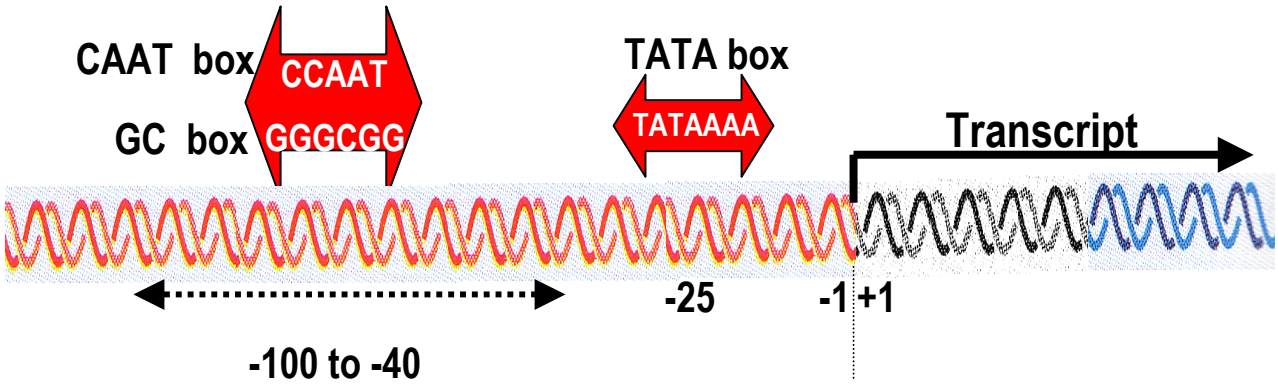
BIOC4004 - Industrial Biochemistry
Lecture 02 - Fri Jan 09, 04

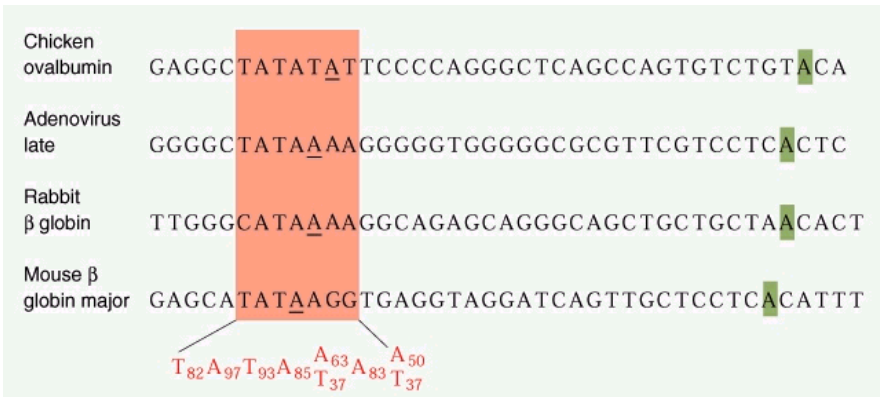
Topics for the Day:

- Promoters
    - Prokaryotic vs Eukaryotic
- Transcription
    - Prokaryotic vs Eukaryotic
- Transcriptional regulation
    - Transcription factors
- Translation
    - Prokaryotic vs Eukaryotic
- The Genetic Code
- Regulation of translation

Generalized view of Eukaryotic Promoters

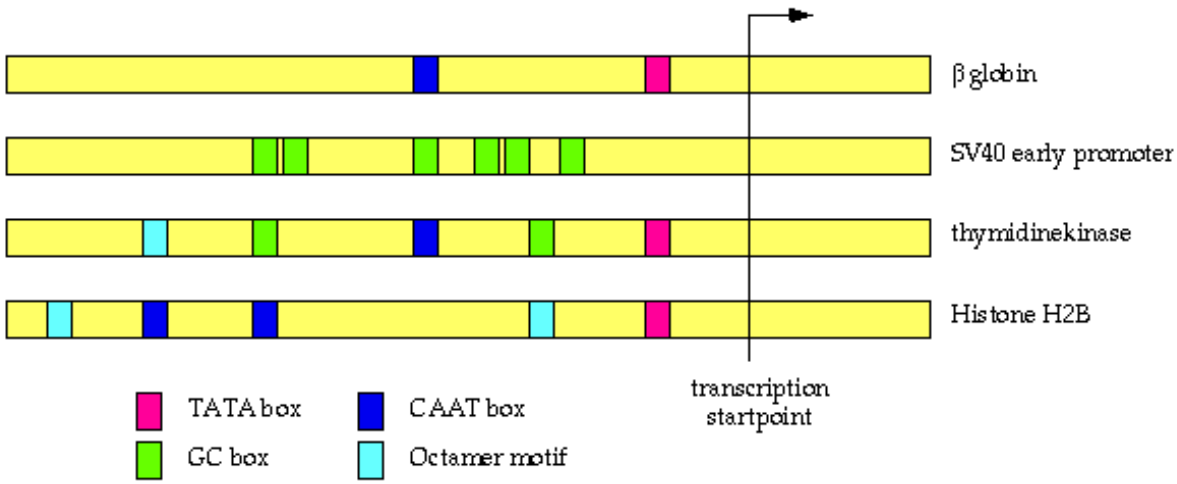


Gannon et al. (1978) looked at transcription start sites...



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**

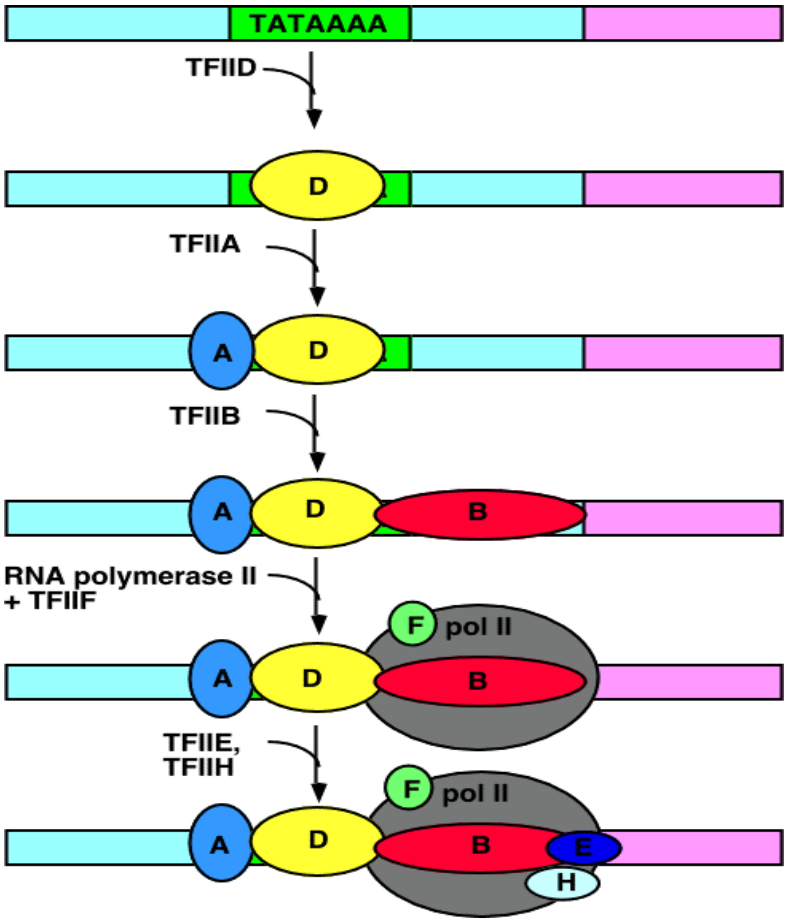
- Eukaryotic RNA polymerases are assembled from numerous subunits
    - as many as 14 different subunits come together

- 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!!!

Generalized view of Eukaryotic Promoters (cont)

Eukaryotic Trancription



TFIID: TBP (TATA box binding protein)
TAFs (interact with regulatory sequences and their TFs

TFIIA: stabilizes TFIID interaction with promoter

TFIIB: interaction between TFIID and PolII-TFIIF

TFIIF: recruits PolII to promoter; helicase activity

TFIIE: binding and stimulation of transcription by recruiting TFIIH

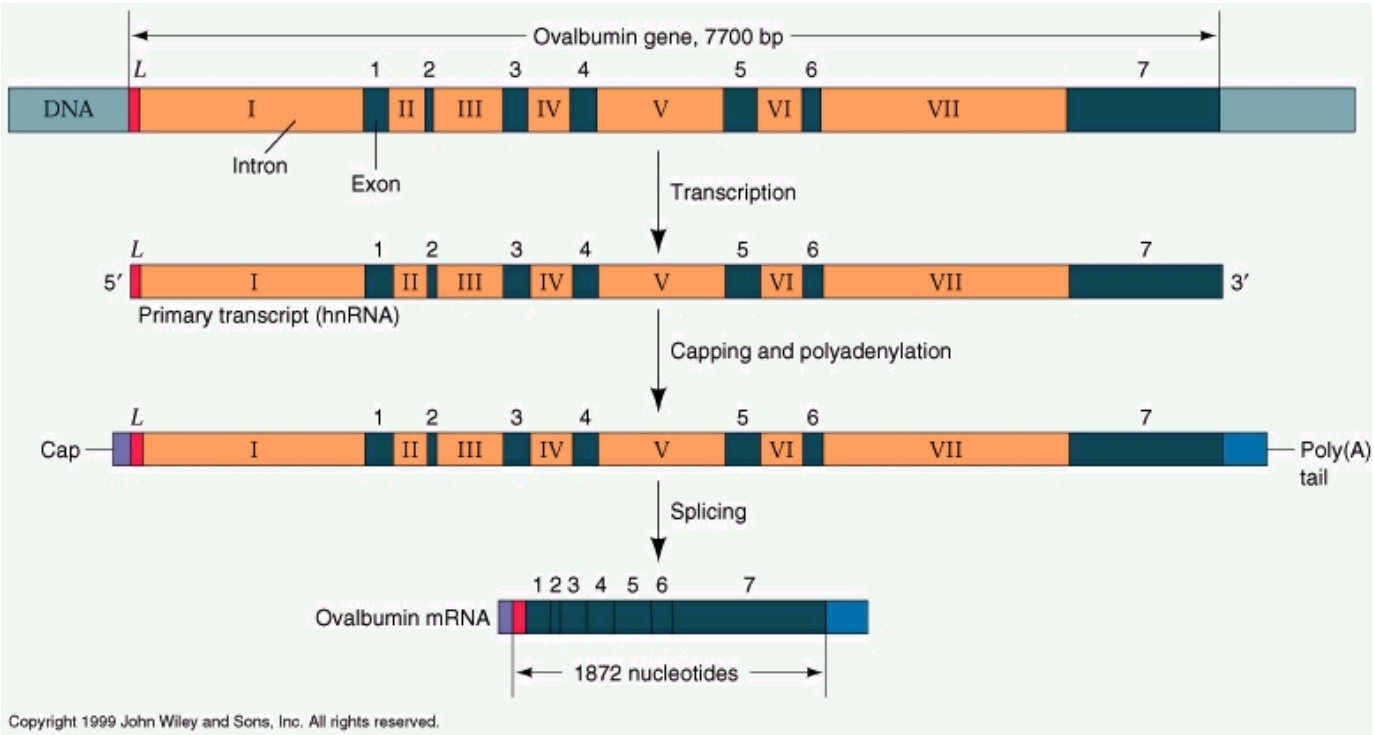
TFIIH: kinase activity (activates PolII by phosphorylation - promoter clearance)

Eukaryotic Transcription Termination:

- process still largely unknown
- polymerase extends ~1000 bp past the mature end of the mRNA
- termination signals cause the polymerase to fall off the DNA template
    - sequences required for termination are unknown
    - polymerase can fall off at one of several different places

        ie. a heterogeneous mixture of mRNAs of different length is produced

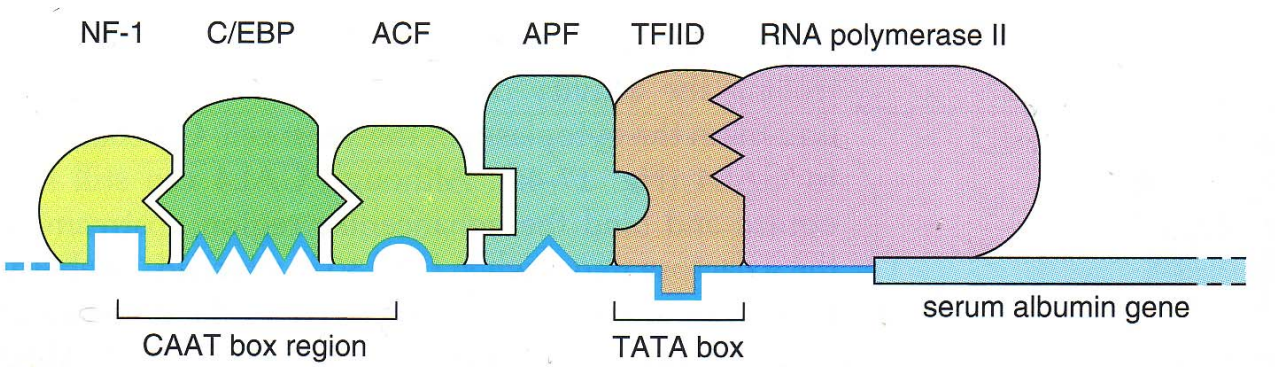
- mature 3'-end is generated by post-transcriptional cleavage
    - followed by polyadenylation(except for non-poly-A mRNAs



Generalized view of Eukaryotic Promoters (cont)

• Promoters contain combinations of "upstream regulatory motifs" to regulate expression

- binding sites for **Transcription Factors (TFs)**
- cooperative interaction between TFs
- combinations allow fine tuning of gene regulation
  - temporal; tissue specific; stimulus-specific (hormone, growth factor, metabolite…)



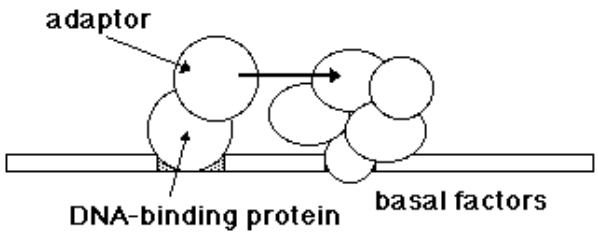
example: the Serum Albumin gene

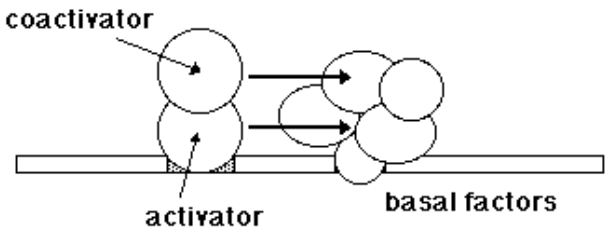
- failure of NF-1, C/EBP or ACF to bind significantly reduces transcription
- APF is very specific for the serum albumin gene
- The combination of TFs increases binding affinity of RNA pol II for promoter

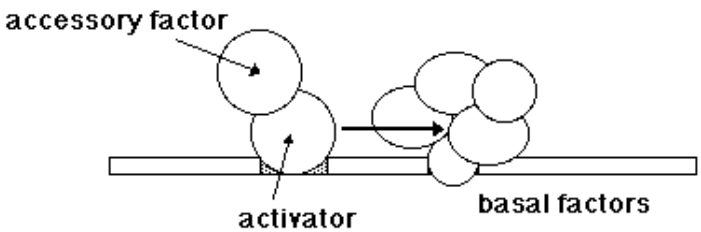
• **Gene regulation networks** can be achieved when different genes share binding sites for a given TF

Generalized view of Eukaryotic Promoters (cont)

Transcription Factors can come together in cooperative fashion:







Transcription Factor Activity is tightly regulated:

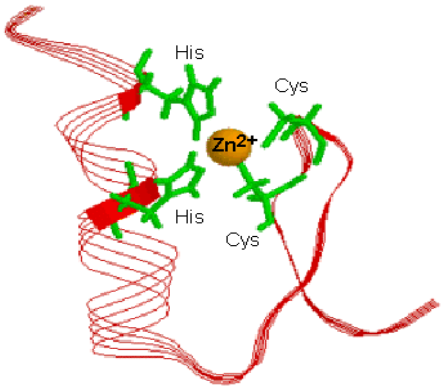
- control of TF levels
- post-translational modification (phosphorylation, dephosphorylation)
- localization
- binding to an inhibitor or activator (a molecule, a protein)
- change in oligomerization state

A quick word on Transcription Factors

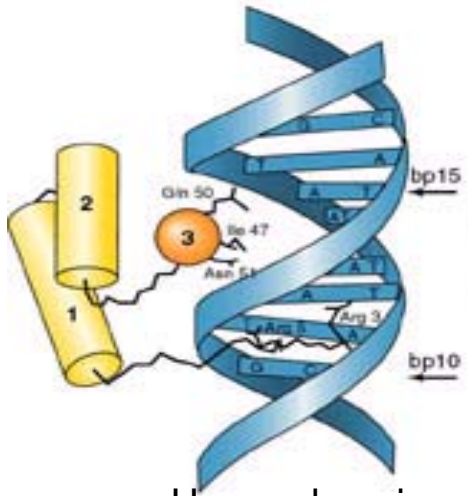
Need two things:

- DNA binding motif
- Activation (or repressing) domain

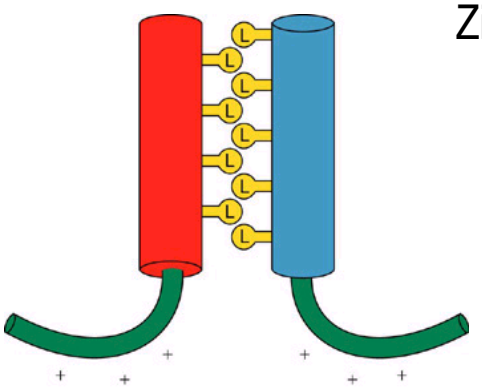
Some DNA Binding Motifs:



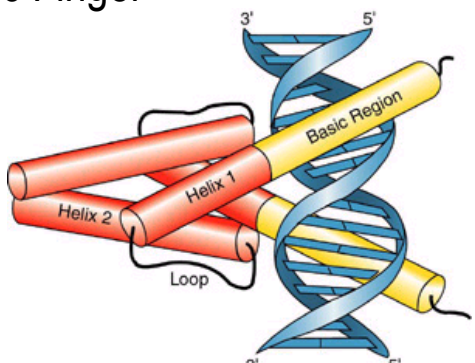
Zinc Finger



Homeodomain



Leucine Zipper



bHLH (basic Helix-Loop-Helix)

Transcription-activating Domains

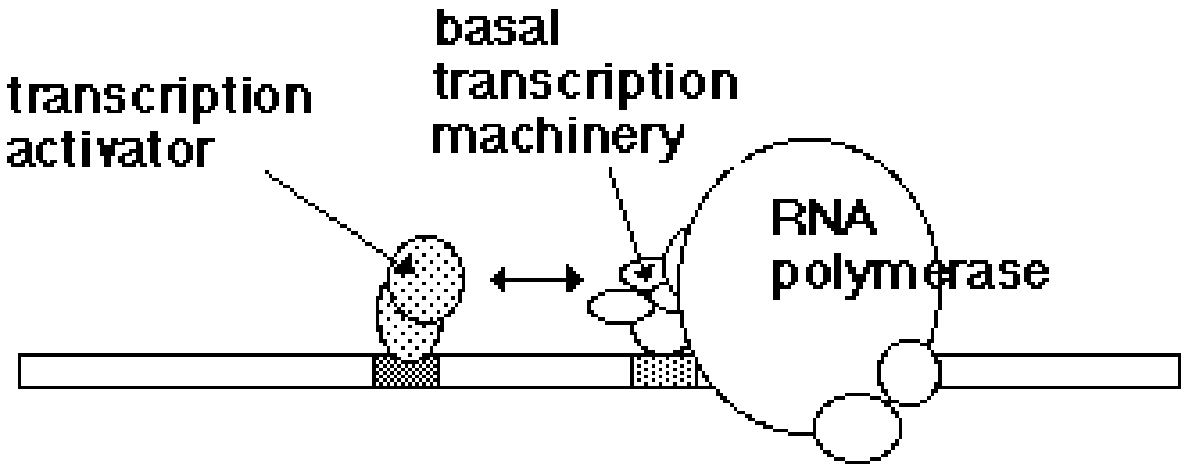
Three kinds of protein domains are observed

1. acidic domains, where the amino acid side chains are acidic in nature (glutamic acid, aspartic acid)

2. glutamine-rich domains, with about 25% glutamine in the sequence

3. proline-rich domains

A quick word on Transcription Factors (cont)



What do Transcription Factors Do ?

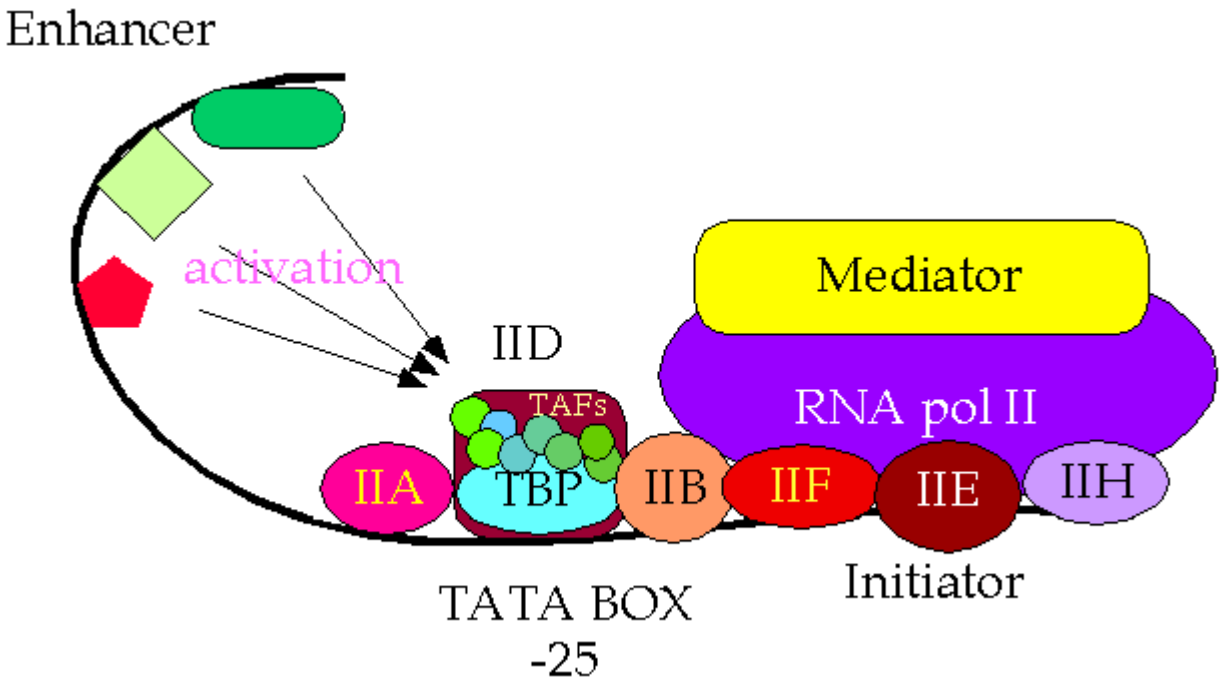
- basal transcription machinery recruits RNA polymerase to the promoter
    - low to nil transcription levels

- transcription factors interact with basal factors to increase the rate of transcription
    - stimulate assembly of preinitiation complexes
    - stimulate formation of an open complex
    - stimulate promoter escape

- binding sites for transcription activators must be made accessible
    - some transcription factors function to re-model chromatin
    - remodelling makes the DNA accesible to other factors
    - "euchromatin" vs "heterochromatin"
        - DNA packaging "silences" gene expression
        - histone modification (acetylation, methylation & phosphorylation) activates DNA

The ever-mysterious "Enhancers"

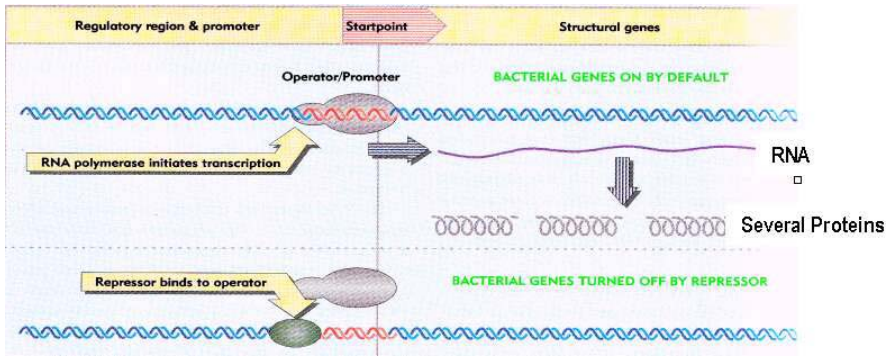
- In eukaryotic cells
  - sequences that influence the rate of transcription
  - interact with specific proteins.
  - can be located within, near, or **very far** (ie. Kbs) from the gene's promoter
  - can be upstream or downstream of gene



- Bind
- Interact with regulatory protein attached to promoter
  - DNA loop formed
- Enhance RNA pol binding

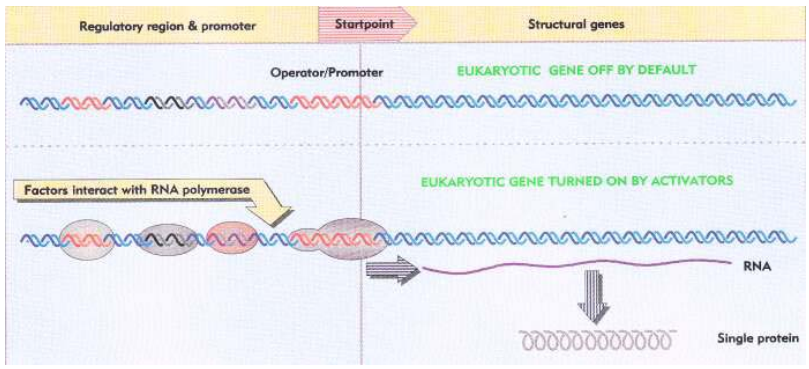
Differences between prokaryotic and eukaryotic genes

• **Prokaryotes:**



- more often than not, gene is ON by default
- most gene regulation is negative (ie. repressors proteins turn off transcription)
- a lot of related genes are arranged in operons
- coding sequences are co-linear with the encoded polypeptide (ie. no introns)
- no nucleus; mRNA can be translated as it's being transcribed

• **Eukaryotes:**



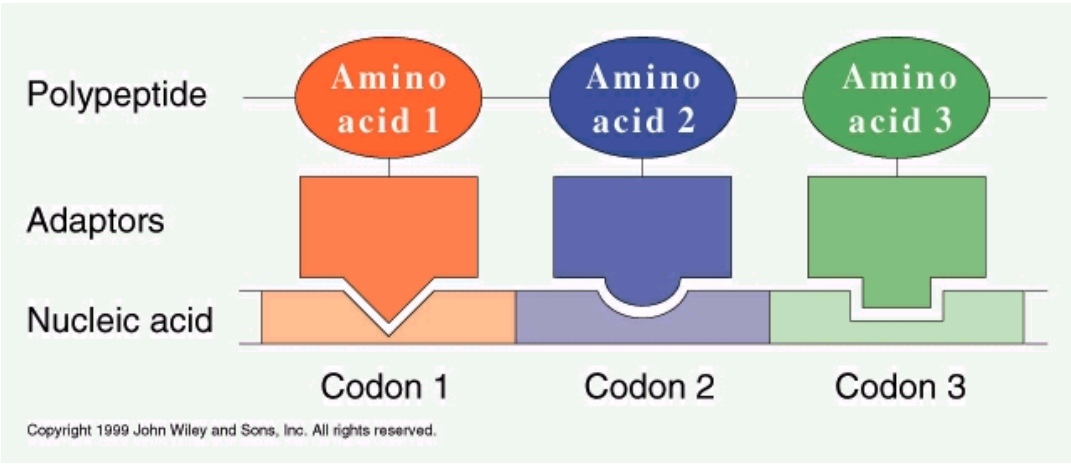
- most genes are OFF by default
- most gene regulation is positive (ie. transcription factors needed for activation)
- promoters are rich with TF-binding motifs
- the majority of genes contain introns
- modifications are made before mRNAs are exported out of the nucleus
    - 5' cap
    - 3' poly[A] tail
    - splicing to remove introns (intervening sequences)
    - **alternative splicing can generate multiple polypeptides from same gene**
- translation occurs outside the nuclus:
    - secreted proteins - rough ER
    - cytosolic proteins - cytoplasm

From RNA to Protein: the process of Translation

- **Translation** decodes the mRNA sequence into the polypeptide that it encodes

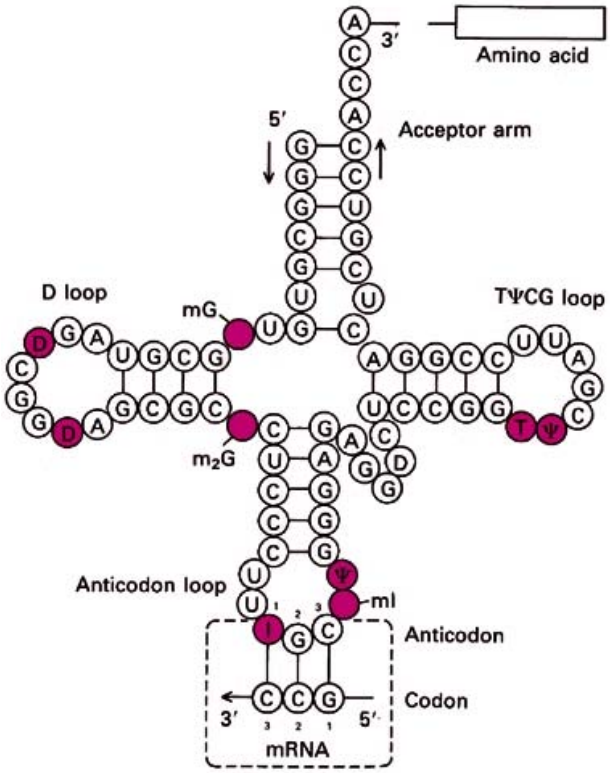
- **Machinery Required**
  - mRNA
  - amino acids, tRNAs & aminoacyl tRNA synthetases
  - ribosomes & other protein factors
  - an initiator tRNA
  - Energy !!! (GTP & ATP)

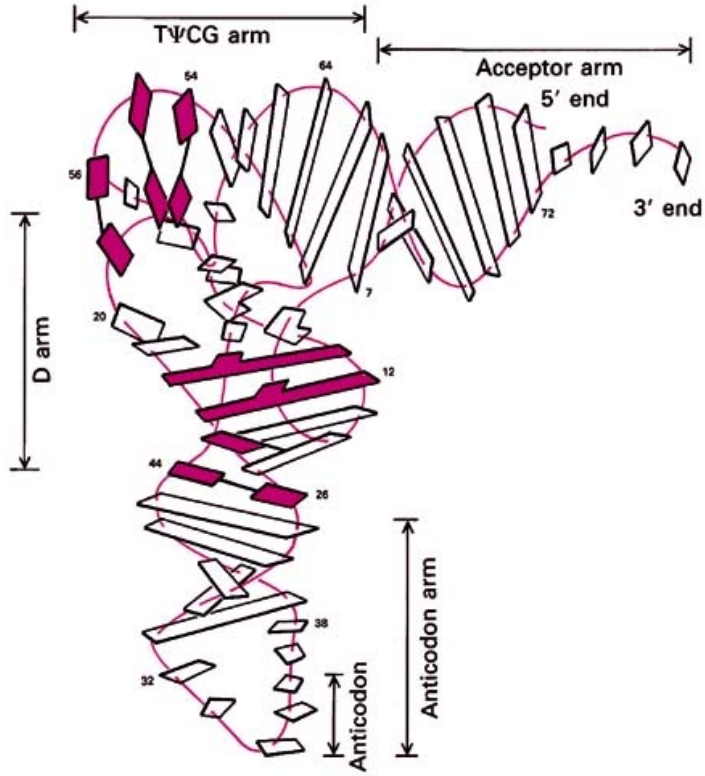
- **Linking the DNA to protein synthesis**
  - FrancisCrick's "Adaptor Hypothesis"



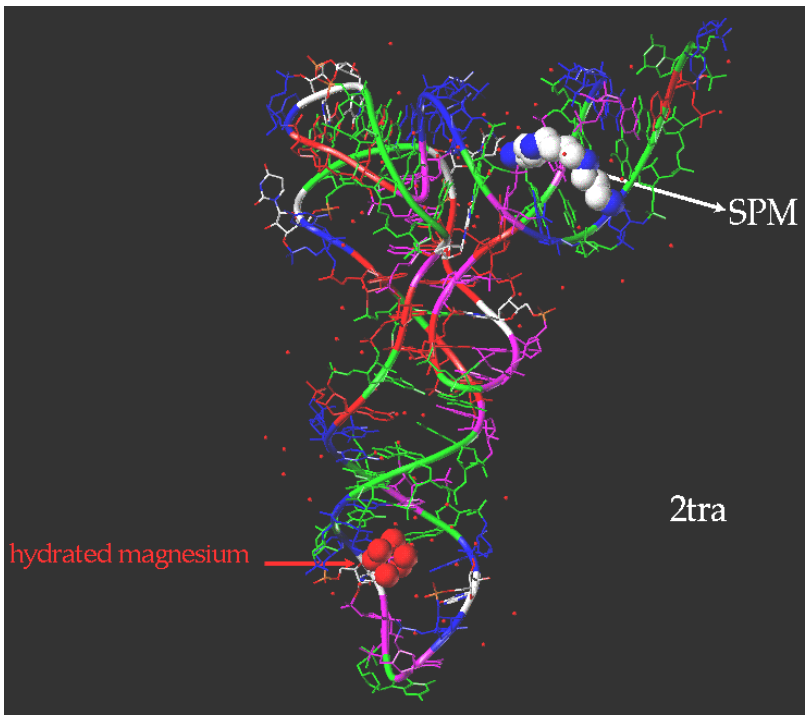
- The "adaptor" is tRNA!!!
- Specific tRNA synthetases "charge" the correct amino acid onto each specific tRNA

A tRNA looks like this...





"An L-shaped 3-D structure"

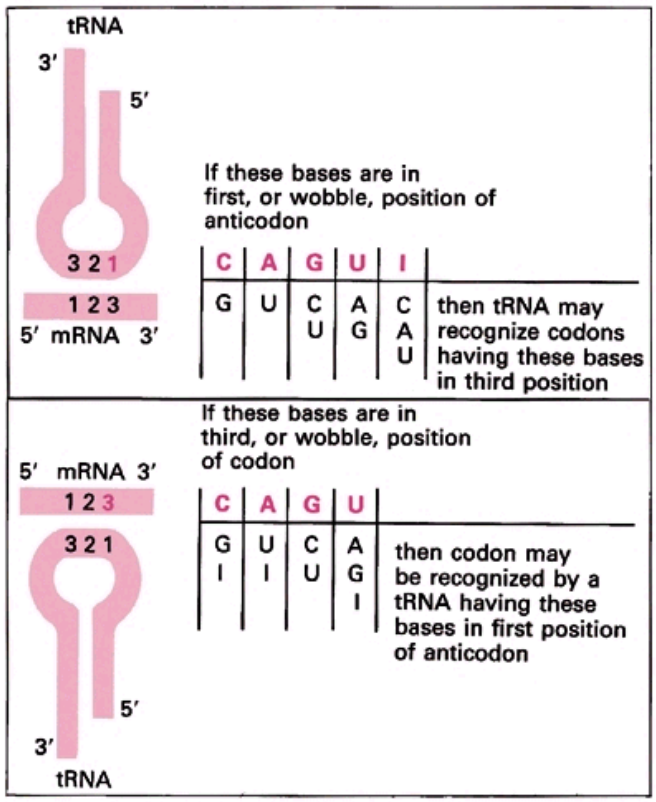


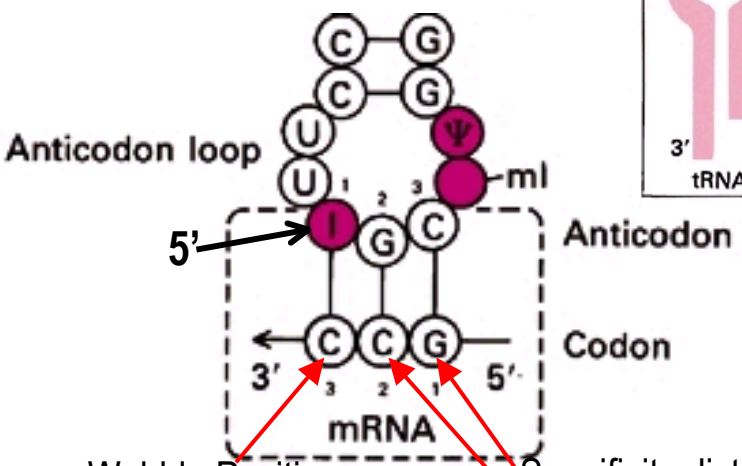
Francis Crick's "Wobble Hypothesis"

We use 20 amino acids but have 61 codons to code for them… what's up with that ????

- **Inosine** was found at the 1st (5') anticodon position (Holley, 1965)
- proposed that codon/anticodon specificity only requires 1st and 2nd codon base-pairing
- 3rd base follows non-standard base pairing with 1st anticodon position

5' anticodon base	3' codon base
A	U
C	G
G	C or U
U	A or G
I	A or C or U

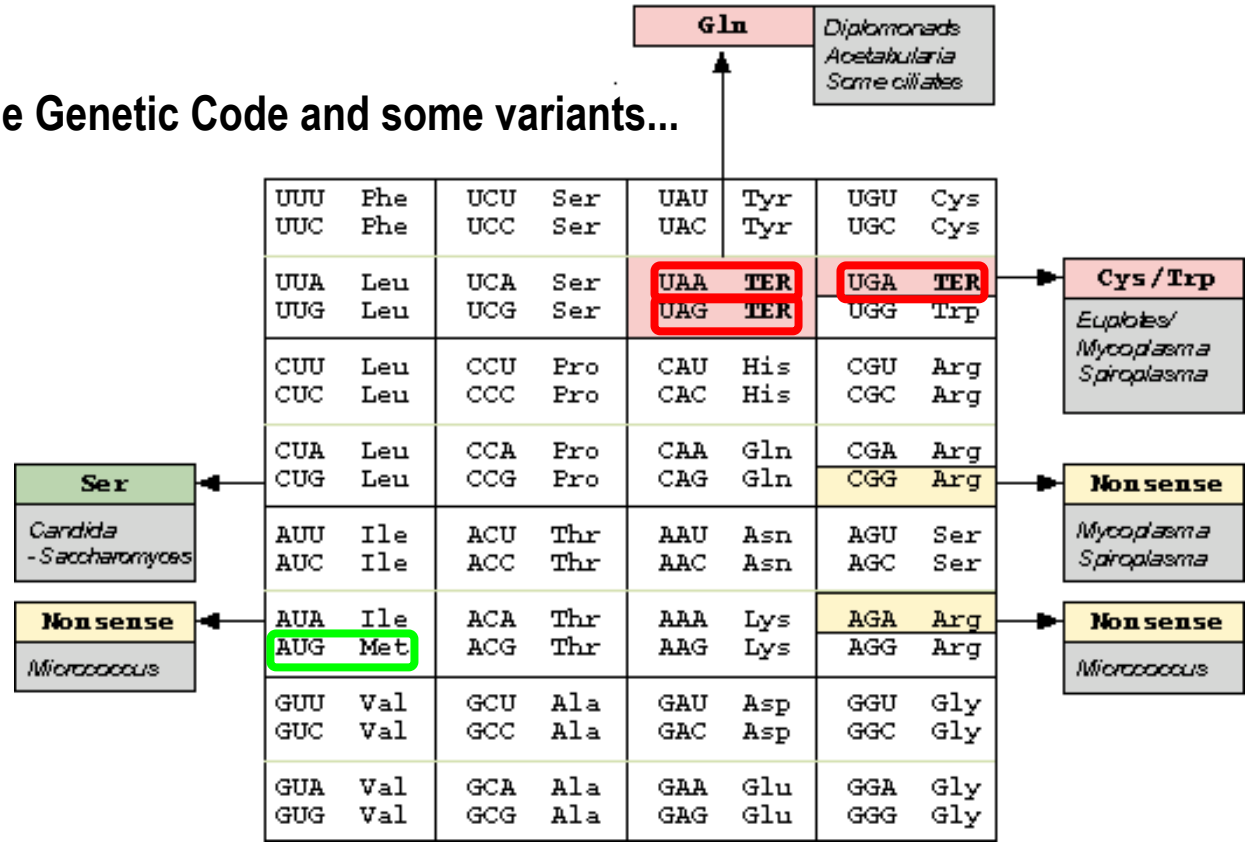




Wobble Position
- 3rd codon position

Specificity dictated here
- 1st & 2nd codon positions

The Genetic Code and some variants...

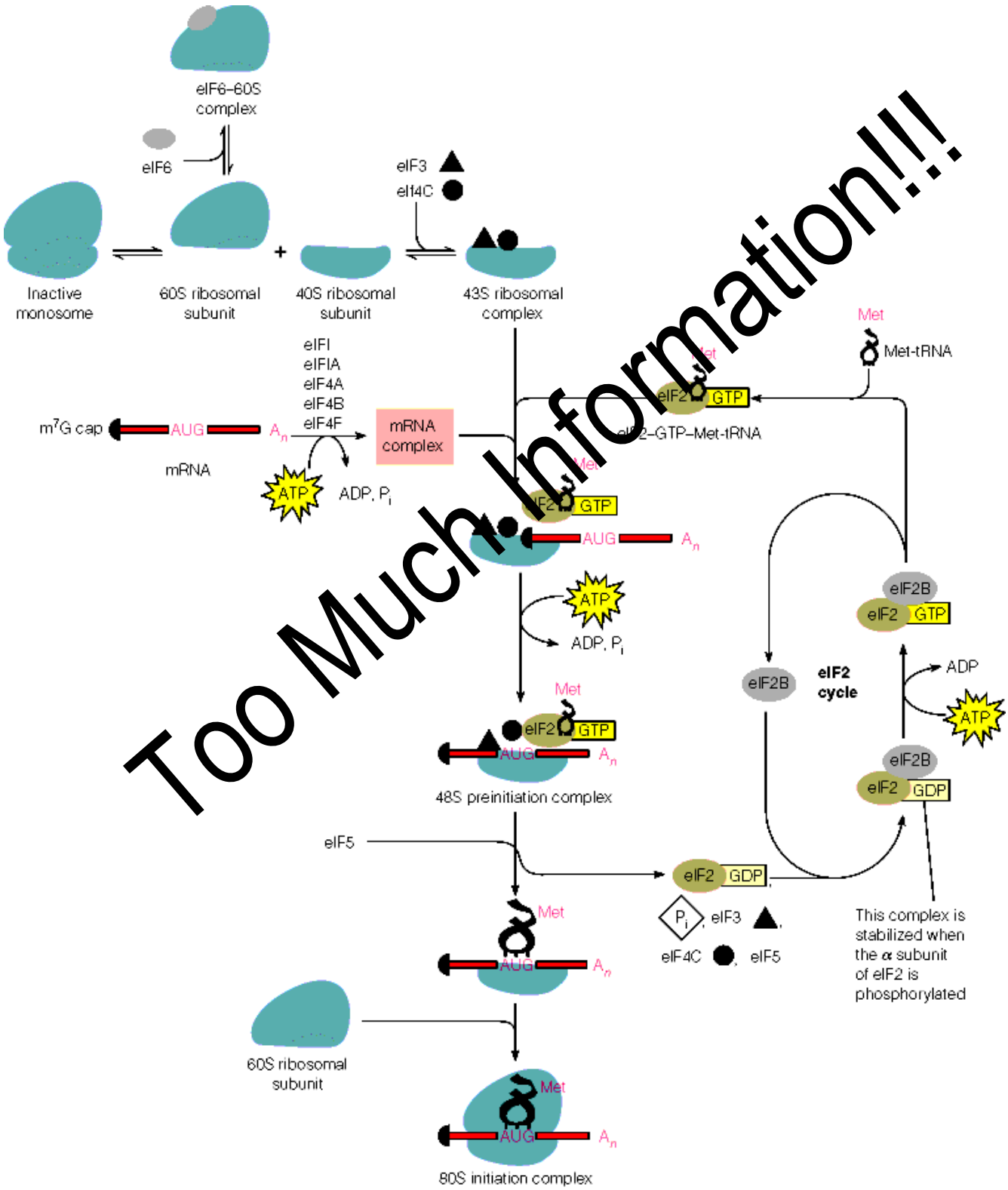


The Genetic Code:

- Based on triplets(codons)
- Codons:
  - no gaps between them
  - are non-overlapping
  - unambiguous (ie. a codon codes for a single amino acid)
- **The code is degenerate**:
  - most amino acids are specified by more than one codon
  - all of the degeneracy is found at the 3rd codon position (Wobble position)
- The code is “universal” (though there are exceptions)

Great implications for genetic engineering !!! Why ?

Eukaryotic Translation:



Translation, the important factoids….

The Mechanics are similar for Prokaryotes and Eukaryotes:

- mRNAs are read from 5' to 3'
- Proteins are synthesized from Amino- to Carboxy- terminus

- **Machinery Required**
  - mRNA
  - amino acids, tRNAs & aminoacyl tRNA synthetases
  - ribosomes & other protein factors
  - an initiator tRNA
  - Energy !!! (GTP & ATP)

- **Initiation**
  - detection of translation start site
  - assembly of the ribosome on the mRNA
  - initiator tRNA-met is brought in
- **Elongation**
  - charged aminoacyl-tRNA brought in
  - new peptide bond gets made
  - ribosome translocates down three positions along mRNA
  - cycle repeats until stop codon is reached
- **Termination**
  - binding of release factors
  - release of the completed polypeptide
  - disassembly of the ribosome from the mRNA

Regulation of Translation

- Ribosome binding is a rate-limiting step
- Ability to initiate translation
- mRNA stability (short in Prokaryotes vs long in Eukaryotes)
- Secondary structure in the 5' UTR can inhibit translation
- Rate of translation
- Rate of re-initiation
- Presence of binding sites for proteins that can block translation
    - Iron Response Element (IRE) & IRE-BP
- Use of "Rare Codons"
    - most organisms do not encode every possible tRNA gene in their genomes
        - certain codons are not used at all !!!! Either made in very low amounts, or not at all
            - **Codon Bias**
    - abundance of specific tRNAs (or of "charged" tRNAs) can influence translation efficiency
    - **Codon Adaptive Index**
        - implication for heterologous gene expression

	FBA1	ENO2	PYK1		FBA1	ENO2	PYK1
TTT	0	1	0	TCT	11	13	13
TTC	15	14	15	TCC	8	17	13
TTA	7	2	3	TCA	0	0	1
TTG	17	36	32	TCG	0	0	0
CTT	0	0	0	CCT	0	0	1
CTC	0	0	0	CCC	0	0	0
CTA	0	0	0	CCA	16	13	24
CTG	0	0	0	CCG	0	0	0
ATT	8	11	11	ACT	10	8	10
ATC	14	9	26	ACC	10	12	28
ATA	1	0	0	ACA	0	0	0
ATG	8	10	11	ACG	0	0	0

Proline codon bias in three different *C. albicans* genes

Protein Synthesis: Prokaryotes vs. Eukaryotes

- **Main differences:**
  - Eukaryotic ribosomes are larger
  - Prokaryotic mRNAs have significantly shorter half-lives (min vs hrs)
  - Eukaryotes require **many** more factors for initiation

  - **Prokaryotes:**
    - ribosomal binding results from 18S subunit binding to the Shine-Delgarno motif (Ribosomal binding site)
    - First MET is formylated
    - no mRNA modifications
    - no nucleus; translation is **concurrent** with transcription
    - polycistronic
  - **Eukaryotes:**
    - Ribosomal binding in Eukaryotes results from 5' Cap recognition
      - In a few cases, IRES (internal ribosome entry sites)
    - ribosomal scanning to correct ATG
      - Kozak motif - **(GCC)(GCC)GCCGCCATGG**
    - need splicing, 5' capping, poly-adenylation
    - done inside the nucleus (concurrent with transcription)
    - transcription in nucleus, translation outside nucleus
      - export mature mRNA before translation can happen
    - monocistronic (however, alternative splicing & polyadenylation)

Next Time:

- protein folding
- post-translational modification
- Targetting and degradation
- quick overview of basic protein structure

- Recombinant DNA methodologies