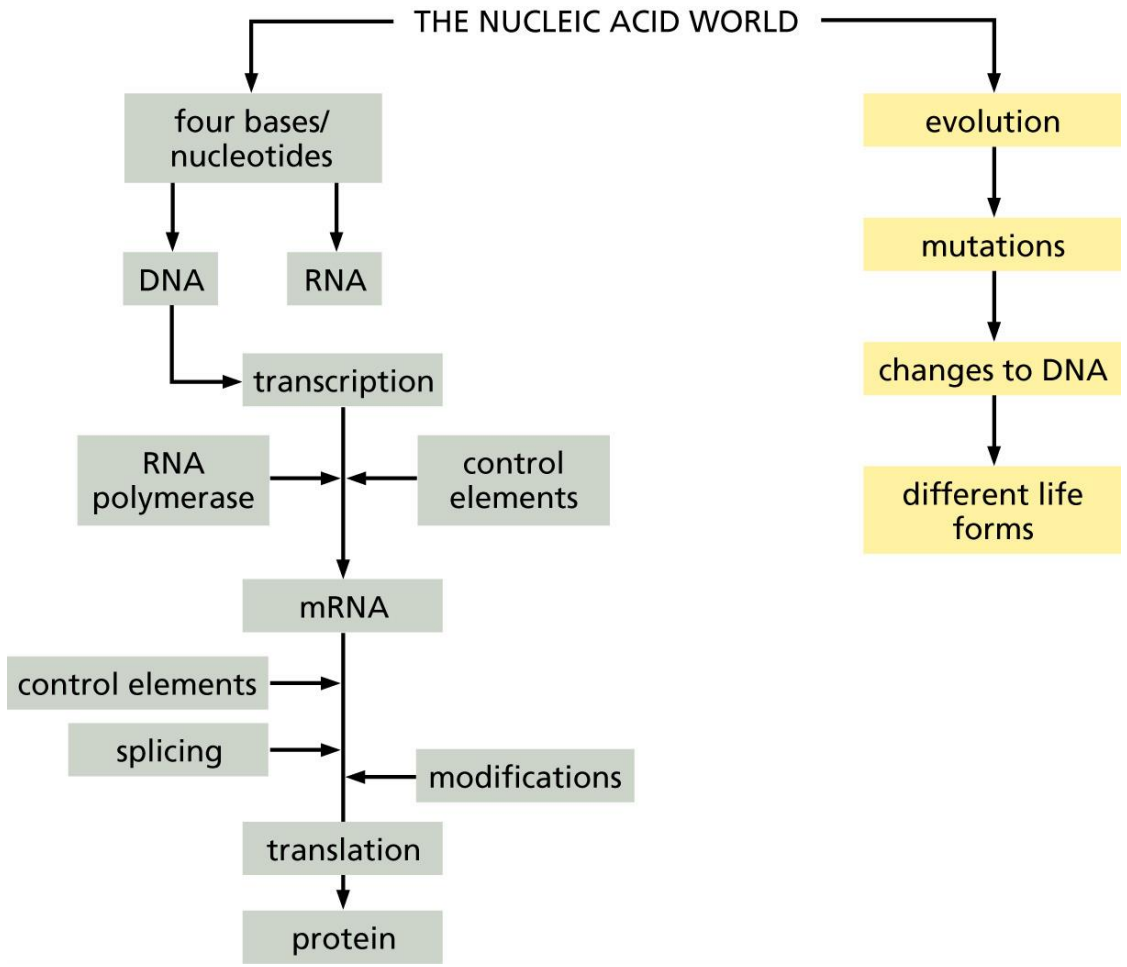
Molecular Biology Basics

by Ahmet Sacan

Images are from "Understanding Bioinformatics" and Wikipedia.

> Overview of Molecular Biology

- Cell
- Macromolecules
- Central Dogma



Living Organisms

- Entropy (a measure of disorder) always increases (2nd Law of Thermodynamics)
- Living organisms have low entropy, accomplished at the expense of increased entropy on the "outside"

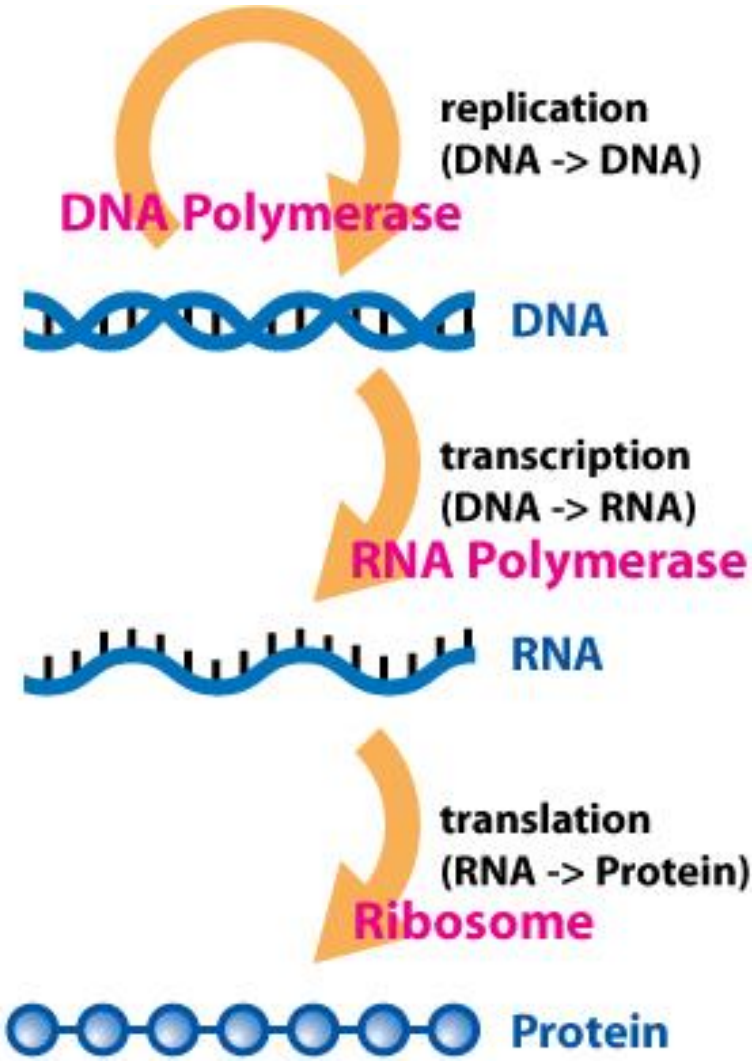
Living Organisms

- "Entropy fighting machines"
- Functions of life facilitate:
  - acquisition and
  - orderly expenditure of energy
- Tasks:
  - Gather energy from environment
  - Use energy to maintain inside/outside
  - Use energy to reproduce
  - Develop strategies to perform above tasks efficiently

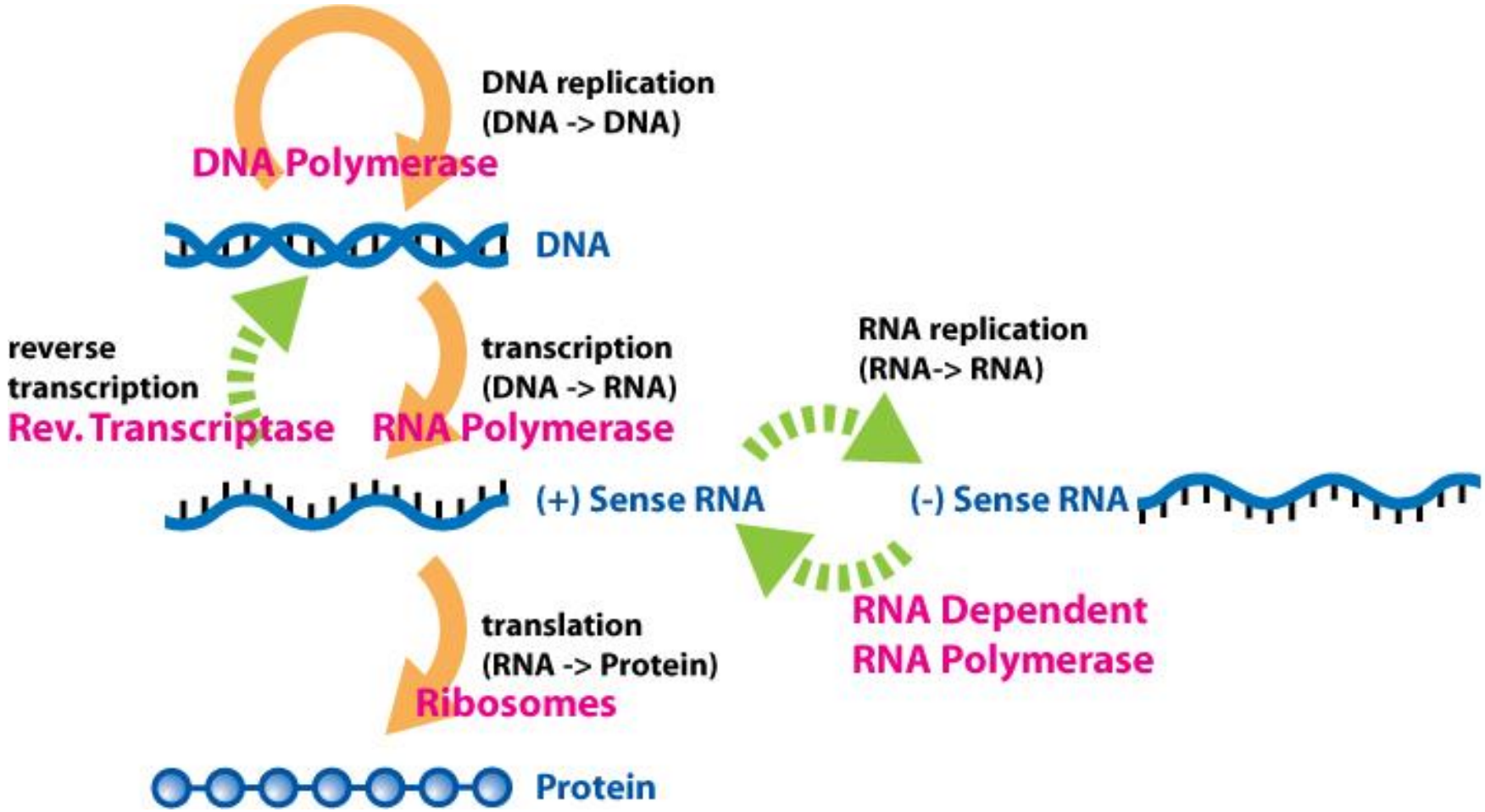
Macromolecules

- Lipids: Separate inside/outside
- Proteins, RNA: perform critical functions
- DNA: encode information about how/when to produce the above molecules

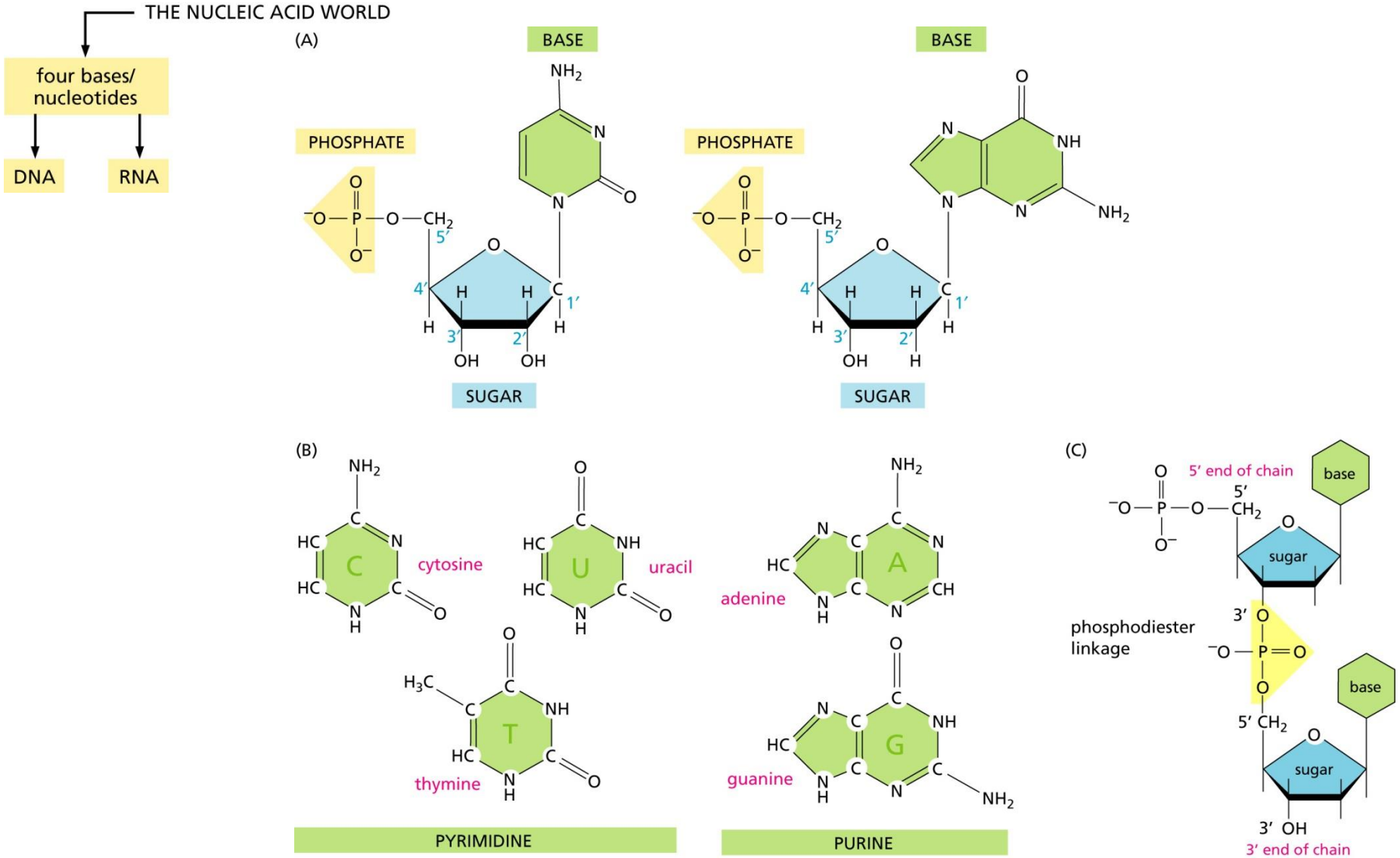
Central Dogma



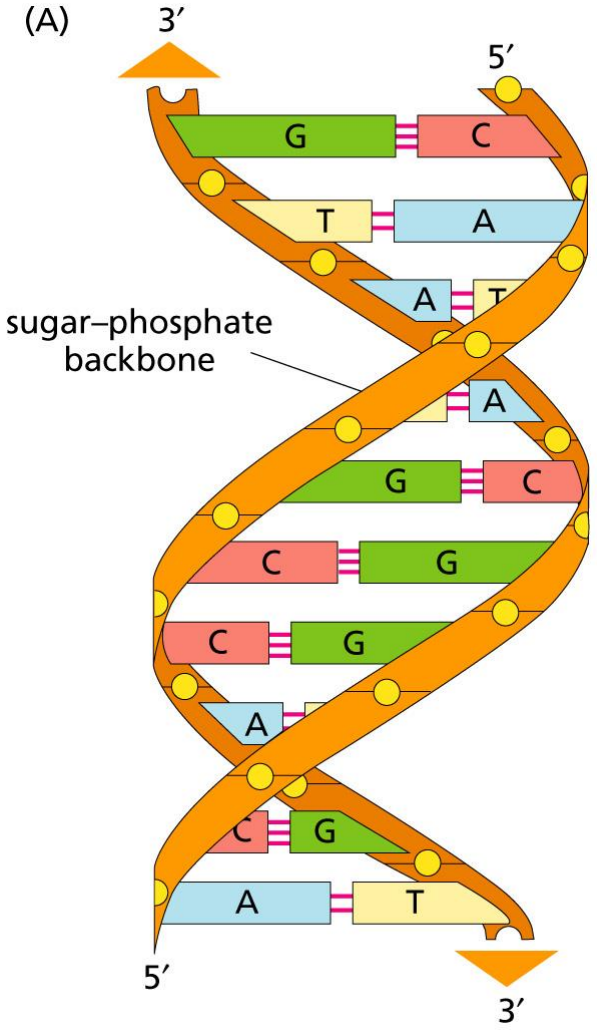
Central Dogma ++



Nucleotide monomers

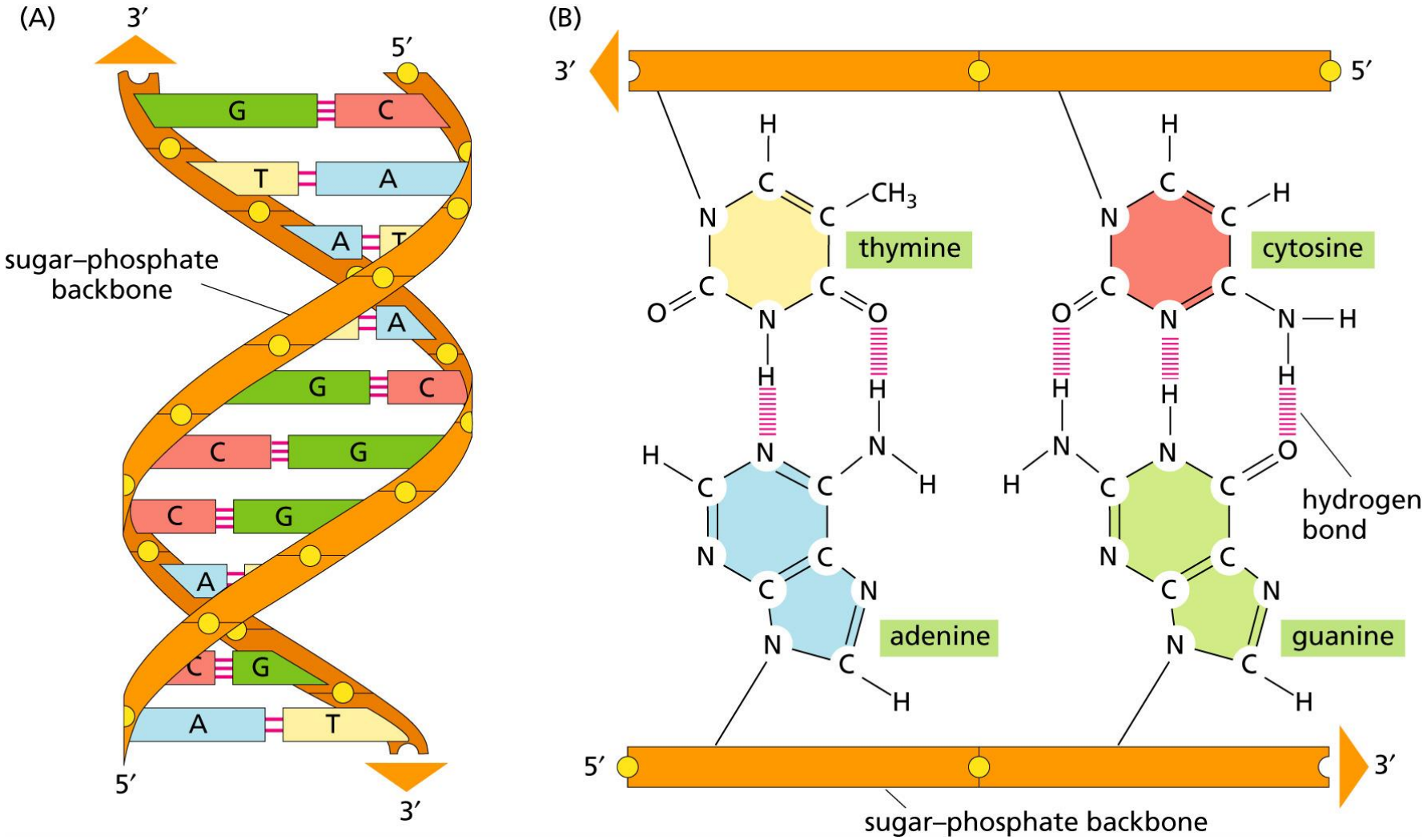


DNA

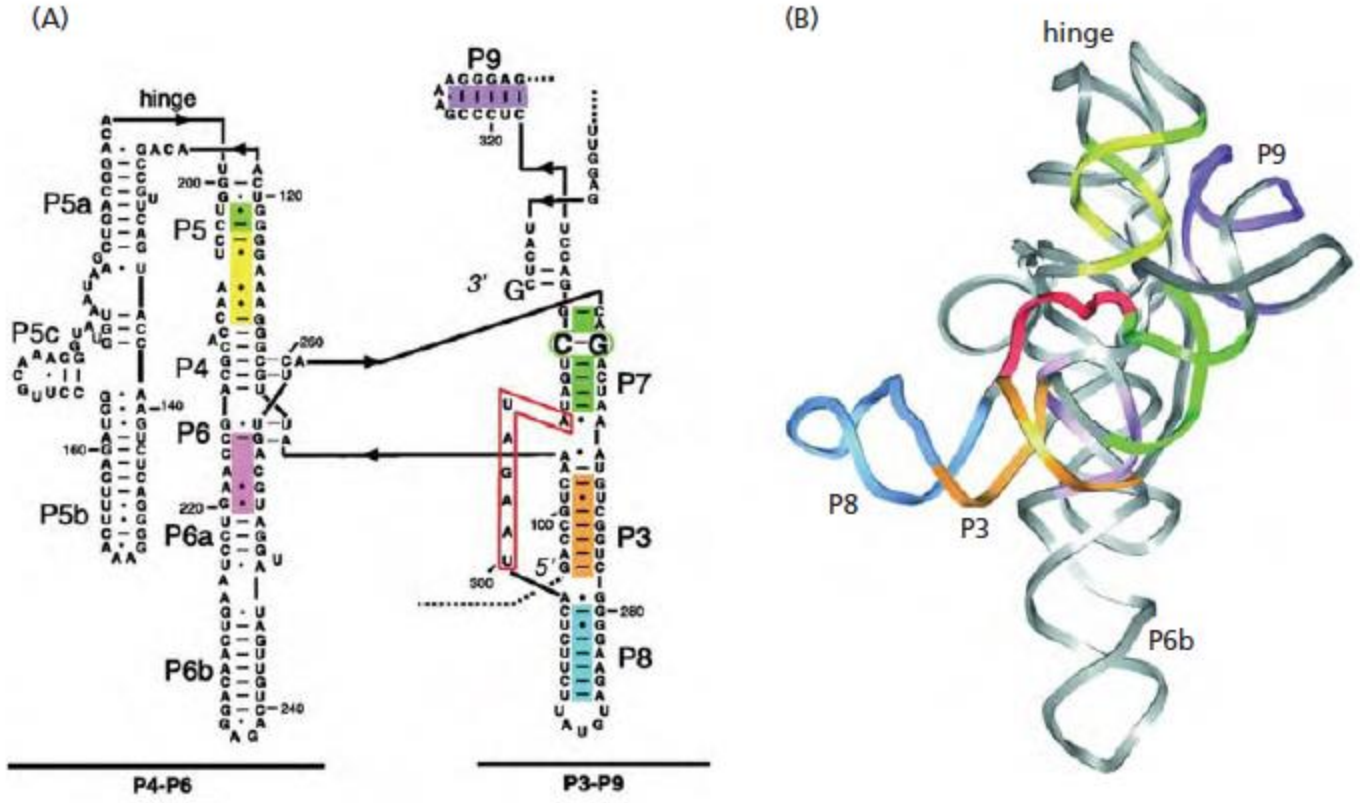


- *Watson-Crick base-pairing*

DNA

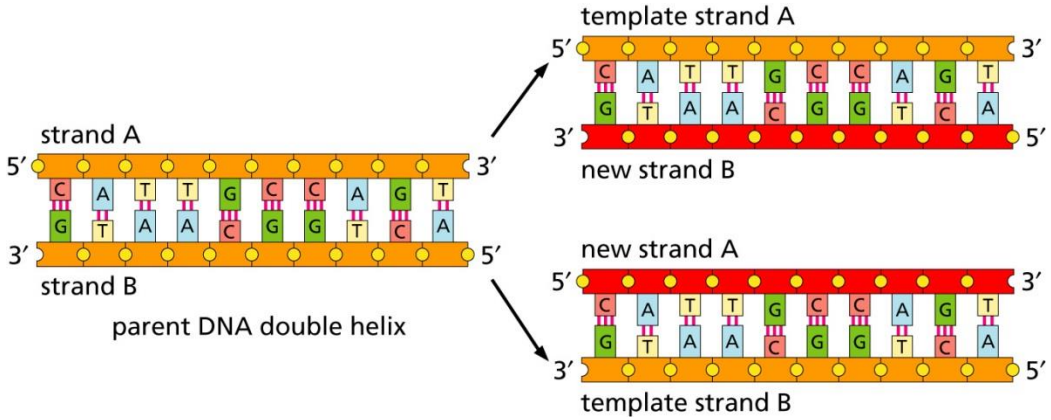

(A)
3′
5′
G
C
T
A
A
T
sugar–phosphate
backbone
A
G
C
C
G
C
G
A
C
G
A
T
5′
3′
(B)
3′
5′
H
C
CH3
N
C
thymine
C
O
C
N
O
H
H
N
H
H
N
N
H
C
C
N
C
N
C
N
adenine
H
H
C
N
cytosine
O
H
N
H
hydrogen
bond
guanine
5′
3′
sugar–phosphate backbone

RNA



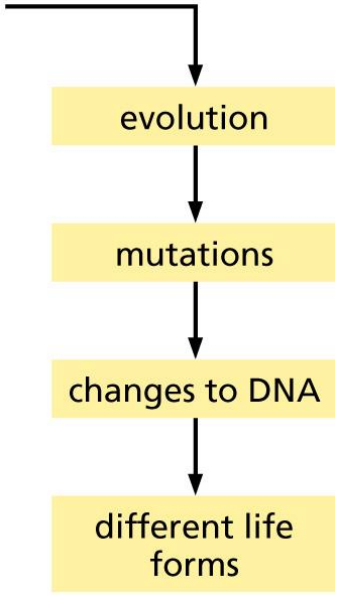
- Three dimensional structure of an RNA

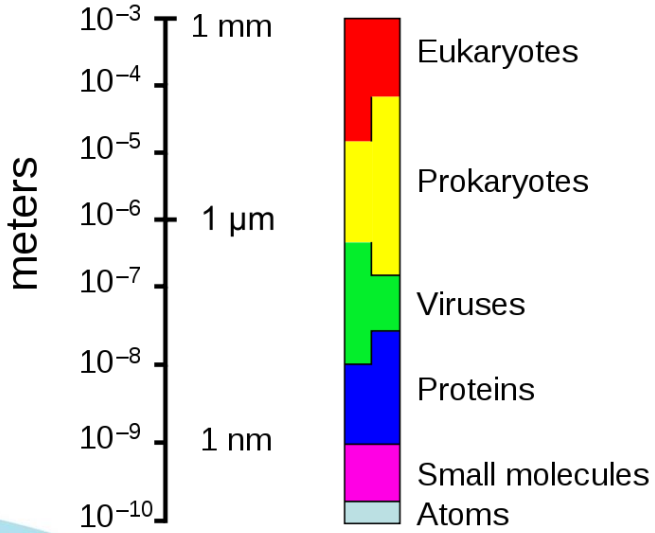
DNA Replication

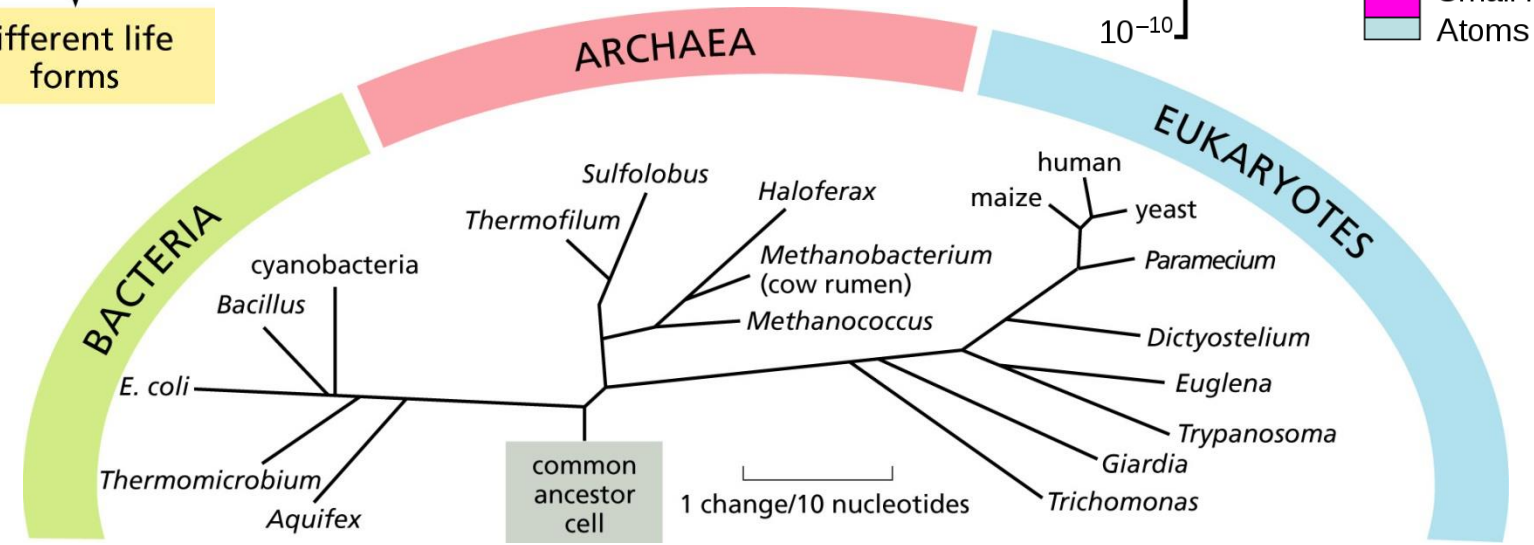


- Error rate: 1 in 10^9

Evolution

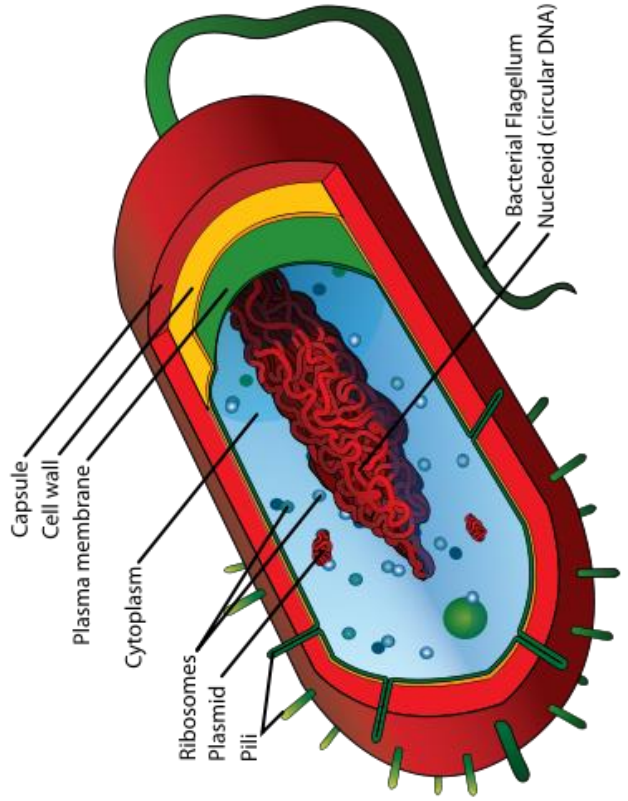

evolution
mutations
changes to DNA
different life forms


meters
10^{-3}
10^{-4}
10^{-5}
10^{-6}
10^{-7}
10^{-8}
10^{-9}
10^{-10}
1 mm
1 μm
1 nm
Eukaryotes
Prokaryotes
Viruses
Proteins
Small molecules
Atoms


ARCHAEA
BACTERIA
EUKARYOTES
Sulfolobus
Thermofilum
Haloferax
Methanobacterium
(cow rumen)
Methanococcus
cyanobacteria
Bacillus
E. coli
Thermomicrobium
Aquifex
common ancestor cell
1 change/10 nucleotides
human
maize
yeast
Paramecium
Dictyostelium
Euglena
Trypanosoma
Giardia
Trichomonas

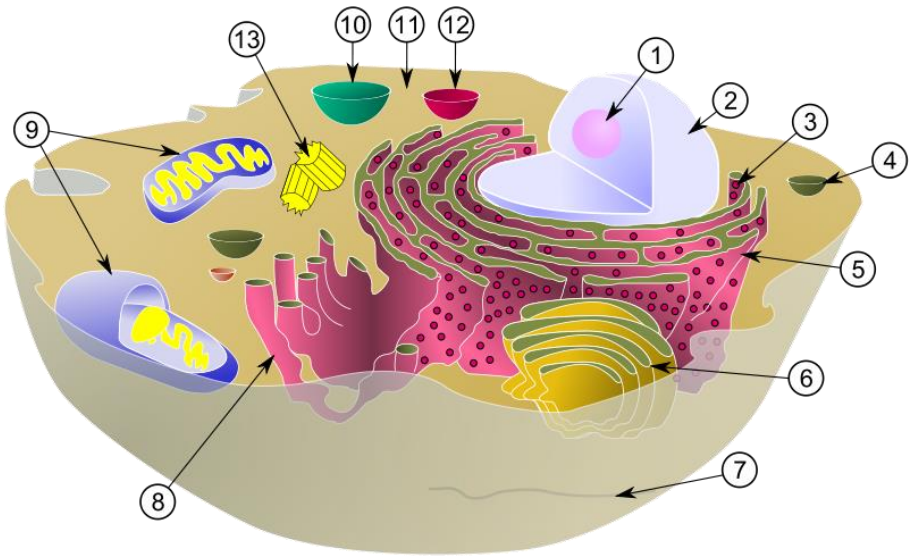
Prokaryotes

- No nucleus or organelles
- Single, circular chromosome
- Plasmids
  - Drug resistance
  - Rare compound metabolism
- Most studied: Escherichia coli (E. coli) - gut bacterium



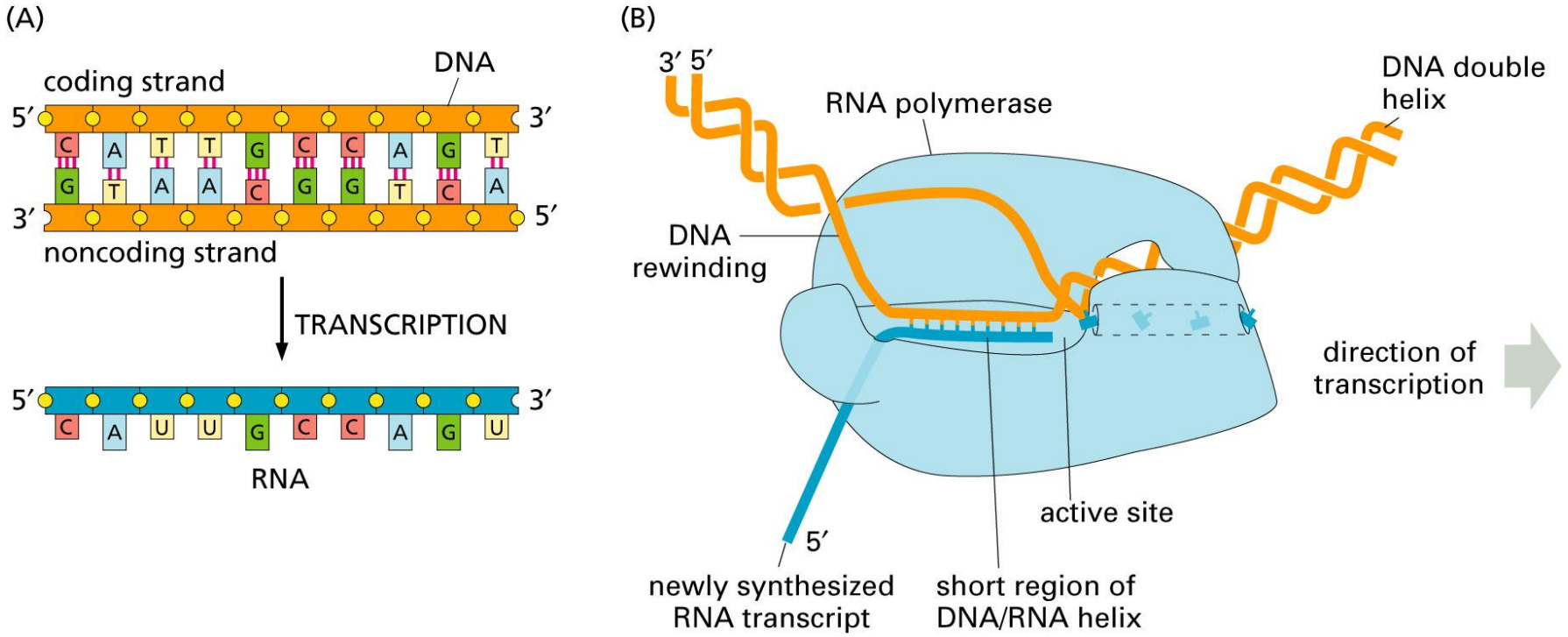
Eukaryotes

- Animals, plants, fungi, algae, protozoa
- Compartmentalized
  - nucleus, mitochondria, chloroplast
- Multiple linear chromosomes
  - Highly packed w/ histones



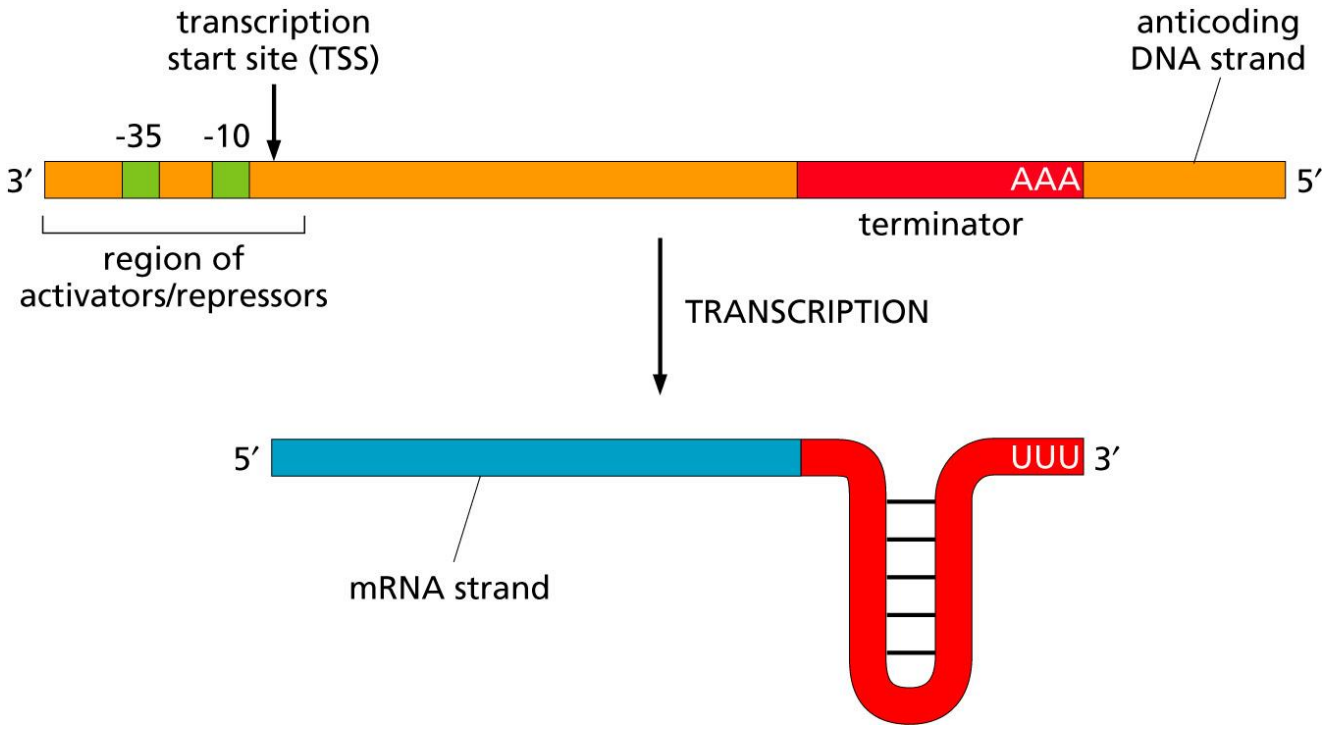
typical animal cell.
(1) nucleolus, (2) nucleus, (3) ribosome, (4) vesicle, (5) rough endoplasmic reticulum (ER), (6) Golgi apparatus, (7) Cytoskeleton, (8) smooth endoplasmic reticulum, (9) mitochondria, (10) vacuole, (11) cytoplasm, (12) lysosome, (13) centrioles within centrosome

Transcription

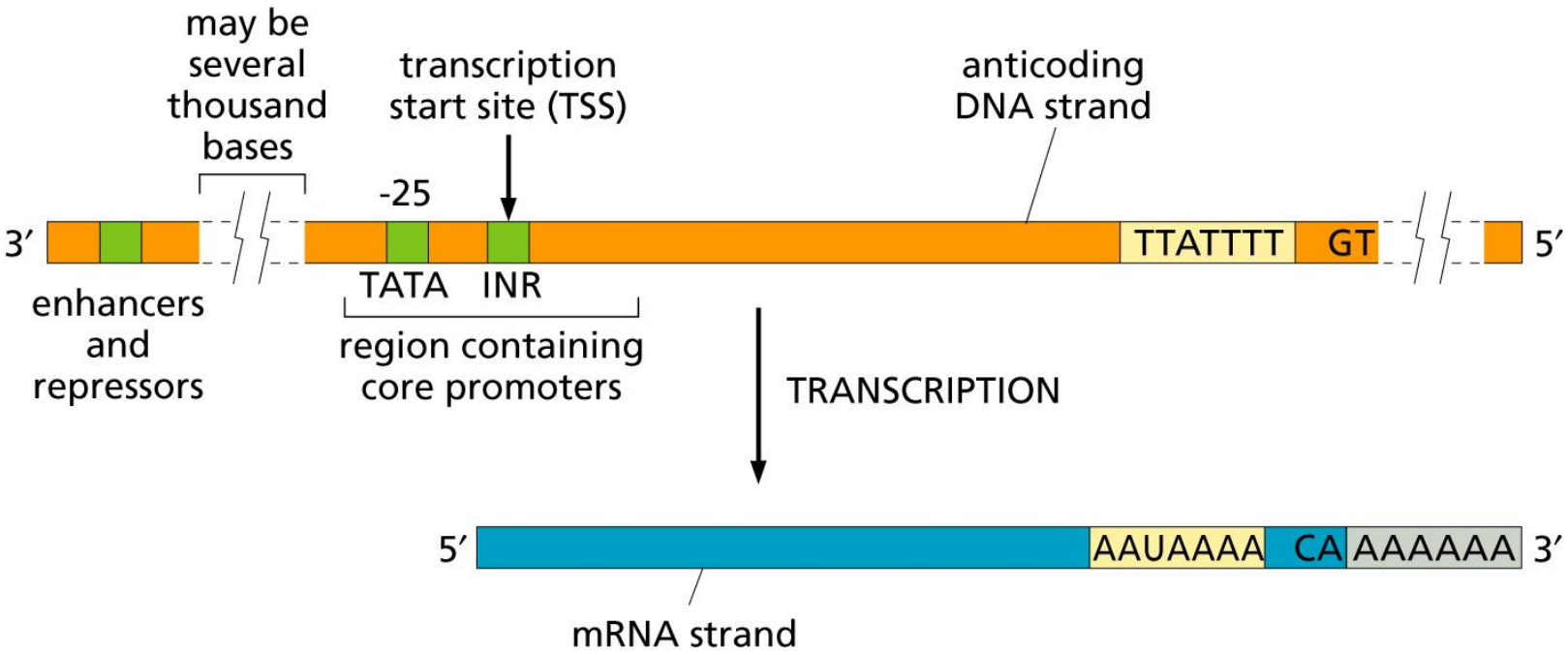


- Noncoding/anticoding/antisense strand serves as the physical template. Sequence of coding strand is identical to mRNA (except for T->U replacement).

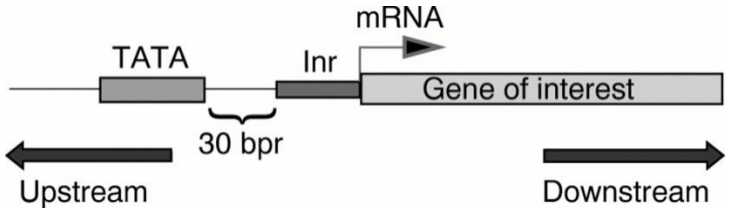
Prokaryotic gene



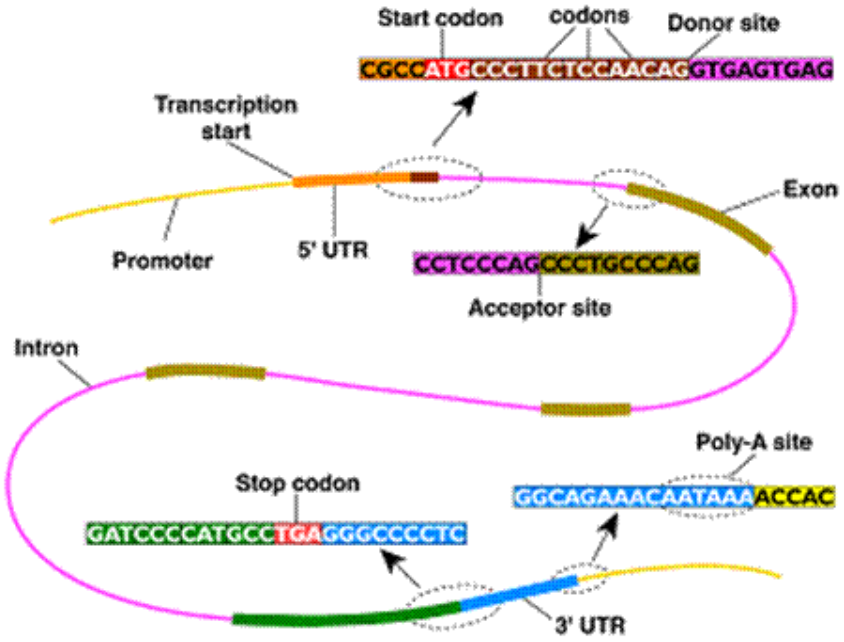
Eukaryotic gene



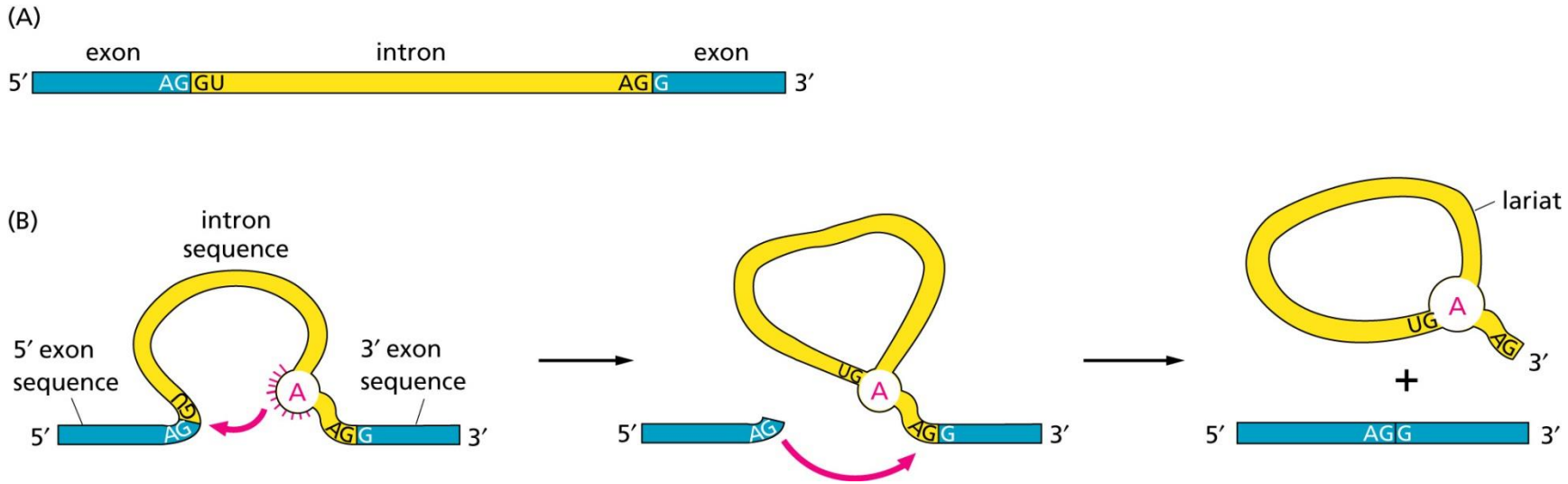
Eukaryotic gene



- Posttranscriptional modifications:
  - RNA capping (5'end)
  - Polyadenylation (~200xA at 3'end)
  - RNA splicing to remove introns - "alternative splicing"



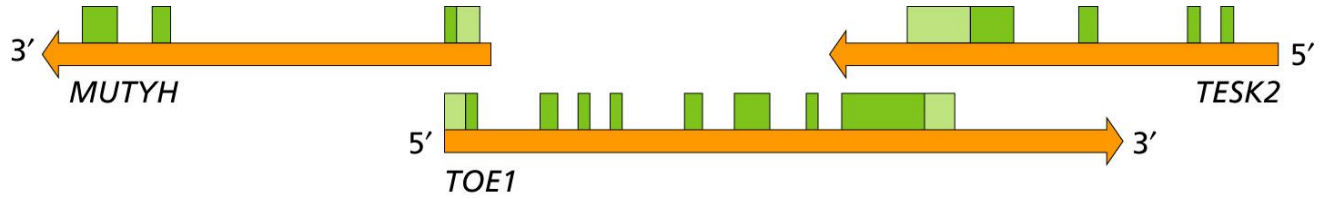
Eukaryotic mRNA

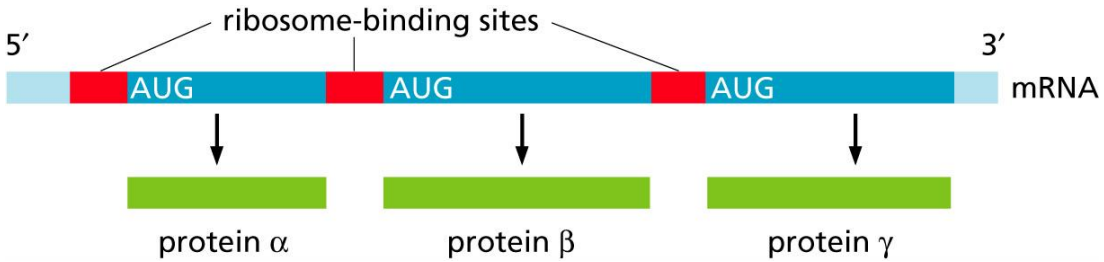


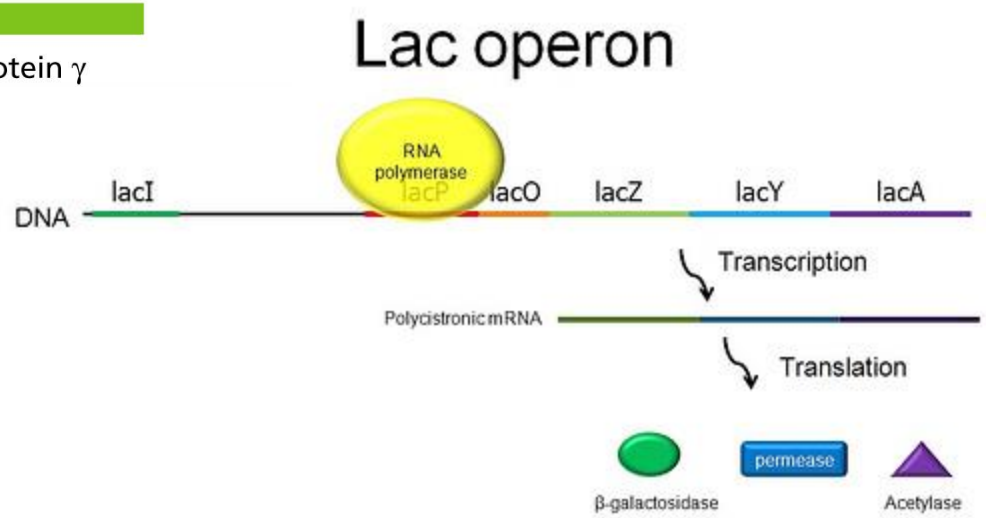
Open Reading Frame (ORF)

- 5'- CATACTAGATAATATTCGATTAAGC -3'
- Frame 1:
  - CAT ACT AGA TAA* TAT TCG ATT AAG
- Frame 2:
  - ATA CTA GAT AAT ATT CGA TTA AGC
- Frame 3:
  - TAC TAG* ATA ATA TTC GAT TAA*

Overlapping ORFs, Operons

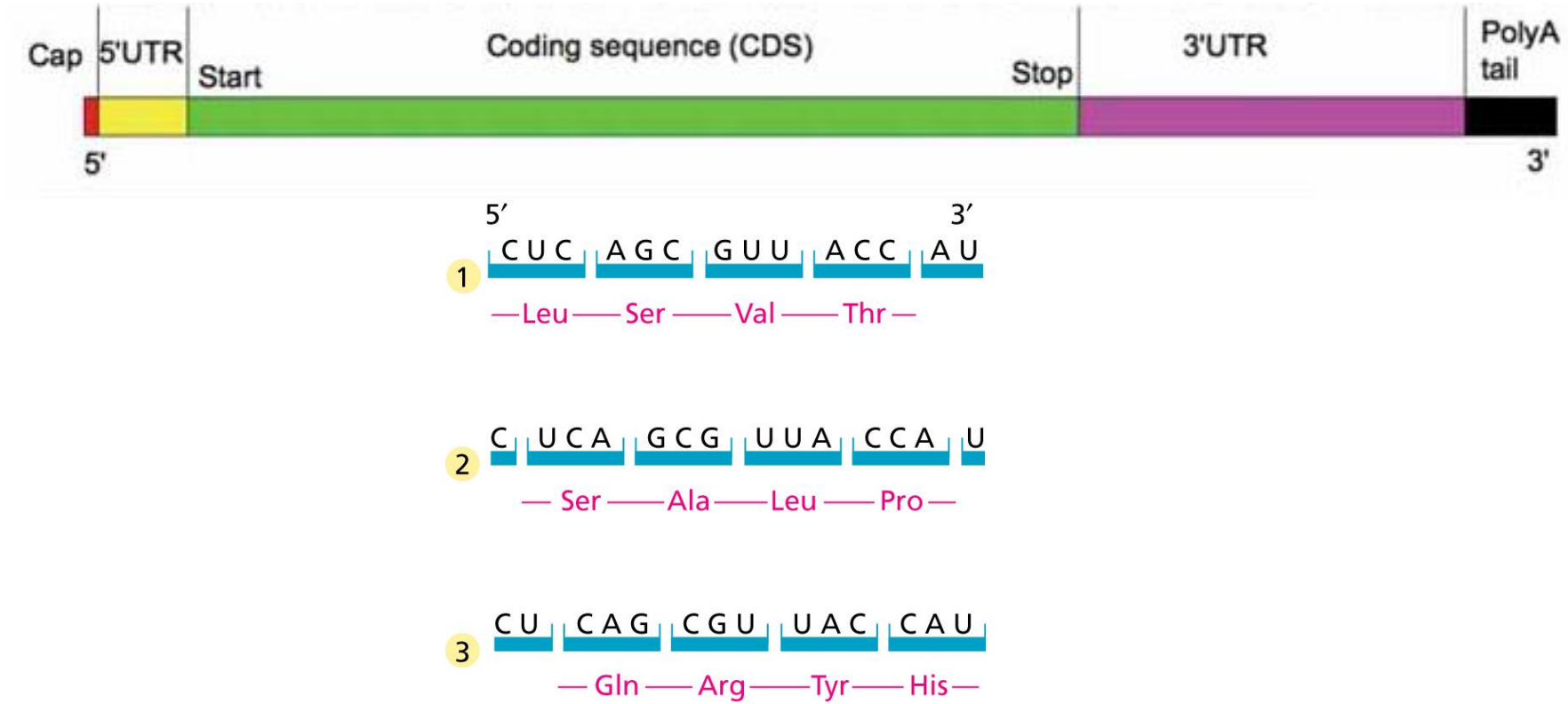






mRNA

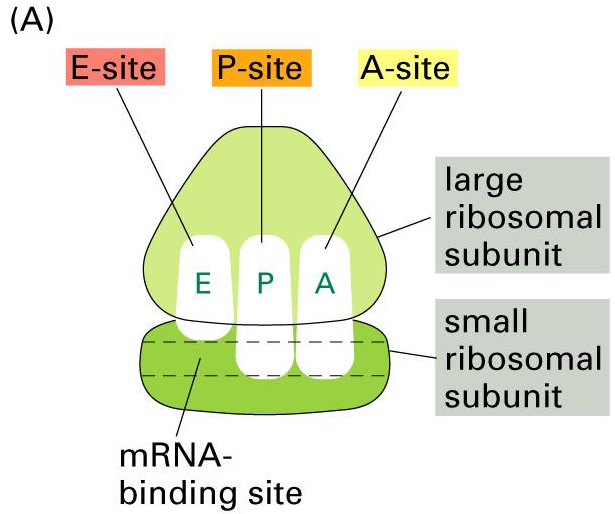
- Coding Sequence: codons

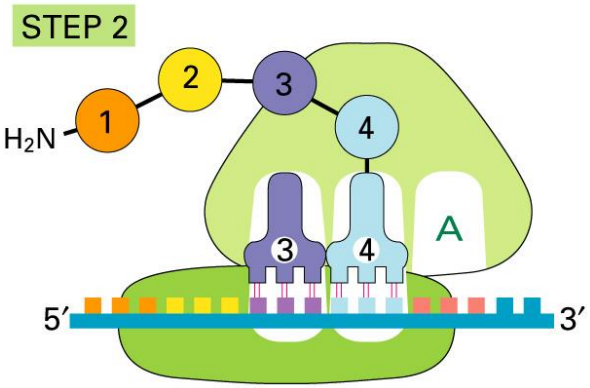


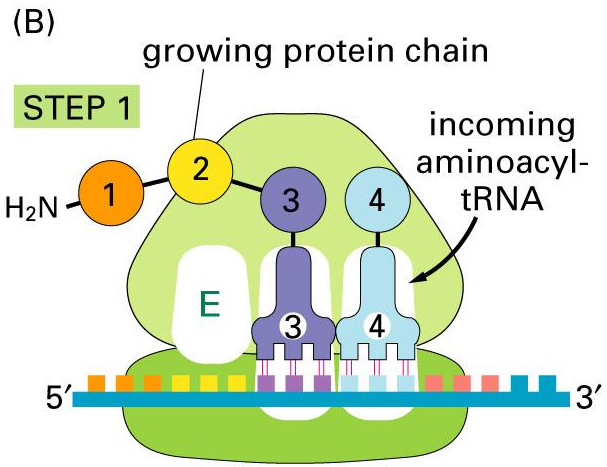
Translation initiation

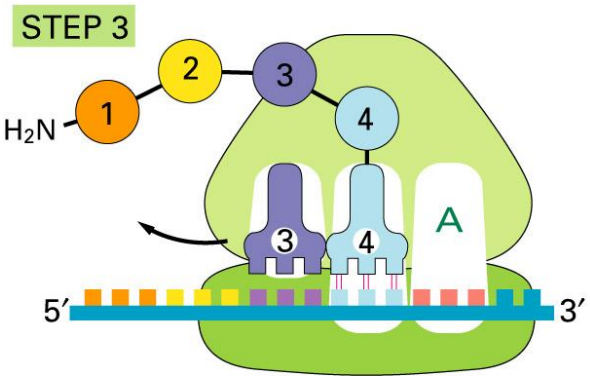
- Bacteria:
  - **Shine–Dalgarno sequence: a consensus** *AGGAGGU* that appear upstream of *AUG*.
- Eukaryotes:
  - Ribosomal components bind 7-methylguanosine nucleotide at the 5′ end and scan for an *AUG* codon.

Translation

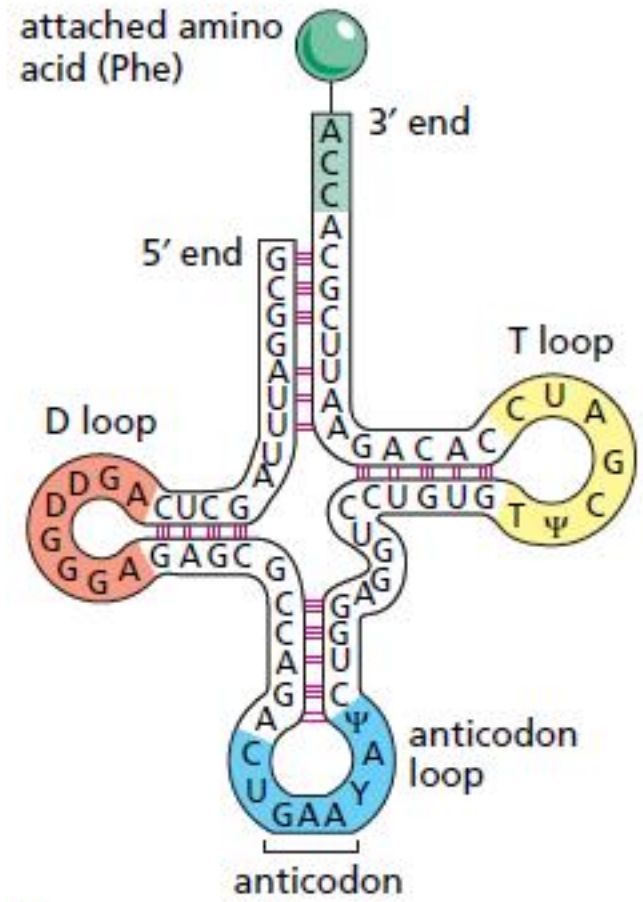








tRNA

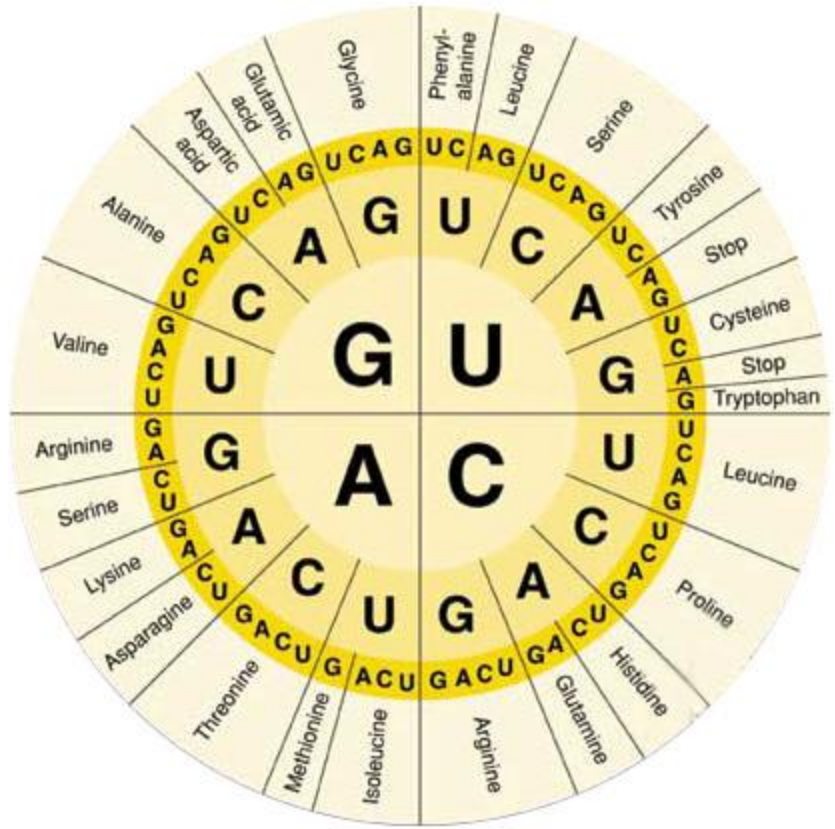

attached amino acid (Phe)
3′ end
5′ end
T loop
D loop
anticodon loop
anticodon

(A)

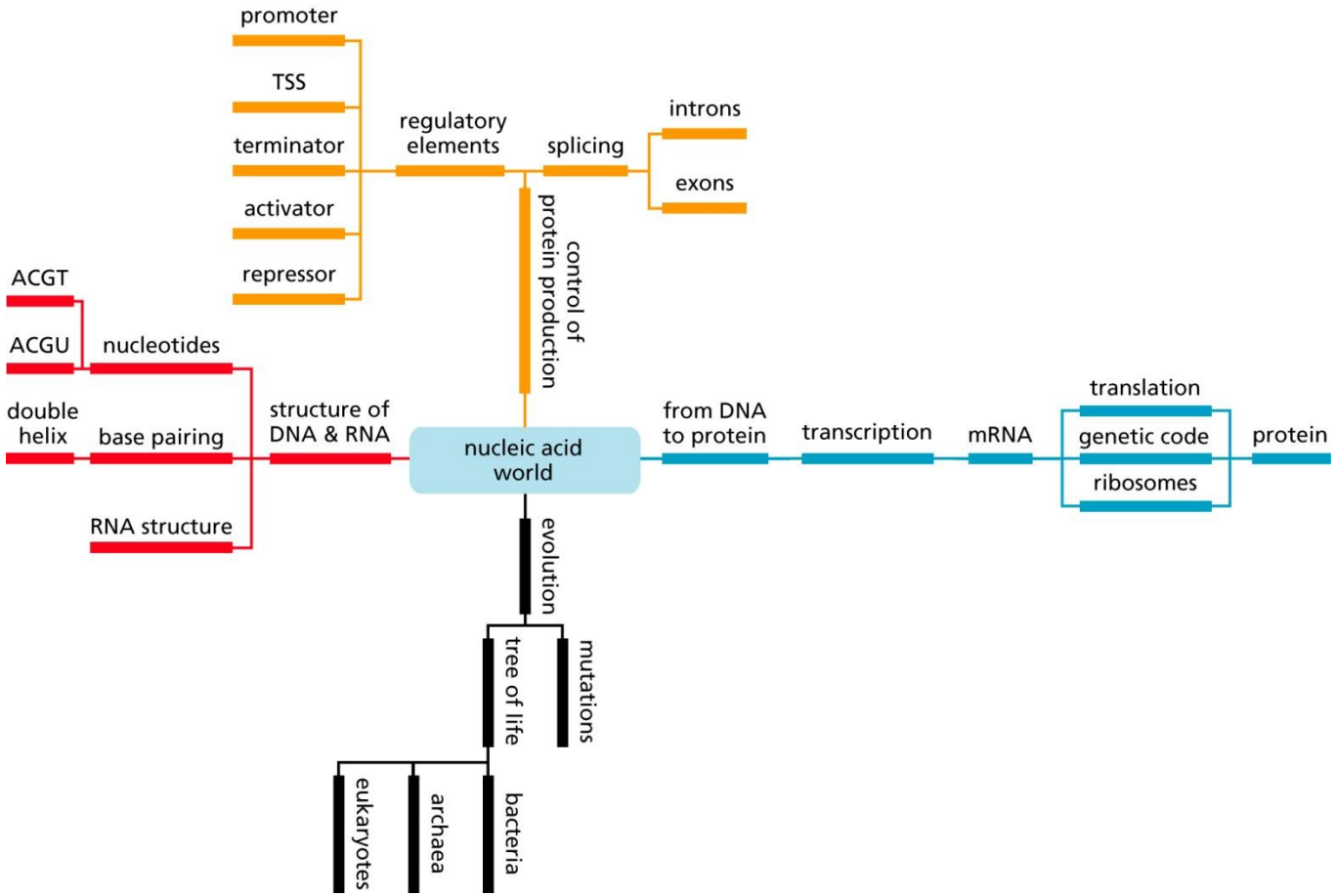
(B)

The Genetic Code

First letter	Second letter: U	Second letter: C	Second letter: A	Second letter: G	Third letter
U	UUU Phe	UCU Ser	UAU Tyr	UGU Cys	U
	UUC Phe	UCC Ser	UAC Tyr	UGC Cys	C
	UUA Leu	UCA Ser	UAA Stop	UGA Stop	A
	UUG Leu	UCG Ser	UAG Stop	UGG Trp	G
C	CUU Leu	CCU Pro	CAU His	CGU Arg	U
	CUC Leu	CCC Pro	CAC His	CGC Arg	C
	CUA Leu	CCA Pro	CAA Gln	CGA Arg	A
	CUG Leu	CCG Pro	CAG Gln	CGG Arg	G
A	AUU Ile	ACU Thr	AAU Asn	AGU Ser	U
	AUC Ile	ACC Thr	AAC Asn	AGC Ser	C
	AUA Ile	ACA Thr	AAA Lys	AGA Arg	A
	AUG Met	ACG Thr	AAG Lys	AGG Arg	G
G	GUU Val	GCU Ala	GAU Asp	GGU Gly	U
	GUC Val	GCC Ala	GAC Asp	GGC Gly	C
	GUA Val	GCA Ala	GAA Glu	GGA Gly	A
	GUG Val	GCG Ala	GAG Glu	GGG Gly	G



Summary


promoter
TSS
terminator
activator
repressor
regulatory elements
splicing
introns
exons
control of protein production
ACGT
ACGU
nucleotides
double helix
base pairing
structure of DNA & RNA
RNA structure
nucleic acid world
from DNA to protein
transcription
mRNA
translation
genetic code
ribosomes
protein
evolution
tree of life
mutations
eukaryotes
archaea
bacteria