

Nucleotides

Mostly RNA

- complement RNA course in summer semester
- more DNA in sequence context
- RNA does more biochemistry
 - RNAszymes, regulators

Assumed

- you remember
 - ACGT in DNA
 - ACGU in RNA
- always write from 5' to 3'

Roles of molecules

	RNA	DNA	proteins
genetic information	X	X	
structure	usually single stranded	duplex	lots
regulation/interactions	X	X	X
ligand binding / catalysis	X		X

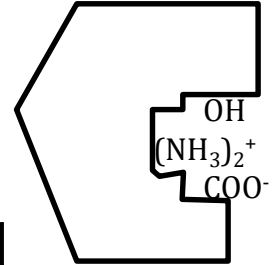
If RNA does all this interesting chemistry

- it has interesting structure

How do proteins work ?

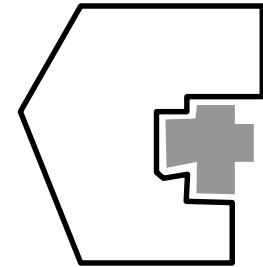
Some site decorated with special groups

+ / -, neutral, polar / non-polar, big / small



Chemical choice ?

- 20 kinds of amino acid
- half a dozen really different types

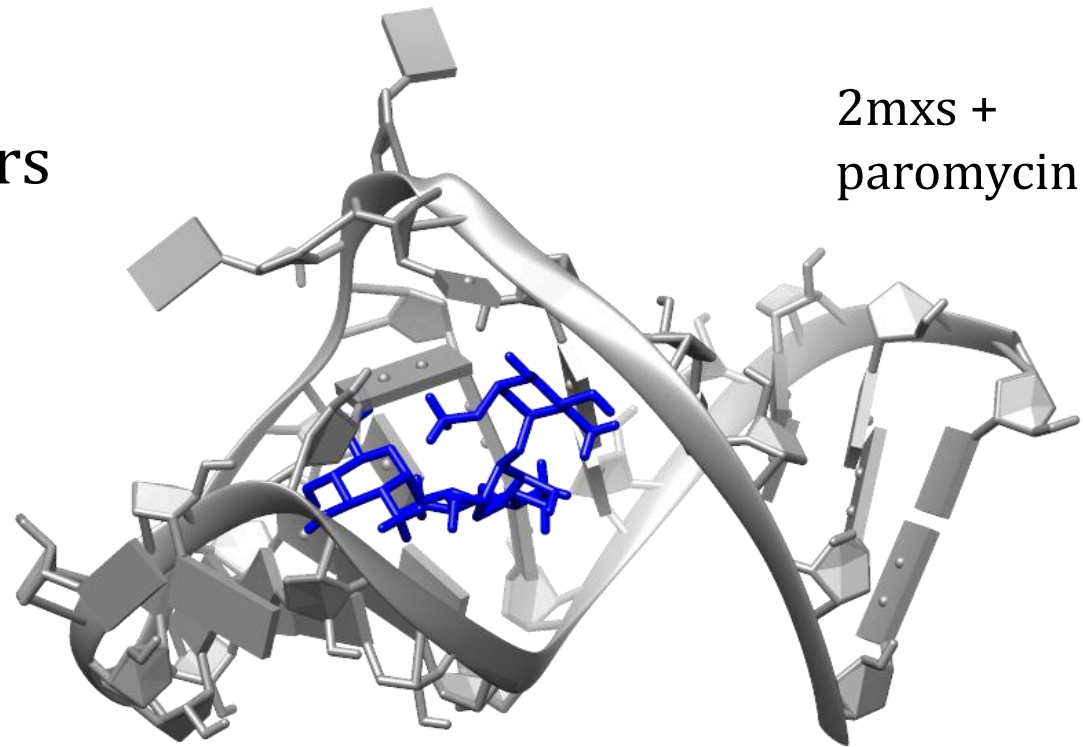


Do you see this with nucleotides ?

RNA binding ligands ?

Examples

- riboswitches / regulators
- catalysts



Two consequences

1. RNA must fold to certain shape
2. Exposed chemical groups give specificity / strength

Do not see this much with DNA

Structures / type of molecule

Protein

- specific structure depends on sequence
- sometimes floppy – not structured

DNA

- double helix

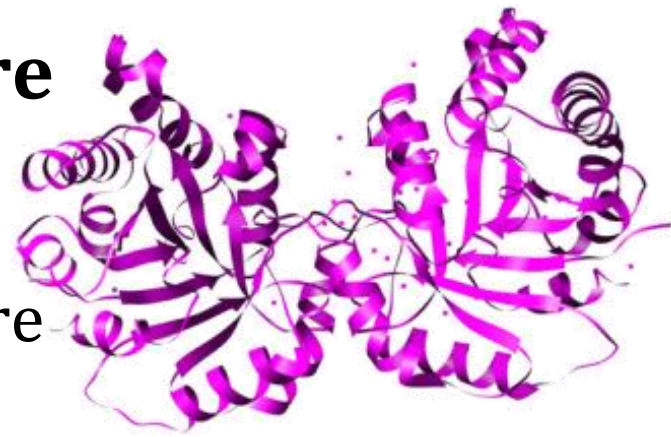
RNA

- do they fold to nice, well-defined shape ?
 - all RNA ?
 - all biologically-interesting RNA ?
 - some ?

Views of structure

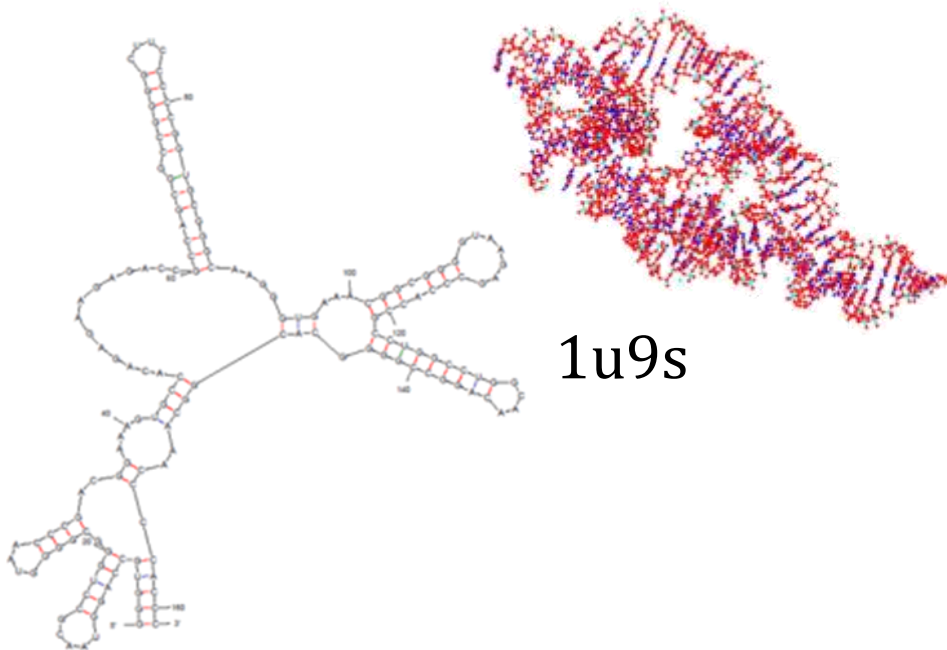
Proteins

- usually 3D – rarely secondary structure



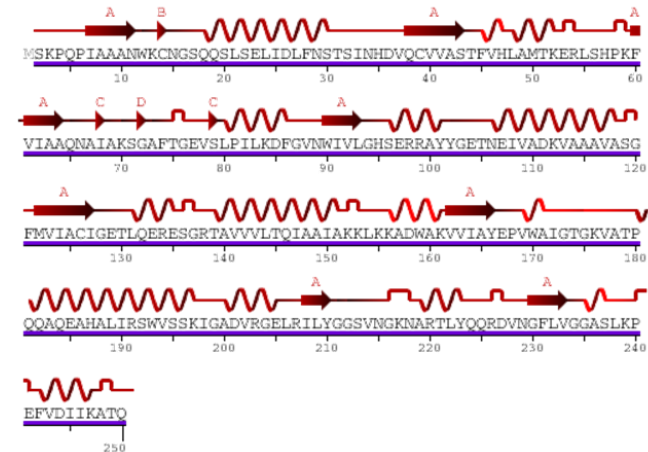
RNA

- usually 2D – less 3D information



1u9s

3tim



RNA – how much information ?

Proteins

- 1.3×10^5 or about 3×10^4 interesting ones

RNA

- 3.5×10^3 structures with some RNA
- 1290 with pure RNA - many small and boring
- 409 pure RNA ≥ 40 bases (really less - lots of redundancy)

Why so few RNA structures ?

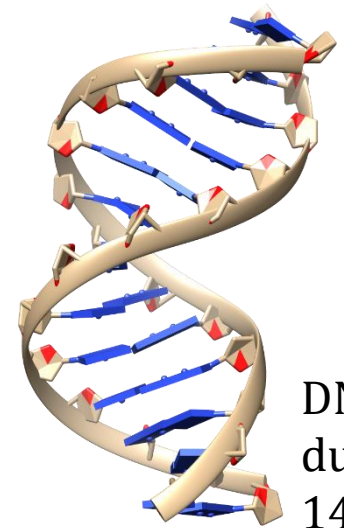
- RNA hard to handle (RNases)
- crystallography
- NMR
 - assignments very difficult (only 4 kinds of base)

Features of RNA

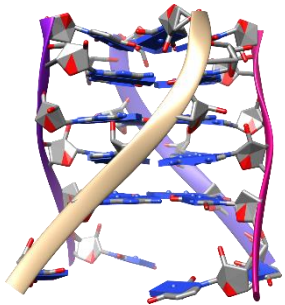
What dominates literature ?

- base pairing

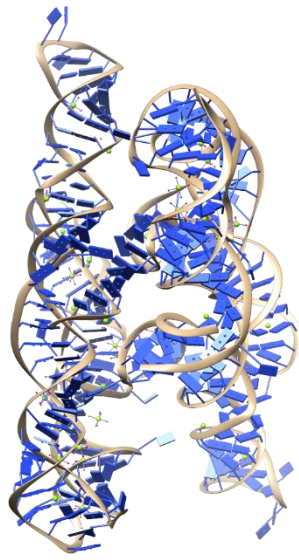
Need more interactions to explain all these shapes



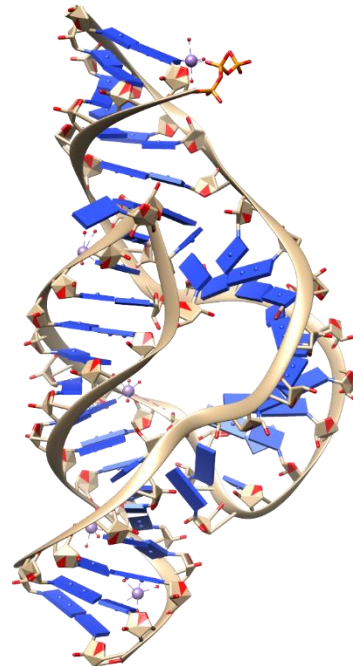
DNA
duplex
140D



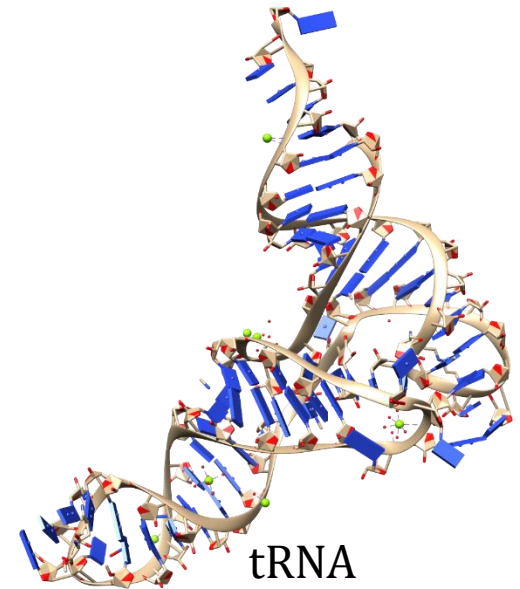
tetraplex
1mdg



group I intron
1hr2



hammerhead
2oeu



tRNA
1evv

Important for RNA structures

Energies ?

- As in previous lectures
- bonds, bond angles, torsion angles
- non-bonded (electrostatics, van der Waals)

Details coming ..

- H-bonds
- charges
- stacking

Is my description consistent ?

- H-bonds/charges/stacking vs electrostatics/van der Waals

non-bonded terms / convenience

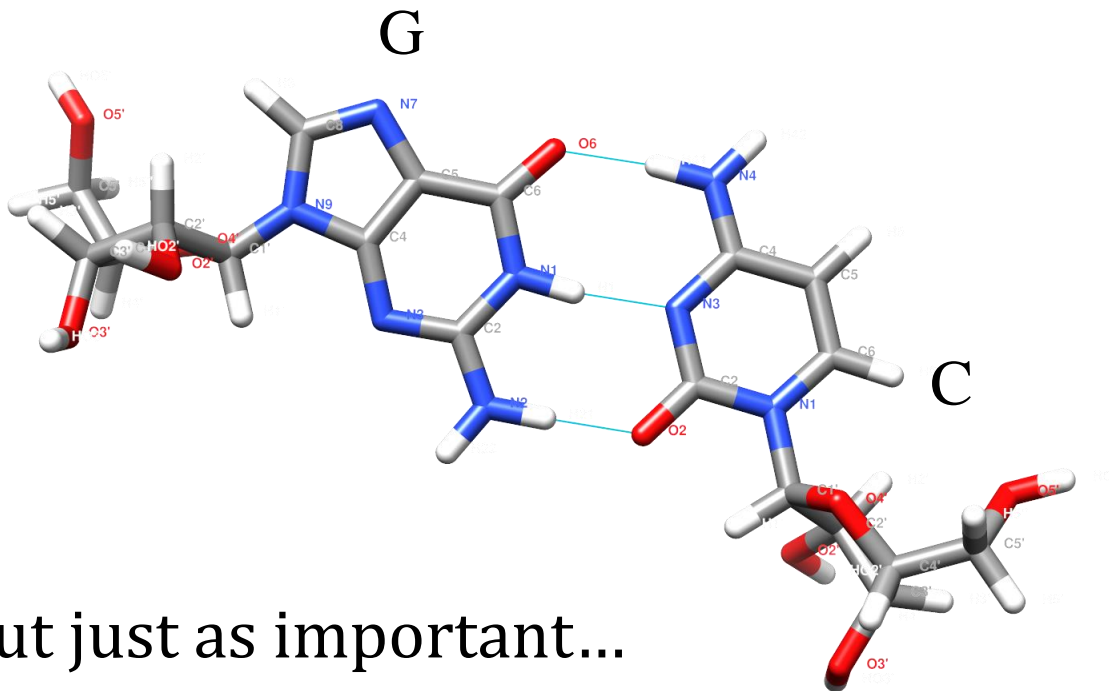
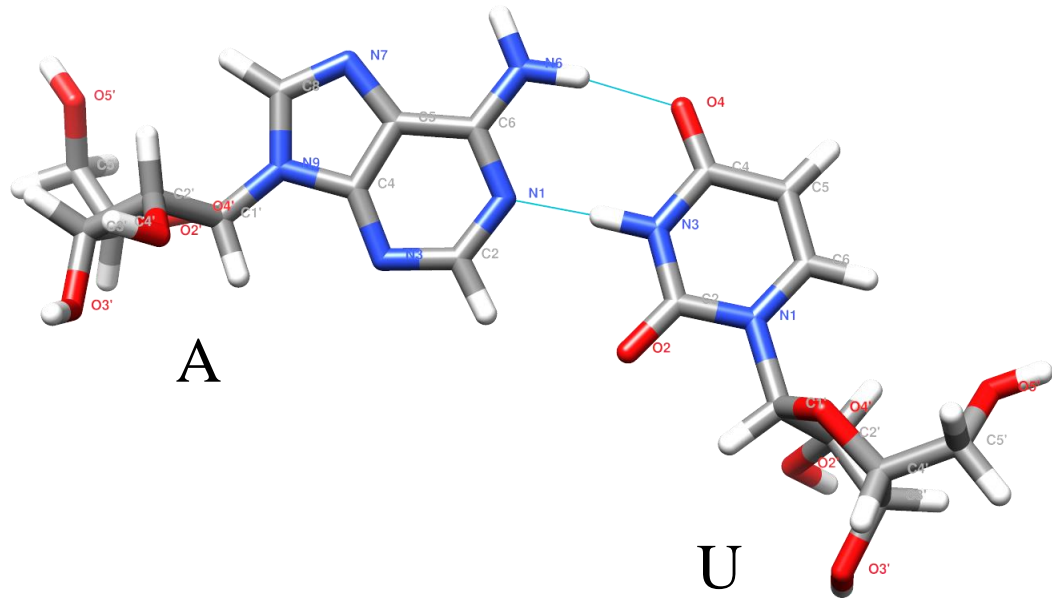
Physics not changed

- convenient to talk in terms of H-bonds, charges and stacking

H-bonds	van der Waals electrostatics	base-pairing + bit more
charges	electrostatics	backbone
stacking	van der Waals	bases

Base-pairing / H-bonds

Historic



but just as important...

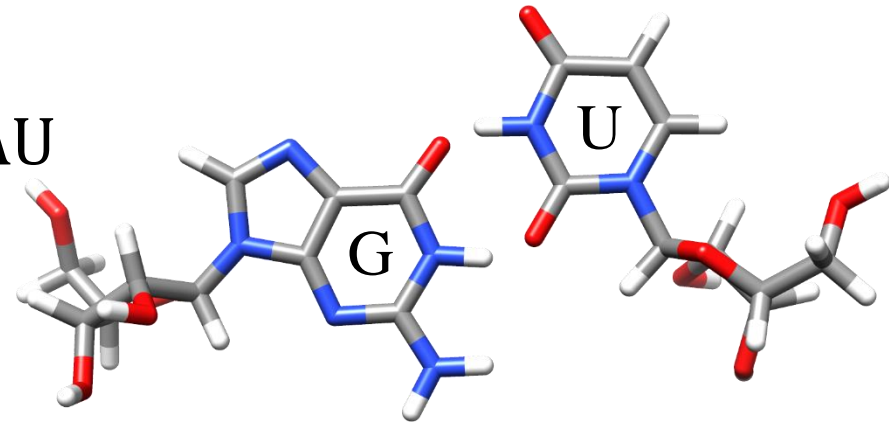
H-bonds wobble pairs

GU

- strength very comparable to AU

Compare with DNA

- mismatches – very rare



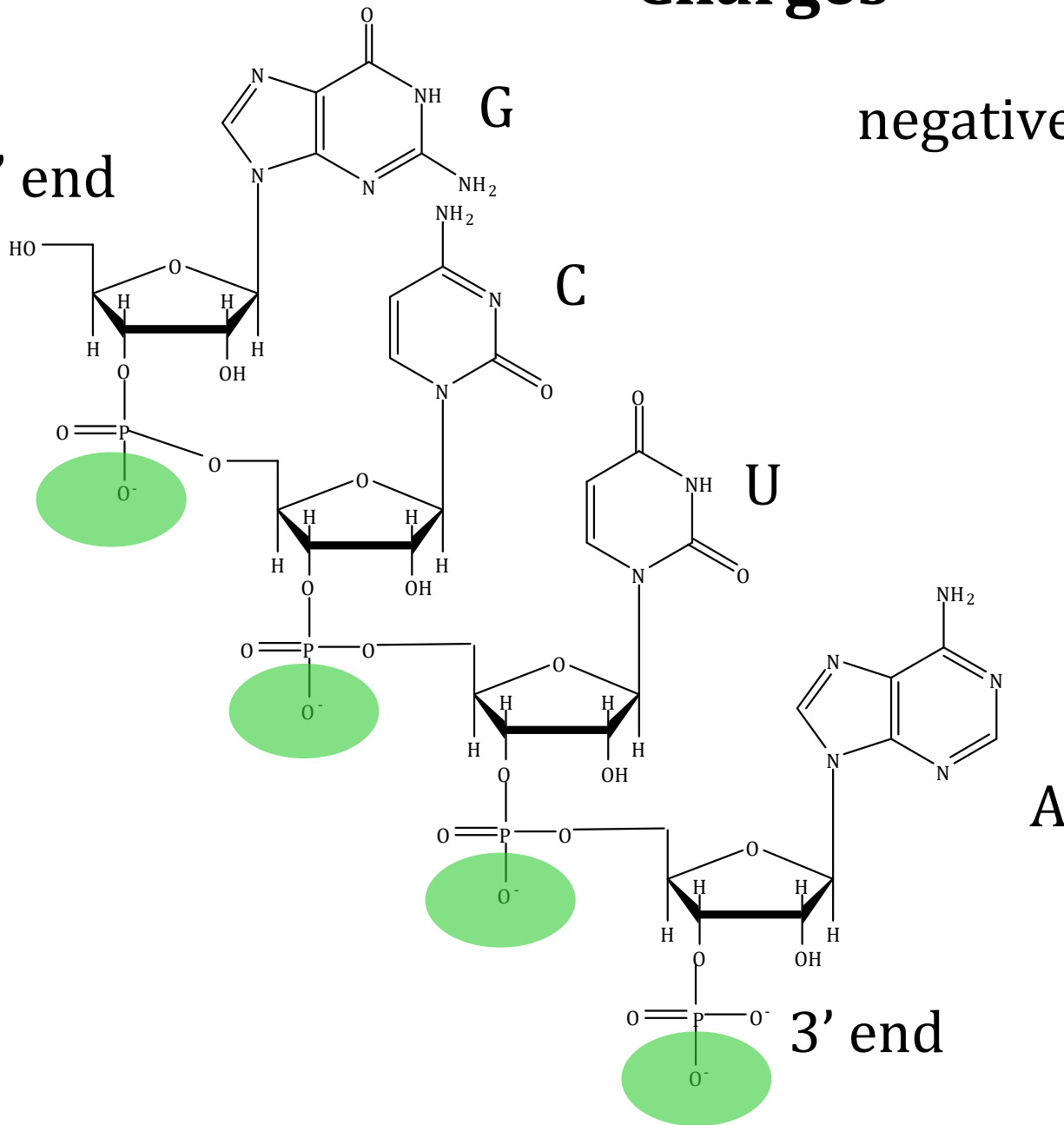
More generally

- count the H-bond donors and acceptors
- many H-bond possibilities
 - not limited to bases

Charges

5' end

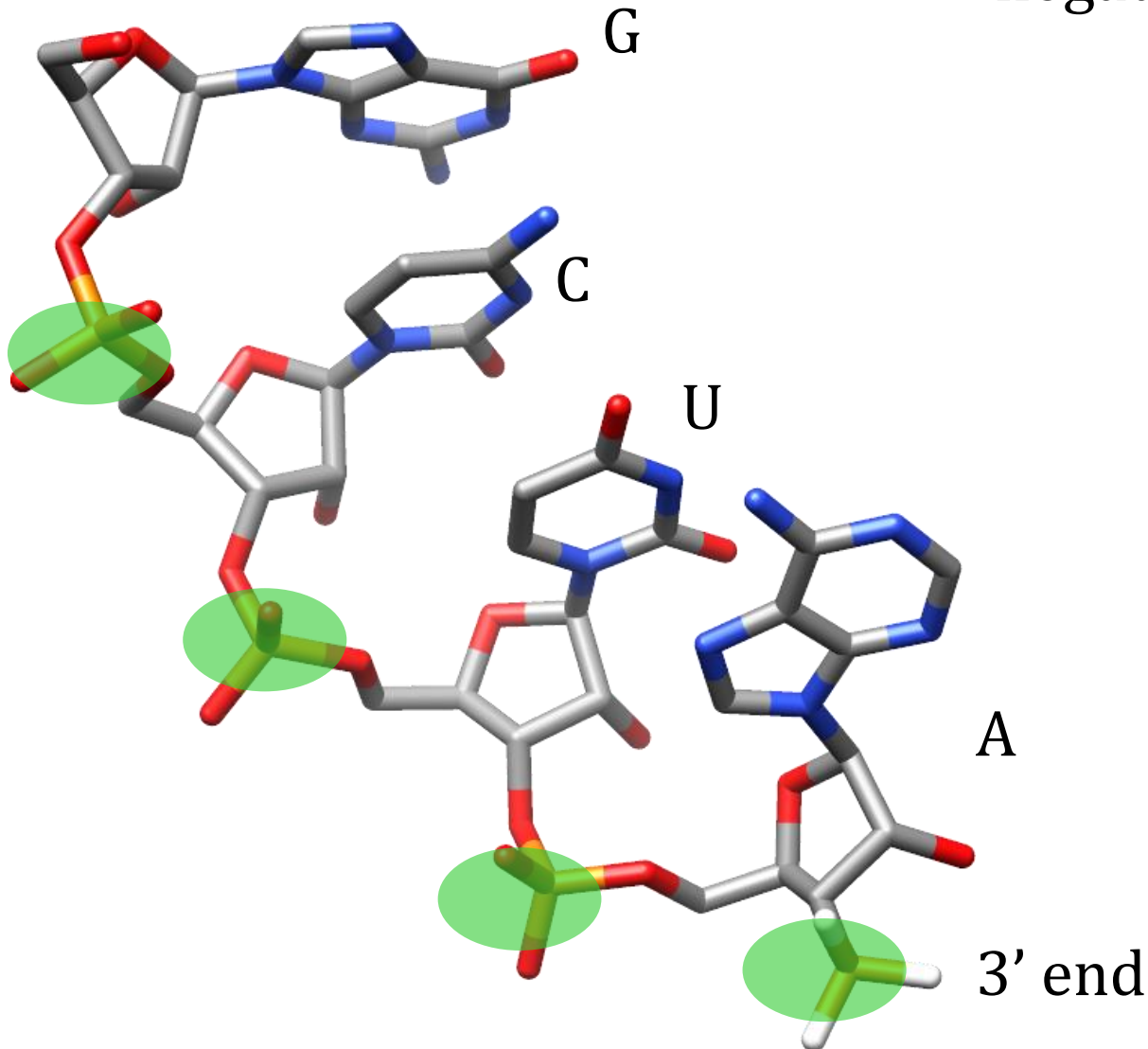
negative charges



Charges

5' end

negative charges



Charges

Contrast with proteins

- mostly neutral, some charged residues

RNA and DNA

- full negative charge every base (at backbone)

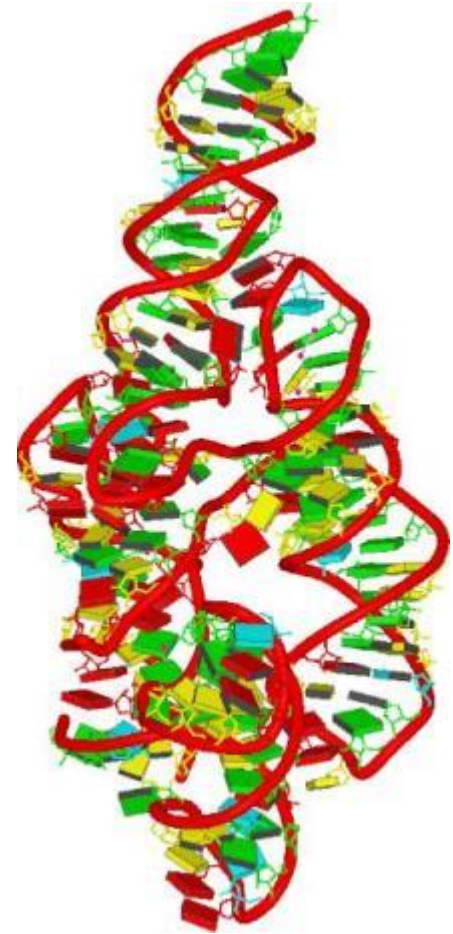
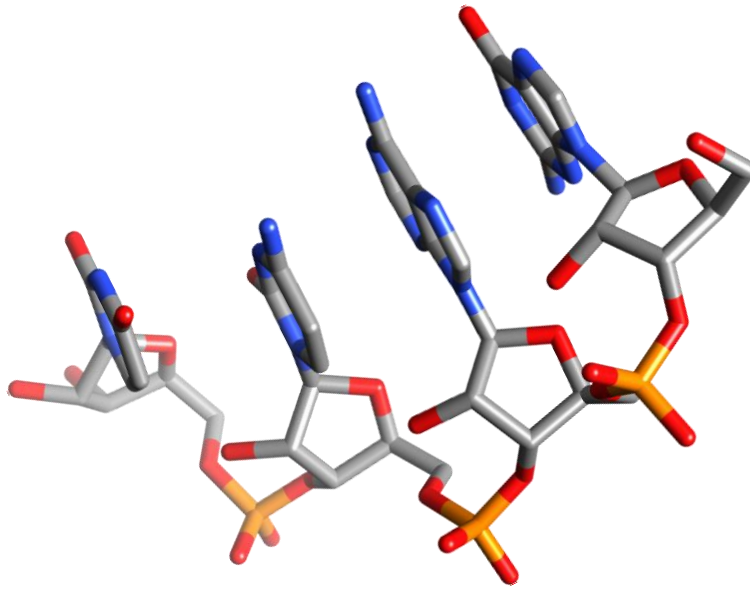
Consequences

- strong interaction with
 - solvent
 - +ve ions
- shape of backbone
 - move PO_4^- away from each other

Stacking

Bases are large aromatic systems

Very strong preference to form stacks



Representation / storing 3D structures

Proteins – conventions and simplifications

- diagrams – ribbon plots
- break into secondary structure and loops
- represent as a set of C^α atoms
- Ramachandran / ϕ, ψ plots

RNA - similar ideas ?

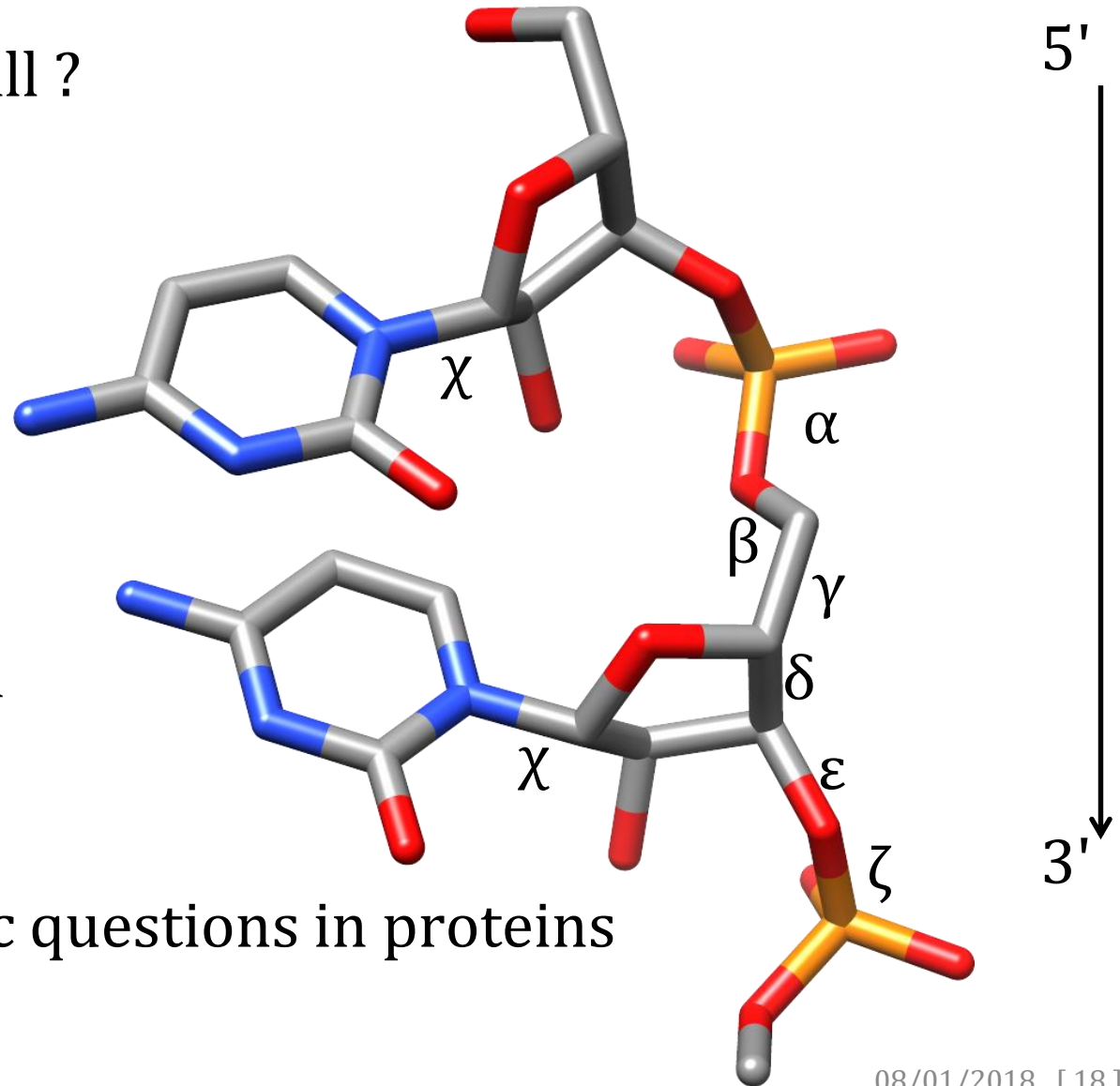
RNA – no Ramachandran plot

Many angles

- do we need them all ?

Two issues

- restricted freedom
consider δ
- correlations
 - partly like steric questions in proteins



Use less than 6 angles

We do not need 6 independent descriptors (angles)

- want to simplify
 - for communication
 - calculations / storage

Easy – but no agreed scheme

- a proposal

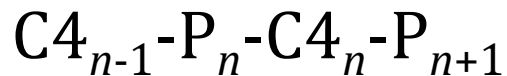
Torsion angles

Use atoms that are not bonded to each other

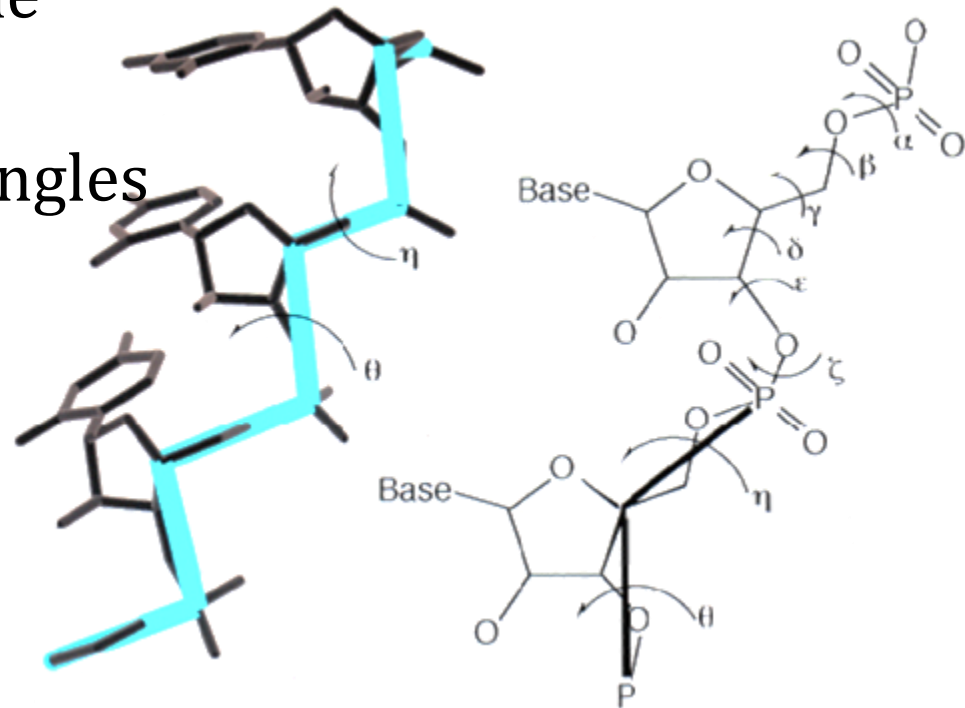
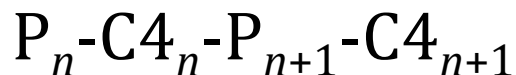
Basic idea

- pick 4 atoms that are not sequential
- define a simplified backbone
 - $P-C_4-P-C_4-P-C_4-\dots$
- leads to "pseudo-torsion" angles

η



θ



End of structure introductions

- Nucleotide history dominated by base-pairing
- single-stranded RNA folds into shapes like an enzyme / receptor
- Energies - we use simplifications
- Must be more than just base-pairing
- Representations - not as nice as for proteins

Remember everything for next topic

- predicting secondary structure