Reconstruction of bacteriophage T4 DNA replication apparatus from purified components: Rolling circle replication following *de novo* chain initiation on a single-stranded circular DNA template

(replication fork/ribonucleoside-triphosphate-dependent priming/DNA strand displacement)

CHARLES F. MORRIS, NAVIN K. SINHA, AND BRUCE M. ALBERTS

Department of Biochemistry, Princeton University, Princeton, New Jersey 08540

Communicated by A. D. Kaiser, September 15, 1975

ABSTRACT The protein products of T4 bacteriophage genes 41, 43, 45, 44, and 62 have been purified to near homogeneity using an assay which measures their stimulation of DNA synthesis in a crude lysate of Escherichia coli cells infected by an appropriate mutant phage. When all of these proteins and T4 gene 32 protein are incubated in the presence of deoxyribonucleoside and ribonucleoside triphosphates, extensive DNA synthesis occurs on both single and double-stranded DNA templates. Analysis of this in vitro system reveals most of the features attributed to in vivo DNA replication: (1) De novo DNA chain initiation is found on a single-stranded DNA template only if ribonucleoside triphosphates are present (as expected for RNA priming of Okazaki pieces on the "lagging" strand of a replication fork). (2) With single-stranded circular DNA as template, synthesis continues for many doublings. The products after extensive synthesis resemble a rolling circle as visualized in the electron microscope, with discontinuous "lagging" strand synthesis generating a long, unbranched double-stranded tail. The fact that all six mutationally identified T4 replication gene products are required for these syntheses suggests the existence of a large multienzyme complex, constituting the T4 replication apparatus.

Since 1971 efforts in this laboratory have been directed towards achieving *in vitro* reconstruction of the T4 bacteriophage DNA replication apparatus from purified components. From the elegant genetic studies on this virus, at least six major T4 gene products appear to function at the replication fork (1-3). We have now purified these six proteins to near homogeneity and report their catalysis of a concerted *in vitro* polymerization reaction which closely resembles that seen in DNA replication *in vivo*.

MATERIALS AND METHODS

The proteins 32, 44/62, and 43 were purified by modifications of published procedures (4-6). The proteins 41 and 45 were isolated with the aid of the "in vitro complementation assay" described previously (4, 7). The exact purification procedures for each of the proteins will be found elsewhere (manuscripts in preparation).

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 comigrated with the appropriate activity in the *in vitro* complementation assay.

DNA Isolation. The bacteriophage fd DNA was isolated from phage by phenol extraction following polyethylene glycol precipitation and banding in a CsCl gradient (8).

Polyacrylamide Gel Electrophoresis. The procedure used was essentially that described by Laemmli (9). The so-

dium dodecyl sulfate-acrylamide slab gel (10) was made with a separating gel of 12.5% acrylamide and a stacking gel of 6% acrylamide in buffers containing 0.1% sodium dodecyl sulfate.

S₁ Nuclease Digestion. S₁ endonuclease was isolated by the procedure of Vogt (11). The DNA sample to be digested was added to a sodium dodecyl sulfate–N,N'-bis(2-hydroxyethyl)glycine (bicine) buffer (0.22% sodium dodecyl sulfate, 11 mM bicine, titrated to pH 9.2 with NaOH) and incubated for 10 min at 25° to dissociate 32 protein from single-stranded DNA. S₁ reaction buffer (to give 60 mM sodium acetate, pH 4.6, 0.1 M NaCl, 5 mM ZnCl₂, 10% glycerol) was then added and the reaction mix was divided into two aliquots, one of which received S₁ nuclease and the other an equal volume of buffer. Following incubation for 10 min at 37°, identical aliquots were removed from each of the two mixes (+ and - S₁ nuclease) and placed on glass fiber filters. Acid-insoluble radioactive DNA remaining on each filter after extensive washing was then determined by standard techniques.

DNA•DNA Hybridization. DNA samples to be hybridized were dialyzed into a high salt buffer (0.5 M NaCl, 10 mM potassium phosphate, pH 7.5, 1 mM Na₃EDTA). The incubation was carried out at 65° for 3 hr. The hybridized DNA samples were then diluted 10-fold and analyzed by S₁ nuclease digestion as described above.

Electron Microscopic Analysis. The DNA spreading procedure used was essentially that described by Inman (12) except that no formaldehyde was used. The reaction mixture containing the DNA to be analyzed was either diluted directly to a concentration of 1-5 µg/ml of DNA or passed through a Sepharose 4B column (0.5 × 12 cm) to remove free protein. To 5 µl of freshly prepared buffer (4 ml of H₂O, 0.3 ml of 1 M Na₂CO₃, 0.4 ml of 0.13 M Na₂EDTA, pH 10) were added successively 5 µl of diluted DNA sample, 10 μ l of formamide, and 2 μ l of 0.1% cytochrome c. The solutions were gently mixed after each addition. The high pH of this buffer and the deletion of formaldehyde are needed to eliminate 32 protein binding to single-stranded DNA. The DNA was then spread, platinum shadowed, and placed on copper grids for examination in a Philips EM-300 electron microscope.

RESULTS

The T4 replication proteins

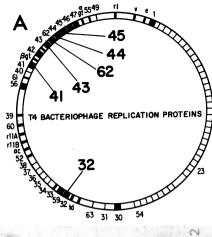
Fig. 1A presents a highly schematized version of the current T4 genetic map (13), emphasizing the six T4 replication gene products, henceforth designated as the genes 32, 41, 43, 44, 45, and 62 proteins, respectively. These components have been purified by taking advantage of the fact that each

gene product stimulates DNA synthesis in a concentrated crude extract prepared from a bacterial culture infected with a T4 bacteriophage mutant in that gene (4, 7). With this in vitro complementation assay, we have been able to purify the active form of each replication protein to near homogeneity, as analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (Fig. 1B). All of the proteins currently in use are 95–99% homogeneous by this criterion, and their mixture displays no detectable endonuclease activity at functional concentrations (as tested on either double-stranded, supercoiled PM2 DNA or fd DNA single strands).

Studies on the individual proteins, conducted in this and other laboratories, provide the basis from which mechanistic information will ultimately be drawn: (A) the 43 protein (110,000 daltons) is T4 DNA polymerase (6, 15-19). This enzyme has both a polymerase (6, 15-19) and a proof-reading exonuclease (20) activity. On a primed, single-stranded DNA template, it is capable of polymerizing at a rate of approximately 250 deoxyribonucleotides per second at 37° (ref. 21; D. Mace and B. Alberts, manuscript in preparation). (B) the 32 protein (35,000 daltons) is a DNA-unwinding protein (5, 22, 23) which is thought to aid replication by causing a local DNA helix opening just preceding the replication fork (24). The 32 and 43 proteins form a specific complex (23). Moreover, the 32 protein increases the rate at which 43 polymerase utilizes single-stranded DNA templates (23), as well as being essential for its activity on a double-stranded DNA template (25). (C) the 45 protein (27,000 dalton monomer) is isolated as a dimer: its function in DNA replication has yet to be understood. (D) the proteins 44 and 62 are isolated as a tight complex (180,000 daltons) which contains four molecules of 44 protein (34,000 daltons each) and two molecules of 62 protein (20,000 daltons each) (4). The exact function of the 44/62 complex in DNA replication is not known. However, it displays two activities which are dependent upon or are stimulated by the presence of 45 protein and rATP: a DNA-dependent hydrolysis of ATP to ADP and inorganic phosphate and a 3- to 4-fold stimulation of the 43 protein polymerization rate on a primed, single-stranded template (ref. 21; D. Mace and B. Alberts, in preparation). Note that this stimulated rate approximates the in vivo rate, estimated at about 1000 nucleotides per second at 37° (26). (E) the essential gene product most recently isolated in a biologically active form is 41 protein (58,000 daltons) (C. Morris, L. Moran, and B. Alberts, manuscript in preparation). The function of this protein is yet to be established. It is known to catalyze a single-stranded-DNA-dependent hydrolysis of GTP to GDP and inorganic phosphate (unpublished results of M. Davies, this laboratory).

The complete reaction on a single-stranded DNA template

In a number of experiments that will be found elsewhere (ref. 21; manuscript in preparation), we have determined that when the six purified proteins are added to a reaction mixture containing salts, the four deoxyribonucleoside and four ribonucleoside triphosphates, and a DNA template (either single-stranded or double-stranded) extensive DNA synthesis will occur. With a T4 single-stranded DNA template, deletion of any one of the replication proteins results in greater than a 10-fold reduction in the amount of DNA synthesis (data not shown). The deletion of a mixture of rUTP, rCTP rGTP also causes a similar 10-fold decrease. The rATP has more than one role in the stimulation of the rate of DNA synthesis, and its removal also drastically reduces synthesis



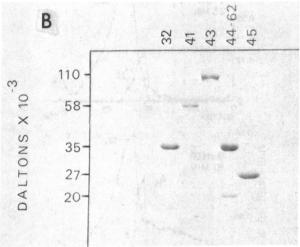


FIG. 1. (A) The genetic map of T4 bacteriophage with replication genes emphasized (1, 13). The broadest black segments indicate the relative locations of mutationally identified genes with a major effect on DNA replication at 37°. Mutants in genes 32, 41, 43, 44, 45, or 62 are unique in synthesizing little or no DNA, even though all four deoxyribonucleotide triphosphate precursors are present (14). (B) Sodium dodecyl sulfate-polyacrylamide gel electrophoresis of the purified T4 replication proteins, stained with Coomassie blue. Subsequent to the preparation of this gel, 41 protein has been purified further to about 95% homogeneity.

(ref. 21; D. Mace and B. Alberts, in preparation). Examination of the reaction products from the complete reaction by sedimentation through alkaline sucrose gradients reveals that the DNA product is not covalently attached to the DNA template (data not shown). These results suggest that RNAprimed de novo initiation of new chains might be occurring in our system. In order to carefully examine this reaction in detail with a minimum of ambiguity, the circular singlestranded DNA from fd bacteriophage was chosen as template. With this template de novo chain initiation is absolutely required, since no 3'-OH template terminus is present to fold back upon itself and act as a primer for synthesis of the complementary (-) strand. Moreover, because of the initial strand asymmetry in the template, the single-stranded regions in product produced at later times of reaction will be of the same strandedness (-) and therefore cannot renature with each other to confuse the analysis (see Fig. 2E below). Using single-stranded fd DNA as the template, the "complete" reaction can typically produce more than twice as much product DNA as original template in 10 min at 30°. whereas the deletion of any of the replication proteins 44/ 62, 43, 41, or 45 reduces the amount of DNA synthesis 100-

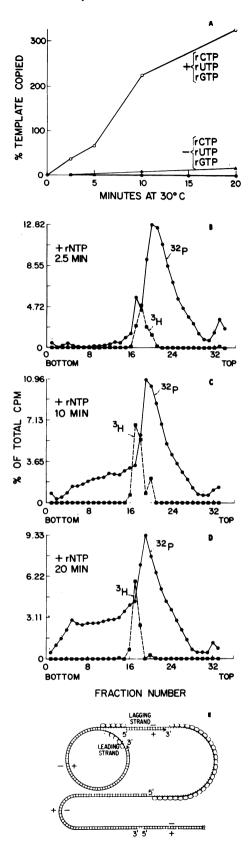


FIG. 2. (A) Time course of the complete reaction using 3 H-labeled fd single-stranded circular DNA template incubated at 30° . The complete reaction components consist of 43 protein, $2 \mu g/ml$; 41 protein, $60 \mu g/ml$; 44-62 protein, $6 \mu g/ml$; 45 protein, $10 \mu g/ml$; 32 protein, $20 \mu g/ml$; $20 \mu g/ml$; magnesium acetate, $20 \mu g/ml$; potassium acetate, $20 \mu g/ml$; $20 \mu g/ml$; $20 \mu g/ml$; magnesium acetate, $20 \mu g/ml$; $20 \mu g/ml$;

fold. As was true for the reaction on T4 DNA single strands, the deletion of the mixture of rUTP, rCTP, and rGTP caused a similar decrease in synthesis (10- to 50-fold). Since synthesis in the complete reaction continues far past the level equivalent to a complete copy of the fd template, this system must continue on after the initial template is converted to a circular double-helical DNA molecule.

Sedimentation of template and product through alkaline sucrose gradients

To enable the physical properties of template and product DNAs to be separately monitored, 3 H-labeled fd DNA was used as a template and α^{32} P-labeled dNTPs were incorporated in an experiment whose time course of DNA synthesis and rNTP requirement are shown in Fig. 2A. Template and product DNA strands were then separately analyzed at each point by sedimenting the reaction mixture through alkaline sucrose gradients, as shown in Fig. 2B-D.

Note first that the sedimentation rate of the ³H-labeled template remains unchanged throughout the incubation; therefore, these DNA strands must remain circular and intact during the entire course of the reaction (i.e., since template never cosediments with the bulk of the product DNA, these DNAs cannot be covalently attached to each other). This is the result expected if product chains start by *de novo* chain initiations.

The second major point concerns the manner in which the average strand length of the ³²P-labeled product grows with time. Whereas the product made during the first 2.5 min is the size of a linear fd genome or smaller (Fig. 2B), by 10 min a small amount of product is seen sedimenting faster than fd circular template (Fig. 2C), and by 20 min almost half the product (45% in this experiment) sediments faster than these template strands (Fig. 2D). These results are those expected if, following the ribo-dependent *de novo* initiation, the (-) strand grows to multiple fd lengths by continuous (-) strand displacement off of the circular (+) strand template. This would represent a "rolling circle" mode of DNA synthesis (27), as schematically diagrammed in Fig. 2E.

Direct S₁ nuclease digestion as a test for singlestrandedness

In order to determine the amount of single-strandedness in the template and product DNAs present during the reaction

acetate (pH 7.8), 20 mM; dithiothreitol, 0.5 mM; ³H-labeled fd single-stranded DNA, 8 μ g/ml. [α - 32 P]dGTP was used to label DNA products made. The +rUTP, +rGTP, +rCTP labeled data represent a time course of the complete reaction and the -rUTP, -rGTP, -rCTP data represent the time course observed if these components are omitted. The percent template copied equals (pmol of DNA product/pmol DNA template) × 100. (B-D) Alkaline sucrose gradient sedimentation of the reaction products. At the indicated times, aliquots were removed from the reaction mix into an equal volume of cold 20 mM Na₃EDTA. 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 hr at 50,000 rpm, in a Spinco SW 50.1 rotor (4°). Fractions were collected from the bottom onto glass fiber filters, washed, and counted by standard techniques. Counts were plotted by computer after overlap corrections. (The fraction of ³H-labeled template counts has been adjusted to 10% of their real value in order to allow the product counts to be better visualized). (E) Schematic representation of a "rolling circle" mode of DNA synthesis. The replication fork is shown with DNA-unwinding protein binding to the single-stranded DNA exposed by fork movement. The "lagging" side of the fork is necessarily discontinuously synthesized, but it appears that in vivo synthesis on the "leading" side of the fork is discontinuous as well (28, 29).

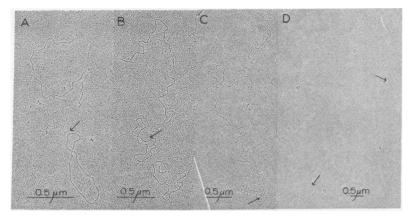


FIG. 3. Electron microscopic analysis of the fd reaction products. From the fd reaction, aliquots were removed and prepared for visualization as described in *Materials and Methods*. Arrows denote origins of rolling circle tails. (A) A double-stranded fd molecule with a rolling circle tail. This tail is single-stranded at the circle-tail junction and again near the top of micrograph. (B) Above a partially duplicated fd circle is another rolling circle with tail. The tail is single-stranded both at the circle-tail junction and at a position midway down the tail. (C) A double-stranded circle with a long double-stranded tail; note again the single-stranded region which connects the circle and tail. (D) Two double-stranded circles with long double-stranded tails (see arrows).

in Fig. 2, aliquots at various times were taken to pH 4.6, 0.1% sodium dodecyl sulfate, and digested with S₁ nuclease (see Materials and Methods). The results indicate that the percent of S₁-resistant (and therefore double-stranded) DNA changes as a function of DNA synthesis time. Whereas only 35% of ³H-labeled template is resistant to S₁ digestion at 2.5 min, by 5 min 69% is resistant, and nearly all of the template has become double-stranded (100% S₁ resistant) by 10 min of synthesis. In contrast, whereas the initial ³²P-labeled product made at 2.5 min is almost completely doublestranded (92% S₁ resistant), at later times more product becomes single-stranded, so that by 5 min an apparent limit has been reached where the S₁-sensitive fraction remains approximately constant at 20% of the total. These results are explained if a complete (-) strand is first made on each (+) strand circle, and then some of this product DNA is displaced as single strands into a rolling circle tail, where it efficiently serves as a template for discontinuous synthesis of its (+) strand complement. If this analysis is correct, one expects a typical replication fork at the junction of circle and tail with both "leading" and "lagging" strand synthesis (Fig. 2E).

DNA hybridization studies

If the "rolling circle" scheme in Fig. 2E is actually being carried out by our system, at late times in the course of the reaction the largest single-stranded product DNA (multiple fd genome lengths) should be predominantly (—) strand and the smallest product (less than one fd genome length) should be predominantly (+) strand.

In order to test this prediction, the reaction products after 15 min of incubation were sedimented through a neutral sucrose gradient (data not shown) to isolate the largest (most rapidly sedimenting) rolling circles. The molecules were then denatured and sedimented through an alkaline sucrose gradient. Fractions across this alkaline gradient were pooled into four size classes and dialyzed into a hybridization buffer. Aliquots of these pooled fractions were then either self-hybridized, or hybridized with an excess of fd viral DNA [(+) strand]. S₁ nuclease digestion was used to determine how much double-stranded DNA had formed. The results were that more than 90% of the single-stranded DNA larger than 12 × 10⁶ daltons is (-) strand, while 75% of the single-

stranded DNA smaller than 2×10^6 daltons is (+) strand, consistent with our expectation.

Electron microscopy of the reaction products

Examination of the reaction products in the electron microscope shows that 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, as shown in Fig. 3. 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 examples shown in Fig. 3A and B. This is again consistent with the view that the (+) strand complement made on the singlestranded tail is synthesized by a discontinuous mechanism. Structures similar to those shown in Fig. 3 are seen whether fd or G4 bacteriophage DNAs (single-stranded), or PM2 bacteriophage or simian virus 40 DNAs (double-stranded) are used as templates (ref. 21; N. Sinha, C. Morris, and B. Alberts, in preparation).

DISCUSSION

In this communication, we have presented evidence that the six proteins from T4 bacteriophage essential for DNA replication fork movement can function together *in vitro* to mimic closely the characteristics of *in vivo* DNA replication. In a subsequent communication we will present evidence that the DNA product is made at near *in vivo* rates (800 nucleotides/second per replication fork) and that it is a faithful copy of its template, lacking the reversibly denaturable hairpin structure generated by other purified systems that can copy double-helical DNAs (25, 30, 31).

Use of a circular single-stranded DNA molecule as a template, suggested by work on host systems (32), is advantageous here. Since the template has no 3'-OH terminus, the only DNA synthesis possible must arise from de novo initiation. When the complementary (—) strand has been completed on the circular template to form a double-stranded circular molecule, synthesis can continue only if the newly synthesized product strand is displaced. This occurs, establishing the rolling circle with a single-stranded tail. De novo initiations occur on the displaced strand leading to a long, predominantly double-stranded tail.

It is important to note that initiation of DNA replication in vivo is likely to involve additional replication machinery besides that needed for fork movement. After a T4 bacteriophage DNA molecule enters an Escherichia 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 seems to remain true for T4 mutant infection in which this host DNA is not significantly degraded (33). It therefore seems likely that special DNA sequences or structures 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 on a wide variety of double- and single-stranded DNA templates (manuscript in preparation). In the case of the fd DNA template it is likely that the fork with both "leading" and "lagging" strand synthesis arises from strand displacement at a nick or small gap. While replication fork initiation by this mechanism may be strongly suppressed in vivo (possibly by special blocking proteins), it conveniently allows us to bypass the normal T4 fork initiation system and enables the mechanism of fork propagation to be examined in a relatively simple in vitro system.

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). N.S. and C.M. were supported by postdoctoral fellowships from the National Cancer Institute (Fellowship nos. 5F22CA01829 and 6F22CA03266, respectively).

- 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. & Lielausis, A. (1963) Cold Spring Harbor Symp. Quant. Biol. 28, 375-394.
- 2. Warner, H. R. & Hobbs, M. D. (1967) Virology 33, 376-384.
- Riva, S., Cascino, A. & Geiduschek, E. P. (1970) J. Mol. Biol. 54, 85-102.
- Barry, J. & Alberts, B. (1972) Proc. Nat. Acad. Sci. USA 69, 2717-2721.
- 5. Alberts, B. & Frey, L. (1970) Nature 227, 1313-1318.
- Nossal, N. G. & Hershfield, M. S. (1971) J. Biol. Chem. 246, 5414-5426.
- Barry, J., Hama-Inaba, H., Moran, L., Wiberg, J. & Alberts, B. (1973) DNA Synthesis In Vitro, eds. Wells, R. D. & Inman, R. B. (Baltimore, Md.).

- Yamamoto, K., Alberts, B., Benzinger, R., Lawhorne, L. & Treiber, G. (1970) Virology 40, 734-744.
- 9. Laemmli, U. K. (1970) Nature 227, 680-685.
- 10. Studier, F. (1973) J. Mol. Biol. 79, 227-236.
- Vogt, V. (1973) Eur. J. Biochem. 33, 192-200.
- Inman, R. B. (1974) in Methods in Enzymology, eds. Grossman, L. & Moldave, K. (Academic Press, New York), Vol. 29, pp. 451-458.
- Wood, W. B. (1974) in Handbook of Genetics, ed. King, R. C. (Plenum Publishing, New York).
- 14. Mathews, C. (1972) J. Biol. Chem. 247, 7430-7438.
- DeWaard, A., Paul, A. V. & Lehman, I. R. (1965) Proc. Nat. Acad. Sci. USA 54, 1241-1248.
- 16. Warner, H. R. & Barnes, J. (1966) Virology 28, 100-107.
- 7. Englund, P. T. (1971) J. Biol. Chem. 246, 5684-5687.
- Goulian, M., Lucas, Z. J. & Kornberg, A. (1968) J. Biol. Chem. 243, 627-638.
- Lehman, I. R. (1974) in Methods in Enzymology, eds. Grossman, L. & Moldave, K. (Academic Press, New York), Vol. 29, pp. 46-53.
- Brutlag, D. & Kornberg, A. (1972) J. Biol. Chem. 247, 241– 248
- Alberts, B., Morris, C., Mace, D., Sinha, N., Bittner, M. & Moran, L. (1975) DNA Synthesis and Its Regulation: ICN-UCLA Symposia on Molecular and Cellular Biology, eds. Goulian, M., Hanawalt, P. & Fox, C. F. (W. A. Benjamin, Inc., Menlo Park, Calif.), in press.
- Delius, H., Mantell, N. & Alberts, B. (1972) J. Mol. Biol. 67, 341-350.
- Huberman, J., Kornberg, A. & Alberts, B. (1971) J. Mol. Biol. 62, 39-52.
- Alberts, B. (1973) in Molecular Cytogenetics, eds. Hamkalo,
 B. & Papaconstantinou, J. (Plenum, New York).
- 25. Nossal, N. G. (1974) J. Biol. Chem. 249, 5668-5676.
- Werner, R. (1968) Cold Spring Harbor Symp. Quant. Biol. 33, 501-503.
- Gilbert, W., and Dressler, D. (1968) Cold Spring Harbor Symp. Quant. Biol. 33, 473-484.
- Sugimoto, K., Okazaki, T. & Okazaki, R. (1968) Proc. Nat. Acad. Sci. USA 60, 1356-1362.
- Sternglanz, R. (1975) DNA Synthesis and Its Regulation: ICN-UCLA Symposia on Molecular and Cellular Biology, eds. Goulain, M., Hanawalt, P. & Fox, C. F. (W. A. Benjamin, Inc., Menlo Park, Calif.), in press.
- Inman, R. B., Schildkraut, C. L. & Kornberg, A. (1965) J. Mol. Biol. 11, 285-292.
- Masamune, Y. & Richardson, C. C. (1971) J. Mol. Biol. 246, 2692-2701.
- Schekman, R., Weiner, J., & Kornberg, A. (1974) Science 186, 987–993
- 33. Snustad, P. & Conroy, L. (1974) J. Mol. Biol. 89, 663-673.