Biological Building Blocks

monomer		macromolecule
√ Amino acids	→	Proteins
Nucleotides	→	Nucleic acids
Fatty acids	→	Lipids
√ Monosaccharides	→	Polysaccharides

A limited number of different monomeric building blocks are used to build a wide variety of different macromolecules

Precursors → Monomeric units → polymeric units: 1°
↓
polymeric units: 2°
↓
polymeric units: 3°
↓
polymeric units: 4°

- *Enzymes catalyze the synthesis of the complex biomolecules required for the assembly and function of cells and organisms*
- *The primary sequence of enzymes determines their structure and function*
- *What determines the primary sequence of a protein?*
- *How are cells faithfully and accurately duplicated?*

Nucleic acids: the molecular repositories for genetic information

Functions:
(1)Define the primary sequence of proteins, direct and accomplish protein synthesis
(2)Store, express, propagate genetic information

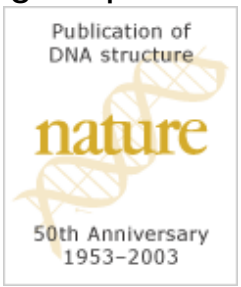
Nucleotides are the building blocks of nucleic acids

Other functions for nucleotides:

- Carriers of chemical energy (ATP)
- Components of enzyme cofactors
- Signalling molecules, cellular communication
- Components of metabolic intermediates
- Modulate enzyme activity- covalent and allosteric

An overview of nucleic acids

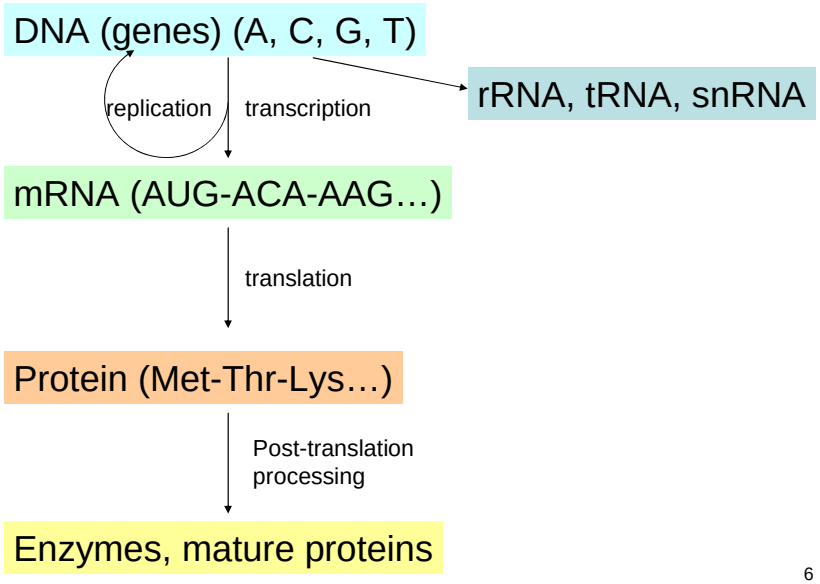
- DNA: deoxyribonucleic acid- the hereditary molecule in all cellular life forms. The storage of information is DNA's only known function
- Genes: segments of DNA that contain the information required for the synthesis of a functional biological product (protein, RNA)

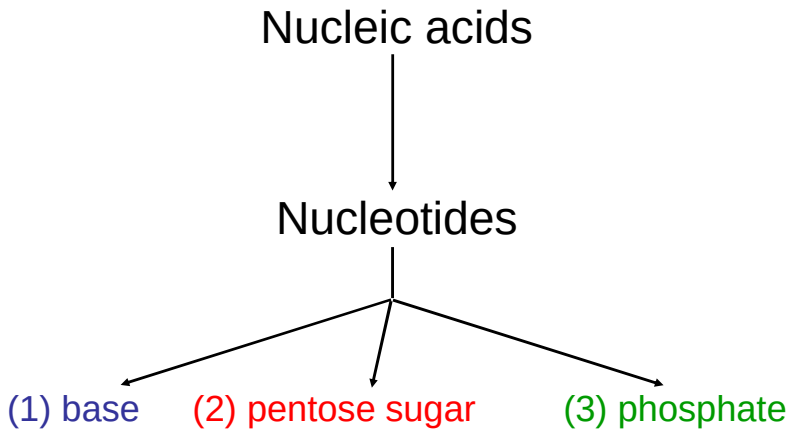


An overview of nucleic acids

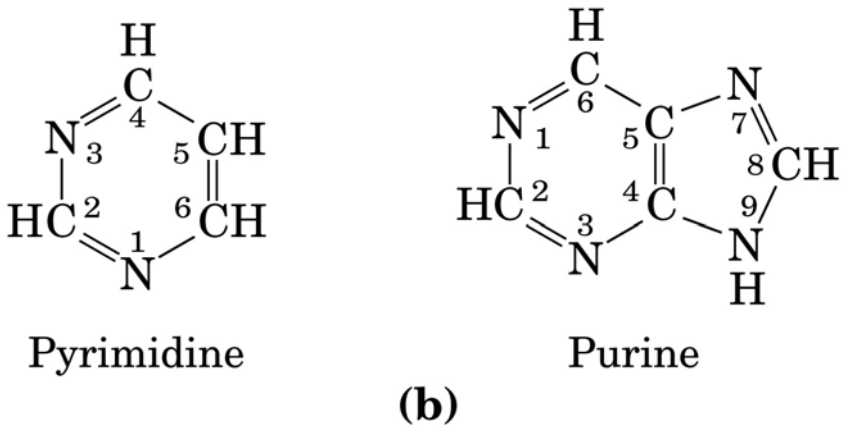
- RNA- ribonucleic acid:
  - ribosomal RNA (rRNA): structural and functional component of ribosomes, the large complexes that carry out protein synthesis
  - Messenger RNA (mRNA): specify primary structure of proteins; carry information from one or more genes to ribosome
  - Transfer RNA (tRNA): transfer information in mRNA into specific sequence of amino acids
  - "genomic" RNA: RNA viruses
  - Special function RNAs: enzymes (ribozymes), snRNA, RNA processing

Overview of flow of genetic information



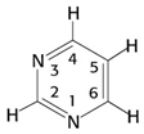


Parent compounds for nitrogenous bases

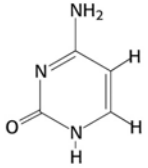


I. Nitrogenous bases

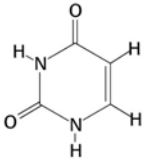
PYRIMIDINES



Pyrimidine

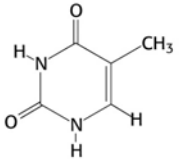


Cytosine



Uracil

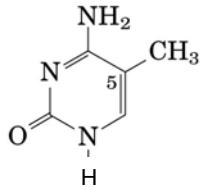
RNA only



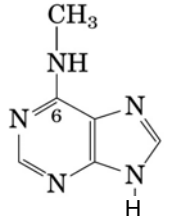
Thymine

DNA only

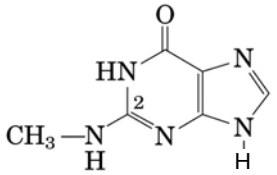
Some modified nucleotide bases



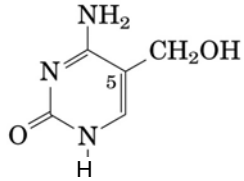
5-methylcytosine



N^6-methyladenine



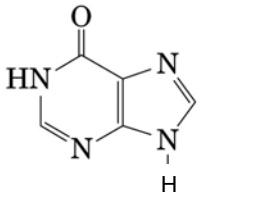
N^2-methylguanine



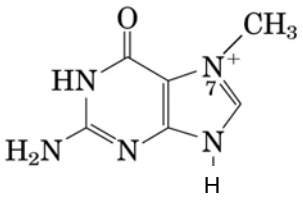
5-hydroxymethylcytosine

(a)

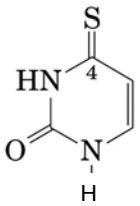
Some modified nucleotide bases



Inosine



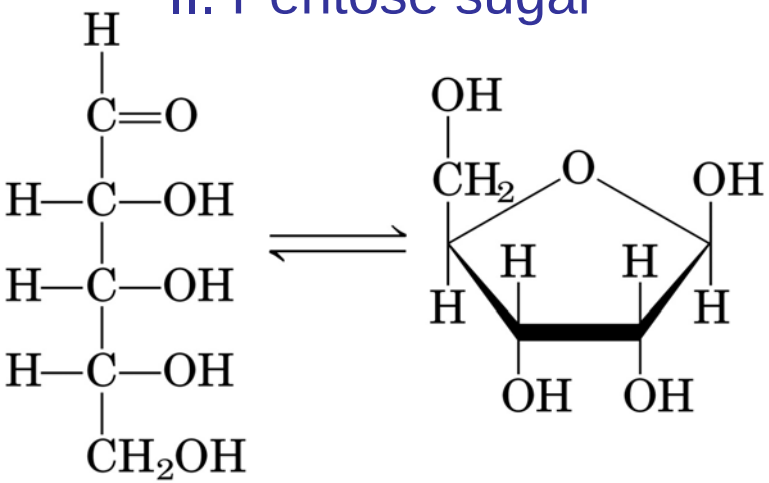
7-methylguanine



4-thiouracil

(b)

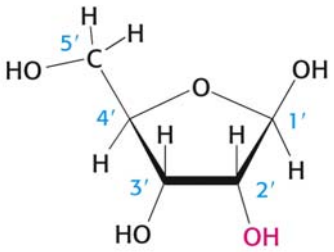
II. Pentose sugar

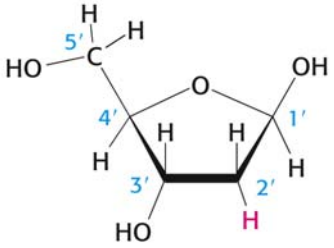


Aldehyde

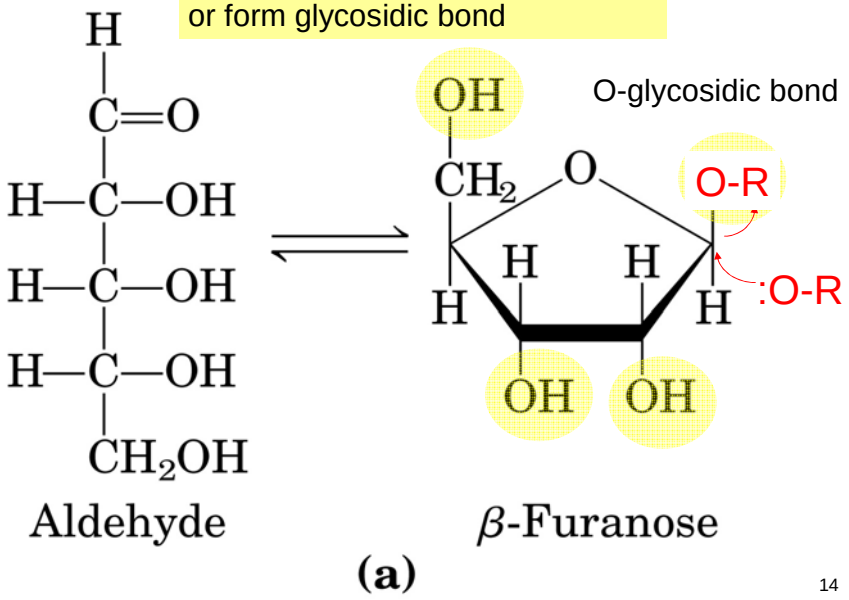
β-Furanose

(a)

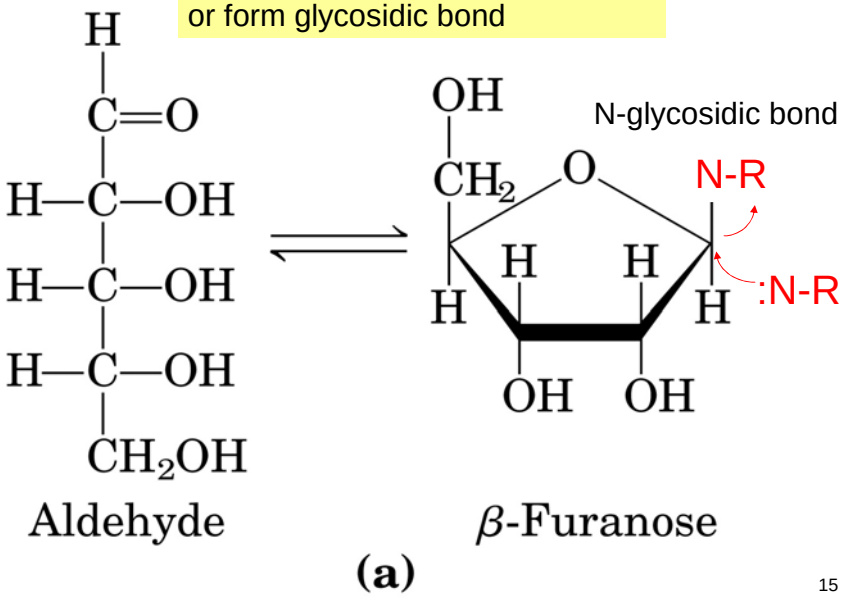




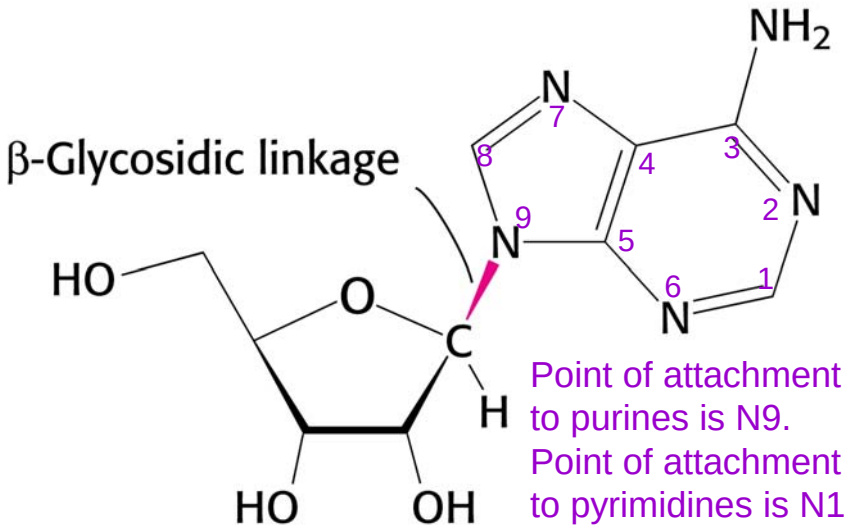
Reactive hydroxyl groups: esterify or form glycosidic bond



Reactive hydroxyl groups: esterify or form glycosidic bond



Nitrogenous base + ribose = nucleoside



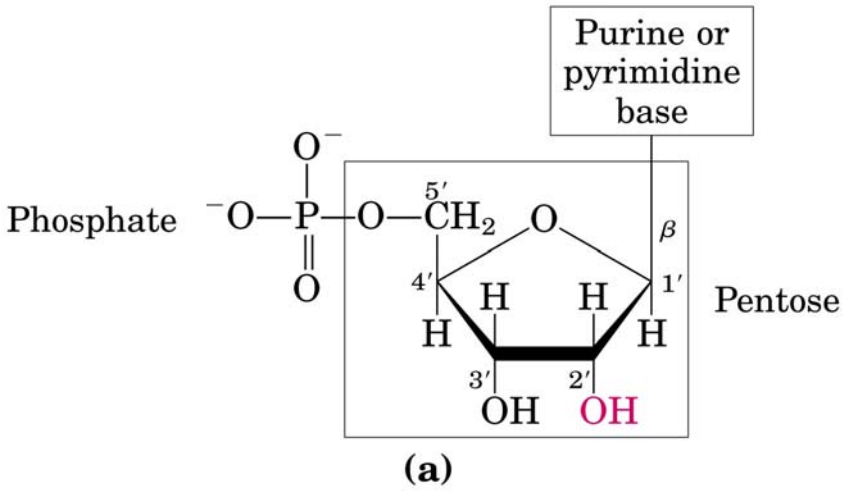
Naming nucleosides

- Adenine + ribose = adenosine
- Adenine + deoxyribose = deoxyadenosine
- guanine + ribose = guanosine
- cytosine + ribose = cytidine
- thymine + deoxyribose = thymidine
- Uracil + ribose = uridine

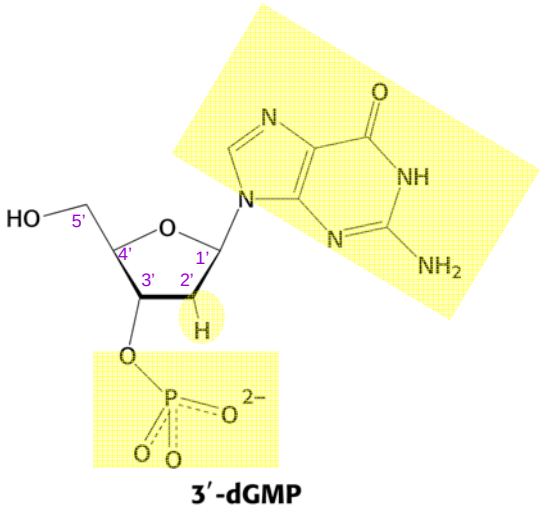
III. Phosphate ester

- Note multiple sites for phosphate esterification
- Monophosphate
- Diphosphate
- Triphosphate
- Cyclic monophosphate

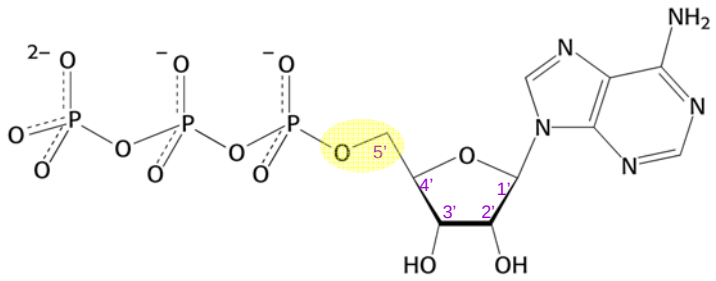
Nucleoside + phosphate = nucleotide



3'-deoxyguanosinemonophosphate

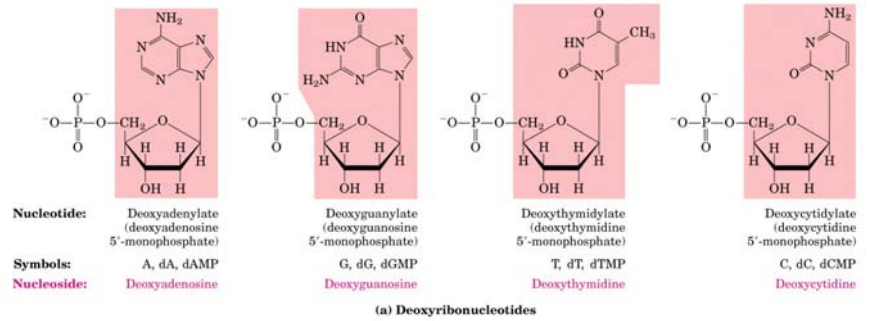


5’-adenosinetriphosphate



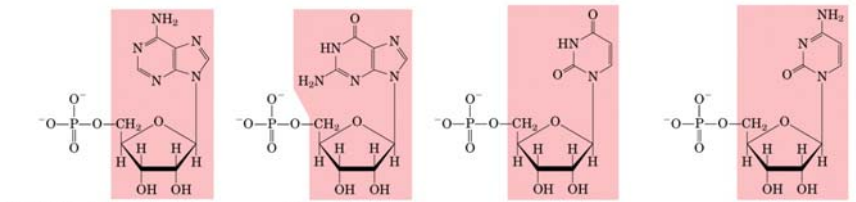
5′-ATP

When point of attachment is the 5’ hydroxyl, the “5’” can be left off when naming



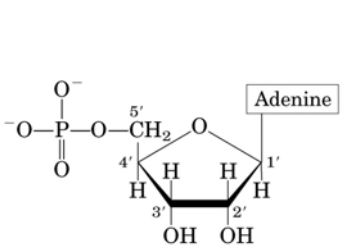
(a) Deoxyribonucleotides

When point of attachment is the 5’ hydroxyl, the “5’” can be left off when naming

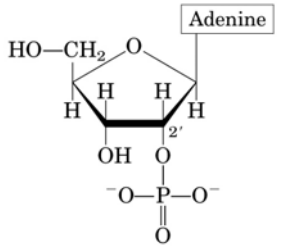


(b) Ribonucleotides

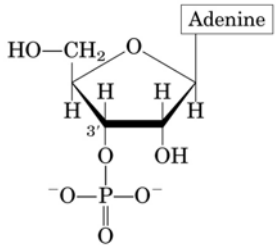
Some more nucleoside monophosphates



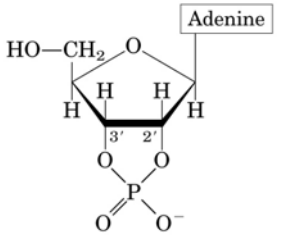
Adenosine 5′-monophosphate



Adenosine 2′-monophosphate



Adenosine 3′-monophosphate



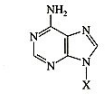
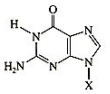
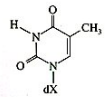
Adenosine 2′,3′-cyclic monophosphate

table 10–1

Nucleotide and Nucleic Acid Nomenclature

Base	Nucleoside*	Nucleotide*	Nucleic acid
Purines			
Adenine	Adenosine	Adenylate	RNA
	Deoxyadenosine	Deoxyadenylate	DNA
Guanine	Guanosine	Guanylate	RNA
	Deoxyguanosine	Deoxyguanylate	DNA
Pyrimidines			
Cytosine	Cytidine	Cytidylate	RNA
	Deoxycytidine	Deoxycytidylate	DNA
Thymine	Thymidine or deoxythymidine	Thymidylate or deoxythymidylate	DNA
Uracil	Uridine	Uridylate	RNA

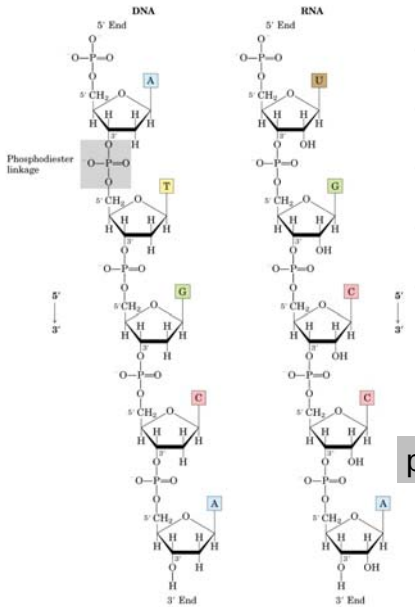
**Nucleoside* and *nucleotide* are generic terms that include both ribo- and deoxyribo- forms. Note that here ribonucleosides and ribonucleotides are designated simply as nucleosides and nucleotides (e.g., riboadenosine as adenosine), and deoxyribonucleosides and deoxyribonucleotides as deoxynucleosides and deoxynucleotides (e.g., deoxyriboadenosine as deoxyadenosine). Both forms of naming are acceptable, but the shortened names are more commonly used. Thymine is an exception; the name ribothymidine is used to describe its unusual occurrence in RNA.

Base Formula	Base X = H	Nucleoside X = ribose[a]	Nucleotide[b] X = ribose phosphate[a]
	Adenine Ade A	Adenosine Ado A	Adenylic acid Adenosine monophosphate AMP
	Guanine Gua G	Guanosine Guo G	Guanylic acid Guanosine monophosphate GMP
	Cytosine Cyt C	Cytidine Cyd C	Cytidylic acid Cytidine monophosphate CMP
	Uracil Ura U	Uridine Urd U	Uridylic acid Urdine monophosphate UMP
	Thymine Thy T	Deoxythymidine dThd dT	Deoxythymidylic acid Deoxythymidine monophosphate dTMP

[a] The presence of a 2′-deoxyribose unit in place of ribose, as occurs in DNA, is implied by the prefixes "deoxy" or "d." For example, the deoxynucleoside of adenine is deoxyadenosine or dA. However, for thymine-containing residues, which rarely occur in RNA, the prefix is redundant and may be dropped. The presence of a ribose unit may be explicitly implied by the prefixes "ribo" or "r." Thus the ribonucleotide of thymine is ribothymidine or rT.
[b] The position of the phosphate group in a nucleotide may be explicitly specified as in, for example, 3′-AMP and 5′- GMP.

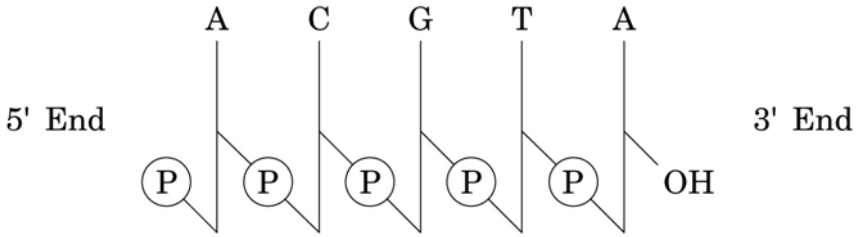
pdApdTpdGpdCpdA

Polynucleotides

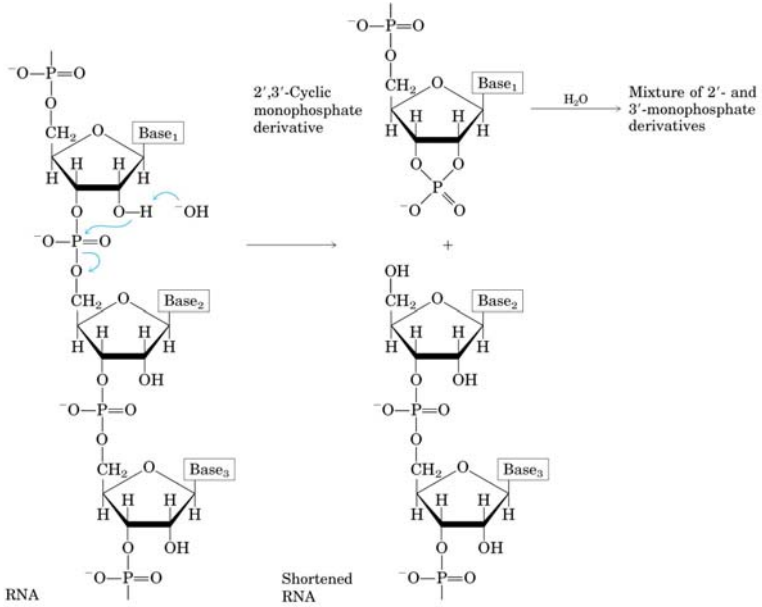


- Phosphodiester bonds
- Same orientation
- 5’ end and 3’ end
- Polyanion, pH 7
- Naming: 5’ → 3’

pUpGpCpCpA



RNA is less stable than DNA



DNA structure

- The linear sequence of nucleotides in a strand is the primary structure of DNA
- What about *higher order structure?* (2˚, 3˚, 4˚??)

Chargaff's rules

- DNA base composition differs from one species to another
- DNA base composition is identical in different individuals and different tissues of same species
- DNA base composition does not change with age, nutritional status, or changing environment
- Quantitation of base composition:

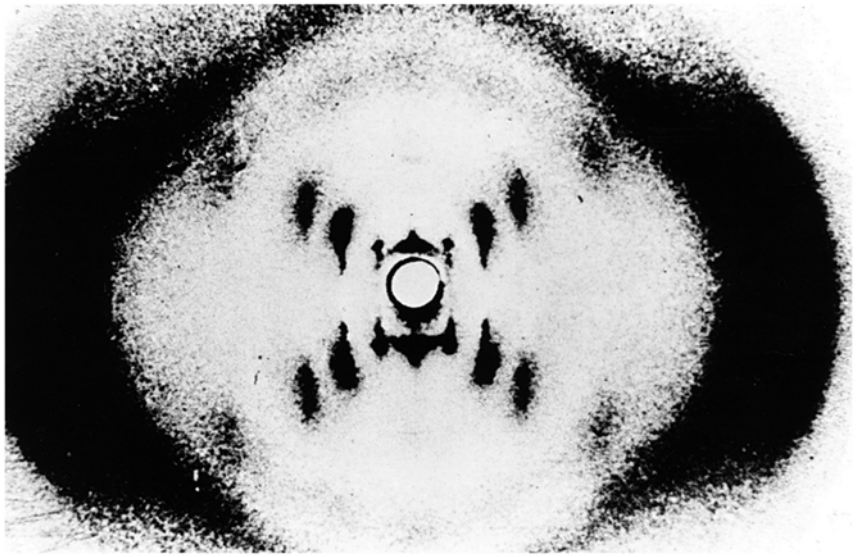
Species	%A	%G	%C	%T
man	30.9	19.9	19.8	29.4
Wheat	27.3	22.7	22.8	27.1
Virus	26.0	24.0	24.0	26.0

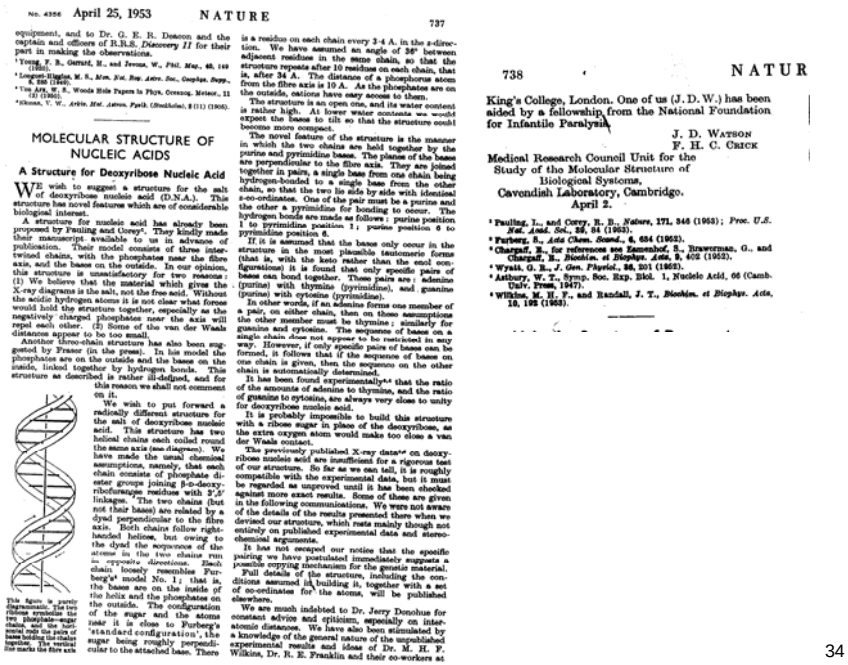
A = T
G = C

Σ purines = Σ pyrimidines

DNA structure

- Satisfy Chargaff's rules (A=T, G=C)
- Build on primary structure (linear polynucleotide, 5' → 3')
- Franklin and Wilkins: x-ray diffraction patterns: periodicities of 3.4 and 36 Å
- A mechanism for rapid duplication?
- Watson and Crick: molecular models





No. 4356 April 25, 1953 NATURE 737

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

We wish to suggest a structure for the salt of deoxyribose nucleic acid (D.N.A.). This structure has novel features which are of considerable biological interest.

A structure for nucleic acid has already been proposed by Pauling and Corey[1]. They kindly made their manuscript available to us in advance of publication. Their model consists of three inter-twined chains, with the phosphates near the fibre axis, and the bases on the outside. In our opinion, this structure is unsatisfactory for two reasons: (1) We believe that the material which gives the X-ray diagrams is the salt, not the free acid. Without the acidic hydrogen atoms it is not clear what forces would hold the structure together, especially as the negatively charged phosphates near the axis will repel each other. (2) Some of the van der Waals distances appear to be too small.

Another three-chain structure has also been suggested by Fraser (in the press). In his model the phosphates are on the outside and the bases on the inside, linked together by hydrogen bonds. This structure as described is rather ill-defined, and for this reason we shall not comment on it.

We wish to put forward a radically different structure for the salt of deoxyribose nucleic acid. This structure has two helical chains each coiled round the same axis (see diagram). We have made the usual chemical assumptions, namely, that each chain consists of phosphate diester groups joining β-D-deoxyribofuranose residues with 3',5' linkages. The two chains (but not their bases) are related by a dyad perpendicular to the fibre axis. Both chains follow right-handed helices, but owing to the dyad the sequences of the atoms in the two chains run in opposite directions. Each chain loosely resembles Furberg's[2] model No. 1; that is, the bases are on the inside of the helix and the phosphates on the outside. The configuration of the sugar and the atoms near it is close to Furberg's 'standard configuration', the sugar being roughly perpendicular to the attached base. There is a residue on each chain every 3·4 A. in the z-direction. We have assumed an angle of 36° between adjacent residues in the same chain, so that the structure repeats after 10 residues on each chain, that is, after 34 A. The distance of a phosphorus atom from the fibre axis is 10 A. As the phosphates are on the outside, cations have easy access to them.

The structure is an open one, and its water content is rather high. At lower water contents we would expect the bases to tilt so that the structure could become more compact.

The novel feature of the structure is the manner in which the two chains are held together by the purine and pyrimidine bases. The planes of the bases are perpendicular to the fibre axis. They are joined together in pairs, a single base from one chain being hydrogen-bonded to a single base from the other chain, so that the two lie side by side with identical z-co-ordinates. One of the pair must be a purine and the other a pyrimidine for bonding to occur. The hydrogen bonds are made as follows: purine position 1 to pyrimidine position 1; purine position 6 to pyrimidine position 6.

If it is assumed that the bases only occur in the structure in the most plausible tautomeric forms (that is, with the keto rather than the enol configurations) it is found that only specific pairs of bases can bond together. These pairs are: adenine (purine) with thymine (pyrimidine), and guanine (purine) with cytosine (pyrimidine).

In other words, if an adenine forms one member of a pair, on either chain, then on these assumptions the other member must be thymine; similarly for guanine and cytosine. The sequence of bases on a single chain does not appear to be restricted in any way. However, if only specific pairs of bases can be formed, it follows that if the sequence of bases on one chain is given, then the sequence on the other chain is automatically determined.

It has been found experimentally[3,4] that the ratio of the amounts of adenine to thymine, and the ratio of guanine to cytosine, are always very close to unity for deoxyribose nucleic acid.

It is probably impossible to build this structure with a ribose sugar in place of the deoxyribose, as the extra oxygen atom would make too close a van der Waals contact.

The previously published X-ray data[5,6] on deoxyribose nucleic acid are insufficient for a rigorous test of our structure. So far as we can tell, it is roughly compatible with the experimental data, but it must be regarded as unproved until it has been checked against more exact results. Some of these are given in the following communications. We were not aware of the details of the results presented there when we devised our structure, which rests mainly though not entirely on published experimental data and stereochemical arguments.

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

Full details of the structure, including the conditions assumed in building it, together with a set of co-ordinates for the atoms, will be published elsewhere.

We are much indebted to Dr. Jerry Donohue for constant advice and criticism, especially on interatomic distances. We have also been stimulated by a knowledge of the general nature of the unpublished experimental results and ideas of Dr. M. H. F. Wilkins, Dr. R. E. Franklin and their co-workers at

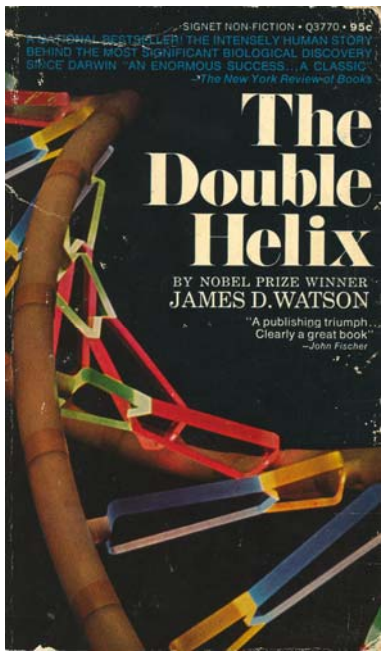
738 NATURE

King's College, London. One of us (J. D. W.) has been aided by a fellowship from the National Foundation for Infantile Paralysis.

J. D. WATSON
F. H. C. CRICK

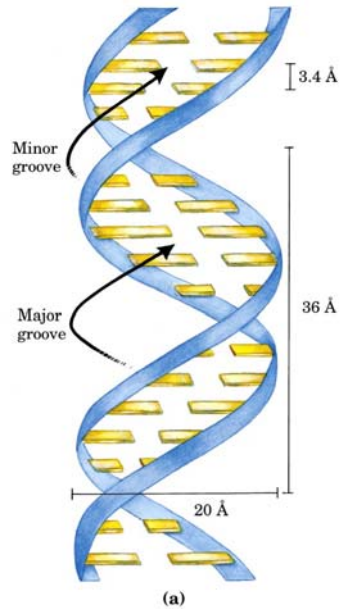
Medical Research Council Unit for the Study of the Molecular Structure of Biological Systems, Cavendish Laboratory, Cambridge. April 2.

[1] Pauling, L., and Corey, R. B., *Nature*, **171**, 346 (1953); *Proc. U.S. Nat. Acad. Sci.*, **39**, 84 (1953).
[2] Furberg, S., *Acta Chem. Scand.*, **6**, 634 (1952).
[3] Chargaff, E., for references see Zamenhof, S., Brawerman, G., and Chargaff, E., *Biochim. et Biophys. Acta*, **9**, 402 (1952).
[4] Wyatt, G. R., *J. Gen. Physiol.*, **36**, 201 (1952).
[5] Astbury, W. T., Symp. Soc. Exp. Biol. 1, Nucleic Acid, 66 (Cambridge Univ. Press, 1947).
[6] Wilkins, M. H. F., and Randall, J. T., *Biochim. et Biophys. Acta*, **10**, 192 (1953).

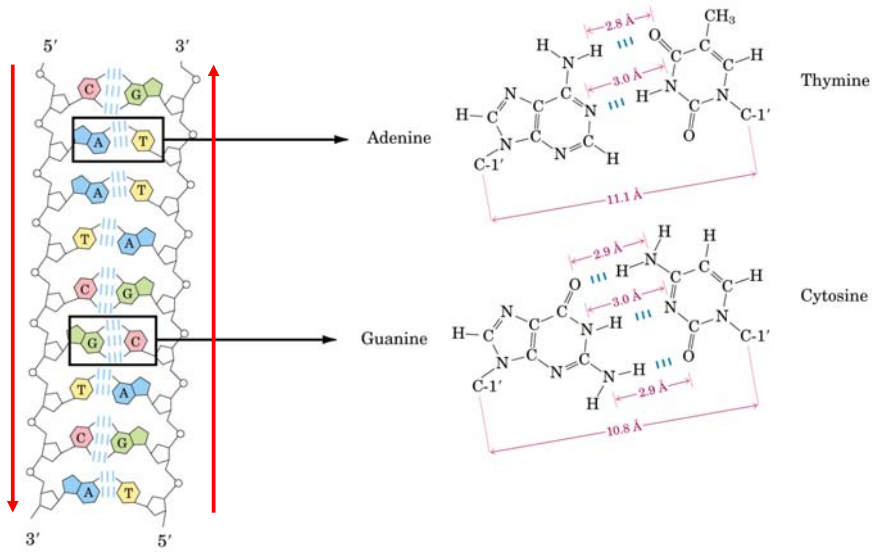


Watson-Crick DNA double helix (B-DNA)

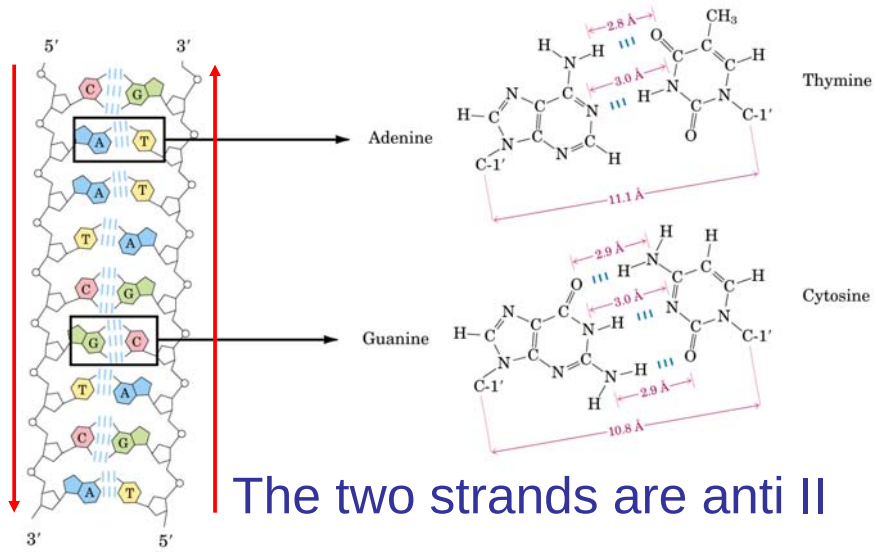
- Two helical chains, coiled about same axis. Right handed double helix
- Hydrophilic backbones of sugars and phosphates outside, hydrophobic bases inside
- Right handed helical staircase: rails are 2 anti II phosphate-sugar chains, rungs are purine and pyrimidine base pairs



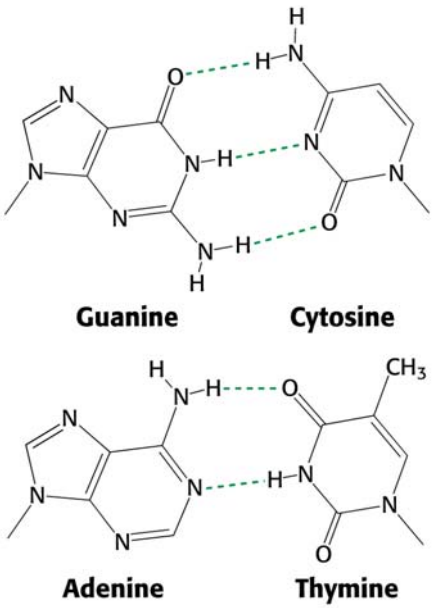
Base pairing in DNA:
A pairs with T, and G pairs with C



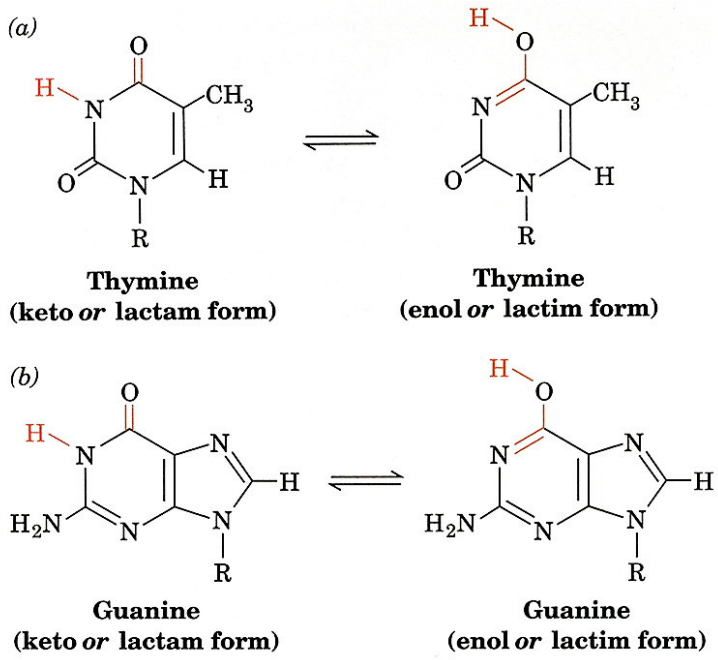
The two DNA strands (chains) are *complimentary*, not *identical*

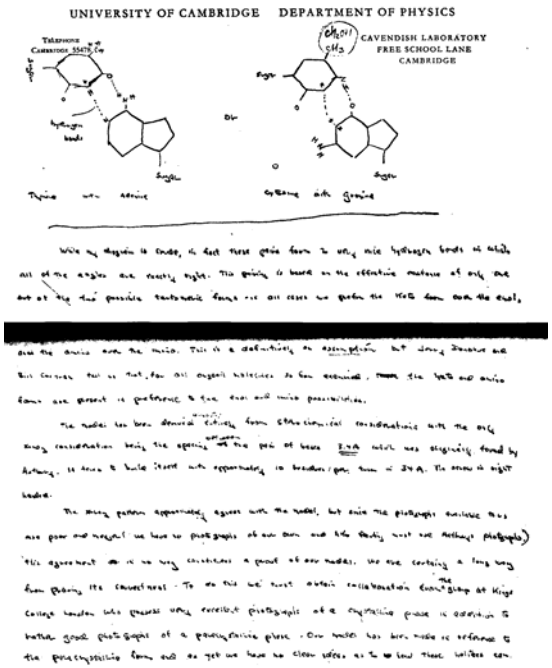


The two strands are anti II



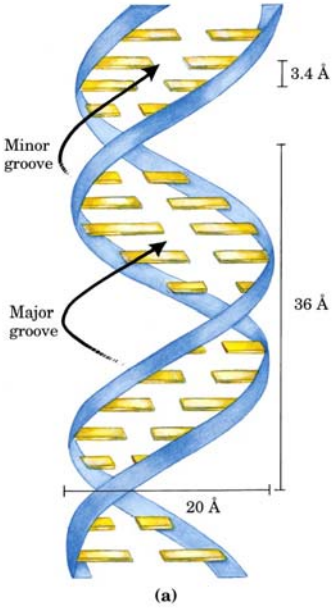
- C-G and A-T base pairs are superimposable.
- Drawing the correct base pairs required predicting the correct tautomeric forms used by DNA
- 2 H-bonds in A-T; 3 H-bonds in G-C
- The double helix will accommodate purine-pyrmidine or pyrimidine-purine





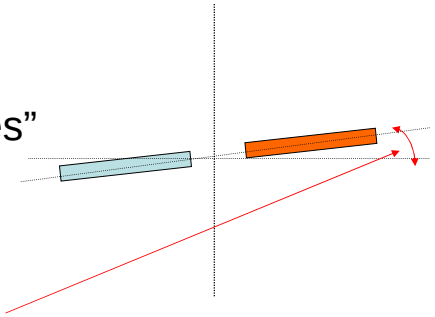
DNA structure

- Two periodicities:
  - 3.4 Å base-base distance
  - 36 Å per one turn of helix (pitch)
- Each base pair is rotated 34˚ relative to previous base pair
- 10.5 bp/turn
- Distance between strands is a constant 20 Å

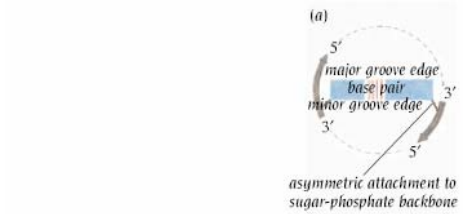


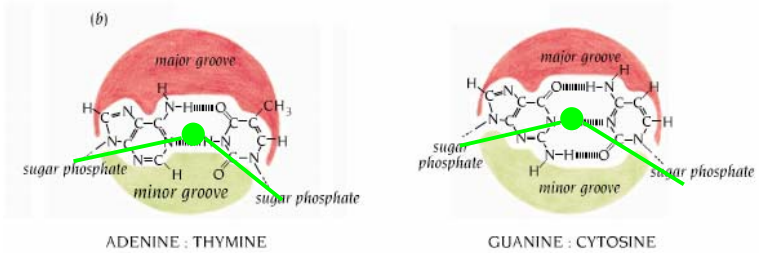
DNA structure

- Each base pair is rotated 34˚ relative to previous base pair
- Successive stacking of base pairs forms “grooves”
- Asymmetry of base pairs results in a major groove and a minor groove
- 6˚ base tilt to helix axis



DNA structure: the major and minor grooves

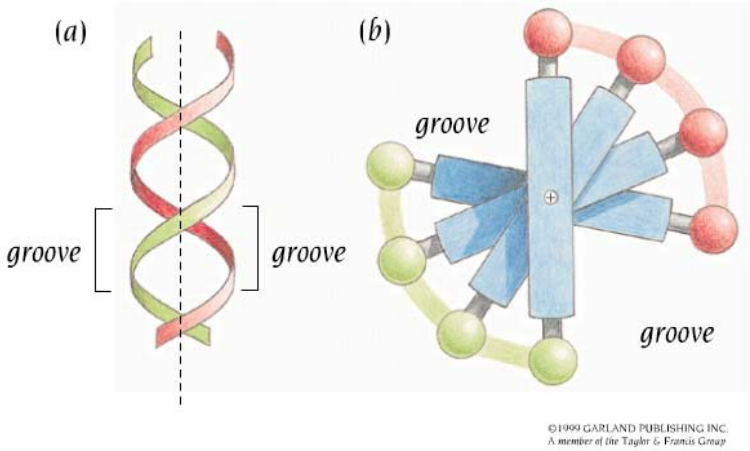




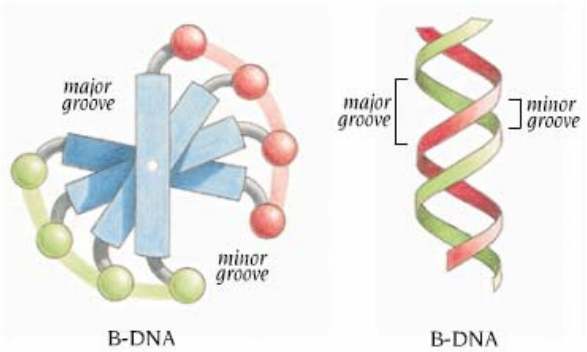
ADENINE : THYMINE GUANINE : CYTOSINE

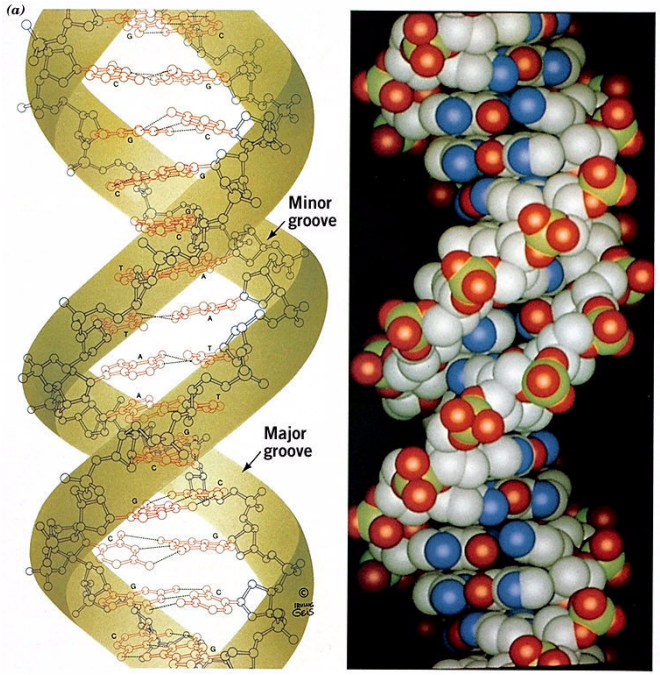
Note lines connecting C1’ carbon of one nucleotide, helix axis, and C1’ of the paired nucleotide

No asymmetry in base pairs: grooves are identical

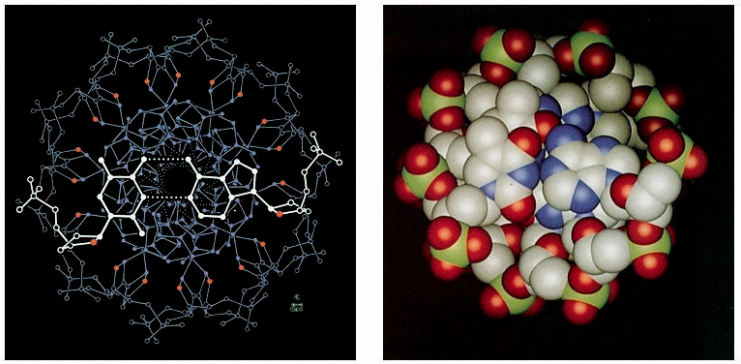


Asymmetry in base pairs





The inside of the helix is a solid core



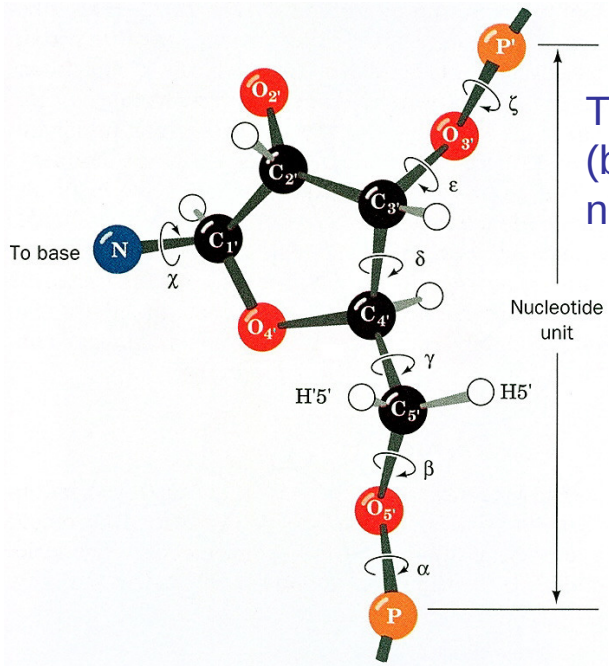
Forces that stabilize DNA double helix

- Base stacking (inside): dipole-dipole, VDW, and hydrophobic effect
- H-bonding (inside): A-T and G-C
- Ionic interactions (outside): positively charged counterion(s)
- H-bonding (outside): sugar, phosphate, water
- Protein-nucleic acid interactions

Conformational subtleties

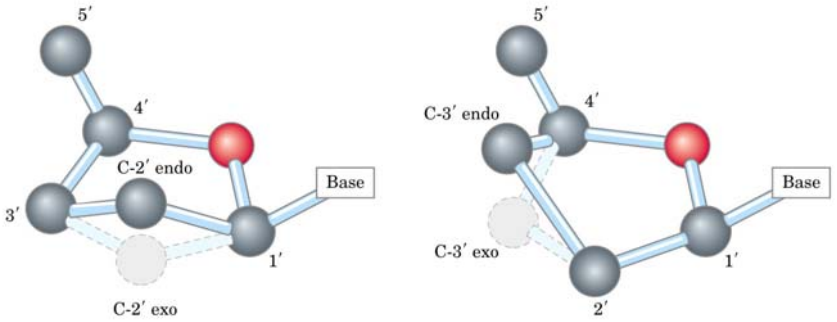
- Base tilt
- Sugar pucker
- N-glycosidic bond orientation

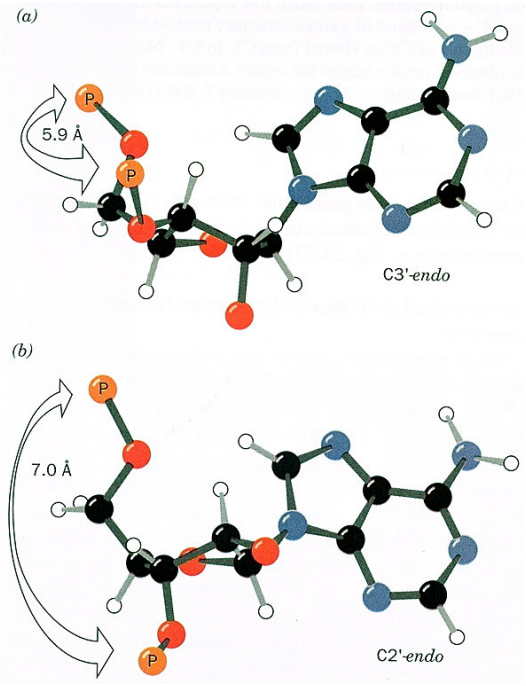
See kinemage exercise 12A, Exploring Molecular Structure



Torsion angles (bonds) in a nucleotide unit

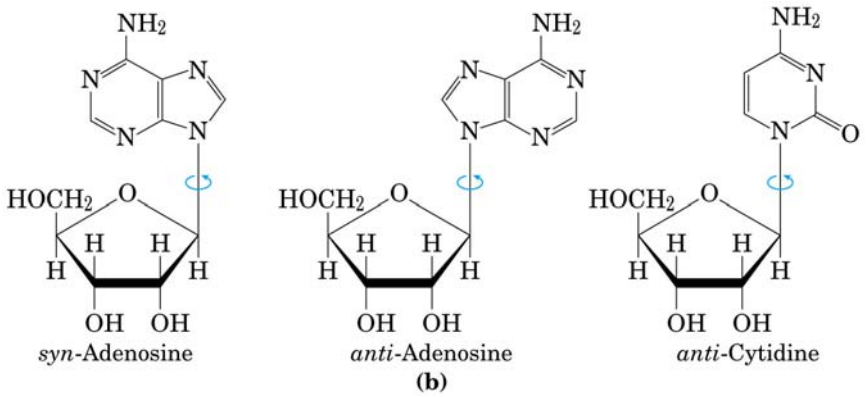
Sugar pucker: C2'-endo in B-DNA





Ramifications of sugar pucker on DNA structure. Note the difference in distance between C3'-Phosphate and C5'-Phosphate with the two different sugar puckers

N-glycosidic bond orientation

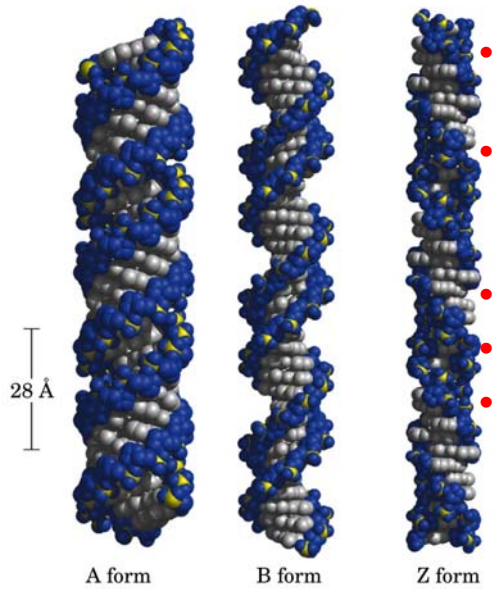


Always anti in B- and A-DNA

Summary of DNA structure

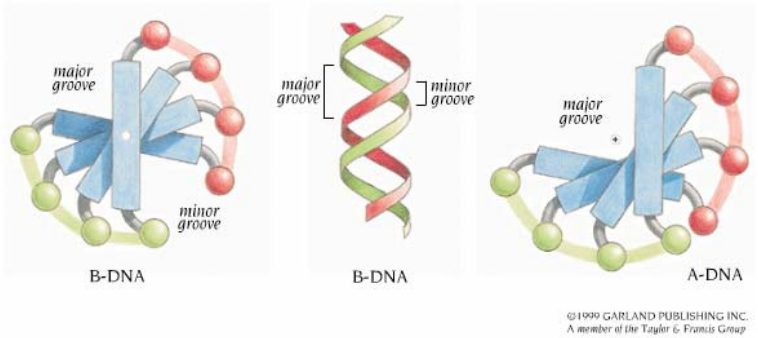
property	B-DNA
Helical sense	Right handed
Diameter	20 Å
Bp/turn	10.5
Helix twist/bp	34˚
Pitch	36 Å
Base tilt	6˚
Major groove	Wide, deep
Minor groove	Narrow, deep
Sugar pucker	C2'-endo
Glyco. bond	Anti
Helix rise/bp	3.4 Å

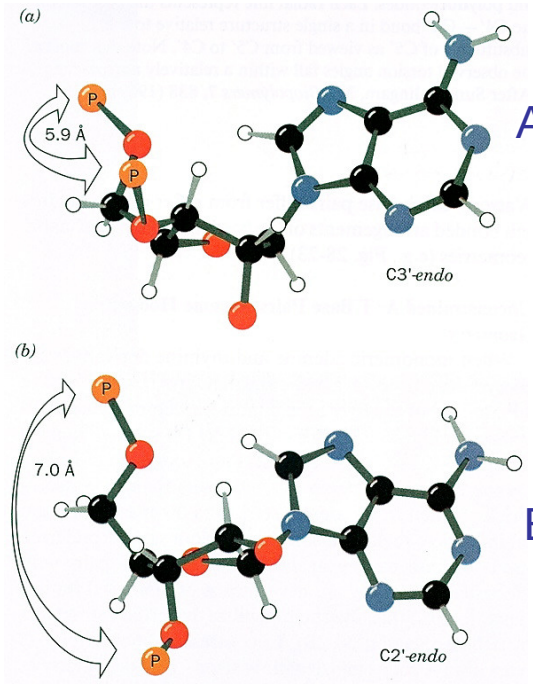
Other DNA structures form under certain circumstances



- A-DNA: the first form studied structurally
- Forms under dehydrated conditions
- Wider, flatter, R-hand
- Axial hole: bp offset
- C3'-endo sugar pucker

Asymmetry in base pairs: axial hole in A-DNA

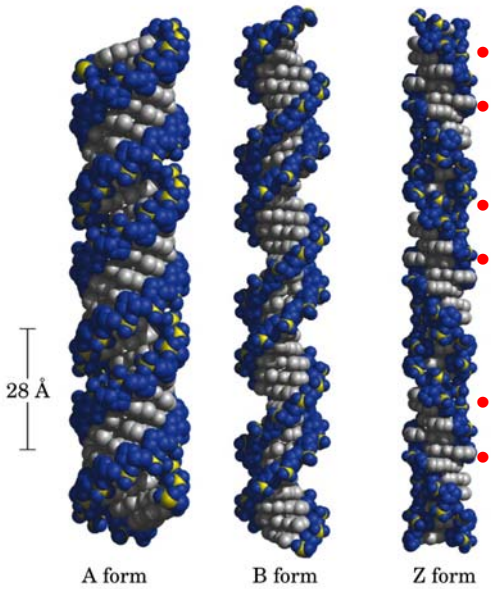




A-DNA

B-DNA

Other DNA structures form under certain circumstances

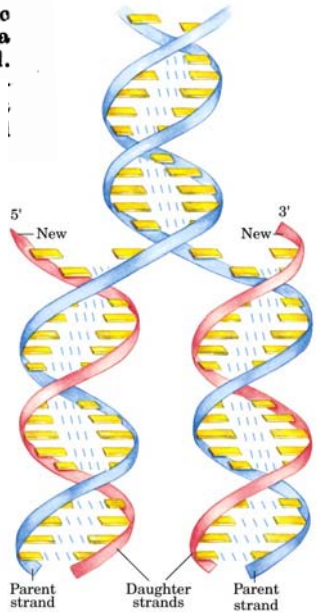


- Z-DNA
- L-handed double helix
- High salt
- alternating purine-pyrimidine repeating unit
- base pairs flipped
- Zig-zag backbone

Comparison of DNA structures (don't memorize A- and Z-parameters!!

property	B-DNA	A-DNA	Z-DNA
Helical sense	Right handed	Right handed	Left handed
Diameter	20 Å	26 Å	18 Å
Bp/turn	10.5	11	12
Helix twist/bp	34˚	33˚	30˚
Pitch	36 Å	28 Å	45 Å
Base tilt	6˚	20˚	7˚
Major groove	Wide, deep	Narrow, deep	flat
Minor groove	Narrow, deep	Wide, shallow	Narrow, deep
Sugar pucker	C2'-endo	C3'-endo	C2'- endo pyrimidines C3'-endo purines
Glyco. bond	Anti	Anti	Anti: pyrimidines Syn: purines
Helix rise/bp	3.4 Å	2.6 Å	3.7 Å

It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material.

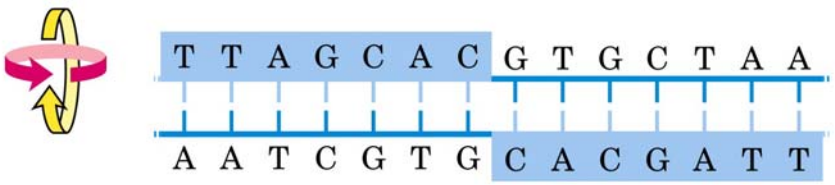


Nucleic acid structure

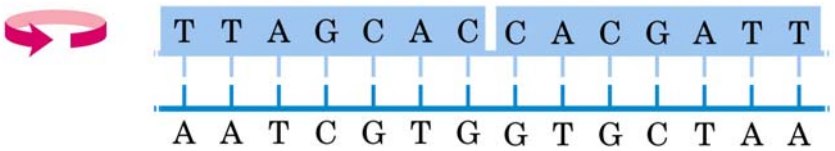
- Linear sequence of nucleotides in a strand is 1° structure
- The DNA double helix is 2° structure
- 2° structure is the highest order of DNA structure
- DNA almost always exists as double stranded double helix
- DNA-RNA hybrids form a double helix of "A" type
- RNA double helices have the "A" conformation
- RNA can form other types of 2° structure
- RNA can in some instances form 3° structure (tRNA, rRNA, snRNA)

Palindromic sequences: inverted repetition of base sequence with two-fold symmetry occurring over two strands of DNA

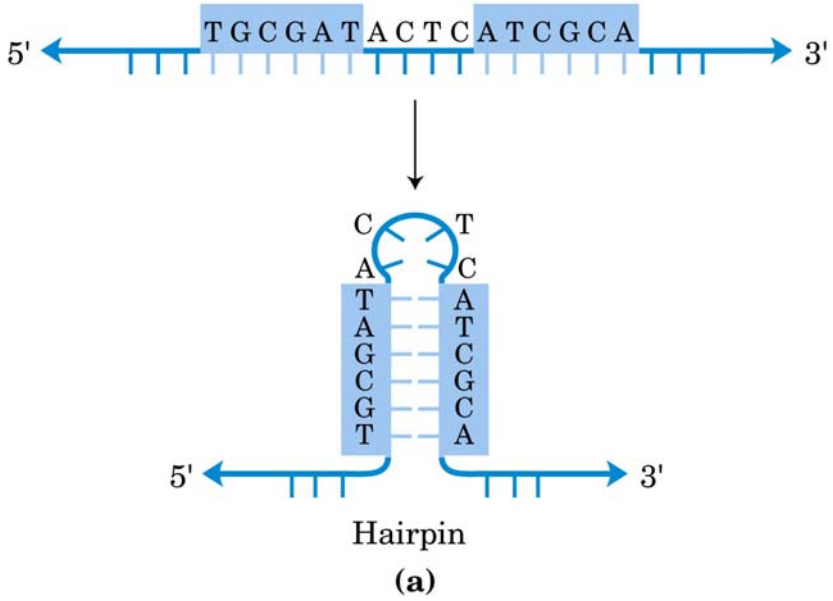
Palindrome



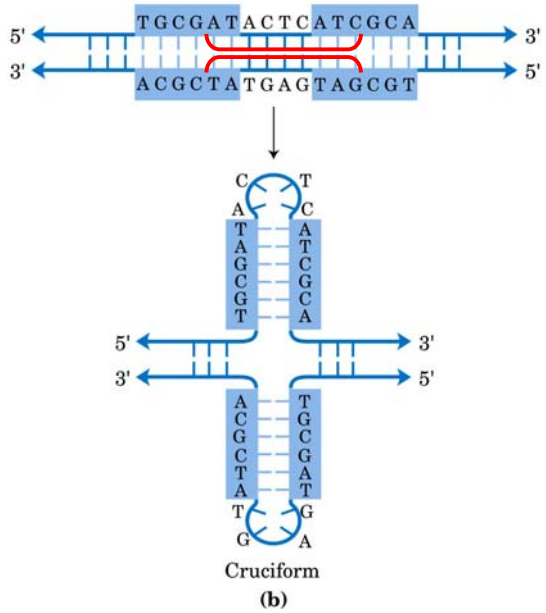
Mirror repeat



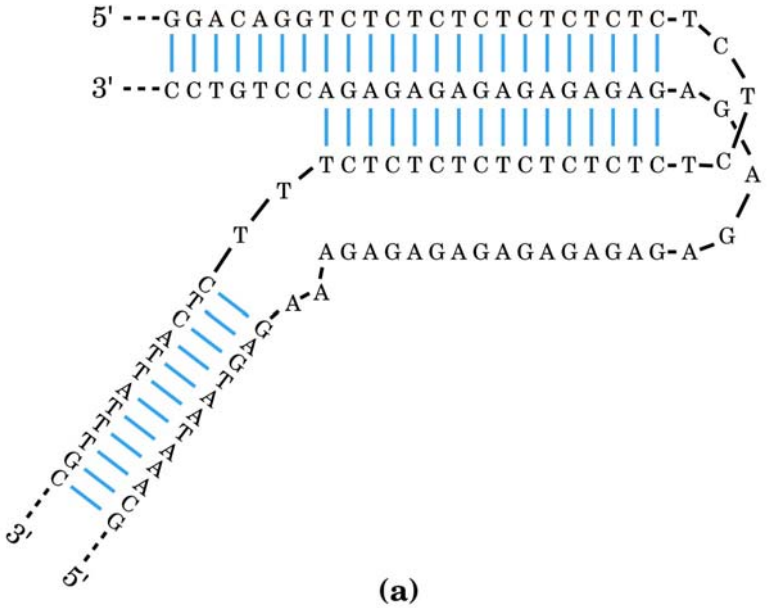
Self complimentarity within a single strand of DNA separated from its palindromic partner

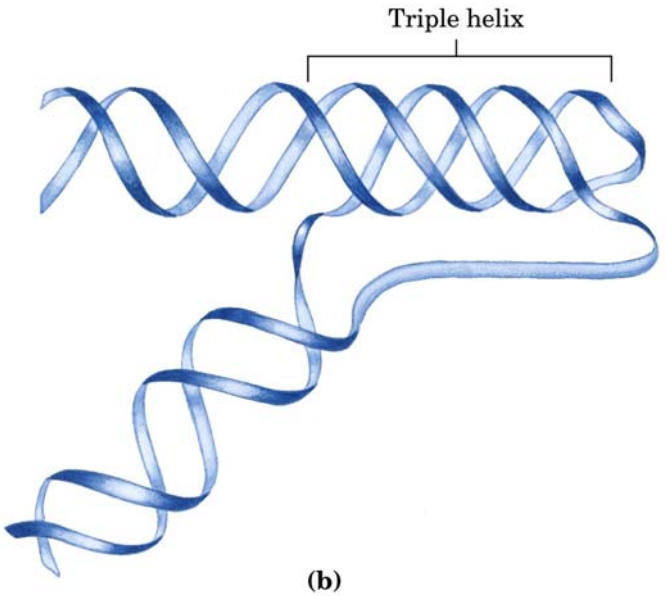


Formation of a cruciform between two strands having a palindrome



A triple helix can form under some circumstances

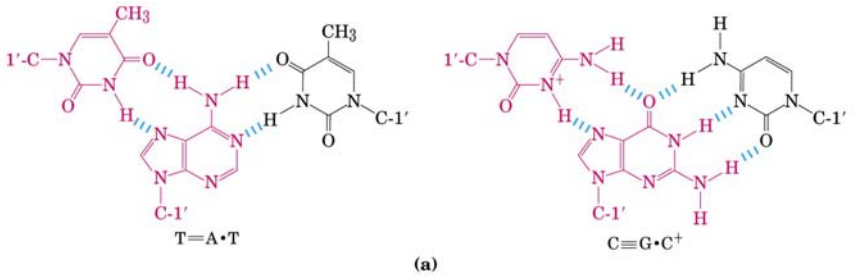




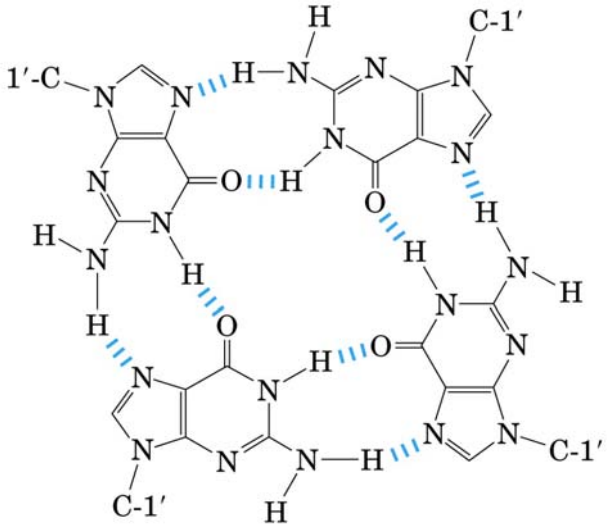
Versatility of RNA structure

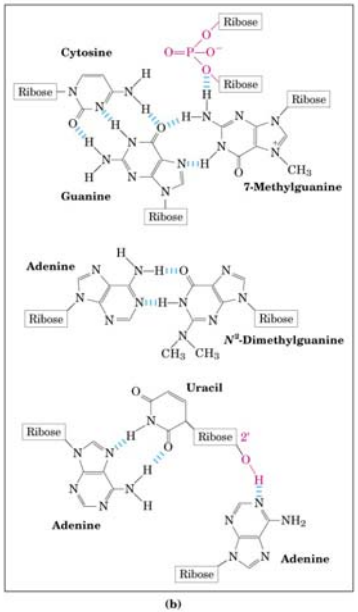
- Usually single stranded (not found with a complimentary partner strand)
- When double standed, has “A” conformation
- ssRNA can adopt a variety of 2˚ structures due to self-complimentary regions
- Non Watson-Crick base pairs
- Modified and different nucleotide bases
- 2’-OH can participate in H-bonding
- May form tertiary structure

Examples of nonstandard base pairing that can occur in RNA

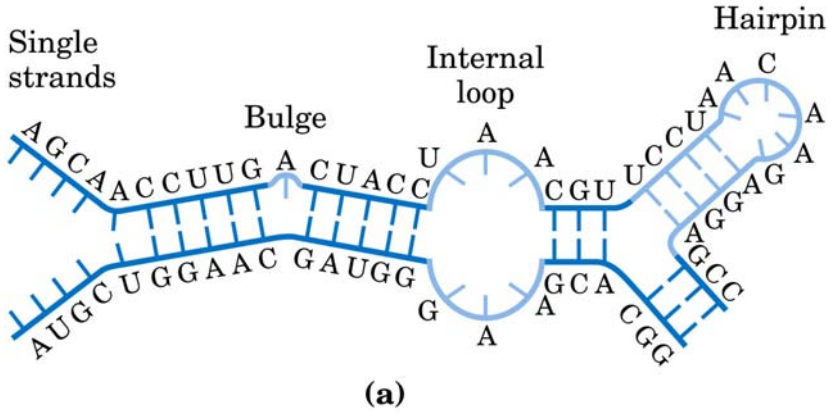


Examples of nonstandard base pairing that can occur in RNA

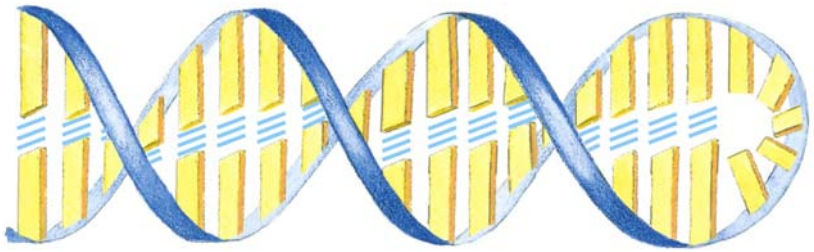




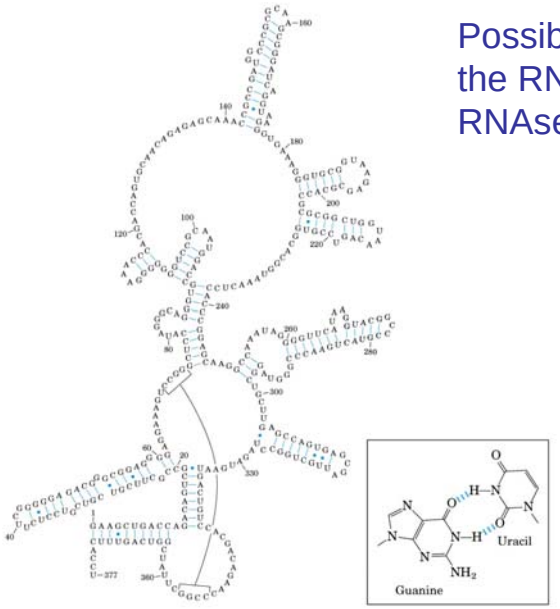
Association of regions of ssRNA to form secondary structure



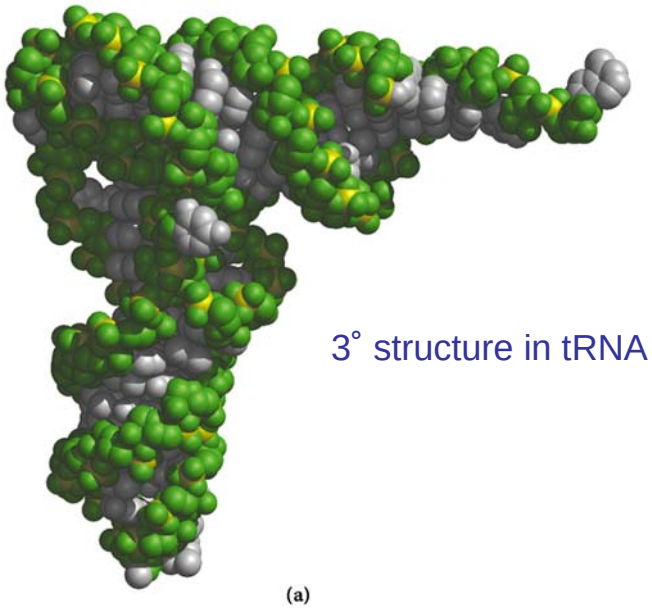
RNA in a hairpin double helix will be in the “A” conformation



Hairpin double helix
(b)

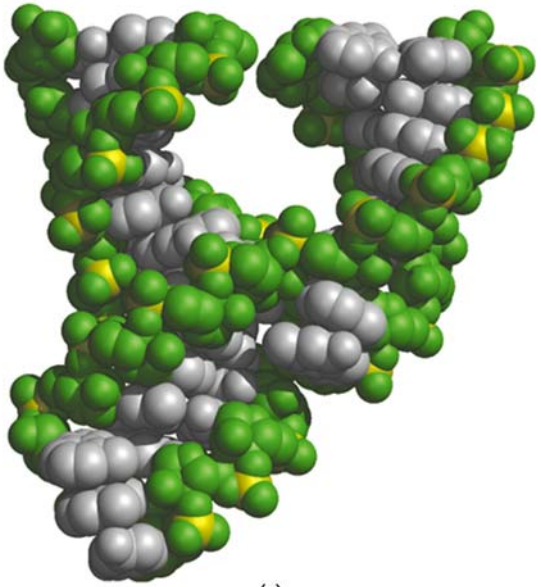


Possible 2° structure in the RNA component of RNAseP

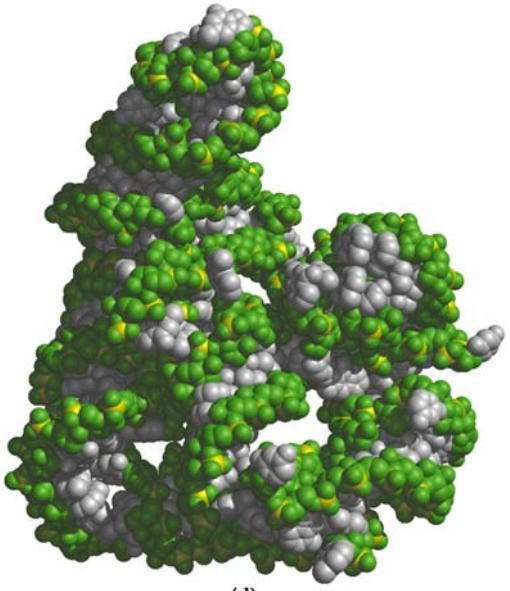


(a)

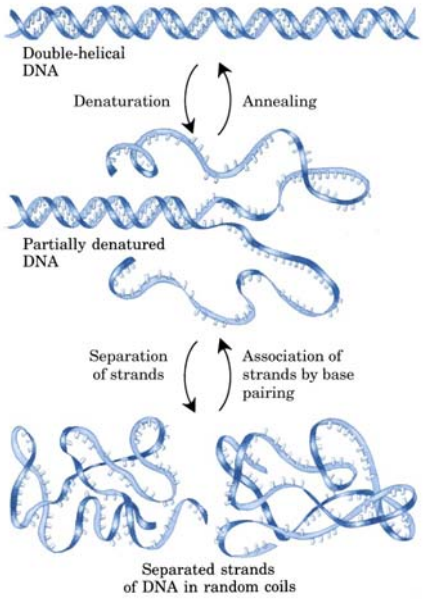
3° structure in tRNA

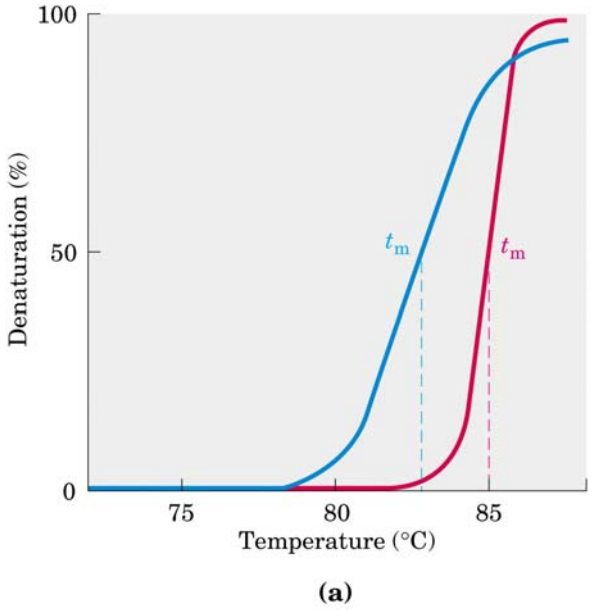


(c)

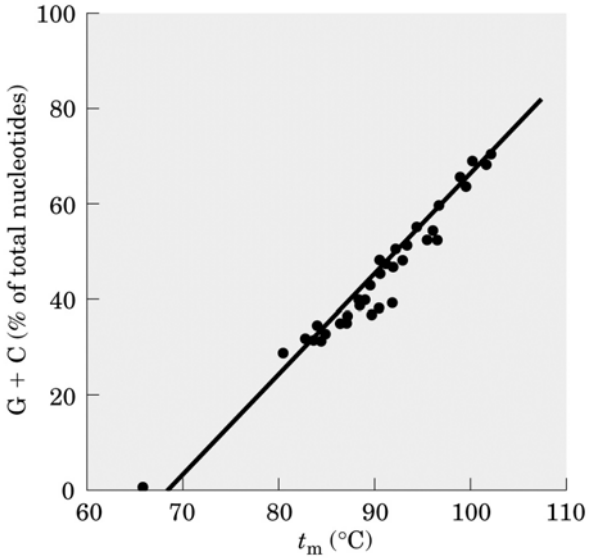


(d)

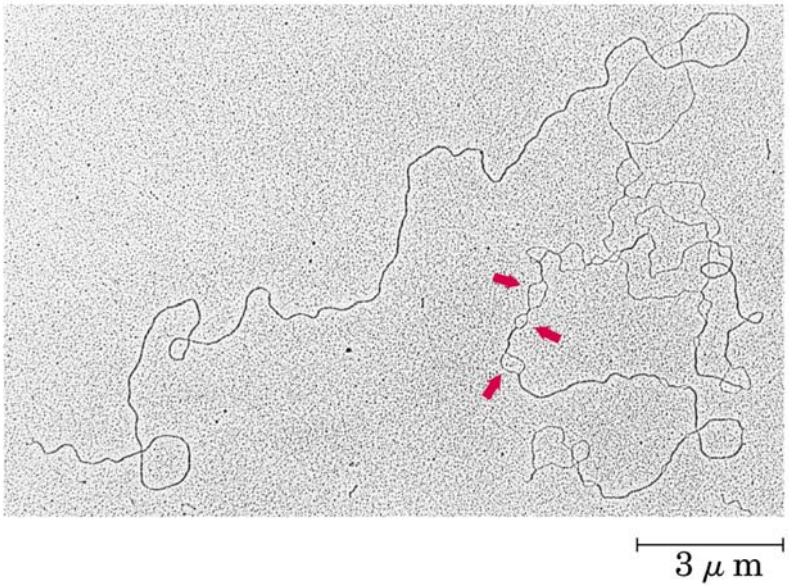

Double-helical DNA
Denaturation
Annealing
Partially denatured DNA
Separation of strands
Association of strands by base pairing
Separated strands of DNA in random coils

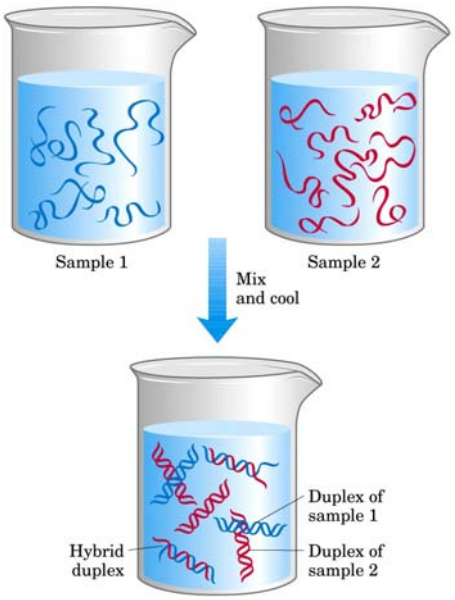

100
50
0
Denaturation (%)
t_m
t_m
75
80
85
Temperature (°C)

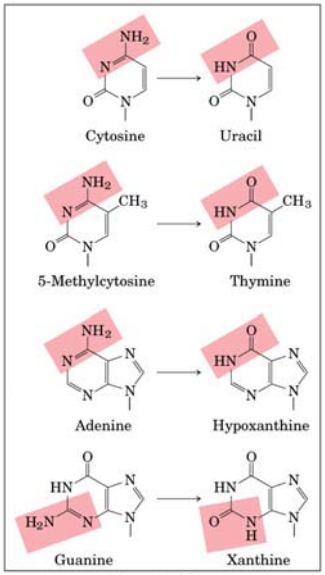
(a)


100
80
60
40
20
0
G + C (% of total nucleotides)
60
70
80
90
100
110
t_m (°C)

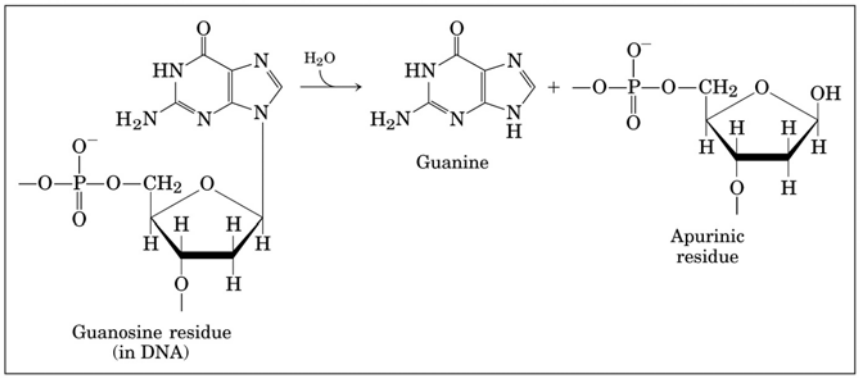
(b)



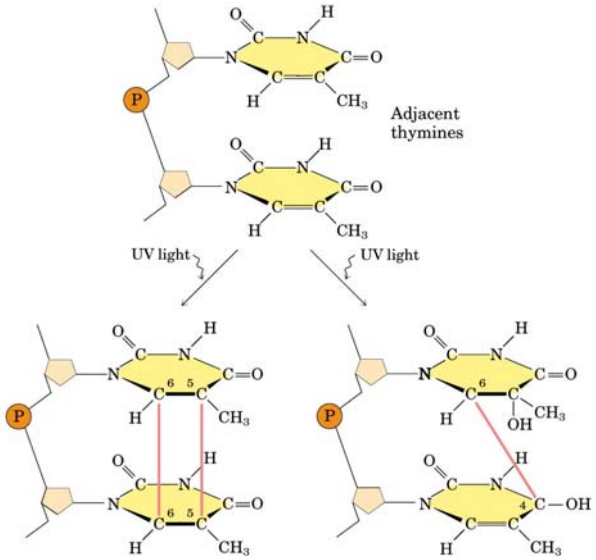


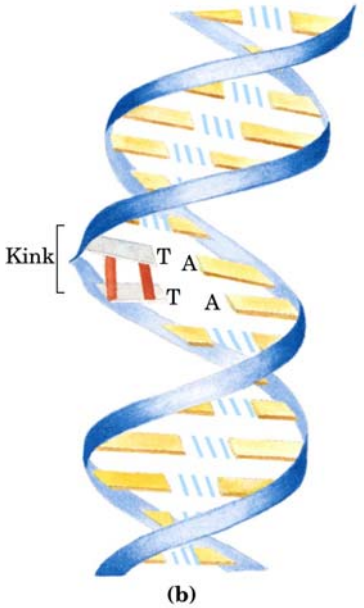


Deamination
(a)



Depurination
(b)

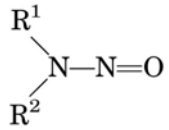




Deaminating agents

$NaNO_2$
Sodium nitrite

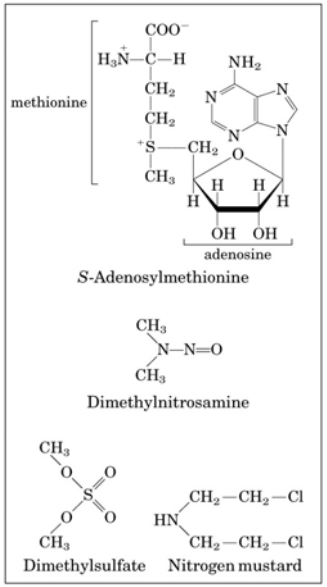
$NaNO_3$
Sodium nitrate

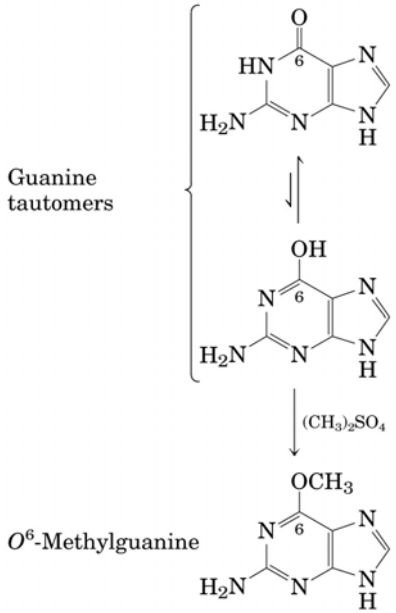


Nitrosamine

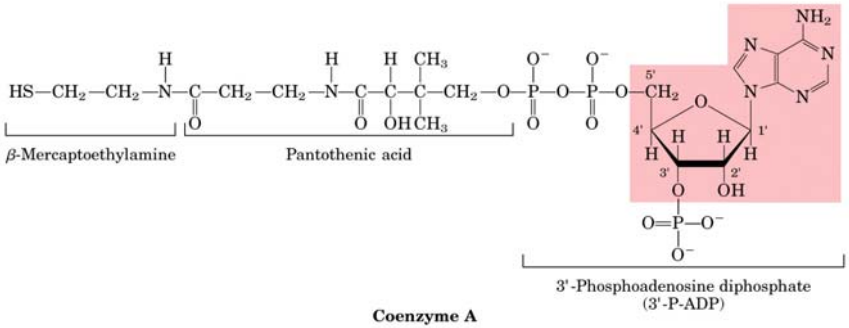
Nitrous acid precursors (a)

Alkylating agents



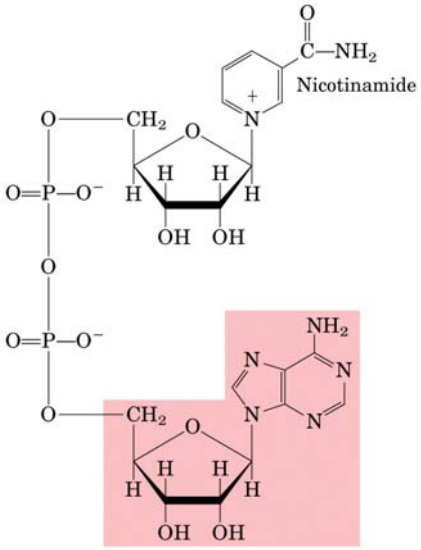


Nucleotides in enzymatic cofactors



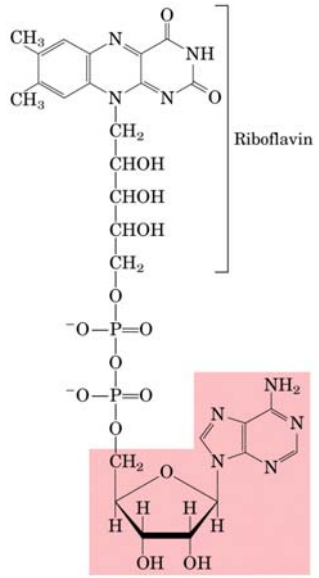
Coenzyme A

Nucleotides in enzymatic cofactors



Nicotinamide adenine dinucleotide (NAD^+)

Nucleotides in enzymatic cofactors



Flavin adenine dinucleotide (FAD)