BE 183 Applied Genomic Technologies

Lecture 1

(a) Introduction

(b) Molecular Biology review

(c) Genome size and the C-value paradox

Trey Ideker

Departments of Bioengineering and Medicine
University of California San Diego

Course Staff

Trey Ideker
Professor of Medicine & Bioengineering
tideker@ucsd.edu
Office: Skaggs Pharmacy Bldg 4244
Office hours: Tuesdays 2pm-3pm

Aarash Bordbar
PhD Candidate, Bioengineering
aabordba@gmail.com
Office: Bioengineering Hall Room 417
Office hours: Mondays 10am-11am

Course Website:
http://chianti.ucsd.edu/BENG183-2010/

Bioengineering 183

- Human genetics research has seen explosive growth thanks to major advances in genomic technologies over the past 5-10 years.
- We will cover exactly what these major technologies are and how they are built.
- This includes biological, chemical, genetic, mechanical, electrical, and computational aspects.
- We expect you to know basic human genetics and a little math.

Course Structure

- Lectures:
  - Center Hall Room 223, 3:30pm Tuesdays & Thursdays
  - 1.5 hours per lecture, 10 weeks x 2 lectures per week
- Problem Sets: One per week, excepting midterm week and final project weeks.
- Working hands-on with genotype data will make learning human genetics more engaging and exciting. Like the difference between taking world history versus spending a year abroad.
- Attendance and participation very important!

Course Grading

- 30% Problem Sets
- 30% Midterm
- 30% Final Project
- 10% Attendance

- Graded on usual A-F scale
- Copying your friends' problem sets not allowed

- Need access to web from laptop or desktop computer

Student Group Projects

- Form groups of ~2 students each (1-3 ok).
- Your group will write and orally present an NIH Grant Proposal.
- Proposal is to develop new genomic instrumentation or their novel use.
- Project grade is based on both oral and written portions. Same grade for each student in a group.
- More info and group sign-ups after the midterm.

Personalized Genomics

- Companies such as 23andme and Navigenics will sequence your own genome at ~500,000 single nucleotide polymorphic positions (SNPs).
- This personalized genomic information may soon revolutionize health care.
- We will study the technology that enables this personalized information to be generated.
- We will also log in to the 23andme website to examine existing genomes uploaded for education.
- The instructors have no ties, fiscal or otherwise, to these companies.

Homework

- For next class, read:

  Eriksson et al. Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits, *PLoS Genetics* **6:6** (2010).

Molecular Biology Review

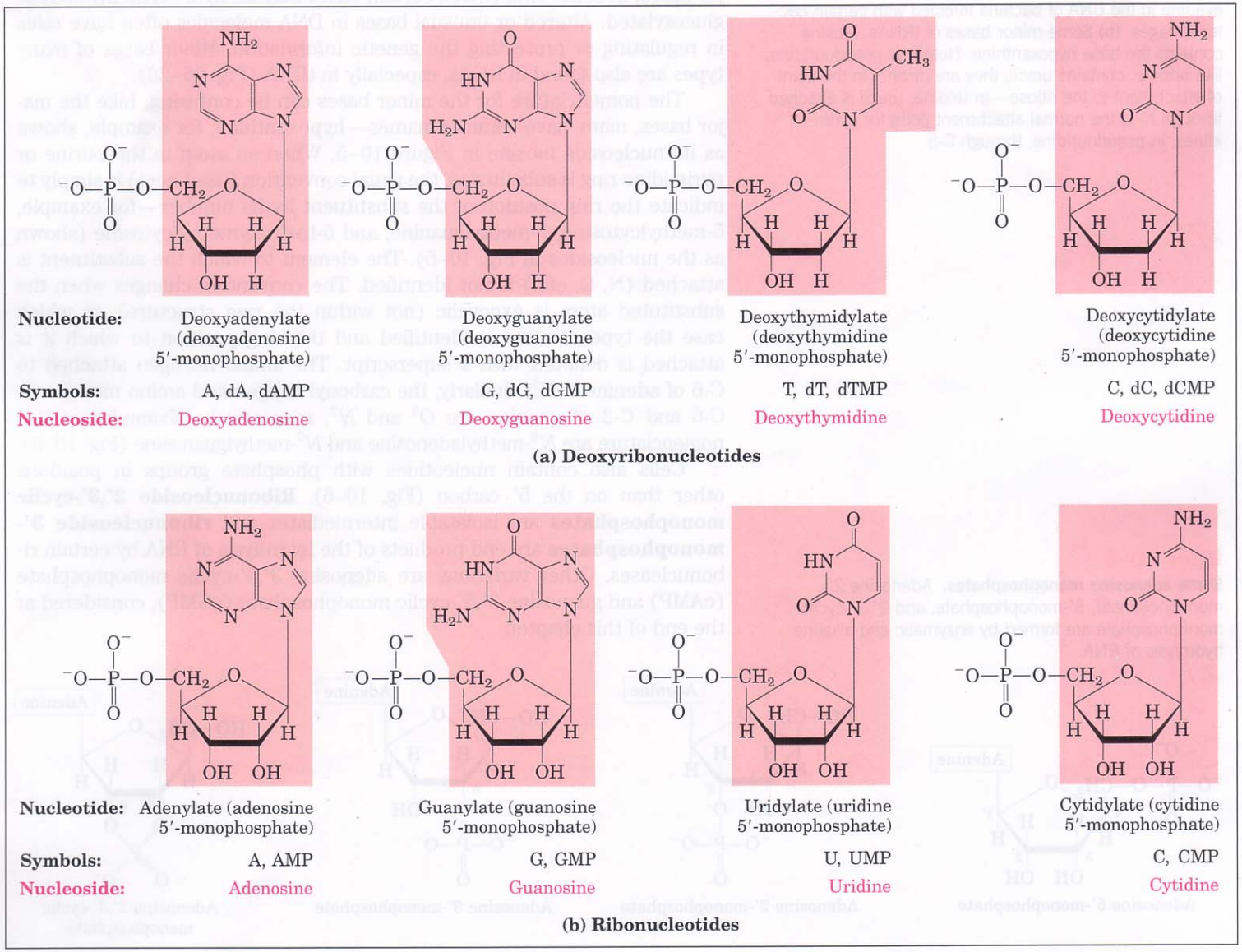
- Structure of nucleic and amino acids
- DNA packaging and locations in cell
- DNA replication
- Transcription and translation
- Hybridization

Who can draw the structure of DNA?

How does RNA differ?

10

Chemical Structure of Nucleic Acids



	Deoxyadenylate (deoxyadenosine 5′-monophosphate)	Deoxyguanylate (deoxyguanosine 5′-monophosphate)	Deoxythymidylate (deoxythymidine 5′-monophosphate)	Deoxycytidylate (deoxycytidine 5′-monophosphate)
Symbols:	A, dA, dAMP	G, dG, dGMP	T, dT, dTMP	C, dC, dCMP
Nucleoside:	Deoxyadenosine	Deoxyguanosine	Deoxythymidine	Deoxycytidine

(a) Deoxyribonucleotides

	Adenylate (adenosine 5′-monophosphate)	Guanylate (guanosine 5′-monophosphate)	Uridylate (uridine 5′-monophosphate)	Cytidylate (cytidine 5′-monophosphate)
Symbols:	A, AMP	G, GMP	U, UMP	C, CMP
Nucleoside:	Adenosine	Guanosine	Uridine	Cytidine

(b) Ribonucleotides

Nucleotides and Nucleic Acids

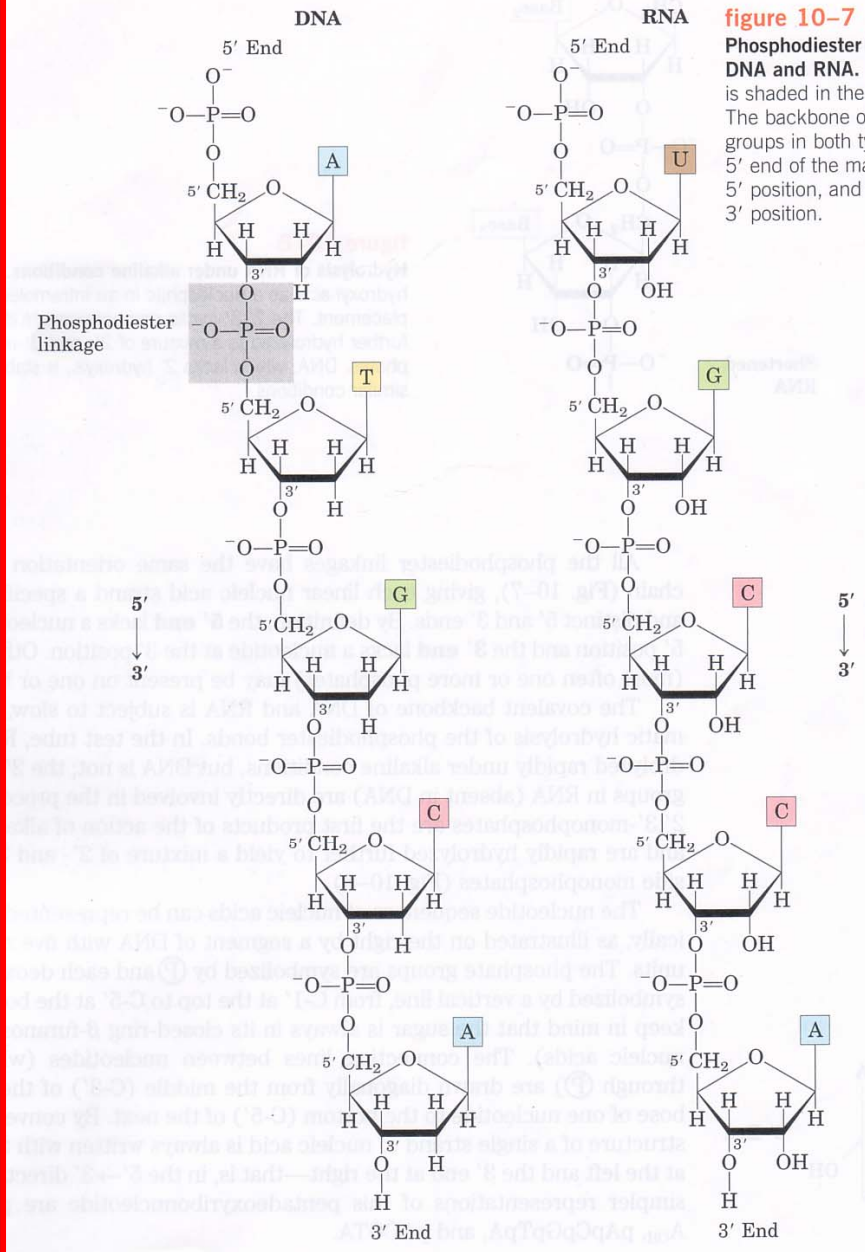
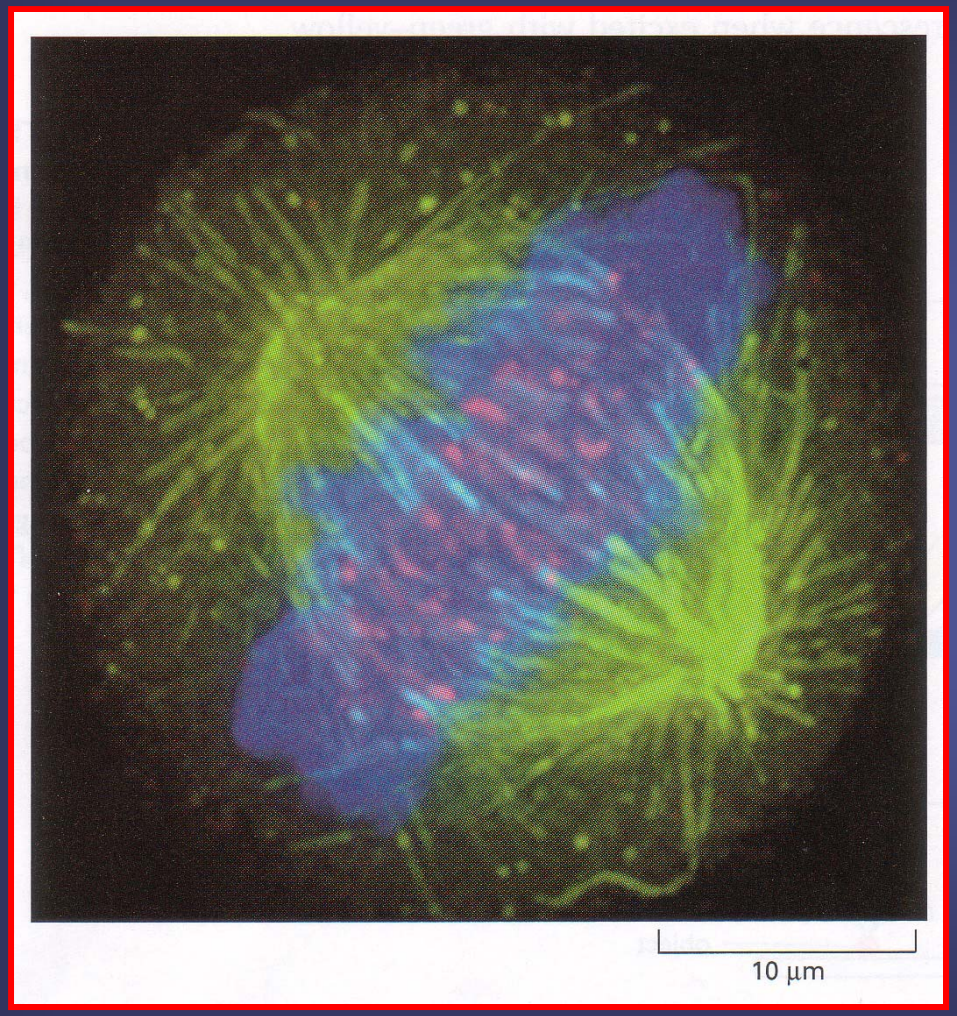


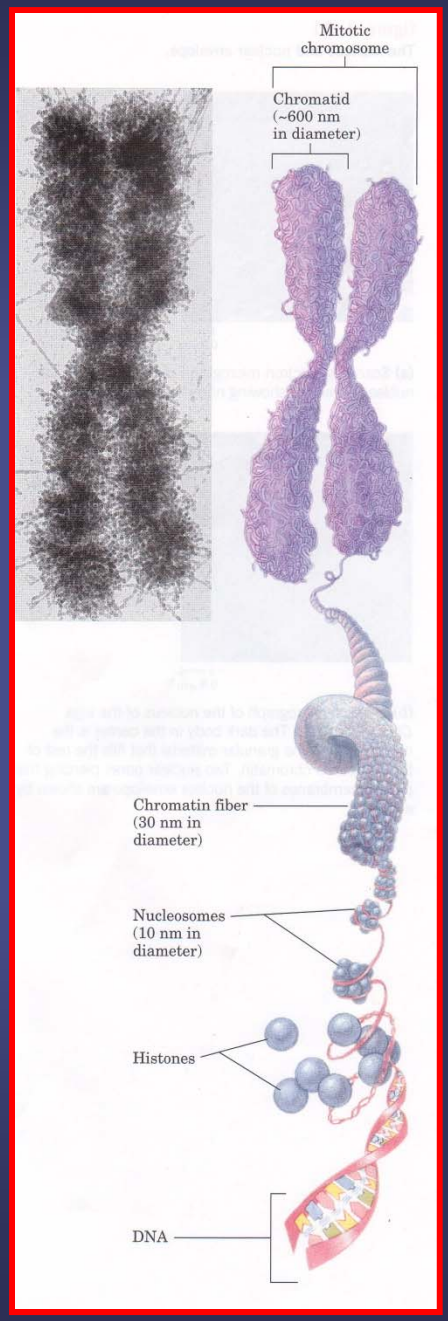
figure 10–7

Phosphodiester linkages in the covalent backbone of DNA and RNA. The phosphodiester bonds (one of which is shaded in the DNA) link successive nucleotide units. The backbone of alternating pentose and phosphate groups in both types of nucleic acid is highly polar. The 5′ end of the macromolecule lacks a nucleotide at the 5′ position, and the 3′ end lacks a nucleotide at the 3′ position.

Packaging of DNA in Chromosomes: Histones, Chromatin

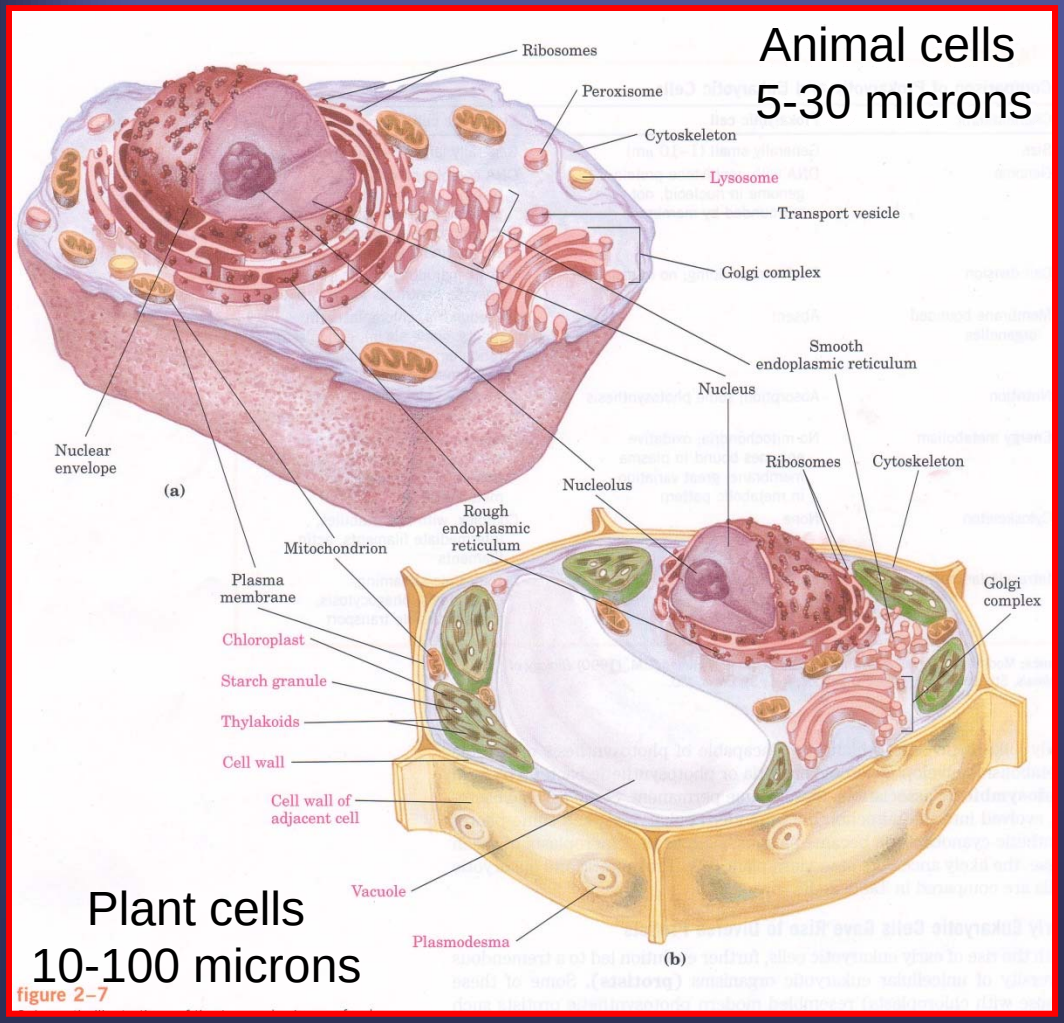


Fluorescent micrograph of cell in mitosis, green F – spindle microtubules, red F – centromeres, blue F – condensed chromosomes. Mol Bio Cell 4ed pp. 556

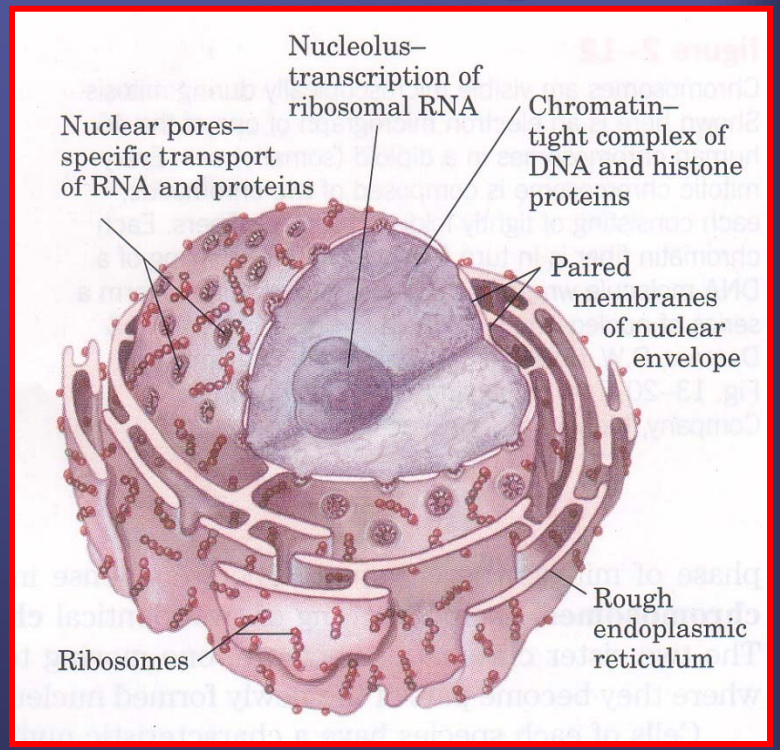


Lehninger pp. 36

Where is DNA Located in the Cells ?



Nucleus

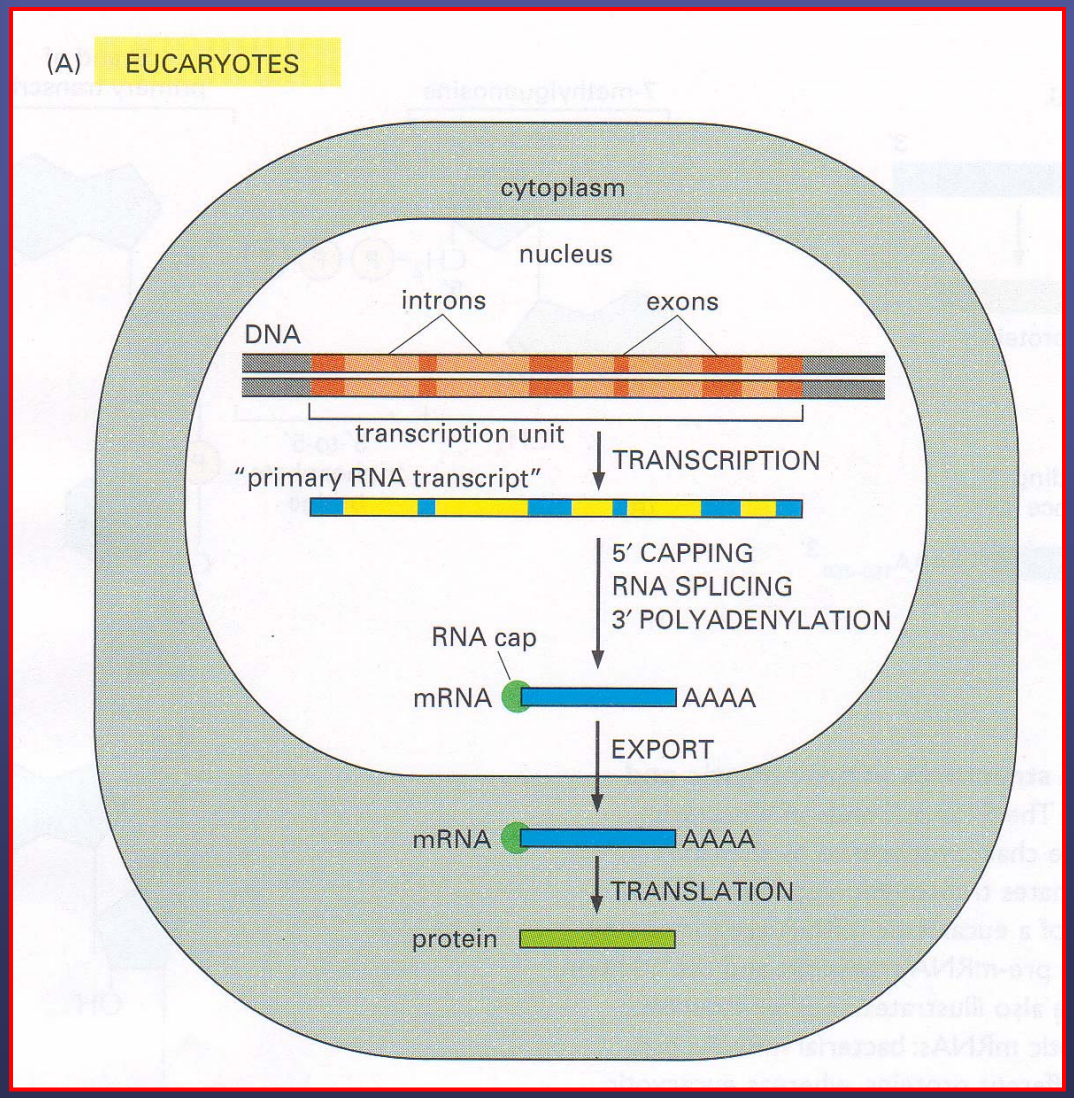


Lehninger pp. 30 & 35

What is the central dogma of molecular biology?

How is RNA processed?

Central Dogma of Biology
"DNA to RNA to Protein"



Mol Bio Cell 4ed pp. 315

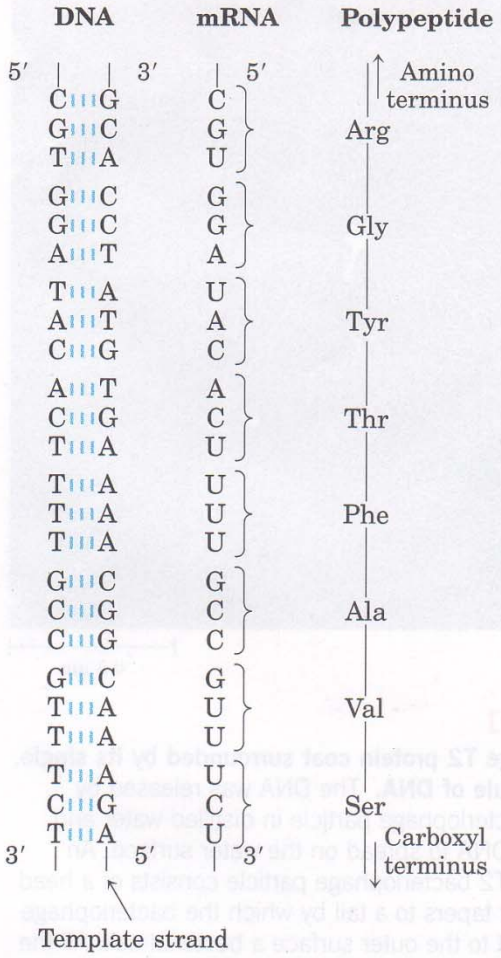
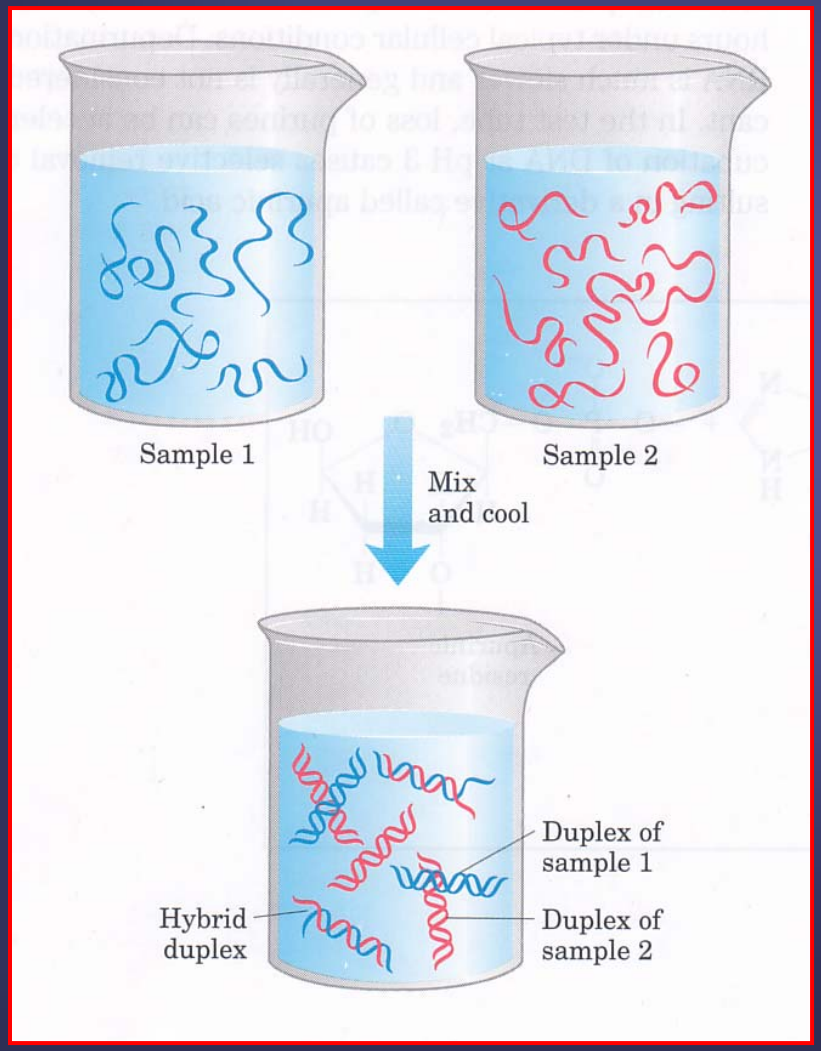


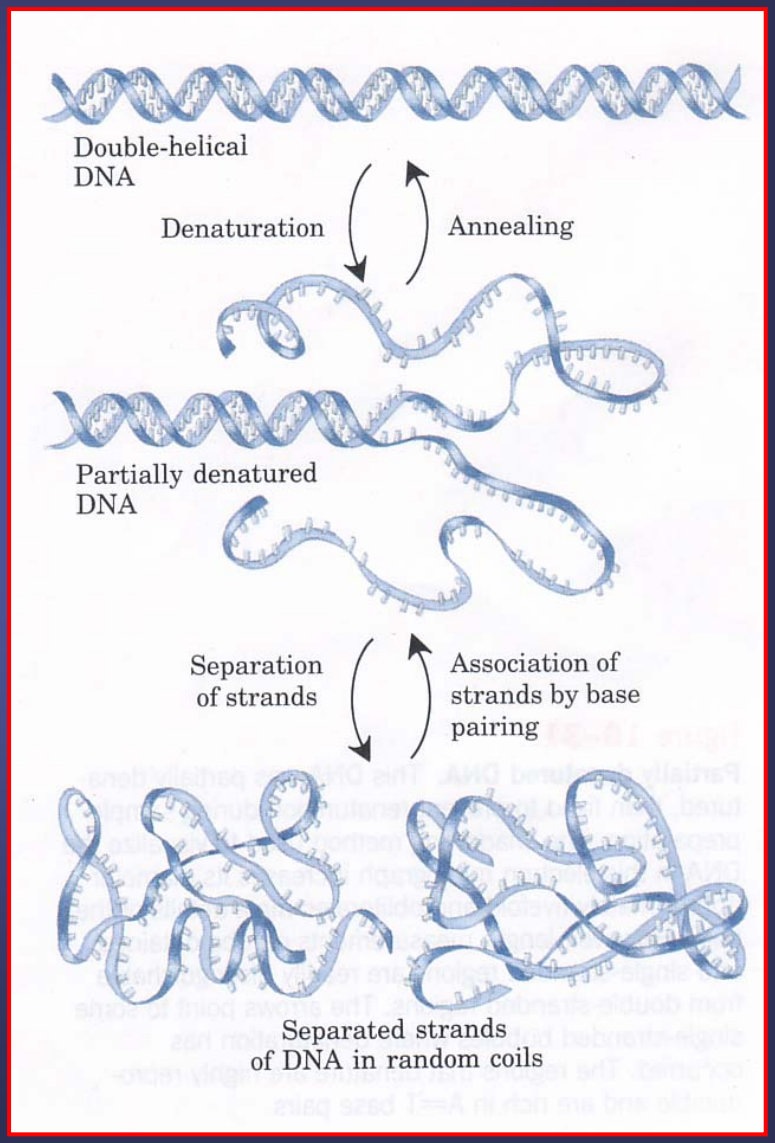
figure 24–2

Colinearity of the coding nucleotide sequences of DNA and mRNA and the amino acid sequence of a polypeptide chain. The triplets of nucleotide units in DNA determine the amino acids in a protein through the intermediary mRNA. One of the DNA strands serves as a template for synthesis of mRNA, which has nucleotide triplets (codons) complementary to those of the DNA. In some bacterial and many eukaryotic genes, coding sequences are interrupted at intervals by regions of noncoding sequences (called introns).

Lehninger pp. 908

The most basic building block of DNA technology: DNA Hybridization, Denaturation and Annealing





DNA Hybridization Scheme

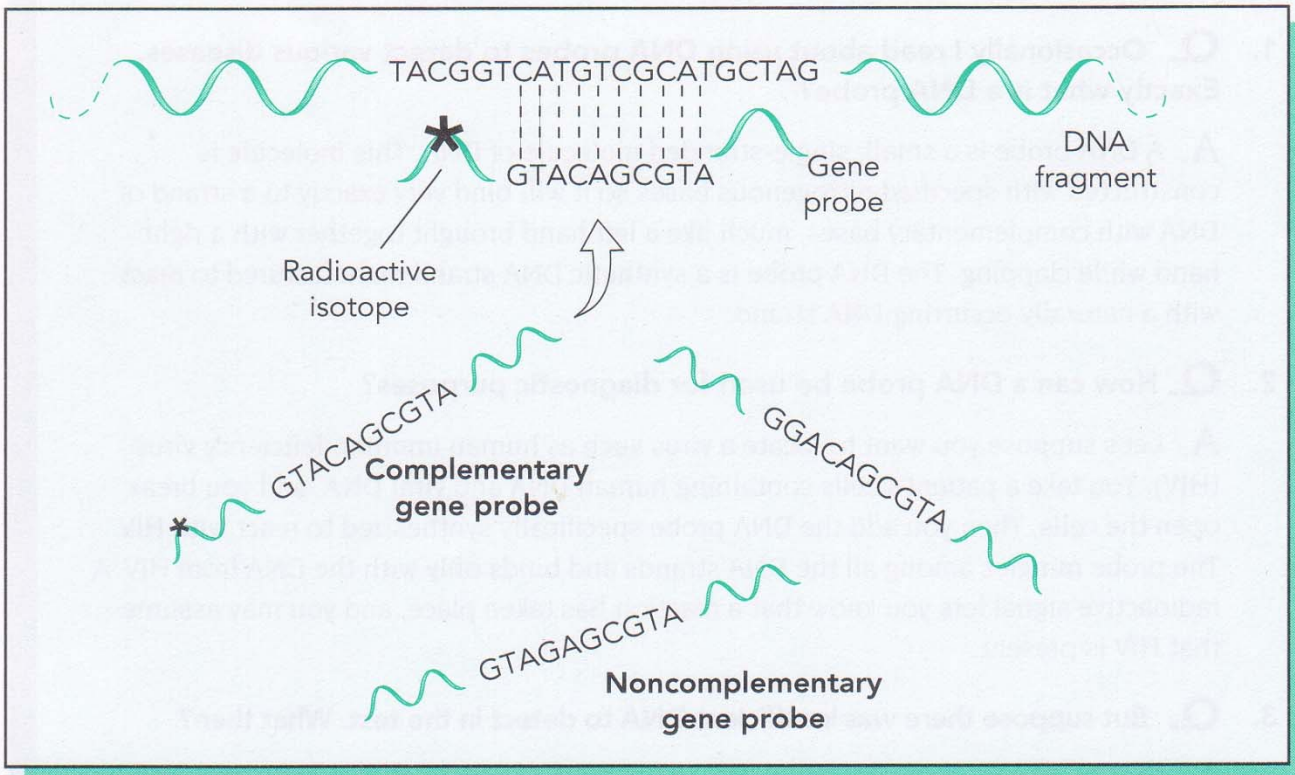
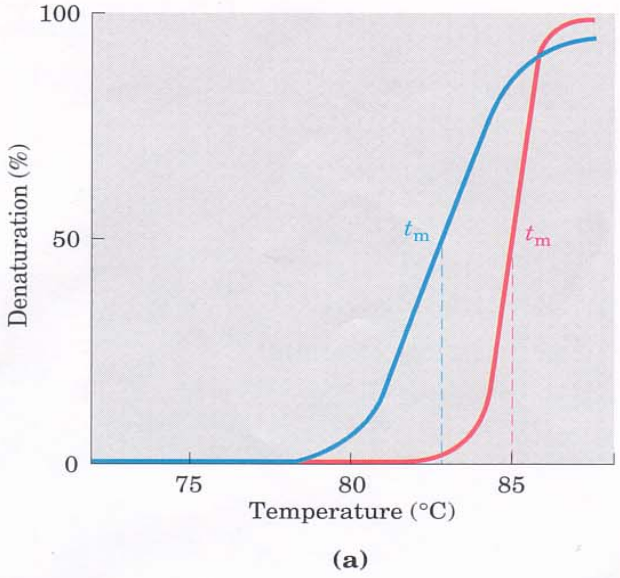


FIGURE 7.1

How a gene probe works. A gene probe is a single-stranded segment of DNA. When combined with a DNA molecule containing a complementary site, the gene probe seeks out the site and binds with it. If a radioactive molecule or atom is attached to the probe, the radioactivity accumulates at the binding site and signals that a reaction has taken place. Note in the diagram how the bases of only one probe complement the bases of the DNA fragment.

Hybridization - Heat denaturation, melting temperature (tm), other factors



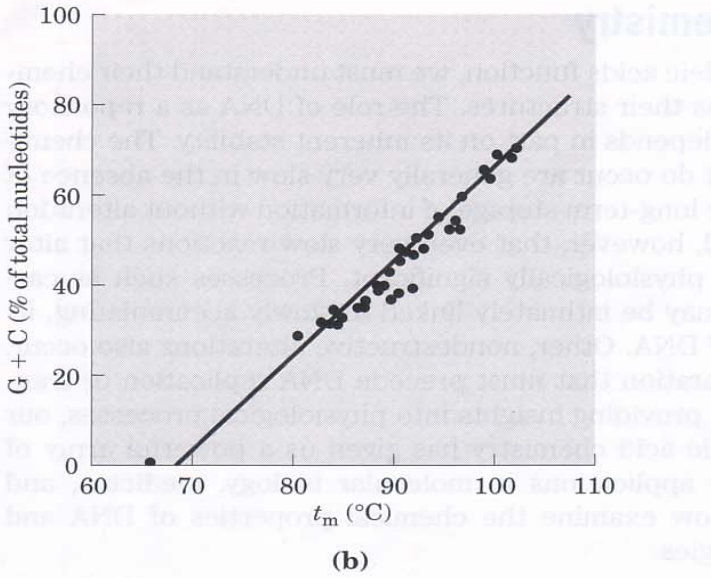


figure 10–30

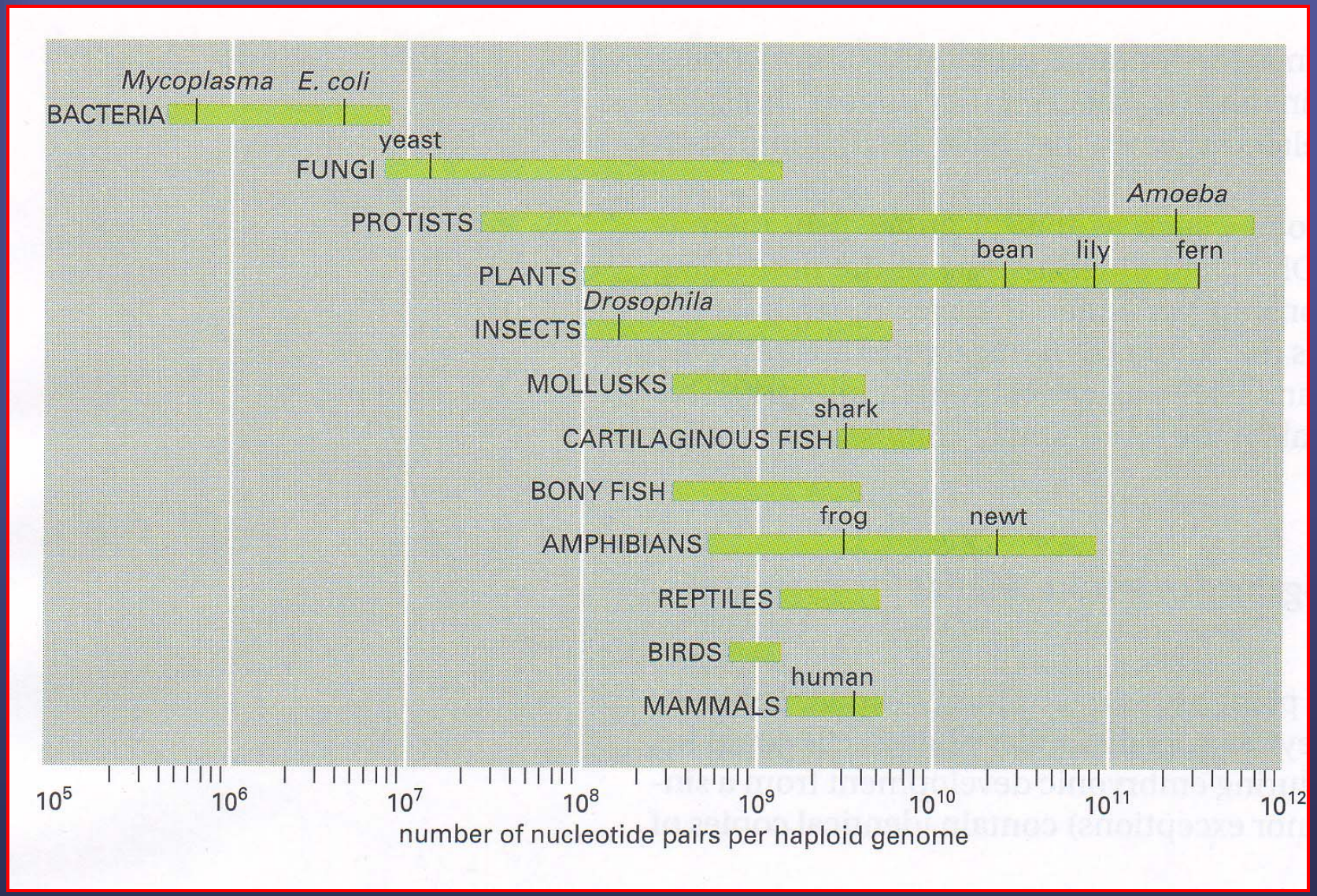
Heat denaturation of DNA. **(a)** The denaturation or melting curves of two DNA specimens. The temperature at the midpoint of the transition (t_m) is the melting point; it depends on pH and ionic strength and on the size and base composition of the DNA. **(b)** Relationship between t_m and the G≡C content of a DNA.

Temperature, pH, size (# bp's),
G/C to A/T ratio, ionic strength,
chem denaturants, detergents, chaotropics
Stringency (high and low)

Genome Organization

- Genome sizes and the C-value paradox
- $C_o t$ curves and genome complexity
- Repeated sequences
- Introns and exons
- Genome structure

Genome sizes and the *C*-value paradox



Mol Bio Cell 4ed pp. 33

Reassociation kinetics- The $C_o t$ curve

C = Concentration of ssDNA

C_0 = Initial ssDNA conc.

k = reassociation rate const.

$t_{1/2}$ = reassociation half time

Big $C_0 t_{1/2}$ = Slow reassociation

This value is proportional to the number of different types of DNA fragments

$$\frac{dC}{dt} = -kC^2$$

$$\frac{C}{C_0} = \frac{1}{1 + kC_0 t}$$

$$C_0 t_{1/2} = \frac{1}{k}$$

Comparison of sequence copy number for two organisms with different genome sizes

	Organism A	Organism B
Starting DNA concentration	10 pg/ml	10 pg/ml
Genome size	0.01 pg	2 pg
# genome copies/ml	1000	5
Relative concentration	200	1

Table 2.1 Primrose and Twyman

Reassociation kinetics- The $C_o t$ curve (resolution of the *C*-value paradox)

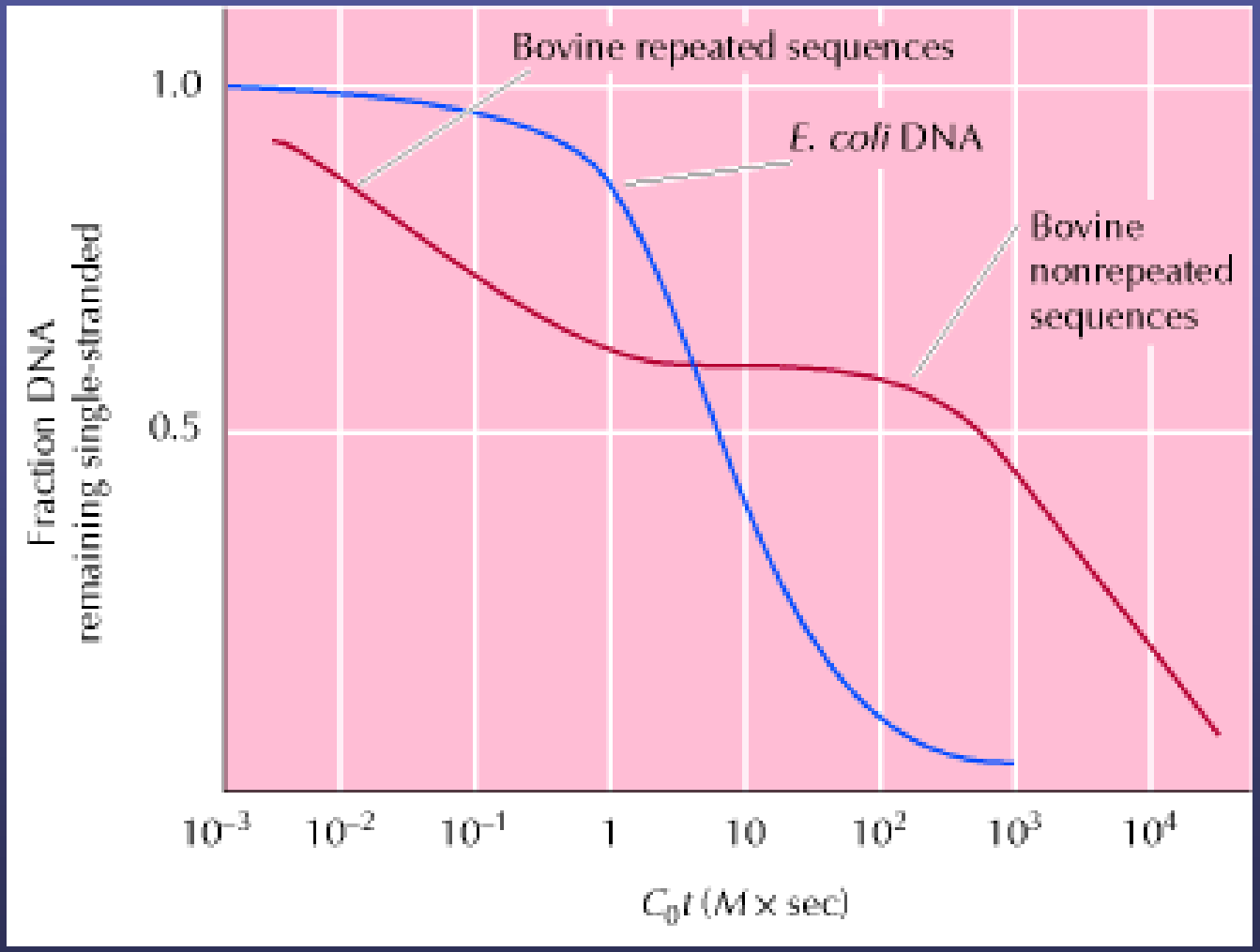


Fig. 4.6
The Cell: A Molecular Approach