7.5A: Elongation and Termination in Eukaryotes

Elongation synthesizes pre-mRNA in a 5’ to 3’ direction, and termination occurs in response to termination sequences and signals.

LEARNING OBJECTIVES

Describe what is happening during transcription elongation and termination

Key Takeaways

Key Points

• RNA polymerase II (RNAPII) transcribes the major share of eukaryotic genes.
• During elongation, the transcription machinery needs to move histones out of the way every time it encounters a nucleosome.
• Transcription elongation occurs in a bubble of unwound DNA, where the RNA Polymerase uses one strand of DNA as a template to catalyze the synthesis of a new RNA strand in the 5’ to 3’ direction.
• RNA Polymerase I and RNA Polymerase III terminate transcription in response to specific termination sequences in either the DNA being transcribed (RNA Polymerase I) or in the newly-synthesized RNA (RNA Polymerase III).
• RNA Polymerase II terminates transcription at random locations past the end of the gene being transcribed. The newly-synthesized RNA is cleaved at a sequence-specified location and released before transcription terminates.
Key Terms

- **nucleosome**: any of the subunits that repeat in chromatin; a coil of DNA surrounding a histone core
- **histone**: any of various simple water-soluble proteins that are rich in the basic amino acids lysine and arginine and are complexed with DNA in the nucleosomes of eukaryotic chromatin
- **chromatin**: a complex of DNA, RNA, and proteins within the cell nucleus out of which chromosomes condense during cell division

Transcription through Nucleosomes

Following the formation of the pre-initiation complex, the polymerase is released from the other transcription factors, and elongation is allowed to proceed with the polymerase synthesizing RNA in the 5′ to 3′ direction. RNA Polymerase II (RNAPII) transcribes the major share of eukaryotic genes, so this section will mainly focus on how this specific polymerase accomplishes elongation and termination.

Although the enzymatic process of elongation is essentially the same in eukaryotes and prokaryotes, the eukaryotic DNA template is more complex. When eukaryotic cells are not dividing, their genes exist as a diffuse, but still extensively packaged and compacted mass of DNA and proteins called chromatin. The DNA is tightly packaged around charged histone proteins at repeated intervals. These DNA–histone complexes, collectively called nucleosomes, are regularly spaced and include 146 nucleotides of DNA wound twice around the eight histones in a nucleosome like thread around a spool.

For polynucleotide synthesis to occur, the transcription machinery needs to move histones out of the way every time it encounters a nucleosome. This is accomplished by a special protein dimer called FACT, which stands for “facilitates chromatin transcription.” FACT partially disassembles the nucleosome immediately ahead (upstream) of a transcribing RNA Polymerase II by removing two of the eight histones (a single dimer of H2A and H2B histones is removed.) This presumably sufficiently loosens the DNA wrapped around that nucleosome so that RNA Polymerase II can transcribe through it. FACT reassembles the nucleosome behind the RNA Polymerase II by returning the missing histones to it. RNA Polymerase II will continue to elongate the newly-synthesized RNA until transcription terminates.

The FACT protein dimer allows RNA Polymerase II to transcribe through packaged DNA: DNA in eukaryotes is packaged in nucleosomes, which consist of an octomer of 4 different histone proteins. When DNA is tightly wound twice
around a nucleosome, RNA Polymerase II cannot access it for transcription. FACT removes two of the histones from the nucleosome immediately ahead of RNA Polymerase, loosening the packaging so that RNA Polymerase II can continue transcription. FACT also reassembles the nucleosome immediately behind the RNA Polymerase by returning the missing histones.

**Elongation**

RNA Polymerase II is a complex of 12 protein subunits. Specific subunits within the protein allow RNA Polymerase II to act as its own helicase, sliding clamp, single-stranded DNA binding protein, as well as carry out other functions. Consequently, RNA Polymerase II does not need as many accessory proteins to catalyze the synthesis of new RNA strands during transcription elongation as DNA Polymerase does to catalyze the synthesis of new DNA strands during replication elongation.

However, RNA Polymerase II does need a large collection of accessory proteins to initiate transcription at gene promoters, but once the double-stranded DNA in the transcription start region has been unwound, the RNA Polymerase II has been positioned at the +1 initiation nucleotide, and has started catalyzing new RNA strand synthesis, RNA Polymerase II clears or “escapes” the promoter region and leaves most of the transcription initiation proteins behind.

All RNA Polymerases travel along the template DNA strand in the 3′ to 5′ direction and catalyze the synthesis of new RNA strands in the 5′ to 3′ direction, adding new nucleotides to the 3′ end of the growing RNA strand.

RNA Polymerases unwind the double stranded DNA ahead of them and allow the unwound DNA behind them to rewind. As a result, RNA strand synthesis occurs in a transcription bubble of about 25 unwound DNA basepairs. Only about 8 nucleotides of newly-synthesized RNA remain basepaired to the template DNA. The rest of the RNA molecules falls off the template to allow the DNA behind it to rewind.

RNA Polymerases use the DNA strand below them as a template to direct which nucleotide to add to the 3′ end of the growing RNA strand at each point in the sequence. The RNA Polymerase travels along the template DNA one nucleotide at a time. Whichever RNA nucleotide is capable of basepairing to the template nucleotide below the RNA Polymerase is the next nucleotide to be added. Once the addition of a new nucleotide to the 3′ end of the growing strand has been catalyzed, the RNA Polymerase moves to the next DNA nucleotide on the template below it. This process continues until transcription termination occurs.
Transcription termination by RNA Polymerase II on a protein-encoding gene.: RNA Polymerase II has no specific signals that terminate its transcription. In the case of protein-encoding genes, a protein complex will bind to two locations on the growing pre-mRNA once the RNA Polymerase has transcribed past the end of the gene. CPSF in the complex will bind a AAUAAA sequence, and CstF in the complex will bind a GU-rich sequence (top figure). CPSF in the complex will cleave the pre-mRNA at a site between the two bound sequences, releasing the pre-mRNA (middle figure). Poly(A) Polymerase is a part of the same complex and will begin to add a poly-A tail to the pre-mRNA. At the same time, Xrn2 protein, which is an exonuclease, attacks the 5’ end of the RNA strand still associated with the RNA Polymerase. Xrn2 will start digesting the non-released portion of the newly synthesized RNA until Xrn2 reaches the RNA Polymerase, where it aids in displacing the RNA Polymerase from the template DNA strand. This terminates transcription at some random location downstream from the true end of the gene (bottom figure).

The termination of transcription is different for the three different eukaryotic RNA polymerases.
The ribosomal rRNA genes transcribed by RNA Polymerase I contain a specific sequence of basepairs (11 bp long in humans; 18 bp in mice) that is recognized by a termination protein called TTF-1 (Transcription Termination Factor for RNA Polymerase I.) This protein binds the DNA at its recognition sequence and blocks further transcription, causing the RNA Polymerase I to disengage from the template DNA strand and to release its newly-synthesized RNA.

The protein-encoding, structural RNA, and regulatory RNA genes transcribed by RNA Polymerase II lack any specific signals or sequences that direct RNA Polymerase II to terminate at specific locations. RNA Polymerase II can continue to transcribe RNA anywhere from a few bp to thousands of bp past the actual end of the gene. However, the transcript is cleaved at an internal site before RNA Polymerase II finishes transcribing. This releases the upstream portion of the transcript, which will serve as the initial RNA prior to further processing (the pre-mRNA in the case of protein-encoding genes.) This cleavage site is considered the "end" of the gene. The remainder of the transcript is digested by a 5′-exonuclease (called Xrn2 in humans) while it is still being transcribed by the RNA Polymerase II. When the 5′-exonuclease "catches up" to RNA Polymerase II by digesting away all the overhanging RNA, it helps disengage the polymerase from its DNA template strand, finally terminating that round of transcription.

In the case of protein-encoding genes, the cleavage site which determines the "end" of the emerging pre-mRNA occurs between an upstream AAUAAA sequence and a downstream GU-rich sequence separated by about 40-60 nucleotides in the emerging RNA. Once both of these sequences have been transcribed, a protein called CPSF in humans binds the AAUAAA sequence and a protein called CstF in humans binds the GU-rich sequence. These two proteins form the base of a complicated protein complex that forms in this region before CPSF cleaves the nascent pre-mRNA at a site 10-30 nucleotides downstream from the AAUAAA site. The Poly(A) Polymerase enzyme which catalyzes the addition of a 3′ poly-A tail on the pre-mRNA is part of the complex that forms with CPSF and CstF.

The tRNA, 5S rRNA, and structural RNAs genes transcribed by RNA Polymerase III have a not-entirely-understood termination signal. The RNAs transcribed by RNA Polymerase III have a short stretch of four to seven U’s at their 3′ end. This somehow triggers RNA Polymerase III to both release the nascent RNA and disengage from the template DNA strand.