**NUCLEIC ACID SYNTHESIS**

Nucleic acids are required for the storage and expression of genetic information. There are two chemically distinct types of nucleic acids: deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). DNA, the repository of genetic information, is present not only in chromosomes in the nucleus of eukaryotic organisms, but also in mitochondria and the chloroplasts of plants. Prokaryotic cells, which lack nuclei, have a single chromosome, but may also contain non-chromosomal DNA in the form of plasmids. The genetic information found in DNA is copied and transmitted to daughter cells through DNA replication (DNA synthesis). The DNA contained in a fertilized egg encodes the information that directs the development of an organism. This development may involve the production of billions of cells. Each cell is specialized, expressing only those functions that are required for it to perform its role in maintaining the organism. Therefore, DNA must be able to not only replicate precisely each time a cell divides, but also to have the information that it contains be selectively expressed. Transcription (RNA synthesis) is the first stage in the expression of genetic information. Next, the code contained in the nucleotide sequence of messenger RNA molecules is translated, thus completing gene expression.

DNA is a polymer of deoxyribonucleoside monophosphates covalently linked by 3'-5'–phosphodiester bonds. With the exception of a few viruses that contain single-stranded (ss) DNA, DNA exists as a double stranded (ds) molecule, in which the two strands wind around each other, forming a double helix. In eukaryotic cells, DNA is found associated with various types of proteins (known collectively as nucleoprotein) present in the nucleus, whereas in prokaryotes, the protein–DNA complex is present in a non-membrane-bound region known as the nucleoid.

Phosphodiester bonds join the 3'-hydroxyl group of the deoxy pentose of one nucleotide to the 5'-hydroxyl group of the deoxy pentose of an adjacent nucleotide through a phosphate group. The resulting long, unbranched chain has polarity, with both a 5'-end (the end with the free phosphate) and a 3'-end (the end with the free hydroxyl) that are not attached to other nucleotides. The bases located along the resulting deoxy ribose–phosphate backbone are, by convention, always written in sequence from the 5'-end of the chain to the 3'-end. Phosphodiester linkages between nucleotides (in DNA or RNA) can be cleaved hydrolytically by chemicals, or hydrolyzed enzymatically by a family of nucleases: deoxyribonucleases for DNA and ribonucleases for RNA.

**STEPS IN PROKARYOTIC DNA SYNTHESIS**

When the two strands of the DNA double helix are separated, each can serve as a template for the replication of a new complementary strand. This produces two daughter molecules, each of which contains two DNA strands with an antiparallel orientation. This process is called semiconservative replication because, although the parental duplex is separated into two halves (and, therefore, is not “conserved” as an entity), each of the individual parental strands remains intact in one of the two new duplexes. The enzymes involved in the DNA replication process are template-directed polymerases that can synthesize the complementary sequence of each strand with extraordinary fidelity. The process of DNA synthesis was first known from studies of the bacterium Escherichia coli (E. coli). DNA synthesis in higher organisms is less well understood, but involves the same types of mechanisms. In either case, initiation of DNA replication commits the cell to continue the process until the entire genome has been replicated.

****

**Semiconservative replication of DNA**

**A. Separation of the two complementary DNA strands**

In order for the two strands of the parental double helical DNA to be replicated, they must first separate (or “melt”) over a small region, because the polymerases use only ssDNA as a template. In prokaryotic organisms, DNA replication begins at a single, unique nucleotide sequence—a site called the origin of replication. This is referred to as a consensus sequence, because the order of nucleotides is essentially the same at each site. This site includes a short sequence composed almost exclusively of AT base pairs that facilitate melting. In eukaryotes, replication begins at multiple sites along the DNA helix. Having multiple origins of replication provides a mechanism for rapidly replicating the great length of the eukaryotic DNA molecules.



**Origin of replication in prokaryotic and eukaryotic DNA**

**B. Formation of the replication fork**

As the two strands unwind and separate, they form a “V” where active synthesis occurs. This region is called the replication fork. It moves along the DNA molecule as synthesis occurs. Replication of dsDNA is bidirectional—that is, the replication forks move in opposite directions from the origin, generating a replication bubble.

**Proteins required for DNA strand separation:**

Initiation of DNA replication requires the recognition of the origin of replication by a group of proteins that form the pre-priming complex. These proteins are responsible for maintaining the separation of the parental strands, and for unwinding the double helix ahead of the advancing replication fork. These proteins include the following:

**a. DnaA protein:** DnaA protein binds to specific nucleotide sequences at the origin of replication, causing short, tandemly arranged AT-rich regions in the origin to melt. Melting is ATP dependent, and results in strand separation with the formation of localized regions of ssDNA.

**b. DNA helicases:** These enzymes bind to ssDNA near the replication fork, and then move into the neighbouring double-stranded region, forcing the strands apart—in effect, unwinding the double helix. Helicases require energy provided by ATP.

**c. Single-stranded DNA-binding (SSB) proteins:** These proteins bind to the ssDNA generated by helicases. They bind cooperatively—that is, the binding of one molecule of SSB protein makes it easier for additional molecules of SSB protein to bind tightly to the DNA strand. The SSB proteins are not enzymes, but rather serve to shift the equilibrium between dsDNA and ssDNA in the direction of the single-stranded forms. These proteins not only keep the two strands of DNA separated in the area of the replication origin, thus providing the single-stranded template required by polymerases, but also protect the DNA from nucleases that degrade ssDNA.

****

**Proteins responsible for maintaining the separation of the parental strands and unwinding the double helix ahead of the advancing replication fork.**

**Solving the problem of supercoils**

As the two strands of the double helix are separated, a problem is encountered, namely, the appearance of positive supercoils (also called supertwists) in the region of DNA ahead of the replication fork. The accumulating positive supercoils interfere with further unwinding of the double helix. To solve this problem, there is a group of enzymes called DNA topoisomerases, which are responsible for removing supercoils in the helix.

**a. Type I DNA topoisomerases:** These enzymes reversibly cut one strand of the double helix. They have both nuclease (strand-cutting) and ligase (strand-resealing) activities. They do not require ATP, but rather appear to store the energy from the phosphodiester bond they cleave, reusing the energy to reseal the strand. Each time a transient “nick” is created in one DNA strand, the intact DNA strand is passed through the break before it is resealed, thus relieving (“relaxing”) accumulated supercoils. Type I topoisomerases relax negative supercoils (that is, those that contain fewer turns of the helix than relaxed DNA) in E. coli, and both negative and positive supercoils (that is, those that contain fewer or more turns of the helix than relaxed DNA) in eukaryotic cells.

**b. Type II DNA topoisomerases:** These enzymes bind tightly to the DNA double helix and make transient breaks in both strands. The enzyme then causes a second stretch of the DNA double helix to pass through the break and, finally, reseals the break. As a result, both negative and positive supercoils can be relieved by this ATP-requiring process. Type II DNA topoisomerases are also required in both prokaryotes and eukaryotes for the separation of interlocked molecules of DNA following chromosomal replication.

**C. Direction of DNA replication**

The DNA polymerases responsible for copying the DNA templates are only able to “read” the parental nucleotide sequences in the 3'→5' direction, and they synthesize the new DNA strands only in the 5'→3' (antiparallel) direction. Therefore, beginning with one parental double helix, the two newly synthesized stretches of nucleotide chains must grow in opposite directions—one in the 5'→3' direction toward the replication fork and one in the 5'→3' direction away from the replication fork. This feat is accomplished by a slightly different mechanism on each strand.

**1. Leading strand:** The strand that is being copied in the direction of the advancing replication fork is called the leading strand and is synthesized continuously.

**2. Lagging strand:** The strand that is being copied in the direction away from the replication fork is synthesized discontinuously, with small fragments of DNA being copied near the replication fork. These short stretches of discontinuous DNA, termed Okazaki fragments, are eventually joined (ligated) to become a single, continuous

strand. The new strand of DNA produced by this mechanism is termed the lagging strand.

**D. RNA primer**

DNA polymerases cannot initiate synthesis of a complementary strand of DNA on a totally single-stranded template. Rather, they require an RNA primer—that is, a short, double-stranded region consisting of RNA base-paired to the DNA template, with a free hydroxyl group on the 3'-end of the RNA strand. This hydroxyl group serves as the first acceptor of a deoxynucleotide by action of DNA polymerase. A specific RNA polymerase, called primase (DnaG), synthesizes these RNA primers (approximately ten nucleotides long) that are complementary and antiparallel to the DNA template. In the resulting hybrid duplex, the U in RNA pairs with A in DNA. These short RNA sequences are constantly being synthesized at the replication fork on the lagging strand, but only one RNA sequence at the origin of replication is required on the leading strand. The substrates for this process are 5'-ribonucleoside triphosphates, and pyrophosphate is released as each ribonucleoside monophosphate is added through formation of a 3'→5' phosphodiester bond.

****

**Elongation of the leading and lagging strands**

**E. Chain elongation**

Prokaryotic (and eukaryotic) DNA polymerases elongate a new DNA strand by adding deoxyribonucleotides, one at a time, to the 3'- end of the growing chain. The sequence of nucleotides that are added is dictated by the base sequence of the template strand with which the incoming nucleotides are paired.

**1. DNA polymerase III:** DNA chain elongation is catalyzed by DNA polymerase III. Using the 3'-hydroxyl group of the RNA primer as the acceptor of the first deoxyribonucleotide, DNA polymerase III begins to add nucleotides along the single-stranded template that specifies the sequence of bases in the newly synthesized chain. DNA polymerase III is a highly “processive” enzyme—that is, it remains bound to the template strand as it moves along, and does not diffuse away and then rebind before adding each new nucleotide. The processivity of DNA polymerase III is the result of its β subunit forming a ring that encircles and moves along the template strand of the DNA, thus serving as a sliding DNA clamp. The new strand grows in the 5'→3' direction, antiparallel to the parental strand. The nucleotide substrates are 5'-deoxy ribo nucleoside triphosphates. Pyrophosphate (PPi) is released when each new deoxynucleoside monophosphate is added to the growing chain. Hydrolysis of PPi to 2Pi means that a total of two high-energy bonds are used to drive the addition of each deoxynucleotide. All four deoxyribonucleoside triphosphates (dATP, dTTP, dCTP,and dGTP) must be present for DNA elongation to occur. If one of the four is in short supply, DNA synthesis stops when that nucleotide is depleted.

**2. Proofreading of newly synthesized DNA:** It is highly important for the survival of an organism that the nucleotide sequence of DNA be replicated with as few errors as possible. Misreading of the template sequence could result in deleterious, perhaps lethal, mutations. To ensure replication fidelity, DNA polymerase III has, in addition to its 5'→3' polymerase activity, a “proofreading” activity (3'→5' exonuclease). As each nucleotide is added to the chain, DNA polymerase III checks to make certain the added nucleotide is, in fact, correctly matched to its complementary base on the template. If it is not, the 3'→5' exonuclease activity corrects the mistake. The enzyme requires an improperly base-paired 3'-hydroxy terminus and, therefore, does not degrade correctly paired nucleotide sequences. For example, if the template base is cytosine and the enzyme mistakenly inserts an adenine instead of a guanine into the new chain, the 3'→5' exonuclease activity hydrolytically removes the misplaced nucleotide. The 5' →3' polymerase activity then replaces it with the correct nucleotide containing guanine. The proofreading exonuclease activity requires movement in the 3'→5' direction, not 5' →3' like the polymerase activity. This is because the excision must be done in the reverse direction from that of synthesis.

**F. Excision of RNA primers and their replacement by DNA**

DNA polymerase III continues to synthesize DNA on the lagging strand until it is blocked by proximity to an RNA primer. When this occurs, the RNA is excised and the gap filled by DNA polymerase I.

**1. 5'**→**3' Exonuclease activity:** In addition to having the 5'→3' polymerase activity that synthesizes DNA, and the 3' →5' exonuclease activity that proofreads the newly synthesized DNA chain like DNA polymerase III, DNA polymerase I also has a 5'→3' exonuclease activity that is able to hydrolytically remove the RNA primer. These activities are exonucleases because they remove one nucleotide at a time from the end of the DNA chain, rather than cleaving the chain internally as do the endonucleases. First, DNA polymerase I locates the space (“nick”) between the 3'-end of the DNA newly synthesized by DNA polymerase III and the 5'-end of the adjacent RNA primer. Next, DNA polymerase I hydrolytically removes the RNA nucleotides “ahead” of itself, moving in the 5'→3' direction (5'→3' exonuclease activity). As it removes the RNA, DNA polymerase I replaces it with deoxyribonucleotides, synthesizing DNA in the 5'→3' direction (5'→3' polymerase activity). As it synthesizes the DNA, it also “proofreads” the new chain using 3'→5' exonuclease activity. This removal/synthesis/proofreading continues, one nucleotide at a time, until the RNA primer is totally degraded and the gap is filled with DNA.

****

**Removal of RNA primer and filling of the resulting “gaps” by *DNA polymerase I*.**

**2. Differences between 5'**→**3' and 3'**→**5' exonucleases:** The 5'→3' exonuclease activity of DNA polymerase I differs from the 3'→5' exonuclease used by both DNA polymerase I and III in two important ways. First, 5'→3' exonuclease can remove one nucleotide at a time from a region of DNA that is properly base-paired. The nucleotides it removes can be either ribonucleotides or deoxyribonucleotides. Second, 5'→3' exonuclease can also remove groups of altered nucleotides in the 5'→3' direction, removing from one to ten nucleotides at a time. This ability is important in therepair of some types of damaged DNA.

**G. DNA ligase**

The final phosphodiester linkage between the 5'-phosphate group on the DNA chain synthesized by DNA polymerase III and the 3'- hydroxyl group on the chain made by DNA polymerase I is catalyzed by DNA ligase. The joining of these two stretches of DNA requires energy, which in most organisms is provided by the cleavage of ATP to AMP + PPi.

**EUKARYOTIC DNA REPLICATION**

The process of eukaryotic DNA replication closely follows that of prokaryotic DNA synthesis. Some differences, such as the multiple origins of replication in eukaryotic cells versus single origins of replication in prokaryotes, have already been noted. Eukaryotic single-stranded DNA-binding proteins and ATP-dependent DNA helicases have been identified, whose functions are analogous to those of the prokaryotic enzymes previously discussed.

**A. The eukaryotic cell cycle**

The events surrounding eukaryotic DNA replication and cell division (mitosis) are coordinated to produce the cell cycle. The period preceding replication is called the G1 phase (Gap1). DNA replication occurs during the S (synthesis) phase. Following DNA synthesis, there is another period (G2phase, or Gap2) before mitosis (M). Cells that have stopped dividing, such as mature neurons, are said to have gone out of the cell cycle into the G0phase. Cells can leave the G0 phase and reenter the early G1phase to resume division. The cell cycle is controlled at a series of “checkpoints” that prevent entry into the next phase of the cycle until the preceding phase has been completed. Two key classes of proteins that control the progress of a cell through the cell cycle are the cyclins and cyclin-dependent kinases (Cdk).



**The eukaryotic cell cycle**

**B. Eukaryotic DNA polymerases**

At least five key eukaryotic DNA polymerases have been identified and categorized on the basis of molecular weight, cellular location, sensitivity to inhibitors, and the templates or substrates on which they act. They are designated by Greek letters rather than by Roman numerals.

**1. Pol α:** Pol **α** is a multisubunit enzyme. One subunit has primase activity, which initiates strand synthesis on the leading strand and at the beginning of each Okazaki fragment on the lagging strand. The primase subunit synthesizes a short RNA primer that is extended by the pol **α** 5'→3' polymerase activity, generating a short piece of DNA.

**2. Pol ε and pol δ:** Pol **ε** is thought to be recruited to complete DNA synthesis on the leading strand whereas pol **δ** elongates the Okazaki fragments of the lagging strand, each using 3'→5' exonuclease activity to proofread the newly synthesized DNA. [Note: DNA polymerase **ε** associates with the protein, proliferating cell nuclear antigen (PCNA), which serves as a sliding DNA clamp in much the same way the β subunit of DNA polymerase III does in E. coli, thus ensuring high processivity].

**3. Pol** β **and pol** γ: Pol β is involved in "gap filling" in DNA repair. Pol γ replicates mitochondrial DNA.

**C. Telomeres**

Telomeres are complexes of noncoding DNA plus proteins located at the ends of linear chromosomes. They maintain the structural integrity of the chromosome, preventing attack by nucleases, and allow repair systems to distinguish a true end from a break in dsDNA. In humans, telomeric DNA consists of several thousand tandem repeats of a noncoding hexameric sequence, AG3T2, base-paired to a complementary region of Cs and As. The GT-rich strand is longer than its CA complement, leaving ssDNA a few hundred nucleotides in length at the 3'-end. The single-stranded region is thought to fold back on itself, forming a loop structure that is stabilized by protein.

**1. Telomere shortening:** Eukaryotic cells face a special problem in replicating the ends of their linear DNA molecules. Following removal of the RNA primer from the extreme 5'-end of the lagging strand, there is no way to fill in the remaining gap with DNA. Consequently, in most normal human somatic cells, telomeres shorten with each successive cell division. Once telomeres are shortened beyond some critical length, the cell is no longer able to divide and is said to be senescent. In germ cells and other stem cells, as well as in cancer cells, telomeres do not shorten and the cells do not senesce. This is a result of the presence of a ribonucleoprotein, telomerase, which maintains telomeric length in these cells.

**2. Telomerase:** This complex contains a protein that acts as a reverse transcriptase, and a short piece of RNA that acts as a template. The CA-rich RNA template base-pairs with the GT-rich, single-stranded 3'-end of telomeric DNA. The reverse transcriptase uses the RNA template to synthesize DNA in the usual 5'→3' direction, extending the already longer 3'-end. Telomerase then translocates to the newly synthesized end, and the process is repeated. Once the GT-rich strand has been lengthened, primase can use it as a template to synthesize an RNA primer. The RNA primer is extended by DNA polymerase, and the primer is removed.

**D. Inhibition of DNA synthesis by nucleoside analogs**

DNA chain growth can be blocked by the incorporation of certain nucleoside analogs that have been modified in the sugar portion of the nucleoside. For example, removal of the hydroxyl group from the 3'-carbon of the deoxyribose ring as in 2',3'-dideoxyinosine (ddI, also known as didanosine), or conversion of the deoxyribose to another sugar such as arabinose, prevents further chain elongation. By blocking DNA replication, these compoundsslow the division of rapidly growing cells and viruses. Cytosine arabinoside (cytarabine, or araC) has been used in anticancer chemotherapy, whereas adenine arabinoside (vidarabine, or araA) is an antiviral agent. Chemically modifying the sugar moiety, as seen in zidovudine (AZT, ZDV), also terminates DNA chain elongation. These drugs are generally supplied as nucleosides, which are then converted to the active nucleotides by cellular “salvage” enzymes.