BCHE2030 Tutorial 3

Group

Contents

Proteins:

amino acid charisteristic

Pka value and ionization properties

Protein structure:

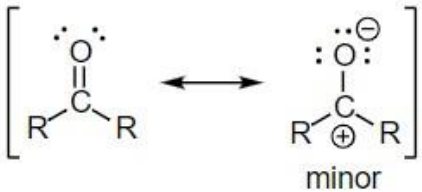
Amino acid sequence & Homologous protein, Protein Folding (Protein II p1-12, p27-35)

Secondary, tertiary structure

1. How can I distinguish whether the carbon atom in an aldehyde group is neutral, positively or negatively charged?

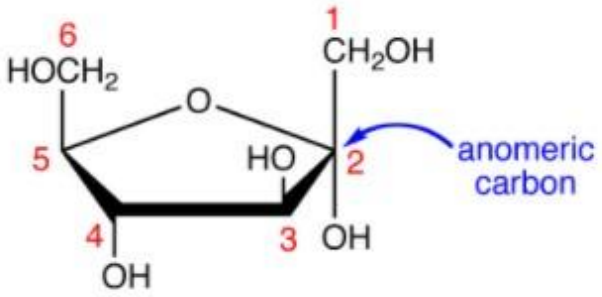
The carbonyl carbon can be neutral or carry a partial positive charge due to the presence of a *more* electronegative oxygen atom

The carbonyl carbon in aldehyde has two resonance structures.

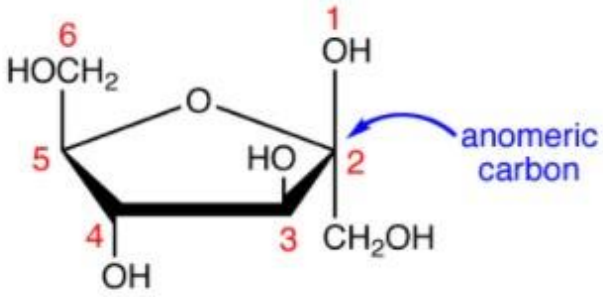


2. What are alpha and beta anomers?

In alpha anomers, the hydroxyl group bonded to the anomeric carbon is pointing downwards whereas in beta anomers, it is pointing upwards.



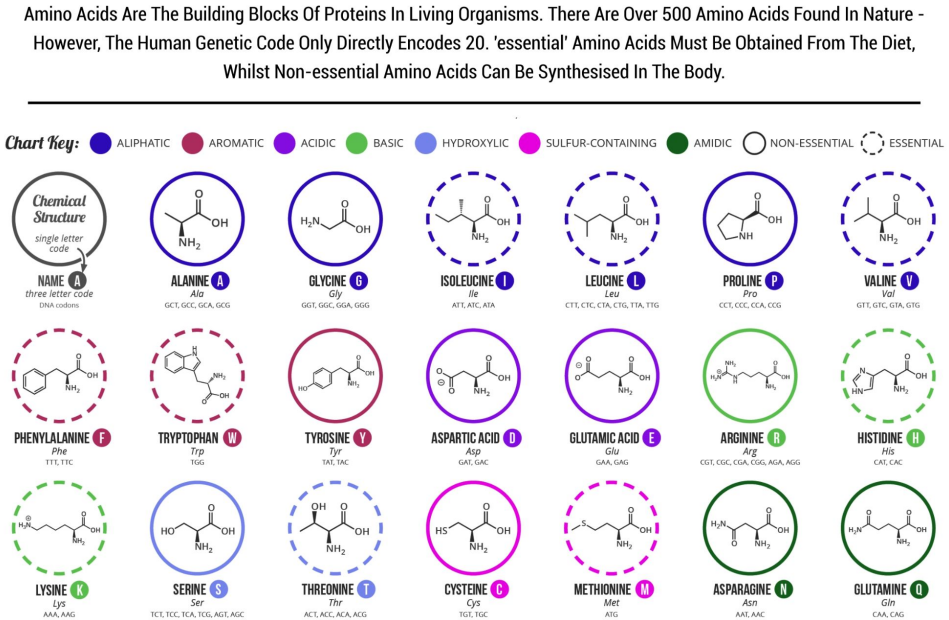
α-D-fructofuranose



β-D-fructofuranose

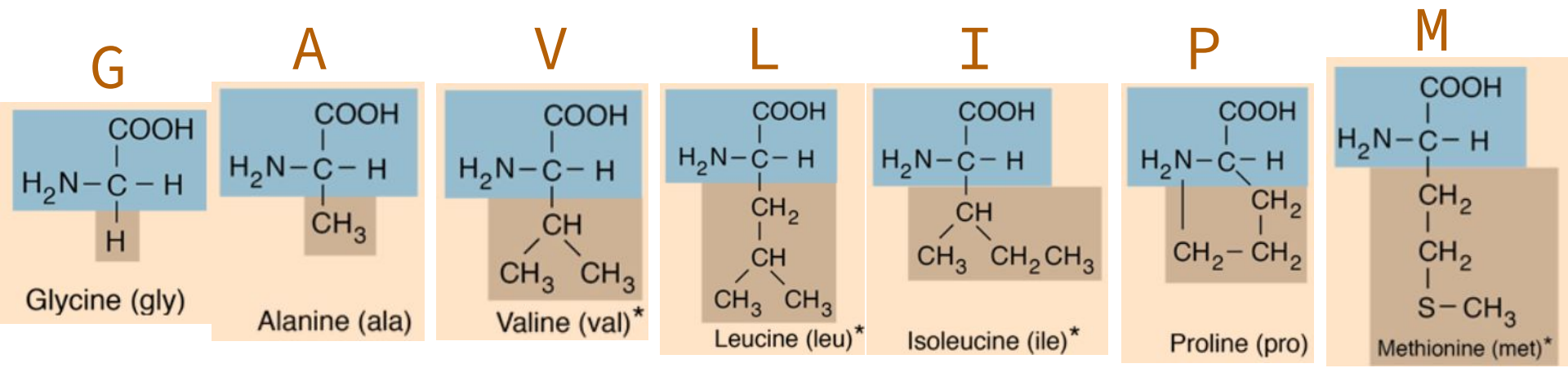
20 AMINO ACID

- Hydrophobic amino acids
- Polar amino acids
- -> +ve charged
- -> -ve charged



HYDROPHOBIC AMINO ACID

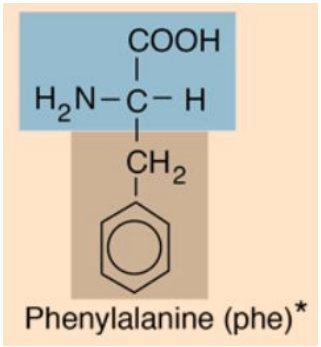
1. R group have similar eletronegativity -> hydrophobic
2. L & I have same chemical component
3. P aliphatic ring restrict in peptide angle
4. M is code for start codon (AUG)



Aromatic amino acid

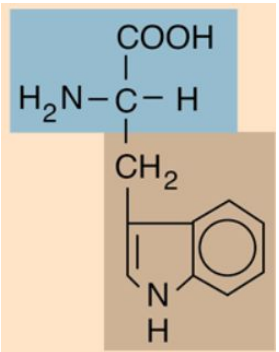
aromatic ring on the side chain of each of these three amino acids

F



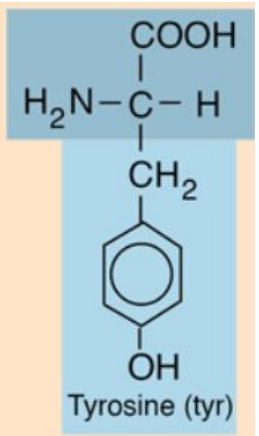
Phenylalanine (phe)*

W



Tryptophan (trp)*

Y



Tyrosine (tyr)

aromatic

hydrophobic

polar

Polar (nerutal) amino acid

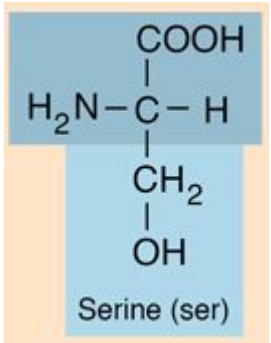
common phosphorylation site of protein

Nitrogen carrier

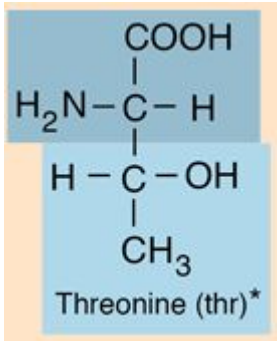
Disulfur bond formation

Intra-Inter bond

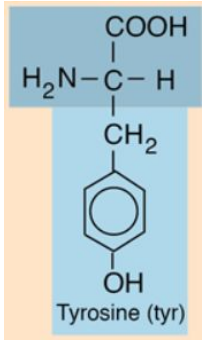
S



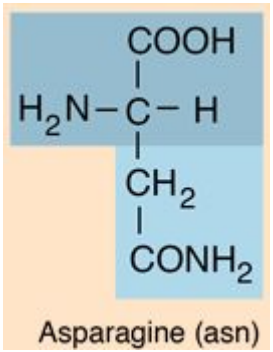
T



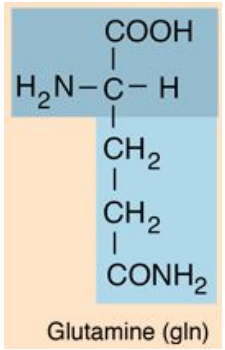
Y



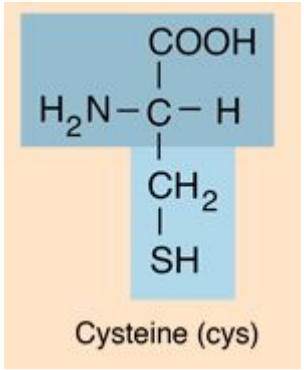
N



Q



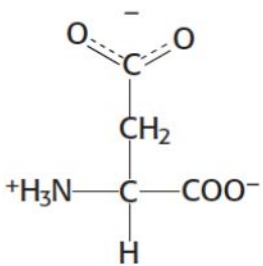
C



Polar (-ve charged) amino acid

pI < pH(7.4)

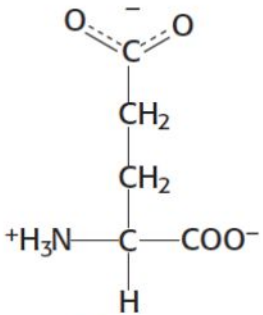
D



Aspartate
(Asp, D)

3.7

E



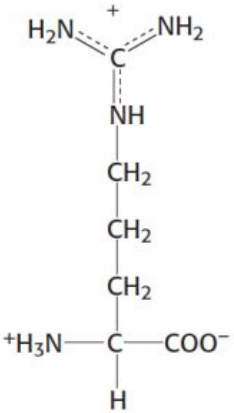
Glutamate
(Glu, E)

4

Polar (+ve charged) amino acid

pI >pH(7.4)

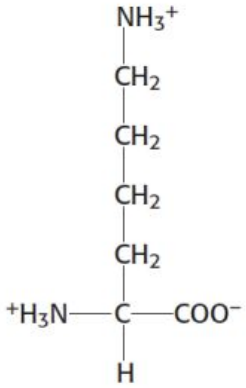
R



Arginine
(Arg, R)

12.5

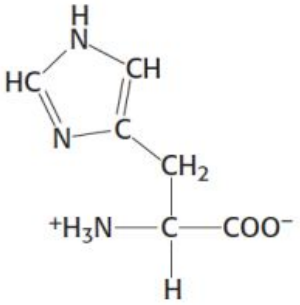
K



Lysine
(Lys, K)

10.5

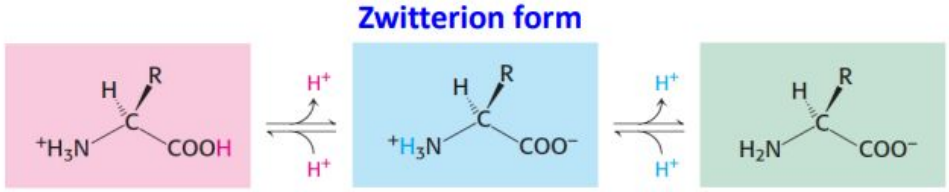
H



Histidine
(His, H)

6

Pka value



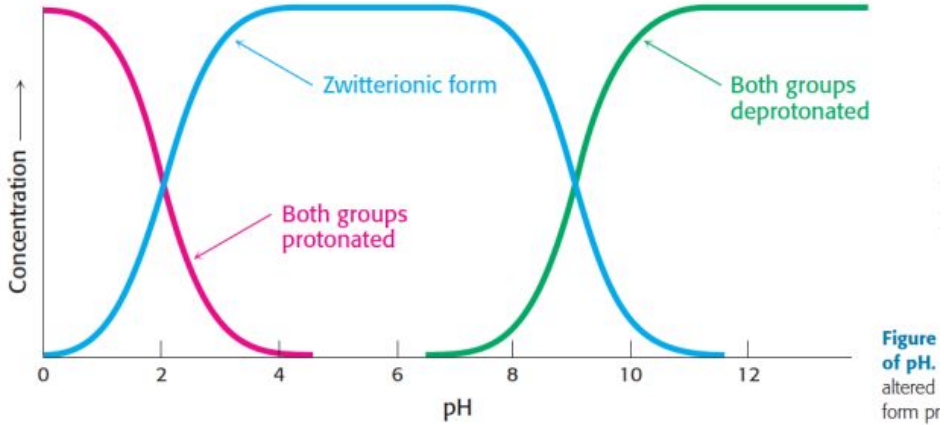
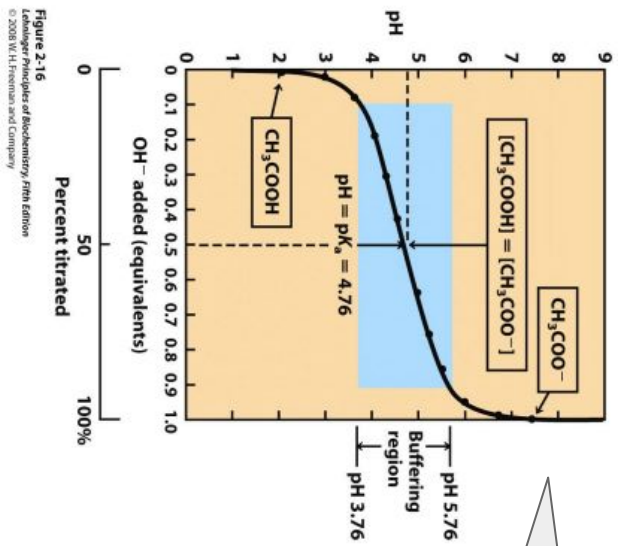
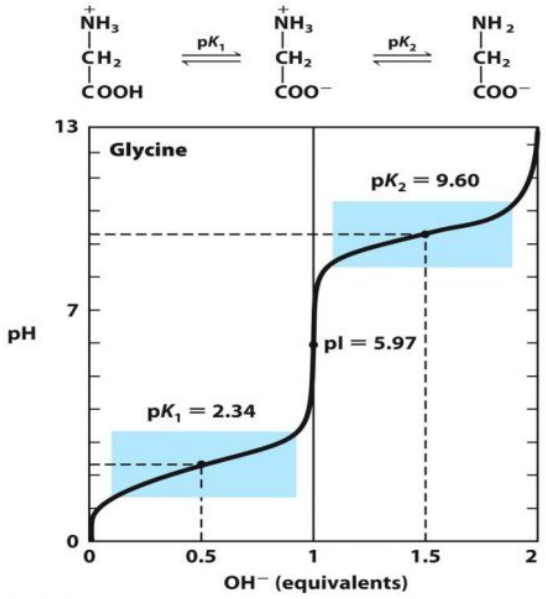


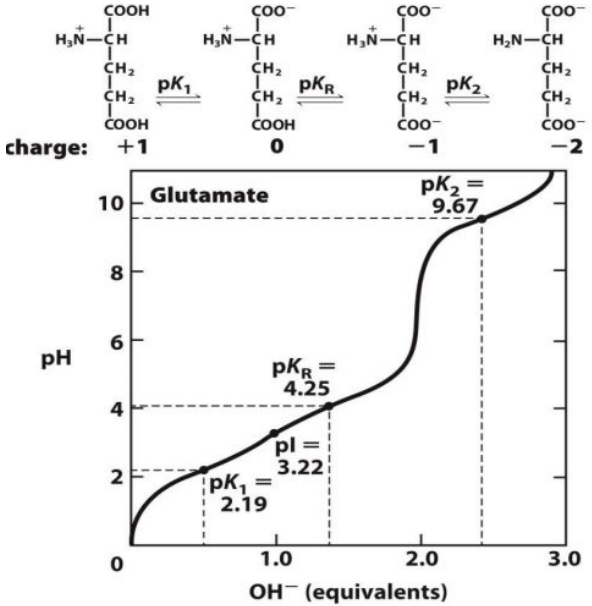
Figure of pH. altered form pr

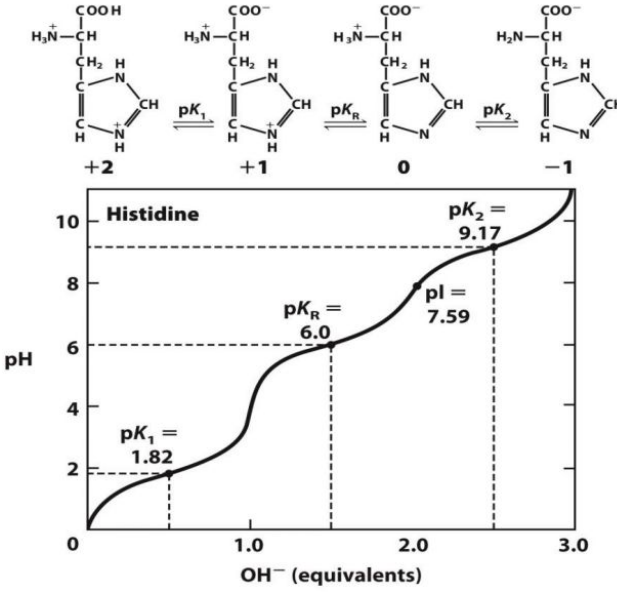


Recall Water Chapter

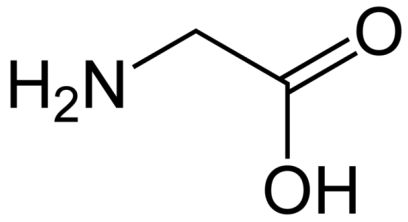
PKA Value


Glycine
pK1 = 2.34
pI = 5.97
pK2 = 9.60
pH
OH⁻ (equivalents)


charge: +1 0 −1 −2
Glutamate
pK1 = 2.19
pI = 3.22
pKR = 4.25
pK2 = 9.67
pH
OH⁻ (equivalents)


+2 +1 0 −1
Histidine
pK1 = 1.82
pKR = 6.0
pI = 7.59
pK2 = 9.17
pH
OH⁻ (equivalents)

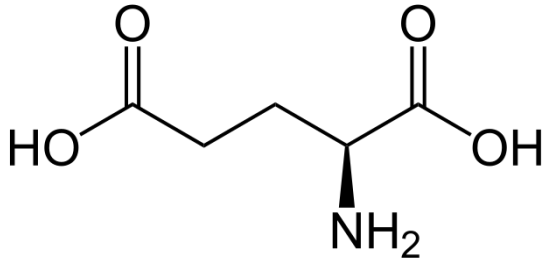
Isoelectric point

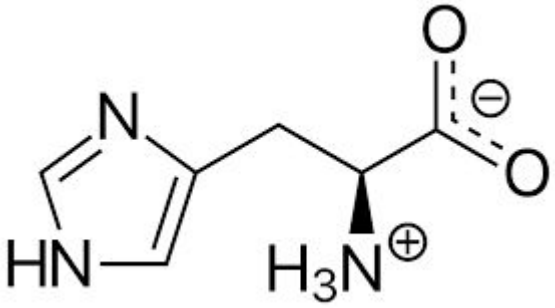


pK1 + pK2 /2

pK1 + pKR /2

pKR + pK2 /2





Isolectric point (Definitions)

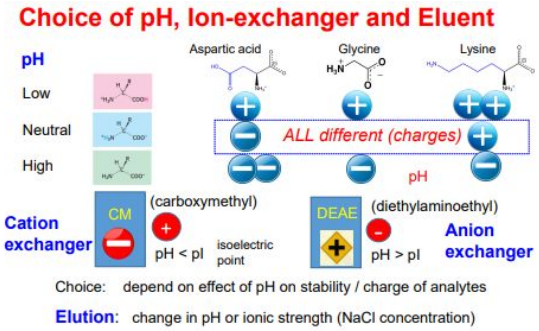
The characteristic pH at which the net charge of the molecule is zero.

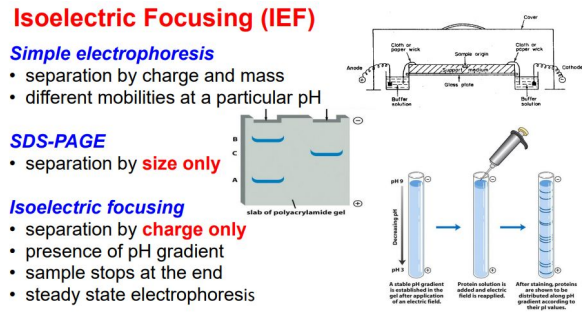
Given by the average of the pKas that involve the zwitterion.

The pH at which the amino acid does NOT migrate in an electric field. (Laboratory applications)

Recall zwitterion: able to act as either an acid or a base (lecture slide 27)

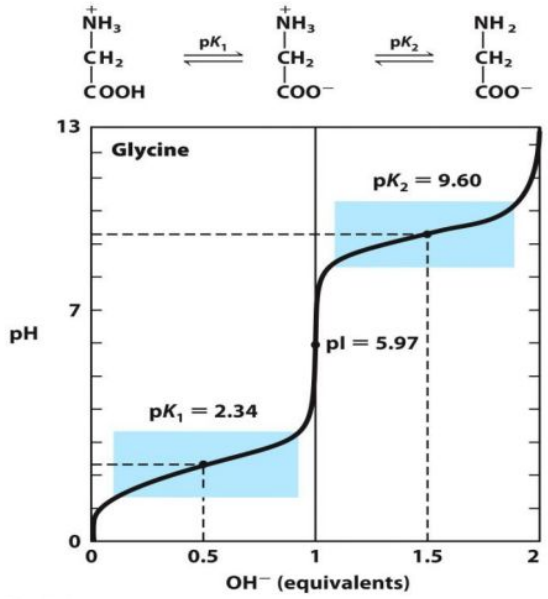
source: BCHE3030



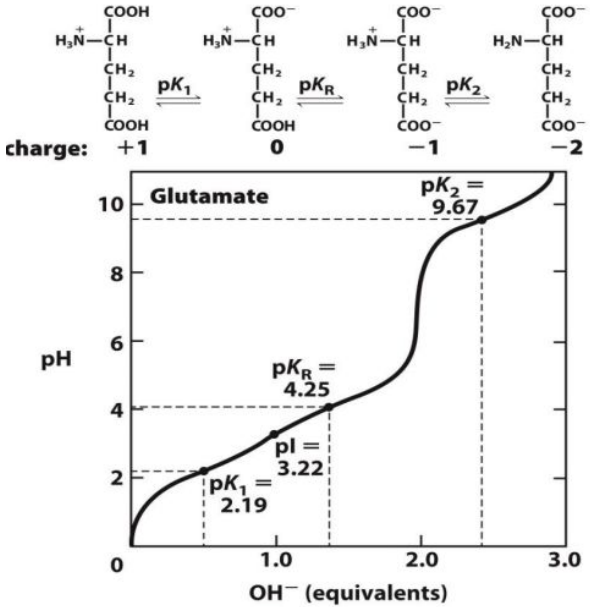


Isoelectric point

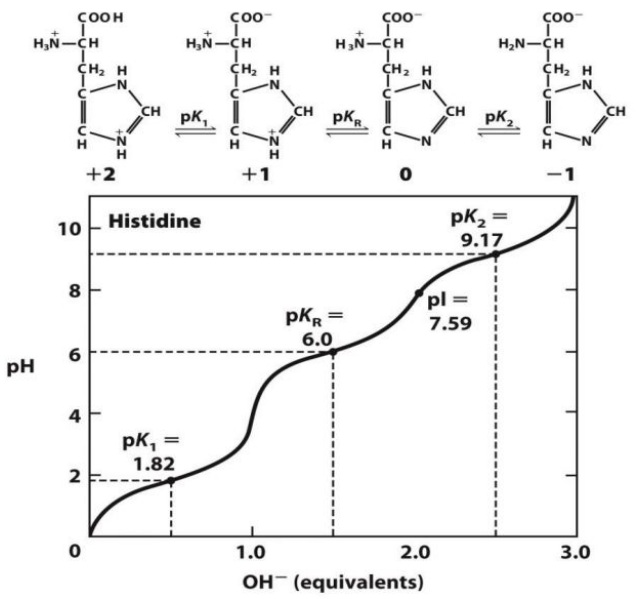
(pk1+ pk2) /2



(pk1+ pkR) /2



(pkR+ pk2) /2



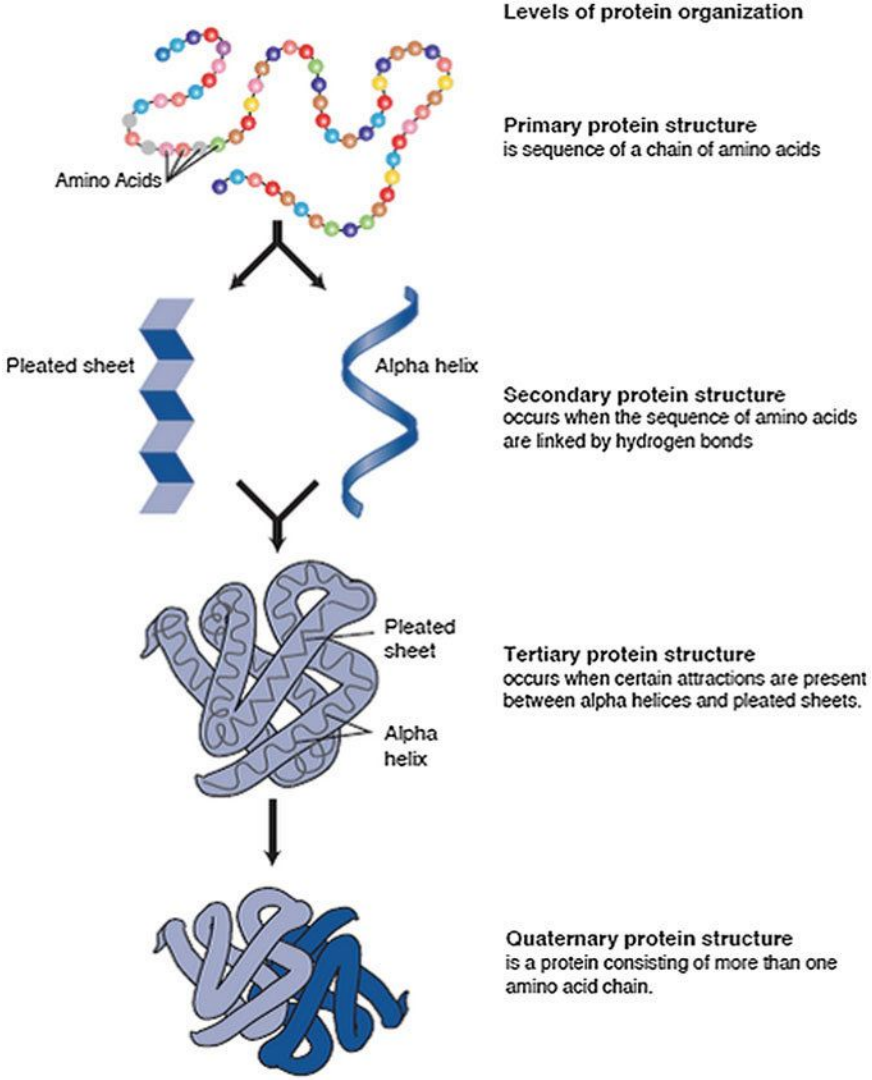
Secondary structure of protein

- conformation of a region of a peptide → **Secondary** (e.g.alpha helix)
- conformation of a peptide as one entity → **Tertiary**

  (e.g. myoglobin)

- conformation of multiple peptide → **Quaternary**

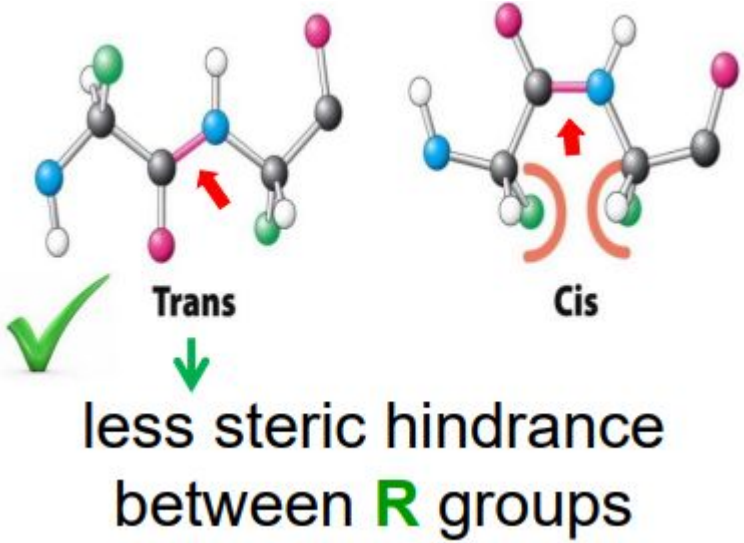
  (e.g. hemoglobin)



Zhang, Longhe & Zhong, Jing & Ren, Xiaofeng. (2017). Natural Fiber-Based Biocomposites. 10.1007/978-3-319-46610-1_3.

Peptide ANGLE

- peptide bond (HN-CO) = semi-planar, no rotation
- Usually, peptide bond adapt trans conformation
- Cis is seldomly allowed (except when Proline involved)



Ramachandran plot

- Phi ϕ :C(O)-N-C(α)-C(O) bonds
- Psi ψ :N-C(α)-C(O)-N bonds
- clockwise: positive
- anticlockwise: negative

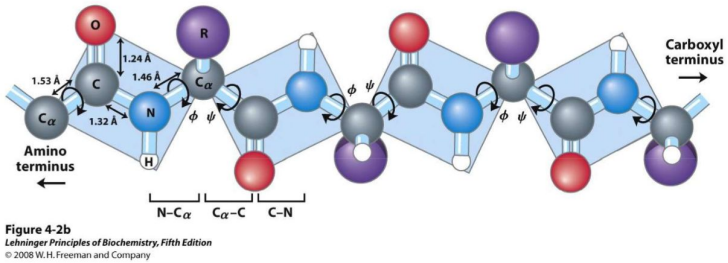
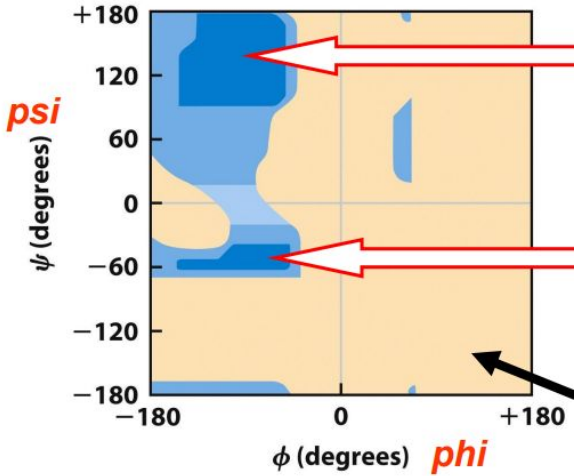
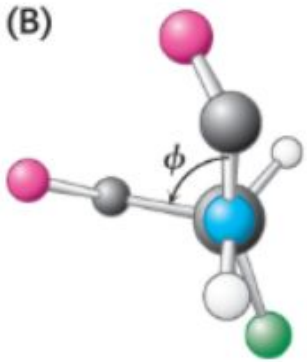
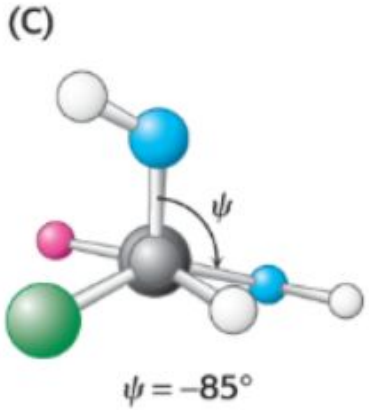


Figure 4-2b
Lehninger Principles of Biochemistry, Fifth Edition
© 2008 W.H. Freeman and Company



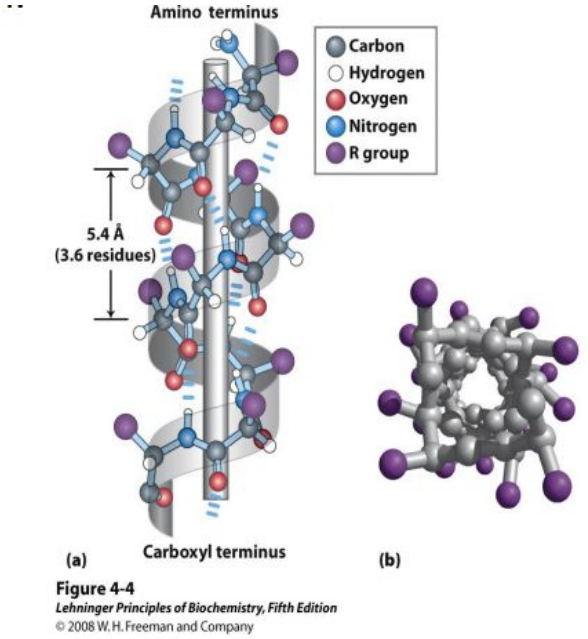


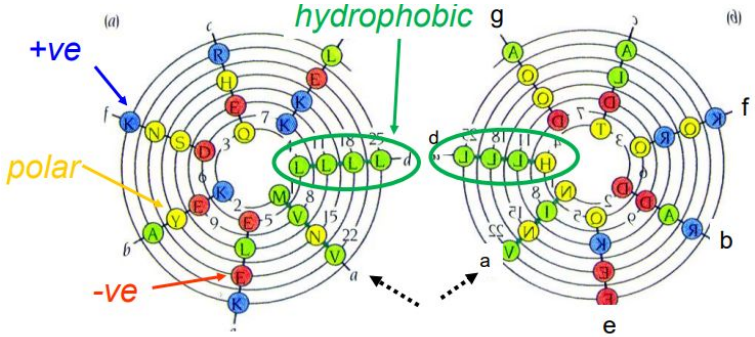


reverse -> negative when clockwise

Alpha helix

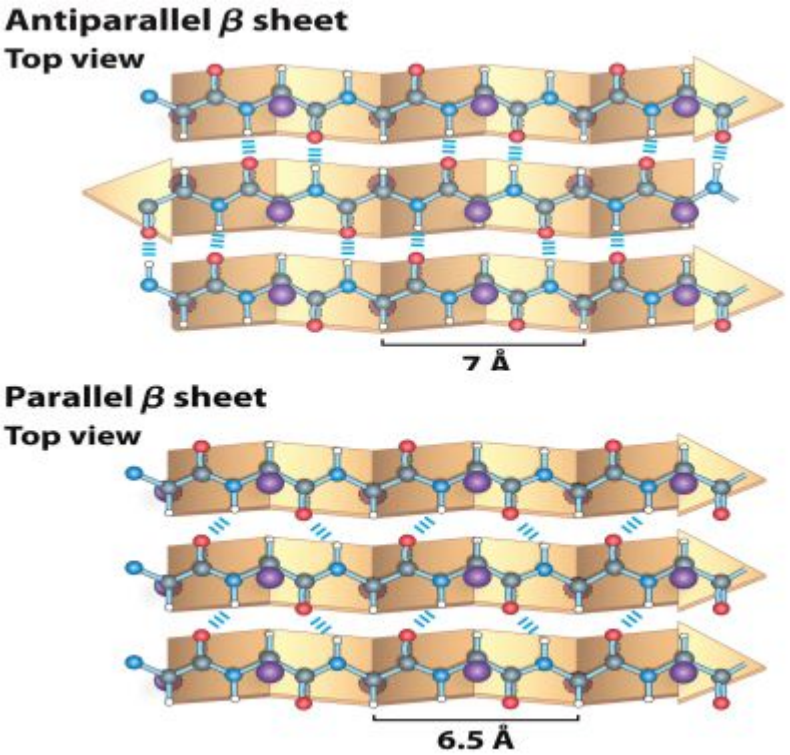
- bury peptide backbone
- side chains pointing outside
- Hydrogen bond between CO and NH among backbone
- Interaction between n and n+3/4 residue (facing same direction)
- compact





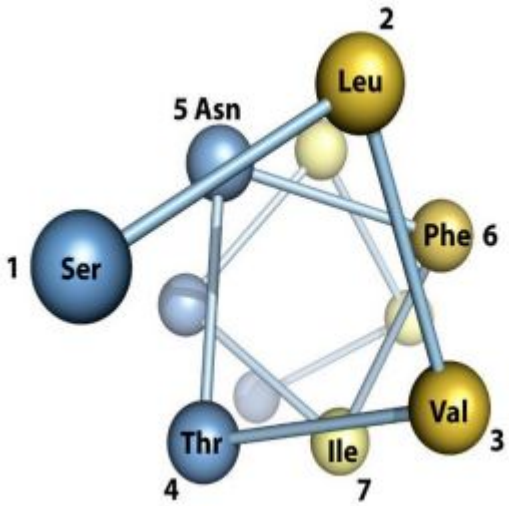
Beta strand

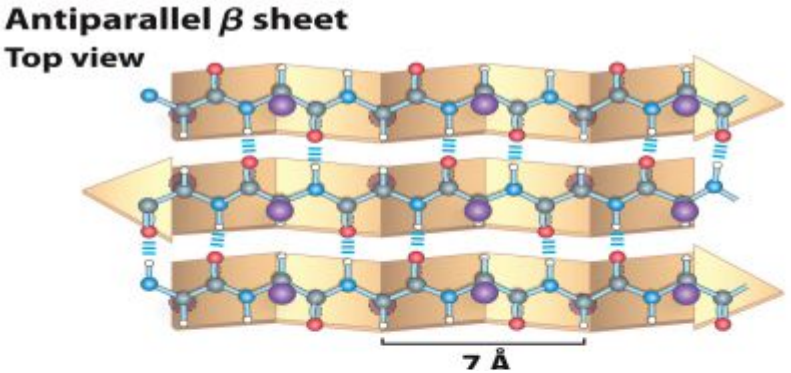
- Parallel or antiparallel(anti more stable)
- side chain pointing alternating direction
- more extended
- zig-zag, not a flat strand



DEDUCE SECONDARY STRUCTURE FROM SEQUENCE

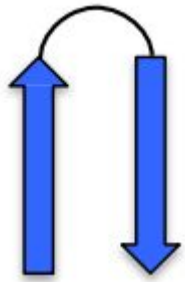
- NOT GUARANTEED!!! (might have exception)
- Consider globular protein only (membrane and fibrous excluded) → amphipathic secondary structure
- Consider distribution of amino acids, not singular AA





Tertiary and quaternary structure

- Secondary structures groups together -> forming motif



beta hairpin

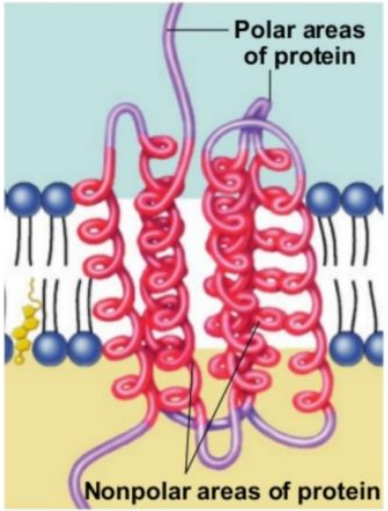
beta-alpha-beta motif

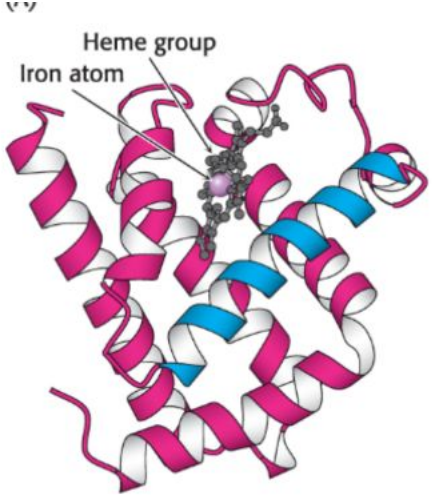
- One protein can have more than 1 motifs

Tertiary and quaternary structure

- covalent (S-S) and non-covalent interaction
  - disulphide bond, hydrophobic interaction, charge interaction, Hydrogen bond between side chains
- Mostly related to side chains, not backbone
- May involved other groups (heme, FAD, lipoate …)

membrane protein (contact with hydrophobic fatty acid)





globular protein (contact with hydrophilic environment)

Quiz 1

Which of the following amino acids would you expect to see in a protein present in the transmembrane region?

a) Arginine

b) Leucine

c) Glutamine

d) Histidine

Quiz 1

Which of the following amino acids would you expect to see in a protein present in the transmembrane region?

a) Arginine (positively charged polar amino acid)

b) Leucine (non-polar amino acid)

c) Glutamine (neutral and polar amino acid)

d) Histidine (positively charged polar amino acid)

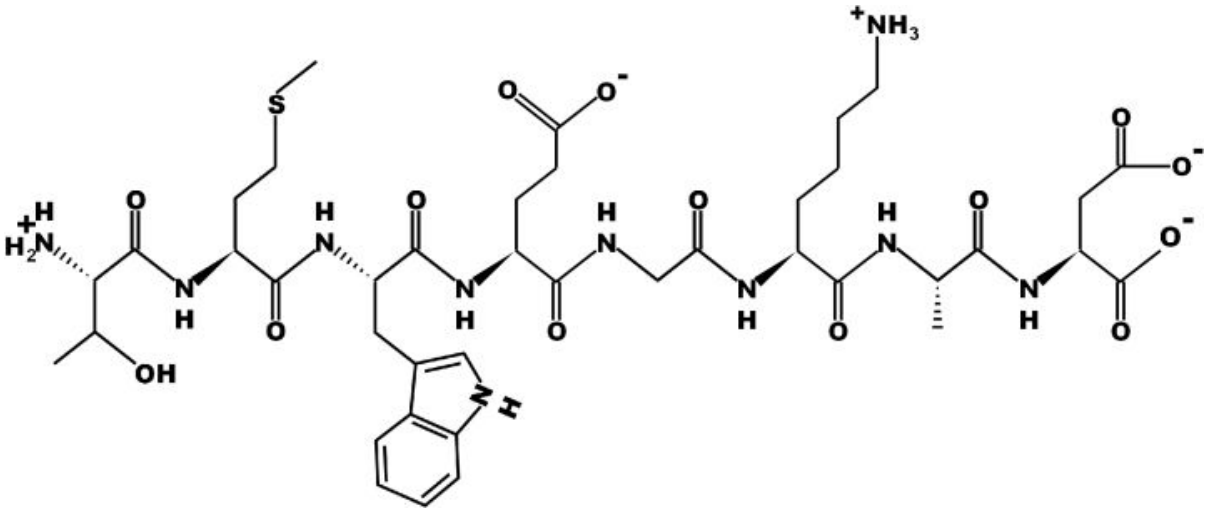
PEPTIDE EXAMPLE

What is the charge of the following peptide chain at physiological pH (pH = 7.4)?

TMWEGKAD

PEPTIDE EXAMPLE

What is the charge of the following peptide chain at physiological pH (pH = 7.4)?



AMINO ACID SEQUENCE

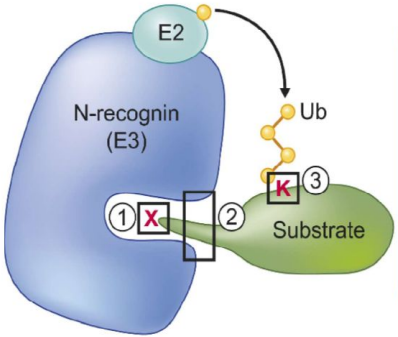
Protein sequence presented using the **abbreviation**s of the 20 amino acids

- It is directional!
  - (N-end)TMWEGKAD(C-end) is different from (N-end)DAKGEWMT(C-end)
  - One example about “end-specificity”: N-end rule
- We can BLAST the sequence of a protein to find other proteins with similar primary structure (and very often, they will be its homologous proteins)

Directional: N-end rule (NO need to know details; it's in CMBI4001)

The N-end Rule: in-vivo half-life of proteins

- The **N-terminus of a protein** can undergo a diverse posttranslational modifications
  - create **degradation signal (N-degron)** that mediate protein destruction via **the N-end rule pathway** of ubiquitin-mediated proteolysis.
- **N-recognin:** The E3 Ub ligases that recognize N-degron



Three key features that determine **an N-degron**:
1. a primary N-terminal **destabilizing amino acid ('X');**
2. an unstructured N-terminal region ensuring the N-residue is exposed and accessible;
3. an appropriately positioned downstream **lysine(s) (K)** to act as a receptor site for ubiquitin conjugation.

Take-away message: N terminus and C terminus has their own function respectively.

So (N-end) TMWEGKAD (C-end) is definitely different from (N-end) DAKGEWMT (C-end).

BLASTP

https://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome

to blast for protein sequence!

BLAST = basic local alignment search tool

Do you know what protein “mkalivlglv llsvtvqgkv fercelartl krlgmdgyrg islanwmcla kwesgyntra tnynagdrst dygifqinsr ywcndgktpg avnachlscs allqdniada vacakrvvrd pqgirawvaw rnrcqnrdvr qyvqgcgv” is? What proteins have similar sequence to it?

Blast it!

BLASTP

Sequences producing significant alignments Download Select columns Show 100

select all *100 sequences selected* GenPept Graphics Distance tree of results Multiple alignment MSA Viewer

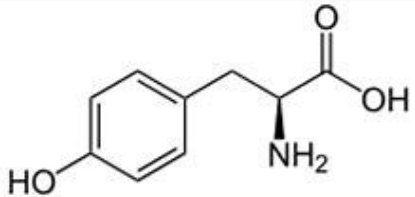
Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
lysozyme C precursor [Homo sapiens]	Homo sapiens	303	303	100%	5e-104	100.00%	148	NP_000230.1
lysozyme precursor (EC 3.2.1.17) [Homo sapiens]	Homo sapiens	303	303	100%	1e-103	99.32%	148	AAA36188.1
LYZ [synthetic construct]	synthetic construct	302	302	100%	3e-103	99.32%	148	AKI70658.1
lysozyme C precursor [Gorilla gorilla]	Gorilla gorilla	301	301	100%	3e-103	99.32%	148	NP_001266591.1
lysozyme [synthetic construct]	synthetic construct	301	301	100%	4e-103	99.32%	148	BAG73364.1
LYZ [synthetic construct]	synthetic construct	301	301	100%	4e-103	99.32%	148	AKI70659.1
lysozyme C [Pongo abelii]	Pongo abelii	300	300	100%	2e-102	97.97%	148	XP_002823550.1
LYZ [synthetic construct]	synthetic construct	300	300	100%	2e-102	99.32%	148	AKI70661.1
LYZ [synthetic construct]	synthetic construct	300	300	100%	2e-102	99.32%	148	AKI70660.1
lysozyme C [Nomascus leucogenys]	Nomascus leuco…	298	298	100%	6e-102	96.62%	148	XP_003259554.1
lysozyme precursor [Homo sapiens]	Homo sapiens	297	297	100%	2e-101	96.62%	148	AAC63078.1
RecName: Full=Lysozyme C; AltName: Full=1,4-beta-N-acetylmuramidase C; Flags: Precursor [Hylobates lar]	Hylobates lar	295	295	100%	2e-100	95.27%	148	P79180.1
lysozyme C [Chlorocebus sabaeus]	Chlorocebus sa…	283	283	100%	6e-96	89.86%	148	XP_008002172.1
RecName: Full=Lysozyme C; AltName: Full=1,4-beta-N-acetylmuramidase C; Flags: Precursor [Nasalis larva…	Nasalis larvatus	280	280	100%	1e-94	87.84%	148	P79811.1
lysozyme C precursor [Macaca mulatta]	Macaca mulatta	280	280	100%	2e-94	88.51%	148	NP_001095203.1
RecName: Full=Lysozyme C; AltName: Full=1,4-beta-N-acetylmuramidase C; Flags: Precursor [Miopithecus t…	Miopithecus tala…	279	279	100%	2e-94	88.51%	148	P79806.1
lysozyme C precursor [Papio anubis]	Papio anubis	279	279	100%	2e-94	88.51%	148	NP_001106112.1
RecName: Full=Lysozyme C; AltName: Full=1,4-beta-N-acetylmuramidase C; Flags: Precursor [Pygathrix ne…	Pygathrix nema…	278	278	100%	4e-94	86.49%	148	P79847.1
RecName: Full=Lysozyme C; AltName: Full=1,4-beta-N-acetylmuramidase C; Flags: Precursor [Semnopithec…	Semnopithecus …	278	278	100%	1e-93	86.49%	148	P67977.1

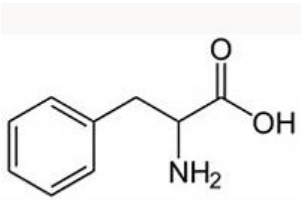
Answer: That sequence is **human's lysozyme C precursor**! You can also see that in the result list, some homologous proteins have good match to it.

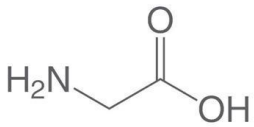
Homologous proteins share similar sequence

Residues:

- **invariant** (totally conserved): the amino acid residue is identical for **ALL** species
- **variant**: the amino acid residue is NOT totally identical
  - **conservative substitutions**: residues are replaced by amino acids of **similar chemical properties**







	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
Ala	4																			
Arg	−1	5																		
Asn	−2	0	6																	
Asp	−2	−2	1	6																
Cys	0	−3	−3	−3	9															
Gln	−1	1	0	0	−3	5														
Glu	−1	0	0	2	−4	2	5													
Gly	0	−2	0	−1	−3	−2	−2	6												
His	−2	0	1	−1	−3	0	0	−2	8											
Ile	−1	−3	−3	−3	−1	−3	−3	−4	−3	4										
Leu	−1	−2	−3	−4	−1	−2	−3	−4	−3	2	4									
Lys	−1	2	0	−1	−3	1	1	−2	−1	−3	−2	5								
Met	−1	−1	−2	−3	−1	0	−2	−3	−2	1	2	−1	5							
Phe	−2	−3	−3	−3	−2	−3	−3	−3	−1	0	0	−3	0	6						
Pro	−1	−2	−2	−1	−3	−1	−1	−2	−2	−3	−3	−1	−2	−4	7					
Ser	1	−1	1	0	−1	0	0	0	−1	−2	−2	0	−1	−2	−1	4				
Thr	0	−1	0	−1	−1	−1	−1	−2	−2	−1	−1	−1	−1	−2	−1	1	5			
Trp	−3	−3	−4	−4	−2	−2	−3	−2	−2	−3	−2	−3	−1	1	−4	−3	−2	11		
Tyr	−2	−2	−2	−3	−2	−1	−2	−3	2	−1	−1	−2	−1	3	−3	−2	−2	2	7	
Val	0	−3	−3	−3	−1	−2	−2	−3	−3	3	1	−2	1	−1	−2	−2	0	−3	−1	4

FYI: The BLOSUM62 matrix is derived from comparing amino acid substitutions among proteins that are no more than 62% similar – it's derived from empirical data

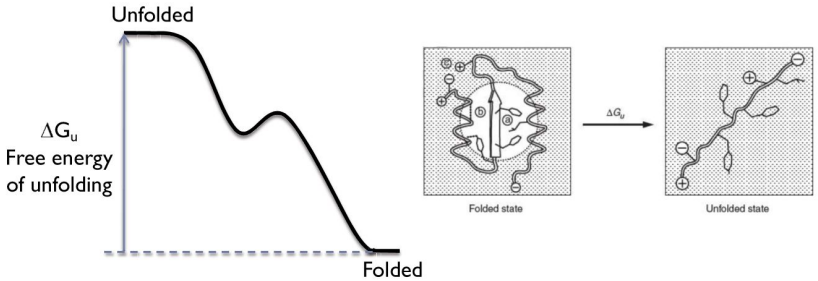
Henikoff, S. & Henikoff, J.G. (1992) Proc. Natl. Acad. Sci. USA 89:10915-10919.

What determines protein tertiary structure?

Protein's **primary structure** has enough information dictating its folding (you will learn a lot about this in CMBI4002)

Thermodynamics hypothesis

- Anfinsen (1972) "*The three-dimensional structure of a native protein in its normal physiological milieu is the one in which the Gibbs free energy of the whole system is lowest; that is, the native conformation is determined by the totality of inter-atomic interactions and hence by the amino acid sequence*"
- The tendency to reach this free energy minimum drives proteins fold.



Anfinsen's Nobel Lecture (1972) (www.nobelprize.org)

Stabilizing factors:

hydrophobic effect

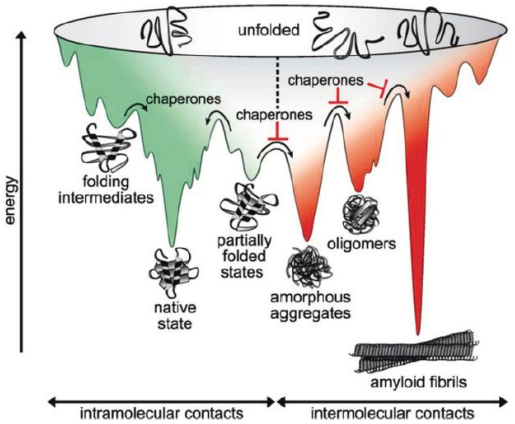
hydrogen bond

electrostatic interaction

(these weak interactions make protein fold properly; **urea can disrupt some of these interactions**)

SOME FEATURES OF PROTEIN FOLDING

- **stepwise** process, NOT random trial-and-error
- **intramolecular** interactions drive proteins towards its functional native state (**thermodynamically favourable**)
- partially folded intermediates may either: 1) turn to aggregates (often disease-causing) due to **intermolecular** interactions, or 2) assisted by chaperones, fold back to the native conformation



Quiz 2

Which statement best explains why antiparallel beta strands are more stable?

a) Antiparallel beta strands are highly packed

b) Parallel beta strands form fewer hydrogen bonds between adjacent strands

c) Antiparallel beta strands form stronger hydrogen bonds between adjacent strands

d) Parallel beta strands experience greater steric hindrance due to the orientation of R groups.

Quiz 2

Which statement best explains why antiparallel beta strands are more stable?

a) Antiparallel beta strands are highly packed

b) Parallel beta strands form fewer hydrogen bonds between adjacent strands

c) Antiparallel beta strands form stronger hydrogen bonds between adjacent strands (because the C=O and NH group of peptide bonds in adjacent strands are closer)

d) Parallel beta strands experience greater steric hindrance due to the orientation of R groups. (R groups extend from sheet and point on alternatively on opposite sides)