Protein Structure, Function and Methods of Analysis I

Dr. Abel Baerga Ortiz
Bioquímica

Oficina: A626
Teléfono: x1603
Email: **abel.baerga@upr.edu**

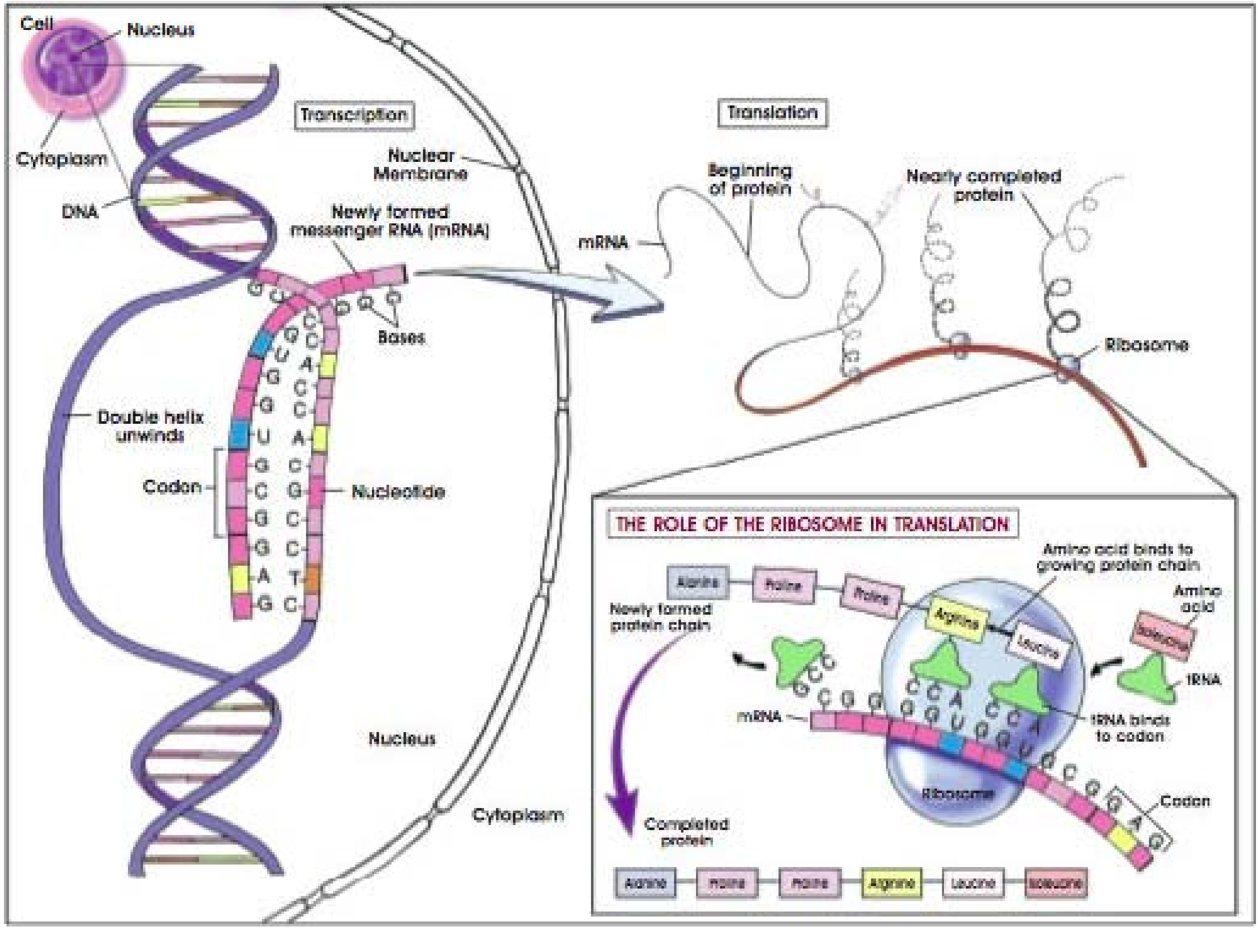
Referencias:

Lieberman, M; Marks, AD. Basic Medical Biochemistry: A Clinical Approach, 3rd Edition, 2009

Devlin, Thomas M. Textbook of Biochemistry with Clinical Correlations, 6th Edition, 2006

Nelson, DL; Cox, MM. Lehninger Principles of Biochemistry, 3rd Edition 2000

Proteins


Cell
Nucleus
Cytoplasm
DNA
Transcription
Nuclear Membrane
Newly formed messenger RNA (mRNA)
Bases
Double helix unwinds
Codon
Nucleotide
Nucleus
Cytoplasm
mRNA
Translation
Beginning of protein
Nearly completed protein
Ribosome
THE ROLE OF THE RIBOSOME IN TRANSLATION
Newly formed protein chain
Alanine
Proline
Proline
Arginine
Leucine
Amino acid binds to growing protein chain
Amino acid
Isoleucine
tRNA
mRNA
tRNA binds to codon
Ribosome
Codon
Completed protein
Alanine
Proline
Proline
Arginine
Leucine
Isoleucine

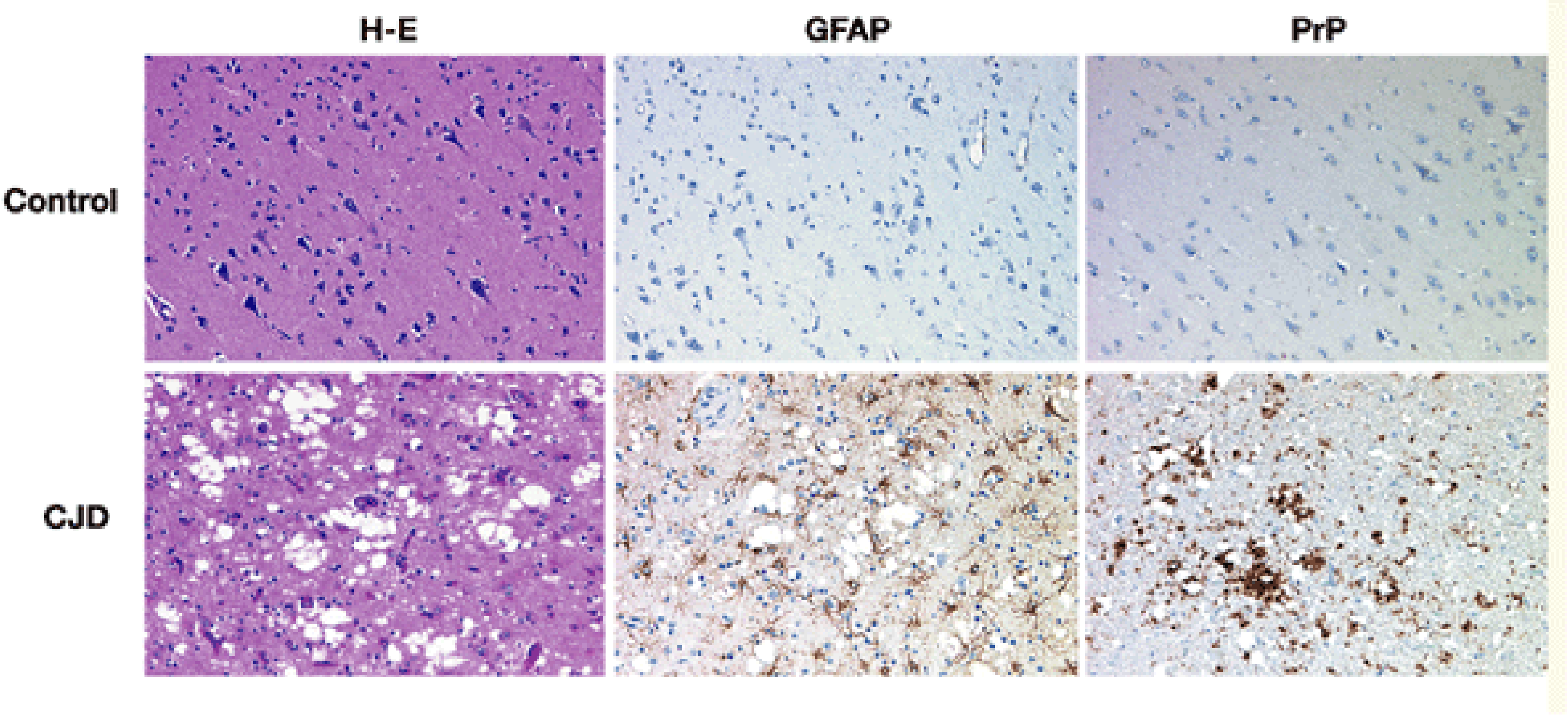
Prion Disease and Protein Misfolding

A) Creutzfeldt-Jacob Disease CJD
B) Bovine Spongiform Encephalopathy
C) Kuru
D) Fatal familial insomnia

Histopathology:

spongiform changes in the brain
degeneration of neurons
astrocytosis

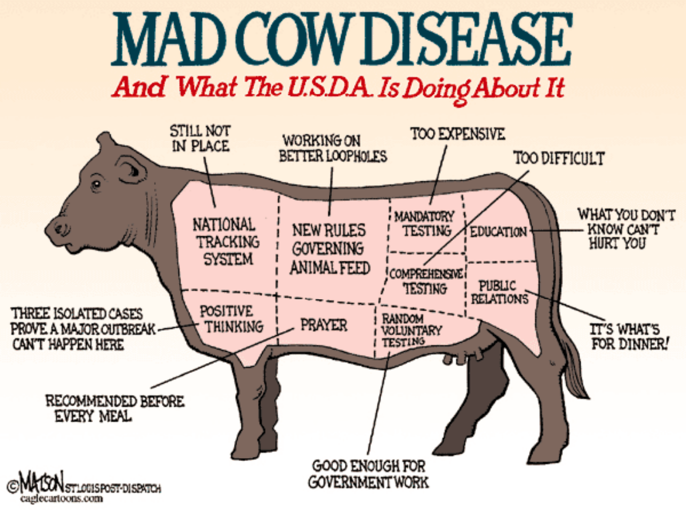
Prion Disease and Protein Misfolding

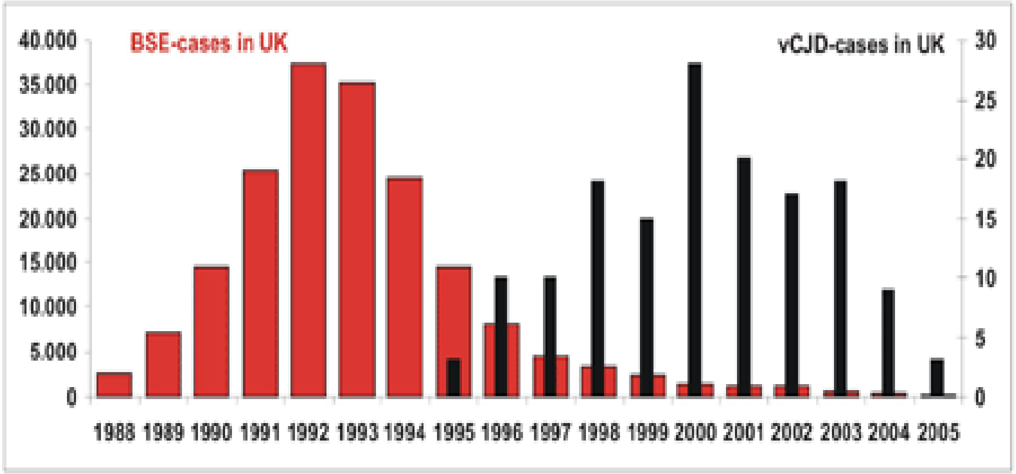


Nature Reviews | Neuroscience

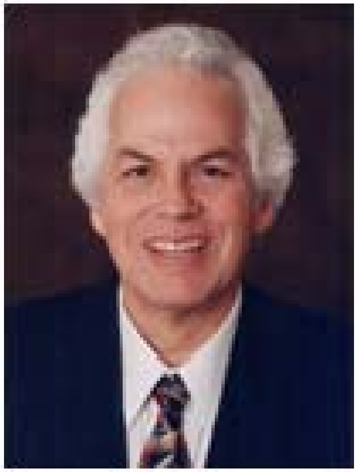
(pictures Adriano Aguzzi, Markus Glatzel, Fabio Montrasio, Marco Prinz & Frank L. Heppner
Nature Reviews Neuroscience **2**, 745-749**)**

Prion Disease and Protein Misfolding



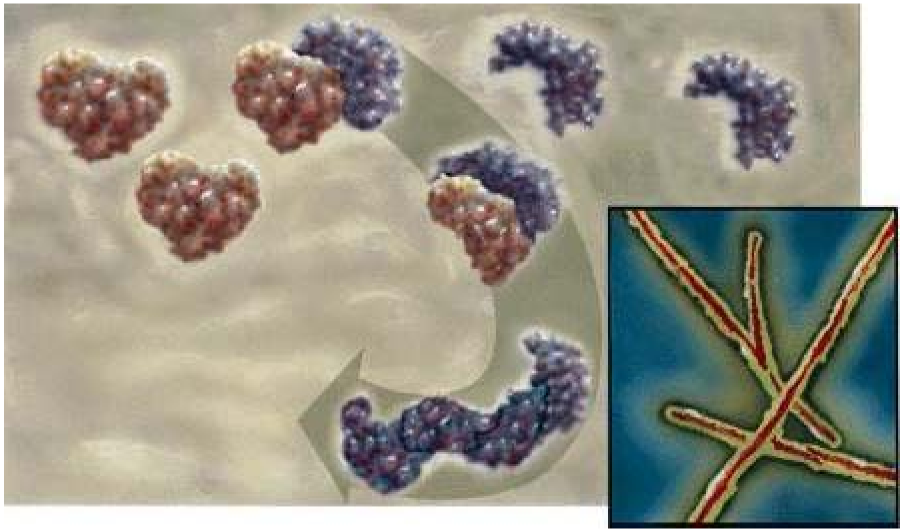


Prion Disease and Protein Misfolding



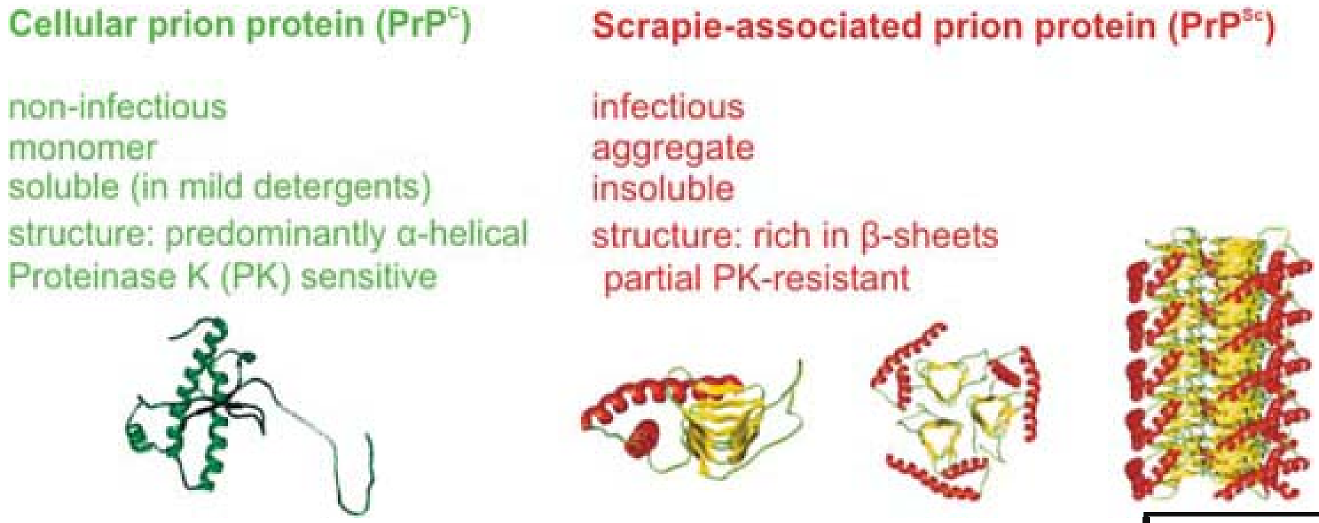
Stanley B. Prusiner

Nobel Prize in Medicine
For the Prion Hypothesis



Irreversible conformational self – replication of prion protein

Prion Disease and Protein Misfolding



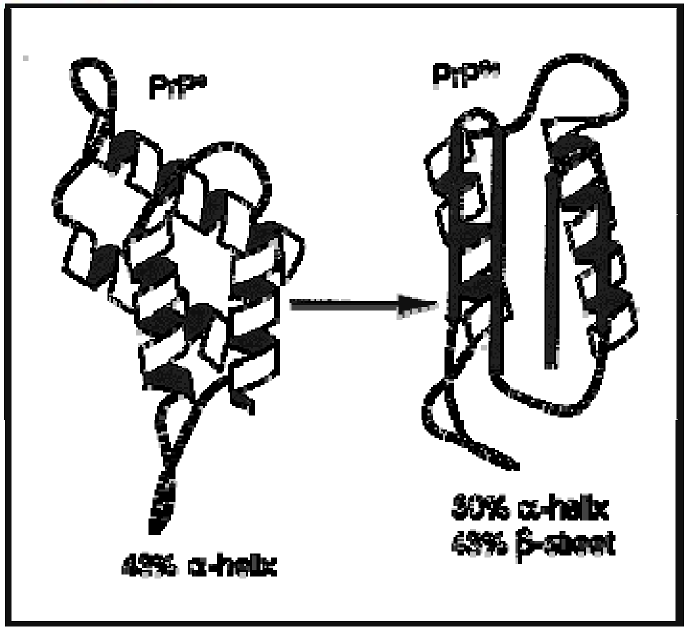
Cellular prion protein (PrP^C)

non-infectious
monomer
soluble (in mild detergents)
structure: predominantly α-helical
Proteinase K (PK) sensitive

Scrapie-associated prion protein (PrP^{Sc})

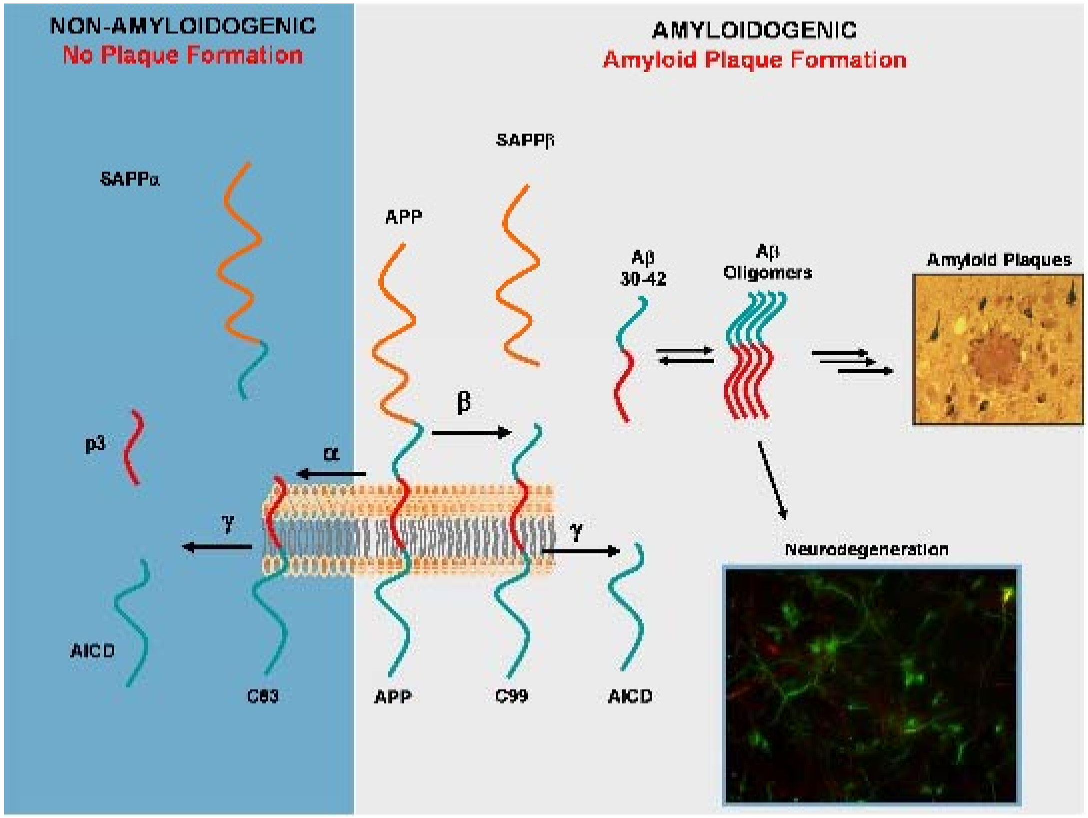
infectious
aggregate
insoluble
structure: rich in β-sheets
partial PK-resistant

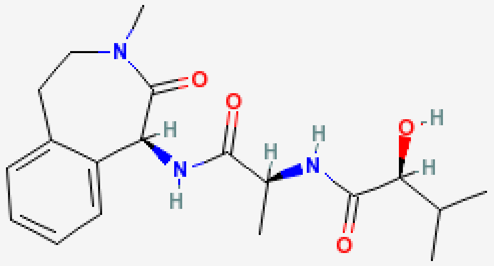
https://www.youtube.com/watch?v=6-Tz8a_vgX0



Amyloid Plaques and Alzheimer's Disease

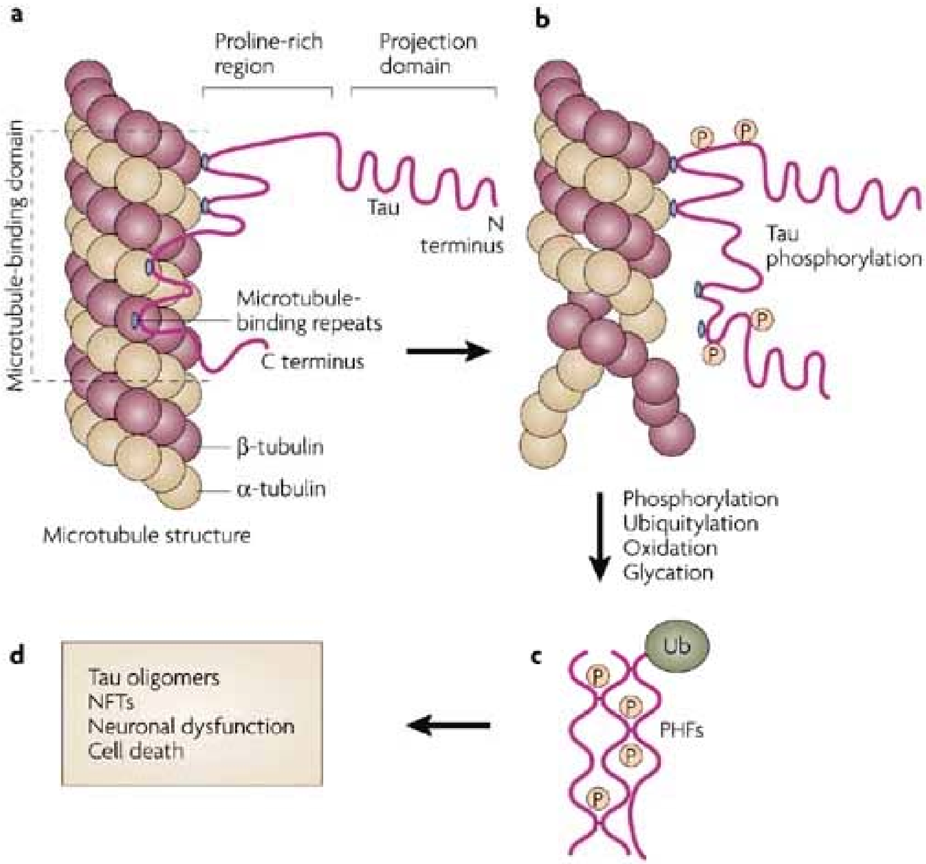
Protein oligomers form the plaque
normally associated with Alzheimer's Disease





Semagacestat
Gamma-Secretase Inhibitor

Neurofibrillary Tangles and Alzheimer's Disease



Nature Reviews | Drug Discovery

Tau is a protein present in neurons only.

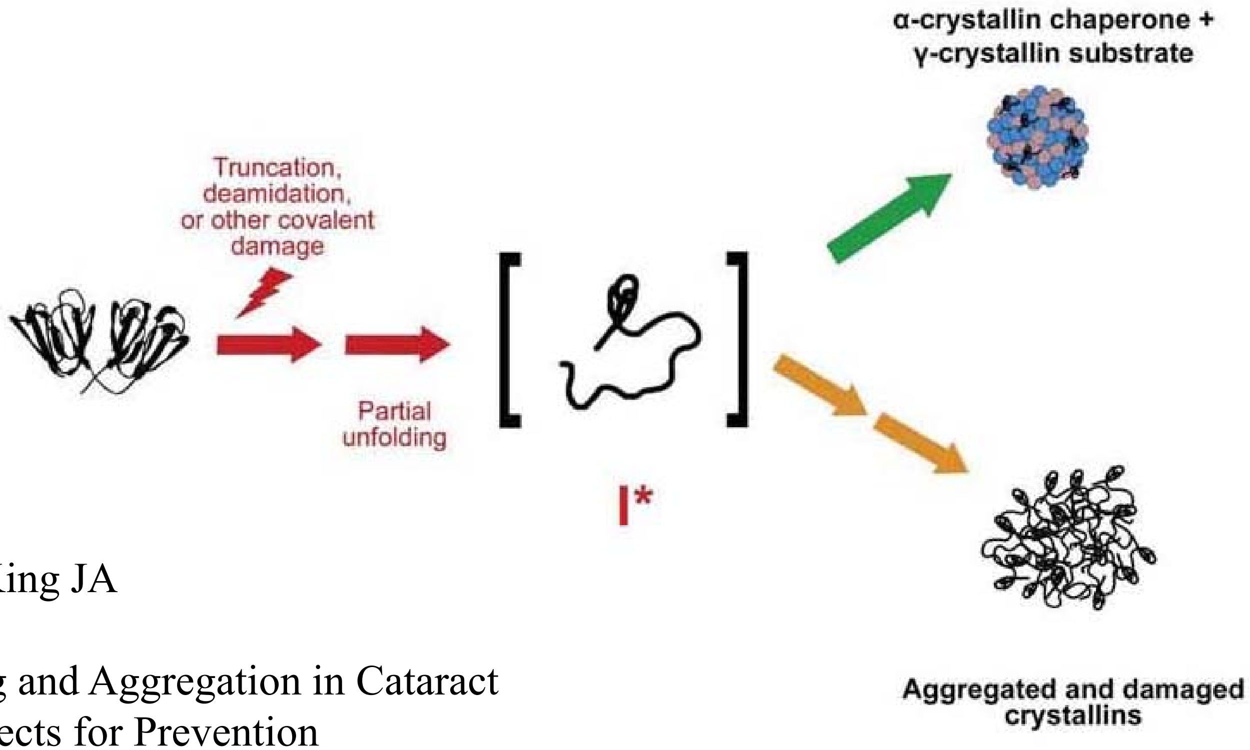
Tau stabilizesmicrotubules

Hyper-phosphorylation results in aggregation

Tau aggregates are associated with Alzheimer and other dementias.

Lens Proteins and Damage Accumulation

With age, covalent protein damage accumulates through pathways thought to include UV radiation, oxidation, deamidation, and truncations.



Moreau KL and King JA

Protein Misfolding and Aggregation in Cataract Disease and Prospects for Prevention

Trends Mol Med 2012 May ; 18(5): 273–282

Protein Structure, Function and Methods of Analysis I

I. Protein Structure – Primary, Secondary, Tertiary, etc

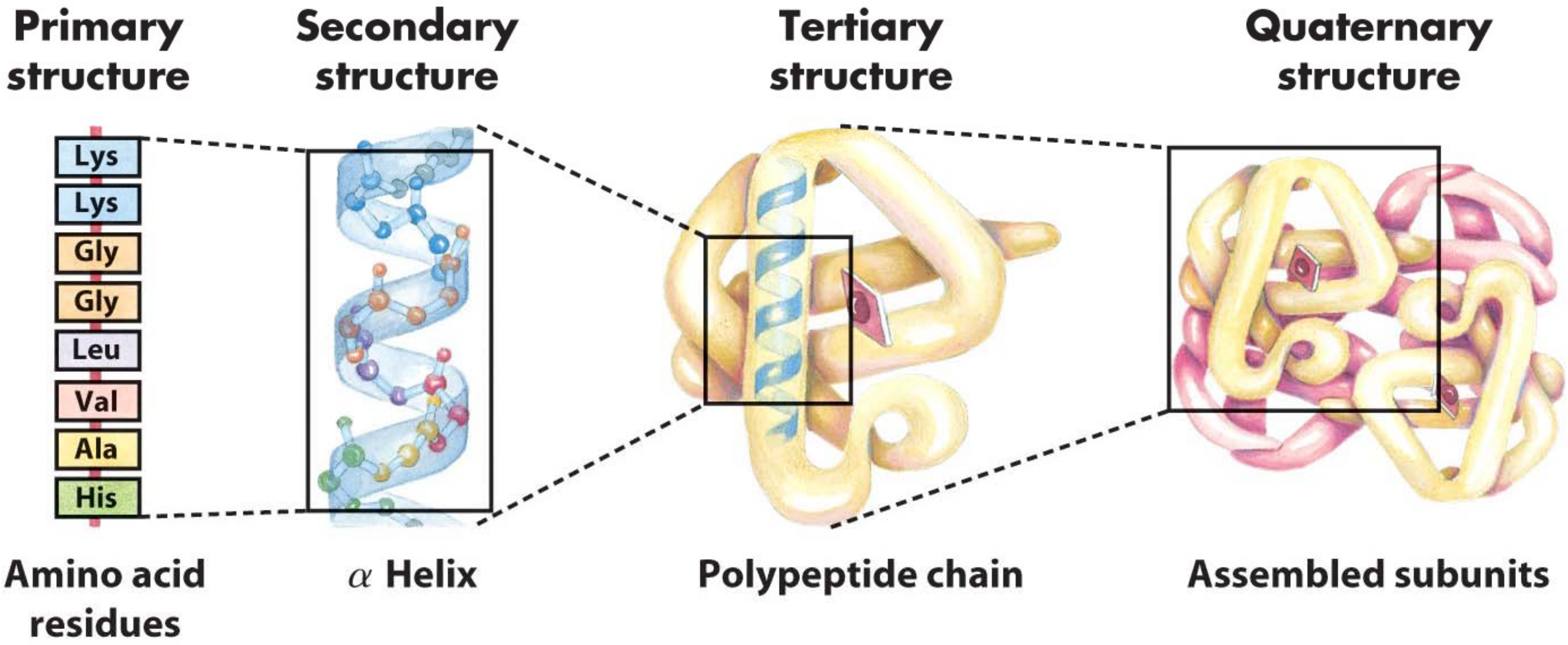
II. Chemical Properties of Proteins – pH, UV absorbance

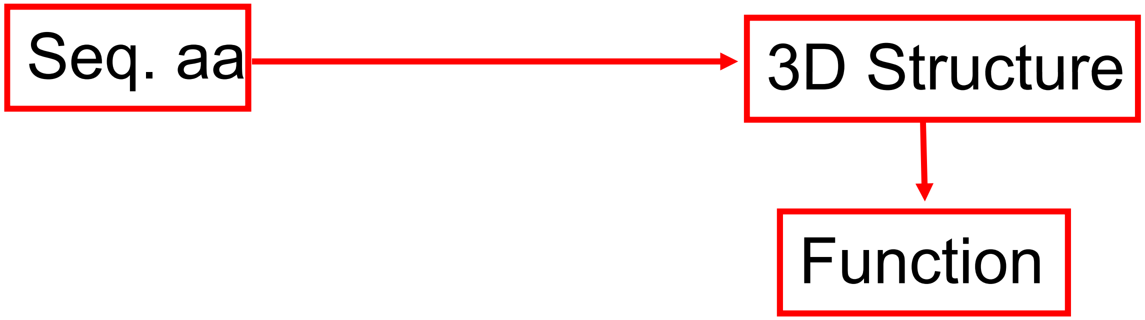
III. Protein Stabilization and Denaturation

IV. Post-translational Modifications

V. Analysis or Proteins
    i. Electrophoresis – PAGE, IEF
    ii. Fluorescence – for sub-cellular localization
    iii. Chromatography
    iv. Primary Structure Determination – ID
    v. Glycosylation

The Basics of Protein Structure





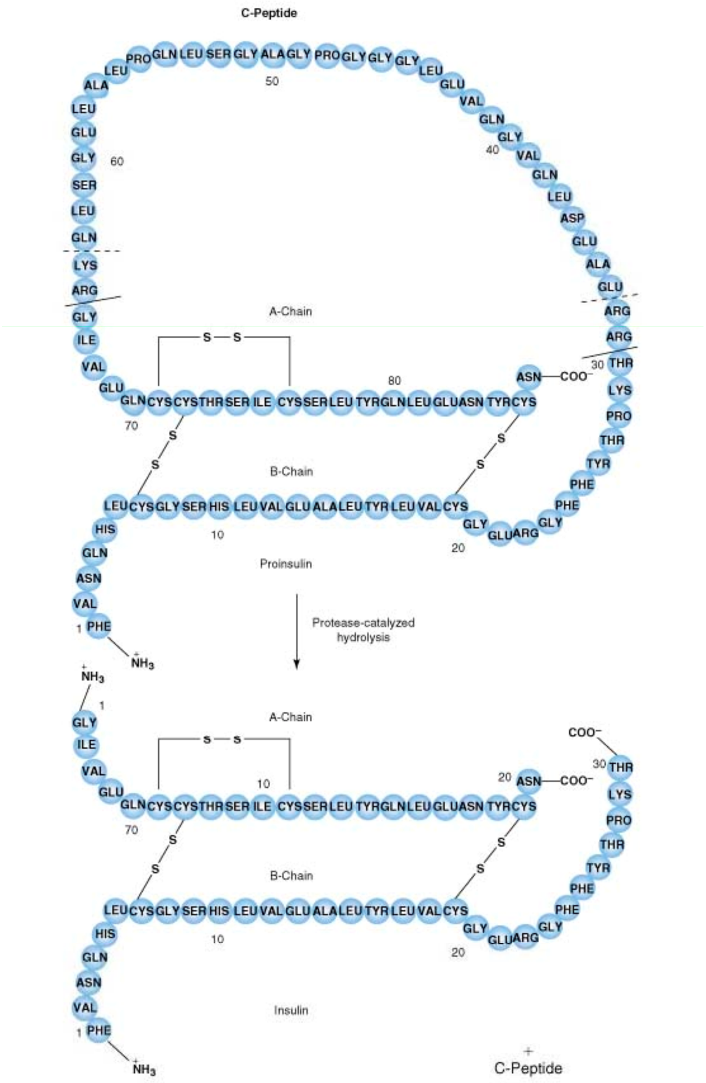
Seq. aa

3D Structure

Function

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

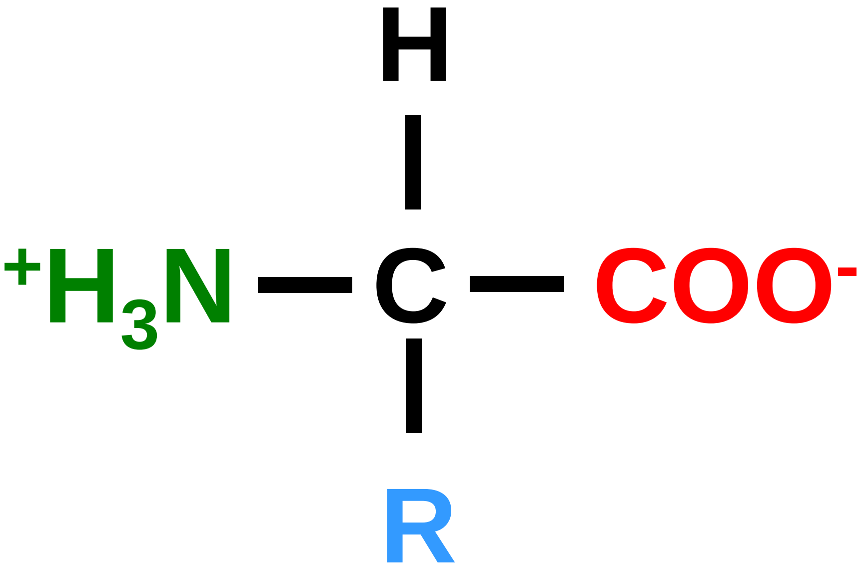
Primary Structure



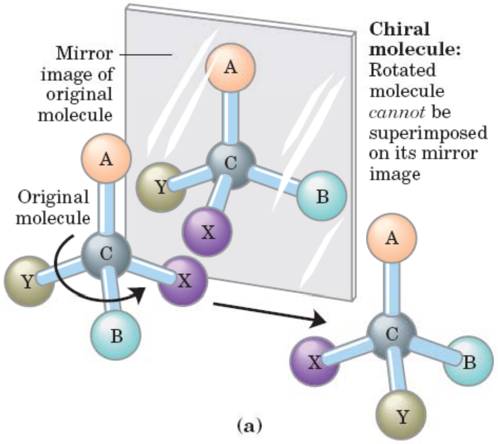
- Refers to the amino acid sequence of the protein
- Stabilized by the peptide bond
- Very stable, to “destroy” need:
  - ✓ 6 N HCl 100 -110°C 18 - 36 hrs.
  - ✓ Problem with acid hydrolysis:
    - Destroys Trp
    - glutamine & asparagine convert into glutamic acid & aspartic acid respectively

Devlin, *Biochemistry 6th ed* (2006)

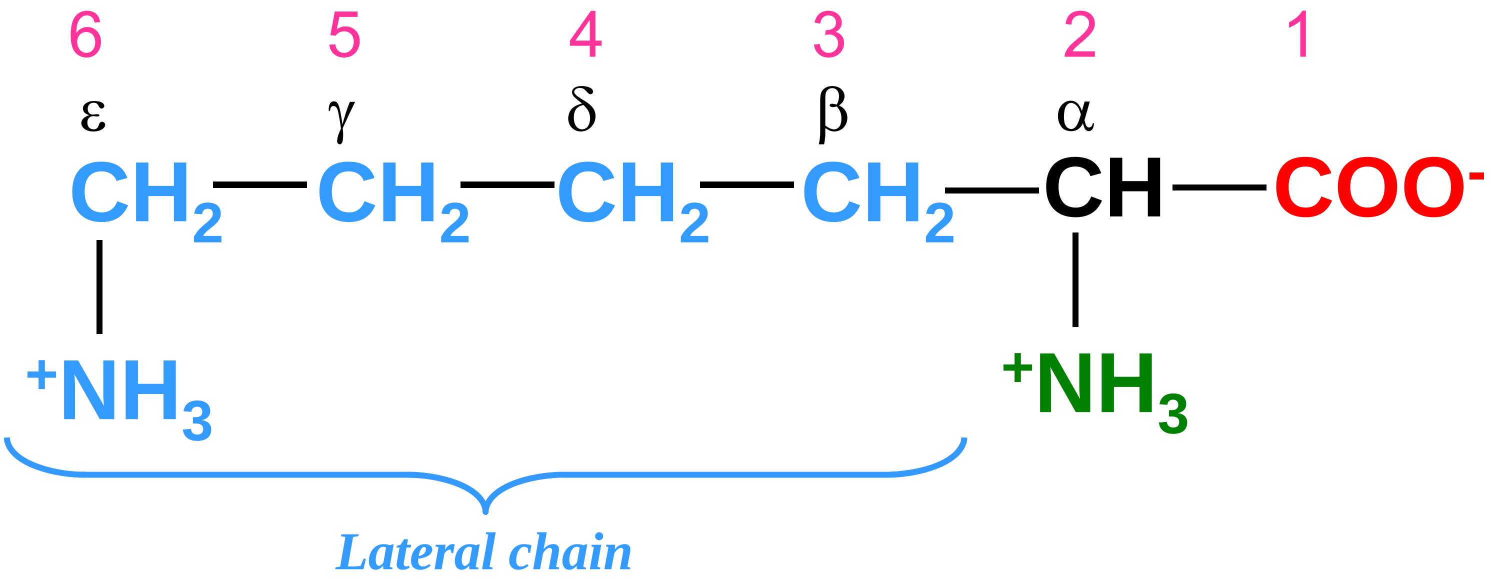
Amino Acids - Structure



$${}^{+}H_3N - \underset{\underset{R}{|}}{\overset{\overset{H}{|}}{C}} - COO^{-}$$



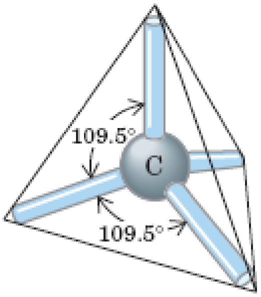
Amino Acids - Nomenclature

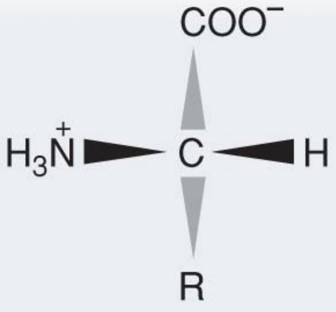


$$\underset{^+NH_3}{CH_2}-CH_2-CH_2-CH_2-\underset{^+NH_3}{CH}-COO^-$$

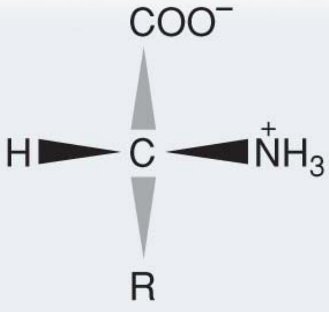
lysine

Absolute Configuration

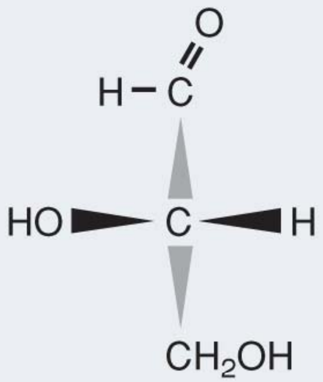




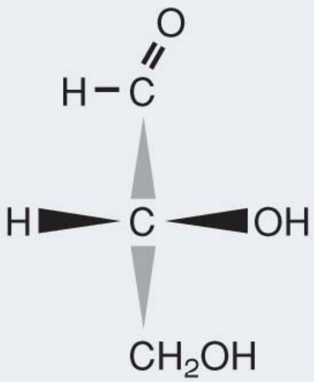
L-amino Acid



D-amino acid

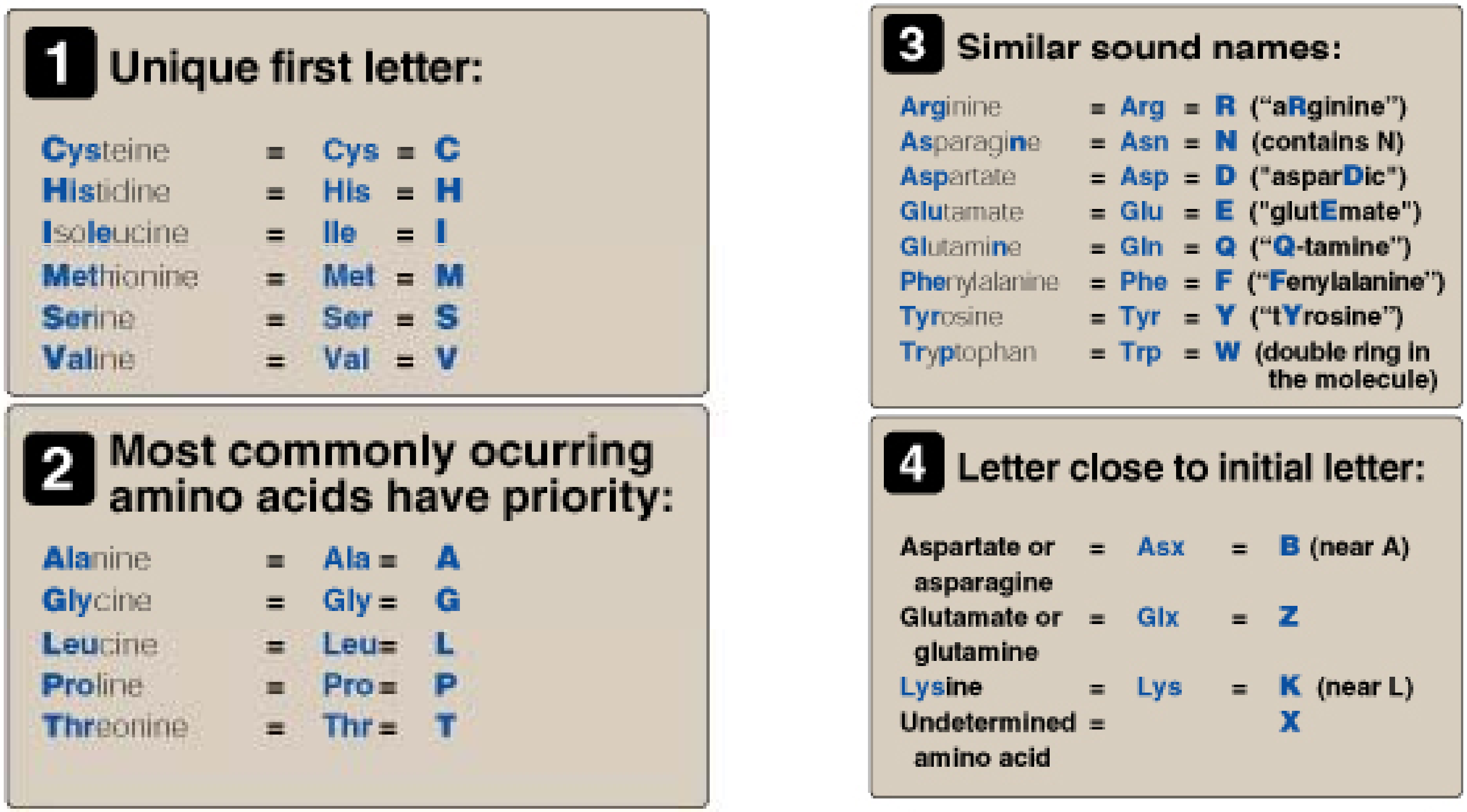


L-glyceraldehyde



D-glyceraldehyde

Amino Acids - Nomenclature



1 Unique first letter:

Cysteine = Cys = C
Histidine = His = H
Isoleucine = Ile = I
Methionine = Met = M
Serine = Ser = S
Valine = Val = V

2 Most commonly ocurring amino acids have priority:

Alanine = Ala = A
Glycine = Gly = G
Leucine = Leu = L
Proline = Pro = P
Threonine = Thr = T

3 Similar sound names:

Arginine = Arg = R ("aRginine")
Asparagine = Asn = N (contains N)
Aspartate = Asp = D ("asparDic")
Glutamate = Glu = E ("glutEmate")
Glutamine = Gln = Q ("Q-tamine")
Phenylalanine = Phe = F ("Fenylalanine")
Tyrosine = Tyr = Y ("tYrosine")
Tryptophan = Trp = W (double ring in the molecule)

4 Letter close to initial letter:

Aspartate or asparagine = Asx = B (near A)
Glutamate or glutamine = Glx = Z
Lysine = Lys = K (near L)
Undetermined amino acid = X

Williams & Wilkins, *"Lippincott's Illustrated Reviews: Biochemistry*", 3rd Ed (2005)

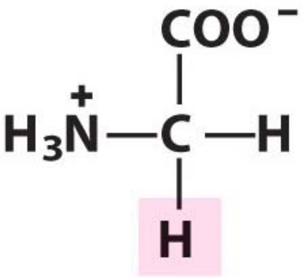
Amino acids found in proteins

- Cα is α-carboxyl & α-amino
- configuration L-
- Only 20 can be incorporated into proteins.
    - ✓ Proteins can have amino acids derivatives, but those modifications are integrated after protein synthesis.

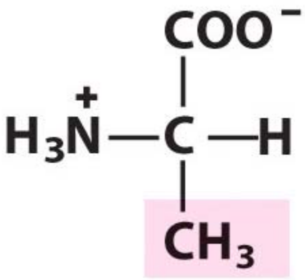
Classification of amino acids

- Properties of group R
  - ✓ Non-polar
  - ✓ Polar, uncharged
  - ✓ Polar, charged
    - Basic
    - Acids
- Synthesis
  - ✓ Essentials
  - ✓ Non- Essentials

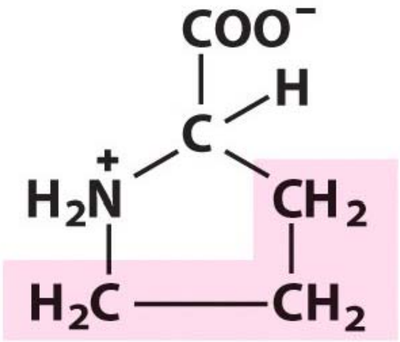
Nonpolar, aliphatic R groups



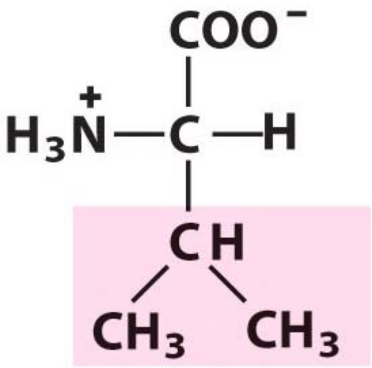
Glycine



Alanine

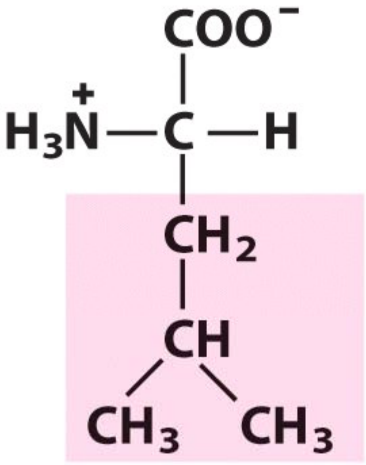


Proline

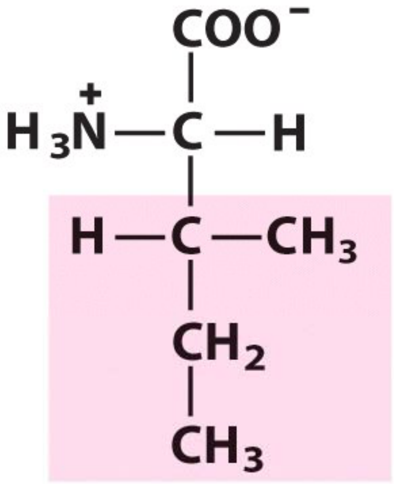


Valine

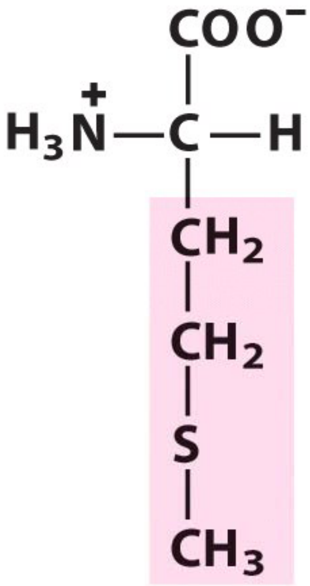
Devlin:
Fig 3.3
Page 78



Leucine



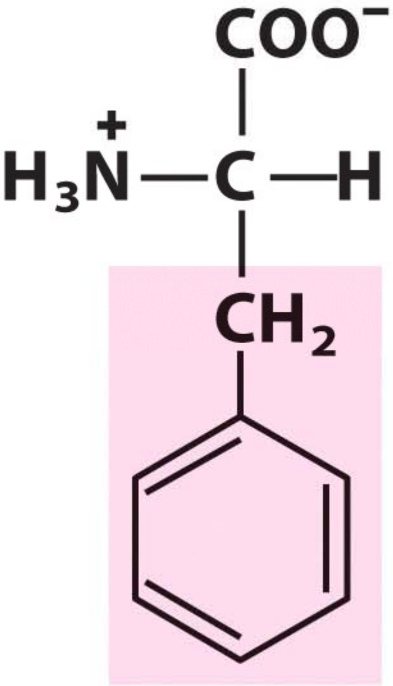
Isoleucine



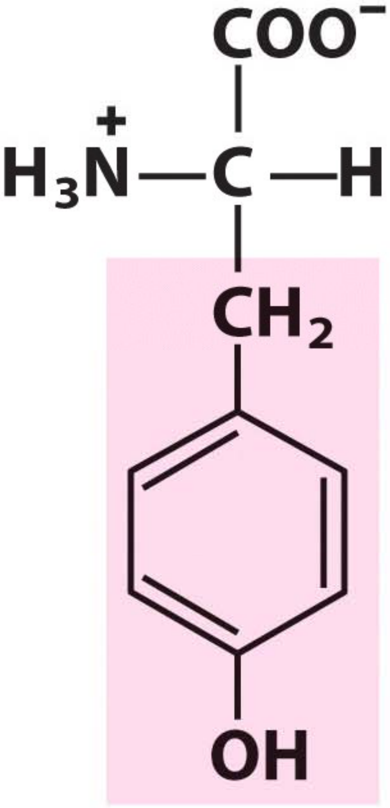
Methionine

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

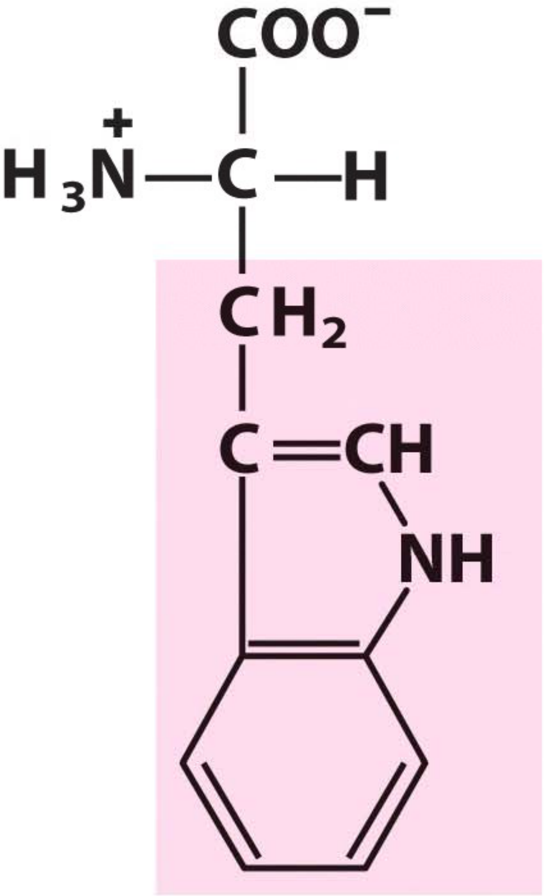
Aromatic R groups



Phenylalanine



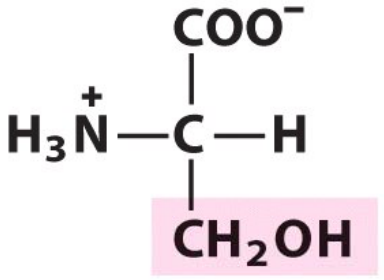
Tyrosine



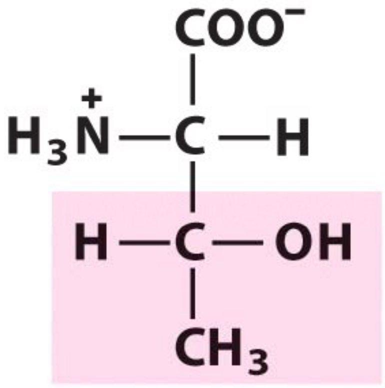
Tryptophan

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

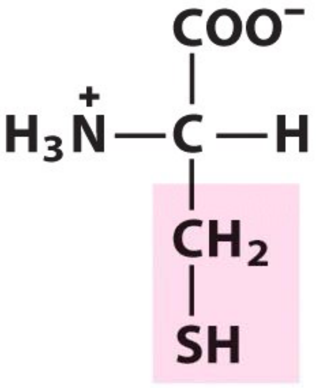
Polar, uncharged R groups



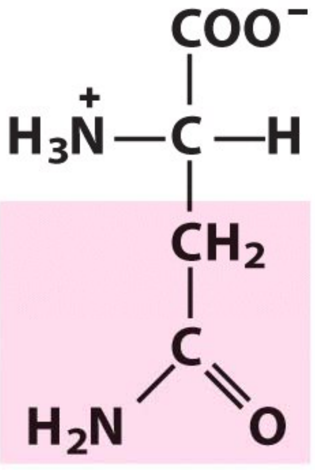
Serine



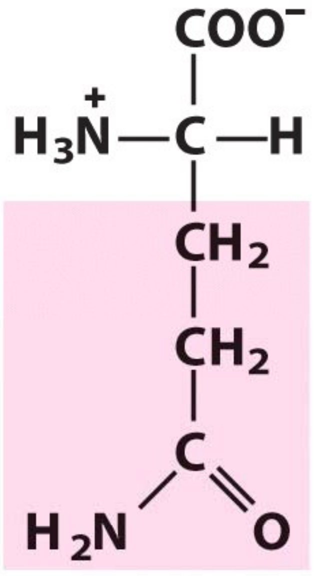
Threonine



Cysteine



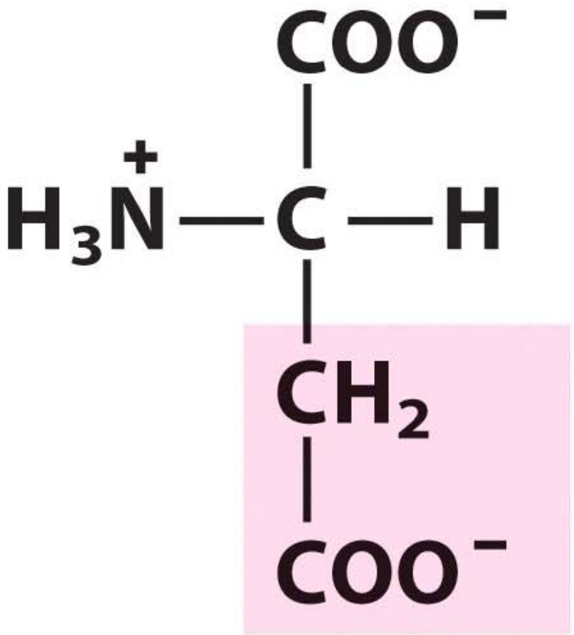
Asparagine



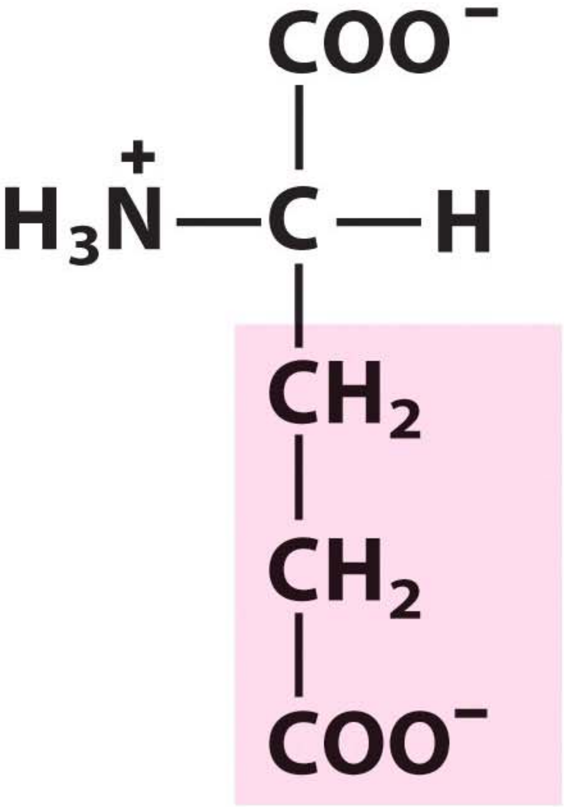
Glutamine

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

Negatively charged R groups



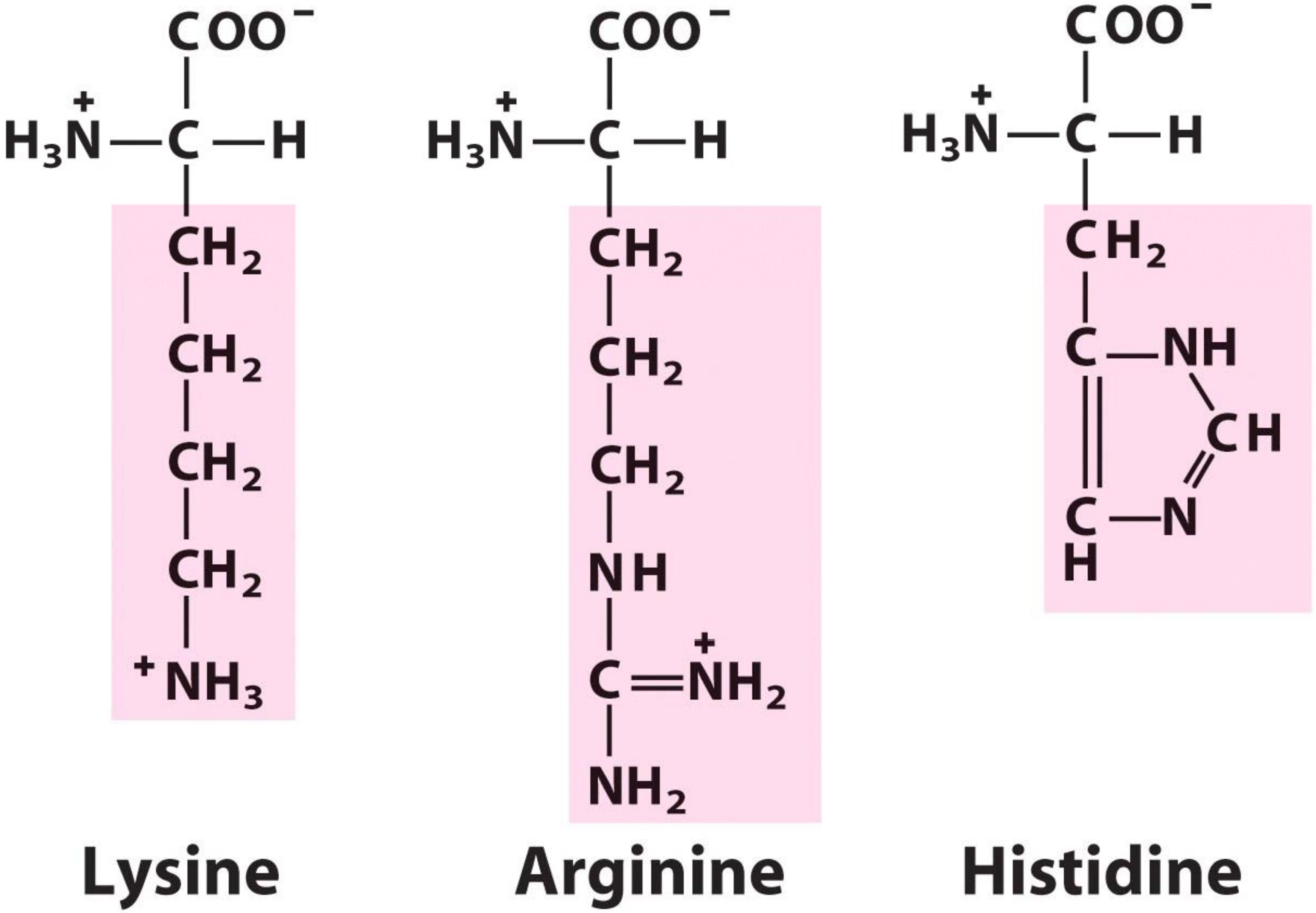
Aspartate



Glutamate

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

Positively charged R groups



Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

Modified amino acids

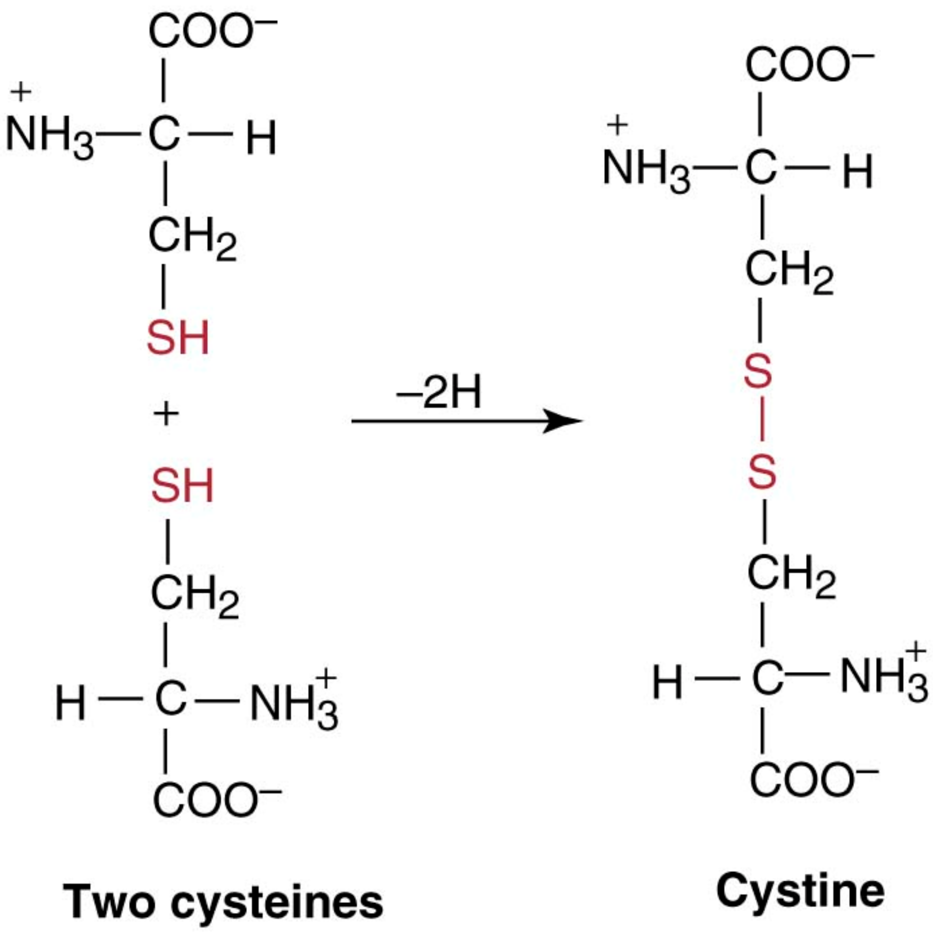


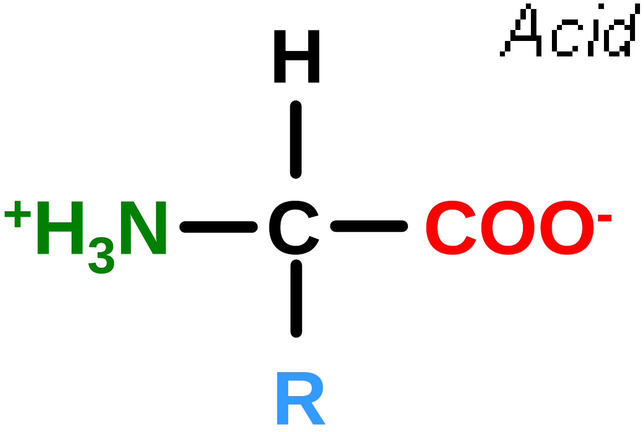
Figure 3.7. Cystine bond formation.

Cystinuria: incomplete reabsorption of cysteine during filtration in the kidney results in high cysteine and cystine in urine.

Excess cystine forms stones.

Devlin, *Biochemistry 6th ed* (2006)

Acid- Base Properties



$$^{+}H_3N-\overset{\displaystyle H}{\underset{\displaystyle R}{C}}-COO^-$$

$$\text{Acid} \rightleftharpoons \text{Base} + H^+$$

$$-COOH \rightleftharpoons -COO^- + H^+$$

$$-NH_3^+ \rightleftharpoons -NH_2 + H^+$$

Acid- Base Properties

Zwitterion

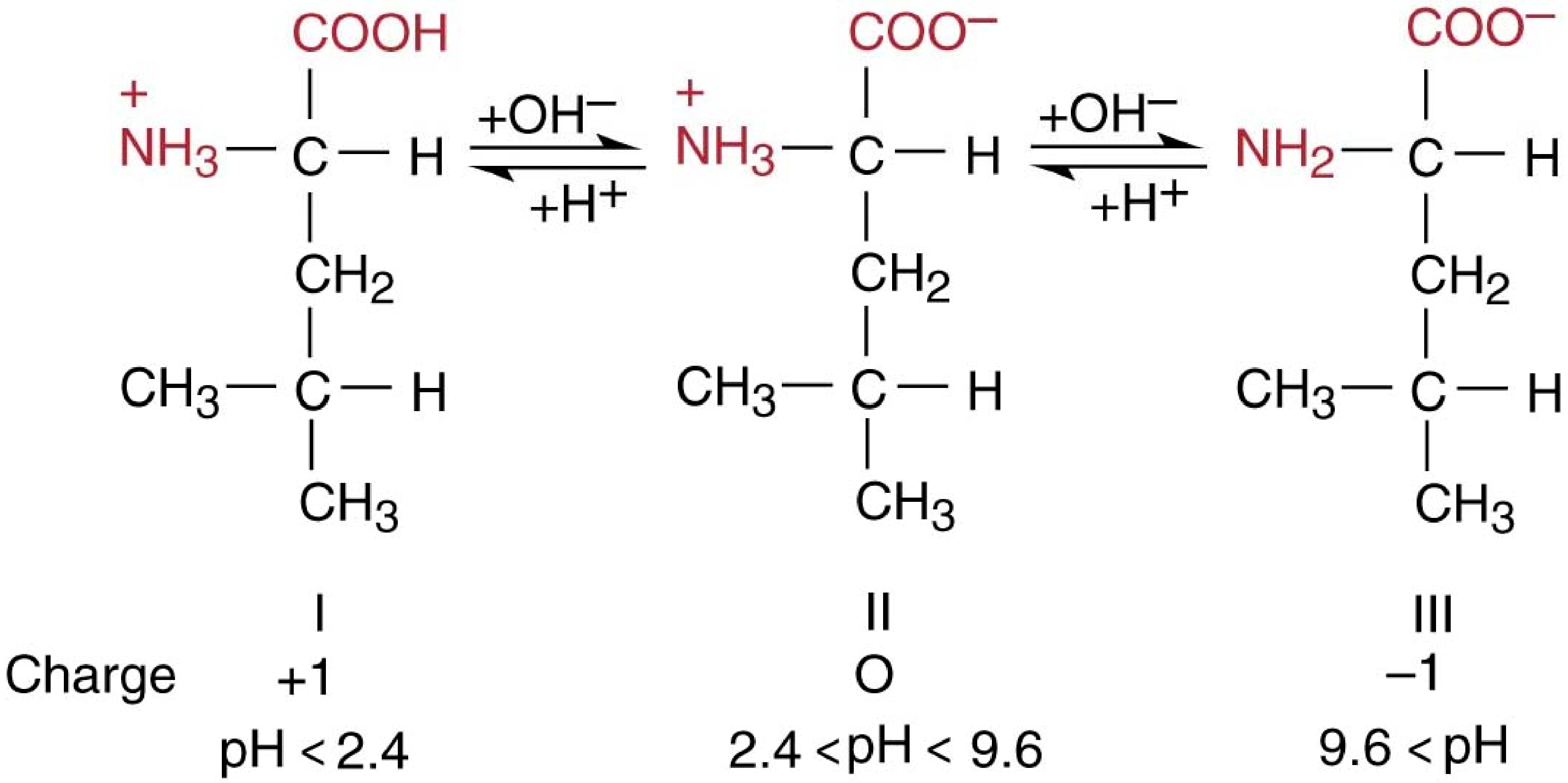


Figure 3.14. Ionic forms of leucine.

Devlin, *Biochemistry 6th ed* (2006)

Acid- Base Properties

Henderson-Hasselbalch

$$pH = pK_a + \log\frac{[\text{conjugate base}]}{[\text{conjugate acid}]}$$

- pK_a = pH acid is 50% deprotonated
- isoelectric point- (pI)
  - ✓ pH where the net charge of a molecule equals zero

$$pI = \frac{pK_aCOOH + pK_aNH_3^+}{2}$$

Acid- Base Properties

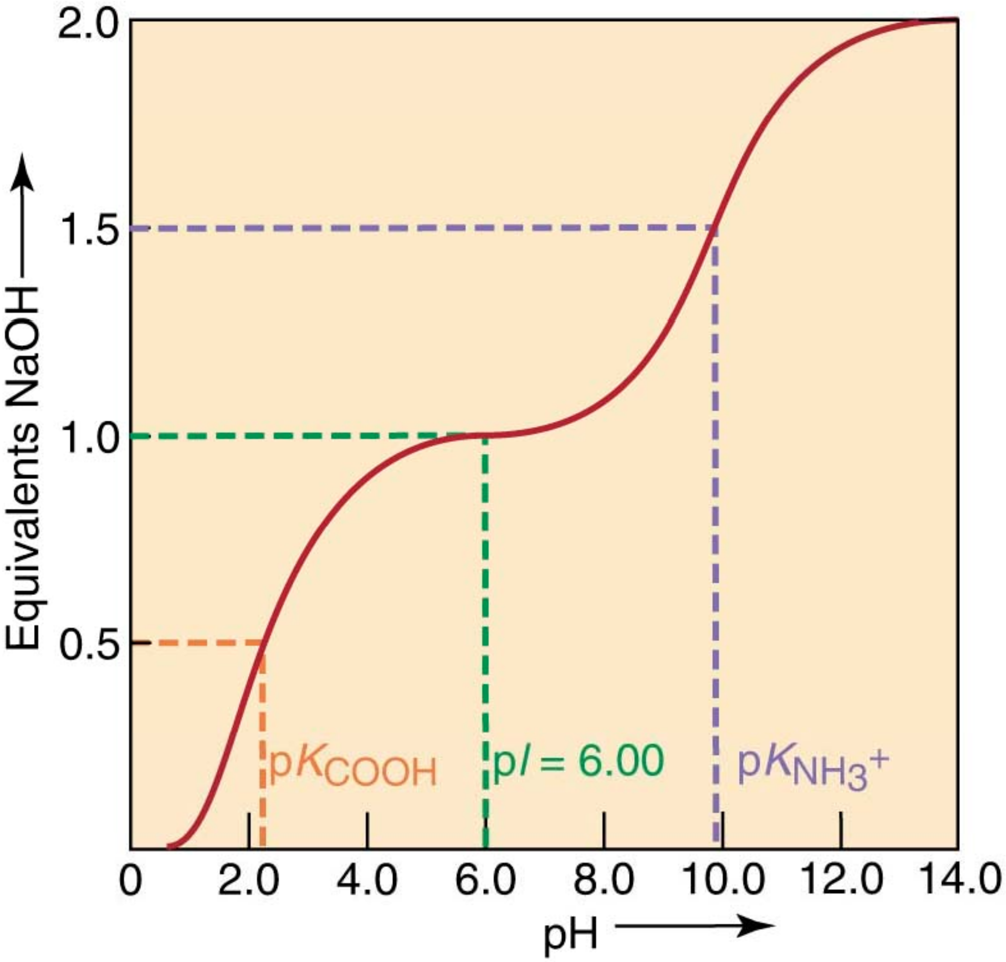
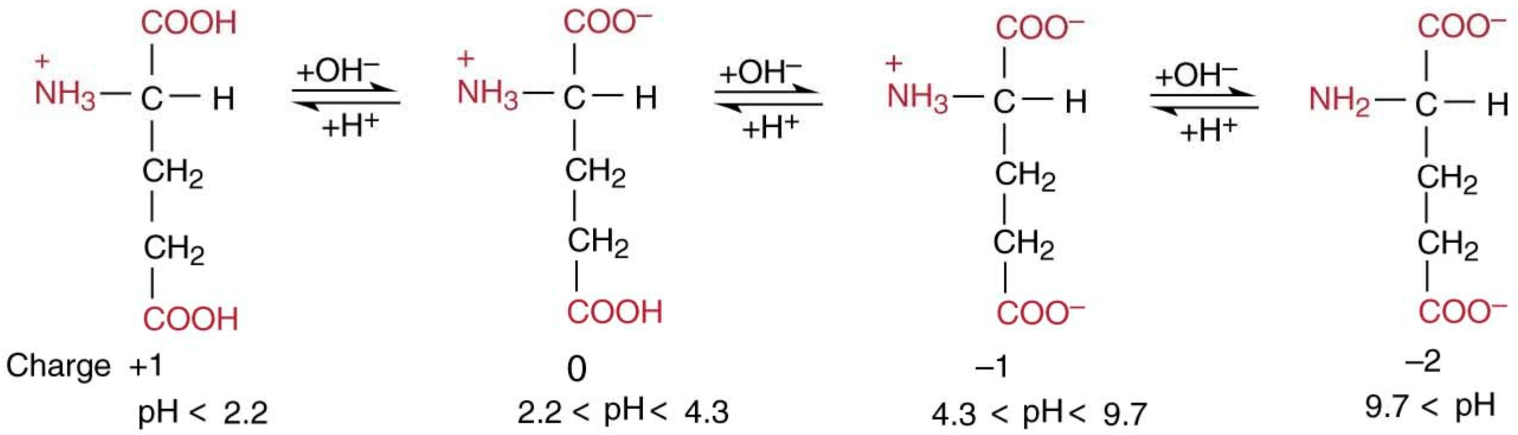
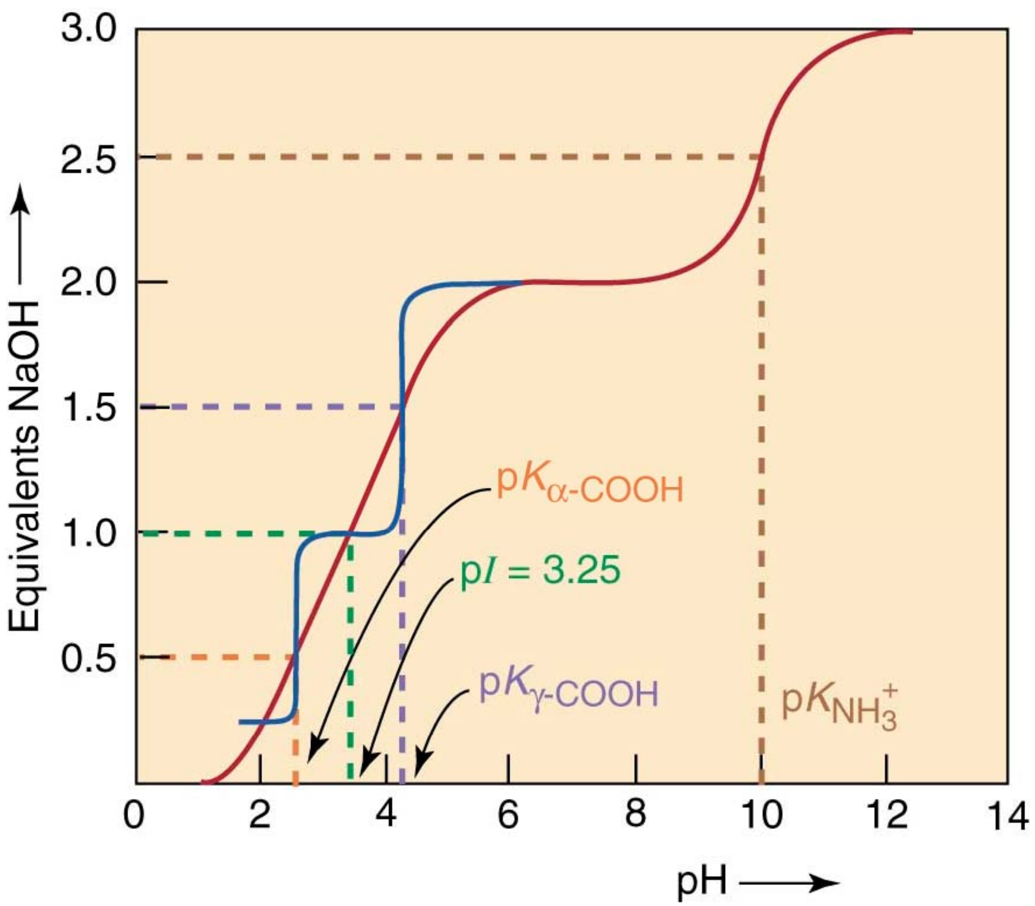


Figure 3.15. Titration curve of leucine.

Devlin, *Biochemistry 6th ed* (2006)





Glutamic Acid

Devlin, *Biochemistry 6th ed* (2006)

Acid- Base Properties

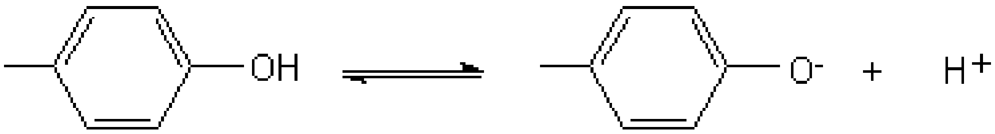
Aspartate / Glutamate

$$-COOH \rightleftharpoons -COO^- + H^+$$

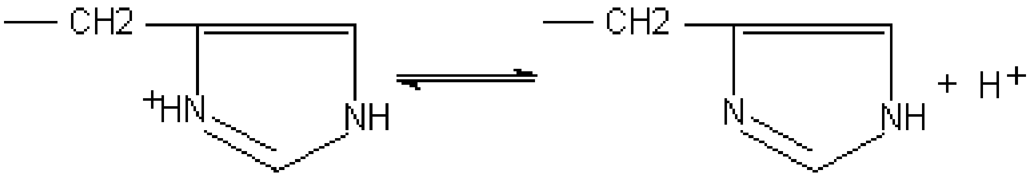
Cysteine

$$-SH \rightleftharpoons -S^- + H^+$$

Tyrosine



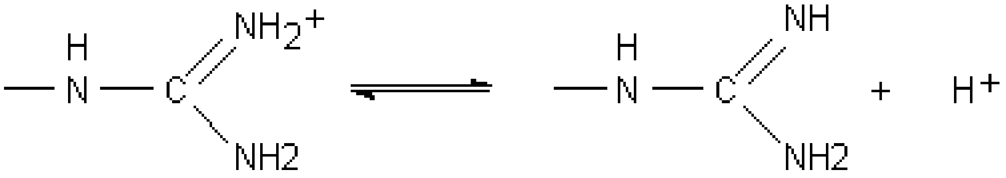
Histidine



Lysine

$$-NH_3^+ \rightleftharpoons -NH_2 + H^+$$

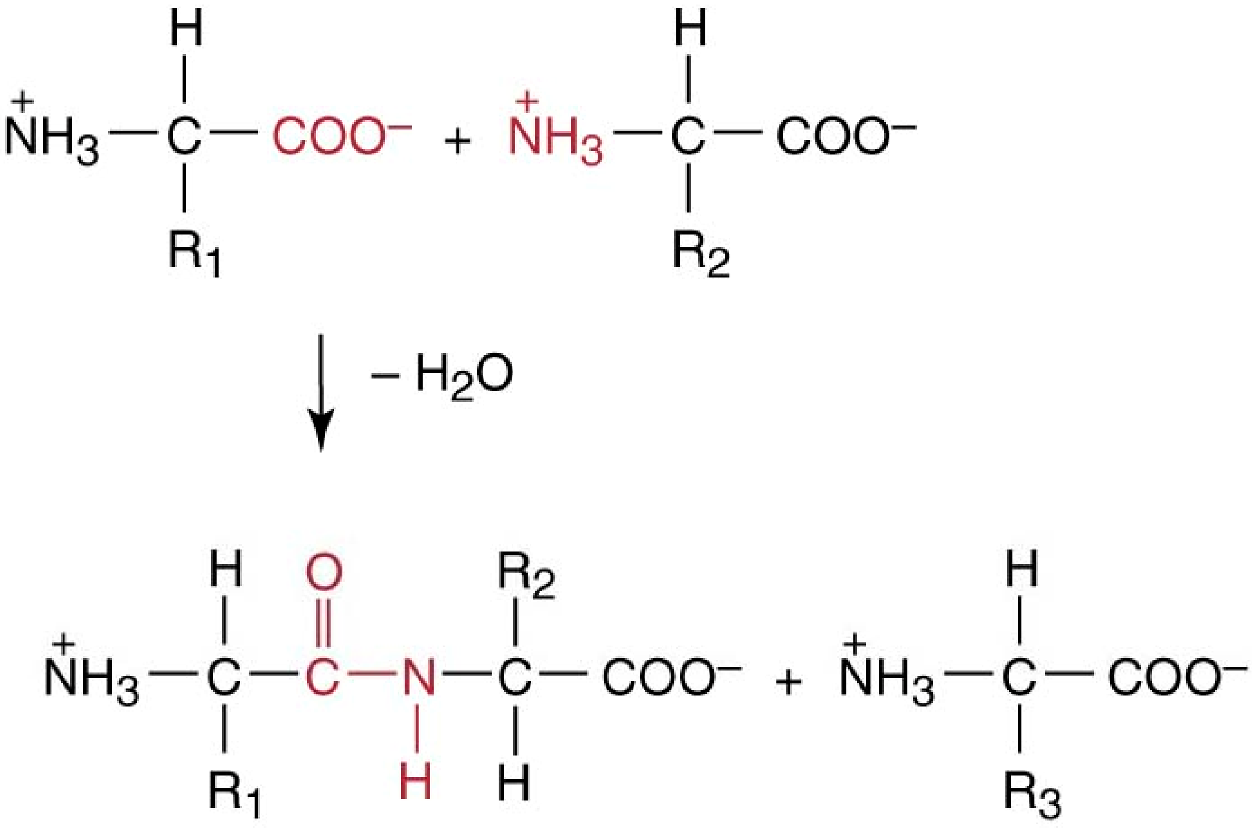
Arginine



http://www.agsci.ubc.ca/courses/fnh/410/protein/pkhis.gif

Devlin: Table 3.3

Peptide bond

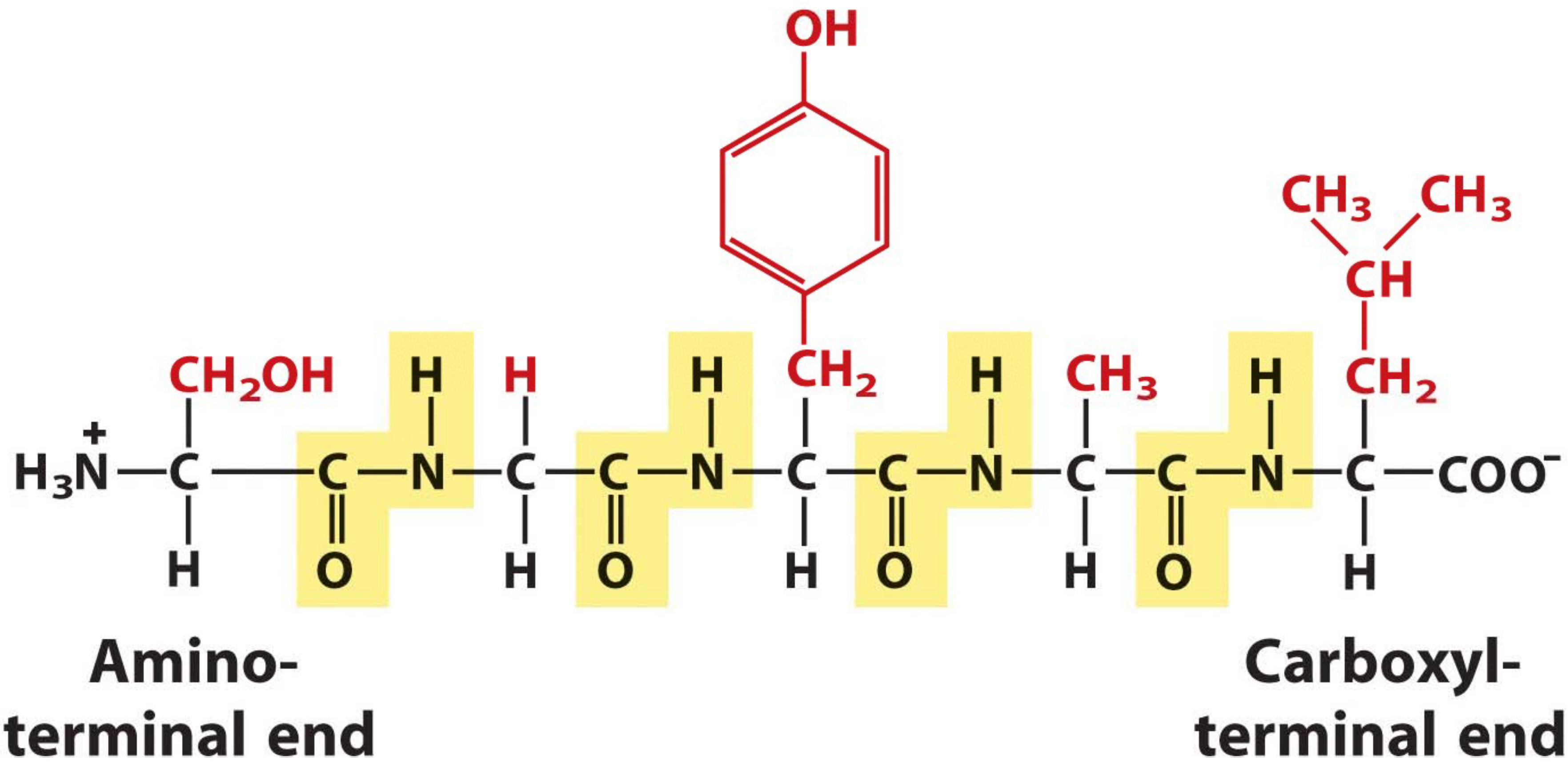


Dipeptide

Devlin, *Biochemistry 6th ed* (2006)

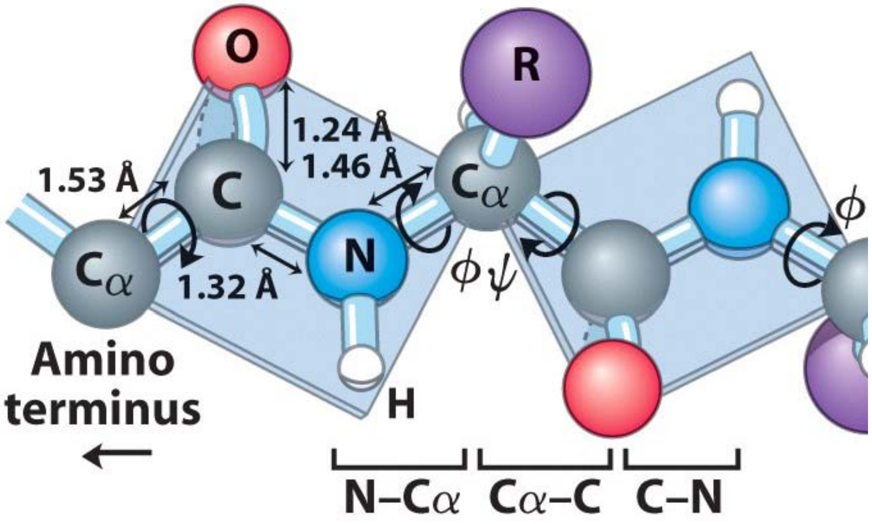
Peptide bond

serylglicyltyrosylalanylleucine

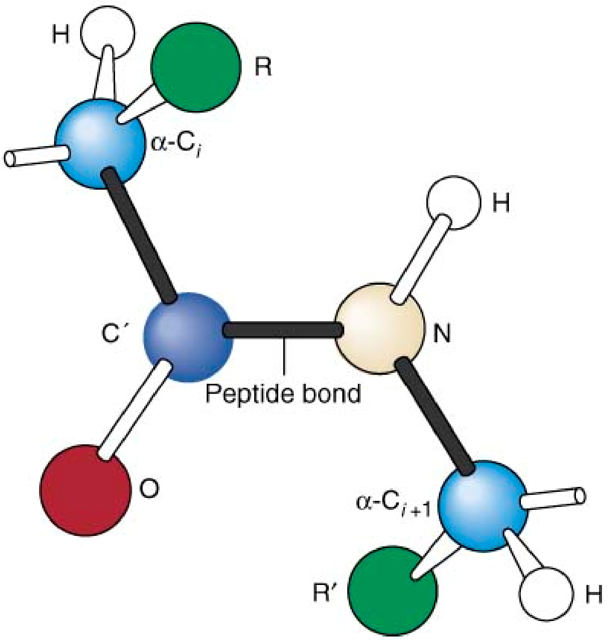


Lehninger, *Principles of Biochemistry* 4th Ed.

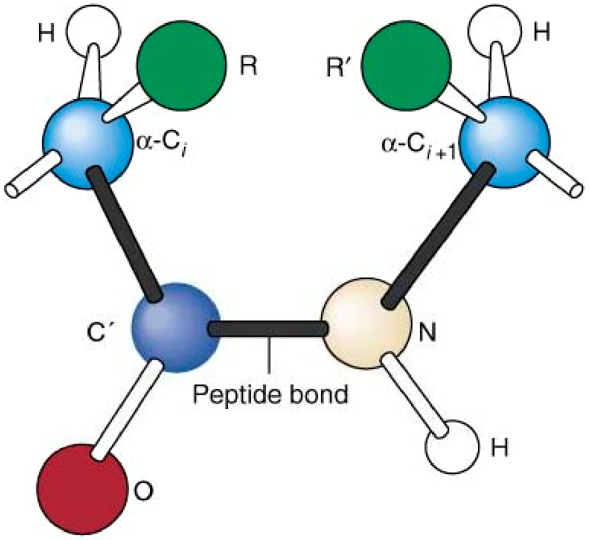
Peptide bond



- Partial double bound character
- Planar
- trans- configuration



(*a*) *trans* configuration

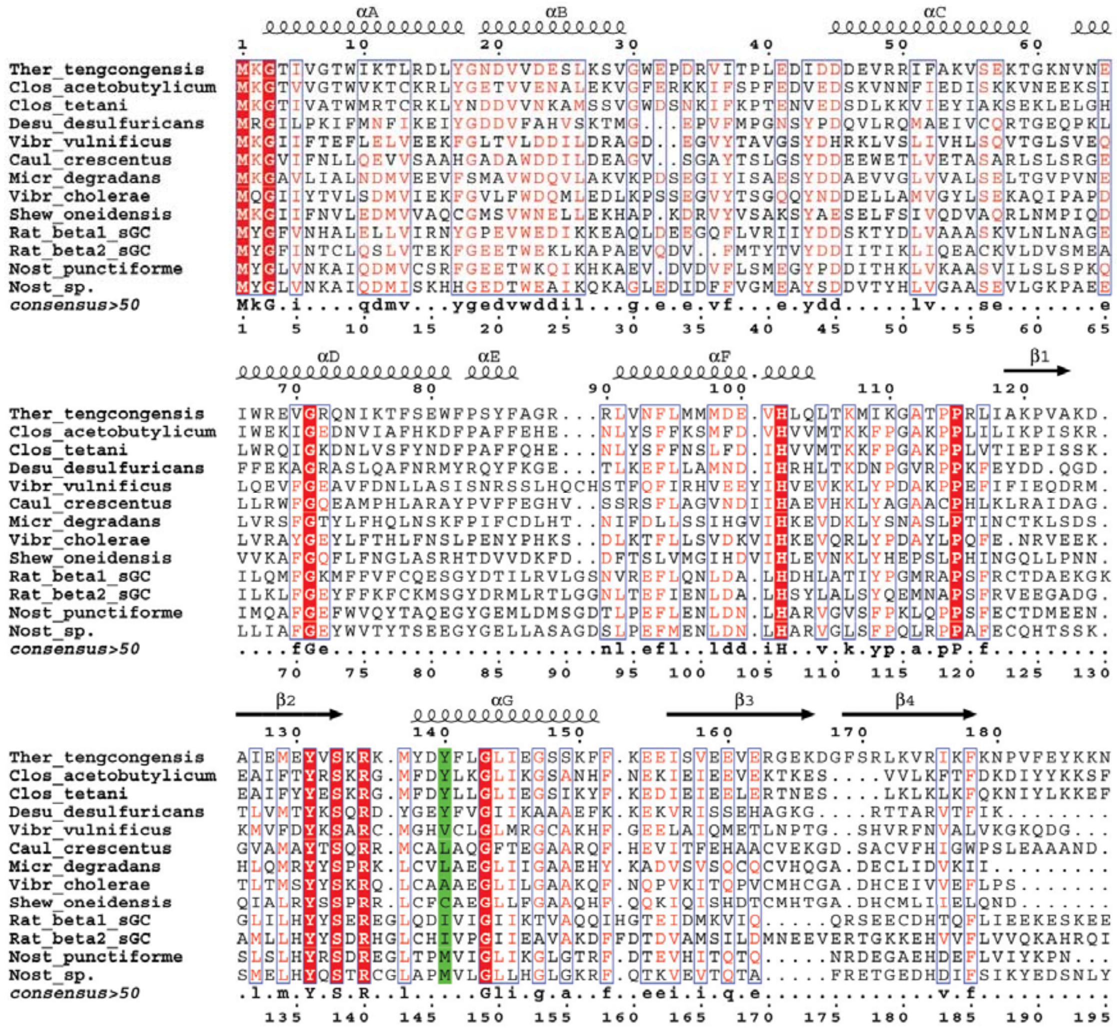


(*b*) *cis* configuration

Lehninger, *Principles of Biochemistry* 4th Ed.

Primary Structure

Sequence Alignment



```
Ther_tengcongensis    MKGTIVGTWIKTLRDLYGNDVVDESLKSVGWEPDRVITPLEDIDDDEVRRIFAKVSEKTGKNVNE
Clos_acetobutylicum   MKGTVVGTWVKTCKRLYGETVVENALEKVGFERKKIFSPFEDVEDSKVNNFIEDISKKVNEEKSI
Clos_tetani           MKGTIVATWMRTCRKLYNDDVVNKAMSSVGWDSNKIFKPTENVEDSDLKKVIEYIAKSEKLELGH
Desu_desulfuricans    MRGILPKIFMNFIKEIYGDDVFAHVSKTMG...EPVFMPGNSYPDQVLRQMAEIVCQRTGEQPKL
Vibr_vulnificus       MKGIIFTEFLELVEEKFGLTVLDDILDRAGD..EGVYTAVGSYDHRKLVSLIVHLSQVTGLSVEQ
Caul_crescentus       MKGVIFNLLQEVVSAAHGADAWDDILDEAGV..SGAYTSLGSYDDEEWETLVETASARLSLSRGE
Micr_degradans        MKGAVLIALNDMVEEVFSMAVWDQVLAKVKPDSEGIYISAESYDDAEVVGLVVALSELTGVPVNE
Vibr_cholerae         MQGIIYTVLSDMVIEKFGVLFWDQMLEDLKPSSEGVYTSGQQYNDDELLAMVGYLSEKAQIPAPD
Shew_oneidensis       MKGIIFNVLEDMVVAQCGMSVWNELLEKHAP.KDRVYVSAKSYAESELFSIVQDVAQRLNMPIQD
Rat_beta1_sGC         MYGFVNHALELLVIRNYGPEVWEDIKKEAQLDEEGQFLVRIIYDDSKTYDLVAAASKVLNLNAGE
Rat_beta2_sGC         MYGFINTCLQSLVTEKFGEETWEKLKAPAEVQDV..FMTYTVYDDIITIKLIQEACKVLDVSMEA
Nost_punctiforme      MYGLVNKAIQDMVCSRFGEETWKQIKHKAEV.DVDVFLSMEGYPDDITHKLVKAASVILSLSPKQ
Nost_sp.              MYGLVNKAIQDMISKHHGEDTWEAIKQKAGLEDIDFFVGMEAYSDDVTYHLVGAASEVLGKPAEE
consensus>50          MkG.i....qdmv...ygedvwddil...g.e.e.vf...e.ydd.....lv...se.......e

Ther_tengcongensis    IWREVGRQNIKTFSEWFPSYFAGR...RLVNFLMMMDE.VHLQLTKMIKGATPPRLIAKPVAKD.
Clos_acetobutylicum   IWEKIGEDNVIAFHKDFPAFFEHE...NLYSFFKSMFD.VHVVMTKKFPGAKPPLILIKPISKR.
Clos_tetani           LWRQIGKDNLVSFYNDFPAFFQHE...NLYSFFNSLFD.IHVVMTKKFPGAKPPLVTIEPISSK.
Desu_desulfuricans    FFEKAGRASLQAFNRMYRQYFKGE...TLKEFLLAMND.IHRHLTKDNPGVRPPKFEYDD.QGD.
Vibr_vulnificus       LQEVFGEAVFDNLLASISNRSSLHQCHSTFQFIRHVEEYIHVEVKKLYPDAKPPEFIFIEQDRM.
Caul_crescentus       LLRWFGQEAMPHLARAYPVFFEGHV..SSRSFLAGVNDIIHAEVHKLYAGAACPHLKLRAIDAG.
Micr_degradans        LVRSFGTYLFHQLNSKFPIFCDLHT..NIFDLLSSIHGVIHKEVDKLYSNASLPTINCTKLSDS.
Vibr_cholerae         LVRAYGEYLFTHLFNSLPENYPHKS..DLKTFLLSVDKVIHKEVQRLYPDAYLPQFE.NRVEEK.
Shew_oneidensis       VVKAFGQFLFNGLASRHTDVVDKFD..DFTSLVMGIHDVIHLEVNKLYHEPSLPHINGQLLPNN.
Rat_beta1_sGC         ILQMFGKMFFVFCQESGYDTILRVLGSNVREFLQNLDA.LHDHLATIYPGMRAPSFRCTDAEKGK
Rat_beta2_sGC         ILKLFGEYFFKFCKMSGYDRMLRTLGGNLTEFIENLDA.LHSYLALSYQEMNAPSFRVEEGADG.
Nost_punctiforme      IMQAFGEFWVQYTAQEGYGEMLDMSGDTLPEFLENLDN.LHARVGVSFPKLQPPSFECTDMEEN.
Nost_sp.              LLIAFGEYWVTYTSEEGYGELLASAGDSLPEFMENLDN.LHARVGLSFPQLRPPAFECQHTSSK.
consensus>50          ....fGe....................nl.efl..ldd.iH..v.k.yp.a.pP.f.........

Ther_tengcongensis    AIEMEYVSKRK.MYDYFLGLIEGSSKFF.KEEISVEEVERGEKDGFSRLKVRIKFKNPVFEYKKN
Clos_acetobutylicum   EAIFTYRSKRG.MFDYLKGLIKGSANHF.NEKIEIEEVEKTKES....VVLKFTFDKDIYYKKSF
Clos_tetani           EAIFYYESKRG.MFDYLLGLIEGSIKYF.KEDIEIEELERTNES....LKLKLKFQKNIYLKKEF
Desu_desulfuricans    TLVMTYKSQRD.YGEYFVGIIKAAAEFK.KEKVRISSEHAGKG....RTTARVTFIK........
Vibr_vulnificus       KMVFDYKSARC.MGHVCLGLMRGCAKHF.GEELAIQMETLNPTG..SHVRFNVALVKGKQDG...
Caul_crescentus       GVAMAYTSQRR.MCALAQGFTEGAARQF.HEVITFEHAACVEKGD.SACVFHIGWPSLEAAAND.
Micr_degradans        HLQMRYYSPRK.LCVLAEGLIIGAAEHY.KADVSVSQCQCVHQGA.DECLIDVKII.........
Vibr_cholerae         TLTMSYYSKRQ.LCAAAEGLILGAAKQF.NQPVKITQPVCMHCGA.DHCEIVVEFLPS.......
Shew_oneidensis       QIALRYSSPRR.LCFCAEGLLFGAAQHF.QQKIQISHDTCMHTGA.DHCMLIIELQND.......
Rat_beta1_sGC         GLILHYYSEREGLQDIVIGIIKTVAQQIHGTEIDMKVIQ.....QRSEECDHTQFLIEEKESKEE
Rat_beta2_sGC         AMLLHYYSDRHGLCHIVPGIIEAVAKDFFDTDVAMSILDMNEEVERTGKKEHVVFLVVQKAHRQI
Nost_punctiforme      SLSLHYRSDREGLTPMVIGLIKGLGTRF.DTEVHITQTQ.....NRDEGAEHDEFLVIYKPN...
Nost_sp.              SMELHYQSTRCGLAPMVLGLLHGLGKRF.QTKVEVTQTA.....FRETGEDHDIFSIKYEDSNLY
consensus>50          .l.m.Y.S.R..l.....Gli.g.a..f..eei.i.q.e.............v.f..........
```

- Analogy – seq. that are structurally similar but no evolutionary relationship has been demonstrated.
- Homology (homologous proteins) – aa sequences are highly alignable (proteins belong to the same family) – evolve from same gene and have similar functions
  - ✓ Paralog – two proteins within a family are present in the same species.
  - ✓ Ortholog – homologs from different species

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

Human and chimp myoglobin are ~100 % Identical

```
               10        20        30        40        50        60
Chimpa MGLSDGEWQLVLNVWGKVEADIPGHGQEVLIRLFKGHPETLEKFDKFKHLKSEDEMKASE
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Human  MGLSDGEWQLVLNVWGKVEADIPGHGQEVLIRLFKGHPETLEKFDKFKHLKSEDEMKASE
               10        20        30        40        50        60

               70        80        90       100       110       120
Chimpa DLKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIPVKYLEFISECIIQVLHSKH
       ::::::::::::::::::::::::::::::::::::::::::::::::::::::::.:::
Human  DLKKHGATVLTALGGILKKKGHHEAEIKPLAQSHATKHKIPVKYLEFISECIIQVLQSKH
               70        80        90       100       110       120

              130       140       150
Chimpa PGDFGADAQGAMNKALELFRKDMASNYKELGFQG
       ::::::::::::::::::::::::::::::::::
Human  PGDFGADAQGAMNKALELFRKDMASNYKELGFQG
              130       140       150
```

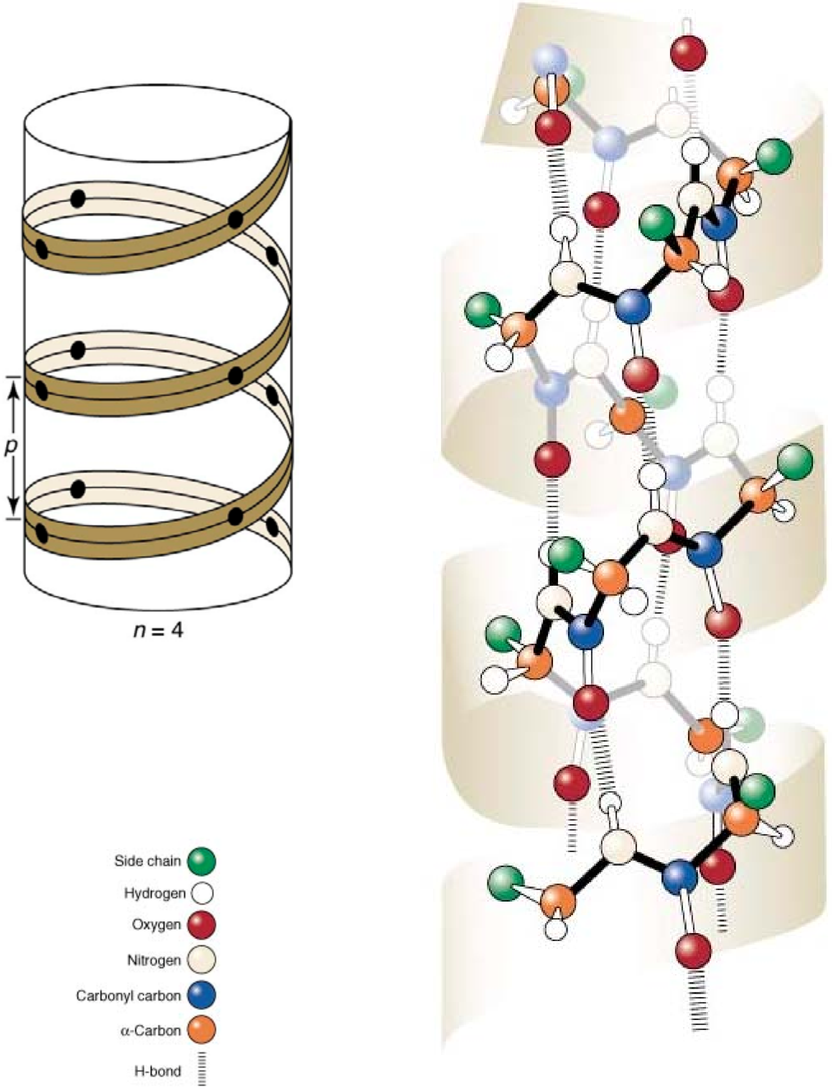
95% of their DNA sequence, and 99% of coding DNA sequences are in common

Secondary Structure

- Refers to local conformation of some part of the polypeptide
- Due to the partial double character of the peptide bound
- Stabilized mainly by H-bonds

- α-helix
- β-sheet
- β-turn

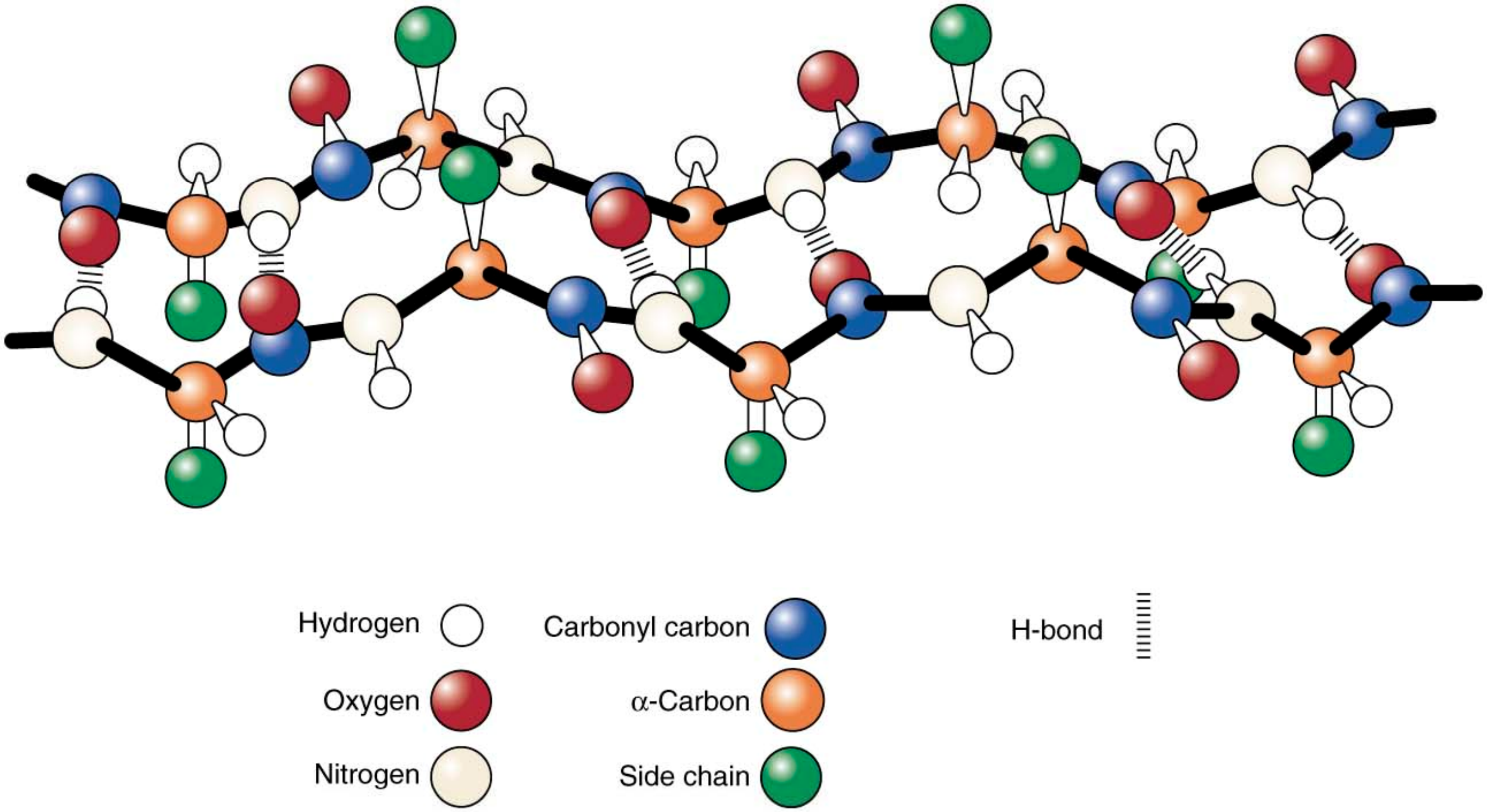
Secondary Structure



- alpha - helix
  - ✓ 3.6 aa per turn
  - ✓ Stabilized by H-bonds
  - ✓ Right handed

Secondary Structure

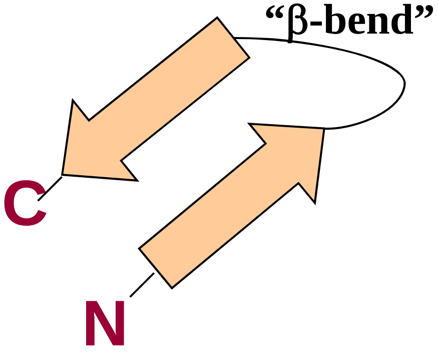
- β-pleated sheets



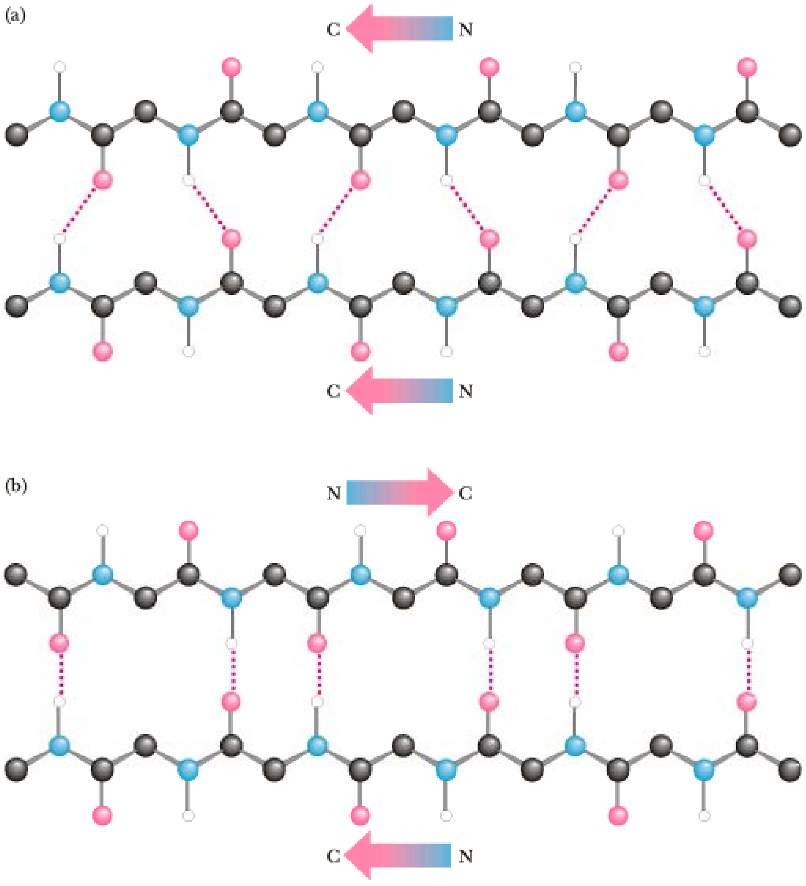
Devlin, *Biochemistry 6th ed* (2006)

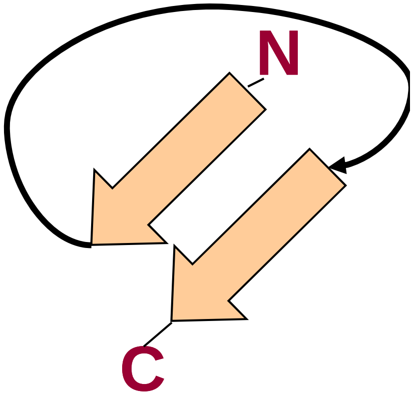
Secondary Structure

- β-pleated sheets



anti-parallel



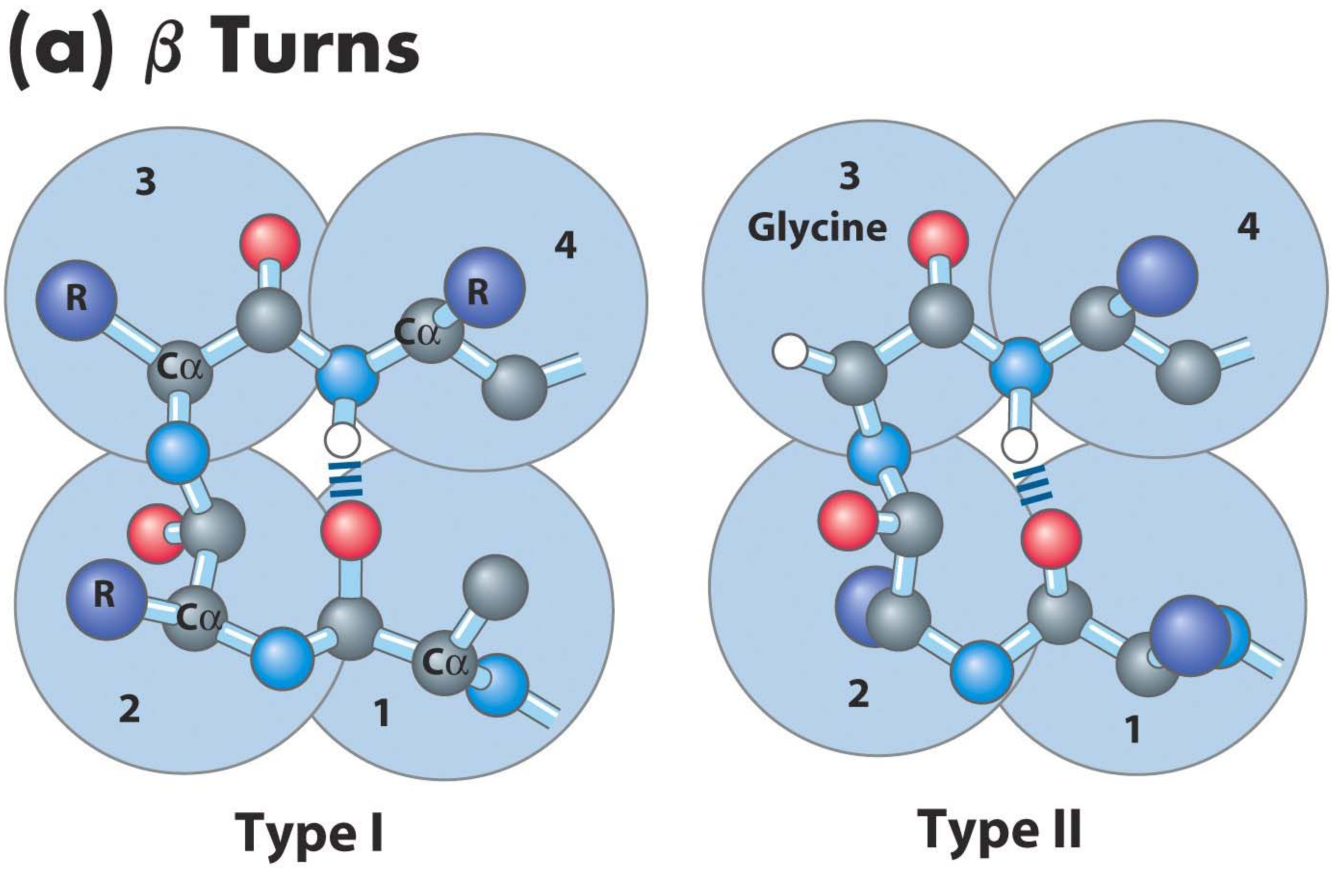


parallel

Garrett & Grisham, ***Biochemistry*** 2nd Ed.

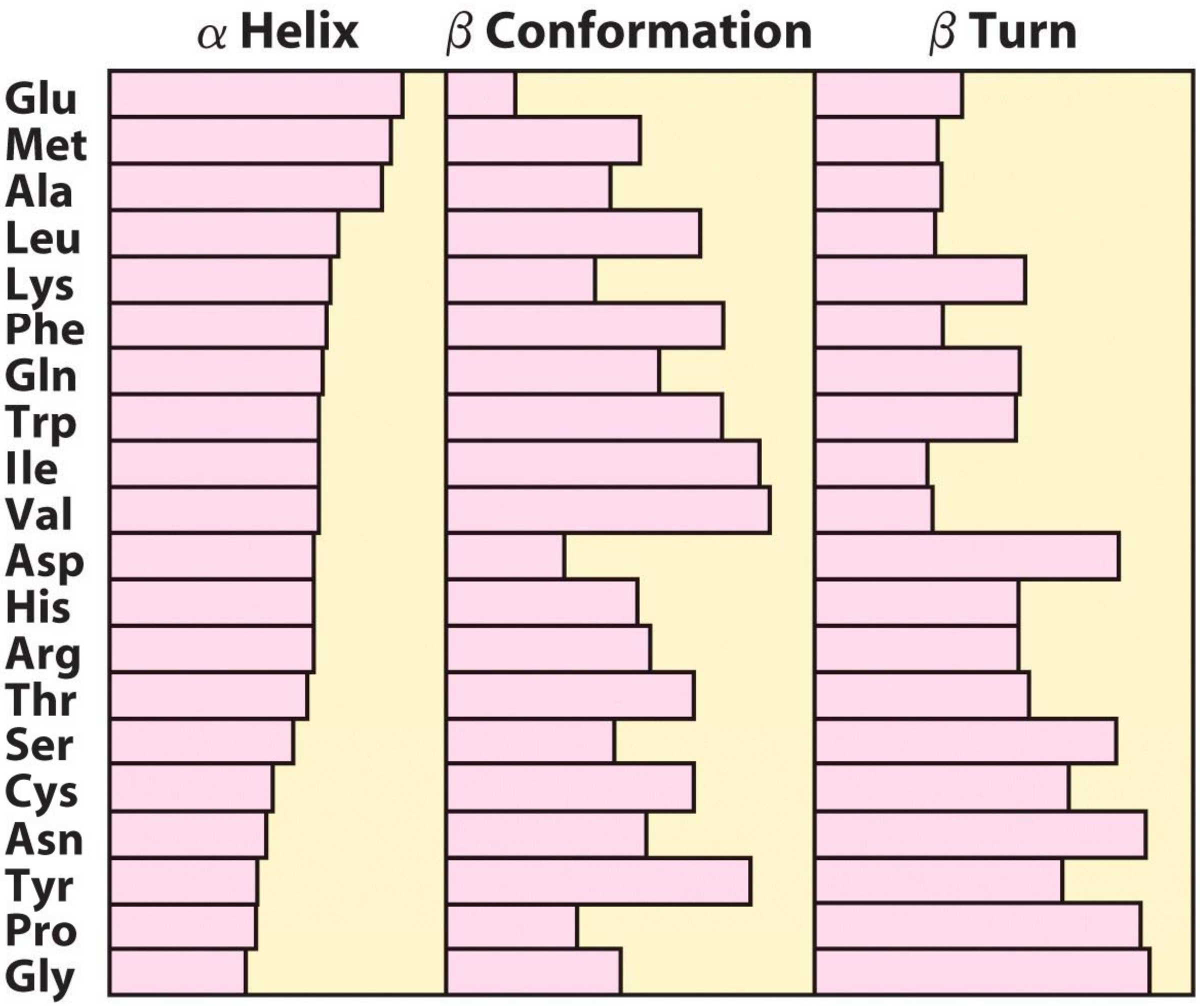
Secondary Structure

- Turns / Loops
  - ✓ Gly & Pro



Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

Amino acids occurrence in secondary structures

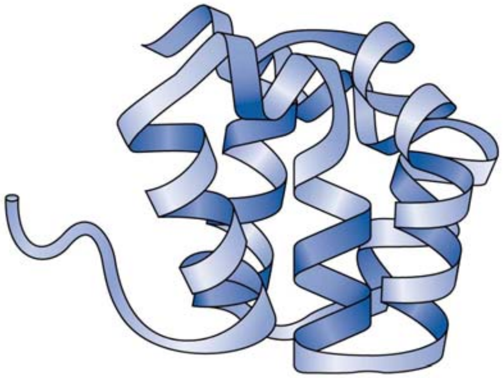


Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

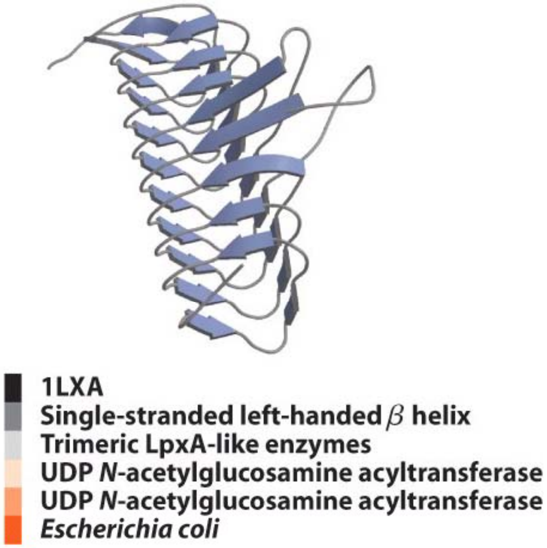
Ultrasecondary Structures

- Motifs / Folds – refers to particularly stable arrangements of several elements of secondary structure and the connections between them.
- Structural Classification

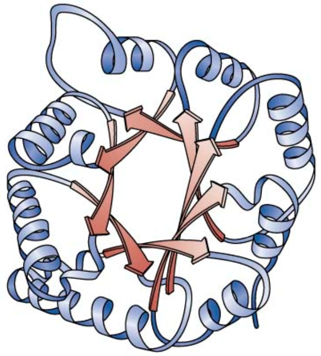
All α



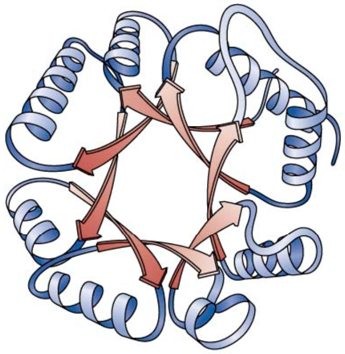
All β



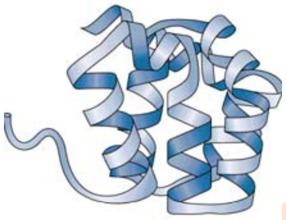
α,β



Triose Phosphate Isomerase

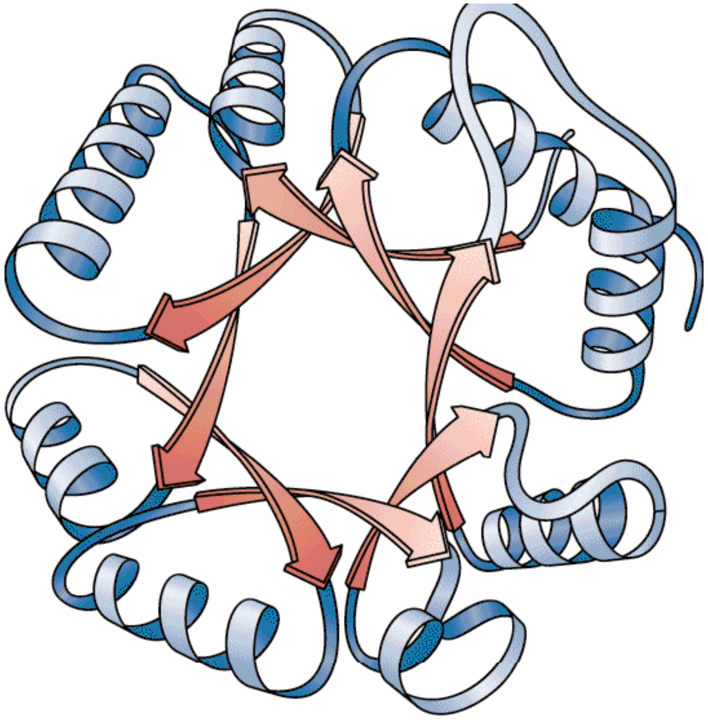


Pyruvate Kinase domain 1



Tertiary Structure

- Refers to 3D conformation (location of each atom in space)
- Stabilized by:
  - ✓ Disulfide bonds (covalente)
  - ✓ Non-covalent interactions
    - Hydrophobic
    - H-bonds
    - Ionic



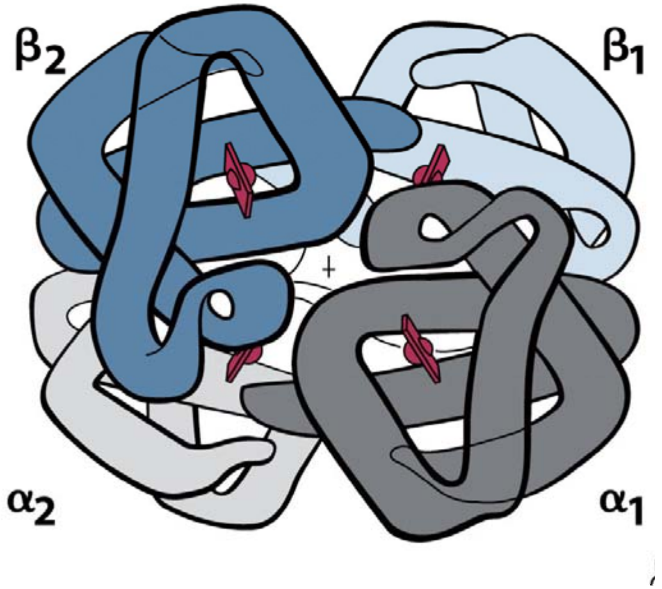
Pyruvate Kinase domain 1

Devlin, *Biochemistry 6th ed* (2006)

Quaternary Structure

- Non-covalent assemblies of two or more monomer subunits.
- Sub-units may work independently of cooperatively

Collagen, a fibrous protein

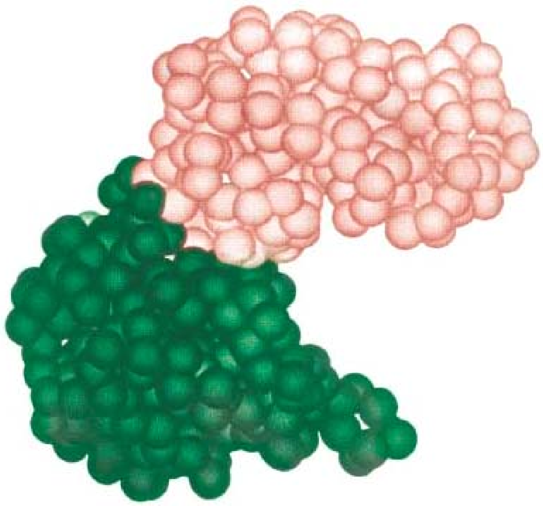


Garrett & Grisham, ***Biochemistry*** 2nd Ed.

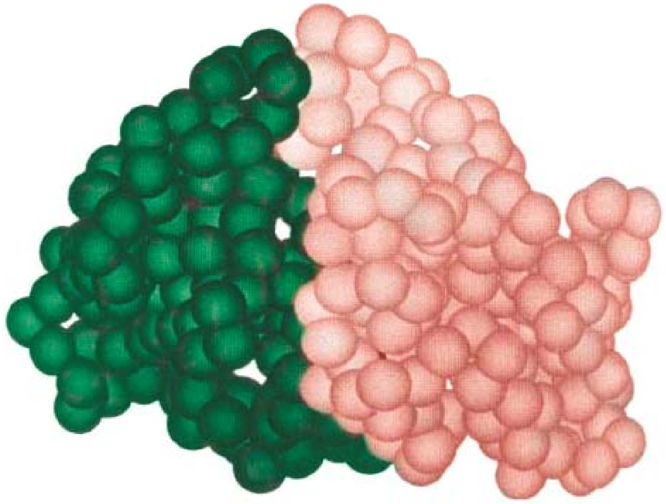
Voet, Voet & Pratt, *Fundamentals of Biochemistry* 2nd ed (2006)

Protein Domains

- Domains = Globular units within proteins
  - Different domains have different functions
  - Small proteins usually have one domain

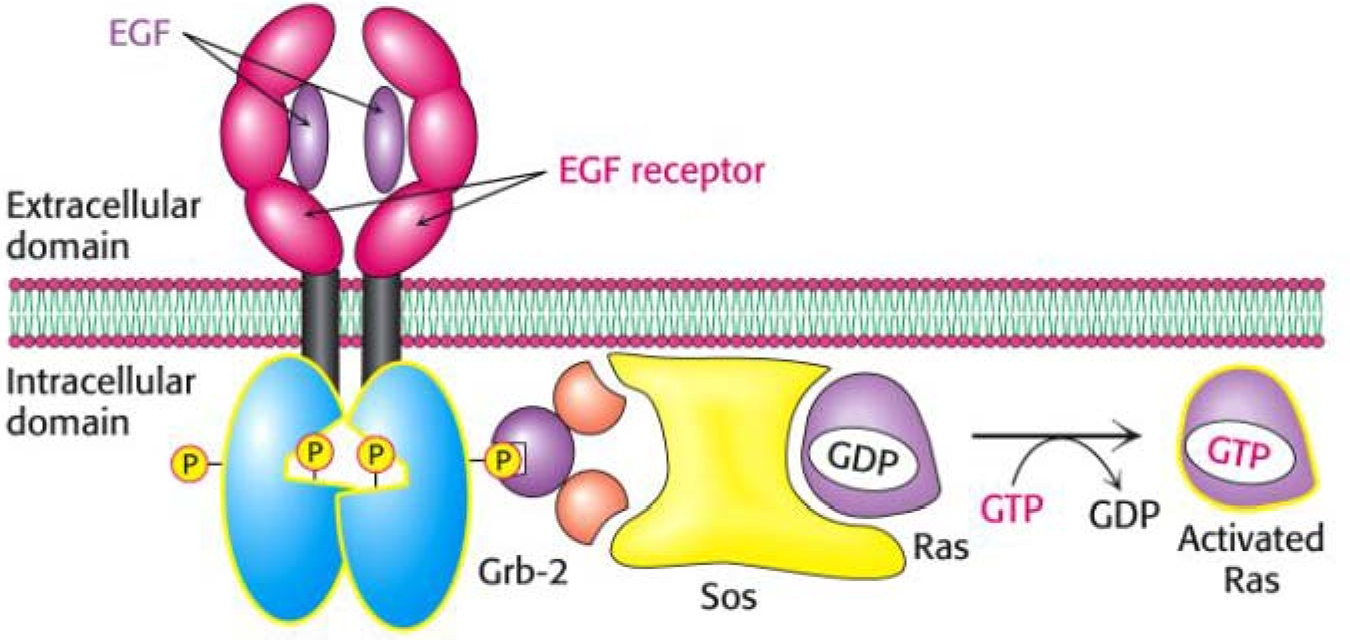


(a)

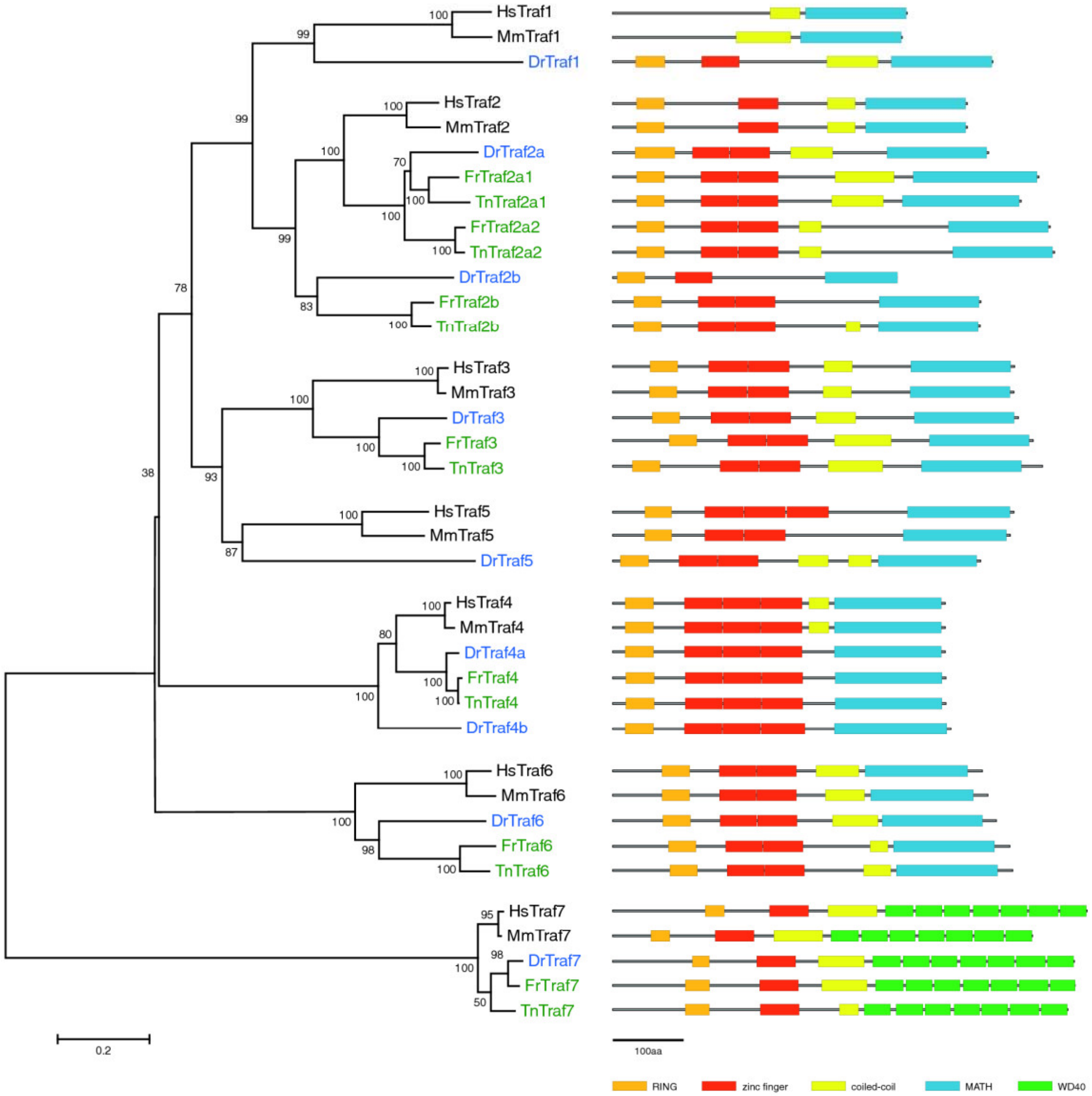


(b)

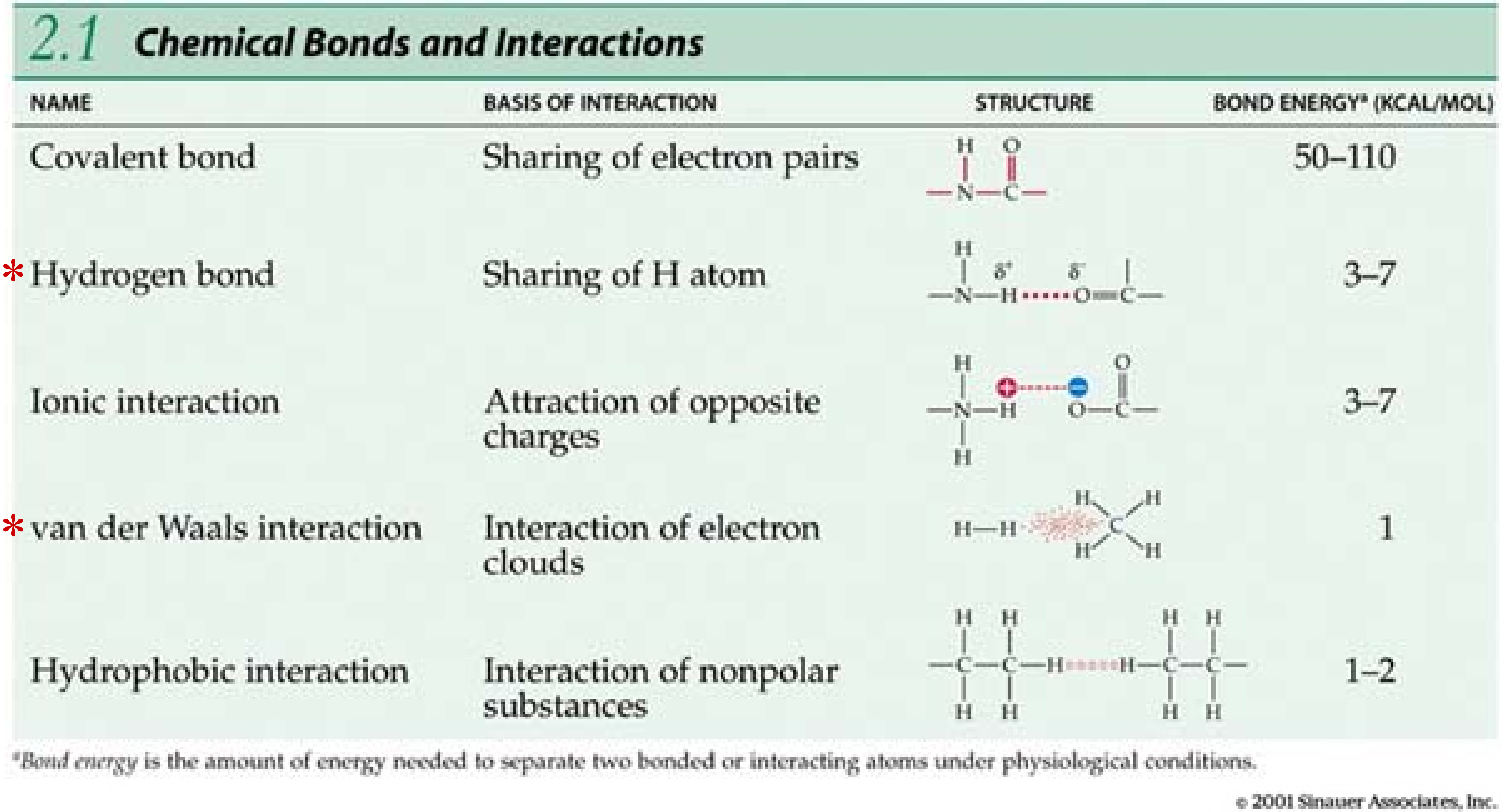
Domains


EGF
EGF receptor
Extracellular domain
Intracellular domain
P
P
P
P
Grb-2
Sos
GDP
Ras
GTP
GDP
GTP
Activated Ras

Domains


HsTraf1
MmTraf1
DrTraf1
HsTraf2
MmTraf2
DrTraf2a
FrTraf2a1
TnTraf2a1
FrTraf2a2
TnTraf2a2
DrTraf2b
FrTraf2b
TnTraf2b
HsTraf3
MmTraf3
DrTraf3
FrTraf3
TnTraf3
HsTraf5
MmTraf5
DrTraf5
HsTraf4
MmTraf4
DrTraf4a
FrTraf4
TnTraf4
DrTraf4b
HsTraf6
MmTraf6
DrTraf6
FrTraf6
TnTraf6
HsTraf7
MmTraf7
DrTraf7
FrTraf7
TnTraf7
0.2
100aa
RING
zinc finger
coiled-coil
MATH
WD40

Forces that stabilize protein structure



2.1 Chemical Bonds and Interactions

NAME	BASIS OF INTERACTION	STRUCTURE	BOND ENERGY* (KCAL/MOL)
Covalent bond	Sharing of electron pairs		50–110
* Hydrogen bond	Sharing of H atom		3–7
Ionic interaction	Attraction of opposite charges		3–7
* van der Waals interaction	Interaction of electron clouds		1
Hydrophobic interaction	Interaction of nonpolar substances		1–2

*Bond energy is the amount of energy needed to separate two bonded or interacting atoms under physiological conditions.

Hydrogen bonds

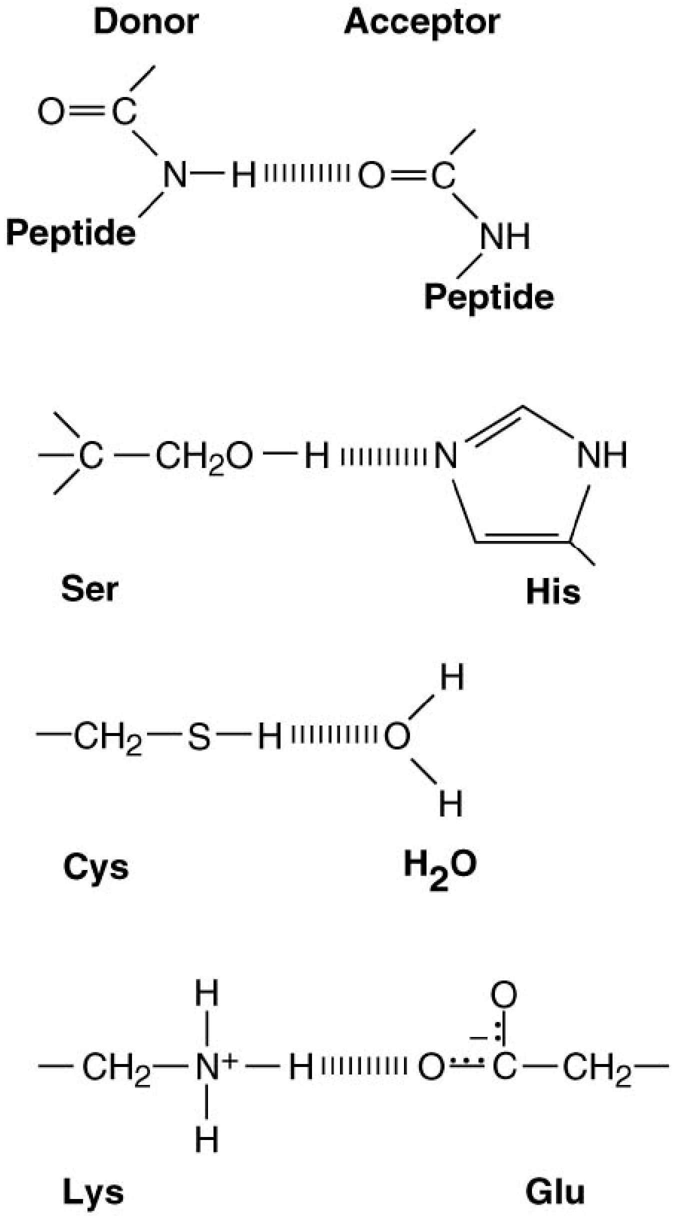
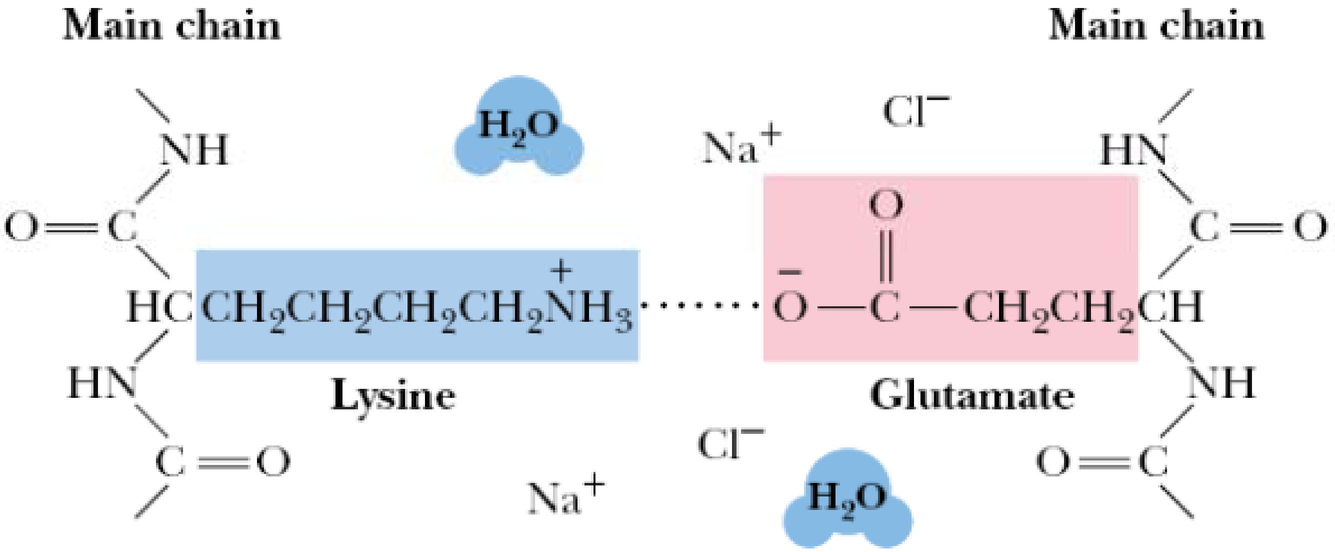


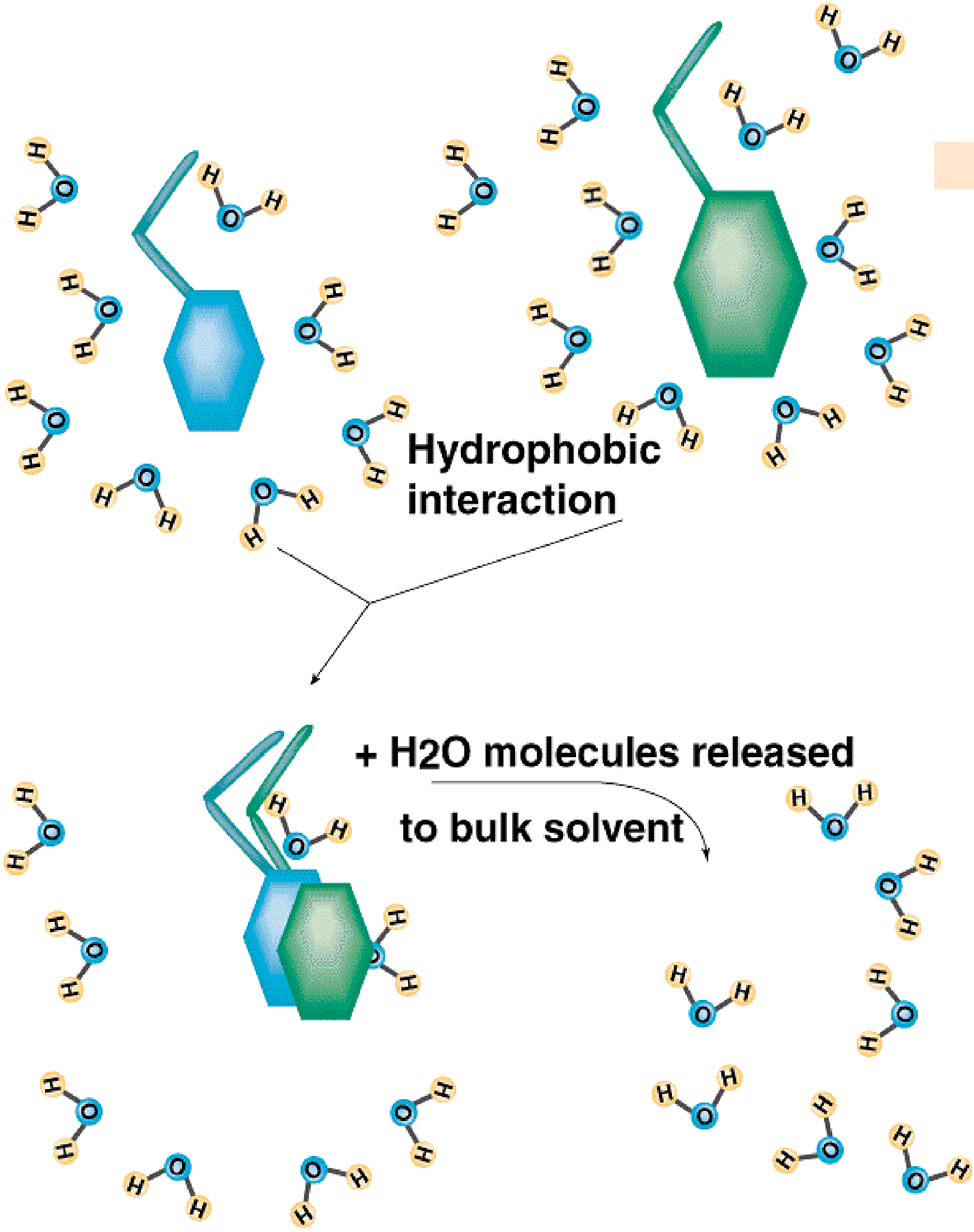
Figure 3.48. Some common hydrogen bonds found in proteins.

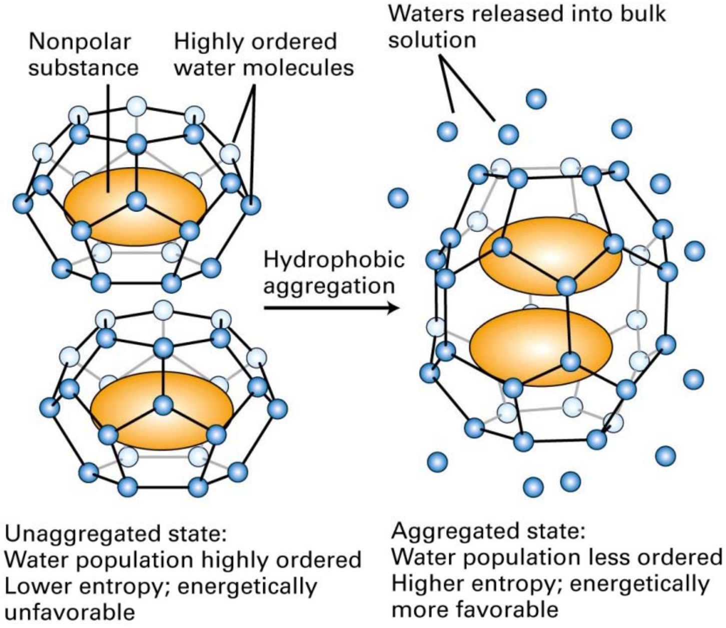
Devlin, *Biochemistry* 6th *ed* (2006)

Ionic Interaction



Hydrophobic Interactions





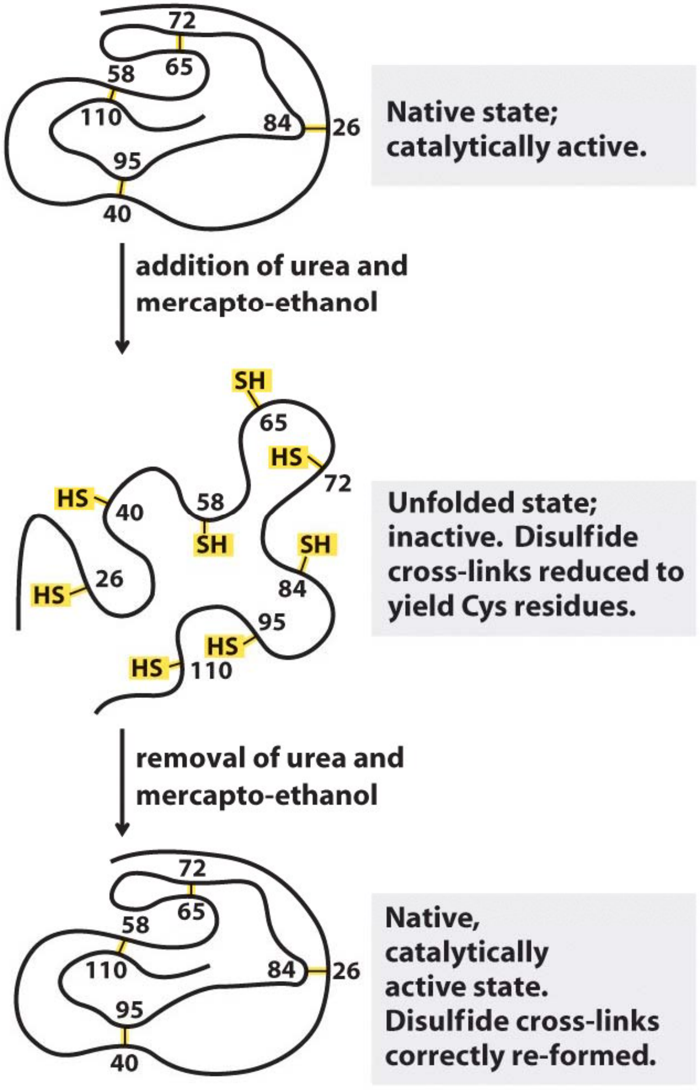
- Increases entropy of H_2O

Devlin, *Biochemistry 6th ed* (2006)

Lodish, *Molecular Cell Biology* 6th Ed.

Protein Denaturation

- Anfisen
  - Denatured ribonuclease using urea and β-ME
  - Removed denaturating agents – protein refolded into its native conformation
- Evidence that 3° structure of a globular protein is determined by its aa sequence


Native state; catalytically active.

addition of urea and mercapto-ethanol

Unfolded state; inactive. Disulfide cross-links reduced to yield Cys residues.

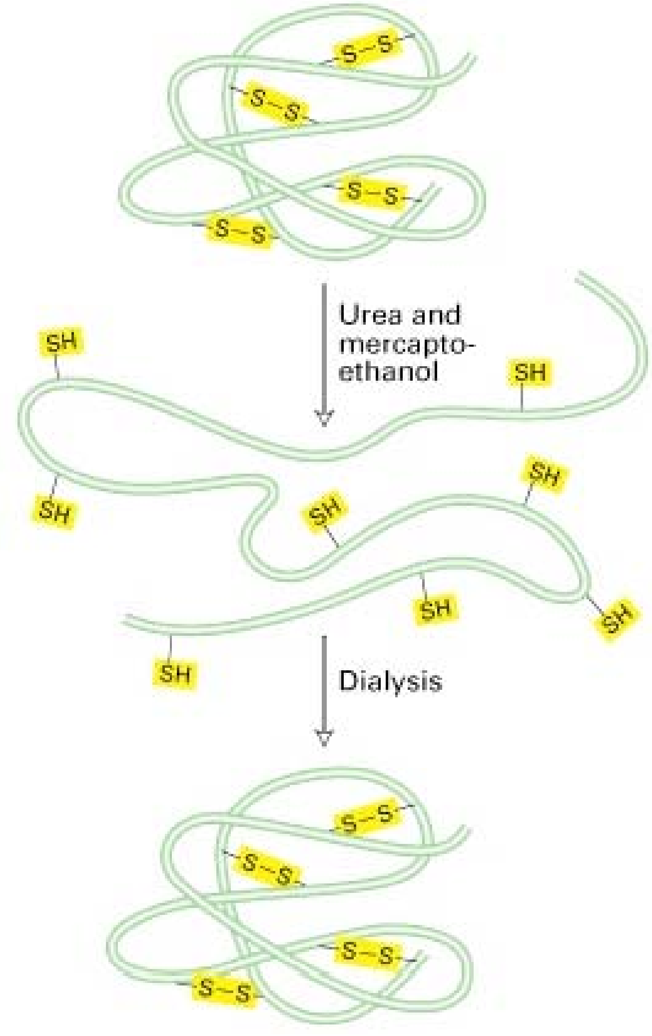
removal of urea and mercapto-ethanol

Native, catalytically active state. Disulfide cross-links correctly re-formed.

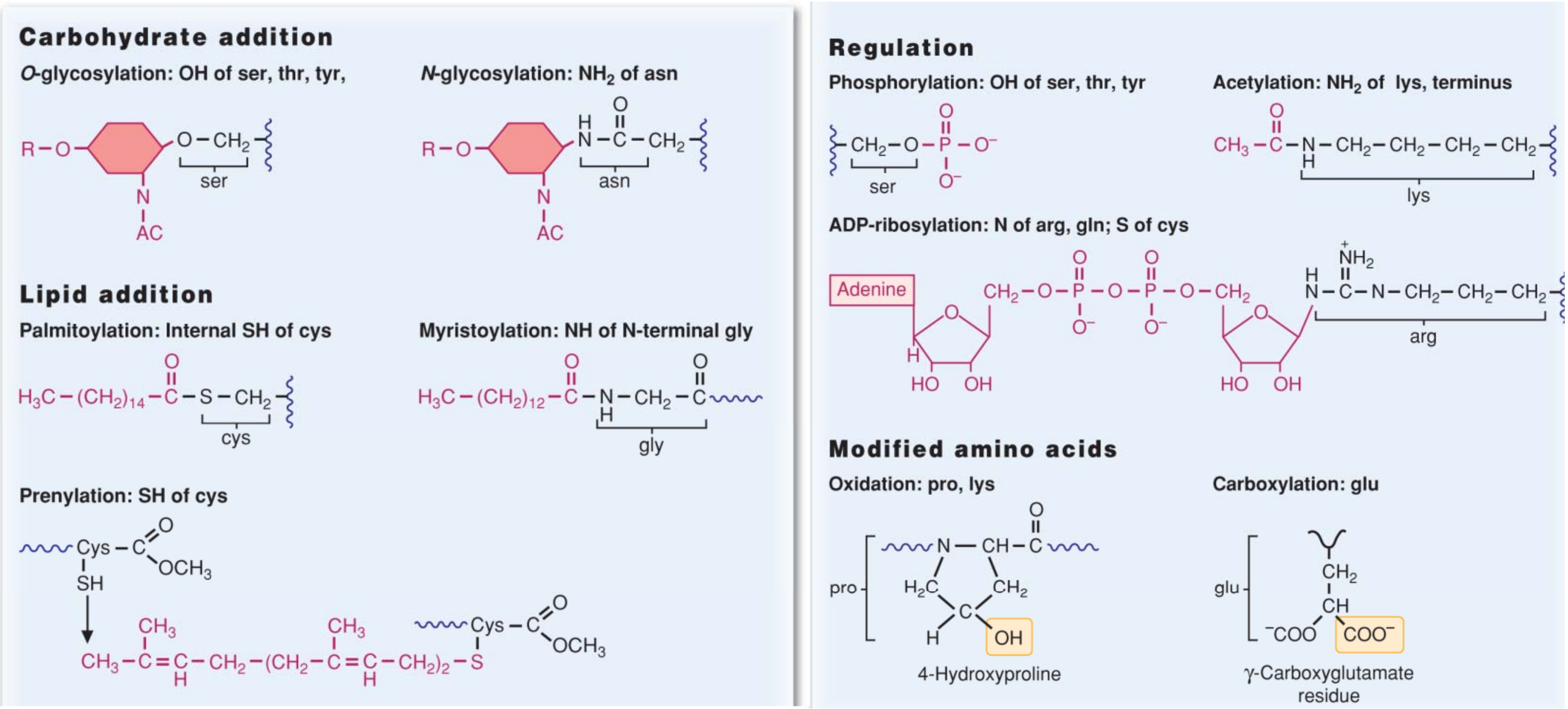
Protein Denaturation

- Temperature (heat)
  - ✓ Affects weak interaction - hydrophobic
- pH (Acids and Bases)
  - ✓ Alters net charge, cause electrostatic repulsion and disruption of H bonds
- Organic Solvents
- Detergents
- Urea
- Heavy metal ions
  - ✓ Lead
  - ✓ Mercury

} Disrupt hydrophobic interactions

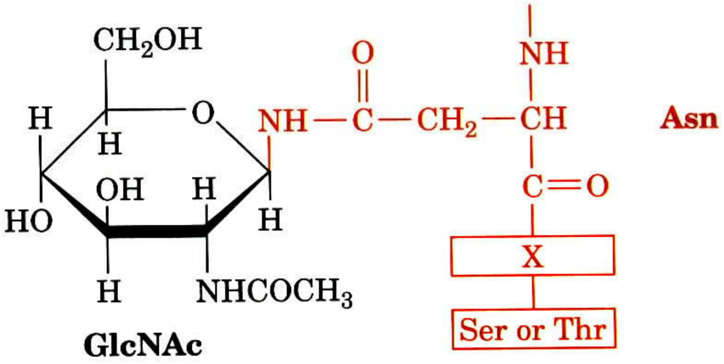


Post-translational modifications

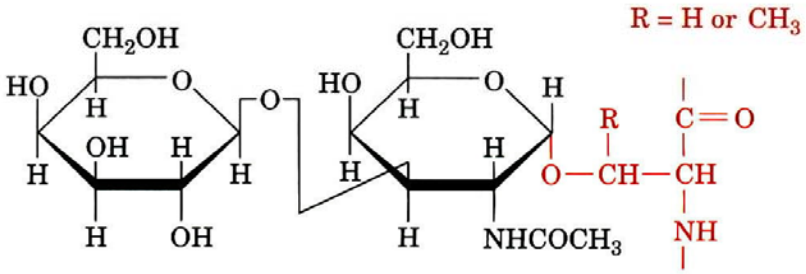


Post-translational modifications

- Glycosylation
  - ✓ Protection against proteases



N-linked

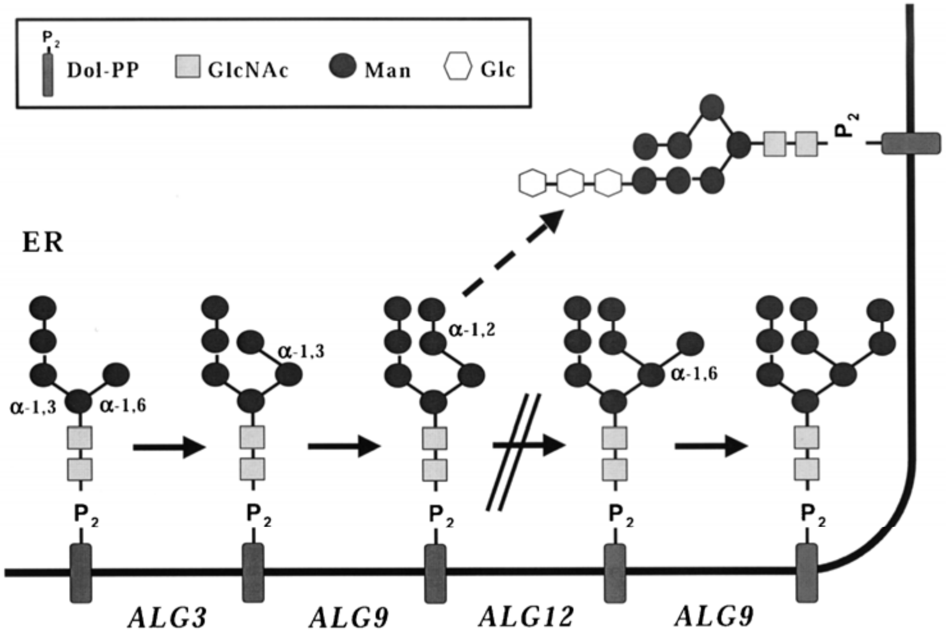


O-linked

Voet, Voet & Pratt, *Fundamentals of Biochemistry* 2nd ed (2006)

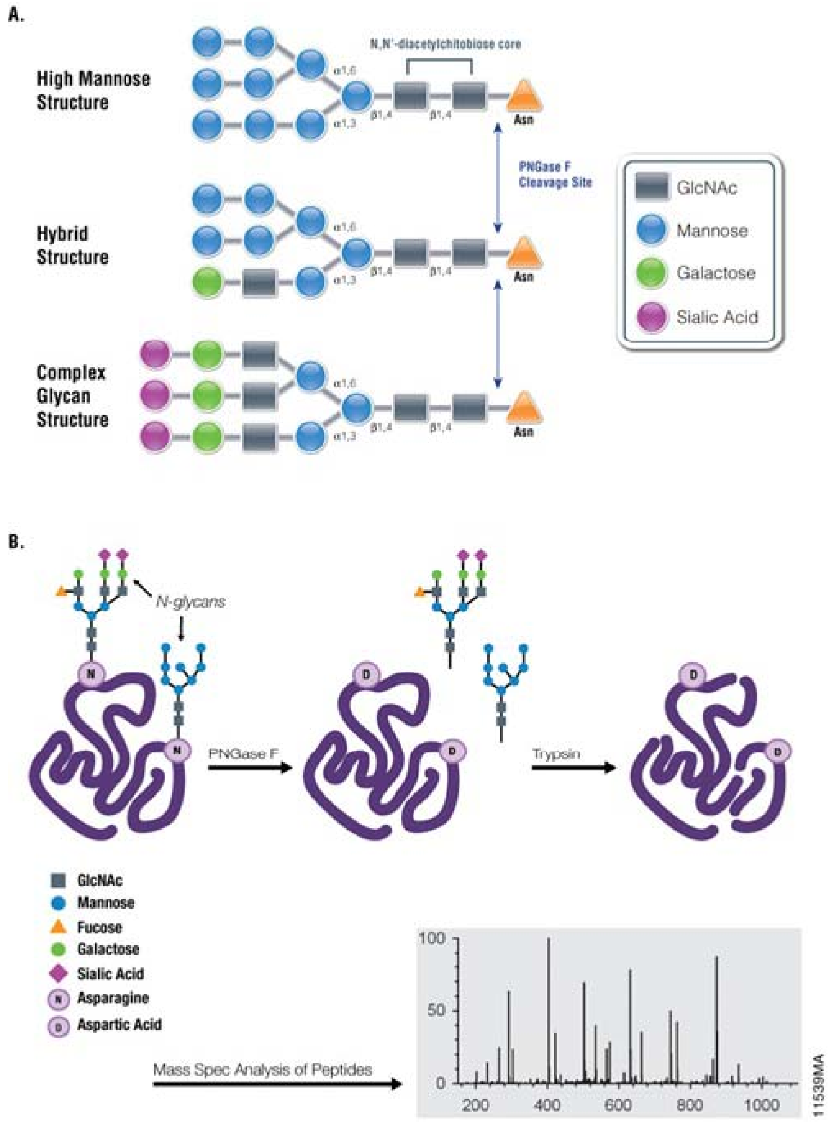
Post-translational modifications

- Glycosilation
  - ✓ Protection against proteases
  - ✓ For the detection of unfolded proteins in cells
  - ✓ Changes in glycosylation patterns in malignant cells
  - ✓ Increased branching pattern associated with malignancy

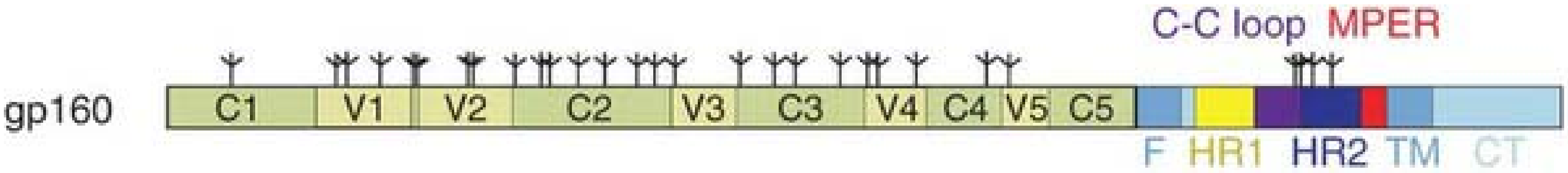


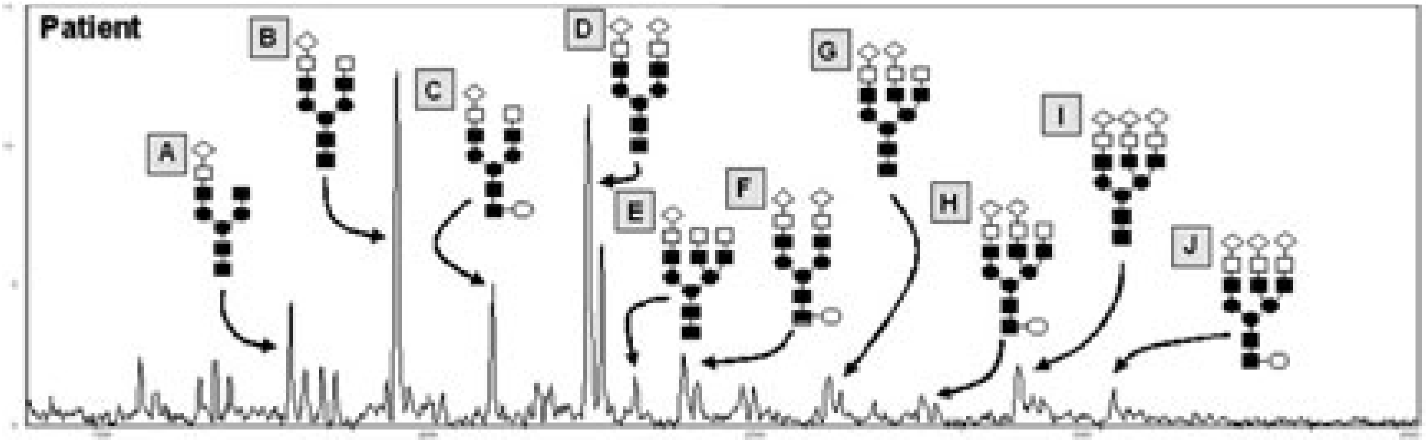
Voet, Voet & Pratt, *Fundamentals of Biochemistry* 2nd ed (2006)

Glycosylation of the HIV Env Protein



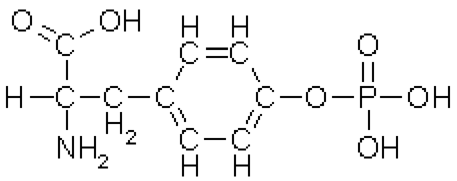
Glycosylation of the HIV Env Protein





Post-translational modifications

- Phosphorylation - regulation
  - ✓ Serine, threonine, tyrosine



Conjugated Proteins

- Apoprotein + prosthetic group = Prot conjugated

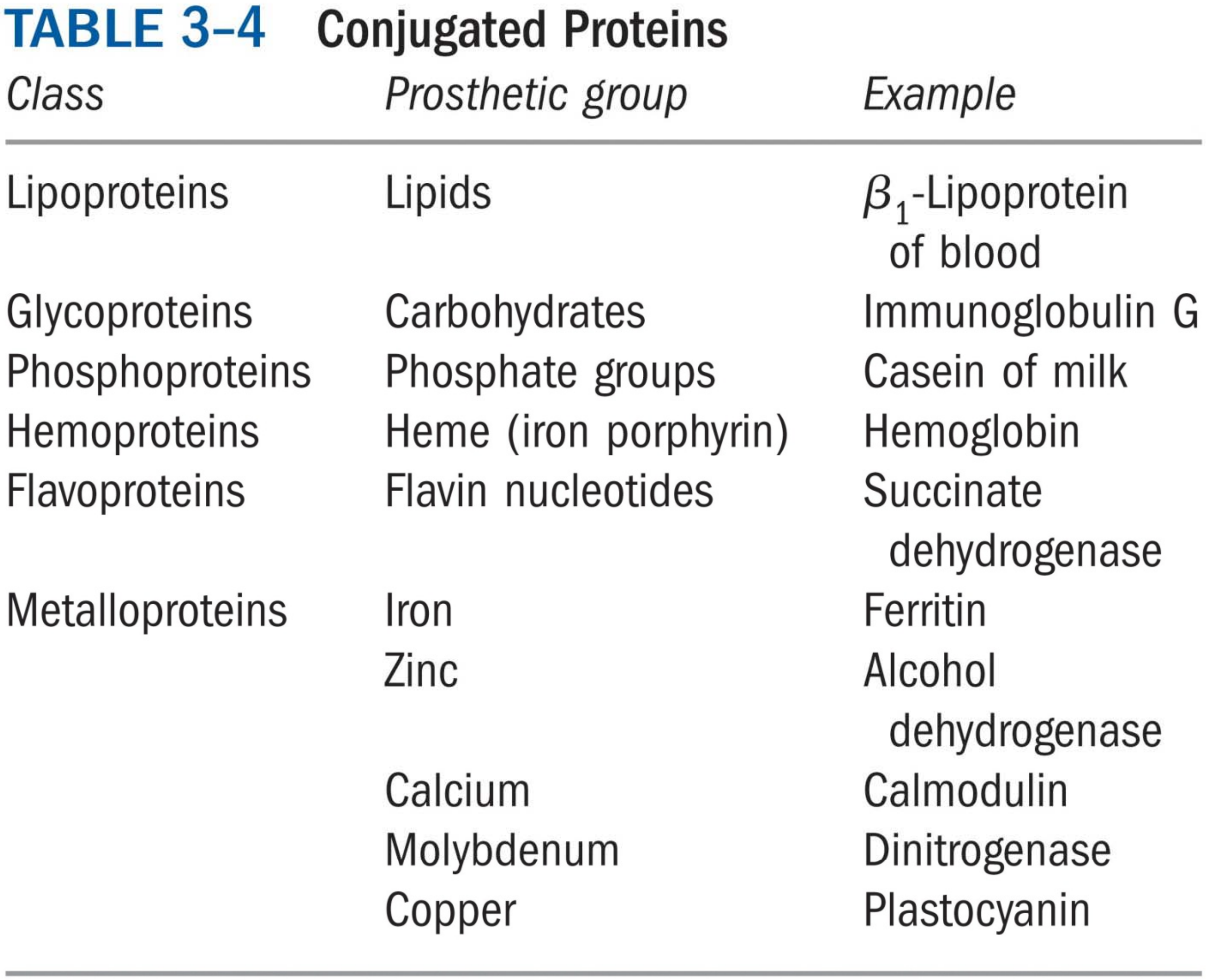


TABLE 3–4 Conjugated Proteins

Class	*Prosthetic group*	*Example*
Lipoproteins	Lipids	β_1-Lipoprotein of blood
Glycoproteins	Carbohydrates	Immunoglobulin G
Phosphoproteins	Phosphate groups	Casein of milk
Hemoproteins	Heme (iron porphyrin)	Hemoglobin
Flavoproteins	Flavin nucleotides	Succinate dehydrogenase
Metalloproteins	Iron	Ferritin
	Zinc	Alcohol dehydrogenase
	Calcium	Calmodulin
	Molybdenum	Dinitrogenase
	Copper	Plastocyanin

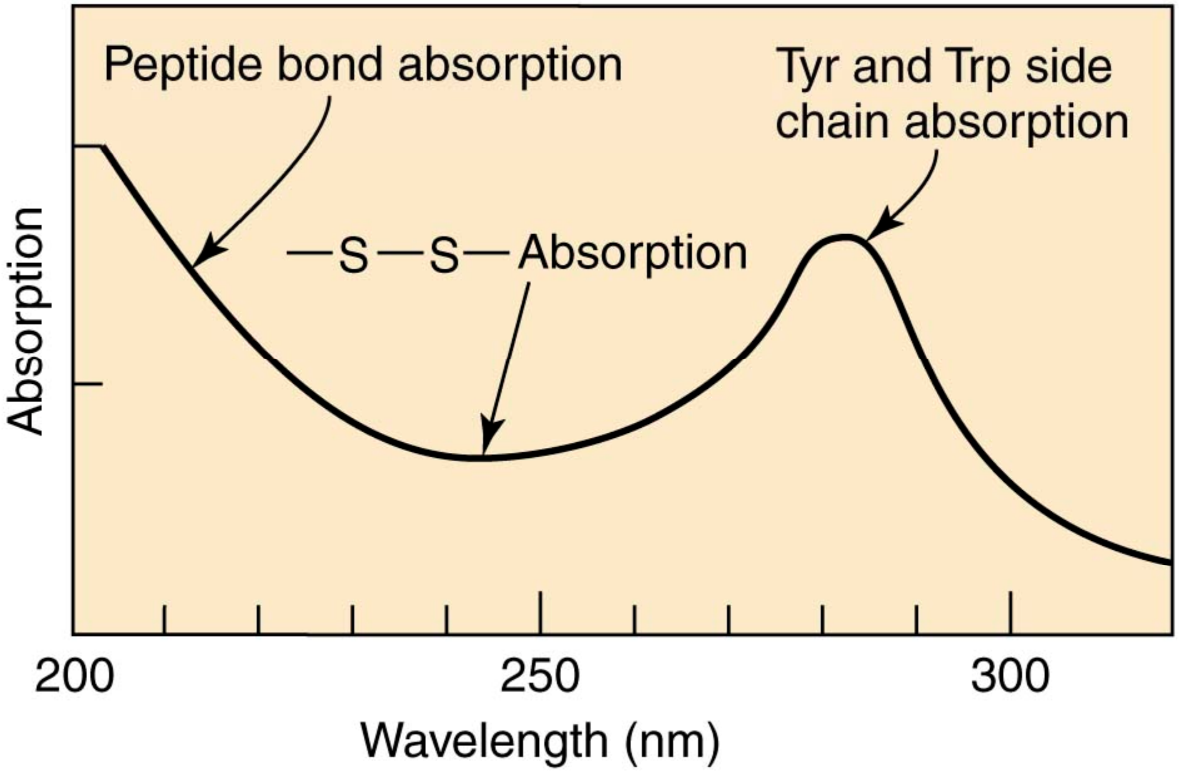
Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

Chemical Properties of Proteins

- UV Light
- Fluorescence
- Optical Rotatory Dispersion & Circular Dichroism

Spectroscopic Properties

- UV



Devlin, *Biochemistry 6th ed* (2006)

Spectroscopic Properties

- **Fluorescence**
  - ✓ Most relevant = Trp
  - ✓ Used to study structural and conformational changes

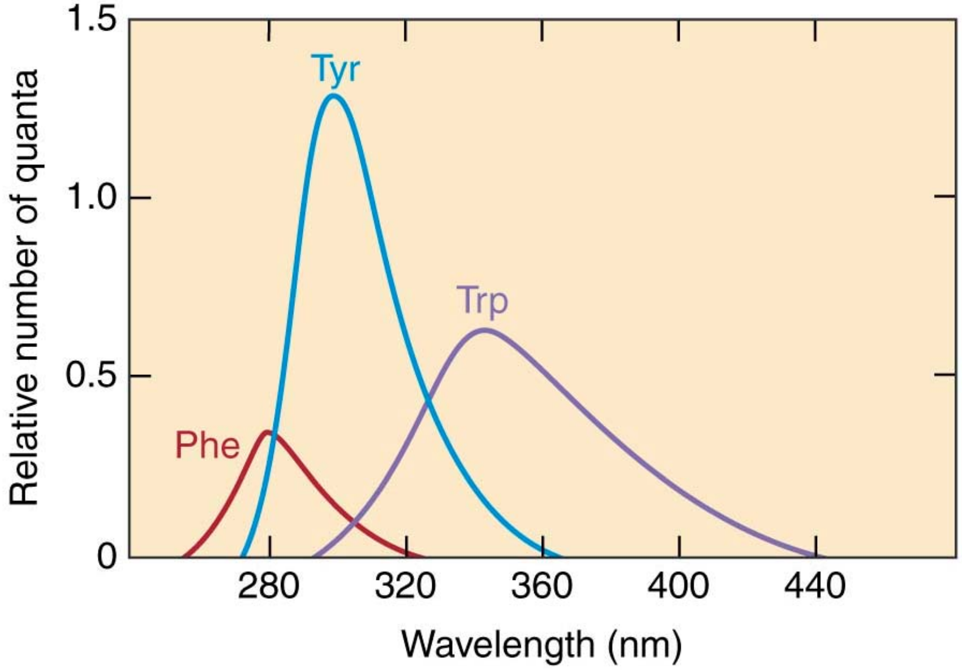
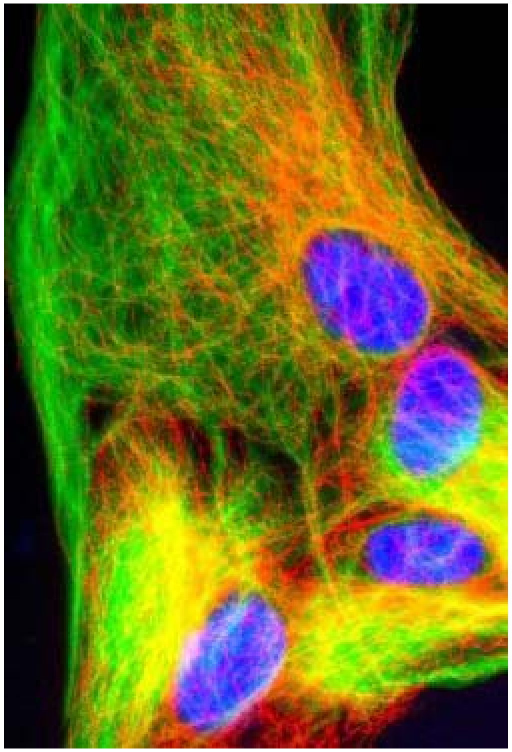


Figure 3.74. Characteristic fluorescence of aromatic groups in proteins. Redrawn from d'

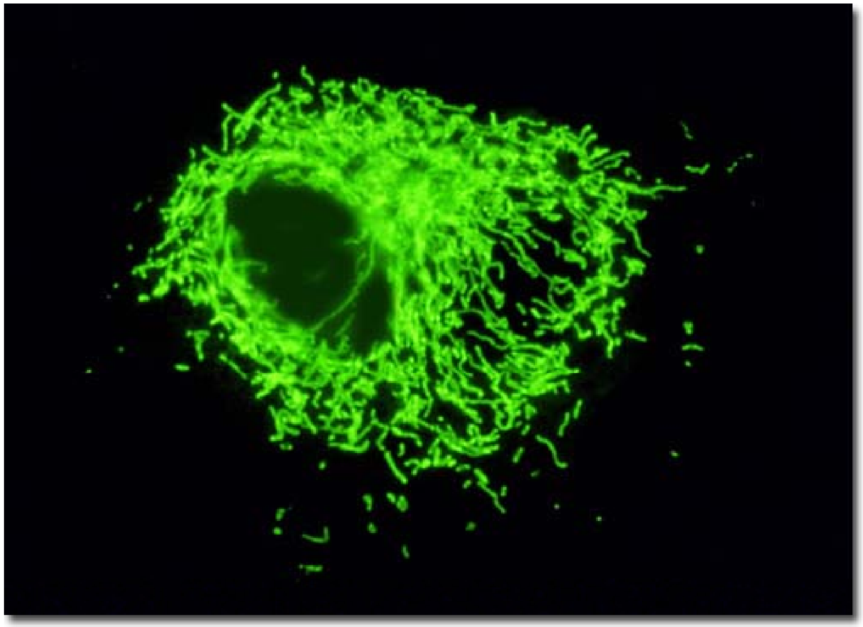
Devlin, *Biochemistry 6th ed* (2006)

Spectroscopic Properties

Fluorescence is used to visualize proteins inside the cell and determine their localization



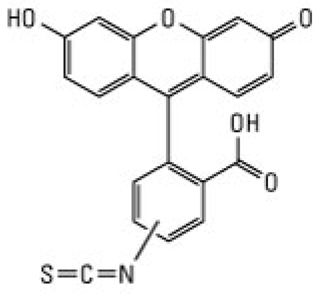
Nucleus



Mitochondria

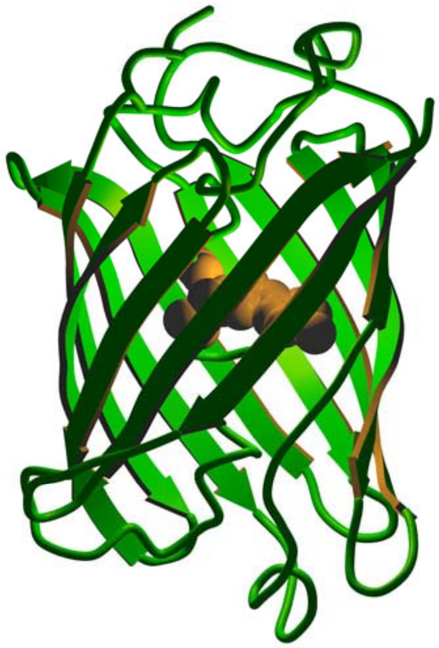
Spectroscopic Properties

Fluorescence is used to visualize proteins inside the cell and determine their localization



FITC
5/6-Fluorescein-isothiocyanate
MW 389.38

Protein Labeling



Green Fluorescent Protein
from jellyfish *Aequorea victoria*

Protein Analysis

- Separation of Proteins
- General Approach to Protein Purification
- Determination of Amino Acid Composition
- Spectroscopic Methods
- Determination of 3D Structure

Protein Analysis

(by Techniques)

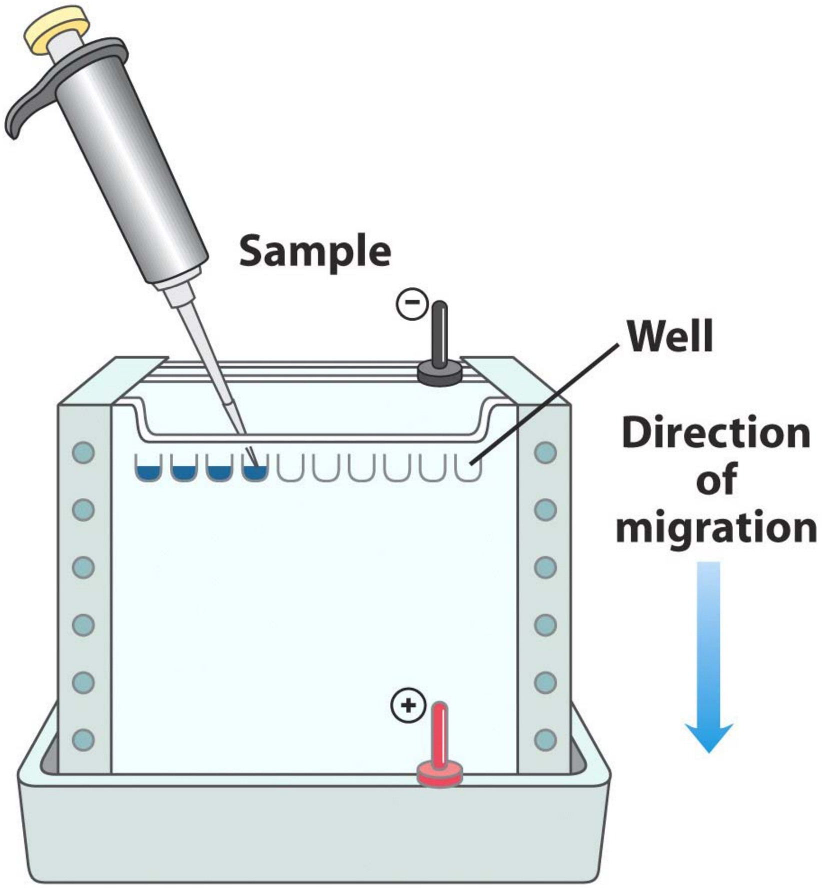
- Precipitation
- Ultracentrifugation
- Electrophoresis
  - ✓ Isoelectric focusing
  - ✓ SDS-PAGE Gel Electrophoresis
  - ✓ 2D gel electrophoresis
- Chromatography
  - ✓ Ion Exchange
  - ✓ Gel Filtration
  - ✓ Affinity
- Others

Electrophoresis

- Based on the movement of charged molecules when an electric field is applied
- Protein migrates according to its mass/charge ratio
- Use inert matrix:
  - ✓ Acrylamide
    - Small pore gels
    - used to separate most proteins
    - 5,000 to 200,000 Da
  - ✓ Agarose

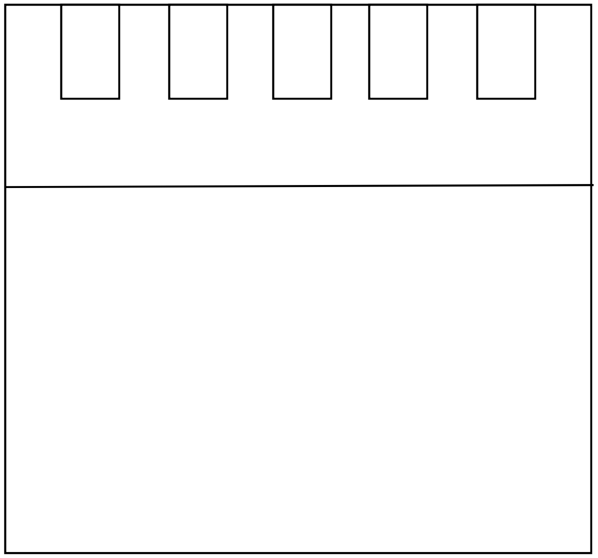
Electrophoresis

- Molecules negatively charged will move towards the anode

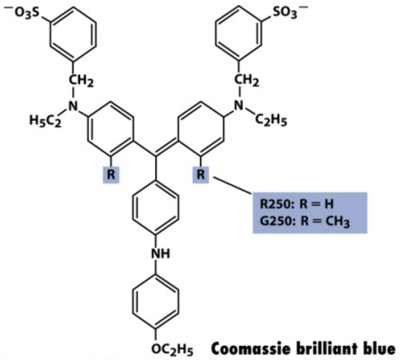


Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

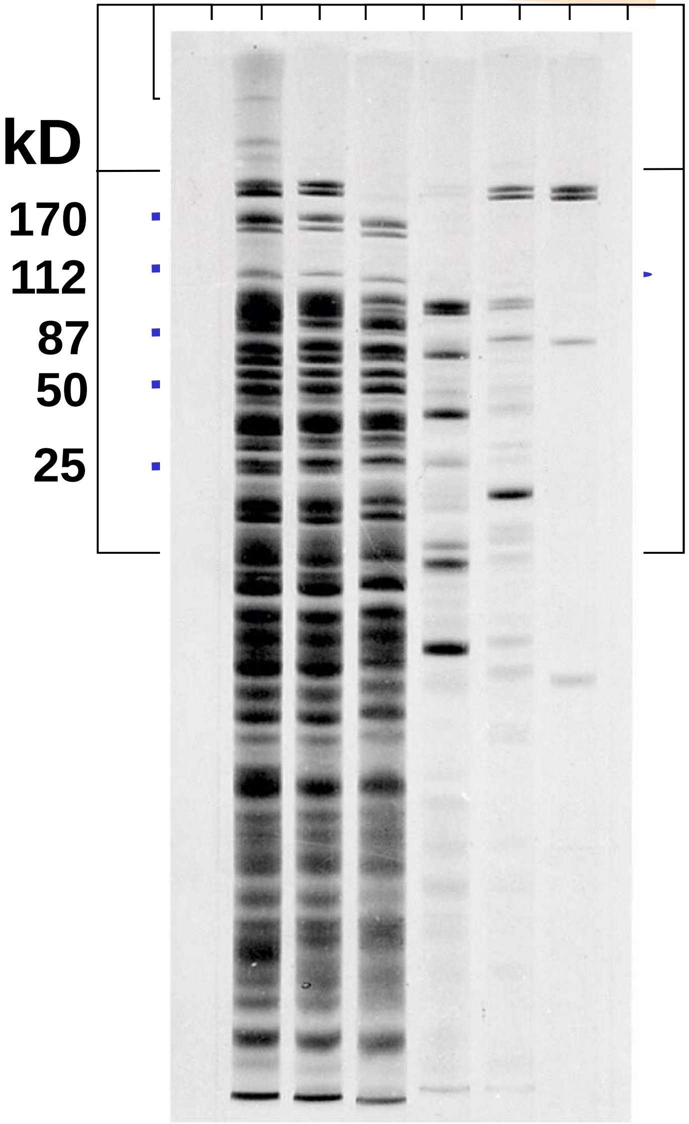
Electrophoresis



Coomassie Blue Staining

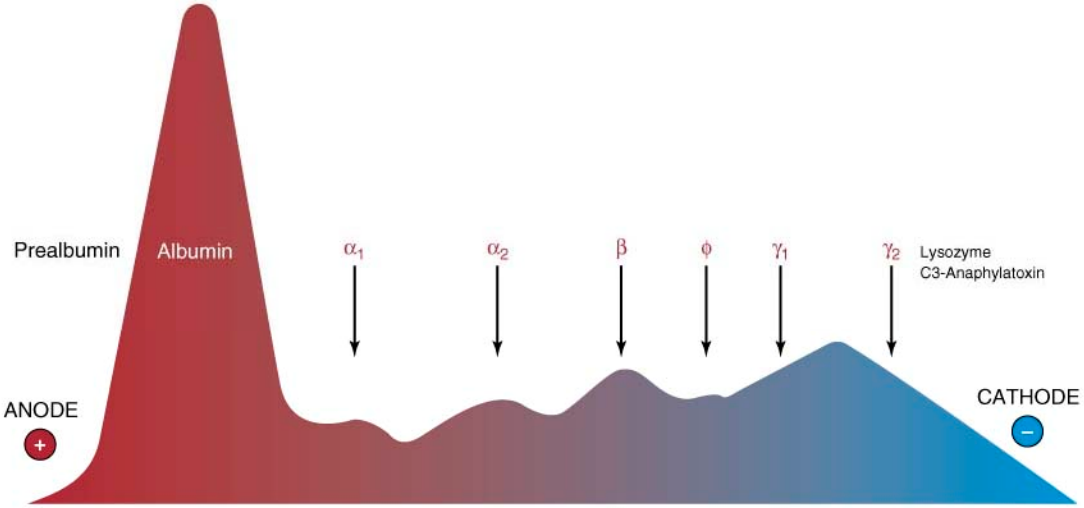


Unnumbered figure pg 100 Fundamentals of Biochemistry, 2/e



- Other Stains:
  - ✓ Silver Stain

Electrophoresis Pattern for Serum Proteins



α_1

- α_1-Acid glycoprotein
- α_1-T Glycoprotein
- α_1-Antitrypsin
- Transcortin
- α_1-Antichymotrypsin
- α_1-B Glycoprotein
- α_1-glycoprotein
- Vitamin D-binding protein
- α_1-Lipoproteins

α_1/α_2

- Thyroxine-binding globulin
- Zn-α_2-glycoprotein
- Gc globulin
- Ceruloplasmin
- Inter-α-trypsin inhibitor
- Antithrombin III
- Factor X (Stuart-Prower)
- Transcobalamin I
- C9

α_2

- Retinol-binding protein
- α_2 HS glycoprotein
- Histidine-rich 3,8 S-α_2-glycoprotein
- Haptoglobin
- Pregnancy zone protein
- α_2-Macroglobulin
- Prothrombin
- Antihemophilic factor
- C1 inactivator
- C1s

α_2/β_1

- Serum cholinesterase
- α_3-glycoprotein
- α_2, β_1-glycoprotein
- Transcobalamin III

β_1

- Hemopexin
- Steroid-binding β-globulin
- Transferrin
- Pregnancy-specific β_1-glycoprotein
- Cold insoluble globulin
- Factor V (Accelerin)
- Factor VII (Proconvertin)
- Factor IX (Christmas)
- Plasminogen
- C3 Proactivator
- Transcobalamin II
- β-Lipoproteins
- C1r, C2
- C4, C5

β_2

- β_2-Microglobulin
- β_2-Glycoprotein III
- β_2-Glycoprotein I
- Fibrinogen
- Factor XI (PTA)
- Factor XII (Hageman)
- Factor XIII (FSF)
- C3, C6, C7
- C3 Activator (β_2II)

γ_1

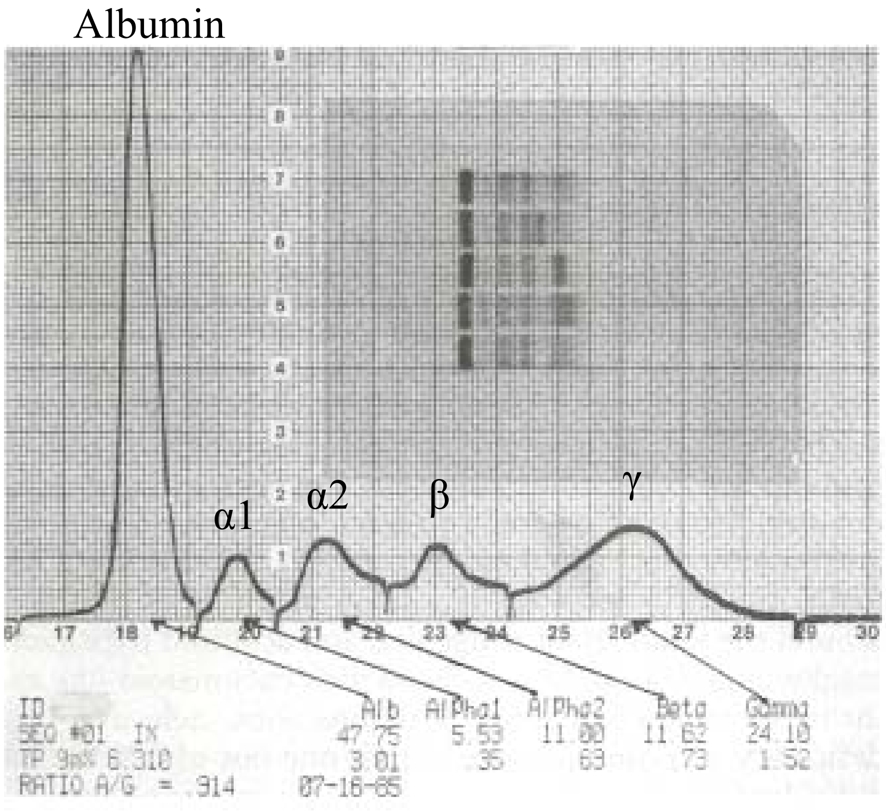
- IgG
- IgA
- IgD
- IgE
- IgM
- Amylase

γ_2

- IgG
- Clq
- Properdin

Devlin, *Biochemistry 6th ed* (2006)

Electrophoresis Pattern for Serum Proteins



```
ID [redacted]                          Alb  AlPha1  AlPha2   Beta   Gamma
SEQ #01  IX                          47.75    5.53   11.00  11.62   24.10
TP 9m% 6.310                          3.01     .35     .69    .73    1.52
RATIO A/G  = .914     07-16-85
```

In response to infections there is an increase in the intensity of the α2 and γ fractions

Hypo-gamma-globulinemia is a virtual lack of gamma-globulins and is symptomatic of immunosuppression.

Hyper-gamma-globulinemia used to diagnose liver disease and lupus erythematosus

In hepatic cirrhosis there is a decrease in total albumin together with an increase in the gamma fractions.

α1 band indicative of antitrypsin deficiency

Devlin, *Biochemistry 6th ed* (2006)

Electrophoresis Pattern for Serum Proteins

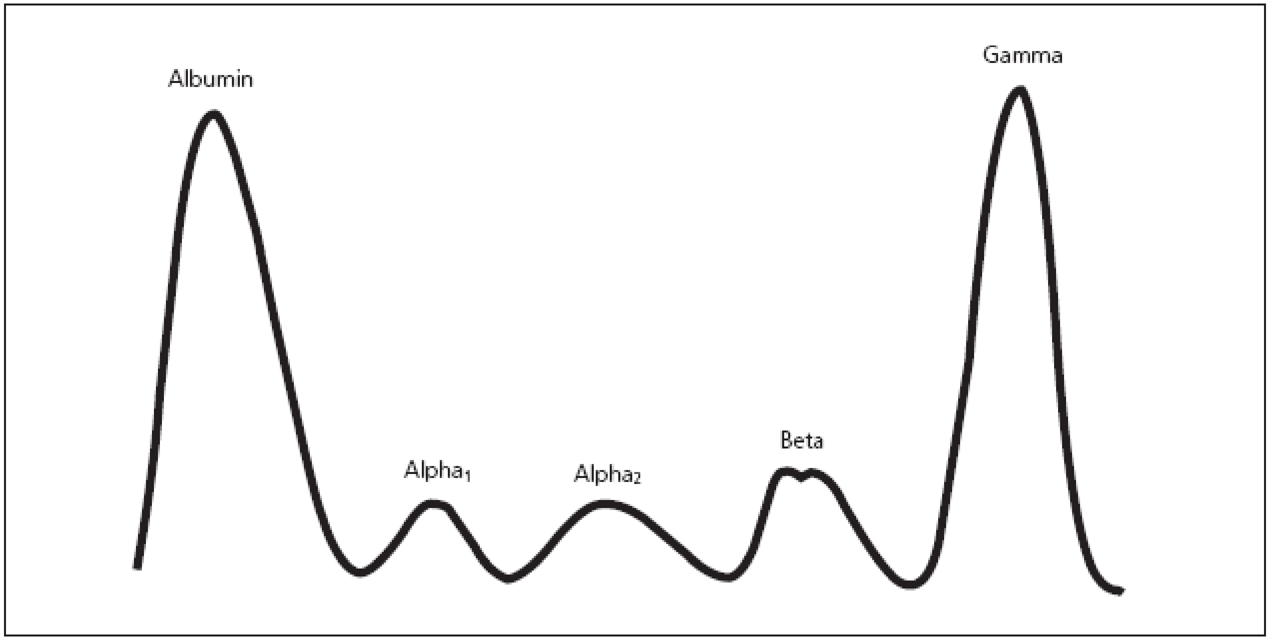


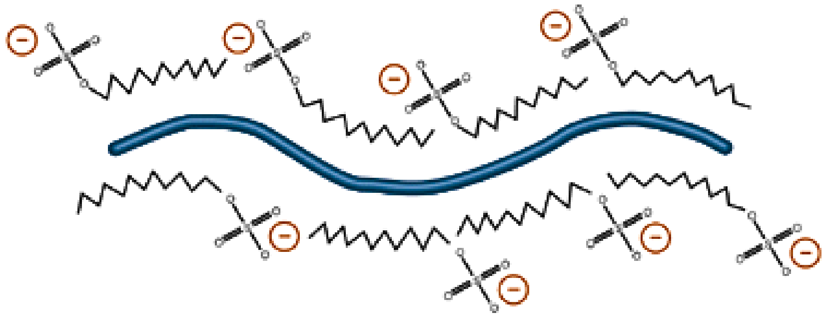
Figure 2. Abnormal serum protein electrophoresis pattern in a patient with multiple myeloma. Note the large spike in the gamma region.

Gamma fraction spikes correlate with multiple myeloma and Hodgkin’s disease

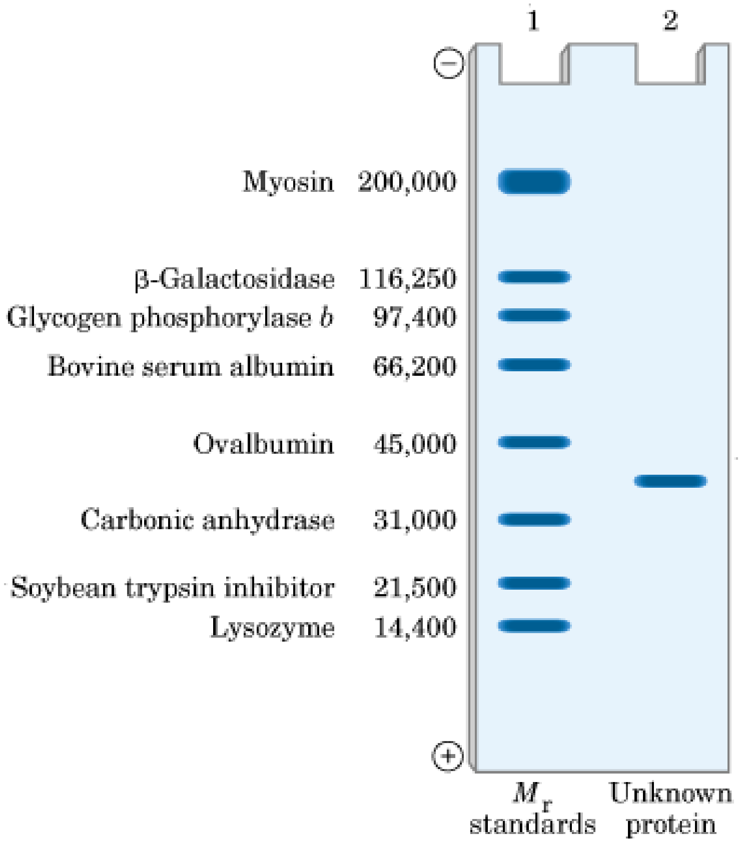
O’Connell *et al.* (2005) American Family Phycician

SDS-PAGE

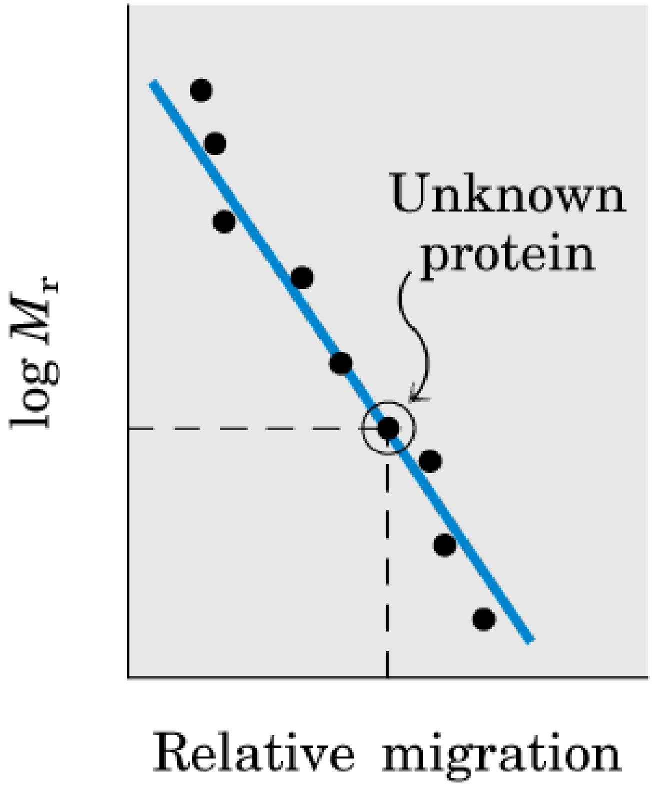
- PAGE = PolyAcrilamide Gel Electrophoresis
- SDS = sodium dodecyl sulfate
- Separate proteins based on **size**
- All proteins will be negatively charged and will migrate based on size (not charge/size)



SDS - PAGE



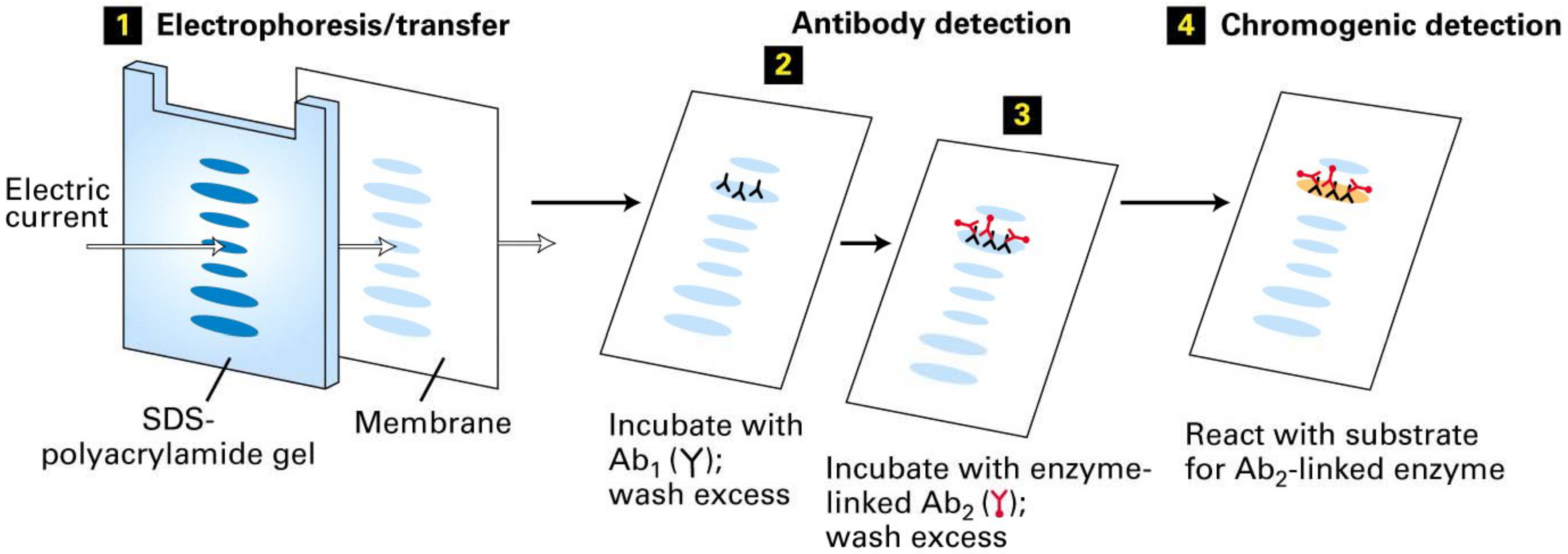
(a)



(b)

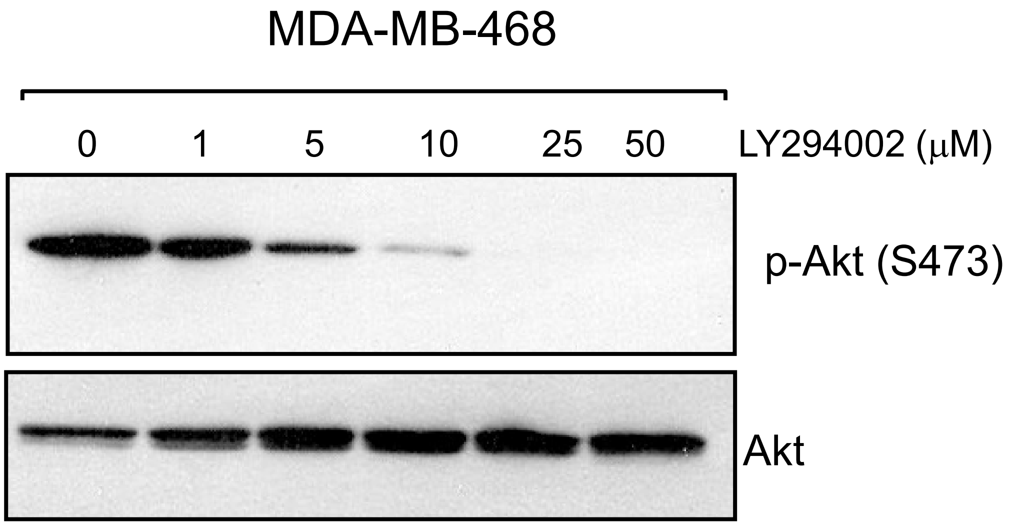
Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

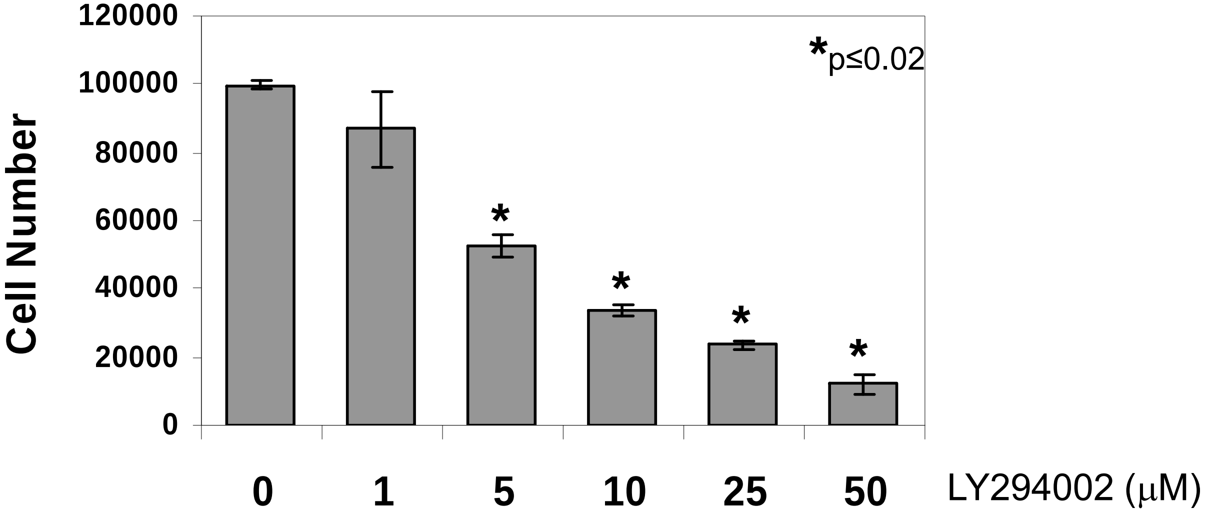
Western Blot



Lodish, *Molecular Cell Biology* 6th Ed.

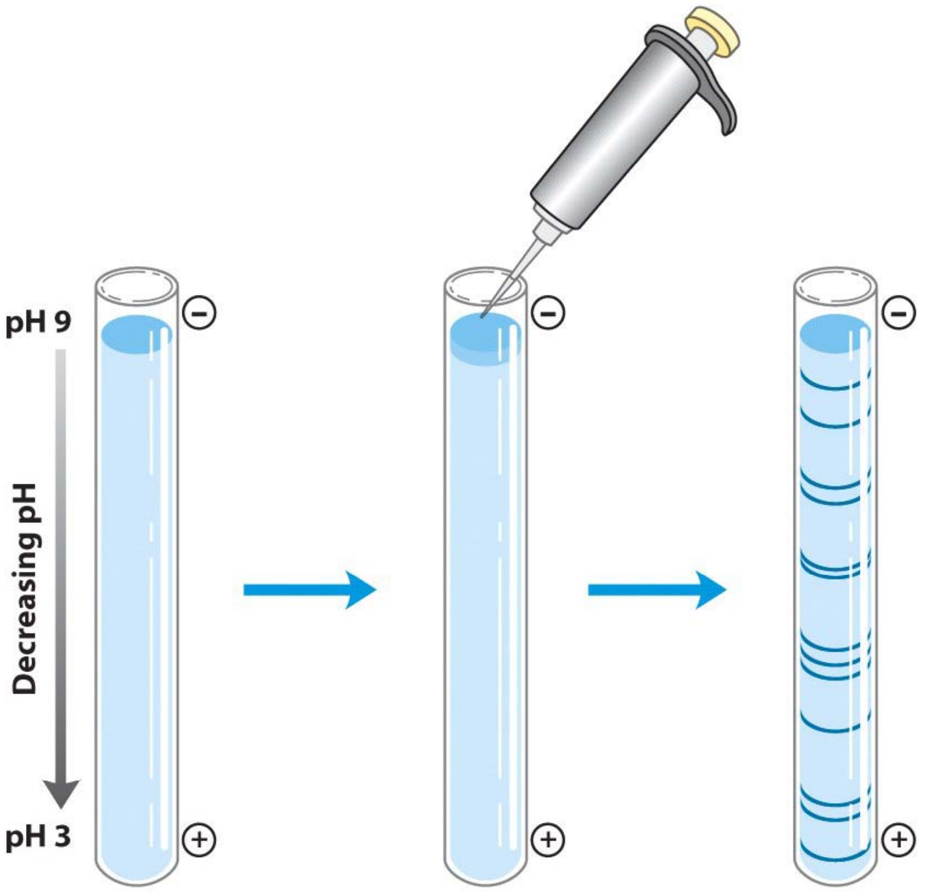
Western Blot - Example





Unpublished data MPT & CLA

Isoelectric focusing



An ampholyte solution is incorporated into a gel.

A stable pH gradient is established in the gel after application of an electric field.

Protein solution is added and electric field is reapplied.

After staining, proteins are shown to be distributed along pH gradient according to their pI values.

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

pI

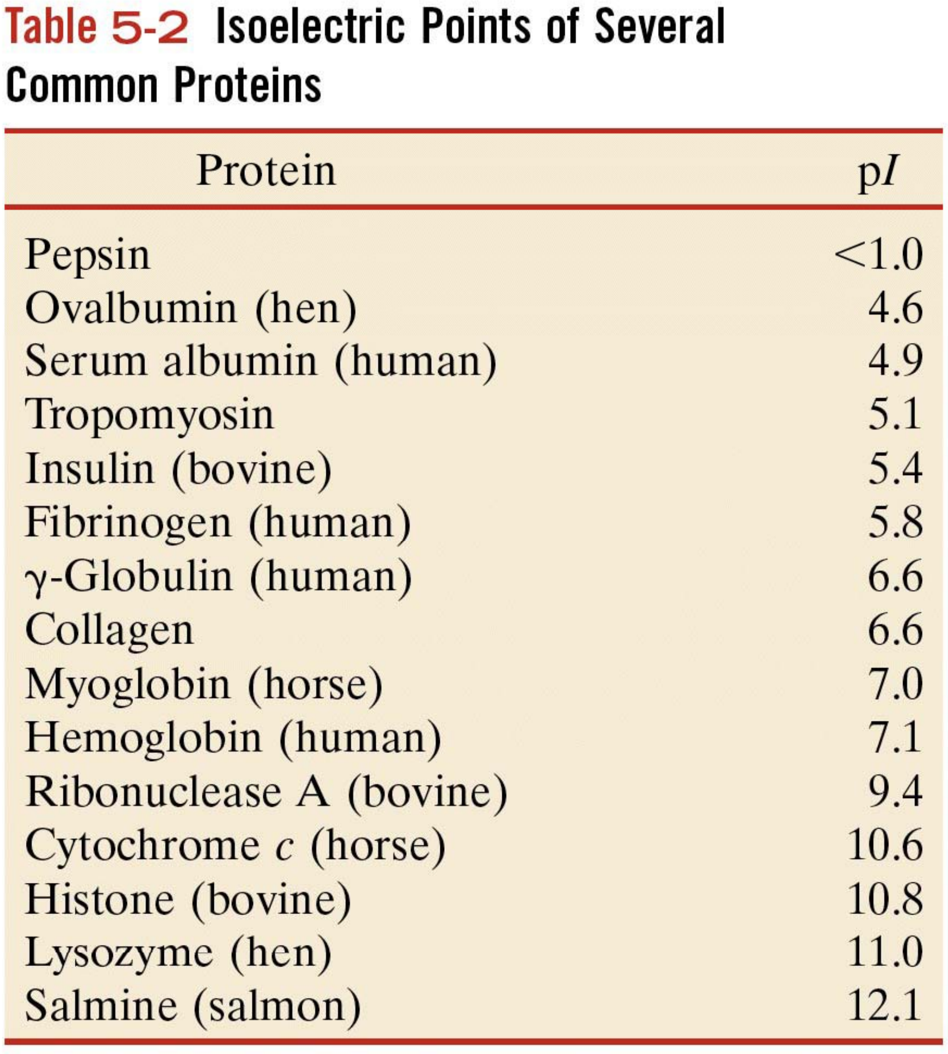


Table 5-2 Isoelectric Points of Several Common Proteins

Protein	p*I*
Pepsin	<1.0
Ovalbumin (hen)	4.6
Serum albumin (human)	4.9
Tropomyosin	5.1
Insulin (bovine)	5.4
Fibrinogen (human)	5.8
γ-Globulin (human)	6.6
Collagen	6.6
Myoglobin (horse)	7.0
Hemoglobin (human)	7.1
Ribonuclease A (bovine)	9.4
Cytochrome *c* (horse)	10.6
Histone (bovine)	10.8
Lysozyme (hen)	11.0
Salmine (salmon)	12.1

pH > p*I*, then protein charge negative
pH < p*I*, then protein charge positive

Voet, Voet & Pratt, *Fundamentals of Biochemistry* 2nd ed (2006)

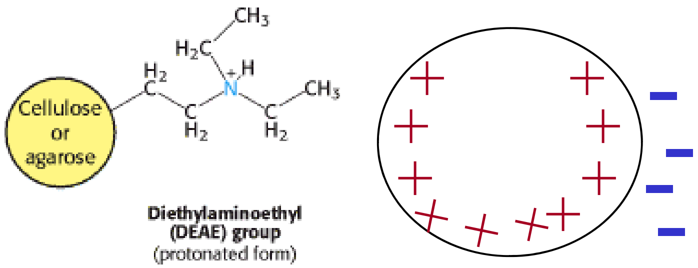
Devlin, *Biochemistry 6th ed* (2006)

Chromatography

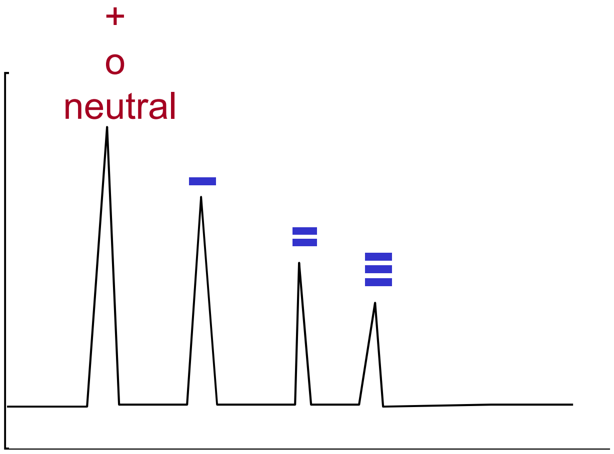
- Group of techniques based in the separation of the components of a mixture and its detection
- Chromatographic techniques are diverse, but all consist of a **mobile phase** (gas, liquid) that moves the sample along a **stationary phase**.
- Each component of the sample will interact in a different manner with the mobile and stationary phases.

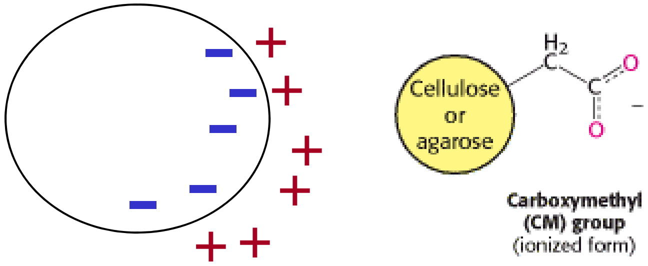
Ion Exchange Chromatography

- Anionic Exchange
  - ✓ resin - positive (+)
  - ✓ Attracts anions (-)
- Cationic Exchange
  - ✓ resin - negative (-)
  - ✓ Attracts cations (+)

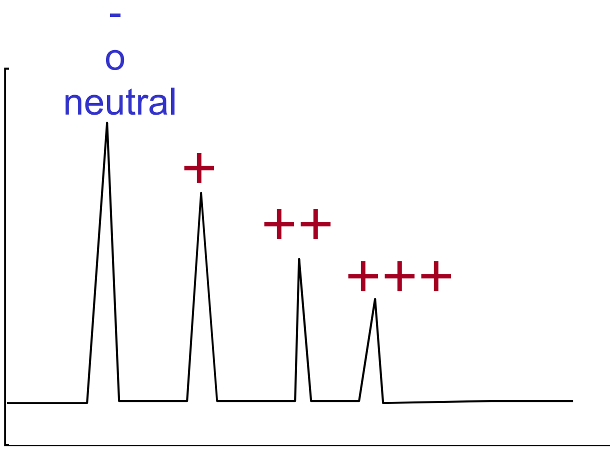


Diethylaminoethyl (DEAE) group
(protonated form)



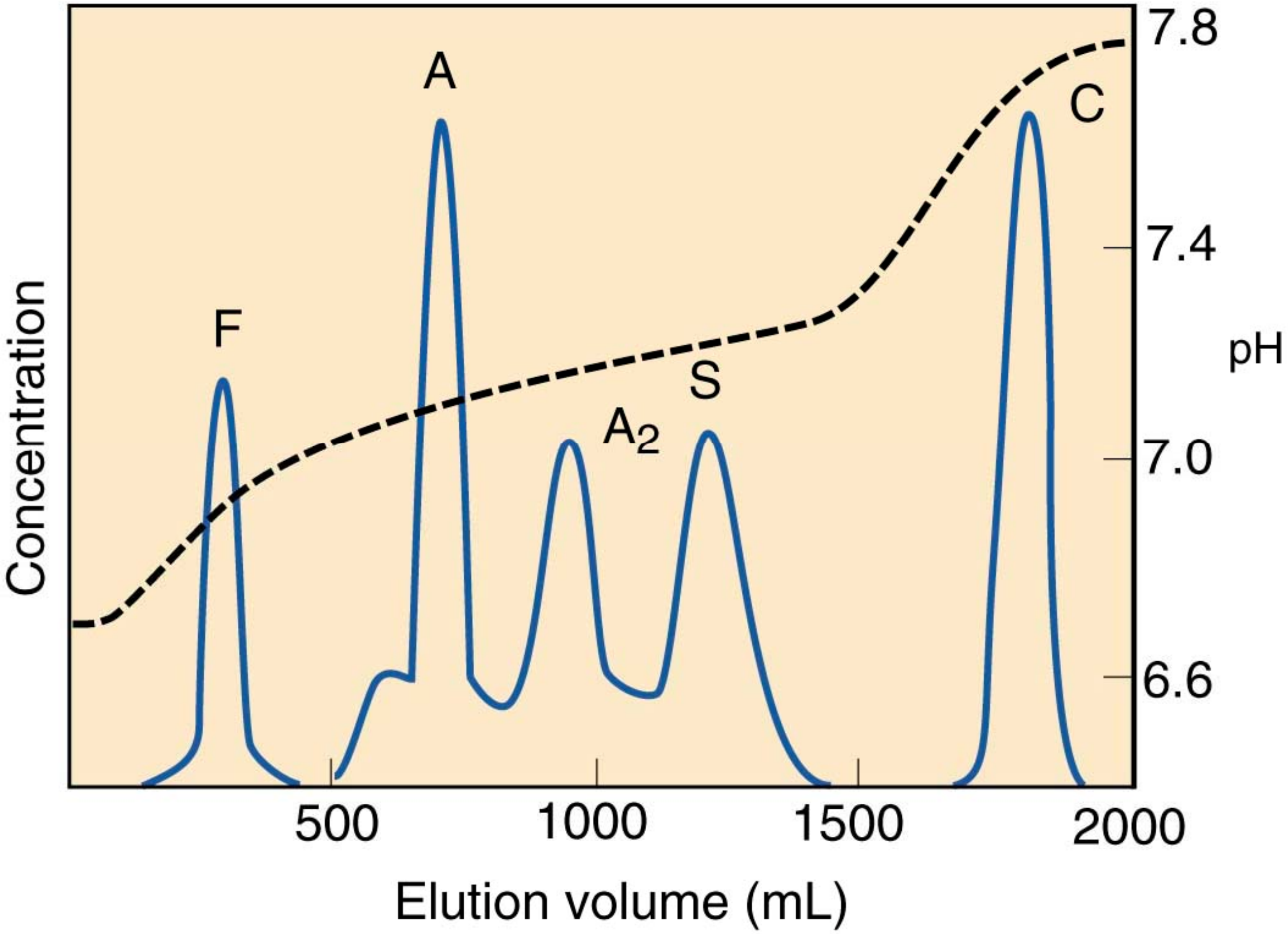


Carboxymethyl (CM) group
(ionized form)



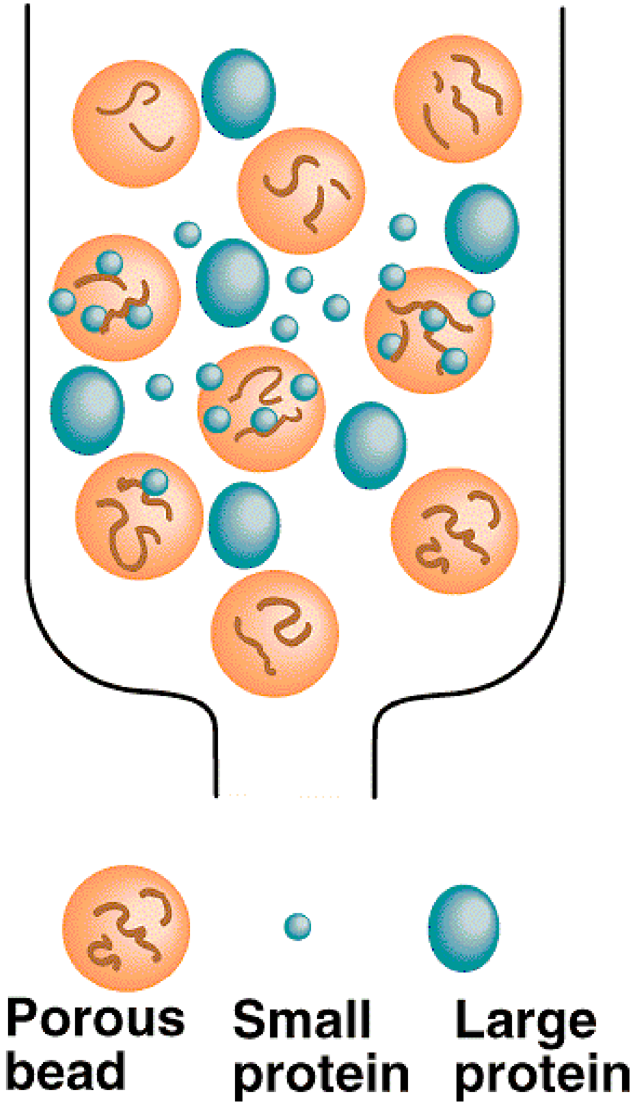
Stryer, ***Biochemistry*** 5th Ed. (2002)

Ion Exchange Chromatography



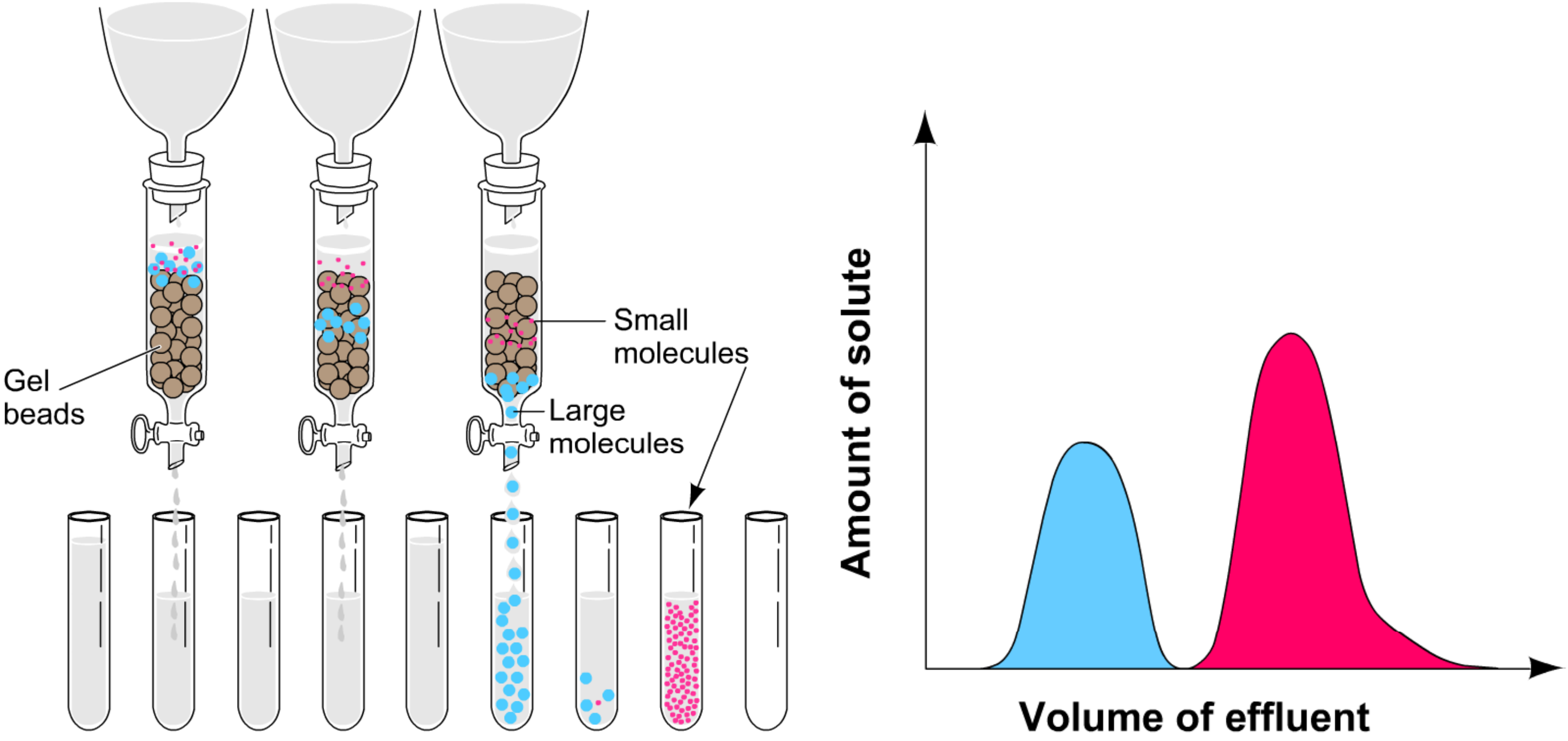
Size-Exclusion Chromatography

- aka: Gel Filtration Chromatography
- Separates by size
- Small proteins penetrate the pores of the gel and have a larger solvent volume through wich to travel in the column thatn larger proteins



Devlin, *Biochemistry 6th ed* (2006)

Size-Exclusion Chromatography



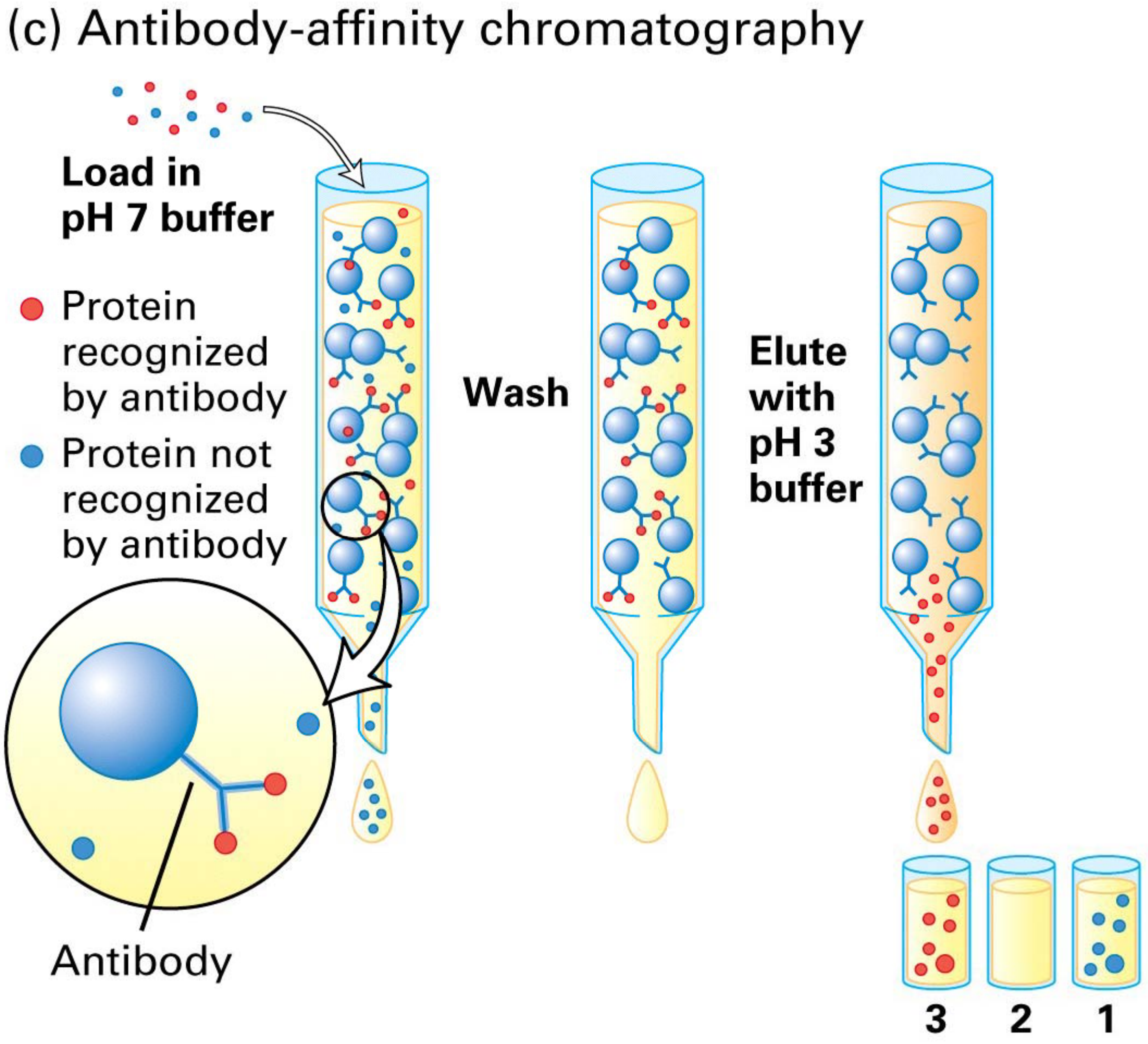
Voet, Voet & Pratt, *Fundamentals of Biochemistry* 2nd ed (2006)

Other Separation Techniques:

Affinity Chromatography

- Take advantage of specific interactions of proteins
- Examples:
  - ✓ enzyme - substrate
  - ✓ antigen - antibody
  - ✓ Any other biological pair

Affinity Chromatography



(c) Antibody-affinity chromatography

Lodish, *Molecular Cell Biology* 6th Ed.

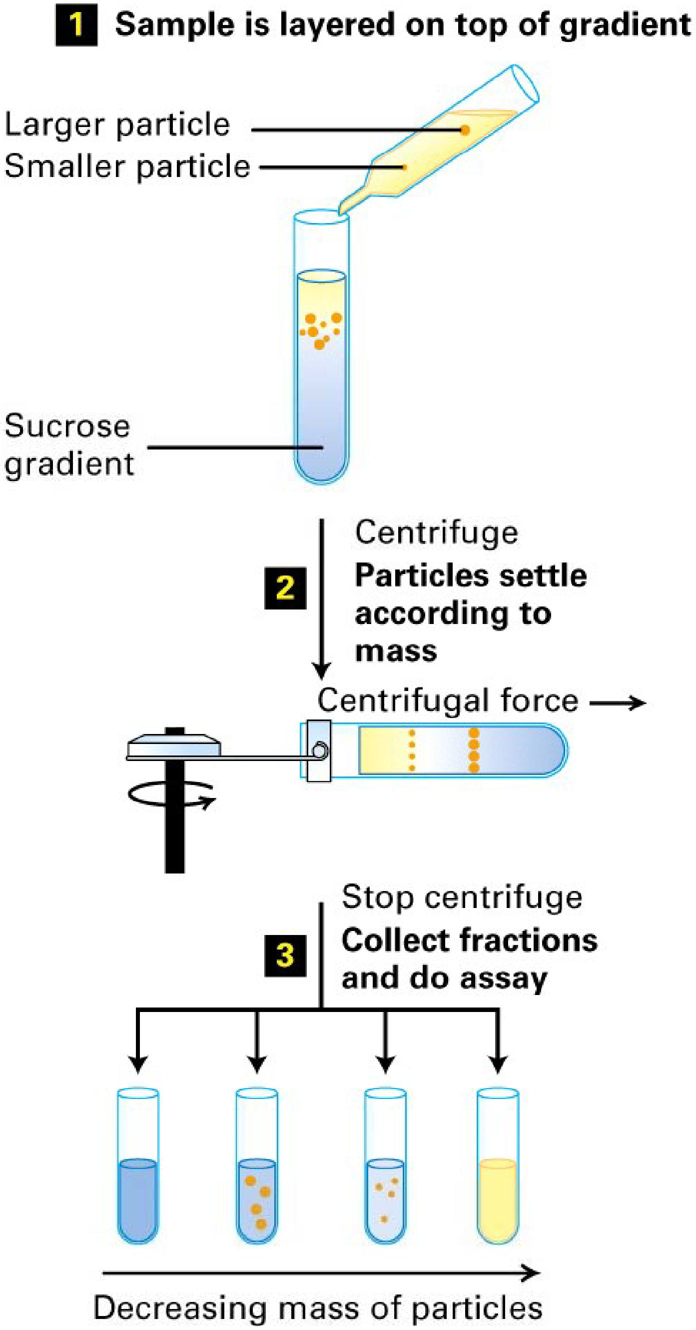
Ultracentrifugation

- Measures sedimentation coefficient
  - ✓ Svedver units (10^{-13} sec)
  - ✓ Method to determine MW
- Use a gradient
  - ✓ Sucrose
  - ✓ Cesium Chloride
- Protein will migrate at a rate controlled by factors that affect sedimentation constant

$$s = \frac{v}{\omega^2 r}$$

Equation for calculation of the Svedberg coefficient.

Devlin, *Biochemistry* 6th ed (2006)



Lodish, *Molecular Cell Biology* 6th Ed.

Ultracentrifugation

Table 3.17 Svedberg Coefficient for some Plasma Proteins

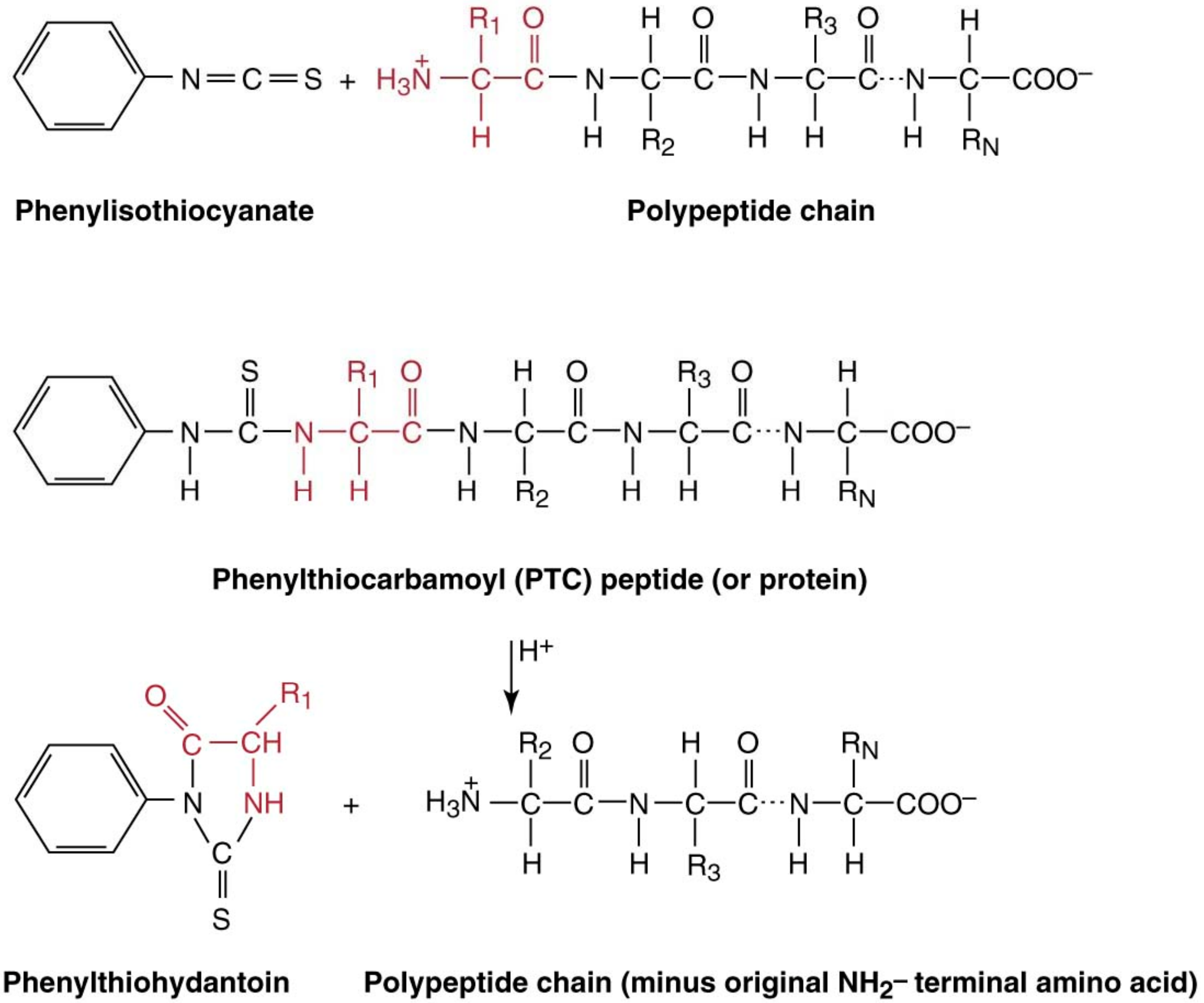
Protein	Svedberg Coefficient	Mol. Weight
Lysozyme	2.19	15,000 – 16,000
Albumin	4.6	69,000
Immunoglobulin G	6.6 – 7.2	153,000
Fibrinogen	7.63	341,000
C1q	11.1	410,000

Determination of Amino Acid Sequence

- Classic Method – Edman Rxn.
  - ✓ Edman Reaction
  - ✓ Enzymatic Fragmentation
  - ✓ Chemical Fragmentation
- DNA Sequencing
- MS (Proteomics)

Determination of Amino Acid Sequence

Edman Rxn



Phenylisothiocyanate

Polypeptide chain

Phenylthiocarbamoyl (PTC) peptide (or protein)

Phenylthiohydantoin

Polypeptide chain (minus original NH_2– terminal amino acid)

Determination of Amino Acid Sequence

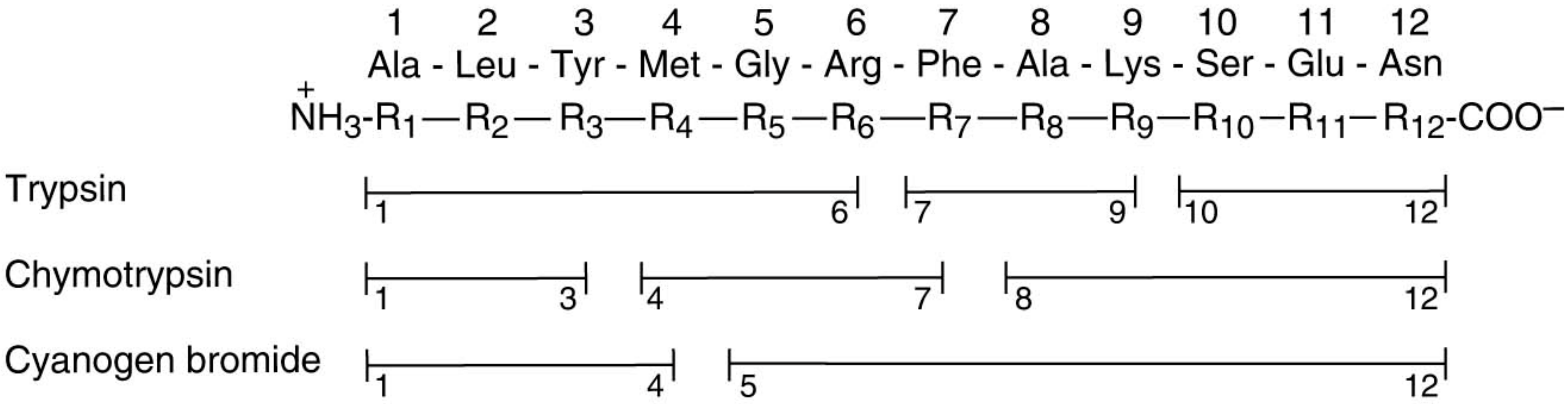
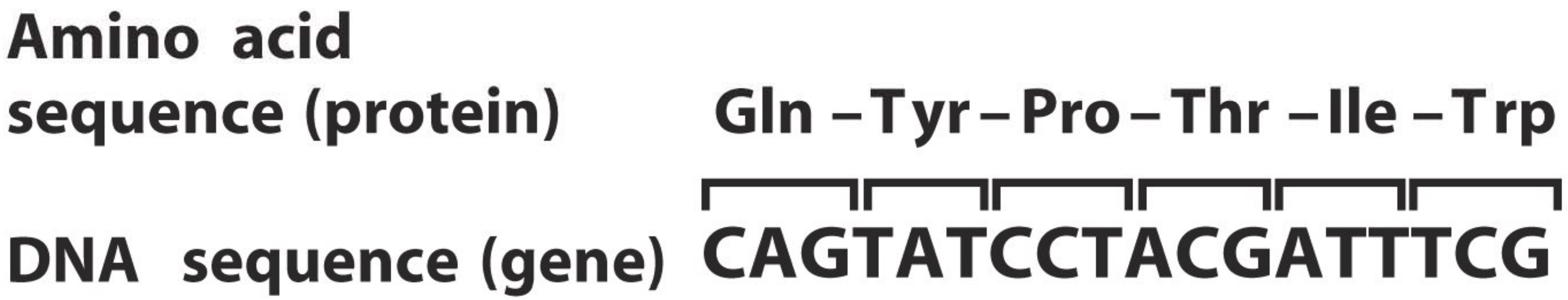


Figure 3.67. Ordering of peptide fragments from overlapping sequences produced by specific proteolysis of a peptide.

In the Post-Genome Era

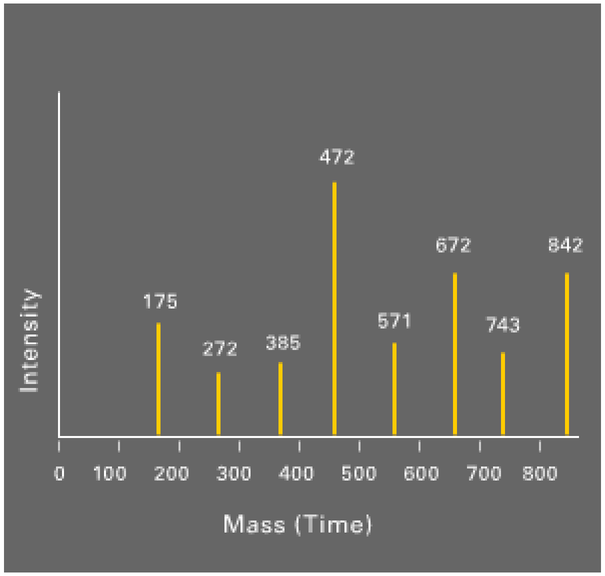
DNA → Transcription → Translation → Protein

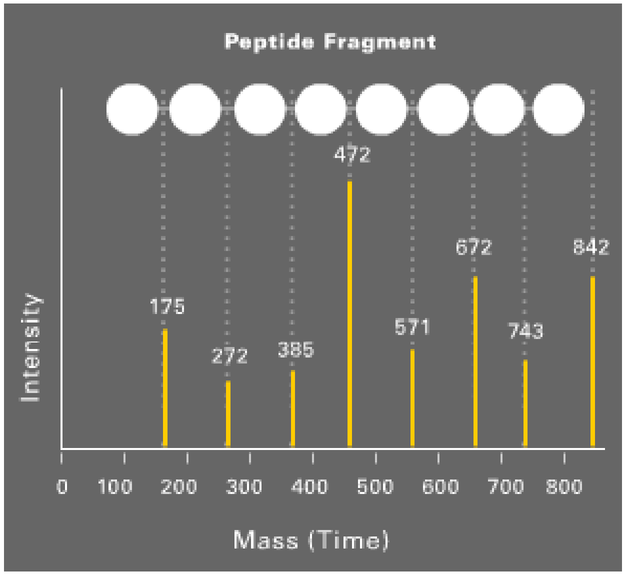


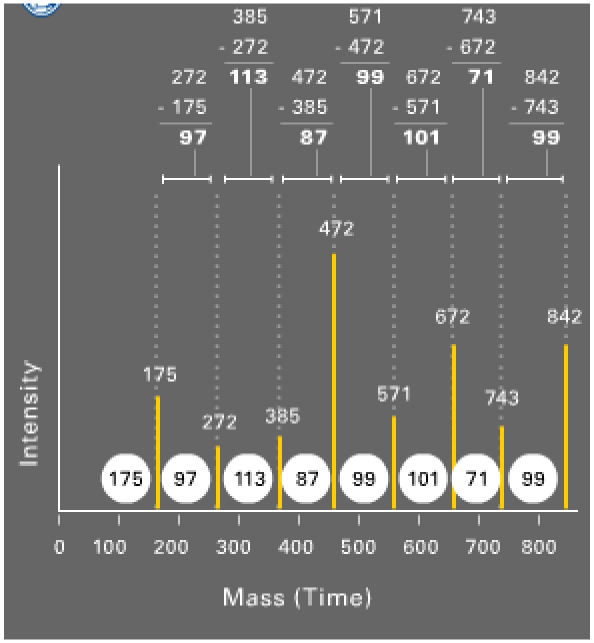
- Method to predict the amino acid sequence of a protein
- Disadvantage:
  - ✓ Do not predict the position of disulfide bonds
  - ✓ Do not identify modified aa (post-translation)

Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

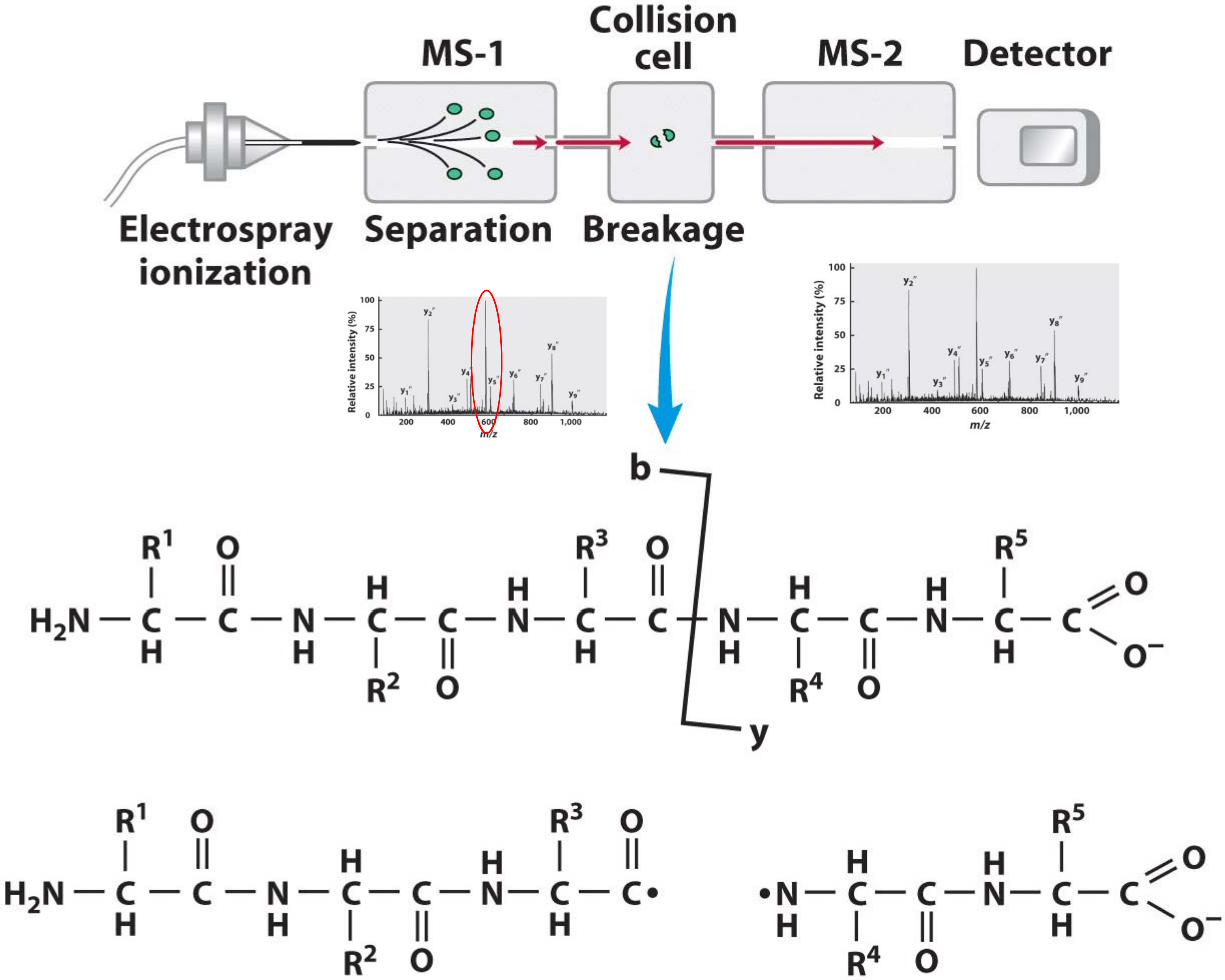
Mass Spectrometry for Protein Sequence Determination



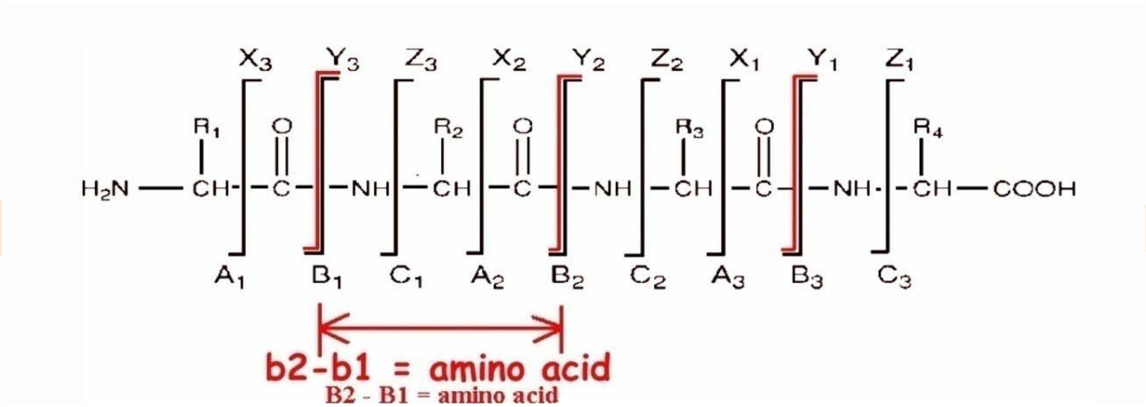


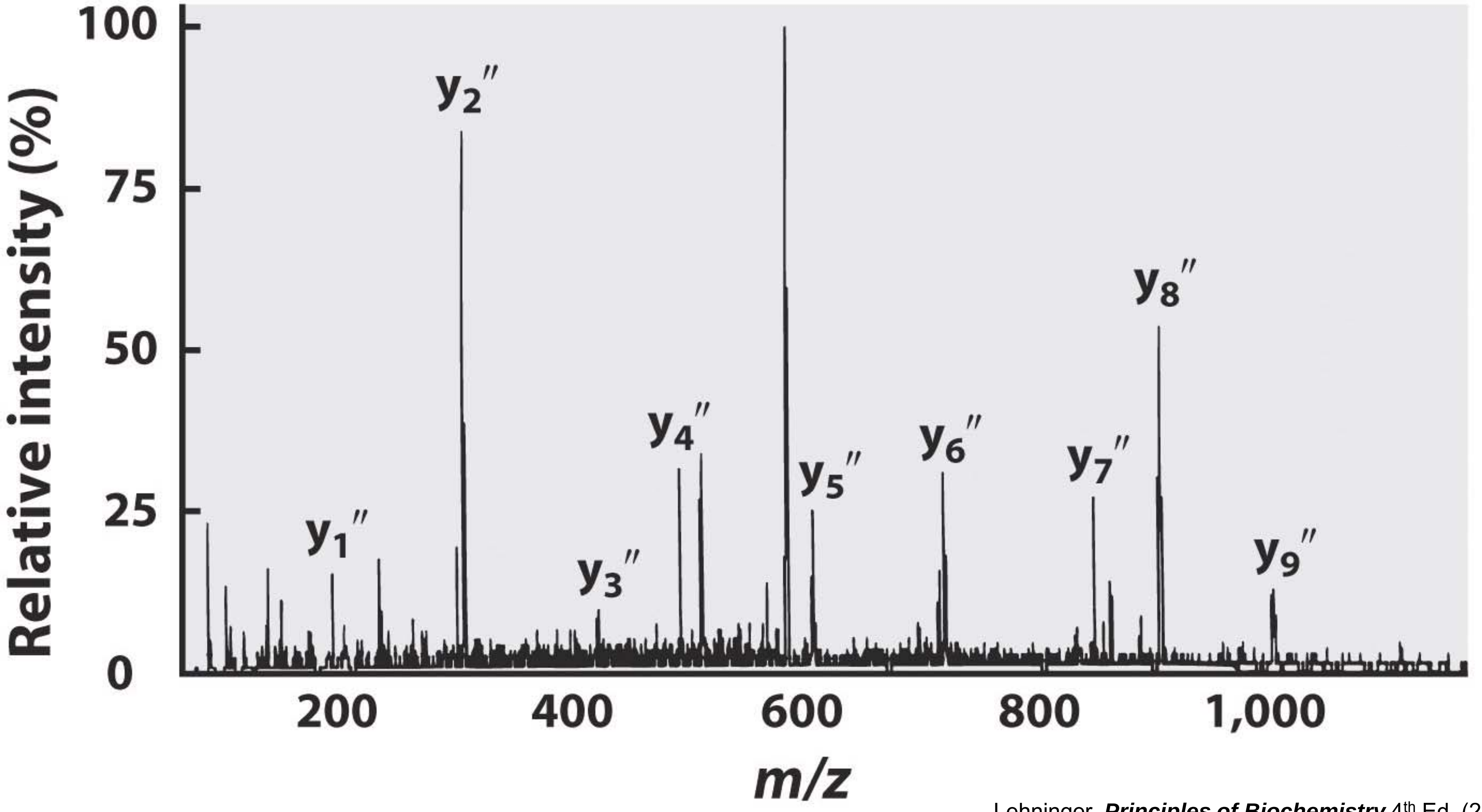


http://www.childrenshospital.org/cfapps/research/data_admin/Site602/mainpageS602P0.htm as on August 2007l



Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)





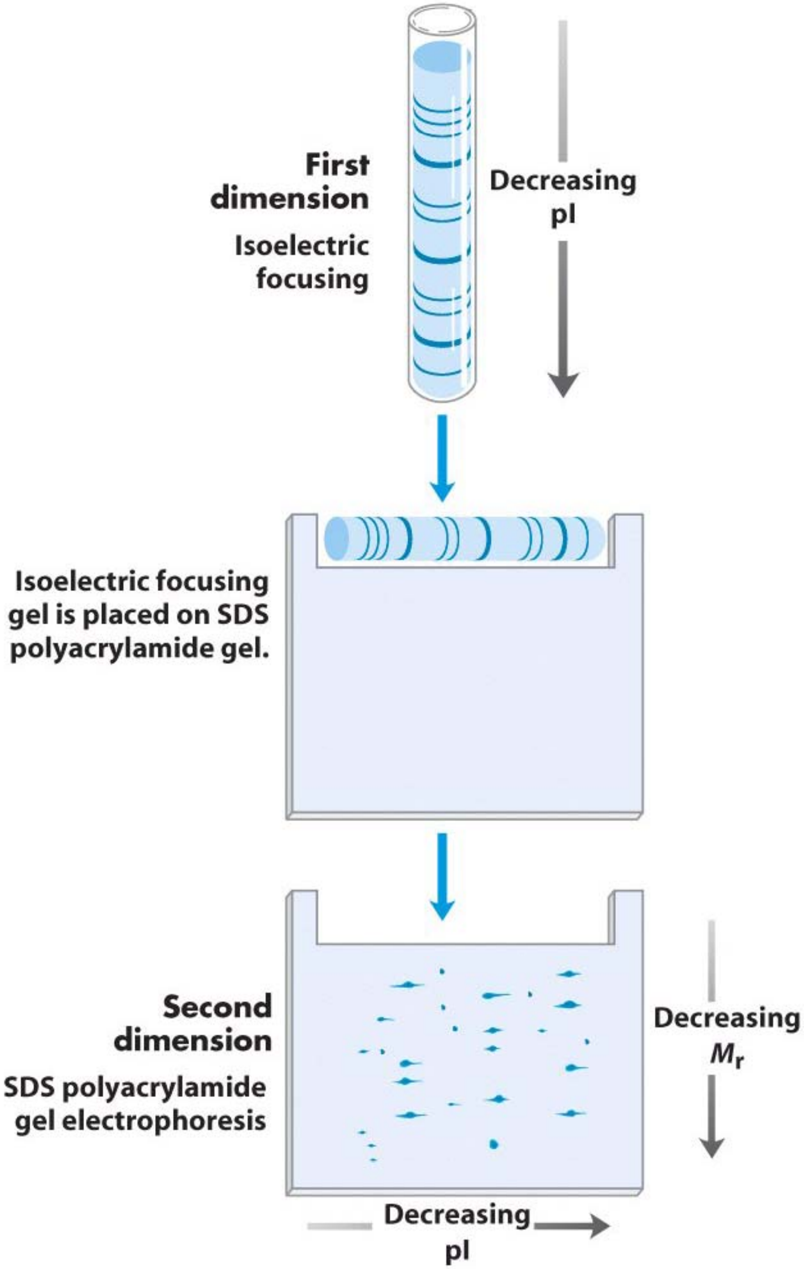
Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

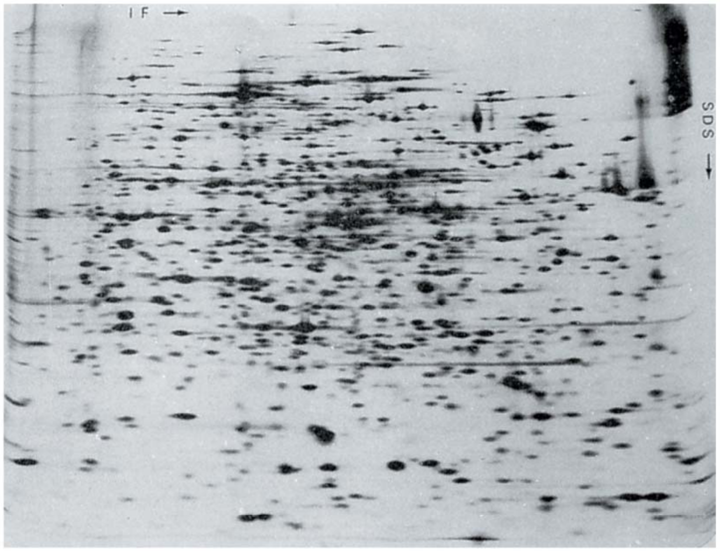
Proteomics

- Systematic study of the amounts, modifications, interactions, localization and function of all or a subsets of proteins at the whole-organism, tissue, cellular and sub-cellular levels.
- Can address questions as:
  - ✓ In a given sample, what fraction of the proteoma is expressed?
  - ✓ Of those present – relative abundance?
  - ✓ Chemically modified?
  - ✓ Protein profile change due to a condition? Cancer changes the profile of serum proteins? Response to tx alter profile?

Lodish, *Molecular Cell Biology* 6th Ed.

2D Electrophoresis

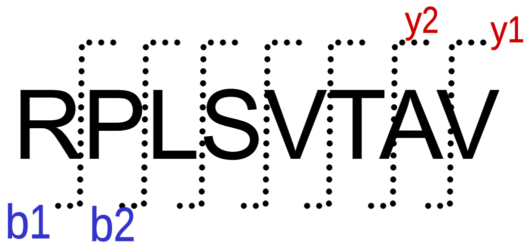


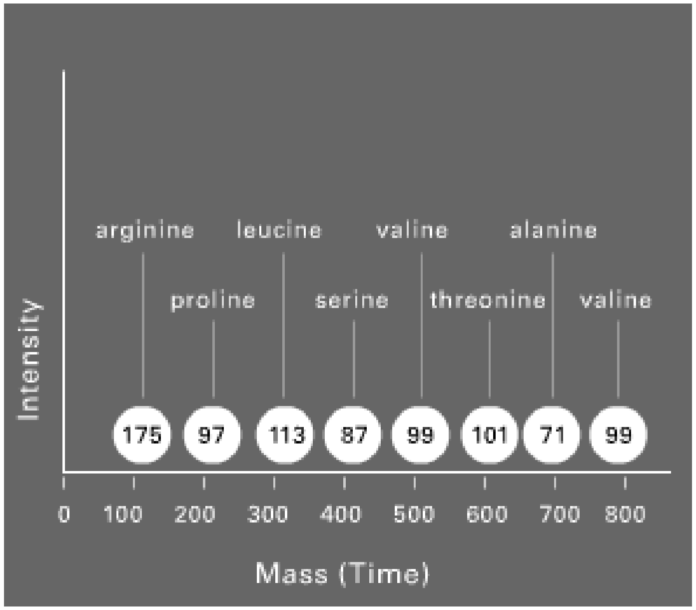


Lehninger, ***Principles of Biochemistry*** 4th Ed. (2005)

Symbol	Structure	Mass (Da)
Ala A	-NH.CH.(CH$_3$).CO-	71.0
Arg R	-NH.CH.[(CH$_2$)$_3$.NH.C(NH).NH$_2$].CO-	156.1
Asn N	-NH.CH.(CH$_2$CONH$_2$).CO-	114.0
Asp D	-NH.CH.(CH$_2$COOH).CO-	115.0
Cys C	-NH.CH.(CH$_2$SH).CO-	103.0
Gln Q	-NH.CH.(CH$_2$CH$_2$CONH$_2$).CO-	**128.1**
Glu E	-NH.CH.(CH$_2$CH$_2$COOH).CO-	129.0
Gly G	-NH.CH$_2$.CO-	57.0
His H	-NH.CH.(CH$_2$C$_3$H$_3$N$_2$).CO-	137.1
Ile I	-NH.CH.[CH.(CH$_3$)CH$_2$.CH$_3$].CO-	**113.1**
Leu	-NH.CH.[CH$_2$CH(CH$_3$)$_2$].CO-	**113.1**
Lys K	-NH.CH.[(CH$_2$)$_4$NH$_2$].CO-	**128.1**
Met M	-NH.CH.[(CH$_2$)$_2$.SCH$_3$].CO-	131.0
Phe F	-NH.CH.(CH$_2$Ph).CO-	147.1
Pro P	-NH.(CH$_2$)$_3$.CH.CO-	97.1
Ser S	-NH.CH.(CH$_2$OH).CO-	87.0
Thr T	-NH.CH.[CH(OH)CH$_3$).CO-	101.0
Trp W	-NH.CH.[CH$_2$.C$_8$H$_6$N].CO-	186.1
Tyr Y	-NH.CH.[(CH$_2$).C$_6$H$_4$.OH].CO-	163.1
Val V	-NH.CH.[CH(CH$_3$)$_2$].CO-	99.1

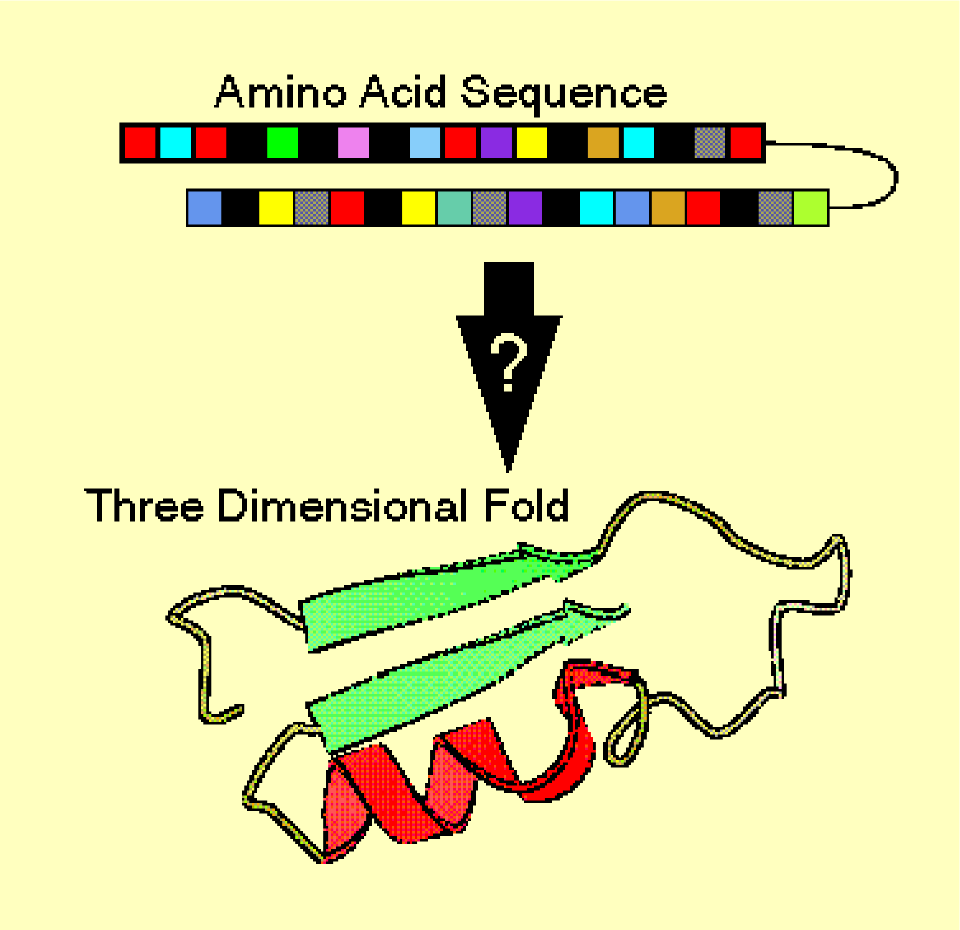
Proteomics





http://www.childrenshospital.org/cfapps/research/data_admin/Site602/mainpageS602P0.html

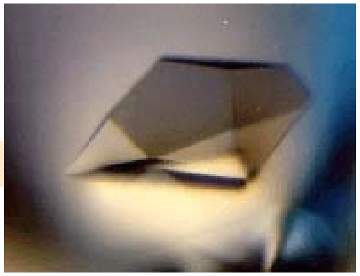
Determination 3D Structure



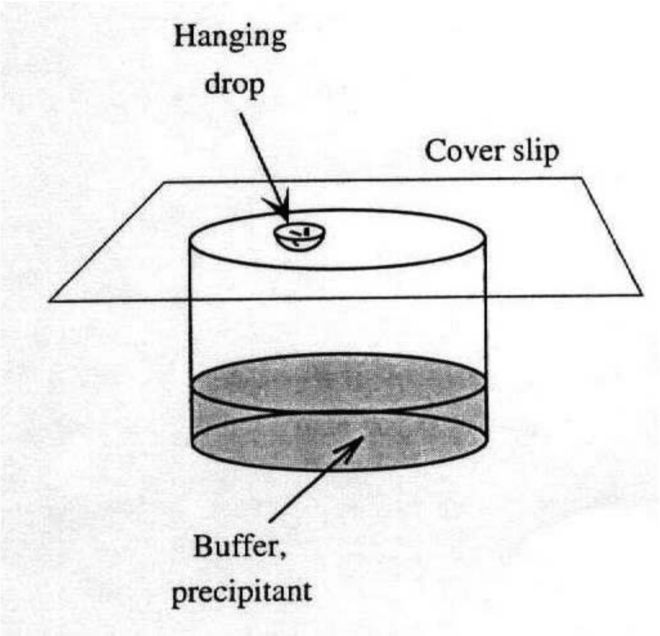
- X-ray Difration
- NMR

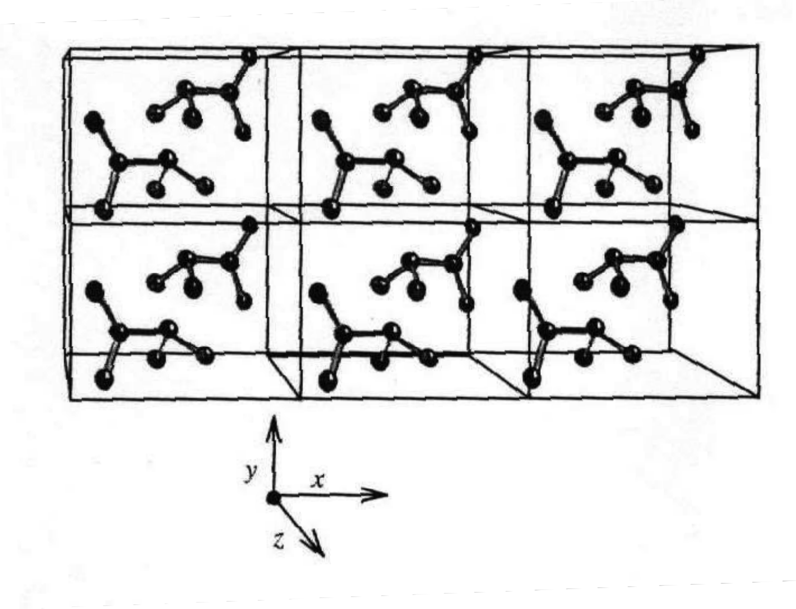
http://biop.ox.ac.uk/www/mol_of_life/PFP_BV.html - as on Aug. 2007

Crystallography



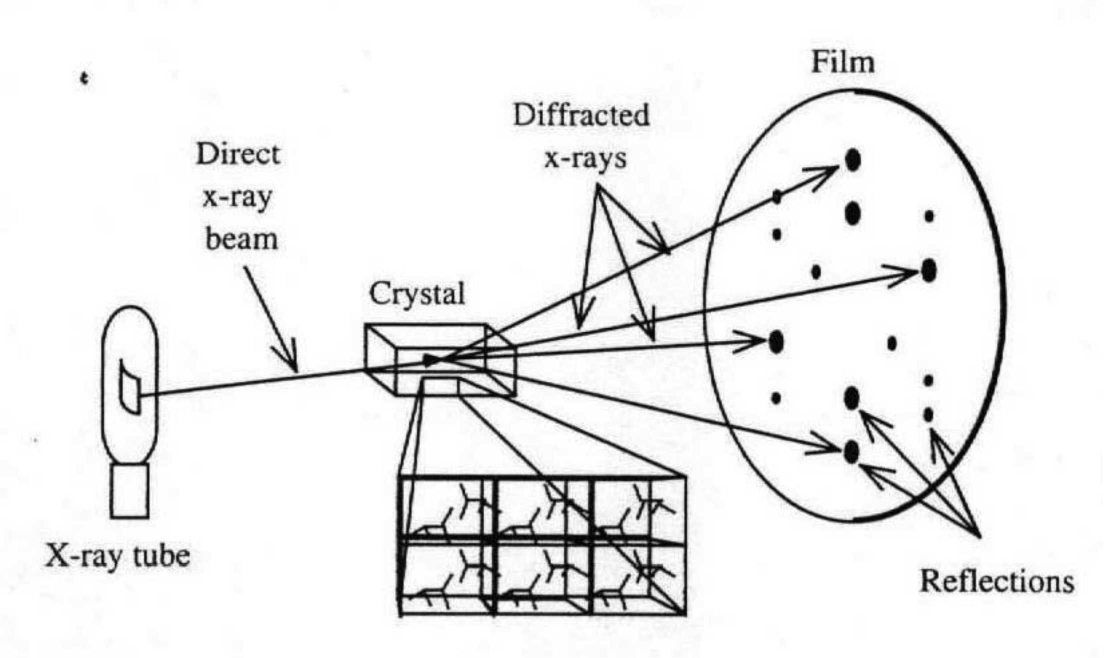
- Crystals
  - ✓ Method: “salting out”

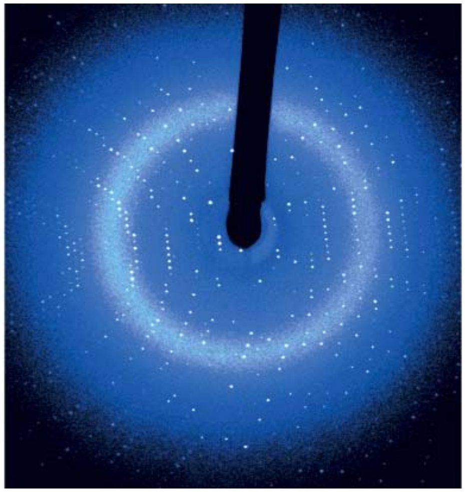




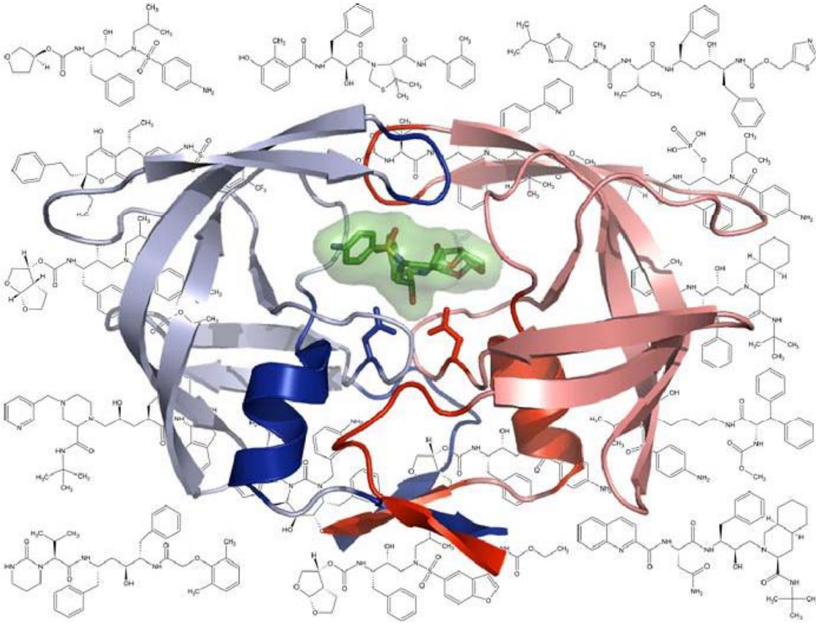
Crystallography – Data Collection

- ✓ Expose crystal to X-rays



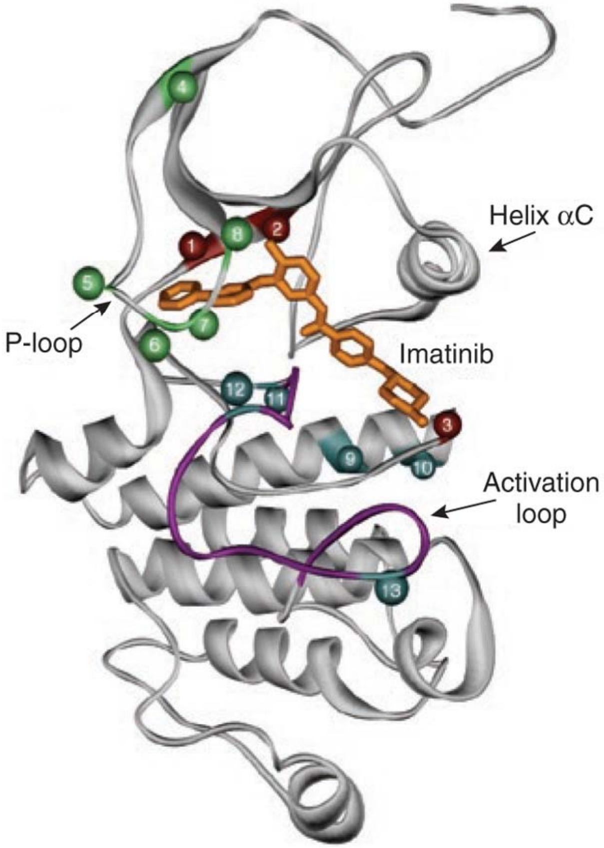


Protein Structures in Medicine



HIV protease

Essential for the maturation of HIV



cAbl kinase

Constitutively active in Chronic myelogenous leukemia

NMR

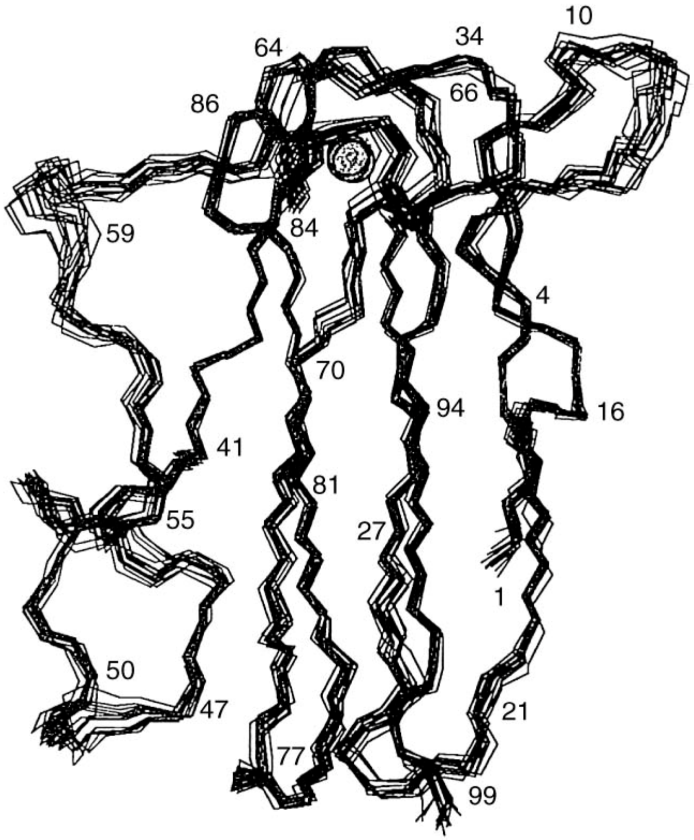


Figure 3.76. NMR structure of plastocyanin from the French bean. From Moore, J.M., Lepre, C. A., Gippert, G. P., Chazin, W. J., Case, D. A., and Wright, P. E. J. *Mol. Biol.* 221:533, 1991. Figure generously supplied by P. E. Wright.

- Advantage:
  - ✓ No crystal needed
- Disadvantage:
  - ✓ limited to small proteins 20kDa

Devlin, *Biochemistry 6th ed* (2006)