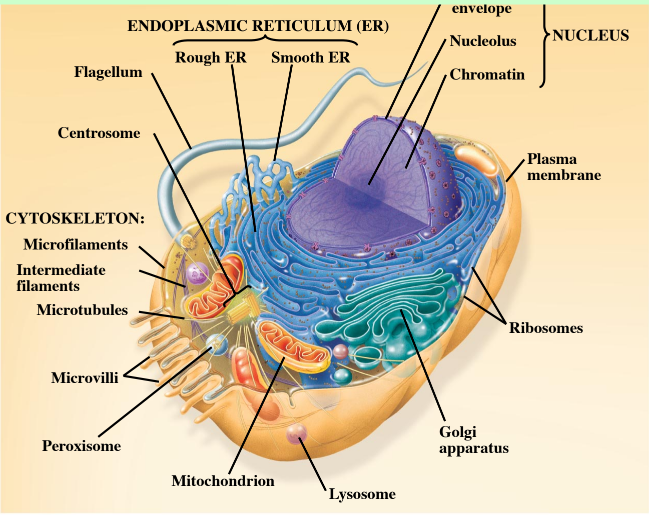
Molecular and Cellular Biology

3. The Cell – From Genes to Proteins

key processes

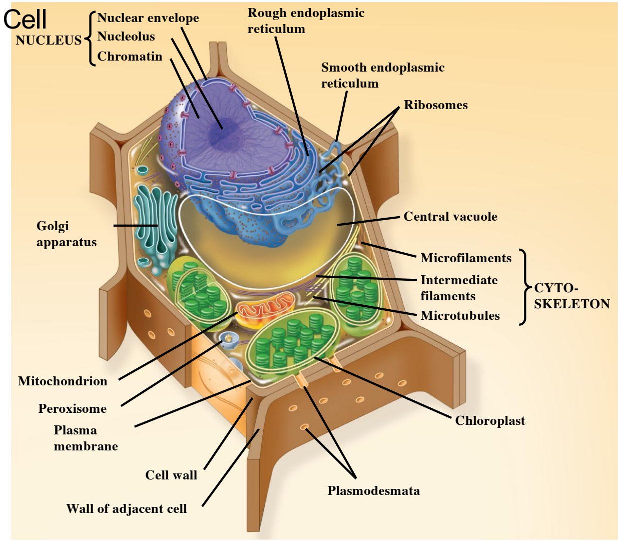
Prof. Dr. Klaus Heese

Animal Cell ((eukaryotic cell) -----> compare with prokaryotic cell)


envelope
ENDOPLASMIC RETICULUM (ER)
Rough ER
Smooth ER
Nucleolus
Chromatin
NUCLEUS
Flagellum
Centrosome
Plasma membrane
CYTOSKELETON:
Microfilaments
Intermediate filaments
Microtubules
Ribosomes
Microvilli
Peroxisome
Golgi apparatus
Mitochondrion
Lysosome

Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

Plant Cell


NUCLEUS
Nuclear envelope
Nucleolus
Chromatin
Rough endoplasmic reticulum
Smooth endoplasmic reticulum
Ribosomes
Golgi apparatus
Central vacuole
Microfilaments
Intermediate filaments
Microtubules
CYTO-SKELETON
Mitochondrion
Peroxisome
Plasma membrane
Chloroplast
Cell wall
Wall of adjacent cell
Plasmodesmata

Copyright © 2008 Pearson Education, Inc., publishing as Pearson Benjamin Cummings.

Monomers

vs

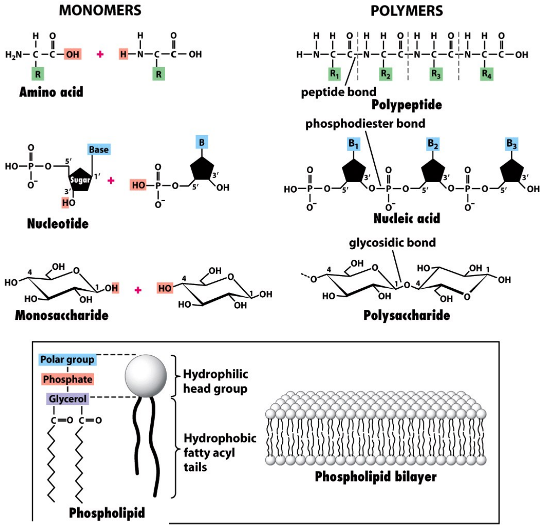
Polymers:

defines

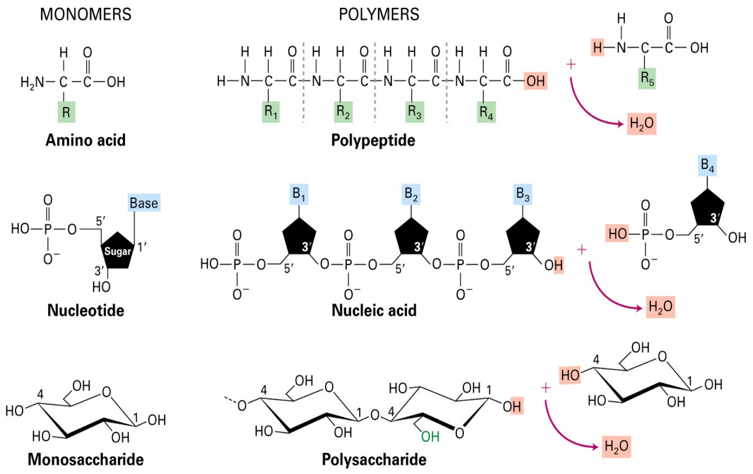
cell feature

&

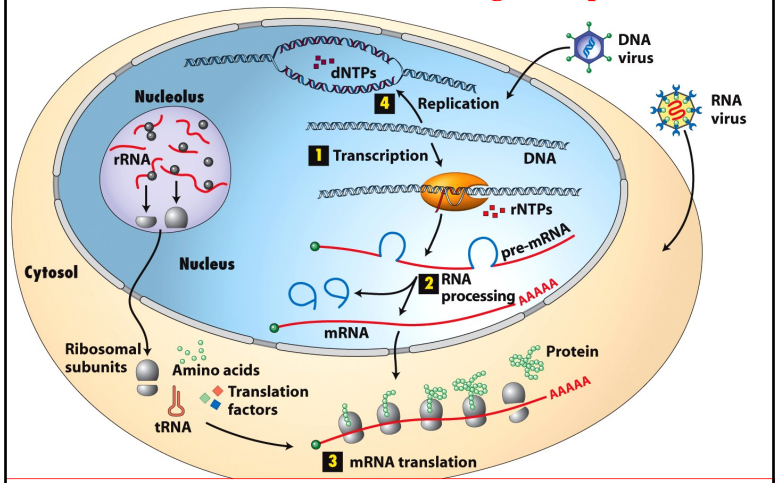
cell metabolism


MONOMERS
POLYMERS
Amino acid
peptide bond
Polypeptide
phosphodiester bond
Base
Sugar
Nucleotide
Nucleic acid
glycosidic bond
Monosaccharide
Polysaccharide
Polar group
Phosphate
Glycerol
Hydrophilic head group
Hydrophobic fatty acyl tails
Phospholipid
Phospholipid bilayer

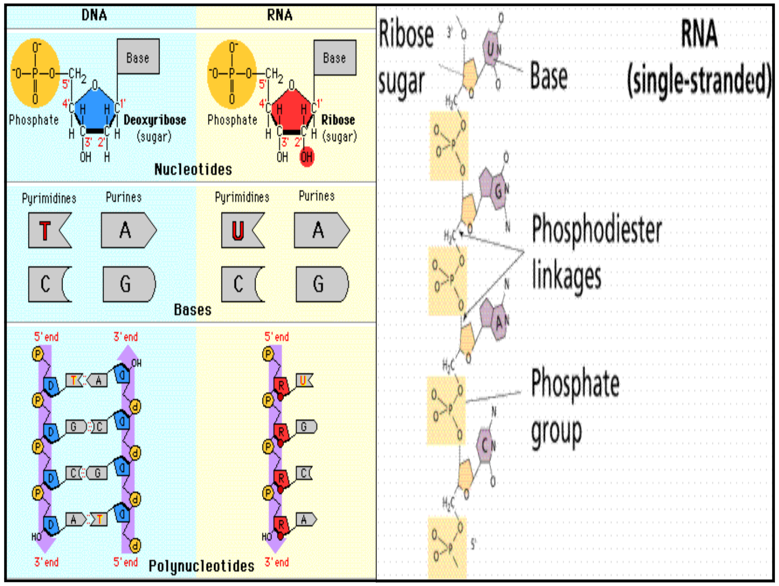
Essential Bio-Engineering-related Bio-Molecules


MONOMERS
POLYMERS
Amino acid
Polypeptide
H_2O
Nucleotide
Base
Sugar
Nucleic acid
H_2O
Monosaccharide
Polysaccharide
H_2O

Overview of four (4) basic molecular genetic processes

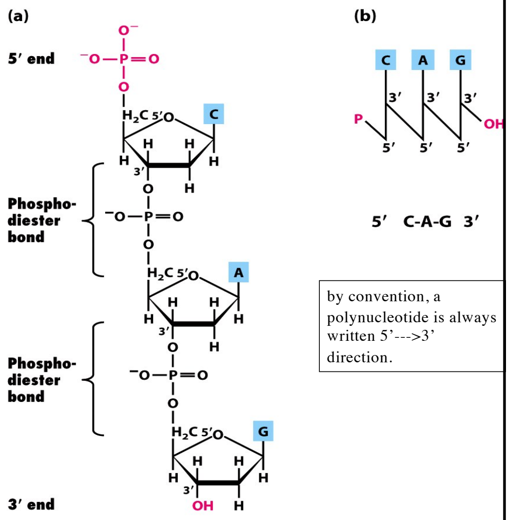

DNA virus
RNA virus
dNTPs
4 Replication
Nucleolus
rRNA
1 Transcription
DNA
rNTPs
pre-mRNA
Nucleus
Cytosol
2 RNA processing
AAAAA
mRNA
Ribosomal subunits
Amino acids
Protein
Translation factors
tRNA
AAAAA
3 mRNA translation

from genes to peptides / proteins: 3 major steps: 1) transcription, 2) processing (splicing), 3) translation

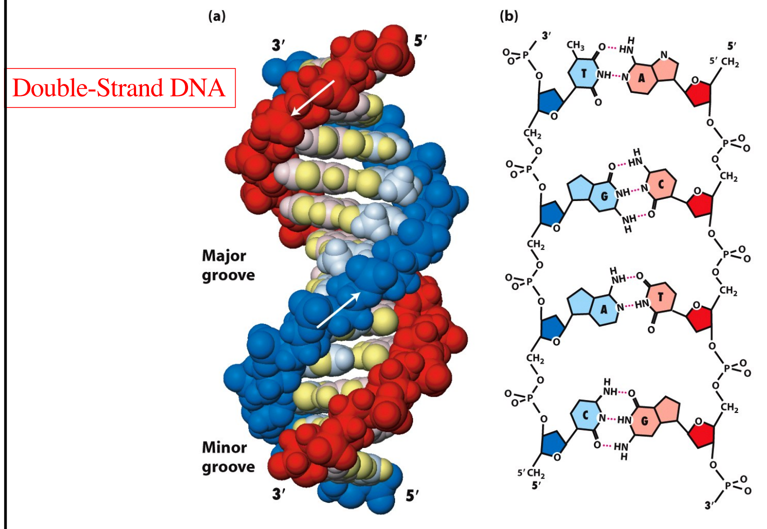

DNA
RNA
Phosphate
Base
Deoxyribose (sugar)
Phosphate
Base
Ribose (sugar)
Nucleotides
Pyrimidines
Purines
Pyrimidines
Purines
T
A
U
A
C
G
C
G
Bases
5' end
3' end
5' end
3' end
5' end
3' end
Polynucleotides
Ribose sugar
Base
RNA (single-stranded)
Phosphodiester linkages
Phosphate group

Single-Strand DNA

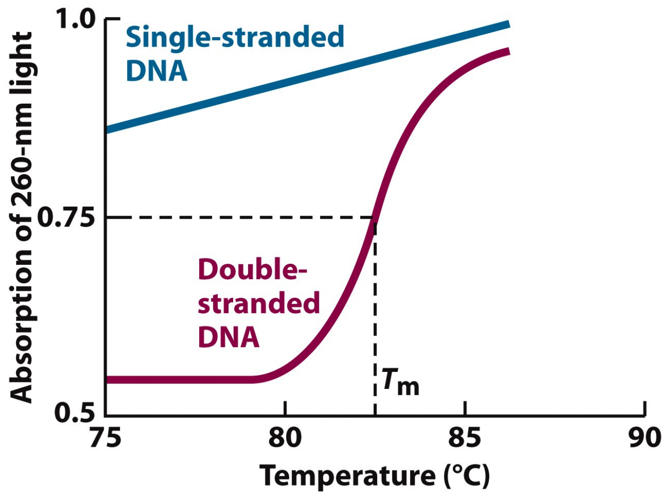
alternative representations of a nucleic acid strand illustrating its chemical directionality.

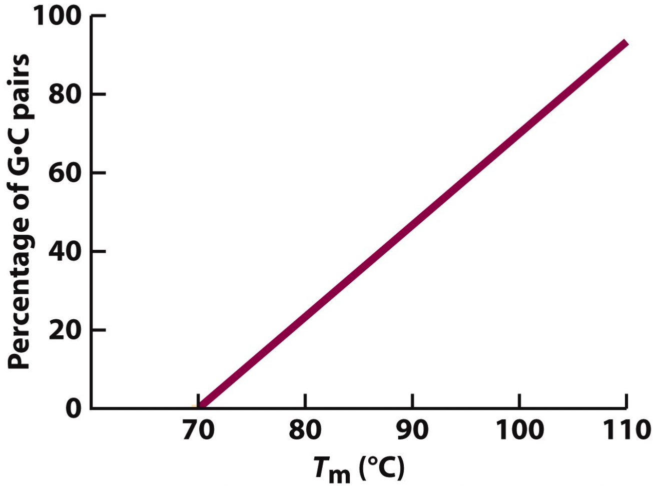

(a)
5′ end
Phospho-diester bond
Phospho-diester bond
3′ end
C
A
G
(b)
C
A
G
P
OH
5′ C-A-G 3′

by convention, a polynucleotide is always written 5'--->3' direction.

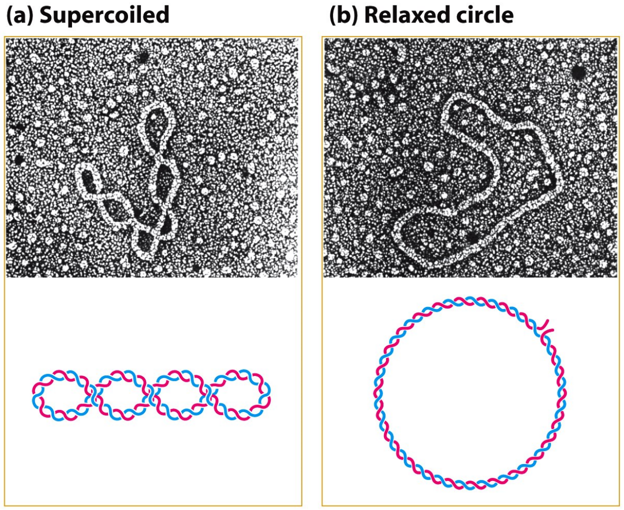

(a)
(b)
3′
5′
3′
5′
Major groove
Minor groove
T
A
G
C
A
T
C
G
CH3
CH2
5′ CH2
O
HN
NH
N

Double-Strand DNA



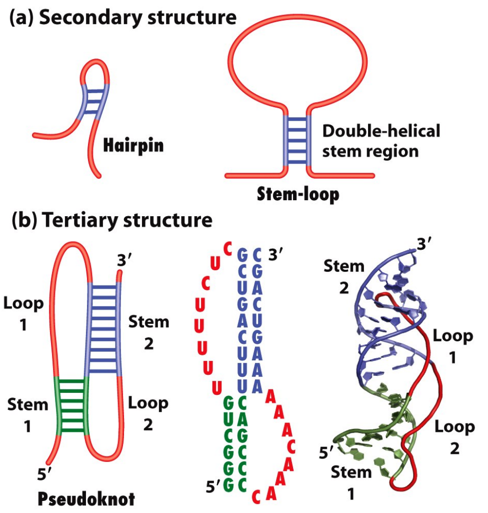


The temperature at which DNA denatures increases with the proportion of GC-pairs. The higher GC%, the higher Tm.


(a) Supercoiled
(b) Relaxed circle

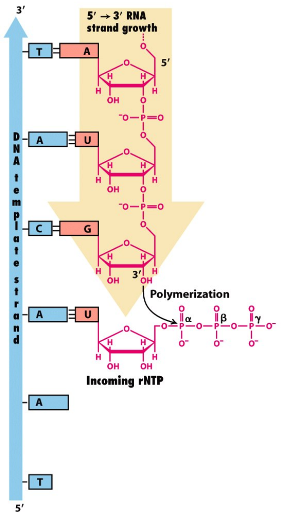
DNA supercoils can be removed by cleavage of one strand

RNA secondary and tertiary structures


(a) Secondary structure
Hairpin
Double-helical stem region
Stem-loop
(b) Tertiary structure
Loop 1
3′
Stem 2
Stem 1
Loop 2
5′
Pseudoknot
C U C U U U U U
G C 3′
C G
U A
G C
A U
C G
U A
U A
U A
G C
U A
C G
G C
G C
5′ G C
A A A C A A A C A
Stem 2
3′
Loop 1
Loop 2
5′
Stem 1

Transcription:

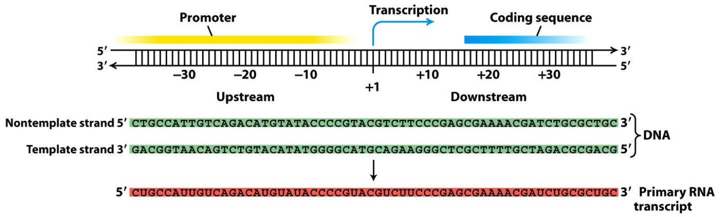
RNA-synthesis: 5’--->3’


3′
5′ → 3′ RNA strand growth
T A
A U
C G
A U
A
T
5′
DNA template strand
H OH
O P O
3′ OH
Polymerization
α β γ
Incoming rNTP

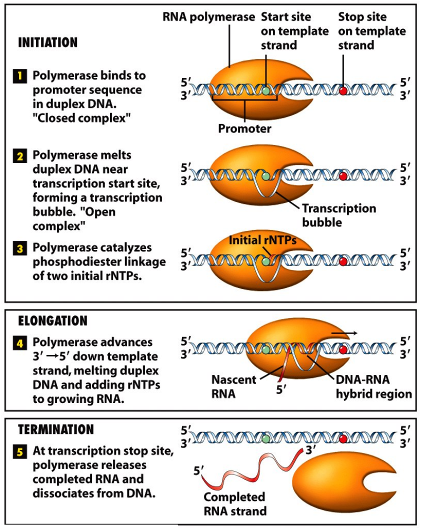
RNA polymerase

Transcription:

DNA ---> RNA


Promoter
Transcription
Coding sequence
5′
3′
3′
5′
−30
−20
−10
+1
+10
+20
+30
Upstream
Downstream
Nontemplate strand 5′ CTGCCATTGTCAGACATGTATACCCCGTACGTCTTCCCGAGCGAAAACGATCTGCGCTGC 3′
Template strand 3′ GACGGTAACAGTCTGTACATATGGGGCATGCAGAAGGGCTCGCTTTTGCTAGACGCGACG 5′
DNA
5′ CUGCCAUUGUCAGACAUGUAUACCCCGUACGUCUUCCCGAGCGAAAACGAUCUGCGCUGC 3′
Primary RNA transcript

Three stages in Transcription


RNA polymerase
Start site on template strand
Stop site on template strand
5′
3′
Promoter
Transcription bubble
Initial rNTPs
Nascent RNA
DNA-RNA hybrid region
Completed RNA strand

INITIATION

1 Polymerase binds to promoter sequence in duplex DNA. "Closed complex"

2 Polymerase melts duplex DNA near transcription start site, forming a transcription bubble. "Open complex"

3 Polymerase catalyzes phosphodiester linkage of two initial rNTPs.

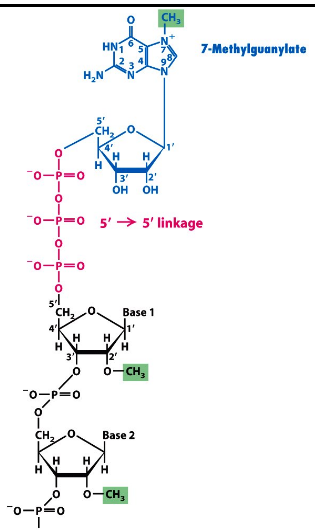
ELONGATION

4 Polymerase advances 3′→5′ down template strand, melting duplex DNA and adding rNTPs to growing RNA.

TERMINATION

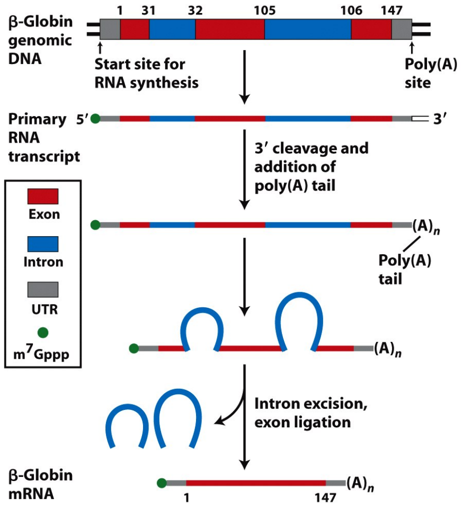
5 At transcription stop site, polymerase releases completed RNA and dissociates from DNA.

Structure of the 5' methylated cap of eukaryotic mRNA


7-Methylguanylate
CH_3
HN
N^+
H_2N
5'
CH_2
1'
4'
3'
2'
OH
OH
$^-O—P=O$
5' → 5' linkage
Base 1
Base 2
$O—CH_3$

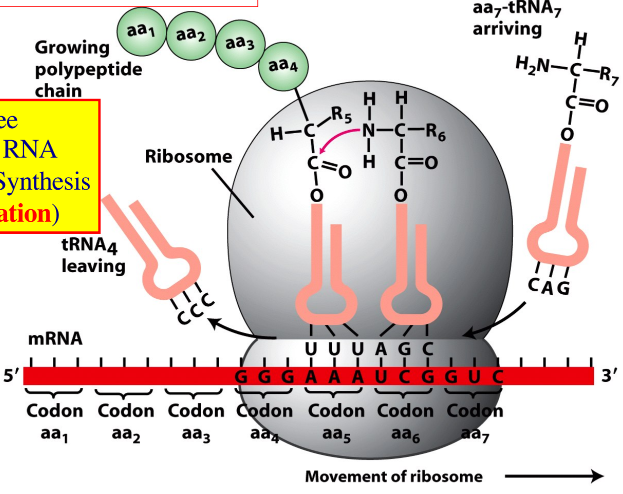
Processing:

Splicing: RNA processing to produce functional mRNA in eukaryotes


β-Globin genomic DNA
1
31
32
105
106
147
Start site for RNA synthesis
Poly(A) site
Primary RNA transcript
5'
3'
3' cleavage and addition of poly(A) tail
$(A)_n$
Poly(A) tail
Exon
Intron
UTR
m^7Gppp
Intron excision, exon ligation
β-Globin mRNA
1
147

mRNA, tRNA & rRNA

Three roles of RNA in Protein Synthesis (**Translation**)


aa_1
aa_2
aa_3
aa_4
Growing polypeptide chain
aa_7-$tRNA_7$ arriving
$H_2N—C—R_7$
R_5
R_6
Ribosome
$tRNA_4$ leaving
CCC
CAG
mRNA
5'
3'
U U U A G C
G G G A A A U C G G U C
Codon aa_1
Codon aa_2
Codon aa_3
Codon aa_4
Codon aa_5
Codon aa_6
Codon aa_7
Movement of ribosome

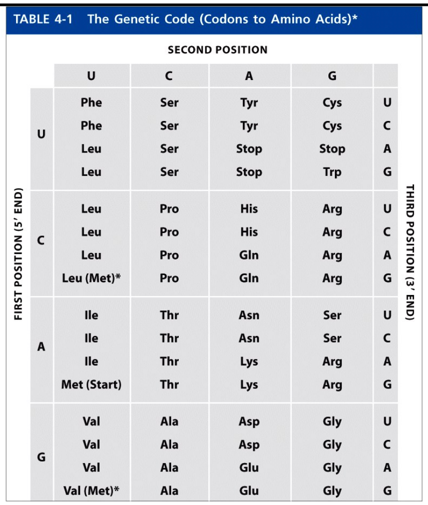


TABLE 4-1 The Genetic Code (Codons to Amino Acids)*

FIRST POSITION (5' END)	SECOND POSITION U	C	A	G	THIRD POSITION (3' END)
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	Stop	Stop	A
	Leu	Ser	Stop	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu (Met)*	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met (Start)	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val (Met)*	Ala	Glu	Gly	G

*AUG is the most common initiator codon; GUG usually codes for valine and CUG for leucine, but, rarely, these codons can also code for methionine to initiate a protein chain.

Cells are governed by a cellular chain of commands

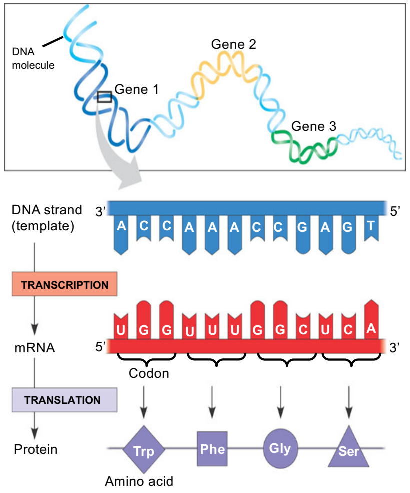
– **DNA —> RNA —> protein**

The Genetic Code

- Genetic **information** is encoded as a sequence of non-overlapping base triplets, or codons

Codons: Triplets of Bases

- How many bases correspond to an amino acid?
- During transcription the gene determines the sequence of bases along the length of an mRNA molecule

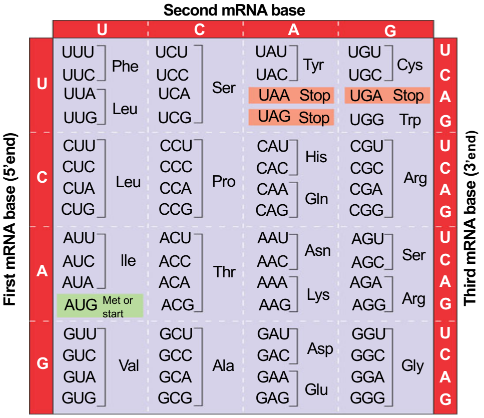

DNA molecule
Gene 1
Gene 2
Gene 3
DNA strand (template)
3'
A C C A A A C C G A G T
5'
TRANSCRIPTION
mRNA
5'
U G G U U U G G C U C A
3'
Codon
TRANSLATION
Protein
Trp
Phe
Gly
Ser
Amino acid

---

64 options

A codon in messenger RNA is either translated into an amino acid or serves as a translational stop signal

- Codons must be read in the correct reading frame for the specified polypeptide to be produced
- The genetic code is nearly universal shared by organisms from the simplest bacteria to the most complex animals



Second mRNA base

First mRNA base (5' end)	U		C		A		G		Third mRNA base (3' end)
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC		UCC		UAC		UGC		C
	UUA	Leu	UCA		UAA	Stop	UGA	Stop	A
	UUG		UCG		UAG	Stop	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Gln	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lys	AGA	Arg	A
	AUG	Met or start	ACG		AAG		AGG		G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glu	GGA		A
	GUG		GCG		GAG		GGG		G

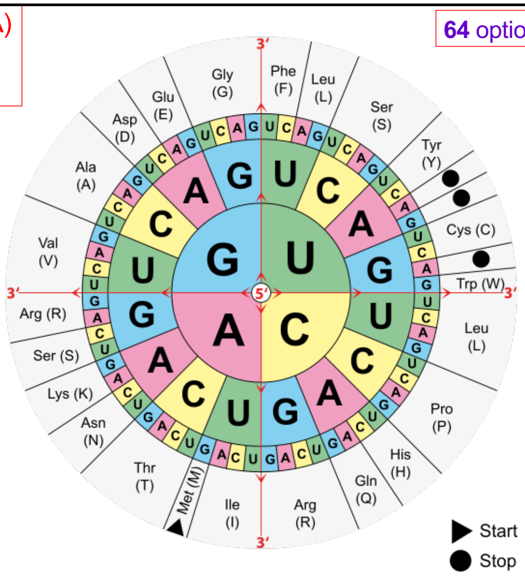
---

from genes (DNA) to proteins – the genetic code

64 options

Amino Acids

Genetic Code

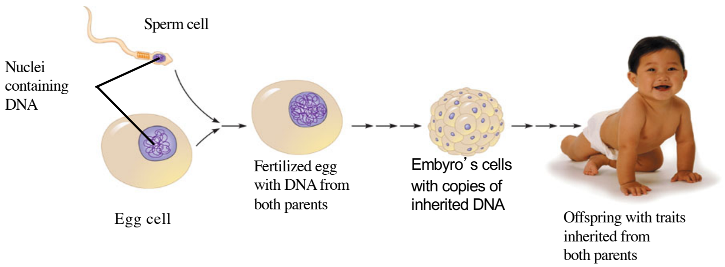

3'
5'
3'
3'
3'
G
U
A
C
Phe (F)
Leu (L)
Ser (S)
Tyr (Y)
Cys (C)
Trp (W)
Leu (L)
Pro (P)
His (H)
Gln (Q)
Arg (R)
Ile (I)
Met (M)
Thr (T)
Asn (N)
Lys (K)
Ser (S)
Arg (R)
Val (V)
Ala (A)
Asp (D)
Glu (E)
Gly (G)
▶ Start
● Stop

read 5' ---> 3'

The Cell's Heritable Information

Cells contain chromosomes made partly of DNA, the substance of genes

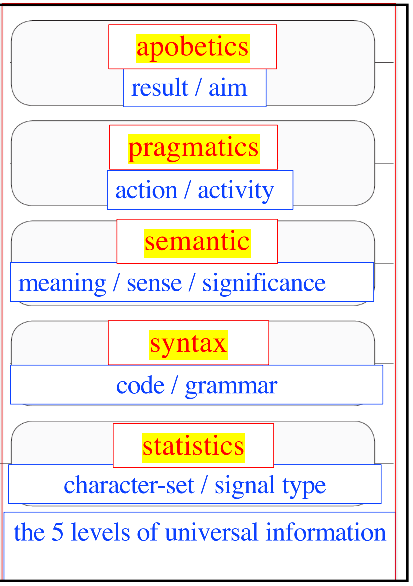
– which program the cells' production of proteins and transmit information from parents to offspring


Sperm cell
Nuclei containing DNA
Egg cell
Fertilized egg with DNA from both parents
Embryo's cells with copies of inherited DNA
Offspring with traits inherited from both parents

Universal Information is an abstract illustrated message (in the form of a code,) abstracted from coded symbols that contains expected action and intended goals.

1) Universal information is indispensable for all life and

2) for humans there are other non-material components of life: e.g. ‘will / volition’ and ‘consciousness’.



apobetics

result / aim

pragmatics

action / activity

semantic

meaning / sense / significance

syntax

code / grammar

statistics

character-set / signal type

the 5 levels of universal information

---

n: odd numbers not possible

with 20 amino acids, minimum info per word: $i_w = \log_2 20 = 4{,}32$ bit

Biological Information Store: The human DNA is about two meters long in an elongated form and contains 6 x 10^9 nucleotides (chemical letters: adenine [A], cytosine [C], guanine [G] and thymine [T]). How much statistical information (i.e. how much information according to Shannon) is that?

Since the N = 4 chemical letters A, C, G and T occur about equally distributed,

is their mean information content according to the equation $H = \mathrm{ld}\, 4 = \log_2 4 = \log 4 / \log 2 = 2$ bit calculable.

For the entire DNA this results in a storage capacity of $I_{total} = 6 \times 10^9$ nucleotides x 2 bits / nucleotide = 12 x 10^9 bits. This corresponds to the amount of information of 750,000 typewritten pages.

L = length of word = # of letters / word n = # of different letters		$L = 2$ doublet	$L = 3$ triplet	$L = 4$ quartet	$L = 5$ quintet	$L = 6$ sextet
		word length ***L*** --->				
Binärcode $n = 2$ $i_B = \mathrm{ld}\, n = 1$ bit	<--- alphabet length ***n***	$m = n^L = 4$ $i_w = L\, \mathrm{ld}\, n$ **C** 2 bit/Wort	$2^3 = 8$ **B** 3 bit/word	$2^4 = 16$ **A** 4 bit/word	$2^5 = 32$ **2** 5 bit/word	$2^4 = 64$ **1** 6 bit/word
Ternärcode $n = 3$ $i_B = 1{,}585$ bits		$3^2 = 9$ **D** 3,170	$3^3 = 27$ **3** 4,755	$3^4 = 81$ 6,340	$3^5 = 243$ 7,925	$3^6 = 729$ 9,510
Quaternärcode $n = 4$ $i_B = 2$ bit		$4^2 = 16$ **E** 4,0	$4^3 = 64$ **4** 6,0	$4^4 = 256$ 8,0	$4^5 = 1024$ 10,0	$4^6 = 4096$ 12,0
Quinärcode $n = 5$ $i_B = 2{,}322$ bit		$5^2 = 25$ **5** 4,644	$5^3 = 125$ 6,966	$5^4 = 625$ 9,288	$5^5 = 3125$ 11,610	$5^6 = 15625$ 13,932
Senärcode $n = 6$ $i_B = 2{,}585$ bit		$6^2 = 36$ **6** 5,170	$6^3 = 216$ 7,755	$6^4 = 1{,}296$ 10,340	$6^5 = 7776$ 12,925	$6^6 = 46656$ 20,680

$i_B = \mathrm{ld}\, n$ information content of *one* letter (bit/letter)
$i_w = L\, \mathrm{ld}\, n$ information content of *one* word (bit/word)
$m = n^L$ number of possible combinations to form a word of length L from n different characters.

The highest information density known to us is in the DNA molecules realized in the living cells. The figure shows the DNA dimensions.

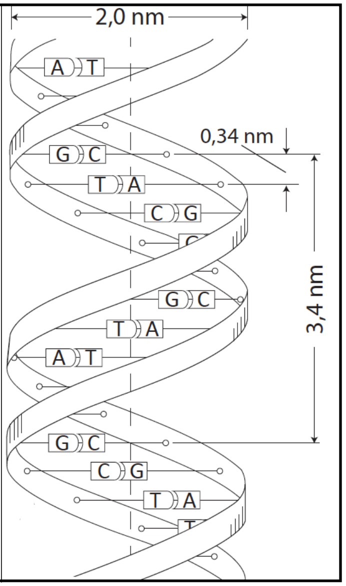
This chemical storage medium with 2 nm = 2 x 10^{-9} m = 2 x 10^{-6} mm diameter and 3.4 nm pitch of the helix (Greek helix = spiral) indicated. From this, the cylinder volume is calculated per turn ($V = h \times d^2 \times \pi / 4$):

$V = 3.4 \times 10^{-6}$ mm x (2×10^{-6} mm) 2 x π / 4 = 1.068×10^{-17} mm^3 per turn.

The human DNA (germ cell) has 3 x 10^9 character pairs. There, one turn comprises 10 pairs of letters, this results in a Volume of (1.068×10^{-17} mm3 / turn) x 3 x 10^9 letters / (10 letters /turn) = 3.2×10^{-9} mm^3.

Per turn of the double spiral are 10 chemical letters (nucleotides), so that a statistical information density from ϱ = 10 letters / ($10{,}68 \times 10^{-18}$ mm^3) = $0{,}94 \times 10^{18}$ letters / mm^3 = 0.94×10^{21} letters / cm^3.

If one considers the determined information requirement of 4.32 bit for one amino acid on a letter (nucleotide) of the genetic codes, that's how you get on with this method of calculation: 4.32: 3 = 1.44 bit / letter. We now want the statistical information density of the DNA with the values we have just established, and here we are for the information content per letter with the value 2 bit / letter : ϱ = 0.94×10^{18} letters / mm^3 Å ~ 2 bit / letter = 1.88×10^{18} bit / mm^3 or 1.88×10^{21} bit / cm^3.

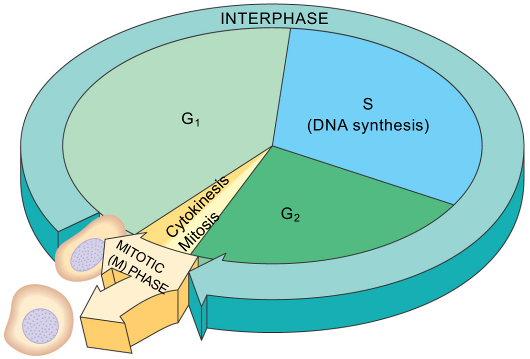


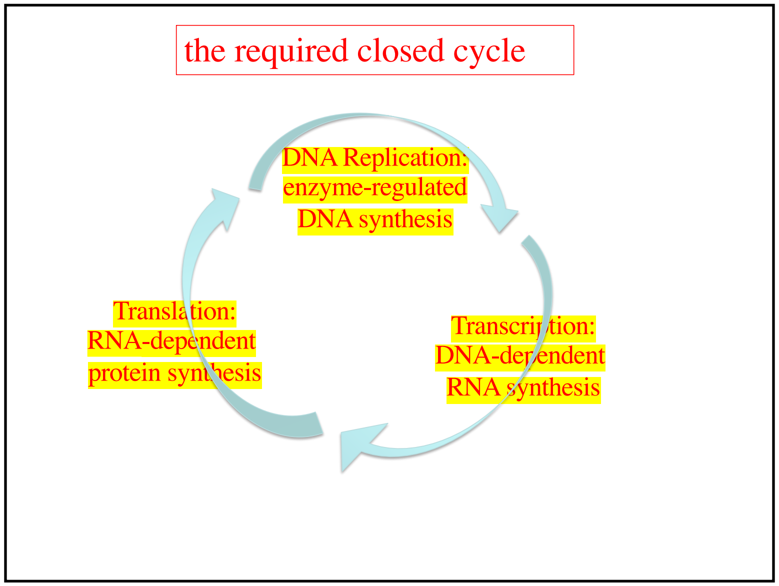
Phases of the Cell Cycle

The cell cycle consists of

- the Mitotic phase (M)
- Interphase (G1, S, G2)

- Interphase can be divided into sub-phases
  - G_1 phase
  - S phase
  - G_2 phase
- The mitotic phase
  - is made up of mitosis and cytokinesis



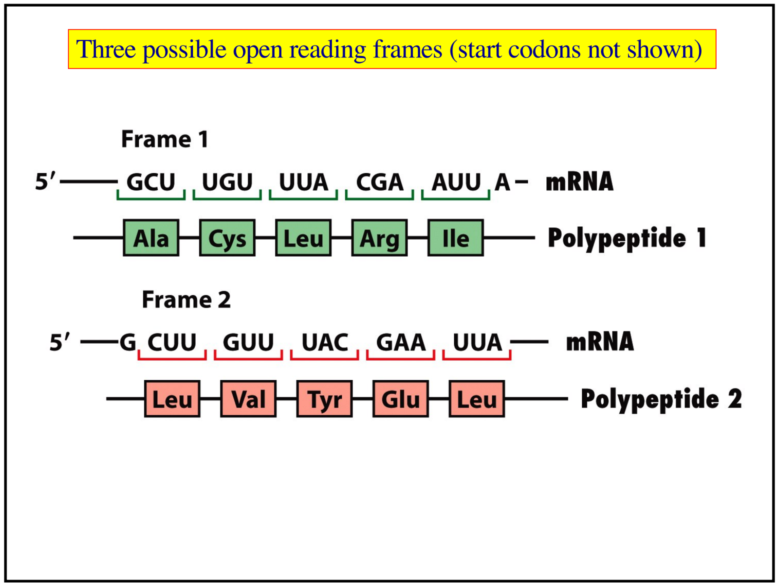


the required closed cycle

DNA Replication:
enzyme-regulated
DNA synthesis

Transcription:
DNA-dependent
RNA synthesis

Translation:
RNA-dependent
protein synthesis



Three possible open reading frames (start codons not shown)

Frame 1
5′ GCU UGU UUA CGA AUU A – mRNA
Ala Cys Leu Arg Ile Polypeptide 1
Frame 2
5′ G CUU GUU UAC GAA UUA mRNA
Leu Val Tyr Glu Leu Polypeptide 2

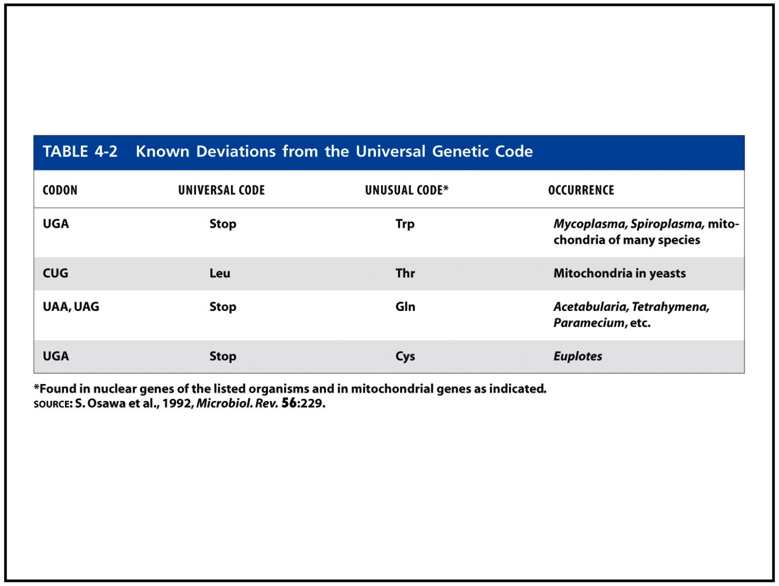
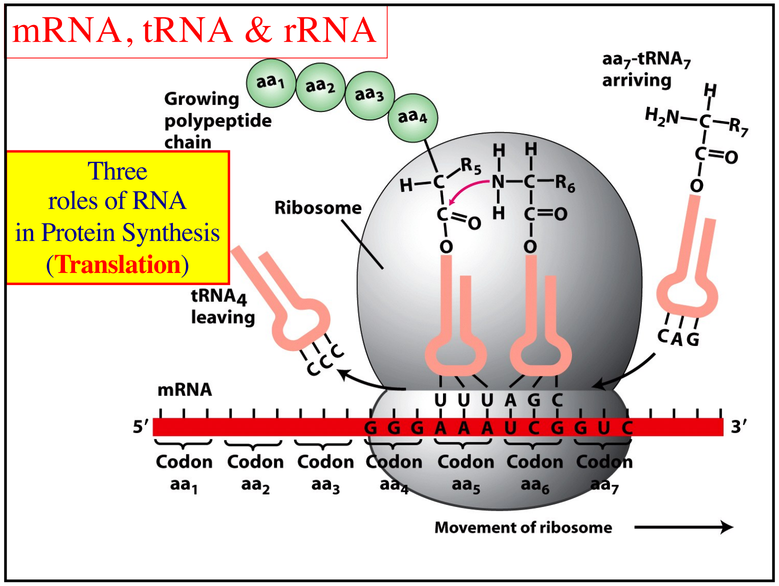


TABLE 4-2 Known Deviations from the Universal Genetic Code			
CODON	**UNIVERSAL CODE**	**UNUSUAL CODE***	**OCCURRENCE**
UGA	**Stop**	**Trp**	***Mycoplasma, Spiroplasma,*** **mitochondria of many species**
CUG	**Leu**	**Thr**	**Mitochondria in yeasts**
UAA, UAG	**Stop**	**Gln**	***Acetabularia, Tetrahymena, Paramecium,*** **etc.**
UGA	**Stop**	**Cys**	***Euplotes***

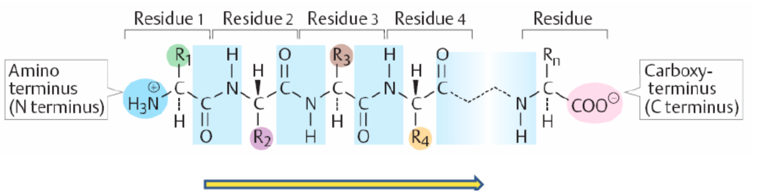
*Found in nuclear genes of the listed organisms and in mitochondrial genes as indicated.
SOURCE: S. Osawa et al., 1992, *Microbiol. Rev.* 56:229.



mRNA, tRNA & rRNA

Three
roles of RNA
in Protein Synthesis
(**Translation**)

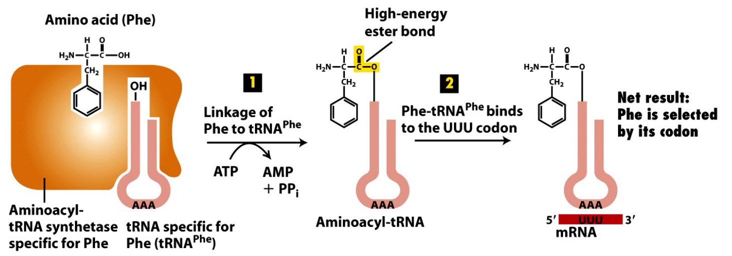
Growing polypeptide chain
aa1
aa2
aa3
aa4
aa7-tRNA7 arriving
H
H2N–C–R7
C=O
O
H–C–R5
C=O
O
H H
N–C–R6
H C=O
O
Ribosome
tRNA4 leaving
CCC
CAG
mRNA
U U U A G C
5′
G G G A A A U C G G U C
3′
Codon aa1
Codon aa2
Codon aa3
Codon aa4
Codon aa5
Codon aa6
Codon aa7
Movement of ribosome


Residue 1
Residue 2
Residue 3
Residue 4
Residue
Amino terminus (N terminus)
$H_3N^{\oplus}$
R_1
R_2
R_3
R_4
R_n
H
O
N
C
$COO^{\ominus}$
Carboxy-terminus (C terminus)

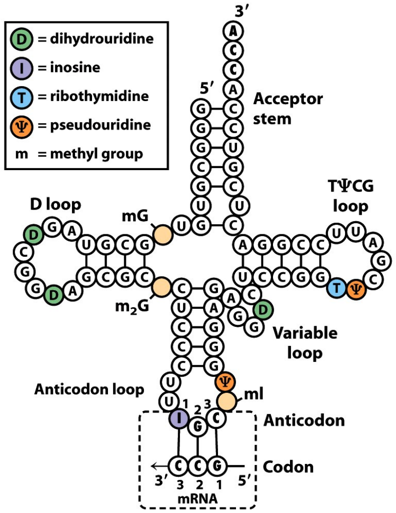
synthesis direction of the protein in ribosoms

plain, plane, layer, platform, planar

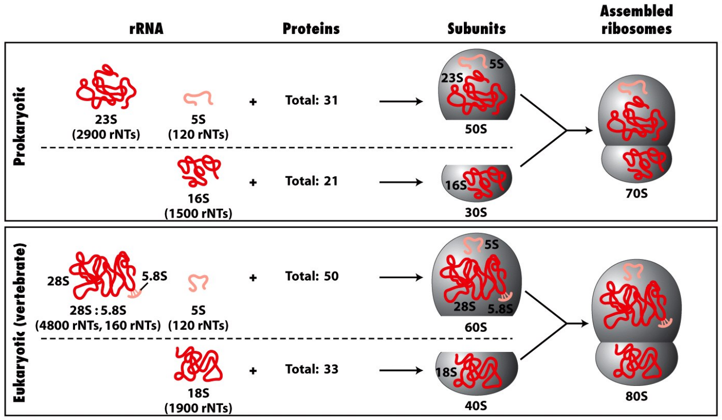
Specific Aminoacyl-tRNA Synthethases activate Amino Acids by covalently linking them to tRNAs


Amino acid (Phe)
OH
Aminoacyl-tRNA synthetase specific for Phe
tRNA specific for Phe (tRNA^Phe)
AAA
1
Linkage of Phe to tRNA^Phe
ATP
AMP + PPi
High-energy ester bond
Aminoacyl-tRNA
2
Phe-tRNA^Phe binds to the UUU codon
Net result: Phe is selected by its codon
5′
UUU
3′
mRNA

Structure of tRNAs


D = dihydrouridine
I = inosine
T = ribothymidine
Ψ = pseudouridine
m = methyl group
3′
5′
Acceptor stem
D loop
mG
TΨCG loop
m₂G
Variable loop
Anticodon loop
ml
Anticodon
Codon
3′ 3 2 1 5′
mRNA

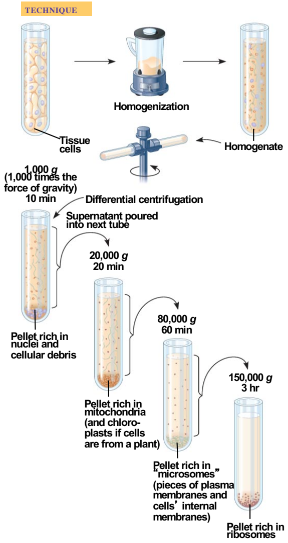
The general structure of ribosomes in prokaryotes and eukaryotes


rRNA
Proteins
Subunits
Assembled ribosomes
Prokaryotic
23S (2900 rNTs)
5S (120 rNTs)
+
Total: 31
23S
5S
50S
16S (1500 rNTs)
Total: 21
16S
30S
70S
Eukaryotic (vertebrate)
28S
5.8S
28S : 5.8S (4800 rNTs, 160 rNTs)
5S (120 rNTs)
Total: 50
5S
28S
5.8S
60S
18S (1900 rNTs)
Total: 33
18S
40S
80S

Lab methods

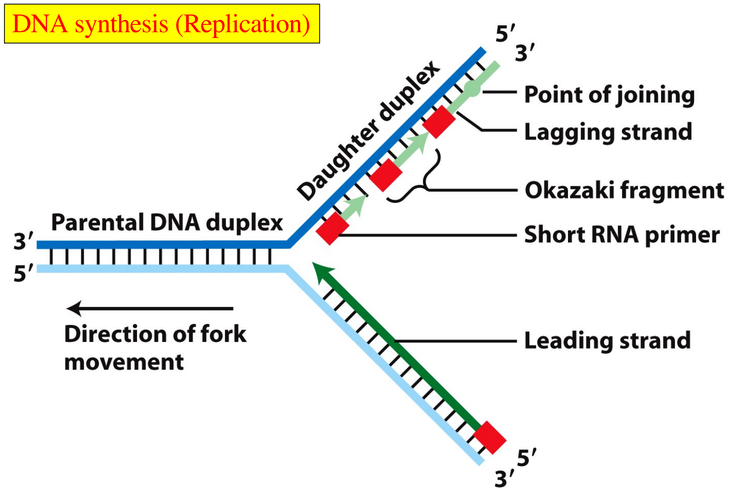
for

Ribosome isolation

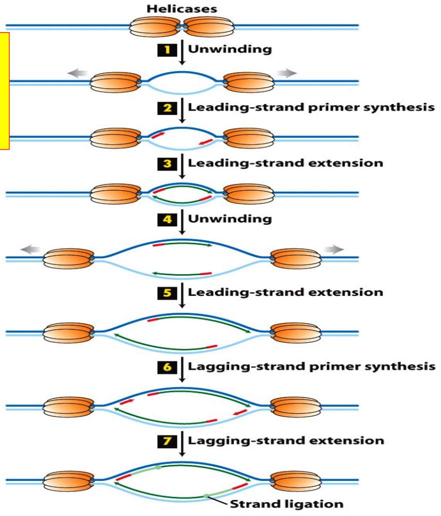

TECHNIQUE
Tissue cells
Homogenization
Homogenate
1,000 *g* (1,000 times the force of gravity) 10 min
Differential centrifugation
Supernatant poured into next tube
20,000 *g* 20 min
80,000 *g* 60 min
150,000 *g* 3 hr
Pellet rich in nuclei and cellular debris
Pellet rich in mitochondria (and chloroplasts if cells are from a plant)
Pellet rich in "microsomes" (pieces of plasma membranes and cells' internal membranes)
Pellet rich in ribosomes

Ph/CHCl3 – H_2O-buffer
Separation of Proteins
vs
28S rRNA
18S rRNA
5S rRNA

DNA synthesis (Replication)


5′
3′
Daughter duplex
Point of joining
Lagging strand
Okazaki fragment
Short RNA primer
Parental DNA duplex
3′
5′
Direction of fork movement
Leading strand
5′
3′

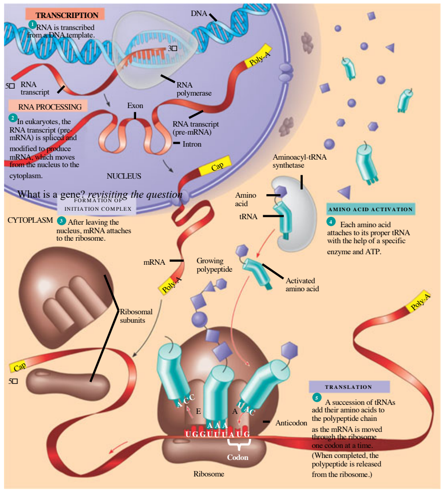
Bidirectional mechanism of DNA replication


Helicases
1 Unwinding
2 Leading-strand primer synthesis
3 Leading-strand extension
4 Unwinding
5 Leading-strand extension
6 Lagging-strand primer synthesis
7 Lagging-strand extension
Strand ligation

What is a gene? *revisiting the question*

- A gene is a region of DNA whose final product is either a polypeptide or an RNA molecule –
- A gene is the molecular unit of heredity (= DNA // or RNA) of a living organism that codes for a protein or a (functional) RNA.

A summary of transcription and translation in a eukaryotic cell


TRANSCRIPTION
1 RNA is transcribed from a DNA template.
DNA
3′
5′ RNA transcript
RNA polymerase
Poly-A
RNA PROCESSING
2 In eukaryotes, the RNA transcript (pre-mRNA) is spliced and modified to produce mRNA, which moves from the nucleus to the cytoplasm.
Exon
RNA transcript (pre-mRNA)
Intron
NUCLEUS
Cap
What is a gene? revisiting the question
FORMATION OF INITIATION COMPLEX
CYTOPLASM
3 After leaving the nucleus, mRNA attaches to the ribosome.
Aminoacyl-tRNA synthetase
Amino acid
tRNA
AMINO ACID ACTIVATION
4 Each amino acid attaches to its proper tRNA with the help of a specific enzyme and ATP.
mRNA
Growing polypeptide
Activated amino acid
Poly-A
Ribosomal subunits
Cap
5′
TRANSLATION
5 A succession of tRNAs add their amino acids to the polypeptide chain as the mRNA is moved through the ribosome one codon at a time. (When completed, the polypeptide is released from the ribosome.)
A C C
E
A
U A C
A A A
U G G U U U A U G
Anticodon
Codon
Ribosome
Poly-A