

RNA Synthesis

The process of synthesizing RNA from the genetic information encoded by DNA is called transcription. The enzymes involved in transcription are called RNA polymerases. Prokaryotes have one type; eukaryotes have three types of nuclear RNA polymerases.

The prokaryotic RNA polymerase consists of a core enzyme and an auxiliary protein factor called sigma (σ factor). The core consists of four subunits, two are identical, α , the other two similar, β and β' . The β' subunit binds the DNA while the β subunit binds the nucleotides that are to be joined together to form the RNA molecule. Sigma factors function in identifying specific DNA sequences known as promoters. Promoters are sites that tell the RNA polymerase where to begin transcription.

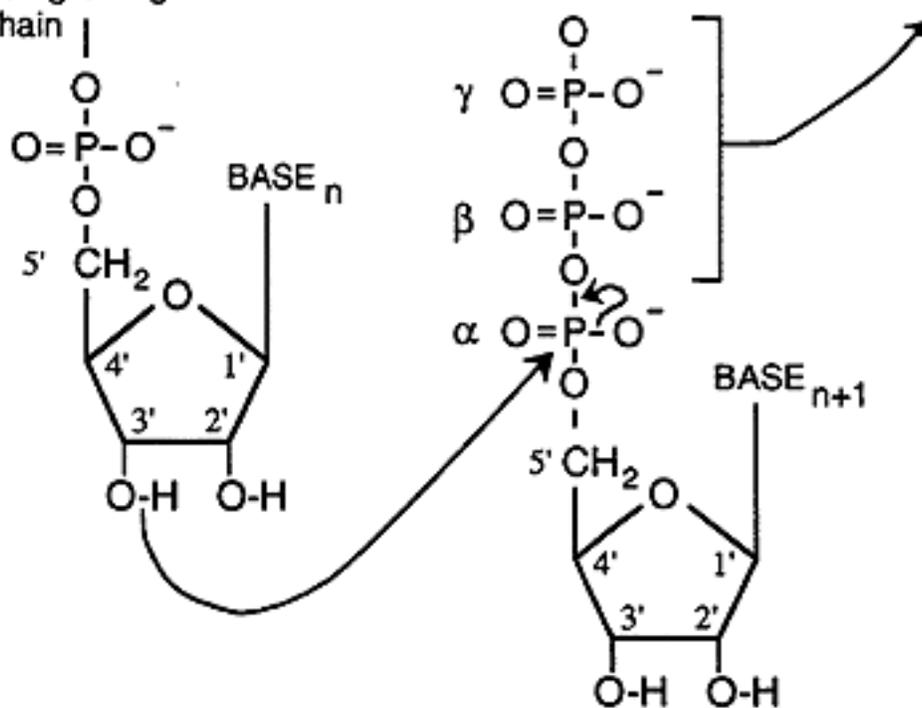
Eukaryotes have four different RNA polymerases (RNA pol). Three are required for transcription of nuclear genes and the fourth for transcription of mitochondrial genes. RNA polymerase I transcribes ribosomal RNA (rRNA), pol II transcribes mRNA and pol III tRNA and several small RNA's. The three polymerases consist of ten or more subunits. All have two large subunits with homology to the β and β' subunits of the prokaryotic RNA polymerase. The three eukaryotic polymerases can be distinguished based on their sensitivity to α -amanitin, a toxin found in some types of mushrooms. RNA pol II activity is severely inhibited, pol III weakly and pol I is insensitive. The antibiotic rifampicin inhibits prokaryotic RNA polymerases.

There are three phases of transcription: initiation, elongation and termination. It is easier to understand the process by first examining elongation then initiation and termination.

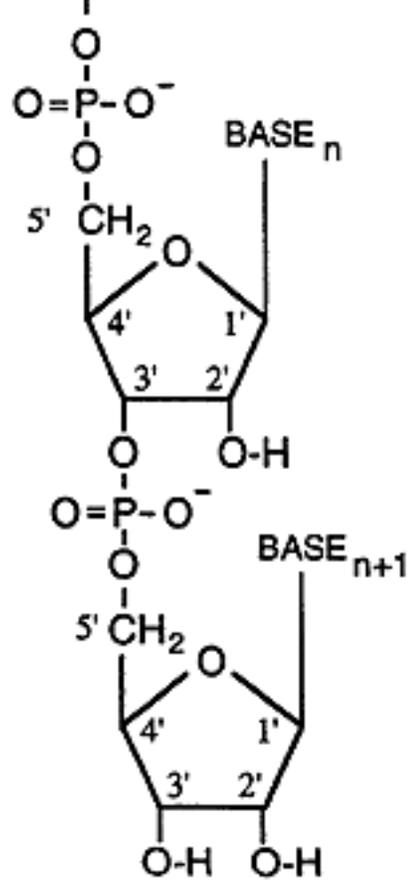
Elongation

RNA polymerase links ribonucleotides together in a 5' to 3' direction. The polymerase induces the 3' hydroxyl group of the nucleotide at the 3' end of the growing RNA chain which attacks (nucleophilic) the a phosphorous of the incoming ribonucleotide. A diphosphate is released and the 5' carbon of the incoming nucleotide is linked through a phosphodiester bond to the 3' carbon of the preceding nucleotide.

5' end of growing
RNA chain



5' end of growing
RNA chain



Nucleotide incorporation is determined by base pairing with the template strand of the DNA. The template is the DNA strand, also called the sense strand, that is copied by the RNA polymerase into a complementary strand of RNA called the transcript. The DNA strand that is not copied is known as the antisense strand. Note that while the RNA chain grows in a 5' to 3' direction the polymerase migrates along the sense strand in a 3' to 5' direction. Thus the 5' to 3' ribonucleotide sequence of the RNA transcript is identical to the 5' to 3' antisense DNA strand with uracil in place of thymidine.

Initiation

The initiation of transcription is directed by DNA sequences called promoters which tell the RNA polymerase where to begin transcription. The subunits that enable RNA polymerases to recognize and bind promoters are called initiation factors. The initiating nucleotide can be either a purine or pyrimidine. There are numerous eukaryotic promoters with multiple promoter sequence elements. Some of the elements specify where transcription is to be initiated, others determine the frequency with which transcription is initiated at a specific gene. The initiation of transcription in eukaryotes is complicated and involves numerous factors (proteins) that must interact with the DNA and with one another to initiate transcription.

Promoters

- 1 Only one strand of the DNA that encodes a promoter, a regulatory sequence, or a gene needs to be written.
- 2 The strand that is written is the one that is identical to the RNA transcript, thus the antisense strand of the DNA is always selected for presentation.
- 3 The first base on the DNA where transcription actually starts is labeled +1.

4 Sequences that precede, are upstream of the first base of the transcript, are labeled with negative numbers. Sequences that follow the first base of the transcript, are downstream, are labeled with positive numbers.

RNA pol II promoters are quite diverse. This enables the cell to choose and regulate the expression of the 50 to 100 thousand different genes encoded by its DNA. There are some sequence elements that are conserved and found in most RNA pol II promoters. There are three "boxes": TATA usually found 25 to 35 base pairs upstream, the CAAT box and the GC box both located from 40 to 200 base pairs upstream.

These three elements provide a basal level of transcription and are found in most "housekeeping" genes. Housekeeping genes encode enzymes and proteins that all cell types require for normal function and are usually expressed at steady state or basal levels. Other sequence elements, which are continually being discovered, serve as regulatory elements. Elements that enable a cell to specifically turn other non-housekeeping genes on or off in response to environmental signals such as hormones, growth factors, metals and toxins. The spacing and orientation of all of the sequence elements are critical for proper functioning. There is a third type of sequence element that can be located either upstream or downstream relative to the initiation site which is called an enhancer or silencer. Enhancers or silencers affect the rate and frequency of initiation of transcription.

RNA pol III promoters for tRNA are found downstream of the initiation point. These promoters consist of two elements, the first of which is located 8 to 30 base pairs downstream and is called Box A. The second element is 50 to 70 base pairs downstream and is called Box B.

RNA pol I promoter consists of a 70 base pair long core element and an upstream element that is about 100 base pairs long. The core spans a

segment of DNA that includes sequences that are both up and downstream of the initiation site.

Termination

Prokaryotes use two means for terminating transcription, factor-independent and factor-dependent. Certain DNA sequences function as signals that tell the RNA polymerase to terminate transcription. The DNA of a terminator sequence encoded an inverted repeat and an adjacent stretch of uracils. Factor-dependent termination involves a terminator sequence as well as a factor or protein called rho. The mechanisms by which eukaryotes terminate transcription are poorly understood. Most eukaryotic genes are transcribed for up to several thousand base pairs beyond the actual end of the gene. The excess RNA is then cleaved from the transcript when the RNA is processed into its mature form.

RNA Processing

Most transcripts must be processed before becoming fully functional. Most eukaryotic RNA must be transported across the nuclear membrane where it is processed then transported to the cytosol. Processing helps stabilize and protect the RNA so it can function in the cytosol and also functions in regulating the expression of certain genes.

Mature mRNA is formed by extensively modifying the primary transcript also called heterogeneous nuclear RNA (hnRNA). The hnRNA must undergo three major modifications before maturing into mRNA: capping, polyadenylation and splicing.

Capping: all mRNA's are capped at their 5' ends with 7-methylguanylate. Guanylyl transferase catalyzes the linking of 7'-methylguanylate to the mRNA through a 5' to 5' triphosphate bridge. The capping positions the mRNA onto the 40S preinitiation complex and protects it from exonuclease

activity.

Polyadenylation: is the addition of a chain of adenylate residues, known as a poly A tail to the 3' terminus of mRNA. After the RNA is cut, an enzyme poly A polymerase, catalyzes the polymerization of adenylates. The poly A tail slows the exonucleolytic degradation of mRNA, once the tail is removed mRNA is quickly degraded.

Splicing: is the removal of noncoding sequences, derived from the DNA template, from the hnRNA to form a functional mRNA. The noncoding sequences are called introns while the coding sequences are known as exons. All introns have the sequence GU at their 5' ends and AG at their 3' ends. The guanyl residue at the 5' end of the intron is linked by a 2' to 5' phosphodiester linkage to an adenylate residue within the intron. The result is a lariat (loop) structure and the release of the 3' end of the first exon. The 3' end of the intron is spliced by an enzyme known as a spliceosome, which releases the loop and frees the 5' end of the second exon. The exons are then joined together.

The **rRNA** of both prokaryotes and eukaryotes are synthesized as large precursors. The precursor rRNA's are processed into their mature form by nucleases and methylases.

The **tRNA's** of both prokaryotes and eukaryotes are also transcribed as precursors which are cleaved and extensively modified.