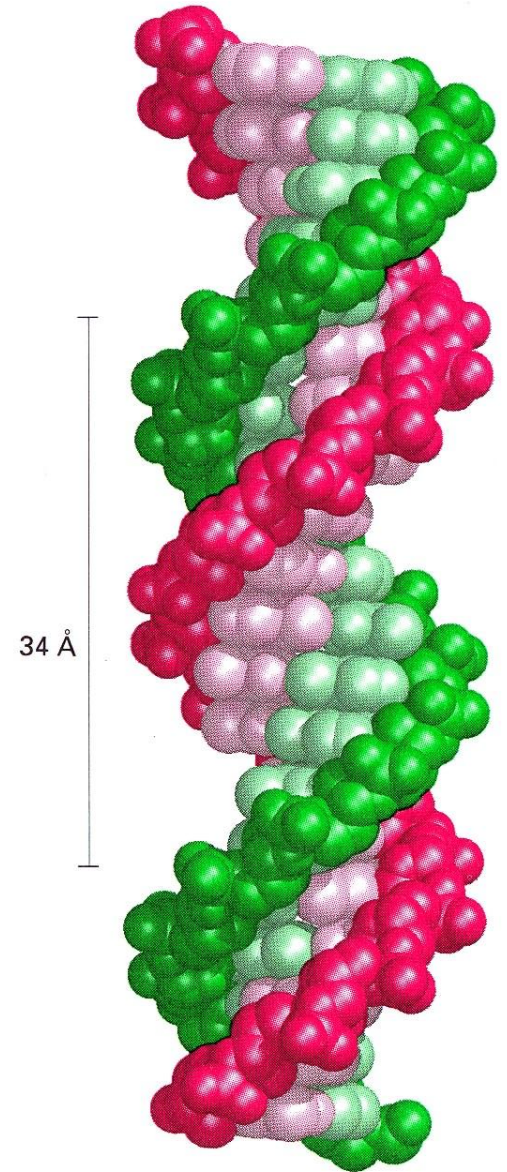




# The Introduction to Molecular Biology

**Jianbin Ye PhD**

**Email: [happye1986@gmail.com](mailto:happye1986@gmail.com)**





# Chemical substance

## OUTLINE

- **History of molecular biology**
- **DNA as the genetic material**
- **RNA as intermediate**
- **Protein as workers**



# Brief History

## **What is molecular biology?**

The attempt to understand biological phenomena in molecular terms

The study of gene structure and function at the molecular level

**Molecular biology is a melding of aspects of genetics and biochemistry**



**Molecular biology is the study of genes and their activities at the molecular level, including transcription, translation, DNA replication, recombination and translocation.**

**--- Robert Weaver**

**Molecular biology seeks to explain the relationships between the structure and function of biological molecules and how these relationships contribute to the operation and control of biochemical processes.**

**---Turner et al.**



# Development in History

- **Stage 1:** Preparation
- **Stage 2:** Establishing and developing of contemporary Molecular Biology
- **Stage 3:** Primary understanding of the nature of life and modification of life-deep development stage
- **Stage 4:** Post-genome era



# Stage 1

**From late 1900s to early 1950s**

- **Confirmed that protein is one of the fundamental materials of life.**
- **Confirmed that DNA is the nature of genetic material.**



**By definition, the early work on genes cannot be considered molecular biology, or even molecular genetics, because early geneticists did not know the molecular nature of genes.**

**In fact, the chemical composition of genes was not known until 1944. At that point, it became possible to study genes as molecules, and the discipline of molecular biology was born.**



## **Stage 2**

**From early 1950s to early 1970s**

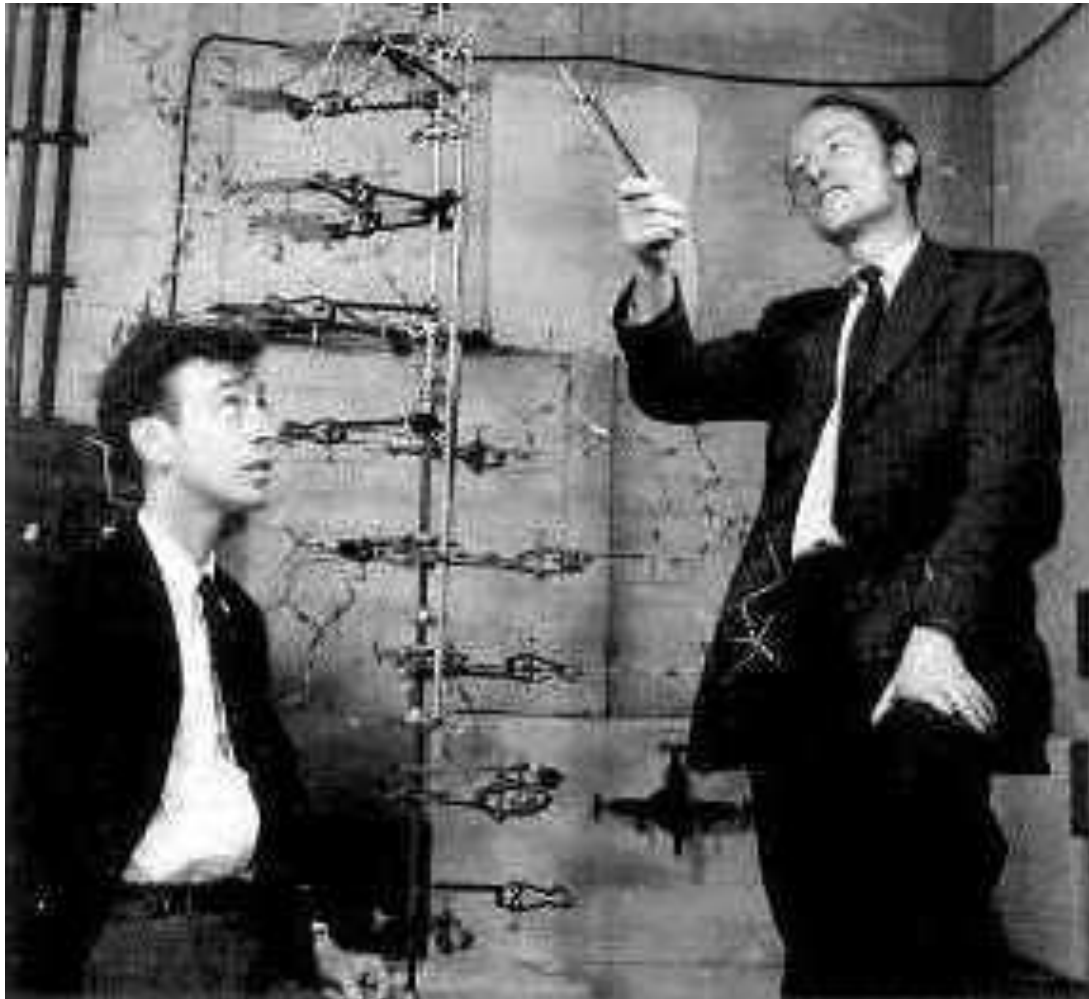
- **Central Dogma**
- **Further understanding on protein structure and function**





# The DNA Double Helix (1953)

The foundation of molecular biology

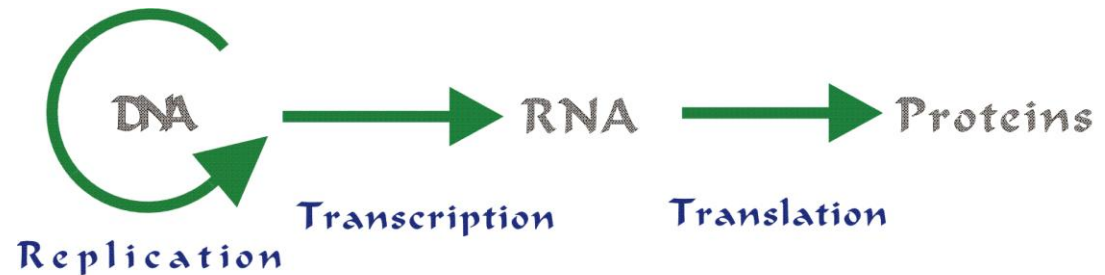


**Francis H.  
Crick**

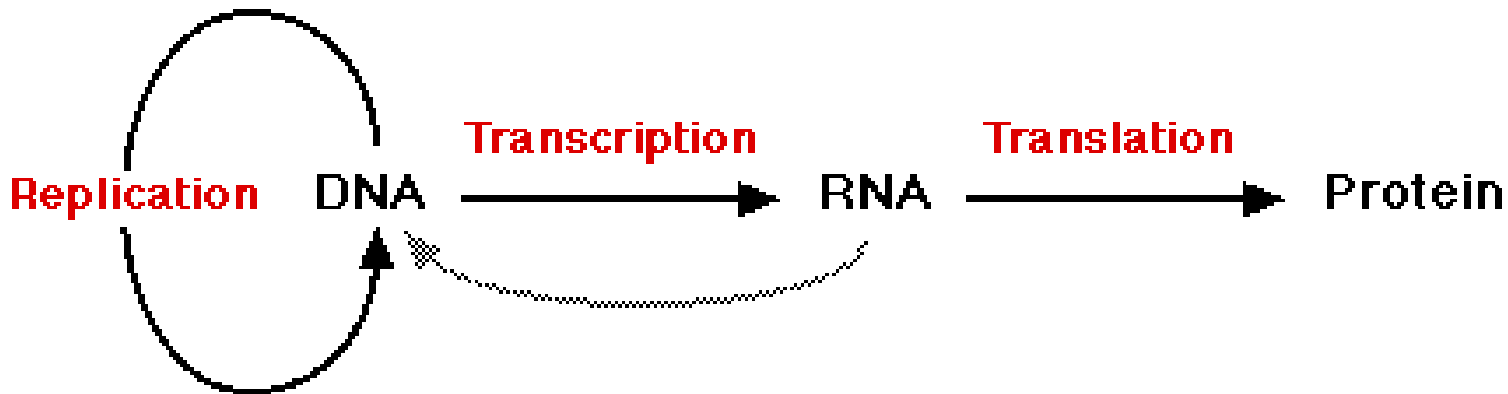
**James D.  
Watson**



# Dogma



## The Central Dogma of Molecular Biology



**Transcription** is carried out by **RNA polymerase**

**Translation** is performed on **ribosomes**

**Replication** is carried out by **DNA polymerase**

Reverse transcriptase copies RNA into DNA



## Stage 3

**From late 1970s to 2000s**

- **DNA recombination**
- **Genomics**
- **Monoclonal antibody**
- **Regulation of gene expression**
- **Mechanisms of signal transduction**



# Stage 4

## **Genomics**

**structural genomics**  
**functional genomics**  
**environmental genomics**  
**nutritional genomics**  
**drug genomics**  
**cancer genomics**  
**toxicological genomics**

## **Transcriptomics**

## **Proteomics**

## **Metabolomics**



# DNA as Genetic Material

## Historical Background

Miescher isolated nuclei from pus (white blood cells) in 1869

Found a novel phosphorus-bearing substance = nuclein

Nuclein is mostly chromatin, a complex of DNA and chromosomal proteins

End of 19<sup>th</sup> century – DNA and RNA separated from proteins

Levene, Jacobs, et al. characterized basic composition of DNA and RNA

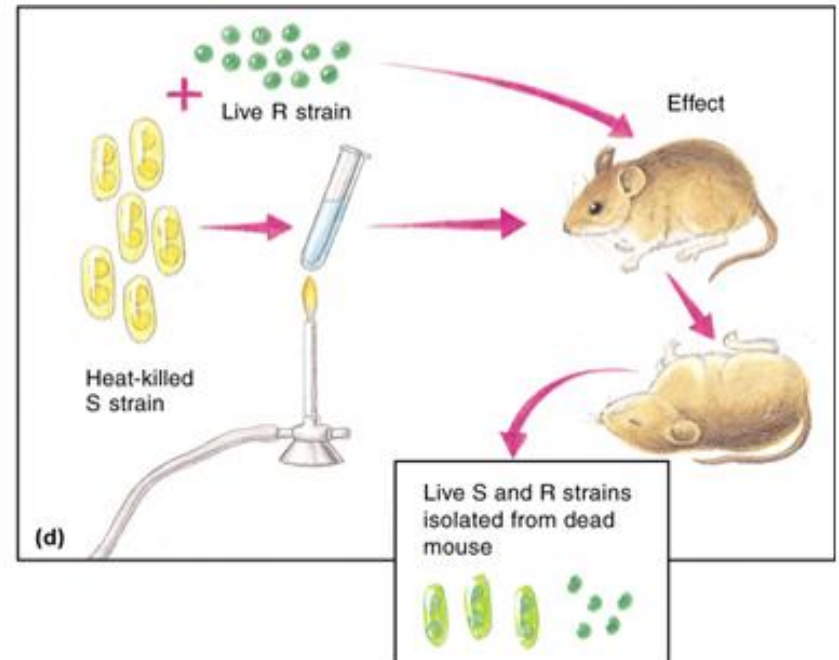
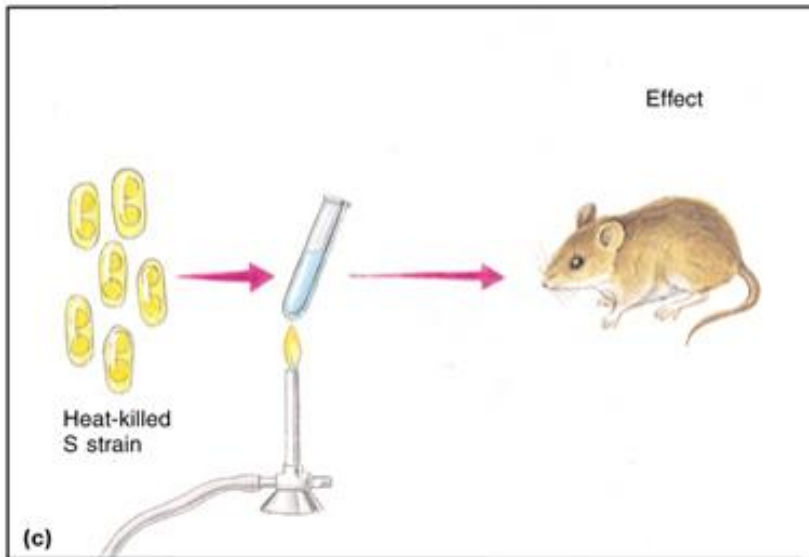
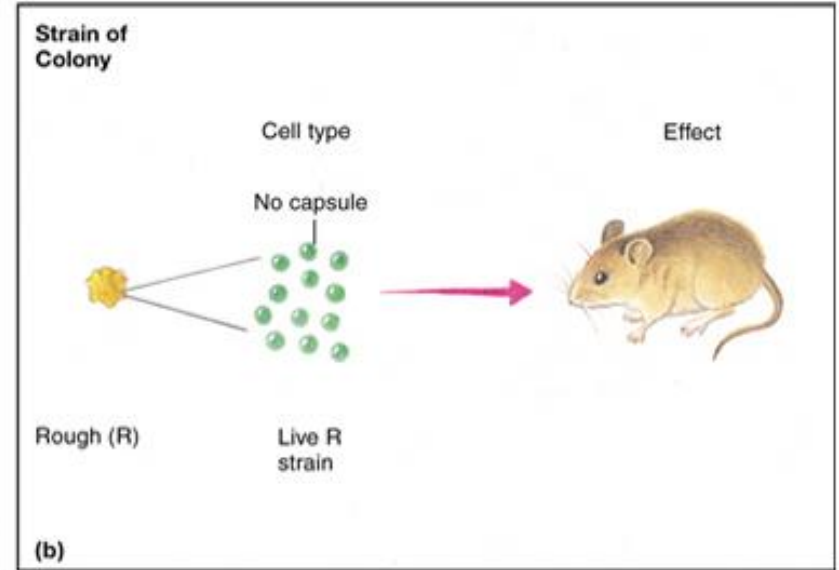
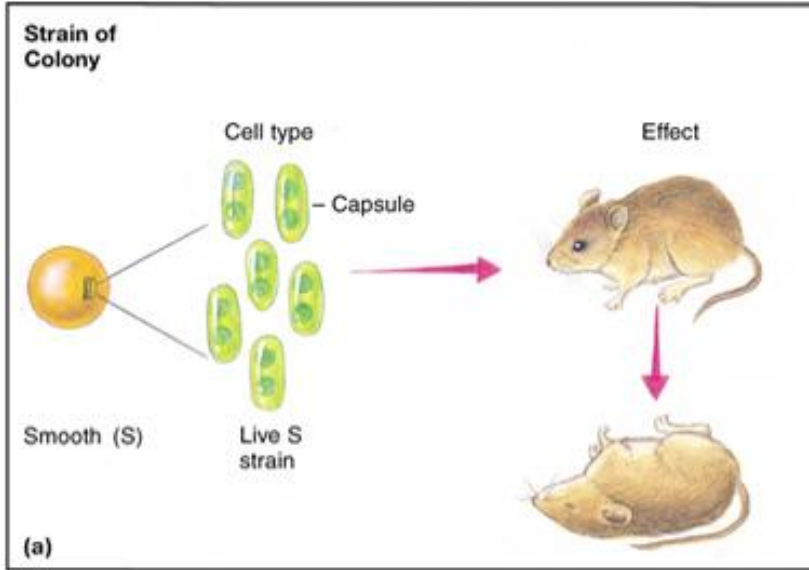


# Transformation in Bacteria

- Key experiments done by Frederick Griffith in 1928
- Observed change in *Streptococcus pneumoniae* — from virulent (S) smooth colonies where bacterial had capsules, to avirulent (R) rough colonies without capsules
- Heat-killed virulent colonies could transform avirulent colonies to virulent ones

# Griffith's Transformation Experiments

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.





# DNA: The Transforming Material

In 1944 a group used a transformation test similar to Griffith's procedure taking care to define the chemical nature of the transforming substance

- Techniques used excluded both protein and RNA as the chemical agent of transformation
- Other treatments verified that DNA is the chemical agent of transformation of *S. pneumoniae* from avirulent to virulent





# Analytical Tools

Physical-chemical analysis has often used:

1. Ultracentrifugation

Used to estimate size of material

2. Electrophoresis

Indicated high charge-to-mass ratio

3. Ultraviolet Absorption Spectrophotometry

Absorbance of UV light matched that of DNA

4. Elementary Chemical Analysis

Nitrogen-to-phosphorus ratio of 1.67, not found in protein

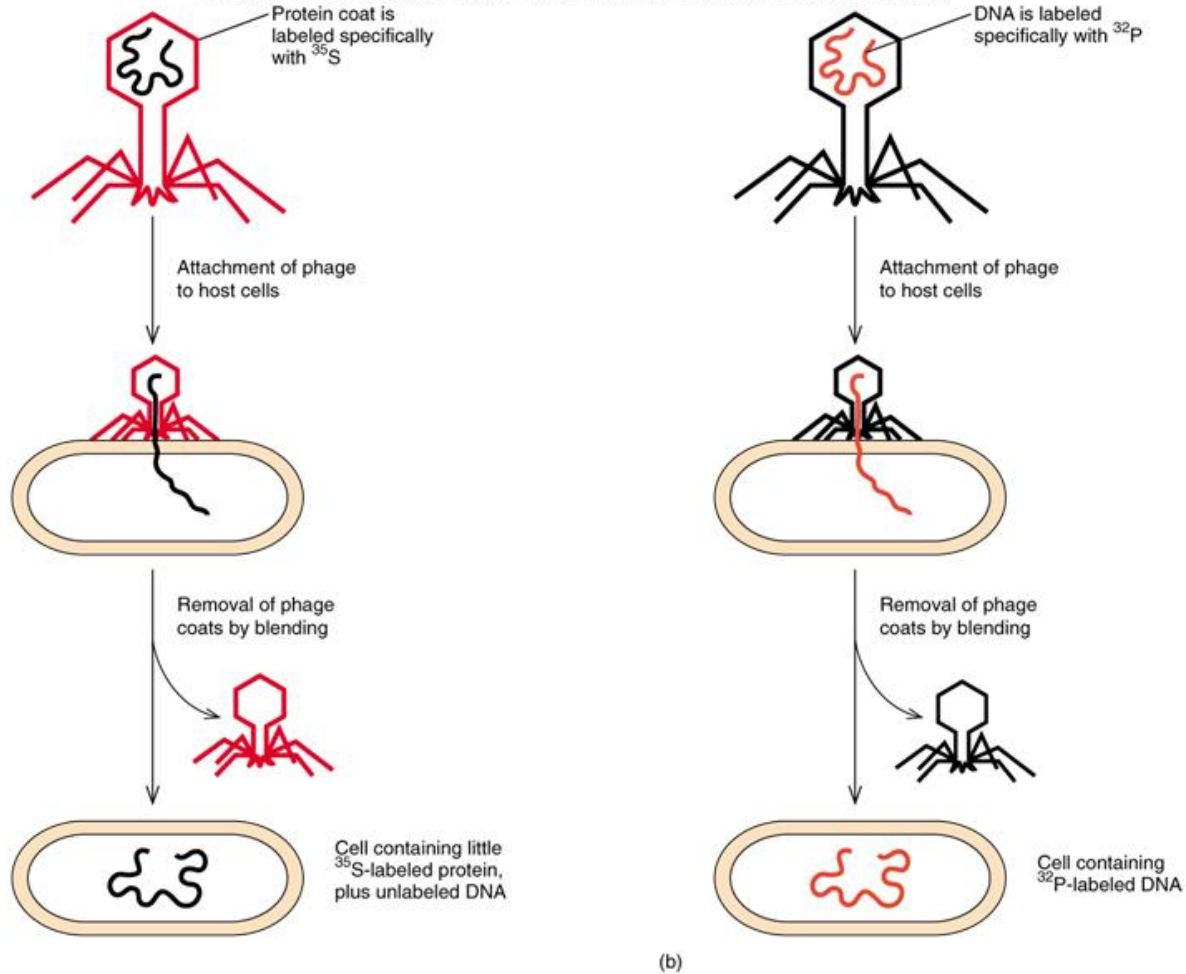


# DNA Confirmation

- In 1940s geneticists doubted use of DNA as it appeared to be monotonous repeats of 4 bases
- By 1953 Watson & Crick published the double-helical model of DNA structure and Chargaff had shown that the 4 bases were not present in equal proportions
- Hershey and Chase demonstrated that bacteriophage infection comes from DNA

# Procedure for the Hershey-Chase Transformation Experiments

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.





# Summary

- Genes are made of nucleic acid, usually DNA
- Some simple genetic systems such as viruses have RNA genes



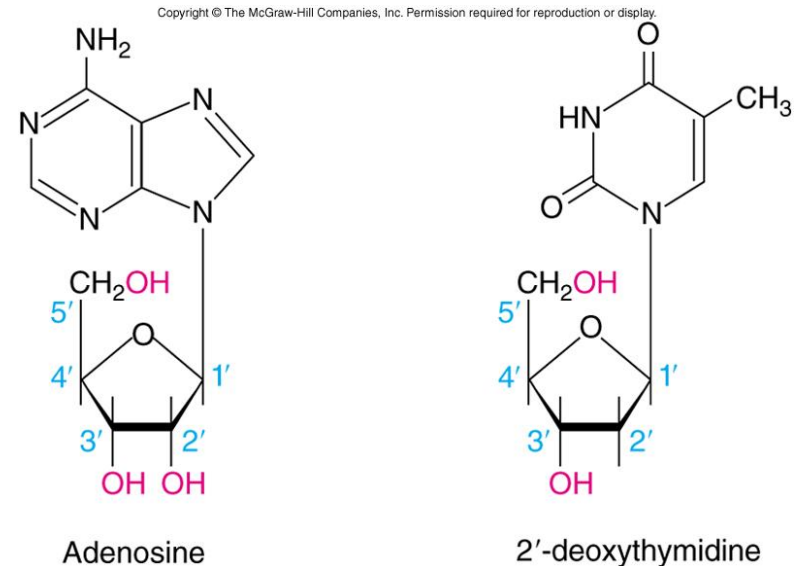
# The Chemical Nature of Polynucleotides

- Biochemists determined the components of nucleotides during the 1940s
- The component parts of DNA
  - Nitrogenous bases:
    - Adenine (A)
    - Cytosine (C)
    - Guanine (G)
    - Thymine (T)
  - Phosphoric acid
  - Deoxyribose sugar



# Nucleotides and Nucleosides

- RNA component parts
  - Nitrogenous bases
    - Like DNA except Uracil (U) replaces Thymine
  - Phosphoric acid
  - Ribose sugar
- Bases use ordinary numbers
- Carbons in sugars are noted as primed numbers
- Nucleotides contain phosphoric acid
- Nucleosides lack the phosphoric acid

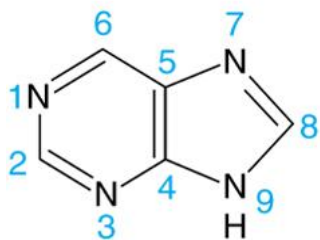




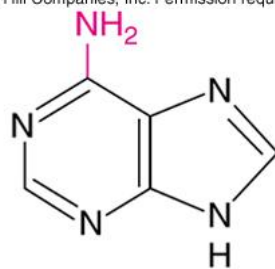
# Purines and Pyrimidines

- Adenine and guanine are related structurally to the parent molecule purine
- Cytosine, thymine and uracil resemble pyrimidine

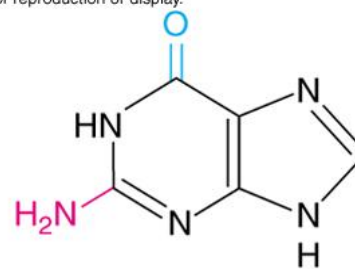
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



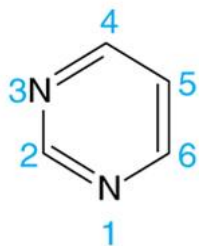
Purine



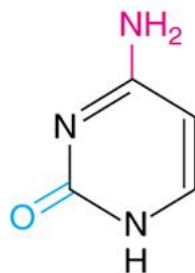
Adenine



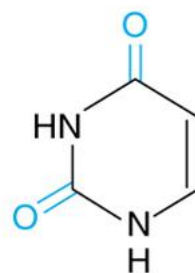
Guanine



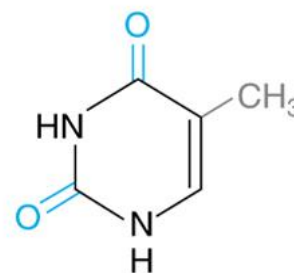
Pyrimidine



Cytosine



Uracil



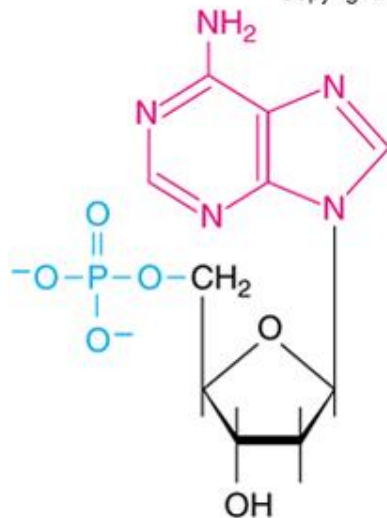
Thymine



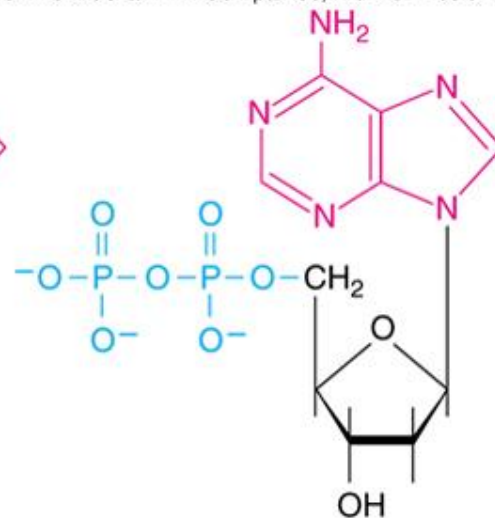
# DNA Linkage

- Nucleotides are nucleosides with a phosphate group attached through a phosphodiester bond
- Nucleotides may contain one, two, or even three phosphate groups linked in a chain

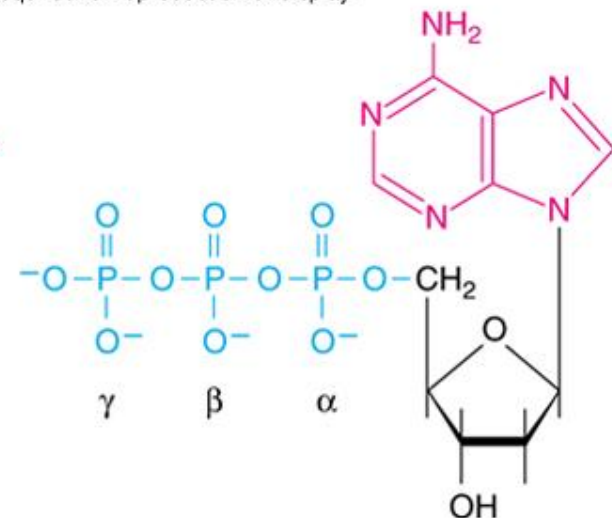
Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.



Deoxyadenosine-5'-  
monophosphate (dAMP)



Deoxyadenosine-5'-  
diphosphate (dADP)



Deoxyadenosine-5'-  
triphosphate (dATP)



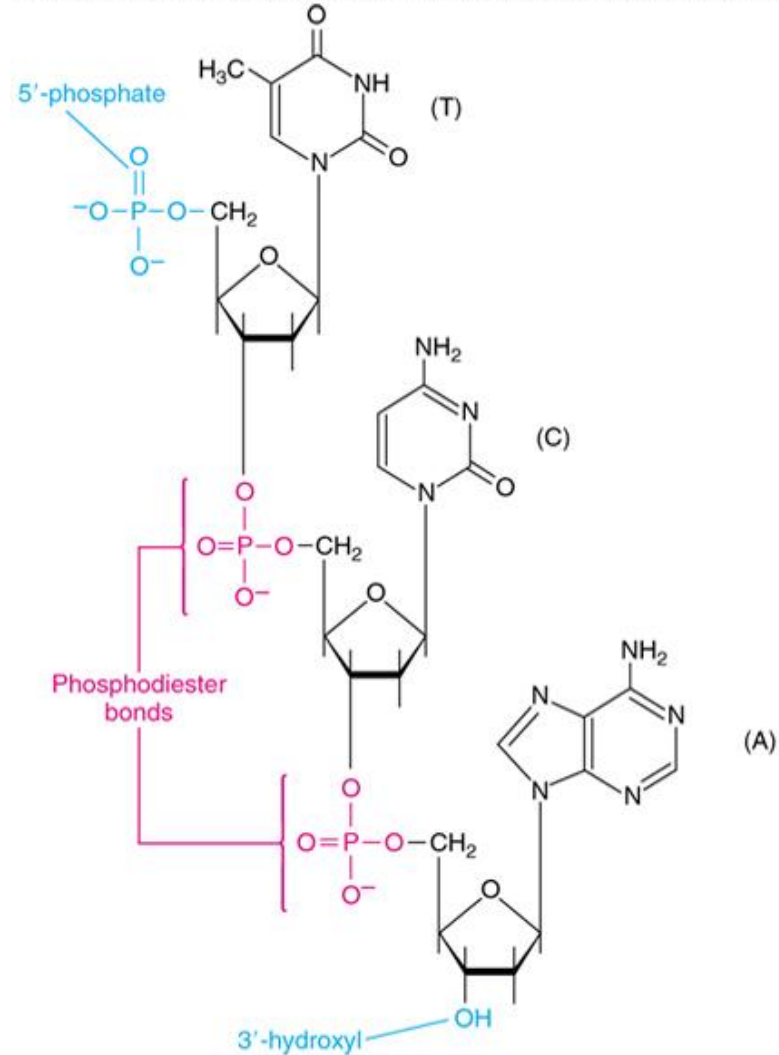


# A Trinucleotide

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

The example trinucleotide has polarity

- Top of molecule has a free 5'-phosphate group = **5' end**
- Bottom has a free 3'-hydroxyl group = **3' end**





# Summary

- DNA and RNA are chain-like molecules composed of subunits called nucleotides
- Nucleotides contain a base linked to the 1'-position of a sugar and a phosphate group
- Phosphate joins the sugars in a DNA or RNA chain through their 5'- and 3'-hydroxyl groups by phosphodiester bonds



# DNA Structure

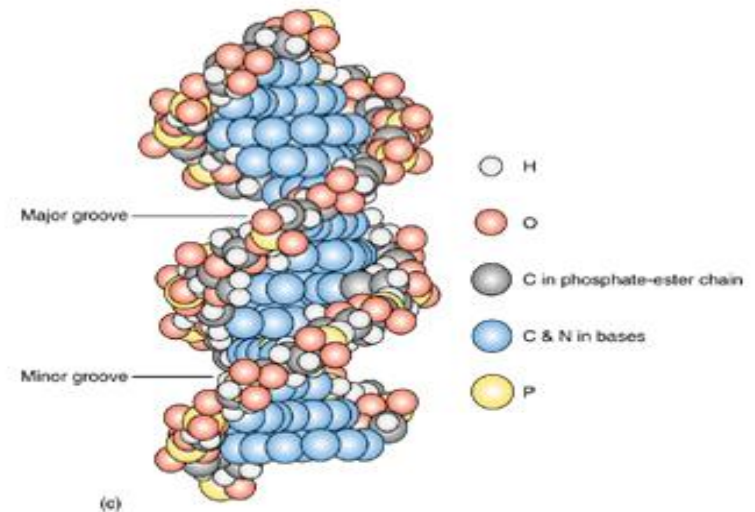
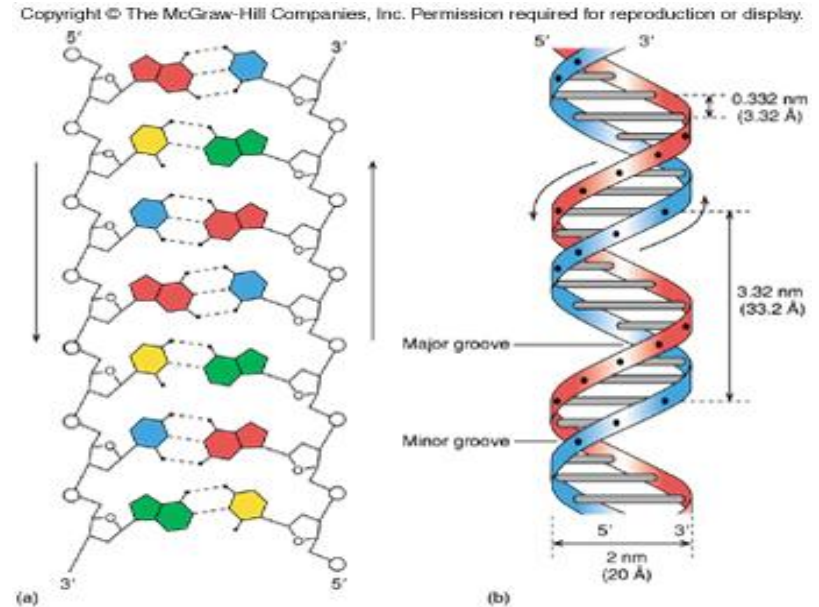
## The Double Helix

- Rosalind Franklin's x-ray data suggested that DNA had a helical shape
- The data also indicated a regular, repeating structure
- DNA was believed to require an irregular sequence
- Watson and Crick proposed a double helix with sugar-phosphate backbones on the outside and bases aligned to the interior



# DNA Helix

- Structure compared to a twisted ladder
  - Curving sides of the ladder represent the sugar-phosphate backbone
  - Ladder rungs are the base pairs
  - There are about 10 base pairs per turn
- Arrows indicate that the two strands are antiparallel





# Summary

- The DNA molecule is a double helix, with sugar-phosphate backbones on the outside and base pairs on the inside
- The bases pair in a specific way:
  - Adenine (A) with thymine (T)
  - Guanine (G) with cytosine (C)



# Genes Made of RNA

Hershey & Chase investigated bacteriophage, virus particle by itself, a package of genes

- This has no metabolic activity of its own
- When virus infects a host cell, the cell begins to make viral proteins
- Viral genes are replicated and newly made genes with viral protein assemble into virus particles

Some viruses contain DNA genes, but some viruses have RNA genes, either double- or single-stranded

# Physical Chemistry of Nucleic Acids

DNA and RNA molecules can appear in several different structural variants

- Changes in relative humidity will cause variation in DNA molecular structure
- The twist of the DNA molecule is normally shown to be right-handed, but left-handed DNA was identified in 1979



# A Variety of DNA Structures

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

**Table 2.2 Forms of DNA**

- High humidity DNA is called the B-form
- Lower humidity from cellular conditions to about 75% and DNA takes on the A-form
  - Plane of base pairs in A-form is no longer perpendicular to the helical axis
  - A-form seen when hybridize one DNA with one RNA strand in solution

Form	Pitch Å	Residues per Turn	Inclination of Base Pair from Horizontal (degrees)
A	24.6	10.7	+19
B	33.2	~10	-1.2
Z	45.6	12	-9

- When wound in a left-handed helix, DNA is termed Z-DNA
- One gene requires Z-DNA for activation





# Variation in DNA between Organisms

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

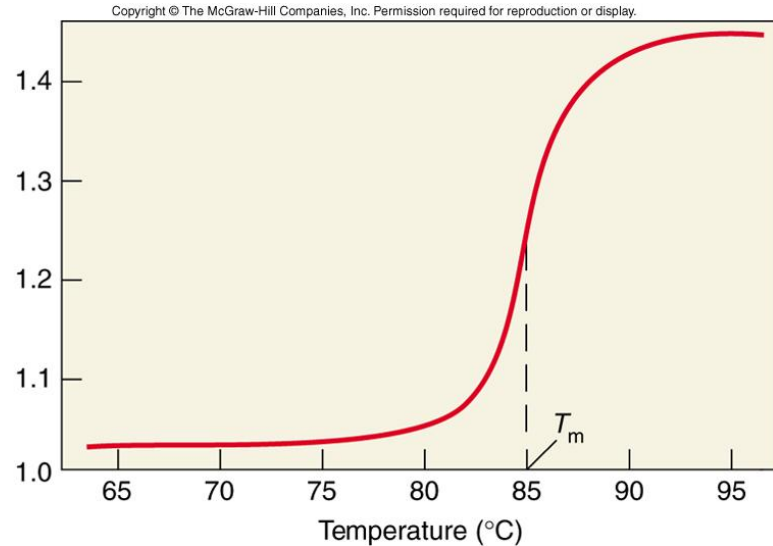
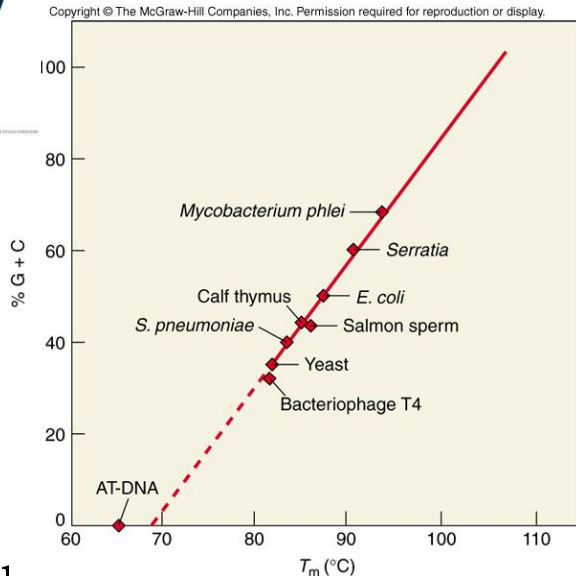
**Table 2.3** Relative G + C Contents of Various DNAs

Sources of DNA	Percent (G + C)
<i>Dictyostelium</i> (slime mold)	22
<i>Streptococcus pyogenes</i>	34
Vaccinia virus	36
<i>Bacillus cereus</i>	37
<i>B. megaterium</i>	38
<i>Haemophilus influenzae</i>	39
<i>Saccharomyces cerevisiae</i>	39
Calf thymus	40
Rat liver	40
Bull sperm	41
<i>Streptococcus pneumoniae</i>	42
Wheat germ	43
Chicken liver	43
Mouse spleen	44
Salmon sperm	44
<i>B. subtilis</i>	44
T1 bacteriophage	46
<i>Escherichia coli</i>	51
T7 bacteriophage	51
T3 bacteriophage	53
<i>Neurospora crassa</i>	54
<i>Pseudomonas aeruginosa</i>	68
<i>Sarcina lutea</i>	72
<i>Micrococcus lysodeikticus</i>	72
Herpes simplex virus	72
<i>Mycobacterium phlei</i>	73

- Ratios of G to C and A to T are fixed in any specific organism
- The total percentage of G + C varies over a range to 22 to 73%
- Such differences are reflected in differences in physical properties



# DNA Melting

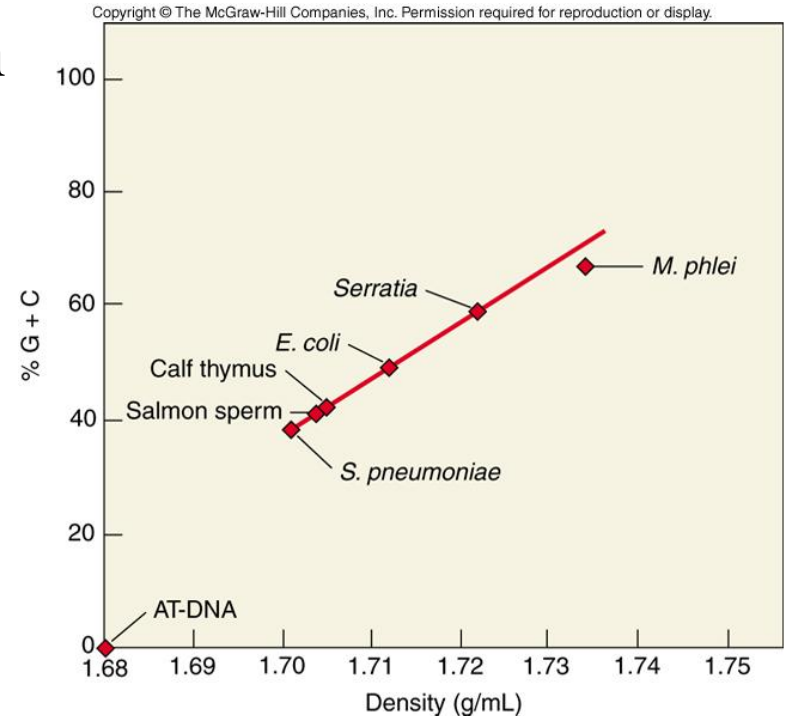


- With heating, noncovalent forces holding DNA strands together weaken and break
- When the forces break, the two strands come apart in denaturation or melting
- Temperature at which DNA strands are 1/2 denatured is the melting temperature or  $T_m$
- GC content of DNA has a significant effect on  $T_m$  with higher GC content meaning higher  $T_m$



# DNA Denaturation

- In addition to heat, DNA can be denatured by:
  - Organic solvents
  - High pH
  - Low salt concentration
- GC content also affects DNA density
  - Direct, linear relationship
  - Due to larger molar volume of an A-T base pair than a G-C base pair





# Summary

- GC content of a natural DNA can vary from less than 25% to almost 75%
- GC content has a strong effect on physical properties that increase linearly with GC content
  - Melting temperature, the temperature at which the two strands are half-dissociated or denatured
  - Density
  - Low ionic strength, high pH and organic solvents also promote DNA denaturation



# DNA Renaturation

- After two DNA strands separate, under proper conditions the strands can come back together
- Process is called **annealing** or **renaturation**
- Three most important factors:
  - Temperature – best at about 25 C below  $T_m$
  - DNA Concentration – within limits higher concentration better likelihood that 2 complementary will find each other
  - Renaturation Time – as increase time, more annealing will occur

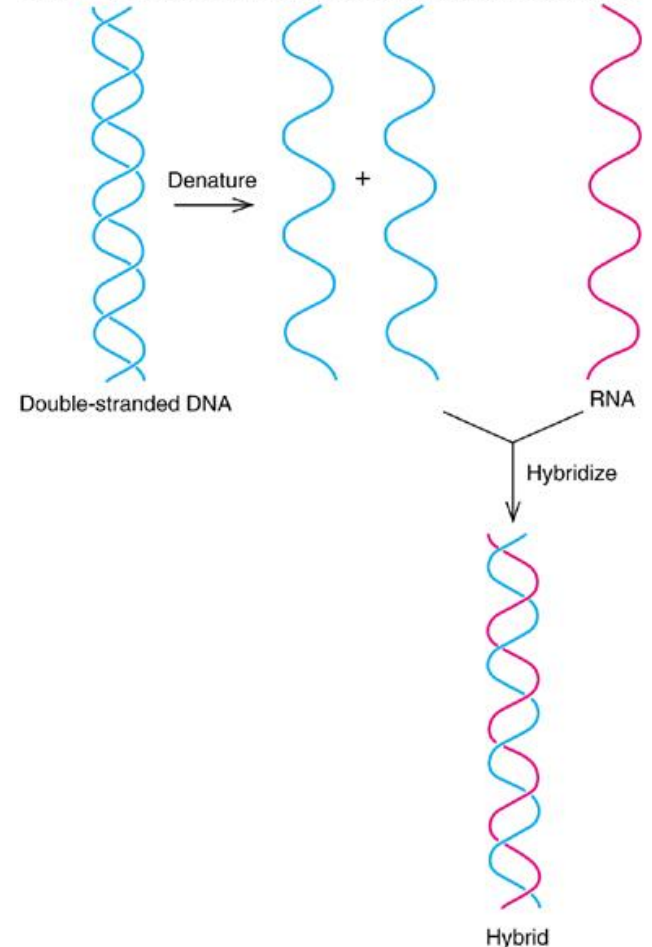


# Polynucleotide Chain Hybridization

Hybridization is a process of putting together a combination of two different nucleic acids

- Strands could be 1 DNA and 1 RNA
- Also could be 2 DNA with complementary or nearly complementary sequences

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.





# DNA Sizes

DNA size is expressed in 3 different ways:

- Number of base pairs
- Molecular weight – 660 is molecular weight of 1 base pair
- Length – 33.2 Å per helical turn of 10.4 base pairs

Measure DNA size either using electron microscopy or gel electrophoresis



# DNAs of Various Sizes and Shapes

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

**Table 2.4** Sizes of Various DNAs

Source	Molecular weight	Base pairs (bp)	Length
<b>Subcellular Genetic Systems:</b>			
SV40 (mammalian tumor virus)	$3.5 \times 10^6$	5226	1.7 $\mu\text{m}$
Bacteriophage $\phi$ X174 (double-stranded form)	$3.2 \times 10^6$	5386	1.8 $\mu\text{m}$
Bacteriophage $\lambda$	$3.3 \times 10^7$	$4.85 \times 10^4$	13 $\mu\text{m}$
Bacteriophage T2 or T4	$1.3 \times 10^8$	$2 \times 10^5$	50 $\mu\text{m}$
Human mitochondria	$9.5 \times 10^6$	16,596	5 $\mu\text{m}$
<b>Bacteria:</b>			
<i>Haemophilus influenzae</i>	$1.2 \times 10^9$	$1.83 \times 10^6$	620 $\mu\text{m}$
<i>Escherichia coli</i>	$3.1 \times 10^9$	$4.64 \times 10^6$	1.6 mm
<i>Salmonella typhimurium</i>	$8 \times 10^9$	$1.1 \times 10^7$	3.8 mm
<b>Eukaryotes (content per haploid nucleus):</b>			
<i>Saccharomyces cerevisiae</i> (yeast)	$7.9 \times 10^9$	$1.2 \times 10^7$	4.1 mm
<i>Neurospora crassa</i> (pink bread mold)	$\approx 1.9 \times 10^{10}$	$\approx 2.7 \times 10^7$	$\approx 9.2$ mm
<i>Drosophila melanogaster</i> (fruit fly)	$\approx 1.2 \times 10^{11}$	$\approx 1.8 \times 10^8$	$\approx 6.0$ cm
<i>Mus musculus</i> (mouse)	$\approx 1.5 \times 10^{12}$	$\approx 2.2 \times 10^9$	$\approx 750$ cm
<i>Homo sapiens</i> (human)	$\approx 2.3 \times 10^{12}$	$\approx 3.2 \times 10^9$	$\approx 1.1$ m
<i>Zea mays</i> (corn, or maize)	$\approx 4.4 \times 10^{12}$	$\approx 6.6 \times 10^9$	$\approx 2.2$ m
<i>Rana pipiens</i> (frog)	$\approx 1.4 \times 10^{13}$	$\approx 2.3 \times 10^{10}$	$\approx 7.7$ m
<i>Lilium longiflorum</i> (lily)	$\approx 2 \times 10^{14}$	$\approx 3 \times 10^{11}$	$\approx 100$ m

- Phage DNA is typically circular
- Some DNA will be linear
- Supercoiled DNA coils or wraps around itself like a twisted rubber band





# Summary

- Natural DNAs come in sizes ranging from several kilobases to thousands of megabases
- The size of a small DNA can be estimated by electron microscopy
- This technique can also reveal whether a DNA is circular or linear and whether it is supercoiled



# Relationship between DNA Size and Genetic Capacity

How does one know how many genes are in a particular piece of DNA?

- Can't determine from DNA size alone
- Factors include:
  - How DNA is devoted to genes?
  - What is the space between genes?
- Can estimate the upper limit of number genes a piece of DNA can hold



# DNA Size and Genetic Capacity

How many genes are in a piece of DNA?

- Start with basic assumptions
  - Gene encodes protein
  - Protein is about 40,000 D
- How many amino acids does this represent?
  - Average mass of an amino acid is about 110 D
  - Average protein –  $40,000 / 110 = 364$  amino acids
  - Each amino acid = 3 DNA base pairs
  - 364 amino acids requires 1092 base pairs



# DNA Genetic Capacity

How large is an average piece of DNA?

- *E. coli* chromosome
  - $4.6 \times 10^6$  bp
  - ~4200 proteins
- Phage  $\lambda$  (infects *E. coli*)
  - $4.85 \times 10^4$  bp
  - ~44 proteins
- Phage  $\phi$ x174 (one of smallest)
  - 5375 bp
  - ~5 proteins



# DNA Content and the C-Value Paradox

- C-value is the DNA content per haploid cell
- Might expect that more complex organisms need more genes than simple organisms
- For the mouse or human compared to yeast this is correct
- Yet the frog has 7 times more per cell than humans



# C-Value Paradox

- The observation that more complex organisms will not always need more genes than simple organisms is called the C-value paradox
- Most likely explanation for the paradox is that DNA that does not code for genes is present when the less complex organism has more DNA



# Summary

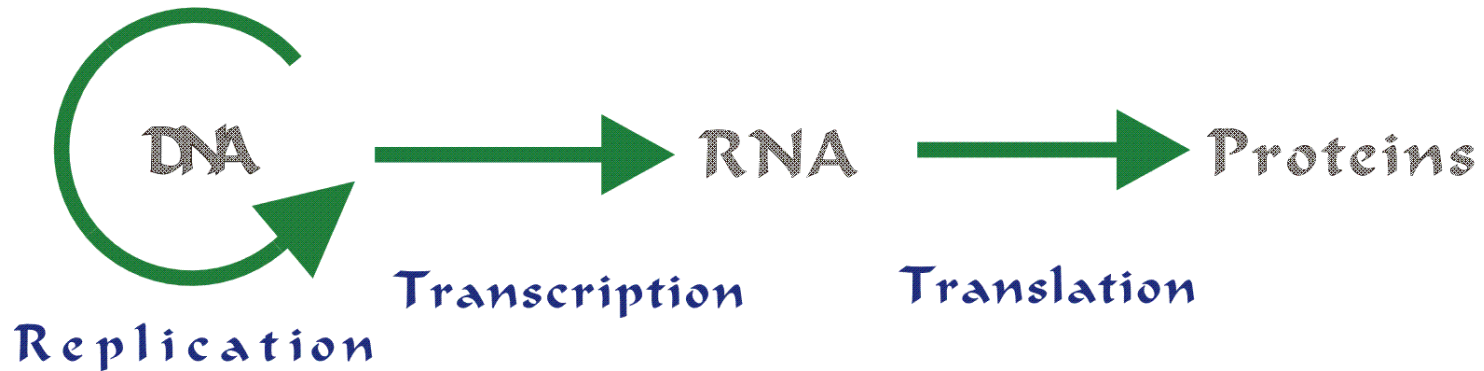
- There is a rough correlation between DNA content and number of genes in a cell or virus
- This correlation breaks down in several cases of closely related organisms where the DNA content per haploid cell (C-value) varies widely
- C-value paradox is probably explained not by extra genes, but by extra noncoding DNA in some organisms



# RNA as Intermediate

From DNA to protein, we need a translator.

RNA is the worker



**Question: Why?**





- 1. DNA is double chain
- 2. one RNA chain could produce many protein
- 3. DNA need to keep stable

**An example: book in the library**



# Players of translation

- Ribosome (RNA, proteins)
- **mRNA**
- **tRNA**
- Aminoacyl-tRNA synthetases
- Translation factors (initiation, elongation, termination)



## Differences between eubacteria and eukaryotes

- Ribosome: 30S+50S → 70S
- Few initiation factors:
  - IF-1(eIF1A), IF-2(eIF5B), IF-3 (?)
- Elongation factors
  - EF1A (EF-Tu), EF1B (EF-Ts), EF2 (EF-G)
- Release factors
  - RF-1, RF2, RF3
- Ribosome recycling factor
  - RRF
- mRNA is not capped
- Direct binding of 30S particle next to initiation codon (AUG) at Shine-Dalgarno sequence, 5'-AGGAGGU-3'
- Translation coupled to transcription

### Eukaryotes

- Ribosome: 40S+60S→ 80S
- Many initiation factors
  - eIF1, eIF1A, eIF2, eIF2B, eIF3, eIF4A, eIF4B, eIF4E, eIF4F, eIF4G, eIF4H, eIF5, eIF5B, eIF6
- Elongation factors
  - eEF1, eEF2
- Release factors
  - eRF1, eRF3
- Most mRNA is capped at 5' end and polyadenylated at 3' end
- 40S particle is recruited to 5' cap structure or poly(A) tail or an internal ribosome entry site (IRES)
- Translation is always (?) in cytoplasm apart from transcription

# tRNA

- Up to 50 (eukaryotes), or 30-35 (bacteria) different tRNAs

- Cloverleaf structure

- Unusual bases - covalent modification after transcription but before tRNAs leave nucleus

- Acceptor arm:

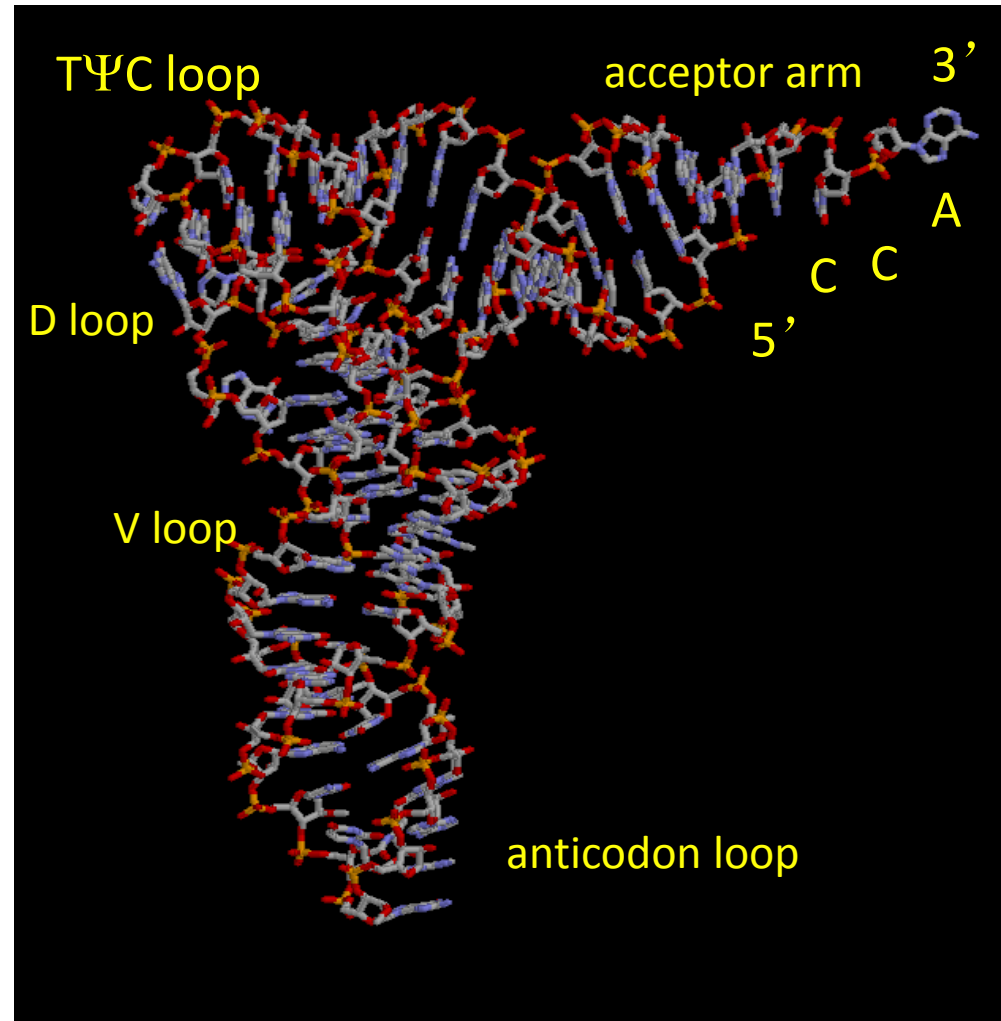
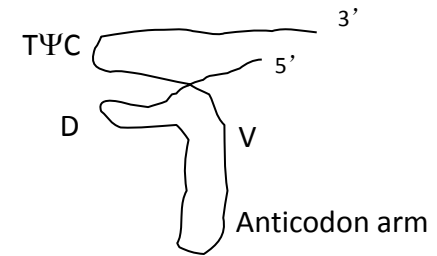
  - 7 base pairs followed by xCCA-3' aa attached to 2' or 3'-OH of terminal A by Class I and Class II aa-tRNA synthetases, respectively

- TYC arm forms one continuous helix with acceptor arm

- D arm (dihydro-uridine) interacts with TYC loop via unusual H-bonds

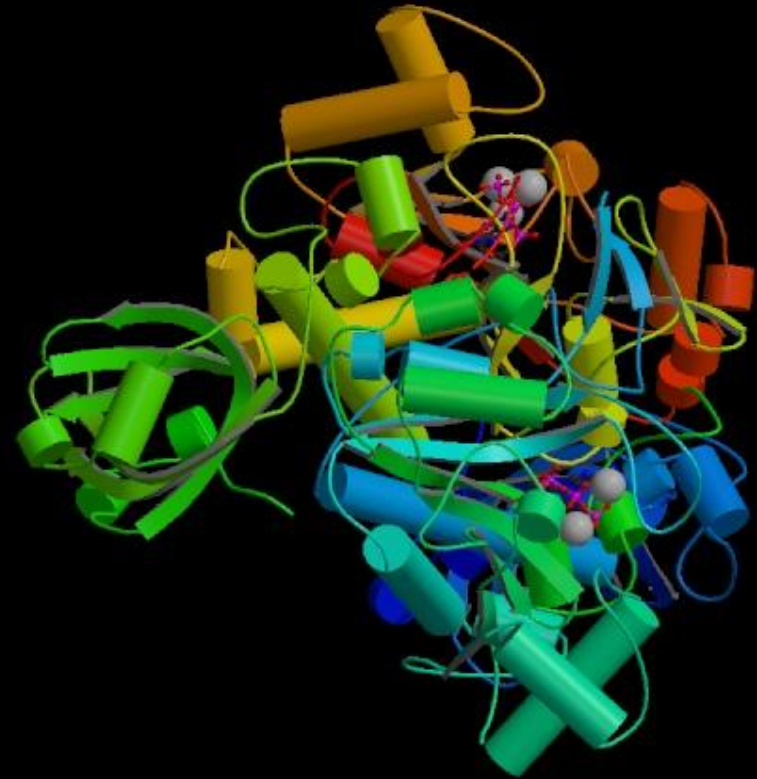
- V loop short in Class I tRNAs, long in Class II.

- Anticodon arm contains base triplet that pairs with mRNA codon

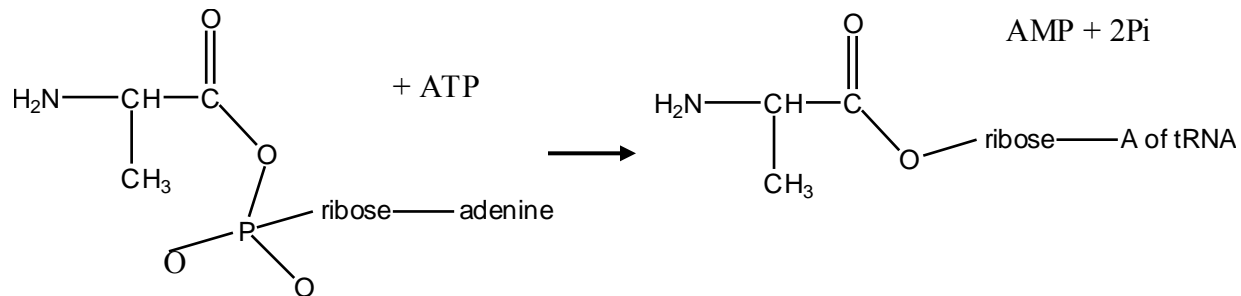


## Aminoacyl-tRNA synthetases

- Synthetase attaches aa to tRNA in a two-step process:
  - adenylation of aa
- 20 aa-tRNA synthetases, one for each aa. Bacteria have often fewer synthetases, and one synthetase attaches different amino acids to tRNA. Another enzyme then chemically modifies the incorrectly attached aa so that it corresponds to the anticodon of the tRNA
- Two classes of aa-tRNA synthetases
- Class I binds minor groove of acceptor arm, Class II binds major groove of acceptor arm (there are newly found exceptions)
- aa-tRNA synthetases have been engineered to incorporate unusual amino acids (P. Schultz, S. Yokoyama)

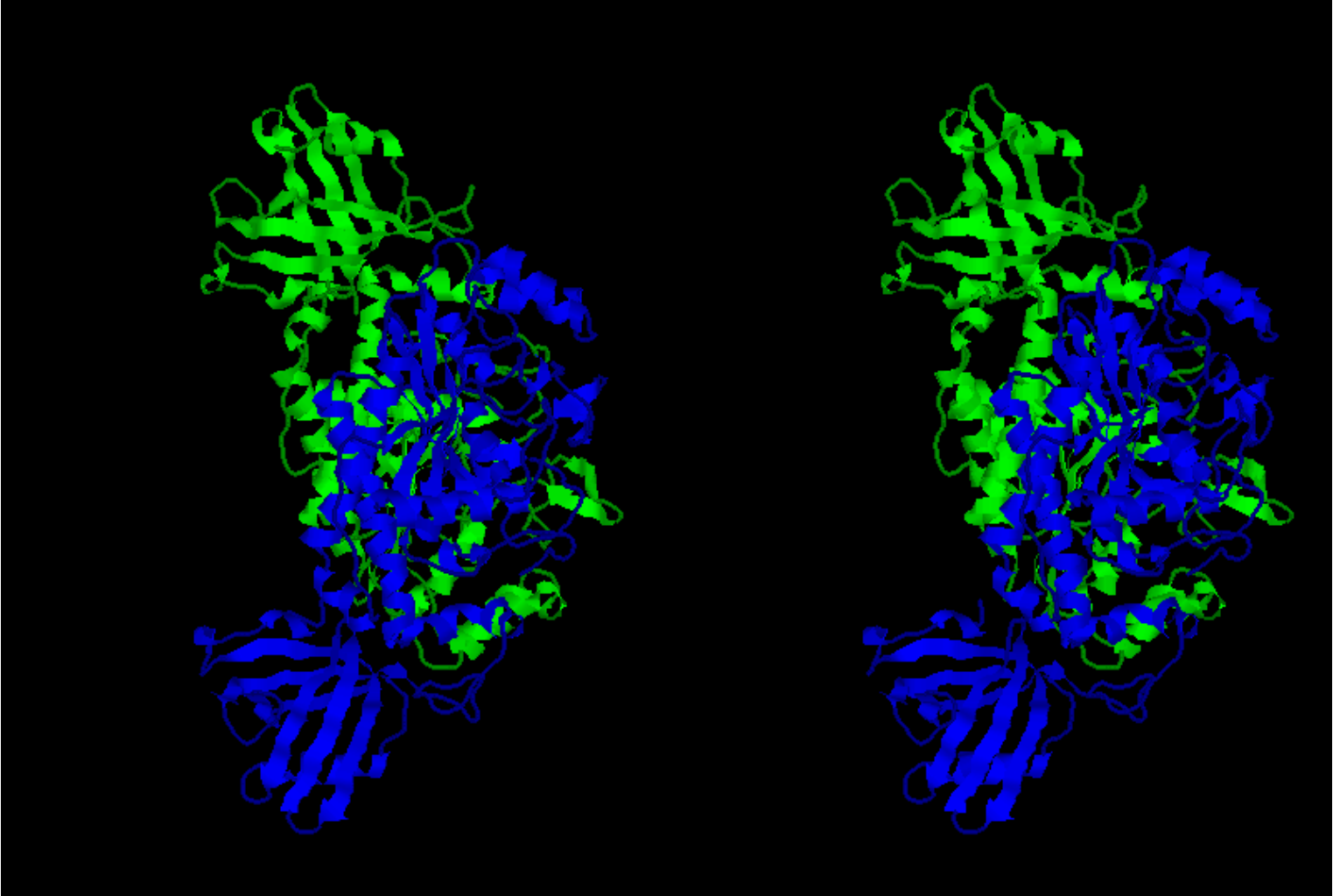


Asp-tRNA-synthetase





# aa-tRNA synthase





## Complex of Class I Tyr-tRNA synthetase with tRNA<sup>tyr</sup>

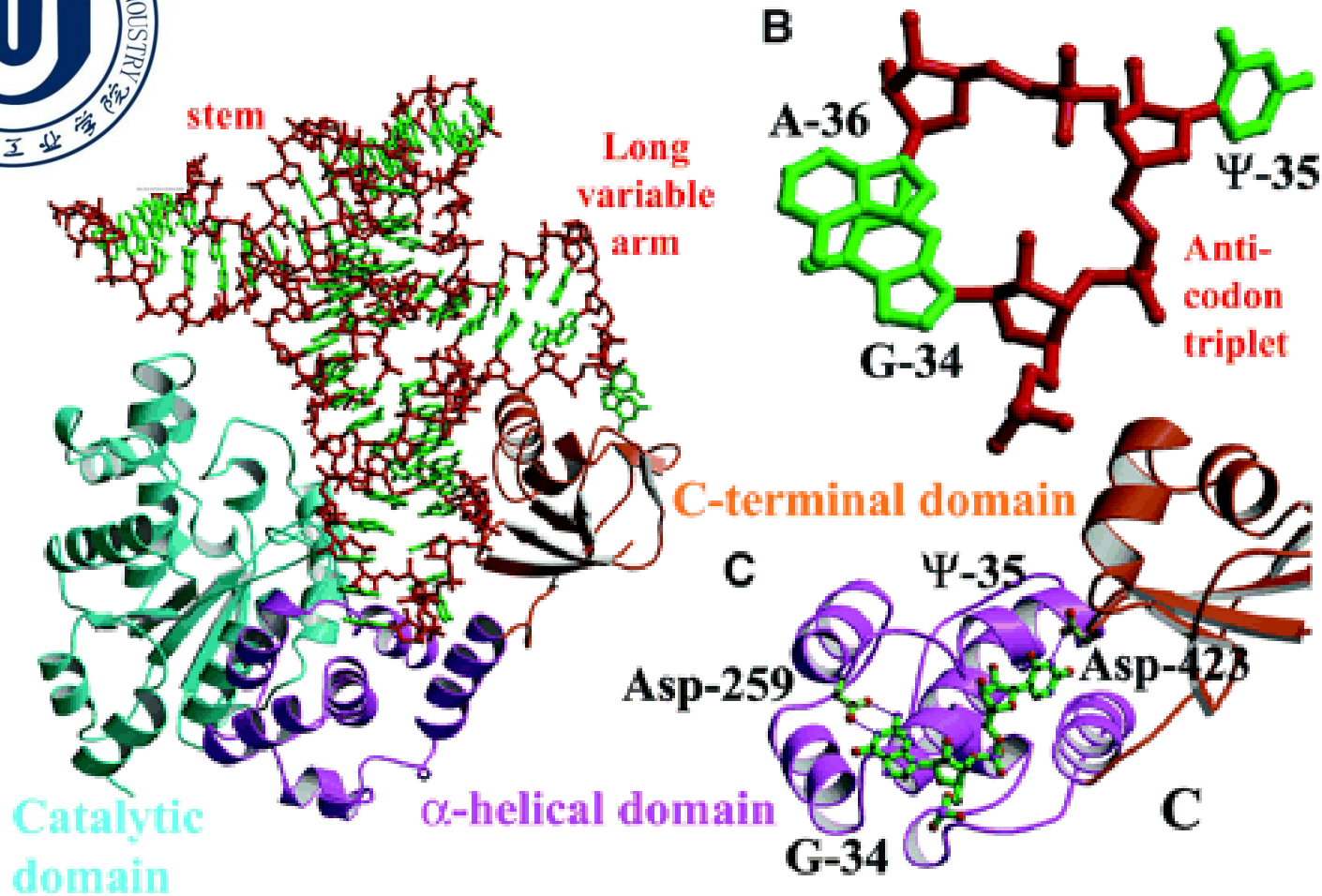
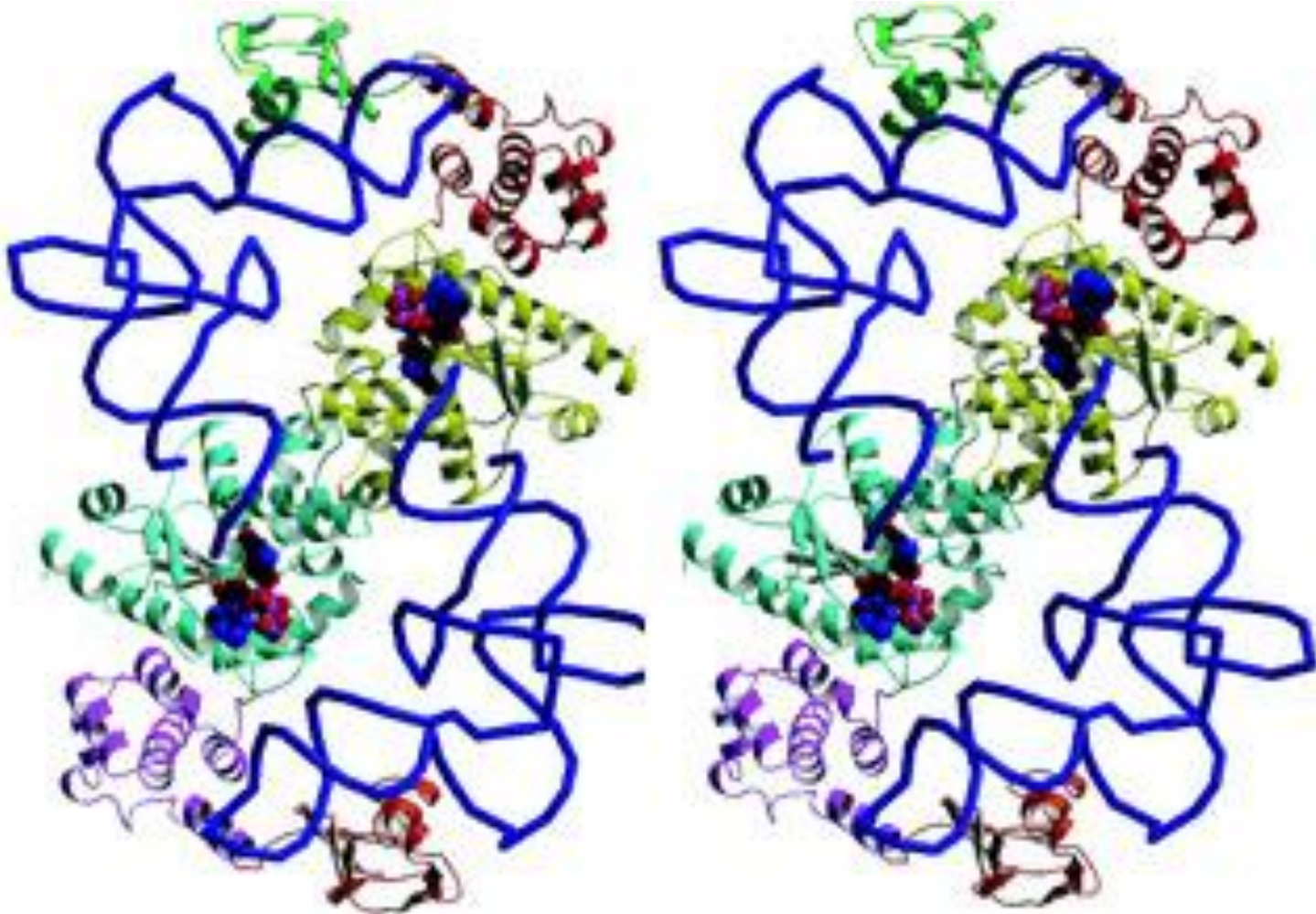


Fig. 3. Interactions between tyrosyl-tRNA synthetase and tRNA<sup>tyr</sup>. (A) The C-terminal domain (orange) binds in the elbow between the long variable arm and the anti-codon stem of the tRNA (red backbone, green bases). The anti-codon stem loop interacts with both the C-terminal domain and the  $\alpha$ -helical domain (pink). The tRNA makes no contact with the catalytic domain of the same subunit (cyan). (B) The unusual conformation of the anti-codon triplet in which Ade-36 is stacked on Gua-34, while Psi-35 bulges out. (C) Base-specific interactions of Asp-259 from the  $\alpha$ -helical domain with Gua-34 and Asp-423 from the C-terminal domain with Psi-35.



**Tyr-tRNA synthase complex with tRNA<sup>Tyr</sup>**

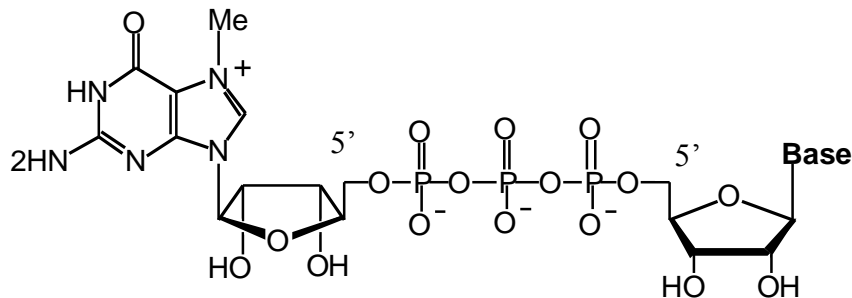






# mRNA

- Linear in bacteria - can circularize in eukaryotes (via Pabp, eIF4G and eIF4E)
- In bacteria, ribosome is recruited to AUG codon via a Shine-Dalgarno sequence
  - 5' AGGAGGU-(X)<sub>3-10</sub>-AUG 3'
- In eukaryotes, mRNA is usually capped and poly-adenylated - a consensus sequence is found around the initiation codon -ACCAAUGG- (Kozak sequence)
- 5'end-5'UTR-AUG-coding region-stop codon-3'UTR-poly(A)tail





# Capped mRNA

- Capping happens right after transcription, after about 25 nucleotides have been synthesized
- Capping by three enzymes:
  - Phosphatase removes one phosphate from 5' end
  - Guanyl transferase adds a GMP in reverse linkage (5' to 5' instead of 5' to 3')
  - Methyl transferase adds a methyl to the guanosine
  - Some RNAs are also methylated at the second nucleotide
  - All three enzymes bind to the phosphorylated RNA polymerase tail
- Cap addition distinguishes mRNA from other RNAs and helps to direct the ribosome to mRNA
- Cap is recognized by the cap-binding complex (CBC), consisting of two proteins, CBP80 and CBP20. Cap is stacked between two tyrosines Y20 and Y43 of CBP20. Binding is achieved via the  $\pi$ -stacking effect. CBC stabilizes the mRNA and interacts with nuclear pore complex during export of mRNA.
- In cytoplasm, CBC is replaced with eIF4E that helps to recruit the ribosome to mRNA. Here the m<sup>7</sup>G is stacked between two tryptophanes ( $\pi$ -stacking )



# Protein as Workers

## Translation:

1. Initiation
2. Elongation
3. Termination



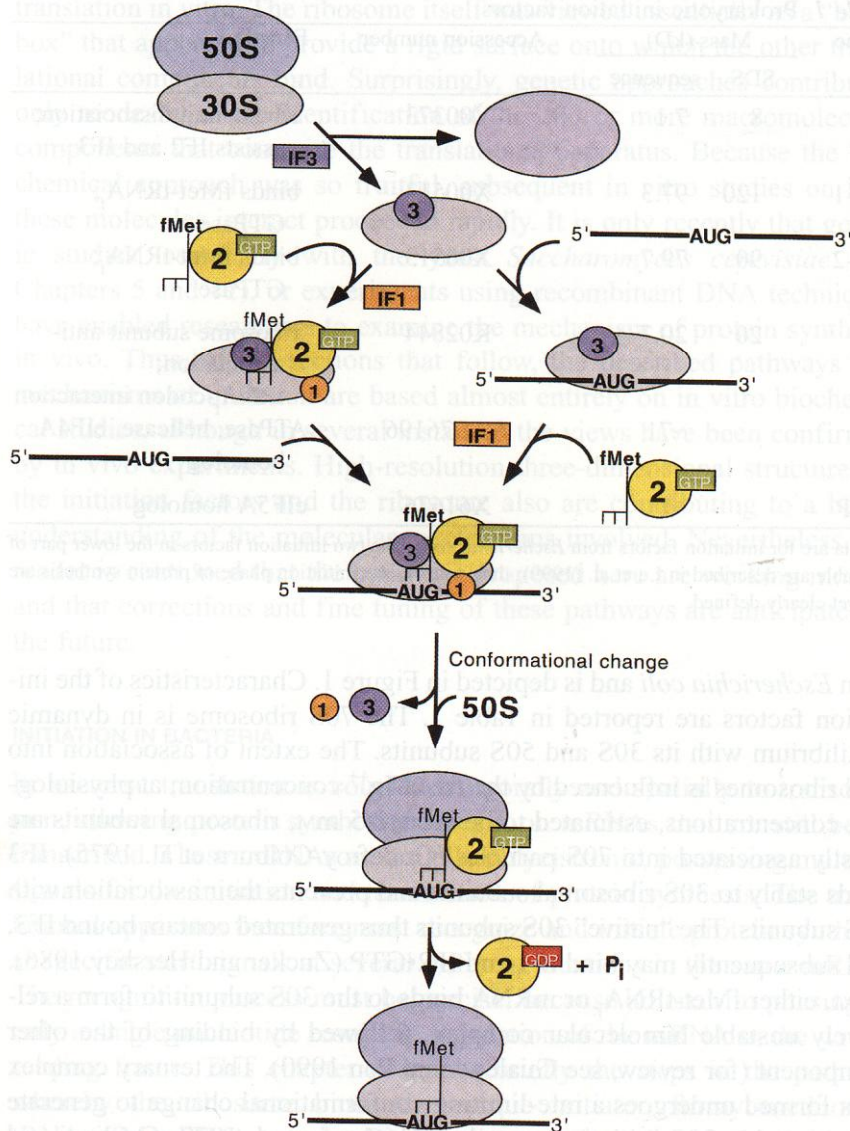
# Initiation

- Two important events must occur before translation initiation can take place
  - Generate a supply of aminoacyl-tRNAs
    - Amino acids must be covalently bound to tRNAs
    - Process of bonding tRNA to amino acid is called **tRNA charging**
  - Dissociation of ribosomes into their two subunits
    - The cell assembles the initiation complex on the small ribosomal subunit
    - The two subunits must separate to make assembly possible

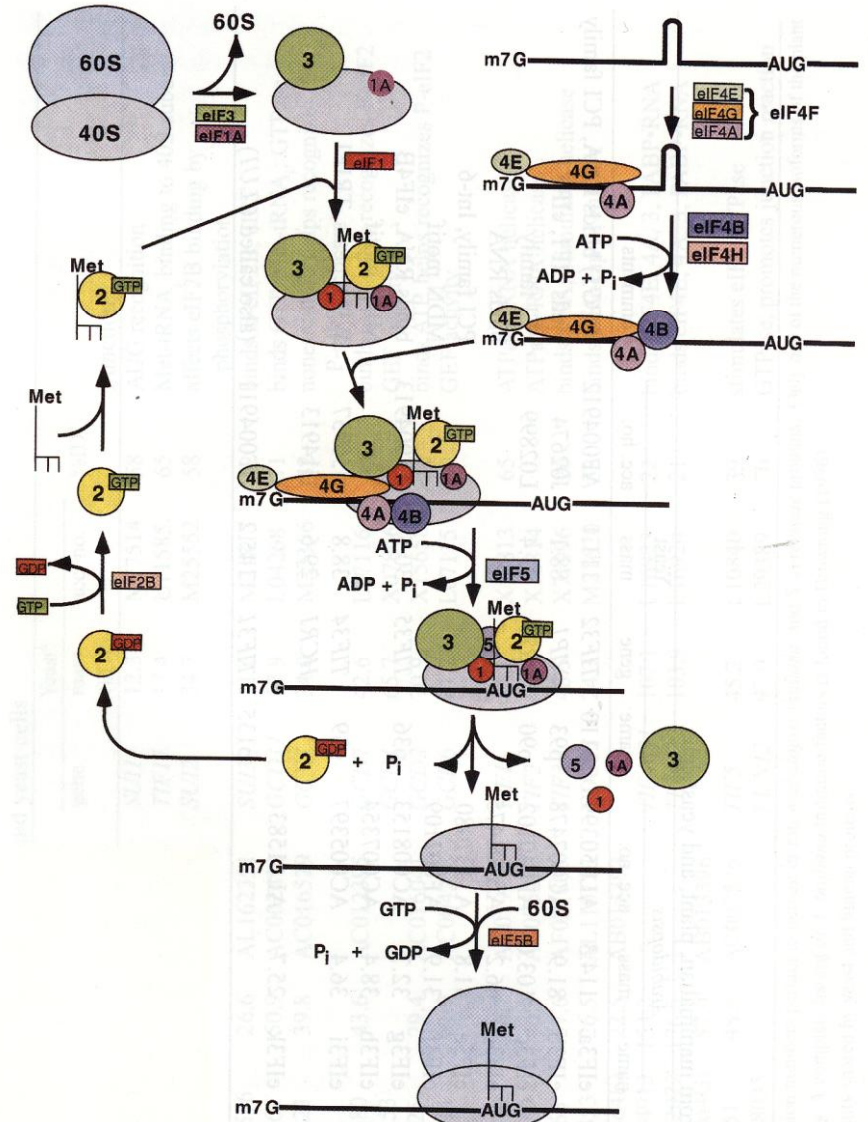
# Initiation prokaryotes vs. eukaryotes

From: Hershey & Merrick in "Translational Control of Gene expression" pp. 33-88", Sonenberg, Hershey, Matthews, eds. CSH Press 2000

## prokaryotes



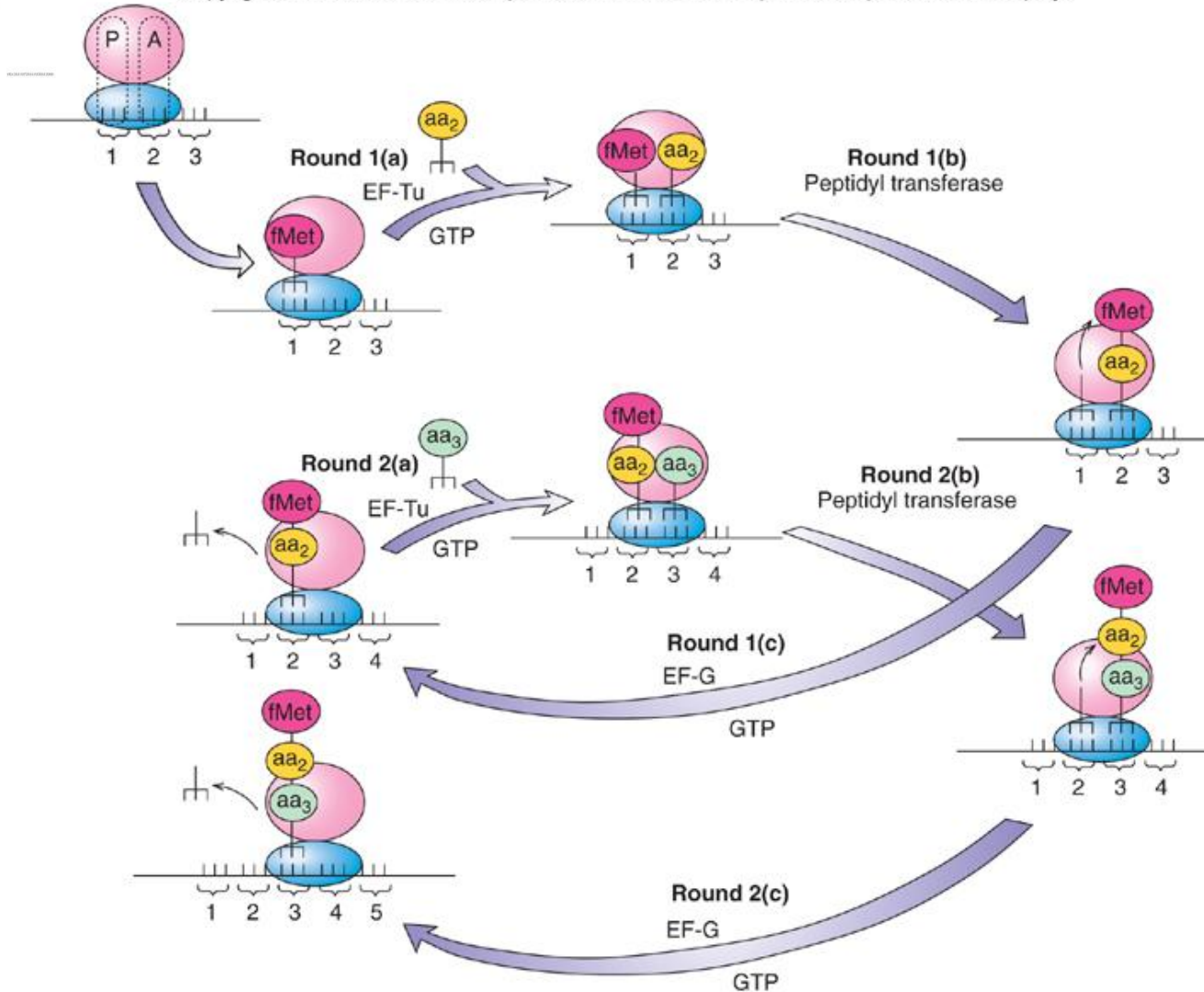
## eukaryotes





# Elongation

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.





# Elongation Step 1

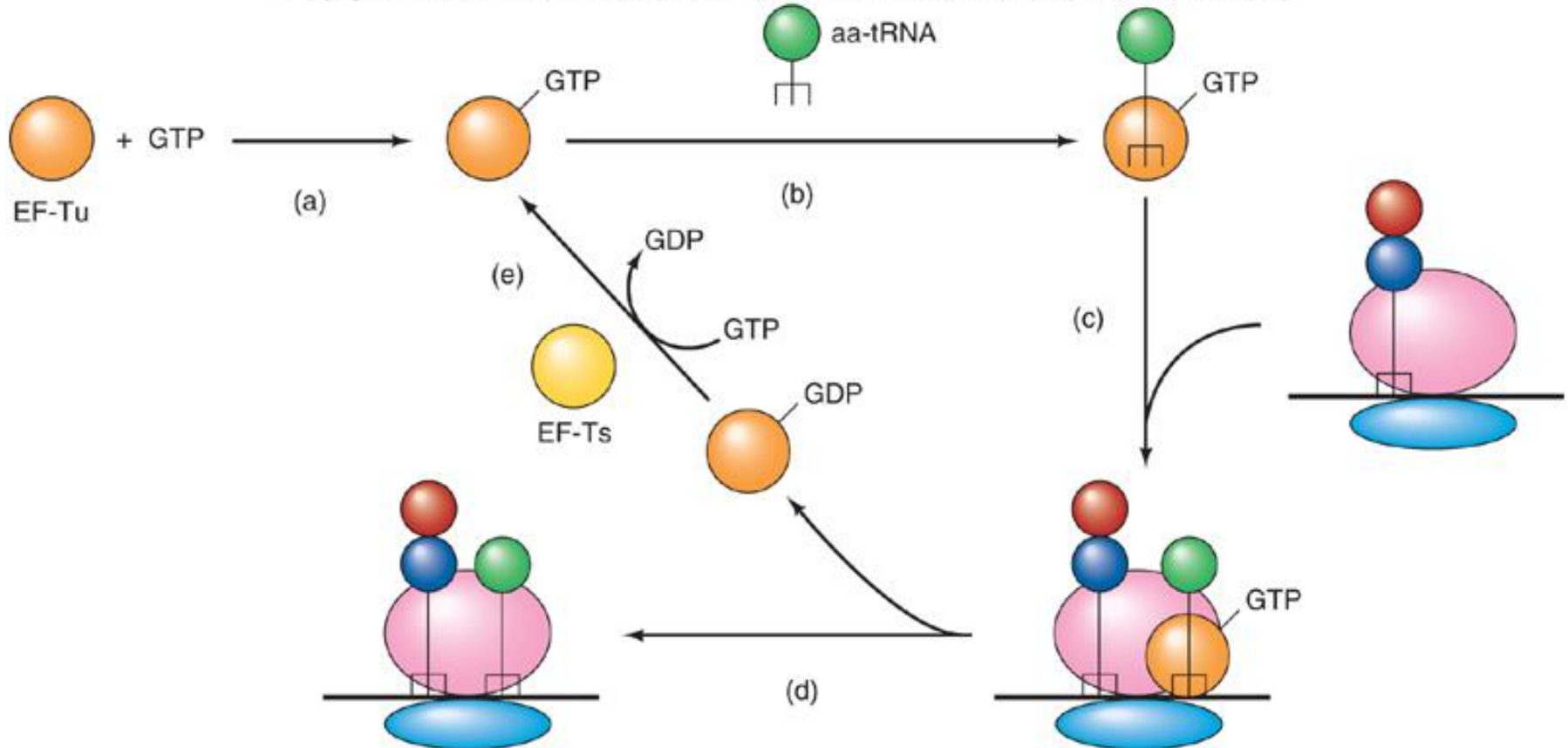
Binding aminoacyl-tRNA to A site of ribosome

- Ternary complex formed from:
  - EF-Tu
  - Aminoacyl-tRNA
  - GTP
- Delivers aminoacyl-tRNA to ribosome A site without hydrolysis of GTP
- Next step:
  - EF-Tu hydrolyzes GTP
  - Ribosome-dependent GTPase activity
  - EF-Tu-GDP complex dissociates from ribosome
- Addition of aminoacyl-tRNA reconstitutes ternary complex for another round of translation elongation



# Aminoacyl-tRNA Binding to Ribosome A Site

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.







# Proofreading

- Protein synthesis accuracy comes from charging tRNAs with correct amino acids
- Proofreading is correcting translation by rejecting an incorrect aminoacyl-tRNA before it can donate its amino acid
- Protein-synthesizing machinery achieves accuracy during elongation in two steps



# Protein-Synthesizing Machinery

- Two steps achieve accuracy:
  - Gets rid of ternary complexes bearing wrong aminoacyl-tRNA before GTP hydrolysis
  - If this screen fails, still eliminate incorrect aminoacyl-tRNA in the proofreading step before wrong amino acid is incorporated into growing protein chain
- Steps rely on weakness of incorrect codon-anticodon base pairing to ensure dissociation occurs more rapidly than either GTP hydrolysis or peptide bond formation



# Proofreading Balance

- Balance between speed and accuracy of translation is delicate
  - If peptide bond formation goes too fast
    - Incorrect aminoacyl-tRNAs do not have enough time to leave the ribosome
    - Incorrect amino acids are incorporated into proteins
  - If translation goes too slowly
    - Proteins are not made fast enough for the organism to grow successfully
- Actual error rate,  $\sim 0.01\%$  per amino acid is a good balance between speed and accuracy



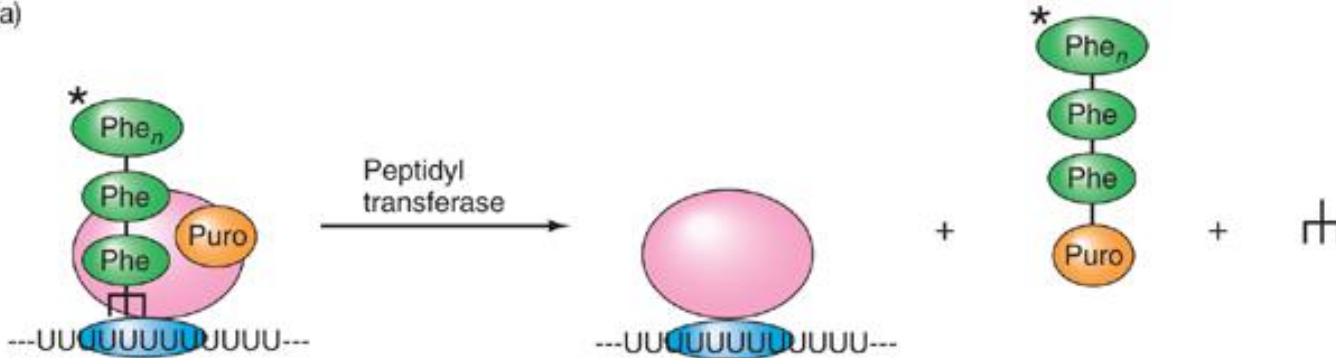
## Elongation Step 2

- One the initiation factors and EF-Tu have done their jobs, the ribosome has fMet-tRNA in the P site and aminoacyl-tRNA in the A site
- Now form the first peptide bond
- No new elongation factors participate in this event
- Ribosome contains the enzymatic activity, **peptidyl transferase**, that forms peptide bond

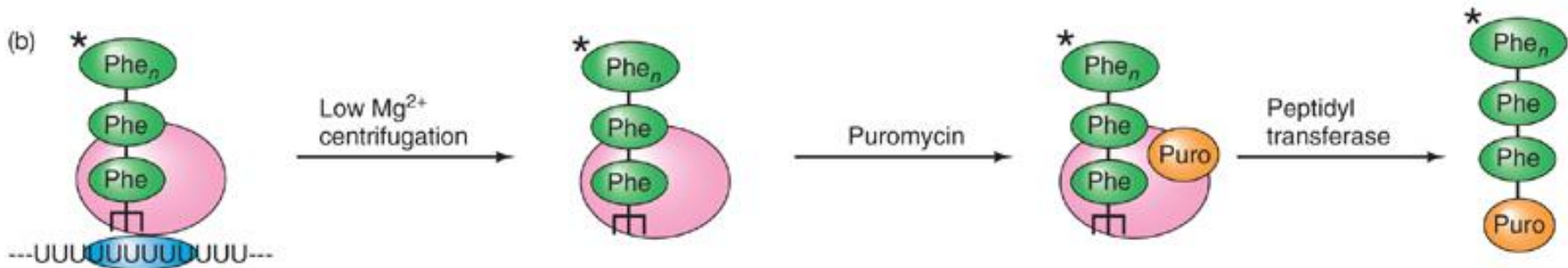
# Assay for Peptidyl Transferase

Copyright © The McGraw-Hill Companies, Inc. Permission required for reproduction or display.

(a)



(b)





# Peptide Bond Formation

- The peptidyl transferase resides on the 50S ribosomal particle
- Minimum components necessary for activity are 23S rRNA and proteins L2 and L3
- 23S rRNA is at the catalytic center of peptidyl transferase



# Elongation Step 3

- When peptidyl transferase has worked:
  - Ribosome has peptidyl-tRNA in the A site
  - Deacylated tRNA in the P site
- Translocation, next step, moves mRNA and peptidyl-tRNA one codon's length through the ribosome
  - Places peptidyl-tRNA in the P site
  - Ejects the deacylated tRNA
  - Process requires elongation factor EF-G which hydrolyzes GTP after translocation is complete



# Termination

- Elongation cycle repeats over and over
  - Adds amino acids one at a time
  - Grows the polypeptide product
- Finally ribosome encounters a stop codon
  - Stop codon signals time for last step
  - Translation last step is termination





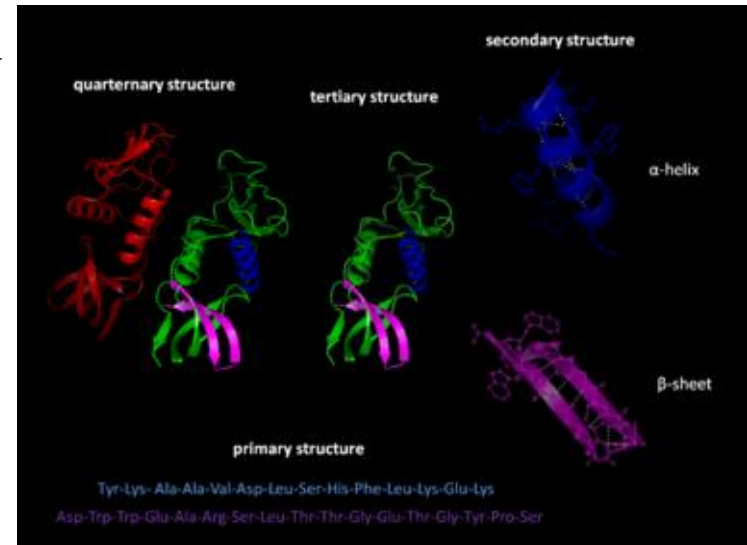
# Termination Codons

- Three codons are the natural stop signals at the ends of coding regions in mRNA
  - UAG
  - UAA
  - UGA
- Mutations can create termination codons within an mRNA causing premature termination of translation
  - Amber mutation creates UAG
  - Ochre mutation creates UAA
  - Opal mutation creates UGA



# Posttranslation

- Translation events do not end with termination
  - Proteins must fold properly
  - Ribosomes need to be released from mRNA and engage in further translation rounds
- Folding is actually a cotranslational event occurring as nascent polypeptide is being made





# Folding Nascent Proteins

- Most newly-made polypeptides do not fold properly alone
  - Polypeptides require folding help from molecular chaperones
  - E. coli cells use a **trigger factor**
    - Associates with the large ribosomal subunit
    - Catches the nascent polypeptide emerging from ribosomal exit tunnel in a hydrophobic basket to protect from water
  - Archaea and eukaryotes lack trigger factor, use freestanding chaperones



Thank You!