BCMB 3100 – Chapter 3 (part 1)

- **Diversity of protein function**
- Complete definition of amino acids
- Memorize complete structure of 20 common amino acids!!!
- pKa's of α amino and α carboxyl groups
- Amino acids with ionizable side groups
- Titration curve

1

PROTEIN STRUCTURE AND FUNCTION

Mulder (1800s): showed albumins (milk and egg) contained ____________

Berzelius (1838): named such substances ________________

from Greek proteios "of first rank"

Proteins recognize and ___________ many different types of molecules & ___________ most of the ____________________ necessary for life.

2

Examples of Protein Function

____________________:
enzymes increase reaction rates by ≥10^6-fold; nearly all known enzymes are proteins

____________________:
many small molecules and ions are transported by proteins; examples: hemoglobin, myoglobin

____________________:
examples: myosin and actin (muscle movement)

3

Examples of Protein Function

__________________: collagen (tensile strength to skin and bones)

__________________: antibodies

____
____________: example: acetylcholine receptor

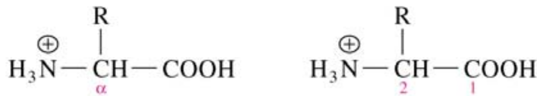
________________________________:
DNA binding proteins (repressors, activators, transcription factors, hormones, regulation of translation

4

WHAT ARE PROTEINS?

Proteins are macromolecules made up of amino acids (____________________)

α-amino acids: consist of an ____________, a _______________, a ________________ and a distinctive __________ bonded to a carbon atom. This carbon is called the _________ because it is adjacent to the carboxyl group.



BCMB 3100 - Chapter 3

- Diversity of protein function
- **Complete definition of amino acids**
- Memorize complete structure of 20 common amino acids!!!
- pKa's of α amino and α carboxyal groups
- Amino acids with ionizable side groups
- Titration curves & pI

6

Amino acids in solution will be in a charged state. The amino group and/or the carboxyl group will be charged depending upon the pH. The R group may also be charged.

At neutral pH amino acids are predominantly __________________ (zwitterions)

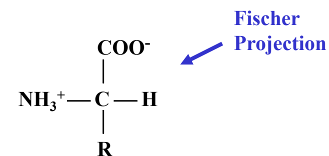
- Under normal cellular conditions amino acids are **zwitterions** (dipolar ions):

  Amino group = **$-NH_3^+$**

  Carboxyl group = **$-COO^-$**

7

General Structure of the ionized from of an amino acid



Fischer Projection

α COO^-: pKa 1.8-2.5

α NH^+: pKa 8.7-10.7

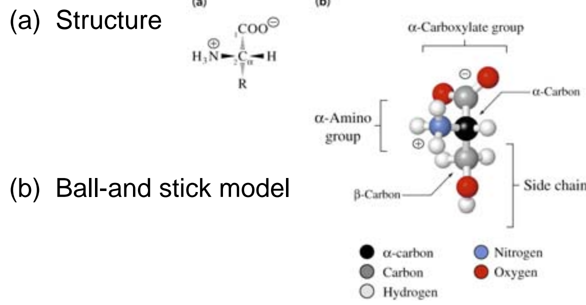
You MUST know this!!!

8

Fig 3.1 Two representations of an amino acid at neutral pH

(a) Structure

(b) Ball-and stick model



Amino acids are __________ (asymmetric) due to the tetrahedral array of 4 different groups around the α-carbon (glycine is an exception).

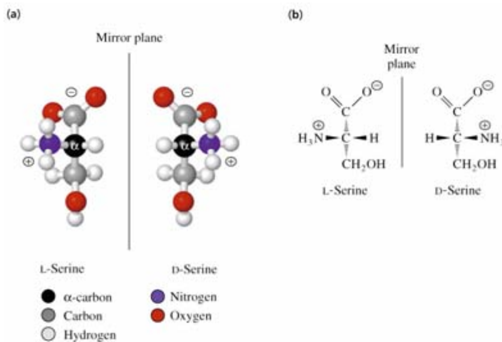
Thus all amino acids except glycine can exist as enantiomers: two sterioisomer that are nonsuperimposable mirror images of each other.

Enantiomers of amino acids are called D (right-handed) or L (left-handed)

L and D refer to absolute configuration

________________ are the only constituents of proteins

Fig 3.2 Mirror-image pairs of amino acids



20 different amino acids are found in proteins

____________________ (side chains) that differ in size, shape, charge, hydrogen-bonding capacity & chemical reactivity → 20 different amino acids found in proteins of all organisms from bacteria to humans

The amino acid alphabet is at least 2 billion years old

The diversity of protein structure & function is due to the sequence and number of amino acids found in a protein (primary structure)

It is essential that a biochemist commit to memory the structure of the 20 amino acids! Know this!! (*structure, & 1 & 3 letter code*)

BCMB 3100 - Chapter 3

- Diversity of protein function
- Complete definition of amino acids
- **Memorize complete structure of 20 common amino acids!!!**
- pKa's of α amino and α carboxyal groups
- Amino acids with ionizable side groups
- Titration curves & pI

13

3.2 Structures of the 20 Common Amino Acids

- You must learn the one and three letter abbreviations
- You must know the properties of their side chains (R groups)
- Classes: , , , , , and
- You must be able to draw the structure of the 20 amino acids in their correct ionized form at any given pH

14

Four __________amino acid structures

Glycine [G] (Gly)

Alanine [A] (Ala)

Valine [V] (Val)

Leucine [L] (Leu)

15

Fig 3.3 Stereoisomers of Isoleucine

Isoleucine [I] (Ile) **Aliphatic amino acid** Page 58

- Ile has 2 chiral carbons, 4 possible stereoisomers

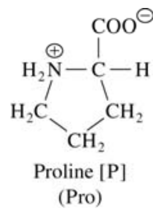
****

L-Isoleucine D-Isoleucine L-Alloisoleucine D-Alloisoleucine

16

Proline has a nitrogen in the ring system

- **Proline** (**Pro**, **P**) - has a three carbon side chain bonded to the α-amino nitrogen
- The heterocyclic pyrrolidine ring restricts the geometry of polypeptides



Proline [P]
(Pro)

17

R Groups

- Side chains have aromatic groups

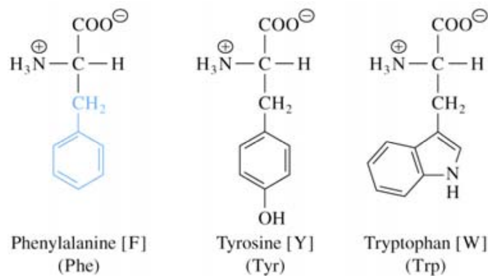
Phenylalanine (**Phe**, **F**) - benzene ring (OD260 nm)

Tyrosine (**Tyr**, **Y**) - phenol ring (OD280 nm)

Tryptophan (**Trp**, **W**) - bicyclic indole group (OD280 nm)

18

Aromatic amino acid structures



Phenylalanine [F] (Phe) Tyrosine [Y] (Tyr) Tryptophan [W] (Trp)

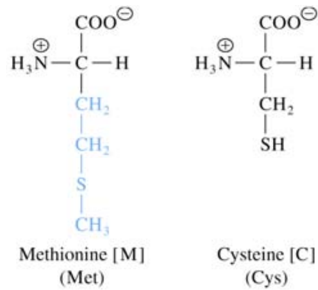
19

R Groups

- **Methionine** (**Met**, **M**) - ($-CH_2CH_2SCH_3$)
- **Cysteine** (**Cys**, **C**) - ($-CH_2SH$)
- Two cysteine side chains can be cross-linked by forming a disulfide bridge ($-CH_2-S-S-CH_2-$)
- Disulfide bridges may stabilize the three-dimensional structures of proteins

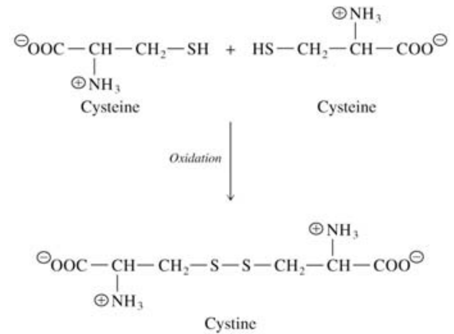
20

Methionine and cysteine



21

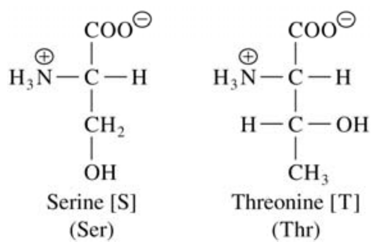
Fig 3.4 Formation of cystine



22

Side Chains with ______________

Serine (**Ser**, **S**) and **Threonine** (**Thr**, **T**) have uncharged polar side chains



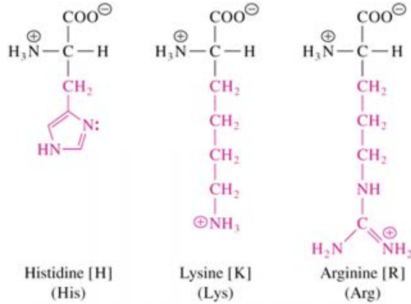
23

__________R Groups

- **Histidine** (**His**, **H**) - imidazole
- **Lysine** (**Lys**, **K**) - alkylamino group
- **Arginine** (**Arg**, **R**) - guanidino group
- Side chains are nitrogenous bases which are substantially positively charged at pH 7 (true for K & R) [NOTE: error in book, pg 59]

24

Structures of histidine, lysine and arginine



Histidine [H] (His) | Lysine [K] (Lys) | Arginine [R] (Arg)

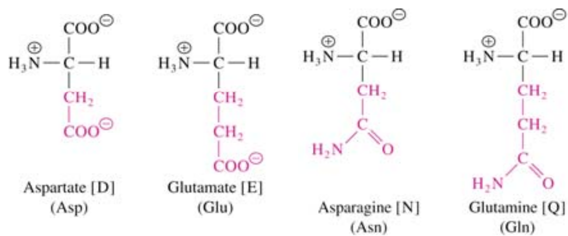
25

________ R Groups and ________ Derivatives

- **Aspartate** (**Asp**, **D**) and **Glutamate** (**Glu**, **E**) are dicarboxylic acids, and are negatively charged at pH 7
- **Asparagine** (**Asn**, **N**) and **Glutamine** (**Gln**, **Q**) are uncharged but highly polar

26

Structures of aspartate, glutamate, asparagine and glutamine



Aspartate [D] (Asp) | Glutamate [E] (Glu) | Asparagine [N] (Asn) | Glutamine [Q] (Gln)

27

______________ of Amino Acid Side Chains

- ______________: the relative hydrophobicity of each amino acid
- The larger the hydropathy, the greater the tendency of an amino acid to prefer a hydrophobic environment
- Hydropathy affects protein folding: ____________ chains tend to be in the ________ ____________ tend to be on the __________

28

Challenge of the Week

given out on 8-20-08

Work with your group of 4 people and find at least one example of a mutation in humans, or in industry-relevant plants, animals or microbes. Present, on a single, one-side, typed page, the amino acid mutated, the phenotype of the effect on the organism, the molecular reason that the mutation causes the effect(s), and the effect that this mutation has on/for humans.

Hand in a single, one-sided, typed sheet of paper with ALL group members names (first and last name correctly spelled) as well as a single sentence behind each name describing their contribution to the answer. **These will be collected ONLY at the Breakout Session on August 31.**

________________ of Amino Acid Side Chains

- ____________: the relative hydrophobicity of each amino acid
- The larger the hydropathy, the greater the tendency of an amino acid to prefer a hydrophobic environment
- Hydropathy affects protein folding:
  - *__________ side chains tend to be in the interior
  - *__________ residues tend to be on the surface

Table 3.1

- Hydropathy scale for amino acid residues

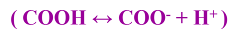
(Free-energy change for transfer of an amino acid from interior of a lipid bilayer to water)

Amino acid	Free-energy change for transfer ($kJ mol^{-1}$)
Highly hydrophobic	
Isoleucine	3.1
Phenylalanine	2.5
Valine	2.3
Leucine	2.2
Methionine	1.1
Less hydrophobic	
Tryptophan	1.5*
Alanine	1.0
Glycine	0.67
Cysteine	0.17
Tyrosine	0.08
Proline	-0.29
Threonine	-0.75
Serine	-1.1
Highly hydrophilic	
Histidine	-1.7
Glutamate	-2.6
Asparagine	-2.7
Glutamine	-2.9
Apartate	-3.0
Lysine	-4.6
Arginine	-7.5

BCMB 3100 - Chapter 3

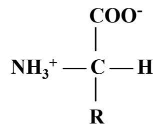
- Diversity of protein function
- Complete definition of amino acids
- Memorize complete structure of 20 common amino acids!!!
- **pKa's of α amino and α carboxyal groups**
- Amino acids with ionizable side groups
- Titration curves & pI

Amino acids in solution will be in a __________. The amino group and/or the carboxyl group will be charged depending upon the pH. The R group may also be charged. **At neutral pH amino acids are predominantly dipolar ions (____________).**



(COOH ↔ COO^- + H^+)

(NH_3^+ ↔ NH_2 + H^+)



Note: pH dependance of the ionization state

α COO^-: pKa 1.8-2.5

α NH^+: pKa 8.7-10.7

Ionization of Amino Acids

- Ionizable groups in amino acids: (1) α-**carboxyl**, (2) α-**amino**, (3) **some side chains**
- Each ionizable group has a specific pK_a

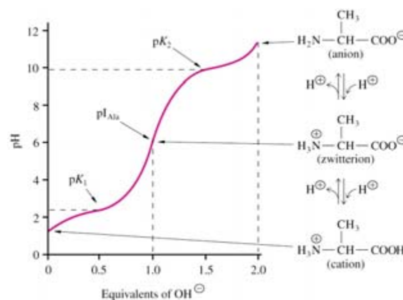
AH ⇌ B + H^+

- For a solution pH below the pK_a, the protonated form predominates (**AH**)
- For a solution pH above the pK_a, the unprotonated (conjugate base) form predominates (**B**)

Fig 3.6 Titration curve for alanine

- Titration curves are used to determine pK_a values
- pK_1 = 2.4
- pK_2 = 9.9
- pI_{Ala} = isoelectric point

(pI = pH when net charge is zero)



pI = ________________

pKa values of amino acids

All amino acids contain two ionizable groups:

α-carboxyl (pKa 1.8-2.5)

and protonated α-amino group (pKa 8.7-10.7)

(Table 3.2, pg 64 gives specific pKa values)

The ____________________ are also ionizable. The pKa of these R groups can range from 3.9 to 12.5. You must memorize the pKa's of the side chains of these seven amino acids!

BCMB 3100 - Chapter 3

- Diversity of protein function
- Complete definition of amino acids
- Memorize complete structure of 20 common amino acids!!!
- pKa's of α amino and α carboxyal groups
- **Amino acids with ionizable side groups**
- Titration curves & pI

37

Table 3.2

pK_a values of amino acid ionizable groups

Amino acid	pKa value		
	α-Carboxyl group	α-Amino group	Side chain
Glycine	2.4	9.8	
Alanine	2.4	9.9	
Valine	2.3	9.7	
Leucine	2.3	9.7	
Isoleucine	2.3	9.8	
Methionine	2.1	9.3	
Proline	2.0	10.6	
Phenylalanine	2.2	9.3	
Tryptophan	2.5	9.4	
Serine	2.2	9.2	
Threonine	2.1	9.1	
Cysteine	1.9	10.7	8.4
Tyrosine	2.2	9.2	10.5
Asparagine	2.1	8.7	
Glutamine	2.2	9.1	
Aspartic acid	2.0	9.9	3.9
Glutamic acid	2.1	9.5	4.1
Lysine	2.2	9.1	10.5
Arginine	1.8	9.0	12.5
Histidine	1.8	9.3	6.0

BCMB 3100 - Chapter 3

- Diversity of protein function
- Complete definition of amino acids
- Memorize complete structure of 20 common amino acids!!!
- pKa's of α amino and α carboxyal groups
- Amino acids with ionizable side groups
- **Titration curves & pI**

39

Fig 3.7 (pg 64) Ionization of Histidine

(a) Titration curve of histidine

$pK_1 = 1.8$
$pK_2 = 6.0$
$pK_3 = 9.3$

pI = pH at which net charge is zero

To determine the pI: pick the pKa above & below point where charge is zero and average those 2 pKa's

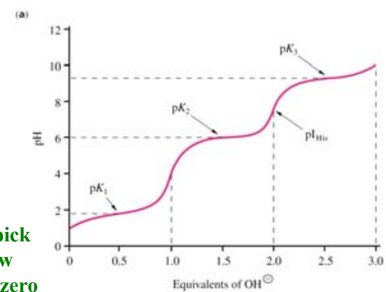
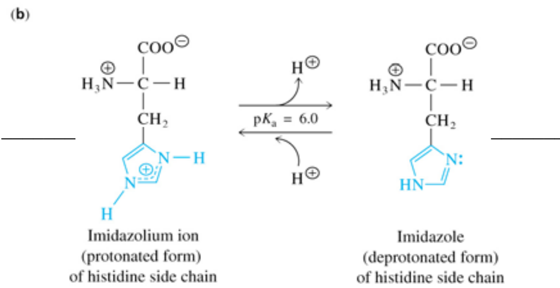


Fig 3.7 (b) Deprotonation of imidazolium ring



41

Fig 3.7 (pg 64) Ionization of Histidine

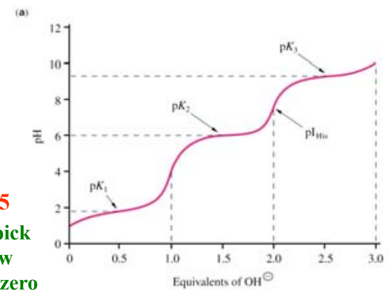
(a) Titration curve of histidine

pK_1 = 1.8
pK_2 = 6.0
pK_3 = 9.3

pI = pH at which net charge is zero

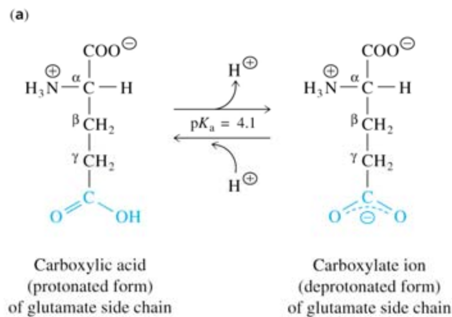
pI = (6 + 9.3)/2 =7.65

To determine the pI: pick the pKa above & below point where charge is zero and average those 2 pKa's



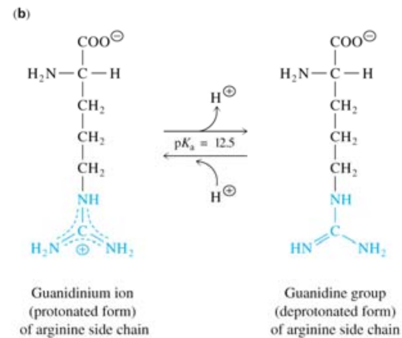
42

Fig 3.8 (a) Ionization of the protonated γ-carboxyl of glutamate



43

Fig 3.8 (b) Deprotonation of the guanidinium group of Arg



44

BCMB 3100 – Chapter 3 (part 2)

- **Summary of amino acids**
- Polypeptides: definition, structure, and direction
- Peptide bond
- Disulfide bonds
- Protein purification
- Methods to determine amino acid composition, cleavage of proteins, protein sequencing
- Diversity of proteins

The 20 amino acid are divided into 7 groups

- **Aliphatic: Gly, Ala, Val, Leu, Ile, Pro**
- **Aromatic: Phe, Tyr, Trp**
- **Sulfur-containing: Met, His**
- **Alcohols: Ser, Thr**
- **Basic: His, Lys, Arg**
- **Acidic: Asp, Glu**
- **Amides: Asn, Gln**

46

BCMB 3100 - Chapter 3

- Quick review on amino acids
- **Polypeptides: definition, structure, and direction**
- Disulfide bonds
- Protein purification
- Methods to determine amino acaid composition, cleavage of proteins, protein sequencing

47

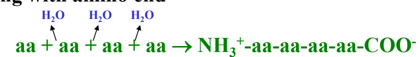
Amino acids are linked by peptide bonds to form polypeptide chains

α carboxyl of one amino acid is joined to α-amino group of another amino acid by a peptide bond (amide bond) with concomitant loss of H_2O

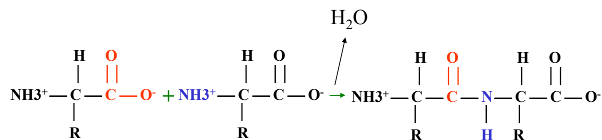
Many amino acids are joined by peptide bonds to form an unbranched polypeptide

Residue: amino acid unit in a polypeptide

By convention the direction of a polypeptide is written starting with amino end



Formation of a dipeptide



49

Most polypeptides contain between ___________ amino acids

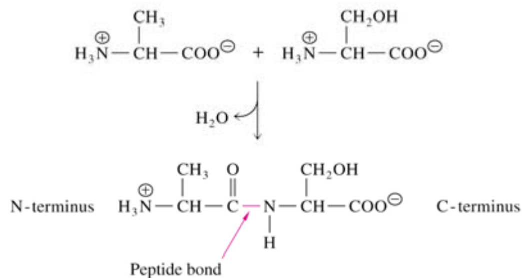
Average M.W. for an amino acid is ___________ so M.W. of most proteins is 5500 to 220,000 daltons. (One dalton equals one atomic mass unit; kilodalton = 1000 daltons). Most proteins have M.W. of 5.5-220 kd.

Some proteins contain disulfide bonds that cross-link between _________residues by the oxidation of ___________. The resulting disulfide is called cystine.

Intracellular proteins often lack disulfides while extracellular proteins often have them.

50

Fig 3.9 Peptide bond between two amino acids

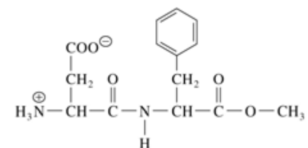


Formation of peptide bonds eliminates the ionizable α-carboxyl and α -amino groups of the free amino acids except for those at the amino and carboxy termini

51

Fig 3.10 Aspartame, an artificial sweetener

- Aspartame is a dipeptide methyl ester (aspartylphenylalanine methyl ester)
- About 200 times sweeter than table sugar
- Used in diet drinks



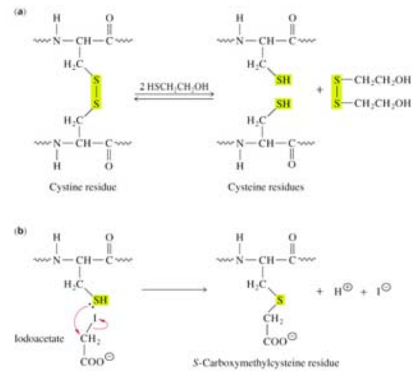
52

BCMB 3100 - Chapter 3

- Memorize complete structure of 20 common amino acids!!!
- pKa's of α amino and α carboxyl groups
- Amino acids with ionizable side groups
- Titration curves & pI
- Polypeptides: definition, structure, and direction
- Peptide bond
- **Disulfide bonds**
- Protein purification
- Methods to determine amino acaid composition, cleavage of proteins, protein sequencing

53

Fig 3.19 Cleaving, blocking disulfide bonds



54

BCMB 3100 - Chapter 3

- **Polypeptides: definition, structure, and direction**
- **Peptide bond**
- **Disulfide bonds**
- **Protein purification**
- **Methods to determine amino acaid composition, cleavage of proteins, protein sequencing**

55

PROTEIN PURIFICATION

General strategy

Tissue → disrupt → crude fractionation → selected fractionation

Proteins can be separated by:

___________: SDS gel electrophoresis, gel filtration chromatography, dialysis, centrifugation

________________: salting out

________________: ion-exchange chromatography

________________: interaction chromatography, reverse phase chromatography

________________: affinity chromatography, antibodies

Purified proteins can be analyzed separated by:

Sequencing: Edman degradation, (proteolysis)

3D structure: X-ray crystallography, NMR

Synthesis: automated solid phase

56

To purify large amounts of proteins one requires:

1. An ______ for the protein (enzyme, antibody, etc)

2. A ________________ to separate desired protein from "all" other proteins and which retains protein "activity"

3. Example strategy:

Salting-out: the "specific" precipitation of a given protein at a specific high-salt concentration

Ion exchange → Gel-filtration → Affinity chromatography

Dialysis often used to remove salts: separation of protein from small molecules through a semipermeable membrane (cellulose)

57

Different types of chromatography (1)

_________________________: proteins passed over a column filled with a hydrated porous beads made of a carbohydrate or polyacrylamide polymer (large molecules exit first)

___________________________: separation of proteins over a column filled with charged polymer beads (+ charge = anion-exchange; - charge = cation exchange. Positively charged proteins bind to beads of negative charge & vice versa Bound proteins are eluted with salt. Least charged proteins will elute first.

__________________________: proteins are passed through a column of beads containing a covalently bound high affinity group for the protein of interest. Bound protein is eluted by free high affinity group.

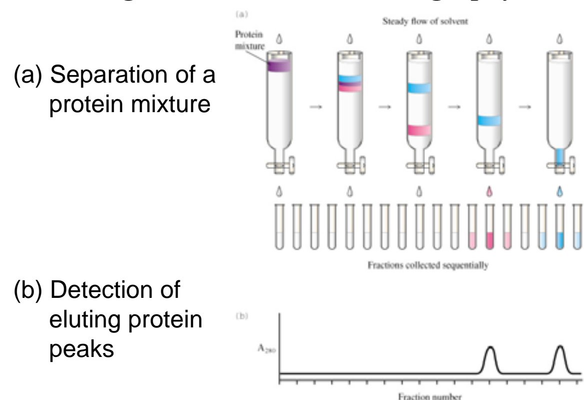
58

Different types of chromatography (2)

Chromatography based on hydrophobicity of protein:
____ (HIC) and __________________________ (RPC) are both based on interactions between hydrophobic patches on the surface of a protein and on the hydrophobicity of ligands (e.g. alkyl groups) covalently attached to a gel matrix. In RPC proteins can bind very strongly to the gel and require non-polar solvents for their elution. In HIC protein binding is promoted by inclusion of salt in the solvent and elution of proteins is caused by decreasing or removing salt from the solvent.

59

Fig 3.11 Column Chromatography

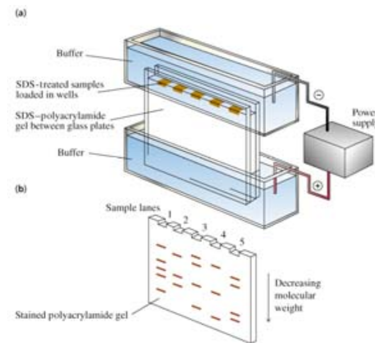


_________________: movement of charged solutes through a gel in response to an electric field

________________________________:
chemically inert; polymerized acrylamide matrix of controlled pore size; allows separation of proteins based on mass and charge

___________: (sodium dodecyl sulfate, page 30): anionic detergent used for polyacrylamide gel electrophoresis. It complexes with proteins (1 SDS/2 amino acids) → denatured protein of negative charge proportional to protein mass. Note: reducing agents (mercaptoethanol, dithiothreitol) are also added to reduce disulfide bonds. Mobility of many proteins under these conditions is linearly proportional to the log of their mass. Smaller proteins migrate faster.

Fig 3.12 (a) SDS-PAGE Electrophoresis (b) Protein banding pattern after run



62

Proteins can also be separated by electrophoresis based on their native charge.

_________________: pH at which net protein charge is zero

_________________: electrophoresis of proteins (w/o SDS) in a pH gradient to a position in the gel at which pH = pI. pH gradient formed by polyampholytes (small multi-charged polymers of many pIs).

63

BCMB 3100 - Chapter 3

- Polypeptides: definition, structure, and direction
- Peptide bond
- Disulfide bonds
- Protein purification
- Methods to determine amino acid composition, cleavage of proteins, protein sequencing
- Diversity of proteins

64

Amino Acid Composition (1)

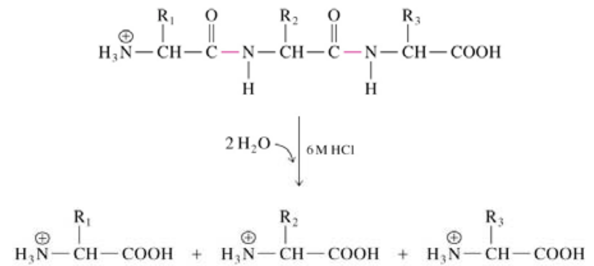
1. Determine amino acid composition

a. peptide —hydrolysis, 6 N HCL, 100°C, 24 hr→→ **free amino acids**

b. Amino acid composition of hydrolysates determined by automated cation-exchange chromatography (amino acid analyzer). Column contains solid granules of sulfonated polystyrene. Amino acids reacted with ninhydrin (yields colored product) & detected by O.D.

65

Fig 3.15 Acid-catalyzed hydrolysis of a peptide



Pg 72

Problem: Asn → Asp; Asn + Asp = Asx or B
Gln → Glu; Gln + Glu = Glx or Z
Loose some: Ser, Thr, Tyr; Loose most Trp

66

Amino Acid Composition (2)

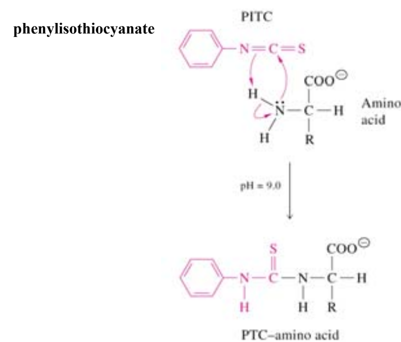
c. Another method for aa composition analysis is to treat protein hydrolysate with phenylisothiocyanate (PITC) at pH 9.0 to yield PITC-aa derivatives, separate by HPLC via hydrophobic attraction of aa side chains to hydrocarbon matrix of column and quantitate by OD 254nm (due to PTC moiety).

(The first pure protein analyzed for a.a. composition was β-lactoblobulin. It took several years of work. Today amino acid analyzer allows composition analysis within 2-4 hours with samples as small as 1 pmole!!!)

67

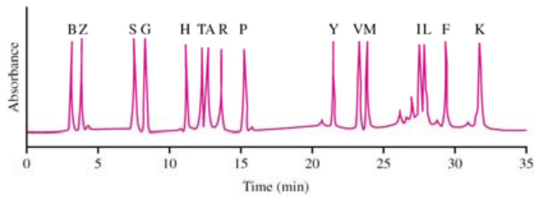
Fig 3.16 Amino acid treated with PITC

phenylisothiocyanate



68

Fig 3.16 Chromatogram from HPLC-separated PTC-amino acids



Note: acid treatment converts Asn→Asp; Gln→Glu; Trp is destroyed, some loss of Ser, Thr, Tyr.
B = (Asn→Asp) + Asp; Z = (Gln→Glu) + Glu

2. Determine amino-terminal residue

Sanger devised the first method to label N-terminal residue by reacting 2,4-dinitrophenylbenzene with α-amino group →→ yellow derivative. Subsequent hydrolysis (6N HCl) hydrolyzes away all other amino acids. N-terminal residue derivative identified by chromatography

Today Dabsyl chloride (→colored derivative) or Dansyl chloride (→ fluorescent derivative) are used.

3. Determination of amino acid sequence by Edman Degradation (Pehr Edman) (1950)

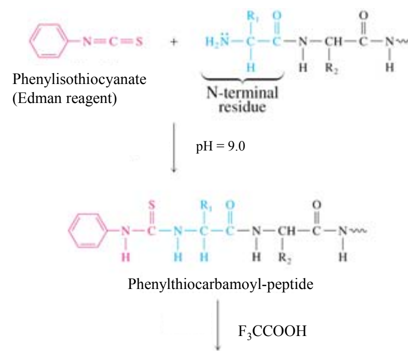
Edman Degradation is now done on sequenators (automated, Liquid phase sequenator →→ analysis by HPLC (one cycle 2 hr)

Gas-phase sequenator: detection of pmole amounts (single SDS-PAGE band)

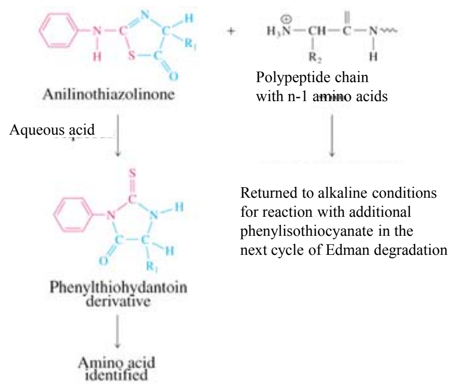
Frederick Sanger determined the first complete sequence of a protein (insulin) in 1953 (51 amino acid long)

Note: Amino acid sequence of small proteins and peptides is now commonly determined by mass spectrometry (e.g. electrospray MS, MALDI MS).

Fig 3.18 Edman degradation procedure



Edman degradation procedure (cont)



Protein Cleavage

Edman degradation limited to polypeptides of 50 a.a.

Chemical cleavage:

________________________ splits on carbonyl side of Met

Enzymatic cleavage:

______________: protease cleaves on carbonyl side of Arg and Lys

_________________: protease cleaves on carbonyl side of bulky hydrophobic and aromatic amino acids

Fig 3.20 Protein cleavage by BrCN

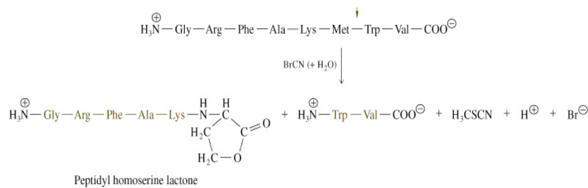
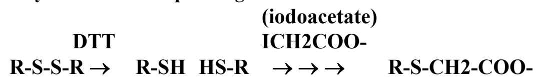


Fig 3.21 Cleavage, sequencing an oligopeptide

Disulfide bonds must be removed from protein by reduction & alkylation before sequencing.



DNA recombinant technology → DNA sequence of nascent protein

____________________: direct polypeptide product of translation (no modification)

Fig 3.19 Cleaving, blocking disulfide bonds

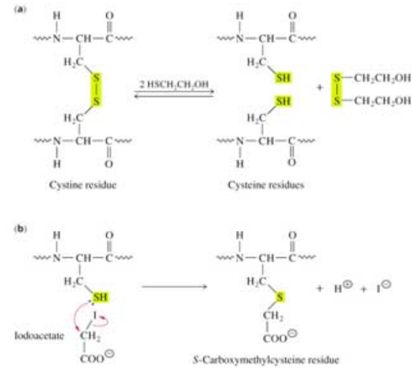
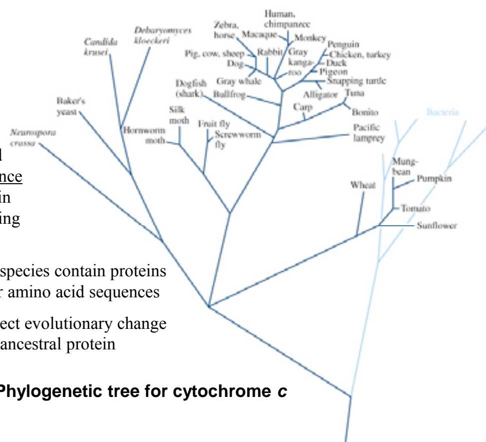


Fig 3.24

(see also Fig. 3.23)

- Protein amino acid sequences can be deduced from the sequence of nucleotides in the corresponding gene
- Closely related species contain proteins with very similar amino acid sequences
- Differences reflect evolutionary change from a common ancestral protein sequence



Phylogenetic tree for cytochrome *c*

Polypeptide chain nomenclature

- **Amino acid "residues"** compose peptide chains
- Peptide chains are numbered from the N (amino) terminus to the C (carboxyl) terminus
- Example: (**N**) **Gly-Arg-Phe-Ala-Lys** (**C**)
  (or **GRFAK**)
- Formation of peptide bonds eliminates the ionizable α-carboxyl and α-amino groups of the free amino acids