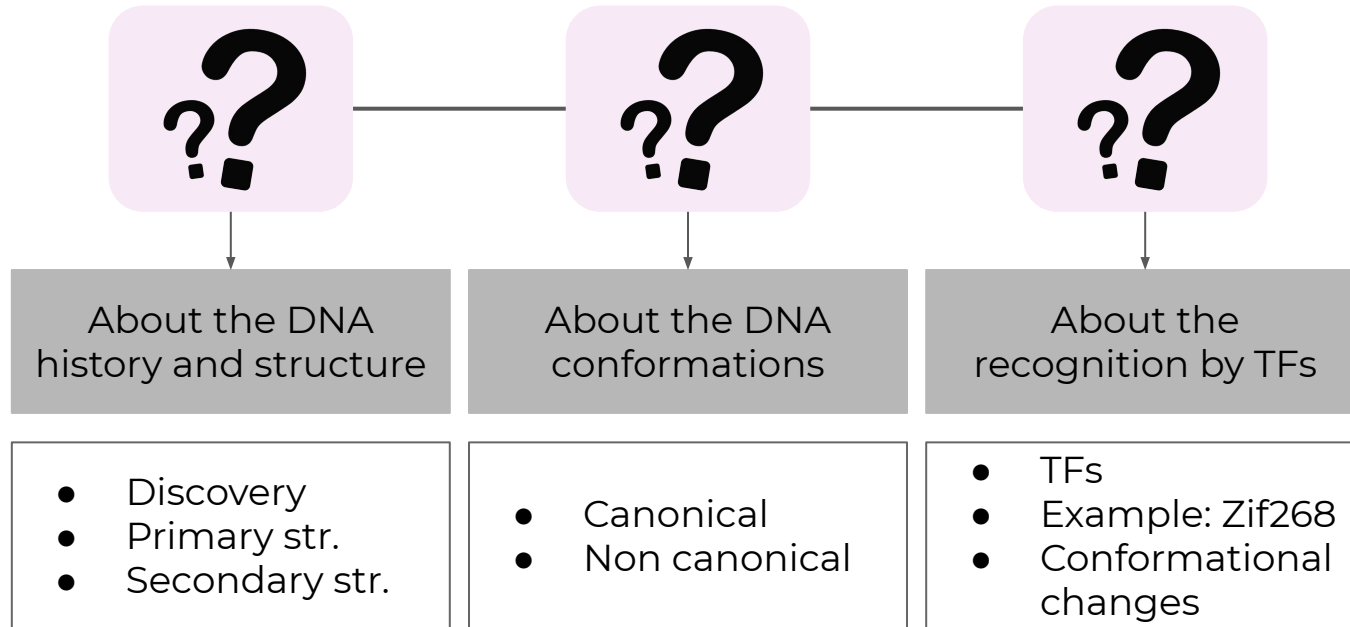




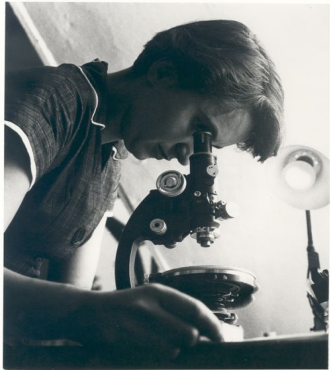
DNA structure and recognition by TF

**Jessica Brillas
Berta Estévez
Carla Garrido**

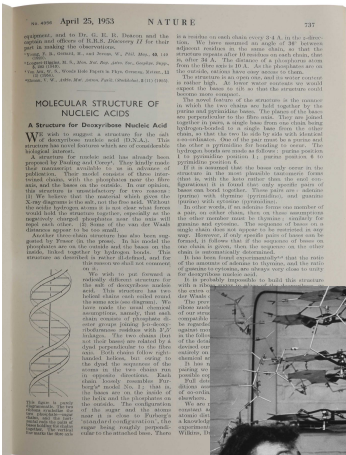
Workflow



DNA structure discovery

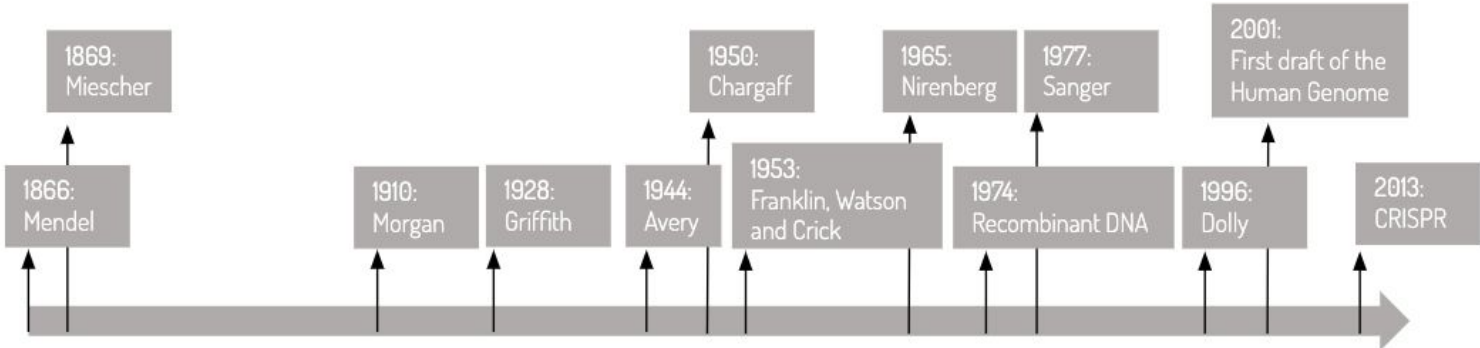
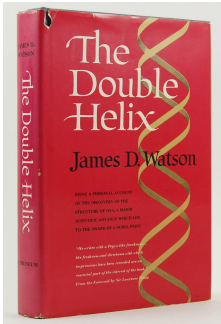


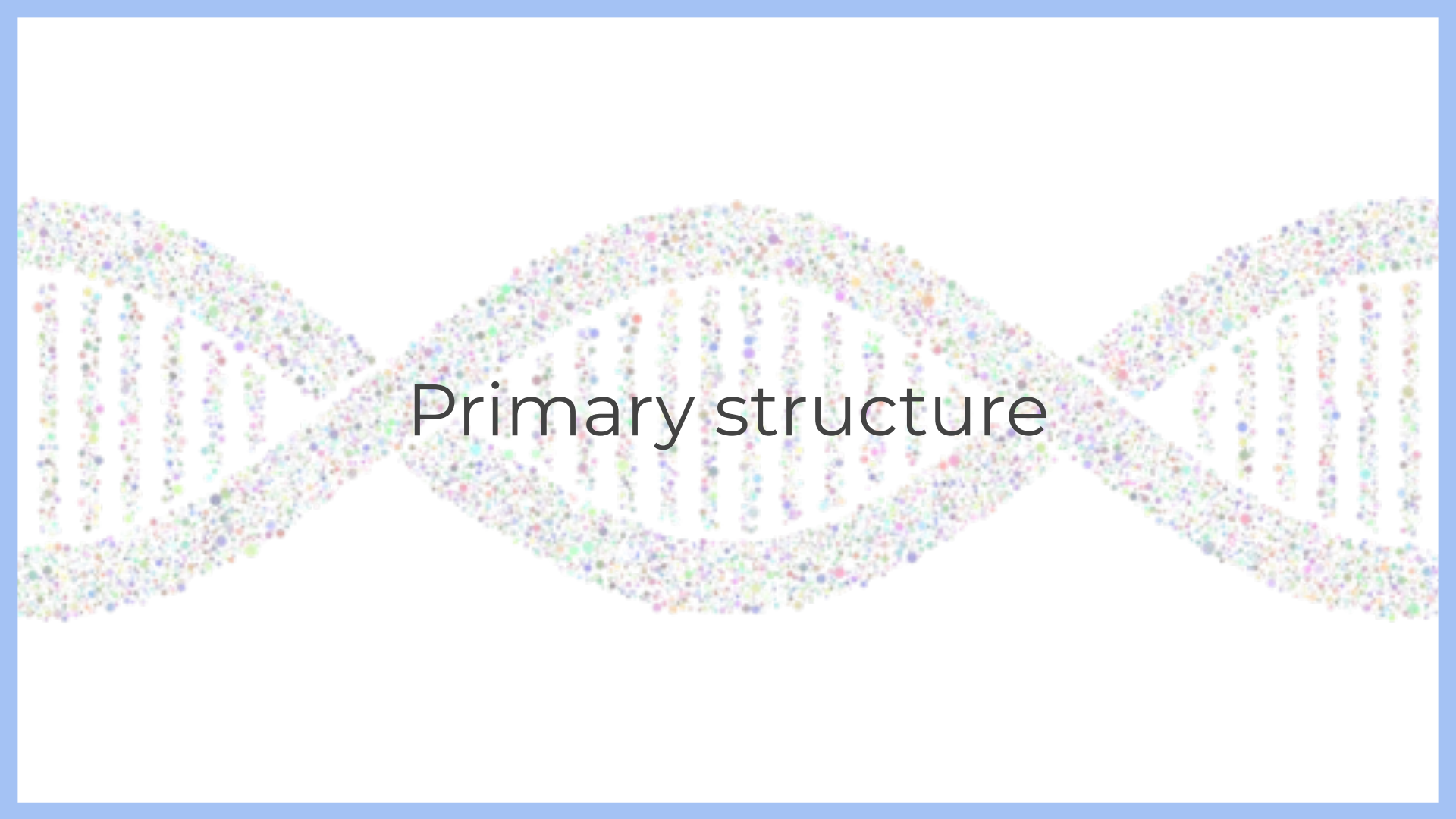
Rosalind Franklin



Watson and Crick

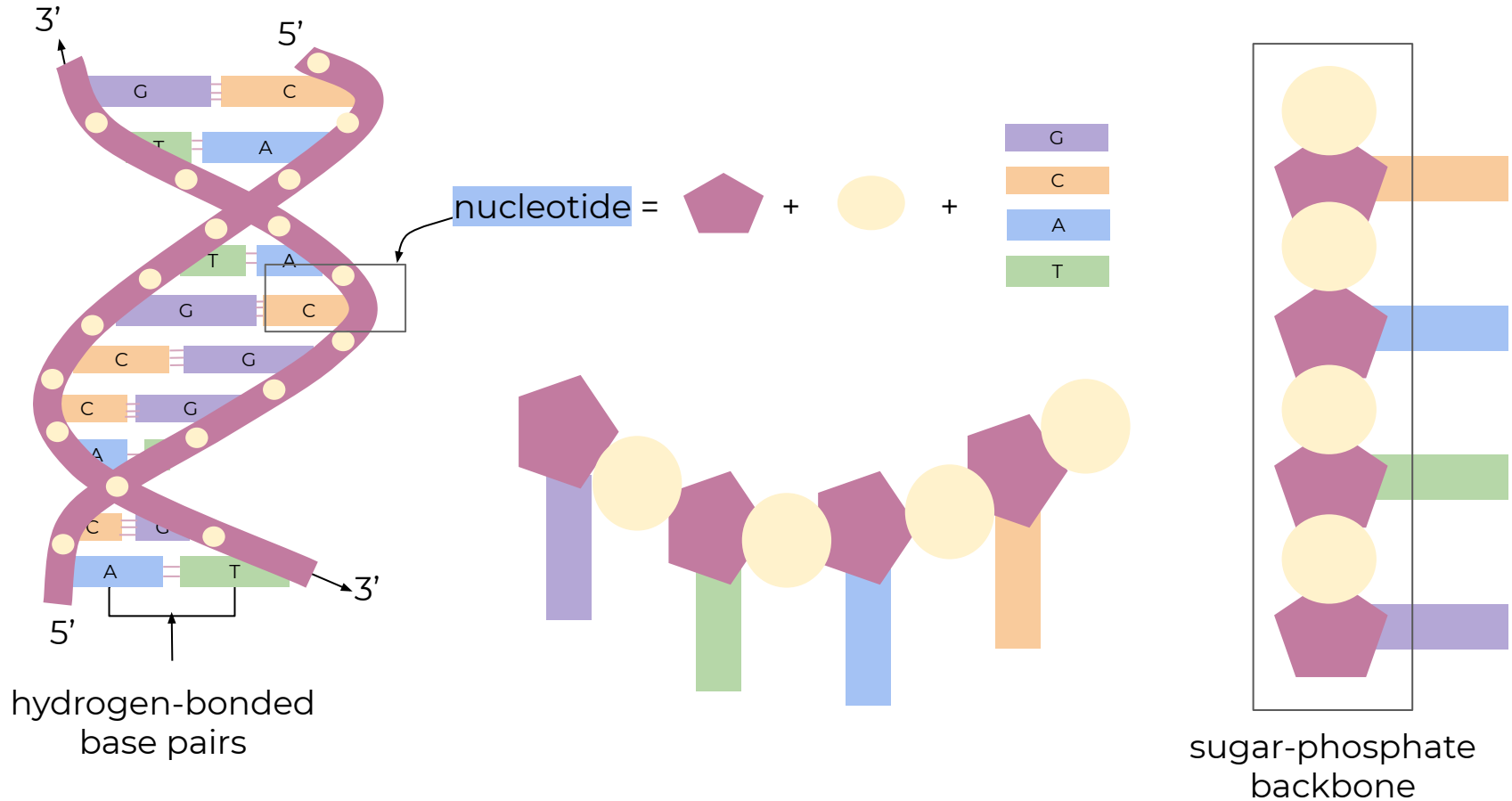
- Other important names:
- Avery and colleagues
 - Hershey and Chase
 - Erwin Chargaff



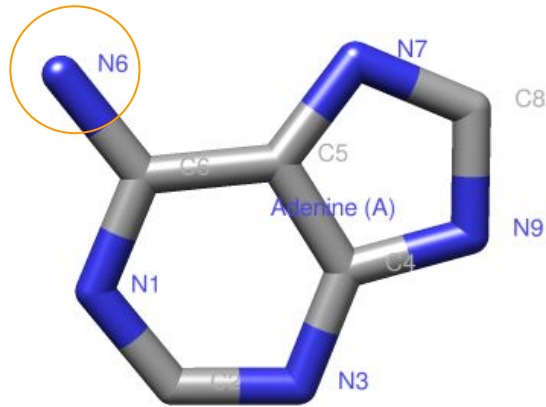
A stylized representation of a DNA double helix. The structure is formed by two intertwined strands, each composed of a dense collection of small, multi-colored dots (red, green, blue, yellow, purple, etc.) that represent individual nucleotides. The dots are arranged in a way that creates a continuous, wavy path for each strand. The two strands are connected by horizontal rungs, also made of dots, representing the base pairs. The overall effect is a textured, three-dimensional appearance of the DNA molecule.

Primary structure

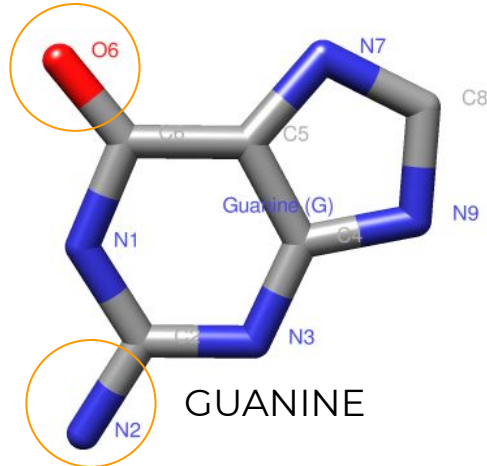
Primary structure: overview



Bases: PURINES and PYRIMIDINES



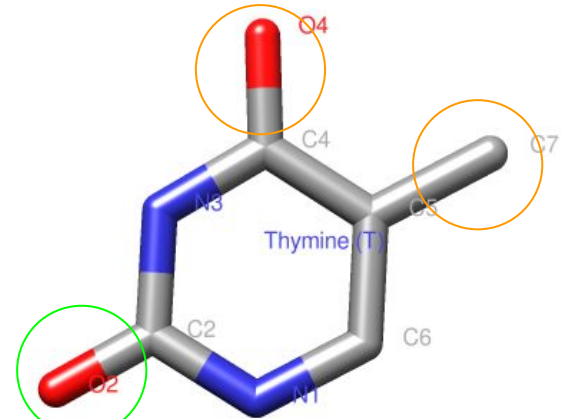
ADENINE



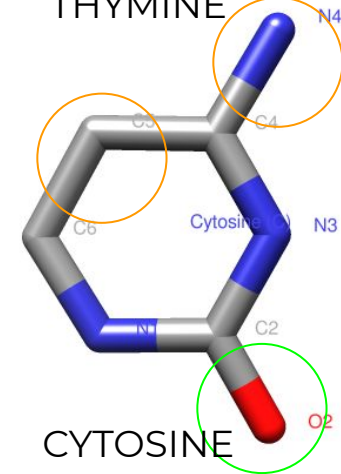
GUANINE

Purines

Pyrimidines

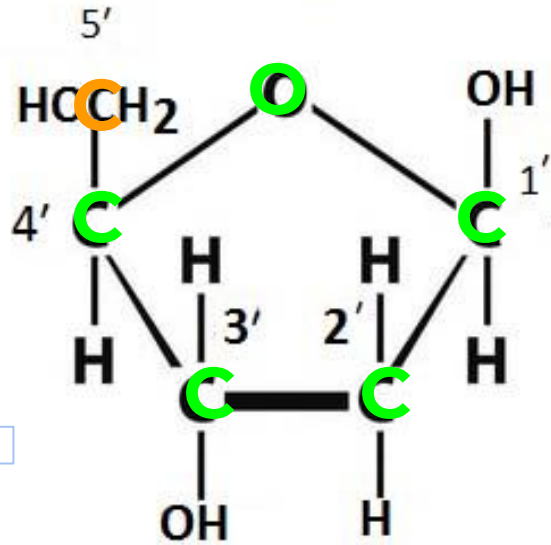


THYMINE

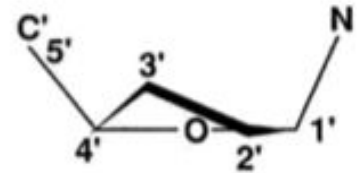
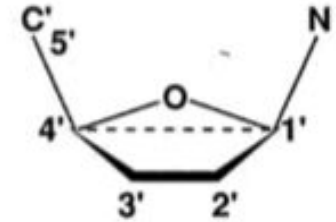
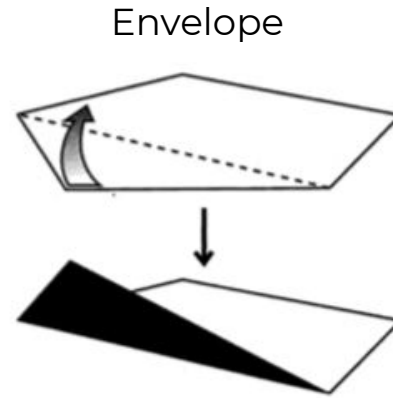


CYTOSINE

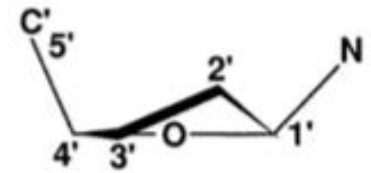
Sugar: DEOXYRIBOSE



beta-D-2-deoxyribose



3' endo



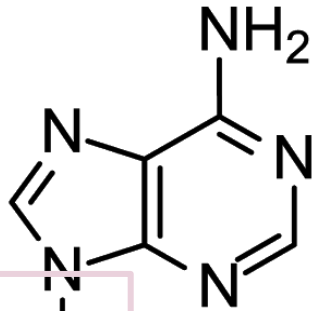
2' endo

Nucleosides

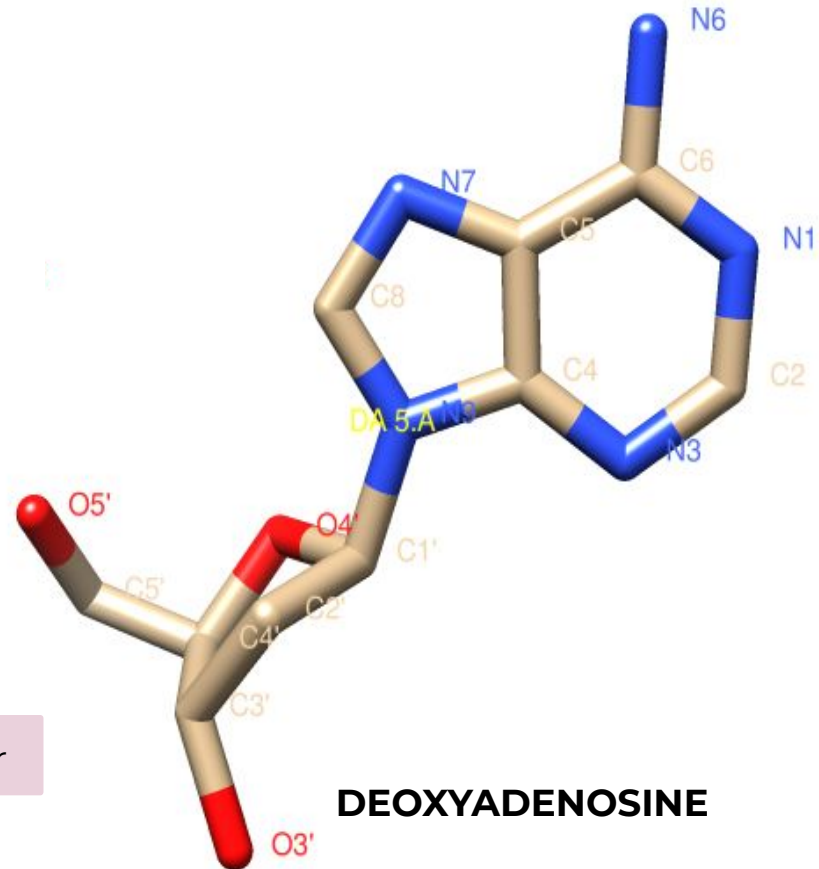
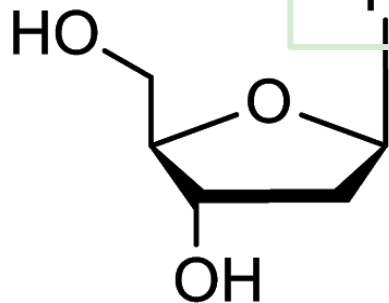
Sugar + Base = **Nucleoside**

Deoxyadenosine
Deoxyguanosine
Deoxythymidine
Deoxycytidine

Base

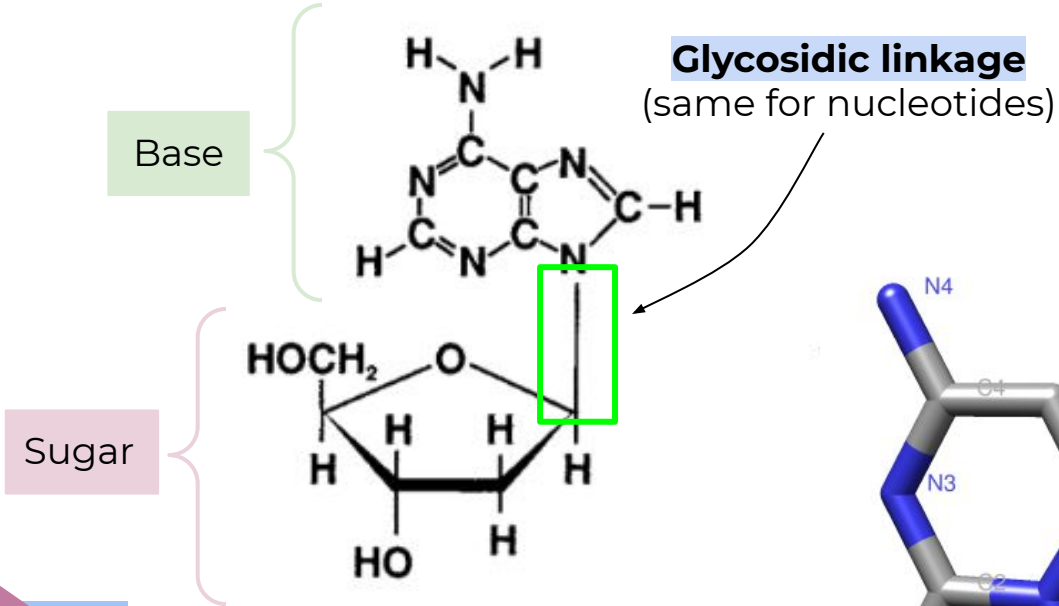


Sugar



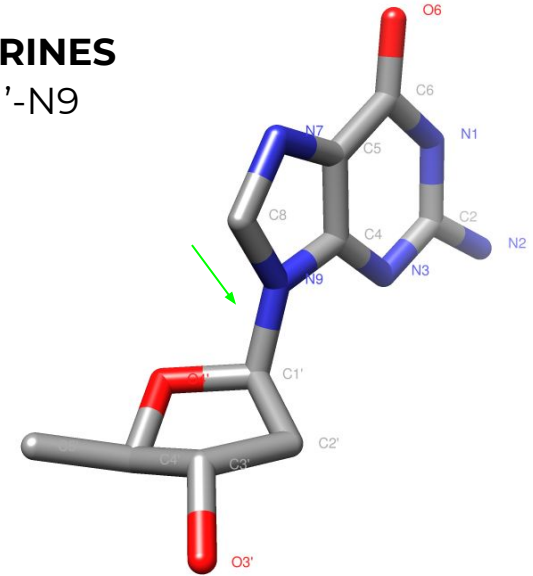
DEOXYADENOSINE

Glycosidic bond



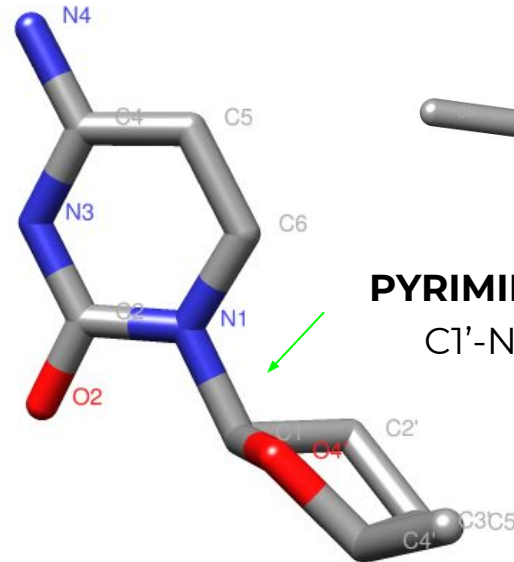
PURINES

C1'-N9

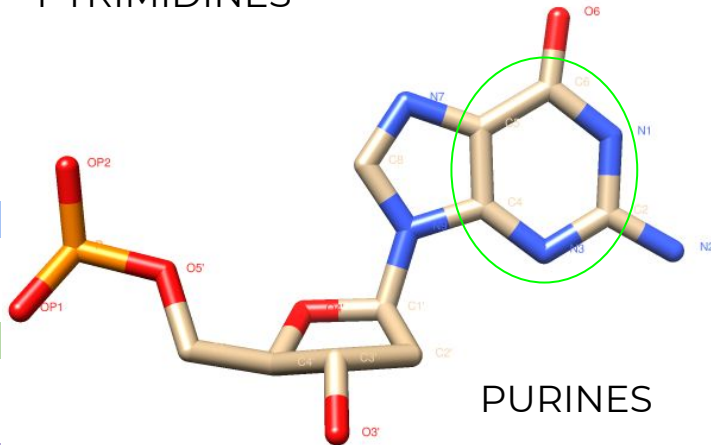
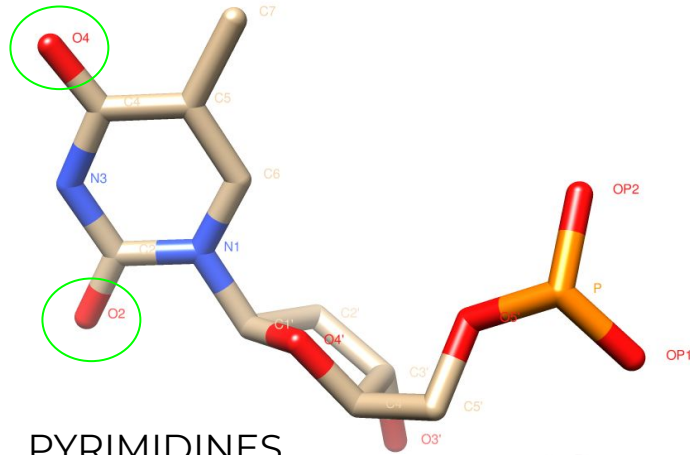


PYRIMIDINES

C1'-N1

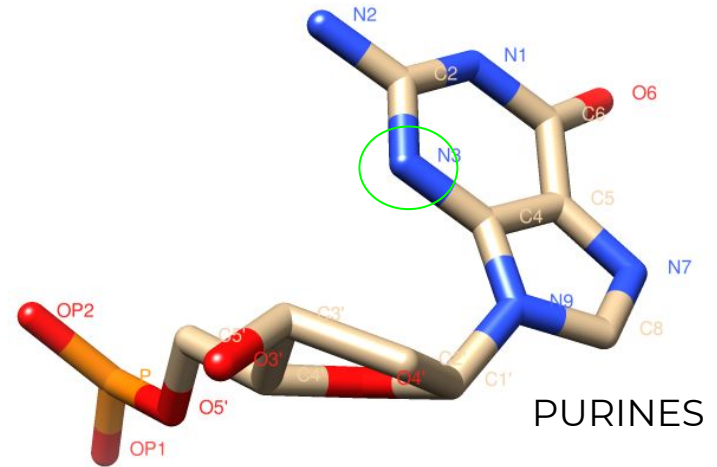


Base conformation



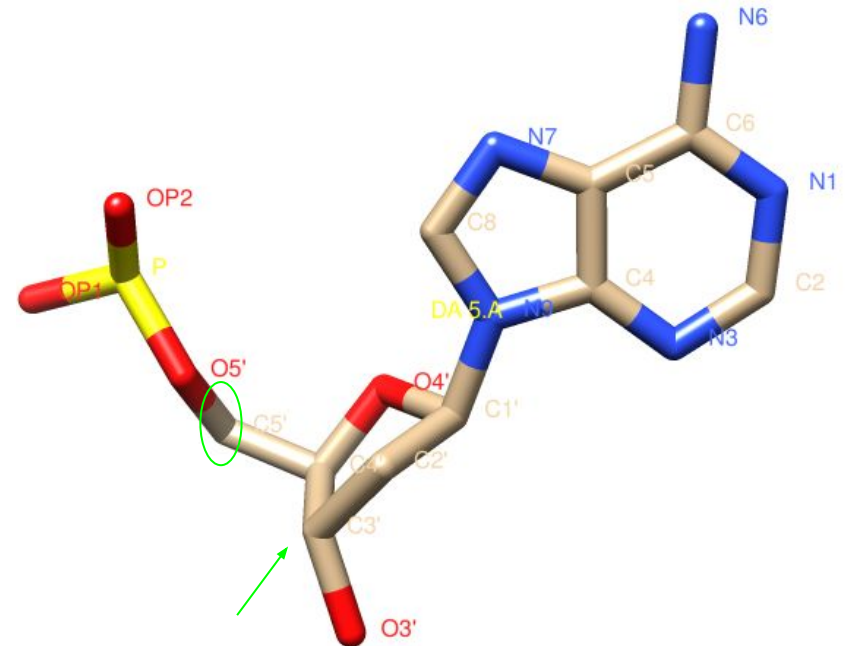
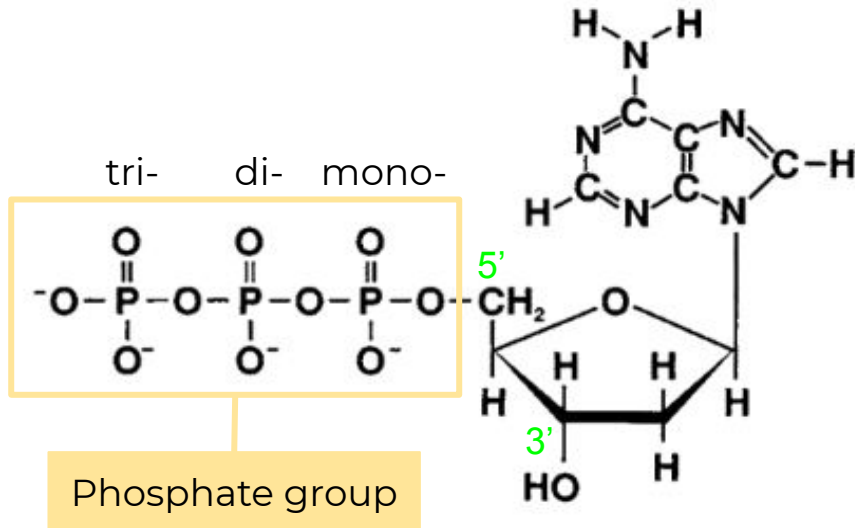
Anti

Syn



Phosphate group and nucleotides

Sugar + Base + Phosphate group = **Nucleotide**



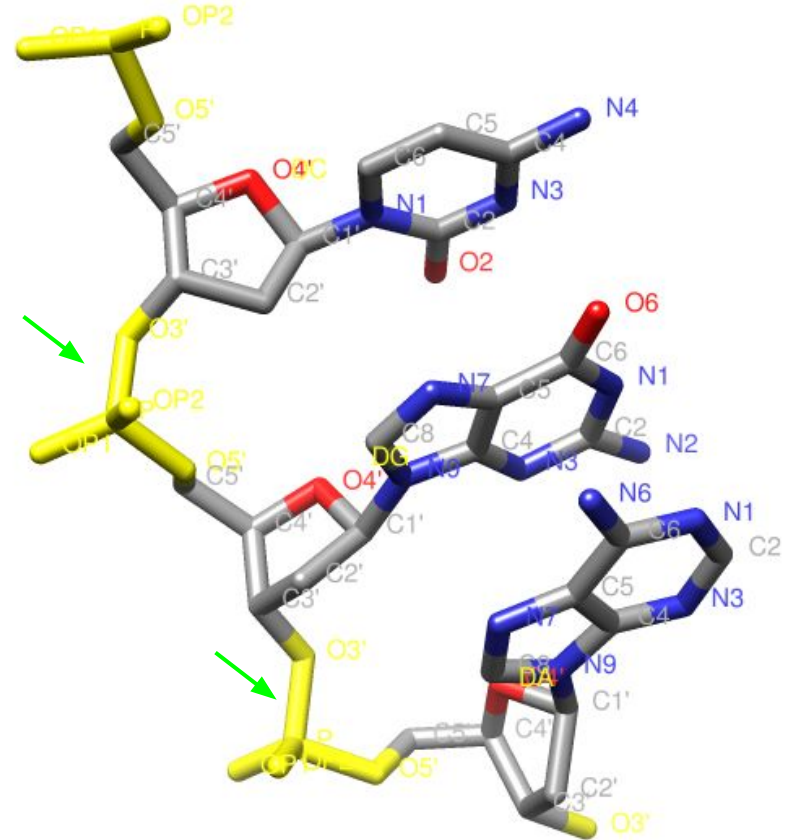
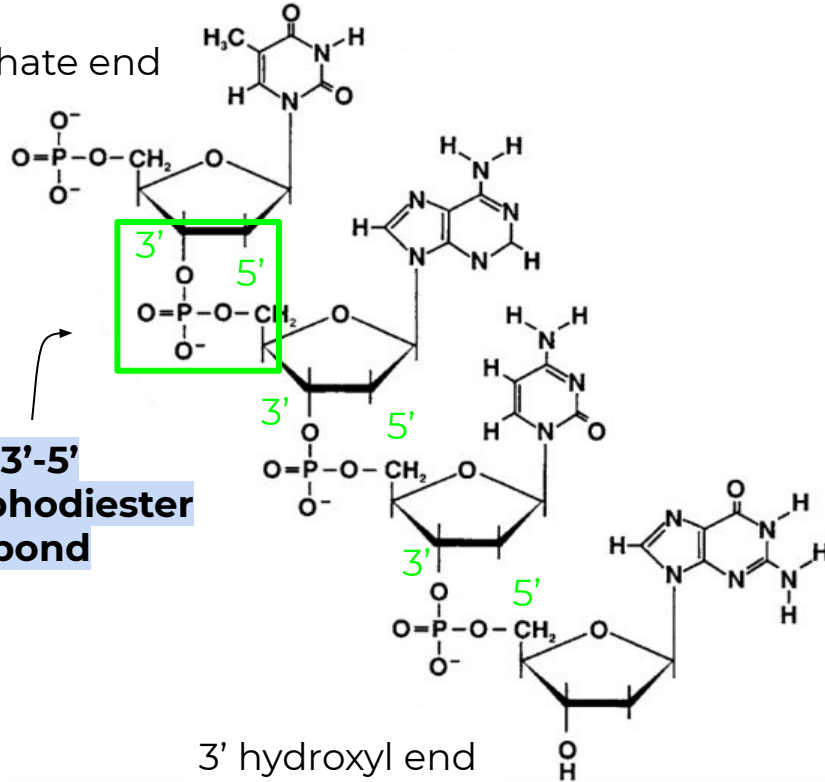
**DEOXYADENOSINE
5'MONOPHOSPHATE**

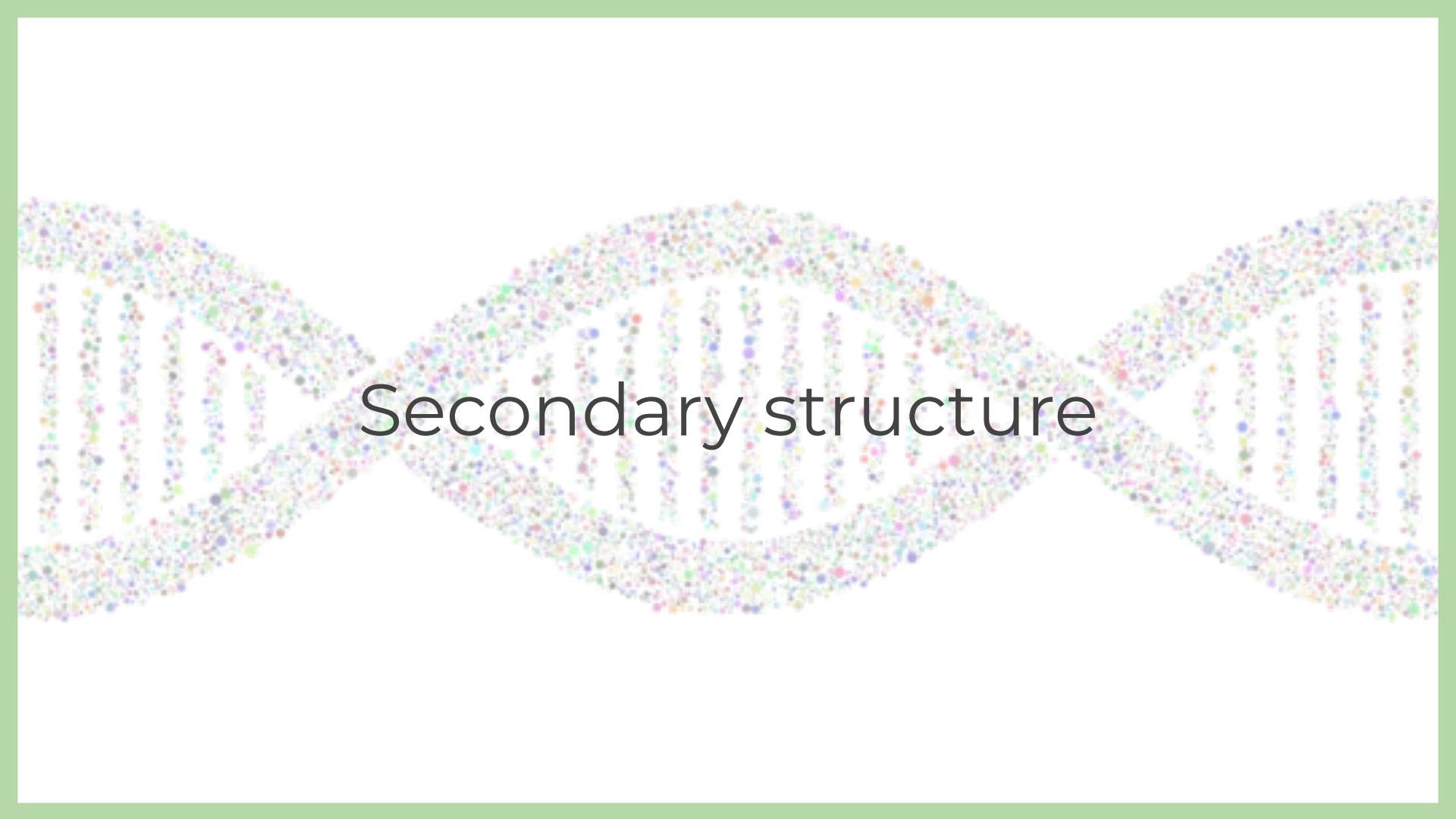
Phosphodiester bond

5' phosphate end

3'-5'
phosphodiester
bond

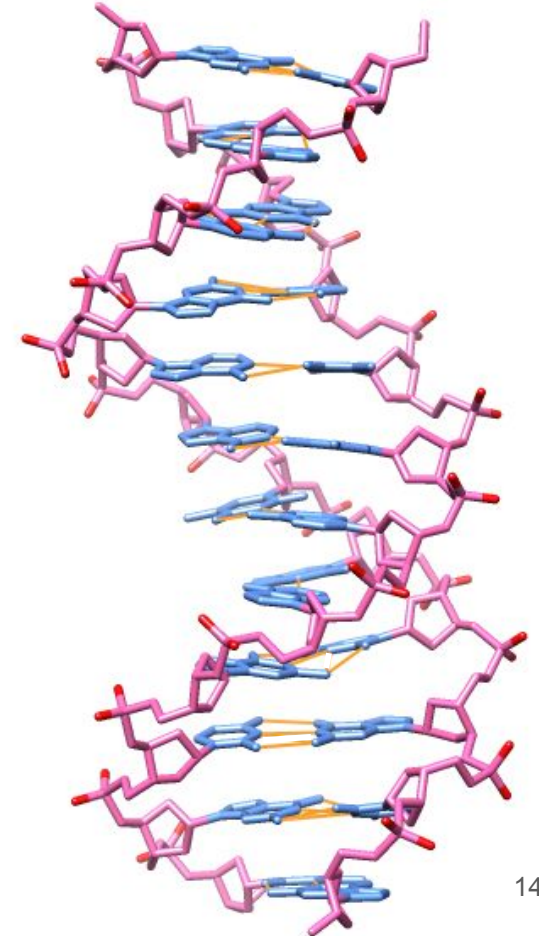
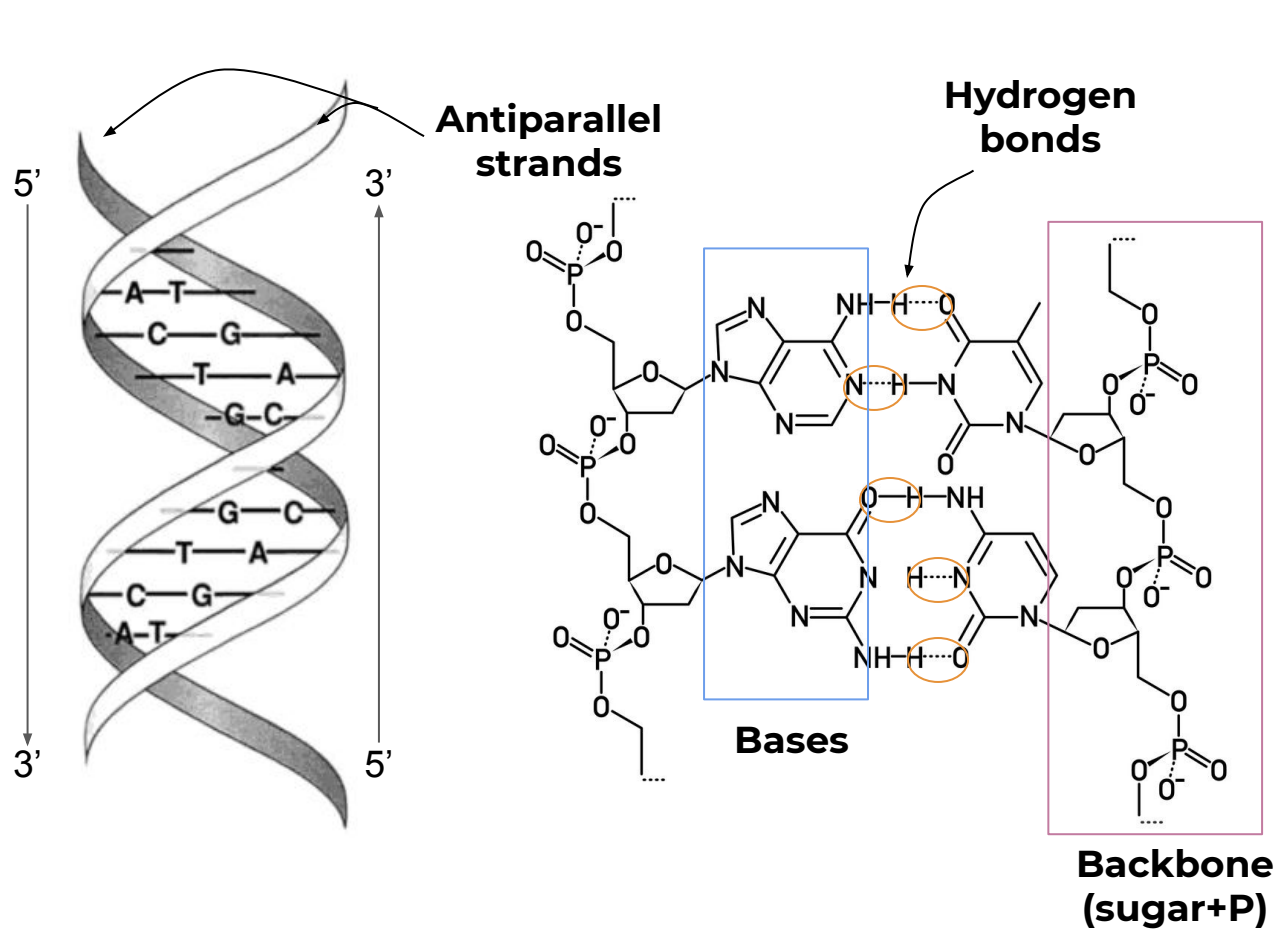
3' hydroxyl end



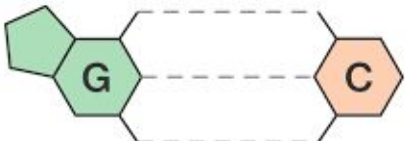


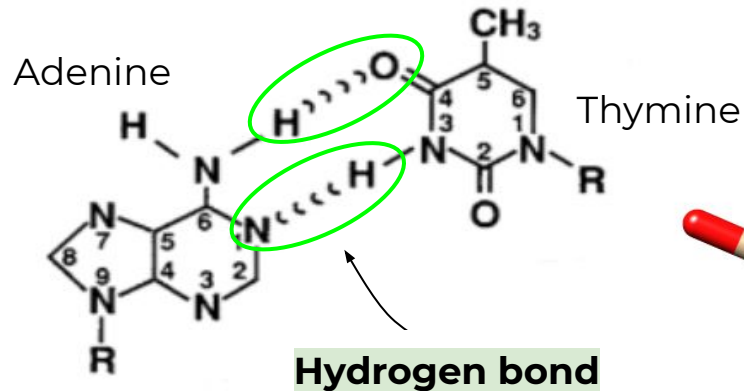
Secondary structure

Secondary structure

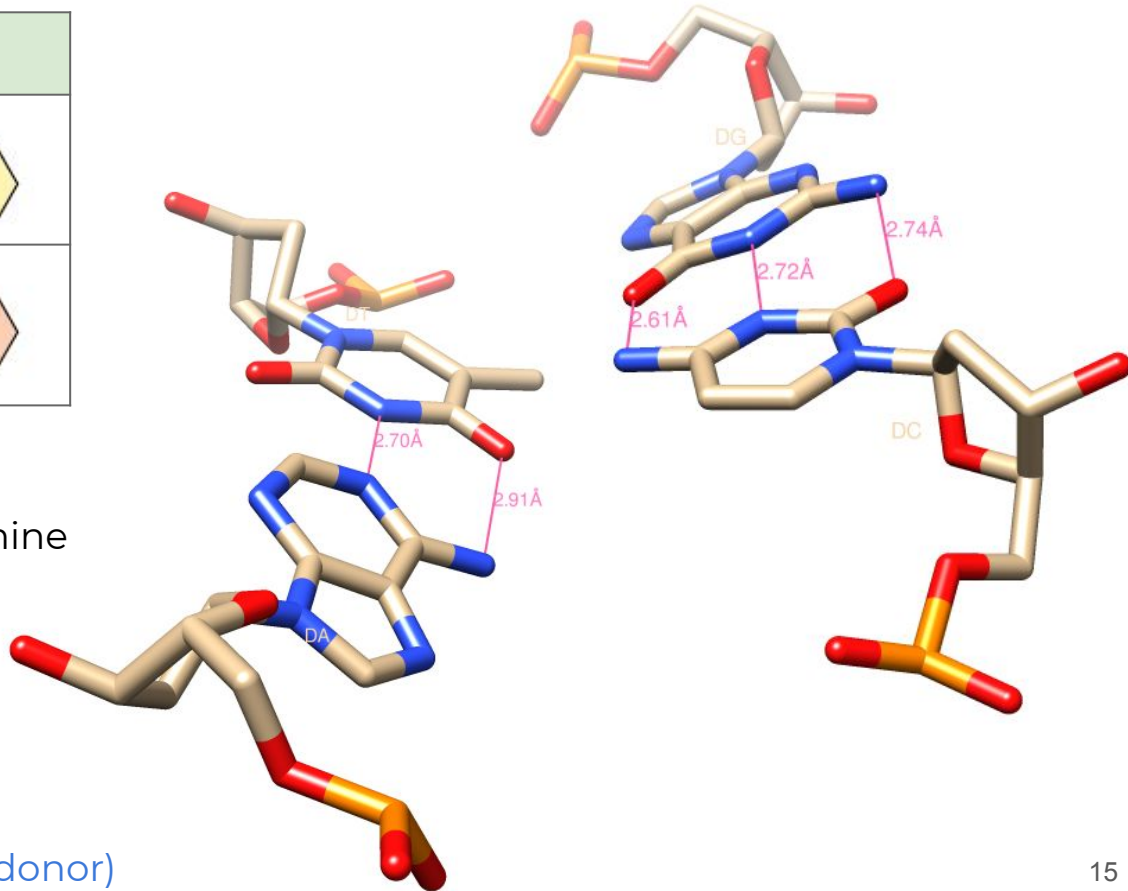


Base pairing and Hydrogen bonding

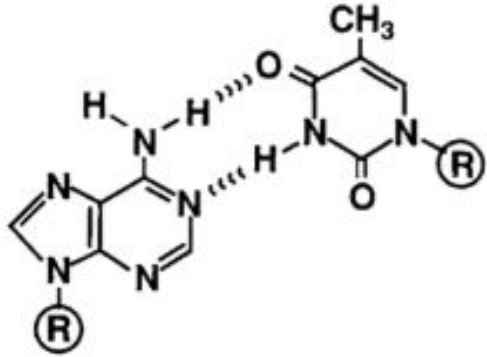
	CHARGAFF RULES
2 H-bonds	
3 H-bonds	



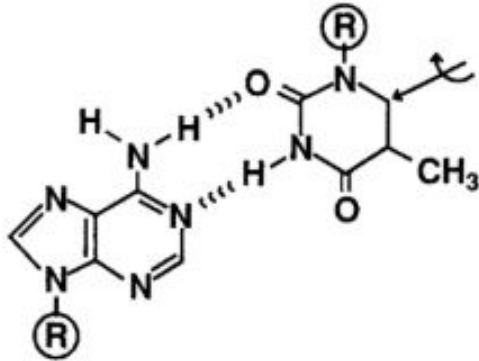
(acceptor) **A** ||||| **H** — **D** (donor)



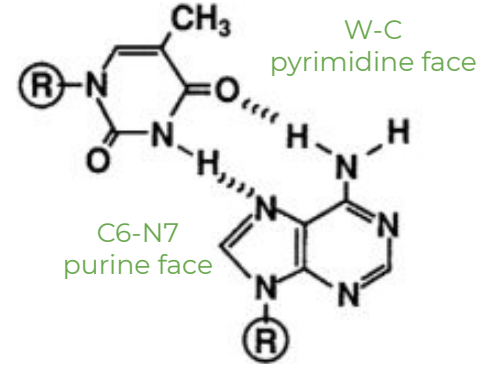
Base pairing



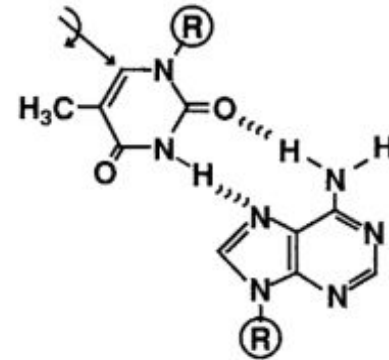
Watson Crick (A·T)



Reverse Watson Crick (A·T)



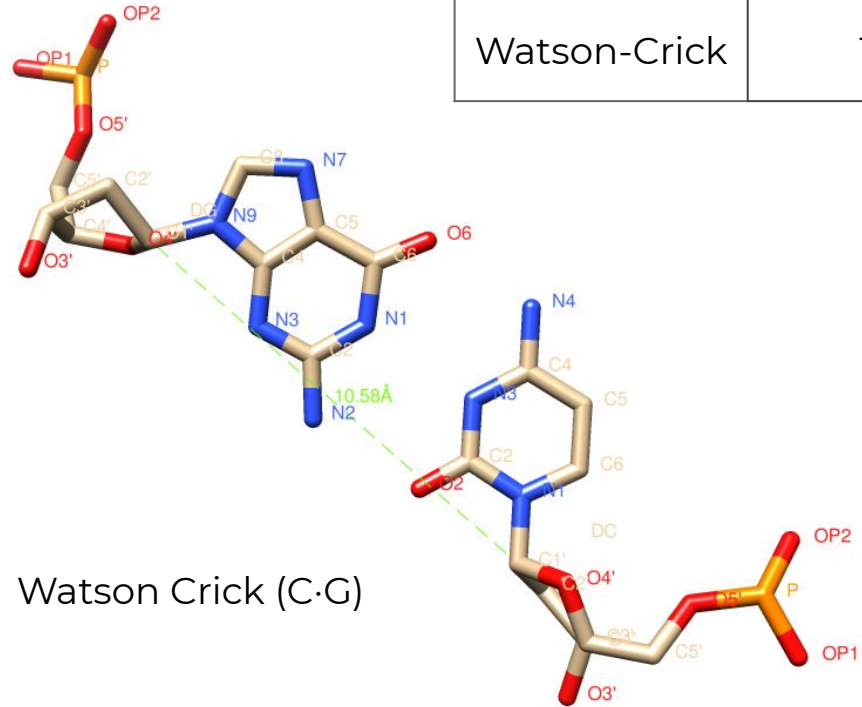
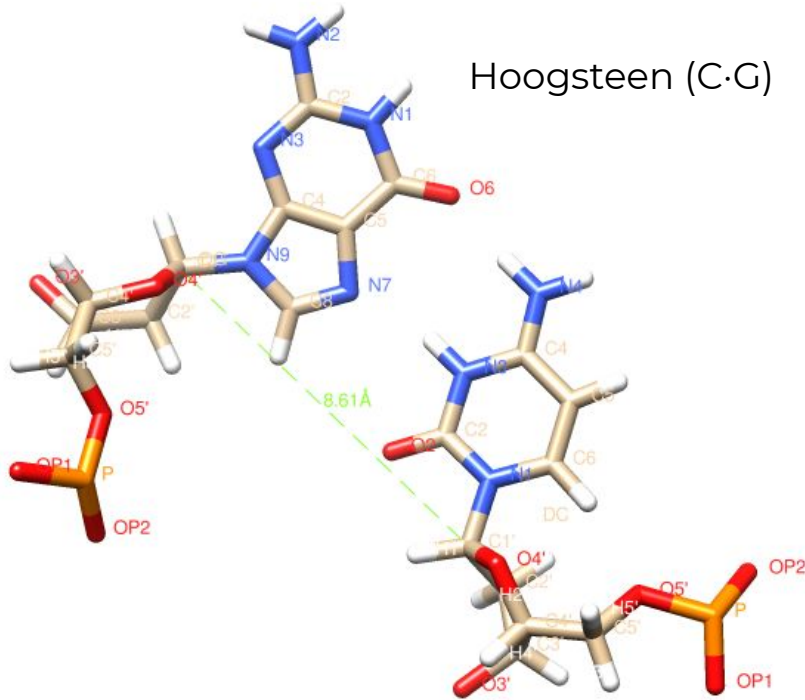
Hoogsteen (A·T)



Reverse Hoogsteen (A·T)

Base pairing

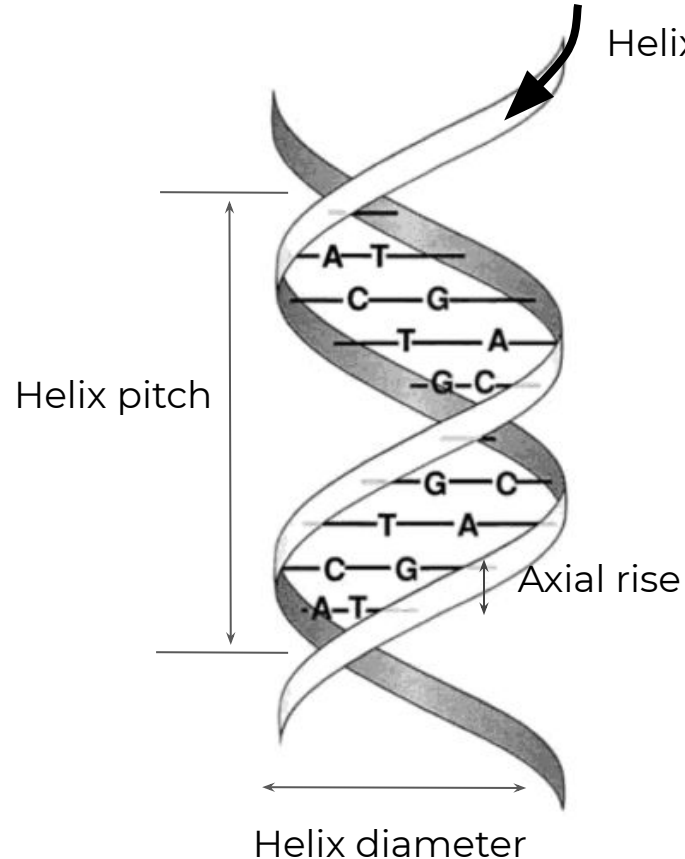
Hoogsteen (C·G)



Watson Crick (C·G)

	C1-C1 distance
Hoogsteen	8.65Å
Watson-Crick	10.5Å

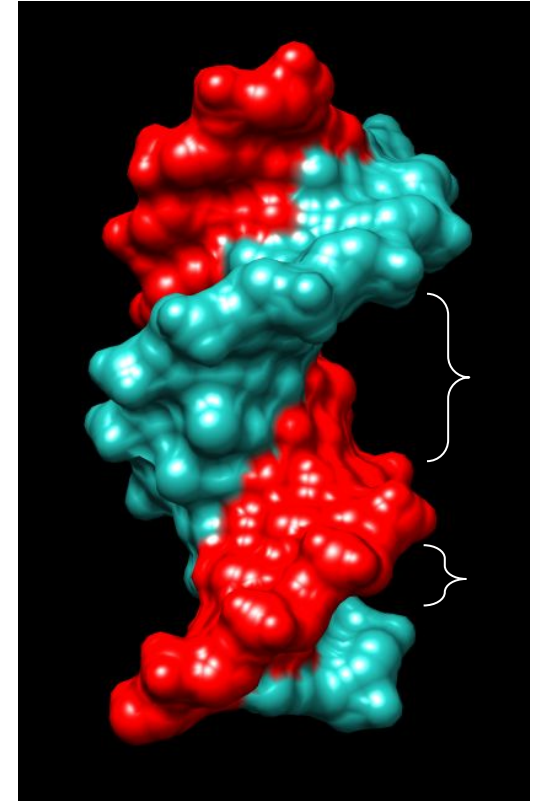
Helix parameters

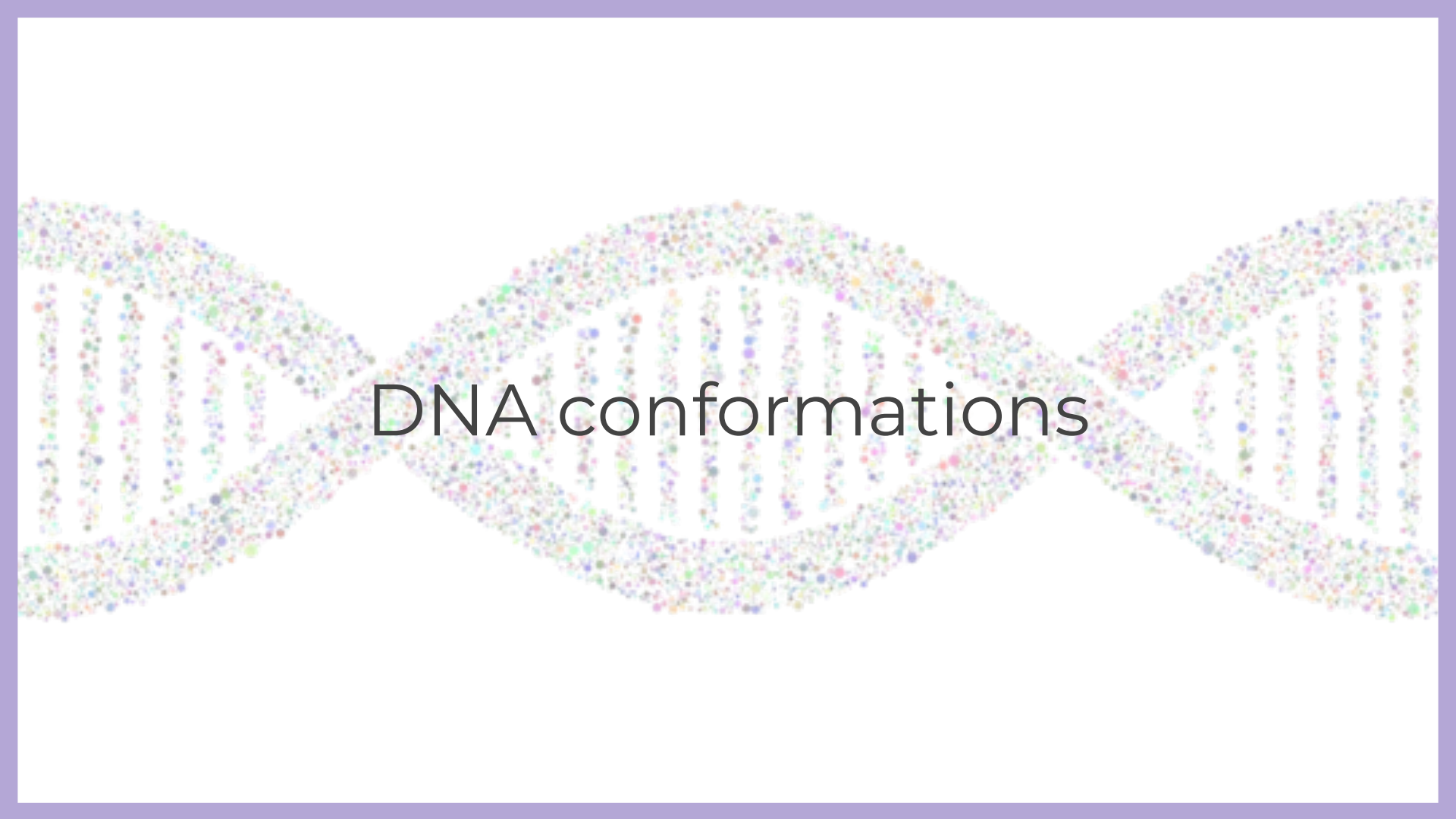


Helix parameters:

- Helix sense
- Residues per turn
- Axial rise
- Helix pitch
- Base pair tilt
- Diameter of the helix
- Rotation per residue

Major groove/Minor groove





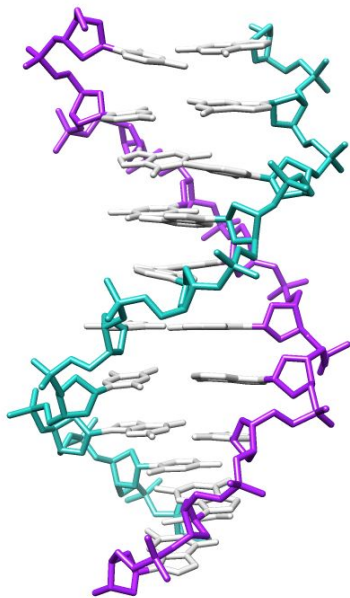
DNA conformations

Canonical forms
B-DNA
A-DNA
Z-DNA

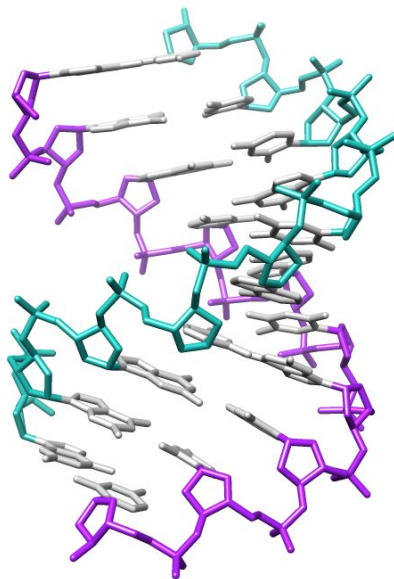
Non canonical forms
Triple helix
G-Quadruplexes

Canonical forms

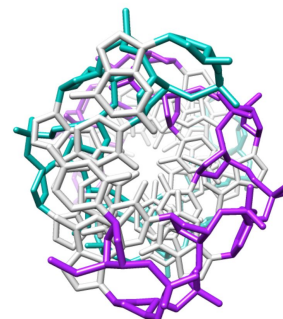
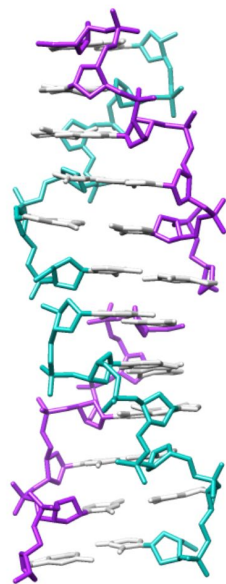
Parameter	B-DNA	A-DNA	Z-DNA
Helix sense	Right	Right	Left
Residue per turn	10 (10,5)	11	12
Axial rise (Å)	3,4	2,55	3,7
Helix pitch (Å)	34	28	45
Base per tilt (°)	-6	20	7
Rotation per residue (°)	36	33	-30
Intrastrand phosphate-phosphate distance (Å)	7,0	5,9	7,0/5,9
Diameter of helix (Å)	20	23	18
Glycosidic bond configuration	Anti	Anti	Anti/Syn
Sugar pucker	C2' endo	C3' endo	C2' endo/C3' endo



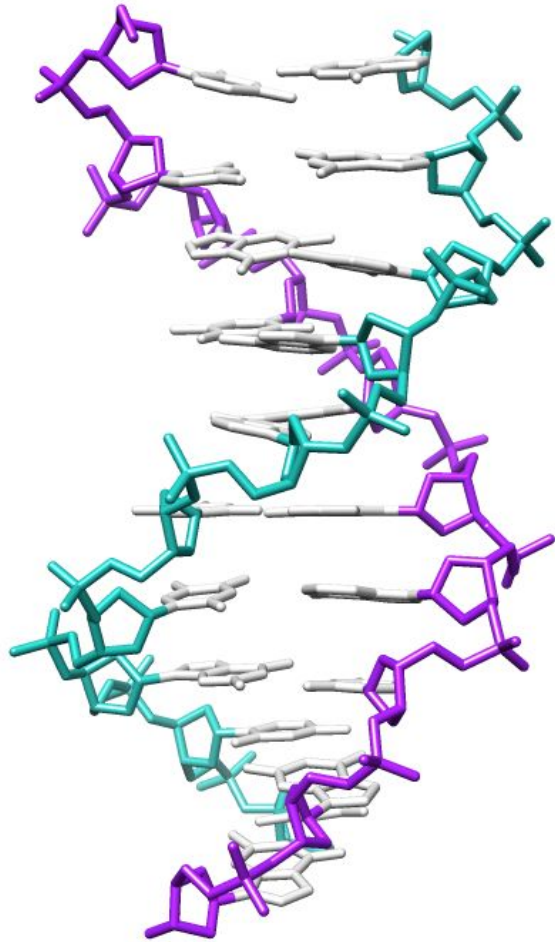
B-DNA
PDBid: 1ZF0



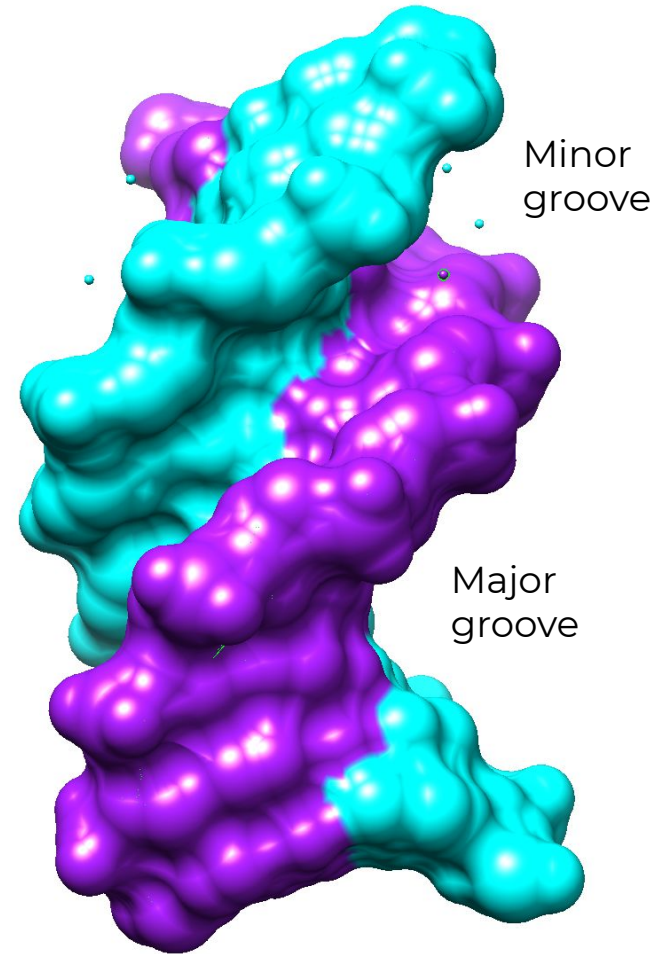
A-DNA
PDBid: 4IZQ



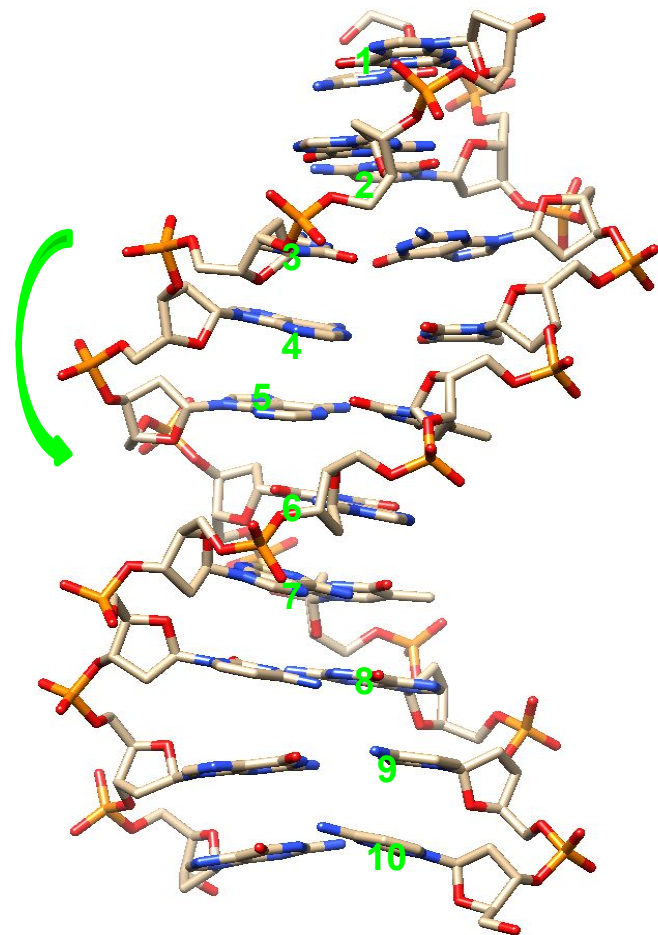
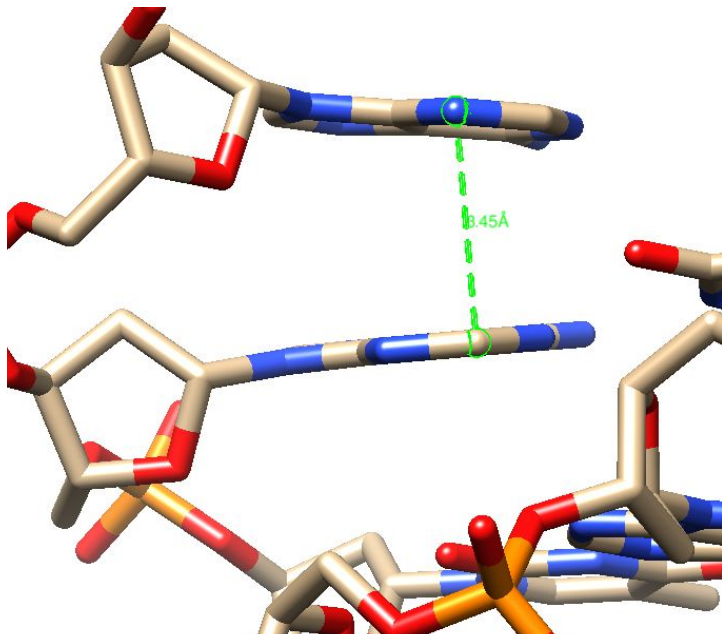
Z-DNA
PDBid: 2IE1



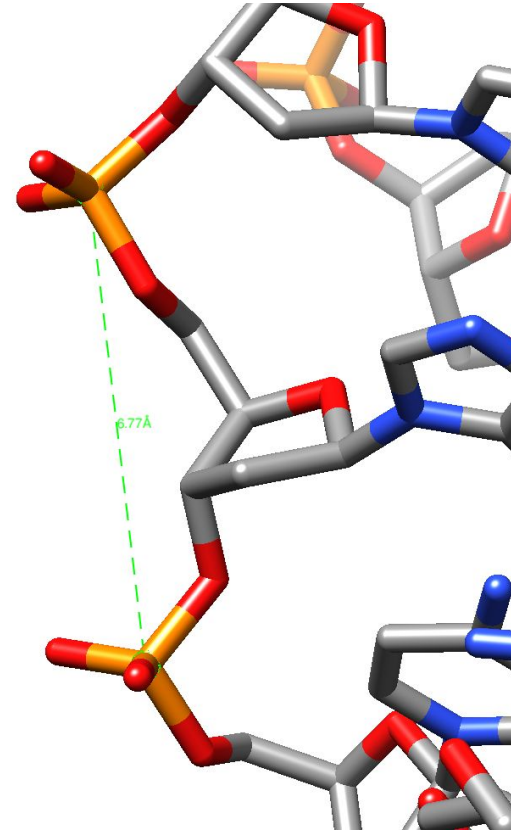
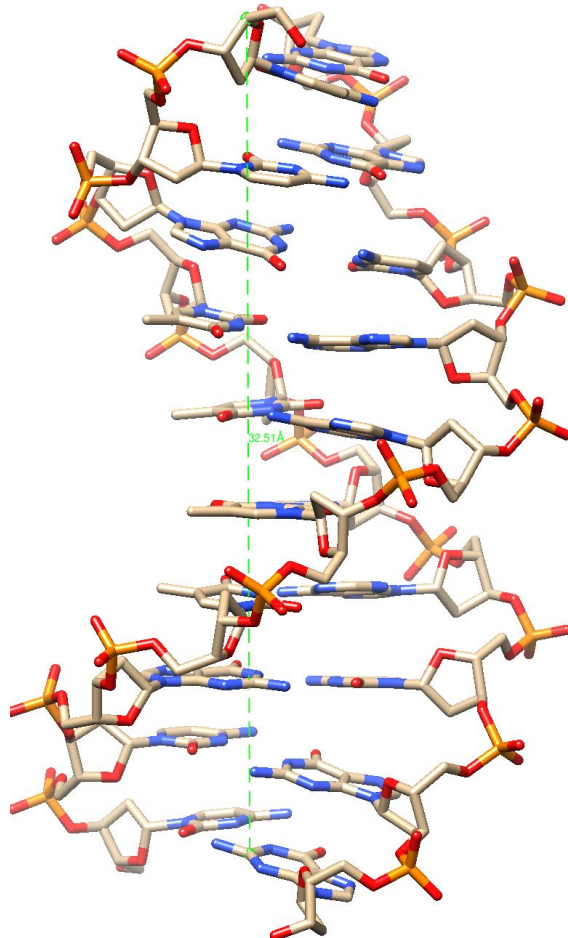
B-DNA



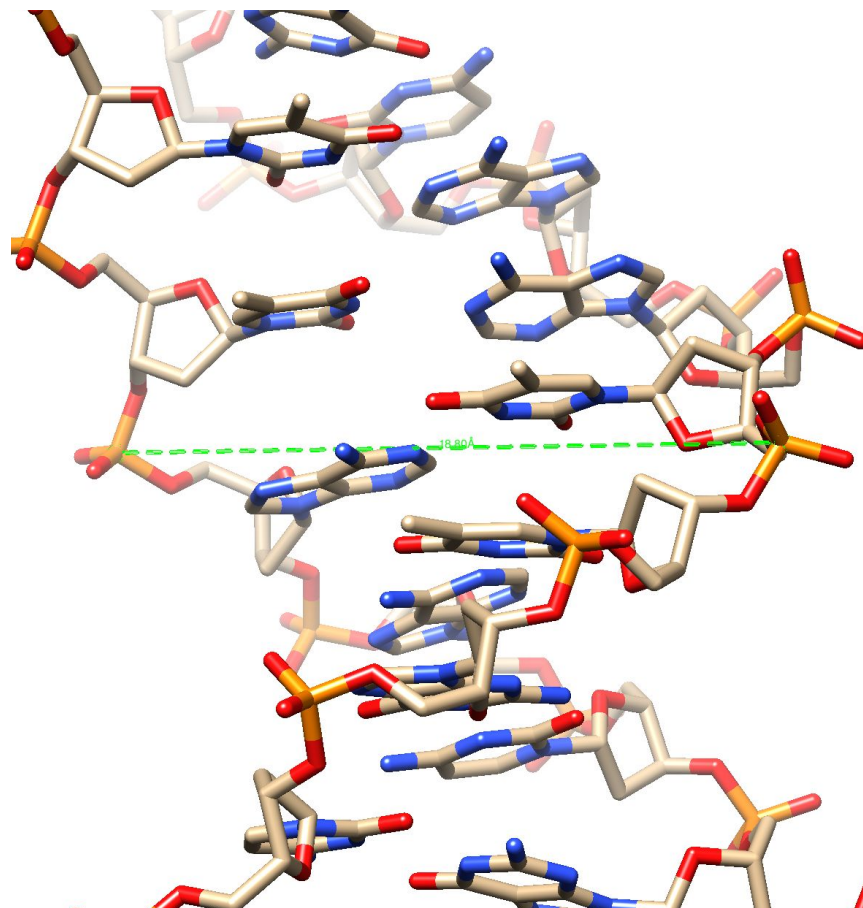
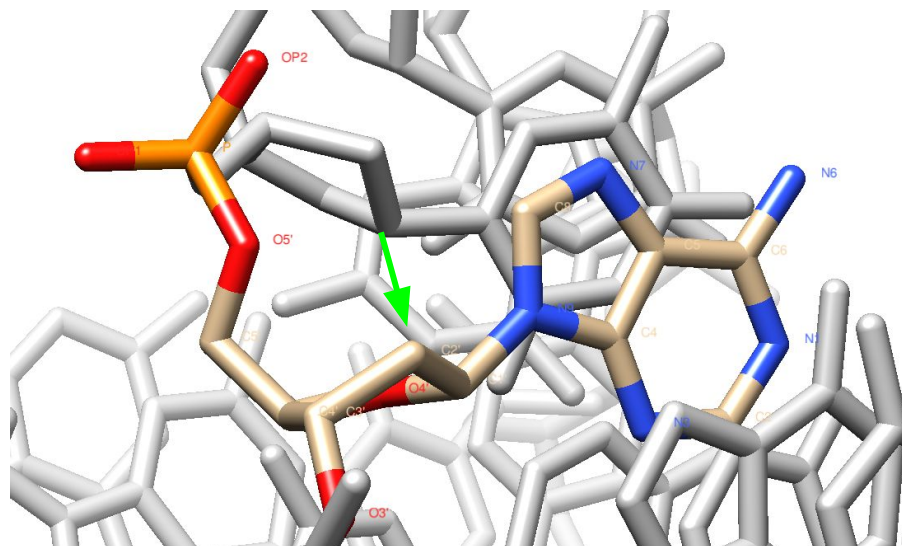
Helix sense	Right
Residue per turn	10 (10,5)
Axial rise (Å)	3,4 → 3,45

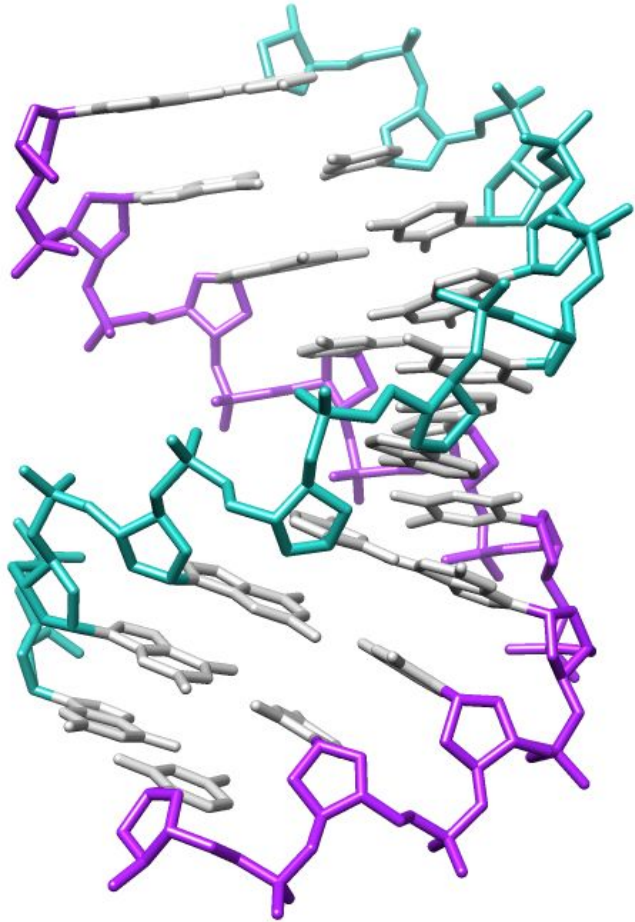


Helix pitch (Å)	34 → 32,51
Base per tilt (°)	-6
Rotation per residue (°)	36
Intrastrand phosphate-phosphate distance (Å)	7,0 → 6,77

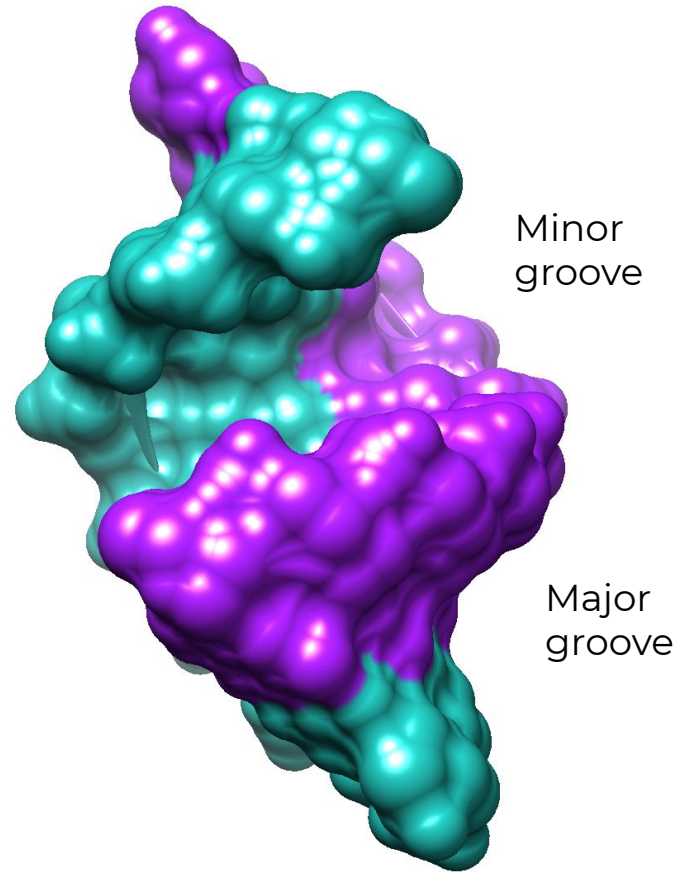


Diameter of helix (Å)	20 → 18.80
Glycosidic bond configuration	Anti
Sugar pucker	C2' endo

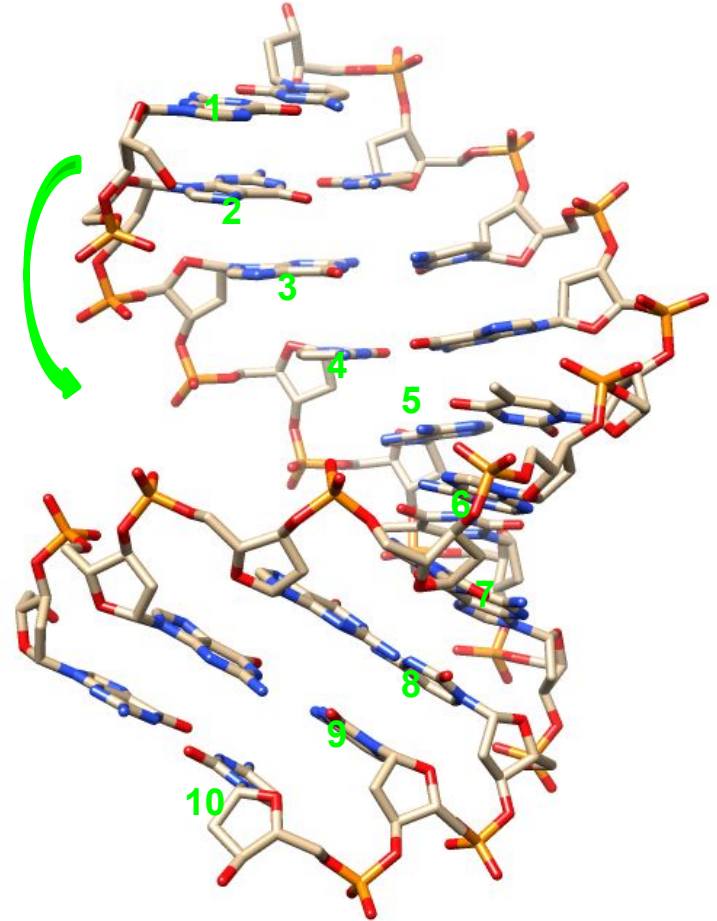
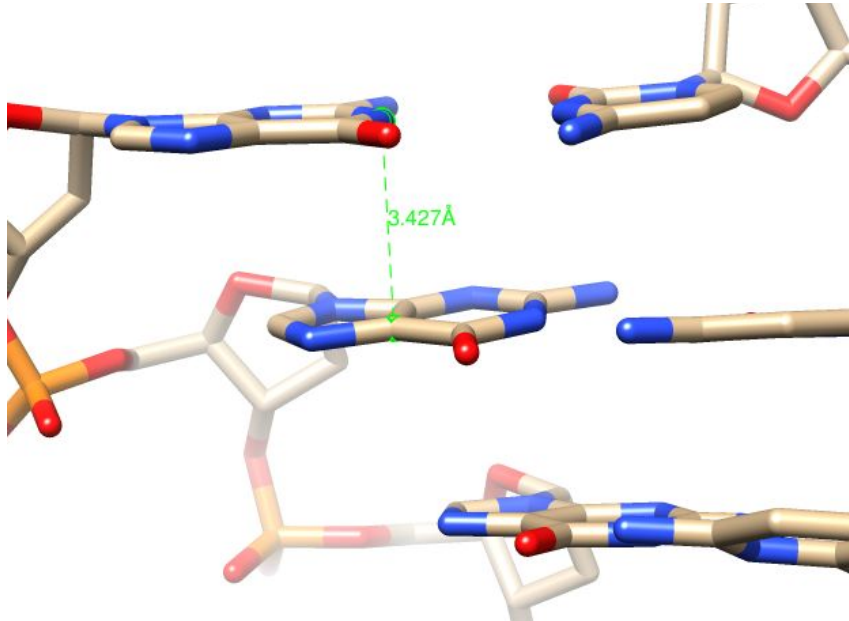




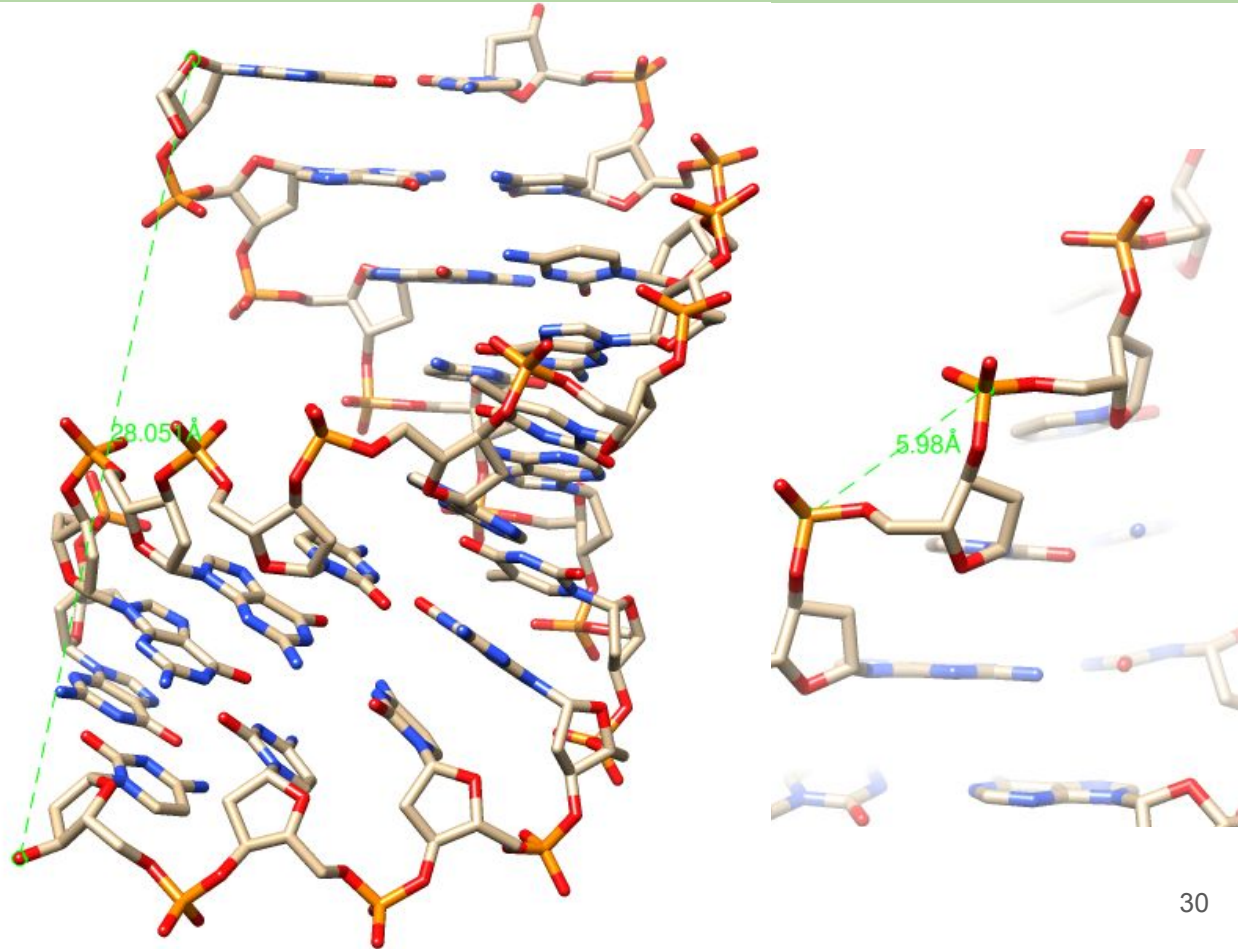
A-DNA



Helix sense	Right
Residue per turn	11
Axial rise (Å)	2,55 → 3,4



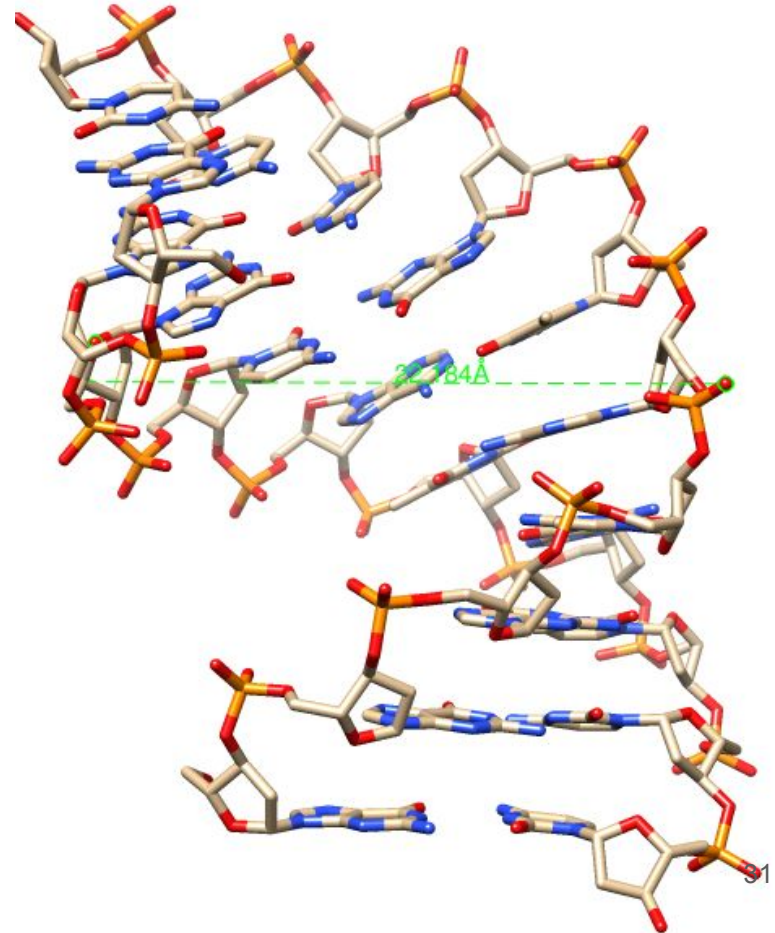
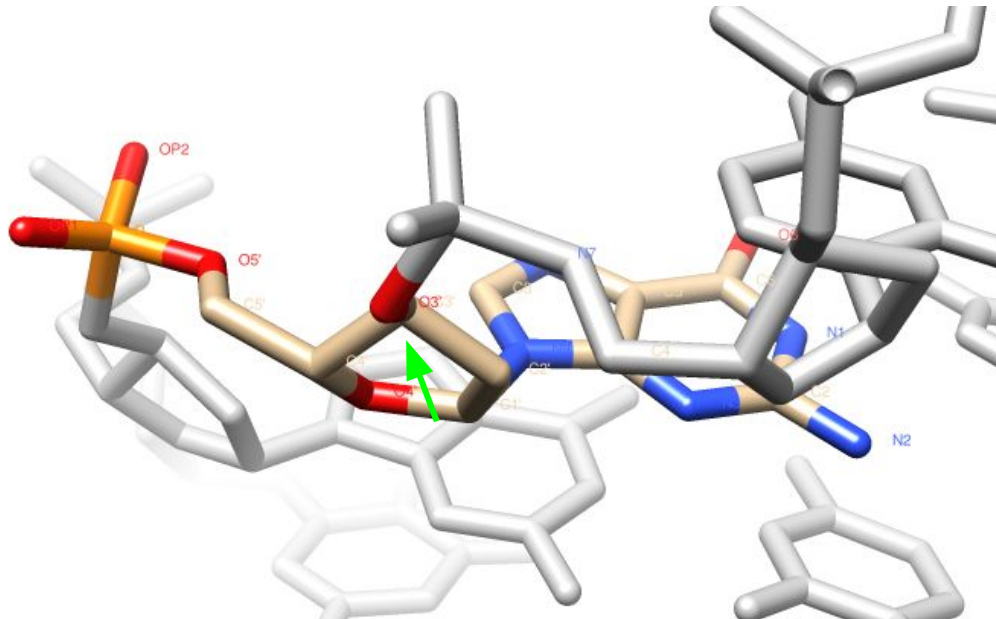
Helix pitch (Å)	28 → 28,05
Base per tilt (°)	20
Rotation per residue (°)	33
Intrastrand phosphate-phosphate distance (Å)	5,9 → 5,98

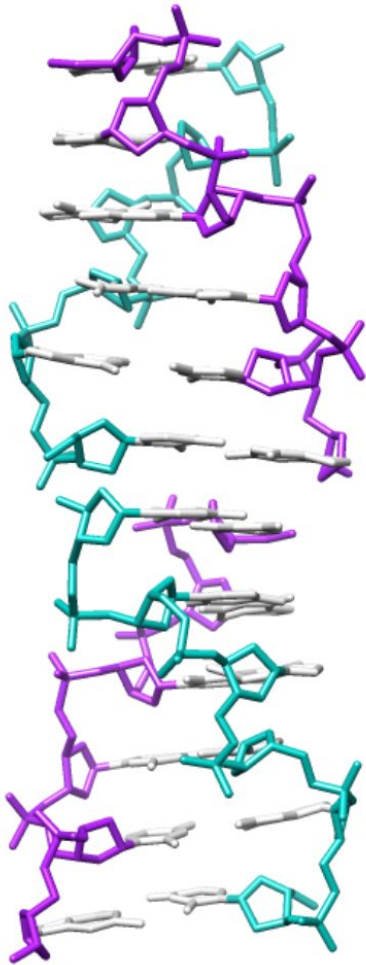


A-DNA

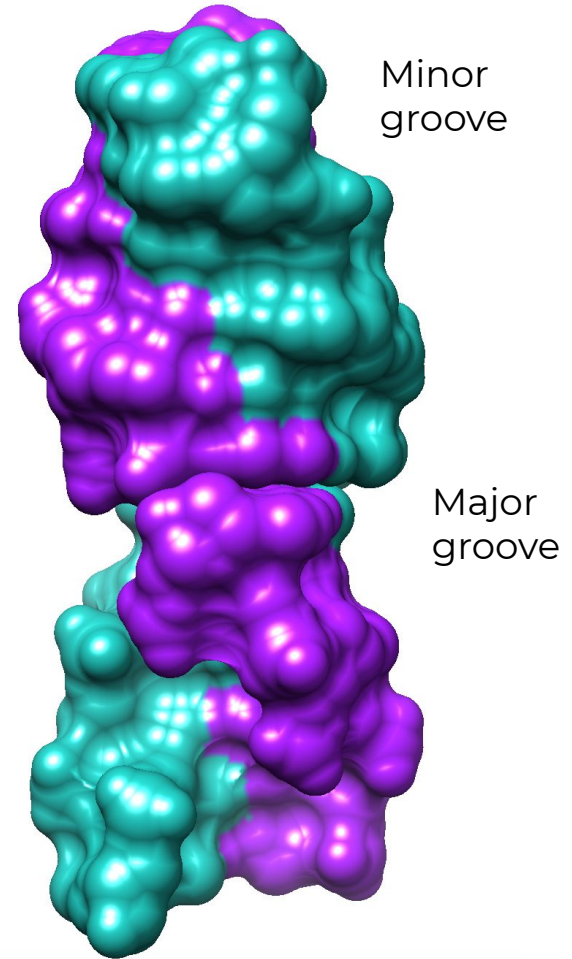
PDBid: 4IZQ

Diameter of helix (Å)	23 → 22,18
Glycosidic bond configuration	Anti
Sugar pucker	C3' endo

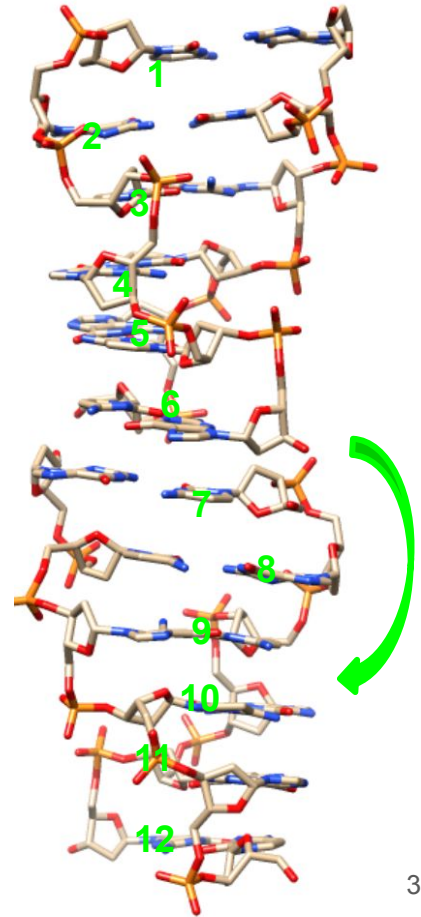
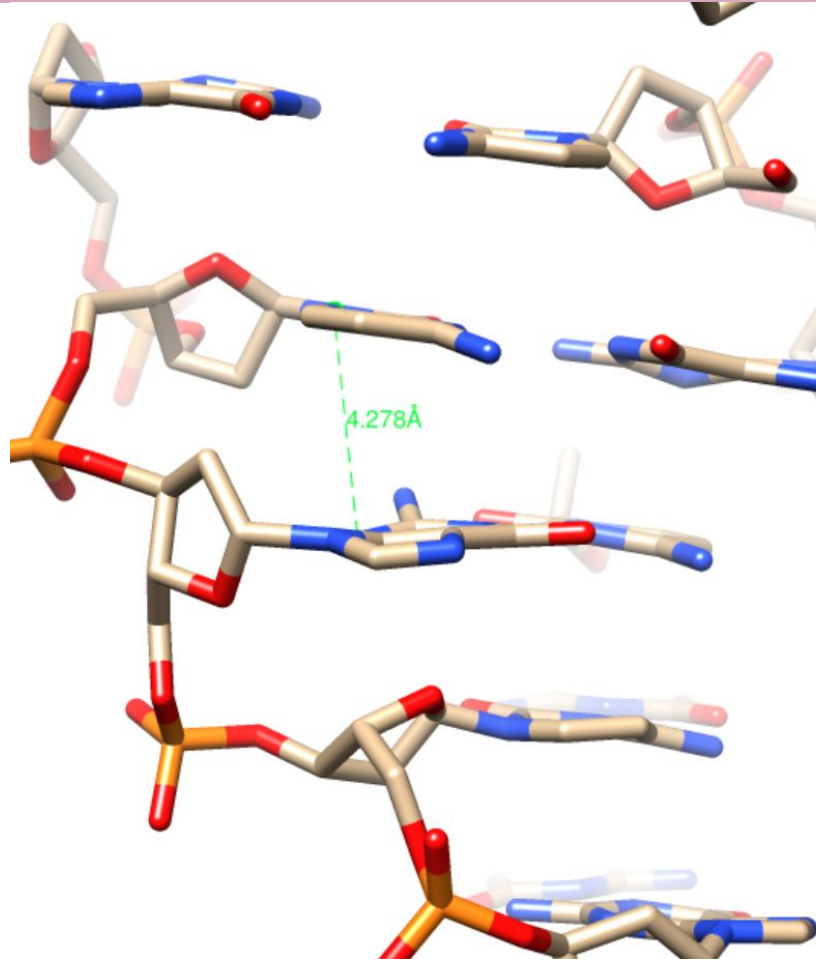




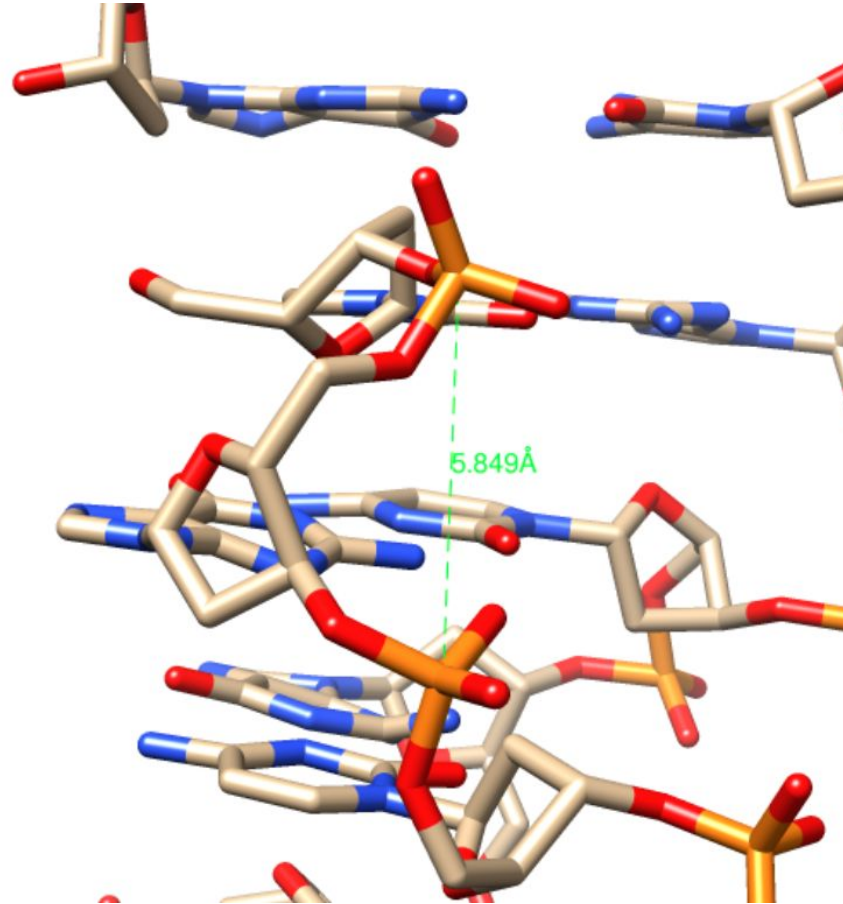
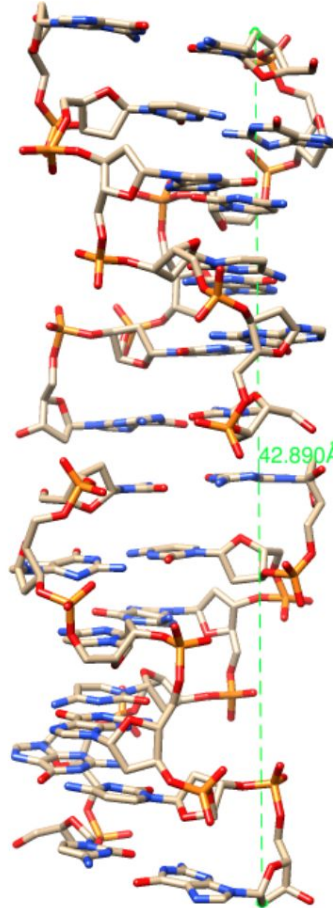
Z-DNA



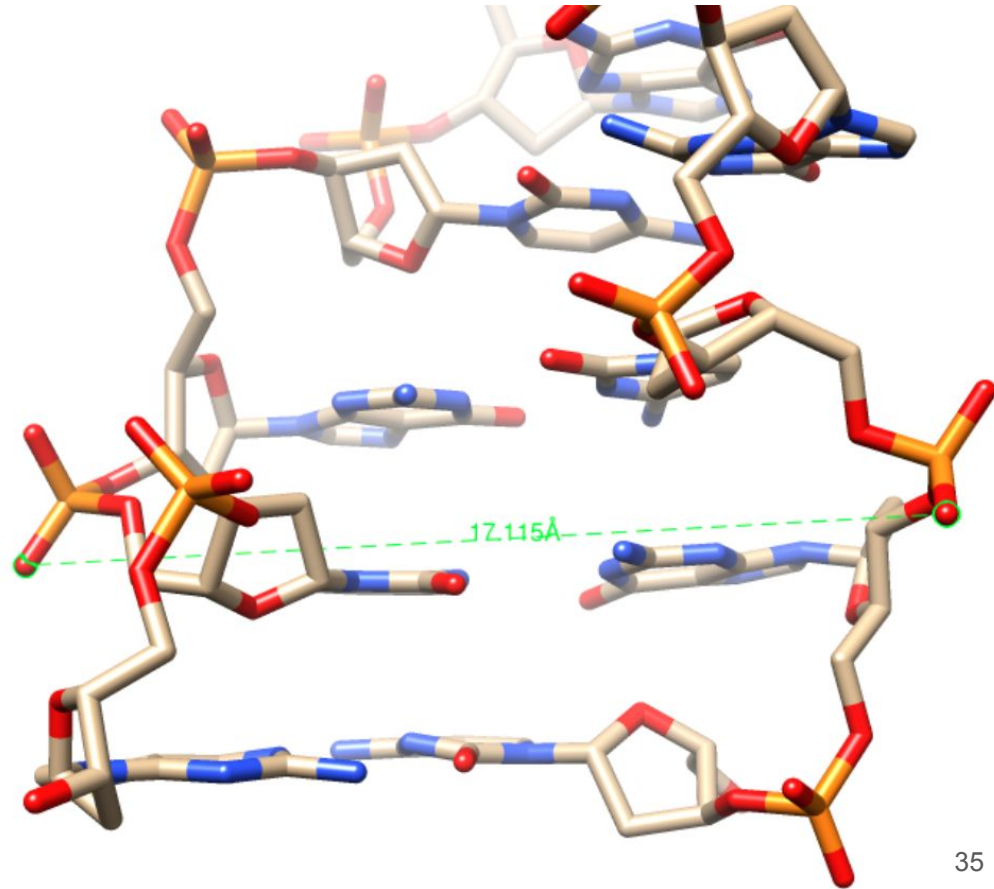
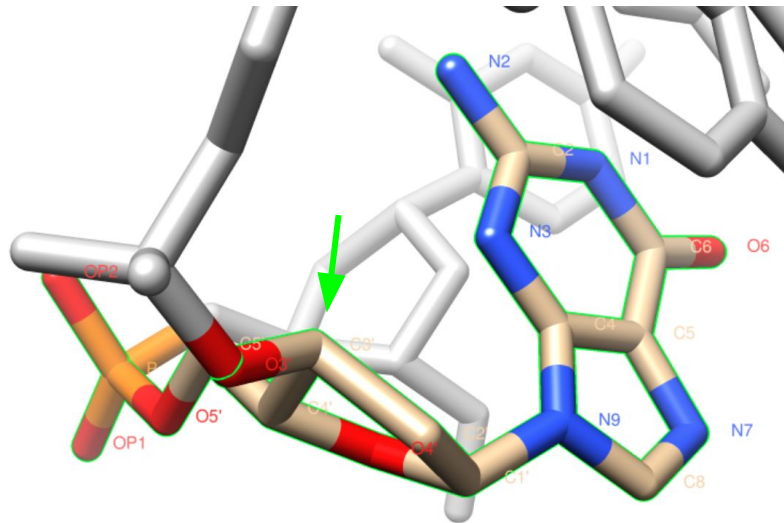
Helix sense	Left
Residue per turn	12
Axial rise (Å)	3,7 → 4,27



Helix pitch (Å)	45 → 42,89
Base per tilt (°)	7
Rotation per residue (°)	-30
Intrastrand phosphate-phosphate distance (Å)	7,0/5,9



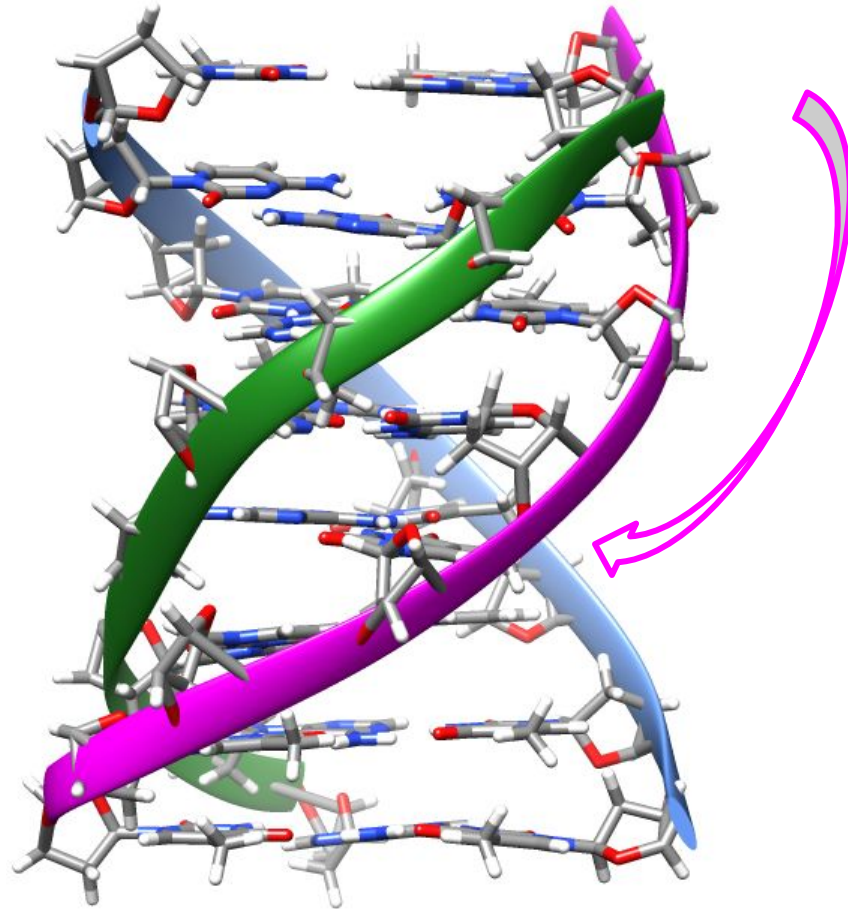
Diameter of helix (Å)	18 → 17,12
Glycosidic bond configuration	Anti/ Syn
Sugar pucker	C2' endo/ C3' endo



Non-canonical forms

Non-canonical forms: **Triple helix**

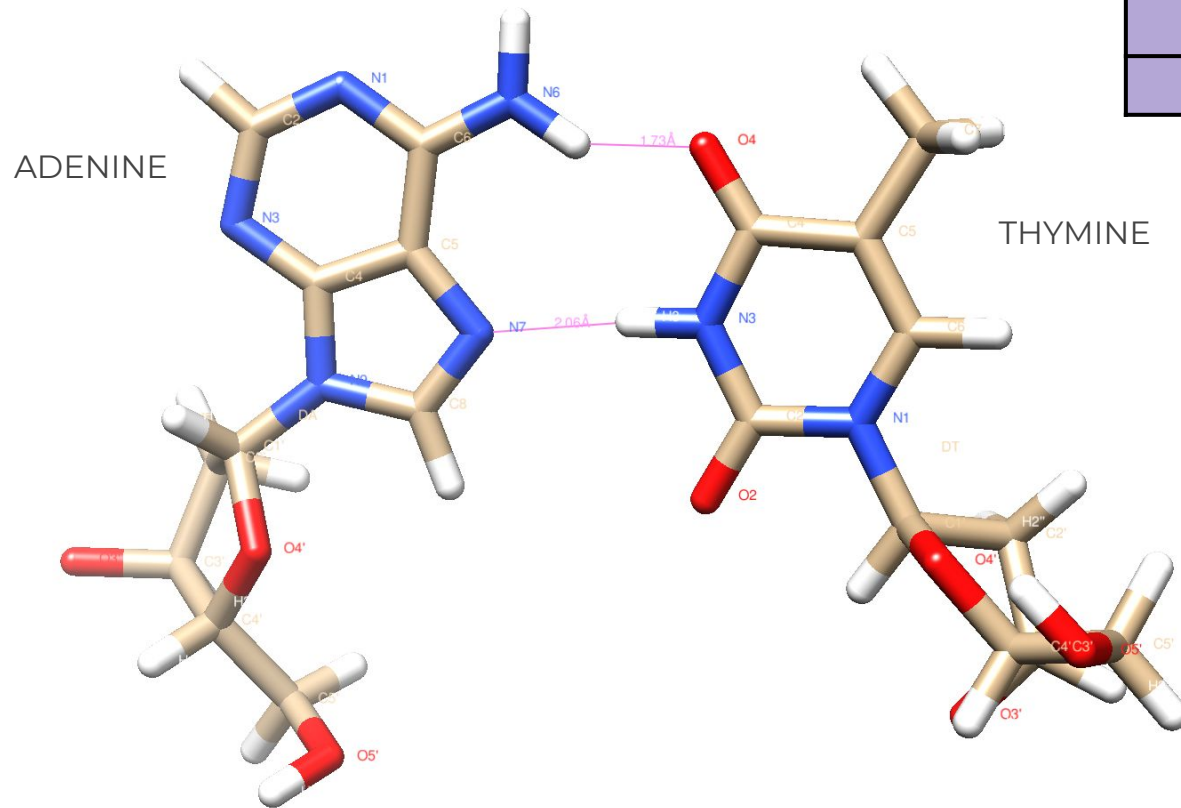
PDBid: 1D3X



Major groove

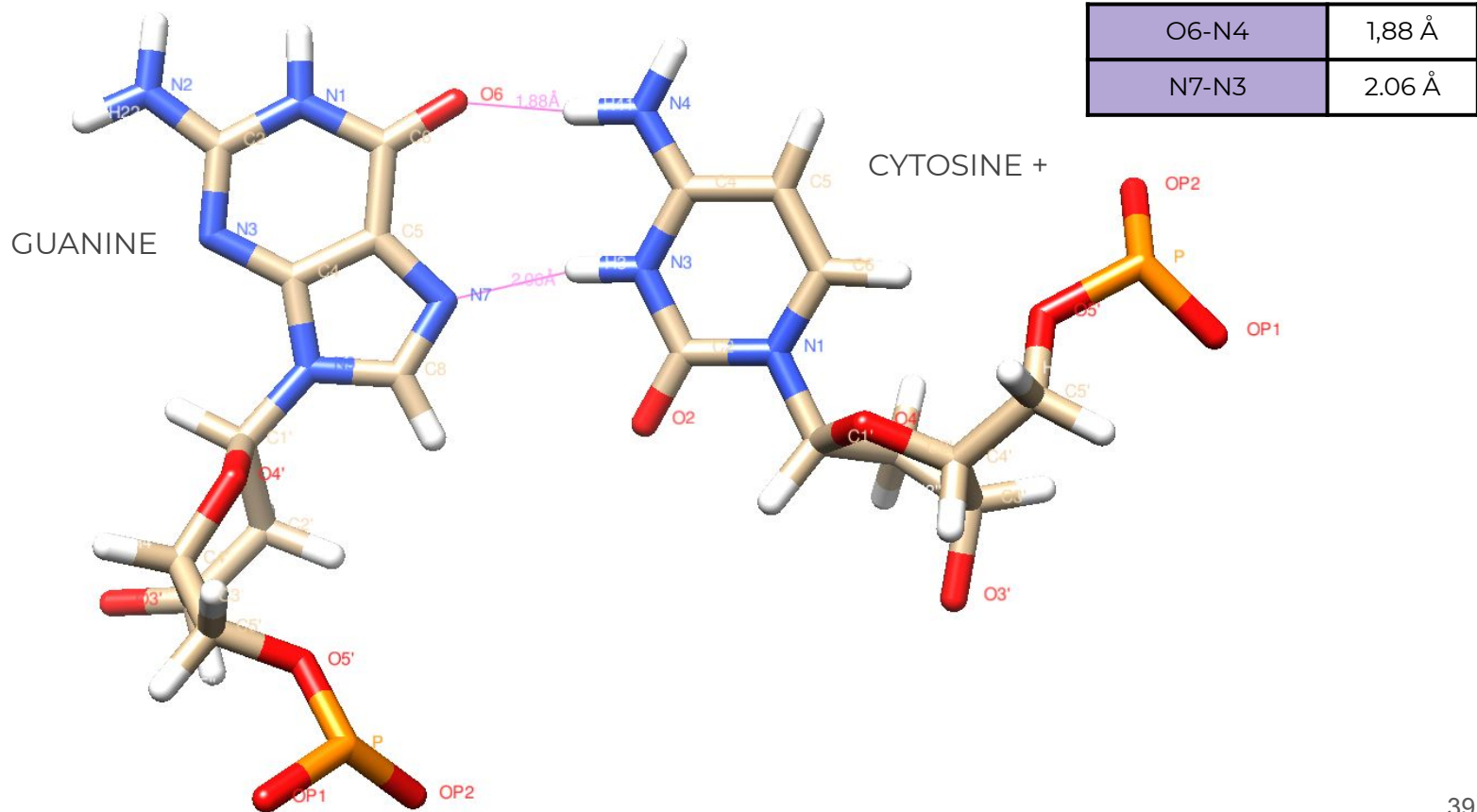
Triple helix: Hoogsteen hydrogen bonds

PDBid: 1D3X



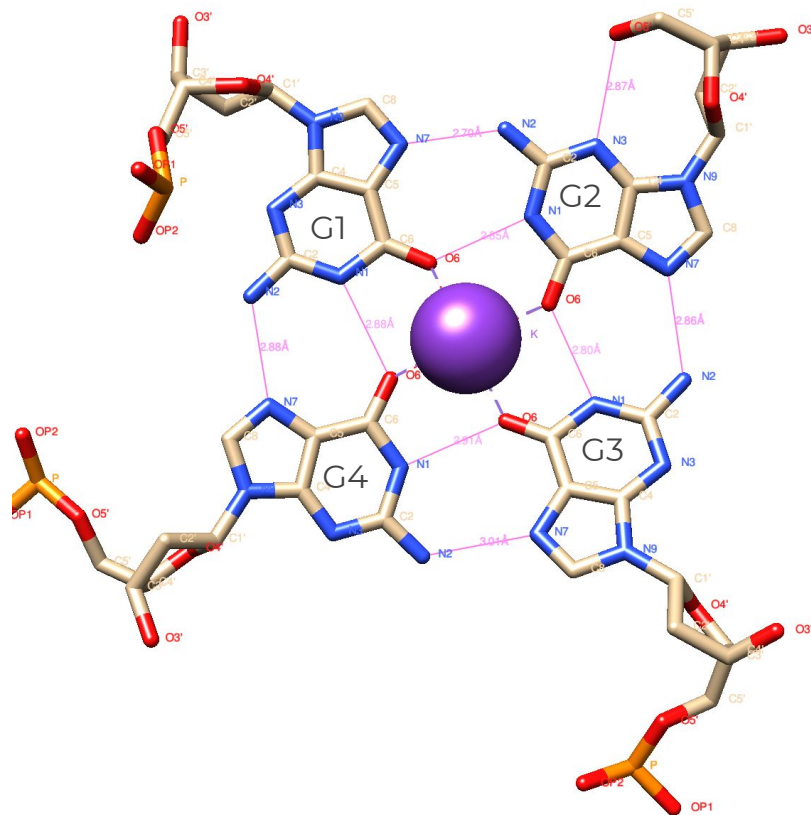
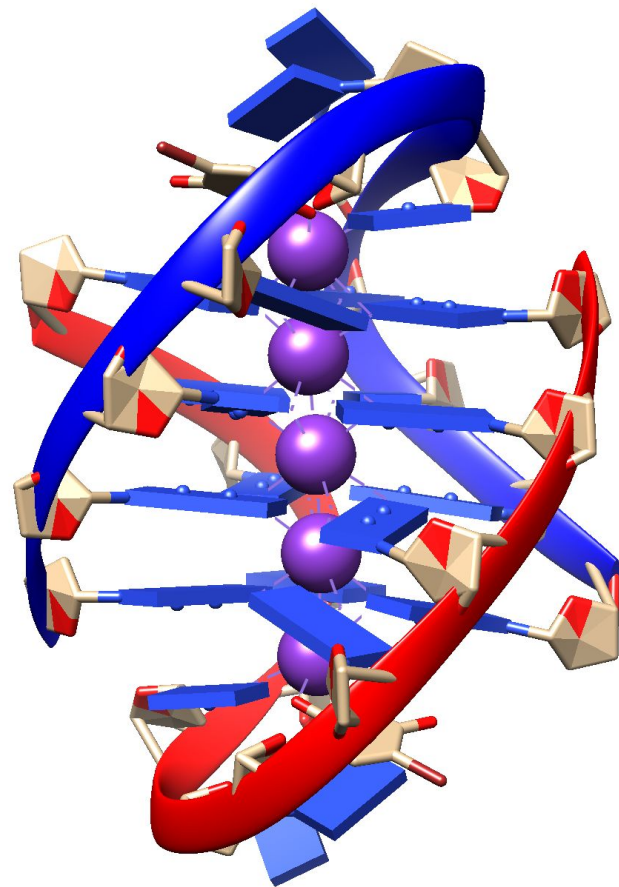
Triple helix: Hoogsteen hydrogen bonds

PDBid: 1D3X

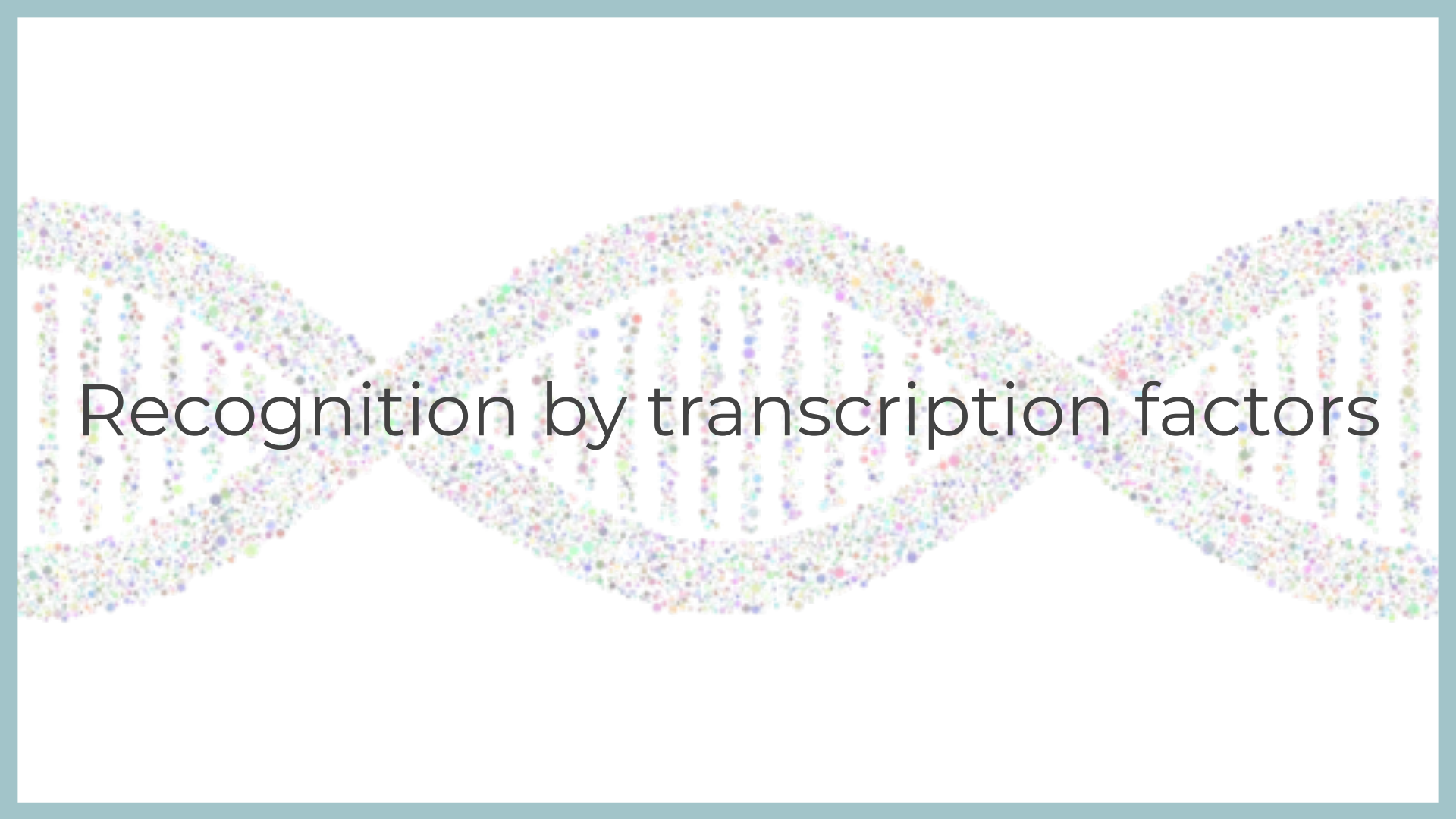


Non-canonical forms: G-quadruplexes / G-tetrad

PDBid: 1JPQ

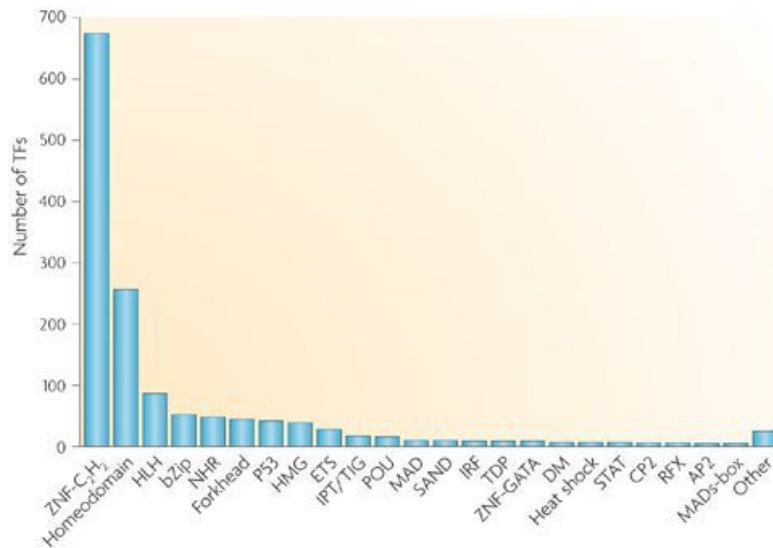
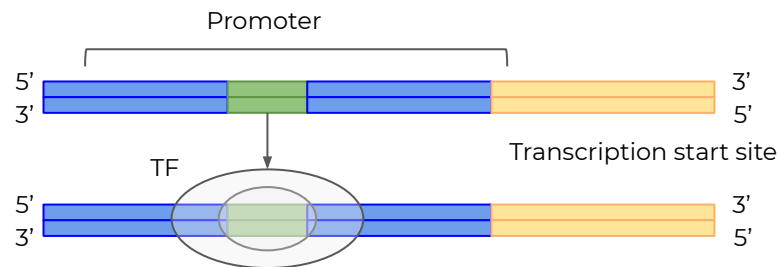


G1-G2 (N7-N2)	2,70 Å
G1-G2 (O6-N1)	2,65 Å
G2-G3 (N7-N2)	2,86 Å
G2-G3 (O6-N1)	2,80 Å
G3-G4 (N7-N2)	3,01 Å
G3-G4 (O6-N1)	2,91 Å
G4-G1 (N7-N2)	2,88 Å
G4-G1 (O6-N1)	2,88 Å



Recognition by transcription factors

Transcription factors



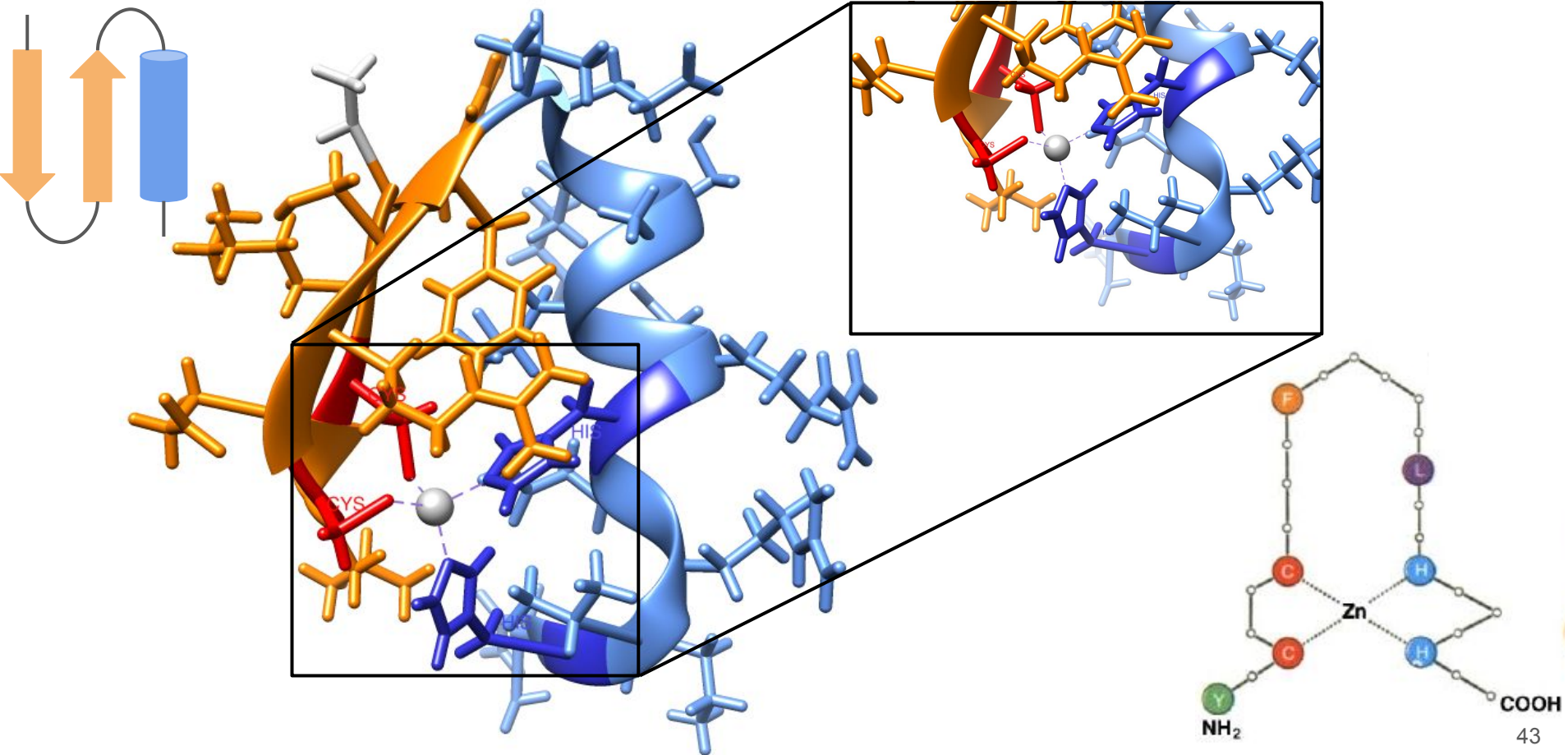
From Vaquerizas JM et al. A census of human transcription factors



Class	Small proteins
Fold	Beta-beta-alpha zinc fingers
Superfamily	Beta-beta-alpha zinc fingers
Family	Classic zinc finger C2H2

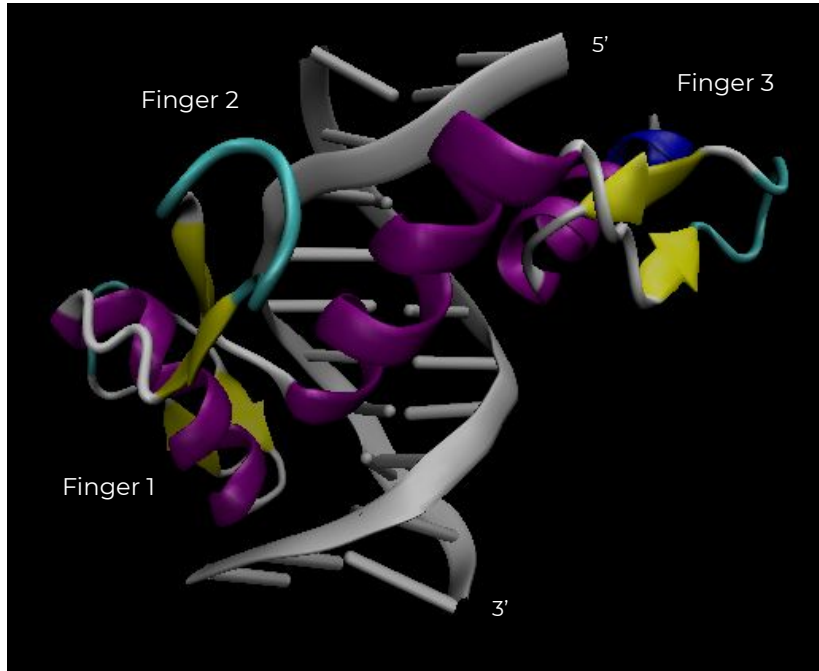
Transcription factors

PDBid: 1ZNF

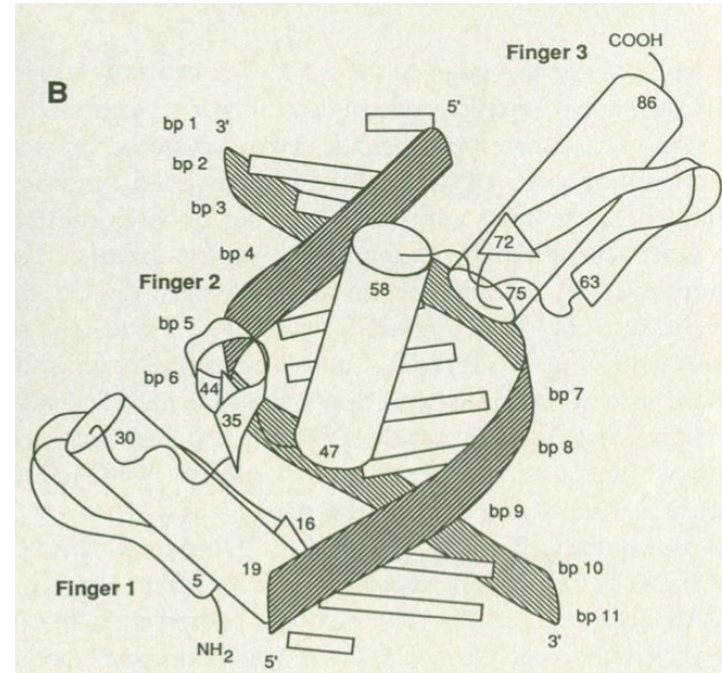


PDBid: 1ZAA



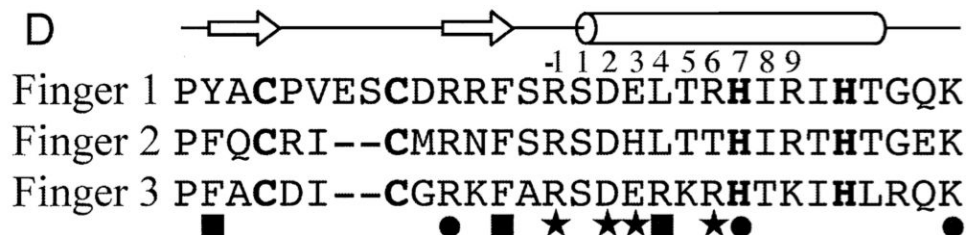
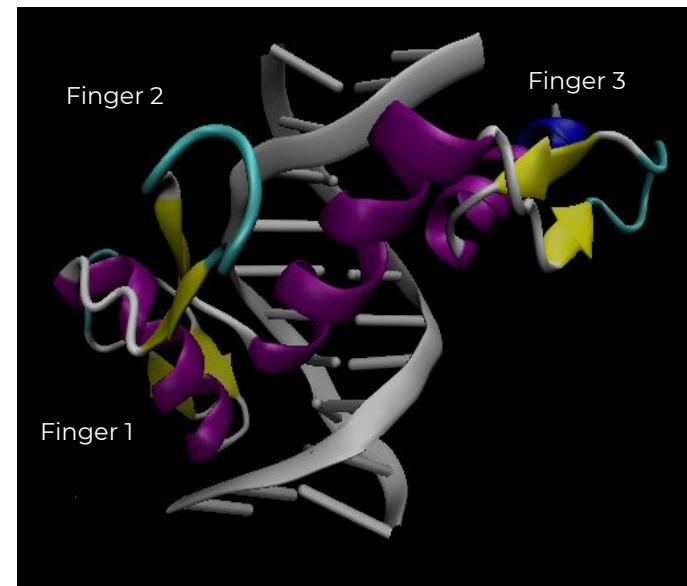
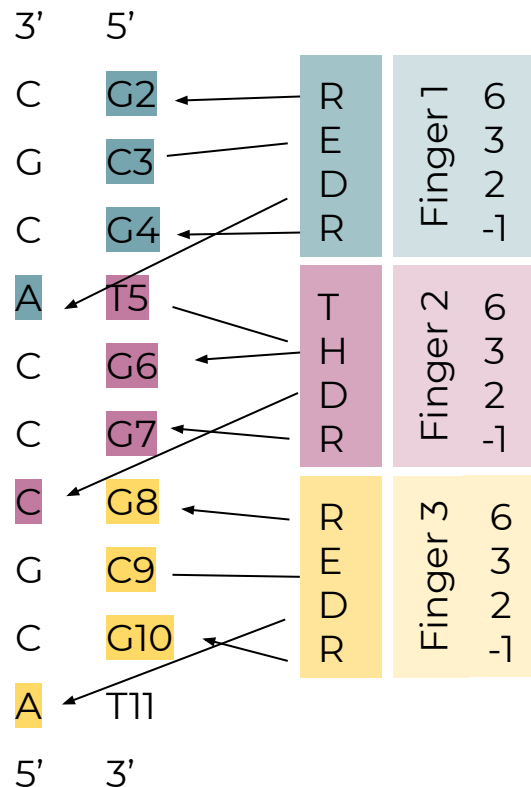


The three zinc fingers are arranged in a semicircular, C-shaped structure.



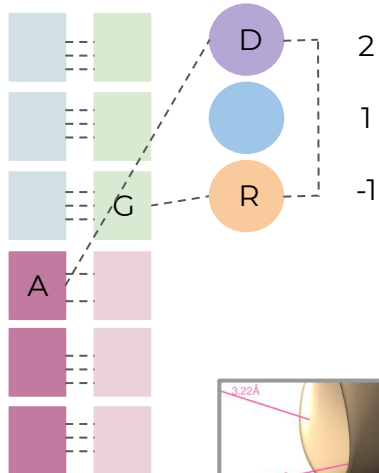
From Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å.

Each zinc finger domain consists of an antiparallel β sheet and an α helix, held together by a Zn ion.

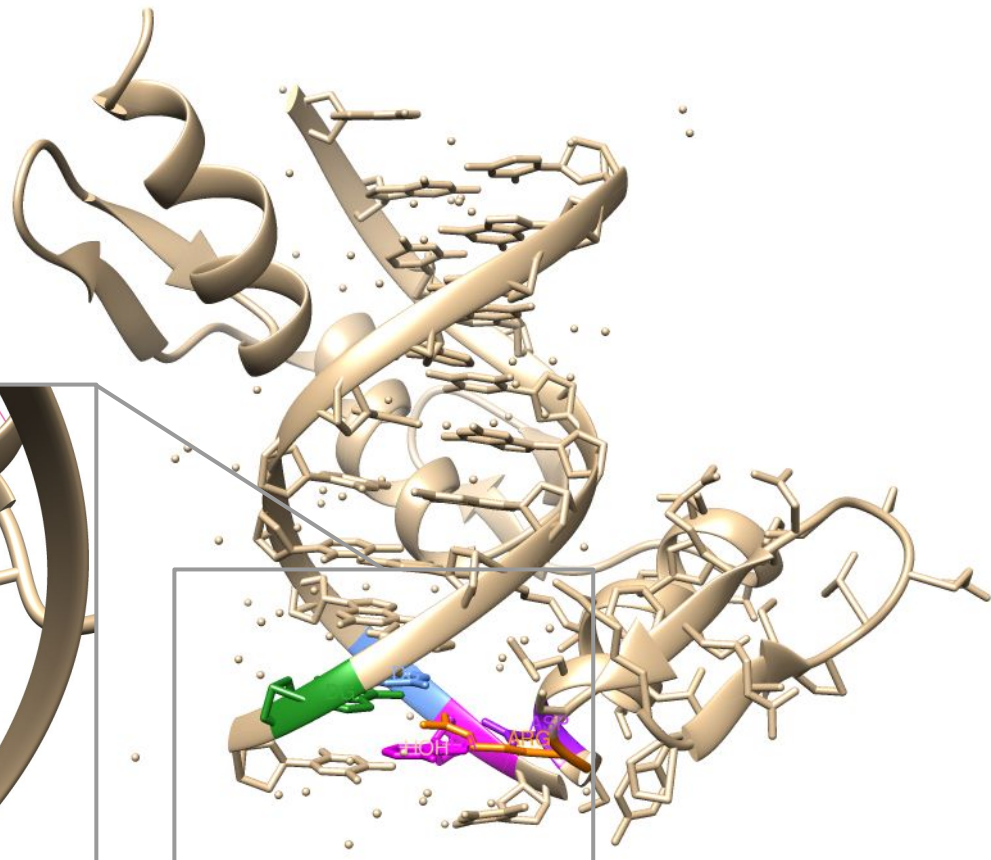
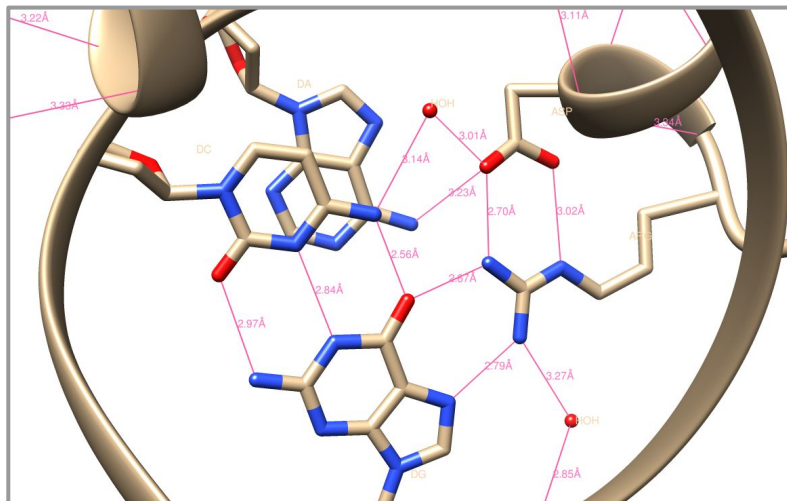


Positions -1 and 2 (all fingers)

PDBid: 1ZAA

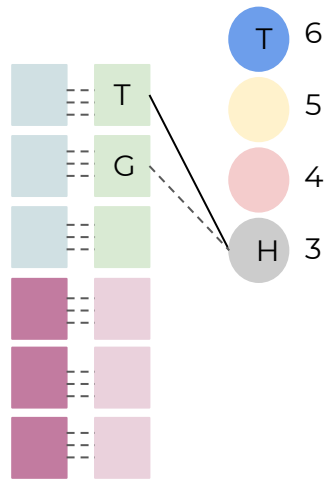


ARG(Nε)-ASP(O)	3.02 Å
ARG(NH2)-ASP(O)	2.70 Å
ARG(NH2)-DG(O6)	2.67 Å
ARG(NH1)-DG(N7)	2.79 Å
ASP(OD2)-DA(N6)	3.23 Å

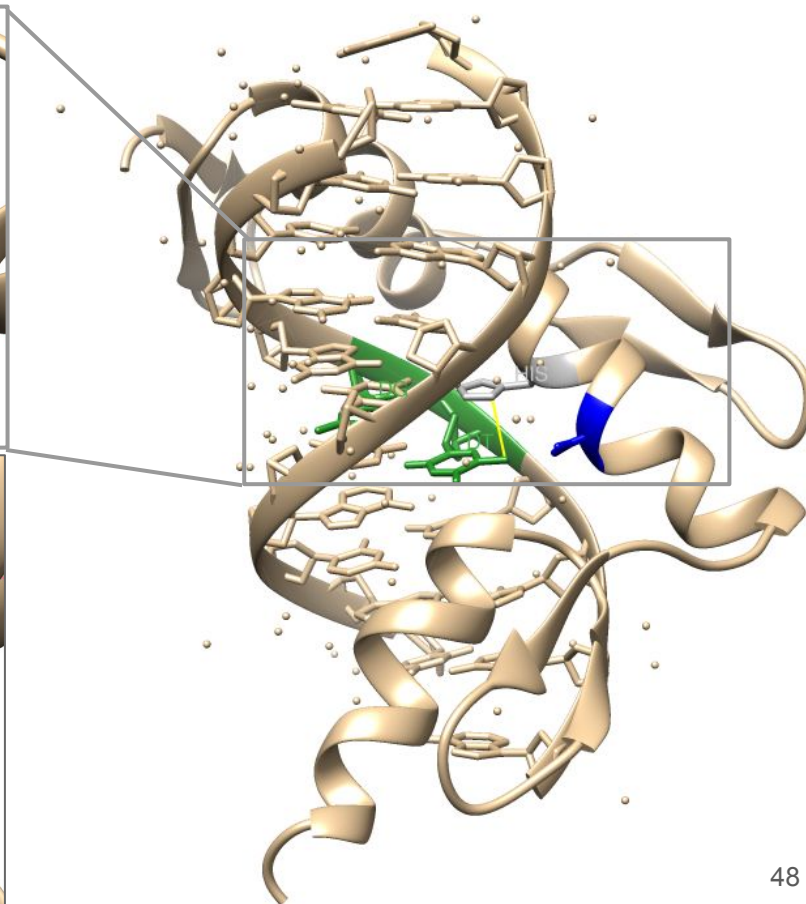
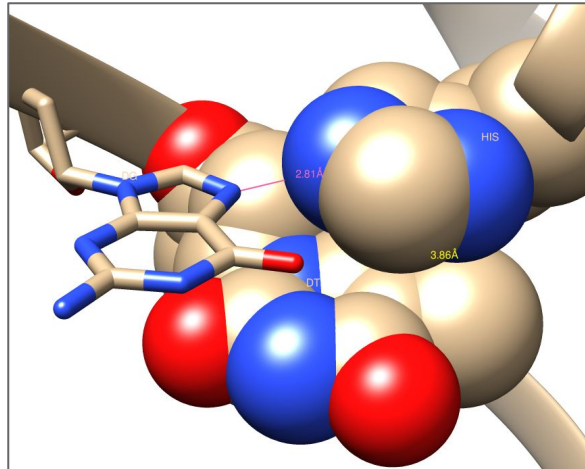
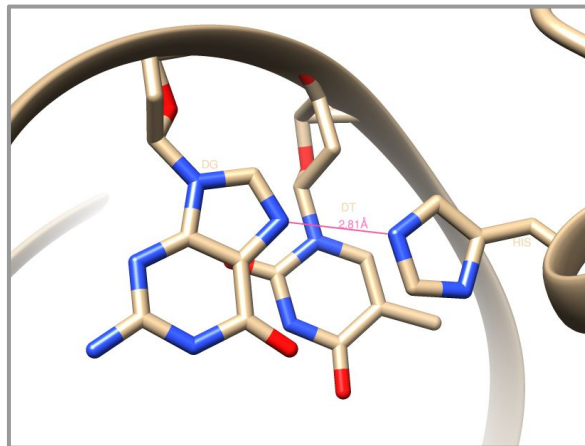


Position 3 (finger 2)

PDBid: 1ZAA

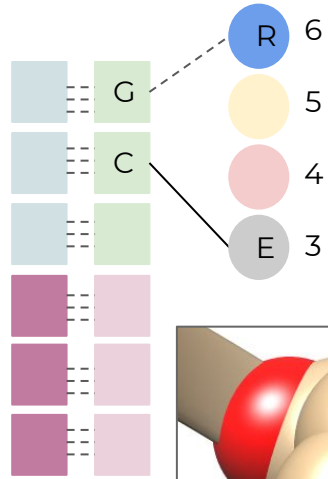


HIS(ϵ)-DG(N7)	2.81 Å
HIS-DT	3.86 Å

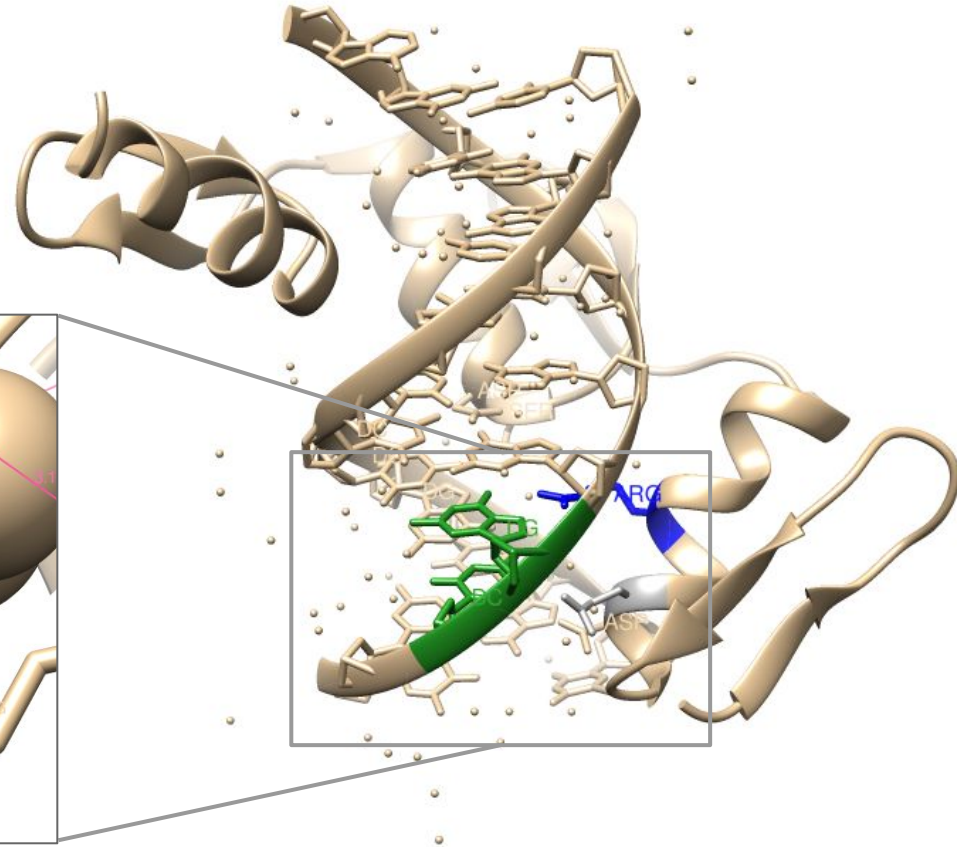
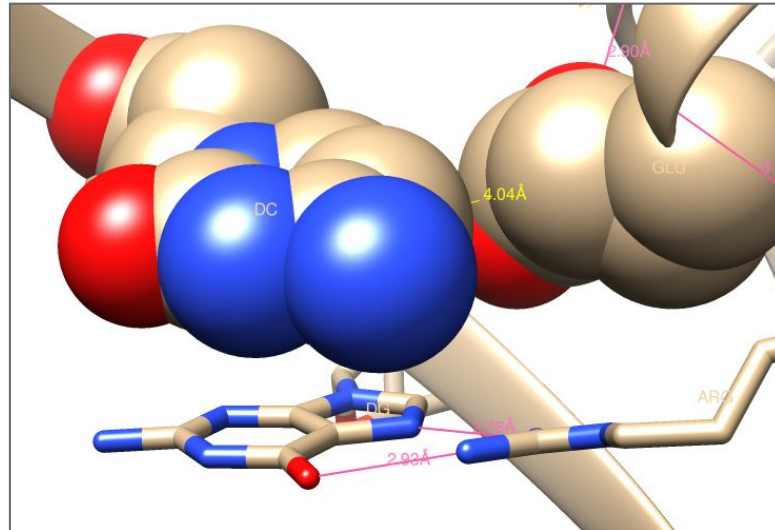


Position 6 (fingers 1 and 3)

PDBid: 1ZAA

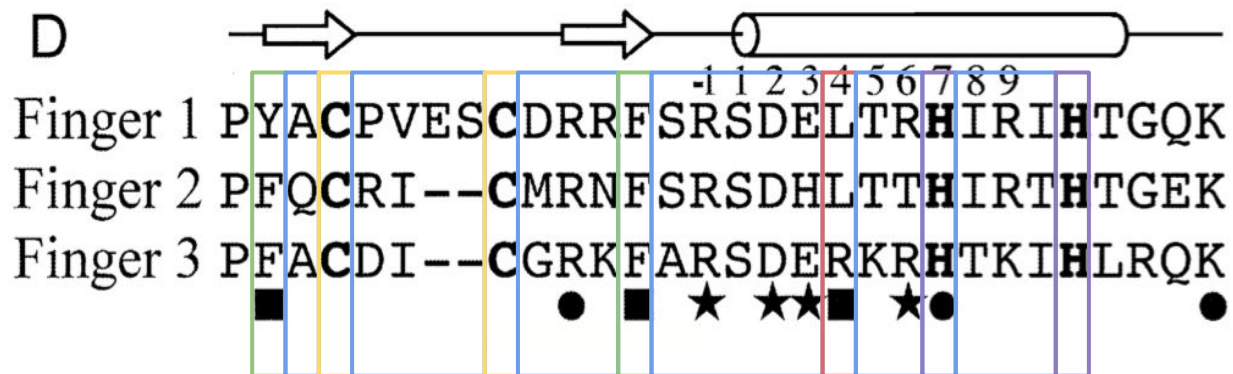


ARG(NH1)-DG(O6)	2.93 Å
ARG(NH2)-DG(N7)	2.78 Å
GLU-DC	4.04 Å



ZNF TF analysis

(F/Y)-X-C-X₂₋₅-C-X₃-(F/Y)-X₅-Ψ-X₂-H-X₃₋₅-H



ZNF TF analysis: CLUSTALW

CLUSTAL 2.1 multiple sequence alignment

(F/Y)-X-C-X₂₋₅-C-X₃-(F/Y)-X₅-ψ-X₂-H-X₃₋₅-H

```
WT1      -----KPYQCDFKDCERRFSRSDHLKRHQR-RHT-GVKP-----FQ-CKTCQRK
ERG1     -----ERPYACPVESCDRRFSRSDDELTRHIR-IHT-GQKP-----FQ-CRICMRN
AEBP2    MGHHHHHHNNIAYNCWDQCQACFNSSPDLADHIRSIHVDGQRGGVFVCLWKGCKVYNTF
GLI1     -----ETDCRWDCGCSQEFDSQEQLVHHINSEHIHGERK-EFVCHWGGCSRELRF
ZFP91    -----HMRDYICEY--CARAFKSSHNLAVHRM-IHT-GEKP-----LQCEICGFT
          * * * * * * * * * * * * * * * * * * * * * * * * * *
```

```
WT1      FSRSDHLKTHTRHTTGEKPFSCRWPSCQKKFARSDELVRHHNMHQR-----
ERG1     FSRSDHLTTHIRHTTGEKPFAC--DICGRKFARSDEKRHTKIHLR-----
AEBP2    STSQSWLQRHMLTHSGDKPFKCVVGGCNASFASQGGLARHVPTHFSQQN-----
GLI1     FKAQYMLVVHMRRTTGEKPHKCTFEGCRKSYSRLENLKTHLRSHTGEKPYMCEHEGCSKA
ZFP91    CRQKASLNWVHMKKHDAADFYQFSCNICGKKFEKKDSVVAHKAKSHPEV-----
          . * * * * . . . . * . . . . *
```

```
WT1      -----
ERG1     -----
AEBP2    -----
GLI1     FSNASDRAKHQNRTHSNEKPYVCKLPCCTKRYTDPSSLRKHVKTVHG
ZFP91    -----
```


DNA is a flexible structure

PDBids: 1ZAA, 6O3T, 6OD5

1.- Superimposition:
Zif268-DNA + Zif268

2.- Superimposition:
DNA in complex with 3 TFs

DNA as a **flexible** structure



do transcription factors
modify DNA structure?



DNA in complex with Zif268

DNA in complex with Fox-C2

DNA in complex with TCF4

DNA is flexible an TFs
modify its structure



Function

RMSD: 1.127 Å

Conclusions

There are several DNA conformations and they can be divided into canonical and non-canonical forms

In zinc fingers, there are a few of fully conserved residues. The others are highly variable

Zif268 is a C2H2 zinc finger protein. They are common in human genome

Not only TFs change their conformation in complex with DNA, but also DNA structure can be modified in complex with them

Bibliography

- A Jain, G Wang, KM Vasquez. DNA triple helices: Biological consequences and therapeutic potential. *BIOCHIMIE*. 2008; 90: 1117-1130.
- Alberts B, Johnson A, Lewis J, et al. *Molecular Biology of the Cell*. 4th edition. New York: Garland Science; 2002.
- Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO. Zif268 protein-DNA complex refined at 1.6 Å: a model system for understanding zinc finger-DNA interactions. *Structure*. 1996 Oct 15;4(10):1171-80.
- Fedotova AA, Bonchuk AN, Mogila VA, Georgiev PG. C2H2 Zinc Finger Proteins: The Largest but Poorly Explored Family of Higher Eukaryotic Transcription Factors. *Acta Naturae*. 2017 Apr-Jun;9(2):47-58.
- Ferry G. The structure of DNA. *Nature*. 2019 Nov;575(7781):35-36.
- Lodish H, Berk A, Zipursky SL, et al. *Molecular Cell Biology*. 4th edition. New York: W. H. Freeman; 2000.
- Neidle S and Balasubramanian S. *Quadruplex Nucleic Acids*. UK: RSC publishing; 2006.
- Pavletich NP, Pabo CO. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science*. 1991 May 10;252(5007):809-17.
- Razin SV, Borunova VV, Maksimenko OG, Kantidze OL. Cys2His2 zinc finger protein family: classification, functions, and major members. *Biochemistry (Mosc)*. 2012 Mar;77(3):217-26.
- Sinden RR. *DNA structure and function*. UK: Academic Press; 1994.
- T.R. Hughes. *A Handbook of transcription factors*. Canada: Springer; 2011.
- Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. *Nat Rev Genet*. 2009 Apr;10(4):252-63.
- Wolfe SA, Nekludova L, Pabo CO. DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct*. 2000;29:183-212.



Thanks for your attention

Questions?

PEM questions

1. There are helix parameters:
 - a. Helix sense
 - b. Diameter of the helix
 - c. Both of them
 - d. Residues per turn
 - e. All of them**
2. There are canonical conformations of the DNA
 - a. A-DNA
 - b. Z-DNA
 - c. Both of them**
 - d. Triple helix
 - e. All of them
3. We can find a right helix sense in:
 - a. A-DNA
 - b. B-DNA
 - c. Both of them**
 - d. Z-DNA
 - e. All of them
4. The most common TF in humans are:
 - a. Homeodomain
 - b. HLH
 - c. NHR
 - d. C2H2 ZNF**
 - e. bZip
5. The essential residues in the interaction Zif268-DNA are in positions:
 - a. 1, 2, 3, 4
 - b. -1, 3, 4, 6
 - c. -1, 1, 2, 3
 - d. -1, 2, 3, 6**
 - e. 1, -2, -3, -6

PEM questions

6. How many zinc fingers form Zif268?
- Only one
 - Two zinc fingers
 - Three zinc fingers**
 - Two zinc fingers and a bZip
 - Zif268 is not a zinc finger protein
7. Zif268 makes a complex with:
- B-DNA**
 - A-DNA
 - Z-DNA
 - Triple helix DNA
 - Quadruplex DNA
8. Specificity and stability of the triple helix structure is afforded via:
- Hoogsteen hydrogen bonds**
 - Watson-Crick hydrogen bonds
 - Both of them
 - Reverse Hoogsteen hydrogen bonds
 - All of them
9. The two different conformations of the bases through the glycosidic bonds are:
- cis and trans
 - syn and anti**
 - left and right
 - endo and exo
 - none of the above
10. Choose the correct sentences:
- DNA is a flexible structure and this can be related with its function
 - DNA structure is very stable: it can be never modulated
 - TFs can modify the conformation of DNA when they make a complex
 - TFs change their conformation in complex with DNA but the DNA does not:
- 1, 2, 3
 - 1, 3**
 - 2, 4
 - 4
 - 1, 2, 3, 4