BI 8 LECTURE 2

NUCLEIC ACID STRUCTURES: CHEMISTRY, SPACE, TIME (EVOLUTION)

Ellen Rothenberg

7 January 2016

Reading for 1st two lectures

Be aware: Alberts et al. has a different game plan than this course…

Alberts, et al. (6th edition)

- Ch. 4: pp. 173-186 & 216-236
- Ch. 6 pp. 299-301
- Panel 2-6 & related discussion in Ch. 2
- (see panels 2-1, 2-2, 2-3 for background; & reference pages at end of these notes)

Reading ahead for Tues Jan 12:

- Ch. 6: pp. 301-333

Three types of nucleic acids in this course

- Genomic DNA
- RNA
- cDNA

Three types of nucleic acids in this course

- Genomic DNA
  - Natural
  - Double-stranded
  - Stable
- RNA

- cDNA

Three types of nucleic acids in this course

- Genomic DNA
- RNA
  - Natural
  - Single-stranded
  - May have variable secondary structure (folding)
  - May be very unstable or moderately unstable
  - Only a subset of genomic sequence
- cDNA

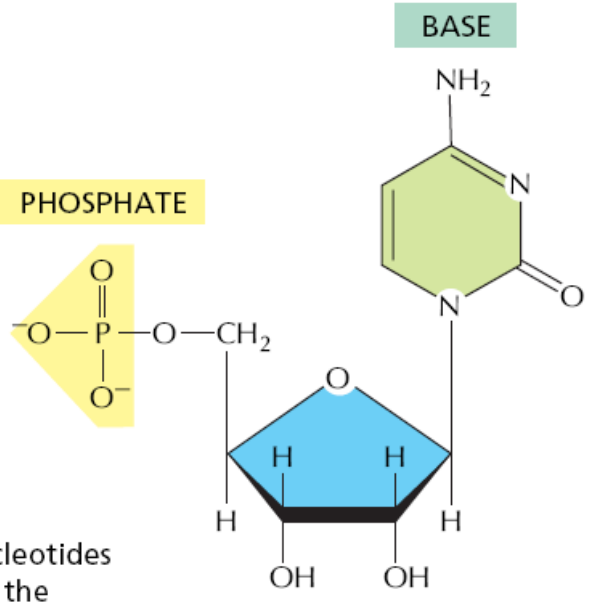
Three types of nucleic acids in this course

- Genomic DNA
- RNA
- cDNA
  - Artificial, not natural
  - An experimental convenience to fabricate a (stable) DNA version of an RNA sequence
  - Initially synthesized as *complementary* strand to RNA, then second strand can be made

Nucleotides: tripartite building blocks of DNA and RNA

NUCLEOTIDES

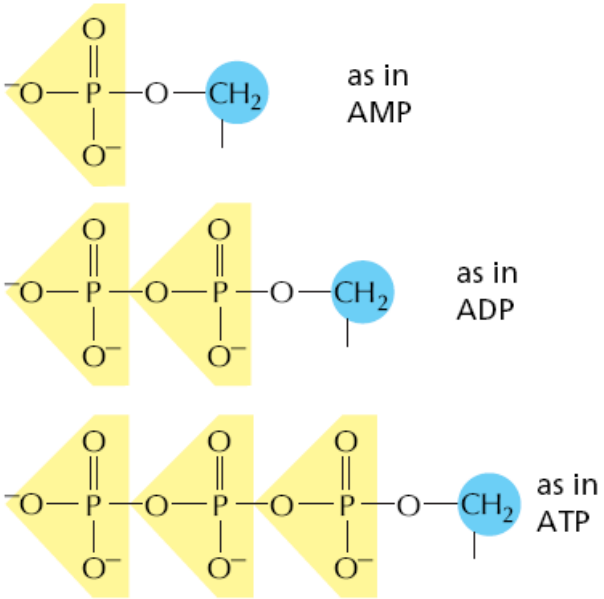
A nucleotide consists of a nitrogen-containing base, a five-carbon sugar, and one or more phosphate groups.



Nucleotides are the subunits of the nucleic acids.

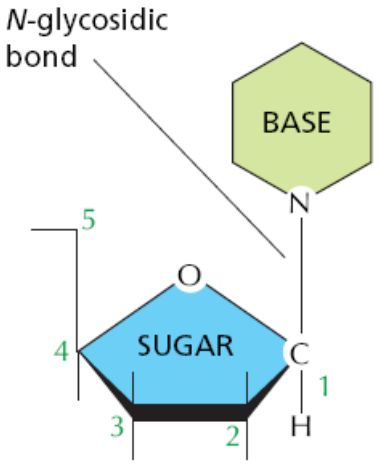
PHOSPHATES

The phosphates are normally joined to the C5 hydroxyl of the ribose or deoxyribose sugar (designated 5'). Mono-, di-, and triphosphates are common.



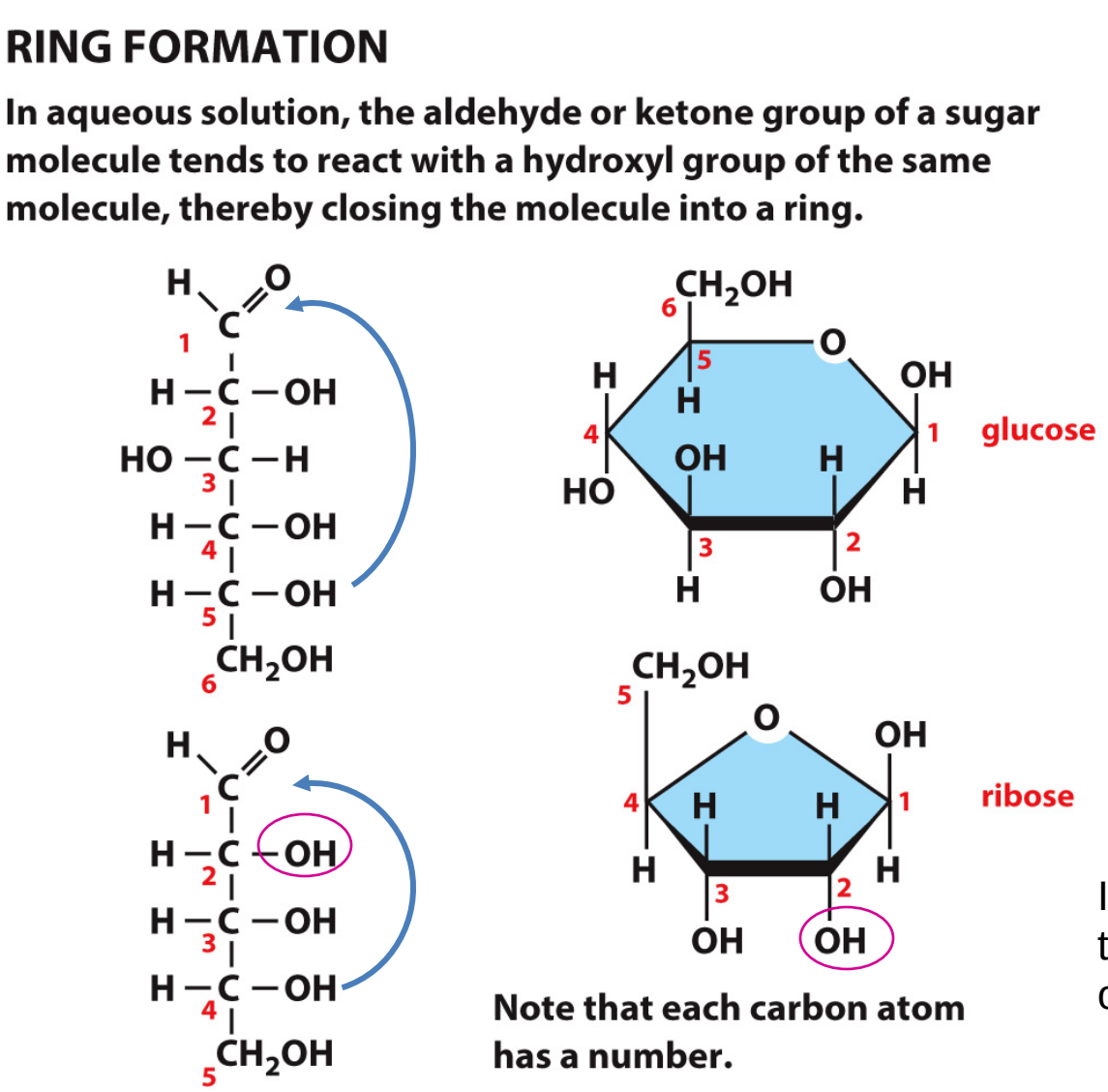
The phosphate makes a nucleotide negatively charged.

BASIC SUGAR LINKAGE



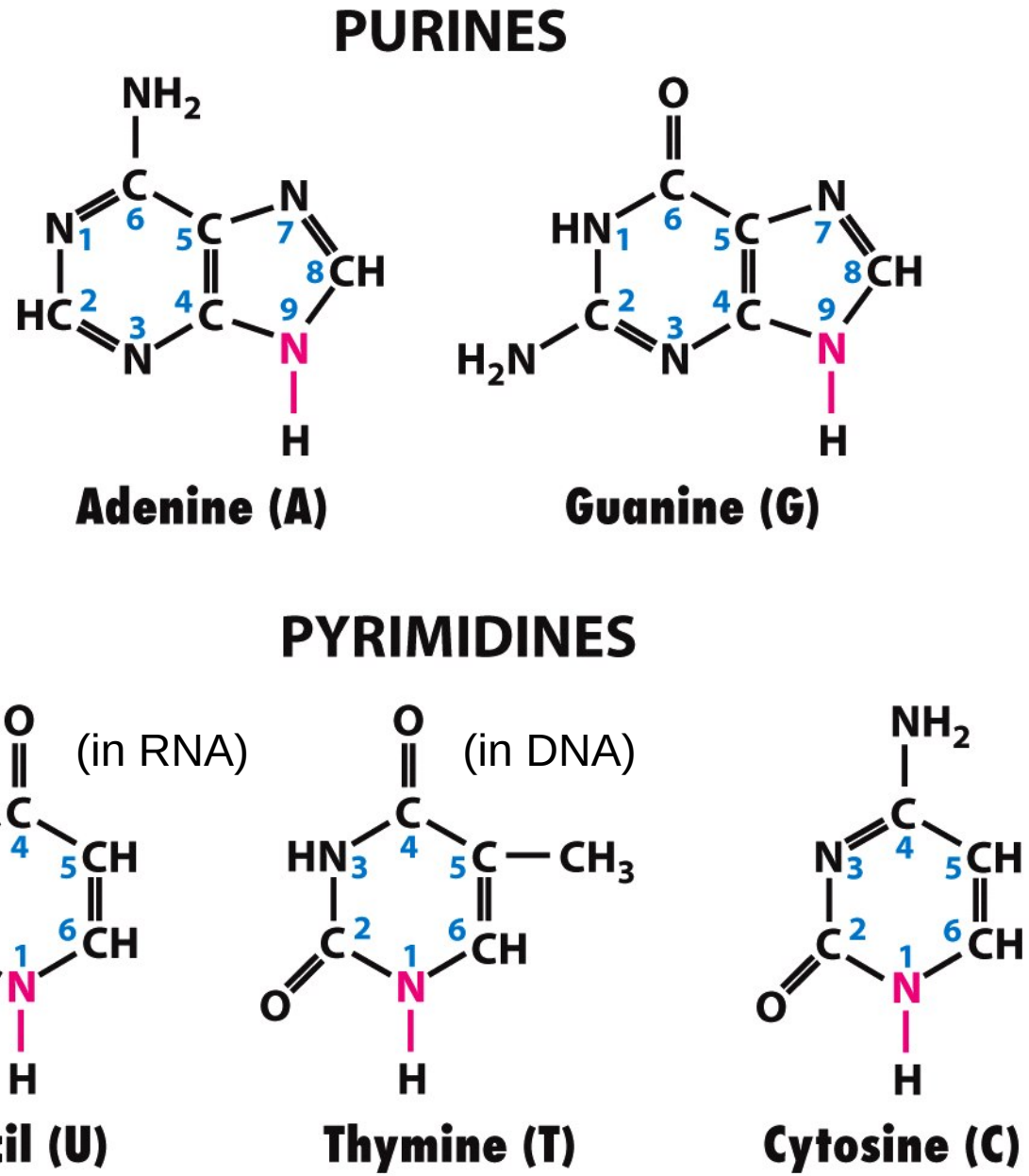
The base is linked to the same carbon (C1) used in sugar–sugar bonds.

Sugar moiety: cyclized ribose or deoxy ribose



In deoxyribose, this is –H instead of –OH

Two families of bases: purines (2-ring) and pyrimidines (1-ring)



Lie “flat” because of resonance in conjugated bond structures

Figure 2-17
Molecular Cell Biology, Sixth Edition

Linkages between nucleotides are asymmetric

A is connected to C through its 5’ hydroxyl but to G through its 3’ hydroxyl

GAC is a different sequence from CAG

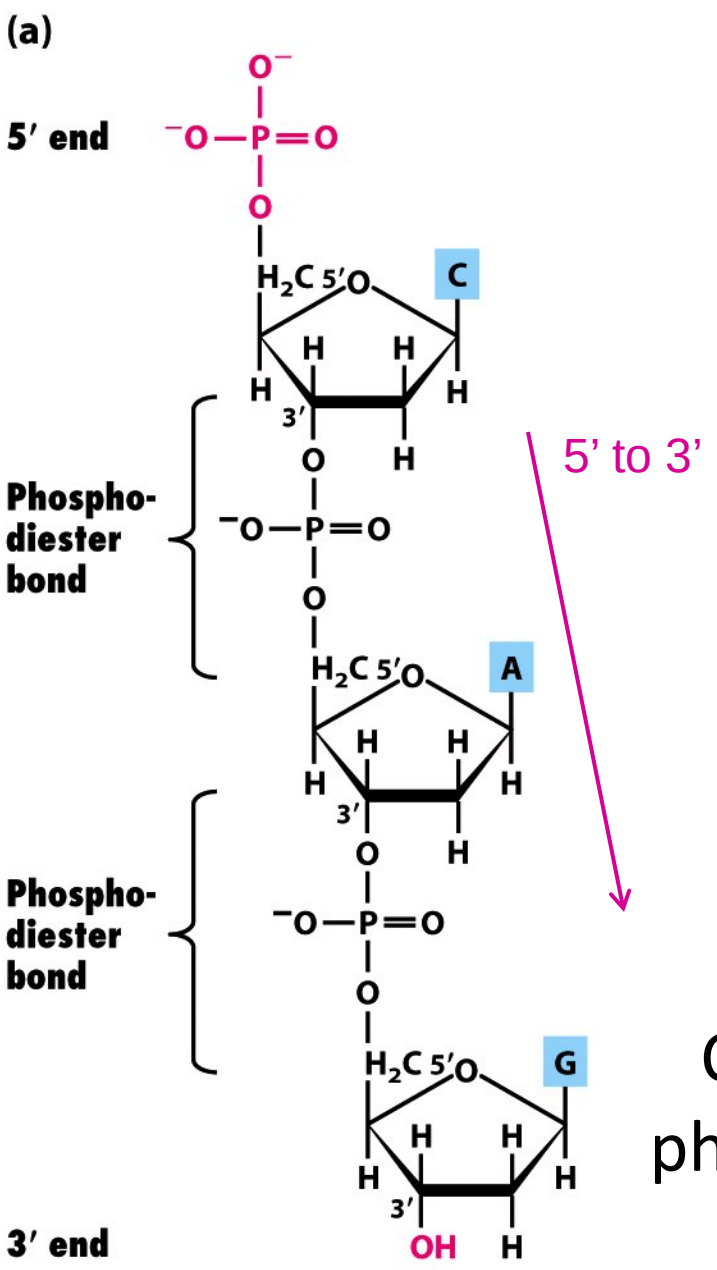
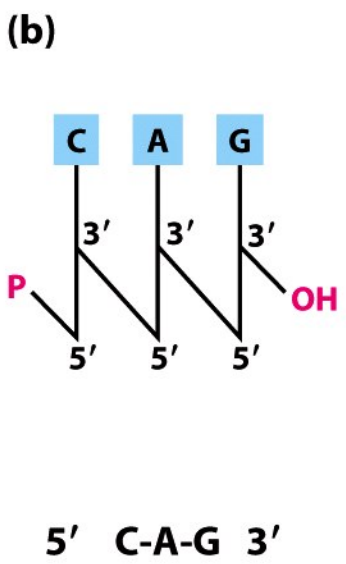


Figure 4-2
Molecular Cell Biology, Sixth Edition

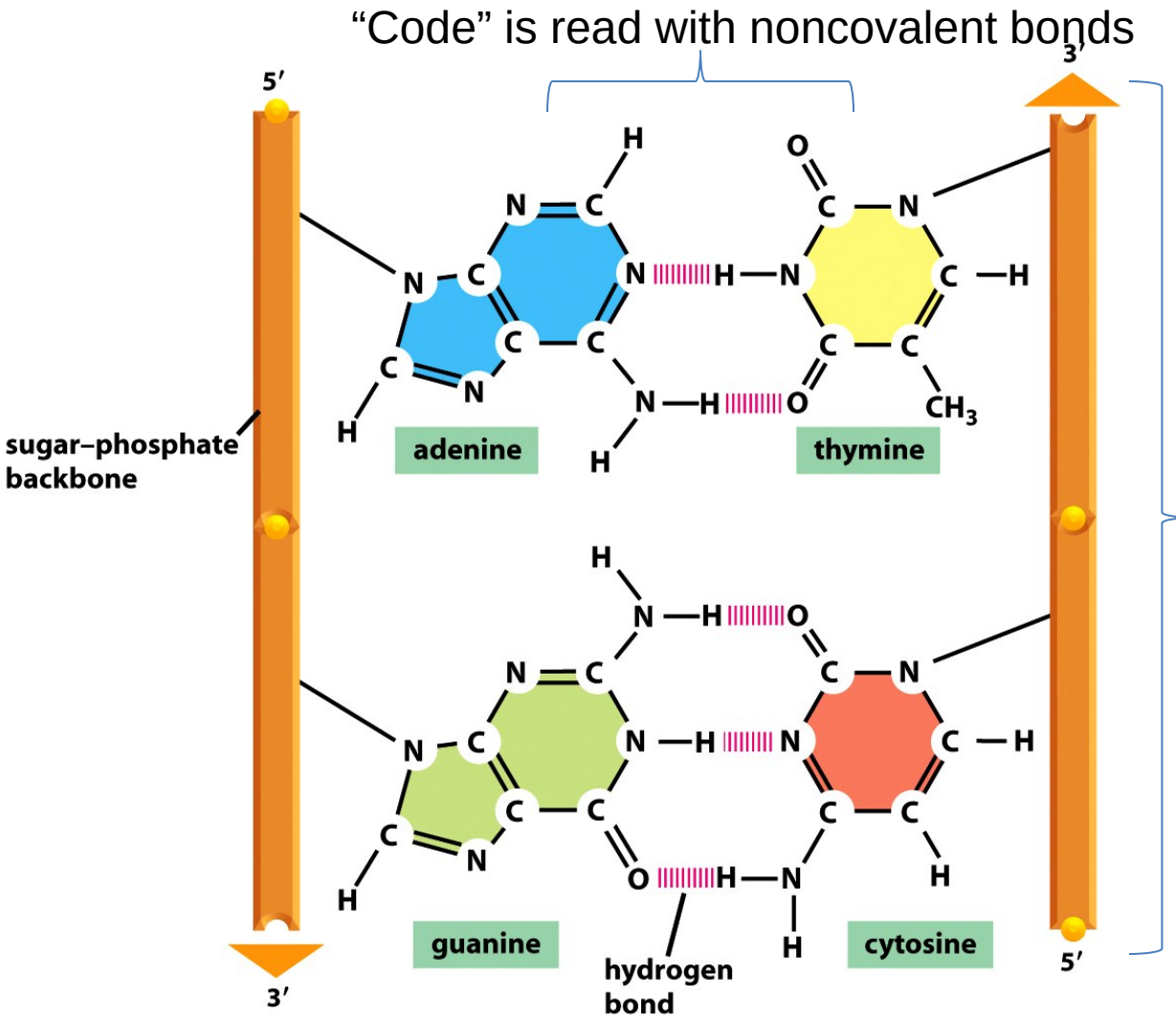


By convention, DNA sequences are written in 5’ to 3’ order

DNA, RNA structure basics

Covalent linkage of 5’ phosphate to 3’ hydroxyl on next sugar ring polymerizes chain

A mixture of covalent and noncovalent bonds hold DNA together



Covalent bonds: Link parts of nucleotides together –

Also link individual nucleotides together in sequence along one strand of DNA or RNA

Figure 4-4 *Molecular Biology of the Cell* (© Garland Science 2008)

Both covalent and different kinds of noncovalent bonds are crucial for macromolecular structure and function

Table 2–1 Covalent and Noncovalent Chemical Bonds

BOND TYPE		LENGTH (nm)	STRENGTH (kcal/mole) * IN VACUUM	IN WATER
Covalent		0.15	90	90
Noncovalent:	ionic*	0.25	80	3
	hydrogen	0.30	4	1
	van der Waals attraction (per atom)	0.35	0.1	0.1

1 kcal ≈ 4.2 kJ

*An ionic bond is an electrostatic attraction between two fully charged atoms.

Ionic: charged (- to +, + to -), like Mg^{2+} with $-PO_3^{2-}$ and $-\overset{\overset{O}{\|}}{C}O^-$ to $-NH_3^+$

Van der Waals: hydrophobic (very important in protein folding and membranes)

H-bonds: the glue that holds DNA and DNA:RNA duplexes together

Table 2-1 *Molecular Biology of the Cell* (© Garland Science 2008)

Bond energies spanning >2 orders of magnitude of strength are involved in molecular biology

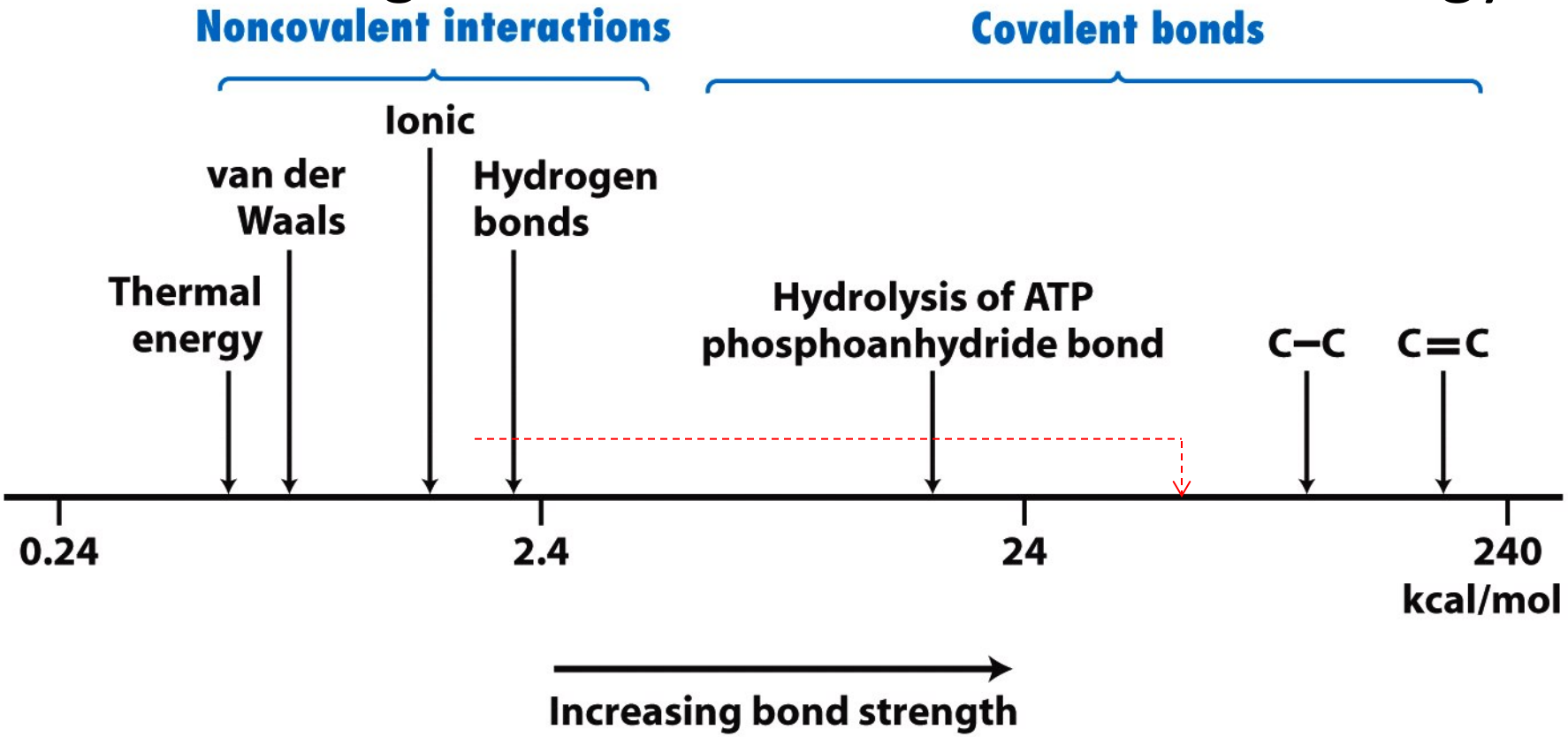


Figure 2-6
Molecular Cell Biology, Sixth Edition

- Diagram for typical biological conditions, in aqueous medium with salt… where OH-, H+, and other charge-balancing ions give ionic group “A” many alternatives to bonding with ionic group “B”
- Ionic interactions ***much*** stronger in nonaqueous, with low salt

Hydrogen bonds to water are secret sharers in all biological structures: included in polar regions, excluded from nonpolar

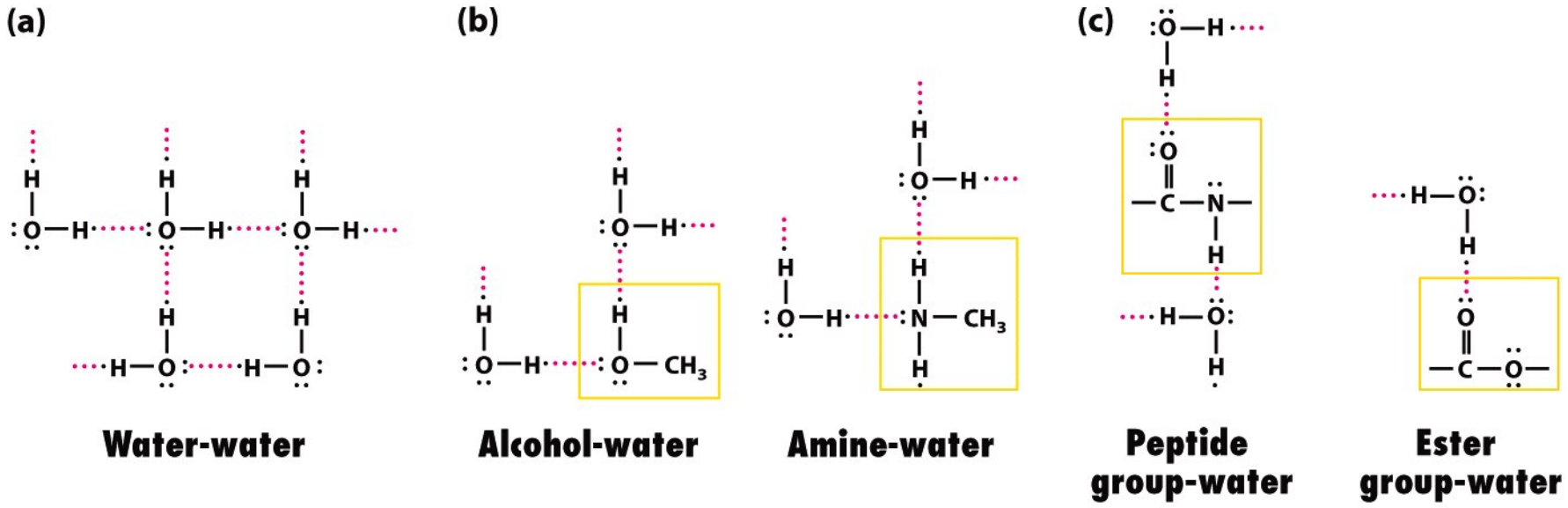
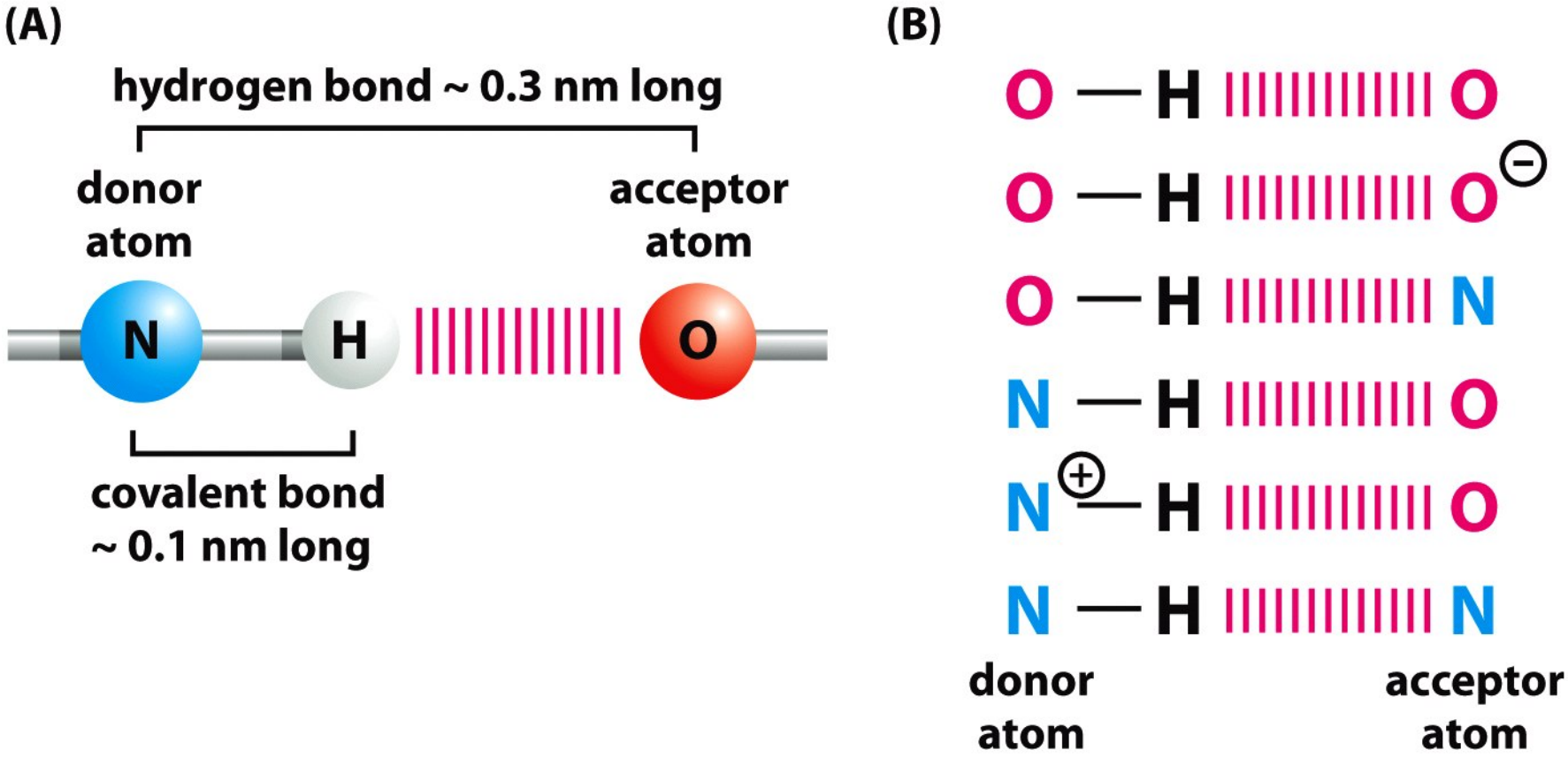


Figure 2-8
Molecular Cell Biology, Sixth Edition

(Amide)

Amine-carboxyl interactions and other donor-acceptor interactions in the right orientation can form H-bonds



0.3 nm = 3 Å

1 Å = 1 Ångström

Figure 2-15 *Molecular Biology of the Cell* (© Garland Science 2008)

Strategically placed NH, ring N, NH_2, and C=O groups distinguish DNA bases

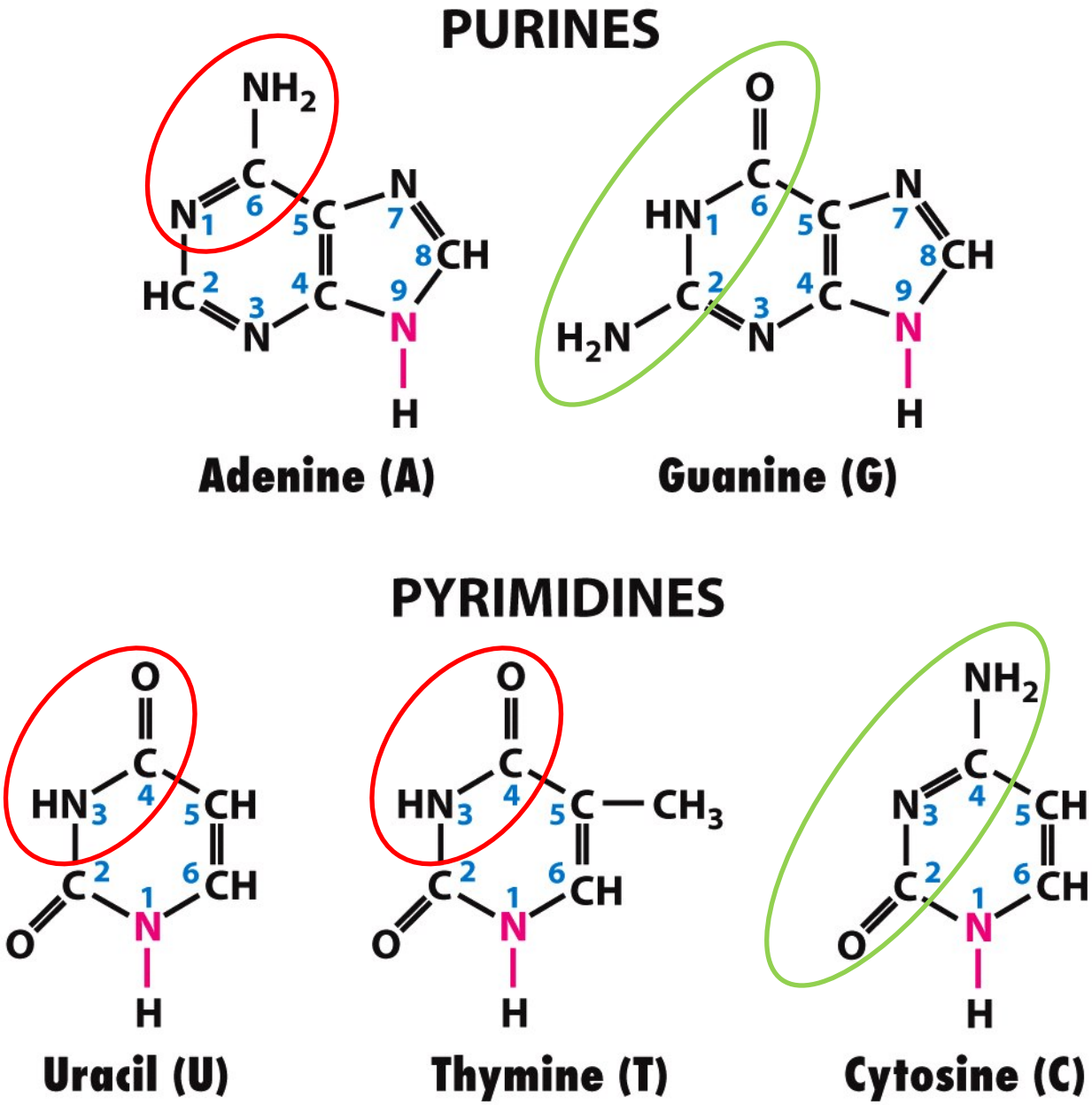


Figure 2-17
Molecular Cell Biology, Sixth Edition

Planar stacking interactions between the flat rings of the base pairs help to stabilize DNA double helix

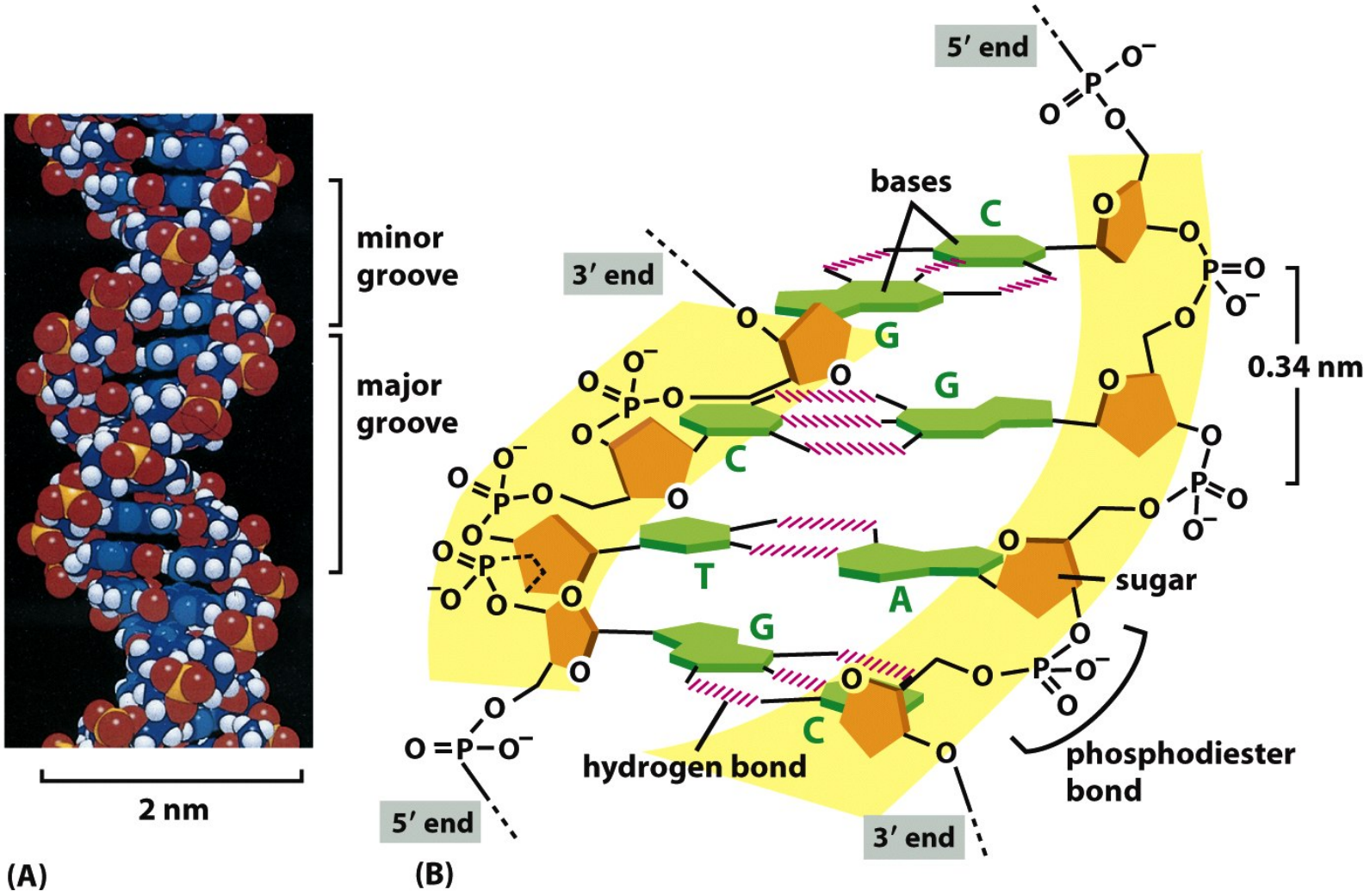
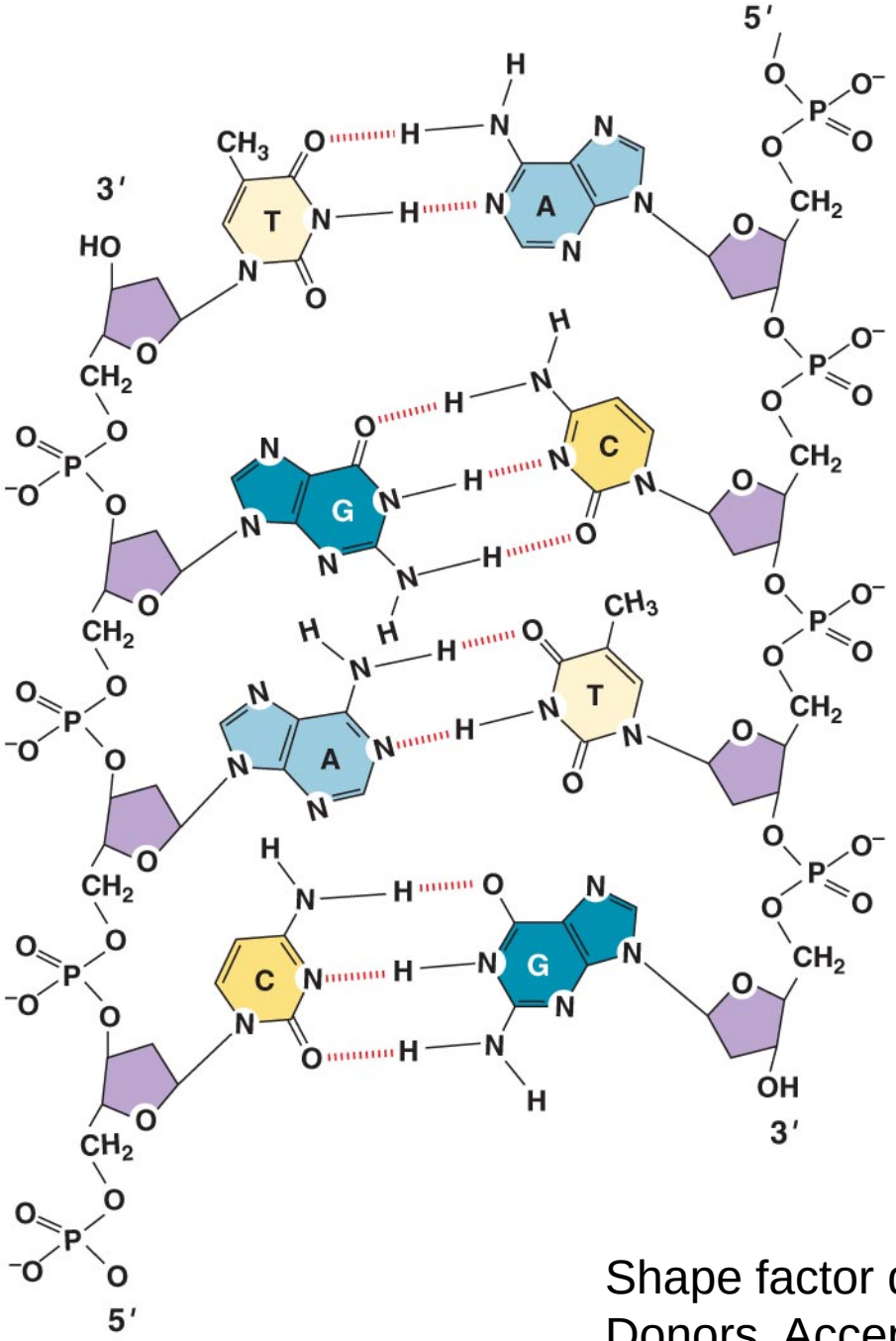
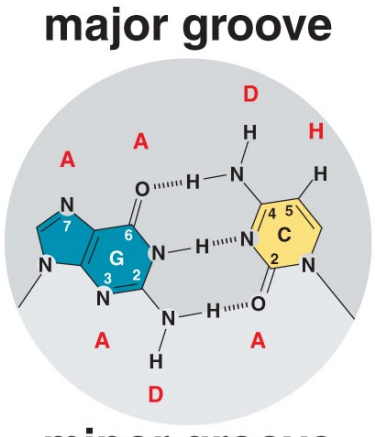
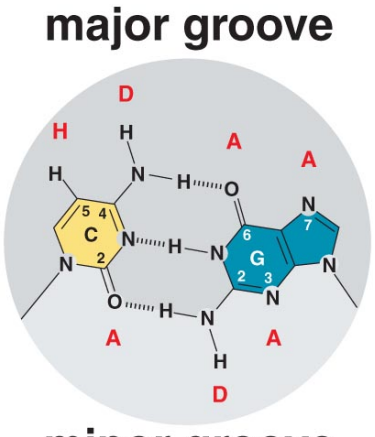


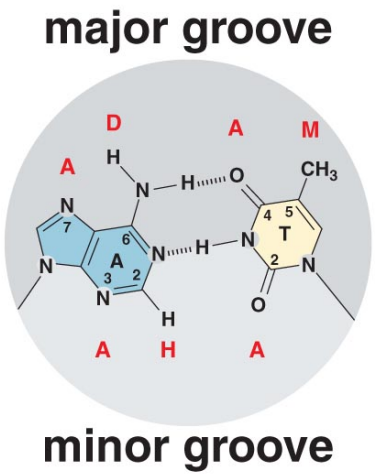
Figure 4-5 *Molecular Biology of the Cell* (© Garland Science 2008)

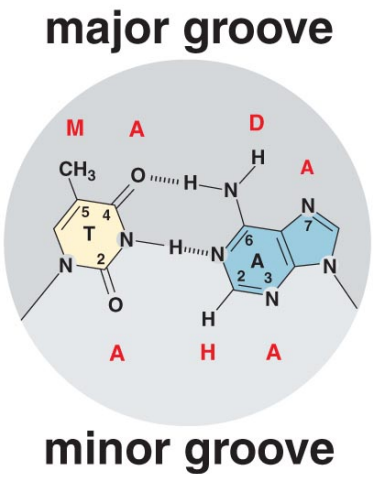
Base pairing up close: A-T hollows in the minor groove











Shape factor differences *and* distinct patterns of H-bond Donors, Acceptors for each base pair

Different sugars are used for backbones of RNA and DNA: one O different

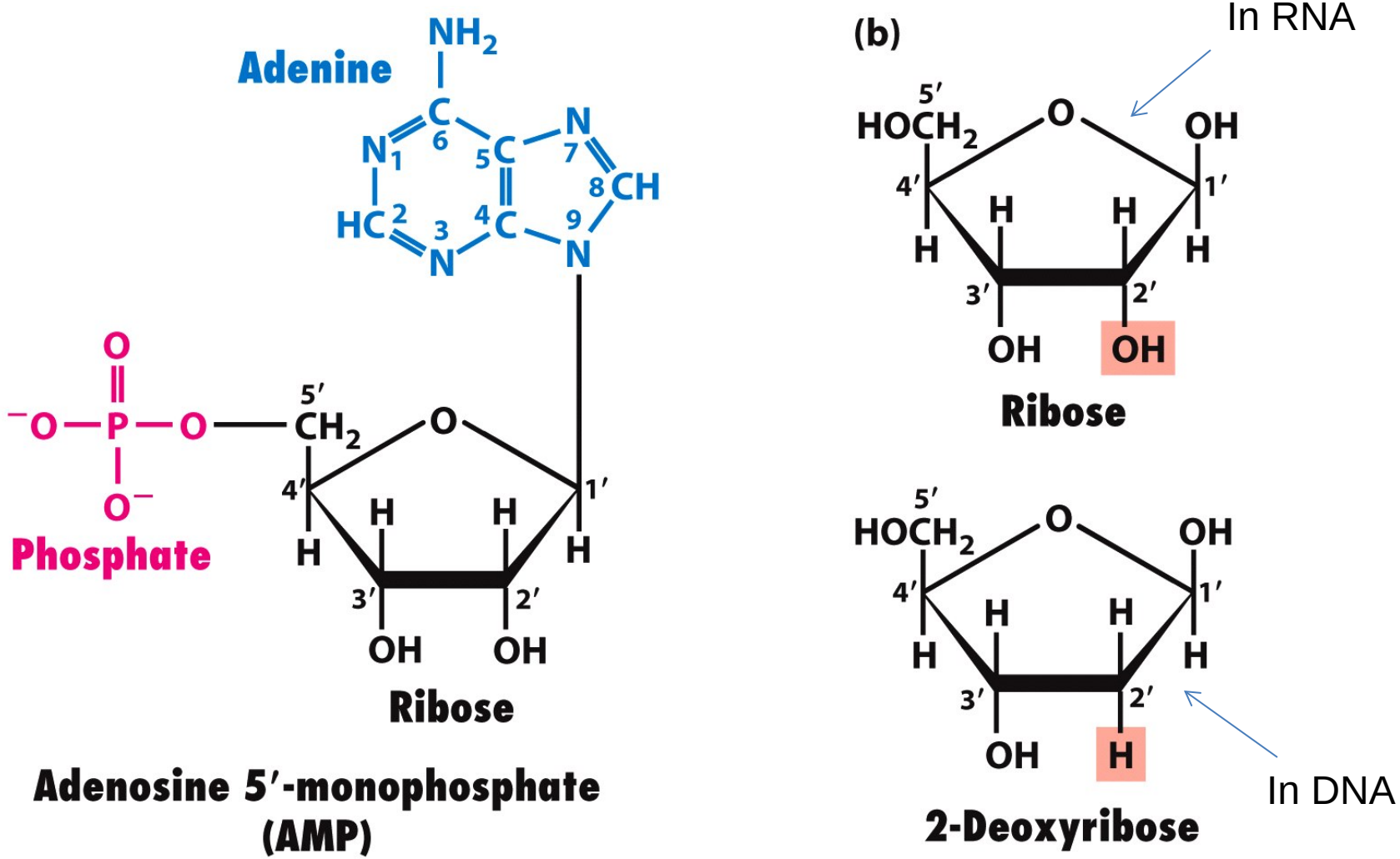
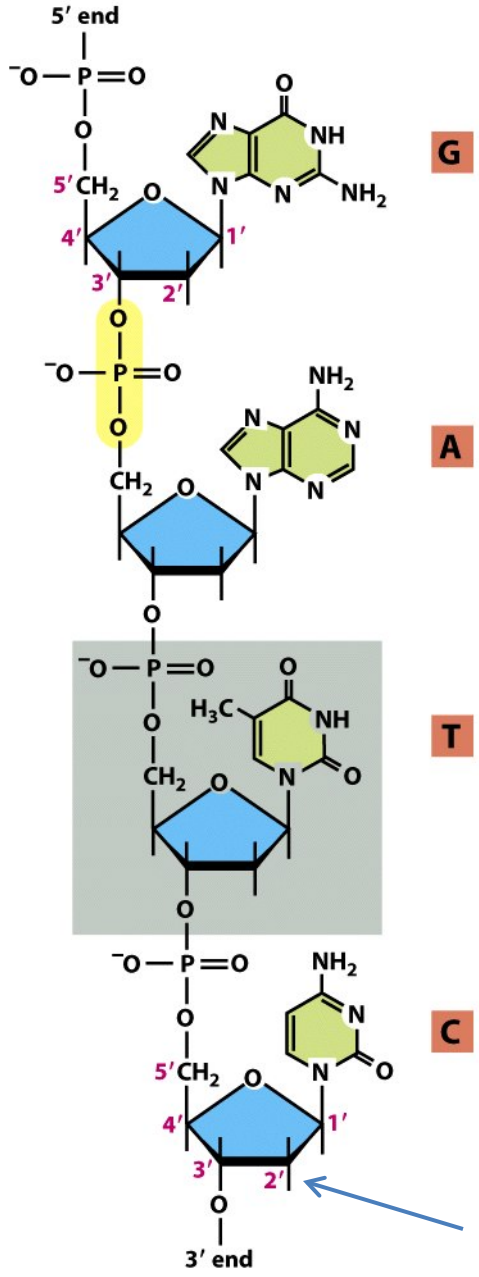
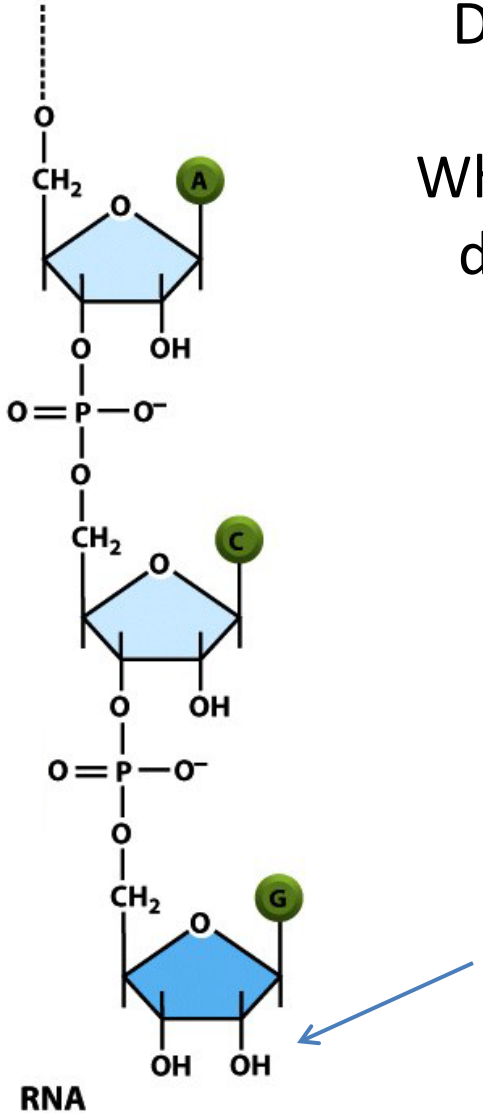


Figure 2-16
Molecular Cell Biology, Sixth Edition





DNA vs. RNA:

What difference does a 2’-OH make?

Figure 2-28 *Molecular Biology of the Cell* (© Garland Science 2008)

RNA 2’-OH , at basic pH, becomes a weapon to attack the phosphodiester linkage and cleave its own chain

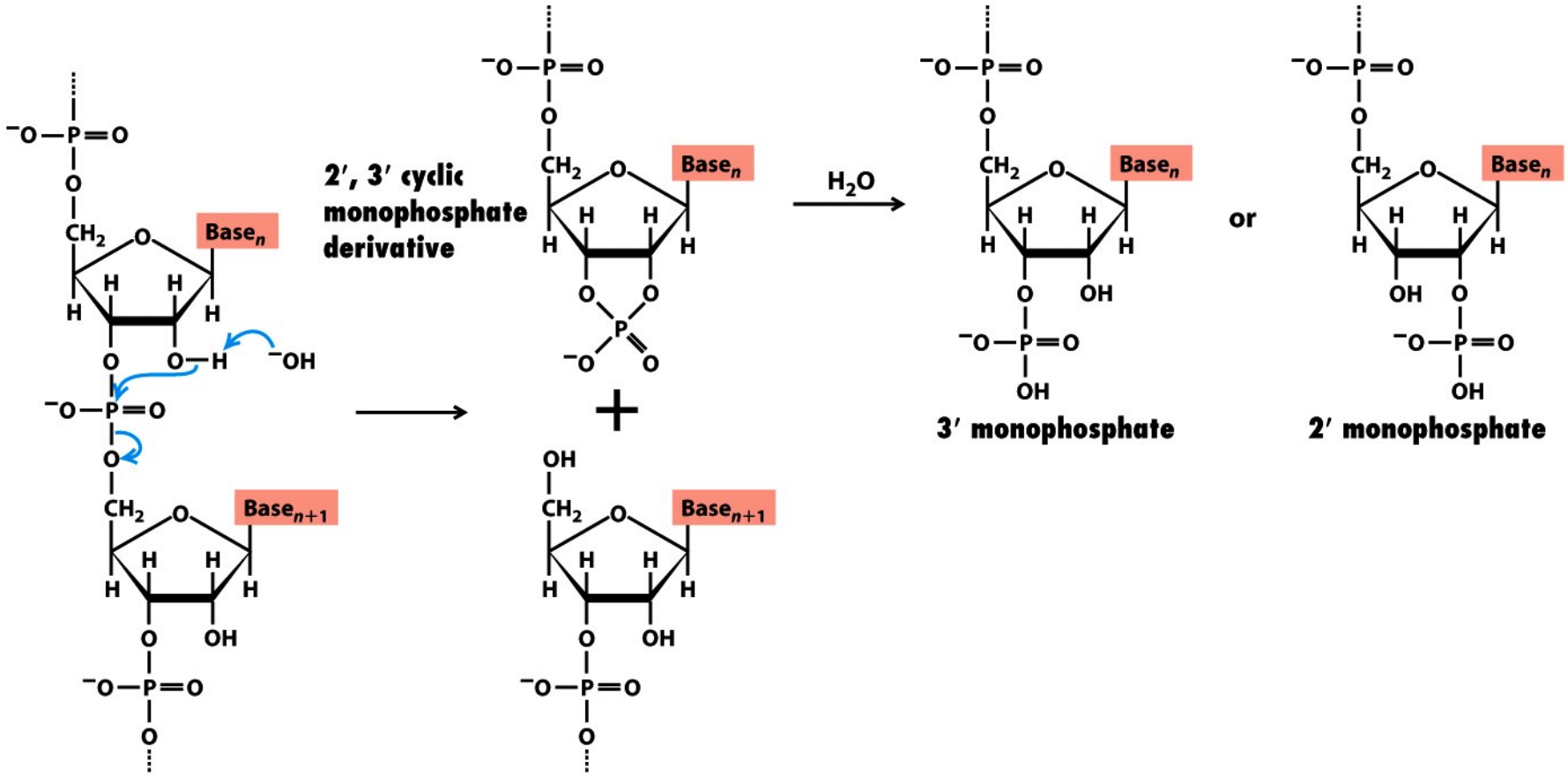
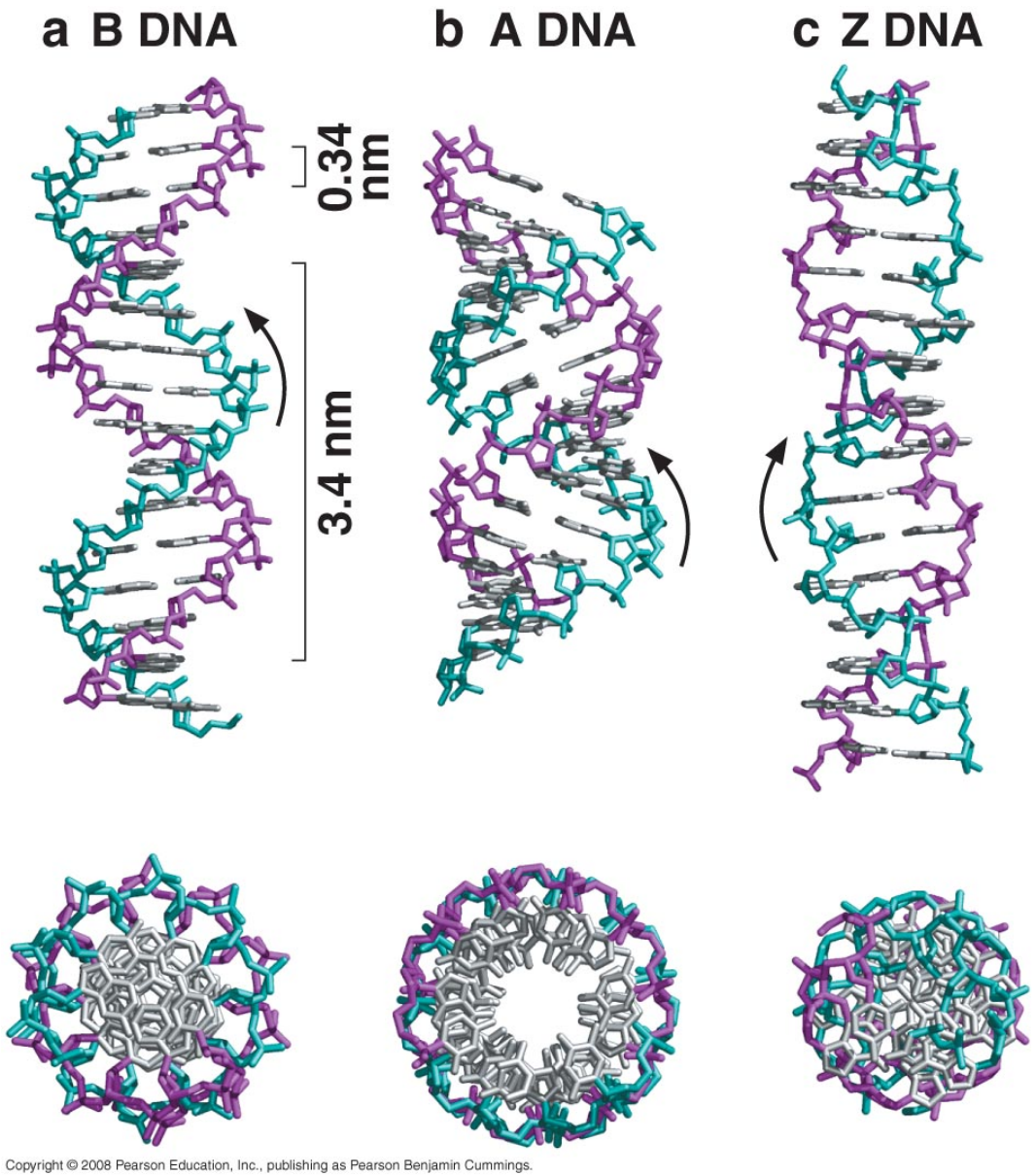


Figure 4-6
Molecular Cell Biology, Sixth Edition

Major form of DNA helix is B-DNA, but bases can be tilted in other angles for other geometries

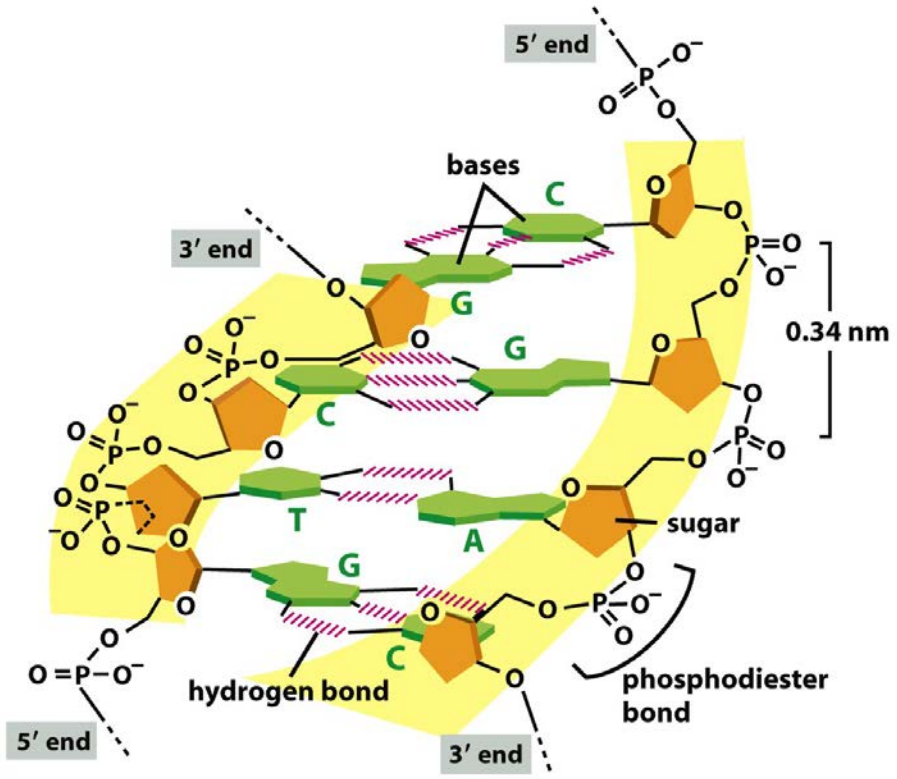
"A" form: typical of RNA-DNA hybrids and RNA-RNA duplexes
(2'-OH torques ribose rings out of hydrophobic interior)

"Z" form: adopted by certain sequences (TGTGTG:CACACA): recognized for regulation??



How does the phosphodiester backbone affect DNA stability?

How do two highly negatively charged polymers stay together?



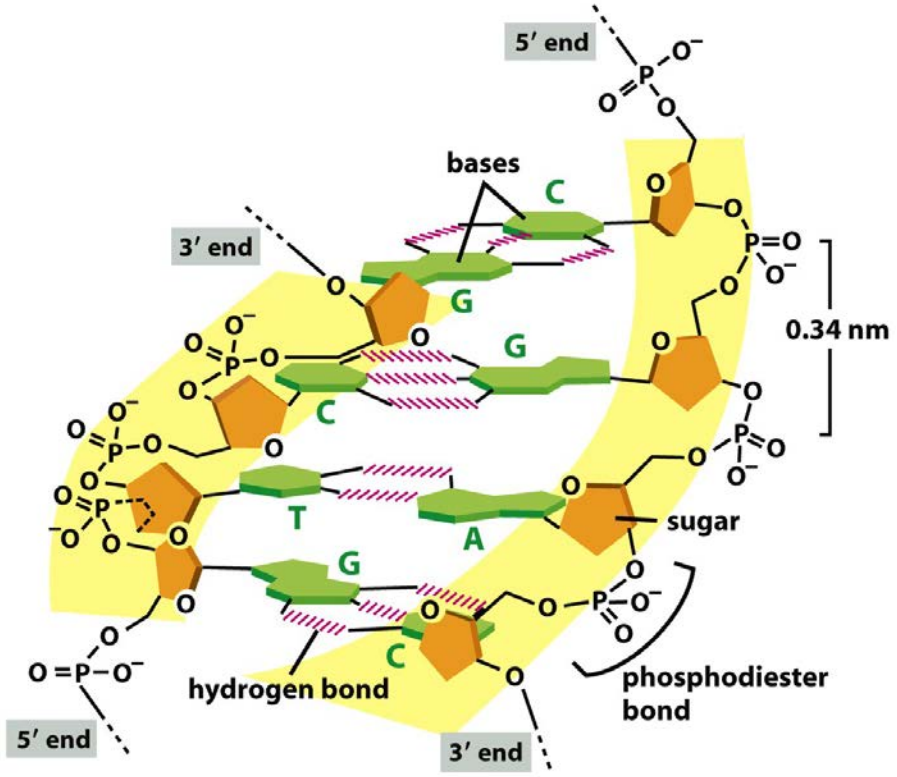
This *is* just a velcro type reaction, many low-energy linkages multiplying their effects to hold DNA together

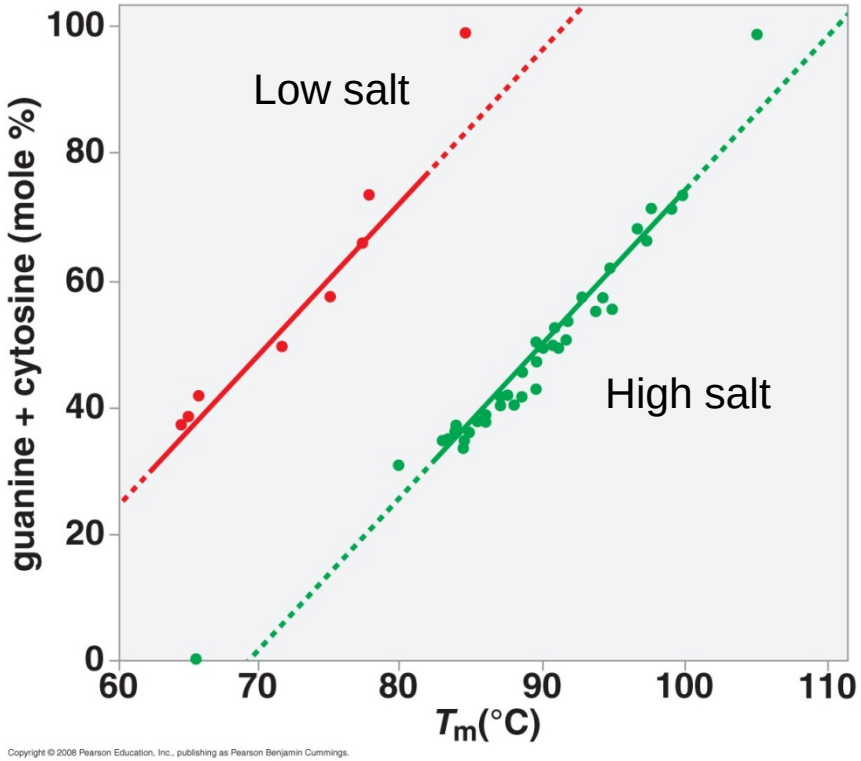
In fact, DNA *can* be melted to single strands when temperature is high

Short DNA duplexes are especially unstable

Figure 4-5b *Molecular Biology of the Cell* (© Garland Science 2008)

Electrostatic repulsion between phosphodiester backbones *always* threatens to push DNA strands apart, especially when salt concentrations are low…
High H-bond density in G+C-rich regions keeps them together





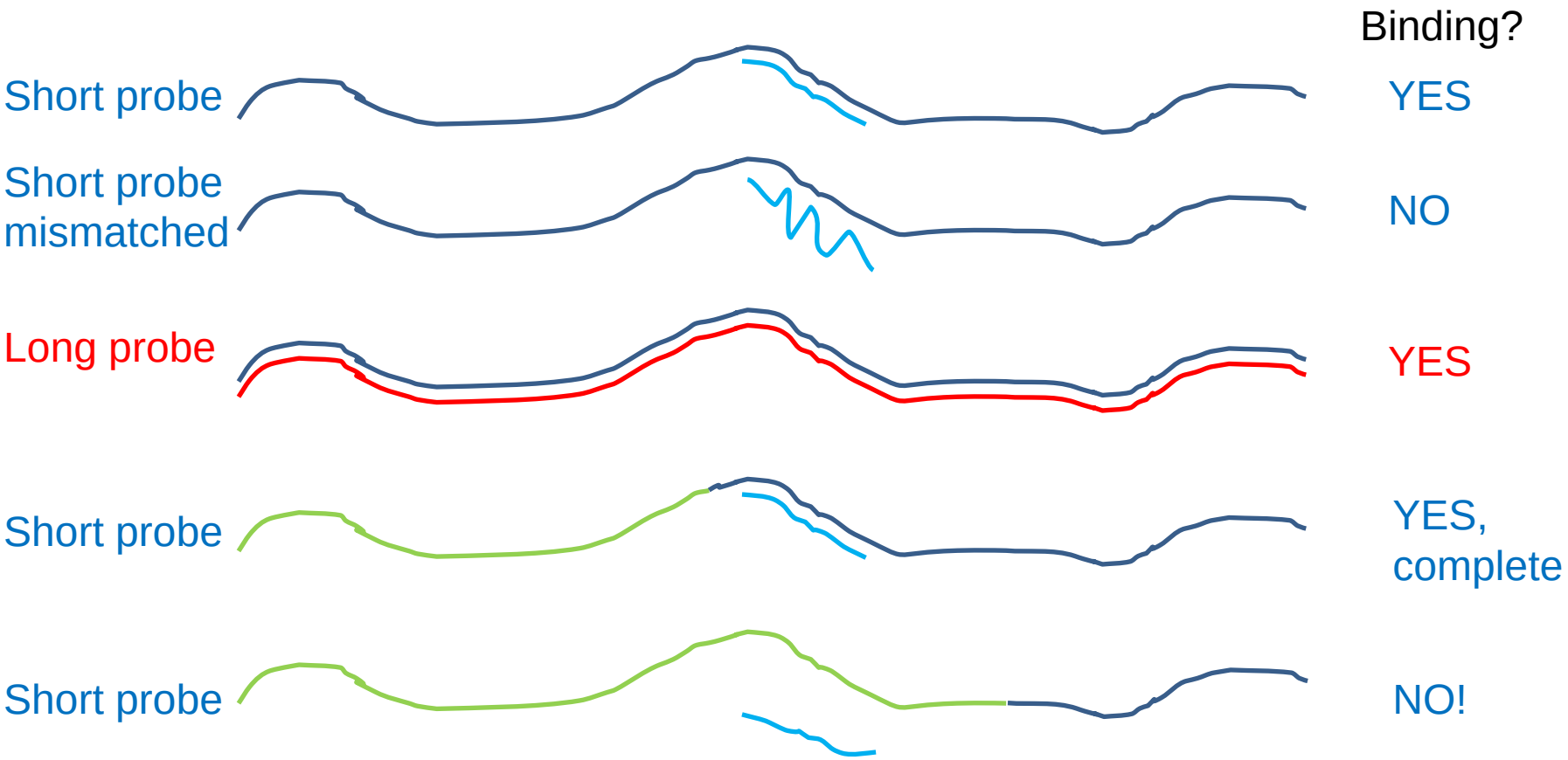
DNA duplex melts to single strands when temperature is high… but high salt can mask negative charge by providing Na+, K+, or Mg++ counterions

Figure 4-5b *Molecular Biology of the Cell* (© Garland Science 2008)

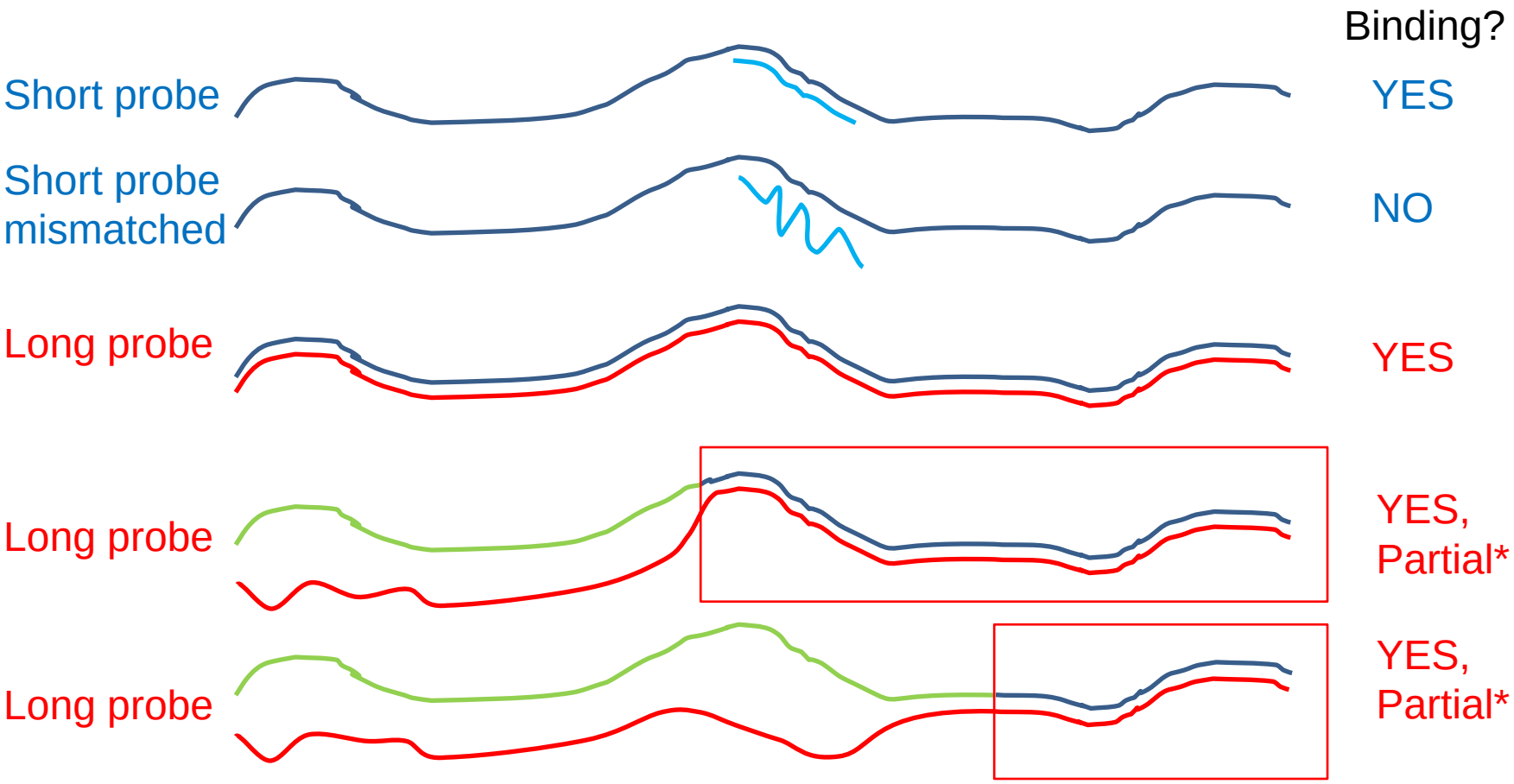
Hybridization and melting in vitro are crucial tools of molecular biology

- Thermal stability of duplex is directly related to perfection of base pair matching
  - approx. 1° C lowering of melting temperature for every % mismatch
  - Longer duplexes hybridize faster and are more stable than short ones because of zippering effect
  - hybridization at reduced temperature, high salt can react probe with target up to ~25% mismatched
  - hybridization at high temperature, low salt (high electrostatic repulsion!) can be highly sequence-specific
- Hybridization of probes is basis of blotting techniques for detection
- Hybridization of primer sequences is basis for targeted DNA replication including PCR

Hybridization: probe binding specifically to its complementary sequence can recognize structurally different DNAs depending on probe relation to target



Hybridization: probe binding specifically to its complementary sequence can recognize structurally different DNAs depending on probe relation to target

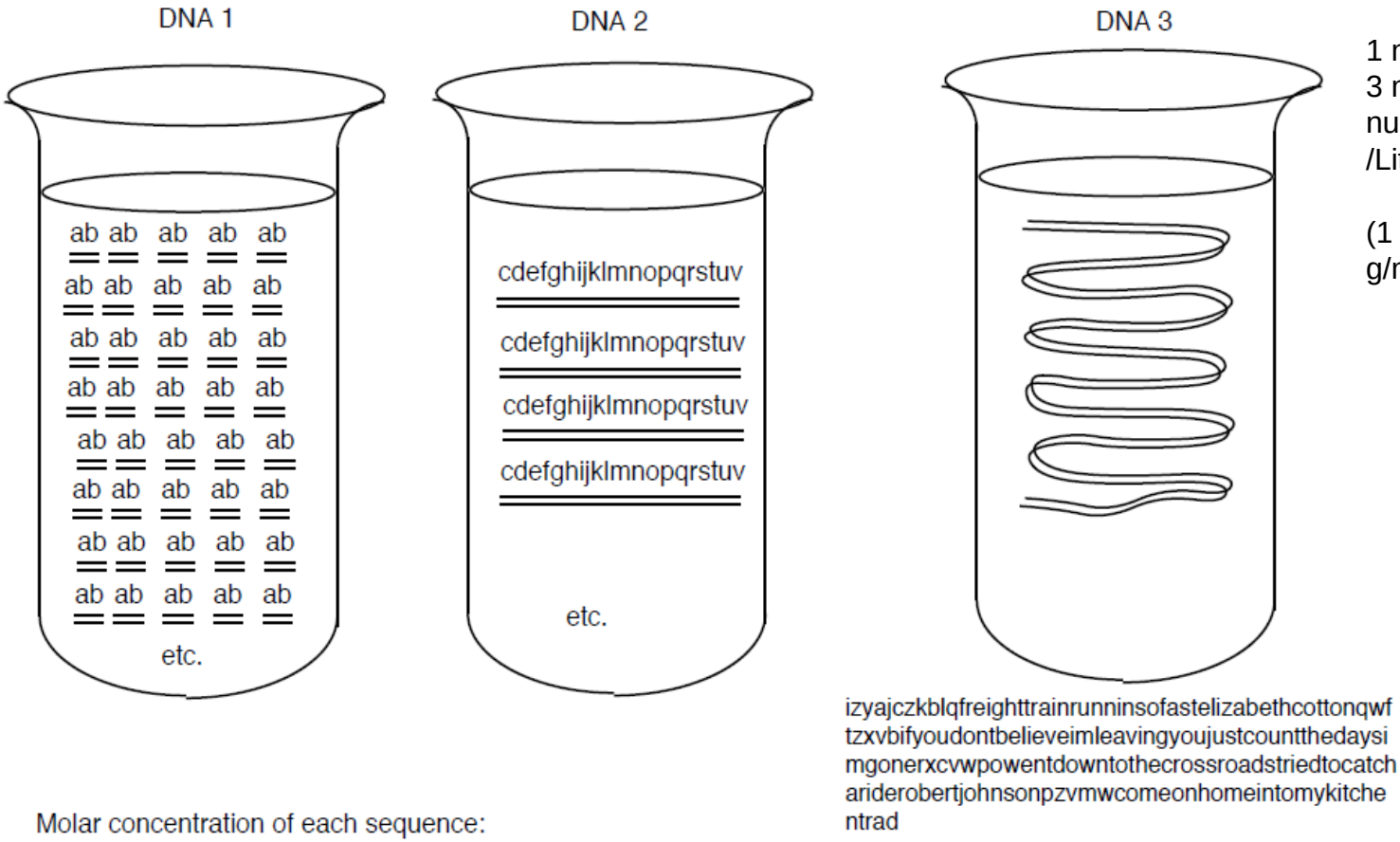


Hybridization and melting in vitro are crucial tools of molecular biology

- *Kinetics* of hybridization depend on “needle in a haystack” principle
  - Probe gets incorporated faster into hybrid when there are a lot of complementary strands around to collide with
- Hybridization kinetics can be used to compare *concentrations of A SPECIFIC SEQUENCE* within different heterogeneous DNA or RNA samples
  - Converts ***CHEMISTRY* to *INFORMATION***
  - Converts physical substance “x number of micrograms of DNA” into a frequency of occurrence of a specific “word” – “x number of copies of this sequence per microgram”

Solutions of the same "amount" of DNA, 1 mg/ml, can contain radically different concentrations of a particular 20-bp sequence depending on the sequence complexity of the DNA

each letter is a specific sequence of 10 nt



1 mg/ml DNA = 3 mM total nucleotides /Liter

(1 nt ≈ 330 g/mole)

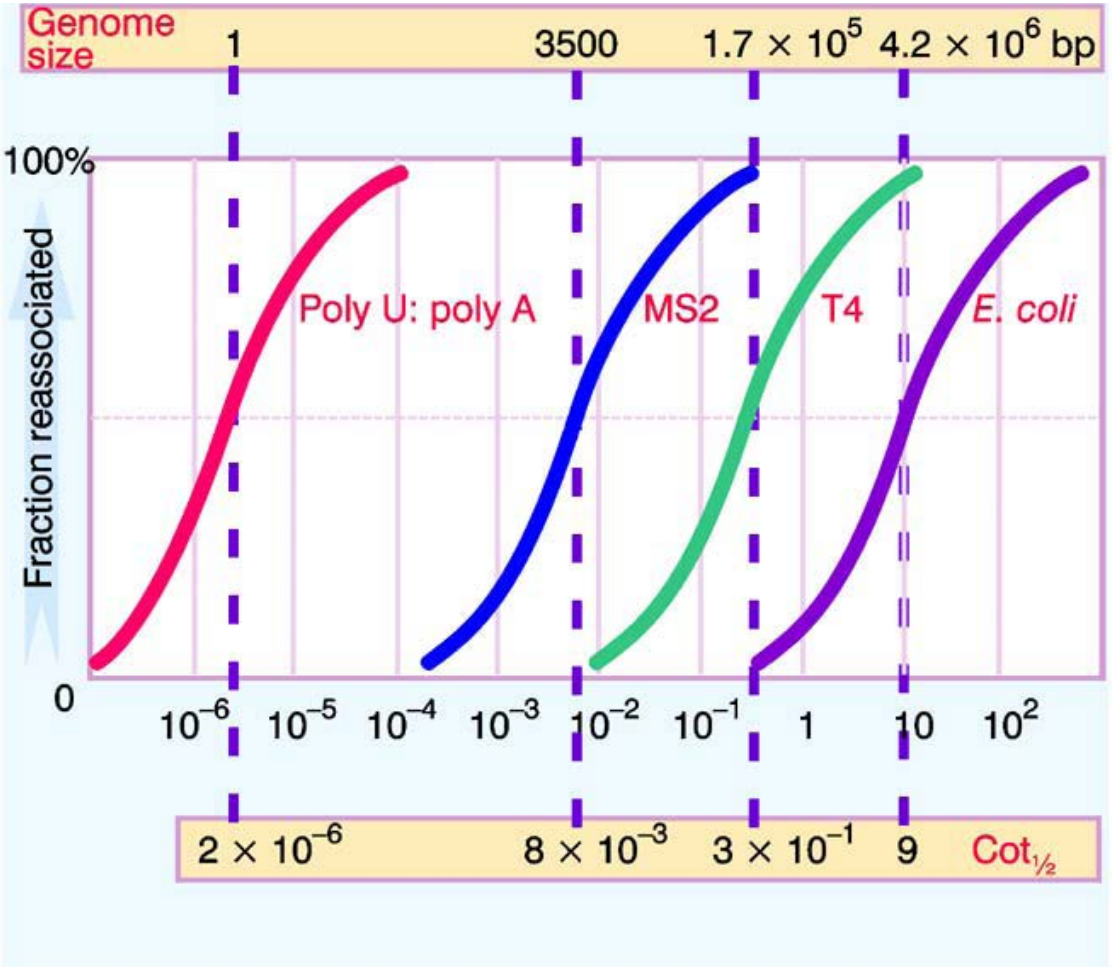
Speed with which a given 20-nt probe finds a complement in solution depends on the ***sequence*** concentration, not just total DNA bulk concentration

Hybridization rate is inversely proportional to the amount of total physical DNA a probe needs to trawl through to find its specific complementary *sequence* in a given physical amount of DNA

C_0t (moles x sec/liter) measures physical mass of DNA allowed to hybridize for a given time – *rate of annealing* tells you if the sequences are all the same or different

$C_0t_{½}$ for a probe seeking a single copy sequence in a bacterial genome: **~10**

For seeking a single copy sequence in a mammalian genome, $C_0t_{½}$ is **~10^4**



(Lewin, Genes VII via Michal Linial)

Hybridization and melting in vitro are crucial tools of molecular biology

- Usefulness
  - Purification: detecting a specific DNA or cDNA or RNA sequence of interest in a particular clone or fraction
  - Purification: using a probe coupled to a solid support to physically pull out all DNA molecules that can hybridize to it
  - Analysis: probing the structure of DNA or cDNA or RNA sequence of interest using different probes
  - In vitro copying of DNA: Basis for all preparative synthesis of DNA . (PCR depends on temperature control of hybridization AND melting of highly specific primers)
  - Quantitative analysis of regulation: measurement of gene expression levels in particular cells or tissues
- Insight: probing genomic organization and evolution

Approx. same number, 25,000, of protein-coding genes are “diluted” by >10x more other DNA in humans vs. Drosophila

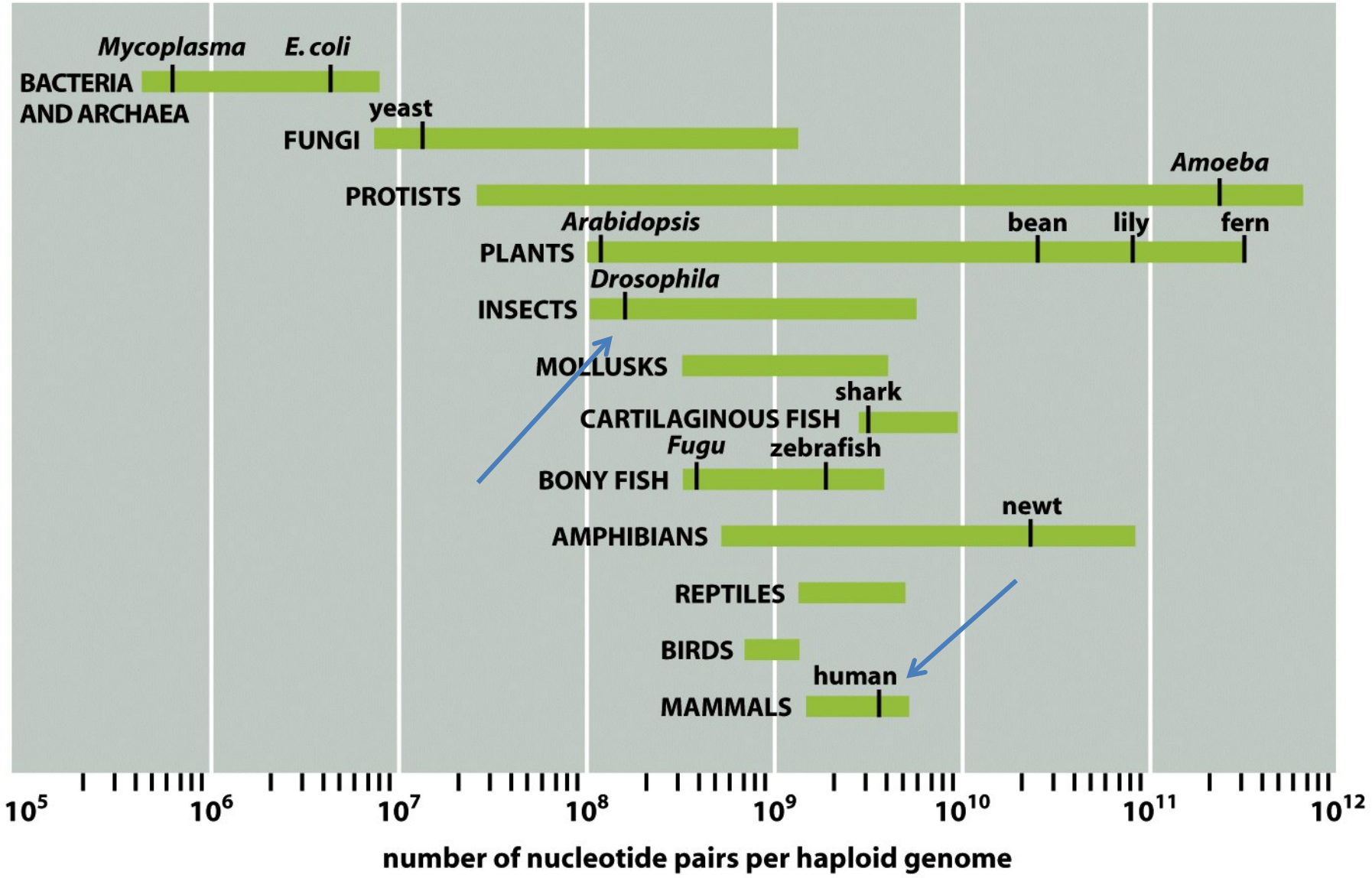
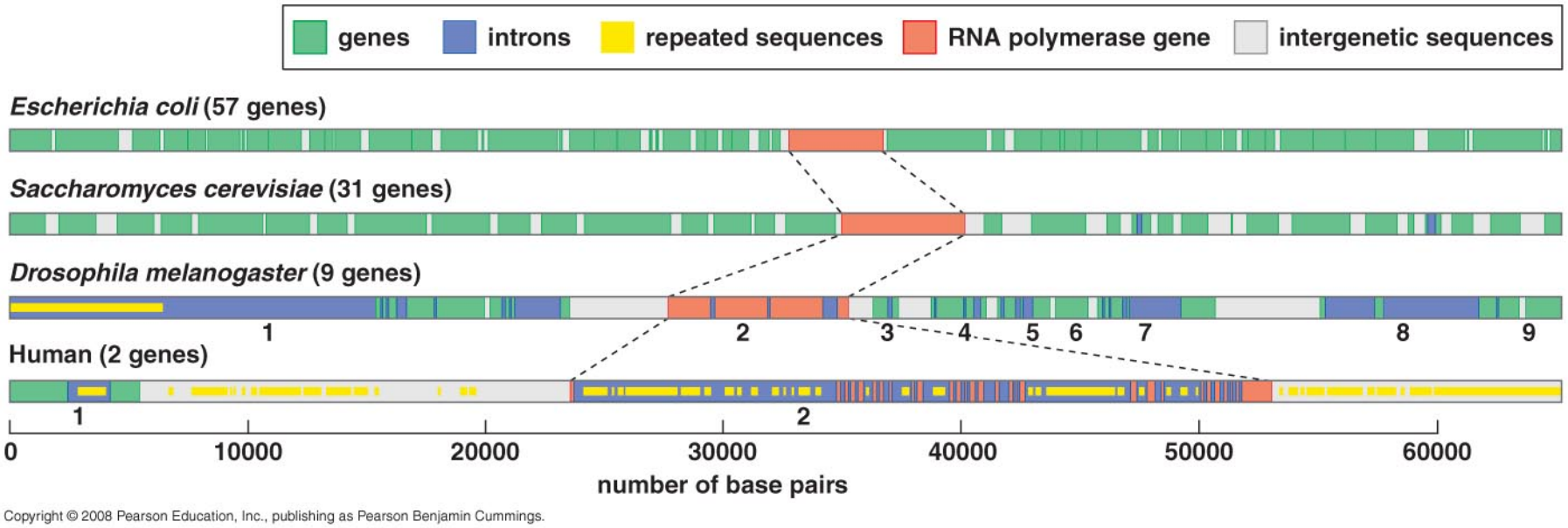


Figure 1-37 *Molecular Biology of the Cell*, Fifth Edition (© Garland Science 2008)

DNA can be copied, miscopied, broken, and rejoined

- Our genomes are normally copied perfectly every time one of our cells divides: A-T, G-C complementarity rules
- Mistakes can, however, occur
  - Putting in the wrong complementary base during copying of DNA
  - Damage to a base making it look like another base when time comes to pick a complement
  - Making an extra copy…
  - Breaking DNA
- Repairs come from rejoining DNA, sometimes to the wrong place
  - INDELS (insertions/deletions) can occur
  - Transpositions: one swath of DNA reconnected to another
- Some DNA sequences "go rogue", get copied many times and inserted around genome → **"repeat sequences"**

A major component of mammalian genomes is highly or moderately repeated – mistakes and “selfish” DNA



These repeat sequences were first discovered by DNA self-hybridization kinetics – if you melt DNA from humans and other mammals, then let it reanneal, about 40% of the total DNA re-hybridizes 10-1000x *too fast* to be “looking” for a single complement in the genome

A useful resource for more information on nucleic acid hybridization

Working with Molecular Genetics
by Ross C. Hardison, 2005
Professor of Biochemistry, The Pennsylvania State University, University Park, PA 16802

Chapter 4

http://www.personal.psu.edu/rch8/workmg/workmolecgenethome.html