Probing DNA Replication Mechanisms with the T4 Bacteriophage In Vitro System

C. C. Liu, R. L. Burke, U. Hibner, J. Barry, and B. Alberts

Department of Biochemistry and Biophysics, University of California, San Francisco, California 94143

We have previously reported the development of a multienzyme in vitro DNA replication system which replicates double-stranded (DS) DNA templates efficiently. This system consists of six highly purified bacteriophage T4-encoded proteins, the products of genes 32, 41, 43, 44, 45, and 62 (Morris et al. 1975; Alberts et al. 1975, 1977). Each of these proteins has been shown to be essential for DNA replication in vivo (Epstein et al. 1964). An amber mutation in any of these genes results in a phenotype showing very little or no DNA synthesis (Warner and Hobbs 1967), even though adequate deoxyribo- and ribonucleoside triphosphate levels are maintained (Mathews 1972). Additionally, temperature-sensitive mutants have been isolated in genes 32, 41, 43, and 45 which cease DNA replication immediately upon transfer to the restrictive temperature (Riva et al. 1970; Curtis and Alberts 1976).

From such in vivo data, each of the above proteins is expected to function directly at a replication fork and, therefore, to be an important component of any T4 in vitro replication system. Our reconstructed in vitro system, in fact, requires each of these six proteins, as previously shown (Morris et al. 1975; Alberts et al. 1975, 1977). In addition, a seventh T4-encoded protein, which has recently been isolated and purified from T4-infected lysates, is shown to have an essential role in the in vitro DNA replication process (see below).

It has been our goal to dissect and reconstruct the T4-encoded replication apparatus in order to understand the protein-protein and protein-DNA interactions generally involved in DNA replication. We wish to obtain a static picture of the architecture of a replication fork (including the arrangement of the multienzyme complex which functions there), as well as a dynamic picture of the events that accompany DNA polymerization. Heretofore much of the work has, of necessity, focused on the isolation and purification of the different replication proteins and on a search for separate, intrinsic protein activities. Recent work from this laboratory has been directed toward understanding the following three major aspects of the in vitro DNA synthesis catalyzed by the complete T4 DNA replication system: (1) How do the T4 proteins interact to initiate de novo synthesis of DNA strands, and what is the nature of the primer involved? (2) Does the rate of replication on duplex templates obtained in vitro approach the rate obtained in vivo, and is the energy of ATP hydrolysis utilized to drive the strand-displacement reaction? (3) How faithful is in vitro template copying, and how is this fidelity generated?

After a brief review of the proteins that comprise the in vitro replication system, the status of these three aspects of our research is discussed in the order indicated.

RESULTS

Proteins of the T4 Replication System

Isolation of six previously described proteins. The proteins encoded by T4 genes 32, 41, 43, 44, 45, and 62 have all been extensively purified from T4-infected *Escherichia coli*. DNA polymerase is the protein product of gene 43, and it has been purified to greater than 95% purity using a standard DNA polymerase assay (Morris et al., in prep.). The products of genes 41, 44, 45, and 62 have all been purified by an in vitro complementation assay (Morris et al., in prep.); this assay is based upon the stimulatory effect of small amounts of each replication protein on DNA synthesis in a crude lysate prepared from cells infected with a T4 mutant deficient in that protein (Barry and Alberts 1972). Using such an assay, it has been possible to purify gene-44, -45, and -62 proteins to greater than 95% purity and gene-41 protein to approximately 90% homogeneity. The gene-44 and -62 proteins are isolated as a tight complex (protein 44/62) which is disruptable only by strong denaturing agents and has a total molecular weight of about 180,000 daltons. The gene-43, -41, and -45 proteins have polypeptide molecular weights of 110,000, 58,000, and 27,000 daltons, respectively, and generally appear as either monomer (43 and 41) or dimer (45) species in solution (Morris et al. 1975; Alberts et al. 1977).

The T4-encoded helix-destabilizing protein (HD protein; for this nomenclature, see the recent review of Alberts and Sternglanz [1977]) is the gene-32 protein. This protein is usually isolated without recourse to an assay for its functionality in the in vitro DNA replication system; it is monitored during purification by merely following the protein absorbance profile after an initial

DNA-cellulose column step (Alberts and Frey 1970).

All of these proteins may be obtained in rather large quantities from T4-infected cells, the final yield/100 g infected-cell pellet being approximately 3 mg, 23 mg, 7 mg, and 7 mg for gene-43, -44/62, -45, and -41 proteins, respectively (Morris et al., in prep.). The gene-32 protein is vastly overproduced in cells infected with a triple T4 amber mutant in genes 33, 55, and 61 (Gold et al. 1976), from which it is obtained in a final yield of 250 mg/100 g infected-cell pellet (Bittner et al., in prep.).

Despite the high nuclease concentrations present in the initial T4-infected cell lysate, each protein has been purified to the point where it is free of detectable DNA exonuclease or endonuclease activities (<1 nick/60,000 nucleotides during a 30-minute incubation at the DNA and protein concentrations used for in vitro DNA replication).

Identification and purification of a seventh required replication protein. Combining our original purified preparations of six T4 DNA-replication proteins, striking reactions were seen on circular DNA templates; here a replication complex, once formed, can continue for very long times, producing a "rolling-circle" DNA molecule with a long, double-stranded DNA tail (Alberts et al. 1975, 1977; Morris et al. 1975). On a single-stranded (SS) circular DNA template (such as bacteriophage fd DNA), the reaction starts by the de novo synthesis of a small RNA primer, as was first suggested by the strong requirement for the ribonucleoside triphosphates rCTP, rUTP, and rGTP in the reaction, in addition to rATP. After the synthesis of one strand equivalent of DNA product, rolling-circle DNA molecules are generated (these are indistinguishable from the products formed on a nicked, double-stranded, circular template, such as bacteriophage PM2 DNA; Morris et al. [1975]).

These reactions require only catalytic amounts (10 μg/ml or less) of five of the six replication proteins. However, the gene-32 protein is required in stoichiometric amounts (100–300 μg/ml). This high concentration is presumably needed because gene-32 protein must not only coat all of the SS DNA intermediates generated in the reaction, but also help destabilize the template double helix (Alberts and Frey 1970; Alberts and Sternglanz 1977). If another protein were hidden in the gene-32 protein preparation and needed for DNA synthesis at levels of only 0.1 μg/ml or less, it would constitute only one part in 10^3–10^4 of the total protein present and therefore would escape detection, even on vastly overloaded polyacrylamide gels.

For some time, we have recognized that the activity of gene-32 protein in this replication system is very labile during a variety of chromatographic procedures designed for its further purification. The inactive gene-32 protein recovered seemed identical to the active fraction in its SS DNA-binding properties, as well as in its ability to denature the synthetic DNA double helix, poly[d(A-T)]. These results suggested that an essential minor component might be present, despite the fact that the active gene-32 protein preparation was about 99% pure (Bittner 1977). However, several attempts to restore the DNA-synthesizing activity of further-purified gene-32 protein by adding back various side fractions from the purification were all unsuccessful. In retrospect, these failures probably reflect instability of the additional protein component when present in dilute solution, free from the gene-32 protein.

A different approach to finding this protein was more successful. An aliquot of active gene-32 protein was passed through a polyacrylamide P-300 gel-filtration column, which destroyed the activity of gene-32 protein completely when assayed in our fd DNA template replication assay, even at a concentration of 260 μg/ml. As shown in Figure 1A, this "inactive" gene-32 protein could be rendered active by supplementing the reaction mix with a very small amount (17 μg/ml) of active gene-32 protein, which was by itself unable to satisfy the gene-32-protein requirement. We reasoned that there was a second stimulatory factor contributed by the "active" gene-32 protein fraction; this factor was present prior to the P-300 fractionation and was missing after this purification step. This essential factor is a seventh T4-induced protein (see below), which has temporarily been designated as "X protein."

The stringent requirement for X protein in all reactions requiring a de novo chain start (in particular the reaction on an fd DNA template, as shown in Fig. 1A) has been used as an assay to follow the purification of this protein, as described in Figure 1B. This protein appears to be unique among the T4 proteins in that only very low concentrations are required in in vitro reactions (~0.1 μg/ml is saturating as compared with >1 μg/ml for the other replication proteins). Correspondingly low levels seem to be present in vivo (perhaps only about 10 copies of X protein per cell, based on our recoveries of activity during protein purifications). Due to this low abundance, our current X-protein fractions, although purified about 5000-fold, are estimated to be about 20% homogeneous and still contain some DNA endonuclease activity.

A better purification of X-protein activity would become possible if we could identify the T4 gene encoding this protein. It should then be possible to amplify protein production, either by selection of appropriate overproducing mutants in vivo or by recombinant DNA cloning of this gene in appropriate vectors. To identify the gene for the X protein, a series of cell lysates infected with various T4 amber mutants were screened for the presence of X-protein activity. The assay measures a ribonucleoside triphosphate (rNTP)-dependent, rifampicin-resistant stimulation of DNA synthesis on an fd DNA template in the presence of the other six highly purified T4 proteins. Since this assay is blocked by inhibitors present in crude cell lysates, each lysate was tested after chromatographic fractionation on a DNA-cellulose column.

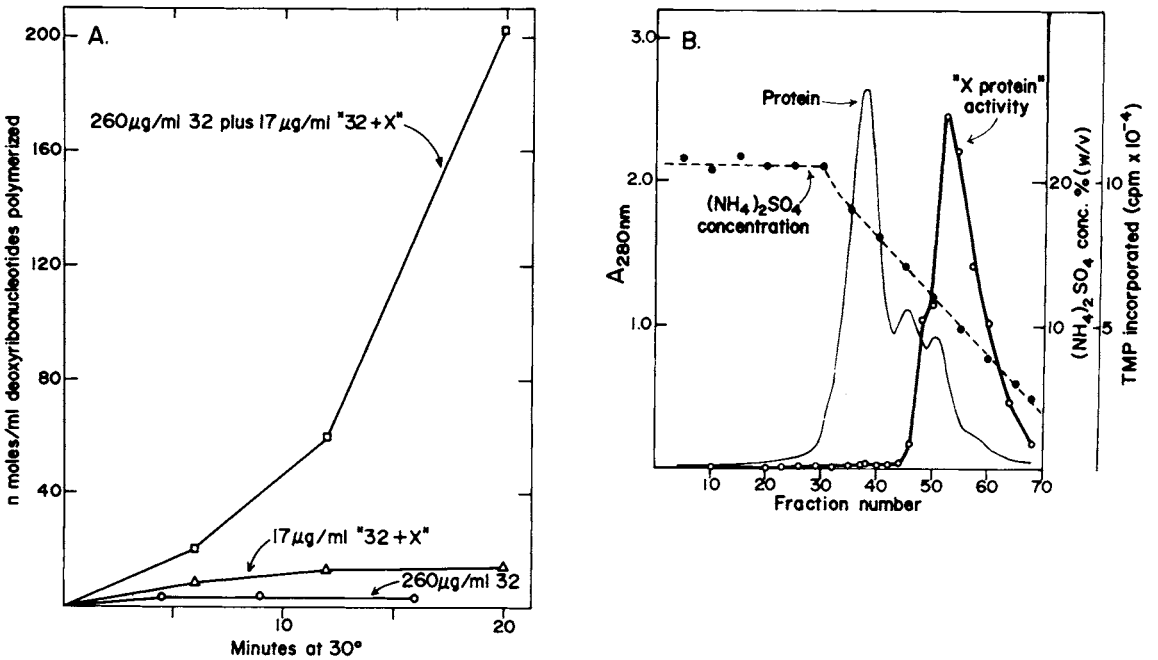

Figure 1. The development of an assay that detects a seventh essential T4 replication protein (X protein), and its use for monitoring protein purification. (*A*) The X protein is present in excess amounts in our previous gene-32 protein preparations. This assay exploits the fact that the gene-32 protein is itself required in stoichiometric levels of at least 100 μg/ml for extensive rolling-circle DNA replication on an fd DNA template. The reaction mixture in 20 μl contained 33 mM Tris-acetate (pH 7.8); 10 mM magnesium acetate; 67 mM potassium acetate; 0.5 mM dithiothreitol (DTT); 0.5 mM rATP; 0.2 mM each of rCTP, rGTP, and rUTP; 0.15 mM each of dATP, dCTP, dGTP, and dTTP; [^{3}H]dTTP (2.4×10^5 cpm/nmole); 20 μg/ml gene 44/62 protein; 10 μg/ml gene-45 protein; 20 μg/ml gene-41 protein; and 2 μg/ml gene-43 protein (henceforth defined as *standard conditions*). This reaction mixture also contained 2.1 μg/ml fd SS DNA as template. The concentrations of the "different" gene-32 proteins used are indicated: the "32 + X" protein is gene-32 protein prior to gel filtration on a polyacrylamide P-300 column; the "32" protein is the same protein after this fractionation step. Both protein preparations are more than 99% pure, as judged by electrophoretic analysis in the presence of SDS (Morris et al. 1975; Bittner et al., in prep.). The maximum synthesis seen at 20 min represents about 30 copies of the original fd DNA template added. (*B*) A hydrophobic chromatography step in the purification of X protein. Prior to this step, a pancreatic DNase-treated lysate from 150 g T4 41$^-$(*am*N81)-infected *E. coli* B was prepared and fractionated on an SS DNA-cellulose column according to previously described procedures (Alberts and Frey 1970). Fractions eluting between 0.2 M and 0.4 M NaCl contained X-protein activity; these were pooled in 36 ml of 20 mM Tris-HCl (pH 8.0), 0.1 M NaCl, 1 mM Na_3EDTA, 1 mM 2-mercaptoethanol, and 10% (w/v) glycerol. Solid $(NH_4)_2SO_4$ was added to 20% (w/v) and the sample was then loaded onto a 1×27 cm norleucine-Sepharose column (Rimerman and Hatfield 1973), equilibrated with the same buffer. Bound protein was eluted with a 125-ml reverse gradient from 20% to 0% in $(NH_4)_2SO_4$, at 25 ml/hr. The activity of X protein was monitored after 60-fold dilution of each column fraction in the assay in *A*, using 10 min of incubation at 30°C and gene-32 protein devoid of X-protein activity at 200 μg/ml. Subsequent to this step, X protein was further purified by successive steps of ion exchange column chromatography (DE-52, DEAE-cellulose; Whatman) and gel filtration (Sephacryl S-200; Pharmacia). The final protein obtained was used in all subsequent experiments with X protein to be reported.

Some results of such screening experiments are presented in Table 1. The X-protein activity is clearly absent from uninfected *E. coli* cells and from cells infected with two different T4 mutants in the "DNA-delay" gene 58-61 (initially identified as two independent complementation groups by Epstein et al. [1964], but since remapped as a single gene by Yegian et al. [1971]). Activity was found to be regained in a spontaneous revertant of one of these mutants. These results, corroborated by those of Silver and Nossal (this volume), indicate that gene 61 is essential for the expression of X-protein activity. Although further work is needed to make this assignment more convincing, the lack of pleiotropic effects in most T4 mutations suggests that the X protein is probably the product of gene 61.

In contrast to the essentially DNA-negative phenotype of mutants in the other replication proteins, substantial T4 DNA synthesis is seen after an initial delay period in cells infected with gene-61 mutants (Yegian et al. 1971). In addition, the first round of T4 DNA synthesis appears to occur normally, as judged by experiments which monitor density-labeled parental phage DNA (Mosig et al., this volume). Therefore, some other *E. coli* host or phage-induced protein seems to be able to substitute for X-protein function in vivo, although rather poorly.

Partial reactions catalyzed by subsets of the seven T4 replication proteins. Some properties of the seven proteins which presently comprise the T4 in vitro system are summarized in Table 2. Also listed in Table 2 is the approximate concentration of each protein which is optimal for in vitro DNA synthesis. It is worthwhile noting that these concentrations are not greater than the estimated in vivo levels for each

Table 1. Assay of X-protein Activity Which Has Been Partially Purified from Uninfected and T4-infected *E. coli* Cells

Source of X activity	Fold stimulation[a]
Uninfected cells	0.7
T4 (wild-type)-infected cells	14.2
41⁻(*am*N81)-infected cells	8.3
30⁻, 42⁻, *regA*⁻ (*am*H39, *am*N55, *sp*62)-infected cells	6.6
45⁻, *regA*⁻ (*am*E10, *sp*62)-infected cells	4.2
61⁻ (*am*HL627)-infected cells	0.5
61⁻(*am*E219)-infected cells	0.6
Cells infected with a spontaneous wild-type revertant from the 61⁻(*am*E219) mutant	3.8
Purified-X-protein control	15
Buffer control	0

As indicated, X-protein activity is detected by the rNTP-dependent stimulation of incorporation obtained in the fd DNA synthesis assay, using a reaction mix containing the other six purified T4 DNA replication proteins. For preparation of the test extract, one liter of *E. coli* B was grown to a cell density of 3×10^8 cells/ml in M9 medium supplemented with casamino acids and infected with the indicated phage at a multiplicity of 5–10. The infection was allowed to proceed for 10 min at 37°C for the wild-type phage or 30 min at 37°C for the amber mutants. Cells infected with the 61⁻ DNA-delay mutants were also collected after 40 min of growth at 25°C (Mufti and Bernstein 1974), with substantially the same results. Cell lysates were prepared by sonic disruption followed by DNase-I digestion. The dialyzed lysate was loaded onto a 1-ml SS DNA-cellulose column and batch-eluted with steps of 0.2 M, 0.6 M, 1.0 M, and 2.0 M NaCl in buffer (Alberts and Frey 1970). Those protein-containing fractions eluting at the 0.6 M step were tested for stimulatory activity in a standard fd DNA-synthesis assay (Fig. 1A) using 150 μg/ml gene-32 protein devoid of X-protein activity. The 25-μl reaction volume also contained rifampicin at 30 μg/ml and either 0.25 or 0.5 μl of the fraction to be tested. Where indicated, purified X protein was added at 0.1 μg/ml. The assay mixture was incubated for 5 and 10 min at 30°C, and aliquots were quantitated for acid-insoluble DNA synthesis by standard liquid scintillation counting techniques.

[a]Fold stimulation is defined as the stimulation of DNA synthesis observed in the fd DNA synthesis assay upon addition of rGTP, rCTP, and rUTP and is equal to the ratio:

$$\frac{\text{Synthesis plus ribonucleotides} - \text{synthesis minus ribonucleotides}}{\text{synthesis minus ribonucleotides}}$$

protein. These proteins function together to propagate a replication fork, whose geometry is schematically illustrated in Figure 2 A and B.

DNA polymerase (gene-43 protein) has an intrinsic 3′→5′ exonucleolytic activity in addition to its 5′→3′ polymerizing activity. This exonuclease is likely to have a "proofreading" function in removing replication errors (Brutlag and Kornberg 1972). The protein products of genes 44 and 62 have an ATPase activity which is stimulated strongly both by SS DNA ends and by the addition of gene-45 protein (Alberts et al. 1975, 1977; Mace, 1975; Piperno and Alberts 1978; Piperno et al. 1978). DNA polymerase activity on SS DNA templates is markedly enhanced by the addition of gene-45 and gene-44/62 proteins. This enhancement absolutely requires ATP (or dATP) hydrolysis, and both the rate of polymerase travel and its processivity are increased (Mace 1975; Alberts et al. 1975). The gene-44/62 and gene-45 proteins are believed to travel along with the DNA polymerase and hence are designated as "polymerase-accessory" proteins.

As described below, the products of genes 41 and X are essential for obtaining rNTP-dependent de novo initiation of new DNA chains; in addition, these two proteins alone catalyze the synthesis of short RNA oligonucleotides on an SS DNA template. The gene-41 protein and the X protein are therefore designated as "RNA-priming" proteins. The gene-41 protein has an intrinsic GTPase and ATPase (and dGTPase and dATPase) activity which is stimulated by SS DNA. However, although the X protein binds strongly to DNA, the gene-41 protein must interact with DNA rather weakly, since its direct association with DNA has not been detected, even in the presence of the other replication proteins.

There is a striking tendency for functionally interacting T4 gene products to map in adjacent genes (Stahl 1967). Thus, among the replication proteins, the genes for the polymerase-accessory proteins (genes 44, 62, and 45) are adjacent both to each other and to the gene for DNA polymerase (gene 43). Identifying X protein with gene 58-61, which maps next to gene 41 (Fig. 2C), correlates with the close functional interaction of these two proteins, to be shown below.

The other protein listed in Table 2 is the HD protein (gene-32 protein). This protein completely coats all SS DNA during replication reactions and

Table 2. Some Properties of Seven Bacteriophage T4 Proteins Required for Efficient In Vitro DNA Replication

	DNA polymerase	Helix-destabilizing protein	Polymerase-accessory proteins		RNA-priming proteins	
	(gene 43)	(gene 32)	genes 44 and 62	gene 45	gene 41	X (gene 61?)
Activities	5′→3′ polymerase 3′→5′ exonuclease	cooperative binding to SS DNA	SS DNA termini-dependent rATPase, dATPase		long SS-DNA-dependent rGTP,dGTPase, rATP,dATPase	binds DNA
Optimal concentration (μg/ml)	2	100	20	10	20	0.1
Current purity (%)	99	99	99	95	90	20?
Molecular weight (kD)	110	35	4 × 34 2 × 20	2 × 27	58	20?

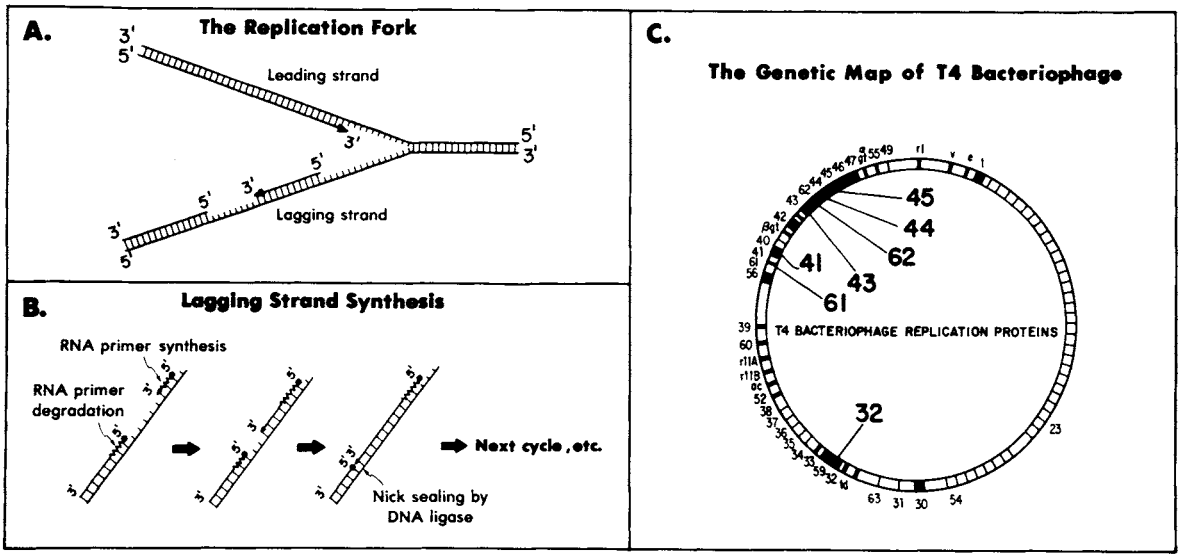

Figure 2. (*A*) Schematic representation of a replication fork to emphasize the asymmetry between the synthesis on the leading and lagging strands. Only lagging-strand synthesis should require de novo chain starts primed by an RNA oligonucleotide. In T4 these starts seem to be spaced about one start every 2000 nucleotides (Okazaki et al. 1969). (*B*) Lagging-strand synthesis in more detail; a cycle of synthesis and degradation of the RNA primer believed to start Okazaki fragments (Kornberg 1974). (*C*) The genetic map of T4 with replication genes emphasized. The broadest black segments indicate the relative locations of mutationally identified genes with the most drastic effect on DNA synthesis in vivo at 37°C. Mutations in genes 32, 41, 43, 44, 45, and 62 are unique in that they synthesize very little DNA, even though all four dNTP precursors are present. Gene-61 mutants (also denoted as gene-58 mutants) synthesize substantial amounts of DNA, but only after a delay period. Note the strong tendency for the genes coding for functionally interacting replication proteins (see Table 2) to cluster on this map. (Modified from Wood and Revel 1976; W. Wood, pers. comm.)

holds it in a conformation which leaves the DNA bases exposed for base pairing with incoming monomers (Huberman et al. 1971). The gene-32-protein–DNA complex would appear to be specifically recognized by the T4 replication apparatus, since the *E. coli* analog of this HD protein fails to substitute for it in any of our reactions (Sigal et al. 1972; R. L. Burke, unpubl.). Genetic analyses suggest that the gene-32 protein is also intimately involved with proteins which function in other aspects of T4 DNA metabolism (Mosig et al., this volume).

It is possible that other T4-induced proteins function at in vivo replication forks in addition to the seven listed in Table 2. Certainly, the present system does not demonstrate selectivity for initiating replication forks on a T4 DNA template, as found in vivo. Moreover, our recent discovery of X protein should alert us to the possibility that other essential factors, required in very low amounts, lie hidden in our present replication-protein mixtures. Finally, it has been suggested that the replication complex in vivo interacts closely with a second large complex of different enzymes, which may supply dNTP substrates directly to the replication fork (Chiu and Greenberg 1969; Chiu et al. 1976; Reddy et al. 1977).

Minimal requirements for DNA synthesis on double-helical templates. Even though the binding strength of gene-32 protein is insufficient by itself to open the DNA helix (Alberts and Frey 1970; Jensen et al. 1976), this protein seems to play an important role in the helix-destabilization process required for rapid DNA synthesis on DS DNA templates. For example, whereas no synthesis occurs on a nicked double helix with the T4 DNA polymerase alone, addition of high concentrations of gene-32 protein at low ionic strength makes a limited amount of DNA synthesis possible at such a nick (Nossal 1974). However, to observe any significant amount of DNA synthesis on double-helical templates at physiological ionic strengths, five proteins (43, 32, 45, 44, and 62), plus ATP hydrolysis to ADP and inorganic phosphate, are required. The requirement for ATP hydrolysis is shown in Figure 3A for a reaction on a nicked T7 DNA template. Further characterization of this "five-protein" reaction demonstrates that all of the DNA product is template-linked and that the polymerase start sites are randomly located on the T7 genome with reference to a T7 restriction map (J. Barry, unpubl.). We therefore believe that the T4 DNA polymerase first binds to a preexisting random nick present in the T7 DNA and then picks up its accessory proteins 44/62 and 45. With the help of gene-32 protein, it then proceeds to polymerize deoxyribonucleoside triphosphates (dNTPs) onto the 3′ end of the DNA chain.

Simultaneously with polymerization, the 5′ end of the parental strand opposite the nick must be displaced from its base-paired partner strand in the parental double helix. It is probably this reaction which requires the presence of substantial quantities of gene-32 protein (100 μg/ml) and the ATP hydrolysis by the gene-44/62 protein. However, unlike the somewhat analogous ATP hydrolyses by the *E. coli rep* protein (Scott et al. 1977; Eisenberg et al., this vol-

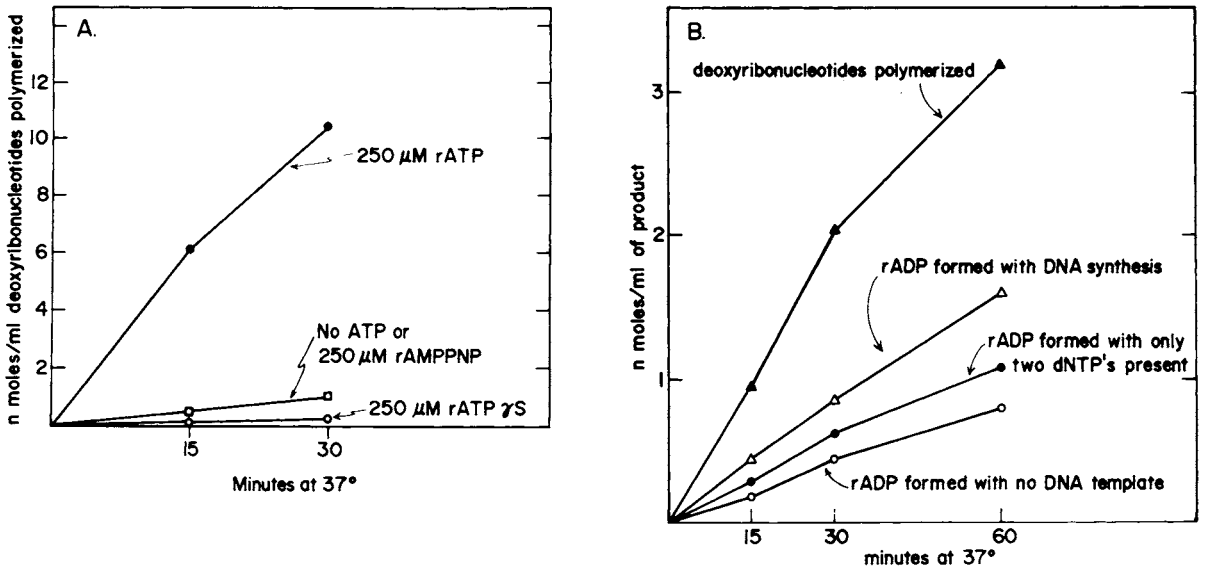

Figure 3. The hydrolysis of ATP in a five-protein reaction on a double-helical DNA template. (*A*) Evidence that hydrolysis of ATP to ADP and inorganic phosphate is required for DNA synthesis. T7 DNA was used as template at 5 μg/ml. DNA was synthesized under standard conditions (Fig. 1A), with gene-43, -32, -44/62, and -45 proteins present, except that 200 μM dAMPPNP (deoxyriboadenyl-5′-yl imidodiphosphate) was used in place of dATP as the fourth deoxyribonucleotide substrate. Since this dNTP analog is hydrolyzable only at its α-β pyrophosphate linkage, it serves as a substrate for DNA polymerase, but not for the gene-44/62 ATPase activity; the gene-44/62 protein would otherwise hydrolyze dATP in place of rATP, masking detection of any ATP requirement (Piperno and Alberts 1978). The labeled nucleotide used to monitor synthesis was [^{3}H]dTTP (0.5 Ci/mmole). As indicated, neither of the nonhydrolyzable ATP analogs, rAMPPNP or rATPγS, substitute for rATP at 250 mM. Further experiments of this type demonstrate that rATPγS is a strong competitive inhibitor of the ATP stimulation of this reaction, with a K_i that is four- to tenfold lower than the K_m for ATP (J. Barry, unpubl.). We conclude, therefore, that ATP hydrolysis, rather than just ATP binding, is required in the five-protein reaction on double-helical templates. (*B*) The stoichiometry of ATP hydrolysis. DNA was synthesized and detected in the five-protein reaction as in *A*, but with gene-44/62 and -45 proteins present at 1.5 μg/ml each. As the ATP source, the reaction mixture contained 250 μM [γ-^{32}P]ATP (0.5 Ci/mmole). To assay for the ATPase activity of gene-44/62 protein, 6-μl aliquots of the reaction mixture were pipetted into norit-HCl and processed essentially as described by Zimmerman and Kornberg (1961). The radioactivity of norit nonadsorbable phosphate released was determined by counting in a Triton X-100 : toluene scintillation fluid (1 : 2). DNA template or dCTP and dGTP were omitted where indicated to block DNA synthesis, revealing a substantial background of uncoupled ATP hydrolysis. All of the ATP hydrolyses shown occur at a low level with gene-44/62 protein alone and are strongly stimulated when gene-45 protein is present (Mace 1975; Piperno and Alberts 1978).

ume) and by the T7 gene-4 protein (Kolodner and Richardson 1977; Lanka et al.; Richardson et al.; both this volume), the number of ATP molecules hydrolyzed per deoxyribonucleotide polymerized is much less than 1, as shown in Figure 3B. As discussed in detail elsewhere, the quantities of ATP hydrolyzed by the gene-44/62 protein in this reaction are insufficient to provide the sole driving force for helix unwinding, and probably the sum of several different contributions forces the DNA helix to open ahead of the fork (Alberts and Sternglanz 1977).

Classification of "five-protein" and "seven-protein" reactions. As just described, the five proteins, 43, 32, 45, 44 and 62, catalyze DNA synthesis on a double-helical DNA template which is accompanied by strand displacement of the parental DNA and is therefore identical in its geometry to the leading-strand DNA synthesis at a replication fork, shown in Figure 2A. Because the RNA-priming proteins (gene-41 protein and X protein) are missing, the displaced parental DNA should remain single-stranded. Without de novo starts on this "lagging strand," the addition of UTP, CTP, and GTP to this five-protein reaction has no effect on DNA synthesis, as shown for synthesis on a T7 DNA template in Figure 4A.

When gene-41 protein and X protein are added to the above proteins (Fig. 4), one sees about a twofold difference in incorporation upon addition of rUTP, rCTP, and rGTP to the seven-protein reaction (as expected from the de novo starts made possible by these ribonucleotides on the lagging strand being displaced). But surprisingly, these proteins also greatly stimulate DNA synthesis in the absence of these ribonucleotides. This stimulation is especially marked with T4 DNA, which serves as a very poor template for synthesis in a five-protein reaction (Fig. 4B) (Alberts et al., 1975, 1977). These data are consistent with the findings (discussed below) that the gene-41 protein (and X protein?) has a second role in leading-strand replication reactions, in addition to its priming function required on the lagging strand.

Synthesis of an RNA Primer by T4 Replication Proteins 41 and X

As previously discussed, the T4 in vitro replication system shows a strong rNTP requirement for DNA

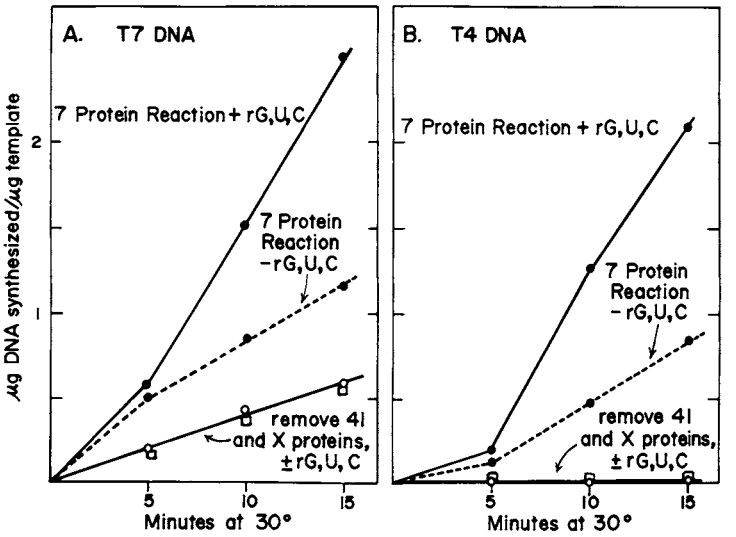

Figure 4. Comparison of seven-protein and five-protein reactions on T4 and T7 DNA templates. The time course of DNA synthesis was measured in a standard DNA polymerization reaction, using the protein, triphosphate, and buffer concentrations in Fig. 1A, with purified gene-32 protein and X protein as in Table 1. (*A*) T7 DNA at 15 μg/ml used as template; (*B*) T4 DNA at 12 μg/ml used as template. To test the rNTP-dependent stimulation of DNA synthesis in the seven-protein reaction, the commercial rATP used was purified by chromatography on DEAE-Sephadex. Subsequent experiments, utilizing a T4 DNA template nicked to varying extents with DNase I, show an increased synthesis in a five-protein reaction. However, a strong effect of addition of gene-41 protein is still obtained, even without rG, U, C present.

synthesis on those templates on which de novo DNA chain starts are needed. These results suggest that a new chain start occurs by RNA priming. To study the priming mechanism directly, a circular, single-stranded fd DNA template was employed, since a de novo initiation of DNA synthesis is required on this template due to its lack of a 3'-OH terminus. Moreover, the complete fd DNA sequence and restriction map are available (Schaller, this volume). A time course of incorporation for the complete system (seven proteins) is shown in Figure 5A. To achieve extensive DNA synthesis (16 copies of the template in 10 min in this example), all the purified replication proteins and rNTPs are absolutely required. The omission of rCTP, rUTP, and rGTP or of any of the protein components except gene-32 protein results in virtually no DNA synthesis. The omission of gene-32 protein leads to a limited amount of synthesis which stops when one equivalent of the template is made. As shown in Figure 5B, this synthesis requires, in addition to rATP, the presence of rCTP, rUTP, and rGTP. (The strongest rNTP requirement is for rATP and rCTP. In the presence of only these two ribonucleotides, the reaction proceeds to the same extent, at about half the rate.) This ribonucleotide requirement suggests that, without gene-32 protein, synthesis also starts with an RNA primer, but stops at the point where a double-helical template is first encountered.

T4 DNA is, of course, the natural template for our replication proteins. Results very similar to those

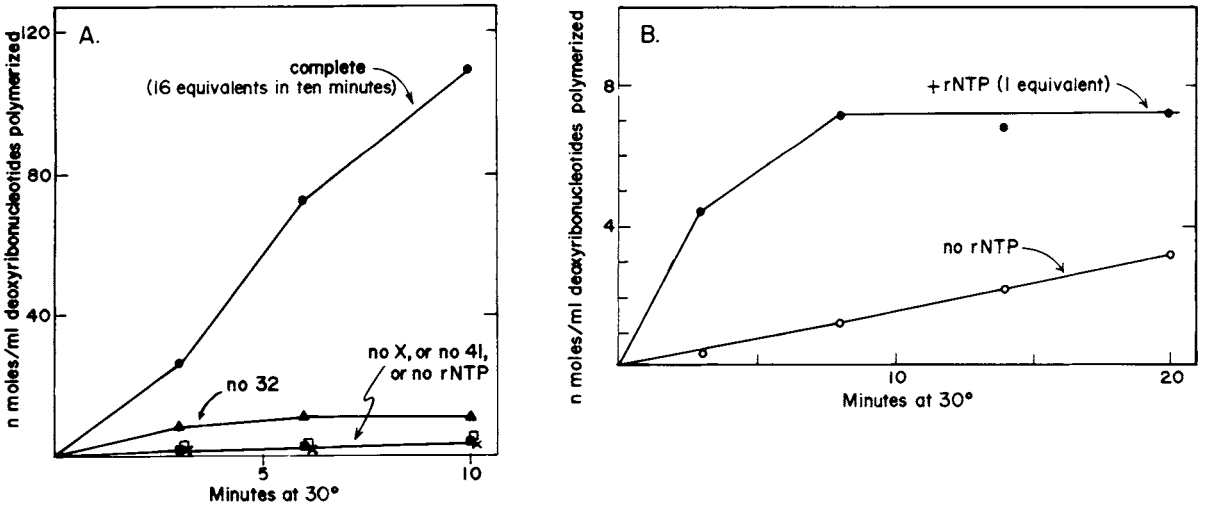

Figure 5. For synthesis of DNA on a single-stranded, circular fd DNA template, gene-41 protein and X protein are required, plus rNTPs in addition to rATP. (*A*) Requirements for the complete reaction. Except for the indicated omissions, the reaction mixture contained those components described in Fig. 1A, with gene-32 protein purified free of X activity added at 120 μg/ml. The amount of X protein used was empirically optimized to 1 μg/ml total protein (about 0.2 μg/ml of X protein). Note that the reaction stops in the absence of gene-32 protein when one template equivalent of DNA has been made. (*B*) The single round of DNA synthesis observed in the absence of gene-32 protein requires rNTPs in addition to rATP. Except for the omission of gene-32 protein, reaction components (including six T4 proteins) were as described for Fig. 1, with fd DNA template present at 5 μg/ml. Note that for the indicated reaction with no rNTP, only rCTP, rGTP, and rUTP were removed; thus, 0.5 mM rATP was still present.

shown in Figure 5 are obtained if T4 SS DNA is used as a template rather than fd DNA (Alberts et al. 1975). However, because of its much greater length (200×10^3 instead of 6×10^3 bases) and the absence of DNA sequence information for T4, we have chosen to concentrate on the simpler fd DNA template for these initial studies.

We find that, if the fd DNA template is artificially primed by purified *E. coli* RNA polymerase, the requirements for gene-41 protein, X protein, and rNTPs for the first round of DNA synthesis in the absence of gene-32 protein can be eliminated (C. C. Liu, unpubl.). This result suggests that one major function of gene-41 protein and X protein is RNA primer synthesis on the fd SS DNA circle. Primer synthesis is shown directly by the experiment in Figure 6, where RNA oligonucleotides are made in a reaction mixture containing only fd DNA, gene-41 protein, X protein, and rNTPs. Omission of any one of these four components abolishes RNA synthesis. The properties of the RNA product are briefly summarized as follows:

1. rATP and rCTP are absolutely required, each with a K_m in the reaction of at least 50 μM. Addition of rGTP and rUTP further stimulates synthesis about twofold.
2. The products are short oligonucleotides, 6–8 nucleotides long, as judged by electrophoresis in 20% polyacrylamide gels according to Maxam and Gilbert (1977); as many as 40 oligonucleotides are made per fd DNA circle in a 10-minute reaction.
3. Synthesis is reduced by rGTPγS (riboguanosine 5′-0-[3-thiotriphosphate]) under conditions in which rGTP itself is not required. The rGTPγS inhibits rATP hydrolysis by gene-41 protein under these conditions. Thus, rATP hydrolysis to rADP plus inorganic phosphate by gene-41 protein (or a similar rGTP hydrolysis when this nucleotide is present) is apparently required for making primer.
4. As judged by the incorporation of [^{3}H]rCTP, the reaction is not affected by the presence of excess gene-32 protein, rifampicin, or rNDPs, and it is only slightly inhibited (~10%) by dNTPs added at equimolar concentrations to the rNTPs.
5. When RNA oligonucleotides are synthesized in the presence of rATP and [α-^{32}P]rCTP and the product is then treated with phosphatase, only eight different species of oligonucleotides are found by one-dimensional paper electrophoresis (DE-81 paper in pyridine-acetate solvent, pH 3.5). Without phosphatase treatment, all of these oligonucleotide products remain at the origin in this system, which is expected if they have a 5′-triphosphate end. Combined alkaline hydrolysis and pancreatic RNase-A digestion on separated, phosphatase-treated oligonucleotides reveals that the most abundant sequences are -ACA- and -ACC-. The complexity of the sequences increases when rGTP and rUTP are included during the RNA synthesis.
6. The RNA made is readily released from the DNA template, since we have not been able to isolate oligoribonucleotides associated with the DNA by standard gel-filtration procedures.

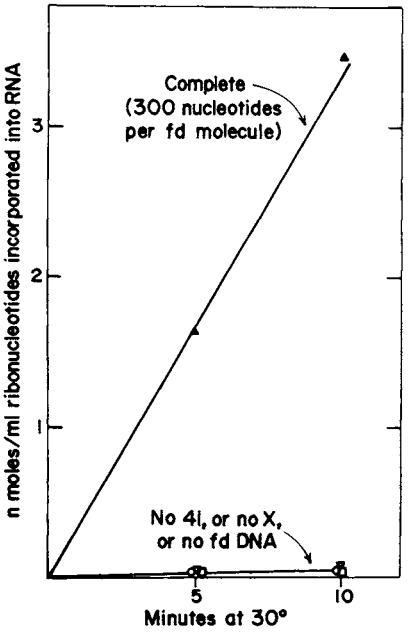

Figure 6. Only X protein, gene-41 protein, SS DNA, and rNTPs are required for the synthesis of small oligoribonucleotides. The reaction mixture contained the standard buffer components (see Fig. 1A); 0.2 mM each of the four rNTPs, including [^{3}H]rCTP at 10 μCi/nmole; 20 μg/ml of single-stranded, circular fd DNA; and, for proteins, only gene-41 protein (80 μg/ml) and X protein (5 μg/ml of total protein). Incorporation was monitored by spotting aliquots onto DE-81 filter paper disks which were then washed free of triphosphates as described by McMacken et al. (1977). This assay was used because the products are too short to be detected by acid insolubility.

To identify the number and location of de novo DNA chain starts on an fd DNA template, we have carried out a DNA-synthesis reaction in the absence of gene-32 protein, where the reaction proceeds to completion when the template has been fully duplicated once (Fig. 5B). This reaction should generate a duplex circular DNA with one nick in the progeny strand where the RNA primer was made. Figure 7A shows the result of agarose gel electrophoresis of the newly synthesized DNA strand. As expected, much of this DNA has an electrophoretic mobility under denaturing conditions indistinguishable from that of a complete linear fd DNA molecule. To locate the position of the nick, which indicates where the de-novo-initiated DNA chain started, this product was digested with restriction endonuclease *Hin*dII, which makes a single cut in the fd DNA genome (Schaller et al., this volume). The cut product DNA was then denatured and subjected to agarose gel electrophoresis, with the result shown in Figure 7B. The unique band seen before digestion (Fig. 7A) has disappeared. In its place, we observe a series of radioactive DNA bands shorter than unit length, each pair representing a

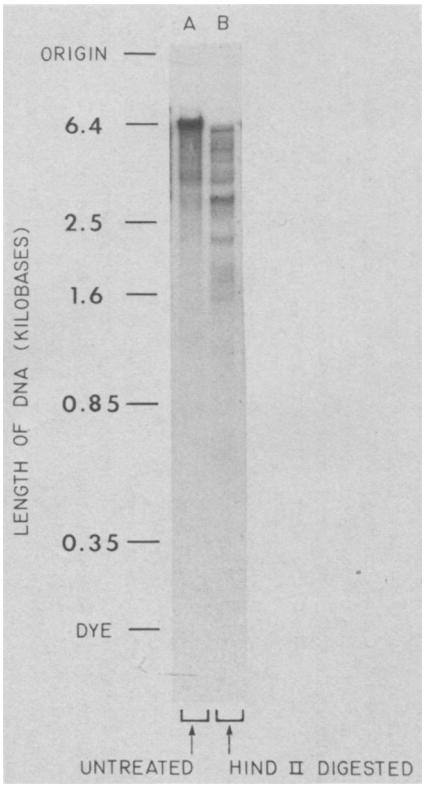

Figure 7. Agarose gel electrophoretic analysis of the DNA product chains made in the complete reaction on an fd DNA template in the absence of gene-32 protein. A reaction was carried out as described in Fig. 5B, but with only rATP and rCTP present as rNTPs and with [α-^{32}P]TTP as the labeled substrate. After 7 min at 30°C, the reaction was stopped by addition of 10 mM Na_3EDTA and 1% Sarkosyl (Geigy Chemical). DNA was deproteinized with phenol and separated from triphosphates by gel-filtration chromatography. After half of the sample was digested to completion with *Hin*dII restriction endonuclease, the digested and undigested aliquots were ethanol-precipitated to remove salt and redissolved in 0.5 mM borate buffer, pH 8.9. These samples were then heated to 95°C for 3 min, quickly cooled, and loaded onto a 1.5% agarose slab gel for electrophoresis in a low-salt buffer designed to minimize reannealing of the heat-denatured DNA chains (Perlman and Huberman 1977). Autoradiography was used to detect the ^{32}P-labeled DNA products in the dried agarose gel. After *Hin*dII digestion, each uniquely started unit-length DNA chain should give rise to a pair of unique DNA fragments, the sum of whose lengths is 6389 bp. Thus, at least four different start sites on the fd DNA circle are revealed by the analysis shown.

different unique start site on the fd DNA genome. From this result, we conclude that the de novo chain starts in this system do not occur randomly, but rather that they occur at more than four different "specific" sites on the fd DNA molecule.

Direct evidence that RNA has been used to prime DNA chain starts can be obtained by following radioactively labeled rNTP incorporation into the DNA product. When experiments of this type are carried out in the presence of [α-^{32}P]rCTP plus the other three rNTPs, label is incorporated into what appears to be unit-length DNA product chains, as judged by an agarose gel electrophoretic analysis identical to that shown for DNA label in Figure 7. This RNA label appears in a set of discrete smaller DNA fragments, when the circular DNA product is cut once with the *Hin*dII enzyme and then analyzed in denatured form. Exhaustive DNase-I treatment of such products releases labeled RNA oligonucleotides, many of which migrate identically on a 20% acrylamide gel to the RNA oligonucleotides made in the absence of DNA synthesis. Finally, DNA strand lengths are not reduced by alkaline hydrolysis of this attached RNA, demonstrating that the RNA is incorporated at one end of a DNA strand (C. C. Liu, unpubl.).

By isolating individual RNA-labeled DNA fragments from a gel such as that shown in Figure 7B, each initiation site on the fd genome can be precisely located by available rapid sequencing techniques (Donis-Keller et al. 1977; Maxam and Gilbert 1977).

Replication Rates on Double-helical DNA Templates: A Major Effect of the Priming Proteins

Except for the polymerase and the helix-destabilizing protein, the T4 replication proteins in Table 2 might be imagined to function as two independent sets, with the polymerase-accessory proteins operating only on the leading strand and the RNA-priming proteins operating only on the lagging strand of a replication fork (see Fig. 2A,B). However, as discussed previously, the polymerase accessory proteins (44/62 and 45) have a major effect on DNA-polymerase activity on single-stranded, as well as on double-stranded, DNA templates. This effect suggests that these three polymerase-accessory proteins function on *both* the leading and lagging strands in vivo. Analogously, the requirement for gene-41 protein (or gene-41 protein plus X protein in Fig. 4B) for substantial DNA synthesis by strand displacement on a T4 DS DNA template (Alberts et al. 1975, 1977) implies that the RNA-priming proteins (and especially the gene-41 protein) play an important role on the leading side of a replication fork, in addition to having a separate primer-making function on the lagging strand.

To test the mechanism of the priming-protein effect directly, we have developed a new method for measuring the in vitro rate at which the leading strand of a replication fork moves through a parental double helix. The principle of this method is schematically outlined in Figure 8A. A very brief polymerization reaction is carried out in which a leading-strand type of DNA synthesis is primed on a preexisting nick on a double-helical template (in the results presented, the 40,000-nucleotide-long T7 DNA molecule was used as template). High-specific-activity dTTP is used to radioactively label the DNA product, and 5-hydroxymethyl dCTP (dHMCTP) replaces dCTP as the fourth DNA precursor nucleotide. (dHMCTP, the natural DNA precursor for the T4 system, is utilized

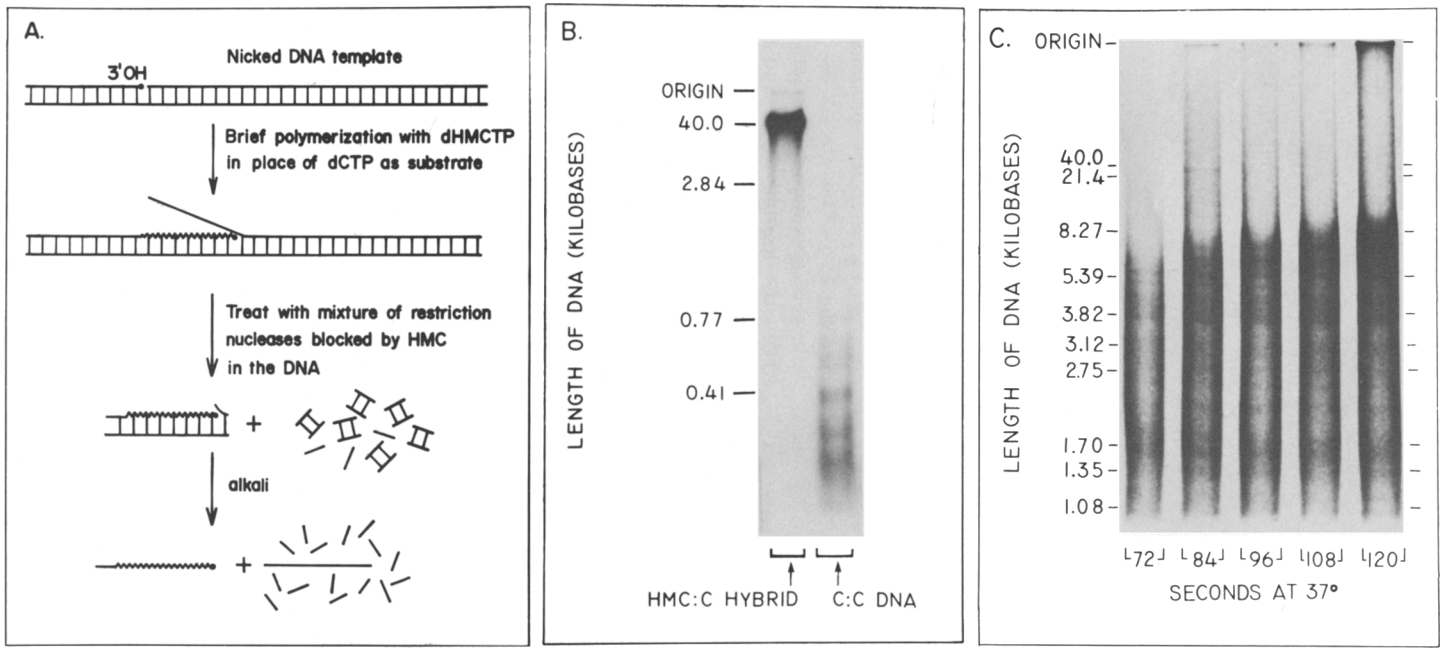

Figure 8. Autoradiographic analysis of various ^{32}P-labeled T7 DNAs electrophoresed through agarose gels following restriction-enzyme treatment. (*A*) Schematic outline of a new method for measuring DNA polymerase rates on a double-helical DNA template. The newly synthesized, radioactive DNA product chain is indicated by the zigzag line. The length of this product chain is measured as a function of polymerization time by exploiting the fact that some restriction endonucleases will not cut DNA unless it has C residues on both DNA strands (see *B* and *C*). (*B*) Demonstration that in-vitro-synthesized double-helical DNA which contains HMC in place of C on one strand is completely resistant to a restriction-enzyme mixture. Using a T7 DNA template at 5 μg/ml, DNA was synthesized for 10 min at 37°C in 1 ml standard reaction mixture containing 500 μCi [^{32}P]dTTP (345 Ci/mmole) and either the standard concentration of 160 μM dCTP or 160 μM 5-hydroxymethyl dCTP. The amount of product DNA was equivalent to only 10% of the input template DNA; thus, nearly all of this DNA will be paired with nonradioactive, C-containing parental template chain. After deproteinization, 300 ng of the DNA in 50 μl was treated with 1 unit of *Hpa*II, 0.1 unit of *Hae*III, and 2 units of *Alu*I restriction enzymes for 1 hr at 37°C in a buffer consisting of 20 mM Tris-HCl (pH 7.4), 0.25 mM Na_3EDTA, 6 mM $MgCl_2$, and 10 mM 2-mercaptoethanol. Next, 50 mM NaCl and 0.4 unit of *Hha*I were added, and the incubation continued for an additional hour. The restricted DNA was electrophoresed on a 2% agarose slab gel for 3.5 hr at 50 mA in 40 mM Tris-acetate (pH 8.1), 20 mM sodium acetate, and 2 mM Na_3EDTA. Nucleases were obtained from New England BioLabs. (*C*) Measurement of the rate of fork movement on a double-helical template in a five-protein reaction. DNA was synthesized in the standard reaction mixture containing 160 μM of each of the four dNTPs (with dHMCTP substituted for dCTP), 5 μg/ml T7 DS DNA, and 300 μCi [^{32}P]dTTP (350 Ci/mmole) per ml reaction mixture. The T4 replication proteins were used in the following concentrations: 4 μg/ml gene-43 protein, 90 μg/ml gene-32 protein, 10 μg/ml gene-44/62 protein, and 40 μg/ml gene-45 protein. Proteins, DNA, buffers, and salts were preincubated separately from the dNTPs for 2 min at 37°C. Triphosphates were then added to start the reactions. Syntheses in a 50-μl volume were stopped at the indicated times by the addition of 20 μl of 50 mM Na_3EDTA and 25 mM NaOH. This DNA was deproteinized by the addition of 0.1% SDS and extracted with redistilled, neutralized phenol. To remove phenol and unincorporated triphosphates, the DNA in the aqueous phase was sedimented through a 0.3-ml column of Sepharose CL6B according to the method of Neal and Florini (1973). The breakthrough was then treated with a mixture of restriction enzymes as described in *A*. Electrophoresis was done for 4 hr at 50 mA on a 0.8% agarose slab gel using the buffer described in *A*. For both *A* and *B*, DNA fragments of known length, ranging from 40,000 bp to 300 bp, were electrophoresed alongside of the samples to calibrate the gel. Gels were air-dried onto DE-81 filter paper under light pressure and then autoradiographed using a Dupont Lightning Plus intensifying screen with Kodak X-Omat film.

indistinguishably from dCTP in the T4 in vitro reactions.) The substitution of HMC for C on one of the two strands of a double helix renders the DNA product resistant to cutting by a group of DNA restriction endonucleases, selected to be specific for C-containing DNA. After the brief reaction has been stopped and the DNA product deproteinized, treatment of the total DNA with these restriction nucleases digests away all but the newly synthesized piece of helix. The length of this nuclease-resistant DNA product is measured by agarose gel electrophoresis, either as DS DNA or as SS DNA, to determine the number of bases polymerized per second at the chain growing point.

An essential aspect of this technique is that a set of DNA restriction endonucleases must be found which will cut normal C-containing DNA into so many different pieces that the vast majority of fragments are less than 400–800 nucleotides long. In addition, these nucleases must be inactive on DNA helices containing HMC in place of C on one strand. All of the commercially available restriction nucleases with a 4-bp recognition site, at least one of which is a G-C base pair, were tested on a special substrate prepared by replicating T7 DS DNA molecules in such a way as to produce long continuous stretches of ^{32}P-labeled product DNA, either with or without HMC substitution. A mixture of four restriction enzymes (*Hpa*II, *Hae*III, *Alu*I, and *Hha*I) was found which cuts C-containing T7 DS DNA into pieces smaller than 770 bp in length, without cutting HMC: C hybrid T7 DNA (Fig. 8B). As indicated in Figure 8A, two of these nucleases (*Hae*III and *Hha*I) also cut C-containing SS DNA, thereby trimming off the displaced parental single-strand from the DNA product.

We first used the method outlined in Figure 8A for measuring polymerization rates in the five-protein reaction. An autoradiograph of a gel is shown in Figure 8C to illustrate the type of data obtained. Discrete products are not seen, since new replication forks continue to form throughout the reaction. Instead one sees a continuum of product sizes, with a fairly clear "front," representing the largest molecules present at each time point. As expected, the ^{32}P-labeled product DNA size, measured from the position of this front, increases continuously during the incubation period.

In Figure 9A, the estimated size of the longest DNA molecules present at each time point is plotted as a function of the time of incubation at 37°C. The slope of this straight line, which gives the rate of fork movement through the double-helical T7 DNA template, is indistinguishable whether measured on neutral or alkaline agarose gels, representing a rate of only 70–80 nucleotides/sec. Extrapolation of each line to near zero chain length at zero time indicates that the chain growth rate is essentially constant throughout the course of this five-protein reaction; in other words, the maximum rate is attained immediately without any lag.

This rate is only about 10% of the estimated in vivo rate of T4 replication fork movement (~800 nucleotides/sec) (McCarthy et al. 1976). Moreover, much greater in vitro rates were estimated earlier for a seven-protein reaction, calculated from the average amount of DNA synthesized per rolling-circle intermediate on a circular DNA template (Morris et al. 1975). This discrepancy is explained by the finding that the addition of gene-41 protein and X protein to the five-protein mixture dramatically increases the rate of fork movement measured by our assay. However, quantitation is made difficult by the fact that there is a lag of more than 1 minute before the five-protein rate of fork movement begins to be noticeably accelerated under normal seven-protein reaction conditions. As a result, the acceleration induced by gene-41 protein and X protein affects growing chains that are already about 5000 nucleotides long and rapidly extends them beyond the size range that can be accurately measured by the agarose gel method (roughly, the span between 10^3 and 10^4 nucleotides).

This problem in measuring the fork rate in the seven-protein reaction has been partly (but not entirely) alleviated by the introduction of a preincubation period in which chain growth rates are artificially slowed by a dNTP limitation. In this way, a more synchronous increase in chain growth rate can be obtained, and about 10% of the forks achieve a rate of nearly 400 nucleotides/sec within 12 seconds after restoration of the normal dNTP concentrations. The results for a seven-protein reaction employing such a preincubation are presented in Figure 9B. Comparison with Figure 9A reveals that a dramatic rate increase is obtained upon addition of X protein and gene-41 protein to the five-protein reaction.

Because an asynchronous increase in rate is still observed, experiments of the type shown in Figure 9B are not a satisfactory way to accurately measure chain growth rates in the seven-protein reaction. Even with the preincubation, there is still a lag of 30 seconds or so before the majority of forks accelerate to the rapid rate, and as a result, the DNA front observed after electrophoresis is heterogeneous and difficult to measure for calculations. Nevertheless, these experiments do show an unmistakable effect of the priming proteins on the rate at which the leading strand of the replication fork moves. Further experiments show that (1) the fork rate is increased even if rCTP, rGTP, and rUTP are omitted, and (2) the fork rate is increased even if gene-41 protein is added without adding X protein (it is not even clear that X protein is directly involved). This effect of priming proteins does not therefore seem to require any RNA-primer synthesis.

One might note that there is also a very substantial delay before de novo RNA-primed starts are seen in a lagging-strand model reaction on a fd DNA template. Both this initial lag and that seen in a seven-protein fork-rate-measurement assay may reflect a similar rate-limiting step; if so, this step might well involve the binding of gene-41 protein to a DNA single strand.

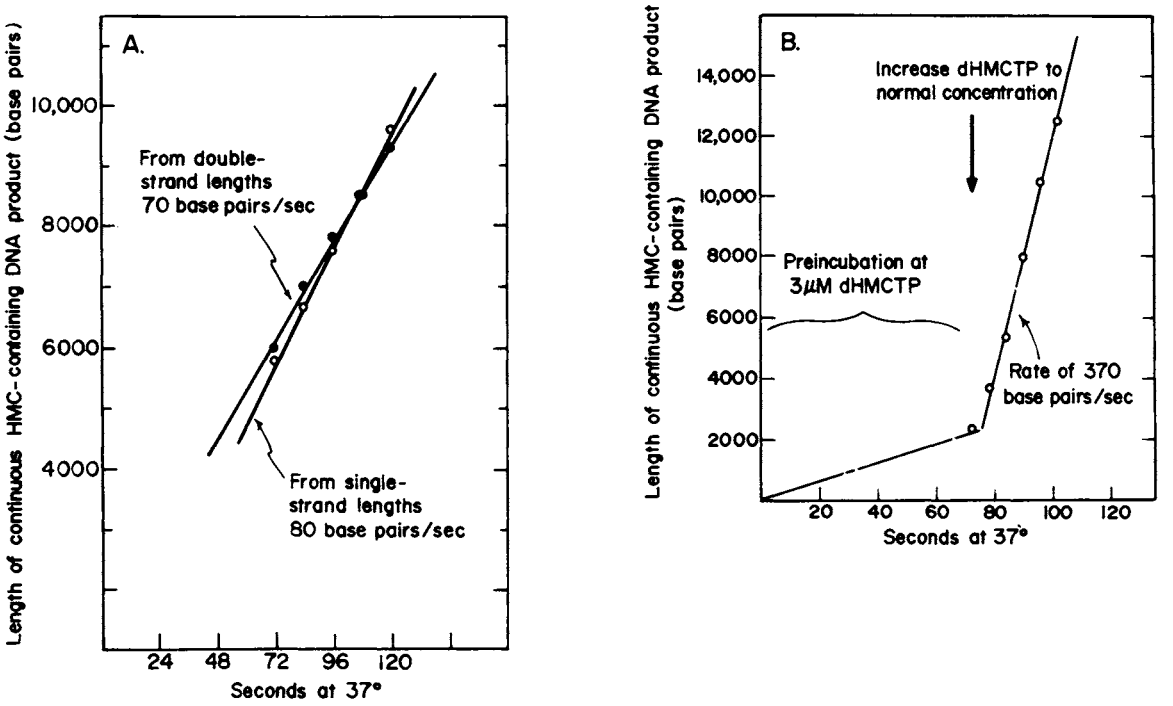

Figure 9. Measurement of the rate of DNA chain growth catalyzed by T4 replication-protein complexes on a double-helical DNA template. (*A*) DNA synthesis in a five-protein reaction (proteins 43, 32, 44/62, and 45 only). The length of the longest continuous HMC-containing DNA product made on a T7 DNA template is plotted against reaction time at 37°C. The data on double-strand lengths is taken from the autoradiograph in Fig. 8C and that on single-strand lengths from a similar autoradiograph of the same DNA samples analyzed by electrophoresis through an agarose gel in alkali, as described by McDonell et al. (1977). The size of the largest continuous HMC-containing DNA at each time point was estimated by measuring the distance from the origin to a point of estimated standard intensity at the "front" of the restriction-enzyme-resistant DNA product (see Fig. 8C). This distance was converted to molecular length by reference to a plot of the migration distances of DNA fragments of known length run in the same gel vs the log of their molecular length in base pairs. (*B*) DNA synthesis in a seven-protein reaction (proteins 43, 32, 44/62, 45, 41, and X). The length of the longest continuous HMC-containing DNA product made on a double-helical T7 DNA template is plotted against reaction time at 37°C. DNA was synthesized as in Fig. 8C, with the addition of 200 μM each of CTP, UTP, and GTP; 30 μg/ml of gene-41 protein; and 0.1 μg/ml of a protein preparation estimated to contain 20% X protein. As described in the text, to improve synchrony dHMCTP was added at only 3 μM for the first 1.2 min of incubation; thereafter, dHMCTP was raised to its normal concentration of 160 μM to start the relevant part of the reaction. The size of the largest restriction-enzyme-resistant DS DNA at each time point was estimated from a neutral agarose gel electrophoresis autoradiograph as in *A*. Because the synchrony with which replication forks accelerate in a seven-protein reaction is poor (even with the indicated preincubation), this rate estimate lacks the accuracy of those obtained for the five-protein reaction in *A* and should be viewed as an approximation only.

In conclusion, at least one of the proteins designated as an RNA-priming protein in Table 2 appears to play a second distinct role in driving the leading strand of a replication fork. It is especially interesting that the gene-41 protein is a DNA-dependent GTPase (and ATPase) and could therefore utilize some of its nucleotide hydrolysis energy to help push the fork ahead through double-helical templates. If so, it is tempting to speculate that this protein runs along on the single strand on the lagging side of the fork, since it could then perform both an RNA-priming role and a fork-driving role from the same point on the DNA. From this viewpoint, gene-41 protein might, for example, function both as proposed for the *E. coli dnaB* protein in RNA priming (McMacken et al. 1977) and as proposed for a DNA-unwinding enzyme in fork movement (Kuhn et al., this volume). An analogy between T4 gene-41 protein and T7 gene-4 protein (Hillenbrand et al.; Richardson et al.; both this volume) could also be made.

On an SS DNA template, the full in vivo rate of about 800 nucleotides/sec is attained with DNA polymerase and the polymerase-accessory proteins alone, with gene-41 protein apparently having no effect on polymerase rate (Alberts et al. 1975; Mace 1975). Thus, one can view the RNA-priming proteins as specifically removing an inhibition of DNA polymerase movement caused by templates that are in a double-helical form. It should be noted that this inhibition cannot be related to the formation of hairpin double-helices, which arise when branch migration or strand switching interrupts the polymerization process (Inman et al. 1965; Masamune and Richardson 1971; Nossal 1974), since more than 99% of the product made in both the five-protein and seven-protein reactions is fully denaturable at all extents of synthesis (Alberts et al. 1975, 1977; Morris et al. 1975).

Our finding of a second role for the priming proteins on the leading side of the fork can be used to explain the nearly instantaneous cessation of replication in a temperature-sensitive mutant in gene 41 after a shift to 42°C (Curtis and Alberts 1976). The *E. coli dnaB*

protein (like gene-41 protein, a DNA-dependent nucleoside triphosphatase with a role in RNA-primer synthesis) might also turn out to have such a dual function at a replication fork, since many temperature-sensitive mutants of this protein also stop replication very quickly after a shift to 42°C (Lark 1972; see also Wickner et al. 1974; Kornberg, this volume).

In conclusion, our results indicate that simple models which postulate a functional and spatial separation of chain-priming and chain-elongation roles in a replication fork are incorrect. Instead, the seven-protein T4 replication system should be viewed as a true multienzyme complex, in which most (if not all) of the proteins interact and, in the process, couple the synthetic events on the leading and lagging sides of the fork to each other. This coupling might be designed to ensure that synthesis of both strands is kept in step, since a temporary deficiency in de novo priming events would otherwise be expected to lead to the potentially deleterious situation in which long lengths of parental SS DNA accumulate opposite the lagging strand (see Fig. 2A).

Studies of the Fidelity of DNA Replication in the T4 In Vitro System

One of the central questions concerning DNA replication mechanisms is how the DNA template is copied with such striking fidelity. For example, as estimated from mutation rates in vivo, only one error is made per 10^9–10^{10} base-pair replications in *E. coli* (Drake 1969, 1970). Low error rates are clearly essential if the genetic information in an organism is to be effectively maintained. It has been argued that much of the enzymatic complexity of replication may be attributable to fidelity-generating processes which are still only poorly understood mechanistically (Alberts and Sternglanz 1977), although general schemes capable of removing errors by "proofreading" have been described (Hopfield 1974; Ninio 1975; see also Brutlag and Kornberg 1972). We have employed two different types of assays for measuring replication fidelity in our T4 in vitro system, with the eventual aim of defining the sources of the fidelity in this system in detail.

Analog-incorporation studies. Following the lead of Bessman et al. (1974), we have studied the ability of our replication system to discriminate between 2-aminopurine deoxyribonucleoside triphosphate (2APTP) and dATP when both triphosphates are present in the reaction mixture. The 2-aminopurine is a mutagen in vivo; it is incorporated rather effectively into DNA by pairing with T, and it causes mutation by mispairing occasionally with C (Freese 1959; Goodman et al. 1977).

Purified T4 DNA polymerase discriminates against 2APTP in an in vitro reaction, incorporating about eight times more dATP than 2APTP into DNA from an equimolar mixture (Bessman et al. 1974). We have confirmed these results and compared them to the discrimination achieved in the synthesis carried out in the five-protein reaction. This five-protein reaction utilizes a nicked, double-stranded T7 DNA template to ensure that all synthesis is carried out entirely by the multiprotein complex required, rather than by the polymerase alone. (In contrast, on an SS DNA template, not all of the polymerase molecules are simultaneously affected by the accessory proteins in model reactions [Mace 1975; Alberts et al. 1975]).

As shown in Table 3, the ratio of dATP to 2APTP incorporated into DNA is increased from about 8 in a reaction carried out by the polymerase alone (on an SS DNA template) to about 30 in the five-protein reaction. A discrimination of 30 is probably sufficient to account for the estimated discrimination seen in vivo after correction has been made for biased intracellular dATP and 2APTP pools (M. Goodman, pers. comm.).

An increase from 8 to 30 in discrimination against 2APTP is a major effect, since a single proofreading step operating at maximum efficiency can increase the discrimination prior to the proofreading step to only its square (Hopfield 1974). In the case of DNA synthesis, this limitation arises because both initial discrimination and proofreading processes can be expected to be constrained by the frequencies of incorrect base-pair formation at the end of the growing DNA helix. Moreover, the T4 DNA polymerase has a 3′→5′ exonuclease function, generally believed to act in a postincorporation proofreading process (Brutlag and Kornberg 1972, Alberts and Sternglanz 1977). Therefore, the "initial" discrimination of 8 by the polymerase alone presumably already represents the product of initial discrimination and some exonuclease proofreading. Our data, in fact, suggest that the initial discrimination against 2APTP by the polymerase alone is only about 5 (Table 3).

By comparing the relative turnover of dATP and 2APTP to the corresponding monophosphates during these two reactions, we have been able to demonstrate that the increased discrimination against 2APTP in the five-protein reaction is due to enhancement of some proofreading step (leading to turnover of triphosphate to monophosphate), rather than to an improved initial selection of the substrates to be polymerized. In fact, the data shown in Table 3 suggest that the proofreading step contributes more to the observed discrimination of 30:1 than the initial base-selection process. Not shown in Table 3 is the fact that the additional T4 proteins increase the ratio of turnover to incorporation even for correct substrates, although to a much lesser degree than for 2APTP (U. Hibner, unpubl.).

Mutant-reversion studies. To study the fidelity of DNA synthesis with regard to discrimination against incorrect base pairs between the four natural bases, an assay much more sensitive than one based on the incorporation of radioactive nucleotides into DNA is needed. As independently recognized by Loeb (Weymouth and Loeb 1978; Loeb et al., this volume), the measurement of the reversion rate of nonsense

Table 3. Effect of Gene-44/62, -45, and -32 Proteins on DNA Polymerase Discrimination against 2APTP

Proteins	Discrimination $\left(\frac{\text{dATP incorporated}}{\text{2APTP incorporated}}\right)$	Contribution to discrimination from nucleotide turnover (observed)[a]	Discrimination during initial polymerization (calculated)
43[b]	8.1	1.6	5.0
43, 44/62, 45, 32 (five-protein reaction)[b]	31.7	8.7	3.6
All[c]	30–40	?	?

The effect of proteins 44/62, 45, and 32 on the discrimination against 2APTP in in vitro DNA synthesis. The DNA was synthesized under the standard conditions described in Fig. 1A. Five dNTPs were present at the following concentrations: 100 μM each dCTP, dGTP, and dTTP; 25 μM [α-^{32}P]dATP (100 cpm/pmole); 50μM [^{3}H]2APTP (66 cpm/pmole). The [^{3}H]2APTP was kindly provided by Dr. M. Goodman. The reaction with polymerase alone was run using gene-43 protein at 5 μg/ml and single-stranded, sonicated, calf thymus DNA (100 μg/ml) as template. In a 5-min reaction at 37°C, 0.49 nmole 2APTP and 2.0 nmoles dATP were incorporated into DNA per ml of reaction. The five-protein reaction utilized T7 DS DNA (13 μg/ml) as template. In addition to gene-43 protein, 20 μg/ml gene-44/62 protein, 28 μg/ml gene-45 protein, and 125 μg/ml gene-32 protein were present. In a 5-min reaction at 37°C, 14 pmoles 2APTP and 240 pmoles dATP were incorporated into DNA per ml of reaction. To assay for DNA synthesis, aliquots were withdrawn from the reaction mixture and DNA synthesis was measured as TCA-precipitable ^{3}H and ^{32}P counts at 0, 5, 10 and 20 min of reaction. Parallel aliquots were assayed for nucleotide turnover from triphosphate to monophosphate by chromatography on PEI cellulose in 1 M LiCl, essentially as described by Hershfield and Nossal (1972). For both reactions, the data presented were calculated from the initial rates observed for DNA synthesis and nucleotide turnover, respectively, averaging two independent experiments.

[a]The relative turnover of 2APTP divided by the relative turnover of dATP, which is the ratio of

$$\frac{\text{pmoles 2APTP turned over}}{\text{pmoles 2APTP incorporated}} \text{ to } \frac{\text{pmoles dATP turned over}}{\text{pmoles dATP incorporated}}$$

[b]In vitro.

[c]In vivo. Preliminary data from M. Goodman, University of Southern California (pers. comm.).

mutants when small bacteriophage DNAs are replicated in vitro provides such an assay.

The outline of our experimental procedure is presented in Figure 10. DS DNA from bacteriophage ϕX174, containing a sequenced amber mutation, is used as a template for in vitro DNA synthesis. The T4 replication proteins replicate such a template as a rolling circle, producing 10–20 new copies of the genome within 30 minutes. The extensive synthesis obtained in our system yields mainly a DS DNA product with both strands as newly synthesized DNA. This fact eliminates the possibility of artifacts due to mechanisms which discriminate new and old DNA within cells when repairing mismatched bases (M. Meselson et al., pers. comm.). After synthesis, the DNA is cut to unit-length fragments by *Pst*I restriction endonuclease treatment, which makes a single cut per ϕX genome. The DNA is treated with *E. coli* DNA ligase to form circles and is then used to transfect *E. coli* spheroplasts. Since DNA containing an amber mutation is used as the template, upon transfection of nonpermissive (Su^-) spheroplasts, only revertants will give plaques. The precise fraction of revertants is determined by comparison with the infectivity of the replicated mutant ϕX DNA on Su^+ spheroplasts, where the original mutant DNA molecules are fully infectious.

In the experiments reported, DNA from the *am3* mutant of ϕX (defective in the product of ϕX gene *E*, which is required for lysis of the host) is used as template. Its DNA sequence compared to wild type is shown in Figure 10 (Sanger et al. 1977) (Although the *am3* mutation lies in a region of overlap of genes *E* and *D*, only the product of gene *E* is affected by the indicated change in the mutant.) Revertants could in theory be produced by eight different single-base changes affecting the *am3* codon, although it is unknown how many of these changes would in fact produce viable phage. As described below, we have chosen to run our in vitro reactions under conditions strongly favoring reversion to the true wild-type genotype (i.e., A-T→G-C at position 587).

Since we rely on nonspecific nicking of the double-stranded template DNA for rolling-circle formation, there is an equal probability that either strand (viral or complementary) may serve as the leading or lagging strand in the synthesis. All of the revertants that appear on the leading strand during the synthesis will result in the production of homoduplex revertant DNA, and these homoduplexes should account for 50% of the revertants. Reversion during lagging-strand synthesis will result in heteroduplex DNA; 25% of the revertants can be expected to be heteroduplexes with wild-type genotype on the viral strand and 25%, on the complementary strand. Thus, neither mismatch repair nor the reported preference for expression of the complementary strand during heteroduplex infection (Denhardt 1975) should significantly affect our results.

Due to the relatively high in vivo reversion frequency of ϕX mutants, the population of *am3* DNA used as template in our experiments is contaminated with what appears to be wild-type revertants at a level of about 5×10^{-6}. Clearly, this background limits our ability to detect new revertants produced during in vitro DNA synthesis with the transfection assay. However, by strongly biasing substrate concentrations dur-

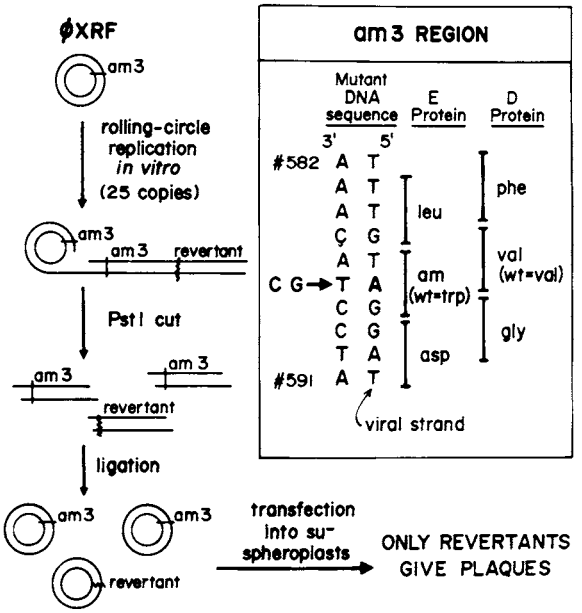

Figure 10. Schematic summary of an infectivity assay for replication fidelity which uses mutant ϕX DNA as template. The DNA is replicated in vitro under the standard conditions described in Fig. 1A, with all seven T4 proteins present. The reaction is stopped after 30 min at 37°C by heating at 70°C for 10 min. After cooling on ice, 0.1 volume of a solution containing 0.2 M Tris-HCl (pH 7.5), 100 mM $MgCl_2$, and 1 mg/ml gelatin is added. The DNA is then digested for 1 hr at 37°C with an excess of *Pst*I restriction endonuclease (10 units/μg DNA). After this reaction is stopped by heating at 70°C for 10 min, 0.1 mM NAD and an excess of *E. coli* DNA ligase are added; the DNA is then incubated overnight at 14°C in order to re-form DNA circles between the overlapping, *Pst*I-generated DNA ends. The efficiencies of restriction cutting and ligation are monitored by electrophoretic analysis on agarose slab gels. The spheroplasts with which this DNA is mixed for transfection assay are prepared from *E. coli* K12 strains according to the method of Benzinger (1978). Strain 594 (Su^-) was used for preparation of nonpermissive spheroplasts, strains CR63 (*su1*) and W3350 (*su3, gal*$^-$, *leu*$^-$, *proA*$^-$) were used interchangeably as a source of ϕX *am3* permissive spheroplasts; the bacterial strains were kindly provided by Dr. R. Benzinger. Infective centers are scored following a 10–12 min period of infection. *E. coli* C strains, either wild type, HF4714 (*su1*) or CQ2 (*su3*) were used as indicator bacteria. The efficiencies of transfection per DNA molecule added varied with different spheroplast preparations between 10^{-7} and 5×10^{-6}.

ing DNA synthesis in favor of wild-type revertants (for either the viral strand [high dGTP/dATP] or the complementary strand [high dCTP/dTTP] or both), we can greatly increase the sensitivity of the assay. For example, the use of a 100-fold excess of dCTP over dTTP to favor wild-type revertants (Fig. 10) increases the sensitivity by a factor of 50 (the factor is not 100, since the reversion frequency on only one of the two strands will be affected).

Before testing the ϕX revertant frequencies following in vitro replication, two control experiments were performed. First, the newly synthesized DNA was tested and found to be essentially as infectious as the original template, for both wild-type and *am3* DNA. (Testing was done by infecting Su^- and Su^+ spheroplasts, respectively.) The second important control demonstrated that the high level of *am3* ϕX molecules present in our replicated DNA (10^{11}–10^{12} DNA molecules/ml) does not interfere with the infectivity of low levels of wild-type DNA present. Thus, the wild-type revertants present in the replicated sample should be efficiently detected.

The results from two independent experiments which measured the reversion frequencies of *am3* mutants of ϕX during in vitro DNA synthesis are presented in Table 4. As indicated, no revertants to wild type were seen above the background level of the in vivo reversion frequency. Taking into account the strong substrate bias in favor of wild-type reversions during the in vitro synthesis, we calculate that the error frequency for an A-T→G-C transition at position 587 is less than 5 nucleotides per 10^7 nucleotides copied in the T4 in vitro system.

This estimate for the fidelity of DNA synthesis by the purified T4 proteins in vitro is impressive, being more than a factor of 10^2 higher than previous best estimates for the fidelity of in vitro synthesis on natural DNA templates (Weymouth and Loeb 1978; Loeb et al., this volume). These other estimates are based on reactions carried out by DNA polymerases alone. Since DNA replication in vivo is carried out by a multiprotein complex, of which the DNA polymerase is only a part, it is not surprising that the estimate of the in vitro fidelity of DNA replication in Table 4 much more closely approaches the fidelity estimated for the complete in vivo process. However, it is difficult to compare the fidelity we observe in vitro with the fidelity found for the T4 replication system in vivo in any exact way. Although the in vivo error frequencies measured for single-base substitutions in T4 average about 10^{-8} for forward mutations in the *r*II locus, specific reverse-mutation rates range from about 10^{-4} to 5×10^{-11} per base-pair replication and strongly depend on the exact neighboring DNA sequence (Drake 1970). This variation is observed even for the same base-pair change (see also Coulondre and Miller 1977). A better comparison with the in vivo situation will be possible when reversion rates for several different sequenced ϕX nonsense mutations have been measured in vitro and compared to in vivo reversion rates for regions of similar sequence on the T4 DNA genome.

CONCLUSIONS

One of the major conclusions of this work is that DNA replication in vitro, as carried out by the T4 bacteriophage multienzyme complex of seven purified proteins, closely mimics the in vivo process. For example, equivalent rates of fork movement and the expected fork configurations are observed (for electron micrographs, see Alberts et al. 1975; Morris et al. 1975), and the impressive faithfulness of the in vivo copying process is maintained. Also, our data on the initiation of DNA pieces with short RNA oligonucleotide primers are in good agreement with recent

Table 4. Reversion Frequency of the ϕX *am3* Mutation during In Vitro DNA Synthesis in a Seven-protein Reaction

	dCTP/dTTP		dGTP/dATP	
dNTP bias during in vitro reaction	10/1 (exp. 1)	100/1 (exp. 2)	10/1 (exp. 1)	50/1 (exp. 2)
No. of infective centers expected if all DNA molecules are wild type	4.1×10^5	4.6×10^4	2.3×10^5	2.0×10^5
No. of infective centers observed	1	0	2	1
Fraction of wild-type molecules in *am3* DNA population[a]	2.4×10^{-6}	$<2.2 \times 10^{-5}$	8.7×10^{-6}	5.0×10^{-6}
Calculated error frequency[b]	$<4.8 \times 10^{-7}$	$<4.3 \times 10^{-7}$	$<8.7 \times 10^{-7}$	$<2.0 \times 10^{-7}$

The experimental procedure used is that described in the legend to Fig. 10. The results from two independent experiments are presented. In experiment 1, the concentration of the nonbiased dNTPs was 100 μM; to obtain the required bias, the indicated triphosphates at 500μM and 50 μM were used. In a typical transfection, duplicate samples of 0.4 ml each containing 2 μg/ml of DNA were incubated with 0.4 ml of spheroplast suspension for 11 min at 37°C. Indicator bacteria were then immediately added, and the spheroplasts were plated on agar. In experiment 2, the unbiased dNTP concentrations were 50 μM, and the two sets of biased concentrations were 2 mM dCTP and 20 μm dTTP, and 1 mM dGTP and 20 μM dATP, respectively.

[a]Not significantly above revertant frequency in the template DNA.

[b]Corrected for dNTP bias favoring wild-type revertants.

data on in-vivo-initiated T4 DNA chains, with the ribonucleotides A and C playing a predominant role in both cases (Okazaki et al., this volume). Finally, genetic evidence demonstrates that all of the T4-encoded proteins used in the in vitro system are also intimately involved in DNA replication in vivo.

There is one important difference between in vitro and in vivo replication. Whereas the process of replication-fork movement occurs in a biologically meaningful manner, in the T4 in vitro system replication forks are initiated at preexisting random nicks on double-helical templates. In contrast, inside the cell it seems likely that any nicks in the double helix are kept sealed by DNA ligase. Moreover, the available data indicate that replication forks begin at a limited number of different specific sites on the T4 genome, seen as a replication bubble from which replication forks propagate bidirectionally (Delius et al. 1971; Mosig et al., this volume). Thus, a special chromosomal initiation process probably exists, which involves additional T4 proteins besides the seven required for the replication-fork propagation reactions discussed here. At a minimum, these proteins need only result in an opening-up of the DNA helix at the specific origin sites, exposing single strands at which RNA-primed de novo chain starts are made.

Completely defined in vitro systems capable of DNA replication have been worked out in order to facilitate mechanistic studies which are not possible in vivo. Thus far, the most extensively developed in vitro replication systems are those from bacteriophage T4 (see also Silver and Nossal, this volume), bacteriophage T7 (Richardson et al., this volume; Hillenbrand et al., this volume), and *E. coli* (Kornberg; Sumida-Yasumoto et al.; Meyer et al.; Eisenberg et al.; Wickner; all this volume). What have we learned to date concerning the central mechanisms involved?

One important question, which has turned out to have an unexpectedly complex answer, concerns the source of the energy needed to unwind the DNA helix ahead of the replication fork (~12–50 kcal/mole/10 bp melted, depending on the DNA sequence) (for review, see Alberts and Sternglanz 1977). In principle, parts of this energy can be derived from (1) the binding of helix-destabilizing protein to the DNA single strands liberated as the parental helix unwinds, (2) the dNTP hydrolysis which accompanies polymerization driving the DNA polymerase ahead, (3) enzymes like DNA gyrase creating an unwinding strain on the helix (Gellert et al. 1976), (4) an ATP-hydrolysis-driven DNA-unwinding enzyme which "walks" along the parental DNA on the lagging side of a replication fork, or (5) a similar DNA-unwinding enzyme which "walks" along the parental DNA on the leading side of a replication fork.

In the T4 system, with its two different DNA-dependent ATPases (the gene-44/62 and -41 proteins), all of the above contributions to helix unwinding except number 3 could be involved (the antibiotics which inhibit *E. coli* DNA replication by acting on DNA gyrase have no affect on T4 replication; nor do our replication proteins contain a new DNA gyrase type of activity) (E. Belikoff and R. Sternglanz, unpubl.). Available data on the use of double-helical templates by the T7 and *E. coli* replication proteins likewise suggest that no one protein is sufficient by itself to accomplish the unwinding task and that different replication systems will turn out to use a different balance of the five possible helix-unwinding mechanisms listed above. Some measurement of the rates with which replication forks move through double-helical templates (analogous to Figs. 8 and 9) would be helpful in evaluating the different contributions to helix unwinding in these other systems.

The details of RNA-priming mechanisms have been most extensively worked out in the *E. coli* replication system (reviewed in Kornberg, this volume); with *E. coli* proteins, RNA priming is very complicated on a ϕX DNA template (involving at least seven proteins) and relatively simple on a G4 DNA template (where only the *dnaG* protein and the *E. coli* HD protein are required). RNA priming appears to involve only one protein in the T7 replication system (the gene-4 protein) (Richardson et al.; Hillenbrand et al.; both this volume). The T4 system, where priming involves the gene-41 protein and the X protein, is intermediate in complexity. Why these two T4 proteins are required (where the use of only one would seem possible based on results in other systems) is not at all clear. Perhaps the use of two proteins increases either the templating fidelity or the reliability of the process of replication (or both). Also, the apparent requirement for ATP or GTP hydrolysis by the gene-41 protein to make RNA primers remains to be explained. These poorly understood features of the priming mechanism may be related to one or more of the following obvious mechanistic problems inherent in priming chain starts on the lagging strand:

1. The primer-making enzyme, which of necessity covers the 3′ end of the de novo chain start, must get out of the way at some point to allow this 3′ end to be picked up by the DNA polymerase. This appears to be accomplished by having the primer-making enzyme self-terminate, with the result that almost all of the oligonucleotides made by the gene-41 and X proteins in vitro are about 7 nucleotides long. But how is this distance measured enzymatically? Also, one might suspect that such short RNA-DNA duplexes need to be specifically stabilized if they are to serve as efficient primers.
2. With the replication fork moving at about 800 nucleotides/sec, new RNA-primed de novo chain starts are needed every few seconds on the lagging strand, since the size of the Okazaki fragments in T4 is about 2000 nucleotides (Okazaki et al. 1969). How and why are specific sites recognized for this purpose, and how is the necessary efficiency of the process guaranteed?
3. Adjacent Okazaki pieces must be prevented from overrunning each other on the lagging strand as the growing 3′ end of one piece meets the 5′ end of the preceding piece (Fig. 2B). If the temporary nick or gap at their junction instead serves as a primer for initiating a strand-displacement synthesis, as occurs in some cases in our in vitro system, the displaced Okazaki piece would serve as a template for further DNA synthesis. As a result, more than one daughter DNA copy would be made on this side of the fork, and the elegant process of DNA replication would become hopelessly complicated.

The final mechanistic question, which the availability of efficient in vitro DNA replication systems now allows one to approach experimentally, is where the enormous fidelity of the DNA-template-copying process arises. By showing that the T4 system is highly faithful, the data in Tables 3 and 4 reveal that the answers can be obtained from in vitro analysis. To date, we have learned that the polymerase-accessory proteins increase fidelity through an effect on deoxyribonucleotide turnover (Table 3). One explanation for this result is that the polymerase-accessory proteins greatly increase the efficiency of the 3′→5′ exonuclease proofreading by the DNA polymerase, especially since this proofreading works rather poorly to discriminate 2APTP in the reaction with polymerase alone (only a factor of 1.6 in the experiment in Table 3). However, other explanations are possible, including an effect on a separate proofreading step. For example, proofreading could occur in the breakdown of a hypothetical high-energy nucleotide-polymerase complex that may form prior to polymerization per se.

Other important questions concern the exact arrangement of the replication proteins with respect to each other and to the DNA. For example, how do the polymerase-accessory proteins fit together with the polymerase at the growing 3′ end, and where in this region does the gene-44/62 protein find a DNA-binding site? Ultimately, the detailed answers to all questions of this type will require not only further biochemical studies, but also knowledge of the detailed three-dimensional structure and chemistry of the T4 replication proteins (including the precise nature of their regions of protein-protein and protein-DNA interaction). Large amounts of the T4 proteins can already be obtained relatively easily, making their crystallization and X-ray diffraction analysis a reasonable goal for the not too distant future.

Acknowledgments

This work was supported by grant GM-24020 from the National Institutes of Health to B. M. A., by a Damon Runyon Fellowship to U. H., and by National Institutes of Health training grant support to R. L. B. We wish to thank particularly Dr. Myron Goodman for his generous gift of [^{3}H]2APTP and Dr. Rolf Benzinger for his advice and materials for the ϕX transfection assay.

REFERENCES

ALBERTS, B. M. and L. FREY. 1970. T4 bacteriophage gene 32: A structural protein in the replication and recombination of DNA. *Nature* **227:** 1313.

ALBERTS, B. M. and R. STERNGLANZ. 1977. Recent excitement in the DNA replication problem. *Nature* **269:** 655.

ALBERTS, B. M., C. F. MORRIS, D. MACE, N. SINHA, M. BITTNER, and L. MORAN. 1975. Reconstruction of the T4 bacteriophage DNA replication apparatus from purified components. In *DNA synthesis and its regulation* (ed. M. M. Goulian et al.), p. 241. W. A. Benjamin, Menlo Park, California.

ALBERTS, B. M., J. BARRY, M. BITTNER, M. DAVIES, H. HAMA-INABA, C. C. LIU, D. MACE, L. MORAN, C. F. MORRIS, J. PIPERNO, and N. K. SINHA. 1977. In vitro

DNA replication catalyzed by six purified T4 bacteriophage proteins. In *Nucleic acid protein recognition* (ed. H. J. Vogel), p. 31. Academic Press, New York.

BARRY, J. and B. ALBERTS. 1972. In vitro complementation as an assay for new proteins required for bacteriophage T4 DNA replication: Purification of the complex specified by T4 genes 44 and 62. *Proc. Natl. Acad. Sci.* **69:** 2717.

BENZINGER, R. 1978. Transfection of Enterobacteriaceae and its applications. *Microbiol. Rev.* **42:** 194.

BESSMAN, M. J., N. MUZYCZKA, M. F. GOODMAN, and R. L. SCHNAAR. 1974. Studies on the biochemical basis of spontaneous mutation. *J. Mol. Biol.* **88:** 409.

BITTNER, M. L. 1977. "Studies on bacteriophage T4 DNA replication proteins." Ph.D. thesis, Princeton University, Princeton, New Jersey.

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

CHIU, C. S. and G. R. GREENBERG. 1969. Evidence for a possible direct role of dCMP hydroxymethylase in T4 phage DNA synthesis. *Cold Spring Harbor Symp. Quant. Biol.* **33:** 351.

CHIU, C. S., P. K. TOMICH, and G. R. GREENBERG. 1976. Simultaneous initiation of synthesis of bacteriophage T4 DNA and of deoxyribonucleotides. *Proc. Natl. Acad. Sci.* **73:** 757.

COULONDRE, C. and J. H. MILLER. 1977. Genetic studies of the *lac* repressor IV. Mutagenic specificity in the *lacI* gene of *Escherichia coli. J. Mol. Biol.* **117:** 577.

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

DELIUS, H., C. HOWE, and A. W. KOZINSKI. 1971. Structure of the replicating DNA from bacteriophage T4. *Proc. Natl. Acad. Sci.* **68:** 3049.

DENHARDT, D. T. 1975. The single-stranded DNA phages. *CRC Crit. Rev. Microbiol.* **4:** 161.

DONIS-KELLER, H., A. M. MAXAM, and W. GILBERT. 1977. Mapping adenines, guanines and pyrimidines in RNA. *Nucleic Acids Res.* **4:** 2577.

DRAKE, J. W. 1969. Comparative rates of spontaneous mutation. *Nature* **221:** 1132.

———. 1970. *The molecular basis of mutation.* Holden-Day, San Francisco.

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. LEILAUSIS. 1964. Physiological studies of conditional lethal mutations of bacteriophage T4 D. *Cold Spring Harbor Symp. Quant. Biol.* **28:** 375.

FREESE, E. 1959. The specific mutagenic effect of base analogues on phage T4. *J. Mol. Biol.* **1:** 87.

GELLERT, M., K. MIZUUCHI, M. H. O'DEA, and H. A. NASH. 1976. DNA gyrase: An enzyme that introduces superhelical turns into DNA. *Proc. Natl. Acad. Sci.* **73:** 3872.

GOLD, L., P. Z. O'FARRELL, and M. RUSSEL. 1976. Regulation of gene 32 expression during bacteriophage T4 infection of *E. coli. J. Biol. Chem.* **251:** 7251.

GOODMAN, M. F., R. HOPKINS, and W. C. GORE. 1977. 2-Aminopurine-induced mutagenesis in T4 bacteriophage: A model relating mutation frequency to 2-aminopurine incorporation in DNA. *Proc. Natl. Acad. Sci.* **74:** 4806.

HERSHFIELD, M. S. and N. G. NOSSAL. 1972. Hydrolysis of template and newly synthesized deoxyribonucleic acid by 3′ to 5′ exonuclease activity of the T4 DNA polymerase. *J. Biol. Chem.* **247:** 3393.

HOPFIELD, J. J. 1974. Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. Natl. Acad. Sci.* **71:** 4135.

HUBERMAN, J. A., 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.

INMAN, R. B., C. L. SCHILDKRAUT, and A. KORNBERG. 1965. Enzymic synthesis of deoxyribonucleic acid. XX. Electron microscopy of products primed by native templates. *J. Mol. Biol.* **11:** 285.

JENSEN, D. E., R. C. KELLY, and P. H. VON HIPPEL. 1976. DNA melting proteins. *J. Biol. Chem.* **251:** 7215.

KOLODNER, R. and C. C. RICHARDSON. 1977. Replication of duplex DNA by bacteriophage T7 DNA polymerase and gene 4 protein is accompanied by hydrolysis of nucleoside 5′-triphosphates. *Proc. Natl. Acad. Sci.* **74:** 1525.

KORNBERG, A. 1974. *DNA synthesis.* W. H. FREEMAN, San Francisco.

LARK, K. G. 1972. Genetic control over the initiation of the synthesis of the short deoxynucleotide chains in *E. coli. Nat. New Biol.* **240:** 237.

MACE, D. 1975. "In vitro interactions of some T4 replication proteins." Ph.D. thesis, Princeton University, Princeton, New Jersey.

MASAMUNE, Y. and C. C. RICHARDSON. 1971. Strand displacement during DNA synthesis at single strand breaks. *J. Biol. Chem.* **246:** 2692.

MATHEWS, C. K. 1972. Biochemistry of deoxyribonucleic acid-defective amber mutants of bacteriophage T4. *J. Biol. Chem.* **247:** 7430.

MAXAM, A. M. and W. GILBERT. 1977. A new method for sequencing DNA. *Proc. Natl. Acad. Sci.* **74:** 560.

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.

MCDONELL, M. W., M. N. SIMON, and F. W. STUDIER. 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. *J. Mol. Biol.* **110:** 119.

MCMACKEN, R., K. UEDA, and A. KORNBERG. 1977. Migration of *E. coli* dnaB protein on the template DNA strand as a mechanism in initiating DNA replication. *Proc. Natl. Acad. Sci.* **74:** 4190.

MORRIS, C. F., N. K. SINHA, and B. M. ALBERTS. 1975. Reconstruction of bacteriophage T4 DNA replication apparatus from purified components: Rolling circle replication following *de novo* chain initiation on single-stranded circular DNA template. *Proc. Natl. Acad. Sci.* **72:** 4800.

MUFTI, S. and H. BERNSTEIN. 1974. The DNA-delay mutants of bacteriophage T4. *J. Virol.* **14:** 860.

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

NINIO, J. 1975. Kinetic amplification of enzyme discrimination. *Biochimie* **57:** 587.

NOSSAL, N. 1974. DNA synthesis on a double-stranded DNA template by the T4 bacteriophage DNA polymerase and the T4 gene 32 DNA unwinding protein. *J. Biol. Chem.* **249:** 5668.

OKAZAKI, R., T. OKAZAKI, K. SAKABE, O. K. SUGIMOTO, R. KAINUMA, A. SUGINO, and N. IWATSUKI. 1969. In vivo mechanism of DNA chain growth. *Cold Spring Harbor Symp. Quant. Biol.* **33:** 129.

PERLMAN, D. and J. A. HUBERMAN. 1977. Preparation of large quantities of separated strands from SV40 DNA restriction fragments by low-temperature low-salt agarose gel electrophoresis. *Anal. Biochem.* **83:** 666.

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.

PIPERNO, J. R., R. G. KALLEN, and B. M. ALBERTS. 1978. Analysis of a T4 DNA replication protein complex. *J. Biol. Chem.* **253:** 5180.

REDDY, G. P. V., A SINGH, M. E. STAFFORD, and C. K. MATHEWS. 1977. Enzyme associations in T4 phage DNA precursor synthesis. *Proc. Natl. Acad. Sci.* **74:** 3152.

RIMERMAN, R. A. and G. W. HATFIELD. 1973. Phosphate-induced protein chromatography. *Science* **182:** 1268.

RIVA, S., A. CASCINO, and E. P. GEIDUSCHEK. 1970. Cou-

pling of late transcription to viral replication in bacteriophage T4 development. *J. Mol. Biol.* **54:** 85.

SANGER, F., G. M. AIR, B. G. BARRELL, N. L. BROWN, A. R. COULSON, J. C. FIDDES, C. A. HUTCHISON III, P. M. SLOCOMBE, and M. SMITH. 1977. Nucleotide sequence of bacteriophage ϕX174 DNA. *Nature* **265:** 687.

SCOTT, J. F., S. EISENBERG, L. L. BERTSCH, and A. KORNBERG. 1977. A mechanism of duplex DNA replication revealed by enzymatic studies of phage ϕX174: Catalytic strand separation in advance of replication. *Proc. Natl. Acad. Sci.* **74:** 193.

SIGAL, N., H. DELIUS, T. KORNBERG, M. L. GEFTER, and B. M. ALBERTS. 1972. A DNA-unwinding protein isolated from *E. coli*: Its interaction with DNA and with DNA polymerases. *Proc. Natl. Acad. Sci.* **69:** 3537.

STAHL, F. W. 1967. Circular genetic map. *J. Cell Physiol.* (suppl. 1) **70:** 1.

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.

WEYMOUTH, L. A. and L. A. LOEB. 1978. Mutagenesis during in vitro DNA synthesis. *Proc. Natl. Acad. Sci.* **75:** 1924.

WICKNER, S., M. WRIGHT, and J. HURWITZ. 1974. Association of DNA-dependent and -independent ribonucleoside triphosphatase activities with *dnaB* gene product of *Escherichia coli*. *Proc. Natl. Acad. Sci.* **71:** 783.

WOOD, W. B. and H. R. REVEL. 1976. The genome of bacteriophage T4. *Bacteriol. Rev.* **40:** 847.

YEGIAN, C. D., M. MUELLER, G. SELZER, V. RUSSO, and F. W. STAHL. 1971. Properties of DNA-delay mutants of bacteriophage T4. *Virology* **46:** 900.

ZIMMERMAN, S. B. and A. KORNBERG. 1961. Deoxycytidine di- and triphosphate cleavage by an enzyme formed in bacteriophage-infected *E. coli*. *J. Biol. Chem.* **236:** 1480.

Cold Spring Harbor Symposia on Quantitative Biology

Probing DNA Replication Mechanisms with the T4 Bacteriophage In Vitro System

C. C. Liu, R. L. Burke, U. Hibner, et al.

Cold Spring Harb Symp Quant Biol 1979 43: 469-487
Access the most recent version at doi:10.1101/SQB.1979.043.01.053