REPLICATION LICENSING FACTOR

Before every round of cell division, the DNA of each cell should be replicated once. There are regulatory mechanisms of the cell to ensure that each replicon replicates only once in one cell cycle of cell division, mediated by a factor known as Replication Licensing Factor, which permit the formation of initiation complex.

A licensing factor is a protein or complex of proteins that allows an origin of replication to begin DNA replication at that site. Licensing factors primarily occur in eukaryotic cells, since bacteria use simpler systems to initiate replication.

Unlike prokaryotes, eukaryotic DNA replicates from different origin (multireplicon) on a single chromosome. Research has shown that the many origins are not all activated at once; instead, clusters of 20-80 adjacent replicons are activated sequentially throughout the S phase until all DNA is replicated.

During the G_1 phase of the cell cycle, all Autonomous Replicating Sequences (ARSs) are initially bound by a group of specific proteins (six proteins in Yeast), forming what is called the Origin Recognition Complex (ORC). Since these recognition complexes are found in G_1, but synthesis is not initiated at their sites until S, there must be still other proteins involved in the actual initiation signal. The most important of these proteins are specific kinases, key enzyme involved in phosphorylation, that are integral part of cell cycle control. When bound along with ORC, a prereplication complex is formed that is accessible to DNA polymerase. After these kinases are activated, they serve to complete the initiation complex, directing localized unwinding by the helicase enzyme and triggering DNA synthesis. Activation also inhibits reformation of the prereplication complexes once DNA synthesis has been completed at each replicon. It ensures that replication only occurs once along each stretch of DNA during each cell cycle.

Now we have a so-called pre-replication complex, which then allows a heterohexameric protein complex of proteins MCM2 to MCM7 to bind. This entire hexamer exerts its activity untill helicase unwind the double stranded DNA.

Helicase loading and its activation are regulated to allow only a single round of replication during each cell cycle. This regulation is done by the level of the protein Cyclin Dependent Kinase (CDK). During G_1 the CDK level is low allowing helicase loading but preventing helicase activation. Entry into S phase is coupled with increase in CDK level and activation

of loaded helicases by CDK, but simultaneously preventing new helicase loading. The high level of CDK during S, G_2 and M phases inhibits the function of ORC and cdc6 and cdt1 (the two helicase loading proteins). At this point Cdc6 leaves the complex and is inactivated, by being degraded in yeast but by being exported from the nucleus in metazoans, triggered by CDK-dependent phosphorylation. The next steps included the loading of a variety of other proteins like MCM10, a CDK, DDK and Cdc45, the latter directly required for loading the DNA polymerase. During this period Cdt1 is released from the complex and the cell leaves G1 phase and enters S phase when replication starts. It is only when cells segregate their chromosome and complete cell division that CDK activity is eliminated, allowing a new round of helicase loading to commence.

Such a CDK-regulated licensing control is conserved from yeast to higher eukaryotes, and ensures that DNA replication takes place only once in a cycle. *Xenopus laevis* and mammalian cells have an additional system to control licensing.

Geminin, whose degradation at the end of mitosis is essential for a new round of licensing, has been shown to bind Cdt1 and negatively regulate it, providing a new insight into the regulation of DNA replication in higher eukaryotes.

The licensing factors present in the cytoplasm of the cell, this factor must get into the nucleus where DNA is scheduled to replicate; however, the factor cannot penetrate the nuclear membrane. So the factor can get entry into the nucleus only when the nuclear membrane is dissolved during mitosis. After mitosis, a new nuclear membrane will be formed, but the factor will be within the nucleus now, since it got in during the time when there was no nuclear wall. Replication can now take place with the help of this factor. After replication the factor is destroyed thus only one round of replication can takes place by the captured factor. If second round of replication is to happen, the factor has to get entry again afresh, which is possible only after another mitosis.

There was a time when it was thought that ORC itself is the key factor for initiation of replication. But ORC is always bound to DNA except in the elongation phase. This quality of ORC argues against its being the initiating/licensing factor.

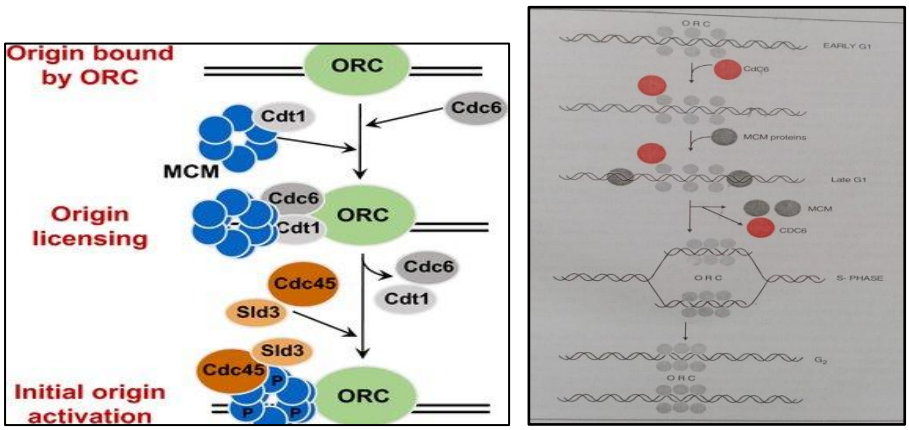
In the yeast, the prime candidates for licensing factor(s) are the MCM proteins – MCM2, 3, 5. These proteins are required for replication and enter the nucleus only during mitosis. In animal cells though, the MCM proteins may just be one of the components of the licensing

factor. This is so because the MCM proteins are inside the nucleus throughout the cell cycle. It may therefore be that some other protein may be required to assist MCM protein in initiating replication. Cdc6 and Cdt1 fulfil the role of licensing factors. They are only produced in G1 phase, in addition to which binding of all the proteins in this process excludes binding of additional copies. In this way their mode of action is limited to starting replication once, since once they have been ejected from the complex by other proteins, the cell enters S phase, during which they are not re-produced or re-activated. Thus they act as licensing factors, but only together they can act.

Recent investigations tell us about involvement of the cell cycle protein Cdc6 in replication licensing. It is highly unstable protein with a half-life less than 5 minutes, thus degrade rapidly which will not allow it to support replication initiation more than once. Its synthesis in G_1 phase and binding with ORC between mitosis and G_1 phase also fits that this is the licensing factor.

The ORC is already bound to the origin. It probably guides Cdc6 to the place of binding at the origin. Once Cdc6 binds, it helps MCM proteins to bind to the complex. The prereplication complex is therefore is made up of ORC, Cdc6 and MCM proteins. This is the situation in the beginning of the S phase.

Immediately upon initiation, Cdc6 and MCM are displaced from the origin leaving just ORC bound. Cdc6 is rapidly degraded, probably through ubiquitination. Cdc6 is therefore not available now throughout the S phase. Without Cdc6, MCM proteins cannot bind again and thus a second initiation becomes impossible unless the cell undergoes mitosis paving way for entry of fresh Cdc6.



Probable model of Replication licensing factor

ASSEMBLY OF NEW NUCLEOSOME

For the cell, the replication has to be the replication of chromatin – faithful replication of the double stranded DNA along with the faithful reconstruction of chromatin; deliberate and correct addition of histones at the growing fork. We know very little about how the cell achieves this aspect. But, the latter process is tightly linked to DNA replication to ensure that the newly replicated DNA is rapidly packaged into nucleosomes.

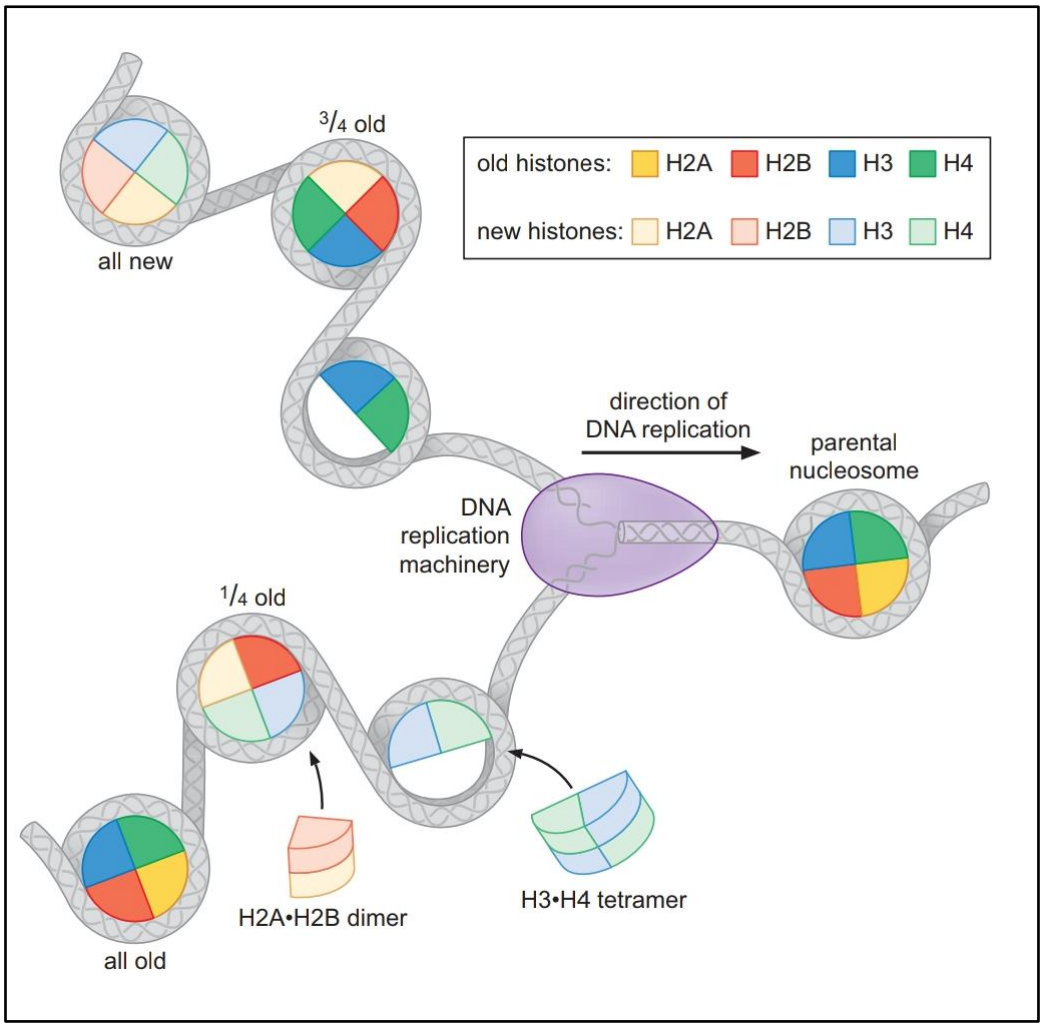
Although the replication of DNA requires the nucleosome disassembly, the DNA is rapidly repackaged into nucleosomes in an ordered series of events. The first step in the assembly of a nucleosome is the binding of an H3.H4 tetramer to the DNA. Once the tetramer is bound, two H2A.H2B dimers associate to form the final nucleosome. H1 joins this complex last, presumably during the formation of higher order chromatin assemblies.

To duplicate a chromosome, at least half of the nucleosomes on the daughter chromosomes must be newly synthesized. If the old histones were lost completely, then chromosome duplication would erase any "memory" of the previously modified nucleosomes. In contrast, if the old histones were retained on a single chromosome, that chromosome would have a distinct set of modifications relative to the other copy of the chromosome.

In experiment, it is found that the old histones are present on both of the daughter chromosomes. Mixing is not entirely random; however, H3.H4 tetramers in one nucleosome are composed of either all new or all old histones. Thus, as the replication fork passes, nucleosomes are broken down into their component subassemblies. H3.H4 tetramers appear to remain bound to one of the two daughter duplexes at random and are never released from DNA into the free pool of histones. In contrast, the H2A.H2B dimers are released and enter the local pool, available for new nucleosome assembly.

Old modified histones will tend to rebind one of the daughter chromosomes at a position near their previous position on the parental chromosome. The old histones have an equal probability of binding either daughter chromosome. This localized inheritance of modified histones ensures that a subset of the modified histones is localized in similar positions on each daughter chromosome.

The ability of these modified histones to recruit enzymes that add similar modifications to adjacent nucleosomes provides a simple mechanism, which plays a critical role in the inheritance of chromatin states from one generation to another.

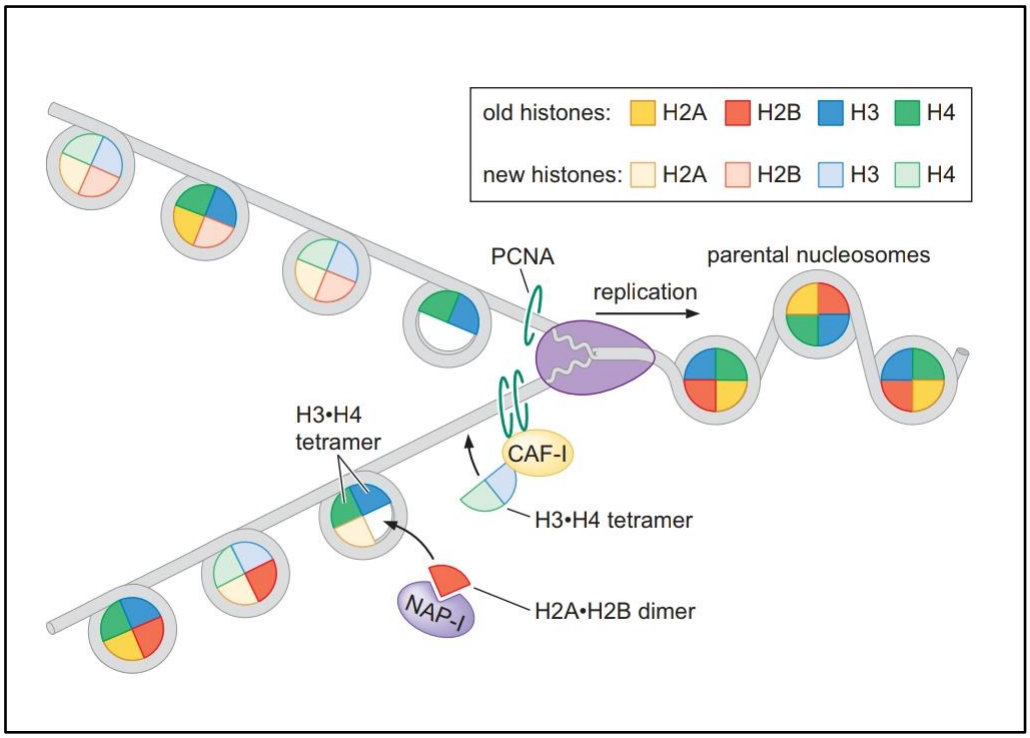


The assembly of nucleosomes is not a spontaneous process. Factors have been identified which are required to direct the assembly of histones onto the DNA. Accessory proteins may be needed in the assembly of nucleosomes. Extracts from *Xenopus* eggs contain two such factors. The protein nucleoplasmin has the ability of binding H2A and H2B while the protein N1 binds H3 and H4. Both N1 and nucleoplasmin are acidic proteins. Their abilities to bind different groups of histones point out to their probable function. Histones bind to DNA by virtue of their positive charges. If the matter is just positive charges meeting the negative charges on the DNA, aberrant assemblies of these histones can form. The job of these two accessory proteins seems to be to hold the different components of octamers and escort them to sites of nucleosome assembly and to release them in a controlled manner leading to the correct assembly of the nucleosomes. Because they act to keep histones from interacting with the DNA nonproductively, these factors have been referred to as histone chaperones.

Acetylation of histones H3 and H4 seems to be an event necessary for chromatin replication. Shortly after their synthesis, histones H3 and H4 associate with each other and are acetylated

at a number of lysine residues within their amino-terminal domain. In higher eukaryotes, this acetylation is transient and residues are rapidly deacetylated after the histones have been packaged into chromatin. The acetylation of histone prior to their packaging into chromatin is carried out by enzyme known as B-type histone acetyltransferases (B-type HATs). Such enzymes are different from other (A-type HATs) acetyltransferases which acetylate chromosomal histones.

The only known and characterised B-type HAT is Hat-1. This enzyme is widely conserved and at least in *Xenopus* and humans can acetylate lysine residues 5 and 12 of histone H4. Though, widely conserved, mutations in this do not put the cells at a disadvantage as far as nucleosome assembly is concerned. This may point out to the fact that there may be other enzymes in the cells which can carry out the same acetylation. We also know that acetylation of lysine 8 of H4 is sufficient for viability of nucleosome assembly. This acetylation must be carried out by a different B-type HAT. We know nothing about the B-type HATs which acetylate H3, except that they must exist. However, unlike H4 where acetylation of lysine 5 and 12 was a conserved feature across many species, the H3 acetylation points differ from species to species.

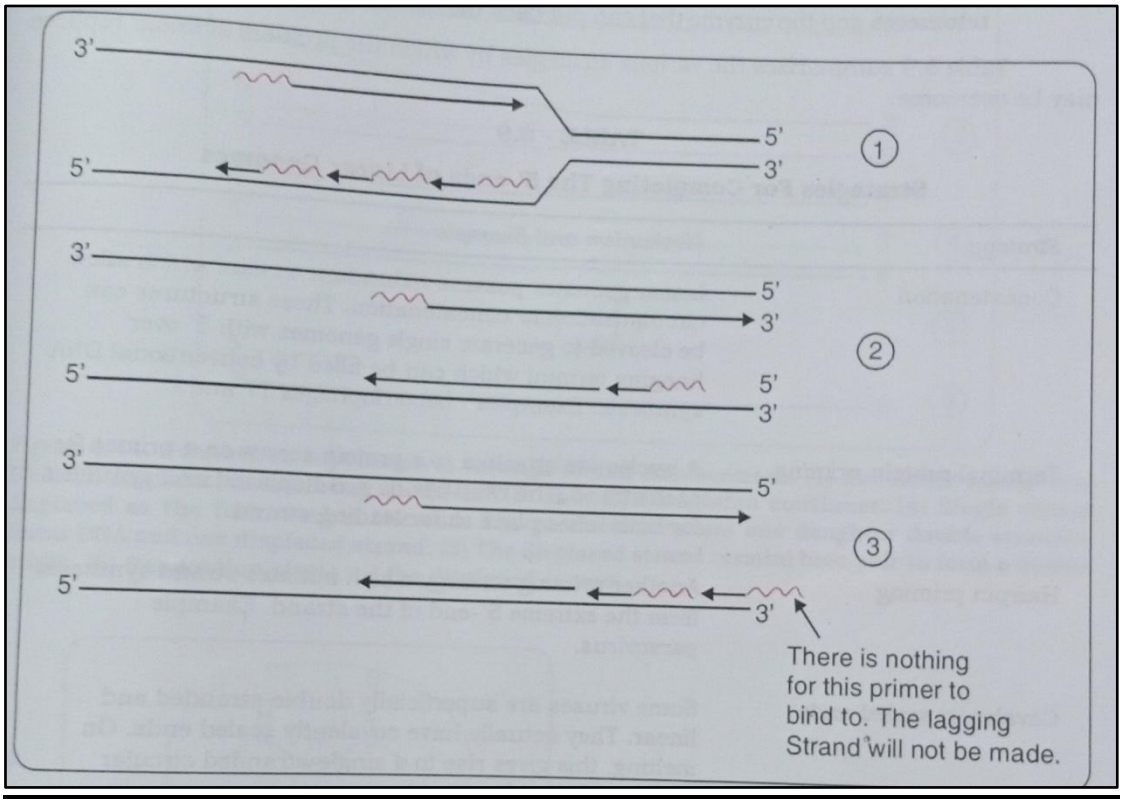


Acetylated histone residues are taken preferentially to the replicating DNA by the Chromatin Assembly Factor 1 (CAF-1). CAF-1 distinguishes between replicating DNA and preformed DNA. CAF-1 has an affinity for PCNA (Proliferating Cell Nuclear Antigen), the DNA

polymerase processivity factor. It is owing to this affinity that CAF-1 can direct its cargo of histones to the replicating DNA. This factor forms a ring shaped sliding clamp around the DNA duplex and is responsible for holding DNA polymerase on the DNA during DNA synthesis. After the polymerase is finished, PCNA is released from the DNA polymerase but still encircles the DNA. In this condition, PCNA is available to interact with other proteins. CAF-1 associates with the released PCNA and assembles H3.H4 preferentially on the PCNA-bound DNA. Thus, by associating with a component of the DNA replication machinery, CAF-1 is directed to assemble nucleosomes at sites of recent DNA replication.

REPLICATION AT THE END CHROMOSOME TELOMERE,

TELOMERASE CONCEPT



The requirement for an RNA primer to initiate all new DNA synthesis creates a dilemma for the replication of the ends of linear chromosomes, called the end replication problem. There is no difficulty in the replication copying the leading strand template. Once initiated from a single internal RNA primer, leading strand can be extended coping to the extreme 5' end of the template. At this end the polymerase will probably fall off. In contrast, the requirement for multiple primers to complete lagging strand synthesis and the direction of the primer towards the direction of replication fork movement creates a problem. If the end of the last

RNA primer for Okazaki fragment synthesis anneals even to the final base of the lagging strand template, once this RNA molecule is removed, there will remain a short region (the size of RNA primer) of unreplicated ssDNA at the end of the chromosome. This means that each round of replication would result in the shortening of one of the two daughter DNA molecules. Obviously this cannot be allowed since with every replication round, the cell will lose chunks of genes located terminally.

The above problem has molecular solution in at least three different ways –

1. Linear chromosome become circular during replication and post-replication specific cuts converts the circularized DNA back to linear. T_4 phage, λ phage adopt this solution.
2. Use a protein instead of RNA as the primer. The 'priming protein' binds to the lagging strand template and uses an amino acid to provide the free 3'OH. Adenovirus solves the problem in this way.
3. Most eukaryotic cells use an entirely different solution to replicate their chromosome ends. The ends of eukaryotic chromosomes are called telomeres. Telomeres are generally composed of 1000 or more tandem repeats of a G-rich DNA sequence (for example in Tetrahymena it is TTGGGG, in human being it is TTAGGG) and 12-16 base overhang at the 3' end of the chromosome which extends beyond the 5' end of complementary strand. These telomeric sequences are replicated by the action of unique enzyme called telomerase. Telomerase solves the end replication problem by extending the 3' end of the chromosome. Using the repeat sequence of its own RNA molecule as template, telomerase can extend the 3' end of one strand by one repeat unit beyond its original length. This action can be repeated several times to lengthen the telomere. In this manner, the linear replicons of metazoans can save themselves from the consequences of shortening chromosome every time they indulge replication.

The telomerase was first discovered in the ciliated protozoan *Tetrahymena*. Telomerase is a remarkable reverse transcriptase enzyme that is composed of multiple protein subunits and a RNA component, i.e., it is a ribonucleoprotein. Like all other DNA polymerase, telomerase acts to extend the 3' end of its DNA substrate. But unlike most DNA polymerase, telomerase does not need an exogenous DNA template to direct the addition of new dNTPs. Instead, the

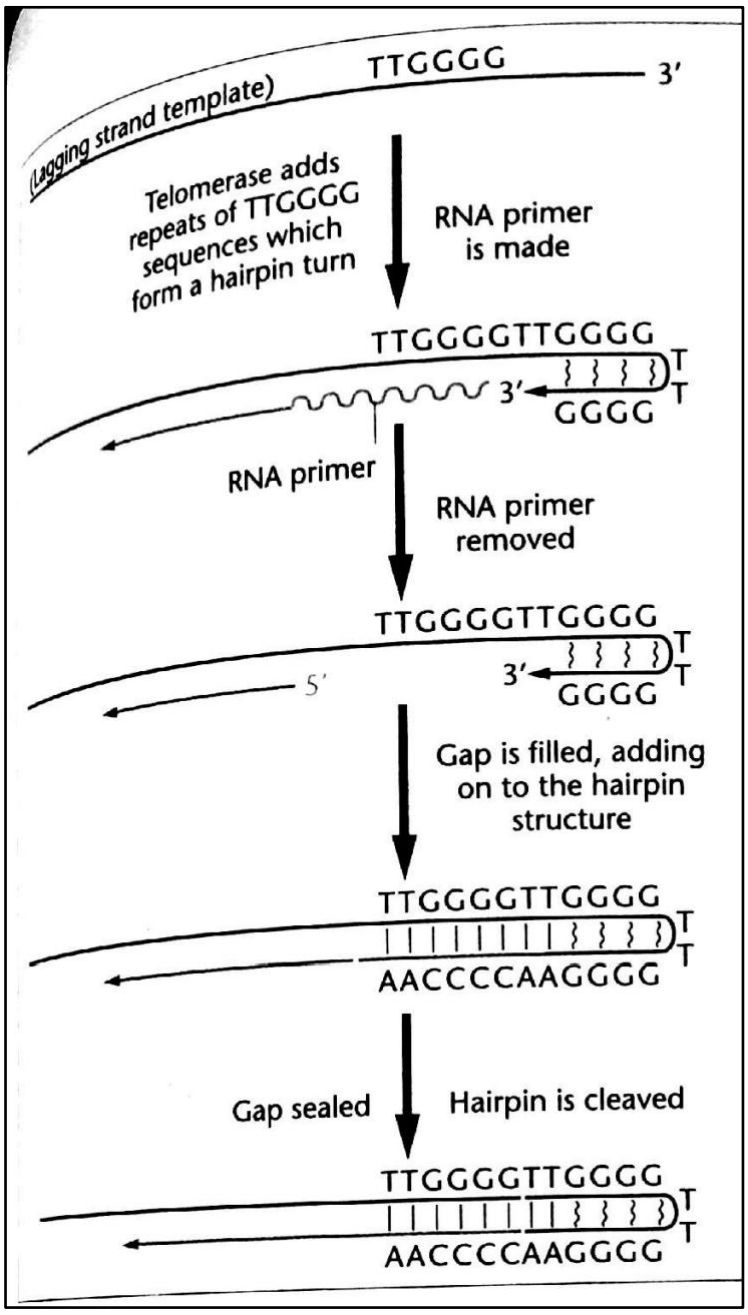
own RNA component of telomerase itself serves as the template for adding the telomeric sequence to the 3' terminus at the end of the chromosome.

The characteristics of telomerase are in some ways distinct and in other ways similar to those of other DNA polymerase. The inclusion of an RNA component, the lack of a requirement for an exogenous template and the ability to use an entirely ssDNA substrate to produce an ssDNA product sets telomerase apart from DNA polymerases. In addition, telomerase must have the ability to displace its RNA template from the DNA product to allow repeated rounds of template-directed synthesis. Formally, this means that telomerase includes an RNA.DNA helicase activity. On the other hand, like all other DNA polymerases, telomerase requires a template to direct nucleotide addition, can only extend a 3'-OH end of DNA, uses the same nucleotide precursors, and acts in a processive manner, adding many sequence repeats each time it binds to a DNA substrate.

Depending on the organism, the telomerase RNA (TER) varies in size from 150 to 1300 bases. In all organisms, the sequence of the RNA includes a short region that encodes about 1.5 copies of the complement of the telomere sequence (for humans, this sequence is 5'-AAUCCCAAUC-3'). This region of the RNA can anneal to the ssDNA at the 3' end of the telomere. Annealing occurs in such a way that a part of the RNA template remains single-stranded, creating a primer:template junction that can be acted on by telomerase. Using the associated RNA template, telomerase reverse transcriptase (TERT) synthesizes DNA to the end of the TER template region but cannot continue to copy the RNA beyond that point. At this point, the RNA template disengages from the DNA product, reanneals to the last four nucleotides of the telomere, and then repeats this process.

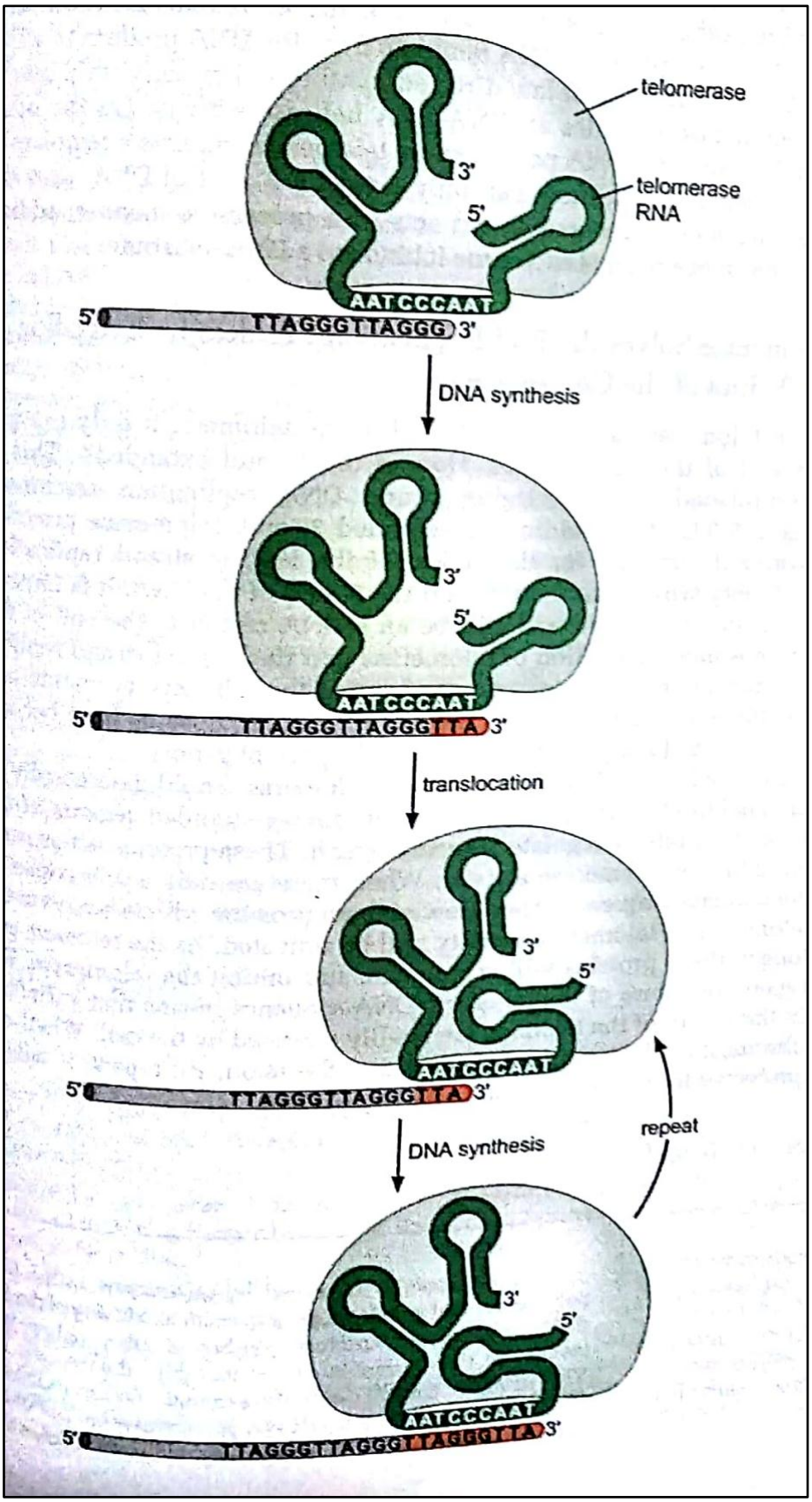
The enzyme is capable of adding several copies of the repeat sequence it has within its RNA molecule to the 3'end of the lagging strand (using 5'-3' synthesis). These repeats are then thought to form a ‘hairpin loop’, which may be stabilized by unorthodox hydrogen bonding between opposite guanine residues, creating a free 3'-OH group on the end of the hairpin that can serve as a substrate for DNA polymerase I. This makes it possible to fill the gap that would otherwise shorten the chromosome. The hairpin loop can then be cleaved off at its terminus and the potential loss of DNA in each subsequent replication cycle is averted. The RNA component serves as both a guide and a template for the synthesis of its DNA complement, a process referred to as reverse transcription. The presence of a 3' overhang may be important for the end protection function of the telomere. Nevertheless, the action of

telomerase and the lagging-strand replication machinery ensures that the telomere is maintained at sufficient length to protect the end of the chromosome from shortening. Because of the repetitive and non-protein-coding nature of the telomeric DNA, variations in the length of the telomere are easily tolerated by the cell.



We can envision that part of the RNA sequence of the enzyme base pairs with the telomeric DNA, with the remainder of the RNA overlapping the end of the lagging strand. Reverse transcription of this overlapping sequence which synthesizes DNA on an RNA template, will lead to the extension of the lagging strand. It is believed that the enzyme is then translocated

toward the end of the lagging strand and the sequence of events is repeated. Specific proteins are responsible for initiating synthesis.



Telomere binding proteins regulate telomerase activity and telomere length. Although extension of telomeres by telomerase could theoretically go on indefinitely, proteins bound to the double-strand regions of the telomere regulate telomere length. In yeast cells, proteins bound to the telomere act as weak inhibitors of telomerase activity. When there are relatively few copies of the telomere sequence repeat, few of these proteins are bound to the telomere, and telomerase can extend the 3'-OH end of the telomere. As the telomere becomes longer, more of the telomere binding proteins accumulate and inhibit telomerase extension of the 3'-OH end of the telomere. This simple negative-feedback loop mechanism (longer telomeres inhibit telomerase) is a robust method to maintain a similar telomere length at the ends of all chromosomes.

Proteins that recognize the single-stranded form of the telomere can also modulate telomerase activity. In yeast, the Cdc13 protein binds to single-stranded regions of the telomere. Studies of this protein indicate that it recruits telomerase to the telomeres. Thus, Cdc13 is a positive activator of telomerase. In contrast, the human protein that binds to single-stranded telomeric DNA, POT1, acts in the opposite manner – that is, as an inhibitor of telomerase activity. In vitro studies show that POT1 binding to the single stranded telomere DNA inhibits telomere activity. Cells that lack this protein show dramatically increased telomere DNA length. Interestingly, this protein interacts indirectly with the double-strand telomere-binding proteins in human cells. It has been proposed that as telomeres increase in length, more POT1 is recruited; thereby increasing the likelihood that it binds to the ssDNA ends of the telomere and inhibits telomerase.

In addition to their role in regulating telomerase function, telomere-binding proteins also play a crucial role in protecting the ends of chromosome. It is possible that protection is conferred simply by coating the telomere with binding proteins.

Telomere shortening has been linked to a molecular mechanism involved in the aging process of cells. In most eukaryotic somatic cells, telomerase is, in fact, not active, and thus, with each cell division, the telomeres of each chromosome shorten. After many divisions, the telomere is seriously eroded and the cell loses the capacity for further division. Malignant cells, on the other hand, maintain telomerase activity and are immortalized.