *et al.* (15) showed that leading and lagging strand synthesis was resistant to dilution of the T7 DNA polymerase and the T7 helicase-primase, and that the average size of Okazaki fragments was independent of the concentration of either protein. Our current model is that a single hexamer of T7 gene 4 protein is responsible for both the helicase activity that catalyzes leading strand DNA synthesis and the primase activity that catalyzes lagging strand DNA synthesis. Two DNA polymerase molecules are coupled to the helicase-primase complex; one catalyzes leading strand DNA synthesis, whereas the other catalyzes lagging strand DNA synthesis. Whereas kinetic and physical analysis of the replication systems described above support the coupling of leading and lagging strands, there has been no direct demonstration to date in any system of the essential DNA loop that must occur as a consequence of this coordination. In this study we have used electron microscopy (EM) to examine the architecture of the DNA and DNA-protein intermediates involved in replication reactions employing the T7 replication proteins. Analysis of these replicating molecules provide direct evidence for looping of the lagging strand at the replication fork.

#### EXPERIMENTAL PROCEDURES

**Materials**—The following proteins and DNA were purified as described: T7 gene 5 protein complexed with thioredoxin (19); wild type and D485G mutant gene 4 protein (31, 41); gene 2.5 protein (36, 42), and *E. coli* SSB (43). All proteins were stored at –20 °C and diluted to a working concentration prior to use with a buffer containing 10 mM HEPES (pH 7.5) and 2 mM  $\beta$ -mercaptoethanol.

Circular dsDNA templates containing a preformed replication fork

(44) were constructed as described by Debyser *et al.* (15). In brief, synthetic 52 nucleotide (nt) primers were synthesized and annealed to either M13 mp6 ssDNA (at map positions 6225–6243) or to M13 mp19 ssDNA (at positions 5489–5507). The first 19 nucleotides from the 3' end of each primer were complementary to M13 sequences, whereas the 33 nucleotides at the 5' end were heterologous. The annealed primer was extended using T7 DNA polymerase at 30 °C for 20 min. Incomplete primer extension products and ssDNA were removed by sedimentation through a 5–25% sucrose gradient containing 1 M NaCl at 40,000 rpm in a SW41 rotor for 7 h at 15 °C.

The linear template containing a G-less cassette was prepared as follows. A 392-bp G-less cassette (45) fused to the adeno major late promoter (–50 to +10) that had been cloned into pUC13 between the *Sma*I and *Sst*I (*Sac*I) sites was removed as a 463-bp fragment using *Bam*HI and *Eco*RI restriction endonucleases and cloned into M13mp18 between the *Bam*HI and *Eco*RI sites to generate the construct M13mp18GL. The DNA was transfected into *E. coli* HB101 cells, and the phage and ssDNA were prepared using standard procedures (46). The orientation of the insert was verified by DNA sequencing.

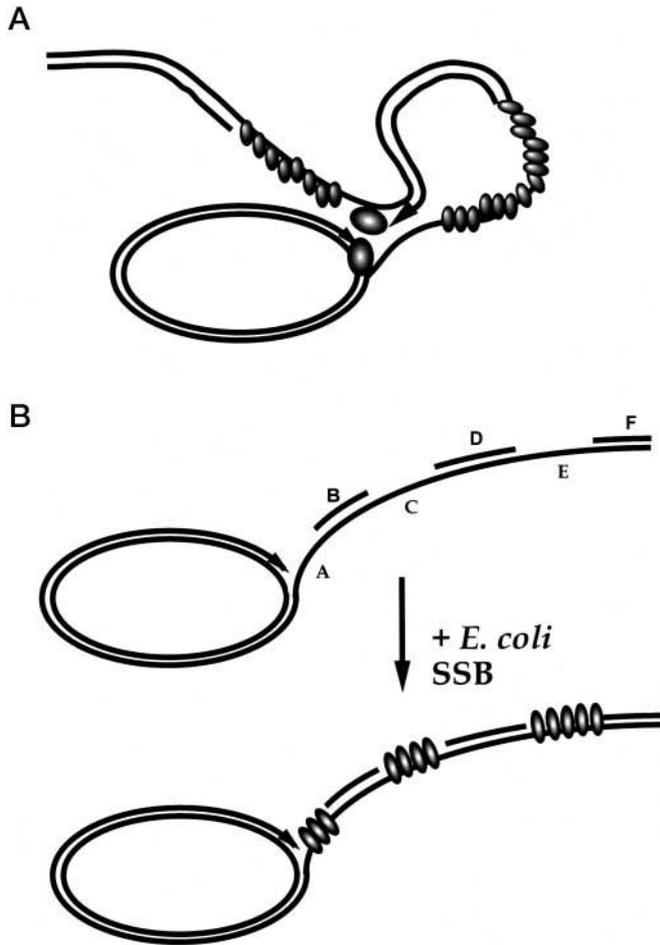
A 52-nucleotide primer (5'-TAAAACCTCTACAAATGTGGTATG-GCTGATTATCCCCTCCATACCCTTCCTC-3') was synthesized such that the nineteen 3' terminal nucleotides are complementary to the sequences just inside the G-less cassette in M13mp18GL (+) ssDNA. The primer was annealed to the template and then extended using T7 DNA polymerase at 30 °C for 20 min. Incomplete primer extension products and ssDNA were removed by sucrose sedimentation as above. The purified circular DNA was cleaved at a single site 747 bp from the beginning of the G-less cassette using *Bgl*II.

**Methods**—Replication reactions were as follows. The dsDNA template was diluted to 17.5  $\mu$ g/ml (3.5 nM) in 10 mM HEPES (pH 7.5), 50 mM NaCl, 10 mM MgCl<sub>2</sub>, 0.1 mM EDTA, 2 mM  $\beta$ -mercaptoethanol, with 0.3 mM ATP and CTP, and 0.6 mM of the 4 dNTPs. To slow the rate of synthesis in some reactions, the concentration of dCTP was decreased to 6  $\mu$ M. When present, T7 gene 5, 4, and 2.5 proteins were added to 88 nM, 88 nM, and 1.5  $\mu$ M each (8, 5.5, and 40  $\mu$ g/ml, respectively). The 63-kDa form of either the wild-type gene 4 protein or the D485G mutant was used unless otherwise noted. The reaction mixtures (25  $\mu$ l) were incubated at 30 °C for 2.5 min with the wild-type gene 4 protein and 3 min with the D485G gene 4 protein. The reactions were terminated by adding EDTA (pH 8.0) to a final concentration of 20 mM. To label the newly synthesized strand with biotin, replication reactions were carried out as above for 2.5 min at 30° with 60 nM dCTP and 15 nM biotinylated dCTP (ENZO Diagnostics), followed by treatment with 0.6% glutaraldehyde for 5 min. The glutaraldehyde was quenched with a molar excess of Tris buffer, and streptavidin was then added (10  $\mu$ g/ml) for 10 min at 21 °C prior to dilution of the sample 10-fold in buffer and preparation for EM.

**Electron Microscopy**—To visualize ssDNA tracts on the DNA replication intermediates, aliquots were removed from the replication reactions, and EDTA and NaCl were added to 20 mM and 1 M, respectively, to release the proteins. Circular M13mp19 ssDNA was added to 5  $\mu$ g/ml to provide a measure of the compaction of the ssDNA tracts, and the mixture was diluted 10-fold with 10 mM HEPES (pH 7.5) and 2 mM EDTA. *E. coli* SSB was then added to 1  $\mu$ g/ml for 2 min at 21 °C, and the samples were fixed with glutaraldehyde and prepared for EM as described below. The length of SSB-covered ssDNA segments was converted to an equivalent length of duplex DNA by comparing its length to that of fully SSB-covered M13mp19 ssDNA circles in the field, which in turn were related to the length of duplex M13 dsDNA circles in the same preparation. In general, the length of ssDNA bound by SSB is one-fourth the length of an equivalent number of base pairs of dsDNA.

DNA-protein complexes were prepared for EM as described in Ref. 47. Aliquots of the reaction mixtures were fixed by the addition of glutaraldehyde to 0.6% for 5 min at 21 °C and then either diluted to 0.5  $\mu$ g/ml (DNA concentration) or chromatographed through Biogel A5M to remove unbound proteins and fixatives. The samples were mixed with a buffer containing 2 mM spermidine and directly adsorbed to glow-charged thin carbon foils for 1 min and then washed in a water-ethanol series, air-dried, and rotary shadowcast with tungsten at  $1 \times 10^{-6}$  torr. Samples were examined in a Philips EM400 or CM12 instrument at 40 kV. Length measurements were made by projecting images on the micrographs onto a Summagraphics digitizing tablet coupled to a McIntosh Quadra 650 computer programmed with software developed in the laboratory of Jack D. Griffith.

<sup>1</sup> The abbreviations used are: ss, single-stranded; ds, double-stranded; EM, electron microscopy; bp, base pair(s); nt, nucleotide(s); bio-dCTP, biotinylated dCTP; kb, kilobase(s).



**FIG. 1. Schematic models showing possible replication intermediates on a duplex circular template containing a replication fork.** *A*, coupling of leading and lagging strand DNA polymerases results in the lagging DNA strand forming a loop. *B*, DNA intermediates after removal of the replication proteins. Initiation of DNA synthesis on a preformed fork on a circular duplex template generates replicative intermediates that consist of a duplex circular molecule that contains a DNA tail. If the displaced strand is synthesized discontinuously by initiation with RNA primers, then the DNA tail will contain duplex regions punctuated by one or more ss gaps. Segment A (ss) corresponds to leading strand DNA synthesis that has occurred since the initiation of the last Okazaki fragment. Segment B (ds) corresponds to lagging strand synthesis that has occurred on the most recently synthesized Okazaki fragment. Segment C (ss) corresponds to the remaining template to be synthesized by the most recently initiated Okazaki fragment. Segments B plus C correspond to the theoretical total length of the most recently initiated Okazaki fragment after complete synthesis to the 5' end of the downstream Okazaki fragment. A molecule with ss regions at A and C represents one in which the proximal Okazaki fragment is still being synthesized, whereas one with only a ss region at A would correspond to one in which the proximal Okazaki fragment has been completely synthesized and initiation of a new Okazaki fragment has not yet occurred. Molecules with three or more ss regions are the result of incomplete synthesis of downstream Okazaki fragments. To visualize the ss regions by EM in these studies, the molecules are complexed with *E. coli* SSB protein to extend and thicken the ss regions as illustrated in the lower diagram.

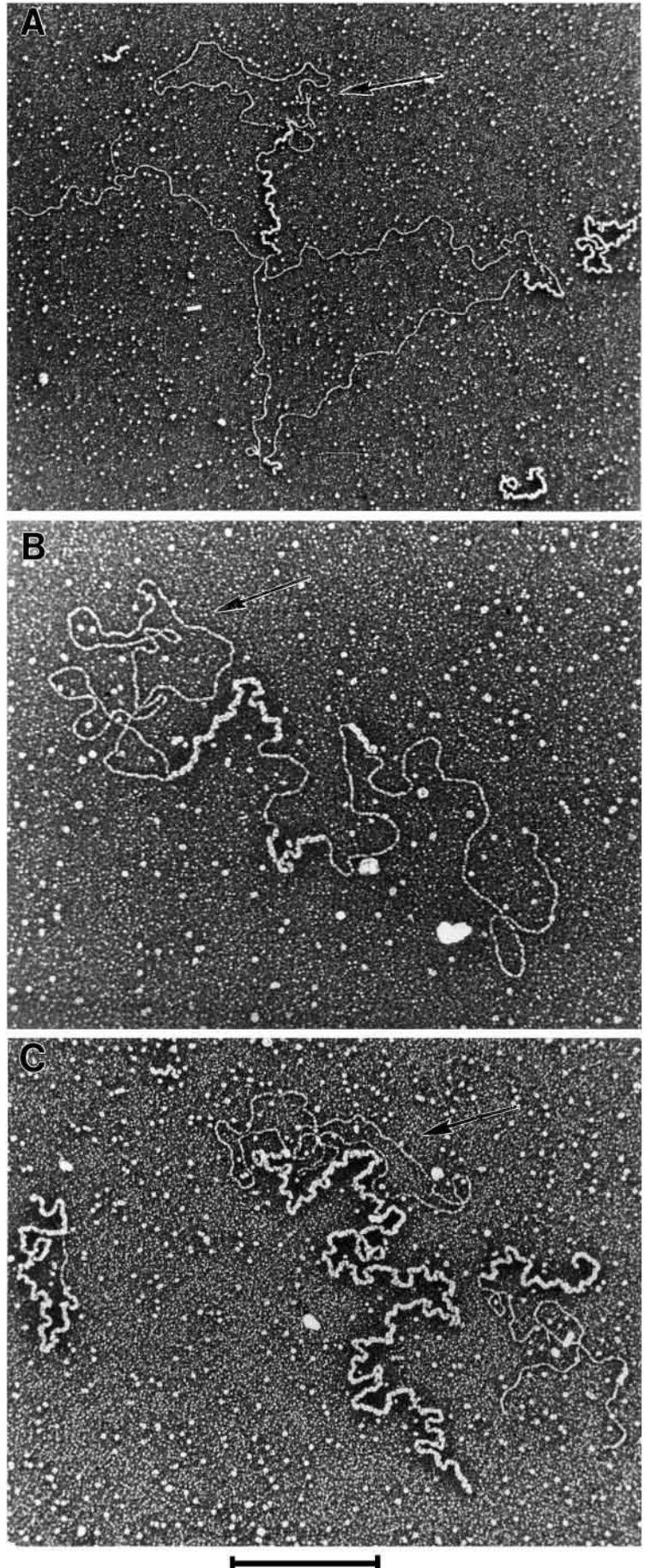
## RESULTS

**Deproteinized Replicative Intermediates Reveal Long DNA Tails with Okazaki Fragments**—An examination of the pattern of ds and ss segments along the rolling circle tails of deproteinized replicative intermediates should provide information about the progress of replication and the nature of the intermediates. The potential types of molecules generated are illustrated in Fig. 1*B*. Incubations were carried out with the circular DNA template containing a preformed fork and the T7

replication proteins (see “Experimental Procedures”). Aliquots were removed from the incubations, the T7 proteins were released with high salt, and the ssDNA segments bound by *E. coli* SSB prior to preparation for EM. SSB extends and coats ssDNA making it easily distinguished from dsDNA (48). We compared the replication intermediates in reactions with wild-type gene 4 protein to those generated using a mutant of gene 4 protein containing a single amino acid residue change at position 485 from Asp to Gly. The D485G mutant, being less efficient as a helicase, facilitated the control of the rate of leading strand DNA synthesis.

Examination of fields of molecules taken from 2- to 3-min incubations employing either gene 4 protein revealed ~30% of the circular M13 dsDNAs with linear tails attached (replicative intermediates) (Fig. 2, *A–C*). In some cases (Fig. 2*A*) the tails measured >4 times the length of the monomer circle, and frequently several alternating tracts of dsDNA and thick SSB-covered ssDNA segments were present. When the 56-kDa gene 4 protein (D485G) was used, which lacks the primase activity present in the 63-kDa form, the tails were fully ss (Fig. 2*C*) since priming is required to initiate Okazaki fragments. Reactions using the 63-kDa gene 4 protein were carried out in parallel with or without gene 2.5 protein, and fields scored for molecules with tails that contained 0, 1, 2, or 3 or more ssDNA tracts. In nearly every case the first ss tract was at the junction of the circle and DNA tail; the very rare exceptions were not scored. Adding gene 2.5 protein reduced significantly the percentage of molecules with two ssDNA tracts; 31% of the molecules scored fall in this class in the absence of gene 2.5 protein, compared with 9% in the presence of gene 2.5 protein (Table I;  $n = \sim 1000$ ). Since a molecule with two ssDNA tracts contains a replication fork in which the Okazaki fragment is still being synthesized, this suggests that gene 2.5 protein increases the rate of lagging strand synthesis. Lagging strand elongation appears to be rate-limiting in the absence of gene 2.5 protein; in its presence 68% of the replicating molecules were “caught” with one ssDNA gap at the junction of the template and tail (Table I). In these molecules, synthesis of the previous Okazaki fragment was complete, and synthesis of a new fragment had not yet started. These molecules may be in the process of recycling in which the lagging strand DNA polymerase has to slide back to a new RNA primer. Alternatively, these complexes may reflect the pausing of the lagging strand polymerase until a new primase recognition site is encountered on the displaced leading strand, and a new RNA primer is synthesized.

**Gene 4 Protein Concentration Has Little Effect on the Length of Okazaki Fragments**—Debyser *et al.* (15) observed that the distribution of Okazaki fragment lengths was independent of the concentration of the helicase-primase in these reactions. To verify that this relation held for these experiments and to provide an additional confirmation, replication reactions were carried out with a fixed concentration of T7 gene 5/thioredoxin and varying concentrations of gene 4 protein either in the presence or absence of gene 2.5 protein. The molar ratio of T7 DNA polymerase to template in these experiments was 4:1 and the molar ratio of gene 4 protein hexamers to template varied from 4:1 to 0.4:1 (see “Experimental Procedures”). Mixtures were incubated for 2.5 min at 30 °C and the samples were prepared for EM as described above. Micrographs of fields of molecules were taken that allowed at least 100 Okazaki fragments to be measured for each concentration of gene 4 protein. For this analysis, an Okazaki fragment was defined as beginning at the start of a duplex segment on the replication tail, continuing toward the distal end of the tail through the duplex segment, and ending at the far end of the adjoining ss gap (Fig. 1*B*). This definition is based on the assumption that continued



**FIG. 2. Architecture of DNA replication intermediates.** Replication reactions employing M13mp6 dsDNA circles (*arrows*) with a preformed fork and T7 gene 5/thioredoxin, gene 4, and gene 2.5 proteins were carried out for 2.5 min at 30 °C (see "Experimental Procedures"). DNA synthesis was stopped by addition of NaCl to 1 M, and *E. coli* SSB was added to thicken and extend the ssDNA. In *panels A and B*, wild-type 63-kDa gene 4 protein was used, whereas in *panel C*, the 56-kDa D485G gene 4 protein (that lacks primase activity) was used. Samples were visualized by adsorption onto thin carbon foils, followed by air-drying and rotary shadowcasting with tungsten (see "Experimental Procedures"). Shown in reverse contrast, The *bar* equals a length of dsDNA equivalent to 0.5 kb (*A*) or 1.0 kb (*B* and *C*).

TABLE I

Effect of T7 single-strand DNA binding protein (gene 2.5 protein) on coupling of leading and lagging strand DNA synthesis

Fraction of the tails on the M13 dsDNA circles containing various numbers of ssDNA gaps. Synthesis reactions were carried out using T7 gene 5 protein/thioredoxin and wild-type gene 4 protein in the presence or absence of gene 2.5 protein for 3 min at 30 °C (see "Experimental Procedures"). The proteins were removed with 1 M NaCl, the ssDNA were coated with *E. coli* SSB, and the DNA examined by EM ( $n = \sim 1,000$ ). Molecules with one gap were almost exclusively ones in which the ssDNA joined the tail to the circle.

	# of ss gaps			Three or more
	None	One	Two	
With gene 2.5 protein (%)	17	68	9	5
Without gene 2.5 protein (%)	15	50	31	4

synthesis would have filled in the ss gap.

EM analysis (Fig. 3A) revealed that in the presence of gene 2.5 protein, the Okazaki fragment length distributions peak near 2300 nt independent of gene 4 protein concentration. Omission of gene 2.5 protein (Fig. 3B) generates Okazaki length distributions that peak at a similar value (2100 nt). However, there were  $\sim 2$ -fold more Okazaki fragments having lengths less than 1000 nt as contrasted to the distributions generated in the presence of gene 2.5 protein. These experiments were repeated using the D485G mutant gene 4 protein. Very similar distributions were obtained over the 10-fold range of D485G gene 4 protein concentration with peaks between 1700 and 1600 nt in the presence or absence of gene 2.5 protein respectively, with a higher number of shorter Okazaki fragments seen in the absence of gene 2.5 protein (data not shown). As discussed below (see "Discussion") these observations support but do not conclusively prove coupling in this system.

**Replication Intermediates Contain a Replication Complex at the Fork That Frequently Contains a DNA Loop**—To visualize the replicative intermediates with the T7 proteins bound, aliquots from the reactions were treated with glutaraldehyde and prepared for EM (see "Experimental Procedures"). Examination by EM revealed that  $\sim 50\%$  of the M13 dsDNA circles contained a single large complex (replication complex). Of these, 10–25% had a long dsDNA tail extending from the replication complex (Fig. 4). The frequency of such forms was not altered significantly by omission of gene 2.5 protein. Occasionally much smaller protein complexes were observed along the dsDNA tails, most commonly in reactions carried out with the mutant D485G gene 4 protein in the absence of gene 2.5 protein (not shown).

The replication complexes were of varied sizes and shapes (Fig. 4). In addition to two copies of gene 5 protein/thioredoxin and at least one gene 4 protein hexamer as suggested in Debyser *et al.* (15), these complexes must also contain a variable amount of ssDNA complexed by gene 2.5 protein since, as shown above,  $\sim 70\%$  of the replicative intermediates contain a ssDNA segment joining the tail to the circle. When gene 2.5 protein was omitted, replication complexes with associated dsDNA tails were still seen and at roughly the same frequency. Qualitatively, however, these complexes were not as large, and small protein complexes were more frequently observed along the dsDNA tails, likely reflecting polymerase molecules engaged in completing Okazaki fragment synthesis. Thus it would appear that in the presence of gene 2.5 protein the ssDNA segments on the lagging strand are condensed rather than stretched out into extended ssDNA-protein filaments. The possible significance of this is discussed below (see "Discussion").

Examination of the replicating circles (dsDNA circle with a replication complex and dsDNA tail) revealed that a dsDNA loop was frequently associated with the replication complex

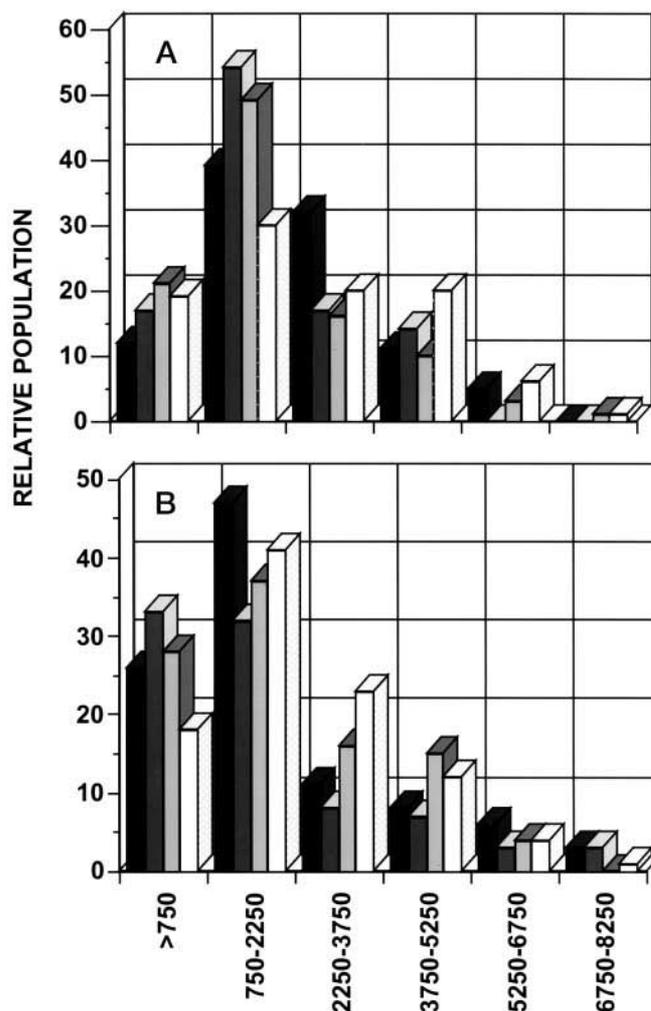


Fig. 3. Dependence of Okazaki fragment length on T7 gene 4 and gene 2.5 proteins DNA synthesis reactions similar to those shown in Fig. 2 containing 88 nM gene 5 protein/thioredoxin were carried out for 2.5 min at 30 °C in the presence (A) or absence (B) of 1.5  $\mu$ M gene 2.5 protein, and varying concentrations of gene 4 protein. The length of Okazaki fragments on the replicating tails was measured from the electron micrographs ( $n = 200$  molecules for each panel) (see Figs. 1 and 2). Gene 4 protein concentrations were 9, 22, 44, or 88 nM (black, dark gray, gray, and white columns; left-to-right, respectively).

(Fig. 4). The fraction of replicating circles containing a dsDNA loop depended on the form of gene 4 protein used and whether gene 2.5 was present. In the absence of gene 2.5 protein, 17% of the replicating circles contained a dsDNA loop with either form of gene 4 protein, whereas in the presence of gene 2.5 protein, 27% and 40% of the replicating circles contained a dsDNA loop when either wild-type or the D485G mutant form of gene 4 protein was used, respectively. These values were determined from at least three separate experiments each, and from analysis of  $>200$  replicating molecules each. It will be shown below that looping occurs only during active displacement of the lagging strand; hence use of a less efficient helicase would be expected to result in a larger fraction of molecules with loops.

The length of the loops varied from 250 bp to 4 kb with a peak near 0.5 kb and mean length close to 1 kb (Fig. 5). Loops were not observed along the dsDNA tails suggesting that looping reflects events at the replication fork. To directly establish that the loops contain newly synthesized DNA, a biotin labeling approach was used and is presented in the next section.

**Loops at the Replication Fork Consist of Newly Synthesized DNA**—To localize the newly synthesized DNA in the replicat-

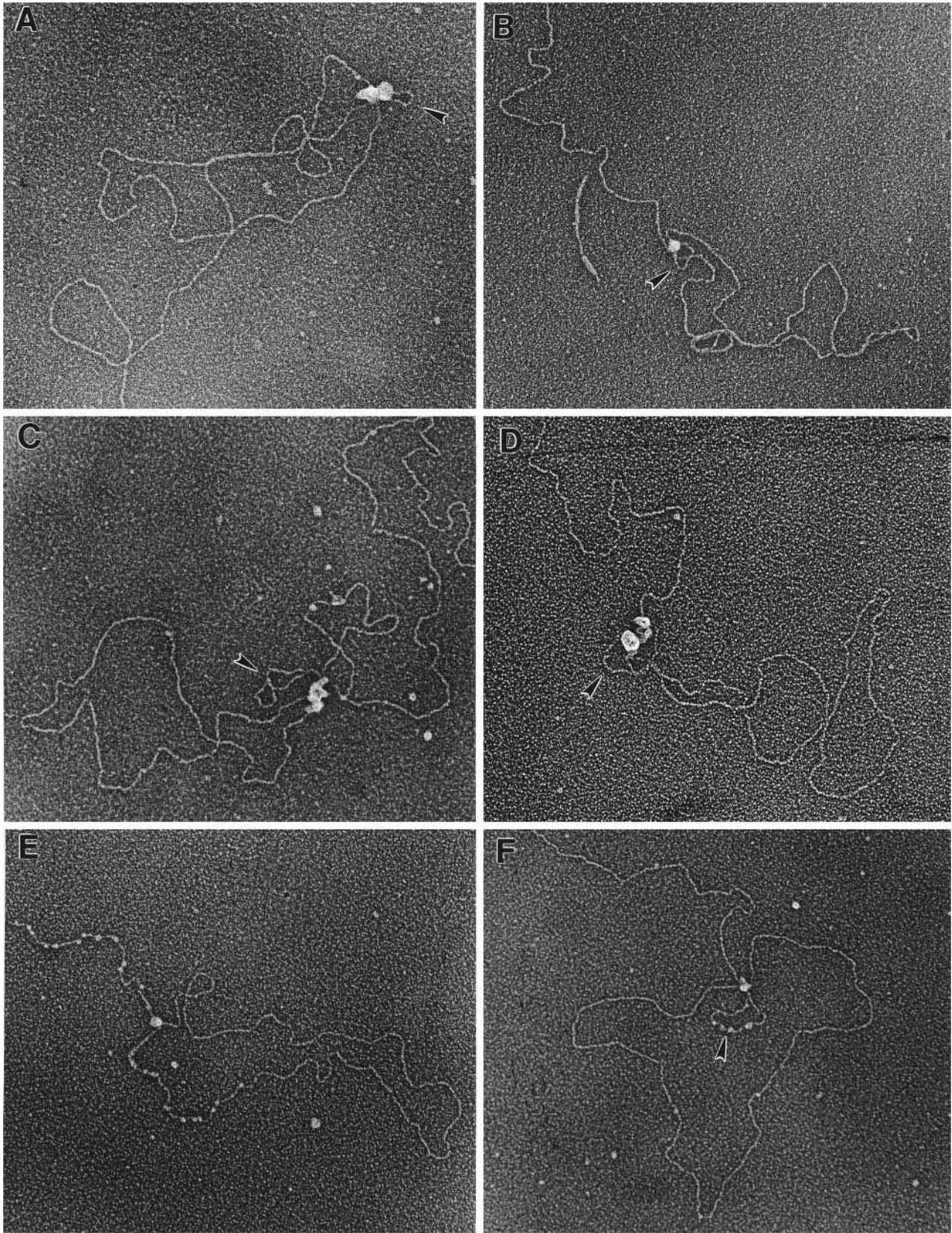


FIG. 4. **Visualization of DNA-protein replication complexes.** In panels A–D, M13 dsDNA circles with a preformed fork were incubated with T7 gene 5/thioredoxin, gene 4, and gene 2.5 proteins for 2.5 min at 30 °C as described in the text. Following fixation of the samples and gel filtration, the samples were prepared for EM by adsorbing onto thin carbon supports, washing, drying, and rotary shadowcasting with tungsten (see “Experimental Procedures”). In panels E and F, reactions were carried out as in panels A–D except that biotin-tagged dCTP was included in the reaction mixture, and after the reaction was stopped, streptavidin was added prior to preparing the samples for EM. Streptavidin balls mark the location of the newly synthesized DNA. Arrows (panels A–D and F) indicate loops at the replication fork. Shown in reverse contrast, the bar equals a length of DNA equivalent to 1 kb.

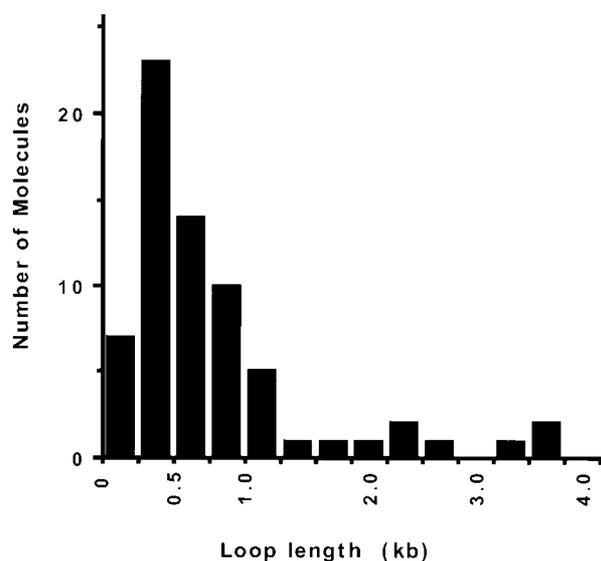


FIG. 5. **Length distribution of DNA loops at the replication fork.** The length of the dsDNA loops at the replication fork were measured from molecules on micrographs as were those illustrated in Fig. 4.

ing circles, biotinylated dCTP (bio-dCTP) was added to the replication reactions (see "Experimental Procedures"). A 4:1 ratio of dCTP to bio-dCTP was employed, resulting in the incorporation of one biotin-dCMP molecule every 100 bp as determined by EM following the addition of streptavidin to mark the location of the biotin. This value reflects some discrimination against bio-dCTP by T7 DNA polymerase. Since the strand of the template circle containing the 5'-extended tail had been synthesized previously in reactions with no bio-dCTP and the incubation time with bio-dCTP was short (2.5 min with a low concentration of dCTP), this approximated a pulse-labeling experiment. Replication reactions were carried out using wild-type gene 4 protein and gene 2.5 protein followed by addition of streptavidin and preparation for EM (see "Experimental Procedures"). As shown in Fig. 4, *E* and *F*, replicating circles with tails and loops were observed, and in 22 of 29 circles with loops scored, the loops were labeled by streptavidin, indicating that those loops were formed from newly synthesized DNA. The labeling was frequently observed to occur in patches with labeled patches flanked by unlabeled patches, and this was most prominent along the linear replicating tails. This patch-wise labeling may be responsible for the 7 molecules scored as not labeled with streptavidin. The patches were similar in length to the Okazaki fragments (about 2 kb) and were observed in numerous experiments. We do not understand the origin of this intriguing variability. One possibility is that there is variable editing by the polymerase of these nucleotides once they are incorporated. This seems unlikely though since the exonuclease activity of T7 DNA polymerase would not be expected to degrade up to several thousand bases before the enzyme switched back to a polymerizing mode. Another possibility is that the polymerase can adopt two slightly different conformations upon initiation of DNA synthesis and that each of these have a different ability to discriminate against the incorporation of the biotin analog.

**Synchronized Replication of a Linear Template: Loops Reflect Lagging Strand Events**—Few replicating circles had been observed following replication reactions of 0.5 to 1 min, suggesting a lack of synchrony in the establishment of active replication complexes. Presumably this is due to a poor efficiency by the T7 gene 4 protein to bind to the relatively short tail on the preformed replication fork. To facilitate the study of early

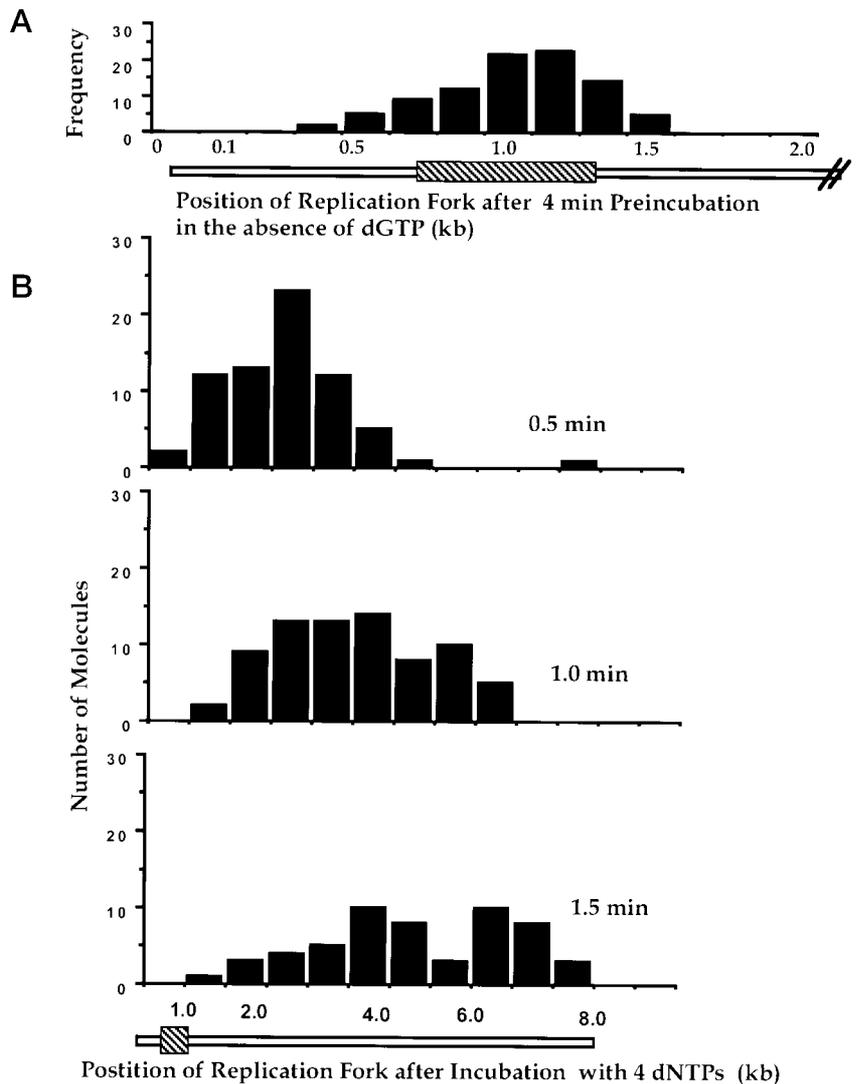
events, a linear DNA template containing a preformed fork at the beginning of a ~400-bp G-less cassette was constructed (see "Experimental Procedures"). Following incubation in the absence of dGTP, the replication proteins were released with high salt, and the ssDNA were complexed with SSB, revealing linear DNAs with short SSB-covered ssDNA stubs. Measurement of the position of these stubs showed that in the absence of dGTP the replication forks were in close proximity to the G-less cassette (Fig. 6A) with the distribution sharpening as the incubation was extended from 1 to 5 min (data not shown). Using a 5-min preincubation period without dGTP, 0.6 mM or 6  $\mu$ M dGTP was added. Aliquots were removed at 0.5, 1.0, 1.5, and 2 min, and the DNA was prepared for EM as above (Fig. 7C). As the incubation time increased from 0.5 to 1.5 min (with 6  $\mu$ M dGTP), the population of forks moved from a region just ahead of the G-less cassette to a broad central distribution, to a population approaching the far end of the DNA (Fig. 6B). When the average position of the replication forks was plotted for reactions containing 0.6 mM or 6  $\mu$ M dGTP, linear regression of the data (not shown) revealed that the plots overlapped between those obtained in the presence or absence of the gene 2.5 protein, suggesting that the presence of gene 2.5 protein does not increase the rate of leading strand synthesis.

Incubation of the linear template with the T7 proteins for 5 min (without dGTP) followed by fixation and preparation for EM showed that 68 to 78% of the templates had a large protein complex located ~750 bp from the closest DNA end (not shown). When the 5-min preincubation period was followed by a 1-min incubation period with 0.6 mM or 6  $\mu$ M dGTP, linear templates containing a replication complex downstream of the origin with an associated loop or tail were common. Some of these replicating molecules exhibited both a dsDNA loop and a dsDNA tail at the replication fork (Fig. 7, *A* and *B*), whereas others had only a dsDNA loop. Fields of molecules were scored (>200 replicating molecules for each condition) from incubations with and without gene 2.5 protein and with 0.6 mM or 6  $\mu$ M dGTP. With either concentration of dGTP, 26% of the replicating DNAs had a dsDNA loop at the replication fork in incubations containing gene 2.5 protein, whereas in the absence of gene 2.5 protein, 21 and 8% of the replicating molecules showed a dsDNA loop at dGTP concentrations of 0.6 mM and 6  $\mu$ M dGTP respectively. The latter value (8%) may also reflect a diminished number of total molecules engaged in replication at the lower dGTP concentration.

For the linear replicating DNA, the position of the replication complex provides a measure of leading strand synthesis from the preformed fork. The extent of lagging strand synthesis is equal to the sum of the linear dsDNA tail, the dsDNA loop, and any ssDNA present on the lagging strand. Measurement of the position of the fork and lengths of the dsDNA tail and dsDNA loops was carried out for 25 linear replicating molecules from incubations containing gene 2.5 protein (Table II). In all 25 molecules containing dsDNA loops (Table II), the sum of the length of the dsDNA loop and dsDNA tail (when present) was less than the total distance that the fork had moved from the origin. This difference must represent the amount of ssDNA present at the fork bound by gene 2.5 protein. The observation that loops are present only when there is a ss segment present on the lagging strand at the fork and not when the tail is fully ds points to the involvement of the loops in lagging strand synthesis.

Of the 25 DNAs, 10 contained only a dsDNA loop and no dsDNA tail. In 8 of the 10, the replication fork had progressed only approximately one-third the length of the DNA template. In the 15 molecules that contained both a dsDNA loop and a dsDNA tail, 12 of the 15 had replication forks that had moved

**FIG. 6. Synchronization of replication on a linear template.** *A*, preincubation reaction. A linear M13 dsDNA template, which contains a ~400 bp G-less cassette, 747 bp internal from the left end of the DNA was constructed; a preformed fork is present at the left end of the G-less cassette. On this template replication will proceed for 400 bases rightward in the absence of dGTP. Replication reactions were carried out as in Fig. 2 for 4 min with dGTP withheld. The position of replication complexes on the DNA was determined by EM and is shown relative to the position of the G-less cassette (*hatched bar*). Due to the size of the SSB-covered-ssDNA complex, some uncertainty in mapping the exact location of the fork was expected. Thus complexes appearing to map beyond the boundary of the G-less cassette may reflect such measurement error. *B*, synchronization reaction. Following the 4-min preincubation period in the absence of dGTP to load the templates with replication complexes, dGTP was added to the replication reactions and incubation continued for 0.5, 1.0, and 1.5 min. Samples were prepared for EM as in Fig. 2, and the distance the fork had moved was determined from the micrographs. The *hatched bar* indicates the position of the G-less cassette on the full-length M13 DNA.



beyond the first one-third of the DNA. This suggests that the first step in generation of a dsDNA loop may be the tethering of the terminus of the displaced ssDNA tail to the replication machinery.

#### DISCUSSION

In this work the architecture of intermediates in DNA synthesis catalyzed by the T7 replication proteins acting on circular and linear ds M13 DNA has been examined by electron microscopy. The study is based on previous biochemical work that provided evidence for coordinated leading and lagging strand synthesis using similar experimental conditions (15). In the present study examination of the intermediates in DNA synthesis reveals a single large complex at the replication fork and in up to 40% of the replicative intermediates, a dsDNA loop is present, attached to the template circle by the replication complex. The mean length of this loop is ~1 kb, roughly one-half the length of the Okazaki fragments. Use of a linear DNA template that provides efficient synchronization of the reactions demonstrates that the dsDNA loops are established early in replication, are associated with Okazaki fragment synthesis, and are present only while lagging strand displacement is underway. These features are consistent with the models of lagging strand looping first proposed for T4 DNA replication (1). One potential departure from the trombone models as they are illustrated is that the ss segments on the lagging strand may be arranged by gene 2.5 protein into compact structures

associated with the replication complex. As discussed below, this could provide several valuable functions in replication. Finally the loops may be initiated by a mechanism involving tethering of the displaced strand to the replication machinery.

*Evidence for Coupled Leading/Lagging Strand Replication*—Coupling of leading and lagging strand synthesis implies that synthesis on both strands is carried out by a single replisome. The complexes observed at the replication fork appeared as single (although polymorphic) structures with no suggestion of two distinct complexes that might be engaged in separate leading and lagging strand synthesis. Using EM, it was observed that the overall effect of decreasing the concentration of gene 4 protein in the reaction was to reduce the number of molecules engaged in replication. Despite this, relatively small differences were observed in the distribution of Okazaki fragment lengths over a 10-fold dilution of gene 4 protein whether the reactions utilized the wild-type or D485G mutant gene 4 proteins. The rate of replication, however, appeared to be qualitatively faster when the wild-type gene 4 protein was employed. This is expected considering that the mutant enzyme is less active as a helicase (31). These EM observations are in excellent agreement with the work of Debyser *et al.* (15) and support but do not conclusively prove coupling in this system. The T7 gene 4 protein contains both the helicase and primase activities; thus in these studies both activities were diluted as contrasted to previous experiments in T4 or *E. coli* where just

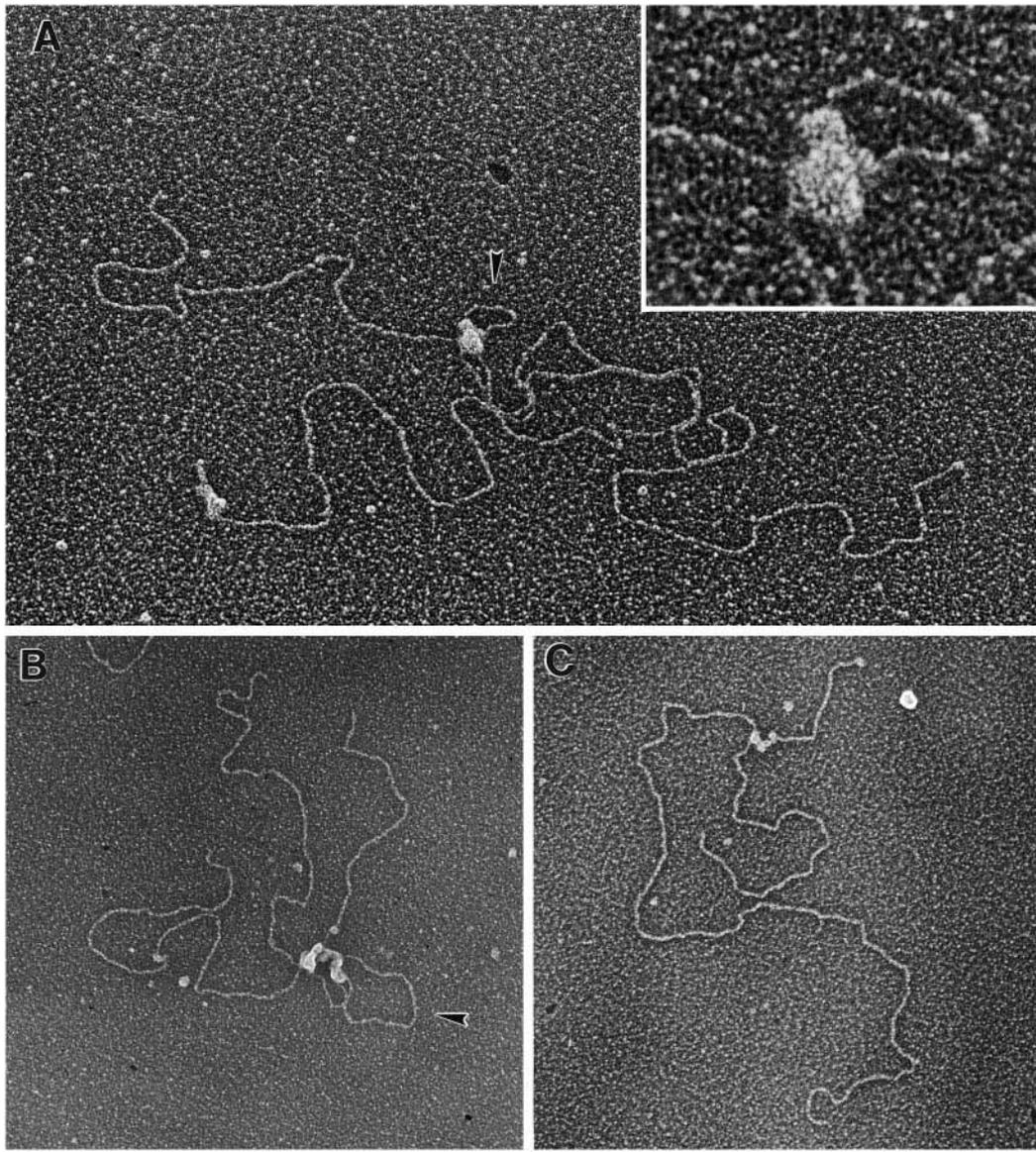


FIG. 7. **Visualization of synchronized replication intermediates.** *A* and *B*, linear M13 DNA containing a  $\sim$ 400-bp G-less cassette with a preformed fork was incubated with T7 gene 5/thioredoxin, T7 gene 4, and T7 gene 2.5 protein for 4 min in the absence of dGTP to synchronize the formation of replication complexes as described in the legend to Fig. 6. Following the preincubation, dGTP was added and replication was allowed to proceed for one min at 30 °C. The samples were fixed and prepared for EM as described in the legend to Fig. 4. The *inset* in *A* is a higher magnification of the complex indicated by the *arrow*. *C*, reactions were carried out as in *A* and *B* above, but after the reactions were terminated, the replication proteins were removed with high salt, and the ssDNA complexed with *E. coli* SSB. The ssDNA coated with SSB can be seen as a thick segment joining the linear tail to the circular template. Shown in reverse contrast, the *bar* is equivalent in length to dsDNA of 1.4 kb (*A*) or 1.0 kb (*B* and *C*).

one component was diluted.

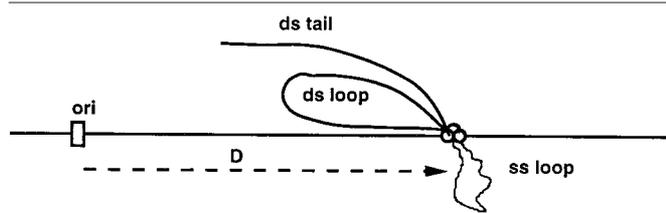
**Role of Gene 2.5 and Gene 4 Proteins in Replication and Looping**—The observation that in the absence of gene 2.5 protein, synthesis of Okazaki fragments are not finished as quickly (Table I) suggests that without gene 2.5 protein synthesis on the lagging strand is retarded, resulting in a higher fraction of the molecules being engaged in Okazaki fragment synthesis. These results are not unexpected since presumably one important role of the ssDNA binding protein is to remove secondary structure on the lagging strand. Further, gene 2.5 protein showed no influence on the rate of leading strand synthesis for the linear templates. The appearance of a larger number of short Okazaki fragments in the absence of gene 2.5 protein suggests that priming is more frequent in this case and that gene 2.5 protein either blocks some of the primer recognition sites or inhibits the use of the primers by T7 DNA polym-

erase. Gene 2.5 protein was not absolutely required for loop formation but did significantly increase the fraction of replicating molecules with loops. These observations point to multiple essential roles for gene 2.5 protein: facilitating lagging strand synthesis by removing secondary structure, reducing the number of priming events, and organizing and stabilizing the replication assembly.

Both gene 4 helicase-primase and gene 2.5 single strand binding protein had a direct influence on the fraction of replicating molecules containing a loop. This fraction was as high as 40% when the reactions contained the D485G mutant gene 4 protein and gene 2.5 protein, which was higher than equivalent reactions using the wild-type gene 4 protein. Looping is a dynamic process involved in synthesis of the lagging strand. Hence the loops may exist only during those phases of the replication process related to Okazaki fragment synthesis.

TABLE II  
Analysis of the length of ds and ss segments in replicative intermediates from a synchronized linear template

D = distance fork has moved from origin (*ori*).



DNA length	D in ds loop	D in ds tail	D calculated to reside in ss loop
		%	
9	67	0	33
15	80	0	20
15	33	0	67
16	56	0	44
21	48	0	44
27	26	0	74
28	25	0	75
31	52	0	48
67	30	0	35
81	65	0	35
9	56	33	11
17	41	41	18
24	38	25	37
36	42	5	53
37	52	11	31
39	8	8	84
52	13	70	18
70	20	26	64
75	5	48	47
83	23	28	49
84	31	56	13
85	11	40	49
87	22	22	56
89	51	34	15
90	44	12	43

Thus a slower helicase-primase, such as the D485G mutant form might be expected to increase the fraction of molecules with a loop by retarding lagging strand synthesis and increasing the time spent in leading strand synthesis.

Indeed in the T4 trombone model, it was suggested largely on the basis of EM results that the lagging strand DNA polymerase proceeds at a higher rate than the leading strand DNA polymerase and also spends 75% of the replication time in the recycling mode (49) during which time the trombone loop is presumably very small or absent. Thus, in sum, our data indicate a pivotal role for the T7 helicase-primase in coordinating DNA replication at the T7 DNA replication fork.

**Arrangement of the ssDNA Segments on the Lagging Strand**—When the replication complexes were prepared for EM by glutaraldehyde fixation to lock the proteins in place, it was observed that the ssDNA segments on the lagging strand were not present as extended filaments of ssDNA bound by gene 2.5 protein but rather were localized to the replication fork in compact bodies. There are several possible explanations for this observation. First prior to fixation they could have been more extended but became condensed during fixation. Whereas filaments of ssDNA bound by T4 gene 32 protein and *E. coli* SSB withstand similar glutaraldehyde treatment appearing as extended filaments, they are likely more robust structurally than gene 2.5-ssDNA complexes. Indeed, preparation of complexes of gene 2.5 protein bound to M13 ssDNA for EM by negative staining (without fixation) in replication buffer<sup>2</sup> revealed more filamentous complexes similar to those shown

before (36). Negative staining itself, however, involves exposure of the sample to salt and low pH and thus is also not totally free of possible artifacts. It is also possible that the way in which gene 2.5 protein is assembled onto the nascent ssDNA at the replication fork by the replication machinery is different from its assembly onto ssDNA when gene 2.5 protein and ssDNA are mixed in bulk. Further detailed experiments will be required to address this issue.

If the ssDNA segments on the lagging strand are arranged into compact structures by gene 2.5 protein, this could serve several functions including protecting the newly synthesized 5' terminus of the Okazaki fragment from nucleases. It could also keep the ssDNA sequestered from *E. coli* proteins such as SSB and *recA*, which might generate 3-stranded synaptic structures were they to assemble onto the newly displaced lagging strand with the preceding duplex leading strand and thus halt the progress of the replication fork. Indeed, a 5' terminus at a ss-ds junction provides a potent initiation site for *recA* polymerization (50).

**A Model of the Looped T7 Replication Fork**—Implicit in the trombone model is the presence of a dimeric DNA polymerase complex, coupling of leading and lagging strand replication via a single replisome, and the presence of a loop in the lagging strand that grows and collapses in tune with the events of Okazaki fragment synthesis, completion, and priming of a new fragment. Whereas direct evidence for a dimeric polymerase in the T7 system has not been obtained, dimeric polymerase complexes have been observed in other similar systems (see the Introduction). We present here direct EM evidence for the presence of a DNA loop that shares most of the predicted properties for the loop in the trombone model. One possible difference from these models as they are drawn relates to the arrangement of the ssDNA segments as described above. In the models, the ss segment of the lagging strand generated by displacement at the fork is drawn as part of an extended loop bound by a ssDNA binding protein. In this work we observed that the segments of ssDNA bound by gene 2.5 protein at the replication fork appear compact and dense.

A further modification of the classic models is suggested by the observation of molecules with loops but no dsDNA tail early in the replication process from the linear templates. This finding was reminiscent of the rolling loop replication in  $\phi$ X174 ds to ss replication (51) in which the end of the displaced strand is tethered to the replication complex. Such a tethering may serve to initiate the looping process.

#### REFERENCES

- Alberts, B., Morris, C. F., Mace, D., Sinha, N., Bittner, M., and Moran, L. (1975) in *DNA Synthesis and Its Regulation* (Goulian, M., Hanawalt P., and Fox, C. F., eds), pp. 241–269, W. A. Benjamin, Menlo Park, California
- Alberts, B. M. (1987) *Philos. Trans. R. Soc. Lond. Biol.* **317**, 395–420
- Sinha, N., Morris, C., and Alberts, B. (1980) *J. Biol. Chem.* **255**, 4290–4303
- Cha, T.-A., and Alberts, B. (1990) *Biochemistry* **29**, 1791–1798
- Nossal, N. G. (1994) in *Molecular Biology of Bacteriophage T4* (Karam, J. D., ed), pp 43–53, American Society for Microbiology, Washington, DC
- Nossal, N. G., and Alberts, B. M. (1983) in *Bacteriophage T4* (Matthews, C. K., Kutter, E. M., Mosig, G., and Berget, P. B., eds), pp 71–81, American Society for Microbiology, Washington, DC
- Wu, C. A., Zechner, E. L., Reems, J. A., McHenry, C. S., and Marians, K. J. (1992) *J. Biol. Chem.* **267**, 4074–4083
- Kim, S., Dallmann, H. G., McHenry, C. S., and Marians, K. J. (1996) *J. Biol. Chem.* **271**, 21406–21412
- Tougu, K., and Marians, K. J. (1996) *J. Biol. Chem.* **271**, 21398–21405
- Alberts, B. M., Barry, J., Bedinger, B. P., Formosa, T., Jongeneel, C. V., and Kruezer, K. N. (1983) *Cold Spring Harbor Symp. Quant. Biol.* **47**, 655–668
- Wu, C. A., Zechner, E. L., Hughes, A. J., Franden, M. A., McHenry, C. S., and Marians, K. J. (1992) *J. Biol. Chem.* **267**, 4064–4073
- Stukenberg, P. T., and O'Donnell, M. (1995) *J. Biol. Chem.* **270**, 13384–13391
- Yuzhakov, A., Turner, J., and O'Donnell, M. (1996) *Cell* **86**, 877–886
- Richardson, R. W., Ellis, R. L., and Nossal, N. G. (1990) *ICN-UCLA Symp. Mol. Cell. Biol.* **127**, 247–259
- Debyser, Z., Tabor, S., and Richardson, C. C. (1994) *Cell* **77**, 157–166
- Richardson, C. C. (1983) *Cell* **33**, 315–317
- Modrich, P., and Richardson, C. C. (1975) *J. Biol. Chem.* **250**, 5508–5514
- Huber, H. E., Tabor, S., and Richardson, C. C. (1987) *J. Biol. Chem.* **262**,

<sup>2</sup> P. Chastain, III and J. D. Griffith, unpublished results.

- 16224–16232
19. Tabor, S., Huber, H. E., and Richardson, C. C. (1987) *J. Biol. Chem.* **262**, 16212–16223
20. Dunn, J. J., and Studier, F. W. (1983) *J. Mol. Biol.* **166**, 477–535
21. Kolodner, R., Masamune, Y., LeClerc, J. E., and Richardson, C. C. (1978) *J. Biol. Chem.* **253**, 566–573
22. Tabor, S., and Richardson, C. C. (1981) *Proc. Natl. Acad. Sci. U. S. A.* **78**, 205–209
23. Matson, S. W., and Richardson, C. C. (1983) *J. Biol. Chem.* **258**, 14009–14016
24. Matson, S. W., Tabor, S., and Richardson, C. C. (1983) *J. Biol. Chem.* **258**, 14017–14024
25. Bernstein, J. A., and Richardson, C. C. (1988a) *J. Biol. Chem.* **263**, 14891–14899
26. Bernstein, J. A., and Richardson, C. C. (1988b) *Proc. Natl. Acad. Sci. U. S. A.* **85**, 396–400
27. Bernstein, J. A., and Richardson, C. C. (1989) *J. Biol. Chem.* **264**, 13066–13073
28. Scherzinger, E., Lanka, E., Morelli, G., Seiffert, D., and Yuki, A. (1977) *Eur. J. Biochem.* **72**, 543–558
29. Patel, S. S., Hingorani, M. M., and Ng, W. M. (1994) *Biochemistry* **33**, 7857–7868
30. Egelman, E. H., Yu, X., Wild, R., Hingorani, M. M., and Patel, S. S. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92**, 3869–3873
31. Notarnicola, S. M., Park, K., Griffith, J. D., and Richardson, C. C. (1995) *J. Biol. Chem.* **270**, 20215–20224
32. Patel, S. S., and Hingorani, M. M. (1993) *J. Biol. Chem.* **268**, 10668–10675
33. Notarnicola, S. M., and Richardson, C. C. (1993) *J. Biol. Chem.* **268**, 27198–27207
34. Washington, M. T., Rosenberg, A. H., Griffin, K., Studier, F. W., and Patel, S. S. (1996) *J. Biol. Chem.* **271**, 26825–26834
35. Scherzinger, E., Litfin, F., and Jost, E. (1973) *Mol. Gen. Genet.* **123**, 247–262
36. Kim, Y. T., Tabor, S., Bortner, C., Griffith, J. D., and Richardson, C. C. (1992) *J. Biol. Chem.* **267**, 15022–15031
37. Kong, D., and Richardson, C. C. (1996) *EMBO J.* **15**, 2010–2019
38. Engler, M. J., and Richardson, C. C. (1983) *J. Biol. Chem.* **258**, 11197–11205
39. Nakai, H., and Richardson, C. C. (1986) *J. Biol. Chem.* **261**, 15208–15216
40. Huber, H. E., Bernstein, J., Nakai, H., Tabor, S., and Richardson, C. (1988) *Cancer Cells* **6**, 11–17
41. Nakai, H., and Richardson, C. C. (1988) *J. Biol. Chem.* **263**, 9818–9830
42. Kim, Y. T., and Richardson, C. C. (1994) *J. Biol. Chem.* **269**, 5270–5278
43. Chase, J. W., Whittier, R. F., Auerbach, J., Sancar, A., and Rupp, W. D. (1980) *Nucleic Acids Res.* **8**, 3215–3227
44. Lechner, R. L., and Richardson, C. C. (1983) *J. Biol. Chem.* **258**, 11185–11196
45. Sawadogo, M., and Roeder, R. G. (1985) *Proc. Natl. Acad. Sci. U. S. A.* **82**, 4394–4398
46. Register, J. C., and Griffith, J. D. (1986) *Proc. Natl. Acad. Sci. U. S. A.* **83**, 624–628
47. Griffith, J. D., and Christiansen, G. (1978) *Ann. Rev. Biophys. Bioeng.* **7**, 19–35
48. Chrysogelos, S., and Griffith, J. (1982) *Proc. Natl. Acad. Sci. U. S. A.* **79**, 5803–5807
49. Selick, H. E., Barry, J., Cha, T.-A., Munn, M., Nakanishi, M., Wong, M. L., and Alberts, B. M. (1987) *ICN-UCLA Symp. Mol. Cell. Biol.* **47**, 183–214
50. Register, J. C., and Griffith, J. D. (1985) *J. Biol. Chem.* **260**, 12308–12312
51. Eisenberg, S., Griffith, J., and Kornberg, A. (1977) *Proc. Natl. Acad. Sci. U. S. A.* **74**, 3198–3202