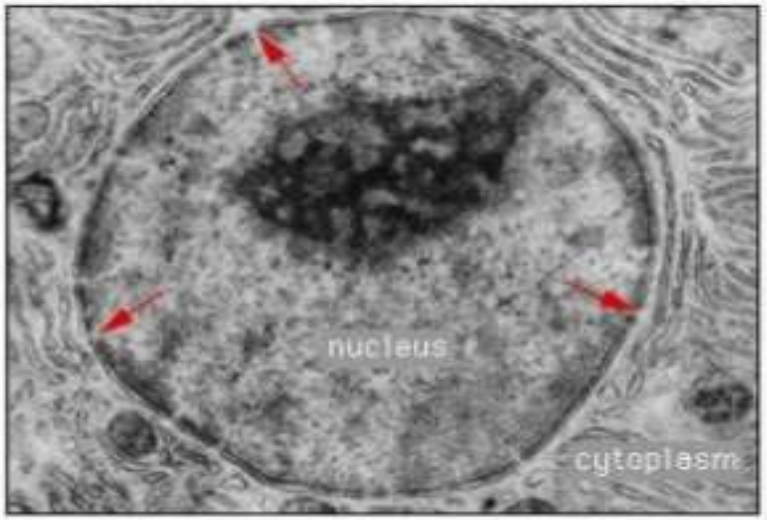
Gene Expression in Eukaryotes

Presented by: Dr. Asma

Outline

- Central dogma in Eukaryotes
- Nature of Genes in Eukaryotes
- Initiation and Elongation of Transcription
- RNA Processing

Eukaryotic Transcription



©1998 GARLAND PUBLISHING

- Transcription occurs in the nucleus in eukaryotes, nucleoid in bacteria
- Translation occurs on ribosomes in the cytoplasm
- mRNA is transported out of nucleus through the nuclear pores

Eukaryotic Central Dogma

In Eukaryotes (cells where the DNA is sequestered in a separate nucleus) the exons must be spliced (many eukaryotes genes contain no introns! Particularly true in ´lower´ organisms).

mRNA (messenger RNA) contains the assembled copy of the gene. The mRNA acts as a messenger to carry the information stored in the DNA in the nucleus to the cytoplasm where the ribosomes can make it into protein.

Eukaryotic Genome - Facts

~6 to 12% of human DNA encodes proteins (higher fraction in nematode)

~90% of human DNA is non-coding

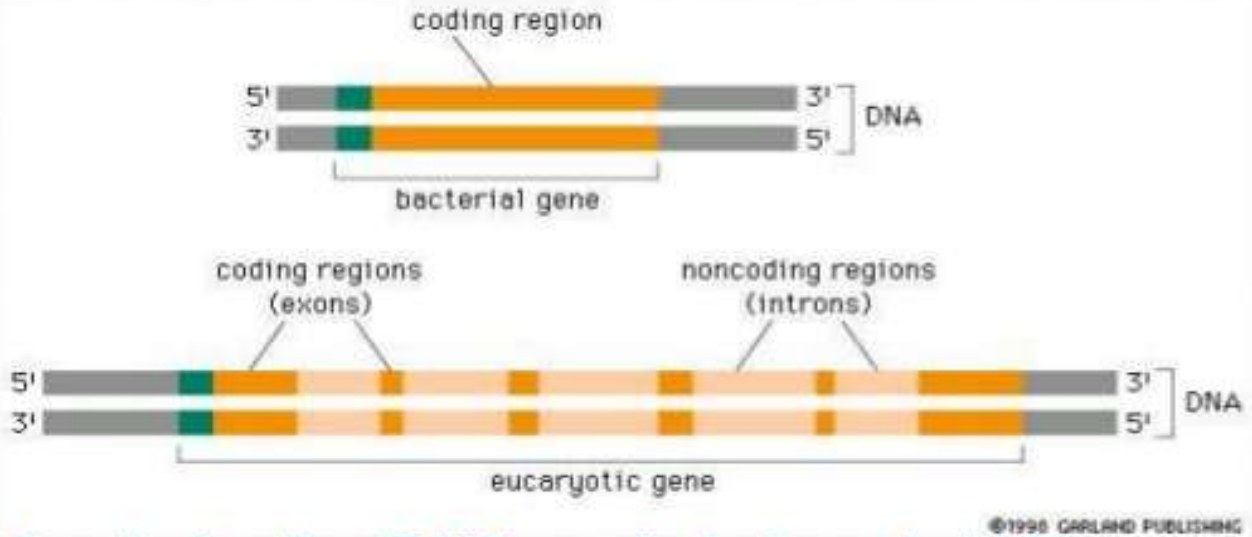
~10% of human DNA codes for UTR

Non-Coding Eukaryotic DNA

Untranslated regions (UTRs)

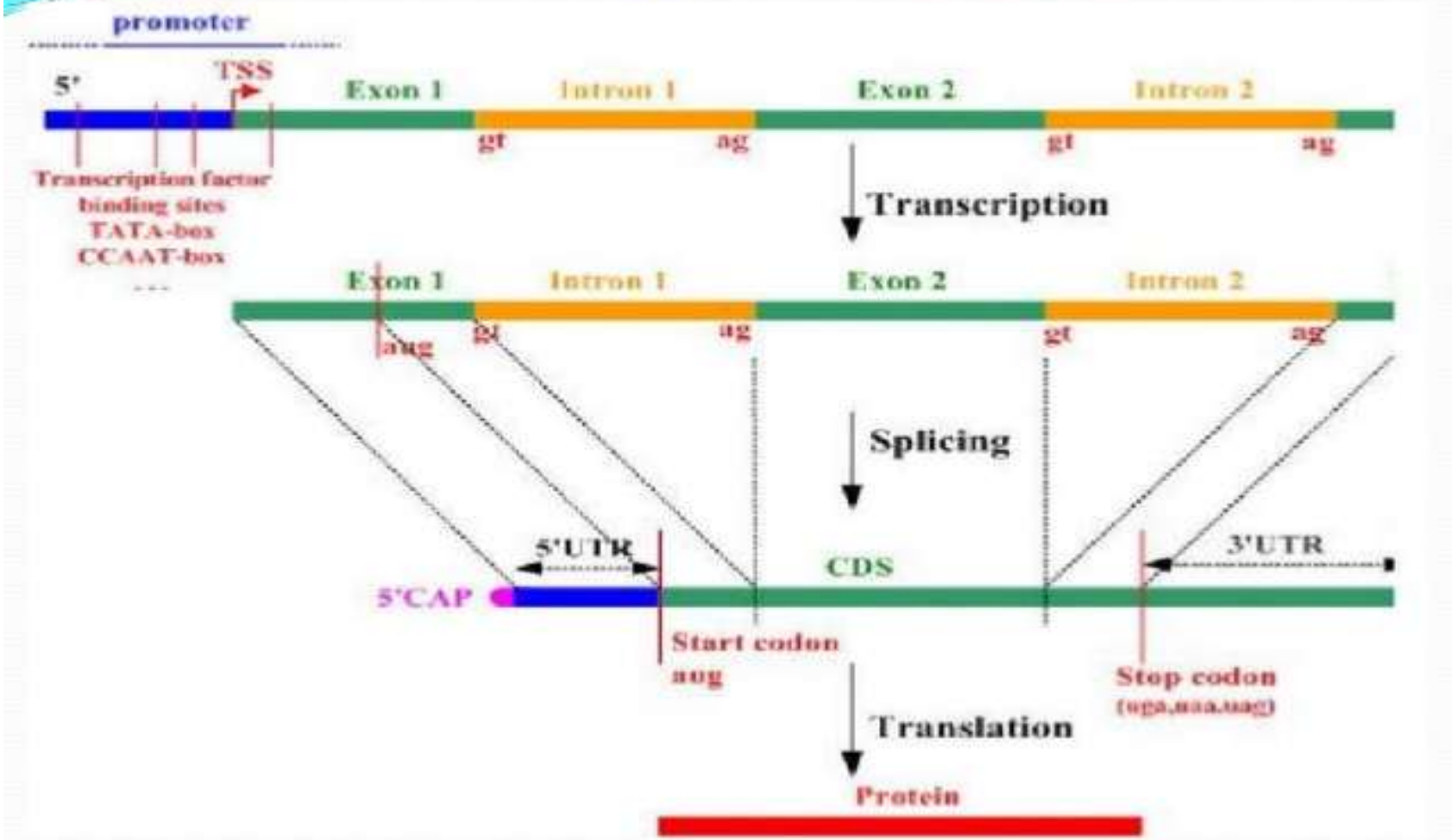
- Introns (can be genes within introns of another gene!)
- Intergenic regions:
  - Repetitive elements
  - Pseudogenes: Dead genes that may (or may not) have been retroposed back in the genome as a single-exon "gene"

Coding and Non-coding Sequences



- In bacteria, the RNA made is translated to a protein
- In eukaryotic cells, the primary transcript is made of coding sequences called exons and non-coding sequences called introns
- It is the exons that make up the mRNA that gets translated to a protein

Eukaryotic Gene



Eukaryotic Nuclear Genes

Genes transcribed by RNA Pol II

- Upstream Enhancer elements.
- Upstream Promoter elements.
- GC box (-90 nt) (20 bp), CAAT box (-75 nt) (22 bp)
- TATA promoter (-30 nt - 70%, 15 nt consensus (Bucher *et al.*, 1990)
- Transcription initiation.
- Transcript region, interrupted by introns. Translation Initiation (Kozak signal – 12 bp consensus: 6 bp prior to initiation codon)
- polyA signal (AATAAA, 99%)

Introns

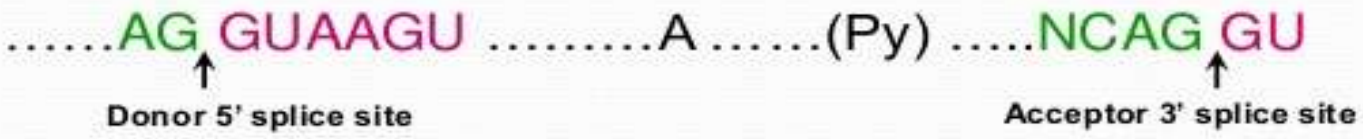
- Transcript region is interrupted by introns

Each intron (on DNA):

✓ starts with a donor site consensus ($G_{100}T_{100}A_{62}A_{68}G_{84}T_{63}$..) – GU on RNA

✓ has a branch site near 3' end of intron (one not very conserved consensus UACUAAC)

✓ ends with an acceptor site consensus. (12Py..$NC_{65}A_{100}G_{100}$)

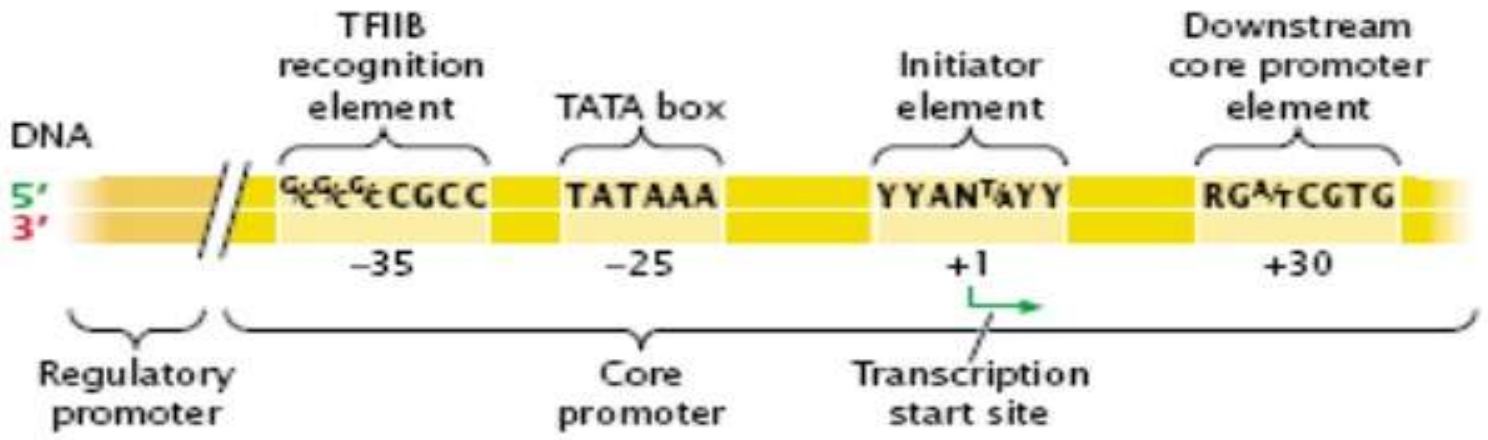


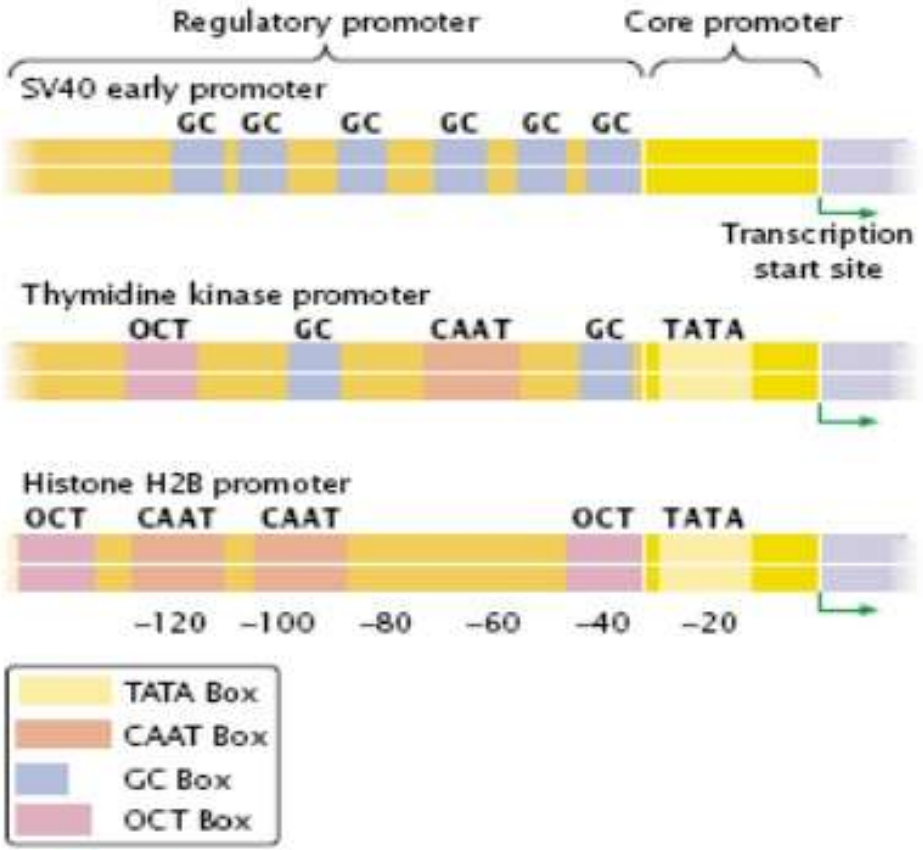
Exons

- The exons of the transcript region are composed of:

- ✓ 5' UTR with a mean length of 769 bp
- ✓ AUG (or other start codon)
- ✓ Remainder of coding region
- ✓ Stop Codon
- ✓ 3' UTR with a mean length of 457 bp

Eukaryotic Promoter

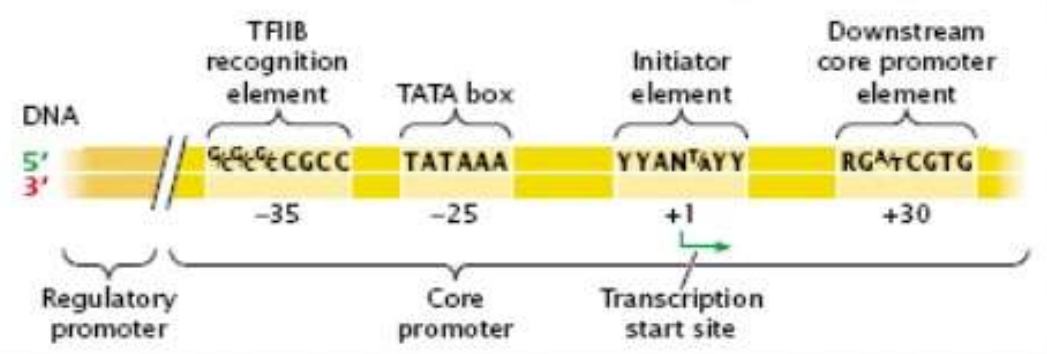




13.19 The consensus sequences in promoters of three eukaryotic genes illustrate the principle that different sequences can be mixed and matched to yield a functional promoter.

Polymerases also use transcription factors
Bind in a specified order, either to promoter or each other

RNA polymerase II must be phosphorylated before it can start synthesizing RNA



Sequences of Eukaryotic promoter

Eukaryotic RNA polymerases

RNA polymerase I- makes precursors for ribosomal RNAs (except for smallest subunit)

RNA polymerase II- mRNA and snRNAs (involved in RNA processing)

RNA polymerase III- variety of RNAs: smallest rRNA subunit, tRNA precursors

Each uses a different promoter (DNA sequences that direct polymerase to begin transcribing there)

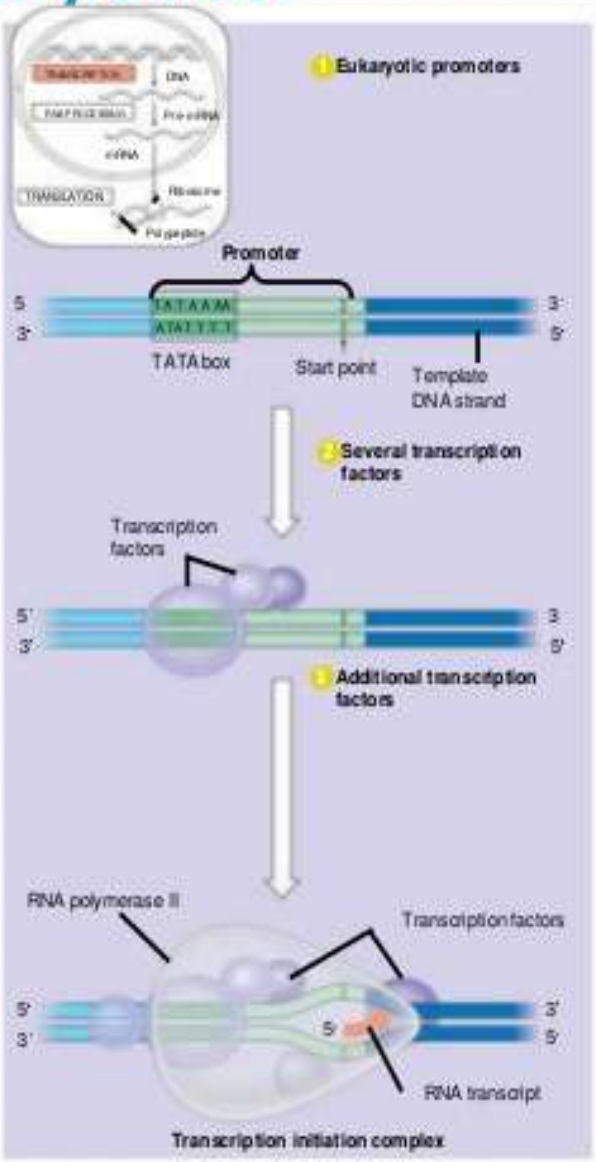
Promoters are "upstream" from coding sequence

Initiation in Eukaryotes

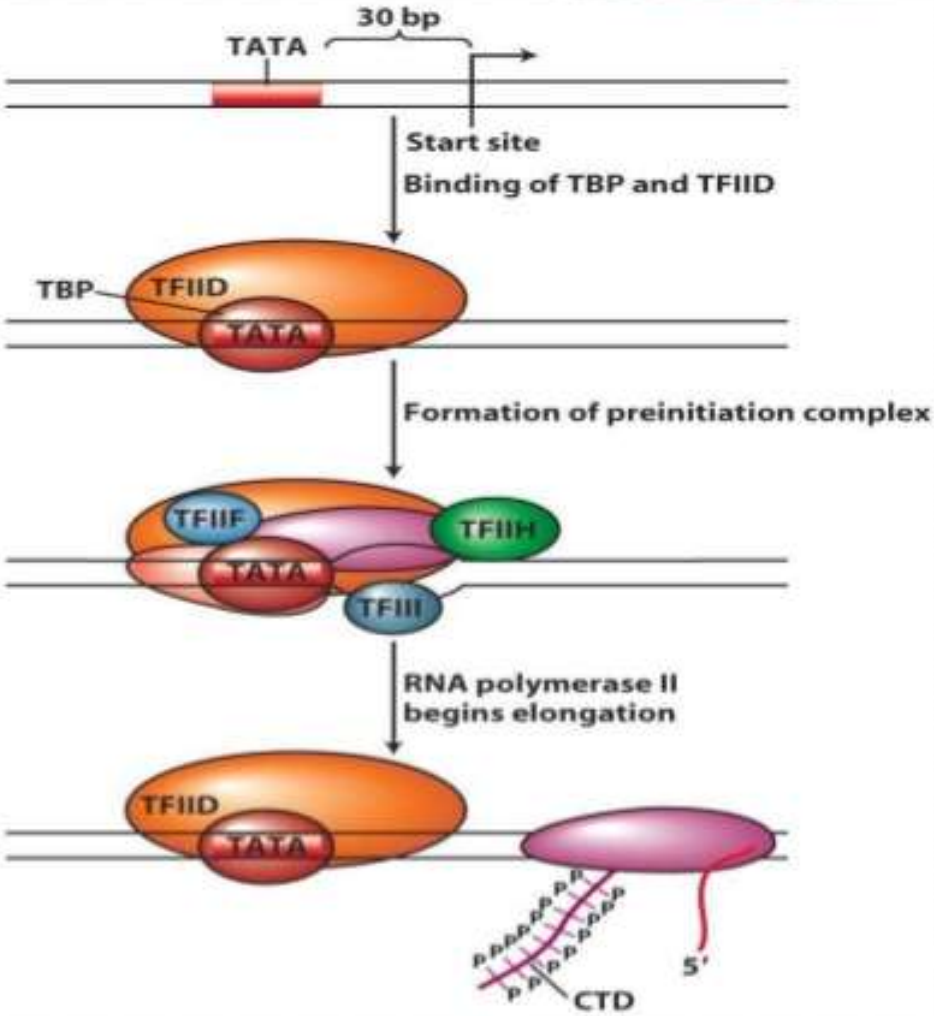
TATA box

Several transcription factors must bind to promoter sequences upstream of the gene

Then RNA polymerase can bind



Requirements for initiation of Transcription



Transcription Factors – Order of their binding

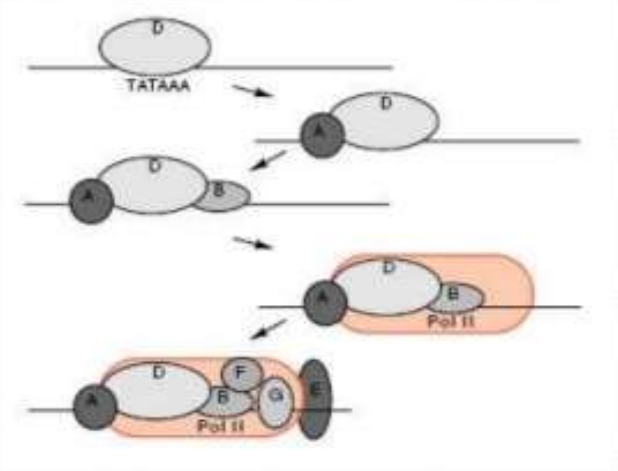
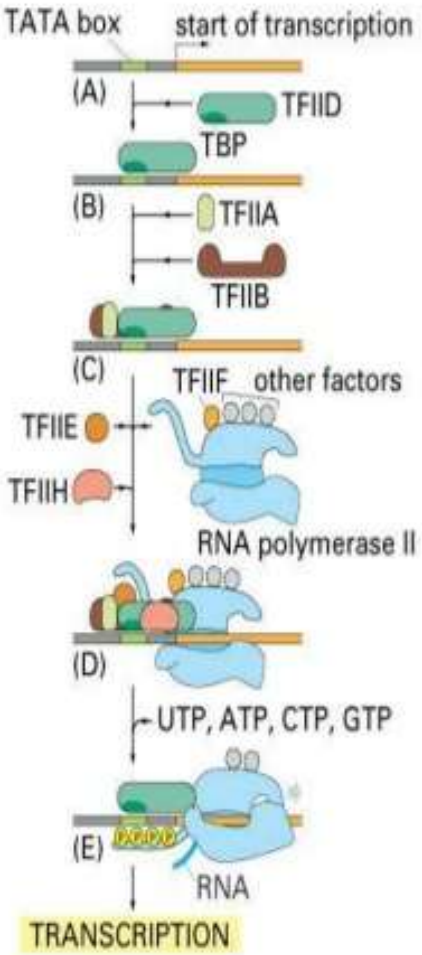


TABLE 26–1 Proteins Required for Initiation of Transcription at the RNA Polymerase II (Pol II) Promoters of Eukaryotes

Transcription protein	*Number of subunits*	*Subunit(s) M_r*	*Function(s)*
Initiation			
Pol II	12	10,000–220,000	Catalyzes RNA synthesis
TBP (TATA-binding protein)	1	38,000	Specifically recognizes the TATA box
TFIIA	3	12,000, 19,000, 35,000	Stabilizes binding of TFIIB and TBP to the promoter
TFIIB	1	35,000	Binds to TBP; recruits Pol II–TFIIF complex
TFIIE	2	34,000, 57,000	Recruits TFIIH; has ATPase and helicase activities
TFIIF	2	30,000, 74,000	Binds tightly to Pol II; binds to TFIIB and prevents binding of Pol II to nonspecific DNA sequences
TFIIH	12	35,000–89,000	Unwinds DNA at promoter (helicase activity); phosphorylates Pol II (within the CTD); recruits nucleotide-excision repair proteins
Elongation*			
ELL†	1	80,000	
p-TEFb	2	43,000, 124,000	Phosphorylates Pol II (within the CTD)
SII (TFIIS)	1	38,000	
Elongin (SIII)	3	15,000, 18,000, 110,000	

Eukaryotic Transcription initiation

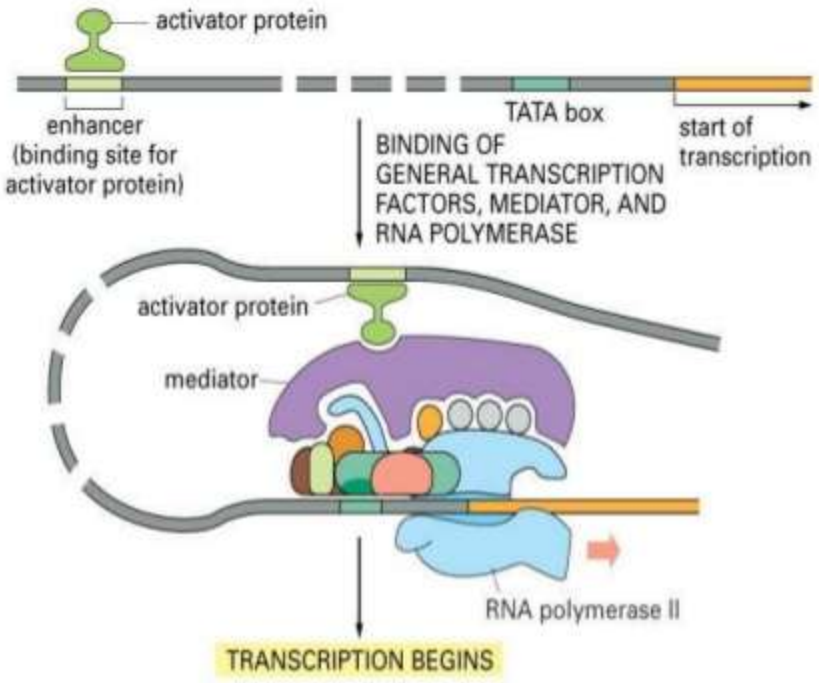
General transcription factors



Fig

TATA binding protein (TBP)/TFIID binds to TATA box (-25)

Role of Enhancers in Initiation



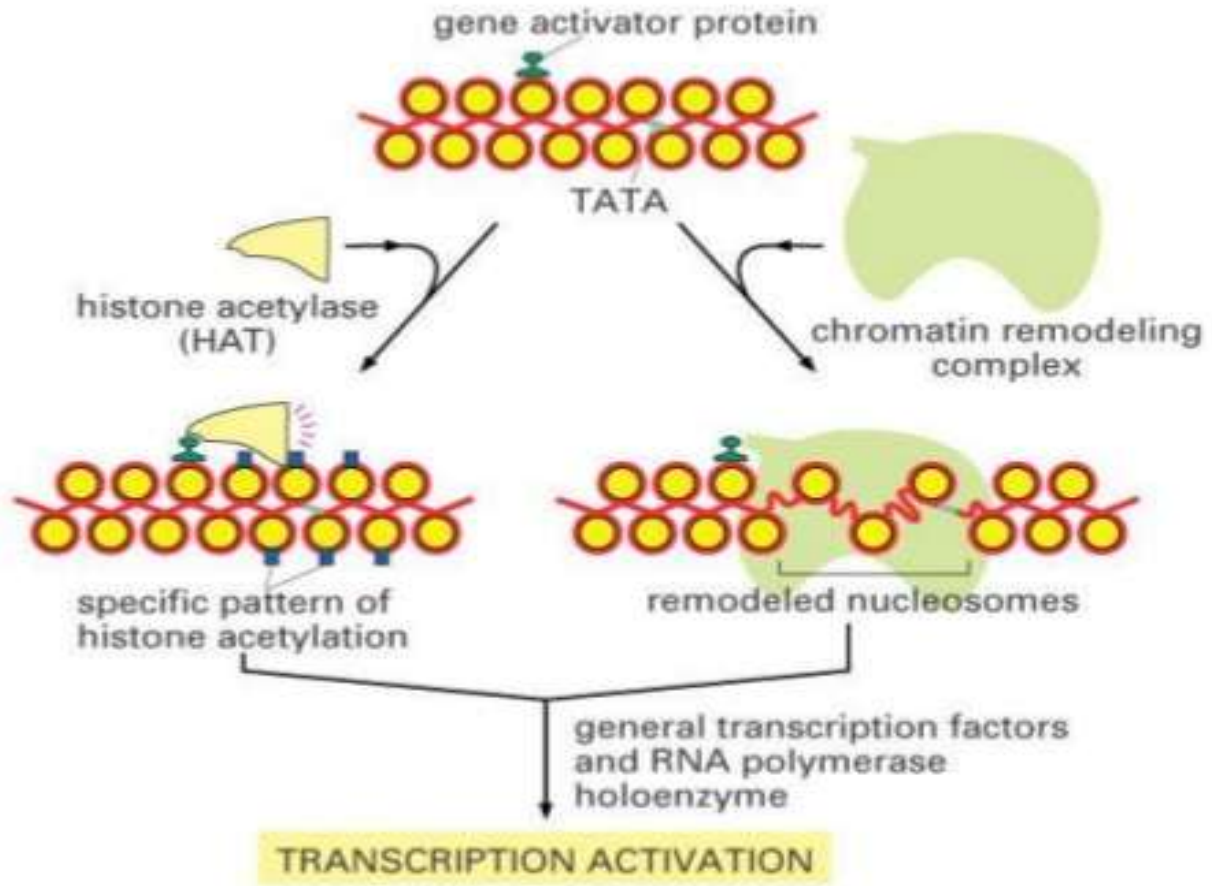
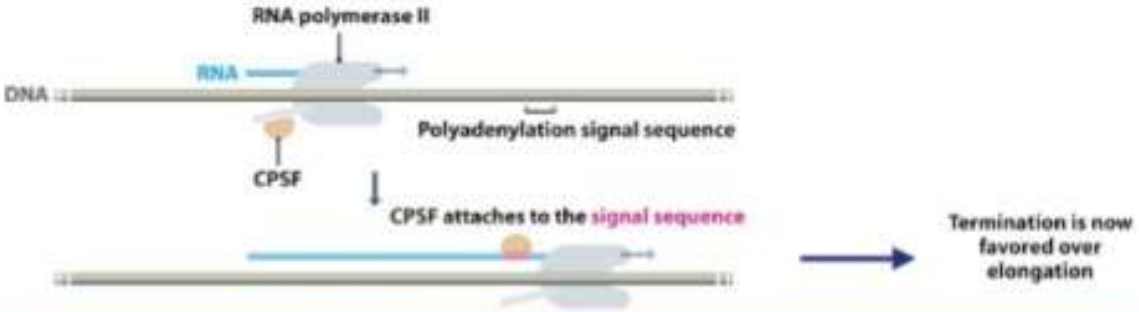

gene activator protein
TATA
histone acetylase
(HAT)
chromatin remodeling
complex
specific pattern of
histone acetylation
remodeled nucleosomes
general transcription factors
and RNA polymerase
holoenzyme
TRANSCRIPTION ACTIVATION

Fig.

Transcription termination

- At the end of transcription the mRNA is cleaved at the poly(A) site by an endonuclease and a poly(A) tail is added to the exposed 3' end
- The release of RNAPII from the DNA template occurs at diffuse positions hundreds of bases downstream of the poly(A) signals and termination requires a functional poly(A) site, but not cleavage of the RNA
- The RNAPII CTD provides a platform for holding RNA binding proteins that recognize the poly(A) signal (CPSF, CstF)
- It appears that these and other 3' processing factors contact elongating RNAPII throughout the entire elongation phase; in fact, CPSF can be recruited to promoters by TFIID and subsequently transferred to RNAPII



Overall Transcription Process