ABSTRACT

A synthetic heptaribonucleotide, GACCCCC, which is complementary to a unique site on fd bacteriophage DNA, primes DNA synthesis of fd by T4 bacteriophage DNA polymerase. The rate of the GACCCCC-primed DNA synthesis was not uniform as reflected by the appearance of discrete DNA fragments as replication intermediates on an alkaline agarose gel. After 10 minutes of synthesis a significant fraction of the DNA product ran as a single band with a length of about 1960 nucleotides. We have isolated this DNA fragment, hybridized back to unlabeled fd DNA template, and mapped the Taq I restriction fragments by urea polyacrylamide gel electrophoresis. This fine mapping procedure has located two major pause sites at fd nucleotide positions 5575 and 5674. These sites reside in the stem of two very stable hairpin helices near the origin of DNA replication of fd. Models for the functional roles of these two hairpin helices are presented.

INTRODUCTION

The filamentous bacteriophage fd contains a circular, single-stranded DNA 6408 nucleotides in length. Schaller et al. (1) have located the origin for the synthesis of the complementary strand of fd within a DNA segment of 125 nucleotides centering around nucleotide position 5690 (see Ref. 2 for the numbering of the fd nucleotide sequence map). When the nucleotide sequence of fd DNA in the region of replication origin was scanned by a computer program which searches for potential secondary structure, a large number of hairpin helices were found (our unpublished result). No definite secondary structure has been assigned to this particular primary sequence of fd at present. Schaller proposed that there are at least four potential hairpin helices in this particular region (3). We have recently detected the presence of secondary structures located near fd nucleotide position 5650 using T4 DNA polymerase (4). The fine structural features of the replication origin of fd, however, were not clear at that time because we were unable to determine the positions of the sites to sequence resolution. While the mechanism by which the onset of DNA
replication occurs has yet to be explored, the detailed structural information presented here should further our understanding of this important biological process.

In this communication, we have used restriction endonuclease Taq I and a polyacrylamide gel to determine the detailed location of the 3'-OH termini of the replication intermediates stalled near fd nucleotide position 5650. The fine mapping procedure has located two strong kinetic pause sites at positions 5575 and 5674. These positions reside in the stems of the two most stable hairpins of fd phage DNA in the region of replication origin. Models for the functional role of these two hairpins are presented.

MATERIALS AND METHODS

Enzymes, DNA, and chemicals

The fd viral DNA and the GACCCCC ribonucleotide primer were prepared as previously described (4). The T4 DNA polymerase was a gift from Chung-Cheng Liu and Bruce M. Alberts. The restriction endonuclease Taq I was purchased from Bethesda Research Laboratories. [α-32P]dATP (410 Ci/mmol) was purchased from Amersham Corporation. PEP1 buffer was prepared by dissolving NaClO4 (obtained from G.F. Smith Chemical Co., Columbus, Ohio) in distilled water to give a 6.5 M solution. To 92.5 ml of this solution 5 ml of 1 M potassium phosphate (pH 7) and 2.5 ml of 0.5 M Na3EDTA were added. The solution was filtered through GF/C filter (Whatman glass fiber) twice and stored at 4°C.

GACCCCC-primed fd DNA synthesis

Solutions (38 μl) containing 1 μl T4 DNA polymerase, 3.5 μg/ml fd viral DNA, 3.5 μg/ml GACCCCC, 0.1 mM dCTP, 0.1 mM dGTP, 0.09 mM dATP, 2 mM dithiothreitol, 33 mM Tris-acetate (pH 7.8), 66 mM potassium acetate and 10 mM magnesium acetate were prepared at room temperature. An incubation was then carried out at 30°C for 15 min to allow the extension of the RNA primer to proceed for two nucleotides to the position where the first TTP is required. The reaction mixture was then transferred to an Eppendorf tube containing [α-32P]dATP which had been dried by a stream of N2. TTP (0.1 mM, final concentration) was added and chain elongation was allowed to proceed for 10 minutes.

32P-labeled products were collected and analyzed by electrophoresis through an alkaline 1% agarose gel.

Elution of DNA from agarose gel

We have modified the method of Yang et al. (5) to isolate DNA from agarose
gels. First, the alkaline agarose gel was brought to neutral pH by soaking in 500 ml of 100 mM Tris (pH 7.5). The slot containing $^{32}$P-labeled DNA fragments was sliced and then counted. The gel slices of interest were located and then dissolved in 4 volumes of PEP buffer at 37°C for 1 hr. A single 24 mm diameter GF/C filter was placed on a sintered-glass filtration set under slight suction and washed with 3 ml of PEP buffer. The dissolved agarose-DNA solution was passed through the filter at a flow rate of 0.5 ml/min. The GF/C filter was then rinsed with 3 ml of PEP buffer and 30 ml of 100% ethanol and finally air-dried (for about 3 min). DNA was eluted from the filter by soaking it in 0.4 ml of 2 mM Tris (pH 7.4) and 0.2 mM Na$_2$EDTA for 5 min at 25°C. The eluate was collected and elution was repeated twice more using 0.3 ml of the same solution. The DNA sample was concentrated 10-fold under N$_2$ flow and then passed twice through an 1 x 8 cm column of Bio-Gel A 0.5 m. The DNA sample was eluted in 2 mM Tris (pH 7.4), 20 mM NaCl, and 0.2 mM EDTA. After desalting, the DNA was concentrated by ethanol precipitation and resuspended in 20 mM Tris (pH 8.4), 12 mM MgCl$_2$, 12 mM 2-mercaptoethanol, and 200 mM NaCl. The solution was directly used for annealing and restriction cleavage as described below.

The replication intermediates eluted from the agarose gels were annealed to a large excess fd template at 65°C. After 1 hr of incubation an equal volume of restriction endonuclease Taq I (9 units) and distilled water were added and the reaction was allowed to proceed at 65°C. After 2 hr the reaction was terminated by the addition of Na$_2$EDTA (20 mM, final concentration) and NaDodSO$_4$ (0.4%, final concentration). The solution was extracted with phenol and then with chloroform / isoamyl alcohol (24:1, vol/vol). The extracted DNA samples were concentrated by ethanol precipitation, resuspended in 0.12 N NaOH and 4 mM Na$_2$EDTA, and then analyzed by polyacrylamide gel electrophoresis in the presence of urea.

Gel electrophoresis

Alkaline agarose (1%, wt/vol) gel electrophoresis was performed as previously described (4). 7 M urea, polyacrylamide (8% acrylamide; 0.4% bis-acrylamide, wt/vol) gel electrophoresis was carried out as reported by Maxam and Gilbert (6), at 1000 V until the bromphenol blue had migrated about 33 cm. $^{32}$P-labeled fragments were detected by autoradiography using Kodak X-Omat film.
RESULTS

Fine mapping of the secondary structures of fd phage DNA in the region of replication origin

A synthetic heptaribonucleotide, GACCCCC, which is complementary to a single site on the fd DNA genome, primes DNA synthesis by T4 DNA polymerase. As shown in Figure 1, T4 DNA polymerase starts by adding a deoxyribonucleotide at position 1194 (4). The products of the GACCCCC-primed DNA synthesis of fd were detected by autoradiography following sizing of the 32P-labeled DNA strand by electrophoresis through an alkaline 1% agarose gel. After 10 minutes of synthesis a significant fraction of DNA products ran as a single band with a length of about 1960 nucleotides (i.e., fragment 1 in Figure 2). This most prominent pause in DNA synthesis is located near fd nucleotide position 5650. Fragments larger than the unit length of fd DNA genome (i.e., 6408 nucleotides) are the result of a template-primed reaction catalyzed on a fraction of linear fd DNA molecules (4) and should be ignored.

The DNA fragment labeled as 1 in Figure 2 was eluted from the alkaline agarose gel as described in MATERIALS AND METHODS. After the isolated DNA fragments had been annealed to excess fd DNA template at 65°C for 1 hr, the resulting

![Diagram of fd bacteriophage DNA](image)

**Fig. 1.** A schematic representation of fd bacteriophage DNA. This DNA genome is 6408 nucleotides long and the single Hin II cleavage site is arbitrarily chosen as the origin of numbering (2). The synthetic heptaribonucleotide, GACCCCC, is complementary to the nucleotides at positions 1195 to 1201, inclusively, with DNA synthesis occurring in the direction of the dashed arrow (→→→). Positions 5650 and 3350 are the strong kinetic pause sites to DNA polymerization previously described (4).
Fig. 2. The products of GACCCCC-primed DNA synthesis of fd by T4 DNA polymerase after 10 min of reaction. Sizing of the new DNA strands was carried out by electrophoresis through an alkaline agarose gel (1%, wt/vol). Fragments designated as 1 and 2 correspond to newly synthesized DNA molecules which have paused near fd sequence positions 5650 and 3350, respectively. The heterogeneous fragments which are longer than 6408 bases are the products of a template-primed reaction catalyzed on a fraction of linear fd DNA molecules present (4) and should be ignored. The numbers on the right side of this autoradiogram denote the length of fd markers obtained by treating the long 'rolling circle' DNA products made on an fd DNA template (4, 7) with restriction endonuclease Hae III.

duplex DNA was incubated with restriction endonuclease Taq I at 65°C for 2 hr. The ^32P-labeled fragments were then detected by autoradiography following sizing of the DNA products by electrophoresis through a urea polyacrylamide gel.

Taq I digestion of the hybridized duplex formed by renaturation of replicated products paused near fd nucleotide 5650 does not only produce restriction fragments labeled C, D, J, and perhaps H in Figure 3, but also small fragments from the region between the starting point of GACCCCC-primed DNA synthesis and the Taq I site at sequence position 1129 (i.e., producing fragments of between 65 and 72 nucleotides long depending on the extent of alkaline hydrolysis of the RNA primer GACCCCC during the alkaline agarose gel electrophoresis and during the alkaline treatment of the Taq I digests prior to layering on polyacrylamide gel; see MATERIALS AND METHODS). More importantly, the DNA fragments corresponding to the sequence between the 3'-OH termini of the replication intermediates paused near the fd nucleotide position 5650 and the Taq I site
Fig. 3. The physical and genetic maps of phage fd. The outer two circles represent the physical map which shows the cleavage sites of restriction endonuclease Taq I. The numbering for this restriction map is the same as that illustrated in Figure 1. Note that the starting point of GACCCCC-primed DNA synthesis of fd (i.e., position 1194) is located within restriction fragment G. The third circle shows the genetic map. Genes are indicated by Roman numerals. IG refers to the intergenic region between genes II and IV. The inner circle is divided into ten equal physical map units with the Hin II cleavage site as the reference point.

at position 5686 or position 6043 would also be revealed. By analyzing the length of the released DNA fragments on a urea polyacrylamide gel, the precise locations of the 3'-OH termini of the stalled replication intermediates have been determined.

The reaction products of Taq I digestion are shown in lane 1 of Figures 4a and 4b. As expected, we observed four restriction fragments, C, D, H, and J (see Figure 3 for the designation of restriction fragments of Taq I reaction) and two fragments of 65 and 72 nucleotides long, respectively (i.e., fragments X and Y). The presence of only 65- and 72-nucleotide fragments suggests that either the alkaline hydrolysis of the RNA primer GACCCCC is highly cooperative or that the sample is somehow not homogeneously exposed to NaOH during alkaline
Fig. 4. Restriction endonuclease Taq I cleaves the replicative form DNA of fd stalled near nucleotide position 5650 into discrete DNA fragments. The products of Taq I reaction were analyzed on an 8% polyacrylamide gel in the presence of 7 M urea. DNA samples were suspended in 6 μl of 0.12 N NaOH and 4 mM Na₂EDTA before layering on the gel (MATERIALS AND METHODS). Electrophoresis was carried out at 1000 V for 3 hr (a) or for 90 min (b). The results are shown in lane 1. In addition to the expected restriction fragments C, D, H and J (see Fig. 3 for the designation of restriction fragments), fragments representing the region between the starting point of GACCCCC-primed DNA synthesis and the Taq I site at position 5629 (Figs. 1 and 3; also see text) were found (i.e., fragments X and Y in lane 1). Furthermore, two additional fragments of 13 and 112 nucleotides were produced, as indicated by arrows (lane 1). We interpret these two additional fragments to correspond to the sequence between the 3'-OH termini of the stalled replication intermediates and the Taq I site at position 5686 (see text). Lane 2 shows fd DNA markers obtained by treating the long "rolling circle" DNA products made on an fd DNA template (4,7) with restriction endonuclease Hpa II. The numbers on the right side of lane 2 denote the length of these markers in number of nucleotides.
treatment (see above). In addition to these expected fragments, two additional
fragments of 13 and 112 nucleotides long were obtained, as indicated by arrows
in lane 1 of Figures 4a and 4b. We thus conclude that T4 DNA polymerase must
have paused at nucleotide position 5575 to produce the 112-nucleotide fragment
and at position 5674 to yield the 13-nucleotide fragment. Inspection of the fd
nucleotide sequence reveals that nucleotides 5575 and 5674 are located within
the stem of the two very stable hairpins of fd, as shown in Figure 5.

In addition to the presence of discrete DNA bands in lane 1 of Figures
4a and 4b, a significant amount of radioactivity remained at the origin of
the gel and a smaller fraction of radioactivity migrated as a fragment of
at least 2500 nucleotides. This result could be due to a partial reassocia-

---

Fig. 5. Potential secondary structures at fd nucleotide positions 5575 and
5674. The horizontal arrows (—0) indicate the positions at which T4 DNA
polymerase paused with a 20 min average delay (7). The stability of structural
features A and B is calculated using the rules developed by Tinoco et al.
(9,10) and shows a $\Delta G = -62.6$ Kcal/mol associated with feature A and a $\Delta G =
-31.9$ Kcal/mol associated with feature B. The GACCCCC-primed DNA synthesis
of fd occurred in the direction of the dashed arrow (---).
tion between the $^{32}$P-labeled (-) strand DNA of fd and unlabeled full length DNA template strand of fd (which is presented in large excess in the solution) early in the course of electrophoresis. To test this possibility, an otherwise identical control experiment was carried out in which the annealing procedure was omitted and the purified (-) strand fd DNA was used directly as the substrate in Taq I reaction (data not shown). The result of this control experiment proved that the holdup at the top of the gel was an artifact caused by the renaturation of complementary strands.

Effect of T4 DNA helix-destabilizing protein

We have previously shown that the addition of T4 DNA helix-destabilizing protein (i.e., gene 32 protein) stimulates the rate of T4 DNA polymerase movement in this system, decreasing the time spent in getting through the helical region located near fd nucleotide position 5650 by about 7-fold (4,7). Furthermore, nearly all of the weak pauses disappear when gene 32 protein is added. Thus, it is of interest to examine the effect of gene 32 protein on the nature of the two pause sites observed above. We have found that T4 DNA polymerase still pauses at the same sites with 32 protein present at 720 μg/ml (data not shown), but for much shorter times. This observation, together with our previous studies (4,7) suggests that gene 32 protein may help to open the hairpin helices just ahead of the two pause sites located at fd sequence positions 5575 and 5674.

Comparison of E. coli DNA polymerase I with T4 DNA polymerase

To determine whether or not the structural features recognized by T4 DNA polymerase were unique to this enzyme, an otherwise identical experiment in which E. coli DNA polymerase I was substituted for T4 DNA polymerase was carried out. The products synthesized by E. coli DNA polymerase I at enzyme concentrations less than or equivalent to those utilized for T4 DNA polymerase form a smear on the alkaline 1% agarose gel (data not shown), suggesting that this polymerase is insensitive to the kinetic barriers seen by T4 DNA polymerase. Our results are in agreement with the recent work of Challberg and Englund (8). Thus, T4 DNA polymerase may be typical of a class of DNA polymerases that recognize secondary structure.

DISCUSSION

In this study, a synthetic RNA primer, GACCCCG, was used to prime the fd DNA template by hybridization at nucleotide positions 1195-1201. As the T4 DNA polymerase moved along the fd DNA template, local regions exhibiting secondary structure functioned as roadblocks since the rate of polymerase
movement through helical region is orders of magnitude slower than along less structured region of the template (4,7). The kinetic barriers to continuous chain extension were mapped by sizing the paused replication intermediates on an alkaline agarose gel. DNA synthesis of fd by T4 DNA polymerase was delayed for a characteristic relaxation time of about 20 min at the region near fd nucleotide position 5650 (7), producing the fragment labeled in Figure 2 as 1 (about 1960 nucleotides long). Restriction fragments were produced by Taq I cleavage of the 32P-labeled replicated strand which had been hybridized back to unlabeled fd template. The precise location of the 3'-OH termini of fragments 1 were analyzed by sizing on a urea polyacrylamide gel. The fine mapping procedure has located strong kinetic pause sites at fd nucleotide positions 5575 and 5674. These sites correspond to the stems of two very stable hairpins postulated from the primary sequence of fd (Fig. 5). Interestingly, T4 DNA polymerase stalls at the same sites with a 3 min average delay when T4 DNA helix destabilizing protein (i.e., gene 32 protein) is present in the solution (data not shown). Inspection of Fig. 5 indicates that T4 DNA polymerase has disrupted 3 and 5 base pairs, respectively, prior to pausing at fd nucleotide positions 5575 and 5674. Since there is some uncertainty in sizing the 3'end fragments released after Taq I reaction (13 ± 2 and 5 ± 1 base pairs). Our results are consistent with the recent work of Challberg and Englund who showed that vaccinia virus DNA polymerase pauses within the stem of a hairpin structure on a single-stranded DNA template strand after disrupting 3 stem base pairs (8). It is of interest to establish how T4 DNA polymerase proceeds through the presumed regions of template structure. Since T4 DNA polymerase halts at one single point in the hairpin as it proceeds further into hairpins A and B in Fig. 5 (Figs 'a and 'b), it is possible that the hairpins occasionally melt and their sequences then are quickly copied by the polymerase. Alternatively, the polymerase occasionally goes through the kinetic barriers by bypassing the sequence in the hairpins.

The nucleotide sequence near fd sequence position 5650 is the site where DNA replication of fd begins (1). Shen and Hearst have previously shown that a compact structure exists in this region of the fd DNA genome (11). More recently, Schaller proposed that there are at least four potential structures at the origin of DNA replication of fd (3). In this communication, we have detected two of the four postulated hairpin helices. The stability of hairpins A and B in Figure 5 reveals that they are very stable hairpins in the fd DNA genome (4). There are two possibilities...
to account for our failure to detect the other two potential hairpin structures proposed by Schaller. It is possible that the presumed DNA fragments were overlooked in our experiments. As described in MATERIALS AND METHODS, the sample track of agarose gel was sliced and then counted. Typically, gel slices corresponding to the main portion of the peak were pooled and used for further treatment. By analyzing only a limited number of gel slices, it is conceivable that some replication intermediates paused near fd sequence position 5650 have been overlooked. Alternatively, the other two hairpins postulated by Schaller may not exist as stable structure under our reaction condition.

Schaller et al. (12) have isolated a unique segment of fd viral DNA that is bound to E. coli RNA polymerase in the presence of DNA binding protein I from an initiation complex of the in vitro reaction. This DNA fragment is also known to contain the origin of minus strand DNA replication of fd (1). In vitro studies of the initiation reaction with fd viral DNA as a template have indicated that a short RNA primer is transcribed from the DNA at the minus strand origin and then elongated by DNA polymerase (13). Inspection of the fd nucleotide sequence reveals that the DNA segment containing hairpin B in Fig. 5 is part of the RNA polymerase-binding site. It is reasonable to speculate that this hairpin helix provides a signal to RNA polymerase for the recognition of the site where the short RNA primer starting at nucleotide position 5757 is synthesized (13). Despite the fact that hairpin B is nearly 80 nucleotides away from the primer position, there is some precedence for such a model. In the case of the single-stranded DNA phage, φK, the binding site of E. coli dnaG protein, which initiates DNA replication by synthesizing primer oligonucleotides, is formed by tertiary interactions between DNA in the region of the upstream hairpin (as far away as 115 bases) and DNA in the vicinity of the starting site of primer synthesis (14).

The biological function of hairpin A, which is the most stable secondary structure in fd DNA, has yet to be established. Ikoku and Hearst (15) have postulated that it is essential for filamentous phage morphogenesis. Very recently, Griffith et al. (16) have shown that exposure to a chloroform-water interface initiates a 20-fold contraction of the filamentous bacteriophage M13 into a hollow protein sphere. Furthermore, two-thirds of the DNA is extruded through a hole of the protein sphere. The portion of DNA remaining inside the protein shell centers about the M13 intergenic region, which contains several large hairpins and the origin of DNA replication. The pro-
tein sphere, which anchors one-third of M13 DNA which remains inside the spheroid, is possibly held through some specific attachment at the intergenic region of M13 (16). Since M13 is almost completely homologous in the intergenic region with fd having the same predicted hairpin structures (17, 18), we favor this attachment site to be hairpin A (15). The attachment may represent the presence of a specific protein complex attaching M13 DNA at hairpin A to the protein shell. Furthermore, hairpin A is also maintained in another related phage f1 (19). If the chloroform-water interface mimics the properties of the bacterial membrane, then a similar change in structure might occur during the entry and uncoating steps of a normal filamentous phage (i.e., fd, f1, or M13) infection.

Alternatively, hairpin A in Fig. 5 may play an important role in the attachment of fd phage to the host receptor and in the penetration of fd DNA into the host cell by orienting the viral DNA during the infection process.

ACKNOWLEDGMENTS

This research was supported in part by American Cancer Society Grant NP 185. We thank Chung-Cheng Liu and Bruce Alberts for providing us with T4 DNA polymerase. Kenneth N. Kreuzer's advice and recipes on the elution of DNA from agarose are especially appreciated. C.-C. Huang is indebted to Howard Gamper for helpful discussion and to Stephen Quong for preparing the manuscript.

FOOTNOTE

1 The abbreviations used are: PEP buffer, 6 M NaClO₄, 10 mM Na₂EDTA, and 50 mM potassium phosphate, pH 7; NaDodSO₄, sodium dodecyl sulfate.

REFERENCES
