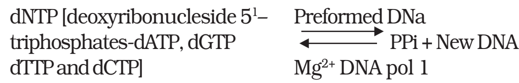
DNA Replication – A Unique Orchestration of Actions of Proteins

JAYDIP MANDAL
Reader in Botany
Regional Institution of Education
(NCERT), Bhopal

WATSON AND Crick's landmark discovery of DNA structure in April 1953 suggested how DNA might be replicated and preserved for hundreds of generations. The self-complementarity of two polynueleotide strands of a DNA double helix discarded the idea that protein template might play a role in DNA replication. The semi-conservative fashion of DNA replication was distinctively demonstrated by Matthew Meselson and Franklin Stahl in 1958. This clearly states that when one DNA molecule duplicates to produce two identical new DNA molecules, each of the new molecules is comprised of one polynucleotide strand from the original DNA molecule bonded to one newly synthesised polynucleotide strand. Replication begins either at a single point or multiple points on gene in procaryotes and cucaryotes respectively. This point of origin on the gene signifies a specific sequence of bases especially A-T rich bases rather than G-C based due to presence of fewer (two) hydrogen bonds in A-T base pair than G-C base pair. DNA synthesis occurs at the replication fork which appears like 'Y' denoting the site at which the DNA molecule is unwound into an open box and new strands are synthesised. The number of replication forks varies among the living organisms from 1 in *E. coli* circular chromosome to many in human linear chromosome.

Molecular Mechanism of DNA Replication

DNA replication involves a careful orchestration of the actions of different enzymes, DNA polymerase I and polymerase III and more than 30 proteins which act in concert to unwind the two polynucleotide strands of the DNA molecule, copy its template and produce two complete daugher DNA molecules. The initiation at an origin starts only after binding of specific major initiator protein dna A protein to the origin and thus, converts closed origin site of the double helix into open double helix. The other protein molecules like dna B, i.e. DNA helicase and dna C (helicase loader) join dna A to open the DNA double helix. Another protein, topoisomerase 11 or DNA gyrase hydrolyzes ATP, and help in unwinding of double helix by removing supertwists at origin site. Once the strands are separated, the unwound portion of DNA is then stabilised by single strand binding protein (SSB), which binds cooperatively to exposed single stranded DNA. The first nucleotide polymerising enzyme, DNA polymerase I was discovered by Arthur Kornberg and his colleagues in 1958 from *E.coli.* This enzyme requires all four deoxyribonucleoside 5^1 - triphosphates – dATP, dGTP, dTTP and dCTP and Mg^{2+} to synthesise DNA. DNA pol-I catalyses



step by step addition of doxyribonucleotide units to the 3[1]-0H end of a performed DNA chain or template DNA strand. Two more enzymes DNA Pol II and Pol III were discovered in 1969 by Paula DeLucia and John Cairns from *E.coli*, which could also analyse a template directed synthesis of DNA from precursors-deoxyribonucleoside 5[1]-triphosphates like that of DNA polymerase I. Now these enzymes have been shown to have clearly defined roles. A multisubunit assembly of 10 kinds of polypeptide chains called DNA polymerase III holoenzyme synthesises most new DNA, whereas DNA polymerase I remove the RNA primer and fills the gaps. DNA polymerase II participates in DNA repair, therefore, is not needed for DNA replication.

Events at Replication Fork: Leading verus Lagging Strand

At a replication fork, both strands of parental DNA serve as templates for the synthesis of both new daughter strands by an assembly complex that is the DNA polymerase III. As DNA structure reveals that parental strands are antiparallel, the overall direction of DNA synthesis must be $5^1 \rightarrow 3^1$ for one daughter strand and $3^1 \rightarrow 5^1$ for the other. In 1960, it was resolved by Reiji Okazaki who found that newly synthesised DNA exists as small fragments containing 1000-2000 nucleotides in dividing bacteria wheras only 100-200 nucleotides in eucaryotes. These fragments are called Okazaki fragments (after the discoverer Reijokazaki and Tsuneko Okazaki), which are polymerised discontinuously only in the $5^1 \rightarrow 3^1$ direction. The strand formed from joining of Okazaki fragments by DNA ligase is called lagging strand whereas the DNA strand that is synthesised continuously without interruption is the leading strand. The synthesis of leading strand precedes the synthesis of lagging strand (Fig. 1). Both the Okazaki fragments and the leading strand are synthesised in the $5^1 \rightarrow 3^1$ direction, though the overall growth of lagging strand if opposite to leading strand, i.e. in the $3^1 \rightarrow 5^1$ direction.

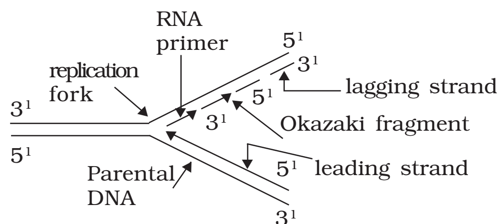


Fig. 1: *Diagram of a replication fork and events on it*

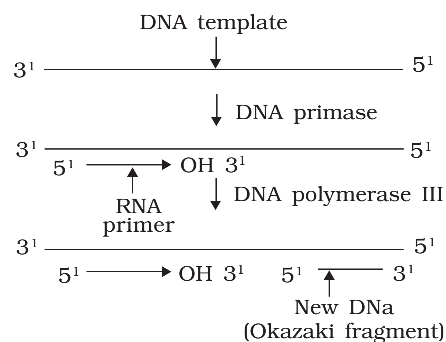


Fig. 2: *DNA replication is primed by a short RNA primer*

Initiation of DNA Synthesis

All DNA synthesis commences with the construction of RNA primer. In 1988, it was Kornberg who found that nascent DNA is covalently linked to a short stretch of RNA. A specialised enzyme called DNA primase or dnaG uses ribonucleoside triphosphates to synthesise short RNA primers on the lagging strand. In procaryotes, DNA primase synthesises a short stretch of RNA primer about 5 ribonucleotides long and are made at intervals of 1000-2000 deoxyribonucleotides whereas in eucaryotes these primers are about 10 ribonucleotides long that are made at intervals of 100-200 deoxyribonucleotides on the lagging strand. For the leading strand, a special RNA primer is needed only at the start of replication. RNA primer, which is complementary to one of the template DNA strands, primes the synthesis of DNA. In *E.coli*, major initiator protein dnaA complexes with preprimosome containing seven proteins n^1, n^{11}, nj, i, dnaB (DNA helicase and DNA unwinding enzyme), dnaC and dnaG (DNA primase) proteins to form a multisubunit assembly called primosome that enables synthesis of RNA primer and initiation of Okazaki fragments (Fig. 2). These protien components function in concert to ensure movement of primosome along the DNA strands, the displacement of SSBs, the recognition of appropriate start sites and the polymerisation of ribonucleotides into RNA primer.

Synthesis of leading and lagging strands: As the RNA primer contains a properly base paired nucleotide with a 3^1-OH group at one end, it can be elongated by the DNA polymerase III holoenzyme at this end to being syntehsis of an Okazaki fragment. The DNA polymerase III holoenzyme (DNA pol III) catalyses the formation of many thousands of phosphodiester bonds before releasing its template compared with only 20 phosphodiester bonds for DNA polymerase I (DNA pol I) thereby revealing capability of DNA pol III in DNA synthesis whereas the DNA pol I is associated with DNA repair. The synthesis of each Okazaki fragment is completed when DNA polymerase III holoenzyme runs into the RNA primer attached to the 5^1 end of the previous DNA fragment. The gaps between Okazaki gragments following removal of RNA primer are filled by polymerising action of DNA polymerase I. This enzyme uses its $5^1 \rightarrow 3^1$ exonuclease activity to remove RNA primer. RNA primer can also be erased by RNase H. A specialised enzyme called DNA ligase joins the 3^1 end of the new DNA fragmment of 5^1 end of the previous one and thus produces a continuous DNA chain (lagging strand) from many Okazaki fragments. The continuous synthesis of leading strand using the RNA primer formed the DNA primase is another significant accomplishment of DNA polymerase III holoenzyme (Fig. 3).

The unwinding of the template DNA helix at a replication fork could in principle be catalysed by two types of DNA helicases acting in concert one running along the leading strand template and one along the lagging strand template. They continue to move along their strands thereby prying apart

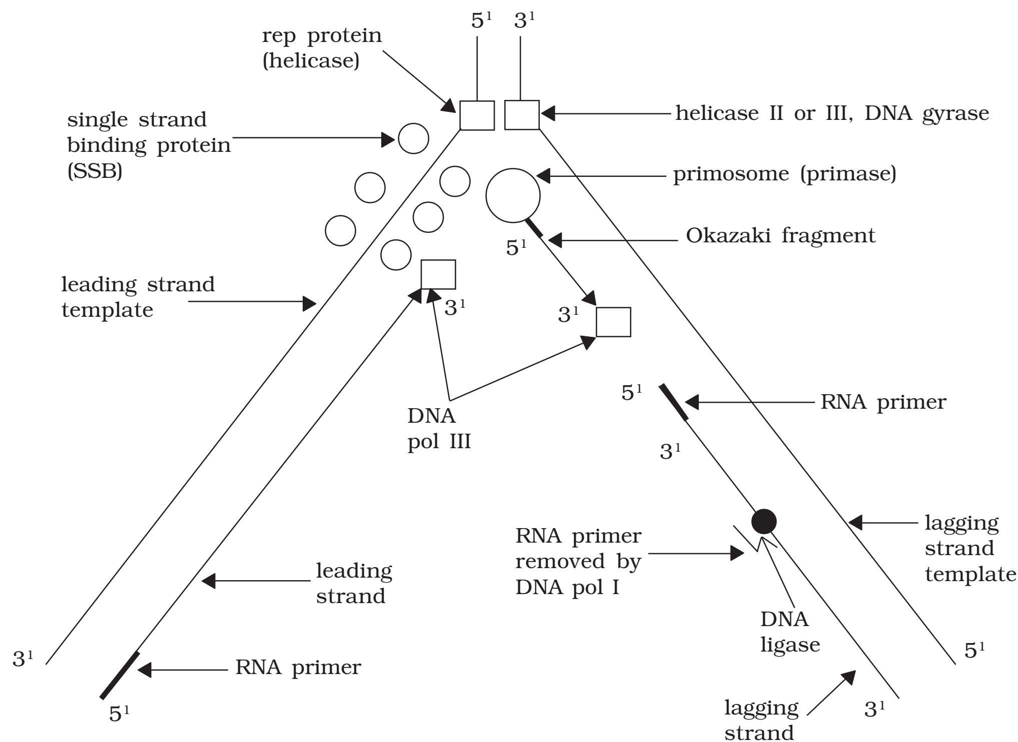


Fig. 3: *Enzymatic events at the replication fork*

the helix at rates upto 1000 nucleotide pairs per second. Single strand DNA binding proteins coat cooperatively the exposed single stranded DNA strands and straighten out the regions of single stranded DNA on the lagging and leading strand template. These proteins prevent formation of short hairpin helices in single stranded DNA thereby enhance the template capability of the single strand template DNA.

The fidelity of copying DNA during replication is such that only about one mistake is made for every 10^9 nucleotides copied. The high fidelity of DNA replication depends not only on coplementary base pairing but also on several proof-reading mechanisms like (i) $5^1 \rightarrow 3^1$ polymerisation of DNA pol I and DNA pol III, (ii) $3^1 \rightarrow 5^1$ exonucleolytic proof-reading of DNA pol I. Polymerases delete errors by acting much like correcting typewriters.

Despite its complexity and accuracy replication occurs very rapidly. In procaryotes, replication rates approach 750 to 1000 base pairs per second per replication fork whereas 50 to 100 base pairs per second per fork in eucaryotic replication. In other words, DNA

synthesis is slower (about 1 um of DNA per minute) in eucaryotes than in procayotes (about 30 um of DNA per minute) due to presence of nucleosomes in eucaryotic chromosomes.

DNA replication stops when the polymerase complex reaches a termination site on the DNA in *E.coli*. The 'Tus' protein binds to these 'Ter' sites and halts replication.