PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Prokaryotic DNA Replication Mechanisms

B. M. Alberts

Phil. Trans. R. Soc. Lond. B 1987 **317**, 395-420
doi: 10.1098/rstb.1987.0068

Phil. Trans. R. Soc. Lond. B **317**, 395–420 (1987)
Printed in Great Britain

Prokaryotic DNA replication mechanisms

By B. M. Alberts

Department of Biochemistry and Biophysics, University of California at San Francisco, San Francisco, California 94143-0448, *U.S.A.*

The three different prokaryotic replication systems that have been most extensively studied use the same basic components for moving a DNA replication fork, even though the individual proteins are different and lack extensive amino acid sequence homology. In the T4 bacteriophage system, the components of the DNA replication complex can be grouped into functional classes as follows: DNA polymerase (gene 43 protein), helix-destabilizing protein (gene 32 protein), polymerase accessory proteins (gene 44/62 and 45 proteins), and primosome proteins (gene 41 DNA helicase and gene 61 RNA primase). DNA synthesis in the *in vitro* system starts by covalent addition onto the 3′OH end at a random nick on a double-stranded DNA template and proceeds to generate a replication fork that moves at about the *in vivo* rate, and with approximately the *in vivo* base-pairing fidelity. DNA is synthesized at the fork in a continuous fashion on the leading strand and in a discontinuous fashion on the lagging strand (generating short Okazaki fragments with 5′-linked pppApCpXpYpZ pentaribonucleotide primers). Kinetic studies reveal that the DNA polymerase molecule on the lagging strand stays associated with the fork as it moves. Therefore the DNA template on the lagging strand must be folded so that the stop site for the synthesis of one Okazaki fragment is adjacent to the start site for the next such fragment, allowing the polymerase and other replication proteins on the lagging strand to recycle.

Introduction

Although the molecular details of the DNA replication process have been investigated in many different biological systems, the most completely understood replication enzyme complexes are responsible for the replication either of the bacterial chromosome or of bacterial viruses (bacteriophages). In the pre-recombinant-DNA era, these were the only organisms that provided a rich-enough source of replication proteins to allow the multienzyme systems responsible for moving a replication fork to be reassembled conveniently in a test tube. Moreover, only for these organisms were extensive mutational analyses done to identify the genes encoding replication proteins. Through the development of '*in vitro* complementation assays', the mutants produced by geneticists enabled numerous replication proteins of unknown function to be purified by biochemists; the prior genetic analysis thereby greatly facilitated the process of reconstructing DNA replication events *in vitro*.

Now that recombinant DNA technologies allow large amounts of proteins to be produced from genes cloned from any source, abundant amounts of replication proteins can in principle be obtained even from higher eukaryotic cells. However, because we lack an extensive collection of DNA replication mutants for such cells, much of the enzymology must be done by using what is known about prokaryotic replication systems as a guide for making an intelligent guess of the types of activities to be purified. It is my role in this symposium to help provide

such a guide. In what follows, I will attempt to outline some basic lessons learned over the past 20 years, emphasizing features that are common to several prokaryotic replication systems. Many of these findings have been made in other laboratories; however, wherever the work from my laboratory is relevant, I will use it preferentially for illustrative purposes.

All of the DNA is synthesized by a self-correcting DNA polymerase enzyme

The replication fork is an asymmetric structure, with distinct modes of synthesis on its 'leading' and 'lagging' strands (Kornberg 1980, 1982). Although the DNA on the leading strand is made continuously, nucleotide by nucleotide, by a rapidly translocating DNA polymerase molecule, the DNA on the lagging strand is made by a discontinuous 'backstitching' process. This results in the production of a series of short DNA fragments (Okazaki fragments) as intermediates in the synthesis of the lagging strand. Consequently, completion of the lagging strand requires the sealing together of approximately 3×10^3 fragments of average size 1500 nucleotides for a bacterial chromosome and 4×10^7 fragments of average size 150 nucleotides for a set of human chromosomes.

The joining of the Okazaki fragments into a continuous DNA strand constitutes the major DNA repair reaction in a growing cell. This repair reaction is complicated by the presence of a short stretch of ribonucleic acid at the 5′ end of each Okazaki fragment (4–10 nucleotides, depending on the organism). This 'RNA primer' is required for the initiation of each Okazaki fragment, and it must be removed by a special type of nuclease (RNase H) before the DNA chain can be completed. A gap is thereby created between adjacent DNA fragments, which is filled in by a repair-type of DNA synthesis. Finally, DNA ligase seals the nick left between the fragments, creating a continuous DNA strand. The extent of repair is especially great in eukaryotes: about 10 of every 150 nucleotides, or 7% of each lagging strand, are normally synthesized during the Okazaki fragment repair process.

Cells do not generally settle for an awkward and inefficient mechanism when carrying out their important biosyntheses. The existence of the above type of replication fork in organisms as diverse as viruses, bacteria, yeast and humans therefore demands an explanation. The asymmetric DNA replication fork has withstood the test of time presumably because it ensures that the DNA is copied by a highly faithful mechanism. To overcome the infidelity inherent in the formation of Watson–Crick base pairs, all of the DNA on both strands is synthesized by the same type of self-correcting DNA polymerase molecule, which is closely associated with a $3' \rightarrow 5'$ exonuclease activity that preferentially excises mispaired nucleotides immediately after their incorporation. As discussed elsewhere (Alberts & Sternglanz 1977), a polymerase enzyme that corrects its own errors efficiently can add nucleotides only in the $5' \rightarrow 3'$ chain direction; moreover, it cannot start synthesizing a polynucleotide chain without a pre-existing primer. These two properties of the DNA polymerase are sufficient to explain the discontinuous mode of synthesis on the lagging strand, as well as the rapidly erased RNA primers that are synthesized by a second enzyme (RNA primase) to start each Okazaki fragment.

The properties expected for a self-correcting DNA polymerase molecule have been clearly established for the prokaryotic enzymes for more than a decade (Brutlag & Kornberg 1972). Although eukaryotic DNA polymerases have often been isolated without the exonuclease activity associated with prokaryotic polymerases, common sense argues strongly for the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

conservation of this central aspect of prokaryotic systems. This suspicion has been strengthened in recent years by the finding, in *Escherichia coli*, that the proofreading exonuclease can be carried on a polypeptide chain that is separate from the polymerase (Scheuermann & Echols 1984); moreover, at least some forms of the mammalian enzyme can be isolated with an associated exonuclease (Lee *et al.* 1984). Prudence suggests that any DNA polymerase that is isolated without proofreading capabilities should be viewed as an enzyme that is missing a subunit, or is defective for some other reason.

A complex of a DNA helicase and an RNA primase, called a 'primosome', is assembled onto DNA to start a replication fork

As I shall discuss subsequently, DNA polymerase can catalyse several different types of DNA synthesis in a cell. The hallmark of a replication fork is not the polymerase, but the presence at the fork of a protein complex, called a 'primosome', which travels rapidly along the lagging-strand DNA template (McMacken *et al.* 1977; Arai & Kornberg 1981). The primosome is a complex of two distinct enzymes. One is the RNA primase molecule, which periodically synthesizes an RNA primer for Okazaki fragment synthesis. The other protein in the complex is a DNA helicase, an enzyme that uses ATP hydrolysis energy to propel itself along a DNA single strand, prying open any region of DNA double helix ahead. Leading-strand DNA synthesis cannot proceed until the helicase opens up the helix ahead of the leading-strand DNA polymerase molecule, whereas lagging-strand DNA synthesis requires the RNA primers that the primase provides for initiation of successive Okazaki fragments. The primosome thus serves to unite and coordinate the DNA synthesis on the two sides of the replication fork; its presence distinguishes DNA replication from DNA repair.

Studies in the bacteriophage T4 system have revealed that the primosome moves at a rate of about 400 nucleotides per second, as required to account for the rate of replication fork movement (Liu & Alberts 1981). Moreover, once bound to the DNA, the primosome moves in a highly processive manner (where processivity is defined as the number of nucleotides that the complex moves along the DNA before dissociating). Thus, in a typical *in vitro* reaction on a double-stranded DNA template, two classes of moving DNA polymerases are detected: those with a primosome driving their synthesis, which move rapidly, and those without a primosome, which move much more slowly (Alberts *et al.* 1980). These two populations can be maintained for a long period of time in a test tube, revealing that a bound primosome rarely, if ever, dissociates from a moving replication fork. It seems to be primarily the tight DNA binding of the DNA helicase that makes the T4 primosome so processive.

If RNA primase could prime DNA synthesis on any single-stranded DNA that it encounters in cells, chromosomes would be unable to maintain single-stranded regions that are useful for special purposes. It is therefore interesting that the RNA primase in both the T4 and *E. coli* replication systems is essentially inactive on its own: although it can bind to DNA, efficient primer synthesis under intracellular conditions requires the addition of the DNA helicase to complete the primosome complex (Kornberg 1980, 1982; Cha & Alberts 1986). The helicase, in turn, is loaded onto DNA only with difficulty, and once bound to it moves along its DNA strand in a highly processive manner. These and other observations focus attention on the DNA helicase as the central element of a replication fork. As we shall see, recent *in vitro* experiments suggest that (1) this helicase is normally loaded onto a DNA strand at a replication origin and

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

immediately attracts a molecule of the RNA primase to complete the primosome, and (2) the primosome may be all that is needed to establish and maintain a replication fork; the assembly of the remaining proteins follows automatically, and the fork then moves rapidly away from the replication origin until all of the DNA has been replicated.

The fork initiation process that occurs at replication origins has been intensively studied in two related biological systems: *E. coli* and λ bacteriophage. In *E. coli*, the replication origin has been identified as a 245 base pair (b.p.) DNA sequence. This *ori*C sequence is highly conserved

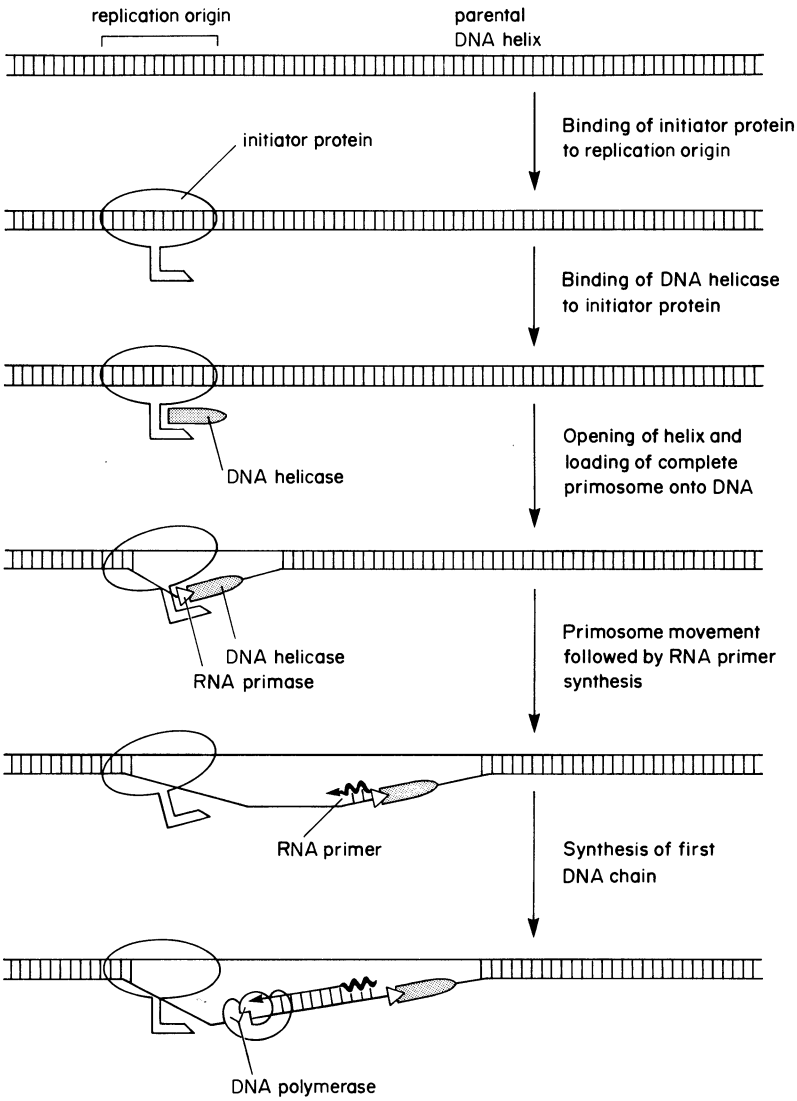

FIGURE 1. An outline of the initial steps leading to the formation of replication forks at the *E. coli* and λ bacteriophage DNA replication origins, as revealed by *in vitro* studies utilizing a mixture of highly purified proteins (Kornberg *et al.* 1987; McMacken *et al.* 1987). Subsequent steps result in the initiation of three more DNA chains by a pathway that is not yet clear, producing two replication forks that move outwards in opposite directions from the origin. For *E. coli* DNA replication, the initiator protein is the dnaA protein; for bacteriophage λ DNA replication the initiator protein is the λ O protein. In both cases, the primosome is composed of the *E. coli* dnaB (DNA helicase) and dnaG (RNA primase) proteins. DNA helicases were first discovered by Abdel-Monem & Hoffmann-Berling (1976); however, the dnaB protein has only recently been recognized as such an enzyme (LeBowitz & McMacken 1986).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

among the Enterobacteriaceae (Zyskind *et al.* 1983). It binds multiple copies of a special *E. coli* initiator protein, the dnaA protein, during the fork initiation process (Fuller & Kornberg 1983). Bacteriophage lambda and related viruses parasitize the *E. coli* host replication machinery by using the same primosome and replication enzymes. But they have their own replication origin and produce their own initiator proteins, as required to attract the *E. coli* enzymes selectively to their own genomes (Furth & Wickner 1983).

By comparing the mechanism of fork initiation at the *E. coli* and λ bacteriophage origins, one would hope to be able to ascertain those features of the initiation process that are most general. From this perspective, the results obtained to date have been quite satisfying. In both systems, the DNA-bound initiator protein seems to bind the DNA helicase (the dnaB protein) to the origin, via a linker protein (either the *E. coli* dnaC protein or the lambda P protein). The two strands of the DNA helix are then opened adjacent to the origin, and the helicase is loaded onto the newly exposed single strands. The DNA primase (the dnaG protein) is added to form a primosome that propels itself away from the origin, opening additional regions of helix as it goes (figure 1). RNA primer synthesis by the primosome then allows a DNA polymerase molecule to begin DNA synthesis (Kornberg *et al.* 1987; McMacken *et al.* 1987). This polymerase apparently heads back towards the origin (Tsurimoto & Matsubara 1984), where it presumably encounters the DNA-bound initiator protein. Exactly what happens next is unclear, but in both the *E. coli* and λ bacteriophage *in vitro* systems a complete replication fork is rapidly formed. This requires the synthesis of a second RNA primer and the recruitment of a second DNA polymerase molecule. It is not yet clear which of the two DNA polymerase molecules ends up synthesizing the leading strand. The experimental analysis is complicated by the fact that two replication forks can be formed at these origins, and *in vivo* each origin normally initiates bidirectional DNA replication.

The seven T4 bacteriophage proteins that catalyse replication fork movement can be grouped into three functional units

In order to discuss the details of replication fork movement, I will describe the replication apparatus for bacteriophage T4. This large virus (genome size of 166000 b.p.) encodes all of its own replication proteins. Although the host cell is parasitized for many other essential metabolic pathways, T4 DNA replication, DNA repair and genetic recombination are essentially unaffected when this bacteriophage infects mutant *E. coli* that are defective in any of these processes (Wood & Revel 1976; Nossal & Alberts 1983).

An extensive genetic analysis of conditionally lethal mutant T4 bacteriophages has revealed seven 'replication genes' whose products are required for DNA synthesis during the 'early' period of T4 bacteriophage infection (Epstein *et al.* 1964). By measuring the stimulation of DNA synthesis in lysates of mutant bacteriophage-infected cells, the activities of five of the missing gene products were detected in wild-type extracts (Barry & Alberts 1972). This '*in vitro* complementation assay' allowed these replication proteins to be purified to homogeneity in an active form, without requiring prior knowledge of their individual functions (Morris *et al.* 1979*a*, *b*). The two other replication proteins in the set could not be detected in this assay; these proteins were identified by using affinity chromatography on DNA cellulose to select them from crude extracts (Alberts & Frey 1971; Burke *et al.* 1985).

The set of proteins that was purified in this way form the core of the T4 replication

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

apparatus, and their names and properties are outlined in table 1. The seven proteins with assigned gene numbers can be placed into three functional units. In the first unit are the products of T4 genes 43, 44/62, and 45, which form the T4 DNA polymerase holoenzyme. By itself, the T4 gene 43 protein (DNA polymerase) is capable of elongating pre-existing primers on single-stranded DNA templates, and its 3′→5′ exonuclease activity provides a proofreading function that removes misincorporated nucleotides (Brutlag & Kornberg 1972). The products of genes 44/62 and 45 are the polymerase accessory proteins and together they display a DNA-dependent ATPase activity (Mace & Alberts 1984*a*). On a primed single-stranded DNA template, the polymerase accessory proteins interact with the DNA polymerase in a reaction that requires ATP hydrolysis and results in a dramatic increase in both the rate and processivity of the DNA synthesis catalysed by the polymerase (Alberts *et al.* 1975; Piperno & Alberts 1978; Newport *et al.* 1980; Huang *et al.* 1981; Mace & Alberts 1984*c*). The accessory proteins appear to form a special 'sliding clamp'. This clamp keeps a moving polymerase molecule attached to the 3′ end of a growing DNA chain for a prolonged period (Huang *et al.* 1981; Mace & Alberts 1984*b*, *c*) while letting the polymerase dissociate rapidly when it stalls (as it does on the lagging strand after the completion of each Okazaki fragment (Selick *et al.* 1987)).

Table 1. Properties of bacteriophage T4 replication proteins

type of protein	T4 Gene	molecular mass/kDa	activities
1. DNA polymerase	43	103.5	5′→3′ polymerase; 3′→5′ exonuclease
2. polymerase accessory proteins	44/62; 45	35.7/21.4; 24.7	DNA–termini-dependent ATPase, dATPase
3. helix-destabilizing protein (SSB protein)	32	34.5	cooperative binding to single-stranded DNA
4. primosome components	41	53.8	processive DNA helicase
	61	39.5	RNA primase
5. template-clearing helicase	*dda*	48	DNA helicase

The second functional unit is the T4 primosome. Its DNA helicase component is the 41 protein, which utilizes the energy of GTP hydrolysis to unwind the DNA helix ahead of the advancing replication complex (Alberts *et al.* 1980; Liu & Alberts 1981); its direction of movement along the DNA is consistent with its association with the template on the lagging strand (Venkatesan *et al.* 1982). This protein interacts with the gene 61 RNA primase to form the primosome (Liu *et al.* 1979; Liu & Alberts 1980).

The final functional unit is the gene 32 protein. This helix-destabilizing protein (or single-stranded DNA binding protein, SSB) binds cooperatively to all single-stranded regions of the DNA template, eliminating short stretches of helical structure while leaving the bases freely accessible to the polymerase (Alberts *et al.* 1969; Alberts & Frey 1971). The 32 protein thereby stimulates DNA synthesis by the polymerase; it also exerts part of its effect through a direct interaction with the polymerase holoenzyme (Huberman *et al.* 1971; Burke *et al.* 1980).

It is important to note that the *E. coli* replication proteins can be divided up into the same three functional groups: DNA polymerase holoenzyme, primosome, and helix-destabilizing protein (SSB). However, the *E. coli* polymerase holoenzyme is more complex than its T4

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

analogue, containing more than seven different peptide chains (McHenry 1985). Because little or no homology in amino acid sequence can be detected between the proteins of homologous function in these two organisms, the T4 and *E. coli* replication systems would seem to be separated by a great span of evolutionary time. Nevertheless, these two organisms seem to use nearly identical mechanisms to replicate their DNA, which suggests that the mechanisms to be described below are general ones, likely to apply in eukaryotes as well.

Positioning the primosome, DNA polymerase holoenzyme, and helix-destabilizing protein at a replication fork

The position of the seven T4 replication proteins at a replication fork, as inferred from the results of numerous experiments, is illustrated in figure 2. There are two DNA polymerase holoenzymes at the fork at all times: one on the leading strand and one on the lagging strand. These two enzyme complexes are apparently identical, each consisting of one polymerase

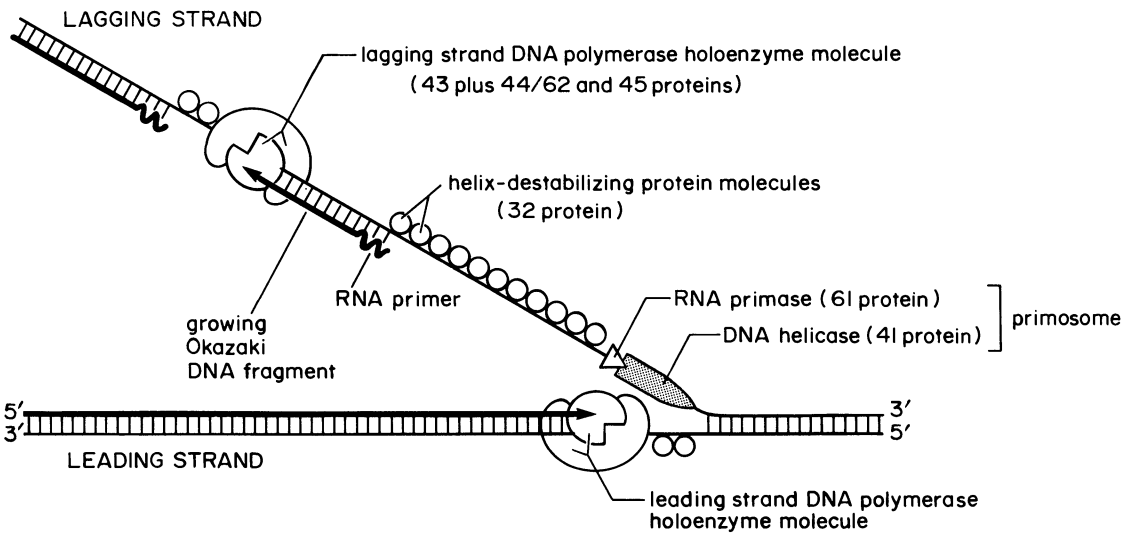

Figure 2. A two-dimensional view of the DNA replication fork of bacteriophage T4. The T4 DNA polymerase is the product of gene 43, and the gene 44/62 and 45 proteins that function with it are designated as 'polymerase accessory proteins'. The gene 32 protein binds to all of the single-stranded DNA at the fork; it is a helix-destabilizing protein, also described as a 'single-strand binding protein'. The gene 41 protein is both a DNA helicase and a protein that is required, together with the gene 61 RNA primase, for RNA primer synthesis on the lagging strand. The average length of the Okazaki fragments formed on the lagging strand is about 1200 nucleotides.

subunit (gene 43 protein), and an associated complex of the polymerase accessory proteins (gene 44/62 and 45 proteins). The primosome is bound to the lagging strand, where it both forces open the helix just ahead of the leading strand DNA polymerase molecule and synthesizes RNA primers for the polymerase on the lagging strand. Finally, all of the single-stranded DNA at the fork is covered by cooperatively bound clusters of the helix destabilizing protein (gene 32 protein).

Detailed studies of RNA primer synthesis in the T4 system have revealed that the RNA primers synthesized on T4 DNA begin at the central nucleotide in the DNA sequence GTT

(Cha & Alberts 1986). This trinucleotide in the template is necessary and sufficient to cause the synthesis of primers with the sequence pppApCpXpYpZ, where X, Y and Z can be any of the four ribonucleotides (figure 3). The observation that all the RNA primers made in the T4 system are five nucleotides long suggests that the RNA primase contains a built-in length-measuring mechanism that causes it to stop synthesis, freeing the 3′ end for the DNA polymerase as soon as a pentaribonucleotide has been made (Liu & Alberts 1980; Nossal 1980). A mechanism of this type makes sense; because the primase product is 'handed off' to the polymerase as soon as it is long enough to serve as a primer, the two enzymes work sequentially instead of competing.

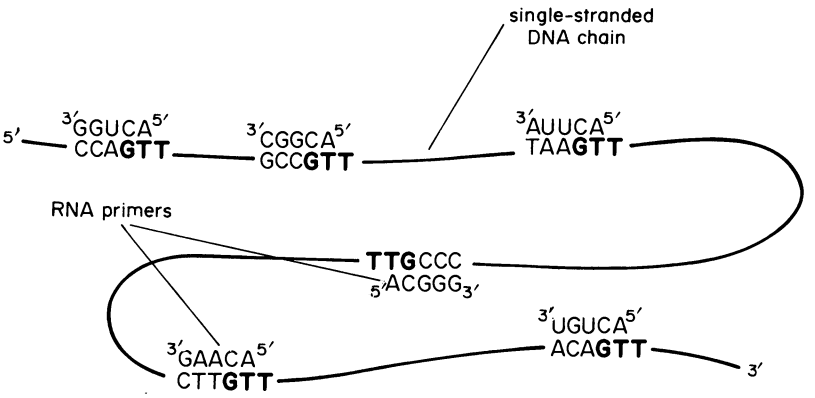

FIGURE 3. Illustration of some of the GTT sites that catalyse RNA primer synthesis by the T4 primosome on a single strand of T4 DNA. The primers made are all pentaribonucleotides that start with the sequence pppApC, as shown.

The fork shown in figure 2 provides what can be called a 'two-dimensional view' of the replication process. In the early 1980s, this view of the fork seemed quite satisfying: the role of each of the genetically identified T4 replication proteins had been defined, and the fork formed *in vitro* by a mixture of these seven proteins could be shown to mimic the *in vivo* fork quite well. Thus DNA synthesis requires the same proteins *in vitro* that the genetics shows are required *in vivo*, the fork moves at about the same rate of 400 nucleotides per second (Sinha *et al.* 1980), and the copying fidelity is close to the very high fidelity found inside the cell (Hibner & Alberts 1980). Moreover, although Okazaki fragments are very heterogeneous in size, their mean length is about the same (1000–1500 nucleotides), and they start with RNA primers of the identical length and sequence as found inside the cell (Kurosawa & Okazaki 1979). However, as further experiments were done with the *in vitro* system, it gradually became clear that this view of the DNA replication process was inadequate in other respects.

SOME PROBLEMS WITH A SIMPLE VIEW OF A REPLICATION FORK

An early result suggesting that something important might be missing from the above view of replication concerned the very efficient use of RNA primers by the fork. As expected for their small size, only a small fraction of the RNA primers are used when added exogenously to the DNA polymerase. In contrast, most of the primers that are synthesized by a moving fork seem to end up at the 5′ end of Okazaki fragments (Liu & Alberts 1980). This suggests that the

primase is very close to the lagging strand DNA polymerase molecule so that it can hand off its primer, which is not an obvious feature of the figure 2 scheme.

A second relevant observation concerns the spacing of the primers synthesized along the lagging strand. Inspection of sequenced regions of the T4 genome reveals that in most regions the trinucleotide sequence GTT that triggers the synthesis of an RNA primer by the primosome is distributed more or less at random, with an average spacing of about 60 nucleotides between adjacent sites (Selick *et al.* 1987). To account for the average size of an Okazaki fragment observed *in vitro* (about 1200 nucleotides) the probability of using any particular GTT encountered on the lagging strand must be small. However, if the GTT sites were selected completely at random by the primosome to start an Okazaki fragment, the first site passed would have the best chance of being used, so that one would expect the largest number of fragments to have a size of about 60 nucleotides, with an exponentially decreasing frequency of fragments tailing out into an extremely broad size distribution. The observed distribution is very different from this expectation; most notably, Okazaki fragments of very small size are essentially absent. It therefore appears that some 'measuring device' is present that prevents the use of the first GTT sites encountered by the primosome on the lagging strand, allowing at least several hundred nucleotides of single-stranded DNA to accumulate behind the fork before a new Okazaki fragment is started.

The third problem with the two-dimensional view of the replication fork was the one that finally forced its revision. The leading-strand DNA polymerase molecule moves continuously, and experiments suggest that the same molecule can remain in place for the synthesis of hundreds of thousands of nucleotides or more (Sinha *et al.* 1980). However, according to the figure 2 model, a lagging-strand DNA polymerase molecule would associate with a newly synthesized RNA primer, move about 1200 nucleotides in a few seconds, collide with the 5′ end of the previous Okazaki fragment, and then dissociate completely from the DNA. DNA synthesis on the lagging strand would therefore require that the fork be in equilibrium with a source of free DNA polymerase holoenzyme molecules at a concentration sufficient to supply a new lagging strand polymerase molecule approximately once every three seconds by a random collision process. In the *in vitro* replication system, this idea can be directly tested by progressively reducing the concentration of the DNA polymerase (gene 43 protein) to levels where polymerase association events would be unable to keep up with needs of the lagging strand. In this case, one would expect an unusually large amount of single-stranded DNA to accumulate on each lagging strand before the start of an Okazaki fragment, causing the average length of these fragments to become much greater than the length observed at high concentrations of polymerase.

A folded replication fork allows recycling of the DNA polymerase molecule that synthesizes each Okazaki fragment on the lagging strand

Suppose that the DNA in the replication fork is folded in a way that brings the start site for each succeeding Okazaki fragment in close juxtaposition with the place where the previous Okazaki fragment will end. In this case, the DNA polymerase on the lagging strand could be 'recycled', thereby making DNA synthesis on the lagging strand completely insensitive to the

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

free concentration of the DNA polymerase, just like the DNA synthesis on the leading strand.

To determine whether the replication fork is folded, a series of *in vitro* replication reactions were done in which the concentration of DNA polymerase was varied with all other components in excess. When alkaline gel electrophoresis was used to determine the size of the Okazaki fragments synthesized, they were found to be essentially unaffected by the polymerase dilution (Alberts *et al.* 1983). Similar samples were examined by electron microscopy to determine the length of the single-stranded connection that joins the double-stranded DNA tail on the lagging strand to the rest of the fork. The median length of this connection remained constant at about 1000 nucleotides, independent of the polymerase dilution. Both of these results are explicable only if the DNA polymerase molecule on the lagging strand is recycled for multiple rounds of Okazaki fragment synthesis, thereby remaining bound to the fork at all times, just like the leading-strand DNA polymerase molecule.

The structure of the replication apparatus proposed to account for these results is schematically illustrated in figure 4. Because of protein–protein interactions detected by protein affinity chromatography (Formosa *et al.* 1983; Alberts *et al.* 1983), as well as other results, I believe that the two DNA polymerase holoenzymes, one on the leading and one on the lagging strand, are directly held together. This leads to the model for DNA synthesis shown in figure 5, in which the synthesis of each Okazaki fragment involves the enlargement of a large loop of DNA, part of which is single-stranded and part of which is double-stranded. The crucial points in the proposed cycle are the 'termination' and 'restart' steps, which involve movements of the DNA on the lagging side of the fork around a fixed lagging-strand DNA polymerase molecule, as indicated.

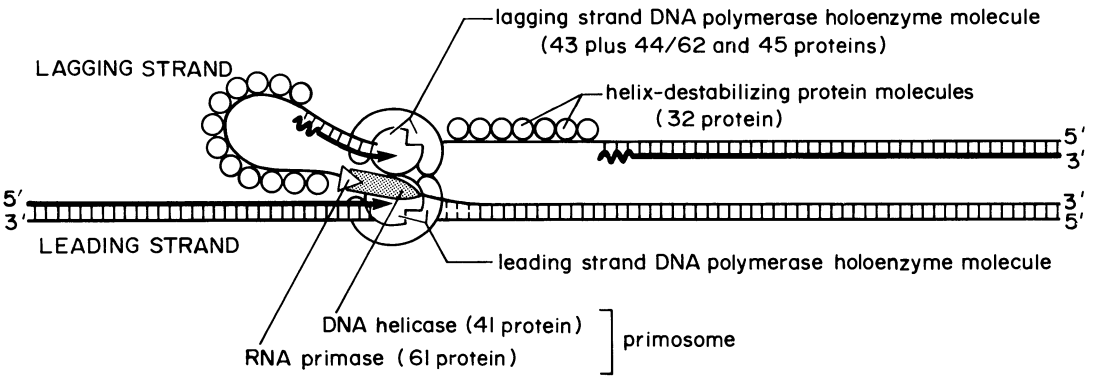

FIGURE 4. The seven proteins of the T4 DNA replication apparatus as they are thought to exist in an actual replication fork. The two-dimensional replication fork in figure 2 has been converted into the structure shown by folding the DNA on the lagging strand in such a way as to bring the DNA polymerase molecule on the lagging strand close to the DNA polymerase molecule on the leading strand. The lagging strand DNA polymerase molecule is thereby held to the rest of the replication proteins, allowing it to be retained for many successive cycles of Okazaki fragment synthesis, as shown in figure 5.

The model just proposed readily accounts for the results of the polymerase dilution experiments, because once a lagging strand DNA polymerase molecule is assembled on the lagging strand, its behaviour is expected to be independent of the concentration of additional polymerase molecules that are free in solution. The model also accounts for the absence of very short Okazaki fragments. Because a new fragment cannot be started until the lagging strand

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

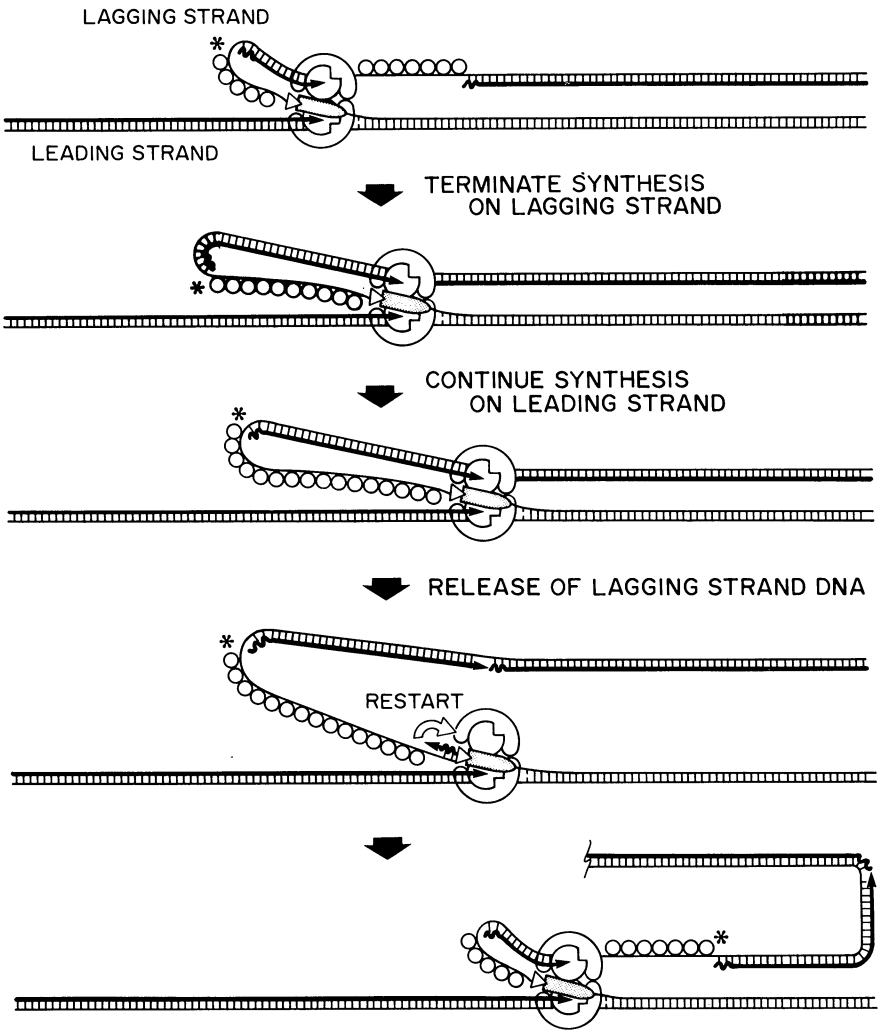

FIGURE 5. A model for DNA synthesis at a replication fork that allows the same DNA polymerase molecule to be repeatedly recycled to synthesize all of the Okazaki fragments on the lagging strand. The crucial step here occurs when the synthesis of each Okazaki fragment is completed. Rather than falling off the entire fork assembly, the DNA polymerase molecule on the lagging strand remains bound to the assembly of replication proteins, letting go only of the DNA on the lagging strand. For the different symbols used to represent the T4 replication proteins see figures 2 and 4. Note that the rate of polymerase synthesis on the lagging strand is faster than the rate of polymerase synthesis on the leading strand (Selick *et al.* 1987).

polymerase molecule completes the synthesis of the previous Okazaki fragment and dissociates from the 3′ end of its DNA chain, the primosome is unable to start two DNA fragments in rapid succession. In general, many GTT sites will be passed by the primosome before the lagging strand DNA polymerase is released from its previous commitment. We suggest that the release of the lagging strand template by its polymerase serves as a signal that activates the primosome to synthesize a new primer at the next available GTT site. With this primer, the lagging strand polymerase then restarts another cycle of Okazaki fragment synthesis. This model can therefore account for the efficient utilization of RNA primers observed.

The size of Okazaki fragments is determined by relative polymerization and dissociation rates

If the rates of polymerization on the leading and lagging strands were identical, with the two polymerases forced to march exactly in step at all times, the length of each successive Okazaki fragment on any particular template molecule would be set equal to the length of the previously synthesized fragment (Alberts *et al.* 1983; Nossal & Alberts 1983). In this case, the polydisperse size-range of Okazaki fragments observed would reflect a corresponding difference in the size of the first Okazaki fragment synthesized on each template DNA molecule *in vitro*. An attractive aspect of this proposal is that the total amount of single-stranded DNA on the lagging strand would remain constant, so that it would be possible for the many molecules of 32 protein at the fork to be recycled within the replication complex.

We have tested whether such a templating mechanism determines Okazaki fragment size in several experiments, and all have produced negative results (Selick *et al.* 1987). For this and other reasons, our current view of coupled leading and lagging-strand DNA synthesis is that depicted previously in figure 5. Because the DNA helix has already been opened to expose the single-stranded lagging strand template, the lagging strand DNA polymerase is free to move at a rate that is faster than that of the leading strand polymerase, even though these two polymerase molecules are physically linked together; measurements of the rate of DNA synthesis on a single strand (Mace & Alberts 1984*c*), as well as the electron microscopical examination of rapidly quenched replicating molecules (Selick *et al.* 1987) support this idea. However, once the lagging-strand polymerase reaches the previously synthesized primer, there is a delay before it releases its DNA template. During this pause, the leading strand polymerase continues its translocation, displacing an additional amount of single-stranded template for the next round of lagging strand synthesis. Finally, the release of the DNA by the lagging strand polymerase allows the associated primosome to synthesize an RNA primer at the next available primer site, so that the polymerase can restart the next cycle of Okazaki fragment synthesis.

What determines the length of Okazaki fragments in the model in figure 5? Note that the primosome becomes 'activated' to make a primer as a consequence of the release of the DNA by the lagging-strand polymerase molecule. Because the potential GTT primer sites will be encountered with a frequency of about once in every 60 nucleotides, we assume that primer site selection by such an activated primosome is a very fast event in the synthesis reaction. If we therefore ignore the brief time required to find a primer site, the major parameters that will influence the spacing of primed DNA starts and hence Okazaki fragment size will be (1) the different rates of movement of the leading and lagging strand polymerases and (2) the length of time that the lagging strand polymerase pauses before releasing the lagging strand template. These parameters can be used to derive an equation for the size of Okazaki fragments, which predicts that, as observed, the ultimate size of the Okazaki fragments synthesized on a particular template will quickly reach an equilibrium value that is independent of the size of the first fragment synthesized on that template (Selick *et al.* 1987).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

The lagging-strand polymerase must dissociate rapidly from its DNA template upon completion of each Okazaki fragment

The recycling of the DNA polymerase on the lagging strand raises a new complication, in that we must now explain how it is possible for this polymerase molecule to dissociate within a second or so after it finishes each Okazaki fragment. In contrast, a very tight binding of the polymerase is required on the leading strand. For example, the observation that this polymerase can remain in place for the synthesis of hundreds of thousands of nucleotides demands a half-time for its dissociation of 5 min or more.

The requirement for very tight binding of the polymerase on the leading strand and much less tight binding of the polymerase on the lagging strand could be solved, in principle, by the use of different polymerases on the two strands. This might explain the existence of both DNA polymerase α and DNA polymerase δ in mammalian cells (see Fairman *et al.*, this symposium). In *E. coli*, there is evidence for a functional asymmetry between the two halves of the dimeric polymerase holoenzyme, despite the fact that these two halves are at least largely formed from the same polypeptide chains (McHenry *et al.* 1987). In the T4 system, there is no evidence for such asymmetry, and recent studies have suggested a different solution to this problem, which will require testing by further experiments.

In these recent studies, the interactions between the T4 DNA polymerase holoenzyme and a primer–template DNA molecule were probed by DNA footprinting techniques (Munn 1986). A footprint was observed for the DNA polymerase at the 3′ terminus of the primer, as expected. However, the affinity of the polymerase for this site was weak, and it did not increase upon addition of the accessory proteins and ATP. To do such footprinting experiments, the T4 DNA polymerase must be kept in place ('stalled') at a single site, which is done by omitting one of the four deoxyribonucleoside triphosphates. These experiments therefore show that a stalled molecule of T4 DNA polymerase is very weakly bound to the DNA and that it cannot be tied down by addition of the accessory protein complex. Yet kinetic studies show that the same accessory proteins greatly stabilize a moving DNA polymerase molecule, with both the polymerase and one molecule of the 44/62 protein seeming to be very tightly bound at each primer template terminus (Piperno & Alberts 1978; Bedinger & Alberts 1983). We have therefore proposed that the replication complex is designed to dissociate from the template in the absence of a moving polymerase. The assembly of the holoenzyme complex is thought to begin when the primer–template junction binds an ATP-bound accessory protein complex, activating its ATP hydrolysis. The polymerase then interacts with this complex, positioning itself at the primer–template junction and displacing the accessory proteins to an adjacent position on the duplex (Munn 1986). The resulting accessory protein complex, which presumably remains bound to ADP + Pi, is postulated to be in a high-energy state that clamps down a moving DNA polymerase molecule. However, when the polymerase stalls, the decay of this complex leads to its dissociation from the DNA molecule, which explains our inability to detect stabilization of the polymerase by accessory proteins in footprinting analyses.

In the above view, active polymerase translocation causes the periodic reactivation of the accessory protein complex, thereby keeping it 'spring-loaded' in the high-energy states that are required to keep the polymerase tightly bound (Selick *et al.* 1987). As soon as the polymerase pauses, however, the activated accessory protein complex begins to decay through a series of intermediate energy states, each characterized by a particular half-life. So long as the pause

time is less than the sum of the half-lives for these decays, the accessory protein complex will be restored to its original high-energy state as soon as the polymerase restarts (presumably by ATP hydrolysis) and the polymerase will therefore remain tightly bound. For longer pauses, however, such as when the polymerase holoenzyme encounters the 5′ end of a previously synthesized Okazaki fragment during lagging strand DNA synthesis, when it encounters a region of DNA damage over which it cannot replicate, or when it is stalled on the primer template junction by the omission of a nucleoside triphosphate (as in footprinting experiments), the polymerase accessory protein complex decays to its lowest energy state and dissociates, allowing the DNA polymerase to leave the template rapidly.

The postulated decay process that begins when the polymerase pauses represents a clock that enables the conformational changes in the accessory proteins to measure the time of polymerase stalling, thereby converting the polymerase from a tight-binding- to a weak-binding enzyme after a minimum stalling time is exceeded. Such a timing mechanism is biologically useful, because it allows the lagging-strand polymerase to dissociate rapidly when it finishes its synthesis. The clock mechanism also prevents the unnecessary dissociation of a moving DNA polymerase that has only transiently paused, thereby allowing the leading strand DNA polymerase molecule to move processively for many minutes.

It is the unidirectional nature of energetically favourable reactions that is capable of measuring time, and, as pointed out by Hopfield (1974), any biological clock mechanism requires an input of energy. Here, the energy required comes initially from the ATP hydrolysis that establishes the spring-loaded complex, and it is subsequently provided by coupling the reactivation to the polymerization and excision cycles of the DNA polymerase (probably involving new rounds of ATP hydrolysis). The usefulness of such a clock, which enables the holoenzyme complex to distinguish between significant and insignificant barriers to its continued translocation, is that it allows the polymerase to shift between loose and tight binding states, as appropriate. Assuming that this mechanism is used in T4, it would not be surprising to find that it also operates in many other replication systems. Indeed, ATP-dependent steps have been characterized in the assembly and function of the *E. coli* DNA polymerase holoenzyme (Kornberg 1982; McHenry 1985); whether they have a related function will require further studies of both the T4 and *E. coli* enzyme complexes.

Some general implications of DNA polymerase recycling at the replication fork

The control of the rate of dissociation of a stalled DNA polymerase molecule just discussed is not a major problem in a simple two-dimensional replication fork, because DNA polymerase molecules, like RNA primers, can be left to accumulate behind the fork for a while (see figure 2). However, in a recycling model the polymerase dissociation rate becomes an important determinant of the length of Okazaki fragments. If the type of model in figure 5 is correct, a mutant polymerase holoenzyme that dissociates abnormally slowly from DNA should increase the Okazaki fragment size; finding and characterizing such mutants therefore represents a crucial test of these ideas.

There are several other general implications of a polymerase recycling model. Without recycling, the same DNA polymerase molecule that starts an Okazaki fragment can in principle finish it completely by waiting until the RNA primer at the 5′ end of the previous fragment is

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

removed (see figure 2). With recycling, the mobile replication apparatus is expected to lay down Okazaki fragments that still contain their attached RNA primers; because there is no time for the fragment joining to occur in the complex, the joining must involve separate DNA polymerase molecules that fill in the gaps created by RNA primer removal. Thus the DNA made on the lagging strand is thought to be completed by a separate DNA repair reaction, involving RNA excision, gap filling, and strand sealing by DNA ligase, which occurs after the fork has passed each point on the chromosome.

A third implication concerns the nature of fork initiation and termination processes. For T4 bacteriophage, the replication apparatus consists of a large moving complex of seven polypeptide chains, in which the entire DNA replication fork is embedded. The proteins have a total mass of at least 900 kDa, not including about 5000 kDa of 32 protein that are bound to the regions of single-stranded DNA at a fork. As a group, the proteins at the fork proceed unidirectionally along the DNA at a rate of about 400 nucleotides per second, faithfully replicating both strands of the template DNA helix as they go. In spirit therefore, the DNA replication apparatus is like a tiny sewing machine, composed of protein parts and powered by several different types of nucleoside triphosphate hydrolyses. Replication therefore involves the assembly of such a machine at a replication origin, and its specific disassembly at the replication terminus. For most replication origins, two replication forks are formed and move in opposite directions outwards, stopping only when they collide with another fork (Huberman & Riggs 1968; Inman & Schnös 1971). Two separate replication machines must therefore be formed at these bidirectional origins, and their disassembly must be triggered when two forks that are moving in opposite directions collide. Just as fork initiation centres on loading the primosome onto DNA (see figure 1), it is likely that the process of fork termination requires poorly understood steps of primosome disassembly.

DNA replication inside a cell requires additional proteins

The seven T4 proteins thus far described, when mixed with pure double-stranded DNA, seem to be able to catalyse replication fork movement with *in vivo* rates and fidelities. But it is clear that these proteins constitute only the core of the replication machinery. For example, the forks in the *in vitro* system are initiated at random nicks in the DNA double helix (Morris *et al.* 1975; Nossal & Peterlin 1979), whereas they are initiated at special replication origins inside the cell (Kozinski 1983). Additional T4 replication proteins, including initiator proteins analogous to the *E. coli* dnaA protein or the λ O protein, are almost certainly required for the biological fork initiation process.

There are also less obvious deficiencies in the T4 *in vitro* system. Thus replication inside the cell must occur on DNA to which many other proteins, including RNA polymerase, are bound. It was therefore surprising to find that a single RNA polymerase molecule bound to its promoter sequence completely blocks the replication fork in the T4 *in vitro* system (Bedinger *et al.* 1983). Similarly, a transcribing molecule of RNA polymerase stops the fork in a head on collision while slowing the fork rate to the rate of transcription when approached from behind (Jongeneel *et al.* 1984*c*). DNA replication on such templates *in vitro* requires an eighth T4 protein, the product of the *dda* gene (Bedinger *et al.* 1983; Jongeneel *et al.* 1984*c*). The dda protein, like the 41 protein, is a DNA helicase that can utilize the energy of nucleotide hydrolysis to unwind the helix in front of the growing fork (Krell *et al.* 1979;

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Jongeneel *et al.* 1984*a*, *b*). Unlike the 41 protein, however, the dda protein functions in a non-processive manner, rapidly dissociating and reassociating with a DNA strand, and it is unable to interact with the 61 protein to form a primosome. But this helicase possesses the unique ability to push the fork past bound RNA polymerase molecules, whether the replication complex is colliding head-on with, or travelling in the same direction as, their growing RNA chains (Bedinger *et al.* 1983; Jongeneel *et al.* 1984*c*). This 'snowplough' effect of the dda helicase seems to be a general one, because it has recently been shown to be required for the rapid passage of the replication complex through a lac operator DNA sequence that is bound by the lac repressor protein (J. Barry, unpublished results). Because the dda protein moves in the same direction on DNA as the 41 protein, we suggest that it runs just ahead of the 41 protein on the lagging strand, thereby pushing other template-bound protein molecules out of the way as the fork advances.

Other replication proteins must exist that are less well defined. For example, we have already discussed how the self-correcting nature of DNA polymerase allows it to correct most of its base-pairing errors. But other accidents must occur during intracellular DNA synthesis that are not so easily corrected. Among these are 'geometrical errors', such as the turning around of the DNA polymerase molecule on the leading strand, which spell disaster for the chromosome being replicated. Inside the cell such geometrical errors must either be stringently avoided or immediately recognized and reversed. In either case, special replication fork proteins are likely to be required; how they function is unknown.

The next section discusses an important event in intracellular DNA replication that is not yet understood: what happens at the fork when its leading strand encounters a damaged region in the template that cannot be replicated directly, such as a thymine dimer or some other lesion incompatible with base pairing? Here, additional proteins are required to allow the growing chain end to switch to a second DNA template molecule.

Linked recombination and DNA synthesis reactions

Many bulky lesions in DNA such as thymine dimers disrupt DNA replication *in vivo*, presumably because they cause the replication machinery to stall (Setlow *et al.* 1963; Moore *et al.* 1981; Villani *et al.* 1978). The replication machinery is often able to proceed past such damaged DNA eventually and continue normal replication without the repair of the lesion. Because *E. coli* can live with a large number of UV-induced pyrimidine dimers in its genome without a major increase in replication errors (Walker 1985), any mechanism used for trans-lesion DNA synthesis must produce the strand complementary to the dimer by copying an undamaged region of a second DNA template. Genetic studies in *E. coli* suggest that this type of DNA synthesis requires genetic recombination functions, including the direct participation of the *E. coli* recA protein (see Friedberg 1985). The linkage between recombination and DNA synthesis is also likely to be important in other aspects of DNA metabolism (Meselson & Radding 1975).

In T4, a single-stranded 3′ extension will remain on the end of a replicated chromosome, because an Okazaki fragment cannot start at the very tip as a DNA chain (Watson 1972; Broker 1973). The complete copying of this end is thought to require DNA synthesis that is templated from a homologous region of double helix (Mosig 1983; Dannenberg & Mosig 1983) in a reaction that involves synapsis with the end. Also, the undamaged genetic information from several heavily damaged copies of the T4 genome that coinfect a cell can be

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

efficiently pooled to produce one good copy: a phenomenon called 'multiplicity reactivation' (Luria & Dulbecco 1949). Both the mode of T4 terminal replication and the efficiency of multiplicity reactivation suggest that this bacteriophage produces an efficient system for the coupling of replication to recombination events.

The T4 analogue of the *E. coli* recA protein (Flory *et al.* 1984) is the uvsX protein, a DNA-dependent ATPase that binds cooperatively to DNA to form long filaments (Fujisawa *et al.* 1985; Yonesaki *et al.* 1985; Griffith & Formosa 1985). The uvsX protein catalyses two important reactions: the synapsis between a single strand and a homologous region of double helix to form a 'D-loop', and a directional branch migration that leads to strand exchange (Yonesaki *et al.* 1985; Yonesaki & Minagawa 1985; Hinton & Nossal 1986; Formosa & Alberts 1986*a*). In characterizing *in vitro* DNA synthesis reactions, we have discovered that this protein also plays a central role in a novel, conservative form of DNA synthesis (Formosa & Alberts 1986*b*).

A CONSERVATIVE FORM OF *IN VITRO* DNA SYNTHESIS

In *in vitro* reactions, DNA synthesis can be initiated by priming from the base-paired 3′OH DNA chain end in a D-loop. Formation of this intermediate *in vitro* requires, in addition to homologous single- and double-stranded DNA molecules, ATP hydrolysis by the uvsX protein

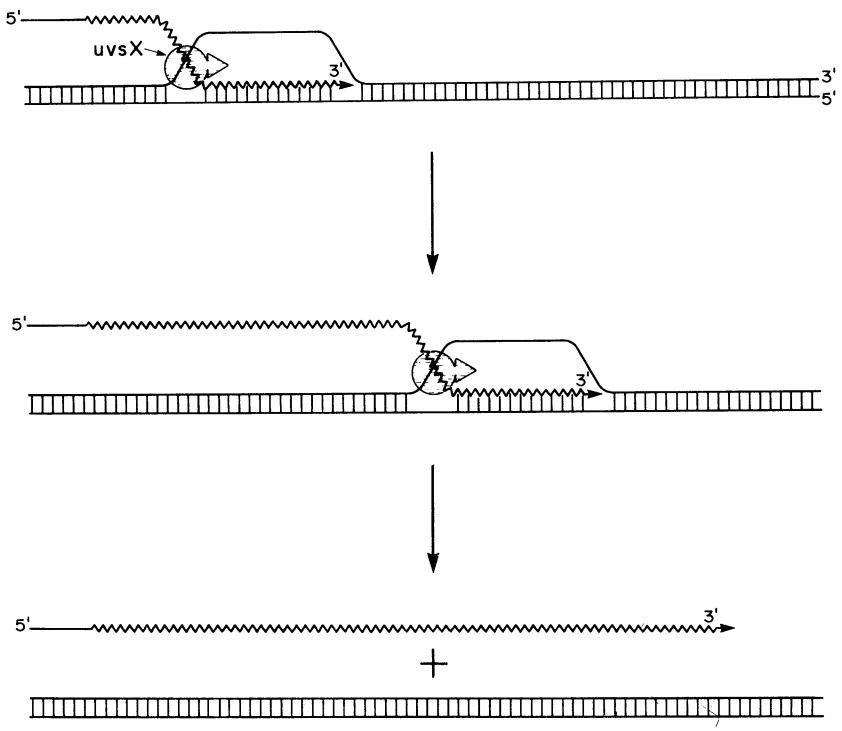

FIGURE 6. The conservative nature of synapsis-dependent DNA synthesis. DNA synthesis begins at a D-loop that is formed by uvsX protein-mediated synapsis. The trailing edge of the D-loop bubble is continuously reannealed because of branch migration catalysed by the uvsX protein, as shown. Strand separation at the leading edge of the bubble is caused by DNA synthesis, and the reannealing at the trailing edge creates a migrating 'replication bubble' that maintains a constant size as long as the strand-separation and reannealing reactions proceed at equivalent rates, as shown at top and centre. However, if the reannealing reaction rate exceeds the rate of DNA synthesis, as observed at high uvsX protein concentrations, the replication bubble shrinks, and DNA synthesis halts when the growing strand is displaced from the template (bottom). The termination of chain elongation is expected to occur predominantly at those sites on the DNA template where the DNA polymerase tends to pause, and it releases a free single strand.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

and the presence of the 32 protein. Upon addition of the T4 DNA polymerase holoenzyme (43, 44/62 and 45 proteins) and the *dda* DNA helicase, an efficient conservative form of DNA synthesis is initiated. Here the D-loop bubble moves rapidly along the DNA, and the new DNA strand being elongated is released as a DNA single strand. As in RNA synthesis, the DNA double helix being used as the template is left unperturbed. This type of reaction, which requires the cooperation of recombination and replication enzymes, seems likely to be a general feature of DNA metabolism (Formosa & Alberts 1986*b*).

One would expect the directional branch migration reaction catalysed by the uvsX protein to reanneal the back of the D-loop, as observed. Moreover, when *in vitro* reactions are done in the presence of intracellular concentrations of the uvsX protein, the reannealing at the back of the moving 'replication bubble' is fast enough to catch up with the moving DNA polymerase molecule, periodically collapsing the bubble and causing DNA synthesis to be transient (Formosa & Alberts 1986*b*). Because the reaction studied *in vitro* tends to elongate the 3′ ends of a single strand by copying a homologous double-stranded template only transiently, an elongated single strand is released as the major product (figure 6). As for the DNA replication fork, specific protein complexes are involved (Formosa & Alberts 1984), and the T4 uvsX protein cannot be substituted by its *E. coli* homologue, the recA protein.

A model for trans-lesion DNA synthesis

The above results suggest that an accurate bypass of DNA damage by trans-lesion DNA synthesis may involve a 'copy choice' type of DNA synthesis, in which an encounter with damage on the leading strand triggers a transient switch to a second DNA template. In the model illustrated in figure 7, whenever leading-strand DNA synthesis stalls, spontaneous branch migration displaces the 3′ end of the growing DNA chain and thereby converts it to a single-stranded form (step 1). As a single-strand, this 3′ end eventually invades a homologous DNA sequence on another DNA template and becomes extended by the conservative form of DNA synthesis just described. Because this synthesis is transient, it releases the 3′ end in an elongated form (step 2). Protein-directed branch migration (which acts to 'pin down' the 3′ end and release the 5′ end for both the *E. coli* recA protein and the T4 uvsX protein) will then rapidly reanneal the newly synthesized DNA strand back to the damaged DNA template (step 3). Once the 3′ end has been moved past the DNA damage in this way, the normal replication fork can restart, having accurately bypassed the initial lesion.

During multiplicity reactivation, the conservative DNA synthesis shown in figure 7 would automatically cease at any site of DNA damage on a second template, releasing a free single strand and forcing the synthesis to switch back to the original template. In this way, a single undamaged DNA molecule could eventually be generated by many cycles of template switching, which would have the effect of summing only the good nucleotide sequences present in different DNA molecules. Although the product is genetically a recombinant (as observed), extensive breakage and rejoining of the parental DNA strands is not required.

Although differing in detail, several earlier models for trans-lesion DNA synthesis contain similar features (Higgins *et al.* 1976; Fujiwara & Tatsumi 1976; Clark & Volkert 1978). An attractive feature of the Clark and Volkert model is that the damage bypass occurs only after the undamaged parental DNA strand (the lagging strand) has been copied past the point of

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

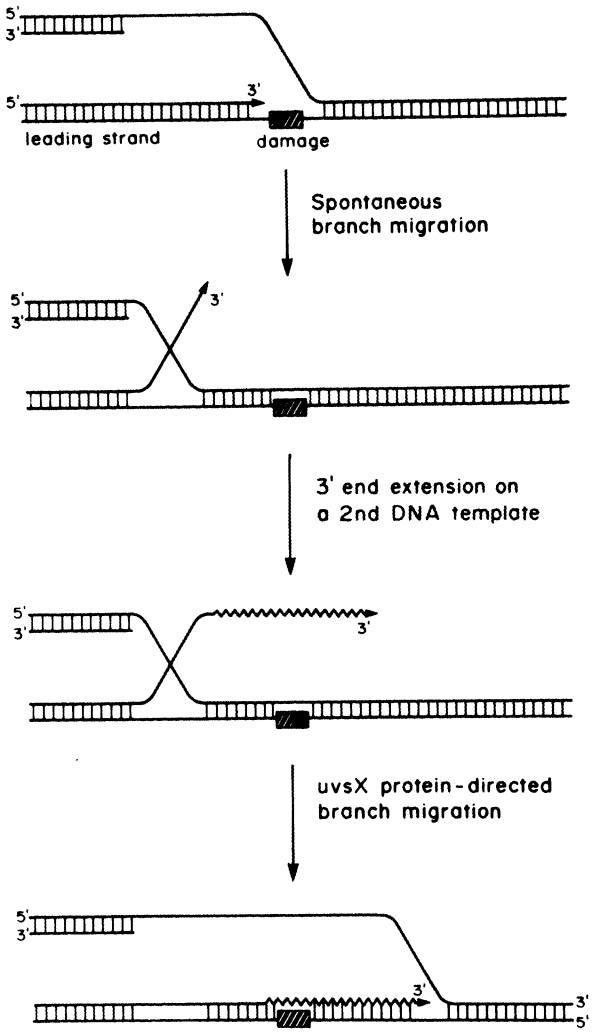

FIGURE 7. A model for trans-lesion DNA synthesis. DNA damage is encountered by the leading-strand DNA replication complex, causing both leading- and lagging-strand DNA synthesis to halt. After the 3′ end of the leading strand is displaced as a single strand by spontaneous branch migration, this 3′ end is transiently extended by synapsis-dependent synthesis on a second DNA template (as in figure 6). The model requires that an intact double-stranded copy of the damaged sequence be present in the cell to serve as this second template. Neither the second template nor the position of the primosome is indicated on this diagram. (From Formosa & Alberts (1986*b*).)

damage, which would make the daughter DNA double helix available as the template used for conservative DNA synthesis. (Note that in this case, leading- and lagging-strand synthesis must become uncoupled, with lagging-strand synthesis temporarily extending ahead of leading-strand synthesis). Although possibly not important in a T4-infected cell, where multiple copies of each DNA sequence should be present to serve as the template, such a switch to the lagging strand helix as the template best explains trans-lesion DNA synthesis in higher eukaryotic cells, where finding a homologous chromosome is difficult, owing to the complexity of the genome and the low copy number of most sequences.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Is recombination-initiated DNA synthesis involved in telomere maintenance?

Telomeres are specialized DNA sequences occurring at the ends of linear eukaryotic chromosomes; they are required to stabilize linear DNA molecules, presumably by providing a means for completely copying their ends (see Blackburn & Szostak 1984). The mechanism that we have described for 3′ end extension by conservative DNA synthesis presumably represents at least part of the mechanism that T4 bacteriophage uses to solve its end replication problem. In eukaryotes, a similar mechanism could be used, in which the 3′ end of each linear DNA chromosome is extended by a synapsis-initiated synthesis that uses either another homologous telomere or a repeated DNA sequence in the same telomere as the template. Once the 3′ end is extended in this way, retrograde synthesis could be primed on the resulting single-stranded tail to complete the end of the double helix.

DNA can regulate its own synthesis by controlling the proteins that function with DNA polymerase

Many different forms of DNA synthesis can occur in a cell, each with a different biological function. To control and catalyse these reactions requires a network of interacting proteins, whose nature and complexity are only beginning to be appreciated. Inside a cell infected with a T4 bacteriophage, for example, both DNA replication and DNA repair utilize the same DNA polymerase molecule, the product of T4 gene 43 (Bernstein & Wallace 1983). Yet DNA replication and DNA repair are very different types of processes. Much of the synthesis in DNA repair is thought to involve simple gap filling by the DNA polymerase, and it is important to avoid the type of DNA strand-displacement reaction that can generate a complete replication fork. Consider the most frequent DNA repair event in the cell, the resealing of adjacent Okazaki fragments on the lagging strand. If even a small fraction of these repair reactions continues too far and creates a replication fork, innumerable new forks would form on each lagging strand, making the DNA replication process hopelessly complex.

Figure 8 illustrates different types of protein complex containing the T4 DNA polymerase that are formed at different sites on a DNA molecule. Those special DNA sequences that direct the assembly of the T4 primosome onto a DNA strand are replication origins (Kreuzer & Alberts 1985) and they start replication forks (figure 8*a*). Other sites on the DNA allow the DNA polymerase to become associated with the uvsX protein and the *dda* DNA helicase; the ensuing DNA synthesis can produce single-stranded DNA by a conservative form of DNA synthesis (figure 8*b*). Finally, other types of sites on the DNA, including each junction between a pair of unsealed Okazaki fragments, will bind a DNA polymerase molecule that is prevented from using any type of DNA helicase. This DNA polymerase is unable to perform strand-displacement DNA synthesis inside the cell, and it falls off the DNA after completing gap repair (figure 8*c*).

The above example provides one illustration of how a single protein (the T4 DNA polymerase) can participate in several different types of discrete protein complexes. Another clear case is that of the T4 gene 32 protein, which plays major roles in DNA replication, DNA repair, and genetic recombination (Alberts & Frey 1971). This helix-destabilizing protein has been shown by both genetic (Mosig *et al.* 1979) and biochemical (Formosa *et al.* 1983) methods

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

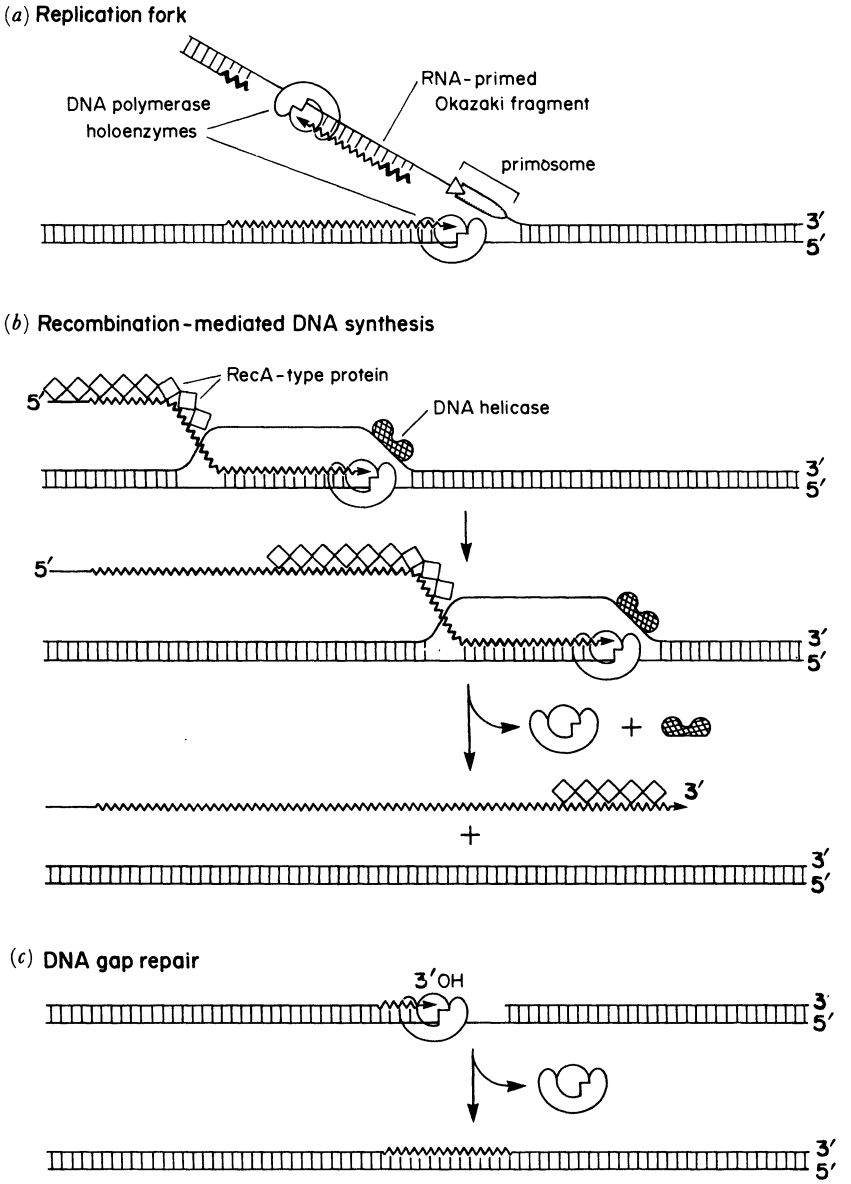

FIGURE 8. Assembly of the T4 DNA polymerase into different replication protein complexes at different sites on the DNA. In (*a*) a specific DNA sequence corresponding to a DNA replication origin directs the assembly of the T4 primosome onto the lagging strand. As a result, a true replication fork is established with both leading- and lagging-strand DNA synthesis; for simplicity, the fork is illustrated here in its unfolded form. In (*b*) a D-loop DNA structure has allowed a DNA polymerase molecule to co-assemble into a protein complex that contains the uvsX protein and the dda DNA helicase. Because this DNA helicase is unable to participate in RNA primer synthesis, the DNA product remains single-stranded. The directional branch migration catalysed by the uvsX protein, here assumed to be caused by its directional assembly onto a DNA single strand (see Register & Griffith 1985), produces the conservative form of DNA synthesis indicated. In (*c*), a typical DNA repair reaction is shown in which no DNA helicase is involved. The helix-destabilizing protein also participates in all three of those reactions (not shown).

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

to interact with a large number of different T4 proteins. These interactions are thought to reflect alternate associations that the 32 protein undergoes in a variety of different multiprotein complexes.

An understanding of eukaryotes will require model systems

In the T4 bacteriophage-infected cell, which is unusually simple, there are likely to be about 50 proteins in the interacting set that catalyse DNA replication, DNA repair and genetic recombination events at early times of infection (Alberts 1984). To date, the functions of about 15 of these proteins are understood at an elementary level. What the other 30–40 proteins do is not known, although many possibilities can be imagined. As just one example, the model for trans-lesion DNA synthesis in figure 7 neglects the primosome, which presumably must be kept in place on the lagging strand during the entire process. In T4, the chromosomal region located between the two genes that encode the primosome components (genes 61 and 41) contains open reading frames for five unknown proteins (Selick *et al.* 1987). These proteins have been highly conserved during the evolution of the T-even bacteriophages, and one or more of them could be required to alter the primosome in a way that makes trans-lesion DNA synthesis possible. Mutations in each of these genes have been made by reverse genetic techniques to test for these and other functions (H. E. Selick, unpublished results).

In my opinion, further detailed studies of relatively simple organisms like T4 bacteriophage will be crucial for understanding the basic genetic processes in eukaryotes. In the T4 chromosome, nature has concentrated the genes for a large number of interesting proteins that function together on DNA. The impending completion of the nucleotide sequence of the entire chromosome will provide ready access to a network of interacting genetic mechanisms that have been developed and refined over the course of billions of years. Although there is still much to be done, the combination of technologies available makes assigning both genes and functions to all of the proteins that are involved in T4 DNA replication, repair and recombination a reasonable and accessible goal. The complete description of such a system is an important task because once one thinks of molecular genetics in terms of 'protein machines' (rather than in terms of sequential reactions carried out by single proteins), finding and studying every piece of a multienzyme complex and understanding the relation of each complex to others on the DNA becomes essential for fully appreciating the mechanisms involved.

The work from my laboratory reviewed here was supported by grant GM24020 from the National Institutes of Health. It reflects the efforts of the many co-workers listed on the publications cited from this laboratory, whom I thank for their many contributions. I also acknowledge Dr Art Sussman for helpful comments on this manuscript.

References

Abdel-Monem, M. & Hoffmann-Berling, H. 1976 Enzymatic unwinding of DNA. I. Purification and properties of a DNA-dependent ATPase from *E. coli*. *Eur. J. Biochem.* **65**, 431–440.

Alberts, B. M. 1984 The DNA enzymology of protein machines. *Cold Spring Harb. Symp. quant. Biol.* **49**, 1–12.

Alberts, B. M., Amodio, F. J., Jenkins, M., Gutmann, E. D. & Ferris, F. L. 1969 Studies with DNA–cellulose chromatography. I. DNA-binding proteins from *Escherichia coli*. *Cold Spring Harb. Symp. quant. Biol.* **33**, 289–305.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Alberts, B. M. & Frey, L. 1971 T4 bacteriophage gene 32: a structural protein in the replication and recombination of DNA. *Nature, Lond.* **227**, 1313–1318.

Alberts, B., Morris, C. F., Mace, D., Sinha, N., Bittner, M. & Moran, L. 1975 Reconstruction of the T4 bacteriophage DNA replication apparatus from purified components. In *DNA synthesis and its regulation*, vol. 3 of ICN–UCLA Symposium on Molecular and Cellular Biology (ed. M. Goulian, P. Hanawalt & C. F. Fox), pp. 241–269. Menlo Park, California: W. H. Benjamin.

Alberts, B. & Sternglanz, R. 1977 Recent excitement in the DNA replication problem. *Nature, Lond.* **269**, 655–661.

Alberts, B. M., Barry, J., Bedinger, P., Burke, R. L., Hibner, U., Liu, C.-C. & Sheridan, R. 1980 Studies of replication mechanisms with the T4 bacteriophage in vitro system. In *Mechanistic studies of DNA replication and genetic recombination*, vol. 19 of ICN–UCLA Symposium on Molecular and Cellular Biology (ed. B. M. Alberts), pp. 449–473. New York: Academic Press.

Alberts, B. M., Barry, J., Bedinger, P., Formosa, T., Jongeneel, C. V. & Kreuzer, K. N. 1983 Studies on DNA replication in the T4 bacteriophage *in vitro* system. *Cold Spring Harb. Symp. quant. Biol.* **47**, 655–668.

Arai, K. & Kornberg, A. 1981 Unique primed start of phage φX174 DNA replication and mobility of the primosome in a direction opposite chain synthesis. *Proc. natn. Acad. Sci. U.S.A.* **78**, 69–73.

Barry, J. & Alberts, B. 1972 *In vitro* complementation as an assay for new proteins required for bacteriophage T4 DNA replication: purification of the complex specified by T4 genes 44 and 62. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2717–2721.

Bedinger, P. & Alberts, B. M. 1983 The 3′–5′ proofreading exonuclease of bacteriophage T4 DNA polymerase is stimulated by other T4 DNA replication proteins. *J. biol. Chem.* **258**, 9649–9656.

Bedinger, P., Hochstrasser, M., Jongeneel, C. V. & Alberts, B. M. 1983 Properties of the T4 bacteriophage DNA replication apparatus: the T4 dda DNA helicase is required to pass a bound RNA polymerase molecule. *Cell* **34**, 115–123.

Bernstein, C. & Wallace, S. S. 1983 DNA repair. In *Bacteriophage* T4 (ed. C. Mathews, E. Kutter, G. Mosig & P. Berget), pp. 138–151. Washington, D.C.: American Society for Microbiology.

Blackburn, E. H. & Szostak, J. W. 1984 The molecular structure of centromeres and telomeres. *A. Rev. Biochem.* **53**, 163–194.

Broker, T. R. 1973 An electron microscopic analysis of pathways for bacteriophage T4 DNA recombination. *J. molec. Biol.* **81**, 1–16.

Brutlag, D. & Kornberg, A. 1972 Enzymatic synthesis of deoxyribonucleic acid: a proofreading function for the 3′ to 5′ exonuclease activity in DNA polymerase. *J. biol. Chem.* **247**, 241–248.

Burke, R. L., Alberts, B. M., & Hosoda, J. 1980 Proteolytic removal of the COOH terminus of the T4 gene 32 helix-destabilizing protein alters the T4 in vitro replication complex. *J. biol. Chem.* **255**, 11484–11493.

Burke, R. L., Munn, M., Barry, J. & Alberts, B. M. 1985 Purification and properties of the bacteriophage T4 gene 61 RNA priming protein. *J. biol. Chem.* **260**, 1711–1722.

Cha, T.-A. & Alberts, B. M. 1986 Studies on the DNA helicase – RNA primase unit from bacteriophage T4. *J. biol. Chem.* **261**, 7001–7010.

Clark, A. J. & Volkert, M. R. 1978 A new classification of pathways repairing pyrimidine dimer damage in DNA. In *DNA repair mechanisms*, vol. 9 (ed. P. C. Hanawalt, E. C. Friedberg & C. F. Fox), pp. 57–72. London and New York: Academic Press.

Dannenberg, R. & Mosig, G. 1983 Early intermediates in bacteriophage T4 DNA replication and recombination. *J. Virol.* **45**, 813–831.

Epstein, R. H., Bolle, A., Steinberg, C., Kellenberger, E., Boy de la Tour, E., Chevalley, R., Edgar, R., Susman, M., Denhardt, C. & Lielausis, I. 1964 Physiological studies of conditional lethal mutants of bacteriophage T4D. *Cold Spring Harb. Symp. quant. Biol.* **28**, 375–392.

Flory, S. S., Tsang, J., Muniyappa, K., Bianchi, M., Gonda, D., Kahn, R., Azhderian, E., Egner, C., Shaner, S. & Radding, C. M. 1984 Intermediates in homologous pairing promoted by recA protein and correlations of recombination *in vitro* and *in vivo*. *Cold Spring Harb. Symp. quant. Biol.* **49**, 513–523.

Formosa, T., Burke, R. L. & Alberts, B. M. 1983 Affinity purification of T4 bacteriophage proteins essential for DNA replication and genetic recombination. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2442–2446.

Formosa, T. & Alberts, B. M. 1984 The use of affinity chromatography to study proteins involved in bacteriophage T4 genetic recombination. *Cold Spring Harb. Symp. quant. Biol.* **49**, 363–370.

Formosa, T. & Alberts, B. M. 1986*a* Purification and characterization of the T4 bacteriophage uvsX protein. *J. biol. Chem.* **261**, 6107–6118.

Formosa, T. & Alberts, B. M. 1986*b* DNA synthesis dependent on genetic recombination: characterization of a reaction catalyzed by purified bacteriophage T4 proteins. *Cell* **47**, 793–806.

Friedberg, E. C. 1985 *DNA repair*. San Francisco: W. H. Freeman.

Fujisawa, H., Yonesaki, T. & Minagawa, T. 1985 Sequence of the T4 recombination gene, uvsX, and its comparison with that of the recA gene of *Escherichia coli*. *Nucl. Acids Res.* **13**, 7473–7481.

Fujiwara, Y. & Tatsumi, M. 1976 Replicative bypass repair of ultraviolet damage to DNA of mammalian cells: caffeine sensitive and caffeine resistant mechanisms. *Mut. Res.* **37**, 91–110.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Fuller, R. S. & Kornberg, A. 1983 Purified *dnaA* protein in initiation of replication at the *E. coli* chromosomal origin of replication. *Proc. natn. Acad. Sci. U.S.A.* **80**, 5817–5821.

Furth, M. E. & Wickner, S. H. 1983 Lambda DNA replication. In *Lamba II* (ed. R. W. Hendrix, J. W. Roberts, F. W. Stahl & R. Weisberg), pp. 145–173. New York: Cold Spring Harbor Laboratory.

Griffith, J. & Formosa, T. 1985 The uvsX protein of bacteriophage T4 arranges single-stranded and double-stranded DNA into similar helical nucleoprotein filaments. *J. biol. Chem.* **260**, 4484–4491.

Hibner, U. & Alberts, B. M. 1980 The fidelity of DNA replication catalyzed in vitro on a natural DNA template by the T4 bacteriophage multienzyme complex. *Nature, Lond.* **285**, 300–305.

Higgins, N. P., Kato, K. & Strauss, B. 1976 A model for replication repair in mammalian cells. *J. molec. Biol.* **101**, 417–425.

Hinton, D. M. & Nossal, N. G. 1986 Cloning of the bacteriophage T4 uvsX gene and purification and characterization of the T4 uvsX recombination protein. *J. biol. Chem.* **261**, 5663–5673.

Hopfield, J. J. 1974 Kinetic proofreading: a new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proc. natn. Acad. Sci. U.S.A.* **71**, 4135–4139.

Huang, C.-C., Hearst, J. E. & Alberts, B. M. 1981 Two types of replication proteins increase the rate at which T4 DNA polymerase traverses the helical regions in a single-stranded DNA template. *J. biol. Chem.* **256**, 4087–4094.

Huberman, J. A., Kornberg, A. & Alberts, B. M. 1971 Stimulation of T4 bacteriophage DNA polymerase by the protein product of T4 gene 32. *J. molec. Biol.* **62**, 39–52.

Huberman, J. A. & Riggs, A. D. 1968 On the mechanism of DNA replication in mammalian chromosomes. *J. molec. Biol.* **32**, 327–341.

Irman, R. B. & Schnös, M. 1971 Structure of branch points in replicating DNA: presence of single-stranded connections in lambda DNA branch parts. *J. molec. Biol.* **56**, 319–325.

Jongeneel, C. V., Bedinger, P. & Alberts, B. M. 1984*a* Effects of the bacteriophage T4 dda protein on DNA synthesis catalyzed by purified T4 replication proteins. *J. biol. Chem.* **259**, 12933–12938.

Jongeneel C. V., Formosa, T. & Alberts, B. M. 1984*b* Purification and characterization of the bacteriophage T4 dda protein, a DNA helicase that associates with the viral helix destabilizing protein. *J. biol. Chem.* **259**, 12925–12932.

Jongeneel, C. V., Formosa, T., Munn, M. & Alberts, B. M. 1984*c* Enzymological studies on the T4 bacteriophage DNA replication proteins. In *Proteins involved in DNA replication* (ed. U. Hübscher & S. Spadari) *Adv. exp. Med. Biol.* vol. 179, pp. 17–33. New York: Plenum Publishing Corporation.

Kornberg, A. 1980 *DNA replication*. San Francisco: W. H. Freeman.

Kornberg, A. 1982 *1982 supplement to DNA replication*. San Francisco: W. H. Freeman.

Kornberg, A., Baker, T. A., Bertsch, L. L., Bramhill, D., Funnell, B. E., Lasken, R. S., Maki, H., Maki, S., Sekimizu, K. & Wahle, E. 1987 Enzymatic studies of replication of *oriC* plasmids. In *DNA replication and recombination*, vol. 19 of ICN–UCLA Symposium on Molecular and Cellular Biology (ed. T. Kelly & R. McMacken), pp. 137–149. New York: Alan R. Liss.

Kozinksi, A. W. 1983 DNA metabolism in vivo: origins of T4 DNA replication. In *Bacteriophage T4* (ed. C. Mathews, E. Kutter, G. Mosig & P. Berget), pp. 111–119. Washington, D.C.: American Society for Microbiology.

Krell, H., Durwald, H. & Hoffman–Berling, H. 1979 A DNA-unwinding enzyme induced in bacteriophage-T4-infected *Escherichia coli* cells. *Eur. J. Biochem.* **93**, 387–395.

Kreuzer, K. N. & Alberts, B. M. 1985 A defective phage system reveals bacteriophage T4 replication origins that coincide with recombination hotspots. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3345–3349.

Kurosawa, Y. & Okazaki, T. 1979 Structure of the RNA portion of the RNA-linked DNA pieces in bacteriophage T4-infected E. coli cells. *J. molec. Biol.* **135**, 841–861.

LeBowitz, J. H. & McMacken, R. 1986 The *Escherichia coli* dnaB replication protein is a DNA helicase. *J. biol. Chem.* **261**, 4738–4748.

Lee, M. Y., Tan, C. K., Downey, K. M. & So, A. G. 1984 Further studies on calf thymus DNA polymerase delta purified to homogeneity by a new procedure. *Biochemistry, Wash.* **23**, 1906–1913.

Liu, C. C. & Alberts, B. M. 1980 Pentaribonucleotides of mixed sequence are synthesized and efficiently prime *de novo* DNA chain starts in the T4 bacteriophage DNA replication system. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5698–5702.

Liu, C.-C. & Alberts, B. M. 1981 Characterization of the DNA-dependent GTPase activity of T4 gene 41 protein, an essential component of the T4 bacteriophage DNA replication apparatus. *J. biol. Chem.* **256**, 2813–2820 (1981).

Liu, C.-C., Burke, R. L., Hibner, U., Barry, J. & Alberts, B. M. 1979 Probing DNA replication mechanisms with the T4 bacteriophage *in vitro* system. *Cold Spring Harb. Symp. quant. Biol.* **43**, 469–487.

Luria, S. E. & Dulbecco, R. 1949 Genetic recombination leading to production of active bacteriophage from ultraviolet inactivated bacteriophage particles. *Genetics* **43**, 93–125.

Mace, D. C. & Alberts, B. M. 1984*a* The complex of T4 bacteriophage gene 44 and 62 replication proteins form an ATPase that is stimulated by DNA and by T4 gene 45 protein. *J. molec. Biol.* **177**, 279–293.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Mace, D. C. & Alberts, B. M. 1984*b* T4 DNA polymerase: rates and processivity on single-stranded DNA templates. *J. molec. Biol.* **177**, 295–311.

Mace, D. C. & Alberts, B. M. 1984*c* Characterization of the stimulatory effect of T4 gene 45 protein and the gene 44/62 protein complex on DNA synthesis by T4 DNA polymerase. *J. molec. Biol.* **177**, 313–327.

McHenry, C. 1985 DNA polymerase III holoenzyme of *E. coli*: components and functions of a true replication complex. *Molec. cell. Biochem.* **66**, 71–85.

McHenry, C. S., Oberfelder, R., Johanson, K., Tomasiewicz, H. & Franden, M. A. 1987 Structure and mechanism of the DNA polymerase III holoenzyme. In *DNA replication and recombination*, vol. 47 of ICN–UCLA Symposium on Molecular and Cellular Biology (ed. T. Kelly & R. McMacken), pp. 47–61. New York: Alan R. Liss.

McMacken, R., Alfano, C., Gomes, B., LeBowitz, J. H., Mensa-Wilmot, K., Roberts, J. D. & Wold, M. 1987 Biochemical mechanisms in the initiation of bacteriophage λ DNA replication. In *DNA replication and recombination*, vol. 47 of ICN–UCLA Symposium on Molecular and Cellular Biology (ed. T. Kelly & R. McMacken), pp. 227–245. New York: Alan R. Liss.

McMacken, R., Veda, K. & Kornberg, A. 1977 Migration of *E. coli* dnaB protein on the template DNA strand as a mechanism in initiating DNA replication. *Proc. natn. Acad. Aci. U.S.A.* **74**, 4190–4194.

Meselson, M. S. & Radding, C. M. 1975 A general model for genetic recombination. *Proc. natn. Acad. Sci. U.S.A.* **72**, 358–361.

Moore, P. D., Bose, K. K., Rabkin, S. D. & Strauss, B. S. 1981 Sites of termination of *in vitro* DNA synthesis on ultraviolet- and *N*-acetylaminofluorene-treated ϕX174 templates by prokaryotic and eukaryotic DNA polymerases. *Proc. natn. Acad. Sci. U.S.A.* **78**, 110–114.

Morris, C. F., Hama-Inaba, H., Mace, D., Sinha, N. K. & Alberts, B. 1979*a* Purification of the gene 43, 44, 45 and 62 proteins of the bacteriophage T4 DNA replication apparatus. *J. biol. Chem.* **254**, 6787–6796.

Morris, C. F., Moran, L. A. and Alberts, B. M. 1979*b* Purification of gene 41 protein of bacteriophage T4. *J. biol. Chem.* **254**, 6797–6802.

Morris, C. F., Sinha, N. K. & Alberts, B. M. 1975 Reconstruction of the bacteriophage T4 DNA replication apparatus from purified components: rolling circle replication following *de novo* chain initiation on a single-stranded circular DNA template. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4800–4808.

Mosig, G., Luder, A., Garcia, G., Dannenberg, R. & Bosk, S. 1979 In vivo interactions of genes and proteins in DNA replication and recombination of phage T4. *Cold Spring Harb. Symp. quant. Biol.* **43**, 501–515.

Mosig, G. 1983 Relationship of T4 DNA replication and recombination. In *Bacteriophage T4* (ed. C. Matthews, E. Kutter, G. Mosig & P. Berget), pp. 120–130. Washington, D.C.: American Society for Microbiology.

Munn, M. 1986 Analysis of the bacteriophage T4 DNA replication complex. Ph.D. dissertation, University of California.

Newport, J. W., Kowalczykowski, S. C., Lonberg, N., Paul, L. S. & von Hippel, P. H. 1980 Molecular aspects of the interactions of T4-encoded gene 32-protein and DNA polymerase (gene 43-protein) with nucleic acids. In *Mechanistic Studies of DNA Replication and Genetic Recombination*, vol. 19 of ICN–UCLA Symposium on Molecular and Cellular Biology (ed. B. M. Alberts), pp. 485–505. New York: Academic Press.

Nossal, N. G. & Alberts, B. M. 1983 The mechanism of DNA replication catalyzed by purified bacteriophage T4 DNA replication proteins. In *Bacteriophage T4* (ed. C. Mathews, E. Kutter, G. Mosig & P. Berget), pp. 77–81. Washington, D.C.: American Society for Microbiology.

Nossal, N. G. 1980 RNA priming of DNA replication by bacteriophage T4 proteins. *J. biol. Chem.* **255**, 2176–2182.

Nossal, N. G. & Peterlin, B. M. 1979 DNA replication by bacteriophage T4 proteins: the T4 43, 32, 44–62, and 45 proteins are required for strand displacement synthesis at nicks in duplex DNA. *J. biol. Chem.* **254**, 6032–6037.

Piperno, J. R. & Alberts, B. M. 1978 An ATP stimulation of T4 DNA polymerase mediated via T4 gene 44/62 and 45 proteins: the requirement for ATP hydrolysis. *J. biol. Chem.* **253**, 5174–5179.

Register, J. C. & Griffith, J. 1985 The direction of recA protein assembly onto single strand DNA is the same as the direction of strand assimilation during strand exchange. *J. biol. Chem.* **260**, 12308–12312.

Scheuermann, R. & Echols, H. 1984 A separate editing exonuclease for DNA replication. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7747–7751.

Selick, H. E., Barry, J., Cha, T.-A., Munn, M., Nakanishi, M., Wong, M. L. & Alberts, B. M. 1987 Studies on the T4 bacteriophage DNA replication system. In *DNA replication and recombination*, vol. 47 of ICN–UCLA Symposium on Molecular and Cellular Biology (ed. T. Kelly & R. McMacken), pp. 183–214. New York: Alan R. Liss.

Setlow, R. B., Swenson, P. A. & Carrier, W. L. 1963 Thymine dimers and inhibition of DNA synthesis by ultraviolet irradiation of cells. *Science, Wash.* **142**, 1464–1466.

Sinha, N. K., Morris, C. F. & Alberts, B. M. 1980 Efficient *in vitro* replication of double-stranded DNA templates by a purified T4 bacteriophage replication system. *J. biol. Chem.* **255**, 4290–4303.

Tsurimoto, T. & Matsubara, K. 1984 Multiple initiation sites of DNA replication flanking the origin region of λdv genome. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7402–7406.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES

Venkatesan, M., Silver, L. L. & Nossal, N. G. 1982 Bacteriophage T4 gene 41 protein, required for the synthesis of RNA primers, is also a DNA helicase. *J. biol. Chem.* **257**, 12426–12434.
Villani, G., Boiteux, S. & Radman, M. 1978 Mechanisms of ultraviolet-induced mutagenesis: extent and fidelity of *in vitro* DNA synthesis on irradiated templates. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3037–3041.
Walker, G. C. 1985 Inducible DNA repair systems. *A. Rev. Biochem.* **54**, 425–457.
Watson, J. D. 1972 Origin of concatameric T7 DNA. *Nature, new Biol.* **239**, 197–201.
Wood, W. B. & Revel, H. R. 1976 The genome of bacteriophage T4. *Bact. Rev.* **40**, 847–868.
Yonesaki, T. & Minagawa, T. 1985 T4 phage gene uvsX product catalyzes homologous DNA pairing. *EMBO J.* **4**, 3321–3327.
Yonesaki, T., Ryo, Y., Minagawa, T. & Takahashi, H. 1985 Purification and some of the functions of the products of the bacteriophage T4 recombination genes, uvsX and uvsY. *Eur. J. Biochem.* **148**, 127–134.
Zyskind, J. W., Cleary, J. M., Brusilow, W., Harding, N. & Smith, D. W. 1983 Chromosomal replication origin from the marine bacterium *Vibrio harveyi* functions in *E. coli*: *oriC* consensus sequence. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1164–1168.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B BIOLOGICAL SCIENCES