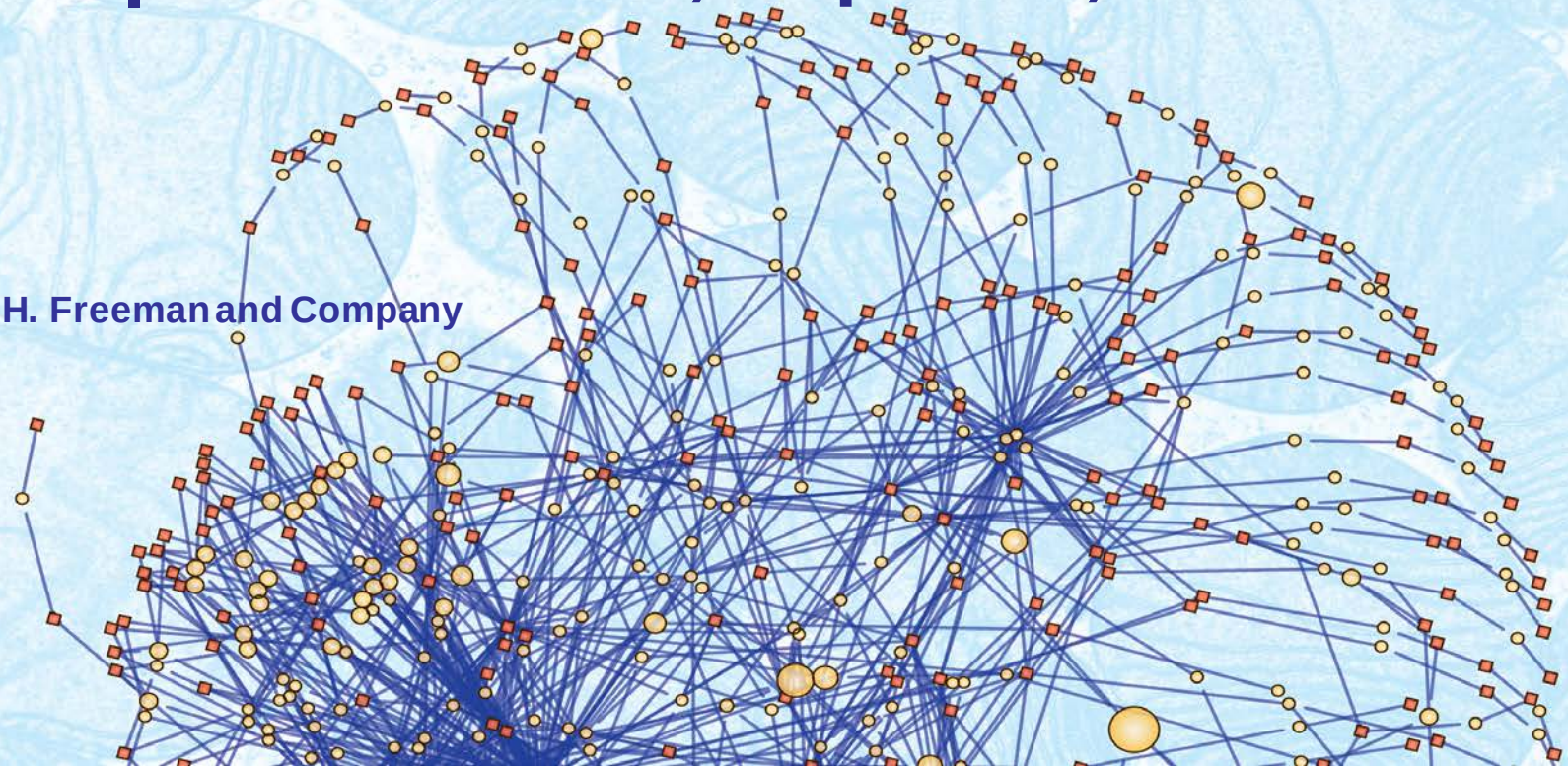
Lehninger

SIXTH EDITION

Principles of Biochemistry

David L. Nelson | Michael M. Cox

3 | Amino Acids, Peptides, Proteins



Natural selection favors some mutations

- Mutations occur more or less randomly
  - Balance between too many (below sustainability threshold) and too few (generate enough genetic variations)
- Mutations that give organisms an advantage in a given environment are more likely to be propagated
  - Natural selection (survival of the fittest)

From last lecture

A most recent example in literature

RESEARCH ARTICLE | BIOCHEMISTRY

RNA-catalyzed evolution of catalytic RNA

Nikolaos Papastavrou, David P. Horning, and Gerald F. Joyce Authors Info & Affiliations

Edited by Jack Szostak, University of Chicago Department of Chemistry, Chicago, IL; received December 7, 2023; accepted January 25, 2024

March 4, 2024 | 121 (11) e2321592121 | https://doi.org/10.1073/pnas.2321592121

- RNA acts as an enzyme (ribozyme)
  - It has RNA polymerase activity
  - Its fidelity is not 100%
  - It enables evolution of hammerhead ribozyme

Relevant example in literature

An RNA polymerase ribozyme that was obtained by directed evolution can propagate a functional RNA through repeated rounds of replication and selection, thereby enabling Darwinian evolution. Earlier versions of the polymerase did not have sufficient copying fidelity to propagate functional information, but a new variant with improved fidelity can replicate the hammerhead ribozyme through reciprocal synthesis of both the hammerhead and its complement, with the products then being selected for RNA-cleavage activity. Two evolutionary lineages were carried out in parallel, using either the prior low-fidelity or the newer high-fidelity polymerase. The former lineage quickly lost hammerhead functionality as the population diverged toward random sequences, whereas the latter evolved new hammerhead variants with improved fitness compared to the starting RNA. The increase in fitness was attributable to specific mutations that improved the replicability of the hammerhead, counterbalanced by a small decrease in hammerhead activity. Deep sequencing analysis was used to follow the course of evolution, revealing the emergence of a succession of variants that progressively diverged from the starting hammerhead as fitness increased. This study demonstrates the critical importance of replication fidelity for maintaining heritable information in an RNA-based evolving system, such as is thought to have existed during the early history of life on Earth. Attempts to recreate RNA-based life in the laboratory must achieve further improvements in replication fidelity to enable the fully autonomous Darwinian evolution of RNA enzymes as complex as the polymerase itself.

Macromolecules Are Major Constituents

TABLE 1–1 Molecular Components of an *E. coli* Cell

	Percentage of total weight of cell	Approximate number of different molecular species
Water	70	1
Proteins	15	3,000
Nucleic acids		
DNA	1	1–4
RNA	6	>3,000
Polysaccharides	3	10
Lipids	2	20
Monomeric subunits and intermediates	2	500
Inorganic ions	1	20

From last lecture

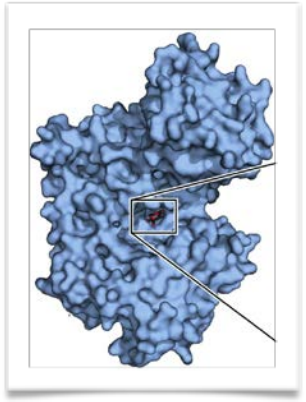
CHAPTER 3
Amino Acids, Peptides, Proteins

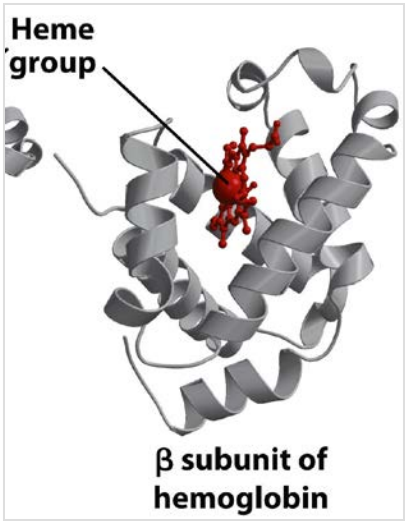
Learning goals:

- Structure and naming of **amino acids**
- Structure and properties of **peptides**
- **Ionization** behavior of amino acids and peptides
- Methods to **characterize** peptides and proteins

Proteins:
Main Agents of Biological Function

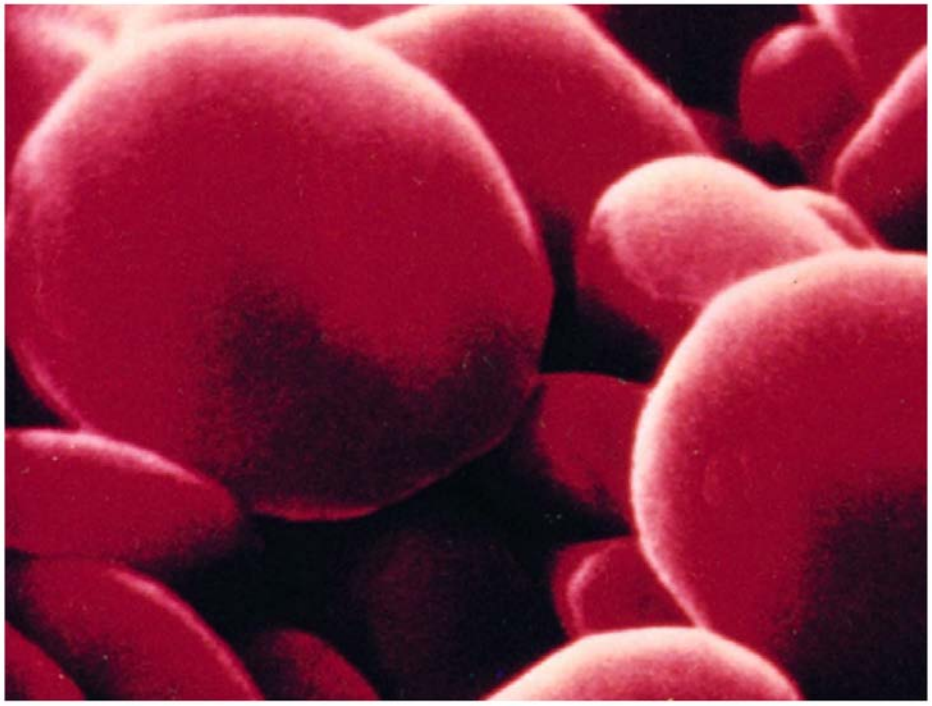
- Catalysis
  - **Hexokinase** (in the glycolytic pathway)
  - DNA polymerase (in DNA replication)
- Transport
  - **Hemoglobin** (transports O_2 in the blood)
  - Lactose permease (transports lactose across the cell membrane)
- Structure
  - Collagen (connective tissue)
  - Keratin (hair, nails, feathers, horns)
- Motion
  - Myosin (muscle tissue)
  - Actin (muscle tissue, cell motility)





生物发光：萤火虫，荧光素酶，需要氧气和ATP

Proteins Serve A Wide Range of Biological Functions



- Red blood cells contain lots of O_2-transporting protein hemoglobin
- Keratin protein is chief component of hair, scales, wool, feather and horn

发鳞毛甲羽

Amino Acids, Peptides, and Proteins

3.1 Amino Acids

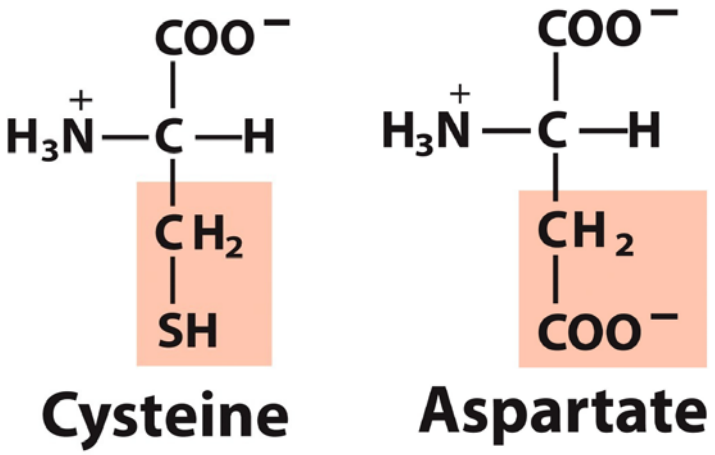
3.2 Peptides and Proteins

3.3 Working with Proteins

3.4 Protein Primary Structure

Amino Acids: Building Blocks of Protein

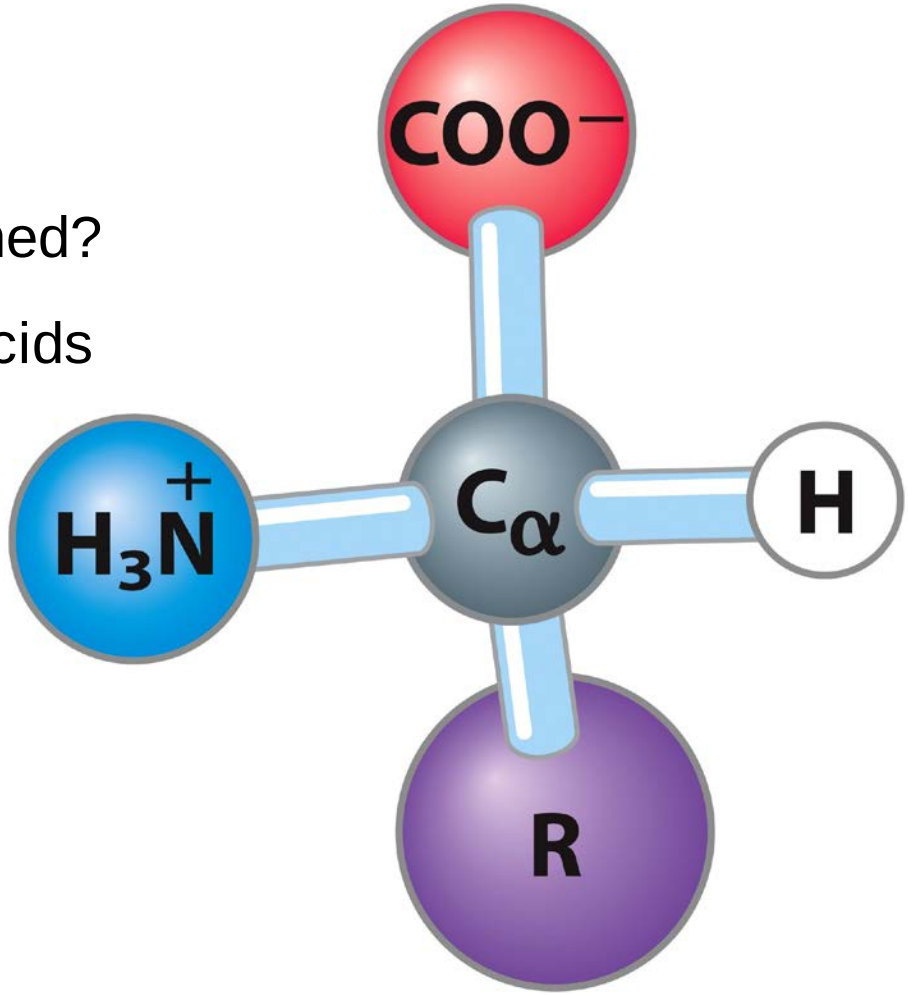
- Proteins are **linear heteropolymers** of α-amino acids
- Amino acids have properties that are well-suited to carry out a variety of biological functions
  - Capacity to polymerize
  - Useful acid-base properties
  - Varied physical properties
  - Varied chemical functionality



Amino Acids Differ Only at R substituent

α-amino acid

- What is α-carbon?
- Where is amino group attached?
- Difference between amino acids
- Other types of amino acids?



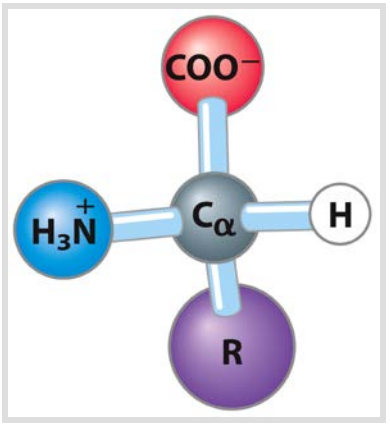
Three-Letter and One-Letter Code

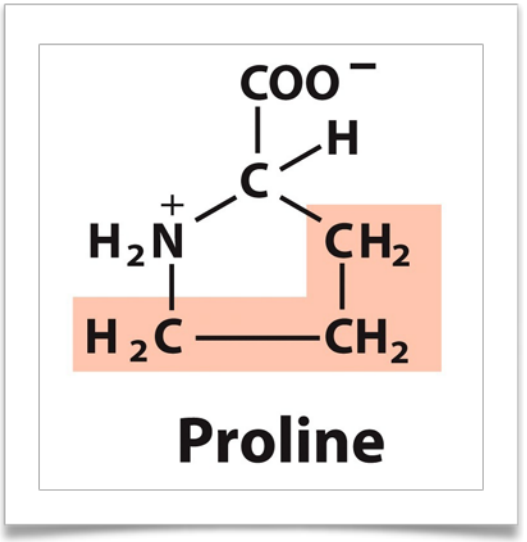
Full Name	*Three-Letter*	*One-Letter*
Cysteine	Cys	C
Histidine	His	H
Isoleucine	Ile	I
Methionine	Met	M
Serine	Ser	S
Valine	Val	V
Alanine	Ala	A
Glycine	Gly	G
Leucine	Leu	L
Proline	Pro	P
Threonine	Thr	T

Full Name	*Three-Letter*	*One-Letter*
Arginine	Arg	R
Phenylalanine	Phe	F
Tyrosine	Tyr	Y
Tryptophan	Trp	W
Aspartate	Asp	D
Asparagine	Asn	N
Glutamate	Glu	E
Glutamine	Gln	Q
Lysine	Lys	K

Most α-Amino Acids Are Chiral

- The α-carbon always has four substituents and is tetrahedral
- All (except proline) have:
  - an acidic carboxyl group
  - a basic amino group
  - an α-hydrogen connected to the α-carbon
- The fourth substituent (R) is unique
  - In glycine, the fourth substituent is hydrogen
  - In alanine, the fourth substituent is methyl group

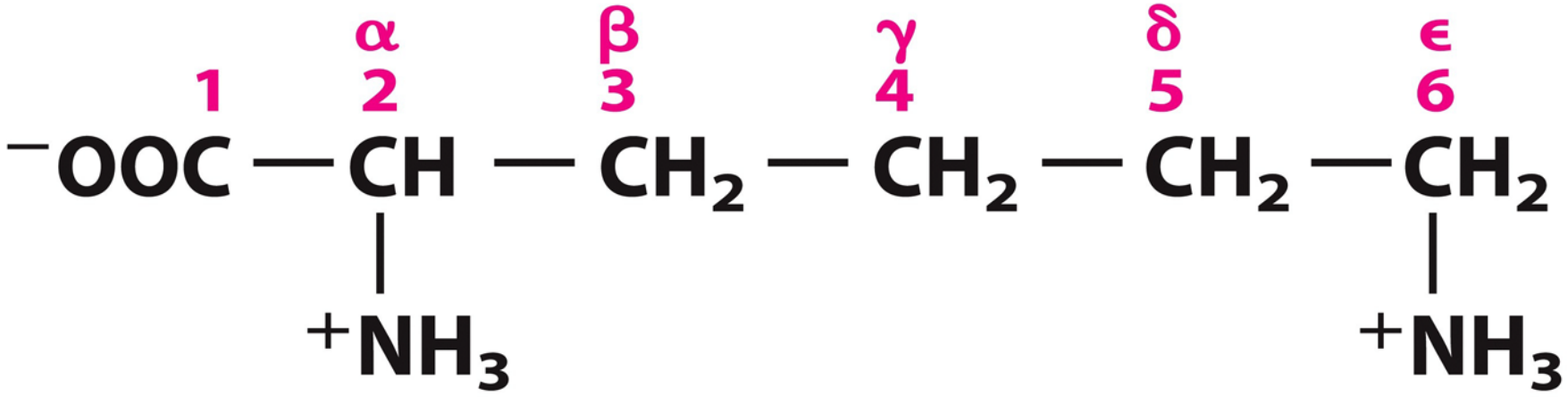




手性的重要性：阿斯巴甜与苦味异构体，生物分子相互作用的立体选择性

Amino Acids Atom Naming

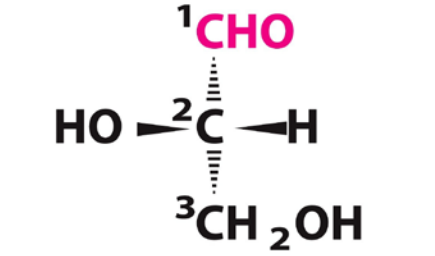
- Organic nomenclature: start from one end
- Biochemical designation:
  - Start from α-carbon and go down the R-group



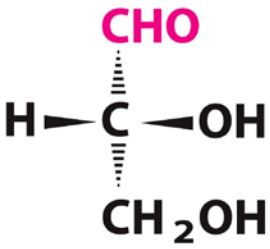
Lysine

All Amino Acids Are Chiral Except Gly

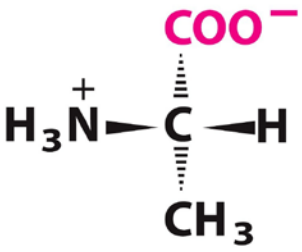
- DL system
  - Compare with glyceraldehyde
  - Absolute configuration only
  - Not optical properties
- RS system



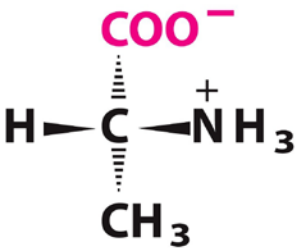
L-Glyceraldehyde



D-Glyceraldehyde

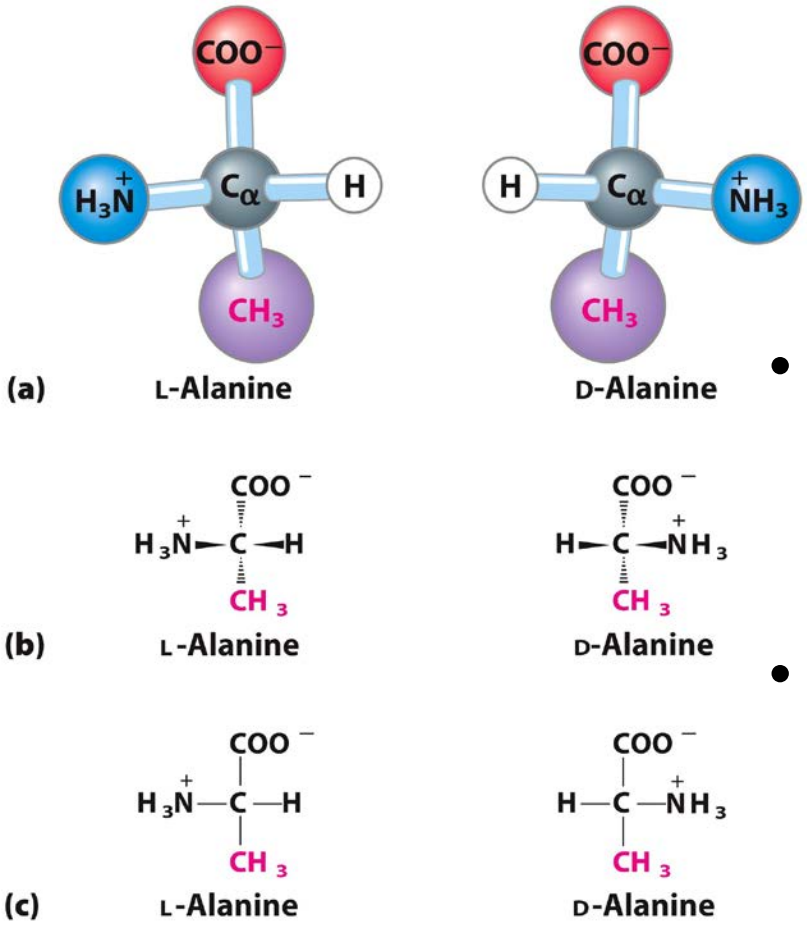


L-Alanine



D-Alanine

Amino Acids in Proteins Are L-Stereoisomers



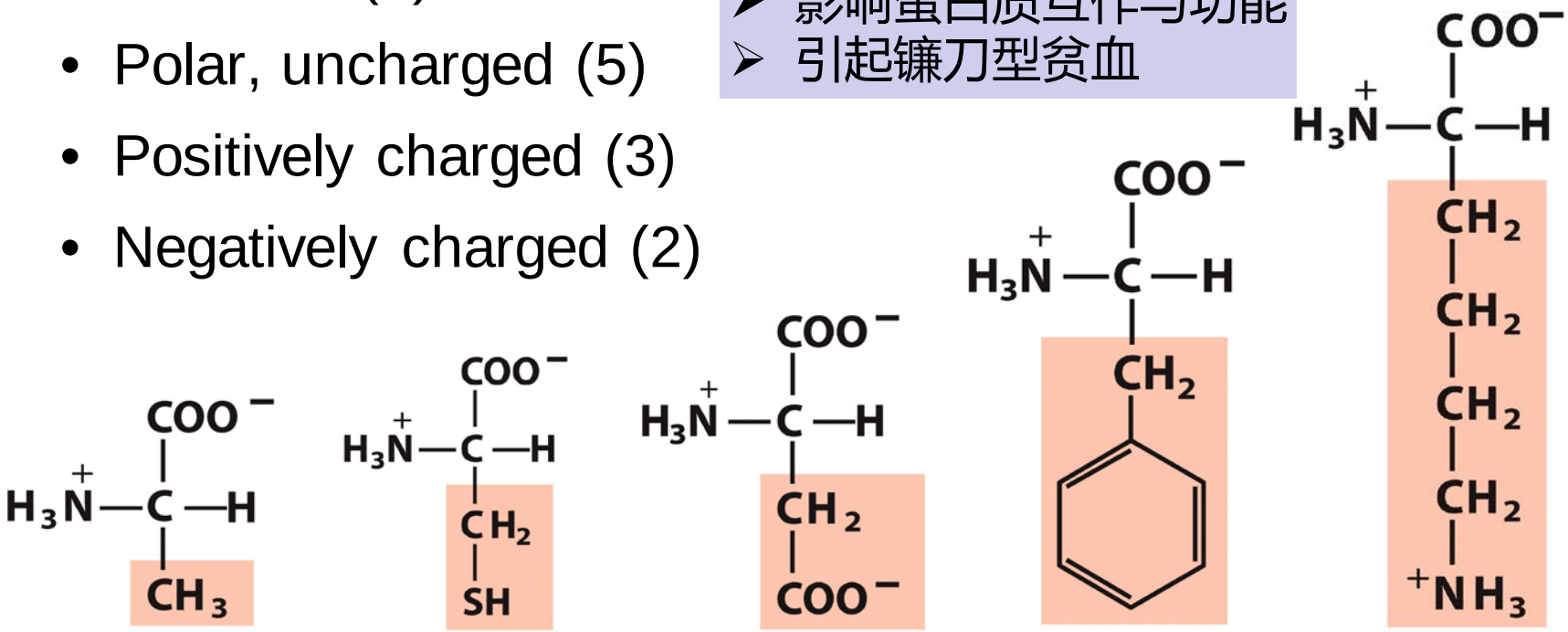
- Perspective formula
  - Wedge-shaped bond: towards viewer
  - Dashed bond: behind paper
- Projection formula
  - Horizontal bond: towards viewer
  - Vertical bond: behind paper
  - Not always specify stereochemistry

Amino Acids Classification

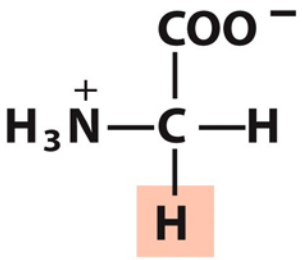
Common amino acids can be placed in five basic groups depending on their R substituents:

- Nonpolar, aliphatic (7)
- Aromatic (3)
- Polar, uncharged (5)
- Positively charged (3)
- Negatively charged (2)

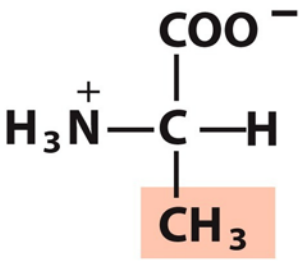
一个氨基酸突变：
- 谷氨酸变成缬氨酸
- 影响蛋白质互作与功能
- 引起镰刀型贫血



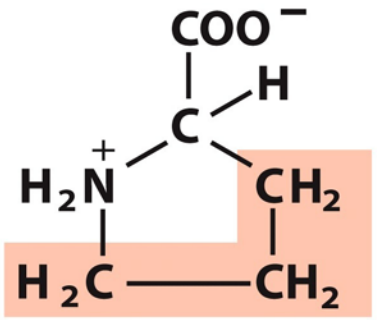
Nonpolar, aliphatic R groups



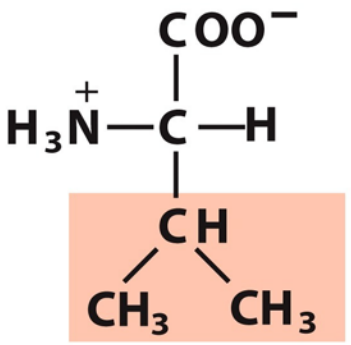
Glycine



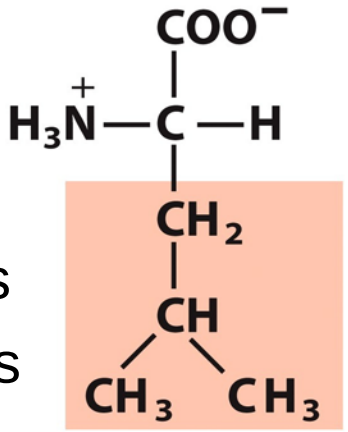
Alanine



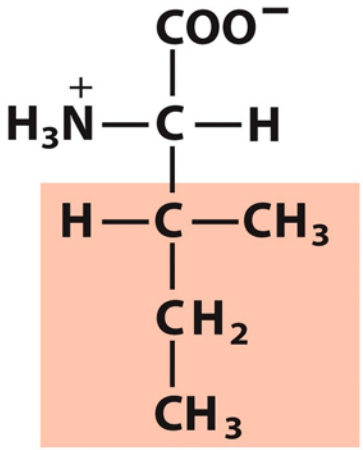
Proline



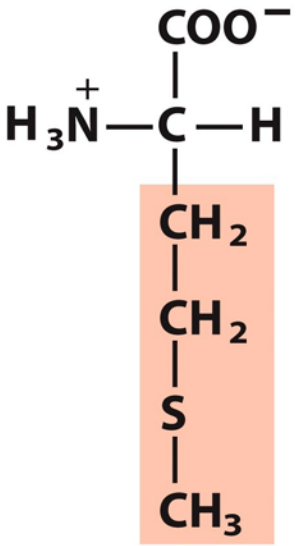
Valine



Leucine



Isoleucine

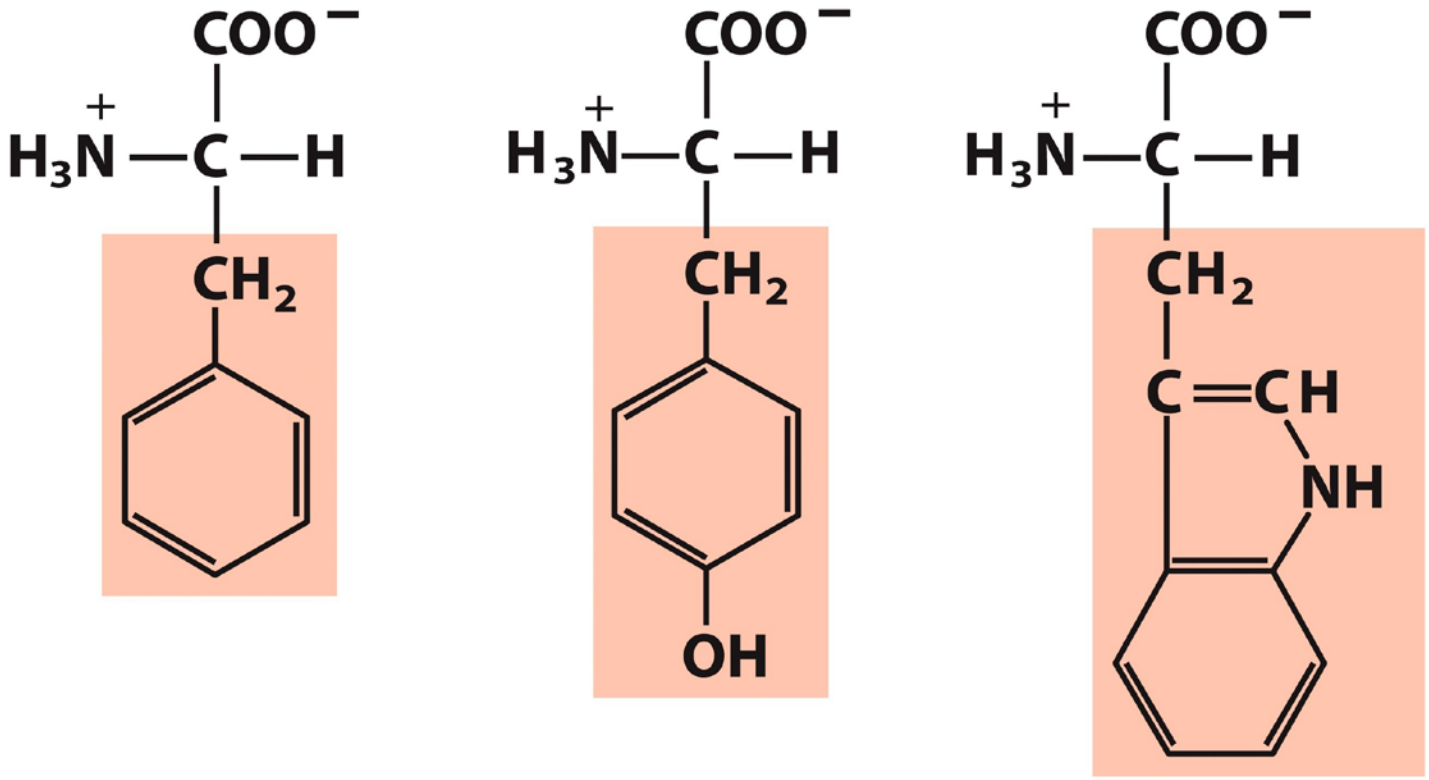


Methionine

<u>Hydrocarbons</u>

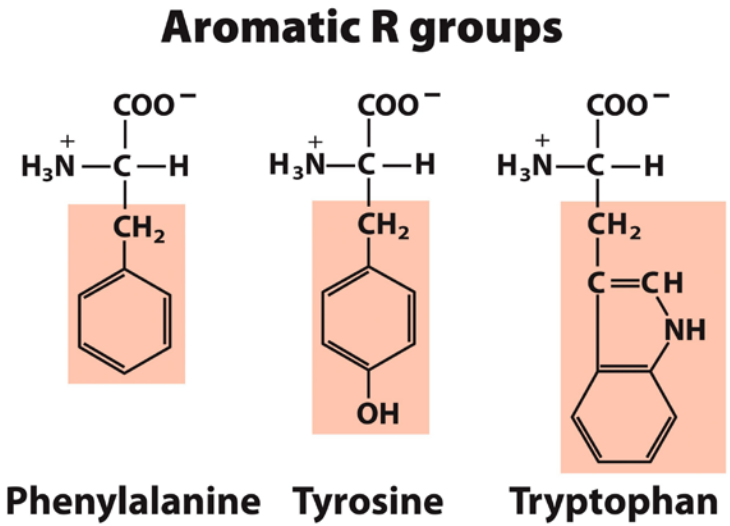
- aliphatic compounds
- aromatic compounds

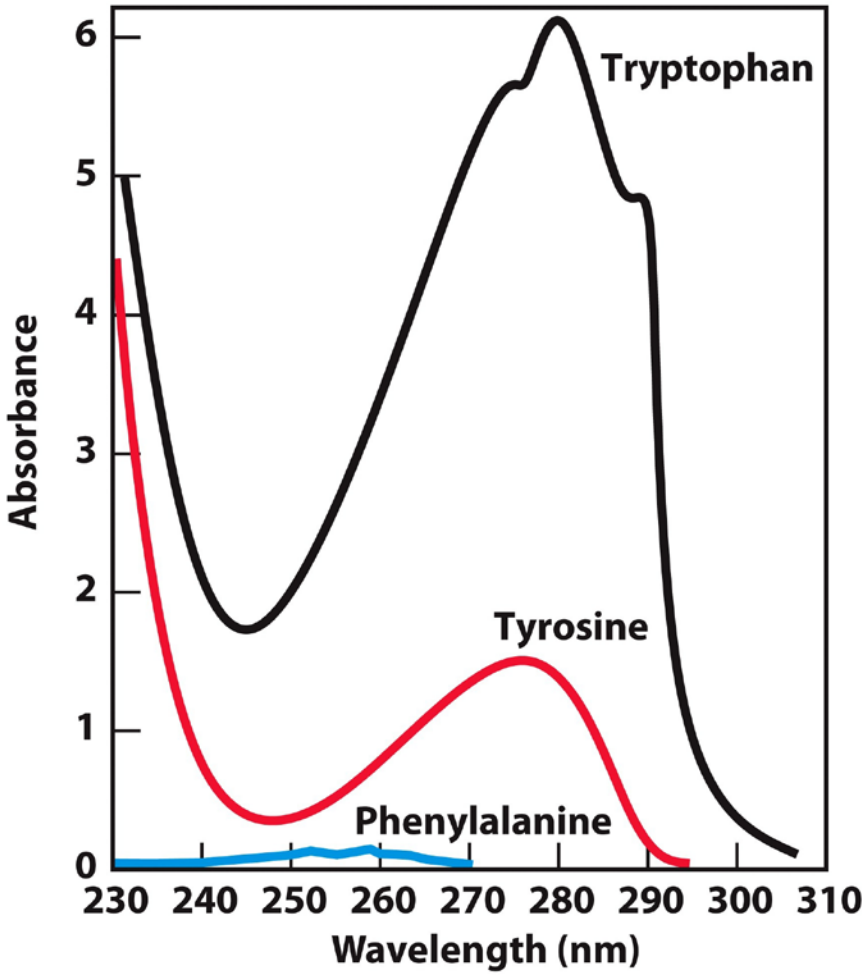
Aromatic R groups



Phenylalanine Tyrosine Tryptophan

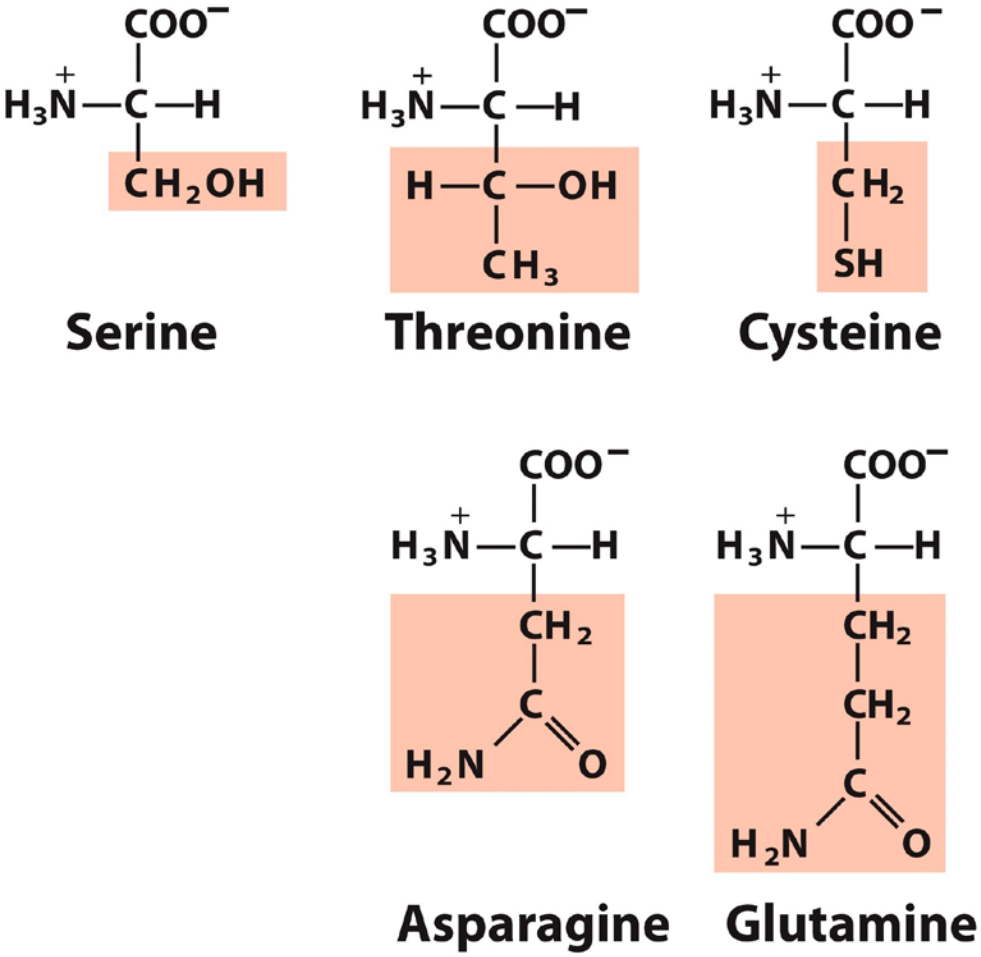
Absorption of UV Light by FYW





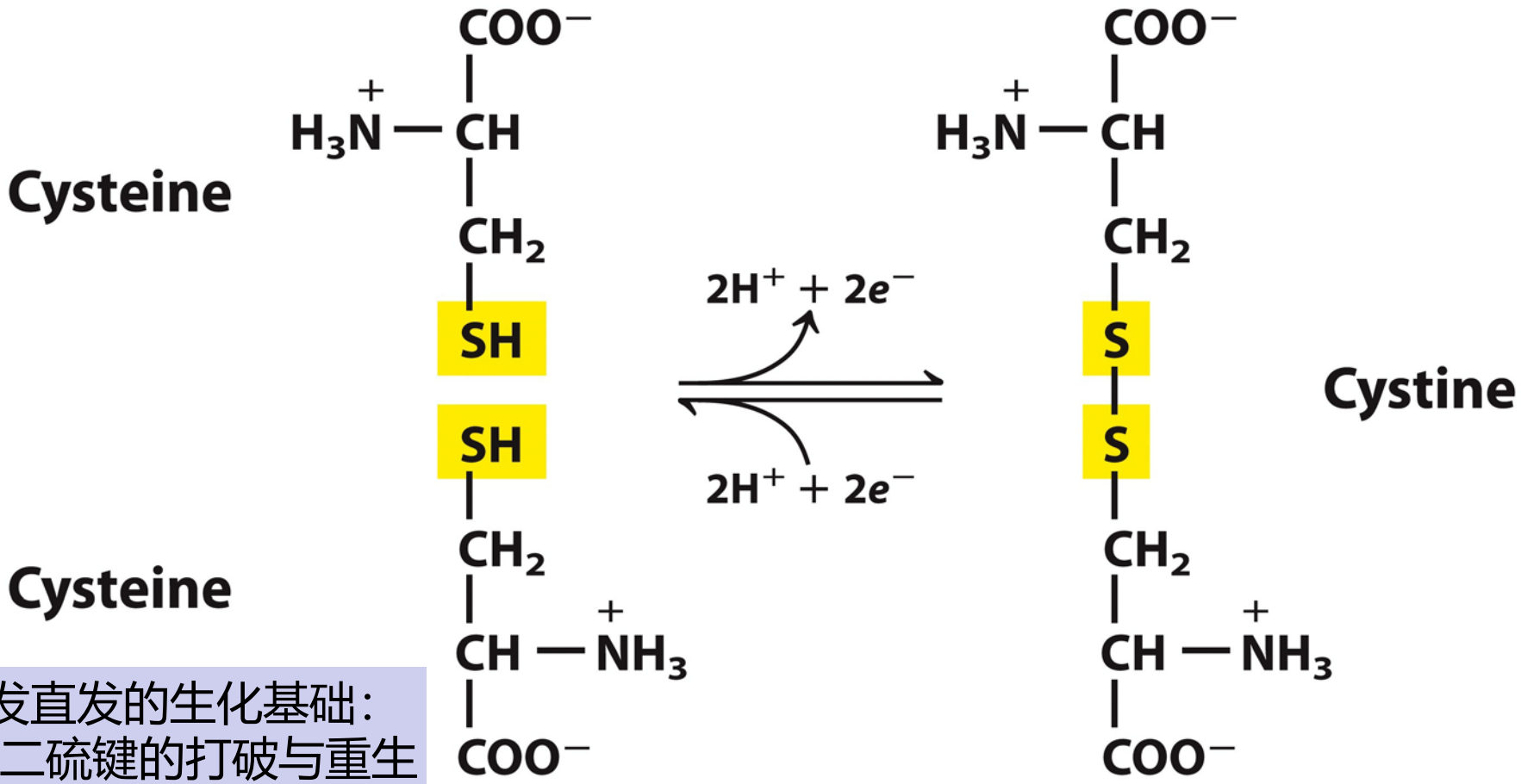
These amino acid side chains absorb UV light at 270 - 280 nm

Polar, uncharged R groups



These amino acids side chains can form hydrogen bonds

Reversible Formation of S-S Bond

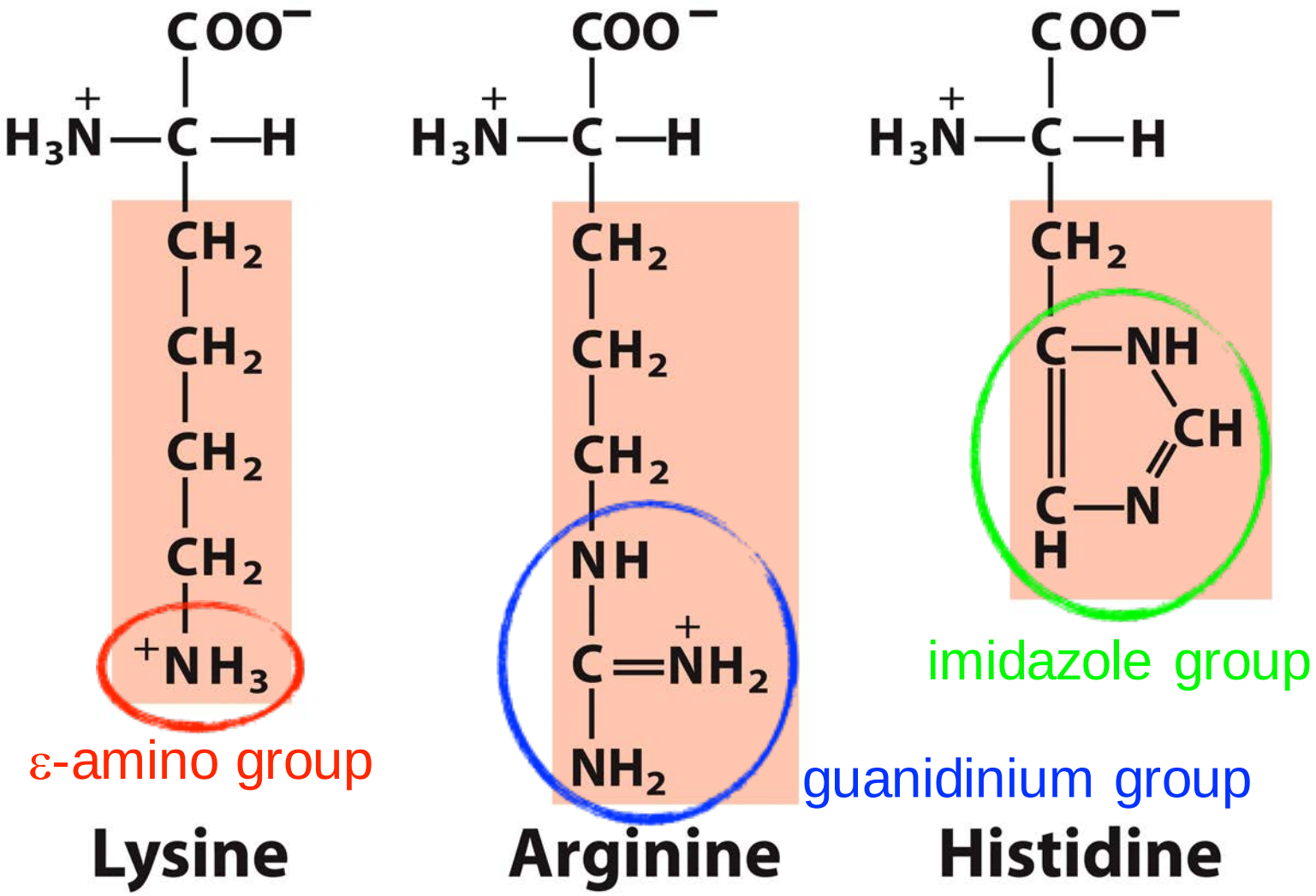


烫发直发的生化基础：

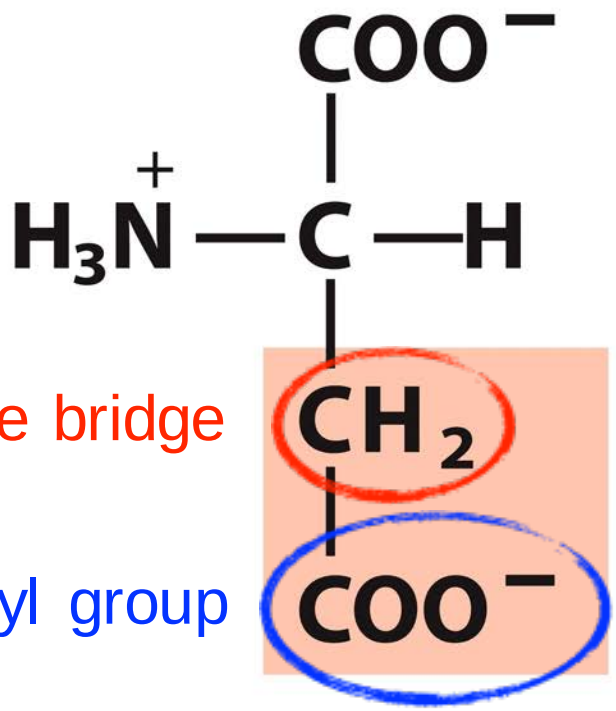
- 二硫键的打破与重生

Cysteine can form disulfide bonds

Positively charged R groups



Negatively charged R groups



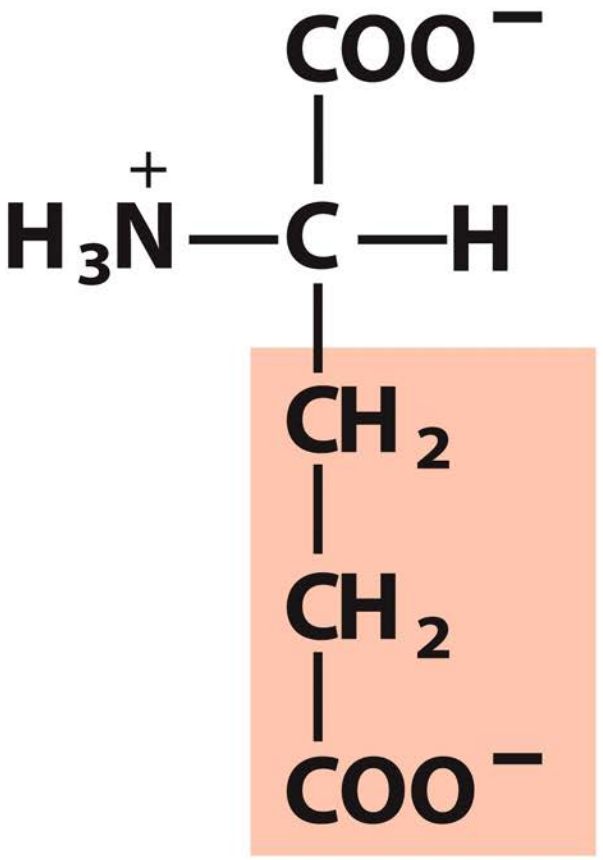
methylene bridge

carboxyl group

Amino acid for 2nd week

Aspartate

Asp, D

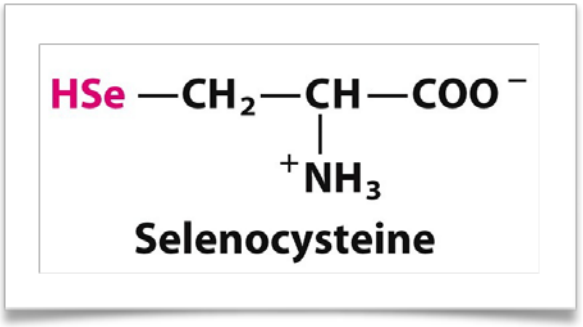


Glutamate

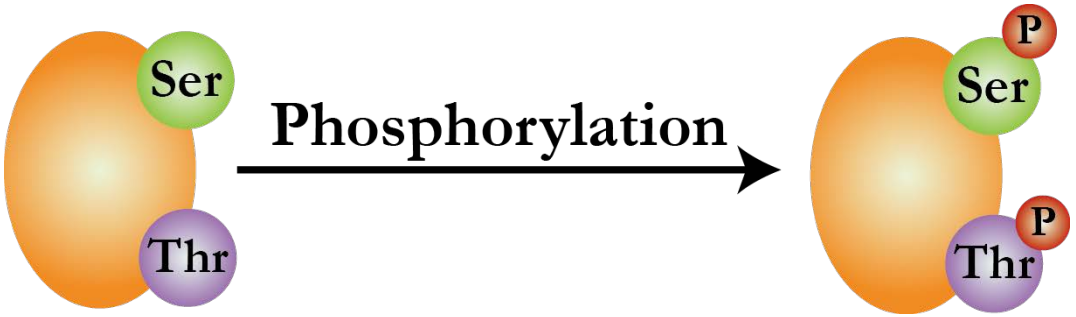
Glu, E

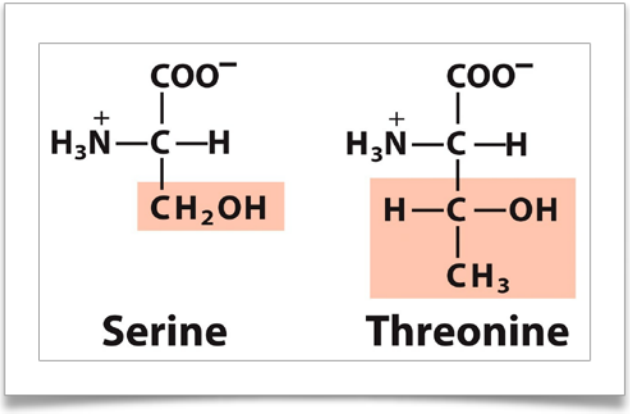
Uncommon Amino Acids in Proteins

- Not incorporated by ribosomes
  - Except for selenocysteine

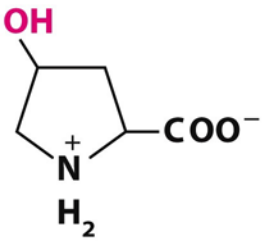


- Arise by post-translational modifications of proteins
- Reversible modifications, especially phosphorylation, are important in regulation and signaling

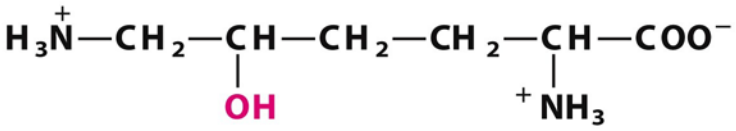




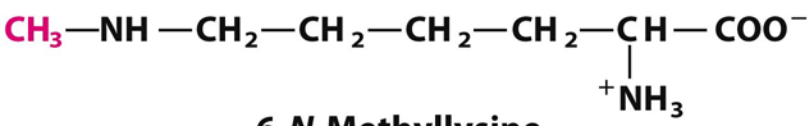
Modified Amino Acids Found in Proteins



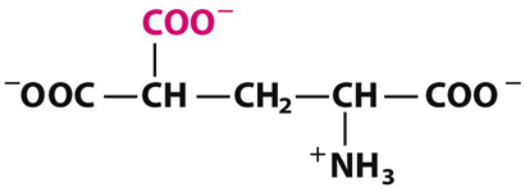
4-Hydroxyproline



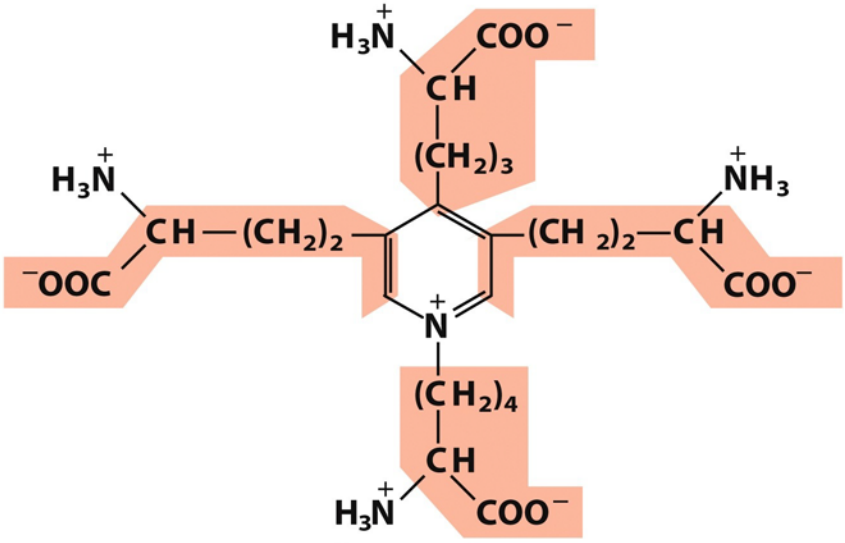
5-Hydroxylysine



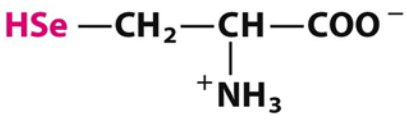
6-*N*-Methyllysine



γ-Carboxyglutamate

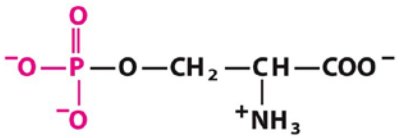


Desmosine

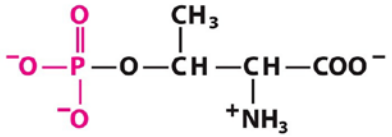


Selenocysteine

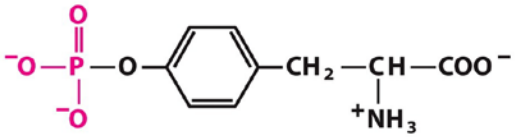
Reversible Modifications of Amino Acids



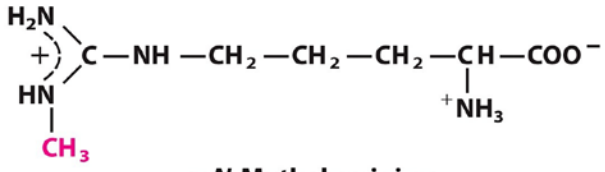
Phosphoserine



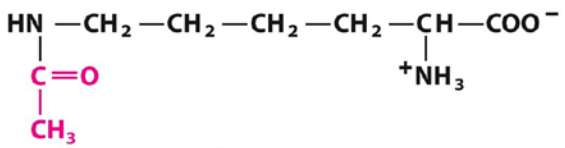
Phosphothreonine



Phosphotyrosine

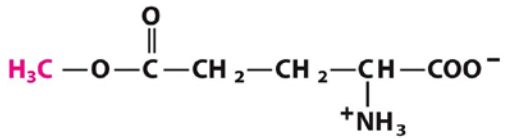


σ-*N*-Methylarginine

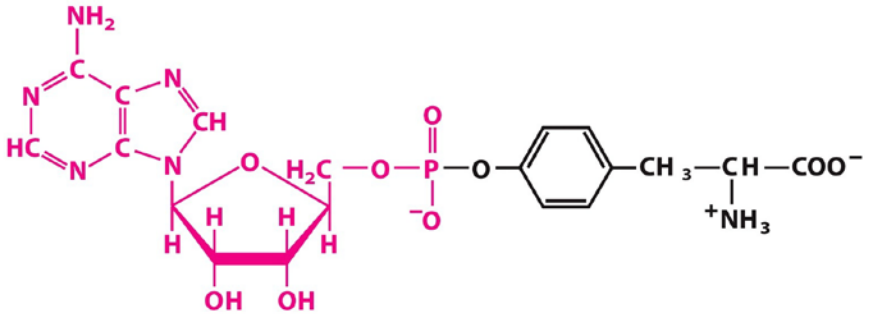


6-*N*-Acetyllysine

Transient modification



Glutamate γ-methyl ester

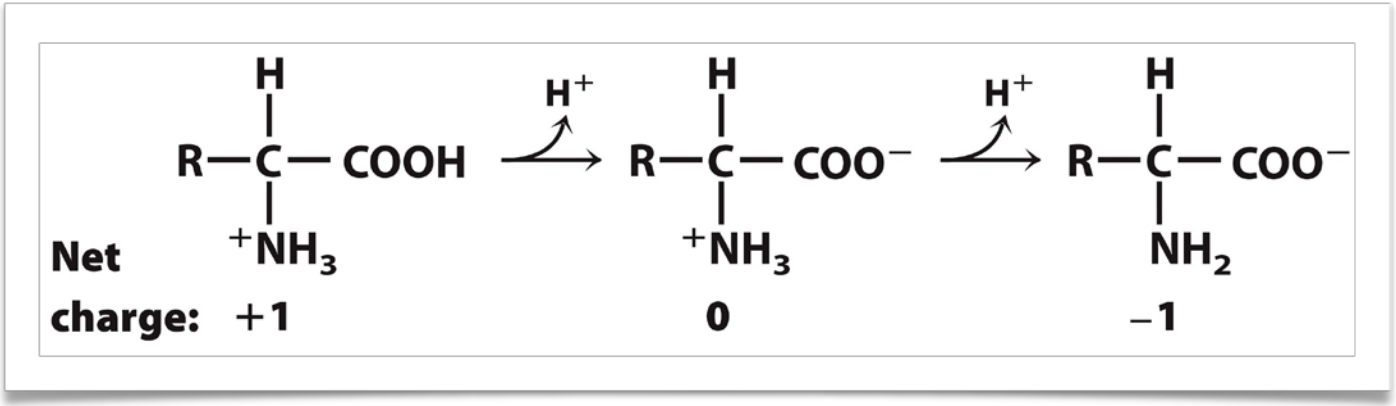


Adenylyltyrosine

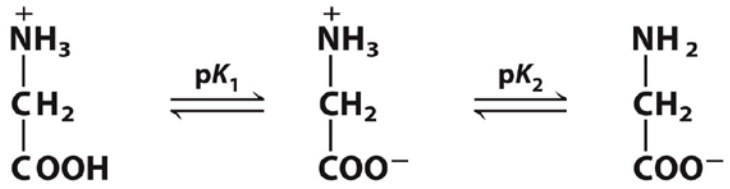
- Phosphorylation
- Methylation
- Acetylation
- …

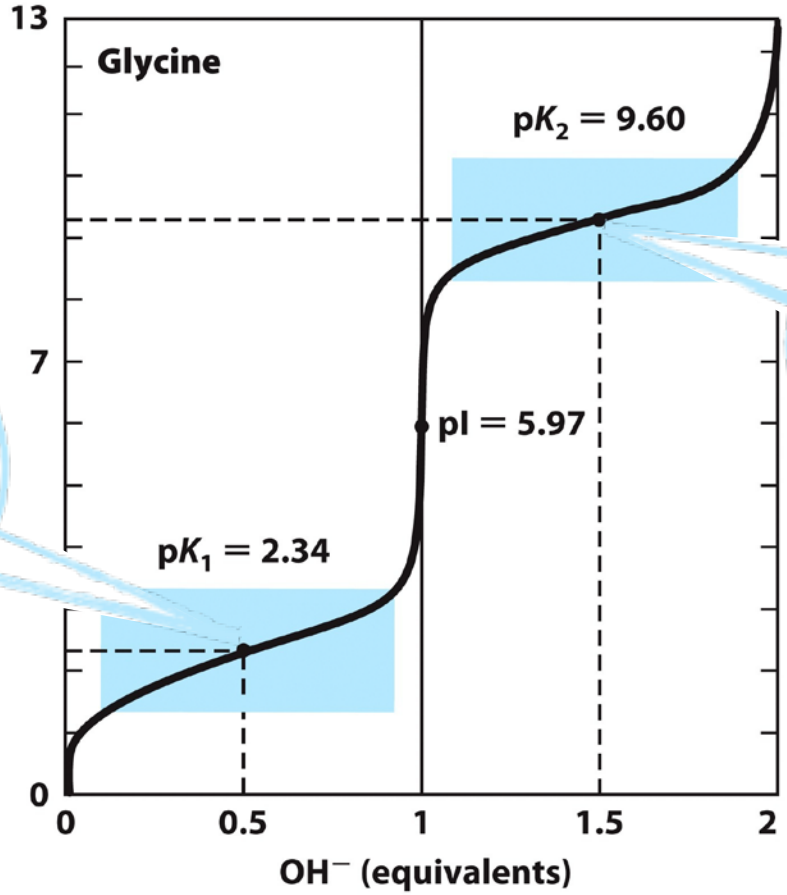
Ionization of Amino Acids

- At acidic pH, both carboxyl and amino groups are protonated and amino acid is in cationic form
- At neutral pH, carboxyl group is deprotonated but amino group is still protonated. The net charge is zero; such ions are called **Zwitterions**
- At alkaline pH, amino group is neutral -NH_2 and amino acid is in the anionic form



Cation **Zwitterion** **Anion**



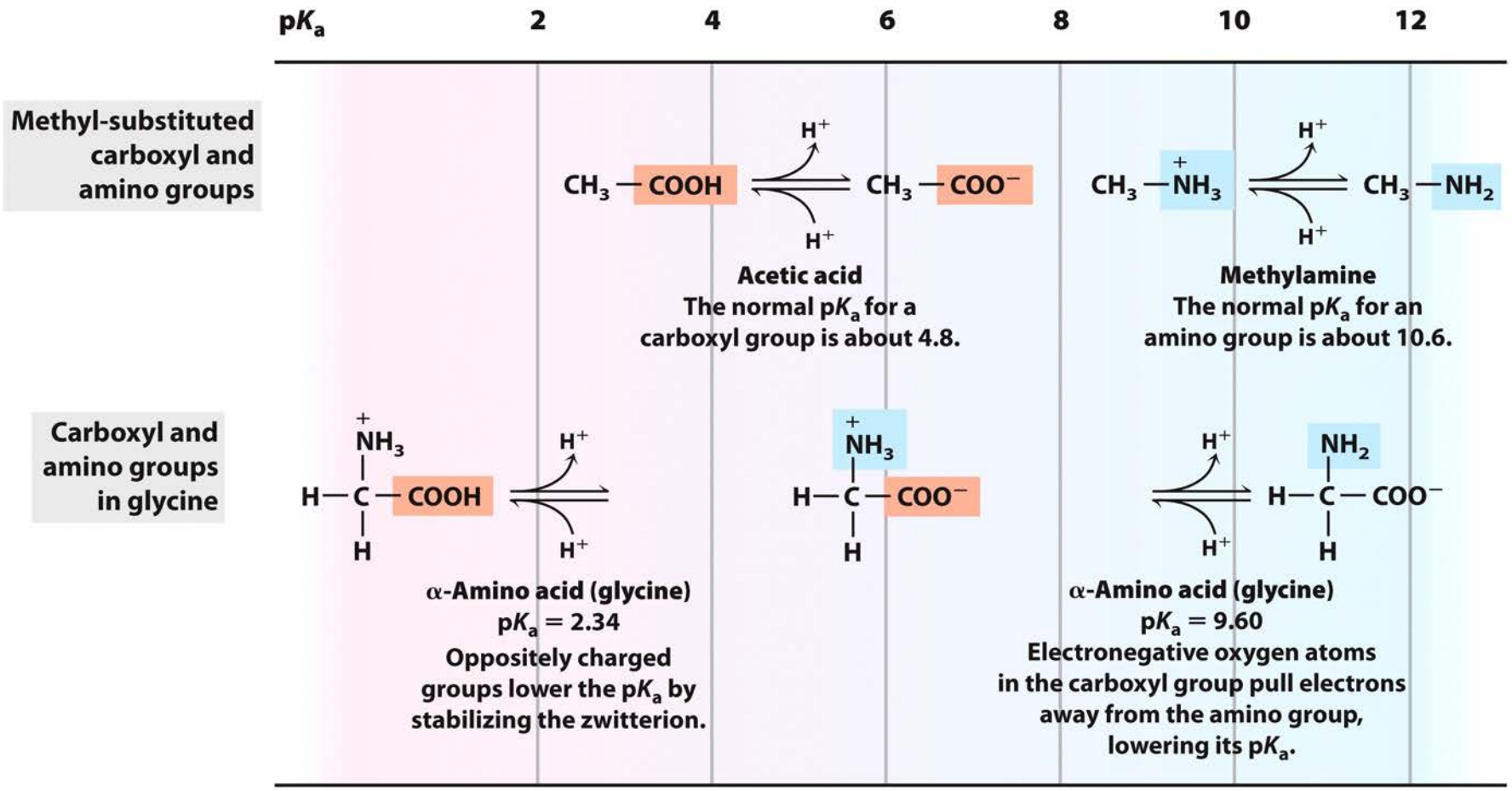


50% 50%

50% 50%

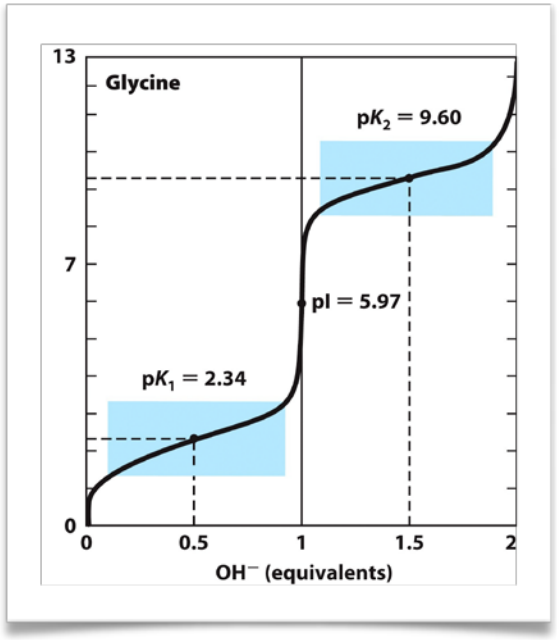
Chemical Environment Affects pK_a Values

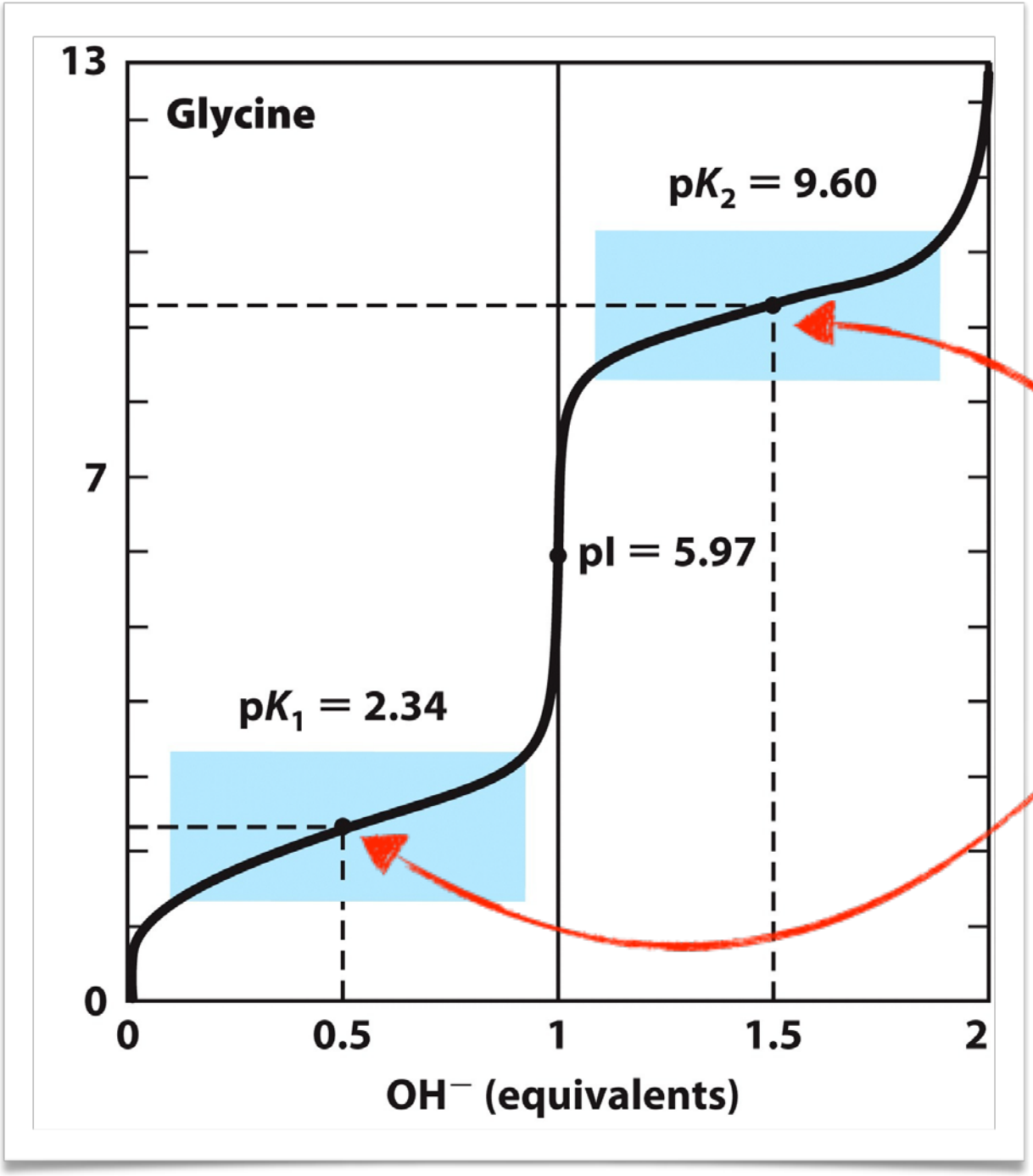
- α-carboxyl group is much more acidic than in carboxylic acids.
- α-amino group is slightly less basic than in amines.



Amino Acids Can Act as Buffers

- Amino acids with uncharged side chains, such as glycine, have two pK_a values:
  - The pK_a of the α-carboxyl group is 2.34
  - The pK_a of the α-amino group is 9.6
- It can act as a buffer in two pH regions
  - Not a good buffer near pH 6

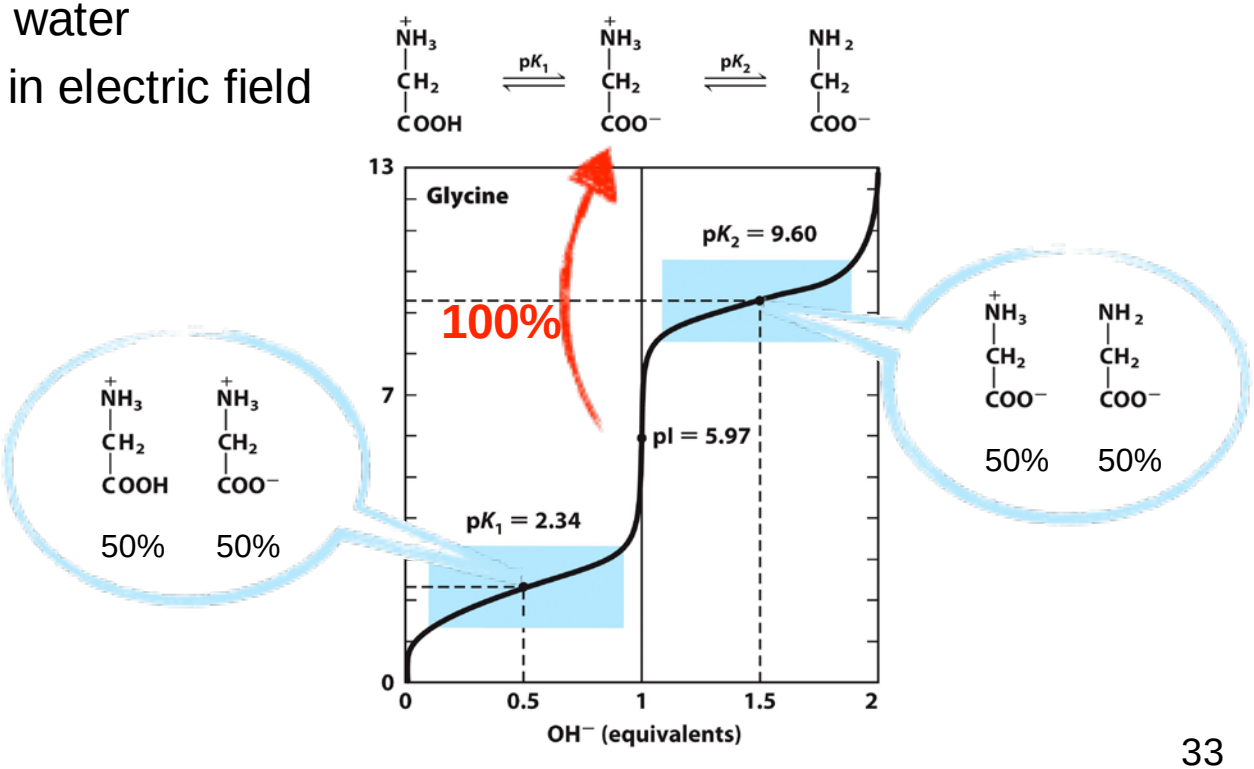



13
Glycine
$pK_2 = 9.60$
7
pI = 5.97
$pK_1 = 2.34$
0
0
0.5
1
1.5
2
OH^- (equivalents)
Buffer Regions

Net Charge of Zero at A Specific pH (pI)

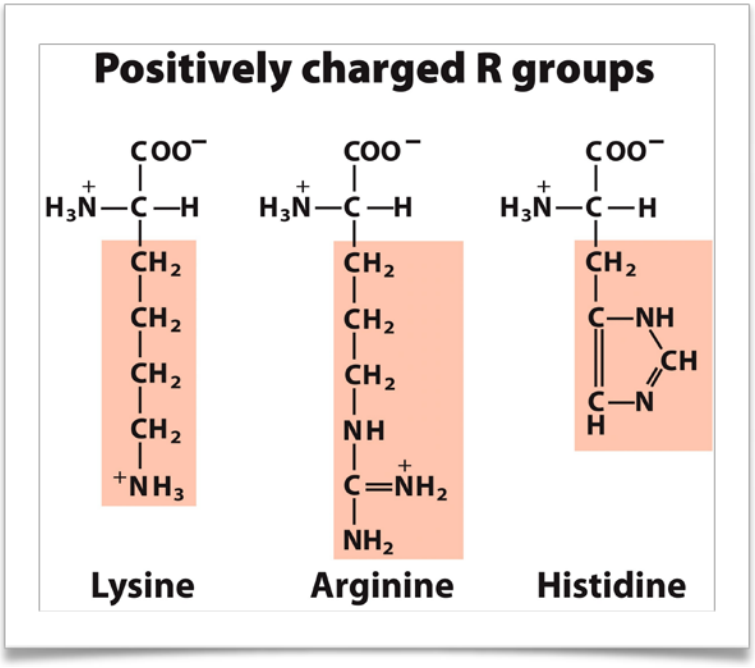
- Zwitterions predominate at pH values between the pK_a values of the amino and carboxyl groups
- For amino acids without ionizable side chains, the Isoelectric Point (equivalence point, pI) is the mean of two pK_a values
- At this point, the net charge is zero
  - AA is **least soluble** in water
  - AA does **not migrate** in electric field

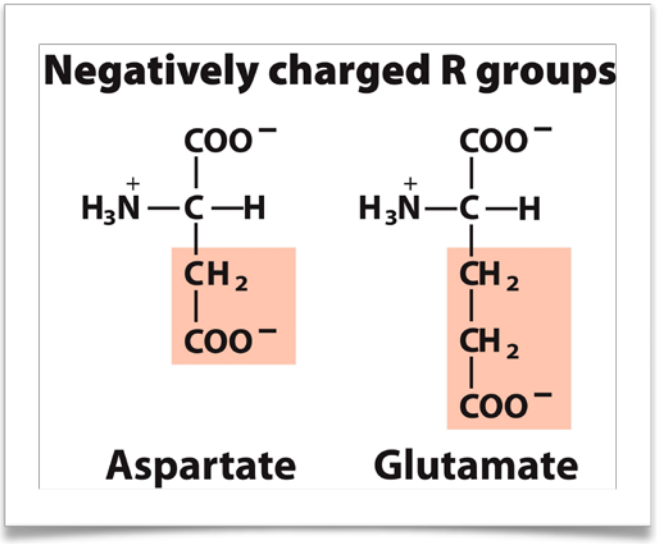
$$pI = \frac{pK_1 + pK_2}{2}$$

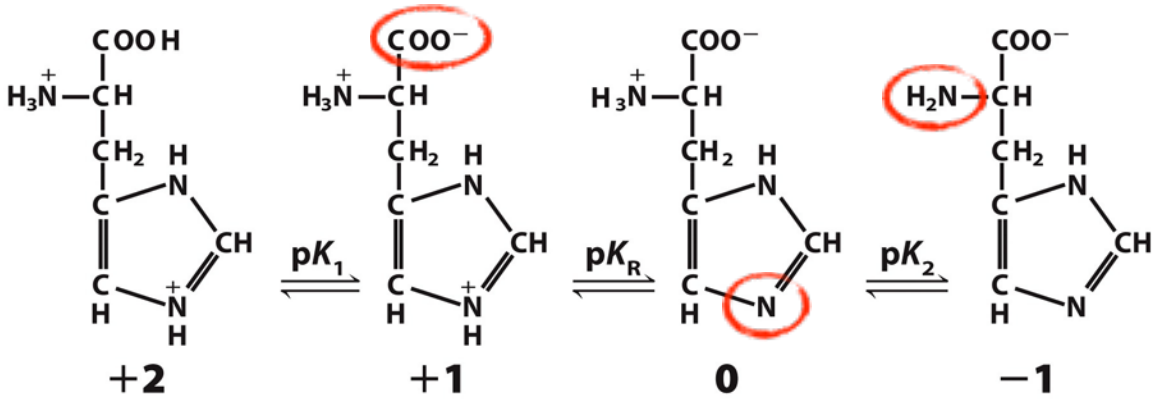


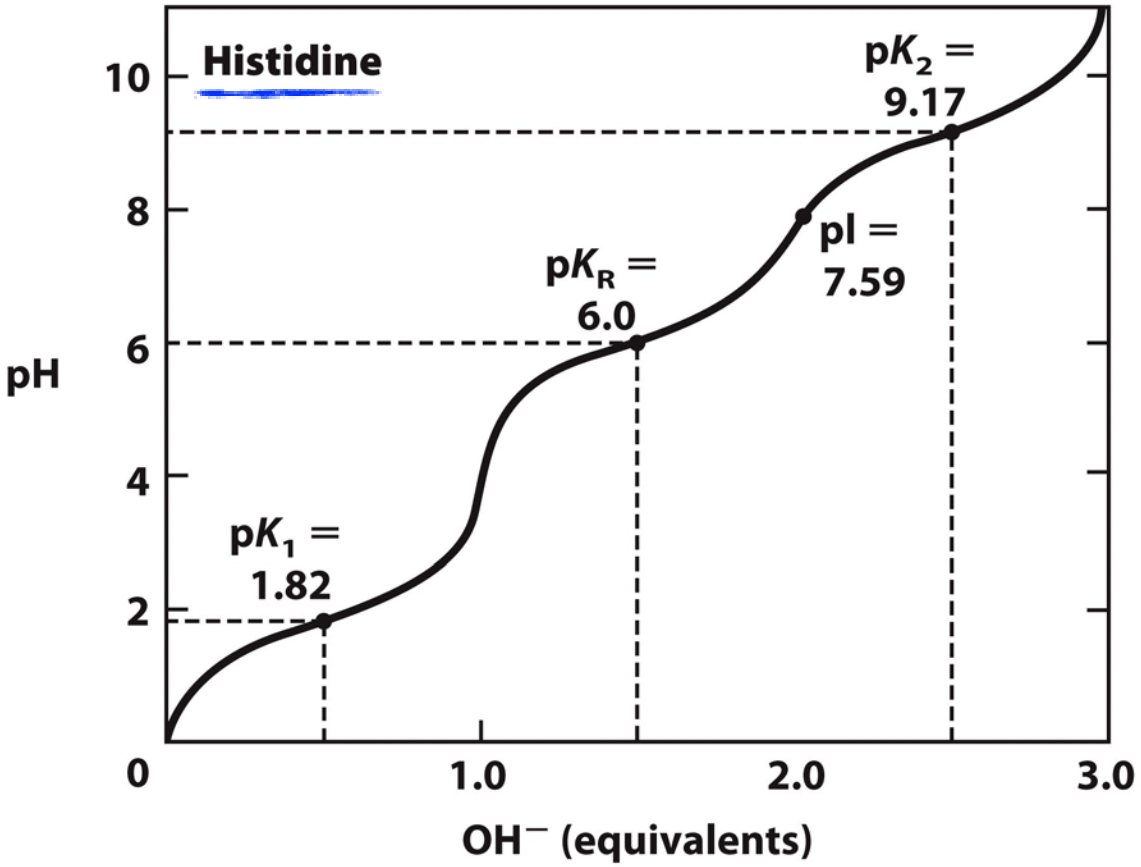
Ionizable Side Chains Show in Titration

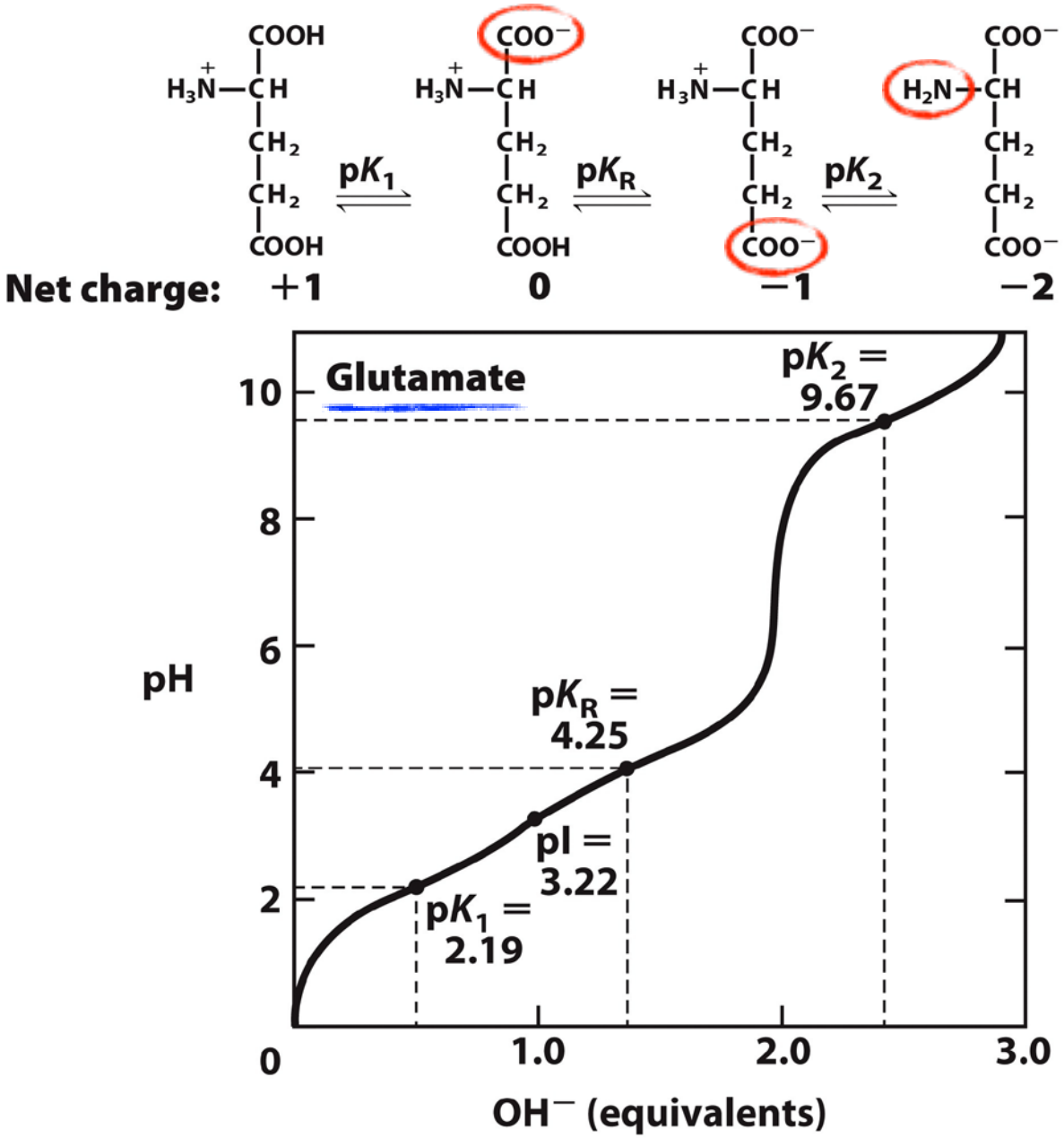
- Ionizable side chains can be also titrated
- Titration curves are now more complex
- pK_a values are discernable if two pK_a values are more than two pH units apart





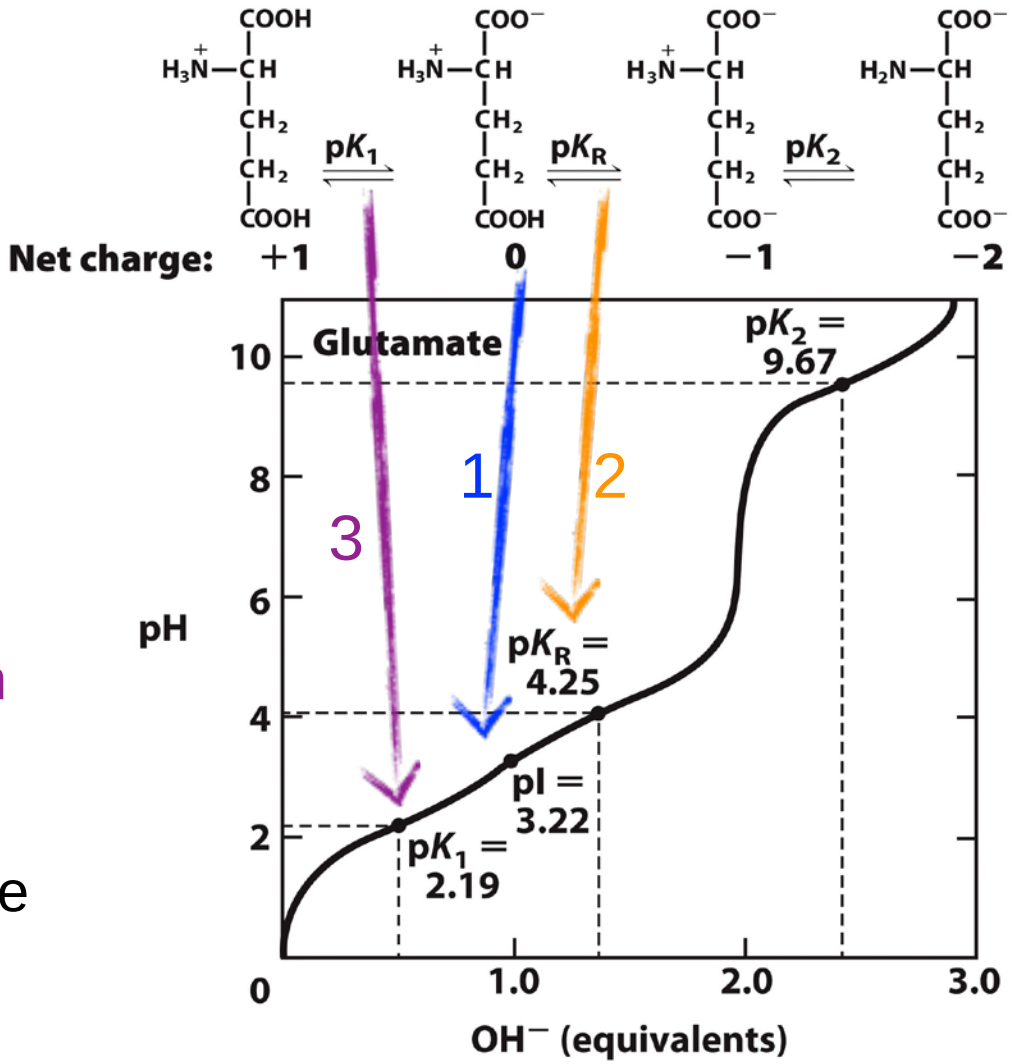

COOH
COO⁻
COO⁻
COO⁻
H₃N—CH
H₃N—CH
H₃N—CH
H₂N—CH
CH₂
H
N
CH
C
C
C
C
H
N
H
N
H
pK1
pKR
pK2
+2
+1
0
−1


Histidine
pK2 = 9.17
pI = 7.59
pKR = 6.0
pK1 = 1.82
pH
10
8
6
4
2
0
1.0
2.0
3.0
OH⁻ (equivalents)


Net charge: +1 0 −1 −2
pK1
pKR
pK2
Glutamate
pK2 = 9.67
pKR = 4.25
pI = 3.22
pK1 = 2.19
pH
OH− (equivalents)
0
2
4
6
8
10
1.0
2.0
3.0

How to Calculate the pI When the Side Chain is Ionizable

1. Identify species that carries a net zero charge
2. Identify pK_a value that defines the acid strength of this zwitterion
3. Identify pK_a value that defines the base strength of this zwitterion
4. Take the average of these two pK_a values



Summary 3.1 Amino Acids

- 20 amino acids commonly found in proteins contain an α-carboxyl group, an α-amino group, and a distinctive R group
- All amino acids except glycine have a chiral α-carbon center. Only the L stereoisomers are found in proteins
- Zwitterion and pI value

TABLE 3–1 Properties and Conventions Associated with the Common Amino Acids Found in Proteins

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

Example Question

The chirality of an amino acid results from the fact that its α carbon:

A) has no net charge.

B) is a carboxylic acid.

C) is bonded to four different chemical groups.

D) is in the L absolute configuration in naturally occurring proteins.

E) is symmetric.

Example Question

Two amino acids of the standard 20 contain sulfur atoms. They are:

A) cysteine and serine.

B) cysteine and threonine.

C) methionine and cysteine

D) methionine and serine.

E) threonine and serine.

Example Question

Which two amino acids differ from each other by only one atom?

A) Ser and Thr

B) Leu and Ile

C) Ala and Ser

D) Asp and Asn

E) Ser and Cys

Example Question

For amino acids with neutral R groups, at any pH below the pI of the amino acid, the population of amino acids in solution will have:

A) a net negative charge.

B) a net positive charge.

C) no charged groups.

D) no net charge.

E) positive and negative charges in equal concentration.

Example Question

At pH 7.0, converting a proline to hydroxyproline, will have what effect on the overall charge of the protein containing it?

A) It will become more negative

B) It will become more positive.

C) It will stay the same.

D) There is not enough information to answer the question.

E) The answer depends on the salt concentration.

Example Question

At pH 7.0, converting a glutamic acid to γ-carboxyglutamate, will have what effect on the overall charge of the protein containing it?

A) It will become more negative.

B) It will become more positive.

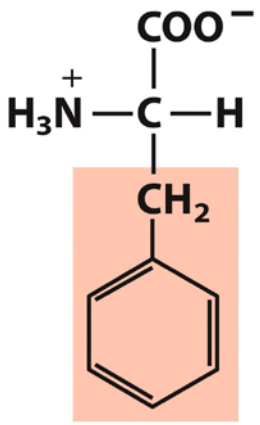
C) It will stay the same.

D) There is not enough information to answer the question.

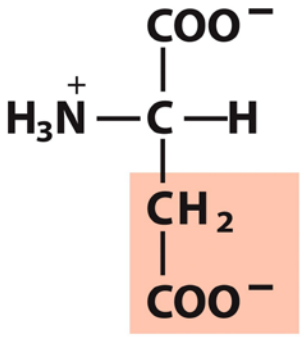
E) The answer depends on the salt concentration.

Example Question

Draw the structures of the amino acids phenylalanine and aspartate in the ionization state you would expect at pH 7.0. Why is aspartate very soluble in water, whereas phenylalanine is much less soluble?



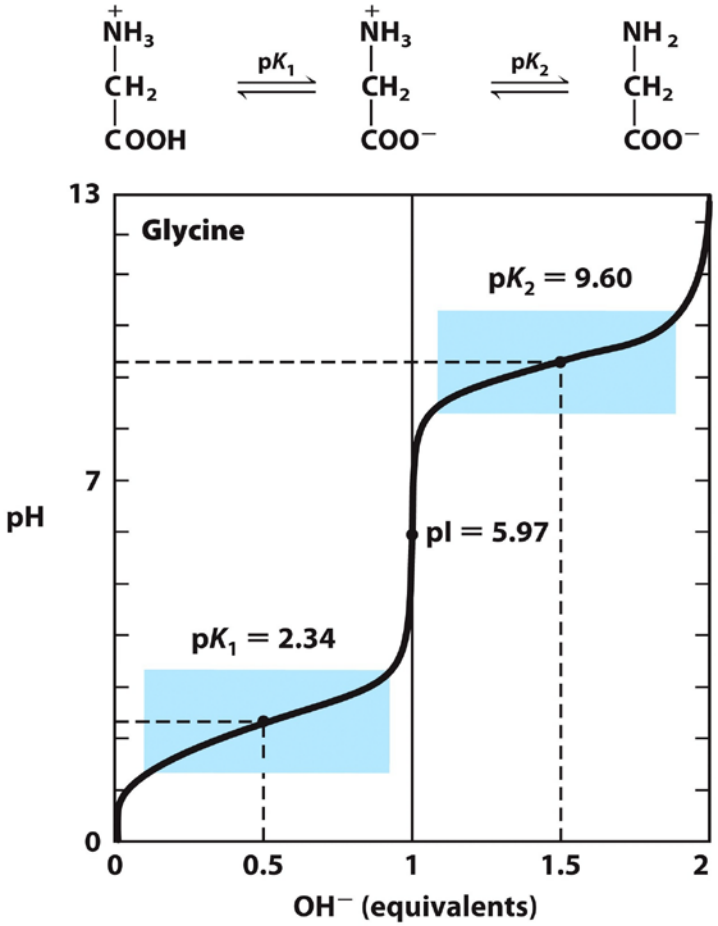
Phenylalanine



Aspartate

Example Question

As more OH^- equivalents (base) are added to a glycine solution, what titration reaction will occur around pH = 9.5?



Example Question

In the amino acid glycine, what effect does the positively charged -NH_3^+ group have on the pK_a of an amino acid's -COOH group?

1) The amino group **repels** the departing H^+ thereby promoting deprotonation.
2) The positively charged amino group **stabilizes** the negatively charged ionized form of the carboxyl group, -COO^-.
3) The effect is to **lower** the pK_a of the carboxyl group.

Example Question

1. **Draw the structure of Gly–Ala–Glu in the ionic form that predominates at pH 7.**

2. **The artificial sweetener aspartame is a simple dipeptide, aspartylphenylalanine methyl ester, on which the free carboxyl of the dipeptide is esterified to methyl alcohol. Draw the structure of aspartame, showing the ionizable groups in the form they have at pH 7. (The ionizable group in the side chain of aspartate has a pK_a of 3.96)**

Amino Acids, Peptides, and Proteins

3.1 Amino Acids

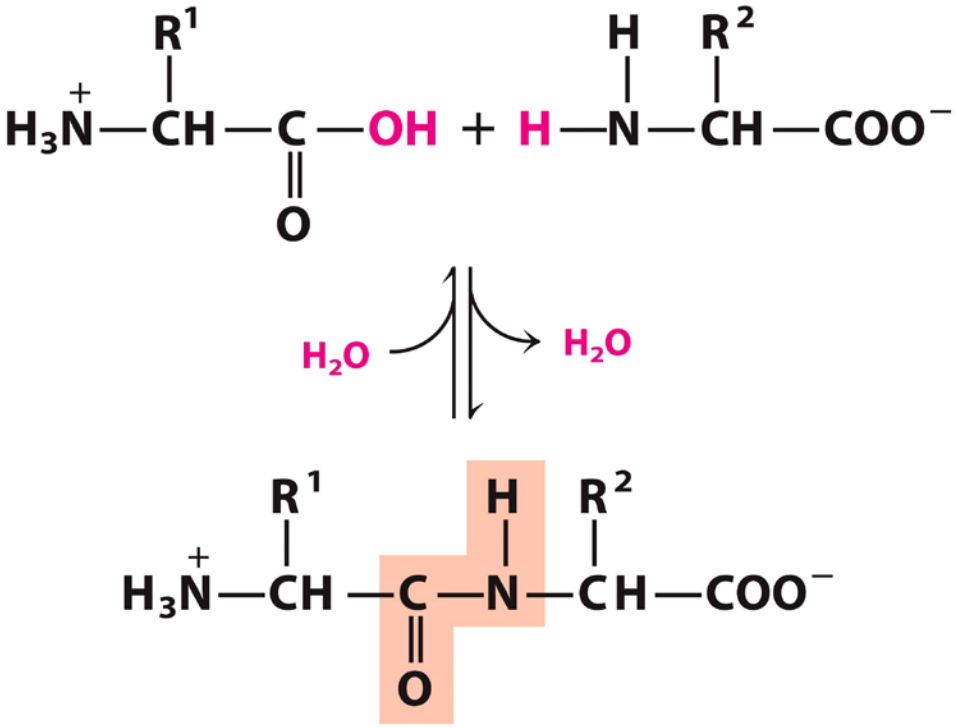
3.2 Peptides and Proteins

3.3 Working with Proteins

3.4 Protein Primary Structure

Formation of Peptides

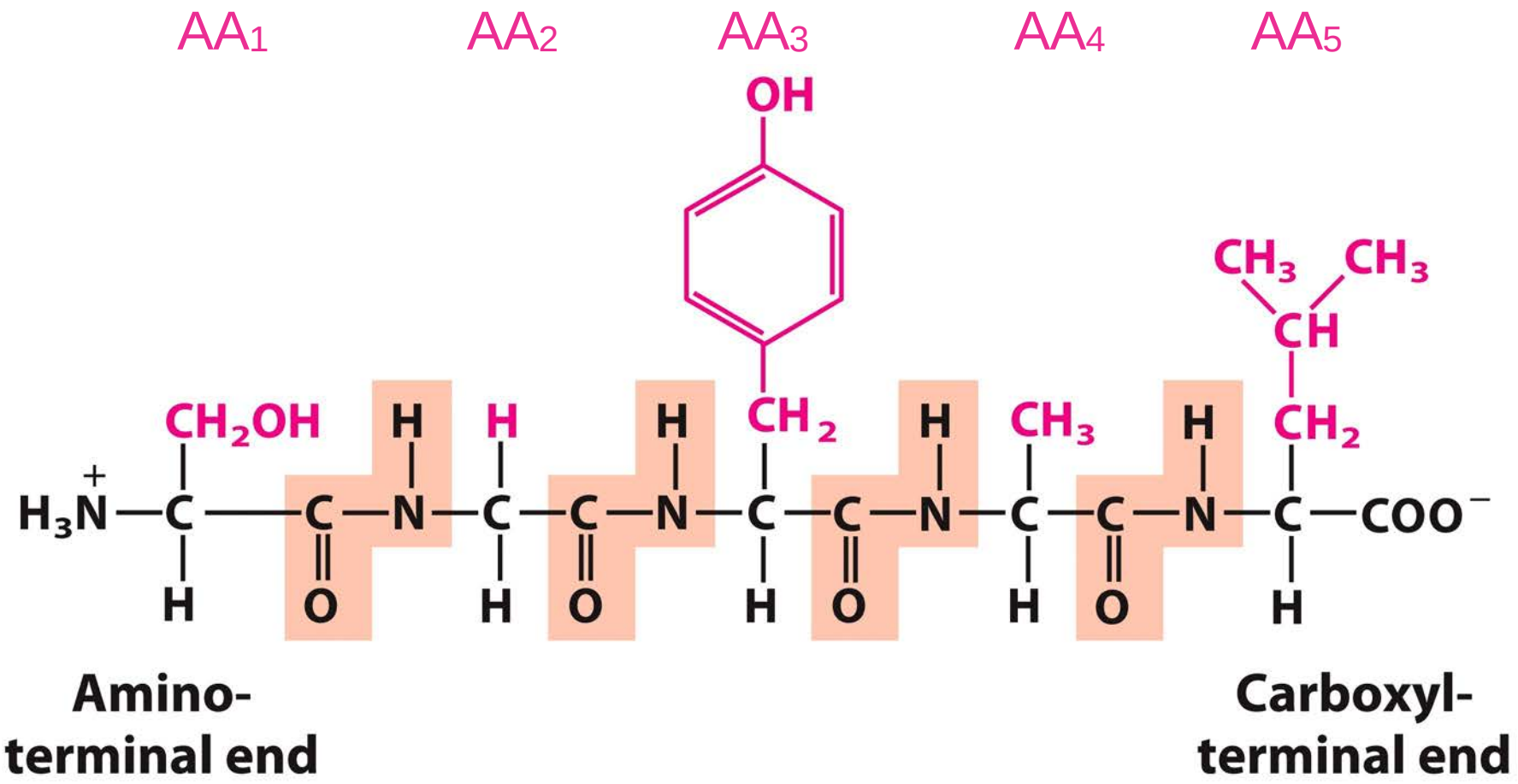
- Peptides are small condensation products of amino acids
- They are “small” compared to proteins (MW < 10 kDa)



AA #	*Product*
two	**di**peptide
three	**tri**peptide
four	**tetra**peptide
five	**penta**peptide
a few	**oligo**peptide
many	**poly**peptide

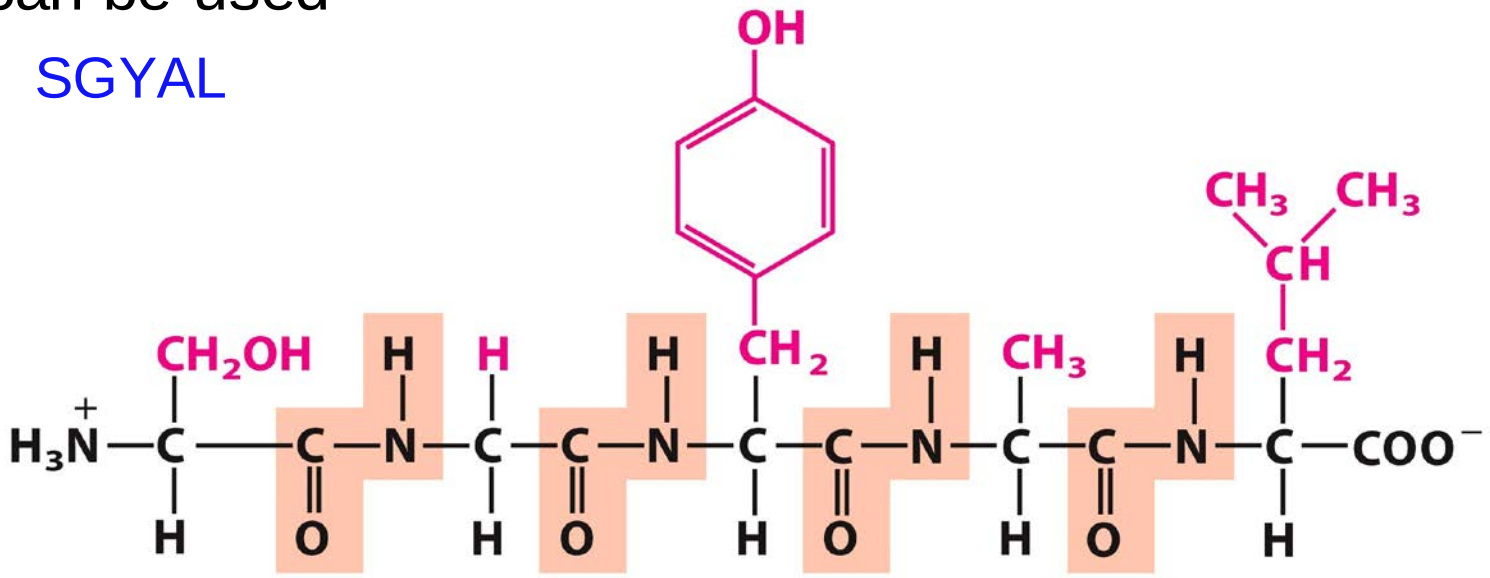
Peptide Ends Are Not the Same

Numbering (and naming) starts from the amino terminus



Naming Peptides: Start at N-terminus

- Using full amino acid names
  - Serylglycyltyrosylalanylleucine
- Using the three-letter code abbreviation
  - Ser-Gly-Tyr-Ala-Leu
- For longer peptides (like proteins), the one-letter code can be used
  - SGYAL



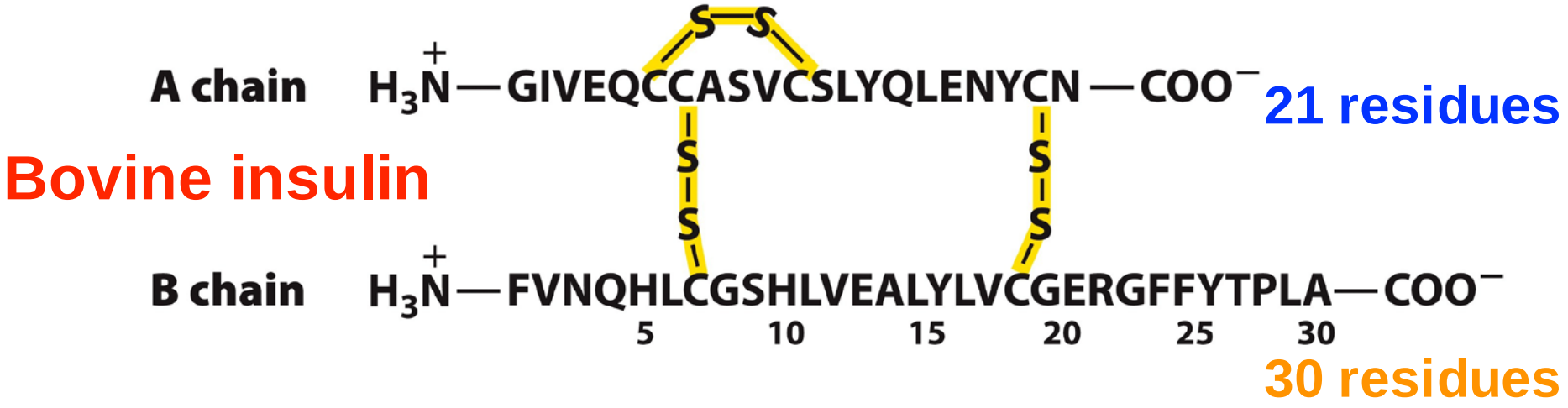
Peptides: A Variety of Functions

- Hormones and pheromones
  - **Insulin**
  - Glucagon
  - Oxytocin
- Neuropeptides
  - Substance P (pain mediator)
- Antibiotics
  - Polymyxin B (for Gram- bacteria)
  - Bacitracin (for Gram+ bacteria)
- Protection, e.g., toxins
  - Amanitin (mushrooms)
  - Conotoxin (cone snails)
  - Chlorotoxin (scorpions)

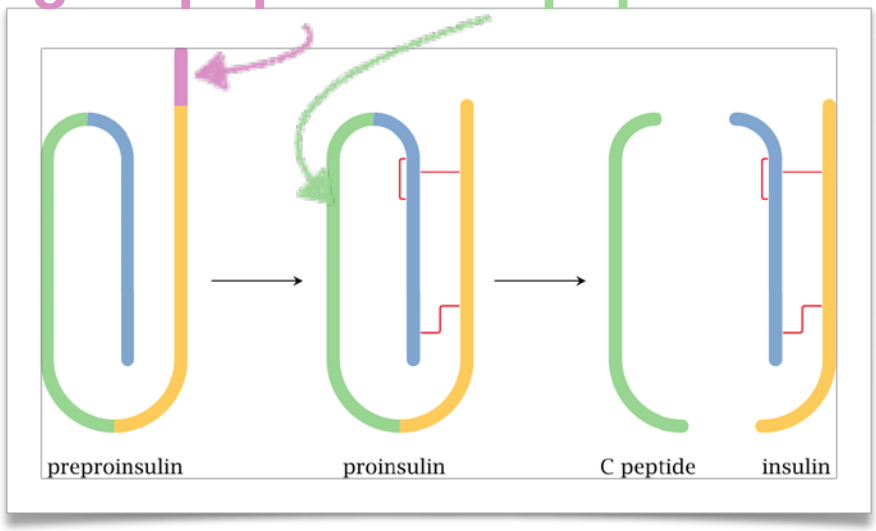
Human insulin

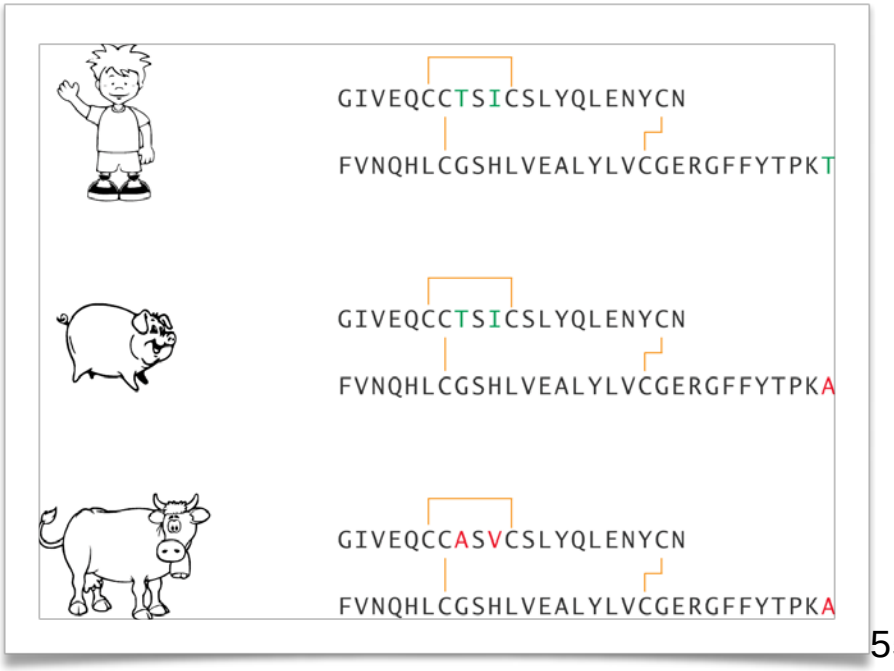
- peptide hormone
- 110 residues (5800 Da)
- two polypeptide chains
  - A-chain and B-chain
  - linked by disulfide bonds
- produced by pancreatic beta cells

Insulin

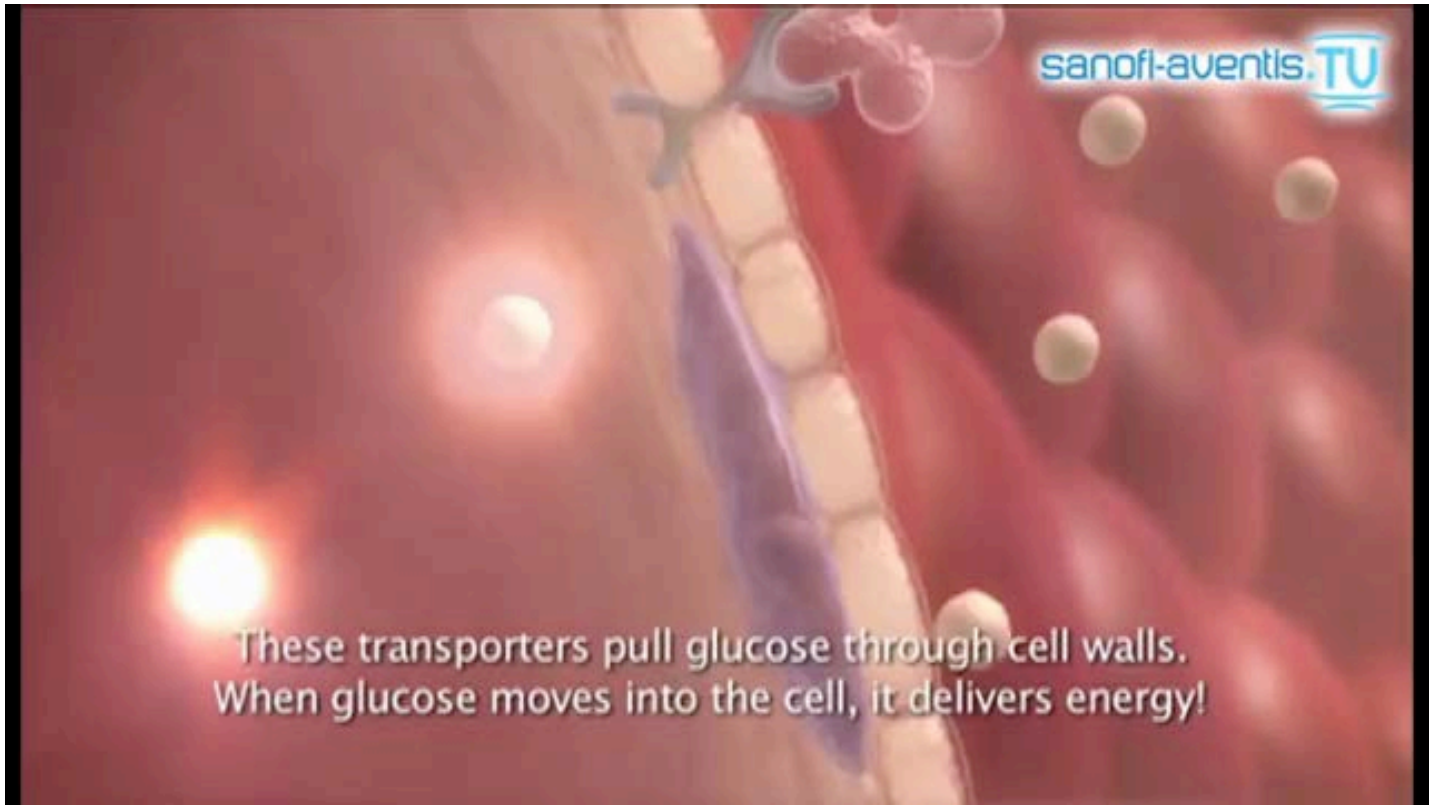


signal peptide C peptide





Insulin, Glucose and Diabetes



insulin 胰岛素
glucose 葡萄糖
small intestine 小肠
pancreas 胰腺
diabetes 糖尿病
insulin deficiency 胰岛素缺乏症
insulin resistance 胰岛素抗性

Polypeptide Size and Number Varies Greatly in Proteins

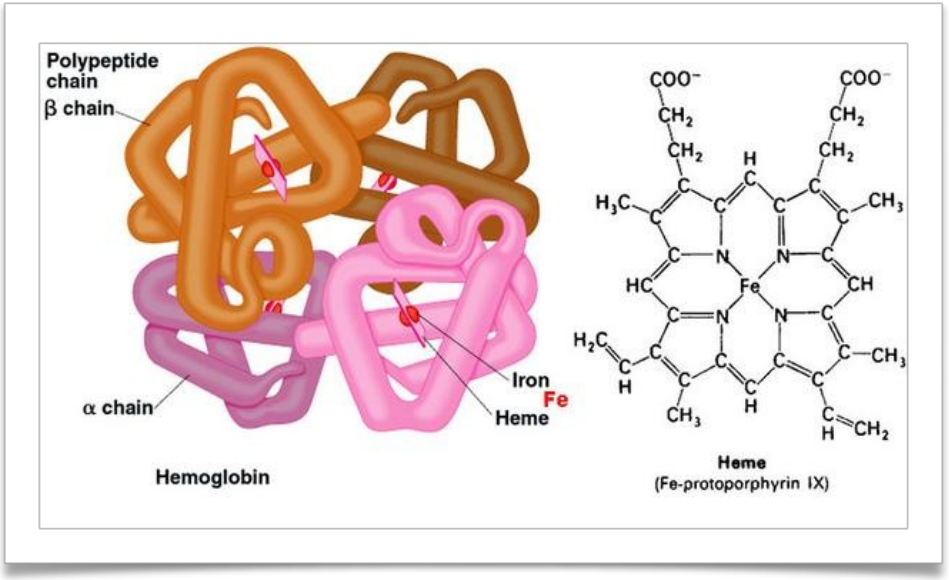
TABLE 3–2 Molecular Data on Some Proteins

	Molecular weight	Number of residues	Number of polypeptide chains
Cytochrome *c* (human)	12,400	104	1
Ribonuclease A (bovine pancreas)	13,700	124	1
Lysozyme (chicken egg white)	14,300	129	1
Myoglobin (equine heart)	16,700	153	1
Chymotrypsin (bovine pancreas)	25,200	241	3
Chymotrypsinogen (bovine)	25,700	245	1
Hemoglobin (human)	64,500	574	4
Serum albumin (human)	66,000	609	1
Hexokinase (yeast)	107,900	972	2
RNA polymerase (*E. coli*)	450,000	4,158	5
Apolipoprotein B (human)	513,000	4,536	1
Glutamine synthetase (*E. coli*)	619,000	5,628	12
Titin (human)	2,993,000	26,926	1

multisubunit protein

Simple Proteins and Conjugated Proteins

- Simple proteins:
  - amino acid residues only
  - no other chemical constituents
- Conjugated proteins:
  - polypeptide chain(s)
  - prosthetic groups
    - lipids
    - carbohydrates
    - phosphate groups
    - metal ions
    - others



Classes of Conjugated Proteins

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

Summary 3.2 Peptides and Proteins

- Amino acids are joined through peptide bonds to form peptides and proteins. Cell contains thousands of different proteins
- Peptides have only a few amino acid residues. Proteins can be very long chains of 100 to several thousands residues. Multisubunit proteins contain several polypeptide chains
- Conjugated proteins contain prosthetic groups

Example Question

Lysine has a molecular weight of 146. In protein ribonuclease, Lys residues make up 10.5% of the weight of ribonuclease. The ribonuclease molecule contains 10 Lys residues. Calculate the molecular weight of ribonuclease.

A) 146 * 10 / 10.5%

B) (146 - 18) * 10 / 10.5%

Example Question

The peptide alanylglutamylglycylalanylleucine has:

A) a disulfide bridge.

B) five peptide bonds.

C) four peptide bonds.

D) no free carboxyl group.

E) two free amino groups.

Example Question

In a conjugated protein, a prosthetic group is:

A) a fibrous region of a globular protein.

B) a nonidentical subunit of a protein with many identical subunits.

C) a part of the protein that is not composed of amino acids.

D) a subunit of an oligomeric protein.

Example Question

Once we know the gene sequence of a protein, we can deduce the amino acid sequence of the protein. But even when we know the gene sequence of a protein, chemical studies of the protein are still required to determine:

A) the molecular weight of the protein.
B) the amino-terminal amino acid.
C) the location of disulfide bonds.
D) the number of amino acids in the protein.
E) whether the protein has the amino acid methionine in its sequence.

Amino Acids, Peptides, and Proteins

3.1 Amino Acids

3.2 Peptides and Proteins

3.3 Working with Proteins

3.4 Protein Primary Structure

What to Study about Proteins

What is its sequence and composition?

What are its physico-chemical properties?

What is its three-dimensional structure?

How does it find its native fold?

How does it achieve its biochemical role?

Where is it localized within the cell?

How does it interact with other macromolecules?

How is its function regulated?

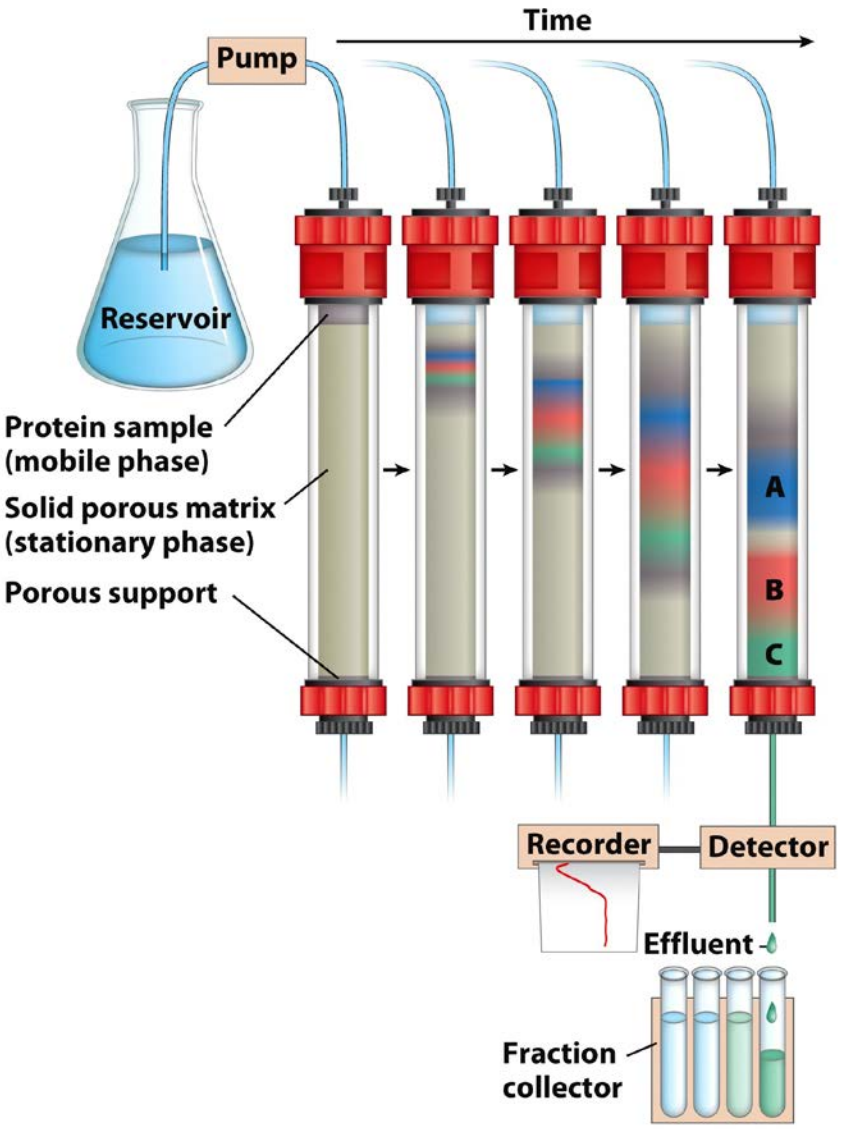
How is it related to other proteins?

FIRST STEP: separate in pure form

A mixture of proteins can be separated

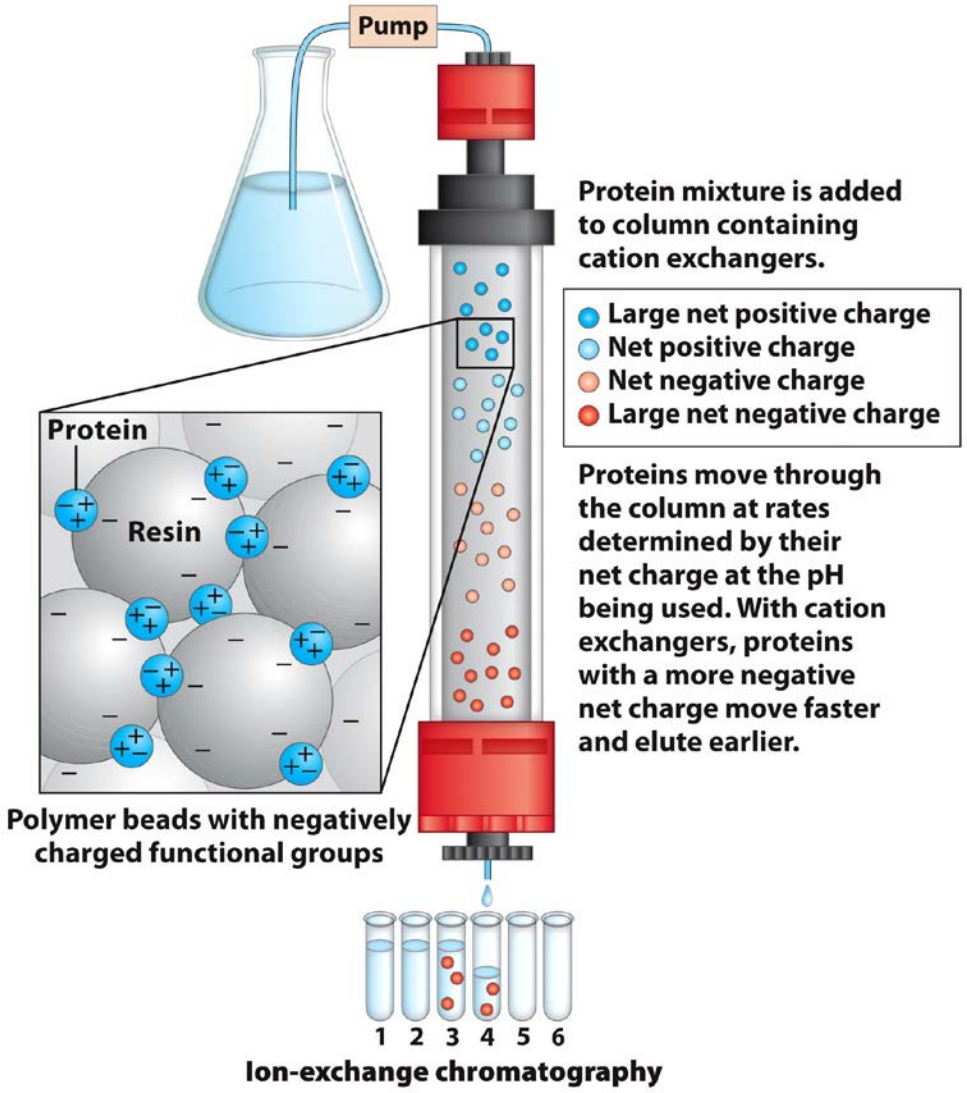
- Separation relies on differences in physical and chemical properties
  - Size
  - Charge
  - Solubility
  - Affinity for a ligand
  - Hydrophobicity
  - Thermal stability
- Chromatography is commonly used for preparative separation

Column Chromatography



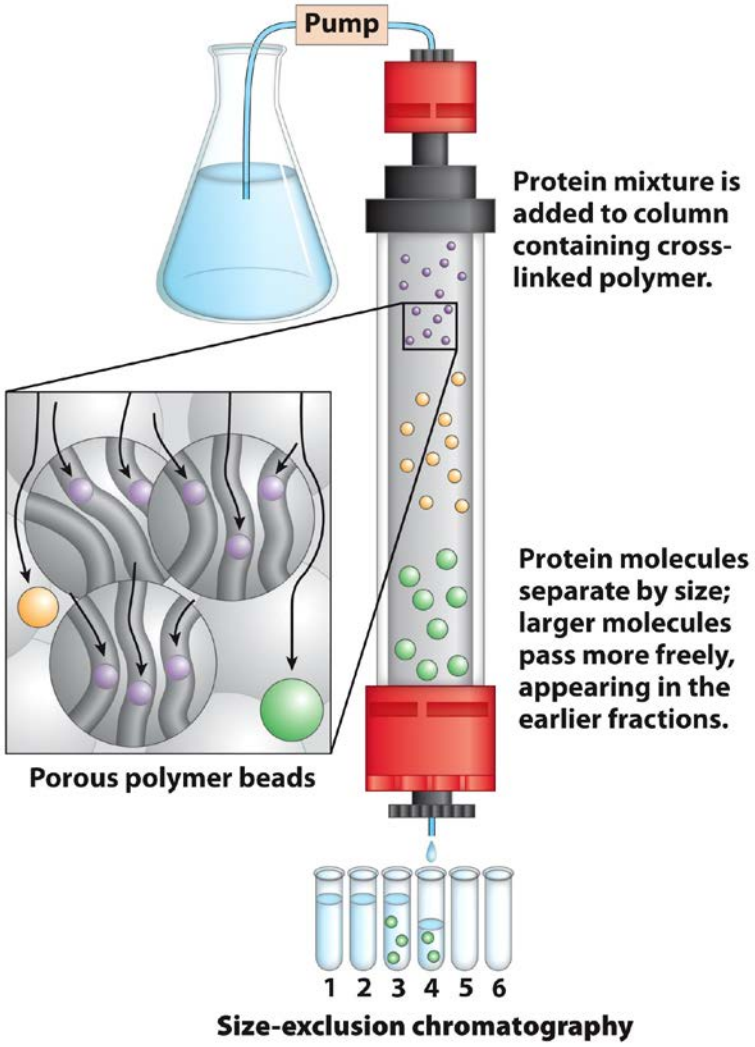
- Stationary phase
  - solid matrix
- Mobile phase
  - solution
- Protein sample
  - brought into mobile phase
  - interact with stationary phase
  - different migration speeds

Separation by Charge



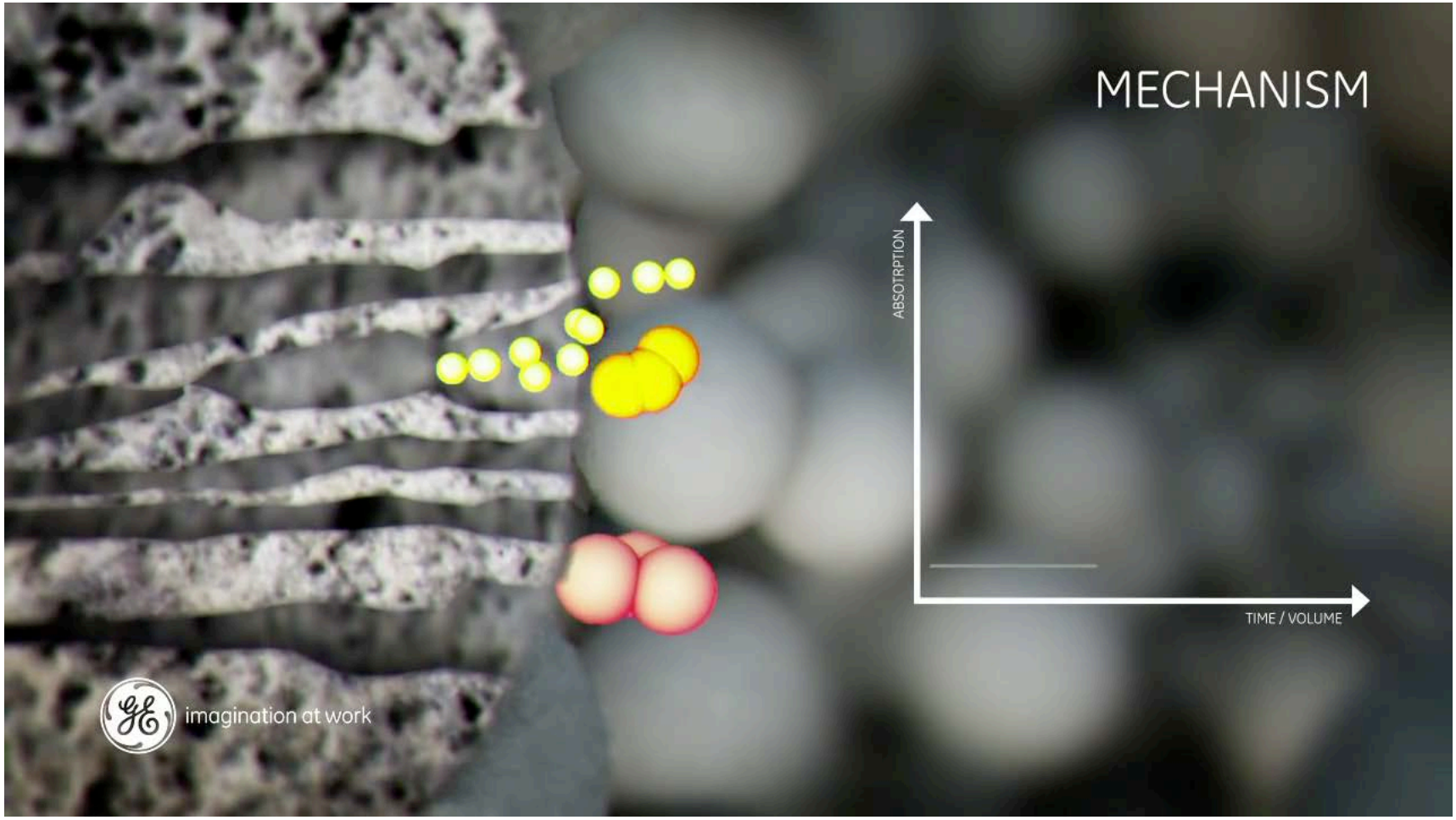
- Ion-exchange chromatography
- Sign and magnitude of net charge
- Affected by pH
- Cation exchangers
  - matrix with anionic groups
  - proteins with negative charges move faster
  - proteins with positive charges move slower
- Anion exchangers
  - …

Separation by Size

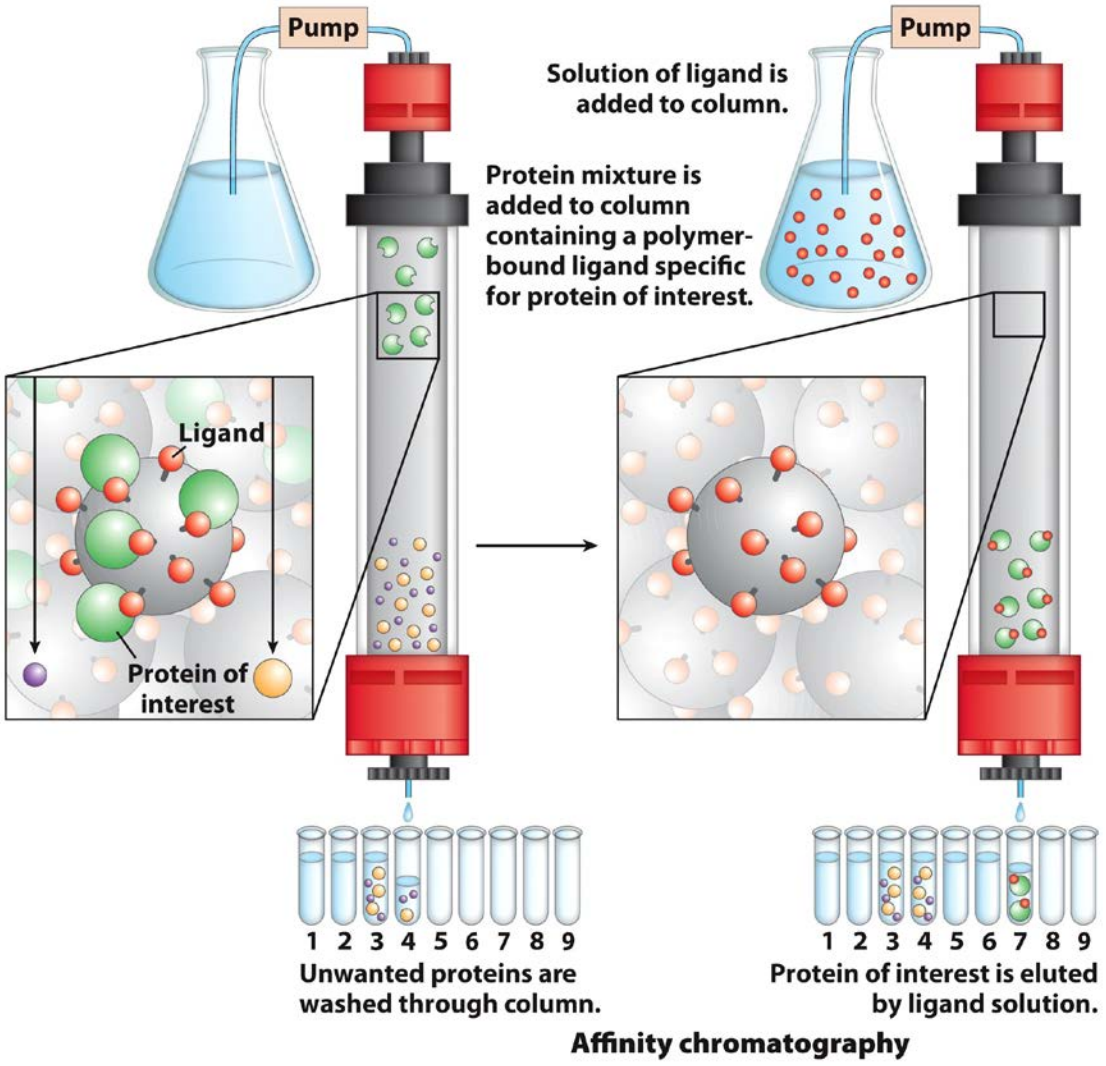


- Size-exclusion chromatography
- Or: Gel filtration chromatography
- Solid phase consists of beads with pores or cavities
- Large proteins cannot enter cavities and so move through the column faster
- Smaller proteins enter cavities and take a longer and slower path through the column

Size-Exclusion Chromatography



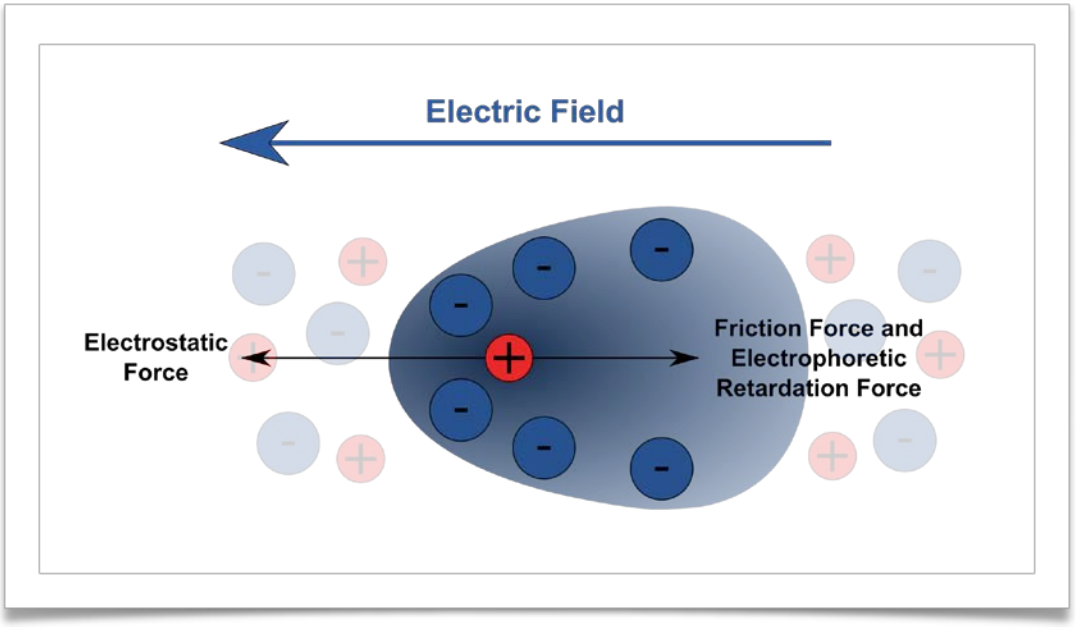
Separation by Affinity

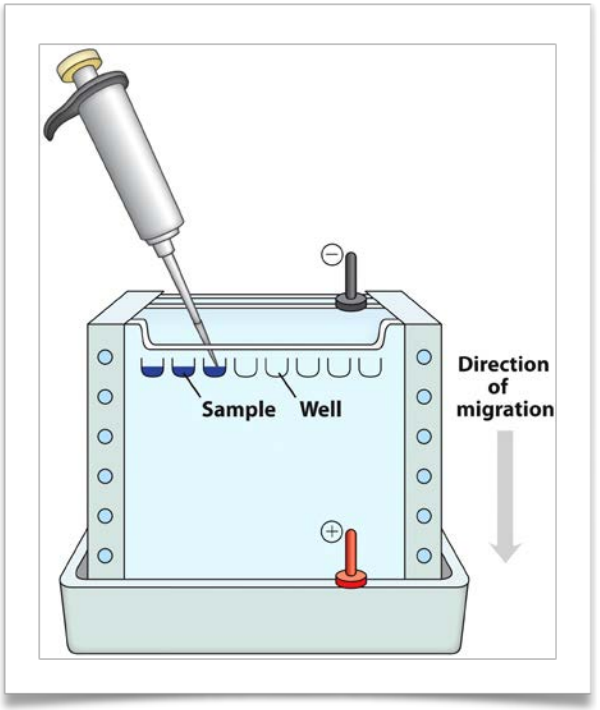


- Affinity-chromatography
- Beads covalently attached to ligands
- Proteins with affinity to ligand bind to beads
- Bound proteins **eluted** by a solution of high concentrations of salt or free ligand

Electrophoresis for Protein Analysis

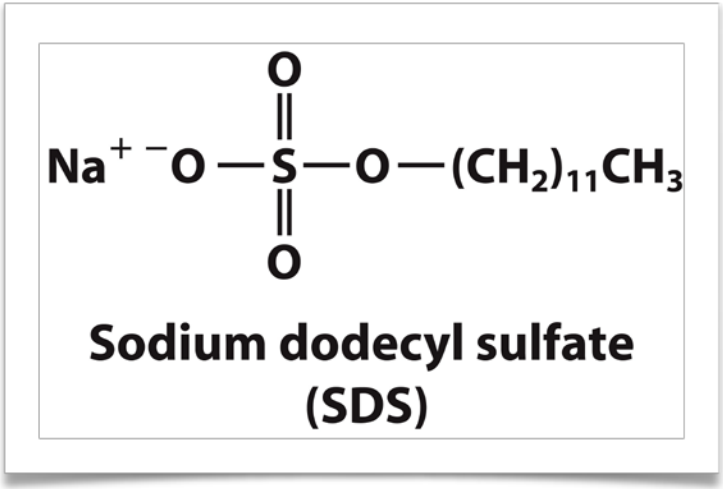
- Separation in analytical scale is commonly done by electrophoresis
  - Electric field pulls proteins according to their charge
  - Gel matrix hinders mobility of proteins according to their size and shape





SDS-PAGE: Molecular Weight

- SDS: sodium dodecyl sulfate (a detergent)
- SDS micelles bind to and unfold all the proteins
  - SDS gives all proteins an uniformly negative charge
  - The native shape of proteins does not matter
  - Rate of movement will only depend on size: small proteins will move faster



AA #	*Product*
two	dipeptide
three	tripeptide
four	tetrapeptide
five	pentapeptide
twelve	dodecapeptide
a few	oligopeptide
many	polypeptide

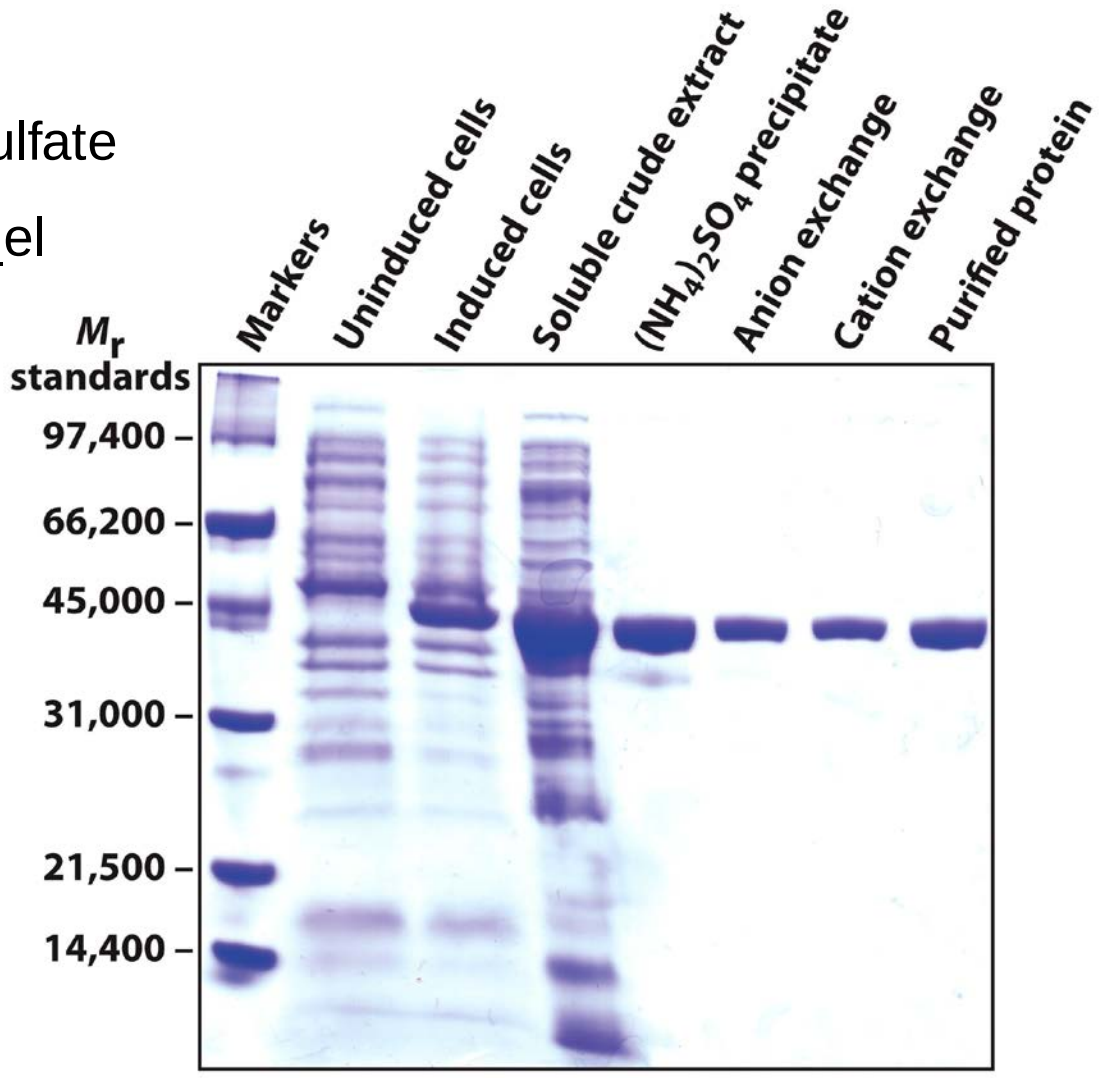
Protein Visualized by Staining

SDS-PAGE

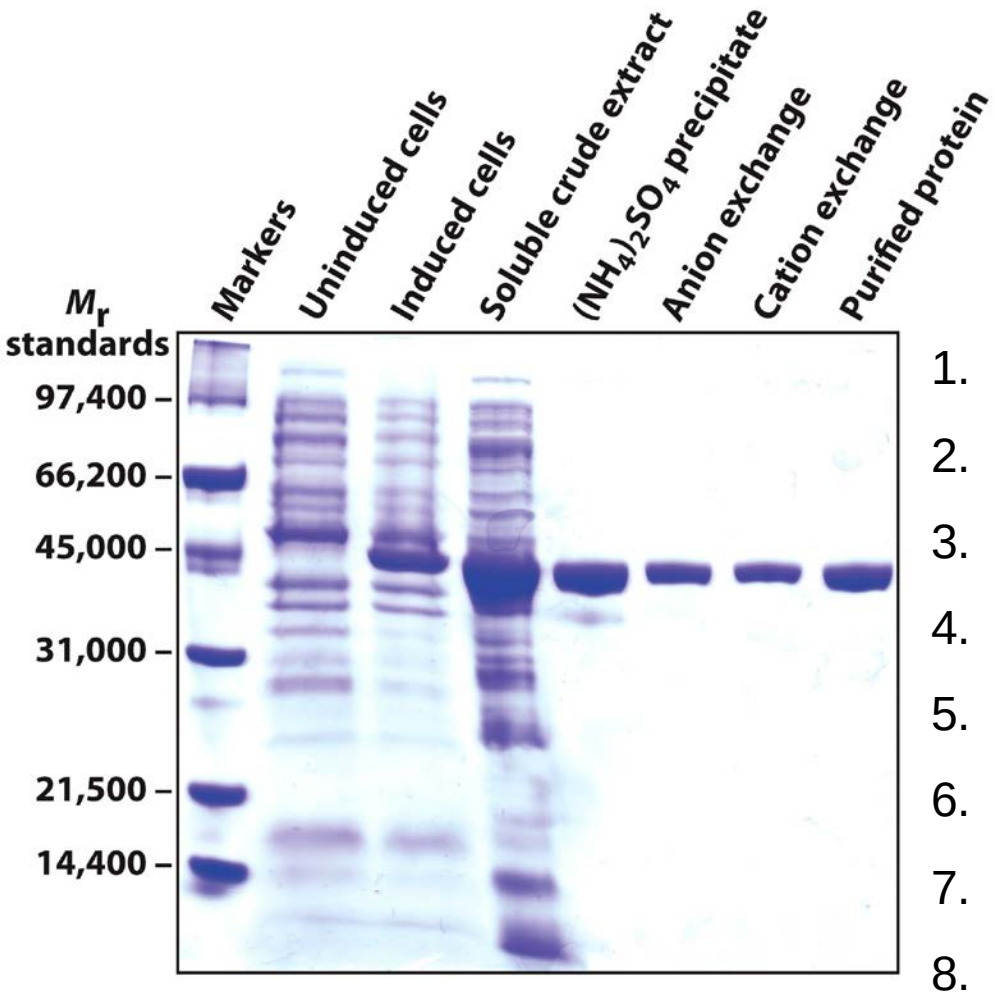
SDS: Sodium Dodecyl Sulfate

PAGE: Polyacrylamide Gel

Electrophoresis

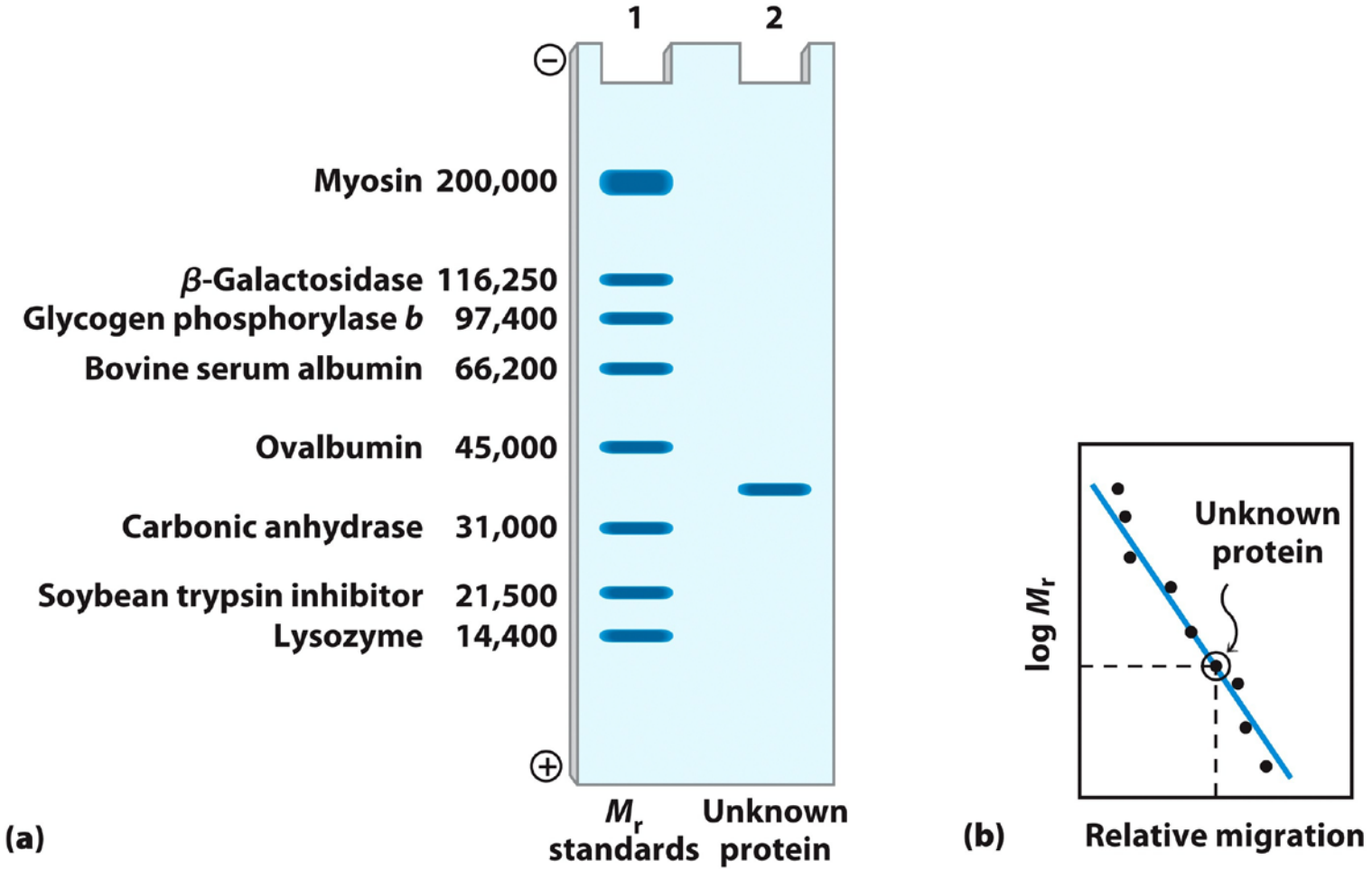


Protein Visualized by Staining



1. proteins with known molecular weight
2. cells NOT expressing target protein
3. cells expressing target protein
4. contents from lots of broken cells
5. separation method 1 - precipitate
6. separation method 2 - ion exchange
7. separation method 3 - ion exchange
8. purified protein

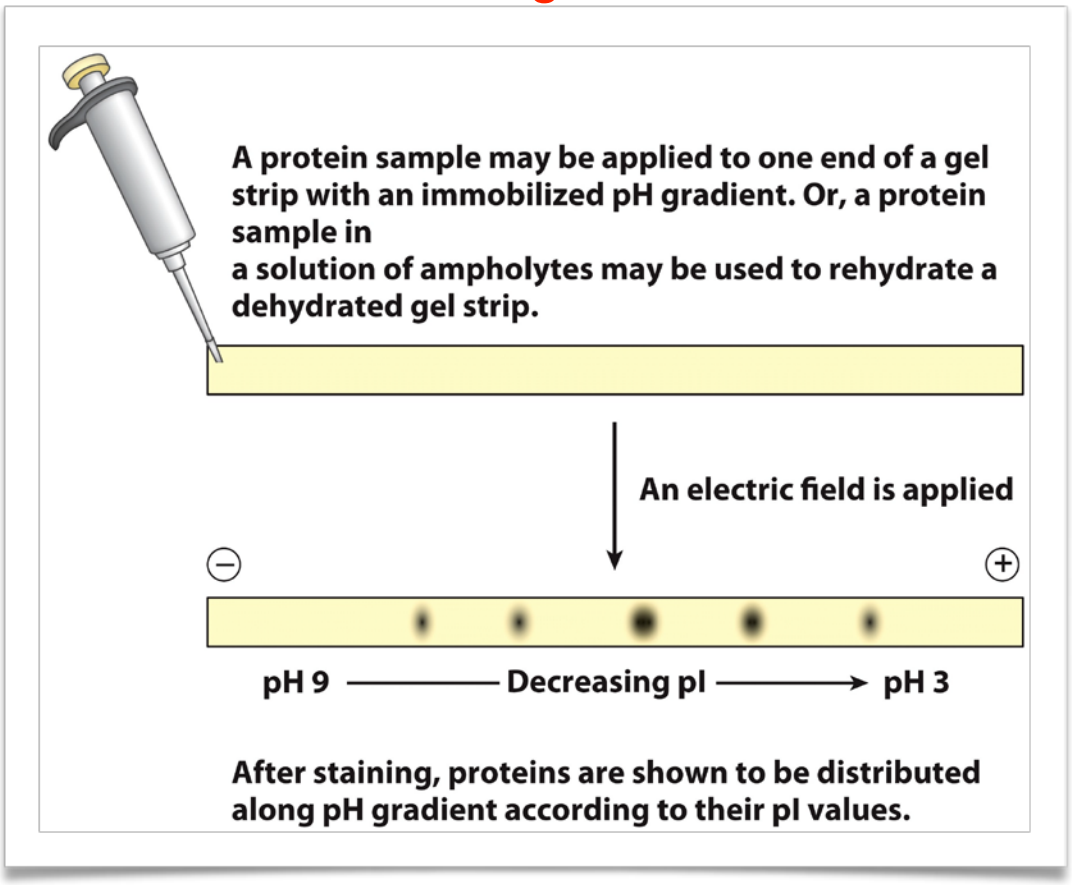
Estimate Protein Molecular Weight



Isoelectric Focusing Determines pI

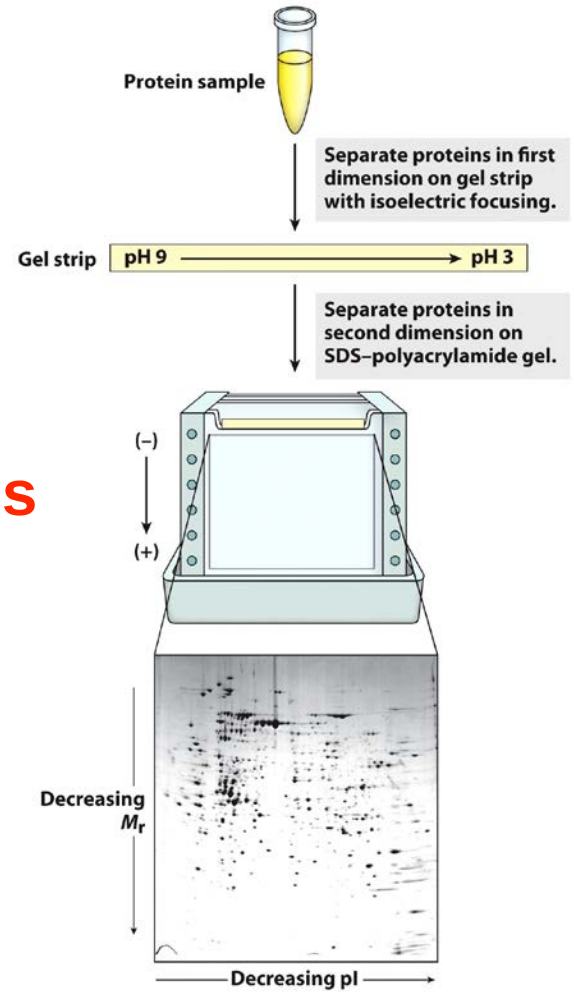
At pI

- Net charge is zero
- Does not migrate in electric field

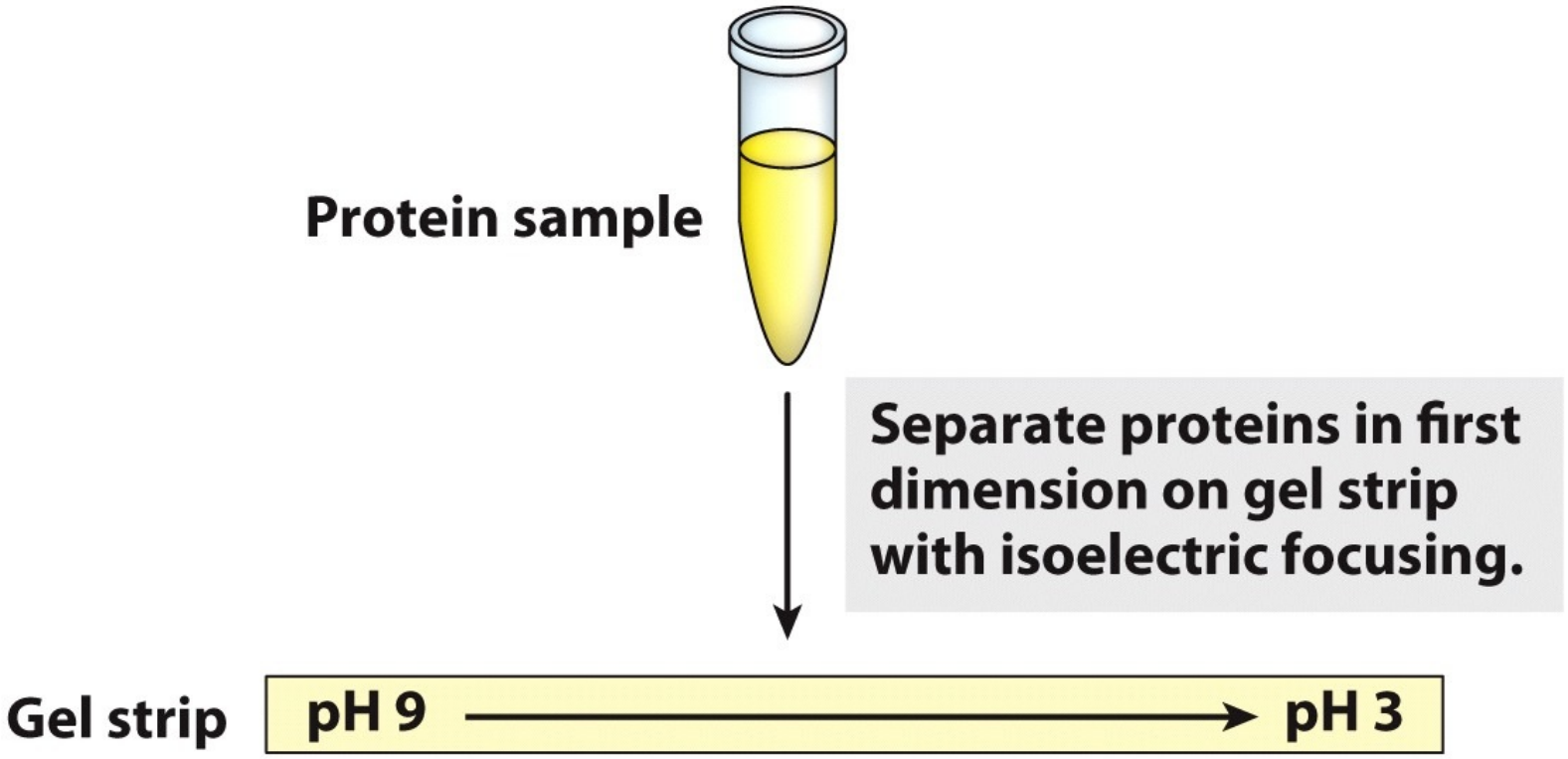


IEF and SDS-PAGE Combined

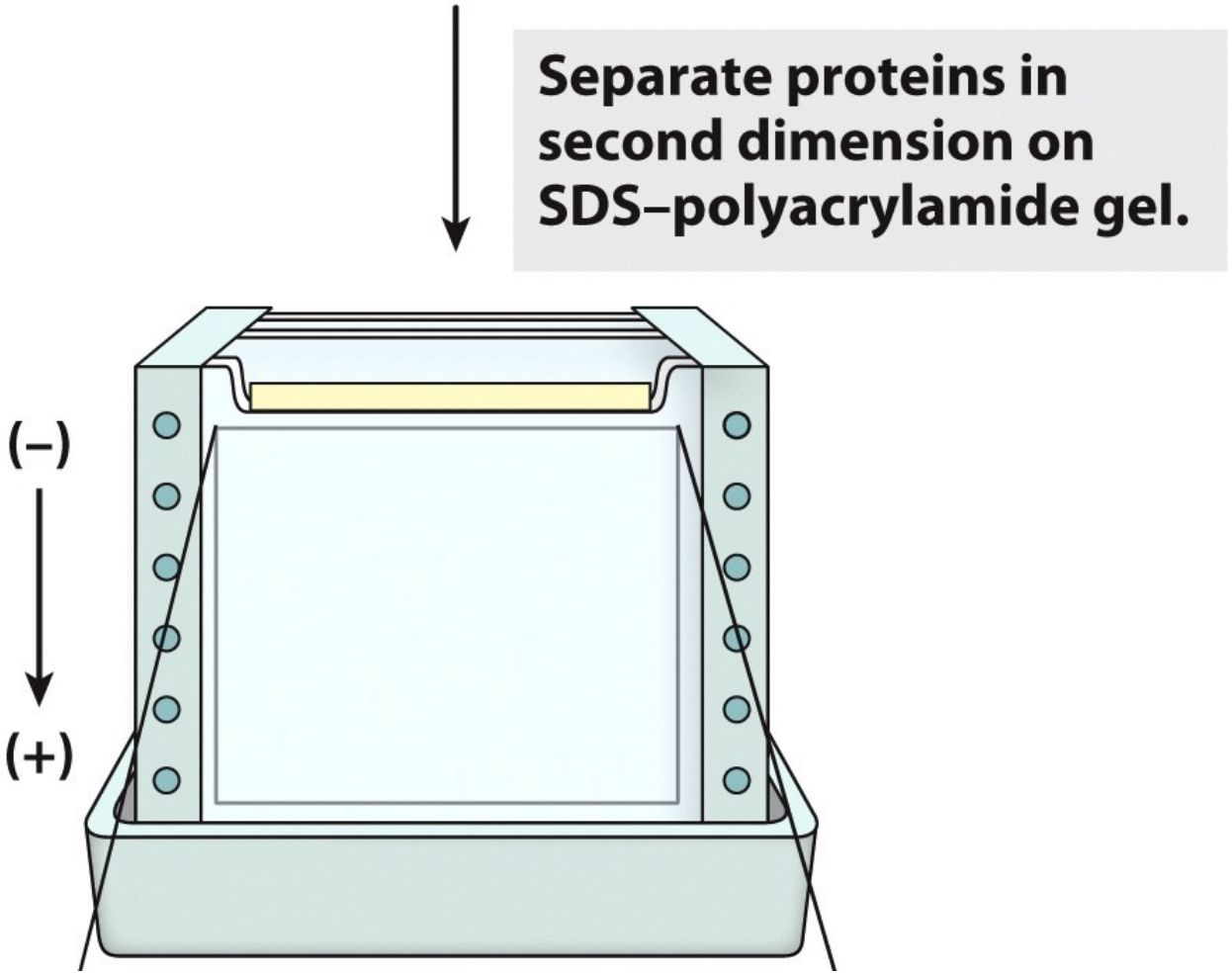
Two-dimensional gel electrophoresis or 2D gel



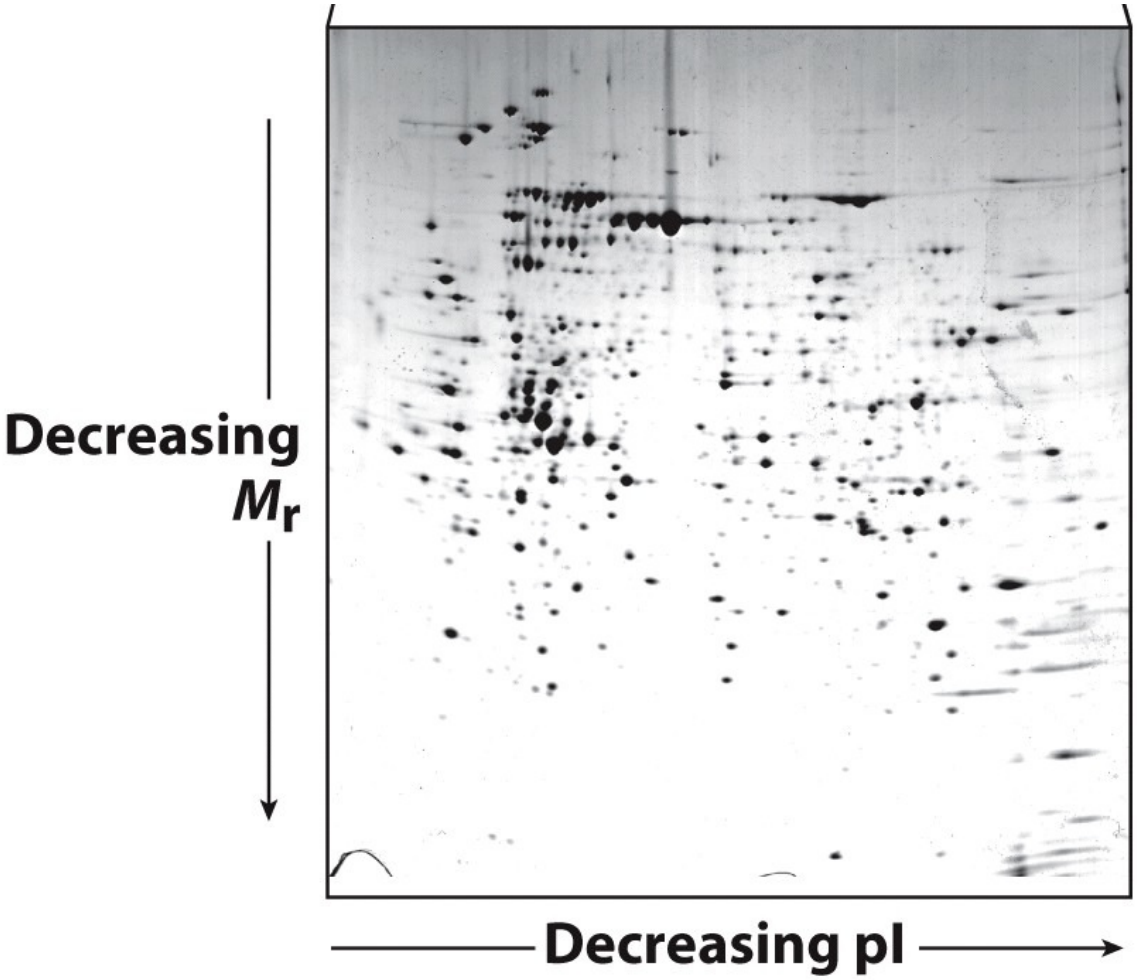
IEF and SDS-PAGE Combined



IEF and SDS-PAGE Combined



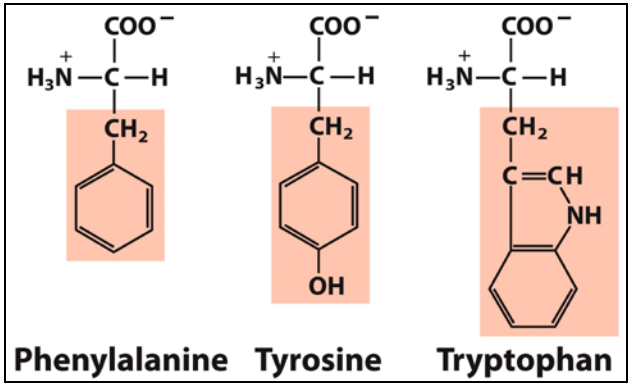
IEF and SDS-PAGE Combined

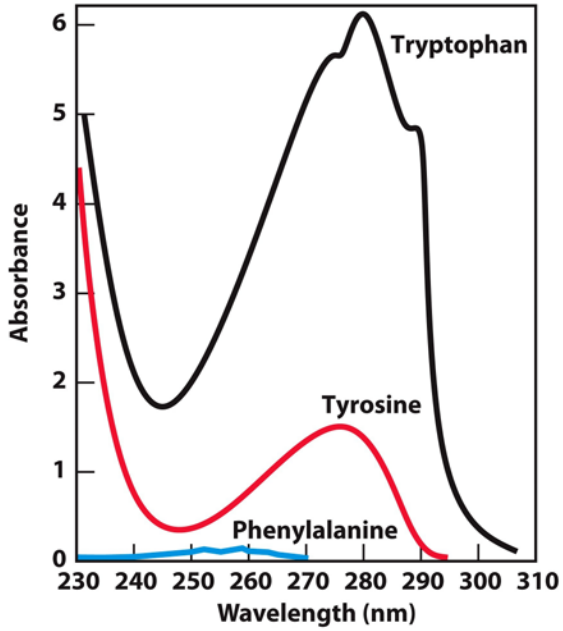


Spectroscopic Detection of Aromatic AA

- Aromatic AAs absorb light in the UV region
- Proteins have strong UV absorbance around 280 nm
- Trp and Tyr are the strongest chromophores
- Concentration can be determined by UV-visible spectrophotometry using Lambert-Beer law: A = ·c·l

A	absorbance
ε	molar extinction coefficient
c	concentration
l	path length





Summary 3.3 Working with Proteins

- Proteins are separated and purified on the basis of their physical and chemical properties
- There are a range of chromatographic procedures, including ion-exchange, size-exclusion, affinity and high-performance liquid chromatography (HPLC)
- Electrophoresis separates proteins on the basis of mass and charge, such as SDS-PAGE and IEF

Example Question

What factors would make it difficult to interpret the results of a gel electrophoresis of proteins in the absence of sodium dodecyl sulfate (SDS)?

Without SDS, protein migration through a gel would be influenced by:

1) the protein's intrinsic net charge—which could be positive or negative.
2) its molecular weight.
3) its unique three-dimensional shape.
4) Thus, it would be difficult to ascertain the difference between proteins based upon a comparison of their mobilities in gel electrophoresis.

Amino Acids, Peptides, and Proteins

3.1 Amino Acids

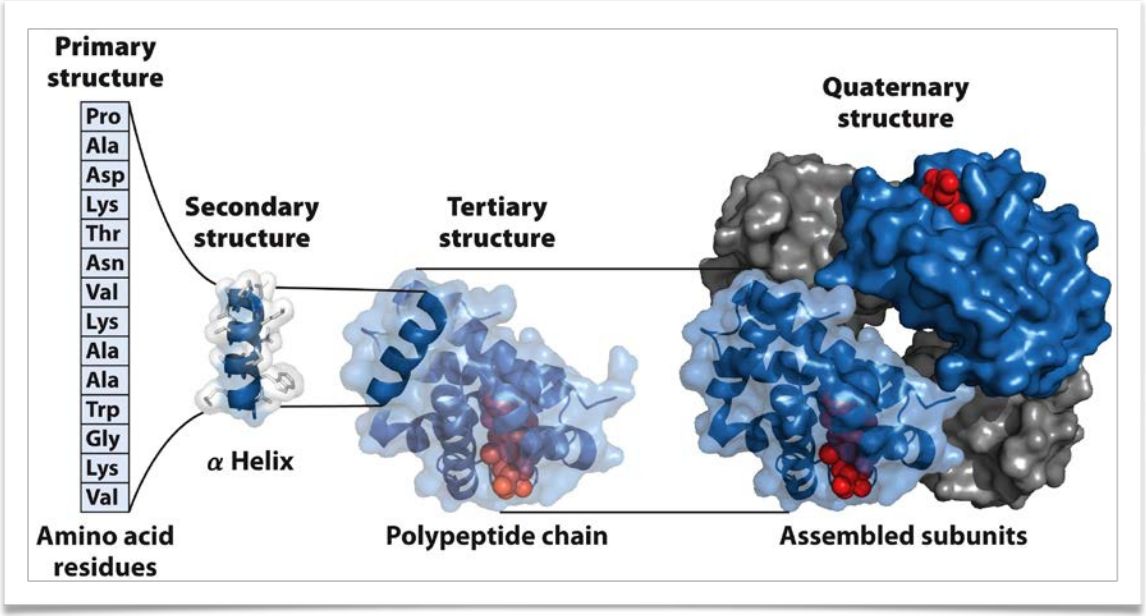
3.2 Peptides and Proteins

3.3 Working with Proteins

3.4 Protein Primary Structure

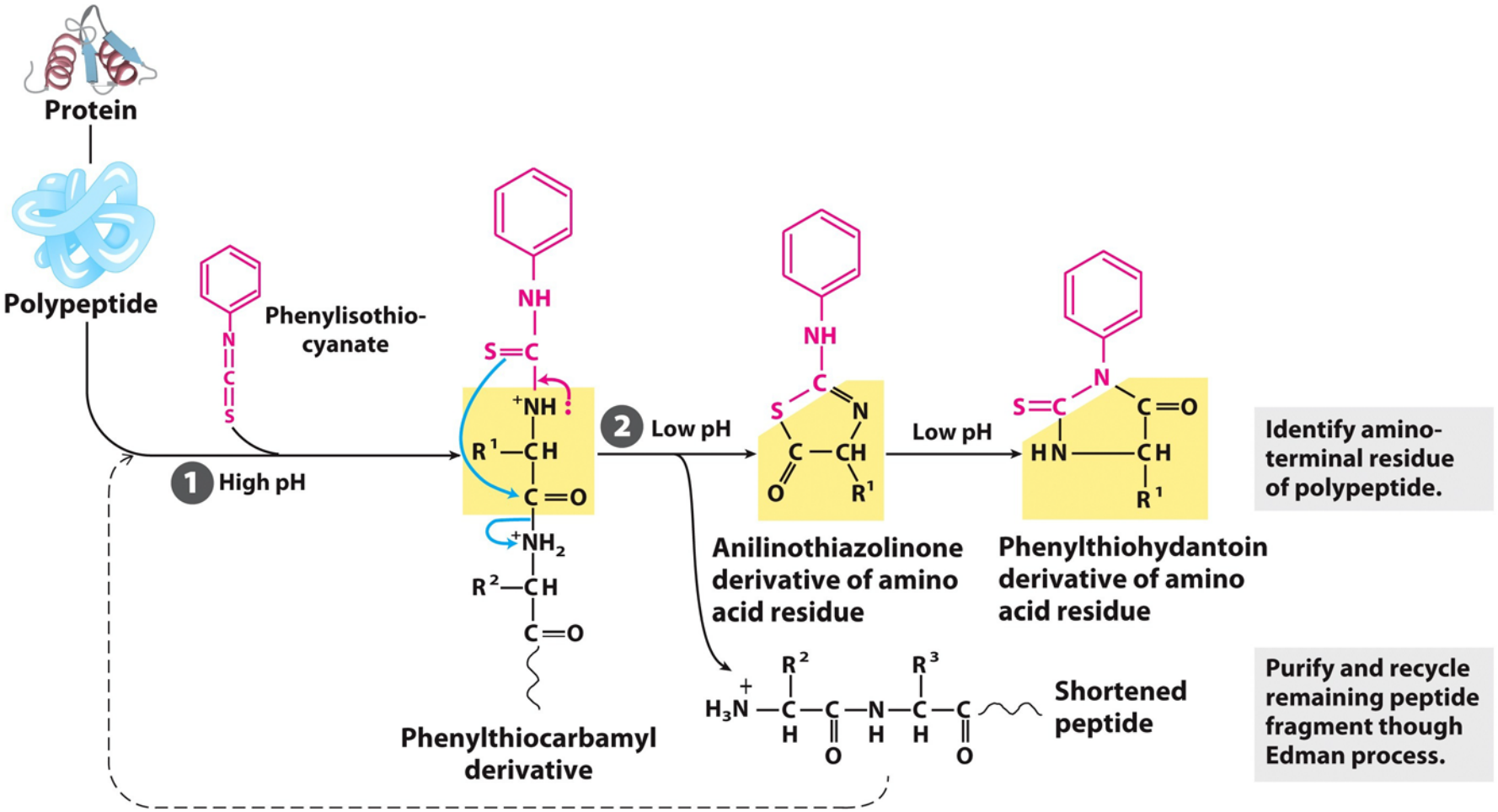
Protein Sequencing

- Essential to know the protein sequence
  - Actual sequence determined from DNA sequence
- Edman Degradation (Classical method)
  - Successive rounds of modification, cleavage, and identification
- Mass Spectrometry (Modern method)
  - MALDI MS and ESI MS can precisely identify sequence



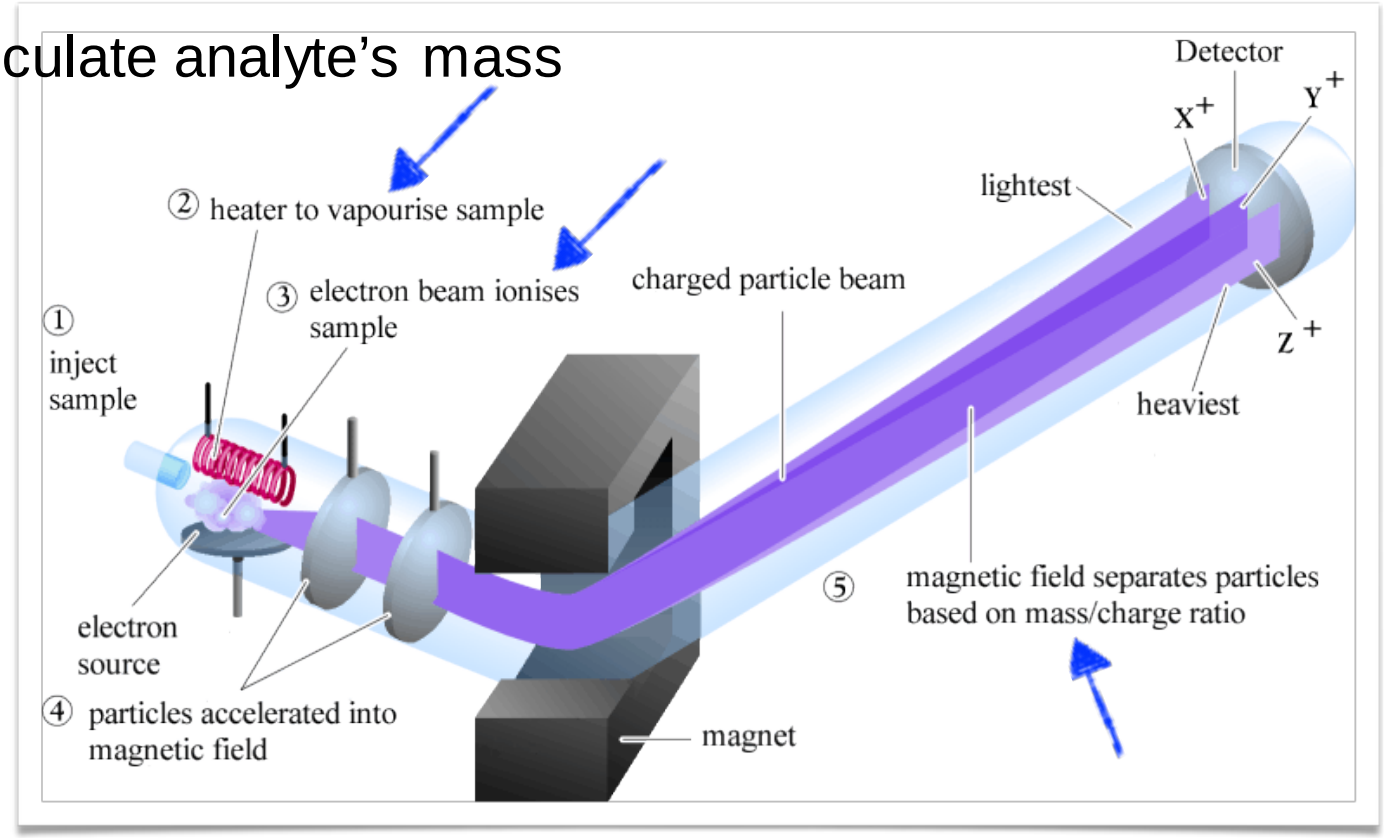
- amino acid sequence determines protein structure
- protein structure defines protein function

Edman's Degradation

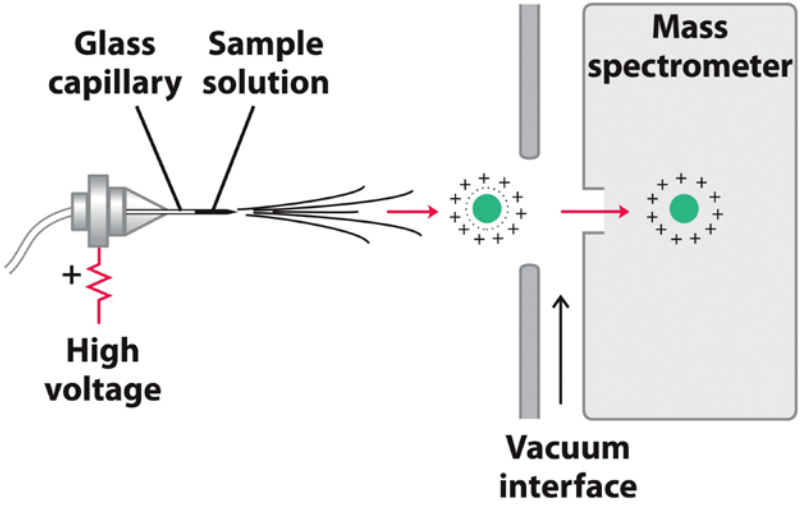


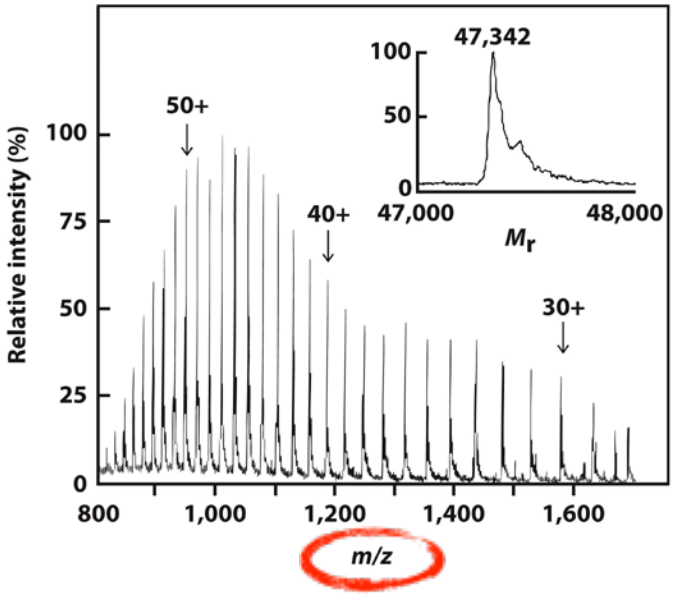
Mass Spectrometry

1. Vaporize and ionize the molecule to be analyzed (analyte)
2. Analyte travels through an electric field and/or magnetic field
3. Monitor analyte's path in the field
4. Path is directly related to analyte's mass-to-charge ratio (*m/z*)
5. Calculate analyte's mass



MS Procedures for Sequence IDs

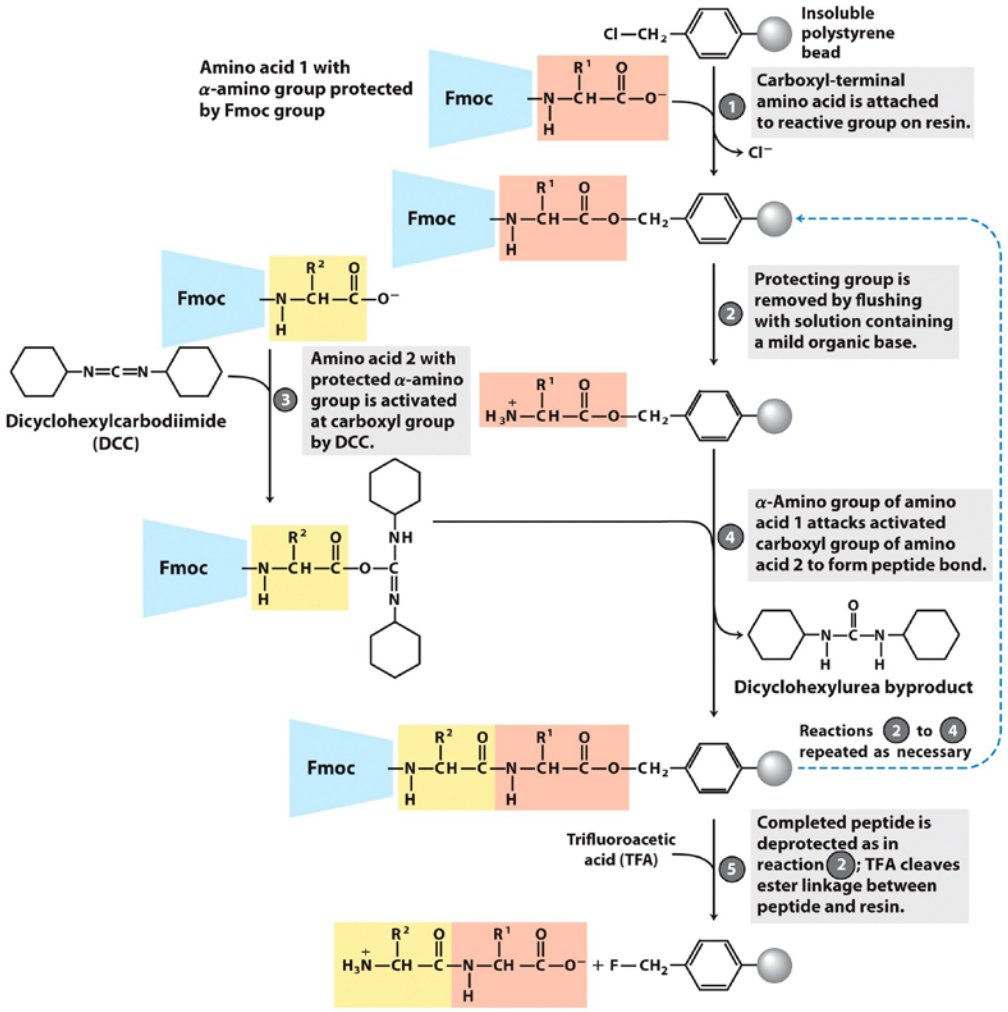




Electrospray ionization mass spectrometry

- Charged microdroplets
- Solvent rapidly evaporates
- Multiply charged protein ion in gas phase
- Protein with different m/z ratios
  - …
  - (47342 + 34) / 34 = 1393
  - (47342 + 35) / 35 = 1354
  - (47342 + 36) / 36 = 1316
  - (47342 + 37) / 37 = 1280
  - (47342 + 38) / 38 = 1247
  - (47342 + 39) / 39 = 1215
  - …

Chemically Synthesize Small Peptides



- Synthesis proceeds from C terminus to N terminus
- Limited by efficiency of each chemical cycle

TABLE 3–7 Effect of Stepwise Yield on Overall Yield in Peptide Synthesis

Number of residues in the final polypeptide	Overall yield of final peptide (%) when the yield of each step is:	
	96.0%	99.8%
11	66	98
21	44	96
31	29	94
51	13	90
100	1.8	82

Table 3-7
Lehninger Principles of Biochemistry, Sixth Edition

Summary 3.4 Protein Primary Structure

- Differences in protein structure/function result from differences in amino acid sequence
- Amino acid sequences can be deducted from the nucleotide sequence of its corresponding gene DNA. they can also be deducted by Edman's degradation method or mass spectrometry
- Short peptides (up to 100 residues) can be chemically synthesized

Example Question

Describe the differences between chemical synthesis of polypeptides and synthesis of polypeptides in the cell.

1) Chemical synthesis proceeds from carboxyl terminus to amino terminus; in the living cell, the process starts at the amino terminus and ends at the carboxyl terminus.
2) In the living cell, synthesis occurs under physiological conditions; chemical synthesis does not.
3) Chemical synthesis is only capable of synthesizing short polypeptides; cells can produce proteins of several thousand amino acids.
4) It takes a few days to synthesize a 100-residue peptide; the same protein would be synthesized in 5 seconds in a cell.

Chapter 3: Summary

In this chapter, we learned about:

- The structures and names of amino acids found in proteins
- The ionization properties of amino acids and peptides
- The many biological functions of peptides and proteins
- The methods for separation and analysis of proteins