

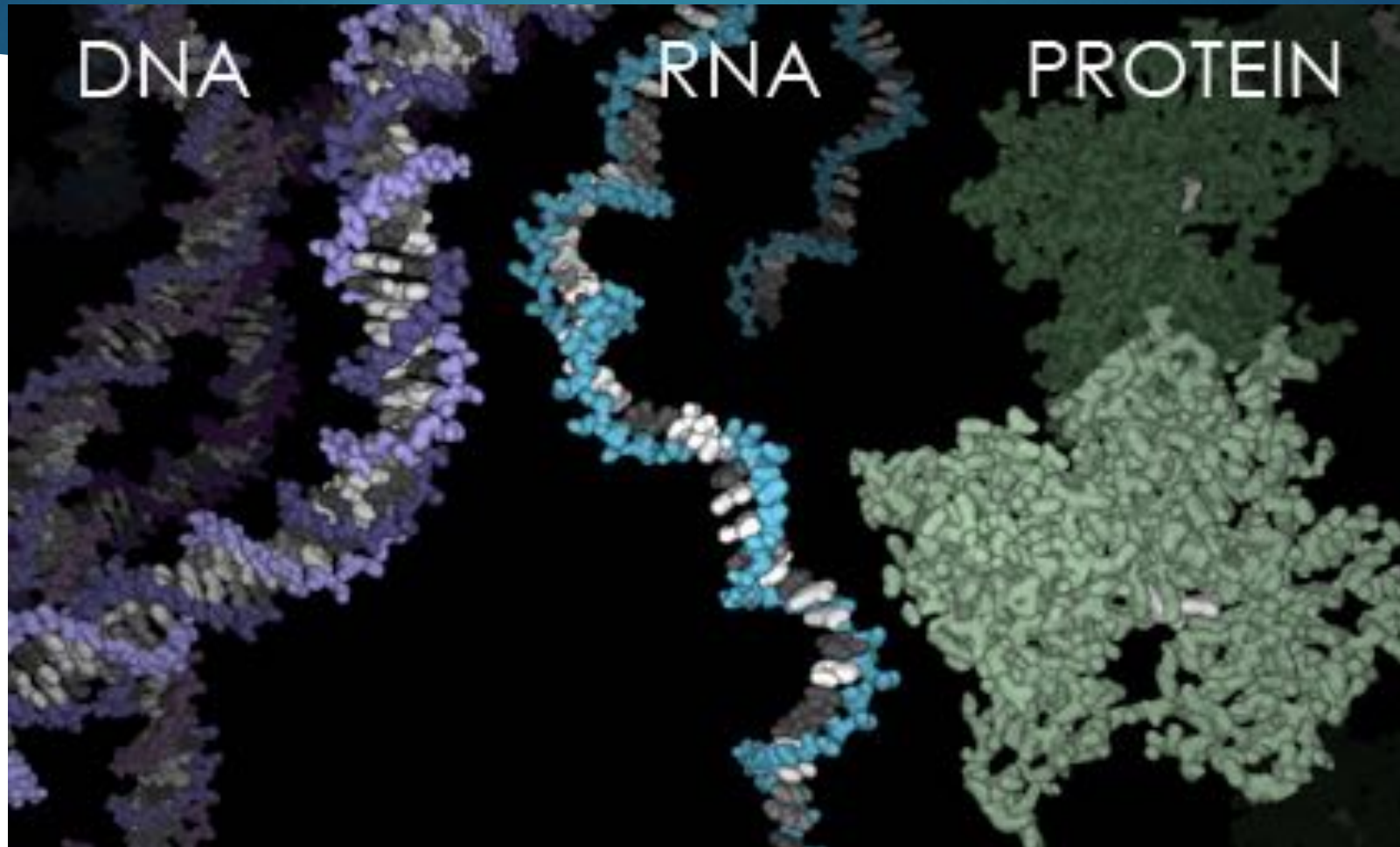


CAP 5510/CGS 5166:
Bioinformatics &
Bioinformatic Tools

GIRI NARASIMHAN, SCIS, FIU

Molecular Biology Background

Drama of Mol. Biology ... the actors



Polymeric Actors

Molecule	Unit Name	Unit Composition
DNA	Nucleotide or Base	A, C, G, T
RNA	Nucleotide or Base	A, C, G, U
Protein	Amino acid residues	amino acids represented by 20-letter alphabet missing {B, J, O, U, X, Z}

Typical DNA Sequence

```
1 gggagaacac cgggagaagg aggaggaggc gaagaaaagc aacagaagcc cagttgctgc
61 tccaggtccc tcggacagag ctttttccat gtggagactc tctcaatgga cgtgccccct
121 agtgcttctt agacggactg cggctctccta aaggctcgacc atggtggccg ggaccgcctg
181 tcttctagtg ttgctgcttc cccaggtcct cctgggcggc gcggccggcc tcattccaga
241 gctgggcccgc aagaagtctg ccgcggcatc cagccgacct ttgtcccggc cttcgggaaga
301 cgtcctcagc gaatttgagt tgaggctgct cagcatgttt ggcctgaagc agagaccac
361 ccccagcaag gacgtcgtgg tgccccctta tatgctagat ctgtaccgca ggcactcagg
421 ccagccagga gcgcccgcc cagaccaccg gctggagagg gcagccagcc gcgccaacac
481 cgtgcgccagc ttccatcacg aagaagccgt ggaggaactt ccagagatga gtgggaaaac
541 ggcccggcgc ttcttttca atttaagttc tgtccccagt gacgagtttc tcacatctgc
601 agaactccag atcttccggg aacagataca ggaagctttg ggaaacagta gtttccagca
661 ccgaattaat atttatgaaa ttataaagcc tgcagcagcc aacttgaaat ttctgtgac
721 cagactattg gacaccaggt tagtgaatca gaacacaagt cagtgggaga gcttcgacgt
781 caccaccagct gtgatgcggt ggaccacaca gggacacacc aaccatgggt ttgtggtgga
841 agtggcccat ttagaggaga acccaggtgt ctccaagaga catgtgagga ttagcaggtc
901 tttgcaccaa gatgaacaca gctggtcaca gataaggcca ttgctagtga cttttggaca
961 tgatggaaaa ggacatccgc tccacaaacy agaaaagcgt caagccaaac acaaacagcg
```

The building blocks of DNA & RNA

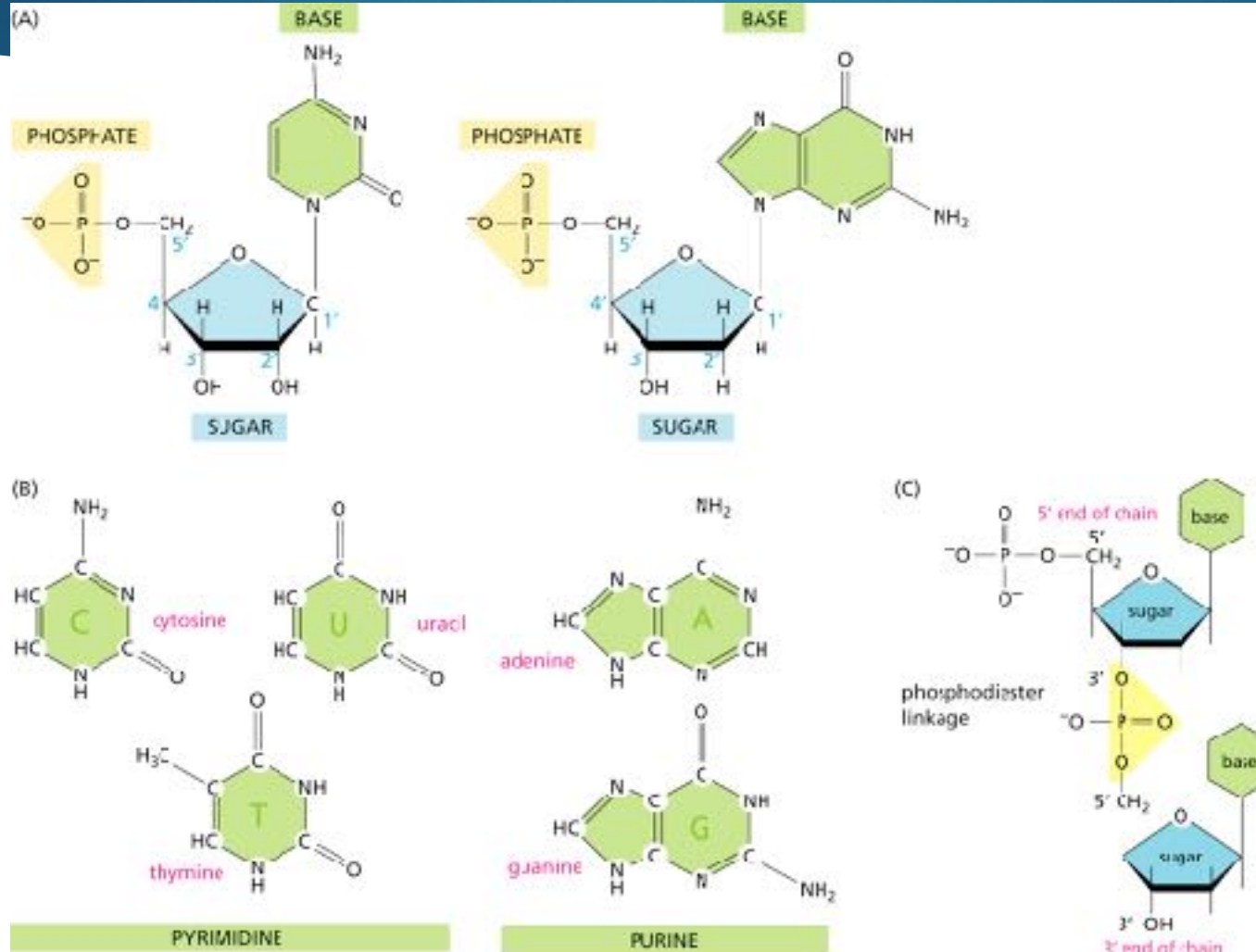


Fig 1.1, Zvelebil/Baum

DNA double helix structure

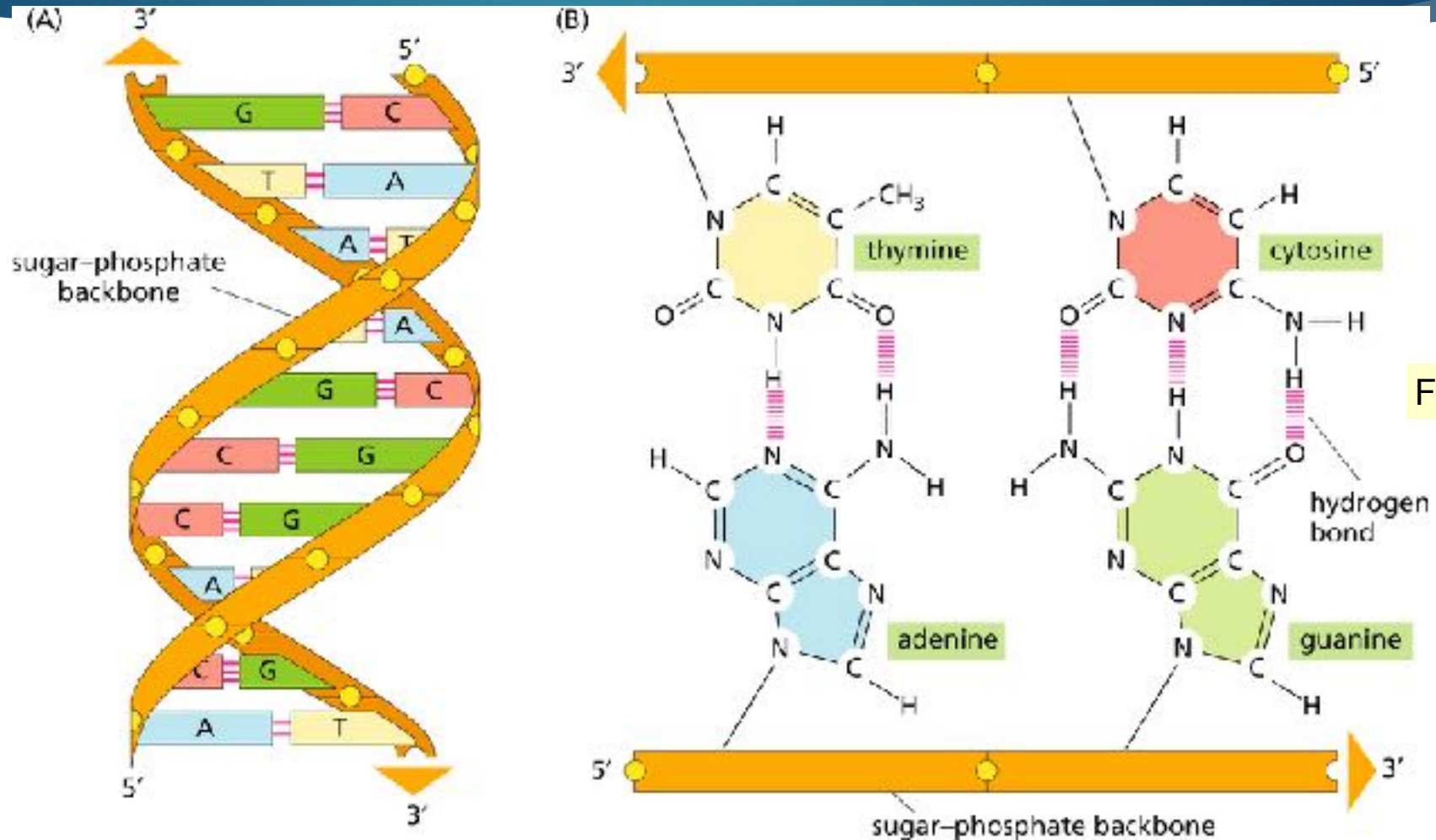


Fig 1.3, Zvelebil/Baum

RNA molecule

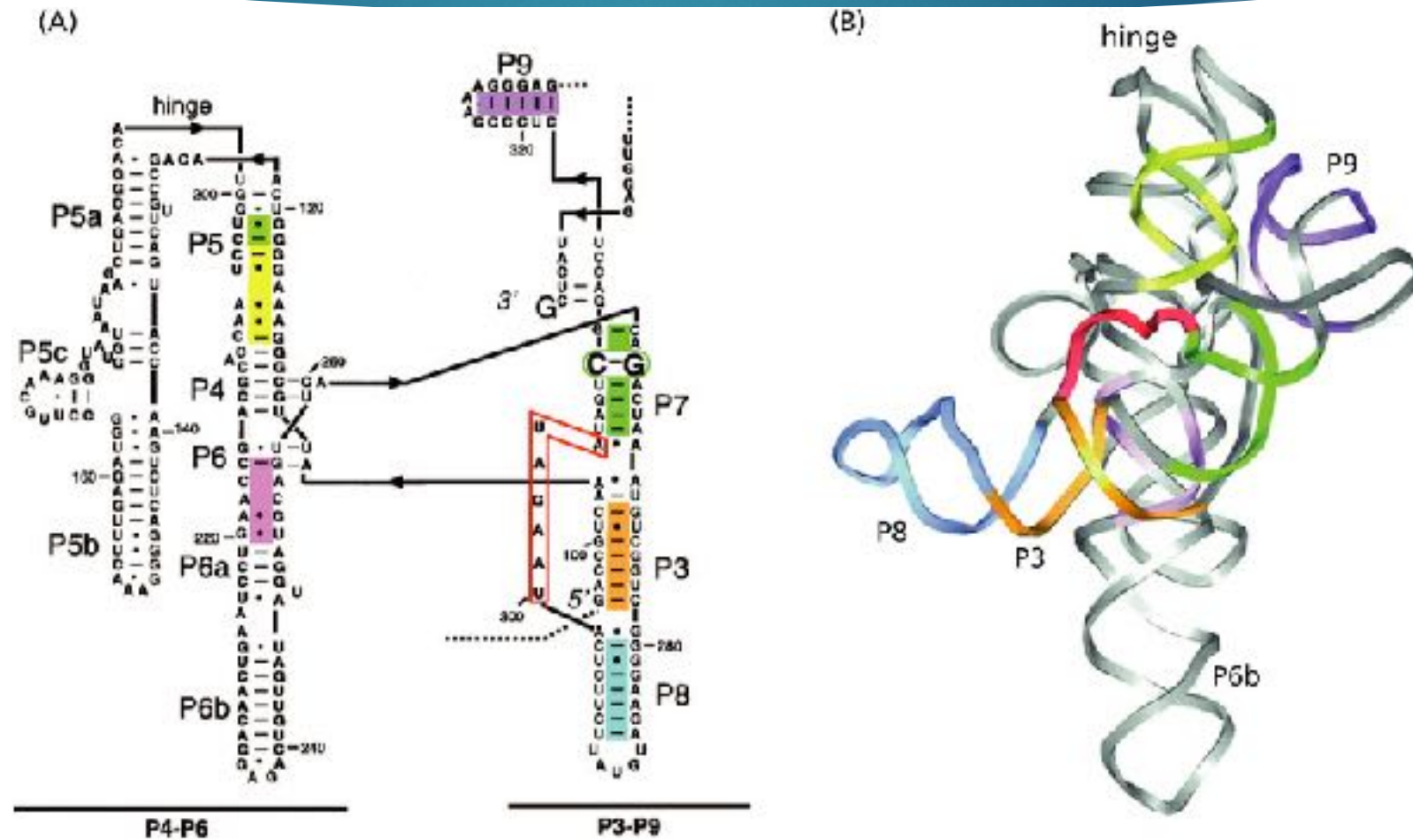


Fig 1.5, Zvelebil/Baum

Proteins: Amino acids

Letter Code	3 Letter Code	Amino Acid	Letter Code	3 Letter Code	Amino Acid
A	Ala	Alanine	M	Met	Methionine
C	Cys	Cysteine	N	Asn	Asparagine
D	Asp	Aspartic Acid	P	Pro	Proline
E	Glu	Glutamic Acid	Q	Gla	Glutamine
F	Phe	Phenylalanine	R	Arg	Arginine
G	Gly	Glycine	S	Ser	Serine
H	His	Histidine	T	Thr	Threonine
I	Ile	Isoleucine	V	Val	Valine
K	Lys	Lysine	W	Trp	Tryptophan
L	Leu	Leucine	Y	Tyr	Tyrosine

Typical protein sequence

```
/translation="MVAGTRCLLVLLLPQVLLGGAAGLIPELGRKKFAAASSRPLSRP  
SEDVLSEFELRLLSMFGLKQRPTPSKDVVVPPYMLDLYRRHSGQPGAPAPDHLERAA  
SRANTVRSFHHEEAVEELPEMSGKTARRFFFNLSSVPSDEFLLTSAELQIFREQIQEAL  
GNSSFQHRINIYEI IKPAAANLKF PVTRLLDTRLVNQNTSQWESFDVTPAVMRWTTQG  
HTNHGFVVEVAHLEENPGVSKRHVRI SRSLHQDEHSWSQIRPLLVTFGHDGKGHPLHK  
REKRQAKHKQRKRLKSSCKRHPLYVDFSDVGWNDWIVAPPGYHAFYCHGECPFPLADH  
LNSTNHAI VQTLVNSVNSKIPKACCVPTELSAISMLYLDENEKVVLKKNYQDMVVEGCG  
CR"
```

Missing letters of the alphabet: B, J, O, U, X, Z

Protein 3D Structure

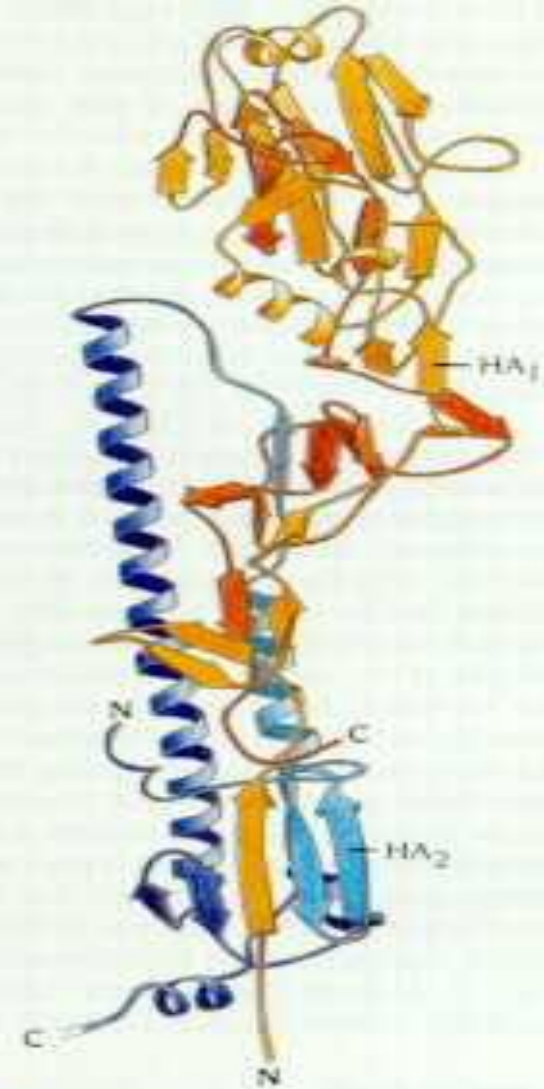
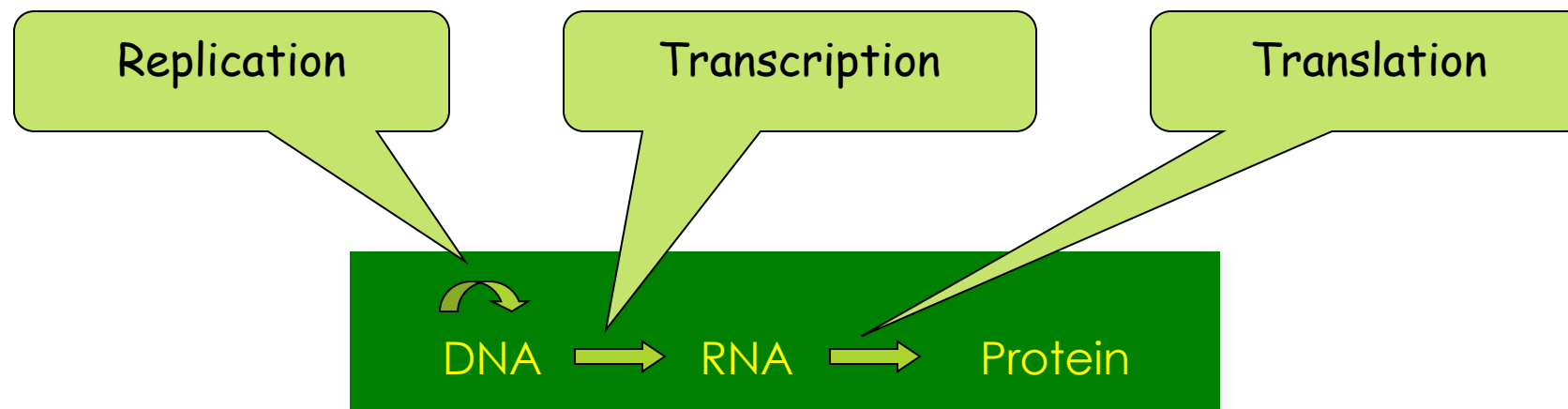


Figure 5.21 Schematic diagram of the subunit structure of hemagglutinin from influenza virus. The structure comprises about 550 amino acids arranged in two chains HA₁ (red) and HA₂ (blue). The first half of each chain has a lighter color in the diagram. The subunit is very elongated with a long stemlike region built up by residues from both chains and includes one of the longest α helices known in a globular structure, about 75 Å long. The globular head is formed by residues only from HA₂. (Courtesy of Don Wiley, Harvard University.)

Central Dogma

- ▶ DNA acts as a template to replicate itself.
- ▶ DNA is transcribed into RNA.
- ▶ RNA is translated into **Protein**.



Central Dogma

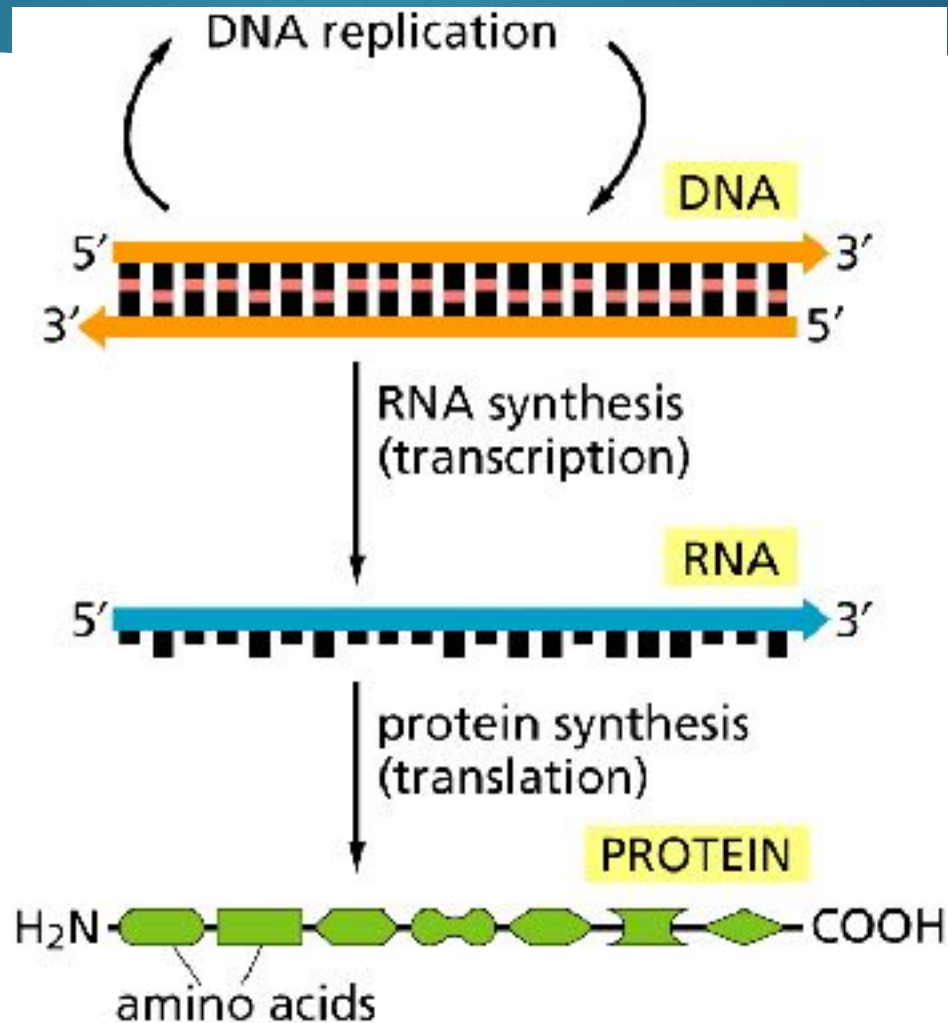


Fig 1.6, Zvelebil/Baum

DNA Replication

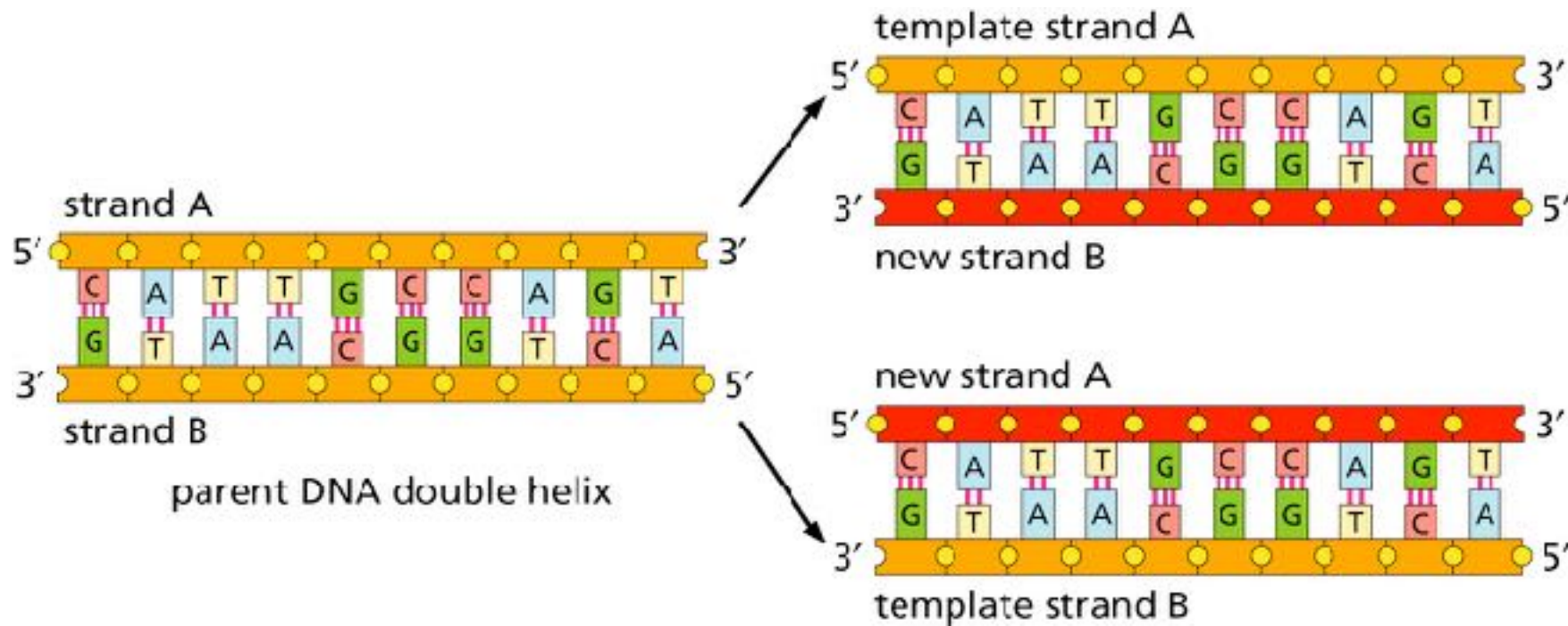
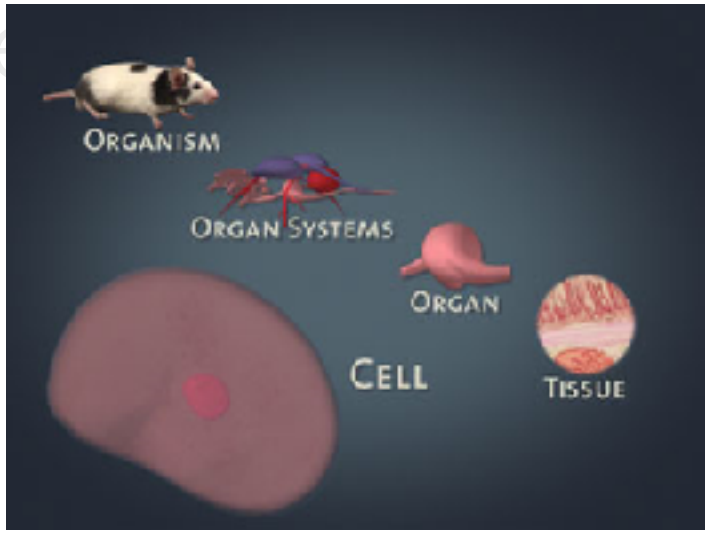
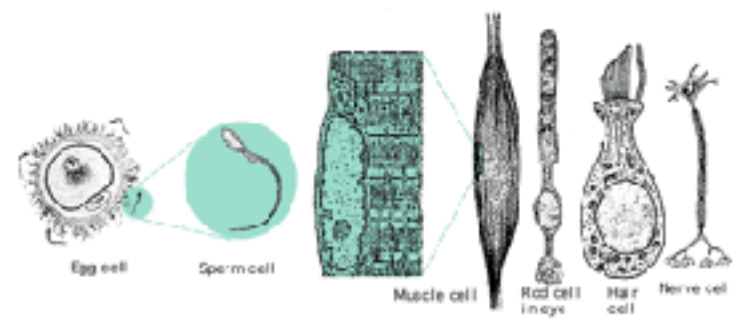


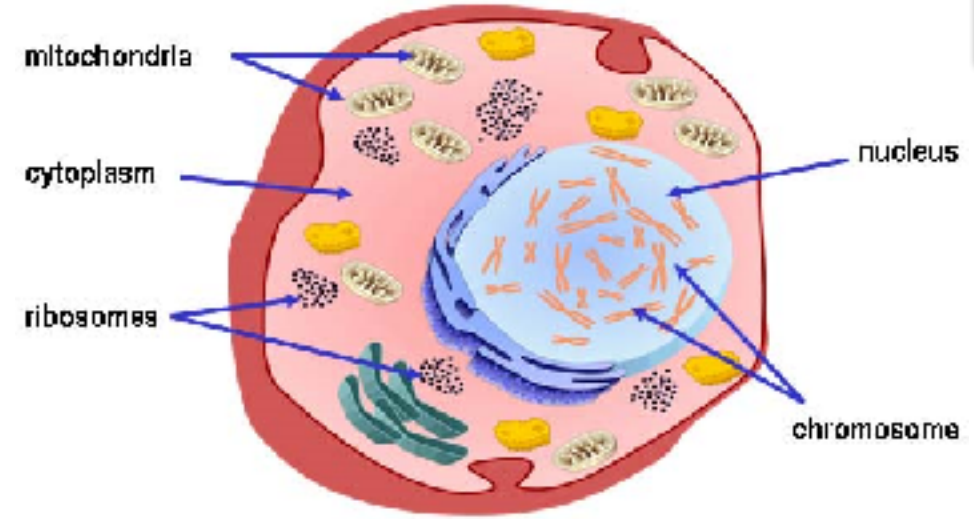
Fig 1.4, Zvelebil/Baum



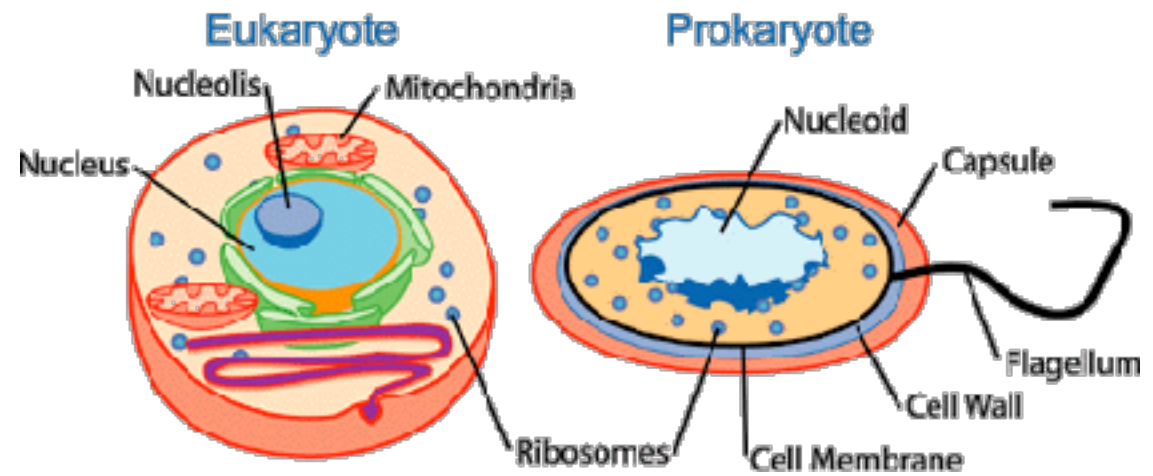
<http://www.learner.org/channel/courses/essential/life/session1/closer1.html>



<http://www.biology.eku.edu/RITCHISO/301notes1.htm>



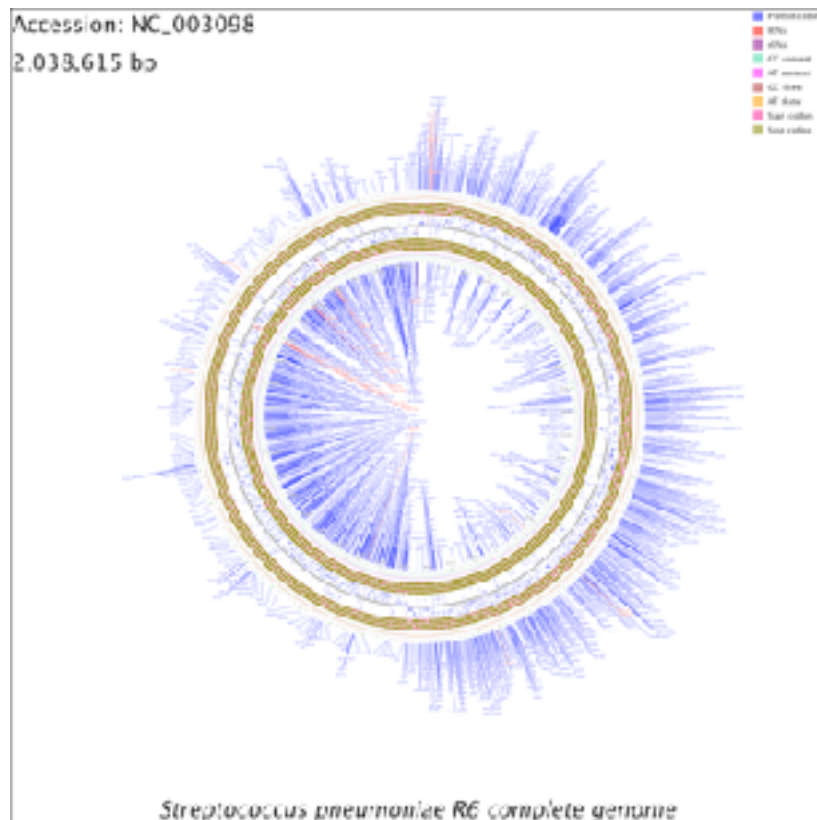
http://www.biotechnologyonline.gov.au/popups/img_cellwithlabels.cfm



<http://en.wikipedia.org/wiki/File:Celltypes.png>

http://www.cellsalive.com/cells/cell_model.htm

Chromosomes: Bacterial & Human



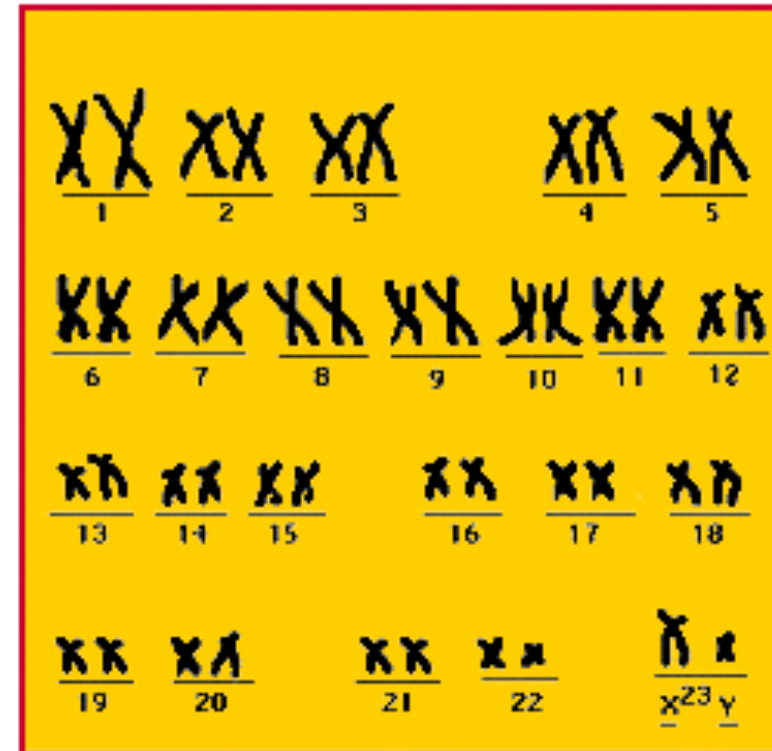
Human chromosomes!

centromere



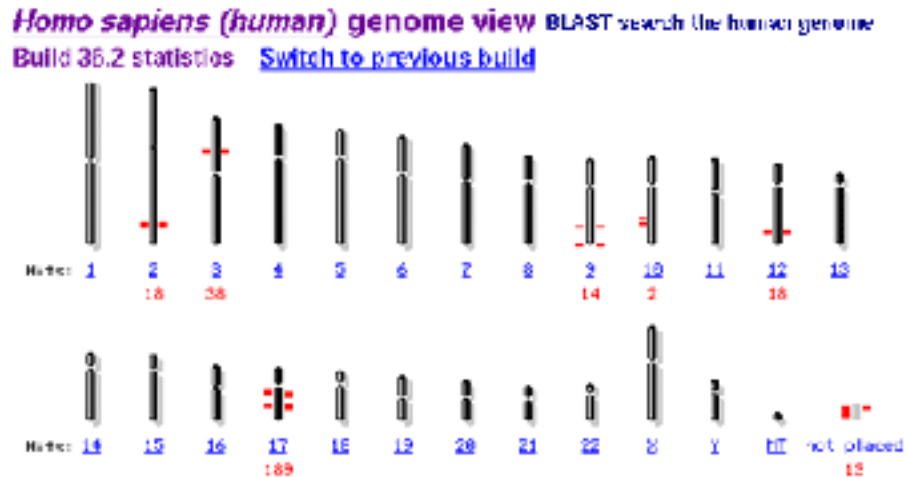
a

chromatid

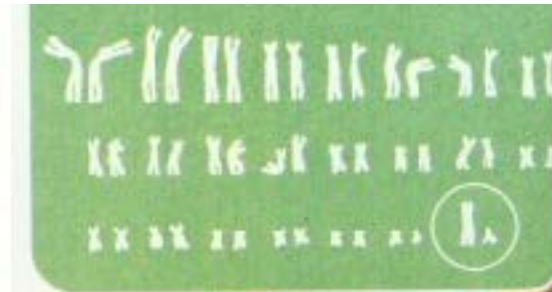


b

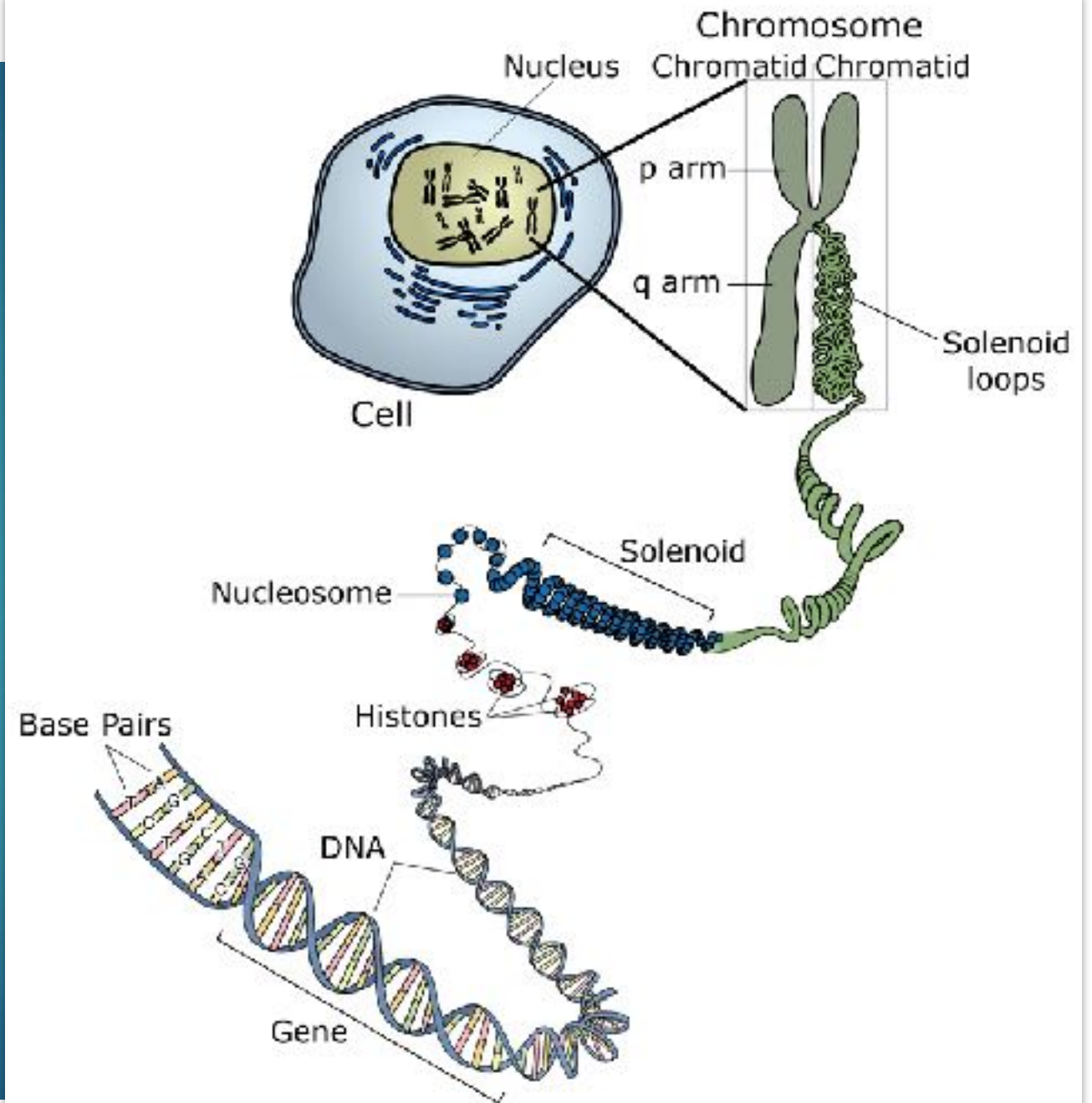
Chromosomes



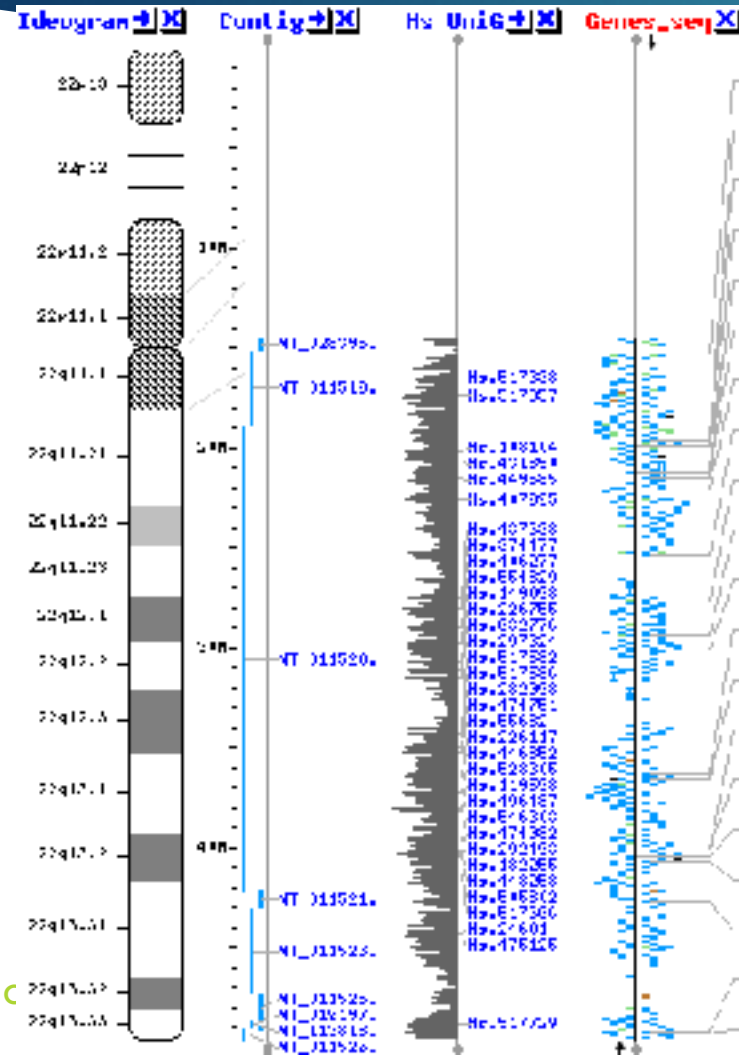
The chromosomal locations of several genes believed to be associated with the human BRCA1 gene implicated in breast cancer are highlighted.



From Cell to Genes



Human Chr 22



Symbol Position Description

ABCD1P4

22q11 ATP-binding cassette, sub-family D (ALD)

SNAP29

22q11.21

synaptosomal-associated protein

-
-
-

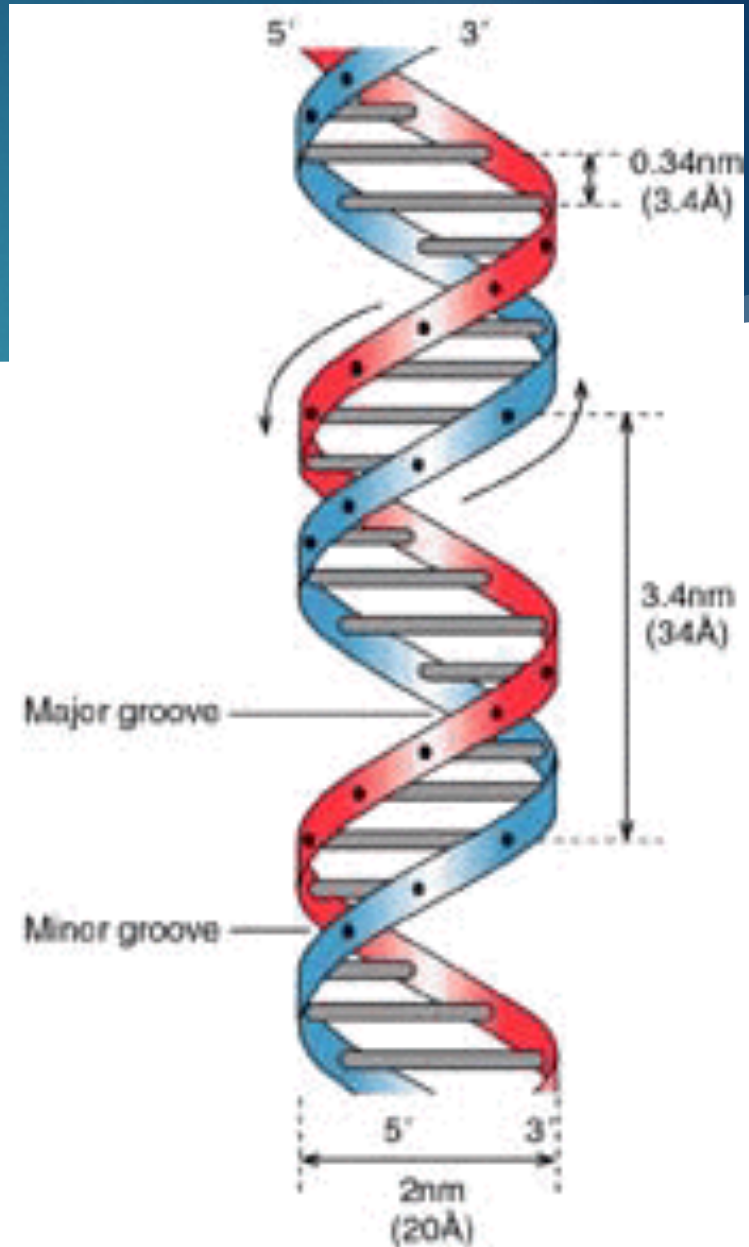
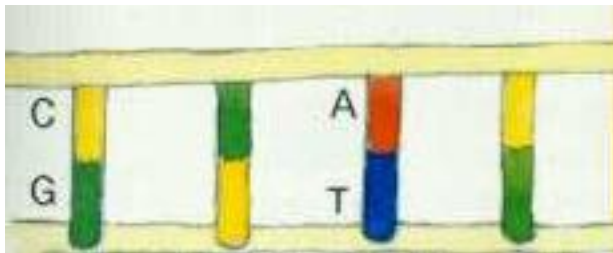
2nd smallest chr – 50MB

855 genes

First chr to be sequenced (1999)

DNA Molecule

Complementary Bases



Genes



Basic Genetic Processes

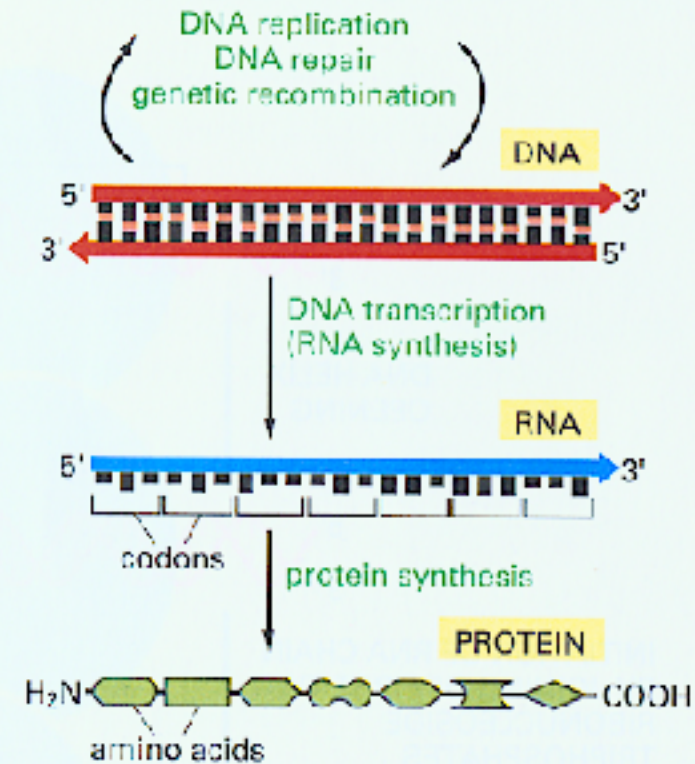
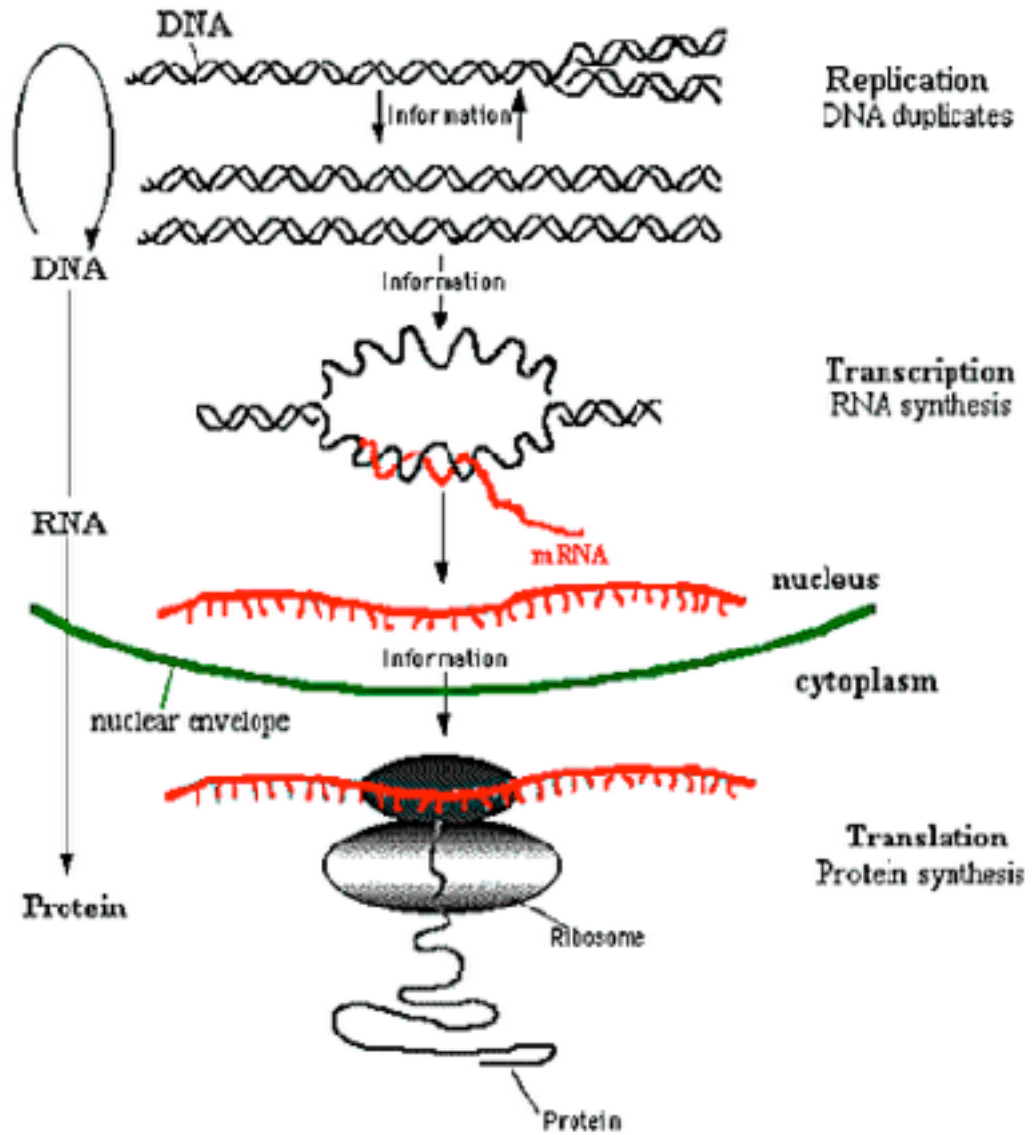


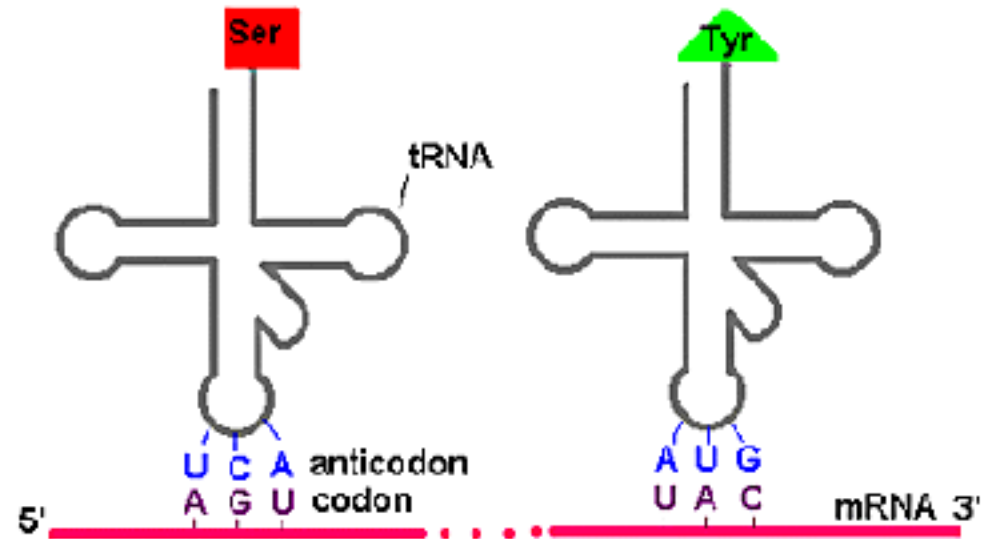
Figure 6–1 The basic genetic processes. The processes shown here are thought to occur in all present-day cells. Very early in the evolution of life, however, much simpler cells probably existed that lacked both DNA and proteins (see Figure 1–11). Note that a sequence of three nucleotides (a codon) in an RNA molecule codes for a specific amino acid in a protein.

Basic Genetic Processes



The Central Dogma of Molecular Biology

RNA and Genetic Code



		2nd base in codon					
		U	C	A	G		
1st base in codon	U	Phe Phe Leu Leu	Ser Ser Ser Ser	Tyr Tyr STOP STOP	Cys Cys STOP Trp	U C A G	3rd base in codon
	C	Leu Leu Leu Leu	Pro Pro Pro Pro	His His Gln Gln	Arg Arg Arg Arg	U C A G	
	A	Ile Ile Ile Met	Thr Thr Thr Thr	Asn Asn Lys Lys	Ser Ser Arg Arg	U C A G	
	G	Val Val Val Val	Ala Ala Ala Ala	Asp Asp Glu Gln	Gly Gly Gly Gly	U C A G	

The Genetic Code

Proteins and their functions

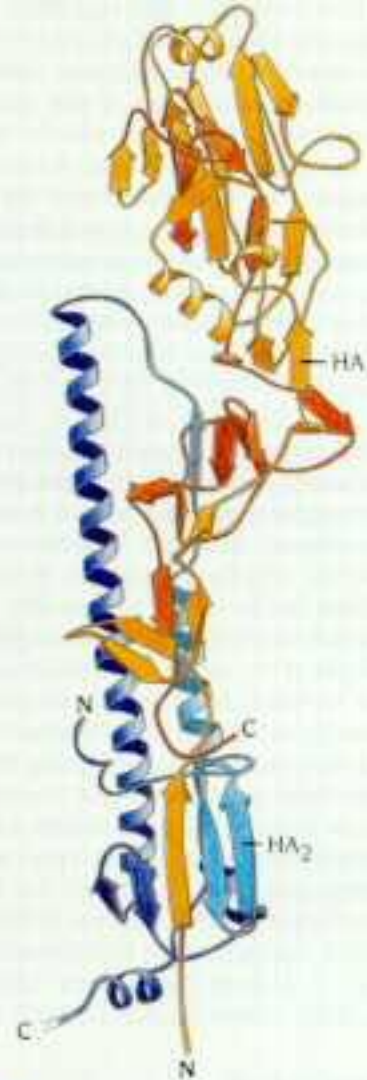
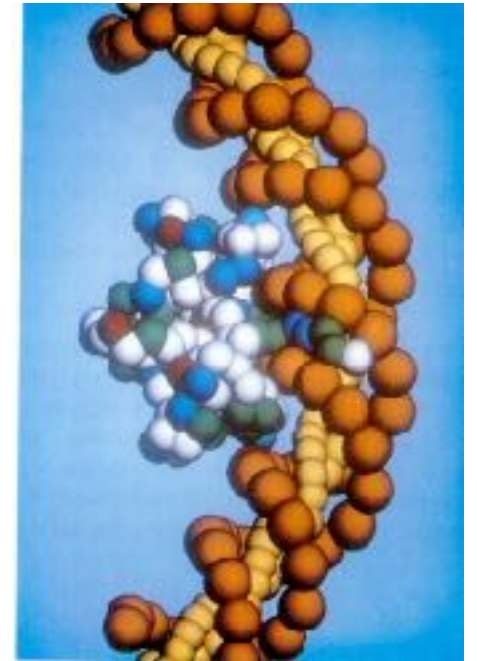
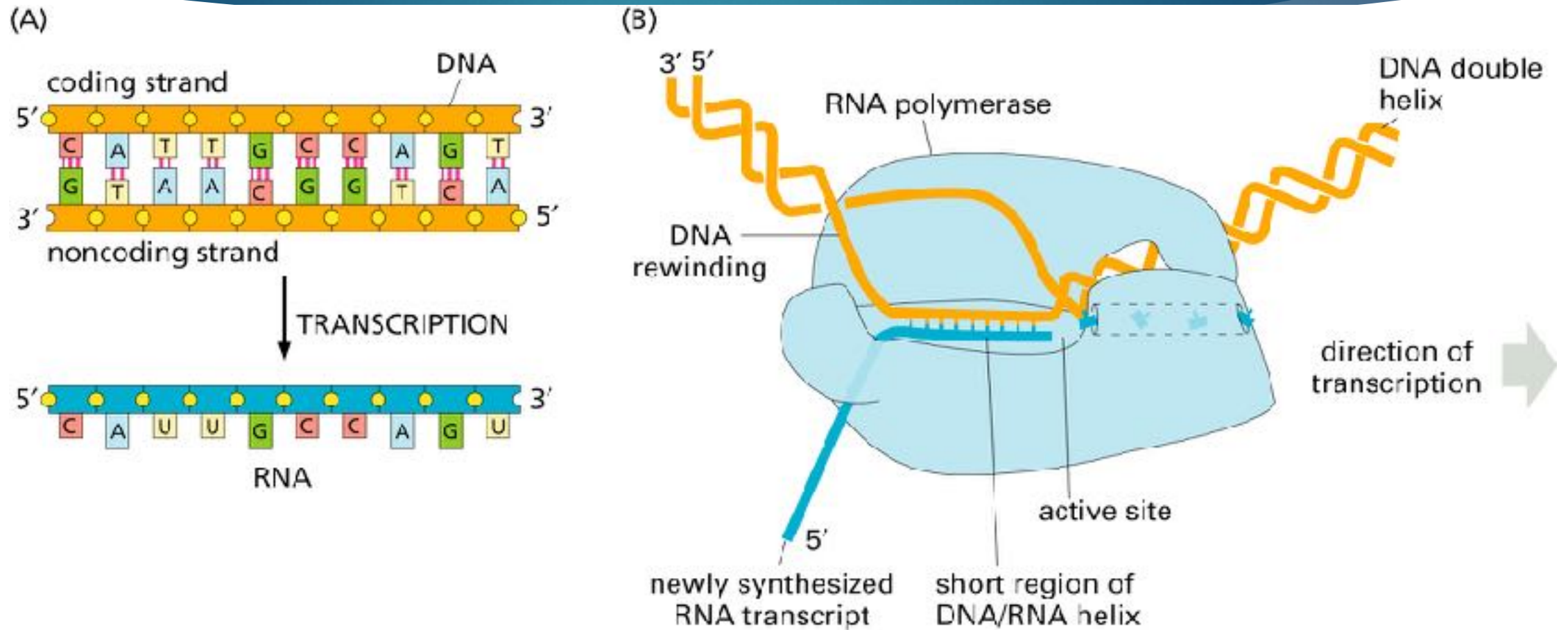


Figure 5.21 Schematic diagram of the subunit structure of hemagglutinin from influenza virus. The structure comprises about 550 amino acids arranged in two chains HA₁ (red) and HA₂ (blue). The first half of each chain has a lighter color in the diagram. The subunit is very elongated with a long stemlike region built up by residues from both chains and includes one of the longest α helices known in a globular structure, about 75 Å long. The globular head is formed by residues only from HA₁. (Courtesy of Don Wiley, Harvard University.)

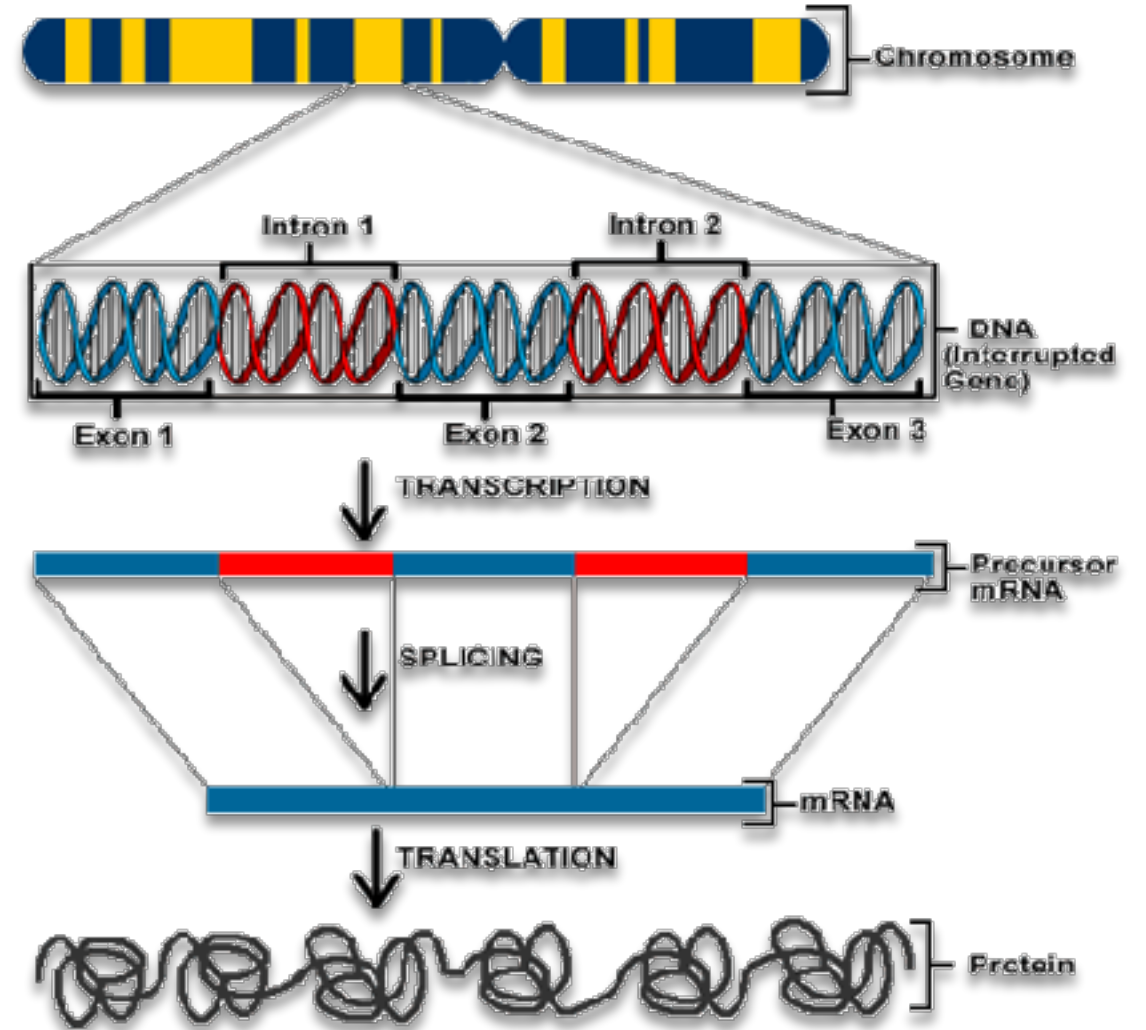


Transcription

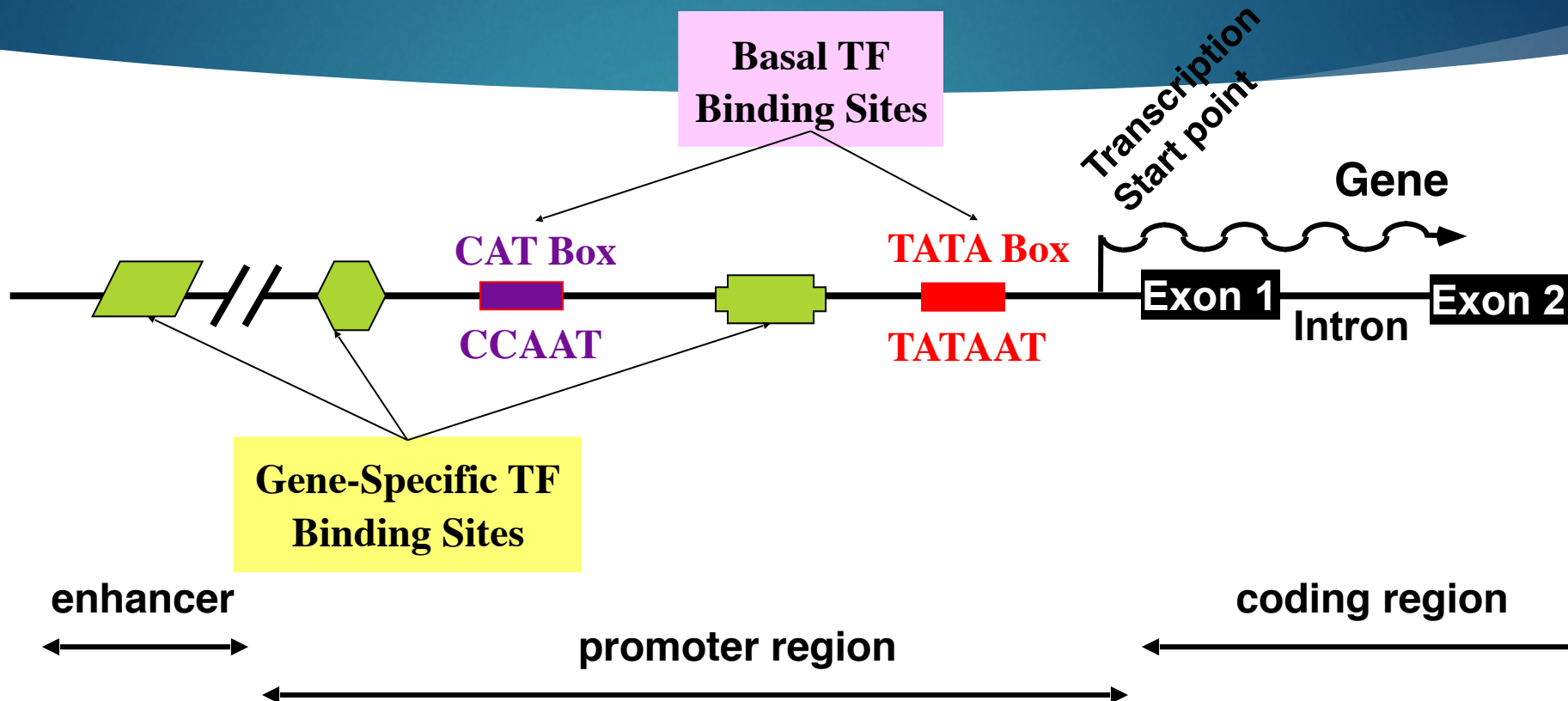
Fig 1.7, Zvelebil/Baum



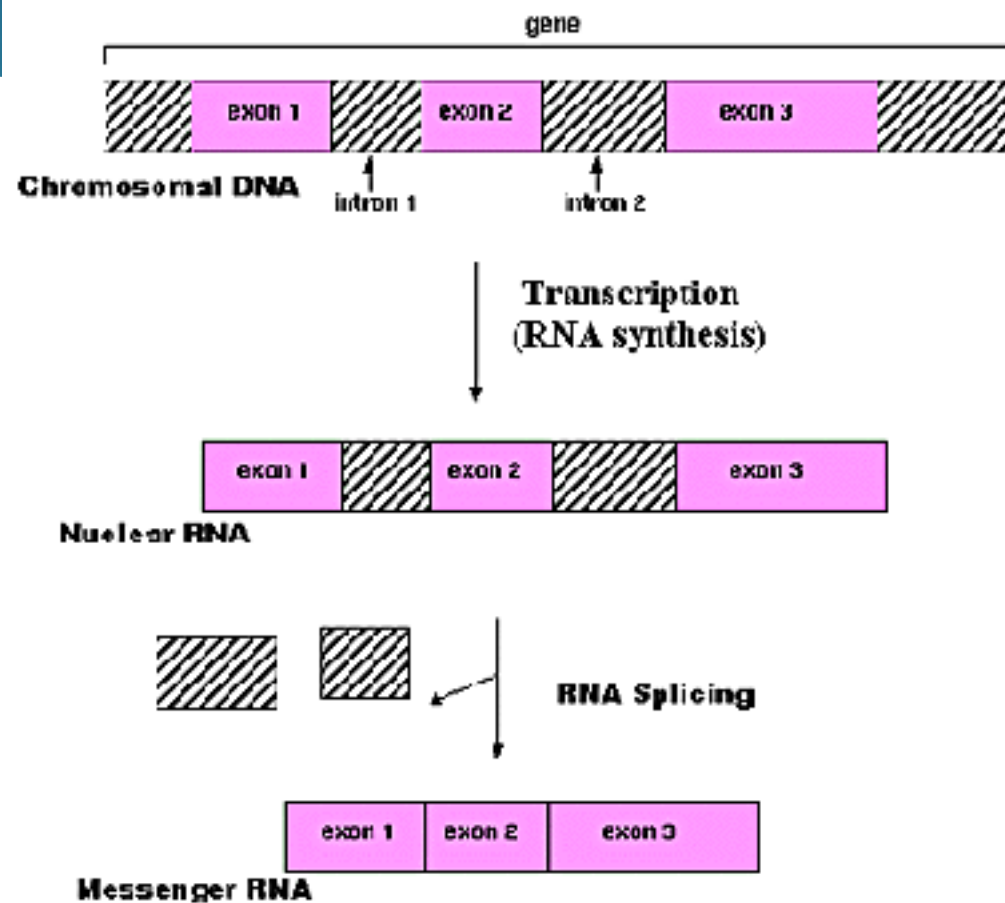
Central Dogma, for Eukaryotes ...



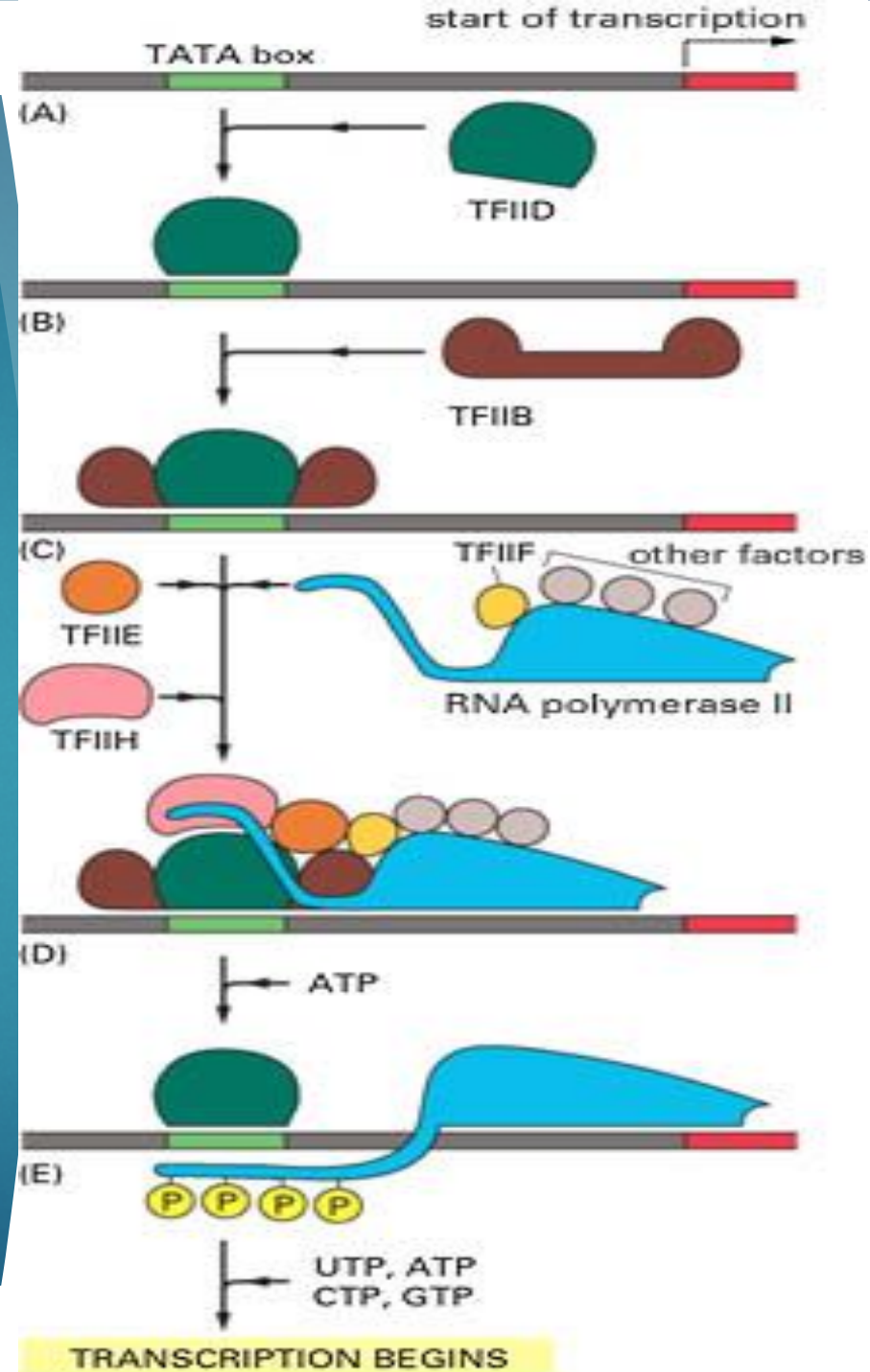
Transcription Regulation



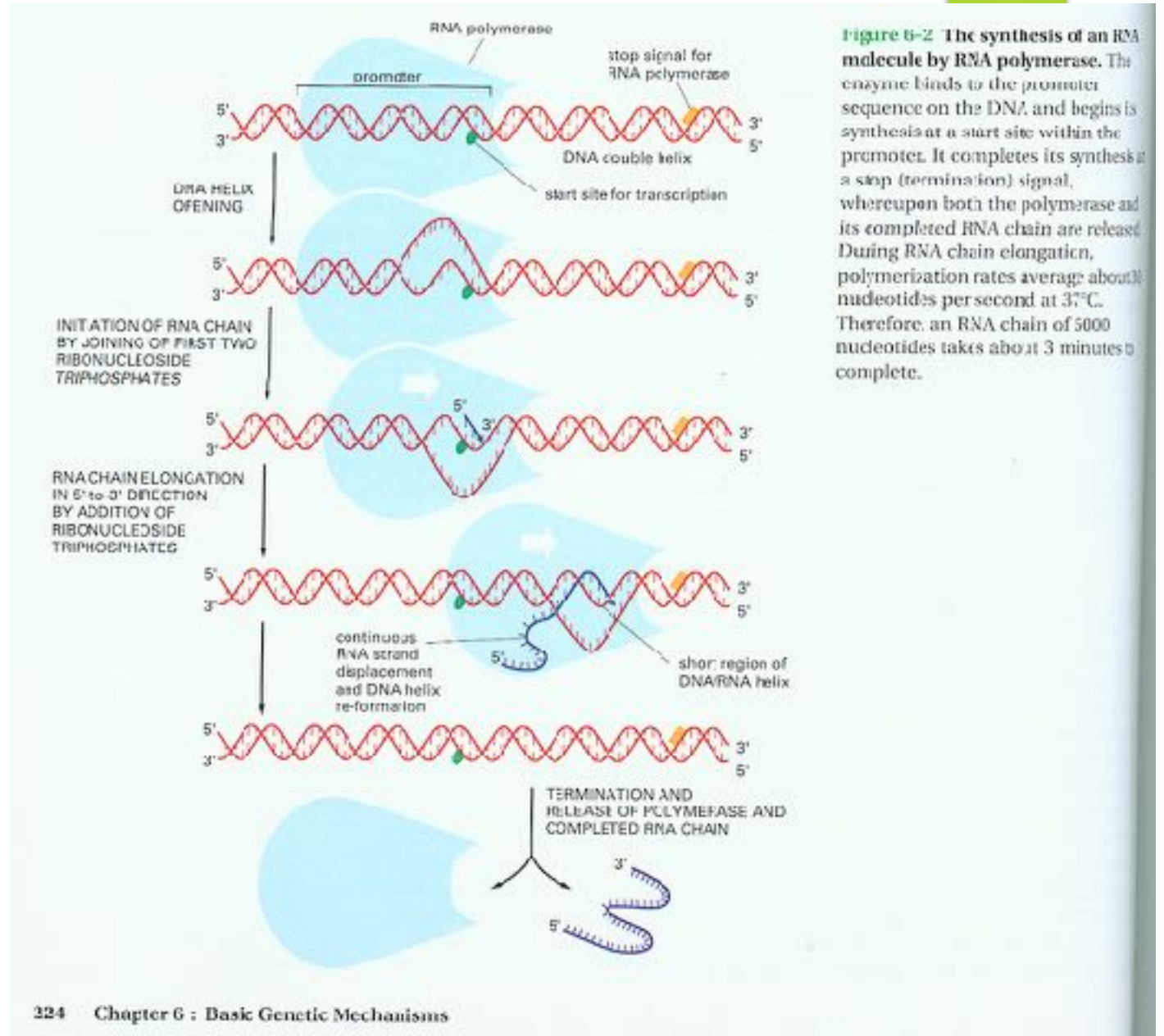
DNA Transcription



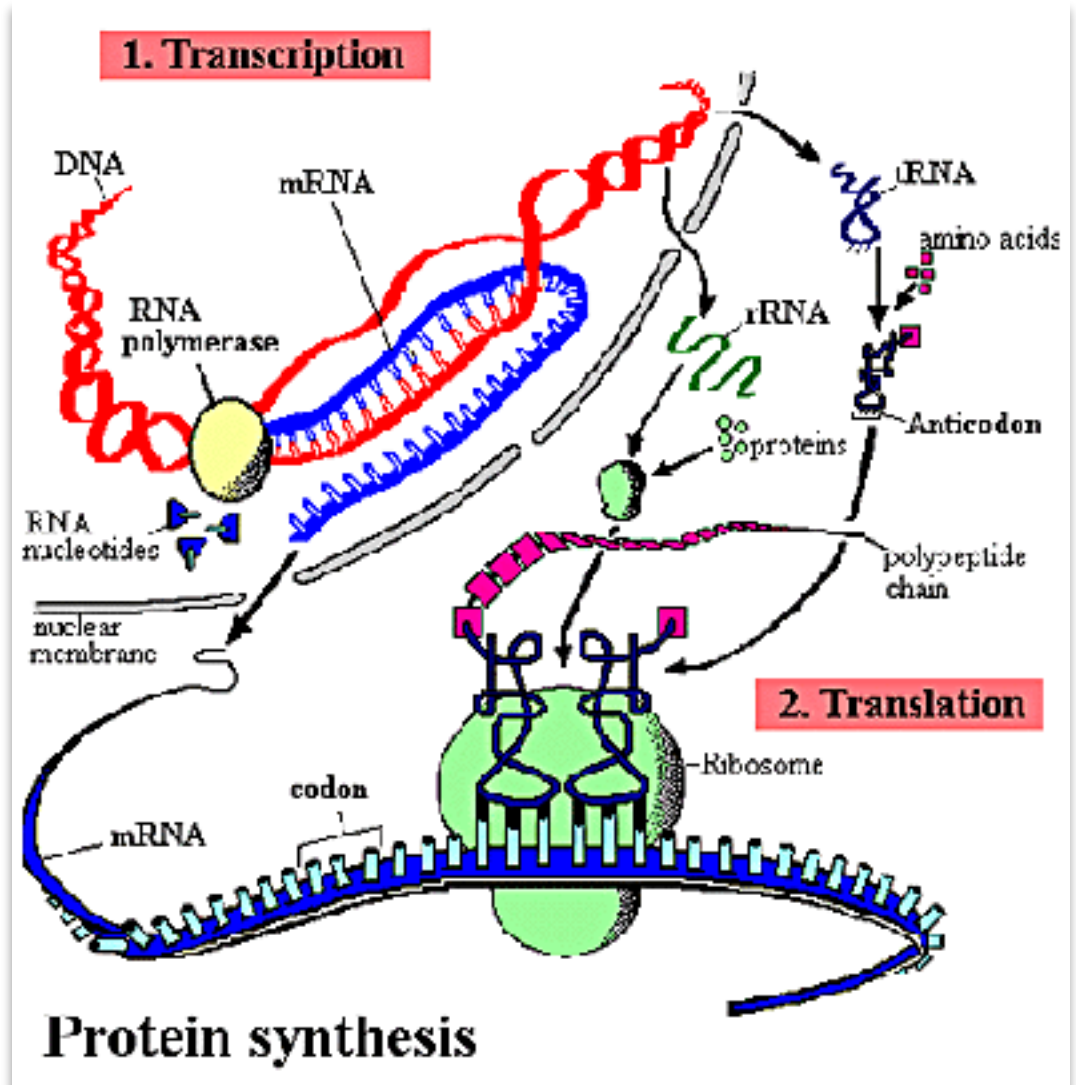
Transcription Initiation



Transcription, revisited ...



Protein Synthesis



Protein Synthesis: Incorporation of amino acid into protein

