



Estructura y biofísica de ácidos nucleicos y proteínas

David Dufour Rausell
ddufour@pcb.ub.cat

Summary

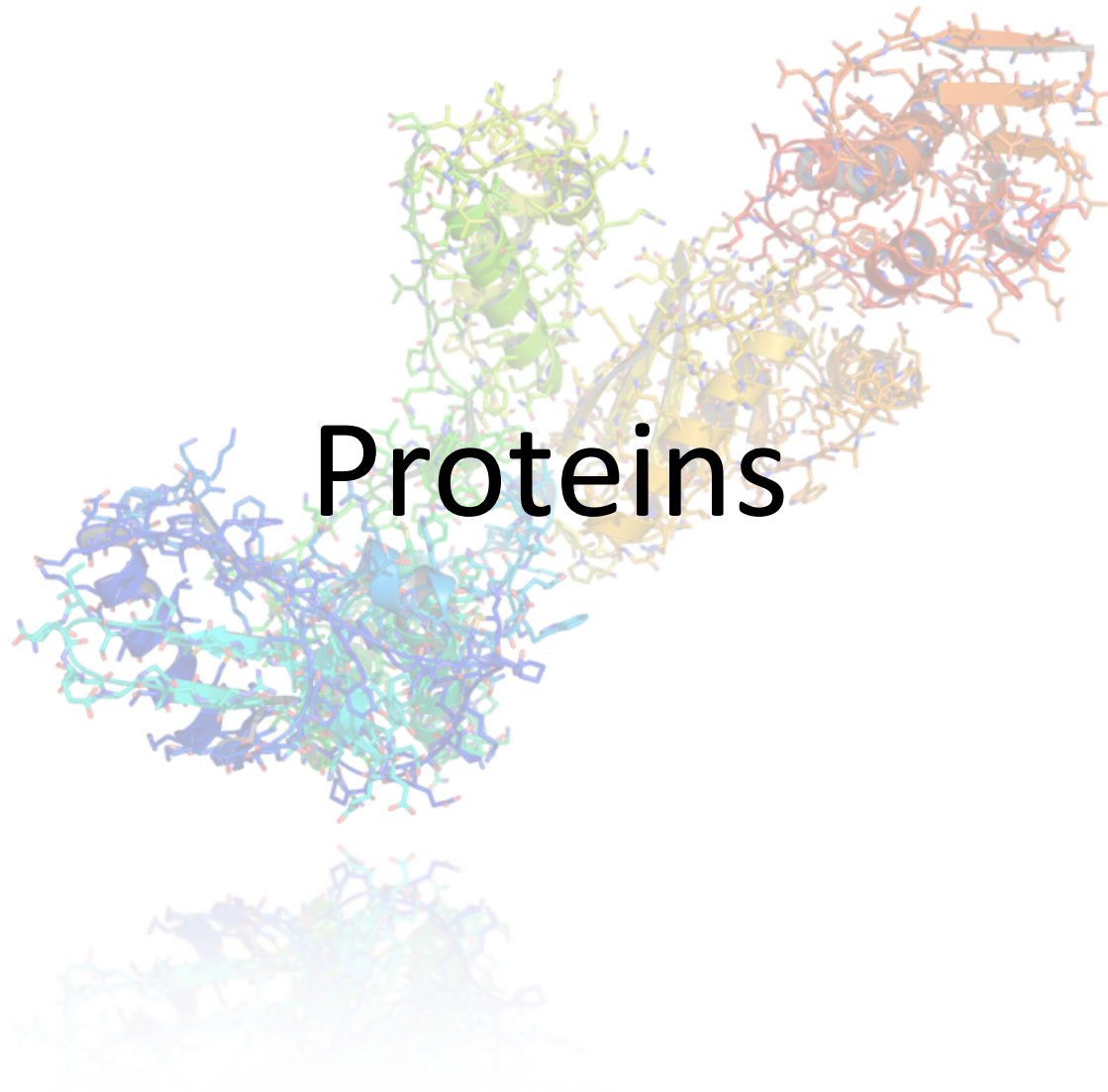
Proteins

- Primary to quaternary structure
- Peptide Bond
- Amino acids
- Torsion angles
- Ramachandran plot
- Secondary structures: α -helix, β -sheets
- Fold space
- SCOP, CATH, PDB

Summary

Nucleic Acids (DNA, RNA)

- Bases, sugar, phosphate
- Phosphodiester bond
- Numbering system
- Sugar Puckering
- Orientation around glycosidic angle
- C4'-C5' torsion angle
- RNA rotamers
- Base pairs
- Triples
- Helical parameters
- Helical grooves
- Secondary structures in RNA
- Forces that drive RNA folding
- NDB



Nomenclature

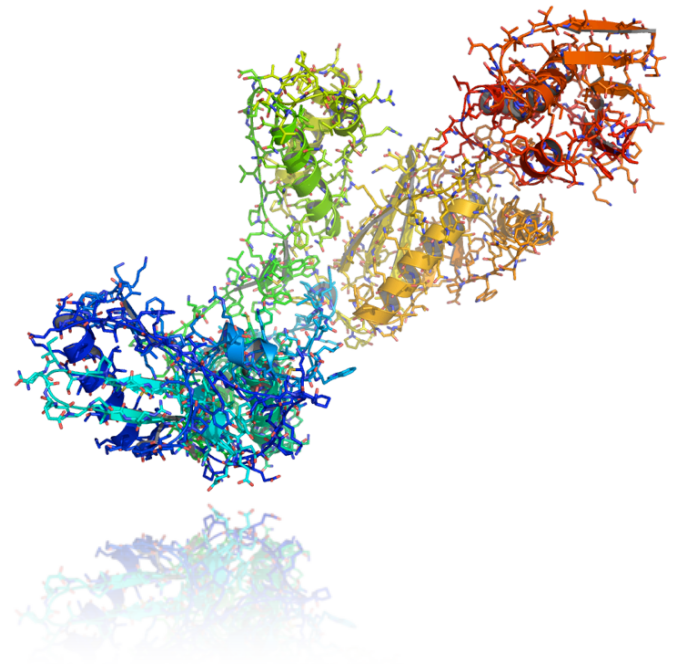
Fold: Three dimensional conformation of a protein sequence (usually at domain level).

Domain: Structurally globular part of a protein, which may independently fold.

Secondary Structure: Regular sub-domain structures composed by alpha-helices, beta-sheets and coils (or loops).

Backbone: Protein structure skeleton composed by the carbon, nitrogen and oxygen atoms.

Side-Chain: Specific atoms identifying each of the 20 residues types.



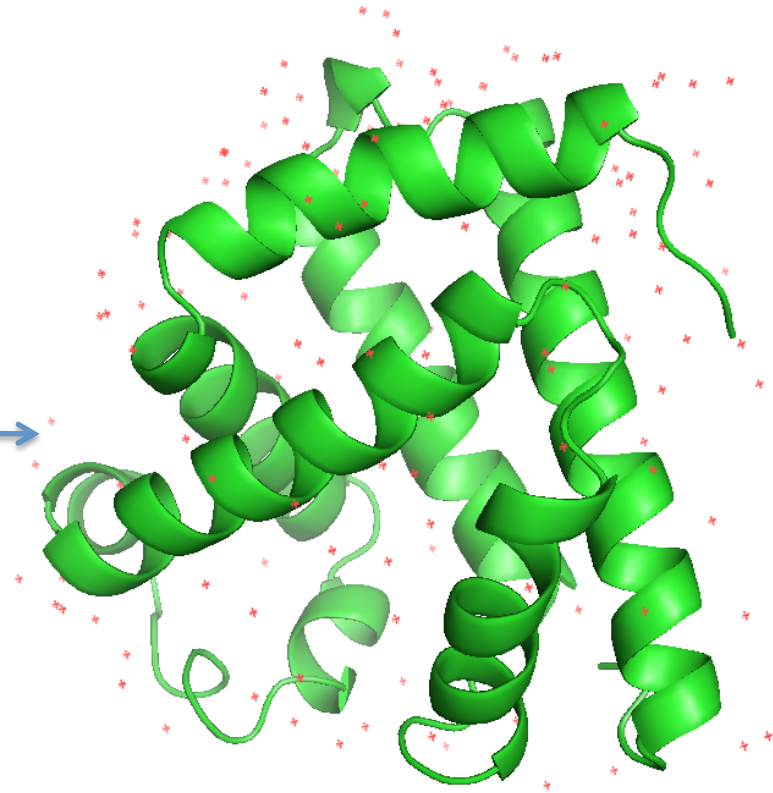
Primary structure

MVLSEGEWQLVLHVWAKVEA
DVAGHGQDILIRLFKSHPETLEK
FDRFKHLKTEAEMKASEDLKKH
GVTVLTALGAILKKKGHHEAELK
PLAQSHATKHKIPKYLEFISEAII
HVLHSRHPGNFGADAQGAMN
KALELFRKDIAAKYKELGYQG

Secondary structure

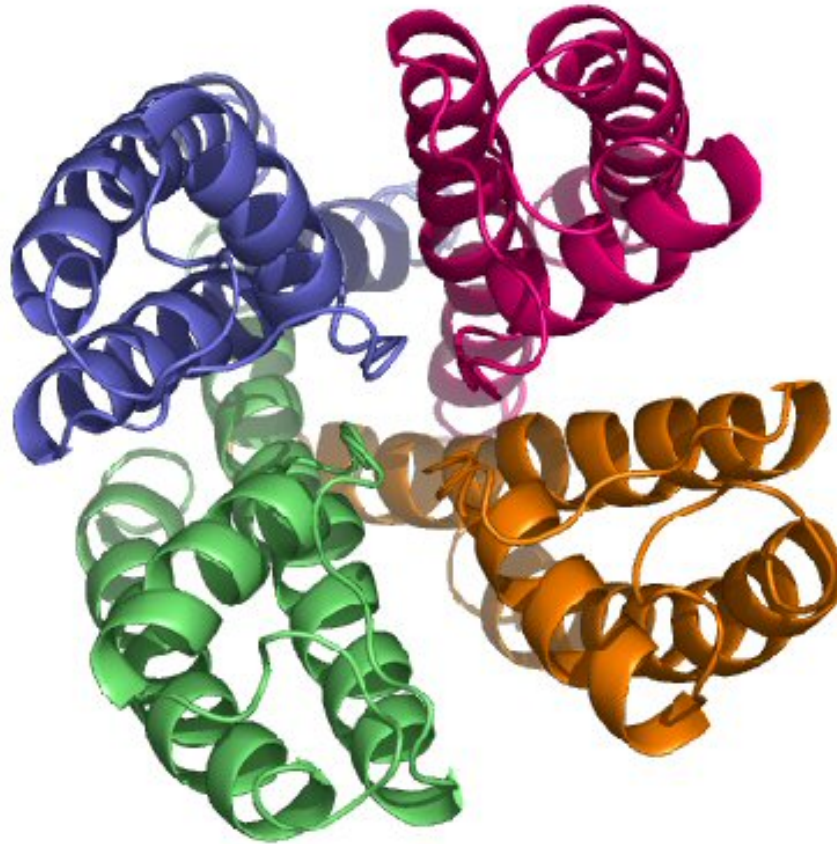


Tertiary structure

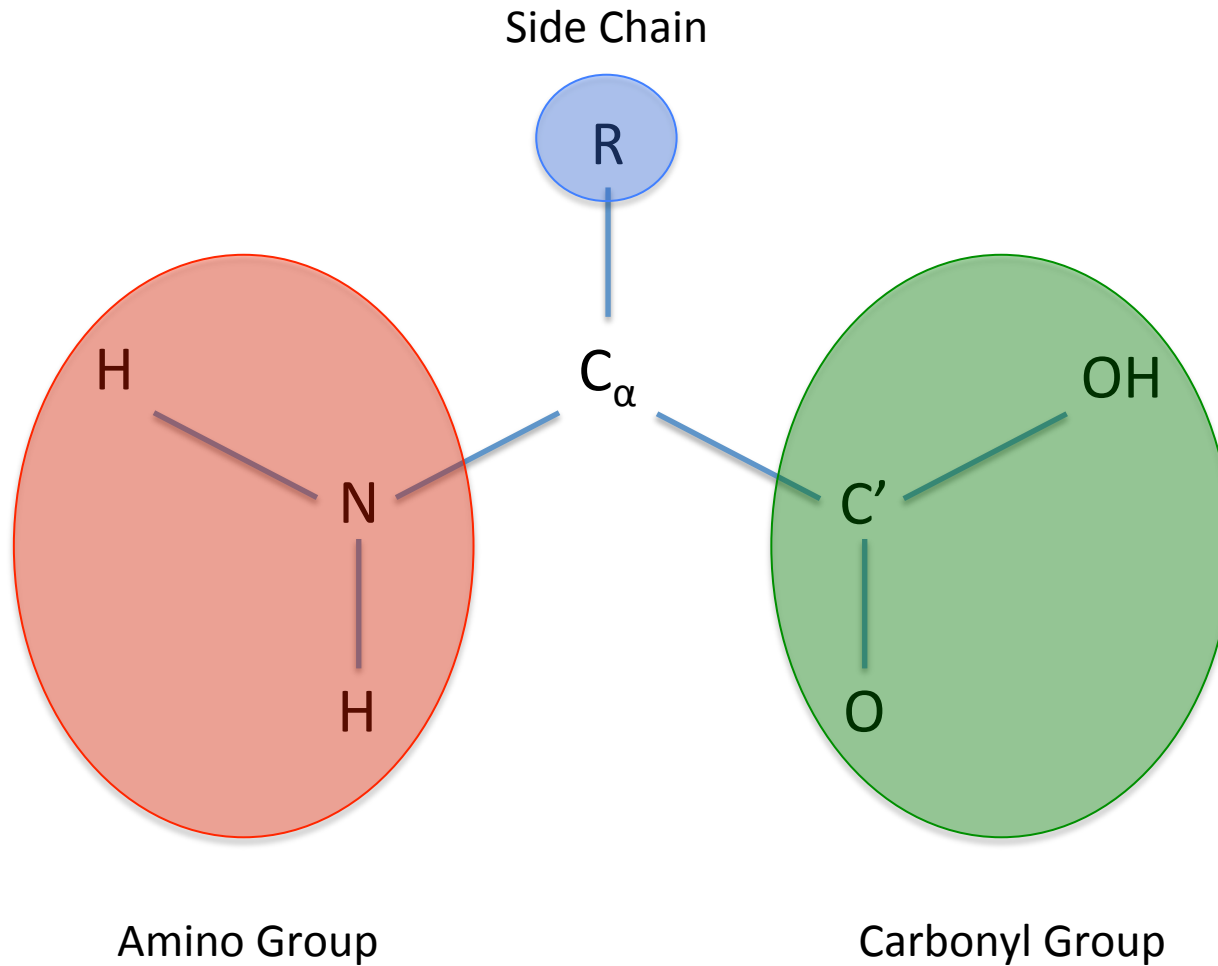


(Myoglobin)

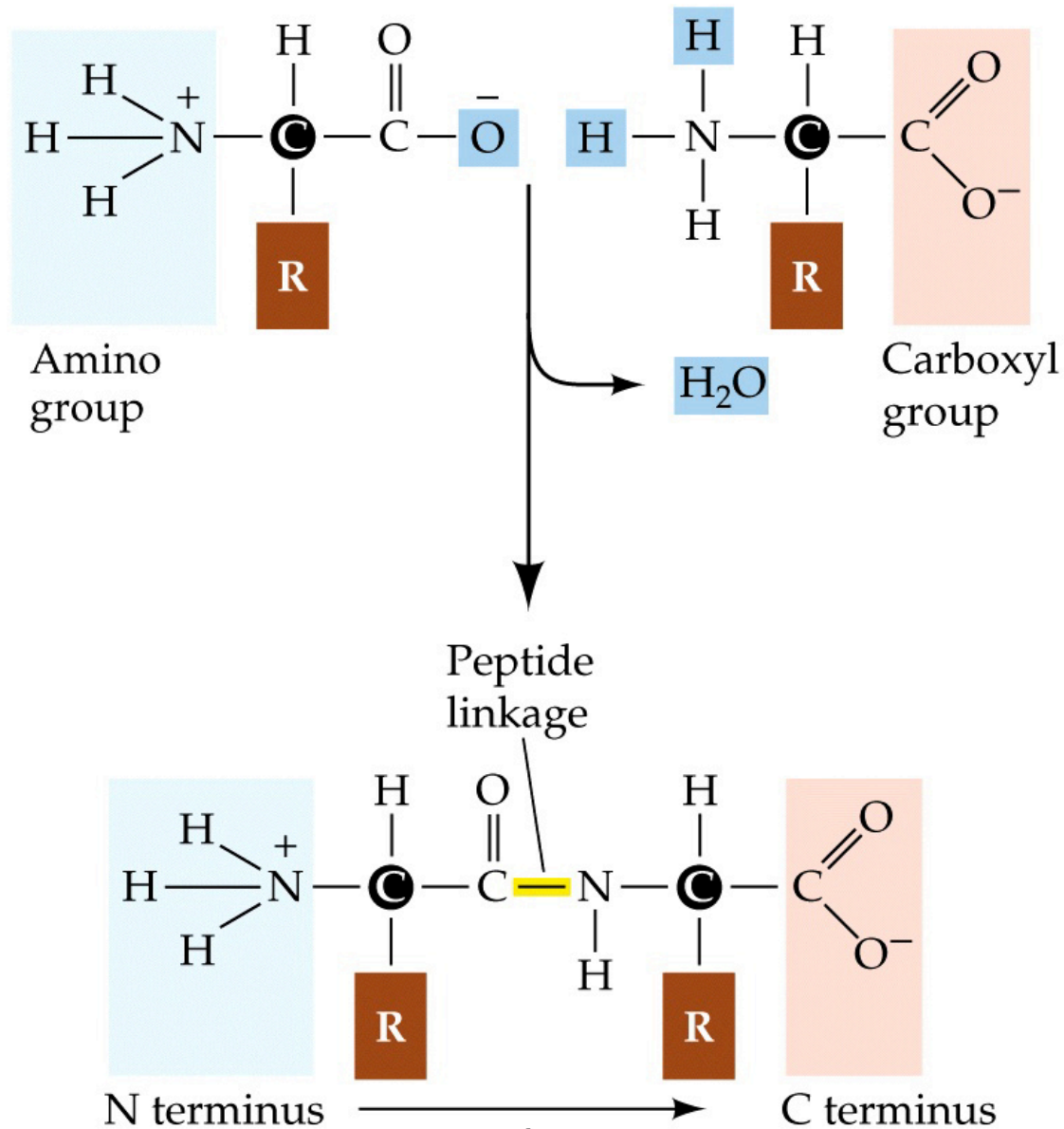
Quaternary structure



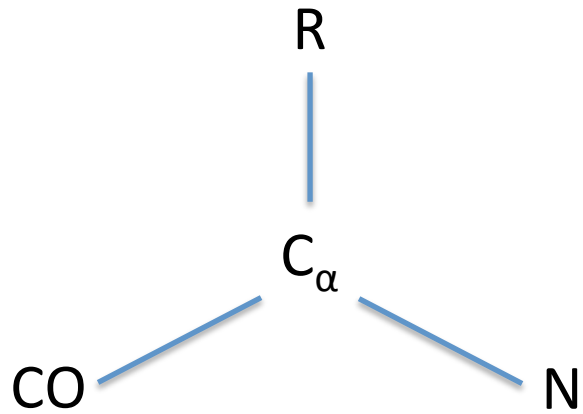
Proteins are polypeptide chains



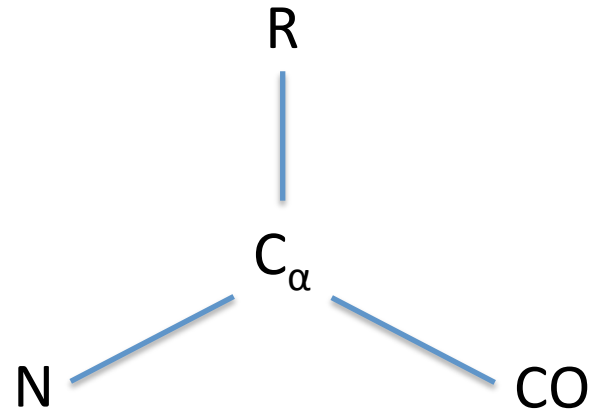
Peptide Bond



L- and D- forms

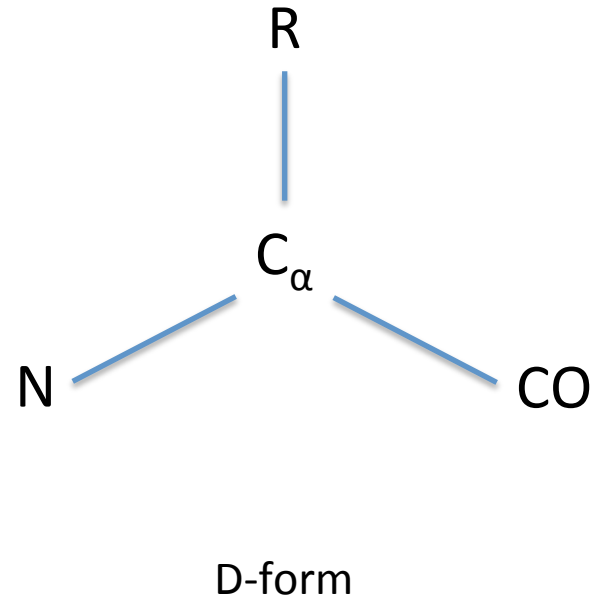
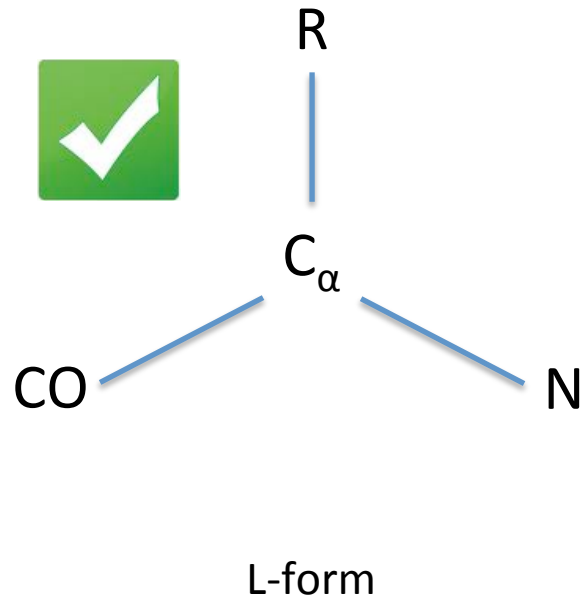


L-form



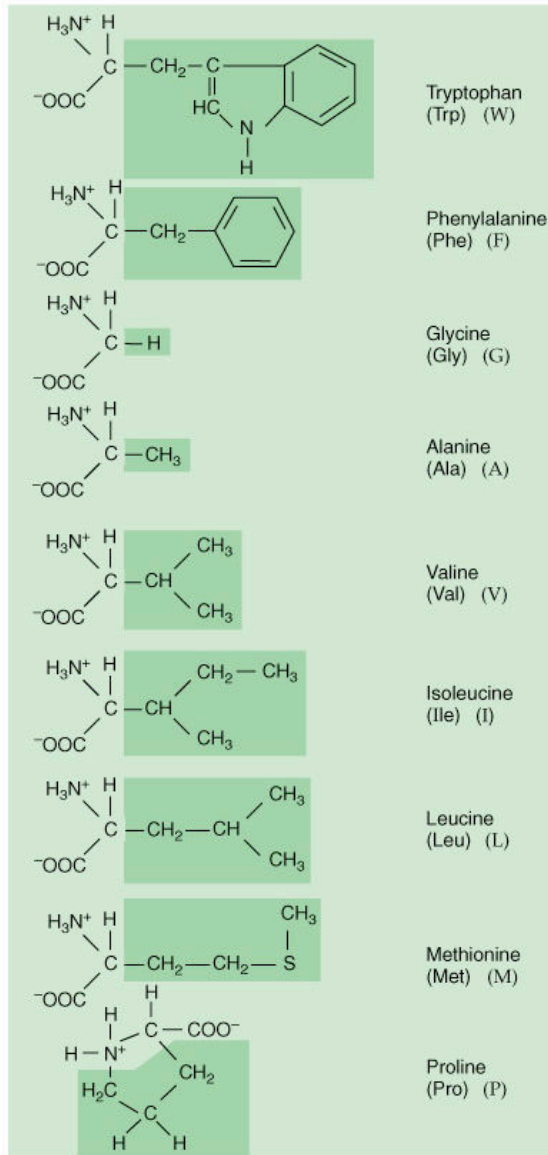
D-form

L- and D- forms

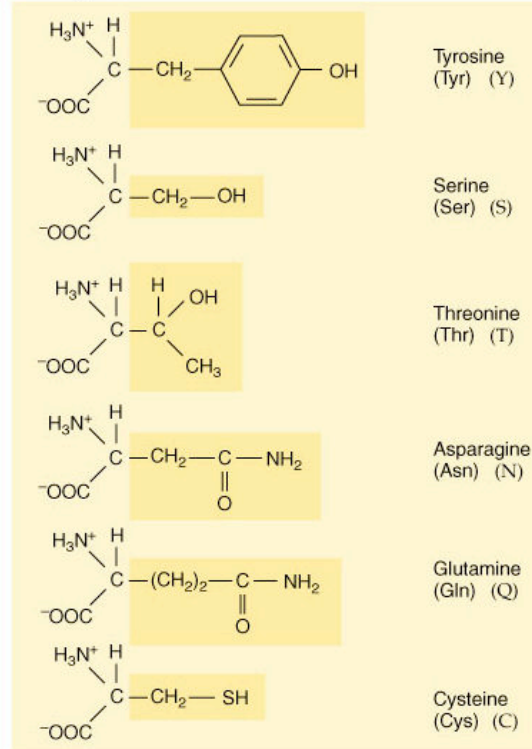


The 20 amino acids

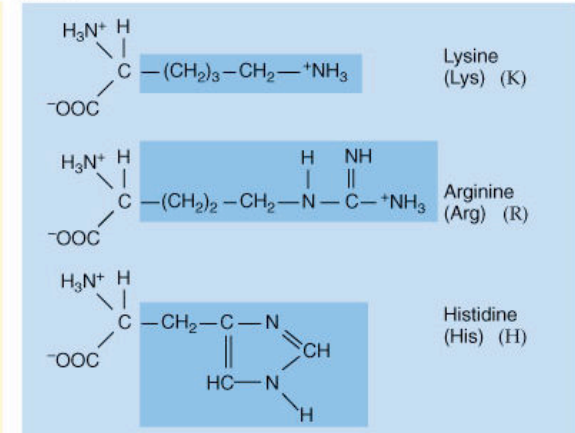
Neutral, nonpolar



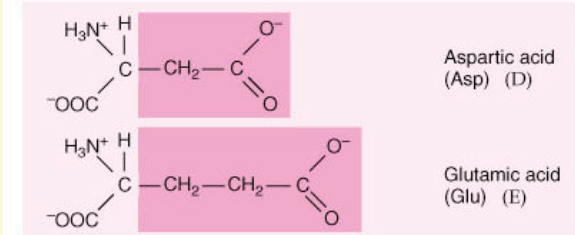
Neutral, polar



Basic

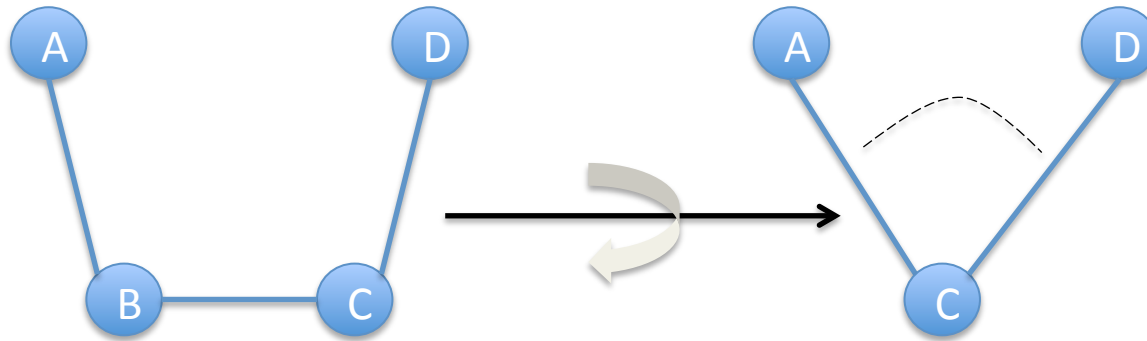


Acidic

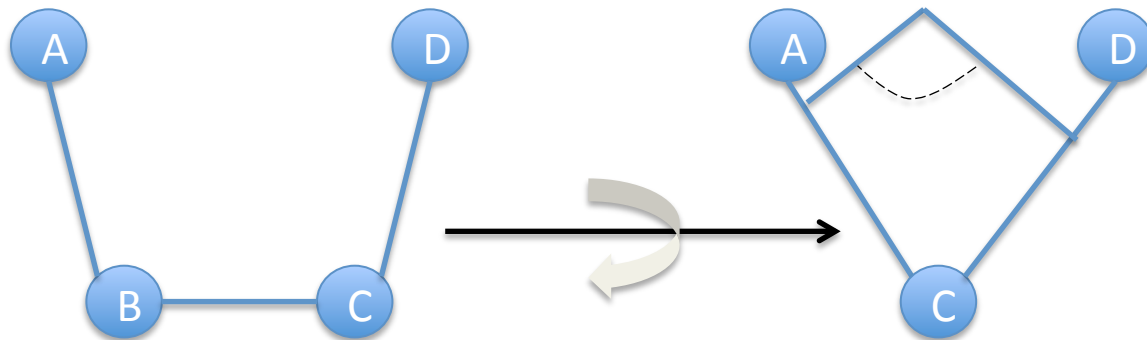


Torsion and dihedral angles

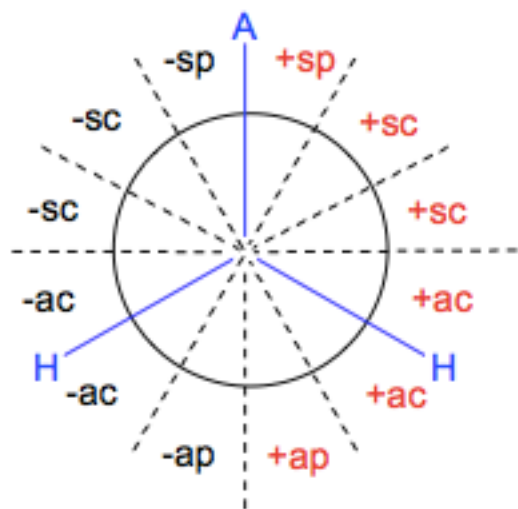
Torsion
angle



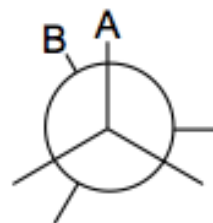
Dihedral
angle



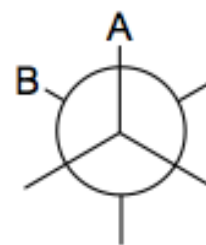
Torsion angle classification



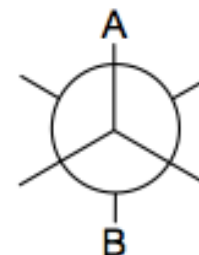
s: syn
p: periplanar
a: anti
c: clinal



Cis



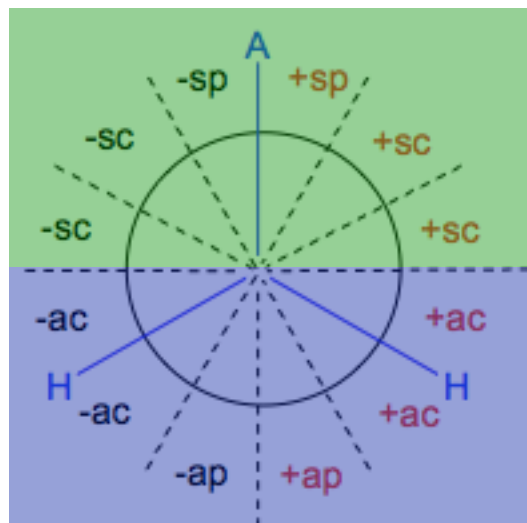
Gauche



Trans

Klyne-Prelog system

Torsion angle classification



s: syn

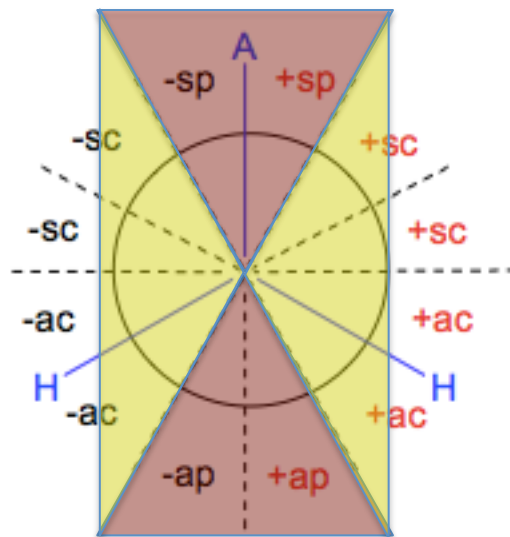
p: periplanar

a: anti

c: clinal

Klyne-Prelog system

Torsion angle classification



s: syn

p: periplanar

a: anti

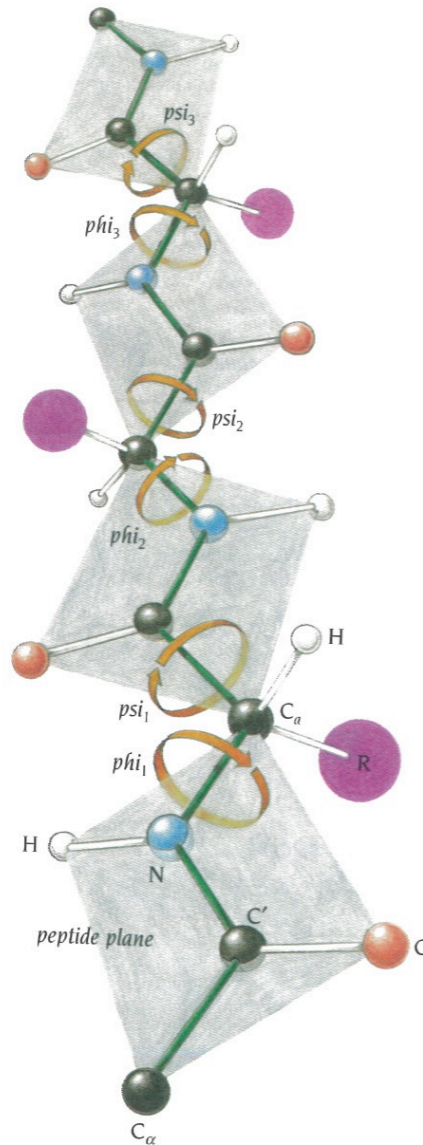
c: clinal

Klyne-Prelog system

Peptide units = suits

Phi (φ) = $C'-C_{\alpha}-N-C'^{+1}$,
rotation around $C_{\alpha}-N$

Psi (Ψ) = $N-C'-C_{\alpha}-N^{+1}$,
rotation around $C'-C_{\alpha}$



Ramachandran plot

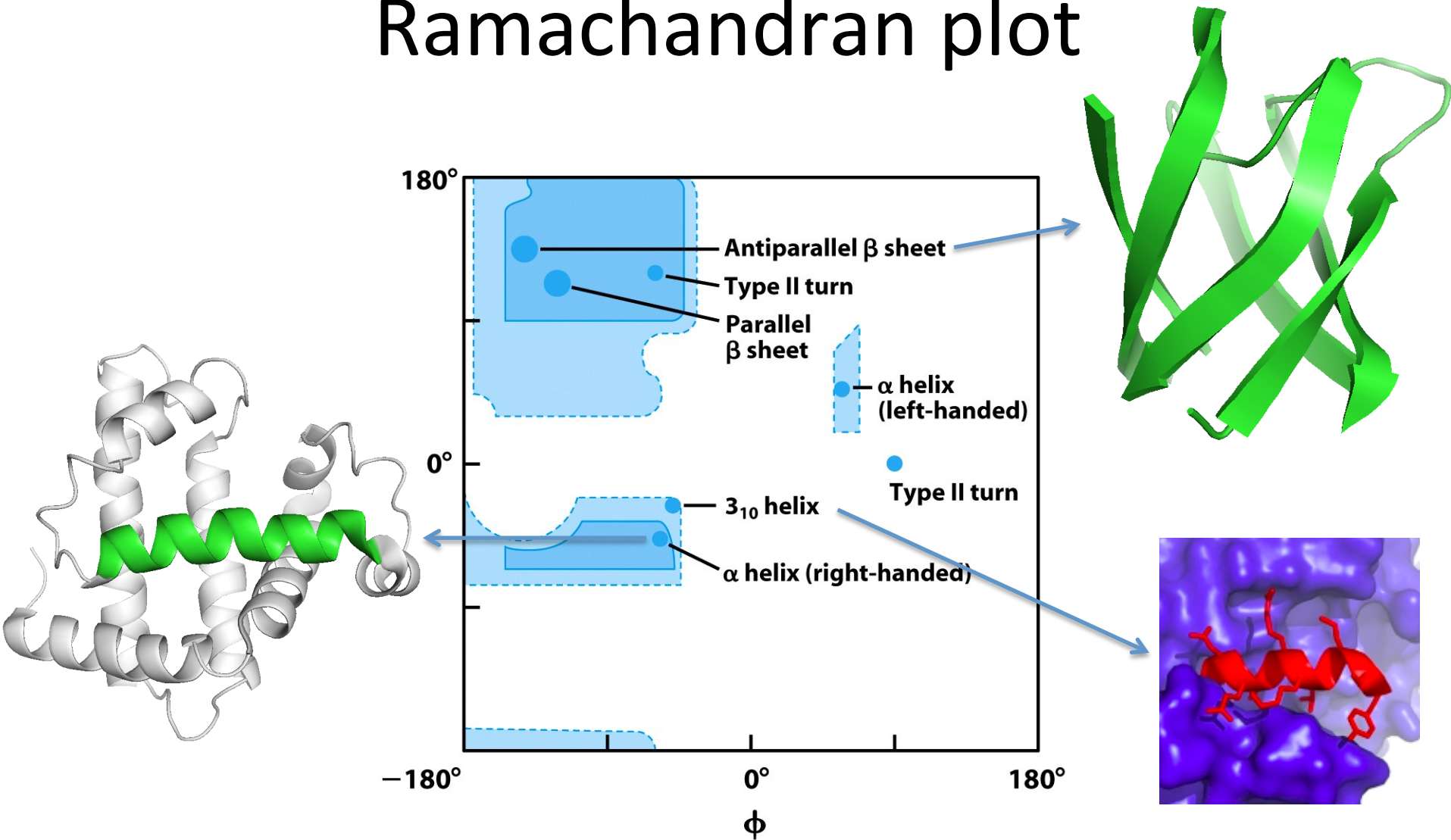
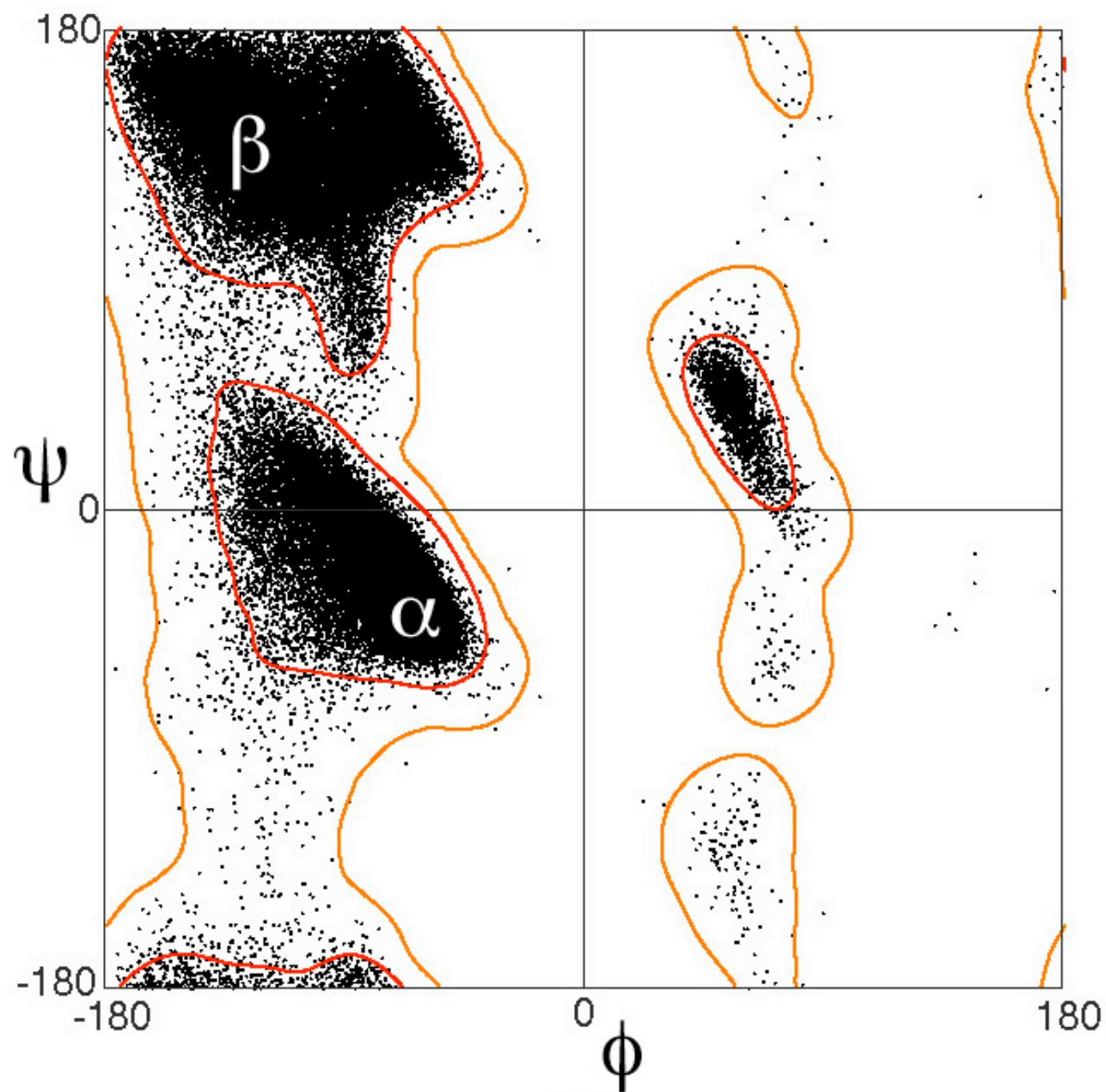
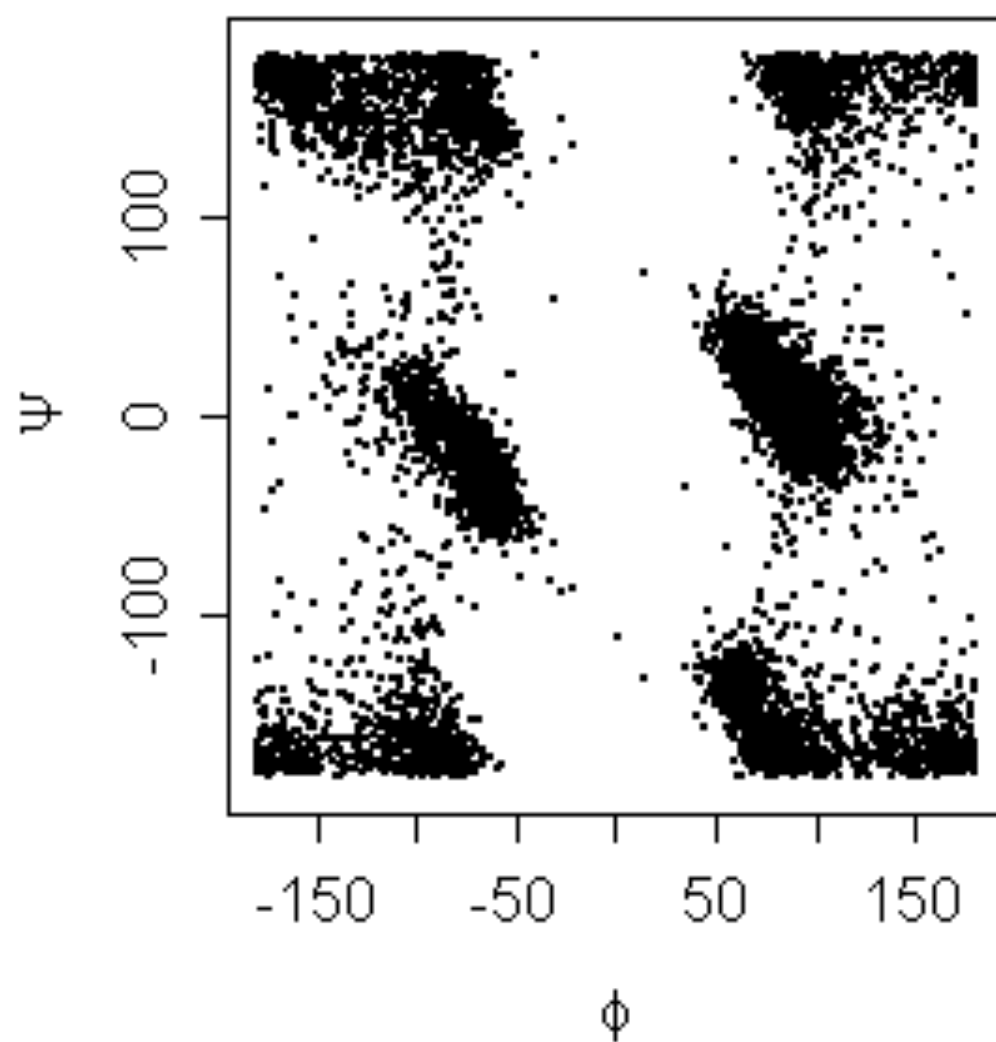


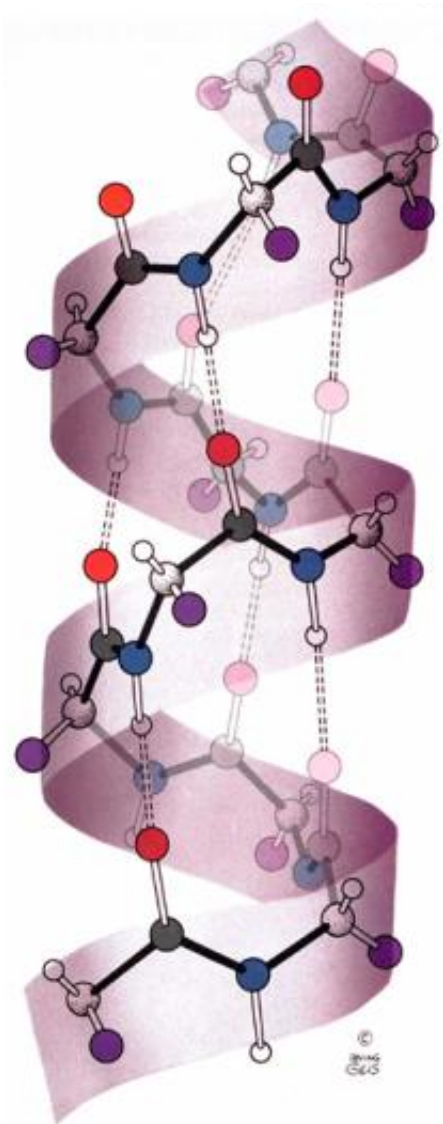
Figure 4-9a Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.



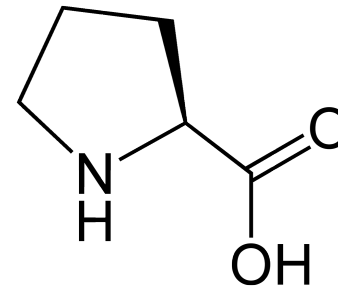
Glycine



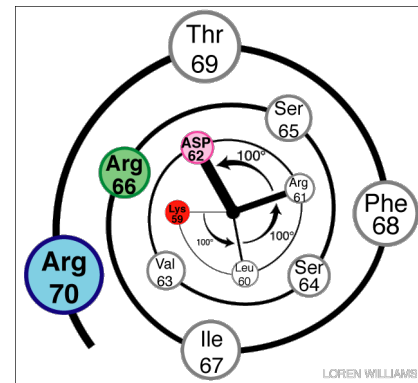
α -helix



- Backbone angles:
 - $\psi = -50^\circ$
 - $\phi = -60^\circ$
- Dipole moment
- Proline bending

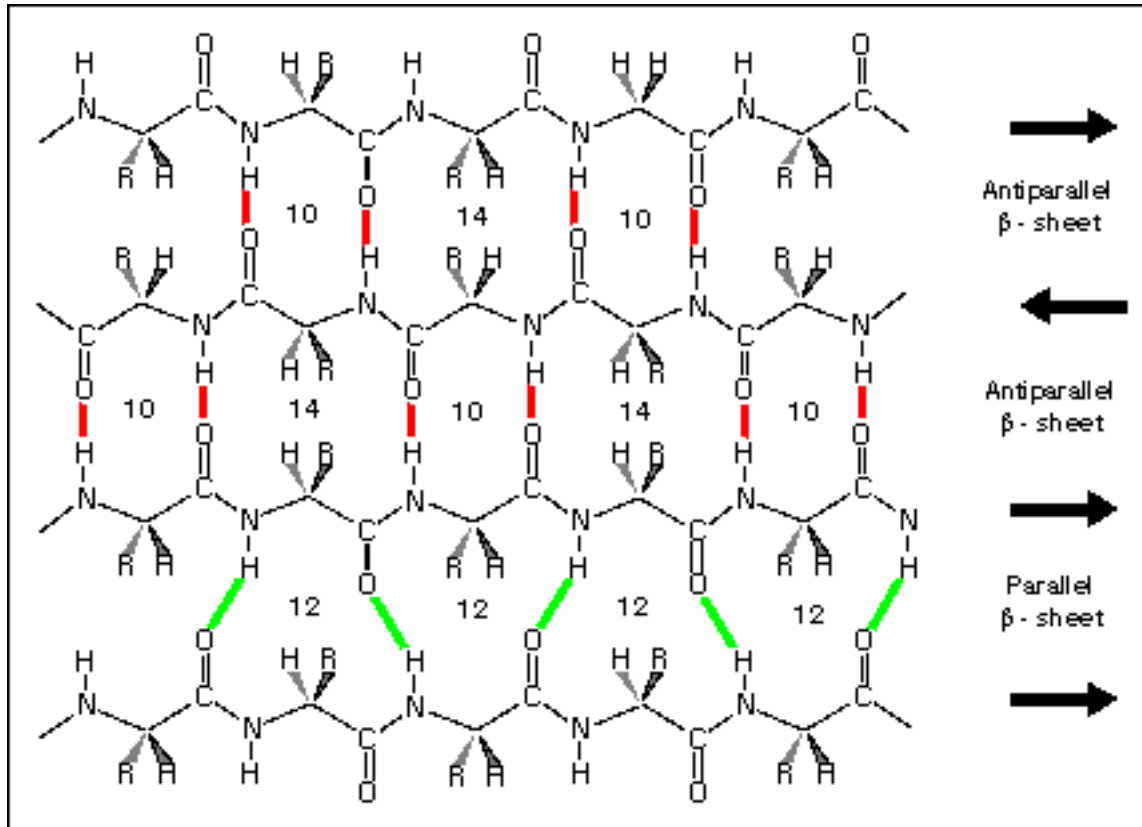


Proline

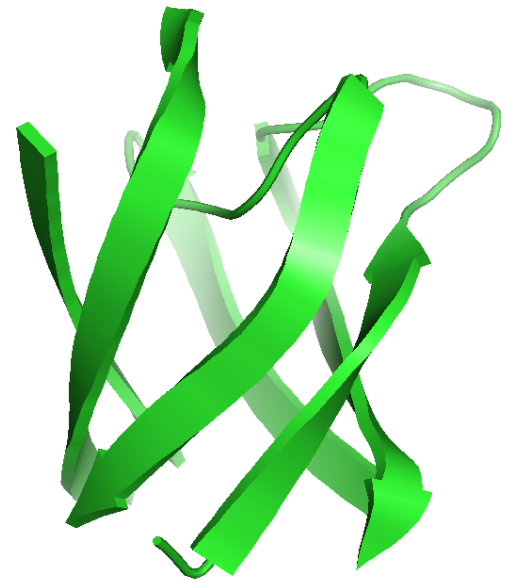


Helical wheel projection

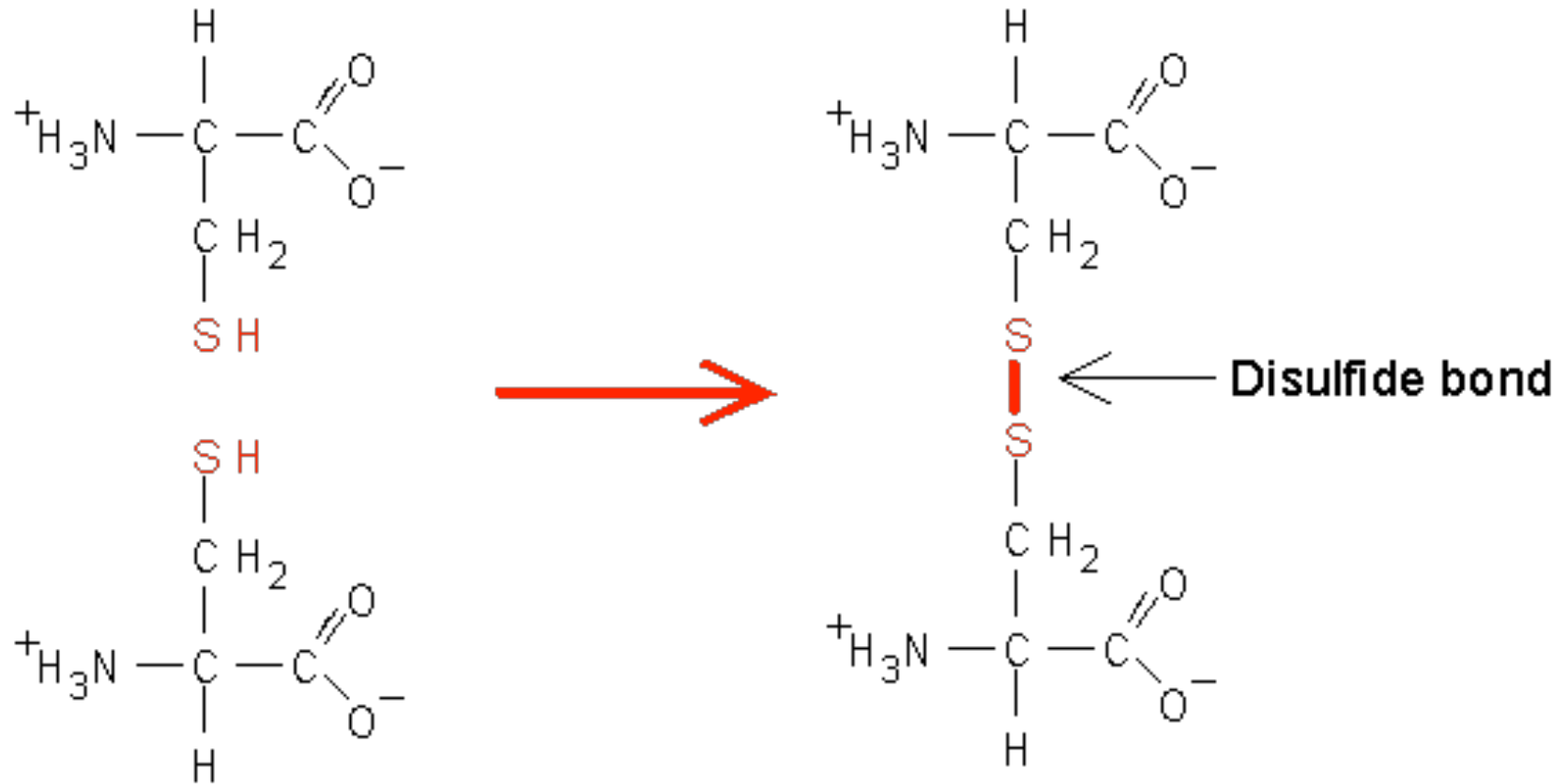
β -sheet



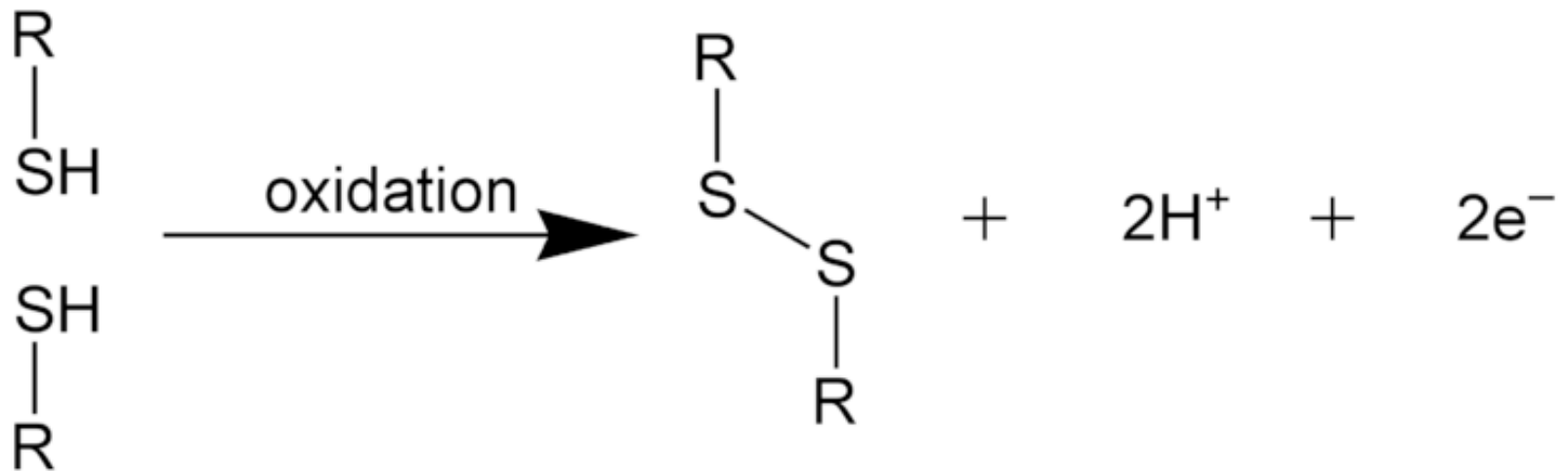
- First quadrant in Ramachandran's plot
- Parallel or antiparallel



Cysteine disulfide bond



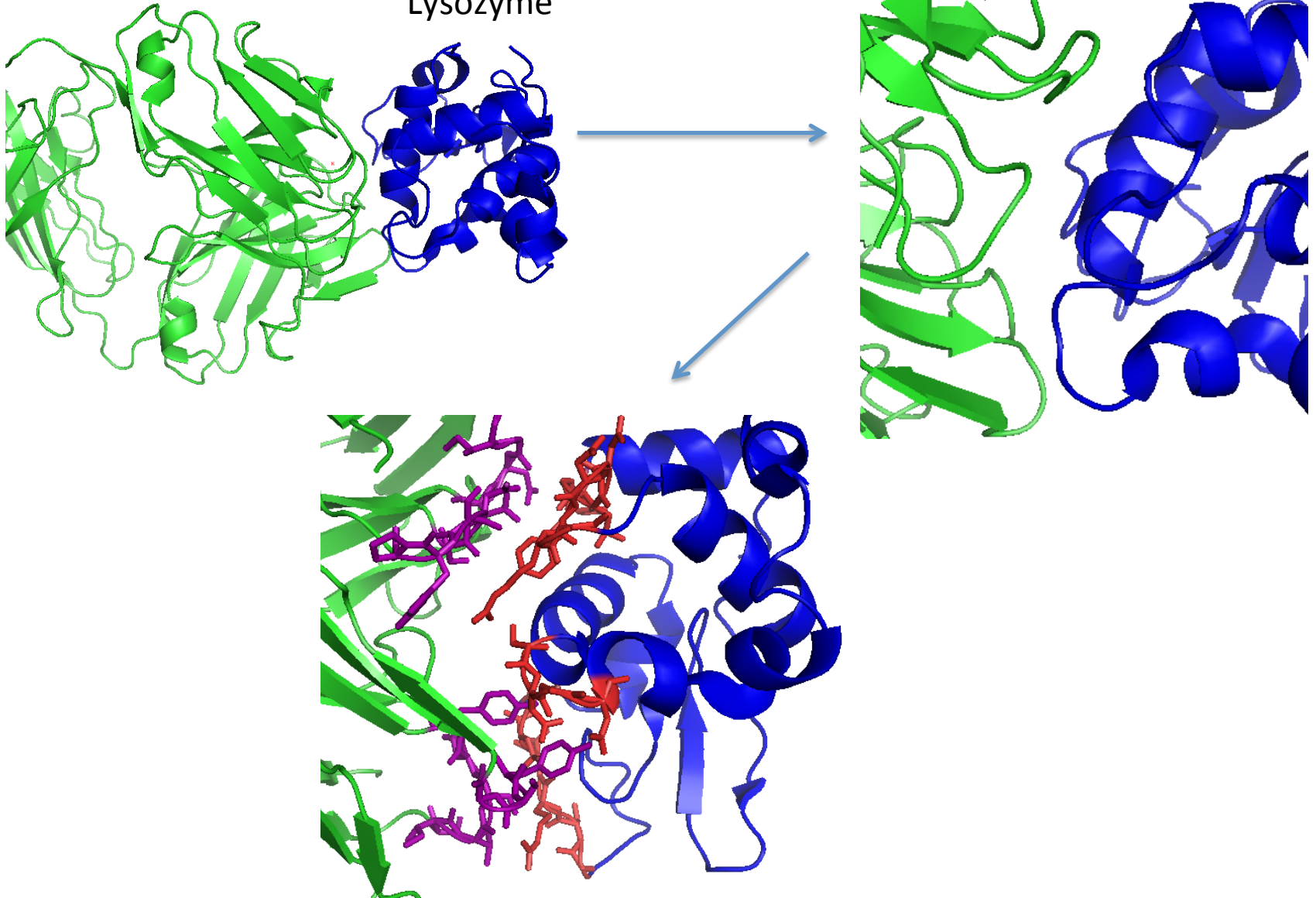
Cysteine disulfide bond



Loops and function

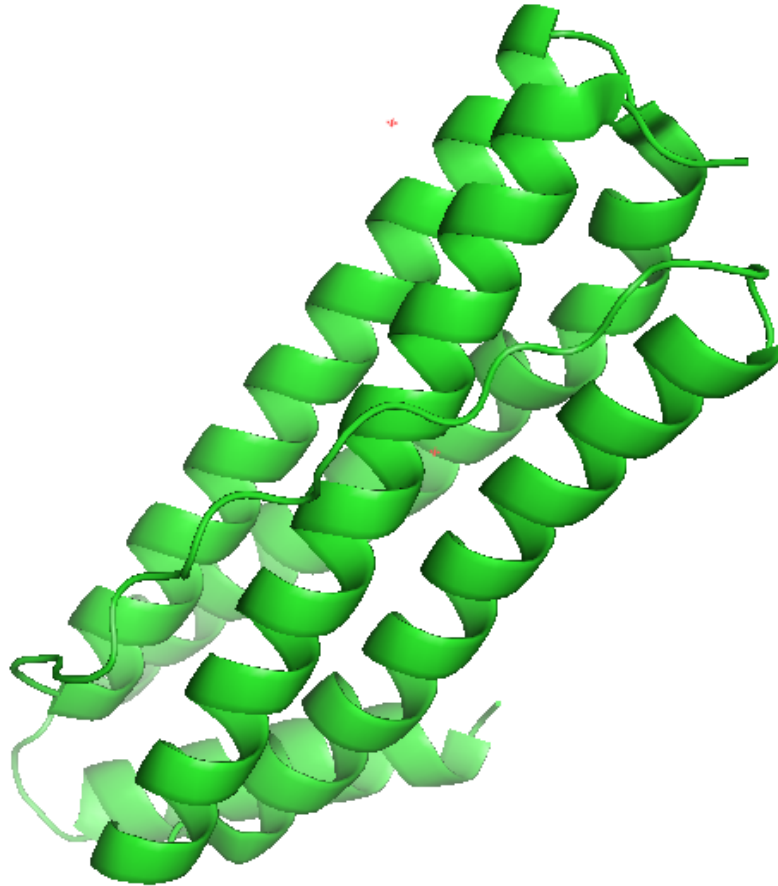
Anti-lysozyme HyHEL-10 Fab

Lysozyme



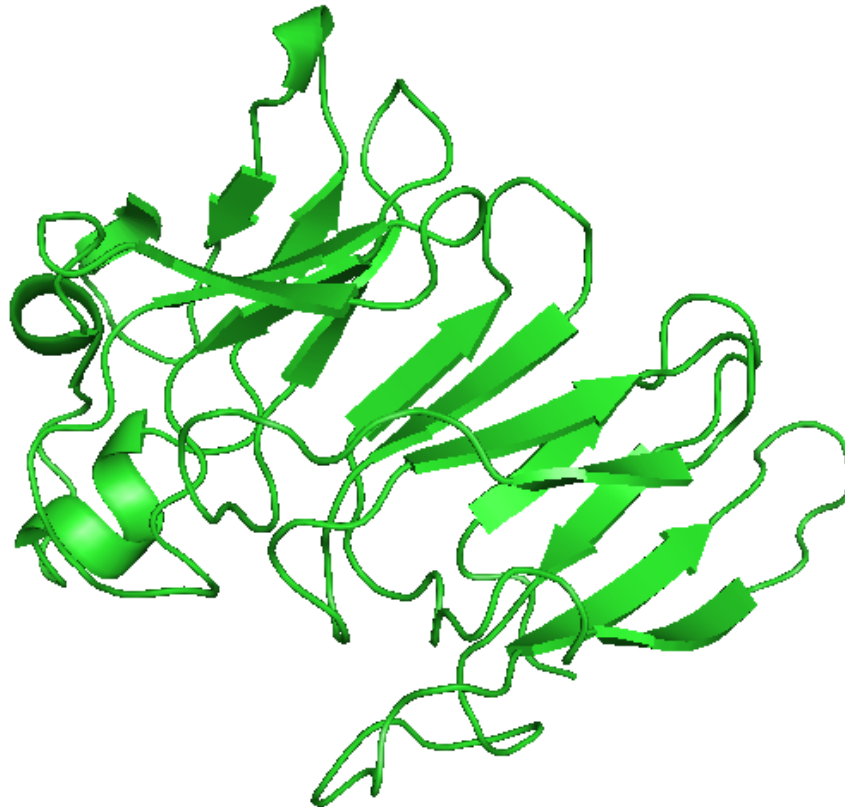
Tertiary structures - Fold space

Mainly alpha:



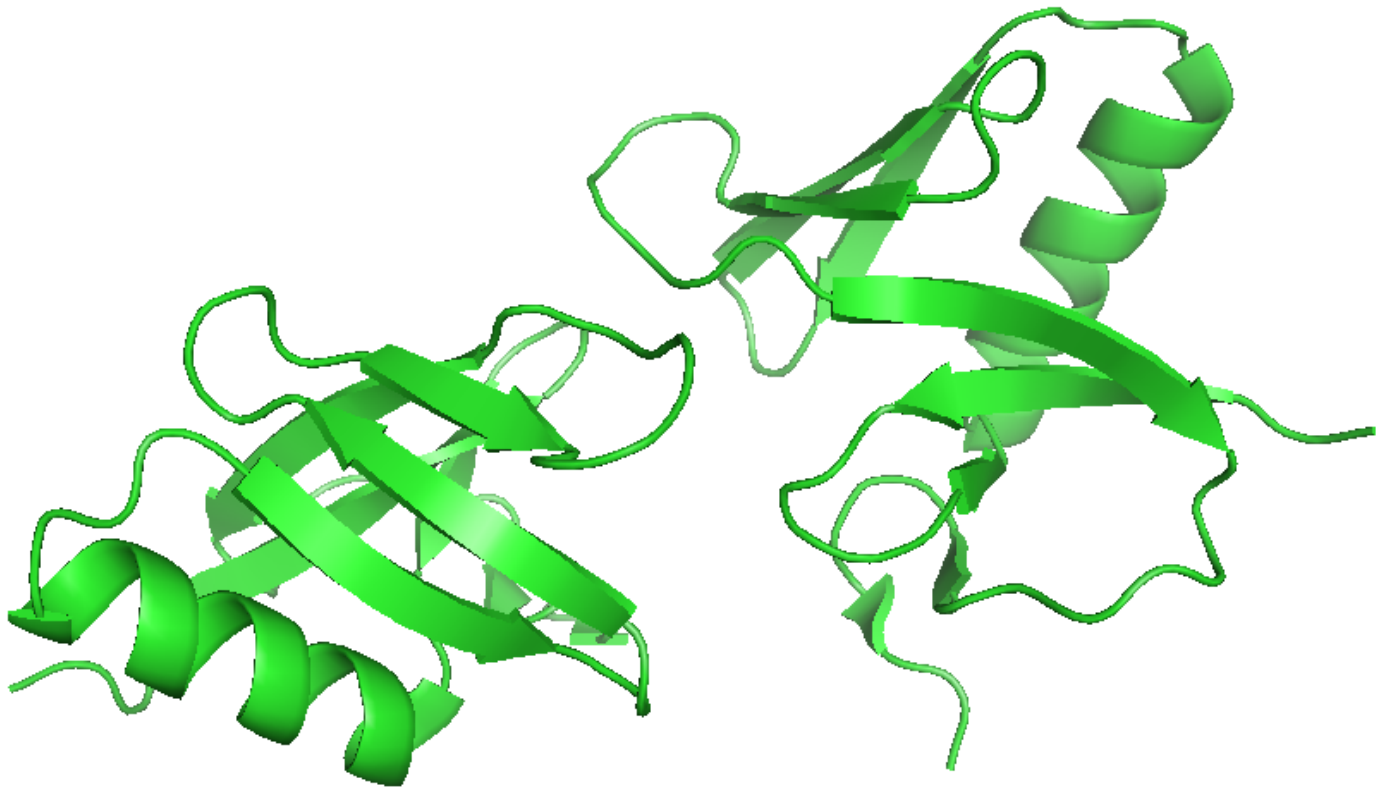
Tertiary structures - Fold space

Mainly beta:



Tertiary structures - Fold space

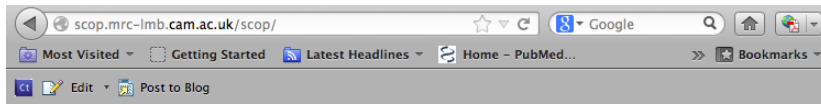
Alpha beta:



SCOP_{1.75} database

<http://scop.mrc-lmb.cam.ac.uk/scop/>

Murzin A. G., et al. (1995). *J. Mol. Biol.* **247**, 536-540.



Structural Classification of Proteins



Welcome to **SCOP: Structural Classification of Proteins**.

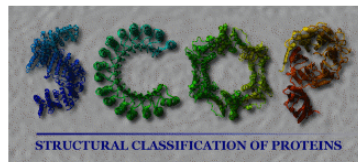
1.75 release (June 2009)

38221 PDB Entries. 1 Literature Reference. 110800 Domains. (excluding nucleic acids and theoretical models).

Folds, superfamilies, and families [statistics here](#).

[New folds](#) [superfamilies](#) [families](#).

[List of obsolete entries and their replacements](#).



Authors. Alexey G. Murzin, John-Marc Chandonia, Antonina Andreeva, Dave Howorth, Loredana Lo Conte, Bartlett G. Ailey, Steven E. Brenner, Tim J. P. Hubbard, and Cyrus Chothia. scop@mrc-lmb.cam.ac.uk

Reference: Murzin A. G., Brenner S. E., Hubbard T., Chothia C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* **247**, 536-540. [\[PDF\]](#)

Recent changes are described in: Lo Conte L., Brenner S. E., Hubbard T.J.P., Chothia C., Murzin A. (2002). SCOP database in 2002: refinements accommodate structural genomics. *Nucl. Acid Res.* **30**(1), 264-267. [\[PDF\]](#), Andreeva A., Howorth D., Brenner S.E., Hubbard T.J.P., Chothia C., Murzin A.G. (2004). SCOP database in 2004: refinements integrate structure and sequence family data. *Nucl. Acid Res.* **32**:D226-D229. [\[PDF\]](#), and Andreeva A., Howorth D., Chandonia J.-M., Brenner S.E., Hubbard T.J.P., Chothia C., Murzin A.G. (2007). Data growth and its impact on the SCOP database: new developments. *Nucl. Acids Res.* **2008** **36**: D419-D425; doi:10.1093/nar/gkm993 [\[PDF\]](#).

Postdoc Wanted

- Want to help us design and build the next generation of SCOP and ASTRAL?
[Get more details and apply here.](#)

Access methods

- Enter SCOP at the [top of the hierarchy](#)
- [Keyword search of SCOP entries](#)
- [SCOP parseable files](#)
- [All SCOP releases and reclassified entry history](#)
- [pre-SCOP - preview of the next release](#)
- SCOP domain sequences and pdb-style coordinate files ([ASTRAL](#))
- Hidden Markov Model library for SCOP superfamilies ([SUPERFAMILY](#))
- Structural alignments for proteins with non-trivial relationships ([SISYPHUS](#))

- [Online resources](#) of potential interest to SCOP users

SCOP [mirrors](#) around the world may speed your access.

- ✓ Largely recognized as “standard of gold”
- ✓ Manually classification
- ✓ Clear classification of structures in:
CLASS
FOLD
SUPER-FAMILY
FAMILY
- ✓ Some large number of tools already available

Manually classification
Not 100% up-to-date
Domain boundaries definition

Class	Number of folds	Number of superfamilies	Number of families
All alpha proteins	284	507	871
All beta proteins	174	354	742
Alpha and beta proteins (a/b)	147	244	803
Alpha and beta proteins (a+b)	376	552	1055
Multi-domain proteins	66	66	89
Membrane and cell surface proteins	58	110	123
Small proteins	90	129	219
Total	1195	1962	3902

SCOP2 database

<http://scop2.mrc-lmb.cam.ac.uk/>



Structural Classification of Proteins 2

[About](#) | [Browser](#) | [Graph](#) | [Download](#) | [Support](#)

MRC

Laboratory of
Molecular Biology

News

November, 2013

During the development of SCOP2, we have identified a new, previously unrecognised type of alpha-alpha superhelix. Unlike other alpha-alpha superhelices..
[More...](#)

January, 2014

SCOP2 article in NAR is published
[More...](#)

January, 2014

The structure of the month
[More...](#)

Welcome to SCOP2!

Citation

Antonina Andreeva, Dave Howorth, Cyrus Chothia, Eugene Kulesha, Alexey Murzin, SCOP2 prototype: a new approach to protein structure mining (2014) Nucl. Acid Res., 42 (D1): D310-D314. [\[PDF\]](#)

Description of the SCOP2 database

SCOP2 is a successor of Structural classification of proteins (SCOP). Similarly to SCOP, the main focus of SCOP2 is on proteins that are structurally characterized and deposited in the PDB. Proteins are organized according to their structural and evolutionary relationships, but, in contrast to SCOP, instead of a simple tree-like hierarchy these relationships form a complex network of nodes. Each node represents a relationship of a particular type and is exemplified by a region of protein structure and sequence.

In SCOP2, we try to put in use the knowledge we acquired over the past years and the lessons we have learned during the classification of protein structures. We believe that there are many

Search Browser

Add an asterisk to search free text (e.g. serine*)

Search Graph

Add an asterisk to search free text (e.g. protein*domain)

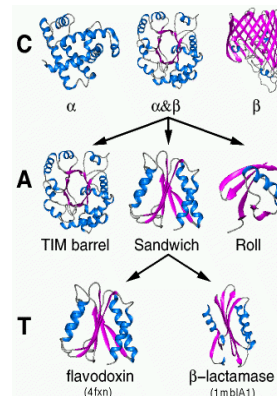
CATH_{3.5} database

<http://www.cathdb.info>

Uses FSSP for superimposition

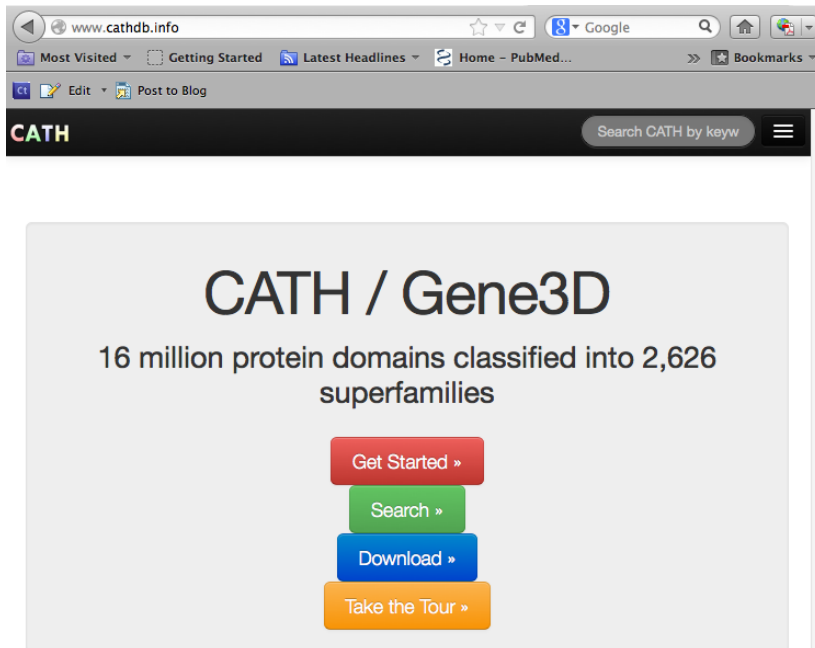
- ✓ Recognized as “standard of gold”
- ✓ Semi-automatic classification
- ✓ Clear classification of structures in:
CLASS
ARCHITECTURE
TOPOLOGY
HOMOLOGOUS SUPERFAMILIES
- ✓ Some large number of tools already available
- ✓ Easy to navigate

Semi-automatic classification
Domain boundaries definition



173,536	CATH Domains
2,626	CATH Superfamilies
51,334	PDBs

Orengo, C.A., et al. (1997) *Structure*. **5**. 1093-1108.



What's New?

The CATH website has recently undergone a big overhaul. We really hope you find the new pages more useful, easier to use and quicker to load. Please [get in touch](#) and let us know what you think.

Searching CATH

- Search by ID / keyword
- Search by FASTA sequence
- Search by PDB structure

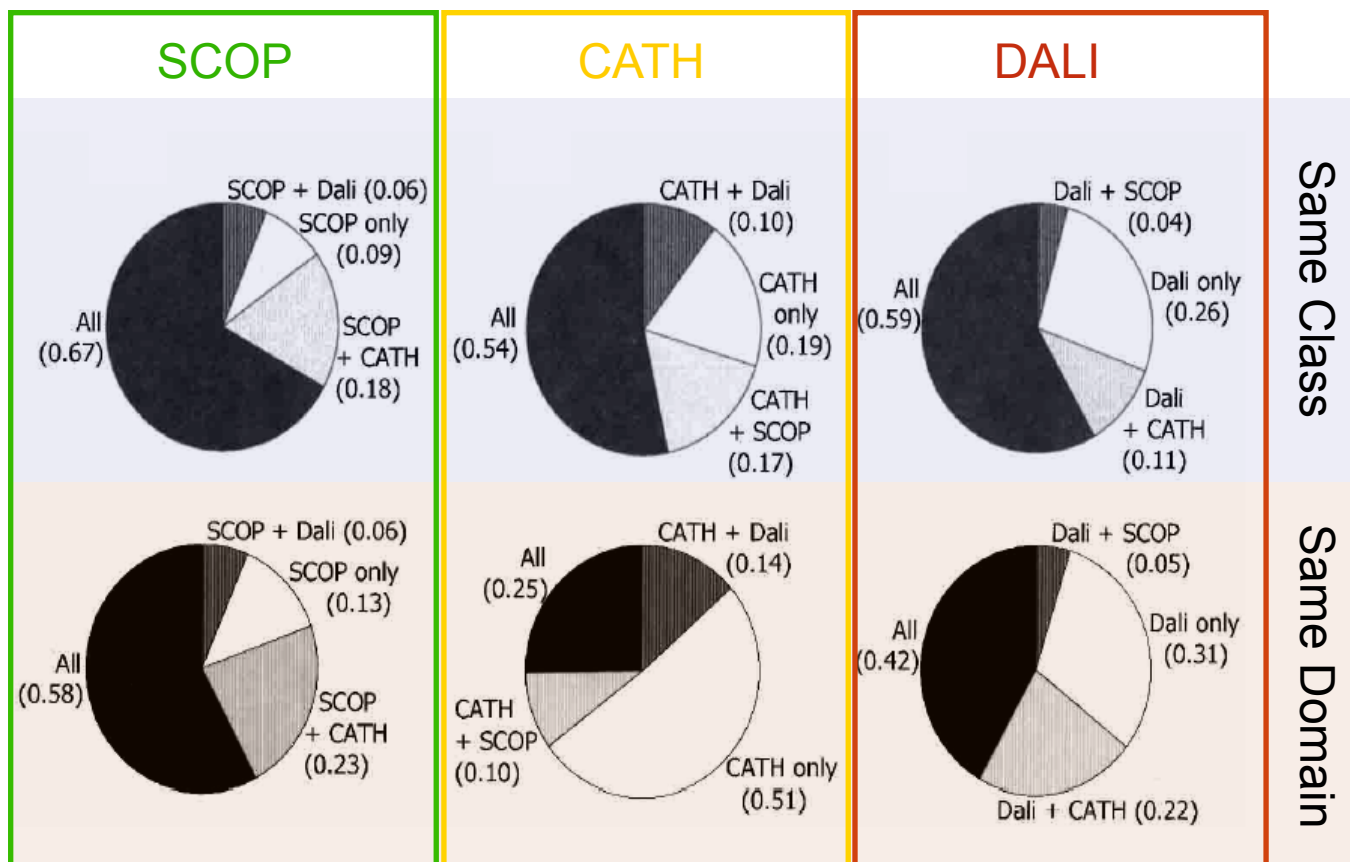
Example pages

- PDB "1dan"
- Domain "1cukA01"
- Relatives of "1cukA01"
- Superfamily "HUPs"
- Functional Family

Classification of the structural space

Not an easy task!

Domain definition AND domain classification



PDB

RCSB Protein Data Bank - RCSB PDB

www.pdb.org/pdb/home/home.do

An Information Portal to Biological Macromolecular Structures

As of Tuesday Feb 19, 2013 at 4 PM PST there are 68325 Structures | PDB Statistics

Search | All Categories: e.g., PDB ID, molecule name, author

Customize This Page

Available on the App Store

PDB-101

Structural View of Biology

Molecule of the Month

Proton-Gated Urea Channel

The acid in your stomach helps to digest food, but it also helps protect you from bacterial infection. However, one type of bacteria, *Helicobacter pylori*, is able to live in the acidic environment of the stomach. It is one of the most common bacterial infections, found worldwide in half of the population. It causes a continued inflammation of the stomach, which leads in some cases to stomach ulcers and stomach cancer.

Full Article

Protein Structure Initiative Featured System

Designer Proteins

The engineering of new proteins with novel structures and functions is one of the grand challenges facing the scientific community. This goal is particularly tempting, because we can look to nature to see thousands of working examples of proteins that spontaneously fold and perform diverse functions. By looking at natural proteins, scientists have discovered many of the features that are required to create a functional protein, and now, researchers at PSI have proven that these rules may be used for design.

Full Article | Archive | PSI Structural Biology Knowledgebase

Explore Archive

Organism

Exp. Method

Release Date

Enzyme Classification

Taxonomy

X-ray Resolution

Polymer Type

SCOP Classification

Organism

- Homo sapiens (21905)
- Escherichia coli (4695)
- Mus musculus (3772)
- Saccharomyces cerevisiae (2219)
- Bos taurus (2128)
- Rattus norvegicus (1850)
- Escherichia coli K-12 (1518)
- Other (47689)

Latest Structures

4G4R : MutM containing F114A mutation bound to oxoG-containing DNA

Structural and biochemical analysis of DNA helix-invasion by the bacterial 8-oxoguanine DNA glycosylase MutM.

View in 3D (3mol)

check out all of the latest structures released.

New Structures

Latest Release

New Structure Papers

Search Unreleased Entries

New Features

Macromolecule Names: Synonyms

Latest features released:

Website Release Archive:

RCSB PDB News

Weekly | Quarterly | Yearly

2013-02-19

Build complex queries with Advanced Search

Advanced Search

Combine different searches to find structures and refine search results. New options include quick searches by experimental and/or molecule type, structure determination method, and intermediate connectivity (LINK records). more

Poster Prize Awarded at AsC

- Author Profiles: Timeline
- Display of a Researcher's Structures
- News Reader Survey
- Visit the Biomolecular Discovery Dome at Biophysical Society
- Join the RCSB PDB Team

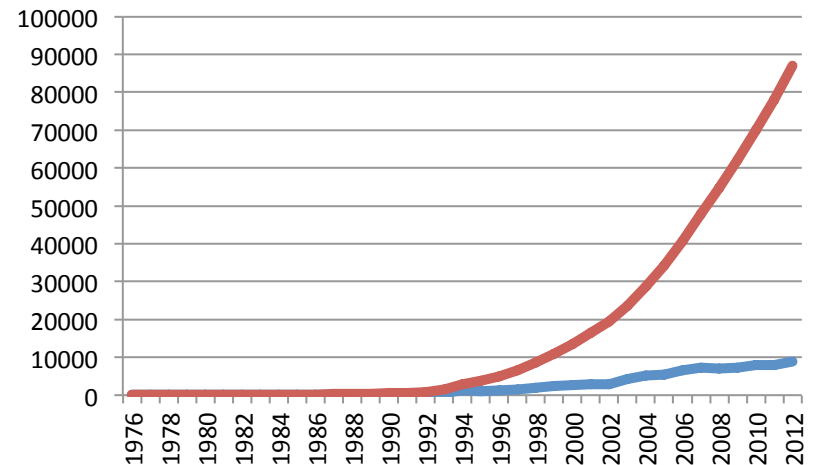
wwPDB News

X-ray Validation Task Force Report Published more

2012-12-21

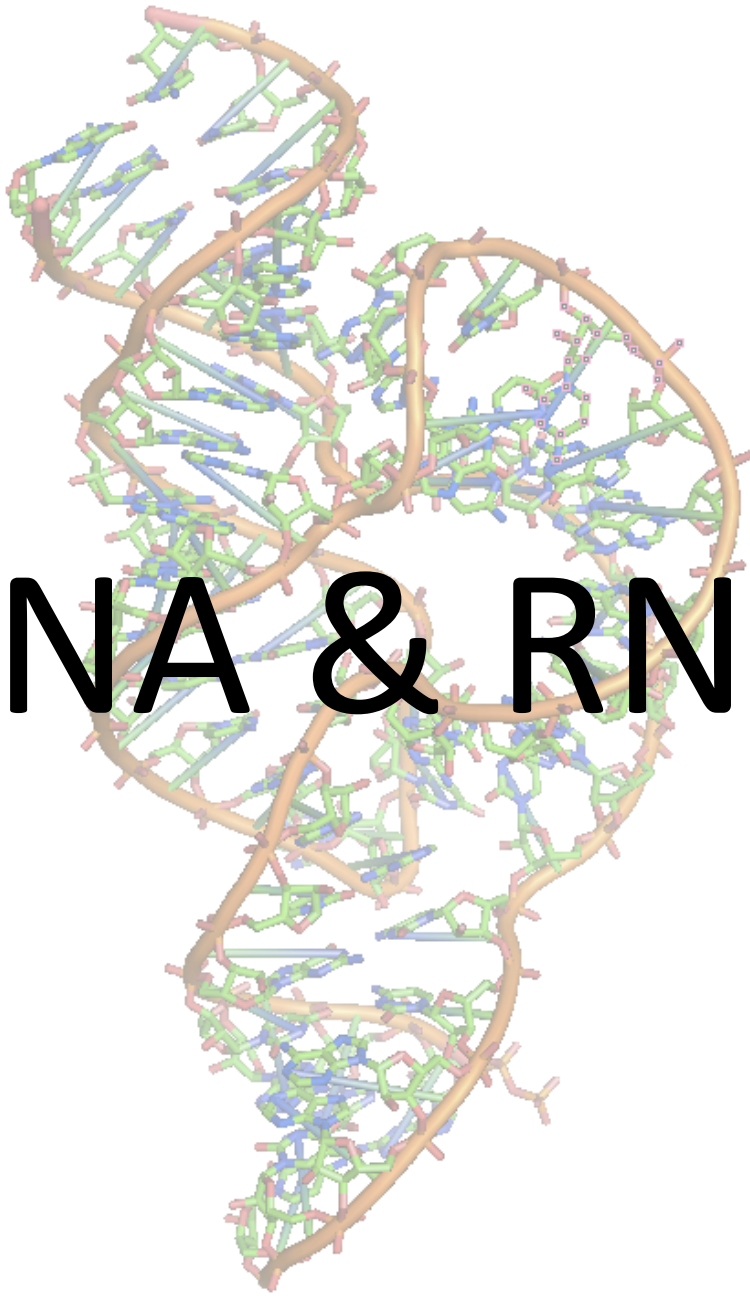
Announcement: Remediation of Structure

Yearly and total PDB structures per year



Yearly

Total

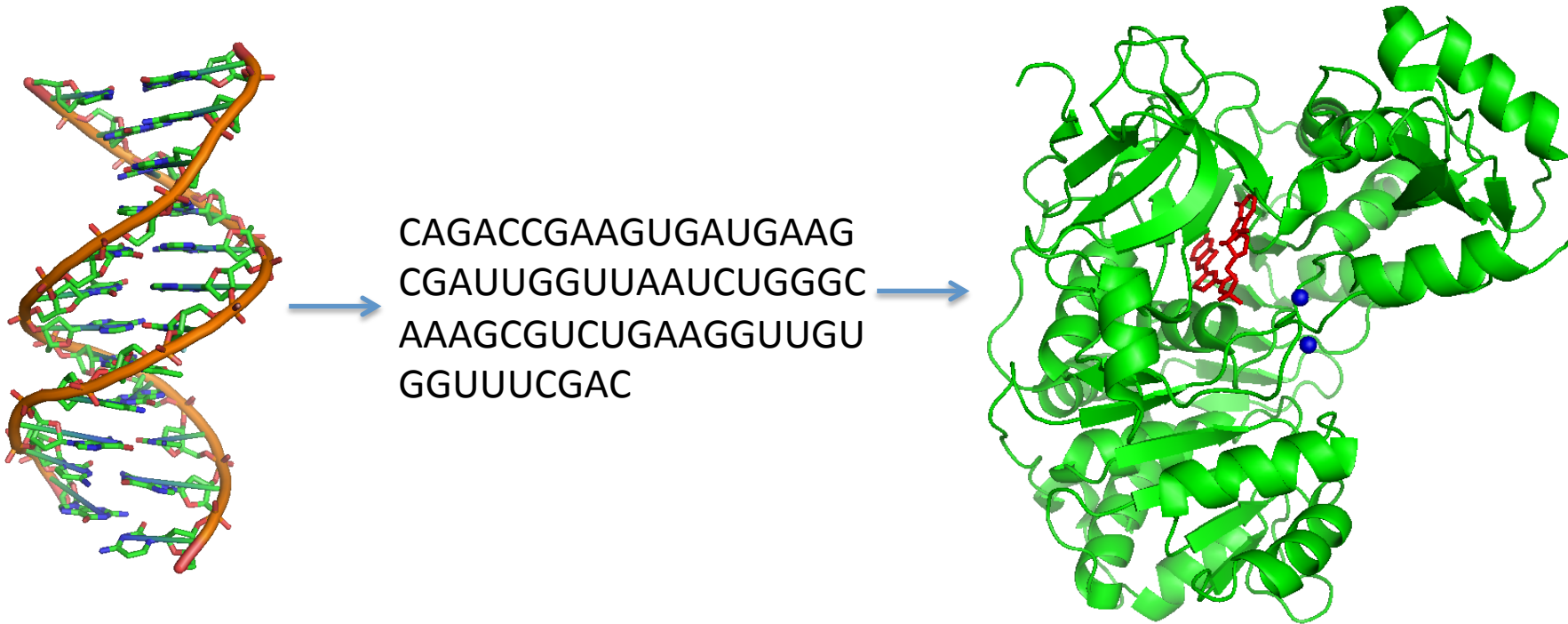


DNA & RNA

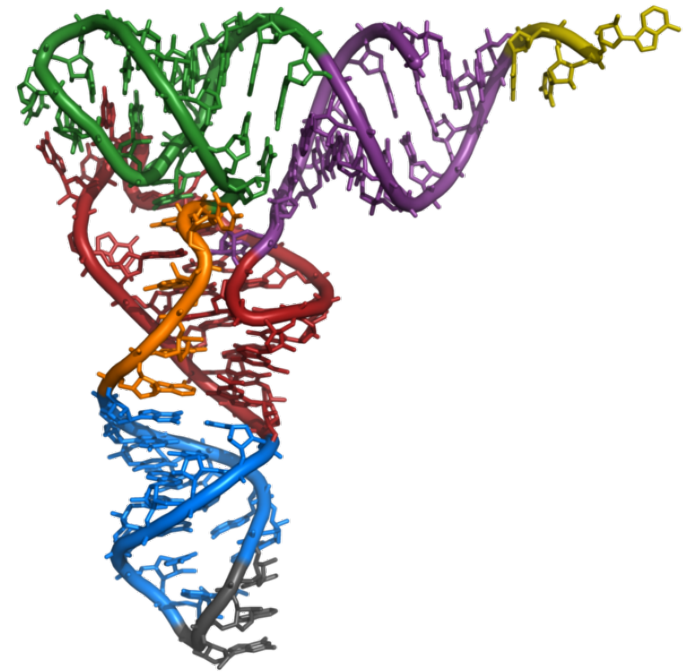
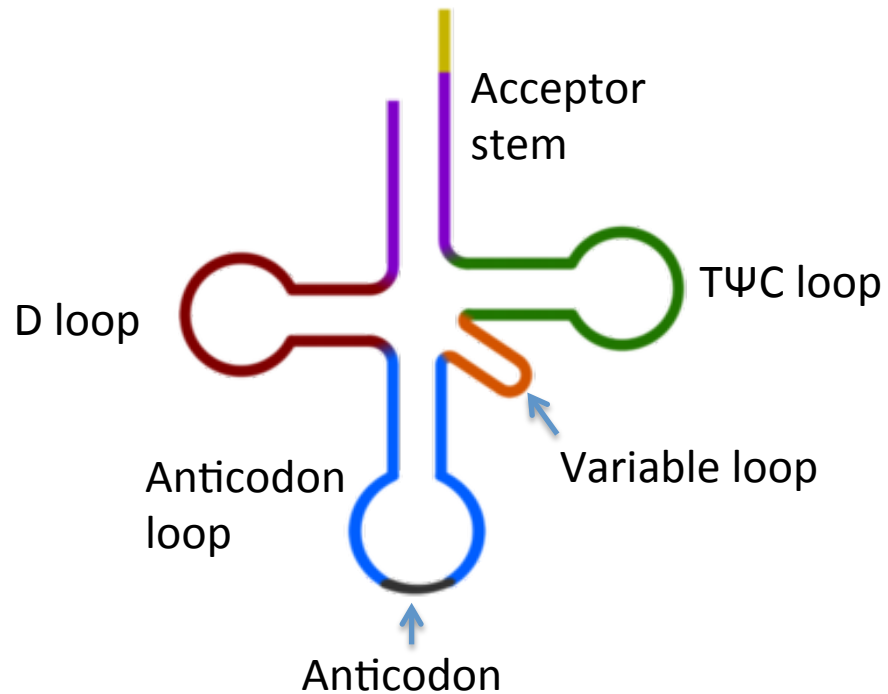
RNA function

- 1) Informative
- 2) Structural
- 3) Catalytic
- 4) Regulatory

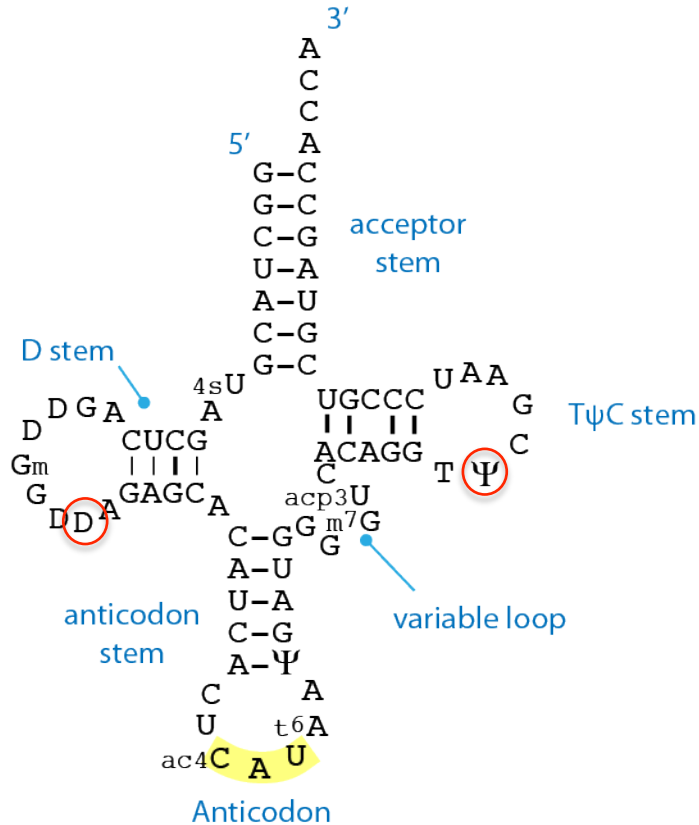
- 1) Messenger (mRNA)



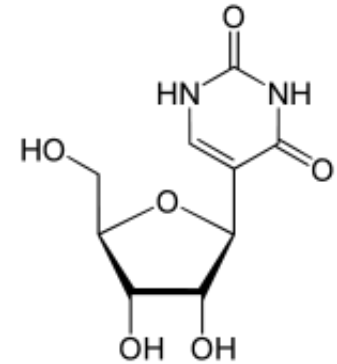
- 2) Transfer (tRNA)



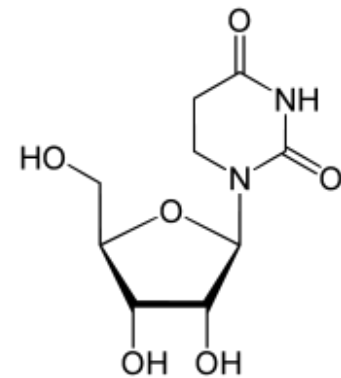
- tRNA modifications



ψ = Pseudouridine

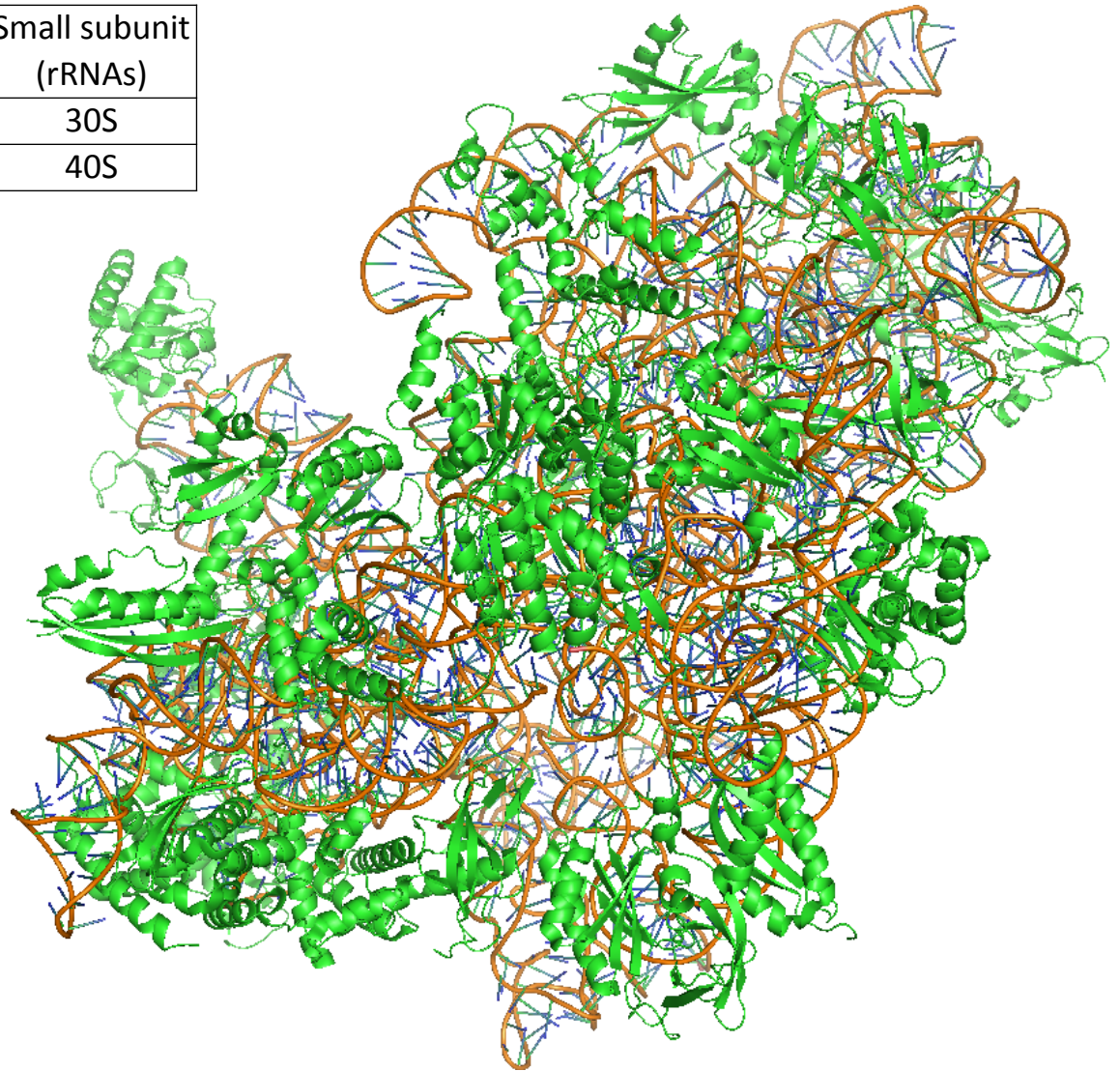


D = Dihydrouridine



- 3) Ribosomal (structural) (rRNA)

Type	Size	Large subunit (rRNAs)	Small subunit (rRNAs)
prokaryotic	70S	50S	30S
eukaryotic	80S	60S	40S



- 4) Genomic

- RNA virus

- Single stranded RNA virus (ssRNA)

- Positive strand RNA virus

- Negative strand RNA virus

- Double stranded RNA virus (dsRNA)

- RNA satellites

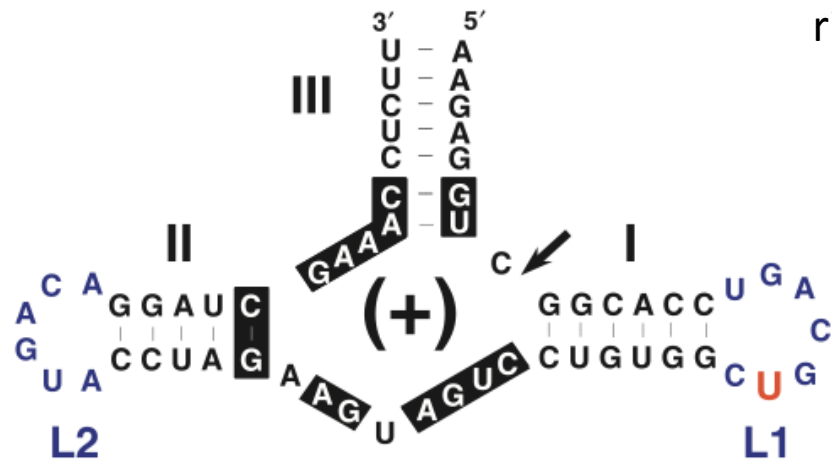
- SatRNAs of *Cucumber Mosaic Virus*

- SatRNAs of *Turnip Crinkle Virus*

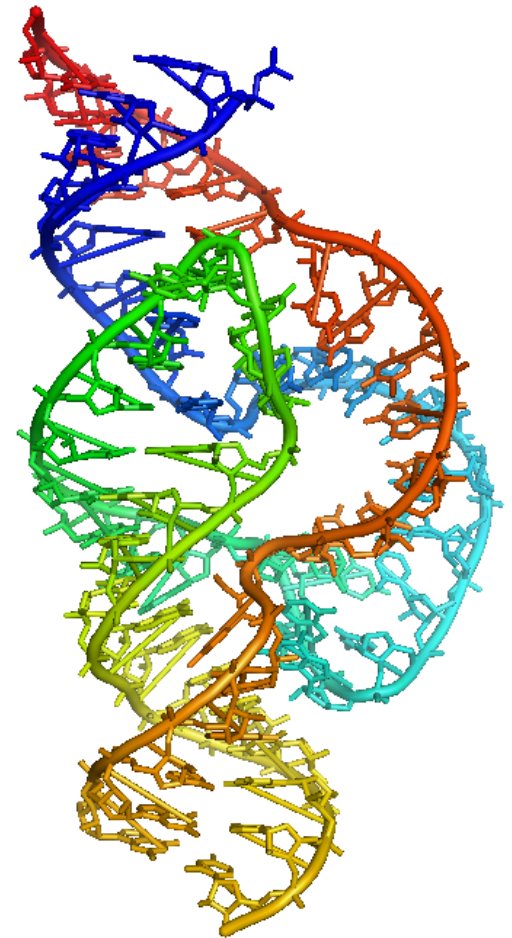
- ...

- Retroviruses

- 5) Catalytical (ribozymes)



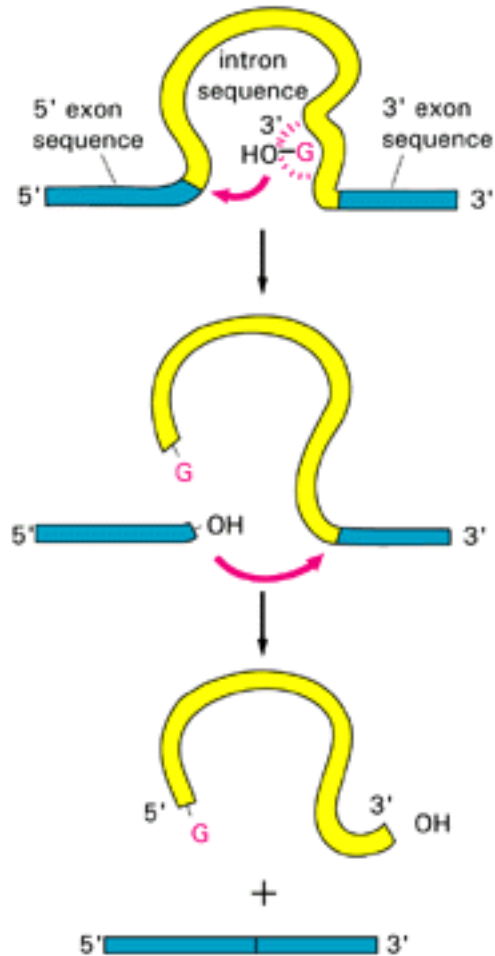
Hammerhead
ribozyme



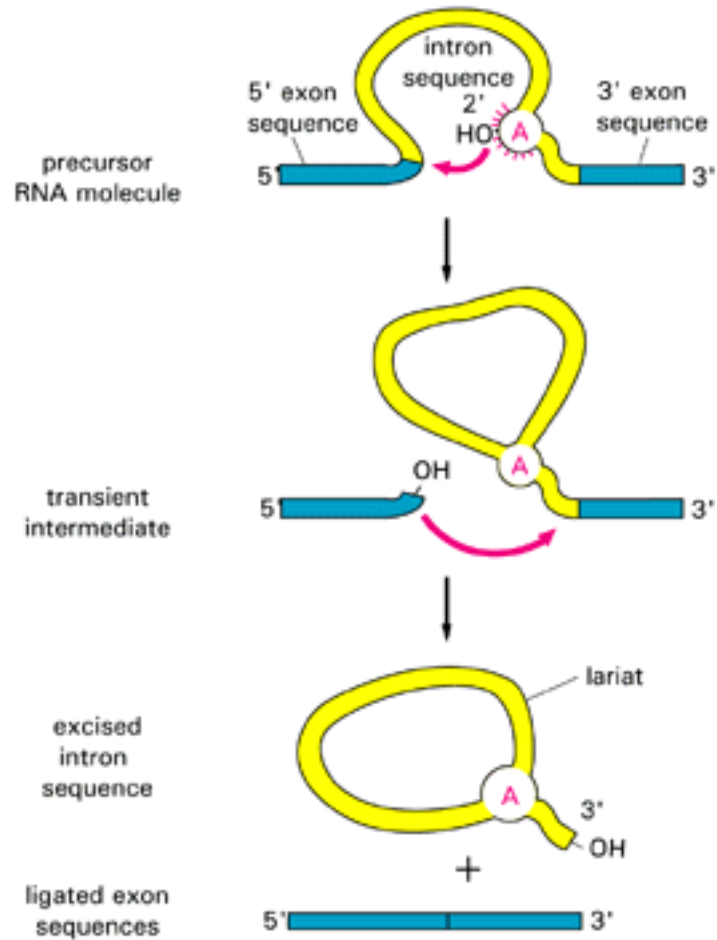
- 5) Catalytical (ribozymes)
- Examples:
 - Peptidyl transferase 23S rRNA
 - Group I and II introns
 - RNase P
 - Hammerhead ribozyme
 - Hairpin ribozyme
 - ...

Introns

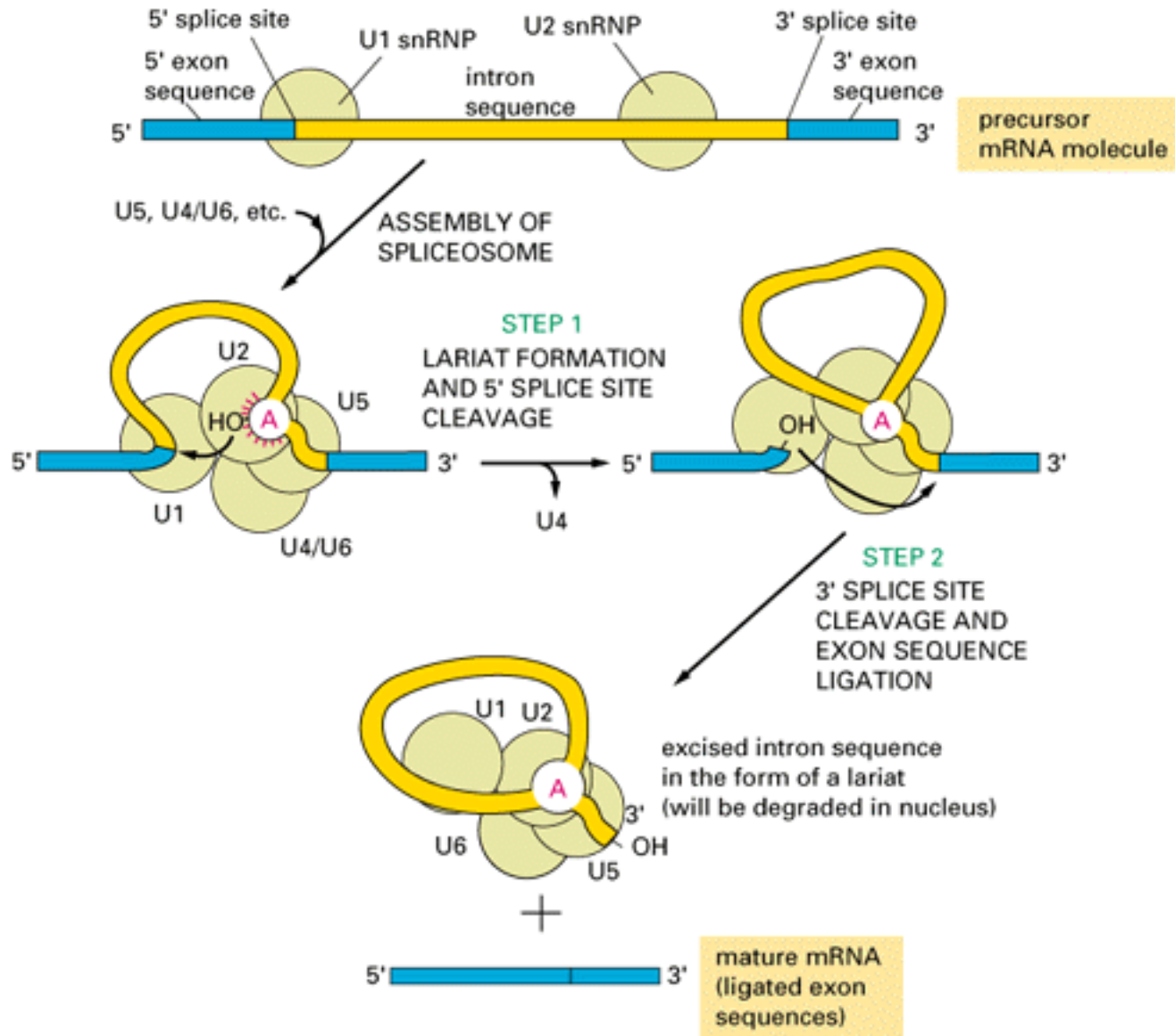
Group I self-splicing intron sequences



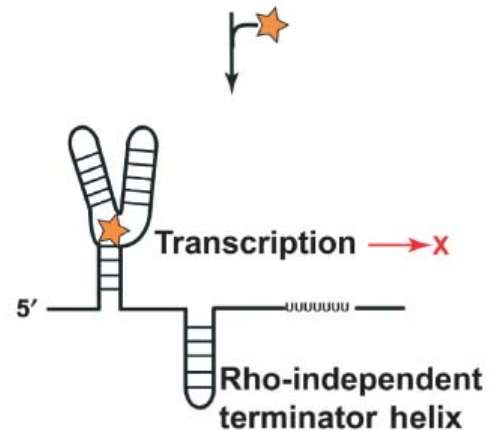
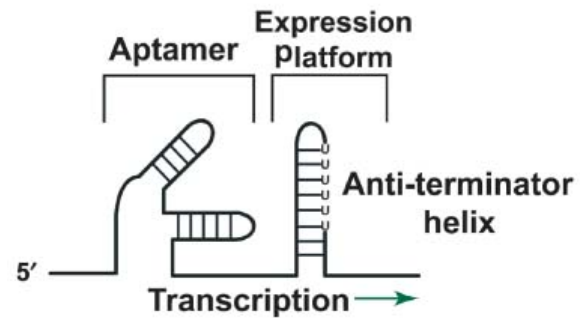
Group II self-splicing intron sequences



Type III introns



- 6) Molecular sensor (riboswitches)

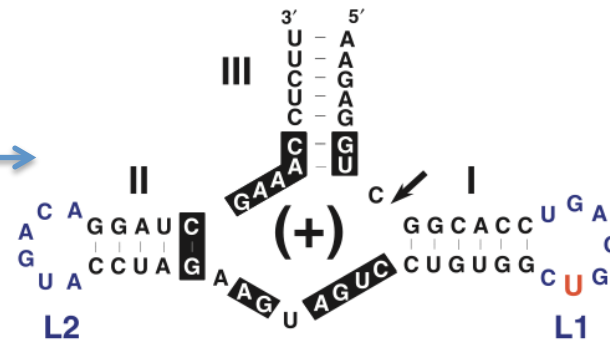


- Small RNAs (sRNA)
 - Small nuclear RNAs (snRNA)
 - Small nucleolar RNAs (snoRNA)
 - Micro RNAs (miRNA)
 - Silencing RNAs (siRNA)
 - Piwi-interacting RNAs (piRNA)
- Long non-coding RNAs (lncRNA)
- others...
- To be discovered...
- RNA therapeutics

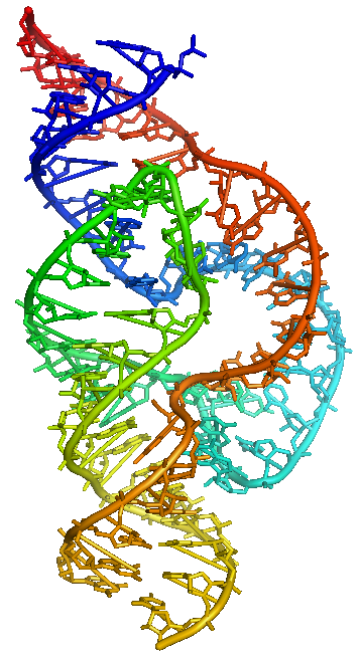
Primary structure

CAGACCGAAGUGAUGAAGC
GAUUGGUUAAUCUGGGCA
AAGCGUCUGAAGGUUGUG
GUUUCGAC

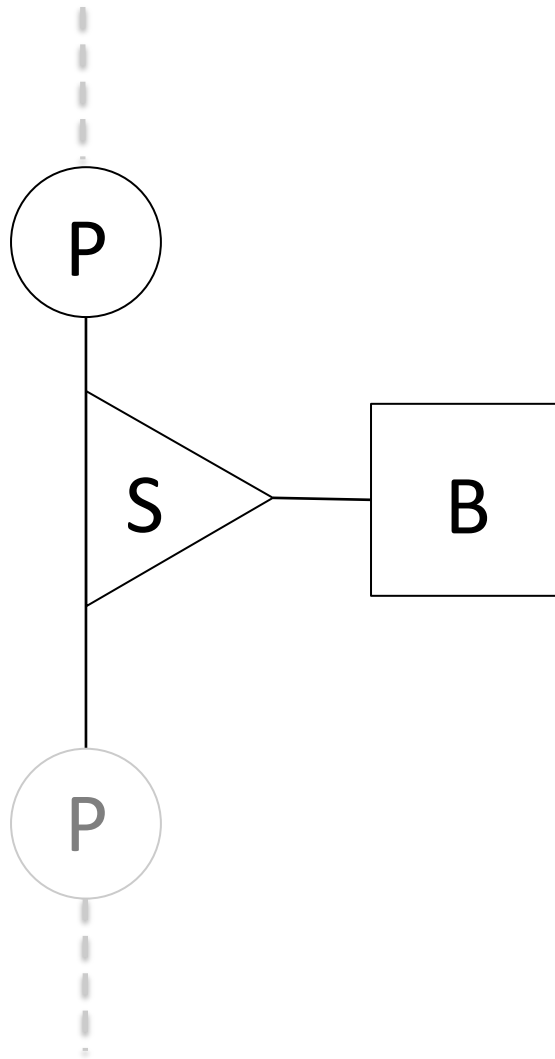
Secondary structure



Tertiary structure (3D)



Nucleotide structure



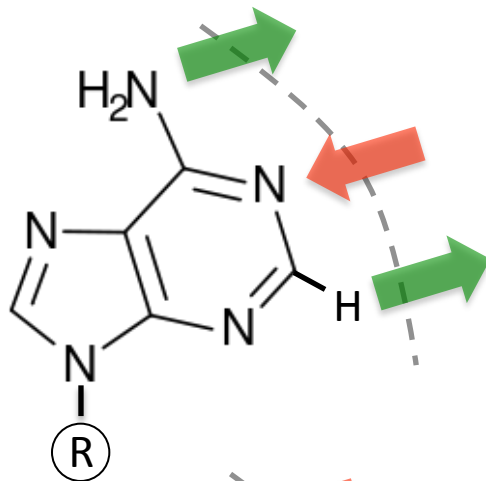
P = Phosphate
S = Sugar (ribose,
deoxyribose)
B = nucleotidic
Base

Bases

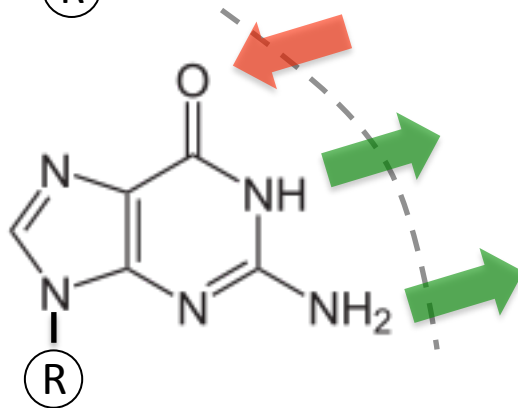
(aka nucleotidic bases, aromatic bases)

Purines

Adenine
(A)

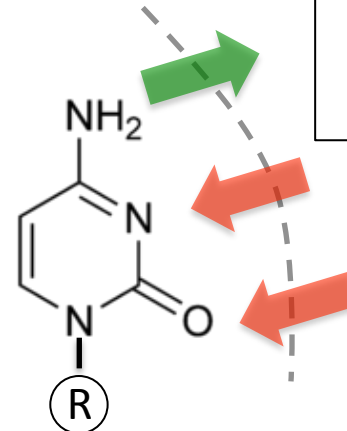


Guanine
(G)

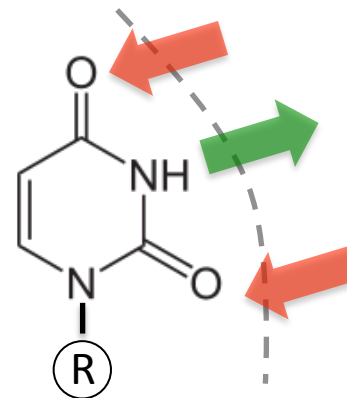


Pyrimidines

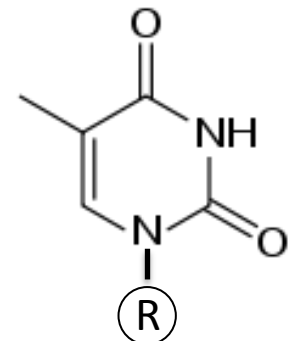
Cytosine
(C)



Uracil
(U)



Thymidine
(T)

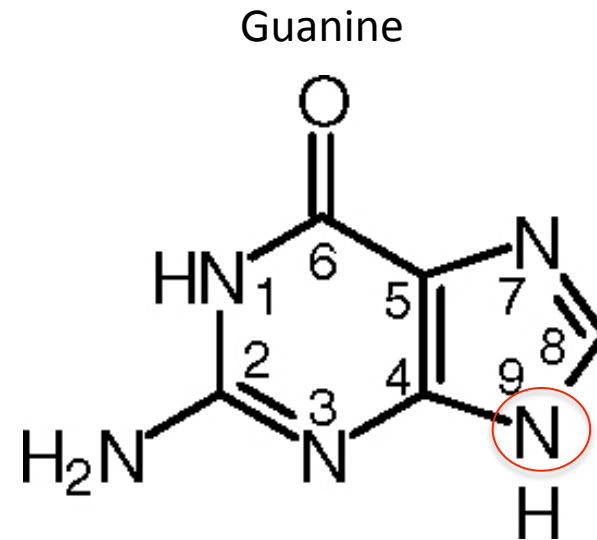
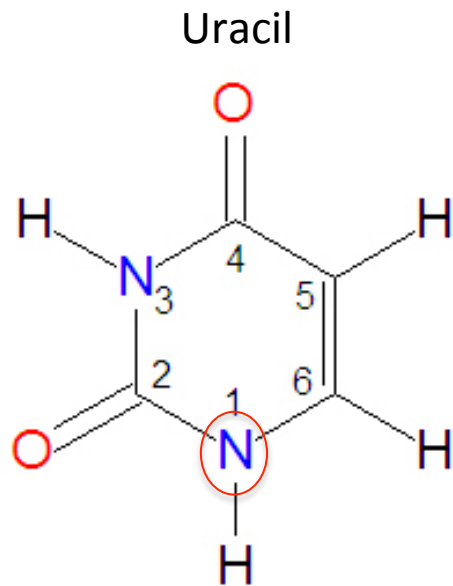


H bond

Donor

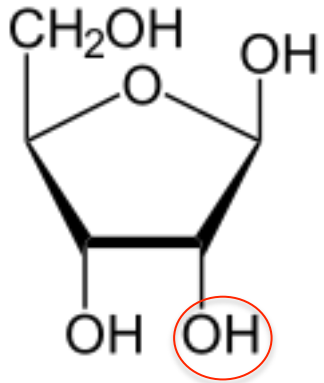
Acceptor

Numbering system - bases

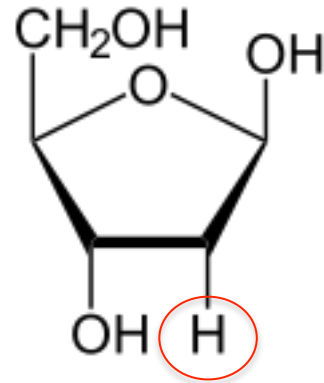


Sugars

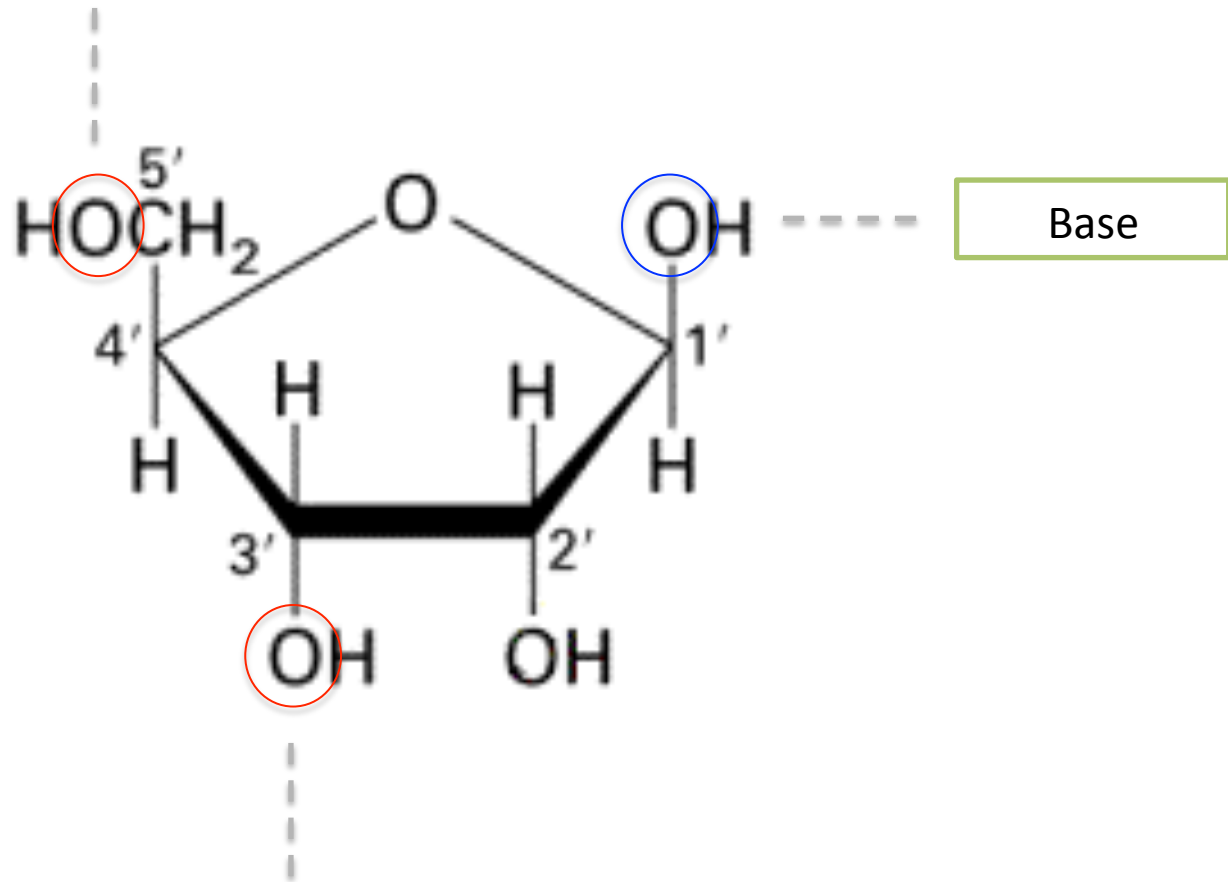
β -D-ribose



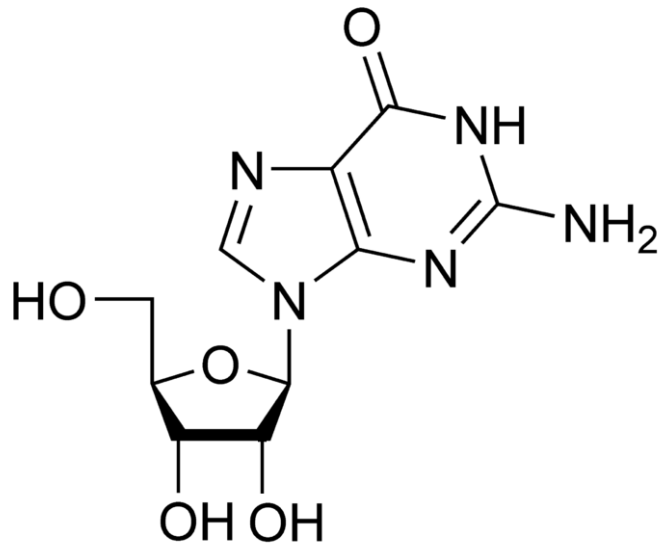
β -D-2-Deoxyribose



Numbering system -ribose

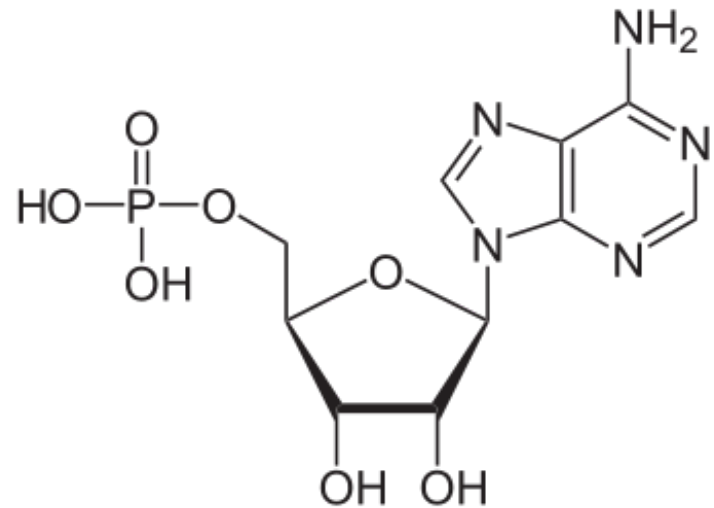


Nucleoside



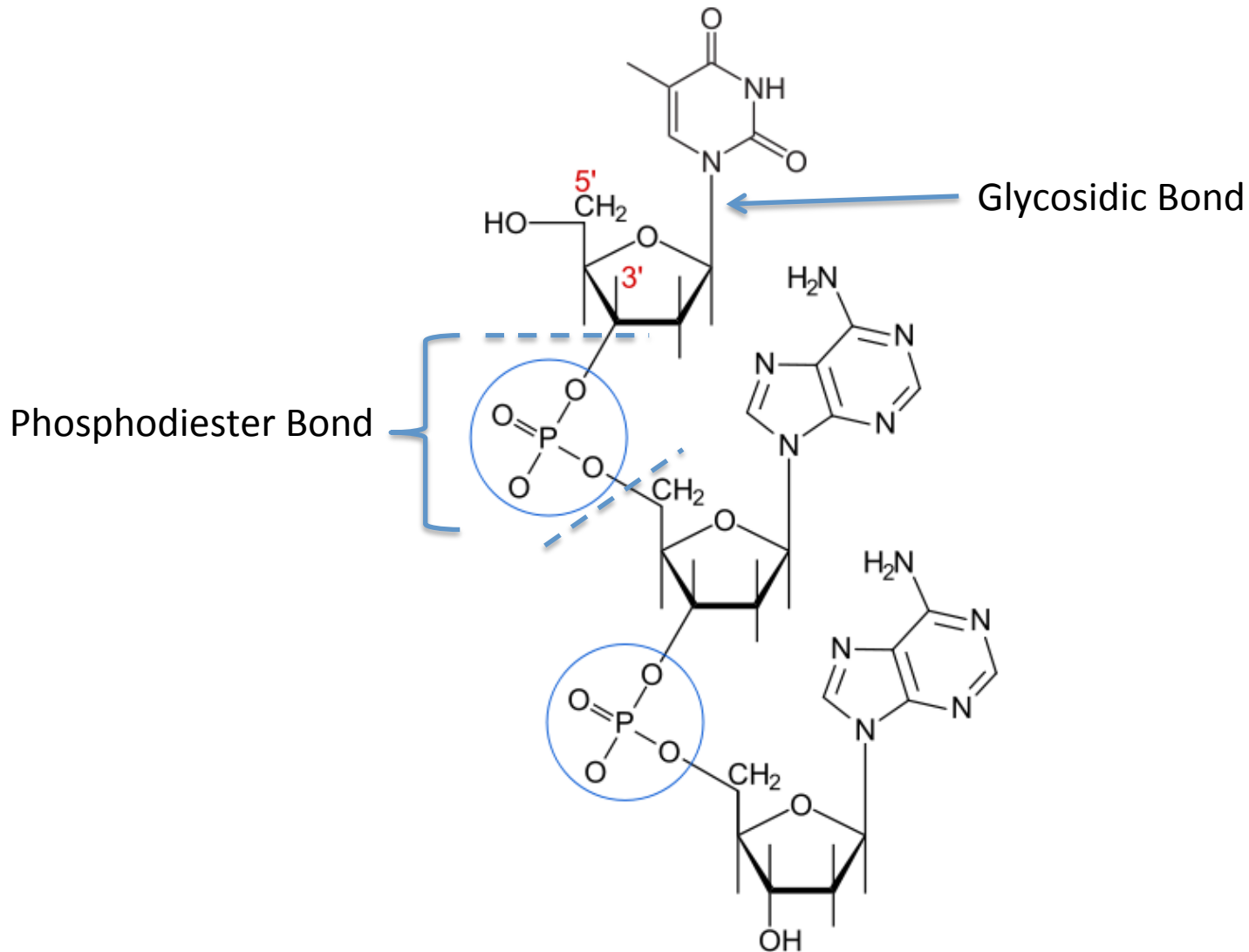
Sugar + base

Nucleotide monophosphate

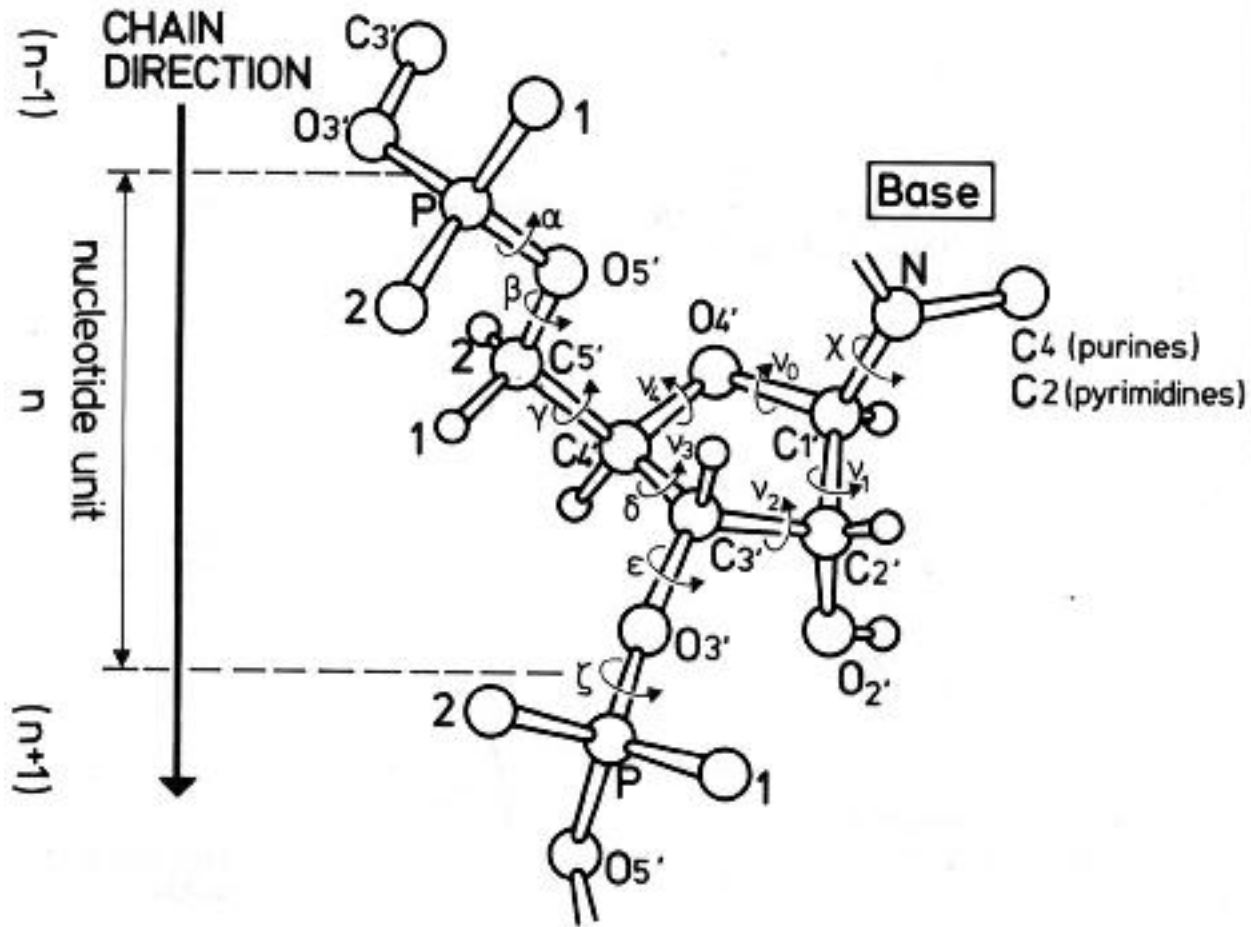


Phosphate + sugar + base

Phosphodiester Bond



Nucleic acid torsion angles



Ramachandran plot

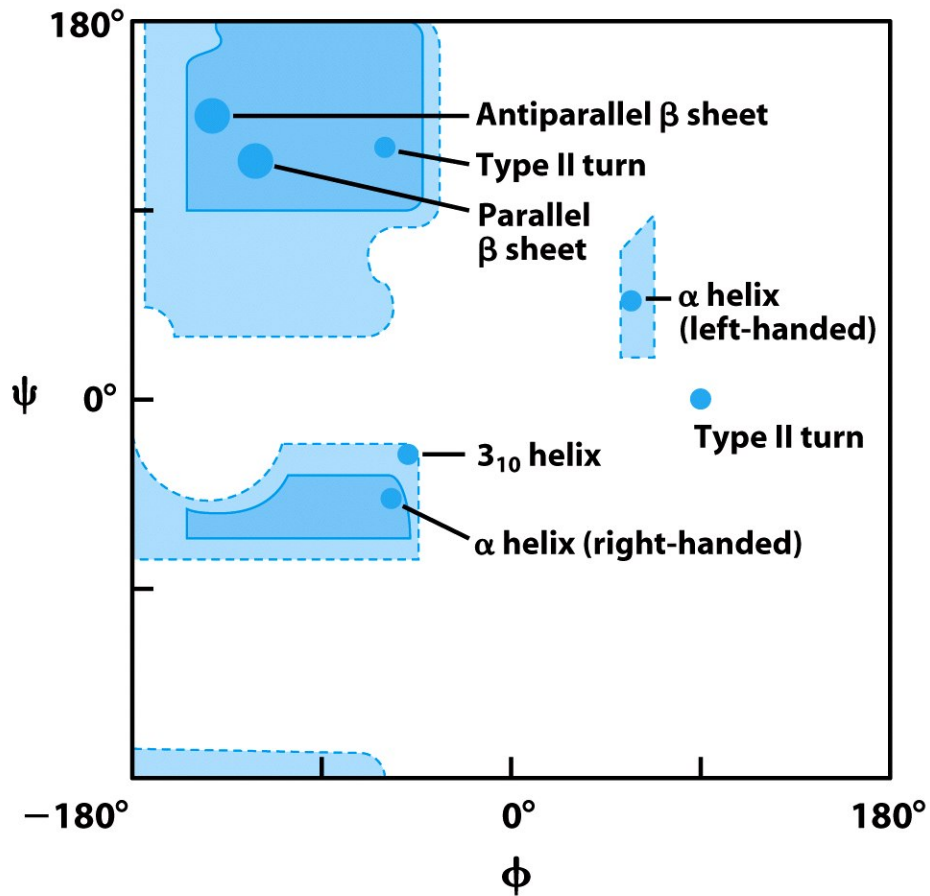
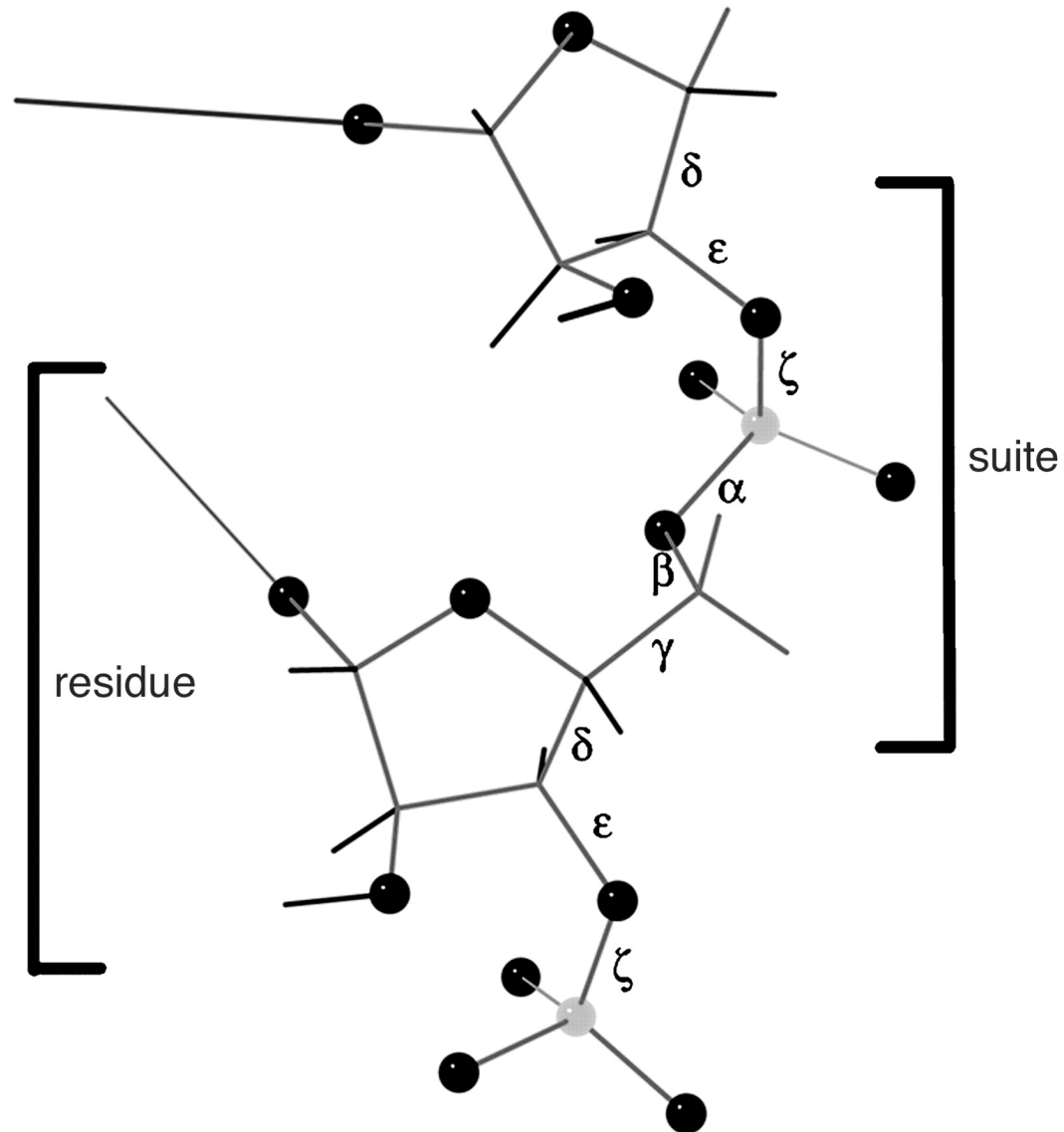
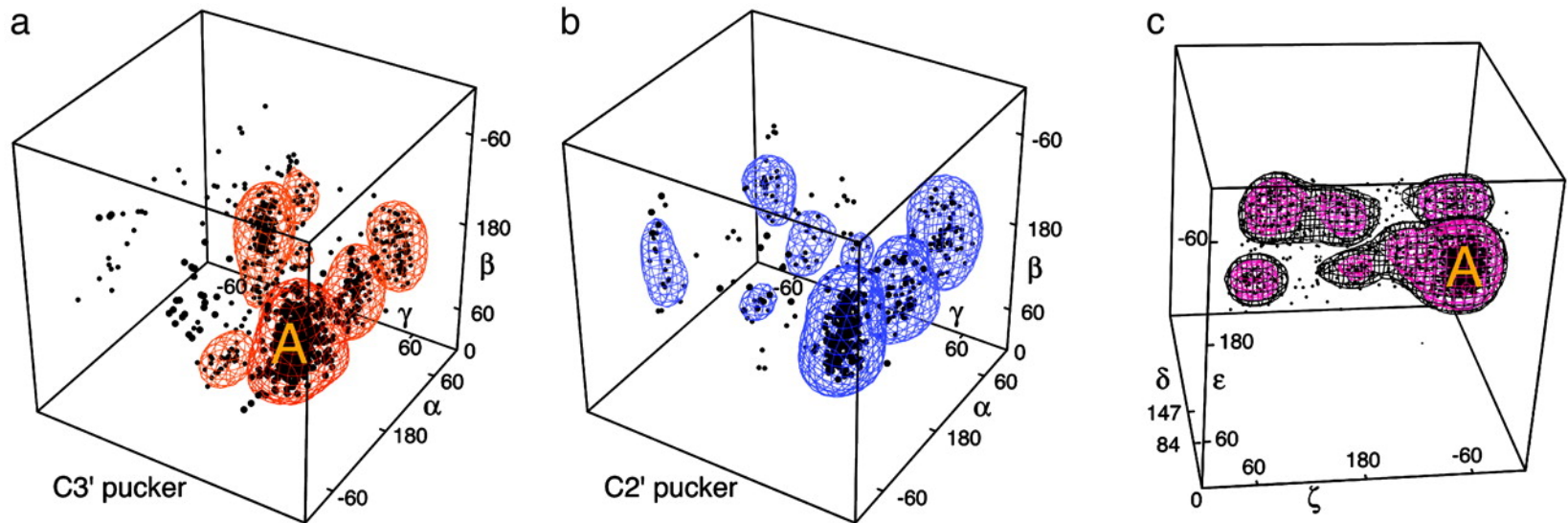


Figure 4-9a Principles of Biochemistry, 4/e
© 2006 Pearson Prentice Hall, Inc.

RNA backbone is rotameric



Plots of the heminucleotide angle triplets for data filtered by clashes and $B > 60$, with smoothed contours enclosing the top 7–10 peak clusters: (a) α – β – γ plot for adjacent sugar pucker C3' endo.



Murray L J W et al. PNAS 2003;100:13904-13909

RNA backbone rotamers

$\alpha \backslash \begin{matrix} \delta \\ \beta \quad \epsilon \\ \gamma \quad \zeta \end{matrix}$		C3' e p	C3' e t	C3' e -140	C3' e m	C2' e p	C2' e t	C2' e m
C3' (2nd)	p t p	*		*	*	**	*	*
	p 110 t	*						
	p t t					* ^a		
	t t p				*** ^b	*	**	*
	t 135 t				*			
	t t t				*** ^c			
	m t p		**	**	A ^d		** ^e	***
	m -135 p				*			
	-110 80 t				** ^f			
	m t t							*
C2' (2nd)	p t p	*	*	*		**	*	
	p 110 t	*						
	p t m			*				
	t t p	** ^g			* ^h			
	t t t				*			
	m t p		*	*	***	*	*	* ⁱ
	m -135 p				*** ^j			*
	m t m				*			*

Conformer frequencies:

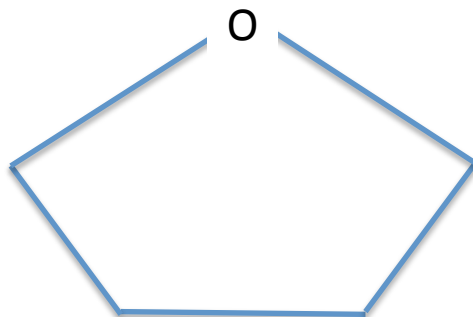
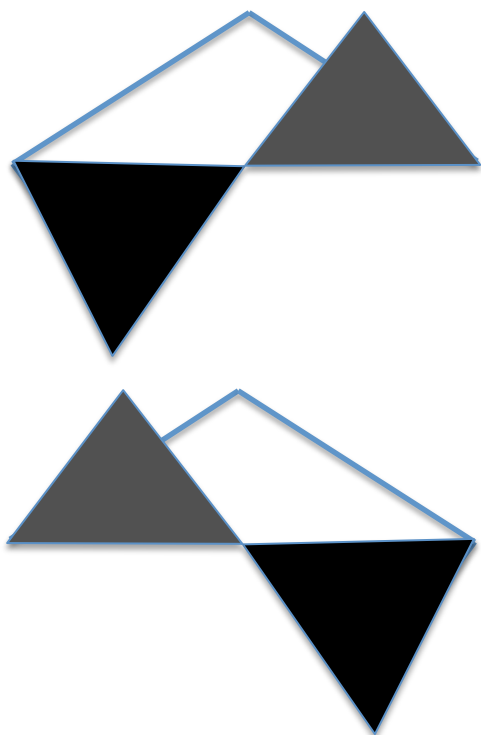
- * ~1% of non-A
- ** 2-3% of non-A
- *** 4-20% of non-A

Angle codes:

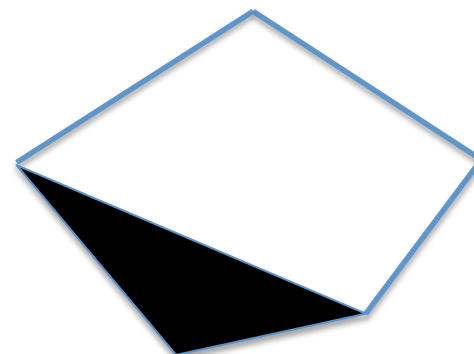
- t trans
- m gauche minus
- p gauche plus
- e eclipsed

Furanose ring

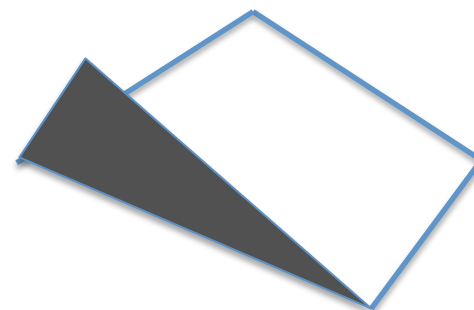
Twisted
(T)



Envelope
(E)



C3' exo



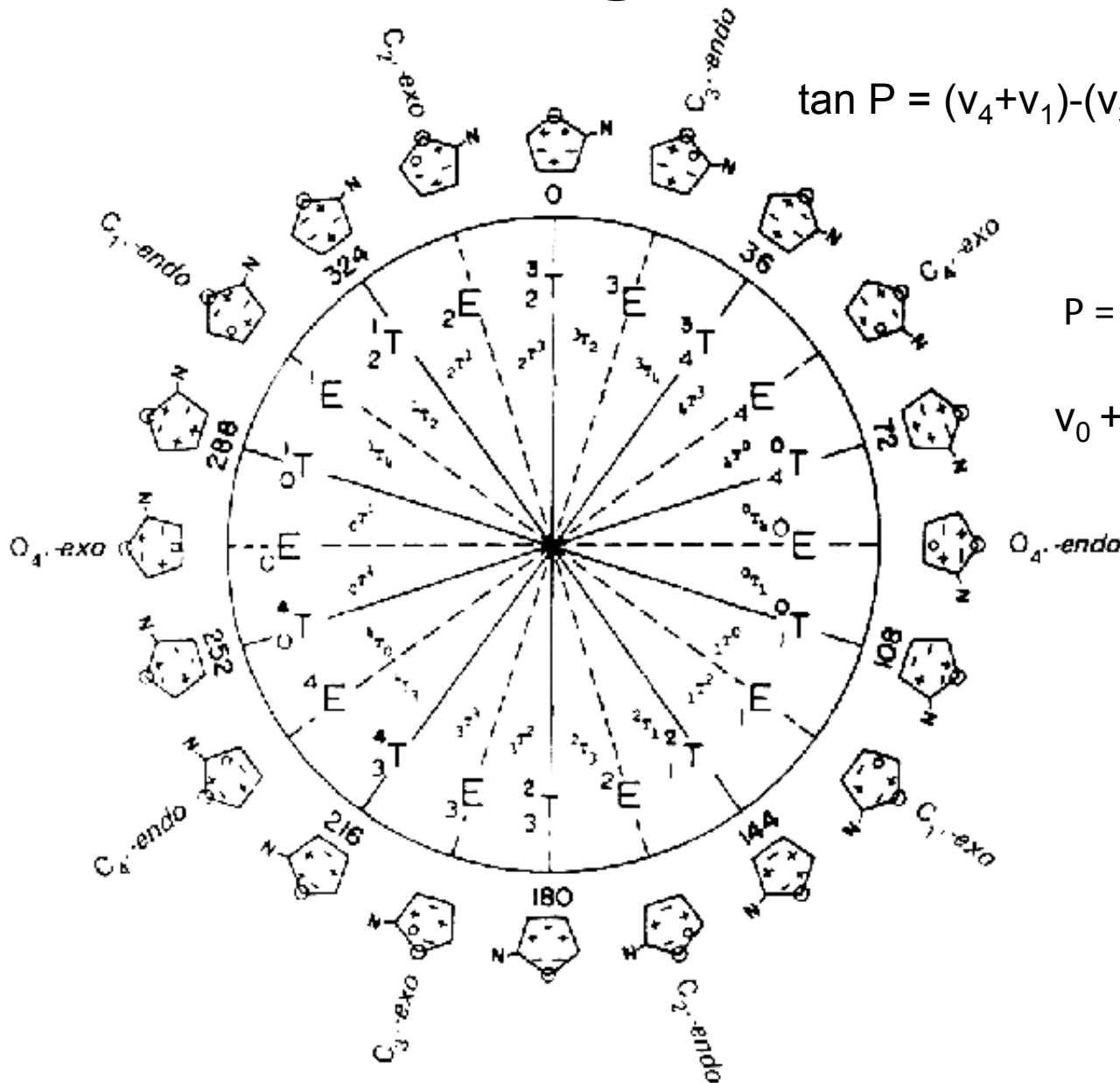
C3' endo

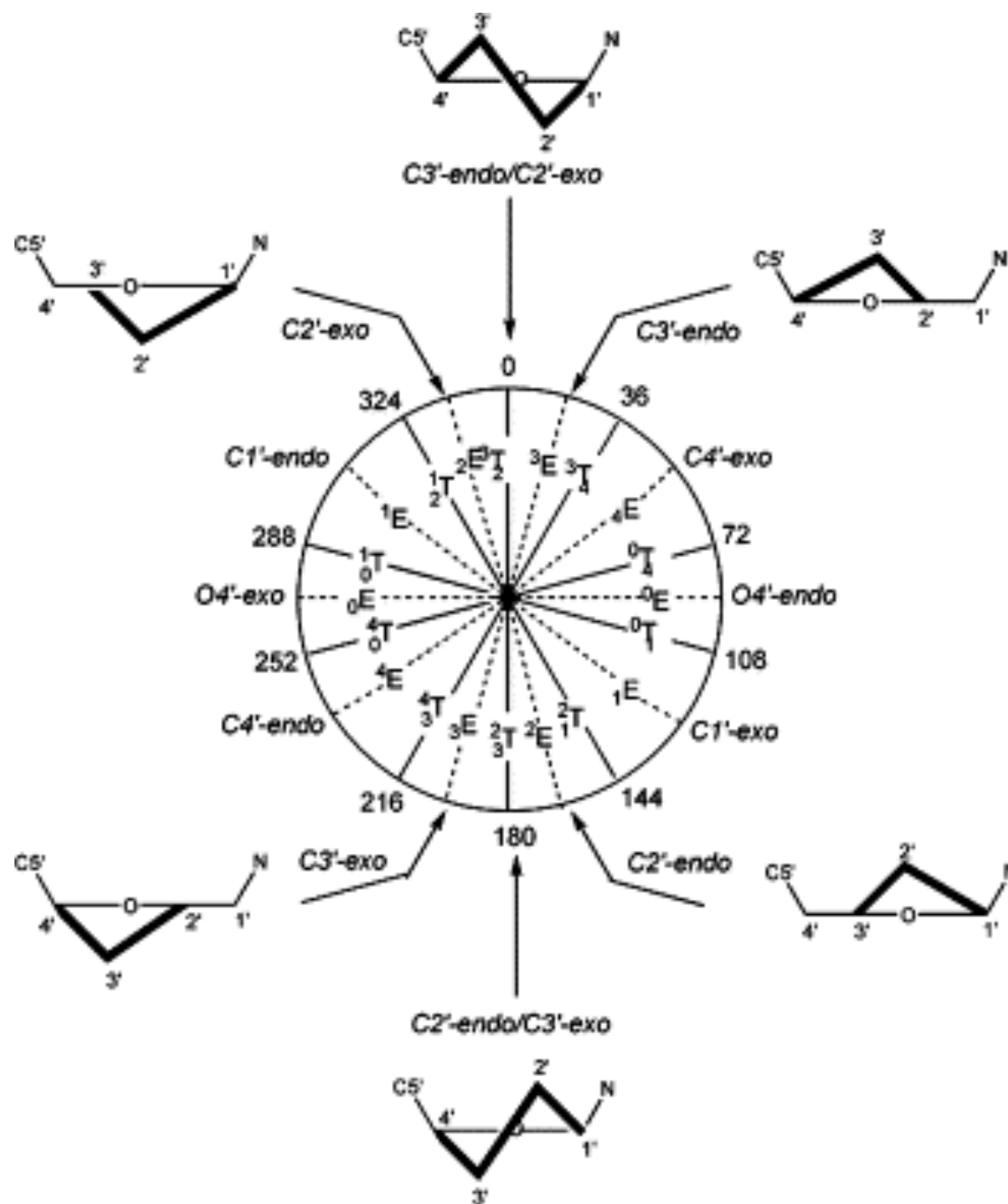
Sugar Puckering

$$\tan P = (v_4 + v_1) - (v_3 + v_0) / [2v_2(\sin 36^\circ + \sin 72^\circ)]$$

P = pseudorotation angle

$$v_0 + v_1 + v_2 + v_3 + v_4 = 0$$

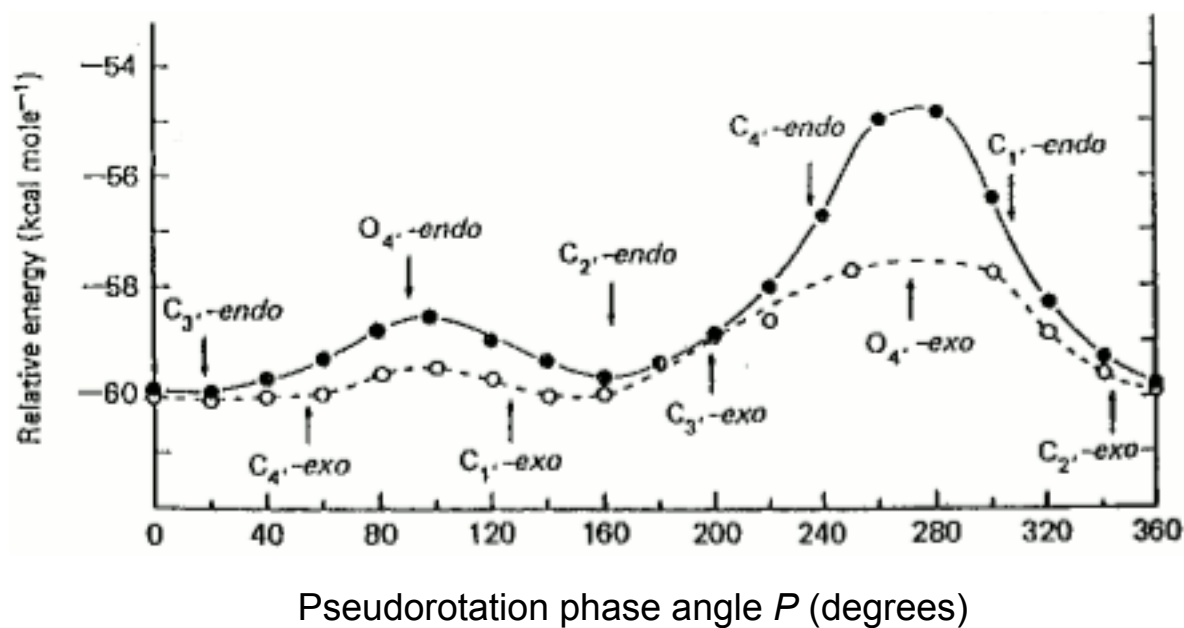




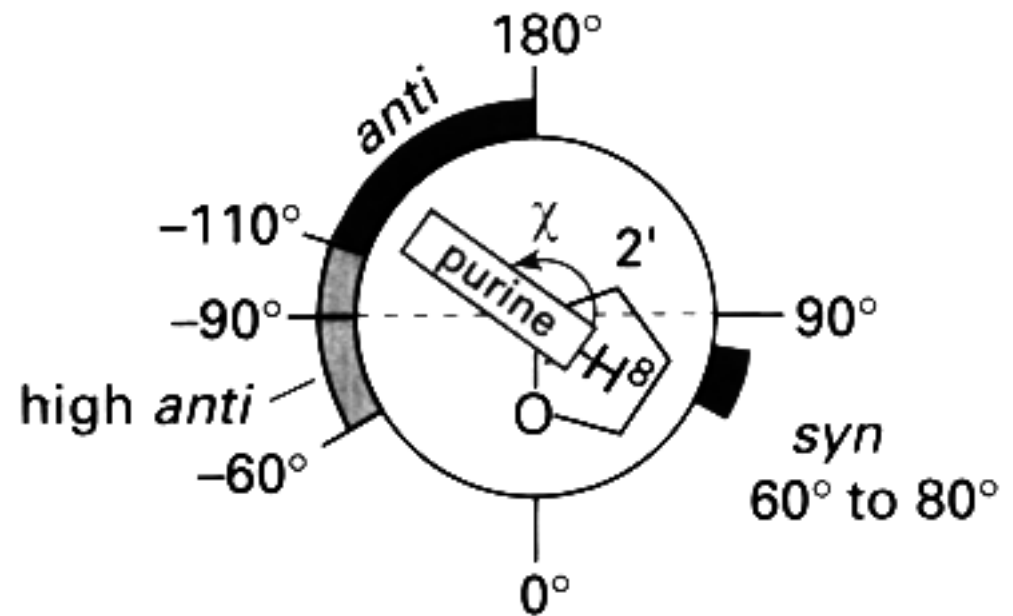
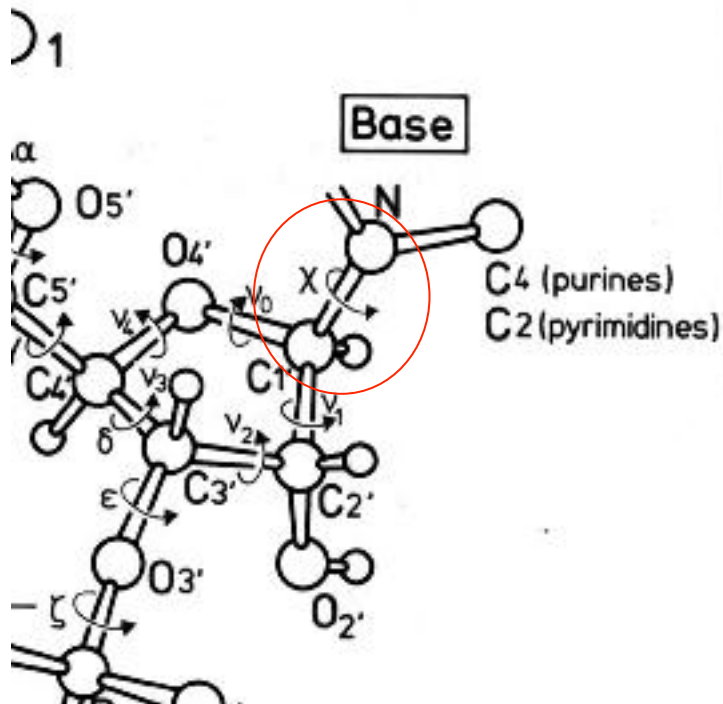
RNA

DNA

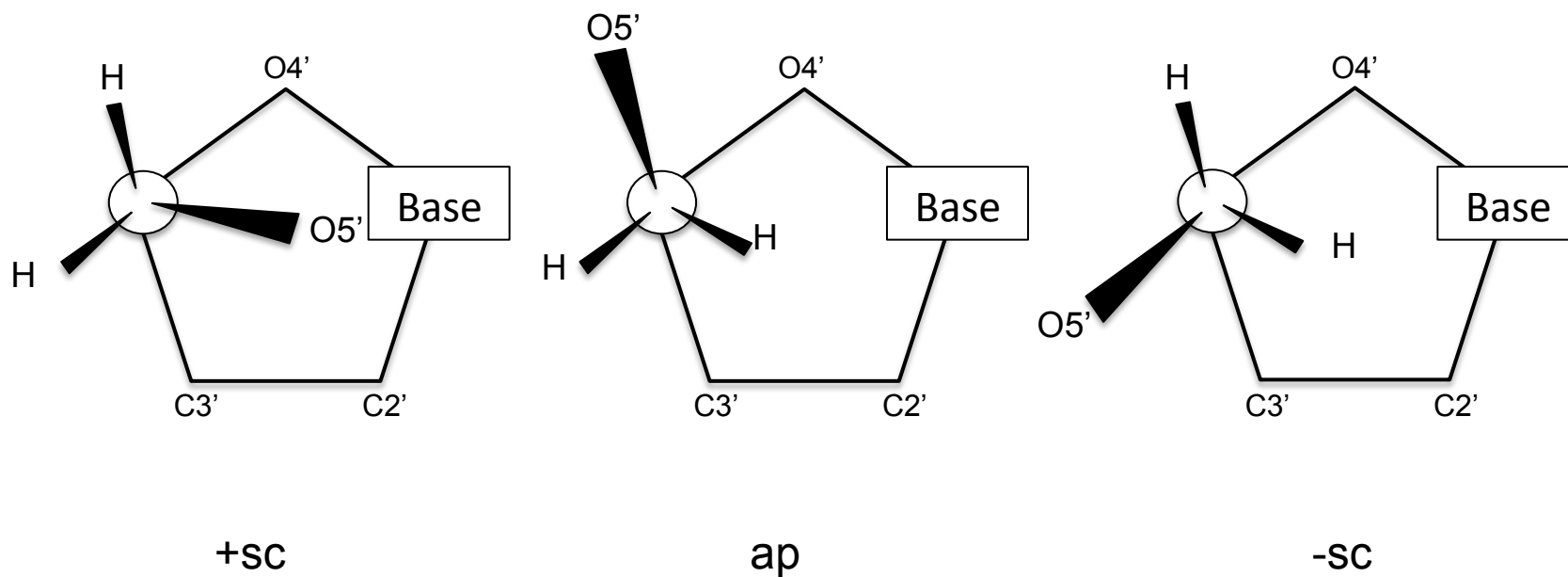
Energy variation related to P



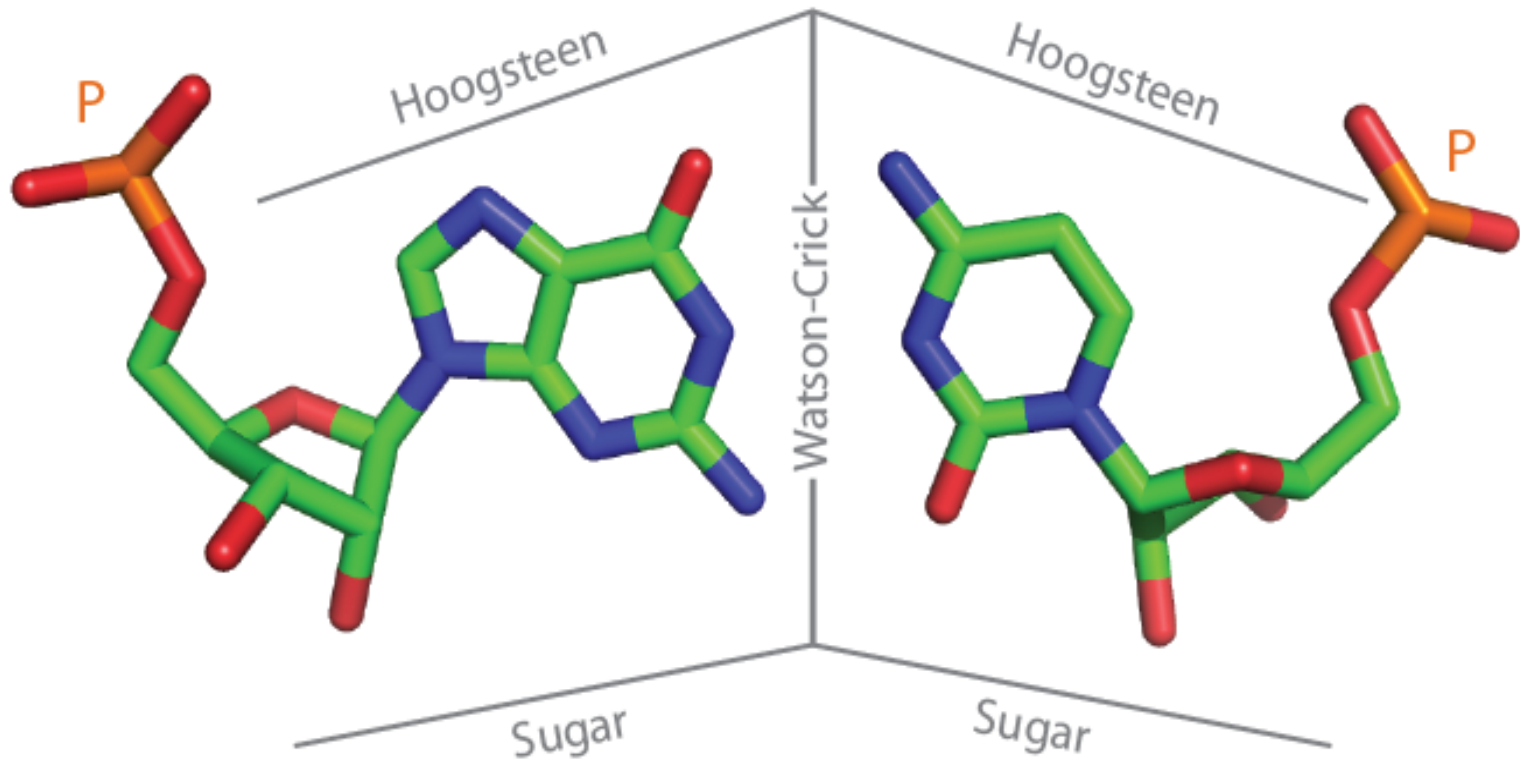
Orientation around glycosidic angle



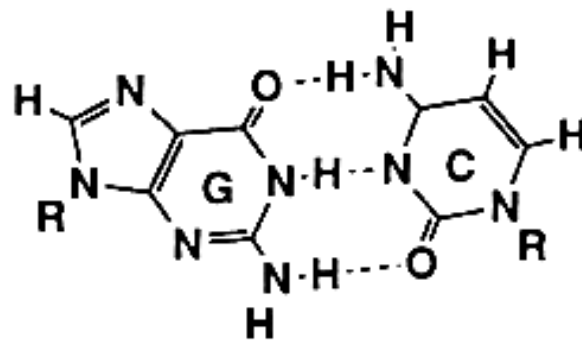
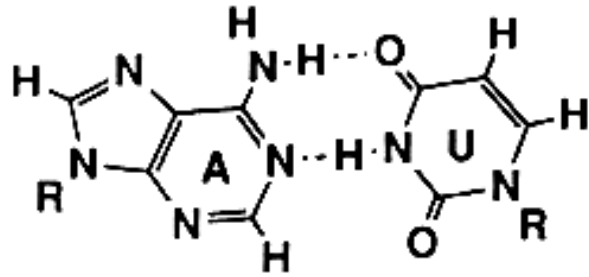
C4'-C5' torsion angle



Base pairs



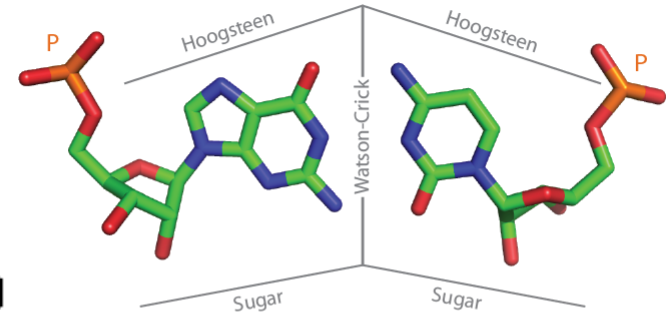
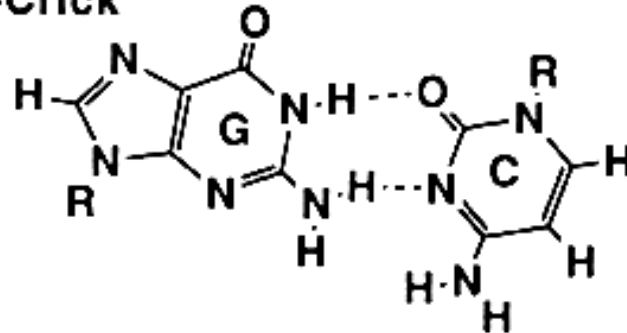
Base pairs



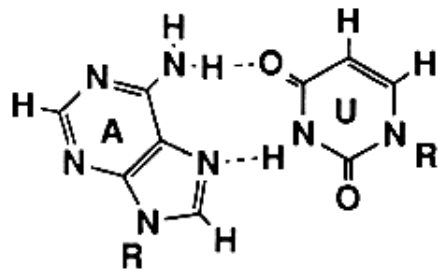
Watson-Crick



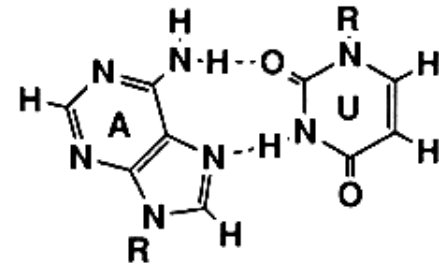
Reverse Watson-Crick



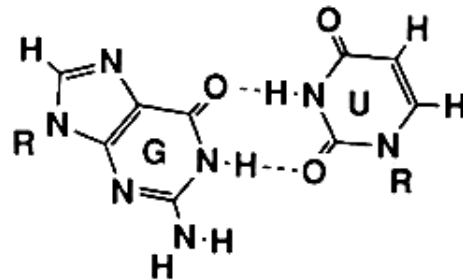
Hoogsteen, Wobble



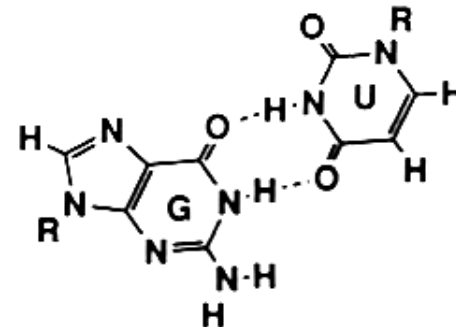
A•U Hoogsteen



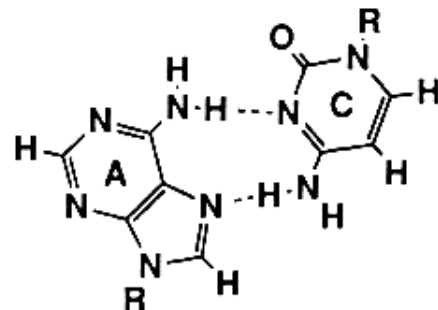
A•U Reverse Hoogsteen



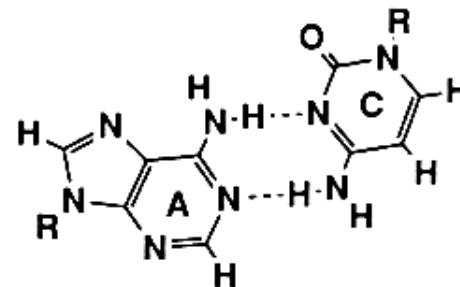
G•U Wobble



G•U Reverse Wobble

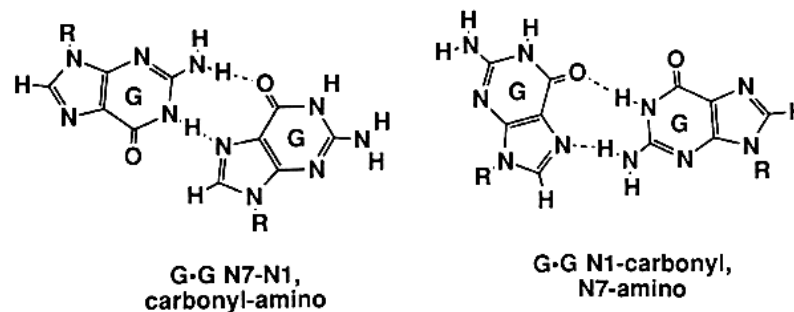
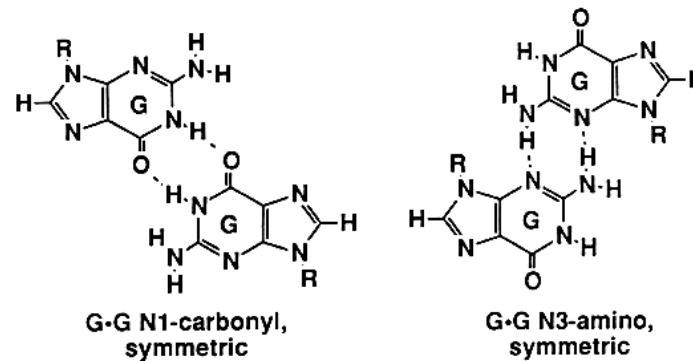
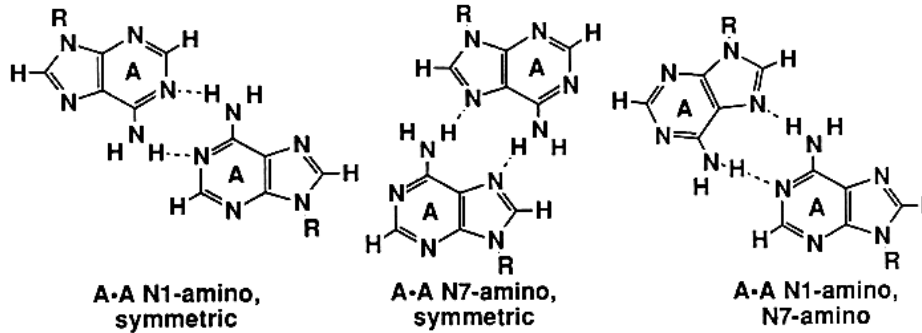


A•C Reverse Hoogsteen

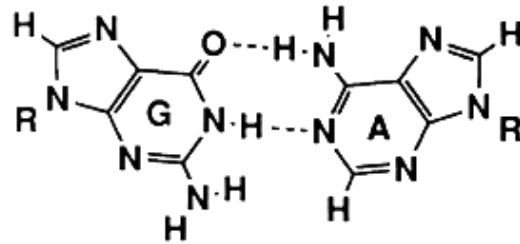


A•C Reverse Wobble

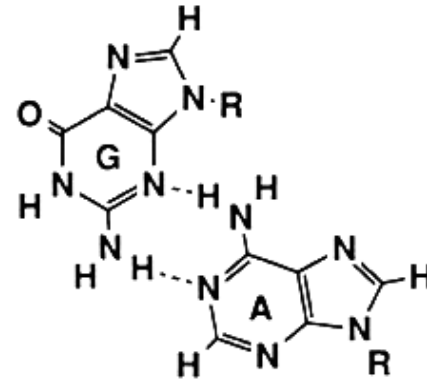
Homopurines



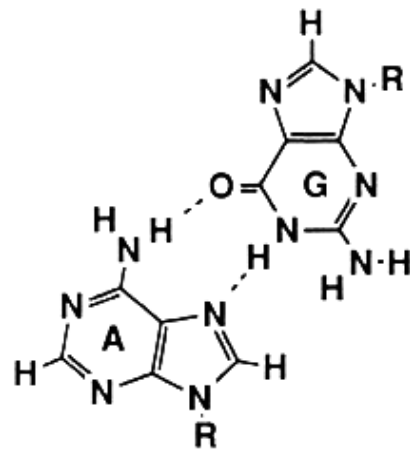
Heteropurines



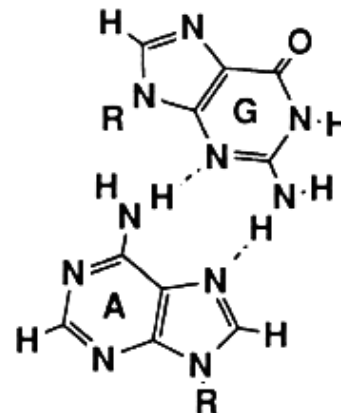
**G•A N1-N1,
carbonyl-amino**



**G•A N3-amino,
amino-N1**

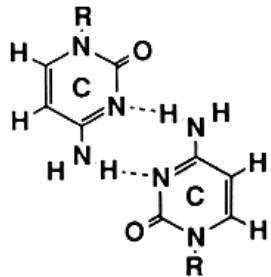


**A•G N7-N1,
amino-carbonyl**

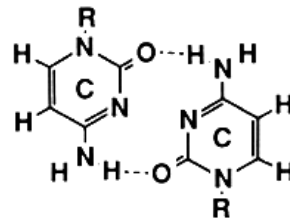


**A•G N7-amino,
amino-N3**

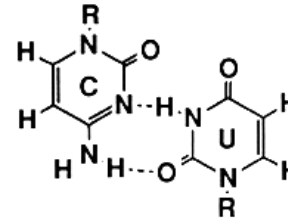
Pyrimidines



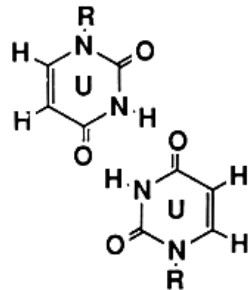
C·C N3-amino,
symmetric



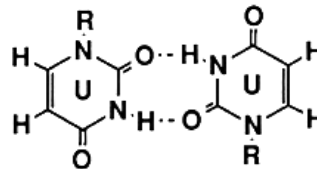
C·C carbonyl-amino,
symmetric



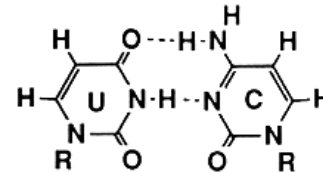
C·U N3-N3,
2-carbonyl-amino



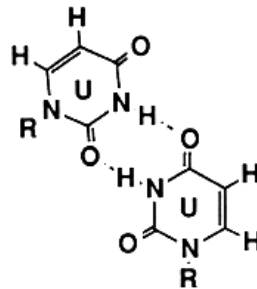
U·U 4-carbonyl-N3,
symmetric



U·U 2-carbonyl-N3,
symmetric

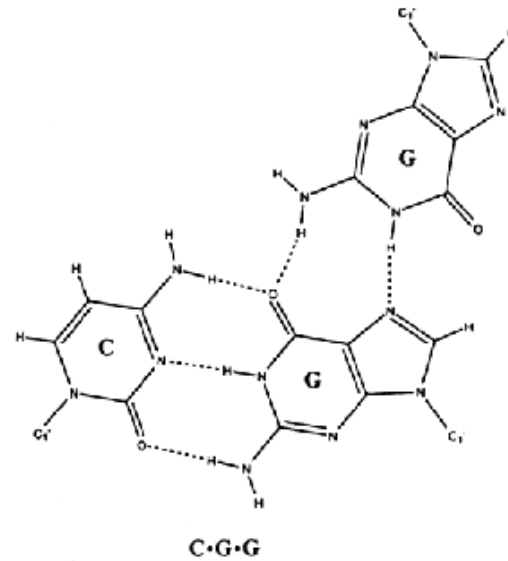
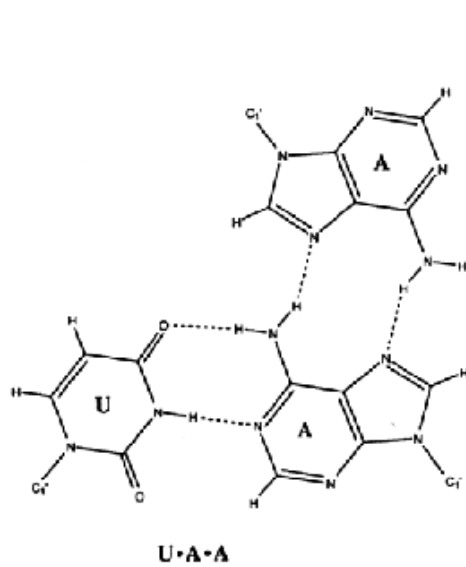
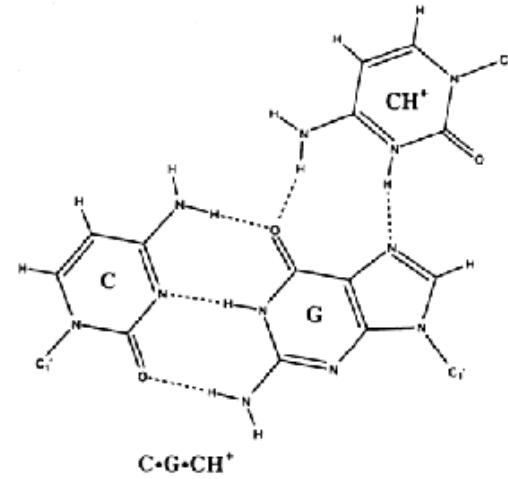
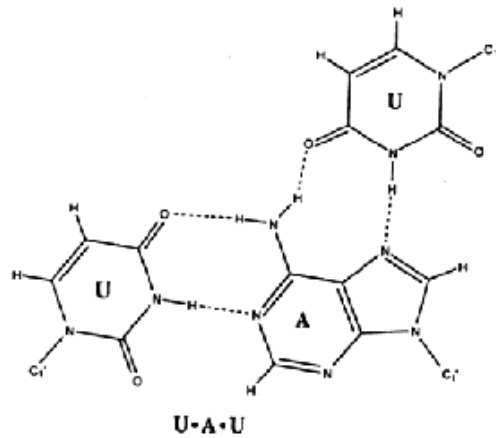


U·C N3-N3,
4-carbonyl-amino

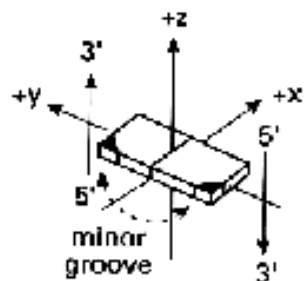


U·U 2-carbonyl-N3,
4-carbonyl-N3

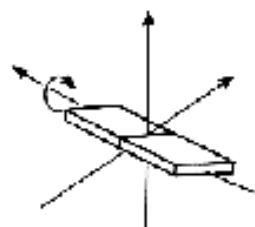
Base triples



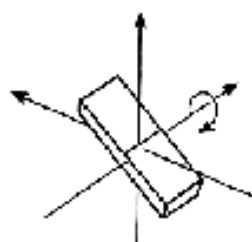
Movements of base pairs



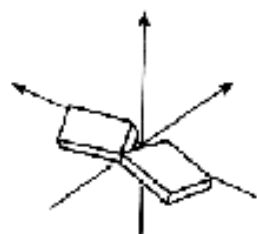
co-ordinate frame



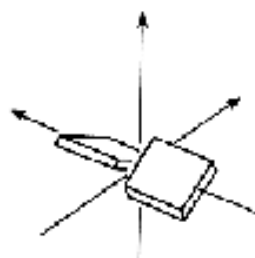
tip (θ)



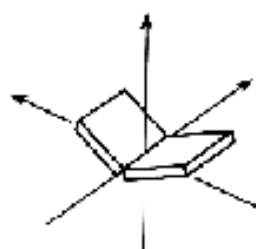
inclination (η)



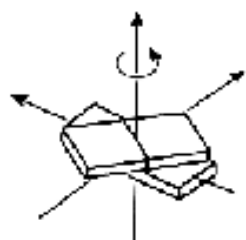
opening (σ)



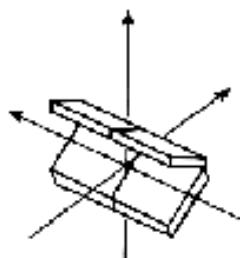
propeller (π)



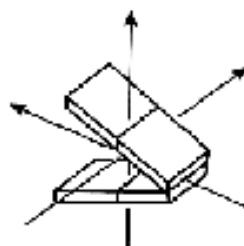
buckle (κ)



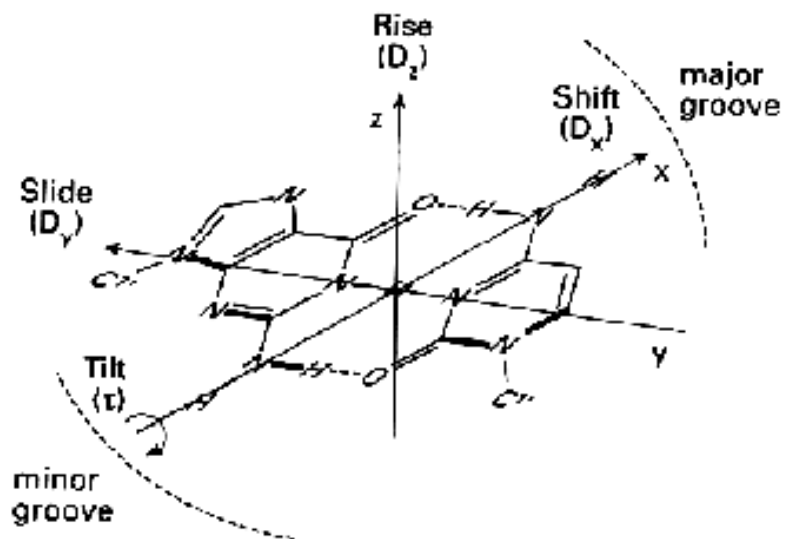
twist (Ω)



roll (P)



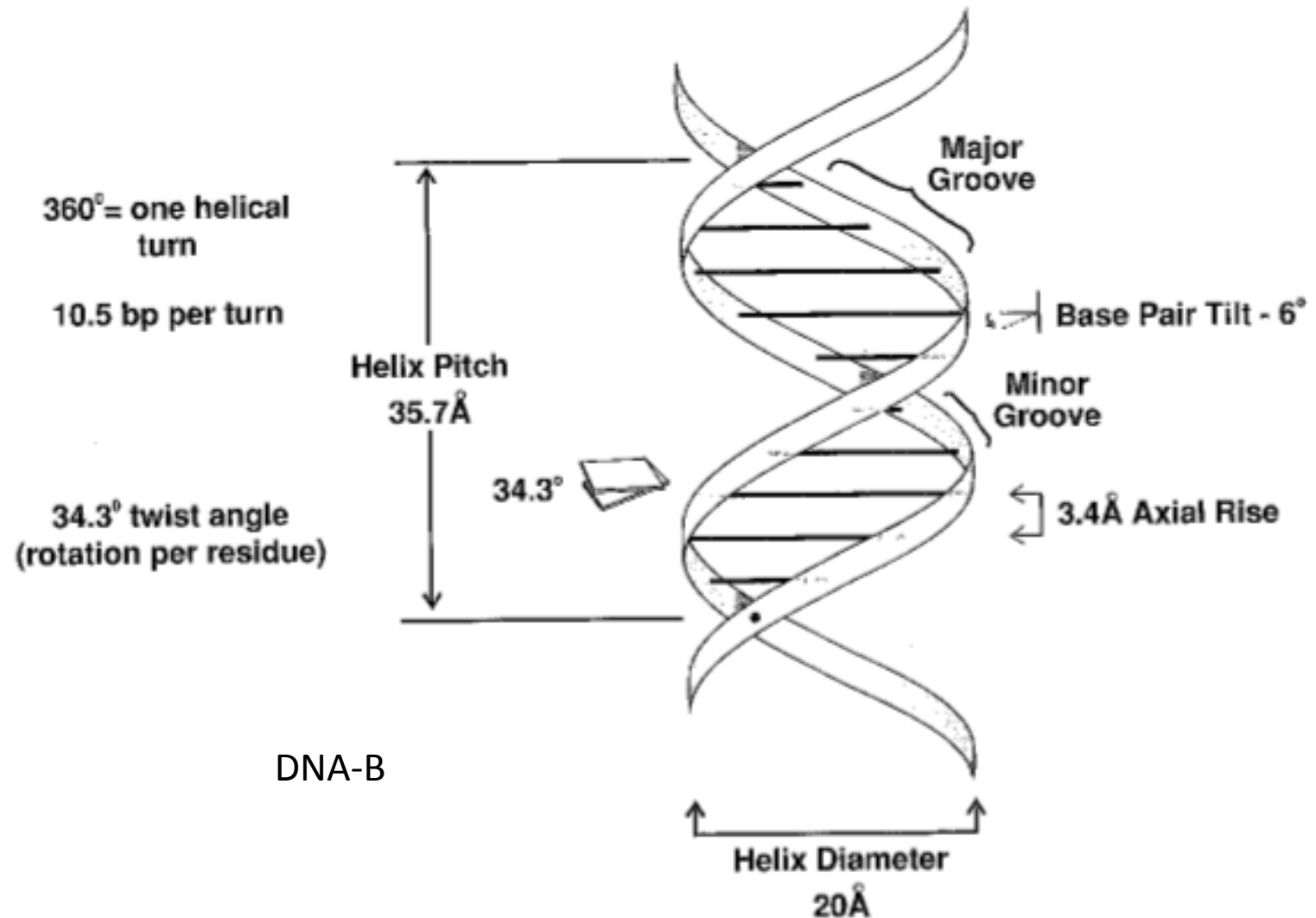
tilt (τ)



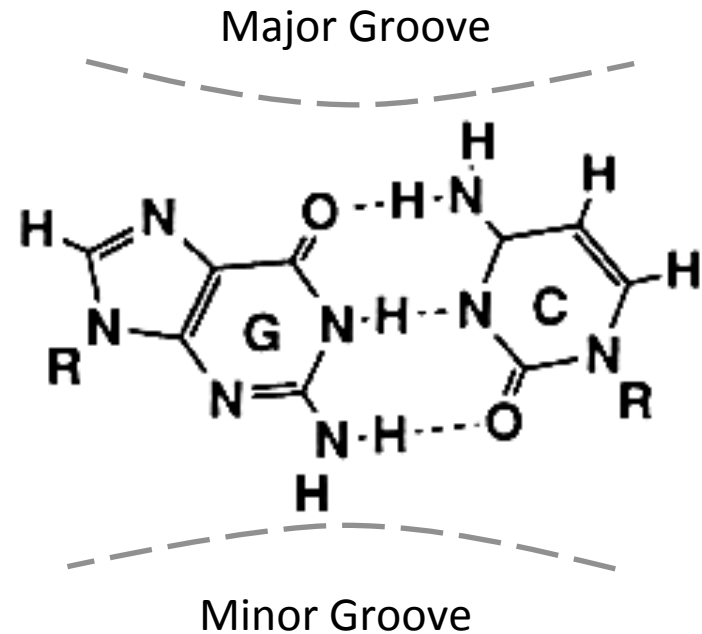
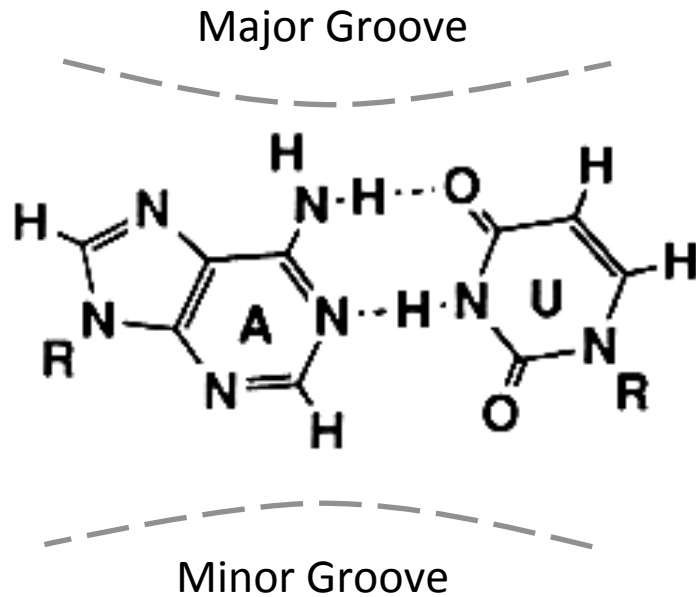
Helical parameters

- Sense. Helical rotation of the double helix. Clock- or anticklockwise
- Residues per turn. Number of residues in one helical turn.
- Axial rise. Distance between adjacent planar bases.
- Helix pitch. Length of one helical turn.
- Diameter of the helix. Width across the helix.
- Rotation per residue or Twist angle. Angle between two adjacent base pairs.

Helical parameters

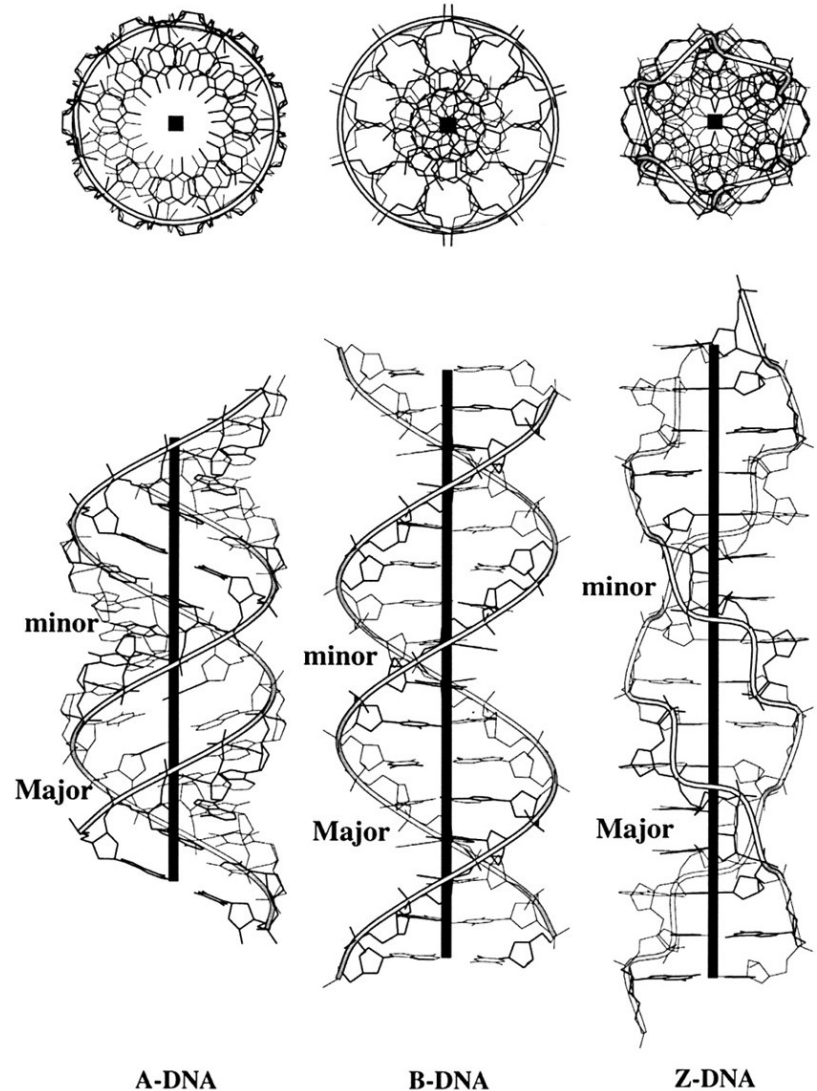


Major VS Minor Groove



Nucleic Acid Helical Structures

Geometry attribute:	A-form	B-form	Z-form
Helix sense	right-handed	right-handed	left-handed
Repeating unit	1 bp	1 bp	2 bp
Rotation/bp	33.6°	35.9°	60°/2
Mean bp/turn	11	10.5	12
Inclination of bp to axis	+19°	-1.2°	-9°
Rise/bp along axis	2.4 Å (0.24 nm)	3.4 Å (0.34 nm)	3.7 Å (0.37 nm)
Rise/turn of helix	24.6 Å (2.46 nm)	33.2 Å (3.32 nm)	45.6 Å (4.56 nm)
Mean propeller twist	+18°	+16°	0°
Glycosyl angle	anti	anti	pyrimidine: anti,
			purine: syn
Sugar pucker	C3'-endo	C2'-endo	C: C2'-endo,
			G: C2'-exo
Diameter	23 Å (2.3 nm)	20 Å (2.0 nm)	18 Å (1.8 nm)

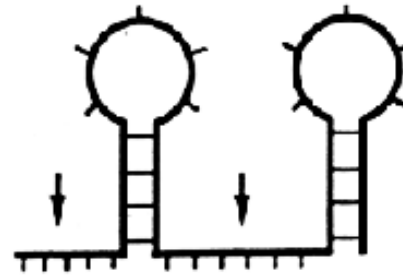


Secondary structures in RNA

a. DUPLEXES



b. SINGLE STRANDED REGIONS



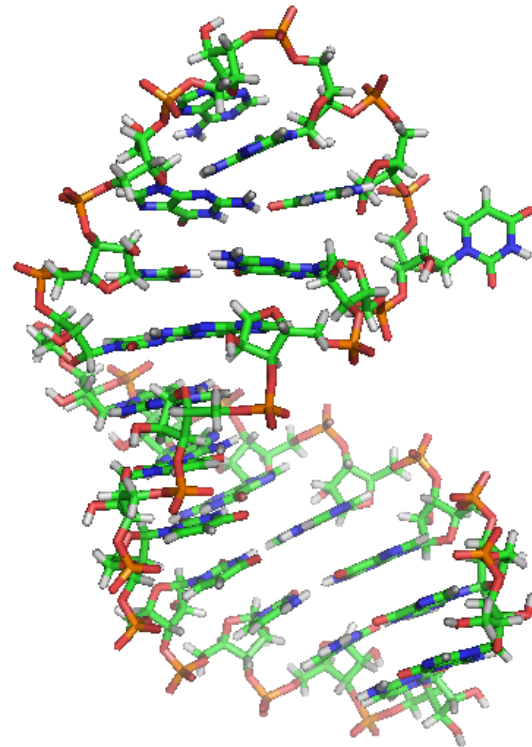
Apical loops (hairpins)

c. HAIRPINS



HAIRPIN LOOP

HAIRPIN STEM

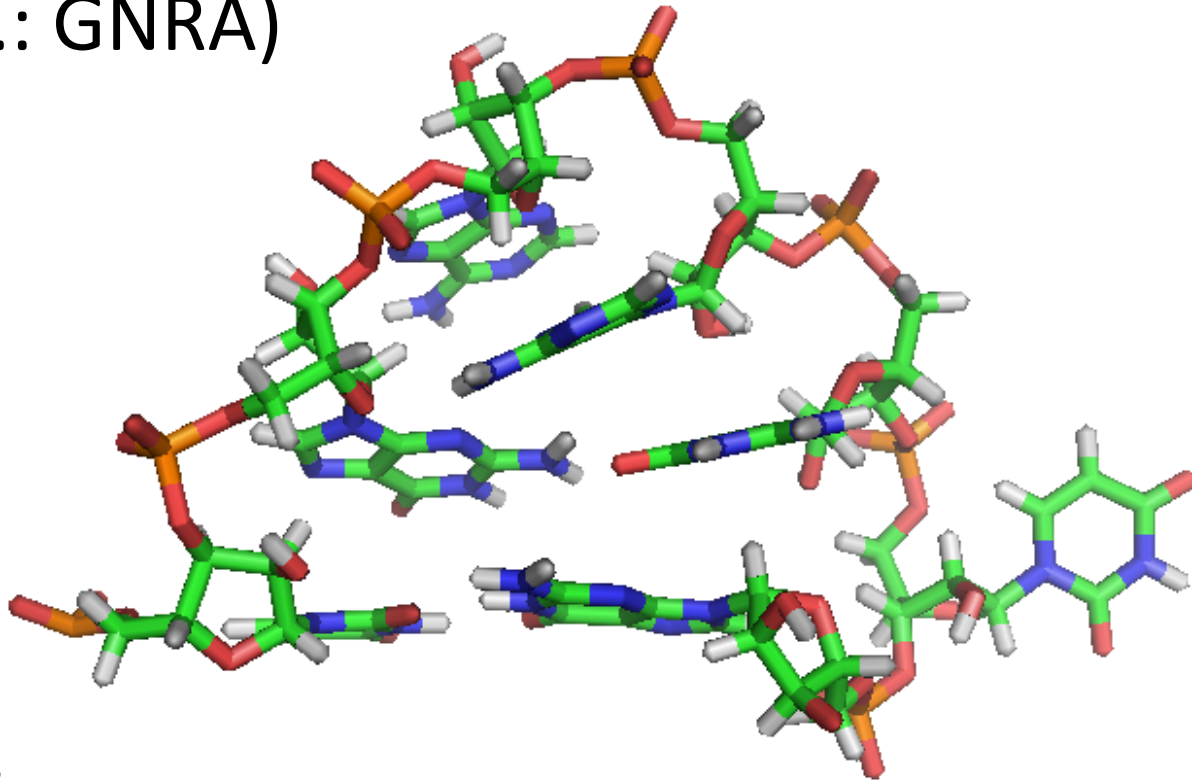


Several kinds of apical loops

- Minimum: triloops
- Tetraloops (e.g.: GNRA)
- Pentaloops
- Hexaloops

- .
- .
- .

- whatever loops



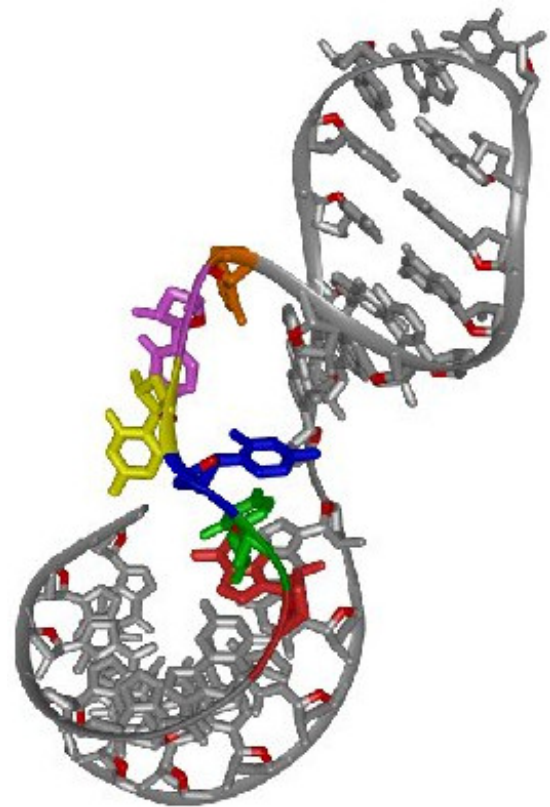
d. BULGES



BULGE



SINGLE-BASE BULGE



e. INTERNAL LOOPS



MISMATCH

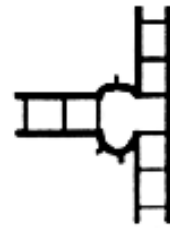


**SYMMETRIC
INTERNAL LOOP**

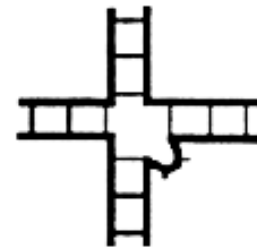


**ASYMMETRIC
INTERNAL LOOP**

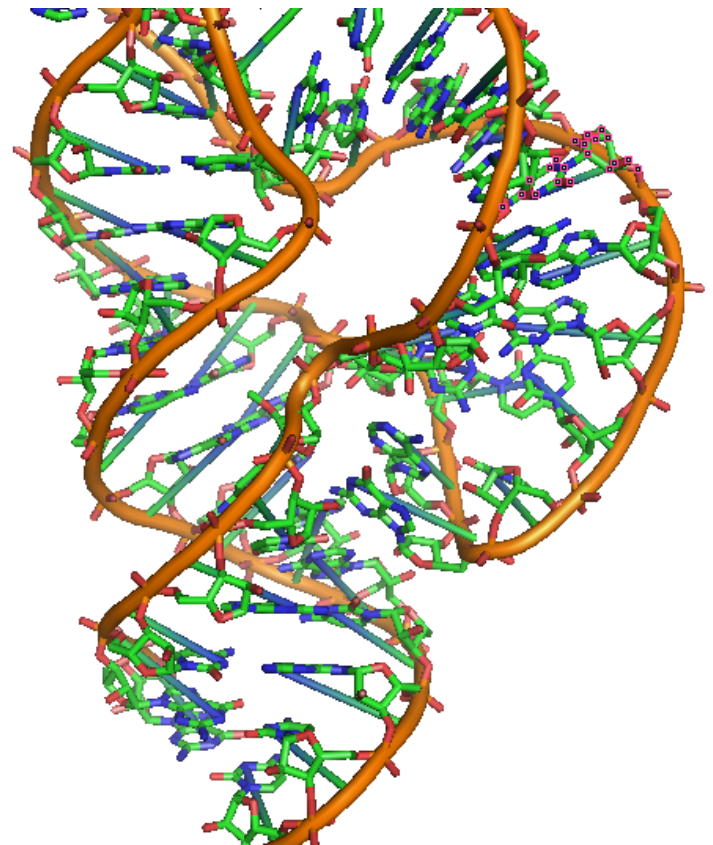
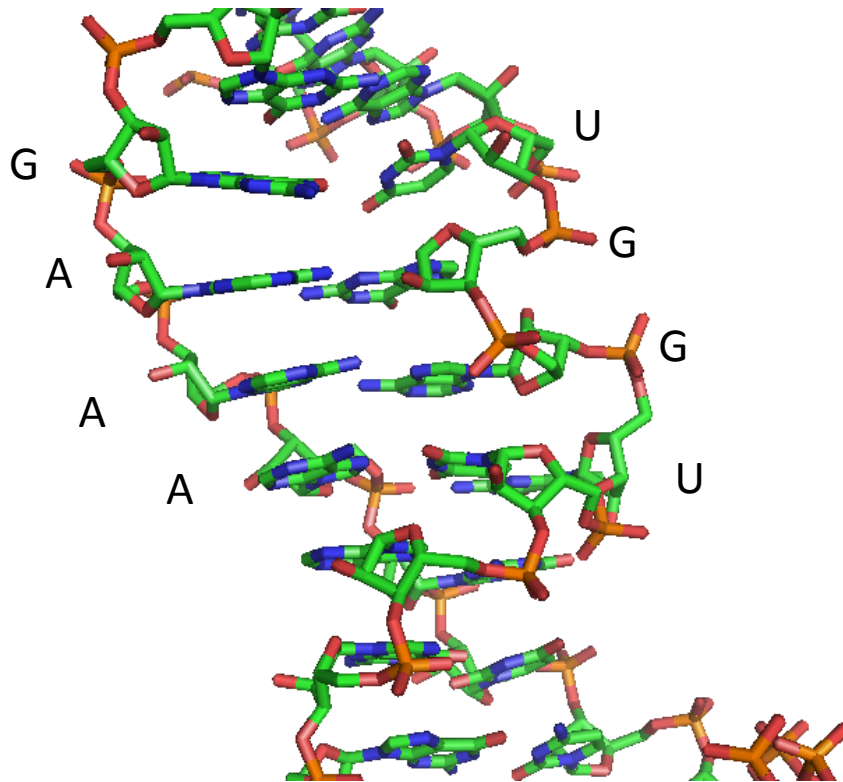
f. JUNCTIONS



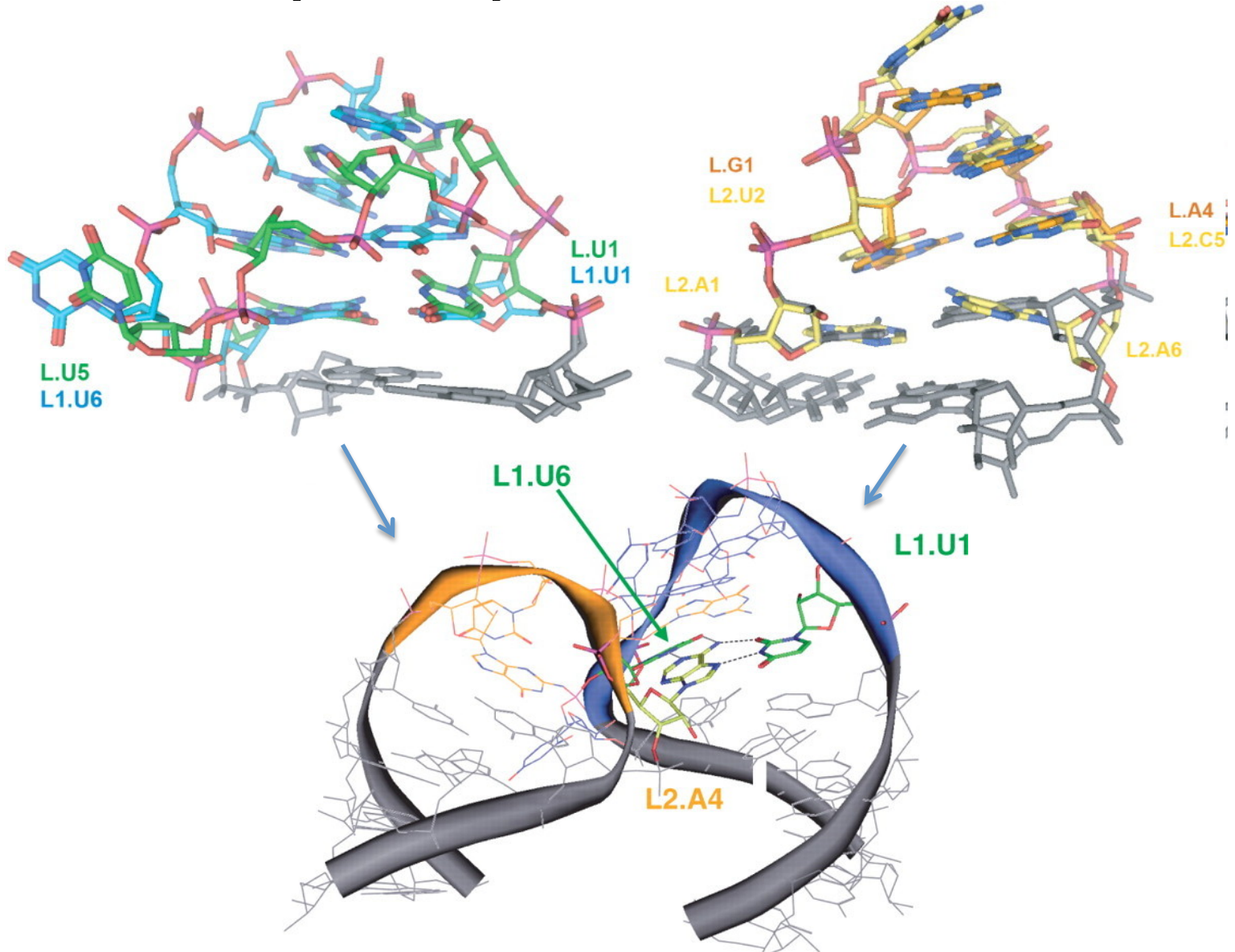
THREE STEM



FOUR STEM



Loop-loop interactions

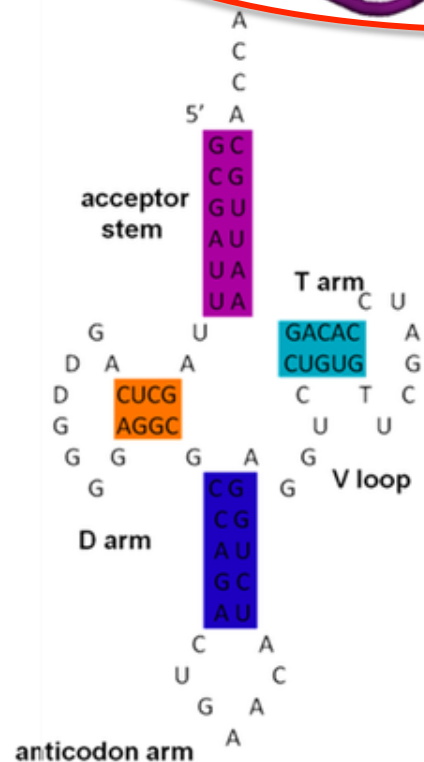


A 3D ribbon diagram of a tRNA molecule. The acceptor stem is highlighted in purple, the T arm in cyan, the D arm in orange, and the anticodon arm in blue. A red circle highlights the acceptor stem and the beginning of the T arm.

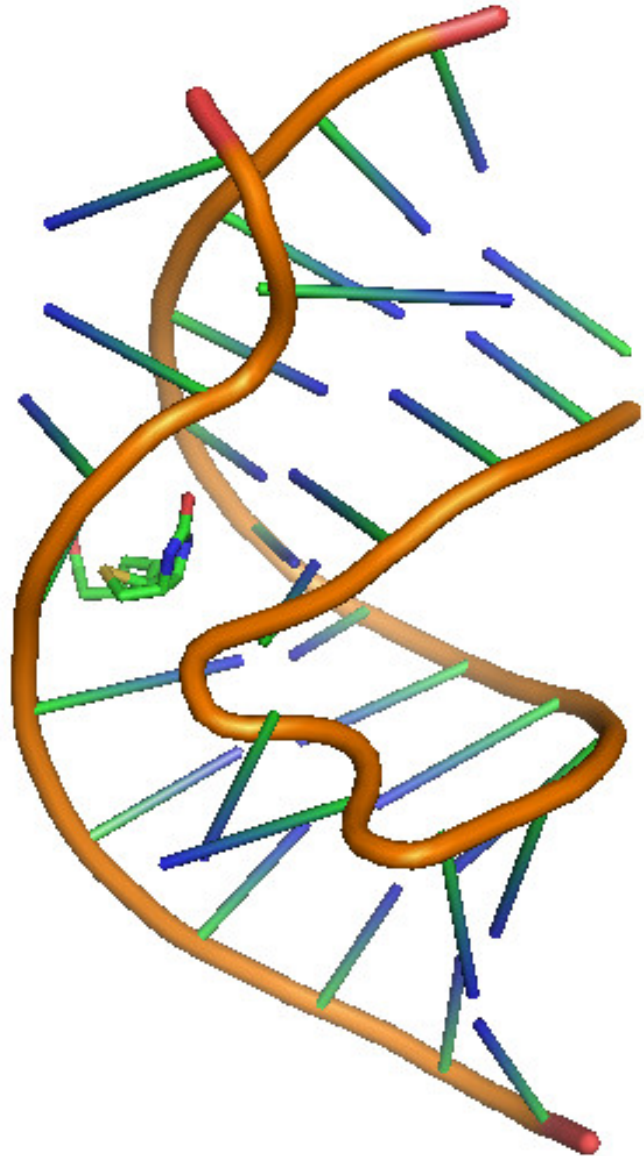
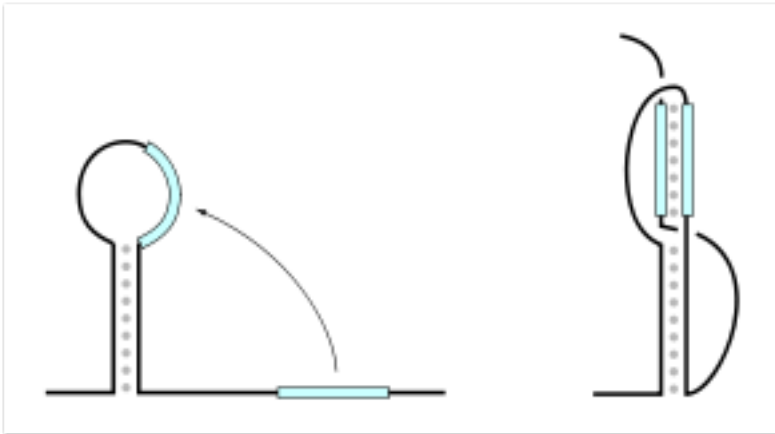
Sequence Diagram:

```

      A
      C
      C
      A
5'  A
    GC
    CG
    GU
    AU
    UA
    UA
    acceptor stem
      T arm
      C U
      GACAC
      CUGUG
      C T C
      U U
      G
      V loop
      G
      D arm
      CG
      CG
      AU
      GC
      AU
      C A
      U A
      G A
      A
anticodon arm
  
```

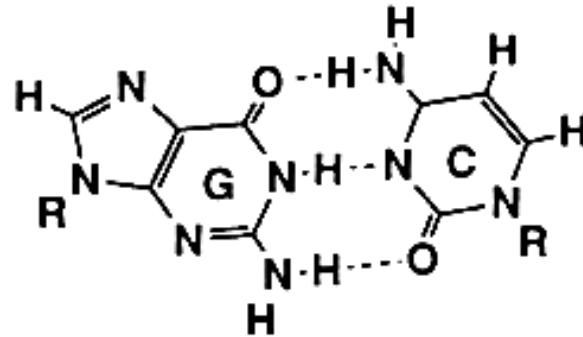
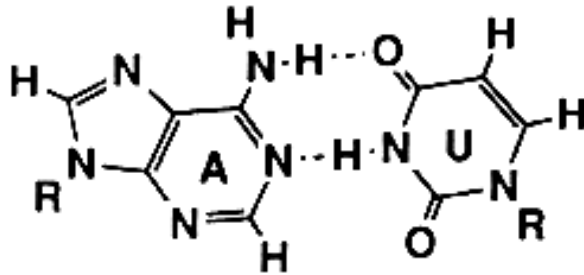


Pseudoknots



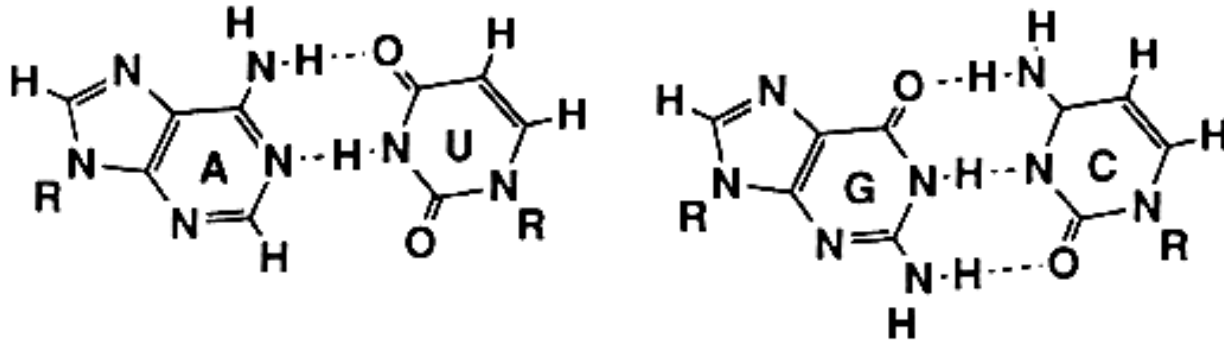
Forces that drive RNA folding

- 1) Hydrogen bonds

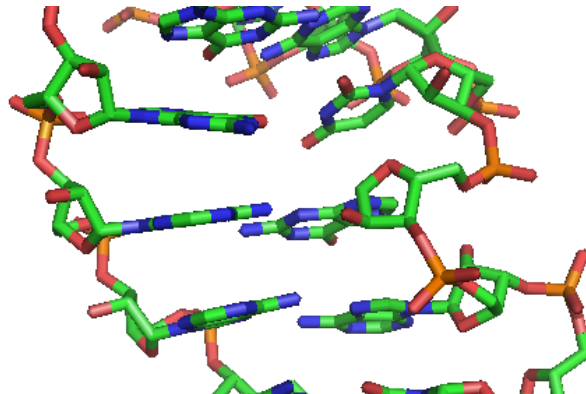


Forces that drive RNA folding

- 1) Hydrogen bonds

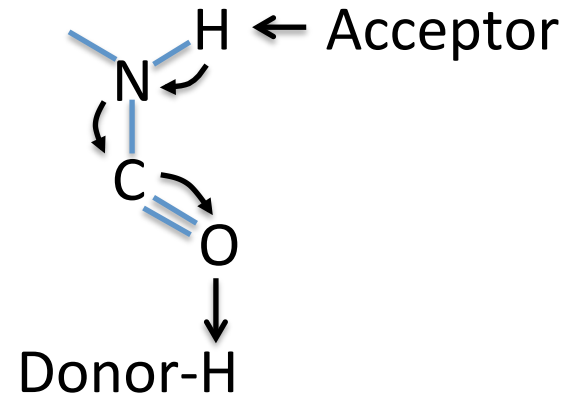


- 2) Stacking interactions



Hydrogen Bonding

- Short, noncovalent, highly directional between an electronegative atom (-C=O, -N:) and a H atom bound to an electropositive atom (-N-H).
- 20-30 times weaker compared to covalent bonds (3-7 kcal/mol VS 80-100 kcal/mol)
- Additive and cooperative



Base Stacking

- Base stacking = Van der Waals forces + Hydrophobic interactions
- 4-15 kcal/dinucleotide
- Van der Waals = dipole-dipole interactions + London dispersion forces

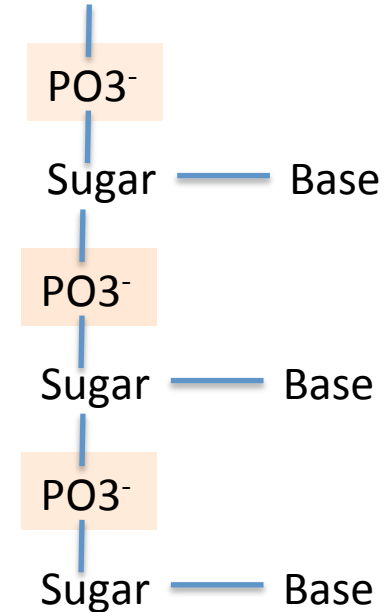
Dinucleotide base pairs	Stacking energies (kcal/mol/stacked pair)
GC · GC	-14.59
AC · GT	-10.51
TC · GA	-9.81
CG · CG	-9.69
GG · CC	-8.26
AT · AT	-6.57
TG · CA	-6.57
AG · CT	-6.78
AA · TT	-5.37
TA · TA	-3.82

Base Stacking

- Isodesmic reaction
- Additive
- Diffusion controlled reaction
- Weak force controlled process
- $E_{\text{tot}} = E_{\text{d}} + E_{\text{p}} + E_{\text{el}}$
- E_{tot} = Total energy
- E_{d} = Dispersion energy
- E_{p} = Polarization energy
- E_{el} = Electrostatic energy ($\approx 80\% E_{\text{tot}}$)

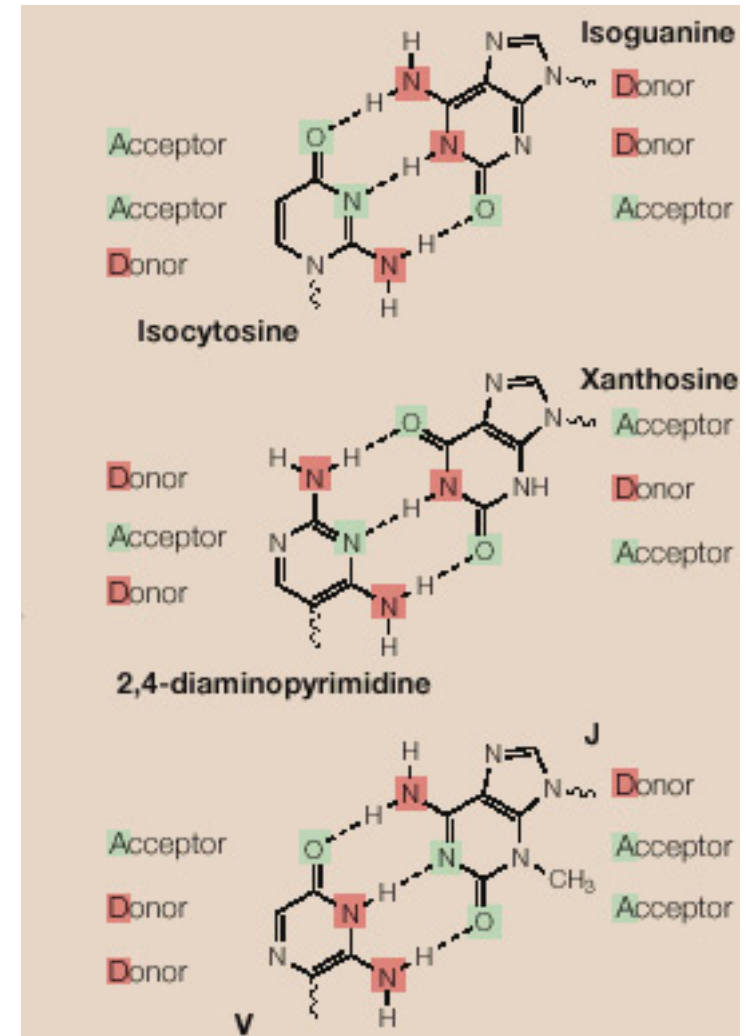
Other interactions

- RNA is a polyanion
- Metal ion coordination (Mg^{2+})
 - Diffusion controlled
 - Though water or direct contact
 - PO_3^- or bases



Bonus track

- Why are there four letters in the genetic alphabet?



NDB

<http://ndbserver.rutgers.edu/>



WELCOME TO THE NUCLEIC ACID DATABASE

a repository of three-dimensional structural information about nucleic acids

[Site Index](#)

- [Atlas](#)
- [Deposit Data](#)
- [Download Data](#)
- [Search](#)
- [Education](#)
- [Standards](#)
- [Tools](#)
- [Links](#)

Number of Released Structures:

6425 Structures

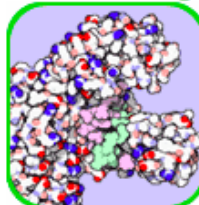
Last Update: 20-Feb-2013

Search the NDB by ID

Enter an NDB ID or PDB ID

Search for Released Structures

Nucleic Acids Highlight



About NDB

The NDB follows the dictionaries and formats used by the Worldwide Protein Data Bank. Please see www.wwpdb.org for format announcements and documentation.

Archive of NDB newsletters

The NDB is supported by funds from the [National Science Foundation](#) and the [Department of Energy](#).

In citing the NDB please refer to: H. M. Berman, W. K. Olson, D. L. Beveridge, J. Westbrook, A. Gelbin, T. Demeny, S.-H. Hsieh, A. R. Srinivasan, and B. Schneider. (1992) The Nucleic Acid Database: A Comprehensive Relational Database of Three-Dimensional Structures of Nucleic Acids. *Biophys. J.*, 63, 751-759.

ndbadmin@ndbserver.rutgers.edu

©1995-2012 The Nucleic Acid
Database Project Rutgers, The State
University of New Jersey