UCSF
UC San Francisco Electronic Theses and Dissertations

Title
Studies on the properties of bacteriophage T4 DNA replication complexes

Permalink
https://escholarship.org/uc/item/0p522334

Author
Bedinger, Patricia Anne

Publication Date
1982

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Studies on the Properties of Bacteriophage T4 DNA Replication Complexes

by

Patricia Anne Bedinger

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

[redacted]

Date University Librarian

Degree Conferred: SEP 8 1982

STUDIES ON THE PROPERTIES OF BACTERIOPHAGE T4 DNA REPLICATION COMPLEXES

DEDICATION

At the risk of being prosaic, I would like to dedicate this thesis to my parents, Jon and Bette Bedinger, in thanks for their love and support for these many years.

ACKNOWLEDGMENTS

The acknowledgment section is the greatest pleasure to write in a thesis for it gives the author an opportunity to thank the people who have provided the much-needed professional and personal support in the arduous struggle through graduate school.

First and foremost, I wish to acknowledge my thesis advisor, Bruce Alberts. Bruce's incredibly creative and clear thinking make him a constant source of intellectual inspiration to the members of his laboratory and to the scientific community in general. His generosity, warmth and integrity make him a personal inspiration. I have especially appreciated Bruce's ability to consistently pinpoint the essential strengths and weaknesses of an experiment or a piece of writing, and his ability to salvage meaningful information from seemingly hopeless morass of data. It has been a great honor to be Bruce's student, and I hope that in some way I can effectively emulate him throughout my scientific career.

I would also like to acknowledge the members of Bruce's laboratory who have contributed to this work and my education throughout the years with provocative discussions and helpful criticism. I especially appreciate the support and example of the other women in the laboratory.

I wish to acknowledge the support of my friends and family; especially that of Ken, who provided a loving, peaceful haven.

Lastly, I wish to acknowledge Roget and his Thesaurus, for loosening the strait jacket of my scientific vocabulary.

TABLE OF CONTENTS

Abstract

The replication of DNA is accomplished by multienzyme complexes. An *in vitro* DNA replication system composed of highly purified bacteriophage T4 proteins, shown to be an accurate model of *in vivo* replication in many respects, has been developed in the laboratory of Dr. Bruce Alberts. This thesis presents studies of *in vitro* replication reactions using the complete and partial (missing one or more protein components) T4 enzyme systems with a variety of defined DNA templates in order to probe protein-protein and protein-DNA interactions.

The first chapter of this thesis contains a description of experiments that show a marked stimulation of the 3'-5' proofreading exonuclease activity of the T4 DNA polymerase by other T4 replication proteins. This work suggests a possible mechanism by which proteins in a complex may enhance the fidelity of replication.

Experiments demonstrating the pausing of the *in vitro* replication fork at specific DNA sequences on a double-stranded DNA template are described in the second chapter of this thesis. The strongest pause sites are DNA sequences that, if single-stranded, could form "hairpin" secondary structures, which are known to block replication on a single-stranded template. These results suggest that the multienzyme replication fork creates a single-stranded region ahead of the actual site of polymerization.

Experiments establishing that the movement of the *in vitro* replication fork can be blocked by purified *E. coli* RNA polymerase bound to DNA at transcription promoter sequences are reported in the final

chapter. The addition of T4 dda protein, a DNA helicase, allows the fork to progress through the sites where RNA polymerase is bound. These studies indicate that DNA helicases in a replication fork could serve the important function of removing protein barriers on a DNA template.

The experiments described in this thesis affirm the remarkable complexity of DNA replication in terms of the interaction of replication proteins with each other, with the DNA template, and with proteins bound to the DNA template. The intricate molecular mechanisms involved in DNA replication, a process of primary importance in the study of biology, are becoming amenable to scrutiny in the laboratory with the development of *in vitro* systems. Such studies should continue to be extremely productive in the foreseeable future.

1. BACKGROUND

1.1. DNA replication by multienzyme complexes

The replication of nucleic acid genetic material is theoretically a process that should be a relatively simple one. The specificity of the system is inherent, given the double helical base-paired structure of DNA, elucidated thirty years ago (Watson and Crick,1953). Each of the two parental strands serves as a template which is "matched" within the built-in constraints of normal nucleotide base-pairing through hydrogen bond formation; adenine pairs only with thymine and guanine with cytosine. One could postulate that DNA elongation could be accomplished with a minimum of two enzymatic activities; one to "unzip" the double helix in order to expose single-stranded parental templates and a polymerizing activity to covalently join the incoming nucleotides together through a phosphodiester linkage, forming the semiconserved daughter double helix.

This kind of elementary mechanism for DNA replication is, of course, naively simplistic. Replication is performed *in vivo* by seemingly baroque multiprotein replication complexes, or "replisomes" (Alberts and Sternglanz, 1977). Multienzyme complexes seem to be the rule rather than the exception for most important biological processes in prokaryotes and eukaryotes. Well known examples include glycolysis (pyruvate dehydrogenase complex), ATP production in mitochondria (F1 ATPase),fatty acid synthesis, RNA synthesis, and protein synthesis. In general the enzyme complexes involved in the above reactions are more stable than DNA replication complexes, presenting a mixed blessing to the DNA researcher; while complete replication complexes are

difficult to isolate in entirety, it is relatively simple to examine the role of individual components by dissecting the system into simpler subsets. Such studies have provided valuable clues about the complexity of DNA synthesis. Table I contrasts the functions of T4 DNA polymerase alone with those of the multiprotein T4 DNA replication apparatus, the components of which are described in detail below.

It is readily apparent from Table I that replication proteins are crucial in the establishment and maintenance of the special structure of a replication fork. When associated with the central enzyme of DNA replication- DNA polymerase-these proteins increase the rate, processivity, and fidelity of replication.

Generally, DNA replication can be divided into the distinct processes of replication fork initiation and replication fork movement. Initiation requires the recognition of specific DNA sequences known as replication origins, where the replication fork is first established. Elongation is the process by which an established fork is propagated along the template DNA. Elongation can be further divided into "leading" strand elongation (continuous 5'-3' synthesis), and "lagging" strand elongation,a process which invokes a special discontinuous mechanism to allow net 3'-5' synthesis (all known DNA polymerases synthesize only in a 5'-3' direction). This thesis focuses on the most "elementary" of the replication processes; that of the elongation phase of leading strand DNA synthesis.

The development of <u>in vitro</u> DNA replication systems has greatly expedited our understanding of the mechanisms involved in replication. The "divide and conquer" tactic of constructing partial replication

TABLE I

T4 DNA Polymerase	T4 Replication Apparatus
-largely distributive mechanism of polymerization at physiological salt concentrations (D. Mace, thesis, Newport, et al., 1980)	-highly processive mechanism of polymerization, relatively salt insensitive
-long pauses (as long as 24 min) at sites of secondary structure on a single-stranded DNA template (Huang et al., 1980)	-progresses efficiently through regions of secondary structure with no detectable pausing
-requires a primed ssDNA template; cannot use double-stranded DNA as a template	-replicates double-stranded DNA by a strand displacement mechanism (Nossal and Peterlin, 1979, Sinha, et al., 1980)
	-rates of synthesis on a dsDNA template are close to in vivo
-reduced base pairing fidelity from in vivo levels	-fidelity of the same order of magnitude as in vivo fidelity (Hibner and Alberts, 1980)
-cannot start new DNA chains de novo	-starts DNA chains de novo after synthesis of a pentaribonucleotide RNA primer (Liu and Alberts, 1980, Liu and Alberts, 1981)

reactions in which purified DNA polymerase and various subsets of other replication proteins are combined with appropriately defined DNA substrates enables the researcher to investigate the roles of the individual protein components of the larger more intricate replication ensemble in a readily interpretable way.

There are several types of proteins that participate in the construction of a replication fork. The core enzyme is, of course, DNA polymerase. All prokaryotic polymerases contain a 3'-5' exonuclease activity in addition to the polymerizing function, which acts in a proofreading mode during replication by preferentially removing mismatched nucleotides at a primer end prior to further extension of the growing DNA chain (Brutlag and Kornberg, 1972, Bessman et al., 1974).

Another generic replication protein is the Helix Destabilizing protein (HD protein), sometimes known as Single-Strand Binding (SSB) protein. Proteins of this type bind tightly and cooperatively to single-stranded DNA and hold the DNA strand in an extended conformation with its bases fully exposed. They thereby produce a DNA structure essential for replication and recombination. In addition, the HD proteins have been shown to physically associate with several of the proteins involved with DNA replication, recombination, and repair (Formosa, et al., manuscript in preparation).

Numerous other proteins have been designated as replication fork components either by genetic studies of appropriate mutants or by isolation as a requisite factor for an <u>in vitro</u> replication activity. Many of these proteins are nucleoside triphosphatases; the energy

derived from the hydrolysis of high energy phosphodiester bonds is utilized toward a variety of different ends. DNA helicases, for example, are thought to use the energy from nucleotide triphosphate hydrolysis to do mechanical work; in a manner analogous to the movement of myosin along actin, they move along a DNA single strand to accomplish unidirectional melting of the duplex DNA ahead of the fork (Yarronton and Gefter, 1979, Abdel-Monem and Hoffman-Berling, 1980). Nucleoside triphosphate hydrolysis can also serve in assembly reactions, triggering the formation of functional protein-protein and protein-DNA complexes (Wickner, 1978, and below). In addition, models have been proposed wherein the hydrolysis of nucleoside triphosphates provides a timing mechanism to enhance the fidelity of replication (Hopfield, 1974, Ninio, 1975).

In the remainder of this introductory chapter I will describe the bacteriophage T4 *in vitro* DNA replication system developed in the laboratory of Dr. Bruce Alberts in some detail. Subsequently I will discuss the *in vitro* replication systems derived from proteins of bacteriophage T7 and *E. coli*, drawing analogies with the T4 system where relevant. This chapter then ends with a discussion of the current state of eukaryotic *in vitro* replication systems.

1.2. The T4 replication system

The successful development of the T4 replication system relied on earlier genetic studies that have designated each of the T4 gene products indicated in Table II as essential for the replication fork (Epstein et al., 1963, Warner and Hotts, 1967, Riva et al., 1970). The bacteriophage T4 *in vitro* system presently consists of the seven

highly purified proteins described in Table II, each of which is more than 90% pure and available in milligram amounts. The availability of mutations in the T4 genes coding for these proteins allowed an *in vitro* complementation assay to be developed that was used in the purification to homogeneity of the T4 gene 44/62, 45, and 41 proteins, without knowledge of their true enzymatic activities (Barry and Alberts, 1972, Morris, et al., 1979a, Morris et al., 1979b). In this complementation assay, "receptor cell lysates" are made by gently lysing concentrated *E.coli* cells that have been infected with a T4 phage deficient in the T4 replication protein to be purifed. Such a lysate is deficient in incorporating radioactive precursor nucleotide into a acid-precipitable (DNA) form, compared to a wild-type lysate, unless the missing gene product is supplied exogenously (either in the form of a crude lysate, a column fraction, or a purified protein). Using this complementation assay and standard enzymatic activity assays, all seven of the T4 proteins described in Table II had been purified to >90% homogeneity by the time that I began my studies.

The extent to which the individual proteins of the T4 replication system have been characterized in terms of structure and function varies; the T4 DNA polymerase and helix destabilizing protein (gene 32 protein) are the two most thoroughly studied of the seven proteins.

The T4 DNA polymerase is the product of T4 gene 43, and has a molecular weight of 110,000 daltons. It is an extremely complex enzyme, containing binding sites for DNA, deoxyribonucleoside triphosphates, 3'OH ends of DNA primers, and sites for interacting with several proteins, including 32 protein (Huberman et al., 1971), 45

TABLE II The Proteins of the T4 Replication System

Gene	activities	optimal conc. (μg/ml)	% purity	MW (kilodaltons)
43	5'-3' polymerase 3'-5' exonuclease	2	99	110
32	helix destabilizing	100	99	33.5
44/62	polymerase accessory proteins DNA-dep ATPase	20	99	4 X 34 2 X 20
45	polymerase accessory protein greatly stimulates 44/62 protein	10	95	2 X 24.5
41	DNA-dep GTPase helicase? mobile promoter?	6	90	2 X 58
61	primase?	0.1	90	44?
dda	DNA-dep ATPase 3'-5' helicase	5	60-70 (major contaminant is 32 protein)	48

protein, and (probably) itself (T. Formosa, unpublished observations). Like other prokaryotic DNA polymerases, this protein contains a 3'-5' exonuclease activity in addition to its polymerizing function. Genetic studies have shown that the DNA polymerase plays an important role in determining mutation frequencies in T4 (Drake, 1973). Both mutator and antimutator phenotypes are attributed to different 43 mutants, which in some cases can be traced to a decrease (mutator) or increase (antimutator) in the level of the exonuclease activity relative to polymerase activity of the mutant polymerase proteins (Lo et al., 1976, Bessman et al., 1974). These results are consistent with the idea that the 3'-5' exonuclease functions to edit out mistakes made during replication. This exonuclease activity of the T4 DNA polymerase and the influence of other T4 replication proteins on it will be discussed in detail in this thesis.

The T4 helix destabilizing protein, the product of T4 gene 32, was first described twelve years ago (Alberts and Frey, 1970). Genetic studies have revealed that this protein is essential for T4 DNA replication, DNA repair, and genetic recombination (Mosig et al., 1978). The 32 protein has a monomer molecular weight of 34,500 daltons and , like all helix destabilizing proteins, binds strongly and cooperatively to single-stranded DNA. The function of 32 protein is considerably more subtle than simply that of protecting single-stranded DNA from nuclease attack, or even that of providing a suitable protein-coated single-stranded DNA substrate for DNA metabolic processes. In addition, 32 protein seems to mediate the assembly of groups of enzymes at sites of replication or recombination through

specific protein: protein interactions. Specific binding of T4 genes 45 and 61 proteins, T4 DNA polymerase, and at least eight other proteins to 32 protein has been demonstrated (Burke, et al., 1980, Formosa, et al., manuscript in preparation). Most of these protein interactions are eliminated when the cartoxyl terminal region (8000 daltons) of 32 protein has been removed by proteolytic digestion. The remaining large fragment of 32 protein, known as 32*I protein, retains its DNA binding properties (Greve et al.,1978). Thus, intact 32 protein can tentatively be schematically divided into two functional domains- the amino terminal region which binds DNA in part through electrostatic interactions, and the cartoxyl terminal region that binds specifically to proteins involved in DNA metabolism (Hosoda et al., 1980, Newport, et al., 1980).

The functions of the other T4 proteins involved in DNA synthesis on the leading strand- the 45,44/62, and 41 proteins- are less well understood. The above described in vitro complementation assay allowed for their purification prior to assignment of a biological activity, and their mode of action has been inferred from their affects on a variety of partial reactions carried out with the purified proteins.

The 44/62 and 45 proteins have been classified as "polymerase accessory proteins". The 44/62 protein is a tight complex with a molecular weight of 180,000 daltons containingthe products of T4 genes 44 (34,000 daltons) and 62 (20,000 daltons) in a ratio of either 4:2 or 5:1. The 45 protein functions as a dimer with a molecular weight of 54,400 daltons. Together these three proteins constitute a DNA-dependent ATPase; the 44/62 protein, specifically the 44 subunit (data

of Maureen Munn), contains the nucleoside triphosphatase activity which is greatly stimulated by the 45 protein and DNA. The polymerase accessory proteins increase the processivity of the T4 DNA polymerase, probably by acting as a "sliding clamp" that strengthens the binding of the polymerase to the 3'OH primer terminus (D. Mace, thesis, Newport, et al., 1980, Huang et al., 1981). It appears that ATP hydrolysis is required for the formation of this accessory protein: polymerase complex at DNA ends.

The 45 protein is remarkable in that it, like the 32 protein, is involved in several different DNA-related processes in the infected cell; namely replication, DNA repair, and the switch from early to late transcription. The 45 protein has been shown to physically interact with both 32 protein and with E. coli RNA polymerase (Formosa et al., manuscript in preparation, Ratner, 1974).

The product of T4 gene 41, also purified using the in vitro complementation assay, plays a role in both leading and lagging strand DNA synthesis. This enzyme is a DNA-dependent GTPase, with a monomer molecular weight of 58,000 daltons; it apparently functions as a dimer (Liu and Alberts, 1981). The addition of 41 protein to DNA synthesis reactions on a double-stranded DNA template causes a dramatic increase in the rate of fork movement, in a reaction requiring its GTP hydrolysis activity (Liu et al., 1978, Alberts et al., 1980, Barry and Alberts, manuscript in preparation). The 41 protein is also required for lagging strand DNA synthesis, acting in conjunction with gene 61 protein to synthesize RNA primers at specific sites on the single-stranded DNA displaced during leading strand synthesis. In the

current view of 41 protein action, this protein uses GTP hydrolysis energy to move along the lagging strand, where it directs the 61 protein to correct priming sites; in addition, it acts as a DNA helicase at the replication fork using energy derived from its GTP hydrolysis to push open the double helix ahead of the fork.

Recently the T4 dda protein, a DNA-dependent ATPase and 3'-5' DNA helicase has been purified in our laboratory (Jongeneel and Alberts, manuscript in preparation). Its role in the replication system is unclear. Some data on the effects of the T4 dda protein in some *in vitro* replication reactions will be reported in this thesis.

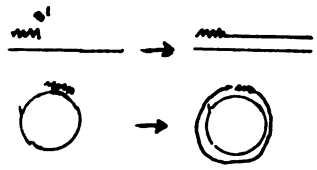
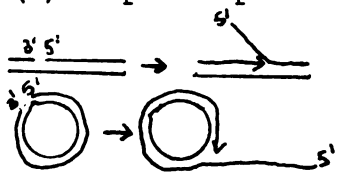
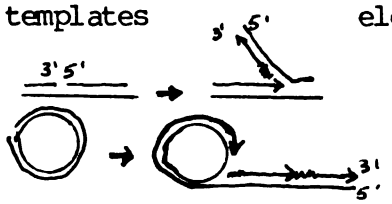
1.3. Partial reactions carried out by the T4 DNA replication system

The goal of studies utilizing *in vitro* systems is to achieve a detailed understanding of the process of DNA replication by reconstituting an activity from purified components that accurately mimics *in vivo* replication. The elongation of T4 DNA *in vivo* takes place on double-stranded DNA, and is continuous on the leading strand and discontinuous on the lagging strand, where *de novo* synthesized RNA primers are made to start Okazaki fragments. The result, after RNA primer removal and DNA ligase sealing of Okazaki fragments on the lagging strand, is two semiconserved daughter DNA molecules. The rate of fork movement is very rapid (~5000 nucleotides/second) and the fidelity of copying very high- about one error made per 10^7 to 10^9 bases incorporated (McCarthy et al., 1976, Drake, 1969). *In vitro* replication by the T4 seven protein system closely simulates *in vivo* replication in all of the above respects (Liu et al., 1978, Alberts et al., 1980).

By utilizing "partial" reactions, in which one or several of the seven T4 replication proteins is omitted, it is possible to evaluate the contribution of the individual proteins of the system with regard to different parameters such as processivity and rate of synthesis. Table III lists various partial reactions of the T4 system according to protein components, DNA substrates, and characteristics of synthesis. It can be seen from these partial reactions that the T4 replication proteins invest the T4 DNA polymerase with far more flexibility in terms of its "active" DNA templates, and, further, increase the speed, processivity and accuracy with which the polymerase synthesizes DNA. The enhancement of replication by these proteins can be attributed to both protein:protein and protein:DNA interactions, as will be dicussed in greater detail in this thesis.

It is worth noting that these partial reactions may be of some relevance in reactions involving DNA *in vivo*. For example, the T4 DNA radiation repair pathway requires some, but not all, of the replication proteins (Maynard-Smith and Symonds, 1973). Genetic studies have also shown that replication proteins are neccessary for recombination (Mosig et al., 1978). One model for T4 replication initiation *in vivo* proposes a limited amount of replication that is then extended in a second stage of replication, somewhat analogous to the two-stage replication of the plasmid ColE1. For ColE1, a small region of DNA is replicated by *E.coli* polymerase I, and then elongated by the more complex DNA polymerase III in conjunction with *E. coli* SSB protein (Tomizawa, 1978). Partial replication complexes may have an important function in a variety of DNA metabolic reactions.

TABLE III

Proteins	DNA Templates	Characterstics
a. DNA polymerase only		requires a pre-existing 3'OH primer end, copies ssDNA only, with long pauses at sites of secondary structure, dissociates frequently from 3'OH end
b. DNA polymerase plus T4 gene 32 protein	(a) templates	stimulates polymerase movement on ssDNA in part by helping to remove helical hairpins, effects on template base stacking increases fidelity of copying (Topal and Sinha, 1980
c. DNA polymerase plus T4 genes 44,62 and 45 proteins	(a) templates	stimulates polymerase movement on ssDNA by creating a "sliding clamp" that stabilizes the polymerase:3'OH primer end interaction in an ATP requiring step
d. DNA polymerase plus 32, 45 and 44/62 proteins (the five-protein reaction)	(a) templates plus 	synthesis on ssDNA becomes rapid and extremely processive, synthesis can proceed from nicks on a dsDNA template by strand displacement (requires ATP hydrolysis by 44/62 protein
e. DNA polymerase plus 32, 45, 44/62, and 41 proteins	(a), (d) templates	rate of DNA polymerase movement on dsDNA template increases 6- to 8-fold (GTP hydrolysis by the 41 protein is required)
f. 41 and 61 proteins (RNA priming reaction)		synthesis of RNA pentamers of sequence ACCXX; these can act as primers for the initiation of new DNA chains
g. DNA polymerase plus 32, 45, 44/62, 41, and 61 proteins (the seven protein reaction)	(a), (d), and (f) templates 	accurately mimics in vivo elongation of DNA

1.4. Other well-characterized *in vitro* replication systems

1.4.1. Prokaryotic systems

Two other prokaryotic *in vitro* DNA replication systems have been highly developed and analyzed in detail; that of bacteriophage T7 and that of the bacterium *E. coli*, the host cell for both bacteriophages T4 and T7.

Bacteriophage T7 has a double-stranded DNA genome of about 40,000 base pairs, a little less than one fourth of the size of the T4 genome. The entire genome has been sequenced, and genetic studies identifying the genes involved in replication are extensive. From *in vitro* studies it has been suggested that both leading and lagging strand DNA synthesis are accomplished by a very streamlined system of only two enzymes, T7 DNA polymerase and T7 gene 4 protein (Tamanoi, et al., 1980).

The T7 DNA polymerase is composed of two subunits; T7 gene 5 protein with a molecular weight of 84,000 daltons and, surprisingly, host thioredoxin, with a molecular weight of 12,000 daltons. Neither subunit alone exhibits any polymerase activity, but when complexed together the resultant enzyme contains both a 5'-3' polymerizing activity and a 3'-5' exonuclease activity. Recently it has been found that the T7 DNA polymerase can be isolated in two active forms, depending on the method of preparation (Fischer and Hinkle, 1980). Although the exact nature of the difference between the two forms is not understood, it is clear that they differ greatly with respect to their interaction with the second T7 replication protein, the gene 4

protein.

The T7 gene 4 protein was first isolated using the same type of *in vitro* complementation assay described above for T4 proteins, has been shown to contain both a DNA helicase activity and an RNA primase activity (Richardson et al., 1978). Given a special "frayed end" DNA entry site, the gene 4 protein helicase can melt duplex DNA in a 5'-3' direction utilizing energy from nucleoside triphosphate hydrolysis. One form of the T7 DNA polymerase can provide such an entry site for the helicase by limted strand displacement synthesis at a nick in duplex DNA. The helicase can subsequently act in conjunction with the polymerase to efficiently replicate a double-stranded DNA template. attaining a synthesis rate of about 250 bases/ second, close to that of the *in vivo* rate (R. Leschner, personal communication).

Studies of the replication system of *E. coli* (which has a genome length of 4 million base pairs) have revealed a far more complex picture of replication (Kornberg, 1980). At least 20 different proteins are involved in leading and lagging strand DNA synthesis. It appears that much of the complexity of the *E. coli* system arises from the mechanism of RNA primer synthesis on the lagging strand, which requires the formation of an intricate multienzyme "primasome". Leading strand elongation can be accomplished by three groups of proteins: a large multiprotein polymerase III holoenzyme, SSB protein (*E. coli* HD protein), and a DNA helicase.

The *E. coli* polymerase III holoenzyme is composed of seven polypeptides, as described in Table IV. It is possible to isolate forms of this enzymes lacking one or more components, thereby allowing

TABLE IV **Components of DNA polymerase III Holoenzyme**

Subunit	MW (kdal)	Subsets
α	140	pol III core, pol III', pol III*
ε	25	pol III core, pol III', pol III*
θ	10	pol III core, pol III', pol III*
τ	83	pol III', pol III*
γ	52	pol III', pol III*
δ	32	pol III', pol III*
β	37	pol III*

the characterization of partial reactions such as those described for T4 in order to investigate the role of each factor in replication. Recent studies with these incomplete complexes have shown that the core polymerase III- containing only subunits alpha, epsilon, and sigma- is far less processive than the complete holoenzyme; the former adds only 10-15 nucleotides before dissociating from the primer end, while the holoenzyme can add thousands of bases before dissociating (Fay, et al., 1981). Reactions with polymerase III core plus the tau subunit, or plus both delta and gamma subunits,show an increase in the processivity of the polymerase from 10-15 bases to 30-40 bases per primer end association. However, the most dramatic increase in processivity is seen when the beta subunit is added to polymerase III*, which is defined as the enzyme containing all of the holoenzyme subunits except for beta (Fay, et al., 1982). This final form of the complex remains tightly bound to the template DNA even after the addition of thousands of bases. The gamma and delta subunits act to assemble the beta subunit into the holoenzyme:DNA complex in an ATP-utilizing reaction.

While the polymerase III holoenzyme alone can efficiently replicate a primed single-stranded DNA substrate, the addition of E. coli SSB protein and a DNA helicase are necessary to achieve replication by strand displacement on a double-stranded DNA substrate. Currently the E. coli rep protein is the helicase used in in vitro replication, but since this enzyme is not essential for in vivo E. coli chromosome replication, there is must be another DNA helicase more commonly used at E. coli replication forks; possibly the aforementioned primasome

complex contains such a DNA helicase activity (similar to that of T4 gene 41 protein and T7 gene 4 protein). Table V summarizes the activities of the leading strand replication proteins of T7, T4, and E. coli , arranging them into suggested comparable groups according to function.

1.4.2. Eukaryotic in vitro DNA replication systems

In general the development of in vitro DNA replication systems derived from eukaryotic cells has lagged far behind that of prokaryotic systems. Starting more than twenty years ago, attention has been focused on the purification and characterization of alpha DNA polymerases from various sources (Chang, 1980).

The alpha DNA polymerases have been designated as the polymerases involved in nuclear DNA replication by both physiological and drug inhibition studies. In contrast to the beta and gamma DNA polymerases also found in eukaryotes, the level of alpha polymerase activity increases as the rate of DNA synthesis increases during S phase of the cell cycle (Falaschi and Spadari, 1978). Furthermore, aphidocolin, a drug that specifically inhibits the alpha polymerase, can block cells in the G1 phase of the cell cycle, preventing the onset of DNA replication (Ikegami, et al., 1978).

Several laboratories have purified alpha polymerase from different sources to suspected homogeneity and have found that the enzyme is composed of several subunits, much like the E. coli DNA polymerase III holoenzyme. For example, the alpha polymerase from Drosophila embryos is composed of at least four distinct polypeptides. The DNA

TABLE V Comparison of Prokaryotic In Vitro Systems

Activity	T7	T4	E. coli
DNA polymerase	T7 gene 5 protein plus host thioredoxin	T4 gene 43	DNA polymerase III (dna E gene) plus holoenzyme subunits and
Helix destabilizing protein	Either T7 gene 2.5 or E. coli SSB protein	T4 gene 32 protein	E. coli SSB protein
DNA helicase	T7 gene 4 protein	T4 genes dda and (probably) 41 proteins	E. coli rep protein and/or others (helicase II? primasome?)
Formation of tight complex tying polymerase to 3'OH DNA end	None?	T4 genes 45 and 44/62 proteins	holoenzyme subunits

polymerizing activity of the Drosophila alpha polymerase has been assigned to the largest (148,000 dalton) subunit by two types of experiments. The first uses glcerol gradient sedimentation in the presence of urea to separate the large subunit of the polymerase from the other subunits; a small but significant amount of polymerization activity is detected only in the gradient fractions containing the large subunit (Villani et al.,1980). The second type of experiment demonstrates that the large subunit is the catalytic one by the use of an electrophoretic technique (Spanos et al., 1981). Purified alpha DNA polymerase is electrophoresed into an SDS- polyacrylamide gel that was impregnated with substrate DNA. Subsequent incubation of the gel with radioactively labeled deoxyribonucleotide triphosphate precursors, washing out of free nucleotide, and autoradiography, indicates that the large subunit was the only one with the capacity to incorporate nucleotides into DNA.

Now that the central polymerizing activity of the large subunit of alpha DNA polymerase seems well established, efforts to find and characterize other enzymatic activities associated with the polymerase have intensified. The presence of an RNA primase activity in highly purified alpha polymerase has recently been established (I.R. Lehman, personal communication). Moreover, antibodies to the alpha and beta subunits of the alpha polymerase have recently been obtained, and should prove useful in defining the roles of the different polypeptides of the Drosphila alpha polymerase in replication. Initial experiments with these antibodies have shown that the subunits of the purified Drosophila alpha polymerase are in fact proteolytic degrada-

tion products of larger, hitherto undetected, proteins (I.R. Lehman, personal communication). While no ATPase, nuclease, or helicase activities have as yet been detected in this purified alpha polymerase, it seems likely that these activities are present at the eukaryotic replication fork, given what we have learned from prokaryotic systems. These activities could either be present in the nondegraded alpha polymerase (not yet isolated) or could be provided by more weakly associated proteins, analogous to the accessory proteins of the bacteriophage T4 replication system.

One area of eukaryotic in vitro DNA replication research in which recent progress has been exceptionally rapid is that of adenovirus DNA replication (Challberg, et al.,1980, Ikeda et al., 1981, Tamanoi and Stillman, 1982). Adenovirus has a double-stranded linear DNA genome of approximately 35,000 base pairs, about the size of the bacteriophage T7 genome. The replication of adenovirus is remarkable in several ways. For example, in vivo replication apparently requires only leading strand replication. A replication fork established at one end of the parental DNA molecule proceeds to the opposite end by strand diplacement synthesis; the diplaced genome length single-stranded molecule is then primed at its 3' end and used as a template to make the second daughter DNA molecule.

Perhaps the most unusual aspect of adenovirus replication is the mechanism of this priming at DNA ends, apparently accomplished by a protein:nucleotide primer complex (Rekosh et al., 1977). A protein with a molecular weight of 55,000 daltons is covalently attached to the adenovirus genome at each 5' end. This 55k "terminal protein" is

actually the proteolytic processing product of a protein with a molecular weight of 80,000 daltons, known as the "pre-terminal" protein. The 80k protein, which associates with nascent DNA, is required for the *in vitro* initiation of adenovirus DNA replication. This 80k preterminal protein forms a covalent complex with the priming nucleotide, dCMP, in a reaction requiring dCTP, ATP, Mg++, and an adenovirus origin DNA sequence located at or near the end of a linear duplex DNA molecule (Tamanoi and Stillman, 1982). The 80k:dCMP complex associates with the adenovirus DNA ends, where it provides a primer for elongation by an apparently unique aphidocolin-resistant DNA polymerase of 140,000 daltons. The 80k protein and DNA polymerase remain tightly associated in a 1:1 complex during subsequent elongation; this complex contains an ATPase activity and may act as a DNA helicase during the elongation reaction.

Another protein involved with adenovirus replication is the adenovirus DNA binding protein (AdDBP), which has a molecular weight of 72,000 daltons. Mutants in the gene for this protein are defective in DNA replication *in vivo* and this protein is clearly required for *in vitro* adenovirus replication. Both nuclear and cytosol fractions from uninfected cells are additionally required for the initiation and full-length replication of adenovirus DNA *in vitro*.

Clearly, the adenovirus system is a very attractive one for researchers hoping to develop a well-characterized eukaryotic counterpart to the *in vitro* DNA replication systems of the bacteriophages T7 and T4, and the *E. coli* bacterium.

REFERENCES

Abdul-Monem, M., and Hoffmann-Berling, H. (1980) Trends in Biochemical Sciences 5: 128-130

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

Alberts, B. M., and Sternglanz R, (1977) Nature 269: 655-661

Alberts, B. M., Barry J., Bedinger, P. ,Burke, R. L., Hitner, U., Liu, C.-C., and Sheridan, R. (1980) in Mechanistic Studies of DNA Replication and Genetic Recombination (Alberts, B. M., ed)Vol 19 pp 449-471 Academic Press

Barry, J. and Alberts, B. M., (1972) Proc. Nat. Acad. Sci. 69: 2717-2721

Bessman, M. J., Muzyczka, N., Goodman, M., and Schnaar, R. L. (1974) J. Mol. Biol. 88:409-421

Bittner, M., Burke, R.L., and Alberts, B. M. (1979) J. Biol. Chem. 254: 9565-9572

Burke, R. L., Alberts, B. M., and Hosoda, J. (1980) J. Biol. Chem. 255: 11484-11493

Brutlag, D., and Kornberg, A. (1972) J. Biol. Chem. 247: 241-248

Challberg, M. D., Desiderio, S. V., and Kelly, T. J., Jr. (1980) Proc. Nat. Acad. Sci. 77: 5105-5109

Chang, L. M. S. (1980) in Mechanistic Studies of DNA Replication and Genetic Recombination (Alberts, B. M., ed) Vol 19 pp 35-40 Academic Press, NY

Drake, J.W., (1973) Genetics supplement 73: 45-64

Epstein, R. H., Bolle, A., Steinberg, C. M., Kellenberger, E., Boy de la Tour, E., Chevalley, R., Edgar, R. S., Susman, M., Denhardt, G. H., and Lielausis, A. (1963) Cold Spring Harbor Symp. Quant. Biol. 28: 375-394

Fay, P. J., Johanson, K. D., McHenry, C. S., and Bambara, R. A. (1981) J. Biol. Chem,. 256: 976-983

Fay, P. J., Johanson, K. O., McHenry, C. S., and Bambara, R. A. (1982) Submitted to J. Biol. Chem.

Fischer, H., and Hinkle, D. C. (1980) J. Biol. Chem. 255: 7956-7964

Greve, J., Maestre, M. F., Moise, H., and Hosoda, J. (1978) Biochemistry 17: 893-898

Hitner, U., and Alberts, B. M. (1980) Nature 285: 300-305

Hopfield, J. J. (1976) Proc. Nat. Acad. Sci. 71: 4135-4139

Hosoda, J., Burke, R. L., Moise H., Kubota, I. and Tsugita, A. (1980) in Mechanistic Studies of DNA Replication and Genetic Recombination (Alberts, B. M., ed.) Vol. 19 pp 505-513 Academic Press, NY

Huang, C.-C., Hearst, J. E., and Alberts, B. M., (1981) J. Biol. Chem. 256: 4087-4094

Huberman, J. A., Kornberg, A., and Alberts, B. M. (1971) J. Mol. Biol. 62: 39-52

Ikeda, J.-E., Enomoto, T., and Hurwitz, J. (1981) Proc. Nat. Acad. Sci. 78: 884-888

Ikegama, S., Taguchi, T., Ohashi, M., Oguro, M., Nagano, H., and Mano, Y. (1978) Nature 275: 458-460

Kornberg, A. (1980) DNA Replication, W. H. Freeman, San Francisco

Krell, H., Durwald, H., and Hoffmann-Berling, H. (1979) Eur. J. Biochem. 93: 387-395

Liu, C.-C., Burke, R. L., Hitner, U., Barry J., and Alberts, B. M. (1978) Cold Spring Harbor Symp. Quant. Biol. 43: 469-487

Liu, C.-C., and Alberts B. M. (1980) Proc. Nat. Acad. Sci. 77: 5698-5702

Liu, C.-C., and Alberts, B. M. (1981) J. Biol. Chem. 256: 2821-2829

Liu, C.-C., and Alberts, B. M. (1981) J. Biol. Chem. 256: 2813-2820

Lo, K.-Y., and Bessman, M. J. (1976) J. Biol. Chem. 251: 2475-2479
Maynard-Smith, S., and Symonds, N. (1973) J. Mol. Biol. 74: 33-44

McCarthy, D., Minner, C., Bernstein, H., and Bernstein, C. (1976) J. Mol. Biol. 106: 963-981

Mosig, G., Luder, A., Garcia, G., Dannenberg, R., and Bock, S. (1978) Cold Spring Harbor Symp. Quant. Biol. 43: 501-515

Newport, J. W., Kowalczykowski, S. C., Lonberg, N., Paul, L. S. and Von Hippel, P. H. (1980) in Mechanistic Studies of DNA Replication and Genetic Recombination (Alberts, B. M., ed.) pp 485-505 Academic Press, NY

Ninio, J. (1974) Biochimie 57: 587-595

Ratner, D. (1974) J. Mol. Biol. 88: 373-383

Rekosh, D. M. K., Russell, W. C., Bellet, A. J. D., and Robinson, A. J. (1977) Cell 11: 283-295

Richardson, C. C., Romano, L. J., Kolodner, R., LeClerc, J. E., Tamanoi,

F., Engler, M., Dean, F. B., and Richardson, D. S. (1978) Cold Spring Harbor Symp. Quant. Biol. 43: 427-440

Riva, S. A., Cascino, A., and Geiduschek, E. P. (1970) J. Mol. Biol. 54: 85-102

Spanos, A., Sedgewick, S. G., Yarronton, G. T., Hutscher, U. and Banks, G. R. (1981) Nuc. Acids Res. 9: 1825-1839

Tamanoi, F., Engler, M. J., Ieschner, R., Orr-Weaver, T., Romano, L. J., Saito, H., Tabor, S. and Richardson, C. C. (1980) in Mechanistic Studies of DNA Replication and Genetic Recombination (Alberts, B . M., ed.) Vol 19 pp 411-428 Academic Press, NY

Tamanoi, F. and Stillman, B. W. (1982) Proc. Nat. Acad. Sci. 79: 2221-2225

Tomizawa, J. (1978) in DNA Synthesis; Present and Future. (Molineux, J. and Kohiyama, M., eds.) Plenum Publishing pp 797-825

Topal, M. D., and Sinha, N. K. (1982) Submitted to J. Biol. Chem.

Villani, G., Sauer, B., and Lehman, I. R. (1980) J. Biol. Chem. 255: 9479-9483

Warner, H. R., and Hotts, M. D. (1967) Virology 33: 376-384

Watson, J. D., and Crick, F. H. C. (1953) Nature 171: 964

Yarronton, G. T., and Gefter, M. L. (1979) Proc. Nat. Acad. Sci. 76: 1658-1662

CHAPTER ONE

THE 3'-5' PROOFREADING EXONUCLEASE OF BACTERIOPHAGE T4 DNA POLYMERASE IS STIMULATED BY OTHER T4 DNA REPLICATION PROTEINS

ABSTRACT

The bacteriophage T4 DNA polymerase has an intrinsic 3'-5' proofreading exonuclease activity that plays a central role in determining the fidelity of T4 DNA replication. In order to monitor this activity, we have measured the rate at which the polymerase decreases the size of a double-stranded DNA substrate in the absence of deoxyribonucleoside triphosphates. With this assay, we find that the addition of the polymerase accessory proteins, 45 protein and 44/62 protein, increases the rate at which the polymerase exonuclease digests the DNA substrate 3 to 4-fold. This stimulation requires the continuous hydrolysis of ATP catalyzed by the accessory protein complex. When added alone, the T4 helix-destabilizing protein, 32 protein, inhibits the exonuclease rate at high concentrations (>100 μg/ml), while stimulating about 3-fold at low concentrations. The 32 protein and the accessory proteins together increase the exonuclease rate 8- to 10-fold above that found for the polymerase alone. These exonuclease stimulations are likely to play a role in enhancing the fidelity of T4 DNA replication. The bacteriophage T7 DNA polymerase displays a similar 3'-5' exonuclease activity, but this exonuclease is not stimulated by any of the T4 replication proteins. It therefore appears that specific protein-protein interactions are involved.

The replication of DNA is accomplished by a multiprotein "replication apparatus," in which other proteins act in conjunction with the DNA polymerase to increase the speed and accuracy of replication (Alberts and Sternglanz, 1977). Several procaryotic in vitro DNA replication systems have been reconstructed from purified protein components, including that of bacteriophage T4 (Liu et al. 1978, Alberts et al., 1980). The T4 in vitro system is made up of seven highly purified replication proteins, most of which were identified and purified by using an in vitro complementation assay developed for this purpose that measures DNA synthesis in cell lysates (Barry and Alberts, 1972).

Current efforts in this laboratory are directed toward defining the function of each protein in the T4 DNA replication apparatus. Fortunately, it is possible to study partial reactions which involve only a subset of the seven purified proteins. This allows the complex replication process to be dissected into simpler elements, as required to investigate the role that each protein performs at the replication fork.

In this study we have monitored the intrinsic 3'-5' exonuclease activity of the T4 DNA polymerase (T4 gene 43 protein) on a double-stranded DNA substrate using an assay developed by Patrick O'Farrell (manuscript in preparation). This assay has provided a probe for analyzing the interactions of several T4 replication proteins with each other and with the DNA. The polymerase-associated exonuclease activity functions in a proofreading capacity during DNA replication due to its preferential excision of non-base paired (and therefore misincorporated) nucleotides from the 3'OH end of the

elongating DNA chain (Brutlag and Kornberg, 1972; Bessman et al., 1974). The results demonstrate that two types of T4 replication proteins - the helix destabilizing protein (gene 32 protein) and the polymerase accessory proteins (gene 44/62 and 45 proteins) - appreciably enhance the rate of DNA digestion by the exonuclease. The stimulations seem to require specific protein-protein interactions between these proteins and the DNA polymerase.

MATERIALS AND METHODS

Proteins - The bacteriophage T4 gene 43 protein, 45 protein and 44/62 proteins were purified as previously described to greater than 90% homogeneity (Morris et al., 1979). The T4 gene 32 protein was purified by published procedures to greater than 99% homogeneity (Bittner et al., 1979). The 32*I protein was generously provided by Junko Hosoda and Herb Moise, purified as published (Burke et al., 1980). Purified 44 protein free of 62 protein was provided by Maureen Munn in this Department. All of the purified replication proteins were free of detectable endonuclease and (excepting polymerase) of exonuclease activity under the assay conditions used (Morris et al., 1979; Bittner et al., 1979). T7 DNA polymerase was the generous gift of Steve Matson and Charles Richardson, purified by the method of Fischer and Hinkle (1980). Purified SSB protein was a gift from Arthur Kornberg, purified as published (Weiner et al., 1975). Endonuclease EcoRI was purified in our laboratory using published methods (Sumegi et al., 1977). The nuclease-free bovine serum albumin used in these reactions was prepared as described (Bittner et al. 1979).

Nucleotides and DNA - Unlabeled nucleotides were purchased from Sigma and stored frozen in 20 mM Tris acetate, pH 7.8. Radioactively labeled nucleotides were purchased from New England Nuclear or Amersham. Plasmid pBR322 DNA was prepared by published methods (Clewell, 1972).

Exonuclease assay - Plasmid pBR322 DNA was digested with EcoR1, producing a linear monomer of 4362 base pairs (each end having a 5' phosphate- terminated single-stranded extension four bases long). The linear DNA was then treated with polynucleotide kinase (New England Nuclear) and $\gamma[^{32}P]$ATP to label the 5' termini with ^{32}P. Free radioactive nucleotide was removed by spin dialysis (Neal and Florini, 1973) through a CL6B gel column (Pharmacia). Reactions contained the DNA (2.5 1/m2/g/ml unless noted) and 2.5 1/m2/g/ml of T4 DNA polymerase (43 protein) - plus 33 mM Tris acetate, pH 7.8, 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM dithiothreitol, and 100 1/m2/g/ml nuclease-free bovine serum albumin. Those reactions containing accessory proteins in addition contained 500 1/m2/M rATP, 15 µg/ml 45 protein and 20 1/m2/g/ml 44/62 protein. Various concentrations of 32 protein were used, as noted in figure legends. Reactions were carried out at 37°C and started by addition of either the DNA or the DNA polymerase. At varying intervals, reactions were stopped by the addition of sodium dodecyl sulfate to 1%, and the product DNA was sized on 1% agarose (Seakem) gels (0.4 x 14 x 21.5 cm) run in TEA buffer (40 mM Tris acetate, pH 8.1, 20 mM sodium acetate, 2 mM Na_3EDTA) for 15 hours at 50 V. Gels were then dried onto filter paper and exposed to Kodak XAR film with Dupont Lighting Plus intensifying screens at -70°C. Rates of exonuclease digestion were determined from the resulting autoradiograms by measuring the decrease in DNA size with time, using restriction fragments of known molecular weight as size standards (method of Dr Patrick O'Farrell, manuscript in preparation).

Assays for DNA synthesis and deoxyribonucleoside triphosphate turnover - DNA synthesis reactions were carried out under the conditions described for exonuclease reactions, except that native bacteriophage T7 DNA was used as substrate and 200 μM each of dATP, dGTP, dCTP and (α-^{32}P)TTP were included. Reactions were spotted on PEI cellulose sheets (Baker Flex) along with cold marker nucleotides and then chromatographed in 1 M LiCl. Spots visualized by UV light corresponding to polynucleotide and to nucleoside monophosphate were scraped off the plastic backing, eluted with 0.1 N HCl, and then counted in a triton-toluene scintillation fluid.

RESULTS

The accessory proteins for the T4 DNA polymerase stimulate the polymerase-associated 3'-5' exonuclease activity- During polymerization, the T4 DNA polymerase dissociates frequently from the 3'OH of the elongating DNA chain at the salt concentrations used in these studies (J. Newport, 1980; D. Mace, 1975). The 3'-5' exonuclease associated with the T4 DNA polymerase likewise acts distributively on a single-stranded DNA substrate, removing only a few nucleotides from each DNA chain end before dissociating (Nossal and Hershfield, 1971; Thomas and Olivera, 1978). A double-stranded DNA molecule treated with the exonuclease degraded at both of its 3'OH termini, decreasing in size as a function of time. An analysis of the DNA product length by agarose gel electrophoresis reveals that a narrow size distribution of products is maintained at all polymerase concentrations, as expected if the exonuclease acts distributively and dissociates frequently from each individual DNA molecule (P. O'Farrell, manuscript in preparation; see also lanes c to e of Fig. 1).

T4 genes 44, 62, and 45, which are essential for in vivo T4 DNA replication (Epstein et al., 1963), encode proteins known as "polymerase accessory proteins." These proteins have been highly purified in this laboratory and shown to have a DNA-dependent ATPase activity (D. Mace, 1975; Piperno and Alberts, 1978; Piperno et al., 1978). In a reaction requiring ATP hydrolysis, the accessory proteins increase the rate of DNA synthesis by an individual T4 DNA polymerase molecule (D. Mace, 1975) and allow this polymerase to

more rapidly traverse helical hairpin structures in a single-stranded DNA template by greatly increasing its processivity (Huang et al., 1981; Roth et al., 1982).

Previous work indicated that the polymerase accessory proteins may also affect the exonuclease activity of T4 DNA polymerase, since these accessory proteins cause an increase in the ratio of turnover to incorporation of deoxyribonucleoside triphosphates substrates during in vitro DNA synthesis by the polymerase (Liu et al., 1978). In order to test this point more directly, an experiment was performed in which we measured the effect of the addition of 45 protein, 44/62 protein and ATP on the digestion of double-stranded DNA by the polymerase-associated exonuclease. The results in Fig. 1 (lanes f to h) reveal that this accessory protein mixture increases the rate of digestion of a double- stranded linear DNA molecule by about 3 to 4-fold. A similar type of result is obtained lwhen T4 gene 32 protein is added to polymerase-associated exonuclease reactions (see below). There is no effect of the accessory proteins unless all three components are added: 45 protein, 44/62 protein and rATP (Fig. 1). Moreover, purified 44 protein, which contains the full ATPase activity of 44/62 protein (unpublished results of Maureen Munn), does not substitute for the 44/62 protein complex (Fig. 1, lane m).

There is a much greater degree of variability in the DNA product size when the accessory proteins (plus ATP) are included in the reactions, all of which are performed in the presence of an excess proteins over DNA, as can be seen from the "smearing out" of the DNA bands in Fig. 1 (compare lanes f to h with lanes c to e, also see Fig. 2, compare lanes b and d). This suggests that the exonuclease

is now acting more processively, as well as moving more rapidly along the DNA.

With all of the other components in excess, the activity of the polymerase-associated exonuclease was measured as the concentration of one of the accessory proteins was varied, in order to estimate the number of 44/62 protein and 45 protein molecules required per DNA end. The results obtained at a constant DNA termini concentration of 1.8×10^{-9} M are shown in Fig. 2. The half-maximal stimulation by the 44/62 protein complex (native molecular weight of 180,000) is reached at 1.1×10^{-9} M, corresponding to a ratio of 0.6 44/62 protein molecules per DNA terminus (Fig. 2A). For 45 protein (a dimer with a molecular weight of 49,000 daltons) the concentration necessary to achieve maximal stimulation is at least 1.5 to 2×10^{-8}M, or 8 to 11 dimers per DNA end. However, unlike the response to 44/62 addition, further 45 protein additions continue to increase the exonuclease rate somewhat (Fig. 2B).

The results in Fig. 2 suggest very different affinities of the two accessory proteins for the polymerase-DNA complex. Thus, only one 44/62 protein molecule seems to interact with each DNA terminus, and a the form of the titration curve indicates that tight complex is formed. In contrast, many 45 protein molecules are required per DNA end for maximal stimulation of the polymerase-associated exonuclease, which is probably due to the much lower affinity of 45 protein for the active complex. Data on the accessory protein-promoted DNA synthesis on a primed single-stranded template, likewise suggests that only one 44/62 molecule per primer end is

required, compared to many 45 protein molecules (Newport et al., 1980).

The exonuclease stimulation by the polymerase accessory proteins requires the continuous hydrolysis of ATP - All of the various activities thus far detected for the accessory proteins require their concomitant hydrolysis of ATP (or dATP) to ADP (or dADP) and inorganic phosphate. In DNA synthesis reactions, the frequency of hydrolysis required varies widely depending on the type of reaction and is seemingly correlated with the difficulty of moving the replication complex forward in the polymerizing direction. Thus, given an ample dNTP supply for polymerization and a single-stranded DNA template, the accessory proteins need to hydrolyze ATP only once every 5 to 10 minutes in order to function as a "sliding clamp" for the polymerase (Huang et al., 1981). In contrast, when the DNA polymerase is confronted with a double-stranded DNA template to copy (or with a very low dNTP substrate concentration on a single-stranded DNA template), ATP hydrolysis is required many times each minute for the accessory proteins to help the polymerase (Alberts et al., 1980; Newport et al., 1980). These results are consistent with the view that ATP is hydrolyzed in order to assemble a polymerase-accessory protein complex, whose stability (and therefore need for reassembly varies depending on the DNA polymerase environment (Alberts et al., 1980).

In order to examine the nature of the ATP hydrolysis requirement for the accessory-protein exonuclease stimulation, the competitive inhibitor rATPγS was added after a time delay to block sub-

sequent ATP hydrolysis by the accessory proteins. The ATPγS is not hydrolyzed by the 44/62 protein and 45 protein complex (Piperno and Alberts, 1978), and under our conditions the K_i for the inhibition of the accessory protein ATPase is about 40 μM (unpublished results of P.B.), a value about 3-fold lower than the K_m for rATP (Piperno et al., 1978). At the ATPγS to ATP ratio of 10:1 used, the ATPγS blocks all of the in vitro replication reactions requiring accessory proteins (Piperno and Alberts, 1978; Alberts et al. 1980), as well as preventing any exonuclease stimulation when present at the start of a reaction (data not shown). Figure 3 shows the results of adding excess rATPγS (500 μM) to an ongoing exonuclease digestion in the presence of accessory proteins and 50 μM rATP. This addition caused a rapid cessation of the accessory protein stimulation of the reaction (compare lanes b and c), revealing that a frequent hydrolysis of ATP is required to maintain the stimulation. However, the heterogeneous product size seen after rATPγS addition (compare lanes e and f) suggests that a small fraction of the polymerase-accessory protein complexes survive for a period of minutes.

The effect of the gene 32 protein on the polymerase-associate exonuclease activity - The T4 gene 32 protein is required in vivo for all of the major DNA reactions in a T4-infected cell, including the replication, recombination and repair of DNA. This 33,500 dalton protein binds tightly and cooperatively to single-stranded DNA (Alberts, 1970), and it can also be shown to bind directly to several T4 replication proteins, including the T4 DNA polymerase (Huberman et al., 1971) and the gene 61 protein (Burke et al., 1980).

Previous studies have shown that the addition of saturating levels of 32 protein strongly inhibits the action of the T4 polymerase-associated exonuclease on single-stranded DNA (Huang and Lehman, 1972). We find that a high concentration of 32 protein likewise inhibits the digestion of double-stranded DNA by this exonuclease. However, at a lower concentration of 32 protein, the rate of exonuclease digestion is stimulated. Figure 4 illustrates the effects of different 32 protein concentrations on the exonuclease rate: in this series, a maximal stimulation of about 3-fold is obtained at 5.5 μg/ml 32 protein, while there is a strong inhibition at 32 protein concentrations above 100 μg/ml.

The core T4 in vitro DNA synthesis system contains 43 protein, 32 protein, 45 protein, and 44/62 protein, which is the minimum set of replication proteins necessary for efficient in vitro DNA synthesis on a double-stranded DNA template. Synthesis begins by covalent addition of nucleotides onto the 3'OH end at a nick in the DNA double helix, with simultaneous strand displacement starting from the 5' end at the nick (Liu et al., 1978; Nossal and Peterlin, 1979; Sinha et al., 1980). This type of DNA synthesis is rapid and processive.

We have measured the rate of the polymerase-associated exonuclease degradation of double-stranded DNA in the five-protein system by omitting the deoxyribonucleoside triphosphates required for DNA synthesis. When 32 protein is added to a reaction mixture containing accessory proteins plus the DNA polymerase, the rate of exonuclease action is increased a further 2-fold (producing a total 6 to 8-fold stimulation above the exonuclease rate found for DNA polymerase alone). Further, the accessory proteins prevent the

32 protein from acting in a mode that interferes with the polymerase-exonuclease, since digestion is of the exonuclease is stimulated even at high 32 protein concentrations (Fig. 5).

The 32*I and E. coli SSB proteins lack the capacity to stimulate the exonuclease - 32*I protein is a proteolytic product of the 32 protein that is missing the first 60 amino acids from the COOH terminus of the intact 32 protein (Hosoda and Moise, 1978; Williams and Koningsberg, 1978). This 32*I protein binds with a somewhat higher affinity than 32 protein to single-stranded DNA, and it has apparently lost the capacity to bind directly to the T4 DNA polymerase (Greve et al., 1978; Burke et al., 1980). When we added 32*I protein to the T4 DNA polymerase, we found that its exonuclease activity was strongly inhibited at all of the 32*I concentrations tested, including those equivalent to the stimulatory concentrations of 32 protein (Fig. 5A). This implies that the 32*I protein is much more efficient than 32 protein in its inhibitory mode, and/or that it lacks the necessary structure(s) to stimulate the exonuclease action of the polymerase. The protein that plays a role analogous to 32 protein as a helix destabilizing protein in E. coli is the SSB protein (18,500 daltons). The effect of adding highly purified SSB protein to the T4 DNA polymerase is also shown in Fig. 5A. This protein, although very similar in function to the T4 32 protein, cannot stimulate the polymerase-associated exonuclease, and it inhibits at all concentrations tested.

Both the 32*I protein and SSB protein also inhibit the polymerase-associated exonuclease when normal concentrations of the

polymerase accessory proeins are present (Fig. 5B). At higher accessory protein concentrations this inhibition is relieved; however, the rate of exonuclease digestion never exceeds that seen with the polymerase and accessory proteins alone. Hence, we conclude that neither 32*I protein nor SSB protein are able to substitute for 32 protein in reactions in which the 32 protein stimulates the rate of exonuclease digestion.

Previous work has demonstrated that 32*I protein can efficiently substitute for 32 protein in a DNA synthesis reaction on a double-stranded DNA template catalyzed by the core replication system (Burke et al., 1980). Levels of exonuclease activity during polymerization can be ascertained by measuring the amount of dNTP "turned over" into the monophosphate formed during the reaction, using thin layer chromatography: the dNMPs represent nucleotides that have been incorporated into DNA by the polymerase and then excised by its 3'-5' exonuclease function (Hershfield and Nossal, 1972). Because of the observed inhibitory effect of 32*I protein on the polymerase-associated exonuclease in the absence of DNA synthesis, we compared the level of nucleotide turnover in DNA synthesis reactions containing either intact 32 protein or 32*I protein. As shown in Table I, there is about a 1:1 ratio of turned-over (excised) nucleotide to stably incorporated nucleotide in both cases. Thus, the inhibition of the polymerase-associated exonuclease exerted by 32*I protein in the absence of DNA synthesis is no longer seen when DNA synthesis is allowed. This result suggests that the exonuclease reaction is somehow altered when the polymerase is in the forward (polymerizing) direction (see Discussion).

The stimulation of the T4 DNA polymerase-associated exonuclease by the T4 replication proteins requires specific interactions -The effects of the T4 replication proteins on the rate of exonuclease digestion described above could be mediated either through the DNA or via protein-protein interactions. In the former case, the replication proteins might stimulate the exonuclease simply by altering the structure of the 3'OH DNA end, and the 3'-5' exonucleases of other DNA polymerases could to be similarly affected. However, if direct protein-protein interactions are involved, the T4 replication proteins would not be expected to stimulate other polymerase-associated exonucleases. We therefore repeated our experiments with the bacteriophage T7 DNA polymerase, which contains a 3'-5' exonuclease activity similar to that of the T4 enzyme (Hori et al., 1979; Adler and Modrich, 1978).

In contrast to the results obtained with the T4 DNA polymerase, addition of the T4 gene 45 and 44/62 proteins inhibited the exonuclease rate of the T7 DNA polymerase (Fig. 6, lanes a through g). Likewise, the 32 protein strongly inhibited the 3'-5' exonuclease activity of the T7 enzyme, even at the low 32 protein concentrations that stimulate the T4 polymerase-associated exonuclease. The proteolytic fragment of 32 protein, 32*I protein, also strongly inhibited the exonuclease activity of T7 DNA polymerase (Fig. 6, lanes h through m).

These results imply that the stimulation of the exonuclease rate caused by the T4 accessory proteins and that caused by the 32 protein both involve specific protein-protein interactions with the T4 DNA polymerase. In contrast, the inhibitions of polymerase-

associated exonuclease activities by high 32 protein concentrations and by 32*I protein would seem to reflect a less specific protein-induced perturbation in the structure of the DNA substrate.

DISCUSSION

Different polymerase-associated exonuclease rates are observed with and without nucleotide polymerization - Specific protein-protein and protein-DNA interactions are essential to mediate the construction of a multienzyme replication fork. We have probed some of these macromolecular interactions by measuring the effects of several purified T4 replication proteins on the 3'-5' exonuclease activity of the T4 DNA polymerase.

From previous work in our laboratory we expected that the T4 replication proteins that we tested -45 protein, 44/62 protein and 32 protein - might have an effect on the polymerase-associated exonuclease, since Dr. Ula Hibner had observed that in a reaction containing polymerase accessory proteins and 32 protein (double-stranded DNA template), the ratio of deoxyribonucleoside triphosphate turnover to nucleotide incorporation into DNA approximately 4-fold greater than in a reaction containing DNA polymerase alone (single-stranded DNA template; Liu et al. 1978). In agreement with this expectation, the rate of digestion of double-stranded DNA by the polymerase-exonuclease in the absence of deoxyribonucleoside triphosphates is stimulated 6 to 8-fold in the core replication system (Fig. 5B).

However, the absolute rate that we have observed for the exonuclease is 30 to 40-fold slower than the rate one might expect from the ratio of polymerization to turnover found in a DNA synthesis reaction. Thus, the rate of nucleotide polymerization catalyzed by the core replication system on a double-stranded DNA template under our conditions is about 80 nucleotides/second (Liu et al., 1978).

Inasmuch as the synthesis-dependent turnover of deoxyribonucleoside triphosphates to monophosphates in this reaction is approximately equal to the amount incorporated (see Table I), an exonuclease rate of about 80 nucleotides/sec would be expected. Instead, in the absence of polymerization we find a maximal rate of only 2 to 3 nucleotides digested per second with the 43, 44/62, 45, and 32 proteins present (see Fig. 5B). It therefore seems that the DNA polymerase molecule has a greatly enhanced exonuclease rate when it is actively polymerizing deoxyribonucleoside triphosphates. Other evidence that the exonuclease activity of the polymerase is qualitatively different during synthesis comes from our experiemtns with 32*I protein. In our assays, this protein strongly inhibits exonuclease activity, even when the 44/62 and 45 proteins are present (Fig. 5B). However, in a DNA synthesis assay on a double-stranded DNA template (which differs from our exonuclease assay only in that deoxyribonucleoside triphosphates are included), the 32*I protein can efficiently substitute for 32 protein, and normal rates of nucleotide turnover by the exonuclease are observed (Table I).

A third difference between "backward" and "forward" moving polymerase-accessory protein complexes lies in the different frequency of ATP hydrolysis that is required for maintenance of an accessory protein effect. This ATP hydrolysis is thought to be required for the assembly of an accessory protein: DNA complex (or a tertiary complex between DNA) accessory proteins, and T4 DNA polymerase) that acts as a "sliding clamp" that ties down the polymerase at a DNA primer end (Huang et al., 1981). In this view, frequent ATP hydrolysis is required only if the accessory protein complex is unstable

(Alberts et al., 1980). In order to maintain the observed stimulation of the polymerase-associated exonuclease by the accessory proteins, frequent ATP hydrolysis is necessary (Fig. 3). In contrast, during DNA synthesis on a primed single-stranded DNA template, an active accessory protein lifetime of up to 10 minutes is observed in the presence of 100 μM deoxyribonucleoside triphosphates (Huang et al., 1981). An intermediate lifetime has been reported using a primed single-stranded DNA template at a low deoxyribonucleoside triphosphate concentration (4 μM) (Newport et al., 1980). In general, this data suggests that the stability of the polymerase accessory protein complex is increased by the presence of deoxyribonucleoside triphosphates.

Evidence for specific protein-protein interactions - The interactions between proteins at the T4 replication fork appear to be quite specific. For example, a helix destabilizing protein from E. coli, the SSB protein, cannot substitute for 32 protein in stimulating the T4 polymerase exonuclease even though these two proteins have a similar functional role in vivo (Fig. 5). Further, none of the T4 replication proteins stimulate the similar exonuclease activity of the T7 DNA polymerase (Fig. 6). Direct evidence for specific protein-protein interactions between the various T4 replication proteins has been obtained by protein affinity chromatography including discovery of a weak 45 protein-T4 DNA polymerase complex (T. Formosa, R.L. Burke and B.M. Alberts, manuscript in preparation).

Our current concept of the activities of the various T4 replication proteins investigated here at DNA ends is presented in Fig. 7.

The polymerase accessory proteins form a complex that ties down the DNA polymerase to the 3' OH end of the DNA (or RNA) primer. In this complex, the 44 protein subunit is likely to interact with the DNA, since when this subunit is dissociated from its tightly bound 62 protein subunit, it displays the full DNA-dependent ATPase activity of the complex (unpublished results of Maureen Munn). Both the 44 protein subunit (with or without 45 protein) and the 44/62 protein (without 45 protein) inhibit the exonuclease activity of T4 DNA polymerase to some extent (Fig. 1 and unpublished results). This suggests that all three polypeptide chains (44, 62 and 45) are required to form a complex that is useful for the polymerase. Figure 2A reveals that the 44/62 protein molecules form a 1:1 complex with the 3'OH ends of the DNA. Since direct binding occurs between the 45 protein and the T4 DNA polymerase and between the 45 protein and the 44/62 protein, (T. Formosa, R.L. Burke and B.M. Alberts, in preparation) the 45 protein may act to join the 44/62 protein and the polymerase. However, saturation of the exonuclease stimulation by the accessory proteins requires a large number of 45 protein dimers for each 3'OH DNA end (Fig. 2B). Therefore, the 45 protein is either very weakly bound in this complex, or it acts catalytically in the assembly process in a rather inefficient manner.

Our results suggest that the binding of the polymerase accessory proteins to a DNA end prevents a nonproductive (inhibitory) interaction of 32 protein at this site. Since the 32 protein inhibits the polymerase-associated exonuclease activity on a single-stranded DNA molecule (Huang and Lehman, 1972), we propose that a high concentration of 32 protein inhibits the polymerase exonuclease on double-

stranded DNA by causing a local melting of the 3'OH and that allows the adjacent nucleotides to bind to 32 protein (Fig. 7). The inhibitory effects of the (32*I protein and the E. coli SSB protein would seem to be due to a similar type of protein-DNA interaction. By "clamping" down the 3'OH primer end of the DNA, the accessory proteins would make it inaccessible to 32 protein binding. This makes possible a stimulation of the polymerase-associated exonuclease by 32 protein, which appears to depend on specific protein-protein interactions that are not available to either E. coli SSB protein or the 32*I protein. The intact 32 protein binds to the T4 DNA polymerase, while 32*I protein (lacking the COOH-terminal "A-peptide" of intact 32 protein) does not (Burke et al., 1981, T. Formosa, R.L. Burke and B.M. Alberts, in preparation). This contact of the COOH-terminal region of 32 protein with the T4 DNA polymerase may serve to enhance the activity of its associated exonuclease (Fig. 7).

Possible Relevance to the fidelity of DNA replication - The exonuclease:polymerase ratio of 43 protein is considered to be an important factor in the expression of the mutator and antimutator phenotypes of various gene 43 mutants, because in vitro studies have shown that this activity ratio is unusually high for a purified antimutator DNA polymerase and unusually low for some of the purified mutator DNA polymerases (Lo and Bessman, 1976; Bessman et al., 1974). Our results suggest that the other T4 replication proteins may contribute to the fidelity of DNA replication through their effect on the 3'-5'exonuclease activity of the T4 DNA polymerase.

Genetic studies have shown that replication proteins other than the DNA polymerase play a role in determining the fidelity of replication, since mutations in genes 32, 41, 44, 45 and 62 can alter mutation frequencies in vivo (Watanabe and Goodman, 1978; Mufti, 1979). The seven-protein T4 replication system developed in our laboratory replicates DNA with a high fidelity (Hibner and Alberts, 1980; Sinha and Haimes, 1980; Sinha and Haimes, 1981). One mechanism by which the T4 replication proteins might be suspected to contribute to this fidelity is by their enhancement of the 3'-5' proofreading exonuclease activity of the T4 DNA polymerase. The data presented in this paper show that several of the T4 replication proteins do in fact have a marked stimulatory effect on the exonuclease of the T4 DNA polymerase in vitro, and we suggest that a similar effect at a replication fork serves to increase the effectiveness of the proofreading process that is carried out by the T4 DNA polymerase.

REFERENCES

Adler, S., and Modrich, P. (1979) J. Biol. Chem. 254, 11605-11614.

Alberts, B.M. (1970) Fed. Proc. 29, 1154-1163.

Alberts, B., and Sternglanz, R. (1977) Nature 269, 655-661.

Alberts, B.M., Barry, J., Bedinger, P., Burke, R.L., Hibner, U., Liu, C.-C., and Sheridan, R. (1980) In Mechanistic studies of DNA Replication and Genetic Recombination, Vol. 19 ICN-UCLA Symposium on Molecular and Cellular Biology. (Alberts, B.M., ed) pp. 449-471, Academic Press, New York.

Barry, J., and Alberts, B. (1972) Proc. Natl. Acad. Sci. USA 69, 2717-2721.

Bessman, M.J., Muzyczka, N., Goodman, M.F., and Schnaar, R.L. (1974) J. Mol. Biol. 88, 409-421.

Bittner, M., Burke, R.L., and Alberts, B.M. (1979) J. Biol. Chem. 254, 9565-9572.

Brutlag, D., and Kornberg, A. (1972) J. Biol. Chem. 247, 241-248.

Burke, R.L., Alberts, B.M., and Hosoda, J. (1980) J. Biol. Chem. 255, 11484-11493.

Clevell, D.B. (1972) J. Bacteriol. 110, 667-676.

Epstein, R.H., Bolle, A., Steinberg, C.M., Kellenberger, E., Boy de la Tour, E., Chevalley, R., Edgar, R.S., Susman, M., Dehardt, G.H., and Lielausis, A. (1963) Cold Spring Harbor Symp. Quant. Biol. 28, 375-394.

Fischer, H., and Hinkle, D.C. (1980) J. Biol. Chem. 255, 7956-7964.

Greve, J., Maestre, M.F., Moise, H., and Hosoda, J. (1978) Biochemistry 17, 893-898.

Hershfield, M.S., and Nossal, N.G. (1972) J. Biol. Chem. 247, 3393-3404.

Hibner, U., and Alberts, B.M. (1980) Nature 285, 300-305.

Hori, K., Mark, D.F., and Richardson, C.C. (1979) J. Biol. Chem. 254, 115989-11604.

Hosoda, J., and Moise, H. (1978) J. Biol. Chem. 253, 7547-7555.

Huang, C.-C., Hearst, J.E., and Alberts, B.M. (1981) J. Biol. Chem. 256, 4087-4094.

Huang, W.M., and Lehman, I.R. (1972) J. Biol. Chem. 247, 3139-3146.

Huberman, J.A., Kornberg, A., and Alberts, B.M. (1971) J. Mol. Biol. 62, 39-52.

Liu, C.-C., Burke, R.L., Hibner, U., Barry, J., and Alberts, B.M. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 469-487.

Lo, K.-Y., and Bessman, M.J. (1976) J. Biol. Chem. 251, 2475-2479.

Mace, D. (1975) Ph.D. thesis. Princeton University.

Morris, C.F., Hama-Inaba, H., Mace, D., Sinha, N.K., and Alberts, B.M. (1979) J. Biol. Chem. 254, 6787-6796.

Mufti, S. (1979) Virology 94, 1-9.

Neal, M.W., and Florini, J.R. (1973) Anal. Biochem. 55, 328-330.

Newport, J.W., Kowalczykowski, S.C., Lonberg, N., Paul, L.S., and Von Hippel, P.H. (1980) Mechanisms in Mechanistic Studies of DNA Replication and Genetic Recombination, Vol. 19 ICN-UCLA Symposium on Molecular and Cellular Biology (B.M. Alberts, ed) pp 485-505. Academic Press, New York.

Newport, J.W. (180) Ph.D. Thesis, University of Oregon.

Nossal, N.G., and Hershfield, M.S. (1971) J. Biol. Chem. 246, 5414-5426.

Nossal, N.G., and Peterlin, B.M. (1979) J. Biol. Chem. 254, 6032-6037.

O'Farrell, P.H., Kutter, E., and Nakanishi, M. (1980) Molec. gen. Genet. 179, 421-435.

Piperno, J.R., and Alberts, B.M. (1978) J. Biol. Chem. 253, 5174-5179.

Piperno, J.R., Kallen, R.G., and Alberts, B.M. (1978) J. Biol. Chem. 253, 5180-5185.

Roth, A.C., Nossal, N.G., and Englund, P.T. (1982) J. Biol. Chem. 257, 1267-1273.

Sinha, N.K., Morris, C.F., and Alberts, B.M. (1980) J. Biol. Chem. 255, 4290-4303.

Sinha, N.K., and Haimes, M.D. (1980) in Mechanisms of DNA Replication and Genetic Recombination, Vol. 19 ICN-UCLA Symposium on Molecular and Cellular Biology (Alberts, B.M., ed.) pp. 707-723. Academic Press, New York.

Sinha, N.K., and Haimes, M.D. (1981) J. Biol. Chem. 256, 10671-10683.

Sumegi, J., Breedveld, D., Hossenlopp, P., and Chambon, P. (1977) Biochem. Biophys. Res. Commun. 76, 78-85.

Thomas, K.R., and Olivera, B.M. (1978) J. Biol. Chem. 253, 424-429.

Watanbe, S.M., and Goodman, M. (1978) J. Virology 25, 73-77.

Weiner, J.H., Bertsch, L.L., and Kornberg, A. (1975) J. Biol. Chem. 250, 1972-1980.

Williams, K.R. and Konigsberg, W. (1978) J. Biol. Chem. 253, 2463-2470.

Table I. The 32*I protein does not inhibit turnover of deoxyribonucleoside triphosphates to monophosphates during in vitro replication on a double-stranded DNA template.

Replication reactions including polymerase, accessory proteins and either 32 protein or 32*I protein at 100 µg/ml were performed as described in Materials and Methods. Reactions were started by the addition of a nicked double-stranded T7 DNA template to a final concentration of 7.5 µg/ml. After 5 and 10 minutes at 37°C, aliquots were removed from the reactions and stopped by the addition of 66 mM Na_3EDTA. Subsequent analysis of deoxyribonucleotide incorporation and turnover was carried out as described in Methods.

Protein Components in Reaction	minutes at 37°C	turnover of dNTPs to dNMPs (pmoles)	dNTPs incorporated (pmoles)	ratio of turnover to incorporation
43,44/62,45 and 32 proteins	5	86.2	105	0.82
	10	159	144	1.1
43,44/62,45 and 32*I proteins	5	85.4	79.4	1.08
	10	110	136	0.81

FIGURE LEGENDS

Fig. 1. Gel electrophoretic analysis of the stimulation of the 3'-5' exonuclease activity of T4 DNA polymerase by the polymerase accessory proteins (44/62 protein and 45 protein).
Exonuclease assays were performed as described in Materials and Methods with rATP present. Lanes a and n show an Xba I digest of cytosine-containing T4 DNA used as molecular weight markers. Lanes b and i contain linear pBR322 DNA, the substrate in the exonuclease assay. Lanes c through e show the substrate DNA after digestion for 5,10 and 15 minutes with the T4 DNA polymerase alone. Lanes f through h show DNA from identical exonuclease reactions that include 44/62 protein, 45 protein and rATP; the rate of the exonuclease under these conditions is increased about 4-fold. Lanes j through l show that all three components are required to achieve stimulation by the accessory proteins, and that the purified 44 subunit of 44/62 protein cannot substitute for the intact 44/62 protein (lane m).
Note that the exonuclease digests inward from both 3'OH ends of the double-stranded linear DNA molecule; thus, it reaches the center of the substrate molecule after the excision of about 2000 bases, causing the DNA to fall apart into two single-stranded species (lane h).

Fig. 2. Titration of amounts of each accessory protein required for stimulation of the polymerase-associated exonuclease.
Exonuclease reactions were carried out as described in Materials and Methods with varying concentrations of either 44/62 protein (A) or 45

protein (B) at a DNA terminus concentration of 1.85 x 10^{-9}M. Rates of exonuclease digestion were calculated from a gel analysis such as that shown in Fig. 1, as described in Methods. Note even that low concentrations of 44/62 protein showed a substantial effect, whereas concentrations of 45 protein below 2 µg/ml are inactive; this lag in the stimulation by 45 protein may reflect a minimal concentration required for its dimerization. The accessory protein stimulation is saturated in terms of 44/62 protein at 5 x 10^{-9}M, suggesting that a single protein molecule is bound per DNA end in this reaction (A). In contrast, the stimulation is not saturated in terms of 45 protein even at a 100-fold excess of 45 protein dimers per DNA terminus (B).

Fig. 3. Frequent hydrolysis of rATP by the accessory proteins is required for exonuclease stimulation.

Exonuclease reactions were performed as described in Materials and Methods, except that rATP was present at 50 µM instead of 500 µM in the reactions that included polymerase and accessory proteins. Reactions with polymerase only (lanes a and d) and polymerase plus accessory proteins with or without 500 µM rATPɣS addition at 2 minutes were carried out in parallel. Within one minute after ATPɣ addition, a dramatic reduction in the accessory protein stimulation of the exonuclease is evident (compares lanes b and c). However, there is an unusually broad distribution of DNA product size 8 minutes after rATPɣS addition (lane f), revealing that the accessory protein stimulationm has not been completely blocked by the ATPɣS addition.

Fig. 4. The effect of 32 protein concentration on the rate of the polymerase-associated exonuclease.
Exonuclease reactions were carried out as described in Materials and Methods for 10 minutes at 37°C with T4 DNA polymerase present at 2.5 µg/ml and 32 protein present at indicated concentrations. At the concentrations tested above 88 µg/ml, 32 protein inhibits the exonuclease, whereas it stimulates the exonuclease at lower concentration. The apparent maximal stimulation of 3-fold is seen at a concentration of 32 protein of about 5 µg/ml.

Fig. 5. Helix destabilizing proteins other than intact 32 protein inhibit the T4 polymerase-associated exonuclease. Exonuclease reactions were performed as described in Materials and Methods for 8 min at 37°C. The results are presented as microdensitometer tracings from autoradiographs of 0.8% agarose gels elecrophoresed as described in Materials and Methods. (A). Reactions containing only the T4 DNA polymerase and various helix-destabilizing proteins. The concentration of 32*I protein and E. coli SSB protein used was 10 µg/ml. In contrast to the 32 protein stimulation seen at an equivalent concentration of intact 32 protein (Fig. 4), these other proteins strongly inhibit the exonuclease activity (B). Reactions including the T4 polymerase accessory proteins. These reactions contained 44/62 protein at 20µg/ml, 45 protein at 18 µg/ml and rATP at 0.5 mM. The 32 protein and 32*I protein were present at 100 µg/ml and the E. coli SSB protein at 60 µg/ml in the indicated reactions. Only intact 32 protein stimulated the polymerase-associated exonuclease. 32*I protein and E. coli SSB were inhibitory (compare to panel A).

Fig. 6. The T4 replication proteins fail to stimulate the 3'-5' exonuclease activity of the T7 bacteriophage DNA polymerase.
Exonuclease reactions were performed as described and were stopped after 2 minutes at 37°C. T7 DNA polymerase was present in the indicated reactions at 3 μg/ml. Lanes a, g, and h contain molecular weight markers. Lanes b and i contain undigested linear pBR322 DNA. A comparison of lanes c and d, which show products of reactions with T4 and T7 DNA polymerases respectively, indicate that the 3'-5' exonuclease of the T7 DNA polymerase is more active than that of T4 DNA polymerase. Addition of T4 accessory proteins 44/62 protein and 45 protein, along with rATP, stimulates the exonuclease of T4 polymerase (lane e), but inhibits the exonuclease of the T7 enzyme (lane f). Likewise the addition of either the T4 32 protein or the 32*I protein (lanes j through m) inhibits the T7 exonuclease. For 32 protein, this inhibition occurs even at the low concentrations that stimulate the exonuclease of the T4 DNA polymerase (lane l).

Fig. 7. Models for T4 DNA replication protein action at DNA ends.
(A). Exonuclease stimulation a low concentrations of 32 protein requires specific DNA polymerase: 32 protein interactions. Neither E. coli (SSB protein nor the proteolytic fragment of 32 protein, 32*I protein, can stimulate the exonuclease. These proteins do not physically interact with the polymerase whereas intact 32 protein binds to the T4 DNA polymerase. Presumably, the correct protein-protein interactions require the presence of the 60 amino acids at the COOH terminus of 32 protein, which are missing in the 32*I protein.

(B). Exonuclease inhibition at high concentrations of 32 protein is due to destabilization of the 3'OH DNA terminus. While 32 protein cannot melt the DNA duplex, we propose it can destabilize DNA ends sufficiently to impart a single-stranded character to the DNA ("fraying" of ends). The 32 protein is known to inhibit the activity of the polymerase-associated exonuclease on single-stranded DNA (Huang and Lehman, 1972).

(C). With polymerase accessory proteins present, exonuclease stimulation occurs at both low and high concentrations of 32 protein. In this view, the accessory proteins stabilize polymerase: DNA end interactions and prevent the "fraying" of the DNA end by the 32 protein.

FIGURE ONE

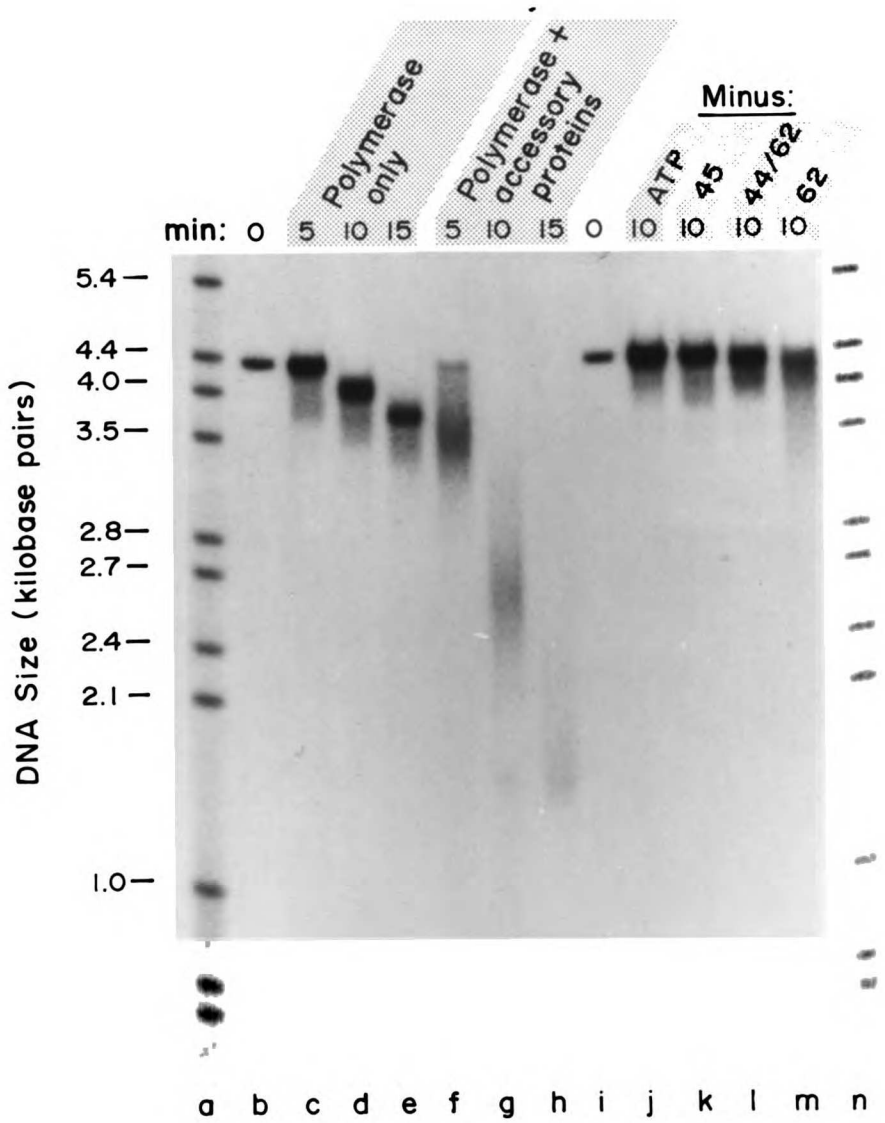

Polymerase only
Polymerase + accessory proteins
Minus:
ATP
45
44/62
62
min: 0 5 10 15 5 10 15 0 10 10 10 10
DNA Size (kilobase pairs)
5.4
4.4
4.0
3.5
2.8
2.7
2.4
2.1
1.0
a b c d e f g h i j k l m n

FIGURE TWO
A

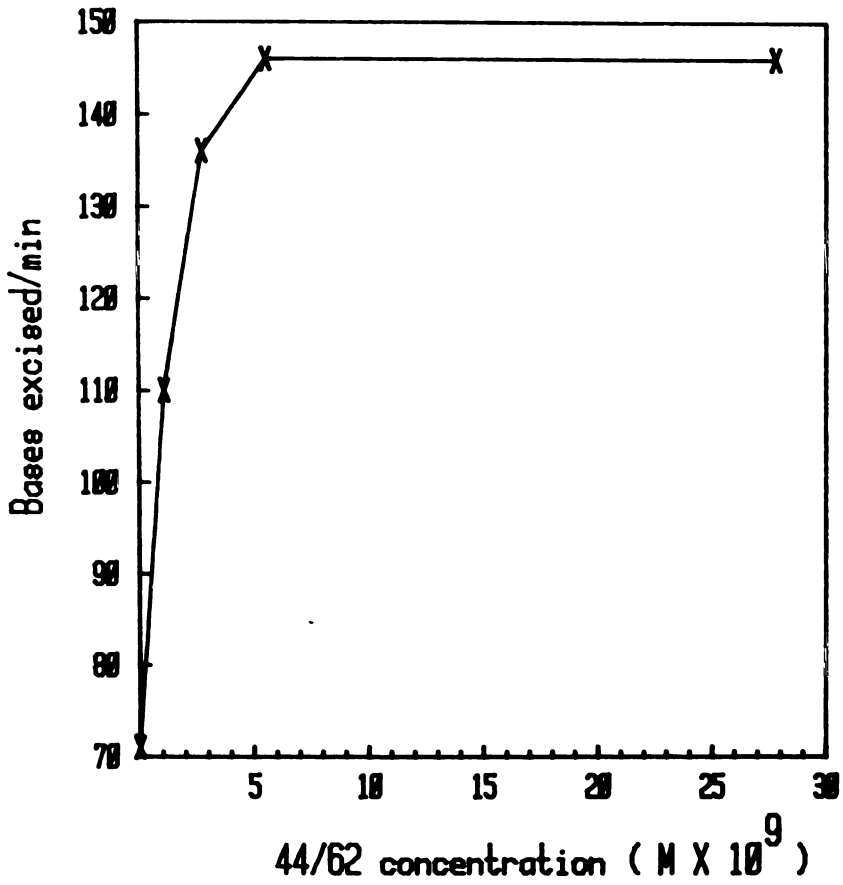

150
140
130
120
110
100
90
80
70
Bases excised/min
5
10
15
20
25
30
44/62 concentration (M X 10^9)

FIGURE TWO
B

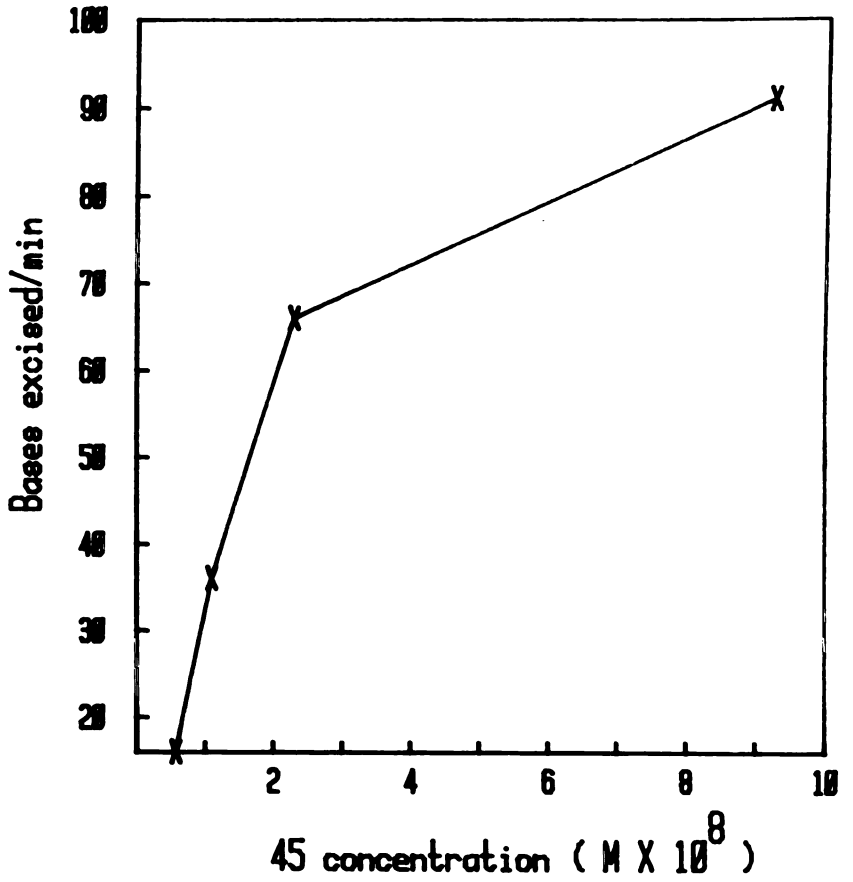


FIGURE THREE

FIGURE FOUR

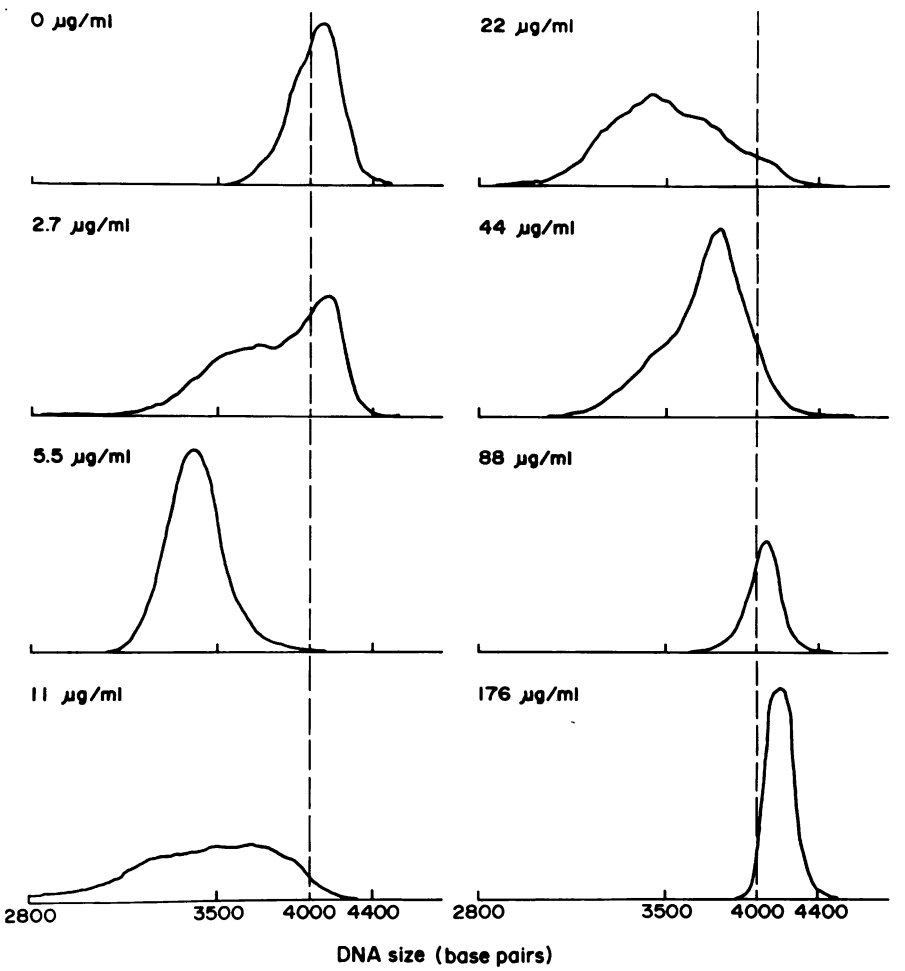

0 µg/ml
22 µg/ml
2.7 µg/ml
44 µg/ml
5.5 µg/ml
88 µg/ml
11 µg/ml
176 µg/ml
2800
3500
4000
4400
DNA size (base pairs)

FIGURE FIVE

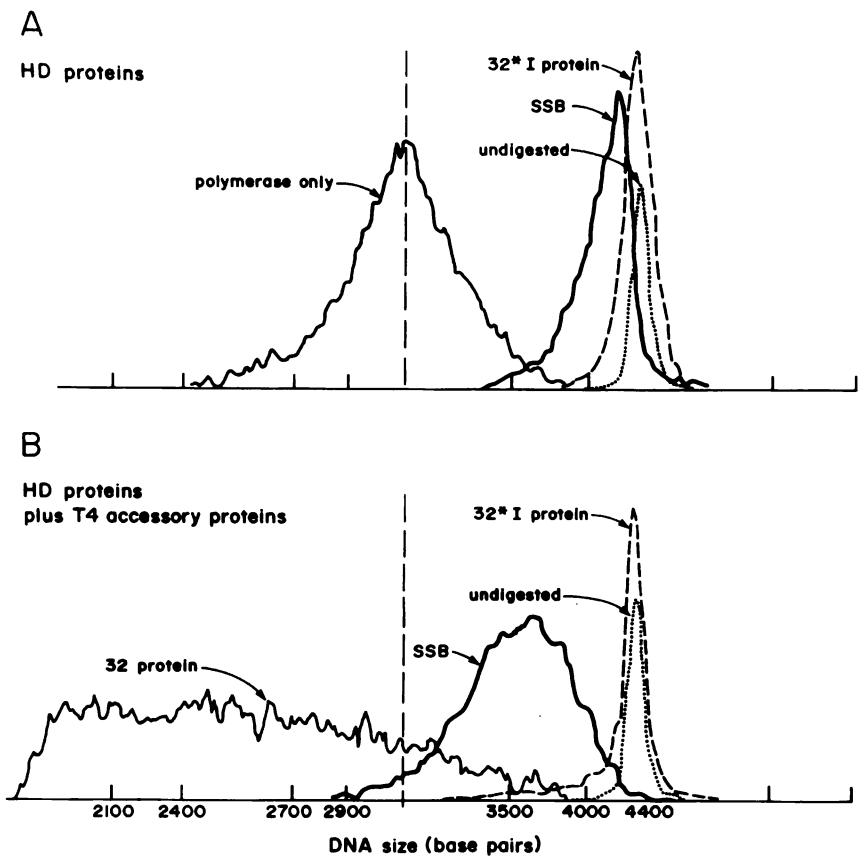

A
HD proteins
32* I protein
SSB
undigested
polymerase only
B
HD proteins
plus T4 accessory proteins
32* I protein
undigested
SSB
32 protein
2100
2400
2700
2900
3500
4000
4400
DNA size (base pairs)

FIGURE SIX

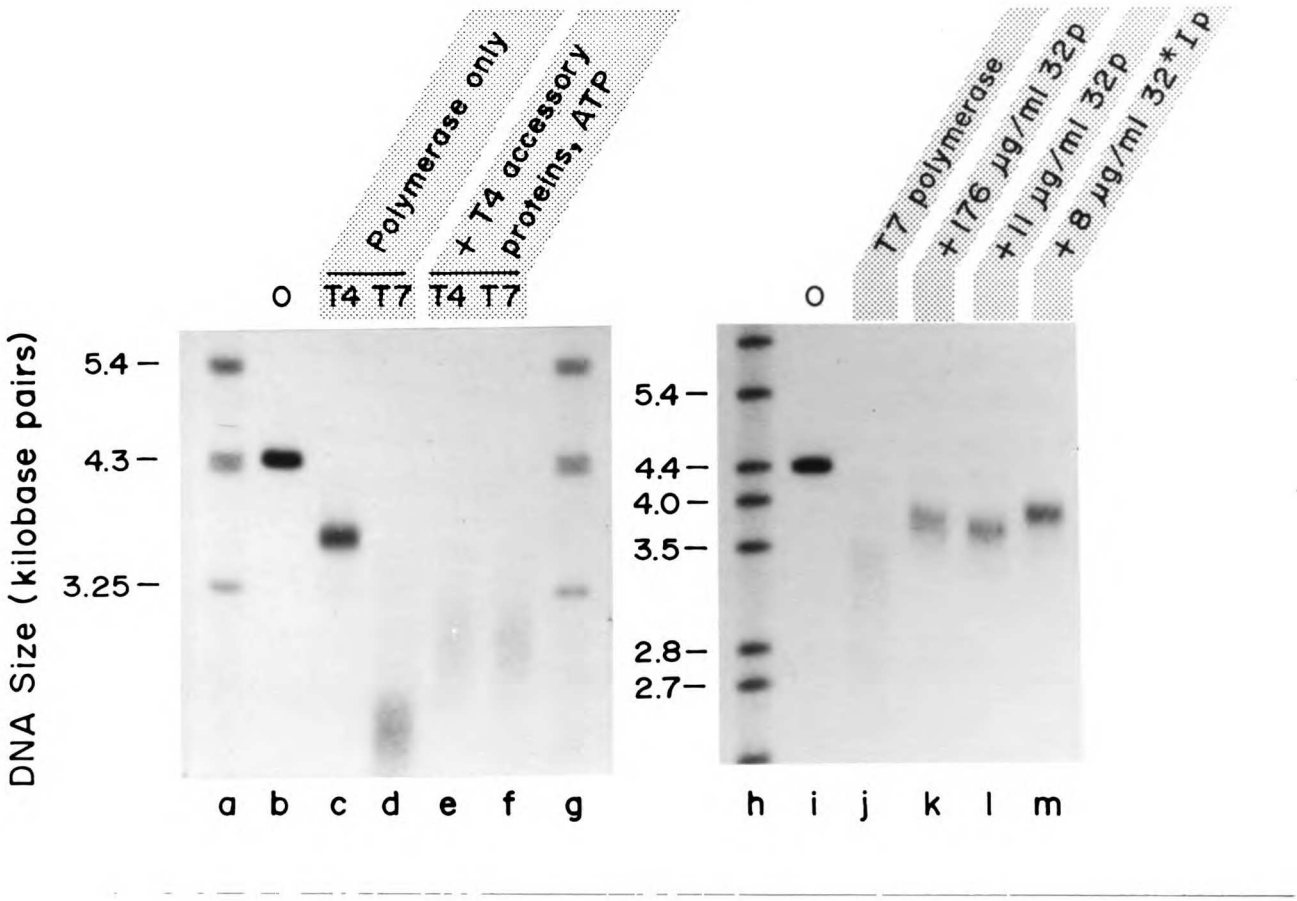

DNA Size (kilobase pairs)
5.4
4.3
3.25
0
T4 T7
Polymerase only
T4 T7
+ T4 accessory proteins, ATP
a b c d e f g
0
T7 polymerase
+176 µg/ml 32p
+11 µg/ml 32p
+8 µg/ml 32*Ip
5.4
4.4
4.0
3.5
2.8
2.7
h i j k l m

FIGURE SEVEN

ADDENDUM TO CHAPTER ONE

STIMULATION OF THE 3'-5' EXONUCLEASE OF T4 DNA POLYMERASE BY T4 DNA HELICASE dda PROTEIN

The effects of the T4 DNA helicase coded for by the dda gene on the T4 polymerase-associated exonuclease - Recently another T4 protein, the product of the dda gne (DNA-dependent ATPase gene), has been purified to near homogeneity and characterized in our laboratory (C.V. Jongeneel and B.M. Alberts, manuscript in preparation). This protein was previously shown by others to be a DNA helicase that moves in the 3' to 5' direction along a DNA single-strand melting an adjacent duplex DNA molecule utilizing the energy of ATP hydrolysis (Krell et al., 1979; C.V.J., personal communication).

When 3 1/m2/g/ml of the purified dda protein, which itself contains no exonuclease activity , was added to the polymerase-associated exonuclease reactions (only DNA-polymerase present), we observed a substantial increase in the rate of DNA digestion. This stimulation of the exonuclease, illustrated in the autoradiograph tracings shown in Fig. 1, is dependent on the presence of ATP in the reaction.

A possible model for the mechanism leading to this stimulation of exonuclease activity is shown in Fig. 2. The polymerase-associated exonuclease is much more active on a single-stranded 3'OH DNA end than on a base-paired 3'OH DNA end; this allows the exonuclease to "proofread" whether a newly incorporated deoxyribonucleotide is correct (base-paired) or incorrect (not base-paired) during replication and thereby to excise it preferentially (Brutlag and Kornberg, 1972). As shown in Fig. 2, the DNA helicase activity of the dda protein is likely to be generating a single-stranded 3'OH DNA end as a substrate for the polymerase-associated exonuclease in our system.

It should be noted that another T4 protein, thought to function as a DNA helicase, the T4 gene 41 protein (Liu and Alberts, 1981), has no effect on the activity of the polymerase-associated exonuclease in our system (data not shown). It is possible that the DNA substrate in our system is not appropriate for allowing 41 protein to function as a helicase; for instance, the polarity of the 41 protein helicase may be the 5'-3', or the 41 protein helicase may require a partially melted duplex DNA end in order to act.

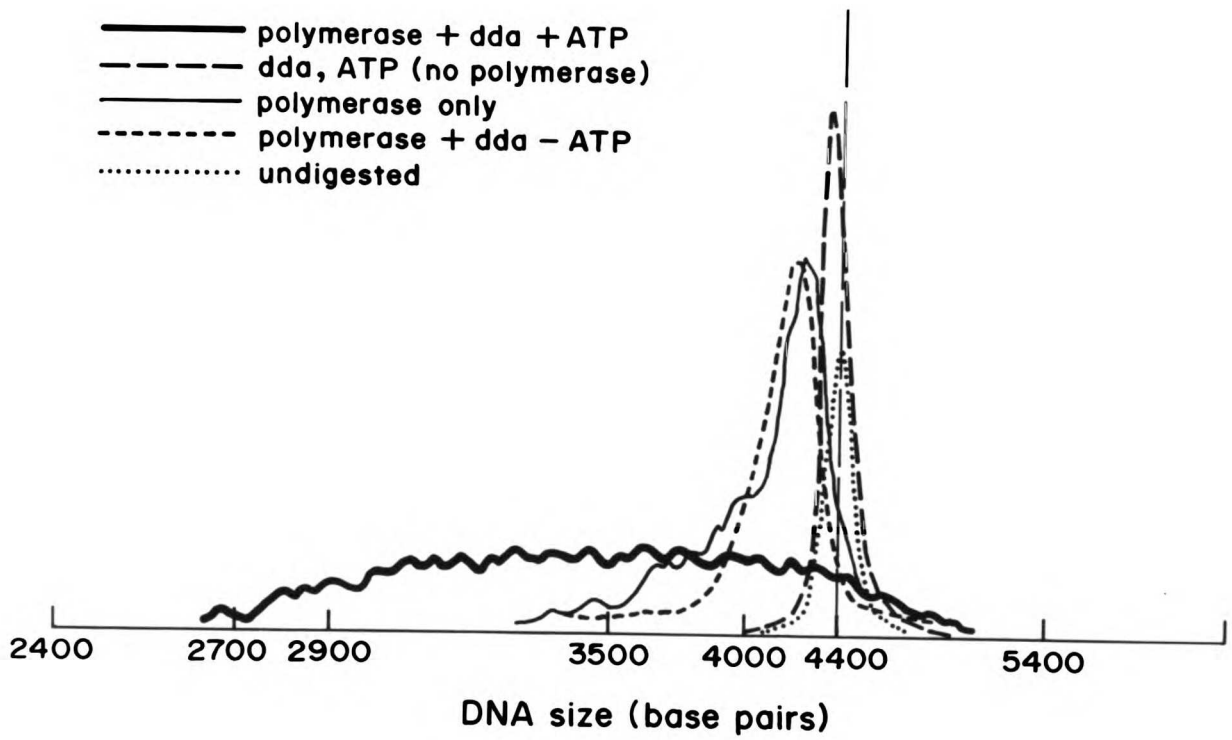
REFERENCES

Brutlag, D. and Kornberg, A. (1972) J. Biol. Chem. 247, 241-248.

Krell, H., Durwald, H. and Hoffmann-Berling, H. (1979) Eur. J. Biochem. 93, 387-395.

Liu, C.C., and Alberts, B.M. (1981) J. Biol. Chem. 256, 2813-2820.

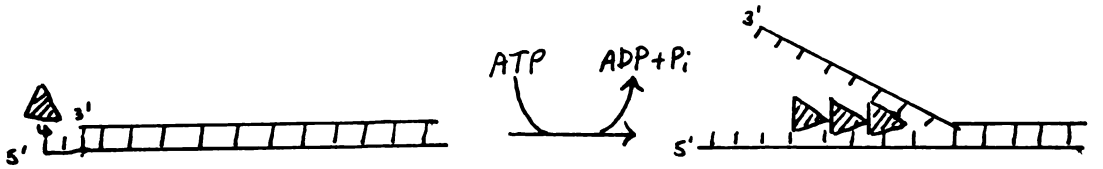
ADDENDUM FIGURE ONE



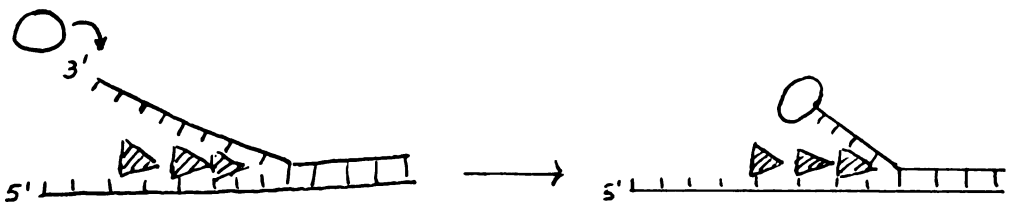
ADDENDUM FIGURE TWO

Possible Mechanism for the Stimulation of the 3'-5' Exonuclease of the T4 DNA Polymerase by T4 dda Protein

The T4 dda protein unwinds a DNA duplex unidirectionally in a reaction requiring ATP hydrolysis, acting as a DNA helicase



This dda protein-catalyzed melting of the double-stranded DNA provides a single-stranded 3'-OH DNA end as a substrate for the polymerase-associated exonuclease, which is 10-fold more active on single-stranded DNA than on double-stranded DNA



=T4 DNA polymerase

=T4 dda protein

CHAPTER TWO

SPECIFIC SEQUENCES ACT AS PAUSE SITES DURING THE REPLICATION OF DOUBLE-STRANDED DNA CATALYZED BY PURIFIED T4 REPLICATON PROTEINS

ABSTRACT

We have been able to detect a site-specific pausing of the replication fork formed *in vitro* by the bacteriophage T4 core replication system on a double-stranded DNA template, by using a specifically nicked replicative form (RF) fd DNA as the template in order to synchronize DNA synthesis. The two strongest pause sites correlate with regions of hairpin structures predicted to form if the DNA were single-stranded. Other pause sites, including two that become prominent at low 32 protein concentrations, are in regions that are not obviously involved in secondary structure. The addition of the T4 gene 41 protein (helicase-primase) to the replication system greatly increases the rate of fork movement and eliminates detectable pausing. In contrast, the addition of the T4 dda protein, another DNA helicase, increases the rate of fork movement to a similar extent without affecting replication fork pausing. The natural terminator sequence of the plasmid R6K, which functions as a general replication terminantor in *E. coli* is not recognized by T4 replication fork.

While the initiation of DNA replication at specific "origin" sequences has been the subject of intense research in recent years, the phenomena of DNA replication fork pausing and termination have been largely neglected. Through the use of deletion mutants it has been demonstrated that neither bacteriophage λ nor the mammalian virus SV40 require specific DNA sequences for replication termination, which occurs where the two replication forks meet 180° from the bidirectionally elongated origin (Lai and Nathans, 1975; Valenzuela, Freifelder and Inman, 1976). Recently, however, specific replication termination sequences have been identified in E. coli (Kuempel and Duerr, 1979), plasmid R6K (Kolter and Helinski, 1978), plasmid ColE1 (Tomizawa, 1978), and mammalian mitochondrial DNA (Doda et al., 1981). The mechanism of replication termination at these sites in vivo is not currently understood. We have investigated replication fork pausing utilizing the well-characterized bacteriophage T4 in vitro DNA replication system with a natural, double-stranded DNA template, the circular replicative form (RF) of bacteriophage fd. The T4 core-41-61 replication system, consisting of seven highly purified (›90% pure) proteins, closely mimics in vivo replication in terms of substrate utilization (Nossal and Peterlin, 1979; Sinha et al. 1980), fidelity (Hibner and Alberts, 1980; Sinha and Haimes, 1981), RNA primer synthesis (Liu and Alberts, 1980) and rate of fork movement (Alberts et al. 1980; Barry and Alberts, in preparation).

We report here that the multienzyme T4 replication apparatus formed by this system pauses at specific DNA sequences as the replication fork moves.

MATERIALS AND METHODS

Enzymes - Bacteriophage fd gene 2 protein was the generous gift of T. Meyer and K. Geider, purified as described (Meyer and Geider, 1979). The T4 DNA replication proteins corresponding to genes 32, 44/62, 45, 41 and 43 (T4 DNA polymerase) were purified using published procedures (Bittner et al. 1979; Morris et al. 1979a, Morris et al. 1979b). The T4 dda protein (DNA-dependent ATPase) was purified in our laboratory by C. Victor Jongeneel according to a new protocol (manuscript in preparation). All of these preparations were nearly homogeneous and free of detectable nuclease contaminations. Restriction enzymes were purchased from New England Biolabs.

DNA - Double-stranded supercoiled replicative form (RF) fd DNA was purified from fd-infected E. coli cells using the method of Clewell (1972). To obtain specifically nicked fd RF DNA, fd gene 2 protein was incubated with 2 µg fd RF DNA in 16 mM Tris acetate pH 7.8, 33 mM potassium acetate, 5 mM magnesium acetate, 0.5 mM dithiothreitol, and 10% glycerol at 30°C for 45 min. The gene 2 protein was then inactivated by heating the reaction to 65°C for 10 min. Analysis of the products of this reaction by agarose gel electrophoresis showed that more than 80% of the fd DNA was nicked in the reaction. The replicative forms of the cloning vector Mp9 DNA and of Mp9 DNA containing an insert of the plasmid R6K replication terminator site (Bastia et al. 1981a) were kindly supplied by Dr Depak Bastia; both of these DNAs contain the gene 2 nicking site.

Replication Reactions - Unless stated otherwise, in vitro DNA replication was carried out in the presence of 33 mM Tris-acetate, pH 7.8, 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM dithiothreitol, 0.5 mM rATP, 0.1 mM each dATP, dGTP, dCTP, and $\alpha[^{32}P]$-TTP; and 2 1/m2/g/ml DNA. In "core replication system" reactions, the following T4 DNA replication proteins were present at the indicated concentrations: T4 DNA polymerase, 2.5 µg/ml; T4 gene 32 protein, 80 to 100 µg/ml, T4 gene 44/62 protein, 20 µg/ml; and T4 gene 45 protein, 18 µg/ml. For "core-41 replication system" reactions, the T4 gene 41 protein 6 µg/ml was also added. Where indicated, the T4 dda protein was added to reactions to a final concentration of 3 µg/ml.

Replication reactions were started synchronously at a specific nick in the fd DNA template, as outlined in Fig. 1. Reaction mixtures lacking only replication proteins and dCTP were prepared at 4°C. All of the proteins were added and the mixture incubated at 37°C for 1 min; dCTP (0.1 mM) was then added to allow replication forks to proceed. At various intervals, aliquots were removed from the reaction into tubes containing sodium dodecyl sulfate (SDS) to produce a final concentration of 3% SDS. After 10 min at room temperature, each aliquot was "spin dialyzed" to remove unincorporated dNTPsthrough sepharose CL6B (Pharmacia) as described by Neal and Florini (1973). Buffer was then added to each aliquot to produce a final concentration of 20 mM Na_3EDTA, 10% sucrose and 0.1% bromocresol green dye and the DNA analyzed by electrophoresis through a 0.6% agarose gel, using 30 mM NaOH and 2 mM Na_3EDTA as the running buffer. Gels measuring 0.4 x 14 x 21.5 cm were electro-

phoresed at 30 volts for 40 hr. The gels were then dried onto Whatman 3 MM filter paper and autoradiographed at -70°C using Kodak XR-2 film with a Dupont "Lightning-Plus" intensifying screen. The sizes of the radioactively-labeled, newly synthesized DNA strands were determined by comparison with ^{32}P-labeled restriction fragments of bacteriophage T4 DNA of known size (O'Farrell et al. 1980).

RESULTS

Replication forks pause at specific sites on a double-helical DNA template - In vitro DNA replication reactions catalyzed by the T4 core replication system include T4 gene 32, 45, and 44/62 proteins in addition to T4 DNA polymerase (the gene 43 product). Each of these proteins is required to achieve efficient DNA synthesis starting from a nick on a double-stranded DNA template (Nossal and Peterlin, 1979, Sinha et al., 1980). The replication process begins by the covalent addition of deoxyribonucleotides to the 3'OH terminus at the DNA nick; the strand containing the 5' phosphate at the nick is displaced ahead of the growing DNA chain during this reaction, producing long single-stranded DNA "tails." On a circular DNA template, the length of these tails can greatly exceed the length of the parental circular DNA, since replication proceeds in a "rolling circle" mode that allows many rounds of copying.

In the experiments to be described, we have used the double-stranded replicative form (RF) of the bacteriophage fd genome as a template for replication. We treated this DNA with purified fd gene 2 protein in order to create a uniquely nicked DNA template (nicked at nucleotide 5781; Meyer et al. 1979). Previous work has shown that this DNA is an effective template for the T4 in vitro replication system, and that replication proceeds in a rolling circle mode (Meyer et al. 1981). In an attempt to synchronize the starts on this template, the procedure shown in Fig. 1 was adopted. At various intervals after starting DNA synthesis, the lengths of the growing DNA strands were measured by agarose gel electrophoresis in alkali a

shown in Fig. 2A. It can be seen that discrete DNA product sizes are observed, representing unique sites where synthesis had stopped. The smaller length molecules are "chased" into molecules of greater length at the longer incubation times, demonstrating that the discrete DNA bands observed represent replication pause sites; i.e., sequences where the replication fork has been temporarily arrested. An analysis of the DNA band sizes reveals that the replication fork pauses at the same DNA sequences during each round of rolling circle replication (Table I). For example, a strong pause site that is located just "upstream" from the gene 2 cutting site appears first produces a DNA band at about 12,600 nucleotides, then again at about 19,000 nucleotides (12,600 nucleotides, and plus 6408 nucleotides, the unit length of the fd genome) when the replication fork encounters this sequence for the second time.

Lowering the concentration of the T4 helix-destabilizing protein (gene 32 protein) from 100 μg/ml to 20 μg/ml decreases the rate of the replication fork formed by the core system about 3 to 4-fold, as previously reported (Alberts et al., 1980). In addition to slowing the net rate of fork movement, the lowering of 32 protein concentration enhances the pausing of the replication fork at specific sites, as shown in Fig. 2B (see arrows).

The initiation of replication forks at a specific site is required to detect discrete pauses - If the replication fork described above is pausing at a restricted set of DNA sequences, rather than pausing after fixed time intervals of synthesis, a specific initiation site should be required to detect pauses by our methods. As a test, replication

was initiated at the random nicks that are present at a level of about 10% in our fd RF DNA preparation, rather than at the specific gene protein nicking site. Fig. 3 displays densitometer tracings of autoradiographs that compare the length distribution of reaction products of two replication reactions; one initiated at these random nicks and the other initiated from the specific nick used previously. It can be seen that the broad distribution of DNA sizes in the randomly-initiated reaction becomes sharply punctuated with defined size classes of DNA products when a specific start site is used. Presumably the randomly-initiated replication fork still pauses at the same specific sites, but the size of the resulting DNA products is now heterogeneous.

The nature of the replication pause sites - Previous studies have revealed that sites of secondary structure in a single-stranded DNA template can act as kinetic blocks to in vitro DNA synthesis catalyzed by the T4 DNA replication proteins (Huang and Hearst, 1980; Huang et al., 1981; Roth et al. 1982). Fig. 4 shows the approximate location of replication pause sites on a double-stranded fd DNA template. We have found that the two strongest pause sites (A and B on Fig. 4) correlate with the strong pauses seen on a single-stranded fd DNA template at sites of hairpin-like secondary structures in the DNA. However, the core replication system also consistently paused (albeit more weakly) at sites not thought to be involved in secondary strucutre (sites C-F, Fig. 4). The definitive characteristics of these sequences that cause the replication fork to pause are currently not understood.

Recently the specific replication termination site in the plasmid R6K has been cloned and shown to function as a terminator site during the in vivo replication of several other DNA molecules, a cloned sequence as short as 216 base pairs retaining functional terminator activity (Kolter and Helinski, 1978; Bastia et al. 1981a). We tested whether this terminator sequence would exert an effect on the above in vitro DNA replication system by replicating a specifically nicked DNA template that contained an inserted R6K terminator sequence. The result was unambiguous: the R6K replication termination sequence does not cause the T4 replication fork to pause (data not shown).

If the stalling of the replication fork that we observe in vitro is caused by an obstructive secondary structure in single-stranded DNA produced ahead of the leading strand at the fork (see Discussion below), the failure of the R6K terminator to function is not surprising. The cloned terminator region does not contain any obvious palendromic sequences that could potentially form stable hairpin structures (Bastia et al., 1981b). It is currently thought that the R6K replication terminator sequence functions in conjunction with unidentified regulatory proteins that specifically recognize and bind to the terminator region (Germino and Bastia, 1981).

Effects of the addition of other T4 replication proteins on replication fork pausing - Two additional T4 proteins that function in DNA replication in vivo have been shown to influence the rate of replication fork movement in in vitro reactions; the T4 gene 41 protein and the T4 dda protein. Using the same assay described in

Fig. 1, we have examined the effect of these proteins on the replication fork pausing, as shown in Figs. 5 and 6.

The T4 gene 41 protein is a DNA-dependent GTPase (and ATPase) that seems to function as a dimer with a molecular weight of 116,000 (Liu and Alberts, 1981). The addition of this protein to the T4 core replication system leads to a dramatic increase in the rate of fork movement on a double-stranded DNA template (Alberts et al. 1980; Barry and Alberts, manuscript in preparation). Since continuous GTP hydrolysis is required for this effect, the stimulation of fork rate is believed to reflect the action of the 41 protein as a DNA helicase that uses the energy of GTP hydrolysis to run along the lagging strand template and melt the duplex DNA ahead of the fork (Liu and Alberts, 1981).

As shown in the autoradiogram in Fig. 5B, an analysis of nascent DNA lengths in reactions including 41 protein reveals a uniform smear of long DNA product lengths from those forks containing 41 protein, very unlike the discrete products seen when the 41 protein is omitted (Fig. 5A). This result indicates that the addition of the T4 gene 41 protein to our reactions eliminates (or greatly reduces) the pausing of the replication fork seen previously. The result in Fig. 5B is complicated by the fact that reactions catalyzed by the core-41 system always contain a mixture of replication forks with and without 41 protein; however, the products from these two types of forks can be readily distinguished by their very different lengths, as described in the Fig. 5 legend. Thus, the removal of kinetic barriers in front of the replication fork by the T4 gene 41 protein can be most clearly seen by comparing the size distributions

of the long DNA products observed after a brief incubation with 41 protein present with the products of the same size produced at later times in the slower reaction catalyzed by the core system without 41 protein. (Compare the DNA in the size range of 6 to 20 kilobases in Fig.5A and B).

The T4 dda (DNA-dependent ATPase) protein has been shown to be a DNA helicase that uses ATP hydrolysis energy to drive a DNA duplex melting reaction while moving in the 3' to 5' direction along an adjacent DNA singgle-strand (Krell et al., 1979). This protein has recently been purified and further characterized in our laboratory (Jongeneel and Alberts, manuscript in preparation). Similar to the effect obtained by addition of the gene 41 protein, the addition of the purified T4 dda protein to the core replication system results in a 4-fold increase in the rate of fork movement, as shown in Fig. 6. However, in contrast with the results of 41 protein addition, replication pause sites are detected in these faster forks that seem to be identical in strength and position with those found without dda protein (Fig. 6, lane c). Therefore, we conclude that the elimination of kinetic barriers to replication is a specific effect caused by the gene 41 protein, rather than being either a general property of DNA helicase action at the fork, or the result of a faster-moving DNA polymerase molecule per se (see Discussion).

DISCUSSION

The experiments reported here have shown that the T4 DNA replication apparatus pauses at specific sites during the *in vitro* replication of double-stranded DNA templates. Previous *in vitro* experiments have demonstrated that purified DNA polymerases pause at specific sites when replicating a single-stranded DNA template (Sherman and Gefter, 1976; Challberg and Englund, 1979; Huang and Hearst, 1980; Weaver and DePamphilis, 1982; Kaguni and Clayton, 1982). Our conditions should more closely resemble *in vivo* replication because a multi-enzyme replication system and a double-stranded DNA template have been employed.

Nature of the Pause Sites - Previous studies have shown that regions of hairpin-like secondary structure in a single-stranded DNA template can act as barriers to DNA synthesis catalysed by procaryotic and eucaryotic DNA polymerases: replication is abruptly arrested just before, or a few nucleotides into the base-paired "stem" of a downstream hairpin structure (Sherman and Gefter, 1976; Callberg and Englund, 1979; Huang and Hearst, 1980). More recent detailed studies with purified T4, DNA polymerase, *E. coli* DNA polymerase II and polymerase III holoenzyme, and polymerase α from either *Drosophila* embryos or mammalian tissue culture cells have revealed that other DNA sequences that are not involved in obvious secondary structures can also arrest DNA synthesis (Weaver and DePamphilis, 1982; Kaguni and Clayton, 1982). The abundance of G-C rich sequences at some (but not all) of these sites is suggestive,

particularly in light of the profound effects of G-C rich sequences during the movement of E. coli RNA polymerase (Gilbert, 1976).

In our experiments, the two strongest pause sites are in regions thought to be folded into hairpin helicies in single-stranded fd DNA (Fig. 4). Given that the template used in these experiments was double-stranded, it is somewhat surprising that the replication fork is able to detect these sites. The result suggests that the double-stranded DNA template is opened up far enough ahead of the growing DNA chain on the leading strand to allow these structures - with a 9 to 20 basepair hairpin stem - to form. Alternatively, it is possible that some recognizable structure can form within the double-stranded DNA template in such regions. Other sequences, that are seemingly not involved with the formation of secondary structure also caused the replication fork to pause (Fig. 4). Therefore, in agreement with previous studies, a unique mechanism for replication fork pausing is not apparent.

The effects of other T4 replication proteins on the character of fork movement - Addition of the T4 gene 41 protein to replication reactions catalyzed by the T4 core system increases the rate of fork movement up to 8-fold (Alberts et al. 1980; Barry and Alberts, in preparation). When we analyzed the lengths of the DNA products formed in reactions including gene 41 protein, we found that the rapidly moving replication forks are no longer detectably arrested (Fig. 5).

Addition of the T4 dda protein, a DNA helicase, to the T4 core replication system increases the rate of fork movement about 4-fold

under our conditions (Fig. 6). However, in contrast to the effect of 41 protein, the same pause sites are still recognized as in the absence of the dda protein.

Both the dda protein and the gene 41 protein appear to be DNA helicases that utilize energy from nucleoside triphosphate hydrolysis to move unidirectionally and melt duplex DNA, thereby speeding the progress of the replication fork. The different effects of these two enzymes on fork pausing indicate that despite their suspected common activities, they somehow act differently at the replication fork.

Possible importance of replication pausing - We have found that in vitro T4 replication forks that lack gene 41 protein pause at specific DNA sequences. Such "incomplete" replication forks could also exist in vivo. The strong pauses seen during in plasmid ColE1 and mammalian mitochondrial DNA replication could in fact be due to lack of particular protein factors in the replication forks involved in proceeding past replication barriers.

While the extent of specific replication fork pausing in vivo is far from clear, several potential functions for such events are conceivable. For example, the pausing of replication forks could serve to synchronize replication with some other cellular process in the same manner that RNA polymerase pausing after a specific amount of transcription is thought to allow binding of ribosomes to "leader" RNA: the subsequent synchrony of transcription and translation plays an important role in the gene expression of trp operon in E. coli (Yanofsky, 1981; Winkler and Yanofsky, 1982). A possible process that could be coordinated in a similar way during T4 DNA replication

would be the assembly and attachement of the multienzyme T4 nucleotide-producing complex to each newly-formed replication fork (Reddy et al., 1977).

Replication forks that have paused may also function by creating a site with a special conformation required for enzymatic activity. The fd gene 2 protein, for example, only nicks at the specific gene 2 recognition sequence if the DNA is either supercoiled or is being replicated (Meyer et al. 1981). In the same manner, the DNA at a paused replication fork may provide an appropriate substrate for genetic recombination enzymes or for topoisomerases.

Acknowledgments

The work described here formed part of a disseration submitted by PB to the Faculty of the Department of Biochemistry and Biophysics of the University of California, San Francisco in partial fulfillment of the requirement for the Ph.D. degree.

REFERENCES

Alberts, B.M., Barry, J., Bedinger, P., Burke, R.L., Hibner, U., Liu, C.-C., and Sheridan, R. (1980) In Mechanistic Studies of DNA Replication and Genetic Recombination, Vol. 19 ICN-UCLA Symposia on Molecular and Cellular Biology (Alberts, B.M., ed) pp. 449-471. Academic Press, New York.

Bastia, D., Germino, J., Crosa, J.H., and Ram, J. (1981) Proc. Natl. Acad. Sci. USA 78, 2095-2099.

Bastia, D., Germino, J., Crosa, J.H. and Hale, P. (1981) Gene 14, 81-89.

Bittner, M., Burke, R.L., and Alberts, B.M. (1979) J. Biol. Chem. 254, 9565-9572.

Challberg, M.D., and Englund, P.T (1979) J. Biol. Chem. 254, 7820-7826.

Clewell, D.B. (1972) J. Bacteriol. 110, 667-676.

Germino, J. and Bastia, D. (1981) Cell 23, 681-687.

Gilbert, W. (1976) In RNA Polymerase (Losick, R. and Chamberlin, M., eds) pp. 193-205. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. pp. 196-205.

Hibner, U. and Alberts, B.M. (1980) Nature 285, 300-305.

Huang, C.-C., and Hearst, J.E. (1980) Analytical Biochem. 103, 127-139.

Huang, C.-C., Hearst, J.E., and Alberts, B.M. (1981) J. Biol. Chem. 256, 4087-4094.

Kaguni, L.S. and Clayton, D.A. (1982) Proc. Natl. Acad. Sci. USA 79, 983-987.

Kolter, R., and Helinski, D.R. (1978) J. Mol. Biol. 124, 425-441.

Krell, H., Durwald, H., and Hoffmann-Berling, H. (1979) Eur. J. Biochem. 93, 387-395.

Kuempel, P.L., and Duerr, S.A. (1979) Cold Spring Harbor Symp. Quant. Biol. 43, 563-567.

Lai, C.-J., and Nathans, D. (1975) J. Mol. Biol. 97, 113-118.

Liu, C.-C., and Alberts, B.M. (1980) Proc. Natl. Acad. Sci. USA 77, 5698-5702.

Liu, C.-C., and Alberts, B.M. (1981) J. Biol. Chem. 256, 2813-2820.

Meyer, T.F., and Geider, K. (1979) J. Biol. Chem. 254, 12636-12641.

Meyer, T.F., Geider, K., Kurz, C., and Schaller, H. (1979) Nature 278, 365-367.

Meyer, T.F., Baumel, I., Geider, K., and Bedinger, P. (1981) J. Biol. Chem., 256, 5810-5813.

Morris, C.F., Moran, L.A., and Alberts, B.M. (1979) J. Biol. Chem. 254, 6797-6802.

Morris, C.F., Hama-Inaba, H., Mace, D., Sinha, N.K., and Alberts, B.M. (1979) J. Biol. Chem. 254, 6787-6796.

Nossal, N.G., and Peterlin, B.M. (1979) J. Biol. Chem. 254, 6032-6037.

O'Farrell, P.H., Kutter, E., and Nakanishi, M. (1980) Molec. gen. Genet. 179, 421-435.

Reddy, G.P.V., Singh, A., Stafford, M.E., Mathews, C.K. (1977) Proc. Natl. Acad. Sci. USA 74, 3152-3156.

Roth, A.C., Nossal, N.G., and Englund, P.T. (1982) J. Biol. Chem. 257, 1267-1273.

Sherman, L.A., and Gefter, M.L. (1976) J. Mol. Biol. 103, 61-76.

Sinha, N.K., Morris, C.F., and Alberts, B.M. (1980) J. Biol. Chem. 255, 4290-4303.

Tomizawa, J.I. (1978) In DNA Synthesis: Present and Future (Molineux, I. and Kohiyama, M., ed) pp. 797-825, Plenum Press, New York.

Valenzuela, M.S., Freifelder, D., and Inman, R.B. (1976) J. Mol. Biol. 102, 569-589.

Weaver, D.T., and DePamphilis, M.L. (1982) J. Biol. Chem. 257, 2075-2086.

Winkler, M.E., and Yanofsky, C. (1981) Biochemistry 20, 3738-3744.

Yanofsky, C. (1981) Nature 289, 751-758.

Table I. The approximate locations of replication pause sites on double-stranded fd DNA.

Sizes of nascent DNA molecules were assigned by comparison with restriction fragments of cytosine-containing T4 DNA of known size. Map locations were calculated by subtracting the length of a fd monomer of 6408 bases (or dimer of 12,816 bases after a second round of rolling circle replication) from the size of each DNA product observed. Data is averaged from eight experiments.

Product size in nucleotides	Approximate map site	Letter designation on Fig. 4 map
6900 ± 100	6270	F
8150 ± 50	1120	E
8350 ± 50	1320	D
9500 ± 100	2470	C
10400 ± 100	3370	B
12600 ± 50	5570	A
13250 ± 50	6230	F
14500 ± 50	1070	E
16600 ± 100	3170	B
18800 ± 100	5370	A

LEGENDS TO THE FIGURES

Fig. 1. An assay that detects specific replication pause sites. The conditions used are described in Materials and Methods. Supercoiled fd RF was nicked at nucleotide 5781 by the fd gene 2 protein. Addition of T4 replication proteins with dATP, dGTP and ^{32}P(TTP) allows the insertion of 12 bases at the nick site as indicated. The delayed addition of the missing deoxyribonucleotide, dCTP, allows further synthesis to occur in a near synchronous manner, and replication proceeds in a rolling circle mode. After unincorporated radioactive nucleotides are removed, the products of synthesis are analyzed by alkaline agarose gel electrophoresis, as described in Materials and Methods.

Fig. 2. Rolling circle replication catalyzed by the T4 core replication system pauses at specific DNA sites. Reactions were performed as described in Fig. 1 and Materials and Methods. Aliquots of 10 µl were processed and electrophoresed as described. Lanes a and b show an autoradiograph of a gel analyzing DNA products of a standard reaction after 5, 10 and 20 min (32 protein concentration of 100 µg/ml). Lanes c and d show the results of changing the concentration of gene 32 protein on the products of the reactions.

Fig. 3. Replication pause sites are detectable only when replication is initiated at a specific site. This figure presents a microdensitometer tracing of autoradiographs similar to those shown in

Fig. 2, and compares products of reactions beginning at the random versus a specific nick on the same fd DNA template.

Fig. 4. Approximate location of replication pause sites on RF fd DNA. Sites of pausing were located by measuring the sizes of nascent DNA chains. The strongest pause sites, A and B, are in regions that are thought to contain secondary structure in a single-stranded DNA molecule, while the other pause sites (C-F) do not correlate with regions of secondary structure. Sites E and F are particularly prominent in reactions with low 32 protein concentrations (see Fig. 2).

Fig. 5. The T4 gene 41 protein eliminates pausing of the replication fork. Products of replication reactions catalyzed by the core replication system without (lanes a to c) and with (lanes d to f) T4 gene 41 protein are compared. Gene 41 protein greatly increases the rate of fork movement, so in order to compare DNA products of equivalent size, the core replication reaction ("5-protein reaction") included a high concentration (300 μg/ml) of 32 protein (the rate of 5-protein fork movement increases with increasing 32 protein concentration in this system). Additionally, aliquots from the core replication reaction were taken at later times than those from the core 41 replication reaction (5, 10, 15 min versus 1, 2, 5 min). The latter reactions ("6 protein reaction") include some slower-moving DNA products, made on forks that lack 41 protein; these display characteristic pausing, as indicated.

Fig. 6. The T4 dda protein increases the rate of fork movement, but does not eliminate replication fork pausing. Addition of purified T4 dda protein to catalyzed by the core replication system reactions, increases the rate of fork movement by about 4-fold under our conditions (compare lanes a and b with lanes c and d). Products of core "5 protein" reactions after 8 min and reactions including dda after 2 min indicate similar, if not identical, pausing patterns.

FIGURE ONE

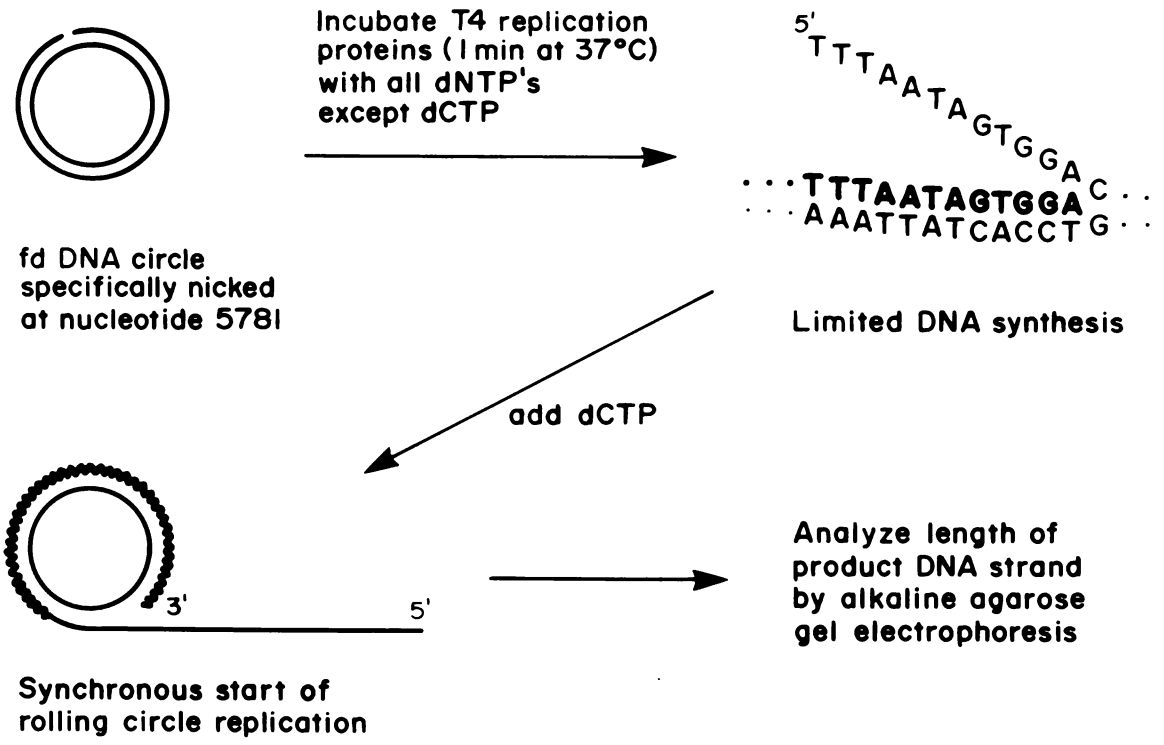

Incubate T4 replication proteins (1 min at 37°C) with all dNTP's except dCTP
5'
TTTAATAGTGGAC
AAATTATCACCTG
fd DNA circle specifically nicked at nucleotide 5781
Limited DNA synthesis
add dCTP
3'
5'
Analyze length of product DNA strand by alkaline agarose gel electrophoresis
Synchronous start of rolling circle replication

FIGURE TWO

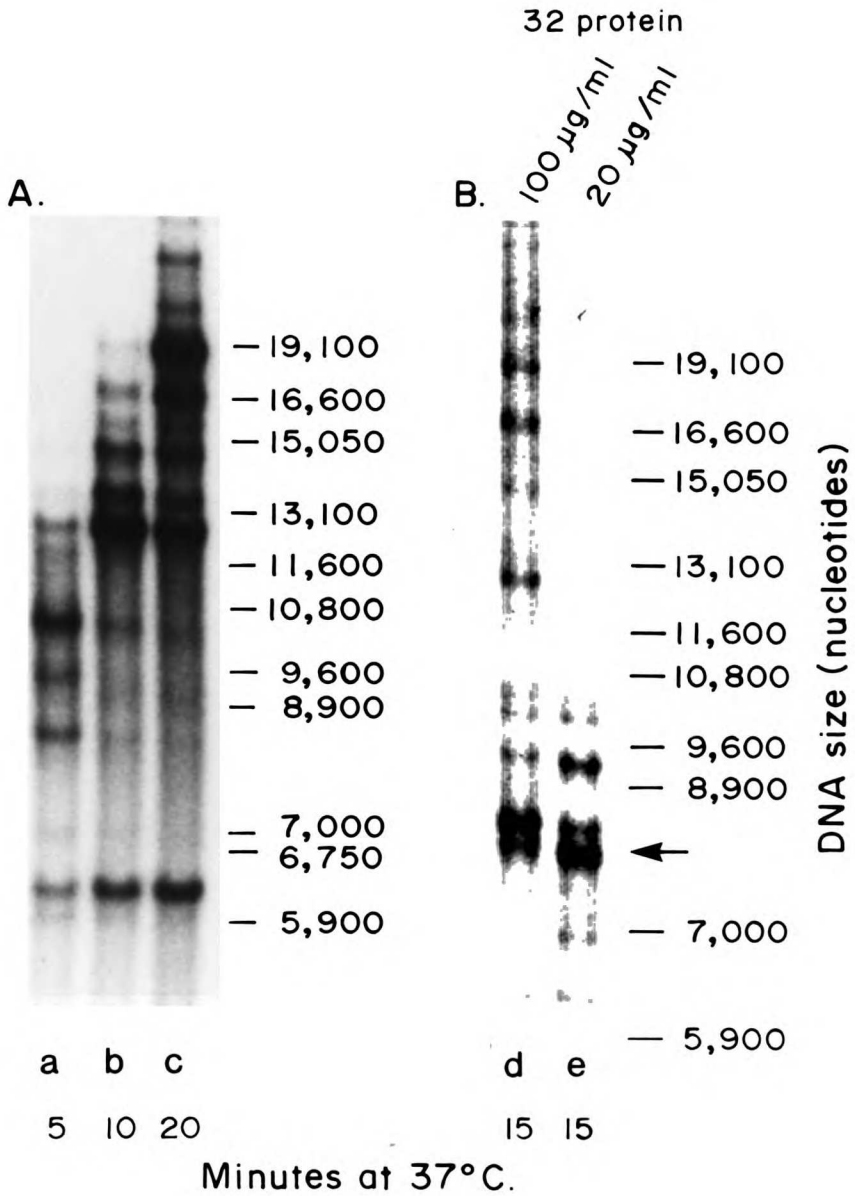

32 protein
100 µg/ml
20 µg/ml
A.
B.
19,100
16,600
15,050
13,100
11,600
10,800
9,600
8,900
7,000
6,750
5,900
DNA size (nucleotides)
a b c
d e
5 10 20
15 15
Minutes at 37°C.

FIGURE THREE

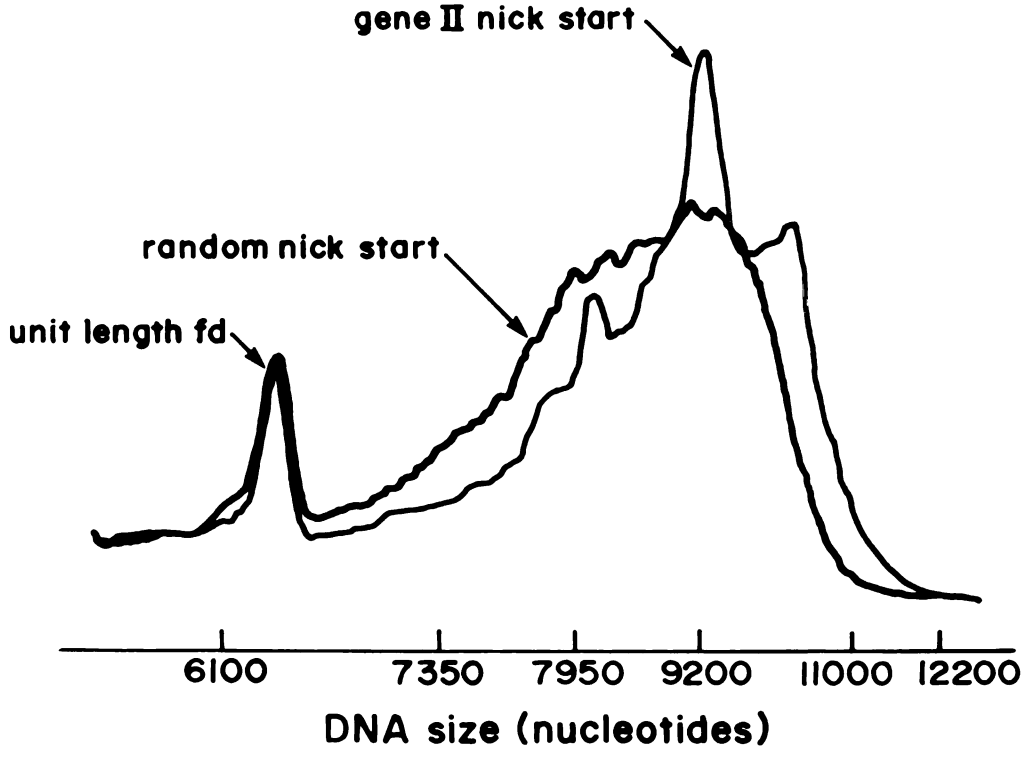


FIGURE FOUR

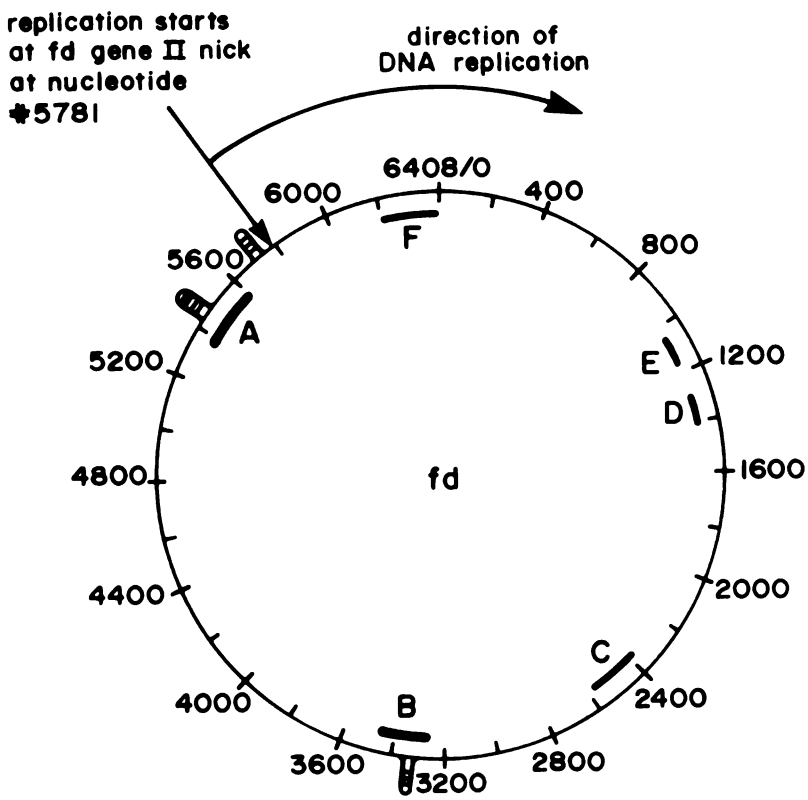

replication starts
at fd gene II nick
at nucleotide
#5781
direction of
DNA replication
6408/0
400
800
1200
1600
2000
2400
2800
3200
3600
4000
4400
4800
5200
5600
6000
A
B
C
D
E
F
fd

FIGURE FIVE

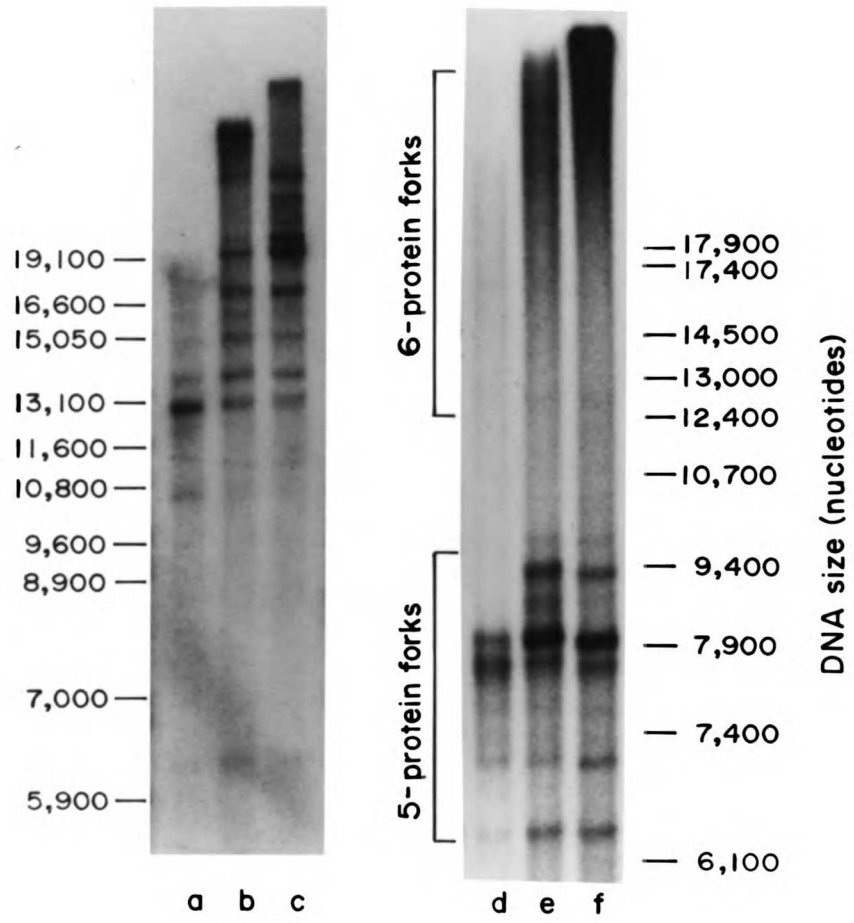
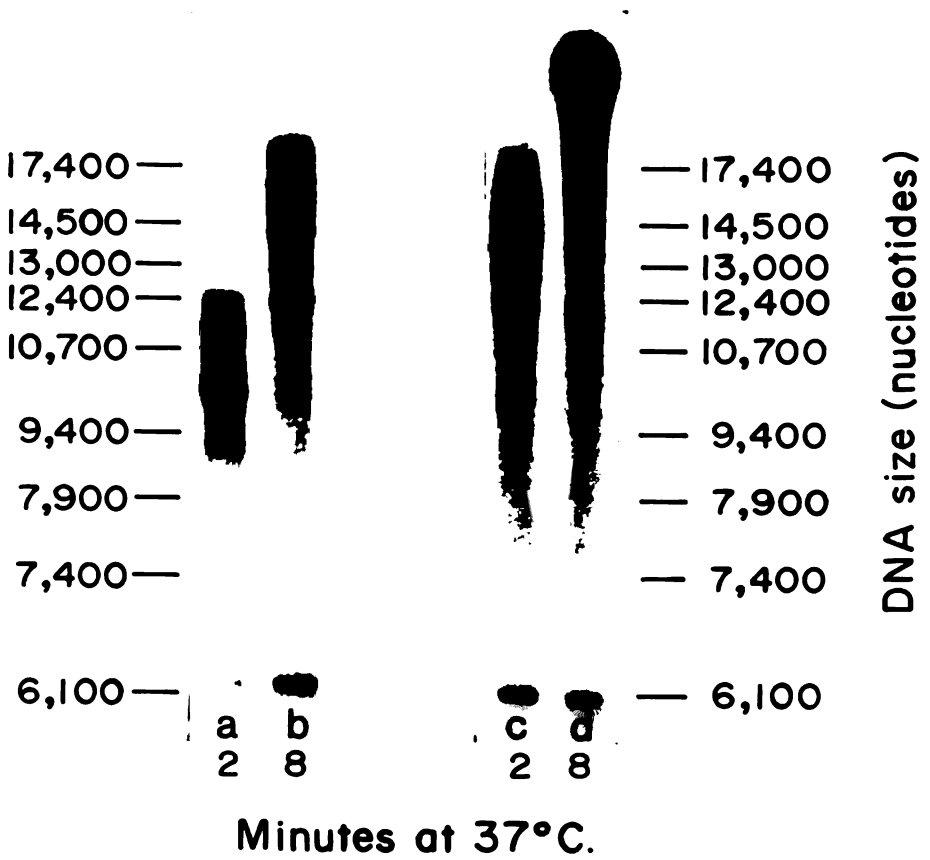

19,100
16,600
15,050
13,100
11,600
10,800
9,600
8,900
7,000
5,900
a b c
6-protein forks
5-protein forks
17,900
17,400
14,500
13,000
12,400
10,700
9,400
7,900
7,400
6,100
d e f
DNA size (nucleotides)

FIGURE SIX


17,400
14,500
13,000
12,400
10,700
9,400
7,900
7,400
6,100
a
b
c
d
2
8
2
8
Minutes at 37°C.
DNA size (nucleotides)

CHAPTER THREE

PROPERTIES OF THE T4 BACTERIOPHAGE DNA REPLICATION APPARATUS: A SINGLE, BOUND RNA POLYMERASE MOLECULE CAN BLOCK REPLICATION FORK MOVEMENT

ABSTRACT

The interaction of DNA replication forks with both stationary and transcribing RNA polymerase molecules has been examined in vitro, using the multienzyme T4 DNA replication system (Alberts et al. 1980) and purified E. coli RNA polymerase. We have found that a single stationary RNA polymerase molecule can block the movement of the T4 replication fork when bound to a promoter on a double-stranded fd DNA template. When transcription is allowed (in the same direction as replication), the replication fork appears to follow the moving RNA polymerase molecule at the relatively slow rate of transcription. The replication barriers formed by E. coli RNA polymerase are eliminated by the addition of purified dda protein, a T4-encoded DNA helicase.

These experiments were performed in collaboration with Mark Hochstrasser.

The multienzyme DNA replication apparatus active at a DNA replication fork must certainly encounter many proteins that are tightly bound to the DNA template inside the cell. How the replication fork deals with these bound proteins is currently not understood. Is the replication fork able to bypass such bound proteins? If so, do these proteins remain bound to the DNA template as the replication fork passes or do they instead dissociate? We have begun to study these questions by using two well-characterized in vitro systems; that of the bacteriophage T4 (Liu et al. 1978; Alberts et al., 1980) and the E. coli RNA polymerase bound to double-stranded bacteriophage fd DNA at specific transcription promoter sequences (Schaller et al., 1978; Konings and Shoenmakers, 1978).

MATERIALS AND METHODS

Enzymes - Bacteriophage fd gene protein was the generous gift of T. Meyer and K. Geider, purified as described (Meyer and Geider, 1979). The T4 DNA replication proteins corresponding to genes 32, 44/62, 45, 41 and 43 (T4 DNA polymerase) were purified using published procedures (Bittner et al. 1979; Morris et al. 1979a, Morris et al. 1979b). The T4 dda protein (DNA-dependent ATPase) was purified in our laboratory by C. Victor Jongeneel according to a new protocol (manuscript in preparation). All of these preparations were nearly homogeneous and free of detectable nuclease contaminations. Restriction enzymes were purchased from New England Biolabs. Purified E. coli RNA polymerase saturated with sigma factor was the generous gift of M. Chamberlin, purified as described (Gonzalez, Wiggs and Chamberlin, 1977).

DNA - Double-stranded supercoiled replicative form (RF) fd DNA was purified from fd-infected E. coli cells using the method of Clewell (1972). To obtain specifically nicked fd RF DNA, fd gene 2 protein was incubated with 2 µg fd RF DNA in 16 mM Tris acetate pH 7.8, 33 mM potassium acetate, 5 mM magnesium acetate, 0.5 mM dithiothreitol, and 10% glycerol at 30°C for 45 min. The gene 2 protein was then inactivated by heating the reaction to 65°C for 10 min. Analysis of the products of this reaction by agarose gel electrophoresis showed that more than 80% of the fd DNA was nicked in the reaction.

Replication Reactions - Unless stated otherwise, in vitro DNA replication was carried out in the presence of 33 mM Tris-acetate, pH 7.8, 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM dithiothreitol, 0.5 mM rATP, 0.1 mM each dATP, dGTP, dCTP, and $\alpha[^{32}P]$-TTP; and 2 µg/ml DNA. In "core replication system" reactions, the following T4 DNA replication proteins were present at the indicated concentrations: T4 DNA polymerase, 2.5 µg/ml; T4 gene 32 protein, 80 to 100 µg/ml, T4 gene 44/62 protein, 20 µg/ml; and T4 gene 45 protein, 18 µg/ml. For "core-41 replication system" reactions, the T4 gene 41 protein 6 µg/ml was also added. Where indicated, the T4 dda protein was added to reactions to a final concentration of 3 µg/ml. E. coli polymerase was present in indicated reactions at 3.2 µg/ml unless otherwise noted.

Replication reactions were started synchronously at a specific nick in the fd DNA template, as outlined in Fig. 1. Reaction mixtures lacking only replication proteins and dCTP were prepared at 4°C. All of the proteins were added and the mixture incubated at 37°C for 1 min; dCTP (0.1 mM) was then added to allow replication forks to proceed. At various intervals, aliquots were removed from the reaction into tubes containing sodium dodecyl sulfate (SDS) to produce a final concentration of 3% SDS. After 10 min at room temperature, each aliquot was "spin dialyzed" to remove unincorporated dNTPs through sepharose CL6B (Pharmacia) as described by Neal and Florini (1973). Buffer was then added to each aliquot to produce a final concentration of 20 mM Na_3EDTA, 10% sucrose and 0.1% bromocresol green dye and the DNA analyzed by electrophoresis through a 0.6% agarose gel, using 30 mM NaOH and 2 mM Na_3EDTA as the

running buffer. Gels measuring 0.4 x 14 x 21.5 cm were electrophoresed at 30 volts for 40 hr. The gels were then dried onto Whatman 3 MM filter paper and autoradiographed at -70°C using Kodak XR-2 film with a Dupont "Lightning-Plus" intensifying screen. The sizes of the radioactively-labeled, newly synthesized DNA strands were determined by comparison with ^{32}P-labeled restriction fragments of bacteriophage T4 DNA of known size (O'Farrell et al. 1980).

The kinetics of nucleotide incorporation were followed by spotting of 4 µl aliquots of reactions at indicated times onto GF/A glass fiber filters (Whatman). The filters were subsequently washed at 4°C for 5 min in 20 ml/filter of 5% trichloroacetic acid containing sodium pyrophosphate washed 3 times at 4°C for 5 min in 1 M HCl, 2 times in ethanol and then dried and counted in a toluene-based scintillation fluid.

RESULTS

Replication forks pause at specific sites on a double-helical DNA template - In vitro DNA replication reactions catalyzed by the T4 core replication system include T4 gene 32, 45, and 44/62 proteins in addition to T4 DNA polymerase (the gene 43 product). Each of these proteins is required to achieve efficient DNA synthesis starting from a nick on a double-stranded DNA template (Nossal and Peterlin, 1979, Sinha et al., 1980). The replication process begins by the covalent addition of deoxyribonucleotides to the 3'OH terminus at the DNA nick; the strand containing the 5' phosphate at the nick is displaced ahead of the growing DNA chain during this reaction, producing long single-stranded DNA "tails." On a circular DNA template, the length of these tails can greatly exceed the length of the parental circular DNA, since replication proceeds in a "rolling circle" mode that allows many rounds of copying. The "core-41" replication system includes in addition the protein product of T4 gene 41. The gene 41 protein is a DNA-dependent GTPase (and ATPase) that seems to function as a dimer with a molecular weight of 116,000 (Liu and Alberts, 1981). The addition of this protein to the replication reaction on a double-stranded DNA template leads to a dramatic increase in the rate of fork movement (Liu et al., 1978; Alberts et al., 1980; Barry and Alberts, manuscript in preparation). Continuous GTP hydrolysis is required for this effect, and the stimulation of fork rate is believed to reflect the action of the 41 protein as a DNA helicase that uses the energy of GTP hydrolysis to run along the lagging strand template and melt the duplex DNA ahead of the fork.

In the experiments to be described, we have used the double-stranded replicative form (RF) of the bacteriophage fd genome as a template for replication. We treated this DNA with purified fd gene 2 protein in order to create a uniquely nicked DNA template (nicked at nucleotide 5781; Meyer et al. 1979). Previous work has shown that this DNA is an effective substrate for the T4 in vitro replication system, and that replication proceeds in a rolling circle mode (Meyer et al. 1981). In an attempt to synchronize the starts on this template the procedure shown in Fig. 1 was adopted. Uniquely nicked RF fd DNA was briefly incubated with purified E. coli RNA polymerase, which binds tightly to the double-stranded at known transcription promoter sites (Schaller et al., 1978; Konings and Shoenmakers, 1978). The promoter site that binds RNA polymerase the most strongly is located 1000 base pairs "downstream" (in the direction of replication) from the site of replication initiation, the unique fd gene 2 nick.

When the kinetics of DNA synthesis were followed by the incorporation of radioactively labeled DNA precursors into an acid precipitable form, the results shown in Fig. 2 were obtained. It can be seen that both core ("5-protein") and core-41 ("6-protein") replication reactions are efficiently inhibited by the addition of purified E. coli RNA polymerase. The inhibition was partially relieved when all of the nucleotide precursors for RNA synthesis (ATP, UTP, GTP and CTP), were included in the reaction.

The effect of RNA polymerase concentration - The RNA polymerase fraction used in these experiments was extremely pure (as judged by SDS gels) and had a high specific activity. It is important to determine whether the dramatic block of DNA replication observed upon addition of RNA polymerase is indeed due to this molecule rather than to some contaminant. We could calculate the ratio of RNA polymerase molecules to fd template molecules from the known concentrations, and vary this ratio using serial dilutions of the RNA polymerase. As the polymerase:DNA ratio decreases, more and more DNA templates should be left free of RNA polymerase molecules, and these RNA polymerase-free templates should support unimpeded replication. If a single DNA-bound RNA polymerase molecule is causing the block in DNA replication on each template, then the percentage of unobstructed templates in the reaction mixture should equal the percentage of the total replication that remains. The amount of RNA polymerase-free template ratio can be predicted from a Poisson probability distribution for each RNA polymerase:template ratio. If the inhibition is due to some contaminant instead, then in general the predictions calculated from the RNA polymerase:template ratio will not reflect the actual inhibition seen.

The effect of the serial dilutions of the RNA polymerase on total DNA synthesis, measured as acid-precipitable counts, is shown in Fig. 3. There is very good agreement between expected and observed values of synthesis, if we make the reasonable assumption that ~50% of the RNA polymerase molecules are active for effective binding (M. Chamberlin, personal communication).

RNA polymerase blocks the replication fork at specific sites - When we analyzed the length of in vitro DNA replication products using alkaline agarose gel electrophoresis (Figs. 4 and 5), we found that replication by the T4 core system was stably blocked by RNA polymerase at specific sequences that appear to be the fd promoter sites downstream from the site of replication initiation rather than the much weaker replication pause sites seen on a protein-free template (see Chapter 2). Replication reactions that include the T4 gene 41 protein, or both the 41 and 61 proteins, behave identically (data not shown). Therefore neither the core replication system, nor the core-41 or core-41-61 replication systems can produce a replication fork capable of bypassing an RNA polymerase molecule that is tightly bound to a transcription promoter sequence.

The T4 replication fork will "follow" a transcribing RNA polymerase molecule - We investigated the interaction of in vitro DNA replication forks with transcribing - and therefore moving - RNA polymerase molecules by including ribonucleotide triphosphates in our reactions. In the in vitro system described in Fig. 1, the direction of transcription for all RNA polymerases is the same as the direction of leading strand replication. As judged by total DNA synthesis, we observe a partial relief of the DNA synthesis inhibition caused when its transcription is allowed (Fig. 2). The analysis of DNA product length in Fig. 4 shows that blocked replication forks can proceed once transcription is allowed, albeit at a rate that is only about one-fourth the rate observed in the absence of RNA polymerase. Since it is known that the rate of transcription is much slower than

the rate of replication, this result is expected if the rate of replication fork movement is being limited by the rate of RNA polymerase movement in our reactions. In fact, the rate at which the fork moves is about 10 nucleotides per second, a rate comparable with rates of rRNA synthesis during transcription (Chamberlin, 1976).

The T4 DNA dda protein removes the replication fork barrier presented by promoter-bound RNA polymerase - The T4 dda (DNA-dependent ATPase) protein has been shown to be a DNA helicase that hydrolyzes ATP to drive its DNA duplex melting reaction (Krell et al., 1979). This protein has recently been purified and further characterized in our laboratory (Jongeneel and Alberts, manuscript in preparation). Similar to the effect obtained by addition of the gene 41 protein, addition of the purified T4 dda protein to the core (5-protein) replication reaction results in a 4-fold increase in the rate of fork movement (see Chapter 2).

The alkaline agarose gel data presented in Fig. 5 shows the results of including the T4 dda protein in replication reactions in the absence and presence of E. coli RNA polymerase. It can be seen that whereas the replication fork is otherwise strongly blocked by RNA polymerase, the rates of DNA elongation in the presence and absence of RNA polymerase are nearly identical in those replication reactions that include T4 dda protein.

Whether the dda helicase allows the replication fork to proceed through sites of bound protein by releasing the protein in front of the replication fork or by some other mechanism that allows protein binding to persist while the fork passes is currently being examined.

DISCUSSION

The genomes of procaryotes and eucaryotes are associated with a multitude of structural and regulatory proteins. Replication of DNA *in vivo*, therefore, takes place not on naked DNA templates such as those standardly used in *in vitro* replication studies, but rather on complex, protein-coated DNA molecules. An important type of DNA-binding protein is RNA polymerase, which not only binds to DNA (at specific sequences known as promoters), but also actively utilizes the DNA as a template for the enzymatic synthesis of RNA.

Since the fundamental biological processes of replication and transcription both occur on the same DNA template molecules, it is important to understand how these two processes interact. At least in *E. coli* transcription and replication are not spatially (or temporally) compartmentalized, since both take place on one very large DNA molecule. Given that the rate of replication fork movement inside the cell is more than an order of magnitude greater than the rate of RNA polymerase movement, replication forks must pass through regions of the DNA template that are being actively transcribed. There is currently little information concerning how this occurs.

The only relevant data would seem to be derived from the electron microscopic studies of McKnight, Bustin and Miller (1977). These pioneering studies of the interaction of replication forks and ribosomal transcription units in Drosophila embryos suggest that the outcome depends on whether the replication forks and the transcribing RNA polymerase molecules are moving in the same or in

opposite directions. When the replication fork and RNA polymerase move in the same direction, the replication fork seems to invade the region being transcribed either by knocking off RNA polymerase molecules and their attached transcripts or by passively following behind the RNA polymerases; the techniques employed in these studies cannot distinguish between these two possibilities. It is important to note that the rates of replication fork and RNA polymerase movement are very similar in eucaryotes (unlike procaryotes), and therefore the passive following of RNA polymerase by the replication fork would not necessarily be detrimental to the cell. When the directions of transcription and replication were opposite, McKnight et al. found that the replication fork could not enter the ribosomal transcription unit. Therefore it seems that either replication forks cannot move past an oncoming RNA polymerase molecule, or that replication termination sites exist before one reaches the transcription termination site on these ribosomal genes. The results obtained by examining nonribosomal genes using the same techniques were preliminary and rather equivocal (McKnight and Miller, 1979).

These kinds of electron microscopic studies, while informative, are limited in that they offer a static rather than a dynamic view of active processes. Further, the system studied is complex and largely undefined in terms of protein components. The in vitro studies reported here utilize well characterized systems of purified proteins that can be readily manipulated in terms of components and reaction conditions. The two systems we have employed (the T4 in vitro replication system and purified E. coli RNA polymerase on a RF fd DNA template) should provide a valid model in that each accurately

simulates the corresponding in vivo reaction and together they constitute a system homologous to that found at early stages in a T4-infected E. coli cell (Liu et al., 1978; Alberts et al. 1980; Schaller et al., 1978; Konings and Schoenmakers, 1978).

We have found that purified E. coli RNA polymerase, when bound to promoter regions on fd DNA, constitutes a significant barrier to replication by the T4 replication fork produced by both the core replication system and the system to which the gene 41 and/or 61 proteins are added (Figs. 2 and 4). The site where fork movement is most strongly inhibition correspond to the position of an unusually strong in vitro promoter on the fd genome-located between nucleotides 378 and 418, 1000 nucleotides downstream from the site of replication initiation in these experiments. This promoter precedes the fd gene X, an open reading frame within the nucleotide sequence of fd gene II; the function of this X gene in vivo is currently not known. We are presently investigating the interaction of the in vitro replication fork with RNA polymerase bound to other promoters, in order to ascertain whether the blocking of replication forks by promoter-bound RNA polymerase molecule occurs only at those promoters that show especially strong RNA polymerase binding.

When we add ribonucleoside triphosphates to our system and thereby allow transcription to occur, replication proceeds past the gene X strong promoter site, although the rate of fork movement is about 4-fold slower than the rate of the fork formed by the core replication system in the absence of RNA polymerase (Fig. 4). It seems that the replication fork passively follows a transcribing RNA polymerase under these conditions, which is possible because tran-

scription and replication are proceeding in the same direction on this template (Fig. 6). Further studies are required to assess the interactions of a replication fork with a RNA polymerase molecule moving in the opposite direction.

We found that the addition of purified T4 dda protein, an ATPase and DNA helicase, to our system completely eliminates the arrest of fork movement by RNA polymerase molecules (Fig. 5). This is true whether or not transcription is allowed and at RNA polymerase to DNA molecular ratios of up to 17 to 1 (data not shown) The dda helicase could possibly play a unique role at the replication fork by removing protein barriers as it drives the unwinding of the duplex DNA template. This could explain why this protein is required for T4 DNA replication: although the dda protein is non-essential for viral growth in a wild-type E. coli host, it is required for growth in an E. coli optA mutant strain (P. Gauss, D.H. Doherty and L. Gold, personal communication). Its function would therefore seem to be essential but to be bypassed in vivo by a functionally equivalent host protein. Studies are in progress to elucidate the mechanism by which the dda protein eliminates the RNA polymerase barriers and to follow the fate of the RNA polymerase and its transcripts subsequent to passage of a replication fork through the transcribed region.

REFERENCES

Alberts, B.M., Barry, J., Bedinger, P., Burke, R.L., Hibner, U., Liu, C.-C., and Sheridan, R. (1980) In Mechanistic Studies of DNA Replication and Genetic Recombination, Vol. 19 ICN-UCLA Symposia on Molecular and Cellular Biology (Alberts, B.M., ed) pp. 449-471. Academic Press, New York.

Bittner, M., Burke, R.L., and Alberts, B.M. (1979) J. Biol. Chem. 254, 9565-9572.

Chamberlin, M.J. (1976) in RNA Polymerase (Losick, R. and Chamberlin, M.J., ed). pp. 17-68. Cold Spring Harbor Laboratory, New York

Clewell, D.B. (1972) J. Bacteriol. 110, 667-676.

Gonzalez, N., Wiggs, J. and Chamberlin, M.J. (1977) Arch. Bicohem. Biophys. 182, 404-408.

Konings, R.N.H. and Schoenmakers, J.G.G. (1978) In The Single-Stranded DNA Phages (Denhardt, D.T., Dressler, D. and Ray, D.S., eds). p 507-530. Cold Spring Harbor, New York.

Krell, H., Durwald, H., and Hoffmann-Berling, H. (1979) Eur. J. Biochem. 93, 387-395.

Liu, C.-C., Burke, R.L., Hibner, U., Barry, J., and Alberts, B.M. (1978) Cold Spring Harbor Symp. Quant. Biol. 43, 469-487.

McKnight, S.L., Bustin, M., and Miller, O.L., Jr. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 741-754.

McKnight, S.L. and Miller, O.L., Jr. (1979) Cell 17, 551-563.

Meyer, T.F., and Geider, K. (1979) J. Biol. Chem. 254, 12636-12641.

Meyer, T.F., Geider, K., Kurz, C., and Schaller, H. (1979) Nature 278, 365-367.

Meyer, T.F., Baumel, I., Geider, K., and Bedinger, P. (1981) J. Biol. Chem., 256, 5810-5813.

Morris, C.F., Moran, L.A., and Alberts, B.M. (1979) J. Biol. Chem. 254, 6797-6802.

Morris, C.F., Hama-Inaba, H., Mace, D., Sinha, N.K., and Alberts, B.M. (1979) J. Biol. Chem. 254, 6787-6796.

Neal, M.W. and Florini, J.R. (1973) Anal. Biochem. 55, 328-330.

Nossal, N.G., and Peterlin, B.M. (1979) J. Biol. Chem. 254, 6032-6037.

O'Farrell, P.H., Kutter, E., and Nakanishi, M. (1980) Molec. gen. Genet. 179, 421-435.

Schaller, H., Beck, E. and Takanami, M. (2978) in The Single-Stranded DNA Phages (Denhardt, D.T., Dressler, D. and Ray, D.S., eds) pp 139-163. Cold Spring Harbor, New York.

Sinha, N.K., Morris, C.F., and Alberts, B.M. (1980) J. Biol. Chem. 255, 4290-4303.

FIGURE LEGENDS

Fig. 1. Schematic representation of the assay used to test replication fork movement in the presence of E. coli RNA polymerase. The conditions utilized are those described in Materials and Methods. Specifically nicked RF fd DNA (2 μg/ml) was incubated for 1 min at 37°C with 3.2 μg/ml purified E coli RNA polymerase. Since the RNA polymerase preparations used in these experiments were saturated with sigma factor, the RNA polymerase molecules became bound to specific fd promoter sequences. To synchronize the start of replication forks, a limited amount of replication was allowed (1 min at 37°C) with the addition of T4 replication proteins and three deoxyribonucleotides only; subsequent addition of the missing deoxyribonucleotide, dCTP, led to further synchronous elongation of nascent DNA chains. Aliquots were removed from reactions at specified times to analyze incorporation of radioactive precursors into an acid-precipitable form (Figs. 2 and 3) or the length of nascent DNA chains by alkaline agarose gel electrophoresis (Figs. 4 and 5), as described in Materials and Methods.

Fig. 2. Time course of DNA synthesis in replication reactions carried out in the presence and absence of RNA polymerase. Reactions were performed as described in Materials and Methods. Reactions catalyzed by the core replication system ("5-protein reactions") include T4 genes 43, 44/62, 45 and 32 proteins. The "6-protein reactions" also include the T4 gene 41 protein in addition, which greatly enhances the rate of fork movement (Alberts et al.

1980). At 0, 3, 6 and 9 min, aliquots were removed and acid-precipitated onto glass fiber filters. The number of pmoles of nucleotide incorporated was determined from the specific radioactivity of the reaction mix. Five-protein and six-protein reactions that included E. coli RNA polymerase (17 molecules per fd DNA template molecule) are largely inhibited. A partial relief of this inhibition is seen when transcription was allowed by the addition of four ribonucleotide triphosphates, CTP, ATP, GTP and UTP.

Fig. 3. The effect of the RNA polymerase:DNA ratio on in vitro synthesis. Replication reactions including T4 genes 43, 44/62, 45 41 and 32 proteins were performed as described in Materials and Methods, with the indicated concentrations of RNA polymerase present (DNA concentration of 2 μg/ml). Aliquots were removed for acid precipitation at 0, 5 and 10 min. Even at RNA polymerase concentrations equivalent to 1.5 RNA polymerase molecules per fd RF DNA molecule a significant amount of inhibition is observed. The ratios were determined using molecular weights of 5×10^5 for RNA polymerase and 4×10^6 for the double-stranded fd DNA.

Fig. 4. Analysis of the length of the nascent DNA molecules synthesized in reactions carried out in the presence and absence of RNA polymerase. Aliquots of equal volume from reactions performed as described in Materials and Methods were removed after 2, 5 and 10 min and analyzed by alkaline agarose gel electrophoresis and followed by autoradiography. Numbers to the left of the auto-

radiogram indicate the positions of DNA restriction fragments of known size that were electrophoresed in the same gel. It can be seen that reactions including RNA polymerase (lanes d to f) were strongly inhibited, with no nascent DNA chains longer than 7400 nucleotides. In reactions where all four rNTPs were included so that transcription could proceed (lanes g to i) replication was blocked, although the forks moved at a slower rate than than in the absence of RNA polymerase (lanes a to c)).

Fig. 5. Effects of the addition of T4 dda protein on replication fork movement. Reactions were performed as described in Materials and Methods. Lanes a to c show reaction products of 5-protein reactions (core replication system) after 1, 2 and 5 min of incubation. Lanes d to f show that RNA polymerase blocks 5-protein replication at specific sites (as in Fig. 4). Lanes g to i and j to l compare products of reactions including T4 dda protein in the absence and presence of RNA polymerase, and show two effects: first, the fork rate without RNA polymerase present is four times faster than without dda protein, and second, the replication rate is now unaffected by RNA polymerase even though 17 molecules of RNA polymerase have been added per template DNA molecule.

Fig. 6. Schematic view of the observed interactions of E. coli RNA polymerase with T4 replication forks that lack dda protein.

FIGURE ONE

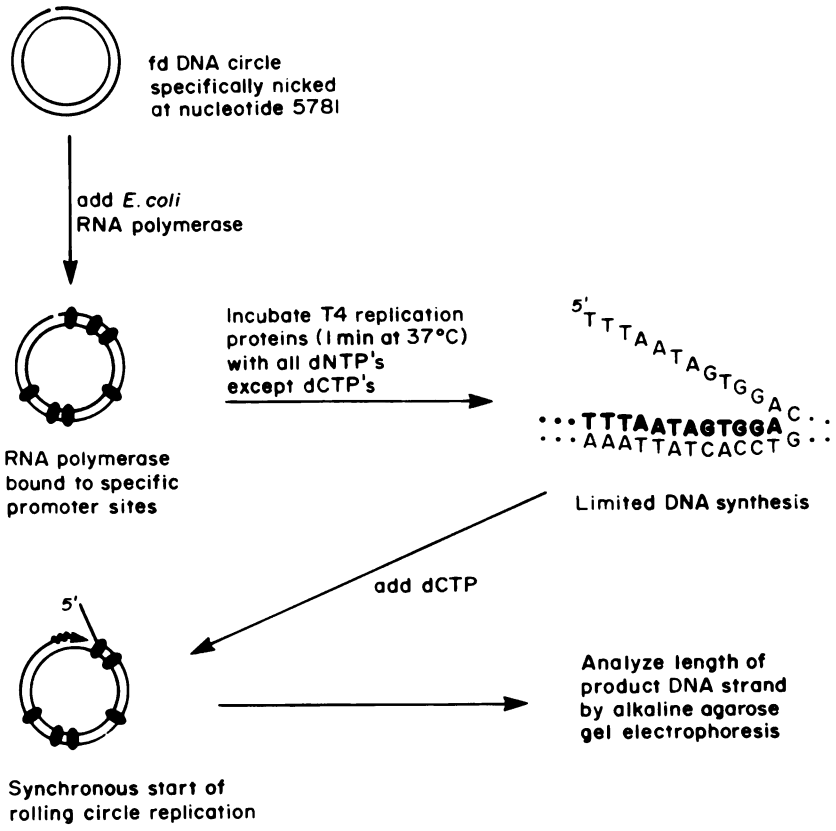

fd DNA circle
specifically nicked
at nucleotide 5781
add E. coli
RNA polymerase
RNA polymerase
bound to specific
promoter sites
Incubate T4 replication
proteins (1 min at 37°C)
with all dNTP's
except dCTP's
5'TTTAATAGTGGAC
TTTAATAGTGGA
AAATTATCACCTG
Limited DNA synthesis
add dCTP
5'
Synchronous start of
rolling circle replication
Analyze length of
product DNA strand
by alkaline agarose
gel electrophoresis

FIGURE TWO

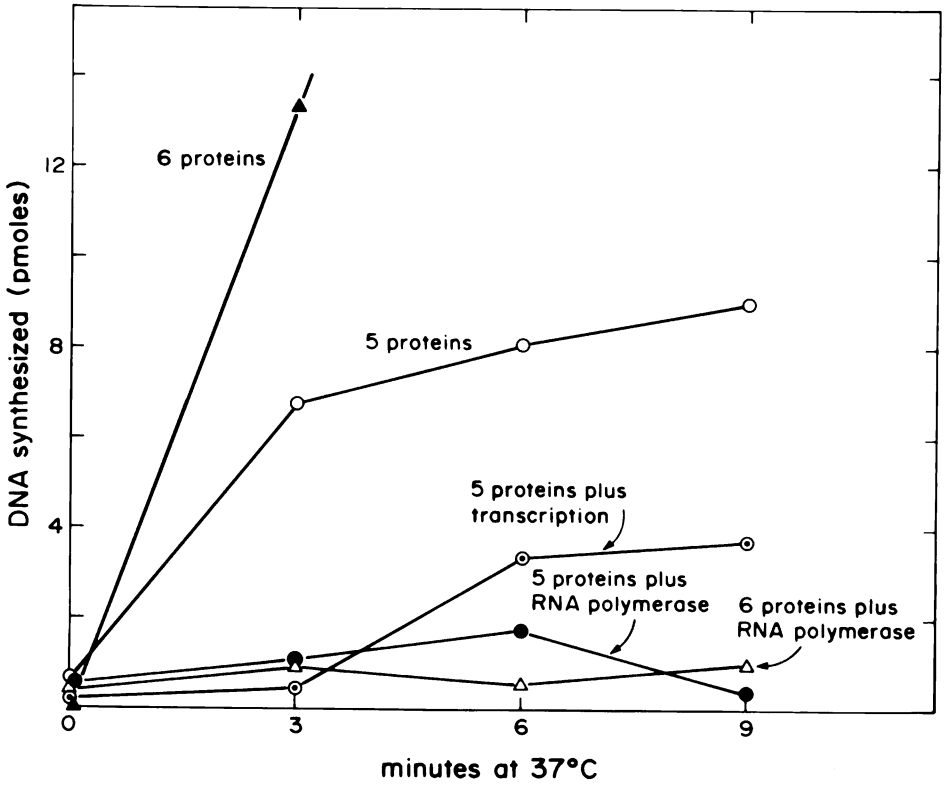

DNA synthesized (pmoles)
12
8
4
0
3
6
9
minutes at 37°C
6 proteins
5 proteins
5 proteins plus transcription
5 proteins plus RNA polymerase
6 proteins plus RNA polymerase

FIGURE THREE

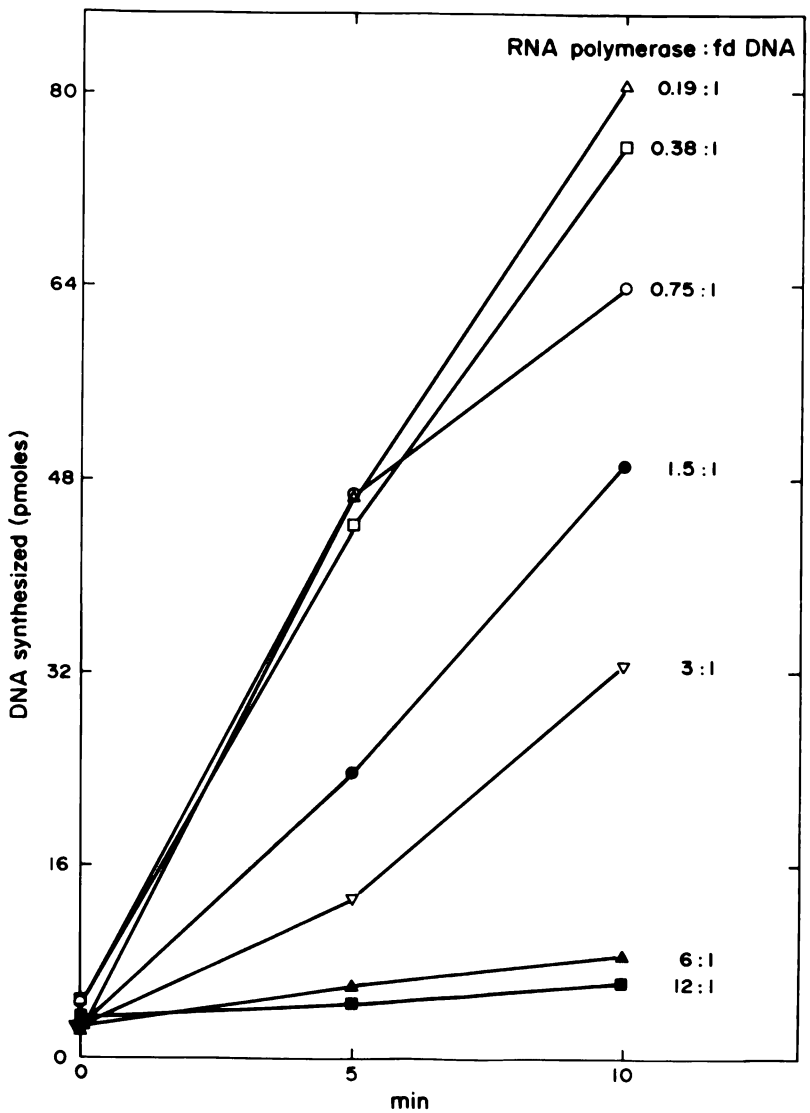


FIGURE FOUR

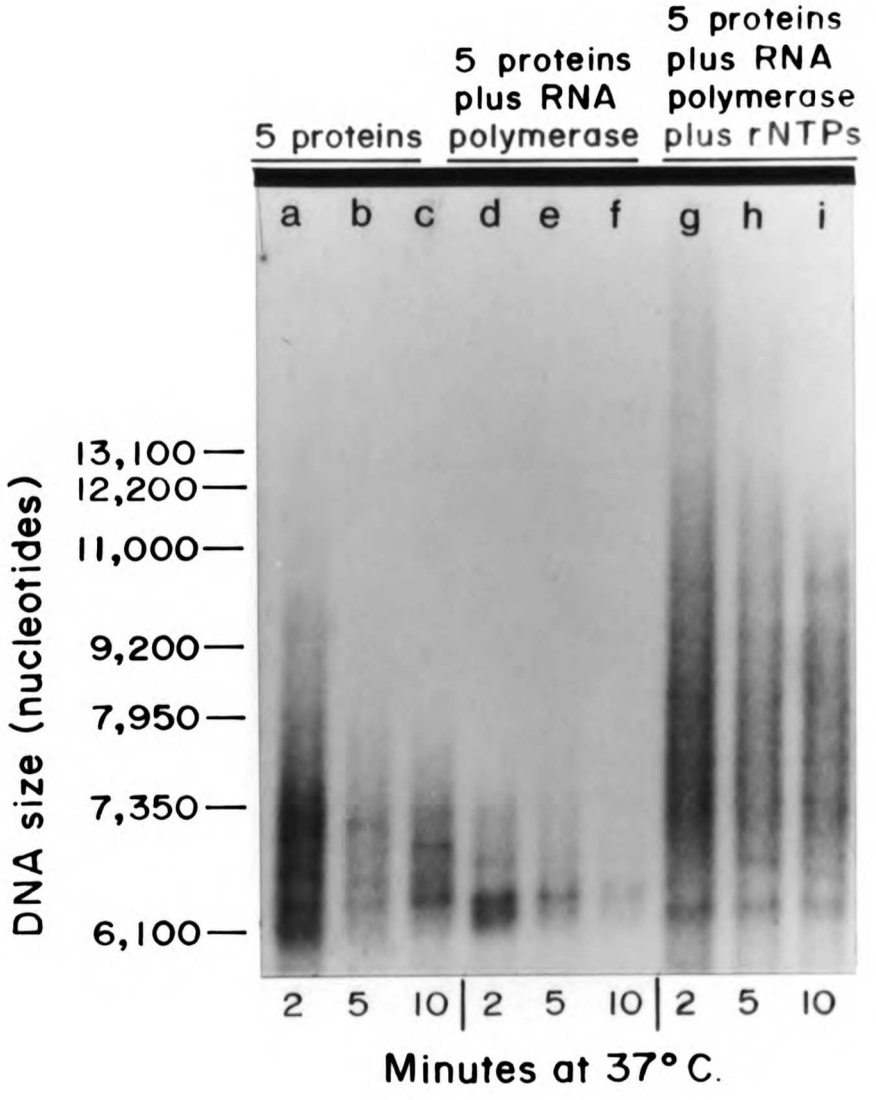

5 proteins
5 proteins plus RNA polymerase
5 proteins plus RNA polymerase plus rNTPs
a b c d e f g h i
DNA size (nucleotides)
13,100
12,200
11,000
9,200
7,950
7,350
6,100
2 5 10 2 5 10 2 5 10
Minutes at 37° C.

FIGURE FIVE

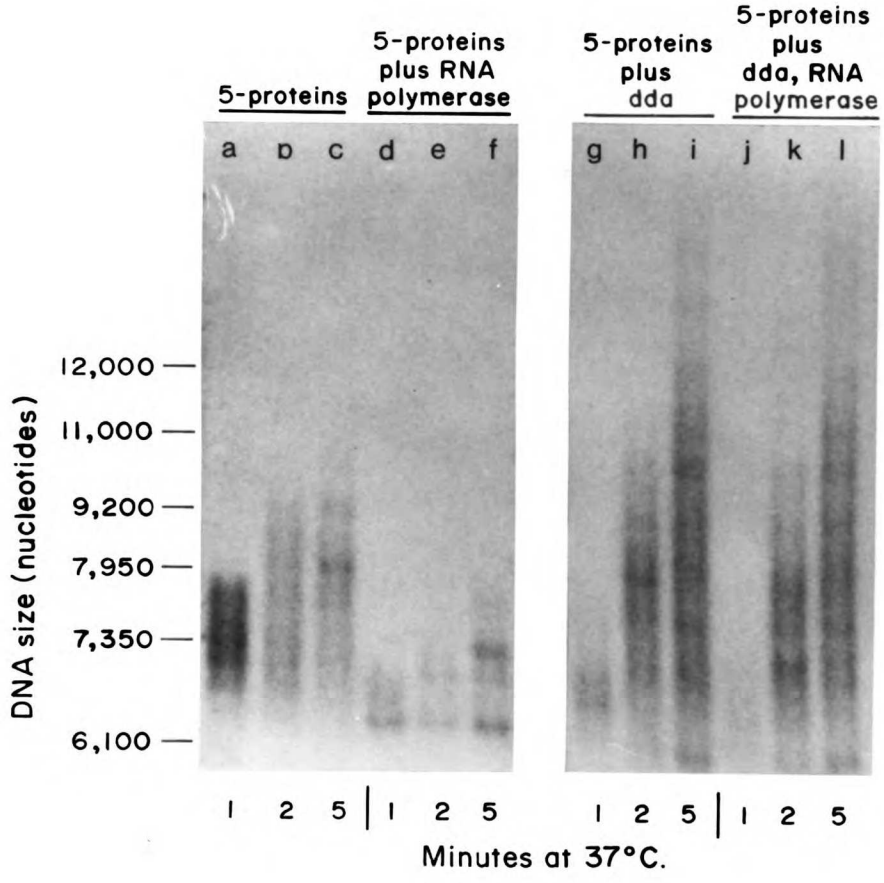
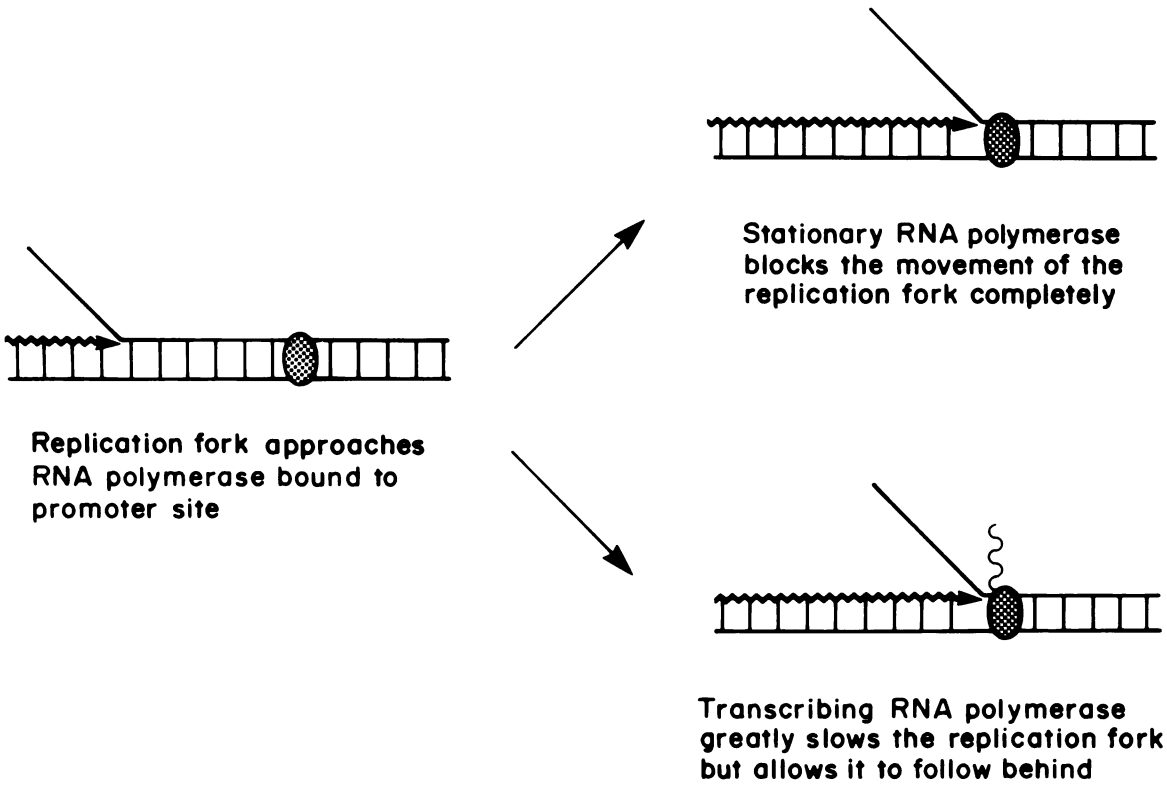

5-proteins
5-proteins plus RNA polymerase
5-proteins plus ddα
5-proteins plus ddα, RNA polymerase
a b c d e f g h i j k l
DNA size (nucleotides)
12,000
11,000
9,200
7,950
7,350
6,100
1 2 5 | 1 2 5 1 2 5 | 1 2 5
Minutes at 37°C.

FIGURE SIX


Stationary RNA polymerase blocks the movement of the replication fork completely
Replication fork approaches RNA polymerase bound to promoter site
Transcribing RNA polymerase greatly slows the replication fork, but allows it to follow behind

UC San Francisco LIBRARY
UNIVERSITY OF CALIFORNIA

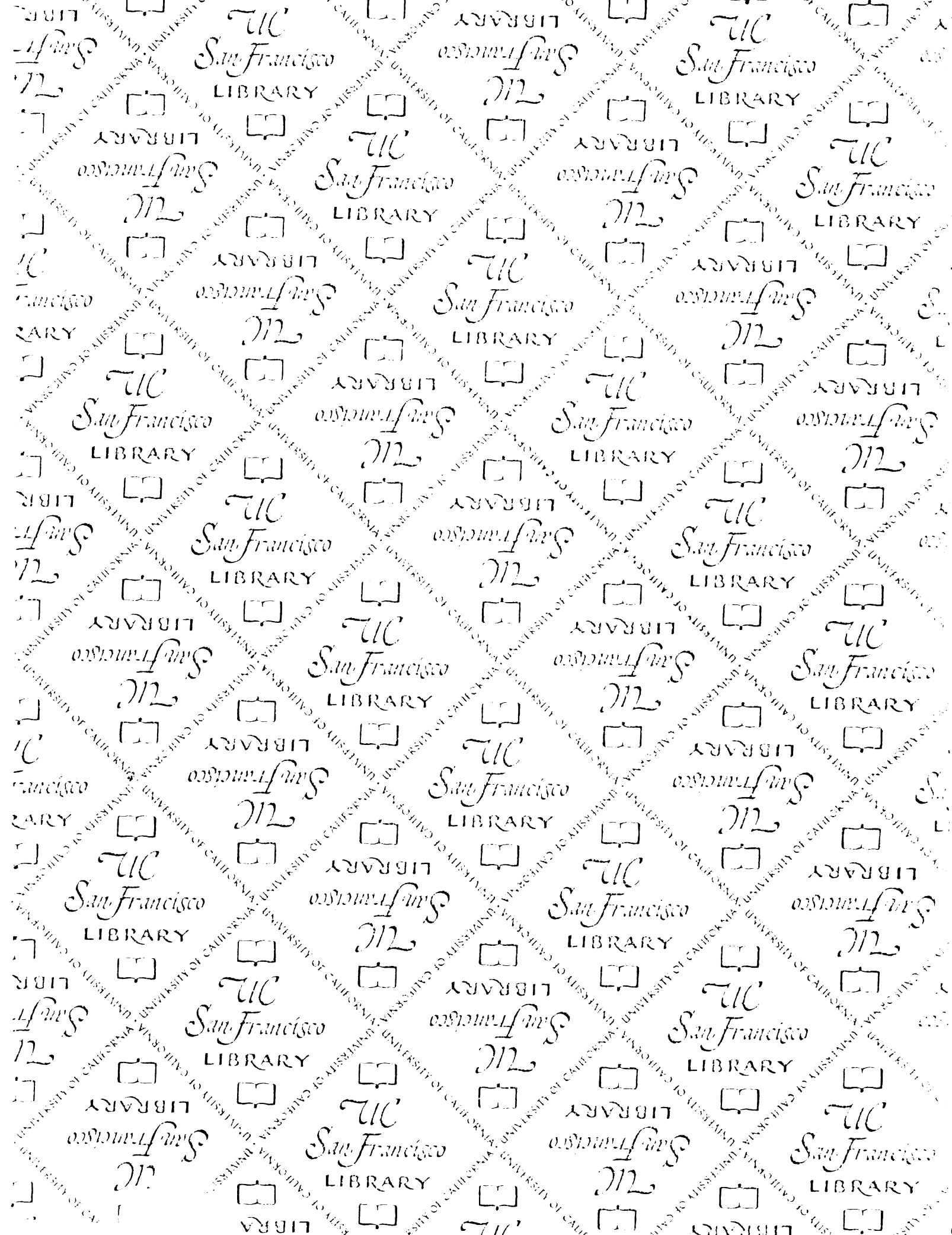
UC San Francisco LIBRARY

FOR REFERENCE

NOT TO BE TAKEN FROM THE ROOM

CAT. NO. 23 012

PRINTED IN U.S.A.


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA

UC San Francisco LIBRARY

FOR REFERENCE

NOT TO BE TAKEN FROM THE ROOM

CAT. NO. 23 012

PRINTED IN U.S.A.

UC San Francisco LIBRARY
UNIVERSITY OF CALIFORNIA

UC San Francisco LIBRARY

FOR REFERENCE

NOT TO BE TAKEN FROM THE ROOM

BRODART CAT. NO. 23 012 PRINTED IN U.S.A.

FIGURE TWO

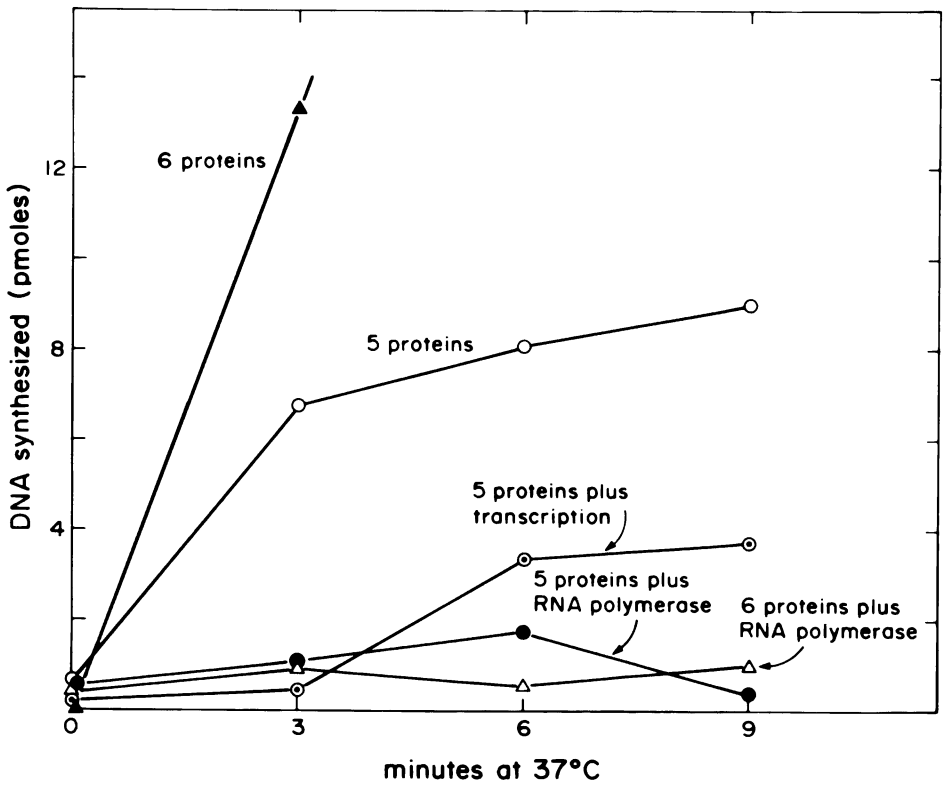

DNA synthesized (pmoles)
12
8
4
0
3
6
9
minutes at 37°C
6 proteins
5 proteins
5 proteins plus transcription
5 proteins plus RNA polymerase
6 proteins plus RNA polymerase

FIGURE THREE

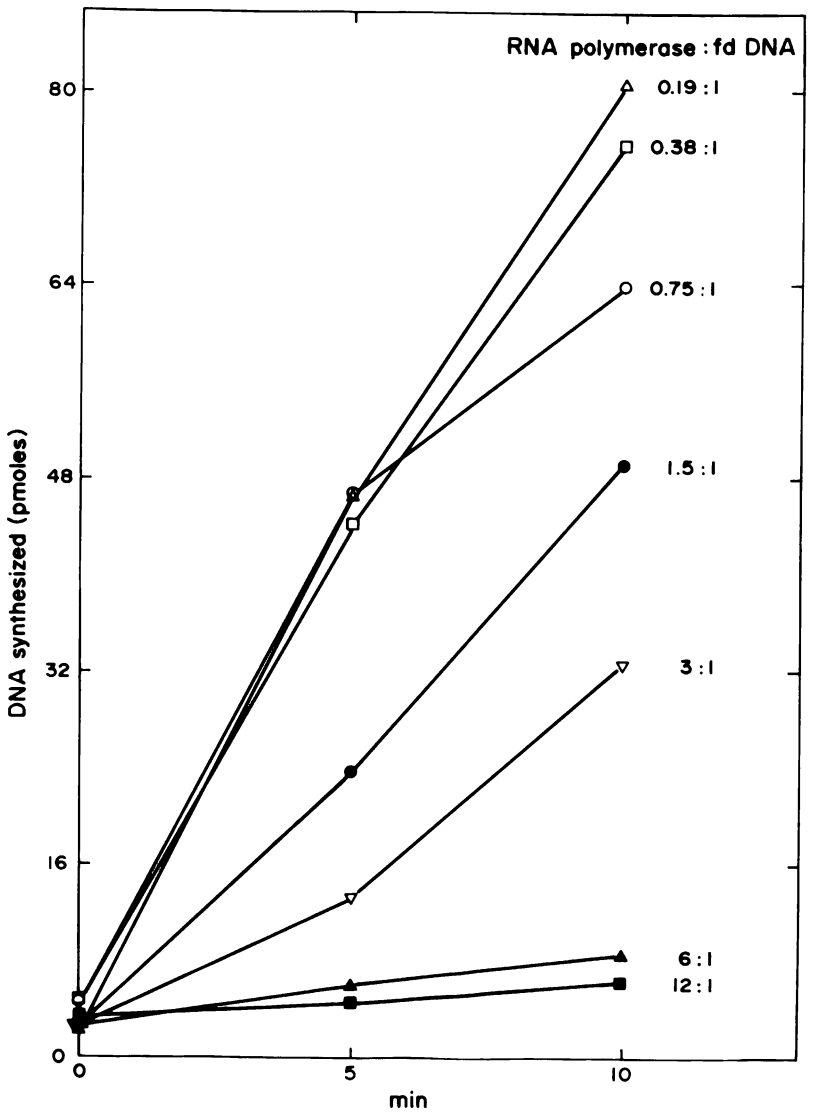


FIGURE FOUR

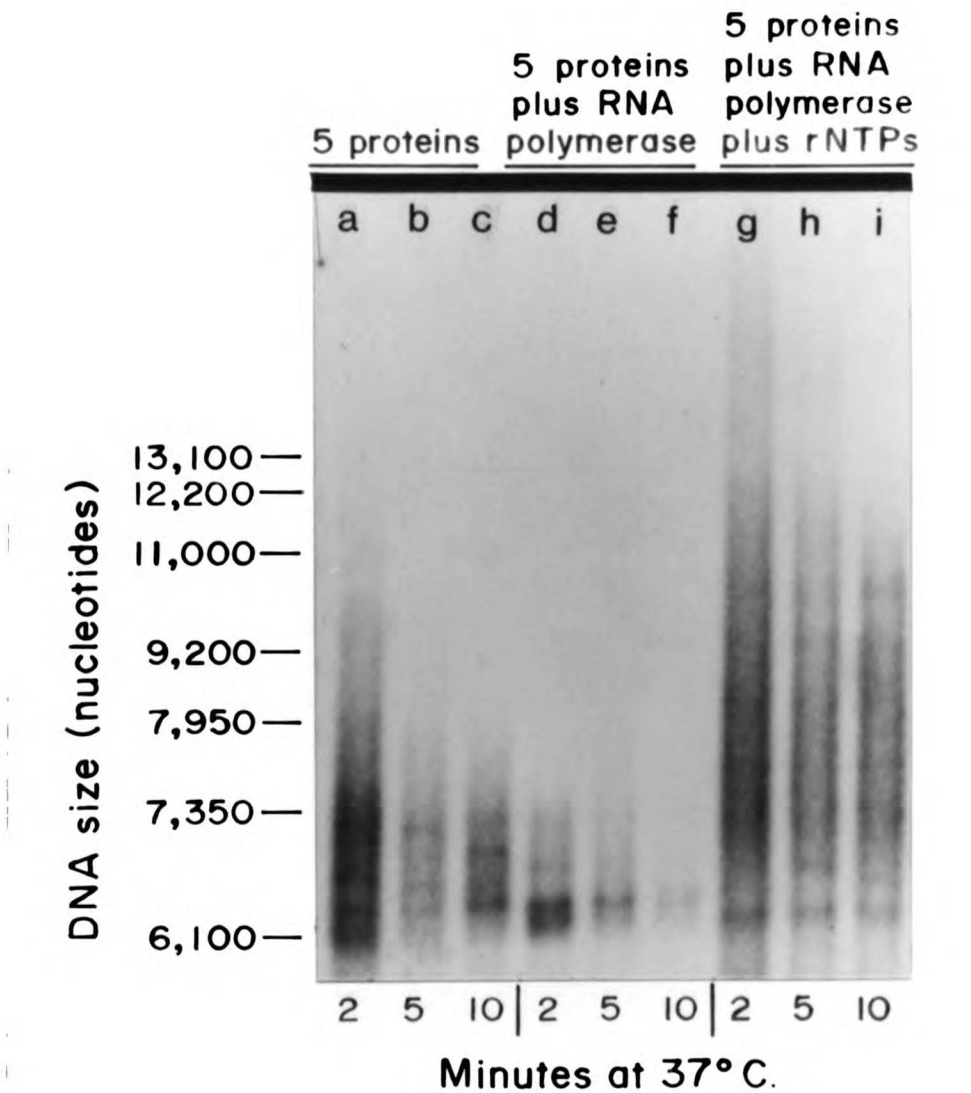

5 proteins
5 proteins plus RNA polymerase
5 proteins plus RNA polymerase plus rNTPs
a b c d e f g h i
DNA size (nucleotides)
13,100
12,200
11,000
9,200
7,950
7,350
6,100
2 5 10 2 5 10 2 5 10
Minutes at 37° C.

FIGURE FIVE

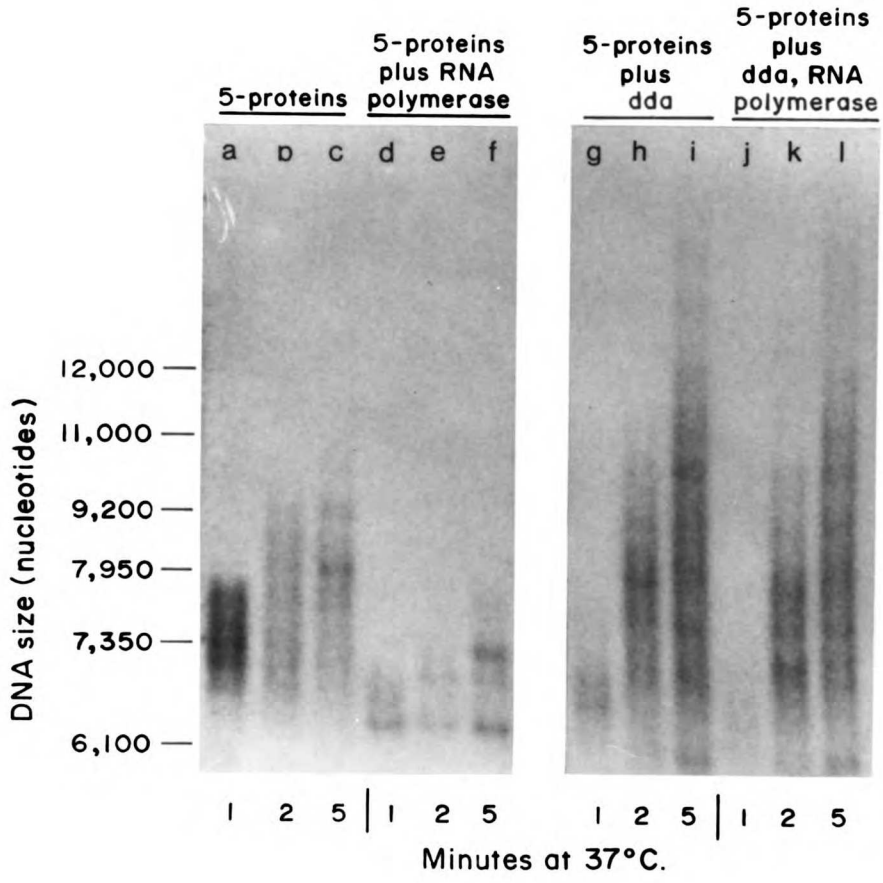
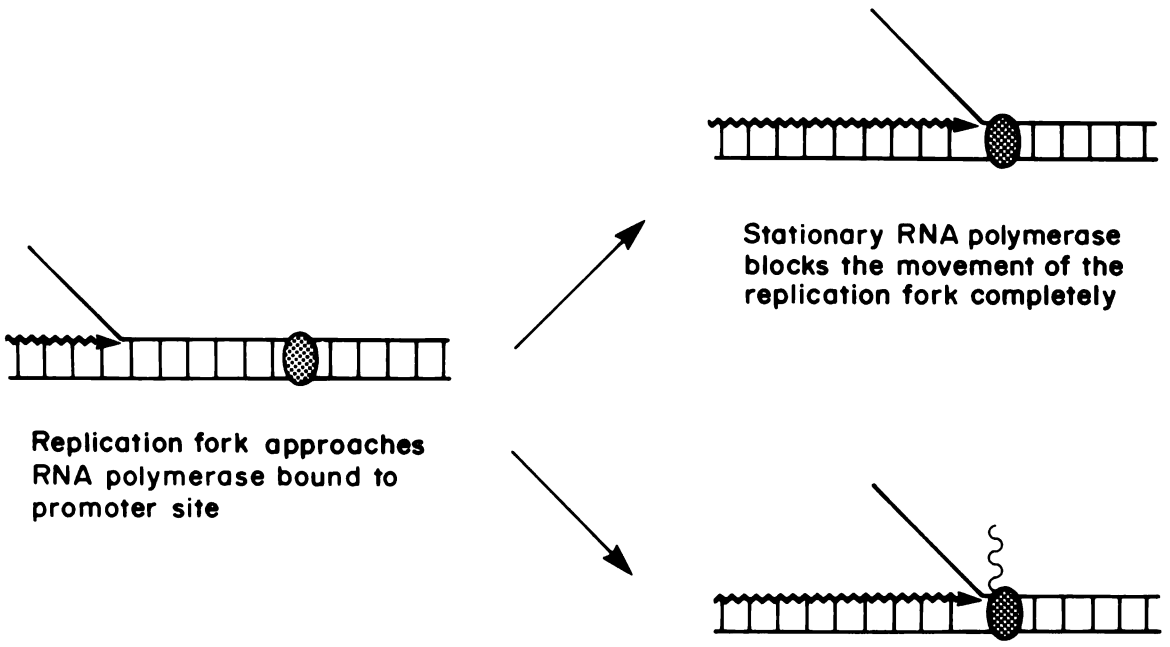

5-proteins
5-proteins plus RNA polymerase
5-proteins plus ddα
5-proteins plus ddα, RNA polymerase
a b c d e f
g h i j k l
DNA size (nucleotides)
12,000
11,000
9,200
7,950
7,350
6,100
1 2 5 | 1 2 5
1 2 5 | 1 2 5
Minutes at 37°C.

FIGURE SIX


Stationary RNA polymerase blocks the movement of the replication fork completely
Replication fork approaches RNA polymerase bound to promoter site
Transcribing RNA polymerase greatly slows the replication fork, but allows it to follow behind

UC San Francisco LIBRARY
UNIVERSITY OF CALIFORNIA