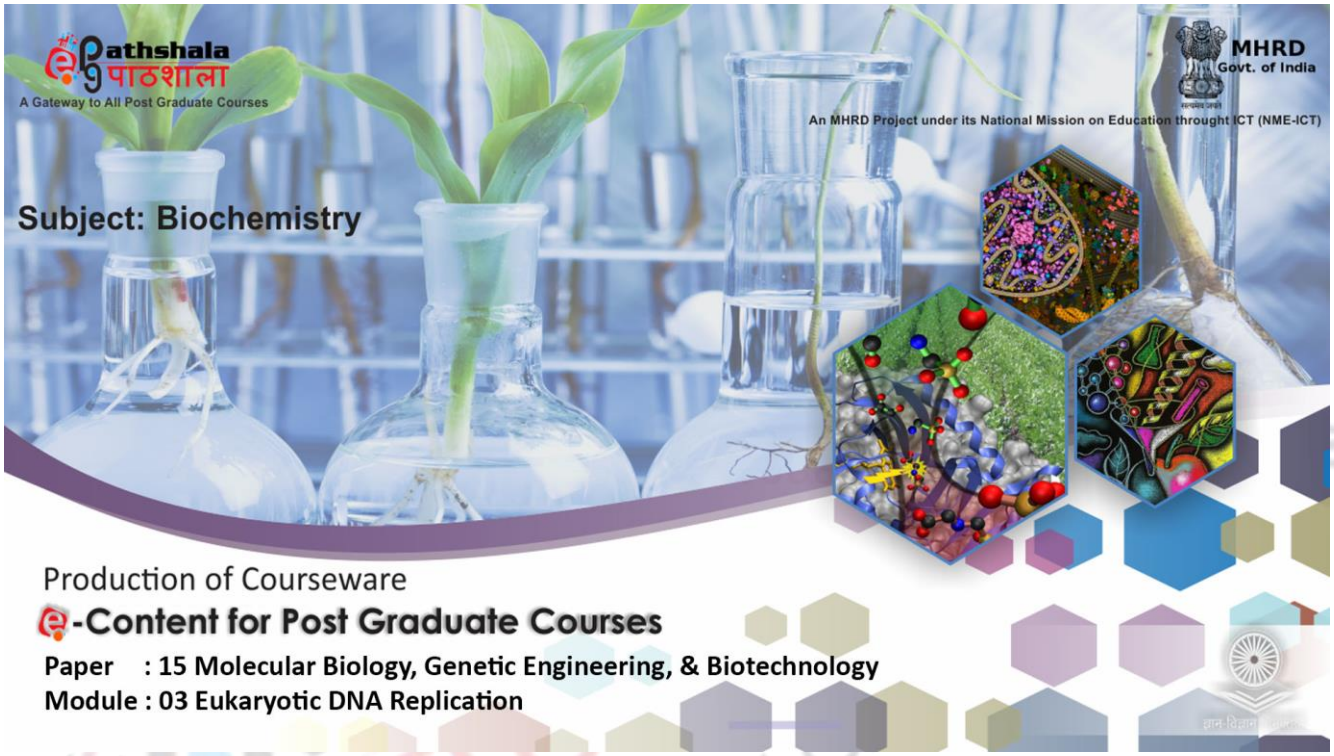


Production of Courseware

e-Content for Post Graduate Courses

Paper : 15 Molecular Biology, Genetic Engineering, & Biotechnology

Module : 03 Eukaryotic DNA Replication

athshala पाठशाला

MHRD Govt. of India

Development Team

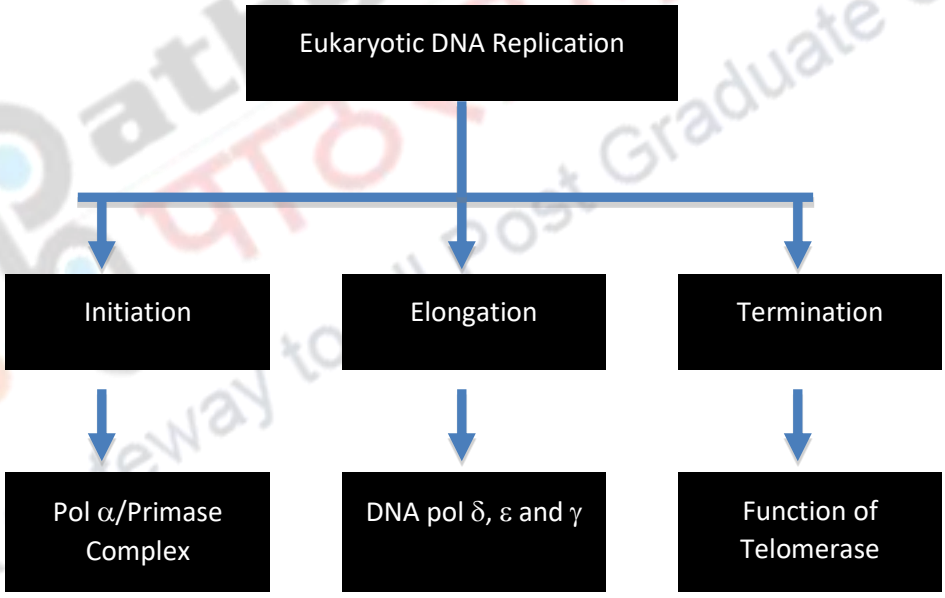
Principal Investigator	Dr. Sunil Kumar Khare, Professor, Department of Chemistry, IIT-Delhi
Content Writer:	Dr. Shashi Sharma (Asst. Prof.) Amity Institute of Biotechnology (AIB), Amity University
Content Reviewer:	Dr. Lata, Principal Scientist, Division of microbiology, IARI, New Delhi
Paper Coordinator:	Dr. Tanmay Dutta, Department of Chemistry, Indian Institute of Technology, Delhi

Description of Module	
Subject Name	Biochemistry
Paper Name	Molecular Biology, Genetic Engineering and Biotechnology
Module Name/Title	Eukaryotic DNA Replication

Eukaryotic DNA Replication

1. **Objective:**
   (a) Understanding of eukaryotic replication process.
   (b) How does eukaryotic replication differ from that of prokaryotes?
   (c) How different DNA polymerases are operative in DNA replication.
   (d) How telomerase functions at the end of eukaryotic chromosome.

2. **Concept map:**



3. Description

DNA replication mechanisms of both prokaryotes and eukaryotes are identical. However, eukaryotic system is more complex in terms of the length of DNA to be replicated and the number of proteins required. In eukaryotes, DNA is compartmentalized in different organelle and hence a vast array of proteins is involved in different mode of replication in different organelle (e.g. nuclear DNA, mitochondrial DNA, chloroplast DNA etc.). In this section, we discuss eukaryotic DNA replication mechanism and the challenge of replicating the ends of linear chromosomes.

3.1 DNA Replication Occurs Bidirectionally from Each Origin

Unlike prokaryote, eukaryotic chromosome contains multiple origin of replication. However, similar to prokaryotic replication, in eukaryotes the replication happens to be bidirectional from each origin. Origins of Replication in eukaryotic chromosome are separated by tens to hundreds of kilo bases. In yeast, the initiation of DNA replication occurs at autonomously replicating sequences (ARS), which are conserved 11-bp sequences adjacent to easily unwound DNA. In mammalian genomes, replication origins do not exhibit sequence conservation, yet these sites all support the binding of a six subunits of origin recognition complex. A hexa-subunit protein called origin recognition complex (ORC), binds to the replication origin. ORC in association with other proteins helps loading of hexameric helicases consisting of six homologous MCM proteins, which separate the parental strands, at an origin in an ATP dependent manner, with RPA proteins (eukaryotic homologue of bacterial SSB protein) binding to the resulting single-stranded DNA. The resulting pre-replication complex (pre-RC) is not competent to initiate replication until it has been activated by other factors that control progress through the cell cycle. Presumably, the separate control of pre-RC assembly and activation allows cells to select replication origins before MCM unwinds the template DNA and replication commences. Once DNA synthesis is initiated, the formation of pre-RC complexes is stopped, thereby ensuring that the DNA replication is occurred only once per cell cycle. Initiation sites are uniformly distributed across the genome in early embryogenesis, when cell division is rapid. However, after cells have differentiated, the distribution of replication origins changes, possibly reflecting patterns of gene expression and/or alterations in DNA packaging in different cell types. Cytological observations indicate that the various chromosomal regions are not all replicated simultaneously. Rather, clusters of 20 to 80 adjacent replicons (replicating units; segments of DNA that are each served by an origin of replication) are activated simultaneously. New sets of replicons are activated until the entire chromosome has been replicated. DNA replication proceeds in each direction from the origin of replication until each replication fork collides with a fork from the adjacent replicon. Eukaryotes appear to lack termination sequences analogous to the Ter sites in *Escherichia coli*.

Primer synthesis and subsequent steps in replication of cellular DNA are discussed in the following sections. Chromosomal DNA replication and other physiological processes leading to the proliferation of the cells are tightly regulated, so that the appropriate numbers of cells constituting each tissue are produced during development and throughout the life of an organism.

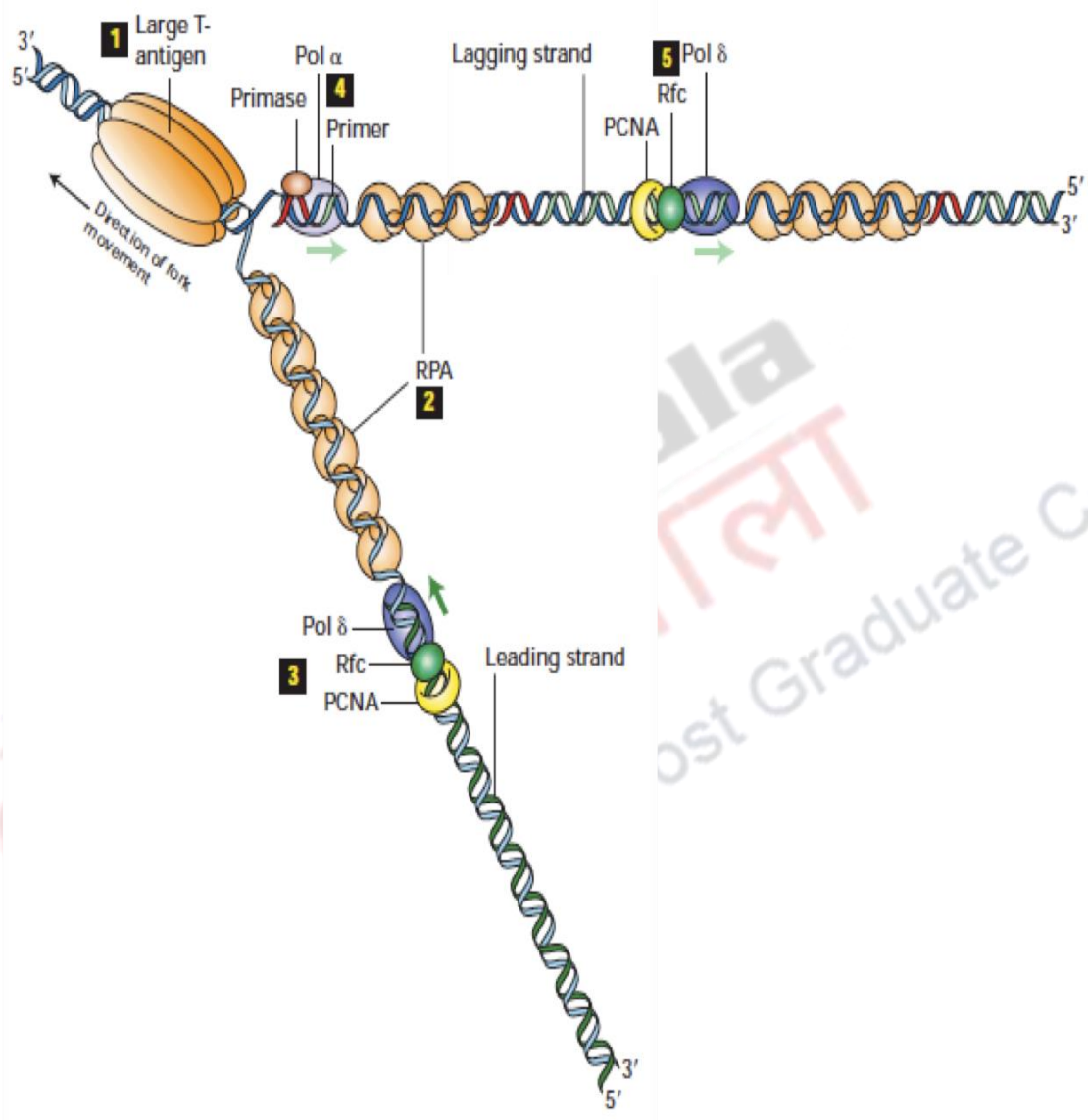


Figure 1: Model of an SV40 DNA replication fork and assembled proteins. Hexameric T-antigen (1), a viral helicase protein unwinds the parental DNA strands. Multiple copies of protein RPA (2) binds to the single stranded region of parental template strands. A protein complex of DNA polymerase ε (Pol ε), PCNA, and Rfc (3) are responsible for the synthesis of leading strand. DNA polymerase α (Pol α) and primase (4) synthesizes lagging strand primers. The 3′-end of each primers are extended by DNA polymerase δ to synthesize each Okazaki fragment (5).

3.2 Several DNA Polymerases are involved in Eukaryotic Replication

Nearly fifteen different DNA polymerases, till date, have been discovered in animal cell. Based on sequence homology, eukaryotic and prokaryotic polymerases are classified into six families: A, B, C, D, X, and Y. Although, several different proteins are actively participate in eukaryotic replication, three among them are very important in replication of nuclear DNA: **polymerases α, δ, and ε**, all of them belongs to B-family of polymerases. Like all DNA polymerases, **DNA polymerase α** replicates DNA by extending a primer in the 5'-3' direction. However, this enzyme lacks exonuclease activity and hence cannot proofread its polymerization product. DNA polymerase α, which is moderately processive and can polymerize 100 nt at a time, remains in a complex with an enzyme called primase. Primase/α complex synthesizes small RNA primer on template strand and extend it to make nearly 20 nt long DNA. Even though the DNA polymerase α does not have any proofreading activity, these few nt are actually replaced along with the elimination of RNA primer. **DNA polymerase δ**, unlike DNA polymerase α, is not associated with a primase and contains proofreading activity (3'-5' exonuclease activity). The processivity of this polymerase is dependent on its association with a sliding clamp protein called proliferating cell nuclear antigen (PCNA), which forms trimeric ring similar to *Escherichia coli* β2 sliding clamp although they rarely have any sequence similarity. In association with PCNA, DNA polymerase δ is required for lagging strand DNA synthesis. A replication factor C is required to load PCNA on to the template strand near primer, which displaces DNA polymerase α from the strand. This process is called Template switching.

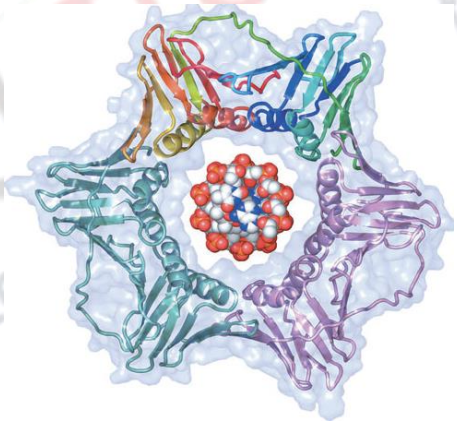


Figure 2: Crystal structure of PCNA. The three subunits, which form a threefold symmetric ring, are drawn in ribbon form embedded in their semi-transparent surface diagram. One of the subunits is colored in rainbow order from its N-terminus (blue) to its C-terminus (red), another is pink, and the third is light green. A space filling model of B-DNA viewed along its helix axis has been drawn in the center of the PCNA ring (PDB ID 1AXC)

A heterotetrameric nuclear enzyme **DNA polymerase ε** is the most enigmatic participant in DNA replication. Even in absence of PCNA, Pol ε is highly processive. It has a 3'- 5' exonuclease activity which instead of replacing mononucleotide, degrades single-stranded DNA to 6- or 7-residue, as does that of pol δ. Although pol ε is necessary for the viability of yeast and its essential function can be carried out by only the noncatalytic C-terminal half of its 256-kD catalytic subunit, which is unique among B-family DNA polymerases. This suggests that the C-terminal half of the pol ε catalytic subunit is required for the assembly of the replication complex. Nevertheless, Thomas Kunkel has shown that pol ε is probably the leading strand replicase, although it may also contribute to lagging strand synthesis. Moreover, pol ε may also participate in leading strand synthesis.

Table 1: Some Important Eukaryotic DNA Polymerases and their properties

DNA Polymerases	Exonuclease (3'-5')	Requires PCNA	Processivity	Primase Association
α	No	No	Moderate	Yes
δ	Yes	Yes	High	No
ε	Yes	No	High	No

Reports show that there are some DNA polymerases like DNA polymerase γ, are organelle specific. DNA polymerase γ carries out polymerization of deoxyribonucleotides in mitochondria. Likewise, chloroplast in plant cell contains a specific DNA polymerase for replication.

3.3 Retrovial DNA Synthesis

Reports show that the host cell genome can be integrated by genomes of a number of animal viruses. Retroviruses, which are enveloped viruses, contain two identical RNA strands in their genome. In retroviruses, RNA strand acts as a template for the formation of a DNA molecule. This mechanism, on the contrary of most common transcription of DNA into RNA, involves the reverse flow of genetic information. Retroviral genome contains a viral enzyme called reverse transcriptase initially copies the viral RNA genome into complementary single stranded DNA to the viral RNA, resulting in the generation of a RNA-DNA hybrid. The viral RNA strand is then nucleolytically degraded by an RNase H (an RNase activity that hydrolyzes the RNA of an RNA–DNA hybrid helix). Same enzyme then catalyzes the synthesis of a complementary DNA strand using the existing DNA strand. The resulting double-stranded DNA is integrated into the chromosome of the infected cell, thus making a way to escape from host defense machinery. The integrated proviral DNA, is finally transcribed in to RNA by the host's transcription machinery. Viral genome specific RNA is either translated into viral proteins or

packaged within viral coat proteins resulting in the formation of infectious progeny virus that are released from the host-cell by budding out from the membrane. Majority of the retroviruses do not kill their host cells. Once infected, cells can replicate, and divide producing daughter cells containing integrated proviral genome. These daughter cells continue transcribing the proviral genome and bud progeny virions. RT has been a particularly useful tool in genetic engineering because it can transcribe mRNAs to complementary strands of DNA (cDNA). Messenger RNA-derived cDNAs can be used, for example, to express eukaryotic structural genes in *E. coli*. Since *E. coli* lacks the machinery to splice out introns, the use of genomic DNA to express a eukaryotic structural gene in *E. coli* would require the prior excision of its introns—a technically difficult feat.

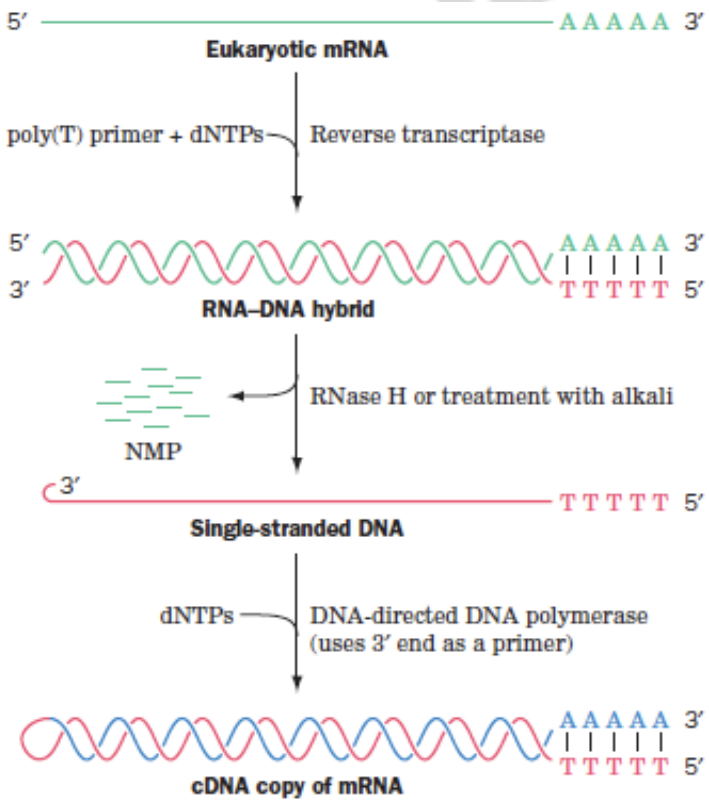


Figure 3: Process of Retroviral Complementary DNA synthesis: Retroviral RNA is used as a template for complementary DNA (cDNA) synthesis. RNase H removes viral RNA in RNA-DNA hybrid and then DNA directed DNA polymerase synthesizes another stand of DNA forming double helix.

3.4 Removal of RNA primers in Eukaryotes

In mammalian systems where the DNA polymerase does not have 5'-3' exonuclease activity (like DNA polymerase I in *E. coli*), Okazaki fragments are connected by a two-step process. RNA primer of the preceding fragment is displaced by the synthesis of Okazaki fragments in the form of "flap". In the first step RNase H1 cleaves major portion of the RNA, leaving only several 5'-ribonucleotides adjacent to the DNA. An enzyme called "flap endonuclease I" (FEN) removes rest of the bases of flap by endonuclease action. It has been seen in eukaryotic system that pol α extends the RNA primer by ~15 nt of DNA before pol δ displace it from template DNA. The lack of proofreading ability in pol α results in an erroneous primer extension which is not observed in the polymerization reaction of pol δ. However, PCNA recruits FEN1 which provides what is, in effect, pol α's proofreading function: It is also an endonuclease that excises mismatch-containing oligonucleotides up to 15 nt long from the 5' end of an annealed DNA strand. Moreover, FEN1 can make several such excisions in succession to remove more distant mismatches. The excised segment is later replaced by pol δ as it synthesizes the succeeding Okazaki fragment

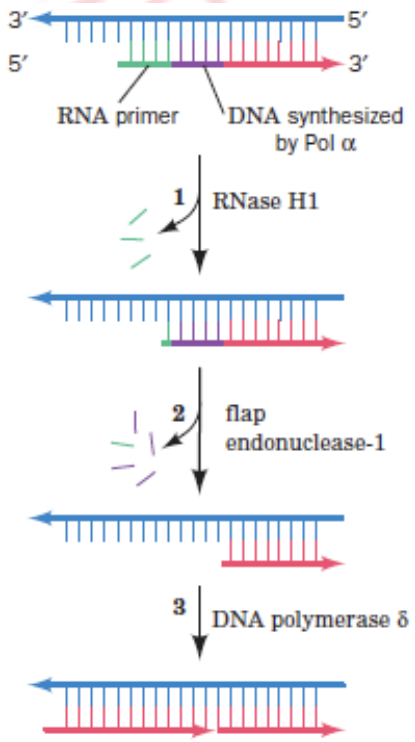


Figure 4: Elimination of RNA primers in eukaryotes: (1) RNase H1 removes all but the 5'-ribonucleotide of the RNA primer. (2) FEN1, a 5'-3' endonuclease, then removes the remaining ribonucleotide along with a segment of adjoining DNA if it contains mismatches. (3) The excised nucleotides are replaced as DNA polymerase δ completes the synthesis of the next Okazaki fragment. The nick is eventually sealed by DNA ligase.

3.5 Telomerase solves the end-replication problem

The replication machinery always have a problem replicating the end of a linear chromosome, specifically, DNA polymerase is unable to synthesize the extreme 5' end of the lagging strand. Even if an RNA primer were paired with the 3' end of the DNA template, it could not be replaced with DNA (recall that DNA polymerase operates only in the 5'-3' direction; it can only extend an existing primer, and the primer must be bound to its complementary strand). Consequently, in the absence of a mechanism for completing the lagging strand, linear chromosomes would be shortened at both ends by at least the length of an RNA primer with each round of replication. Telomeres Are Built from an RNA Template. The ends of eukaryotic chromosomes, the telomeres, have an unusual structure. Telomeric DNA consists of 1000 or more tandem repeats of a short G-rich sequence (TTGGGG in the protozoan Tetrahymena and TTAGGG in humans) on the 3'-ending strand of each chromosome end. Elizabeth Blackburn, Carol Greider, and Jack Szostak have shown that telomeric DNA is synthesized and maintained by an enzyme named telomerase, which is a ribonucleoprotein (a complex of protein and RNA). The RNA component (451 nt in humans) includes a segment that is complementary to the repeating telomeric sequence and acts as a template for a reaction in which nucleotides are added to the 3' end of the DNA. Telomerase translocates a repetitive G-rich sequence to the new 3' end of the DNA strand, thereby adding multiple telomeric sequences to the DNA. The sequences complementary to the telomeric-G-rich strand is apparently synthesized by the normal cellular machinery for lagging strand synthesis, leaving a 100- to 300-nt single-stranded overhang on the G-rich strand. Telomerase functions similarly to reverse transcriptase; in fact, its highly conserved catalytic subunit, called TERT, is homologous to reverse transcriptase. In association with finger, palm and thumb domain, TERT includes an additional N-terminal RNA-binding domain, which is not present in normal polymerases. The absence of telomerase, which allows the gradual truncation of chromosomes with each round of DNA replication, contributes to the normal senescence of cells. Conversely, enhanced telomerase activity permits the uncontrolled replication and cell growth that occur in cancer.

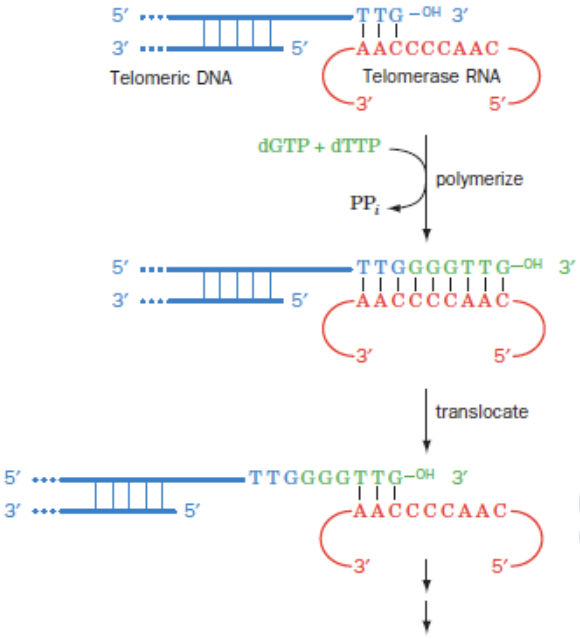


Figure 5: Mechanism for the synthesis of telomeric DNA by Tetrahymena telomerase. The red strand contains the G-rich repetitive sequence. The telomere's 5'-ending strand is later extended by normal lagging strand synthesis.

Cells from knockout mice that cannot produce the telomerase-associated RNA exhibit no telomerase activity, and their telomeres shorten successively with each cell generation. Such mice can breed and reproduce normally for three generations before the long telomere repeats become substantially eroded. Then, the absence of telomere DNA results in adverse effects, including fusion of chromosome termini and chromosomal loss. By the fourth generation, the reproductive potential of these knockout mice declines, and they cannot produce offspring after the sixth generation.

The human genes expressing the telomerase protein and the telomerase-associated RNA are active in germ cells but are turned off in most adult tissues, in which cells replicate only rarely. However, these genes are reactivated in most human cancer cells, where telomerase is required for the multiple cell divisions necessary to form a tumor. This phenomenon has stimulated a search for inhibitors of human telomerase as potential therapeutic

agents for treating cancer. While telomerase prevents telomere shortening in most eukaryotes, some organisms use alternative strategies. Drosophila species maintain telomere lengths by the regulated insertion of non-LTR retrotransposons into telomeres. This is one of the few instances in which a mobile element has a specific function in its host organism.

3.6 G-Quartets are formed in Telomere

The G-quartet is a cyclic hydrogen bonded array of four guanine nucleotides that can be formed by G-rich sequences in both DNA and RNA. Several G-quartets can stack upon each other to form quadruple helical structures. G-rich polynucleotides are essentially difficult to work with due to their propensity to aggregate via Hoogsteen-type base pairing to form cyclic tetramers known as G-quartets. Indeed, the G-rich overhanging strands of telomeres fold back on themselves to form a hairpin, two of which associate in an antiparallel fashion to form stable complexes of stacked G-quartets. Such structures presumably serve as binding sites for capping proteins, which may help regulate telomere length and prevent activation of DNA repair mechanisms that recognize the ends of broken DNA molecules.

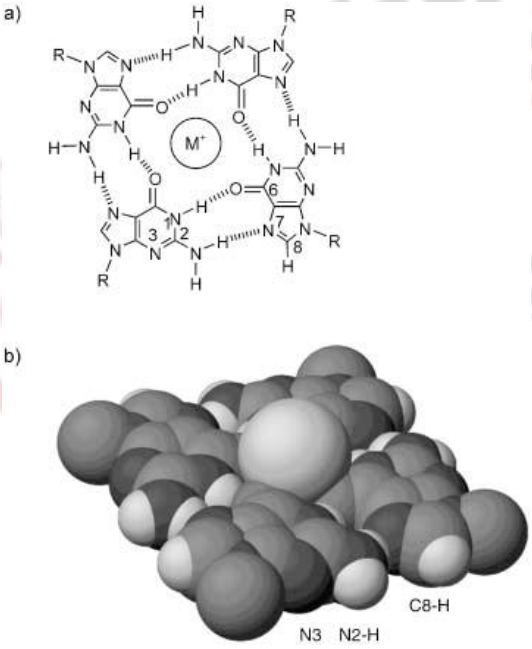


Figure: 6: a) The G-quartet; b) space-filling model from the crystal structure of a typical G-quartet with a K+ ion bound above the planar assembly. The riboses have been removed for clarity.

3.7 Nucleosomes Reassembly happens behind the Replication Forks

In contrast to prokaryotic DNA, eukaryotic DNA is packaged in nucleosomes. Some alteration of this structure is probably necessary for initiation, but once replication is under way, nucleosomes do not seem to impede the progress of DNA polymerases. Experiments with labeled histones indicate that nucleosomes just ahead of the replication fork disassemble and the freed histones, either individually or as dimers or tetramers, immediately reassociate with the emerging daughter duplexes. The parental histones randomly associate with the leading and lagging duplexes. DNA replication (which occurs in the nucleus) is coordinated with histone protein synthesis in the cytosol so that new histones are available in the required amounts.

4. Summary

This lecture includes-

(a) Eukaryotic DNA Replication
(b) Properties of different DNA polymerases and their functions
(c) Retrovial DNA synthesis
(d) Telomerase and its function to maintain the chromosomal integrity
(e) G-quartet structure of telomere