TRANSCRIPTION. The process of RNA synthesis directed by a DNA template is termed transcription, and it occurs in three phases: initiation, elongation, and termination.

Elongation. The basic requirements and fundamental mechanism of the elongation phase of RNA synthesis is the **same in prokaryotes and eukaryotes** and must be understood before initiation and termination can be understood.

1. Template. A single strand of DNA acts as a template to direct the formation of complementary RNA during transcription.

2. Substrate. The substrates for RNA synthesis are the four ribonucleoside triphos phates: adenosine triphosphate (ATP), guanosine triphosphate (GTP), cytidine triphosphate (CTP), and uridine triphosphate (UTP). Cleavage of the high-energy phosphate bond between the a and (3 phosphates of these nucleoside triphosphates provides the energy for the addition of nucleotides to the growing RNA chain.

3. Direction of synthesis. After the first nucleoside triphosphate, subsequent nucleotides are added to the 3'-hydroxyl of the preceding nucleotide. **Therefore, RNA chain growth proceeds in the 5' to 3' direction.**

Enzyme

a. Prokaryotes have a single RNA polymerase responsible for all cellular RNA synthesis. The structure of this enzyme is complex.

(1) **A core enzyme** with the subunit structure α_2 (40,000 MW), β (155,000 MW), and β' (160,000 MW) is required for the **elongation** steps of RNA synthesis.

(2) **A holoenzyme,** which is a core enzyme with an additional subunit σ (85,000 MW), is required for proper **initiation** of transcription.

b. Eukaryotes have one mitochondrial and three nuclear RNA polymerases. The **nuclear RNA polymerases are distinct enzymes that function to synthesize different RNAs.**

(1) Structure. The nuclear RNA polymerases are large enzymes of greater than 500,000 MW. Each has a very complex subunit composition with more than 10 subunit polypeptides.

(2) Location and function. The subnuclear localization and the RNA product of each type of RNA polymerase are determined by the RNA polymerase's differential sensitivity to α**-amanitin** (a toxic bicyclic octapeptide from the poisonous mushroom *Amanita phalloides).*

Other enzymatic activities. Unlike some of the DNA polymerases, **the RNA polymerases have no other associated enzyme activities,** such as proofreading.

Initiation of transcription

Promoter sequences. Unlike the initiation of replication, transcriptional initiation does not require a primer. **Promoter sequences are responsible for directing RNA polymerase to initiate transcription at a particular point.** Promoter sequences differ between prokaryotes and eukaryotes.

Nomenclature and numbering conventions are used to avoid confusion in the description of sequences, such as promoter sequences, and their location in genes.

(1) Because DNA is double-stranded, the **sequence of only one strand is presented.**

(2) Because RNA synthesis occurs in a 5' to 3' direction, and sequences are written in a 5' to 3' direction, the sequence of the DNA strand that is **identical to the RNA transcript** is presented.

(3) For any one gene, the location of particular sequences related to the expression of that gene is given relative to its transcriptional startpoint. Therefore, **position 1 of a gene is the base that is equivalent to the first base of the 5' end of the RNA transcript of that gene.**

(4) Sequences preceding the first base are numbered negatively and said to be **upstream** of the initiation point. Sequences following the first base are numbered positively and are said to be **downstream** of the initiation point.

Prokaryotic promoters. For most prokaryotic genes, there are **conserved sequences** that are necessary to promote accurate initiation of transcription. The **promoters for most prokaryotic genes have three sequence elements** (Figure 9-3A).

(1) Initiation site (i.e., startpoint). Transcription for most genes always starts at the same base (position 1). **The startpoint is usually a purine.**

(2) Pribnow box. For all prokaryotic genes there is a sequence called the Prib-now box that lies 9-18 base pairs upstream of the startpoint.

(a) The sequence of a typical Pribnow box is either identical to or very similar to the sequence **TATAAT.**

(b) The Pribnow box also has been called the **-10 sequence** because it is usually found 10 base pairs upstream of the startpoint.

(3) The **-35 sequence** is a component of typical prokaryotic promoters. It is a sequence that is either identical to or very similar to the sequence **TTGACA.** It is named the -35 sequence because it is typically found 35 base pairs upstream of the startpoint.

Eukaryotic promoters. Each type of eukaryotic RNA polymerase uses a different promoter. The promoters used by RNA polymerase I and II are similar to the prokaryotic promoter in that they are upstream of the startpoint. However, the promoters used by RNA polymerase III are unique because they are usually downstream of the startpoint.

(1) **Initiation site (i.e., startpoint).** The promoters of all the eukaryotic RNA polymerases direct the initiation of transcription to a particular startpoint. As in prokaryotes, the startpoint is usually a purine, although pyrimidine startpoints are not uncommon in eukaryotes.

(2) **Multiple RNA polymerase II promoter sequence elements.** Unlike RNA polymerase I promoters, which are all the same, and prokaryotic promoters, which are similar, **there is much diversity among RNA polymerase II promoters.** Several sequence elements are conserved and common to many of these RNA polymerase II promoters.

(a) The TATA box is a sequence either identical to or very similar to the sequence TATA(T or A)T. It is typically found 25-35 base pairs upstream of the startpoint.

(b) **The CAAT box** and the **GC box** are sequences similar to CCAAT and GGGCG, respectively, that are found one or more times anywhere from 40-200 base pairs upstream of the startpoint.

(3) The **RNA polymerase III promoters** are of interest because they usually occur downstream of the startpoint.

(a) The **5S rRNA promoter** is made of two sequence elements; one is the sequence 50-70 base pairs downstream of the startpoint and the other is the sequence 80-90 base pairs downstream of the startpoint. Both are in the transcribed portion of the gene.

(b) The **tRNA promoters** are also made of two sequence elements that are downstream of the startpoint and within the transcribed portion of the gene; one is between +8 and +30, and one is between +50 and +70.

(4) Other sequence elements—**response elements and enhancers—affect the rate of initiation of transcription** but are not strictly promoter elements because they do not have an effect on the accuracy of the initiation of transcription.

Initiation factors are needed to initiate transcription. In prokaryotes, only a single factor, **sigma** *(σ),* is needed to initiate transcription. In eukaryotes, multiple factors are required, in part because of the diversity of promoters.

The prokaryotic σ factor is required for accurate initiation of transcription.

Function. The σ factor enables the RNA polymerase holoenzyme to recognize and bind tightly to the promoter sequences.

Process

- Upon binding, the σ factor facilitates the opening or melting of the DNA double helix.
- The enzyme then catalyzes the formation of a phosphodiester bond between the first two bases. The first base is usually a purine nucleoside triphosphate (pppA or pppG).
- Elongation proceeds after the formation of the first phosphodiester bond. By the time 10 nucleotides have been added, the σ factor dissociates. The core enzyme then continues the elongation of the transcript.
- The released σ factor can combine with free core enzyme to form another holoenzyme that can initiate transcription.

Rifampin is an effective antibacterial drug and one of the few therapeutic drugs that affects only transcription. Actions of rifampin include the following.

- **Rifampin binds to the β subunit of RNA polymerase when the polymerase is in the holoenzyme form.**
- Through binding of the p subunit of the holoenzyme, rifampin **specifically inhibits initiation of transcription** and not elongation.
- Rifampin has **no effect on eukaryotic nuclear RNA polymerases.**

Eukaryotic initiation factors. The initiation of transcription in eukaryotes is considerably more complex than in prokaryotes, partly because of the increased complexity of eukaryotic RNA polymerases and partly because of the diversity of their promoters.

(1) Multiple factors and RNA polymerase II are needed to initiate transcription from TATA box promoters. Including RNA polymerase II, more than 40 different polypeptides are needed to initiate transcription.

    (a) Transcription factor IID (TFIID) recognizes and binds to the TATA box sequences independently of RNA polymerase II.

        (i) The **TATA box-binding protein (TBP)** is the subunit of TFIID that binds to the TATA-box DNA sequence.
        (ii)TFIID is made of eight other subunits called **TBP-associated factors** (TAP).

    (b) Five other transcription factors are required for the proper initiation of transcription by RNA polymerase II (TFIIA, TFIIB, TFIIF, TFIIE, and TFIIH).

(2) RNA polymerase I and III also need specific transcription factors to initiate transcription from their respective promoters. Other factors are needed for the recognition of response elements and enhancers, which modulate the rate of transcription.

Termination. Prokaryotes and eukaryotes use an identical mechanism of synthesizing RNA and share many similarities in the way that they initiate transcription. They appear, however, to have very little in common in the way they terminate transcription.

Prokaryotes. There are two basic classes of termination events in prokaryotes.

Factor-independent termination. Particular sequences can cause the core enzyme to terminate transcription. These sequences share several common **features**.

> (1) All of these sequences have a sequence that codes for a self-complementary sequence that **can form a stable stem-loop structure**.
>
> (2) The DMA codes for a **stretch of uracils** (Us) to be formed just after the stem-loop region.
>
> **(3) Termination becomes favorable** when the transcript forms a stable stem-loop structure. This structure causes the RNA polymerase to slow its synthesis, and the transcript is displaced when the slowed RNA polymerase synthesizes the U-rich segment. Displacement occurs easily because only weak adenine-uracil bonds hold the transcript to the template.
>
> (4) Because of the nature of this displacement, there is **no specific base where transcription stops** (i.e., different transcripts have a different number of uracil residues on their 3' end).

Factor-dependent termination. Particular sequences act as termination sequences in the presence of factor rho (p).

> **(1) Rho-dependent termination sequences** do not appear to share common structural features as do the factor-independent termination sequences.
>
> **(2) Rho binds as a hexamer** to the forming transcript at these unique sequences.
>
> **(3) Rho is an ATPase.** The exact mechanism that rho uses to terminate transcription is unknown, but it requires the cleavage of ATP by rho.

Eukaryotes. Compared with prokaryotes, very little is known about how eukaryotes terminate transcription.

a.   **RNA polymerase I** terminates transcription in a factor-dependent manner at a particular sequence.

b.   **RNA polymerase III** terminates transcription by an unknown mechanism after the synthesis of a series of U residues.

c.   There is no known transcriptional termination signal for **RNA polymerase II. For** most genes, it transcribes up to several thousand base pairs beyond the point of polyadenylation.

POST-TRANSCRIPTIONAL RNA PROCESSING. Once a gene transcript has been synthesized, numerous post-transcriptional modification or processing events may be needed before the transcript is functional.

Prokaryotes. Post-transcriptional processing of RNA is not as extensive in prokaryotes as in eukaryotes; however, some processing does occur.

1. In prokaryotes, **mRNA is not post-transcriptionally processed.** Prokaryotic mRNA is functional immediately upon synthesis. In fact, its translation often begins before transcription is complete.
2. Seven genes produce rRNA. Each gene produces a **SOS precursor rRNA** that is processed to discrete, functional rRNAs.
   a. All seven genes contain the sequences that become 23S, 16S, **and** 5S rRNA. Within the transcribed portion of these genes are some of the **tRNA genes.** Different rRNA genes contain different tRNA genes.
   b. **Cleavage.** Upon formation of the SOS rRNA precursor, the nonfunctional spacer sequences are removed by a series of specific endonucleolytic cleavages by the enzymes **ribonuclease P and ribonuclease III.**
   c. **Base modification.** In addition to the removal of the spacer sequences, some of the bases in the final rRNAs are **methylated.** Methylation is needed for the rRNAs to be functional.
3. The tRNAs not formed from processing of the precursor rRNA arise from **large precursor transcripts,**
   a. The tRNA genes are clustered, and each transcript contains sequences for two to seven tRNAs.
   b. **Cleavage.** The portions of the transcript that form functional tRNAs are removed by the enzymes ribonuclease P and ribonuclease D.
   c. **Addition of the sequence CCA to the 3' end.** After excision from the precursor, some of the tRNAs are left without the sequence CCA, which is common to all tRNAs, on the 3' end. This sequence is added to these tRNAs by the enzyme **tRNA nucleotidyl transferase.**
   d. **Base modification.** Many of the bases of the tRNAs are modified. These modifications are various methylations and other more extensive modifications of some of the bases. These modifications are necessary for the tRNAs to adopt their unique, functional conformations.