Studies on DNA Replication in the Bacteriophage T4 In Vitro System

B.M. ALBERTS, J. BARRY, P. BEDINGER, T. FORMOSA, C.V. JONGENEEL, AND K.N. KREUZER
Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143

In vitro systems reconstituted from the purified components of a DNA replication apparatus have provided us with much of our knowledge about the mechanistic aspects of DNA replication. The large DNA bacteriophages are especially suited for this approach, since they code for most of the proteins required for their own replication. This makes the task of defining essential replication genes and isolating their products much more straightforward than in bacteria or eukaryotic cells.

Extensive genetic analysis of bacteriophage T4 has identified 11 genes whose products appear to be directly involved in the formation and subsequent movement of DNA replication forks: genes *32, 39, 41, 43, 44, 45, 52, 60, 61, 62,* and *dda* (Epstein et al. 1964; Warner and Hobbs 1967; Curtis and Alberts 1976; McCarthy et al. 1976; P. Gauss et al., pers. comm.). The proteins specified by these genes have been isolated and purified to near homogeneity in our laboratory (Bittner et al. 1979; L.F. Liu et al. 1979; Morris et al. 1979a, b; R. L. Burke and C.V. Jongeneel, unpubl.). Table 1 summarizes the basic properties and functions of these T4 replication proteins as we understand them at the present time.

The viral DNA polymerase is the product of gene *43*. When present by itself, it is capable of elongating preexisting primers on single-stranded DNA templates (Aposhian and Kornberg 1962; Goulian et al. 1968). It also has an intrinsic 3′→5′ proofreading exonuclease activity (Brutlag and Kornberg 1972). Regions of secondary structure in the single-stranded DNA template act as barriers to continued polymerization by the enzyme, causing it to pause (Challberg and Englund 1979; Huang and Hearst 1980). The addition of stoichiometric amounts of gene-*32* protein (the T4 helix-destabilizing protein) increases the rate of polymerase movement on such templates, presumably by melting the secondary structure of the template strand and presenting it to the polymerase in the proper configuration (Huberman et al. 1971; Huang et al. 1981). The products of genes *44* and *62* copurify as a tight complex (hereafter referred to as the *44/62* protein). The gene-*44/62* and gene-*45* proteins are collectively known as the polymerase accessory proteins, and they exhibit a DNA-dependent ATPase activity (Piperno and Alberts 1978). When these proteins are added to the polymerase, they increase its rate and processivity in an ATP-dependent fashion (Alberts et al. 1975; Huang et al. 1981). On a single-stranded DNA template, the stimulatory effects of the gene-*32* protein and the polymerase accessory proteins are multiplicative, and both types of proteins are required to enable the polymerase to use a double-stranded DNA template efficiently (Nossal and Peterlin 1979; Alberts et al. 1980).

On a double-stranded template, the mixture of gene-*32*, -*43*, -*44/62*, and -*45* proteins constitutes a minimal or "core" replication system (Fig. 1). DNA synthesis begins by strand displacement, starting from a

Table 1. Properties of Bacteriophage T4 Replication Proteins

Type of protein	T4 gene	Molecular weight ($\times 10^{-3}$)	Current purity (%)	Activities[a]
DNA polymerase	*43*	110	99	5′→3′ polymerase; 3′→5′ exonuclease
Helix-destabilizing protein	*32*	34.5	99	cooperative binding to SS DNA
Polymerase accessory proteins	*44/62;45*	34 + 20; 24.5	99;95	SS DNA-termini-dependent ATPase, dATPase
RNA-priming proteins	*41*	58	90	long SS DNA-dependent GTPase, ATPase
	61	40	70	binds DNA
Type-II topoisomerase	*39/52/60*	56 + 46 + 17	99	DS DNA strand passage; DS DNA-dependent ATPase
DNA helicase	*dda*	47	95	SS DNA-dependent ATPase; DNA unwinding

[a]SS indicates single-stranded; DS indicates double-stranded.

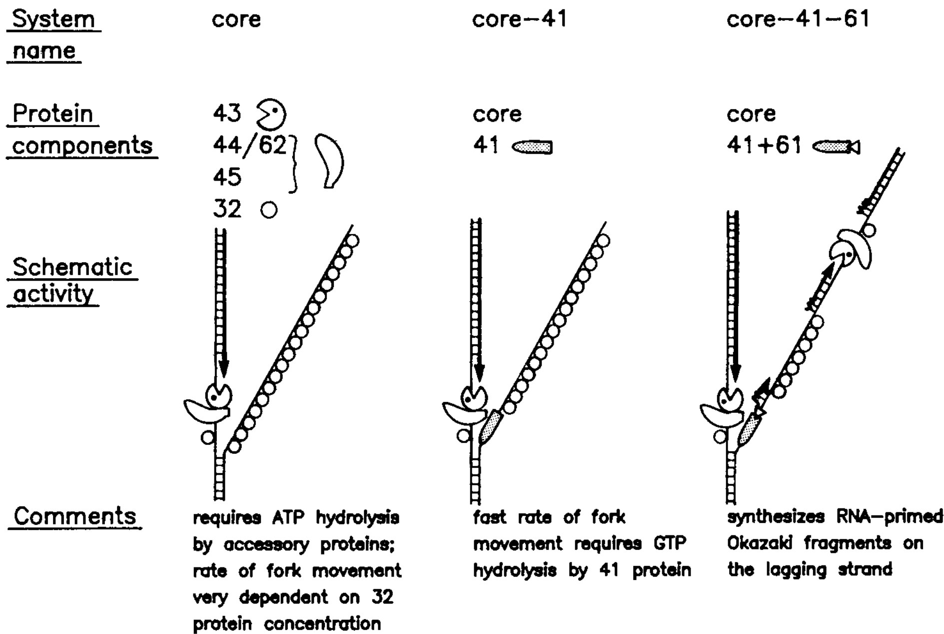

Figure 1. Outline of the properties of three T4 DNA replication systems of increasing complexity reconstituted from purified T4-encoded proteins. As indicated, the core replication system contains five different T4 gene products; this is the simplest system that can replicate double-stranded DNA templates efficiently.

nick on a double-stranded DNA template (C.-C. Liu et al. 1979; Nossal and Peterlin 1979). Omission of one of these five proteins, or of ATP, obliterates the strand-displacement reaction and almost no synthesis is detected. The rate of fork movement observed with the core replication system is highly dependent on the concentration of gene-*32* protein, and more than 0.5 mg/ml is required to approach the fork rates observed in vivo. To achieve such rates on a double-stranded template at lower concentrations of gene-*32* protein, a sixth protein (the product of gene *41*) is required (Morris et al. 1979a; Alberts et al. 1980). This "core-41" replication system provides a reasonable in vitro model for leading-strand synthesis at a replication fork (Fig. 1). The gene-*41* protein has a DNA-dependent GTPase and ATPase activity, and it uses the energy of nucleotide hydrolysis to move rapidly along a single-stranded DNA molecule (Liu and Alberts 1981a). At a replication fork, this protein appears to act as a DNA helicase and, as expected, its enhancement of the fork rate requires that it continuously hydrolyze nucleoside triphosphates.

In conjunction with the gene-*61* protein, the gene-*41* protein also has a primase function, producing the pentaribonucleotides that serve as primers for the synthesis of Okazaki fragments on the lagging strand (Nossal 1980; Liu and Alberts 1980, 1981b). A mixture of the seven proteins described so far constitutes the "core-41-61" system, whose activity is schematically depicted in Figure 1. This system produces replication forks that closely resemble the replication forks that function in vivo (Hibner and Alberts 1980; Liu and Alberts 1980; Sinha et al. 1980; Sinha and Haimes 1981): (1) The template DNA is replicated semiconservatively at about the in vivo rate and with high fidelity. (2) Replication forks are very stable once they are assembled, proceeding for more than 50,000 nucleotides. (3) Okazaki fragments with an average length of about 2000 nucleotides are synthesized on the lagging strand (see below). (4) The RNA primers that start the Okazaki fragments have the same length and sequence (pppApCpNpNpN) as in vivo.

Despite the above properties, the core-41-61 replication system is incomplete. It cannot remove the RNA primers from the ends of Okazaki fragments or seal adjacent fragments together to create a continuous chain on the lagging strand, inasmuch as RNase H and DNA ligase are missing. Moreover, it is unable to initiate DNA replication by forming a replication bubble at a defined origin on a double-stranded T4 DNA molecule.

Recent data have suggested than an eighth protein, a T4-encoded DNA-dependent ATPase and DNA helicase (the product of the *dda* gene) (Purkey and Ebisuzaki 1977; Krell et al. 1979), may also be associated with the replication fork. The *dda*-gene protein binds tightly to the gene-*32* protein (T. Formosa et al., in prep.), and when added to the core replication system, it strongly stimulates strand-displacement synthesis at low concentrations of gene-*32* protein (C.V. Jongeneel, unpubl.). Although the *dda*-gene protein is not essential for virus growth in a wild-type *Escherichia coli* host, it is required for the growth in an *E. coli optA* mutant strain (P. Gauss et al., pers comm.). Therefore, it would seem that its function is essential but is bypassed in vivo by a functionally equivalent host protein.

The three remaining T4 replication genes, *39, 52,* and *60,* code for the three subunits of a type-II DNA topoisomerase (also denoted the *39/52/60* protein com-

plex; L.F. Liu et al. 1979; Stetler et al. 1979). Mutations in these "DNA-delay" genes seem to affect the initiation of DNA replication rather than the rate of fork movement (McCarthy et al. 1976), and it has been suggested that the topoisomerase is involved in the localized melting of the double helix thought to be necessary for the formation of replication bubbles on an intact double-stranded DNA template (L.F. Liu et al. 1979). However, in the cell, two different pathways (mechanisms) seem to be used for the initiation of new T4 replication forks (Mosig et al. 1979, 1981; Luder and Mosig 1982). The primary mechanism, which is used only early in infection, generates replication bubbles at one or a few specific origins on the parental DNA molecule and requires RNA synthesis by *E. coli* RNA polymerase (Snyder and Montgomery 1974; Luder and Mosig 1982). One of these primary origins has been mapped to the region of the T4 genome containing genes *56, 61,* and *41* (C.F. Morris; G. Mosig; both pers. comm.), and it is still unclear whether there are others. Later in infection, initiation of new forks becomes independent of the host RNA polymerase (Luder and Mosig 1982). This secondary mechanism for fork initiation requires an active T4 genetic recombination system, and it has been suggested that recombination intermediates directly prime the initiation of secondary replication forks (Mosig et al. 1981; Luder and Mosig 1982). The mode of T4 topoisomerase involvement in either pathway of replication initiation is still unclear.

The long-term goal of our group is to obtain a detailed, three-dimensional picture of a functional replication fork. In this paper we present the results of recent in vitro studies designed to answer the following questions: (1) As discussed above, the core-41-61 replication system, which performs both leading- and lagging-strand syntheses, produces forks that move very rapidly and processively, in which the leading-strand DNA polymerase molecule seems to remain permanently bound. Does the DNA polymerase molecule that synthesizes each Okazaki fragment on the lagging strand also remain permanently bound, being recycled for multiple rounds of Okazaki-fragment synthesis? (2) What are the direct protein:protein interactions involved in building and stabilizing the replication complex? (3) Since replication and transcription are thought to coexist on the same DNA molecule, what happens when a replication fork encounters an RNA polymerase molecule bound to its DNA template? (4) Does the T4 DNA topoisomerase recognize T4 DNA with any specificity and, if so, how might this relate to its function in the initiation of replication forks?

Evidence for a Recycling of the DNA Polymerase Molecule That Synthesizes Each Okazaki Fragment on the Lagging Strand

Recently, we described the T4 core-41-61 replication system in detail and proposed the view of the T4 DNA replication fork illustrated in the rightmost panel of Figure 1 (Alberts et al. 1980). This fc ' contains two DNA polymerase holoenzyme complexes, each consisting of a single subunit of the gene-*43* protein complexed to a multimer of the gene-*44/62* and gene-*45* proteins. One of these units synthesizes the DNA made on the leading strand and the other synthesizes the DNA made on the lagging strand.

The view of the replication fork presented in Figure 1 implies that the DNA polymerase molecule that synthesizes each Okazaki fragment functions only once, starting the 5′ end of a new DNA fragment when it collides with an RNA primer and dissociating from its template about 4 seconds later, as it finishes the 3′ end of the fragment and encounters the 5′ end of the previous Okazaki fragment. In contrast, the DNA polymerase molecule on the leading strand remains with the fork for the synthesis of at least 50,000 nucleotides (Sinha et al. 1980; J. Barry, unpubl.). Therefore, if the model is correct, the DNA synthesis on the lagging strand should be uniquely sensitive to extreme dilutions of the DNA polymerase, since such dilutions should delay the start of each Okazaki fragment by making the collision of a free DNA polymerase molecule with an RNA primer very unlikely.

Alternatively, the DNA in the replication fork may be folded in a way that brings the start site for each succeeding Okazaki fragment in close juxtaposition to the place where the previous Okazaki fragment will end. In this case, the DNA polymerase on the lagging strand could be "recycled," thereby making DNA synthesis on the lagging strand insensitive to the free concentration of the DNA polymerase, as is true for DNA synthesis on the leading strand.

In Figure 2, the contrasting expectations for a recycling and a nonrecycling mechanism are illustrated diagrammatically for DNA synthesis on a double-

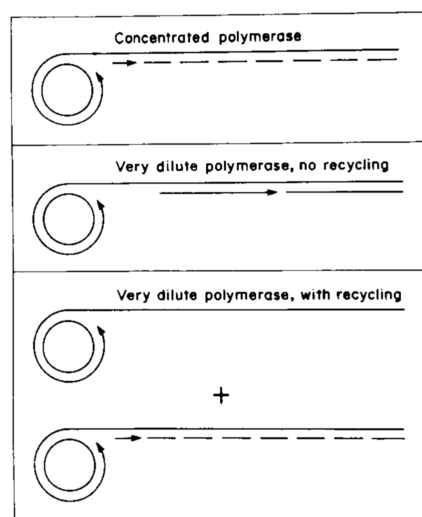

Figure 2. Schematic illustration of the pattern of DNA synthesis expected on the lagging strand of a replication fork at very low DNA polymerase concentrations, with and without recycling of the lagging-strand DNA polymerase molecule. For comparison, the top panel shows the result expected at high polymerase concentrations, which is the same whether or not there is recycling.

stranded, circular DNA template. Without recycling, polymerase dilution should cause a large increase in the length of the single-strand connection between the double-stranded tail that forms on such templates and the circle, and it should cause a corresponding increase in the average length of each Okazaki fragment. With recycling, there should be no such change.

To determine which model is correct, we have carried out a series of in vitro replication reactions with the core-41-61 replication system, in which the concentration of DNA polymerase has been varied with all other components in excess. The amount of DNA synthesis obtained decreases in direct proportion to the DNA polymerase concentration below about 0.2 μg/ml, making measurements below 0.02 μg/ml difficult. However, a dilution of the polymerase to this level should be more than sufficient to determine whether the lagging-strand DNA polymerase molecule is being recycled; if we consider the measured association rate constant of the T4 DNA polymerase for a primer-template end (6×10^6 $M^{-1}sec^{-1}$) (Newport 1980), new Okazaki fragments should not begin more often than once every 2 minutes at the lowest polymerase concentration tested. Since the leading strand moves at a rate of several hundred nucleotides per second irrespective of the polymerase concentration (Alberts et al. 1980; J. Barry and B.M. Alberts, in prep.), if there is no recycling, the average length of an Okazaki fragment should increase from the commonly observed value of 1700 nucleotides to more than 10,000 nucleotides in these experiments.

The method that we have used to measure the lengths of the Okazaki fragments is illustrated in Figure 3, where the sizes of the DNA products of the in vitro DNA synthesis reaction catalyzed by the core-41 and the core-41-61 systems have been compared by gel electrophoresis in alkali. Although the Okazaki fragments produced are quite heterogeneous in size (ranging from 300 to 3000 nucleotides in length), they are cleanly separated on the gel from the much larger DNA chains that are synthesized on the leading strand. As expected, no Okazaki fragments are made in the absence of the gene-*61* protein and rNTPs.

Table 2 presents the results of one of several different experiments that we have performed to discriminate between the two possibilities illustrated in Figure 2. In this experiment, alkaline gel electrophoresis was used to compare the sizes of the Okazaki fragments synthesized at four different DNA polymerase concentrations after a brief period of DNA synthesis. As indicated, the size of these fragments was essentially unaffected by the polymerase dilution. Similar samples were examined by electron microscopy to determine the length of the single-strand connection that joins the double-stranded DNA tail on the lagging strand to the rest of the fork. The median length of this connection remained constant at about 1000 nucleotides, independent of the polymerase dilution (J. Barry and B.M. Alberts, in prep.). Both of these results are explicable only if the DNA polymerase molecule on the lagging strand is recycled for multiple rounds of Okazaki-fragment synthesis, thereby remaining bound to the fork at all times, as in the case of the leading-strand DNA polymerase molecule.

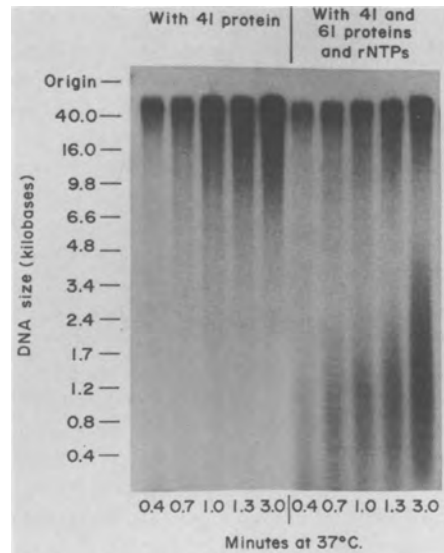

Figure 3. Analysis of Okazaki-fragment synthesis by alkaline gel electrophoresis. DNA was synthesized in the standard reaction mixture containing 160 μM each of dATP, dCTP, and dGTP, 40 μM dTTP, 0.5 mM dithiothreitol (DTT), 66 mM potassium acetate, 33 mM Tris-acetate (pH 7.8), 10 mM magnesium acetate, 250 μM ATP, 250 μM GTP, 200 μg/ml of nuclease-free albumin, 20 μg/ml of gene-*44/62* protein, 15 μg/ml of gene-*45* protein, 3.2 μg/ml of gene-*43* protein (DNA polymerase), 30 μg/ml of gene-*41* protein, 1 mCi/ml of [α-^{32}P]dTTP, and where indicated, 200 μM each of UTP and CTP. The template used was 10 μg/ml of cytosine-containing T4 DNA (C T4 DNA), which is a long double-stranded molecule with a few random nicks that act as start sites for replication forks. Reactions were started after a 4-min preincubation at 37°C by adding gene-*32* protein to a final concentration of 20 μg/ml and, where indicated, gene-*61* protein to a final concentration of 0.4 μg/ml. At the indicated times, reactions were stopped by mixing 50-μl aliquots of reaction mixture with 20 μl of 90 mM Na_3EDTA, 50 mM NaOH, and 0.36% SDS; phenol extracting with 70 μl of neutralized phenol; and purifying the DNA products by rapid filtration through Bio-Gel P-30 (100–200 mesh). DNA was then analyzed by electrophoresis on horizontal 0.4% alkaline agarose gels run in 30 mM NaOH and 2 mM Na_3EDTA with a potential gradient of 1 V/cm. After electrophoresis, the gels were dried and autoradiographed.

The structure of the replication apparatus proposed to account for these results is schematically illustrated in Figure 4. Because of the protein-protein interactions described below (see Table 4), as well as kinetic results (J. Barry and B.M. Alberts, in prep.), we believe that the two DNA polymerase complexes—one on the leading strand and one on the lagging strand—are held together in the indicated manner. This leads to the "trombone model" for DNA synthesis (Fig. 4), in which the synthesis of each Okazaki fragment involves the enlargement of a large loop of DNA, half of which is single-stranded and half of which is double-stranded. The crucial points in the proposed cycle are the "termination" and "restart" steps, which involve movements of the DNA on the lagging side of the fork around

Table 2. Effect of Extreme DNA Polymerase Dilutions on DNA Synthesis and Okazaki Fragment Size

DNA polymerase concentration (μg/ml)	Average Okazaki fragment size (nucleotides at 2 min)	Total synthesis (pmoles dNTP/ml at 2 min)
3.2	1700	430
0.8	1700	330
0.2	1700	200
0.04	1900	50

DNA was synthesized as described in Fig. 3, with gene-*61* protein, CTP, and UTP included in all reactions. In addition, the level of gene-*43* protein (DNA polymerase) was varied as shown. After analysis by alkaline gel electrophoresis and autoradiography, the films were scanned with a densitometer to determine the weight-average Okazaki fragment size.

a fixed lagging-strand DNA polymerase molecule, as indicated (Fig. 4). A major implication of the model is that the replication apparatus itself only lays down *unsealed* Okazaki fragments; the subsequent fragment-ligation reactions occur later and must involve a separate DNA polymerase molecule that fills in the gap that will be created by RNA primer removal.

An especially interesting aspect of the model in Figure 4 is that the length of single-stranded DNA that accumulates in the loop (which will determine the size of the *next* Okazaki fragment) is equal to the length of the *previously synthesized* Okazaki fragment. Therefore, the lengths of the Okazaki fragments made on an individual lagging strand should be similar to each other and approximately equal to the length of the first Okazaki fragment that was made on that strand. This proposition remains to be tested.

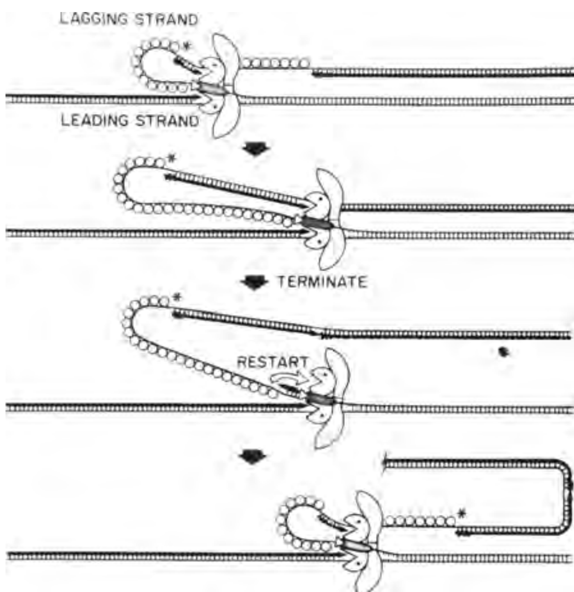

Figure 4. Trombone model for DNA synthesis at a replication fork. This mode of synthesis allows the same DNA polymerase molecule to be repeatedly recycled to synthesize all of the Okazaki fragments on the lagging strand. See Fig. 1 for the different symbols used to represent the replication proteins; an unfolded Y-shape form of this fork structure is shown in Fig. 1.

Experimental evidence similar to that just described for the DNA polymerase suggests that both the gene-*41* protein and the gene-*61* protein remain bound to the fork for prolonged periods (J. Barry, unpubl.), and it may be that all of the protein components at the fork recycle without ever leaving the complex. In any case, the picture that emerges is one of a large replication machine, whose component parts are protein molecules that move relative to each other without disassembling from the complex.

The model for fork movement shown in Figure 4 has several important implications with regard to the mechanism by which replication forks are initiated by replication bubble formation at unique origins: (1) Because the fork structure being created is a complex one, the rate-limiting step in fork initiation may involve the precise assembly of a large replication protein aggregate, rather than reflecting the difficulty of creating a localized helix opening at the origin. (2) The fork-propagation mechanism seems designed to make Okazaki fragments only after the first such fragment has been synthesized by some other mechanism. This other mechanism should be unique to the initiation process, and how it operates is unknown. If it measures the length of the first fragment by reference to start and stop sites that are encoded in the DNA sequence, a functional T4 replication origin would need to span about 2000 bp of DNA. (3) The model implies that the two "partner" replication forks that diverge from a bidirectional origin are initiated separately, inasmuch as the DNA polymerase molecule that synthesizes the Okazaki fragment for the first of the two forks formed will continue to function on that lagging strand. Therefore, this polymerase cannot be used as a leading-strand DNA polymerase molecule for a second partner fork, as often assumed.

Protein-Protein Interactions between T4 Replication Proteins Detected by Affinity Chromatography

Specific protein-protein interactions must be involved in the assembly and regulation of the large machine that constitutes the T4 DNA replication apparatus. However, these interactions turn out to be relatively weak ones, and they have therefore been difficult to detect by conventional means. The products of genes *43* (DNA polymerase) and *32* (helix-destabilizing protein) have been shown to form a complex through interactions detected by sucrose gradient centrifugation (Huberman et al. 1971), and similar experiments have shown that the gene-*32* protein also interacts with the gene-*61* protein (primase component) (R.L. Burke et al., in prep.). Several attempts to detect other complexes by this technique have not been successful, presumably because the binding is too weak.

As an alternative approach, we have used the technique of affinity chromatography, which is able to detect even relatively weak protein-protein interactions when the concentration of the protein immobilized on

the column matrix is high (~ 1 mg/ml). The products of genes *43* (T4 DNA polymerase), *45* and *44/62* (polymerase accessory proteins), *41* (helicase-primase), and *32* (helix-destabilizing protein) have each been covalently attached to separate agarose matrices. Each matrix was then exposed to a cleared lysate of T4-infected *E. coli* cells that contained radioactive T4 early proteins; after being washed extensively, the columns were eluted by increasing the salt concentration. As shown by the elution profiles in Figure 5, each column bound more radioactive protein than control columns. Except for the column that contained immobilized gene-*32* protein (data not shown), the largest differences between experimental and control chromatographs are found in the 50 mM NaCl wash, presumably indicating that fairly weak interactions are being detected. A control column containing a plain agarose matrix bound the same amount of protein as a column containing immobilized albumin (Fig. 5), implying that the increased binding detected on each of the T4 replication protein columns is of biological significance.

Although the columns shown in Figure 5 bind only 1–6% of the total applied radioactivity in a specific manner (i.e., as protein not bound to a control albumin column), the gene-*32*-protein columns bind 8–12% of the applied radioactivity, consistent with the wide-ranging importance of this protein in DNA-mediated processes (Mosig et al. 1979). The gene-*32*-protein column is also unusual in binding many specific proteins tightly, with some of the associations being disrupted only by salt concentrations greater than 1 M NaCl (T. Formosa et al., in prep.).

Various fractions of eluted radioactivity have been analyzed by SDS-polyacrylamide gel electrophoresis in order to identify the radioactive T4 proteins eluting from the replication protein columns. These proteins have been further characterized by two-dimensional polyacrylamide gel electrophoresis, as shown for proteins binding to the T4 DNA polymerase column in Figure 6. In addition, interactions of host-encoded proteins with the immobilized T4 replication proteins were cataloged by comparing the eluted proteins detected by autoradiography (T4-encoded proteins only) with the eluted proteins detected by a sensitive silver-staining technique (both T4- and host-encoded proteins) (Merril et al. 1981). We have thereby shown that the *E. coli* RNA polymerase binds to gene-*45* protein (see also Ratner 1974) and that several unidentified, low-molecular-weight host proteins bind specifically to the various protein columns (data not shown).

The results of our studies are summarized in Table 3, which lists the large number of different proteins with DNA-related functions that bind to a gene-*32*-protein column, and in Table 4, which lists the characterized proteins that bind specifically to the other replication protein columns. It should be noted that the catalog of identified proteins is incomplete and that several specific protein-protein interactions have been detected that do not appear in these two tables. These are mostly T4-encoded proteins that have not yet been identified

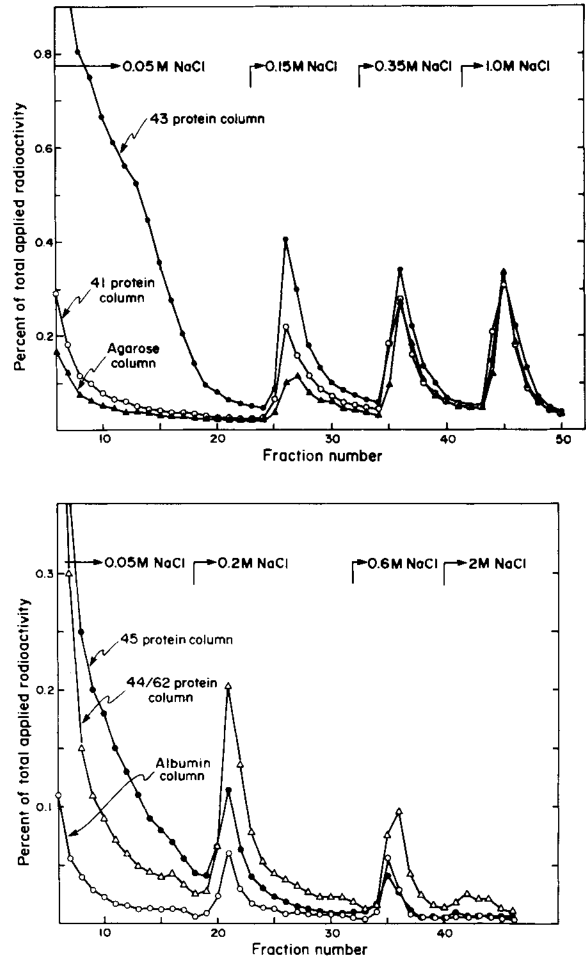

Figure 5. Elution profiles of radioactive T4 early proteins that have bound to affinity columns containing various immobilized T4 DNA replication proteins. Affinity columns were prepared by coupling purified replication proteins to Affi-Gel 10 (Bio-Rad). Proteins were dialyzed into coupling buffer (0.1 M HEPES [pH 7.5], 0.25 M NaCl) and mixed with matrix washed according to the manufacturer's instructions. Coupling was allowed to proceed for about 12 hr with mild agitation at 4°C. Remaining active groups were then blocked with ethanolamine. The final bound concentration of protein on the respective columns was 1.2 mg/ml (gene-*43* protein), 0.28 mg/ml (gene-*41* protein), 1.1 mg/ml (gene-*45* protein), and 1.4 mg/ml (gene-*44/62* protein). Columns of 2-ml packed volume were washed with column buffer (20 mM Tris-HCl [pH 8.1], 1 mM Na_3EDTA, 1 mM 2-mercaptoethanol, 10% (w/v) glycerol, 5 mM $MgCl_2$) containing 2 M NaCl and then equilibrated with the same buffer containing 50 mM NaCl. For the experiment with the *43*-protein and *41*-protein columns, the column buffer was supplemented with an additional 5 mM $MgCl_2$ plus 0.5% (w/v) Triton X-100. *E. coli* B (5×10^8 cells/ml in M9 minimal medium) was infected at an moi of 10 with T4D at 37°C and labeled with [^{35}S]methionine between 3 and 8 min after infection. Cell pellets were stored frozen. For lysis, cells were thawed, resuspended at a concentration of about 5×10^{10} cells/ml in column buffer containing 50 mM NaCl, 10 mM benzamidine, and 1 mM phenylmethylsulfonyl fluoride, and lysed by sonication. After the lysate was cleared by high-speed centrifugation, it was treated with a mixture of DNase I (9 μg/ml) and micrococcal nuclease (3 μg/ml) in the presence of 10 mM $CaCl_2$ for 30–60 min at 0°C. Aliquots of 3–8 ml were then chromatographed on the columns described above (2-ml packed volume), along with control columns containing either albumin-agarose or plain agarose matrix. The columns were washed with column buffer containing 50 mM NaCl and then eluted in steps of increasing NaCl in column buffer as shown. All flow rates were 2 ml/hr at 4°C and fractions of 1 ml were collected.

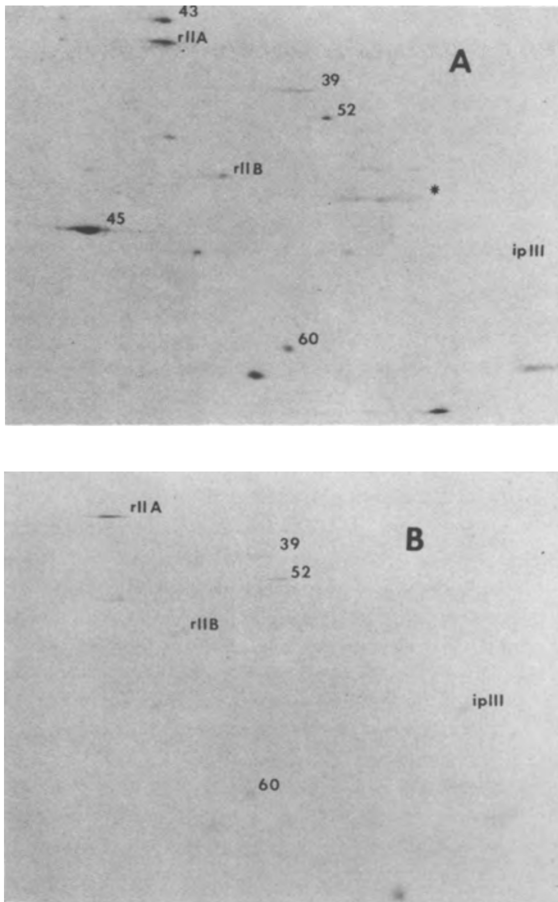

Figure 6. Two-dimensional electrophoretic identification of radioactive T4 early proteins that bind specifically to a replication protein affinity column. Samples from each eluted peak of radioactivity from a T4 DNA polymerase (gene-*43* protein) column (*A*) or from an albumin control column (*B*) were pooled, and the proteins were precipitated with TCA. The pellets were washed successively with acetone and ether, dried, resuspended in lysis buffer, and displayed by the two-dimensional nonequilibrium pH gradient electrophoresis (NEPHGE)-SDS system of O'Farrell et al. (1977). Protein spots were detected by autoradiography and identified by comparing them with a standard T4 early protein pattern (Cook and Seasholtz 1982; R.L. Burke et al., in prep.). Gene products are labeled at the upper right-hand corner of each protein spot. Asterisk indicates a specifically bound, but unidentified, T4 early protein.

with specific genes or activities, and for each column, they involve relatively minor amounts of the total protein (see Fig. 6).

As shown in Table 4, the gene-*43*, gene-*45*, and gene-*41* proteins (and possibly the *44*/*62*-protein complex) can each assemble into multimers. (The binding of the gene-*44* and gene-*62* proteins that we have detected on the *44*/*62*-protein column could instead be due to the known interaction between the separate gene products.) Self-aggregations of the gene-*43* and gene-*41* proteins are disrupted by prolonged washing with the low-salt column buffer and are therefore relatively labile; however, the lysate that passes through the column is completely cleared of gene-*43* protein by the *43*-protein column and the gene-*45* protein is likewise completely bound by the *45*-protein column. The labeled gene-*45* protein that bound to the latter column was not released until the column was stripped with 4 M urea and 0.3% SDS (data not shown), indicating that the *45* protein–*45* protein interaction is an especially strong one. Some or all of these self-aggregates may be involved in holding together the leading-strand and lagging-strand replication machinery (see Fig. 4).

The *43*-protein and *45*-protein columns each quantitatively remove the other protein from a lysate, and we believe that the interaction between gene-*45* protein and gene-*43* protein is one that is central to the replication apparatus.

A limitation of this technique is revealed by the fact that, although the gene-*32* and gene-*43* proteins associate in solution, the *32*-protein column binds gene-*43* protein without the *43*-protein column binding a significant amount of gene-*32* protein. This presumably reflects a nonrandom attachment of proteins to the agarose matrix, with the binding site on the gene-*43* protein being thereby masked for gene-*32*-protein interaction. Possibly, this problem could be overcome by altering the coupling conditions used. We have also failed so far to detect a specific interaction between the gene-*32* and the gene-61 proteins using this technique. However, this may be due to the fact that only trace amounts of the gene-*61* protein are produced in the T4-infected cell, making it especially difficult to detect by gel analysis.

Table 3. Proteins That Bind to a *32*-Protein Affinity Column

Gene product	Function
T4 gene *32*	helix-destabilizing protein (SSB protein)
T4 gene *43*	DNA polymerase
T4 gene *45*	DNA polymerase accessory protein
T4 gene *X*	*recA* analog in genetic recombination[a]
T4 gene *Y*	unknown function in genetic recombination
T4 gene *46*/*47*	exonuclease activity in genetic recombination
T4 gene *dda*	DNA-dependent ATPase (DNA helicase)
T4 RNase H	excises RNA from RNA-DNA hybrids
E. coli host protein	function unknown; 32,000 daltons

[a]T. Minagawa, pers. comm.

Table 4. Replication Protein Interactions as Measured by Affinity Chromatography

Protein affinity column	Proteins that bind to column
43 (DNA polymerase)	*43*
	45
41 (DNA helicase/primase)	*41*
45 (polymerase accessory protein)	*43*
	44/*62* (minor amount)
	45
	30
	E. coli RNA polymerase (minor amount)
44/*62* (polymerase accessory protein)	*45*
	44/*62* (minor amount)
	39/*52*/*60* (T4 DNA topoisomerase)

A Single *E. coli* RNA Polymerase Molecule Bound to a Promoter Site Can Block the Movement of the Replication Fork Formed by the T4 Core Replication System

Inside the cell, the replication fork must encounter many different proteins bound to the DNA template; how the replication fork deals with such bound proteins is currently not understood. We have begun to approach this question by studying the interaction of the replication fork formed by the T4 core replication system described above (Fig. 1) with *E. coli* RNA polymerase molecules that are bound to a double-stranded DNA template. We have chosen to use the replicative form of bacteriophage fd DNA as the template, in part because all the transcription promoter sites on this circular, double-stranded DNA molecule have been mapped and characterized in terms of their relative strength of RNA polymerase binding (Konings and Schoenmakers 1978; Schaller et al. 1978).

The assay we have used is described in Figure 7. The leading strand of a replication fork is initiated by the T4 core replication system at a specific site, the fd gene-2 protein nick at nucleotide 5781 (Meyer et al. 1979), and the start of DNA synthesis is partially synchronized by a brief preincubation with only three deoxyribonucleoside triphosphates present. There are at least ten promoter sites that have been mapped on the fd genome, and all of them are oriented with the same polarity, so that the *E. coli* RNA polymerase molecule that starts at each promoter will move in the same direction as the leading-strand DNA polymerase molecule (Fig. 7).

In our initial experiments, we measured the total amount of DNA synthesis obtained in the assay in Figure 7 as a function of the quantity of purified *E. coli* RNA polymerase added. A strong inhibition of DNA synthesis was observed, and a quantitative analysis suggested that a single RNA polymerase bound to a promoter site on the DNA template completely blocks the movement of the T4 replication fork (P. Bedinger et al., in prep.). We therefore analyzed the length of the nascent DNA chains made in these reactions by alkaline gel electrophoresis (Fig. 8, lanes d–f). These data reveal that when RNA polymerase is included in a DNA replication reaction, fork movement is blocked at specific promoter sequences, most notably at a strong promoter located 1000 bases downstream from the site of replication initiation. (This promoter has been designated the X promoter, and its function during fd infection is unknown [Konings and Schoenmakers 1978; Schaller et al. 1978].)

Very similar results are obtained whether or not RNA transcription is allowed by the addition of all four rNTPs to the reaction mixture. However, assays such as that in Figure 7 suggest that the replication fork stops permanently at a promoter site when transcription is blocked, whereas the same replication fork "follows" a moving RNA polymerase molecule at the relatively

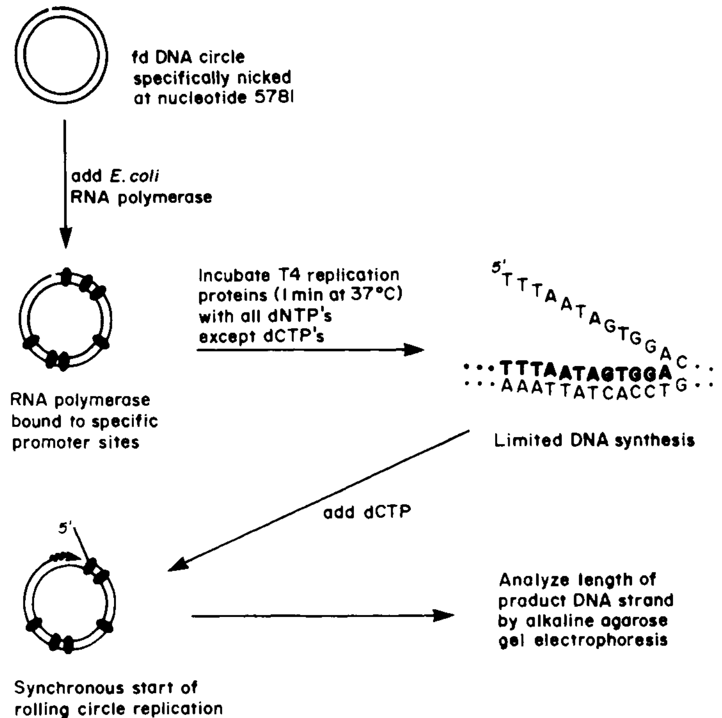

Figure 7. Schematic summary of the assay used to analyze the interaction of T4 replication forks with DNA-bound *E. coli* RNA polymerase molecules. Double-stranded, circular, replicative form (RF) bacteriophage fd DNA (6408 bp) was singly nicked at nucleotide 5781 by purified fd gene-2 protein and used at 2 μg/ml. Purified *E. coli* RNA polymerase (3.2 μg/ml) was incubated for 2 min at 37°C with the specifically nicked fd DNA under the standard replication conditions described in Fig. 3 (except that dCTP was absent) in order to form stable RNA polymerase:DNA complexes at the several known transcription-promoter sequences. The T4 gene-*43*, -*45*, -*44/62*, and -*32* proteins were added 1 min later to final concentrations of 2.5, 18, 20, and 100 μg/ml, respectively. T4 *dda*-gene protein was included in indicated reactions at 3 μg/ml. As indicated, incubation at 37°C for 1 min in the absence of dCTP allowed a limited amount of DNA synthesis to take place; the subsequent addition of dCTP (0.1 mM) starts a synchronous rolling-circle replication reaction. Aliquots of reaction mixtures were transferred into 3% SDS at the indicated times. Unincorporated nucleotides were removed by filtration through Sepharose CL-6B. The length of the nascent DNA chains was then analyzed by alkaline agarose (0.6%) gel electrophoresis. Gels measuring 0.4 × 14 × 21.5 cm were electrophoresed at 30 V for 40 hr in a running buffer containing 30 mM NaOH and 2 mM Na_3EDTA. The gels were dried onto Whatman 3MM filter paper and autoradiographed at −70°C using Kodak XR-2 film with a DuPont Lightning-Plus intensifying screen. The sizes of the radioactively labeled, newly synthesized DNA strands were determined by comparison with ^{32}P-labeled restriction fragments of T4 DNA.

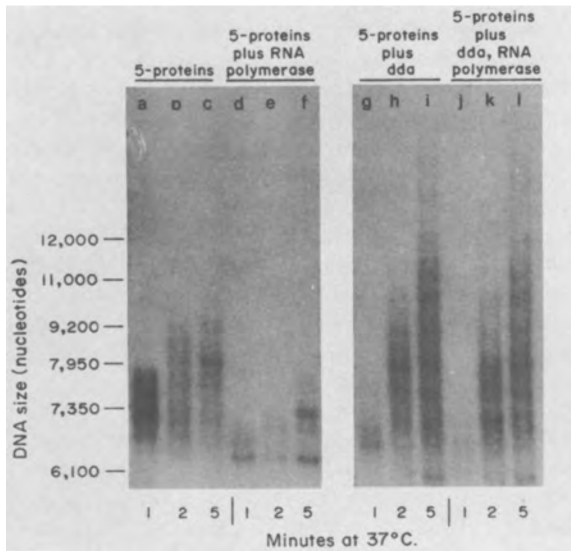

Figure 8. An analysis by alkaline agarose gel electrophoresis of the sizes of newly synthesized DNAs made in the assay of Fig. 7. (*a–c*) Products of a reaction catalyzed by the core replication system (Fig. 1) after 1, 2, and 5 min at 37°C in the absence of RNA polymerase; bands are seen at natural replication "pause" sites on the double-stranded fd DNA template (P. Bedinger and B. Alberts, in prep.). (*d–f*) Products of the same reaction carried out in the presence of *E. coli* RNA polymerase (17 molecules per fd DNA molecule; enzyme was a gift from M. Chamberlin, Dept. of Biochemistry, University of California, Berkeley). Well-defined bands corresponding to replication blocks are seen at about 100 and 1000 bases from the site of replication initiation, where the two promoters closest to the replication start site are located. (*g–i*) Products of a reaction catalyzed by the core replication system plus the T4 *dda*-gene protein; a substantial increase in the rate of replication is seen. (*j–l*) Same reaction in the presence of *E. coli* RNA polymerase (17 molecules per fd DNA molecule). Replication is no longer blocked at sites of RNA polymerase binding, and the net fork rate is equal to that observed without RNA polymerase present.

slow rate of RNA synthesis (about 10 nucleotides/sec) if transcription is allowed. These results reveal that the T4 core replication system is unable to bypass other proteins that are tightly bound to the DNA, suggesting that additional replication proteins are required for this purpose in the cell.

T4 *dda*-gene Protein Relieves the RNA Polymerase Inhibition of Replication

As discussed above, the T4 *dda*-gene protein is a DNA helicase that can unwind a DNA double helix in an ATP-dependent reaction. The gene-*41* protein is also thought to function, in part, as a DNA helicase. Since both proteins appear to be essential for in vivo DNA replication, we have examined the effect of their addition on the ability of the core replication system to bypass a bound RNA polymerase molecule. Neither addition of the gene-*41* protein (to produce the core-*41* system in Fig. 1) nor addition of both the gene-*41* and gene-*61* proteins (to produce the core-41-61 system in Fig. 1) removed the RNA polymerase block to replication in this system (data not shown). In contrast, the addition of the *dda*-gene protein to the T4 core replication system had two striking effects: (1) Like the gene-*41* protein, it caused a threefold to fourfold increase in the rate of replication fork movement in the absence of RNA polymerase (Fig. 8: compare lanes a–c and lanes g–i). (2) Whereas previously a single RNA polymerase molecule completely blocked fork movement, now the presence of more than eight *E. coli* RNA polymerase molecules bound to each fd DNA circle was without effect on fork-movement rates (Fig. 8, lanes g–l). We therefore propose that the removal of barriers to DNA replication caused by such DNA-bound proteins is an important function of the type of DNA helicase encoded by the *dda* gene. This might explain why more than one DNA helicase seems to be required at the T4 replication fork. How these two enzymes cooperate (if they do) and where they are located relative to each other in the replication complex remain to be investigated. However, both helicases would be expected to be in intimate contact with the single-stranded region of the DNA template on the lagging side of the fork: the gene-*41* protein because of its central role in RNA primer synthesis (Liu and Alberts 1980) and the *dda*-gene protein because of its defined polarity of helix unwinding while moving along a DNA single strand (Krell et al. 1979).

Two other important questions need to be addressed: (1) Do the RNA polymerase molecules remain bound to the DNA as the replication fork traverses transcription complexes and, if so, to which daughter DNA helix are they passed? (2) What happens to a replication fork when it encounters an RNA polymerase molecule that is moving in the *opposite* direction to that tested here, thus colliding with the leading-strand DNA polymerase molecule "head-on"?

Site Specificity of T4 Topoisomerase Recognition of T4 DNA

As stated above, genetic studies suggest a role for the T4-induced type-II DNA topoisomerase in the initiation of T4 DNA replication. Using a very sensitive in vitro assay, which should detect on the order of 1 initiation event per 10^4 input template DNA molecules, we have thus far been unable to detect any origin-specific initiation on a native T4 DNA template upon mixing various combinations of the replication proteins listed in Table 1, with and without the host RNA polymerase present. Apparently, there is at least one additional initiation factor that we are not providing in these in vitro reactions. While continuing to pursue this approach, we have also characterized the interaction of the T4 topoisomerase with T4 DNA in order to screen for a possible origin-specific topoisomerase reaction.

Type-II topoisomerases create a transient break in the double-stranded DNA helix through which an intact segment of DNA helix is passed (Cozzarelli 1980; L.F.

Liu et al. 1980; Gellert 1981). The presumptive reaction intermediate, broken DNA with protein covalently attached to each of the newly formed 5′ ends at the break, can be detected at a low frequency after detergent treatment of T4 topoisomerase reactions (Liu 1980). Studies of this intermediate in the *E. coli* DNA gyrase reaction were greatly facilitated by the finding that addition of either of the two closely related antibacterial agents, nalidixic acid or oxolinic acid, results in nearly quantitative recovery of the enzyme in the covalent complex (Gellert et al. 1977; Sugino et al. 1977). Oxolinic acid acts in a mechanistically similar fashion on the T4 topoisomerase, although higher drug concentrations are required (K.N. Kreuzer et al., in prep.). This high-efficiency DNA cleavage reaction allows us to readily map preferred recognition sites for the topoisomerase. Since T4 DNA is quite large (165 kb) and the topoisomerase binds to many sites with widely varying relative affinities, we have developed a filter-binding procedure to identify those sites that react covalently with the topoisomerase. Briefly, radioactive restriction enzyme digests of T4 DNA are treated with topoisomerase in the presence of oxolinic acid, and SDS is added to trap the covalent complex of enzyme and cleaved DNA. The SDS is then removed by a rapid gel filtration, and the covalent complexes are purified away from unreacted substrate DNA fragments by using a modification of the glass-fiber (GF/C) filter-binding technique of Coombs and Pearson (1978) and Thomas et al. (1979). Each strong topoisomerase cleavage site, being located a unique distance from each of the two ends of a DNA restriction fragment, produces two discretely sized DNA cleavage fragments that bind to the filter.

A major complication in analyzing topoisomerase-binding sites is that native T4 DNA contains glucosylated 5-hydroxymethylcytosine (HMC glu) residues, making it refractory to nearly all restriction nucleases. An analysis using restriction digests of a multiply mutant T4 DNA containing unmodified cytosine (C) residues indicated that the T4 topoisomerase recognizes on the order of 15 relatively high-affinity cleavage sites on this DNA (K.N. Kreuzer et al., in prep.). However, since the mutant T4 cytosine-containing phage is seriously disturbed with respect to various aspects of nucleic acid metabolism in vivo, it seemed imperative to compare topoisomerase recognition of C T4 DNA with that of the native HMC glu T4 DNA. We therefore began a search for a commercially available restriction enzyme that would cleave the native DNA, and discovered that *Taq*I is such an enzyme (D. Coit, unpubl.).

Using radioactively labeled *Taq*I restriction fragments and the filter-binding assay described above, the topoisomerase-binding sites on C T4 DNA and HMC glu T4 DNA were compared. From the *Taq*I-digested C T4 DNA substrate, four very prominent cleavage fragments were generated by the T4 topoisomerase along with numerous weaker fragments (Fig. 9a, lanes 3–5). At the

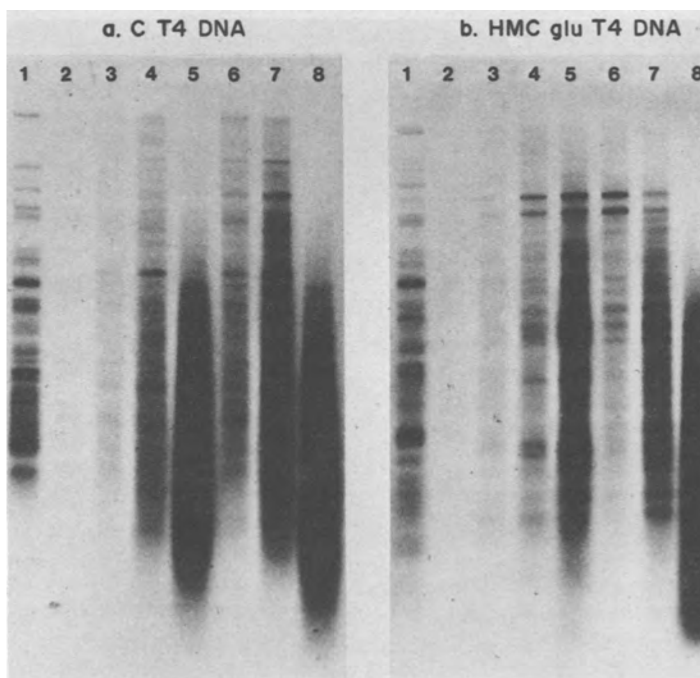

Figure 9. T4-topoisomerase-induced cleavage of *Taq*I-digested T4 DNAs. The substrates for these topoisomerase cleavage reactions were ^{32}P-end-labeled, *Taq*I-digested C T4 DNA (*a*) and HMC glucose T4 DNA (*b*). Besides the indicated amount of topoisomerase and 0.2 μg of the *Taq*I-digested DNA, each reaction also contained 40 mM Tris-Cl (pH 7.8), 60 mM KCl, 10 mM $MgCl_2$, 0.5 mM DTT, 0.5 mM Na_3EDTA, nuclease-free albumin at 30 μg/ml, and oxolinic acid at 500 μg/ml (a total volume of 20 μl). After incubating for 15 min at 30°C, SDS was added to 0.2%, and the volume was increased to 60 μl. The SDS was removed by rapid gel filtration through Sepharose CL-6B (Neal and Florini 1973; K. Kreuzer et al., in prep.), and one-fifth final volume of 5X Binding Buffer was added to the eluate (Binding Buffer: 50 mM Tris-Cl [pH 7.8], 200 mM KCl, 10 mM $MgCl_2$, 0.5 mM Na_3EDTA). Covalent protein-DNA complexes were then collected on GF/C filters as follows. Two GF/C filters (7-mm dia; one on top of the other) were placed onto a larger GF/A filter, and 150 μl of Binding Buffer was added slowly to the top filter (the buffer soaks through the filter pair by capillary action). After this prewash, the filter pair was moved to a dry area of the GF/A filter, and the gel-filtration eluate (about 50 μl) was applied, followed immediately by 150 μl of Binding Buffer. The filter pair was then washed four more times with 150 μl of Binding Buffer. The top GF/C filter disk was removed and eluted twice with 20 μl of 10 mM Tris-Cl (pH 7.8) and 0.1% SDS. The filter eluate was treated with proteinase K at 100 μg/ml for 30 min at 37°C, and the DNA fragments present were separated by electrophoresis through a 1.2% agarose gel. For both sets of reactions, lane *1* contains the unfractionated restriction digest substrate and the other lanes contain the GF/C filter eluates of reactions with the following compositions: (*2*) no enzyme; (*3–5*) increasing amounts of T4 topoisomerase (16, 80, and 400 ng, respectively) in the absence of ATP; (*6–8*) increasing amounts of T4 topoisomerase (16, 80, and 400 ng, respectively) in the presence of 0.5 mM ATP.

highest enzyme level, a smear of low-molecular-weight fragments was generated (Fig. 9a, lane 5), indicating that many of the *Taq*I restriction fragments could be cleaved more than once by the topoisomerase. The addition of ATP had little or no effect on the particular fragments recovered but increased the efficiency of cleavage slightly (Fig. 9a, lanes 6–8). These results demonstrate that the T4 topoisomerase recognizes a large number of cleavage sites, including at least two very strong ones, on C T4 DNA.

The results of the filter-binding assay with *Taq*I-digested HMC glu T4 DNA were quite different (Fig. 9b). Even at the highest enzyme level tested (Fig. 9b, lane 5), only about eight strong and eight weaker cleavage fragments were generated, and the pattern was not obscured by multiple cutting of the larger fragments. Thus, under our conditions, the T4 topoisomerase shows a marked increase in specificity when recognizing HMC glu T4 DNA, as compared with C T4 DNA. Moreover, the addition of ATP has a major effect on the results obtained with HMC glu DNA, increasing the frequency of topoisomerase-induced cleavage approximately fivefold (Fig. 9b, lanes 6–8). At the highest enzyme level, multiple cutting of the largest DNA cleavage fragments is now observed, but about 12–14 specific cleavage fragments are still evident (Fig. 9b, lane 8). The strong cleavage fragments generated in the absence of ATP are also seen at the lower enzyme levels in the presence of ATP, and therefore the same strong sites are recognized in both cases.

DNA fragments generated from substrates of C T4 DNA and HMC glu T4 DNA are not directly comparable with respect to electrophoretic migration, and thus it is not clear whether some of the strong topoisomerase cleavage sites recognized on the two DNAs are the same. T4 DNA containing hydroxymethylated, but not glucosylated, cytosine residues is recognized by the topoisomerase in almost the same way that C T4 DNA is recognized (data not shown). Thus, the glucose moieties attached to the hydroxymethyl groups must be responsible for the increased specificity of topoisomerase action on native T4 DNA.

The topoisomerase cleavage fragments produced from native T4 DNA have been mapped on the genome to explore the possibility that the enzyme specifically recognizes T4 replication origins. Mapping was carried out by hybridization to previously characterized restriction digests of C T4 DNA, using the Southern cross-blot technique (Sato et al. 1977; Wensink et al. 1979). Briefly, an unlabeled restriction digest of C T4 DNA was electrophoresed through an agarose gel and then, after denaturation, transferred to a nitrocellulose sheet. A second agarose gel containing separated, radioactive topoisomerase cleavage fragments of HMC glu T4 DNA was then transferred to the same nitrocellulose sheet under hybridization conditions, with the second gel rotated 90° with respect to the first. The band corresponding to each topoisomerase cleavage fragment intersects with the band corresponding to every unlabeled restriction fragment, and a radioactive spot is produced wherever the cleavage fragment intersects a restriction fragment containing a complementary DNA sequence.

Figure 10 displays a Southern cross-blot of the radioactive HMC glu DNA cleavage fragments mapped against an *Xho*I digest of C T4 DNA. Each strong topoisomerase cleavage fragment (labeled A–F) hybridized to a specific *Xho*I restriction fragment (labeled A–H); e.g., topoisomerase fragment C hybridized

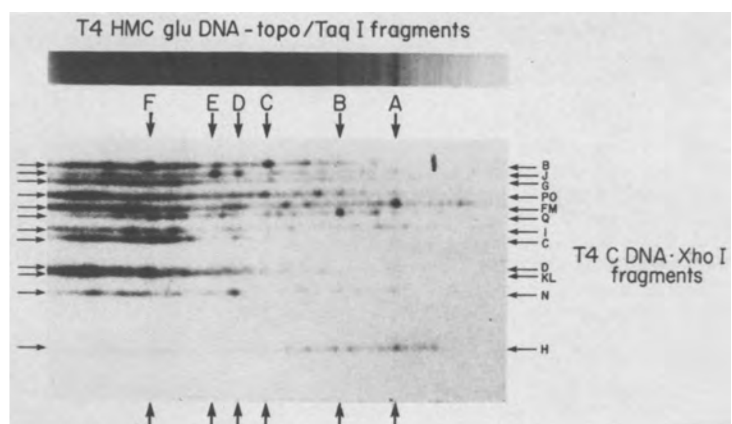

Figure 10. Hybridization of radioactive T4 topoisomerase cleavage fragments of native T4 DNA to an *Xho*I restriction digest of C T4 DNA. To prepare conveniently enough radioactive topoisomerase cleavage fragments for this procedure, intact T4 DNA containing HMC glu was treated with topoisomerase in the presence of oxolinic acid, and the reaction was terminated by adding SDS to 0.2%. Following proteinase-K treatment (to remove the covalently attached topoisomerase), the reaction products were subjected to gel filtration to remove SDS, and the recovered DNA was labeled at the 3′ end with T4 DNA polymerase (O'Farrell et al. 1980), which thereby specifically labels the positions of topoisomerase cleavage, as well as the randomly placed ends of the circularly permuted T4 genome. This labeled DNA was then cleaved with *Taq*I to generate discretely sized fragments in high yield for hybridization. The band pattern of cleavage fragments generated using this procedure is very similar to that obtained using the method described in Fig. 9 (data not shown). The Southern cross-blots were prepared as described by Wensink et al. (1979), with the following exceptions. The first dimension was a horizontal 0.6% agarose gel loaded with 10–15 μg of an (unlabeled) restriction digest of C T4 DNA. The gel was stained with ethidium bromide and photographed to record the positions of the restriction fragment bands. The DNA was then denatured and transferred to nitrocellulose as described previously (Wensink et al. 1979). The second dimension was a horizontal 1.2% agarose gel, loaded with radioactively labeled topoisomerase/*Taq*I cleavage fragments. A segment of this gel was dried down directly for autoradiography to determine the positions of the radioactive cleavage fragments, and the rest of the gel was transferred under hybridization conditions to nitrocellulose (rotated 90° with respect to the first gel). The gel at the top shows the radioactive cleavage fragments without hybridization, and the major bands are labeled *A* through *F* in order of increasing electrophoretic mobility. Arrows indicate the positions of unlabeled *Xho*I restriction fragments of C T4 DNA (designations are those of O'Farrell et al. [1980]). The small *Xho*I fragment A did not show any significant hybridization and is not shown, and *Xho*I fragment E ran off the gel.

to *Xho*I fragment B. The topoisomerase cleavage fragments were also mapped against an *Xba*I digest and against a *Bgl*II, *Bam*HI, *Sal*I triple digest of C T4 DNA (data not shown). By comparing the restriction fragments that hybridized in each of these DNA digests, the locations of the strong topoisomerase cleavage fragments A through E were determined (Fig. 11). Fragments D and E always hybridized to the same cold restriction band, and we presume that these two are partner fragments generated from cleavage at a single topoisomerase-binding site. The partner fragments for cleavage products A, B, and C were not observed.

In the Southern cross-blot shown in Figure 10, there are clearly several spots produced from hybridization of low-molecular-weight cleavage fragments. These fragments (all smaller than about 300 bp) could not be assigned a definitive map position by comparing several Southern cross-blots, but they presumably include some or all of the missing partner fragments, as well as potential new cleavage sites. The mapping analysis is therefore incomplete, but we are confident of the locations of the four strong cleavage sites shown in Figure 11, since the map positions inferred from several Southern cross-blots were consistent.

The locations of the strong topoisomerase cleavage sites on native T4 DNA suggest that they could play a role in the initiation of T4 DNA synthesis. To date, five preferential origins of in vivo T4 DNA synthesis have been mapped by several groups (Halpern et al. 1979; Mosig et al. 1981; C.F. Morris, pers. comm.). For various technical reasons, the map locations of these origins have been defined only to within about 5–10 kbp each, and the possibility of other less efficient origins has not been ruled out. The topoisomerase DNA cleavage fragment C maps in the vicinity of genes *61* and *56*, which has been found to be the region containing the strong primary origin for T4 DNA replication (C.F. Morris; G. Mosig; both pers. comm.). The results of Luder and Mosig (1982) indicate that initiation from this region requires the recognition of an early promoter by (unmodified) host RNA polymerase. In preliminary experiments, we have mapped the topoisomerase cleavage site to a location about 300 bp downstream from a strong RNA-polymerase-binding site, located between genes *56* and *61*. Niggemann et al. (1981) have mapped the strongest early T4 promoter to the same location. It seems possible that an interaction between the T4 topoisomerase and the host RNA polymerase at their respective adjacent binding sites is involved in the initiation of T4 DNA synthesis at this presumptive primary replication origin. Experiments are currently in progress to test for such an interaction.

As described previously, initiation of T4 DNA replication from secondary replication origins occurs

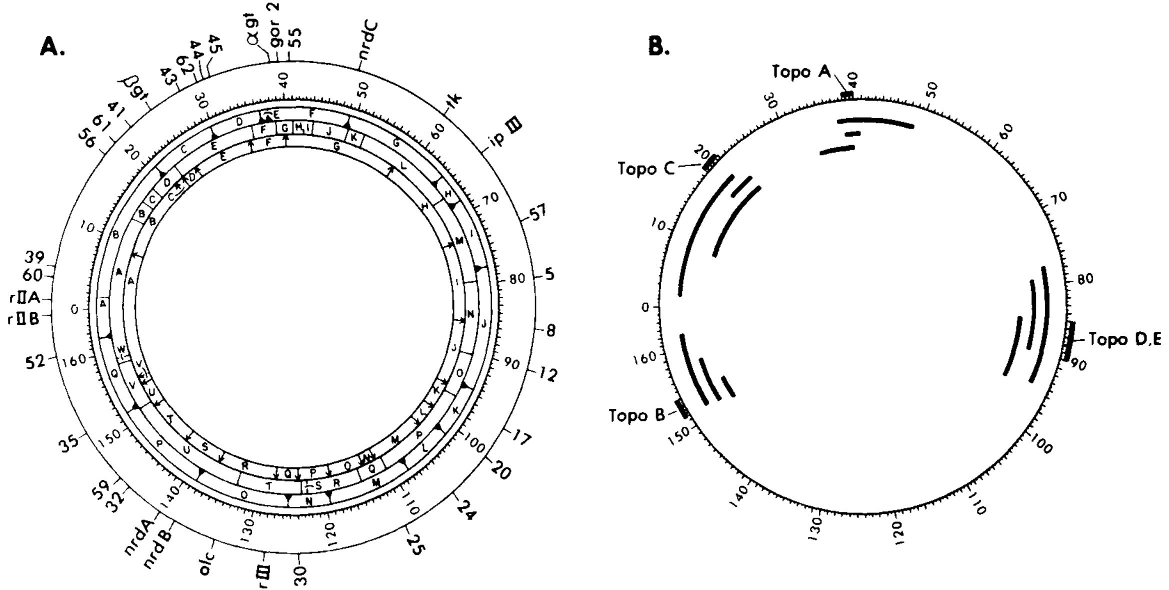

Figure 11. Map of strong T4 topoisomerase cleavage sites on native T4 DNA. Results of several Southern cross-blots (see Fig. 10) using different restriction digests of the unlabeled C T4 DNA are summarized. The restriction map in *A* is modified from O'Farrell et al. (1980). It shows the positions of various T4 genes and map locations in kilobase pairs from 0 to 165 on the outer circles, and restriction sites and fragments of *Xho*I (▲), *Xba*I (I), and the combination of *Bgl*II, *Bam*HI, and *Sal*I (↑) on the three inner circles, respectively. The inferred locations of the strong topoisomerase cleavage fragments are shown just outside the circular map in *B*. Bars inside this map indicate the locations of the restriction fragments that hybridized to each cleavage fragment, with *Xho*I fragments outermost, followed by *Xba*I fragments, and then fragments from a *Bgl*II, *Bam*HI, *Sal*I triple restriction enzyme digest. Cleavage fragment F hybridized to the same restriction fragments as fragment C in some of the blots, but its hybridization was ambiguous in others. Thus, fragment F may be the partner fragment of fragment C. The only other possible ambiguity in the mapping involved the location of cleavage fragment A. This fragment hybridized to *Xba*I restriction fragment G, either *Xho*I restriction fragment F or M (a doublet on the gel), and either restriction fragment F, O, or P in the triple digest (which formed a triplet on the gel). The only consistent mapping of this cleavage fragment is that shown in *B*: hybridization to *Xho*I fragment F, to *Xba*I fragment G, and to restriction fragment F in the triple digest.

later in infection, and it is independent of host RNA polymerase, but dependent on phage-induced recombination proteins (Luder and Mosig 1982). The locations of these origins of DNA synthesis are close to regions previously identified as recombinational hot spots (C.F. Morris, pers. comm.). One such hot spot has been mapped near the gene-*34/35* border (Beckendorf and Wilson 1972), and the strong topoisomerase cleavage fragment B has been mapped to a region of about 2500 bp that includes this gene-*34/35* border (Fig. 11). In addition, recombination from this hot spot is dependent on glucosylation of the participating DNA molecules (Levy and Goldberg 1980), and we have shown here that topoisomerase recognition of T4 DNA is markedly altered by the glucose modification present in native T4 DNA. It therefore appears plausible that recombination at this hot spot and initiation of DNA synthesis in this region are both dependent on the topoisomerase recognition of the strong topoisomerase-binding site mapped here. The DNA cleavage activity of the enzyme could be involved in both processes, since DNA ends are known to be recombinogenic in T4 (Doermann and Parma 1967), and the type of end produced by the topoisomerase, a free 3′-hydroxyl end, could serve as a primer for DNA synthesis.

The correlation of two strong topoisomerase cleavage sites with known origins of replication suggests that such sites may be functionally associated with replication initiation. However, the two other strong cleavage sites mapped here do not seem to be associated with known origins, since the two nearest origins (mapped in the intervals of genes *nrdC* to *tk* and genes *1* to *5*) are several kilobase pairs removed from them. Further experiments will be necessary to establish whether the topoisomerase specificity studied here is actually involved in the selection and utilization of T4 replication origins by one or both of the two different modes of T4 DNA replication initiation.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant GM-24020 from the National Institute of General Medical Sciences. C.V.J. is a fellow of the Leukemia Society of America, K.N.K. is a senior fellow of the California Division of the American Cancer Society, and T.F. is supported by National Institutes of Health training grant CA-09270. We thank Michael Chamberlin for his generous gift of *E. coli* RNA polymerase; some of the expreriments reported using this polymerase were performed in collaboration with Mark Hochstrasser.

REFERENCES

Alberts, B., C.F. Morris, D. Mace, N. Sinha, M. Bittner, and L. Moran. 1975. Reconstruction of the T4 bacteriophage DNA replication apparatus from purified components. *ICN-UCLA Symp. Mol. Cell. Biol.* **3:** 241.

Alberts, B.M., J. Barry, P. Bedinger, R.L. Burke, U. Hibner, C.-C. Liu, and R. Sheridan. 1980. Studies of replication mechanisms with the T4 bacteriophage in vitro system. *ICN-UCLA Symp. Mol. Cell. Biol.* **19:** 449.

Aposhian, H.V. and A. Kornberg. 1962. Enzymatic synthesis of deoxyribonucleic acid. IX. The polymerase formed after T2 bacteriophage infection of *Escherichia coli*: A new enzyme. *J. Biol. Chem.* **237:** 519.

Beckendorf, S.K. and J.H. Wilson. 1972. A recombination gradient in bacteriophage T4 gene 34. *Virology* **50:** 315.

Bittner, M., R.L. Burke, and B.M. Alberts. 1979. Purification of the T4 gene 32 protein free from detectable deoxyribonuclease activities. *J. Biol. Chem.* **254:** 9565.

Brutlag, D. and A. Kornberg. 1972. Enzymatic synthesis of deoxyribonucleic acid. *J. Biol. Chem.* **247:** 241.

Challberg, M.D. and P.T. Englund. 1979. The effect of template secondary structure on vaccinia DNA polymerase. *J. Biol. Chem.* **254:** 7820.

Cook, K.S. and A. F. Seasholtz. 1982. Identification of some bacteriophage T4 prereplicative proteins on two-dimensional gel patterns. *J. Virol.* **42:** 767.

Coombs, D.H. and G.D. Pearson. 1978. Filter-binding assay for covalent DNA-protein complexes: Adenovirus DNA-terminal protein complex. *Proc. Natl. Acad. Sci.* **75:** 5291.

Cozzarelli, N.R. 1980. DNA gyrase and the supercoiling of DNA. *Science* **207:** 953.

Curtis, M.J. and B.M. Alberts. 1976. Studies on the structure of intracellular bacteriophage T4 DNA. *J. Mol. Biol.* **102:** 793.

Doermann, A.H. and D.H. Parma. 1967. Recombination in bacteriophage T4. *J. Cell. Physiol.* (suppl. 1) **70:** 147.

Epstein, R.H., A. Bolle, C.M. Steinberg, E. Kellenberger, E. Boy de la Tour, R. Chevalley, R.S. Edgar, M. Susman, G.H. Denhardt, and A. Lielausis. 1964. Physiological studies of conditional lethal mutants of bacteriophage T4D. *Cold Spring Harbor Symp. Quant. Biol.* **28:** 375.

Gellert, M. 1981. DNA topoisomerases. *Annu. Rev. Biochem.* **50:** 879.

Gellert, M., K. Mizuuchi, M.H. O'Dea, T. Itoh, and J.-I. Tomizawa. 1977. Nalidixic acid resistance: A second genetic character involved in DNA gyrase activity. *Proc. Natl. Acad. Sci.* **74:** 4772.

Goulian, M., Z.J. Lucas, and A. Kornberg. 1968. Enzymatic synthesis of deoxyribonucleic acid. XXV. Purification and properties of deoxyribonucleic acid polymerase induced by infection with phage T4⁺. *J. Biol. Chem.* **243:** 627.

Halpern, M.E., T. Mattson, and A.W. Kozinski. 1979. The origins of phage T4 DNA replication as revealed by hybridization to cloned genes. *Proc. Natl. Acad. Sci.* **76:** 6137.

Hibner, U. and B.M. Alberts. 1980. Fidelity of DNA replication catalyzed in vitro on a natural DNA template by the T4 bacteriophage multienzyme complex. *Nature* **285:** 300.

Huang, C.-C. and J.E. Hearst. 1980. Pauses at positions of secondary structure during in vitro replication of single-stranded fd bacteriophage DNA by T4 DNA polymerase. *Anal. Biochem.* **103:** 127.

Huang, C.-C., J.E. Hearst, and B.M. Alberts. 1981. Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. *J. Biol. Chem.* **256:** 4087.

Huberman, J.S., A. Kornberg, and B.M. Alberts. 1971. Stimulation of T4 bacteriophage DNA polymerase by the protein product of T4 gene 32. *J. Mol. Biol.* **62:** 39.

Konings, R.N.H. and J.G.G. Schoenmakers. 1978. Transcription of the filamentous phage genome. In *The single-stranded DNA phages* (ed. D.T. Denhardt et al.), p.507. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

Krell, H., H. Durwald, and H. Hoffman-Berling. 1979. A DNA-unwinding enzyme induced in bacteriophage

T4-infected *Escherichia coli* cells. *Eur. J. Biochem.* **93:** 387.

LEVY, J.N. and E.B. GOLDBERG. 1980. Region-specific recombination in phage T4. I. A special glucosyl-dependent recombination system. *Genetics* **94:** 519.

LIU, C.-C. and B.M. ALBERTS. 1980. Pentaribonucleotides of mixed sequence are synthesized and efficiently prime de novo DNA chain starts in the T4 bacteriophage DNA replication system. *Proc. Natl. Acad. Sci.* **77:** 5698.

——. 1981a. Characterization of the DNA-dependent GTPase activity of T4 gene 41 protein, an essential component of the T4 bacteriophage DNA replication apparatus. *J. Biol. Chem.* **256:** 2813.

——. 1981b. Characterization of RNA primer synthesis in the T4 bacteriophage in vitro DNA replication system. *J. Biol. Chem.* **256:** 2821.

LIU, C.-C., R.L. BURKE, U. HIBNER, J. BARRY, and B.M. ALBERTS. 1979. Probing DNA replication mechanisms with the T4 bacteriophage in vitro system. *Cold Spring Harbor Symp. Quant. Biol.* **43:** 469.

LIU, L.F. 1980. DNA strand passing and the function of type II DNA topoisomerases. *ICN-UCLA Symp. Mol. Cell. Biol.* **19:** 817.

LIU, L.F., C.-C. LIU, and B.M. ALBERTS. 1979. T4 DNA topoisomerase: A new ATP-dependent enzyme essential for initiation of T4 bacteriophage DNA replicaton. *Nature* **281:** 456.

——. 1980. Type II DNA topoisomerases: Enzymes that can unknot a topologically knotted DNA molecule via a reversible double-strand break. *Cell* **19:** 697.

LUDER, A. and G. MOSIG. 1982. Two alternative mechanisms for initiation of DNA replication forks in bacteriophage T4: Priming by RNA polymerase and by recombination. *Proc. Natl. Acad. Sci.* **79:** 1101.

MCCARTHY, D., C. MINNER, H. BERNSTEIN, and C. BERNSTEIN. 1976. DNA elongation rates and growing point distributions of wild-type phage T4 and a DNA-delay amber mutant. *J. Mol. Biol.* **106:** 963.

MERRIL, C.R., D. GOLDMAN, S.A. SEDMAN, and M.H. EBERT. 1981. Ultra-sensitive stain for proteins in polyacrylamide gels shows regional variation in cerebrospinal fluid proteins. *Science* **211:** 1437.

MEYER, T.F., K. GEIDER, C. KURZ, and H. SCHALLER. 1979. Cleavage site of bacteriophage fd gene II-protein in the origin of viral strand replication. *Nature* **278:** 365.

MORRIS, C.F., L.A. MORAN, and B.M. ALBERTS. 1979a. Purification of gene 41 protein of bacteriophage T4. *J. Biol. Chem.* **254:** 6797.

MORRIS, C.F., H. HAMA-INABA, D. MACE, N.K. SINHA, and B.M. ALBERTS. 1979b. Purification of the gene 43, 44, 45, and 62 proteins of the bacteriophage T4 DNA replication apparatus. *J. Biol. Chem.* **254:** 6787.

MOSIG, G., A. LUDER, G. GARCIA, R. DANNENBERG, and S. BOCK. 1979. In vivo interactions of genes and proteins in DNA replication and recombination of phage T4. *Cold Spring Harbor Symp. Quant. Biol.* **43:** 501.

MOSIG, G., A. LUDER, L. ROWEN, P. MCDONALD, and S. BOCK. 1981. On the role of recombination and topoisomerase in primary and secondary initiation of DNA replication. *ICN-UCLA Symp. Mol. Cell. Biol.* **22:** 227.

NEAL, M.W. and J.R. FLORINI. 1973. A rapid method for desalting small volumes of solution. *Anal. Biochem.* **55:** 328.

NEWPORT, J. 1980. "A study of the proteins involved in bacteriophage T4 replication." Ph.D. thesis, University of Oregon, Eugene, Oregon.

NIGGEMANN, E., I. GREEN, H.P. MEYER, and W. RUGER. 1981. Physical mapping of bacteriophage T4. *Mol. Gen. Genet.* **184:** 289.

NOSSAL, N.G. 1980. RNA priming of DNA replication by bacteriophage T4 proteins. *J. Biol. Chem.* **255:** 2176.

NOSSAL, N.G. and B.M. PETERLIN. 1979. DNA replication by bacteriophage T4 proteins: The T4 43, 32, 44-62, and 45 proteins are required for strand displacement synthesis at nicks in duplex DNA. *J. Biol. Chem.* **254:** 6032.

O'FARRELL, P.H., E. KUTTER, and M. NAKANISHI. 1980. A restriction map of the bacteriophage T4 genome. *Mol. Gen. Genet.* **179:** 421.

O'FARRELL, P.Z., H.M. GOODMAN, and P.H. O'FARRELL. 1977. High resolution two-dimensional electrophoresis of basic as well as acidic proteins. *Cell* **12:** 1133.

PIPERNO, J.R. and B.M. ALBERTS. 1978. An ATP stimulation of T4 DNA polymerase mediated via T4 gene 44/62 and 45 proteins. *J. Biol. Chem.* **253:** 5174.

PURKEY, R.M. and K. EBISUZAKI. 1977. Purification and properties of a DNA-dependent ATPase induced by bacteriophage T4. *Eur. J. Biochem.* **75:** 303.

RATNER, D. 1974. The interaction of bacterial and phage proteins with immobilized *E. coli* RNA polymerase. *J. Mol. Biol.* **88:** 373.

SATO, S., C.A. HUTCHINSON III, and J.I. HARRIS. 1977. A thermostable sequence specific endonuclease from *Thermus aquaticus*. *Proc. Natl. Acad. Sci.* **74:** 542.

SCHALLER, H., E. BECK, and M. TAKANAMI. 1978. Sequence and regulatory signals of the filamentous phage genome. In *The single-stranded DNA phages* (ed. D.T. Denhardt et al.), p. 139. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.

SINHA, N.K. and M.D. HAIMES. 1981. Molecular mechanisms of substitution mutagenesis: An experimental test of the Watson-Crick and Topal-Fresco models of base mispairings. *J. Biol. Chem.* **256:** 10671.

SINHA, N.K., C.F. MORRIS, and B.M. ALBERTS. 1980. Efficient in vitro replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. *J. Biol. Chem.* **255:** 4290.

SNYDER, L.R. and D.L. MONTGOMERY. 1974. Inhibition of T4 growth by an RNA polymerase mutation of *Escherichia coli*: Physiological and genetic analysis of the effects during phage development. *Virology* **62:** 184.

STETLER, G.L., G.J. KING, and W.M. HUANG. 1979. T4 DNA-delay proteins, required for specific DNA replication, form a complex that has ATP-dependent DNA topoisomerase activity. *Proc. Natl. Acad. Sci.* **76:** 3737.

SUGINO, A., C.L. PEEBLES, K.N. KREUZER, and N.R. COZZARELLI. 1977. Mechanism of action of nalidixic acid: Purification of *Escherichia coli nalA* gene product and its relationship to DNA gyrase and a novel nicking-closing enzyme. *Proc. Natl. Acad. Sci.* **74:** 4767.

THOMAS, C.A., JR., K. SAIGO, E. MCLEOD, and J. ITO. 1979. The separation of DNA segments attached to proteins. *Anal. Biochem.* **93:** 158.

WARNER, H.R. and M.D. HOBBS. 1967. Incorporation of uracil-^{14}C into nucleic acids in *E. coli* infected with bacteriophage T4 and T4 amber mutants. *Virology* **33:** 376.

WENSINK, P.C., S. TABATA, and C. PACHL. 1979. The clustered and scrambled arrangement of moderately repetitive elements in *Drosophila* DNA. *Cell* **18:** 1231.