From "DNA Synthesis and Its Regulation" Vol. 3 of ICN-UCLA Symposium on Molecular & Cellular Biology (M. Goulian, P. Hanawalt and C.F. Fox, eds) W.H. Benjamin, Inc., Menlo Park (1975)

RECONSTRUCTION OF THE T4 BACTERIOPHAGE DNA REPLICATION APPARATUS FROM PURIFIED COMPONENTS

Bruce Alberts, C. Fred Morris, David Mace, Navin Sinha, Michael Bittner and Larry Moran

Department of Biochemical Sciences
Princeton University
Princeton, New Jersey 08540

ABSTRACT: The protein products of T4 genes 41, 43 (T4 DNA polymerase), 45, 44 and 62 have been purified to near homogeneity, using individual "complementation assays" which measure their ability to stimulate DNA synthesis in an appropriate mutant-infected crude lysate. The gene 44 and 62 proteins are copurified as a tight complex which displays DNA-dependent ATPase activity. In a reaction requiring 45 protein, this complex speeds up *in vitro* polymerization by T4 DNA polymerase on primed single-stranded templates to near *in vivo* rates. Further addition of gene 32 and gene 41 proteins allows two additional reactions:

(1) Extensive synthesis by strand-displacement on double-stranded DNA templates (as expected for synthesis on the "leading side" of a replication fork),

(2) *De novo* DNA chain initiation on single-stranded DNA templates, dependent on the presence or ribo C, G and U nucleoside triphosphates (as expected for RNA-priming of Okazaki pieces on the "lagging side" of the replication fork).

With the complete system operating on circular fd DNA templates, the products visualized in the electron microscope resemble classical rolling circles, with discontinuous DNA synthesis clearly evident on the long, unbranched double-stranded tails.

INTRODUCTION: Since 1971, efforts in this laboratory have been directed towards achieving *in vitro* reconstruction of the T4 bacteriophage DNA replication apparatus from purified components. With this communication, we report our first real success.

The work to be described was inspired by the elegant genetic studies begun by R.H. Epstein, R. Edgar and colleagues, which indicated that at least six major T4 gene

products are necessary for in vivo replication fork movement (1-3). Fig. 1A presents a highly schematized version of the current T4 genetic map (4), emphasizing the six "replication proteins," henceforth designated as the gene 32, gene 41, gene 43, gene 44, gene 45, and gene 62 proteins, respectively. Of these gene products, only the proteins corresponding to gene 43 (T4 DNA polymerase (5-10)) and gene 32 (DNA-unwinding protein (11-13)) had been purified when our work began.

Isolation of the Replication Proteins In our current view, each of the above replication proteins constitutes part of a large multienzyme complex or "machine" for replicating DNA; thus, individual catalytic functions need not exist for them outside of this complex, and some may even have a purely structural role. Nevertheless, by quantitating the small amounts of DNA synthesis observed in a concentrated crude extract, and exploiting the T4 mutants isolated by the geneticists, these proteins can be detected via an "in vitro complementation assay" (14,15). With this assay, we have been able to purify the four previously unpurified replication proteins to homogeneity in active form. Fig. 1B shows a SDS-polyacrylamide gel electrophoretic analysis of our present preparations of the above six T4 proteins. Using purification procedures to be described elsewhere [and starting with cells infected with overproducing T4 mutants isolated by J. Wiberg and J. Karam (16,17)], we can obtain at least 10 mg quantities of each protein from 100 g of infected cells. The major protein bands in Fig. 1B have been identified as the indicated gene products by standard protein double-label experiments, using mixed radioactive amber-mutant and wild-type-infected extracts as the starting material for a complete isolation. On further fractionation by sucrose gradient sedimentation, each protein band comigrates with the appropriate activity in the in vitro complementation assay, and with the activities to be described below.

As previously reported (14), the gene 44 and 62 proteins are isolated as a tight complex (180,000 daltons) which appears to contain four molecules of 44-protein (34,000 daltons each) and two molecules of 62-protein (20,000 daltons each). The other four proteins contain only a single type of subunit, although the gene 32-protein appears to exist mainly as a 6-8 S aggregate and the gene

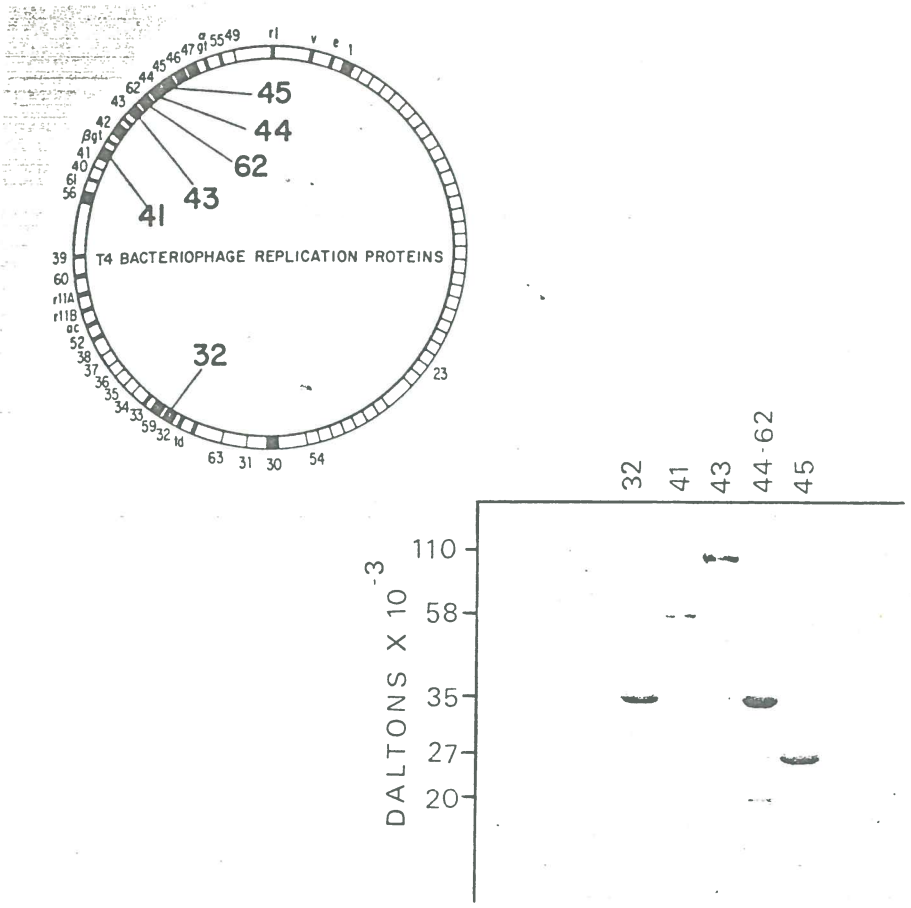


Figure 1. (A). The genetic map of T4 bacteriophage with replication genes emphasized (1,4). The broadest black segments indicate the relative locations of mutationally identified genes with a major effect on DNA replication at 37°C. Mutants in genes 32, 41, 43, 44, 45 or 62 are unique in synthesizing little or no DNA, even though all four deoxyribonucleotide triphosphate precursors are present (42).

(B). Polyacrylamide gel electrophoresis of the purified T4 replication proteins. The separating gel was made with 12.5% acrylamide and the stacking gel with 6% acrylamide, in buffers containing 0.1% sodium dodecyl sulfate (43). The protein samples were diluted from their concentrated stocks directly into SDS-mercaptoethanol sample buffer and boiled for 2 minutes at 100°C just prior to electrophoresis. To estimate purity, the relative intensities of the Coomassie-blue stained protein bands shown were determined with a Joyce-Loebl densitometer.

45-protein as a dimer. The gene 41-protein preparation is the least pure (being about 85% a single species), whereas all of the others appear to be on the order of 99% homogeneous. Tested on supercoiled PM2 native DNA or fd DNA single-strands, the addition of all these proteins at the concentrations at which they are routinely used reveals no detectible endonuclease activity. However, it should be noted that even our best preparations of T4 DNA polymerase exhibit low levels of DNA ligase activity (as do the preparations of others (10)).

Partial Reactions The most striking in vitro activities of the T4 replication proteins require that all six of them be present simultaneously. But since analysis of individual functions in such a multicomponent system is difficult, it is instructive to study in detail as many "partial reactions" as possible. Here we include the DNA polymerase (7) and proof-reading exonuclease (18) activities of 43-protein (which require single-stranded DNA templates), the denaturation (and renaturation) of DNA catalyzed by 32-protein (11), and the specific interaction of this polymerase and DNA-unwinding protein with each other (13). Additionally, the 44-62 protein complex is an ATPase which can hydrolyze either ribo- or deoxyribo-ATP to the corresponding diphosphate and inorganic phosphate. As shown in Fig. 2A, this reaction is strongly stimulated by DNA and by addition of 45-protein. The same proteins stimulate T4 DNA polymerase utilization of long single-stranded DNA templates (Fig. 2B) and enhanced stimulation is obtained upon rATP addition. However, as illustrated in Fig. 3, no polymerase stimulation is seen when very short, single-stranded DNA molecules are used as templates.

Analysis by classical methods reveals that all the DNA product in Fig. 2B and 3 is template-linked in both stimulated and control reactions and that no new primer sites are being generated. Therefore, considering the results in Fig. 3, it seemed likely that the 44-62 and 45 proteins interact with T4 DNA polymerase to increase its average "sticking-distance," or processiveness. That this will give stimulation of net synthesis only on long template strands is clear from the model shown in Fig. 4. This model assumes that initiation of synthesis by polymerase is rate-limiting under our conditions and that the polymerase then continues along the DNA template for some distance be-

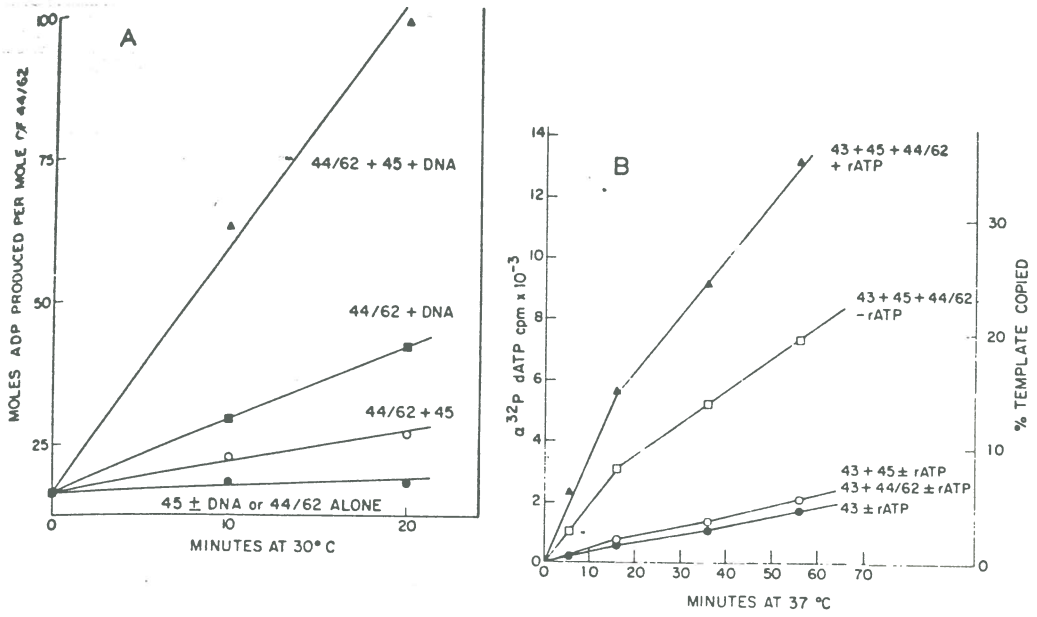


Figure 2. (A). Time course for the hydrolysis of rATP by 44-62 protein. The complete reaction contained 0.25 mM [^{3}H]-rATP, 10 mM Tris-Cl pH 7.4, 25 mM KCl, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 400 µg/ml BSA, 7.5 µg/ml calf-thymus single-stranded DNA, 10 µg/ml 45 protein, and 30 µg/ml 44-62 protein. Aliquots were spotted directly on to PEI plates, overlayered with a mixture of carrier AMP, ADP and ATP, and then developed in 1 M LiCl. The spots seen under U.V. light were cut out and counted.

(B). Time course for DNA synthesis on T4 single-stranded DNA. The reaction contained 10 mM Tris-Cl (pH 7.4), 25 mM KCl, 5 mM $MgCl_2$, 5 mM 2-mercaptoethanol, 40 µM each dNTP (including [α^{32}P]-dATP), 250 µg/ml BSA, 5 µg/ml T4 single-stranded DNA, 0.1 µg/ml T4 DNA polymerase, and where indicated: 5 µg/ml 45 protein, 30 µg/ml 44-62 protein, and 0.25 mM rATP. Aliquots were removed at the indicated times and assayed for acid-insoluble ^{32}P.

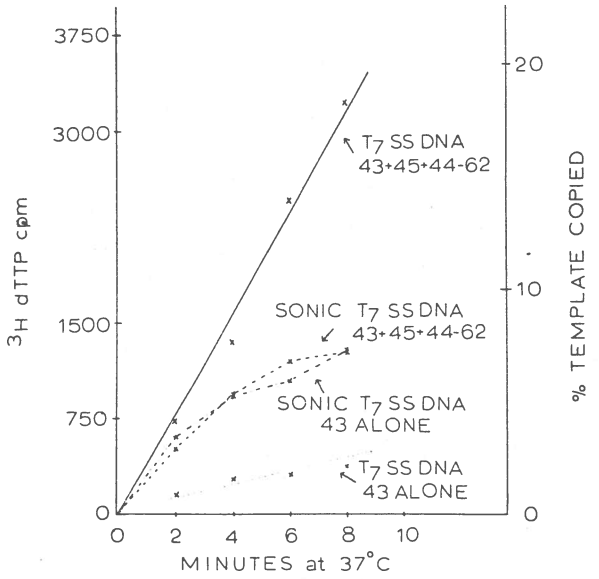
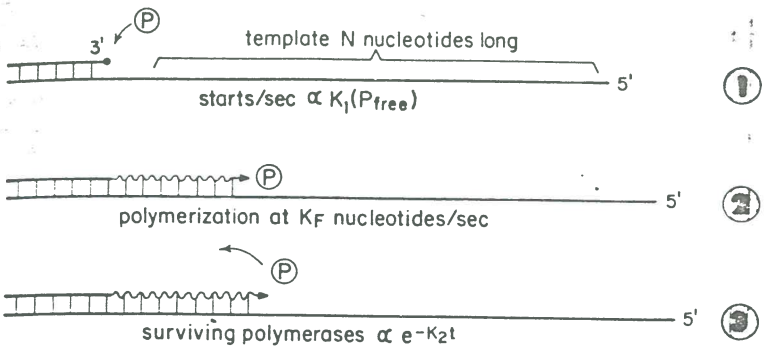


Figure 3. Time course for DNA synthesis on T7 single-stranded DNA molecules of different size. All reactions contained 10 mM Tris-Cl pH 7.4, 25 mM KCl and 2 mM $MgCl_2$, 0.5 mM dithiothreitol, 0.5 mM rATP, 50 μM each dNTP, 75 μg/ml BSA and 0.7 μg/ml T4 DNA polymerase. In addition, as indicated, the reactions contained as template either 5 μg/ml intact T7 single-stranded DNA or 5 μg/ml sonicated T7 single-stranded DNA, plus 138 μg/ml 45 protein and 85 μg/ml 44-62 protein.

The relatively rapid rates seen with the "sonic DNA" template are due to its enrichment for 3'-OH terminal priming sites, which allows more frequent polymerase starts.



$$\text{Macroscopic Rate} \propto \frac{K_1}{K_2}(P_{free})K_F\left(1-e^{\frac{-K_2N}{K_F}}\right)$$

Figure 4. A proposed microscopic DNA polymerase cycle. The T4 DNA polymerase is viewed as starting on a 3'OH primed single stranded template with a rate proportional to the concentration of such 3'OH termini. Polymerization then proceeds with a uniform rate (K_F)until terminated by one of two alternate paths. The first, direct dissociation from the middle of a template chain, is typically first order, while the second path is due to polymerases running off the template end. On very large templates the latter pathway is insignificant under our reaction conditions. The total rate of incorporation ("macroscopic rate") is determined by the number of polymerases bound to a primer end (i.e., one that can be extended by the addition of at least one more deoxyribonucleotide) multiplied by the rate of movement of polymerase, K_F. The final expression is seen to include an exponential term which reflects the relative rates of the above two termination processes.

fore falling off. It is clear that increasing K_F (the microscopic polymerization rate) and/or decreasing K_2 (the falling off rate) will increase the average polymerase "sticking-distance". This in turn will stimulate the macroscopic rate of DNA synthesis, providing that the template length, N, is longer than the typical sticking-distance for polymerase alone.

Our attention was originally directed to the question of DNA polymerase processiveness by Dr. Lucy Chang, who found that calf DNA polymerases work processively for less than 5 seconds (her limit of resolution) before falling off (19). To explain the differential stimulation of the macroscopic synthesis rate for long templates observed (Fig. 3), calculations from the model in Fig. 4 demand a sticking length of at least 1,500 nucleotides for T4 DNA polymerase alone. This led us into a detailed investigation aimed at determining K_F and K_2 in the T4 system.

Fig. 5 is a histogram of the length of double-stranded DNA product made during 5 minutes of incubation, as measured by electron miscroscopy. Under the conditions used, an average of only one template molecule in five had a polymerase initiation. Thus, Fig. 5 may be viewed as a display of microscopic "sticking lengths" in the stimulated and control reactions. Note that the sticking-length for polymerase alone clearly exceeds 1,500 nucleotides, as required to explain the results in Fig. 3. Although the double-stranded DNA lengths are quite heterogeneous in the stimulated reaction, a clear and striking increase in average product length is seen (apparently only about half of the polymerase molecules are affected by the 44-62 and 45 protein additions). Assuming the same number of polymerase initiations in the stimulated and control reactions (as in Fig. 4), we calculate a 4.5 fold stimulation in macroscopic synthesis rate from the histogram in Fig. 5, in agreement with the 3.9-fold increase actually observed.

Our interpretation of the experiment in Fig. 5 is in agreement with direct measurements of the rate of polymerization in the stimulated and unstimulated reactions obtained with 2 to 20 second incubations, to be described in detail elsewhere (D.M. and B.A., in preparation). Here, using electron microscopy or physical sizing of S_1 nuclease-resistant product as independent measures of product size, we estimate for T4 DNA polymerase alone a microscopic polymerization rate (K_F) of $\sim$250 nucleotides/sec and an

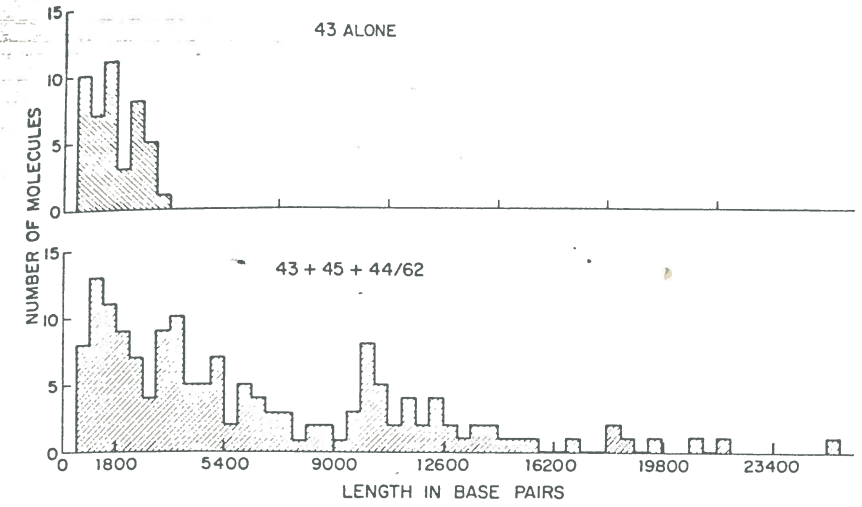


Figure 5. The lengths of double-stranded regions in in vitro DNA polymerase reaction products as measured by electron microscopy. The reactions were carried out essentially as described in Fig. 2B, except that 2 mM $MgCl_2$ was present and 5 μg/ml single-stranded T7 DNA was used as template. For the lower panel, 7.5 μg/ml 45 protein and 9 μg/ml 44-62 protein was added. The reactions were stopped at 5 min by adding an equal volume of cold 0.2 M EDTA, and spread by the Inman procedure (24).

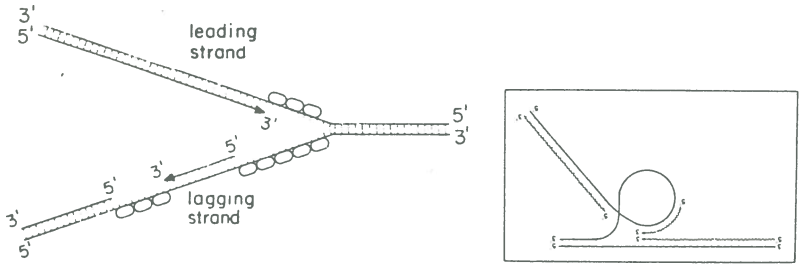
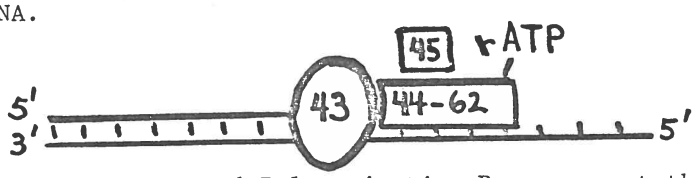


Figure 6. Schematic representation of a general DNA replication fork with DNA-unwinding protein binding to the single-stranded DNA exposed by fork movement. The insert illustrates an alternative visualization of the same replication fork, in which DNA synthesis on the "leading" and "lagging" strands is coupled. In this latter form, one can easily rationalize the existence of discontinuous synthesis on both sides of a fork, as observed (27; see text).

off-rate (K_2) of ~0.1 sec $^{-1}$ under our conditions (see Figs. 2B and 3). Addition of 44-62 and 45 proteins increases the K_F to 800-900 nucleotides/sec, without a major effect on K_2.

Note that these values are much greater than the 10 nucleotides/sec commonly suggested for K_F from standard macroscopic rate measurements (13). This means that T4 DNA polymerase runs at close to the in vivo rate of polymerization (estimated from the data of R. Werner (20) at 1,000 nucleotides/sec in our simple in vitro system, even without the further stimulation expected from addition of 32-protein (13). We conclude, therefore, that the other proteins of the replication apparatus need only decrease K_2 (by tying down polymerase) to account for the observed speed of in vivo fork movement.

In other experiments, we have been able to demonstrate tight binding of 44-62 protein, but not 45 protein, to single-stranded DNA in a reaction requiring 37° incubation (unpublished results of M. Davies). For this reason, our tentative view of the DNA polymerase interaction is that schematically illustrated: 44-62 is viewed as holding onto both the template strand and the DNA polymerase, and its ATP hydrolysis serves to generate protein conformations which help move the DNA polymerase more quickly along the DNA.



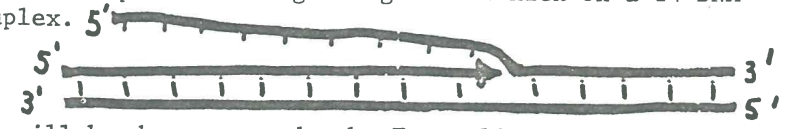
Modeling Suspected Polymerization Processes at the Replication Fork. Based upon work in many laboratories, a schematic representation of a replication fork can be drawn (Fig. 6). Because of the anti-parallel orientation of strands in the DNA double-helix, a discontinuous mode of DNA synthesis is necessary on at least one template strand if all of the DNA is to be made with a proof-reading, 5' → 3' DNA polymerase (21). Biochemical analysis of initial DNA products (22, 23) and electron microscopy of forks (24-25) indicate lengths of 1 to 2 x 10^3 nucleotides between adjacent polymerase starts on the "lagging side" of the fork. On the "leading side" of the fork, synthesis also appears to be discontinuous (26,27). The restarts on the leading strand, however, need not lead to the large single-stranded gaps observed in the lagging strand intermediates. This single-

stranded DNA is almost certain to be covered with tightly bound DNA-unwinding protein. In addition, we propose that 32-protein has a helix-opening role in leading strand synthesis. [Opening of the helix ahead of the fork seems essential, since double-helical templates cannot be copied by the T4 DNA polymerase directly (8)].

The model illustrated in Fig. 6 assumes that two DNA polymerase molecules are working at any one time in the fork. The 44-62 ATPase would be expected to be traveling along with one or both of these. (Note that these two polymerization sites could be physically linked together, as suggested in the Insert to Fig. 6). In addition, because T4 DNA polymerase is unable to start chains de novo, it seems likely [by analogy with observations in other systems (28-30)] that discontinuous synthesis is primed by a special piece of RNA made by the replication proteins on single-stranded template DNA generated at the fork. After this RNA serves as a primer for the DNA polymerase, it must be erased by another enzyme and replaced by DNA. We feel that this mechanism makes good biological sense, since it can be argued that the enzyme that synthesizes the primer will necessarily make a relatively inaccurate copy; thus, the choice of RNA (rather than DNA) as primer automatically marks these sequences as "bad copy" to be removed (for details, see Ref. 21).

In the simplest view, one replication protein (e.g., gene 41 protein?) might function solely to synthesize an RNA primer. To test this idea, we set out to study the predicted reactions on the two sides of the replication fork in Fig. 6 separately. For lagging strand synthesis one might expect to obtain ribonucleoside triphosphate dependent de novo chain initiations on a T4 single-stranded DNA template (see also article by Bouche and Kornberg, this volume):

For leading strand synthesis one might anticipate rapid strand displacement beginning from a nick on a T4 DNA duplex.



As will be demonstrated, the T4 replication proteins cata-

lyze both of the above model reactions in an efficient manner. Although we initially expected that different combinations of replication proteins would be required, both the leading and lagging-strand reactions turn out to require all six of the T4 proteins introduced in Fig. 1 for maximal DNA synthesis.

A "Lagging-Strand" Model Reaction. Incorporation data strongly suggesting RNA-primed DNA chain starts *in vitro* are presented in Fig. 7. Using alkali-denatured T4 single-strands as a template one obtains extensive DNA synthesis which requires the presence of all six of the replication proteins. As shown, this synthesis is completely dependent upon the addition of a mixture of rCTP, rGTP and rUTP. Alkaline sucrose gradient sedimentation reveals that the small amount of product made in the absence of either rCTP, rGTP and rUTP or 32 protein cosediments with the template (as expected for template priming by fold-back at the 3' end). In contrast, essentially all of the product made at early times in the complete reaction sediments with a mean chain length much smaller than the template (data not shown). These results suggest that RNA-primed *de novo* initiation of new DNA chains is occurring. This reaction proceeds under conditions which approximate intracellular concentrations of salts and with protein concentrations which are not greater than those found *in vivo* (see Fig. 7, heading). Therefore, we believe that it is biologically significant. However, the sites at which the presumed primer is being initiated are not restricted to T4 DNA. Shown in Table 1 are results obtained with T4, fd, T7 and synthetic DNA single-strands. We find that the ribo-dependent reaction similar to that shown in Fig. 7 proceeds on many DNA templates (e.g., T4 and fd) but that not all DNAs are active (e.g., T7 and synthetic DNAs). Thus, we conclude that under our conditions, the T4 replication proteins are recognizing some special sequence or structure in this priming reaction. To explain why all six T4 proteins are required, we suggest that a special multienzyme complex is formed between them, and that this is needed to generate the correct environment for the postulated synthesis and utilization of the primer RNA.

Note that the regular DNA-dependent RNA polymerase utilized for transcription is not involved in our reactions: we are unable to detect the activity of this enzyme in any

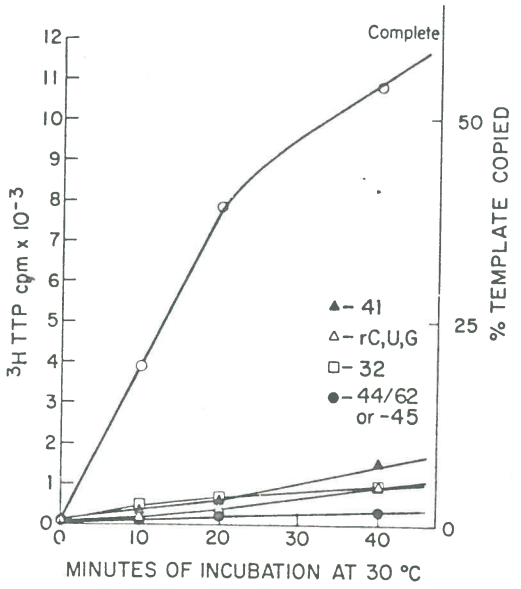


Figure 7. Time course of ribonucleotide-dependent DNA synthesis on a single-stranded T4 DNA template. The "complete" reaction components consist of 43 protein 1 µg/ml, 41 protein 80 µg/ml, 44-62 protein 8 µg/ml, 45 protein 25 µg/ml, 32 protein 125 µg/ml, dNTP's 0.1 mM each, rATP 0.5 mM, nuclease-free BSA 50 µg/ml, $MgSO_4$ 5 mM, KCl 25 mM, Tris-Cl (pH 7.4) 20 mM, dithiothreitol 0.5 mM, alkali-denatured (44) T4 single-stranded DNA 10 µg/ml, and rCTP, rUTP, rGTP 0.1 mM each. [^{3}H]-TTP was used to label DNA products made. For the incomplete reactions, only the specified component was deleted. At the specified times, aliquots were removed and incorporation into acid insoluble material measured.

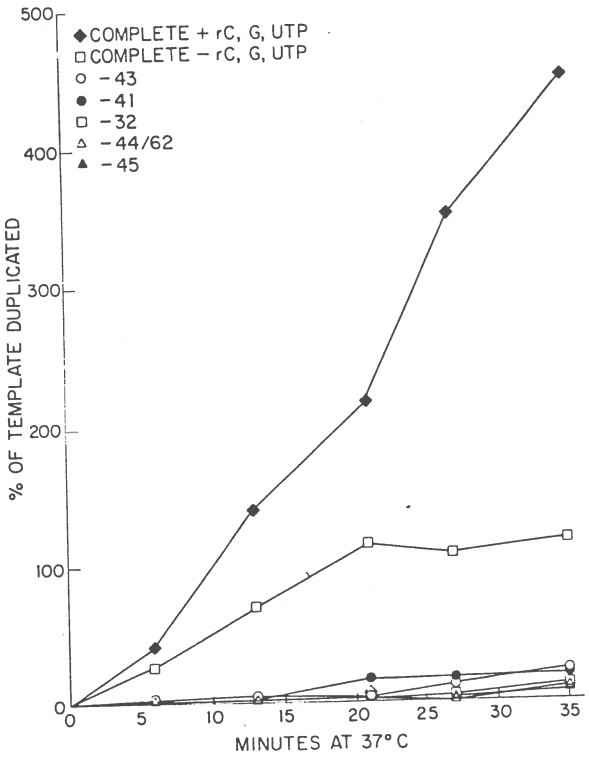
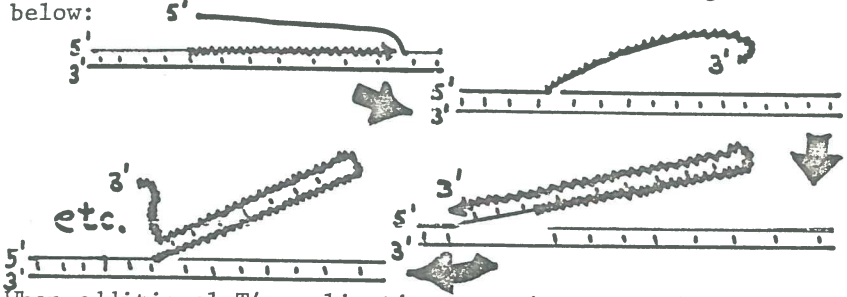


Figure 8. Time course of synthesis on a double-stranded T4 DNA template. The complete reaction components were essentially the same as those specified in Fig. 7, the major changes being addition of K^+ to 67 mM, substitution of 4 μg/ml intact T4 double-stranded DNA as template, and removal of rCTP, rGTP and rUTP. For incomplete reactions, only the specified component was deleted. Note that only the one reaction indicated contained the three ribonucleoside triphosphates.

of our purified proteins; moreover, the ribonucleoside triphosphate requirement for DNA synthesis is retained even after addition of high levels of rifampicin (10 μg/ml) or antibody to T4 modified RNA polymerase.

A "Leading-Strand" Model Reaction The T4 DNA polymerase by itself will not utilize nicked double-helical DNAs as a template, as it is unable to copy base-paired regions (8). However, in 1974 Nancy Nossal made the important observation that addition of high concentrations of 32 protein allows a limited amount of DNA synthesis to begin on such templates (31). Most of the product DNA made is reversibly denaturable, re-forming hairpin-type double-helical structures after heating and quick cooling (31). This is the expected result if branch migration (32) repeatedly interrupts the polymerization process (33,34), as diagrammed below:



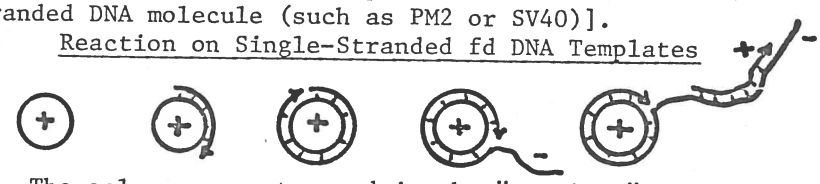
When additional T4 replication proteins are added to this system, such interruptions are apparently prevented. One observes, instead, the rapid and efficient strand-displacement reaction expected for leading strand synthesis. All double-stranded DNAs thus far tested function as templates for this synthesis, including the DNA from bacteriophages T4, T7, lambda, PM2 and G4; E. coli; SV40 virus; and monkey cells. The protein requirements for extensive DNA synthesis are illustrated in Fig. 8, where double-stranded T4 DNA has been used as the template. Once again, all 6 T4 replication proteins are required; note that an amount of product equal to the input template is made within 20 min at 37°. Alkaline sucrose gradient analysis demonstrates that long continuous DNA strands are being synthesized, and test of the S_1 nuclease susceptibility of product and template revealsthat more than 99% of each is fully denaturable at all extents of reaction (data not shown).

Further addition to this system of a mixture of rCTP,

rGTP and rUTP stimulates the observed rate of DNA synthesis between 2- and 5-fold, depending on the exact conditions and template used (e.g., ∿2-fold in Fig. 8). Clearly, part of this stimulatory effect is due to the ribo-dependent lagging strand synthesis which ensues (see below), but the fact that the effect can be greater than 2-fold leaves open the possibility that another ribo-dependent reaction is additionally involved.

Complete Reactions When leading and lagging strand reactions proceed simultaneously, they should generate a replication fork resembling that described earlier in Fig. 6. This is readily seen when circular templates such as PM2 or fd bacteriophage DNA are used. We shall demonstrate that, following de novo chain initiation on a fd template, polymerization proceeds around the circular, single-stranded DNA molecules to yield a circular double-helix with one strand-specific interruption. Further synthesis on some of these double-stranded circles induces strand displacement and generates a "rolling circle" (35) mode of DNA synthesis. Because the initial template is exclusively (+) strand, the 32-protein coated, single-stranded regions (formed as intermediates on these rolling circle tails) are located exclusively on the (-) strand. These regions therefore cannot renature with each other to confuse the analysis. [In contrast, renaturation between rolling circle tails can be a problem when the initial template is a circular double-stranded DNA molecule (such as PM2 or SV40)].

Reaction on Single-Stranded fd DNA Templates



The only components used in the "complete" reaction on fd DNA templates are the 6 replication proteins, rNTPs, dNTPs, bovine serum albumin (BSA), Mg^{++} and salts. If any of the T4 proteins (44/62, 43, 41 or 45) is left out of the reaction mixture, one finds less than 1% the amount of DNA synthesis seen in the complete reaction. Therefore, each is clearly essential. As can be seen in Fig. 9, there is also a dramatic decrease in synthesis when rUTP, rGTP and rCTP are omitted. [The rATP is not omitted, since a role for it in stimulation of DNA synthesis with proteins 43, 44/62 and 45 has already been established (Fig. 2B)]. This

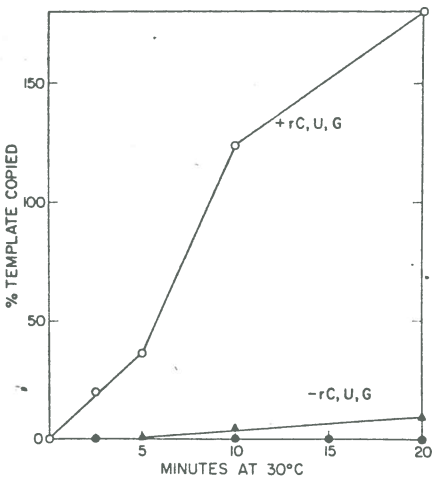


Figure 9. Time course of the complete reaction using [^{3}H]-labeled fd single-stranded circular DNA template. The complete reaction components were essentially the same as those designated in Fig. 7. [α ^{32}P]-dGTP was used to label DNA products made. The +rUTP, +rGTP. +rCTP labeled data represents a time course of the complete reaction and the -rUTP, -rGTP, -rCTP data represents the time course observed if these components are omitted.

TABLE 1

Template	% DNA made relative to template	
	+ rUTP, rGTP, rCTP	− rUTP, rGTP, rCTP
	4' at 37°C	
Fd	101%	3%
T_7 - H strand	5%	4%
T_7 - L strand	5%	7%
	20' at 30°C	
SS T4 DNA	35%	0.6%
Poly dI	0.7%	2%
Poly dG	1%	0.6%
Poly (dI, dT)	0	0

For conditions used, see Fig. 9.

strong requirement for ribonucleotides suggests that an RNA primer is being made. However, in preliminary attempts, it has not been possible to detect a significant fraction of radioactive ribonucleotides incorporated into the DNA product. We suspect that RNA primer is both being made and degraded in our system, perhaps by a RNAse activity intrinsic to the purified proteins.

Since synthesis continues far past the level equivalent to a complete copying of the fd template (Fig. 9), this system must do more than simply convert the initial template to a circular double-helical DNA molecule. To examine this reaction in more detail, ^{3}H-labeled fd DNA was used as template and ^{32}P-labeled deoxyribonucleoside triphosphates incorporated. Template & product could then be separately analyzed by sedimenting the reaction mixture through alkaline sucrose gradients. The results are shown in Fig. 10. As expected, the sedimentation value of the ^{3}H-template remains unchanged during the course of the reaction. Since it never cosediments with the bulk of the product DNA, it cannot be covalently attached to it. The average strand length of the ^{32}P-labeled product, however, is seen to grow with time. The product made during the first 2.5 minutes is the size of a linear fd genome or smaller (Fig. 10a). By 10 minutes, a small amount of product is seen sedimenting faster than fd circular template (Fig. 10b), and by 20 minutes almost half the product (45% in this experiment) is found to sediment faster than the template strands (Fig. 10c). These results are those expected if, following de novo initiation of (-) strand synthesis on the fd circular template [(+) strand], this (-) strand grows to multiple fd lengths by continuous (-) strand displacement as the circular (+) strand template "rolls" (see schematic diagram in text, above). This basic conclusion is supported by separate experiments in which DNA-DNA hybridization has been used to determine the strandedness of the largest and smallest DNA products isolated from alkaline sucrose gradients (unpublished results of C.F.M.).

Direct S_1 nuclease digestion (in 0.2% SDS, pH 4.6) of the DNA products obtained at various times during the course of reaction in Fig. 9 revealed that while only 35% of the ^{3}H template is resistant to S_1 nuclease digestion at 2.5 minutes, 70% is resistant at 5 minutes and nearly 100% at 10 minutes. As might be expected, the initial ^{32}P pro-

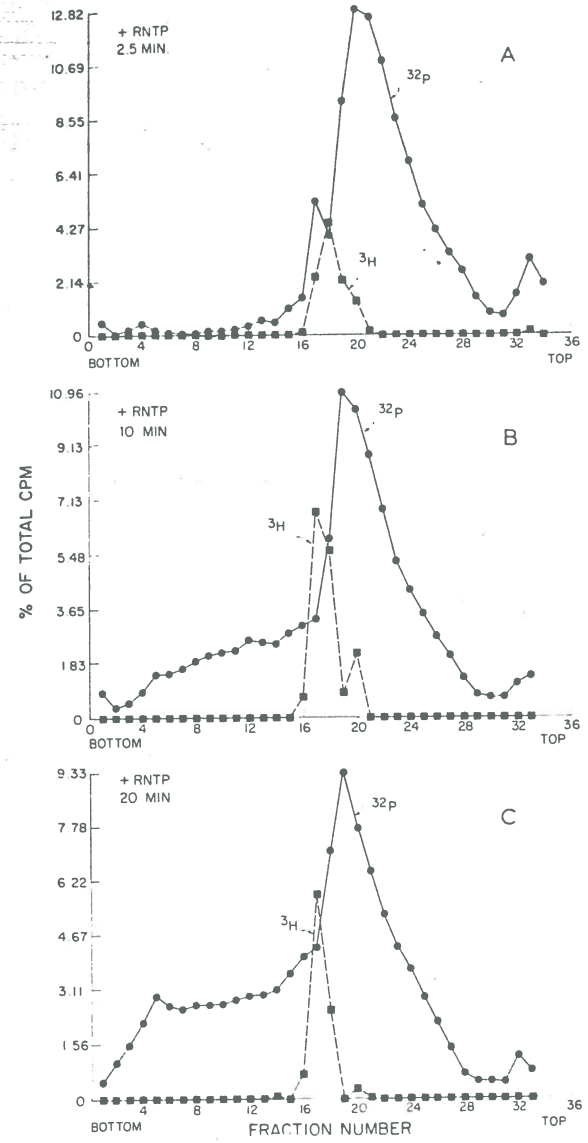


Figure 10. Alkaline sucrose gradient sedimentation of the reaction products made on a [^{3}H]-labeled fd DNA template, as described in Fig. 9. At the indicated times, aliquots were removed from the reaction mix into an equal volume of cold 20 mM Na_3EDTA. Samples were then layered on 5-20% alkaline sucrose gradients (0.8 M NaCl, 5 mM EDTA, 0.2 M NaOH) with 60% sucrose shelves and centrifuged for 3 hours at 50,000 rpm, in a Spinco SW 50.1 rotor (4°C). Fractions were collected from the bottom onto glass fiber filters, and washed and counted by standard techniques. Counts were plotted by computer after overlap corrections. (The fraction of [^{3}H]-template counts has been adjusted to 10% of their real value in order to allow the product counts to be better visualized).

duct made is almost completely resistant (92%) to S_1 digestion. At later times, the fraction of resistant product decreases, but reaches an apparent limit where 75% of it is still double-stranded. Since product DNA would be continually displaced as single-strands from a rolling circle, we conclude that most of it must be used as a template for synthesis of its complementary strand after it enters the tail.

Electron Microscopy of the Reaction Products For electron microscopy, aliquots were removed from reactions on an fd DNA template and examined after cytochrome spreading on a formamide hypophase, using the modified Kleinschmidt procedure developed by R. Inman (36). This type of analysis is not strictly quantitative, since a substantial fraction of DNA molecules are not well enough spread to be scored, and thus a special minor class of product can easily be overlooked. Moreover, the shear forces generated by spreading appear to fragment some of the longer DNA molecules (especially those containing substantial single-stranded interruptions). Therefore, although simple linear double-stranded products are seen, it is not clear that they were present as such before spreading.

Despite these qualifications, it is clear that a major class of product DNA has precisely the geometry predicted from the sedimentation analysis described earlier (Fig. 10). As illustrated in Fig. 11, double-stranded DNA circles of fd length are the predominant structures seen at early times in the reaction. On further incubation, long double-stranded DNA tails are seen attached to some of these circles (Figs. 12 and 13). As expected from the rolling circle model, the DNA at the point of attachment of circle and tail is nearly always single-stranded. Moreover, double- and single-stranded regions frequently alternate in this region, as in the example shown in Fig. 14. We therefore conclude that the (+) strand complement made on the single-stranded tail is synthesized by a discontinuous mechanism, as predicted (see lagging strand in Fig. 6).

Similar structures to those shown in Figs. 12 and 13 are seen whether fd or G4 bacteriophage DNAs (single-stranded), or PM2 bacteriophage or SV40 virus DNAs (double-stranded) provide the initial circular template. In all cases, although an amount of DNA equivalent to multiple copies of the template is made in the first 30 min of incubation,

only a small percentage of double-stranded circular molecules appear as circles with attached tails. In a typical electron microscopic field, many of these tails are greater than 50 microns in length (10^8 daltons). Thus, the mechanism of polymerization at these replication forks appears to be highly processive, with a calculated rate of polymerization per active rolling circle of at least 600 nucleotides/sec at 37°. Note that this is close to the rate at which the T4 replication fork moves *in vivo* (20).

The Fidelity of Synthesis The most sensitive test for fidelity in this system would be to quantitate the biological activity of the DNA product relative to that of the input template. While these studies are not yet completed, we have determined that all (>99%) of the product DNA made is denatured by 3 minutes of heating at 100° followed by rapid cooling, as judged by its becoming sensitive to S_1 endonuclease degradation. Thus, our products do not contain the hairpin inversions found when either *E. coli* DNA polymerase I (33,34) or T4 DNA polymerase plus 32-protein (31) utilize double-stranded DNA as template. Moreover, Fig. 15 shows a sedimentation analysis with and without R_I restriction endonuclease treatment of double-stranded DNA synthesized *in vitro* using single-stranded DNA from G4 phage as template. As expected for faithful copying, the fast-sedimenting rolling circle products made on either G4 or SV40 virus DNA templates [which contain a single $EcoR_I$ restriction site (37,38)] are cut to unit-length linear double-helices by this enzyme treatment (SV40 data not shown). Finally, analysis by CsCl gradient equilibrium sedimentation of products made with deoxybromouridine triphosphate replacing TTP reveal a classical semiconservative pattern of DNA synthesis (unpublished results, N.S.). We conclude therefore that double-helical DNA templates are copied with a reasonable degree of fidelity in our *in vitro* system.

Future Directions After a T4 bacteriophage DNA molecule enters an *E. coli* cell, it synthesizes replication proteins and these are then directed so as to selectively replicate only the T4 genomes present. The *E. coli* DNA is not used as template, even though it is in great excess over phage DNA. This remains true for T4 mutant infections in which this host DNA is not significantly degraded (39). It therefore seems likely that special DNA sequences or structures

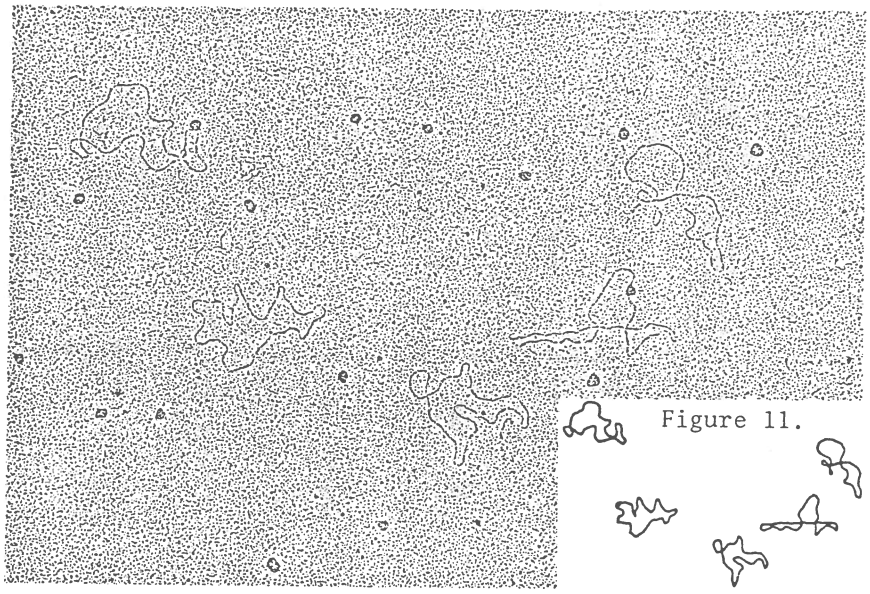


Figure 11.

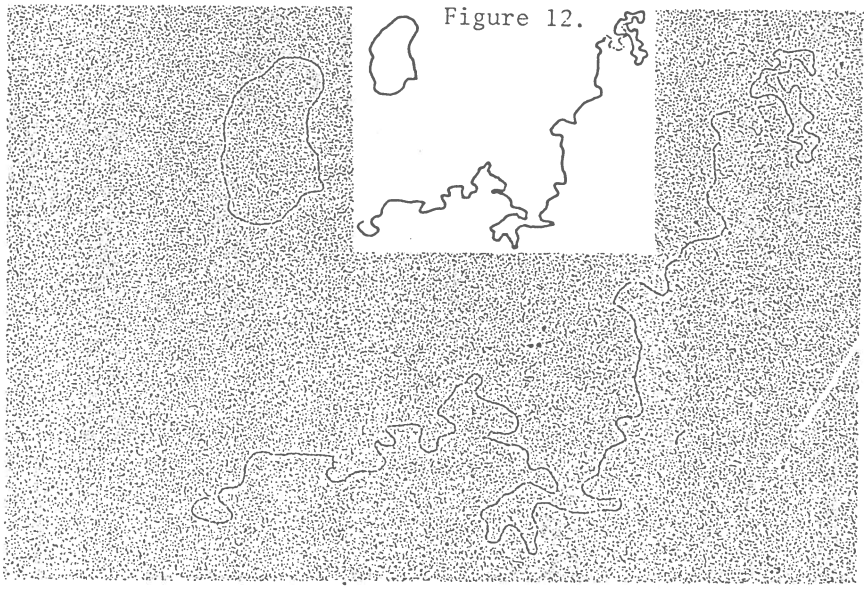


Figure 12.

Figures 11-14. Electron microscopic analysis of reaction products. From the reaction described in Fig. 9, 5 µl of diluted sample was added to 5 µl of freshly prepared buffer [(4 ml H_2O, 0.3 ml 1M Na_2CO_3, 0.4 ml Na_2EDTA 0.13M); the high pH of this buffer is designed to eliminate 32 protein binding to single-stranded DNA]. 10 µl of formamide is then added and gently mixed, followed by

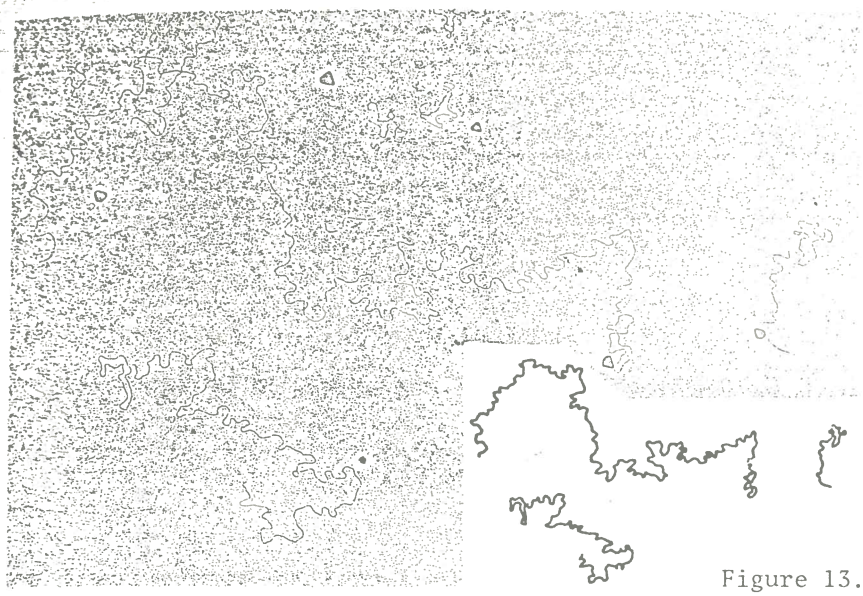
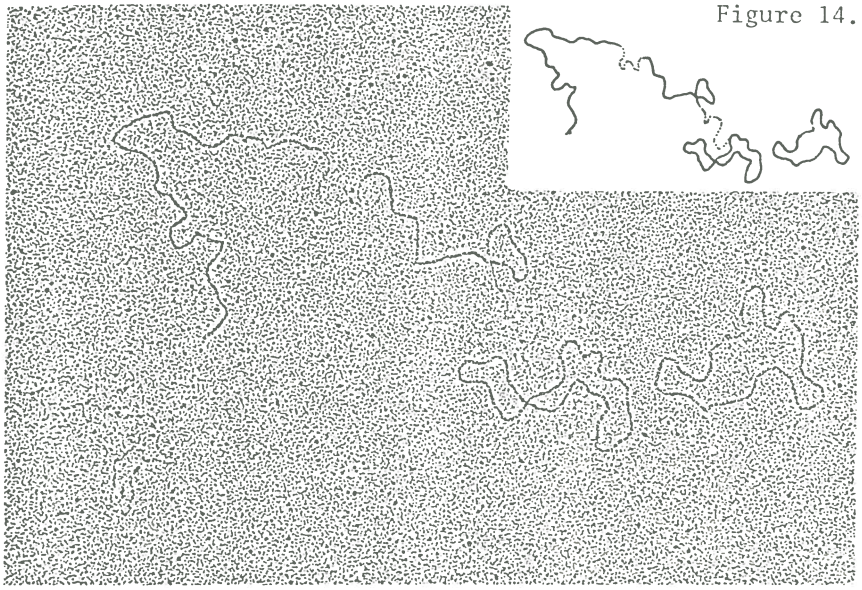


Figure 13.

Figure 14.



addition of 2 µl of 0.1% cytochrome C. The DNA is then spread, platinum shadowed, and placed on copper grids for examination in the electron microscope, essentially as described by Inman (24). Fig. 11 is from the 5 min time point, while the remaining micrographs are from the 10 and 20 min time points.

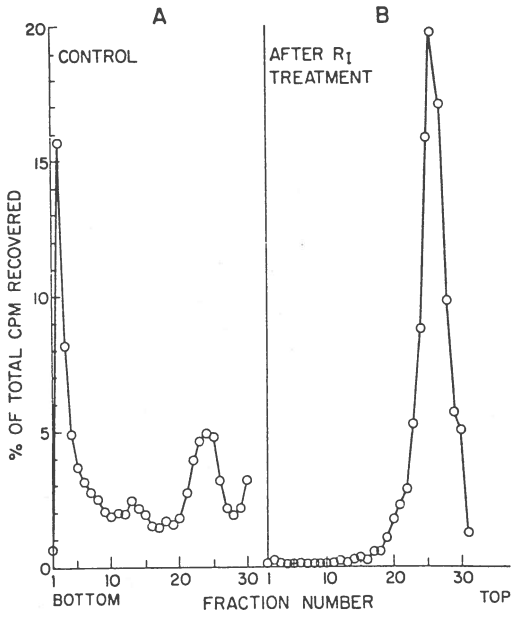


Figure 15. Sedimentation analysis in neutral sucrose gradients of DNA synthesized *in vitro* using single-stranded G4 phage DNA as template under conditions similar to those described in Fig. 8. The reaction mixture contained 0.2 mM each rCTP, rGTP, and rUTP; and α^{32}P-dGTP was used to label the *in vitro* product. After incubation at 37°C for 30 minutes a 10 µl aliquot was withdrawn [for part (a)]. Purified (non-specific endonuclease free) $EcoR_I$ restriction endonuclease was added [for part (b)]. After another 15 minutes of incubation at 37°C the reactions were stopped by adding an equal volume of 20 mM Na_3EDTA and the samples layered on neutral gradients (5-30% in a buffer containing 20 mM Tris-Cl, pH 8.1, 1 M NaCl, 1 mM EDTA and 0.1% Sarkosyl). A 0.3 ml cushion was made by mixing equal volumes of 80% sodium iothalamate (Angio-Conray) and 30% sucrose. The gradients were centrifuged for 90 minutes at 50,000 rpm and 4°C in a Spinco SW 50.1 rotor.

on the T4 genome are required for initiation of new replication forks in vivo.

In contrast, our in vitro T4 system appears to be able to establish replication forks almost equally well on most double helical DNA molecules. Although not yet proven, all our results are consistent with the hypothesis that these forks arise by strand displacement from a nick, followed by de novo initiation of complementary strand synthesis on this displaced single-strand. While this mode of replication fork initiation must be strongly suppressed in vivo (possibly by special blocking proteins), it conveniently allows us to bypass the (presumably complex) biological pathway for T4 fork initiation, and enables the mechanism of fork propagation to be examined in a relatively simple in vitro system.

The characteristics of the DNA synthesis described above make it appear that, despite its apparent non-biological start, the propagation reaction seen in vitro closely mimicks that found in vivo. However, several important questions remain to be resolved:

(1) Our inference that a short RNA primer is made and degraded in this system remains to be proven by direct isolation and characterization of such RNA.

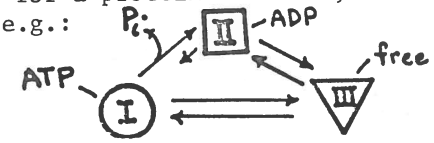
(2) In vivo, the DNA on the leading, as well as on the lagging strand in Fig. 5 appears to be made discontinuously, with about the same average piece size for each (27). This is perhaps best explained by postulating some direct coupling of the synthesis on the two strands, such as that schematically illustrated in the Insert to Fig. 6. In this scheme, both leading and lagging strand DNA polymerase molecules are joined (via other replication proteins). As soon as the lagging strand polymerase runs out of single-stranded template to copy (by reaching the 5' end of the previous Okazaki piece), the postulated structure must rearrange. This could force interruption of both leading and lagging strand synthesis simultaneously, with de novo starts on each strand.

Although we find that most of the (-) strand in the fd reaction rolls out as a continuous molecule of multiple fd length (e.g., Fig. 10), sealing of discontinuous starts on this leading strand could be fast enough to mask their presence(see article by Sternglanz, this volume). Thus, to determine whether strand coupling occurs in our in vitro system will require that we prepare each of the replication

proteins in a completely DNA ligase-free state, a task not yet accomplished.

(3) The strongly synergistic effects of the 6 T4 replication proteins implies that they interact with each other to make a large complex of unique structure. All our attempts to isolate this structure have thus far failed, suggesting that at least some of the interactions are quite weak. Weak interactions might also explain why rather high concentrations of some of the replication proteins are required for maximal activity (see Fig. 7, heading). It is clearly essential to identify all of the interactions and the exact roles of each individual protein in the replication process. This remains as a challenging task for future research.

(4) It is not clear why nature requires so many proteins to make DNA. In particular, one might wonder about the necessity for the 44-62 DNA-dependent ATPase, a large protein of about 180,000 daltons. (Note that DNA-dependent ATPases are also found among the E. coli replication proteins: see article by Wickner and Hurwitz, this volume). From the most general point of view, the efficacy of using bound nucleotide hydrolysis to drive an ordered series of allosteric changes in protein conformation has been recognized for many years (40). For example, consider three conformational states for a protein molecule, determined by bound nucleotide, e.g.:



If hydrolysis of ATP occurs when it is bound to the protein the transition between state I and state II will be strongly biased towards II. Such hydrolysis is thereby sufficient to drive the entire conformational cycle clockwise. In this way, a protein molecule can readily do useful work (e.g., pumping small molecules unidirectionally through a membrane, walking unidirectionally down a DNA chain, etc.). Such miniature "protein machines" play important roles throughout living systems (40). In our case, ATP hydrolysis by the 44-62 complex may serve to set up the clock mechanism needed for a pre-incorporation "kinetic-proofreading" scheme (41). The result would be a substantial increase in the fidelity of the base-pair fit prior to polymerase action, analogous to the fidelity increasing

mechanism proposed by Hopfield for several other important biosynthetic reactions (41). Thus, much of the complexity of the replication machinery may reflect the necessity of achieving a very accurate DNA copy.

ACKNOWLEDGEMENTS

We gratefully acknowledge the skillful technical assistance of Monique Davies, Judy Goldberg, Linda McAfee and Mei Lie Wong. This work was supported by grants from the National Institutes of Health (PHS-GM 14927) and the American Cancer Society (NP 152) to Bruce Alberts. Navin Sinha and Fred Morris were supported by postdoctoral fellowships from the National Cancer Institute (fellowship #'s 5F22CA01829 and 6F22CA03266 respectively).

REFERENCES

1. Epstein, R. H., Bolle, A., Steinberg, C. M., Kellenberger, E., Boy de la Tour, E., Chevallez, R., Edgar, R.S. Susman, M., Denhardt, G. H. and Lielausis, A., Cold Spring Harbor Symp. Quant. Biol. 28, 375 (1963).

2. Warner, H. R. and Hobbs, M. D., Virology 33, 376 (1967).

3. Riva, S., Cascino, A. and Geiduschek, E. P., J. Mol. Biol. 54, 85 (1970).

4. Wood, W. B. In "Survey of Genetics" (King, R. C. ed.), Plenum, N. Y. (1973).

5. DeWaard, A., Paul, A. V. and Lehman, I. R., Proc. Nat. Acad. Sci. U.S.A. 54, 1241 (1965).

6. Warner, H. R. and Barnes, J. E.,Virology 28, 100 (1966)

7. Englund, P. T., J. Biol. Chem. 246, 5684 (1971).

8. Goulian, M., Lucas, Z. J. and Kornberg, A., J. Biol. Chem. 243, 627 (1968).

9. Nossal, N. G. and Hershfield, M. S., J. Biol. Chem. 246, 5414 (1971).

10. Lehman, I. R., Methods in Enzymology 29, 46 (1974).

11. Alberts, B. and Frey, L., Nature 227, 1313 (1970).

12. Delius, H., Mantell, N. and Alberts, B., J. Mol. Biol. 67, 341 (1972).

13. Huberman, J., Kornberg, A. and Alberts, B., J. Mol. Biol. 62, 39 (1971).

14. Barry, J. and Alberts, B., Proc. Nat. Acad. Sci. U.S.A. 69, 2717 (1972).

15. Barry J., Hama-Inaba, H., Moran, L., Wiberg, J. and Alberts, B., In "DNA Synthesis In Vitro" (Wells, R. D. and Inman, R. B. eds.), Baltimore, Md. (1973).

16. Wiberg, J. S., Mendelsohn, S., Warner, V., Hercules, K. Aldrich, C. and Munro, J. L., J. Virol. 12, 775 (1973).

17. Karam, J. D. and Bowles, M. G., J. Virol. 13, 428 (1974).

18. Brutlag, D. and Kornberg, A., J. Biol. Chem. 247, 241 (1972).

19. Chang, L., J. Mol. Biol. 93, 219 (1975).

20. Werner, R., Cold Spring Harbor Symp. Quant. Biol. 33, 501 (1968).

21. Alberts, B. In "Molecular Cytogenetics" (Hamkalo, B. and Papconstantinou, J., eds.), Plenum. N. Y. (1973).

22. Sugino, A. and Okazaki, R., J. Mol. Biol. 64, 61 (1972).

23. Okazaki, R., Okazaki, T., Sakabe, K., Sugimoto, K., Kainuma, R., Sugino, A. and Iwatzuki, N., Cold Spring Harbor Symp. Quant. Biol., 33, 129 (1968).

24. Inman, R. and Schnos, M., J. Mol. Biol. 56, 319 (1971).

25. Wolfson, J. and Dressler, D., Proc. Nat. Acad. Sci. U.S.A. 69, 2682 (1972).

26. Sugimoto, K., Okazaki, T. and Okazaki, R., Proc. Nat. Acad. Sci. U.S.A. 60, 1356 (1968).

27. Sternglanz, R., This Volume (1975).

28. Reichard, P., Eliasson, R. and Söderman, G., Proc. Nat. Acad. Sci. U.S.A. 71, 4901 (1974).

29. Wickner, W., Brutlag, D., Schekman, R. and Kornberg, A., Proc. Nat. Acad. Sci. U.S.A. 69, 965 (1972).

30. Okazaki, R., This Volume (1975).

31. Nossal, N. G., J. Biol. Chem. 249, 5668 (1974).

32. Lee, C. S., Davis, R. W. and Davidson, N., J. Mol. Biol. 48, 1 (1970).

33. Inman, R. B., Schildkraut, C. L. and Kornberg, A., J. Mol. Biol. 11, 285 (1965).

34. Masamune, Y. and Richardson, C.C., J. Biol. Chem. 246, 2692 (1971).

35. Gilbert, W. and Dressler, D., Cold Spring Harbor Symp. Quant. Biol. 33, 473 (1968).

36. Inman, R. B., Methods in Enzymology 29, 451 (1974).

37. Morrow, J. F. and Berg, P., Proc. Nat. Acad. Sci. U.S.A. 69, 3365 (1972).

38. Godson, G. N. and Boyer, H., Virology 62, 270 (1974).

39. Snustad, P. and Conray, L. M., J. Mol. Biol. 89, 663 (1974).

40. Hill, T. L., Proc. Nat. Acad. Sci. U.S.A. 64, 267 (1969).

41. Hopfield, J. J., Proc. Nat. Acad. Sci. U.S.A. 71, 4135 (1974).

42. Mathews, C. K., J. Biol. Chem. 247, 7430 (1972).

43. Laemmli, U. K., Nature 227, 680 (1970).

44. Studier, F. W., J. Mol. Biol. 41, 189 (1969).