Genetics and Jargon

March 29

Intro to summarization

At the end of each class, you, as a group, are going to summarize the important points for the day.

- 2 min to discuss with neighbors what the important points were.
- 2-3 minutes to have class list them, I will add to them as necessary

Summary of class 1

1. Human genome project began in 1990, completed 1st draft in 2000
2. Repeat regions were missing since they are hard to sequence
3. WAY fewer genes than proteins, lots of duplication and repeat regions, very little of the DNA is genes.
4. Moved on to trying to understand how to use this data
5. Linear chromosome needs centromere, origin of replication, and telomere
6. Central Dogma of Biology
7. Humans are diploid and gametes are haploid
8. Many pseudogenes in human genome. Mutate faster than functional genes

Questions from class 1?

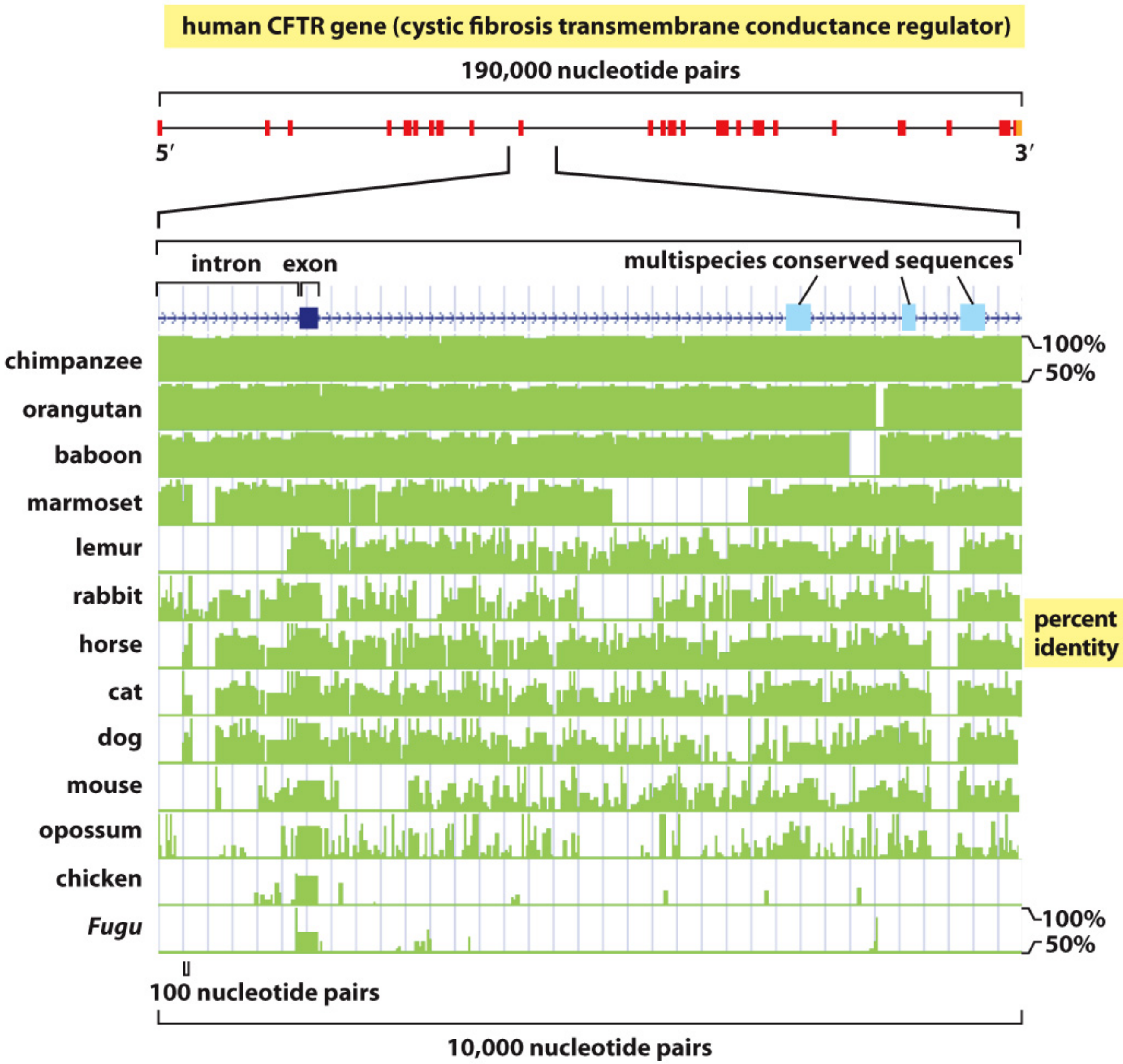
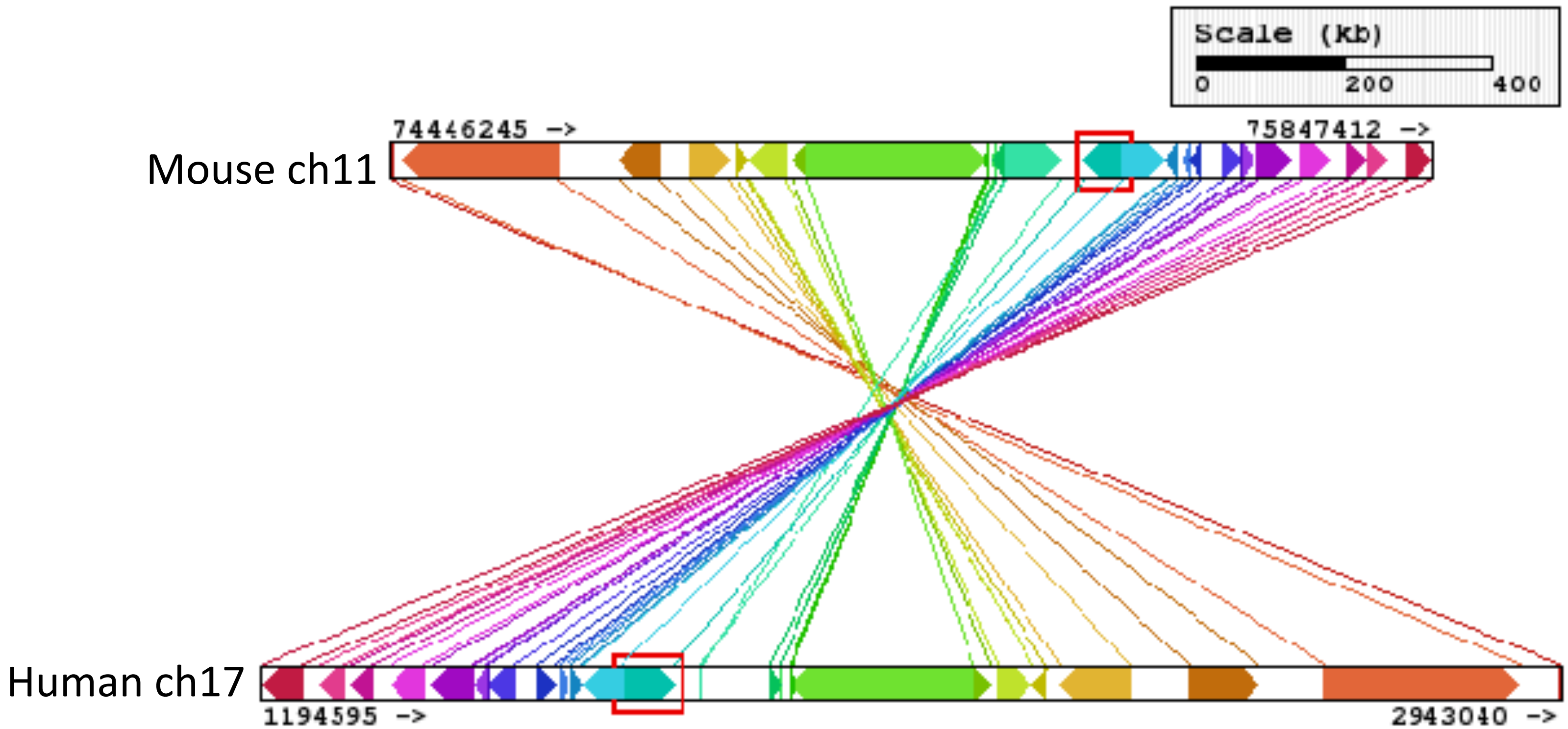


Figure 4-73 Molecular Biology of the Cell 6e (© Garland Science 2015)

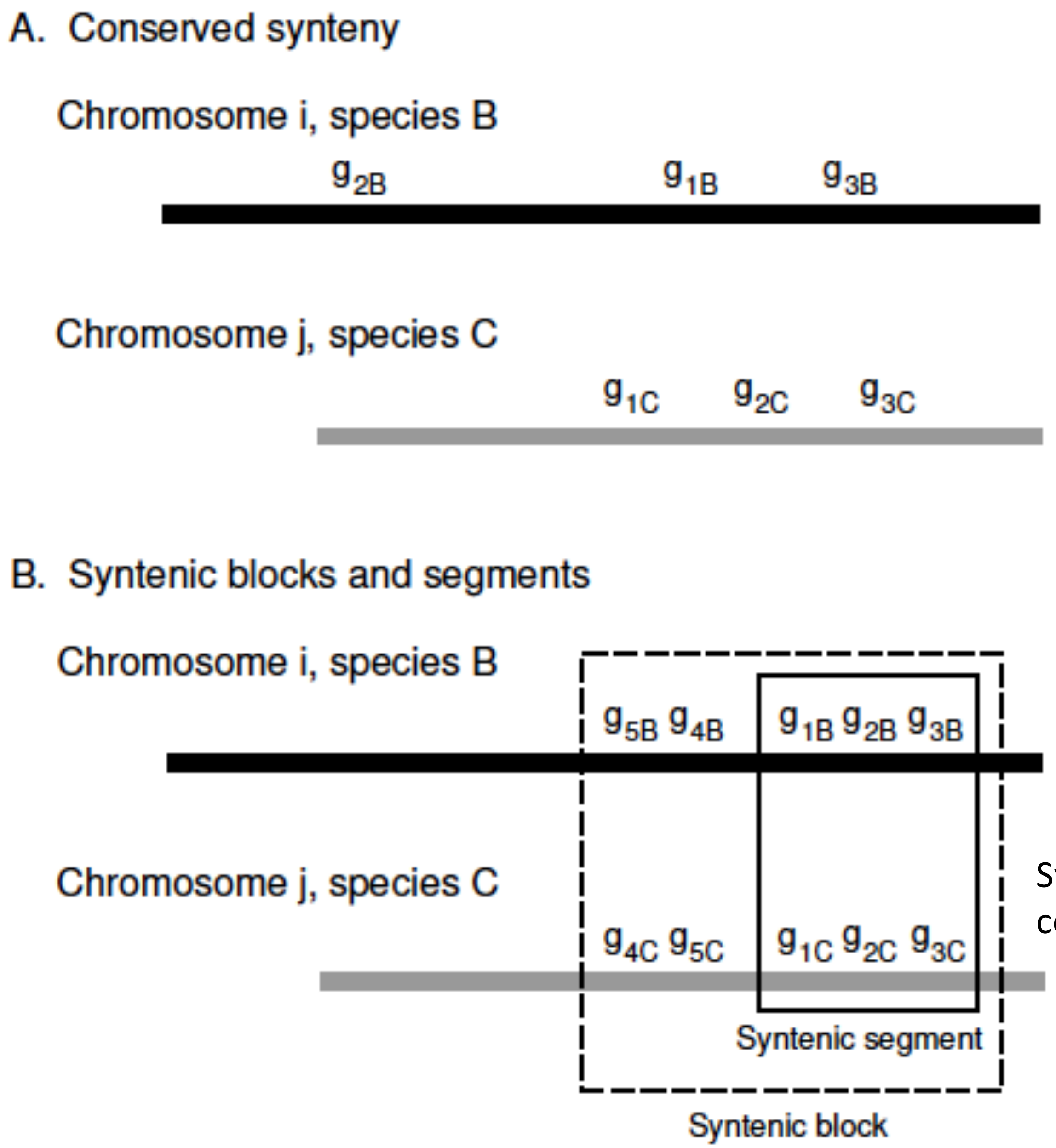
Multi-species conservation can point to functional importance

Conservation can exist beyond single genes. This can result in regions of **synteny.**



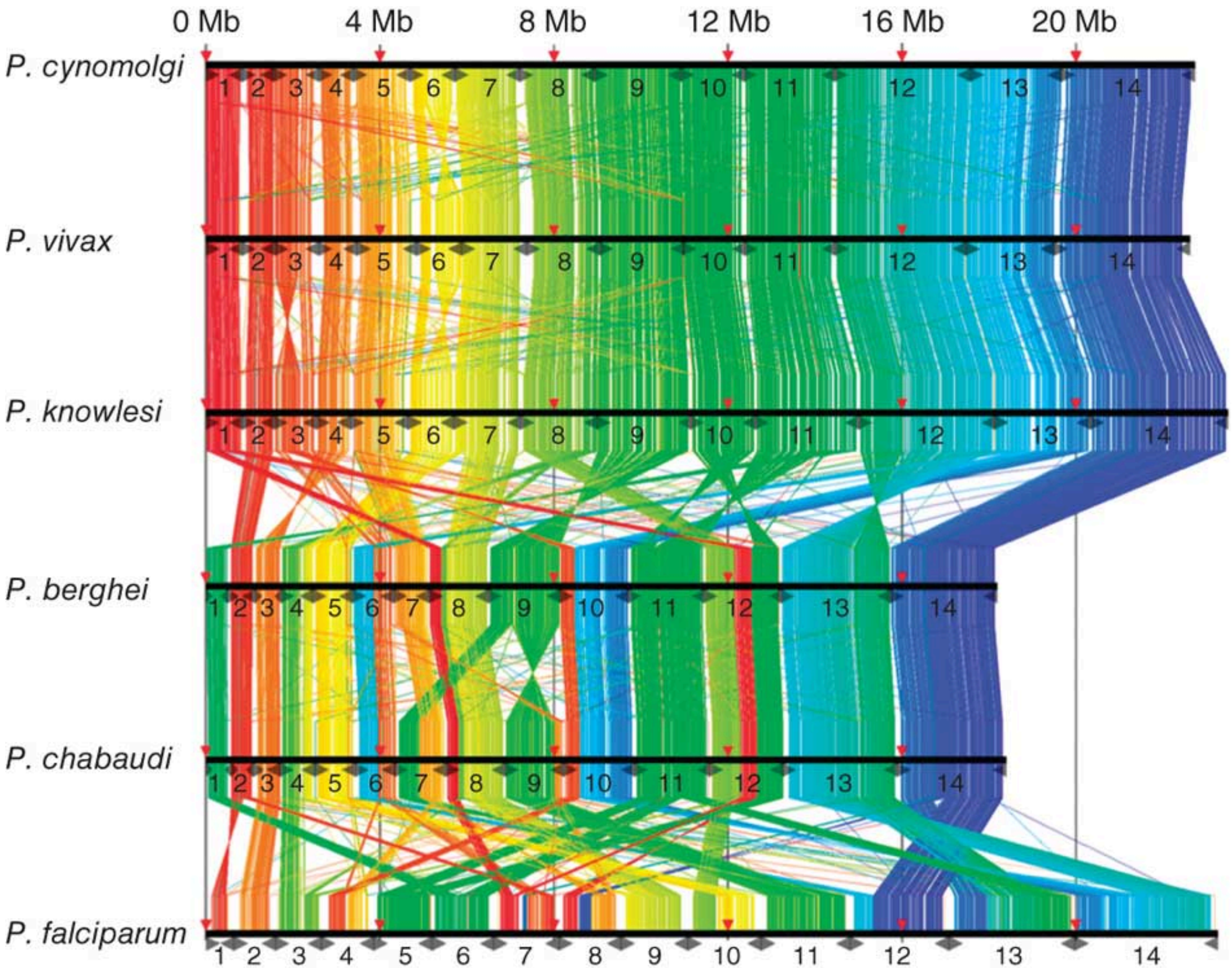
http://rosalind.info/glossary/synteny-block/

Many of the genes are orthologs, whole region is mirror image



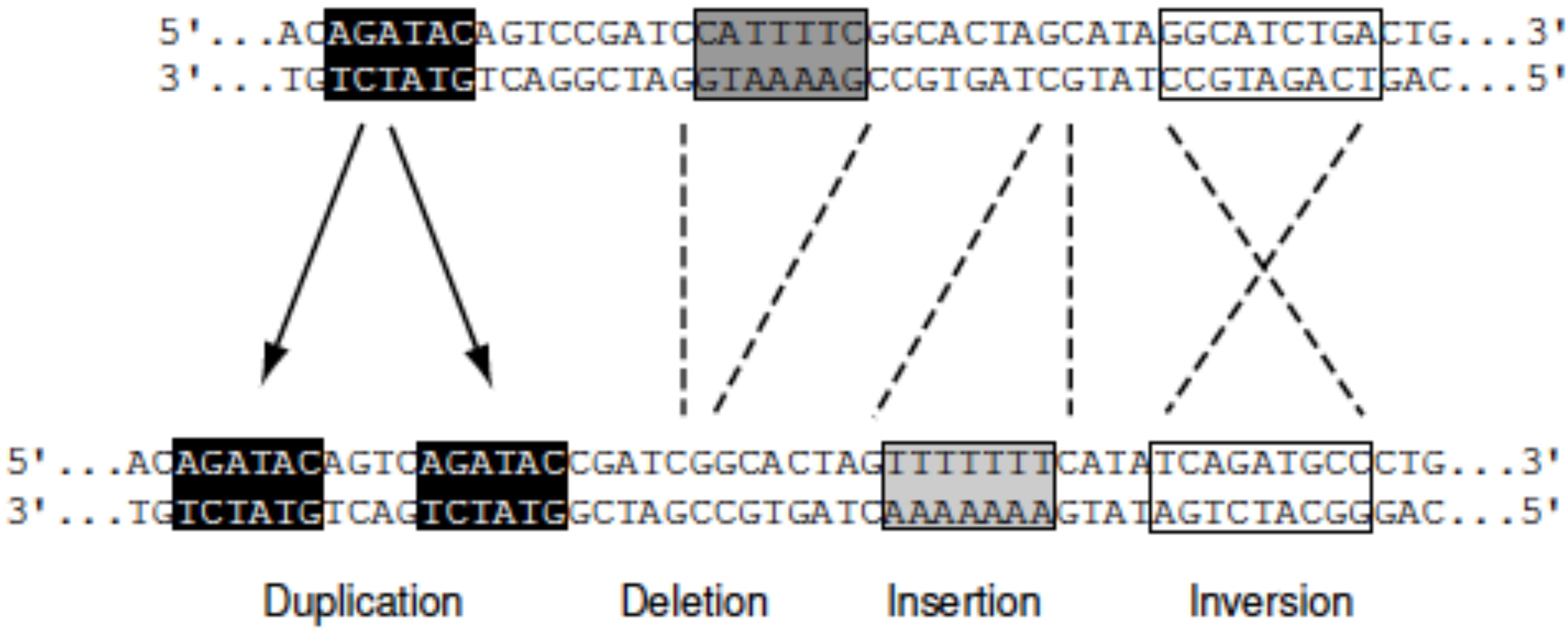
Syntenic segment = conserved segment

Deonier, R.C., et al. *Computational Genome Analysis: An Introduction*. New York, NY: Springer, 2007.



Tachibana, I-S. *Nature Genetics* **44**, 1051–1055 (2012) doi:10.1038/ng.2375

Four types of mutation can lead to these syntenic changes

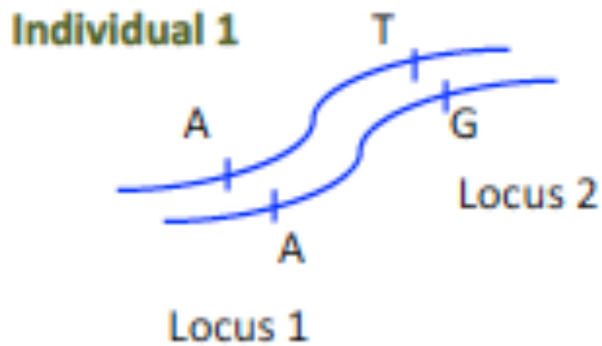


Deonier, R.C., et al. *Computational Genome Analysis: An Introduction*. New York, NY: Springer, 2007.

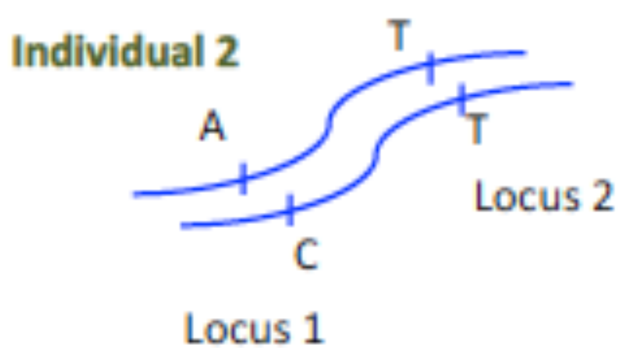
Syntenic Exercises

Humans are diploid organisms,

- Have 2 copies of each gene
- One on each copy of chromosome



Each copy of the gene, or other genetic locus that causes a phenotype is knows as an **allele**



Two options

- Homozygous – have the same allele in each chromosome
- Heterozygous – have different alleles in each chromosome

Either allele may be dominant

For the individuals above, are they homozygous or heterozygous at each locus?

GATCTTCGTACTGAGT
GATCTTCGTACTGAGT

GATTTTCGTACGGAAT
GATTTTCGTACTGAGT

GATCTTCGTACTGAAT
GATTTTCGTACGGAAT

GATTTTCGTACTGAAT
GATCTTCGTACGGAAT

Binary (biallelic) single nucleotide polymorphisms

Each person has a **genotype** that describes the content of their genome at specific alleles

GATCTTCGTACTGAGT
GATCTTCGTACTGAGT
CC/TT/GG

GATTTTCGTACGGAAT
GATTTTCGTACTGAGT
TT/TG/GA

GATCTTCGTACTGAAT
GATTTTCGTACGGAAT
CT/GT/AA

GATTTTCGTACTGAAT
GATCTTCGTACGGAAT
CT/TG/AA

This does not tell us anything about which chromosome the alelle is on

GATCTTCGTACTGAGT CTG
GATCTTCGTACTGAGT CTG

GATTTTCGTACGGAAT TGA
GATTTTCGTACTGAGT TTG

GATCTTCGTACTGAAT CTA
GATTTTCGTACGGAAT TGA

GATTTTCGTACTGAAT TTA
GATCTTCGTACGGAAT CGA

Haplotype is the string of alleles that are on the same chromosome

GATCTTCGTACTGAGT	CTG	CTG/CTG
GATCTTCGTACTGAGT	CTG	
GATTTTCGTACGGAAT	TGA	TGA/TTG
GATTTTCGTACTGAGT	TTG	
GATCTTCGTACTGAAT	CTA	CTA/TGA
GATTTTCGTACGGAAT	TGA	
GATTTTCGTACTGAAT	TTA	TTA/CGA
GATCTTCGTACGGAAT	CGA	

An individual has 2 haplotypes per set of loci

What are potential haplotypes for a mating with orange and yellow?

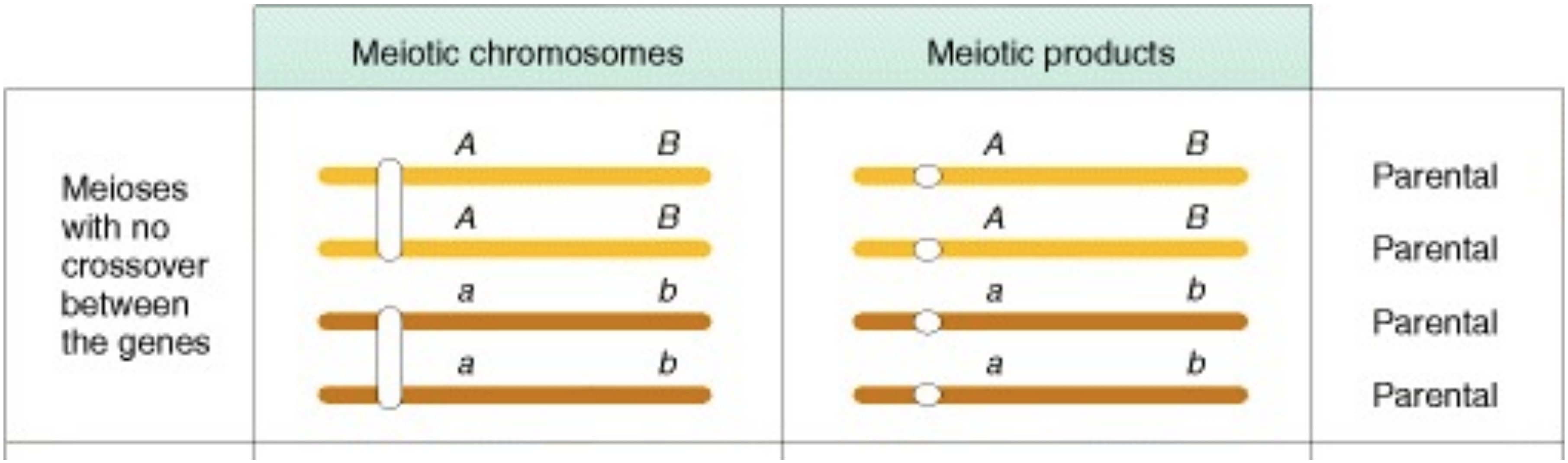
GATCTTCGTACTGAGT	CTG	CTG/CTG
GATCTTCGTACTGAGT	CTG	
GATTTTCGTACGGAAT	TGA	TGA/TTG
GATTTTCGTACTGAGT	TTG	

GATCTTCGTACTGAGT	CTG	CTG/TGA
GATTTTCGTACGGAAT	TGA	

GATCTTCGTACTGAGT	CTG	CTG/TTG
GATTTTCGTACTGAGT	TTG	

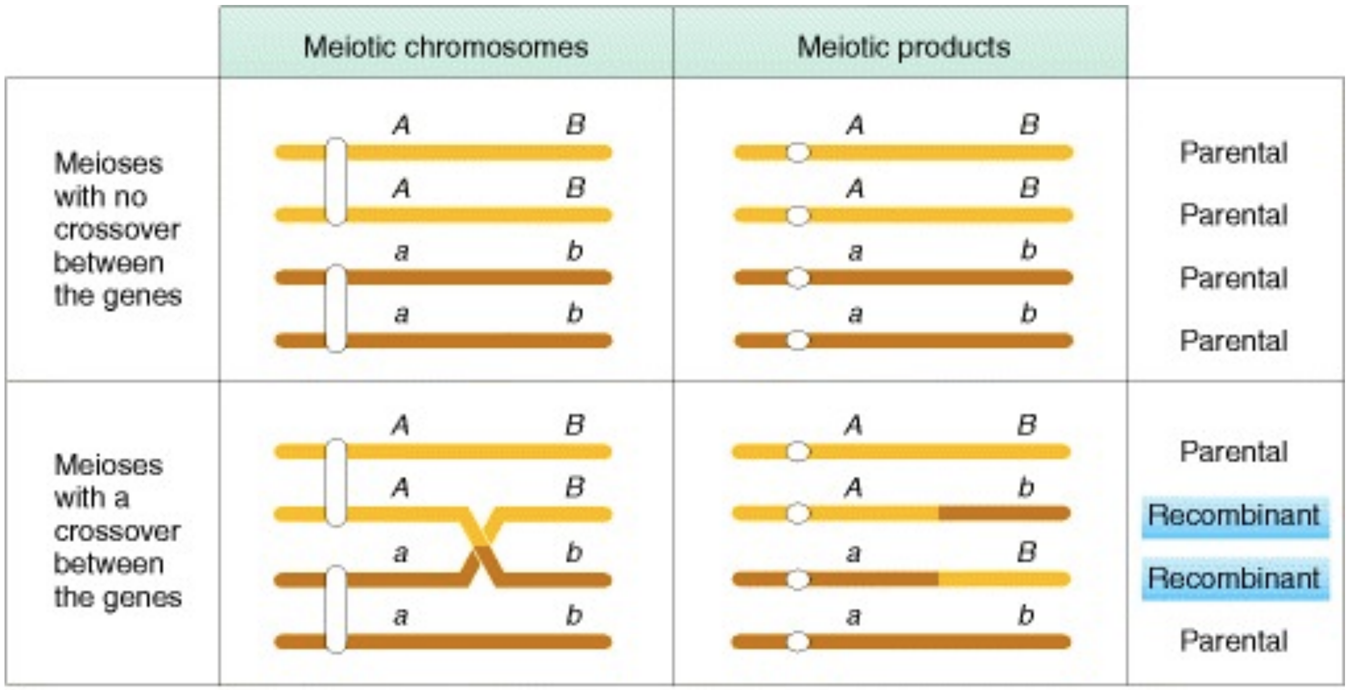
Does this always hold?

On the same Chromosome



Griffiths AJF, Miller JH, Suzuki DT, et al. An Introduction to Genetic Analysis. 7th edition. New York: W. H. Freeman; 2000. Recombination. Figure 5-7.

Recombination via Crossovers



Griffiths AJF, Miller JH, Suzuki DT, et al. An Introduction to Genetic Analysis. 7th edition. New York: W. H. Freeman; 2000. Recombination. Figure 5-7.

We'll come back to this in the population genetics section!

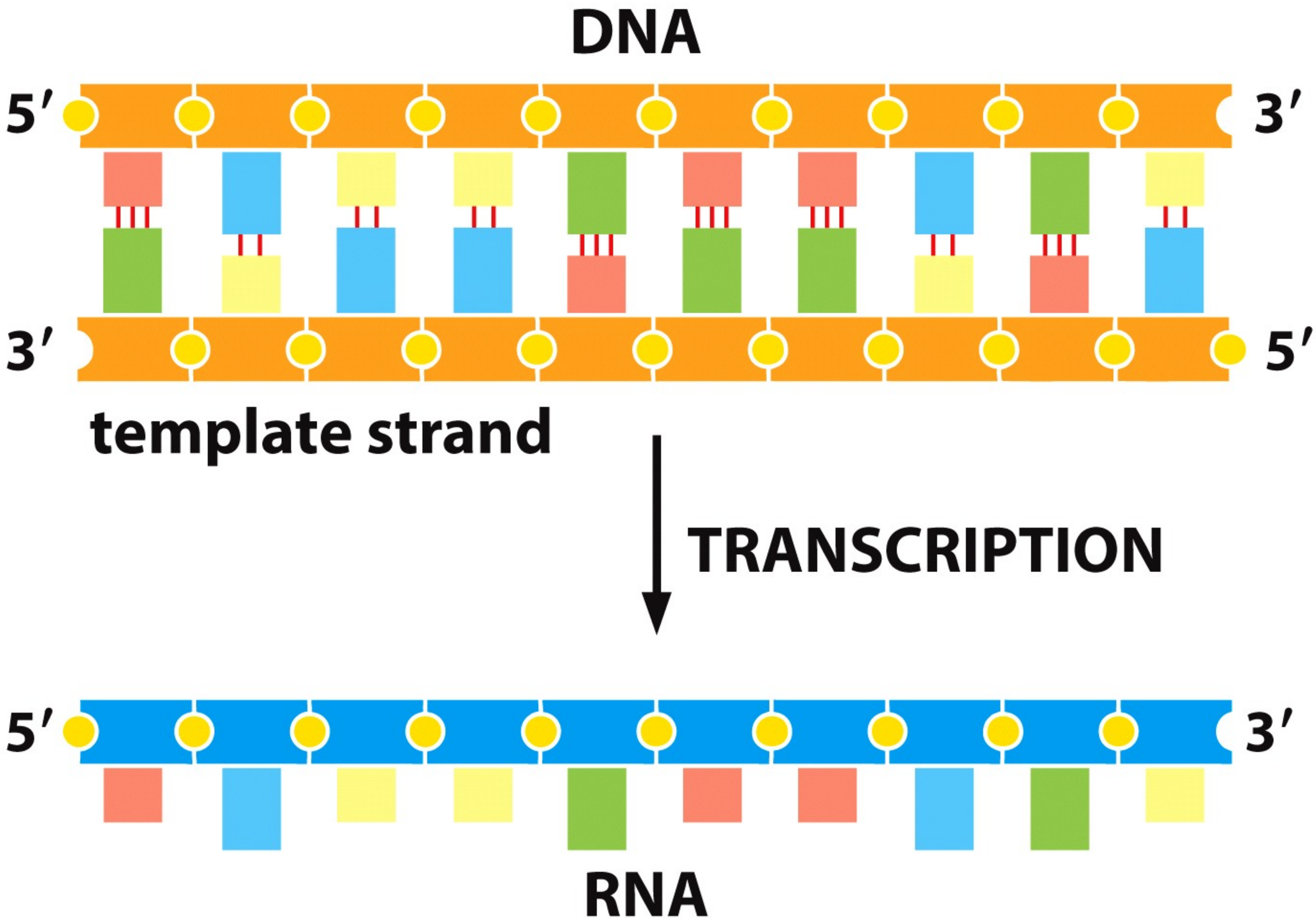


Figure 6-8 Molecular Biology of the Cell 6e (© Garland Science 2015)

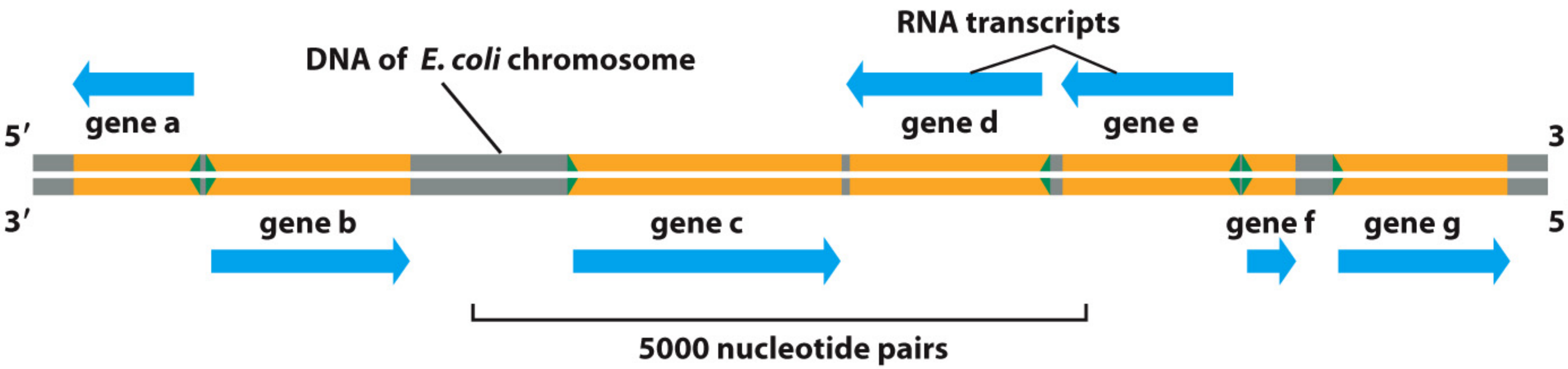
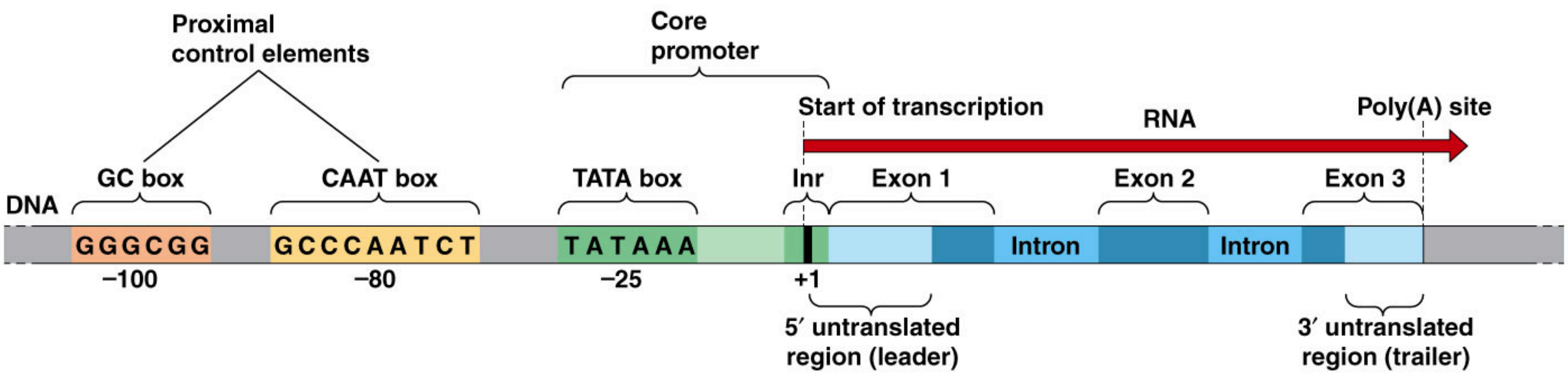


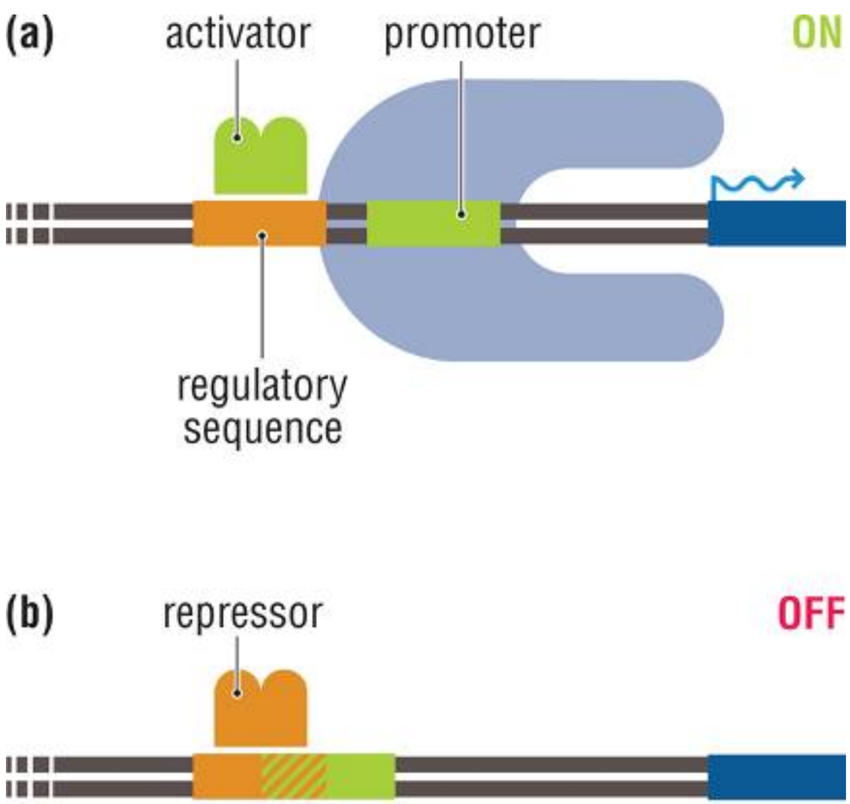
Figure 6-13 Molecular Biology of the Cell 6e (© Garland Science 2015)

Gene has regulatory regions (mainly) upstream of start site



Why are haplotypes relevant to understanding gene expression?

Transcriptional Elements

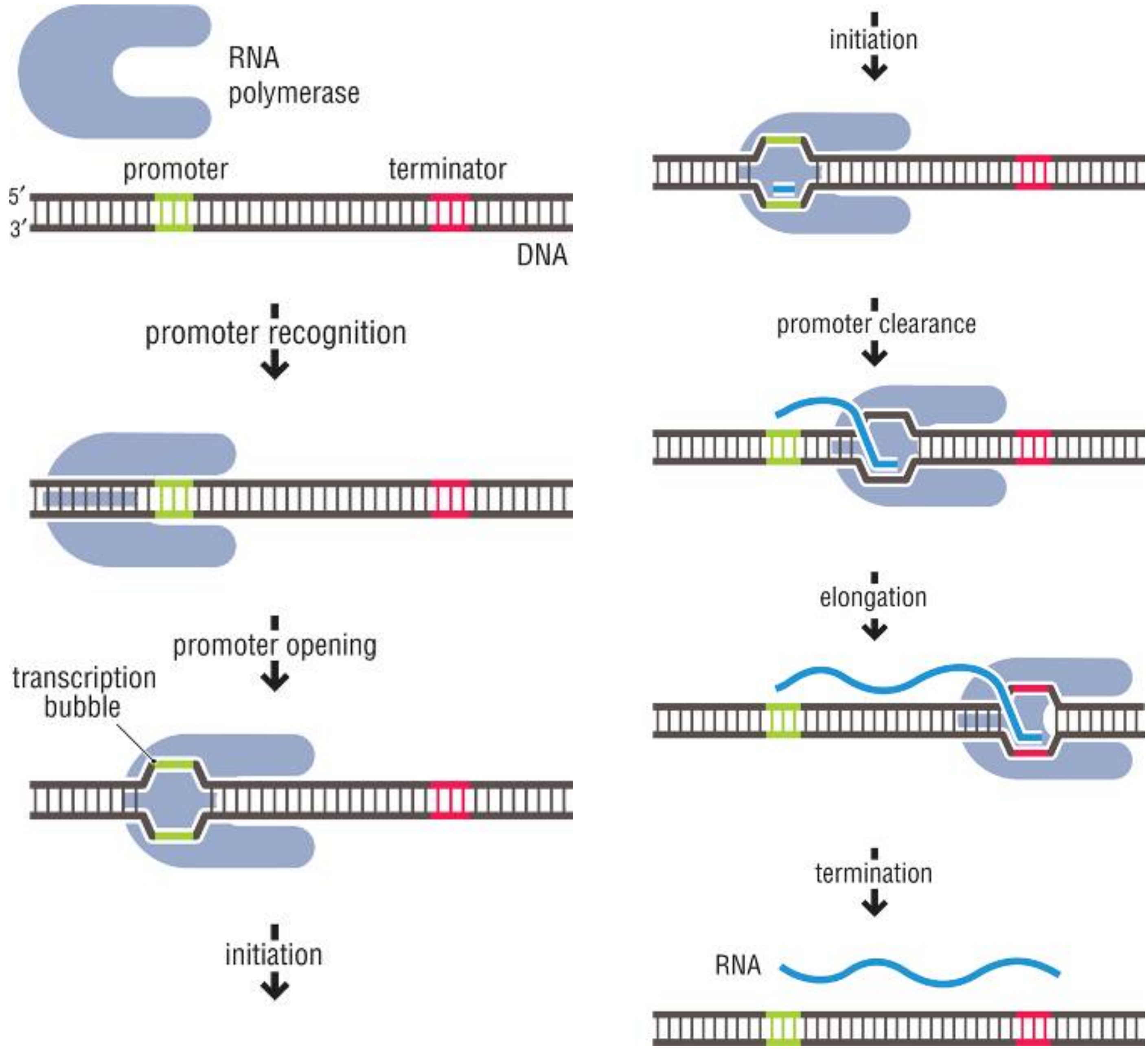


Cis-acting elements

- Nearby
- DNA sequences
- TATA box, UAS, etc.

Trans-acting elements

- From afar
- Usually involve proteins
- General or specific
- GAL4



Craig et. al. *Molecular Biology*, Oxford University Press, 2010

Figure 08-02

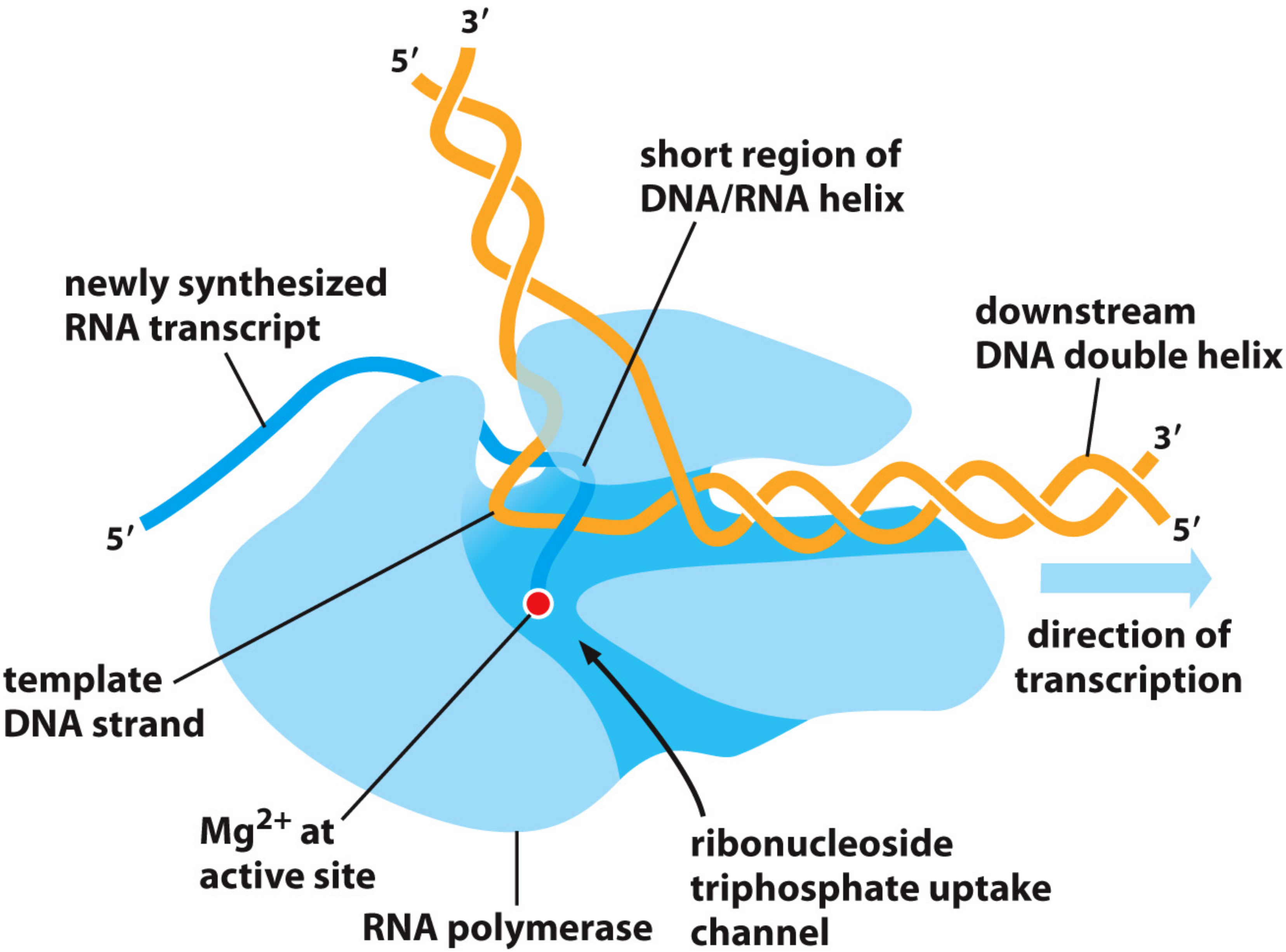


Figure 6-9 Molecular Biology of the Cell 6e (© Garland Science 2015)

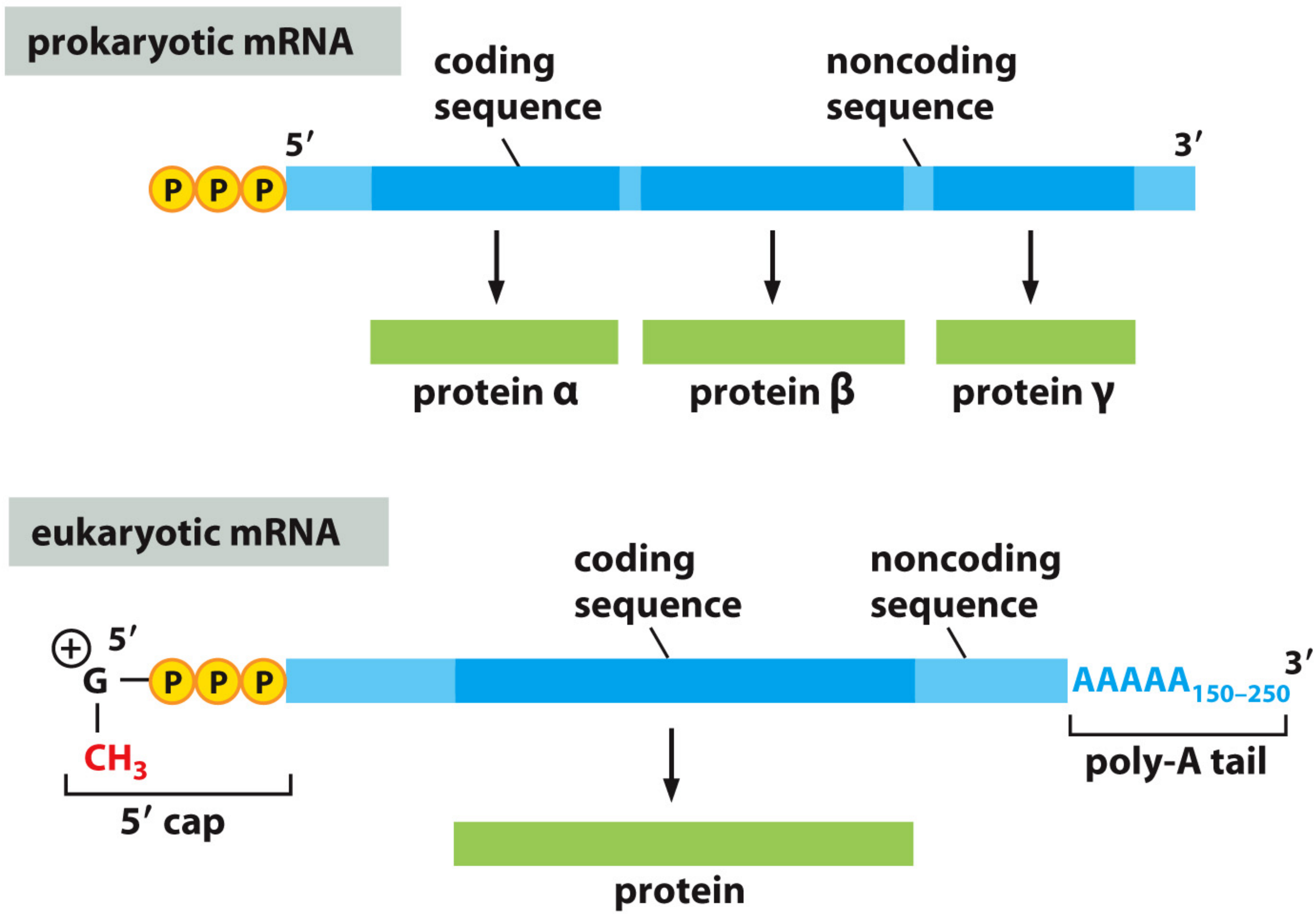


Figure 6-21a Molecular Biology of the Cell 6e (© Garland Science 2015)

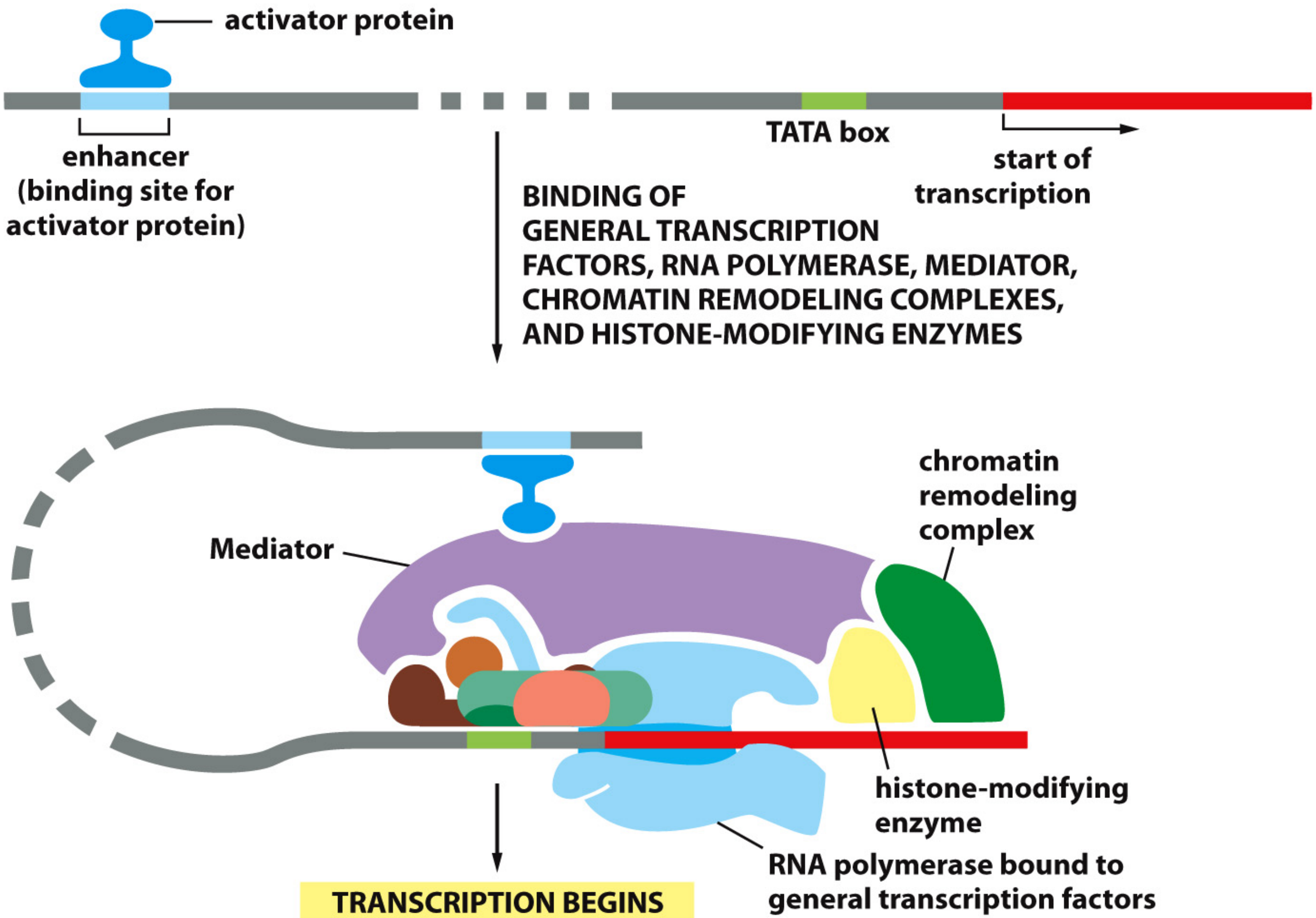


Figure 6-18 Molecular Biology of the Cell 6e (© Garland Science 2015)

Operons in bacteria allow the upregulation of many genes at once

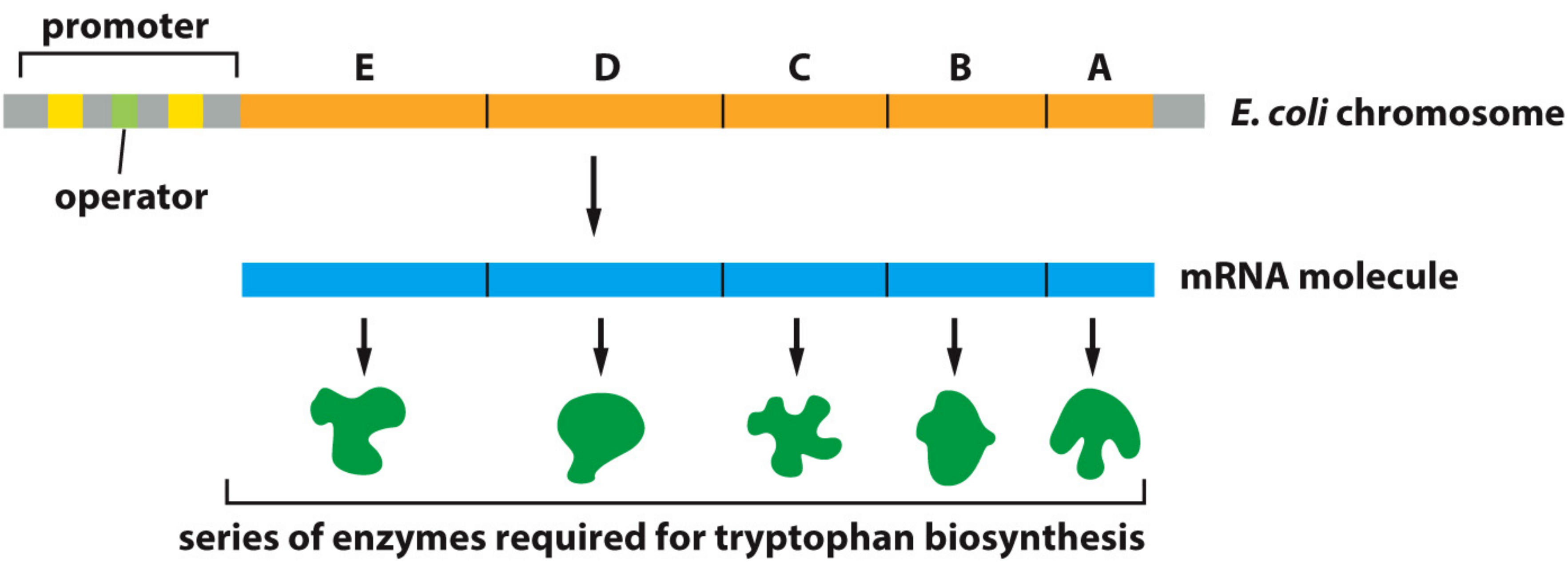


Figure 7-12 Molecular Biology of the Cell 6e (© Garland Science 2015)

Eukaryotes are more complex:

1. No operons
2. Transcription factors will bind multiple sites in genome to turn on/off multiple genes
3. Genes have introns and exons

Multiple type of activators and repressors act in concert based on different cellular signals

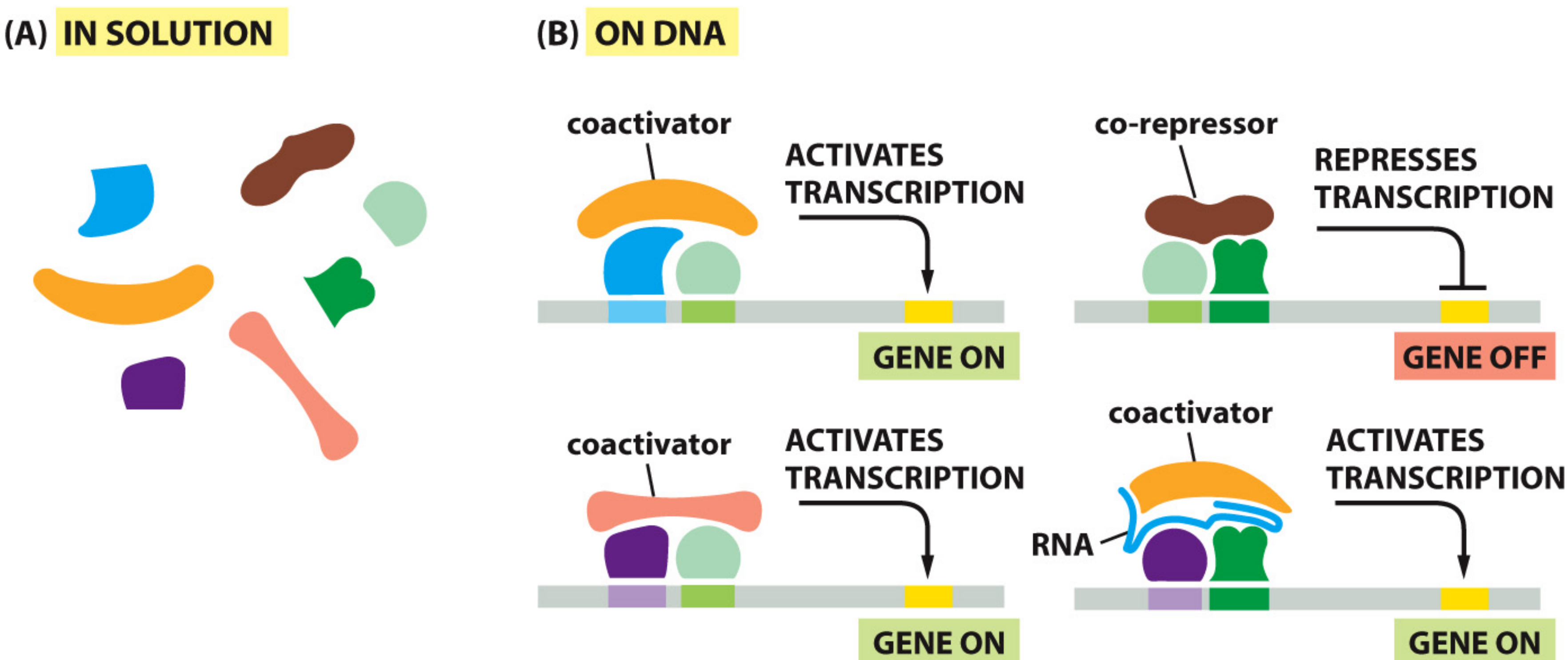


Figure 7-18 Molecular Biology of the Cell 6e (© Garland Science 2015)

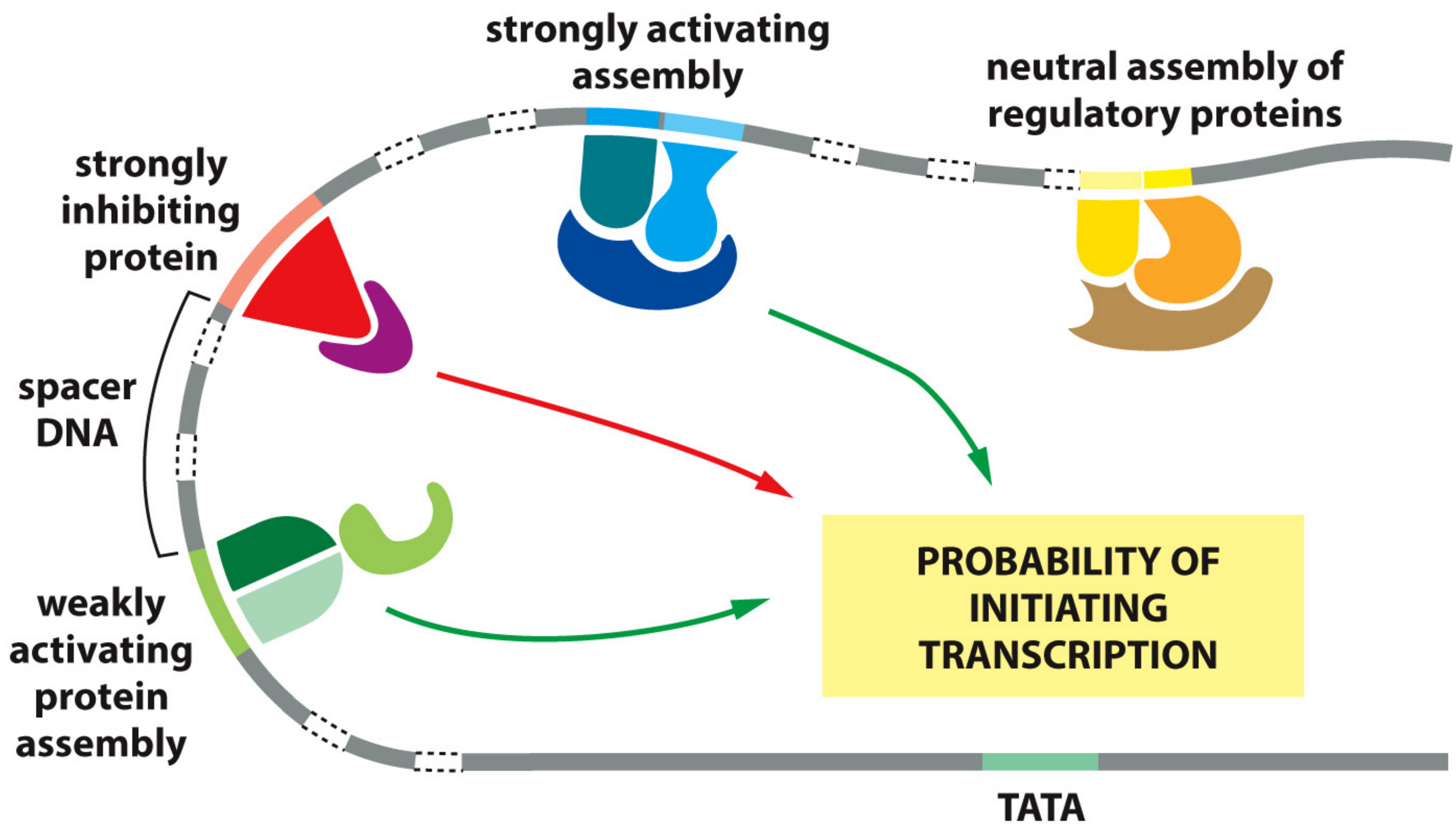


Figure 7-31 Molecular Biology of the Cell 6e (© Garland Science 2015)

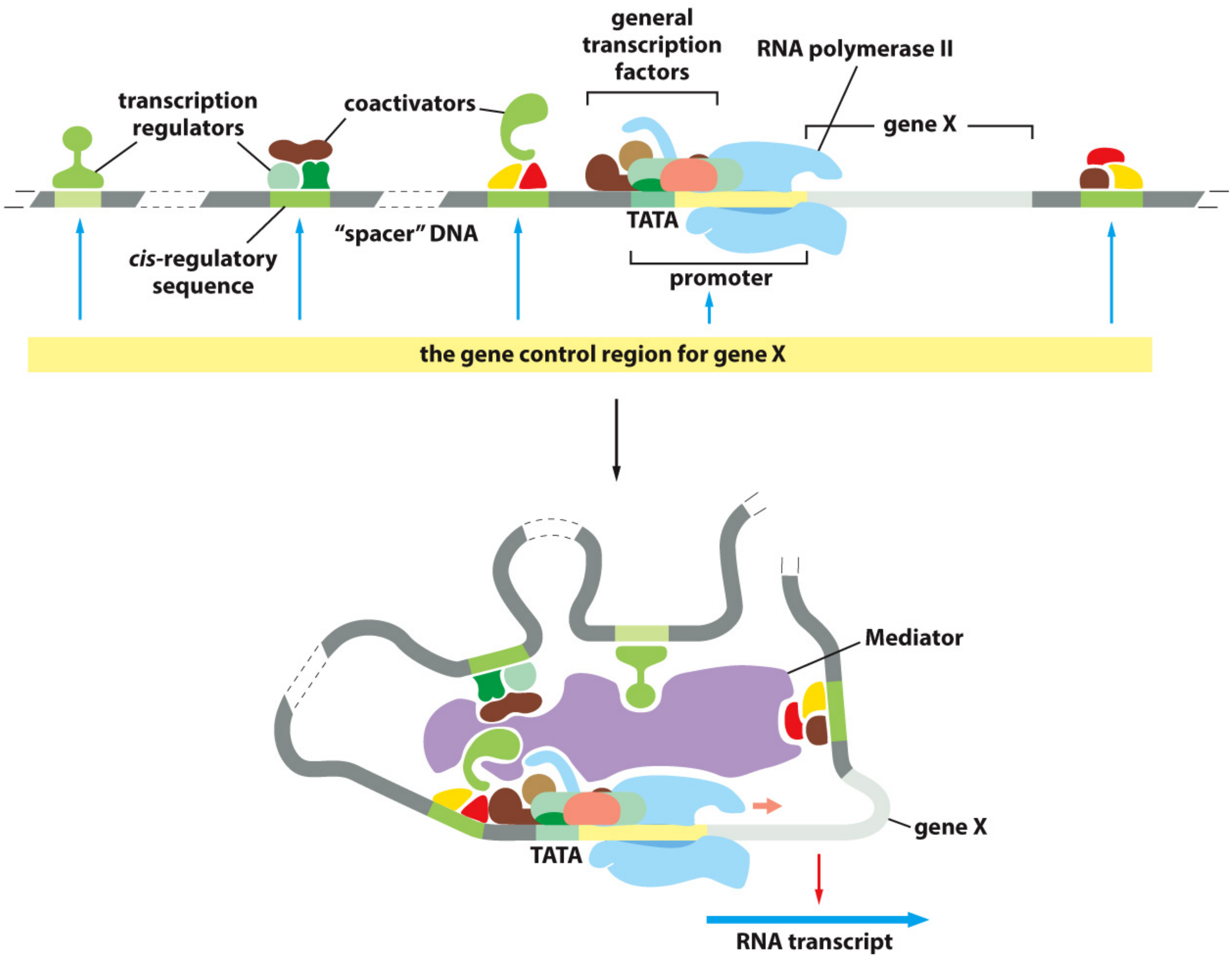


Figure 7-17 Molecular Biology of the Cell 6e (© Garland Science 2015)

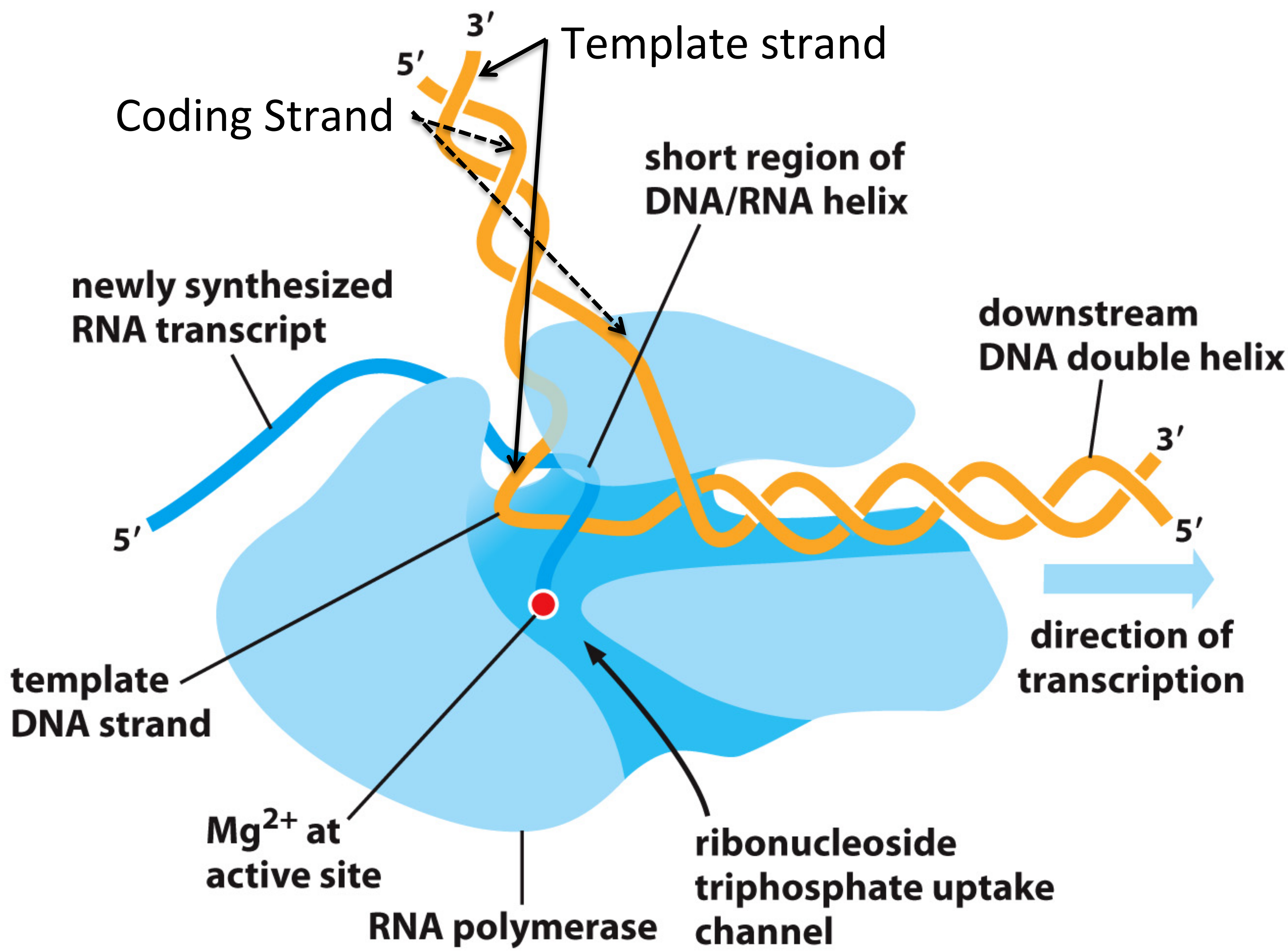


Figure 6-9 Molecular Biology of the Cell 6e (© Garland Science 2015)

DNA is still in nucleosomes and transcription has to happen around and within them

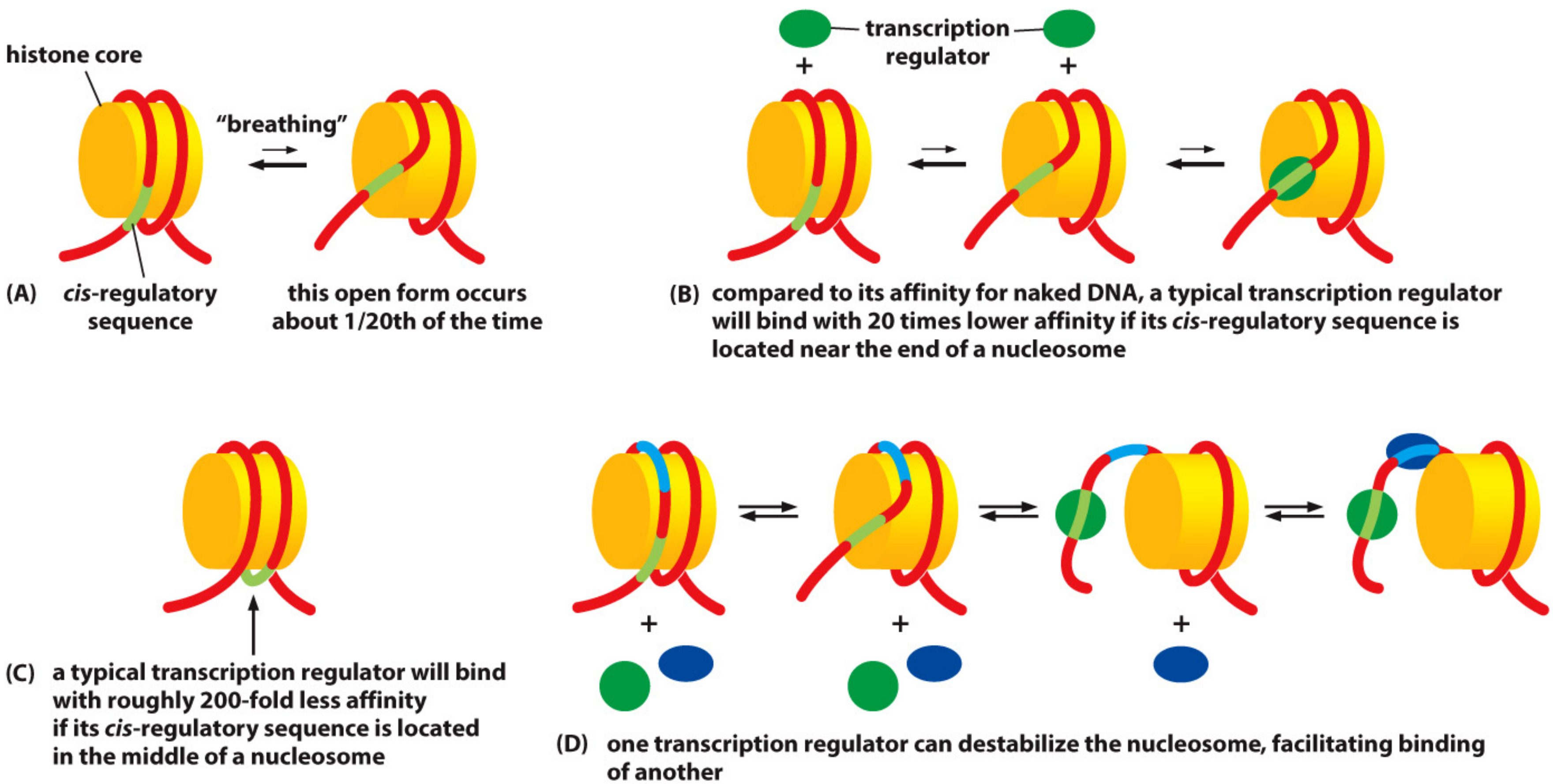


Figure 7-11 Molecular Biology of the Cell 6e (© Garland Science 2015)

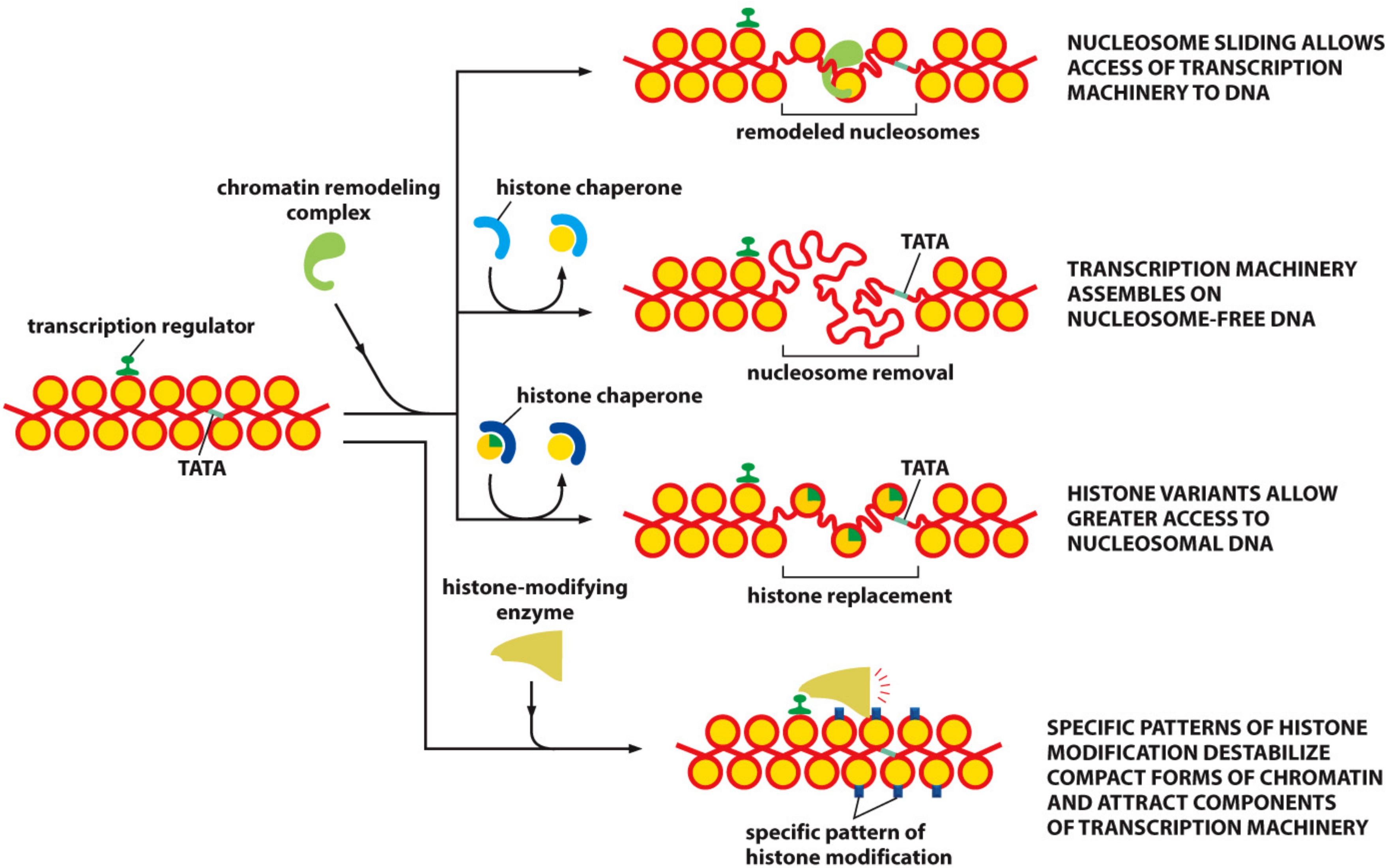
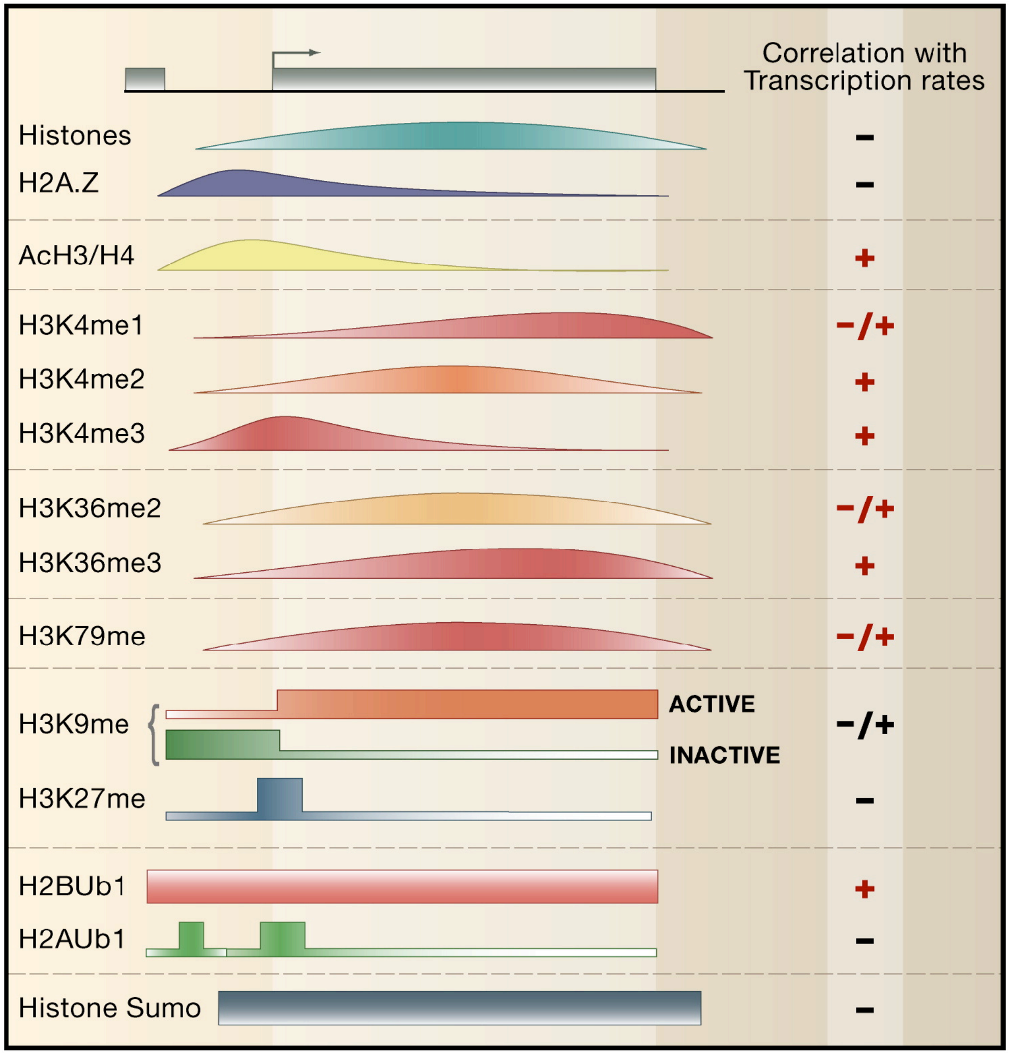
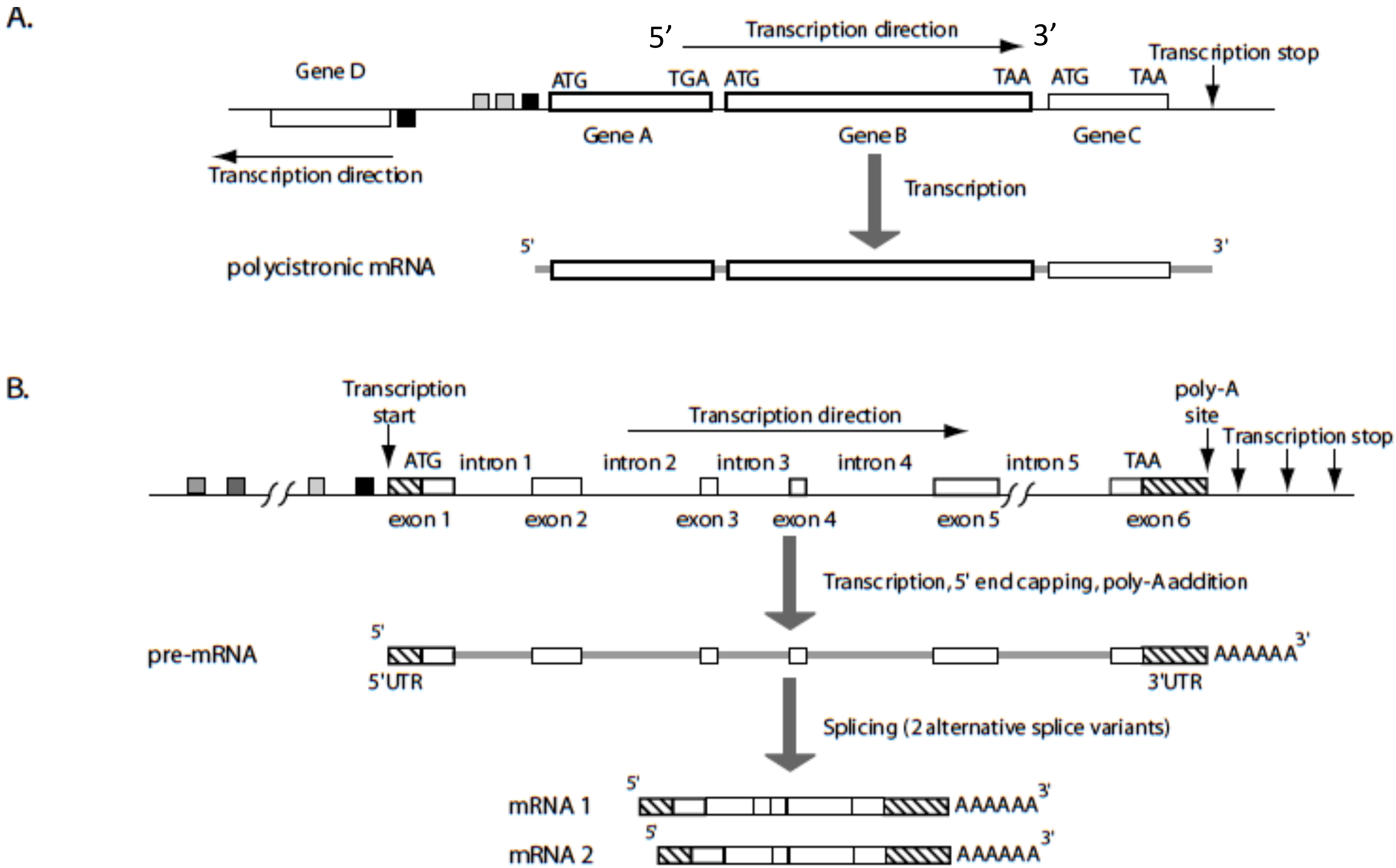


Figure 7-19 Molecular Biology of the Cell 6e (© Garland Science 2015)

Genomic techniques enable us to understand histone modifications



DOI 10.1016/j.cell.2007.01.015

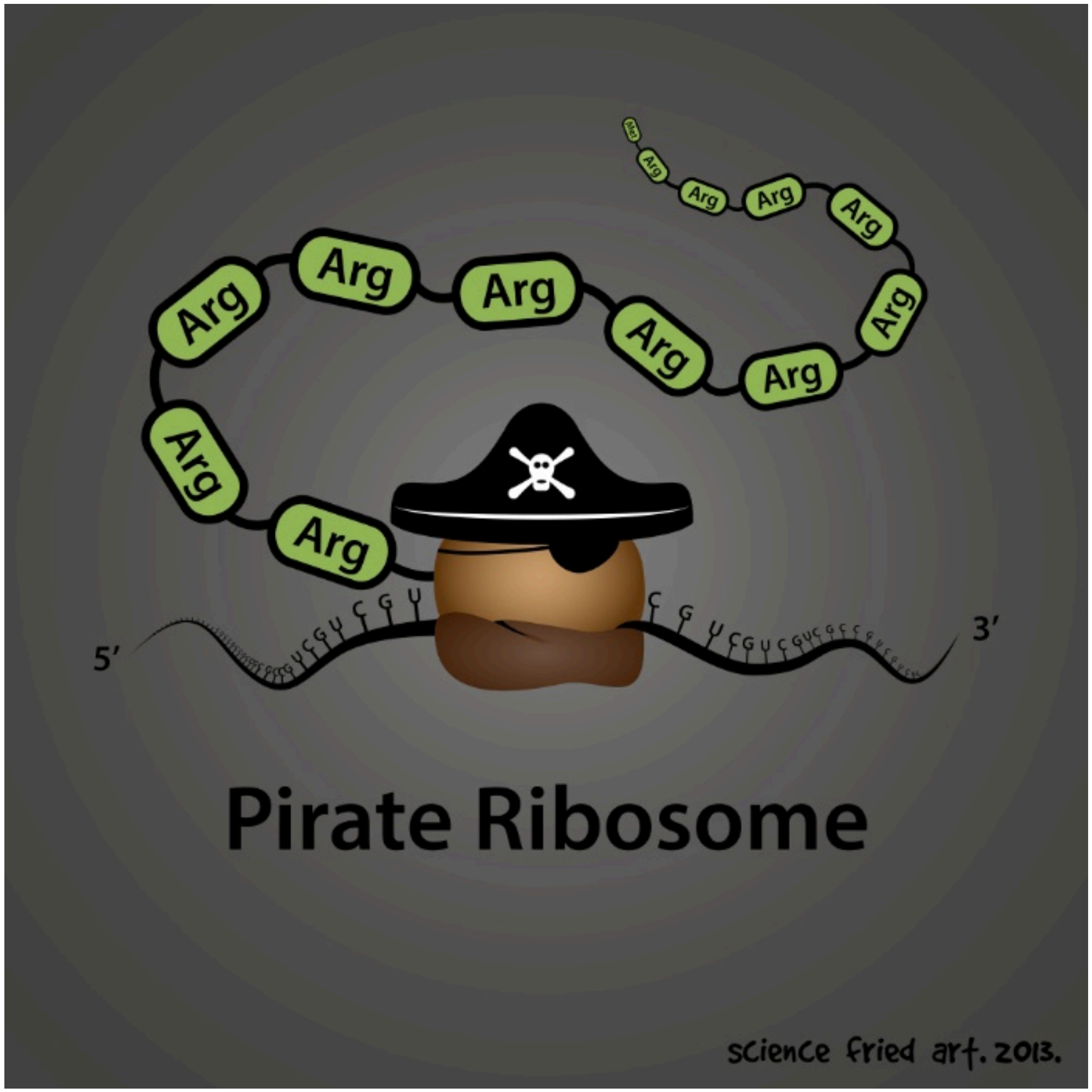


A.
5′
Transcription direction
3′
Transcription stop
Gene D
ATG
TGA
ATG
TAA
ATG
TAA
Gene A
Gene B
Gene C
Transcription direction
Transcription
5'
3'
polycistronic mRNA
B.
Transcription start
ATG
intron 1
intron 2
intron 3
intron 4
intron 5
TAA
Transcription direction
poly-A site
Transcription stop
exon 1
exon 2
exon 3
exon 4
exon 5
exon 6
Transcription, 5' end capping, poly-A addition
5'
pre-mRNA
AAAAAA
3'
5'UTR
3'UTR
Splicing (2 alternative splice variants)
5'
mRNA 1
AAAAAA
3'
5'
mRNA 2
AAAAAA
3'

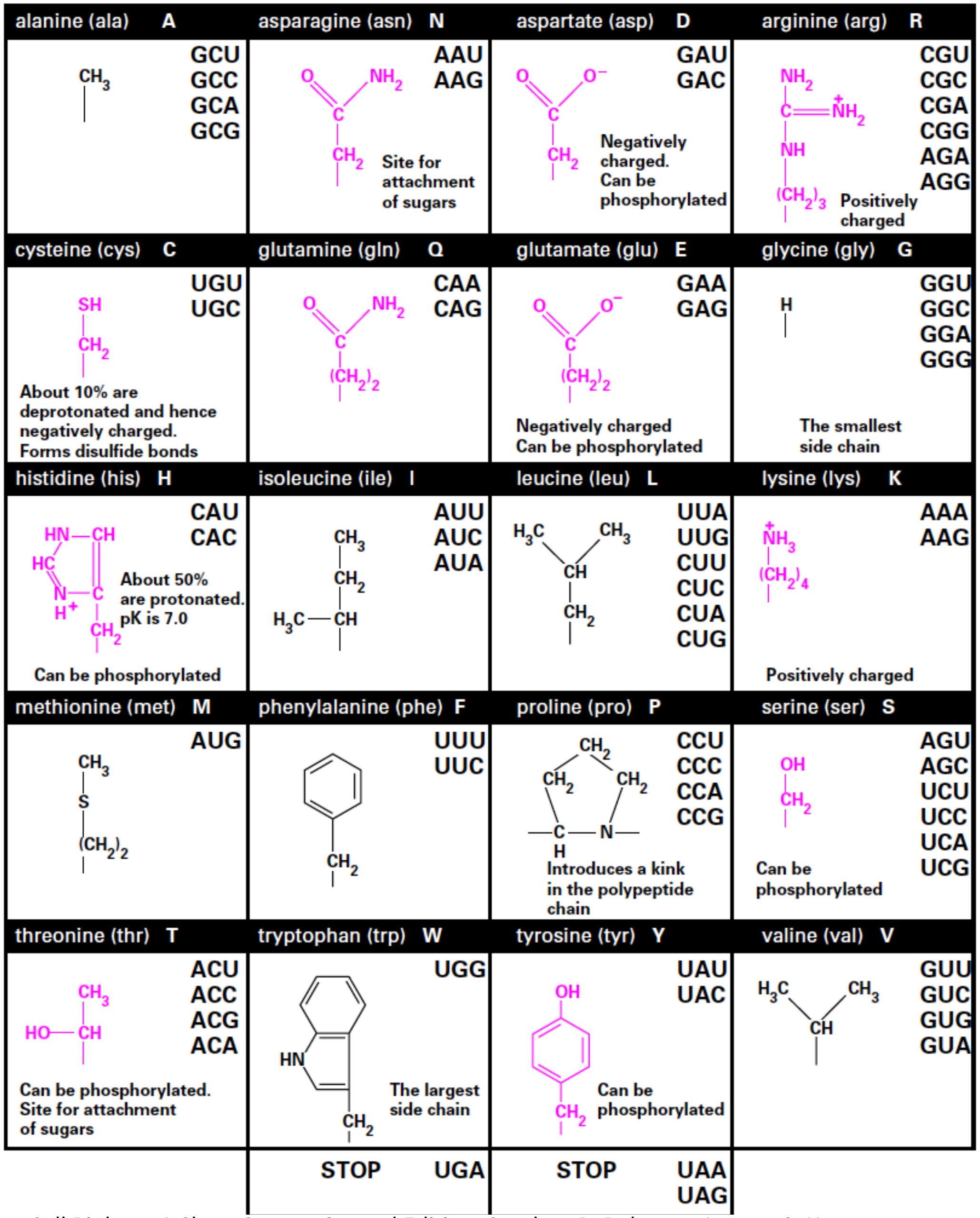
Transcription Exercises

Questions?

Translation!

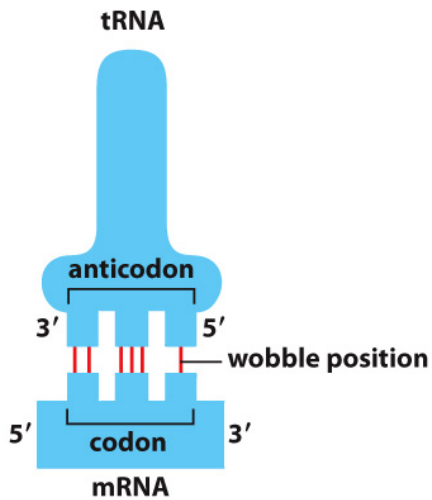


The genetic code and the corresponding amino acid side chains.



Cell Biology: A Short Course, Second Edition; Stephen R. Bolsover, Jeremy S. Hyams, Elizabeth A. Shephard, Hugh A. White, Claudia G. Wiedemann

- Use only 20 amino acids
- Code is Degenerate
  - Multiple codons for many of the amino acids
- Why is this good evolutionarily?
  - Not all changes in genetic code will result in change in protein
  - Organism is more stable to mutation



bacteria

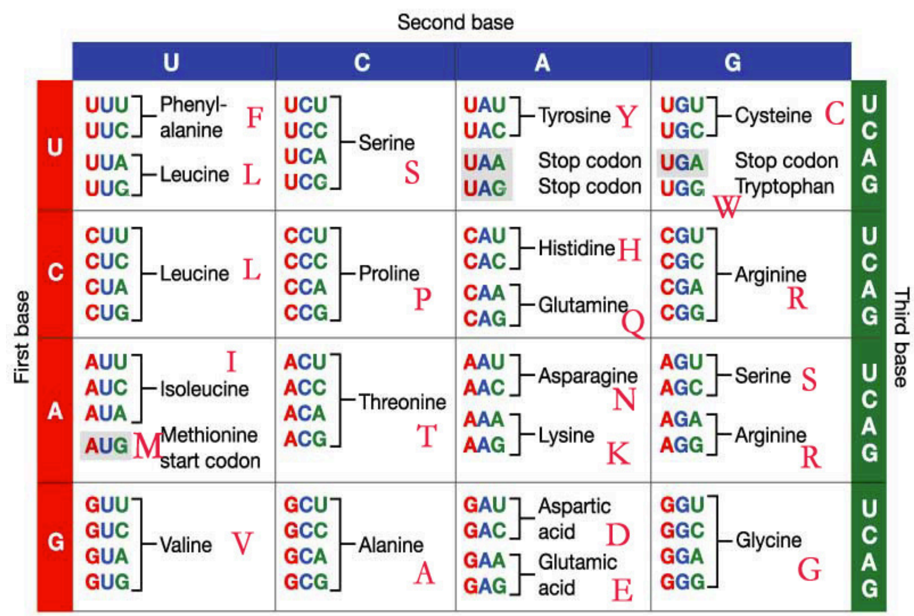
wobble codon base	possible anticodon bases
U	A, G, or I
C	G or I
A	U or I
G	C or U

eukaryotes

wobble codon base	possible anticodon bases
U	A, G, or I
C	G or I
A	U
G	C

Figure 6-51 Molecular Biology of the Cell 6e (© Garland Science 2015)

Degeneracy of codon table is related in part to wobble position in tRNA



Second base

First base	U	C	A	G	Third base
U	UUU, UUC – Phenyl-alanine F UUA, UUG – Leucine L	UCU, UCC, UCA, UCG – Serine S	UAU, UAC – Tyrosine Y UAA Stop codon UAG Stop codon	UGU, UGC – Cysteine C UGA Stop codon UGG Tryptophan W	U C A G
C	CUU, CUC, CUA, CUG – Leucine L	CCU, CCC, CCA, CCG – Proline P	CAU, CAC – Histidine H CAA, CAG – Glutamine Q	CGU, CGC, CGA, CGG – Arginine R	U C A G
A	AUU, AUC, AUA – Isoleucine I AUG M Methionine start codon	ACU, ACC, ACA, ACG – Threonine T	AAU, AAC – Asparagine N AAA, AAG – Lysine K	AGU, AGC – Serine S AGA, AGG – Arginine R	U C A G
G	GUU, GUC, GUA, GUG – Valine V	GCU, GCC, GCA, GCG – Alanine A	GAU, GAC – Aspartic acid D GAA, GAG – Glutamic acid E	GGU, GGC, GGA, GGG – Glycine G	U C A G

Ribosome is a ribozyme

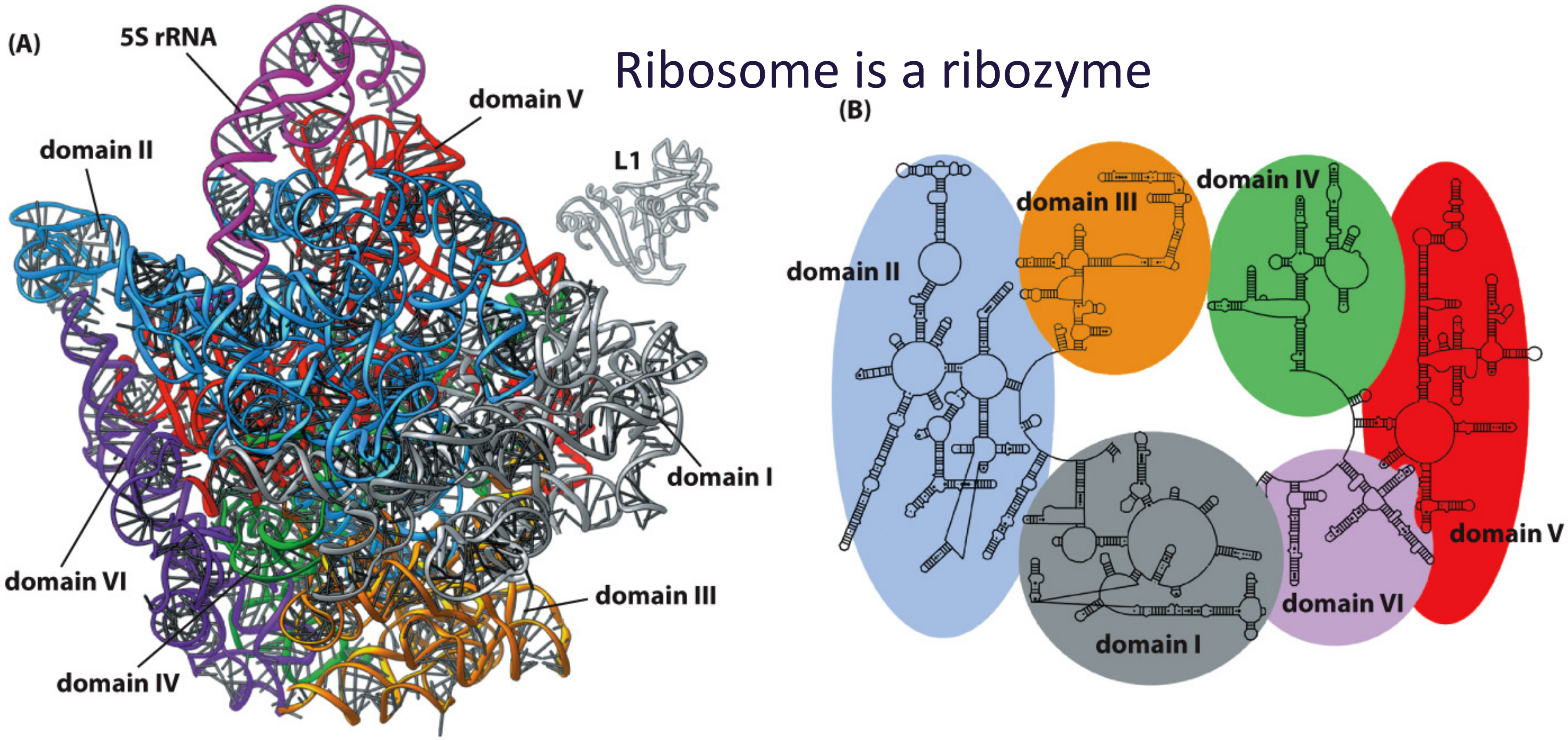


Figure 6-67 Molecular Biology of the Cell 6e (© Garland Science 2015)

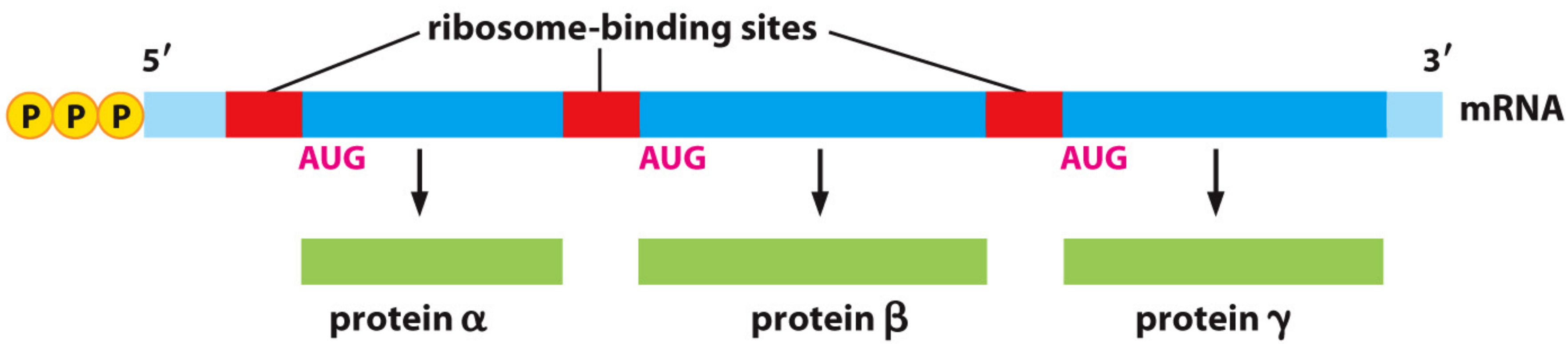
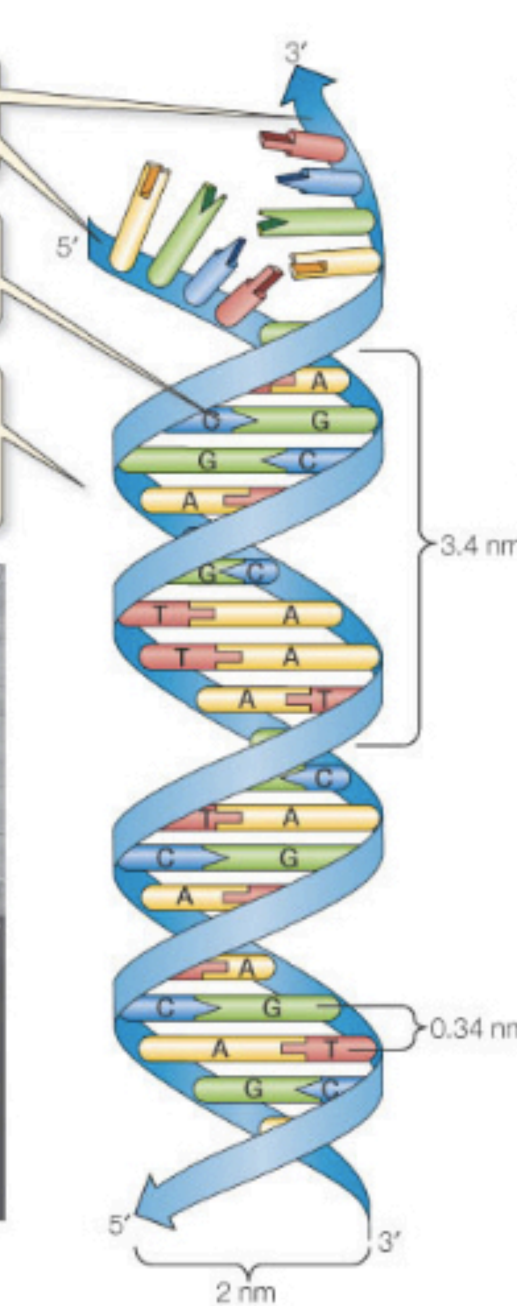
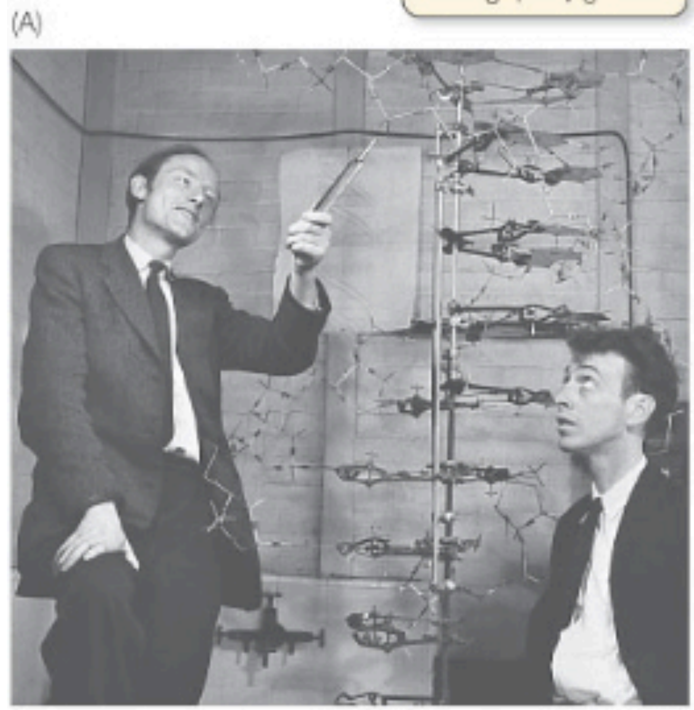


Figure 6-71 Molecular Biology of the Cell 6e (© Garland Science 2015)

Polycistronic bacterial mRNA have multiple internal ribosome entry sites (IRES)

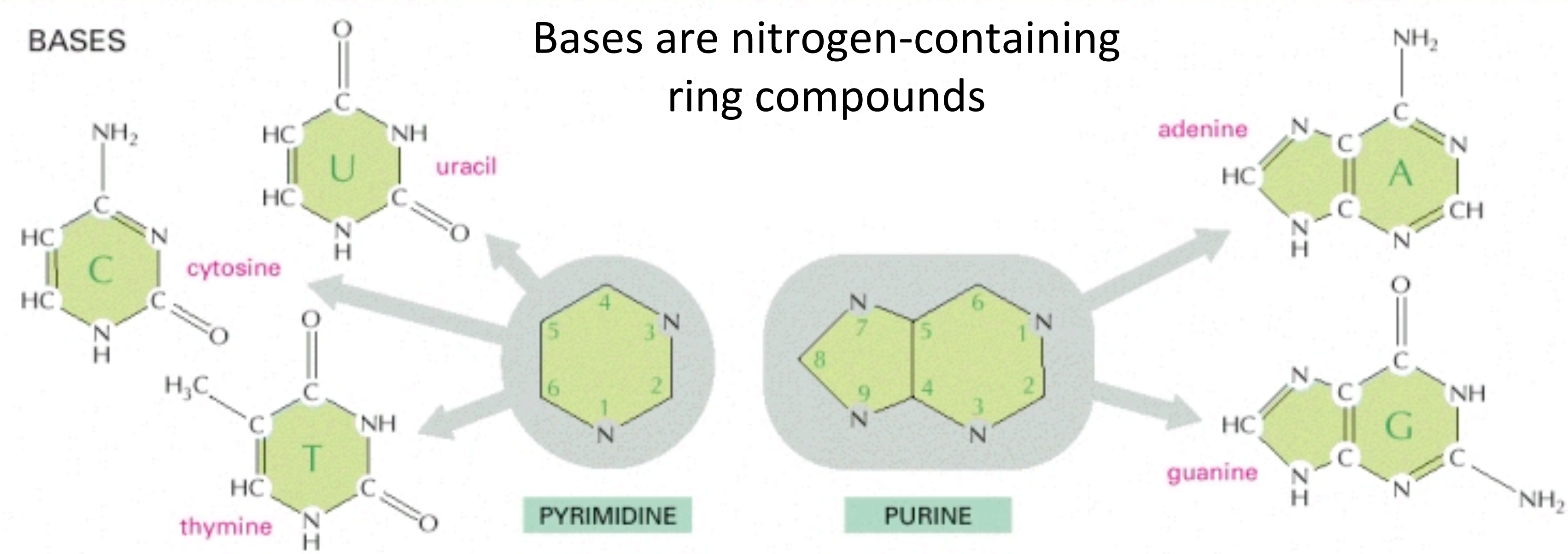
Let’s go EVEN smaller





http://www.nature.com/scitable/content/dna-is-a-double-helix-24263

Bases are nitrogen-containing ring compounds

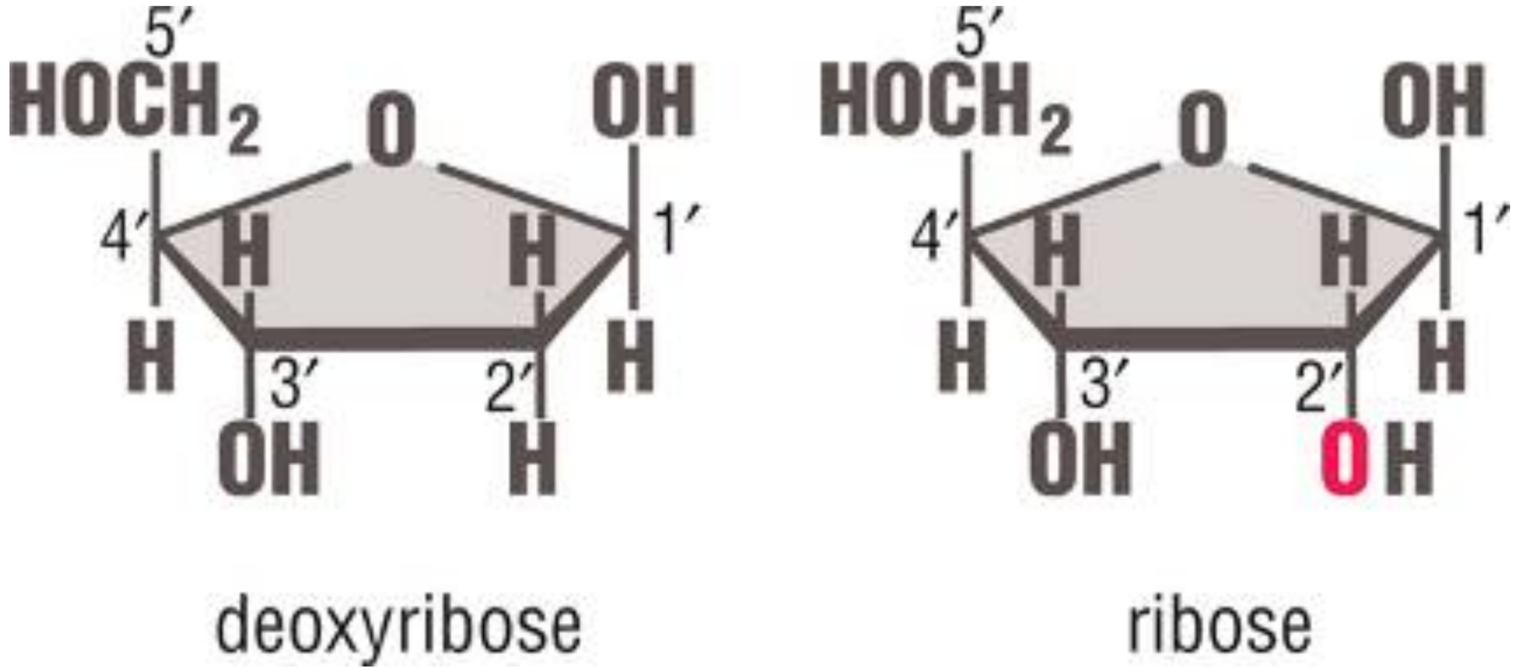


Pyrimidine bases have a single ring

Purine bases have two rings

Molecular Biology of the Cell. 4th edition. Alberts B, Johnson A, Lewis J, et al. New York: Garland Science; 2002.

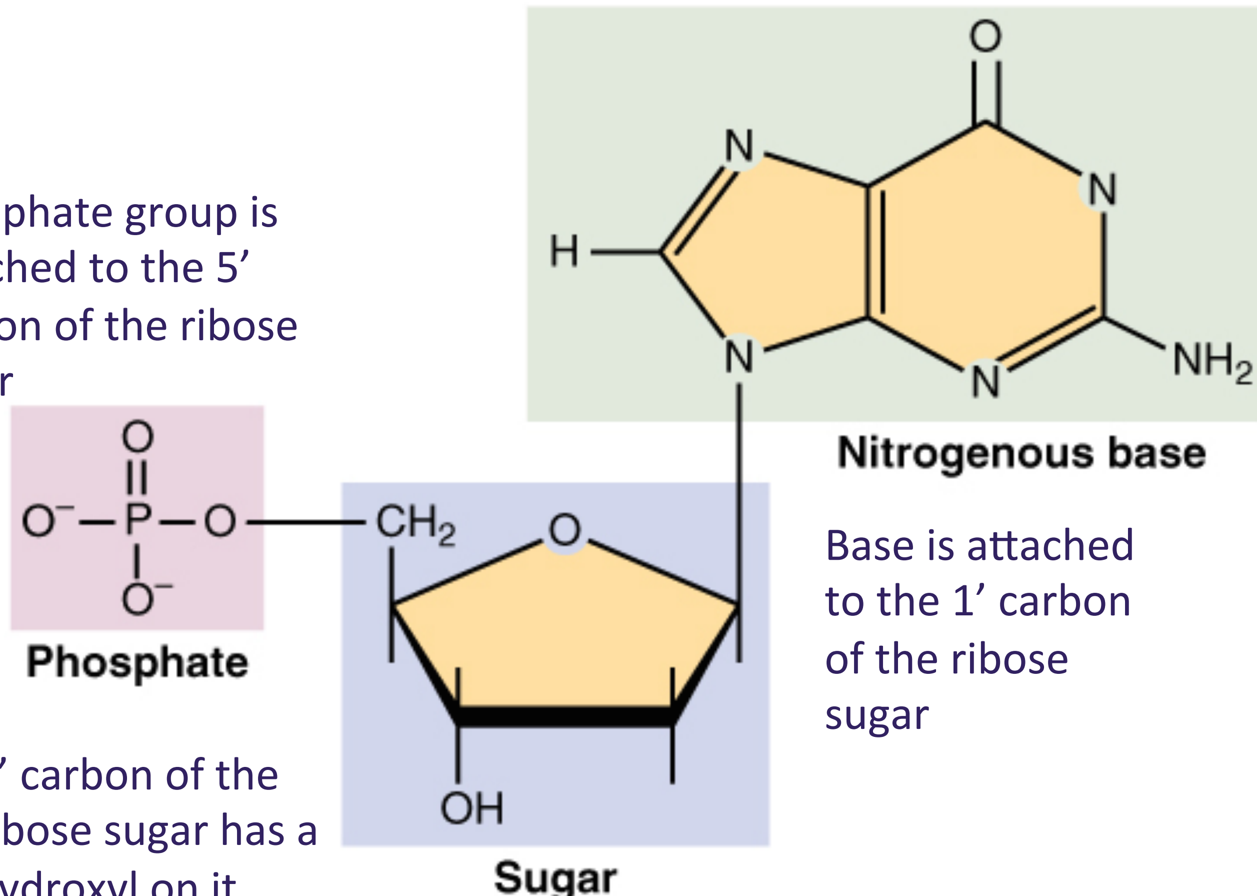
Nucleic acids have a ribose sugar ring as their central component



Craig, et al., *Molecular Biology,* Figure 02-19

Basic building block of DNA is a nucleotide

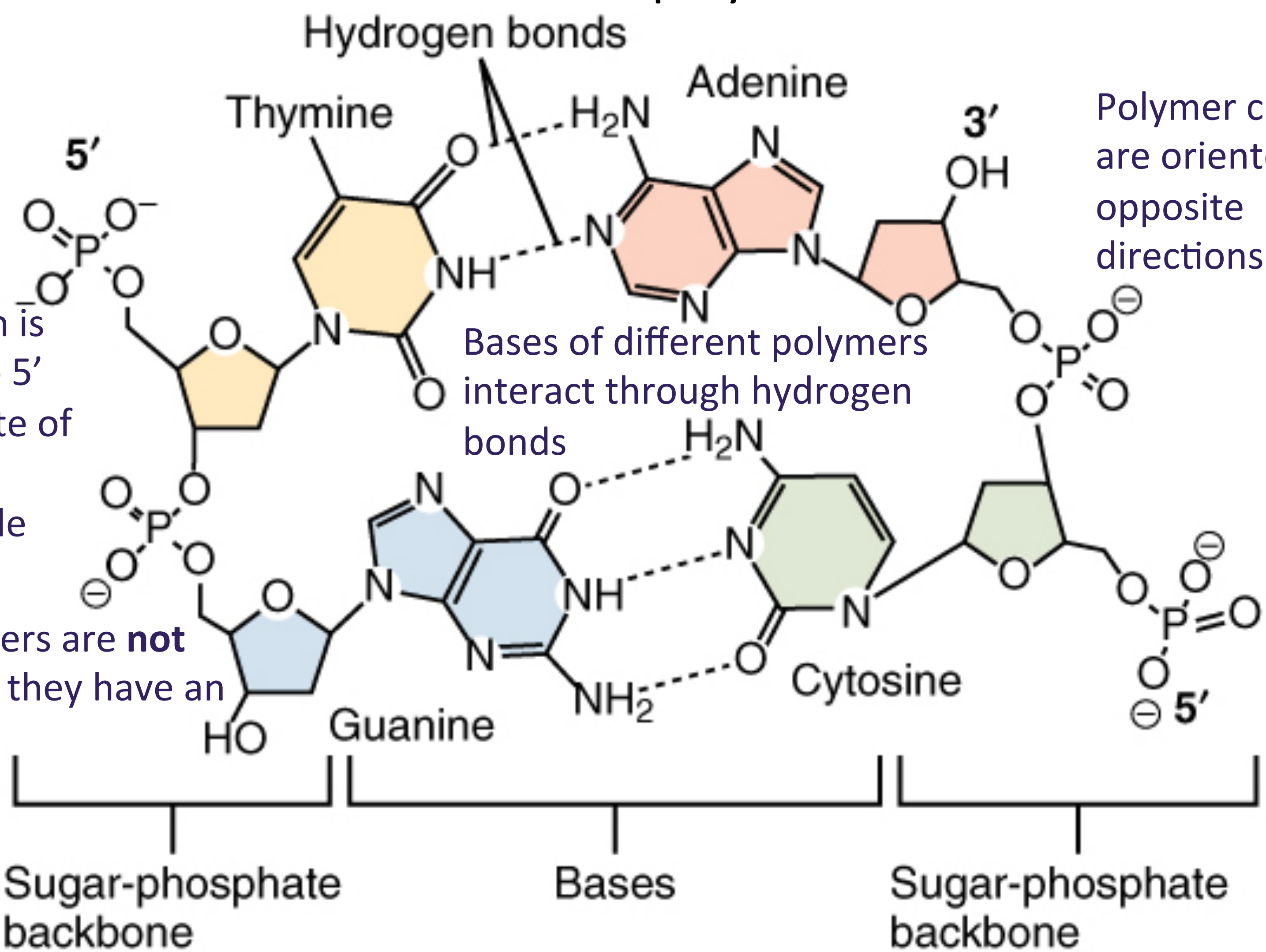
Phosphate group is attached to the 5' carbon of the ribose sugar



Base is attached to the 1' carbon of the ribose sugar

3' carbon of the ribose sugar has a hydroxyl on it

Nucleotides are linked through phosphate bonds to form DNA polymers

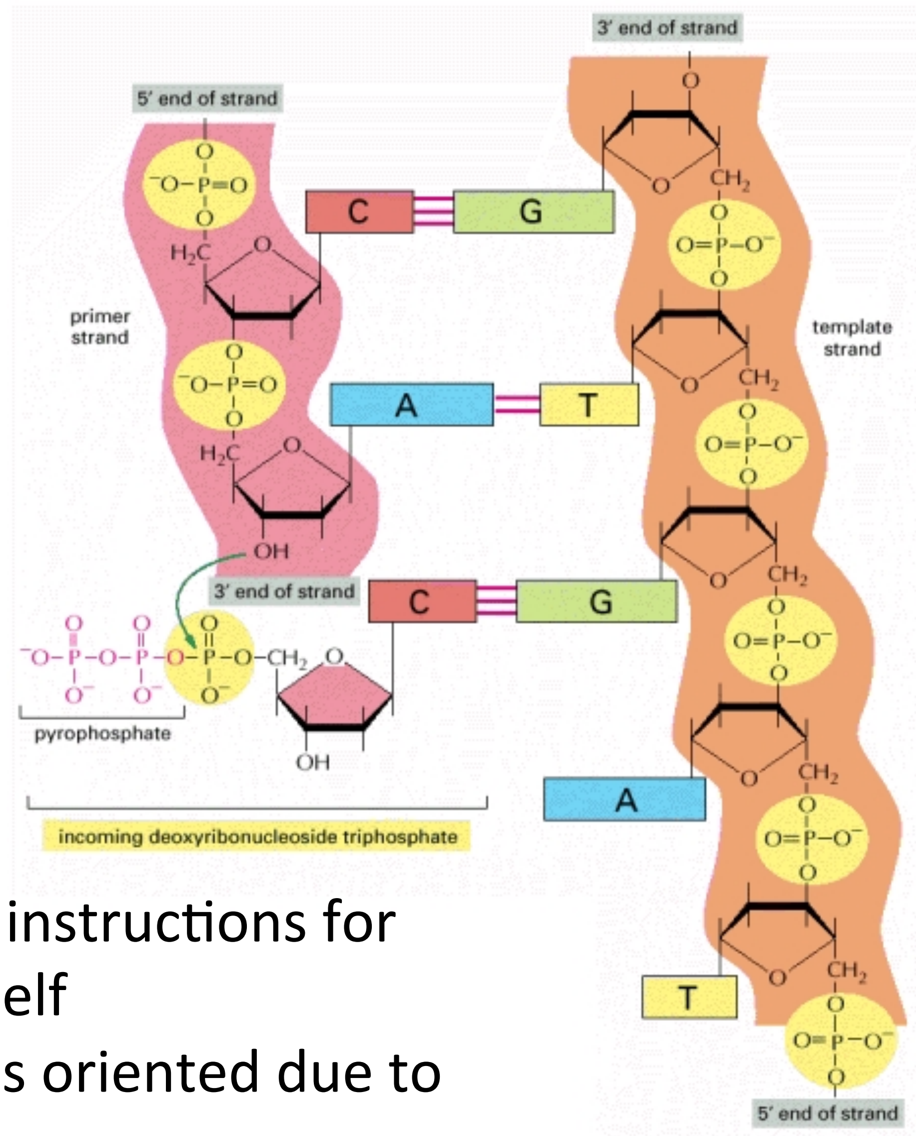


3′ oxygen is bound to 5′ phosphate of the next nucleotide

Bases of different polymers interact through hydrogen bonds

Polymer chains are oriented in opposite directions

DNA polymers are **not** symmetric; they have an orientation



DNA contains instructions for duplicating itself

This reaction is oriented due to chemistry

Can only add to 3’ end of a strand, biologically

Molecular Biology of the Cell. 4th edition.
Alberts B, Johnson A, Lewis J, et al.
New York: Garland Science; 2002.

DNA Polymerase

- Builds polymers of DNA
- Needs a DNA template
- Can only add to growing strand, can't start from scratch
- Many different versions

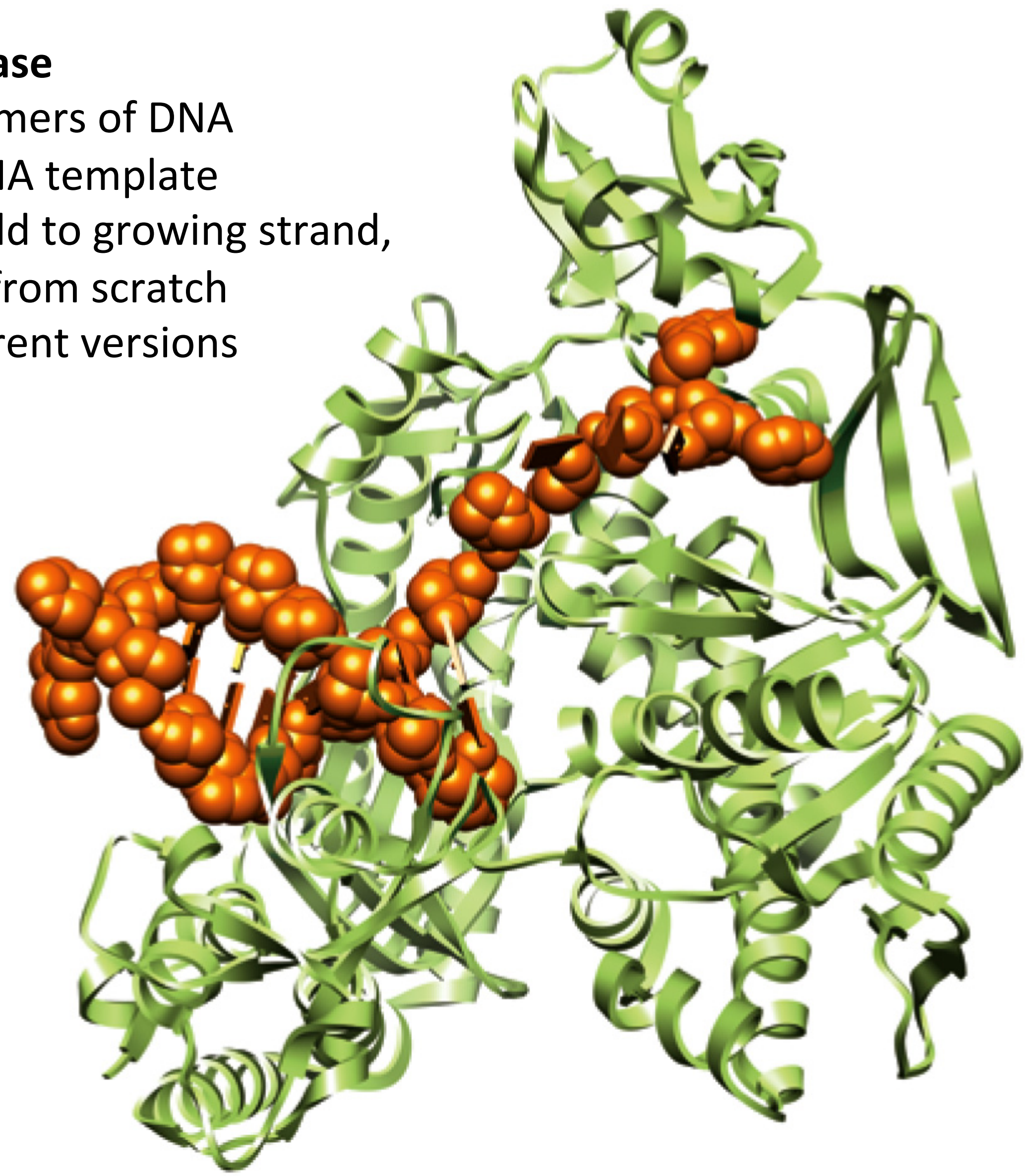
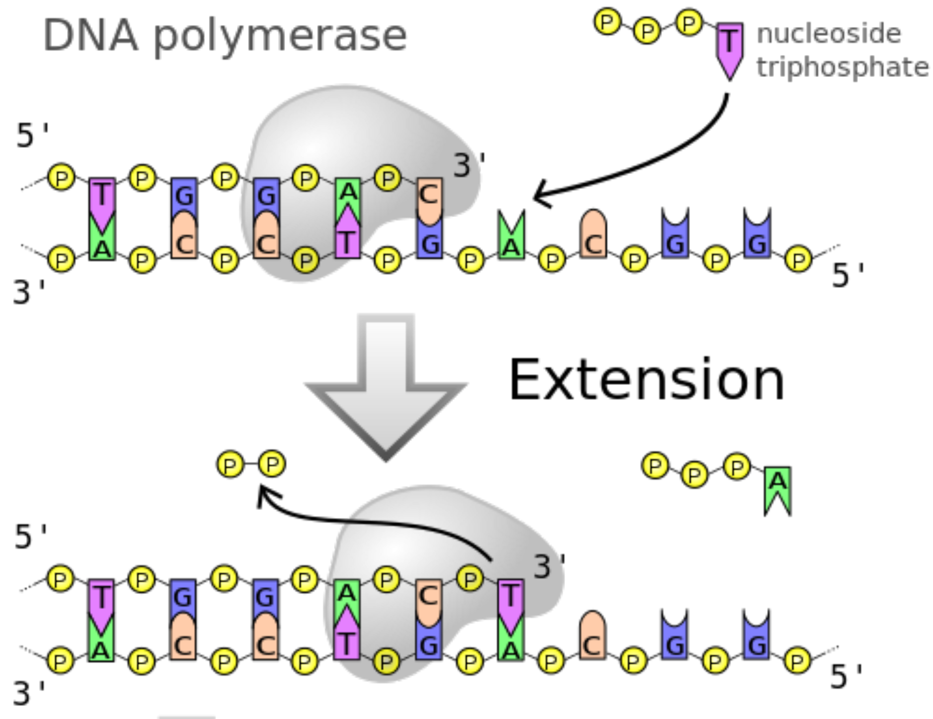
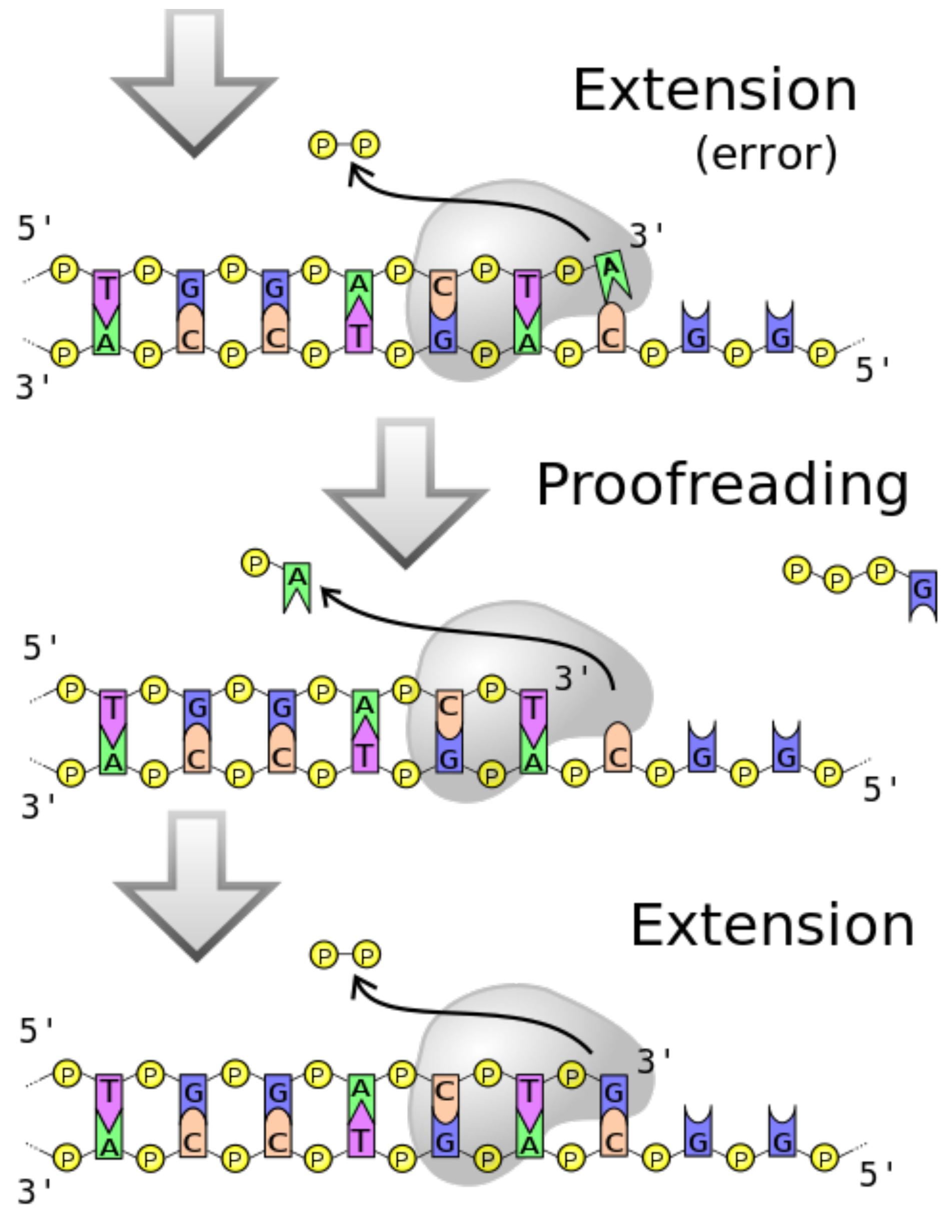


Figure 5-4b Molecular Biology of the Cell 6e (© Garland Science 2015)





Even with proofreading, mistakes made every 10^7-10^9 bases
6 billion bases in human genome!

Translation Exercises