Lecture 24 (11/13/23)

Nucleic Acids

A. Nucleotides
  1. parts
  2. nomenclature
  3. numbering
  4. properties

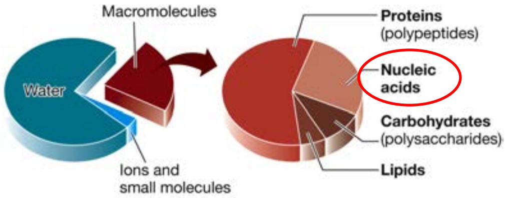
B. Nucleic Acids
  1. Polymer-bond
  2. H-bonds
  3. Roles
    a. Nucleotides
    b. Nucleic acids

C. The 4 S's
  1. Size
    a. genomes
    b. RNAs
  2. Solubility
  3. Shape
  4. Stability

- Reading: Ch8; 294-296
  263-269; 269-274
  Ch24; 885-890, 891-894
- Homework #23 & #24

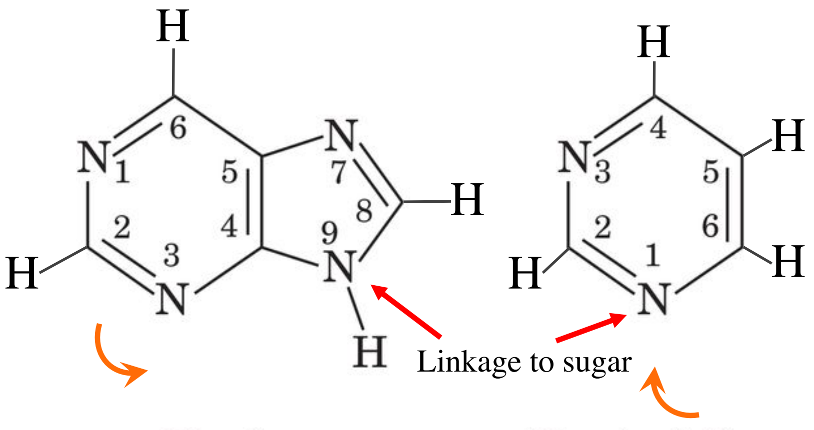
NEXT
- Reading: Ch8; 278-283
  Ch24, 895-897
- Homework #25 & #26

Nucleotides & Nucleic Acids



Definition of Nucleotides

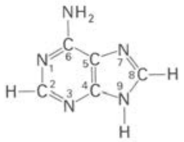
The Nitrogenous Bases of Nucleotides are Derivatives of Purine & Pyrimidine



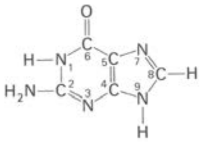
Linkage to sugar

Purine **Pyrimidine**

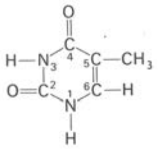
Chemical Structures of Nucleotide Bases



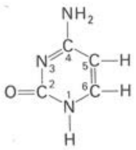
Adenine
(A)



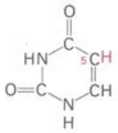
Guanine
(G)



Thymine
(T)

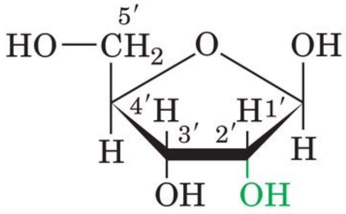


Cytosine
(C)



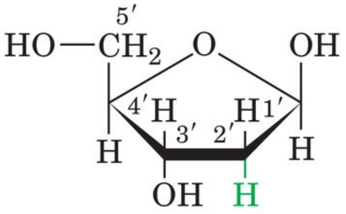
Uracil (U)

Sugars of Nucleotides are Either Ribose & Deoxyribose



Ribose

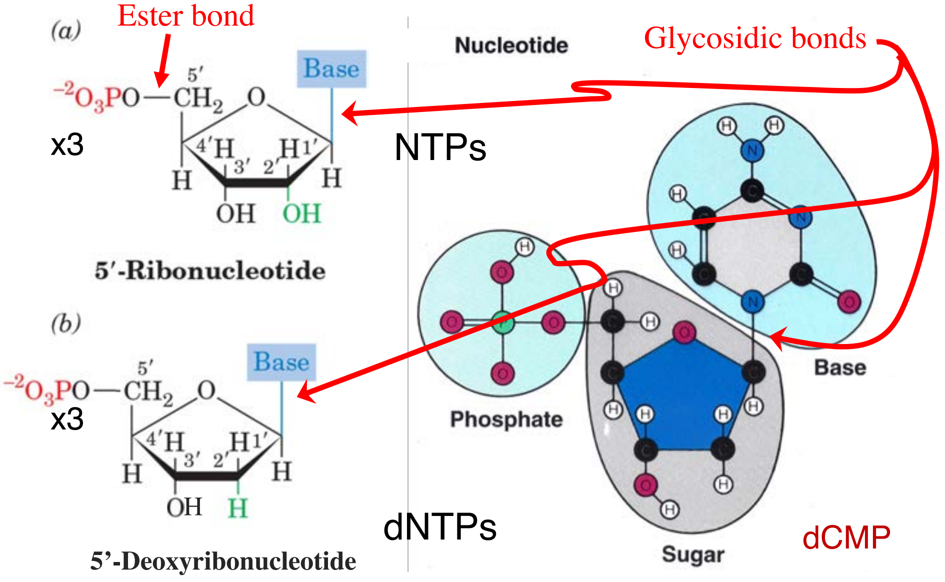
(β-D-ribose)



Deoxyribose

(β-D-2’ deoxyribose)

Chemical Structures of Nucleotides



Ester bond

x3

NTPs

x3

dNTPs

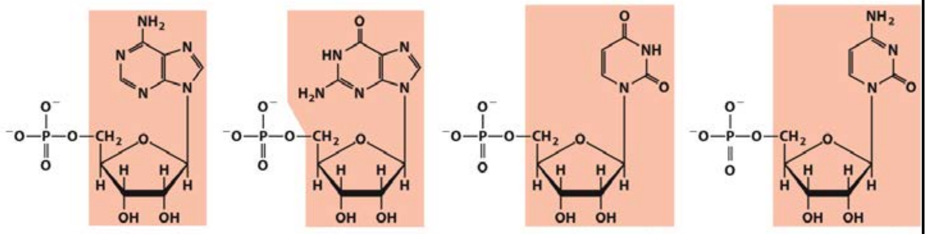
Glycosidic bonds

dCMP

Which nucleotide is this?

Nomenclature: Ribonucleotides

You need to know structures, names, numbering, and symbols (both one-letter (A) and three-letter (AMP) codes).

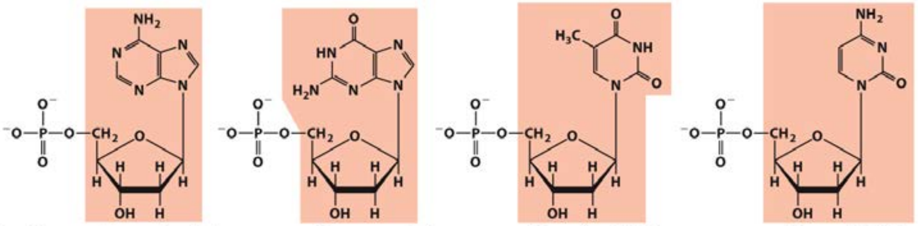


Nucleotide:	Adenylate (adenosine 5′-monophosphate)	Guanylate (guanosine 5′-monophosphate)	Uridylate (uridine 5′-monophosphate)	Cytidylate (cytidine 5′-monophosphate)
Symbols:	A, AMP	G, GMP	U, UMP	C, CMP
Nucleoside:	Adenosine	Guanosine	Uridine	Cytidine

Ribonucleotides

Nomenclature: Deoxyribonucleotides

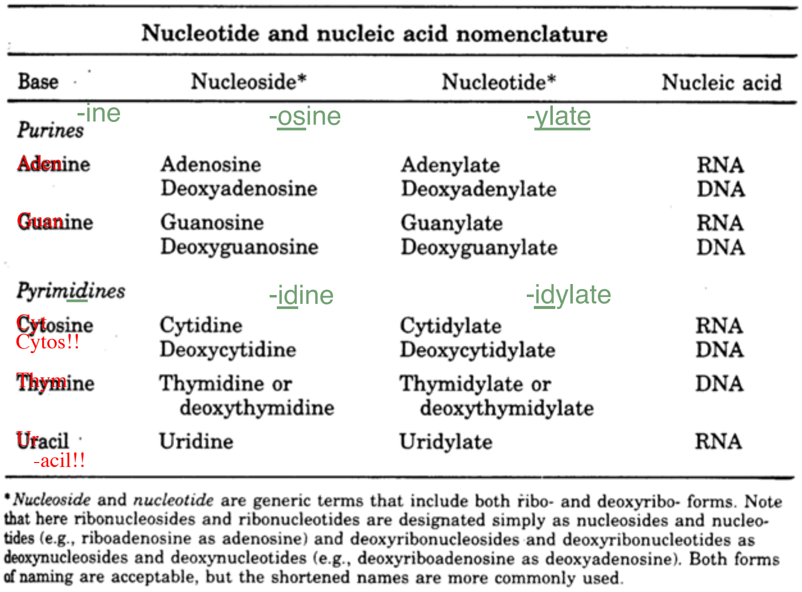
You need to know structures, names, numbering, and symbols (both one-letter (dA) and three-letter (dAMP) codes).



Nucleotide:	**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**

Deoxyribonucleotides

Nucleotides:
Nomenclature

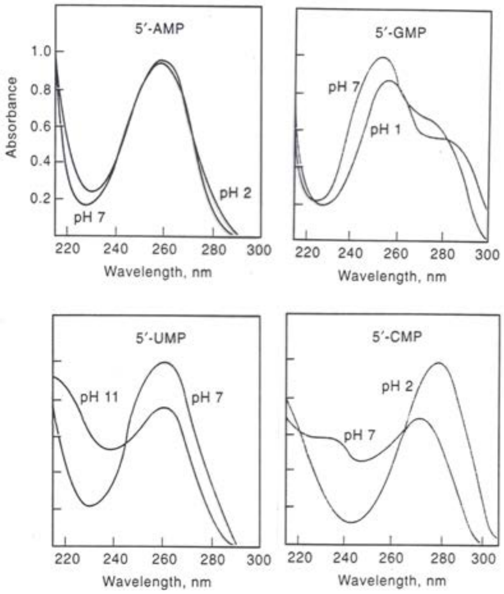


Nucleotide and nucleic acid nomenclature

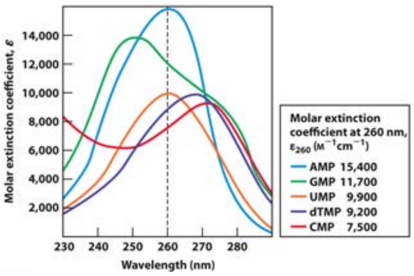
Base	Nucleoside*	Nucleotide*	Nucleic acid
Purines -ine	-osine	-ylate	
Adenine	Adenosine Deoxyadenosine	Adenylate Deoxyadenylate	RNA DNA
Guanine	Guanosine Deoxyguanosine	Guanylate Deoxyguanylate	RNA DNA
Pyrimidines	-idine	-idylate	
Cytosine Cytos!!	Cytidine Deoxycytidine	Cytidylate Deoxycytidylate	RNA DNA
Thymine	Thymidine or deoxythymidine	Thymidylate or deoxythymidylate	DNA
Uracil -acil!!	Uridine	Uridylate	RNA

**Nucleoside* and *nucleotide* are generic terms that include both ribo- and deoxyribo- forms. Note that here ribonucleosides and ribonucleotides are designated simply as nucleosides and nucleotides (e.g., riboadenosine as adenosine) and deoxyribonucleosides and deoxyribonucleotides as deoxynucleosides and deoxynucleotides (e.g., deoxyriboadenosine as deoxyadenosine). Both forms of naming are acceptable, but the shortened names are more commonly used.

Nucleic Acid Bases, Nucleosides, and Nucleotides



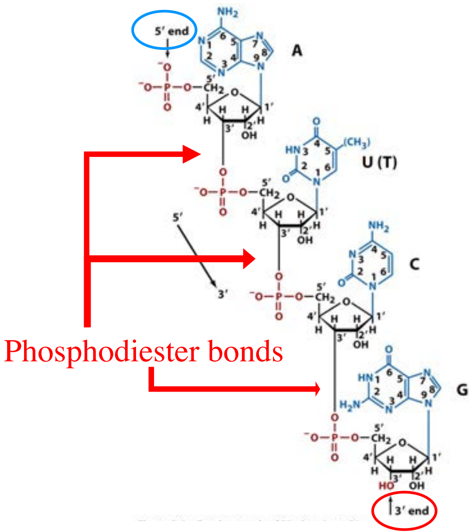
Nucleotide	Extinction Coefficient at pH 7.0 $(M^{-1}cm^{-1}) \times 10^3$	λ_{max} (nm)
AMP	15.4	260
GMP	11.7	248
CMP	7.5	280
dTMP	9.2	270
UMP	9.9	260



Definition of Nucleic Acids

Polymers of nucleotides

Nucleic acid – polymer of nucleotides – directionality 5’→3’



When you write a sequence:

ATCG

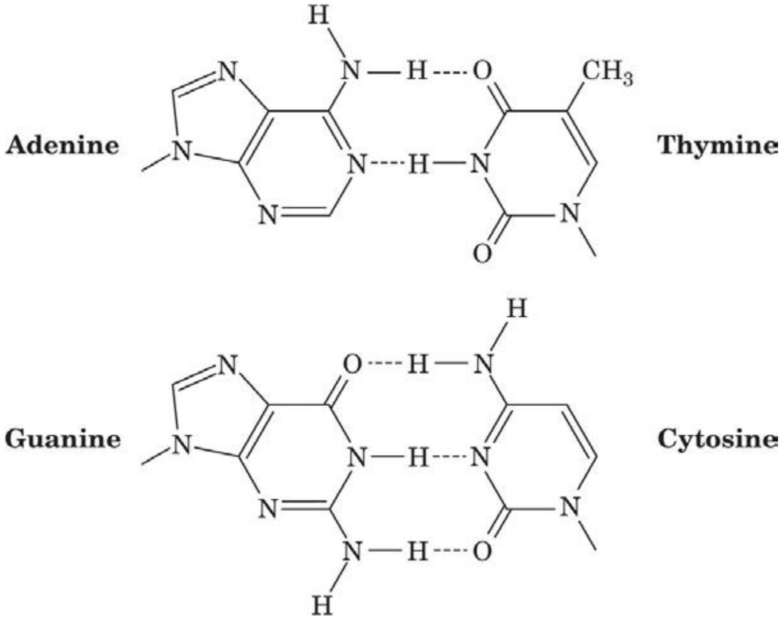
It is assumed that the 5’-end is on the left and the 3’-end is on the right, unless otherwise labeled.

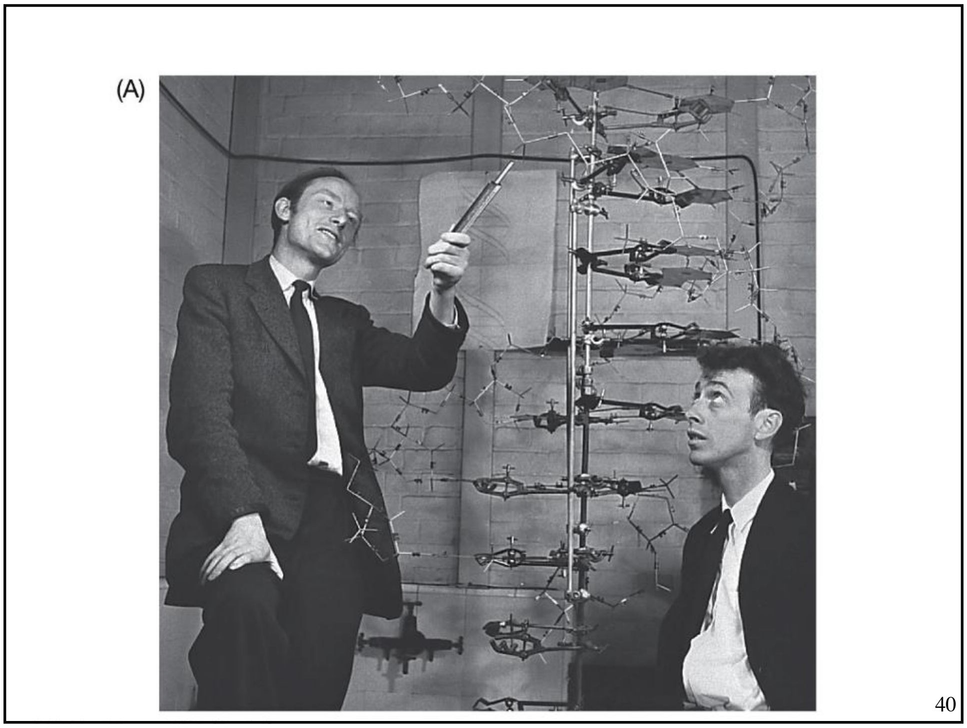
5’-ATCG-3’

3’-GCTA-5’ same molecule

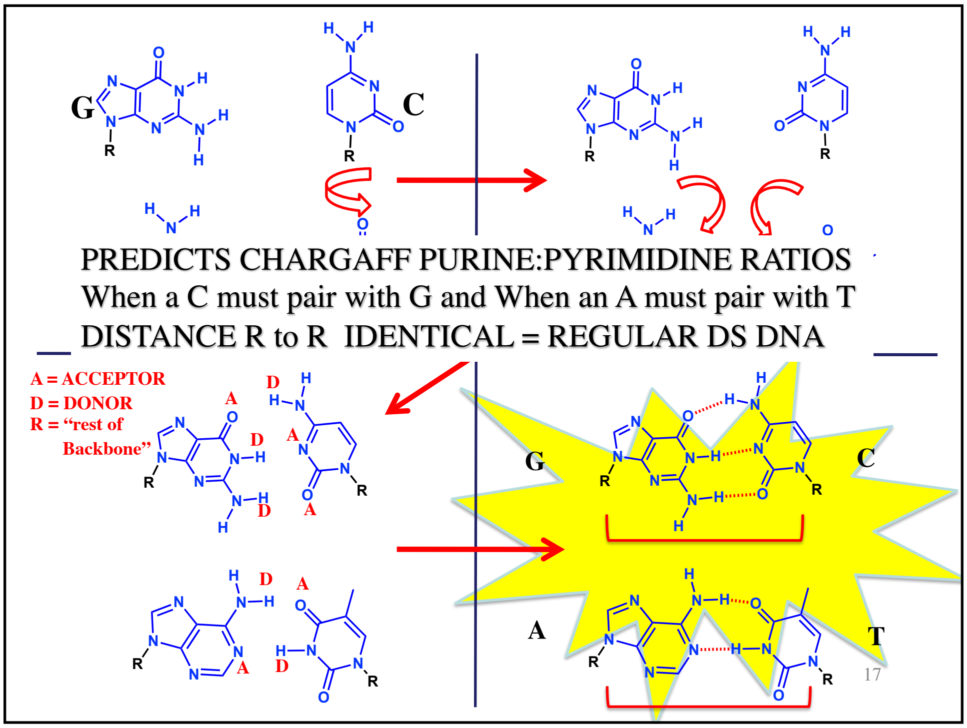
What is the bond that links the nucleotides?

Complementary Base Pairing in DNA





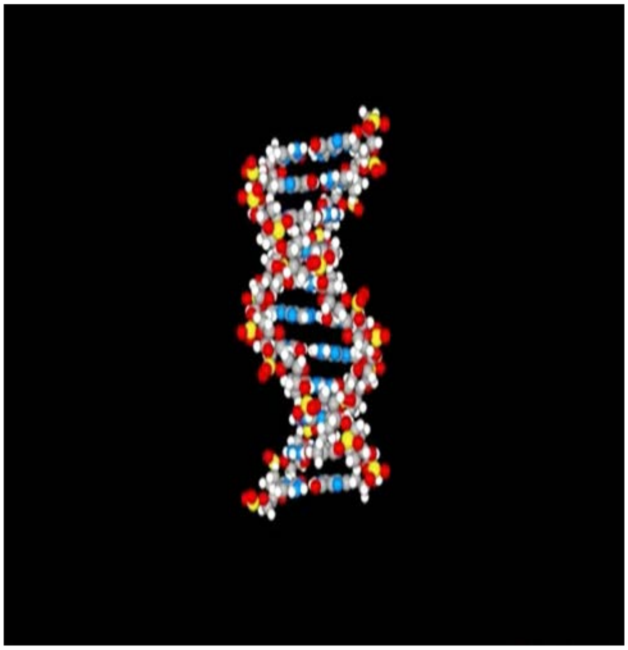
40



PREDICTS CHARGAFF PURINE:PYRIMIDINE RATIOS
When a C must pair with G and When an A must pair with T
DISTANCE R to R IDENTICAL = REGULAR DS DNA

A = ACCEPTOR
D = DONOR
R = "rest of Backbone"

17



Watson-Crick Base Pairs

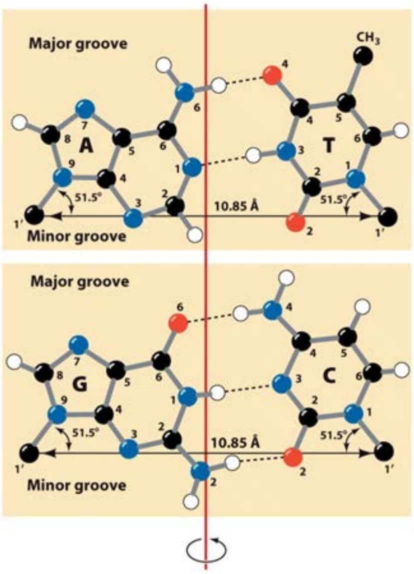
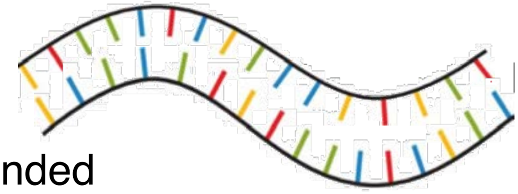
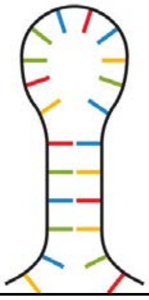


Figure 24-1

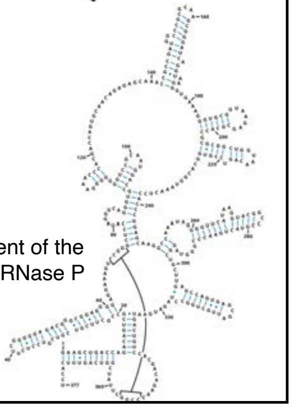


Double-stranded

Single-stranded



Stem-loop



M1 RNA component of the enzyme RNase P

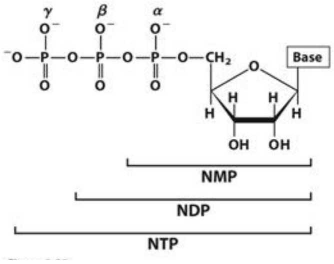
Roles of Nucleotides

The most well known is ATP:

Important roles of other nucleotides:

- Energy rich (high energies of hydrolysis, but kinetically stable)
  besides ATP, includes: GTP, CTP, UTP
- Carrier molecule (key intermediates in metabolism)
  UDP-sugars, CDP-lipids, NADH, FAD
- Secondary messengers (cAMP, cGMP)
- Other cofactors for enzymes

Other Functions of Nucleotides: Energy Source



Abbreviations of ribonucleoside 5′-phosphates

Base	Mono-	Di-	Tri-
Adenine	AMP	ADP	ATP
Guanine	GMP	GDP	GTP
Cytosine	CMP	CDP	CTP
Uracil	UMP	UDP	UTP

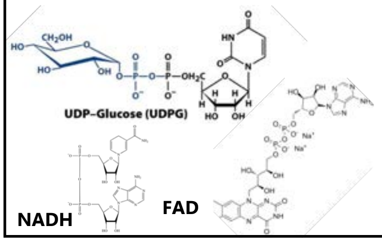
Abbreviations of deoxyribonucleoside 5′-phosphates

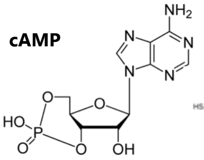
Base	Mono-	Di-	Tri-
Adenine	dAMP	dADP	dATP
Guanine	dGMP	dGDP	dGTP
Cytosine	dCMP	dCDP	dCTP
Thymine	dTMP	dTDP	dTTP

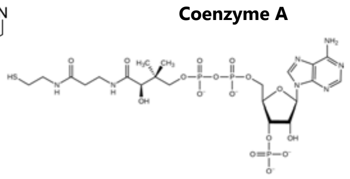
Carrier molecules

Secondary messengers

Other cofactors for enzymes







Roles of Nucleic Acids

- •Information storage
- •Information retrieval
- •Information translation
- •Information processing
- •Information preservation

Key experiments by Griffiths, Avery, and Hershey & Chase*

*on website; not in textbook

The 4 S's

Size
Solubility
Shape
Stability

The 4 S's

Size
Solubility
Shape
Stability

Nucleic Acids: Size

Genome Sizes

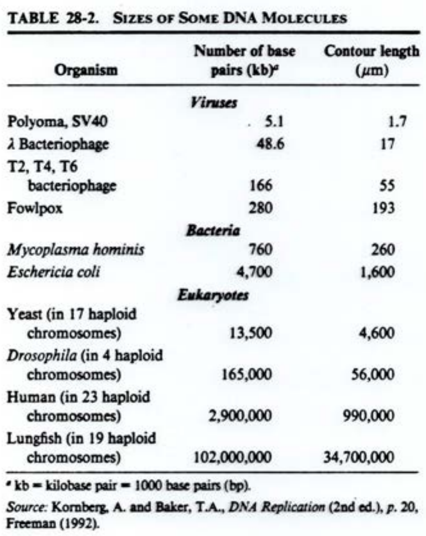


TABLE 28-2. SIZES OF SOME DNA MOLECULES

Organism	Number of base pairs (kb)[a]	Contour length (μm)
	Viruses	
Polyoma, SV40	5.1	1.7
λ Bacteriophage	48.6	17
T2, T4, T6 bacteriophage	166	55
Fowlpox	280	193
	Bacteria	
Mycoplasma hominis	760	260
Eschericia coli	4,700	1,600
	Eukaryotes	
Yeast (in 17 haploid chromosomes)	13,500	4,600
Drosophila (in 4 haploid chromosomes)	165,000	56,000
Human (in 23 haploid chromosomes)	2,900,000	990,000
Lungfish (in 19 haploid chromosomes)	102,000,000	34,700,000

[a] kb = kilobase pair = 1000 base pairs (bp).

Source: Kornberg, A. and Baker, T.A., *DNA Replication* (2nd ed.), *p.* 20, Freeman (1992).

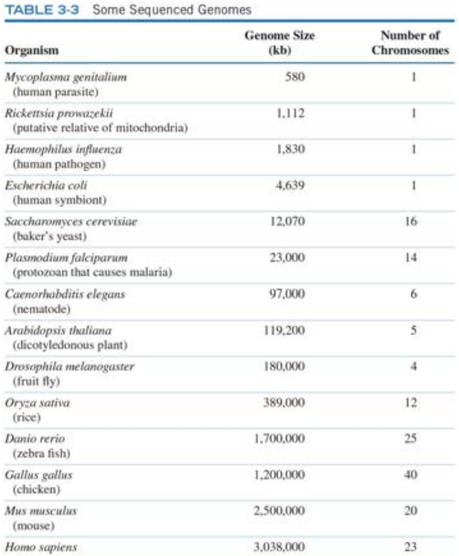


TABLE 3-3 Some Sequenced Genomes

Organism	Genome Size (kb)	Number of Chromosomes
Mycoplasma genitalium (human parasite)	580	1
Rickettsia prowazekii (putative relative of mitochondria)	1,112	1
Haemophilus influenza (human pathogen)	1,830	1
Escherichia coli (human symbiont)	4,639	1
Saccharomyces cerevisiae (baker's yeast)	12,070	16
Plasmodium falciparum (protozoan that causes malaria)	23,000	14
Caenorhabditis elegans (nematode)	97,000	6
Arabidopsis thaliana (dicotyledonous plant)	119,200	5
Drosophila melanogaster (fruit fly)	180,000	4
Oryza sativa (rice)	389,000	12
Danio rerio (zebra fish)	1,700,000	25
Gallus gallus (chicken)	1,200,000	40
Mus musculus (mouse)	2,500,000	20
Homo sapiens	3,038,000	23

Nucleic Acids: Size

Genome Sizes

(from DNA sequence)

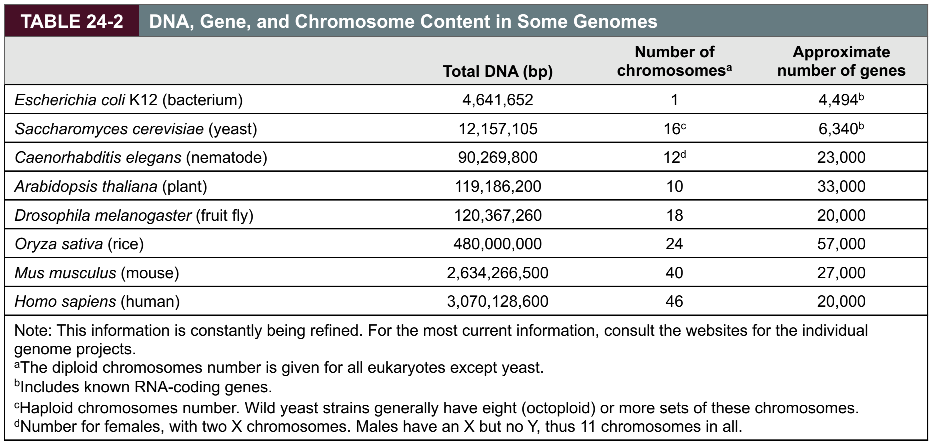


TABLE 24-2 DNA, Gene, and Chromosome Content in Some Genomes

	Total DNA (bp)	Number of chromosomes[a]	Approximate number of genes
Escherichia coli K12 (bacterium)	4,641,652	1	4,494[b]
Saccharomyces cerevisiae (yeast)	12,157,105	16[c]	6,340[b]
Caenorhabditis elegans (nematode)	90,269,800	12[d]	23,000
Arabidopsis thaliana (plant)	119,186,200	10	33,000
Drosophila melanogaster (fruit fly)	120,367,260	18	20,000
Oryza sativa (rice)	480,000,000	24	57,000
Mus musculus (mouse)	2,634,266,500	40	27,000
Homo sapiens (human)	3,070,128,600	46	20,000

Note: This information is constantly being refined. For the most current information, consult the websites for the individual genome projects.

[a] The diploid chromosomes number is given for all eukaryotes except yeast.

[b] Includes known RNA-coding genes.

[c] Haploid chromosomes number. Wild yeast strains generally have eight (octoploid) or more sets of these chromosomes.

[d] Number for females, with two X chromosomes. Males have an X but no Y, thus 11 chromosomes in all.

Nucleic Acids: Size

RNA Sizes

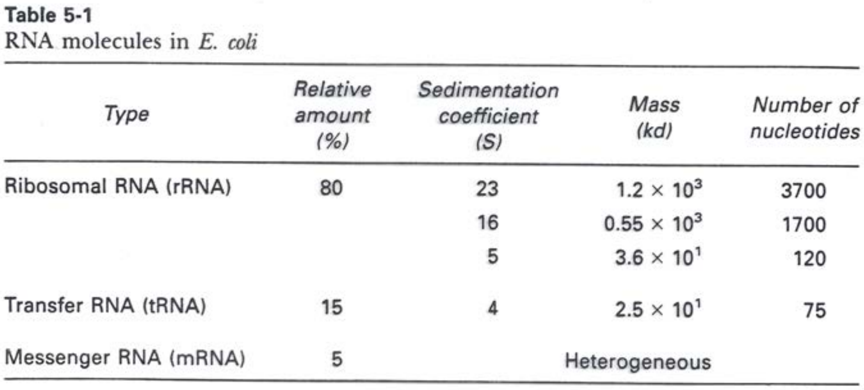


Table 5-1
RNA molecules in *E. coli*

Type	*Relative amount (%)*	*Sedimentation coefficient (S)*	*Mass (kd)*	*Number of nucleotides*
Ribosomal RNA (rRNA)	80	23	1.2×10^3	3700
		16	0.55×10^3	1700
		5	3.6×10^1	120
Transfer RNA (tRNA)	15	4	2.5×10^1	75
Messenger RNA (mRNA)	5		Heterogeneous	

The 4 S's

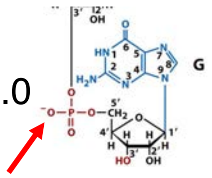
Size

Solubility

Shape

Stability

Nucleic Acids: Solubility



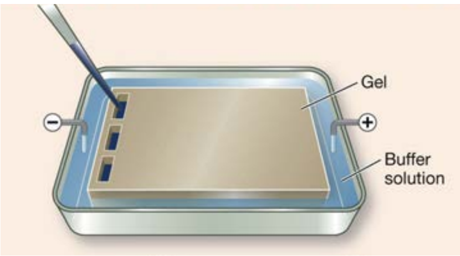
The polymer is a poly-anion

The pK_a of the phosphodiester is ~2.0

Consequences:

1) A counter ion:
   - required for solubility and stability.
   - usually Na^+, K^+, and/or Mg^{+2}
   - in cell, also use of polyamines; spermine & spermidine.
2) Easily separated by electrophoresis:
   - Every nucleotide has one (1) negative charge.
   - so charge/mass ratio is constant.
   - so can separate by size (similar in concept to SDS-PAGE).

Nucleic Acids: Solubility



Nucleic acids can be separated by **gel electrophoresis**.

Polymers can be placed in a well in a semisolid gel and an electric field is applied across the gel.

Negatively charged Nucleic acids move towards positive end.

For polymers of the same shape, smaller polymers travel faster than larger polymers.

For polymers of the same size, more compact topolgies travel faster than less compact ones.

The 4 S's

Size

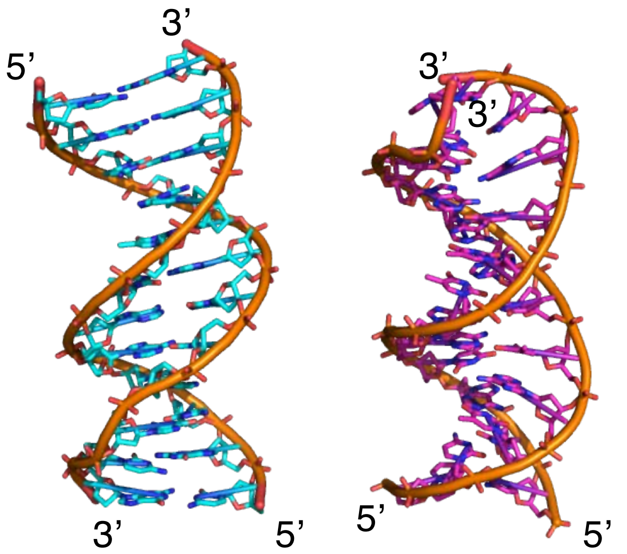
Solubility

Shape

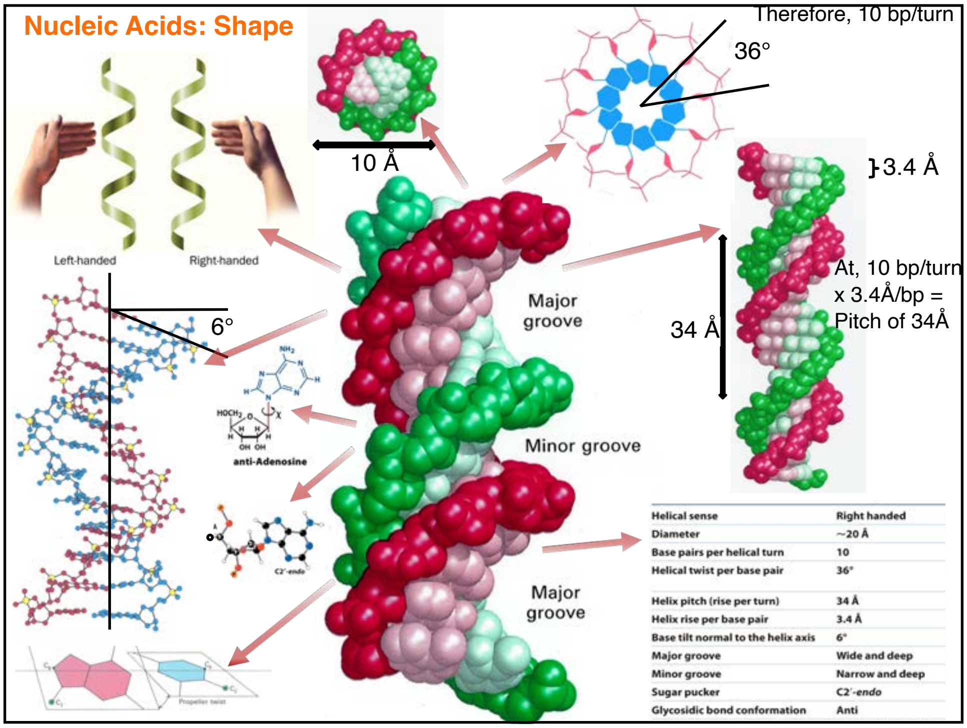
Stability

Nucleic Acids: Shape

ANTIPARALLEL dsDNA **PARALLEL dsDNA**

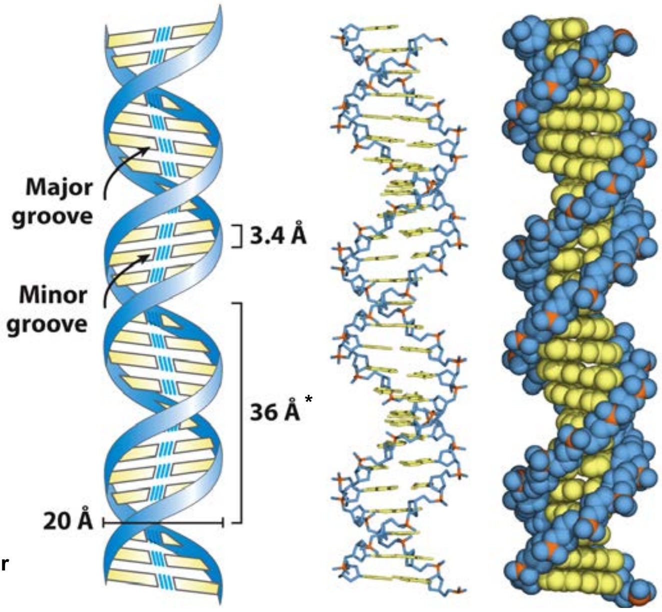


Nucleic Acids: Shape



Helical sense	Right handed
Diameter	~20 Å
Base pairs per helical turn	10
Helical twist per base pair	36°
Helix pitch (rise per turn)	34 Å
Helix rise per base pair	3.4 Å
Base tilt normal to the helix axis	6°
Major groove	Wide and deep
Minor groove	Narrow and deep
Sugar pucker	C2′-*endo*
Glycosidic bond conformation	Anti

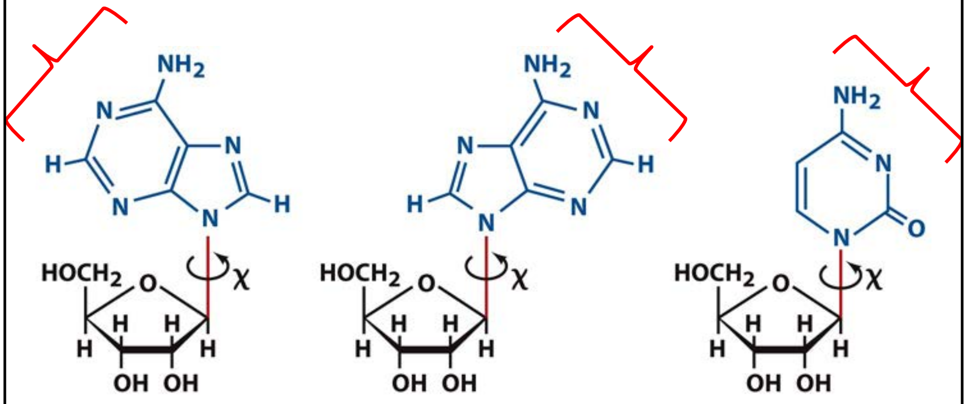
Watson-Crick Model of B-DNA



*(based on 10.5 bases per turn)

Nucleic Acids: Shape

Sterically Allowed Base Orientations



syn-Adenosine **anti-Adenosine** **anti-Cytidine**

B-DNA **B-DNA**

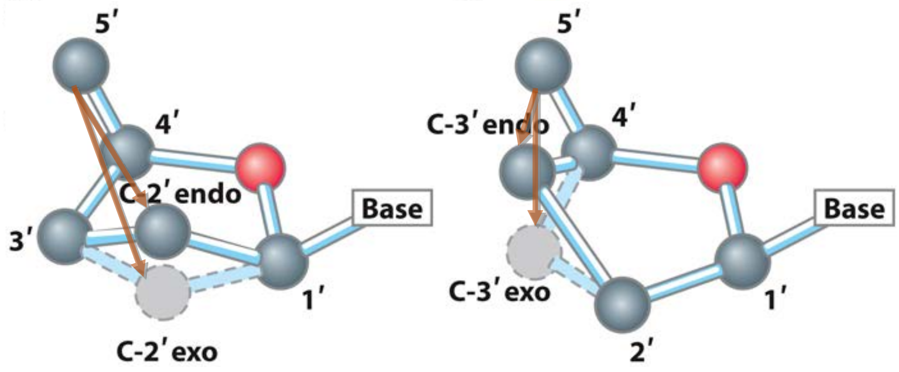
Notice the bp interface

Nucleic Acids: Shape

Nucleotide Sugar Conformations

4 possible pucker conformations:

NUCLEIC ACIDS ARE ALWAYS ***endo***

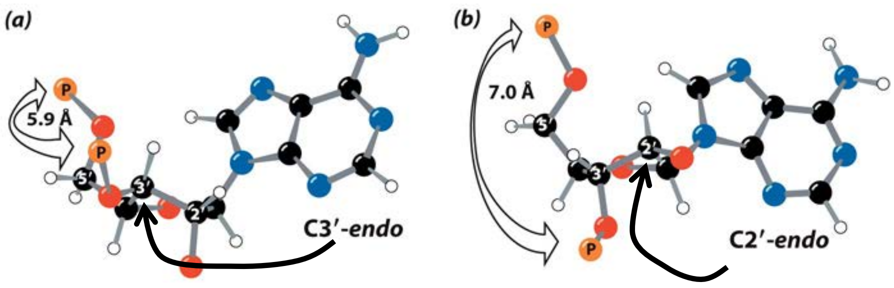


RNA-DNA hybrid
A-DNA

What is this "A-DNA"?

Nucleic Acids: Shape

Nucleotide Sugar Conformations

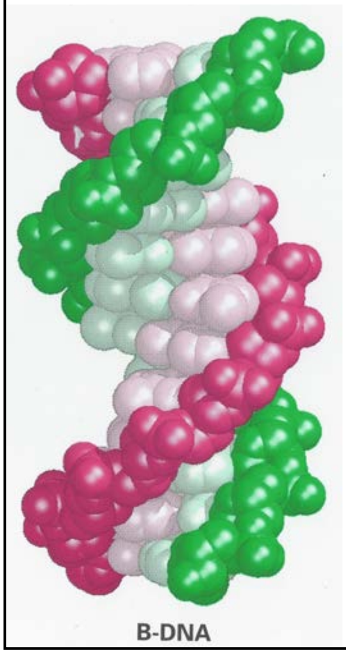


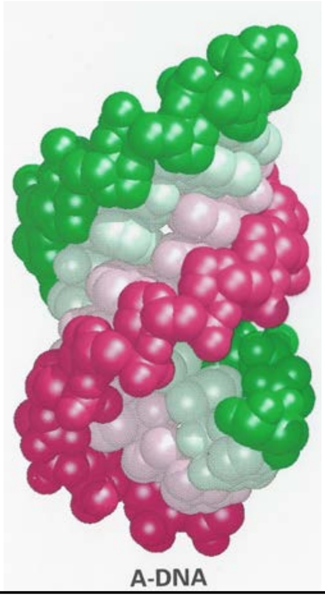
RNA
RNA-DNA hybrid
A-DNA

B-DNA

What is this “A-DNA”?

Nucleic Acids: Shape





Nucleic Acids: Shape

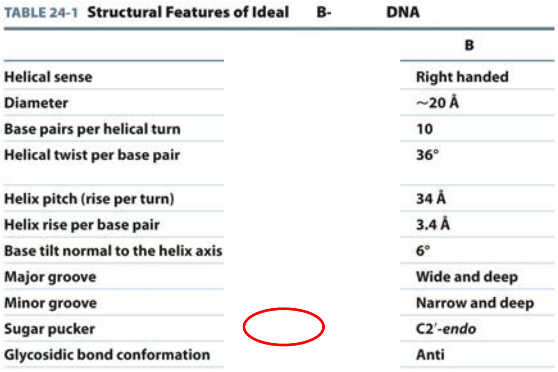
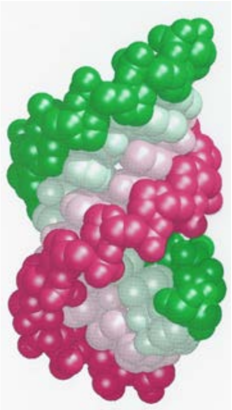


TABLE 24-1 Structural Features of Ideal B- DNA

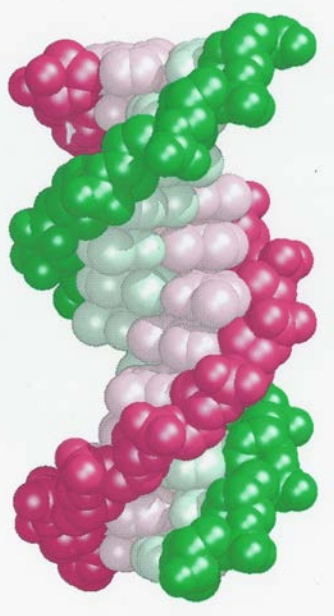
	B
Helical sense	Right handed
Diameter	~20 Å
Base pairs per helical turn	10
Helical twist per base pair	36°
Helix pitch (rise per turn)	34 Å
Helix rise per base pair	3.4 Å
Base tilt normal to the helix axis	6°
Major groove	Wide and deep
Minor groove	Narrow and deep
Sugar pucker	C2′-*endo*
Glycosidic bond conformation	Anti



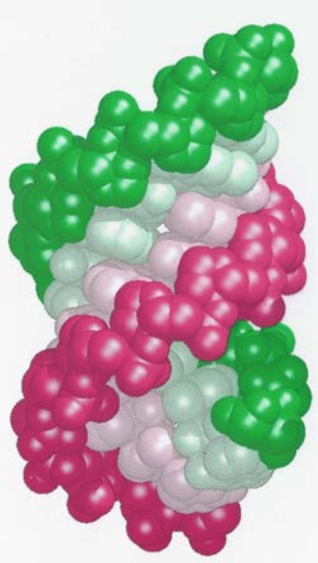
A-DNA

Are there other shapes of DNA?

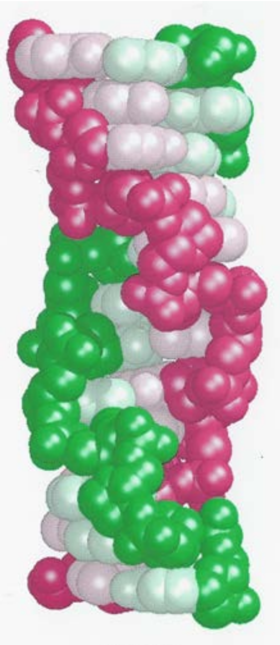
Nucleic Acids: Shape



B-DNA



A-DNA



Z-DNA

Nucleic Acids: Shape

Structural Features of A-, B-, & Z-DNA

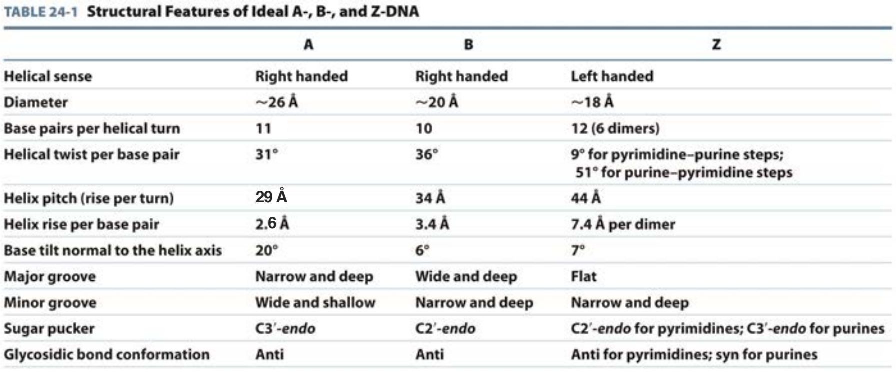
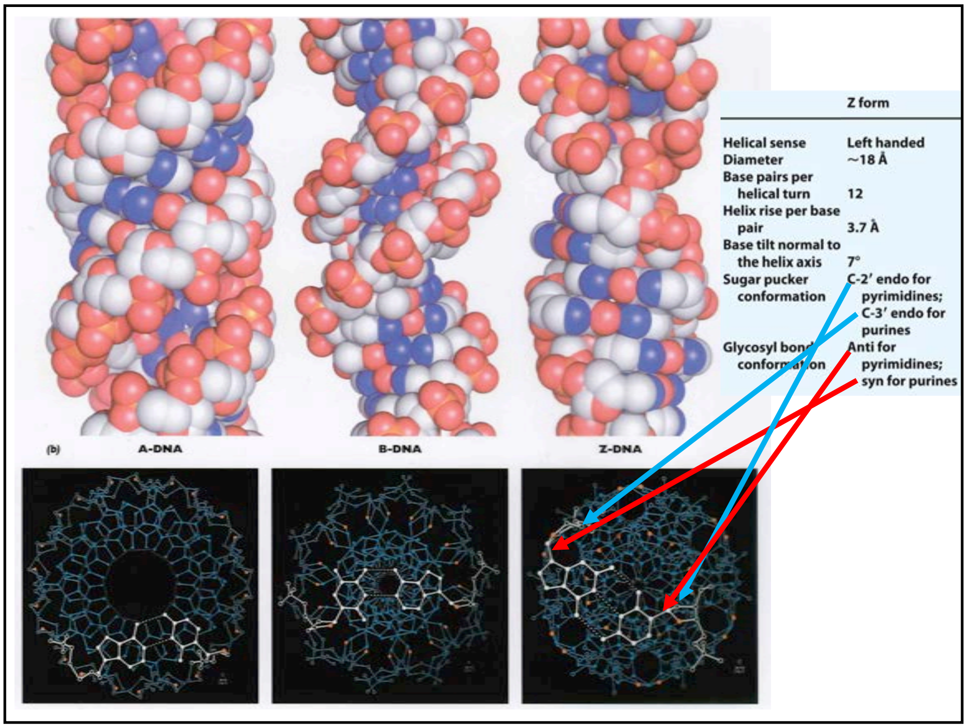


TABLE 24-1 **Structural Features of Ideal A-, B-, and Z-DNA**

	A	B	Z
Helical sense	Right handed	Right handed	Left handed
Diameter	~26 Å	~20 Å	~18 Å
Base pairs per helical turn	11	10	12 (6 dimers)
Helical twist per base pair	31°	36°	9° for pyrimidine–purine steps; 51° for purine–pyrimidine steps
Helix pitch (rise per turn)	29 Å	34 Å	44 Å
Helix rise per base pair	2.6 Å	3.4 Å	7.4 Å per dimer
Base tilt normal to the helix axis	20°	6°	7°
Major groove	Narrow and deep	Wide and deep	Flat
Minor groove	Wide and shallow	Narrow and deep	Narrow and deep
Sugar pucker	C3′-*endo*	C2′-*endo*	C2′-*endo* for pyrimidines; C3′-*endo* for purines
Glycosidic bond conformation	Anti	Anti	Anti for pyrimidines; syn for purines



	Z form
Helical sense	Left handed
Diameter	~18 Å
Base pairs per helical turn	12
Helix rise per base pair	3.7 Å
Base tilt normal to the helix axis	7°
Sugar pucker conformation	C-2′ endo for pyrimidines; C-3′ endo for purines
Glycosyl bond conformation	Anti for pyrimidines; syn for purines

Other Forms of DNA

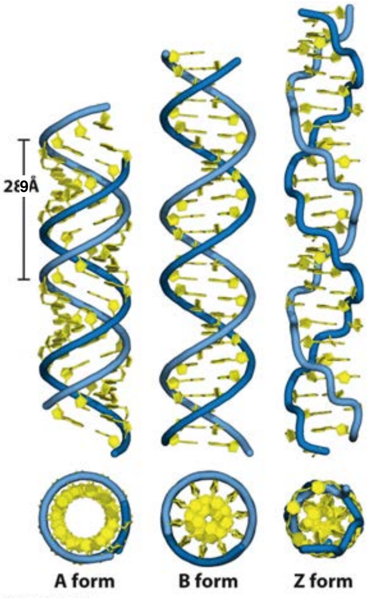
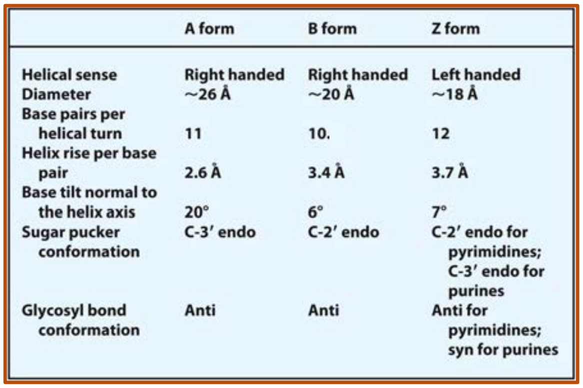


Figure 8-17 part 1



	A form	B form	Z form
Helical sense	Right handed	Right handed	Left handed
Diameter	~26 Å	~20 Å	~18 Å
Base pairs per helical turn	11	10.	12
Helix rise per base pair	2.6 Å	3.4 Å	3.7 Å
Base tilt normal to the helix axis	20°	6°	7°
Sugar pucker conformation	C-3′ endo	C-2′ endo	C-2′ endo for pyrimidines; C-3′ endo for purines
Glycosyl bond conformation	Anti	Anti	Anti for pyrimidines; syn for purines

Nucleic Acids: Shape

Nucleotide Sugar Conformations

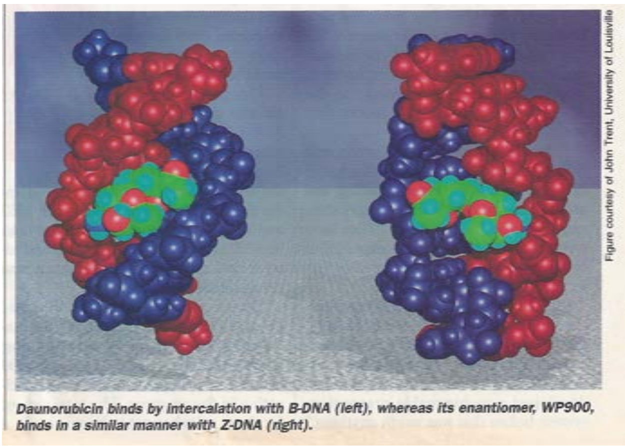


Figure courtesy of John Trent, University of Louisville

Daunorubicin binds by intercalation with B-DNA (left), whereas its enantiomer, WP900, binds in a similar manner with Z-DNA (right).

Nucleic Acids: Global Shape

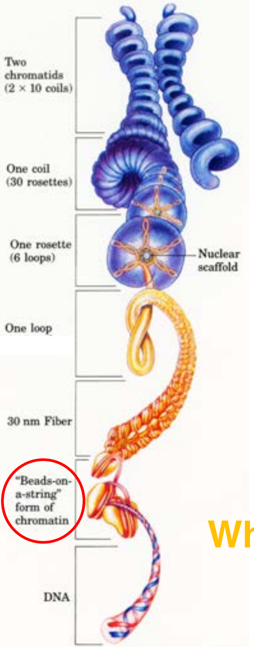
Metaphase Chromosome

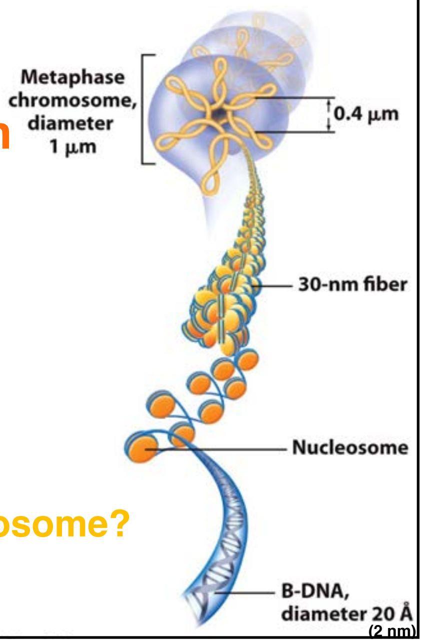
How do we get something that is 2-10 cm long into one of these, which is only 10 μm?

This condensation is 10,000x. Even interphase its 1000x

Nucleic Acids: Global Shape

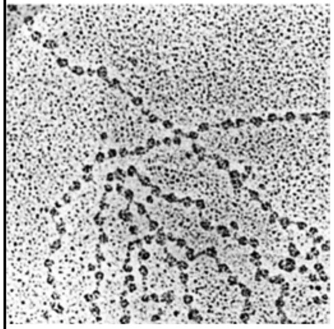
Metaphase Chromatin Organization



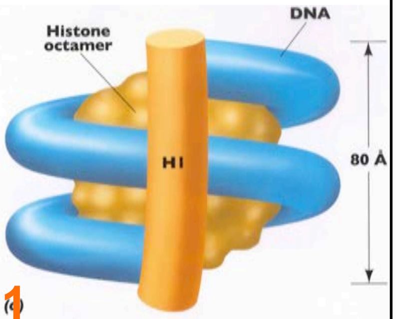


What is this nucleosome?

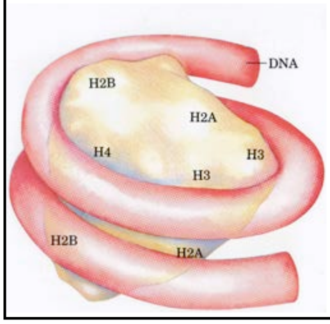
Nucleic Acids: Global Shape



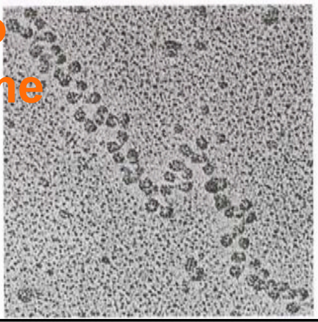
~240 bp (816Å) is now compacted to 80Å = 10x



Histone H1 Bound to Nucleosome



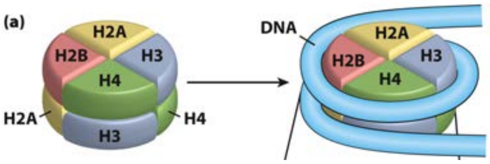
8 Histones make up the core

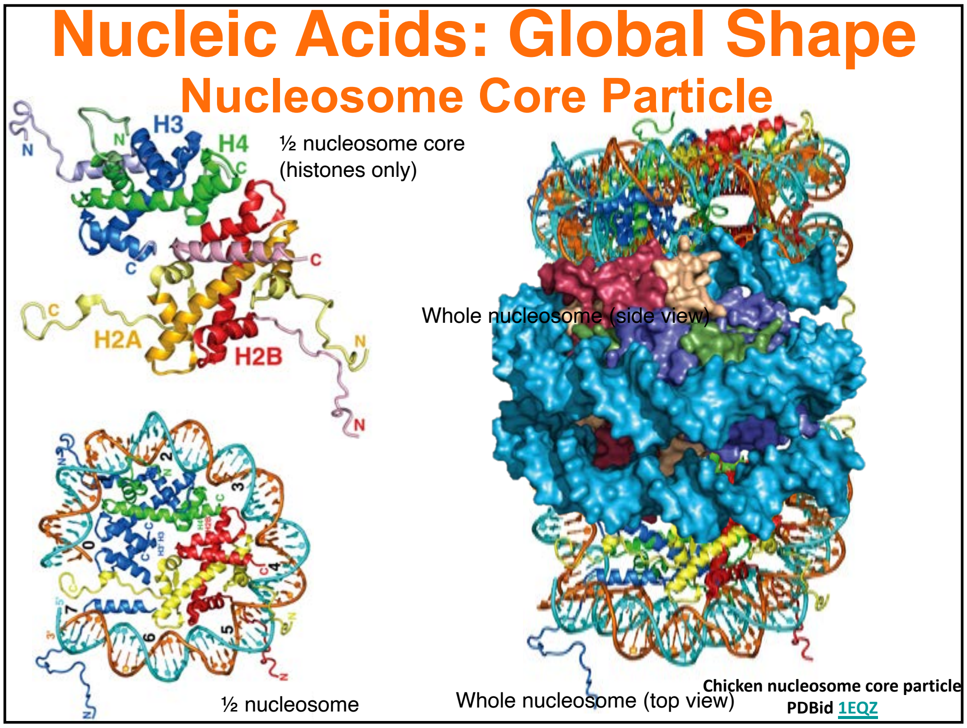


Nucleic Acids: Global Shape

Histones Are Highly Conserved

Histone	Number of Residues	Mass (kD)	% Arg	% Lys	Stoichiometry
H1	215	23.0	1	29	1
H2A	129	14.0	9	11	2
H2B	125	13.8	6	16	2
H3	135	15.3	13	10	2
H4	102	11.3	14	11	2





Nucleic Acids: Global Shape

Nucleosome Core Particle

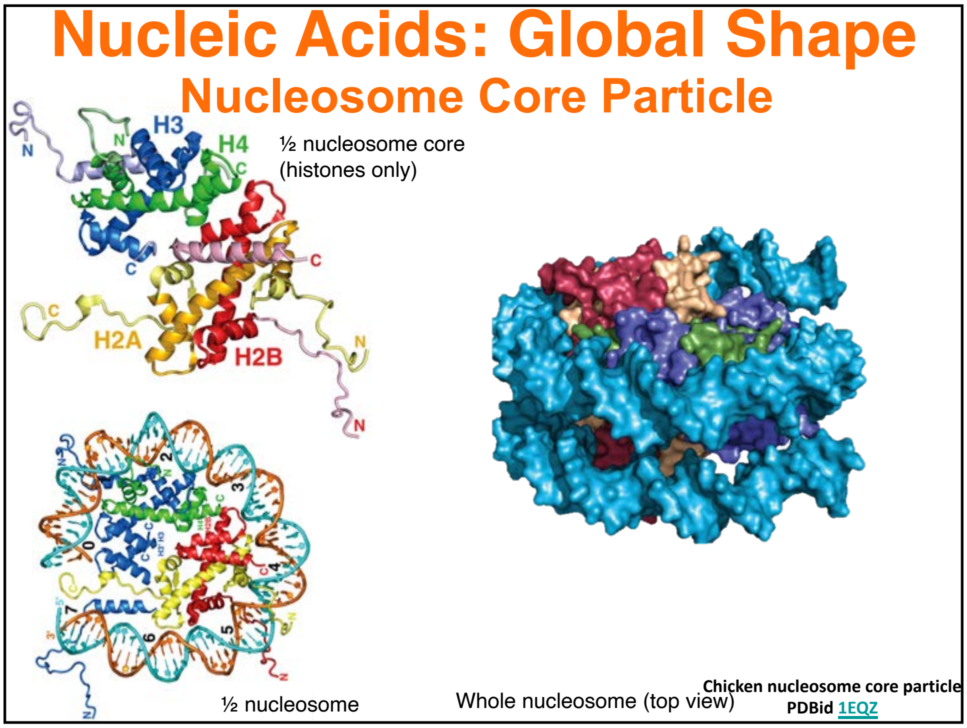
½ nucleosome core (histones only)

Whole nucleosome (side view)

½ nucleosome

Whole nucleosome (top view)

Chicken nucleosome core particle
PDBid [1EQZ]



Nucleic Acids: Global Shape

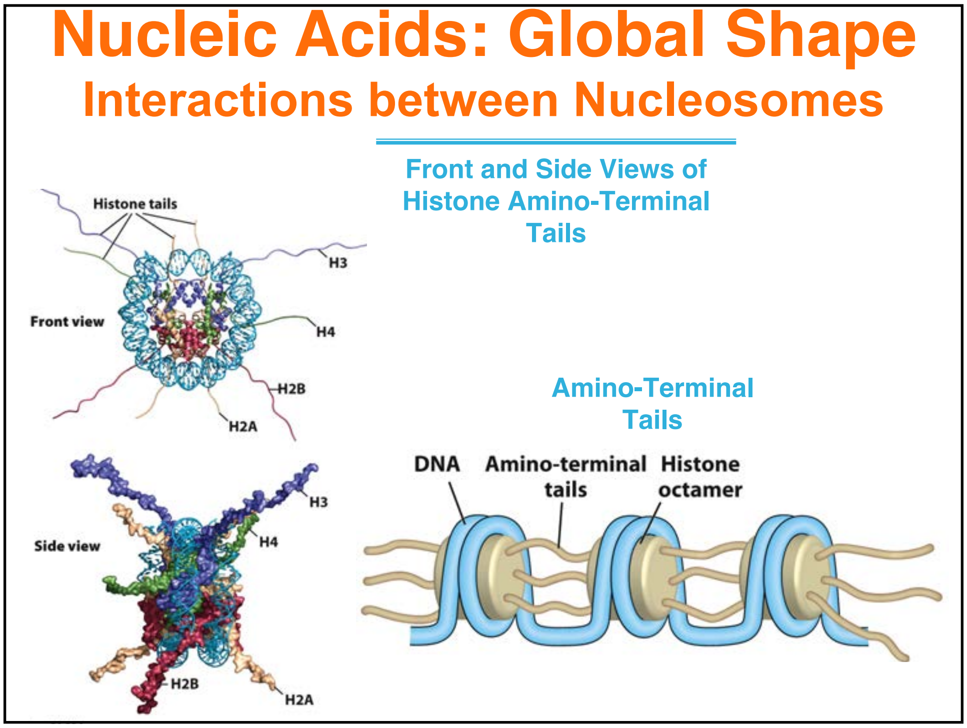
Nucleosome Core Particle

½ nucleosome core (histones only)

½ nucleosome

Whole nucleosome (top view)

Chicken nucleosome core particle
PDBid [1EQZ]

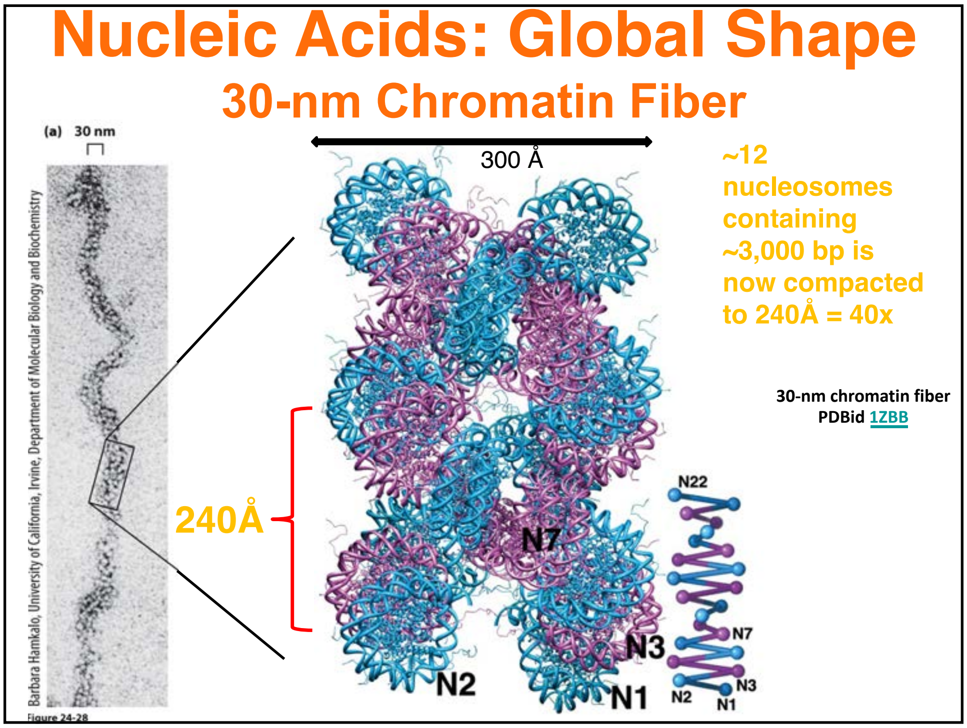


Nucleic Acids: Global Shape

Interactions between Nucleosomes

Front and Side Views of Histone Amino-Terminal Tails

Amino-Terminal Tails



Nucleic Acids: Global Shape

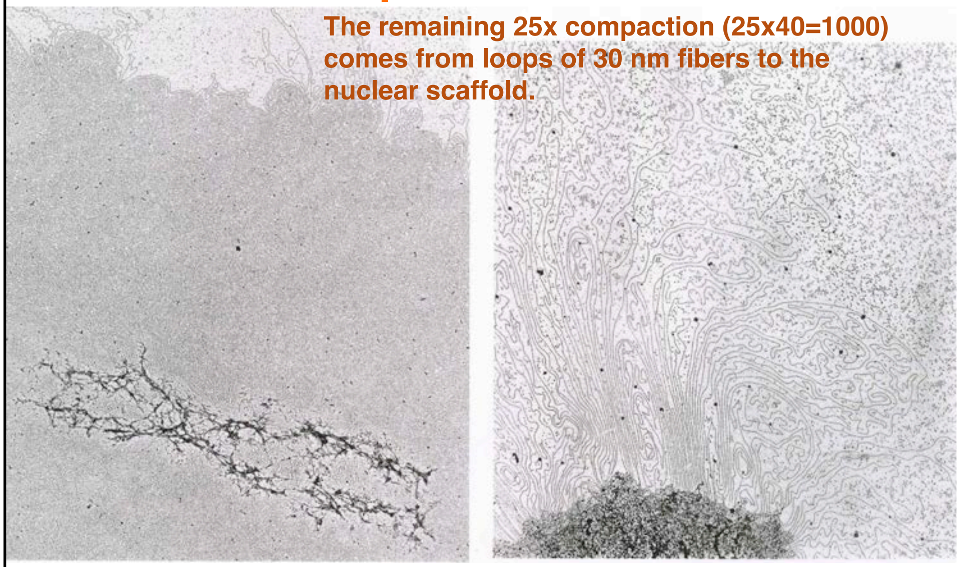
30-nm Chromatin Fiber

~12 nucleosomes containing ~3,000 bp is now compacted to 240Å = 40x

30-nm chromatin fiber
PDBid 1ZBB

Nucleic Acids: Global Shape

Histone-Depleted Chromosome



The remaining 25x compaction (25x40=1000) comes from loops of 30 nm fibers to the nuclear scaffold.

DNA Loops Attached to Scaffold

Nucleic Acids: Global Shape

All this looping leads to a supertwist of the DNA…

Nucleosomal DNA Is Underwound

- Wrapping DNA around the histone core requires removal of one helical turn.
  - The underwinding occurs without a strand break, so a compensatory (+) supercoil forms.
  - This (+) supercoil is relaxed by a topoisomerase, which puts in 2 negative supercoils, leaving DNA with **net (–1) negative supercoil**.

What is Supercoiling: the Coiling of a Coil

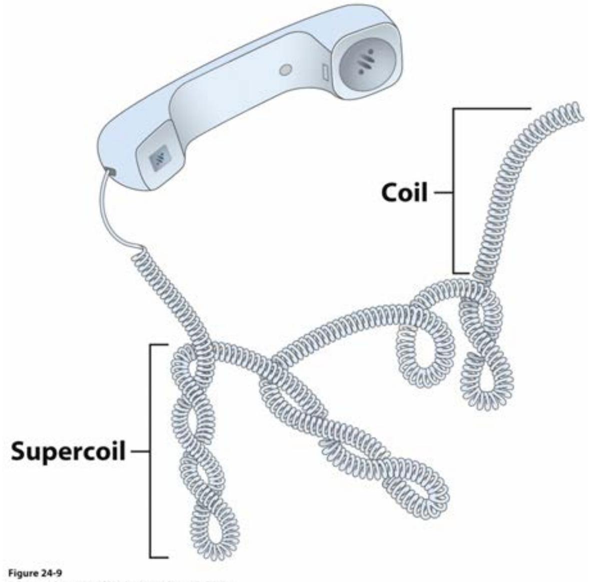


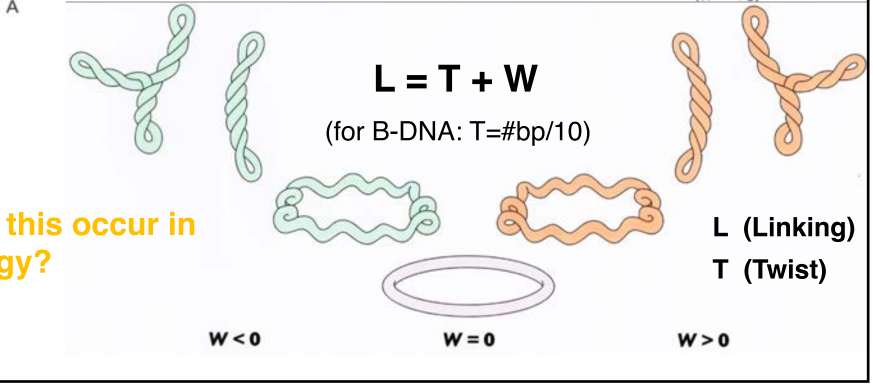
Figure 24-9

Nucleic Acids: Global Shape

All this looping leads to a supertwist of the DNA…

Topology

Supercoiling is called W (writhing number)



$$L = T + W$$

(for B-DNA: T=#bp/10)

L (Linking)

T (Twist)

Does this occur in biology?

Nucleic Acids: Global Shape

Topology

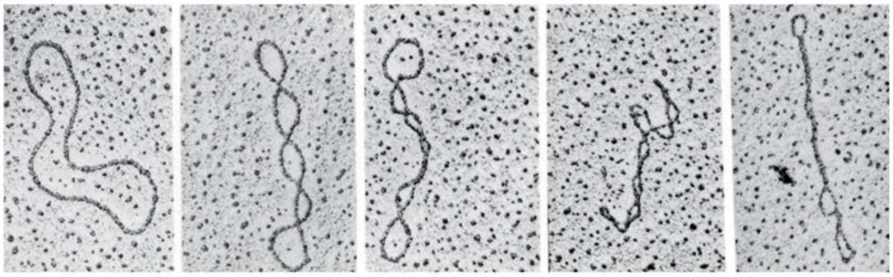


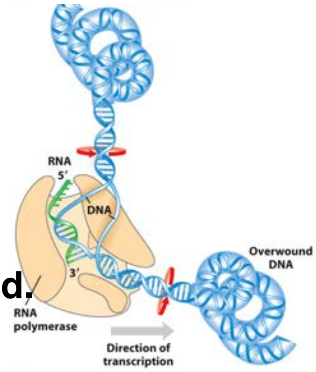
Figure 24-8
Electron micrographs by Laurien Polder. From Kornberg, A. and Baker, T.A., *DNA Replication* (2nd ed.), p. 36, W.H. Freeman (1992). Permission provided courtesy of Roger Kornberg.

Is supercoiling only possible with circular DNA? Eukaryotic chromosomes are linear…..

Supercoiling of Circular Duplex DNA from the adenovirus, SV40

Nucleic Acids: Global Shape

Topology



- Many linear DNAs bound to proteins are supercoiled.
- Supercoiling has great influence on transcription and replication of DNA.
- Supercoiling can be highly regulated.

Nucleic Acids: Global Shape

Consequences of supercoiling:

1) Required for information retrieval; must be negative
2) All circular extra-chromosomal DNAs are negatively supercoiled
3) Can be used for isolation of these DNAs in the laboratory

Nucleic Acids: Global Shape

Techniques to Detect Topology

<u>Separate DNAs with different topology based on density</u>

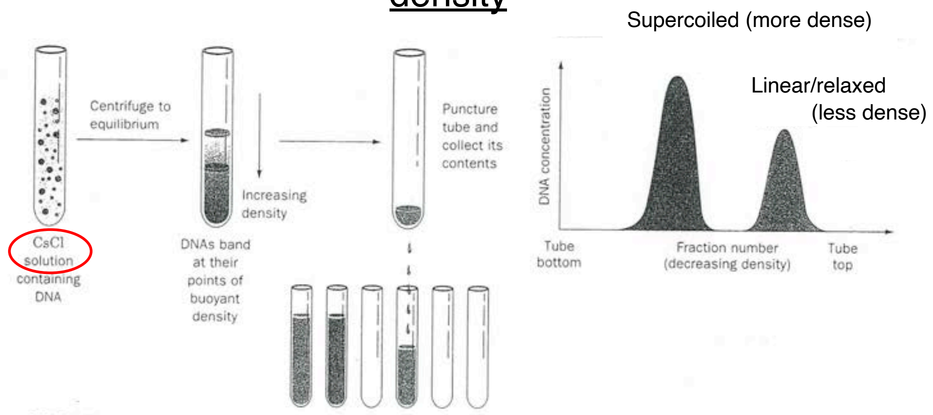
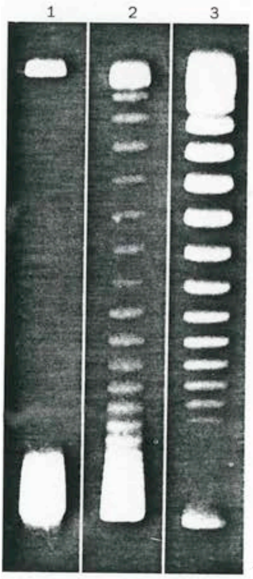


Figure 23-31 The separation of DNAs by equilibrium density gradient ultracentrifugation in CsCl solution. An initially 8 M CsCl solution forms a density gradient that varies linearly from ~1.80 $g \cdot cm^{-3}$ at the bottom of the centrifuge tube to ~1.55 $g \cdot cm^{-3}$ at the top. The sedimentation rates of the DNAs depend on their base composition. The amount of DNA in each fraction is estimated from its UV absorbance, usually at 260 nm.

Nucleic Acids: Global Shape

Techniques to Detect Topology

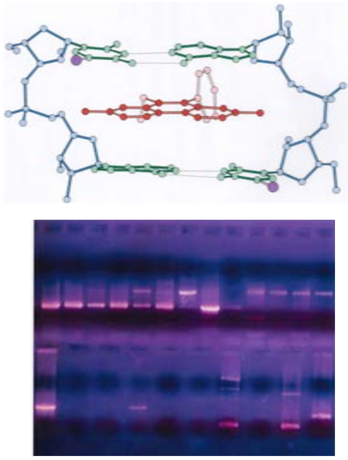
Modify topology with topoisomerase I



What is Topo I?

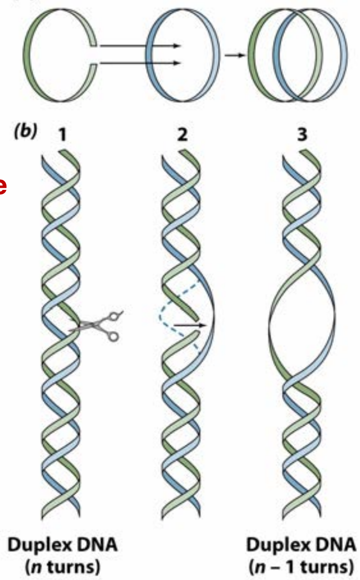
FIGURE 28-40. The agarose gel electrophoresis pattern of SV40 DNA. Lane 1 contains the negatively supercoiled native DNA *(lower band;* the DNA was applied to the top of the gel). In lanes 2 and 3, the DNA has been exposed for 5 and 30 min, respectively, to an enzyme, known as a Type I topoisomerase (Section 28-5C), that relaxes negative supercoils one at a time by increasing the linking number (*L*). The DNAs in consecutively higher bands of a given gel have successively increasing linking numbers ($\Delta L = +1$). [From Keller, W., *Proc. Natl. Acad. Sci.* **72**, 2553 (1975).]

Gel Electrophoresis/Ethidium-Bromide Staining



Nucleic Acids: Global Shape

Type IA Topoisomerase Activity



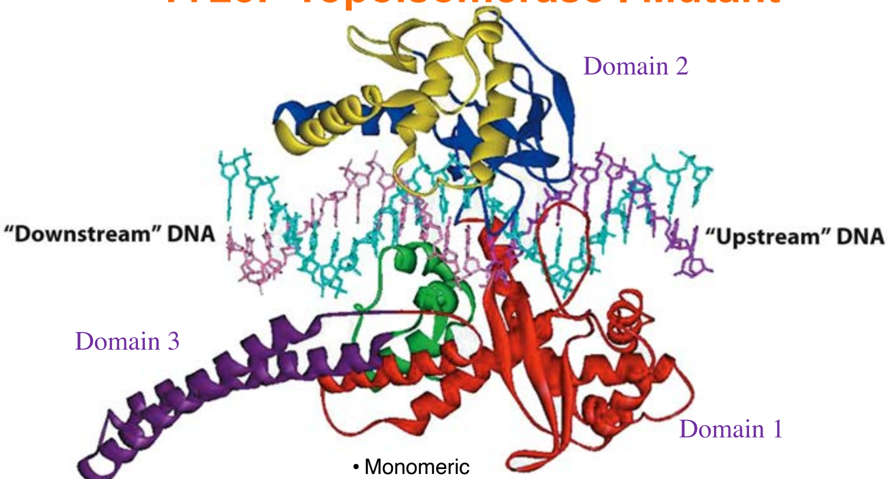
Can generate concatemers of circular ss-DNAs

Can take supercoiled DNA and make it less supercoiled; usually less negative (L = +1)

Type I Topoisomerase also called "nicking-closing" enzyme

Topoisomerase

Y723F Topoisomerase I Mutant

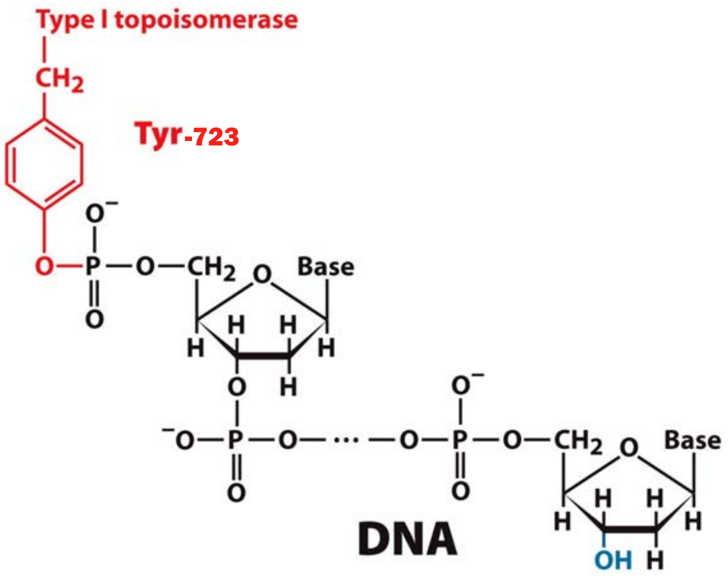


- Monomeric
- Creates single-stranded breaks
- Use of covalent Tyr intermediate
- For negatively supercoiled DNA, it changes L by +1 (if $W \ngeq 0$)
- Uses energy in supercoils (W# ≈ 9 kcal/mole)
- IA & IB types differ in how they relax the cleaved strand

Y723F mutant of human topoisomerase I
PDBid 1A36

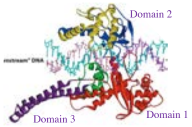
Topoisomerase

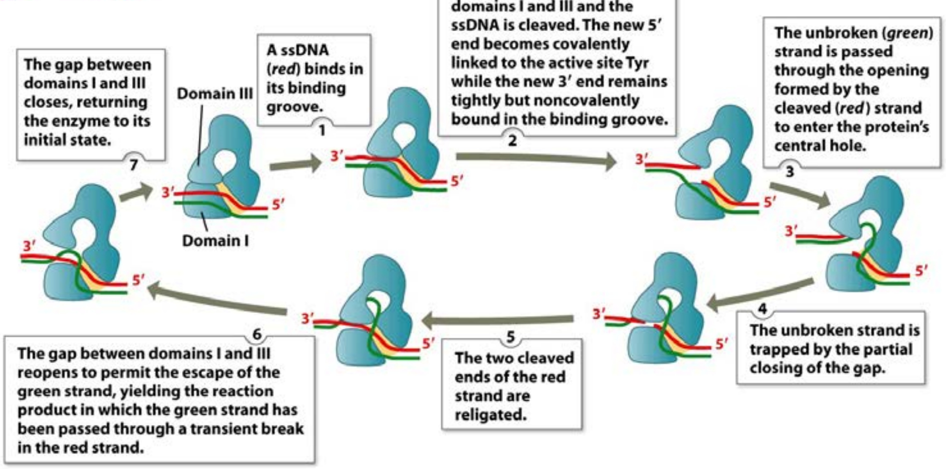
Covalent Topoisomerase-DNA Adduct



Topoisomerase

Type IA Topoisomerase Mechanism





1 A ssDNA (*red*) binds in its binding groove.

2 A gap opens between domains I and III and the ssDNA is cleaved. The new 5′ end becomes covalently linked to the active site Tyr while the new 3′ end remains tightly but noncovalently bound in the binding groove.

3 The unbroken (*green*) strand is passed through the opening formed by the cleaved (*red*) strand to enter the protein's central hole.

4 The unbroken strand is trapped by the partial closing of the gap.

5 The two cleaved ends of the red strand are religated.

6 The gap between domains I and III reopens to permit the escape of the green strand, yielding the reaction product in which the green strand has been passed through a transient break in the red strand.

7 The gap between domains I and III closes, returning the enzyme to its initial state.

How are negative supercoils generated in the first place?