BIOC4004 - Industrial Biochemistry
Lecture 01 - Wed Jan 07, 2004

Topics for the Day:

- About genes…
- Prokaryotic Gene Structure
- Eukaryotic Genes Structure
- Gene Regulation:
    - Promoters
        - Prokaryotic
        - Eukaryotic
    - Transcription
        - Prokaryotic
        - Eukaryotic

Genes

Every gene contains all of the information necessary to produce a "product". The product is the RNA molecule transcribed from the DNA template. The RNA itself is sometimes the end-product (as is the case with structural RNAs - tRNAs, rRNAs, ribozymes). Most of the time the end-product is the protein obtained when the RNA is translated into a polypeptide

The Gene as an engineering problem:

1) Needs to contain necessary "blueprint" for "product"

    ie. the coding sequence

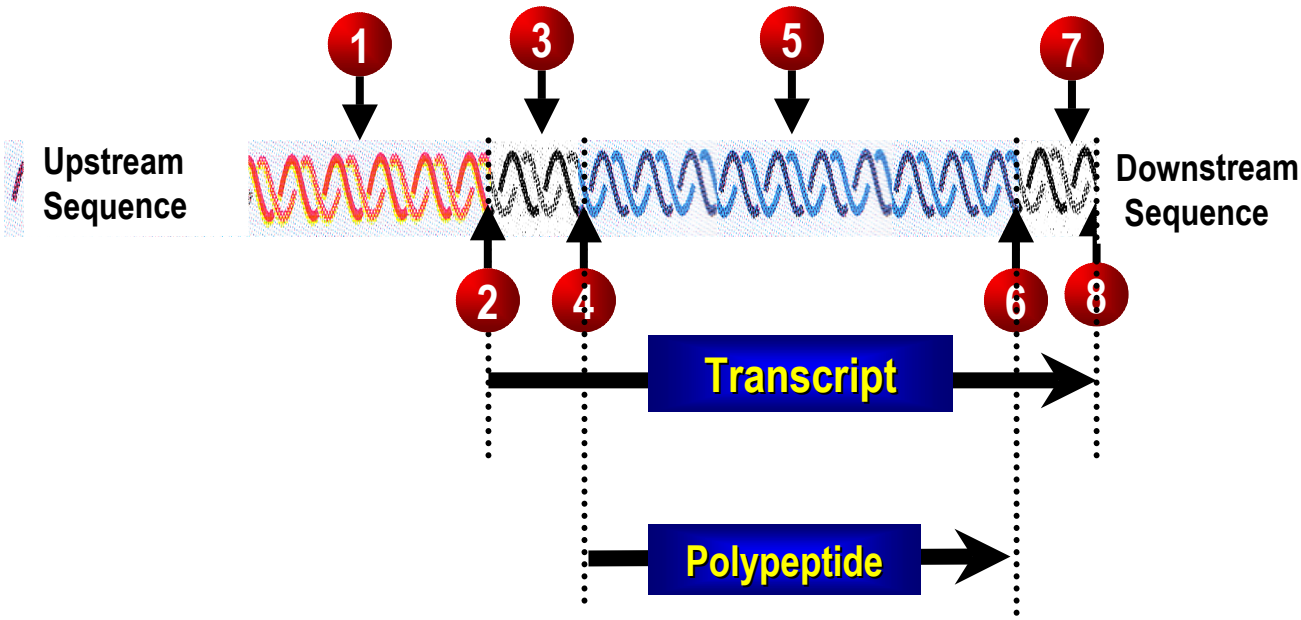
2) Needs to contain information for the mechanics of production including "when" and "how much" to produce

    ie. the promoter

3) Information for turning blueprint into product

    ie. info for Translational Machinery (5' and 3' UTRs)

Prokaryotic Genes:



1) Promoter region (-10 box, -35 box)

2) Transcription initiation site (at +1)

3) 5' UTR

4) Translation initiation (AUG, UUG)

5) Coding region

6) Translation stop site (UGA, UAA, UAG)

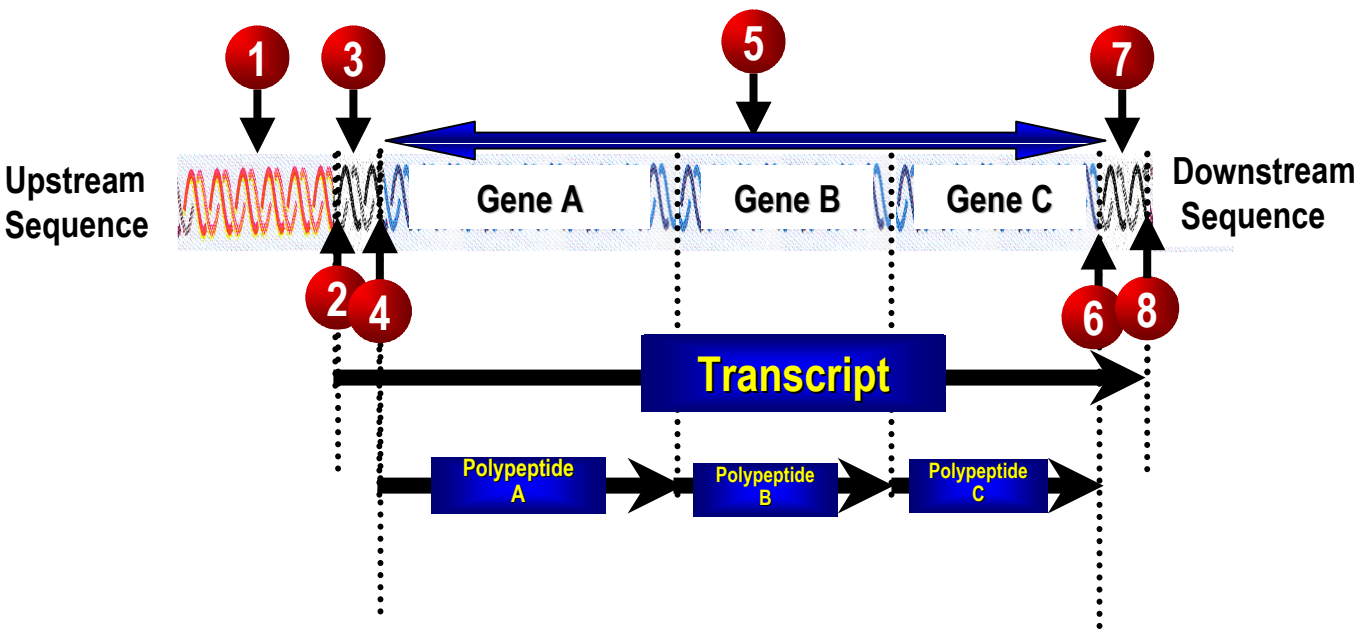
7) 3' UTR

8) Transcription stop site

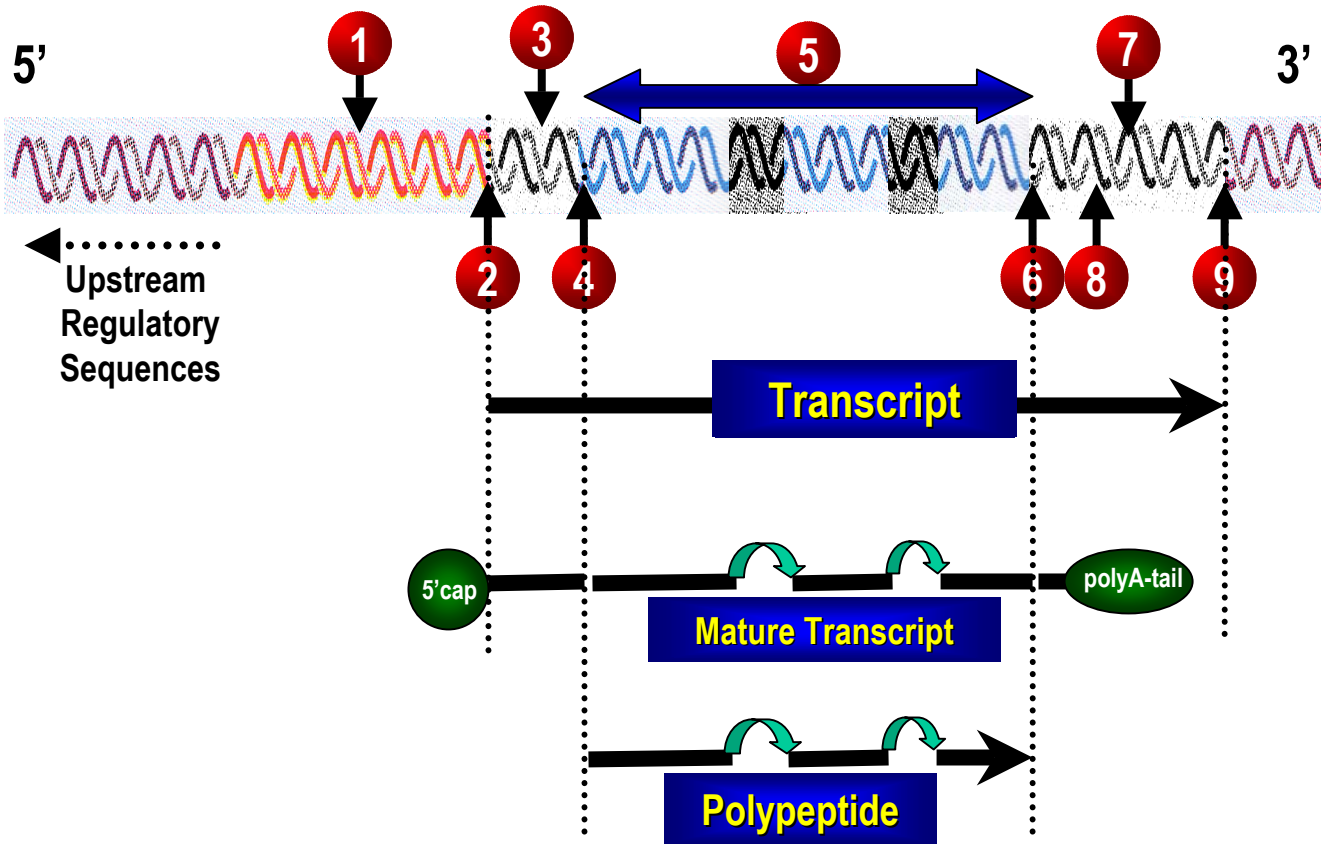
Prokaryotic Genes (cont')

- Often arranged in OPERONS of related function
  - show co-regulation

- Prokaryotic mRNAs are often poly-cistronic
  - A single promoter drives expression of several genes in tandem
  - More than one different protein coded in same mRNA
  - Some are translated as a single polypeptide that is subsequently cleaved
  - Some have "spacers"between successive genes
  - Genes can be partially overlapping and in different frame



Eukaryotic Genes:



1) Promoter (CAAT box, GC box, TATA box)

2) transcription initiation site (at +1)

3) 5' UTR

4) Start Codon (AUG)

5) alternating exons/introns (splice donor/acceptor sites)

6) Terminator Codons (UGA, UAA, UAG)

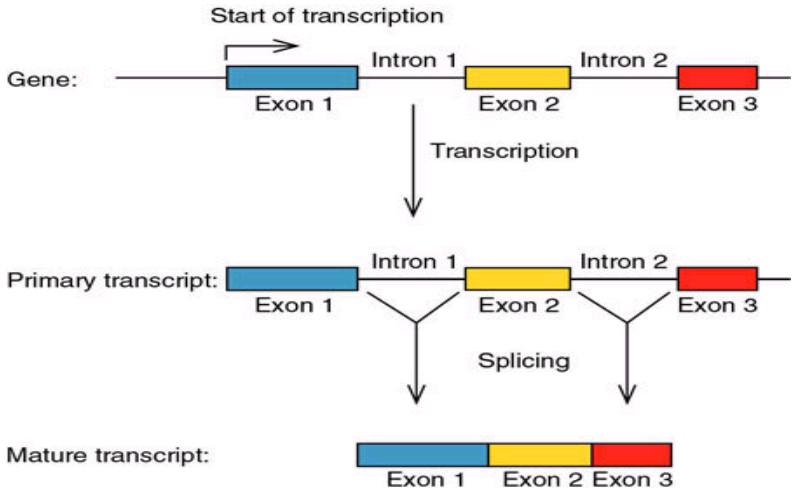
7) 3' UTR

8) polyadenylation signal (AAUAAA)

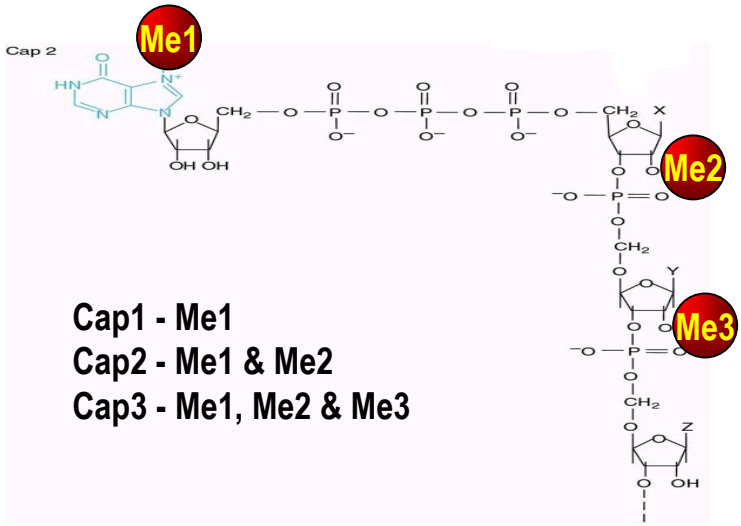
9) transcription termination site

Eukaryotic Genes (cont):

Splicing

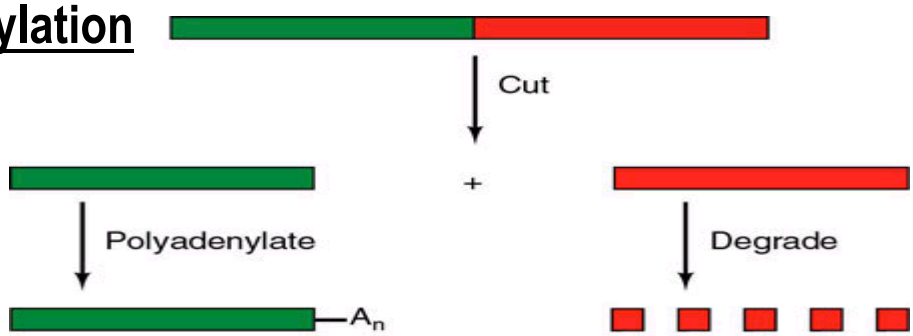


Capping



- Cap affects recognition by Translation machinery (300 fold)

Poly-Adenylation



- poly-A affectsTranslation (~ 20 fold) & RNA stability (somewhat)
- not **required**

Gene Regulation

Gene regulation involves controlling the level of a gene's expression by regulating its transcriptional rate.

Promoters :

- promoters "drive" gene expression

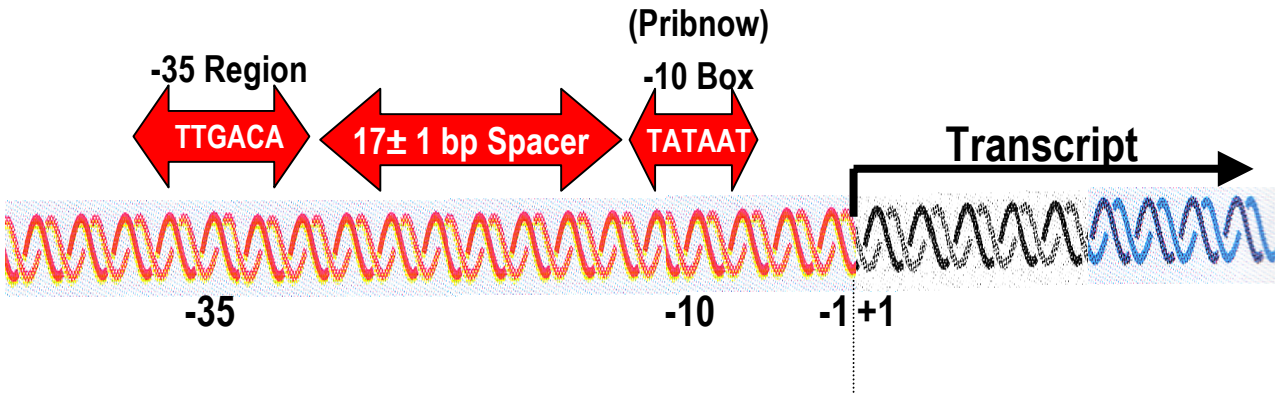
- serve as "assembly points" for the transcriptional machinery
    - binding sites for the RNA polymerase

- contain info for regulation purposes (when and how much)
    - binding sites for regulatory proteins (**transcription factors**)
    - TFs can be positive or negative regulators

    - **Activators**
        - bind to DNA at or near promoters
        - activator binding site is called **Activator Site**
        - help recruit RNA polymerase to the promoter

    - **Repressors**
        - bind to DNA at or near a promoter
        - repressor binding site is called an **Operator**
        - physically block transcriptional machinery
        - Repressors have **effectors** that regulate them
            - **Inducers** bind to repressor, inhibiting ability to bind DNA
            - **Corepressors** bind to repressor, enabling it to bind DNA

Generalized view of Prokaryotic Promoters



Harley & Reynolds (1987) out of McMaster University:

- analysis of 263 **phage**, **plasmid** and **bacterial** promoters
- all bases in the -35 (TTGACA) and -10 (TATAAT) hexamers were highly conserved
- 92% of promoters had inter-region spacing or 17±1 bp
- 75% of the uniquely defined start points initiated 7±1 bp downstream of the -10 box

Lisser and Margalit (1993): looked at *E.coli* promoters only (no viral promoters)

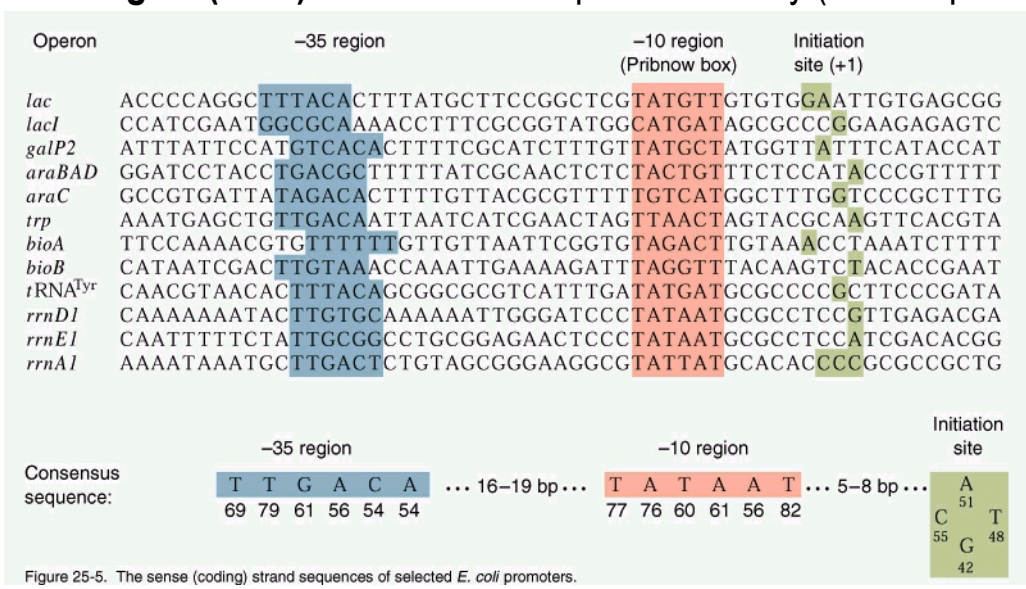


Figure 25-5. The sense (coding) strand sequences of selected *E. coli* promoters.

Generalized view of Prokaryotic Promoters (cont)

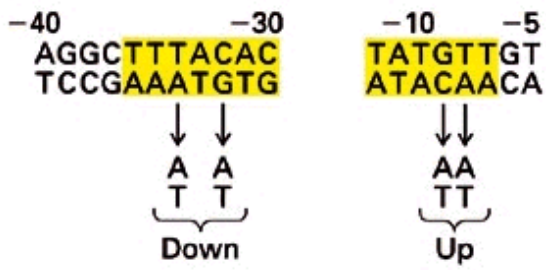
- the -10 and -35 motifs are only “consensus sequences”; some variation allowed
- this variation can impact the “efficiency” of the promoter
  - this is good since not all genes need to be transcribed to the same degree
  - the more similar to the consensus, the stronger the promoter

(a) Strong *E.coli* promoters

tyr tRNA	TCTCAACGTAACACTTTACAGCGGCG··CGTCATTTGATATGATGC·GCCCCGCTTCCCGATAAGGG
rrn D1	GATCAAAAAAATACTTGTGCAAAAAA··TTGGGATCCCTATAATGCGCCTCCGTTGAGACGACAACG
rrn X1	ATGCATTTTTCCGCTTGTCTTCCTGA··GCCGACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
rrn $(DXE)_2$	CCTGAAATTCAGGGTTGACTCTGAAA··GAGGAAAGCGTAATATAC·GCCACCTCGCGACAGTGAGC
rrn E1	CTGCAATTTTTCTATTGCGGCCTGCG··GAGAACTCCCTATAATGCGCCTCCATCGACACGGCGGAT
rrn A1	TTTTAAATTTCCTCTTGTCAGGCCGG··AATAACTCCCTATAATGCGCCACCACTGACACGGAACAA
rrn A2	GCAAAAATAAATGCTTGACTCTGTAG··CGGGAAGGCGTATTATGC·ACACCCCGCGCCGCTGAGAA
λ P_R	TAACACCGTGCGTGTTGACTATTTTA·CCTCTGGCGGTGATAATGG··TTGCATGTACTAAGGAGGT
λ P_L	TATCTCTGGCGGTGTTGACATAAATA·CCACTGGCGGTGATACTGA··GCACATCAGCAGGACGCAC
T7 A3	GTGAAACAAAACGGTTGACAACATGA·AGTAAACACGGTACGATGT·ACCACATGAAACGACAGTGA
T7 A1	TATCAAAAAGAGTATTGACTTAAAGT·CTAACCTATAGGATACTTA·CAGCCATCGAGAGGGACACG
T7 A2	ACGAAAAACAGGTATTGACAACATGAAGTAACATGCAGTAAGATAC·AAATCGCTAGGTAACACTAG
fd VIII	GATACAAATCTCCGTTGTACTTTGTT··TCGCGCTTGGTATAATCG·CTGGGGGTCAAAGATGAGTG

−35 −10 +1

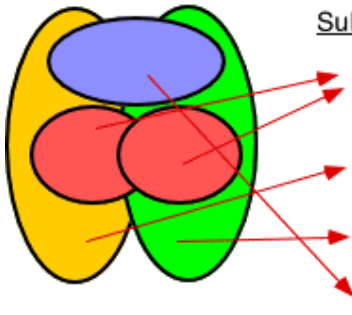
- mutations that affect this similarity affect the promoter strength



- some strong promoters contain the UP motif (at -57 to -47), which enhances binding by RNA polymerase
  - alpha subunit doesn’t normally bind DNA
  - C-terminus binds to the UP motif

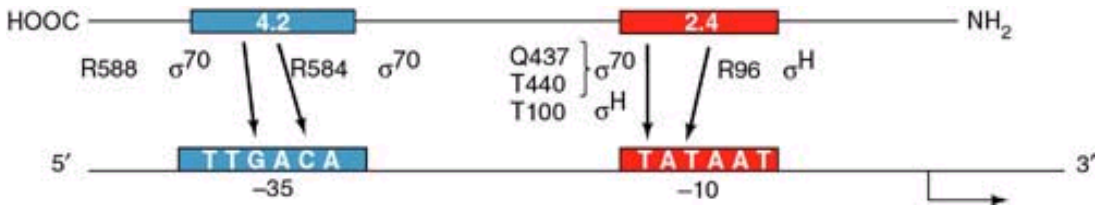
Generalized view of Prokaryotic Promoters (cont)

Prokaryotic RNA Pol holoenzyme



Subunit	Size	#/Molecule	Function
α	36.5 kD	2	chain initiation and interaction with regulatory proteins
β	151 kD	1	chain initiation and elongation
β'	155 kD	1	DNA binding
σ	70 kD	1	promoter recognition

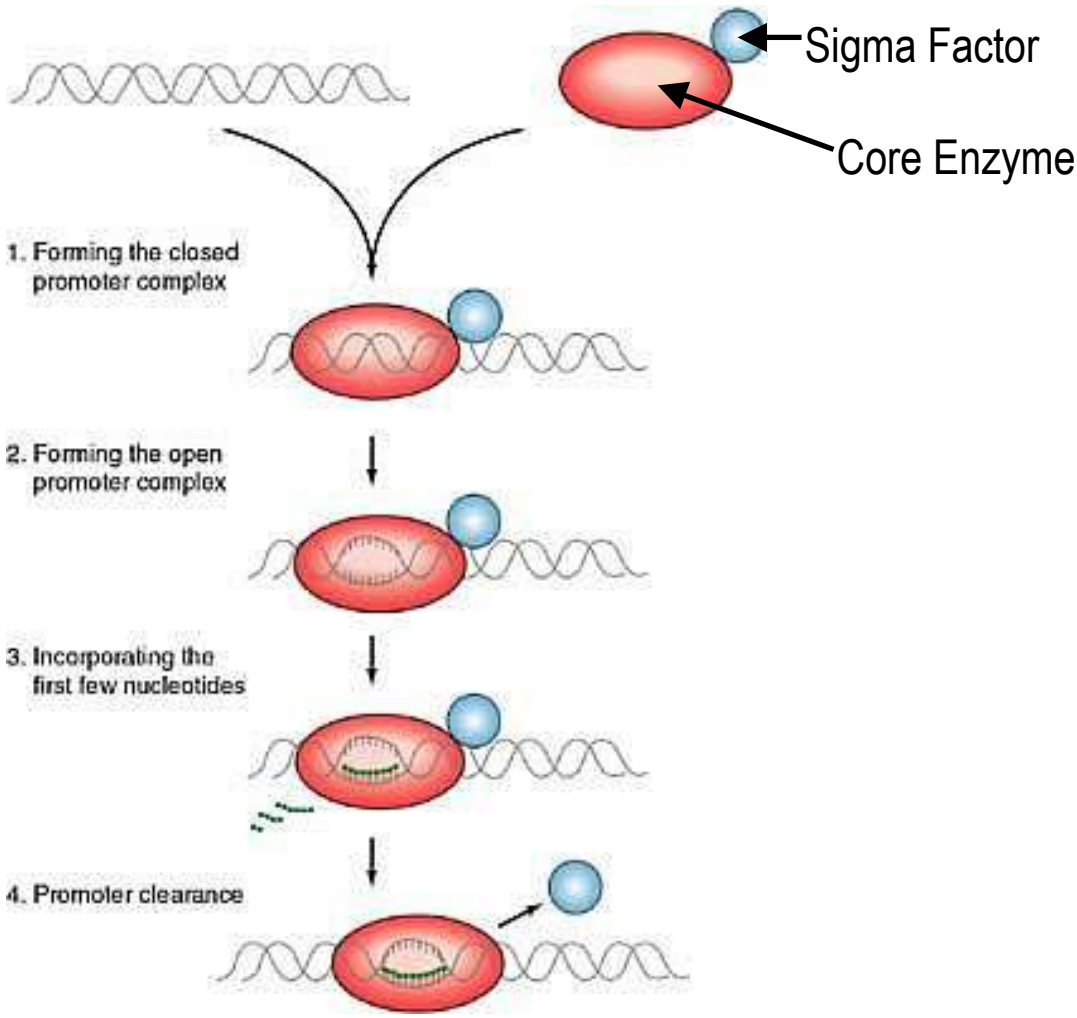
- Without sigma factor, RNA Pol has **no specificity**
- Sigma factors act as specificity factors for bacterial RNA polymerases
  - different sigma factors recruit RNA polymerase to "different types of genes" (nitrogen, flagellar, heat-shock, stationary phase…)
  - different types of genes will have slight variations in their promoter sequences (making them better at attracting a given sigma factor)
- **Contact points between sigma and a typical prokaryotic promoter (–10 and –35):**



- factors exist that bind to, and inactivate, sigma-factors

Generalized view of Prokaryotic Promoters (cont)

Prokaryotic Trancription

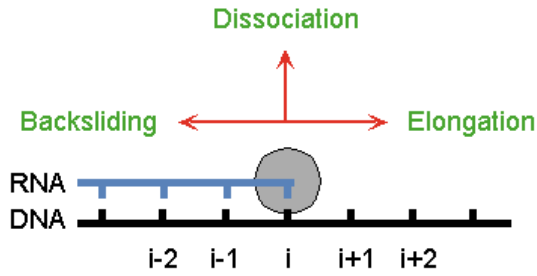


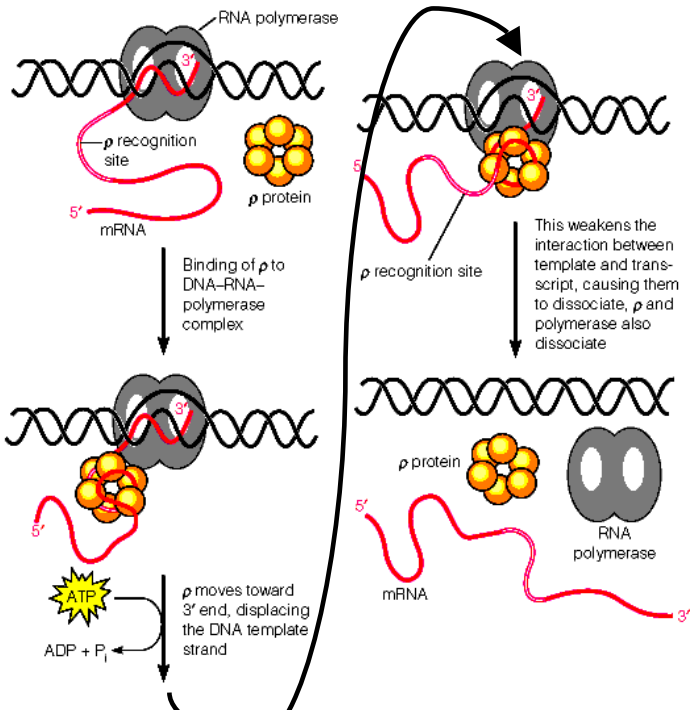
Promoter Clearance:

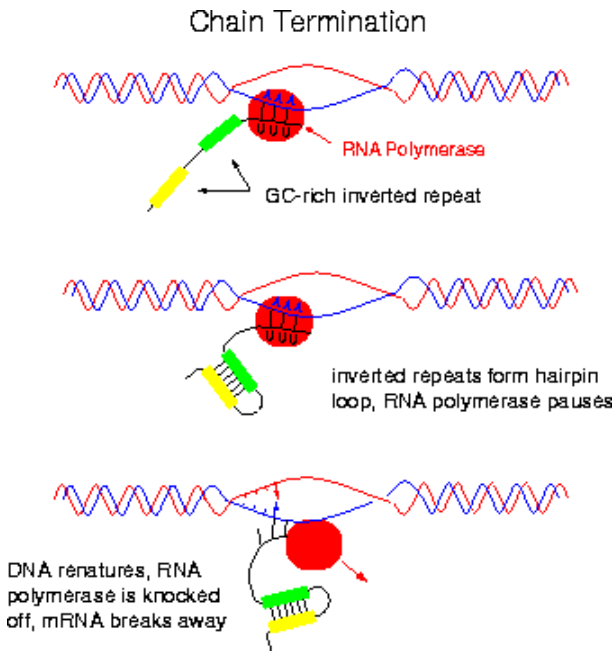
- initially need to bind tightly to promoter
- need to "let go" in order to proceed with transcript extension
    - Sigma factor "needs to come-off"

Prokaryotic Transcription Termination:

What to do ?







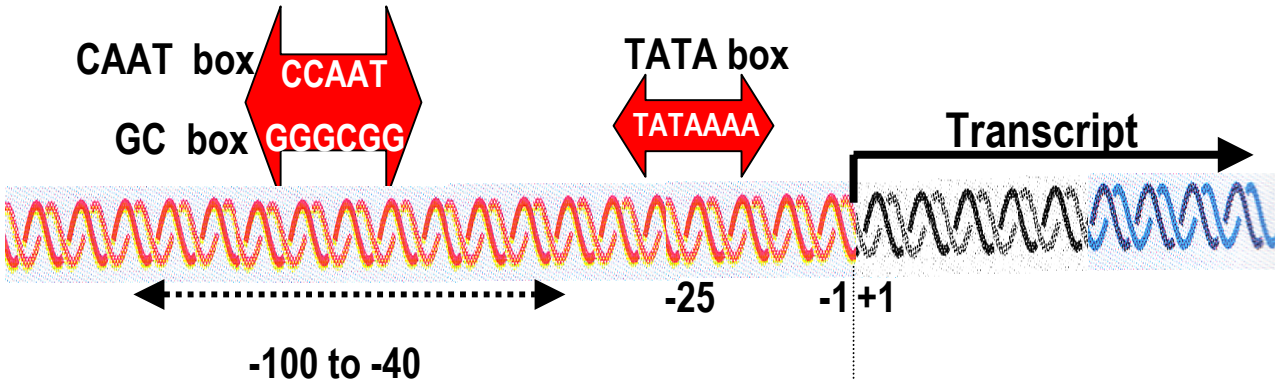
Rho-dependent

- Rho-binding site in the transcript
- A “pause-site” for the transcription complex
- Rho moves towards growing end of transcript, reaches paused Transcription Complex, promotes dissociation

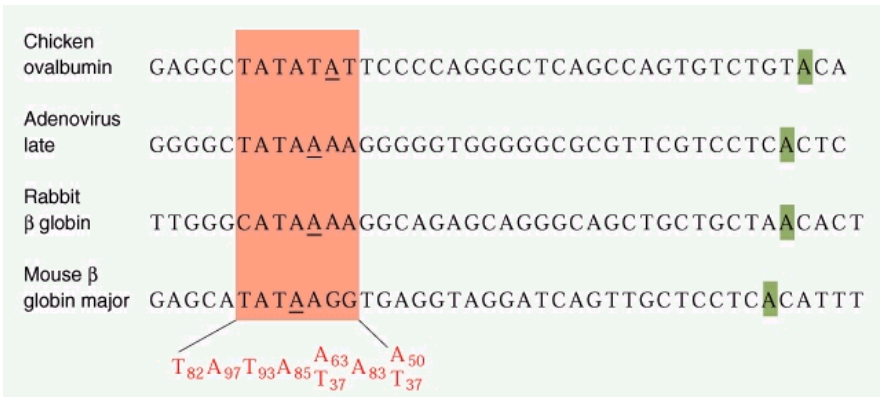
Rho-independent

- Two short GC-rich sequences which are inverted repeats in the DNA sequence.
- A stretch of 6 to 8 adenine nucleotides in the template strand of the DNA
- Transcription Complex stalls at hairpin, transcript melts out, DNA renatures.

Generalized view of Eukaryotic Promoters

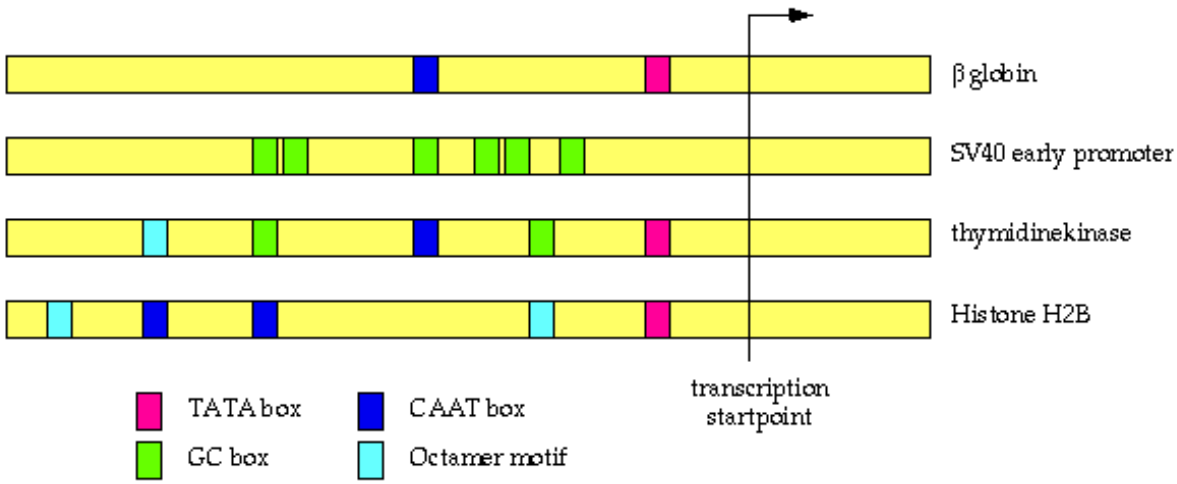


Gannon et al. (1978) looked at transcription start sites...



note: Some genes, including "housekeeping genes" do not have TATA boxes

:



- Promoters contain combinations of "upstream regulatory motifs"

Generalized view of Eukaryotic Promoters (cont)

- Eukaryotes have three types of RNA polymerases

- Protein-encoding genes are transcribed by **RNA pol II**

- Eukaryotic RNA polymerases are assembled from numerous subunits
    - as many as 14 different subunits come together

- At least six **general (or basal) transcription factors** have been characterized
    - TFIIA, TFIIB, TFIID, TFIIE, TFIIF, TFIIH
    - Allow “basal” transcription initiation

- At least 8 - 10 TAF_{II}'s (TBP-associated proteins) associated with TFIID
    - the mixture of TBP-associated proteins is “highly complex”
    - Polymerases assembled at different genes differ in composition!!!