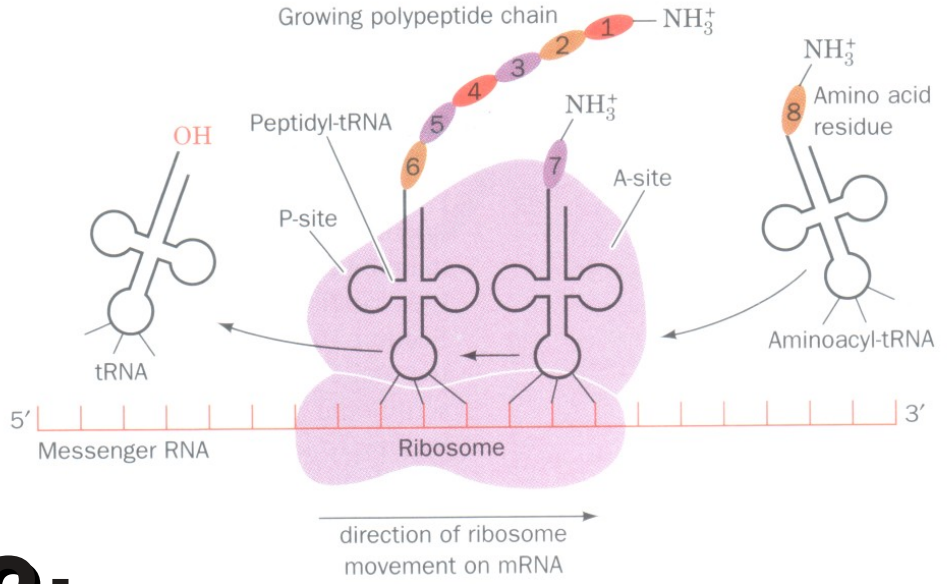
University of Lethbridge



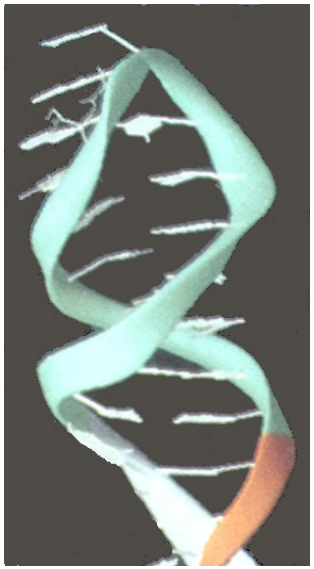
Chapters 31-32: Ribonucleic Acid (RNA)

Short segments from the transcription, processing and translation sections of each chapter

University of Lethbridge

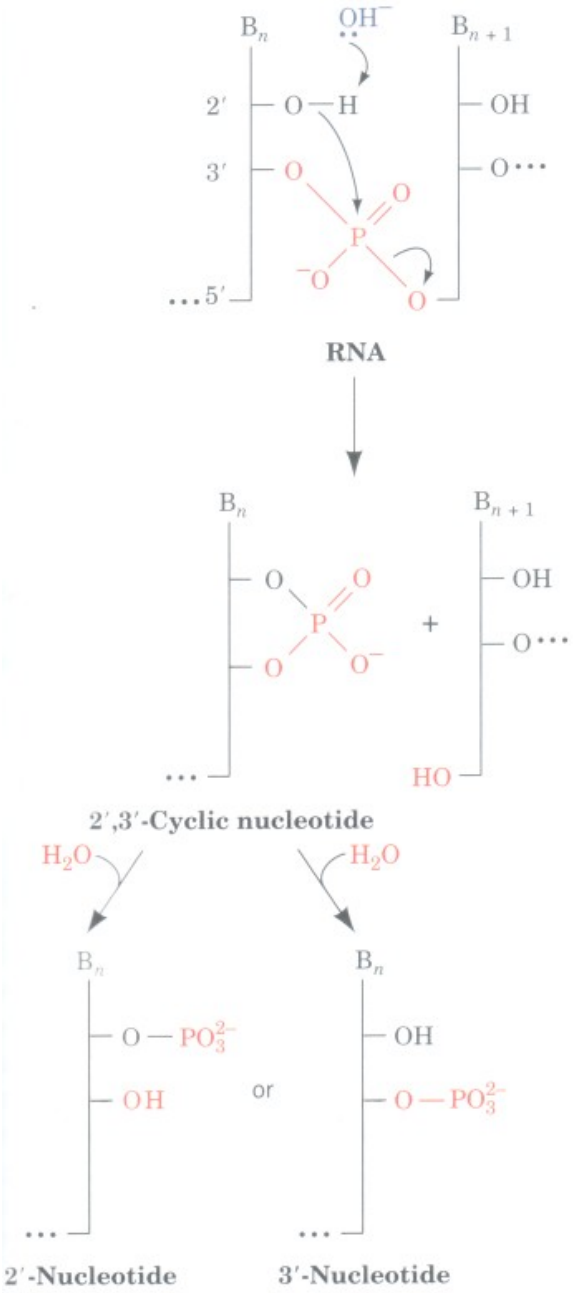
RNA

In comparison with DNA



- RNA utilizes uracil in place of thymine
- RNA species are labile
  - undergo both acid & base catalyzed hydrolysis
  - 2'-OH is central to (all) hydrolysis pathway
- RNA is single stranded
  - forms intramolecular double stranded regions or "stem-loop" structures
  - A-DNA like conformation (RNA-11)
- Catalytic RNAs (ribozymes) are found in all organisms.

Note: DNA can be catalytic but no examples of catalytic DNA have been observed in nature.



University of Lethbridge

Types of RNA

messenger RNA (mRNA)

short-lived, complement of DNA 'coding' strand; directs translation of proteins

ribosomal RNA (rRNA)

most abundant RNA (~90 %); major component of ribosome (protein synthesis)

transfer RNA (tRNA)

adapter molecule; complementary to mRNA codons; numerous modified nucleotides; delivers amino acids to ribosomes

other RNA

involved in maturation of primary RNA transcripts (splicing, editing, post-transcriptional modification, *etc.*)

components of Ribonucleoproteins

University of Lethbridge

mRNA (transcript)

- **Produced by DNA-dependent RNA polymerase**

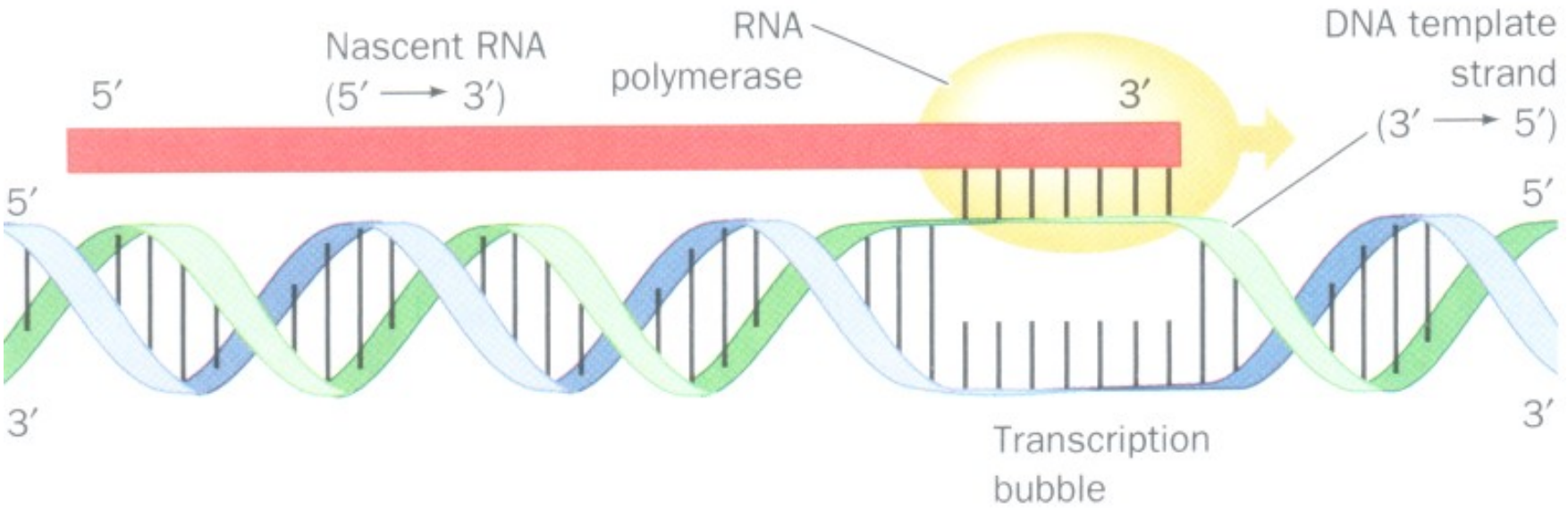
  Synthesized in a 5' -> 3' direction

  Complement of DNA template strand

- **mRNA transcripts includes the gene, upstream (*before the coding region*) and downstream elements (*following the coding region*)**

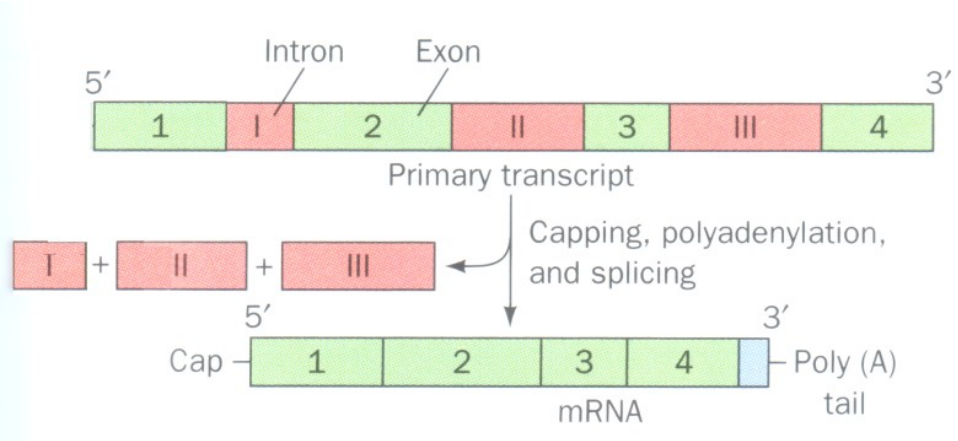
  upstream (*eg.* ribosome binding site)

  downstream (*eg.* termination elements)



(Eucaryotic) mRNA is processed

- **Eucaryotic mRNA is extensively processed before it is exported from the nucleus and is translated by ribosomes**
  - **often contains introns (intervening sequences) which must be removed**
- **Bacterial mRNA is very short lived (1-3 minutes typically) and is often being translated before transcription is completed**
  - **bacterial mRNA processing is uncommon but does occur**



Typical processing of (eucaryotic) primary transcript to mature mRNA

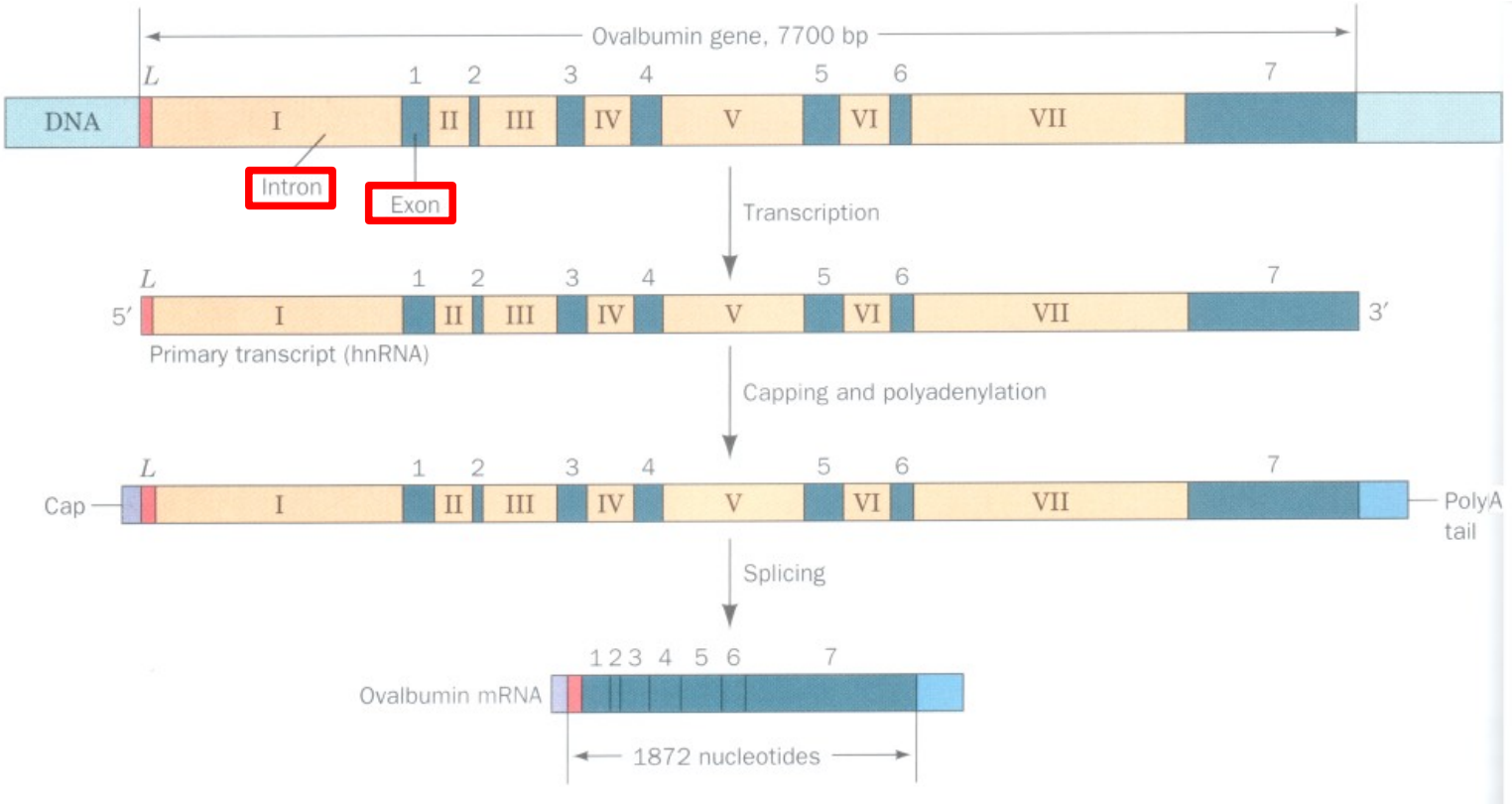
University of Lethbridge

Eucaryotic mRNA processing

Typically involves two steps

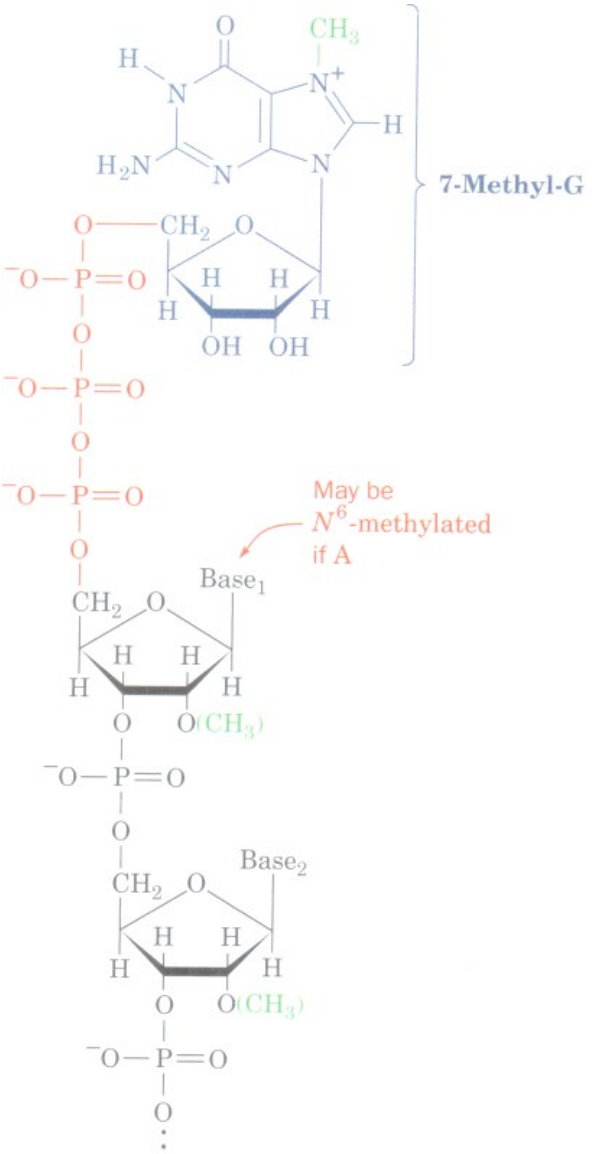
(1) 5'-capping and 3'-polyadenylation of ends

(2) Intron removal or splicing



University of Lethbridge

5' – Capping (eucaryotes)



7-Methyl-guanosine is added to 5' end of transcripts

- **unusual 5',5' linkage with three phosphate groups**

First two nucleotides of transcript may be methylated at 2'-OH

If the first nucleotide of the transcript is A, it may be methylated at N6

Capping is important for mRNA recognition by the ribosome

University of Lethbridge

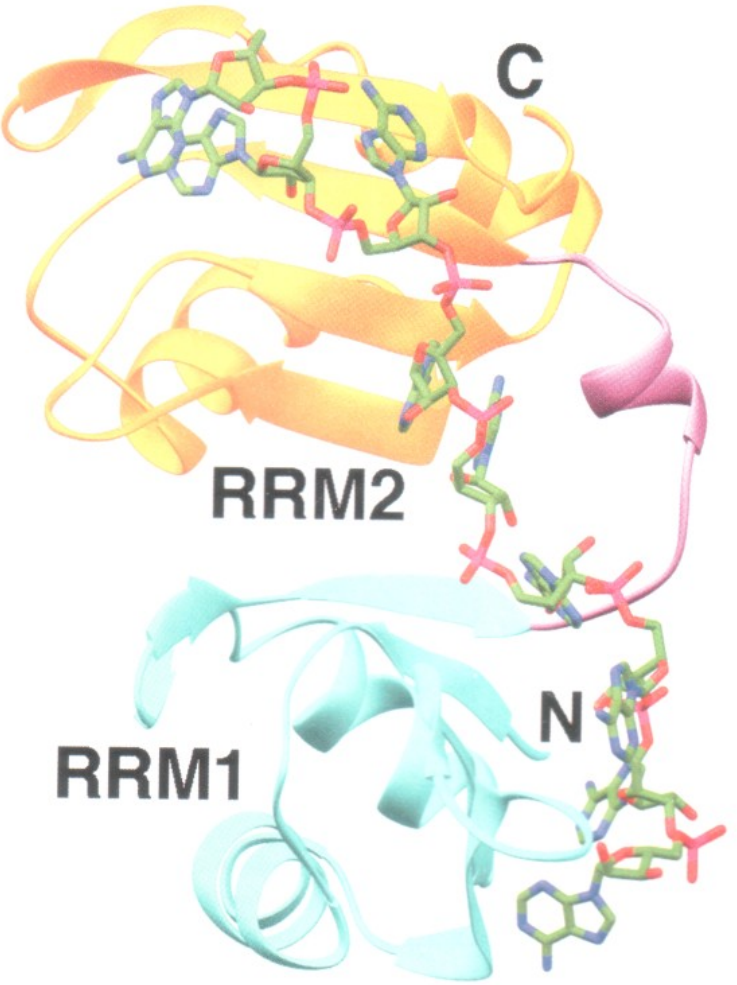
3' - polyadenylation

Poly A phosphorylase adds a variable number of A nucleotides to the transcript

- poly A tail can be up to 100's of nucleotides
- protects mRNA from exoribonucleolytic degradation

Poly A 'tail' can be used to isolate mRNA from eucaryotic cells

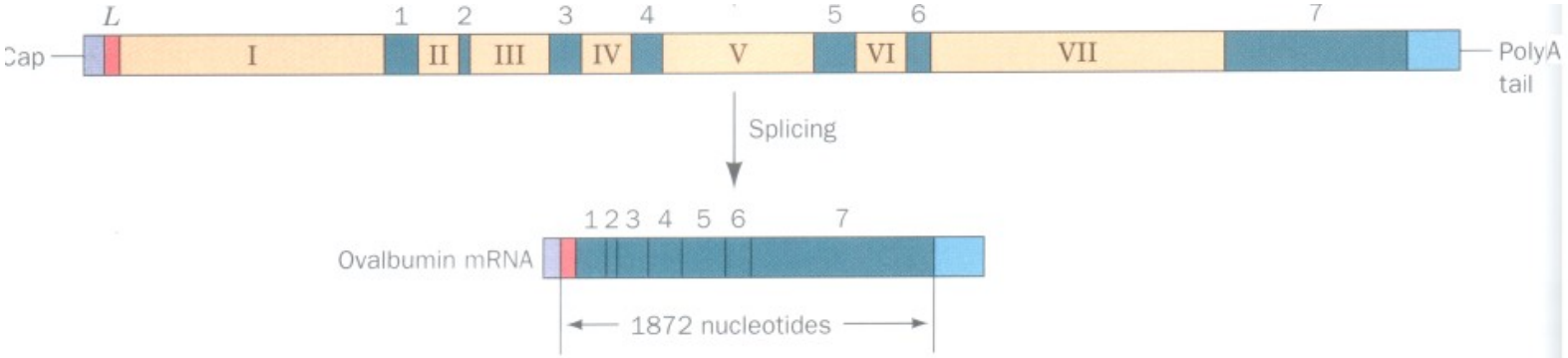
- affinity chromatography where beads are coated with short poly U (or T) oligonucleotides
- allows separation of mRNA from other RNA molecules

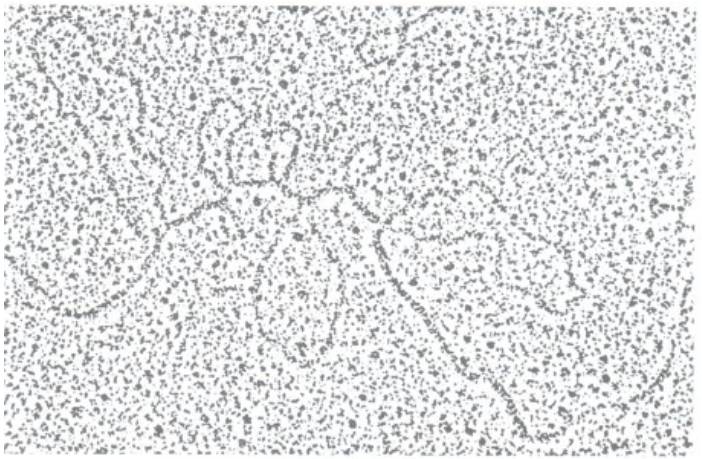


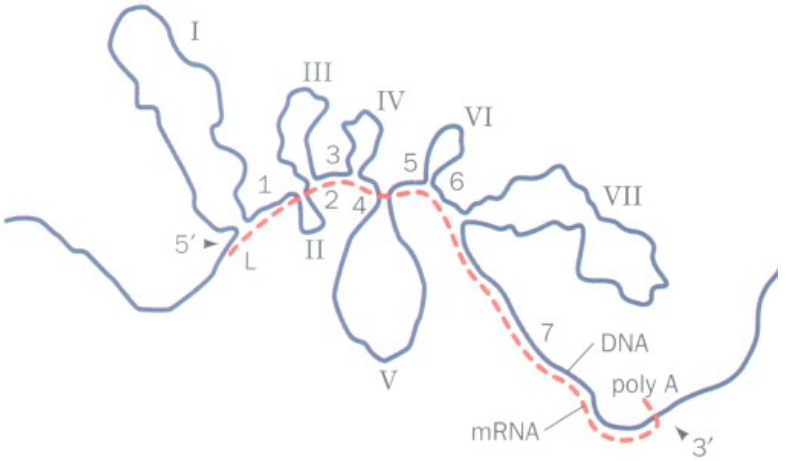
University of Lethbridge

Intron Removal

- Catalyzed by specific ribozymes and ribonucleoproteins (spliceosome)
- Occurs in the nucleus prior to export of mRNA







Mature mRNA hybridized to gene - “loops” are introns

University of Lethbridge

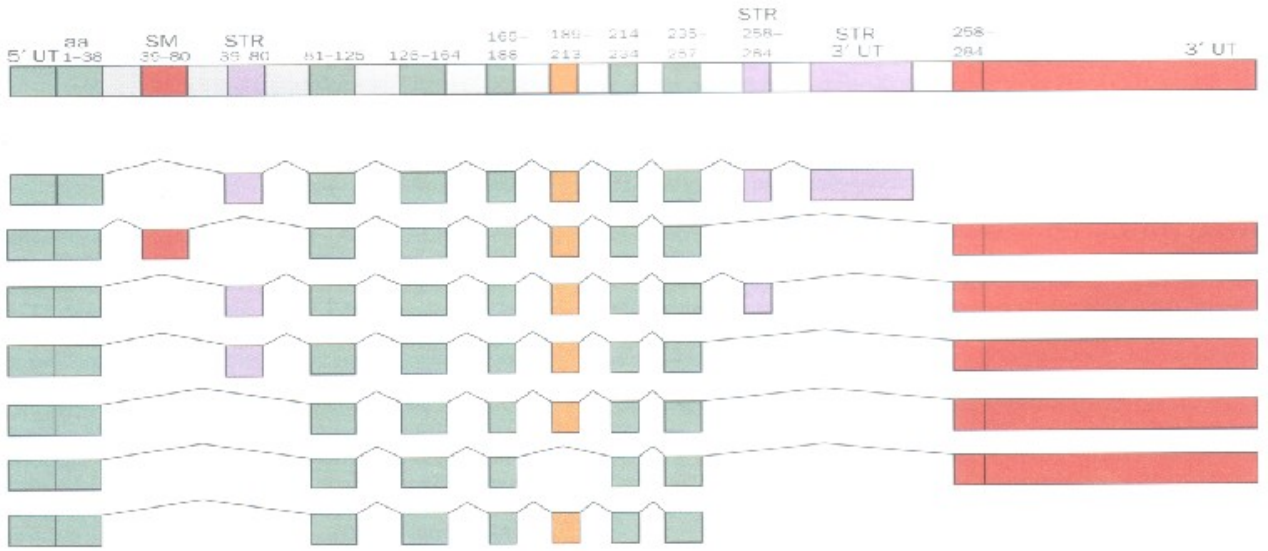
Why do Introns Exist?

Possible Reasons:

Allows for a greater diversity of protein

- **Alternative Splicing** of a single gene into multiple protein products
- **Exon shuffling** within a gene to produce multiple protein products

May increase likelyhood of gene rearrangement producing successfully folded protein (*i.e.* exons are often protein domains)



Multiple proteins produced from a single alternatively spliced gene

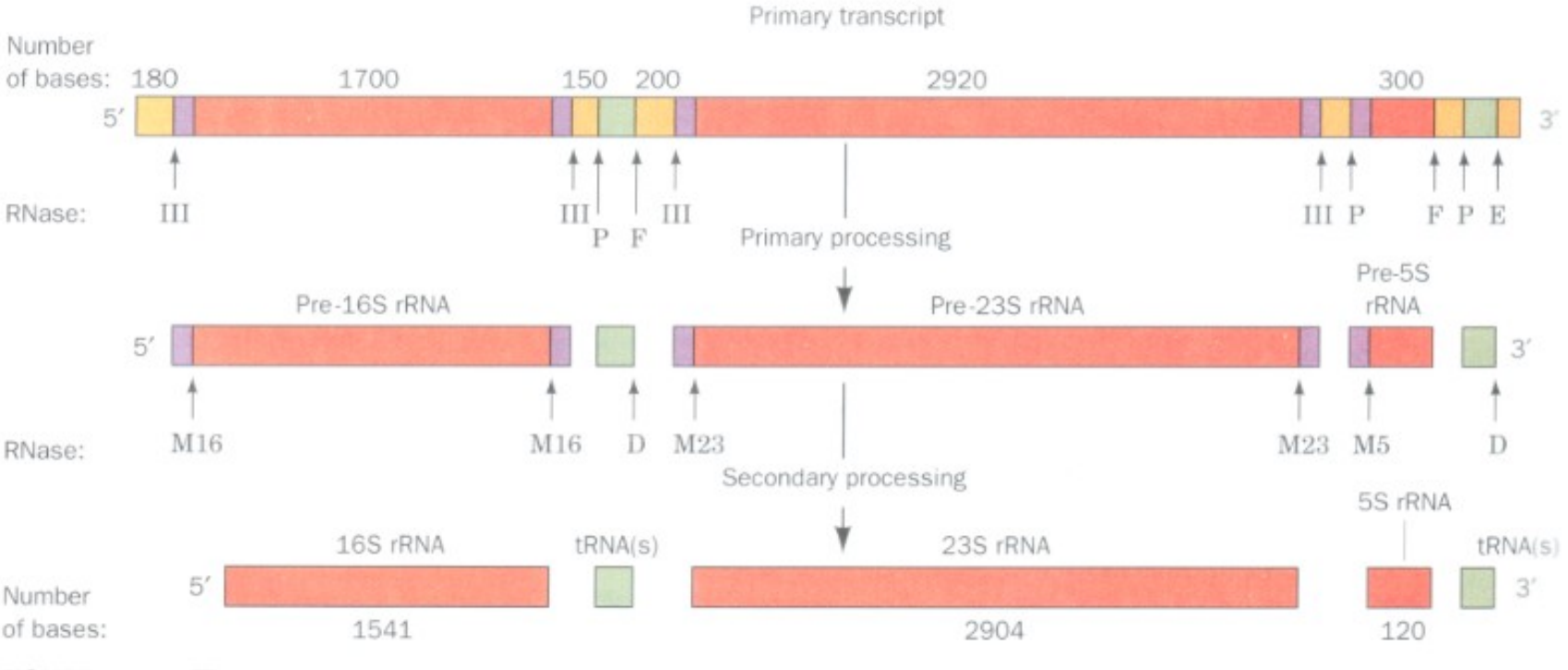
University of Lethbridge

Ribosomal RNA (rRNA)

Large RNA components of ribosomes that are extensively processed. Processing reactions include:

(1) ribonucleolytic cleavage reactions

(2) post-transcriptional modifications (pseudo U, 2'-O methylation, base modification, *etc*)



University of Lethbridge

rRNA and the Ribosome

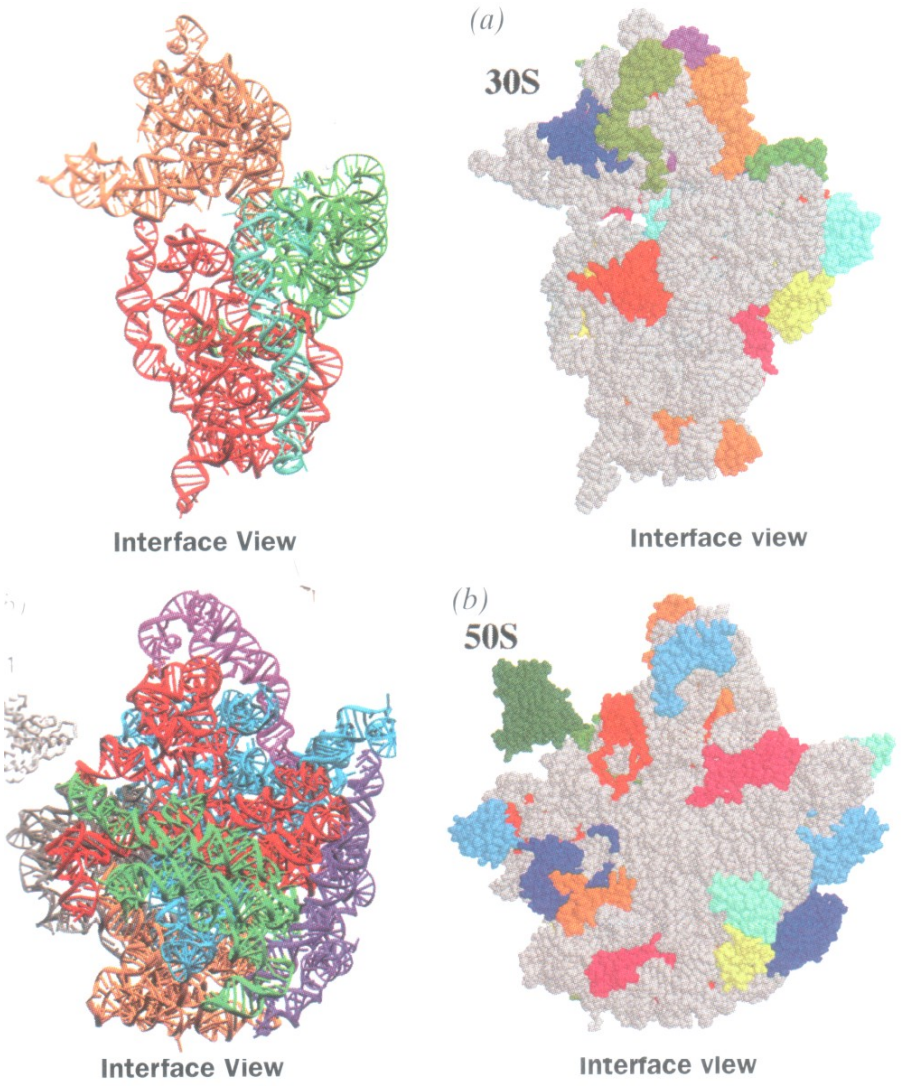
Bulk of the ribosome is rRNA

- RNA shown in left panel
- RNA (grey) and protein shown in right panel

Bacterial ribosomes contain 30S and 50S subunit

- eucaryotic ribosomes are larger and contain 40S and 60S subunit

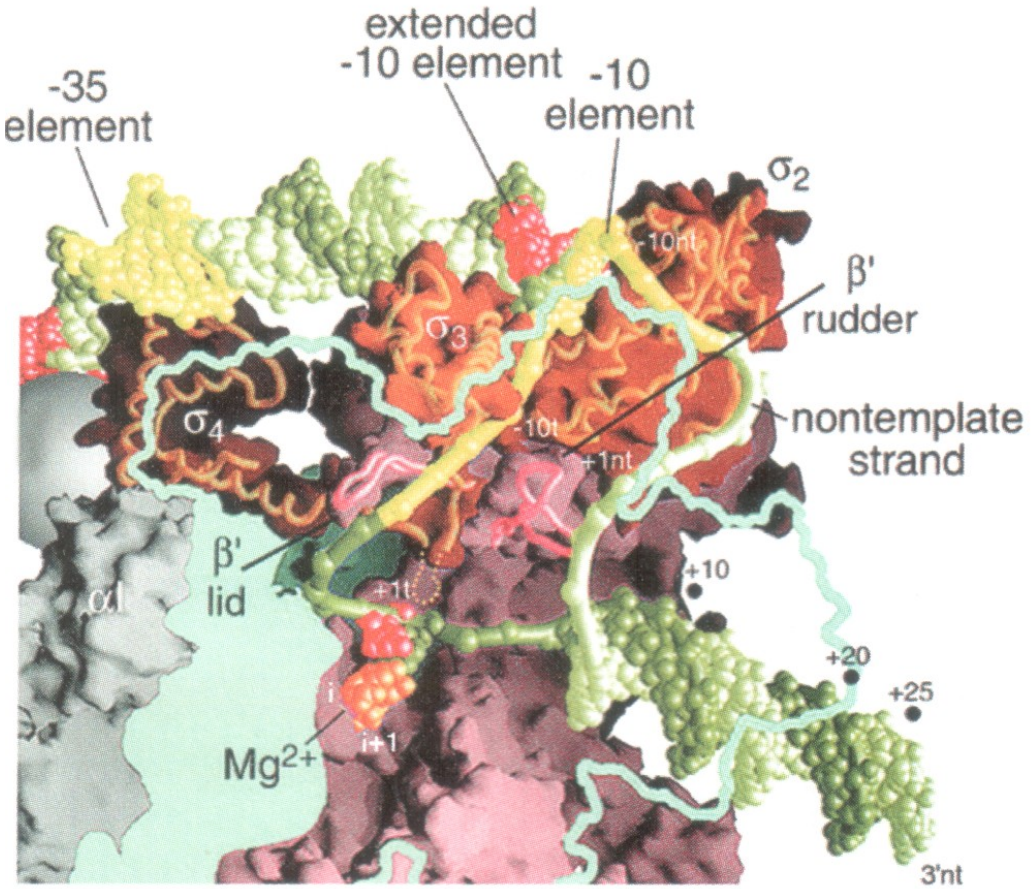
Ribosomes synthesize proteins and are ribozymes (*ie.*the RNA is catalytic)



University of Lethbridge

Chapter 31: Transcription

Voet & Voet: Pages 1216-1231



University of Lethbridge

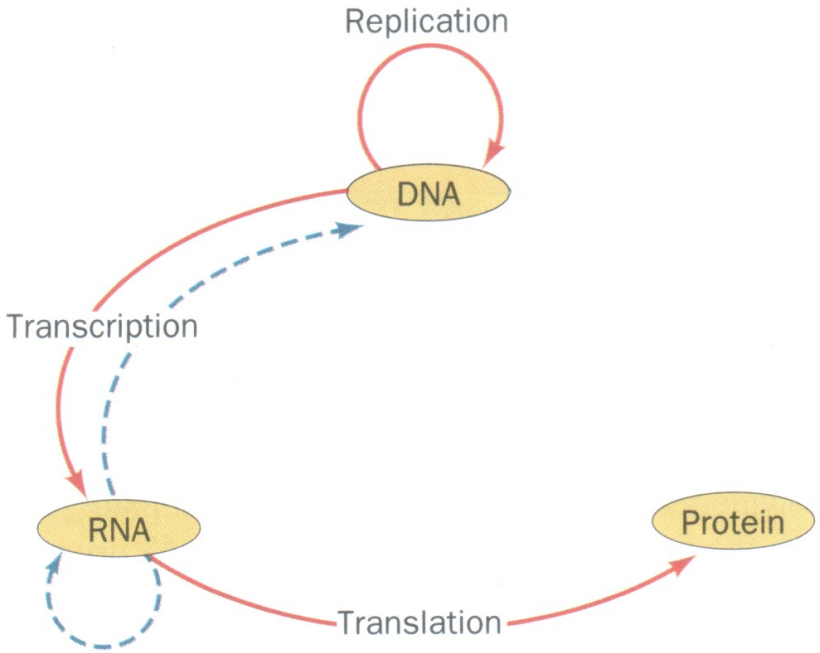
Transcription (discovery)

Casperson & Brachet (1930s)

- DNA (eucaryotes) is located in the nucleus
- RNA is located in the cytosol; primarily associated with protein particles

Multiple researchers (1950s)

- Radioactive ribonucleotides associate with ribosomes (protein particles)



Crick (1958) postulates "Central Dogma"

- DNA directs synthesis of mRNA
- mRNA directs synthesis of proteins

Central Dogma (misnomer) reflects flow of information in living systems

University of Lethbridge

RNA polymerase (RNAP)

DNA dependent RNA polymerase discovered in 1960

- **Bacteria (and some viruses) have a single RNAP while eucaryotes have 4 or 5 RNAPs that each transcribe a different 'class' of RNA**
- **RNAP (tightly) binds DNA and produce mRNA**

Elongation reaction utilizes NTP and releases pyrophosphate (PP_i)

$$(RNA)_n + NTP \leftrightarrow (RNA)_{n+1} + PP_i$$

σ^{70} reversibly binds RNAP

- **$\alpha\beta\beta'\omega\sigma^{70}$ is the RNAP holoenzyme and $\alpha\beta\beta'\omega$ is core RNAP enzyme (or apoenzyme)**

Components of *E. coli* RNA Polymerase Holoenzyme

Subunit	Number of Residues	Structural Gene
α	329	*rpoA*
β	1342	*rpoB*
β'	1407	*rpoC*
ω	91	*rpoZ*
σ^{70}	613	*rpsD*

University of Lethbridge

Bacterial RNAP

Four distinct steps in transcription

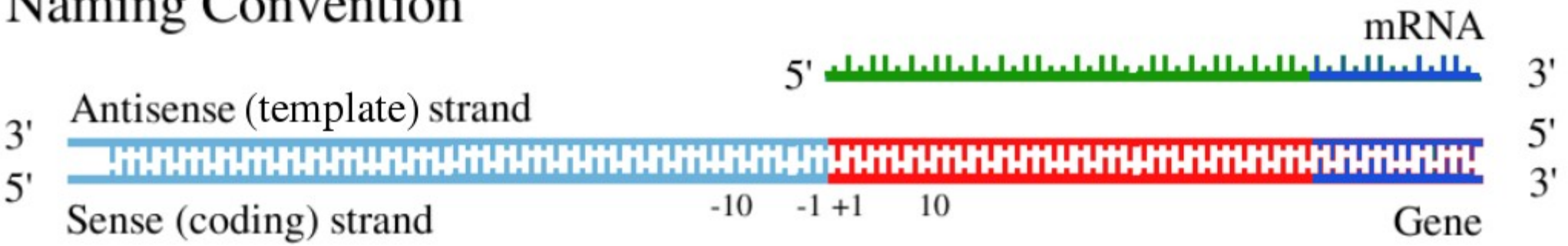
(1) Template Binding — Binding to DNA template near transcription start site

(2) Chain Initiation — Start of transcription

(3) Chain Elongation — 5' → 3' synthesis of mRNA

(4) Chain Termination — End of transcription

Naming Convention



Light Blue – DNA promoter region
Red – Encodes ribosome binding site
Dark Blue – Coding region of gene

Coding strand and mRNA have the same sequence
Template strand and mRNA are complementary

University of Lethbridge

(1) Template (Promoter) Binding

Mutations upstream of gene (~40 nucleotides) affect transcription rates

- **Upstream region referred to as promoter**
- **Bacterial RNAP holoenzyme binds tightly ($K_d = 10^{-14}$ M) to promoter**

RNAP binding protects promoter of genes from DNase I digestion

- **Completely protects region from -20 to +20 dNTPs**
- **Partially protects region from -60 to -20 dNTPs**

Transcription rates are directly proportional to the affinity (K_d) of RNAP for a promoter

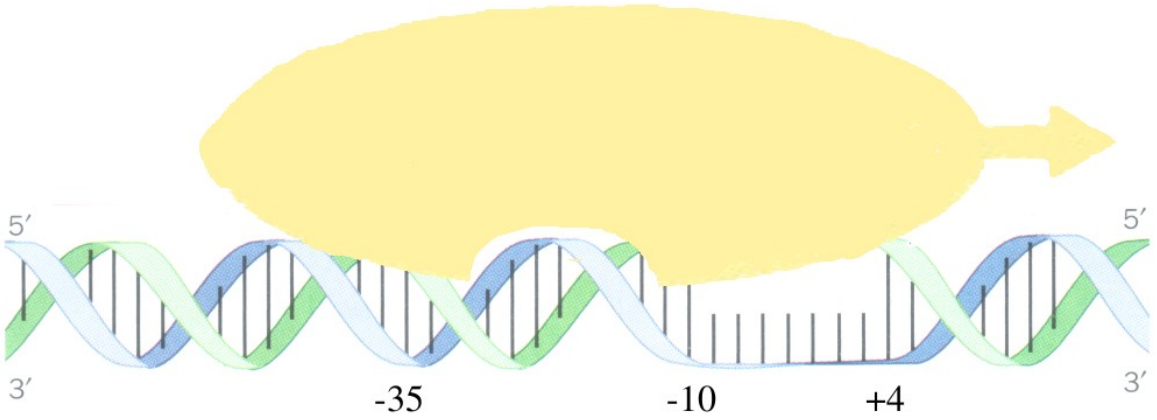
University of Lethbridge

(2) RNAP-promoter Complex

RNAP-promoter complex has two distinct conformations

- **"closed" complex** promoter has a dsDNA conformation
- **"open" complex** promoter has both ssDNA (transcription bubble) and dsDNA regions

"Open" Complex transcription bubble is required for transcription initiation



RNAP-promoter "open" complex

- scale of RNAP and promoter more reasonably sized

University of Lethbridge

(2) Finding the Promoter (σ^{70})

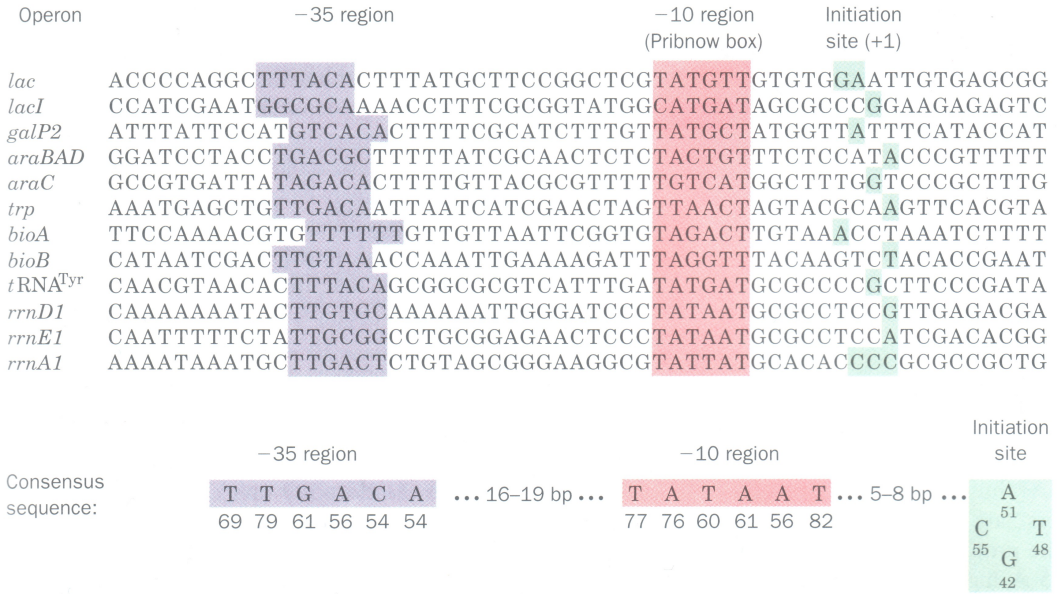
- How does the RNAP locate the promoter?
  - Holoenzyme binds to non-promoter DNA (10^{-7} M) much more weakly than the core enzyme (10^{-14} M)
- σ^{70} subunit allows the holoenzyme to 'slide' along the DNA in search of the promoter
  - Prevents tight binding to all but the promoter region
  - once the promoter is located and transcription is initiated; the σ^{70} dissociates and the core enzyme binds tightly
- Bacteriophage (bacterial viruses) express σ factors
  - bind to RNAP (in place of bacterial σ^{70}) and target viral promoters
  - evidence σ factors are involved in promoter recognition

University of Lethbridge

Consensus Bacterial Promoter

Bacterial promoters share common features that explain experimental results

- conserved regions near -10 (TATAAT) and -35 (TTGACA) involved in RNAP binding
- sequence from -10 to -35 and from +1 to -10 are unimportant BUT their length is important
- poorly conserved CAT sequence about +1



Highly expressed bacterial genes often have UP (upstream promoter) element

- AT rich region upstream of -35

University of Lethbridge

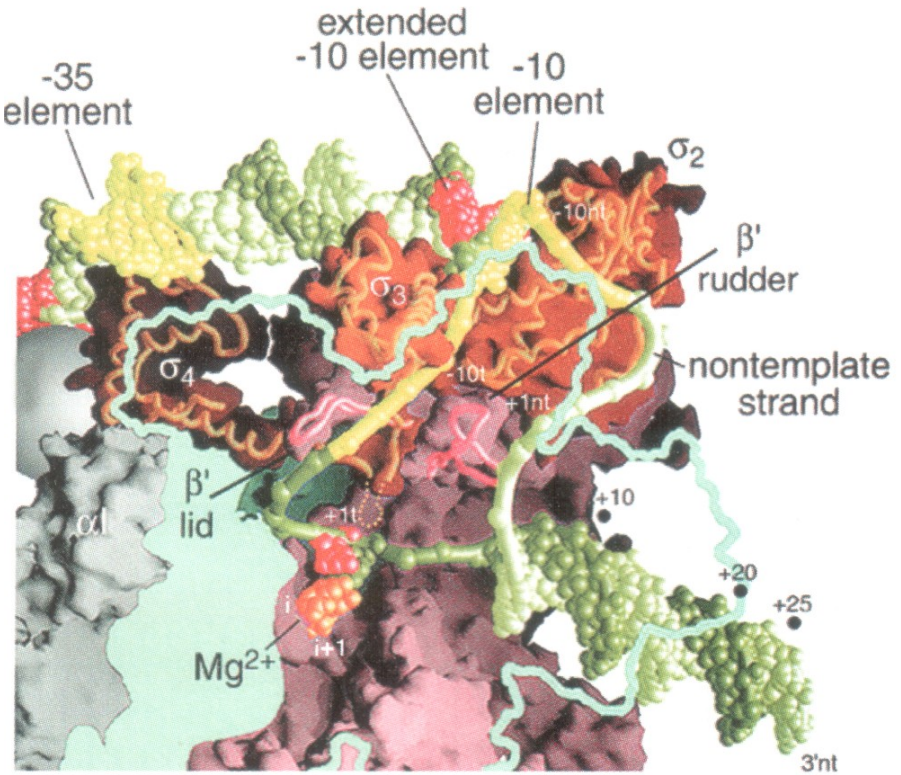
Transcription Initiation

Transcription initiation does not require a primer

$$pppA + pppN \leftrightarrow pppApN + Pp_i$$

Initiation is difficult based upon the number of short abortive transcripts

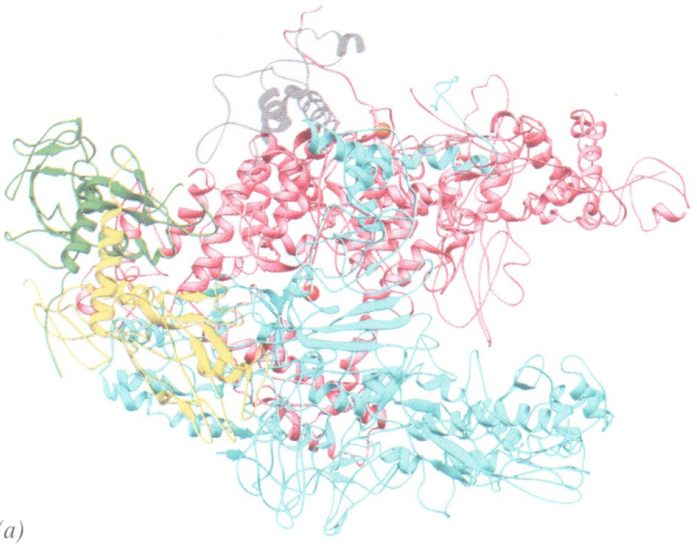
- **Up to five abortive transcripts for every full-length transcript**
- **once σ^{70} dissociates, initiation is complete and transcription rarely aborts**



Direction of RNAP movement

University of Lethbridge

RNAP structure and function

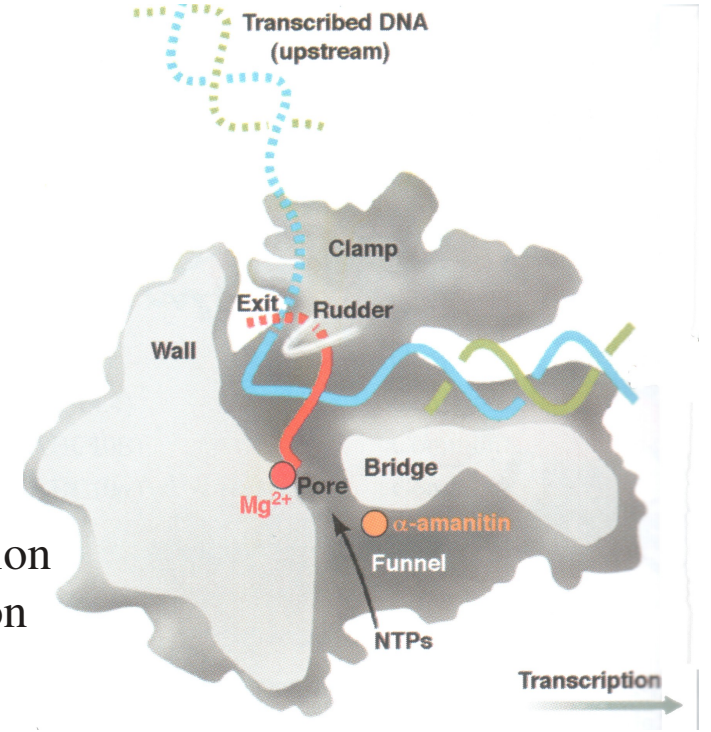


(a)

RNAP β subunits (blue & red) form a "crab claw" shape

Promoter binding at ends of pincers causing the pincers to close (1.0 nm)

- **antisense DNA strand moves to base of pincer in open complex**
- **essential Mg^{2+} at base of pincer that is involved in catalysis**

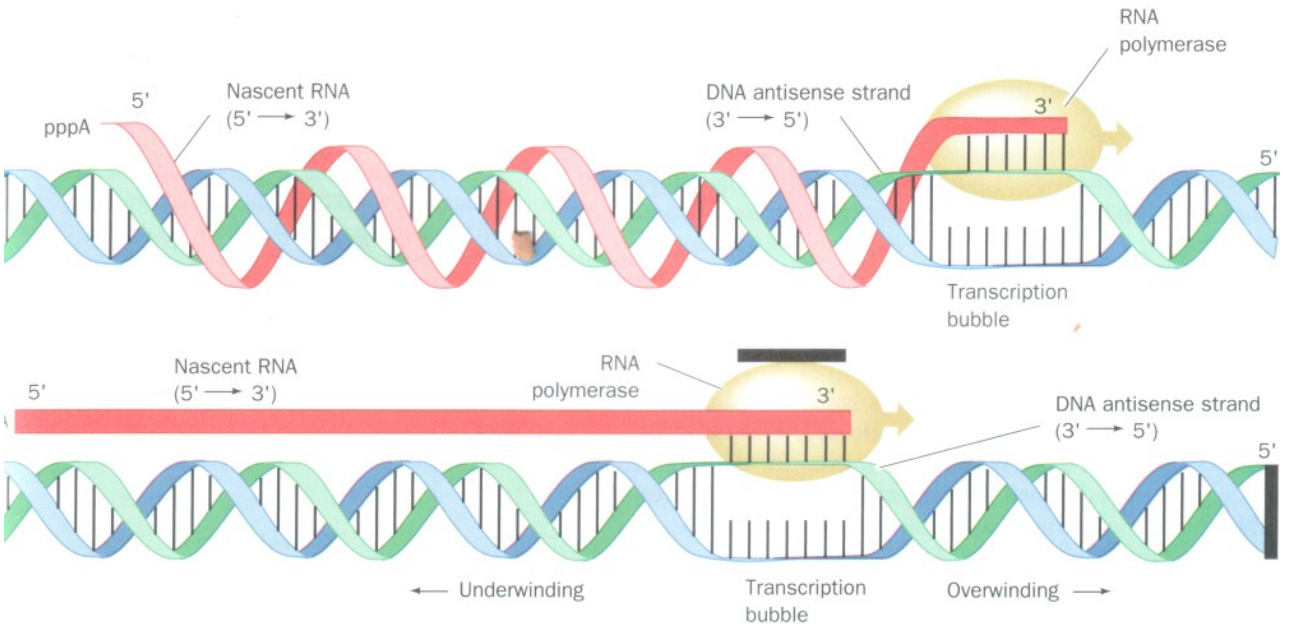


Transcription Elongation

University of Lethbridge

(3) Chain Elongation

- Elongation is rapid (20-50 nucleotides/second)
- 5' → 3' like replication
  - demonstrated using NTP derivatives – 3' deoxyadenosine-5'-triphosphate is a substrate the stops elongation
- Elongation generates supercoils in dsDNA ; consistent with RNAP traveling along a single face of dsDNA



University of Lethbridge

Transcription is Error Prone (in Procaryotes)

Typically 1 in 1000 ribonucleotides are incorrectly transcribed

- Apparently speed is more important than fidelity in bacteria

Errors are not always harmful (surprise!)

(1) multiple transcripts arise from the same gene so a correct copy is typically produced

(2) redundancy of "genetic code" means many mutations do not result in an amino acid change in the resulting proteins

(3) many mutations in proteins are functionally innocuous

(4) (eucaryotes) excision of introns means errors may not appear in mature transcript

University of Lethbridge

(4) Transcription Termination

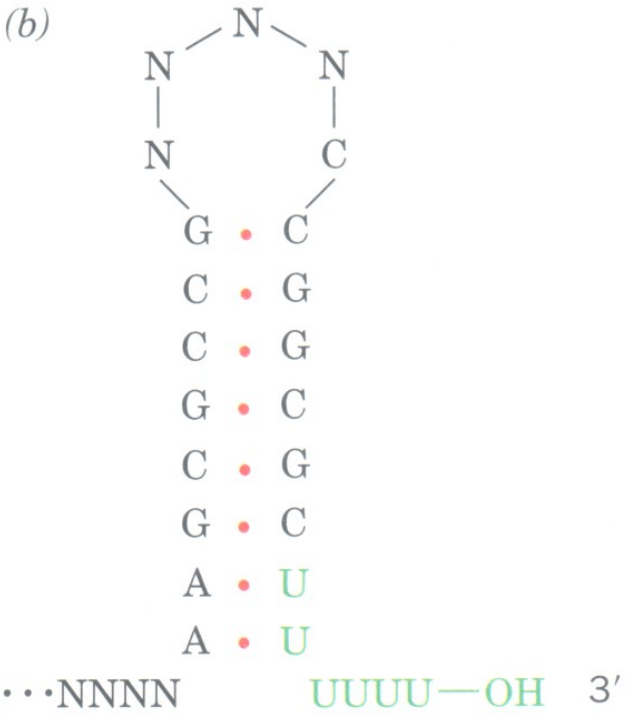
Two common mechanism

ρ (protein) independent (most termination)

- 4-10 AT nucleotides following GC rich region (DNA)
- Destabilize RNAP binding to DNA
    - palindrome (2 fold symmetric sequence) forms in mRNA
    - AT base pairing in DNA

ρ (protein) dependent (many)

- hexameric ρ protein binds near RNAP and unwinds DNA-RNA duplex
- energy dependent



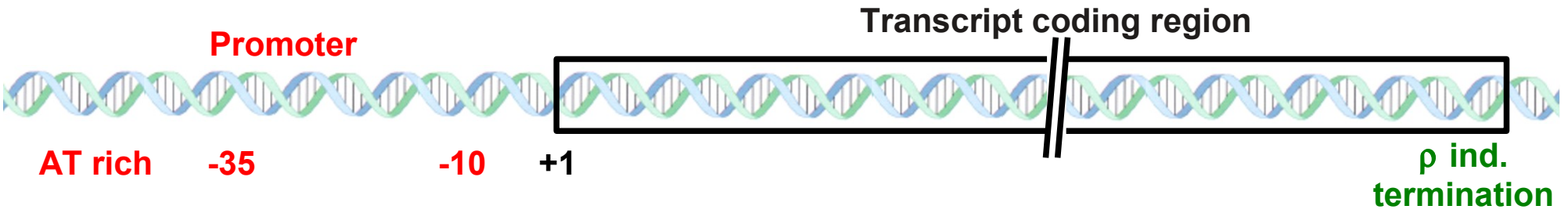
mRNA

University of Lethbridge

Summary

DNA associated with mRNA production (procaryotes)

- RNA polymerase binding site preceeds transcription start site (red)
- ρ independent termination sequence at end of transcript stop site (green)



Transcript (procaryotes)

- 5'-UTR (untranslated region) contains ribosome binding site
- 3'-UTR