Chapter 3 Proteins

Outline

- Classification of proteins
- The basic building blocks of proteins—Amino acids
- Structure of proteins
- Structure and function relationships of proteins
- Properties of proteins
- Separation, purification and determination of proteins

What are proteins?

Proteins are macromolecules composed of amino acids linked together through peptide bonds, which have a stable conformation and a certain biological function.

Proteins are linear polymers聚合物 built of monomer单体 units called amino acids.

Section 1
Classification of proteins

1. Chemical Components of Proteins

- **Major elements: C, H, O, N, S**
- **Trace elements: P, Fe, Cu, Zn, Mo, I, …**
- **The average nitrogen content in proteins is about 16%, and proteins are the major source of N in biological systems.**

Section 1
Classification of proteins

1. Chemical Components of Proteins

- **The protein quantity can be estimated---Kjeldahl determination（凯氏定氮法）**
- **protein in 100g sample = N per gram × 6.25 × 100**

2. Protein Classification

◆**Classification based on the overall shape**

Globular protein（球状蛋白质）:

globular or ellipsoidal (long/short <10), soluble in water; including enzymes, transportors, receptors, regulators, …

Fibrous protein（纤维状蛋白质）:

highly elongated; insoluble in water; including collagen（胶原蛋白）, elastin（弹性蛋白）, α-keratin（α-角蛋白）, …

◆Classification based on chemical compositions

Simple protein（简单蛋白质）:

made up of amino acids completely, without non-protein components

Conjugated protein（缀合蛋白质）:

conjugated protein = apoprotein + prosthetic groups

Prosthetic group （辅基）is non-protein part, binding to protein by covalent bond. This group can be carbohydrates, lipids, nucleic acids, phosphates, pigments（色素）, or metal ions.

Category of conjugated proteins

Category（类别）	Prosthetic group （辅基）	Example（举例）
Nucleoprotein（核蛋白）	nucleic acids	Chromosome，ribosome
Glycoprotein（糖蛋白）	carbohydrates	Immunoglobulin（Ig）
Lipoprotein（脂蛋白）	lipids	High-density lipoprotein
Phosphoprotein（磷蛋白）	phosphates	Casein（酪蛋白）
Chromoprotein（色蛋白）	pigments	Hemoglobin（血红蛋白Hb）
Metalloprotein（金属蛋白）	metal ions	Ferritin（铁蛋白），calmodulin（钙调蛋白）

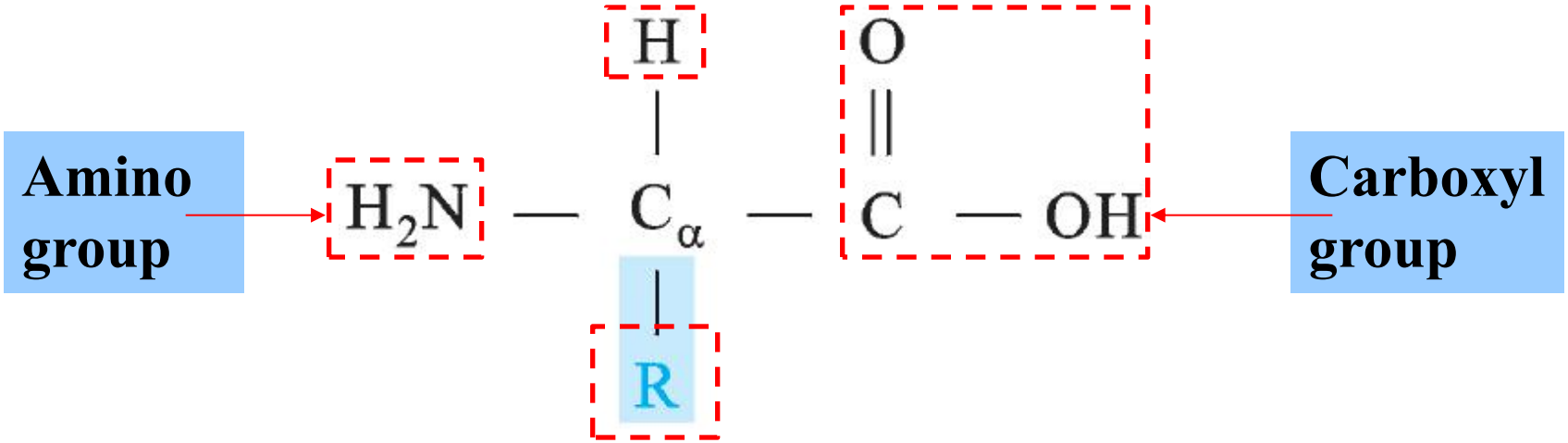
◆Classification based on biological functions

- Enzymes – Ribonuclease（核糖核酸酶）
- Regulatory proteins – Insulin and growth hormone
- Storage proteins – ovalbumin（卵清蛋白）
- Defensive and protective proteins – Antibody, toxin（毒素）
- Transport protein – Hemoglobin
- Structural proteins – α-keratin, Collagen
- Contractile proteins – Actin（肌动蛋白）, Myosin（肌球蛋白）

Section 2
Protein architecture — Amino acids

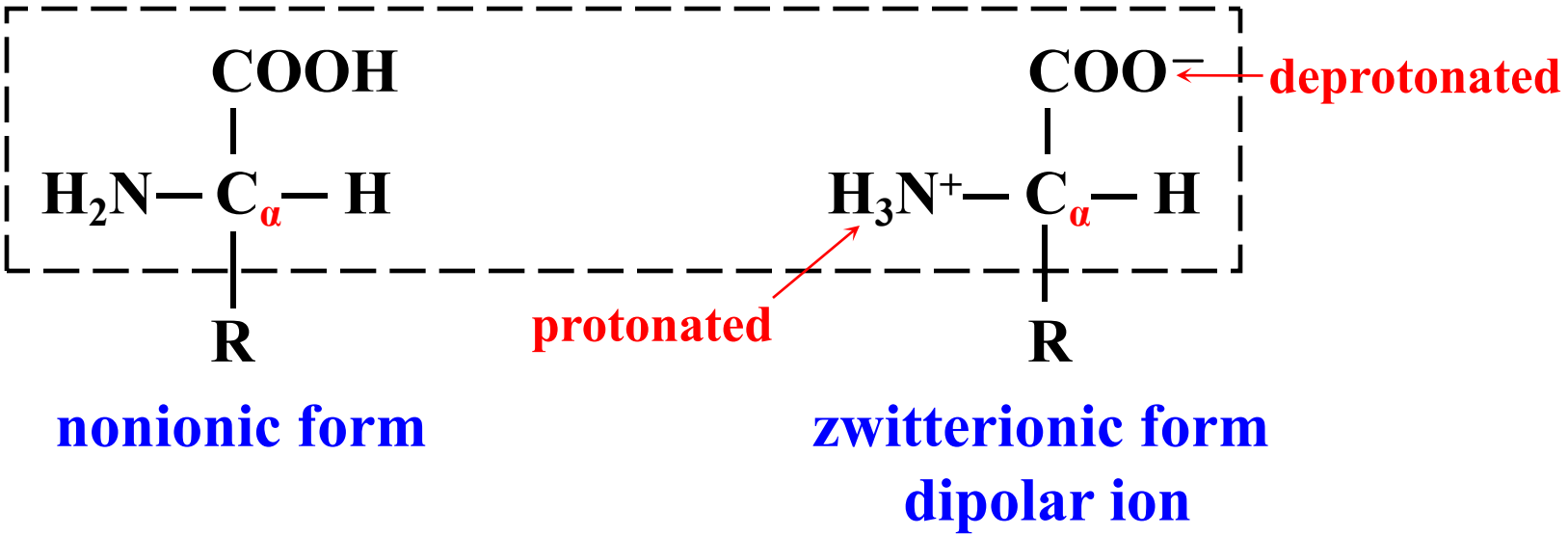
1. General structure of common amino acid (AA)

- The basic building blocks of proteins
- About 300 types of AAs in nature, but only 20 types are used for protein synthesis in biological systems.
- A typical α-amino acid has an amino group, a carboxyl group, a hydrogen atom and a side-chain (R group) attached to the same carbon atom (C_α)



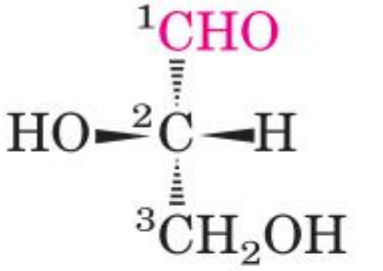
Different side-chain (R group)

Different chemical and physical properties

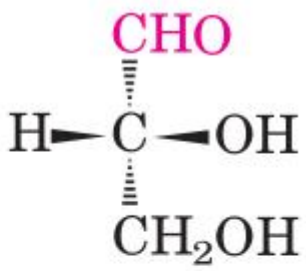


- Except for proline, all 19 of the common AAs are α-amino acids; Proline is an α-imino acids.
- The α-carbon atom is always asymmetric or chiral center except in glycine, hence, all amino acids are optical activity(旋光性) and have two stereoisomers (mirror-image forms) (D or L configuration).

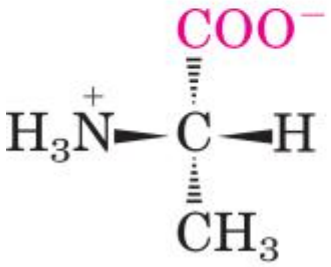
- **The two stereoisomers of each AA are designated by D, L system according to the D- and L-glyceraldehyde**



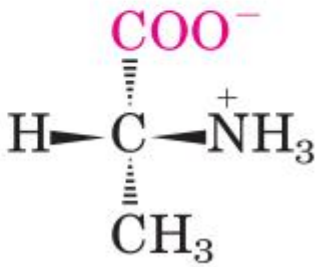
L-Glyceraldehyde



D-Glyceraldehyde



L-Alanine



D-Alanine

- **Only the L-AAs have been found in proteins**
- **D-isomers have been found only in small peptides of bacteria cell walls or in some peptide antibiotics**
- **Racemate(外消旋物): An equimolar mixture of the D- and L-isomers of an optically active compound.**
- **A racemic mixture shows no optical activity.**

2. Amino acid classification

- **Standard amino acids (common AAs, proteinogenic AAs)**
- **Nonstandard amino acids**
- **Non-protein amino acids**

Standard amino acids

- Commonly found in proteins
- Encoded by genetic codes and directly introduced into protein during translation
- Differ in side chain (R group)
- 20 commonly found
- 2 rarely found (selenocysteine and pyrrolysine)

（硒代半胱氨酸，Sec，U）　（吡咯赖氨酸，Pyl，O）

Classification of the standard amino acids

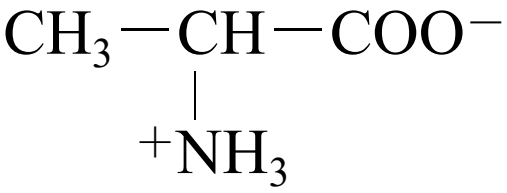
By chemical structure of R groups
- **Aliphatic amino acids: 15**
- **Aromatic amino acids: 3**
- **Heterocyclic amino acids: 3**

By acid-base properties
- **Neutral amino acids: 15**
- **Acidic amino acids: 2**
- **Basic amino acids: 3**

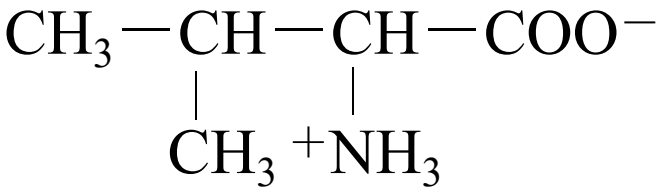
By the polarity of R groups
- **Non-polar amino acids: 9**
- **Polar amino acids: 11**

Non-polar R groups amino acids

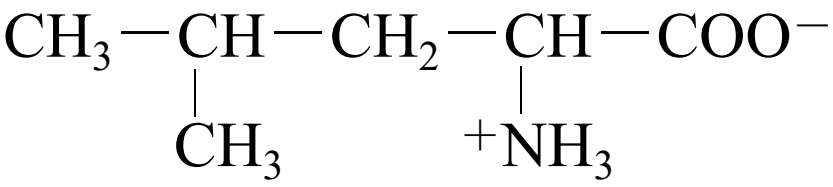
(1)Alanine（Ala，A）



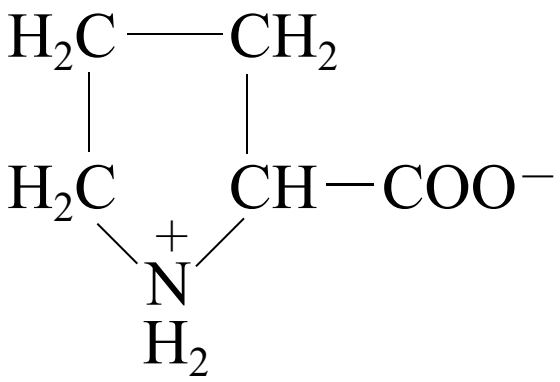
(2) Valine（Val，V）



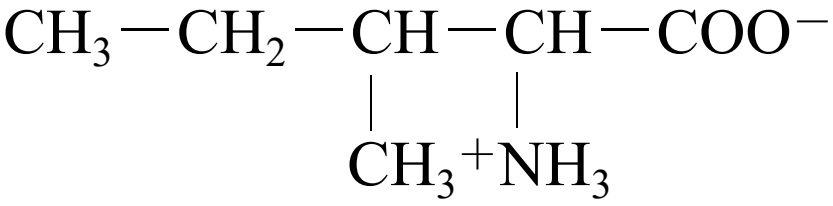
(3)Leucine（Leu，L）



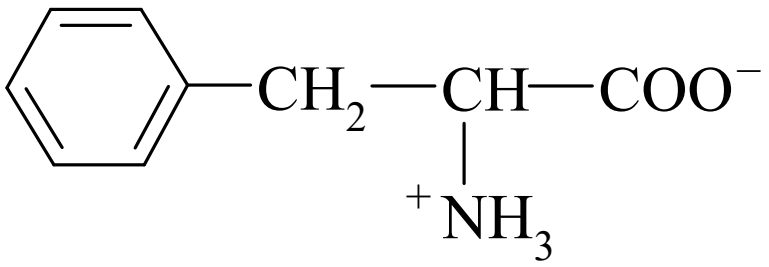
(5)Proline（Pro，P）



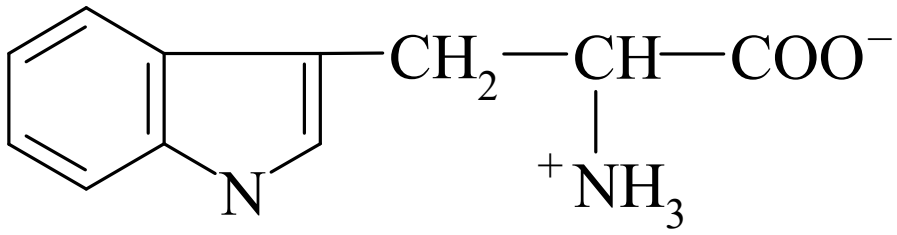
(4)Isoleucine（Ile，I）



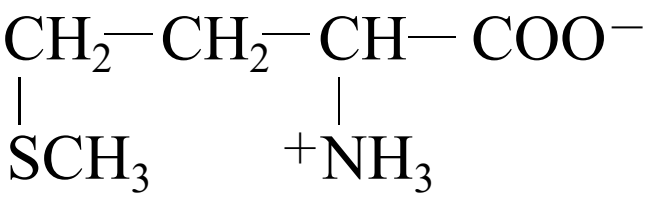
⑹Phenylalanine（Phe，F）



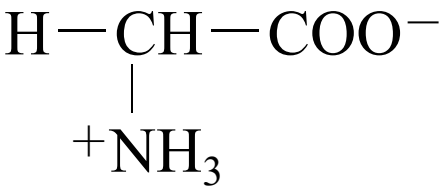
⑺Tryptophan（Trp，W）



⑻Methionine（蛋氨酸，Met，M）

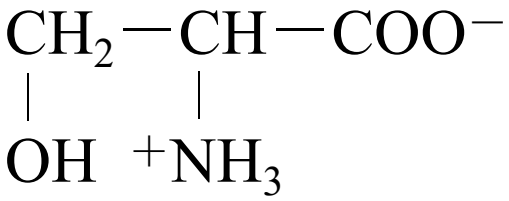


⑼Glycine（Gly，G）

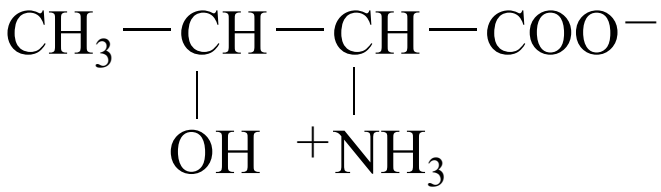


Polar, uncharged R groups amino acids

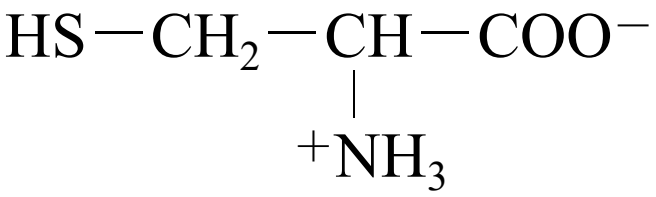
⑩Serine（Ser，S）



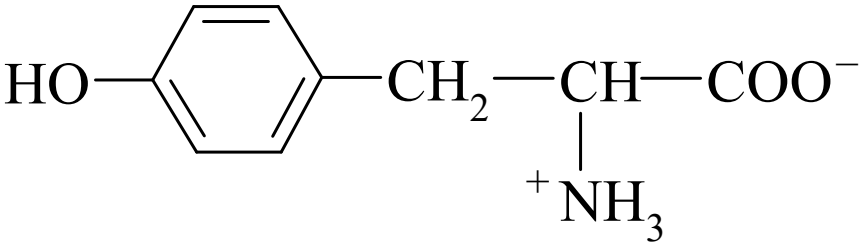
⑪Threonine（Thr，T）



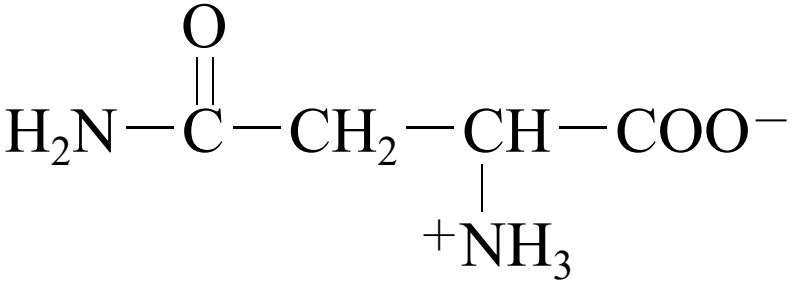
⑫Cysteine（Cys，C）



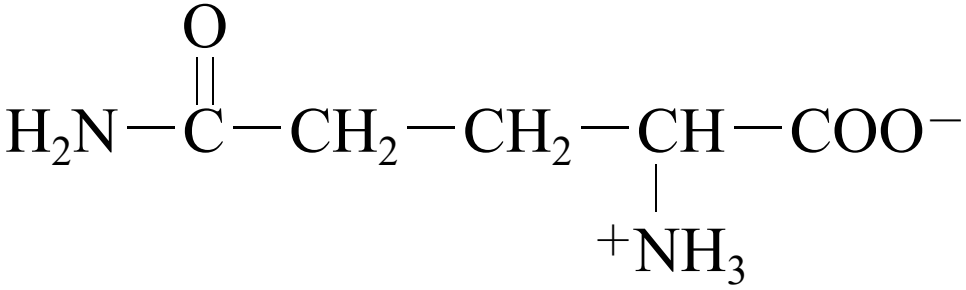
⑬Tyrosine（Tyr，Y）



⑭Asparagine（Asn，N）

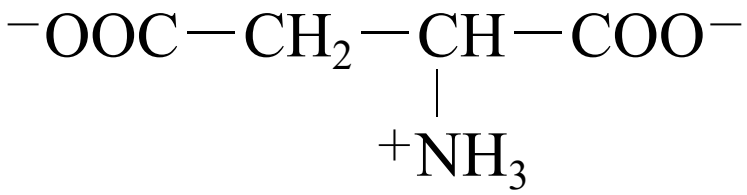


⑮Glutamine（Gln，Q）

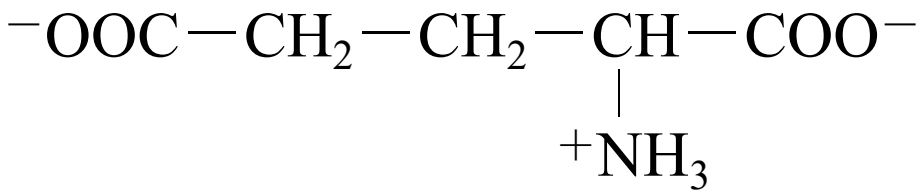


Polar, negatively charged R groups amino acids

⑯Aspartic acid（Asp，D）

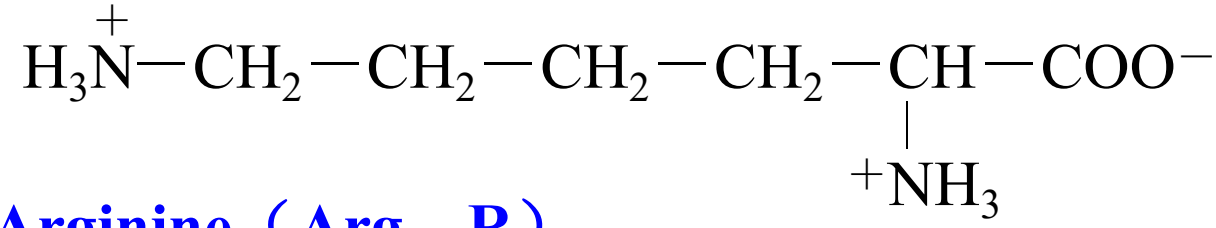


⑰Glutamic acid（Glu，E）

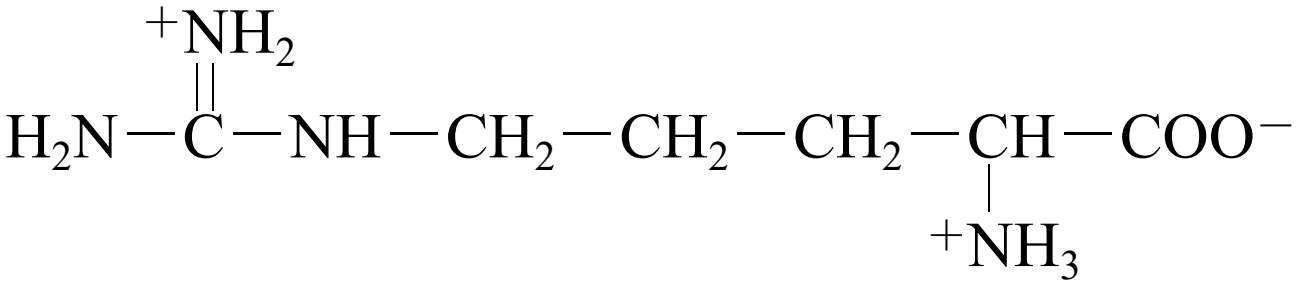


Polar, positively charged R groups amino acids

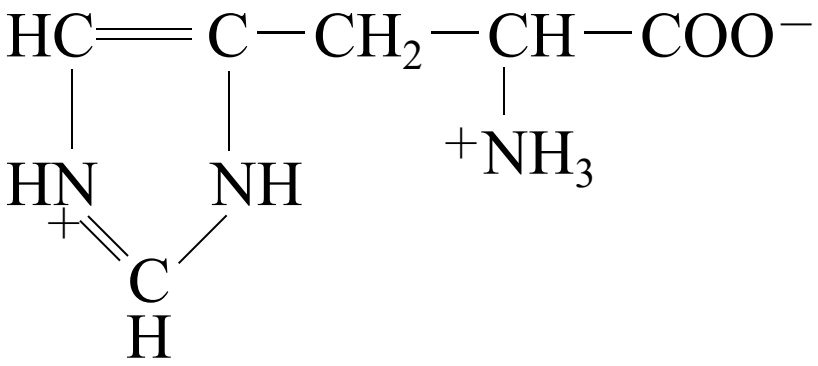
⑱Lysine（Lys，K）



⑲Arginine（Arg，R）



⑳Histidine（His，H）



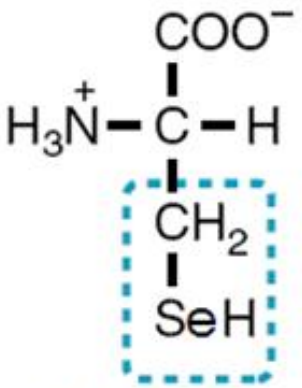
According to whether it can be synthesized { Essential amino acids / Nonessential amino acids

- **Essential amino acids (or indispensable AAs)**
  - **Cannot be synthesized by the humans, must be supplied in the diet**
  - **8: Val, Ile, Leu, Phe, Met, Trp, Thr, Lys**
- **Semi-essential amino acids**
  - **2: His and Arg**
  - **Required by infants and growing children**

Essential and Nonessential Amino Acids in Humans	
Essential	**Nonessential**
Arginine*	Alanine
Histidine*	Asparagine
Isoleucine	Aspartic acid
Leucine	Cysteine
Lysine	Glutamic acid
Methionine	Glutamine
Phenylalanine	Glycine
Threonine	Proline
Tryptophan	Serine
Valine	Tyrosine

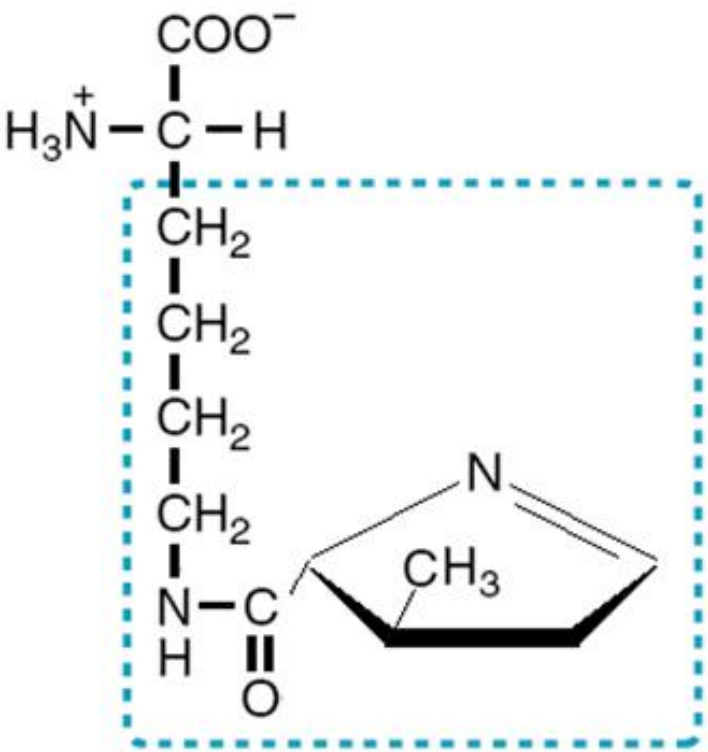
*Arginine and histidine are essential in the diets of juveniles, not adults.

21st & 22nd AAs



Selenocysteine

（Sec，U）



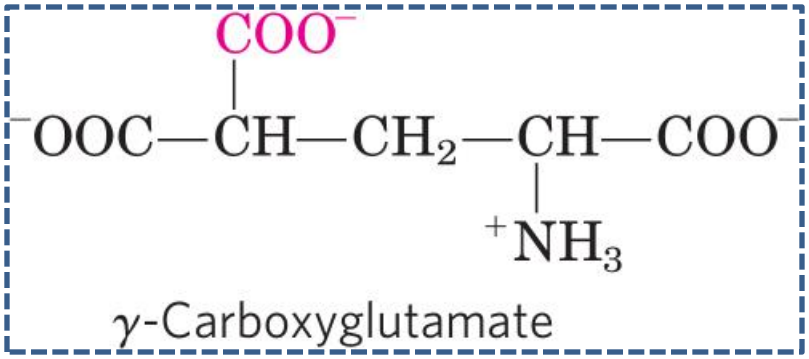
Pyrrolysine

（Pyl，O）

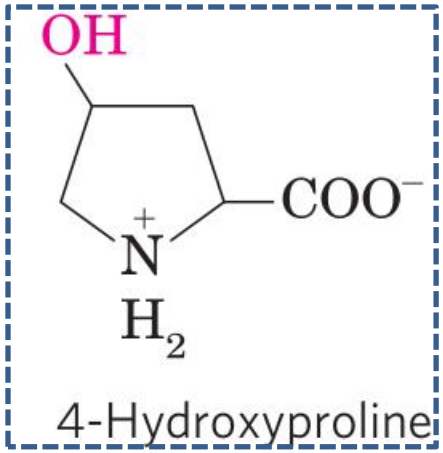
Nonstandard amino acids

- **Occur only rarely in proteins**
- **Amino acid derivatives found in proteins**
- **formed by post-translational modification**

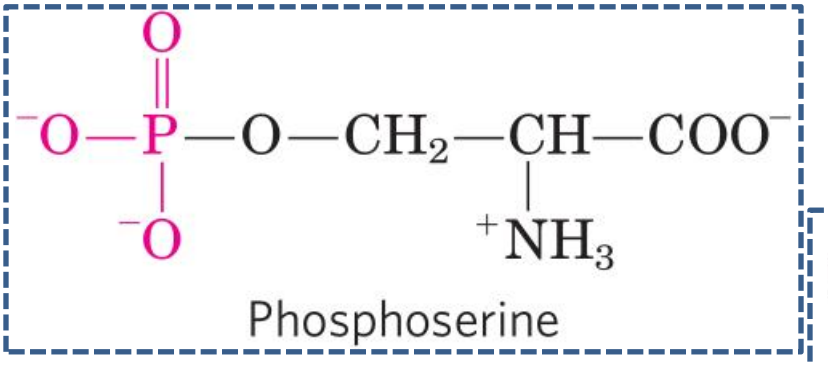
- **γ-Carboxyglutamate (carboxylation of glutamate) found in the blood-clotting protein prothrombin, allows for better binding of calcium cations.**
- **collagen contain 4-hydroxyproline and and 5-hydroxylysine, generated by hydroxylation of proline and lysine respectively.**



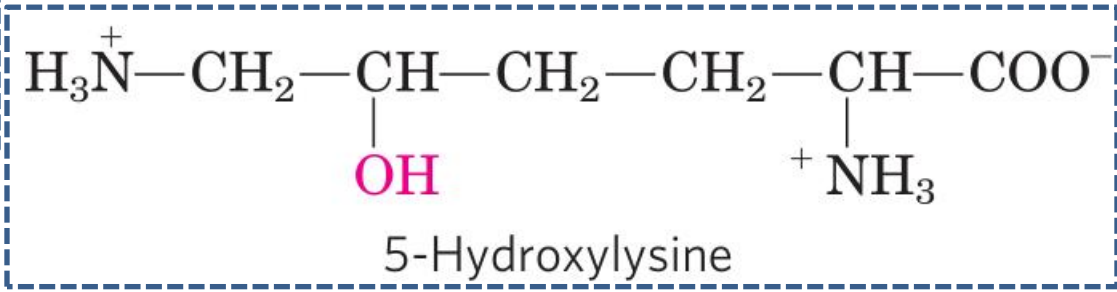
γ-Carboxyglutamate



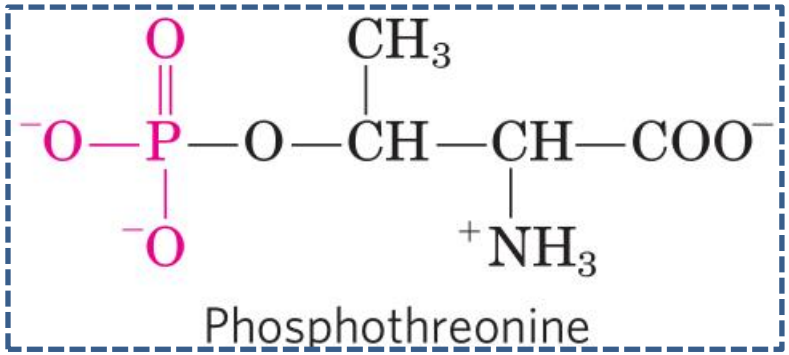
4-Hydroxyproline



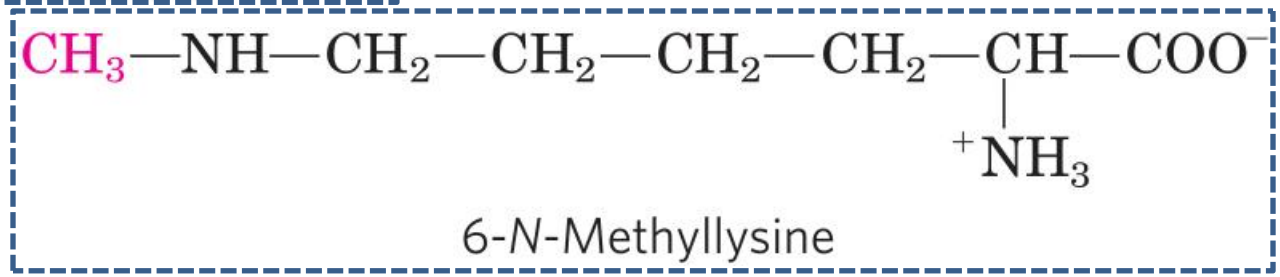
Phosphoserine



5-Hydroxylysine



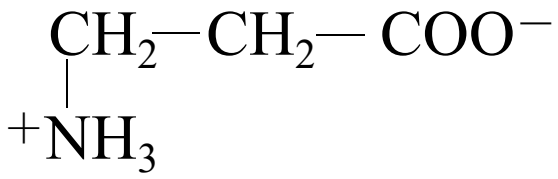
Phosphothreonine



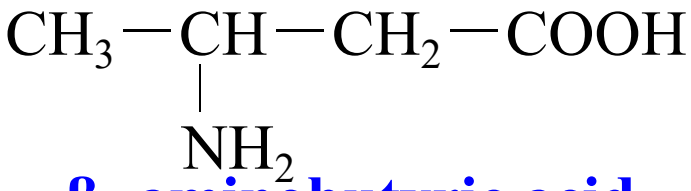
6-*N*-Methyllysine

Non-protein amino acids

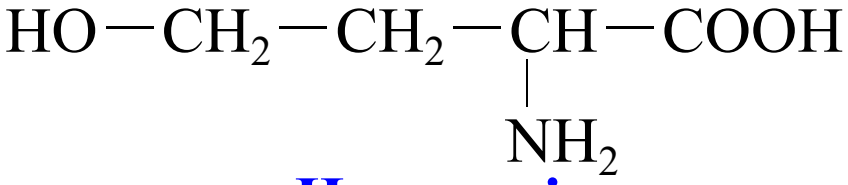
- Present in living organisms, but Not found in proteins
- some are β-AA, γ-AA, or D-AA, etc
- Never directly introduced into proteins during translation
- Can be naturally-occurring or chemical modifications of standard AAs



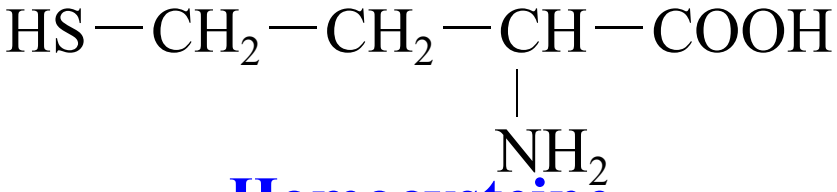
β-Alanine



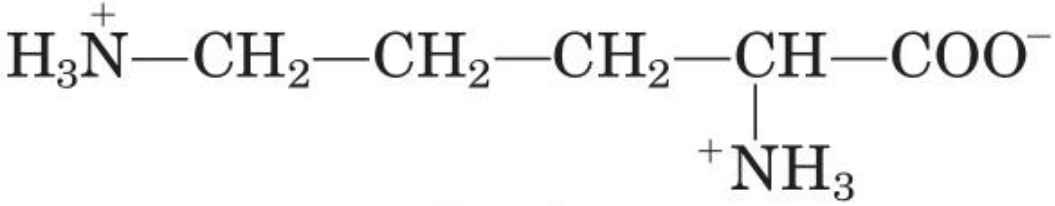
β- aminobutyric acid



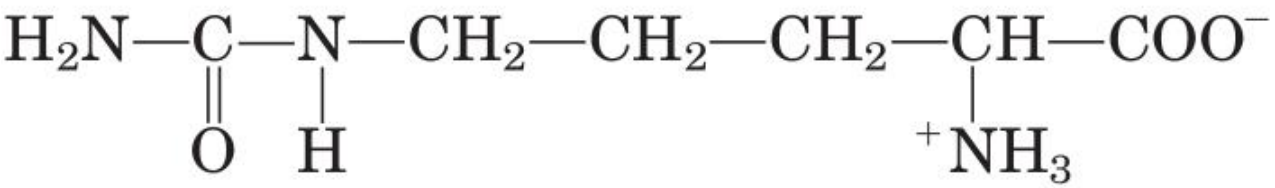
Homoserine



Homocysteine



Ornithine

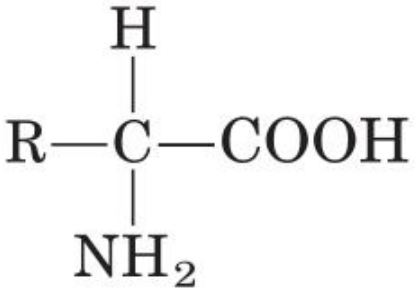


Citrulline

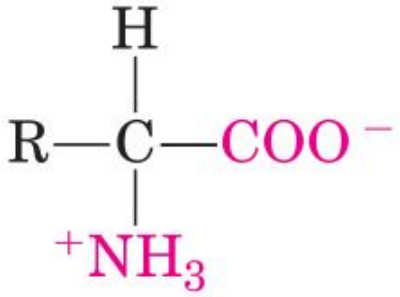
Orn and Cit: Intermediates in AA metabolism

3. Acid-base properties of Amino acids and pI

- **Amino acids has both a basic amine group and an acidic carboxylic acid group**
- **In neutral solution (pH7.0), the amino acid contains a negative charge and a positive charge. It is called a zwitterions(兼性离子) or dipolar ions (偶极离子).**

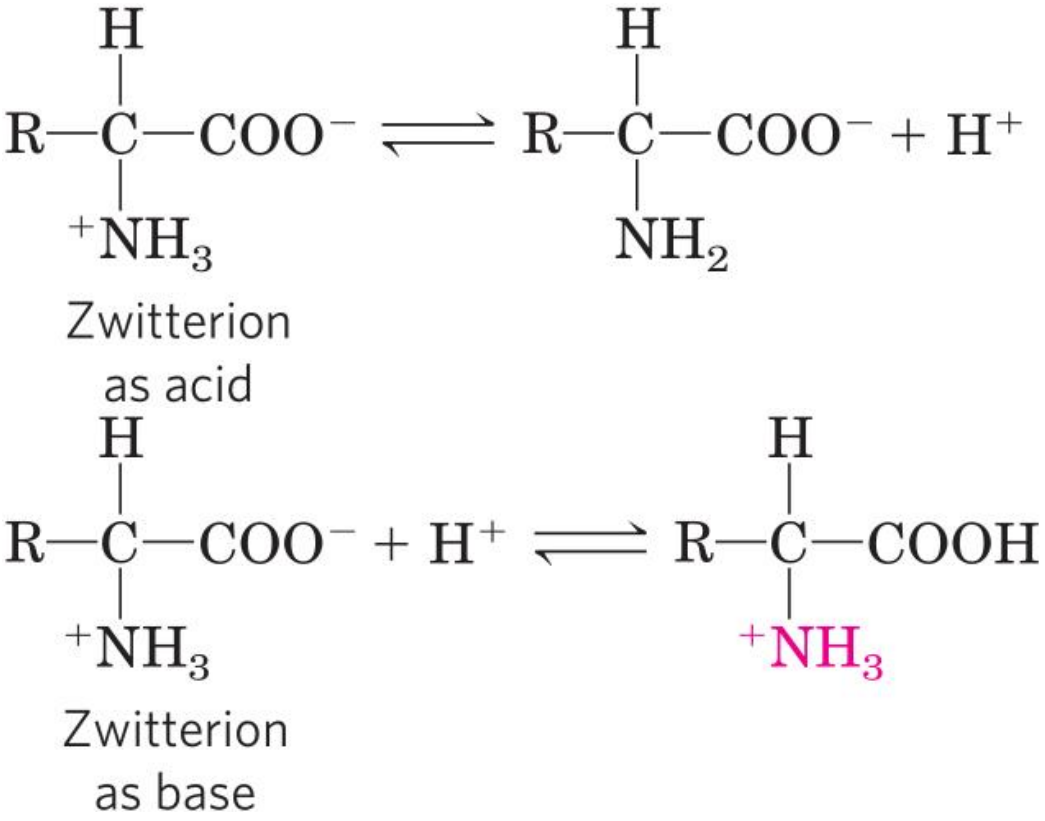


Nonionic form

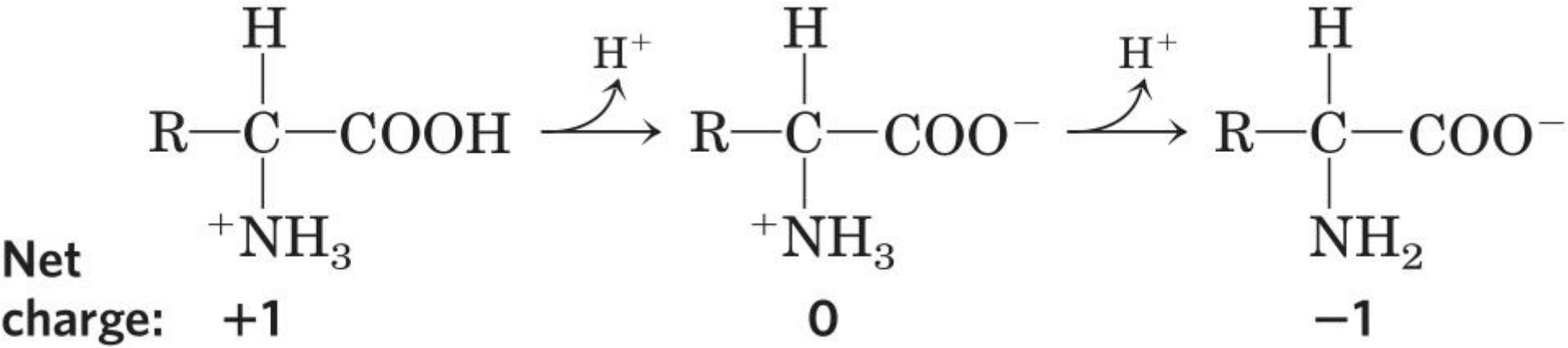


Zwitterionic form

- **Substances having this dual (acid-base) nature are amphoteric 两性的 and are often called ampholytes (两性电解质)**
- **Amino acid zwitterions are amphoteric. They can react as either acids or bases.**

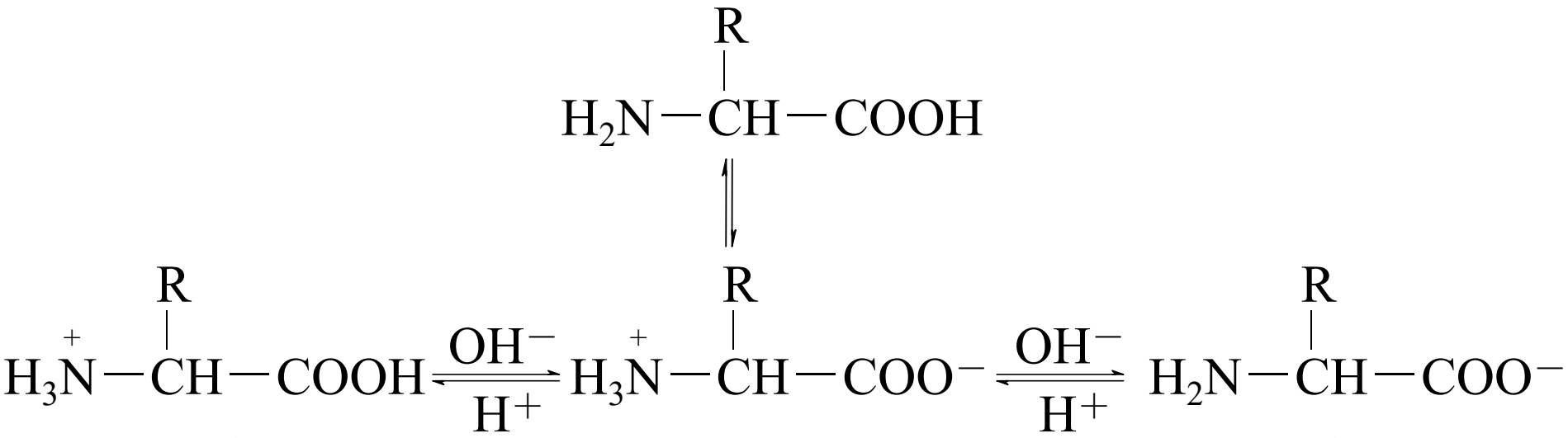


- **AAs are all weak polyprotic acids 多元弱酸**
- **Amino acids have characteristic titration curves (特定的滴定曲线)**



Fully protonated form **Zwitterion** **deprotonated**

- **AAs ionize to various states depending on pH values**
- **Isoelectric Point (pI, 等电点) is the characteristic pH at which an amino acid has equal positive and negative charge (the net electric charge is zero)**
- **AAs in solution at pI are predominantly in dipolar form**
- **pI is determined by pK (K: dissociation constant of the ionizable groups)**



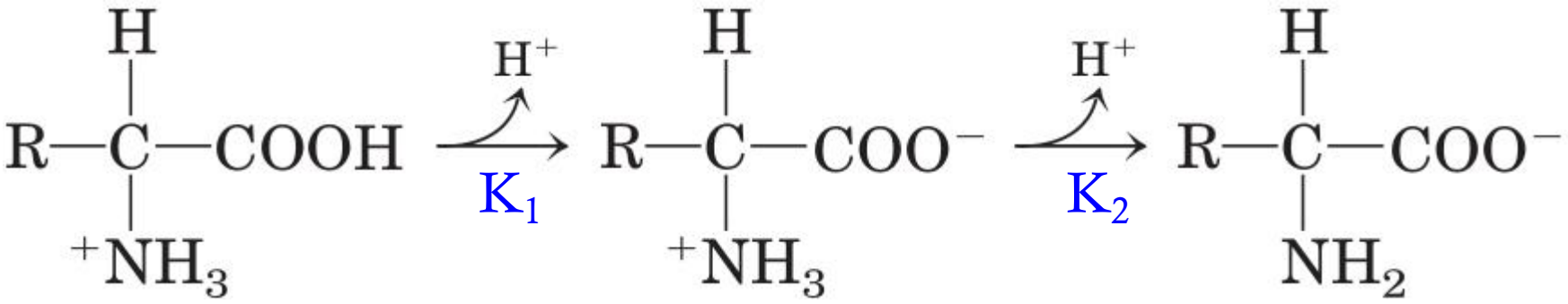
pH＜pI **pH = pI** **pH＞pI**

cation **Zwitterion** **anion**

Move toward the negative electrode **Do not move** **Move toward the positive electrode**

- **At any pH below its pI, AA has a net positive charge and will move toward the negative electrode (the cathode).**
- **At any pH above its pI, AA has a net negative charge and will move toward the positive electrode (the anode).**
- **The farther the pH of a AA solution is from its pI, the greater the net electric charge of the population of AA molecules.**

Calculation of Isoelectric Point



AA^+ **AA^0** **AA^-**

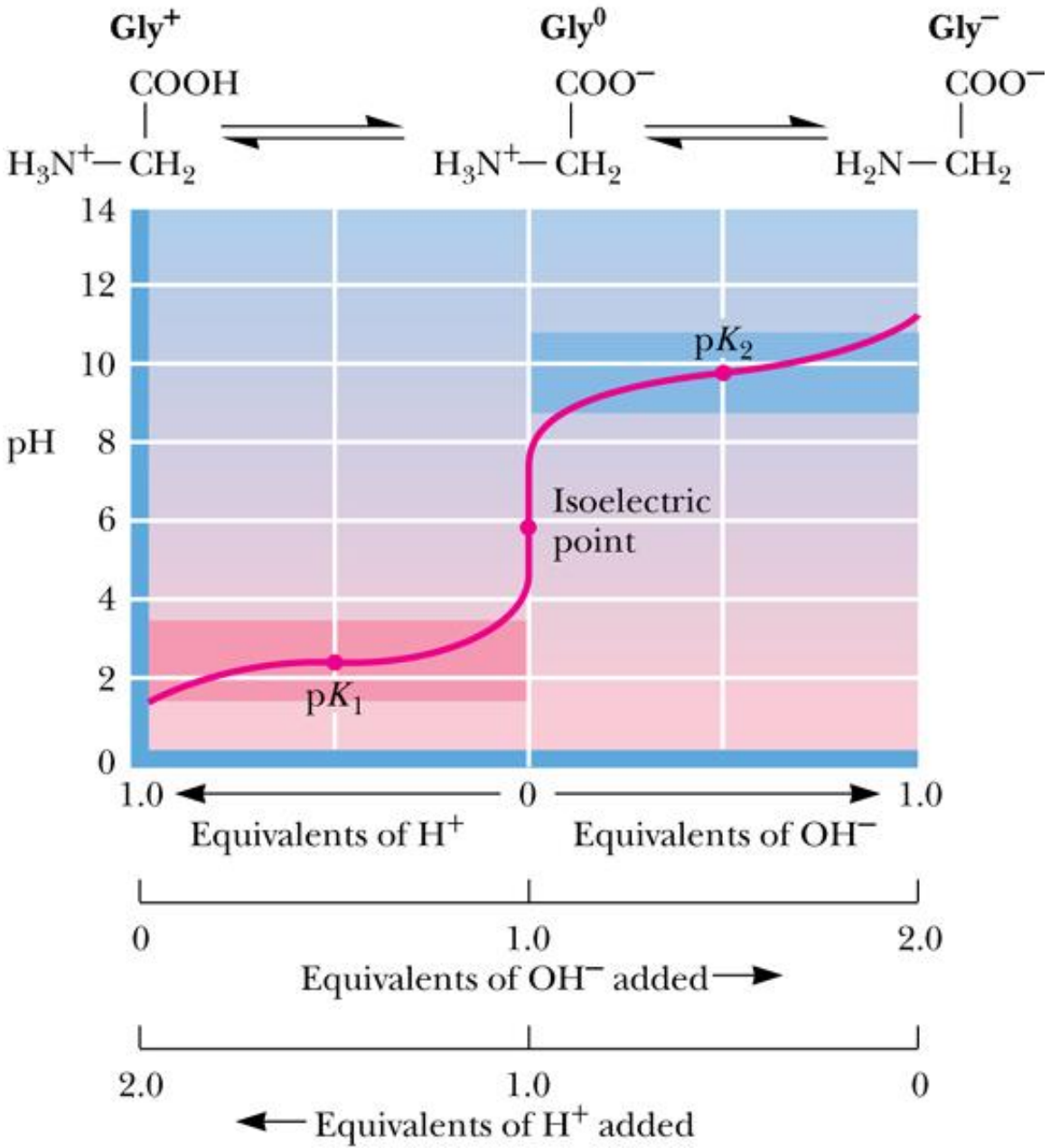
$$K_1=\frac{[H^+]\,[AA^0]}{[AA^+]}\Rightarrow[AA^+]=\frac{[H^+]\,[AA^0]}{K_1}$$

$$K_2=\frac{[H^+]\,[AA^-]}{[AA^0]}\Rightarrow[AA^-]=\frac{K_2\bullet\,[AA^0]}{[H^+]}$$

∵pI时，$[AA^+]=[AA^-]$，$\therefore\ \dfrac{[H^+]\,[AA^0]}{K_1}=\dfrac{K_2\bullet\,[AA^0]}{[H^+]}\Rightarrow[H^+]^2=K_1\bullet K_2$

两边取负对数，得　$pH=(pK_1+pK_2)/2$，此即pI

Titration Curve for Glycine



- **Amino acids with nonionizable R group (R基不电离)**

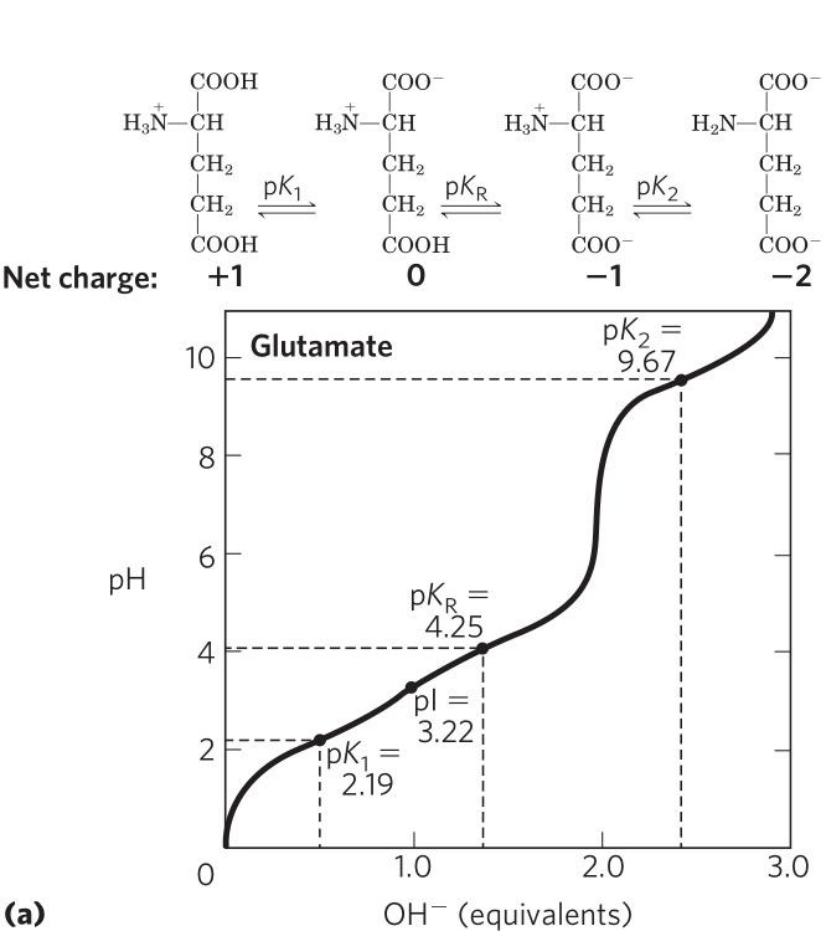
  ——with similar titration curves as that of Gly

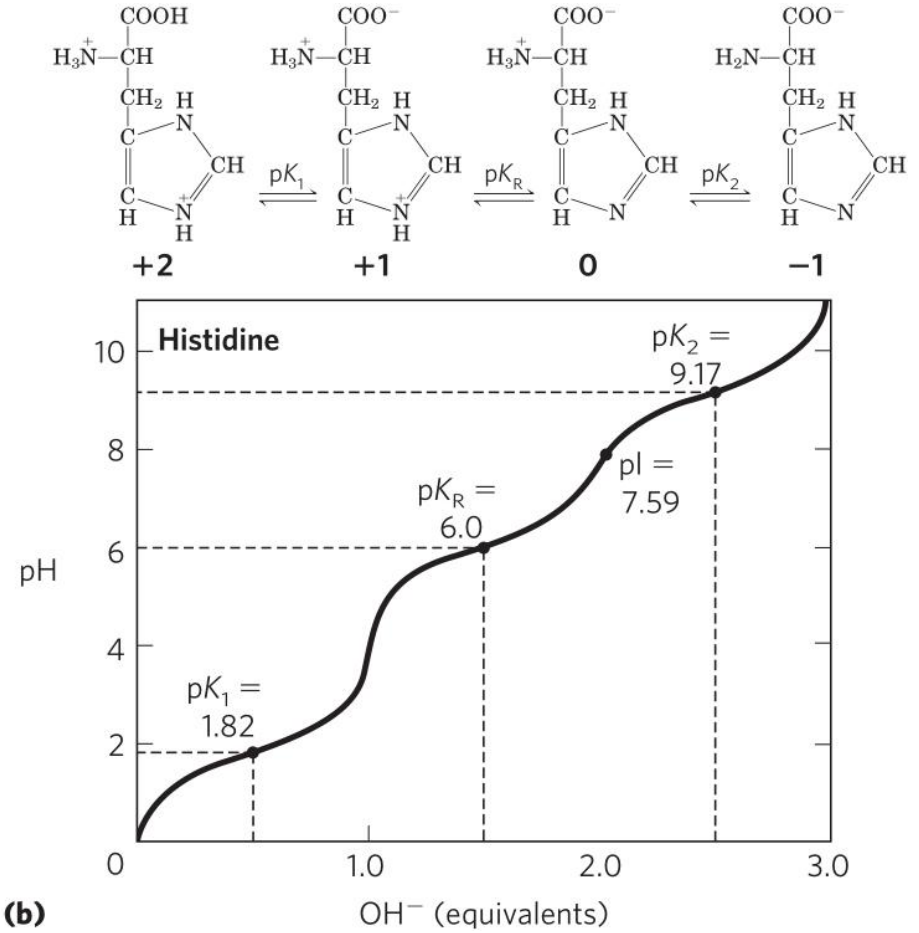
- **Amino acids with ionizable R group (R基电离)**

  ——with more complex titration curves

- **Only His provides buffering power near neutral pH because of the R group (pK_R=6.0) (只有His在中性pH附件有缓冲作用)**

Titration Curve for Glutamic Acid and Histidine





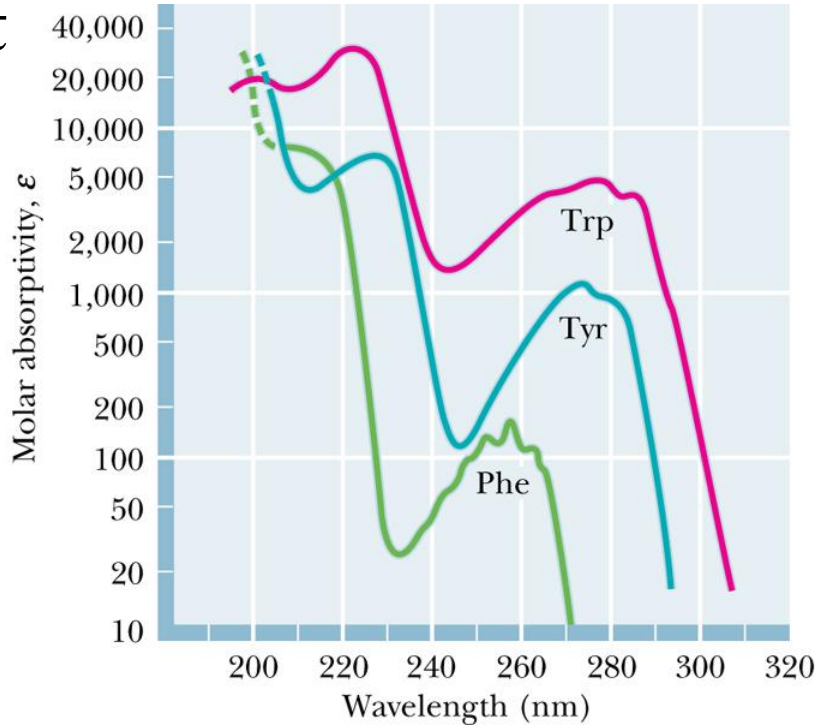
$pI_{Glu} = (pK_1 + pK_R)/2$

$pI_{His} = = (pK_R + pK_2)/2$

4. Spectrophotometric properties of Amino acids

Absorbance of ultraviolet light by aromatic amino acids

AA	λmax	ε(L·mol^{-1}·cm^{-1})
Trp	280nm	5.6×10^3
Tyr	275nm	1.4×10^3
Phe	257nm	2.0×10^2



They are jointly responsible for the light absorption of proteins at 280nm

- **Proteins in solution absorb UV light with absorbance maximum at 280nm**
- **Measuring protein content by photospectrometry**

Absorption of light by molecules: The Lambert-Beer Law

$$A = \lg I_0/I = \lg 1/T = \varepsilon cl$$

The expression $\lg(I_0/I)$ is called the absorbance, designated A
I_0 is the intensity of the incident light
I is the intensity of the transmitted light
I/I_0(the inverse of the ratio in the equation) is the transmittance, T
ε is the molar extinction coefficient (in units of liters per mole-centimeter)
c is the concentration of the absorbing species (in moles per liter)
l is the path length of the light-absorbing sample (in centimeters)

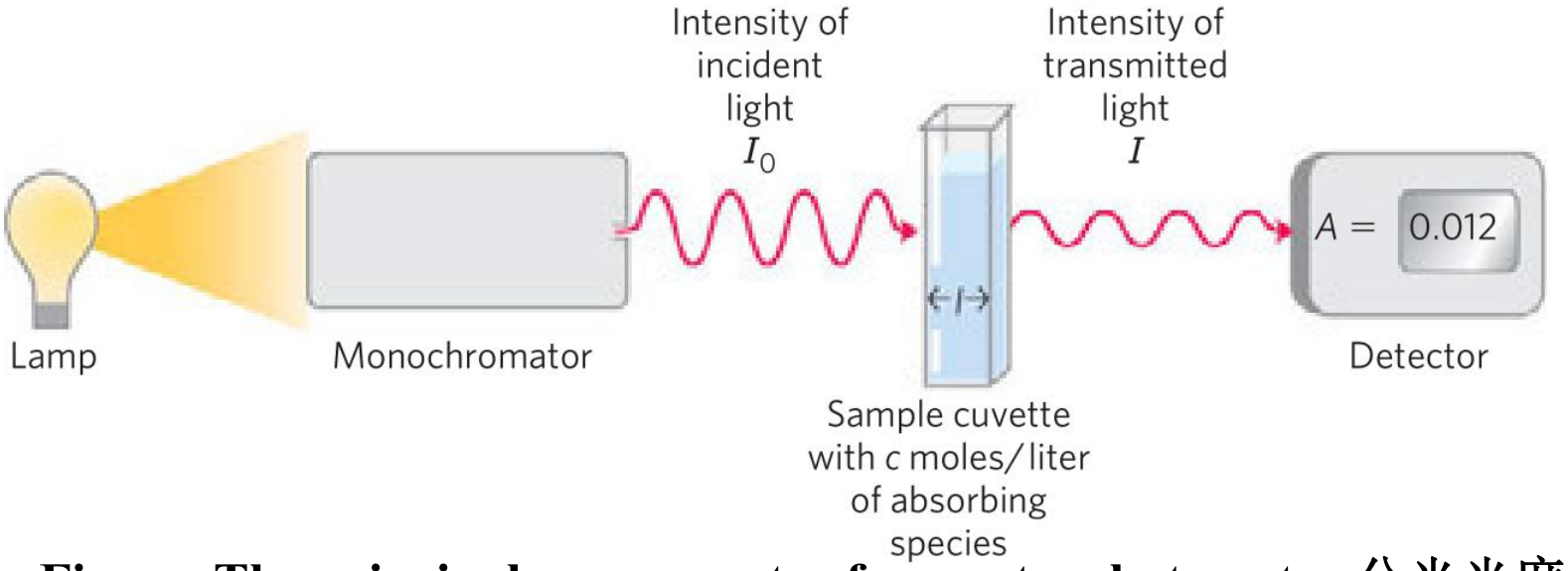


Figure The principal components of a spectrophotometer 分光光度计