

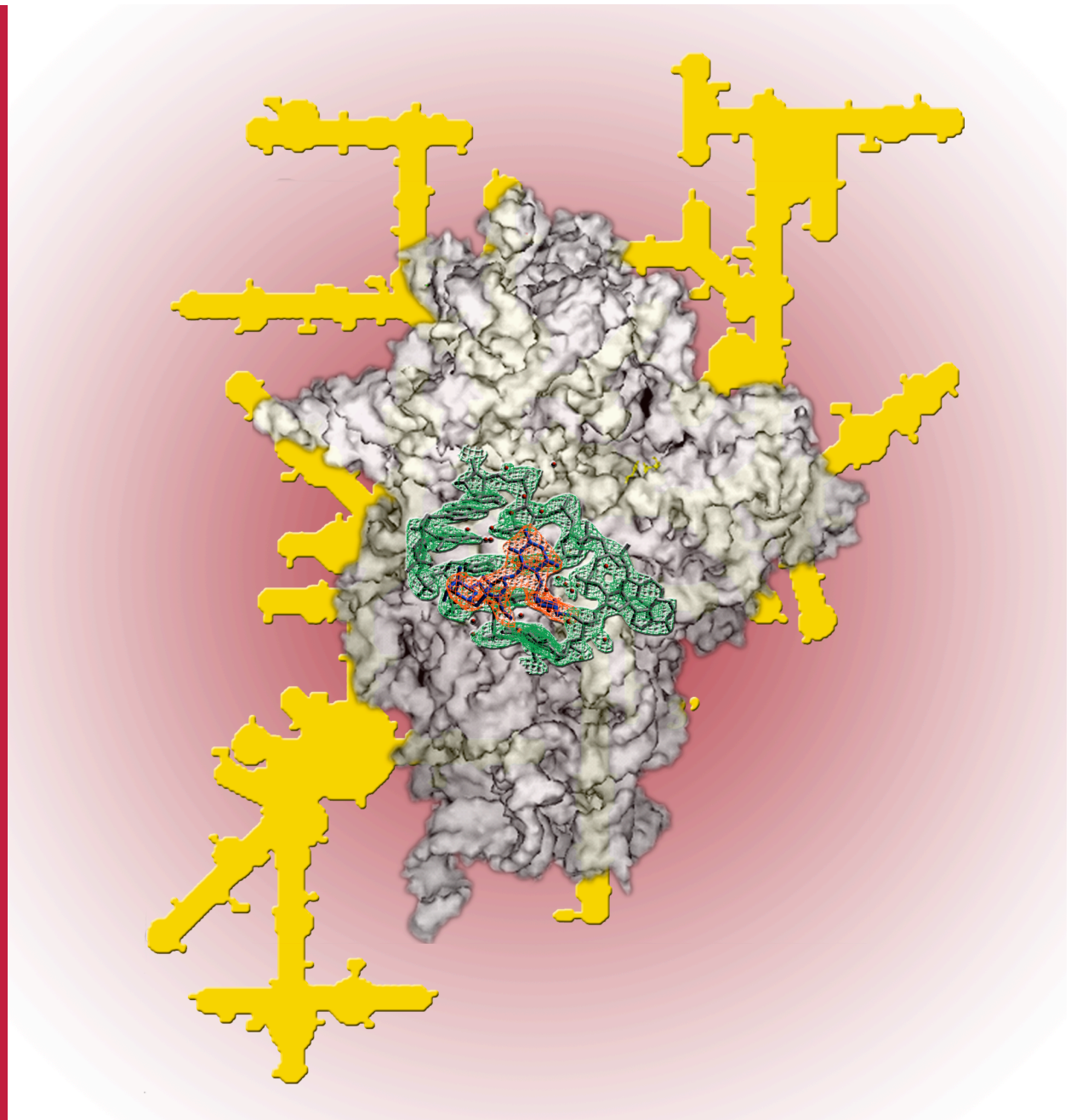
The

RNA

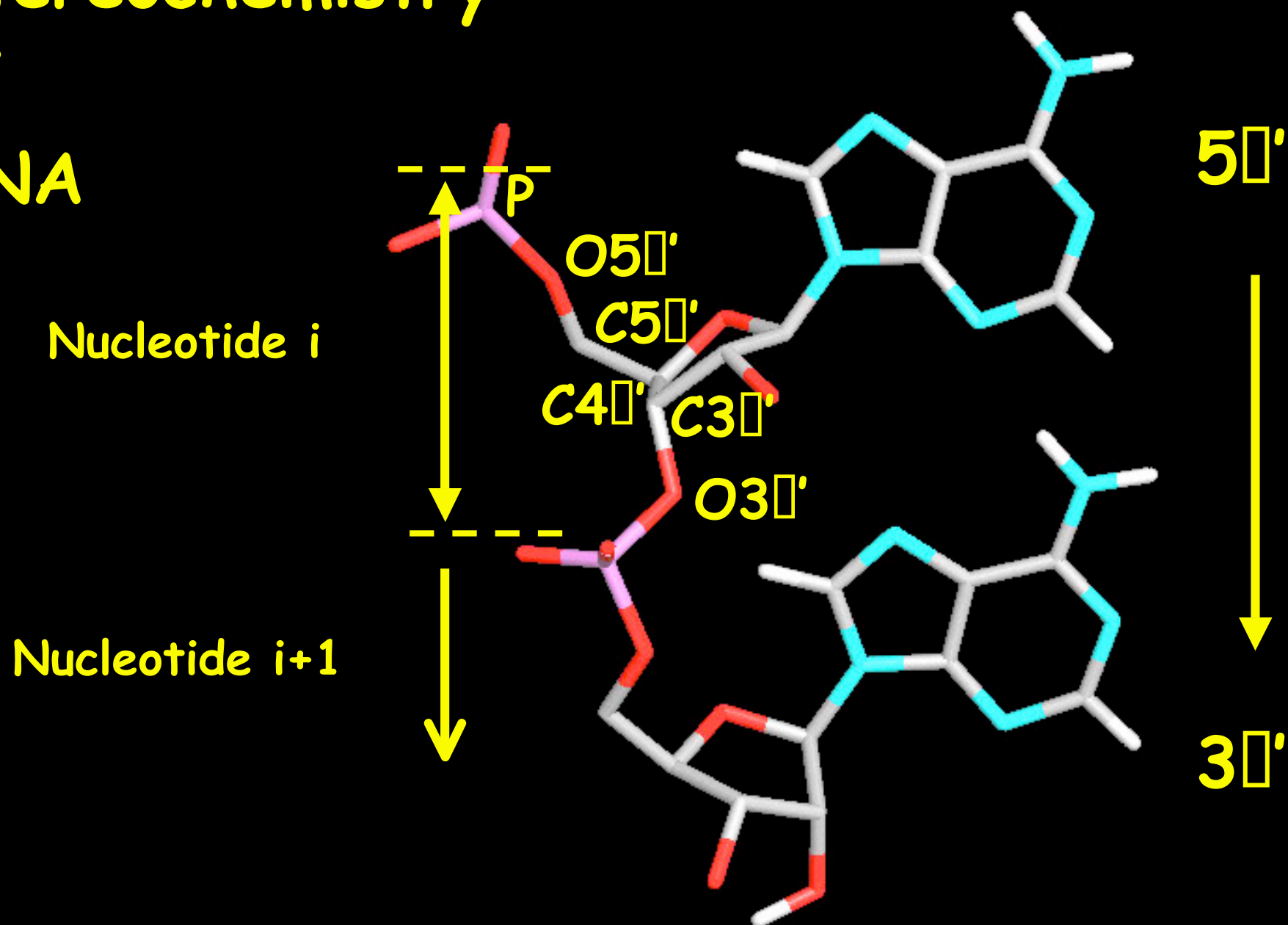
World III

E. Westhof

<http://www-ibmc.u-strasbg.fr/upr9002/westhof/>



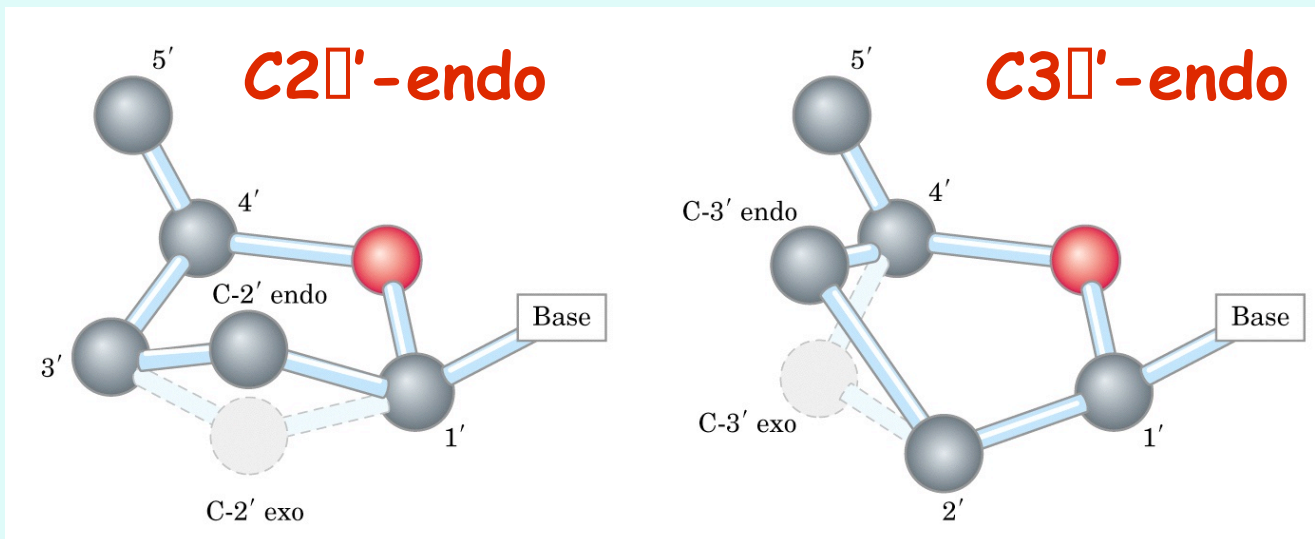
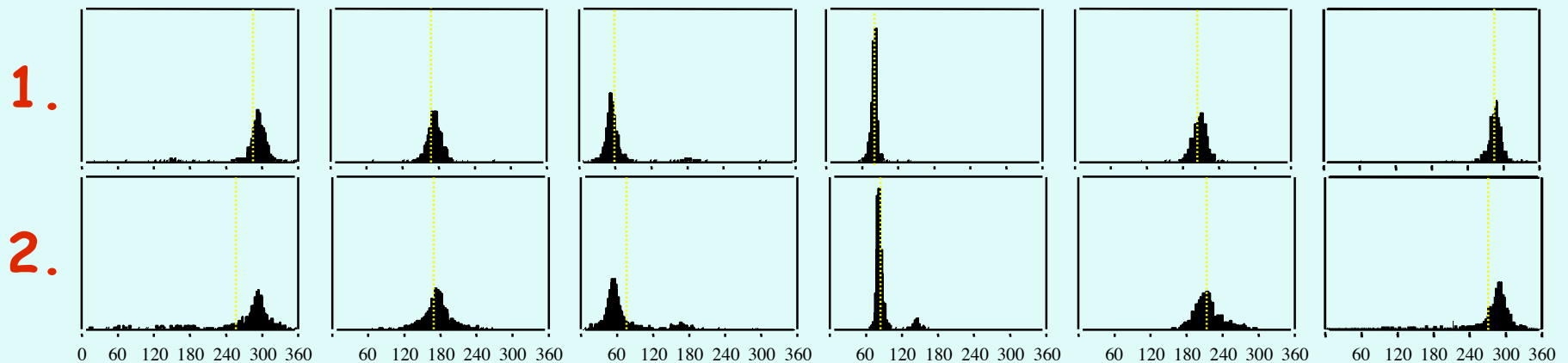
Stereochemistry of RNA



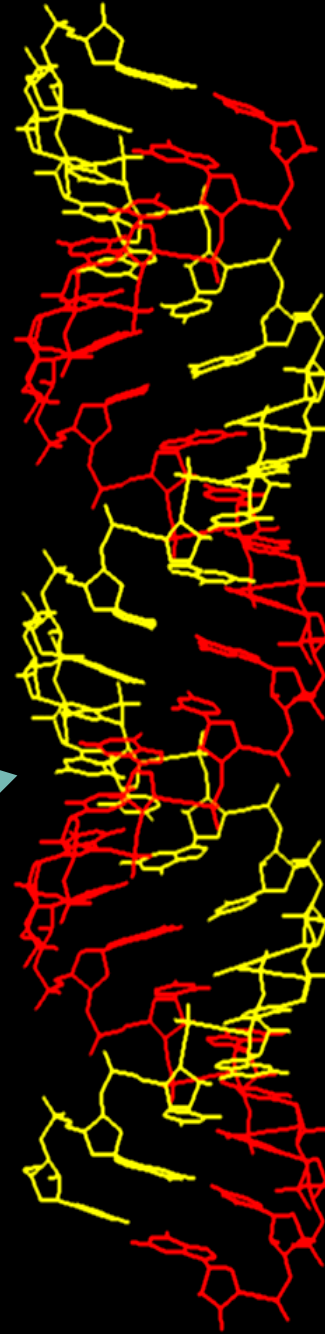
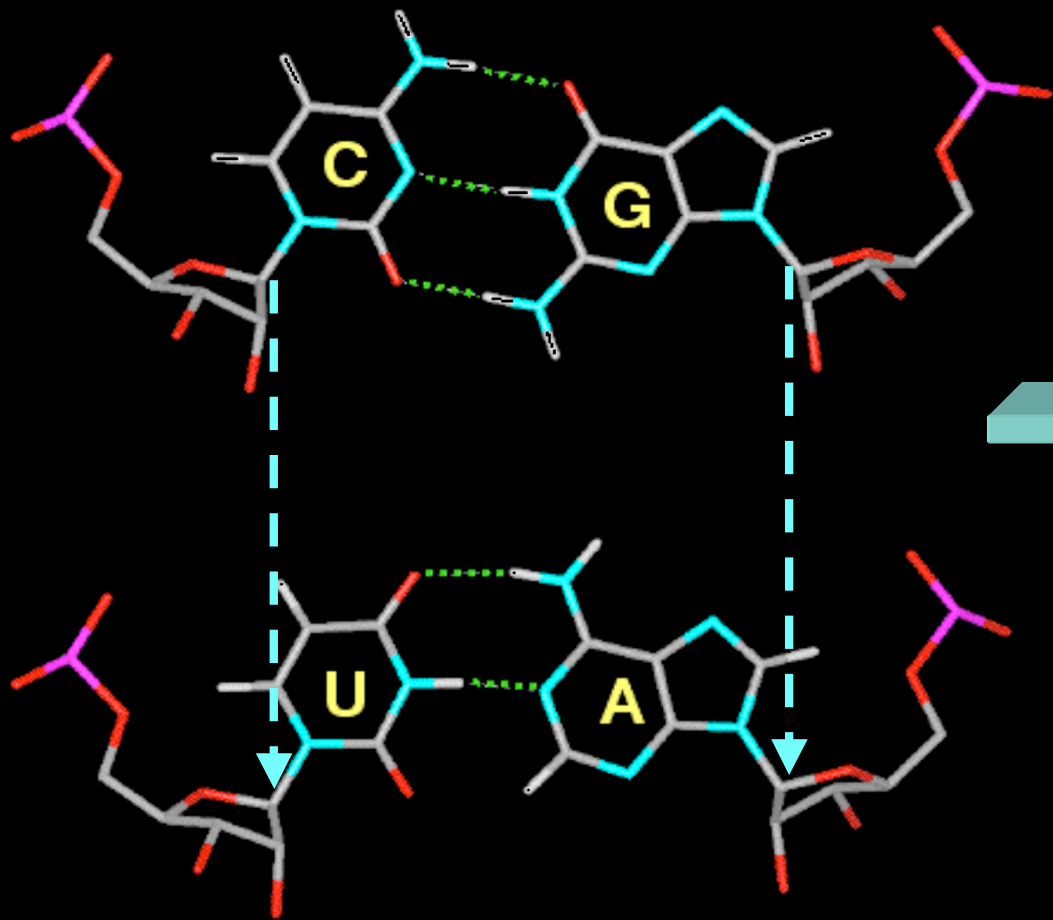
Torsion angles preferences :

1. Helices & 2. Single strands

P-O5'	O5'-C5'	C5'-C4'	C4'-C3'	C3'-O3'	O3'-P
gauche ⁻	trans	gauche ⁺	gauche ⁺	trans	gauche ⁻

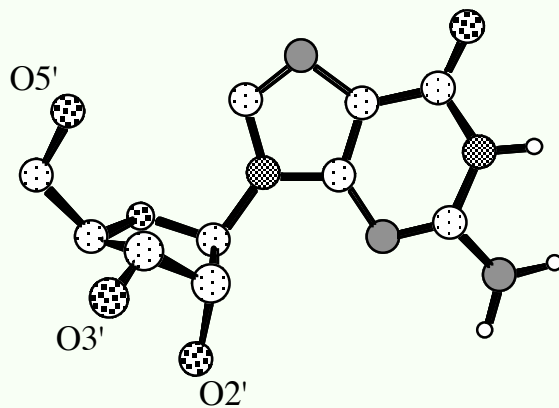


Watson-Crick pairs are
isosteric

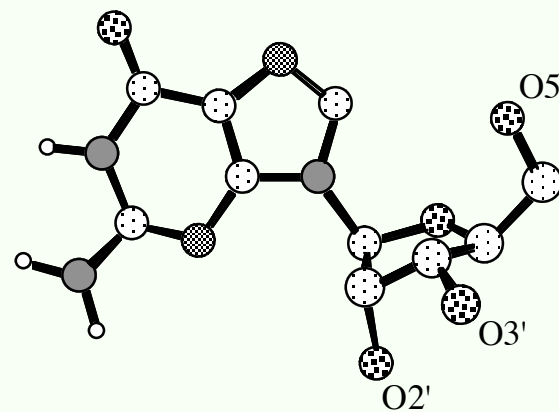


Several asymmetries nucleotides & polynucleotides:

- sugars are chiral
- 5' > 3' polarity in linkages
- strands are antiparallel
- helices are right-handed
- sugars are disposed asymmetrically with respect to the paired bases

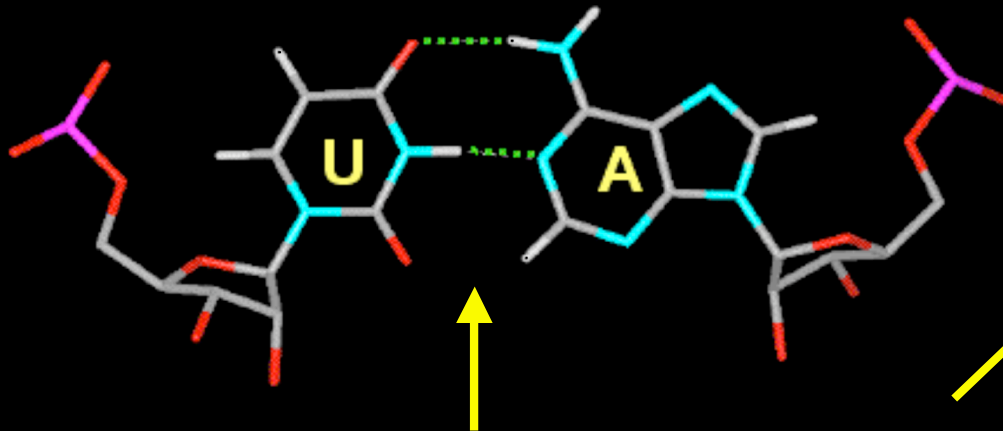
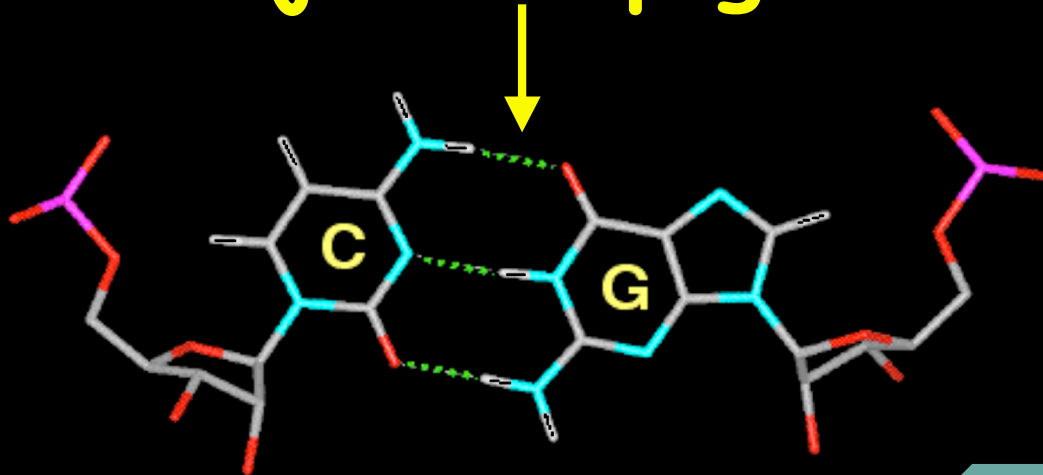


□-D-ribose

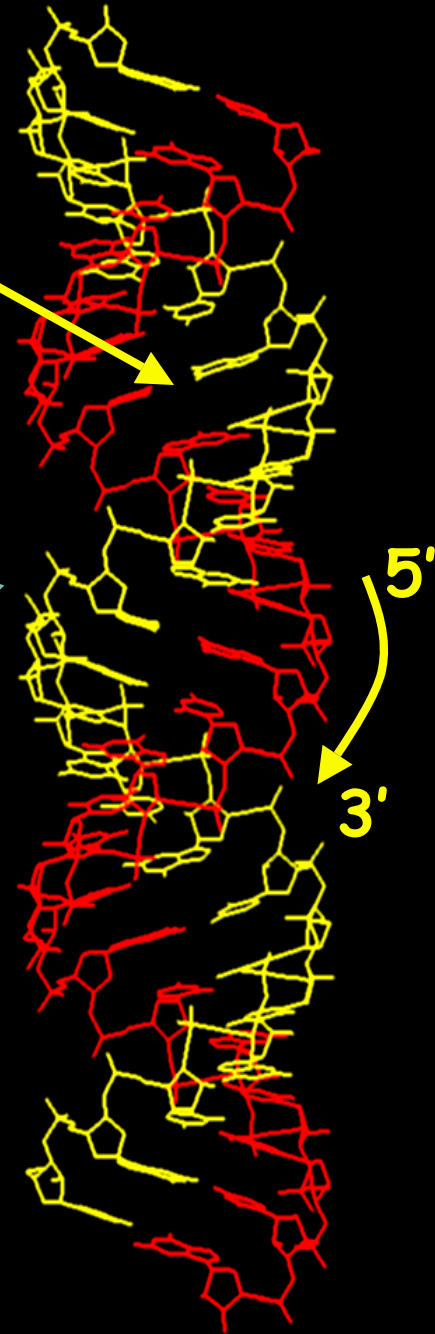


□-L-ribose

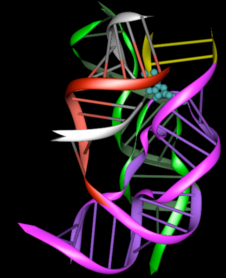
Major/deep groove



Minor/shallow groove



RNA FOLDING PROCESSES

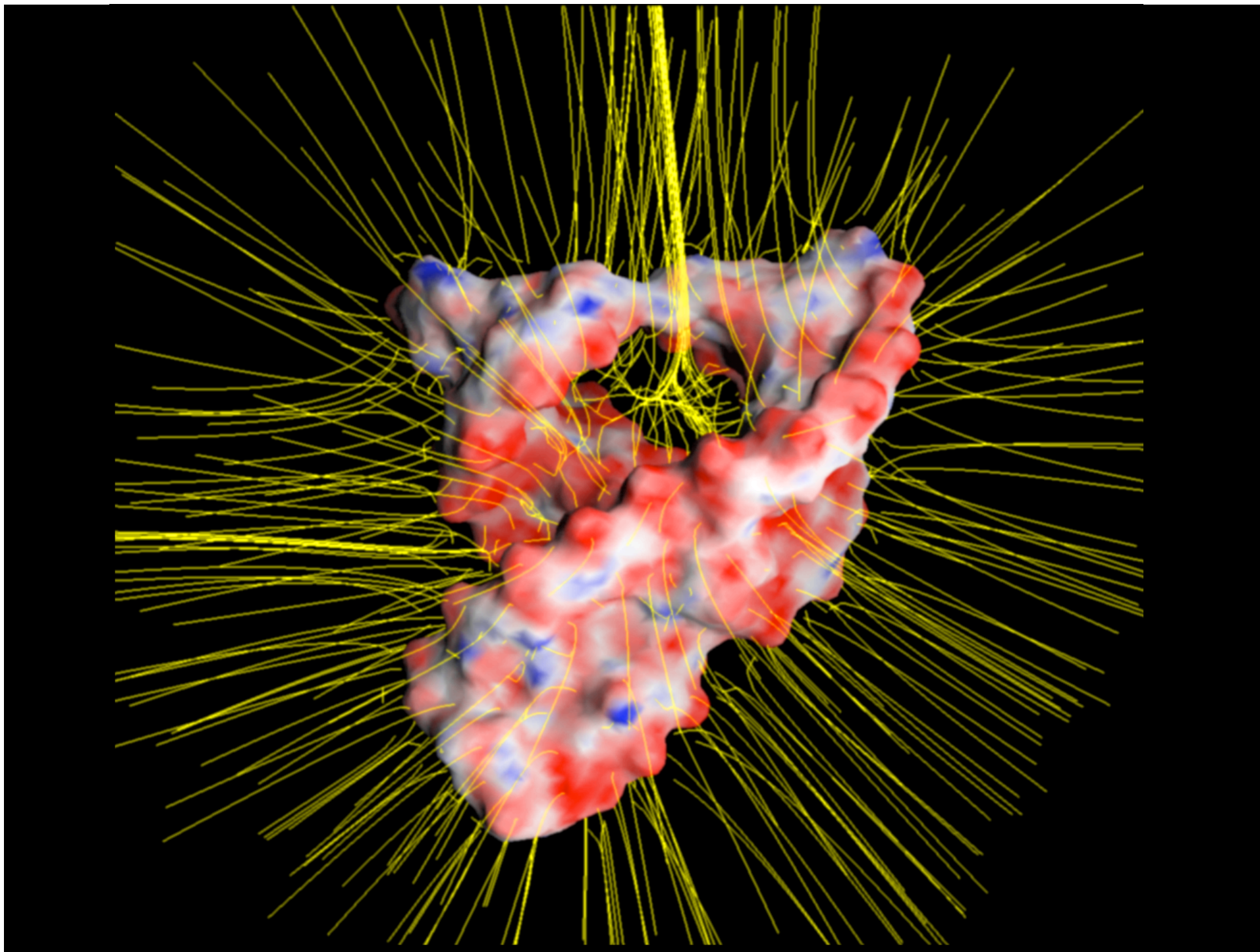


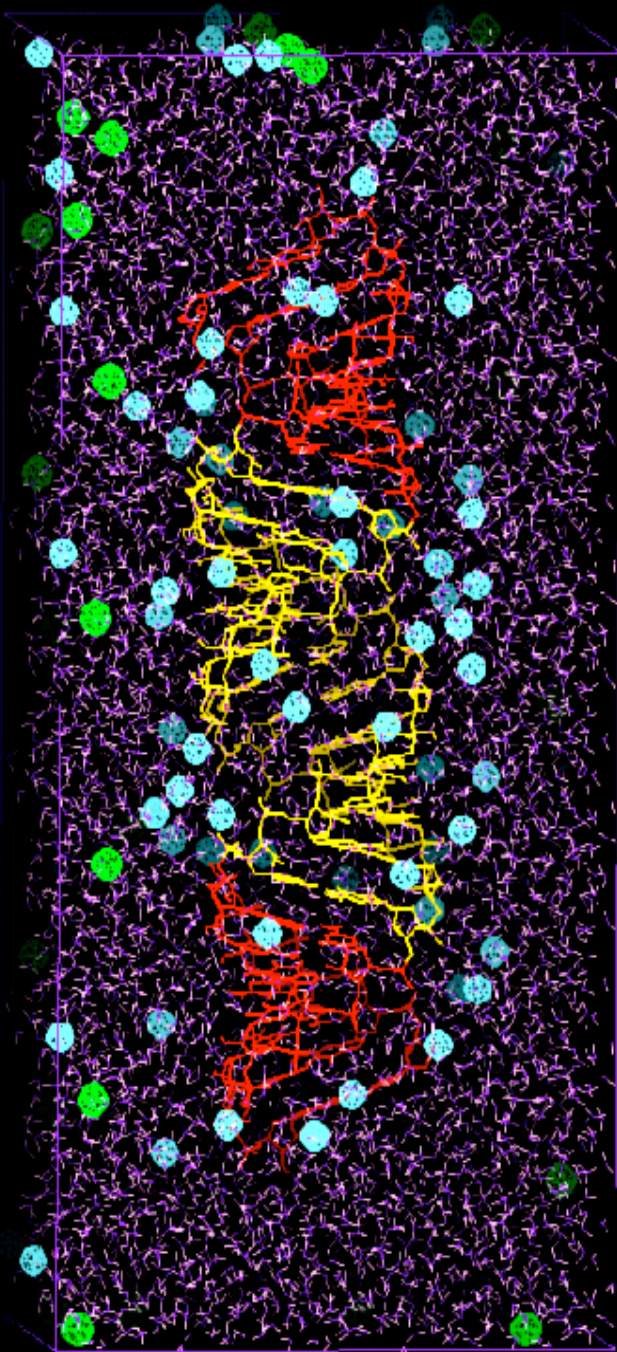
What do we mean ?

- In vitro vs in vivo

- 1D into 2D : in silico

- 2D into 3D : x-ray / modelling



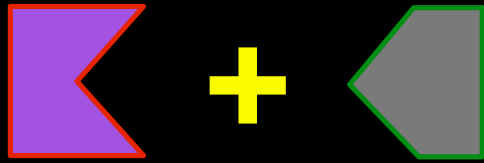


Water molecules
Counter-ions
Co-ions
Polyamines, ...

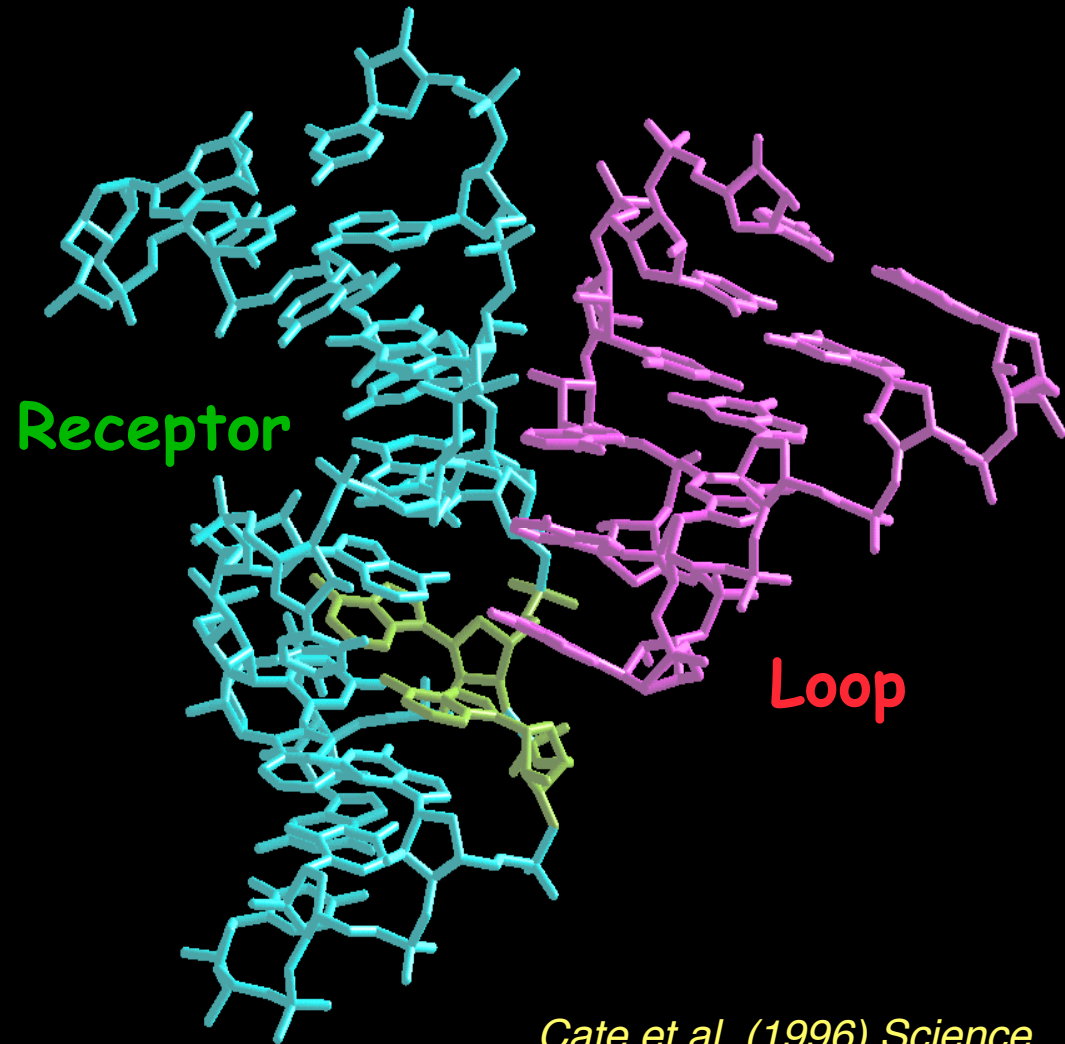
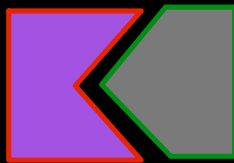
RNA-RNA self-assembly motif requires magnesium ions

Receptor

Loop



+
Magnesium



*Cate et al. (1996) Science
273*

RNA/ion interactions

1) Divalent cations



...

2) Monovalent cations



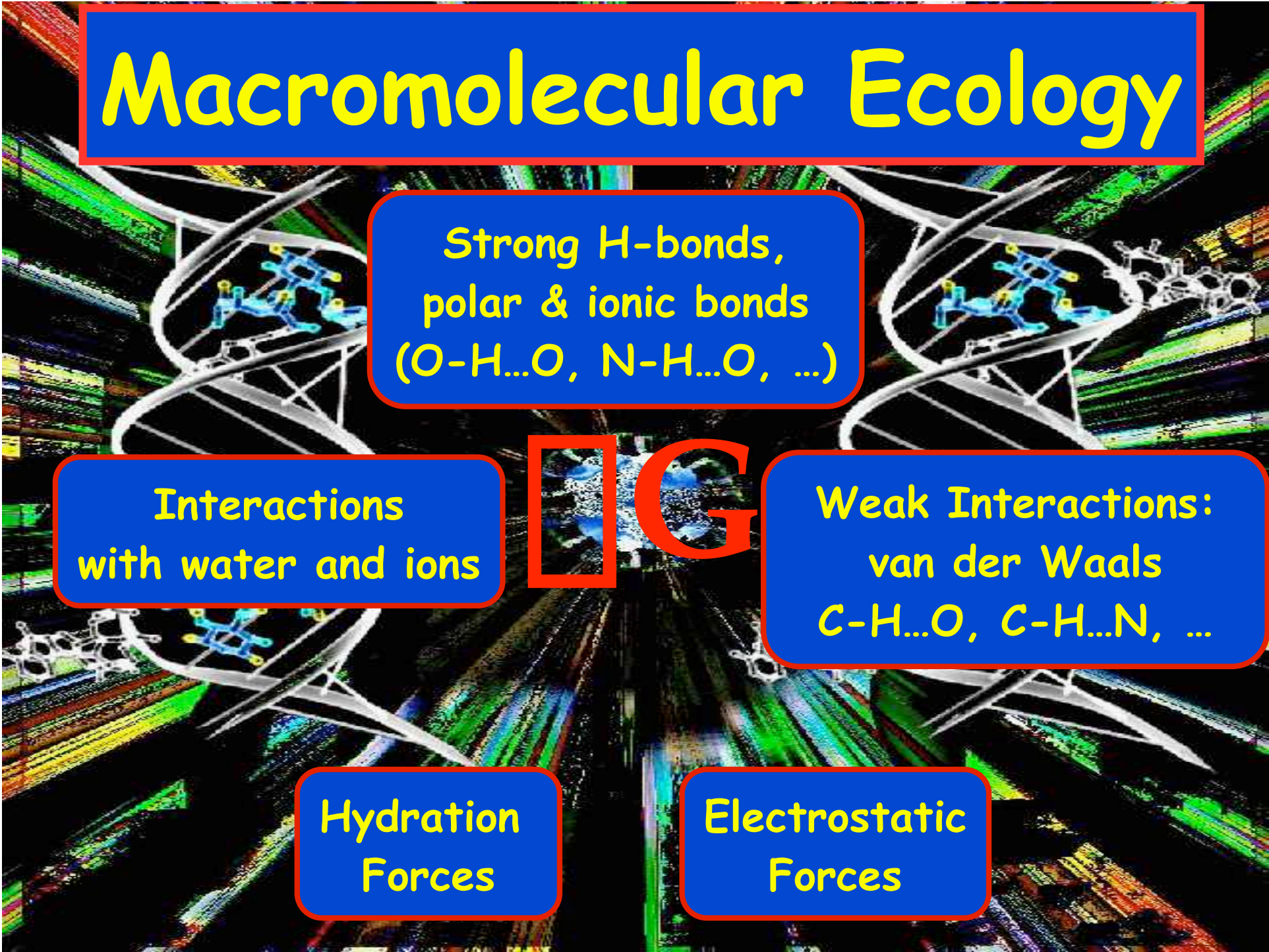
...

3) Anions



...

Macromolecular Ecology



Strong H-bonds,
polar & ionic bonds
(O-H...O, N-H...O, ...)

Interactions
with water and ions



Weak Interactions:
van der Waals
C-H...O, C-H...N, ...

Hydration
Forces

Electrostatic
Forces

In vitro folding:
Kinetic vs
thermodynamic control

In vivo folding:
Sequential 5'→3' or co-transcriptional
Modular and hierarchical

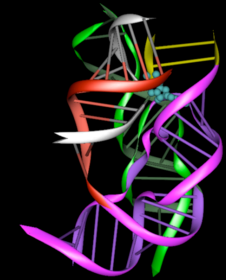
Electrostatic Hierarchy in RNA Folding

- 1. Rapid electrostatic collapse to compact but partially disordered structures ('molten globule') :
 - induced by non-specific outer-sphere binding of cations;
- 2. Cooperative and exothermic transition to the native state :
 - induced by specific inner-sphere binding of cations;

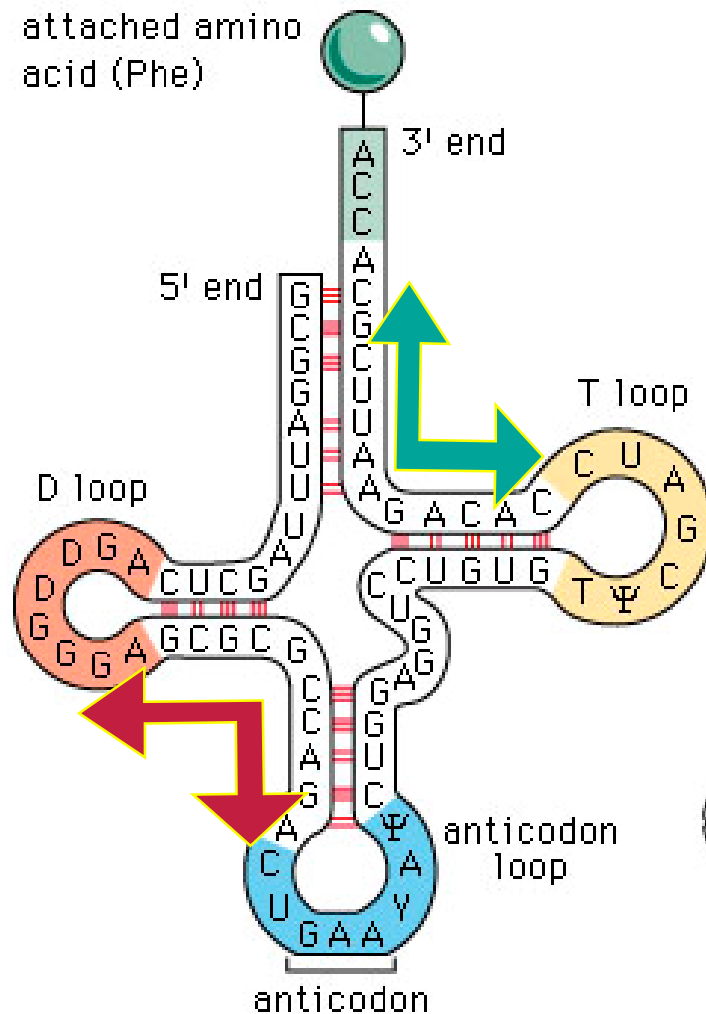
Architectural Hierarchy by modular assembly in RNA

- Helices and hairpin loops first form;
- Helices build sub-domains by parallel or end-to-end packing;
- Local and specific recognition contacts occur cooperatively between preformed sub-domains.

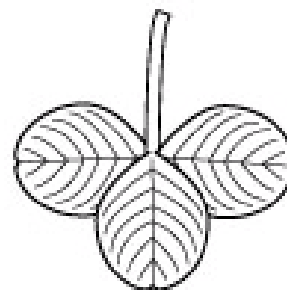
End-to-end stacking of helices



attached amino
acid (Phe)

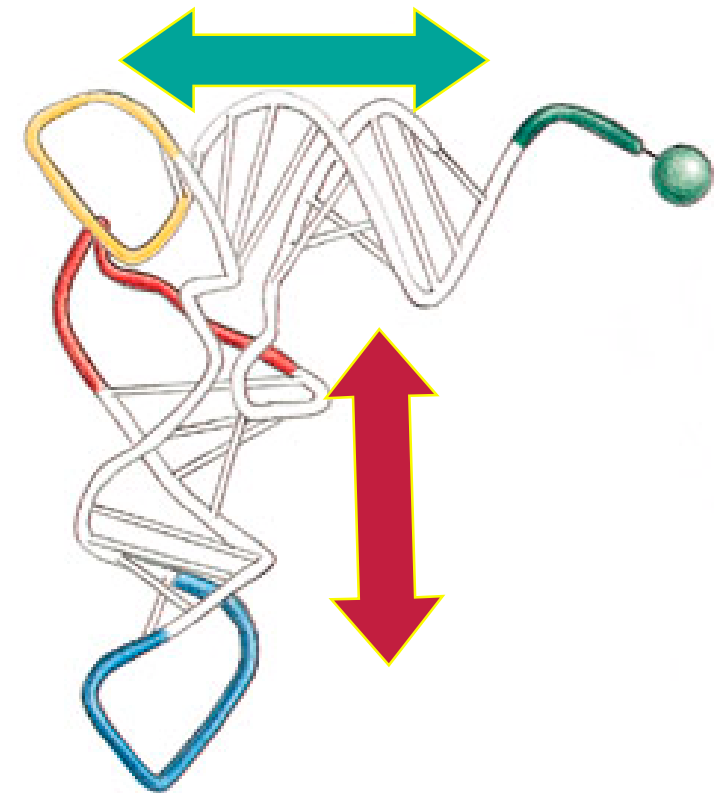


2D

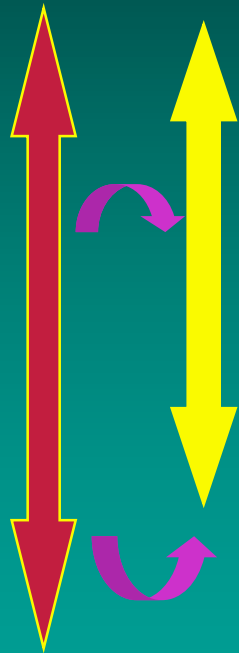
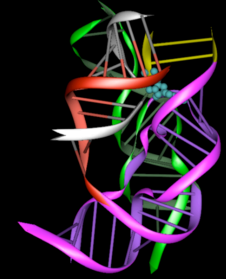


a clover leaf

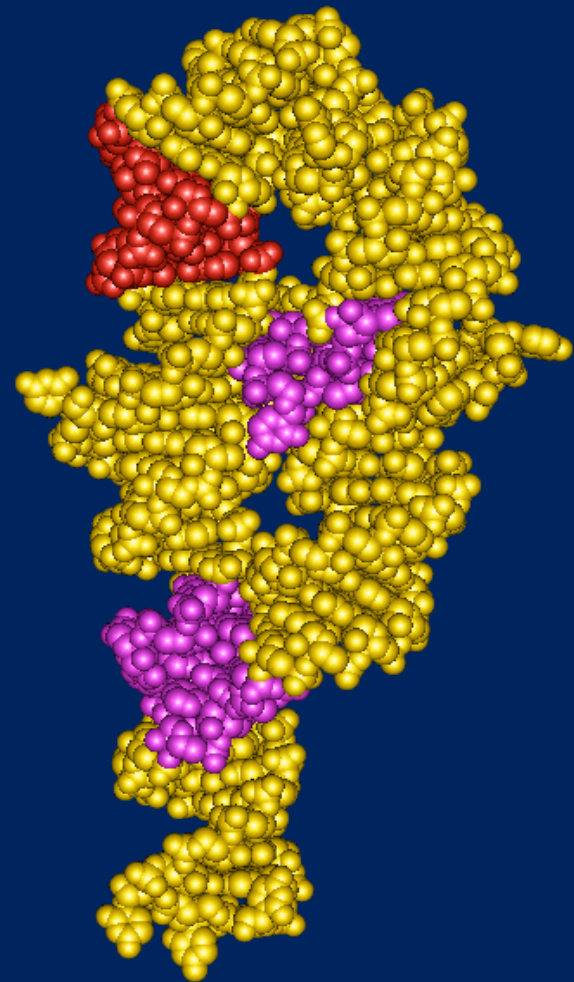
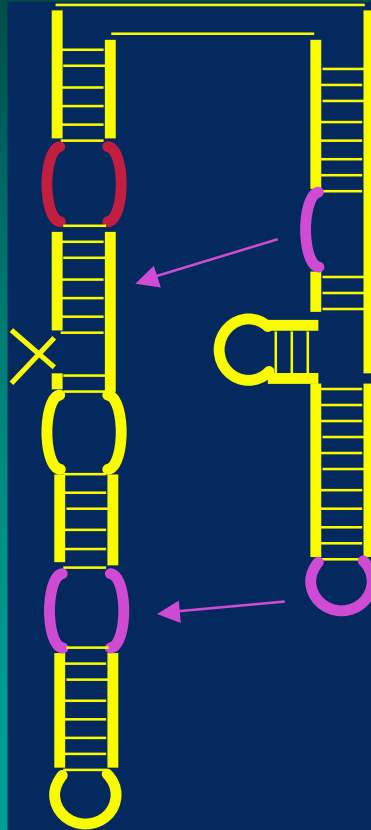
3D



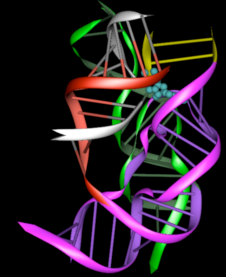
Parallel packing of helices



Specific 3D contacts
Mediated by ions/waters

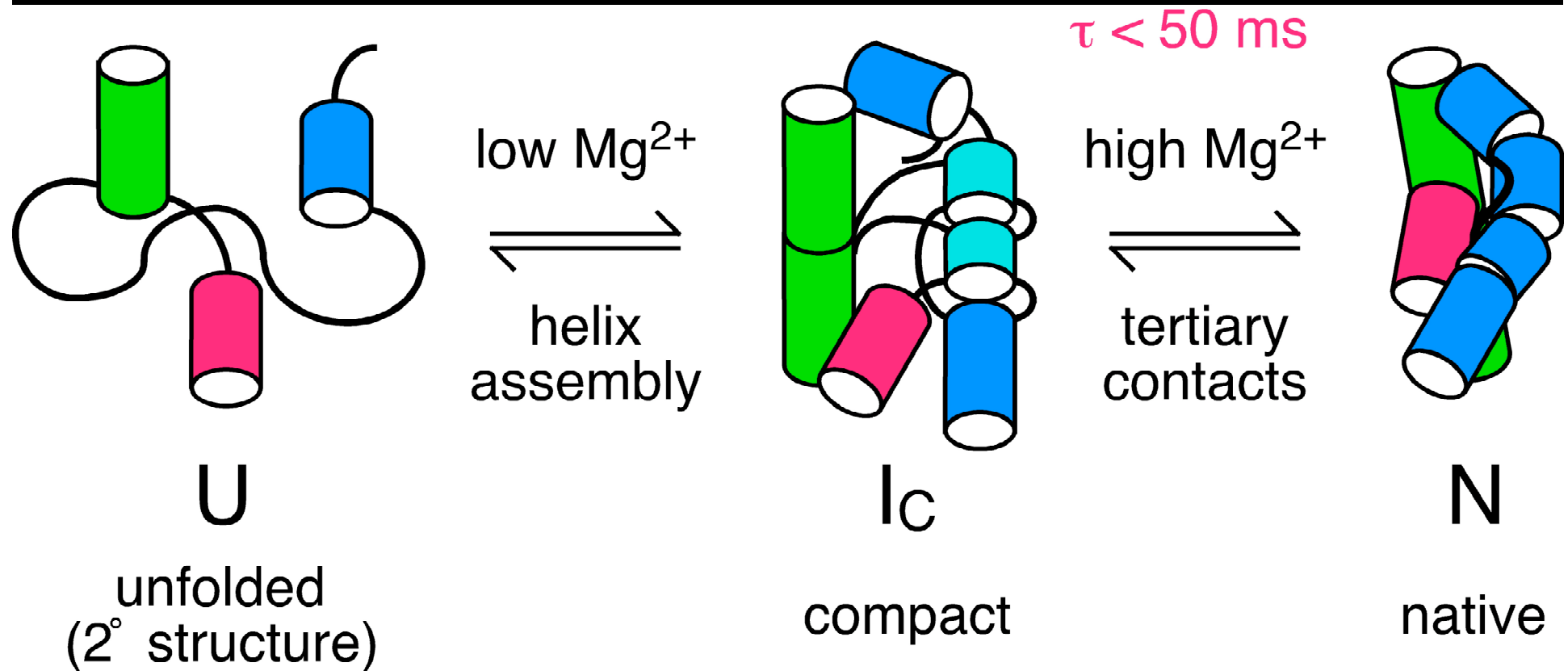


Some kinetics values



- Stacking of single-strands : 1 μ s
- Hairpin formation : 10 -100 μ s
- Tertiary structure formation : 10 -100 ms
- Native state : 1s - 10 min

Coupling between architectural and electrostatic hierarchies



With Rangan and Woodson

RNA self-assembly & folding

Coupled Architectural & Electrostatic Hierarchies

- Formation of helices that build subdomains by parallel or end-to-end packing &

rapid collapse to compact states induced by non-specific ion binding;

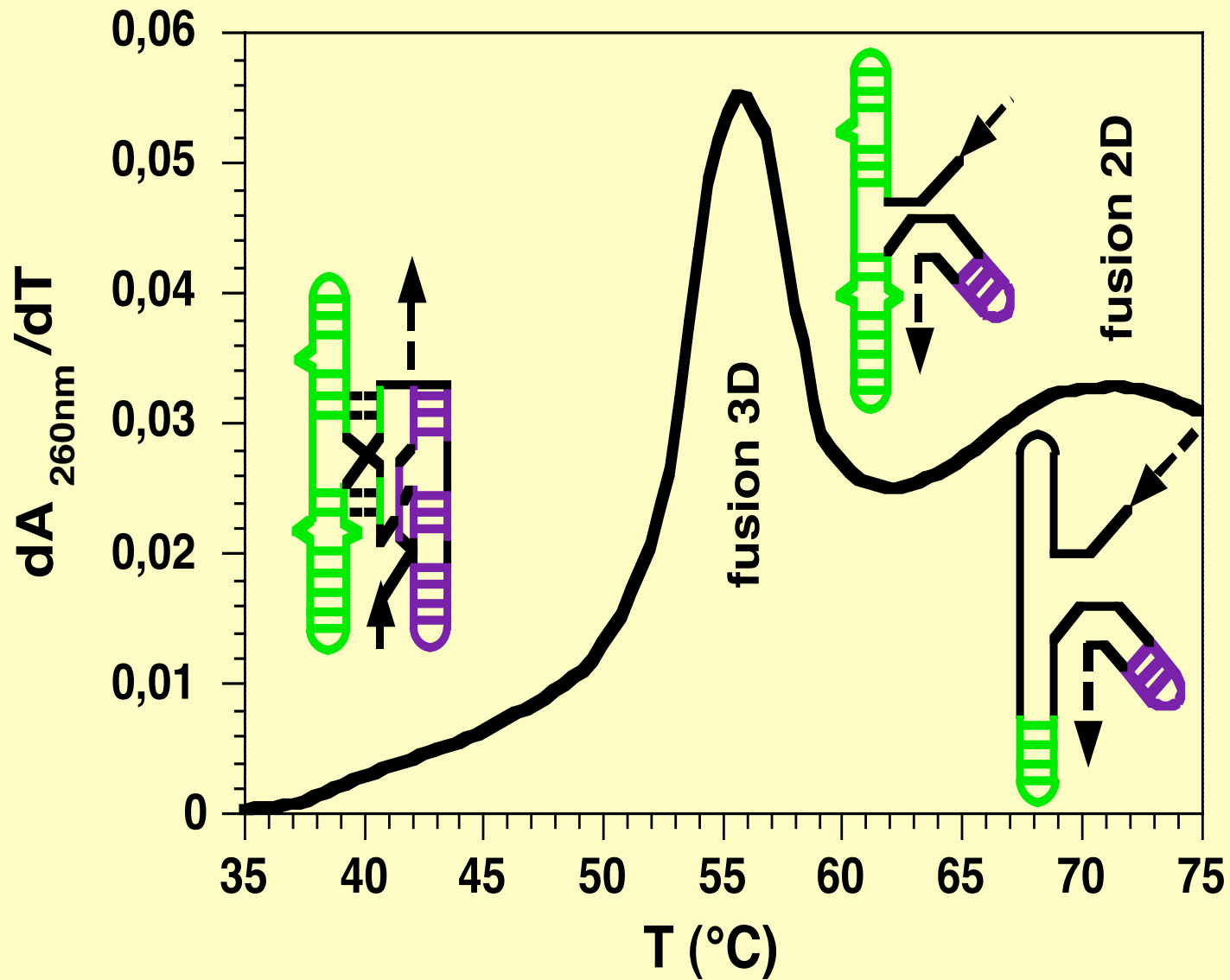
- Specific RNA-RNA recognition & cooperative transitions to native state promoted by specific ion binding.

During sequential RNA folding
(5' > 3'), architectural hierarchy
is coupled to electrostatic
hierarchy.

Favour local contacts

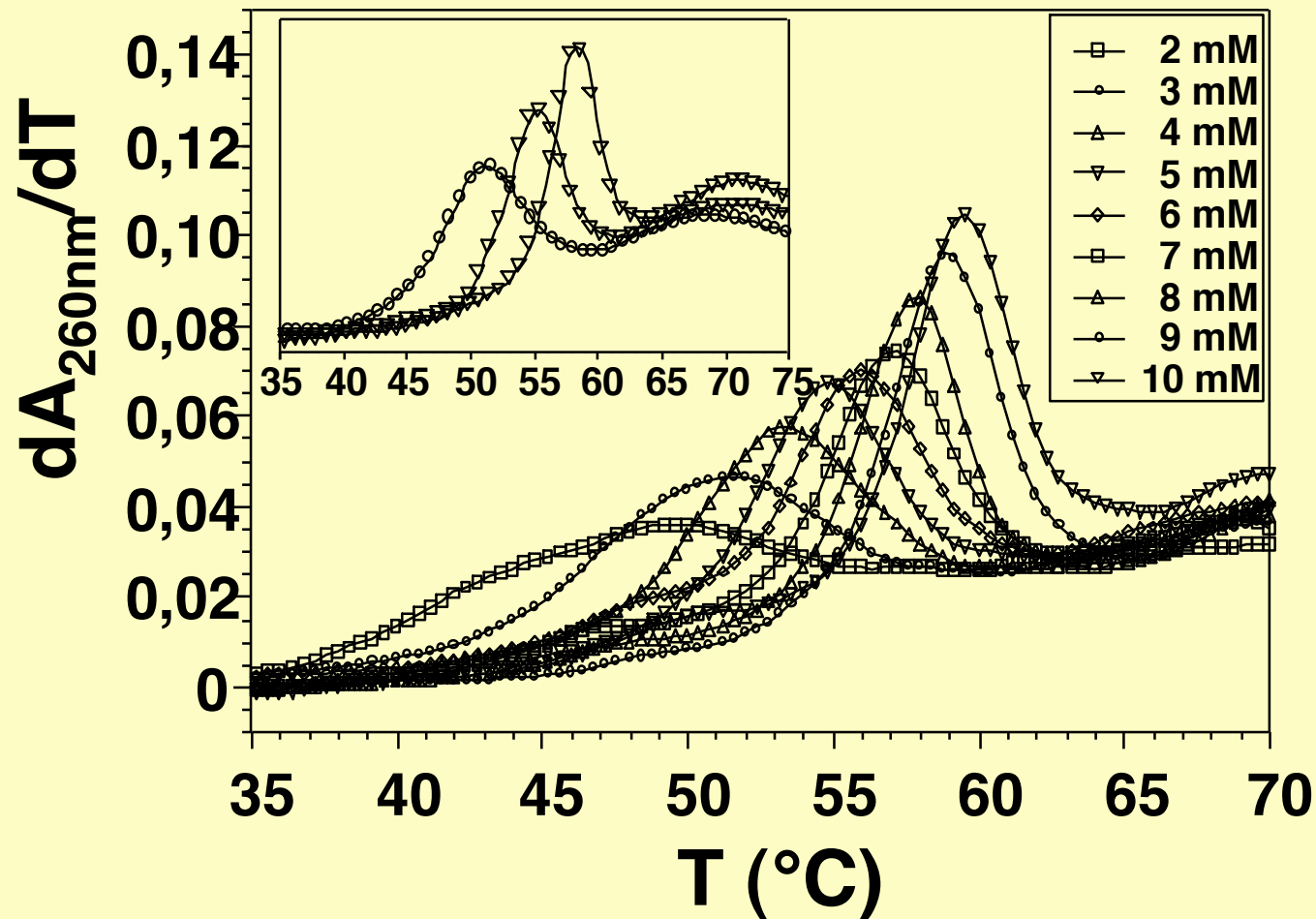
Inhibit distant contacts

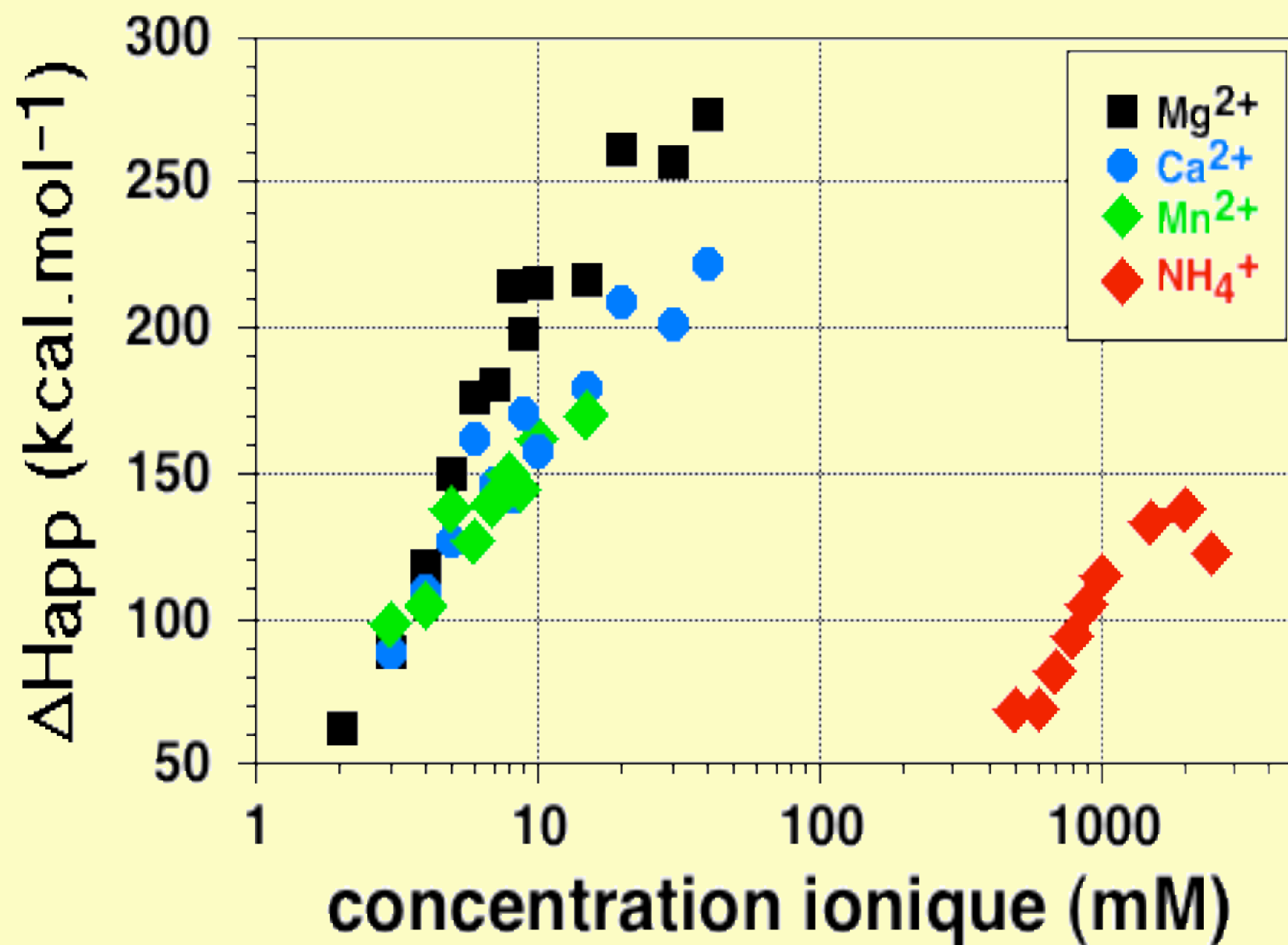
3D melt



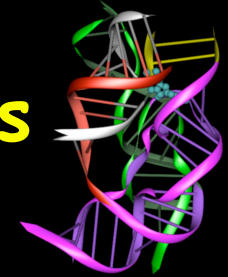
2D
melt

Effects of Mg^{2+} ions on UV-melts





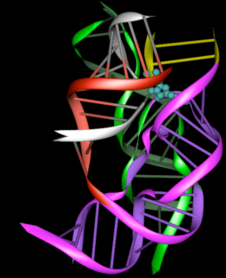
Effects of Mg^{2+} ions on thermodynamic values



0 to 50 mM Mg^{2+}	T_m (°C)	H (kcal.mole ⁻¹)	$G_{37^\circ C}$ (kcal.mole ⁻¹)
Octamers	11.0	10.0	2.2
Euk. Loop E	16.0	9.0	3.0
Eub. Loop E (up to 20 mM)	36.5	60.0	9.0

Metal	R ion (Å)	Number of ligands	Type	Rate of exchange with H ₂ O (s ⁻¹)	pK _a	pH @ 10 mM solubility
Ca ⁺⁺	1.06	6 à 8	hard	10 ⁸	12.9	12.4
Mg ⁺⁺	0.78	6	hard	10 ⁵	11.4	9.6
Mn ⁺⁺	0.91	6	soft	10 ⁷	10.6	7.9
Co ⁺⁺	0.82	6	soft	10 ⁶	10.2	7.2
Ni ⁺⁺	0.78	6	soft	10 ⁴	9.9	8.1
Zn ⁺⁺	0.83	4 à 6	soft	10 ⁷	9.0	6.6
Pb ⁺⁺	1.32	6	soft	10 ⁹	7.7	7.2
Co ⁺⁺⁺	0.64	6	hard	1		

Some values for ions

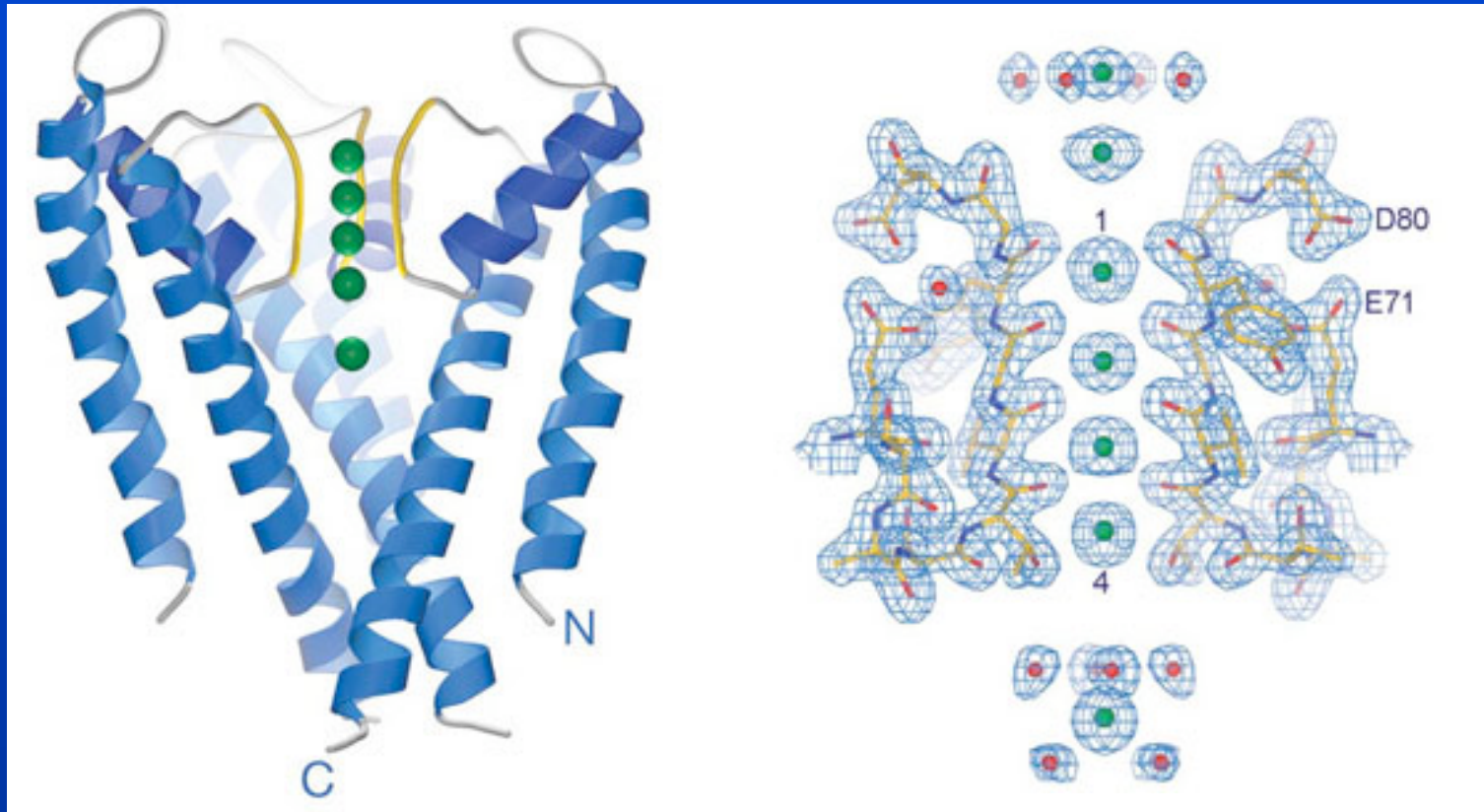


	Mg^{2+}	Ca^{2+}
Ionic radius	0.65 Å	1.0 Å
$M^{2+} \dots O$	2.05 Å	2.35 Å
Rates of water exchange	10^6 s^{-1} Slow	10^8 s^{-1} Fast
$G_{\text{hydration}}$ (kcal.mole $^{-1}$)	-460.	-380.

$$G_{\text{hyd}} = -164Z^2/(r+0.7)$$

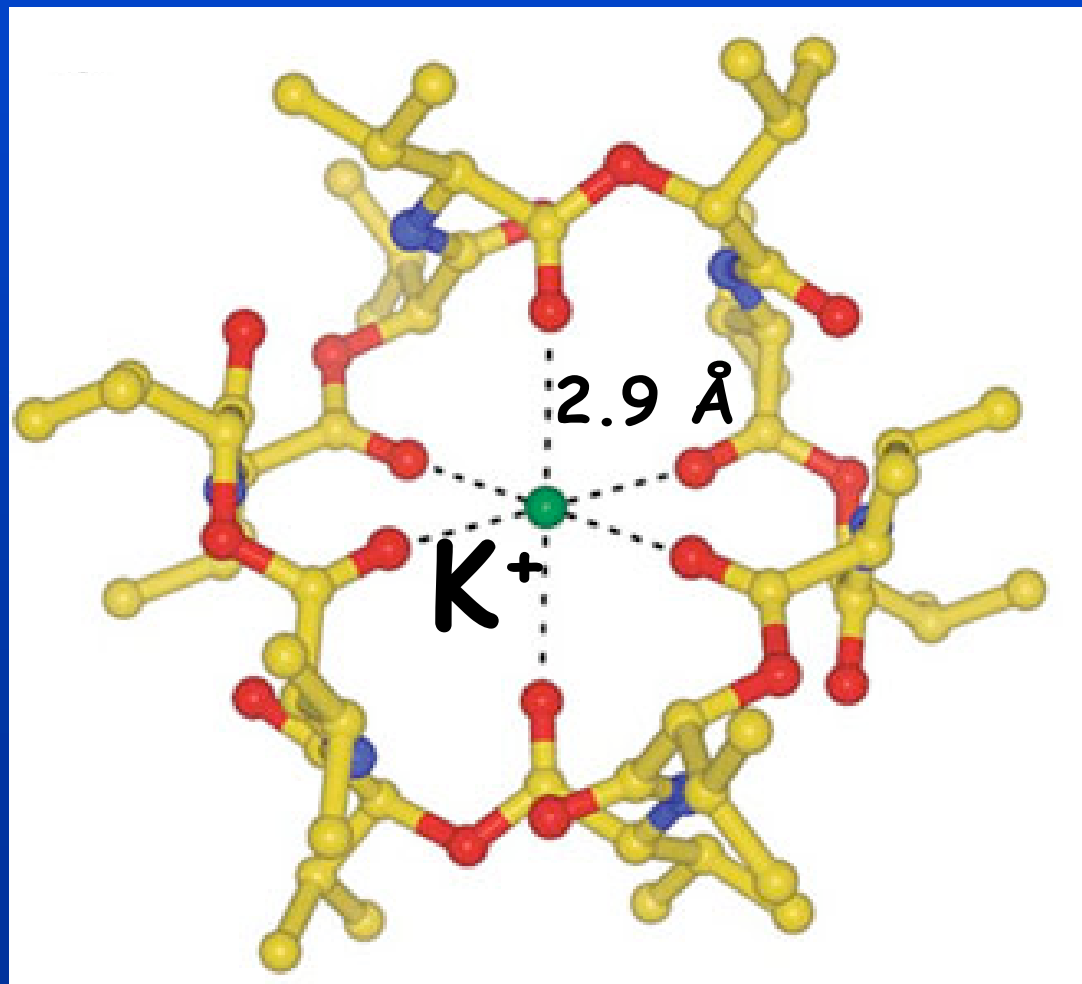
Born

Potassium channels

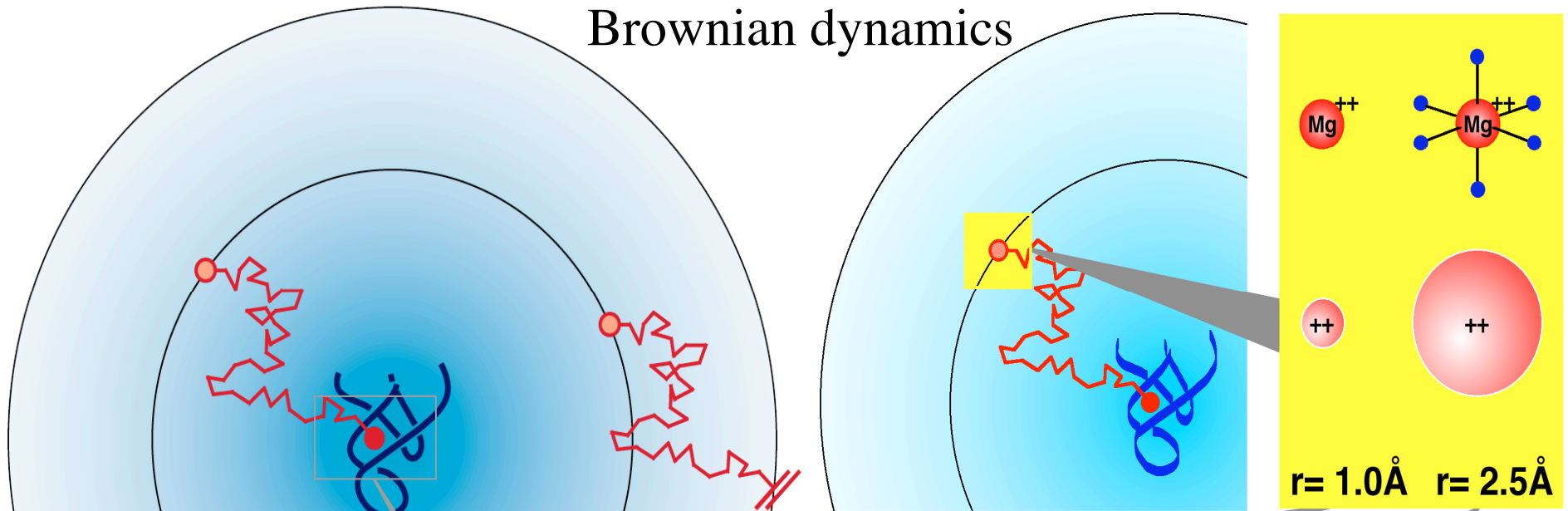


$\text{Na}^+ \dots \text{Ow} \ 2.5 \text{\AA}$ $\text{K}^+ \dots \text{Ow} \ 2.9 \text{\AA}$

Nonactin

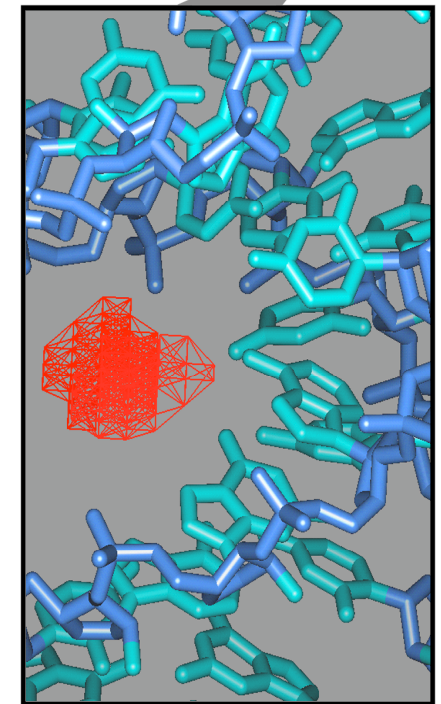
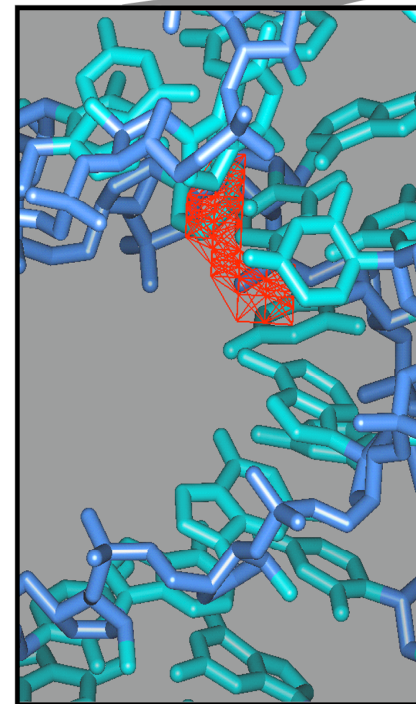
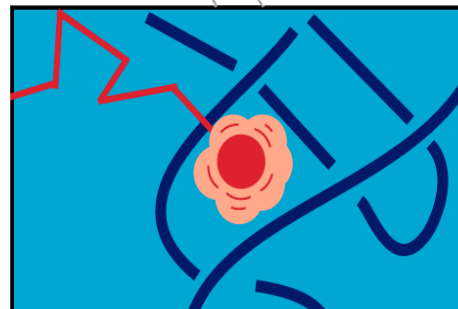


Brownian dynamics



$$r = r_0 + \frac{D_{\text{diffusion}} F(r_0)}{k_B T} \Delta t + R_{\text{random}}$$

IHRD (Madura et al 1994)



Dev. [Å]

E.coli 5S rRNA: loop

0.7 A —

1.3 B —

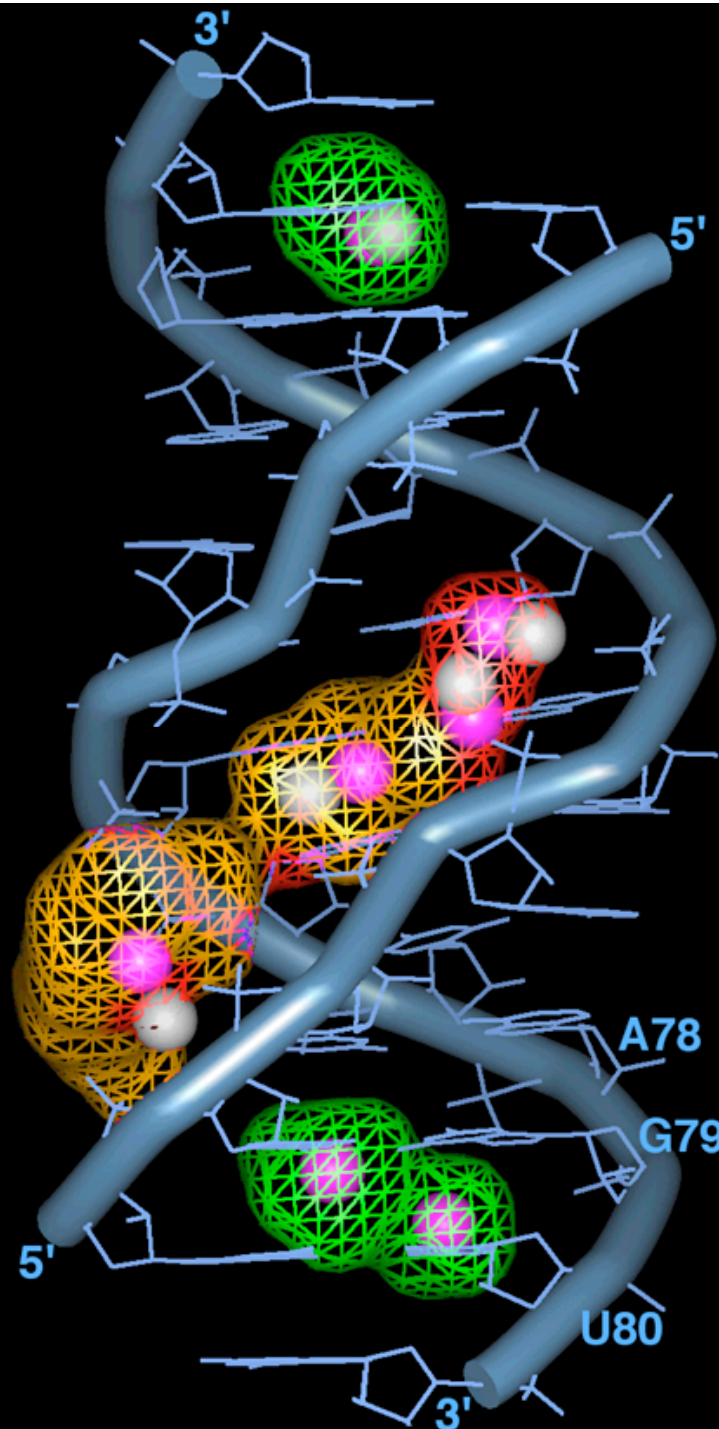
1.4 C —

1.8 D —

2.7 E —

F —

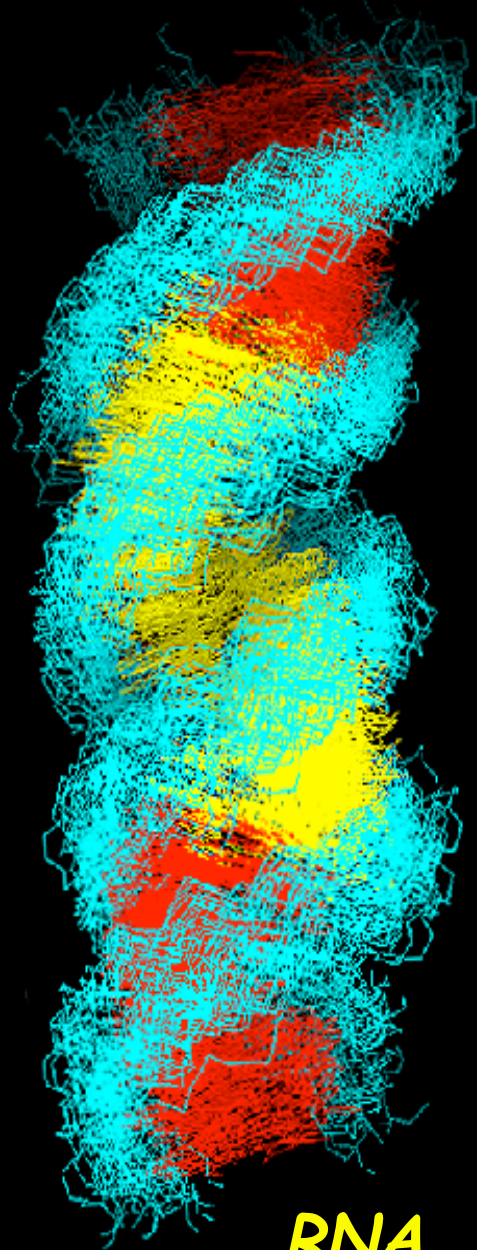
G —



3' 5'
 G - C·70
 G - C
 A G
 U - A
 G · U
 A G
 100· G G
 A - U
 G A
 C - G
 G · U·80
 5' G
 3'

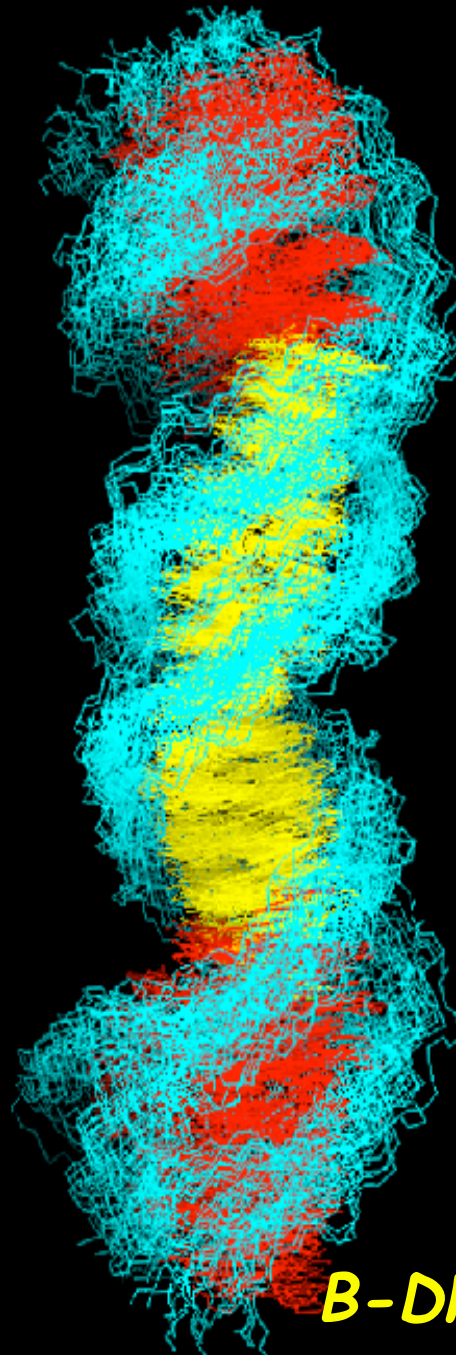
(Correll et al. 199)

MD simulations

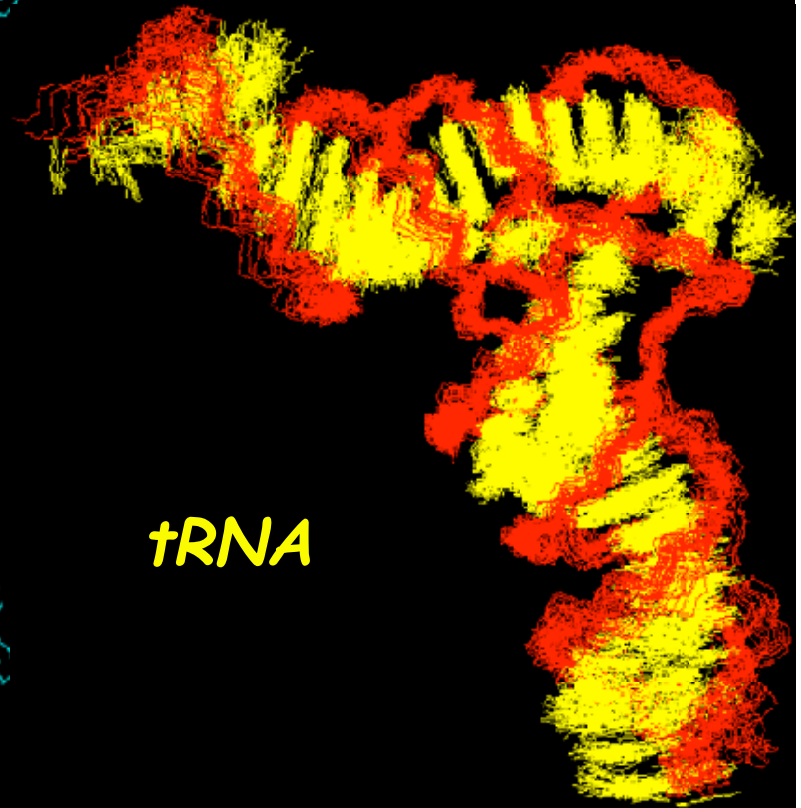


RNA

- | | |
|----|-----|
| 1 | C-G |
| 2 | G-C |
| 3 | C-G |
| 4 | G-C |
| 5 | C-G |
| 6 | G-C |
| 7 | C-G |
| 8 | G-C |
| 9 | C-G |
| 10 | G-C |
| 11 | C-G |
| 12 | G-C |
| 13 | C-G |
| 14 | G-C |
| 15 | C-G |
| 16 | G-C |
| 17 | C-G |
| 18 | G-C |
| 19 | C-G |
| 20 | G-C |
| 21 | C-G |
| 22 | G-C |
| 23 | C-G |
| 24 | G-C |



B-DNA



tRNA

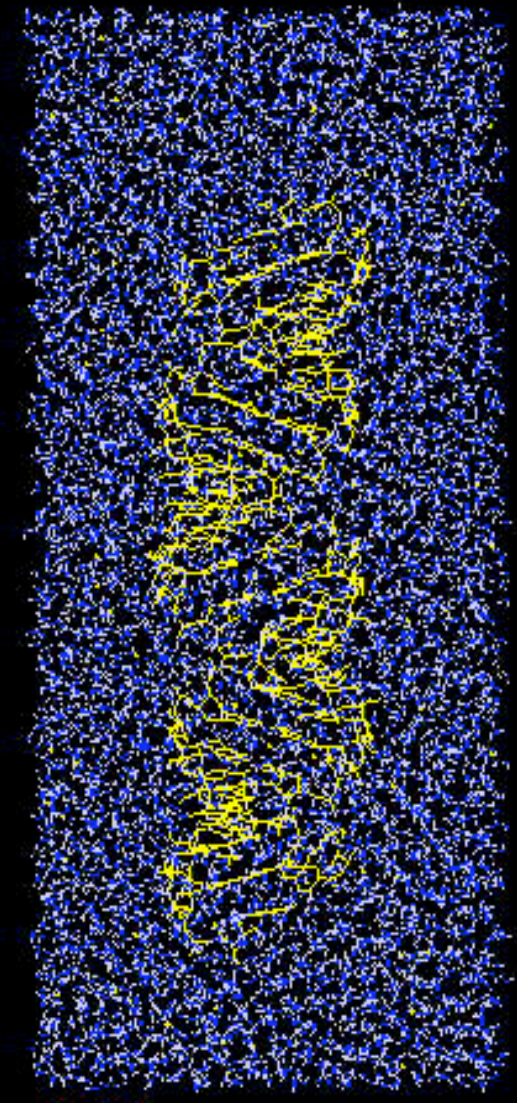
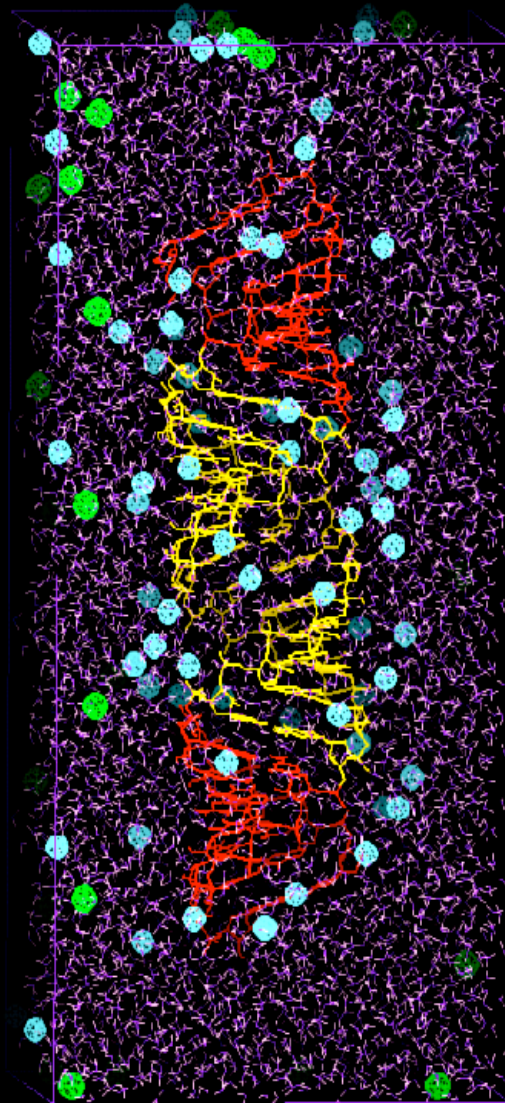
Molecular dynamics simulations

RNA duplexes (with P. Auffinger)

- 24 base pairs
- ≈ 0.25 M KCL
72 K^+ 26 Cl^-
(blue) (green)
- ≈ 5600 H_2O
- 4.4 ns length
- EWALD summation

AMBER - MD
Peter Kollman -
UCSF

1	C-G
2	G-C
3	C-G
4	G-C
5	C-G
6	G-C
7	C-G
8	G-C
9	C-G
10	G-C
11	C-G
12	G-C
13	C-G
14	G-C
15	C-G
16	G-C
17	C-G
18	G-C
19	C-G
20	G-C
21	C-G
22	G-C
23	C-G
24	G-C



0.250

Calculated densities

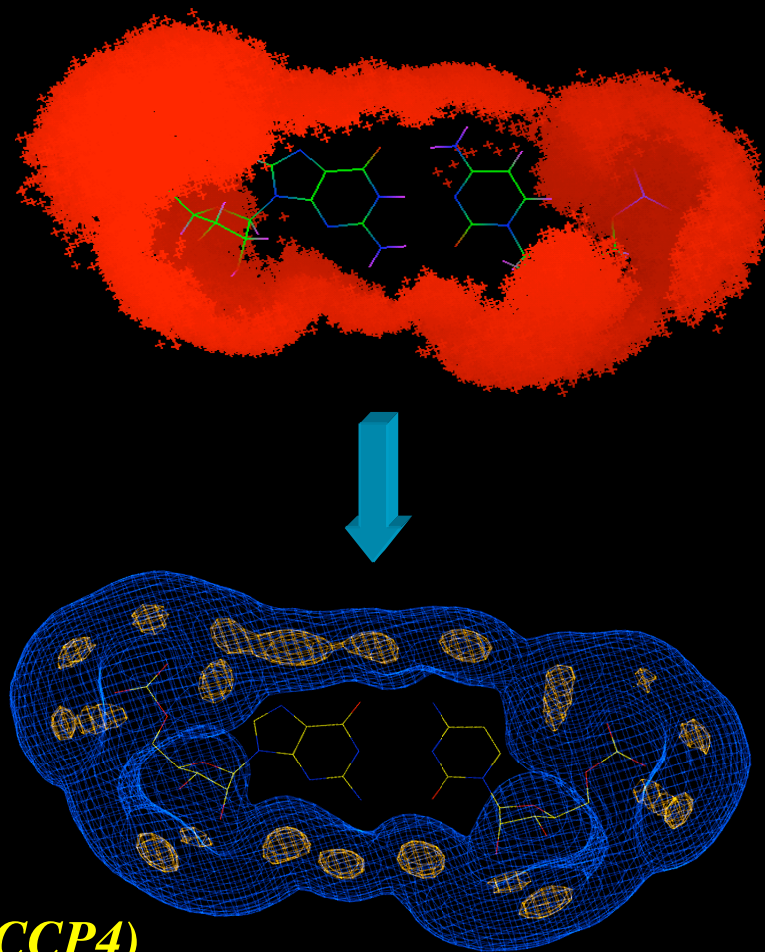
(Example of a G=C pair)

Set of water molecules at less than 3.5 Å from bp atoms

Leads to :

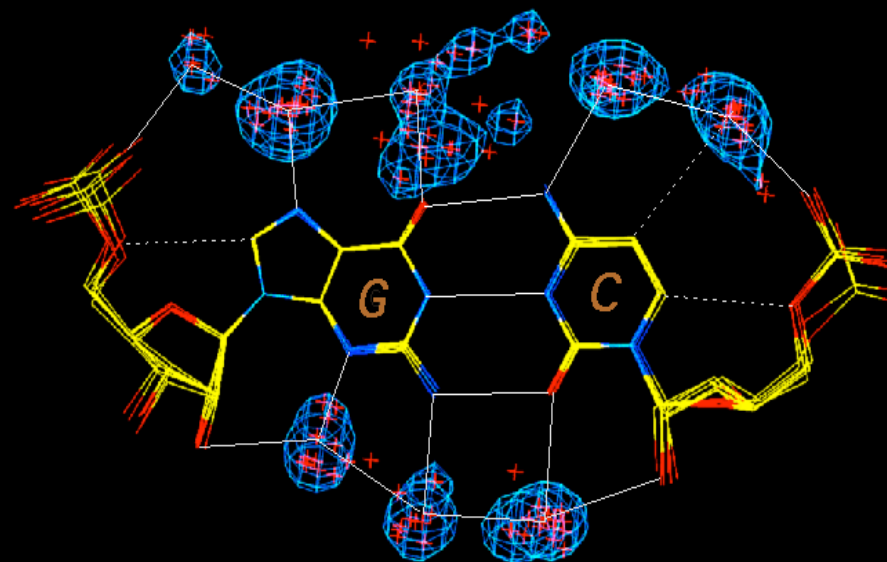
Low and high level
water densities

(generated by using crystallographic toolkits : CCP4)

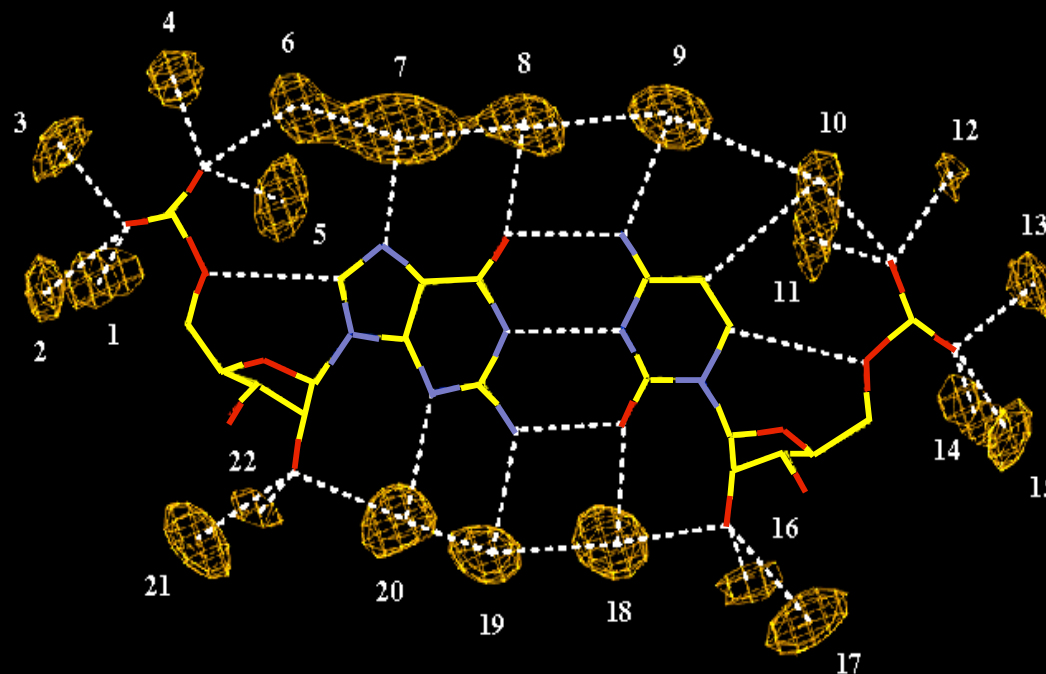


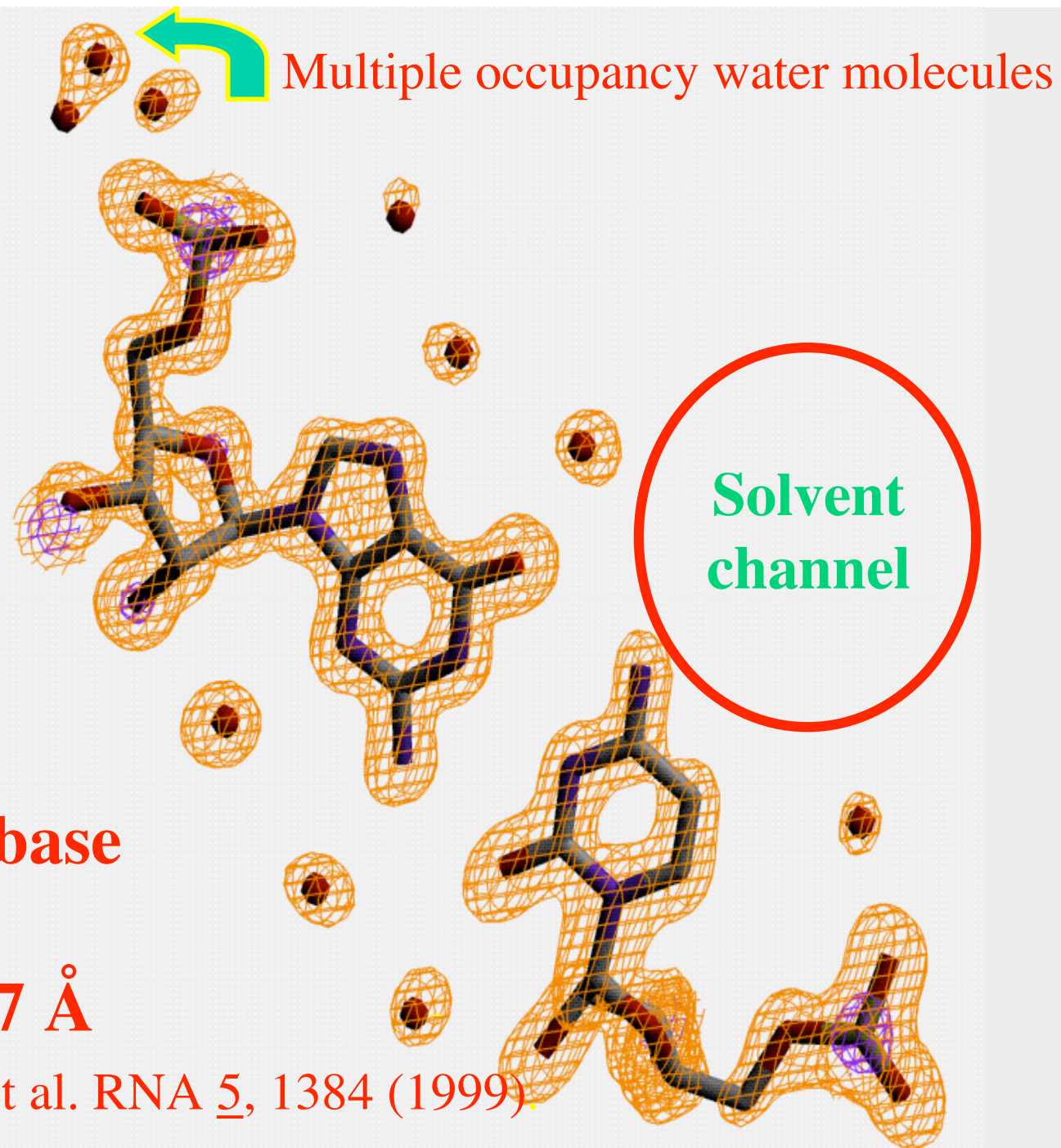
Statistical analysis of X-ray data

25 bp
121 waters



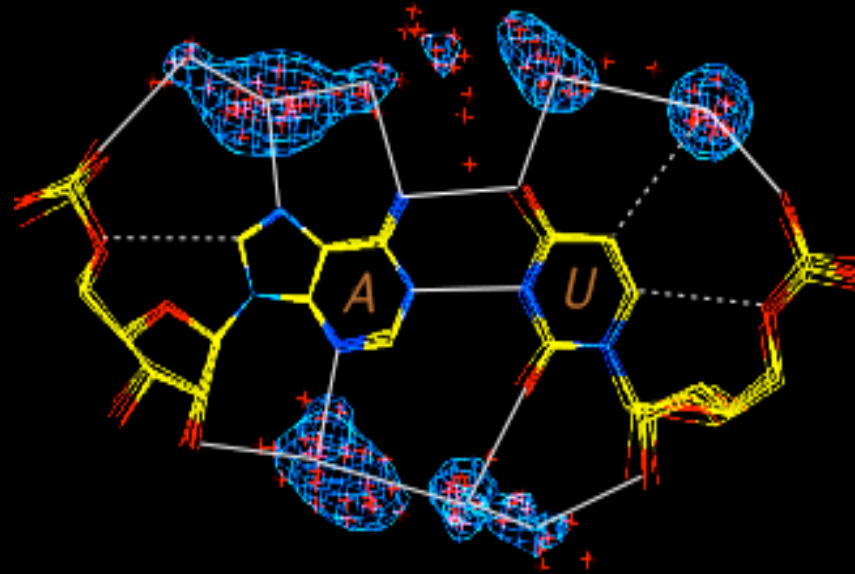
MD results



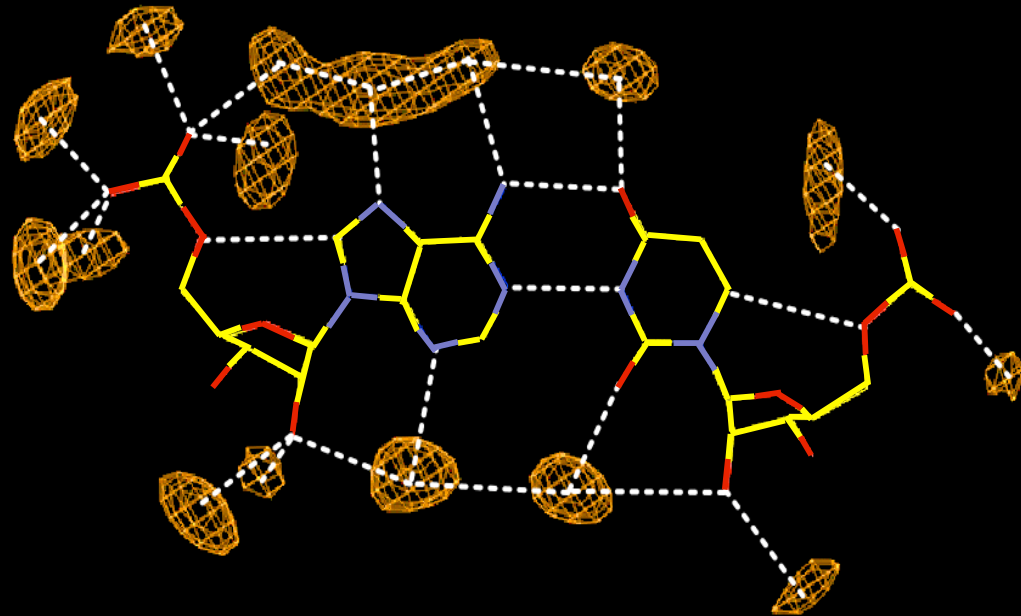


Statistical analysis of X-ray data (*A-U pairs*)

30 bp
108 waters

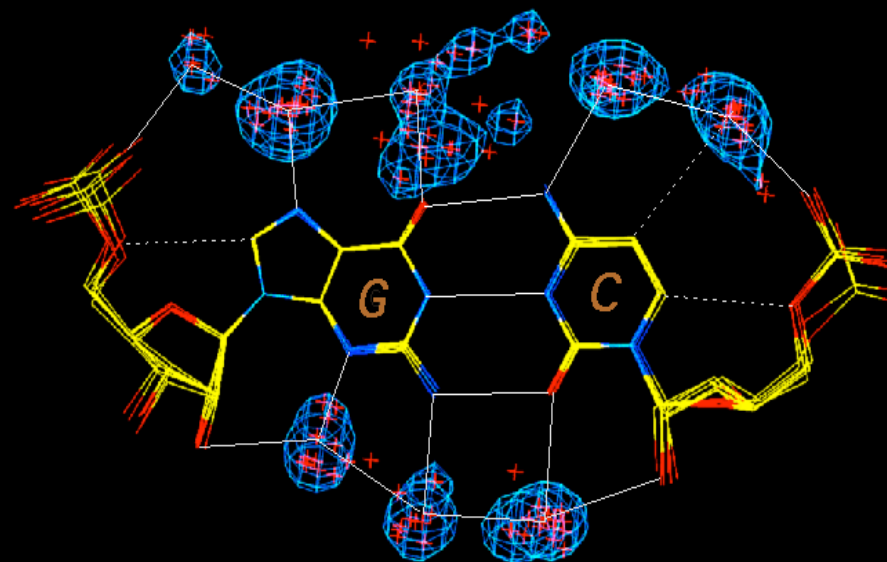


MD results

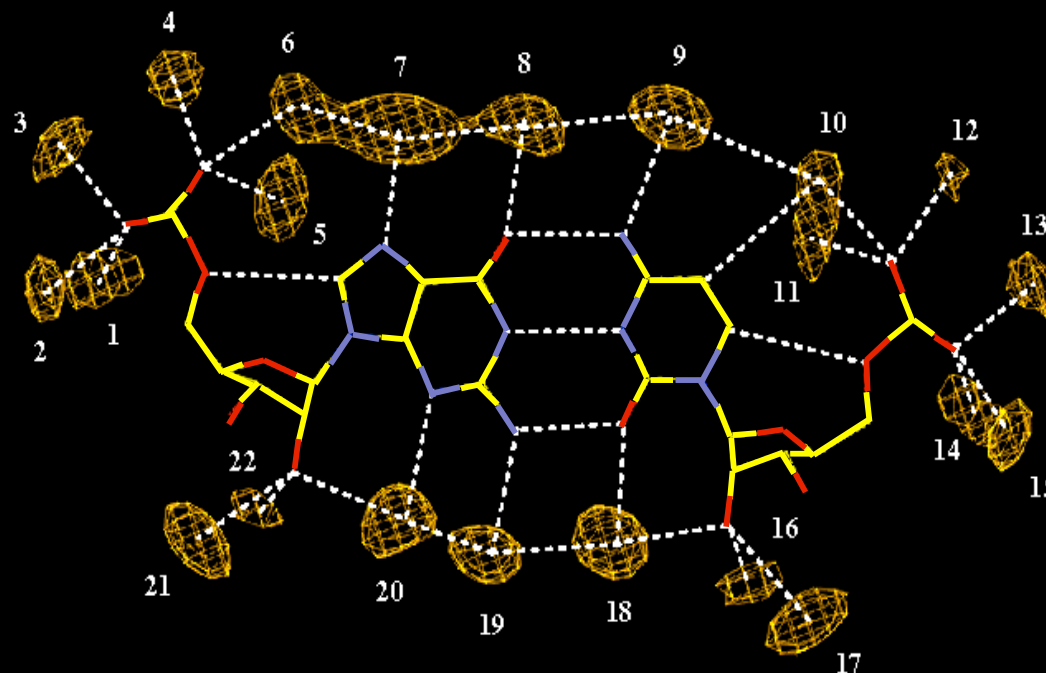


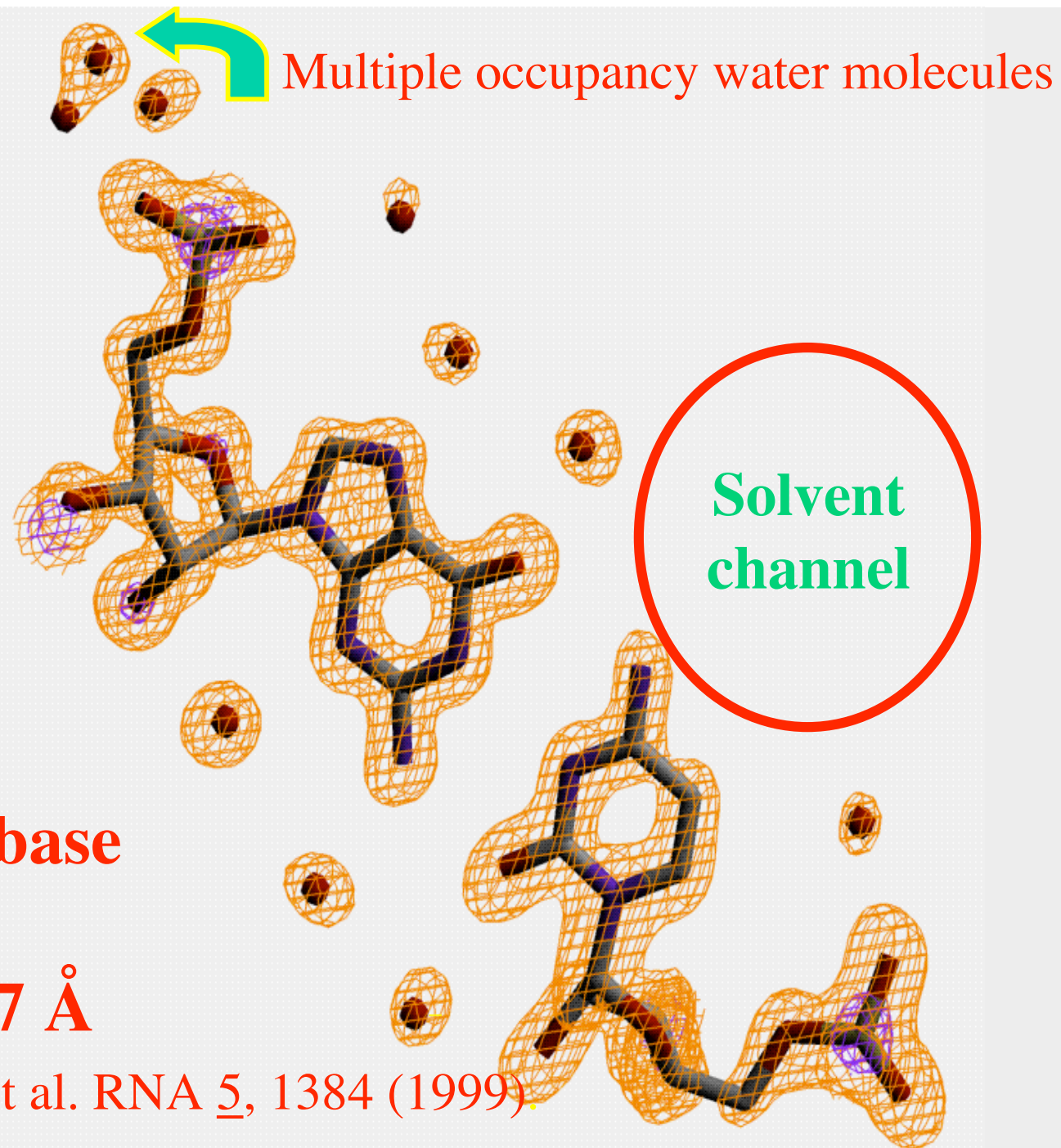
Statistical analysis of X-ray data

25 bp
121 waters



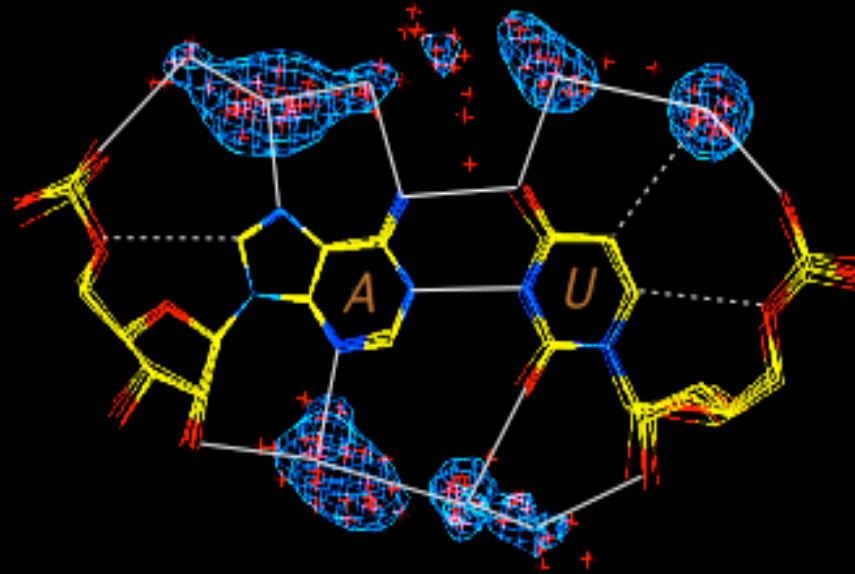
MD results



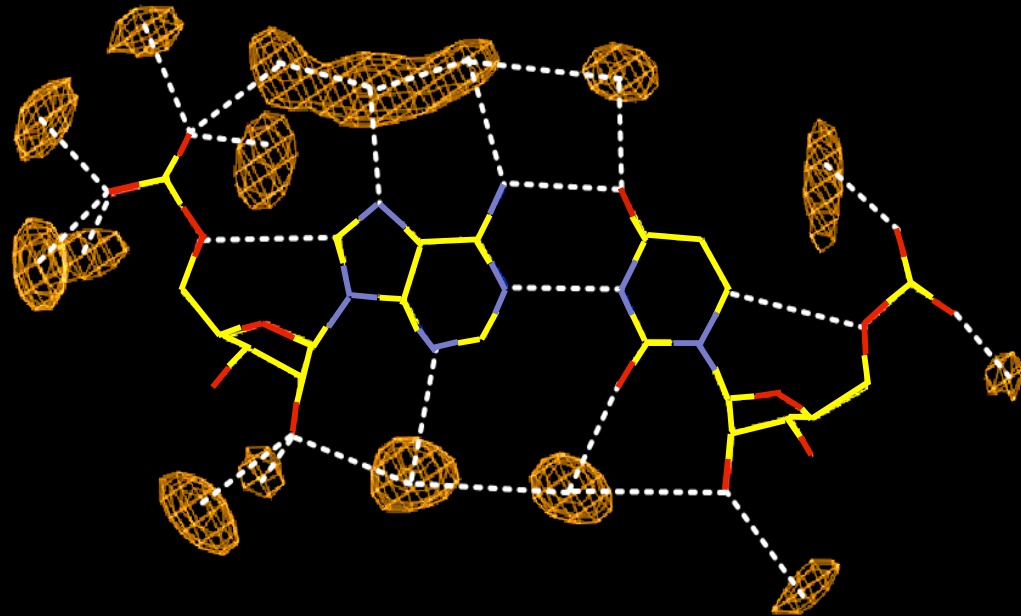


Statistical analysis of X-ray data (*A-U pairs*)

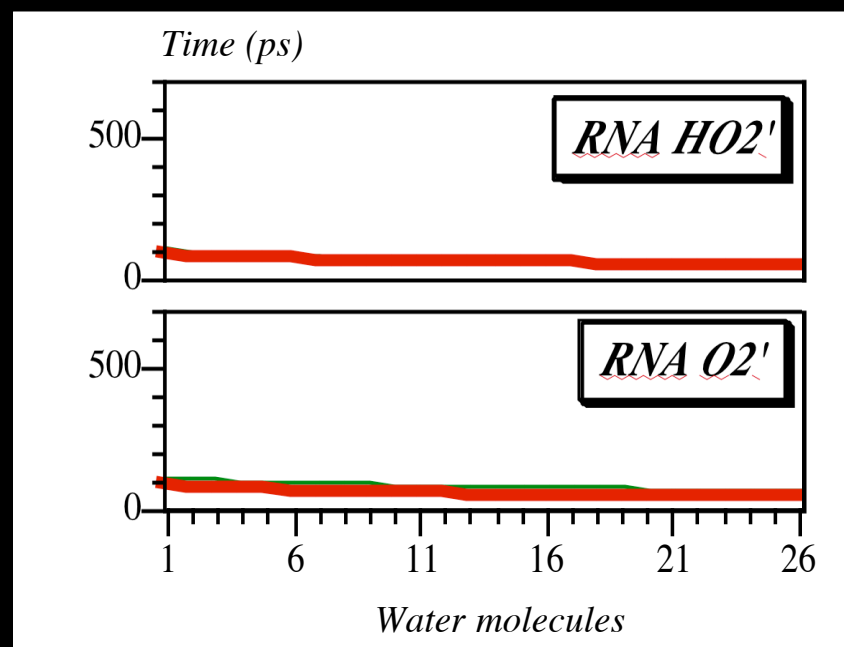
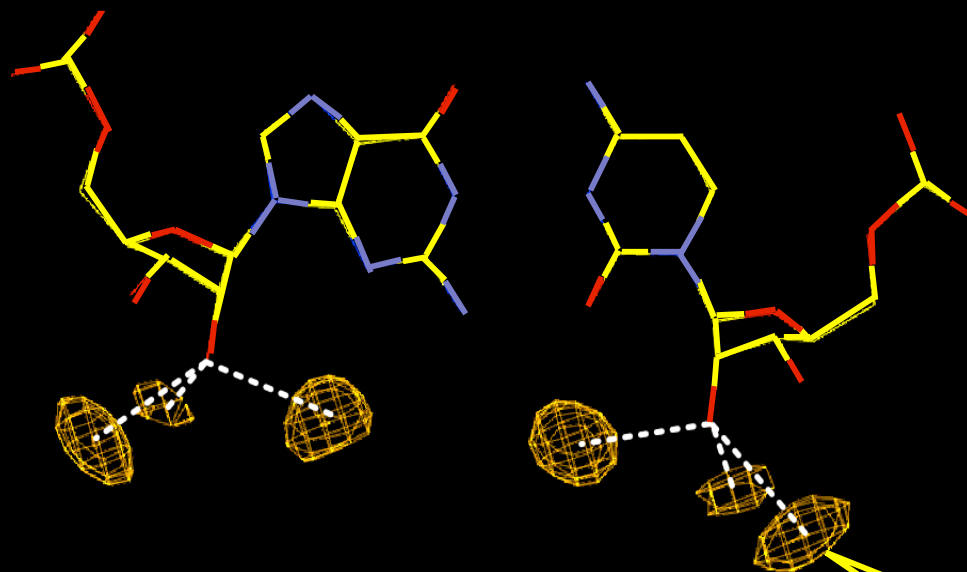
30 bp
108 waters



MD results

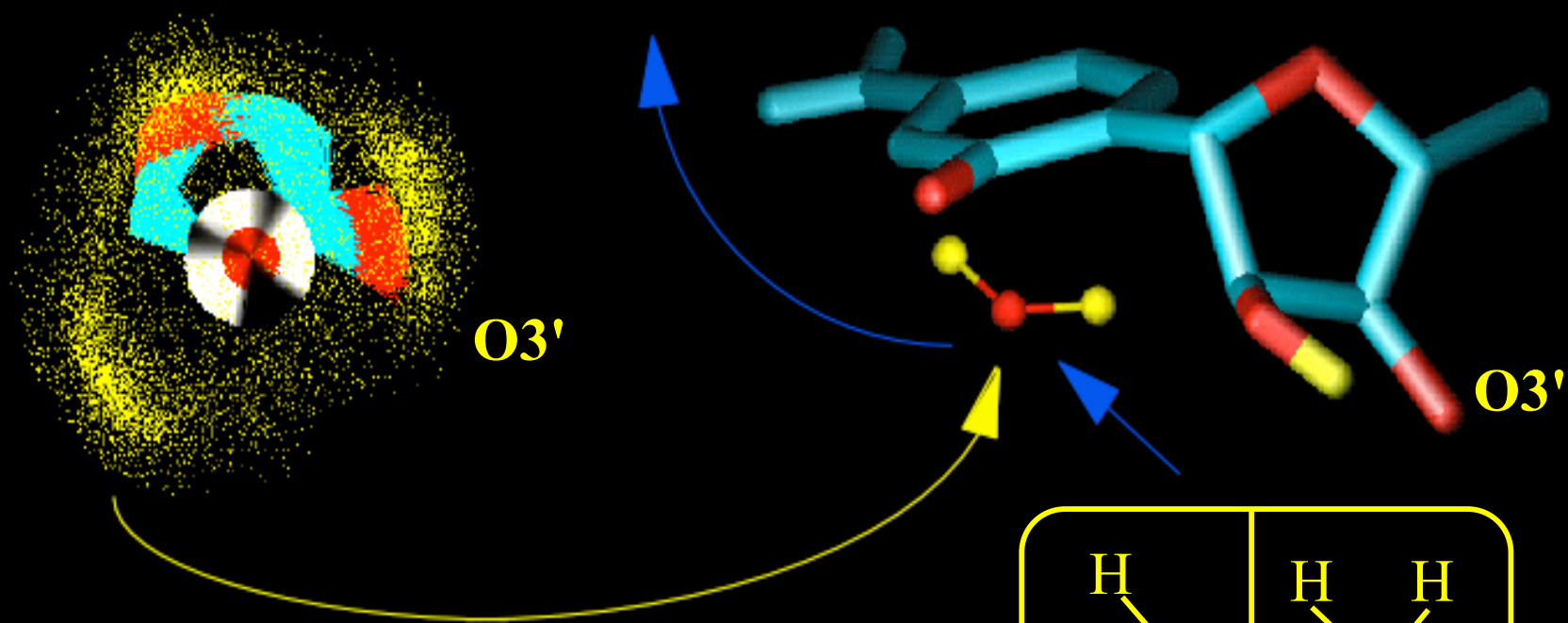


Short residence times of water molecules around 2'-OH group

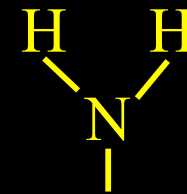


**Hydration
cones**

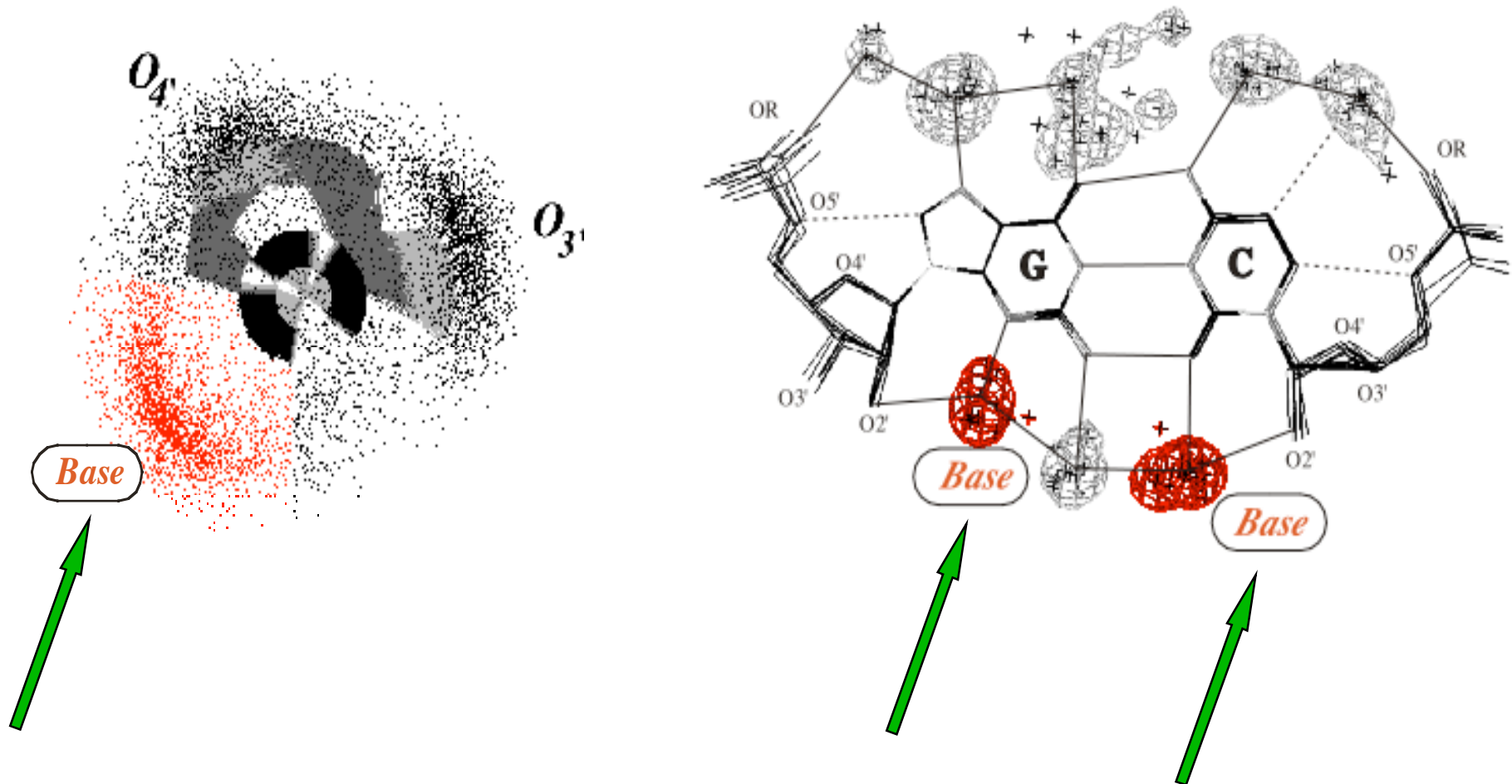
Hydration of the 2'-OH group



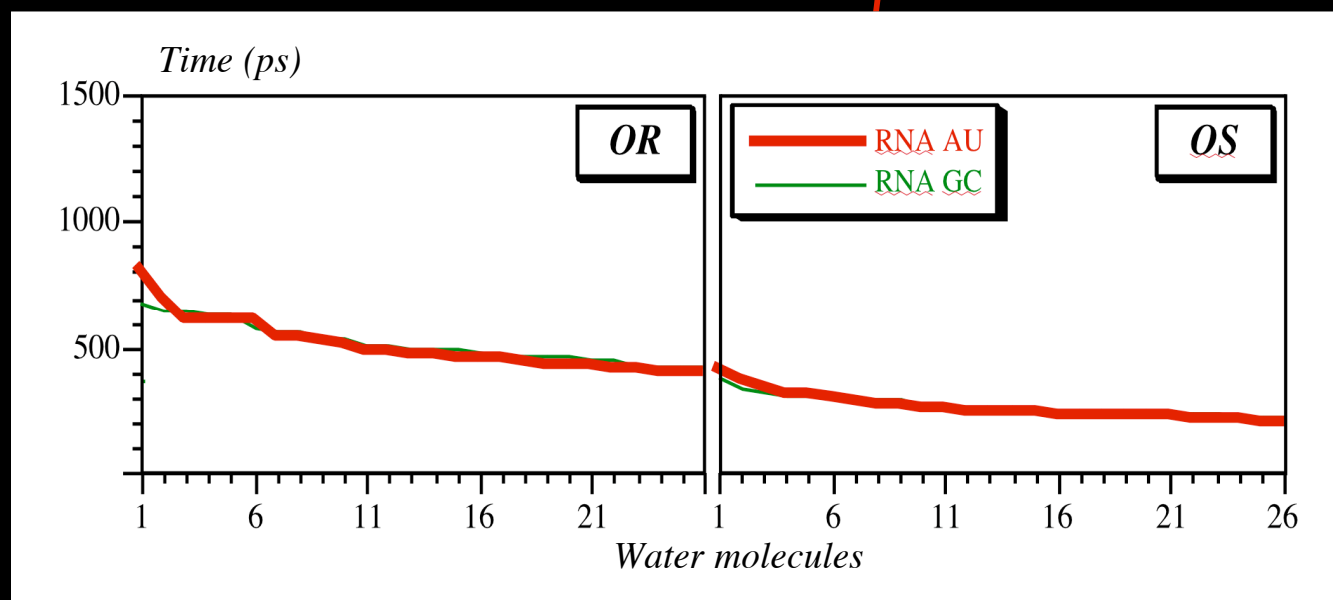
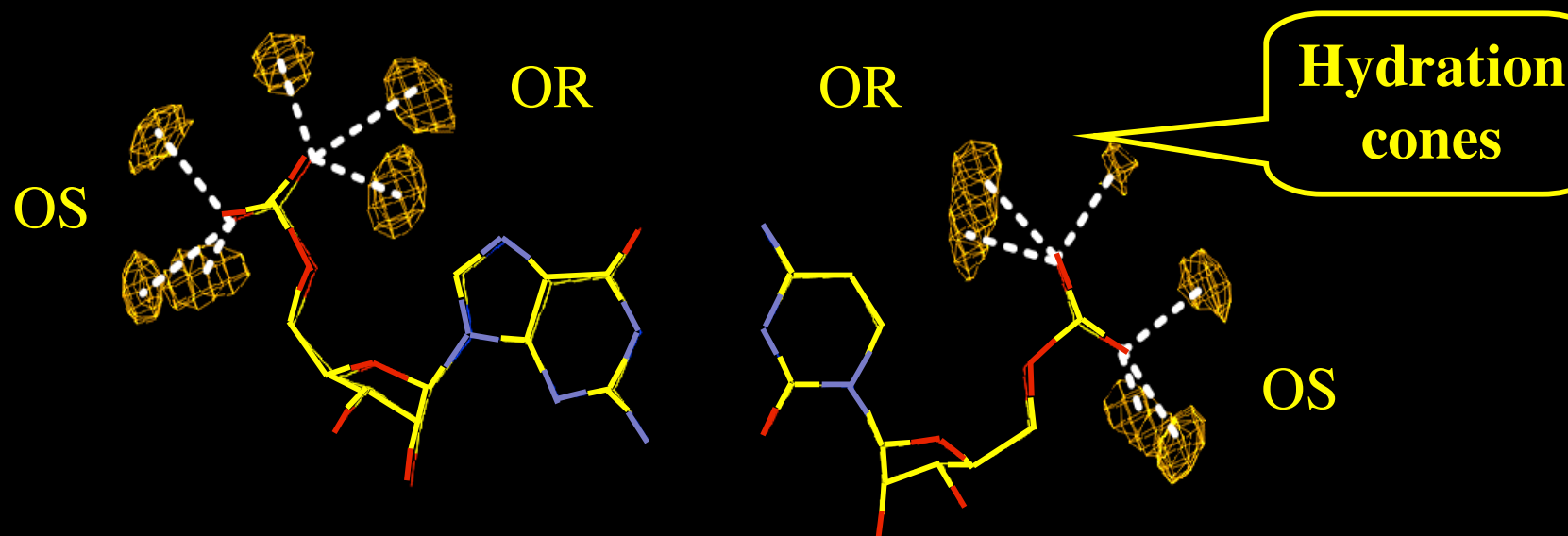
Can be easily exchanged with other groups like :



Comparison between experimental data and simulation results

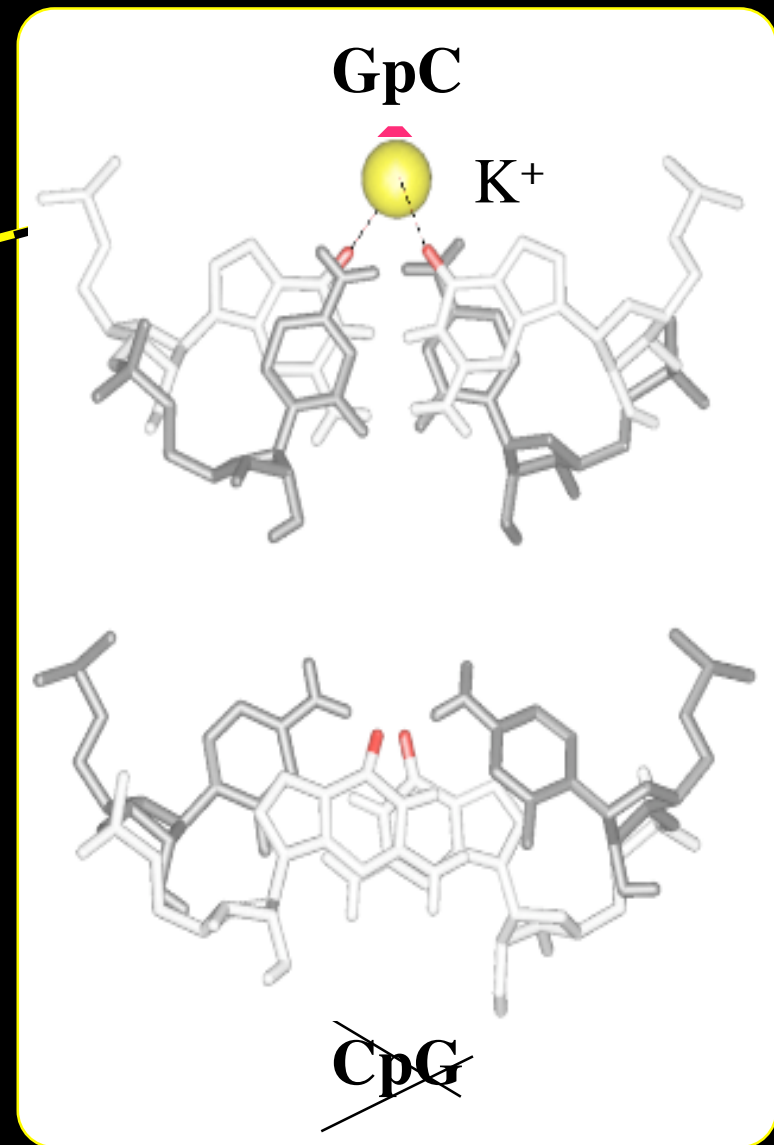
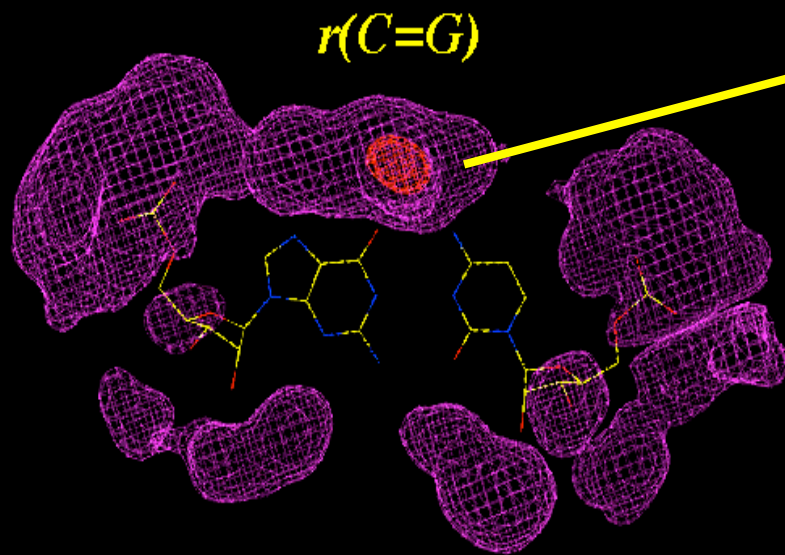


Long residence times of water molecules around phosphate

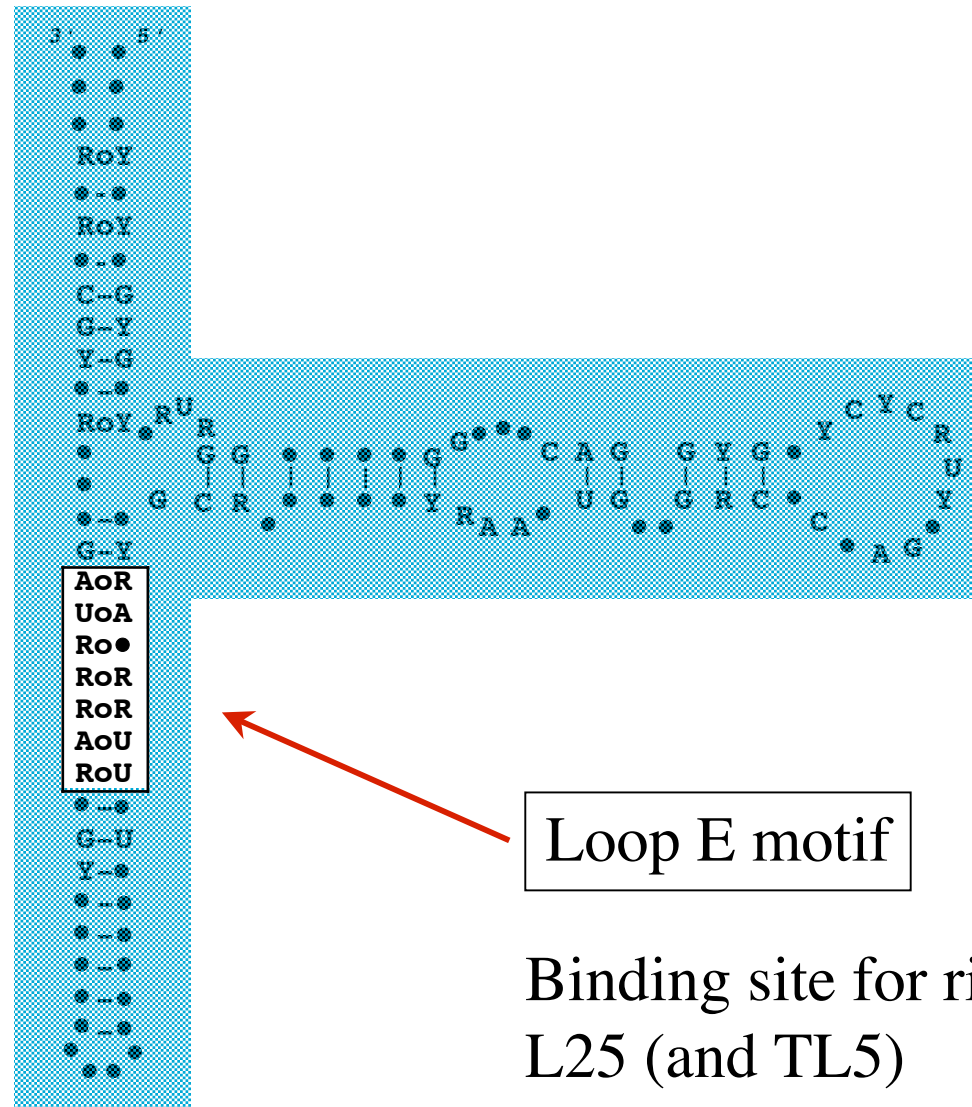


H-bond criteria : $d(\text{H}\dots\text{A}) < 2.5 \text{ \AA}$ & $\angle (\text{D}-\text{H}\dots\text{A}) > 135^\circ$

Monovalent ions penetrate the hydration shell in GpC steps



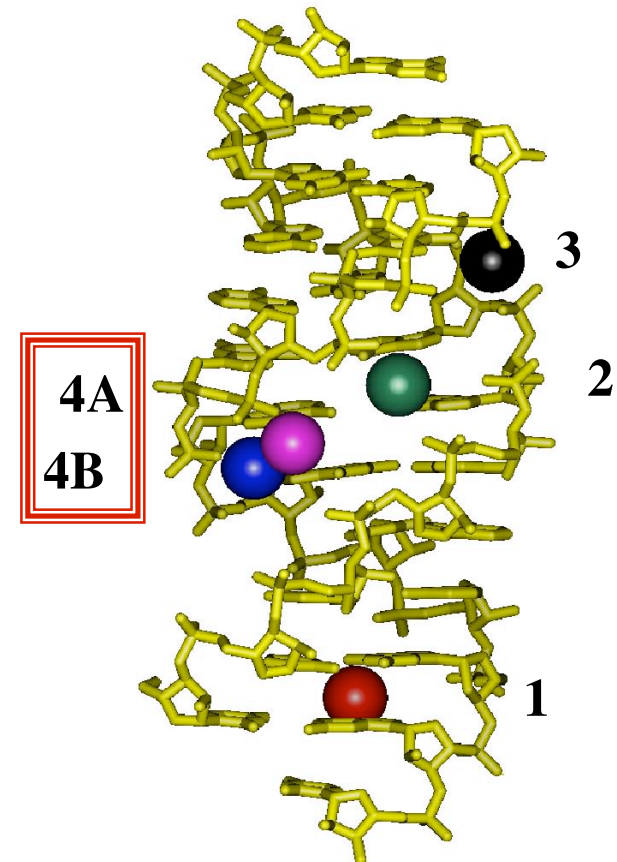
5S rRNA: eubacterial consensus sequence



The *E. coli* loop E

- Seven non-WC pairs
- Pseudosymmetric sequence
- X-ray and NMR structures available
- Stabilized by magnesium
- Five deep groove Mg^{2+} ions
- Binuclear Mg^{2+} cluster
 $d(\text{Mg}^{2+} \dots \text{Mg}^{2+}) \approx 2.7 \text{ \AA}$

G
 G $\circ$ U
 C=G
 G $\leftarrow$ A
 A $\rightarrow$ U
 G $\oplus$ G
 A $\oplus$ G
 G $\oplus$ U
 U $\circ$ A
 A $\rightarrow$ G
 dG=C
 G=C
 C



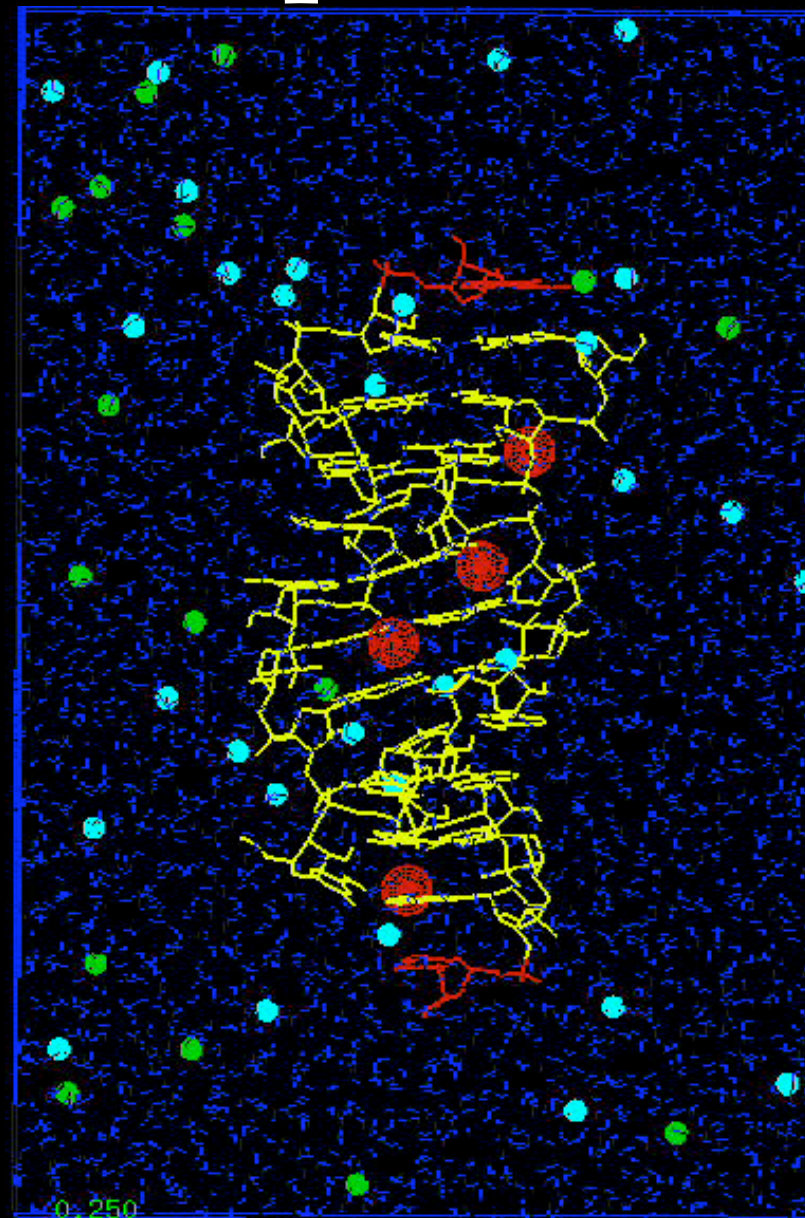
Correll *et al.* **Cell** (1997)
 (res. 1.5 Å)

Simulations setup

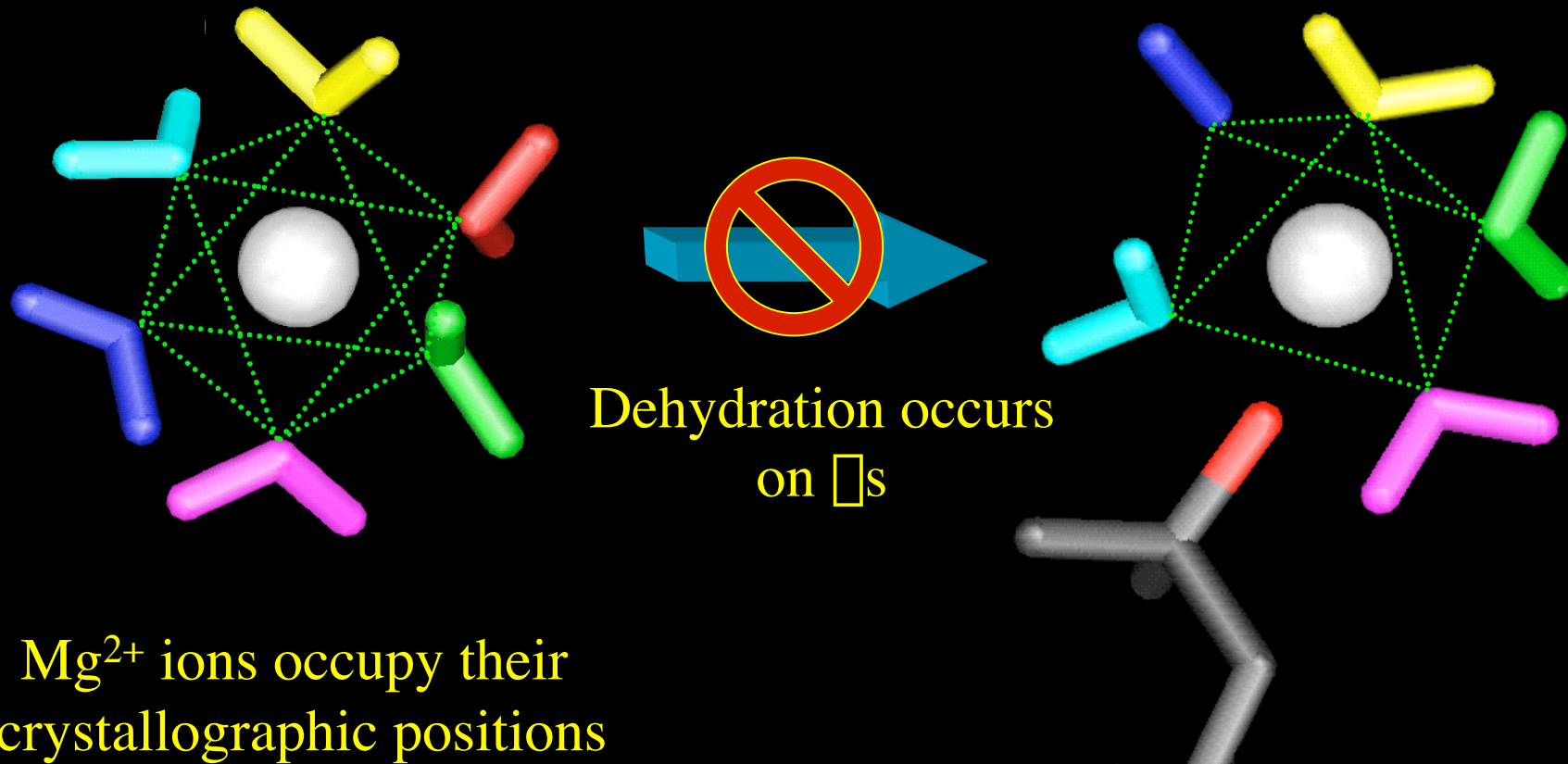
AMBER program and force-field
with PME

10 ns trajectories with and without Mg^{2+}

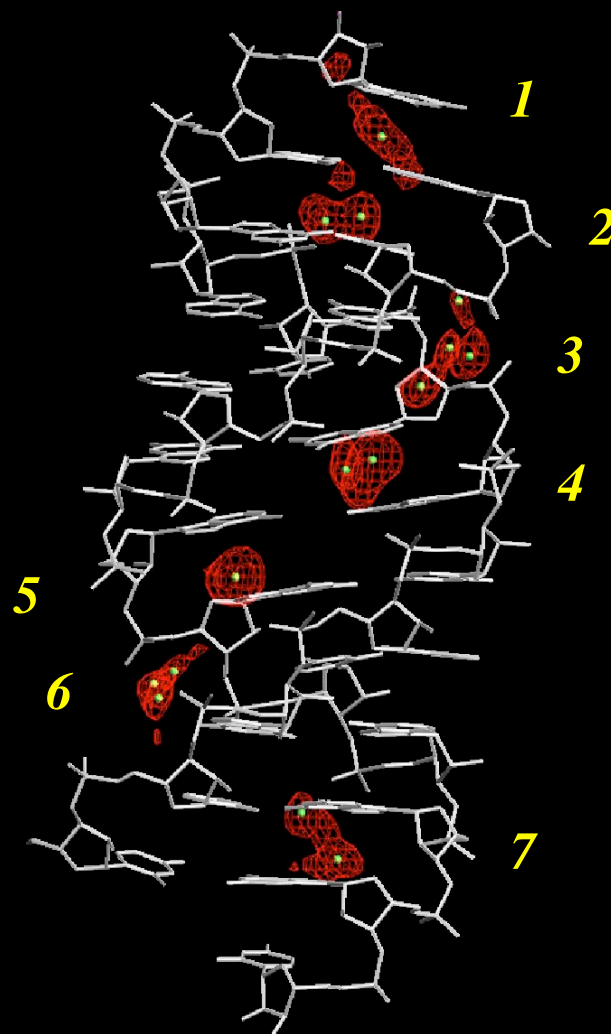
- 11 base pairs
(+2 *nucleotides*)
- ≈ 0.2 M added KCl :
 - 4 Mg^{2+} (*red*)
 - 30 K^{+} (*cyan*)
 - 16 Cl^{-} (*green*)
- ≈ 4000 H_2O



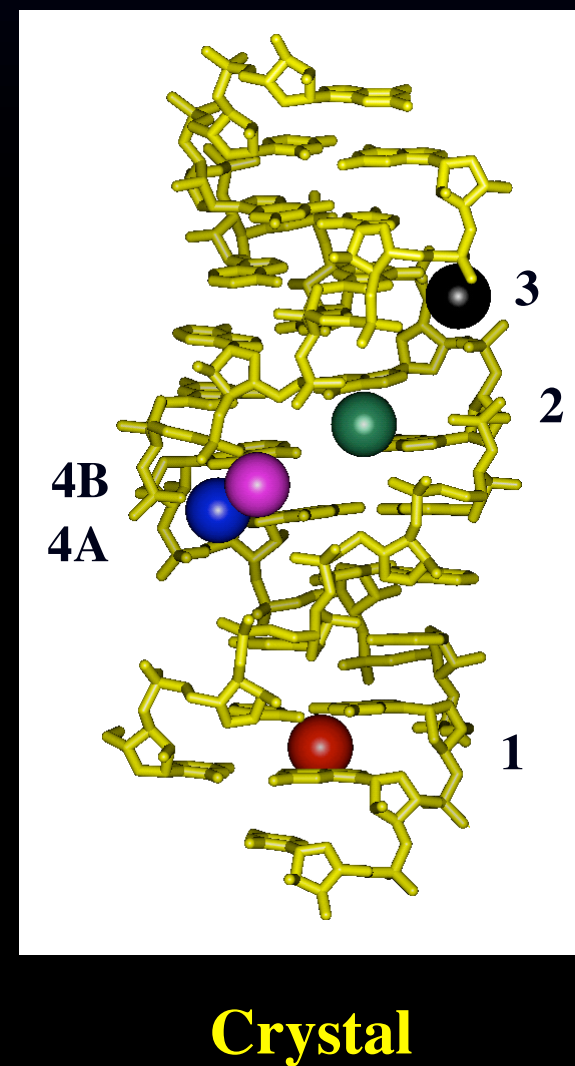
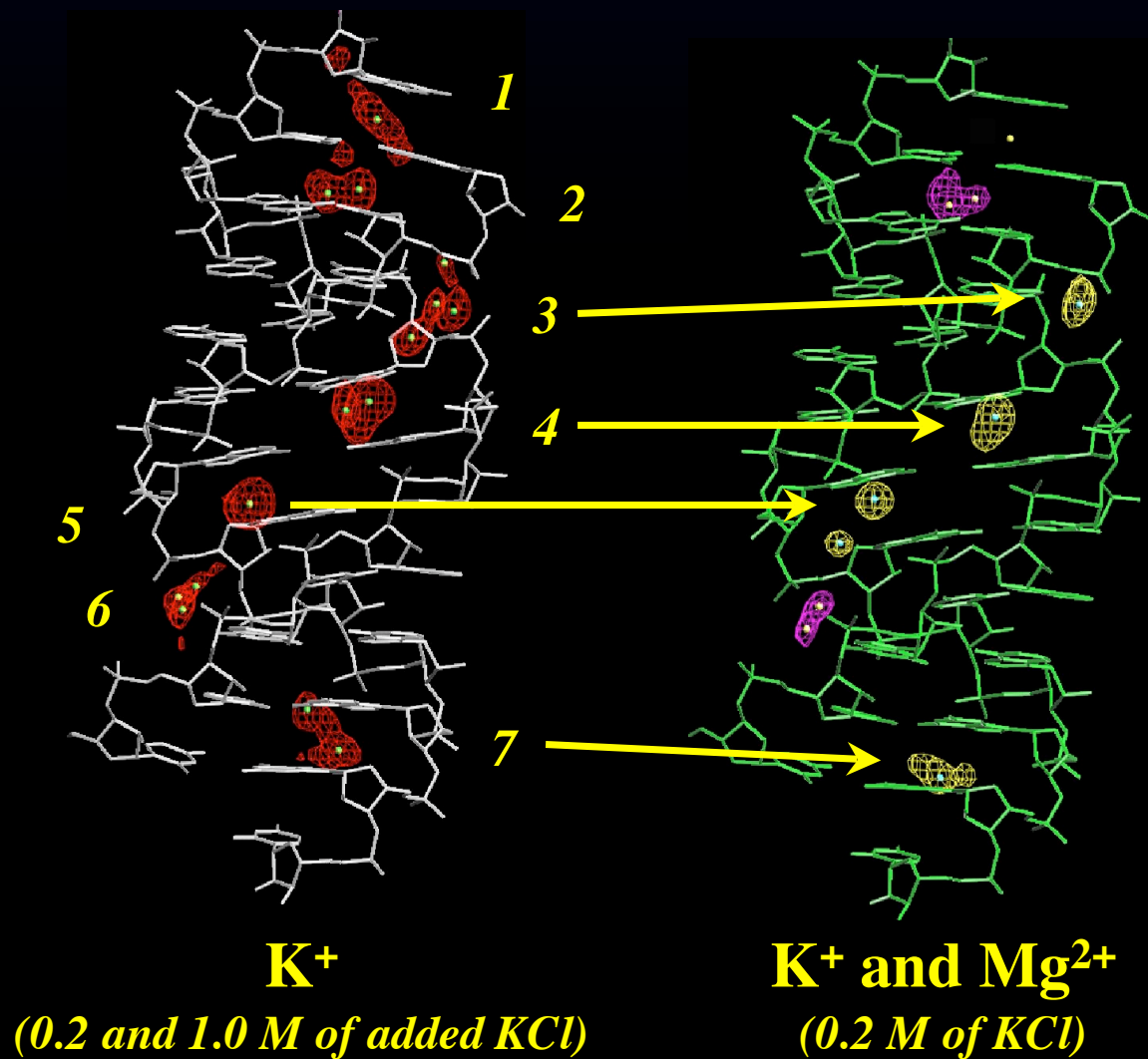
Mg^{2+} initial positions



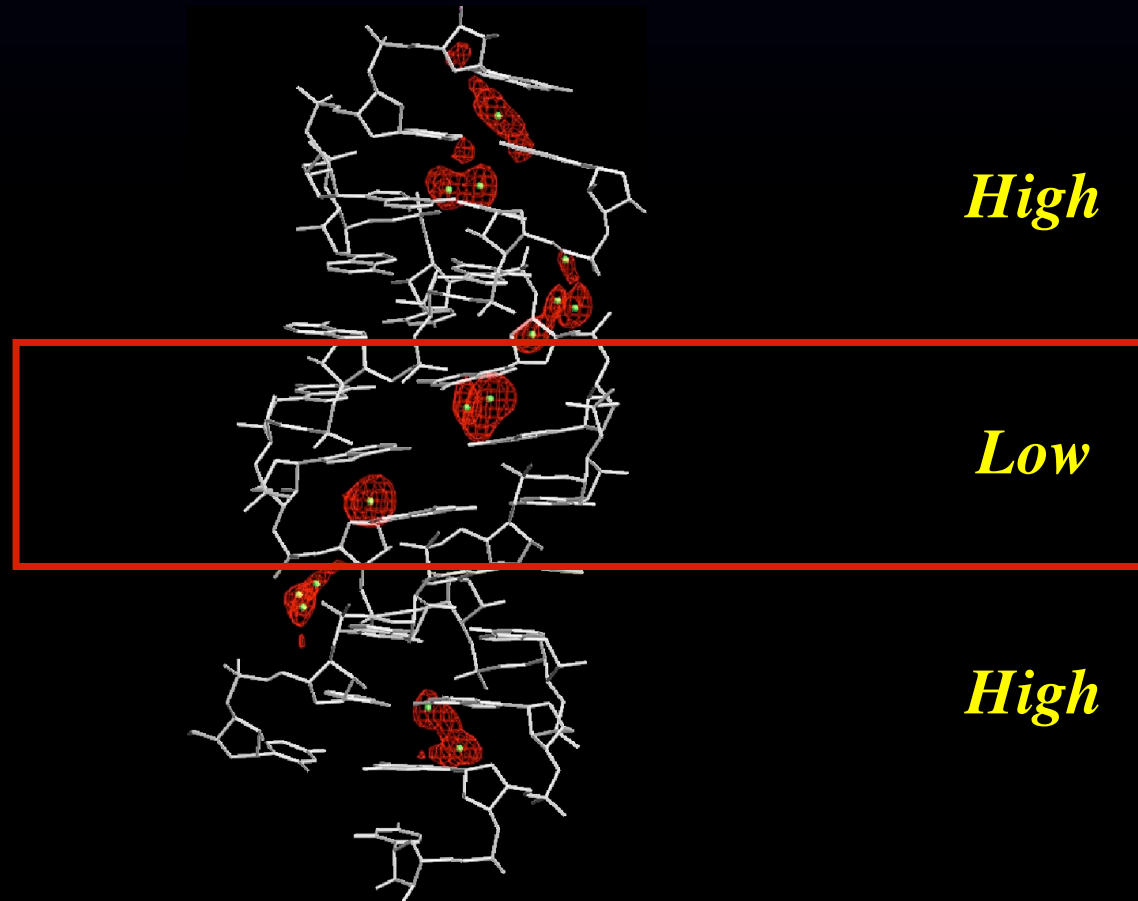
K^+ initial positions



Ion densities

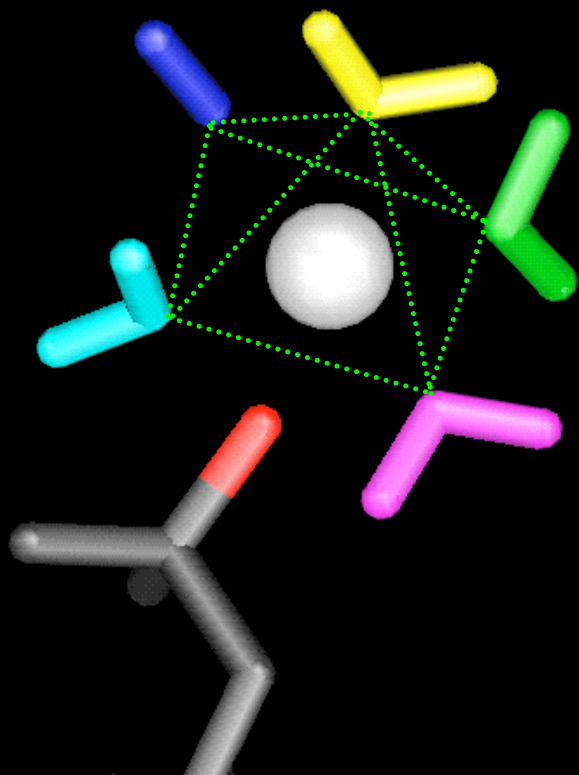


K^+ exchange rates

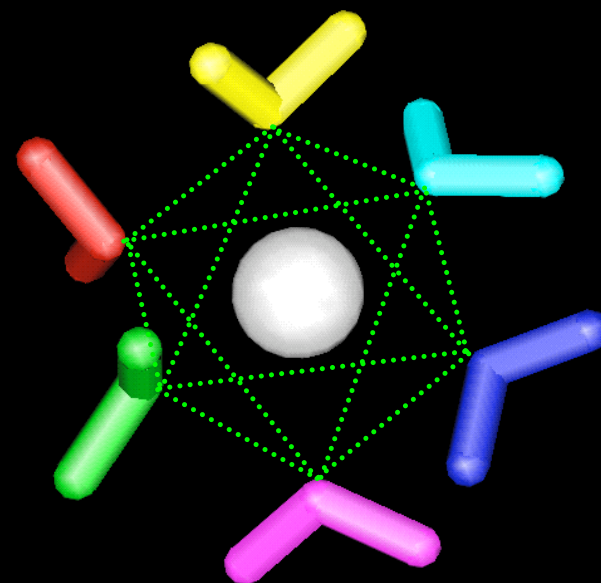


Middle ions are more strongly bound

Two types of hydrated Mg^{2+}



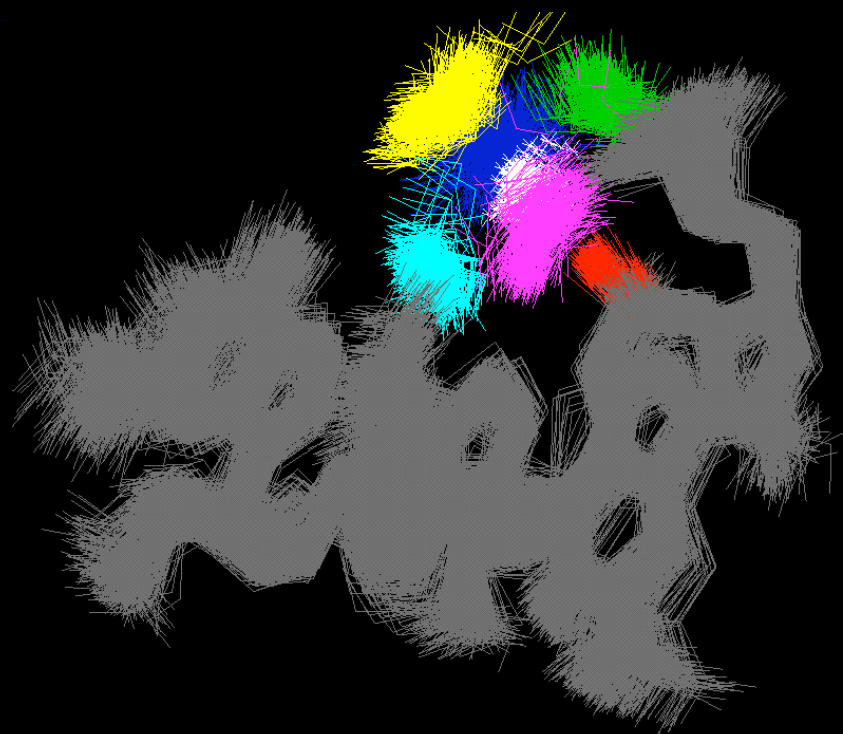
Pentahydrated



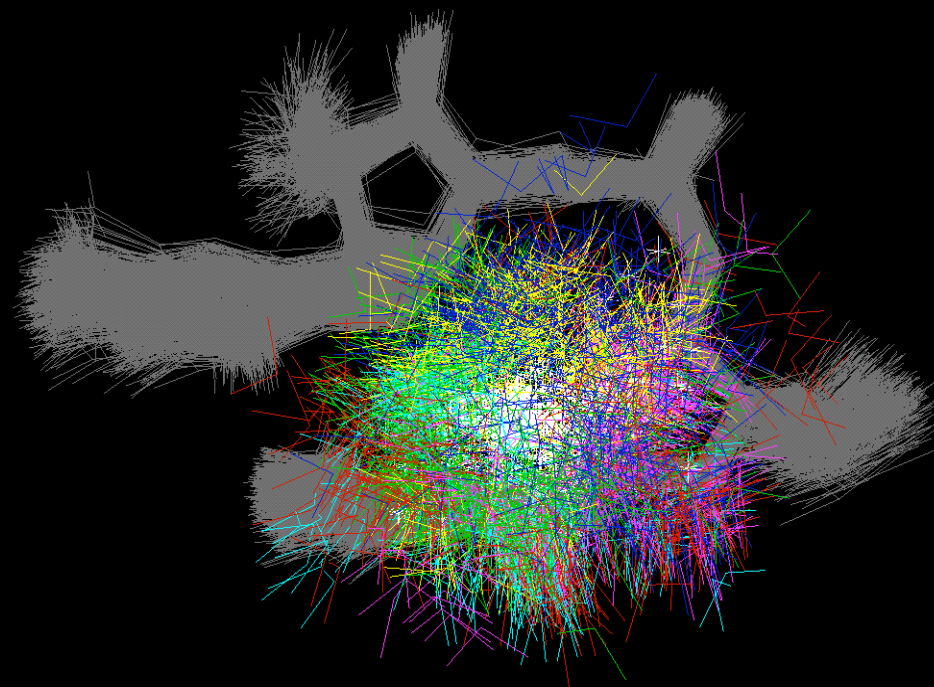
Hexahydrated



Dynamics of hydrated Mg^{2+} ions

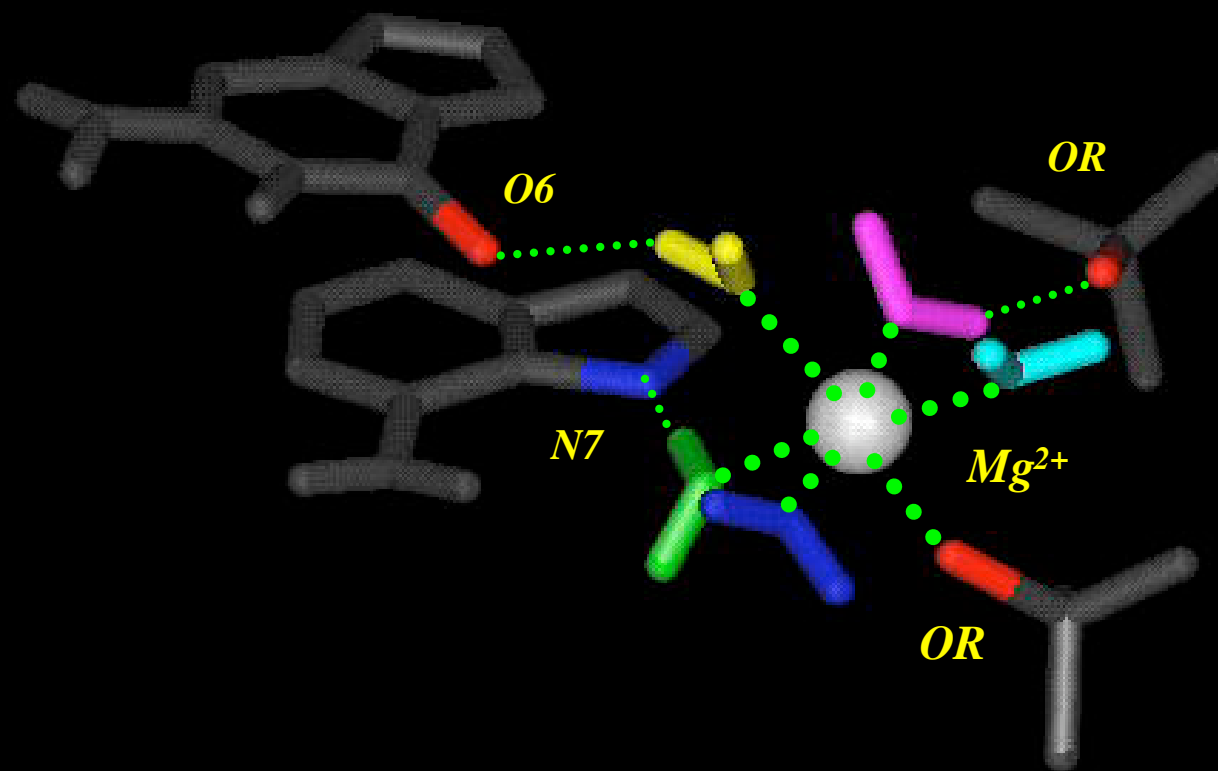


Small mobility: $\text{Mg}(\text{H}_2\text{O})_5^{2+}$



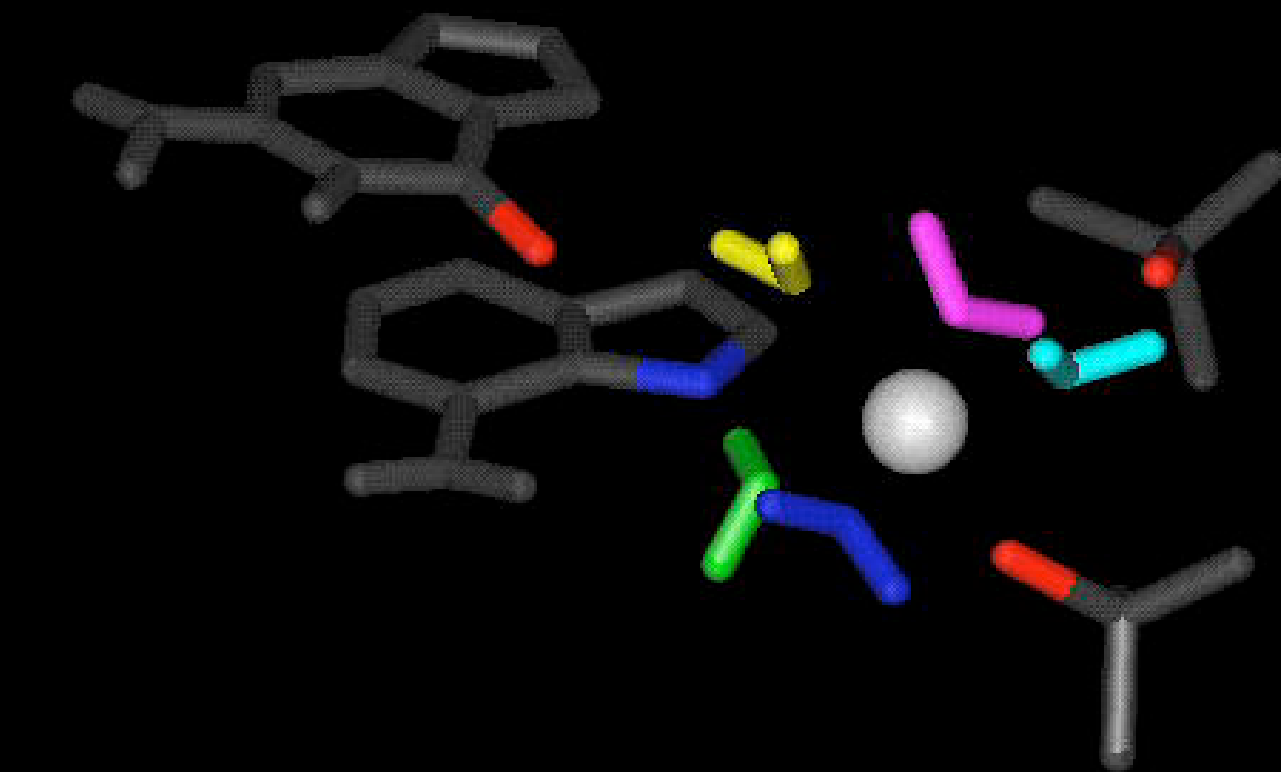
Large mobility: $\text{Mg}(\text{H}_2\text{O})_6^{2+}$

Dynamics of first shell H₂O



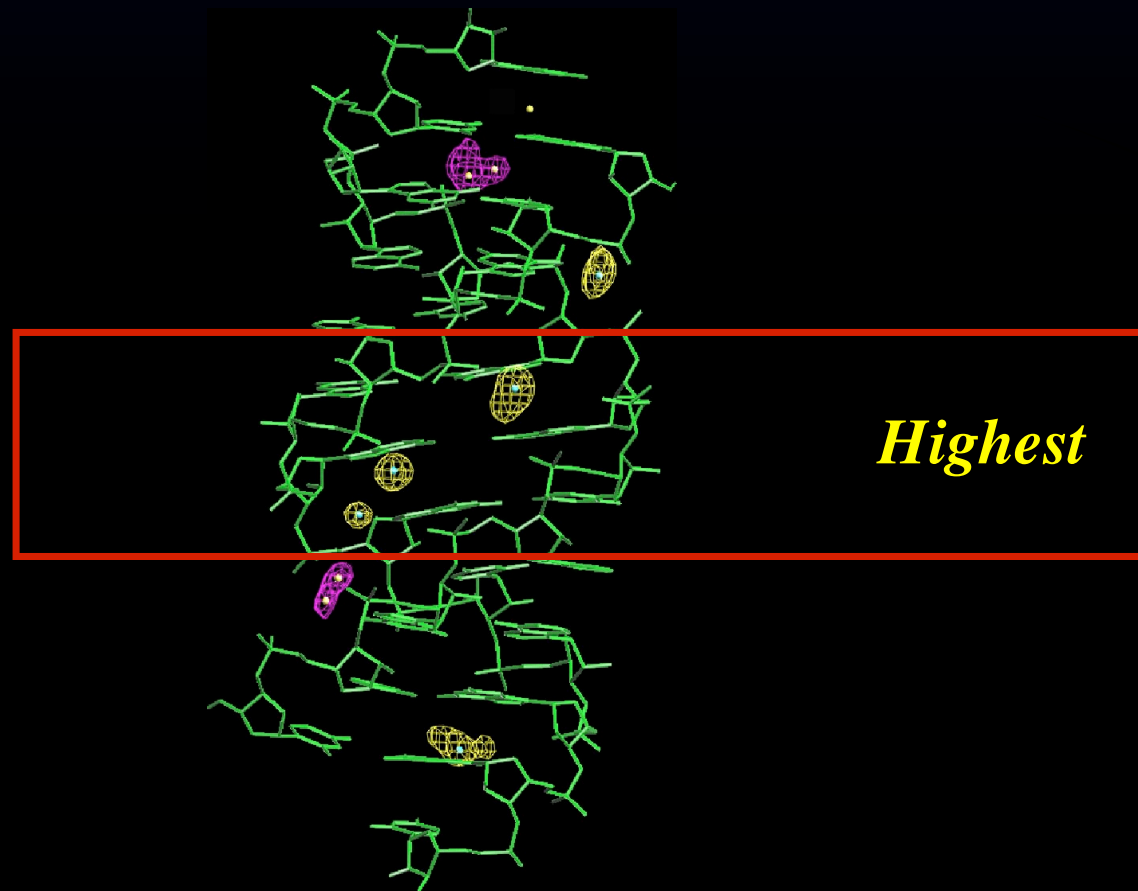
First shell interactions

Dynamics of first shell H₂O



1500,500

Structural complementarity for second shell interactions

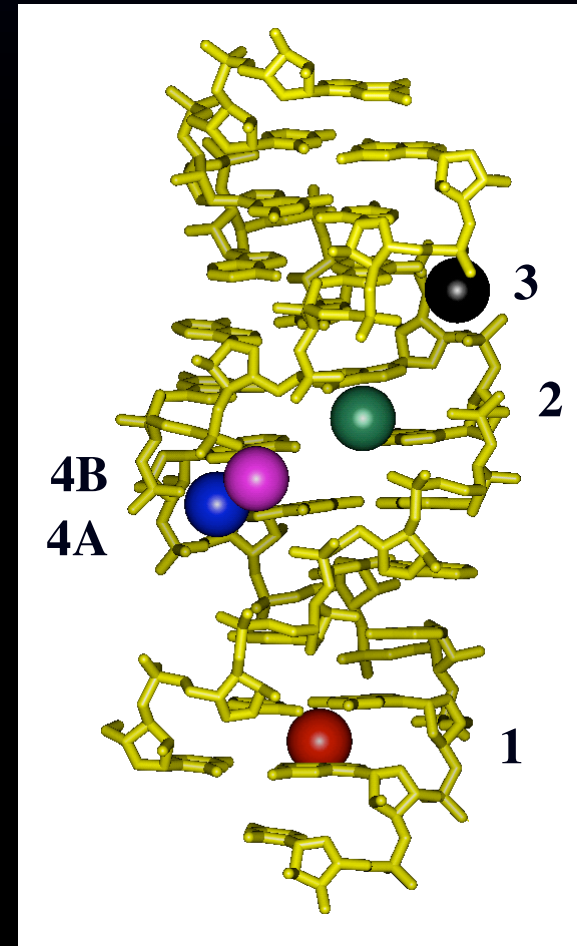


K⁺ and Mg²⁺
(0.2 M of KCl)

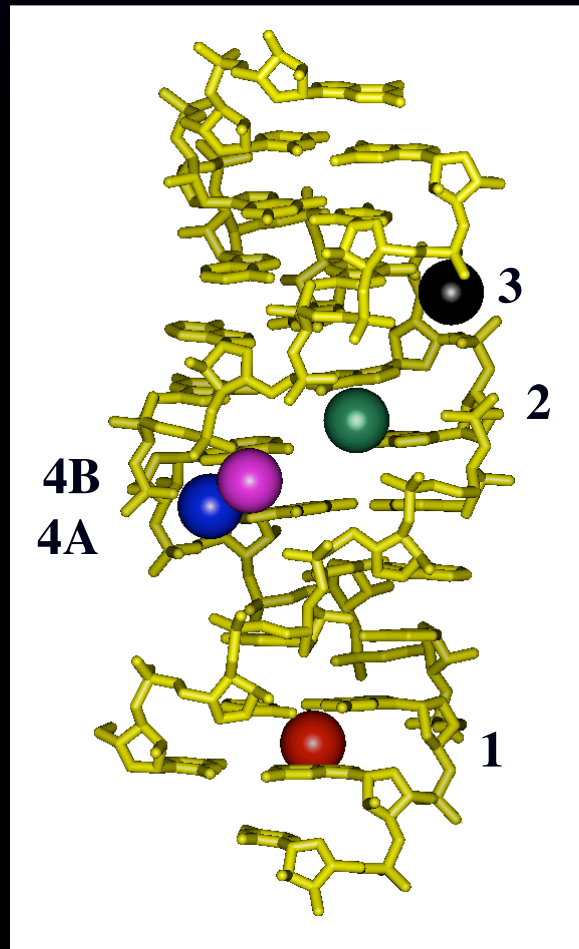
Are all Mg^{2+} ions equally important ?

UV melting study suggest that only 1 or 2 Mg^{2+} ions are strongly bound

(Serra *et al.*, *RNA*, 2003)



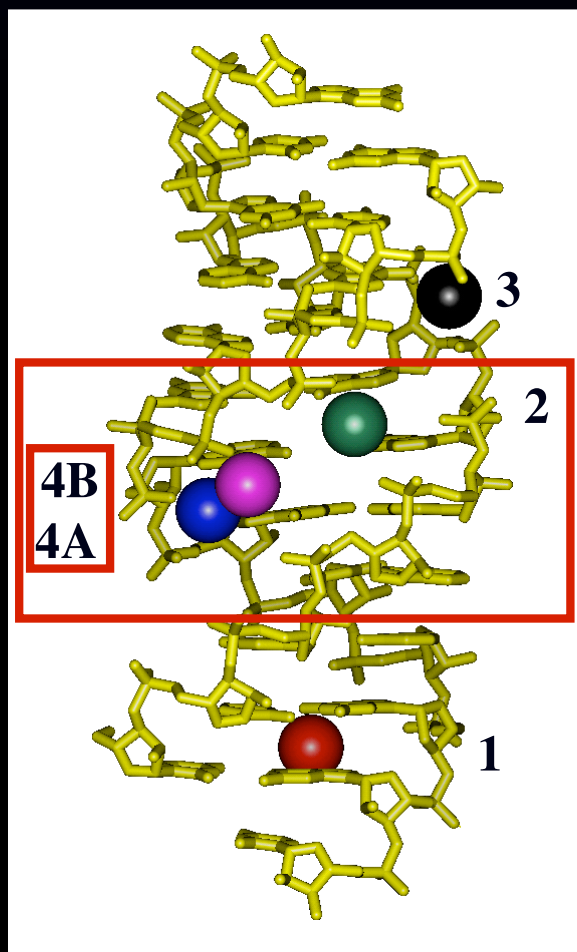
Are all Mg^{2+} ions equally important ?



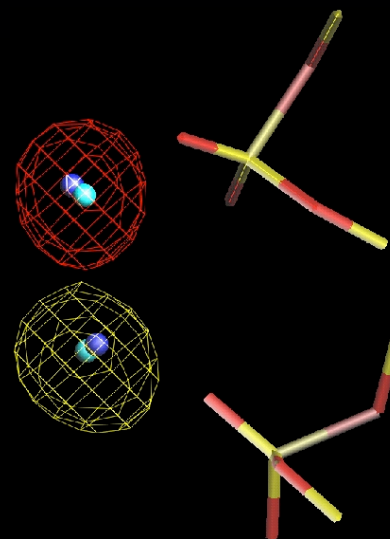
Ion 3 : related to lattice interactions
(non-specific)

Ion 1 : associated with a Watson-Crick GpG step
(non-specific)

The binuclear Mg^{2+} cluster

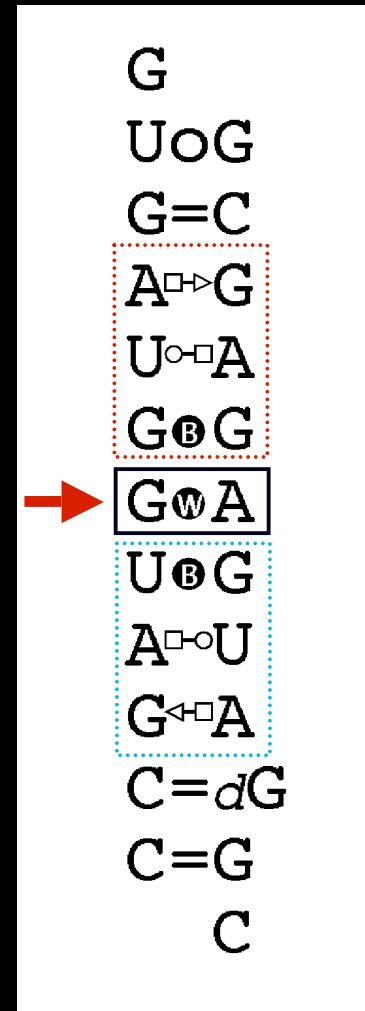
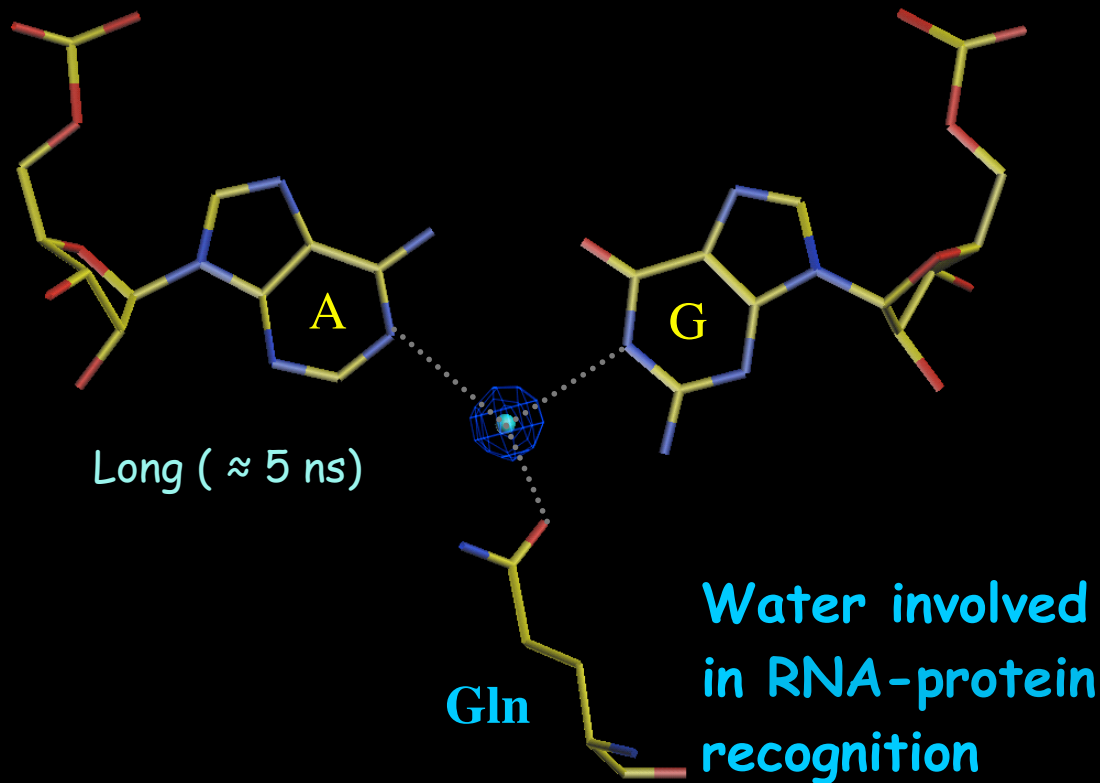


Two simulations were undertaken

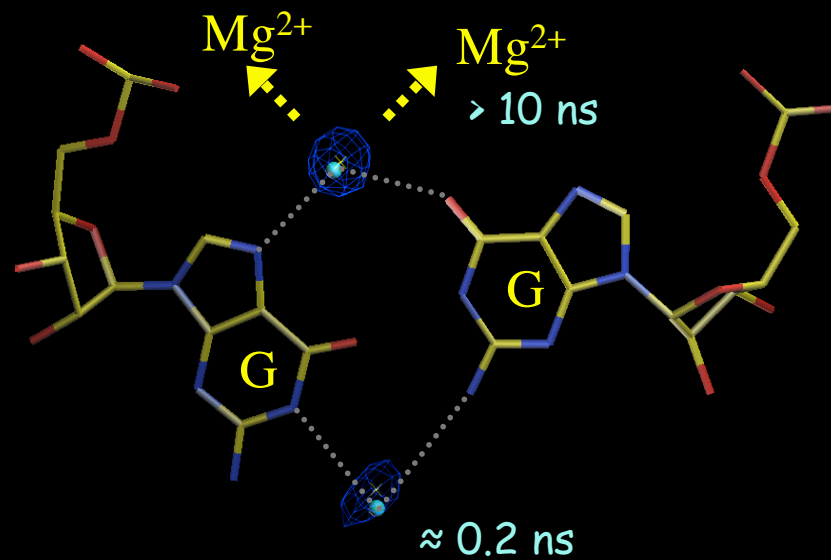


Fractional occupancies well
reproduce experimental densities

Central G•A pair



G•G pair



G

UoG

G=C

A $\square$ $\triangleright$ G

Uo $\square$ A

$\rightarrow$ G $\oplus$ G

G $\oplus$ A

U $\oplus$ G

A $\square$ oU

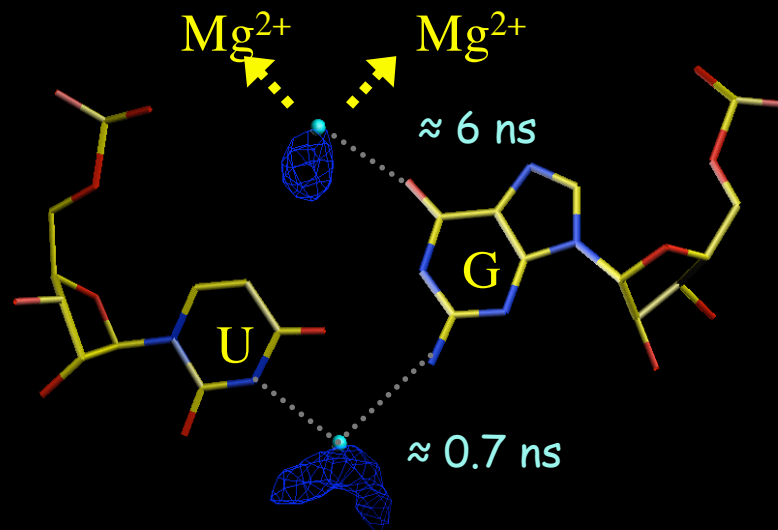
G $\leftarrow$ $\square$ A

C=dG

C=G

C

G•U pair

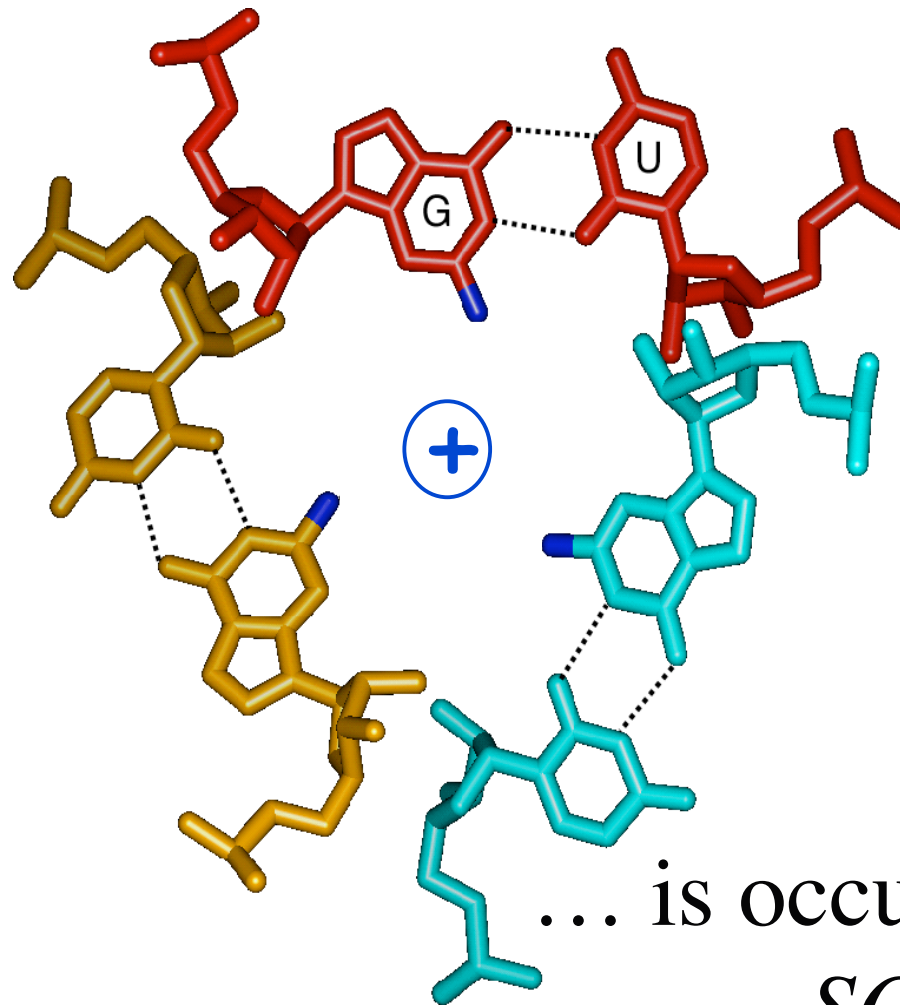


G
 UoG
 G=C
 A $\square$ $\triangleright$ G
 Uo $\square$ A
 G $\oplus$ G
 G $\oplus$ A
 U $\oplus$ G
 A $\square$ oU
 G $\leftarrow$ $\square$ A
 C=dG
 C=G
 C

*Anion coordination geometry
for the Cl^- and SO_4^{2-} anions*

An electropositive cavity in a RNA crystal

...

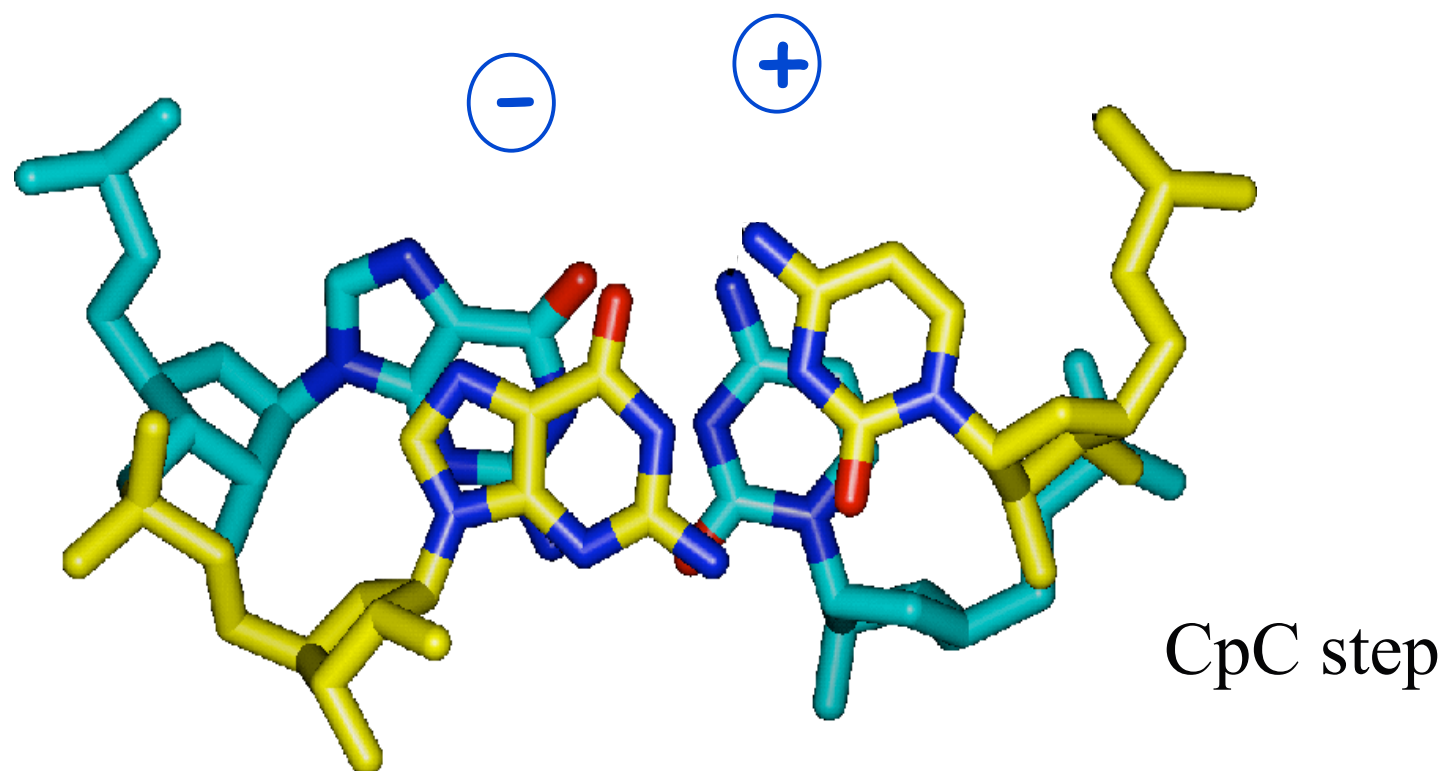


... is occupied by a
 SO_4^{2-} anion

0.98 Å

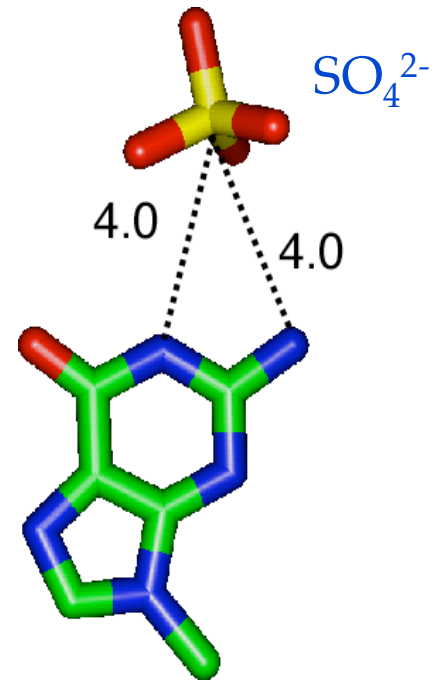
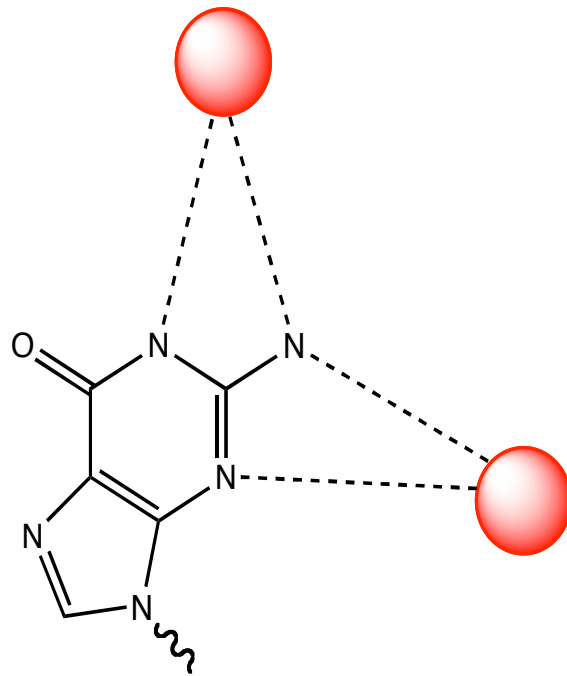
Masquida & Westhof, RNA, 1999

Even deep/major grooves ...

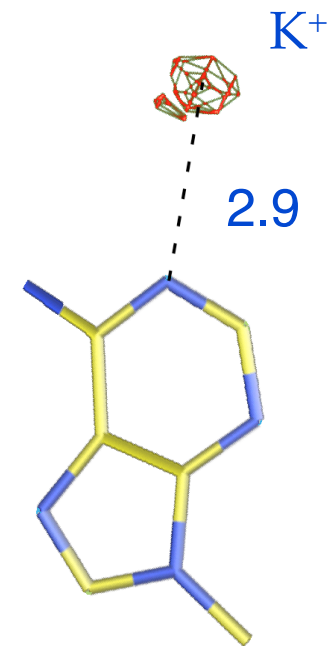
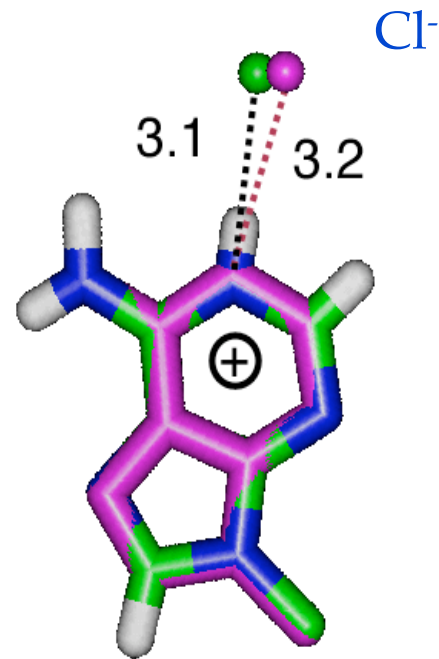
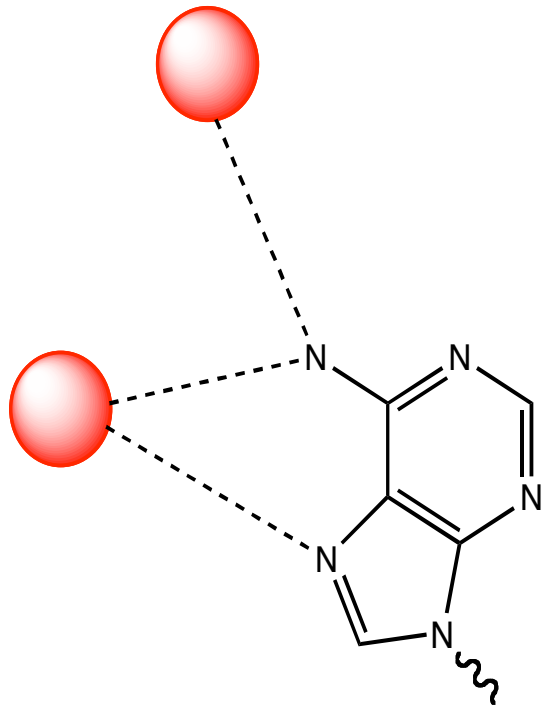


... can accommodate anions

Guanine binding sites

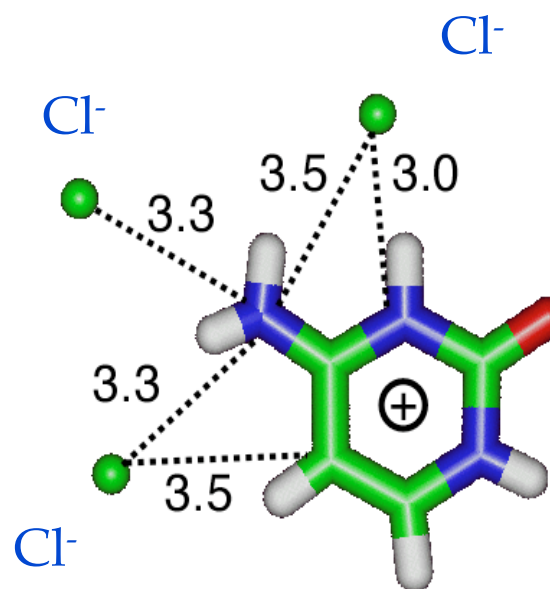
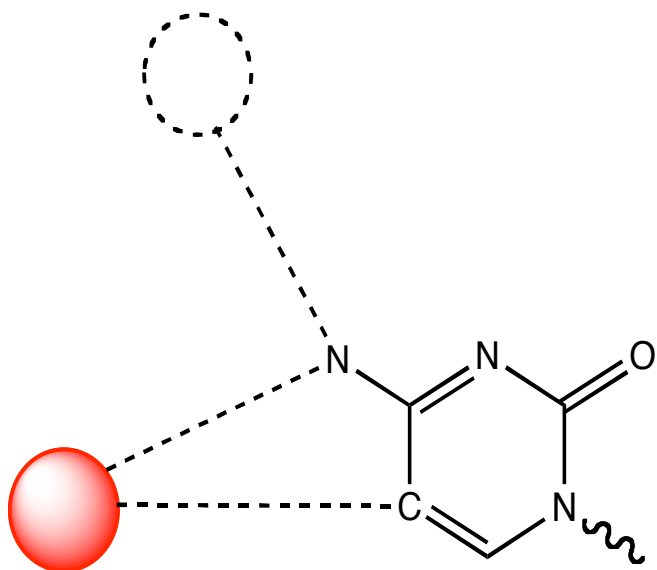


Adenine binding sites



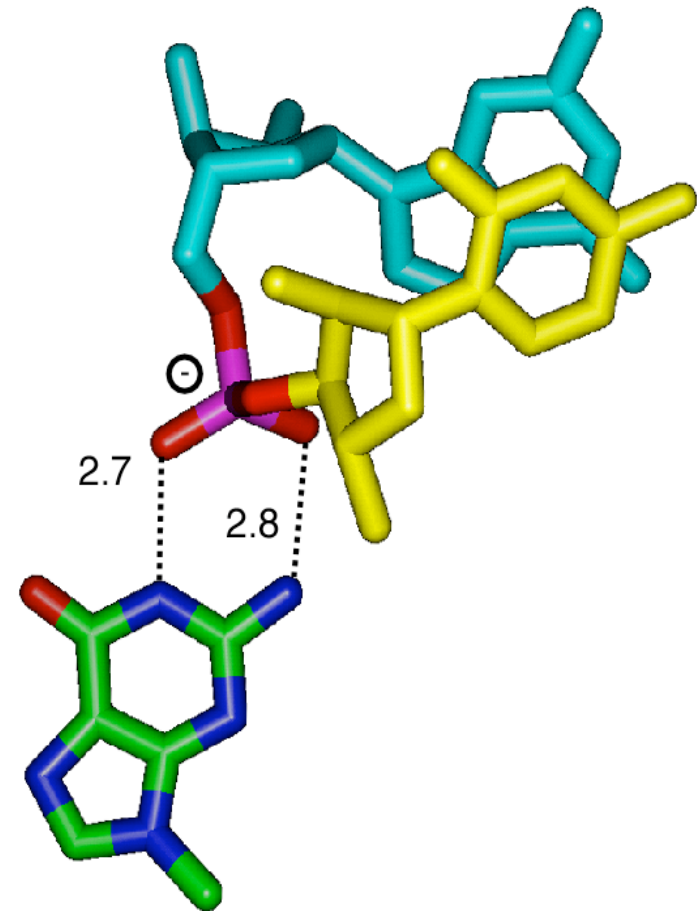
MD results

Cytosine binding sites



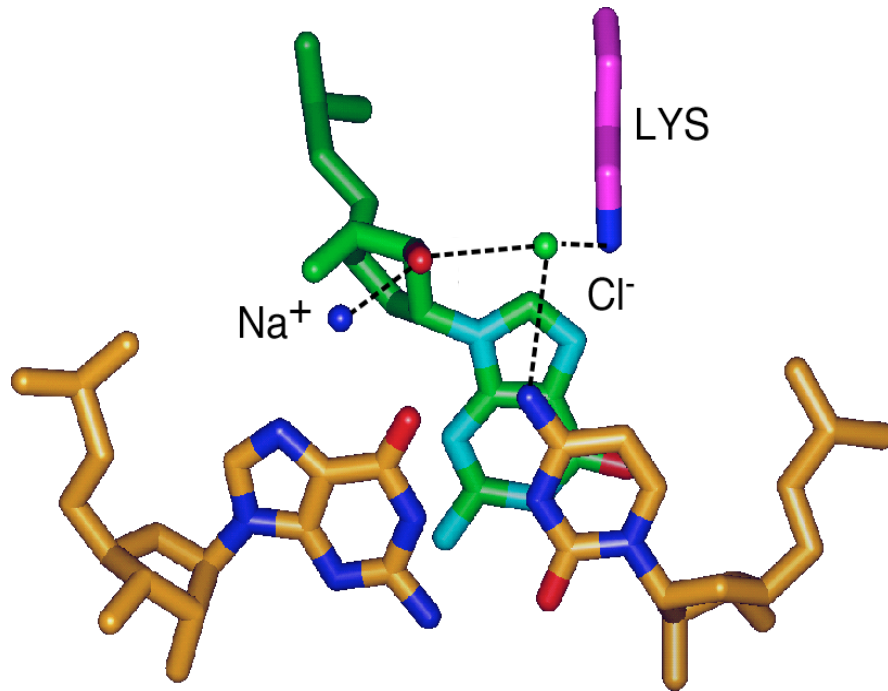
*(Cambridge Structural Database
or CSD)*

RNA/RNA interactions

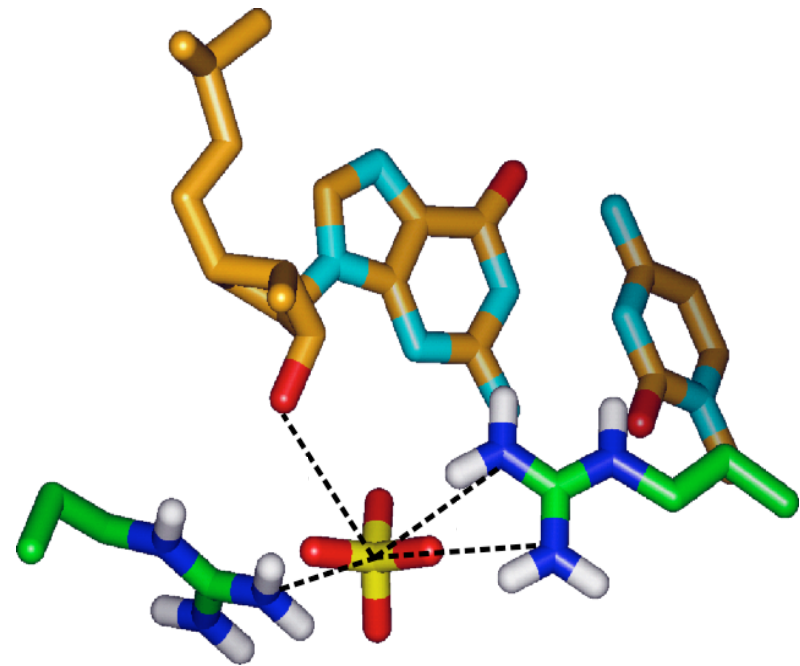


ARN / ARN

RNA/protein interactions mediated by anions

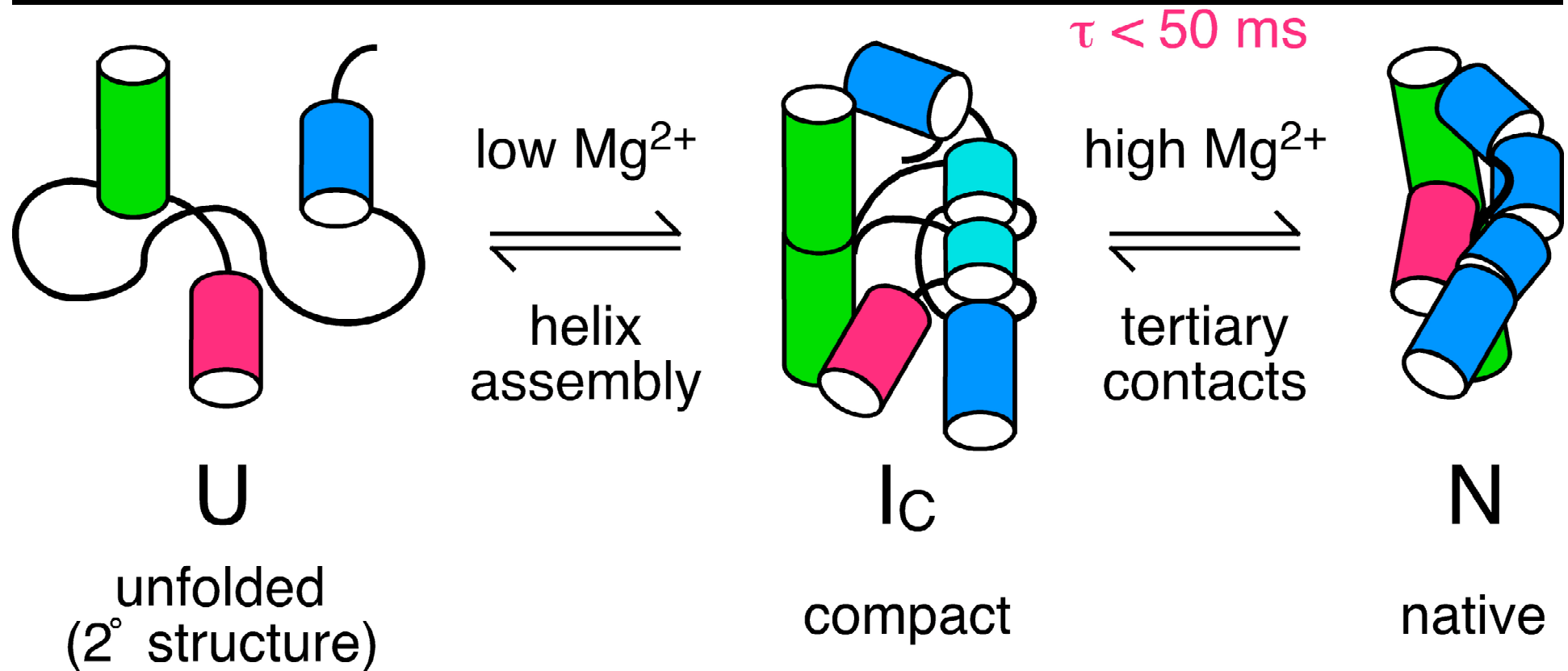


ARN / Cl^- / Protein



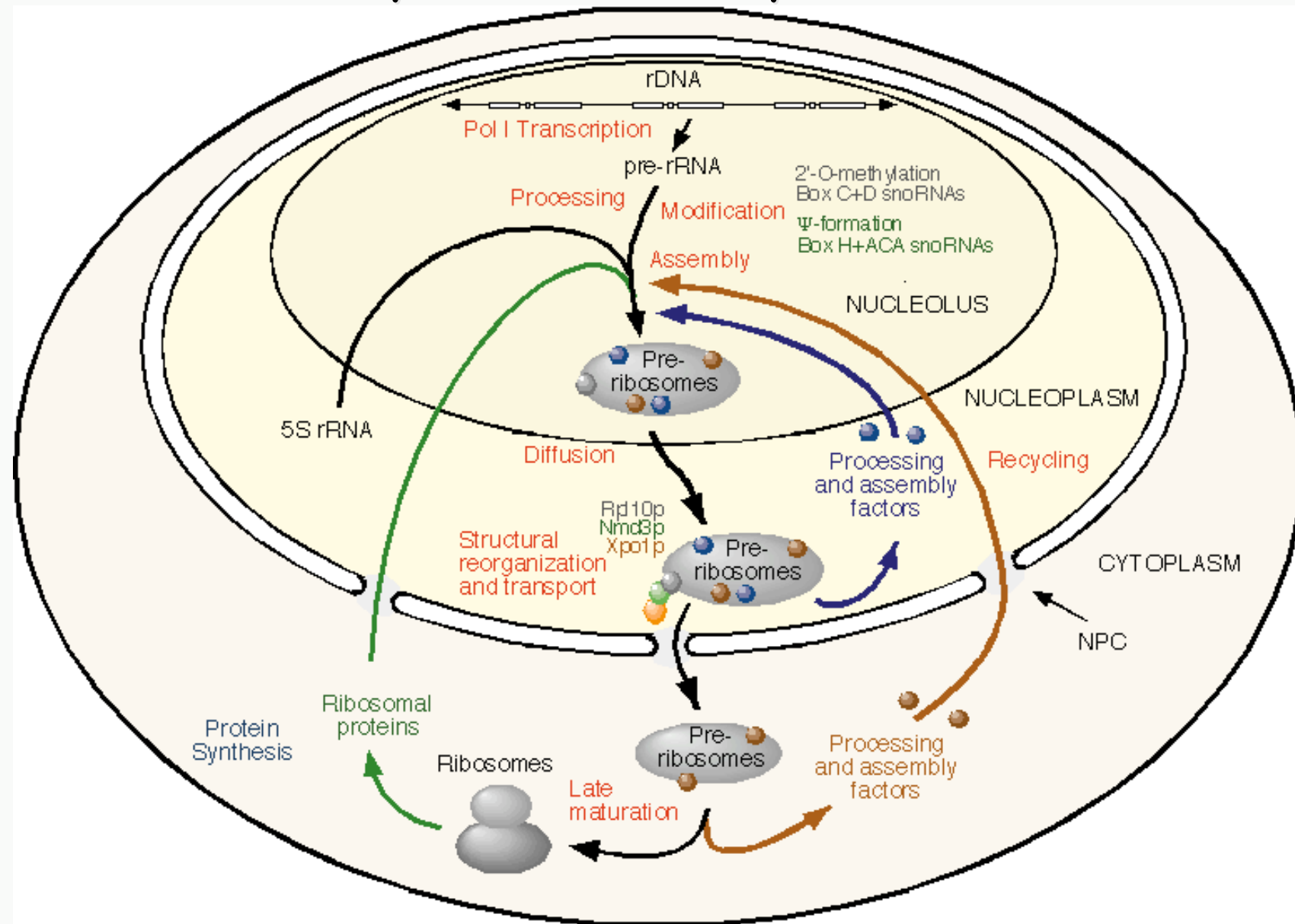
ARN / SO_4^{2-} / Protein

Coupling between architectural and electrostatic hierarchies

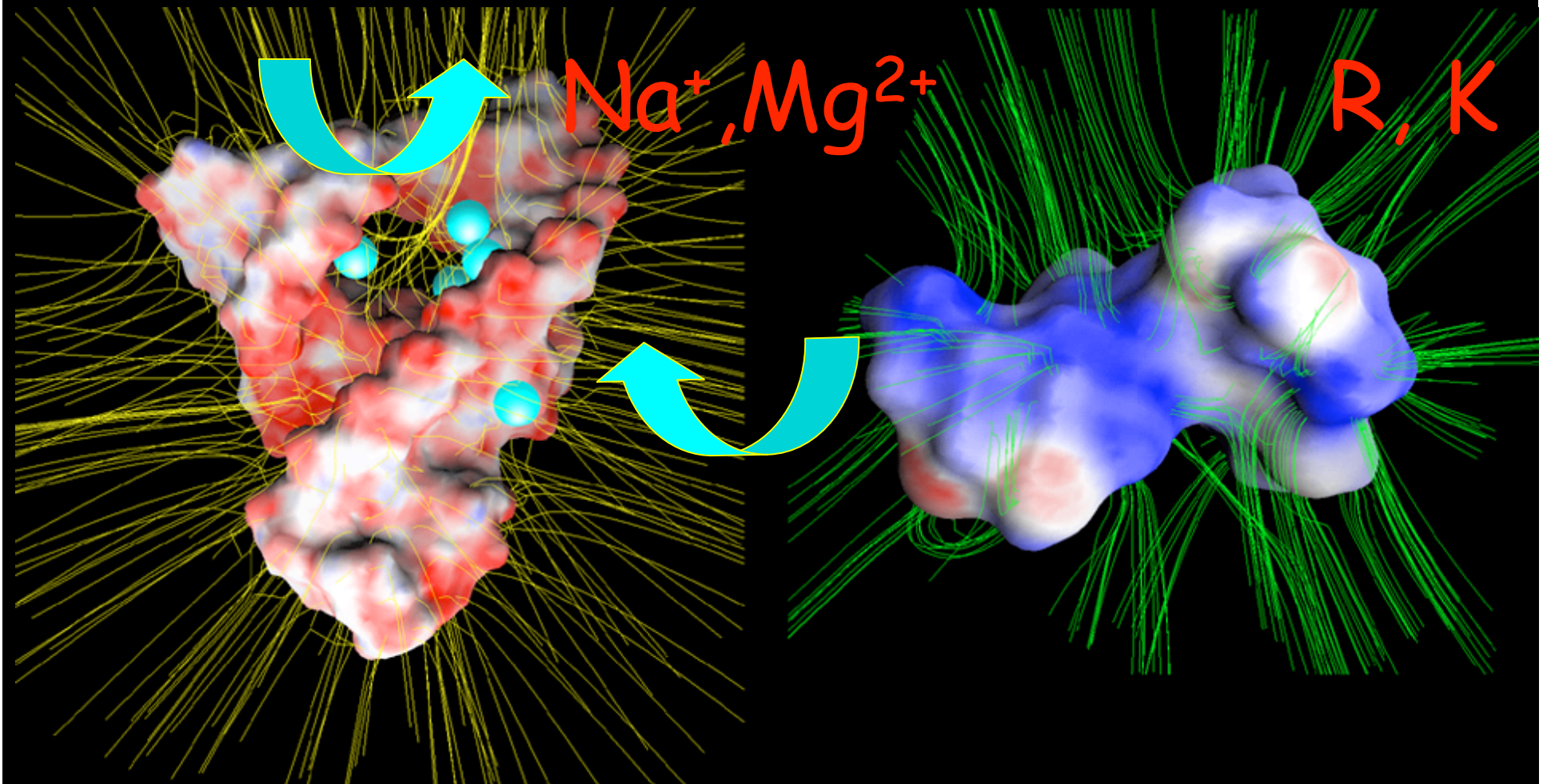


With Rangan and Woodson

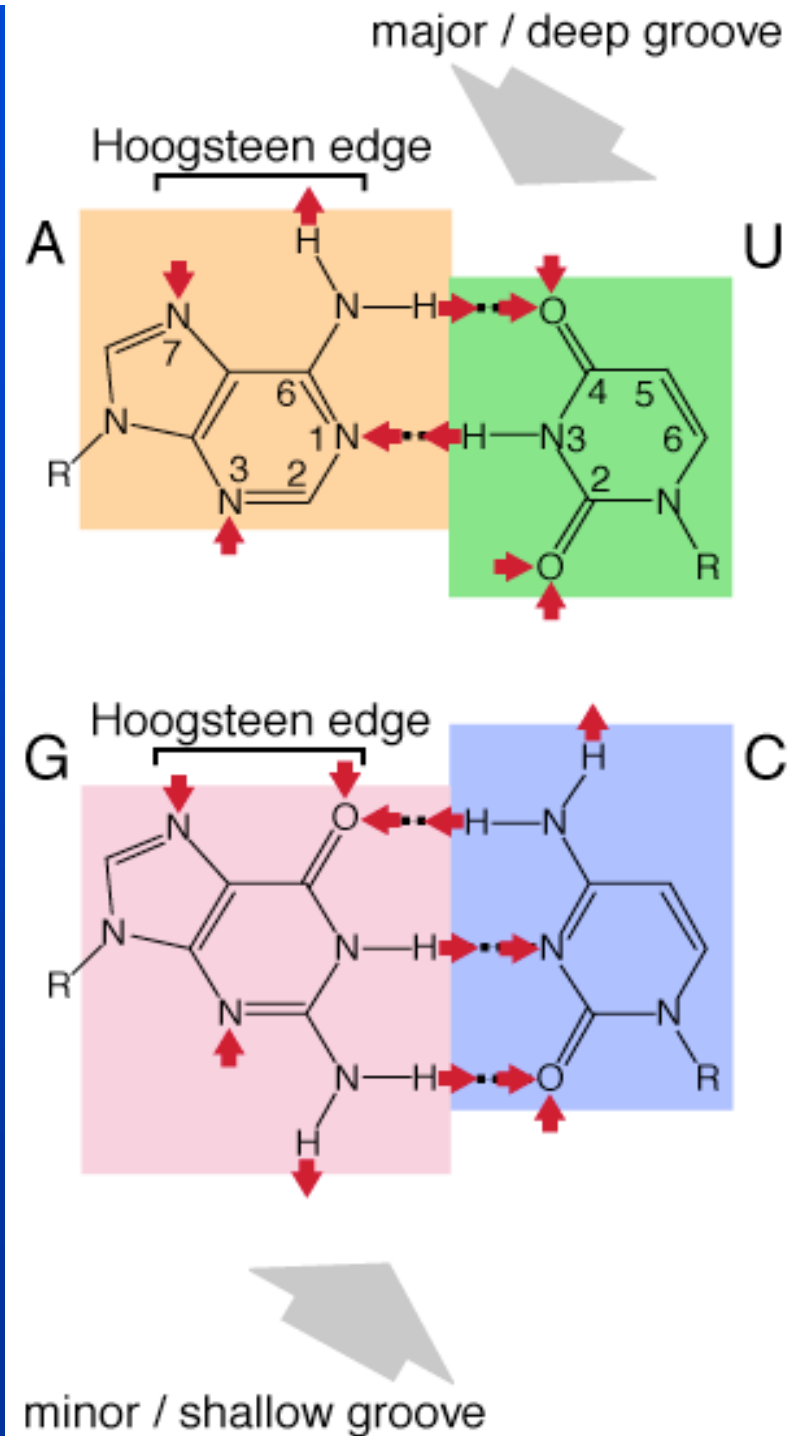
Assembly of the eukaryotic ribosome



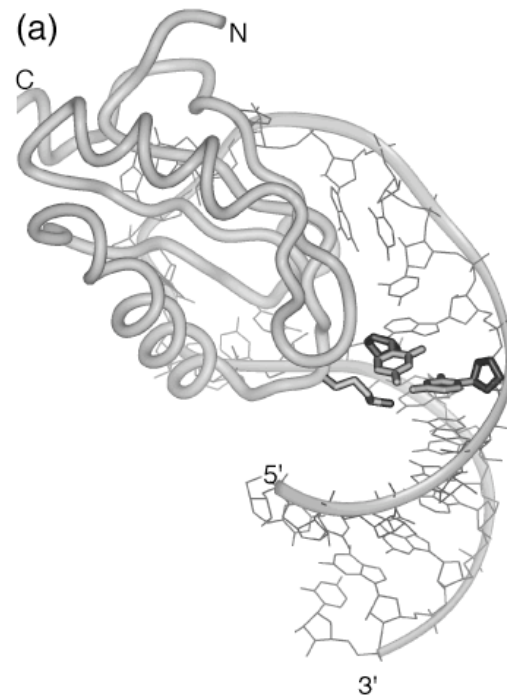
Driving force for protein binding to RNA



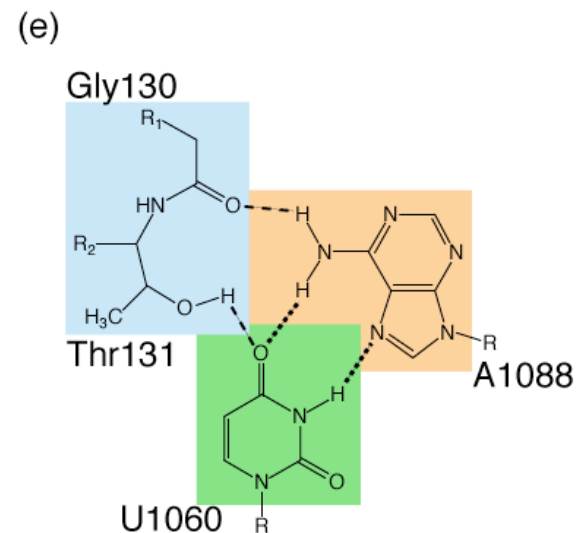
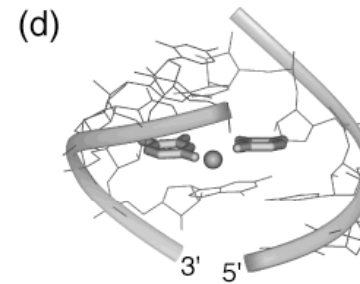
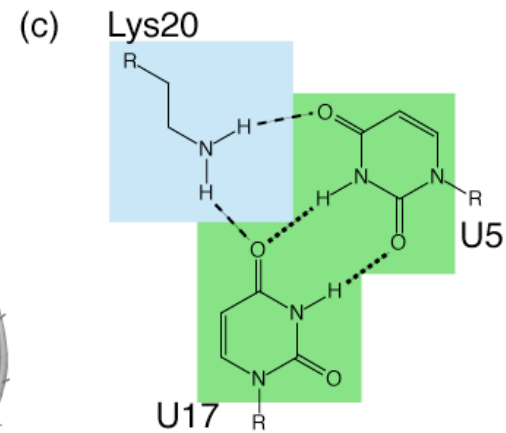
How are the bases recognized?



U2B''-U2A'
+ snRNA hp



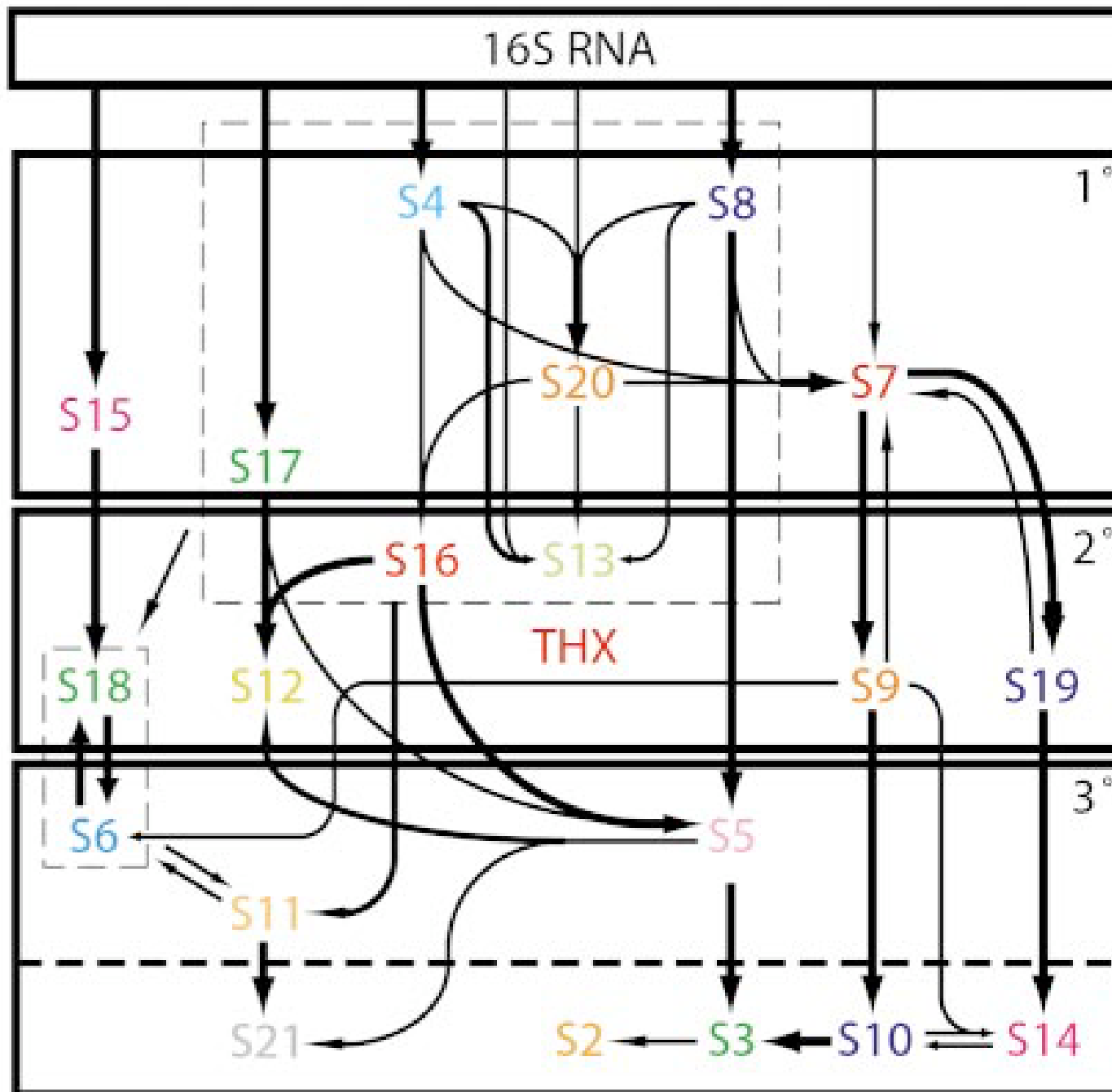
L11 +
23S rRNA frag



Statistics On 45 complexes

Most preferred AAs	R, N, S, K
Least preferred AAs	A, I, L, V
RNA-protein interface	22 % Hydrophobic 30 % Polar 40 % Charged 8 % Gly
Preferred pairs	R/K - phosphate S,F - adenine G - guanine N - uracil
RNA atoms	O2' 21% H-bonds
Protein atoms	Main chain 32% Side chains 68%

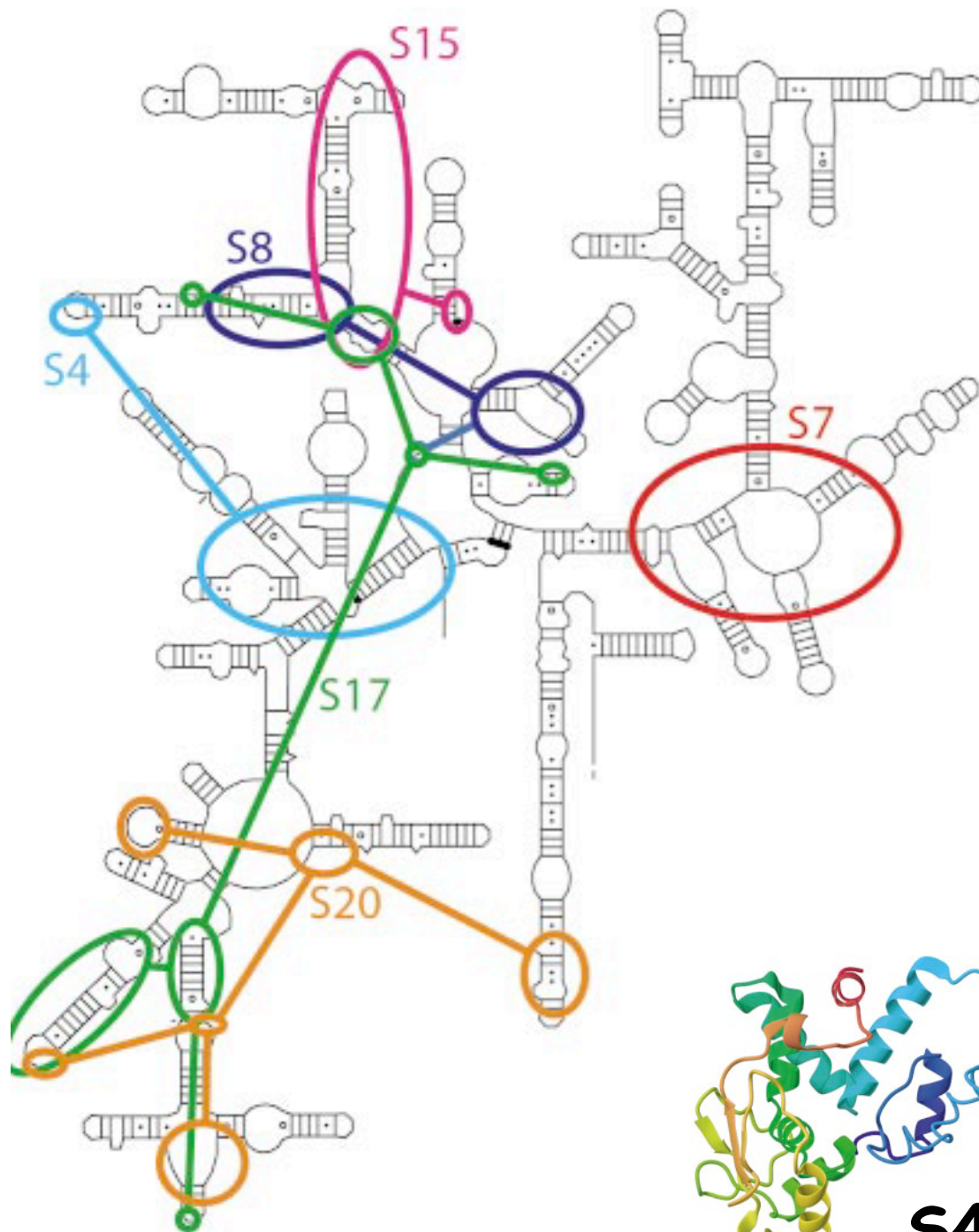
Is Coupling between
Electrostatic and
Architectural Hierarchies
present
in
RNA - Protein Complexes ?



Primary

Secondary

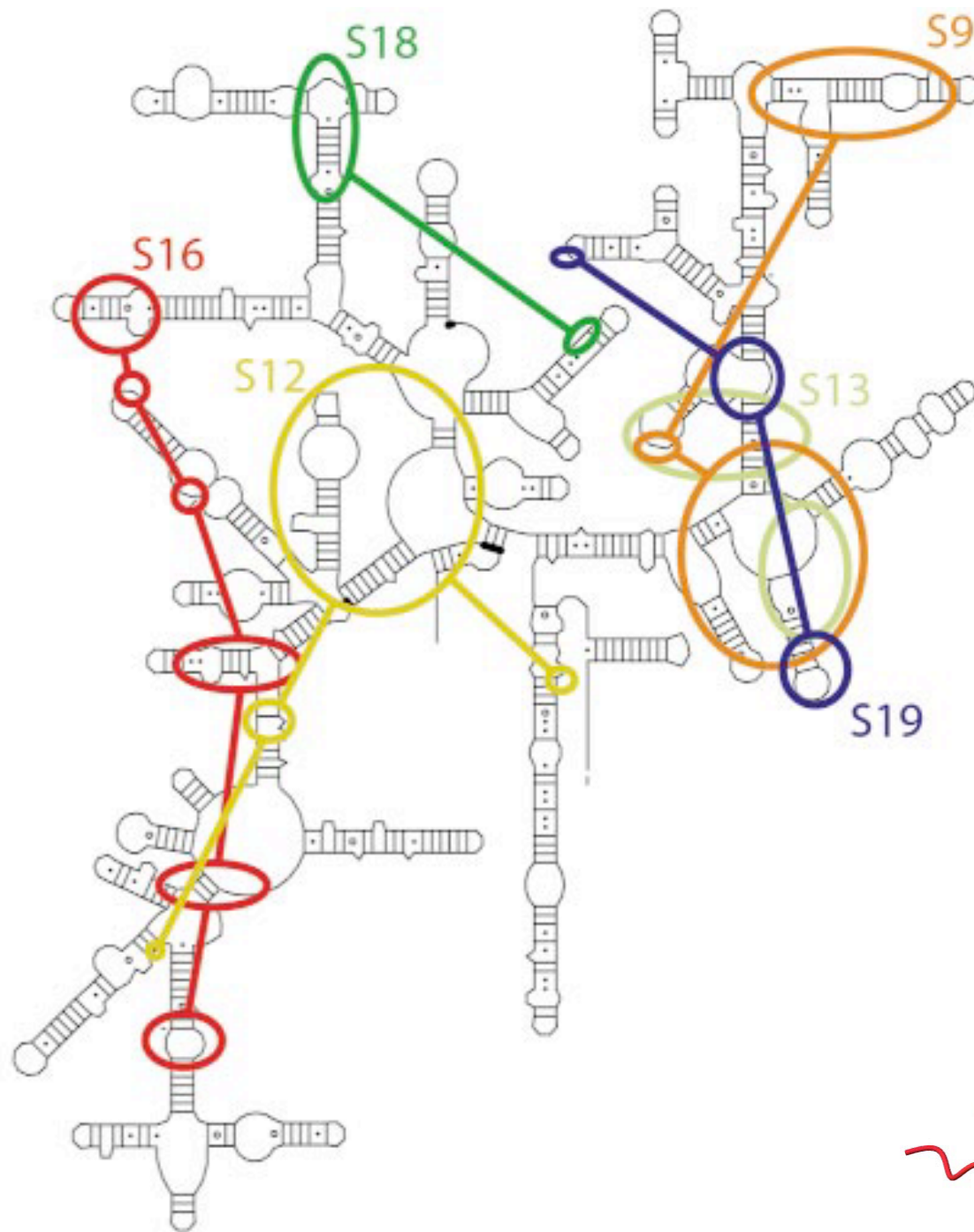
Tertiary



Primary binding proteins

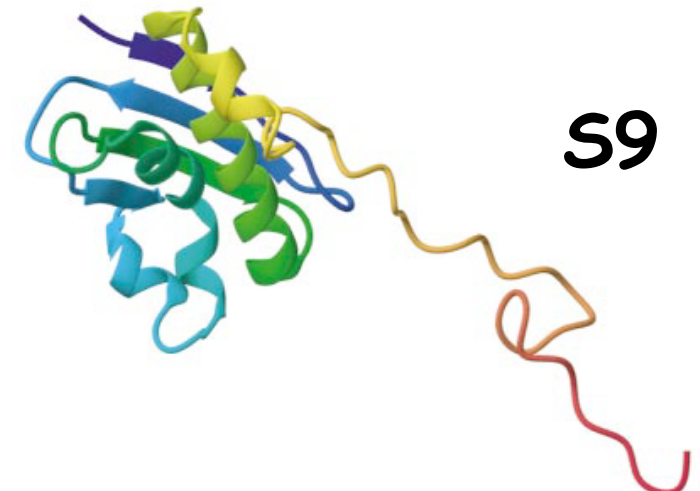
- Globular (no extension)
- Bind to multi-stem junctions
- Tie together different parts of the 30S subunit





Secondary binding proteins

- Arg & Lys-rich extensions



S9



S12

Assembly process: 5' → 3'

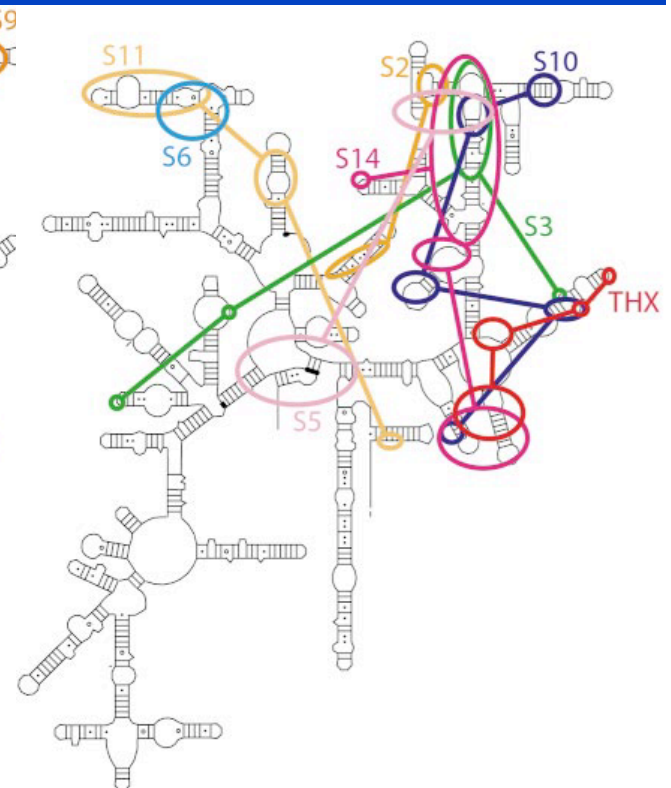
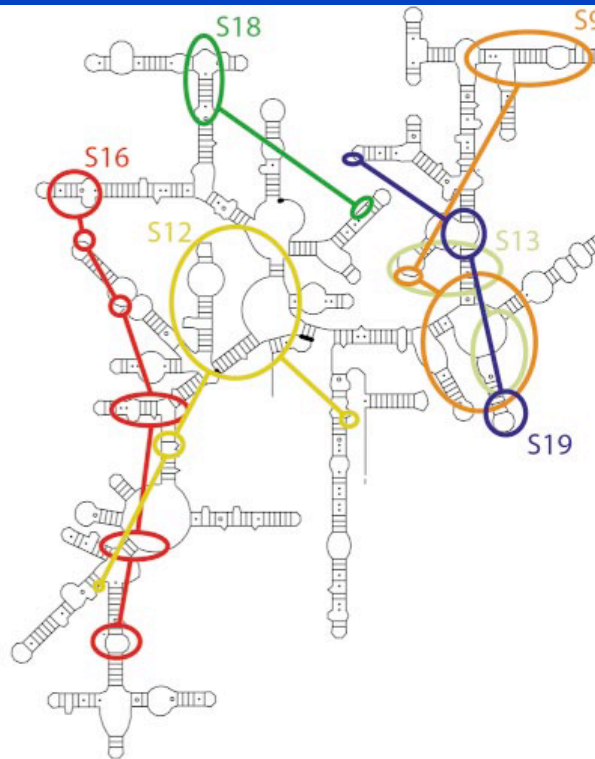
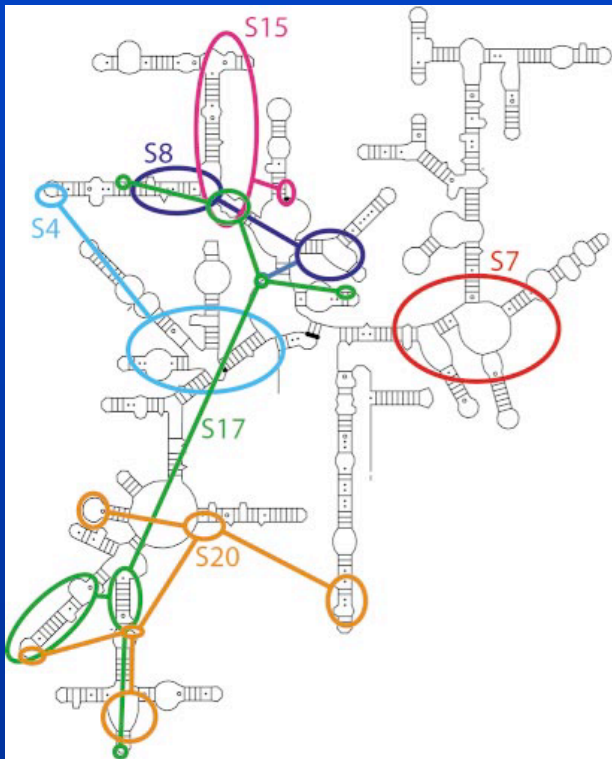


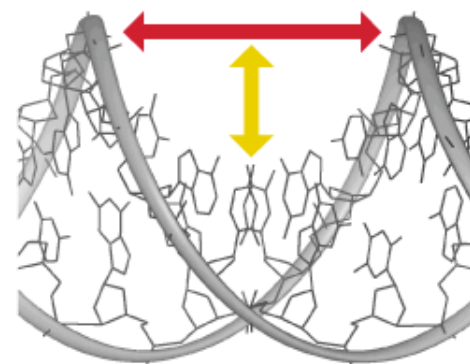
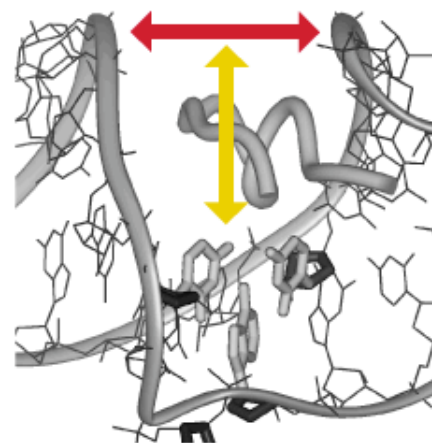
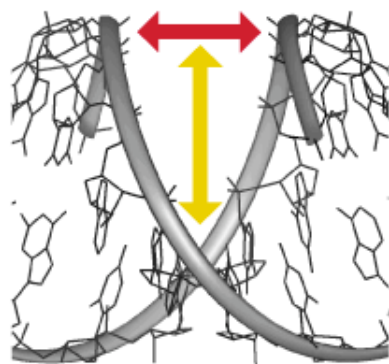
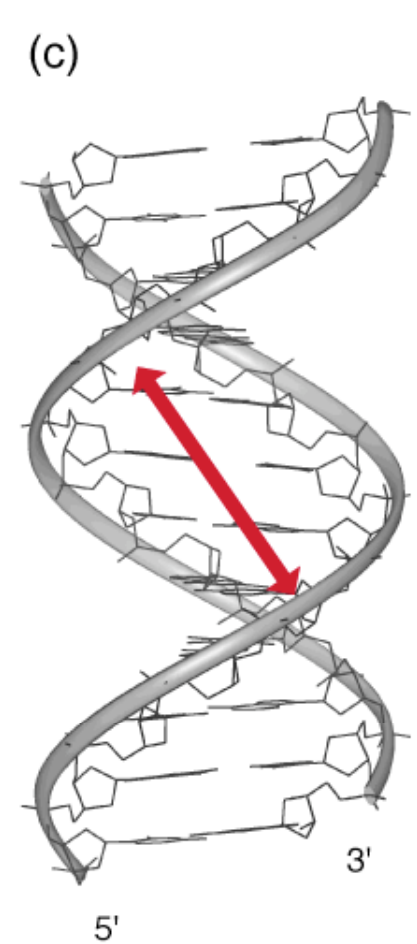
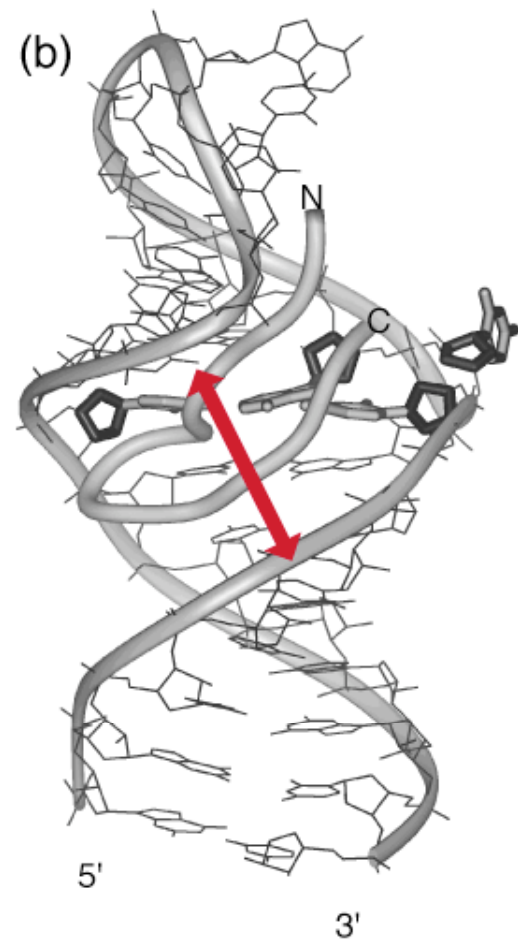
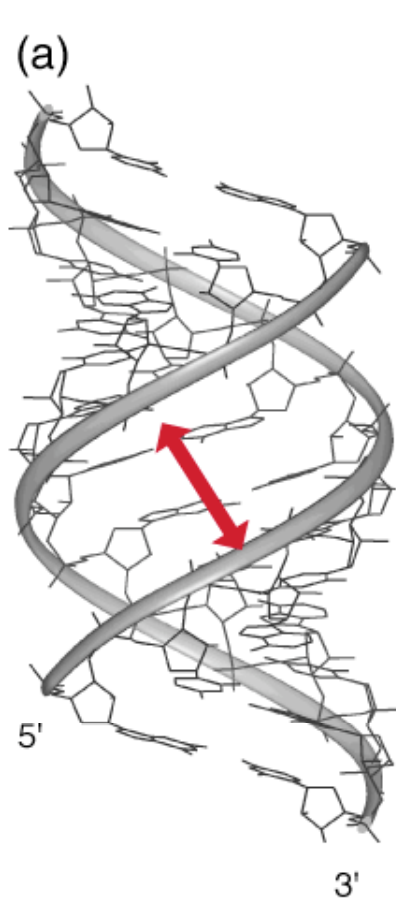
Protein classification as

Primary

Secondary

Tertiary





Proteins that bind both DNA and RNA

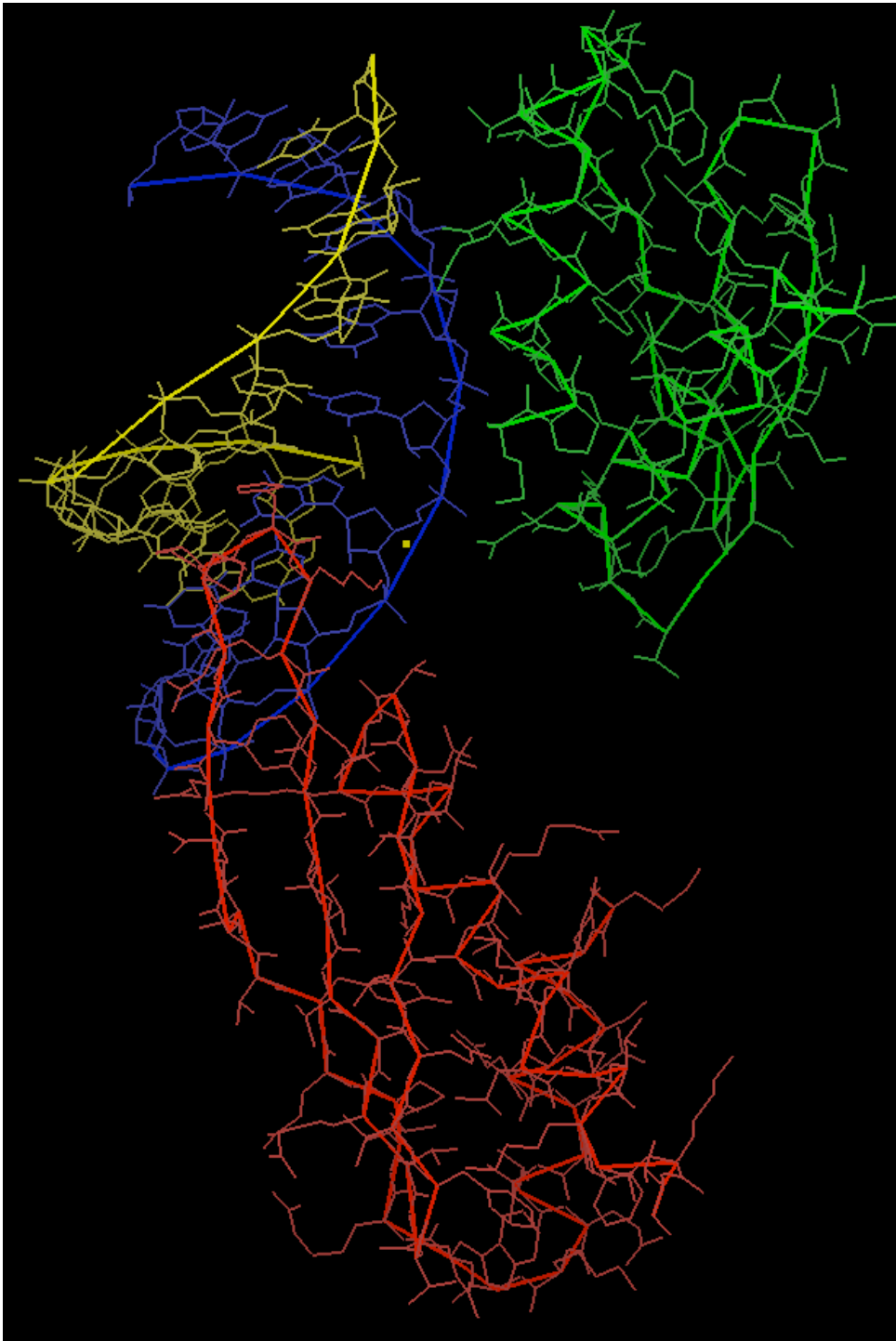
Several transcription factors :

- TFIIIA (5S rRNA/rDNA)

- p53

- Bicoid

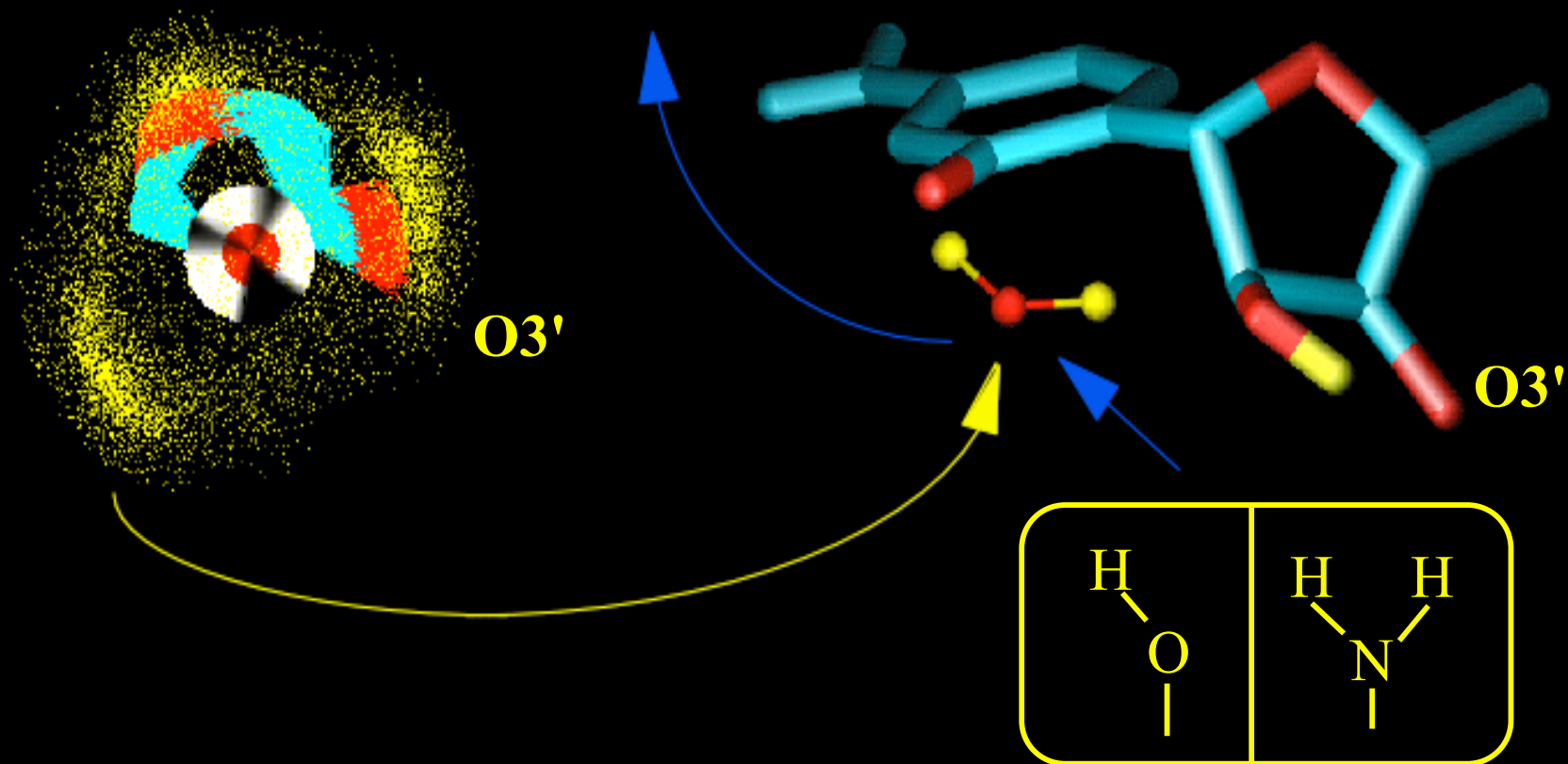
N.A.R. 30, 4118(2002)



X. Laevis Xlrbpa bound
to double-stranded
RNA dsRBM

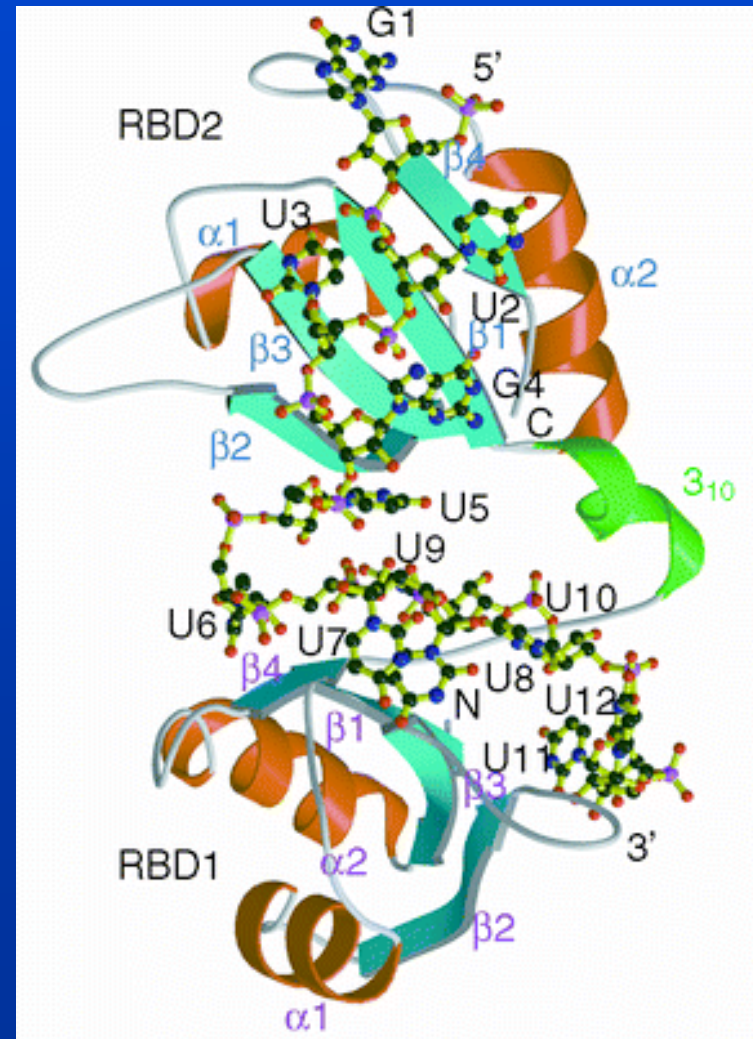
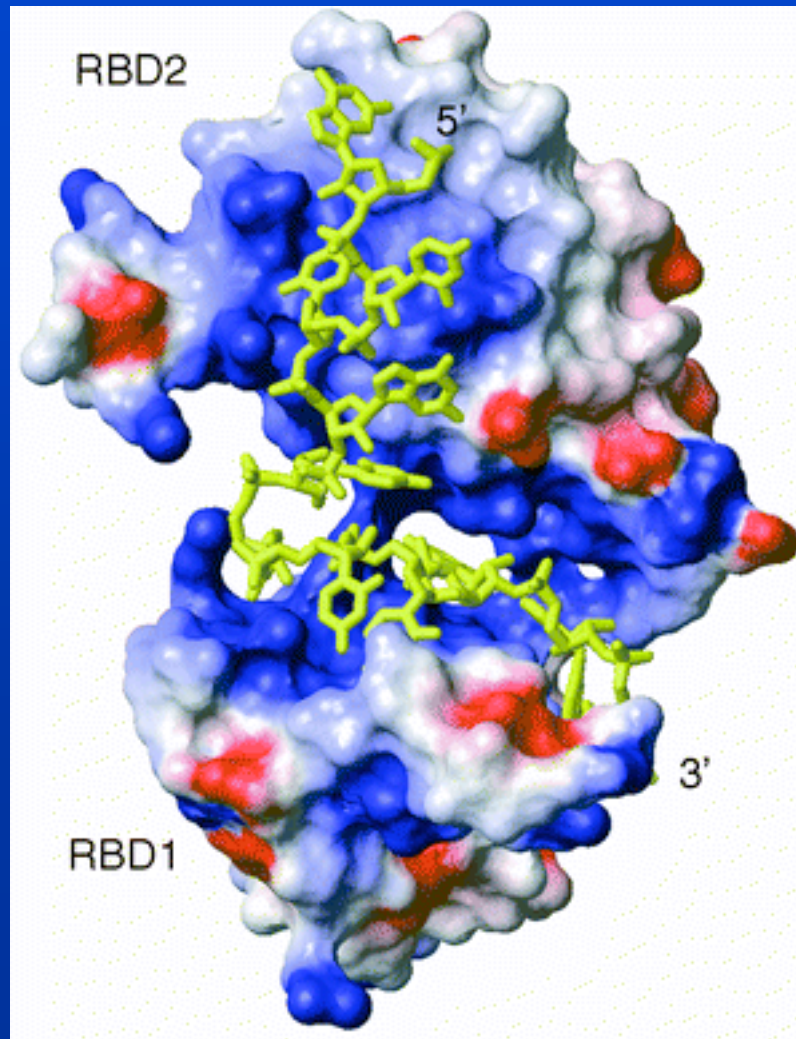
- limited number of
base specific contacts
- most of the contacts
with backbone or
O2'-H hydroxyl
-  fold
(Staufen, RNaseH,
HIV1 integrase, MuA
transposase)

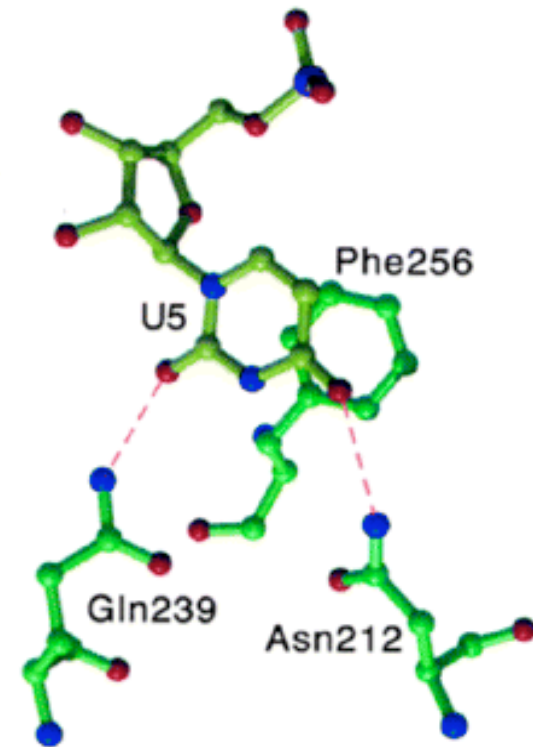
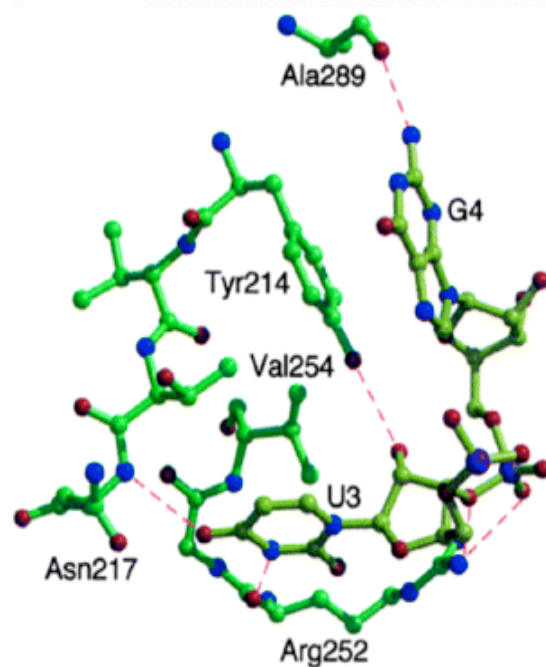
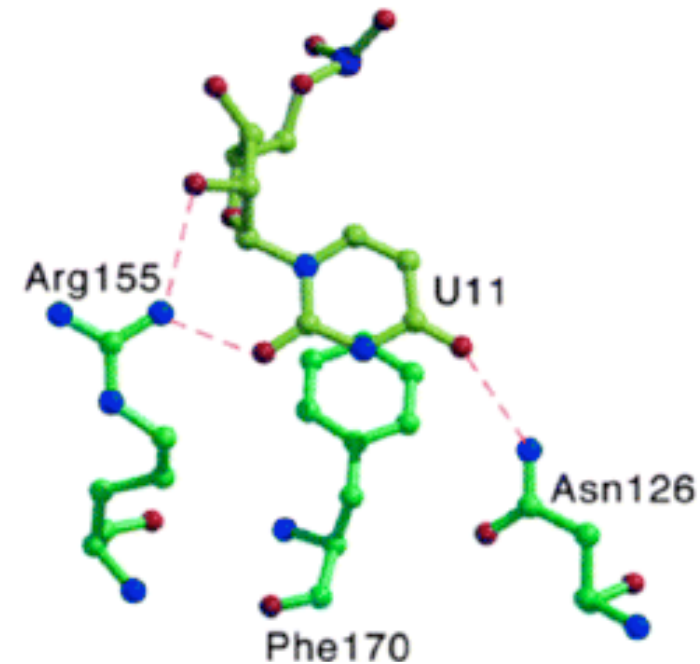
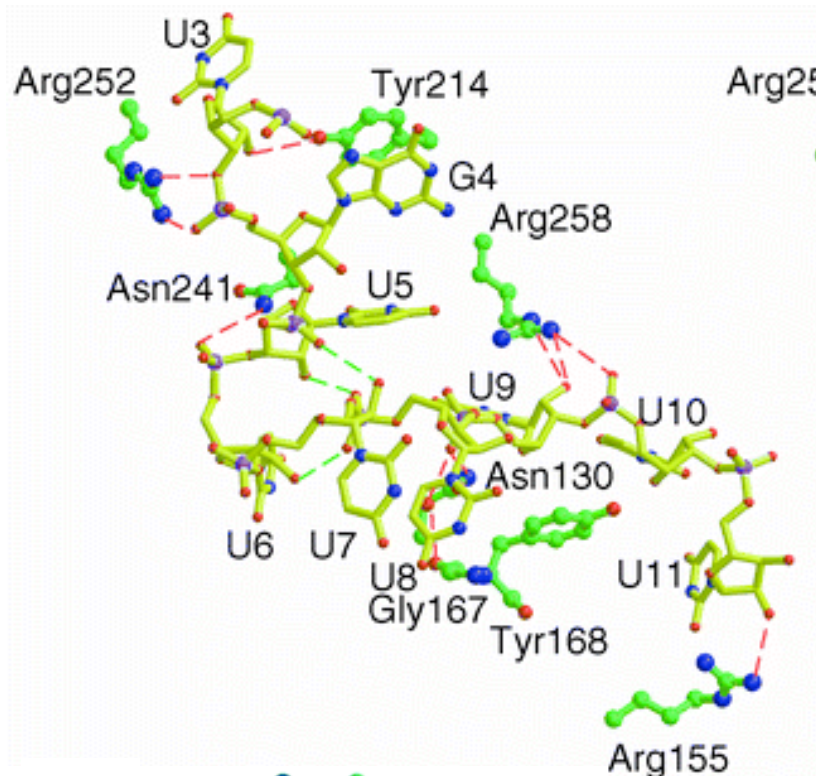
Hydration of the 2'-OH group (short residence times)



Can be easily exchanged with other groups like :

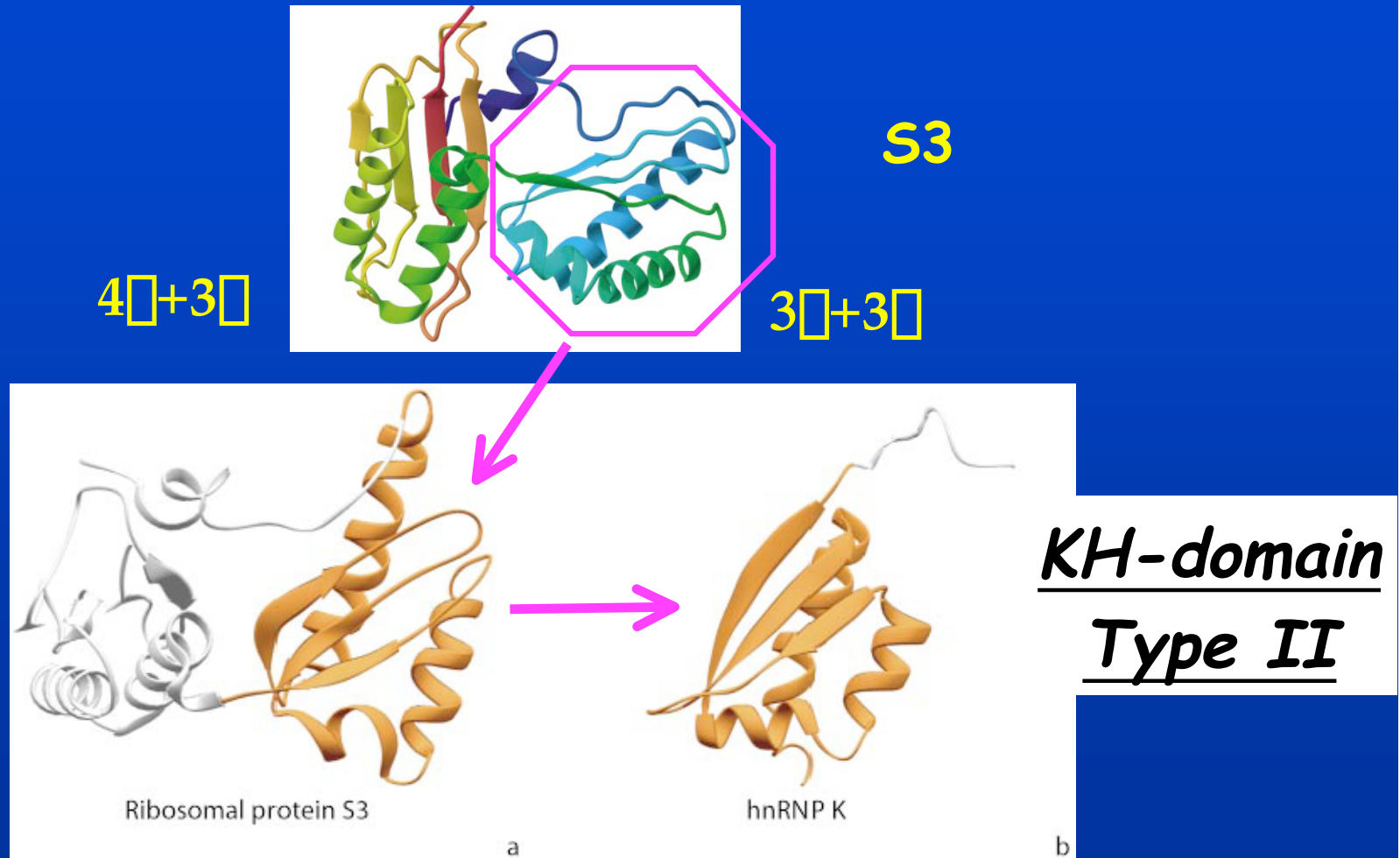
Sxl protein binding to the polypyrimidine tract of the tra messenger UGUUUUUUUU





Nature 398, 579 (1999)

RNA-Protein motifs



Same 3D structure but different topology

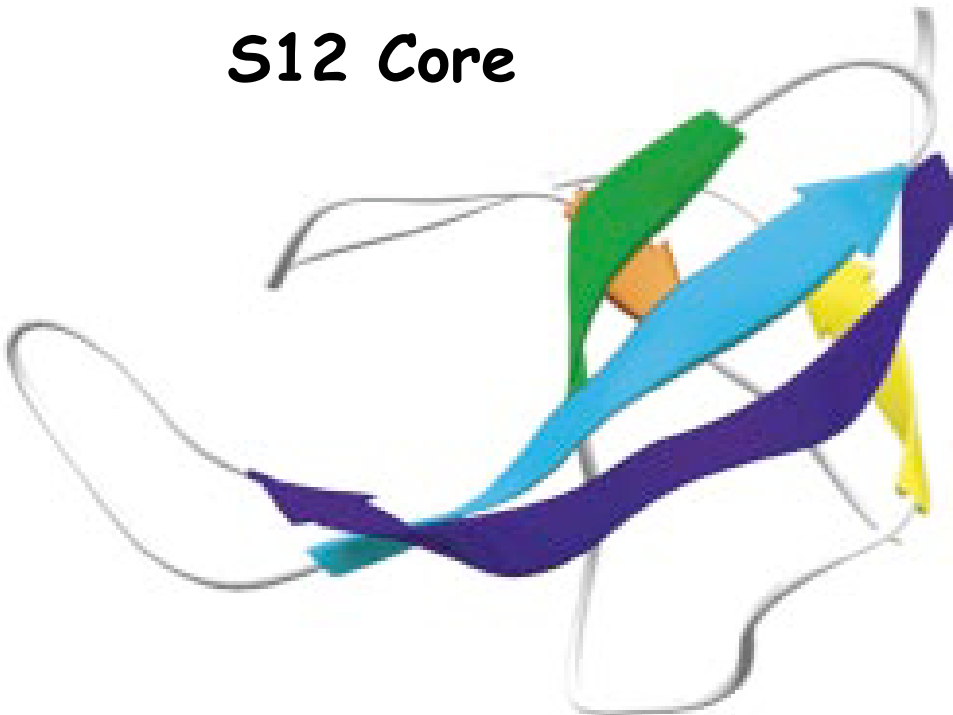
RNA-Protein motifs

OB Fold



Verotoxin

S12 Core



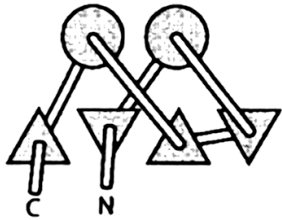
S17



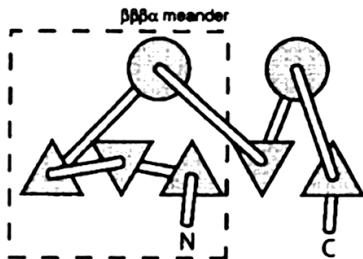
RNA-Protein motifs

2 □
4 □

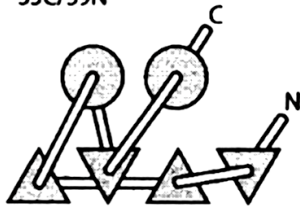
S6/S10/U1A (Classic RNP)



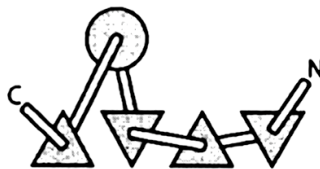
S11



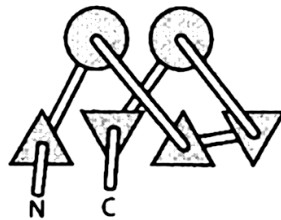
S5C/S9N



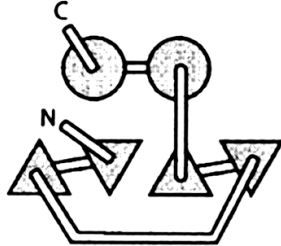
S5N



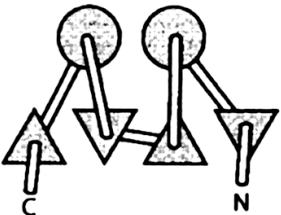
S8C ("Reverse" RNP)



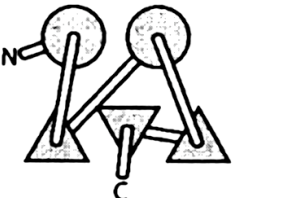
S3C



S3N



S8N



□+□ sandwich

