Reprinted from: MOLECULAR CYTOGENETICS
Edited by Barbara Ann Hamkalo
and John Papaconstantinou
Book available from: Plenum Publishing Corporation
227 West 17th Street, New York, N.Y. 10011

19. STUDIES ON THE REPLICATION OF DNA

Bruce Alberts

Department of Biochemical Sciences, Princeton University, Princeton, New Jersey

The goal of our work is to understand the basic mechanism of DNA replication, with special emphasis on the protein and nucleic acid interactions involved. It appears that replication is carried out by a complex of several proteins (Barry and Alberts, 1972; Schekman *et al.*, 1972; reviewed in Klein and Bonhoeffer, 1972), quite possibly interwoven with the DNA to create a particle with a unique three-dimensional conformation (Alberts, 1971). However, attempts to purify intact DNA replication forks with these proteins attached have thus far failed, apparently reflecting a basic lability of the entire structure (Fuchs and Hanawalt, 1970; Miller and Buckley, 1970; C. Manoil, R. Peniston and B. Alberts, in preparation). This lability of the "replication apparatus" greatly complicates *in vitro* studies of replication mechanisms (see, for example, Schaller *et al.*, 1972). In this context, it is relevant to note how the comparative stability of ribosome and RNA polymerase protein complexes has enabled striking advances in our knowledge of the biochemistry of translation and transcription.

Inspired by the reconstitution of functional ribosomal particles from their individual components (Traub and Nomura, 1968), we have decided to approach the replication problem by individually isolating each of the proteins of one DNA-replication apparatus, trusting that we can eventually get this structure to self-assemble *in vitro* by incubating intracellular concentrations of these proteins with DNA. With this goal in mind, we have focused our attention on the T4 bacteriophage replication system. Its major advantage is that the genetics of T4 replication have been well worked out, beginning with the isolation and characterization of conditional-lethal mutants by Epstein and Edgar and their colleagues (Epstein *et al.*,

PRINTED IN U.S.A.

1963). In addition, about 60 replication forks are established in each infected cell (Werner, 1968); this makes the T4 system a relatively rich source for biochemical isolation of replication proteins, especially with the recent discovery of a T4 mutant which overproduces some of these gene products (J. S. Wiberg *et al.*, manuscript submitted; J. Karam, personal communication).

E. coli cells infected with T4 bacteriophage carrying an amber mutation in gene 32, 41, 43, 45, 44, or 62 synthesize little or no DNA, even though all four deoxyribonucleotide triphosphates are present (Epstein *et al.*, 1963; Warner and Hobbs, 1967; Kozinski and Felgenhauer, 1967; Mathews, 1972). In addition, temperature-sensitive mutants have been isolated for the products of genes 32, 41, 43 and 45, which cause T4 DNA replication to cease within one minute after a shift from a permissive temperature (25°C) to a non-permissive temperature (42°C) (Riva *et al.*, 1970). In fact, replication forks stop completely before moving even one-fifth the length of one mature T4 genome following the shift to 42°C (M. Curtis and B. Alberts, in preparation). Each of these four proteins therefore appears to be involved directly in the DNA polymerization process. The other two gene products (genes 44 and 62) appear to function as a single tight complex of two polypeptide chains (Barry and Alberts, 1972). Phages carrying temperature-sensitive mutations in gene 44 replicate their DNA normally at 42°C if allowed 10 minutes of prior infection at 30°C (Riva *et al.*, 1970; J. Karam, personal communication). It is thus possible that these two gene products function only in the initial stages of T4 DNA synthesis. However, no firm conclusion can be drawn in this respect, especially since the altered polypeptide may be stabilized to heat inactivation as soon as the 44-62 complex forms.

Two of the T4 replication gene products have been extensively characterized. The product of gene 43 is T4 DNA polymerase (DeWaard *et al.*, 1965; Warner and Barnes, 1966) and the product of gene 32 is a DNA-unwinding protein, "32-protein" (Alberts and Frey, 1970; Delius *et al.*, 1972). The gene for DNA polymerase maps in a cluster of replication genes (41, 43, 62, 44, and 45) which surround the T4 replication origin as defined by Mosig and coworkers (Marsh *et al.*, 1971; but see Delius *et al.*, 1971). This tight clustering suggests that these gene products interact with each other (Stahl, 1967); likewise, their close linkage to the replication origin suggests that at least one of these proteins acts with sequence specificity in the initiation of replication forks.

Before proceeding with a discussion of our current work on the protein products of T4 genes 41 and 45 and the 44-62 complex, several properties of DNA polymerases and DNA-unwinding proteins will be reviewed. When combined with data from *in vivo* studies, these properties enable a general view of DNA replication to be derived; this will be useful when considering possible functions for these other proteins, whose mode of action is as yet unknown.

PROPERTIES OF T4 DNA POLYMERASE

The T4 DNA polymerase consists of a single polypeptide chain of 110,000 daltons; its functional properties (Aposhian and Kornberg, 1962; Goulian *et al.*, 1968) are essentially identical to those of the replicative polymerase in uninfected cells, *E. coli* DNA polymerase III (Kornberg and Gefter, 1972). *In vitro* DNA synthesis by these polymerases requires a single-stranded DNA template chain which is paired along part of its length with a shorter complementary chain of DNA (or RNA); the free 3'-OH terminus of this second chain then acts as "primer" (Bollum, 1964). An example of one such template is shown in Fig. 1A.

Three important *in vitro* properties of polymerase which we shall stress when analyzing its *in vivo* function follow. 1) All chain growth is in the 5' to 3' chain-direction only. Each polymerization step involves attack of a 3'-OH of a polymer chain end on the activated 5' end of an incoming deoxyribonucleotide 5' triphosphate monomer. The β and γ phosphates of the monomer are released as pyrophosphate with formation of the new phosphodiester bond. 2) The polymerase cannot start a new DNA chain, since it cannot join two monomers together. DNA polymerases always begin by extending a pre-existing "primer" chain. 3) The T4 DNA polymerase will not utilize double-helical templates. In addition to a primer chain, the polymerase has an absolute requirement for a template DNA strand and the segment to be copied must be single-stranded.

A crucial further requirement for *in vitro* polymerization was discovered recently by Brutlag and Kornberg (1972), who showed that, for polymerization, both T4 and *E. coli* polymerases have an absolute requirement for a Watson-Crick base-paired residue at the 3'OH primer terminus. When confronted with a template with a terminal mismatch (Fig. 1B), these polymerases make use of their built-in 3' → 5' exonuclease activities to clip off unpaired primer residues by hydrolysis; this continues until enough nucleotides are removed to generate an active template (*i.e.*, to convert the template in Fig. 1B to that in Fig. 1A; see also Englund, 1971). As a result, these polymerases will efficiently remove their own polymerization errors. The apparent consequence is fantastic accuracy: one mistake in approximately 10^{10} base-pair replications in *E. coli* (Drake, 1969). Since the self-correcting feature allows polymerase to select for proper template pairing by each added nucleoside triphosphate in two separate reactions - forward and backwards, the mistake frequency at each step could be as large as 1 in 10^5 in this example

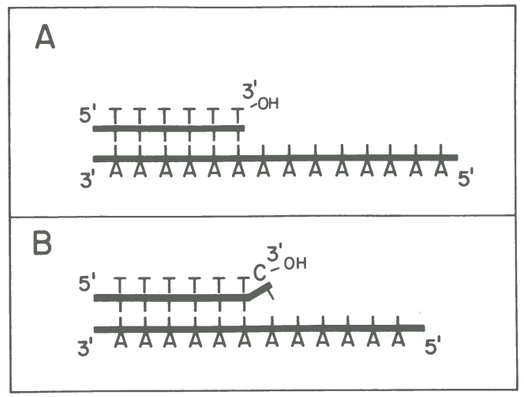

Figure 1. Active and inactive synthetic polynucleotide templates for DNA polymerases. A. Template with base-paired primer terminus (active). B. Template with mismatched primer terminus (inactive) (from Brutlag and Kornberg, 1972).

($10^{-5} \times 10^{-5} = 10^{-10}$). Note that whereas an error frequency of 10^{-10} would require 14 Kcal/mole in discrimination energy if it had to be attained in the polymerization step alone, as little as 7 Kcal/mole would be needed in a two-step mechanism.[1]

I would like to propose that development of such a two-step base-selection mechanism was essential to provide an error frequency in replication low enough to allow any organism with a large number of essential genes to survive (for arguments concerning maximum mutation load, see Crow and Kimura, 1970).[2] Accepting this hypothesis,

---

[1] This is calculated from the standard relationship $W_1/W_2 = \exp(-\Delta G/RT)$, where ΔG is the difference in free energy between $state_1$ and $state_2$ ("discrimination energy"), and W_1 and W_2 are their relative probabilities.

[2] The self-correcting feature of DNA polymerase clearly requires that it reject templates with a mismatched primer terminus (as in Fig. 1B). However, the built-in 3' → 5' exonuclease activity of the prokaryotic polymerases is not essential, since other enzymes could remove the terminal mismatched nucleotide after the polymerase is releasted. In fact, this exonuclease activity appears to be greatly reduced or missing in several eukaryotic DNA polymerases examined (Chang and Bollum, 1973).

it becomes clear why no DNA polymerase is able to start a new polynucleotide chain *de novo*, whereas DNA-dependent RNA polymerases can do so: The necessity to be self-correcting gives DNA polymerase a catholic requirement for a perfectly base-paired primer terminus; to ask this enzyme to start synthesis in the complete absence of primer, without losing any of its discrimination between the two templates in Fig. 1, seems contradictory. In contrast, the RNA polymerases need not be self-correcting, inasmuch as relatively high error rates can be tolerated during transcription.

This line of reasoning likewise suggests an explanation for the failure to find a second DNA polymerase which adds deoxyribonucleotide 5' triphosphates in such a way as to cause chains to grow in the 3' → 5' direction. Despite the relatively simple type of replication fork mechanism that this would allow for such a 3' → 5' polymerase, the growing 5' chain end (rather than the incoming mononucleotide) carries the triphosphate activation. Thus, mistakes in polymerization cannot be hydrolyzed away without a special enzymatic system for reactivating the bare 5' chain end thus created.

PROPERTIES OF T4 DNA-UNWINDING PROTEIN

The T4 gene 32 protein (Alberts and Frey, 1970) has a monomer molecular weight of 35,000 daltons, and at high concentrations it can self-aggregate to form large multimeric complexes (Huberman *et al.*, 1971; Carroll *et al.*, 1972). It closely resembles a 22,000 dalton DNA-unwinding protein isolated from uninfected *E. coli* (Sigal *et al.*, 1972) and has properties which are similar, but not identical, to the gene 5 protein of filamentous bacteriophage (Alberts *et al.*, 1972; Oey and Knippers, 1972) and to helix-unwinding proteins isolated from calf thymus (Herrick and Alberts, 1973 and in preparation).

1) The gene 32-protein binds tightly to single-stranded DNA, but only weakly, if at all, to double-stranded DNA or to RNA. At saturation, one protein monomer of 35,000 daltons is bound per every 10 DNA nucleotides (a weight ratio of protein to DNA of 12:1); this holds the DNA strand in an extended conformation with a 4.6 Å translation distance per nucleotide (Delius *et al.*, 1972). A schematic diagram of this complex is shown in Fig. 2A; as indicated, the protein monomers must be in contact to account for the highly cooperative nature of their binding. The nucleotide bases in the complex appear to be left uncovered and exposed, with little (if any) base specificity in the binding.

2) Because of its specific affinity for single-stranded DNA, 32-protein catalyzes DNA denaturation. This process is schematically illustrated in Fig. 2B; by virtue of the coupled equilibria shown, the denaturation reaction is driven to the right, thus lowering the midpoint (T_m) of the helix-coil transition by about 40°C.

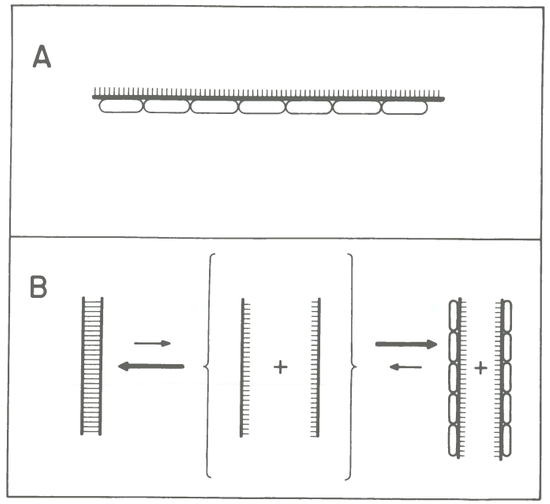

Figure 2. Schematic representation of DNA-unwinding protein interactions with DNA. A. Complex with single-stranded DNA. B. Shift in helix-melting equilibrium caused by DNA-binding specificity.

Since the T_m of T4 DNA under physiological salt conditions *in vitro* is about 85°C (Marmur and Doty, 1962), we would estimate that, with unwinding protein present, helical regions of T4 DNA should melt inside the cell at about 45°C. Thus, while regions of perfect helix should be transiently invaded by 32-protein *in vivo*, the reverse reaction will be faster than the forward reaction, so that an opened helical region rapidly reforms (Fig. 3A). On the other hand, single-stranded DNA is highly folded *in vitro* due to the formation of many short intrastrand "hairpin" helices (Doty *et al.*, 1960; Studier, 1969). Because these helices are weak (due to imperfect base-pairing), they are easy targets for 32-protein induced denaturation, and should be kept permanently open within the cell (Fig. 3B).

3) Using a DNA template like that shown in Fig. 1A, but produced from a natural DNA by limited exonuclease-III degradation, marked stimulations of T4 DNA polymerase by T4 32-protein and of *E. coli* DNA polymerase II by the 22,000 dalton *E. coli* DNA-unwinding protein were obtained. However, there was no noticeable effect of *E. coli* unwinding protein on the T4 polymerase, nor of T4 unwinding protein on the *E. coli* polymerase (Sigal *et al.*, 1972). This specificity may arise from the fact that the two DNA-unwinding proteins hold the DNA template single-strands in quite different conformations; at any rate, it strongly implies that polymerase and unwinding protein function together *in vivo*, as well as *in vitro*.

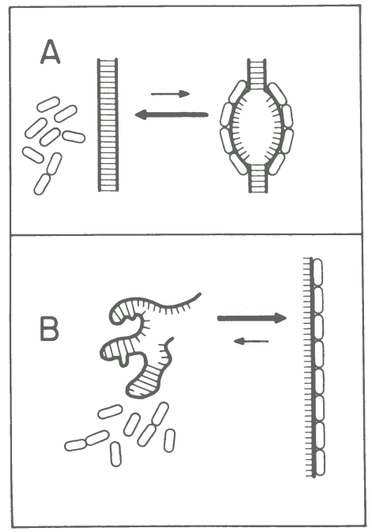

Figure 3. Helix-melting equilibria in the presence of T4 gene 32-protein under physiological conditions. A. Perfect double-helices. B. The imperfectly base-paired helices of single-stranded DNA.

MODELS FOR *IN VIVO* POLYMERIZATION AT THE REPLICATION FORK

In view of the strong genetic evidence that both the T4 DNA polymerase and DNA-unwinding protein are essential for the polymerization process, it is instructive to ask how these two proteins might work together in T4 replication forks. Our current view of a replication fork is based upon work in many laboratories, and is schematically illustrated in Fig. 4. Because of the antiparallel orientation of strands in the DNA double-helix, a discontinuous mode of DNA synthesis is necessary on one template strand if all of the DNA is to be made with a self-correcting, 5' → 3' DNA polymerase. Both biochemical analysis of initial DNA products (Sugino and Okazaki, 1972) and electron microscopy of forks (Inman and Schnos, 1971; Wolfson and Dressler, 1972) indicate a length of about 10^3 nucleo-

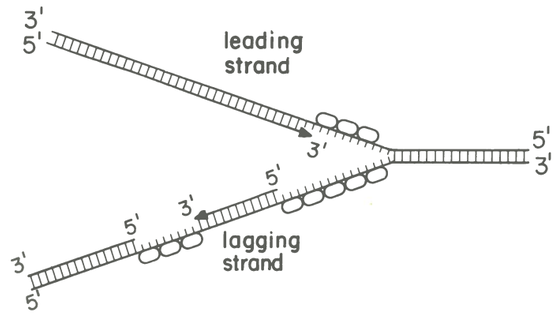

Figure 4. Schematic representation of possible DNA polymerase and unwinding protein cooperation in the T4 DNA replication fork.

tides between adjacent polymerase starts on the "lagging side" of the fork. On the "leading side" of the fork, synthesis can be either continuous (as indicated in Fig. 4) or discontinuous, perhaps depending on physiological conditions. However, the discontinuous synthesis on the leading side is best viewed as a competition between polymerase continuation and polymerase restart mechanisms (Olivera and Bonhoeffer, 1972); since such restarts do not appear to lead to gap formation in the daughter strand, this synthesis can be treated for our purposes as continuous.

A replication fork such as that in Fig. 4 necessarily contains substantial lengths of single-stranded DNA on its lagging side; as indicated, such DNA is almost certain to be covered with tightly bound DNA unwinding-protein. While it is possible that the role of the 32-protein in T4 DNA replication is solely to keep these single-stranded template strands aligned, it seems more likely that it also has a helix-opening role in leading strand synthesis. Some such destabilization of the helix ahead of the fork seems essential, in view of the fact that double-helical templates cannot be copied by the T4 DNA polymerase. An obvious objection to our proposal is that 32-protein should invade double-helical regions only with great difficulty inside the cell (see above). However, the pre-fork region may be much easier to open than intact helices, due to the fact that cooperative 32-protein binding can be nucleated on both the lagging strand 32-protein complexes (Fig. 4), and on the leading strand molecule of DNA polymerase (for which 32-protein has substantial affinity: Huberman *et al.*, 1971).

The model in Fig. 4 assumes that two DNA polymerase molecules are working at any one time in the fork; in such a model, synthesis

on the two sides of the fork could be tightly coupled by some physical linkage between the two polymerase sites. If so, the 32-protein molecules leaving the single-stranded template strand on the lagging side (as the polymerase passes) might be directly recirculated through the fork without mixing with cellular pools. Less speculatively, kinetic arguments can be made for direct reutilization of any 32-protein molecules used to unwind the helix ahead of the fork; in the model shown in Fig. 4, this is perhaps most readily accomplished if these protein molecules can slide, pushed by the polymerization behind (for alternative schemes, see Alberts, 1971).

What evidence is available to support this role for 32-protein at replication forks? Note that the model in Fig. 4 generates a total of one Okazaki-piece length (about 10^3 nucleotides) of single-stranded DNA on the lagging side of each fork. Since each 32-protein monomer binds 10 nucleotides, 100 molecules of 32-protein would be required per fork just to cover this DNA. Genetic experiments, in which 32-protein levels are varied by mixed infections, suggest that 100-200 molecules of 32-protein are indeed necessary to establish each T4 replication fork, in marked contrast to the much smaller amounts of all of the other replication gene products required (Snustad, 1968; Sinha and Snustad, 1971; reviewed by Alberts, 1971). Recent biochemical measurements on DNA-protein complexes isolated from T4-infected cells confirm these estimates; in addition, they demonstrate that the bound 32-protein molecules are present in long clusters, as predicted (C. Manoil, R. Peniston and B. Alberts, in preparation).

The model in Fig. 4 accounts for all of the properties of T4 DNA polymerase discussed earlier, except for its inability to start chains *de novo*. Okazaki and his coworkers have recently shown that in *E. coli* discontinuous synthesis is primed by a special piece of RNA (Sugino *et al.*, 1972; Sugino and Okazaki, 1973). This RNA chain is estimated to be 50-100 nucleotides long, and it appears to be synthesized *de novo* by a special, rifampicin-resistant enzyme. After it serves as a primer for DNA-polymerase, the RNA is erased and replaced by DNA; the specificity of the *E. coli* ligase for a DNA-DNA junction could prevent nick sealing until the last ribonucleotide is removed (Westergaard *et al.*, 1973). Such a cycle of RNA synthesis and degradation is schematically illustrated in Fig. 5.

Why does nature use such an elaborate mechanism to prime DNA polymerase, when a DNA primer would seem more economical? A likely answer follows directly from our previous discussions. The earlier argument that a self-correcting polymerase cannot start chains *de novo* also implies its converse: than an enzyme which starts chains *de novo* cannot do a good job of self-correcting. Thus, any enzyme which primes discontinuous synthesis will of necessity make a relatively inaccurate copy (*e.g.*, one error in 10^5?). Even if the

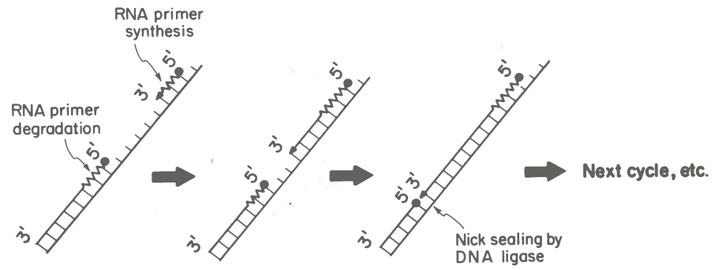

Figure 5. Lagging strand synthesis: a cycle of synthesis and degradation of the RNA primer needed to start Okazaki pieces (after Sugino and Okazaki, 1972). Enzymes likely to function in RNA removal included RNase H (Keller, 1972) and the 5' → 3' exonuclease activity of DNA polymerase I (Westergaard *et al.*, 1973).

amount of this copy which was retained in the final product constituted 1% or less of the total genome (10 nucleotides or less per Okazaki fragment), the resulting increase in overall mutation rate could be enormous. Thus, it seems reasonable to suggest that the use of ribonucleotides to synthesize primer was of great evolutionary advantage, since it automatically marked these sequences as "bad copy" to be removed.

In conclusion, it seems likely that the mechanism of DNA replication sketched earlier (Figs. 4 and 5), sometimes views as the whimsical result of historical accident, was specially selected to ensure that the *entire* DNA sequence is very accurately copied (*i.e.* copied by a self-correcting type of DNA polymerase).

MODELS FOR INITIATION OF NEW REPLICATION FORKS

We have seen that in *E. coli* the priming of Okazaki pieces is an event which occurs about once per second on each DNA strand (*i.e.* once per 10^3 nucleotides polymerized at 37°C), and that it appears to involve a special rifampicin-resistant RNA polymerase which functions only at replication forks. This frequent priming reaction should be clearly distinguished from a second process which also may involve RNA: the *initiation* of chromosomal DNA replication. In bacteria, this latter process is carefully regulated since it occurs only once per generation at a special DNA sequence which specifies the replication origin (Yoshikawa and Sueoka, 1963; reviewed in Lark, 1969). For several genomes, chromosomal initiation appears to require rifampicin-*sensitive* RNA synthesis at this origin (Brutlag *et al.*, 1971; Lark, 1972; Dove *et al.*, 1971; Hayes and Szybalski, this Symposium).

The work of several laboratories has led to the idea that initiation of chromosome replication in bacteria (and thereby the entire cell cycle) is regulated by the accumulation of a special "initiator protein" (reviewed in Helmstetter, 1969). Work with inhibitors suggests that this protein is stable. Since initiation seems to be triggered only when the number of initiator protein molecules *per origin* (X) reaches some threshold level (Helmstetter *et al.*, 1968; Donachie, 1968; Maaløe and Kjeldgaard, 1966), this protein appears to be directly titrated against (and thereby measured by) chromosome origins. Consideration of the statistical fluctuations expected in such a system leads to the conclusion that X is large (Alberts, 1970); a recent quantitative treatment suggests that in fact X must represent 200 initiator molecules per origin or more (Sompayrac and Maaløe, 1973).

In the simplest view, the initiator might be a protein which binds non-cooperatively to the DNA at chromosome origins, and which is needed in at least 200 copies per origin to open the double-helix and establish replication forks there. Since RNA synthesis at the origin is apparently needed for initiation as well, one model for this process is that sketched in Fig. 6. Here an RNA polymerase and a DNA-unwinding type protein are imagined to work in concert in helix opening. The resulting structure resembles the D-loop intermediate visualized in mitochondrial DNA synthesis (Robberson *et al.*, 1972). As indicated, this loop might need to reach a minimum size before subsequent steps can proceed (*e.g.*, in order to expose start sites on the protein-covered strand). In this case, the number of molecules of the unwinding protein could be carefully measured in relation to the number of DNA origins, and this protein thereby becomes a suitable candidate for the "initiator." In this context, it is interesting to note that the gene 32-protein is required in 100-200 copies per T4 fork, and that it is the only replication protein not made in excess in the T4 system (Snustad, 1968). Thus its availability presumably controls the number of T4 replication forks made at early times after infection, when the number of these forks per cell is increasing (Werner, 1968; reviewed in Alberts, 1971).

ATTEMPTS TO ASSEMBLE THE T4 REPLICATION APPARATUS *IN VITRO*

To assemble a replication apparatus *in vitro*, all of its essential protein and nucleic acid components must be available in active form. At a minimum, this means that (in addition to polymerase and unwinding protein) the protein products of T4 genes 41,

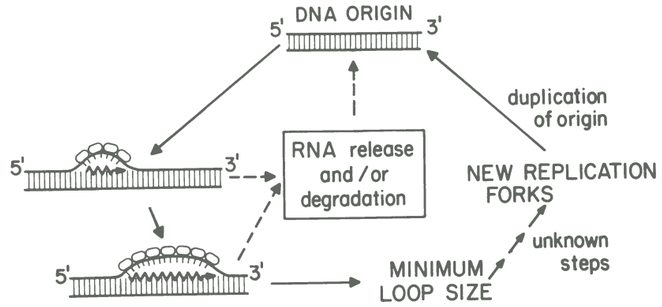

Figure 6. General model for initiator protein-triggering of chromosomal DNA replication. For most genomes the entire initiation process appears to occur without breakage of the parental DNA strands (Jaenisch *et al.*, 1971; Robberson *et al.*, 1972; Inman and Schnos, 1971; Delius *et al.*, 1971; Wolfson and Dressler, 1972). The "initiator" is visualized here as a DNA-unwinding type of protein. By binding tightly to single-stranded DNA, it stabilizes a DNA·RNA hybrid formed at the replication origin by an RNA polymerase. With only low levels of initiator present, the RNA made would be quickly released by strand displacement, and the short cycle indicated by dotted arrows would obtain (see Hayes and Szybalski, this Symposium; Buckley *et al.*, 1972). Eventually enough initiator accumulates per origin to achieve the "minimum loop size" needed to establish replication forks there and a round of DNA replication is triggered. Since such an initiator-binding loop must form only at the DNA origin, some protein component or nucleic acid structure in it must be sequence-specific.

44, 45, and 62 must be purified to the point where they are free of deleterious contaminants (*e.g.* nucleases). Some monitoring for retention of the functional integrity of each gene product seems essential during the purification. For this purpose, we have developed a crude *in vitro* synthesizing system which requires all of these T4-induced proteins for maximal synthesis. This system thereby provides a measure of the activity of these proteins, despite the fact that their function in replication is unknown.

Using such an "*in vitro* complementation" system as an assay, T4 gene 45 protein and a complex of gene 44 and 62 proteins have been purified to electrophoretic homogeneity in active form (Barry and Alberts, 1972; Barry *et al.*, 1973). More recently, active gene 41 protein has been highly purified as well (L. Moran, unpublished data). In all cases, double-label radioisotope techniques (with mixed phage amber-mutant lysates) have been used to confirm the identity of the isolated proteins.

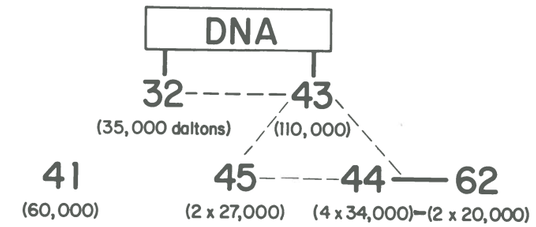

Figure 7. Schematic diagram of some interactions of purified T4 replication proteins with each other and with DNA. Bold numerals are used to denote the gene number of each protein; the subunit molecular weights listed beneath were determined by sodium dodecyl sulfate polyacrylamide gel electrophoresis.

Fig. 7 presents a preliminary diagram that summarizes our current knowledge of the physical properties of the known T4 replication proteins and their interactions with each other and with DNA. A solid connecting-line is used to denote a strong interaction, while a dotted line signifies a weaker affinity. The absence of a particular interaction in the diagram is not necessarily meaningful, since not all possibilities have been tested. Moreover, while none of the new T4 proteins appear to bind to DNA-cellulose columns under our conditions, some of them might do so if these conditions were changed.

The only evidence for the suggested interactions involving 44-62 and 45 proteins comes from their strongly synergistic effect in stimulating T4 DNA polymerase reactions on single-stranded DNA templates *in vitro* (D. Mace and J. Goldberg, unpublished data); the exact mode of action of these proteins in this stimulation is currently under investigation.

We have not yet been able to demonstrate catalysis of RNA primer synthesis by the T4-induced proteins, nor have we been able to obtain significant replication of a double-stranded T4 DNA template with them. It is possible that we have not yet found the proper *in vitro* conditions for their reassembly (see, for example, Lieberman and Oishi, 1973). Alternatively, we may still be missing some essential protein components. Likely candidates in this case include some T4-induced proteins [*e.g.*, the products of genes 42 or 1 (Collinsworth and Matthews, 1973; Chiu and Greenberg, 1968; R. Greenberg, personal communication), gene y (Smith and Symonds, 1973), or the DNA-delay genes (Yegian *et al.*, 1971)]. Although none of the 5 mutationally-defined *E. coli* replication gene products appear to be essential for T4 DNA replication (Gross, 1972; but

see Mosig *et al.*, 1972), a requirement for other host proteins is certainly possible (*e.g.*, the DNA-dependent RNA-polymerase).

It is our hope that the number of additional proteins needed for reconstitution of the T4 DNA-replication apparatus will be small, and that they can therefore be purified from crude extracts with the aid of appropriate assays. Progress along similar lines is being made in both the *E. coli* (Schekman *et al.*, 1972; Nusslein *et al.*, 1973; Wickner *et al.*, 1972) and T7 bacteriophage (Hinkle, 1973) replication systems. Through the combined efforts of all of these groups, a complete picture of the enzymology of DNA replication should emerge in the not too distant future.

ACKNOWLEDGMENT

I am indebted to my colleagues and students at Princeton for discussions and advice on this manuscript and to the National Institutes of Health and the American Cancer Society for grant support.

REFERENCES

Alberts, B. 1970. Function of gene 32-protein, a new protein essential for the genetic recombination and replication of T4 bacteriophage DNA. Fed. Proc. 29: 1154.

Alberts, B. 1971. On the structure of the replication apparatus. In (D. W. Ribbons, J. F. Woessner and J. Schultz, eds.) Nucleic Acid-Protein Interactions. p. 128. North Holland, Amsterdam.

Alberts, B. and L. Frey. 1970. T4 bacteriophage gene 32: a structural protein in the replication and recombination of DNA. Nature 227: 1313.

Alberts, B., L. Frey and H. Delius. 1972. Isolation and characterization of gene 5-protein of filamentous bacterial viruses. J. Mol. Biol. 68: 139.

Aposhian, H. V. and A. Kornberg. 1962. Enzymatic synthesis of deoxyribonucleic acid. IX. The polymerase formed after T2 bacteriophage infection of *E. coli*: a new enzyme. J. Biol. Chem. 237: 519.

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

Barry, J., H. Hama-Inaba, L. Moran, J. Wiberg and B. Alberts. 1973. Proteins of the T4 bacteriophage replication apparatus. In (R. D. Wells and R. B. Inman, eds.) DNA Synthesis *In Vitro*. University Park Press, Baltimore. In press.

Bollum, F. J. 1964. Chemically defined templates and initiators for deoxyribopolynucleotide synthesis. Science 144: 560.

Brutlag, D. and A. Kornberg. 1972. Enzymatic synthesis of DNA. XXXVI. A proofreading function for the 3' → 5' exonuclease activity in DNA polymerases. J. Biol. Chem. 247: 241.

Brutlag, D., R. Schekman and A. Kornberg. 1971. A possible role for RNA polymerase in the initiation of M13 DNA synthesis. Proc. Nat. Acad. Sci. 68: 2826.

Buckley, P. J., L. Kosturko and A. W. Kozinski. 1972. *In vivo* production of an RNA-DNA copolymer after infection of *E. coli* by bacteriophage T4. Proc. Nat. Acad. Sci. 69: 3165.

Carroll, R. B., K. E. Neet and D. Goldthwait. 1972. Self-association of gene-32 protein of bacteriophage T4. Proc. Nat. Acad. Sci. 69: 2741.

Chang, L. M. S. and F. J. Bollum. 1973. A comparison of associated enzyme activities in various DNA polymerases. J. Biol. Chem. 248: 3398.

Chiu, C. S. and G. R. Greenberg. 1968. Evidence for a possible direct role of dCMP hydroxymethylase in T4 page DNA synthesis. Cold Spring Harbor Symp. Quant. Biol. 33: 351.

Collinsworth, W. L. and C. K. Matthews. 1973. DNA synthesis in plasmolyzed T4-infected cells. Fed. Proc. 32: 491 Abs.

Crow, J. F. and M. Kimura. 1970. An Introduction to Population Genetics Theory. Harper and Row, New York.

Delius, H., C. Howe and A. W. Kozinski. 1971. Structure of the replicating DNA from bacteriophage T4. Proc. Nat. Acad. Sci. 68: 3049.

Delius, H., N. Mantell and B. Alberts. 1972. Characterization by electron microscopy of the complex formed between T4 bacteriophage gene 32-protein and DNA. J. Mol. Biol. 67: 341.

DeWaard, A., A. V. Paul and I. R. Lehman. 1965. The structural gene for DNA polymerase in bacteriophage T4 and T5. Proc. Nat. Acad. Sci. 54: 1241.

Donachie, W. D. 1968. Relationship between cell size and time of initiation of DNA replication. Nature 219: 1077.

Doty, P., H. Boedtker, J. Fresco, R. Haselkorn and M. Litt. 1959. Secondary structure in ribonucleic acids. Proc. Nat. Acad. Sci. 45: 482.

Dove, W. F., H. Inokuchi and W. F. Stevens. 1971. Replication control in phage lambda. In (A. D. Hershey, ed.) The Bacteriophage Lambda. p. 747. Cold Spring Harbor Laboratory, New York.

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

Englund, P. T. 1971. The initial step of *in vitro* synthesis of DNA by the T4 DNA polymerase. J. Biol. Chem. 246: 5684.

Epstein, R. H., A. Bolle, C. M. Steinberg, E. Kellenberger, E. Boy de la Tour, R. Chevallez, R. S. Edgar, M. Susman, G. H. Denhardt and A. Lielausis. 1963. Physiological studies of conditional lethal mutants of bacteriophage T4D. Cold Spring Harbor Symp. Quant. Biol. 28: 375.

Fuchs, E. and P. Hanawalt. 1970. Isolation and characteriztion of the DNA replication complex from *Escherichia coli*. J. Mol. Biol. 52: 301.

Goulian, M., Z. J. Lucas and A. Kornberg. 1968. Enzymatic synthesis of DNA. XXV. Purification and properties of DNA polymerase induced by infection with phage T4. J. Biol. Chem. 243: 627.

Gross, J. D. 1972. DNA replication in bacteria. Current Topics in Microbiol. and Immunol. 57: 39.

Helmstetter, C. 1969. Sequence of bacterial reproduction. Ann. Rev. Microbiol. 23: 223.

Helmstetter, C., S. Cooper, O. Pierucci and E. Revelas. 1968. On the bacterial life sequence. Cold Spring Harb. Symp. Quant. Biol. 33: 809.

Herrick, G. and B. Alberts. 1973. A nucleic acid helix-unwinding protein from calf thymus. Fed. Proc. 32: 497 Abs.

Hinkle, D. C. 1973. Bacteriophage T7 DNA replication *in vitro*. Fed. Proc. 32: 492 Abs.

Huberman, J., A. Kornberg and B. Alberts. 1971. The stimulation of T4 DNA polymerase by the protein product of T4 gene 32. J. Mol. Biol. 62: 39.

Inman, R. and M. Schnos. 1971. Structure of branch points in replicating DNA: presence of single-strand connections in lambda DNA brach points. J. Mol. Biol. 56: 319.

Jaenisch, R., A. Mayer and A. Levine. 1971. Replicating SV40 molecules contain closed circular template DNA strands. Nature New Biol. 233: 72.

Keller, W. 1972. RNA-primed DNA synthesis *in vitro*. Proc. Nat. Acad. Sci. 69: 1569.

Klein, A. and F. Bonhoeffer. 1972. DNA replication. Ann. Rev. Biochem. 41: 301.

Kornberg, T. and M. Gefter. 1972. DNA synthesis in cell-free extracts. IV. Purification and properties of DNA polymerase III. J. Biol. Chem. 247: 5369.

Kozinski, A. W. and Z. Felgenhauer. 1967. Molecular recombination in T4 bacteriophage DNA. II. Single-strand breaks and exposure of uncomplemented areas as a prerequisite for recombination. J. Virol. 1: 1193.

Lark, K. G. 1969. Initiation and control of DNA synthesis. Ann. Rev. Biochem. 38: 569.

Lark, K. G. 1972. Evidence for the direct involvement of RNA in the initiation of DNA replication in *E. coli* 15T⁻. J. Mol. Biol. 64: 47.

Lieberman, R. P. and Oishi, M. 1973. Formation of the *recB-recC* DNase by *in vitro* complementation and evidence concerning its subunit nature. Nature New Biol. 243: 75.

Maaløe, O. and N. O. Kjeldgaard. 1966. Control of Macromolecular Synthesis. W. A. Benjamin, Inc., New York.

Marmur, J. and P. Doty. 1962. Determination of the base composition of DNA from its thermal denaturation temperature. J. Mol. Biol. 5: 109.

Marsh, R. C., A. Breschkin and G. Mosig. 1971. Origin and direction of bacteriophage T4 DNA replication. II. A gradient of marker frequencies in partially replicated T4 DNA as assayed by transformation. J. Mol. Biol. 60: 213.

Mathews, C. K. 1972. Biochemistry of DNA-defective amber mutants of bacteriophage T4. III. Nucleotide pools. J. Biol. Chem. 247: 7430.

Miller, R. C. and P. Buckley. 1970. Early intracellular events in the replication of bacteriophage T4 DNA. VI. Newly synthesized proteins in the T4 protein-DNA complex. J. Virol. 5: 502.

Mosig, G., W. Bowden and S. Bock. 1972. *E. coli* DNA polymerase I and other host functions participate in T4 DNA replication and recombination. Nature New Biol. 240: 12.

Nüsslein, V., F. Bonhoeffer, A. Klein and B. Otto. 1973. In (R. Wells and R. Inman, eds.) DNA Synthesis *In Vitro*. University Park Press, Baltimore. In press.

Oey, J. L. and R. Knippers. 1972. Properties of the isolated gene 5 protein of bacteriophage fd. J. Mol. Biol. 68: 125.

Olivera, B. and F. Bonhoeffer. 1972. Discontinuous DNA replication *in vitro*. I. Two distinct size classes of intermediates. Nature New Biol. 240: 233.

Riva, S., A. Cascino and E. P. Geiduschek. 1970. Coupling of late transcription to viral replication in bacteriophage T4 development. J. Mol. Biol. 54: 85.

Robberson, D., H. Kasamatsu and J. Vinograd. 1972. Replication of mitochondrial DNA. Circular replicative intermediates in mouse L cells. Proc. Nat. Acad. Sci. 69: 737.

Schaller, H., B. Otto, V. Nüsslein, J. Huf, R. Herrman and F. Bonhoeffer. 1972. DNA replication *in vitro*. J. Mol. Biol. 63: 183.

Schekman, R., W. Wickner, O. Westergaard, D. Brutlag, K. Geider, L. Bertsch and A. Kornberg. 1972. Initiation of DNA synthesis: synthesis of ϕX174 replicative form requires RNA synthesis resistant to rifampicin. Proc. Nat. Acad. Sci. 69: 2691.

Sigal, N., H. Delius, T. Kornberg, M. Gefter and B. Alberts. 1972. A DNA-unwinding protein isolated from *E. coli*: its interaction with DNA and with DNA polymerases. Proc. Nat. Acad. Sci. 69: 3537.

Sinha, N. K. and D. P. Snustad. 1971. DNA synthesis in bacteriophage T4-infected *E. coli*: evidence supporting a stoichiometric role for gene 32-product. J. Mol. Biol. 62: 267.

Smith, S. M. and N. Symonds. 1973. The unexpected location of a gene conferring abnormal radiation sensitivity on phage T4. Nature 241: 395.

Snustad, D. P. 1968. Dominance interactions in *E. coli* cells mixedly infected with bacteriophage T4 wild-type and amber mutants and their possible implications as to type of gene-product function: catalytic vs. stoichiometric. Virology 35: 550.

Sompayrac, L. and O. Maaløe. 1973. Autorepressor model for control of DNA replication. Nature New Biol. 241: 133.

Stahl, F. W. 1967. Circular genetic maps. J. Cell. Physiol. 70, Suppl. 1: 1.

Studier, F. W. 1969. Conformational changes of single-stranded DNA. J. Mol. Biol. 41: 189.

Sugino, A., S. Hirose and R. Okazaki. 1972. RNA-linked nascent DNA fragments in *Escherichia coli*. Proc. Nat. Acad. Sci. 69: 1863.

Sugino, A. and R. Okazaki. 1972. Mechanism of DNA chain growth. VII. Direction and rate of growth of T4 nascent short DNA chains. J. Mol. Biol. 64: 61.

Sugino, A. and R. Okazaki. 1973. RNA-linked DNA fragments *in vitro*. Proc. Nat. Acad. Sci. 70: 88.

Traub, P. and M. Nomura. 1968. Structure and function of *E. coli* ribosomes. V. Reconstitution of functionally active 30S ribosomal particles from RNA and proteins. Proc. Nat. Acad. Sci. 59: 777.

Warner, H. R. and J. E. Barnes. 1966. DNA synthesis in *E. coli* infected with some DNA polymerase-less mutants of bacteriophage T4. Virology 28: 100.

Warner, H. R. and M. D. Hobbs. 1967. Incorporation of uracil-C^{14} into nucleic acids in *Escherichia coli* infected with bacteriophage T4 and T4 amber mutants. Virology 33: 376.

Werner, R. 1968. Initiation and propagation of growing points in the DNA of phage T4. Cold Spring Harbor Symp. Quant. Biol. 33: 501.

Westergaard, O., Brutlag, D. and A. Kornberg. 1973. Initiation of DNA synthesis. IV. Incorporation of the RNA primer into the phage replicative form. J. Biol. Chem. 248: 1361.

Wickner, R. B., M. Wright, S. Wickner and J. Hurwitz. 1972. Conversion of ϕX174 and fd single-stranded DNA to replicative forms in extracts of *Escherichia coli*. Proc. Nat. Acad. Sci. 69: 3233.

Wolfson, J. and D. Dressler. 1972. Regions of single-stranded DNA in growing points of replicating bacteriophage T7 chromosomes. Proc. Nat. Acad. Sci. 69: 2682.

Yegian, C. D., M. Mueller, G. Selzer, V. Russo and F. W. Stahl. 1971. Properties of the DNA-delay mutants of bacteriophage T4. Virology 46: 900.

Yoshikawa, H. and N. Sueoka. 1963. Sequential replication of *Bacillus subtilis* chromosome. I. Comparison of marker frequencies in exponential and stationary growth phases. Proc. Nat. Acad. Sci. 49: 559.