CY60004 Biophysical Chemistry

Structures of biological macromolecules (proteins and polynucleic acids);

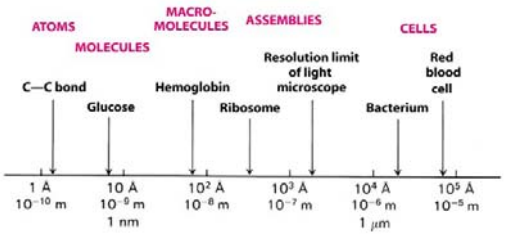
Molecular Mechanics: simulating macromolecular structure;

Spectroscopic (NMR, Fluorescence and Circular Dichroism) methods to study structure of proteins and DNA; solving macromolecular structures by X-ray diffraction; structural transitions in polypeptides, proteins and polynucleic acids.

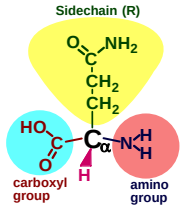
Interactions between macromolecules: thermodynamics of protein folding/stability by fluorescence and circular dichroism techniques.

Binding of small ligands by biological macromolecules: kinetics and energetics of protein-drug, protein-surfactant and DNA-drug interactions by fluorescence, CD and calorimetric methods.

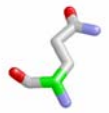
Books :
Biophysical Chemistry, Parts I, II and III, Cantor and Schimmel
Principles of Physical Biochemistry, van Holde, Johnson and Ho



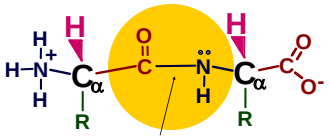
Amino acids



Building blocks

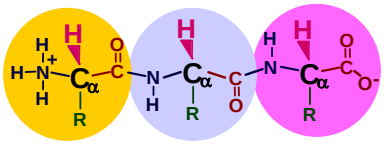


20 common amino acids
R side chain
Three letter code
One letter code

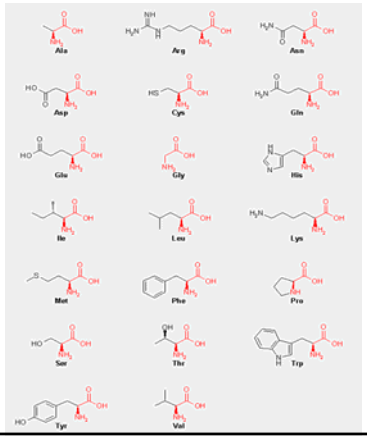


Protein Structure

- Peptide = a short chain of amino acids
- Polypeptide = a longer chain of amino acids
- Protein = a polypeptide that occurs in nature and folds into a defined three-dimensional structure



The beginning of the protein is known as the amino-terminus and the end of the protein is known as the carboxyl-terminus.



Amino Acid Characteristics

- Hydrophobicity
- Size
- Charge
- Secondary structure preference
- Aromaticity

Special characteristics:

- bridge forming by cysteines,
- rigidity of prolines,
- titrating at physiological pH of histidine,
- flexibility of glycines

table 5–1

Properties and Conventions Associated with the Standard Amino Acids

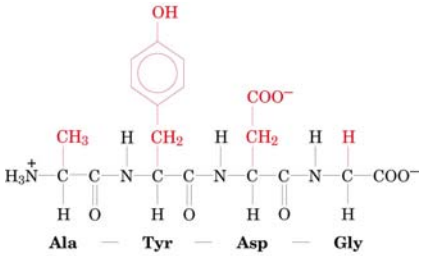
Amino acid	Abbreviated names		M_r	pK_1 (—COOH)	pK_2 (—NH_3^+)	pK_R (R group)	pI	Hydropathy index*	Occurrence in proteins (%)†
Nonpolar, aliphatic R groups									
Glycine	Gly	G	75	2.34	9.60		5.97	−0.4	7.2
Alanine	Ala	A	89	2.34	9.69		6.01	1.8	7.8
Valine	Val	V	117	2.32	9.62		5.97	4.2	6.6
Leucine	Leu	L	131	2.36	9.60		5.98	3.8	9.1
Isoleucine	Ile	I	131	2.36	9.68		6.02	4.5	5.3
Methionine	Met	M	149	2.28	9.21		5.74	1.9	2.3
Aromatic R groups									
Phenylalanine	Phe	F	165	1.83	9.13		5.48	2.8	3.9
Tyrosine	Tyr	Y	181	2.20	9.11	10.07	5.66	−1.3	3.2
Tryptophan	Trp	W	204	2.38	9.39		5.89	−0.9	1.4
Polar, uncharged R groups									
Serine	Ser	S	105	2.21	9.15		5.68	−0.8	6.8
Proline	Pro	P	115	1.99	10.96		6.48	1.6	5.2
Threonine	Thr	T	119	2.11	9.62		5.87	−0.7	5.9
Cysteine	Cys	C	121	1.96	10.28	8.18	5.07	2.5	1.9
Asparagine	Asn	N	132	2.02	8.80		5.41	−3.5	4.3
Glutamine	Gln	Q	146	2.17	9.13		5.65	−3.5	4.2
Positively charged R groups									
Lysine	Lys	K	146	2.18	8.95	10.53	9.74	−3.9	5.9
Histidine	His	H	155	1.82	9.17	6.00	7.59	−3.2	2.3
Arginine	Arg	R	174	2.17	9.04	12.48	10.76	−4.5	5.1
Negatively charged R groups									
Aspartate	Asp	D	133	1.88	9.60	3.65	2.77	−3.5	5.3
Glutamate	Glu	E	147	2.19	9.67	4.25	3.22	−3.5	6.3

*A scale combining hydrophobicity and hydrophilicity of R groups; it can be used to measure the tendency of an amino acid to seek an aqueous environment (− values) or a hydrophobic environment (+ values). See Chapter 12. From Kyte, J. & Doolittle, R.F. (1982) *J. Mol. Biol.* **157**, 105 – 132.

†Average occurrence in over 1150 proteins. From Doolittle, R.F. (1989) Redundancies in protein sequences. In *Prediction of Protein Structure and the Principles of Protein Conformation* (Fasman, G.D., ed) Plenum Press, NY, pp. 599–623.

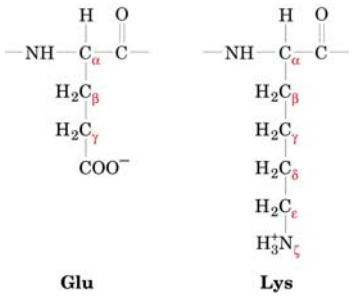
Nomenclature

- **Glx** can be Glu or Gln
- **Asx** can be Asp or Asn
- Polypeptide chains are always described from the N-terminus to the C-terminus



Nomenclature

- **Nonhydrogen** atoms of the amino acid side chain are named in sequence with the Greek alphabet



ISOELECTRIC POINT (pI)

Definition: the pH at which a molecule carries no net electric charge

: using the Henderson-Hasselbalch equation

$pI = \frac{1}{2}(pK_1 + pK_2)$

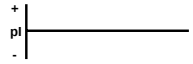
: for amino acids

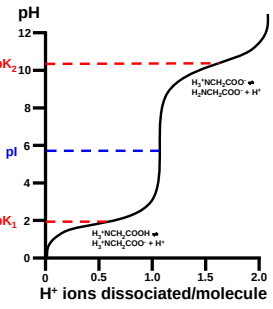
pK_1 = pK α-COOH

pK_2 = pK α-NH_3

Isoelectric Focusing

: separation of proteins according to charge
: pH gradient set up by ampholytes
: electric field applied such that one pole is positively charged and one pole negatively charged
: proteins migrate in the pH gradient until their net charge = 0 (isoelectric point)





Titration of glycine

Ionization of Amino Acids

Equilibrium dissociation constant $K_a = \frac{[H^+][A^-]}{[HA]}$

Henderson-Hasselbach Equation

$$-\log[H^+] = -\log K_a + \log\frac{[A^-]}{[HA]}$$

$$pH = pK_a + \log\frac{[A^-]}{[HA]}$$

$$pH = pK_a \quad @\ [A^-] = [HA]$$

Acid-base properties of amino acids

The dissociation of first proton from the α-carboxyl group is

$Gly^+ + H_2O \longleftrightarrow Gly^0 + H_3O^+$

The dissociation of the second proton from the α-amino group

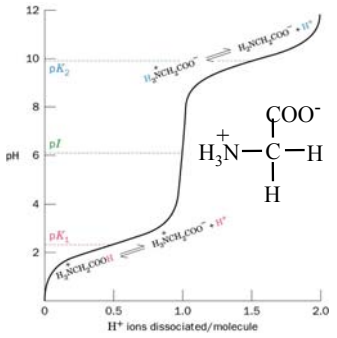
$Gly^0 + H_2O \longleftrightarrow Gly^- + H_3O^+$

$$K_1 = \frac{[Gly^0][H_3O^+]}{[Gly^+]} \qquad K_2 = \frac{[Gly^-][H_3O^+]}{[Gly^0]}$$

The pK_a's of these two groups are far enough apart that they can be approximated by Henderson-Hasselbalch

$$pH = pK_1 + \log\frac{[Gly^0]}{[Gly^+]} \qquad pH = pK_2 + \log\frac{[Gly^-]}{[Gly^0]}$$

Titration curve of glycine



Isoelectric Point of Amino Acids

K_{a1}

$HAH^+ \leftrightarrow HA^0 + H^+$

$pI = pK_{a1} + \log\frac{[HA^0]}{[HAH^+]}$

$pI = pK_{a2} + \log\frac{[A^-]}{[HA^0]}$

K_{a2}

$HA^0 \leftrightarrow A^- + H^+$

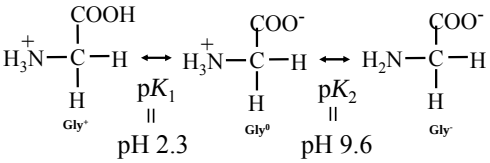
$2pI = pK_{a1} + pK_{a2} + \log\frac{[HA^0]}{[HAH^+]} + \log\frac{[A^-]}{[HA^0]}$

$2pI = pK_{a1} + pK_{a2} + \log\frac{[HA^0]}{[HAH^+]}\frac{[A^-]}{[HA^0]}$

$2pI = pK_{a1} + pK_{a2} + \log\frac{[A^-]}{[HAH^+]}$; @pI, $[A^-] = [HAH^+]$

$pI = \frac{pK_{a1} + pK_{a2}}{2}$

Titration of Gly



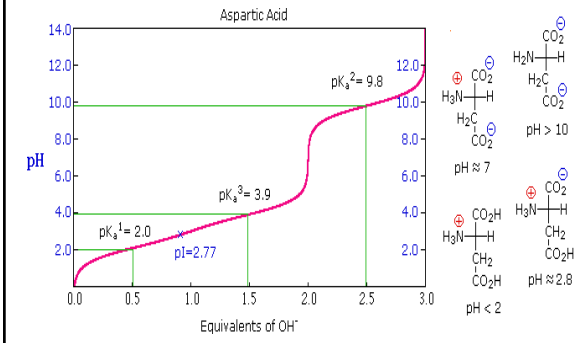
From the p*K* values we can calculate the pI (isoelectric point) where the amino acid is neutral.

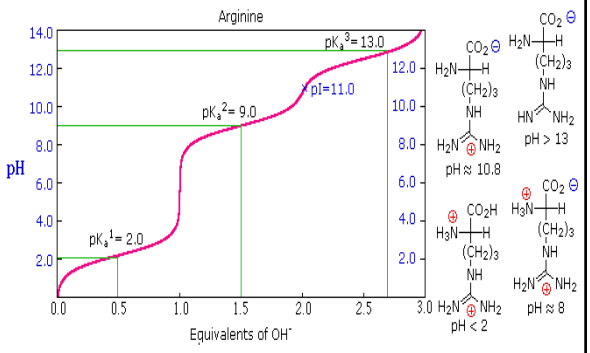
pI ≈ average of (p*K* *below* neutral+ p*K* *above* neutral)

So, for Gly,

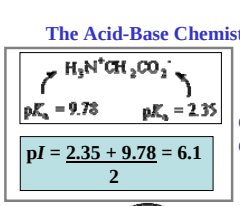
$pI = (pK_1 + pK_2)/2$

$= (2.3 + 9.6)/2 \approx 6$





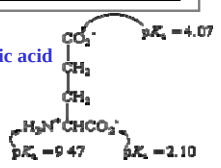
The Acid-Base Chemistry of the Amino Acids



Glycine
Gly

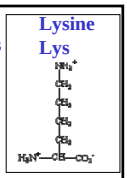
pI = 2.35 + 9.78 = 6.1
2

Glutamic acid
Glu



pI = 2.10 + 4.07 = ~3.1
2

Lysine
Lys



pI = 9.18 + 10.79 = ~10.0
2

Henderson-Hasselbach Equation

$$-\log[H^+] = -\log K_a + \log\frac{[A^-]}{[HA]}$$

$$pH = pK_a + \log\frac{[A^-]}{[HA]}$$

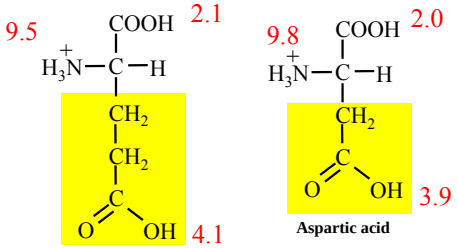
Buffering zone = pK +/- 1 pH unit

At pH = pK, [HA] = [A-]

At pH below pK, [HA] > [A-]

At pH above pK, [HA] < [A-]

Charged polar (acidic) side chains



Glutamic acid

Glu

E

Aspartic acid

Asp

D

The 5 Complex Amino Acids are:

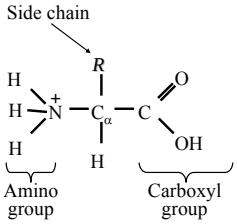
Glutamic acid (Glu)

Aspartic acid (Asp)

Lysine (Lys)

Arginine (Arg)

Histidine (His).



Each of these 5 amino acids has 3 ionizable groups ⇒ 3 pKs.

pI = average of (pK-below) + (pK-above)

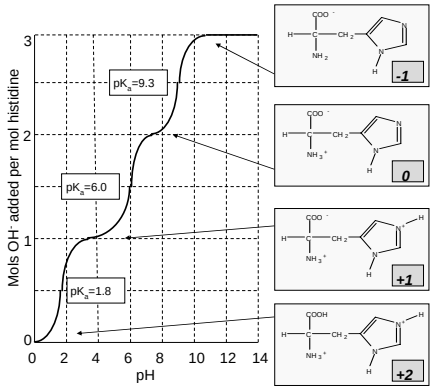
where pK-below is the pK between AA+1 and AA0

and pK-above is the pK between AA0 and AA-1.

- **His has 3 ionizing groups, alpha-carboxylic acid (pK 1.8),**
- **side-chain amino (pK 6.0) and alpha-amino (pK 9.2):**
- **His+2 goes to His+1 via pK 1.8;**
- **His+1 goes to His0 via pK 6.0;**
- **His0 goes to His-1 via pK 9.2.**
- **Therefore, pI = (6.0 + 9.2)/2 = 7.6.**

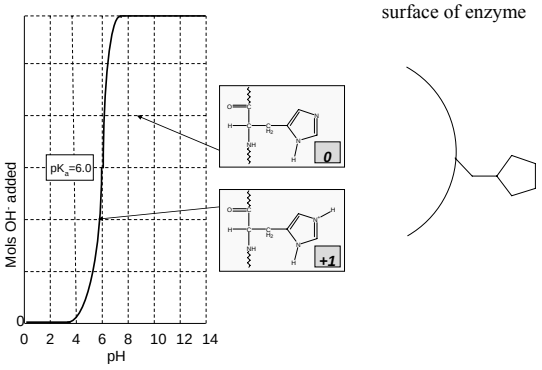
Histidine Titration

Free Amino acid

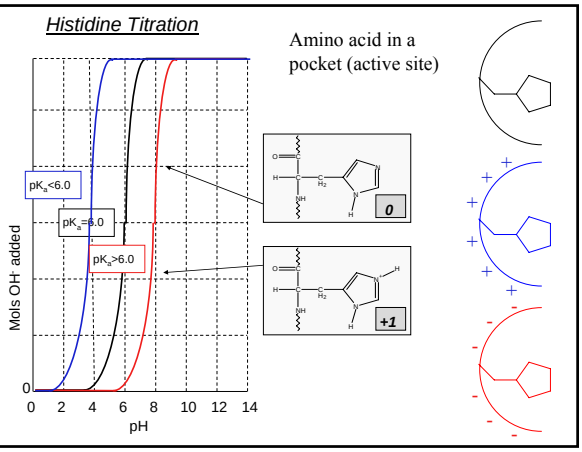


Histidine Titration

Amino acid on surface of enzyme



Histidine Titration



1. Carboxylic acids ionize at acidic pH; ie. Carboxylic acids give up their protons at acid pHs.
2. Amino groups ionize at basic pH; ie. Amines give up their protons at basic or alkaline pHs.
3. Carboxylic acids near an amino group have a more acidic pK than isolated carboxylic acids.
4. Amino groups near a carboxylic acid have a more acidic pK than isolated amines.

For Ala-Lys, there are 3 ionizable groups:

1) alpha-amino group contributed by Ala - assign pK 9.9.

2) alpha-carboxylate group from Lys - assign pK 2.2.

3) side chain amino group from Lys - assign pK 10.8.

	pH 1	pH 5	pH 7	pH 10	pH 12
α−amino					
α-carboxylate					
Side chain amino					
Net Charge					

For the tetrapeptide: **Glu-Ala-Lys-Tyr**

Write out the structure.

With the assigned pK values, determine the net charge at pH 1, 3, 5, 7, 10, 11.

Calculate the pI of this tetrapeptide.

The pK values of the amino acids are:

Ala- 2.4, 9.9

Glu - 2.1, 4.1, 9.5

Lys - 2.2, 9.2, 10.8

Tyr - 2.2, 9.7

Peptide Bond - Structure

1) DOUBLE BOND CHARACTER

: fractional charges
: no rotation between C and N
: HIGH DEGREE OF ROTATION RESTRICTION

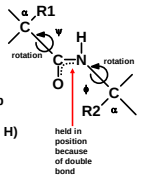
2) COPLANAR NATURE

: αC-N can rotate through angle ϕ
: αC-C can rotate through angle ψ
: R groupings can rotate through a characteristic angle
: as it rotates, gives 3D structure to the protein
: degree of rotation will depend upon R group

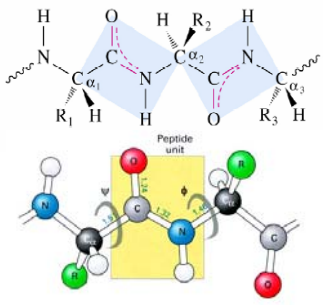
1) BULK
eg. free rotation with glycine (R = H)

2) CHARGED STATE
two positives repel one another

3) WATER REACTION
polarity of R groupings
polar - rotate out to H bond with water
nonpolar - on inside of molecule

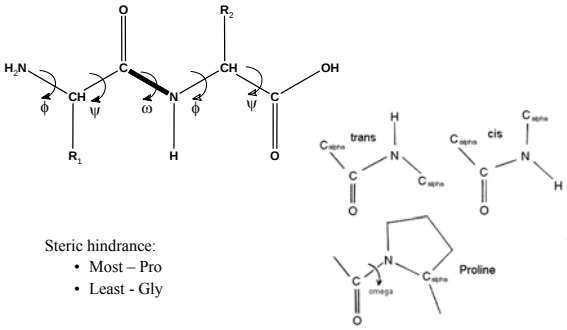


- **Rotation around C-N bond is restricted due to the double-bond nature of the resonance hybrid form**
- **Peptide groups (blue planes) are therefore planar**



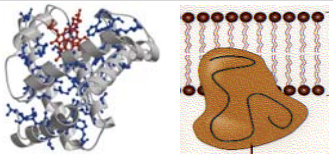
Torsion angles

phi (φ) torsion angle around N-CA bond
psi (ψ) torsion angle around CA-C bond
omega (ω) peptide bond; trans ~180 degrees, *cis* ~0



Steric hindrance:
- Most – Pro
- Least - Gly

Residue	Globular protein	Membrane protein
Non-polar V L I M F Y W	In interior Hydrophobic core	Surface – lipid anchor
Polar charged R K D E H	Surface Catalytic sites	Hydrophilic core
Polar neutral S T N Q Y W	H bond network	Inside surface – part of channel



Amino Acid Partitioning Into Membrane Regions

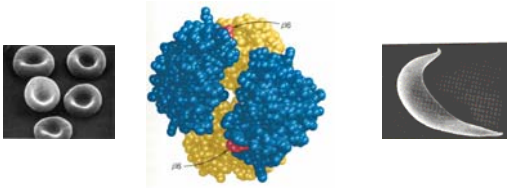
Region	Amino Acids
Bulk water	Arg, Asn, Asp, Gln, Glu, His, Lys, Pro
Bulk water + interfacial	Ala, Cys, Gly, Ser, Thr
Interfacial	Tyr
Hydrophobic	Ile, Leu, Met, Phe, Trp, Val

http://chemed.chem.purdue.edu/genchem/topicreview/bp/1biochem/amino2.html

Importance of Protein Structure

Sequestering hydrophobic residues in the protein core protects proteins from hydrophobic agglutination.

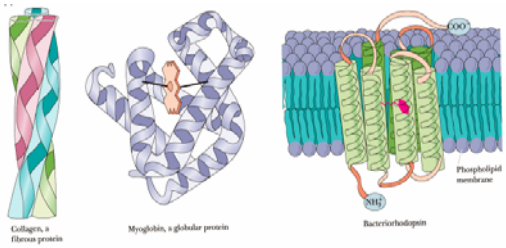
Hemoglobin A: Val-His-Leu-Thr-Pro-Glu-Glu-Lys-
Hemoglobin S: Val-His-Leu-Thr-Pro-Val-Glu-Lys-



"sticky patch" causes hemoglobin S to agglutinate (stick together) and form fibers which deform the red blood cell

Classes of Proteins

Based on structure and solubility,
proteins can be grouped into three large classes:

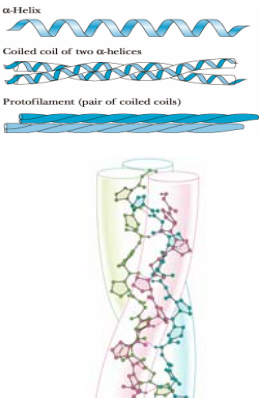


Fibrous Globular Membrane

Fibrous Proteins

- Fibrous proteins contain polypeptide chains organized parallel along a single axis, producing long fibers or large sheets.
- They are mechanically strong, play structural roles in nature;
- Difficult to dissolve in water;

α-Keratins and Collagen are examples of fibrous proteins

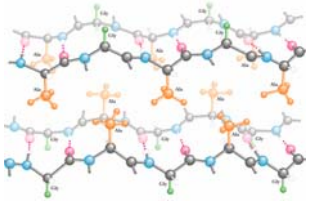


α-keratins are found in hair, fingernails, claws, horns and beaks;

•Sequence consists of long alpha helical rod segments capped with non-helical N- and C-termini

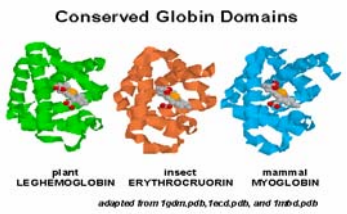
β-keratins are found in silk and consist of gly-ala repeat sequences;

•Ala is small and can be packed within the sheets

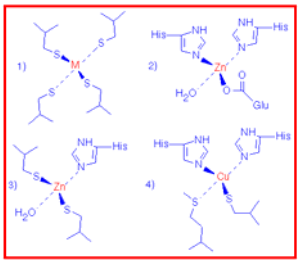


Globular Proteins

- Globular proteins are classified according to the type and arrangement of secondary structure
  - Antiparallel alpha helix proteins
  - Parallel or mixed beta sheet proteins
  - Antiparallel beta sheet proteins



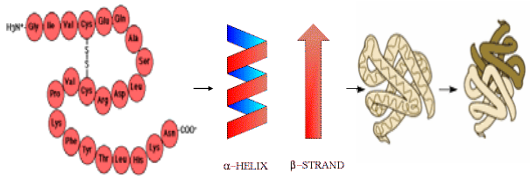
Metal binding sites examples



1) M is Fe (rubredoxin) or Zn (aspartate transcarbamoylase)
2) Carboxypeptidase A
3) Catalytic ion in liver alcohol dehydrogenase
4) Azurin and plastocyanin

Protein Architecture

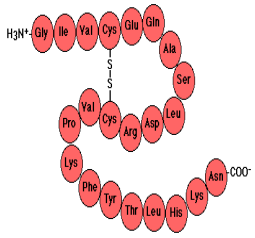
Primary structure (1°) : the amino acid sequence.
Secondary structure (2°) : helices, sheets and turns.
Tertiary structure (3°) : side chain packing in the 3-D structure.
Quaternary structure (4°) : association of subunits.



Protein Structure

- Physical properties of protein that influence stability & therefore, determine its fold:
  - Rigidity of backbone
  - Amino acid interaction with water
    - Hydropathy index for side chains
  - Interactions among amino acids
    - Electrostatic interactions
    - Hydrogen bonds
    - S-S bonds
    - Volume constraints

Primary Structure

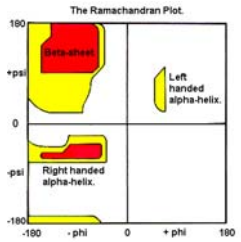


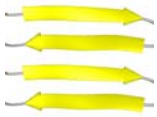
Secondary Structure

- Two major types:
  - Alpha Helical Regions
  - Beta Sheet Regions
- Other classification schemes:
  - Turns
  - Transmembrane regions
  - Internal regions
  - External regions
  - Antigenic regions

Secondary Structure Conformations		
	ϕ	ψ
alpha helix	-57	-47
alpha-L	57	47
3-10 helix	-49	-26
π helix	-57	-80
type II turn	-79	150
β-sheet parallel	-119	113
β-sheet antiparallel	-139	135

Residue conformational preferences:
Glu, Ala, Leu, Met, Gln, Lys, Arg - helix
Val, Ile, Tyr, Cys, Trp, Phe, Thr - strand
Gly, Asn, Pro, Ser, Asp - turn





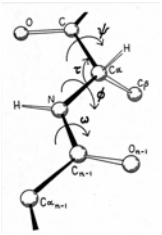
Features

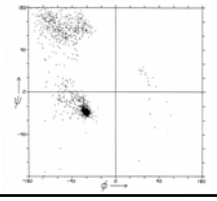
α-helix is one of two secondary structures (the other being the β -sheet) predicted and discovered by Linus Pauling in 1951.

It is a right-handed helix with the following spatial parameters:
ϕ= -57°
ψ = -47°
n = 3.6 (number of residues per turn)
pitch 0.54nm (or 5.4Å)

β-sheet: parallel (ϕ =-119 and ψ =113) or βanti-parallel pleated sheet structures

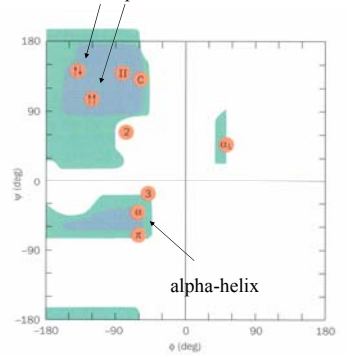
other regular structures
a) 3_{10} helix - i, i+3 hydrogen bonding pattern
b) π-helix - i, i+5 hydrogen bonding pattern
c) turns and loops





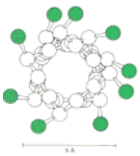
antiparallel beta-sheet parallel beta-sheet

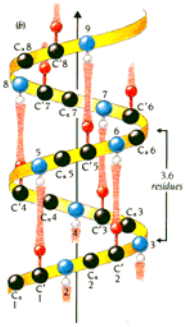
All the common secondary structure conformations fit in the sterically allowed regions of phi-psi space



Properties of the alpha helix

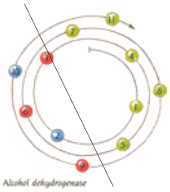
- ϕ ≈ ψ ≈ –60°
- Hydrogen bonds between C=O of residue *n*, and NH of residue *n*+4
- 3.6 residues/turn
- 1.5 Å/residue rise
- 100°/residue turn

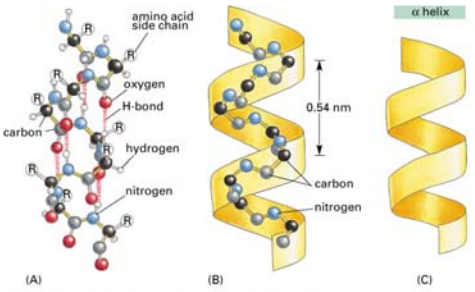




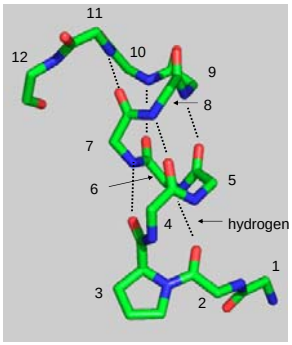
Properties of α-helices

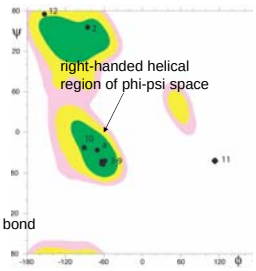
- 4 – 40+ residues in length
- Often *amphipathic* or "dual-natured"
  - Half hydrophobic and half hydrophilic
  - Mostly when surface-exposed
- If we examine many α-helices, we find trends…
  - Helix formers: Ala, Glu, Leu, Met
  - Helix breakers: Pro, Gly, Tyr, Ser





the alpha-helix: repeating i, i+4 H-bonds





Which of residues 1-12 are in the helix? Does this coincide with the residues in the helical region of phi-psi space?

Helix nomenclature: α-helix example

1 Hydrogen bond between

C=O ---------- H-N

(residue i) (residue i+4)

1←5 helix

2 Repeating unit:

- 5 turns
- 18 residues per repeat

18_5 helix

3 Loop formed between

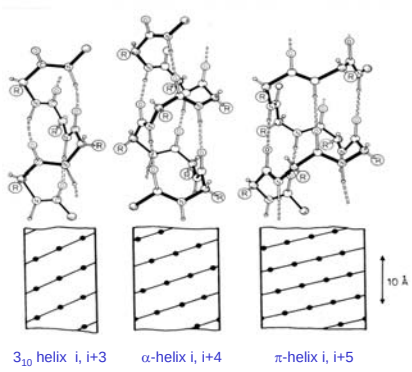
- 13 atoms
- 3.6 residues per turn

C=O H-N

1 ← 5

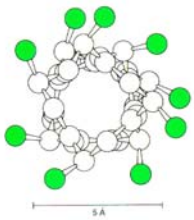
3.6_{13} helix ← nomenclature used for 3_{10} helix

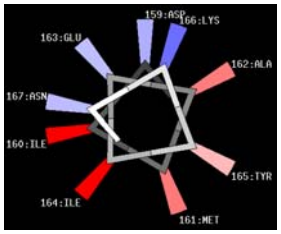
α helices extend with approximately 1.5 Angstrom per residue, 5.4 Angstrom per turn.



3_{10} helix i, i+3 α-helix i, i+4 π-helix i, i+5

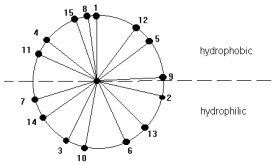
Helical Wheel



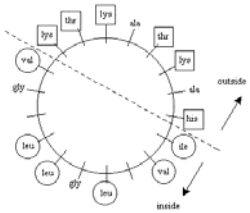


Helical Wheel

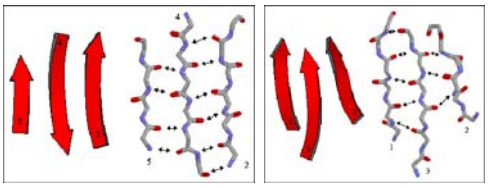
For each residue the rotation is 100^0



lys his gly val thr val leu thr ala leu gly ala ile leu lys lys

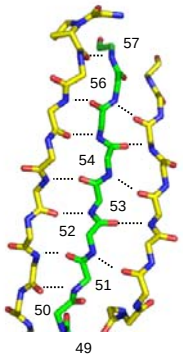


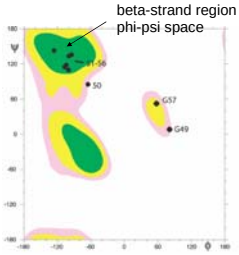
Beta Strands



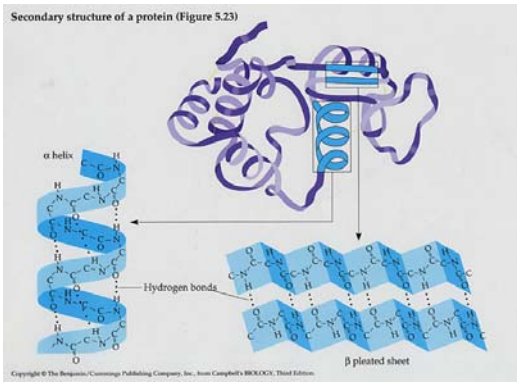
	phi(deg)	*psi*(deg)	*omega* (deg)
β strand	-120	120	180

beta strands/sheets



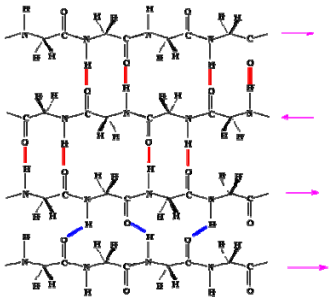


parallel or anti-parallel sheet?



Secondary Structure: Beta Sheets

BETA PLEATED SHEET: a result of H-bonding between polypeptide chains

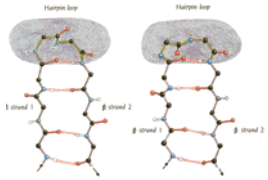


Four-stranded beta sheet that contains three antiparallel and one parallel strand.

Hydrogen bonds are indicated with red lines (antiparallel strands) and blue lines (parallel strands)

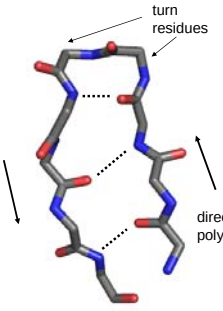
Turns and Loops

- Secondary structure elements are connected by regions of *turns* and *loops*
- Turns – short regions of non-α, non-β conformation



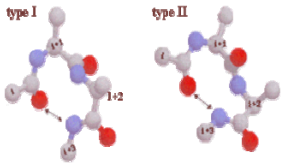
- Loops – larger stretches with no secondary structure. Often disordered.
  - "Random coil"
  - Sequences vary much more than secondary structure regions

β-turns in proteins: reversing the chain direction



A β-turn consists of two residues, where there is a hydrogen bond between the carbonyl of the residue preceding the turn and the amide nitrogen following the turn. There are a number of ways to configure the backbone to achieve this

Secondary Structure: Turns



Gamma-turn
It involves three amino acid residues and the intraturn hydrogen bond for a gamma-turn is formed between the backbone CO(i) and the backbone NH(i+2).

Beta-turn
A beta-turn involves four amino acid residues and may or may not be stabilized by the intraturn hydrogen bond between the backbone CO(i) and the backbone NH(i+3).

Alpha-turn
An alpha-turn involves five amino acid residues where the distance between the Calpha(i) and the Calpha(i+4) is less than 7Å and the pentapeptide chain is not in a helical conformation.

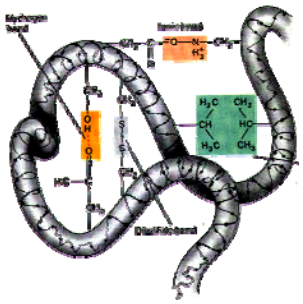
Pi-turn
It is the largest tight turn which involves six amino acid residues.

Tertiary Structure: Aggregation of individual protein.

1. Hydrophobic attraction: the close association attraction of hydrocarbon side-chains.
2. Ionic bond: between positively charged groups and negatively charged groups.
3. Hydrogen bonds
4. Disulfide bonds

A protein has size and shape as well as unique arrangement through hydrogen, ionic, hydrophobic and disulfide bonds.

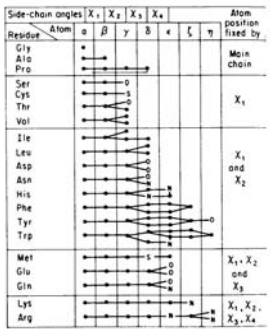
Tertiary Structure of Protein



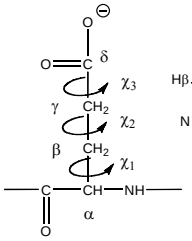
Side chain conformation

• side chains differ in their number of degrees of conformational freedom

•but side chains of very different size can have the same number of χ angles.

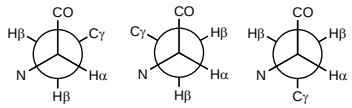


Side chain conformations--canonical staggered forms



glutamate

Newman projections for χ_1 of glutamate:



$\chi_1 = 180°$ $\chi_1 = +60°$ $\chi_1 = -60°$

t *g+* *g–*

t=trans, g=gauche

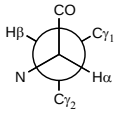
name of conformation

Side chain angles are defined moving outward from the backbone, starting with the N atom: so the χ_1 angle is N–Cα–Cβ–Cγ, the χ_2 angle is Cα–Cβ–Cγ –Cδ ...

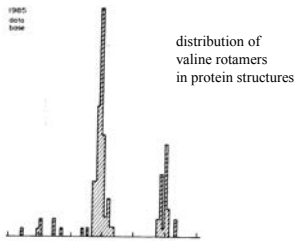
IUPAC nomenclature:
http://www.chem.qmw.ac.uk/iupac/misc/biop.html

Rotamers

- a particular combination of side chain torsional angles χ_1, χ_2, etc. for a particular residue is known as a ***rotamer***.
- for example, for aspartate, if one considers only the *canonical staggered forms*, there are nine (3^2) possible rotamers: *g⁺g⁻*, *g⁺g⁺*, *g-g-*, *g-g⁺*, *tg⁺*, *g⁺t*, *tg-*, *g-t*, *tt*
- not all rotamers are equally likely.
- for example, valine prefers its *t* rotamer

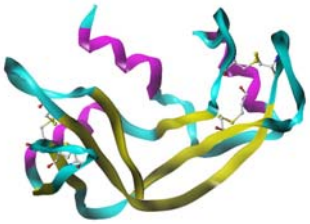


χ_1=180°, *trans or t*



distribution of valine rotamers in protein structures

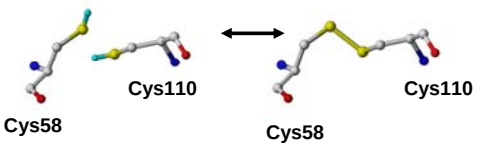
Anfinsen's experiments, late 1950's through 1960's

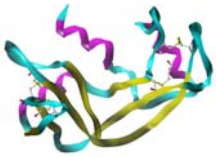


Ribonuclease, an enzyme involved in cleavage of nucleic acids. Structure has a combination of α and β segments and four disulfide bridges

What are Disulfide Bridges?

Cys-SH + Cys-SH ⇌ Cys-S—S-Cys (Oxidation / Reduction)

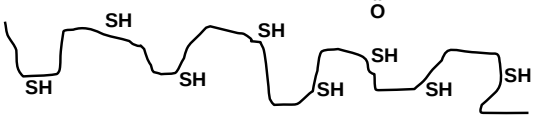




Active, native structure

BME is reducing agent
Urea unfolds proteins

add β-mercaptoethanol (BME)
$HS\text{-}CH_2\text{-}CH_2\text{-}OH$
add urea, $H_2N\text{-}\overset{\overset{O}{\|}}{C}\text{-}NH_2$

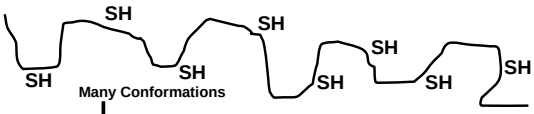


Denatured, inactive, "random coil", many conformations

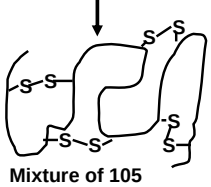
Remove BME
Remove urea

- Native structure
- fully active
- 4 disulfide bond correct

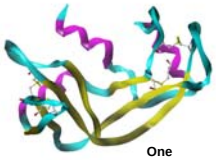
One Conformation



Remove BME



Remove urea



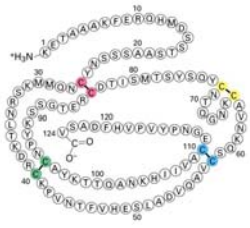
One Conformation

Mixture of 105 different conformations, 1% active

Sequence specifies structure

- Anfinsen's experiments
  - Denatured ribonuclease
  - Spontaneously regained enzymatic activity
  - Evidence that it re-folded to native conformation

The essential structure information is stored in the primary sequence of amino acids

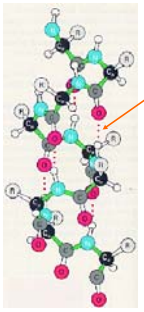


The "native" structure of a protein is the form we find when we isolate that protein in an active state from a natural source.

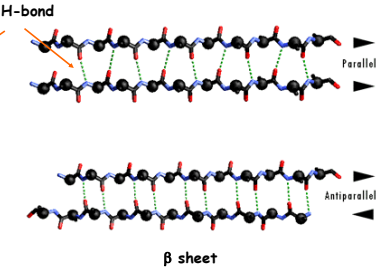
If the protein loses that structure, by unfolding, or unwinding, it loses activity.
Therefore, native = folded
denatured = unfolded

The "native" structure is necessary to create the binding pockets that make up the active site of an enzyme.

Structure Stabilizing Interactions



Higher interaction energy than a simple van der Waals interaction



α helix

β sheet

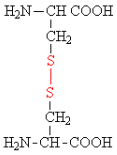
Structure Stabilizing Interactions

- Noncovalent
  - Van der Waals forces (transient, weak electrical attraction of one atom for another)
  - Hydrophobic (clustering of nonpolar groups)
  - Hydrogen bonding

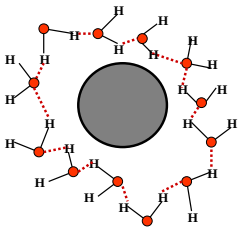
δ- δ+ δ-

D – H A

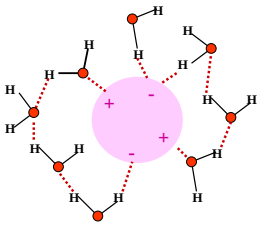
- Covalent
  - Disulfide bonds



The hydrophobic effect



Non-polar solute



Polar solute

The dissolution of a non-polar solute in water is unfavourable – for polar solutes there is favourable hydrogen bonding

Quantification of contribution to the hydrophobic effect (Hydropathy)

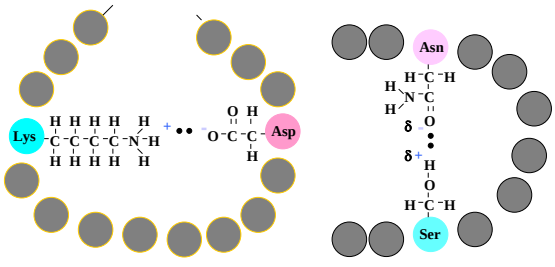
Partitioning of the molecules between water and an organic solvent such as octanol

Amino acid placed in two-solvent system

Hydropathy represented by partition coefficient P, $P = \frac{\chi_{aq}}{\chi_{org}}$

Hydropathy represented by $-\log(P)$

Structure Stabilizing Interactions



Solving Protein Structures

Only 2 kinds of techniques allow one to get atomic resolution pictures of macromolecules

- X-ray Crystallography (first applied in 1961 - Kendrew & Perutz)
- NMR Spectroscopy (first applied in 1983 - Ernst & Wuthrich)

- Structure ⟷ Function
- Structure ⟷ Mechanism
- Structure ⟷ Origins/Evolution
- Structure-based Drug Design
- Solving the Protein Folding Problem

QHTAWCLTSEQHTAAVIWDCETPGKQNGAYQED

HHHHHHCCEEEEEEEEEEECCHHHHHHHCCCCC

Hydropathy plots

•An hydropathy plot is a graphical display of the local hydrophobicity of amino acid side chains in a protein.

•A positive value indicates local hydrophobicity and a negative value suggests a water-exposed region on the face of a protein.

•Hydropathy plots are generally most useful in predicting transmembrane segments and N-terminal secretion signal sequences.

Hydrophobicity scales

	Kyte-Doolittle
Alanine	1.8
Arginine	-4.5
Asparagine	-3.5
Aspartic acid	-3.5
Cysteine	2.5
Glutamine	-3.5
Glutamic acid	-3.5
Glycine	-0.4
Histidine	-3.2
Isoleucine	4.5
Leucine	3.8
Lysine	-3.9
Methionine	1.9
Phenylalanine	2.8
Proline	-1.6
Serine	-0.8
Threonine	-0.7
Tryptophan	-0.9
Tyrosine	-1.3
Valine	4.2

Hydropathy index

A positive value indicates a hydrophobic residue and a negative value a hydrophilic residue

Hydropathy plots

Calculate property for first sub-sequence

I L I K E I R

```
 4.50+3.80+4.50-3.90
-3.50+4.50-4.50 = 5.40
```

= 5.4/7=0.77

Move to the next position

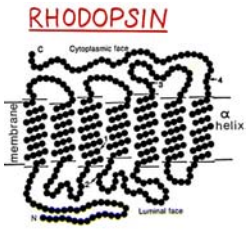
Sliding Window Approach

	Kyte-Doolittle
Alanine	1.8
Arginine	-4.5
Asparagine	-3.5
Aspartic acid	-3.5
Cysteine	2.5
Glutamine	-3.5
Glutamic acid	-3.5
Glycine	-0.4
Histidine	-3.2
Isoleucine	4.5
Leucine	3.8
Lysine	-3.9
Methionine	1.9
Phenylalanine	2.8
Proline	-1.6
Serine	-0.8
Threonine	-0.7
Tryptophan	-0.9
Tyrosine	-1.3
Valine	4.2

Hydropathy plots

•The window size can be changed. A small window produces "noisier" plots that more accurately reflect highly local hydrophobicity.

•A window of 9 or 11 is generally optimal for recognizing the long hydrophobic stretches that typify transmembrane stretches.



•In an α-helix the rotation is 100 degrees per amino acid
•The rise per amino acid is 1.5 Å
•To span a membrane of 30 Å approx.
30/1.5 = 20 amino acids are needed

Transmembrane Helix Predictions

•Not many structures known of transmembrane helix proteins

•Hydropathy analysis can be used to locate possible transmembrane segments

•The main signal is a stretch of hydrophobic and helix-loving amino acids

Hydropathy plot for rhodopsin