BIOCH 590: Biomacromolecules Spring 2010

Introduction to Biological Macromolecules

General Principles and Key Concepts

Copyright: Jianhan Chen

General Introduction

- Course setup: two major parts
  - Part I: physical biochemistry (Jan 15 to March 12)
  - Part II: spectroscopy (March 15 to May 7)
- Lecture notes will be main study materials
  - Will posted ahead of lectures on K-State online
- Two recommended textbooks
  - Not required, but very useful resources for additional reading
  - Available from the Hale Library Reserve
- Four exams (two for each part)
  - Feb 12, Mar 12, April 9 and May 7 during lecture time (Please mark these dates!)
- Grading
  - ~ 70% from exams and ~30% from homework assignments

Course Overview

- Scope: physical properties of biological macromolecules
  - Basic principles of structure at multiple levels
  - Focus on the quantitative aspects
  - Theoretical frameworks: thermodynamics, statistical mechanics
  - Physical biochemistry: molecular interactions and interactions of macromolecules with environments
  - Experimental techniques to measure certain properties, which can be then used to infer about the structural organizations
- For this course
  - Emphasize a qualitative understanding of the quantitative nature
  - Build an appreciation of structural complexity and principles of theoretical and experimental studies of molecular structures
  - In Part II, a survey of basic and practical principles of common spectroscopic techniques for studying molecular structures

Tentative Topics

Part 1: Physical principles of biomolecular structure and dynamics

- Introduction to biological macromolecules
- Classical and statistical thermodynamics; Thermochemistry
- Molecular interactions and biomolecular structures; Computer modeling
- Size and shape of macromolecules: Diffusion and light scattering

Part 2: Techniques for studying biomolecular structure and dynamics

- Absorption spectroscopy (IR, UV)
- Fluorescence spectroscopy
- Circular dichroism (CD) spectroscopy
- Nuclear magnetic resonance (NMR) spectroscopy
- X-ray crystallography
- Mass spectroscopy

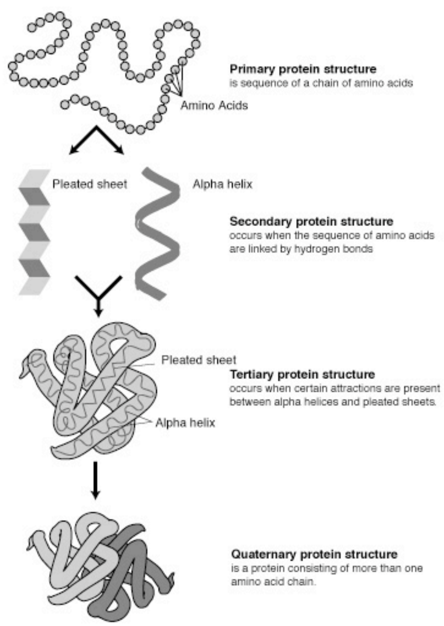
Topic 1

OVERVIEW OF BIOLOGICAL MACROMOLECULES

Reference: Chapter 1 of van Holde

Biological Macromolecules

- Atom: smallest unit of an element (C, N, O, H etc)
- Molecule: smallest unit of an composite (H_2O, CO_2, C_6H_6 etc)
- Macromolecule: a large molecule
  - Smallest molecule: H_2
  - Largest molecules: DNAs (tens of billions of atoms!)
- Biological macromolecule: a large and complex molecule with specific biological functions
  - Typically biopolymers: monomers, oligomers
  - Hierarchical structures
    - Primary (1^o): sequence of monomers (sugar, amino acid, nuclei acid, etc)
    - Secondary (2^o): local regular structures (alpha helix, ect)
    - Tertiary (3^o): global 3D fold
    - Quaternary (4^o): organization of multiple subunits (polymers) in a functional complex

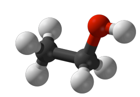


Stoichiometry and Geometry

- Stoichiometry: atomic composition
  - For biological macromolecules: compositions of (monomer) units
  - Relative easy to know
- Geometry: three-dimensional structure
  - More challenging to determine, especially for large, flexible ones
  - Protein folding problem

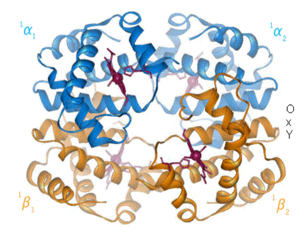
ethanol
a.k.a. alcohol, biofuel

C_2H_6O



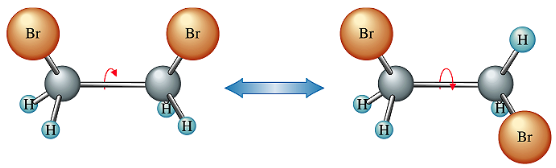
hemoglobin

$\alpha_2\beta_2$

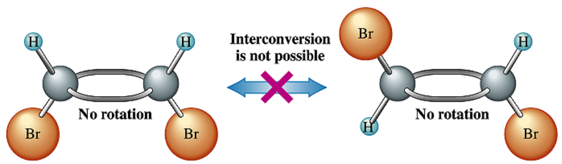


Configuration and Conformation

- Configuration: arrange of atoms around non-rotating bonds or around chiral centers
- Conformation: spatial arrange around rotatable bonds



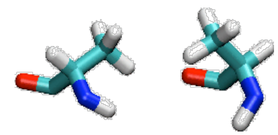
(a) The molecule 1,2-dibromoethane can rapidly interconvert between the two orientations shown, because a carbon–carbon single bond is not structurally rigid.



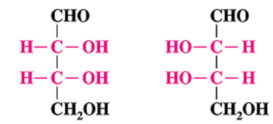
(b) The molecule 1,2-dibromoethene cannot interconvert between the two forms shown, because a carbon–carbon double bond is structurally rigid.

Chirality in Biology

- Most biomacromolecules consist of chiral monomers
- In biology, typically only one isomer is active
  - Virtually all sugars are D isomers
  - Most proteins are made of L amino acids



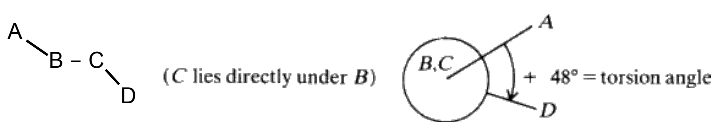
L- and D-alanine



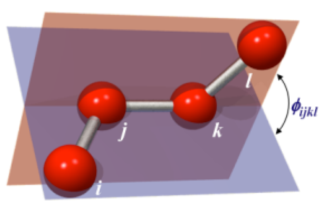
D- and L-erythrose

Conformation/Struture

- Mainly defined by torsion angles
  - Also referred to as dihedral angle: the angle between two groups on either side of a freely rotating chemical bond
  - Defined by four consecutively linked atoms: ϕ(A-B-C-D)



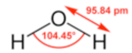
  - Right-hand rule: positive sense is clockwise



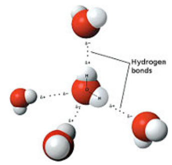
Determinant of Structure

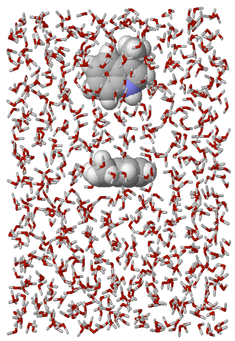
- Probability of observing a particular structure (conformation) is determined by its stability (as defined by the free energy)
  - Thermodynamics and statistical mechanics!
- The stability depends on a range of factors
  - Intramolecular interactions
    - Bonded: chemical bonds, angles, dihedrals etc
    - Nonbonded: "weak" interactions
      - Charged-charged, van der Waals (dispersion and repulsion)
  - Intermolecular interactions: nonbonded/weak interactions
    - Cellular environment: solvent (water), membrane, salt, pH etc
    - Association with other biomolecules, small molecules, ions, etc
- No single structure is *the* structure
  - It is all about probability (statistical mechanics!)
  - Motions and flexibility are important too

Water

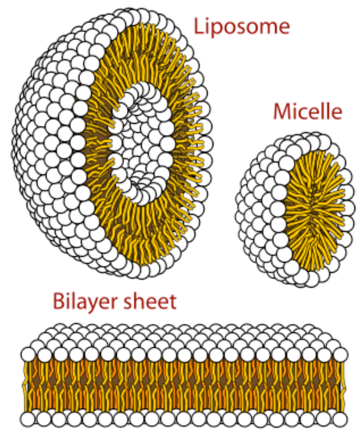


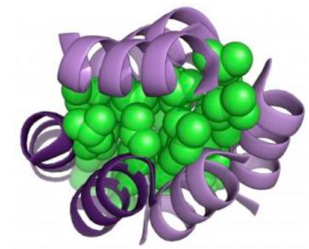
- Solvent of life
- Many unique properties
  - Maximum density at 4 ºC
    - Ice is lighter than liquid water
  - Polar molecule
    - hydrogen bonding network
    - High specific heat capacity
- Hydrophobicity and hydrophilicity
  - Solute polarity (carry partial charges or not)
  - Salts (e.g. NaCl) dissolve in water readily
  - Hydrocarbons (oil) do not mix with water
- Amphipathic molecules
  - Self-assembly to micelles, biological membranes



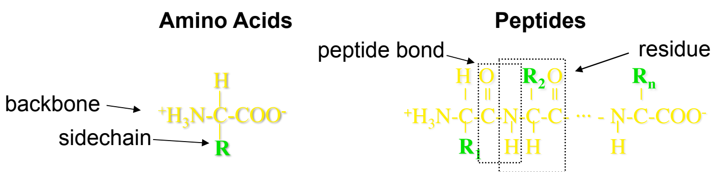


Biological Self Assembly

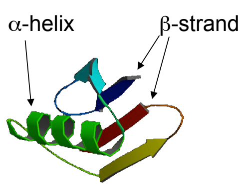




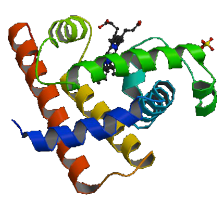
Hierarchical Organization of Proteins



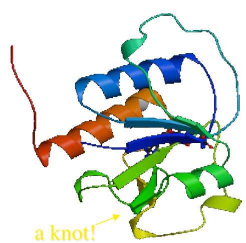
Primary Sequence: MTYKLILNGK TLKGETTTEA VDAATAEKVF KQYANDNGVD GEWTYDDATK TFTVTE



Protein G B1 (3gb1)



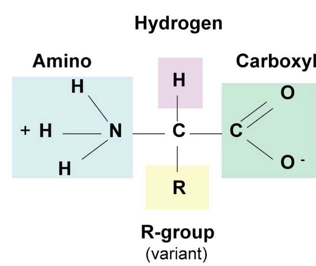
Myoglobin (1mbc)

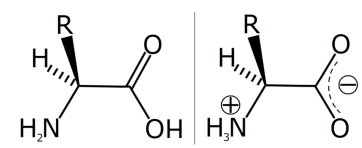


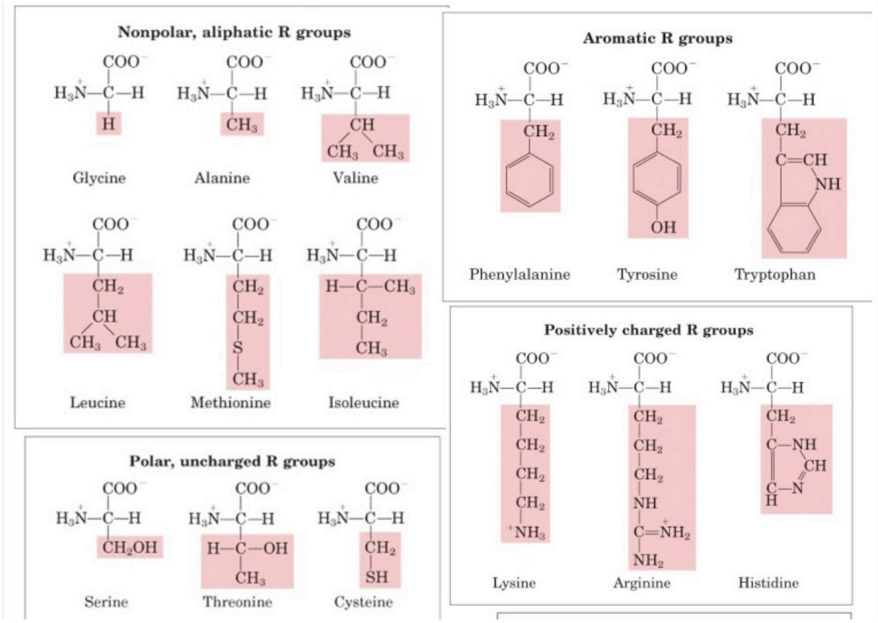
tRNA transferase (1mxi)

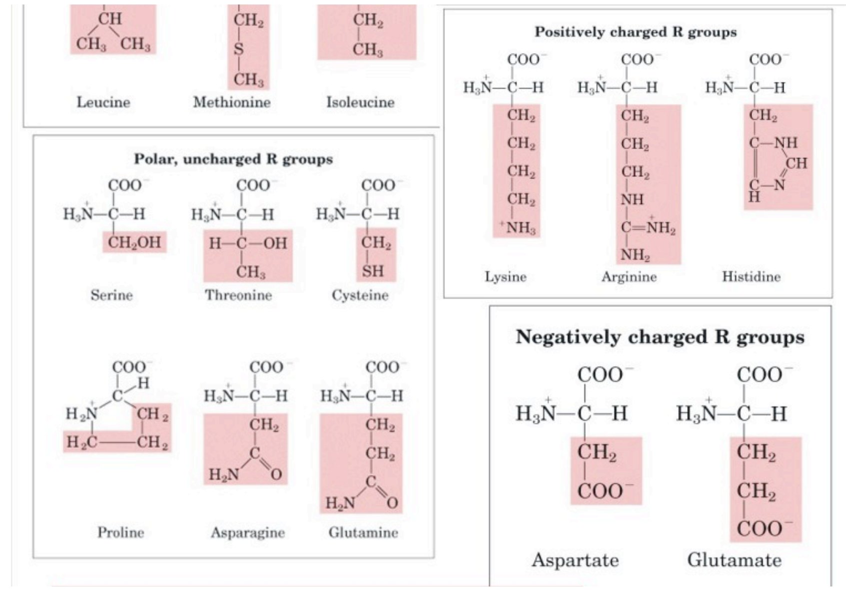
Natural Amino Acids

- All α-amino acids in L-configuration
- Side chains vary
- Essential vs non-essential ones
- Co-exist in two forms
  - Ionic (zwitterionic form) and unionized







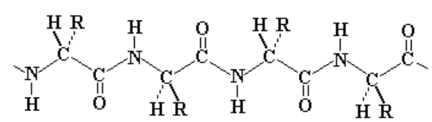


Three-Letter and Single-Letter Codes

Amino Acid	3-Letter	1-Letter	Amino Acid	3-Letter	1-Letter
Alanine	Ala	A	Leucine	Leu	L
Arginine	Arg	R	Lysine	Lys	K
Asparagine	Asn	N	Methionine	Met	M
Aspartate	Asp	D	Phenylalanine	Phe	F
Cysteine	Cys	C	Proline	Pro	P
Histidine	His	H	Serine	Ser	S
Isoleucine	Ile	I	Threonine	Thr	T
Glutamine	Gln	Q	Tryptophan	Trp	W
Glutamate	Glu	E	Tyrosine	Tyr	Y
Glycine	Gly	G	Valine	Val	V

Foldable Sequence Space is Limited

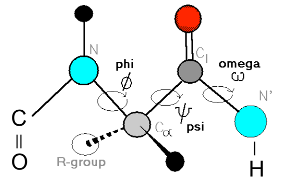
- Huge number of possible sequences: 20^N

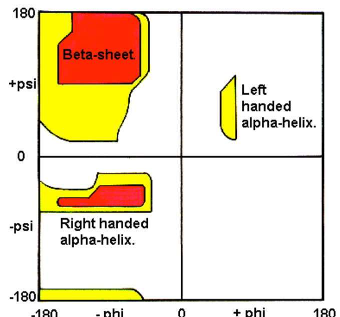


- Only a small subset are explored in biology and most of them give rise to specific 3D structures under the proper conditions
  - Sequence determines structure (or lack or structure)
  - Protein structure-function paradigm: enzymes
  - Some sequences are “intrinsically disordered”! They play particularly important roles in signaling and regulation.

Important Torsions of Proteins

- Backbone torsions
  - ϕ: C(O)-N-Cα-C(O))
  - ψ: N-Cα-C(O)-N
  - ω: H-N-C(O)-O
    - Mostly 180° (trans)
- Ramachandran plot
  - 2D plot of (ϕ,ψ) distribution
  - Allow assignment of secondary structures
  - Statistics of known protein structures reveal common regions of (ϕ,ψ) distributions
  - An important way of structural evaluation

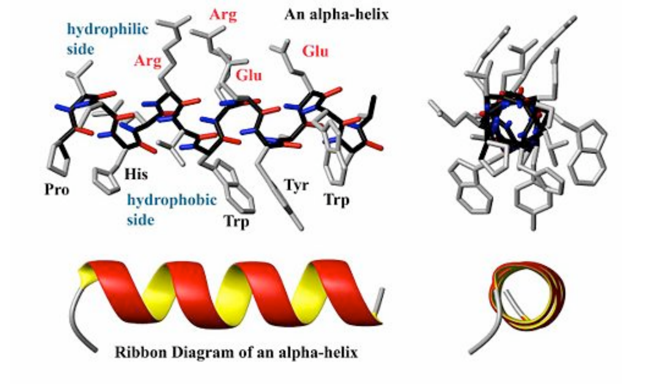




http://www.cryst.bbk.ac.uk/PPS2/course/section3/rama.html

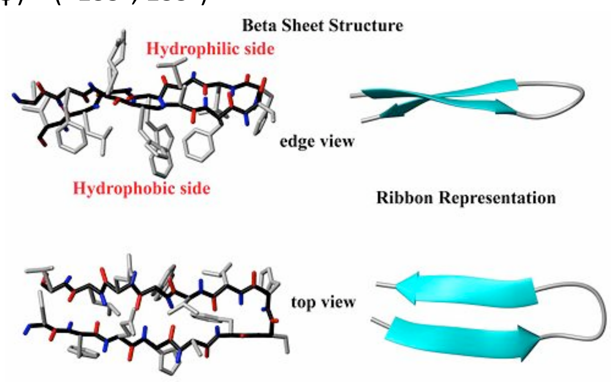
Secondary Structures of Proteins

- α-helix: (i,i+4) hydrogen bonds, 3.6 residues/5.4 Å per turn, (ϕ,ψ) ~ (-55°, -50°)

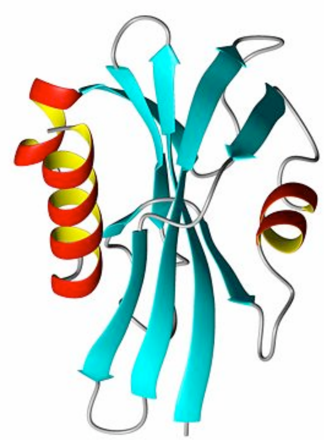


Secondary Structures of Proteins

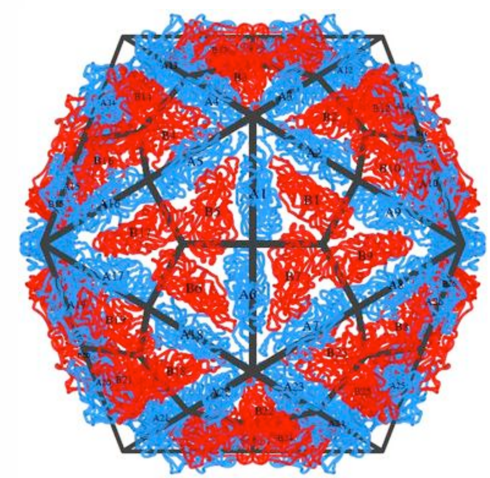
- β-strands: trans- backbone torsions, multiple hydrogen bonds, (ϕ, ψ) ~ (–135°, 135°)



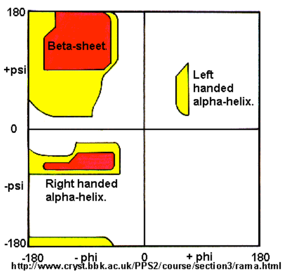
Tertiary and Quaternary Structures

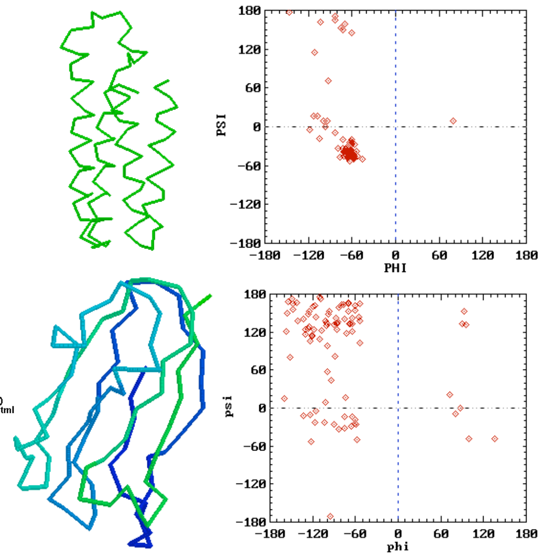


p13 protein from GB virus-B



Structure of the L-A virus Capsid

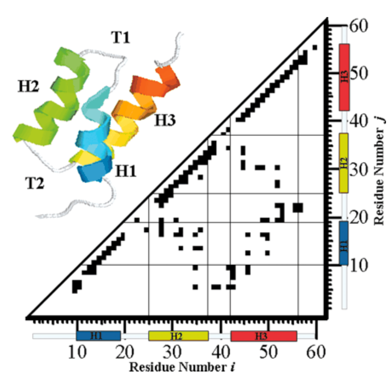


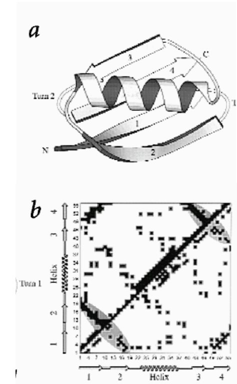


http://www.bmb.uga.edu/wampler/tutorial/prot2.html

Protein Contact Map

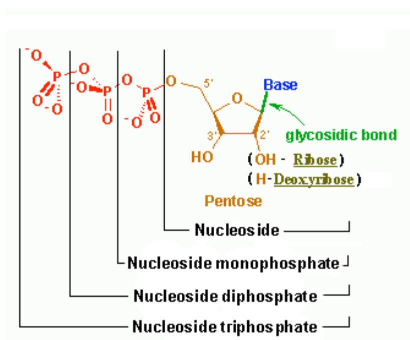
- A simple way of representing protein folds



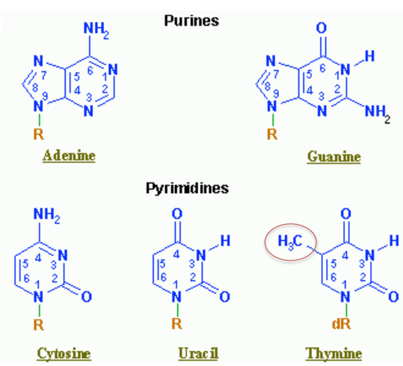


Nuclei Acids

- Polymer of nucleotides: highly flexible (compared to peptides)
- Deoxyribonucleic acid (DNA): A/G/C/T
- Ribonuclei acid (RNA): A/G/C/U



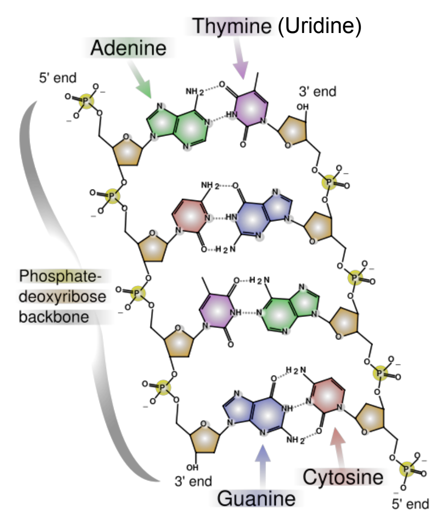
nucleotide



nucleobase

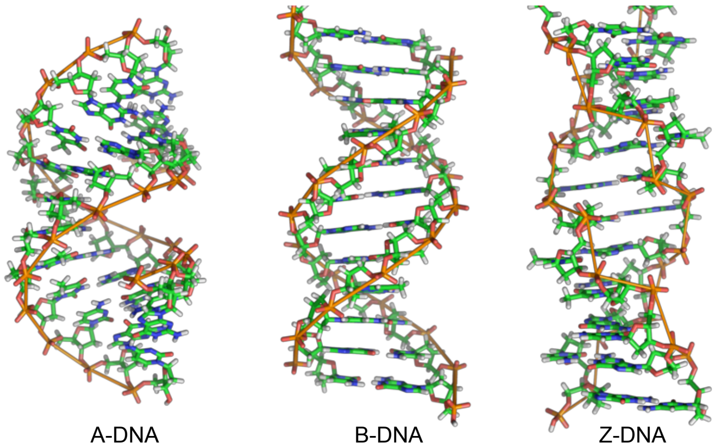
Watson-Crick Base Pairs

- G-C and A-T(U)
- DNA almost exclusively exist as duplex held together by forming Watson-Crick base pairs
- A robust mechanism for faithful replication and translation
- Other pairings possible but mostly in RNA to give rise to non-trivial structures



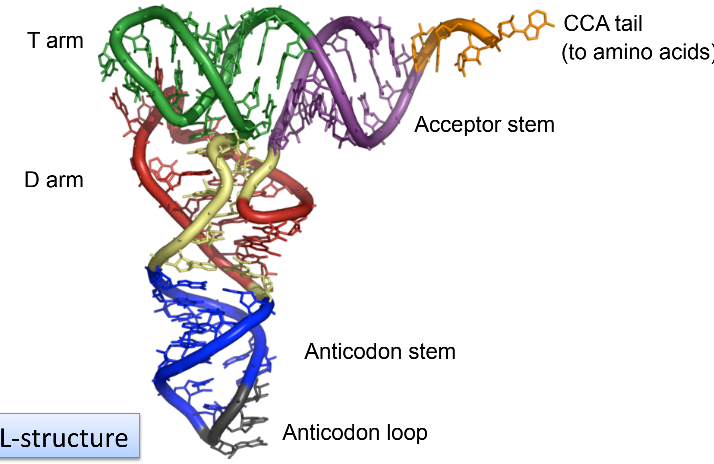
Helical Structures of DNA (RNA)

- Famous DNA double helix: stacking of base pairs, minor/major groves, polymorphic (hydration dependent)



Tertiary Structure of RNA

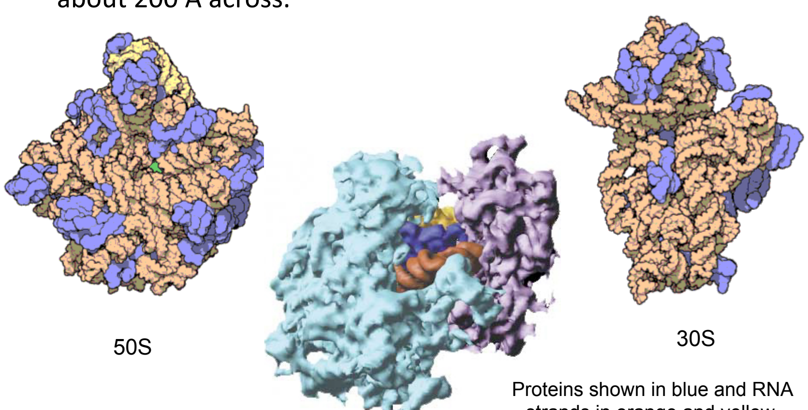
- Single-strand RNA often fold into tertiary structures



tRNA L-structure

Ribosome: the Protein Factory

- Prokaryotic ribosome consists of 42 proteins (~35% weight) and 3 ribosomal RNAs (rRNA) (~65% weight) and measures about 200 Å across.

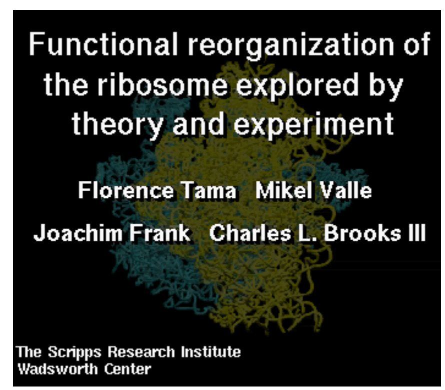


Proteins shown in blue and RNA strands in orange and yellow.

Functional Motions of Ribosome

"This movie depicts a ratchet-like rearrangement of the 70S ribosome. The rotation of the 30S ribosomal subunit relative to the 50S subunit shows high correspondence to motion captured in cryo-EM maps of the ribosome and postulated to be a key mechanical step in the translocation of the mRNA•tRNAs complex. "

http://brooks.chem.lsa.umich.edu/



Tama et al. Proc Nat Acad Sci, **100** 9319 (2003).