

NUCLEIC ACIDS

Basic terms and notions

Presentation by

Eva Fadrná

adapted by Radovan Fiala

Literature

Books

Saenger, W., Principles of Nucleic Acid Structure, Springer 1984.

Bloomfield, V. A., Crothers, D. M., Tinoco, I., Nucleic Acids, Structures, Properties, and Functions, Univ. Sci. Books, 2000.

Wuthrich, K., NMR of Proteins and Nucleic Acids, Wiley, 1986.

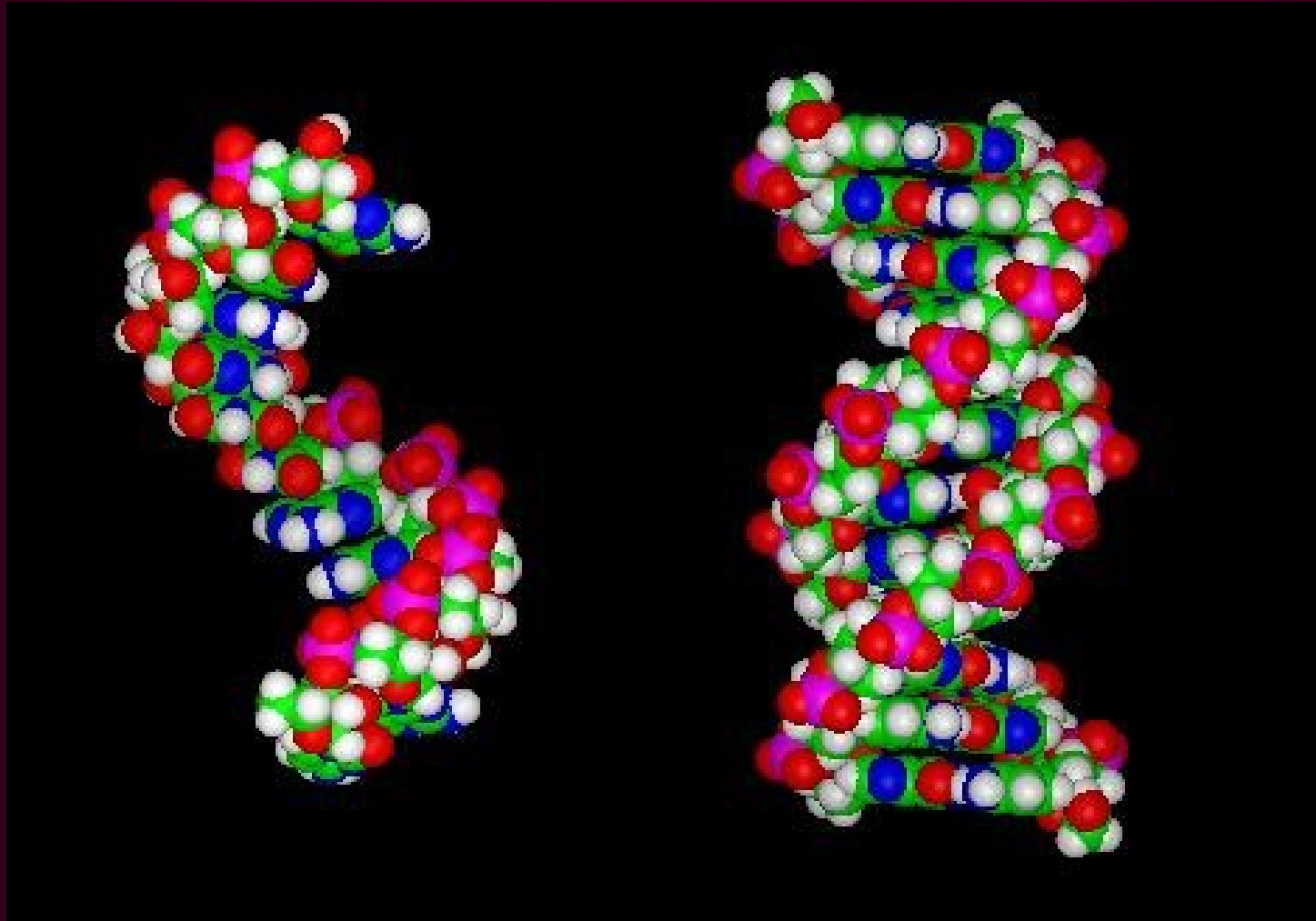
Review articles

Bowater, R. P., Waller, Z. AE., In: eLS. John Wiley & Sons, Chichester, 2014.

Wijmenga, S. S., van Buuren, B. N. M., Progr. NMR Spect. 32, (1998), 287-387.

Furtig, B. et al., ChemBioChem 4 (2003), 936-962.

RNA vs DNA



Single strand A-RNA

B-DNA duplex

Length of NA

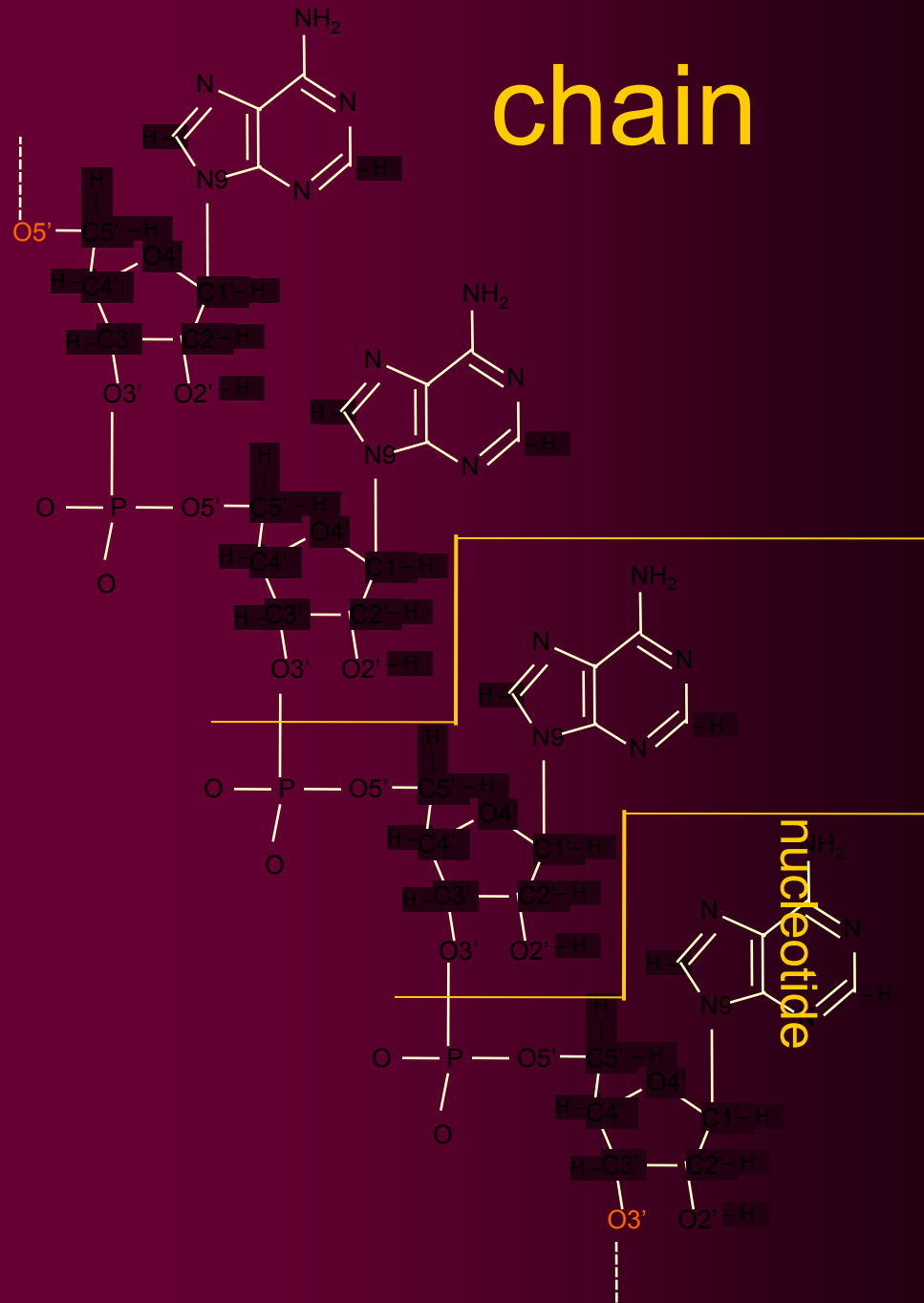
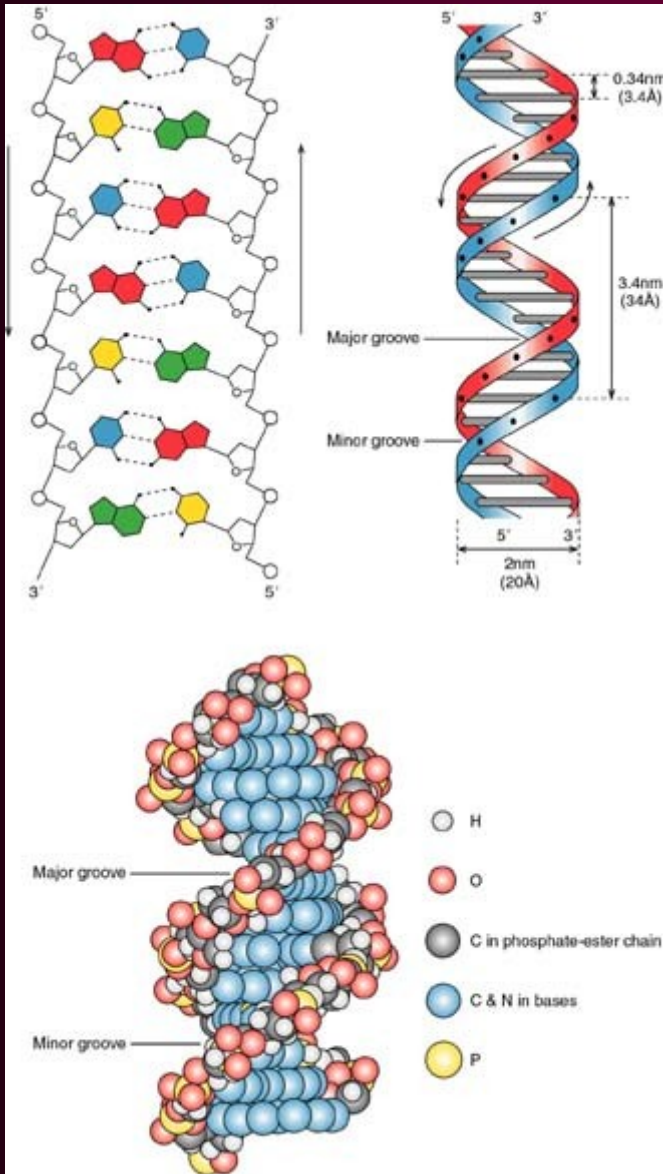
Total length of DNA in a human cell	1 m	(1000 km)
DNA in typical human chromosome	1 cm	(10 km)
DNA from bacterial chromosome	1 mm	
Diameter of typical human cell	0.01 mm	
Diameter of folded DNA	0.1 μm	(0.1 m)
Diameter of DNA fiber	1 nm	(1 mm)
Diameter of atom		1 Å

(multiplied by 10^6)

\Rightarrow 1 chromosome would be 10km long with fiber diameter of 1 mm and it would fold into 10 cm diameter \Rightarrow extraordinary DNA flexibility

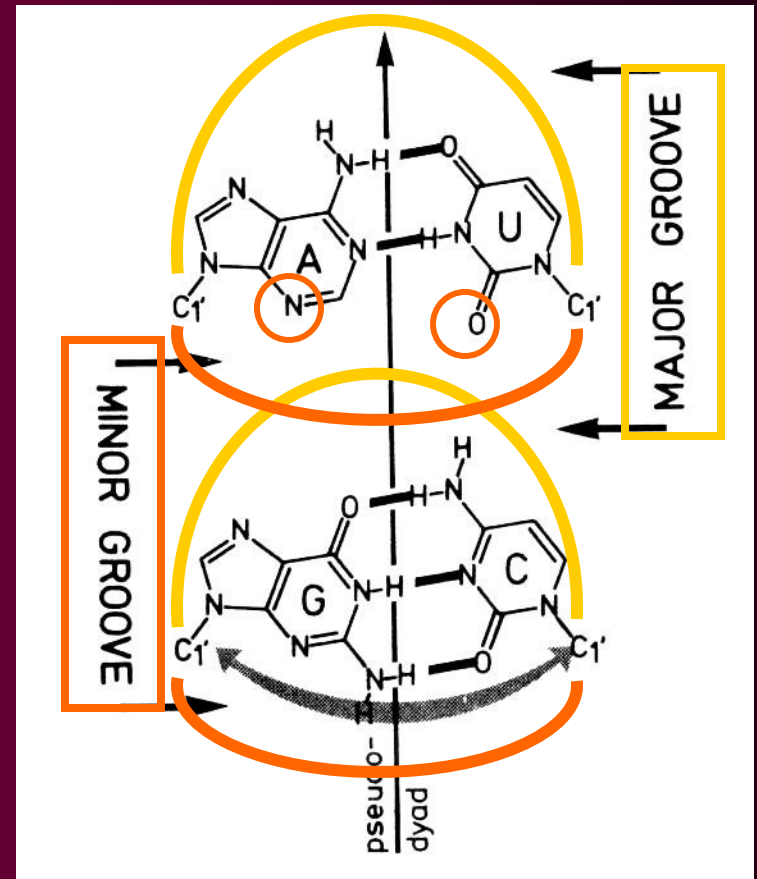
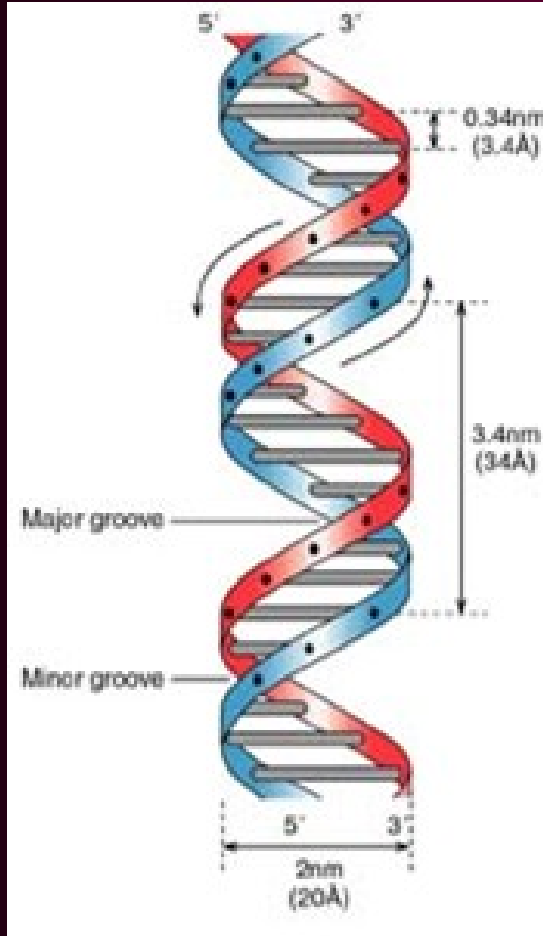
Nukleotide

chain

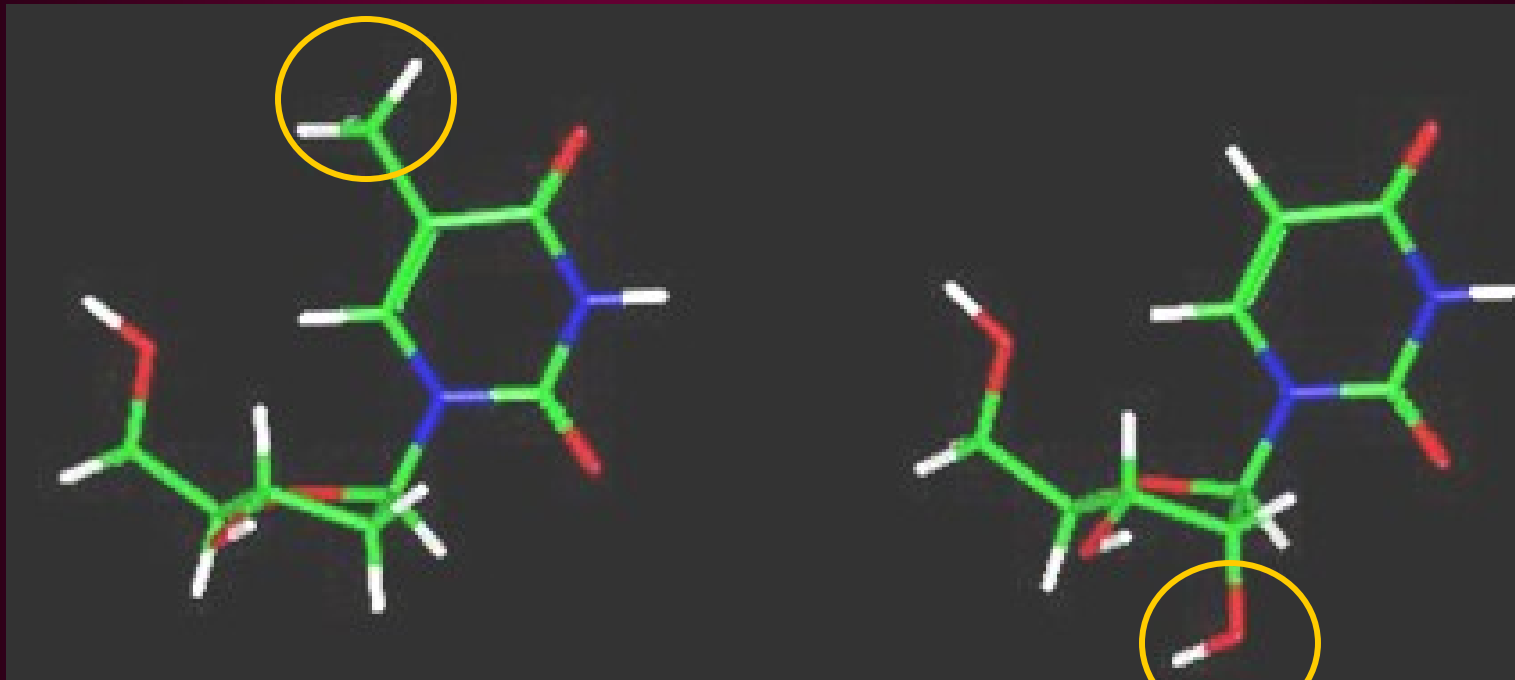
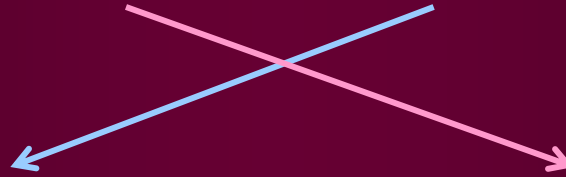


Grooves

major vs. minor



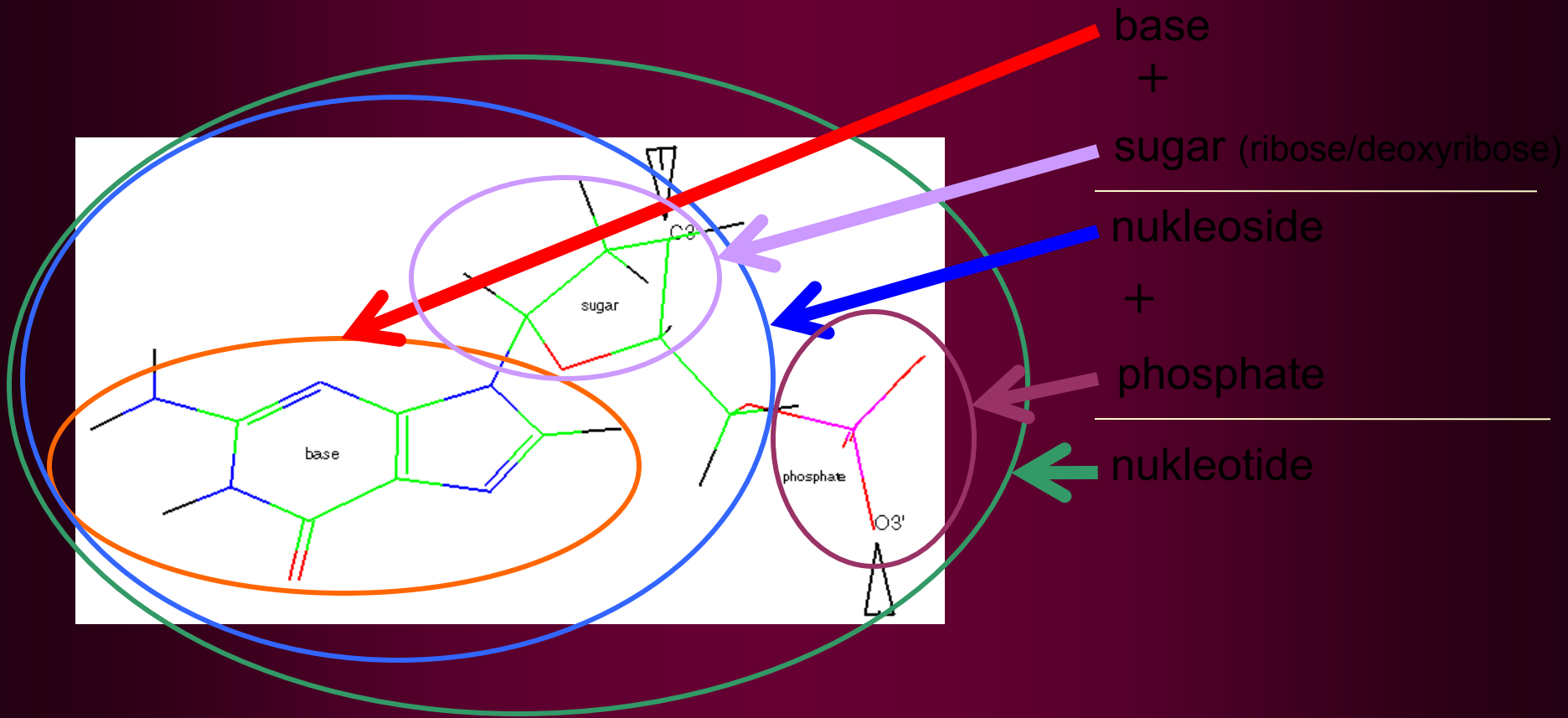
RNA vs DNA



deoxythymidine

uridine

Nukleotide/nukleoside

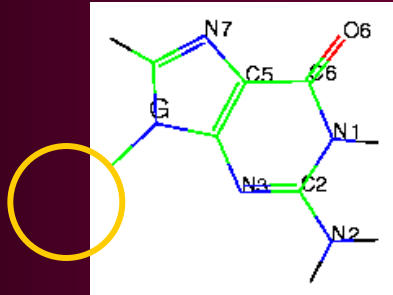


Bases

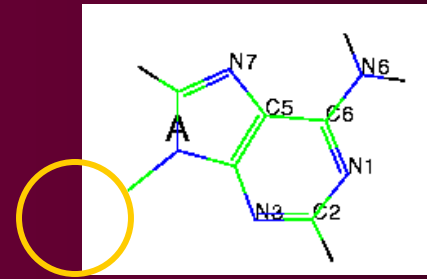
DNA

RNA

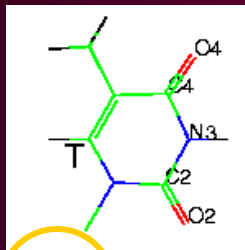
Guanin (Gua)



Adenin (Ade)

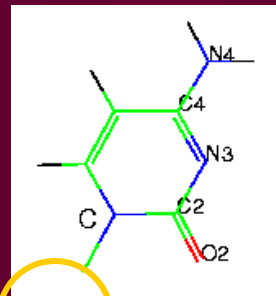


Thymin (Thy)

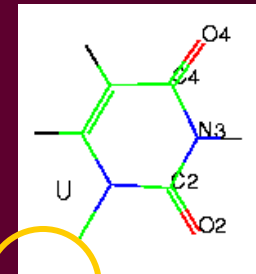


sugar

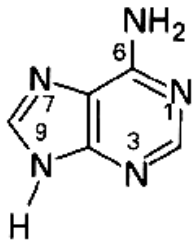
Cytosin (Cyt)



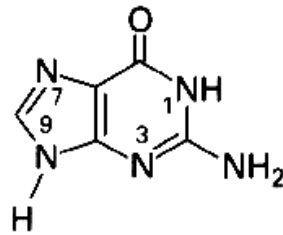
Uracil (Ura)



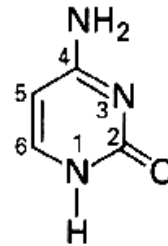
Base numbering



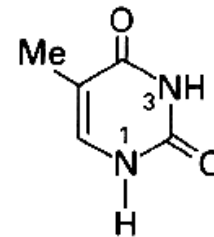
adenine



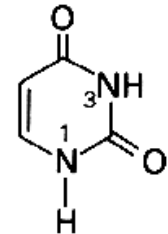
guanine



cytosine

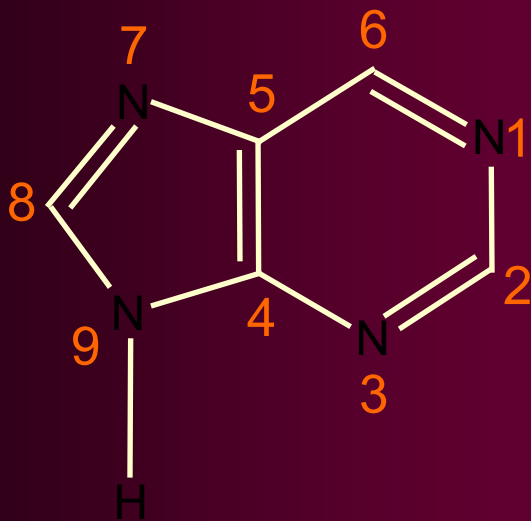


thymine

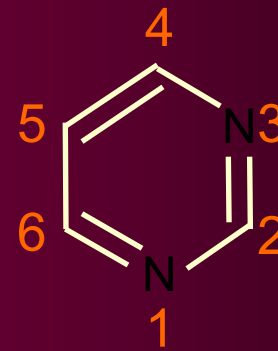


uracil

PURINES



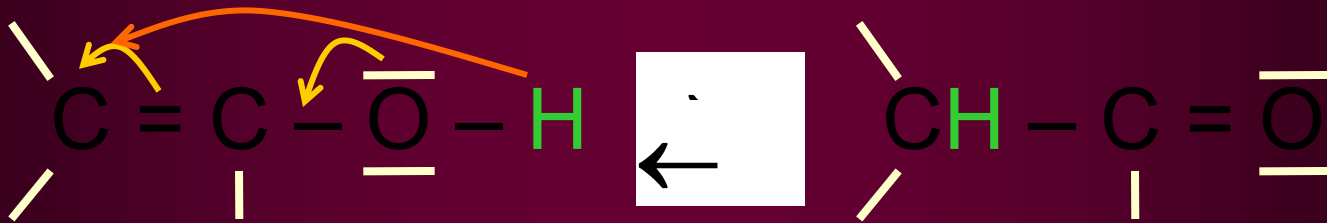
PYRIMIDINES



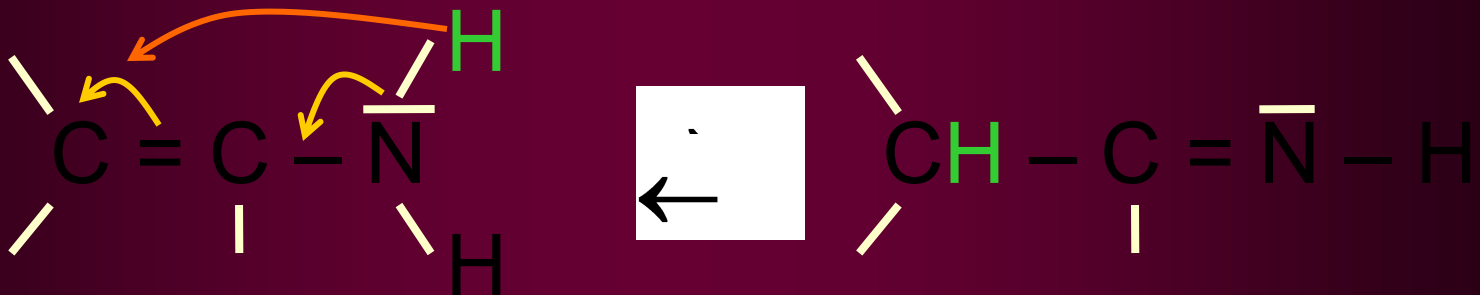
Base tautomerism

fysiolog. conditions

enol \leftrightarrow keto



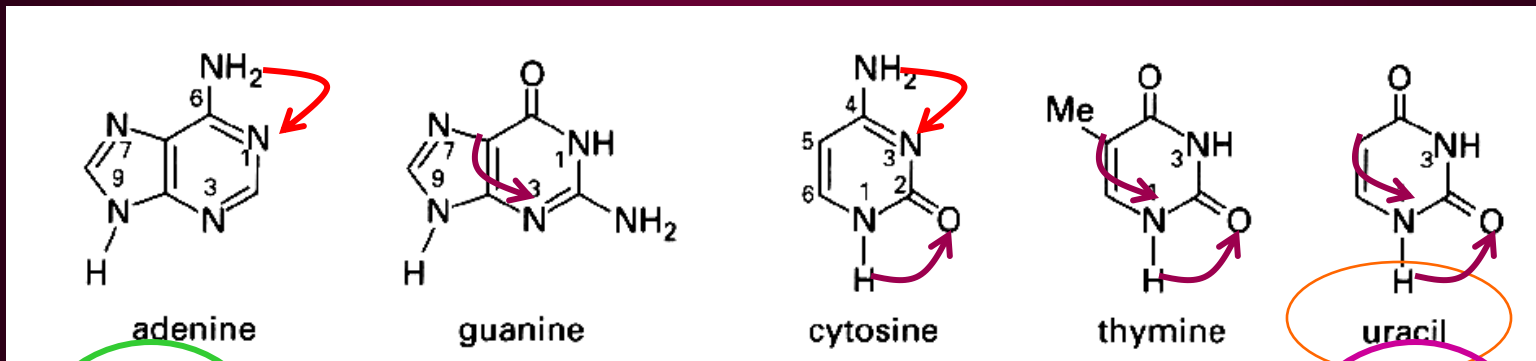
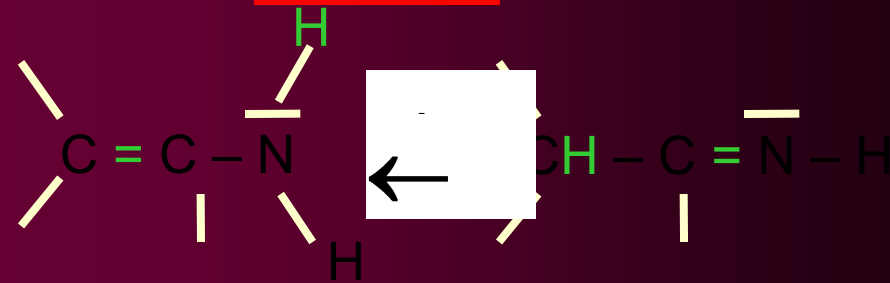
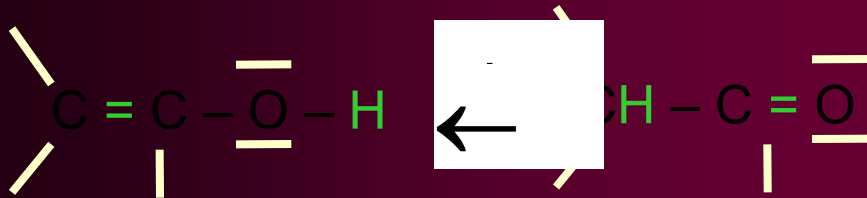
enamin \leftrightarrow imin



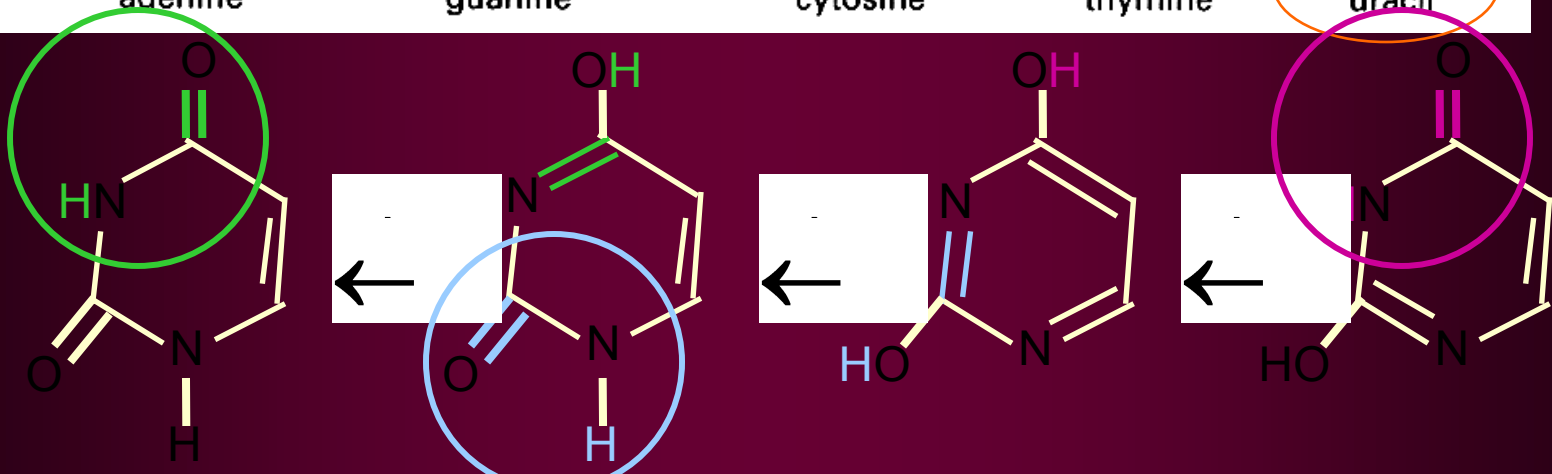
Base tautomerism

enol \leftrightarrow keto

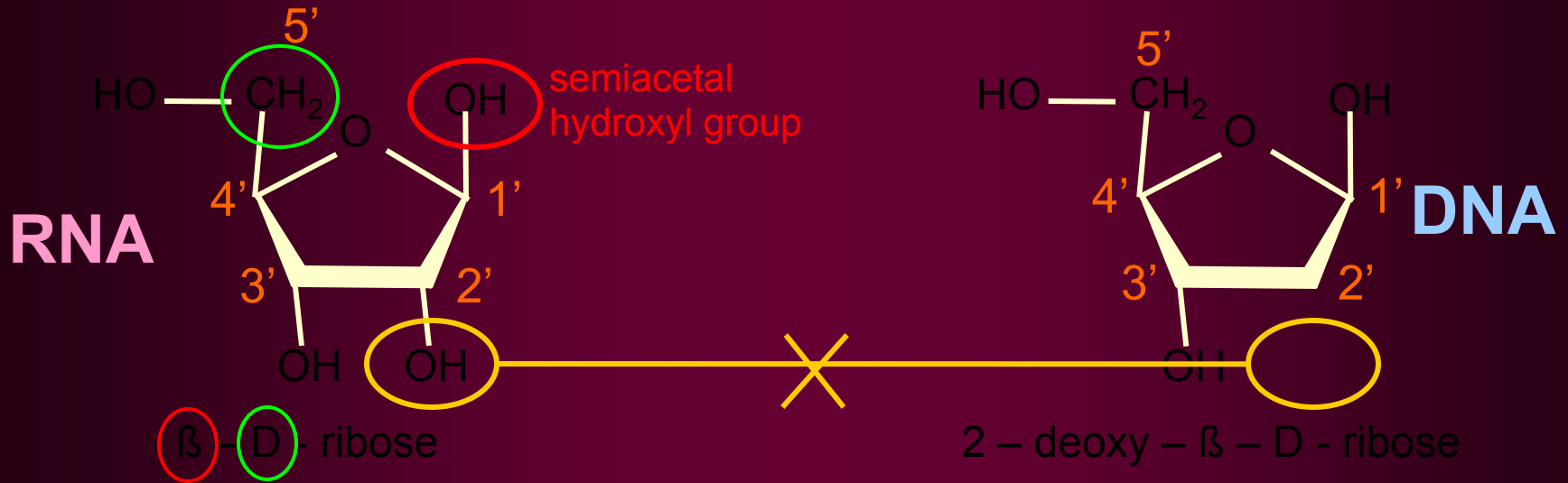
enamin \leftrightarrow imin



4 tautomers



Sugar - pentoses



semiacetal hydroxyl group

+ base



N-glycosidic bond



nukleoside

C1' - N1

pyrimidines

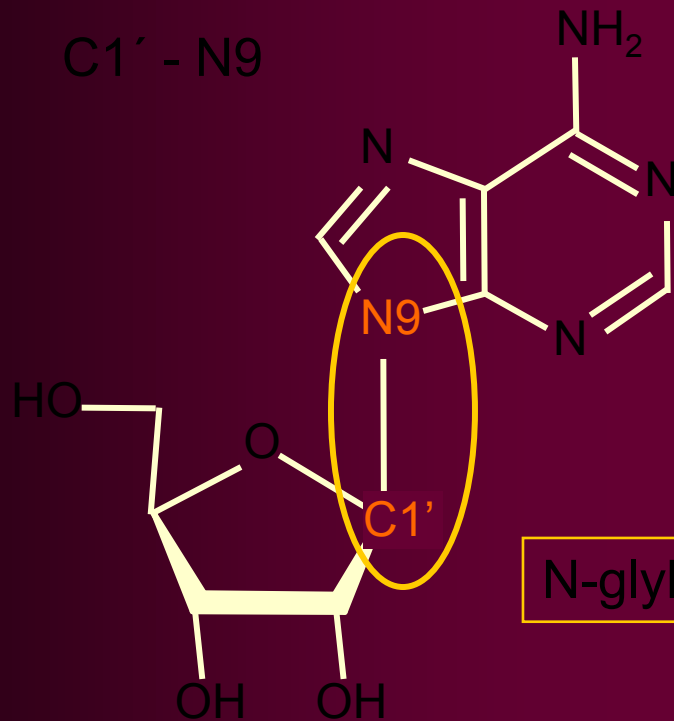
C1' - N9

purines

Nukleosides

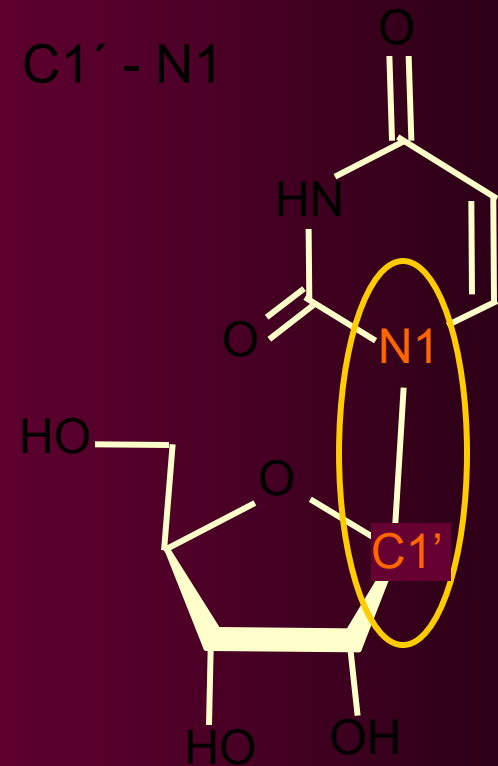
purines

C1' - N9



pyrimidines

C1' - N1



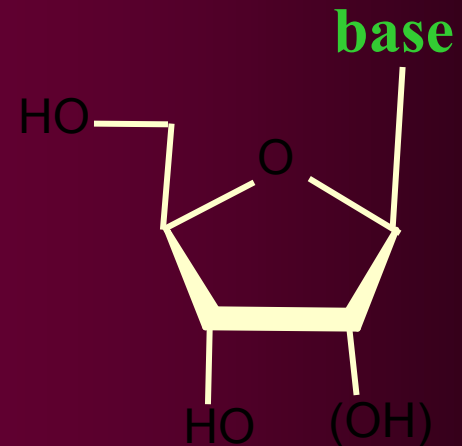
Nukleosides

Ribonukleosides

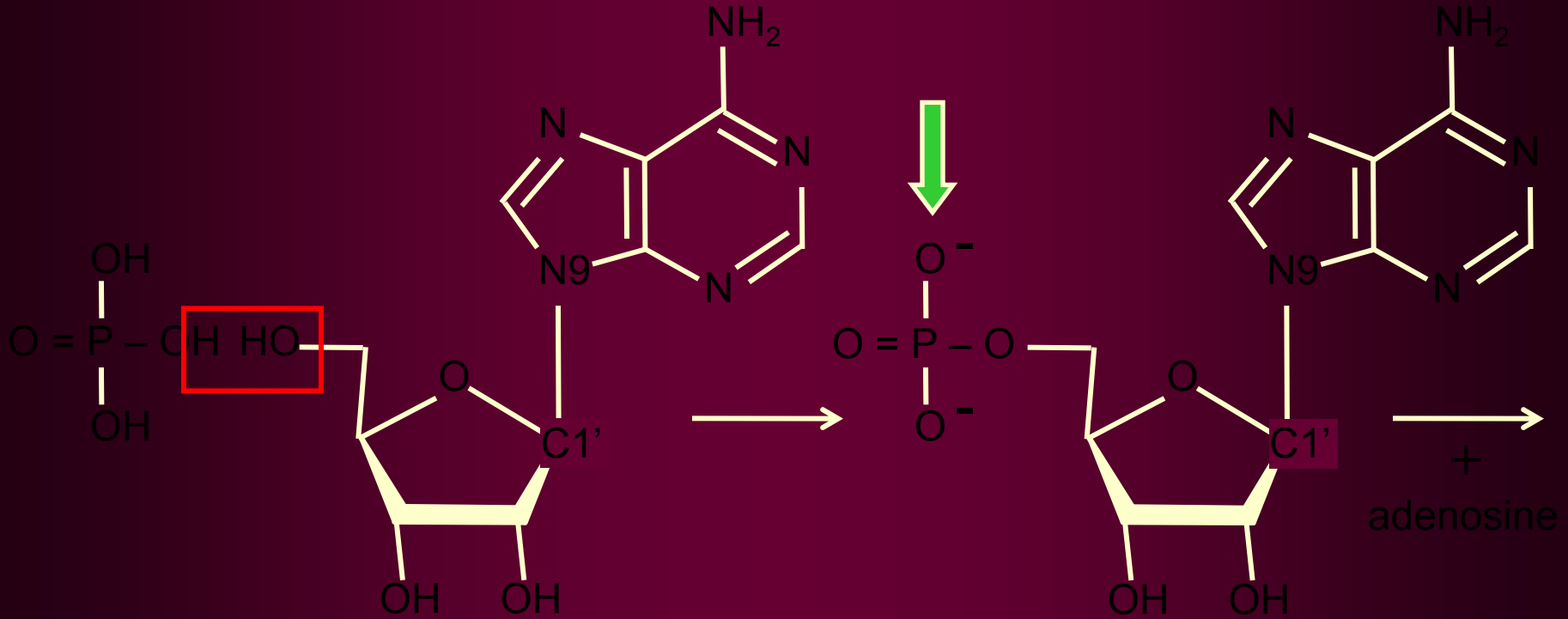
uridine	= U
cytidine	= C
adenosine	= A
guanosine	= G

Deoxyribonukleosides

deoxythymidine	= dT
deoxycytidine	= dC
deoxyadenosine	= dA
deoxyguanosine	= dG



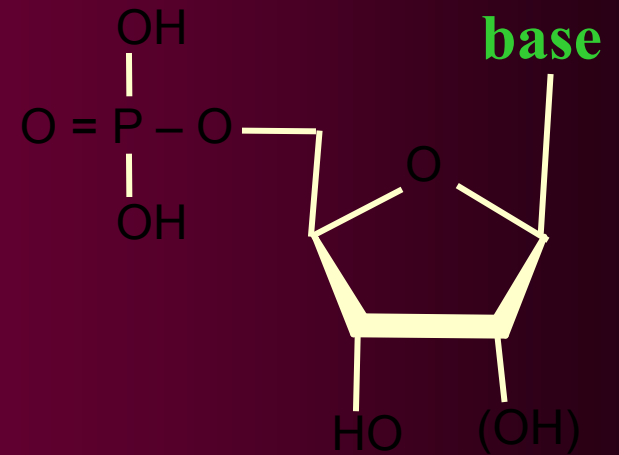
Phosphate group



acid + alcohol \longrightarrow ester

orthophosphoric acid H_3PO_4 + adenosine \longrightarrow adenosine(mono)phosphate (AMP)

Nukleotides



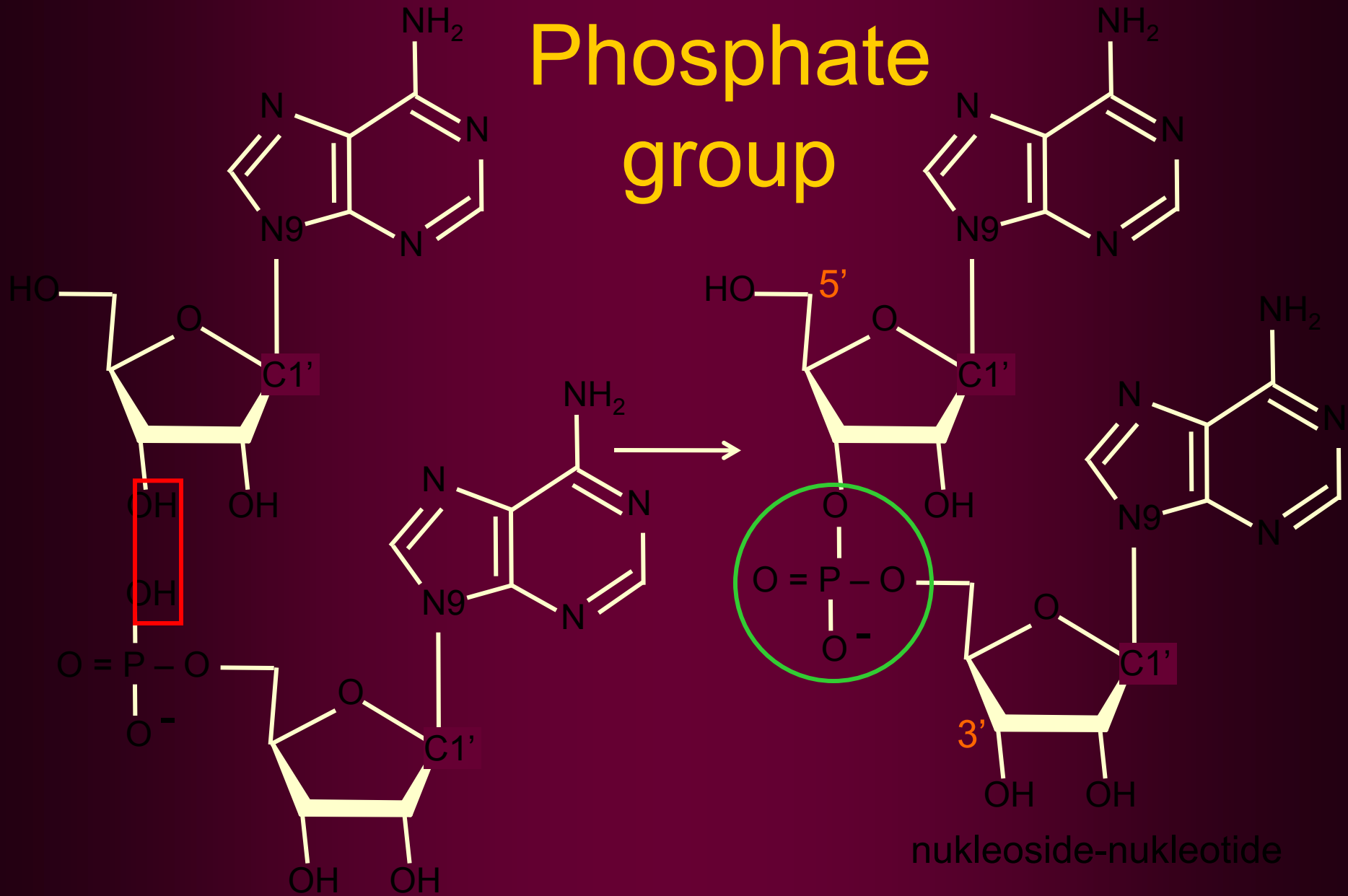
Ribonucleotides

uridyl acid	= uridine – 5' monophosphate	= UMP, pU
cytidyl acid	= cytidin –“–	= CMP, pC
adenyl acid	= adenosin –“–	= AMP, pA
guanyl acid	= guanosin –“–	= GMP, pG

Deoxyribonucleotides

deoxytymidyl acid	= 2' deoxythymidine-5'-monophosphate	= dTMP, pdT
deoxycytidyl acid	= –“– cytidin –“–	= dCMP, pdC
deoxyadenyl acid	= –“– adenosin –“–	= dAMP, pdA
deoxyguanyl acid	= –“– guanosin –“–	= dGMP, pdG

Phosphate group



alcohol

+

acid

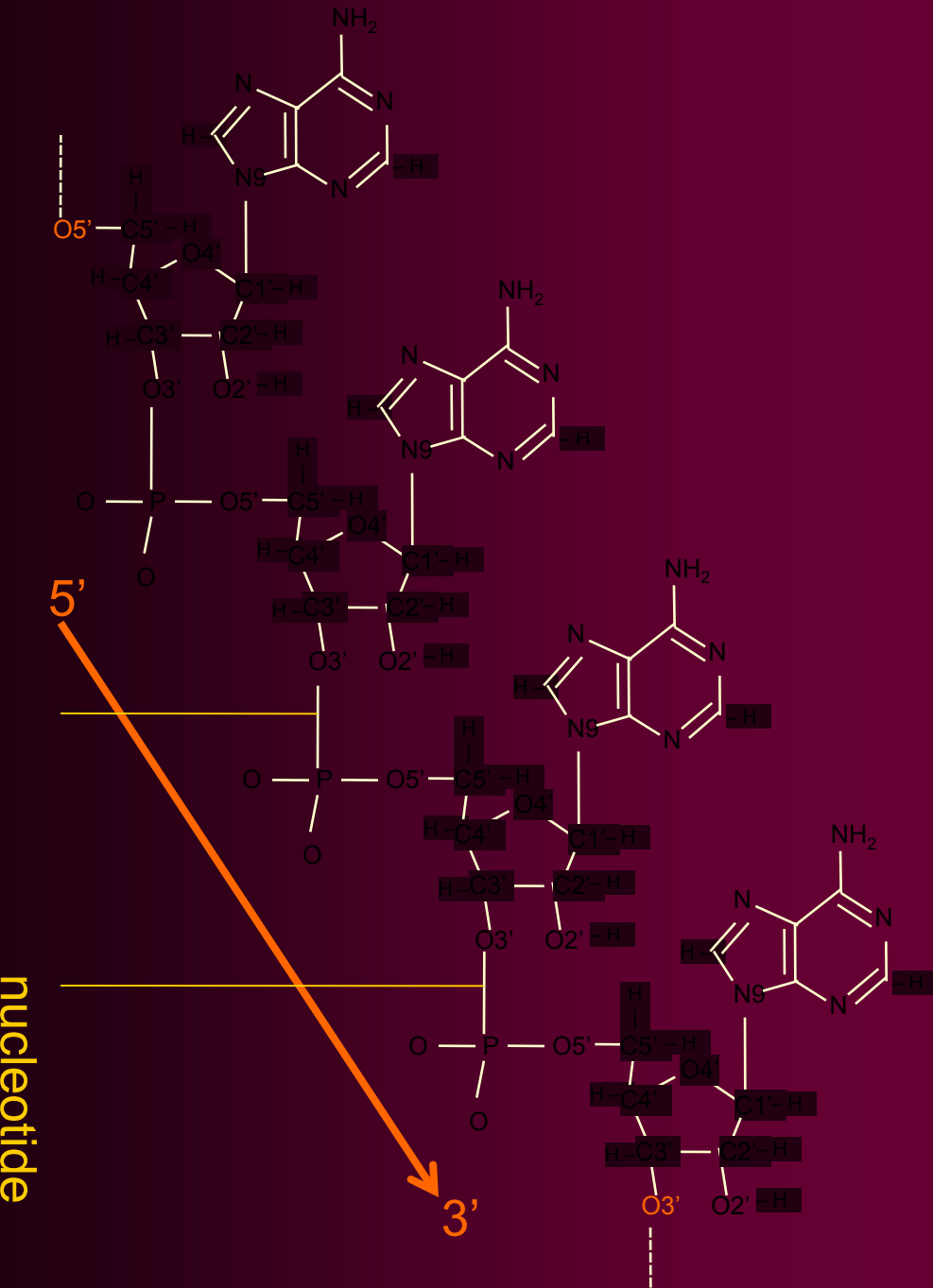
(ester)



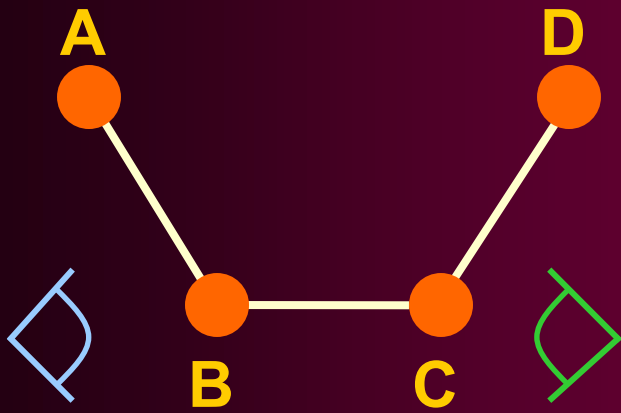
diester

ApA

Nucleotide chain

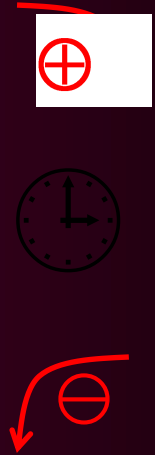
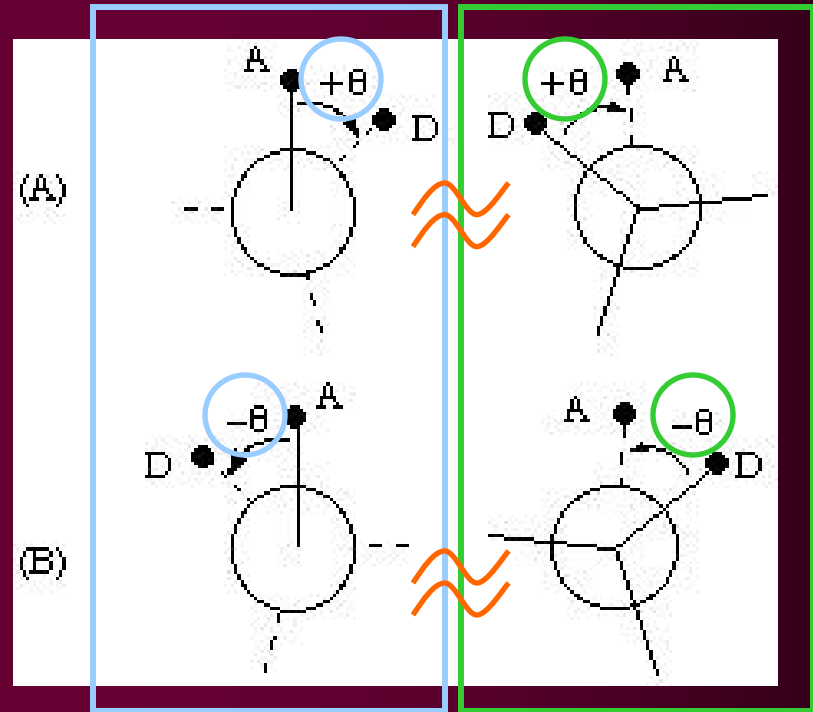


Torsion angle

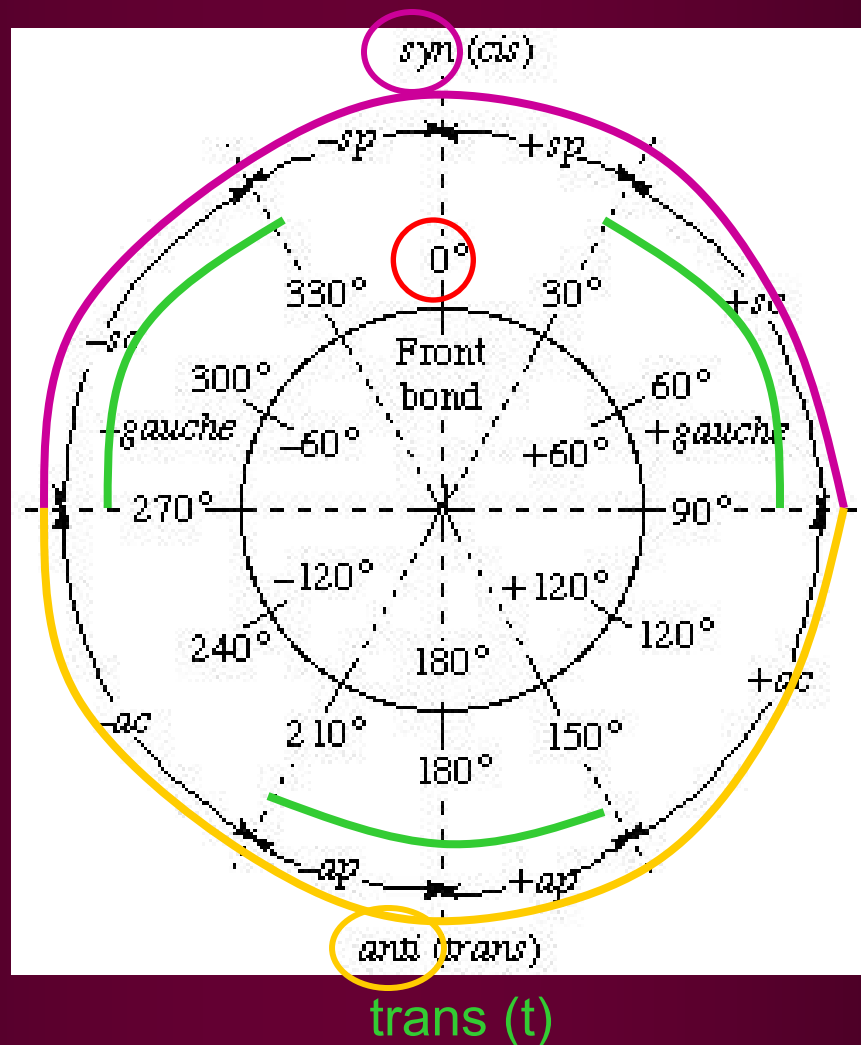


$\langle 0^\circ, 360^\circ \rangle$

$\langle -180^\circ, 180^\circ \rangle$



Torsion angle



synperiplanar (sp)

-gauche (-g)

synclinal (sc)

anticlinal (ac)

antiperiplanar (ap)

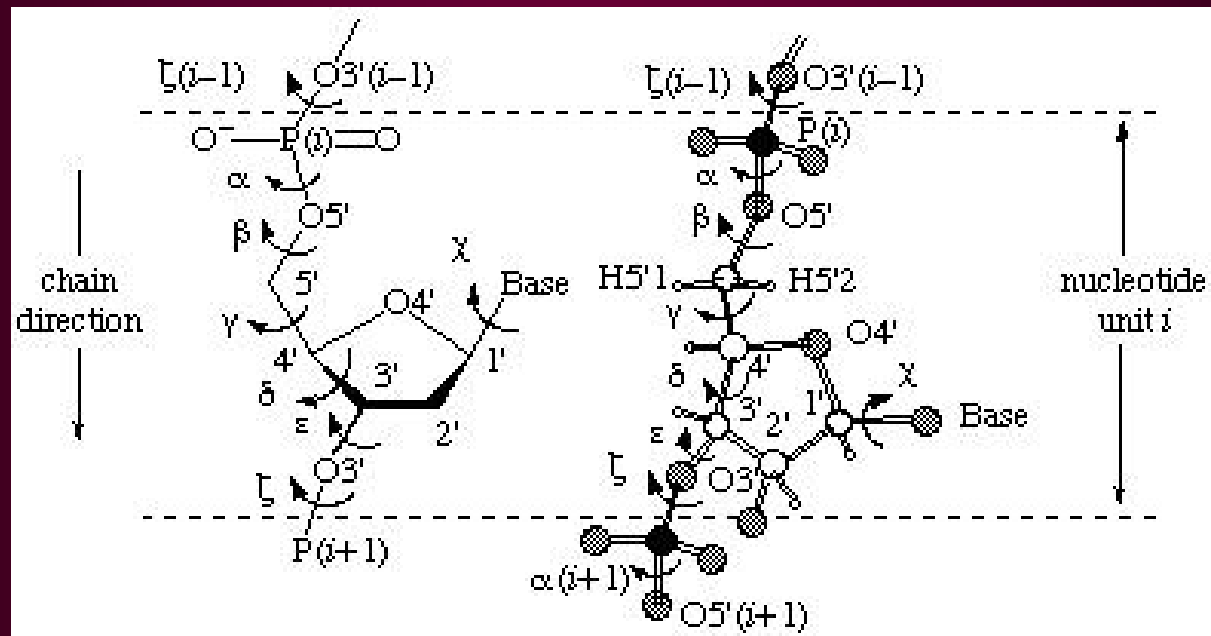
+gauche (+g)

anti (trans)

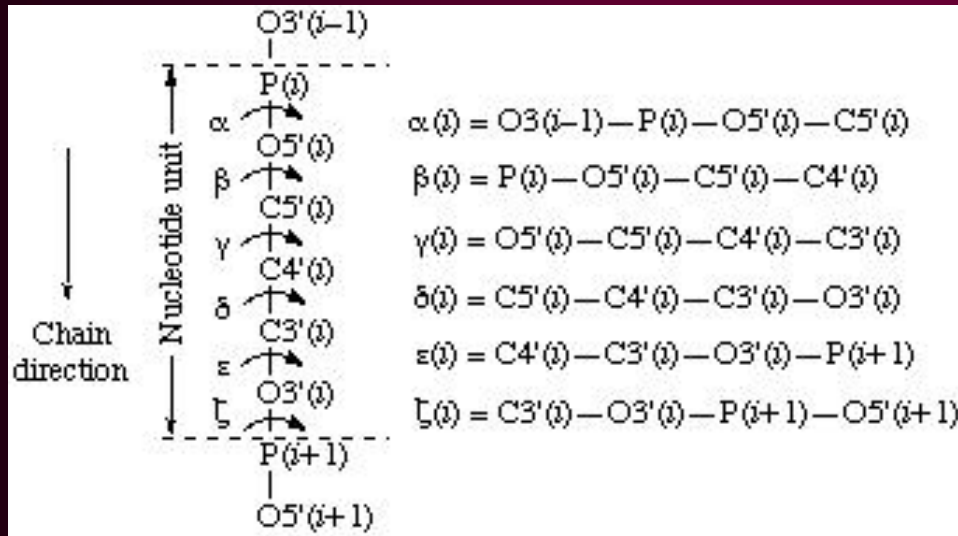
trans (t)

Torsion angles in NA

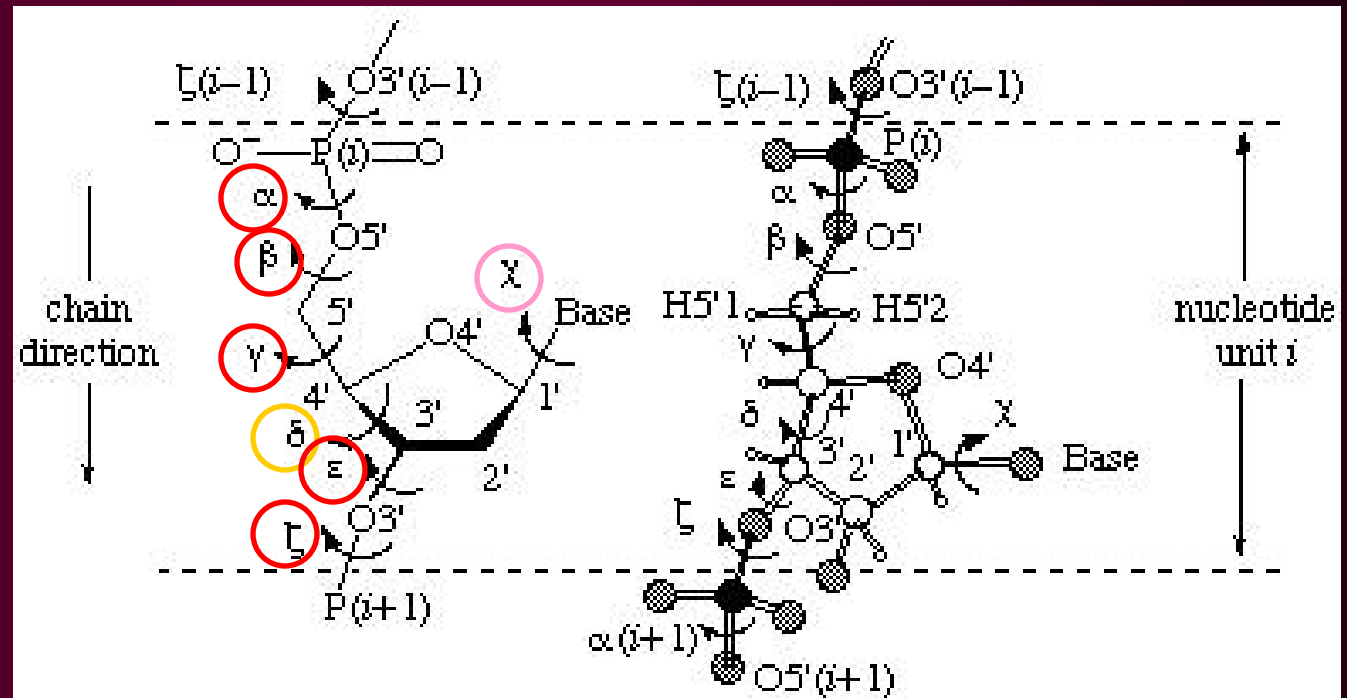
Sugar-phosphate backbone



Torsion angles cont.



α
 β
 γ
 δ
 ϵ
 ζ
 χ

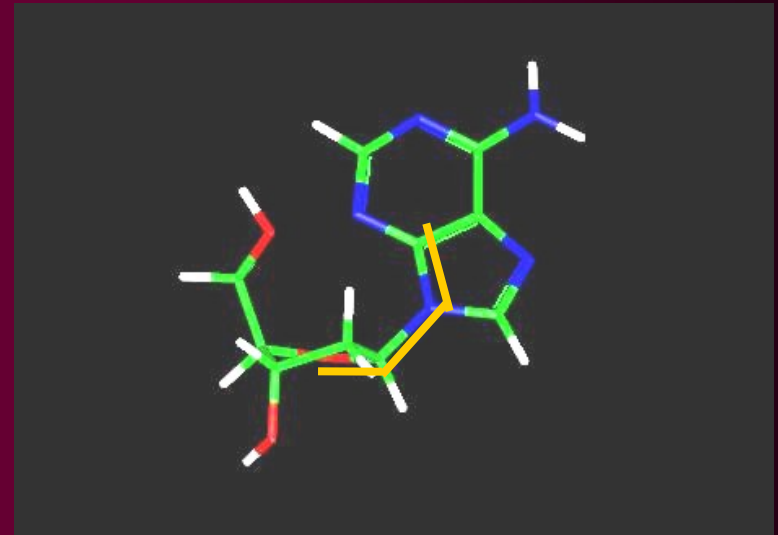
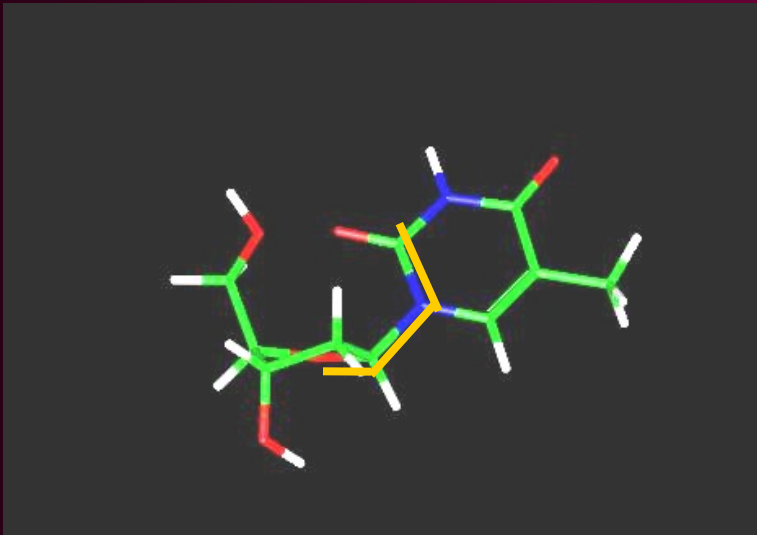


Torsion angle χ

SYN:

Pyrimidines: O2 above the sugar ring

Purines: 6-member purine ring above the sugar ring



Torsion angle χ

Orientation around the C1' – N glycosidic bond

O4' – C1' – N1 – C2 pyrimidines
 O4' – C1' – N9 – C4 purines

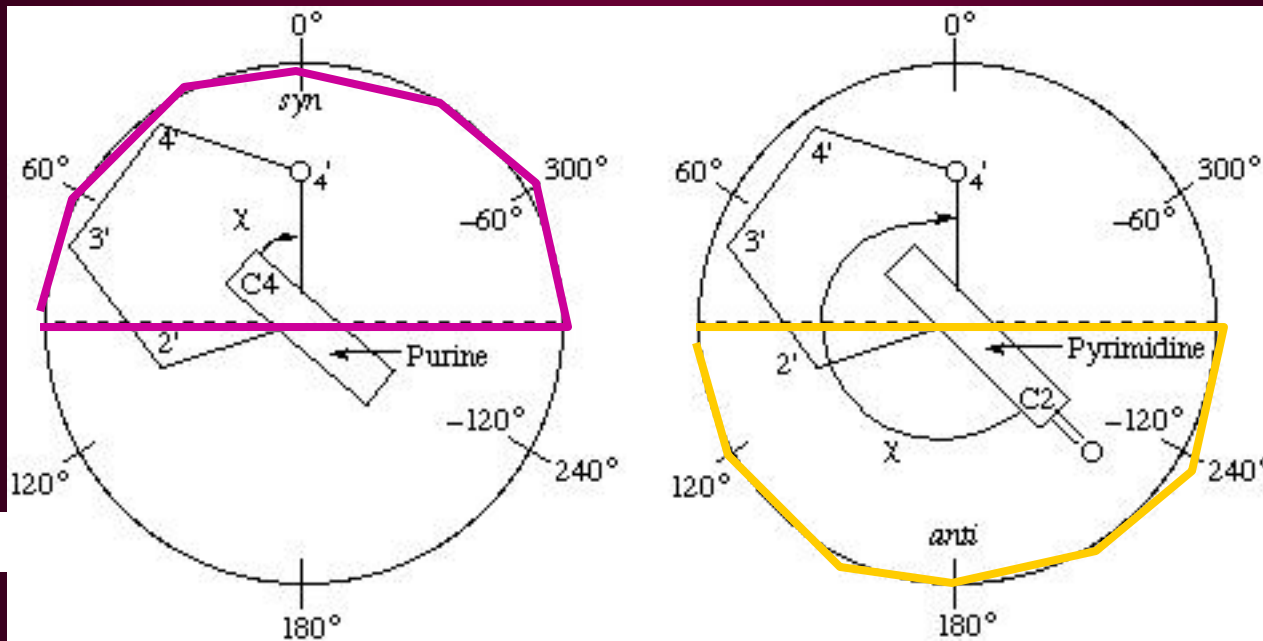
SYN

χ

$\langle 0^\circ, 90^\circ \rangle$

+

$\langle 270^\circ, 360^\circ \rangle$



ANTI

χ

$\langle 90^\circ, 270^\circ \rangle$

Torion χ – border intervals

high-syn (corresponds to +ac) ... $90^\circ +$ intrudes into anti
 high-anti (corresponds to -sc) ... $270^\circ +$ intrudes into syn

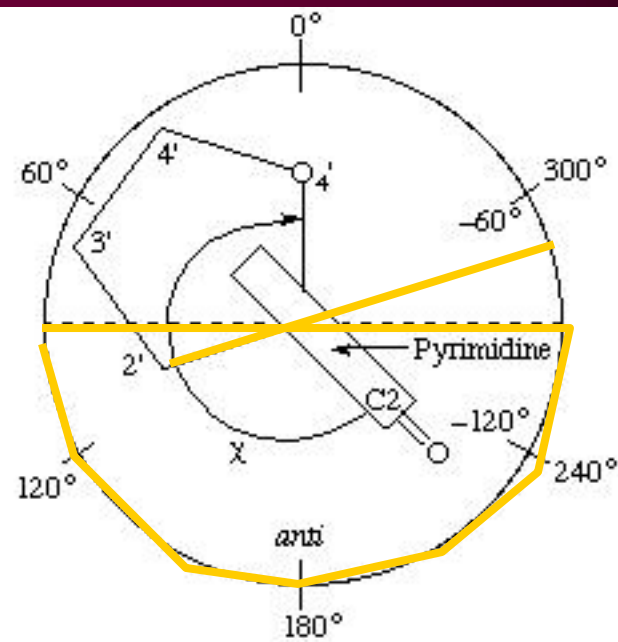
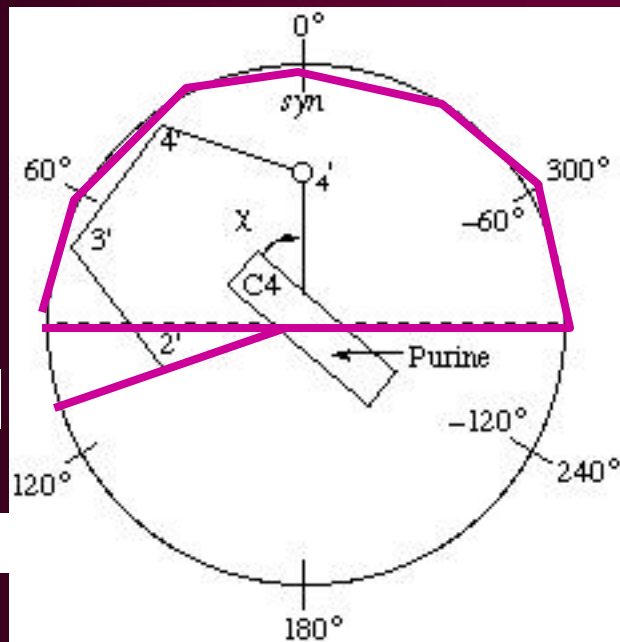
SYN

$\chi \in$

$\langle 0^\circ, 90^\circ \rangle$

+

$\langle 270^\circ, 360^\circ \rangle$



ANTI

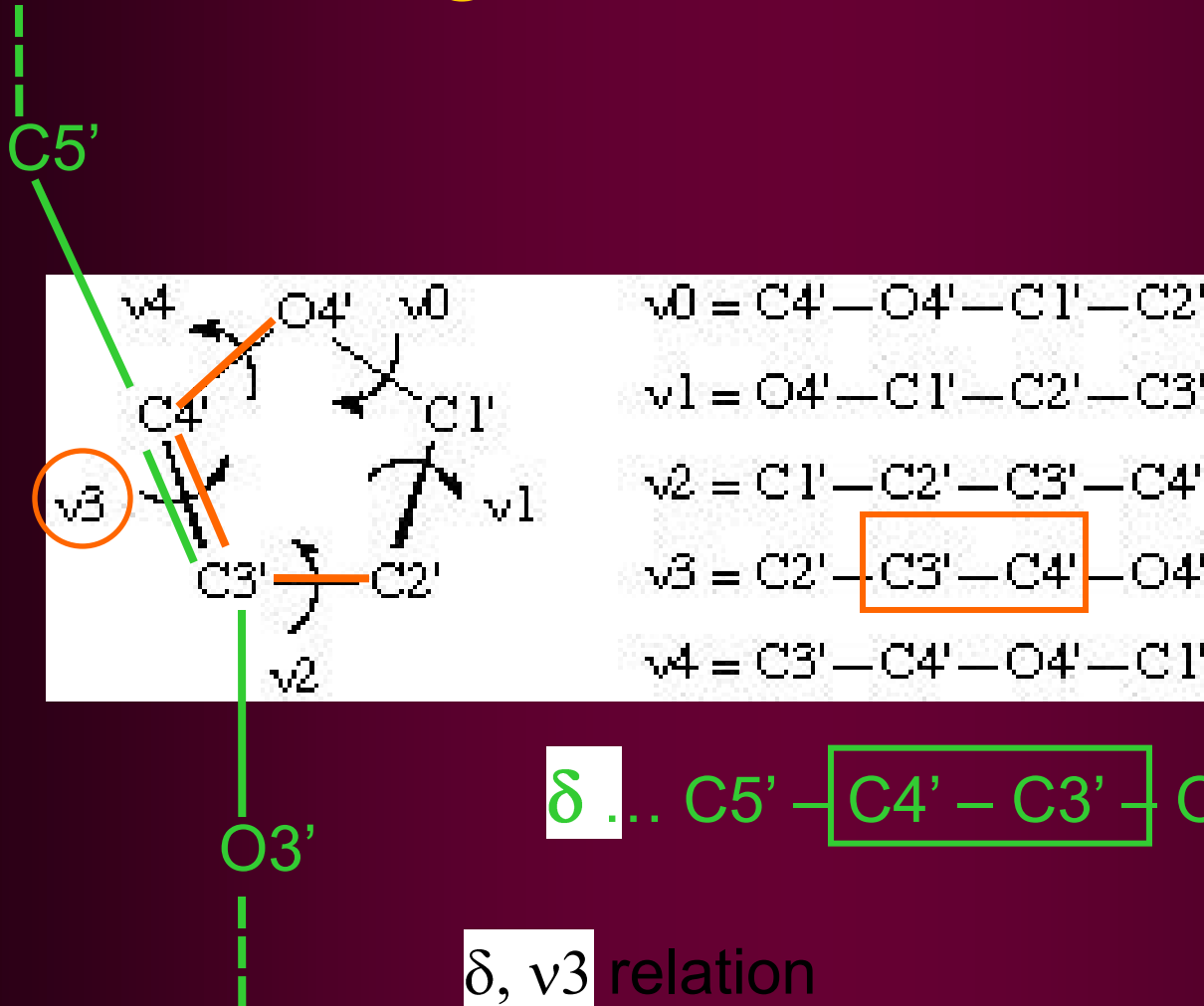
$\chi \in$

$\langle 90^\circ, 270^\circ \rangle$

Torsion angles in DNA

Angle	B-DNA	A-DNA
α	-40.7	-74.8
β	-135.6	-179.1
γ	-37.4	58.9
δ	139.5	78.2
ϵ	-133.2	-155.0
ζ	-156.9	-67.1
χ	-101.9	-158.9

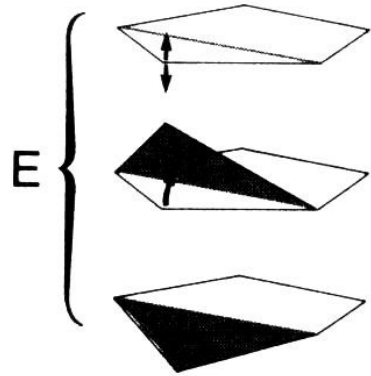
Sugar conformation



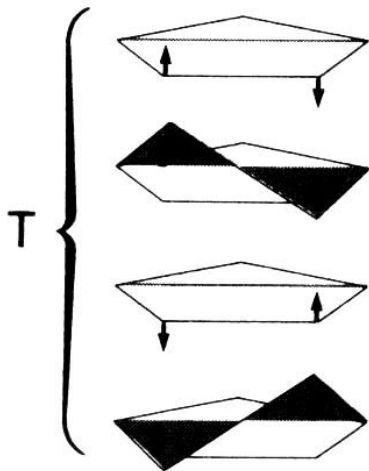
endocyclic

δ, v_3 relation

„Puckering“ of the sugar ring

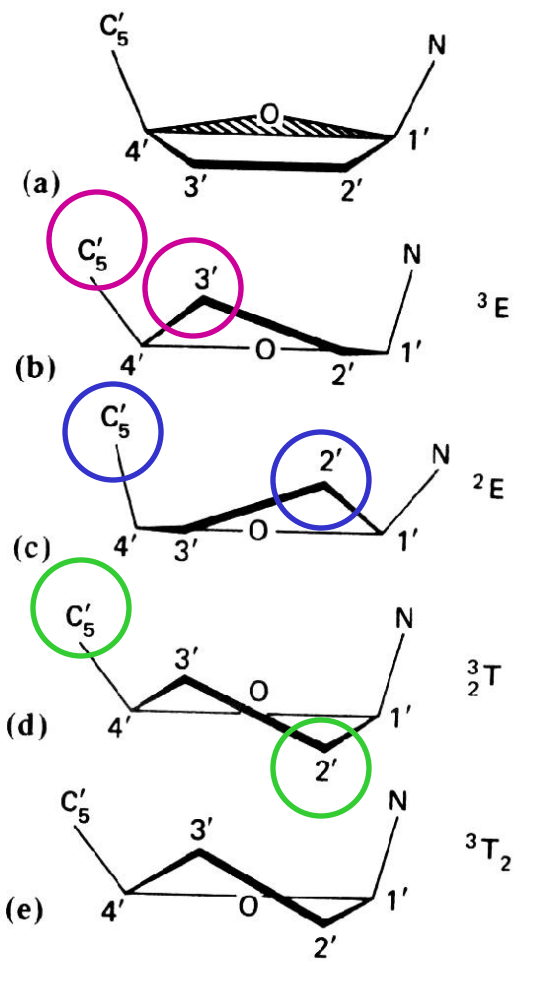


Envelope 4 atoms in a plane,
the 5th above or below



Twist 3 atoms in a plane, the 4th
and the 5th on the opposite
sides of the plane

Definition of the puckering modes



The sugar ring is not planar

C1' – O4' – C4' plane

With respect to C5'

- endo

- exo

Envelope C3'-endo 3E (prevalent in RNA)

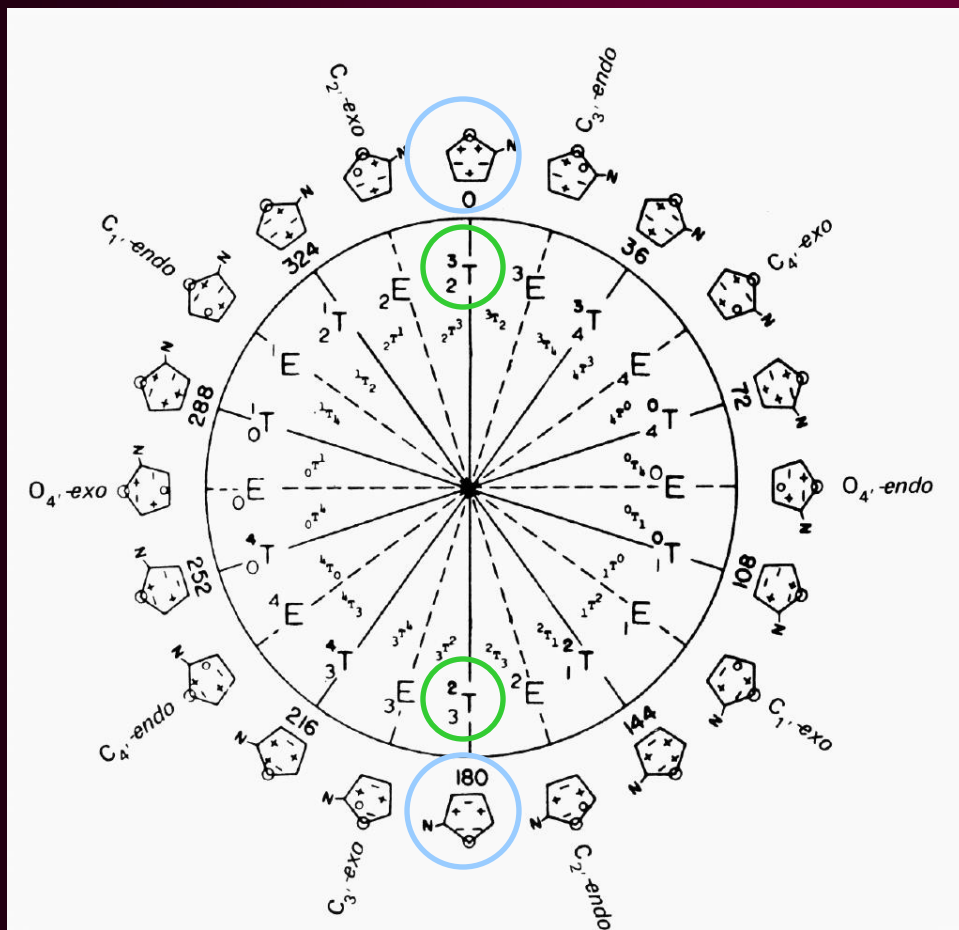
Envelope C2'-endo 2E (prevalent in DNA)

symmetric Twist C2'-exo-C3'-endo 3_2T

Non-symmetric Twist C3'-endo-C2'-exo 3T_2

Pseudorotation cycle

Theoretically – infinite number of conformations, can be characterized by maximum torsion angle (degree of pucker) and pseudorotation phase angle
 Torsion angles are not independent (ring closed)



Pseudorotation phase angle P

$$\tan P = \frac{(v_4 + v_1) - (v_3 + v_6)}{2 \cdot v_2 \cdot (\sin 36^\circ + \sin 72^\circ)}$$

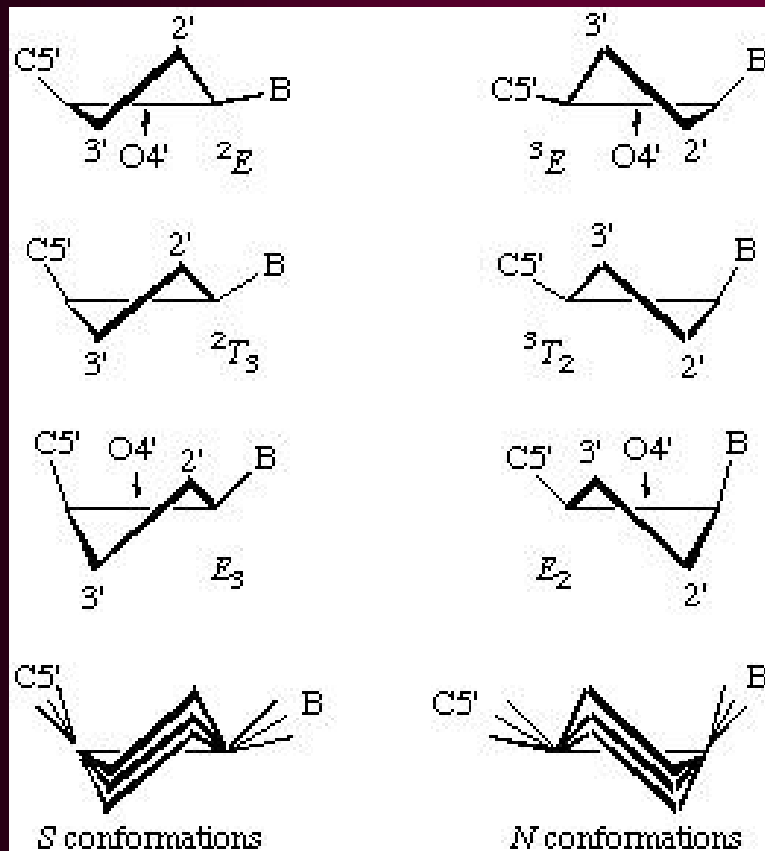
$P = 0^\circ$:

symmetric Twist C2'-exo-C3'-endo 3_2T

$P = 180^\circ$:

asymmetric Twist C2'-endo-C3'-exo 2_3T

v_{\max} amplitude



Maximum out-of-plane pucker

$$v_{\max} = v_2 / \cos(P)$$

P, v_j relation

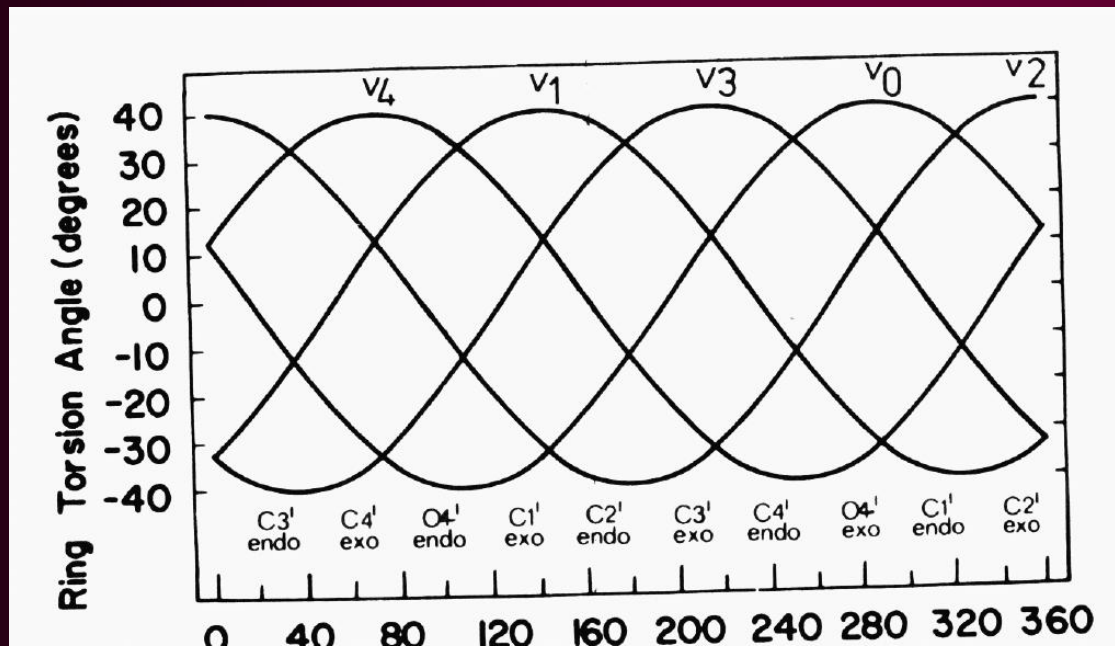
P value defines unambiguously all endocyclic torsion angles v_0 to v_4

$$v_2 = v_{\max} \cdot \cos(P + (j - 2) \cdot 144^\circ)$$

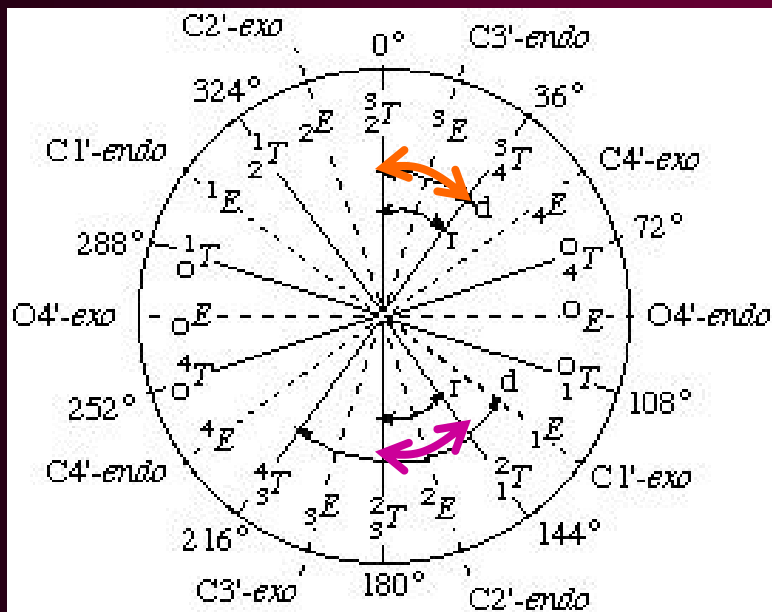
$$v_0 + v_1 + v_2 + v_3 + v_4 = 0$$

$$j = 0 \dots 4$$

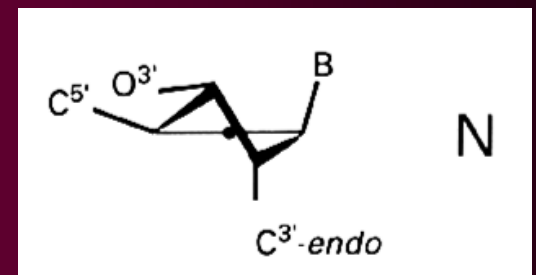
Sum of all 5 $v = 0$



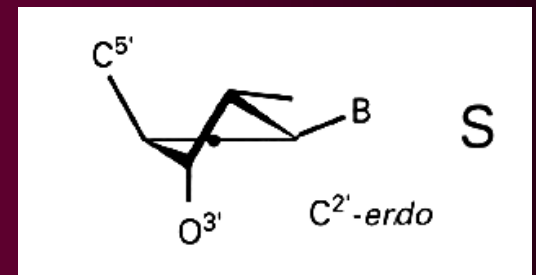
P in nucleic acids



NORTH



SOUTH



$0^\circ \leq P \leq 36^\circ$ north (prevalent in RNA)

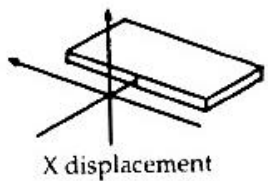
$144^\circ \leq P \leq 190^\circ$ south (prevalent in DNA)

Helical parameters

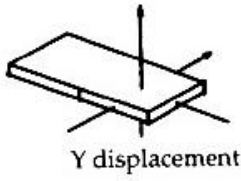
axis-base, axis-base pair

intra-base pair

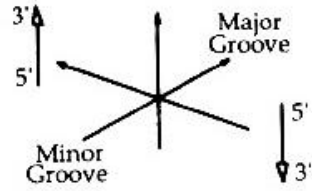
inter-base or inter-base pair



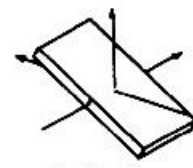
X displacement



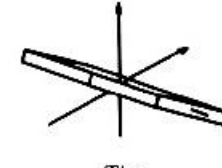
Y displacement



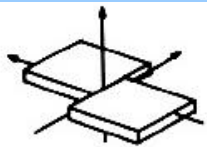
Major Groove
Minor Groove
3°
5°



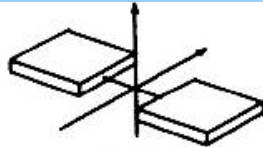
Inclination



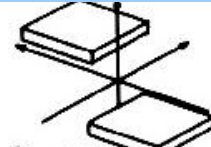
Tip



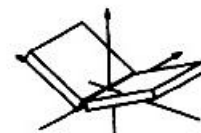
Shear



Stretch



Stagger



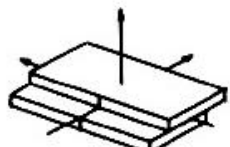
Buckle



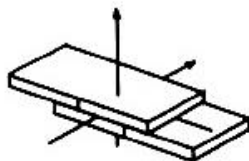
Propeller



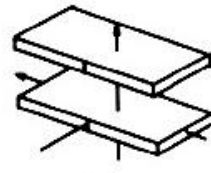
Opening



Shift



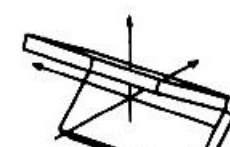
Slide



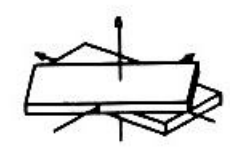
Rise



Tilt



Roll

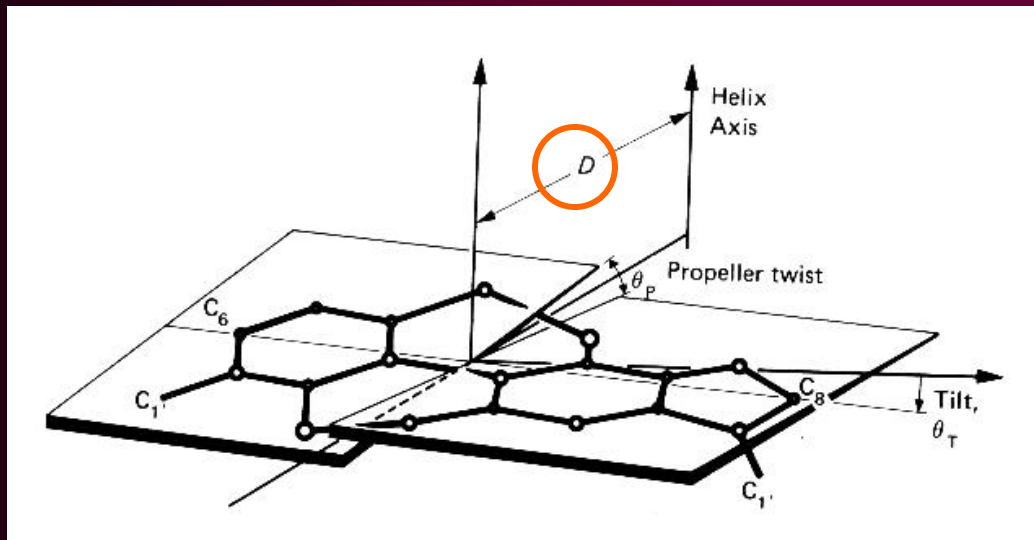


Twist

Distance/shift

Angle

Helical...

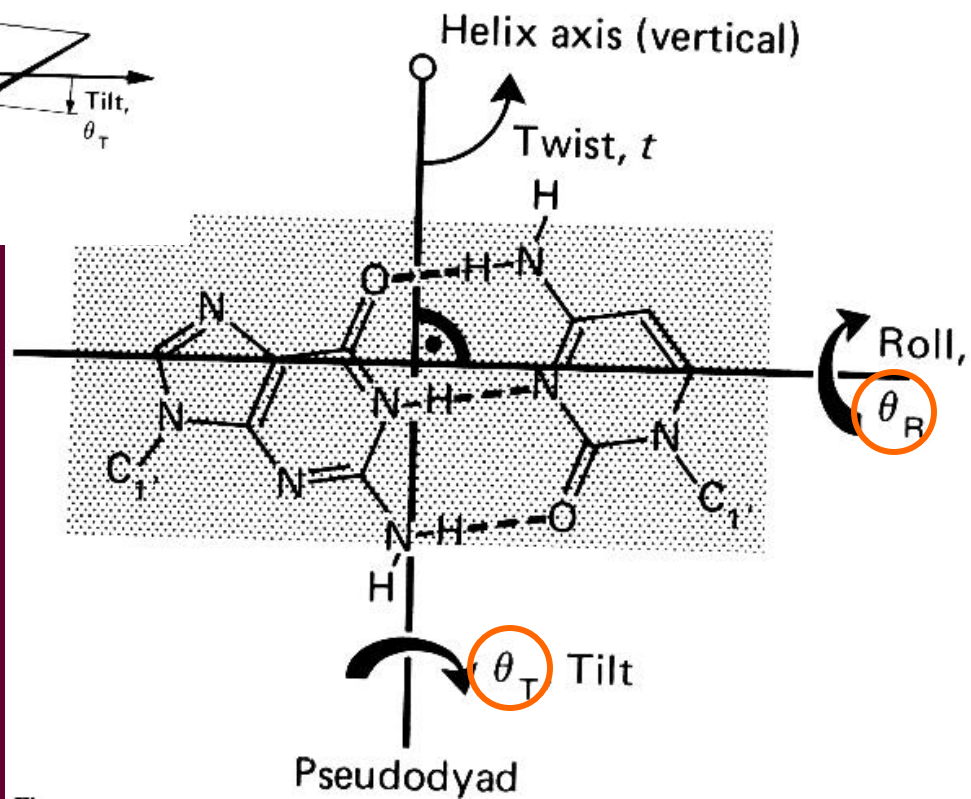


D ... displacement from helical axis

t ... twist = $360^\circ / n$

θ_R ... roll

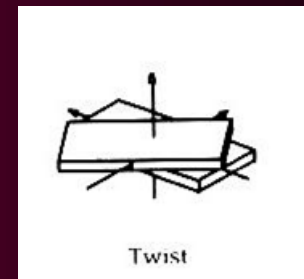
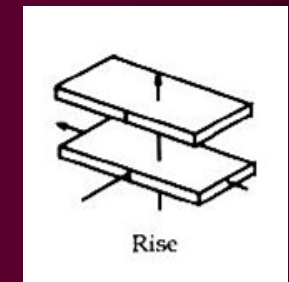
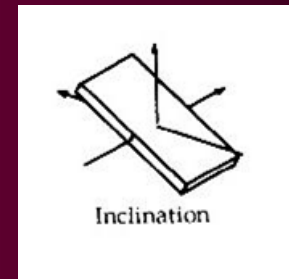
θ_T ... tilt



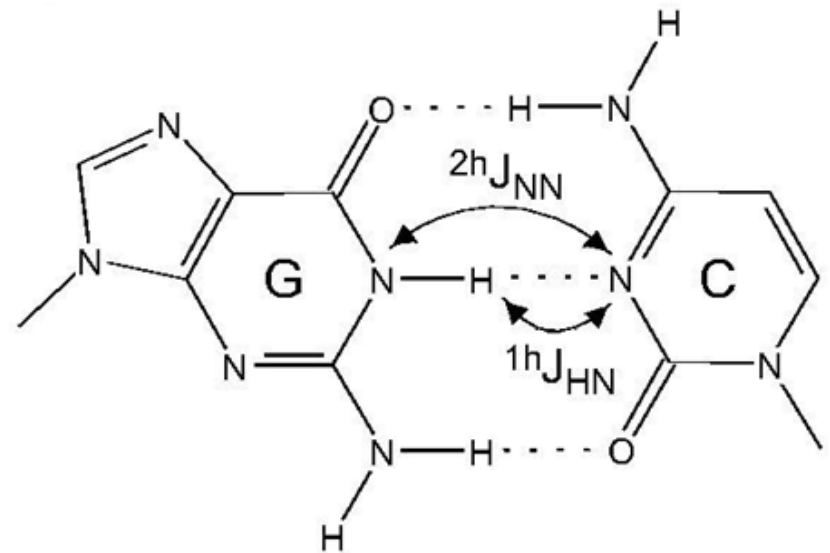
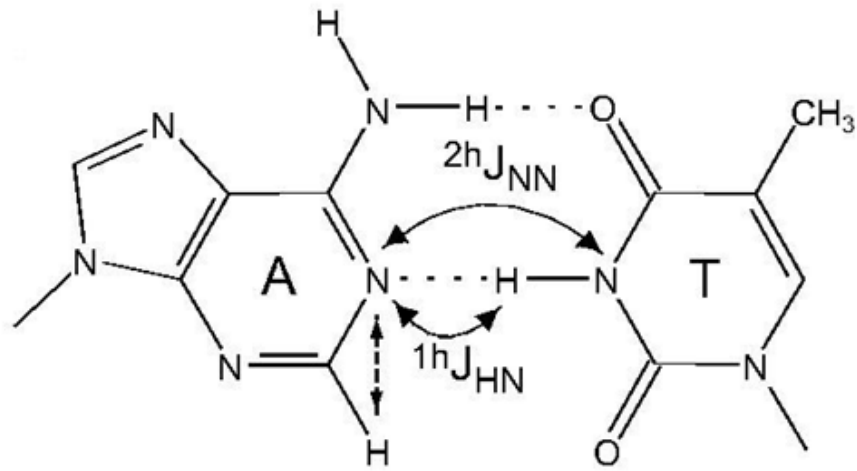
Helical parameters

for A and B DNA

Global	B-DNA	A-DNA	Shifts in Å, angles in degrees
X disp.	0.0	-5.28	
Y disp.	0.0	0.0	
Inclin	1.46	20.73	←
Tip	0.0	0.0	
Shear	0.0	0.0	
Stretch	0.0	0.01	
Stagger	-0.08	-0.04	
Buckle	0.0	0.0	
Propeller	-13.3	-7.5	
Opening	0.0	-0.02	
Shift	0.0	0.0	
Slide	0.0	0.0	
Rise	3.38	2.56	←
Tilt	0.0	0.0	
Roll	0.0	0.0	
Twist	36.00	32.70	
Bases per turn	360/36=10	360/32.7=11	←

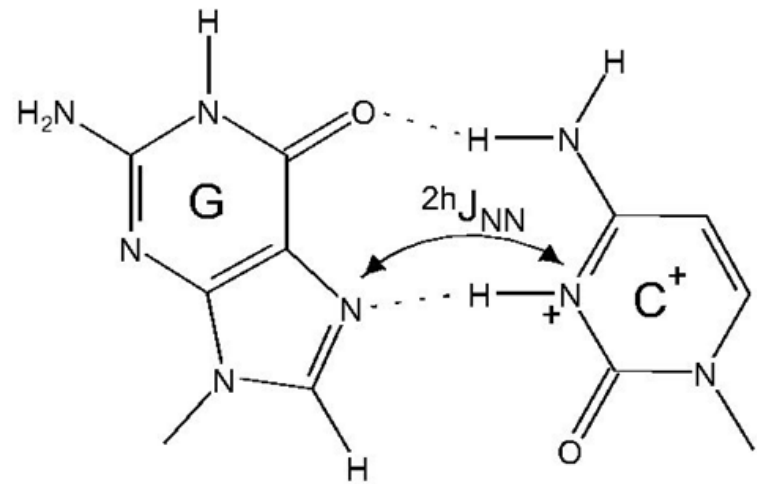
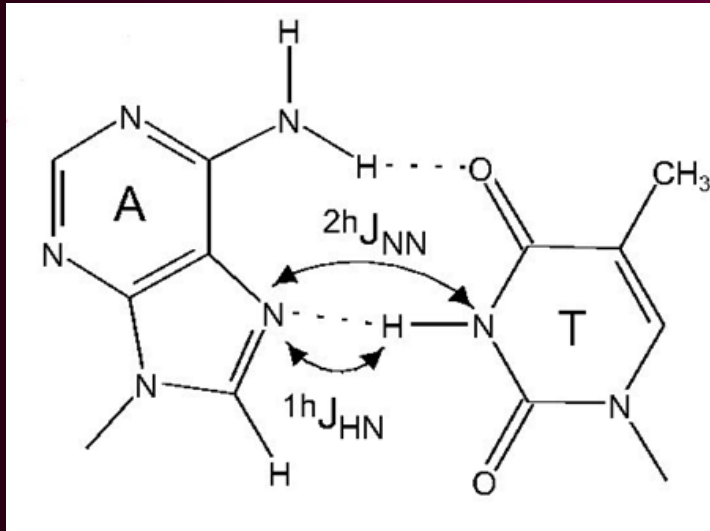


Base pairing

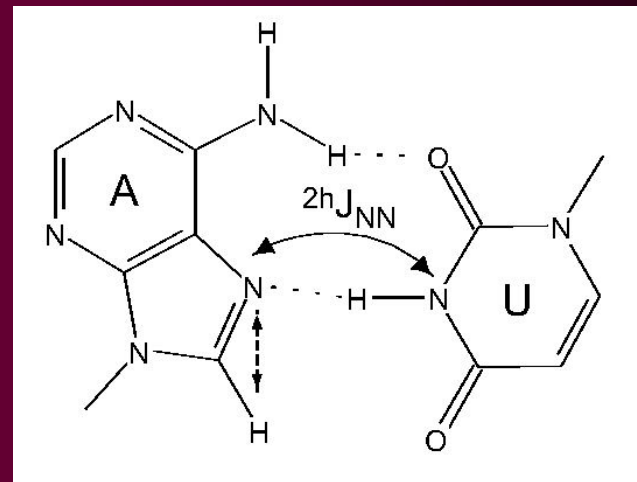


Watson-Crick pairs

Base pairing



Hoogsteen and
reverse Hoogsteen
pairs

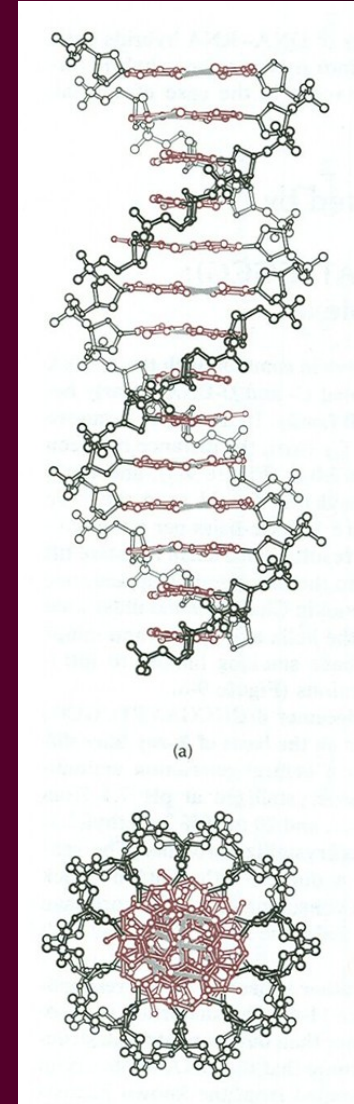
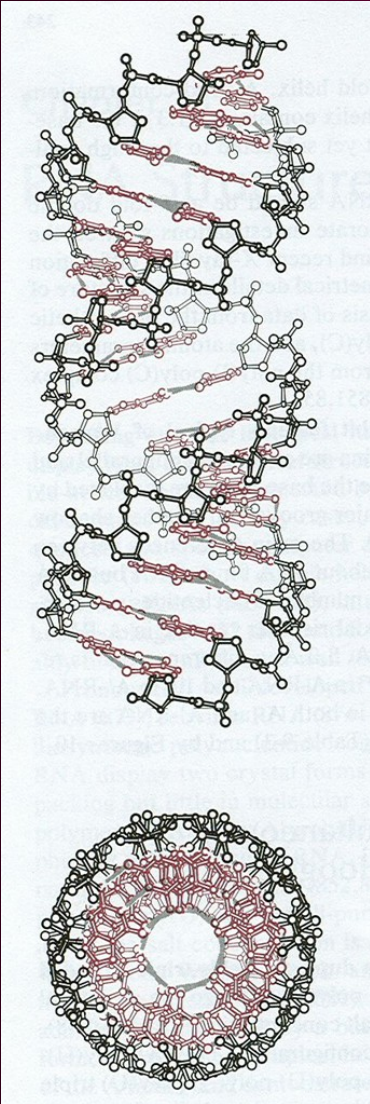


A and B double helix

A-RNA

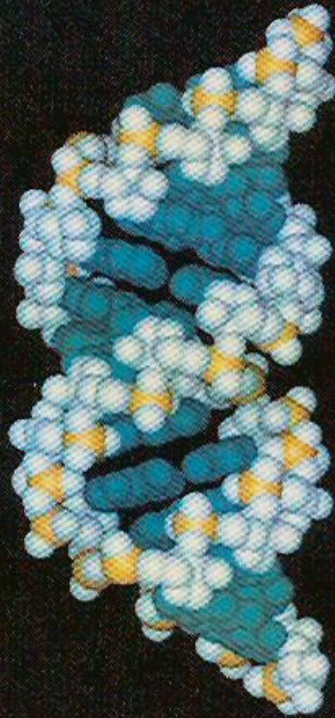
Ball and
stick
models

B-DNA



A and B helices

A-DNA



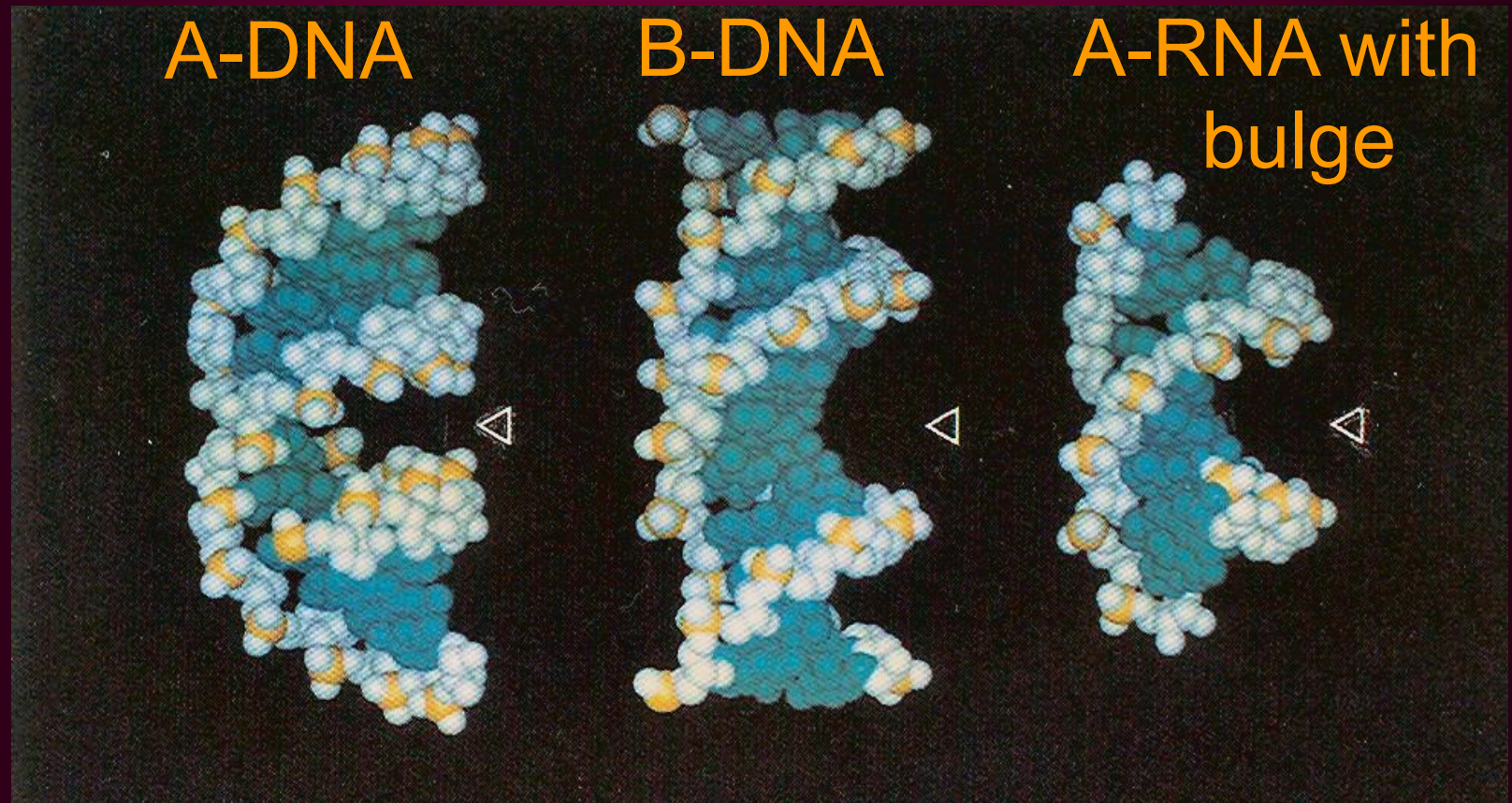
B-DNA



A-RNA
with
bulge



A and B helices



View tilted by 32° to show grooves

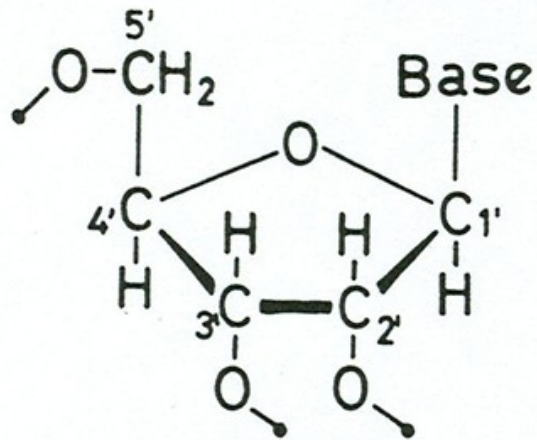
Nuclear properties of selected isotopes

Isotope ($I=1/2$)	$\gamma \times 10^{-7}$ ($\text{rad T}^{-1}\text{s}^{-1}$)	ν at 11.74T (MHz)	Natural Abundance (%)	Sensitivity	
				Rel. ^a	Abs. ^b
^1H	26.75	500.0	99.98	1.00	1.00
^{13}C	6.73	125.7	1.11	1.6×10^{-2}	1.8×10^{-4}
^{15}N	-2.71	50.7	0.37	1.0×10^{-3}	3.8×10^{-6}
^{31}P	10.83	202.4	100	6.6×10^{-2}	6.6×10^{-2}

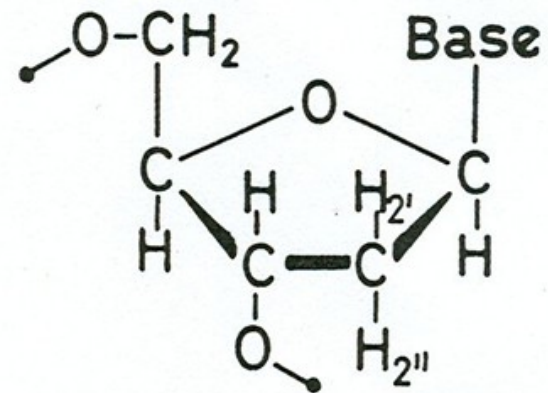
^a Relative sensitivity at constant field for equal number of nuclei.

^b Product of relative sensitivity and natural abundance.

Spin systems in ribose and deoxyribose

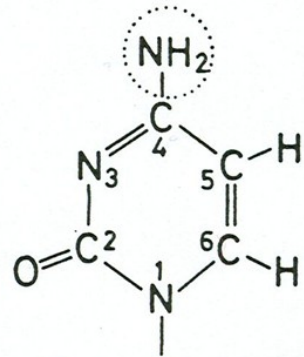


β -D-Ribose
XWTPMA

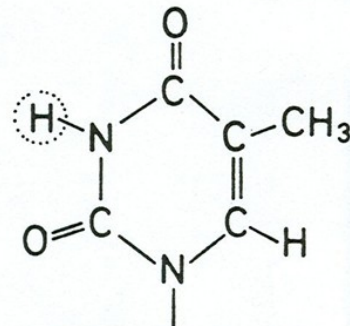


2'-Deoxy- β -D-Ribose
XAMWTNP

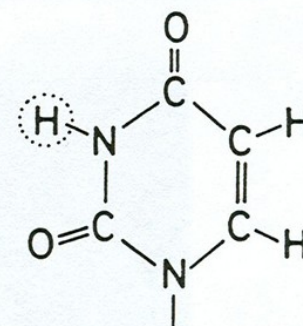
Spin systems in nucleic acid bases



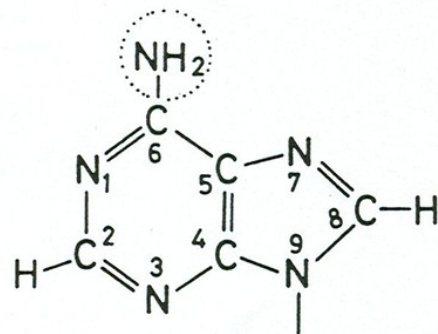
Cytosine, C
AX



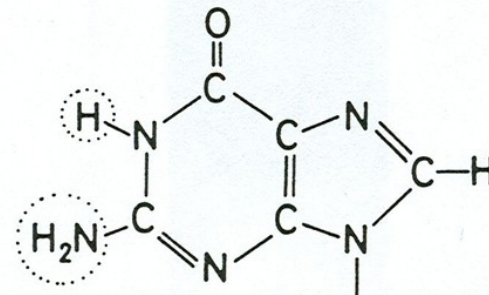
Thymine, T
A₃X



Uracil, U
AX

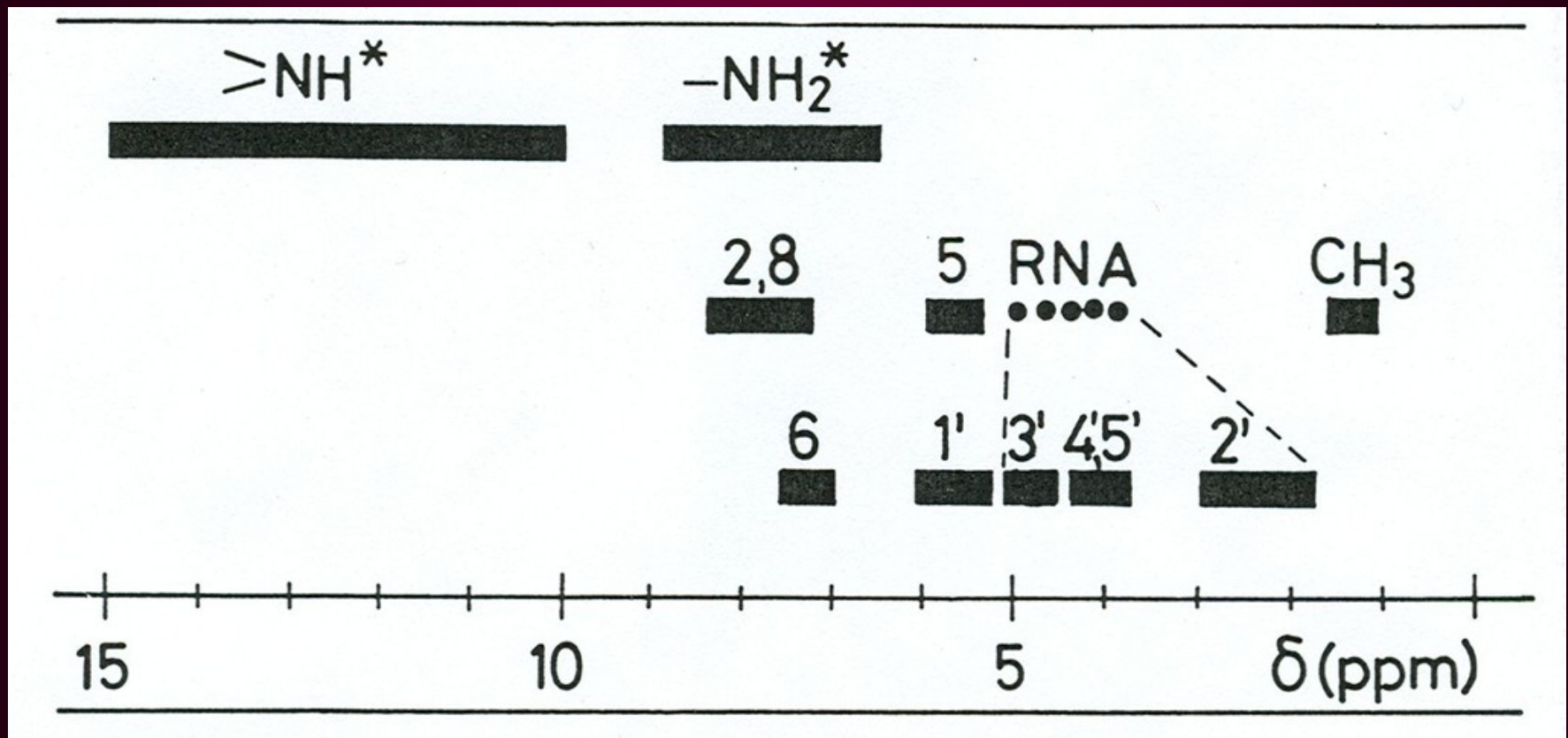


Adenine, A
A + A



Guanine, G
A

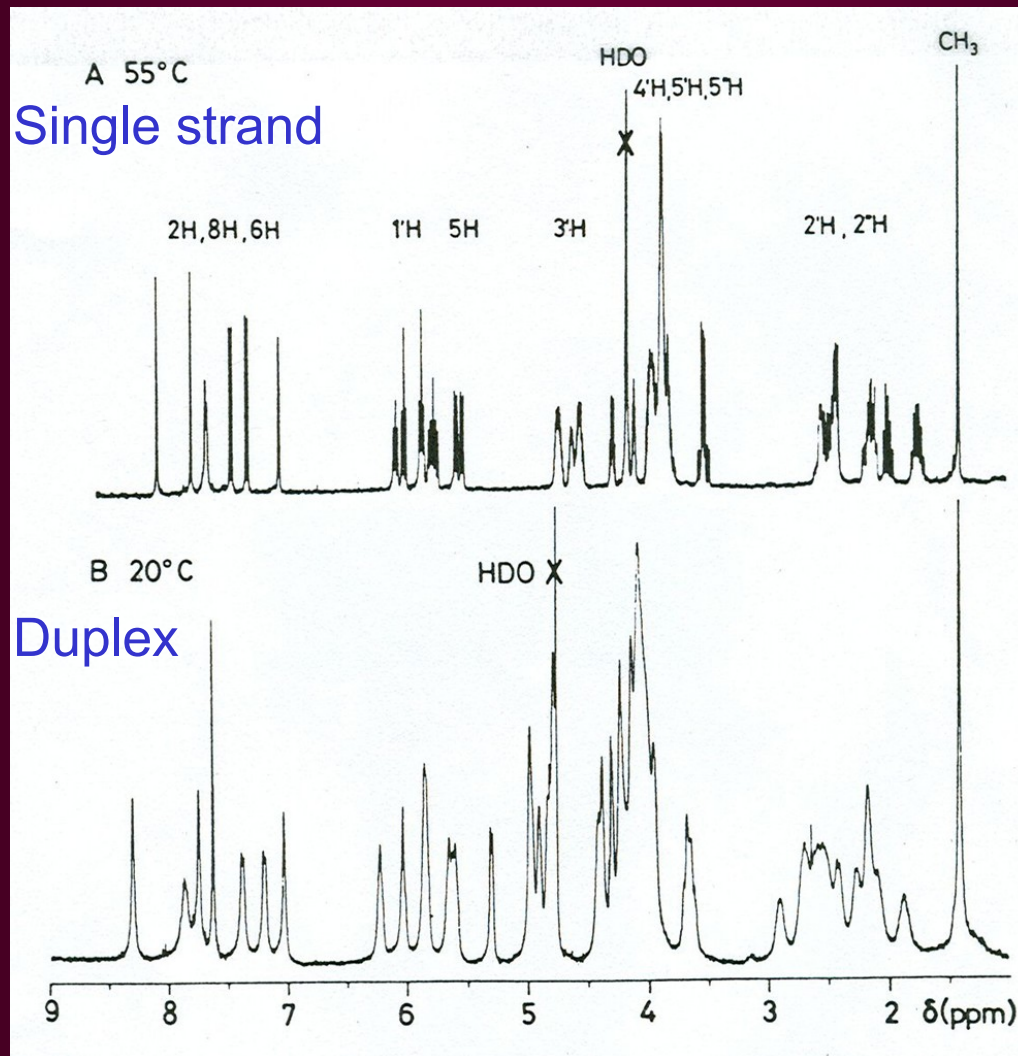
^1H chemical shift ranges in DNA and RNA



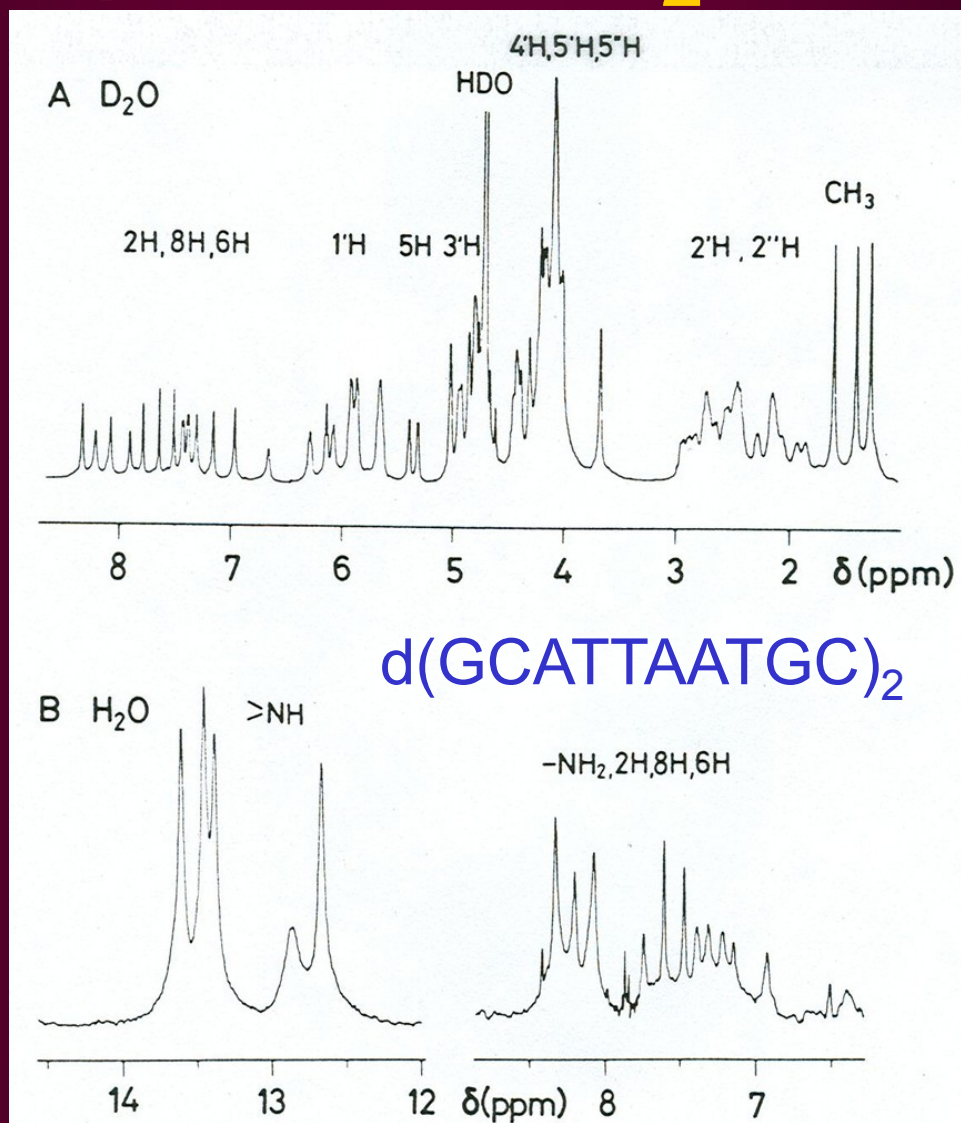
^1H chemical shift ranges in DNA and RNA

Code	δ (ppm)	Comments
2'	1.8-3.0	2'H, 2''H in DNA
4',5'	3.7-4.5	4'H, 5'H, 5''H in DNA
3'	4.4-5.2	3'H in DNA
•••••	3.7-5.2	2'H, 3'H, 4'H, 5'H, 5''H in RNA
1'	5.3-6.3	1'H
CH_3	1.2-1.6	CH_3 of T
5	5.3-6.0	5H of C and U
6	7.1-7.6	6H of C, T and U
2,8	7.3-8.4	8H of A and G, 2H of A
$-\text{NH}_2^*$	6.6-9.0	NH_2 of A, C and G
$>\text{NH}^*$	10 - 15	Ring NH of G, T and U

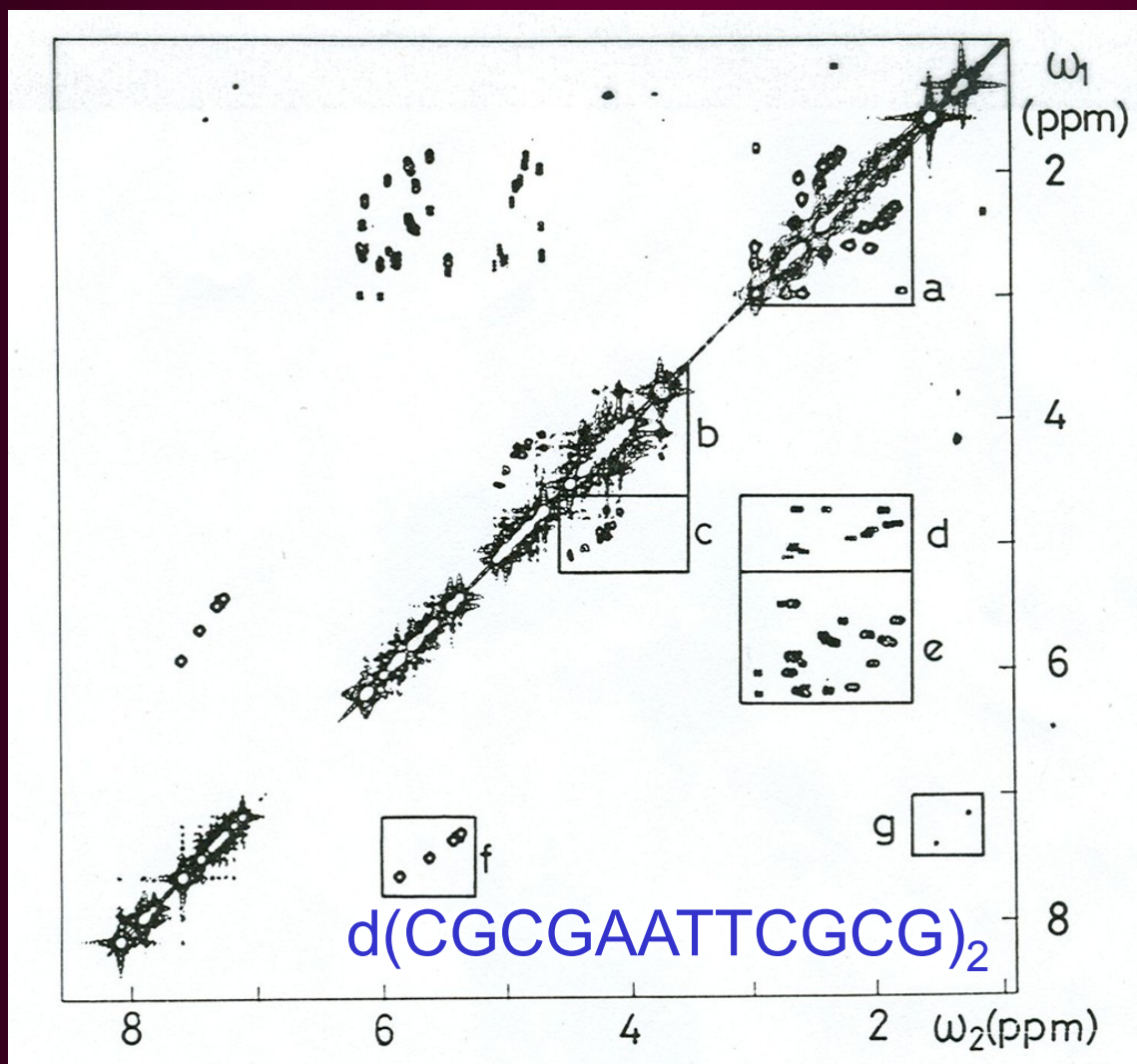
^1H NMR spectra of d(GCATGC)



^1H NMR spectra in D_2O and H_2O

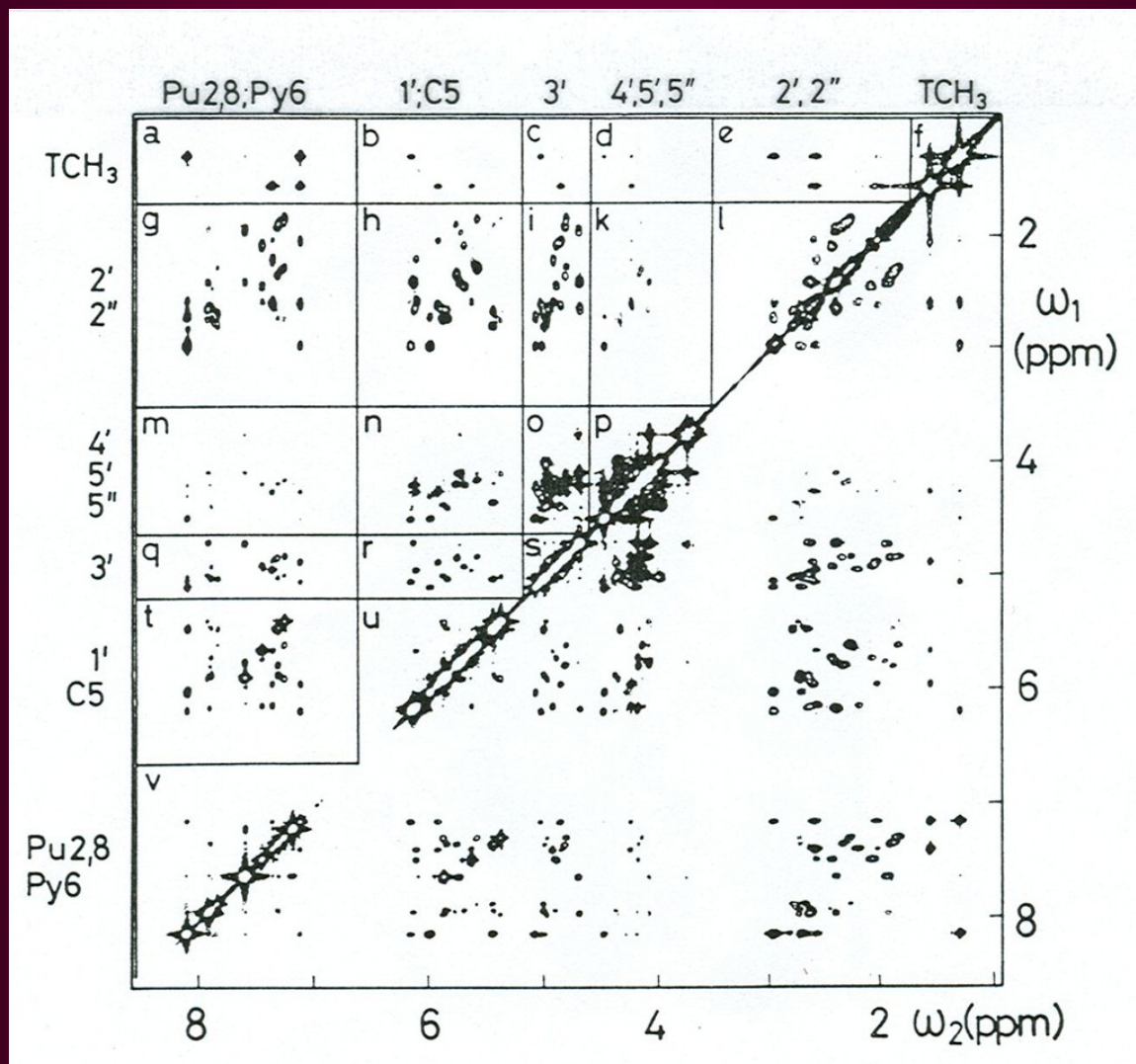


^1H COSY spectrum of DNA



- a H2'-H2''
- b H4'-H5',5''
H5'-H5''
- c H3'-H4'
- d H2',2''-H3'
- e H1'-H2',2''
- f H5-H6 (Cyt)
- g CH₃-H6 (Thy)

^1H NOESY spectrum of DNA



d(CGCGA

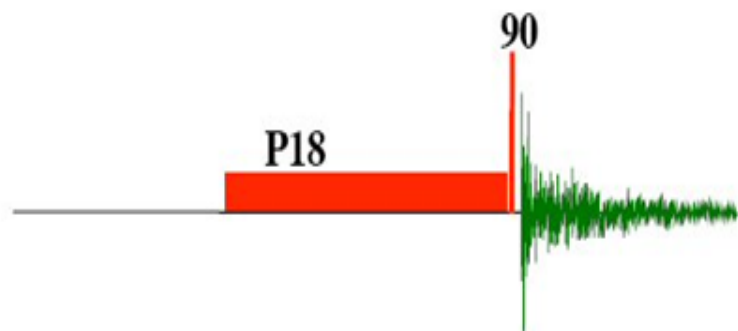
Water Suppression

The presence of an intense solvent resonance necessitates an impractical high dynamic range. **110 M vs <1mM**

To overcome this problem several methods are currently applied:

- 1) Presaturation.
- 2) Observing the FID when the water passes a null condition after a 180 degree pulse.
- 3) Suppression of broad lines based on their T_2 behavior.
- 4) Selective excitation, with and without gradients
- 5a) Use of GRASP to select specific coherences thereby excluding the intense solvent signal. In this case the solvent signal never reaches the ADC. This allows the observation of resonances that are buried under the solvent peak.
- 5b) Use of GRASP to selectively dephase the solvent resonance (WATERGATE)

PRESAT

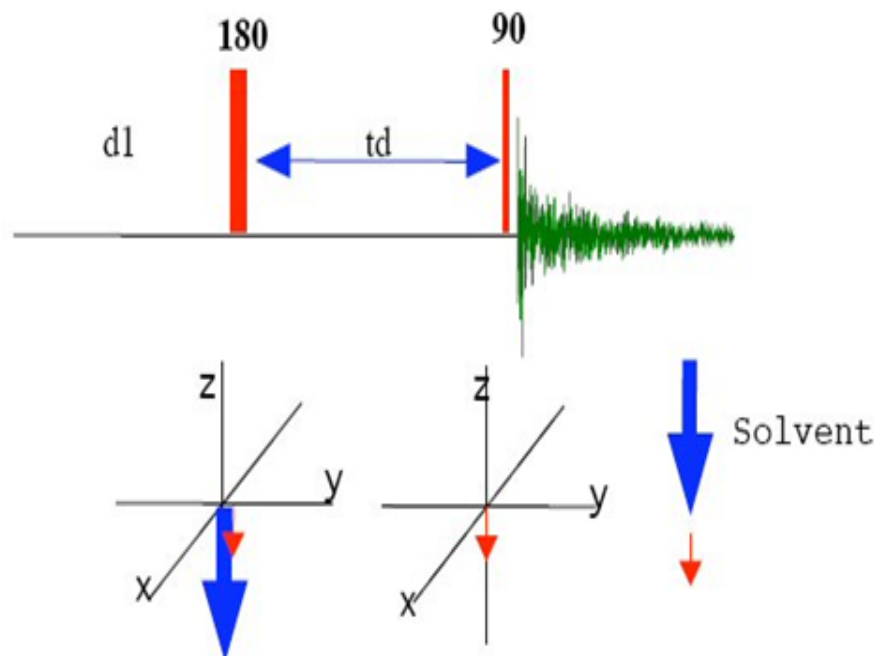


Presaturation field strength:
20-40 Hz corresponds to a
6-12ms 90deg pulse.

Pros: Easy to set up
Excellent water suppression

Cons: Resonances under water signal!
(T variation)
Labile protons not visible
(some GC pairs may be)

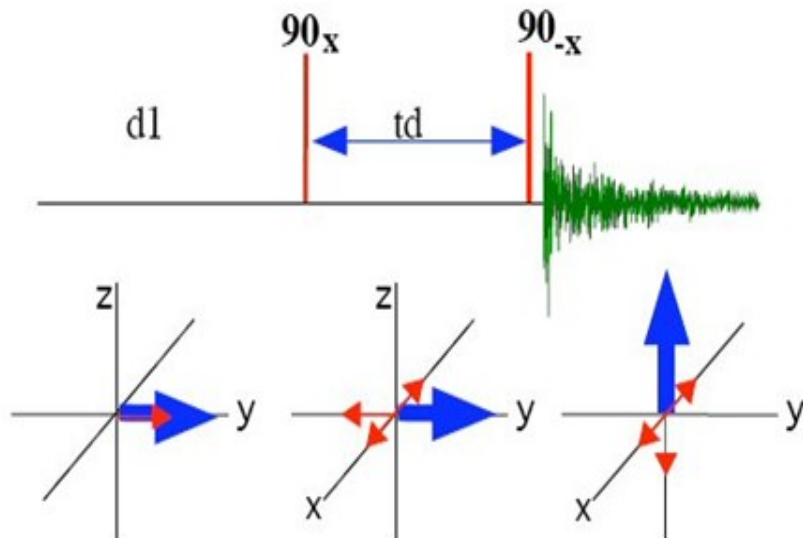
WEFT



Method relies on different T_1 values for
water and solute.

It fails if the relaxation times are similar.
Intensity of the solute resonances may vary.
For a selective 180 degree pulse on the
solvent these problems are largely avoided.

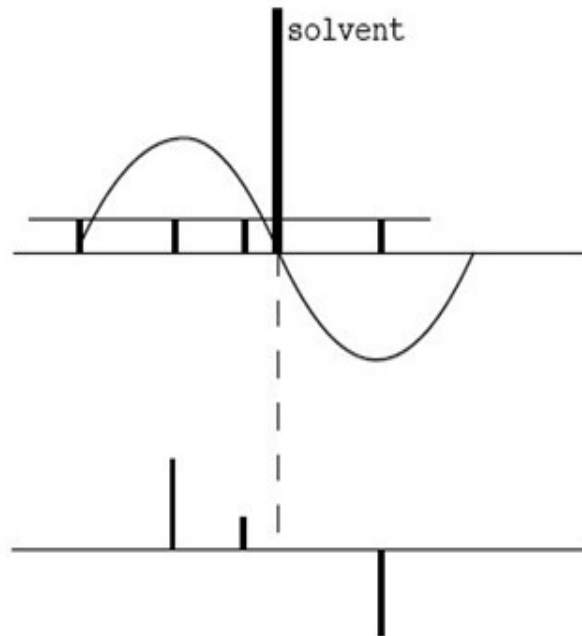
Jump and return



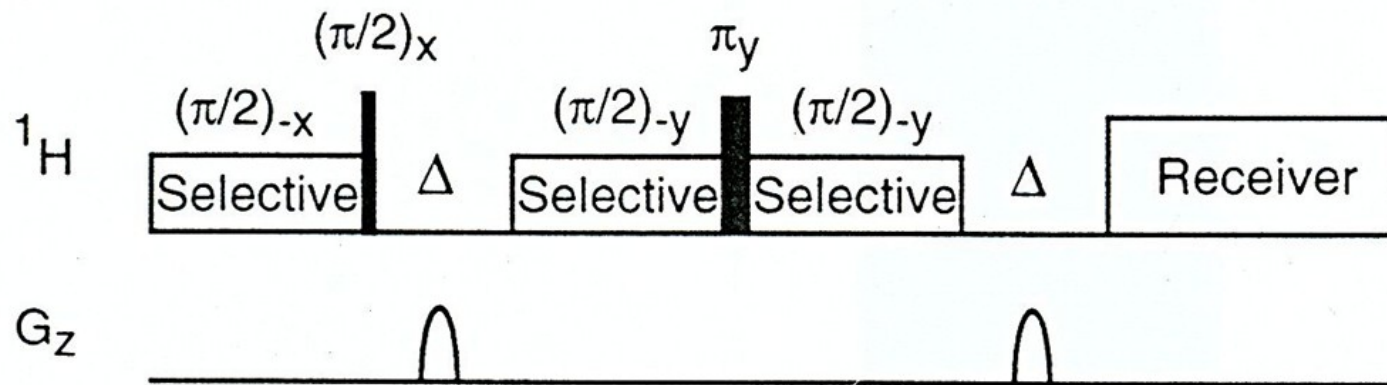
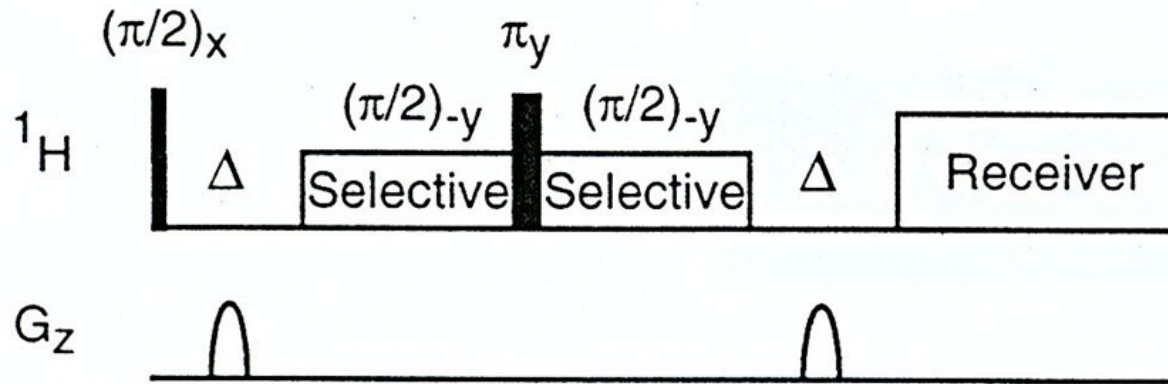
Pros: Easy to set up
Excellent water suppression
(with proper setup as good as presat)
Good for broad signals!

Cons: Non uniform excitation
Baseline not flat

Other sequences: 1331 etc



WATERGATE



Structure Determination Procedure

Structure Determination:

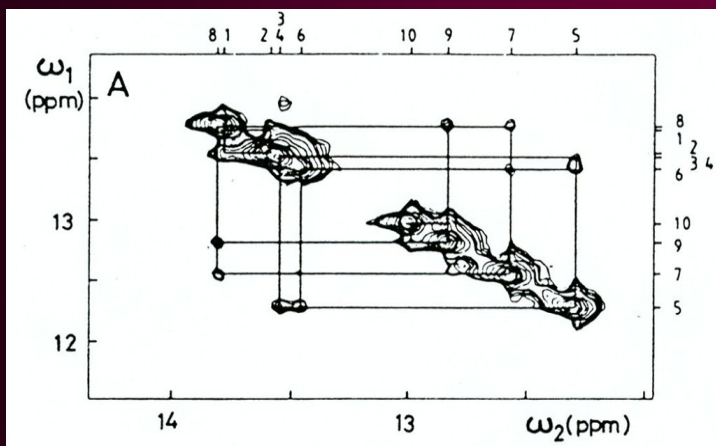
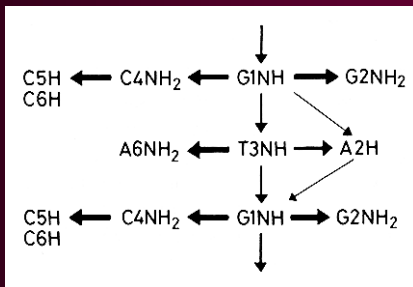
- | | | |
|------|----------------------------|---------------------------------|
| I) | Assignment | NOESY, COSY, HSQC
TOCSY..... |
| II) | Local Analysis | |
| | •glycosidic torsion angle | (NOE, <u>COSY</u>) |
| | •sugar puckering | (COSY, NOE) |
| | •backbone conformation | (COSY) |
| | •base pairing | (NOE, <u>COSY</u>) |
| III) | Global Analysis | |
| | •sequential | (NOE, COSY) |
| | •inter strand/cross strand | (NOE, <u>COSY</u>) |
| | •dipolar coupling | (HSQC, <u>HSQC</u>) |

Resonance Assignment

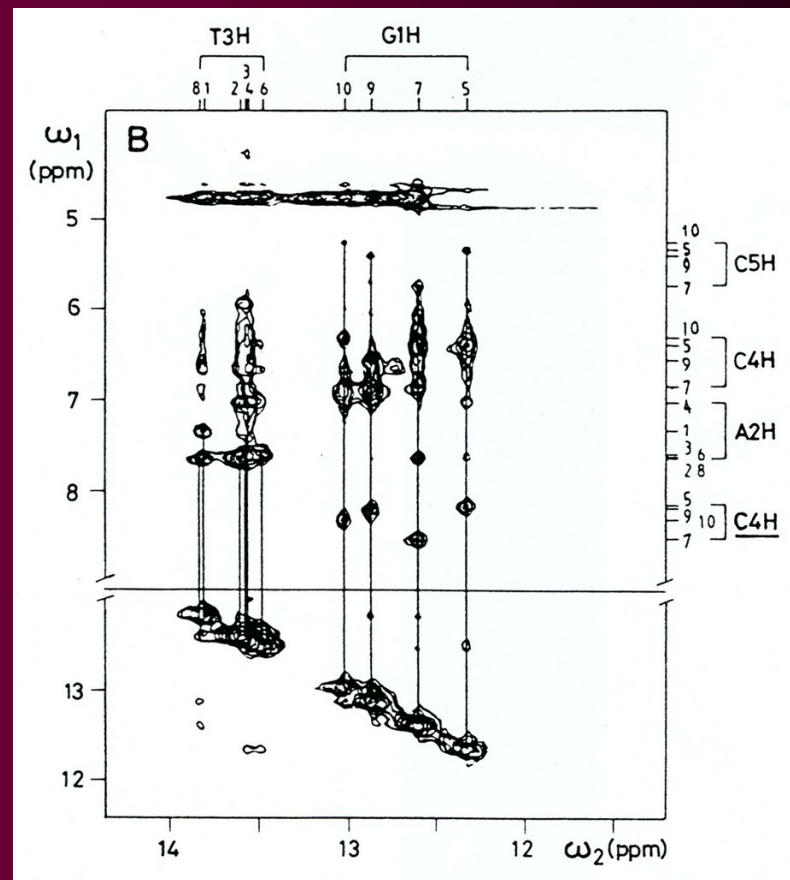
A) Exchangeable protons:	1D ^1H , 2D NOESY
B) Non-exchangeable protons	
• Aromatic Spin Systems:	2D DQF-COSY (H5-H6), 2D NOESY
• Sugar Spin Systems:	2D DQF-COSY 2D TOCSY
• Sequential Assignment:	2D NOESY 2D (^{31}P , ^1H) HETCOR
C) Correlation of exchangeable and non-exchangeable protons:	2D NOESY

Sequential connectivities with exchangeable protons

d(GGAATTGTGAGCGG)
d(CCTTAACACTCGCC)

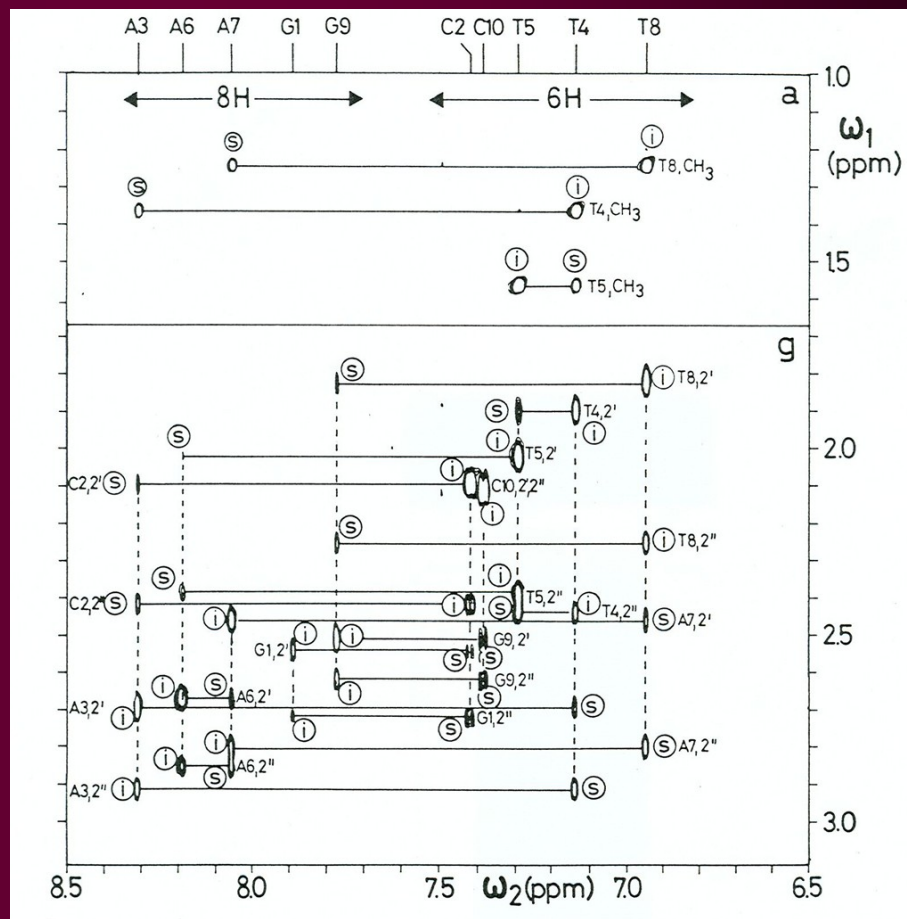
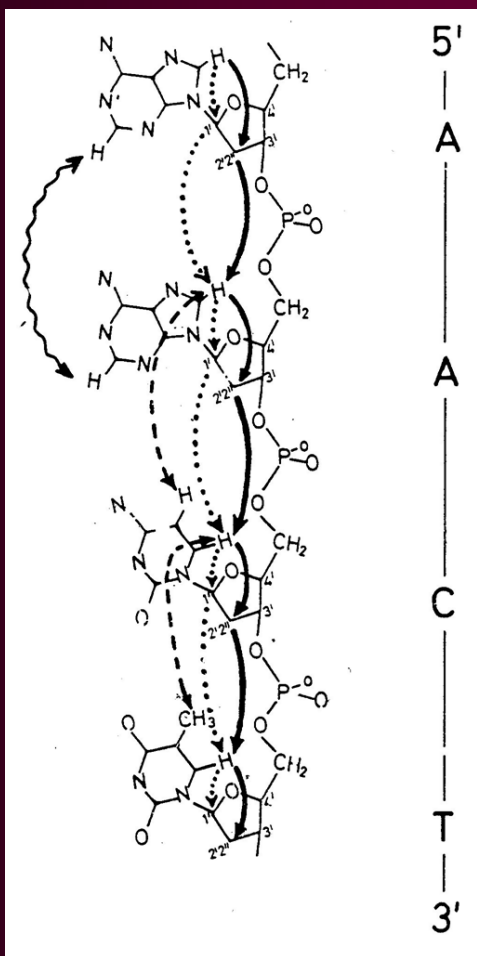


imino-imino



imino-amino

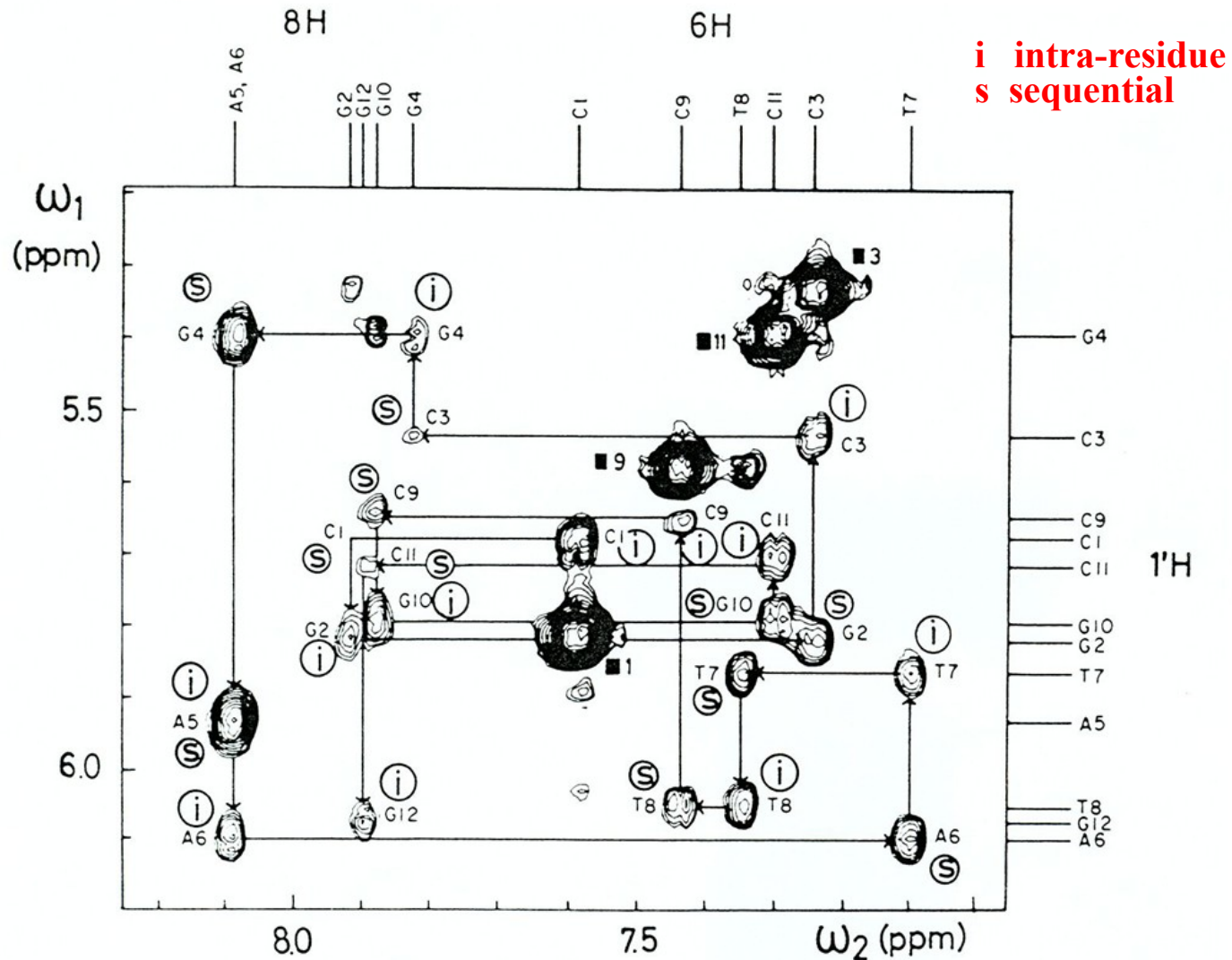
Sequential resonance assignments



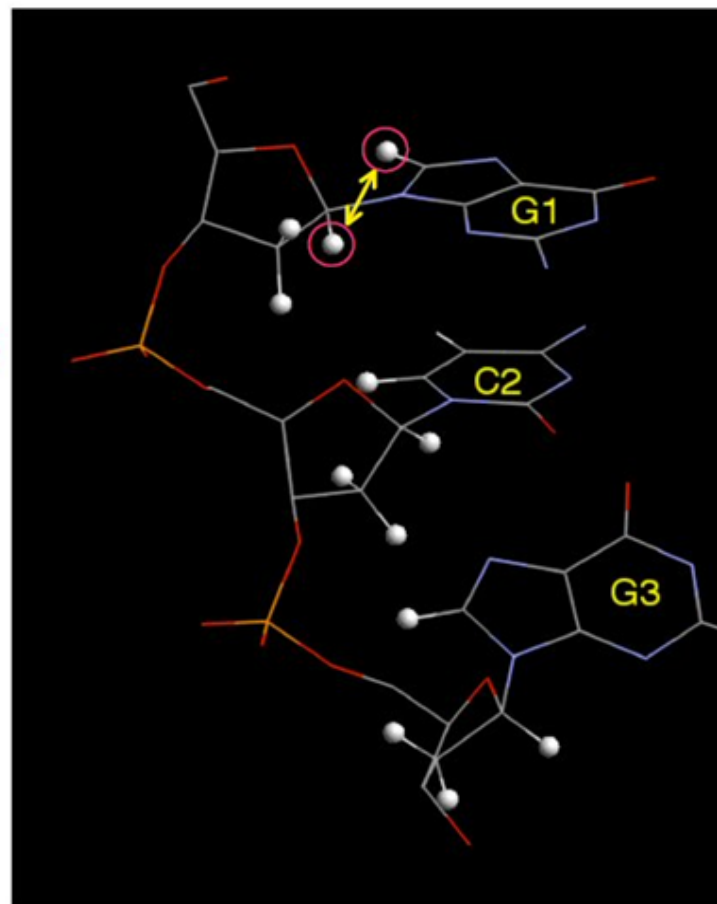
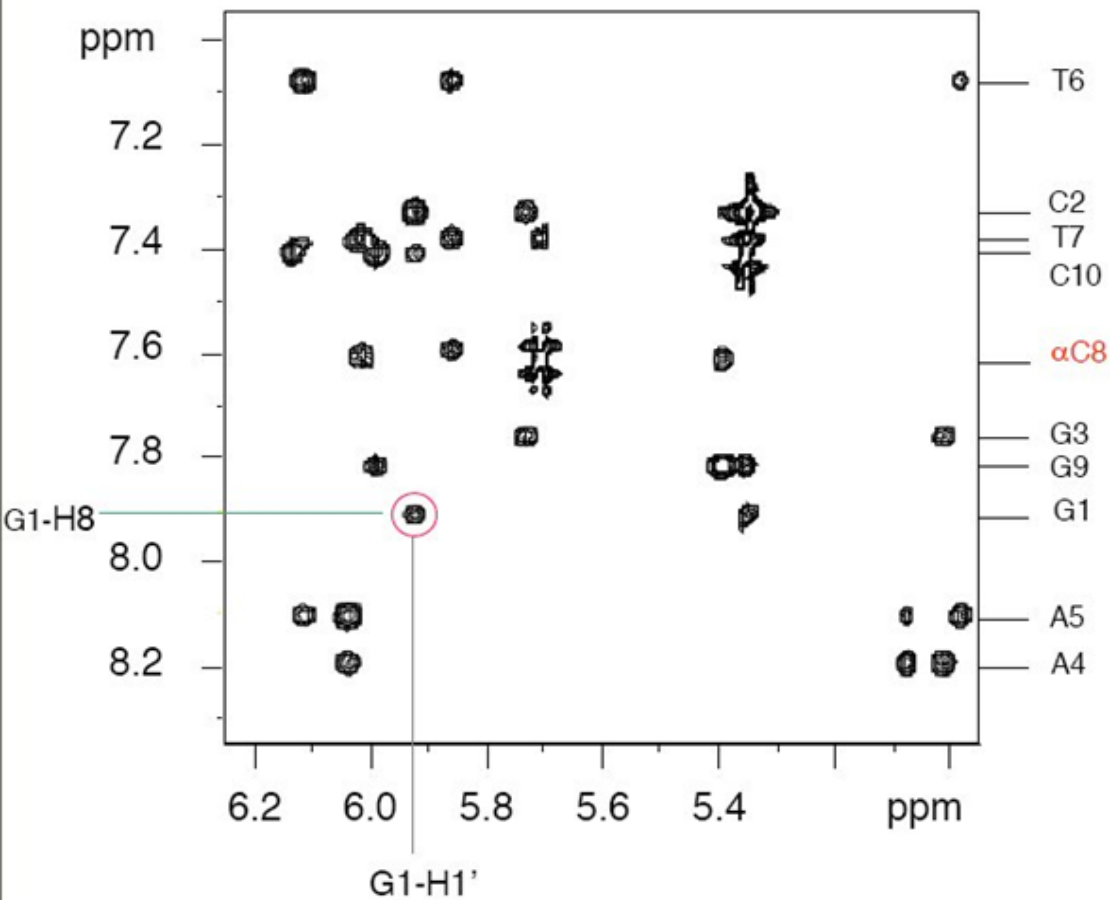
d(GCATTAAATGC)₂

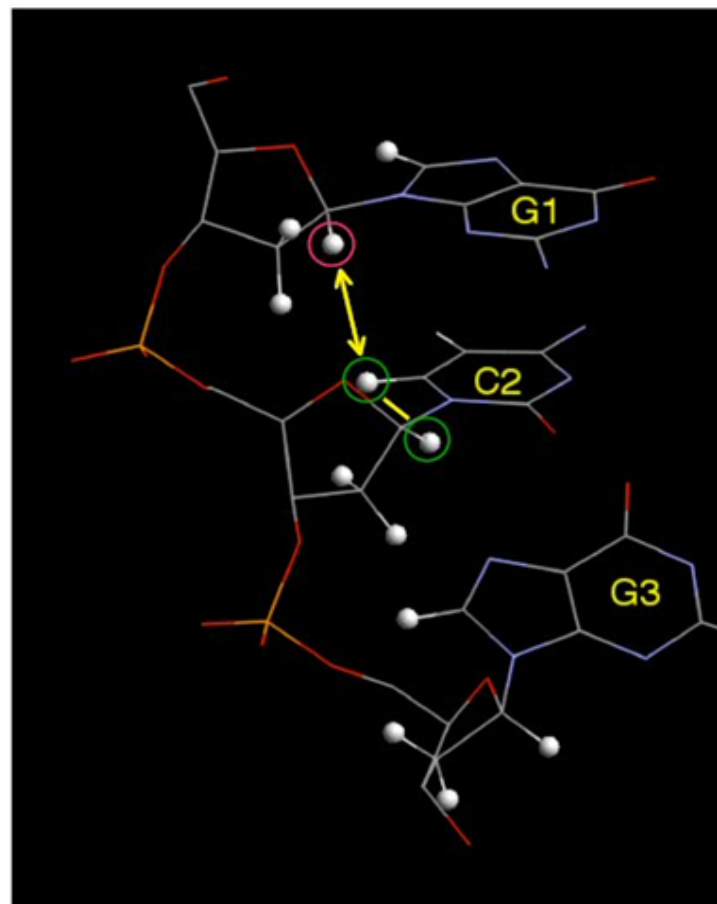
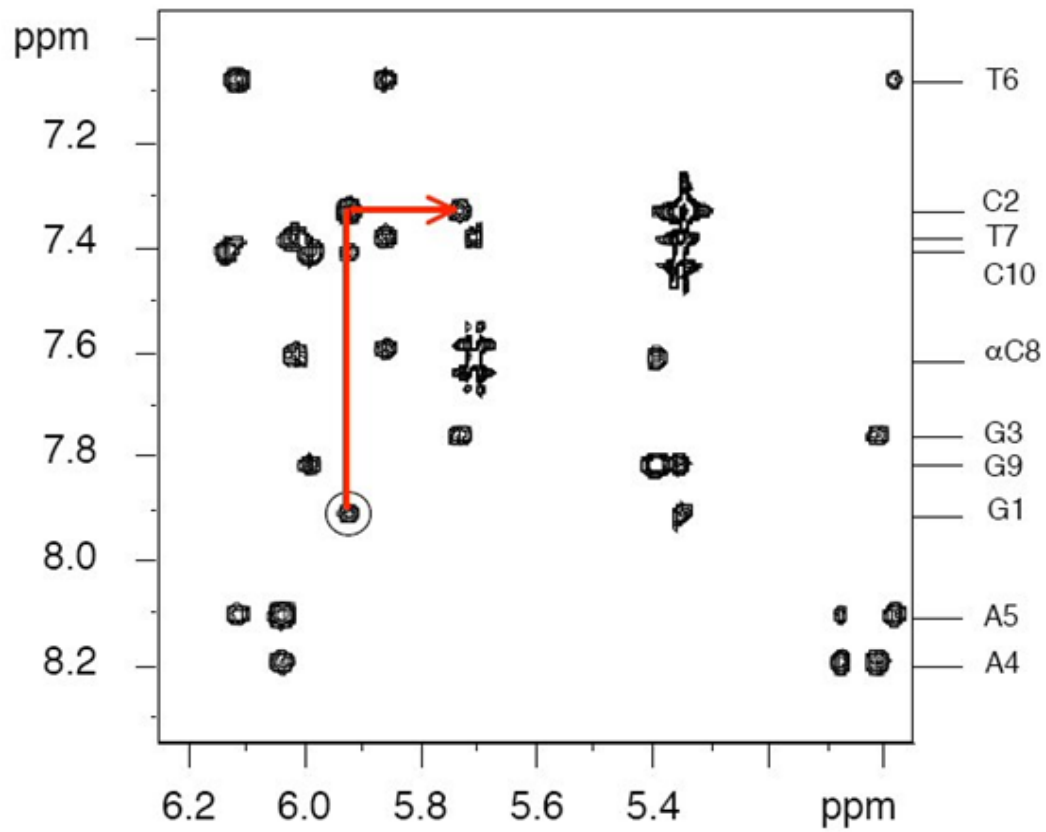
Connectivities between H1' and H6/8

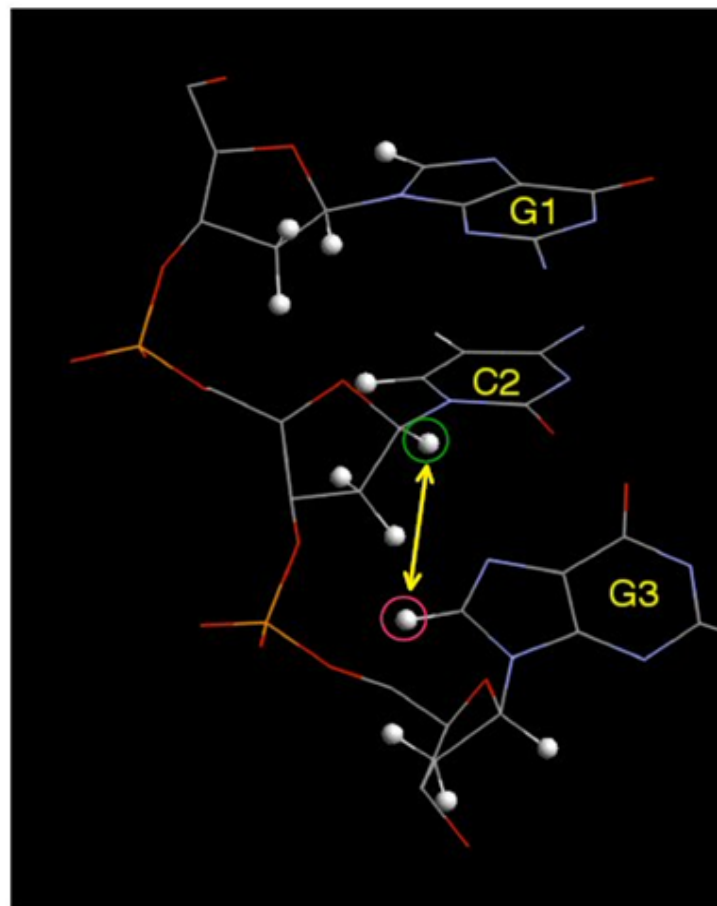
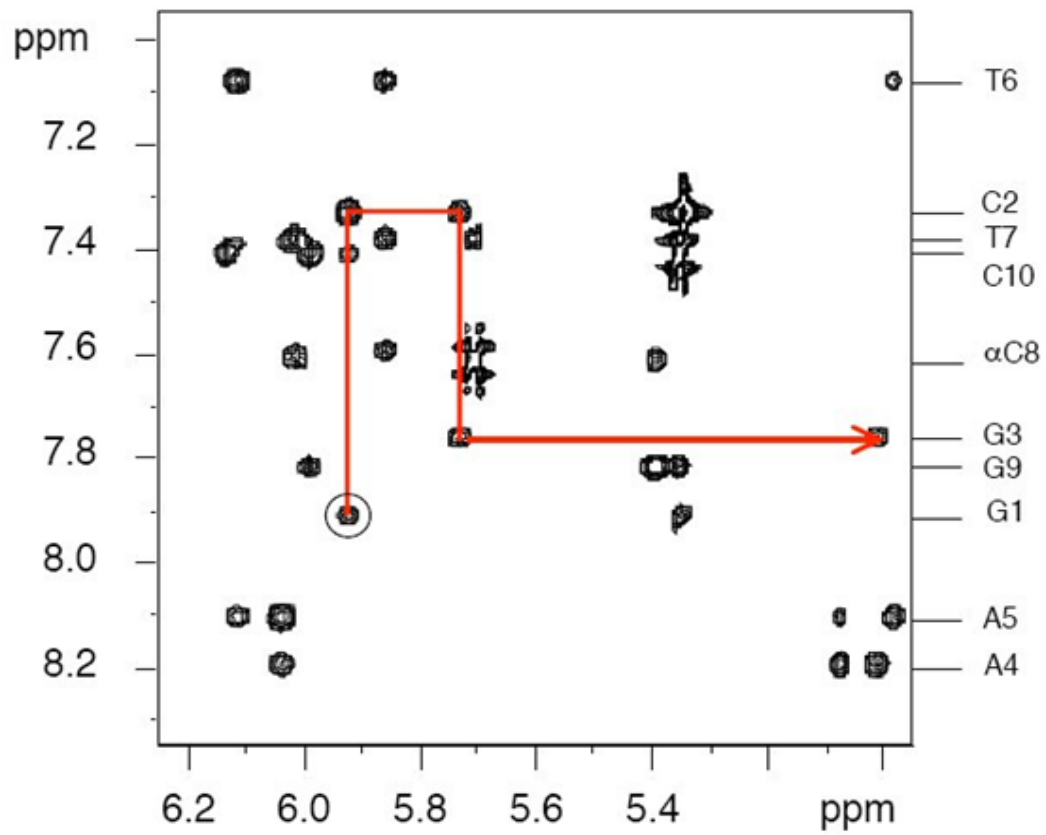
$d(\text{CGCGAATTCGCG})_2$

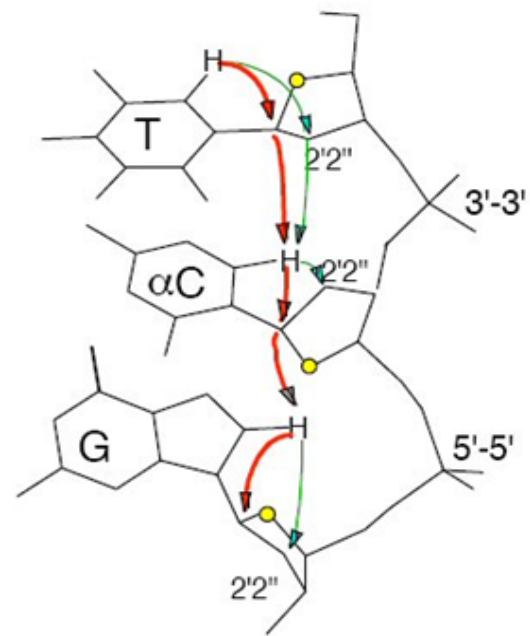
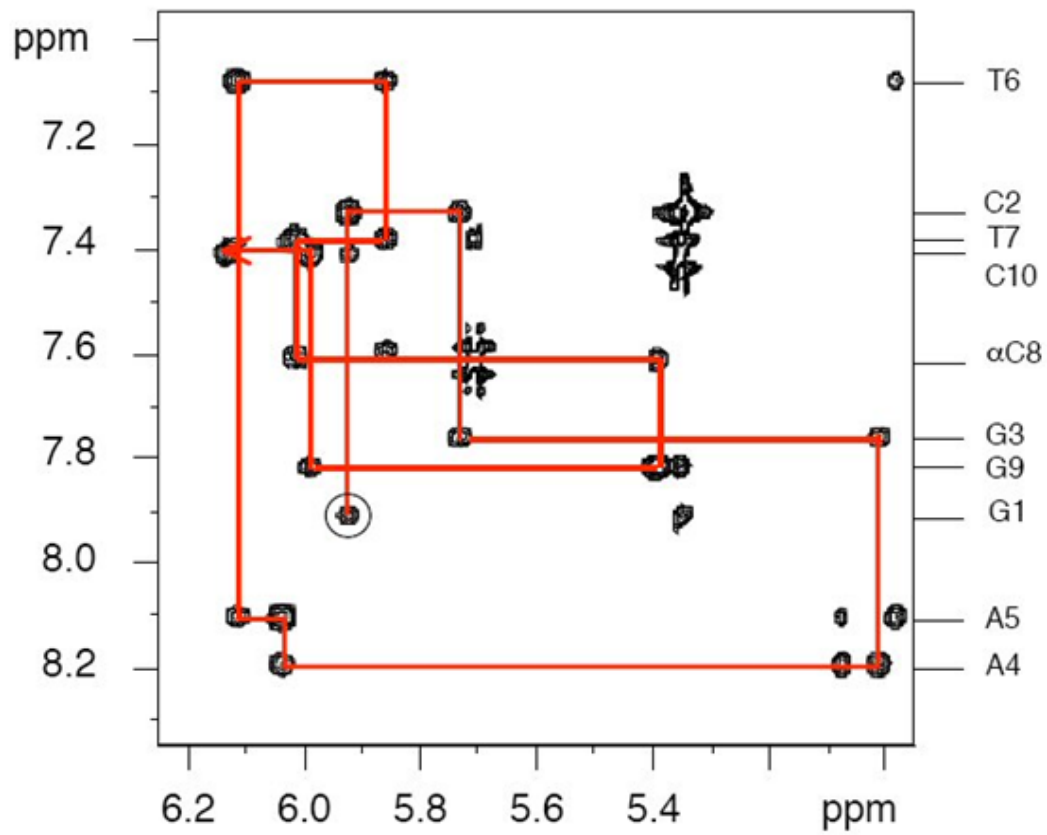


NOESY Connectivity (e.g. α C Decamer)



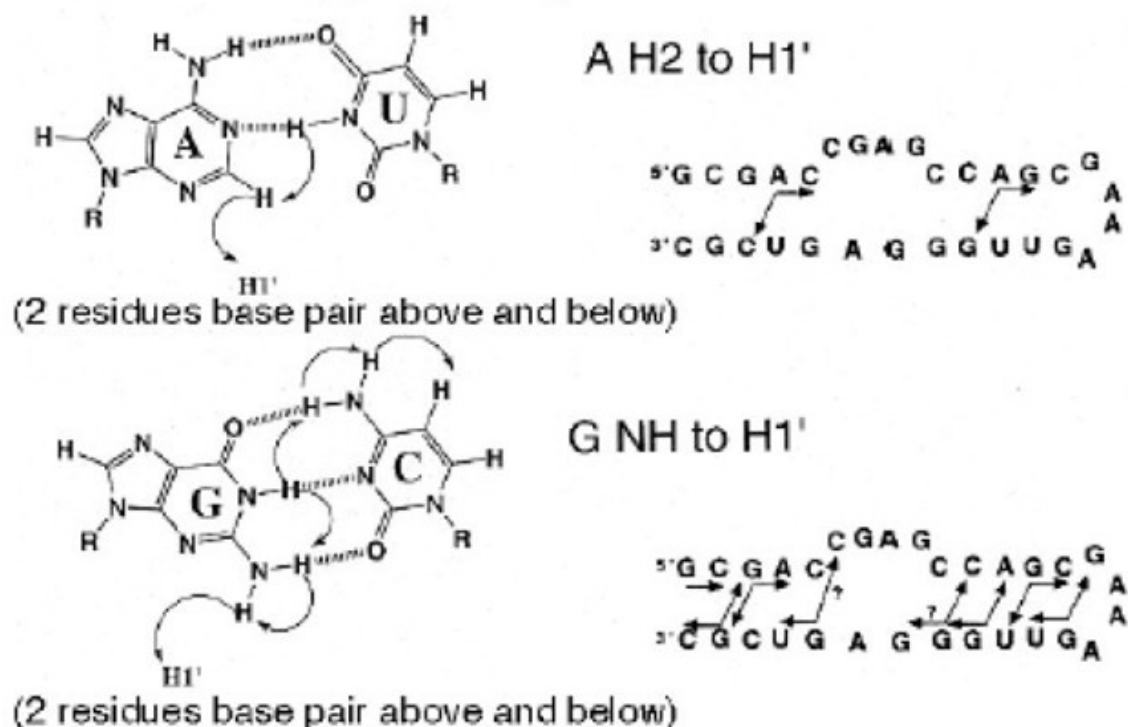






3) Resonance Assignment of RNA by Homonuclear NMR (cont'd)

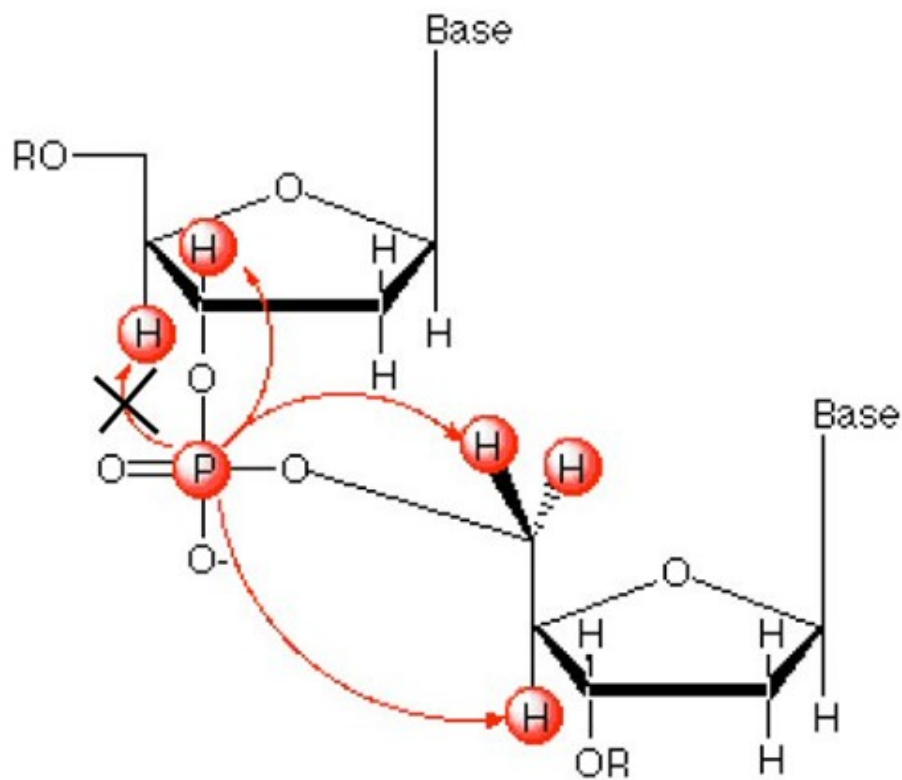
E. Correlations between exchangeable and non-exchangeable protons



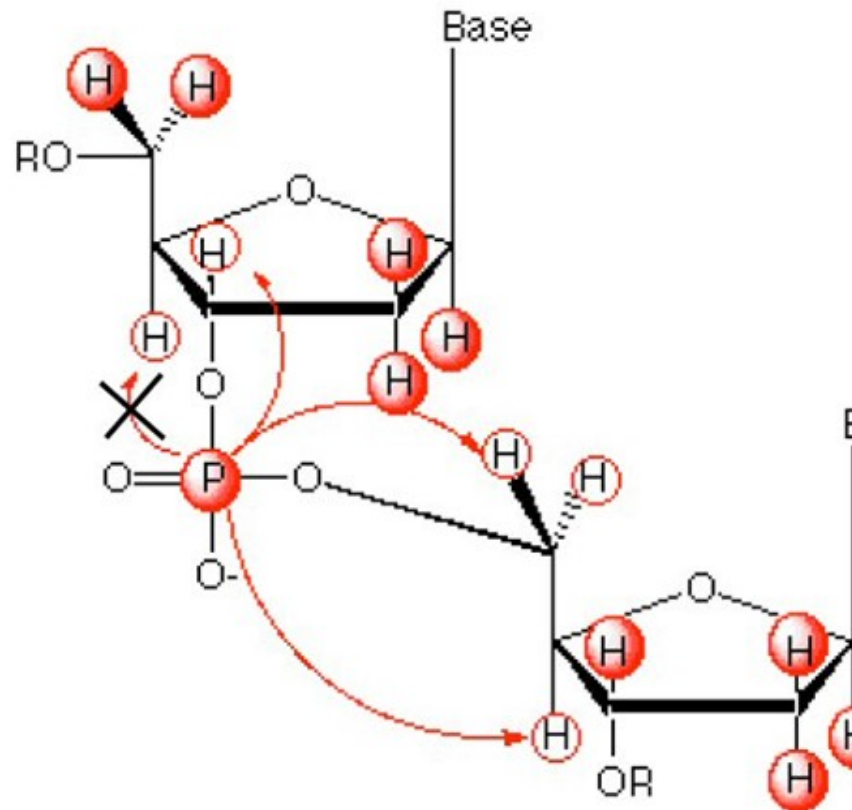
Assignment of Sugar-Phosphate Backbone

^{31}P NMR

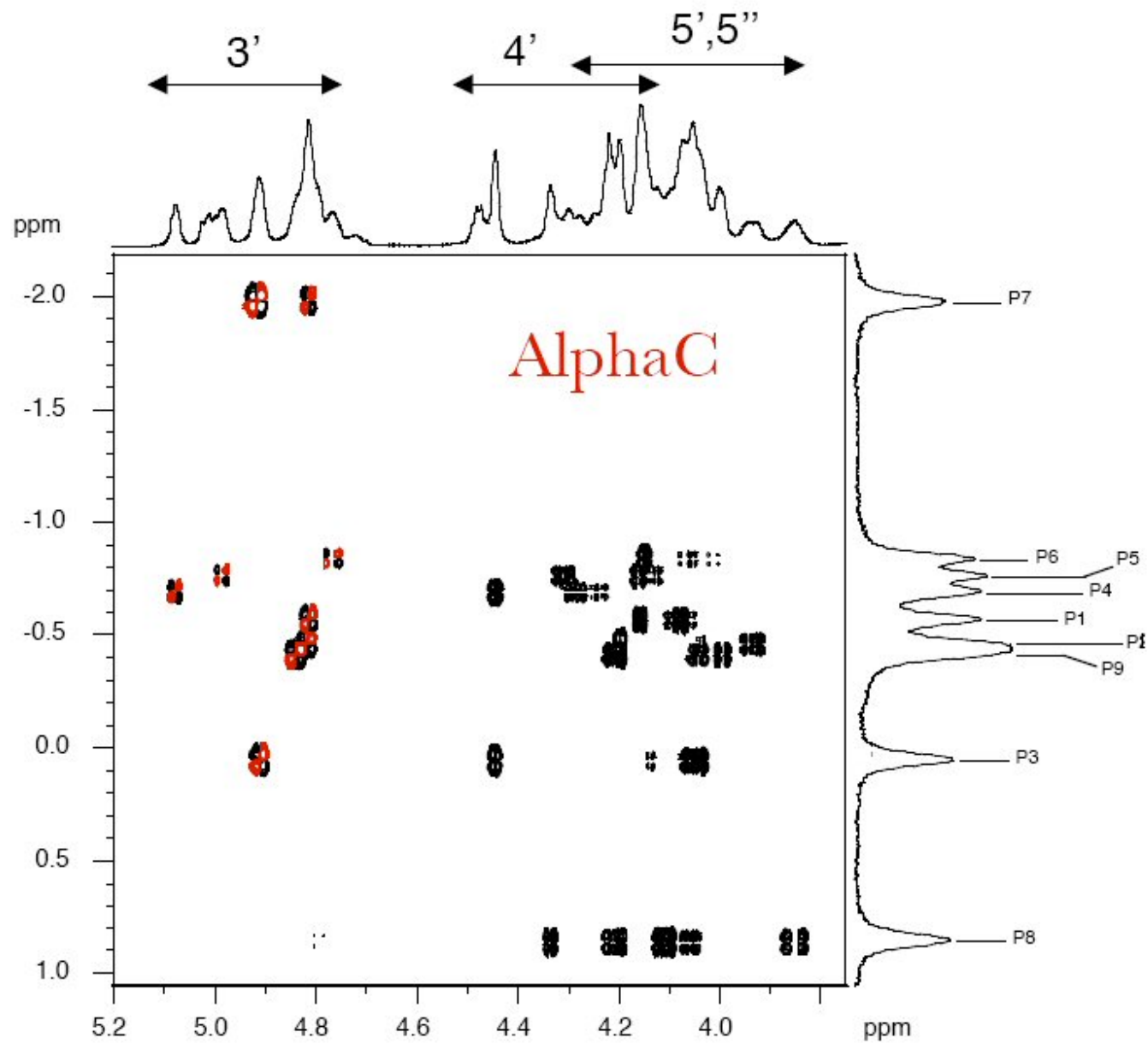
HP-COSY



HP-TOCSY



^{31}P NMR



Sugar pucker

The five membered furanose ring is not planar. It can be puckered in an envelope form (E) with 4 atoms in a plane or it can be in a twist form. The geometry is defined by two parameters: **the pseudorotation phase angle (P)** and the **pucker amplitude (Φ)**.

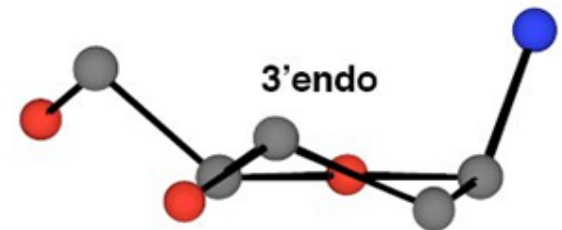
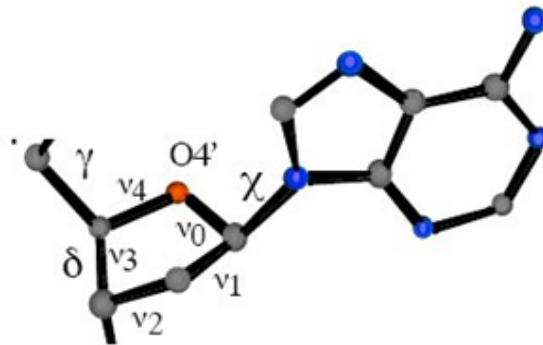
In general:

RNA (A type double helix) C3' endo.

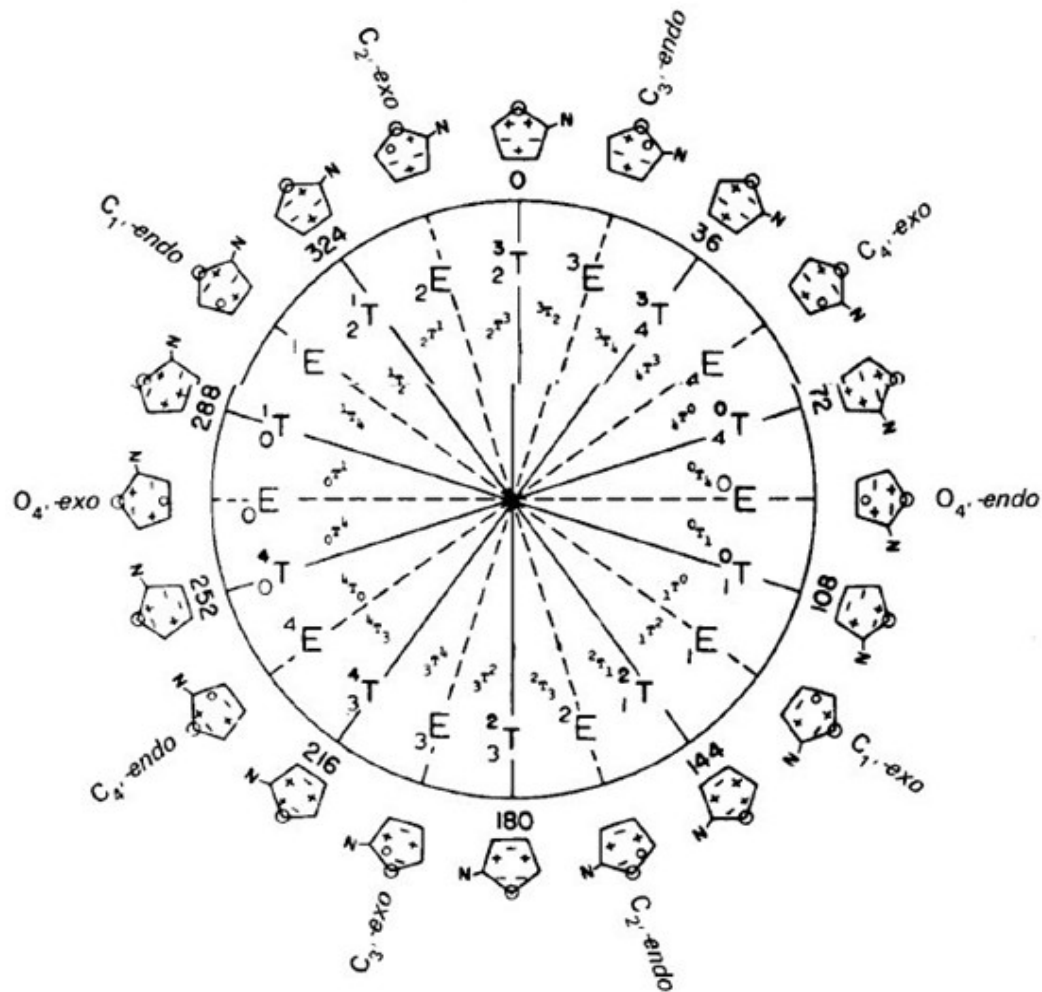
DNA (B type double helix) C2' endo.

$$\nu_i = \Phi_m \cos (P + 144 (j-2))$$

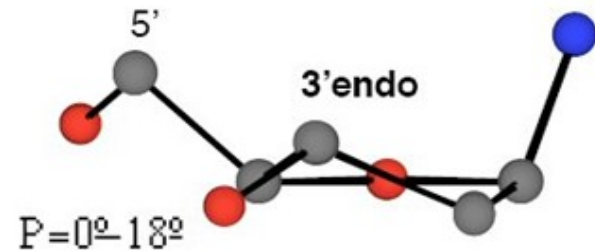
$$\delta = \nu_3 + 125^\circ$$



N (Northern)

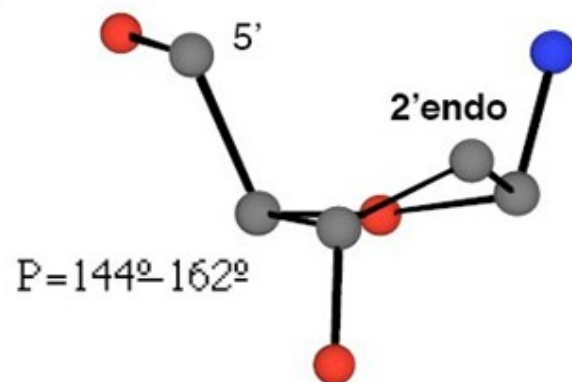


(Southern)



Ribose: $^3J_{H1'-H2'} \approx 1 \text{ Hz}$

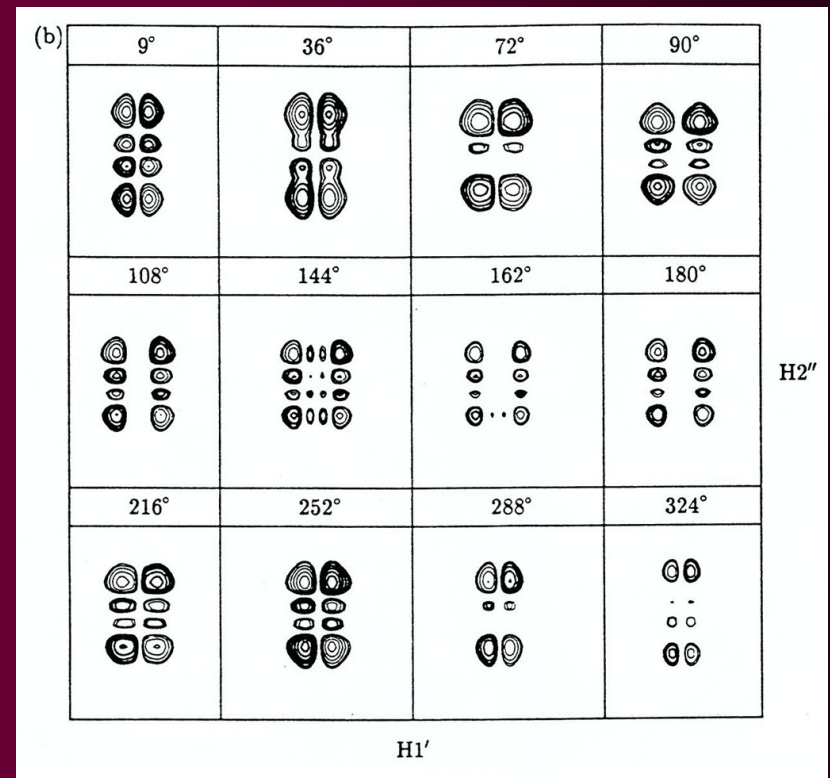
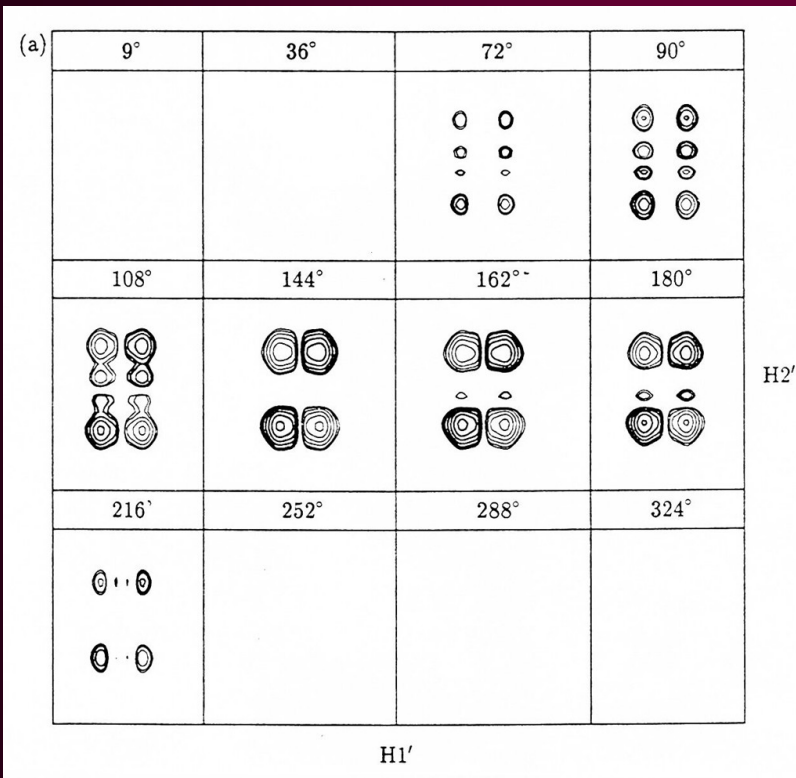
Deoxyribose: $^3J_{H1'-H2'} \approx 1.8 \text{ Hz}$



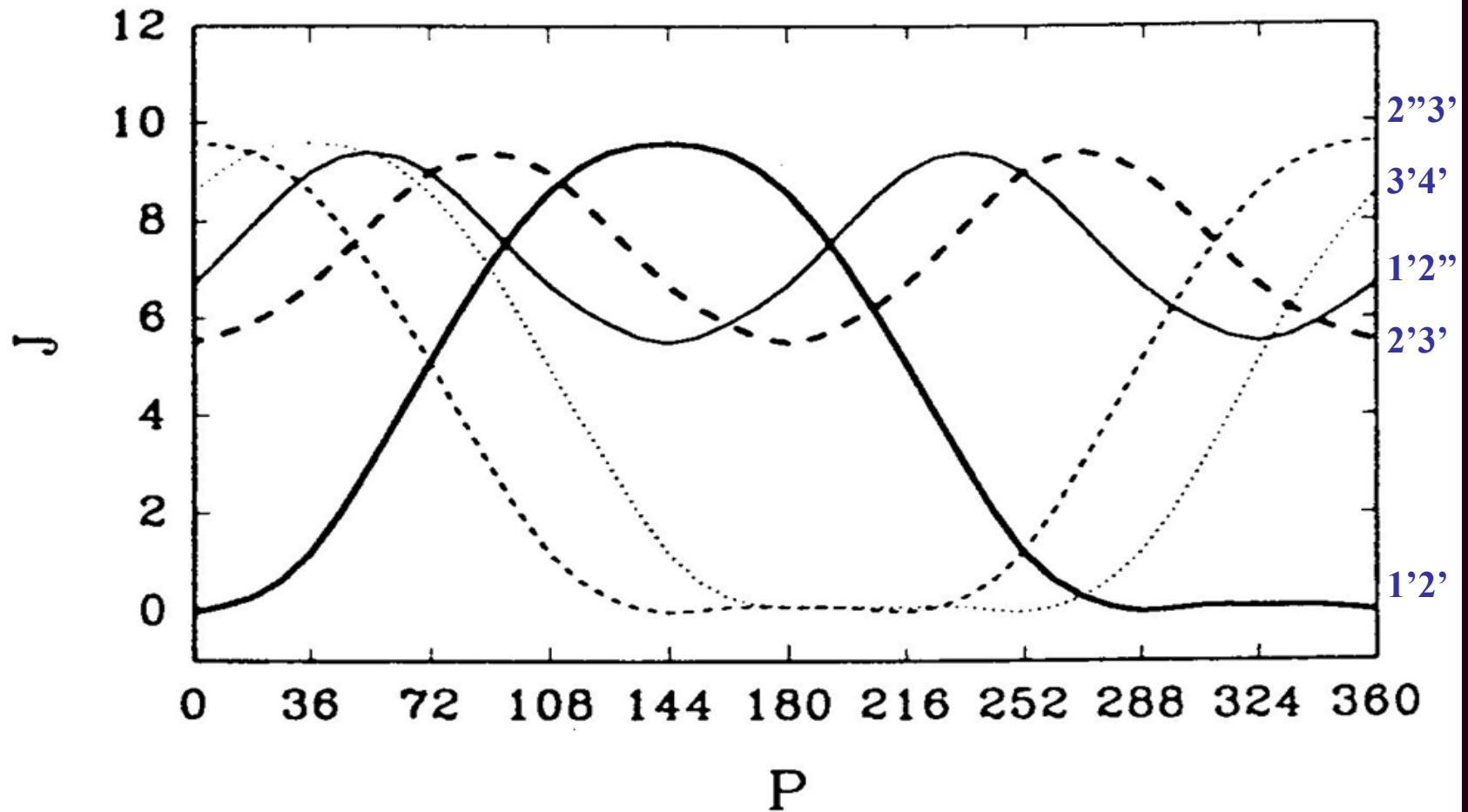
Ribose: $^3J_{H1'-H2'} \approx 7.9 \text{ Hz}$

Deoxyribose: $^3J_{H1'-H2'} \approx 10 \text{ Hz}$

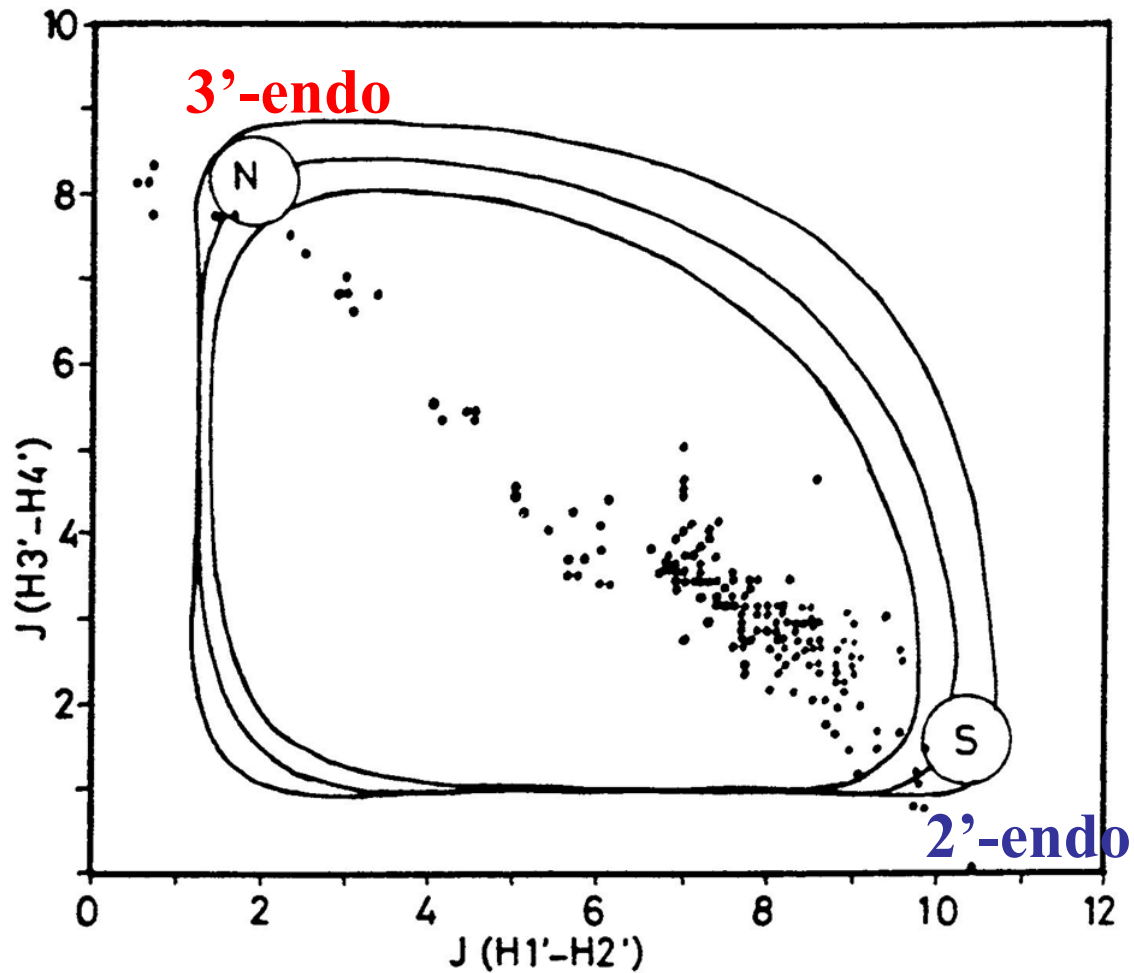
J-couplings from COSY spectra



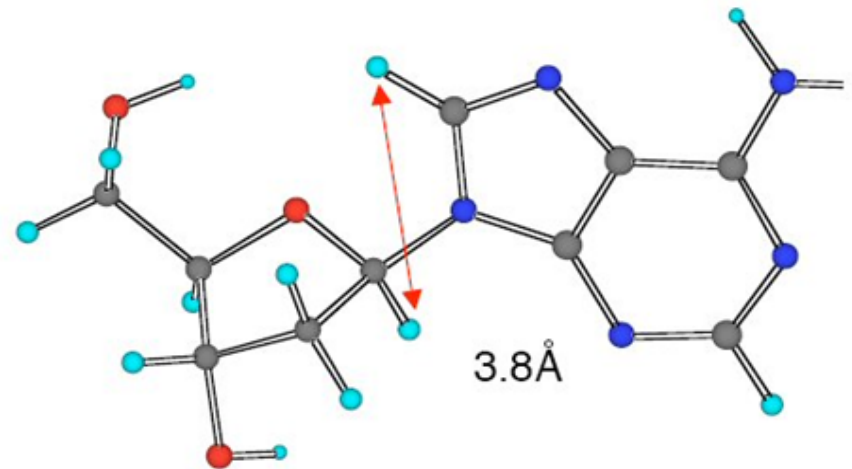
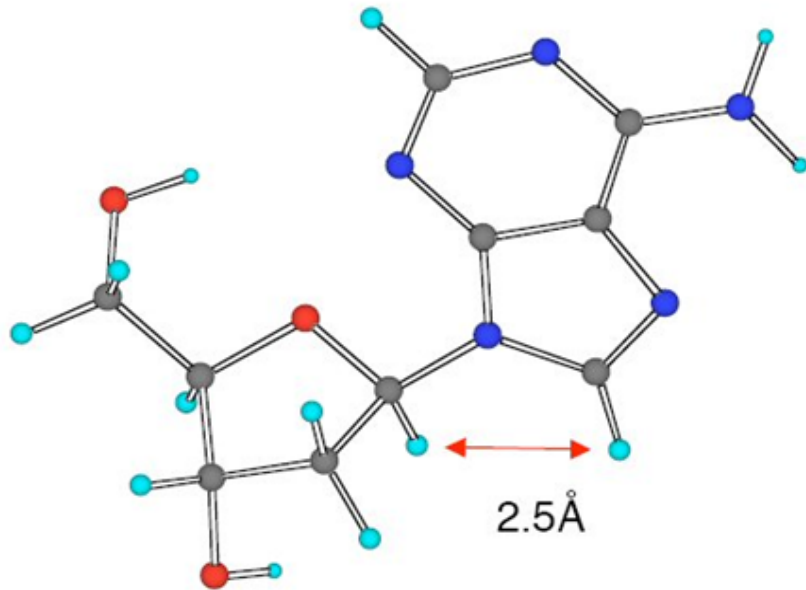
P determination from J-couplings



Equilibrium of N and S conformations

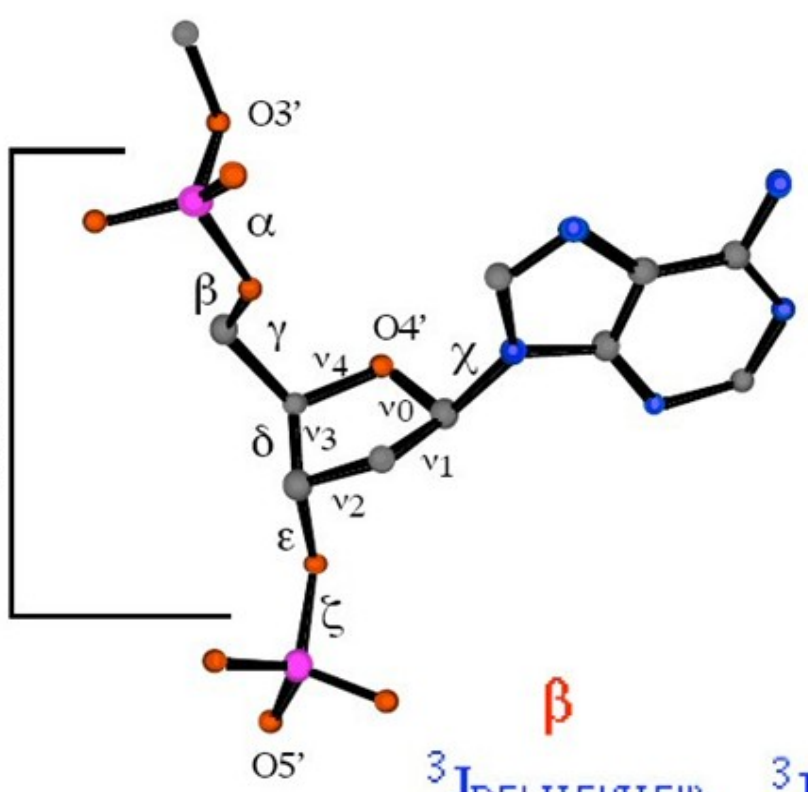


Distance information determines the glycosidic torsion angle



- How do we get distance information?
 - Nuclear Overhauser effect ($< 6\text{Å}$)

nucleotide unit



α and ζ pose problems
 Determinants of ^{31}P chem shift.
 ϵ and ζ correlate. $\zeta = -317 - 1.23 \epsilon$

β	γ	ϵ	χ
$^3\text{J}_{\text{P}5'-\text{H}5'(\text{H}5'')}$	$^3\text{J}_{\text{H}4'-\text{H}5'(\text{H}5'')}$	$^3\text{J}_{\text{P}3'-\text{H}3'}$	$^3\text{J}_{\text{H}1'-\text{C}6}$ (U,C,T)
$^3\text{J}_{\text{P}5'-\text{C}4'}$	$^3\text{J}_{\text{C}3'-\text{H}5'(\text{H}5'')}$	$^3\text{J}_{\text{P}3'-\text{C}2'}$	$^3\text{J}_{\text{H}1'-\text{C}2}$ (U,C,T)
		$^3\text{J}_{\text{P}3'-\text{C}4'}$	$^3\text{J}_{\text{H}1'-\text{C}8}$ (A,G)
			$^3\text{J}_{\text{H}1'-\text{C}4}$ (A,G)

Structure Determination:

- I) Assignment
- II) Local Analysis
 - glycosidic torsion angle, sugar puckering, backbone conformation
 - base pairing
- III) Global Analysis
 - sequential, inter strand/cross strand, dipolar coupling

Nucleic Acids have few protons.....

- NOE accuracy
 - > account for spin diffusion
- Backbone may be difficult to fully characterize
 - > especially α and ζ .
- Dipolar couplings

What do we know?

- Distance, Torsion, H-Bond constraints

What do we want?

- Low energy structures

Methods

- Distance Geometry
- Simulated annealing, rMD
- Torsion angle dynamics (DYANA)
- Mardigras/IRMA/Morass

