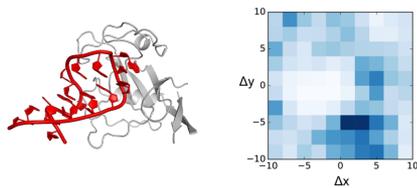# Welcome to CHEM 153B: DNA, RNA, and Protein Synthesis

- This quarter we are going to cover:
  - Nucleic acid structure
  - DNA & RNA polymerases
  - DNA replication
  - DNA repair
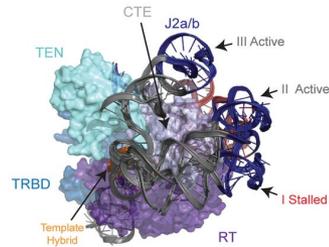 - Transcription
  - RNA processing
  - Translation

My background: modeling the structures and energetics of RNAs and RNA-protein complexes
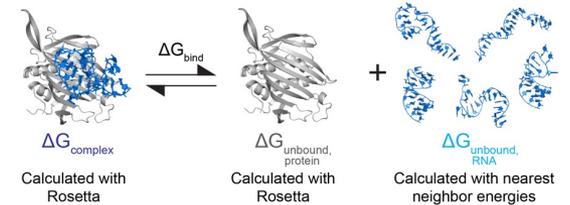
RNP-denovo

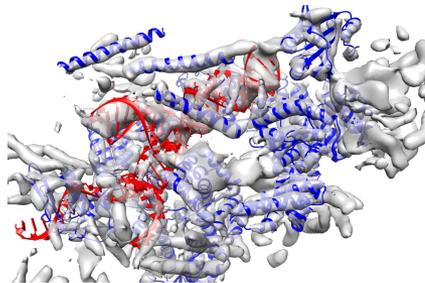K. Kappel, R. Das. *Structure*, 2019.

smFRET-Rosetta

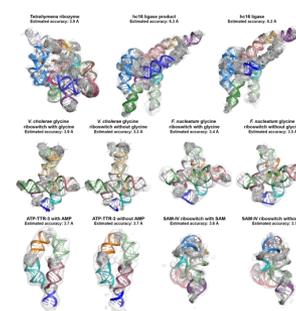J. Parks*, K. Kappel*, et al. *RNA*, 2017.

RNP-ddG

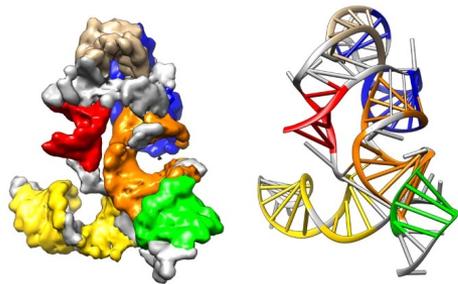K. Kappel, et al. *PNAS*, 2019.

DRRAFTER

K. Kappel, et al. Nature Methods, 2018.
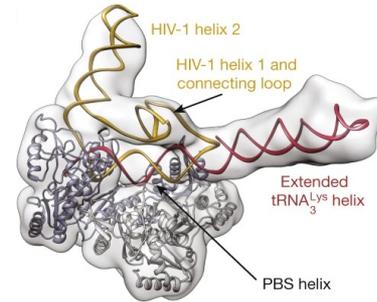
auto-DRRAFTER

K. Kappel*, K. Zhang*, Z. Su*, et al. Nature Methods, 2020.
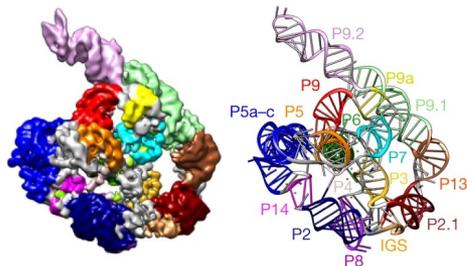
Using cryo-EM and computational modeling to resolve several structures of RNAs and RNA-protein complexes
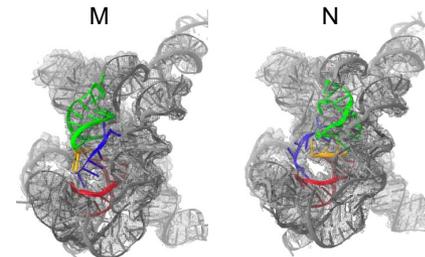

K. Zhang*, S. Li*, K. Kappel*, et al. Nat. Commun. 2018.

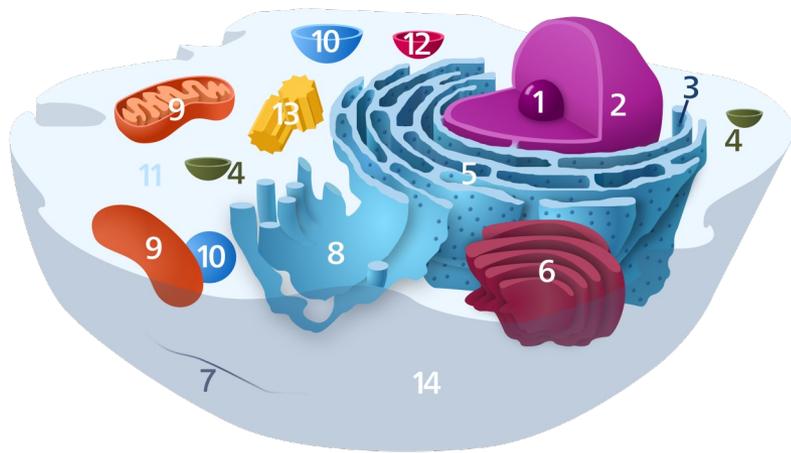K.P. Larsen, Y. Mathiharan, K. Kappel, et al. Nature, 2018.

Z. Su*, K. Zhang*, K. Kappel* et al. Nature, 2021.
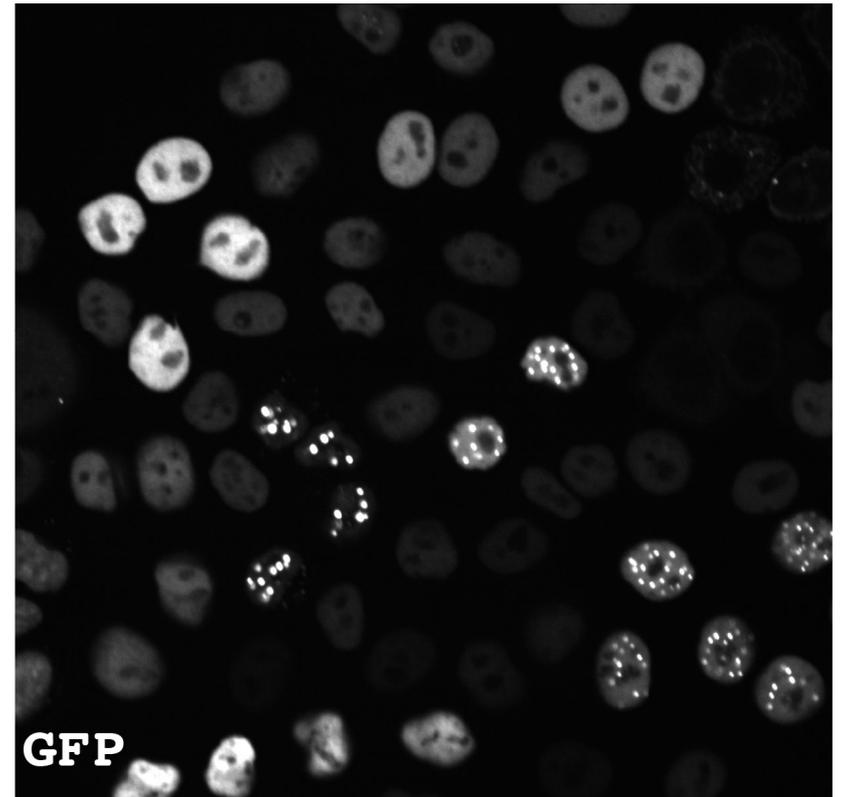

S. Li, et al. PNAS, 2022.

How is the inside of a cell organized?

A human cell contains 5-10 billion protein molecules
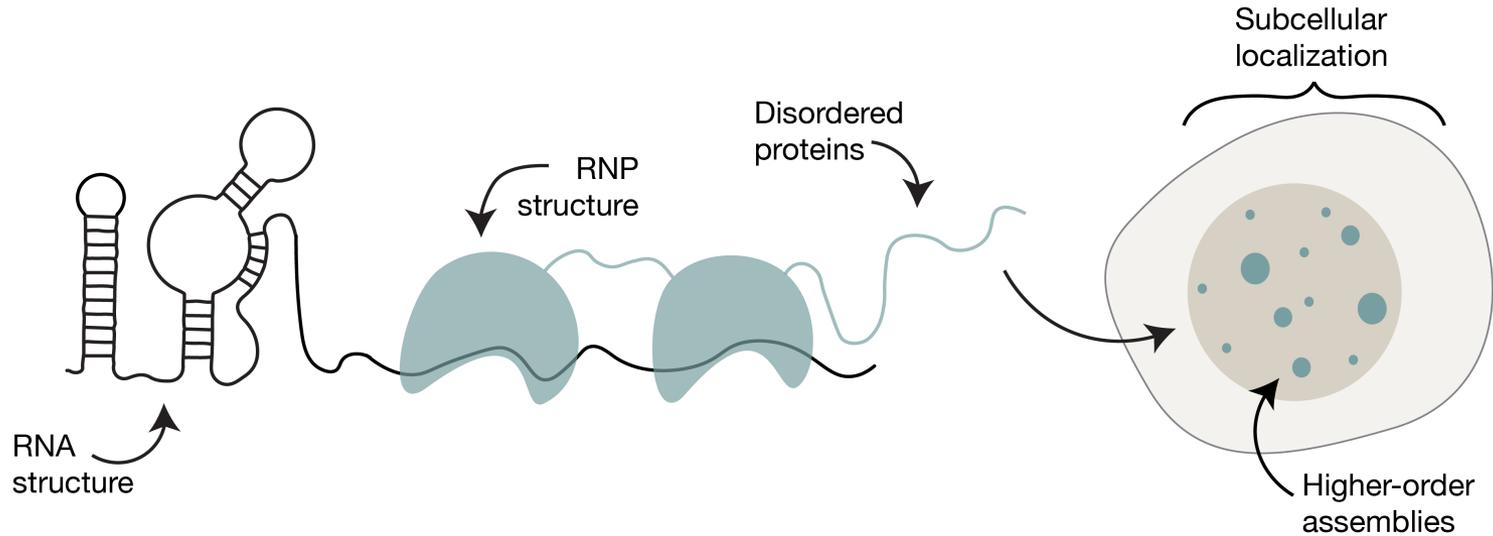

https://en.wikipedia.org/wiki/Organelle

GFP

K. Kappel, et al. Nature Methods, 2025.

The Kappel Lab @ UCLA:
Combining high-throughput experiments and computation to decipher sequence-organization-function relationships of proteins, RNA, and their interactions

Chem 153B Syllabus

- ## The full syllabus is posted on BruinLearn

Course Syllabus
Chem 153B – Winter 2026
Professor Kalli Kappel

Instructor: Professor Kalli Kappel (she/her)

Email: kallikappel@ucla.edu; please use chem153bkappel@gmail.com for course inquiries (see below).

Instructor office hours: 2051 Young Hall, Mondays 2-3 PM and Thursdays 2-3 PM

TAs:
Izaiah Cole (icole@g.ucla.edu)
Jake Cozy (jcozy@g.ucla.edu)
Marcella Mirabelli (mmirabelli@g.ucla.edu)
Vanessa Wiratmo (vanessawiratmo@g.ucla.edu)

TA office hours: Tuesdays 4-5 PM, 4222A Young Hall (except for January 20th, when office hours will be held in 3064 Young Hall)

Course Meetings: M,W,R,F: 1-1:50 PM in Young Hall CS24

Thursday tutorials are optional, except for the days when we have exams.
We will use the tutorials to review course material and go over old exam questions.

Lectures: MWF - Young Hall CS24
Lectures will be in person and will introduce the materials and concepts covered in this course. All materials will be posted on the course BruinLearn page: https://bruinlearn.ucla.edu/courses/220811

Calendar of assignments and exams:
You should submit a homework by 11:59 PM on Monday on all weeks highlighted in Blue.

	Lectures YH-CS24		Tutorials YH-CS24
	Midterms YH-CS24		No Class

Week	Monday	Tuesday	Wednesday	Thursday	Friday
1	1/5		1/7	1/8	1/9
2	1/12		1/14	1/15	1/16
3	1/19		1/21	1/22	1/23
4	1/26		1/28	1/29	1/30
5	2/2		2/4	2/5	2/6
6	2/9		2/11	2/12	2/13
7	2/16		2/18	2/19	2/20
8	2/23		2/25	2/26	2/27
9	3/2		3/4	3/5	3/6
10	3/9		3/11	3/12	3/13

iClickers:
To facilitate active learning and promote engagement with course material, students will answer quizzes shown on some of the lecture slides using iClickers. The iClicker app is now integrated into BruinLearn and you can access it for no additional cost:
https://mhe.my.site.com/iclicker/s/article/Student-Guide-iClicker-Roster-Grade-Sync-Integration
Course Name: Chem 153B
Use this link to register your iClicker for this course (or follow the QR code to the right):
https://join.iclicker.com/WCUO

Please sign up for iClicker!
iClicker participation points will be given for completion regardless of whether or not your answers are correct.

Tutorials (Prof. Kappel):

<mark>• **Tutorials on January 29th and February 19th will be used for midterm exams in Young Hall CS24.**</mark>

• The other tutorials will be in person in Young Hall CS24. These sessions typically consist of a brief review of recent course material followed by working through old exam questions in small groups. These sessions are not recorded and attendance is highly recommended since they are mostly used for problem solving in small groups. <mark>**I reserve the right to cover exam-relevant material during tutorials if I deem it necessary to achieve the course learning objectives by end of quarter.**</mark>

Discussions (TAs):

1A:	T	9:00-9:50 AM	Boelter Hall 4283	Izaiah Cole
1B:	T	1:00-1:50 AM	Young Hall 3069	Izaiah Cole
1C:	W	11:00-11:50 AM	Young Hall 1044	Jake Cozy
1D:	W	3:00-3:50 PM	Young Hall 2200	Jake Cozy
1E:	R	9:00-9:50 AM	Boelter Hall 5252	Marcella Mirabelli
1F:	R	4:00-4:50 PM	Boelter Hall 2444	Marcella Mirabelli

The first discussion section will focus on an introduction to PyMol, and the following discussions will focus on problem sets in preparation of midterms/final exam. Attendance will be taken, and please attend your assigned discussion section.

Grading: Exams, Homework and Participation

Midterms: 30%
Your higher midterm score will count for 20% of your grade, and the lower score as 10%.

Final Exam: 40% <mark>Wednesday, March 18th, 8:00 AM-11:00 AM</mark>
<mark>If you have scheduling conflicts regarding the midterms or final exam, please reconsider your enrollment in the course.</mark>

Homework: 10%
To encourage you to explore the problems presented in the weekly discussions ahead of time, you will be required to complete and submit one (out of the four possible) discussion problems on <u>Monday prior to your discussion section</u> (11:59PM deadline) for completion credit. The question you will answer will be the same as your group number, which will be assigned to you during the first week. Your submission should be a screenshot of a digital version or a photo of a paper version of your single question (<u>not</u> a PDF or word document) to help reduce formatting errors that could arise from various document types. You are welcome to work in groups to answer these questions, but every student should write their own answers. If you need clarification, please ask your TAs to help. You will get 1.25% credit for each completed homework, with a maximum possible percentage of 10%. *This equates to being able to miss one problem set during the quarter without penalty.*

Attendance and Participation Points: 10%
• iClicker lecture questions = 10%. While I expect that you will attend every class, you will receive full credit if you submit all iClicker responses in at least 20 lectures throughout the quarter (≥ 20 would be 10%, 19 would be 9.5%, etc). *In short, this equates to 8 skips that you can use throughout the quarter for mandatory classes without penalty.* iClicker participation points are given for completion regardless of whether or not the answer is correct. Since the class is not on Zoom and we will be using active learning techniques, **you must be physically in the classroom to be eligible for iClicker points on any given day**. If you have trouble with iClicker during any given lecture, please speak with me after class.

Discussions Participations = 10%
Please go to your assigned discussion section. While we expect that you will attend all discussions, you must attend at least 8 for full credit (less than 8 is 1.25% per discussion attended). *This equates to 2 skips during the quarter for mandatory discussion sections without penalty.*

Extra Credit: 1% for submitting an online evaluation for the course by the campus deadline.

Grading Scale: Minimum scores for particular grades are listed below. Scores are based on exceeding a cutoff without rounding (for example, 97.9 is an A, not an A+).

A+ ≥ 98%	A ≥ 90%	A- ≥ 85%	B+ ≥ 80%	B ≥ 75%	B- ≥ 70%
C+ ≥ 66%	C ≥ 63%	C- ≥ 60%	D ≥ 55%	F <55%	

Course Policies:

- **Questions and Course Logistics:** If you need me to answer specific questions related to the materials, **please use the BruinLearn discussion forum rather than emailing me or a TA** as it is beneficial for all students to know the answer to this question. If you have a question about course logistics, **please send an email to chem153bkappel@gmail.com. Emails should only be sent directly to the instructor in rare cases that involve private or personal matters.**

- **Technology Policy**: Students are welcome to take class notes however they prefer: pen and paper, using an iPad, or on a laptop. However, students should remain engaged in the class and refrain from using social media, web browsing, online shopping, etc. during class.

- **Sick Policy**: _**In general, do not come to class sick**_. If you suspect you have COVID-19, please stay home until you have a negative rapid antigen test. If you feel well enough to attend class, please wear a mask.

- **Participation**: All students are expected to be active participants in the course. This includes attending every class meeting, asking questions, and being present and involved in discussions.

- **Regrade policy**: You are allowed two regrade challenges for the entire term-use them wisely. Like in tennis or football, your available challenges do not decrease if your challenge is found to be valid. However, once you use a challenge and it is not upheld, it counts toward your two. Before submitting a regrade request, carefully ask yourself:
  - Does my answer contain all the information requested by the question?
  - Does my answer contain any incorrect or unsupported claims?

  Historically, valid regrades fall into two categories:
  - Clerical errors: Examples include arithmetic mistakes in point totals, overlooked responses, or issues with automated grading (e.g., multiple-choice answers not captured by Gradescope).
  - Fundamental flaws in the question: If a flaw in the question affects multiple students, a broader correction will be made, and no challenge will be deducted from your total.

  Please note that we do not negotiate changes in partial credit. Regrade requests should be submitted on GradeScope, along with a clear explanation that directly addresses the questions above.

- **Course Evaluations**: All students are strongly encouraged to fill out course evaluations.

Textbook:
There is no required textbook for the course as none of the current textbooks cover all the information I give during lectures. You may find it useful to have one of the following texts as a reference to supplement the lecture:

Lehninger Principles of Biochemistry - David Nelson and Michael M. Cox

Many of you may already own the eBook or the hard copy book. For example, if you purchased Achieve/Lehninger for 153A, you already have 4-year access to the eBook. If you can still access your Chem 153A Achieve page, you should be able to read it through the Vital Source website or to download it to your computer for offline viewing (downloading for offline viewing requires that you first install the MacMillan Learning eBook or Bookshelf app on your device).
If you are no longer able to access the previous course, the eBook can be found at macmillan.vitalsource.com. The same email address and password that you used for Achieve must be used and you should then be able to gain free access. If you have any issues with this, contact Macmillan Support at (800) 936-6899 or via their website.

If you already own the book, there is no advantage to purchasing Achieve/Lehninger for this course since I will not be assigning homework in Achieve, and so you should opt out of Inclusive Access before the end of the second week of the course. Access the opt out option from the UCLA Store Course Materials tool in the menu on the left or the Inclusive Access materials list link on program emails or by sending an email to inclusiveaccess@asucla.ucla.edu.

Discussion information week 1

- Please attend your assigned discussion section

- Download PyMol *before* your discussion section this week

- Bring your laptop to your discussion section this week

- Homework: if you are assigned group "1A" – please submit all of question 1

Unit 1: Nucleic Acid Structure

1-Nucleosides and Nucleotides

2- A, B and Z conformations of nucleic acids double helices

3- Principles of DNA recognition by sequence-specific DNA binding proteins nucleic acids/protein interactions

4- Principles of nucleic acids Denaturation

5- DNA tertiary structures G-quadruplex

6- RNA secondary and tertiary structures
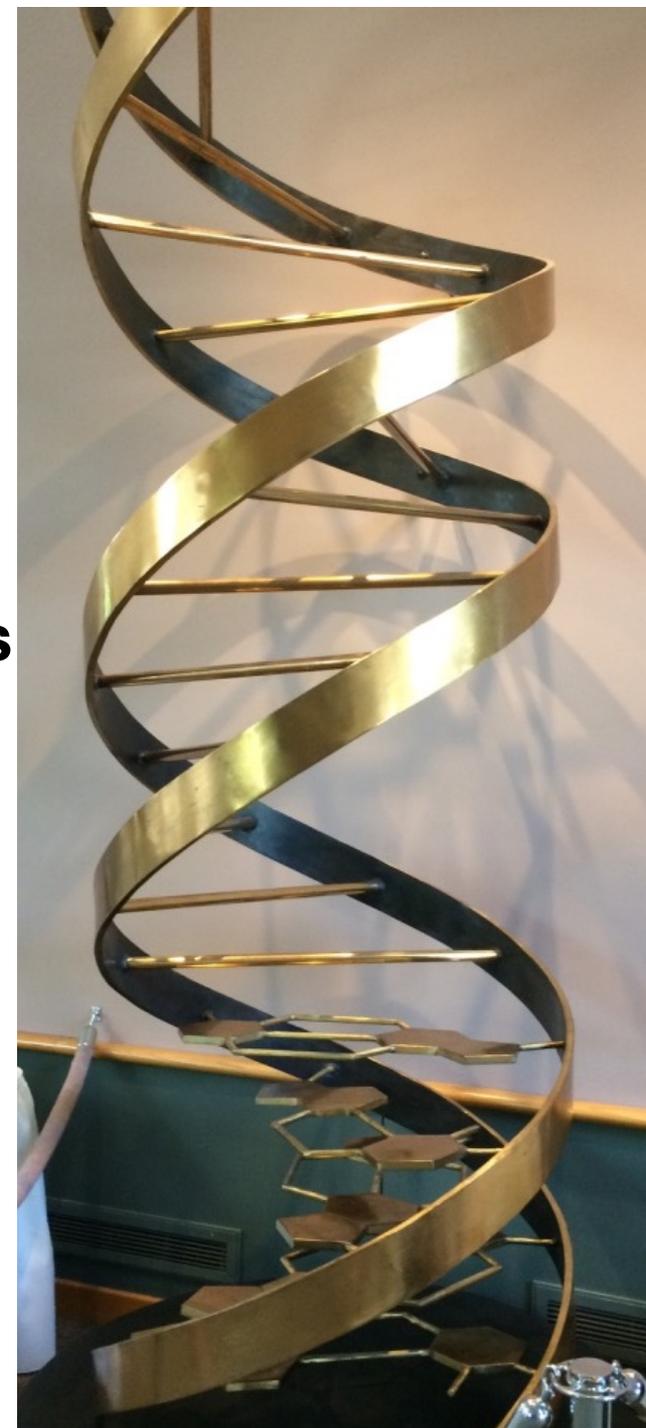
Learning outcomes:
What you need to know/understand after this unit

Memorize structures of nucleosides and understand their chemical specificities

Be able to recognize A/B/Z and quadruplex conformations of DNA/RNA and understand the main structural features in these conformations; recognize variations and differences

Understand and recognize chemical groups in proteins that interact with distinct components of nucleic acids

Understand how proteins interact with DNA in a sequence-specific or non-sequence specific manner; recognize sequence-specific vs. non-sequence specific interactions

Understand the main parameters that stabilize each conformation, as well as the transition from double-stranded to single stranded

Recognize basic elements of secondary and tertiary structure in RNA

Chemical definitions/nomenclature:

Nucleic acids are polymers
This is the monomer

Phosphate
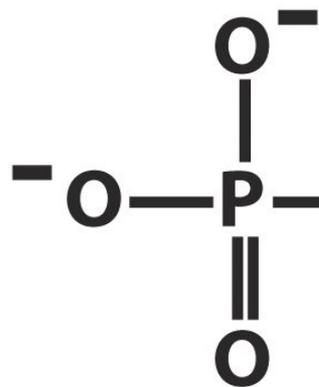
Purine or pyrimidine base

A, C, G, T
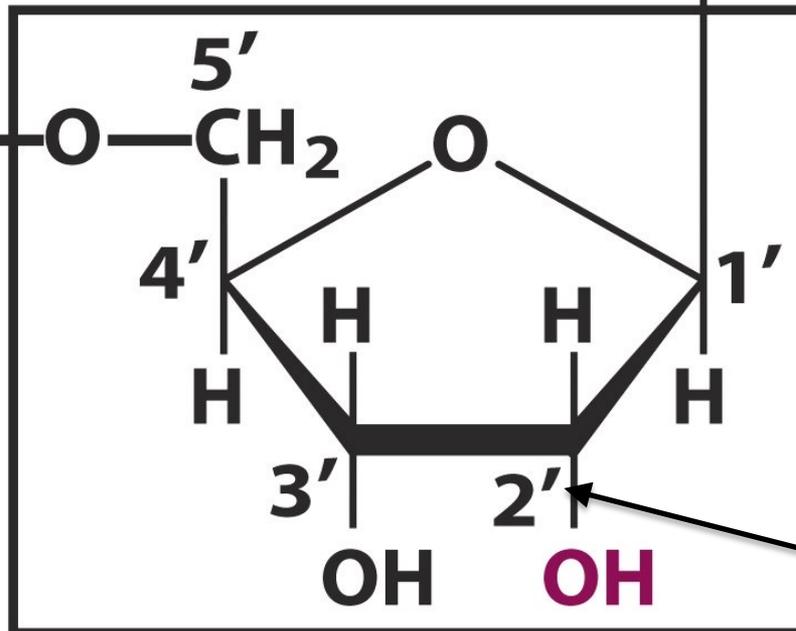
Pentose

Prime (´) notation distinguishes atoms in the sugar from atoms in the bases

Nucleoside = Pentose + base

Nucleotide = Pentose + base + phosphate

Sugar

5' end

Ribose: (RNA)

OH

5'CH₂ O OH

4' H H 1'

H H

3' 2'

OH OH

3' end
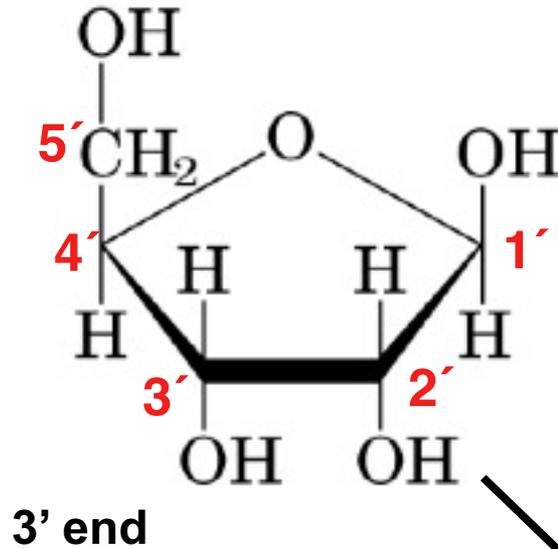
Numbering of carbons: C1´, C2´…

(´ used to prevent confusion with the numbering of atoms in bases).

The base is connected at the 1' position.

The presence of the **2´OH** confers special chemical reactivity and structural properties to RNA compared to DNA

or

Deoxyribose: (DNA)

OH

5'CH₂ O OH

4' H H 1'

H H

3' 2'

OH H
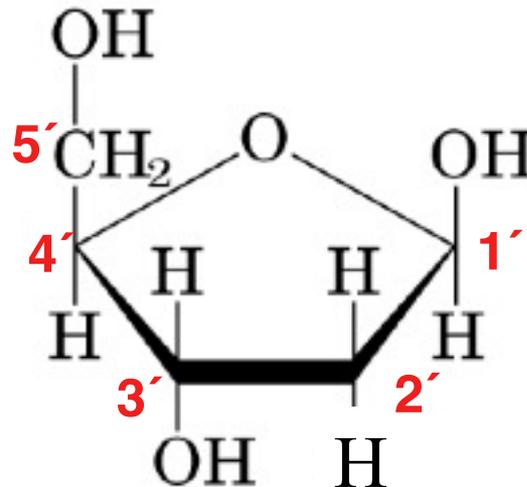
Numbering is the same in DNA and RNA

Sugar puckering: C2′ endo or C3′ endo

Planar Conformations of ribose/deoxyriboses are not stable → Puckering

2 Major Conformations:
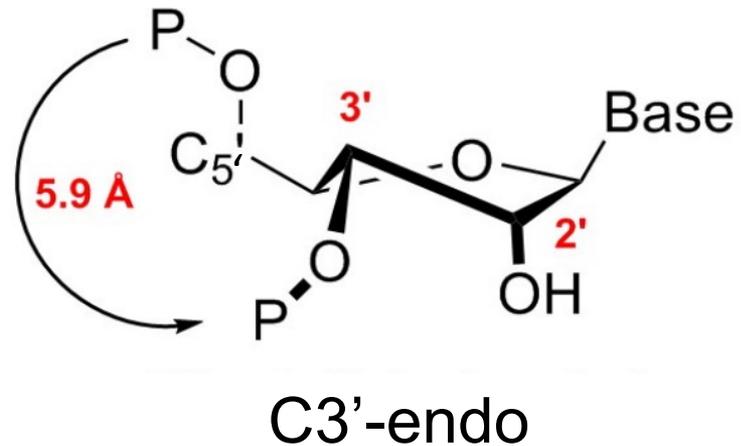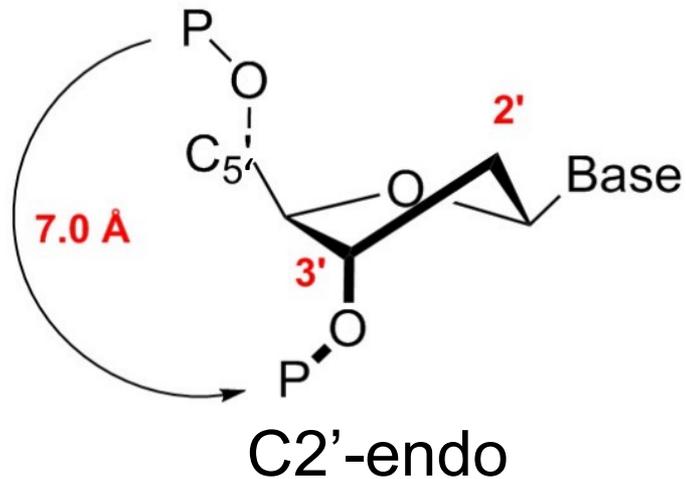
C2'-endo   C3'-endo

Ribose in polymers are constrained in the C3′ endo conformation for steric reasons (in 2' endo, the 2' OH would clash with the phosphate) -->
RNA is nearly always found as C3′ endo

Deoxyriboses in DNA are in the C2′ or C3′ endo conformation (C2' preferred)

Differences in Sugar Pucker impact distances between consecutive phosphate groups! → important to remember for polymer sizes (spacing between 5' and 3'

https://www.ibiology.org/biochemistry/rna-structure/

Phosphoric Acid

$$HO - \overset{\displaystyle O}{\underset{\displaystyle OH}{\overset{\|}{P}}} - OH$$

Pyrophosphate

$$HO - \overset{\displaystyle O}{\underset{\displaystyle OH}{\overset{\|}{P}}} - O - \overset{\displaystyle O}{\underset{\displaystyle OH}{\overset{\|}{P}}} - OH$$

Phosphates

Phosphomonoester

$$OH - \overset{\displaystyle O}{\underset{\displaystyle OH}{\overset{\|}{P}}} - O - R$$
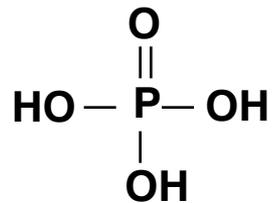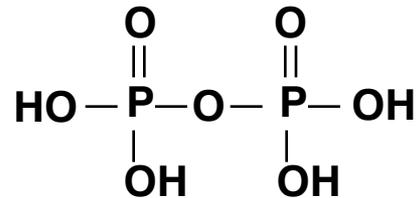
pKas = 1, 6
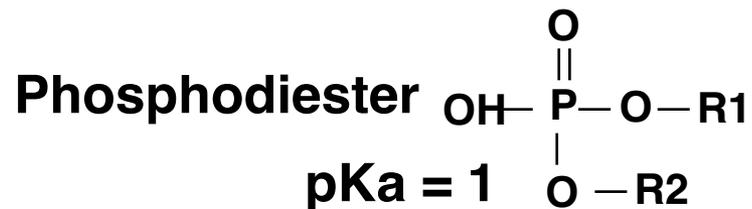
pKa important to remember for ionization of polymers at physiological pH

Phosphodiester

$$OH - \overset{\displaystyle O}{\underset{\displaystyle O - R2}{\overset{\|}{P}}} - O - R1$$

pKa = 1

An atom is protonated at pH < pKa
An atom is unprotonated at pH > pKa

Which amino acids in proteins would typically interact with phosphate groups in nucleic acids?

$$-O - \overset{\overset{\displaystyle O}{\|}}{\underset{\underset{\displaystyle O - R2}{|}}{P}} - O - R1$$

(pKa = 1)

A: Hydrophobic amino acids (e.g. Leucine)

B: Amino acids with Aromatic side chains (eg Tryptophan)

C: Positively charged amino acids (e.g. Lysine)

D: you cannot say in general - it depends whether it´s RNA or DNA
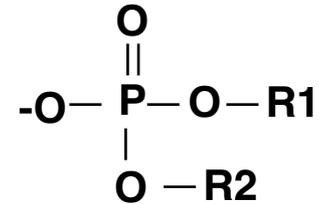
Bases

Aromaticity of bases and consequences:

Bases are planar
(no puckering)
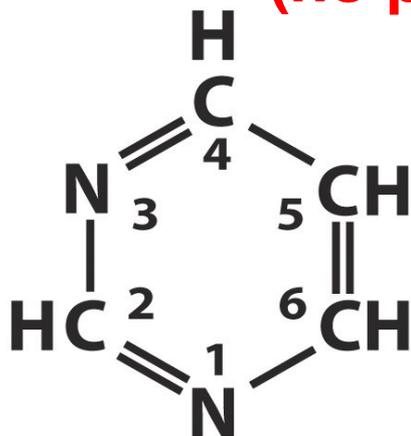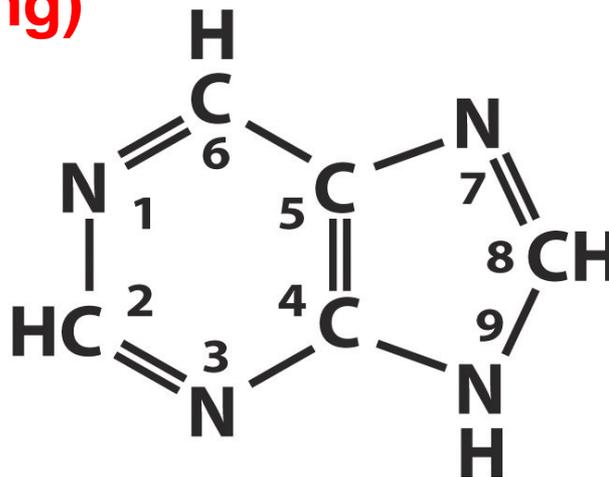
Pyrimidine   **Purine**

Bases have hydrophobic aromatic rings which promotes aromatic stacking (π stacking), an important factor in stabilizing nucleic acids in helical structures

Purine Bases and Nucleosides (abbreviated as R)

Structures need to be memorized
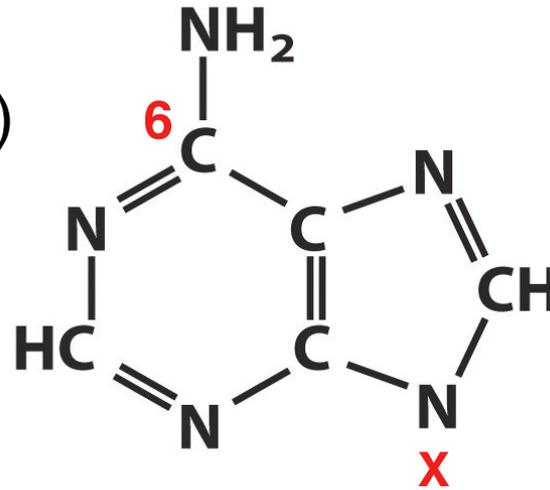
Adenine

Guanine
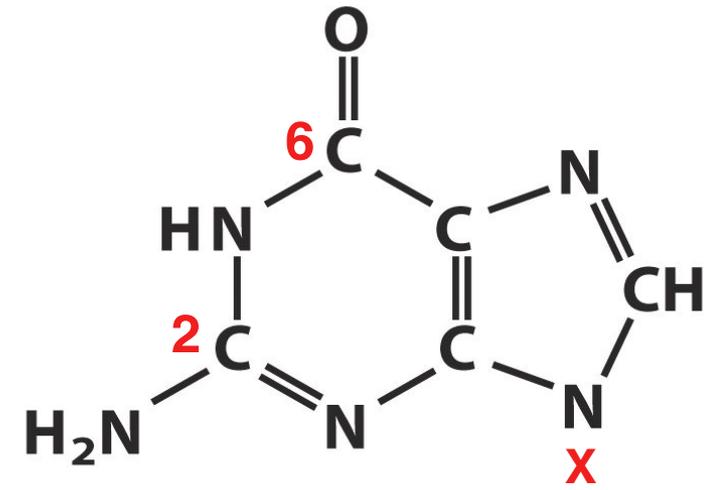
Base (X=H)	Adenine	Guanine
Nucleoside (X=pentose)	Adenosine A/dA/rA	Guanosine G/dG/rG
Nucleotide (X=pentose phosphate)	Adenosine monophosphate (AMP)	Guanosine monophosphate GMP
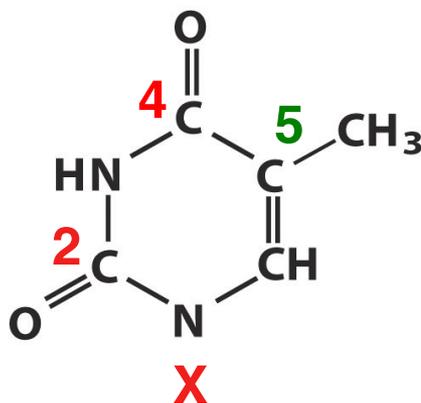
Pyrimidine Bases and Nucleosides (abbreviated as **Y**)

Molecular groups in bases can serve as Hydrogen bond donors or acceptors:

Hydrogen-bond donor

Hydrogen-bond acceptor

$$N—H \quad \text{-----} \quad N$$
$$\delta^- \quad \delta^+ \quad\quad\quad \delta^-$$

$$N—H \quad \text{-----} \quad O$$

$$O—H \quad \text{-----} \quad N$$

$$O—H \quad \text{-----} \quad O$$

These groups provide the chemical basis for interaction between two nucleic acid strands

Uracil

R

Threonine side chain

Serine side chain
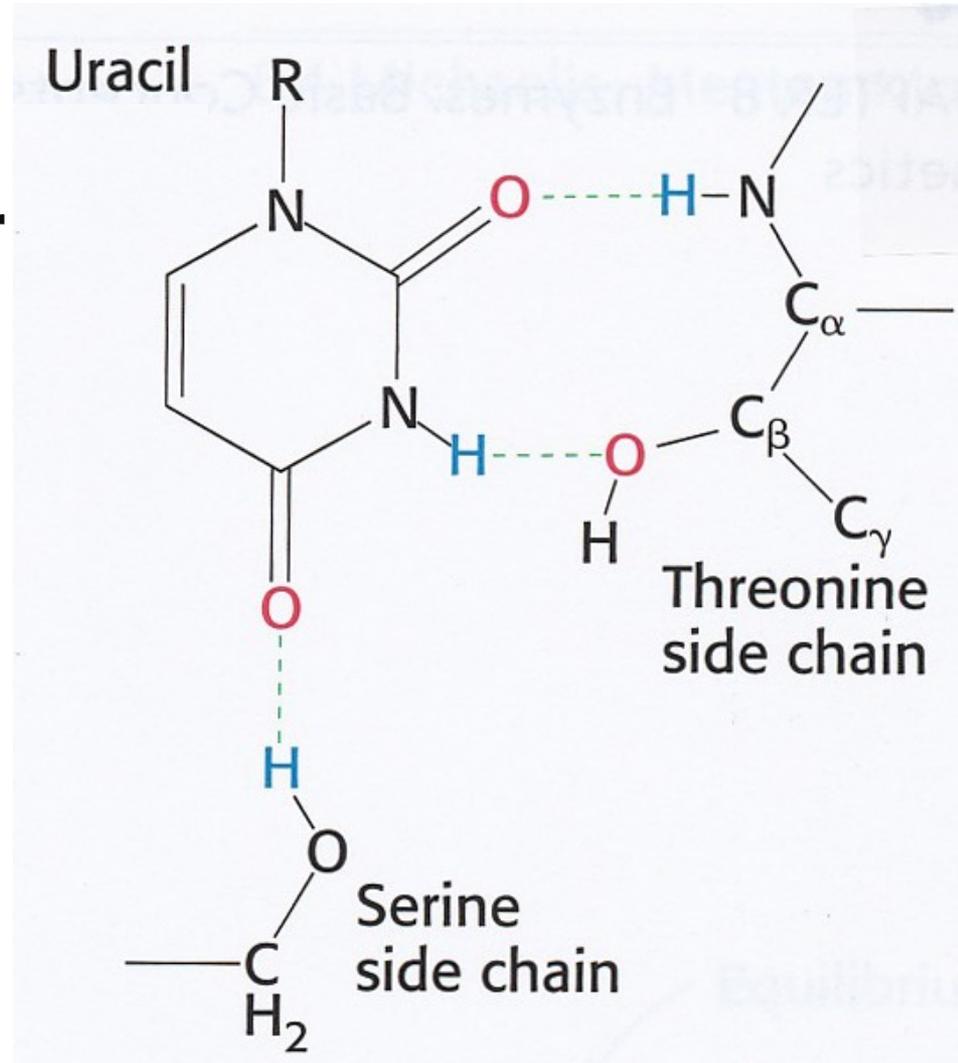
• Also provide some key elements of interactions between protein and nucleic acids

Which chemical group in pyrimidine bases cannot be used as H-bond donor or acceptor?

A: carbonyl at position 2

B: amine at position 4 of C

C: methyl at position 5 of T

D: none – they can all be used
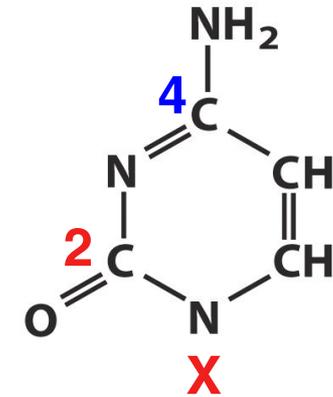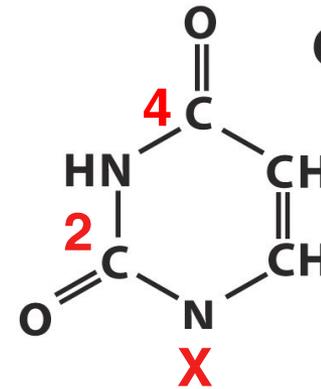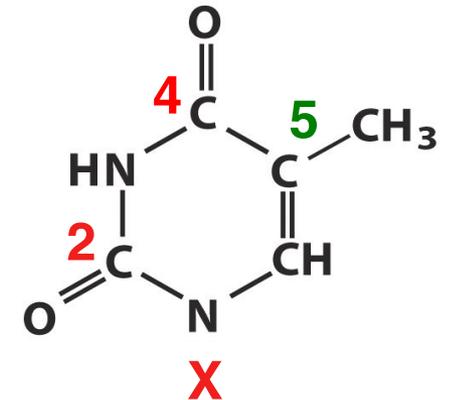
RNA (and to some extent DNA) can contain chemically modified bases which can impact the ability to form H-bonds
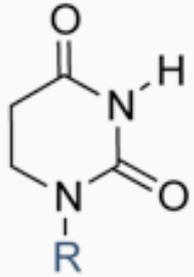
dihydrouridine
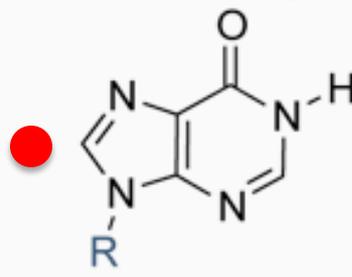
inosine
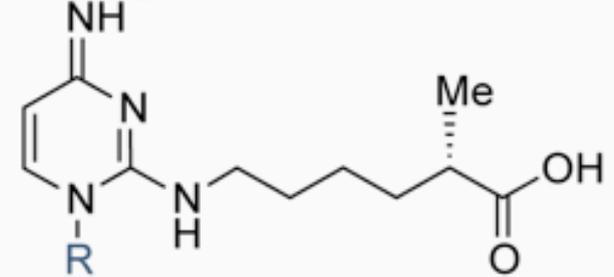
lysidine
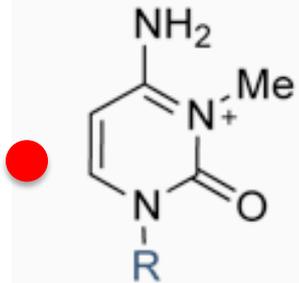
3-methylcytidine
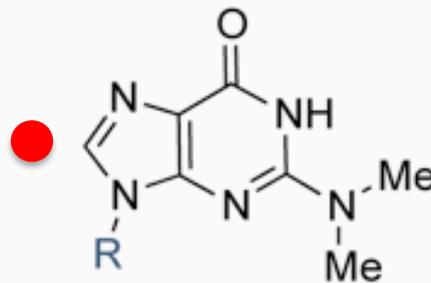
N(2)-dimethylguanosine
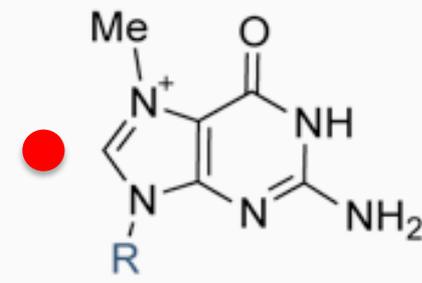
7-methylguanosine
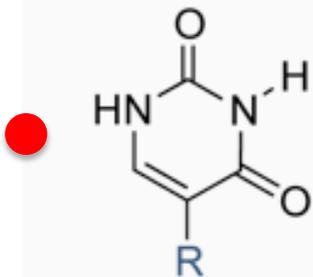
pseudouridine
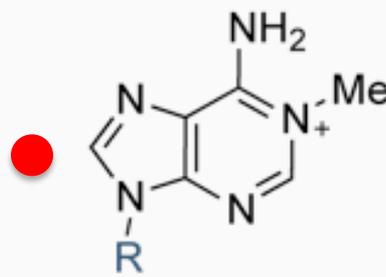
1-methyladenosine

Modifications generated by modifying enzymes on the polymers

● **= structures need to be known**

R = ribose
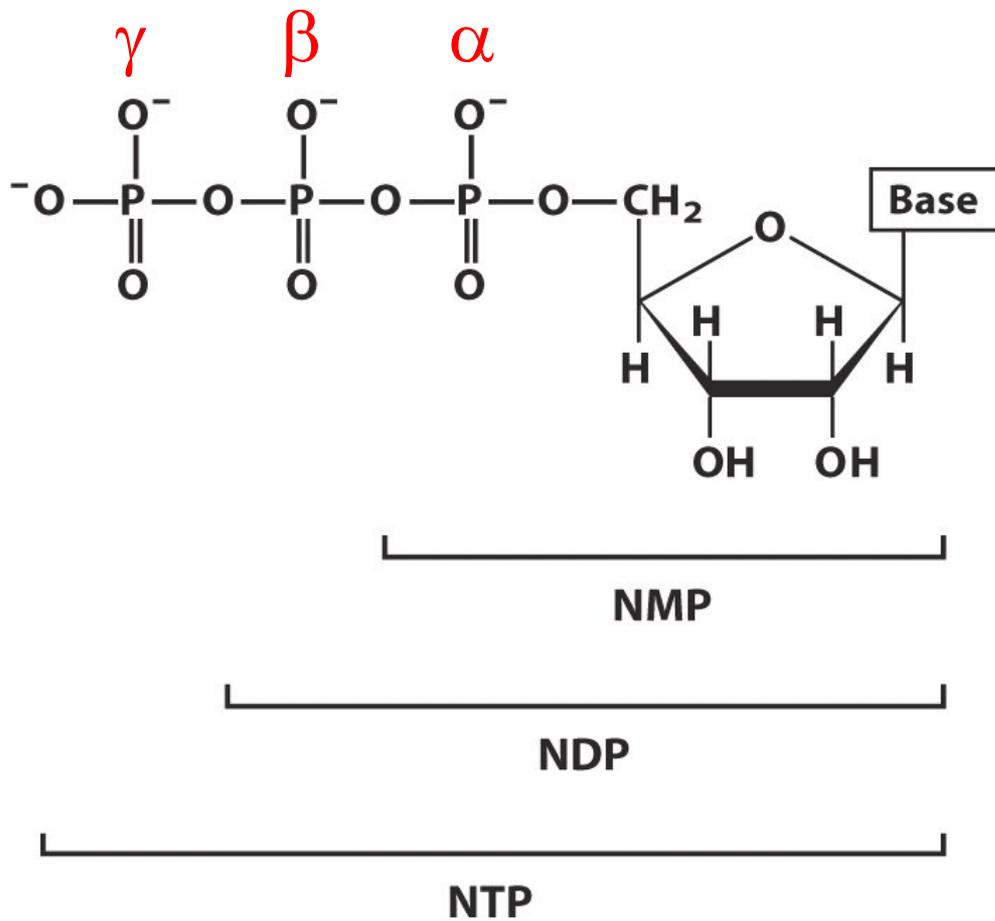
Common Bases and Nucleosides/ Nucleotides Nomenclature and abbreviations

Abbreviations of ribonucleoside 5′-phosphates

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

Abbreviations of deoxyribonucleoside 5′-phosphates

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

Polymeric Structure of Nucleic Acids
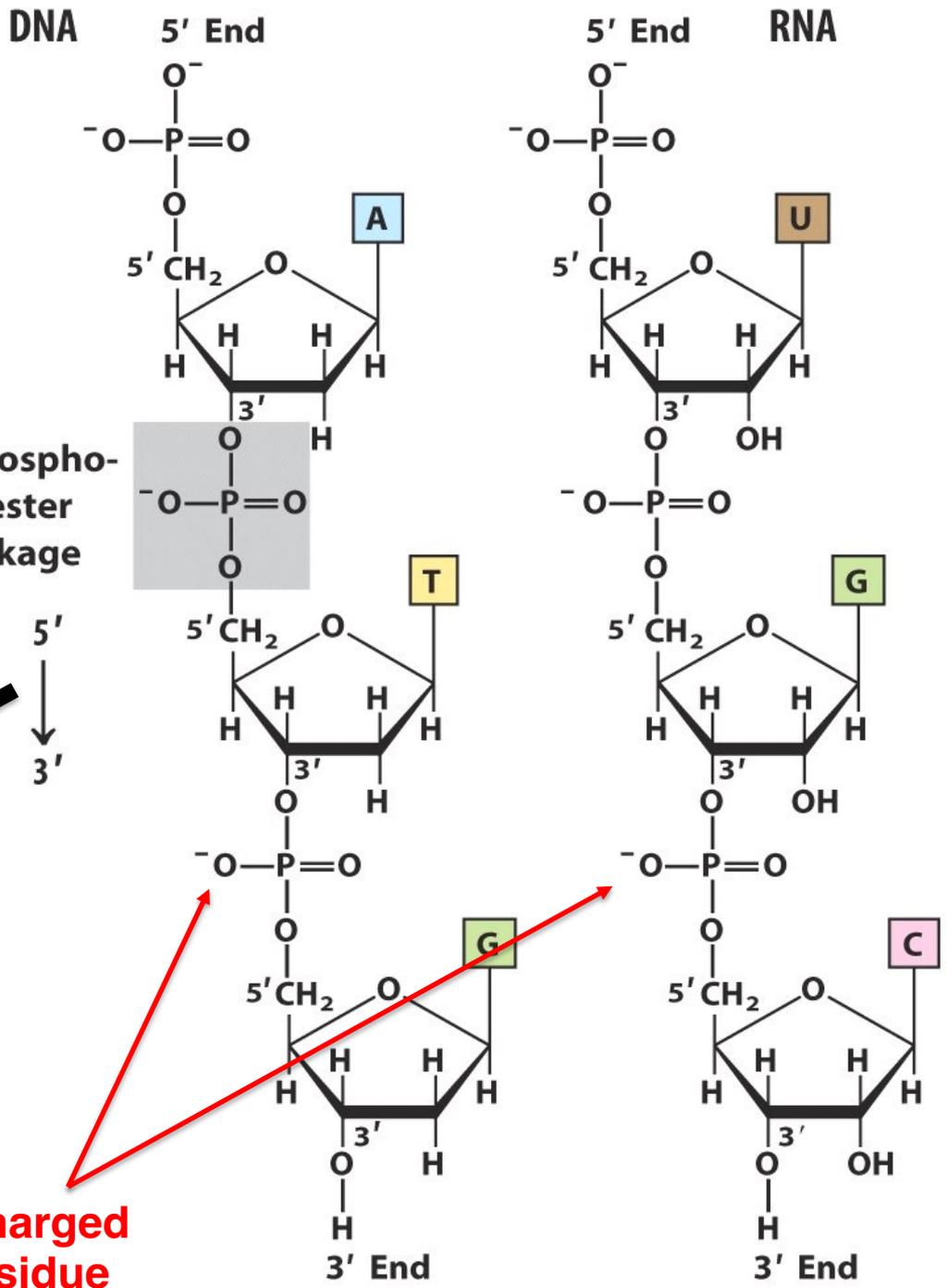
Links 3´ C of preceding nucleoside to 5´ C of next one

Indicates 5´-3´polarity Polarity as in "direction" of the polymer (not in the chemical sense)

· 1 negatively charged Oxygen per residue

DNA

5' End

$^-O-\overset{\displaystyle O^-}{\underset{\displaystyle O}{P}}=O$

A

5' CH_2

H   H

H   H

3'

O   H

Phospho-diester linkage

$^-O-P=O$

O

5'

3'

T

5' CH_2

H   H

H   H

3'

O   H

$^-O-P=O$

O

G

5'CH_2

H   H

H   H

3'

O   H

H

3' End

5' End   RNA

$^-O-\overset{\displaystyle O^-}{\underset{\displaystyle O}{P}}=O$

U

5' CH_2

H   H

H   H

3'

O   OH

$^-O-P=O$

O

G

5' CH_2

H   H

H   H

3'

O   OH

$^-O-P=O$

O

C

5'CH_2

H   H

H   H

3'

O   OH

H

3' End

The discovery of the DNA double helix

3 Side-by-Side papers in Nature in 1953....
the authors of only two got the Nobel Prize in 1962

MOLECULAR STRUCTURE OF NUCLEIC ACIDS

A Structure for Deoxyribose Nucleic Acid

Molecular Structure of Deoxypentose Nucleic Acids

Molecular Configuration in Sodium Thymonucleate

Watson & Crick

Wilkins, Stokes & Wilson

Franklin (died in 1958) & Gosling

Rosalind Franklin (1950-1953)
- **Physical chemist @ King's College London**
- **Used fiber diffraction to study DNA structure**
- **Died at age 37**
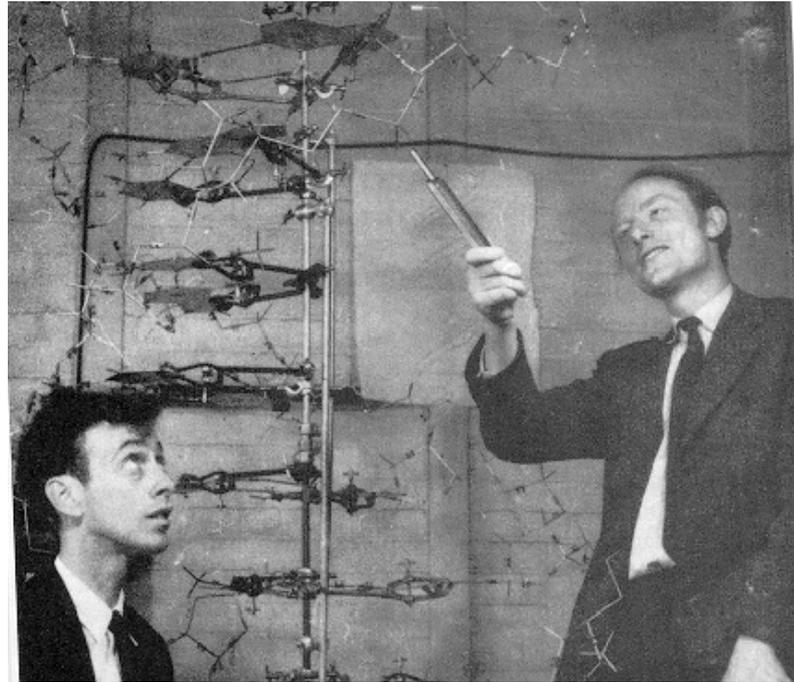
Watson and Crick (1953)
- **Theorists at Cambridge in the protein structure group**
- **Their model may have required Franklin's data…**

<u>**Watson and Crick (1953)**</u>**: "It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material".**

The (almost original) model of DNA structure by Watson & Crick (Science Museum London)
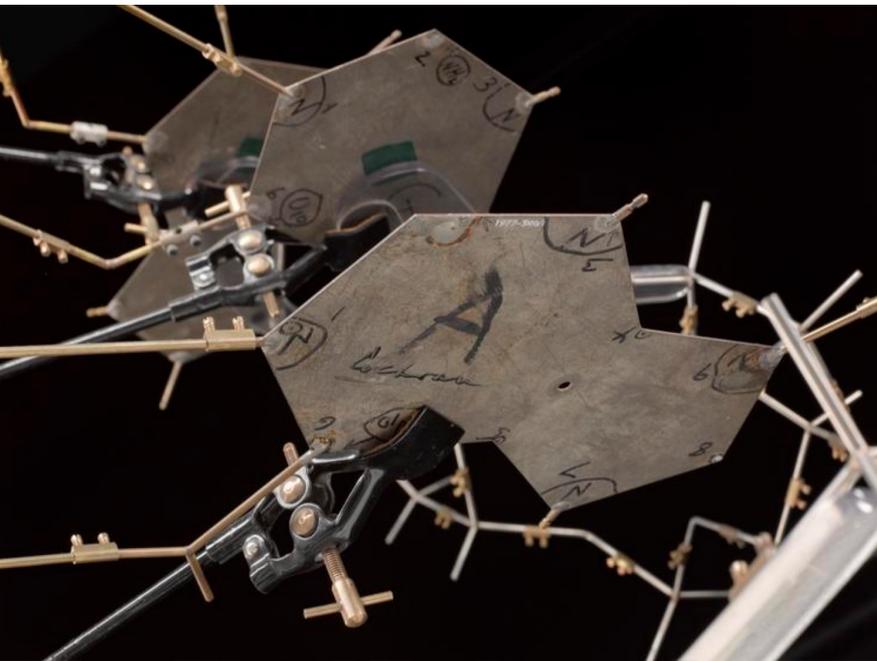
Information that Watson and Crick used to propose the double helix model:

1) R. Franklin DNA fibers
X-ray diffraction data helped to define shape and dimensions

Maltese cross shape indicates Helical pattern

- Spacing between consecutive nucleotide phosphate groups = 3.4A
- Helical Pitch = 34A
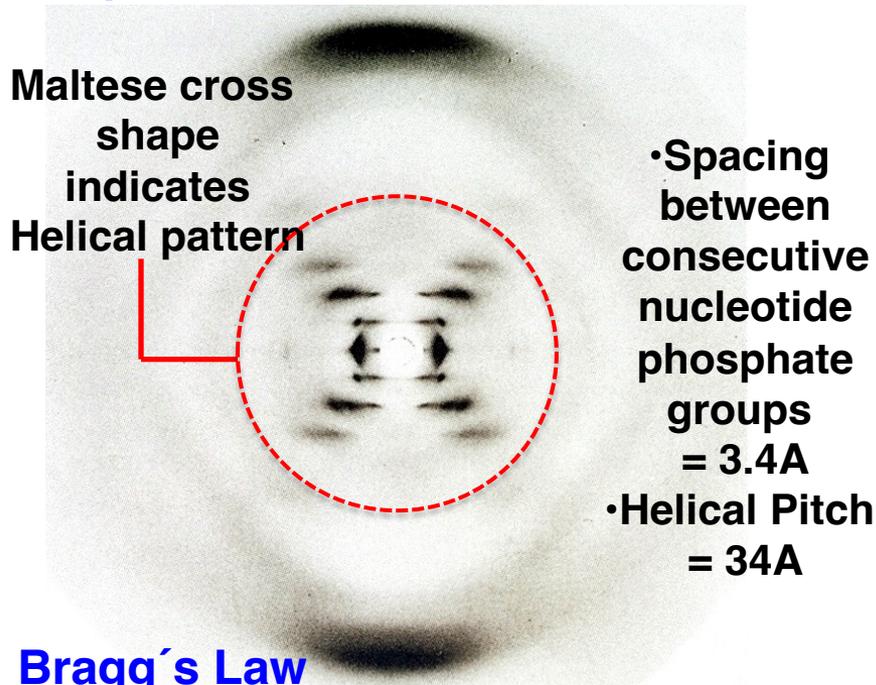
Bragg´s Law used to interpret X-ray diffraction pictures

2) bases are in the keto conformation

keto   *enol*
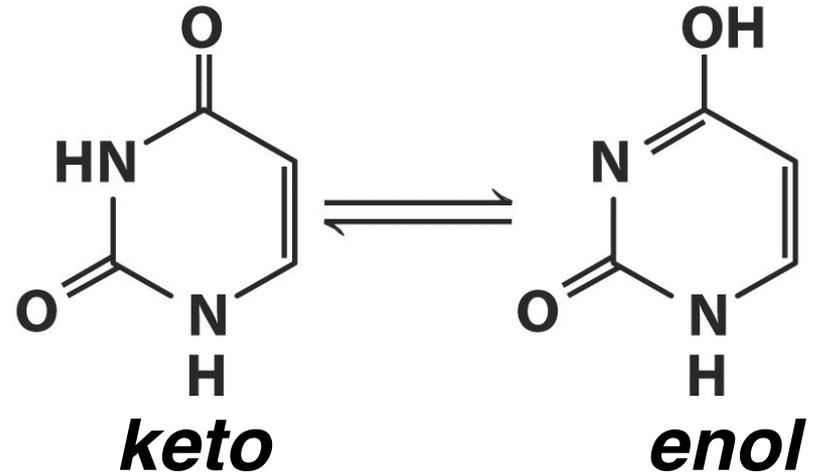
4) Density measurements:
~2 polymers/helix

5) C2´ endo sugar pucker conformation

3) Chargaff´s rules:
(G+C)/(A+T) can vary but (G+A)/(C+T) = G/C = A/T =1

<u>Chargaff. 1950</u>: "It is, however, noteworthy-whether this is more than accidental, cannot yet be said-that in all deoxypentose nucleic acids examined thus far the molar ratios of total purines to total pyrimidines, and also of adenine to thymine and of guanine to cytosine, were not far from 1".

The original model for DNA structure
Watson and Crick (1953)
Nature 171, 964-967

- See posted PDF of original paper

Essential features of the model that proved correct:

1) *Antiparallel* right-handed double helix

2) Strands are linked by complementary sets of donors and acceptor groups on bases

Helical Pitch = 34 A (10 residues/turn)
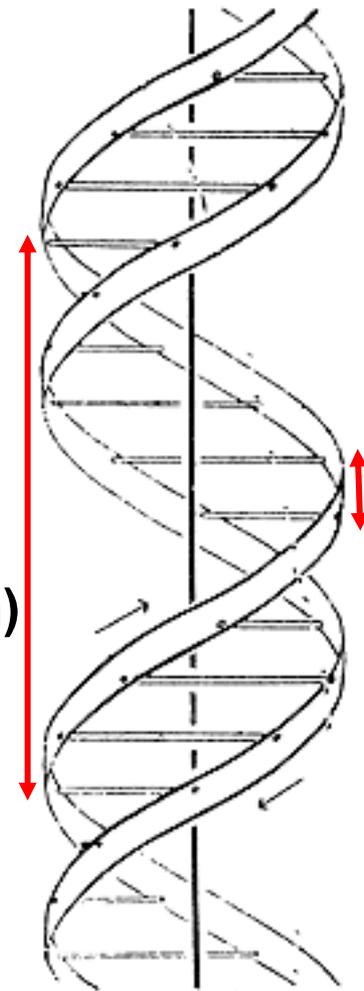
Rise/ residue = 3.4 A

This figure is purely diagrammatic. The two ribbons symbolize the two phosphate—sugar chains, and the horizontal rods the pairs of bases holding the chains together. The vertical line marks the fibre axis

Why is it important to know the exact keto/enol conformation of bases ?

keto   enol

A: The keto/enol conformations influence electrons in the π orbital system and therefore base stacking geometry
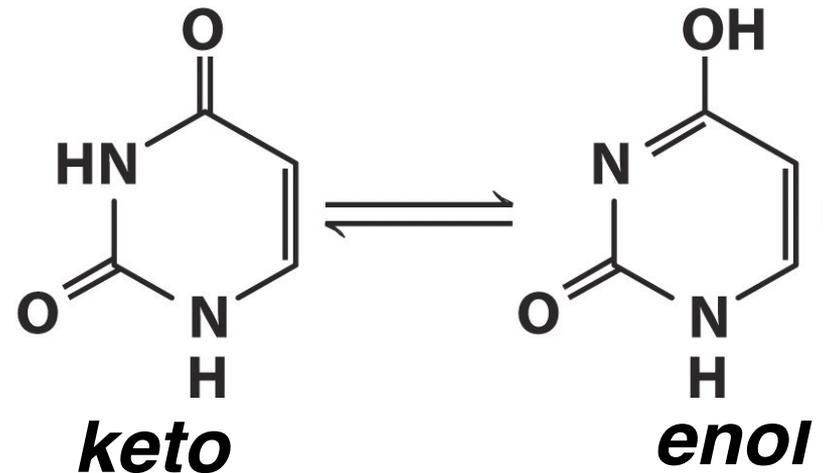
B: The keto/enol conformation indirectly impact sugar puckering because of potential steric hindrances with C2′ endo sugar pucker

C: Because bases in RNA and DNA have different keto/enol conformations

D: The keto/enol conformations affect the ability of specific bases substituents to donate/accept H-bonds and therefore base pairing partners

complementary sets of donors and acceptor groups on bases
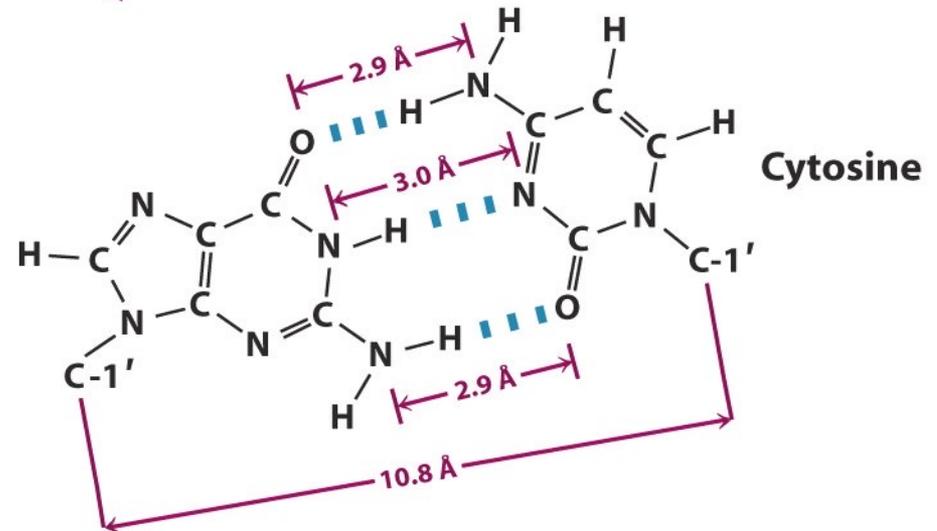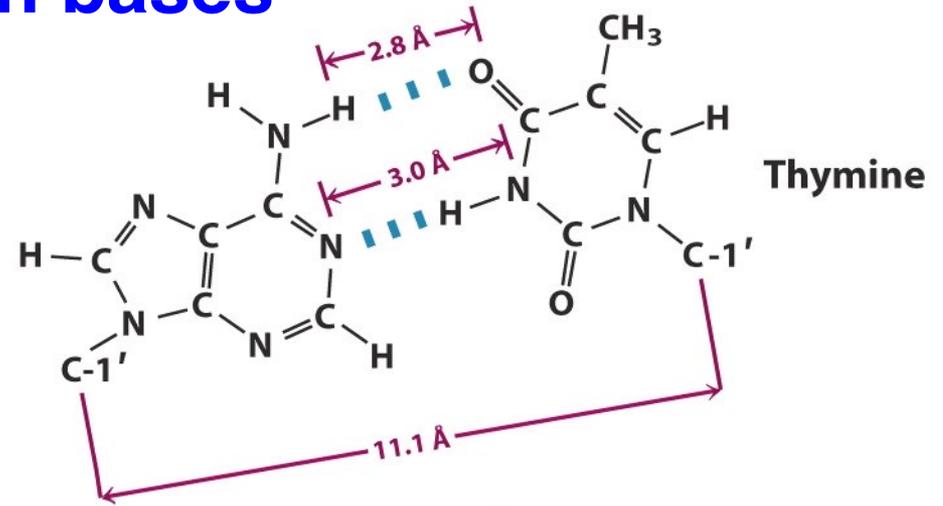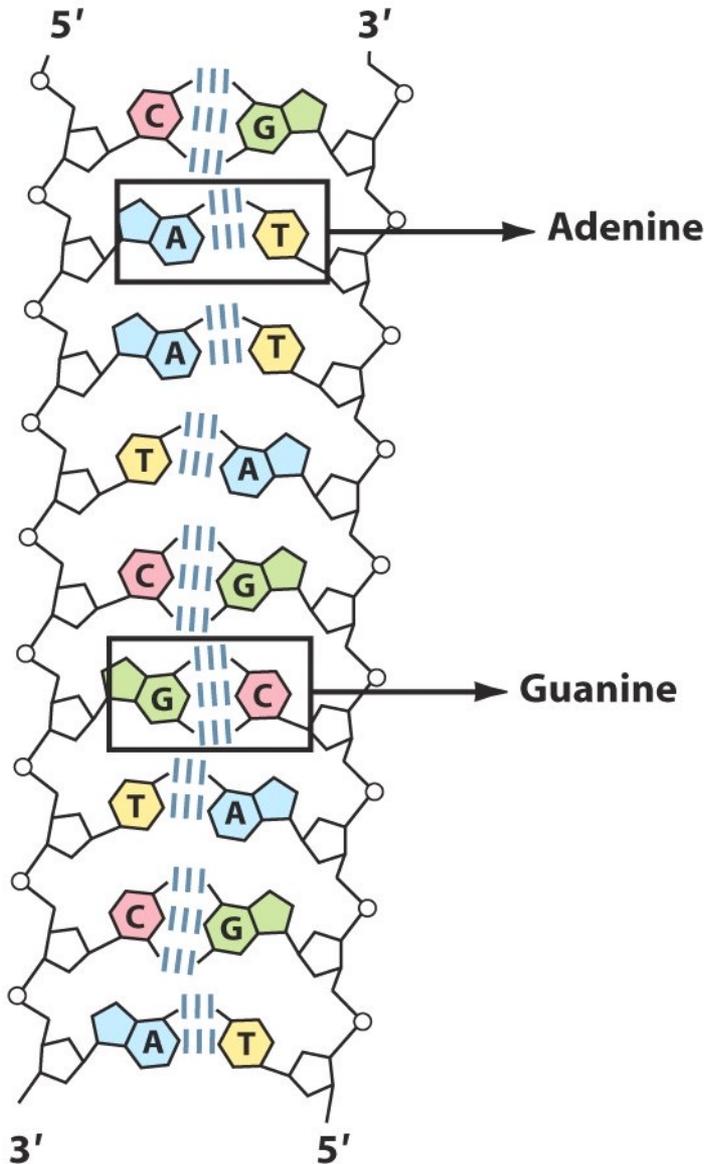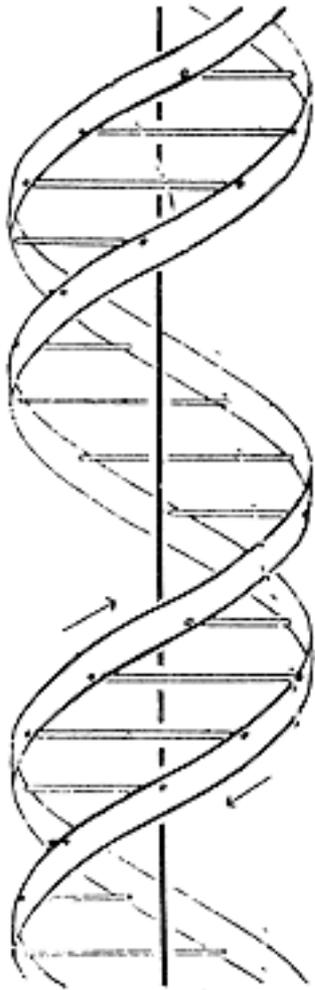
Watson-Crick Model
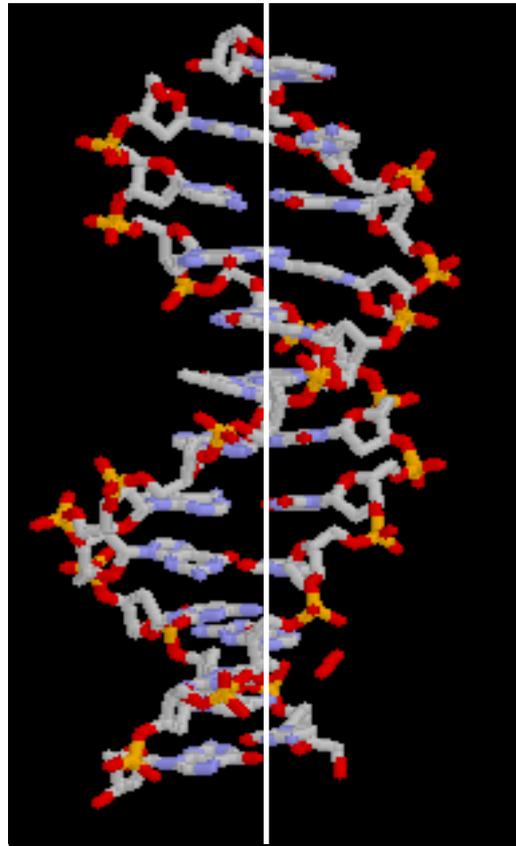
The Dickerson Dodecamer X-ray structure
(CGCGAATTCGCG)

A comparison of the Watson-Crick model (1953) and of the first B-DNA structure solved (1980)

PDB ID: 1BNA

PyMol: DickersonDodecamer.pse

Differences between the WC model and the actual DNA structure solved ?

A: the two bases forming each base pair are not aligned in a plane in the actual structure

B: Base pairs cross the helical axis in the structure while they´re shifted away from the axis in the WC model

C: No major difference

D: Base pairs are not perpendicular to the helical axis in the actual structure, as opposed to the WC model
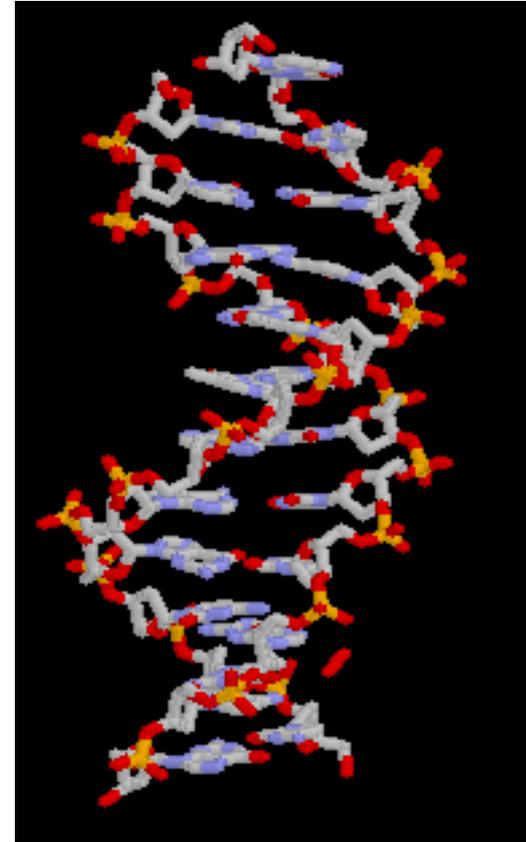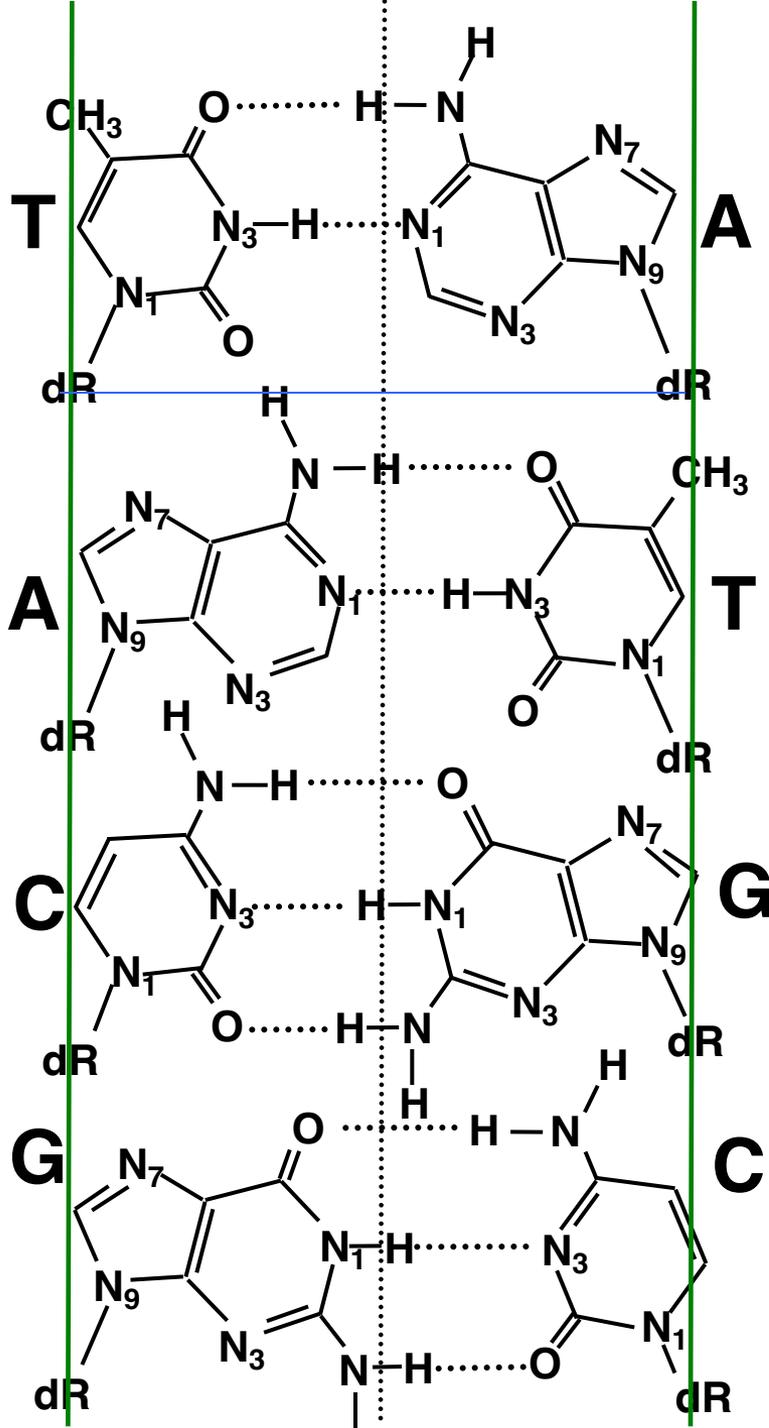
Isostericity of Watson-Crick Base Pairs (and non isostericity of non WC base pairs)

Example of a G-T non WC base pair

360° = one helical turn

10.5 bp per turn

Helix Pitch
35.7Å

34.3°

34.3° twist angle (rotation per residue)

Major Groove

Minor Groove

Helix Diameter
20Å

Helix Axis

D

Propeller twist

θ_P

C_6

C_1'

C_8

C_1'

Propeller twist: base pairs are not planar. The angle between the two planes of the bases defines the propeller twist angle
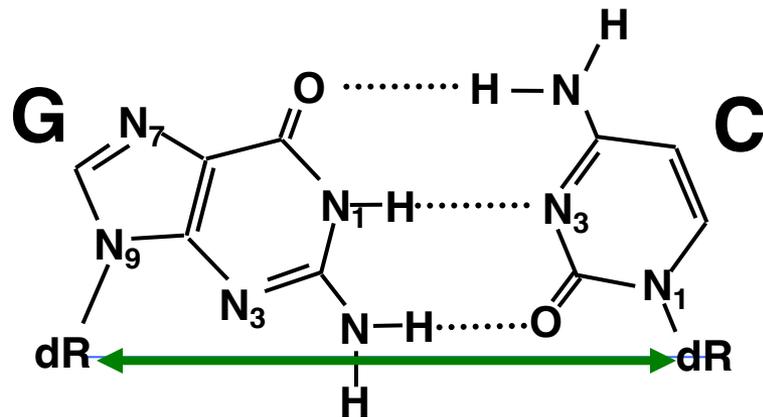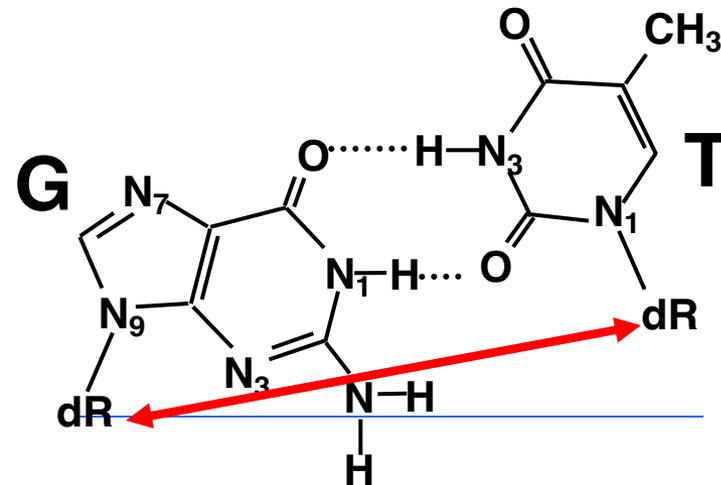
PyMol: DickersonDodecamer.pse
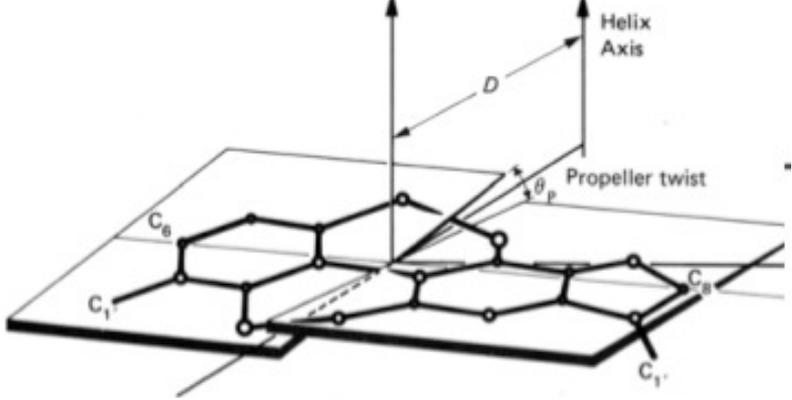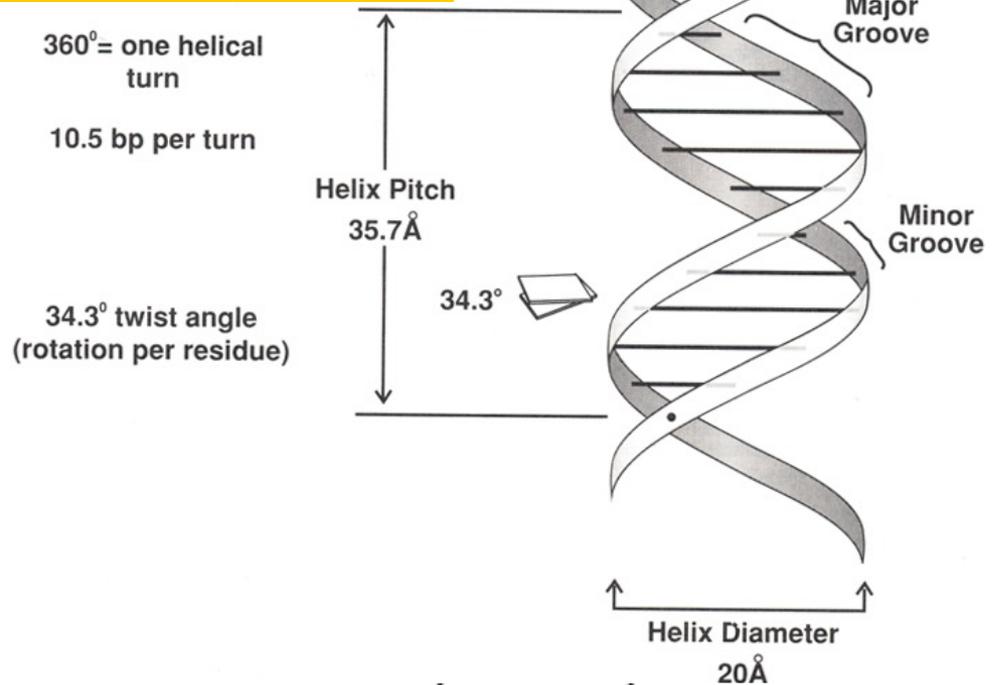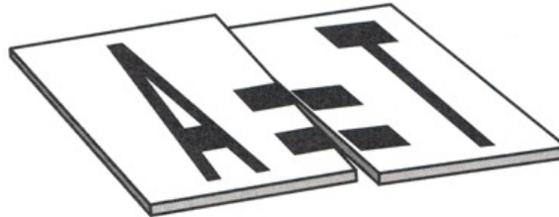
Bases of one strand are stacked onto each other:

Bases (π) Stacking

Bases (π) Stacking

Helix axis (vertical)

Twist, t

H

O----H-N

N

N-H----N

Roll, θ_R

N

N

C_1'

C_1'

N-H----O

H

θ_T, Tilt

Pseudodyad

Base pair tilt: An average mean plane of the base pair can be defined (ignoring the propeller twist) – see gray area above. The rotation of this plane about the pseudo-dyad axis defines base pair tilt.

Base pair roll: the degree of departure of the mean plane of the base pairs from the perpendicular helix axis on the short axis of the base pairs.

Helical twist. Defines the degree of rotation when moving vertically from one base pair to the next.
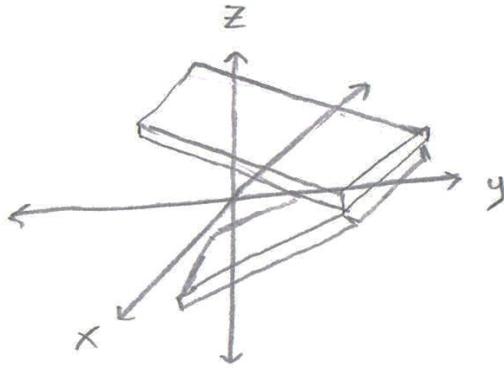
Tilt
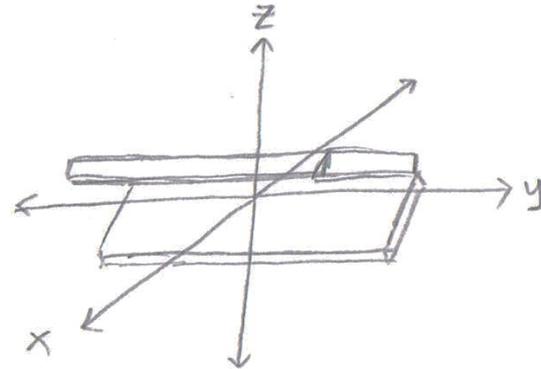
Roll
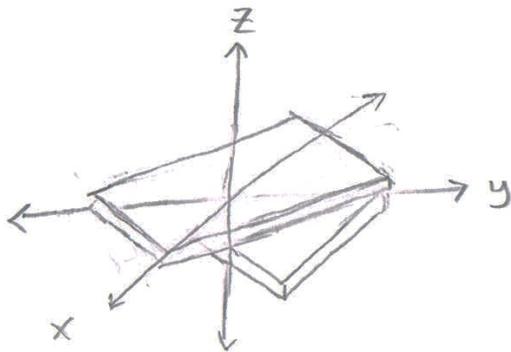
Twist
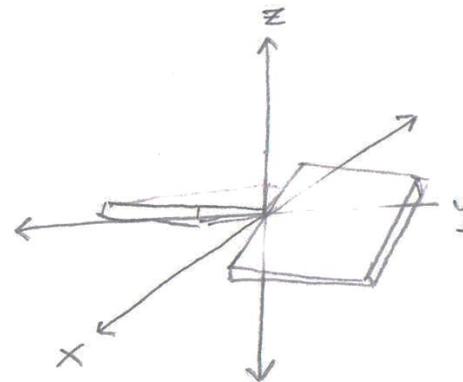
Propeller Twist

Major Groove and Minor Groove in B-form double helices:

Helical Axis

Major Groove Side
α >180°

α
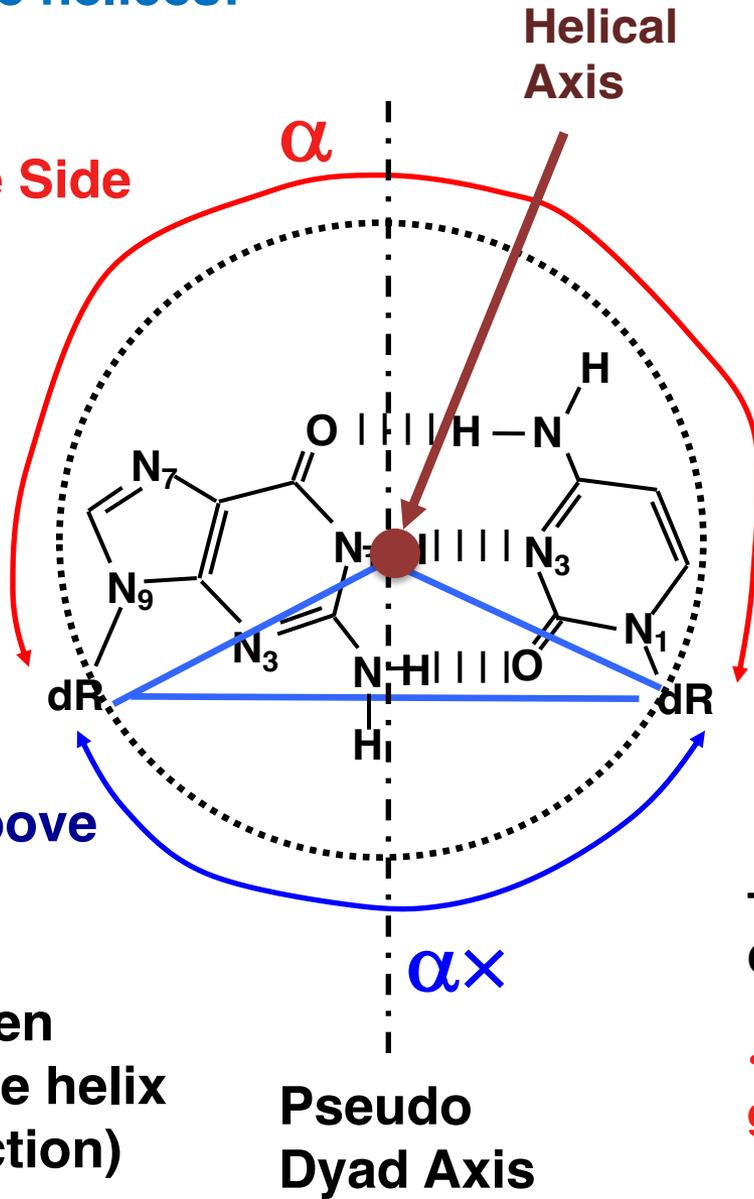
Minor Groove
$\alpha\times$ <180°

$\alpha\times$

Base pairs seen from above the helix (helical projection)

Pseudo Dyad Axis

3.4 Å

Minor groove

Major groove

36 Å

20 Å
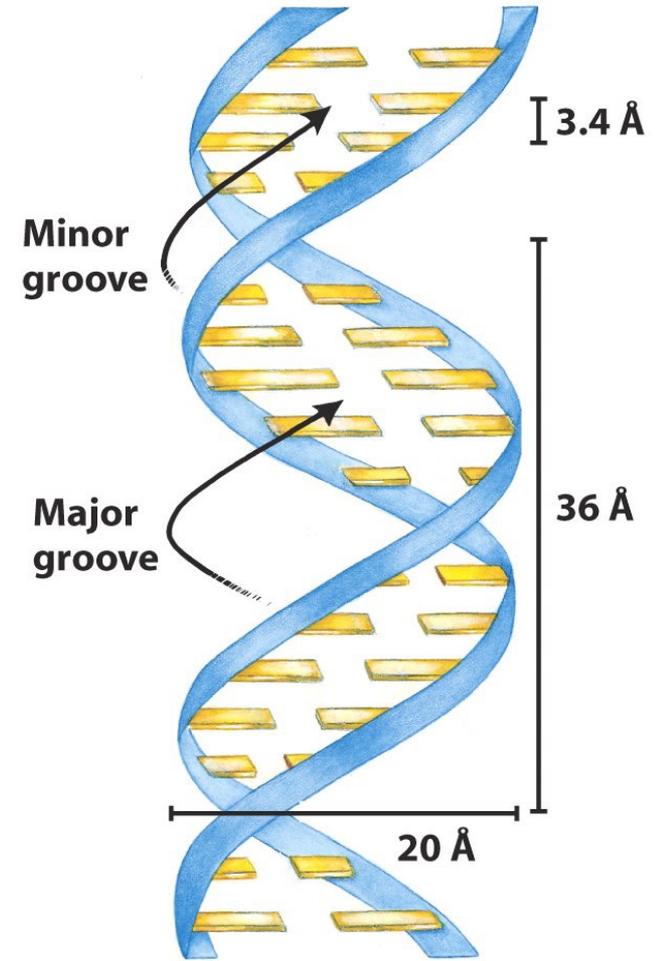
Tip to visualize Major and Minor Groove sides in structures:

·N7 of Purines always on major groove side

· N3 of Purines always on minor groove side

PyMol: DickersonDodecamer.pse

Which grooves are we looking at?

A: 1-Major 2-Major

B: 1-Minor 2-Minor

C: 1-Major 2-Minor

D: 1-Minor 2-Major

A vs B DNA

$$H_2O$$

$$A \rightleftharpoons B$$

A-dsRNA = 2KYD

A

PDB ID = 115D

B

PDB ID 1BNA
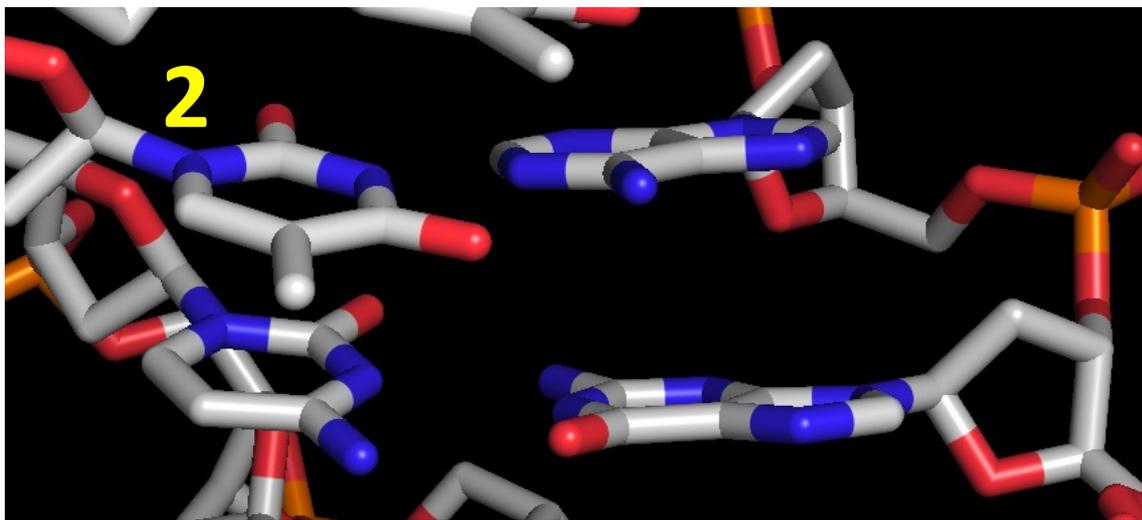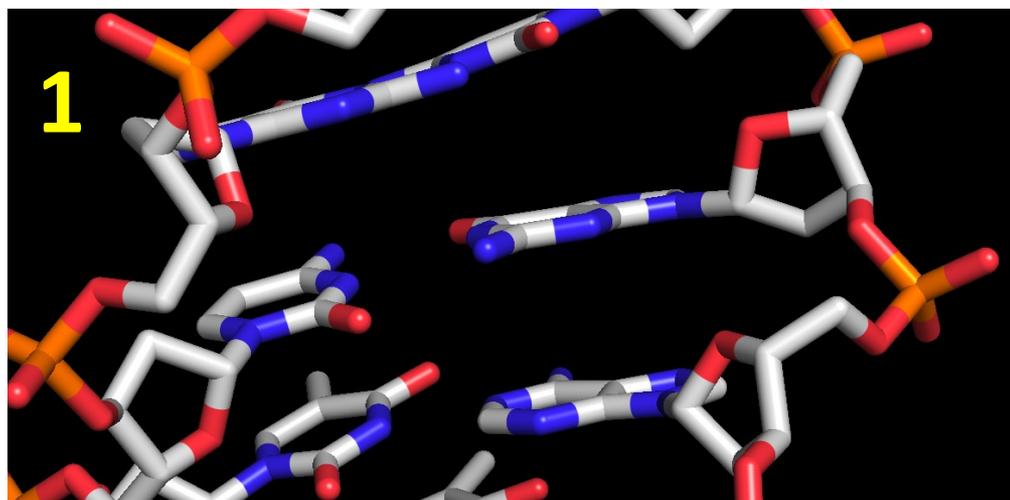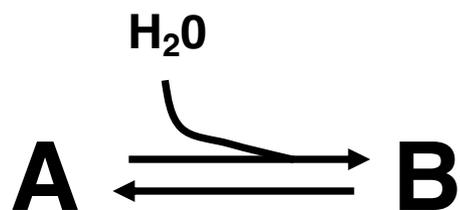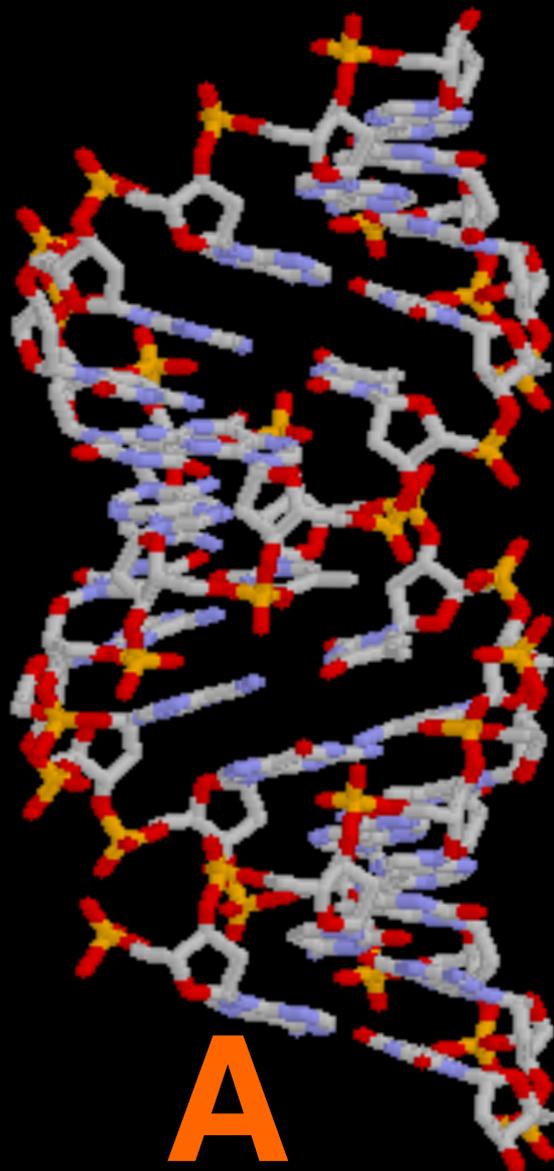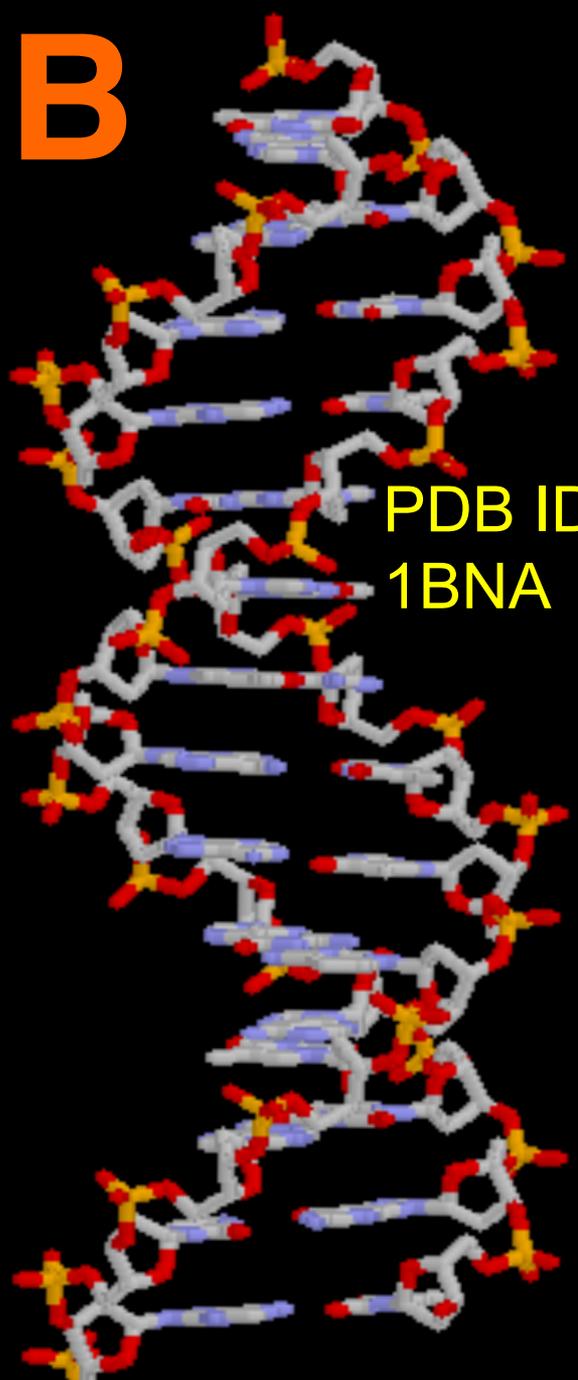
What differences can you see between A and B forms?

A: A is more compact than B (shorter helical pitch)

B: Base pairs are more tilted in A vs. B.

C: Sugar pucker is C3´endo in A, C2´endo in B

D: Major groove is deeper in B than A
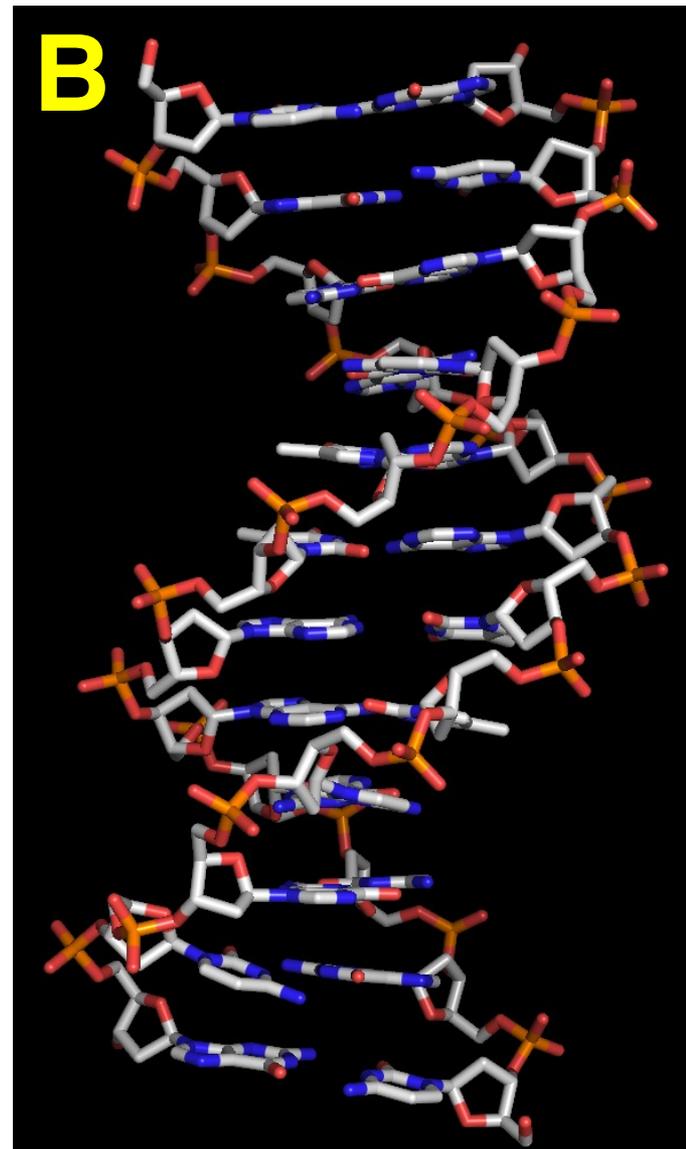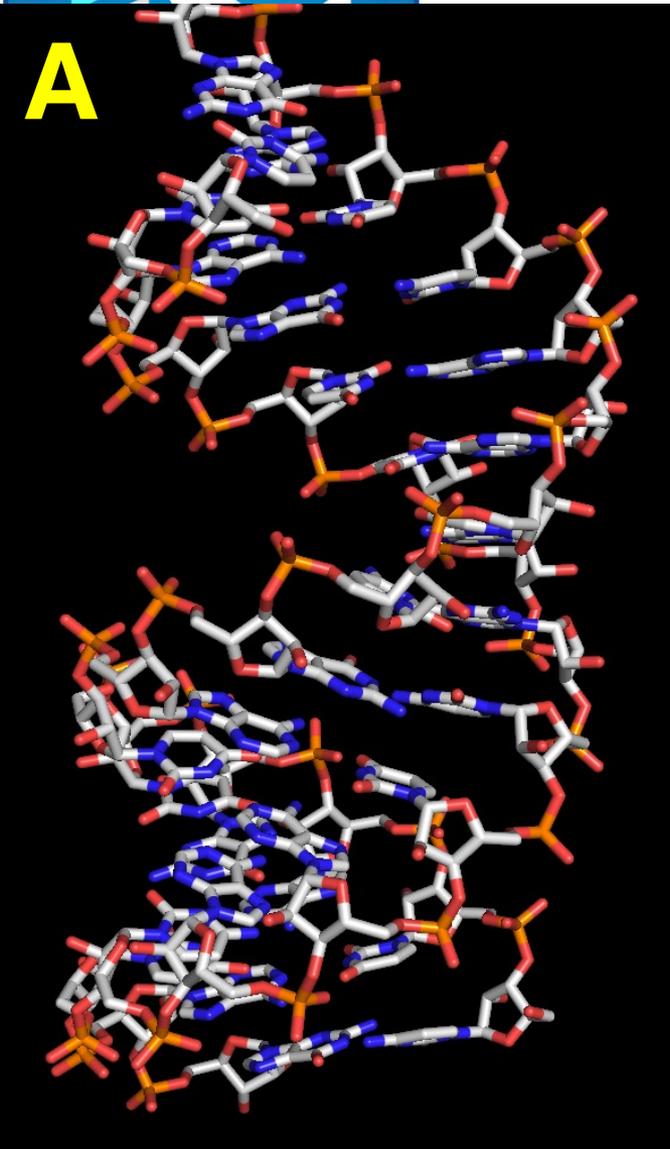
A-DNA

B-DNA

BP rise ~2.5 Å Diameter ~23 Å

BP rise ~3.4 Å Diameter ~18 Å

	B-DNA	A-DNA
Sugar pucker	C2'-endo	C3'-endo
Rise/residue	3.4 Å	2.6 Å
Residues/turn	10.5	11
Helical twist	34°	33°
Diameter	20 Å	26 Å
Tilt	6°	20°
Propellor twist	12°	15°

Major differences :

- A DNA is shorter than B DNA: 1 helix turn is 28.6 A vs 34 A for B DNA. This is due to the 3´ endo sugar pucker in A

- Bases of A-DNA are shifted away from the helical axis. This results in a deep major groove and in a shallow minor groove. There is a 6 A hole in helical projection.

Exact values need not to be remembered…

Impact of Sugar puckering (C2´or C3´ endo) on double helix geometry

~7 Å

P
O5´
C5´
O
2´
Base
3´
O3´
P

C2´-endo

B-Form (DNA)

BP rise ~3.4 Å

~6 Å

P
O5´
C5´
3´
O3´
O
2´
Base
P
O2´H

C3´-endo
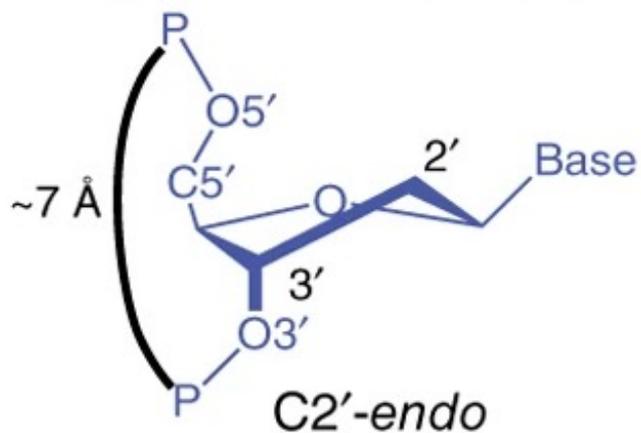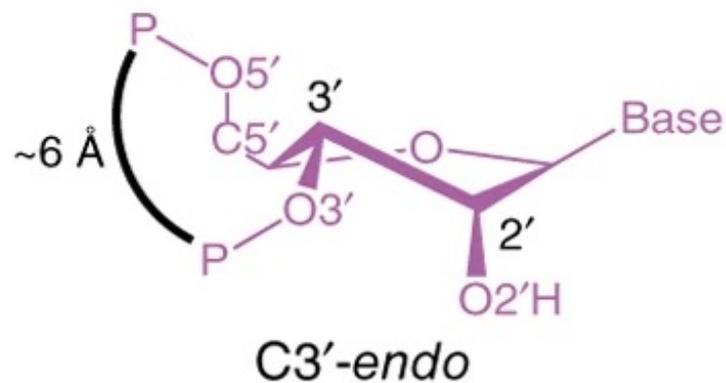
A-Form (DNA or dsRNA)

BP rise ~2.5 Å

Base tilting: A vs. B DNA

A-DNA/dsRNA
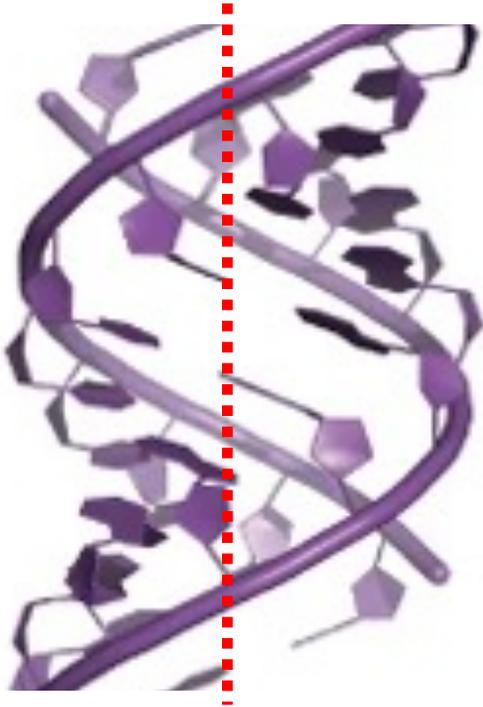
B-DNA
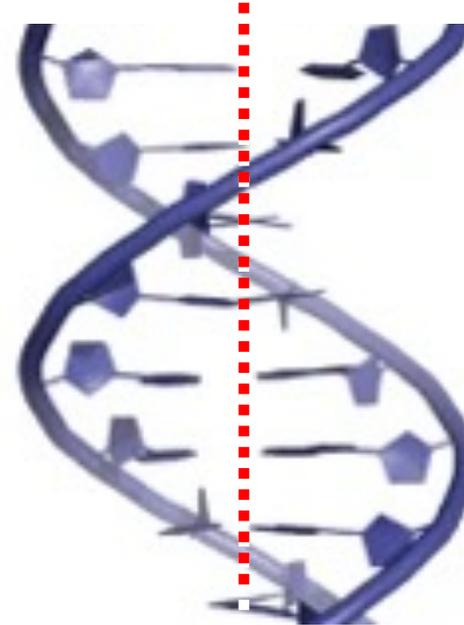
Consider
Angles between
the base pairs
and helical axis
(vertical dashed
lines):

\> or < 90° = tilt

90° = no tilt
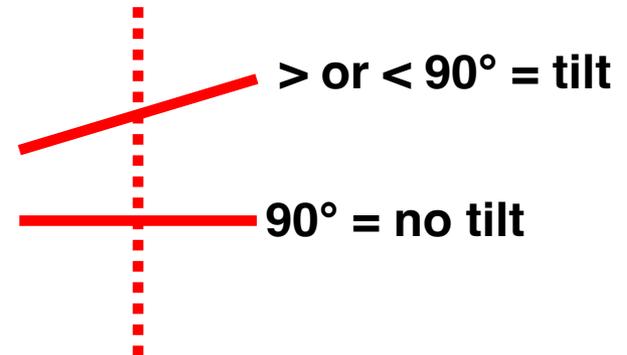
Water is essential in the transition A <--> B DNA

H_2O

A $\rightleftharpoons$ B
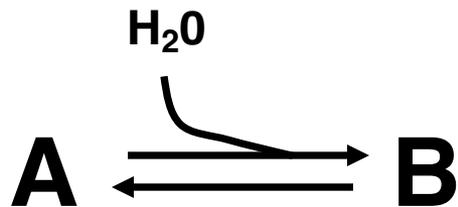
A water spine (green dots) has been proposed to exist in the minor groove of B-DNA that would stabilize the B-form

This concept is controversial and will not be detailed further
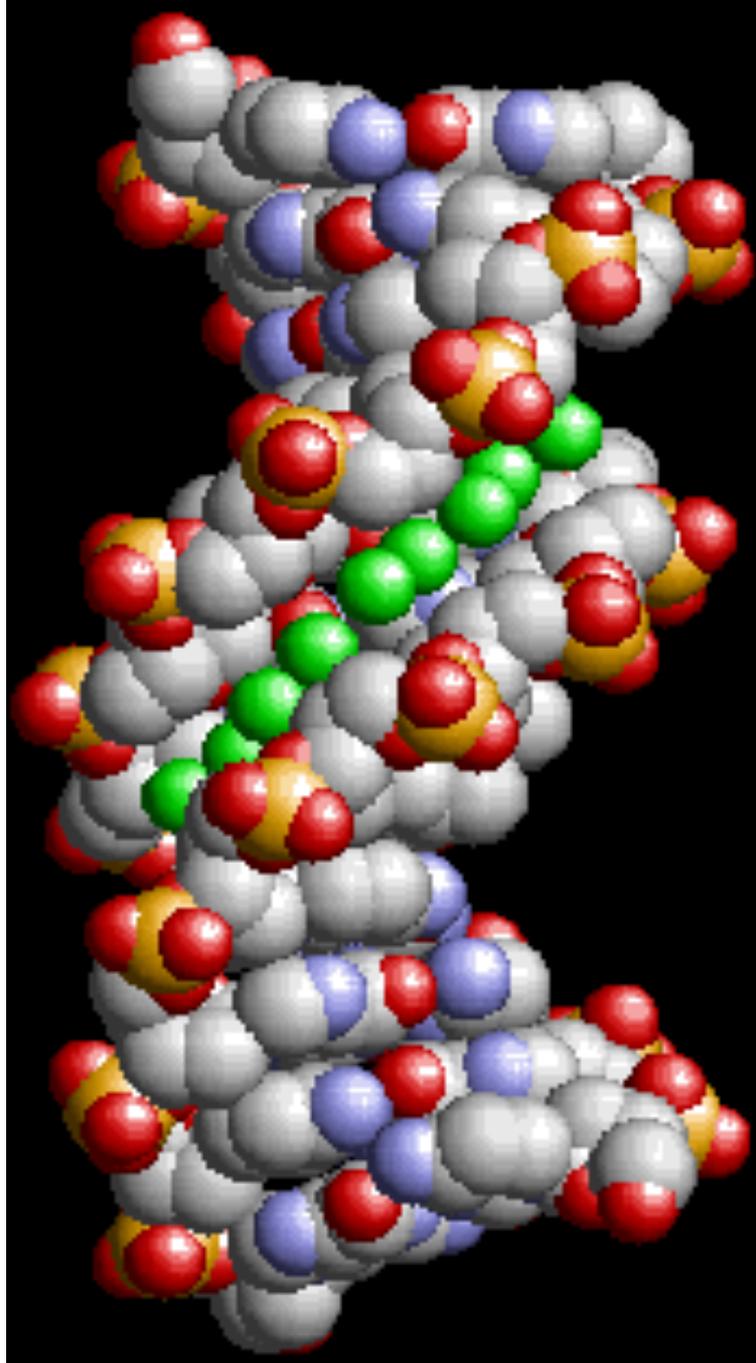
Glycosidic bond Anti /Syn conformations

Anti and Syn conformations are defined based on the torsion angle of the glycosidic bond

The sequence of atoms chosen to define the torsion angle to define anti/syn conformation is: O4'-C1'-N9-C4 for purines - O4'-C1'-N1-C2 for pyrimidines.

Syn A/G:
C1´-O and N9-C4 are pointing in same direction
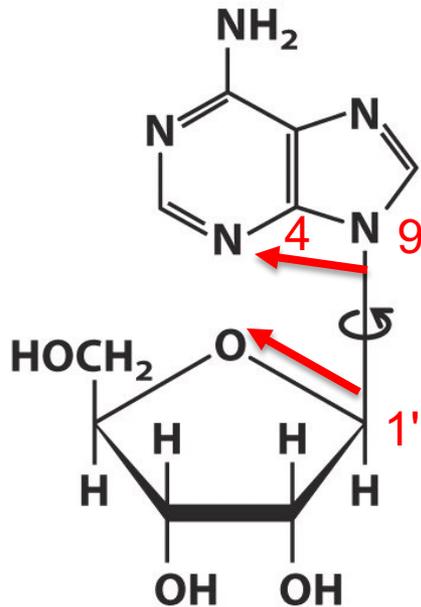
Anti A/G:
C1´-O and N9-C4 are pointing away from each other

Anti C/T:
C1´-O and N1-C2 are pointing away from each other
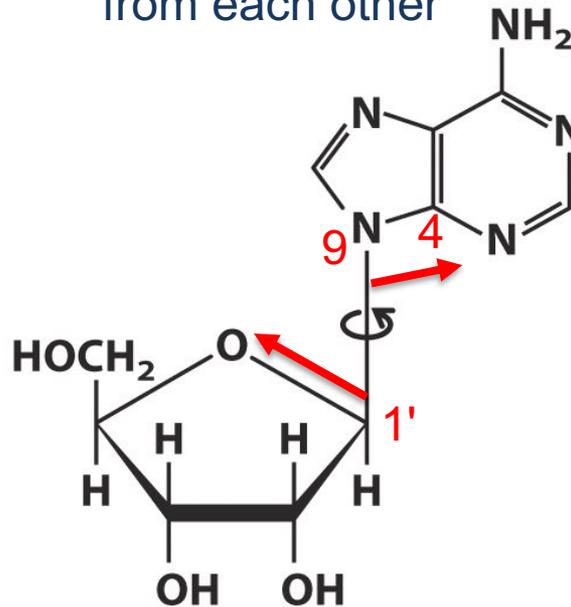
syn-Adenosine
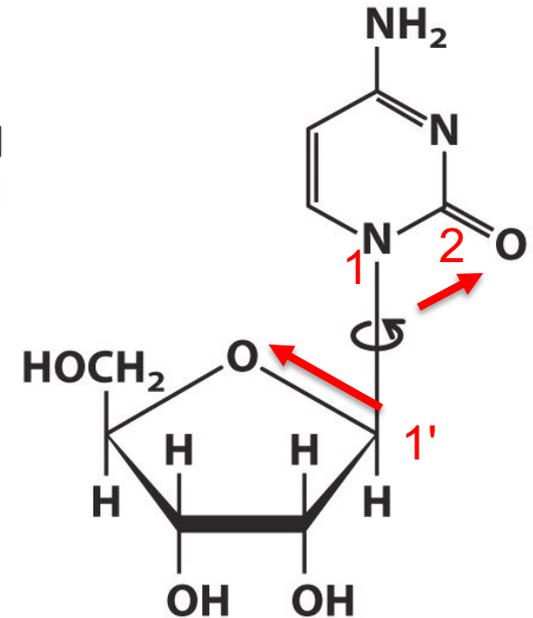
anti-Adenosine

anti-Cytidine

How to recognize Anti /Syn Glycosidic bonds

Syn A/G:
N_3 of base is **close/**
on top of **cyclic O of ribose**
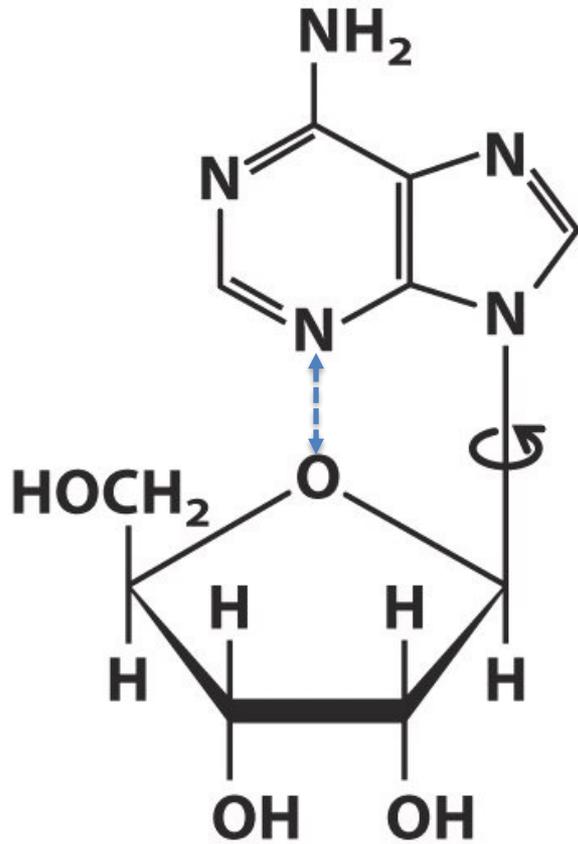
Anti A/G:
N_3 of base is **away** from
the **cyclic O of ribose**
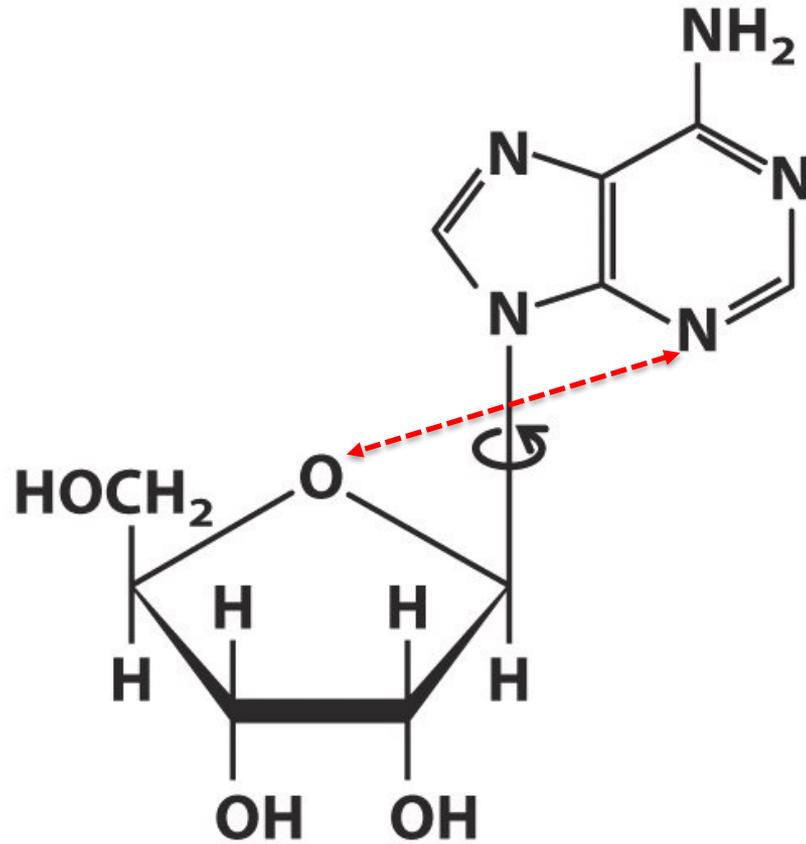
syn-Adenosine

anti-Adenosine

Z-DNA

- **Occurs in DNA sequences with stretches of consecutive G-C base pairs**

- **Requires high salt concentrations for forming stable structures in vitro**
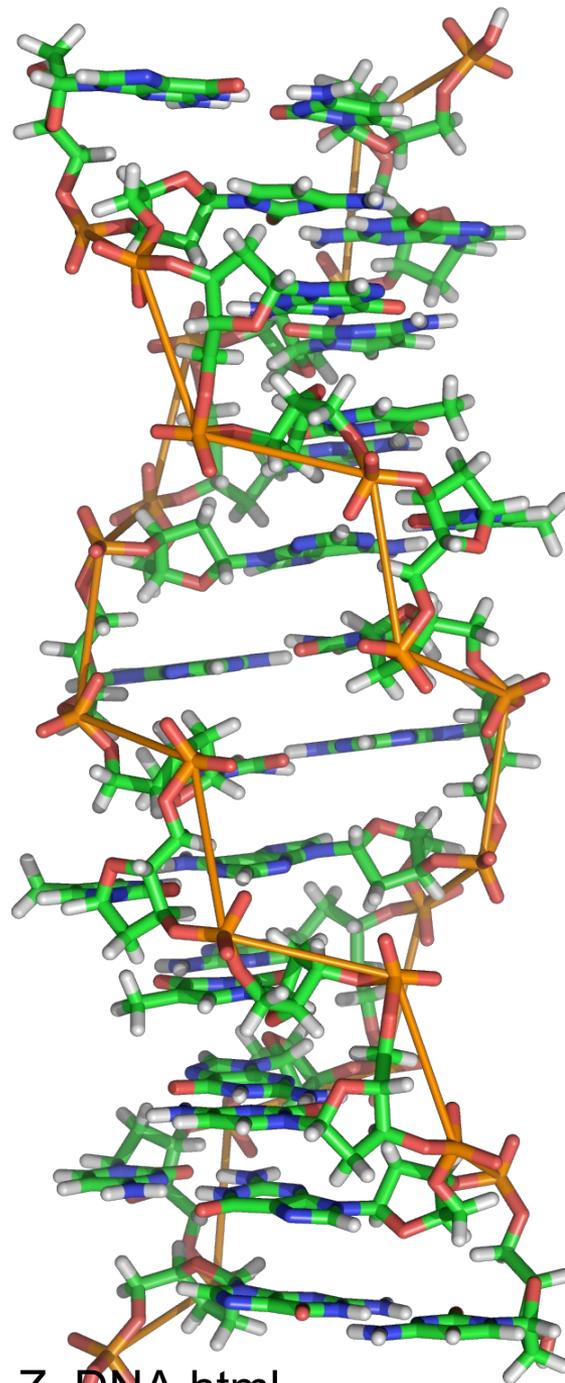
- **Left handed** Helix

- **jagged backbone**

- **G nucleotides switch their glycosidic bond conformations:** *anti* -> *Syn*

- **C nucleotides: No change**

PDB ID: 1DCG

PyMol: ZDNA.pse

www.mun.ca/biology/scarr/A_B_Z_DNA.html

Consecutive Nucleotides flipping in Z-DNA

Focus on one strand

Focus on two consecutive base pairs

G = syn

C = anti
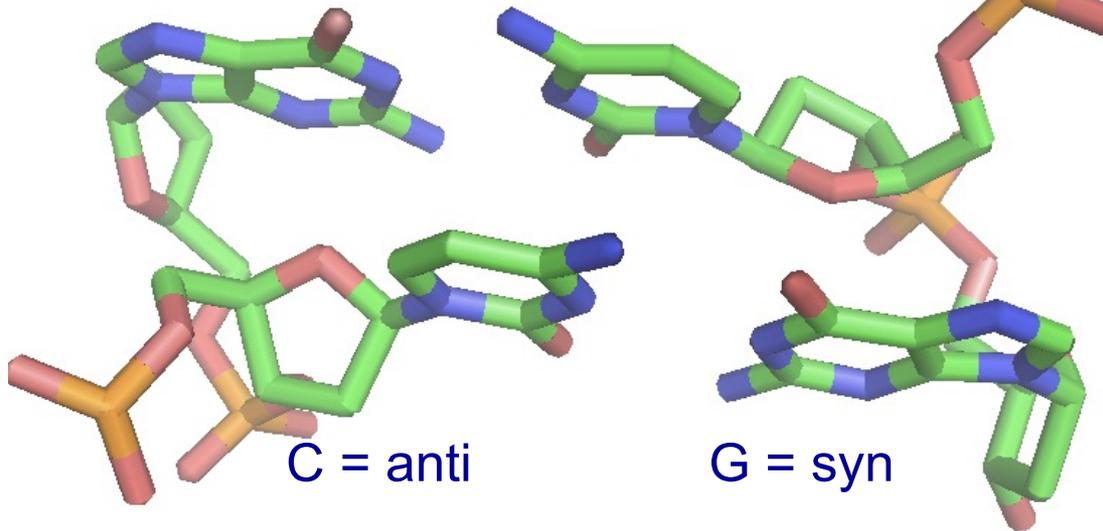
C = anti

G = syn

C = anti/

G = syn

C =anti

G =syn
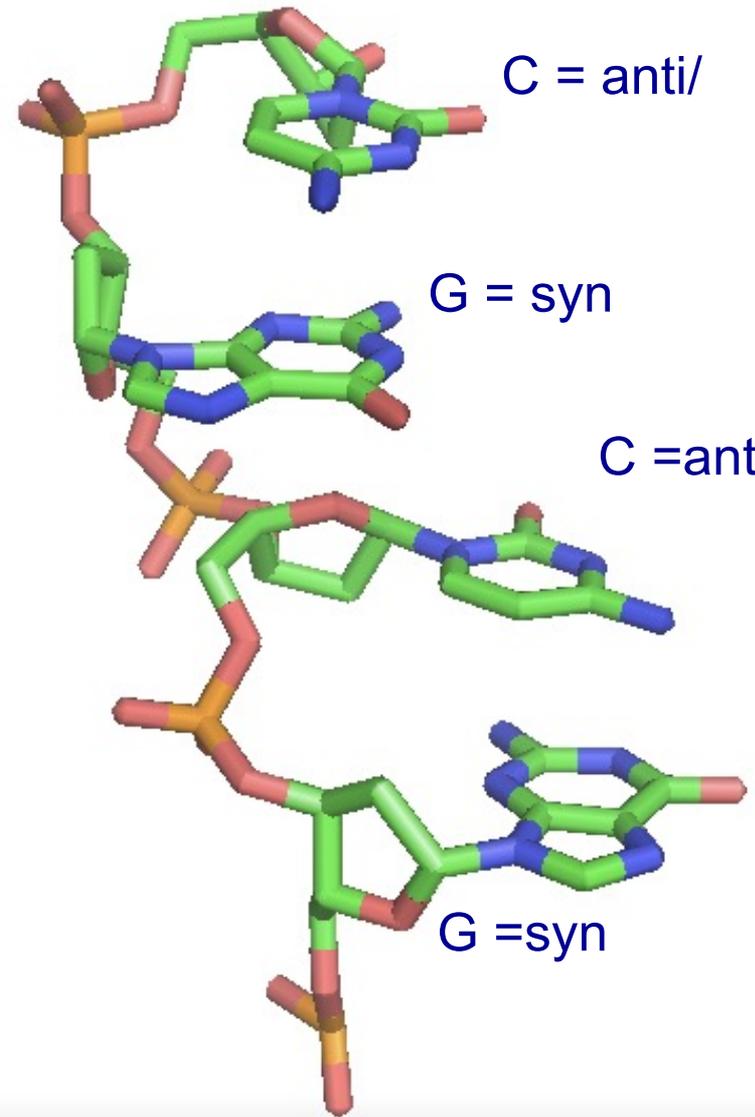
Can you recognize Glycosidic Bonds Conformations?

(Top to bottom)

A: Anti, Syn, Anti, Syn

B: Syn, Syn, Anti, Anti

C: Syn, Anti, Syn, Anti

D: Anti, Anti, Syn, Syn
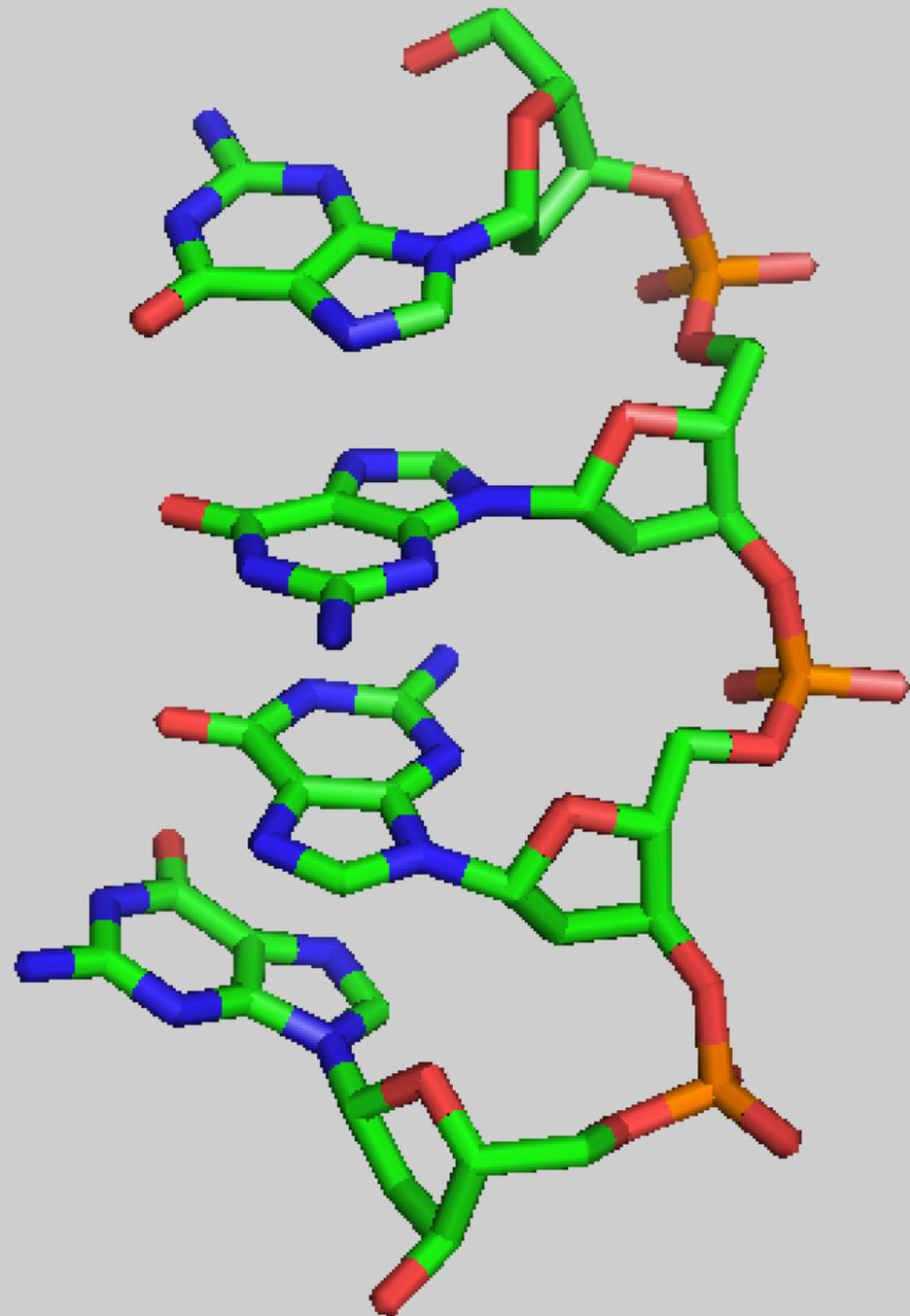
The major groove is very shallow in Z-DNA:

B-DNA

Major

N7

N3

Minor

Z-DNA

Major

N7

N3

Minor
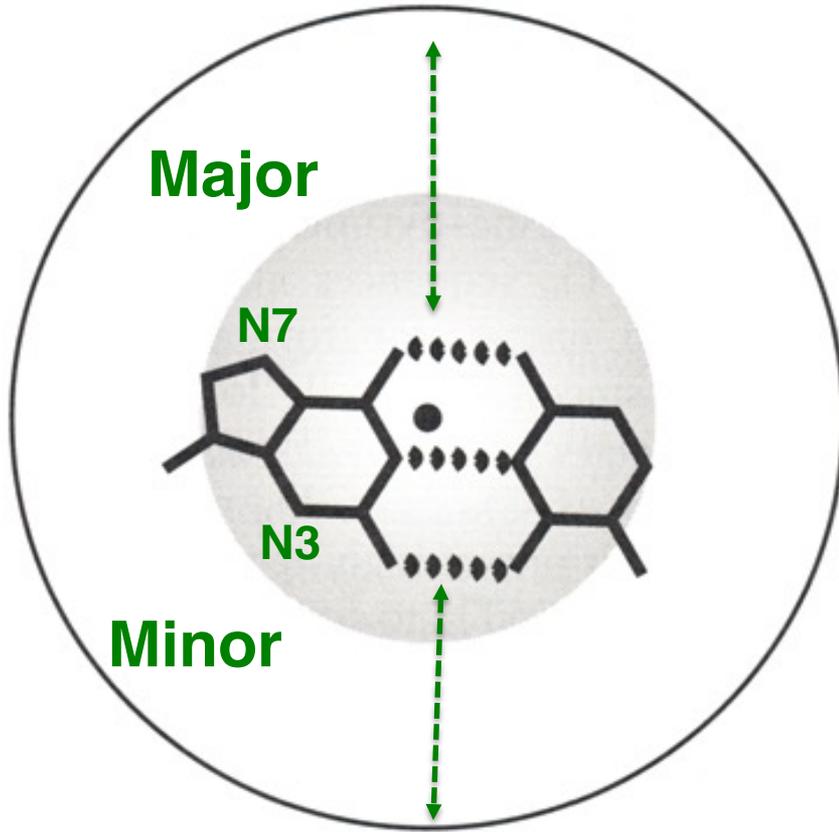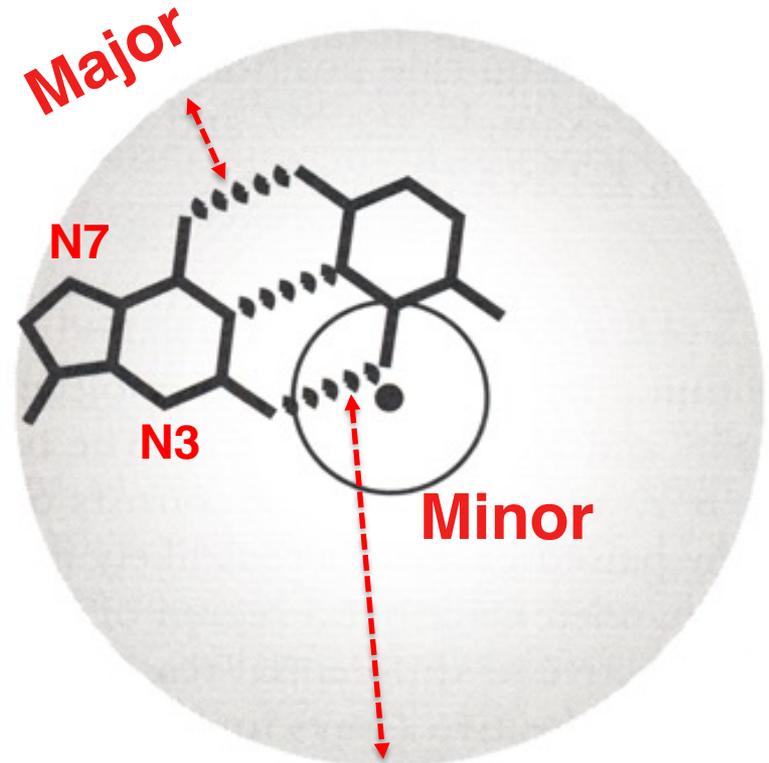
Summary of ABZ structures/ conformations

**!! There are Helical Conformations other than A/B/Z
e.g.: conformations intermediate between A and B**

www.mun.ca/biology/scarr/A_B_Z_DNA.html
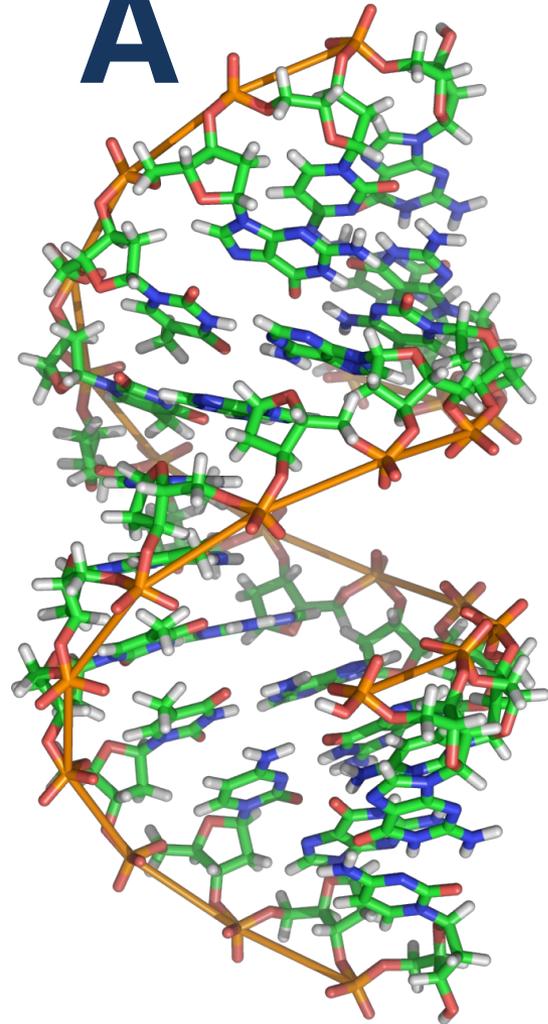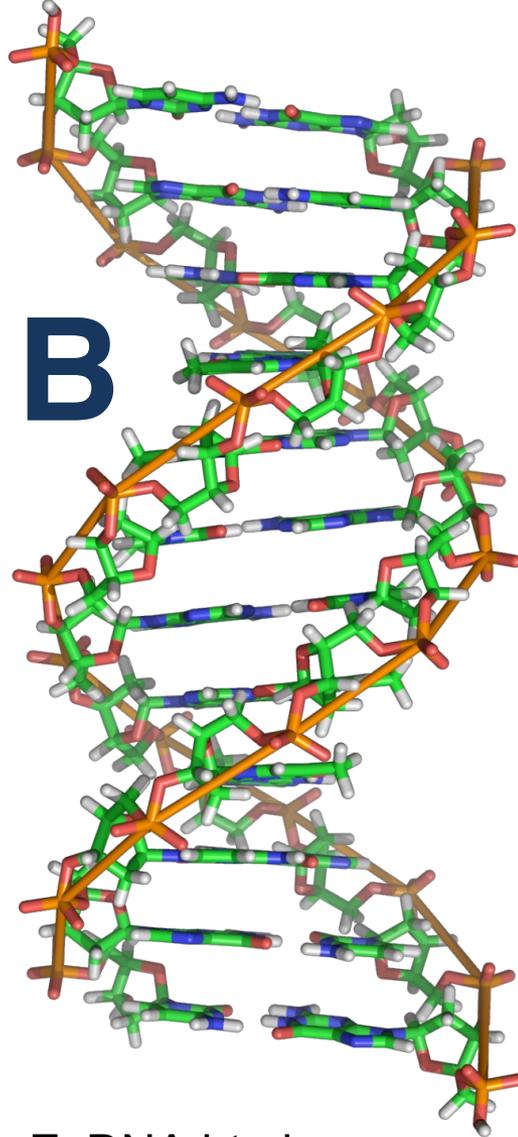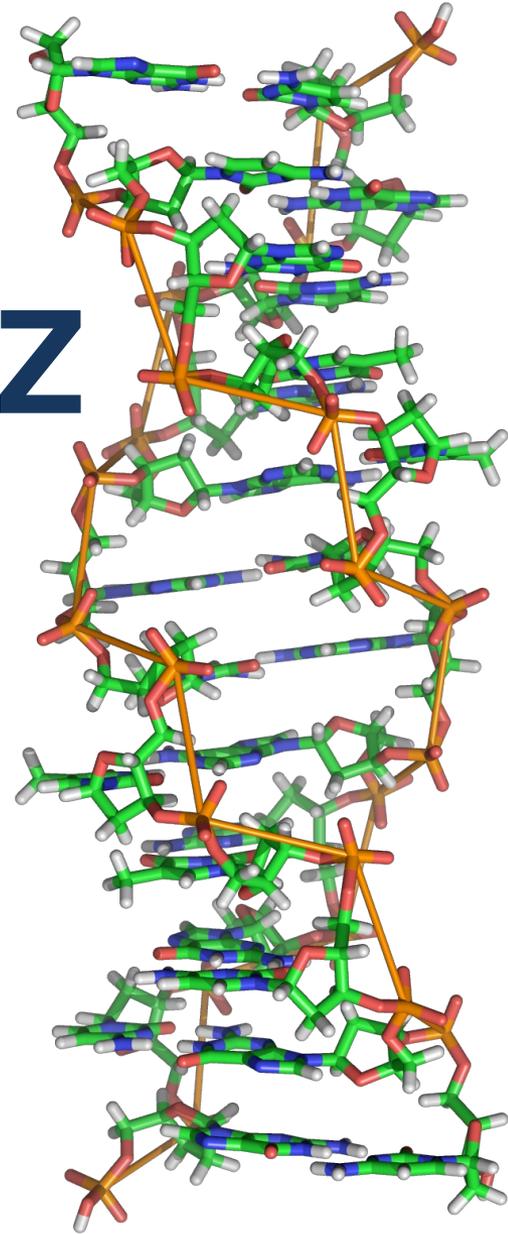
Summary of ABZ structures/ conformations

Backbone Profiles

Helical Projections

A   B   Z
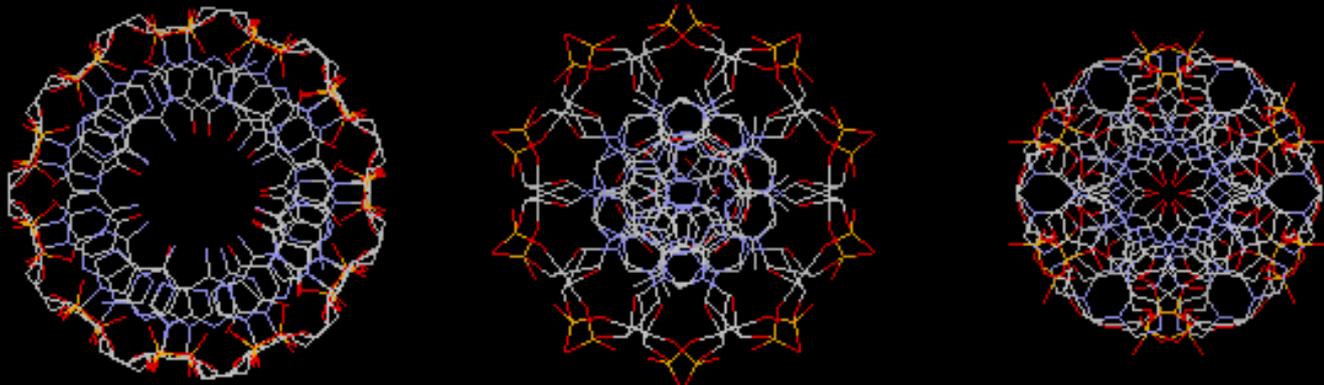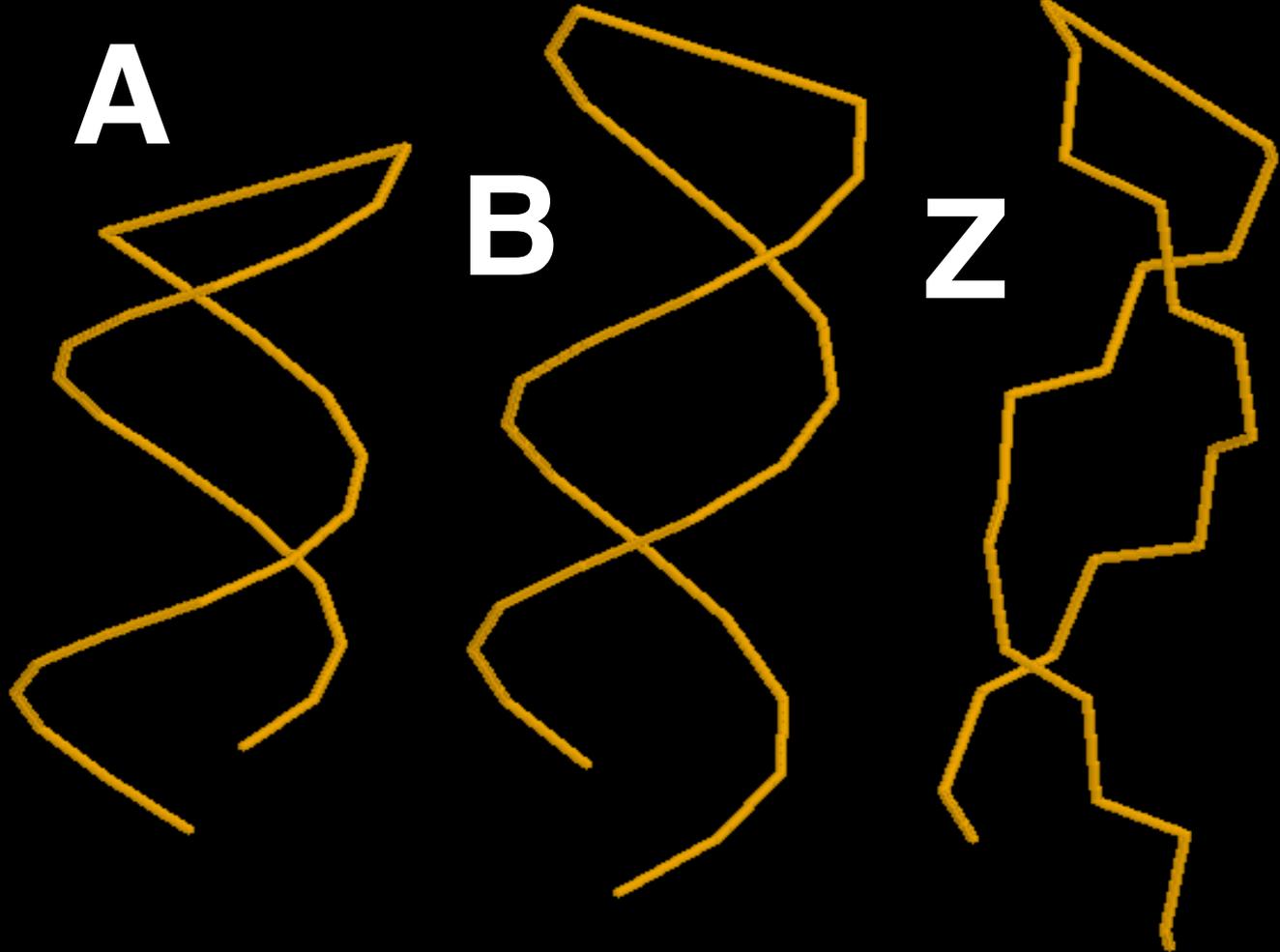
What natural nucleic acid form does this cartoon represent?

A: A-form

B: B-form

C: Z-form

D: None of the above
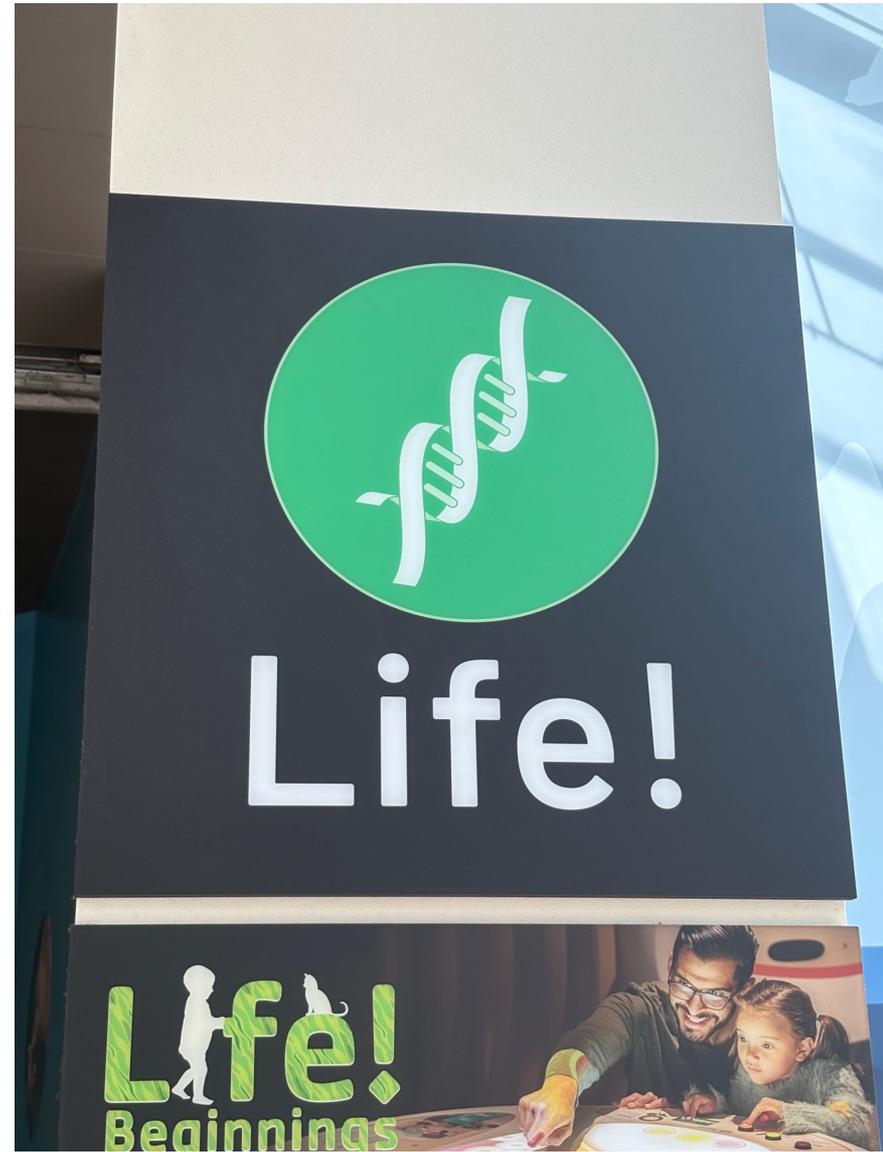
California Science Center, Expo Park

Double-stranded helices can be distorted and form bent/kinked structures:

Bend
(= smooth curve)
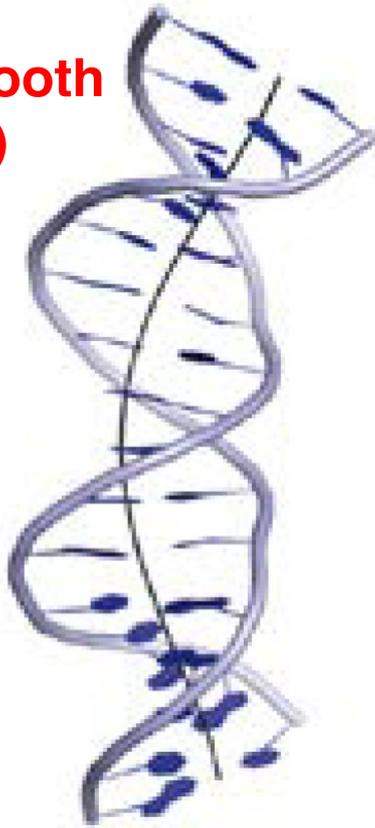
Kink
= sharp change in helical axis

This can be due to:

• Intrinsic DNA sequences/ features that increase the roll parameter, resulting in bend or kinks

• Binding of a protein that forces/favors a bent/kinked conformation

Bend

Kink

Harteis & Schneider
Int. J. Mol. Sci. 2014

Why study nucleic acid structure?

• Structure and Sequence-specific Recognition by DNA binding proteins // see next

• Some non B-DNA structures are biologically relevant:
 - dehydrated living forms
 - dsRNA is A form (see PDB: 2KYD)
 - DNA/RNA duplex (replication, transcription) is A form

• Z-DNA might be associated with sequences that regulate DNA transcription

• Z-DNA is found in bacterial biofilms (Buzzo et al. Cell 2021)

Conversion of B-DNA to Z-DNA

B-DNA = DNase sensitive

Z-DNA = DNase resistant

pre-biofilm
early biofilm
biofilm
mature biofilm

Bacteria   DNABII   Z-DNA   B-DNA   DNase
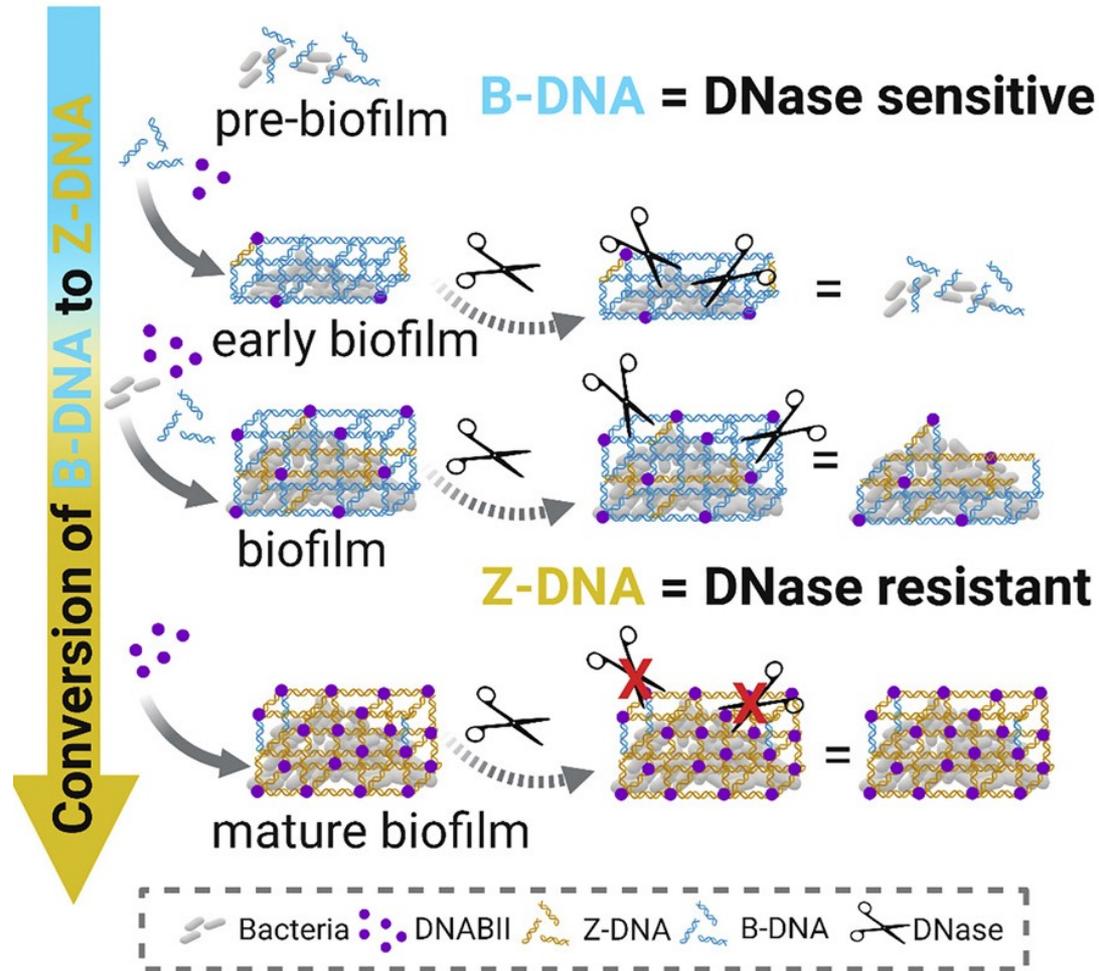
Can you read ?
(the sequence of this DNA)
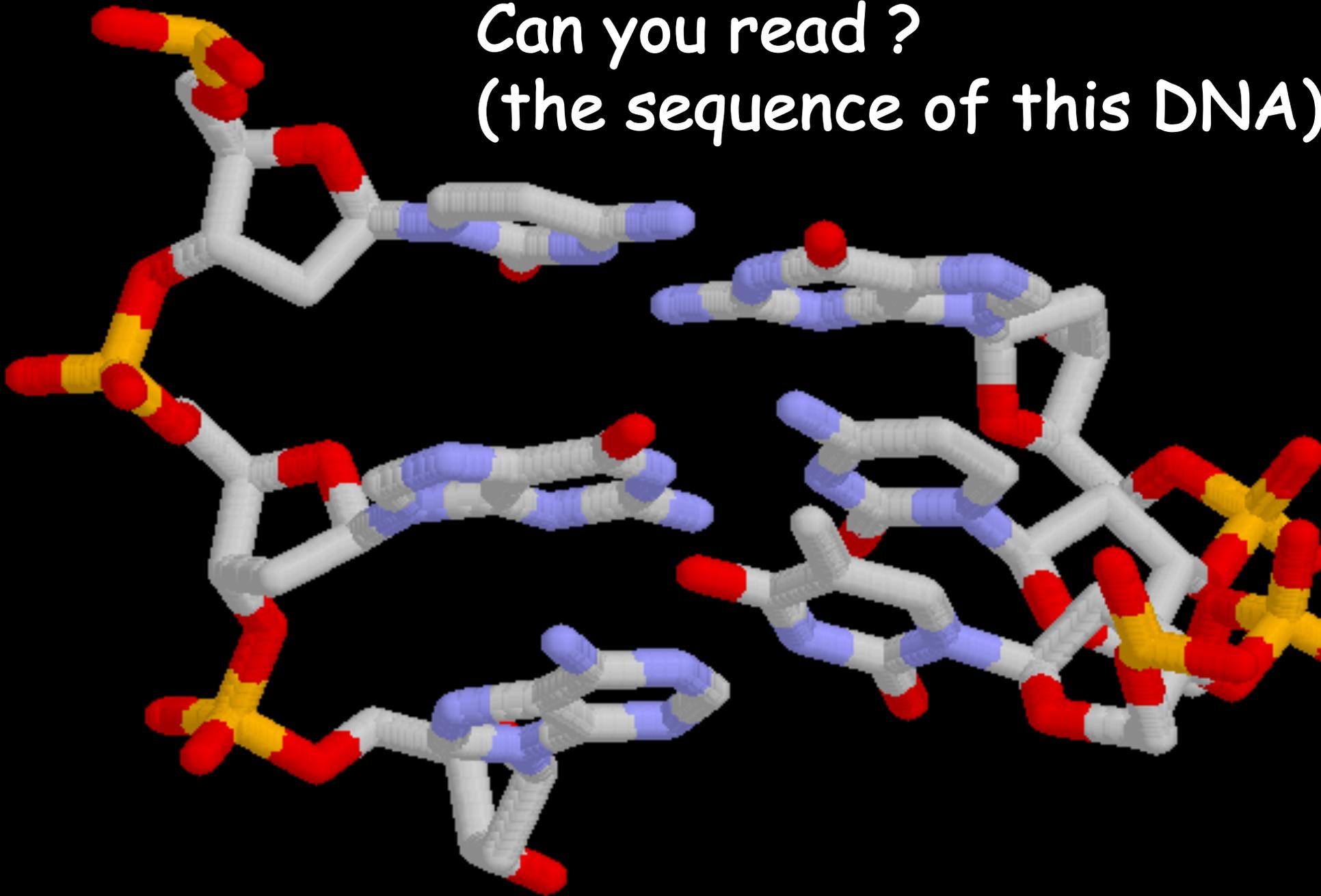
Can you read the sequence of this DNA Strand?

Top to bottom:

A: 5´TGA3´

B: 3´CGA5´

C: 5´CGA3´

D: 5´CAG3´
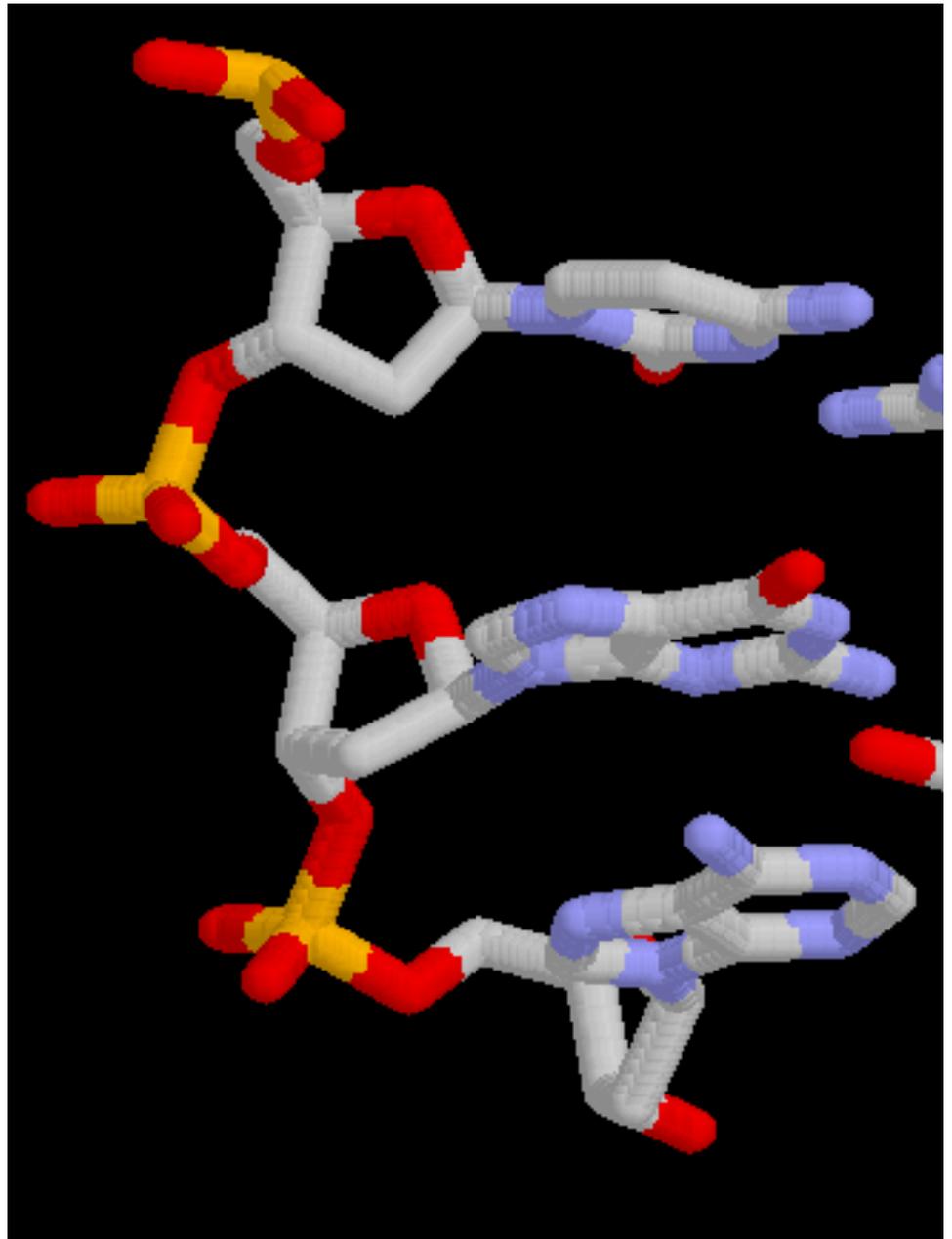
Sequence-Specific Recognition of DNA by proteins: What do proteins "see" ?
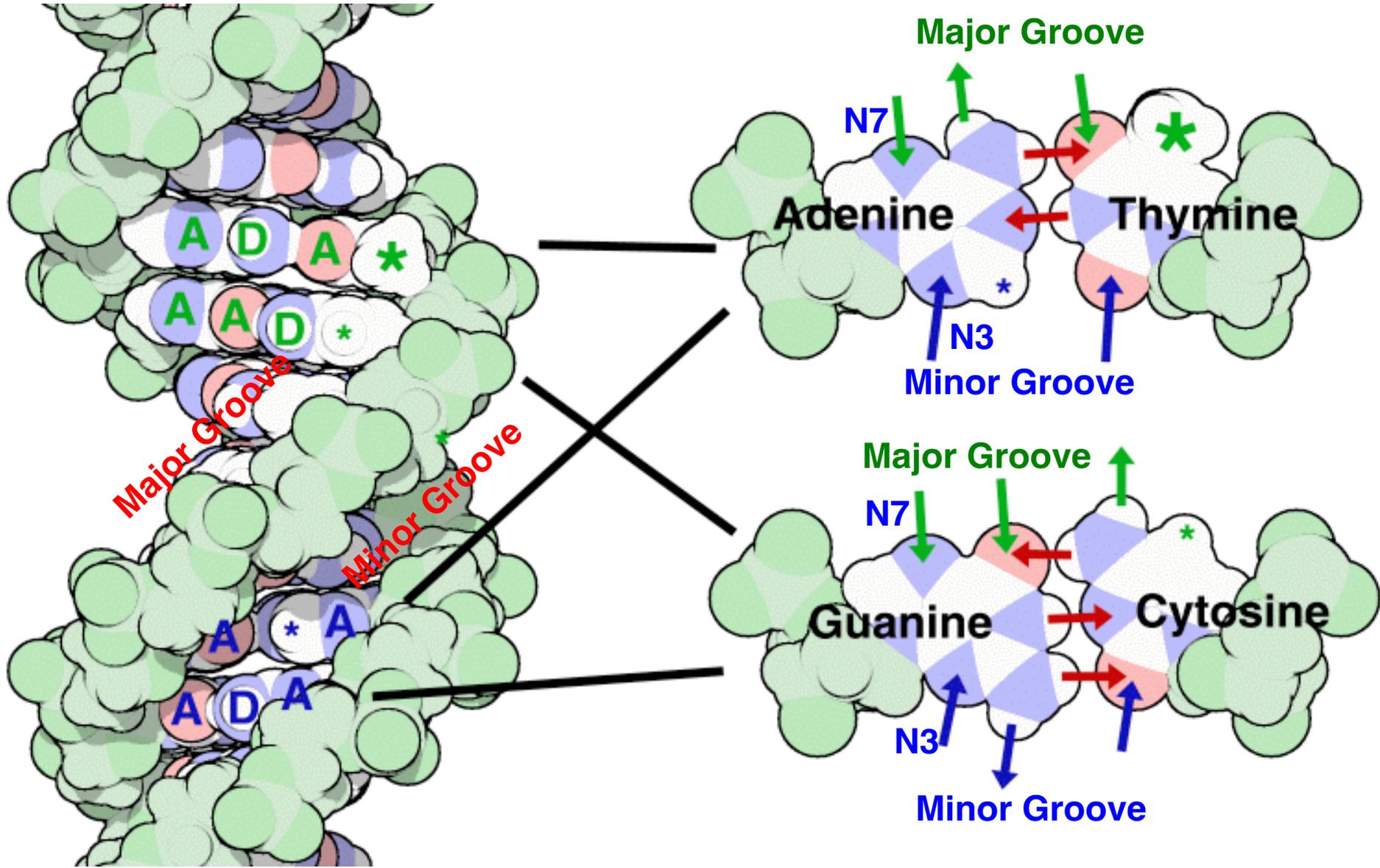
Recognition of Specific sequences by DNA-binding proteins

Distribution of H-bonds Donors (D) Acceptors (A) and Hydrophobic groups (H) available for recognition
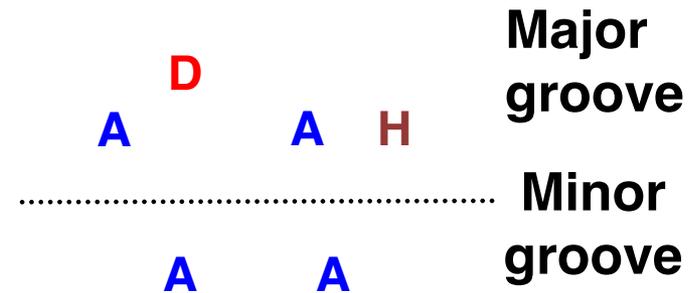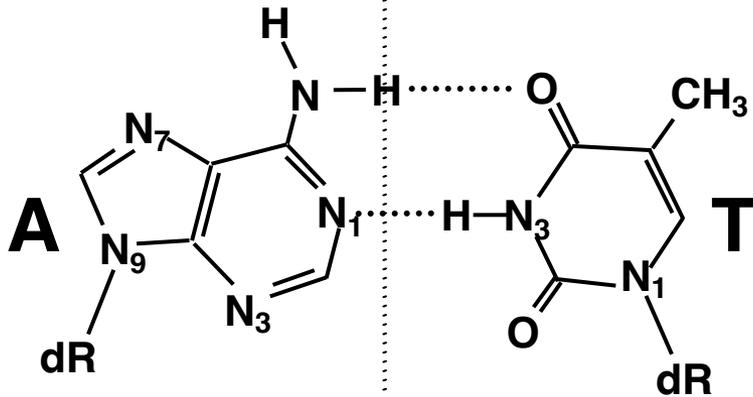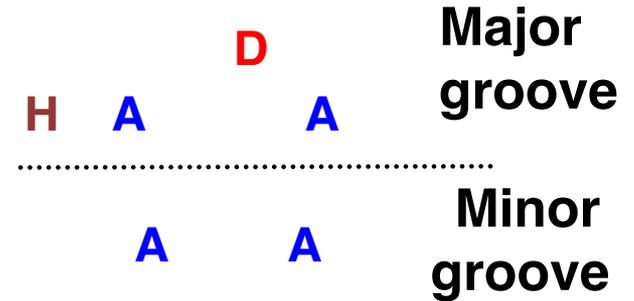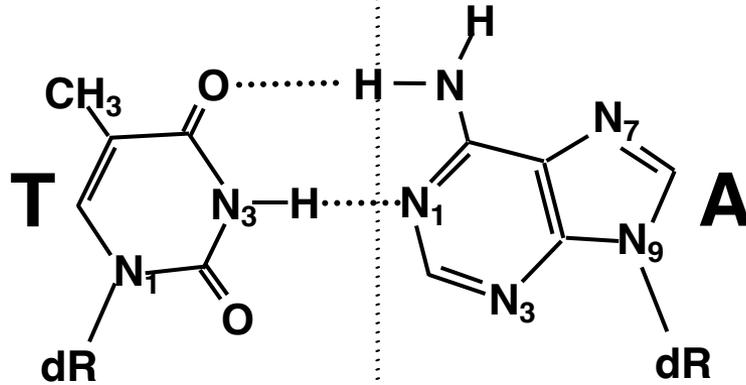
Conclusion: *Sequence-Specific* DNA binding proteins can differentiate A-T base pairs from T-A base pairs if they bind from the major groove side, but not from the minor groove side

Concept is important to remember to understand transcription regulation

Recognition of Specific sequences by DNA-binding proteins

Patterns of H-bonds Donors (D), Acceptors (A), and Hydrophobic groups (H) available for recognition
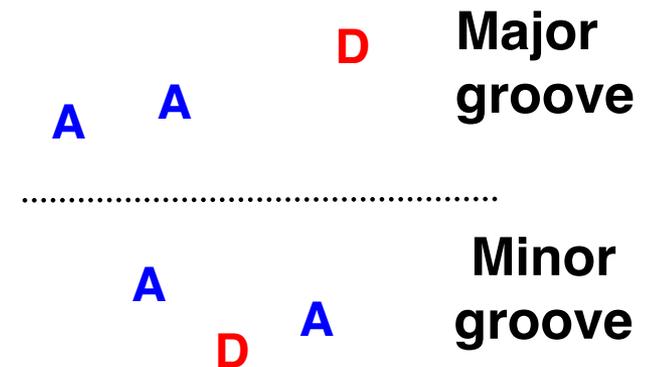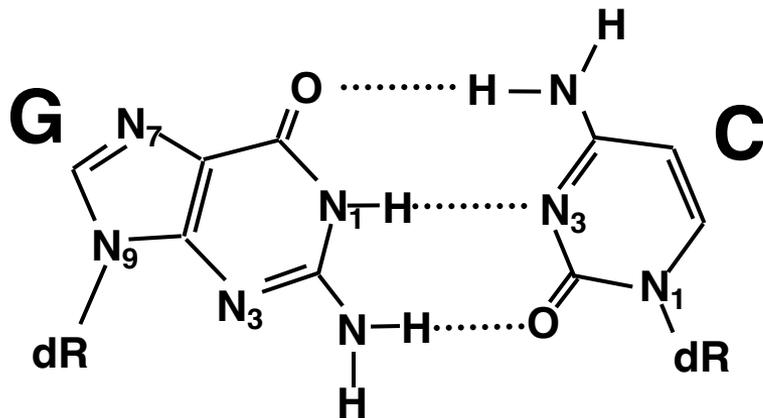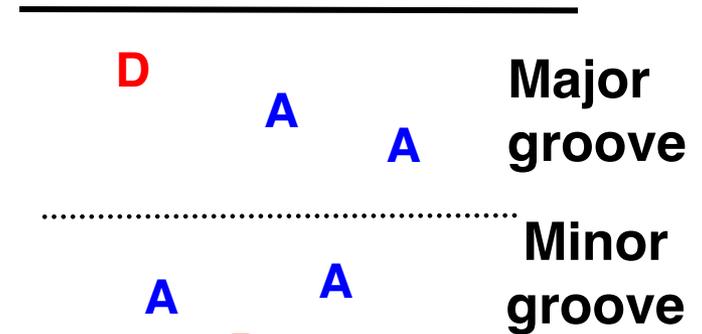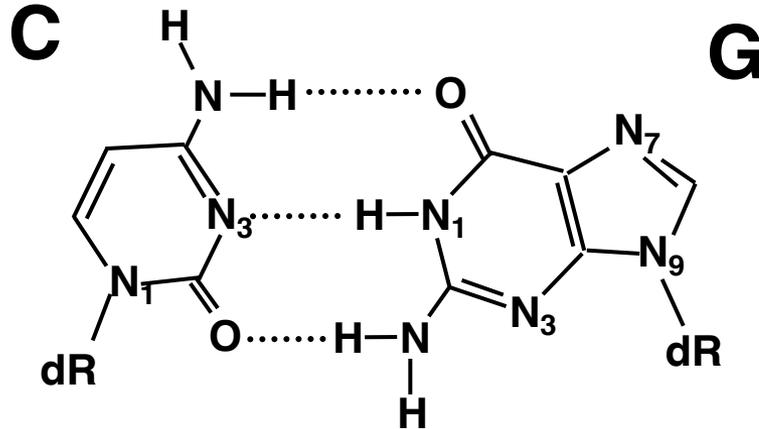
Conclusion: *Sequence-Specific* **DNA binding proteins can differentiate G-C base pairs from C-G base pairs if they bind from the major groove side, but not from the minor groove side**

What is the H-bond donor (D)/Acceptor (A) pattern in the MAJOR GROOVE for this base pair?
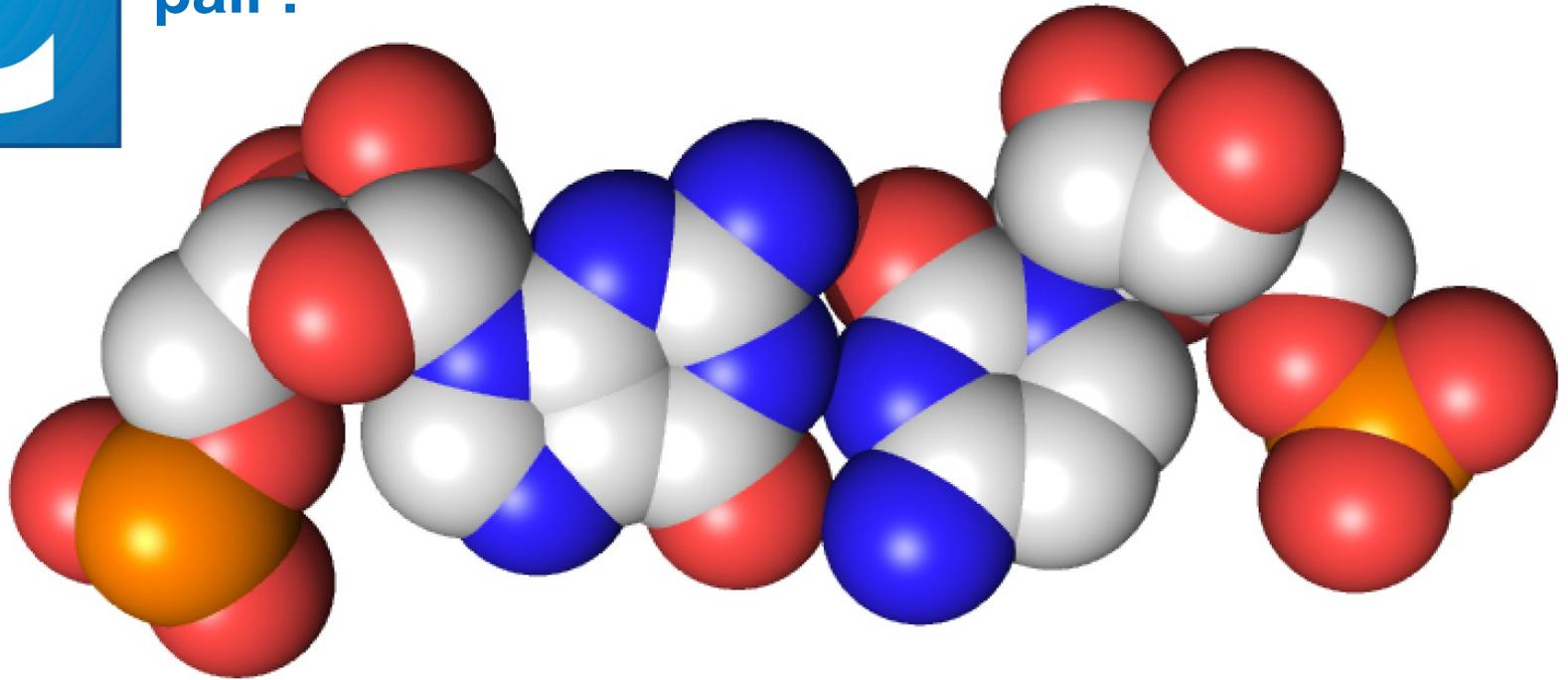
A: A D A D

B: A A D

C: D A A

D: D A A D

Examples of interactions of amino acids with the major groove side of bases or base pairs

• **These will contribute to sequence recognition and generally be found in proteins that bind DNA in a** sequence specific **manner**
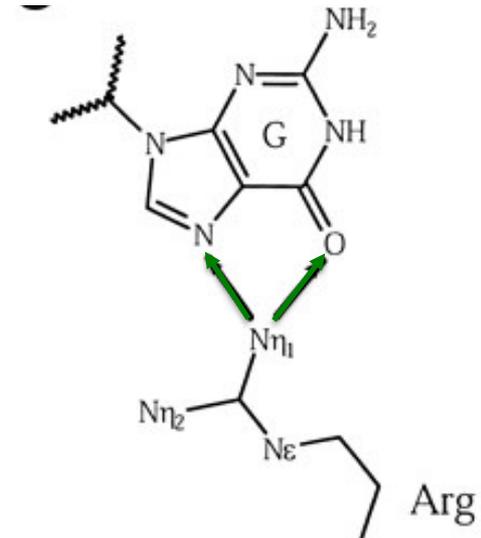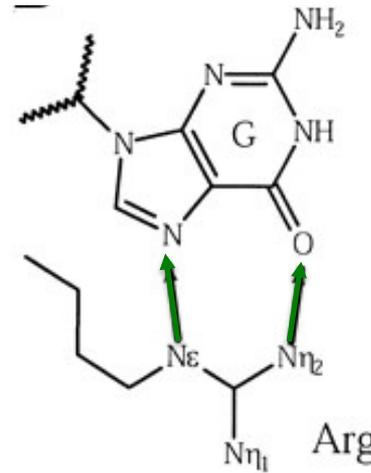
• **There is more than one way to "skin the cat"**: for instance, Arg **can recognize a G from the major groove side using a variety of geometries (see top interactions)**
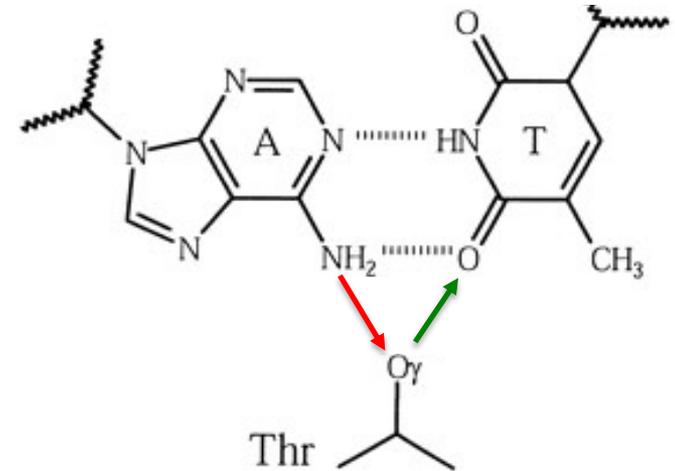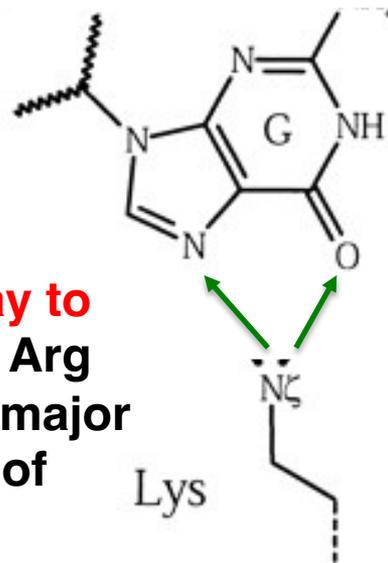
Examples of interactions of amino acids with the minor groove side of bases or base pairs

• In general, these interactions do NOT contribute to sequence recognition

• However these interactions can be made by both sequence-specific and non-sequence specific DNA binding proteins -- any sequence-specific DNA binding proteins also has some level of non sequence-specific binding sites that contribute to general DNA binding). These interactions will just not contribute to sequence specificity
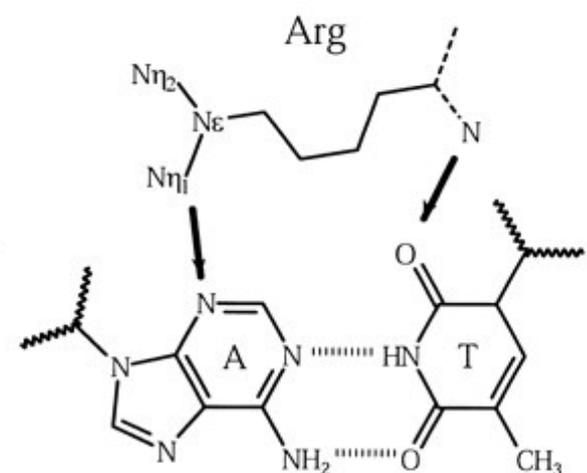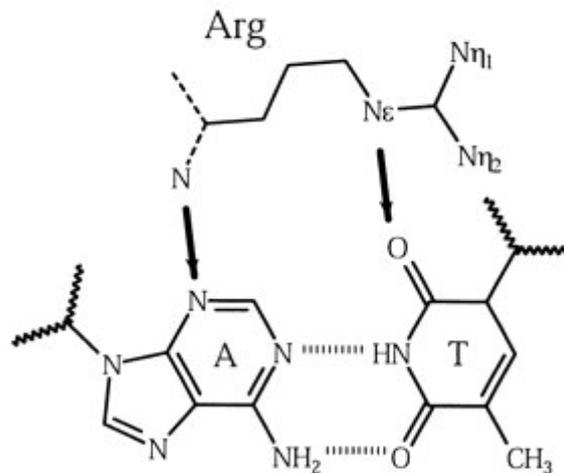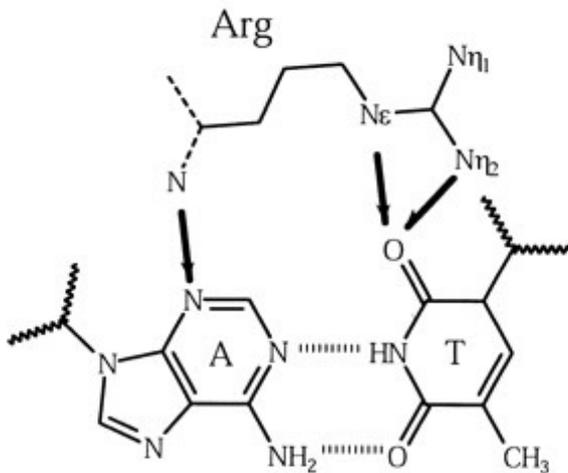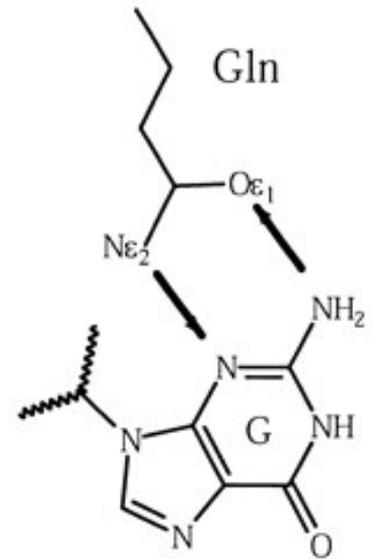
Amino acids side groups and atoms color coded by chemical make up: Hydrogen bond donors or acceptors

Asn   Gln   Cys   His   Ser   Thr   Trp   Tyr
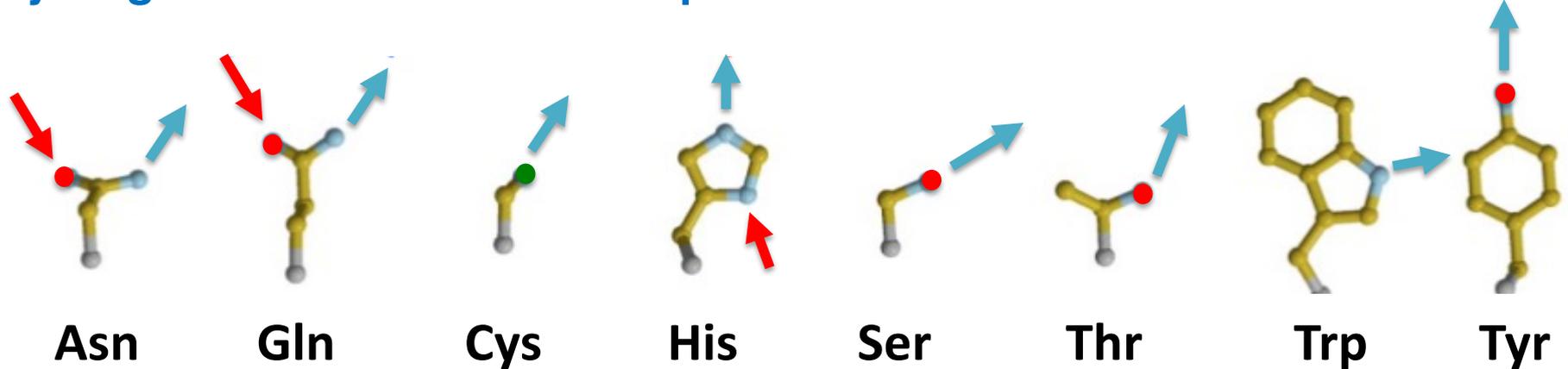
Also: Arg, Lys, Asp, Glu – see 2 slides from now
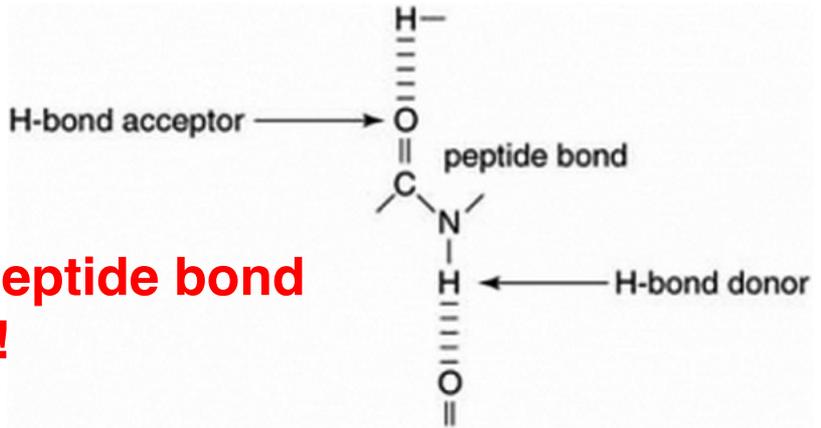
Also: peptide bond atoms!

- **Will establish hydrogen bonds with bases donors or acceptors**

- **depending on the side of the base (major or minor groove), these will contribute to sequence specific recognition… or not.**

For DNA binding proteins:

In what domains of proteins are the main chain (or peptide bond) N-H and C=O groups more likely to be involved in hydrogen bonding with the bound DNA?

A: α-helices

B: β–sheets

C: In unstructured linkers that connect α-helices and/or β-sheets or in loops

D: All the above

DNA

H-bond acceptor ⟶ O

peptide bond

H-bond donor

DNA
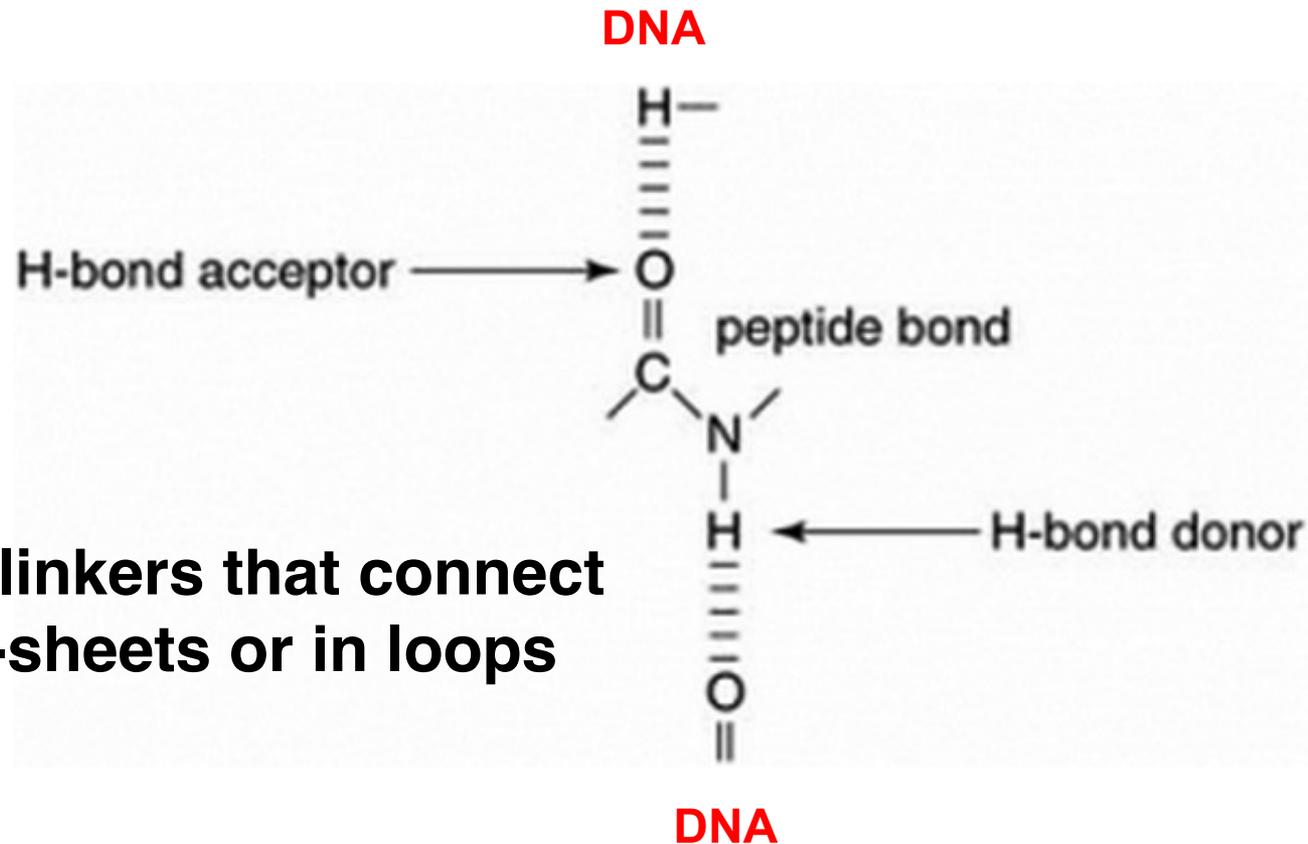
Amino acids side groups and atoms color coded by chemical make up: positively and negatively charged side chains
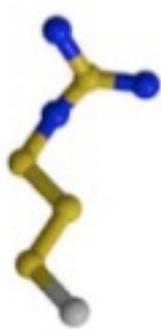
Arg Lys   Asp Glu

Positively Charged:
Will interact electrostatically with negatively charged phosphate oxygens

Negatively Charged:
May interact indirectly with the backbone through binding of positively charged ions (K^+, Mg^{++})

⚠️ **Note that these side chains can also serve as H-bonds donors/acceptors**

Amino acids side groups and atoms color coded by chemical make up: Hydrophobic side chains

Ala   Leu   Ile   Met   Phe   Pro   Val

Trp   Tyr
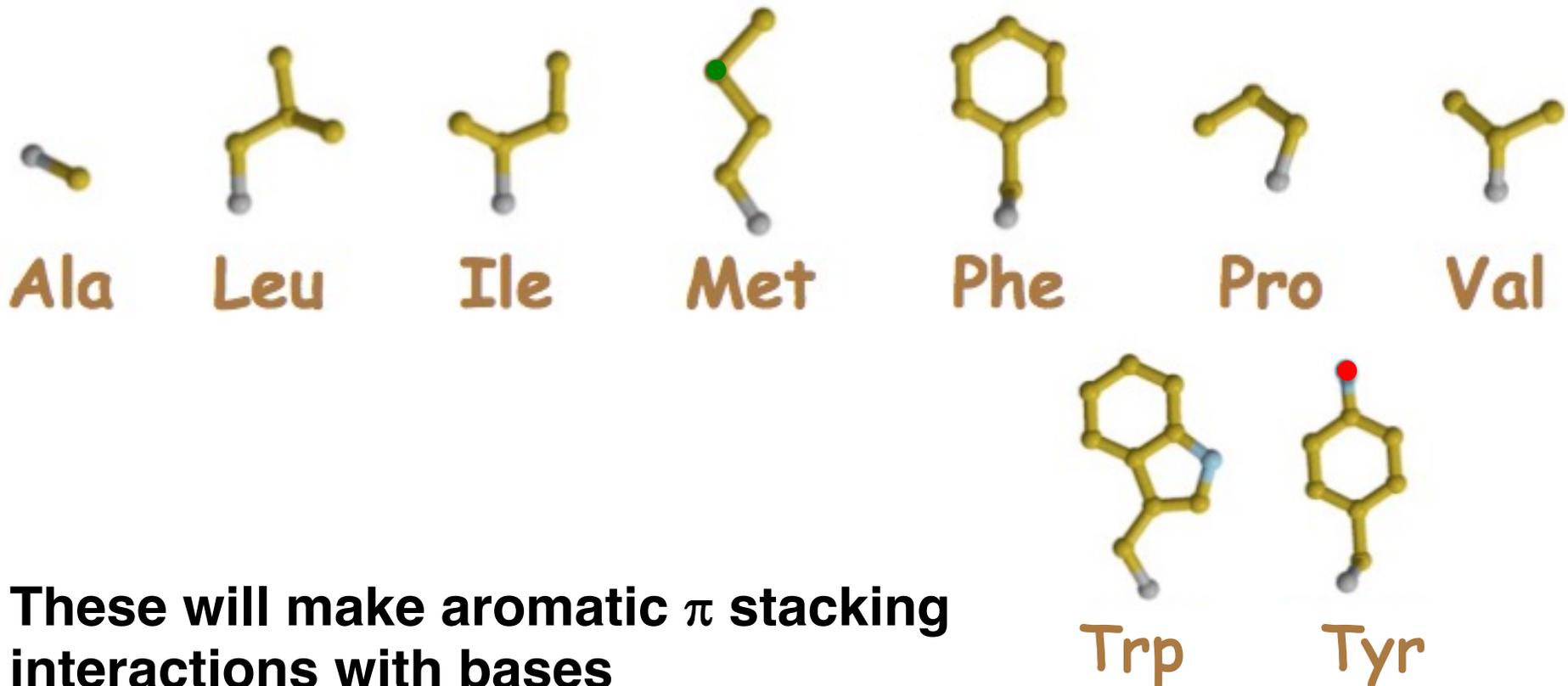
These will make aromatic π stacking interactions with bases
in general, these interactions will not contribute to sequence-specific recognition

DS DNA (Helix)

Double-helical DNA

Denaturation ⟷ Annealing

Partially denatured DNA

Separation of strands ⟷ Association of strands by base pairing

Separated strands of DNA in random coils
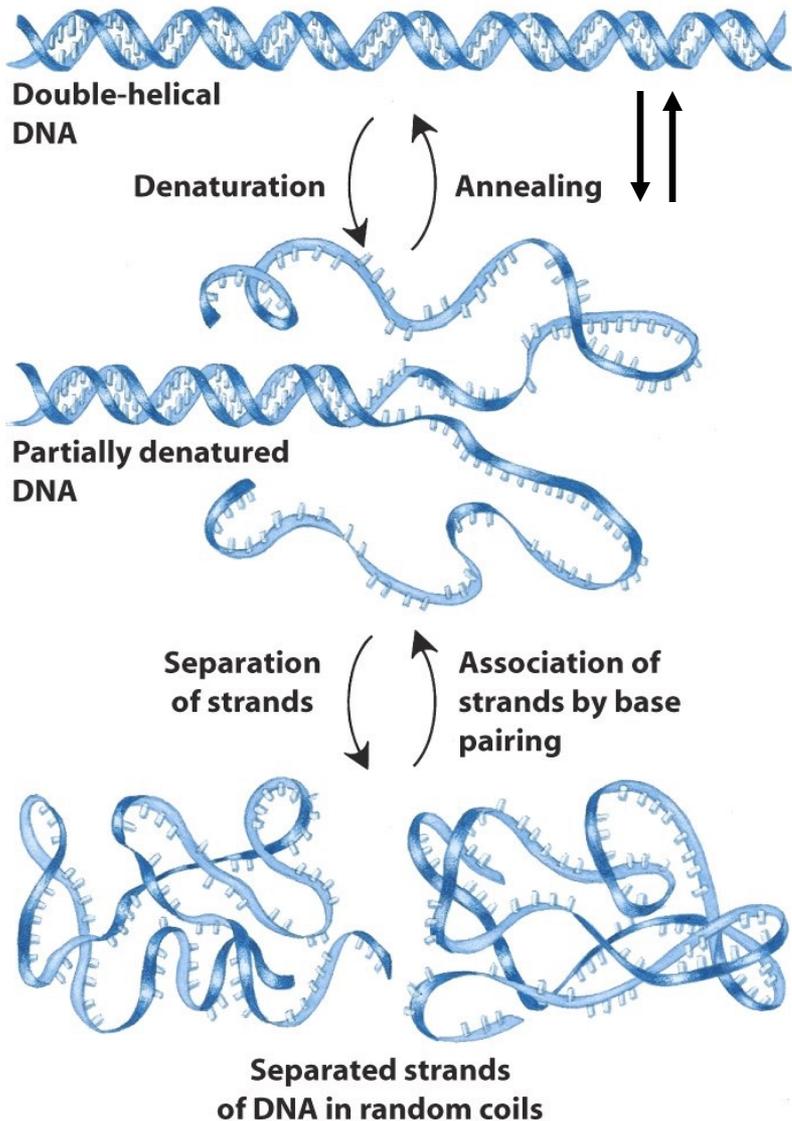
2 SS DNAs ("random coils")

What influences the equilibrium ?
(important because DNA is "opened" during replication and transcription)

In favor of double-stranded DNA

- **Base stacking interactions (major)**
  Pi stacking / dipole-dipole coupling
  Release of otherwise ordered
  water molecules
- **Hydrogen bonds between strands (minor)**

In favor of single-stranded DNA

- **Electrostatic repulsion between strands**

- **Entropic considerations:**
  Increased entropy for ssDNA vs dsDNA
  (increased conformational and translational components of entropic considerations)

Why is the contribution of H-bonds to DNA duplex stability considered minor?

A: Because the ΔG of H-bonds is negligible compared to that of hydrophobic interactions

B: Because scientists are not able to calculate their contribution precisely so they neglect their contributions

C: Because H-bonds are much weaker than covalent bonds

D: because H-bonds between 2 bases are replaced by H-bonds between bases and water during denaturation so the net energy change is close to zero

In favor of single-stranded DNA:
Electrostatic Repulsion between strands

In favor of double-stranded DNA:
Base stacking
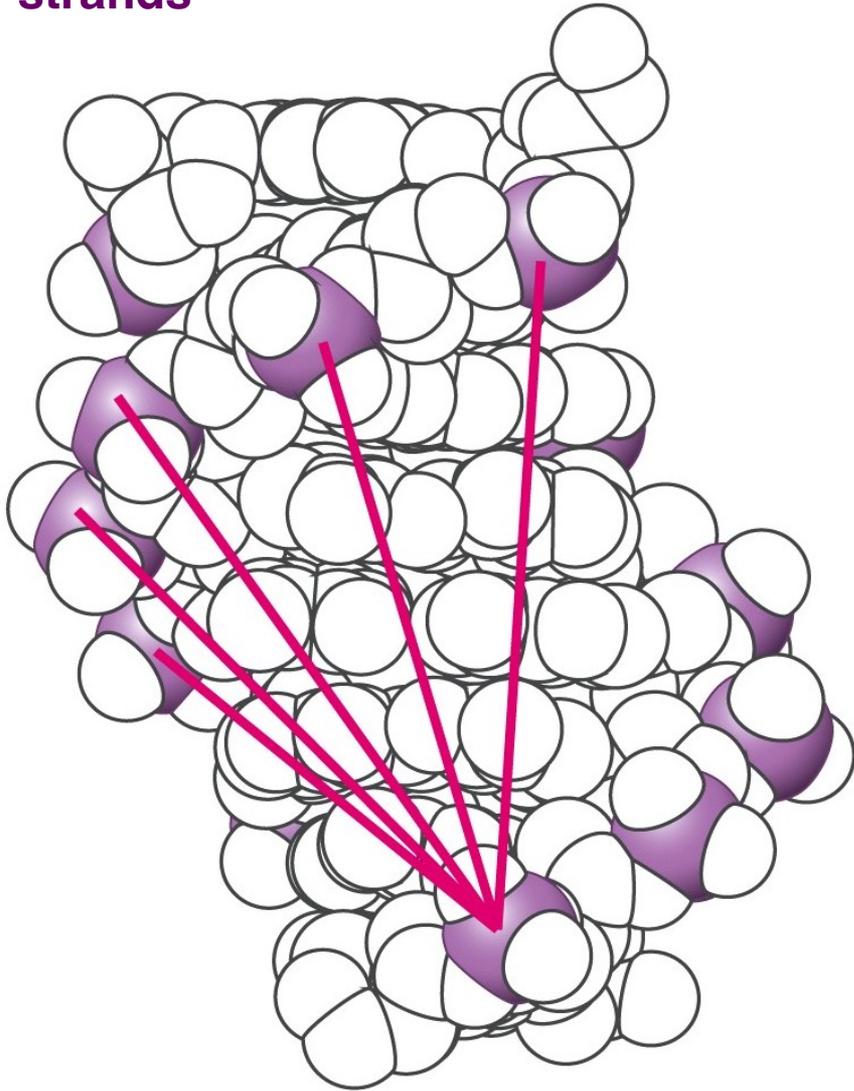
Bases Stacking

Positively charged ions (K+, Na+, Mg++) will counteract this effect
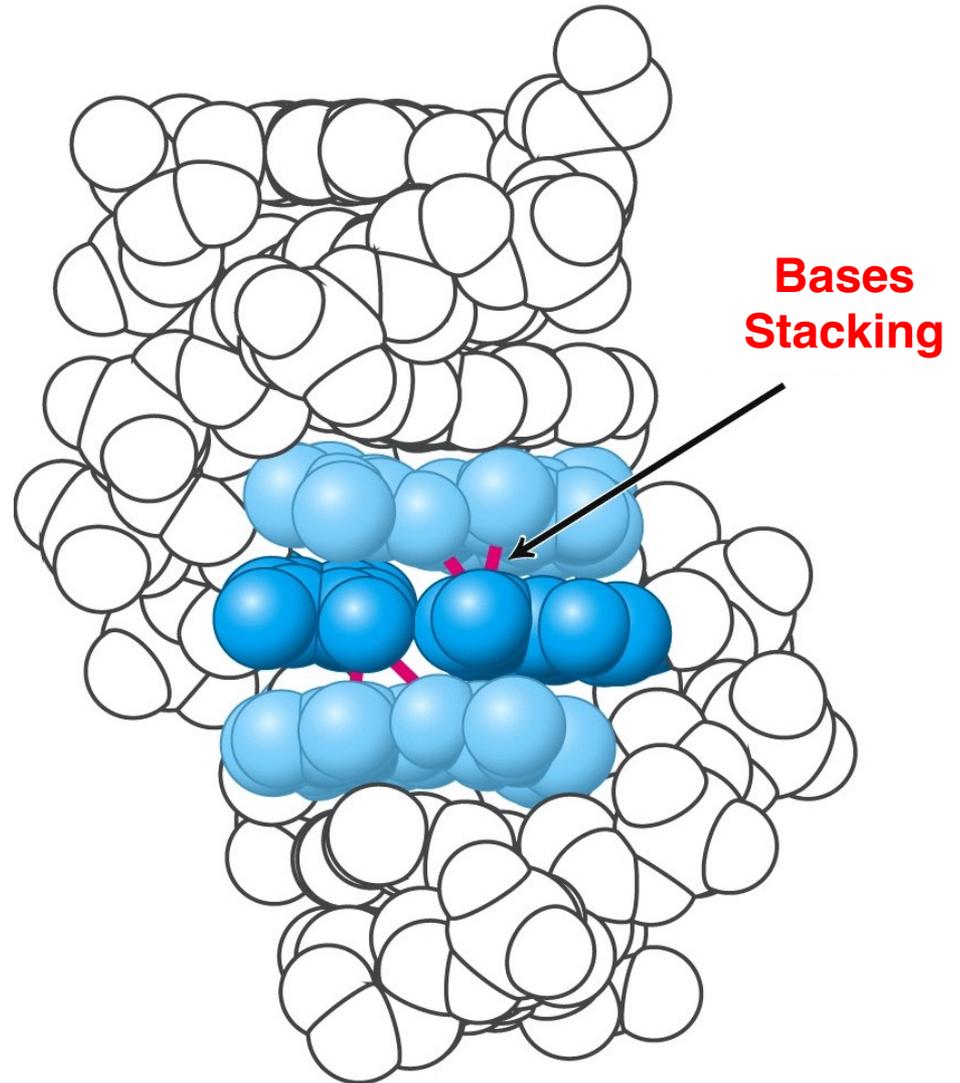
· Conformational Entropy

dsDNA:
Relatively rigid with limited flexibility
$\Rightarrow$ **low conformational entropy**

ssDNA:
More flexible with rotational movement of the strand
$\Rightarrow$ **higher conformational entropy**

· Translational Entropy

Linked to the number of possible states of the thermodynamical system

A·B		A·B
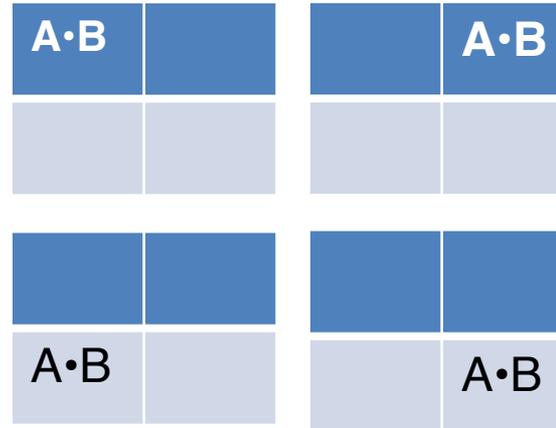
dsDNA:
strands A and B are linked
= 4 possible states

A·B		A·B

A + B		A	B
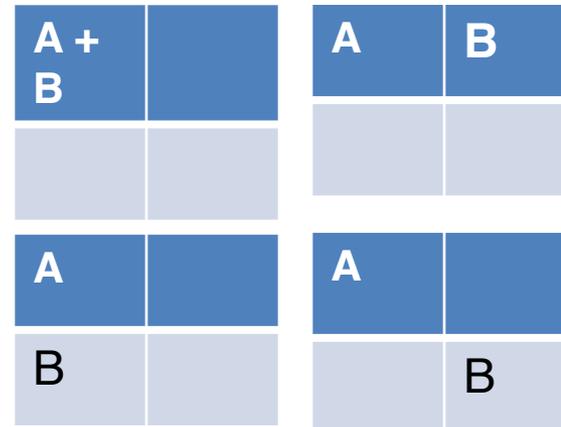
ssDNA:
strands A and B are free from each other
= 16 possible States

A		A	
B			B

(repeat 3x with A in other positions)

Higher Translational Entropy

DNA denaturation can be measured using UV absorbance measurements

Relative Absorbance

UV spectroscopic analysis of SS (denatured) vs DS (native) DNA

Denatured (ss) DNA

Native (ds) DNA

180   200   220   240   260   280

Wavelength (nm)
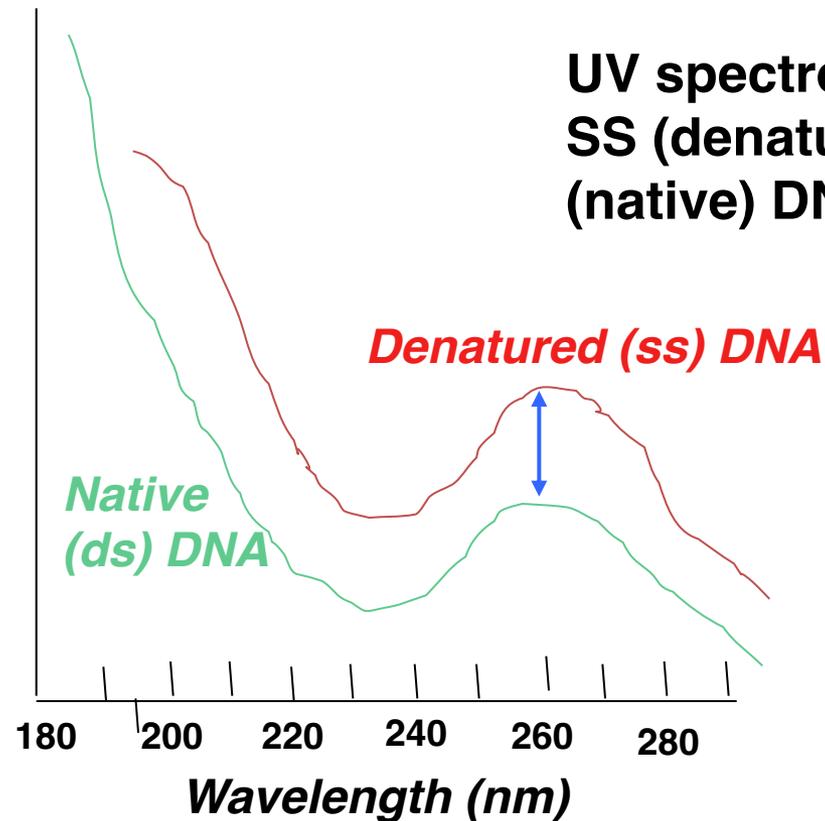
Hyperchromic Effect: UV absorbance of SS DNA > that of native DNA

Why? In DNA found in a single stranded conformation, electrons in the bases orbital systems are more excitable by UV because bases are no longer stacked regularly -> these electrons are less involved in pi-stacking interactions

Measuring DNA denaturation as a function of temperature

% of molecules in single stranded conformation

DNA Molecule 1

DNA Molecule 2

t_m

t_m

Temperature (°C)
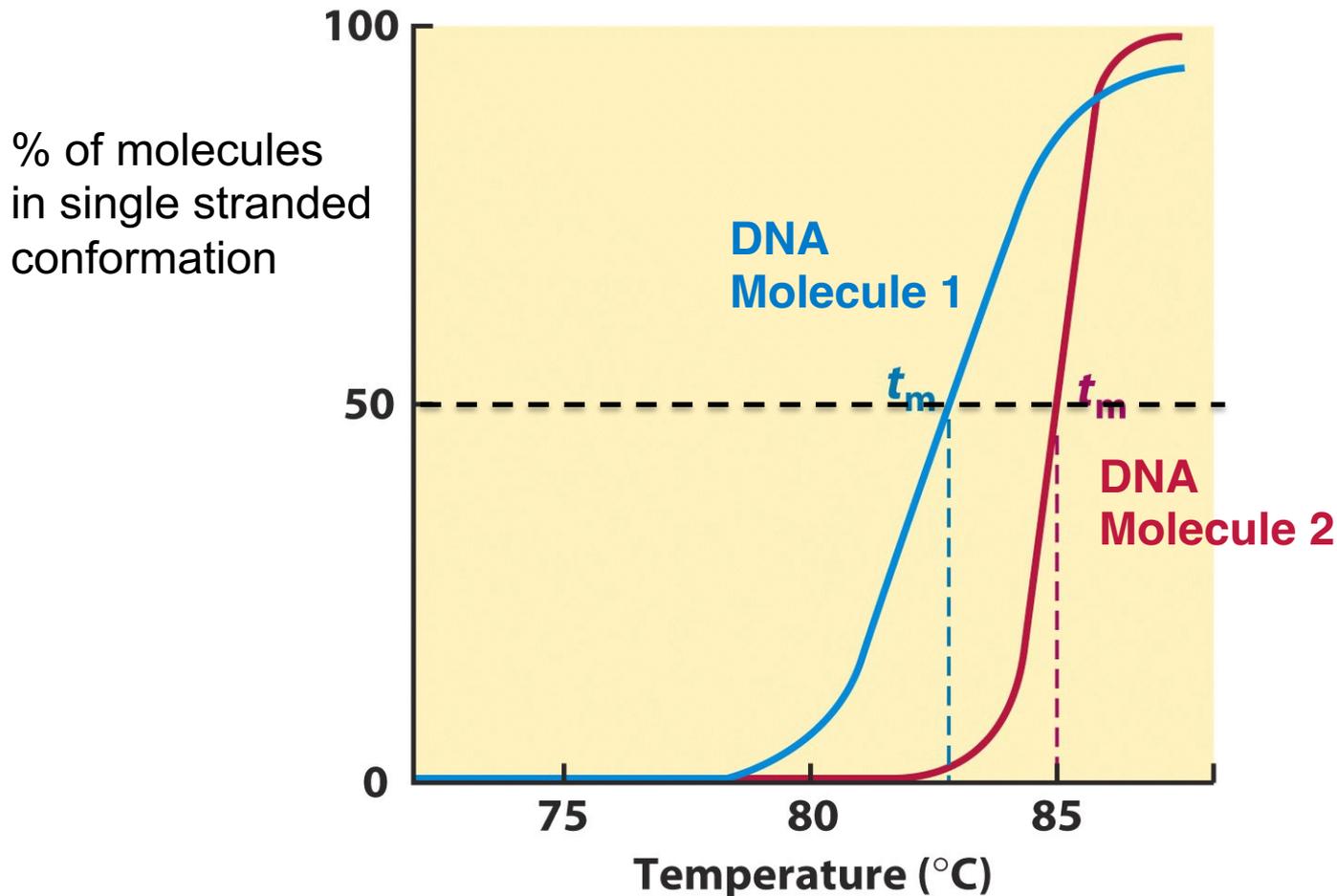
"melting curves" for two different DNA molecules (red and blue) show different "melting points" = 2 different T_m values

What do these DNA melting curves remind you of?

A: Poisson distribution

B: oxygen binding/dissociation by Hemoglobin = sigmoid curve

C: Case counts in the early days of a pandemic = exponential curve

D: Gaussian curve = expected grading for a "curved" class
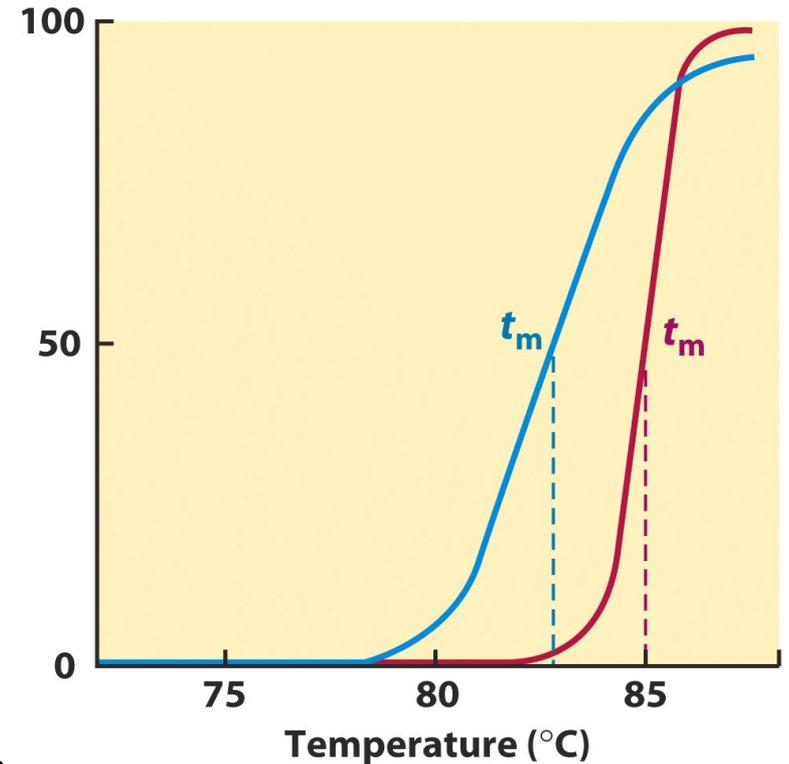
DNA melting is a cooperative process:
this explains the sigmoid shape of DNA denaturation curves

Increasing Conformational Entropy

Increasing
Translational
Entropy
(1 -> 2 molecules)

The Tm of a DNA molecule is a linear function of its G-C content but NOT because of a higher number of H-bonds: 3(GC) vs 2 (AT)

Measuring the ΔG of various consecutive base pairs:

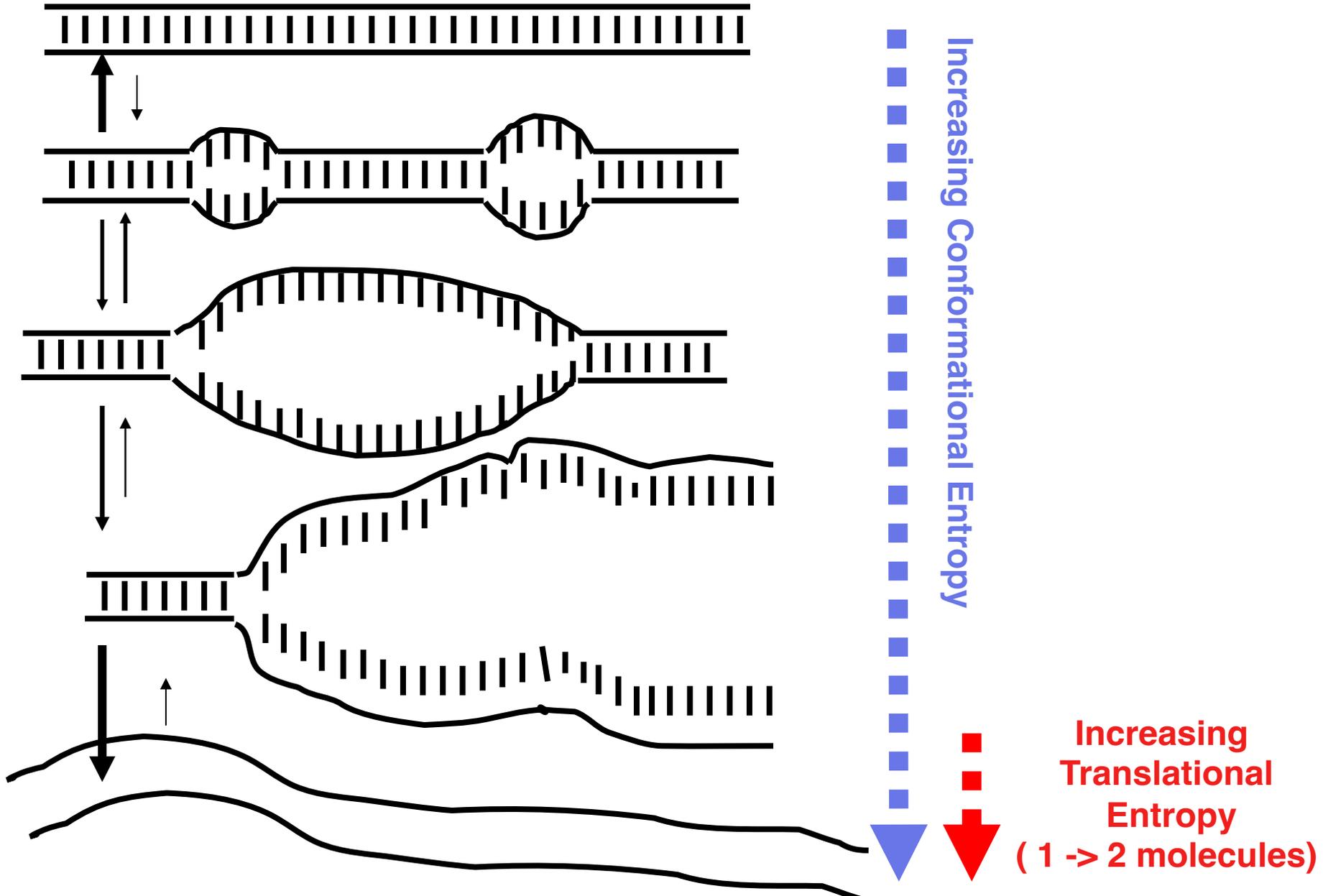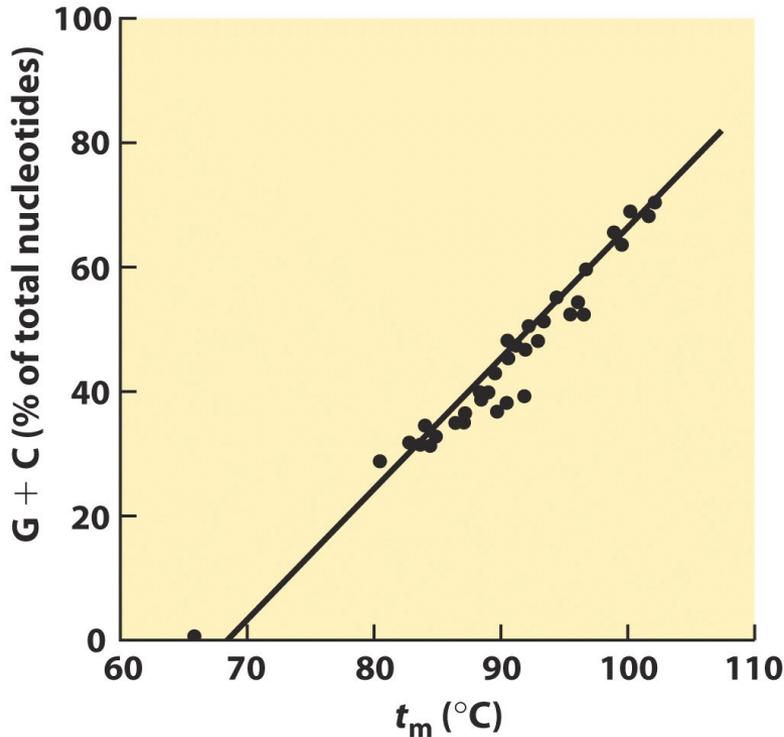|	ΔG of stacking (kcal/mol)
GC·GC	-2.17
AC·GT	-1.81
GG·CC	-1.44
GA·TC	-1.43
AT·AT	-1.34
AA·TT	-1.11
AG·CT	-1.06
CG·CG	-0.91
CA·TG	-0.55
TA·TA	-0.19

Protozanova et al. J.Mol.Biol. 2004

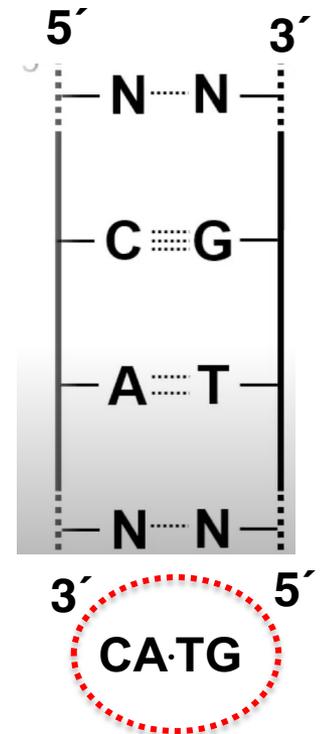How to read this table: Example of a CA•TG Stack:

Effect of G-C content on stability is Due to higher stacking energy provided by G-C base pairs in general compared to AT base pairs – especially cumulative effect of consecutive G-C base pairs

https://www.youtube.com/watch?v=PM3lIvFoLAg&ab_channel=DerekMcLachlin

The identify and orientation of stacked bases affects its favorability
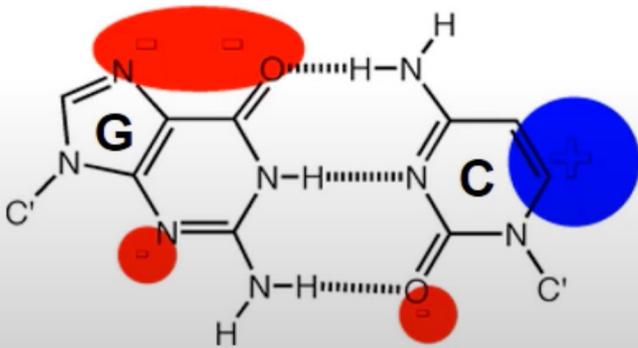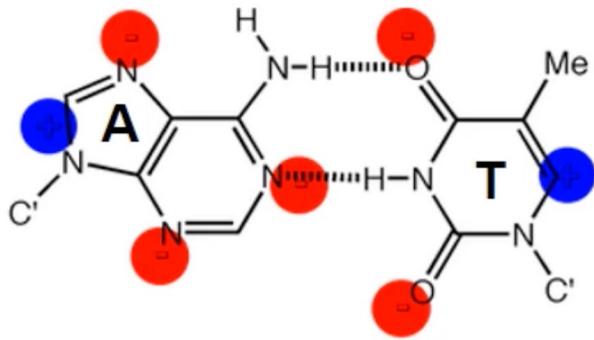
CA Hunter (1996) *BioEssays* 18:157-62

GC/GC

CG/CG
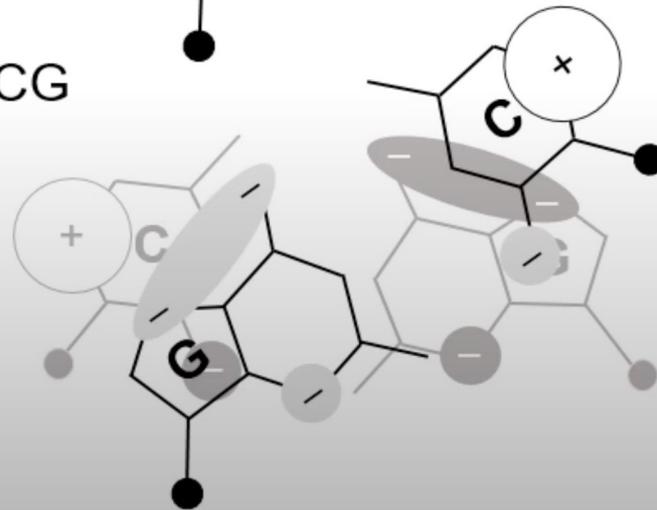
	ΔG of stacking (kcal/mol)
GC·GC	-2.17
CG·CG	-0.91

https://www.youtube.com/watch?v=PM3lIvFoLAg&ab_channel=D

Example of intrinsic DNA Tertiary Structure

G-quadruplex structures in telomeric DNA: case of (T2G4) repeats in telomeric DNA
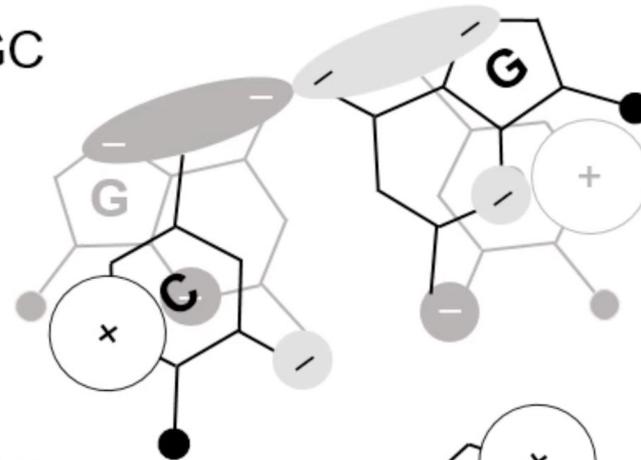
4 G-quartets stack onto each other

K$^+$/Na$^+$

G-quartet = 4 G bases interact with each other
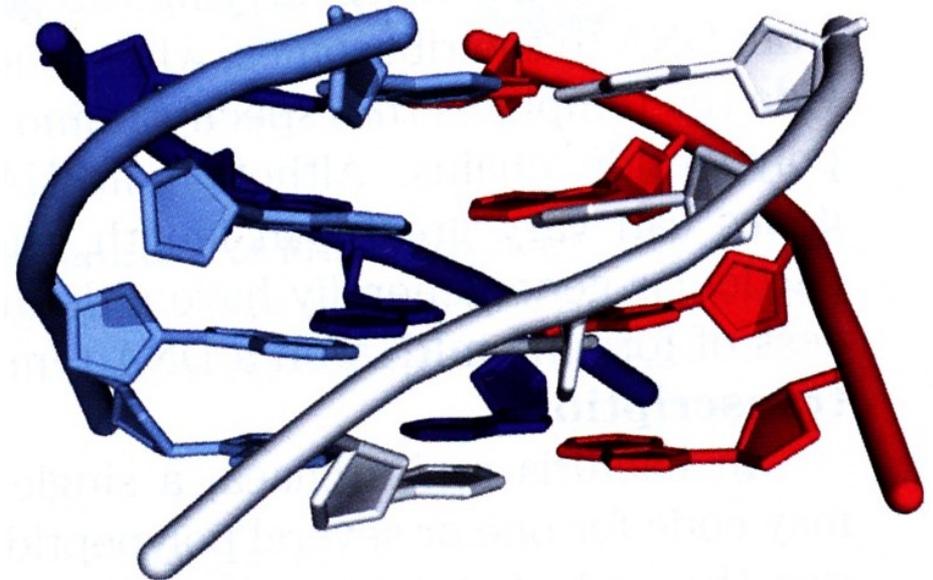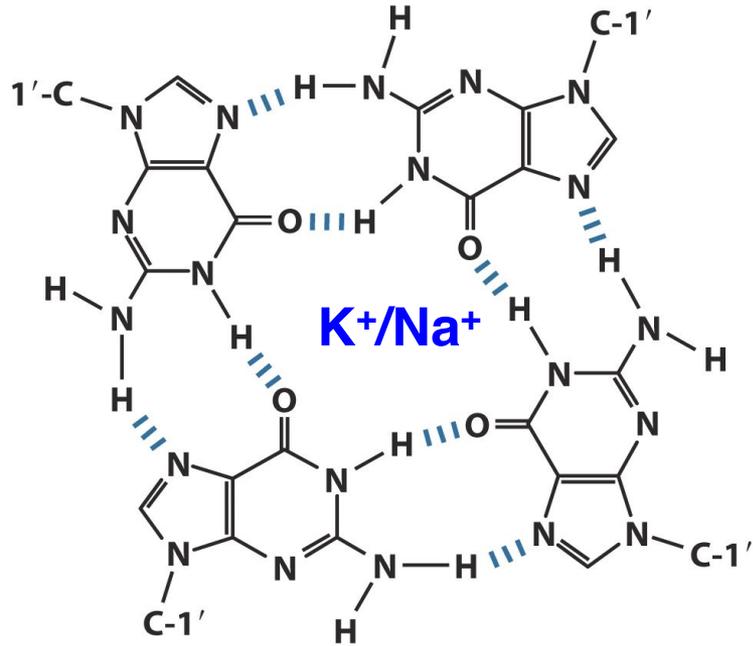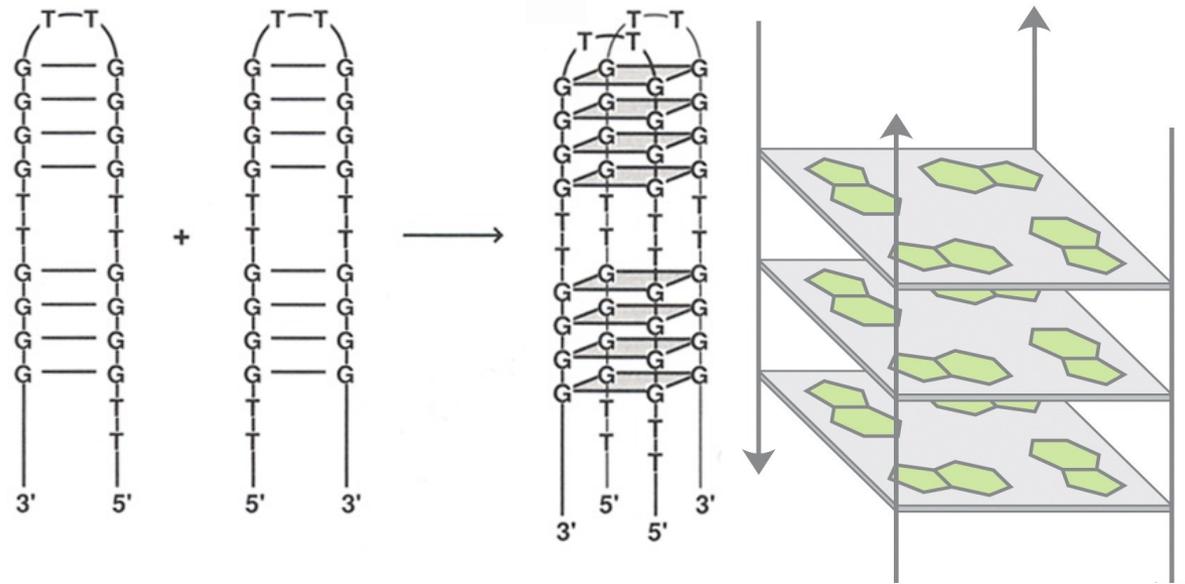
PDB ID:156D

Pseudo 3D of a G4x4 quadruplex motif

PyMol: Quadruplex.pse

G-quadruplex structures found in human chromosomes (blue) revealed by binding of an antibody (red) that recognizes G-quadruplex structures
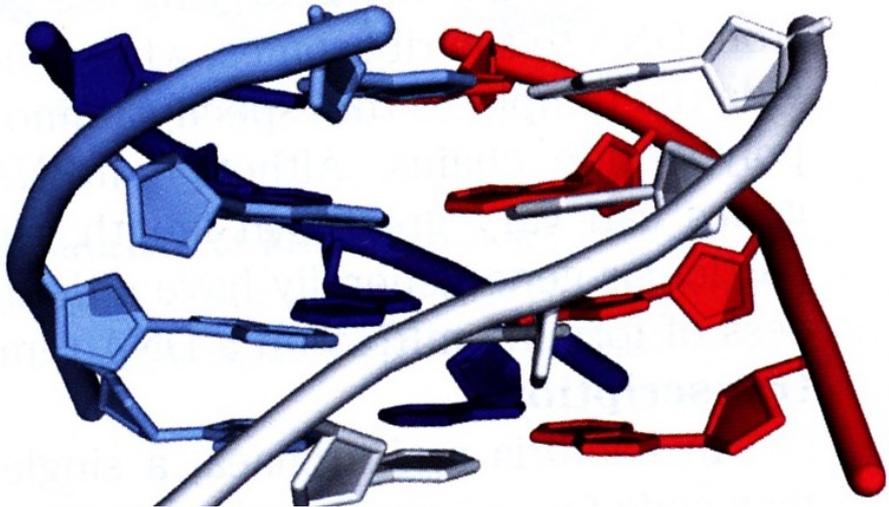
Biffi et al.
Nature Chemistry (2013)
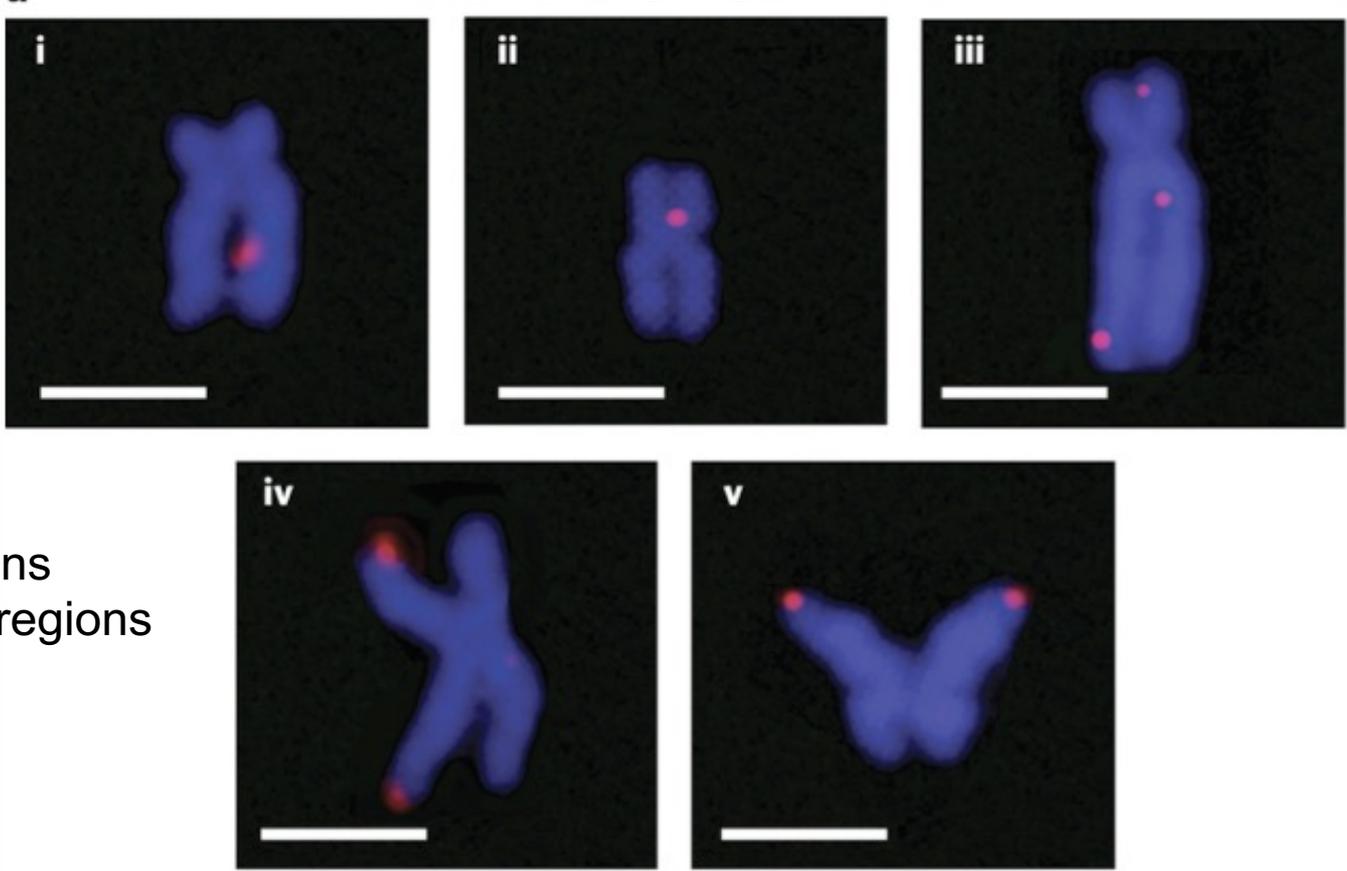
iv, v: Telomeric Regions
i, ii, iii: non telomeric regions

RNA structures

Is this a good representation of RNA structures?

A: Yes for mRNA and tRNA, not for rRNA

B: No mainly because mRNA is shown completely unstructured

C: Yes for tRNA only

D: No they are all oversimplified

Messenger RNA

Ribosome
Ribosomal RNA

Amino acid
Transfer RNA
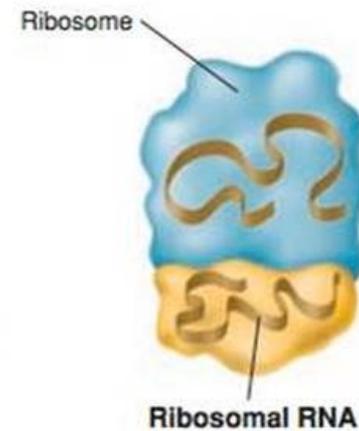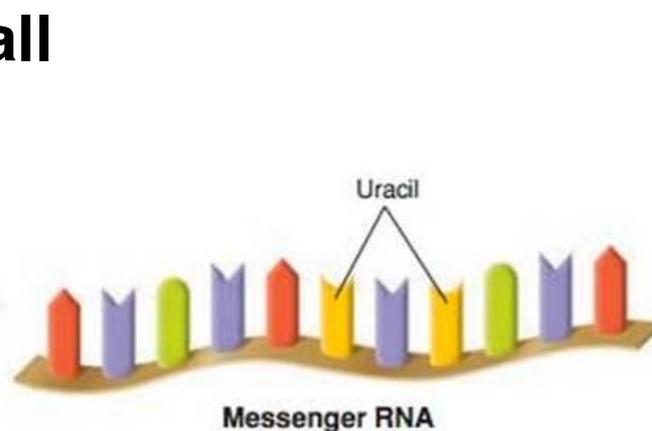
Secondary and Tertiary Structure of RNA

Basic elements of Secondary Structure in RNAs:
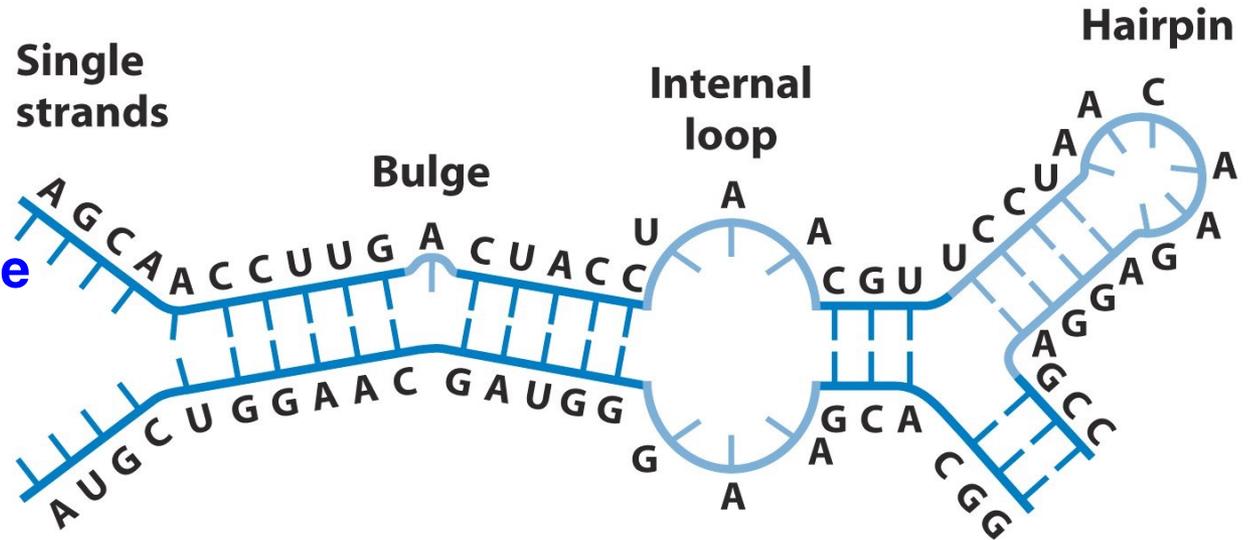
A hairpin segment in pseudo 3D:

Single strandedness nature of RNA makes it able to "fold" on itself and base-pair with complementary segments within the same molecule

• Some General Rules about RNA 3D structures/folding

RNAs are (in general) compact molecules - not limp spaghettis...

• Watson–Crick Pairing is Important but a bit overrated

• Non-Watson–Crick (aka non canonical) pairing is very much underrated

• Base Stacking is a major driver for RNA Structure
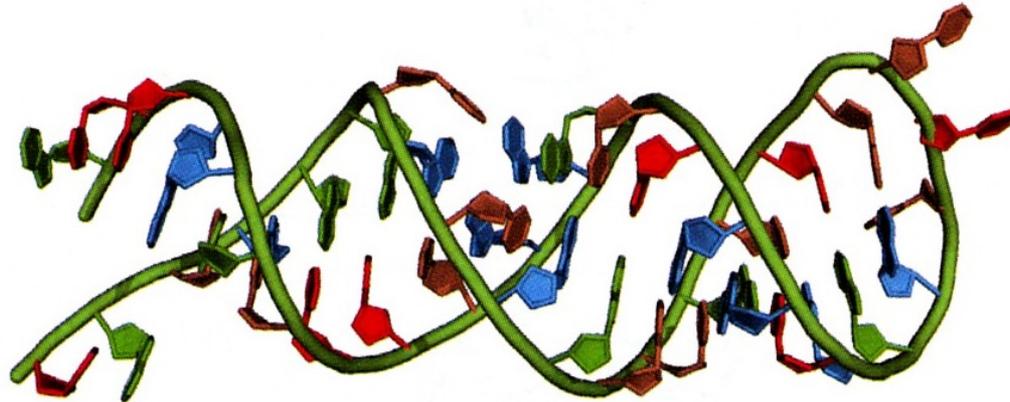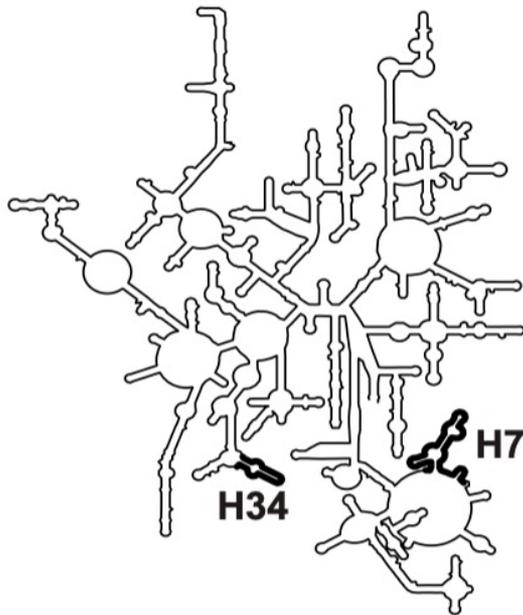
From:
Vicens and Kieft
PNAS 2022)

A — 'ladders, loops & bubbles' how we usually represent the secondary structure of an RNA

B — from up close: what do 'ladders, loops & bubbles' look like? examples around **H7** and **H34**

this region is actually a kink turn

correct secondary

this region actually forms a continuous stack

stem with a tetraloop
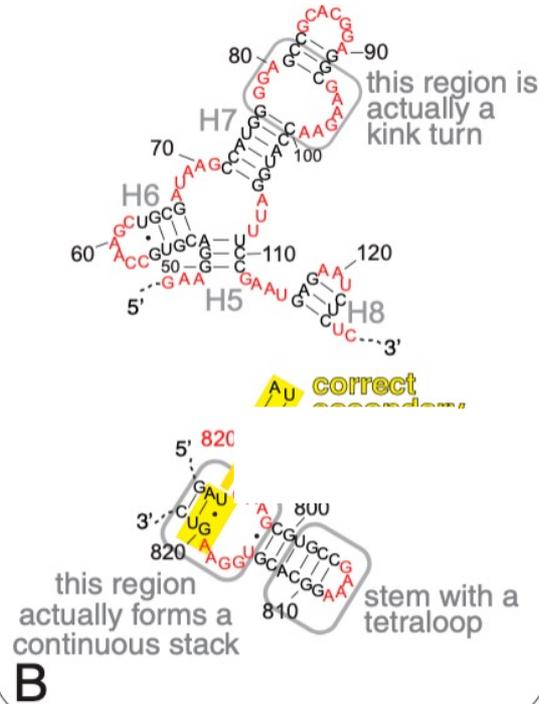
C — in three dimensions: where are the 'loops & bubbles'? nucleotides in red are not 'unstructured' and not even 'unstacked'

double cross-strand stacking

actually organized as:
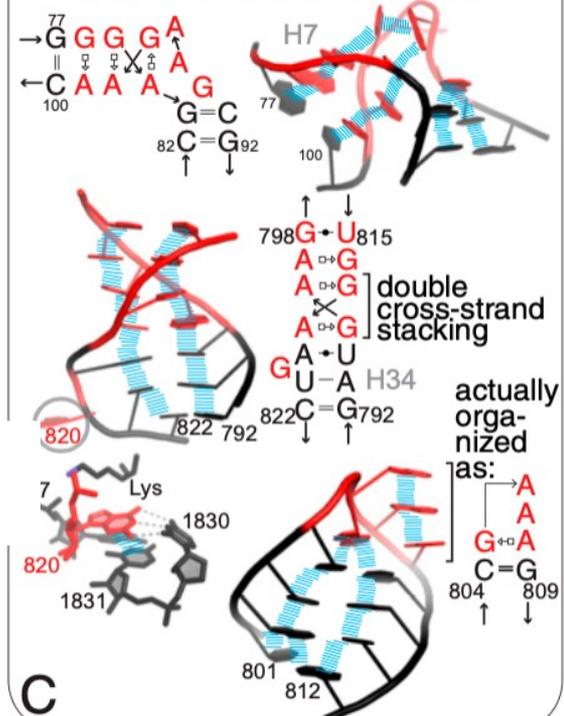
PyMol: GAAA stem loop.pse

Secondary vs 3D Structures: Example of transfer RNA (tRNA)

tRNA cloverleaf secondary structure

"roadkill" version…

Acceptor stem

D-arm

T-arm

D-arm

PyMol: tRNA.pse

Visualize in PyMol:

Anticodon loo
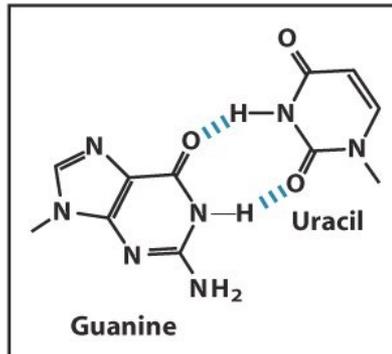
• **Several G:U base pairs: "A" form tolerates the presence of G:U base pairs**

tRNA L-shaped tertiary structure

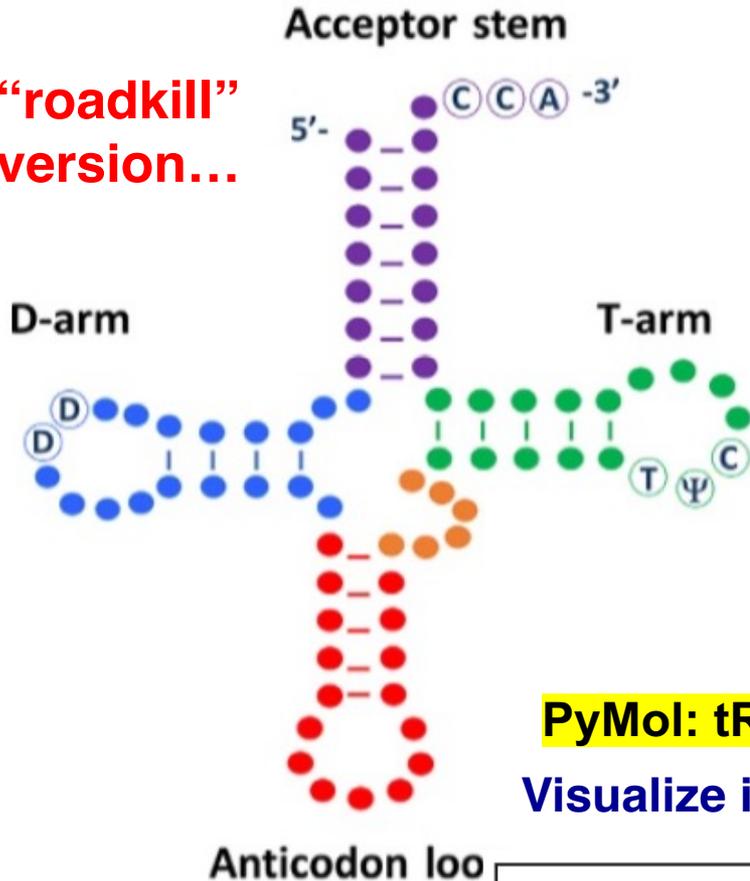Elbow

Amino acid accepting
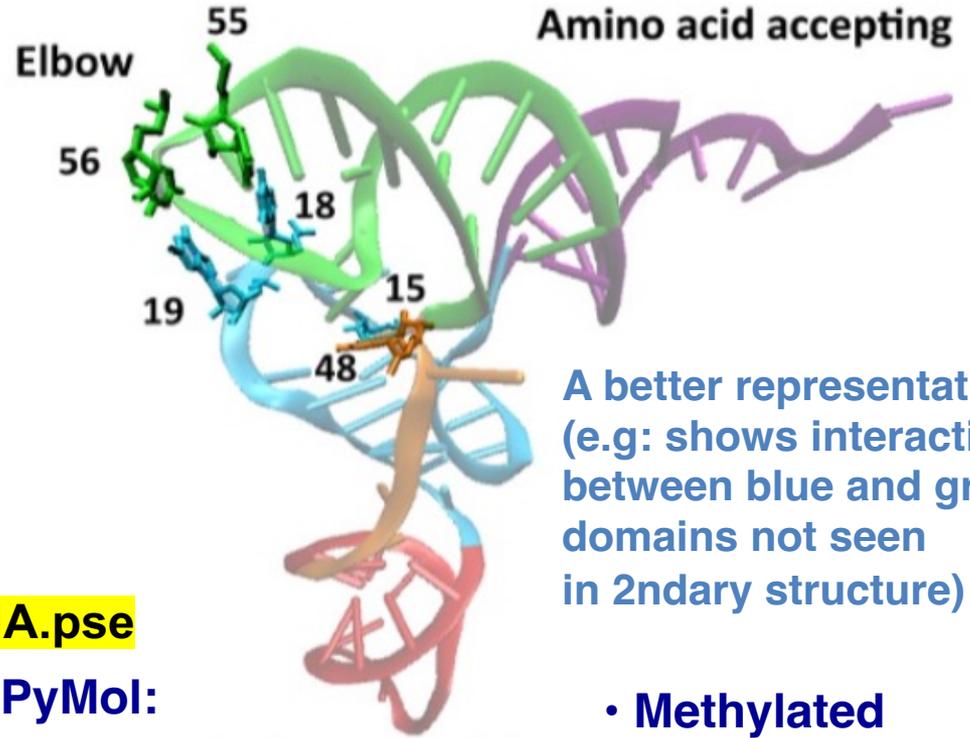
A better representation.. (e.g: shows interactions between blue and green domains not seen in 2ndary structure)

Codon recognition

• **Methylated base**

• **Non canonical base pairs: e.g. A:A**

• **Cations promote 3D folding**

Why does the A-form of dsRNA tolerate the presence of G-U base pairs?

A: Because the major groove is deeper which allows it to fit

B: Because G:U base pairs are isosteric to Watson-Crick base pairs in A-type double helices, but not in B-type.

C: Because divalent cations frequently bind RNA, which stabilize weak base pairs like G:U

D: Because the wider diameter of A-type double helices allows for more flexibility in the geometry of the base pairs that can be included
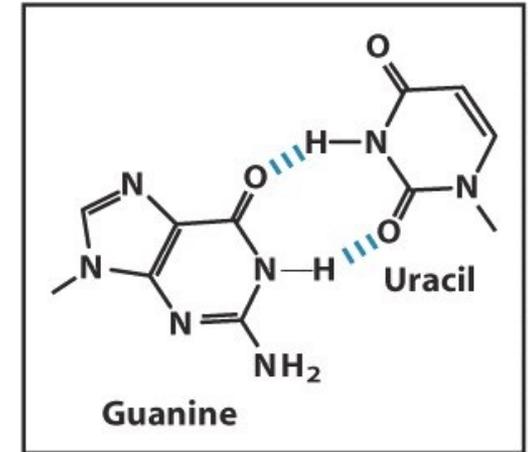
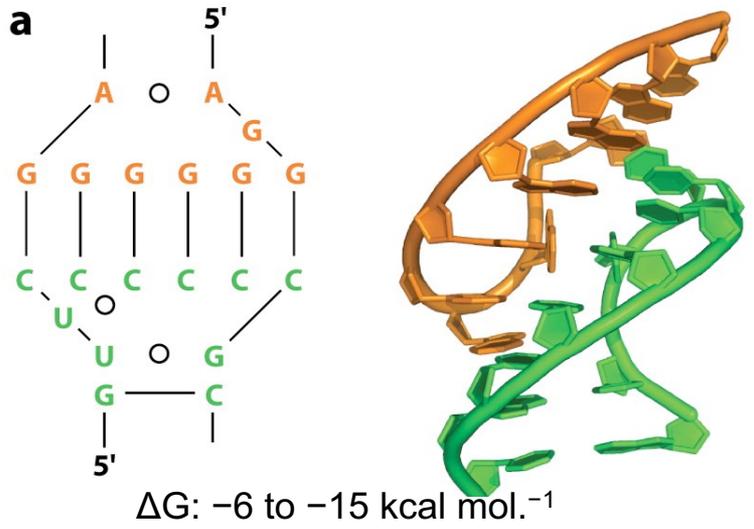# Some structural motifs that help fold RNAs in 3D structures:

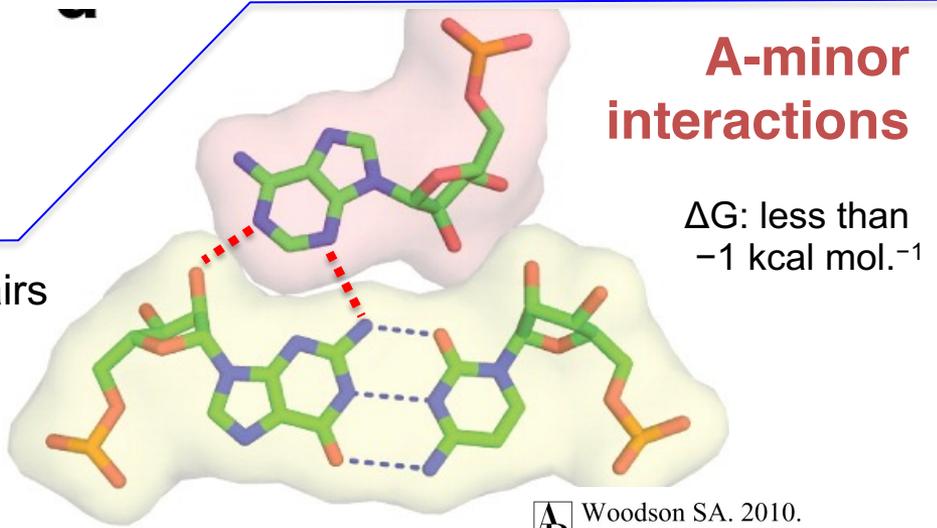## pseudoknot /kissing loops

ΔG: −6 to −15 kcal mol.$^{-1}$

• Canonical (=Watson-Crick) base pairs between remote sequences form **Pseudoknots** or **kissing loops** structures

• Non-Canonical (=non Watson-Crick) base pairs between remote sequences.
One example specific example is the A-minor interaction (found in ~25% of adenosines of small ribosomal RNA

• Interactions between elements of the backbone, for example ribose zippers:

Ribose zippers

ΔG: less than −1 kcal mol.$^{-1}$

Brings two strands/backbone close together
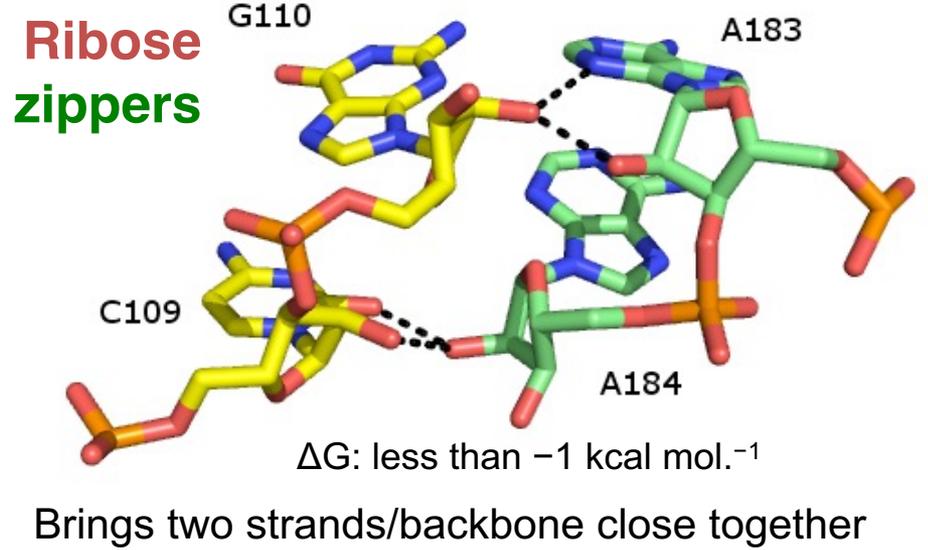
A-minor interactions

ΔG: less than −1 kcal mol.$^{-1}$

Woodson SA. 2010.
Annu. Rev. Biophys. 39:61–77

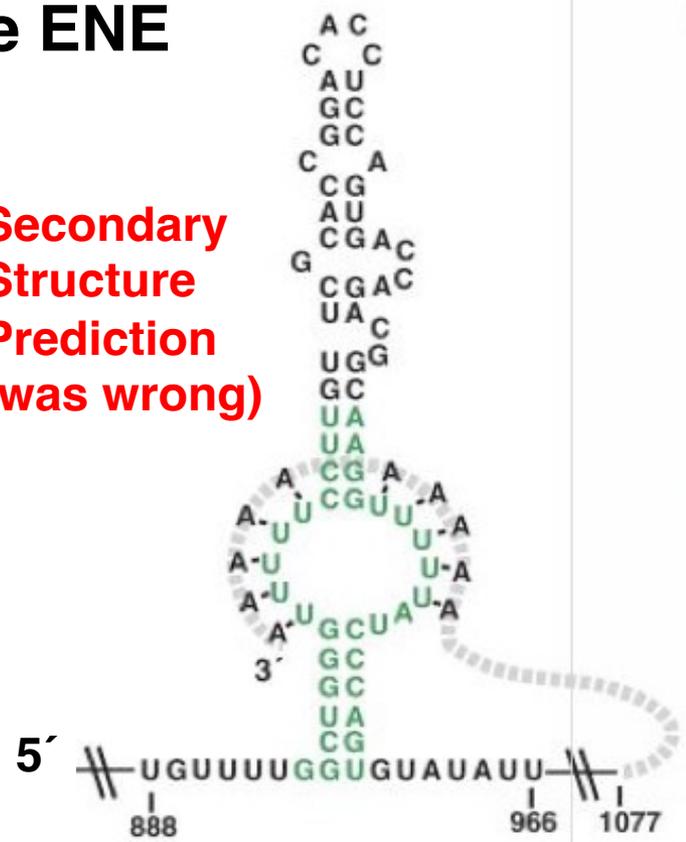# Triple Helical Structures: Example of the ENE Structural element in a viral RNA

Kaposi's sarcoma-associated herpesvirus (KSHV) produces the PAN RNA, a highly abundant, nuclear noncoding RNA
A triple helical structure at the 3´-end of the PAN RNA is important to protect this RNA against degradation by cellular ribonucleases.

Secondary Structure Prediction (was wrong)

Actual Triple Helix structure

5´

PDB ID 3P22

Mitton-Fry et al. Science 2010

U-A-U base pairs