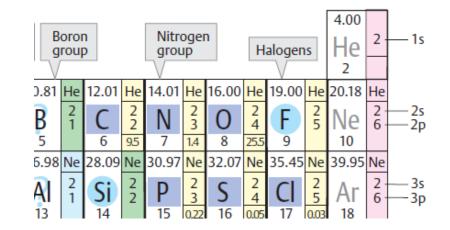
### Chemical properties, structure and interactions of nucleic acids

- DNA structure basics, Watson-Crick and Hoogsteen base pairing, double helix, alternative structures, DNA superhelicity
- Chemical reactivity of DNA, DNA damage, chemical modification of DNA as a tool for structure/interactions studies
- Non-covalent interactions of DNA, outer-sphere electrostatic interactions, groove binding, intercalation, fundamentals of DNA-protein interactions
- Enzymatic processing of nucleic acids, application of enzymes in structure/interactions studies
- Molecular principles of epigenetic regulations.
- Optical spectroscopic methods general introduction
- Principles of circular dichroic (CD) spectroscopy
- Advantages and drawbacks of the use of CD spectroscopy to proteins and nucleic acids studies
- Characteristic CD spectra of particular nucleic acids types
- Structural properties of nucleic acids fresh findings
- Electrochemistry of nucleic acids, electrochemical methods general introduction, electrochemical activity of DNA, DNA structure at electrically charged surfaces, electrochemical sensing of DNA damage, modification and nucleotide sequences.
- Electrochemistry of proteins basics and applications

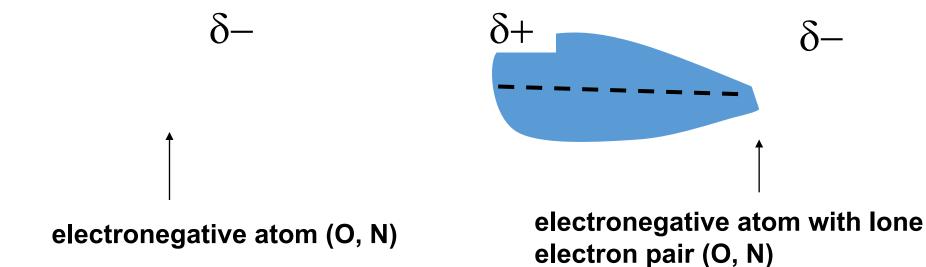
# You and chemistry? ③

# Some basic terms

- hydrogen bond
- ionic interactions
- hydrophobic interactions, stacking
- ester bond
- N-glycosidic bonds
- condensation, hydrolysis
- oxidation, reduction
- electrophile, nucleophile
- tautomers, enol-keto, amino-imino

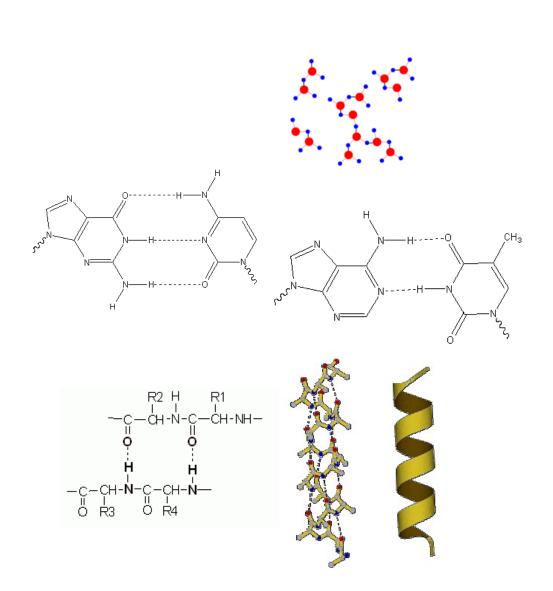


# Hydrogen bond

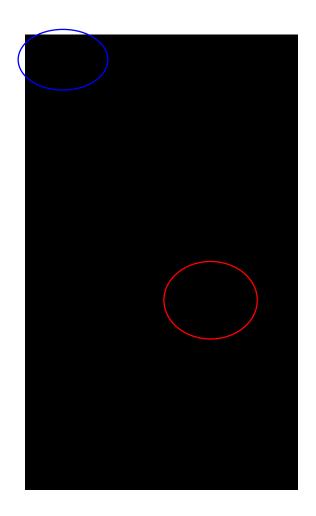


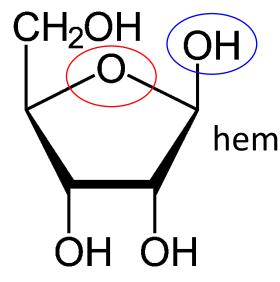
# Hydrogen bond

- crucial importance for biology
- properties of water
- nucleobase pairing
- protein structures
- protein-DNA interactions
- many others



# N-glycosidic bond





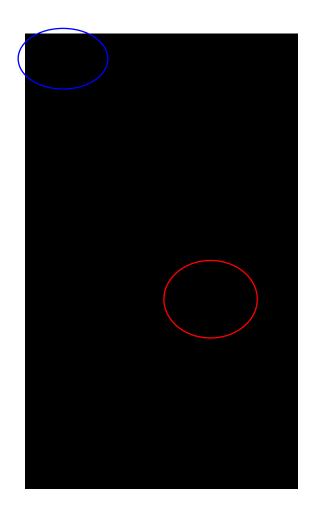
anomers: α (hemiacetal hydroxyl "down") β ("up" as here)

### hemiacetal

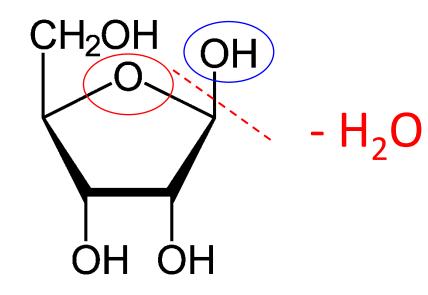
(more exactly: in  $\beta$  anomer the hemiacetal hydroxyl is on the same side of the furanose ring as the -CH<sub>2</sub>OH group)

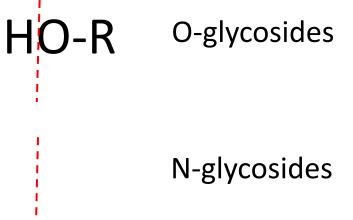
 $\beta$ -D-ribofuranose

# N-glycosidic bond



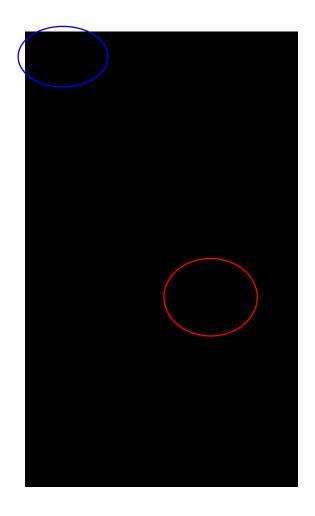
substitution of hemiacetal OH (condensation reaction)



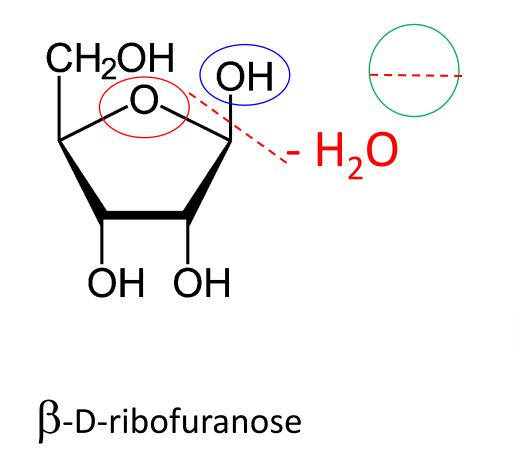


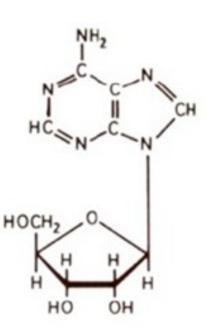
 $\beta$ -D-ribofuranose

# N-glycosidic bond



nucleoside formation

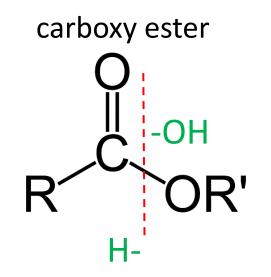


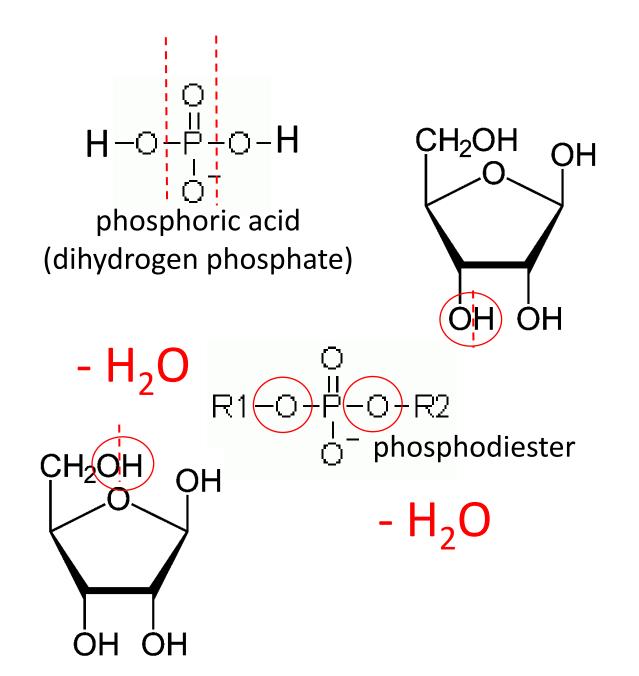


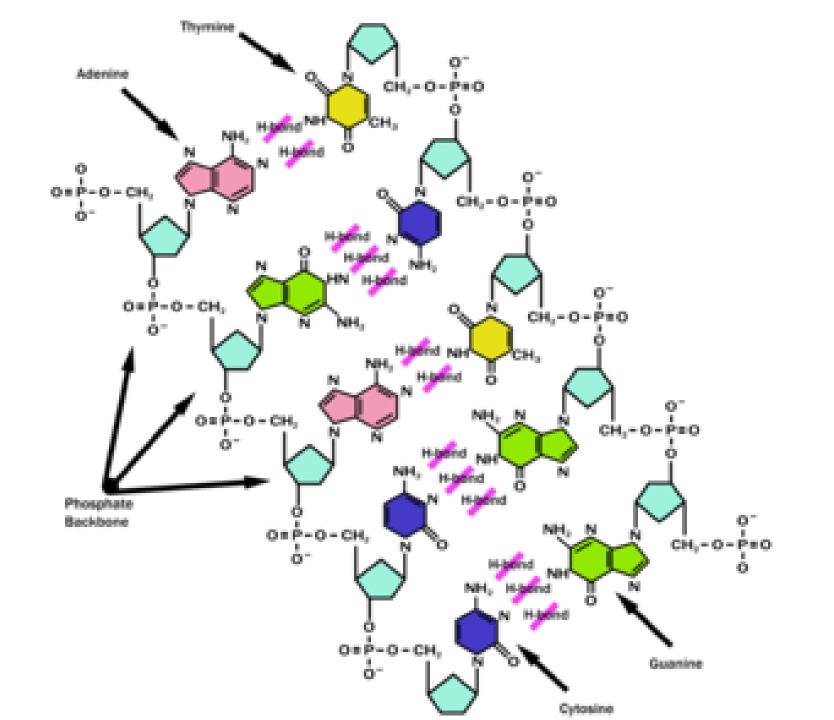
nucleoside

# Ester bonds

- esters: products of condensation of acids with alcohols
- substitution of –OH of the acid with –OR

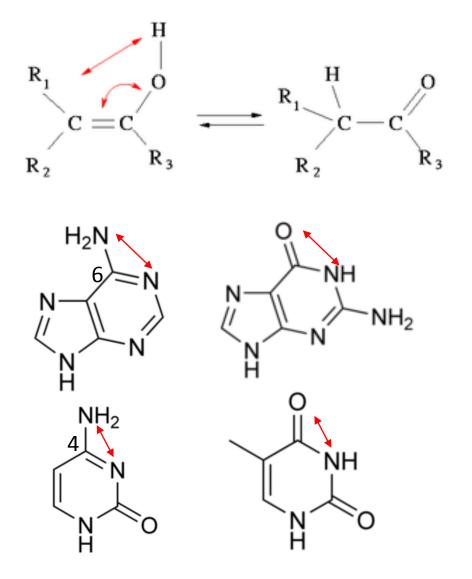




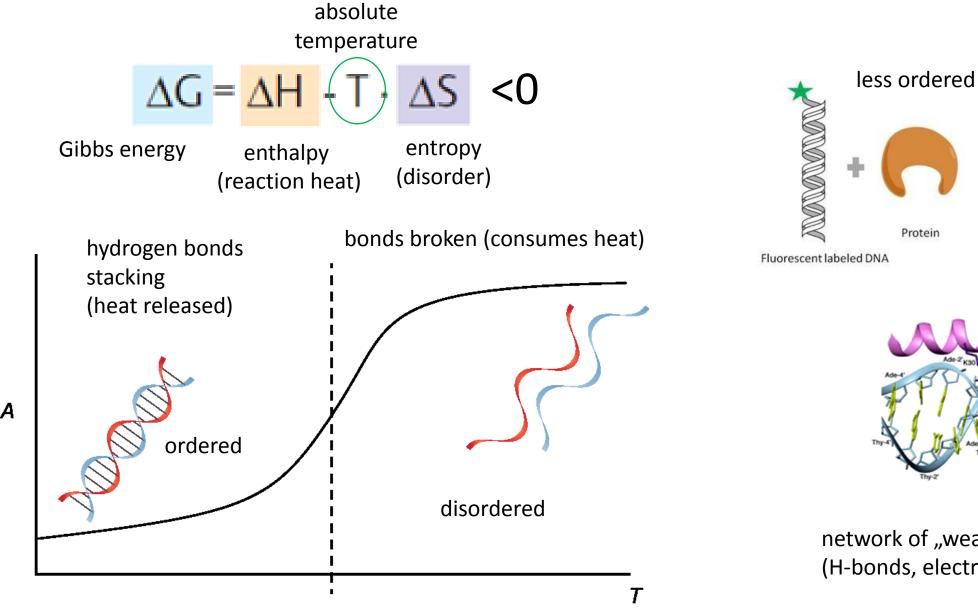


# Tautomers

- isomers enol-keto, amino-imino
- double bond switch plus hydrogen/proton migration
- in nucleobases: critical effect on pairing properties
- 6-substituents in purines + 4substituents in pyrimidines: oxygenous=keto, nitrogenous=amino
- hydrogen yes or not on the neighboring ring nitrogen
- relation to chemical mutagenesis



## Energetics of interactions (including structure)



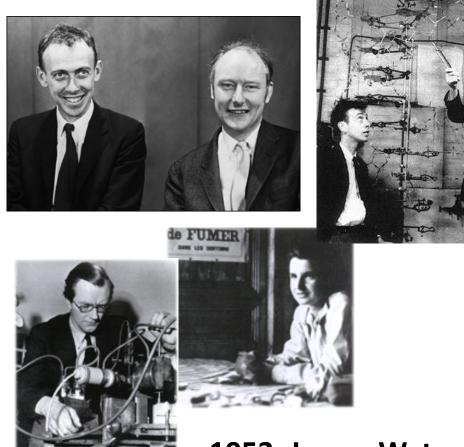
 $T_{\rm m}$ 

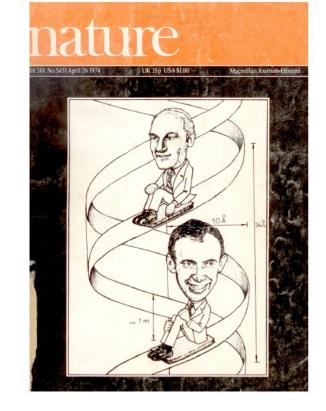
+ Control of the second second

ordered

network of "weak" bonds– released heat (H-bonds, electrostatic...)

# **DNA** structure





1953: James Watson, Francis Crick, Rosalind Franklin, Maurice Wilkins: the DNA double helix

1962: Nobel Prize (JW, FC, MW)

basic principle of the preservation, transfer and expression of genetic information explained

### Mendel 1864: "elements of heredity", Mendel laws





Mendel's Medal, Moravian Museum, Brno





Teachers of Brno gymnasium (High School)

Abbot G. Mendel

G J MENDEL, priest, teacher, scientist and abbot in BRNO

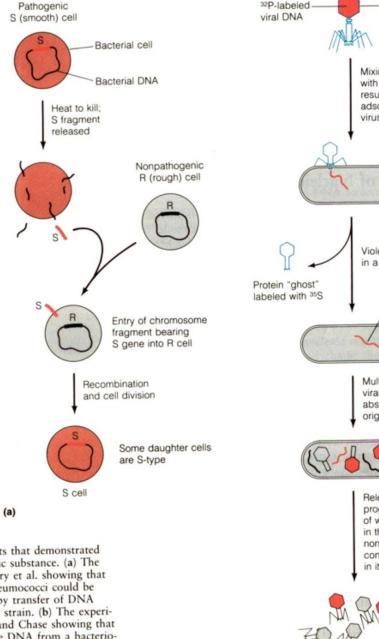
### Miescher 1871: discovered "nuclein", a substance occuring in cell nuclei

Fig. 5. Glass vial containing nuclein isolated from salmon sperm by Friedrich Miescher while working at the University of Basel. The faded label reads Nuclein aus Lachssperma, F. Miescher (Nuclein from salmon sperm, F. Miescher). Possession of the Interfakult-res Institut fqr Biochemie (Interfacultary Institute for Biochemistry), University of Tubingen, Germany; photography by Alfons Renz, University of Tubingen.





# DNA is the genetic material (1944 Avery, 1952 Hershey)



(b)

35S-labeled protein coat Mixing of virus with host cells results in adsorption of virus to host cell E. coli cell Violent agitation in a mixer Viral DNA labeled with 32P Multiplication of viral DNA in the absence of the original protein coat Release of new progeny viruses, some of which contain 32P in their DNA and none of which contains 35S in its coat

(a) 1944: Oswald T. Avery, Colin MacLeod, and Maclyn McCarty demonstrate that Griffith's transforming principle is not a protein, but rather DNA, suggesting that DNA may function as the genetic material

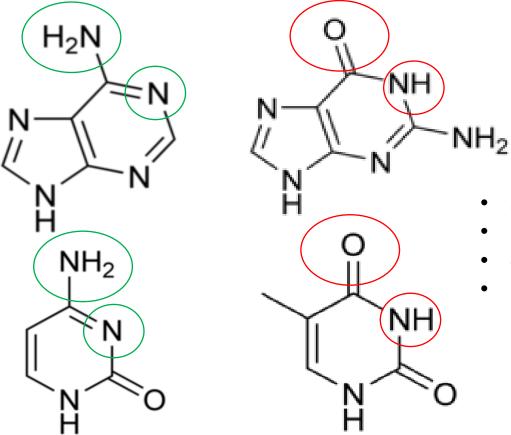
(b) 1952: Alfred Hershey and Martha Chase use viruses (bacteriophage T2) to confirm DNA as the genetic material by demonstrating that during infection viral DNA enters the bacteria while the viral proteins do not and that this DNA can be found in progeny virus particles.

#### Figure 4.8

Crucial experiments that demonstrated DNA as the genetic substance. (a) The experiment of Avery et al. showing that nonpathogenic pneumococci could be made pathogenic by transfer of DNA from a pathogenic strain. (b) The experiment of Hershey and Chase showing that it is transfer of the DNA from a bacteriophage to a bacterium that gives rise to new bacteriophages.

### **Chargaff's Rules**

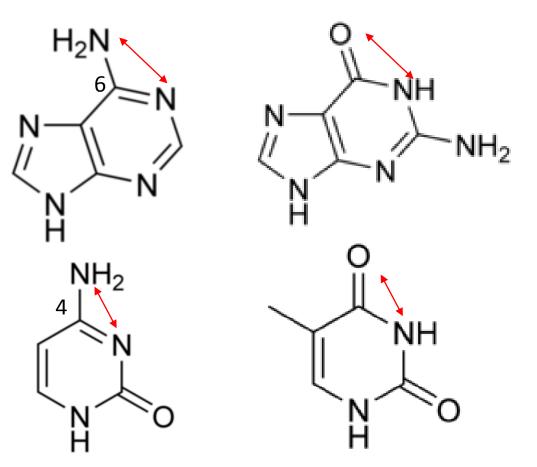
Tetranucleotide hypothesis originated in 1906: DNA is a "statistical tetranucleotide". During the 1950's E. Chargaff showed a number of DNAs, which differ in their base content. Chargaff's rules: 1. amino residues = keto-residues; in another expression A+C = G+T; 2. py = pu; C+T = G+A 3. A/T = G/C = 1 (consequence of combining equations 1 and 2)



- amino pairs with keto
- purine pairs with pyrimidine
- consequently, A pairs with T and G with C
- nitrogen in the ring: donor or acceptor of H bond

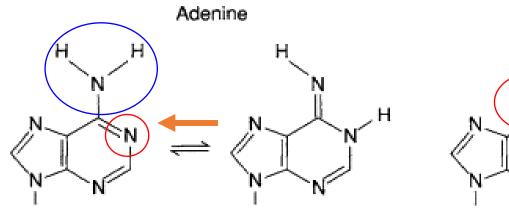
# Tautomers

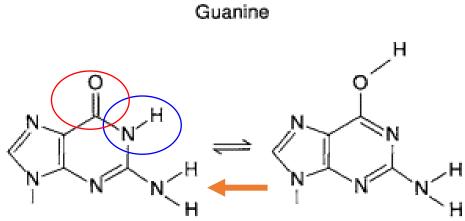
- isomers enol-keto, amino-imino
- double bond switch plus hydrogen/proton migration
- in nucleobases: critical effect on pairing properties
- 6-substituents in purines + 4-substituents in pyrimidines: oxygenous=keto, nitrogenous=amino
- hydrogen yes or not on the neighboring ring nitrogen

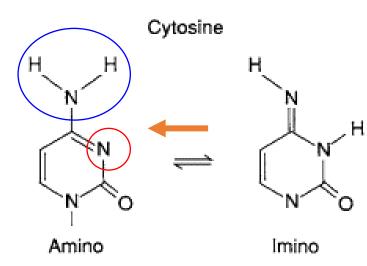


- imino behaves as keto
- enol behaves as amino

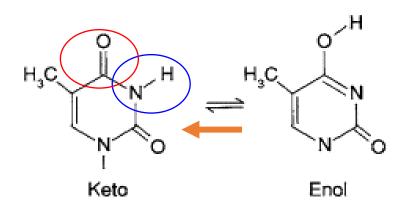
## **Tautomers**

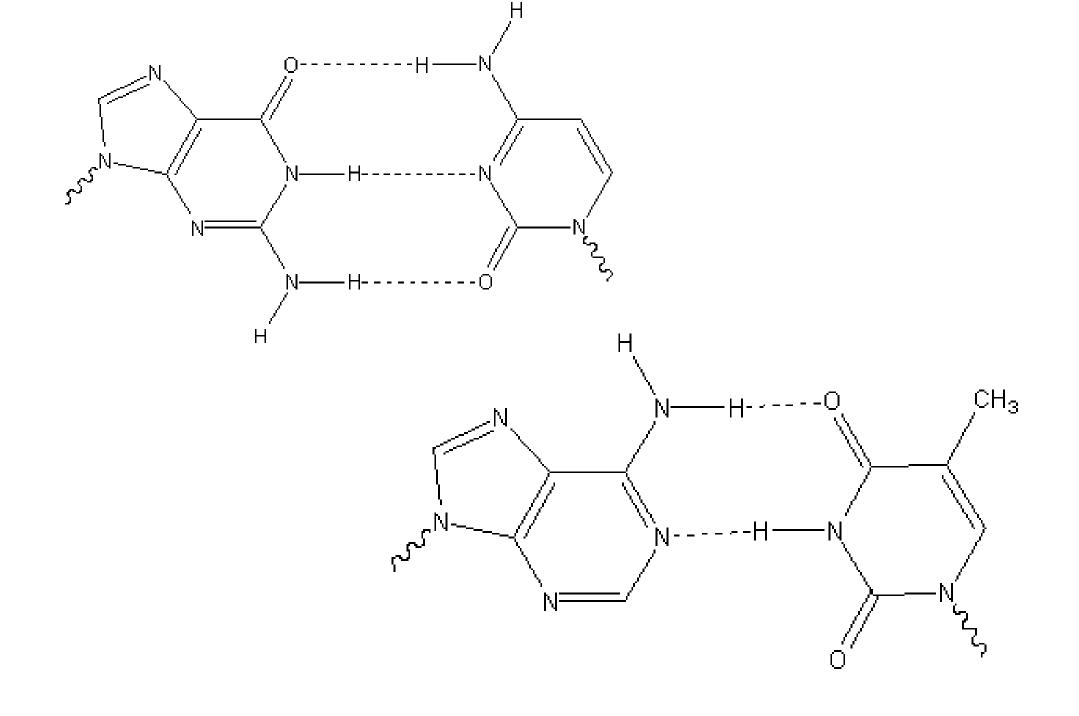






Thymine





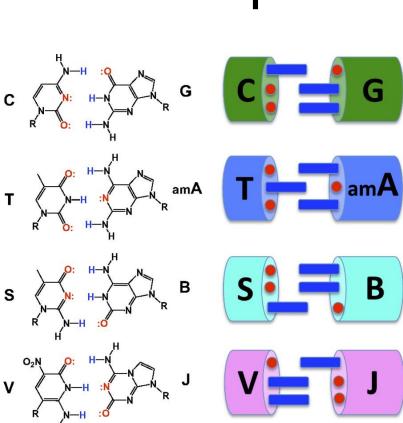
# Unnatural base pairs to expand genetic code

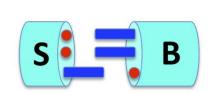
S. Benner: "AEGIS" (Artificially Expanded Genetic Information System)

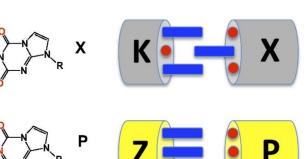
nucleobase analogues with permutated hydrogen bonding donor/akceptor features

Κ

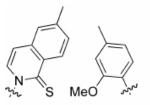
Ζ

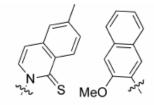






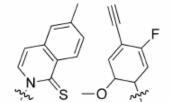
**F.** Romesberg hydrophobic base pairs no hydrogen bonding! shape complementarity only

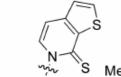


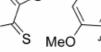


d5SICS-dMMO2

d5SICS-dNaM







d5SICS-dFEMO

dTPT3-dNaM

### X-RAY FIBER ANALYSIS OF DNA

represented the main evidence for the Watson-Crick double helix model This method enabled analysis of high-molecular DNA, but provided only few basic parameters of the helix. such as

Exo

21. 1 - 35

distance between base pairs

number of base residues per turn

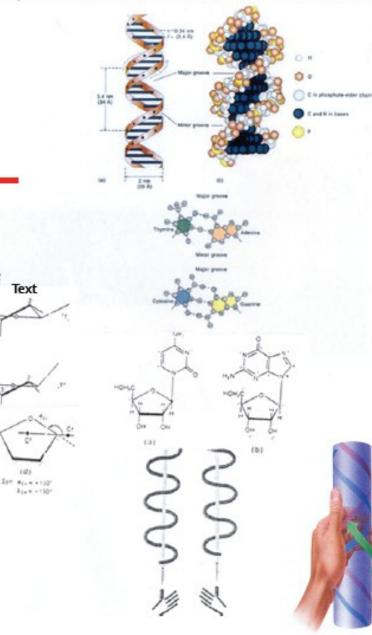
Further data were derived from model building considering the laws of structural chemistry

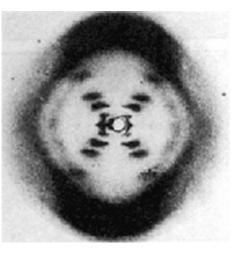
Base pairing from physical-chemical measurements Text

Sugar configuration (PUCKER)

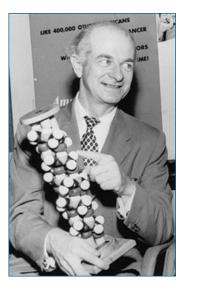
Angles of the glycosidic bonds ( were fixed within certain limits

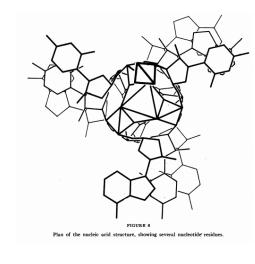
Handedness of the helix The direction of rotation was guessed and then subjected to testing

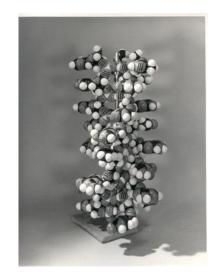




### Linus Pauling – suggested triple helix structure with bases outside - INCORRECT

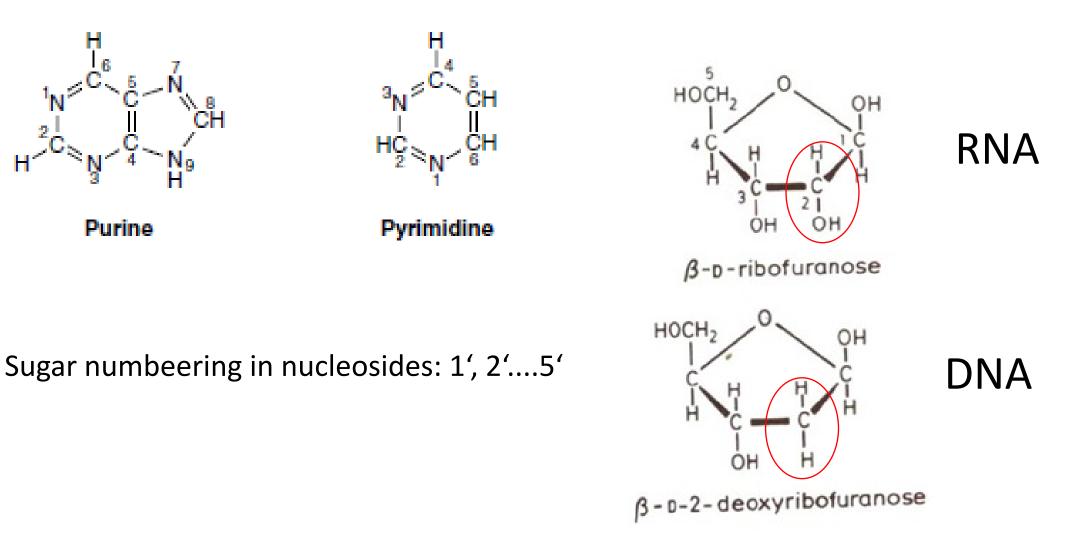




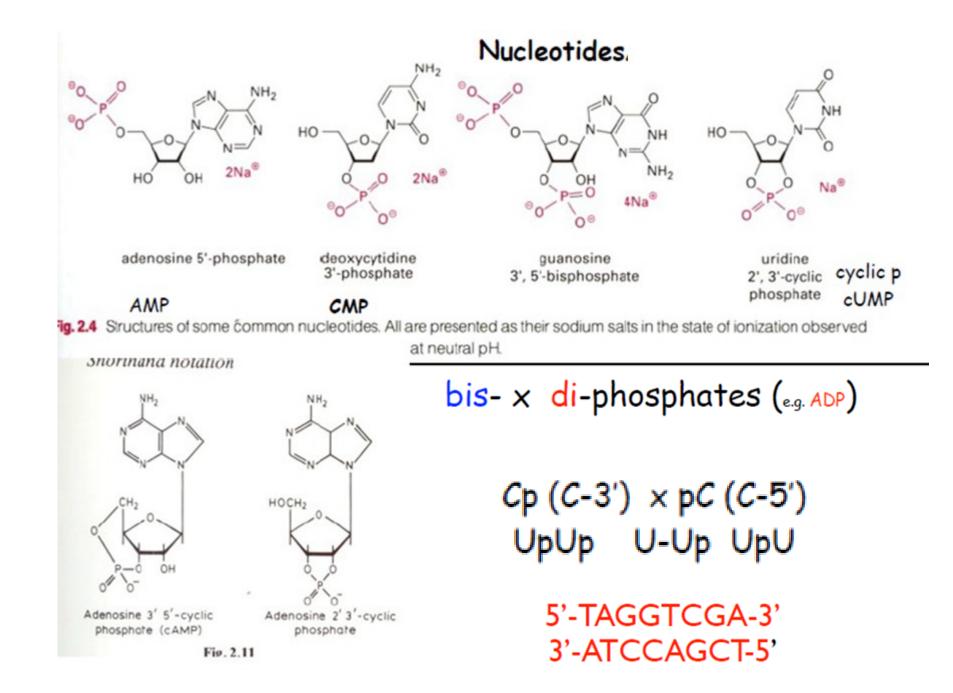


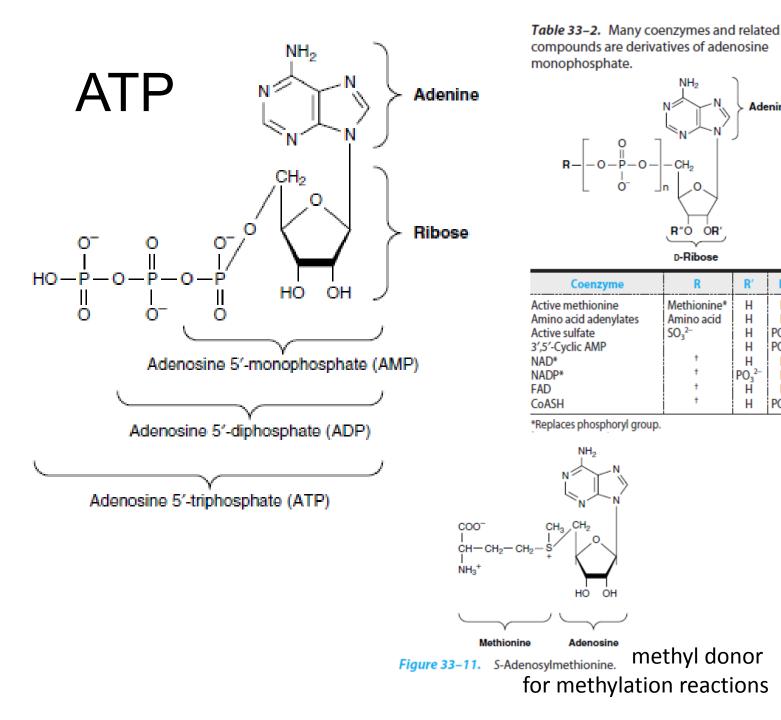
Other concepts: ladder (not interwound) structure (to overcome topological problems with unwinding the double helix)

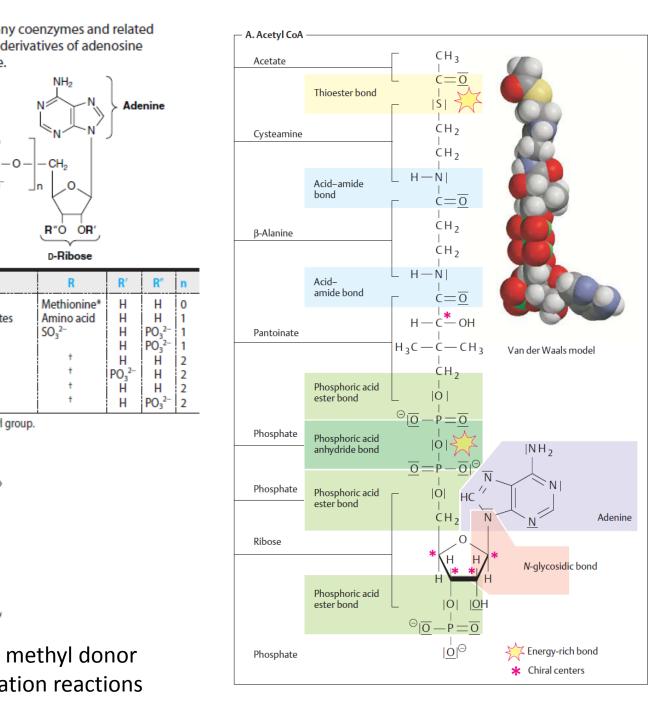
## Building blocks of nucleic acids: bases and pentoses



| Base<br>Formula          | Base<br>X = H | Nucleoside<br>X = Ribose or<br>Deoxyribose | Nucleotide, Where<br>X = Ribose Phosphate |   |   |  |
|--------------------------|---------------|--|---|---|---|--|
|                          | Adenine<br>A  | Adenosine<br>A                             | Adenosine monophosphate<br>AMP            |   |   | Nucleosides                              |
|                          | Guanine<br>G  | Guanosine<br>G                             | Guanosine monophosphate<br>GMP            |   | HOCH2<br>HOCH2<br>HO<br>HO<br>HO<br>H                                     | β anomer                                 |
|                          | Cytosine<br>C | Cytidine<br>C                              | Cytidine monophosphate<br>CMP             | Adenosine<br>(9-ß-D-ribofuranosyl<br>adenine)<br>NH <sub>2</sub><br>N=C<br>CH<br>I<br>I<br>CH | Guanosine<br>(9-β-D-ribofuranosyl<br>guanine)<br>HN<br>C-CH<br>C-CH<br>CH | Guanine riboside<br><sub>Guanosine</sub> |
|                          | Uracil<br>U   | Uridine<br>U                               | Uridine monophosphate<br>UMP              |   |   | Thymine deoxyriboside                    |
| H N CH₃<br>O N CH₃<br>dX | Thymine<br>T  | Thymidine<br>T                             | Thymidine monophosphate<br>TMP            | Cytidine<br>(1-β-D-ribofuranosyl<br>cytosine)   | Thymidine<br>(1-β-D-2-deoxyribo-<br>furanosylthymine)                     | ,  |







CH<sub>2</sub>

In

0

R"O OR

D-Ribose

R

Methionine\*

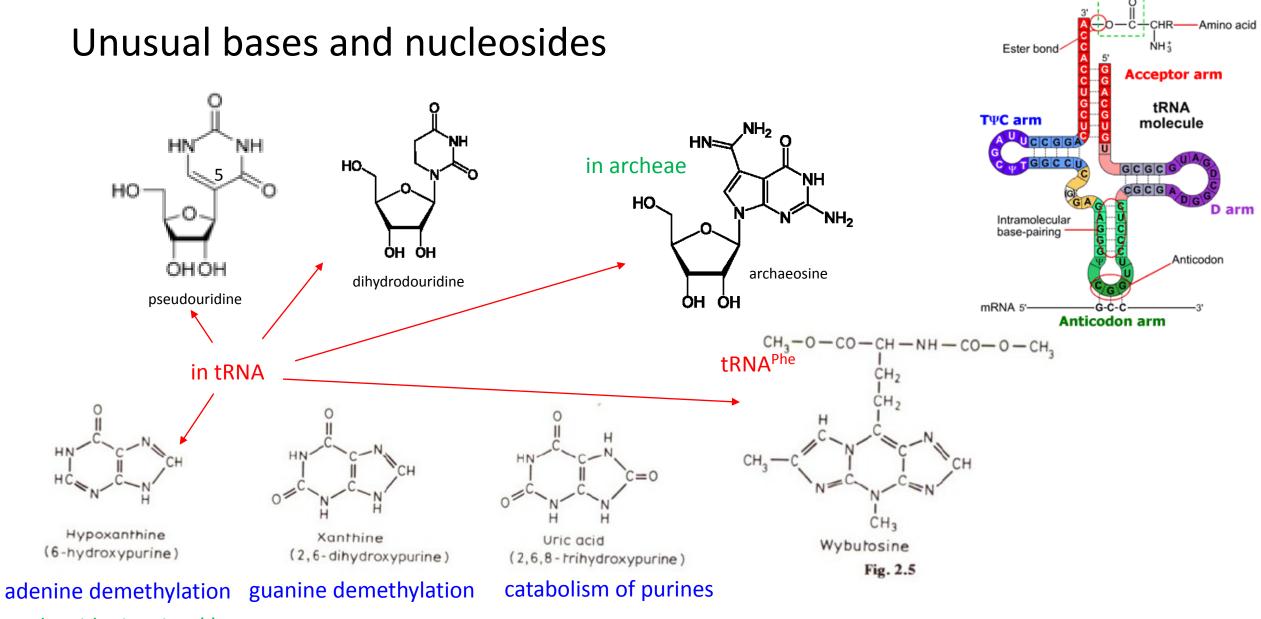
Amino acid

÷

÷

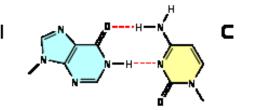
÷

SO32-

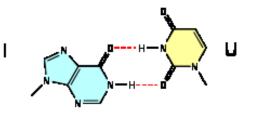


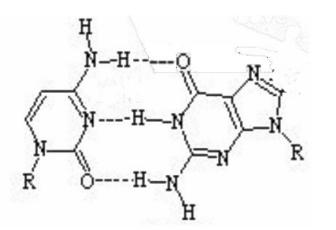
nucleoside=inosine (I)

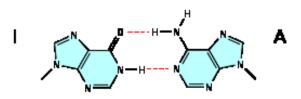
Watson-Crick base pairs ("canonical") wobble pairs (examples)

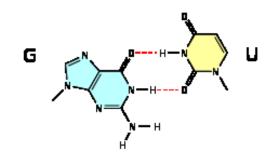


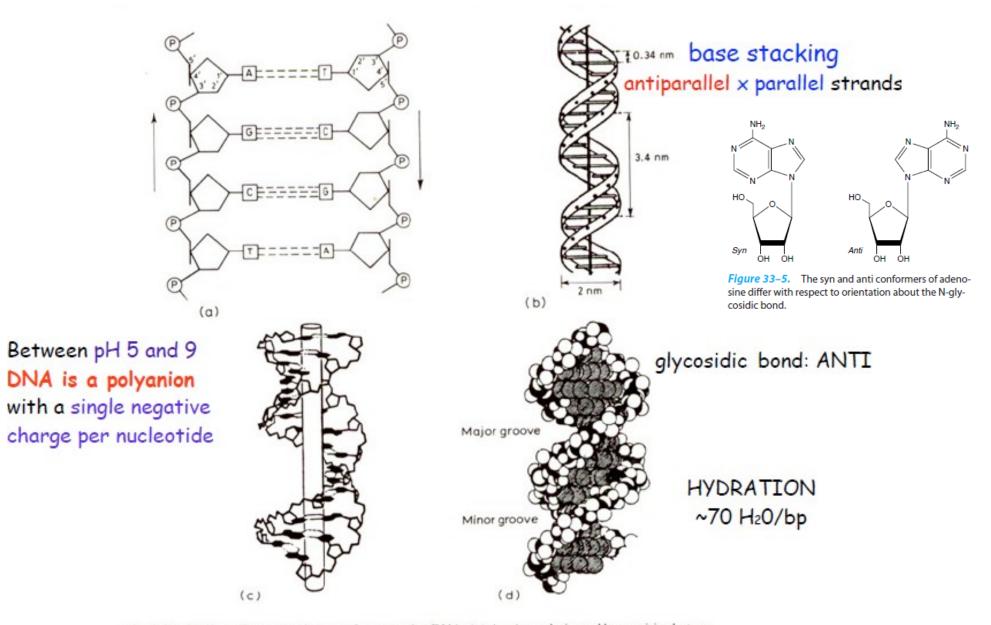
(in fact canonical)







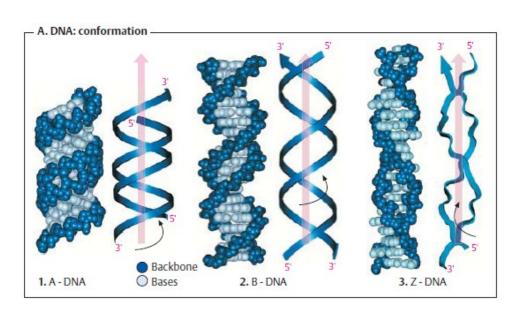


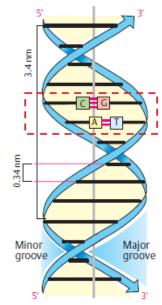


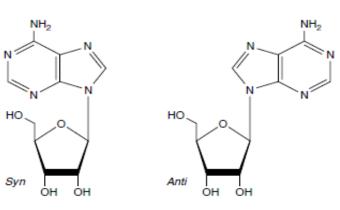
**Fig. 2.15** Various diagrammatic ways of representing DNA: (a) showing polarity and base pairing but no helical twist; (b) showing helical twist and helix parameters but not base pairs; (c) showing helix and base pairs; (d) space-filling representation showing major and minor grooves.

### TABLE 1 Comparison of A-, B-, and Z-DNA

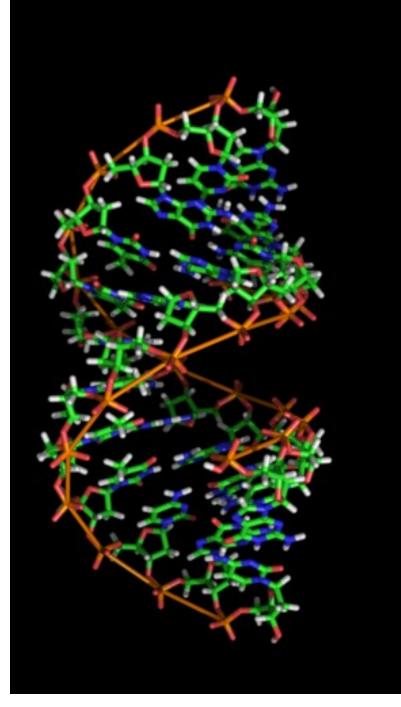
| Helix sense                    | A-DNA <sup>a</sup><br>right-handed | B-DNA <sup>e</sup><br>right-handed | B'-DNA <sup>b</sup><br>right-handed | Z-DNA <sup>o</sup><br>left-handed |
|--------------------------------|------------------------------------|------------------------------------|-------------------------------------|-----------------------------------|
| Base pairs per turn            | 11                                 | 10                                 | 10                                  | 12 (6 dimers)                     |
| Helix twist (°)                | 32.7                               | 36.0                               | 34.1, 36.8                          | -10, -50                          |
| Rise per base pair (Å)         | 2.9                                | 3.4                                | 3.5, 3.3                            | 3.7                               |
| Helix pitch (Å)                | 32                                 | 34                                 | 34                                  | 45                                |
| Base pair tilt (°)             | 13                                 | 0                                  | 0                                   | -7                                |
| P distance from helix axis (Å) | 9.5                                | 9.3                                | 9.1                                 | 6.9, 8.0                          |
| Glycosidic orientation         | anti                               | anti                               | anti                                | anti, syn                         |
| Sugar conformation             | C3'-endo                           | Wide range                         | C2'-endo                            | C2'-endo, C3'-<br>endoª           |

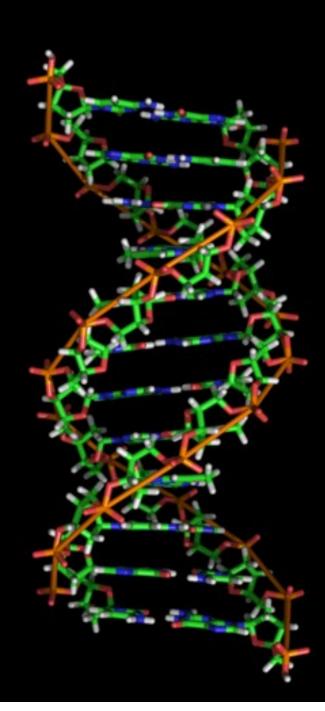


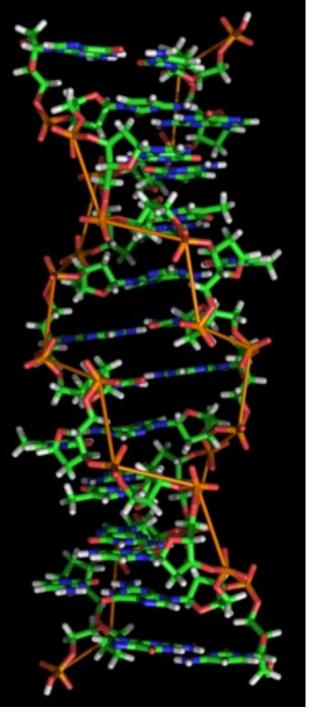




*Figure 33–5.* The syn and anti conformers of adenosine differ with respect to orientation about the N-glycosidic bond.



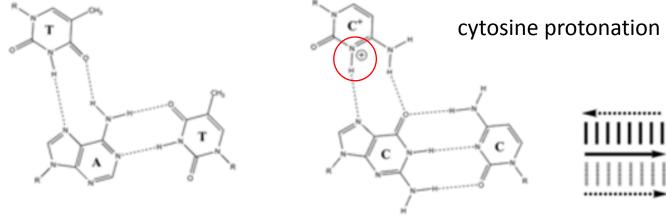




## Multistranded DNA structures

- triplexes
- tetraplexes (quadruplexes)

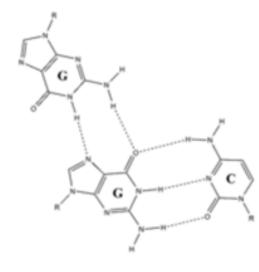
## Hoogsteen base pairs (triads)

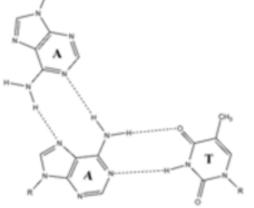


TA\*T

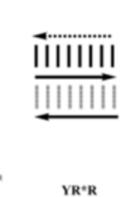


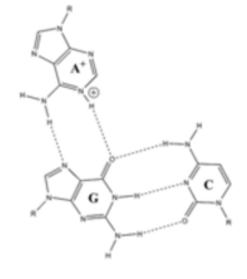






TA\*A

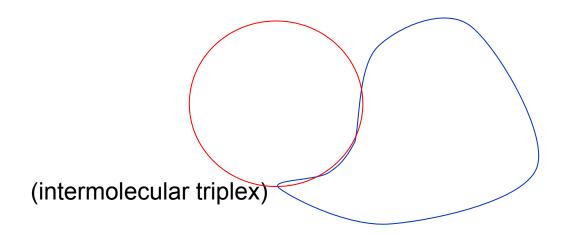




CG\*G

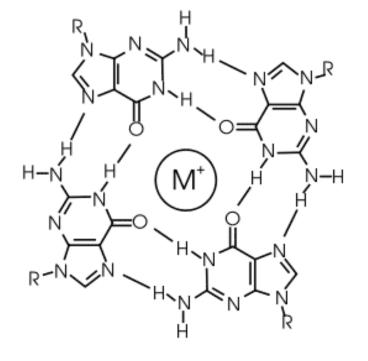
 $CG^*A^*$ 

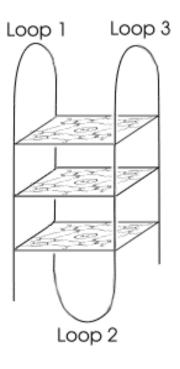
# Triplex DNA (homopurine-homopyrimidine stretch of suitable sequence)

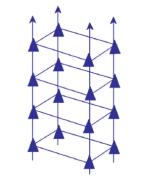


#### Guanine tetrad

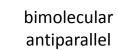
#### Guanine tetraplex



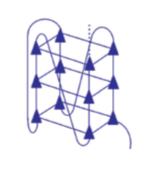




tetramolecular parallel



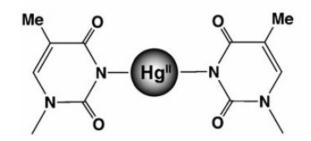
T\_T

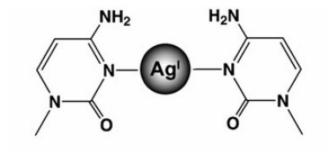




intramolecular parallel intramolecular antiparallel

metal ion-mediated pairing (non-physiological)



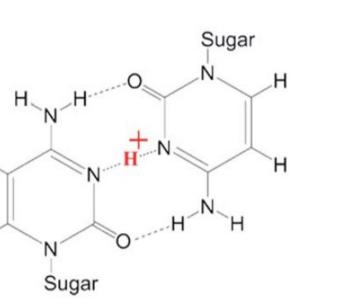


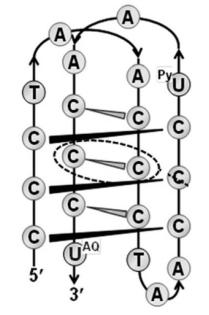
#### cytosine tetraplex (i-motif)

hemiprotonated C<sup>+</sup>•C pair

н

Н

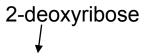




# Chemical reactivity of DNA

## Chemical reactivity of DNA

- DNA chemistry is derived from chemistry of its costituents
- phosphodiester bonds
- N-glycosidic bonds
- deoxyribose
- nitrogenous bases

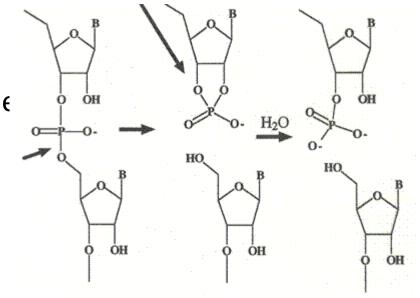


#### Chemical modification of DNA:

- damage to the genetic material
- analytical use

## DNA hydrolysis

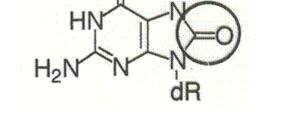
- both phosphodiester and N-glycosidic bonds susceptible to acid hydrolysis
- N-glycosidic bond more stable toward hydrolysis in pyrimidine than in purine nucleosides (and more in ribo- than in deoxynucleosides)
- stable in alkali (unlike RNA)
- alkali-labile sites: upon DNA damage
- enzymatic hydrolysis (N-glycosylases, nucle



## Oxidation

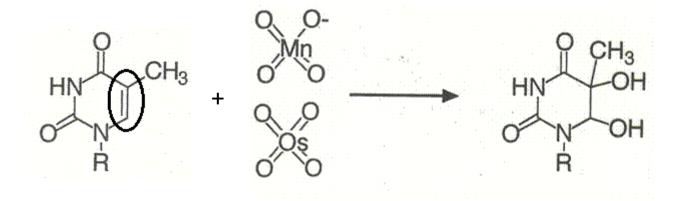
• two main sites susceptible to oxidation attacks:

• C8 of purines (ROS)



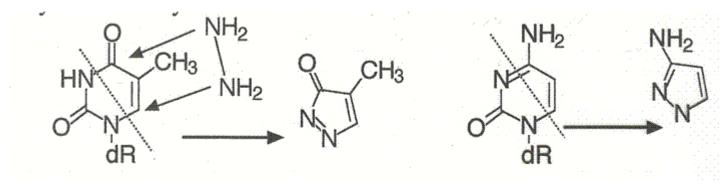
Н

• C5-C6 of pyrimidines



## reactions with nucleophiles

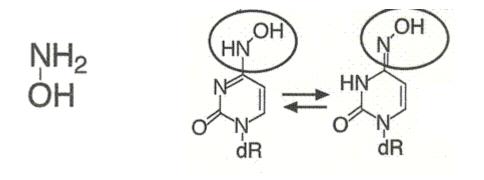
• C4 and C6 are centres of electron deficit in pyrimidine moieties (electrophile centres)



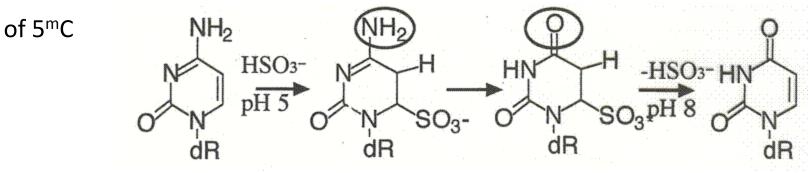
- reaction with hydrazine: pyrazole derivative and urea residue bound to the sugar
- with T the reaction is disfavored in high salt: Maxam-Gilbert sequencing technique

## reactions with nucleophiles

- hydroxylamine: cytosine modification
- the products' preferred tautomer pairs with adenine  $\rightarrow$  mutagenic

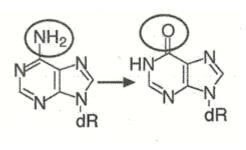


- bisulphite: cytosine modification inducing its deamination to uracil →mutagenic
- 5-methyl cytosine does not give this reaction: genomic sequencing

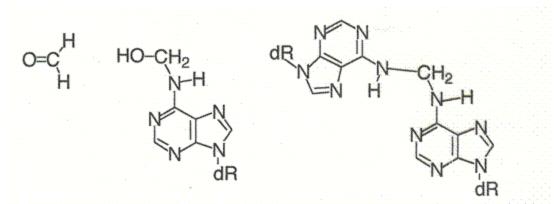


## reactions with electrophiles

- attacking N and/or O atoms
- nitrous acid (HNO<sub>2</sub>) causes base deamination (C→U, A→I) affecting base pairing, mutagenic

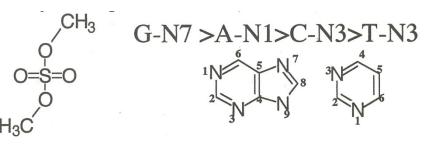


- aldehydes: reactions with primary amino groups
- formaldehyde: two step reaction

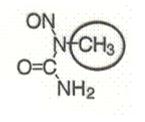


## **DNA** alkylation

- hard or soft alkylating agents
- hard ones attack both N and O atoms, soft only N
- dimethyl sulfate: typical soft alkylating agent

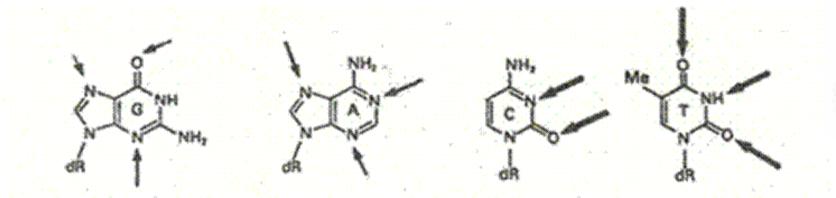


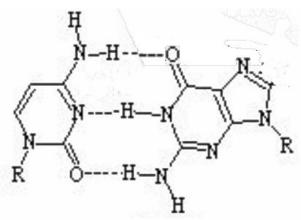
- N-alkyl-N-nitroso urea: typical hard alkylating agent
- modifies all N + O in bases as well as phosphate groups (forming phosphotriesters)
- analytical use (sequencing, foorprinting)



#### Biological consequences of base alkylation

- **N-alkylation**: the primary site = N7 of guanine (accessible in both ss and dsDNA)
  - does not change base pairing; easily repairable
- N3 of adenine or guanine: located in minor groove
  - cytotoxic modification (DNA/RNA polymerization blocked)
- N1 of guanine: interferes with base pairing
- O-alkylation (G-O6, T-O6) the bases "locked" in enol forms → improper base pairing → mutagenic





# Tautomerization, base pairing and chemical mutagenesis

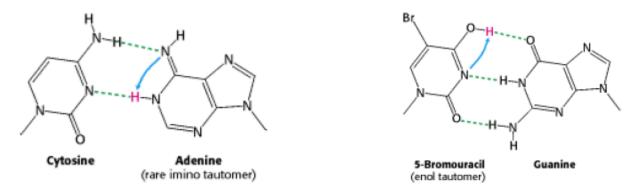
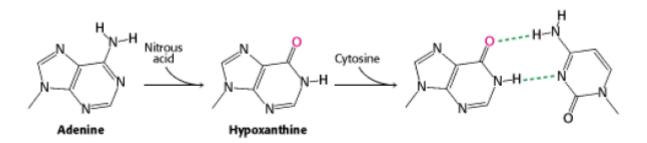


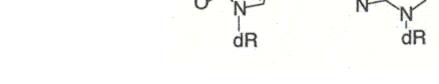
Figure 27.41. Base Pair with Mutagenic Tautomer. The bases of DNA can exist in rare tautomeric forms. The imino tautomer of adenine can pair with cytosine, eventually leading to a transition from A-T to G-C.



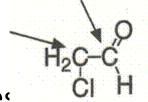
similarly uracil is deamination product of cytosine

**Figure 27.43.** Chemical Mutagenesis. Treatment of DNA with nitrous acid results in the conversion of adenine into hypoxanthine. Hypoxanthine pairs with cytosine, inducing a transition from A-T to G-C.

- chloro- (bromo-) acetaldehyde: two reactive centres (aldehyde and alkylhalogenide)
- reaction with C or A
- chemical probes (react only with unpaired bases,



- diethyl pyrocarbonate: acylation of purines (primarily A) or C
- modification leads to opening of the imidazole ring
- chemical DNA probing  $O = CH_2CH_3 \qquad VH_2CO \qquad VH_2CO$

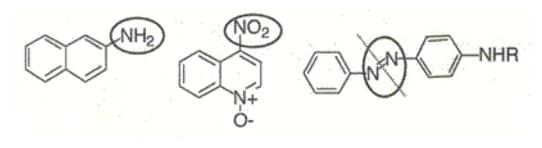


### Metabolically activated carcinogens

- some substances became toxic after their metabolic conversion
- detoxifying machinery of the organism acts here as a bad fellow
- microsomal hydroxylase complex, cytochrome P450
- the role of this system is to introduce suitable reactive groups into xenobiotics enabling their conjugation with other molecules followed by removal from the organism
- but....

### Metabolically activated carcinogens

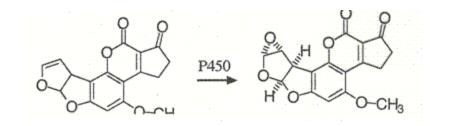
 aromatic nitrogenous compounds (amines, nitro- or azo- compounds):

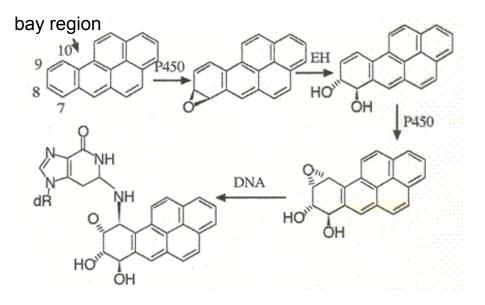


- aromatic amines are converted into either (safe) phenols, or (dangerous) hydroxylamine derivatives
- azo- compounds: "cleaved" into amines
- nitro- compounds: reduced into hydroxylamines

## Metabolically activated carcinogens

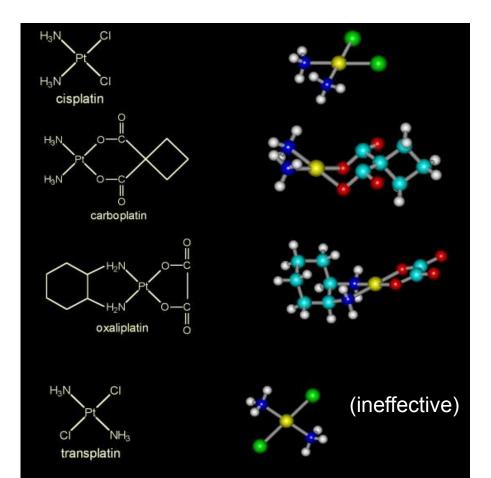
- polycyclic aromatic hydrocarbons like benzo[ $\alpha$ ]pyrene: three-step activation
  - P450 introduces epoxy group
  - epoxide hydrolase opens the epoxide circle
  - P450 introduces second epoxy group
- DNA adduct formation (primarily -NH<sub>2</sub> of guanine, then G-N7, G-O6 and A-N6)
- similar pathway of aflatoxin activation



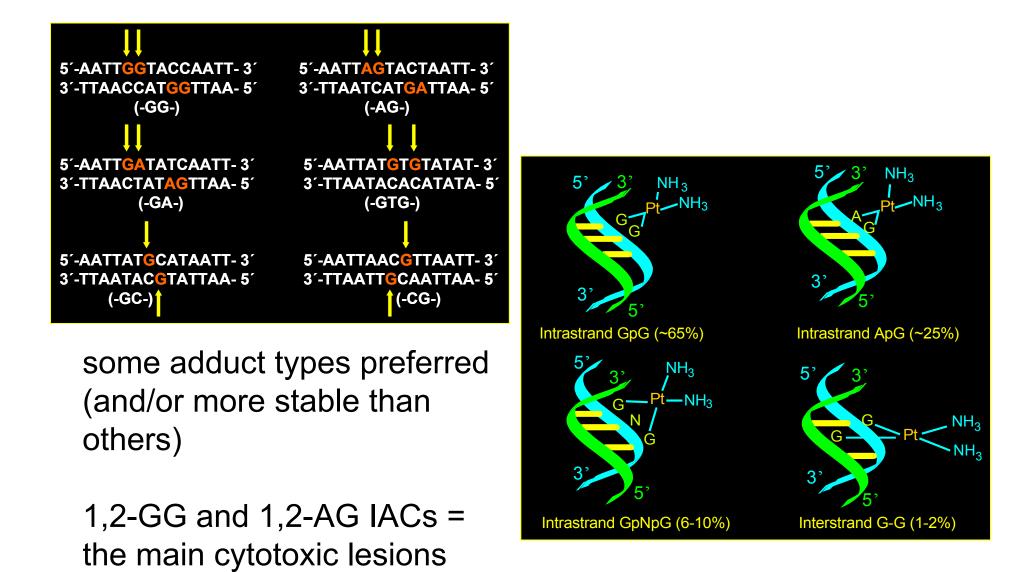


## anticancer drugs

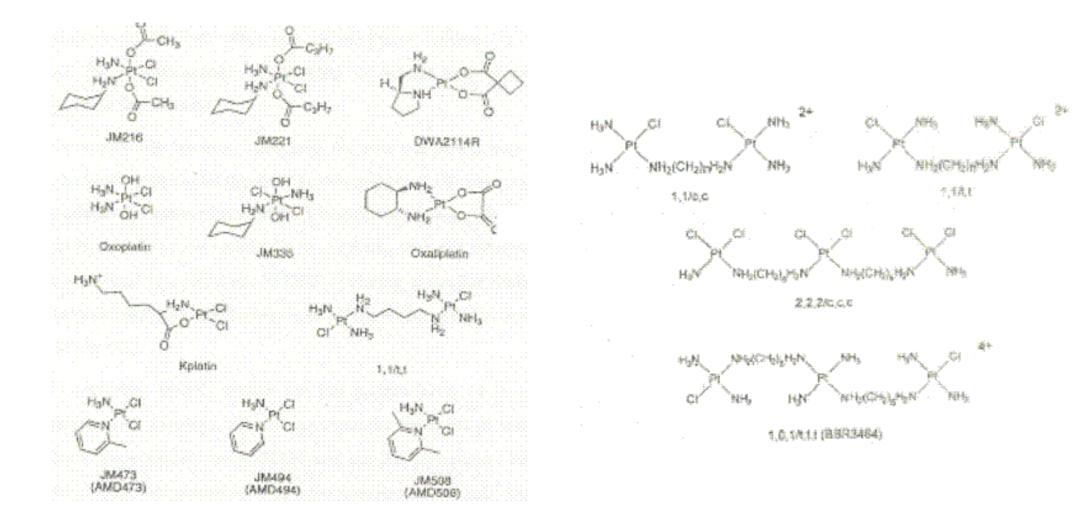
- some types of antineoplastic agents act via formation of DNA adducts
- metallodrugs: mainly platinum complexes



#### cisplatin: reaction with DNA in certain sequence motifs

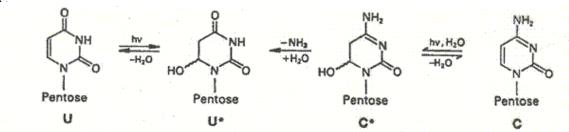


#### other platinum complexes tested as cytostatics

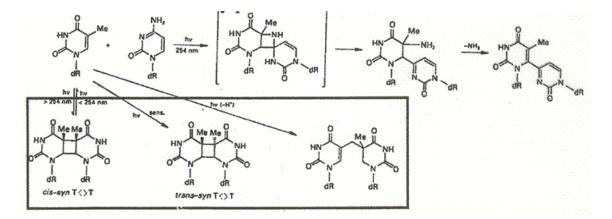


### **Photochemical DNA modifications**

- mainly pyrimidines
- excitation at 240-280 nm: reactive singlet state
- water addition at C5-C6



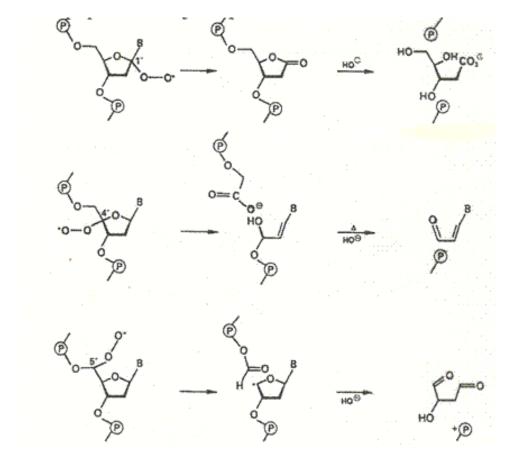
• excitation at 260-280 nm: photodimerization of pyrimidines



• photoproducts of C can deaminate to U (mutagenic effects)

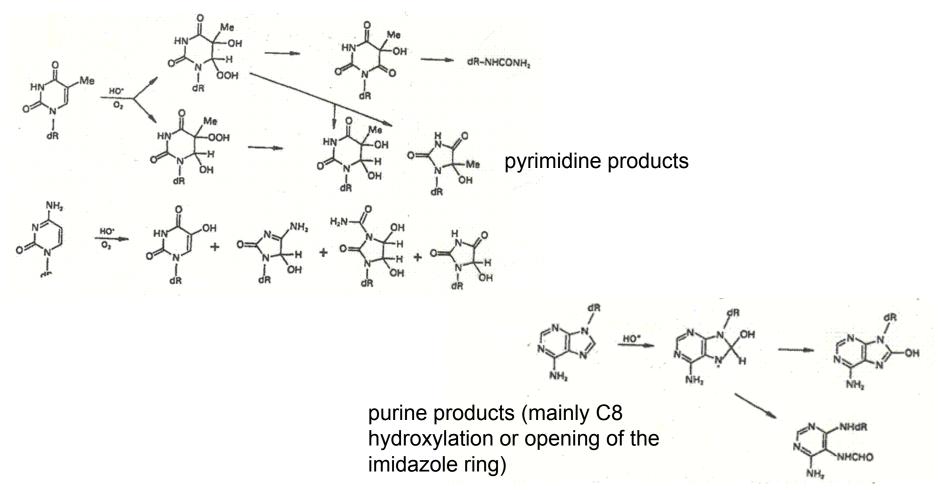
## effects of ionizing radiation

- mostly indirect through water radiolysis
- each 1,000 eV produces ~27 •OH radicals that attack DNA
- sugar damage:abstraction of hydrogen atoms from C-H bonds
- a series of steps resulting in strand breakage



## effects of ionizing radiation

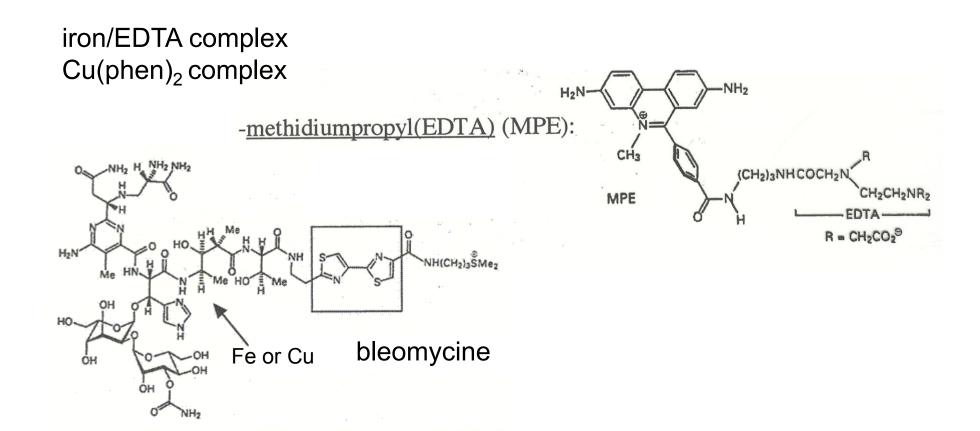
 base damage: hydroxylation and/or (under aerobic conditions) peroxylation



#### chemical nucleases

species containing redox active metal ions mediating production of hydroxyl radicals (or othe reactive oxygen species) via Fenton and/or Haber-Weiss processes

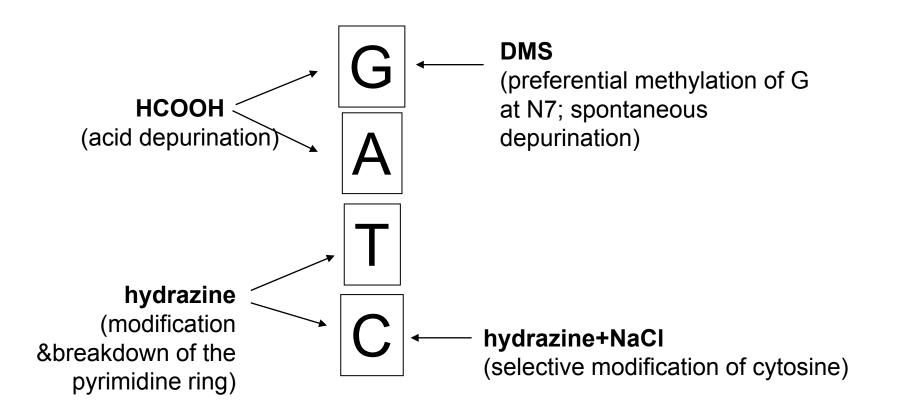
$$Me^n + H_2O_2 \rightarrow Me^{n+1} + \bullet OH + OH^-$$



# Chemical approaches in DNA studies

(several examples)

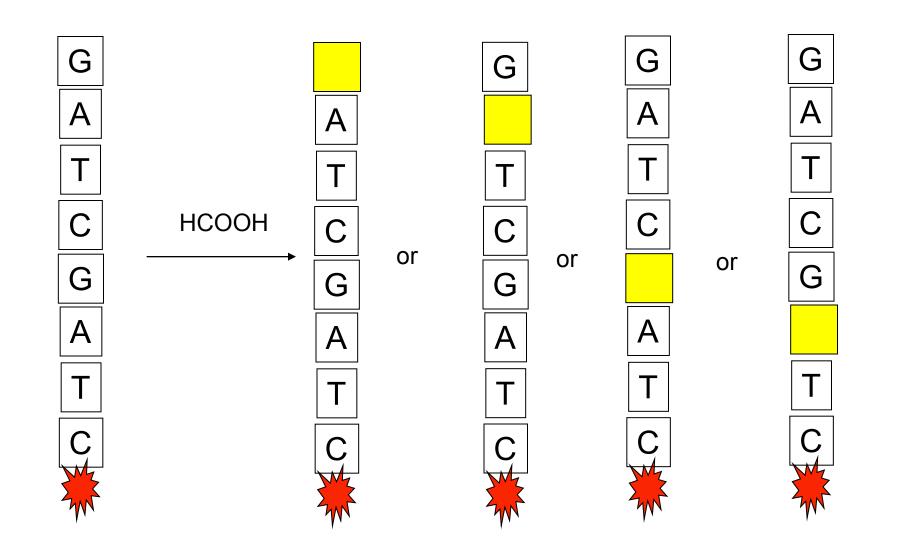
# Maxam and Gilbert method of DNA sequencing

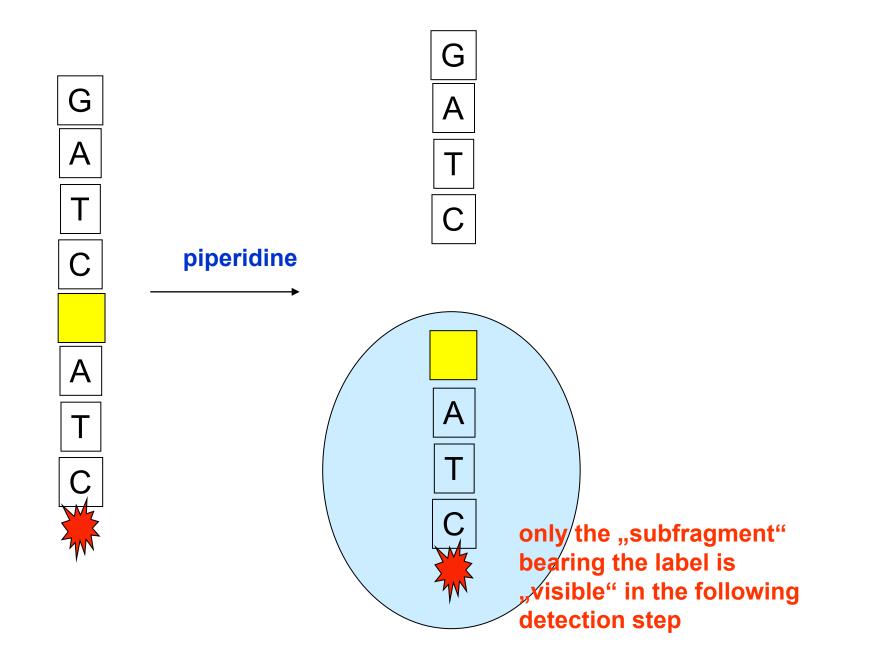


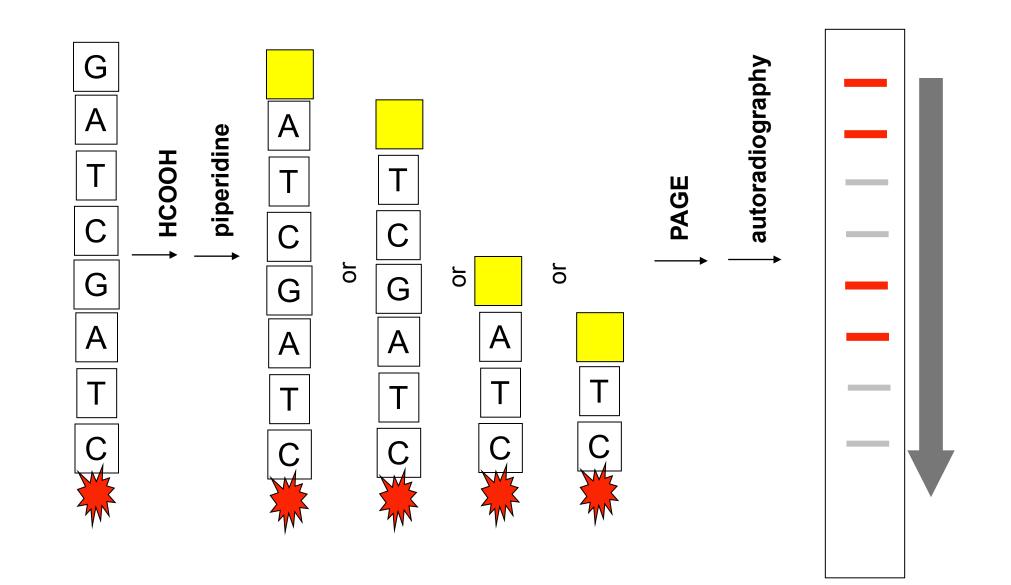
## at sites of base modification (removal) the sugar-phosphate backobone is labile towards alkali

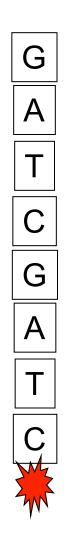
treatment with hot piperidine  $\rightarrow$  cleavage at such sites

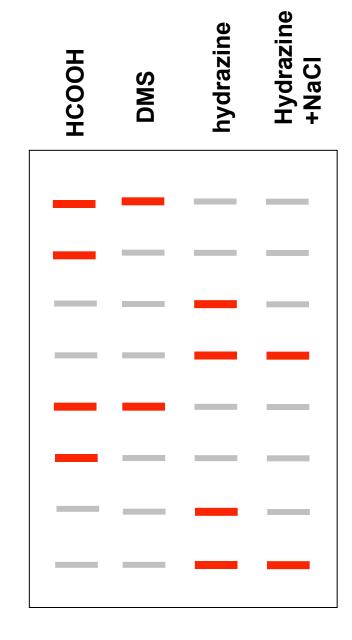
- DNA fragment is end-labeled (radionuclide, fluorophore)
- the sample is divided into four reactions (HCOOH, DMS, hydrazine, hydrazine + NaCl)
- the conditions are chosen to reach only one modification event per DNA molecule











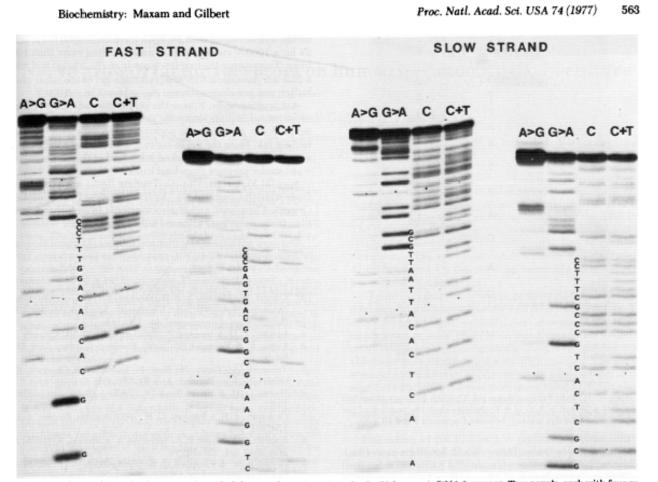
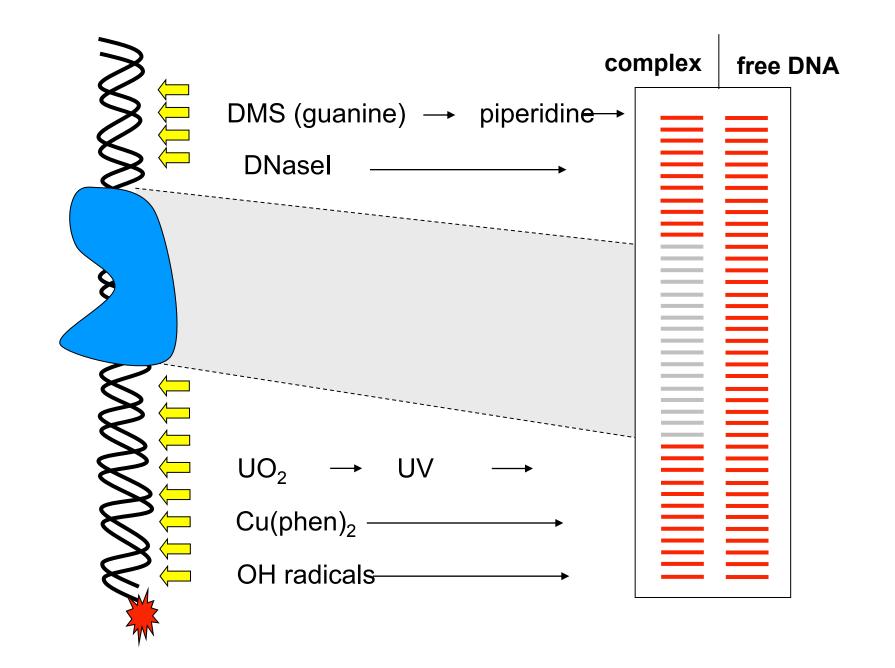


FIG. 2. Autoradiograph of a sequencing gel of the complementary strands of a 64-base-pair DNA fragment. Two panels, each with four reactions, are shown for each strand; cleavages proximal to the 5' end are at the bottom on the left. A strong band in the first column with a weaker band in the second arises from an A; a strong band in the second column with a weaker band in the first is a G; a band appearing in both the third and fourth columns is a C; and a band only in the fourth column is a T. To derive the sequence of each strand, begin at the bottom of the left panel and read upward until the bands are not resolved; then, pick up the pattern at the bottom of the right panel and continue upward. One-tenth of each strand, isolated from the gel of Fig. 1, was used for each of the base-modification reactions. The dimethyl sulfate treatment was 50 mM for 30 min to react with A and G; hydrazine treatment was 18 M for 30 min to react with C and T and 18 M with 2 M NaCl for 40 min to cleave C. After strand breakage, half of the products from the four reactions were layered on a  $1.5 \times 330 \times 400$  mm denaturing 20% polyacrylamide slab gel, pre-electrophoresed at 1000 V for 2 hr. Electrophoresis at 20 W (constant power), 800 V (average), and 25 mA (average) proceeded until the xylene cyanol dye had migrated halfway down. Autoradiography of the gel for 8 hr produced the pattern shown.

DNA "footprinting": determination of binding sites of other molecules (e.g. proteins) at the DNA sequence level



# single strand-selective chemical probes

#### Open local structures in negatively supercoiled DNA

relaxed circular DNA

negatively supercoiled DNA (linking deficit)

stress related to the negative superhelicity (the linking deficit) can be absorbed in local open structures Open local structures in negatively supercoiled DNA

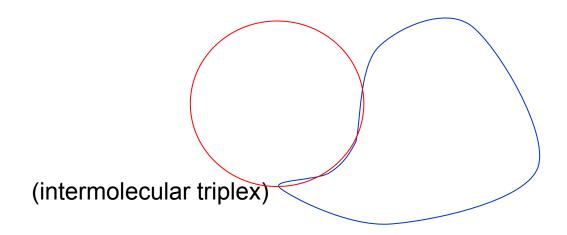
DNA segments of specific sequence can adopt "alternative" local structures

cruciform DNA (inverted repeat)

t unpaired bases Open local structures in negatively supercoiled DNA

#### Triplex DNA

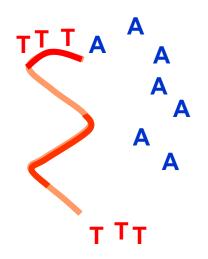
(homopurine homopyrimidine stretch with mirror symmetry)



# Otevřené lokální struktury v negativně nadšroubovicové (sc) DNA

Intramolecular triplex

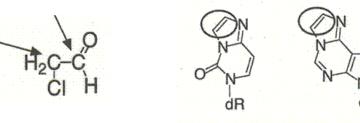
(homoPu•homoPy segment within negatively supercoiled DNA)



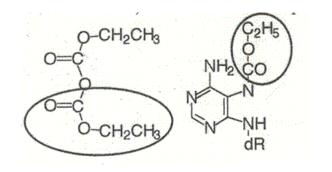
#### Chemicals selectively reacting with unpaired bases:

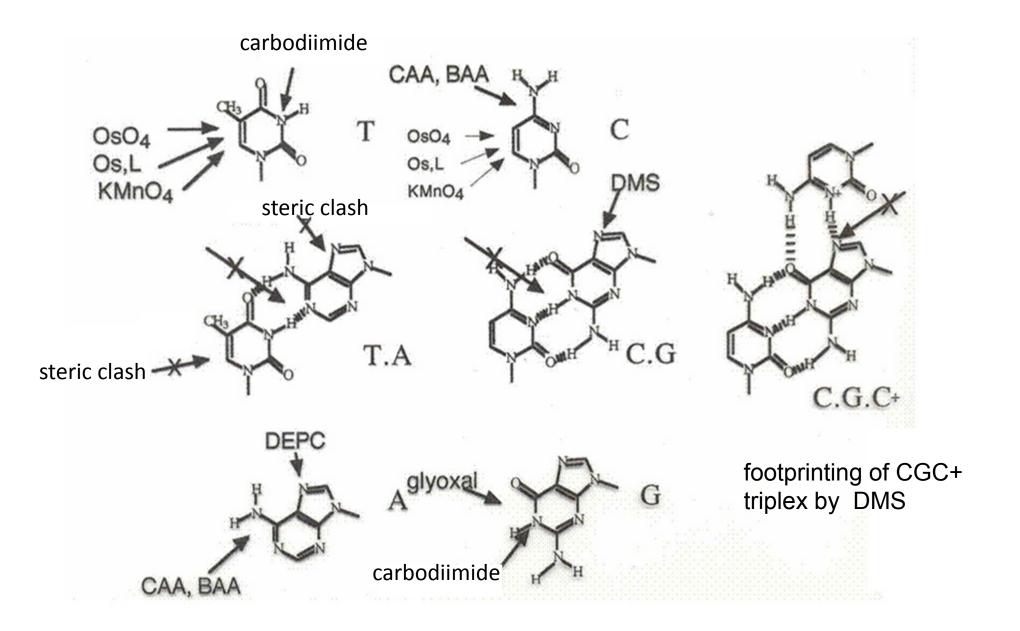
osmium tetroxide complexes (Os,L) (T, more slowly C)

chloroacetaldehyde (CAA) (A, C)



diethyl pyrocarbonate (DEPC) (A, G)





Using the Maxam-Gilbert technique, it is possible to determine with a high preciseness which nucleotides are forming the local structure

- modification of supercoiled DNA
- ➤ restriction cleavage, radiactive labeling
- ➤ hot piperidine
- ➢ sequencing PAGE

the structure can be deduced from the modification pattern

