

RNA Chemistry & Structure

- Roles of molecules

	RNA	DNA	proteins
genetic information	yes	yes	
catalysis	yes		yes
regulation / interactions	yes	yes	yes
structure	usually single stranded	usually duplex	lots

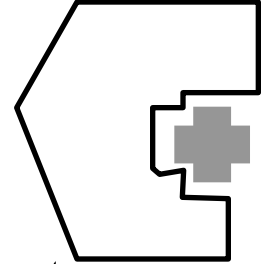
Catalysis

- proteins
 - everywhere
 - classic enzymes
- RNA
 - less common / well established
 - ribosome, hammerhead, ...

Regulation

- Proteins
 - bind substrates, ligands, DNA, RNA
- DNA
 - sequence specific binding – to proteins, RNA, DNA
- RNA
 - same as DNA +
 - specific catalysis implies specific recognition

Recognition / binding specificity



- protein view – via evolution
 - protein scaffold / framework positions groups
 - in binding / reactive region specific groups interact
 - lots of chemical groups to choose from (20 amino acids)
- DNA – not thought of in these terms
 - some specificity
 - regulatory binding proteins are sequence specific
- RNA
 - sequence specificity for binding proteins
 - RNAzymes suggest some protein-like abilities
 - experimental
 - selection of specifically binding RNA (selex later)

Structure

- DNA
 - mostly thought of as double helix
- Protein (simple dogma)
 - from a specific sequence to a well defined structure
 - less often – floppy, unstructured, mobile, alternative folds
- RNA
 - does an RNA sequence fold up to a well defined structure ?
 - all possible RNA's ?
 - biological RNA's ?
 - some RNA's ?

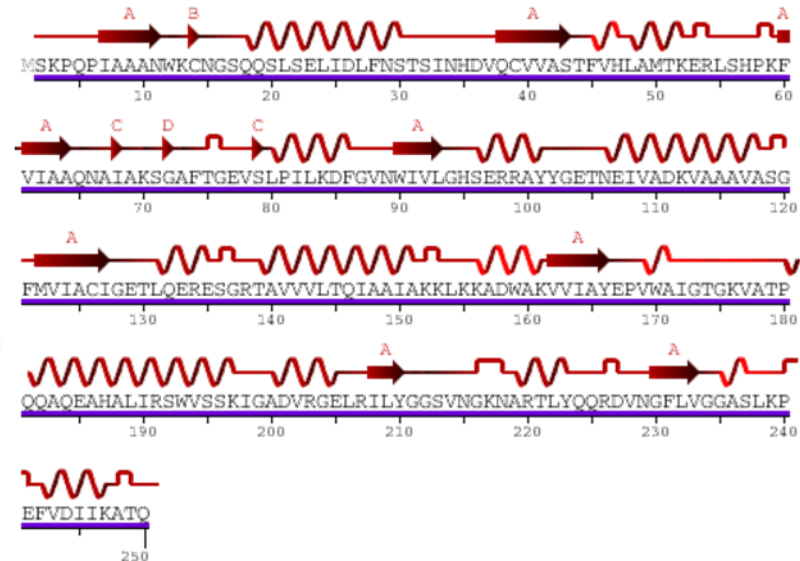
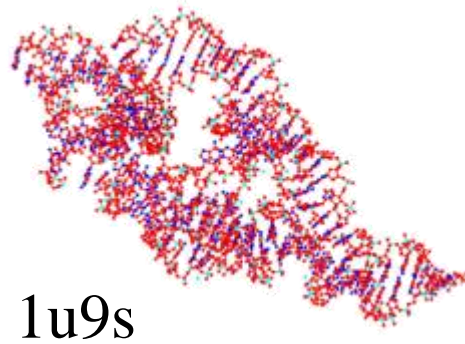
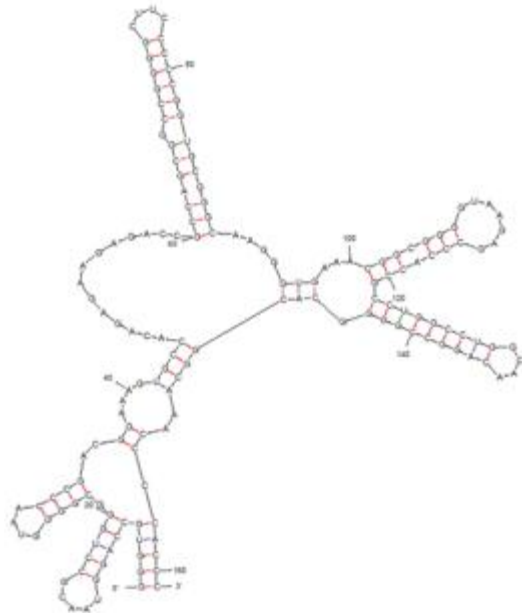
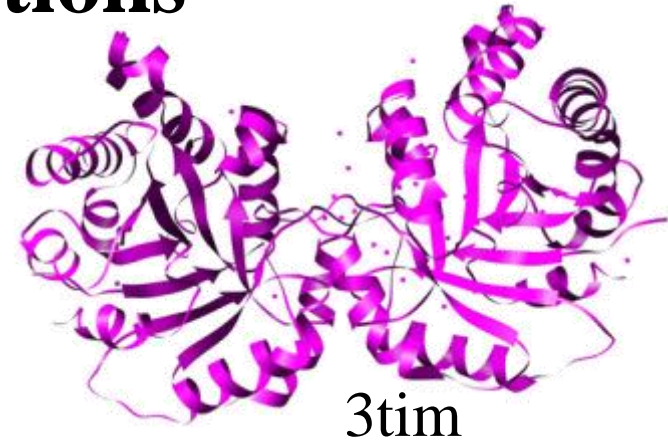
Structure Expectations

Protein

- usually 3D
- rarely secondary structure

RNA

- usually secondary structure
- rarely 3D



Structural Data

Proteins

- 8.1×10^5 or about 2×10^5 interesting ones

RNA

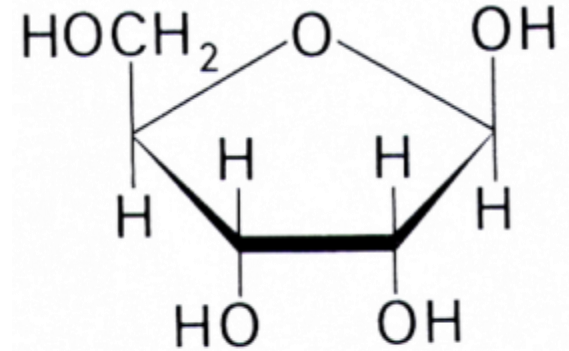
- 2.3×10^3 structures with some RNA
- 145 with RNA + DNA (no protein)
- 9×10^2 with pure RNA

Determining structures

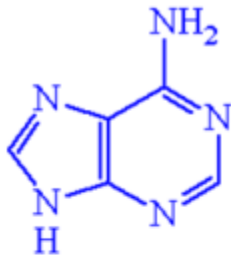
- general – RNA hard to handle (RNases)
- crystallography
- NMR
 - assignments very difficult (only 4 kinds of base)

RNA structure

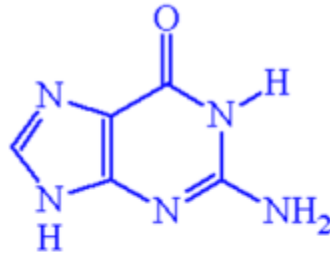
- 3 components
 - ribose (sugar)
 - phosphate (PO_4)
 - base (nucleotide)



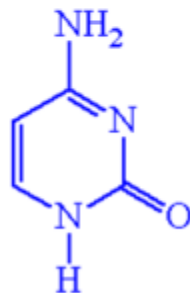
β -D-Ribose



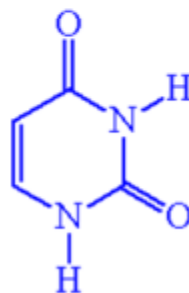
Adenine



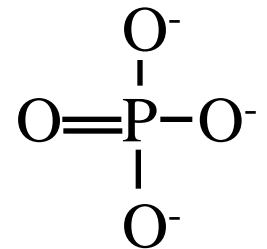
Guanine



Cytosine



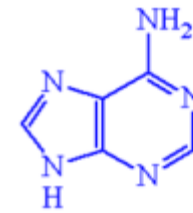
Uracil



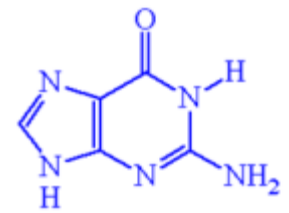
RNA Bases

- Are they like protein residues ?
 - not classified by chemistry
 - do they have interactions ?
 - yes

purines

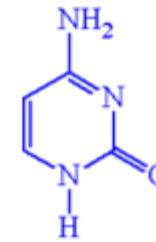


Adenine

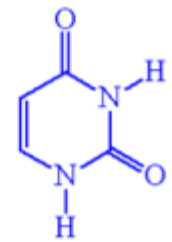


Guanine

pyrimidine

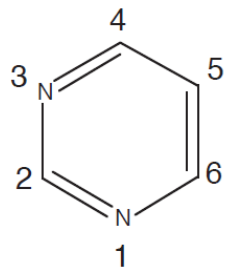


Cytosine

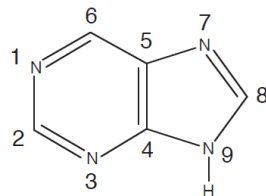


Uracil

- mother shapes



pyrimidine

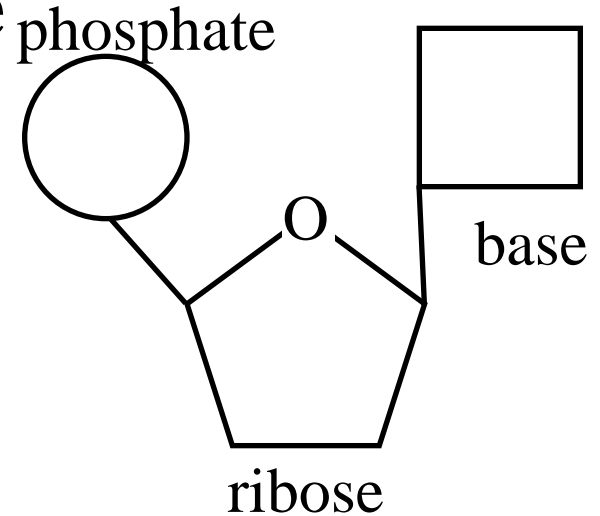


purine

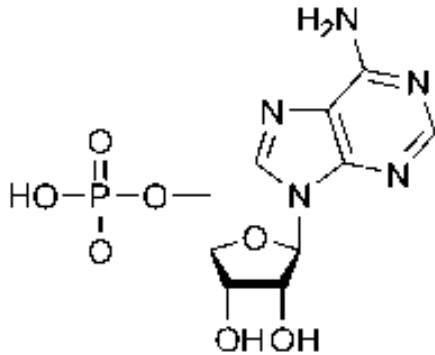
- numbering not used much
- putting pieces together...

RNA structure

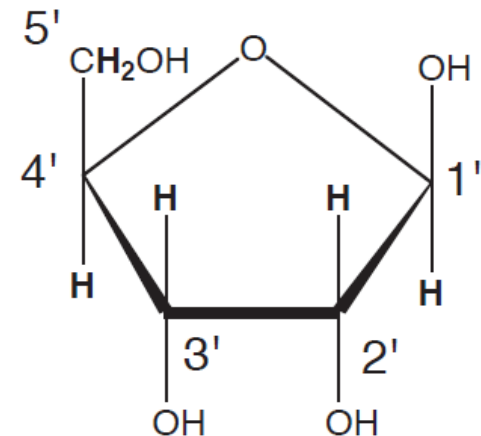
- joining the components



- adenosine 5'-monophosphate
 - not adenine, adenosine, ...

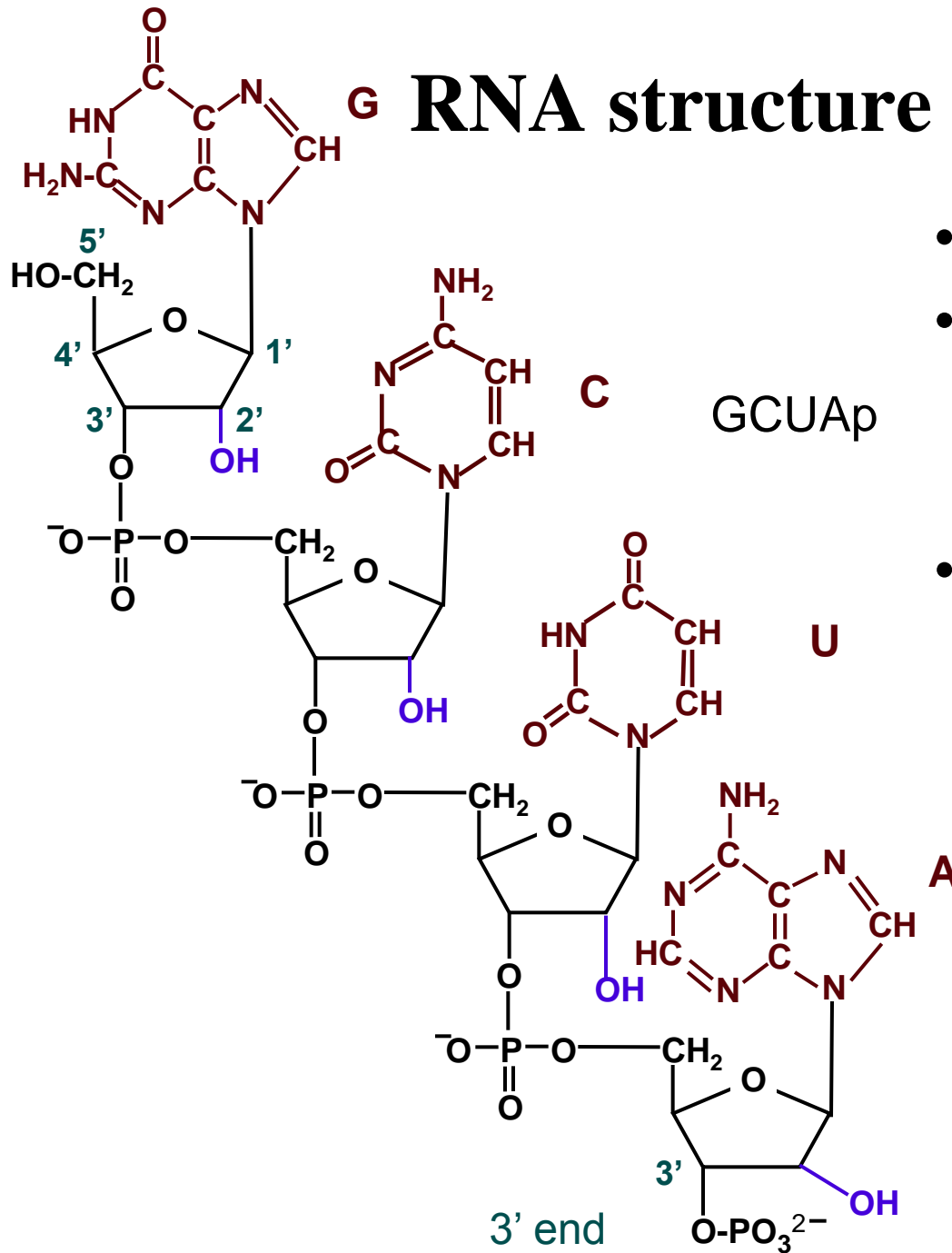


- note numbering on sugar ring



G RNA structure

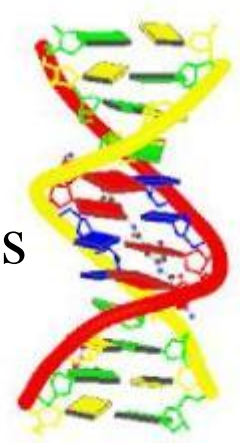
5' end



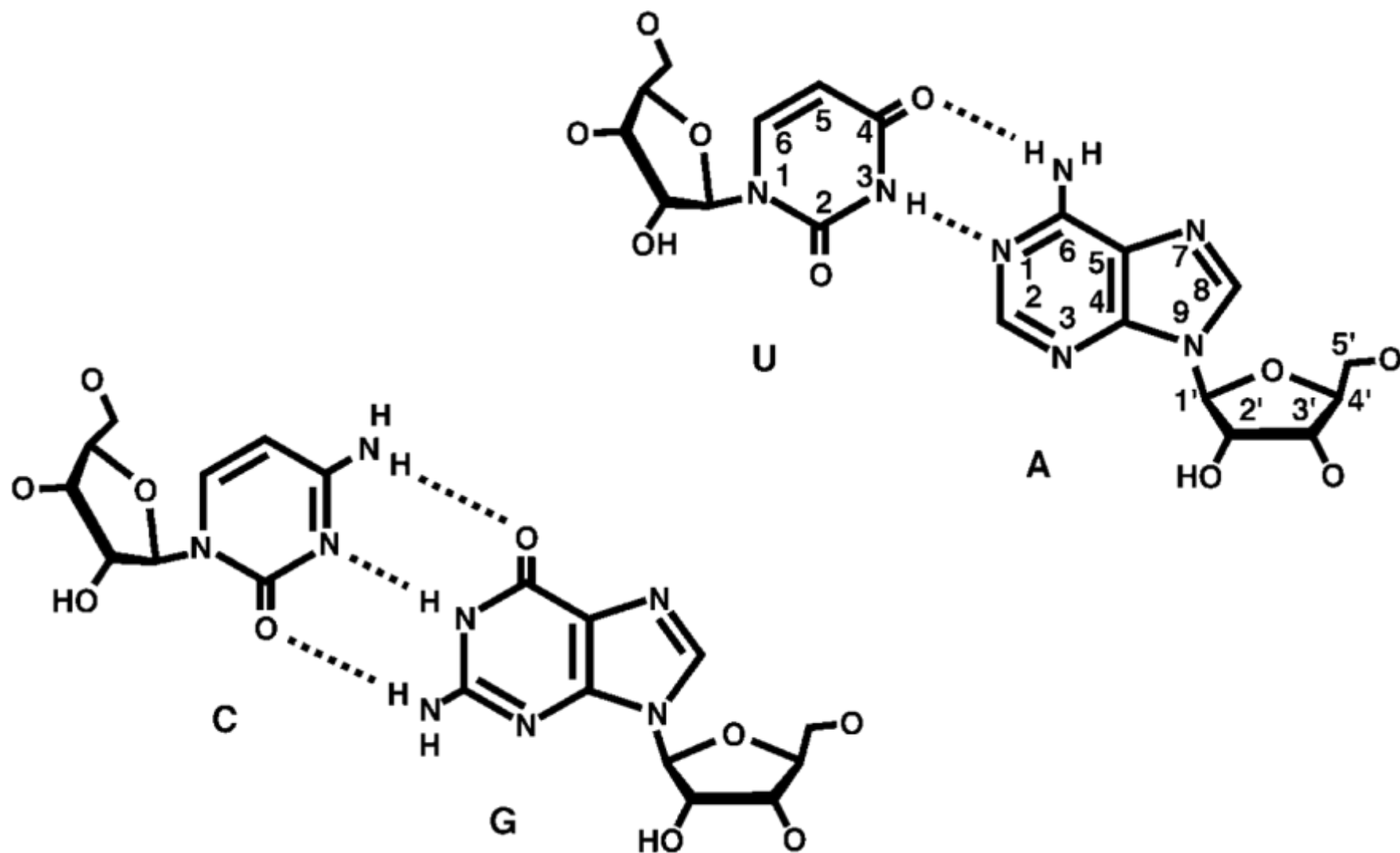
- negative charges
- directional
 - 5' to 3'
- notation
 - always 5' to 3'

H bonding

- What holds the pairs of a helix together ? H-bonds
 - applies to RNA
 - rules from proteins
 - H-bond donors are NH, OH
 - acceptors – anything with partial –'ve
- Historic H-bonding pairs...



Historic H-bonding pairs



Historical point

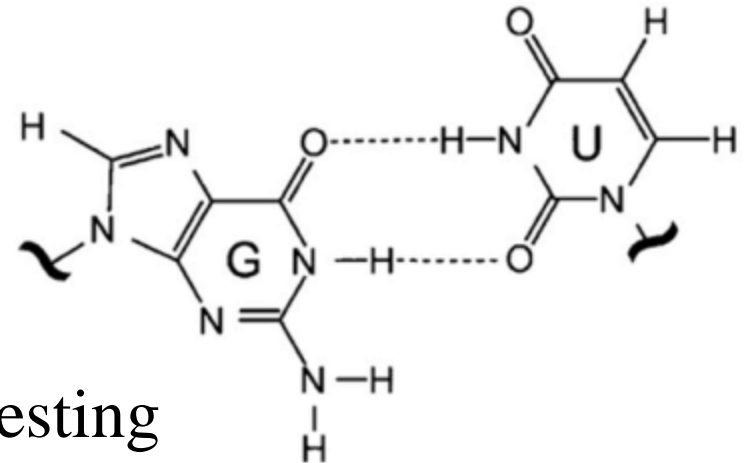
- RNA has 4 bases + GC, AU base pairs
- H-bond pairs look flat

BORING

- other kinds of H-bonds and bases
- base pairs are not perfectly flat

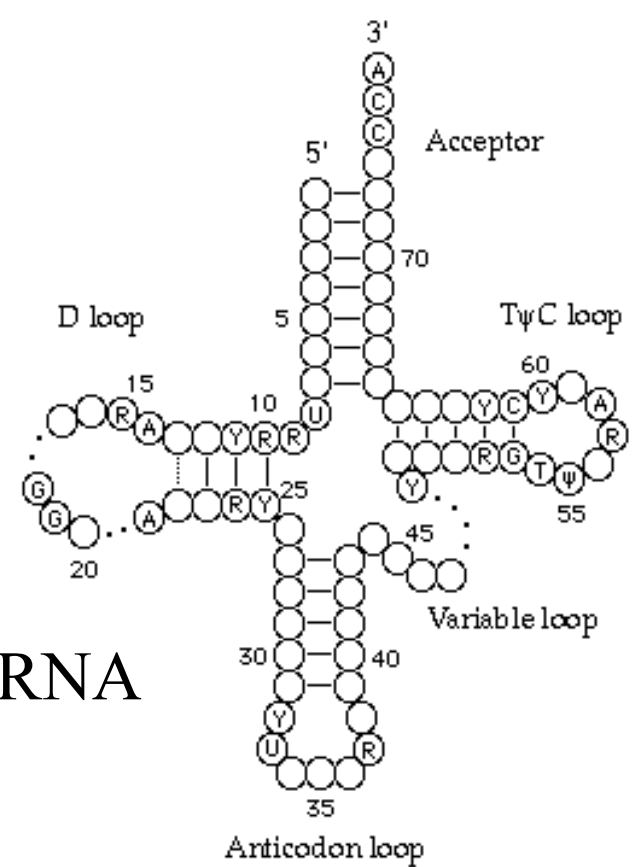
Other common H-bond partner

- Contrast with DNA (GC and AT)
 - rarely violated (mismatch)
- Interesting base pairing
 - RNA (GC, AU) much more interesting
 - third base pair GU (rather common)
 - lots of weaker pairs possible



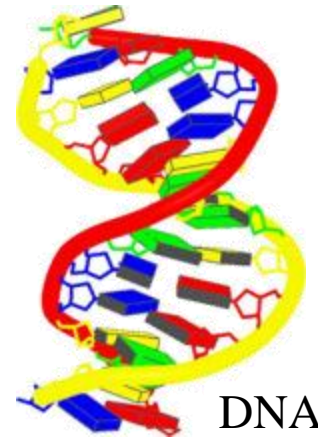
More bases

- standard machinery for copying DNA→RNA (standard base pairs)
- every tRNA has a modified base

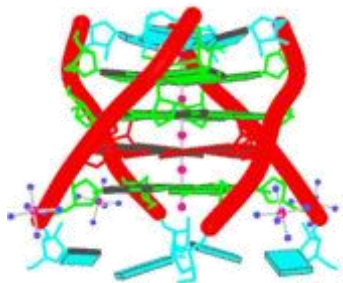


Possible RNA structures

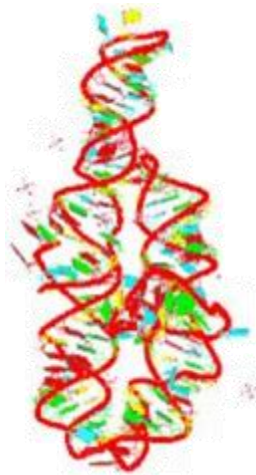
- DNA ? nearly always similar helix
 - some debate about A, B, Z, ..
- RNA
 - lots of varieties known
 - nomenclature..



DNA
duplex
140D



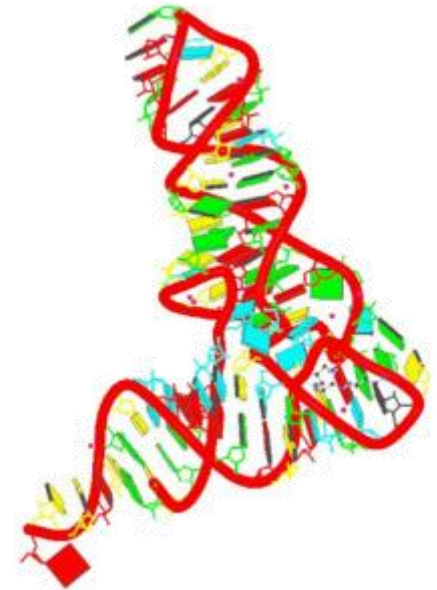
tetraplex
1mdg



group I intron
1hr2



hammerhead
2oeu



tRNA
1evv

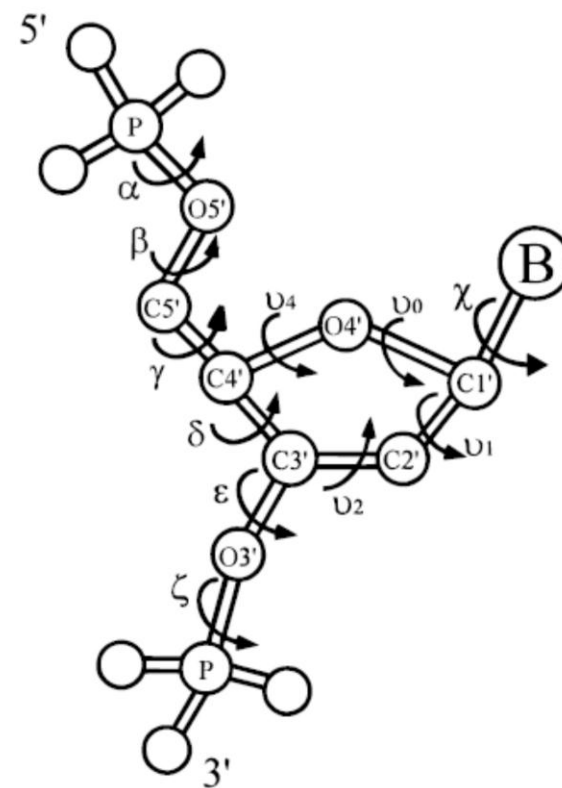
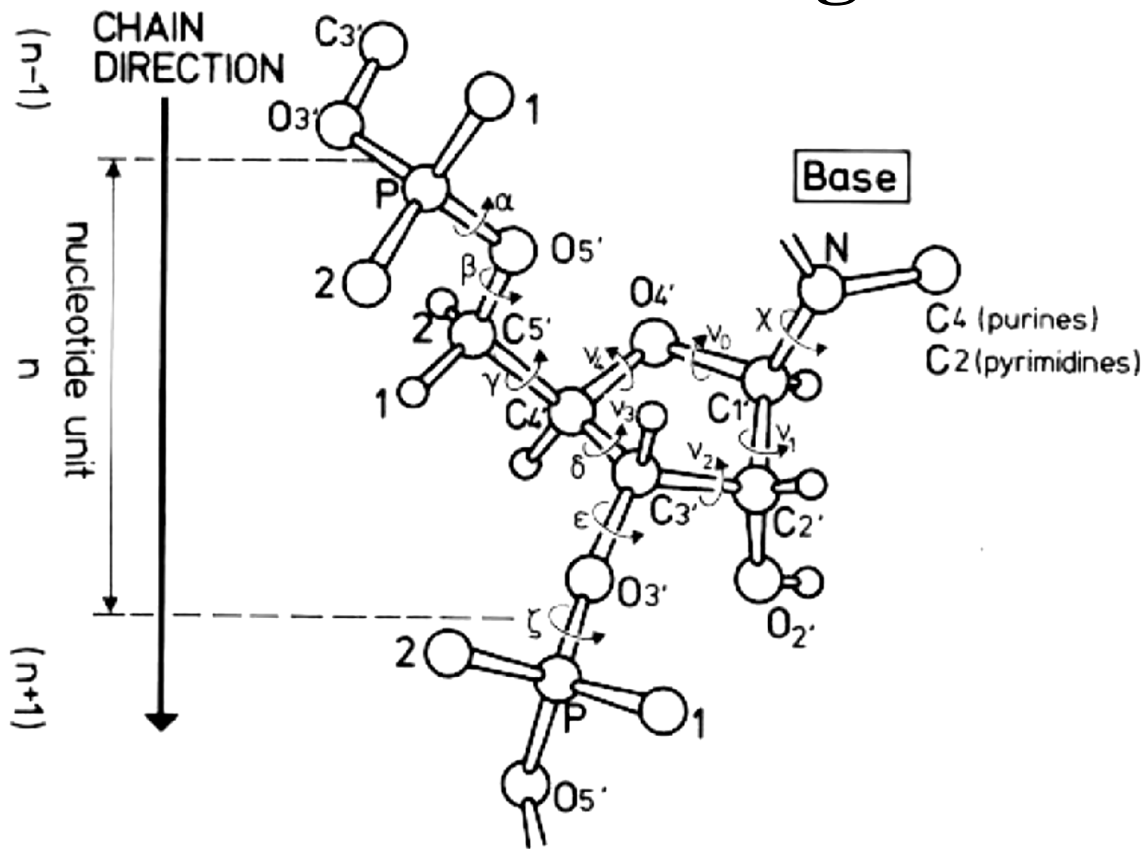
RNA coordinates / nomenclature

- As for proteins: PDB format

ATOM	1	O5*	G A 103	58.355	47.332	91.116	1.00175.32
ATOM	2	C5*	G A 103	57.373	48.210	90.636	1.00175.32
ATOM	3	C4*	G A 103	56.962	47.802	89.224	1.00175.19
ATOM	4	O4*	G A 103	58.148	47.463	88.474	1.00175.34
ATOM	5	C3*	G A 103	56.096	46.543	89.152	1.00175.03

- As for proteins
 - dihedral angles are useful
- Unlike proteins (φ, ψ) there are 8 ($\alpha, \beta, \gamma \dots$)

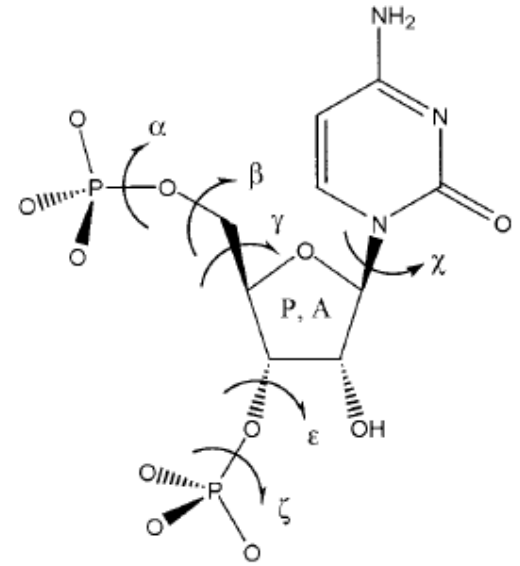
dihedral angle nomenclature



from Marino, JP, Schwalbe, H., Griesinger, C, Acc. Chem. Res. 32, 614-623 (1999)

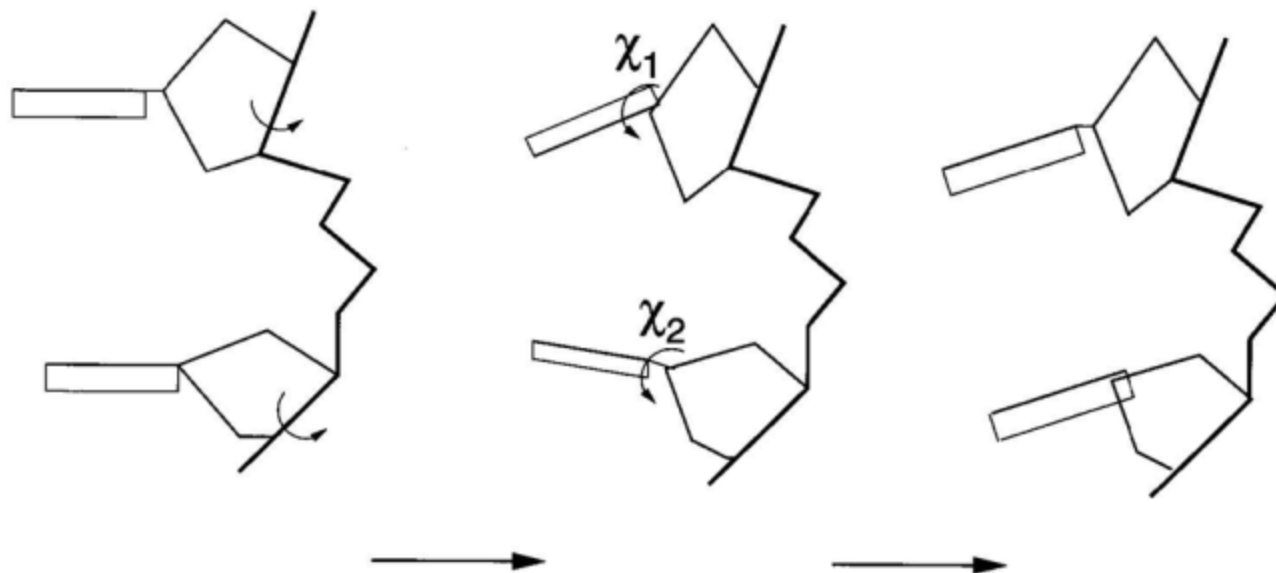
dihedral angle nomenclature

- 8 angles
 - α , β , γ , ε , ζ , χ
 - 2 for sugar (P, A)
- too many for me – how to simplify ?
- what if two angles are highly correlated ?
 - if we know x , then y is probably known
- ideas for classification...

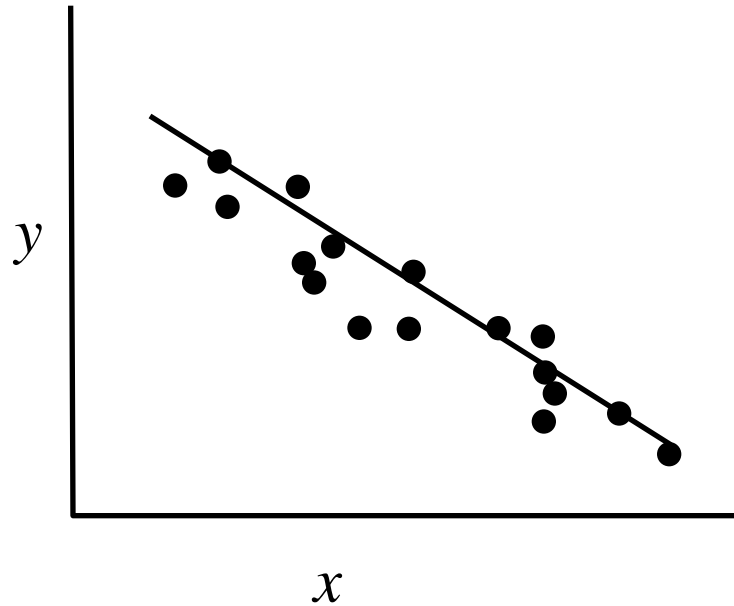


Describing RNA conformation

- Example approach – look for correlations
 - principle component analysis (quick detour if necessary)
- what if sugars move in two residues ?
 - energetically, would like to maintain base pairing...
 - P, A move, χ will compensate
 - χ will be correlated with sugar angles



PCA reminder



I have two dimensional data

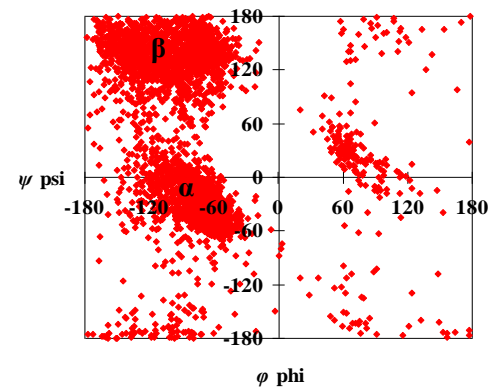
- could well be described by a first (component) and
- maybe second component

n -dimensional data

- how much of variance is described by 1st, 2nd, ... components

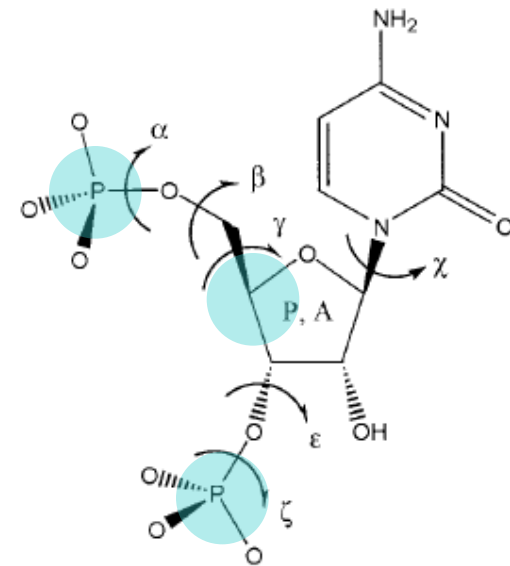
Describing RNA conformation

- Claim – from one DNA paper
 - most conformations are well described by 3 variables
- alternative...
 - do not work in terms of real dihedral angles
 - invent reference points
 - example study...
 - Duarte, CM & Pyle, AM, (1998) 284, 1465-1478
- remember ramachandran plots in proteins
 - can one do something similar in RNA ?



Reduced RNA conformation

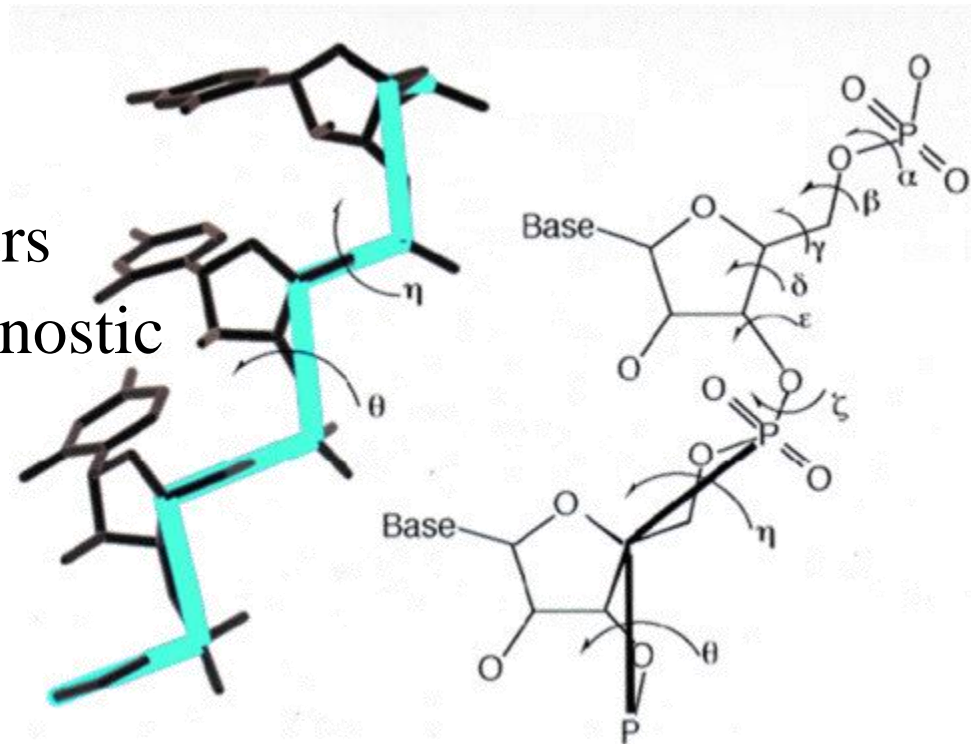
- Basic idea
 - pick 4 atoms that are not sequential
 - define a simplified backbone
 - $P-C_4-P-C_4-P-C_4-\dots$
 - leads to "pseudo-torsion" angles
- η
 - $C4_{n-1}-P_n-C4_n-P_{n+1}$
- θ
 - $P_n-C4_n-P_{n+1}-C4_{n+1}$



Reduced RNA conformation

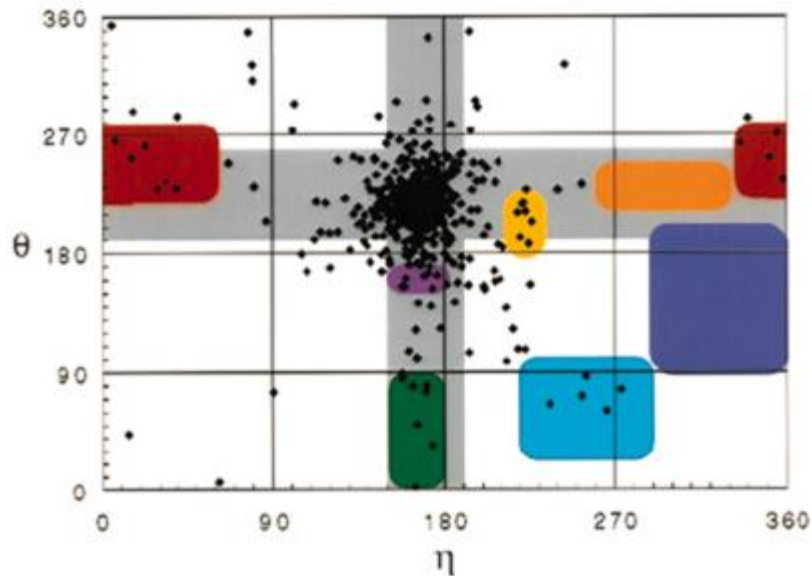
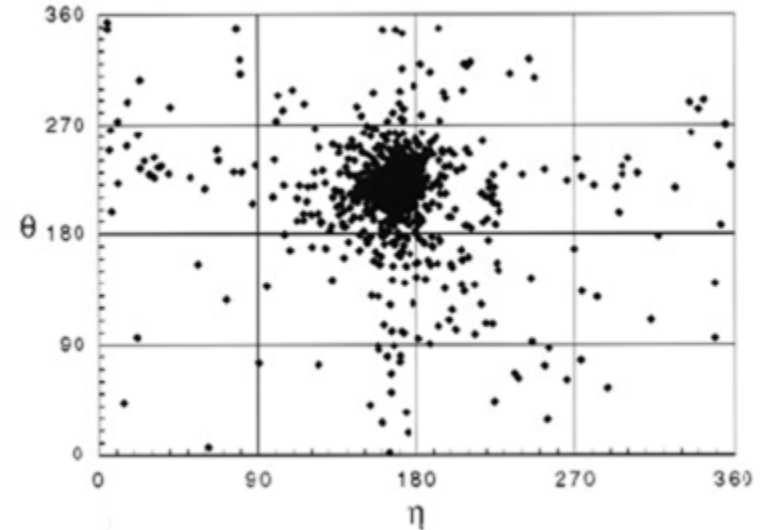
Plan of authors

- take 52 structures
 - (≈ 700 nucleotides)
 - collect η , θ
 - see if there are clusters
 - see if angles are diagnostic

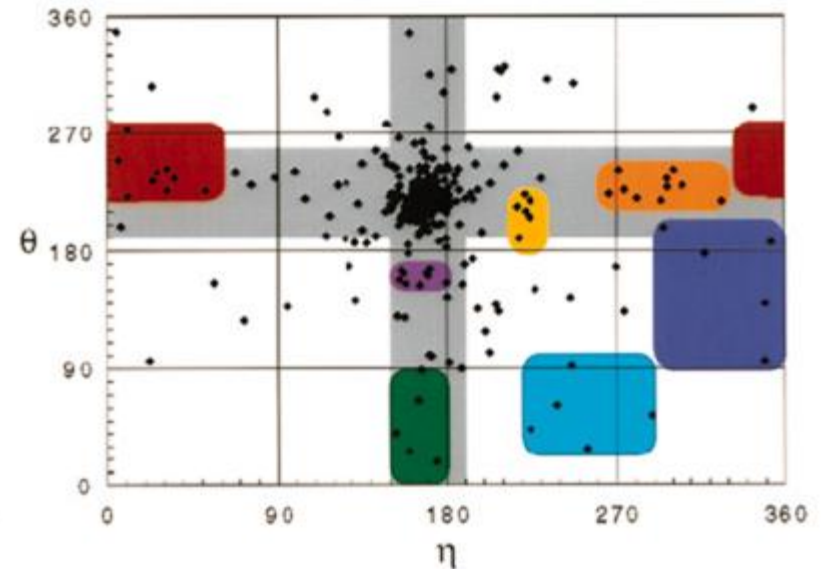


Reduced RNA conformation

- Do you see clusters ?
 - main set of points ...
 - boring RNA helix
- a big claim



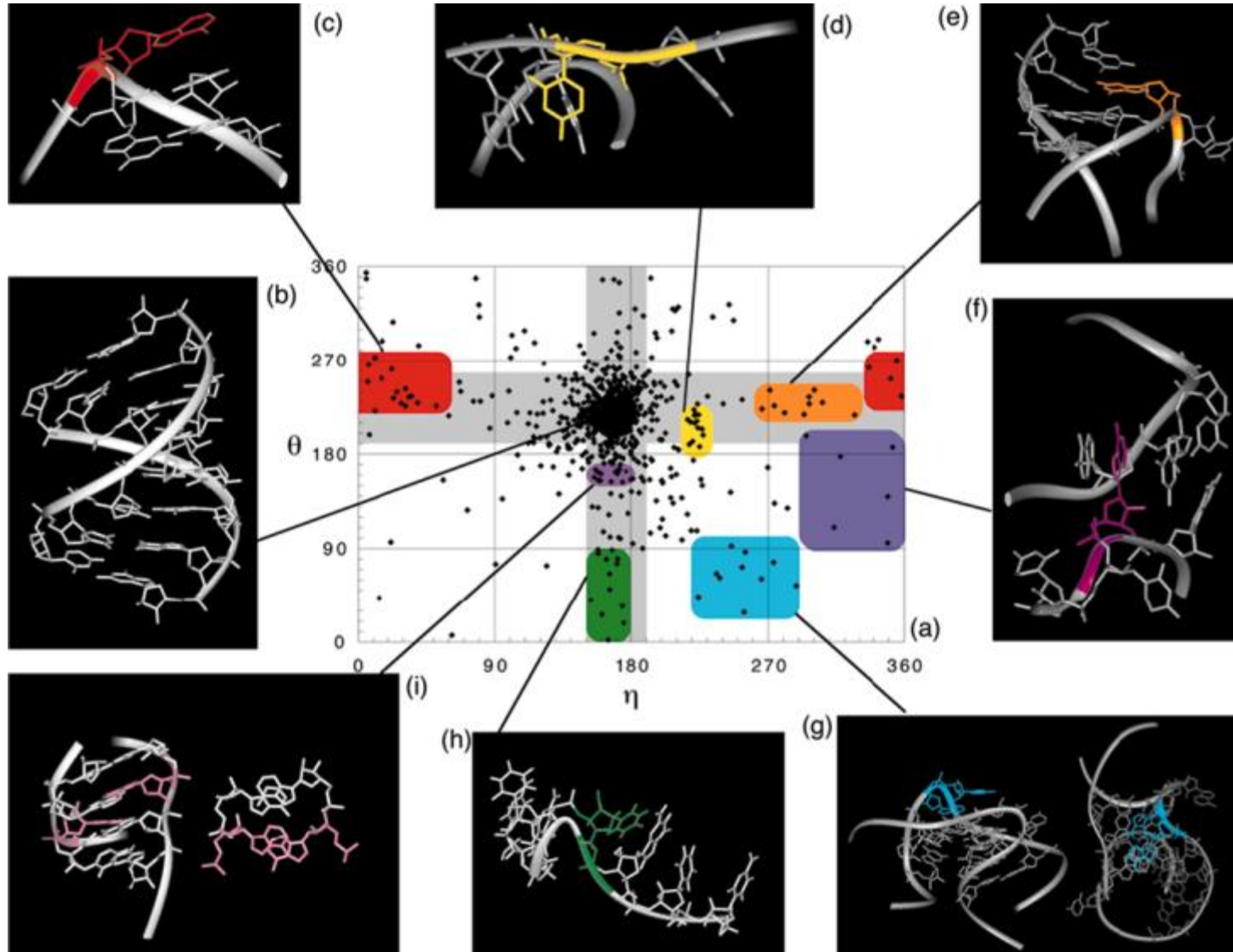
no tertiary interactions



yes tertiary interactions

Reduced RNA conformation

- with a bit more human interpretation



Reduced RNA conformation

We are interested in a critical look at ideas

How to read this...

- if you measure a pair of η , θ pseudo-angles
 - could you guess if something is wrong in structure ?
 - could you use this to categorise the conformation ?
- are there better ways to categorise structure ?

Summary

- RNA structure as per Watson-Crick, old text books
- How are RNA structures different to DNA ?
- What are the biological roles ?
- Can we neatly summarise RNA structures ?
 - see what information (angles) are necessary
 - define alternative angles
- Next..
 - What is life ?