

Protein Structure and Function (Slide 5)

§1 Amino acids

1. Proteins: One class of biopolymers with many functions

1° Enzymes (responsible for the metabolic activities)

2° Transport and storage proteins (hemoglobin and myoglobin)

3° Structure proteins (Fibrous proteins)

4° Immune system proteins (antibodies)

5° Receptors and signal transduction proteins

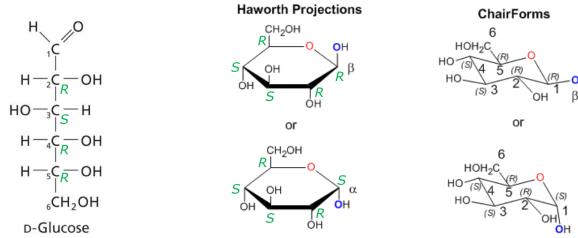
2. Amino acids are the building blocks of biopolymers – proteins

1° Proteins are in fact linear (i.e. unbranched) polymers of amino acids

2° All protein can be hydrolyzed into amino acids.

*: DNA (or RNA) is a linear polymer of nucleotides.

Complex carbohydrates are either linear (i.e. cellulose) or branched polymers (i.e. starch) of simple sugars (monosaccharides)



3. Basic information about amino acids

1° In the structural formula, the symbol R means one side chain.

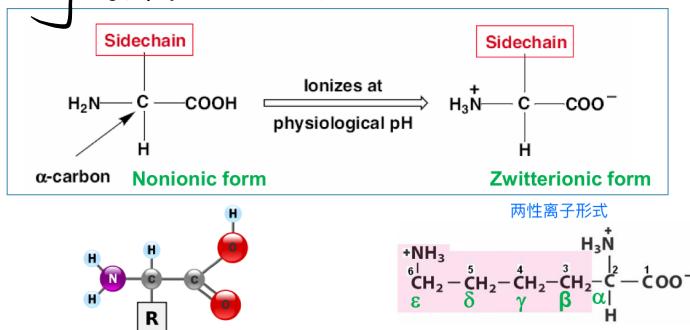
2° In standard amino acids, all have primary amino group and one carboxylic acid bounded to the same carbon, C_{α}

(Exception: proline (脯氨酸))

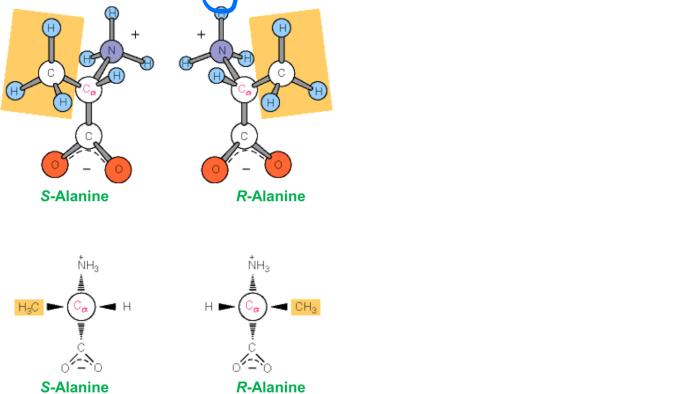
3° At neutral pH, amino acids exist as an internal salt (zwitterion)

This imparts special properties:

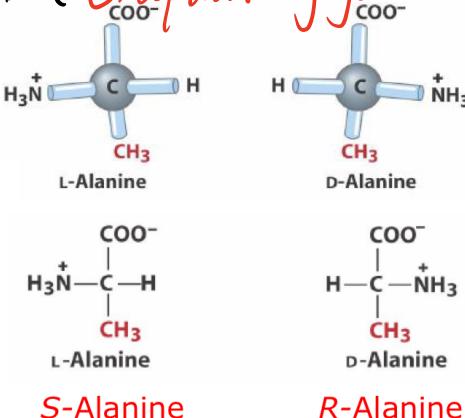
- ① High melting points ($\sim 300^\circ\text{C}$)
- ② Very soluble in water.



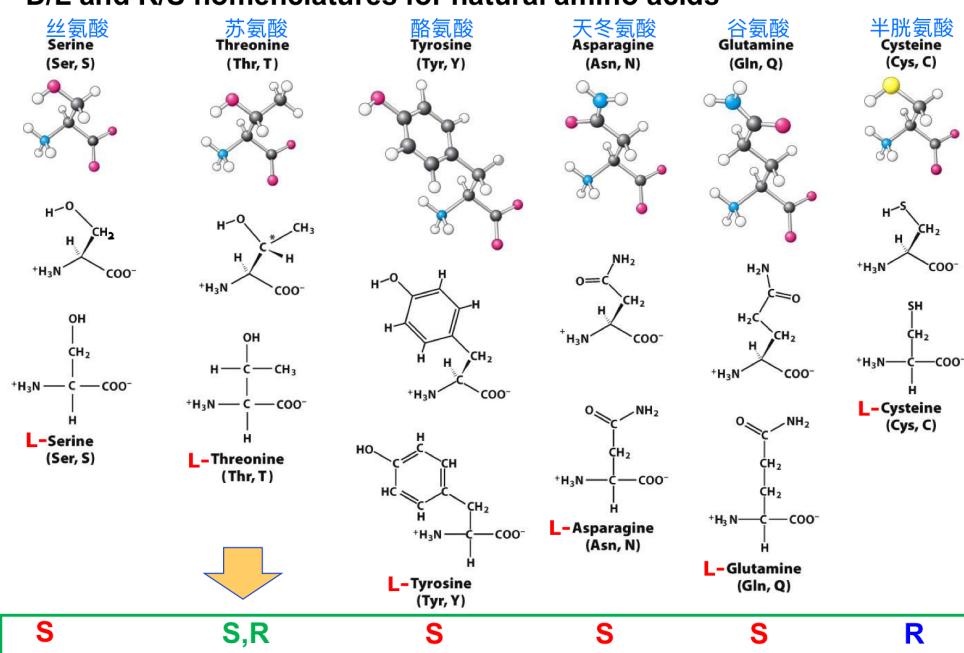
4. α carbon chirality



- 1° Natural carbohydrate: D-isomer
- 2° Natural amino acid α -carbon is usually a chiral center
(Exception: glycine (甘氨酸) (doesn't have chiral center))
- 3° All natural amino acids are defined as L-amino acids
(Exception: glycine)
- 4° Almost all natural amino acid α carbons have the S configuration.
(Exception: glycine and cysteine (半胱氨酸))



D/L and R/S nomenclatures for natural amino acids



5. Classification of amino acids

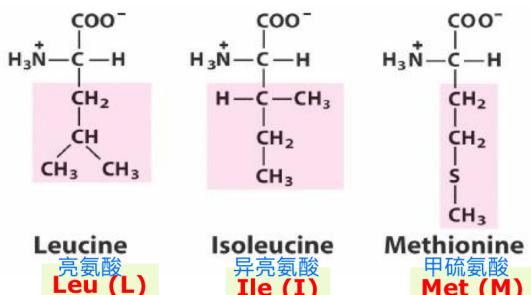
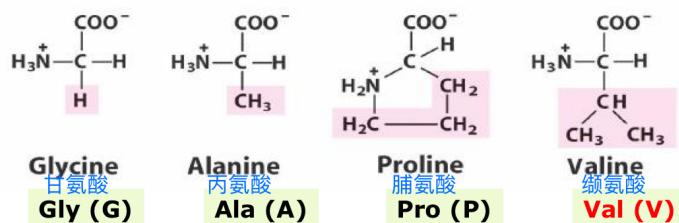
1° Side chains of amino acids are responsible for many of the unique properties of proteins.

2° 3 major classes of side chains:

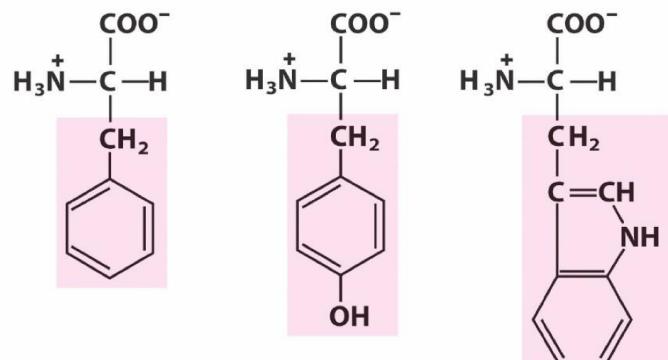
- { Nonpolar (hydrophobic)
- Polar
- Charged

3° The 20 different amino acids are commonly abbreviated with 1- or 3-letter codes.

脂肪族 Nonpolar, aliphatic R groups



Nonpolar aromatic R groups



Phenylalanine

苯丙氨酸

Phe (F)

Tyrosine

酪氨酸

Tyr (Y)

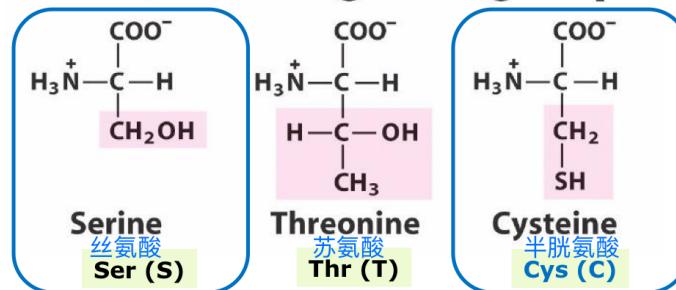
Tryptophan

色氨酸

Trp (W)

considered as non-polar

Polar, uncharged R groups



Serine

丝氨酸

Ser (S)

Threonine

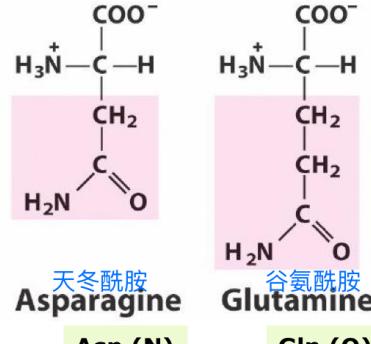
苏氨酸

Thr (T)

Cysteine

半胱氨酸

Cys (C)



Asparagine

天冬酰胺

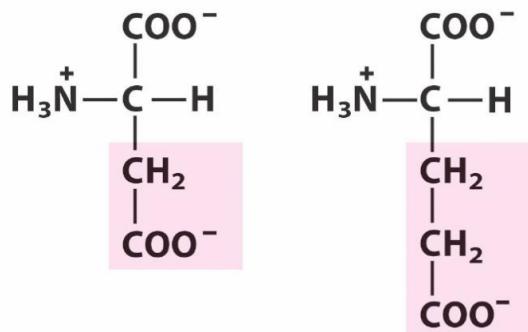
Asn (N)

Glutamine

谷氨酰胺

Gln (Q)

Negatively charged R groups



Aspartate

Asp (D)

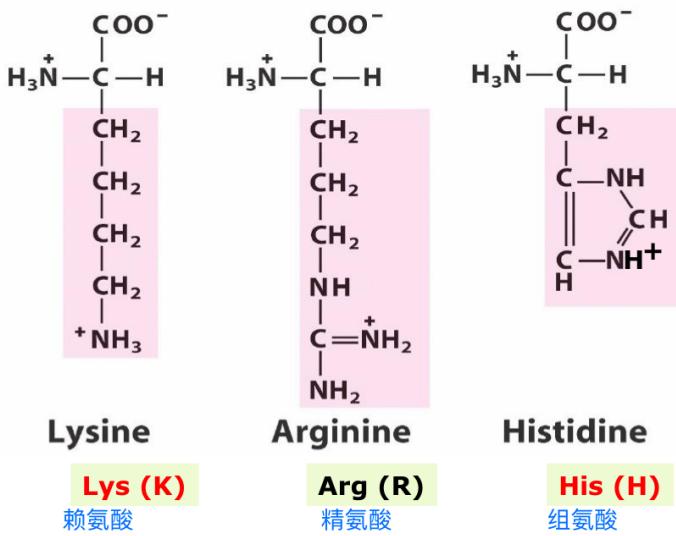
天冬氨酸

Glutamate

Glu (E)

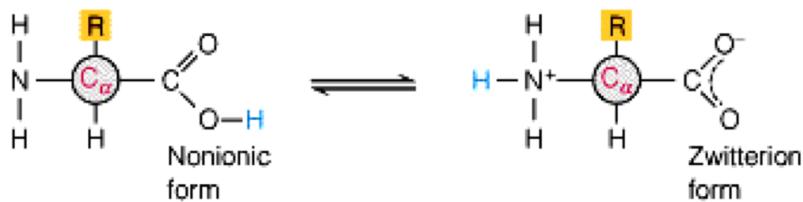
谷氨酸

Positively charged R groups



b. Titration of Amino Acids

- 1° At physiological pH, the carboxyl group is negatively charged and amino group is positively charged.
 - 2° Amino acids **without** charged side chains are **zwitterions** and have no net charge.
 - 3° The pH at this point is called **isoelectric point (pI)**



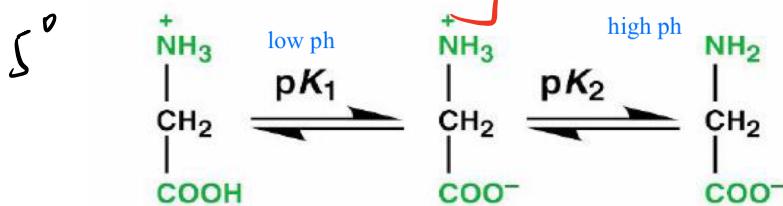
- 4° VS: Natural pH, neutral charge and isoelectric point

 - ① Natural pH:
pH 7. For **whole solution** with $[H^+] = 10^{-7} M$
 - ② Neutral charge:
0 charge. For **one specific molecule**, the positively charged groups are exactly balanced by the negatively charged groups.
 - ③ The isoelectric point, pl:

is the pH of an aqueous solution of an amino acid (or peptide) at which the molecules on average have no net charge.

In other words, the positively charged groups are exactly balanced by the negatively charged groups.

At **isoelectric point**, amino acids have the **lowest solubility**.



Henderson-Hasselbalch equation: $\text{pH} = \text{p}K_a + \log \frac{[\text{A}^-]}{[\text{HA}]}$

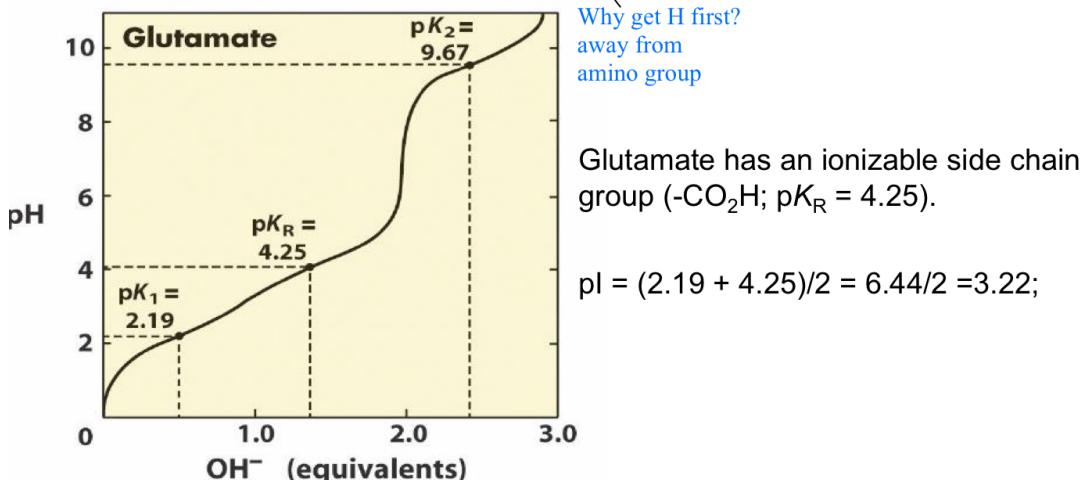
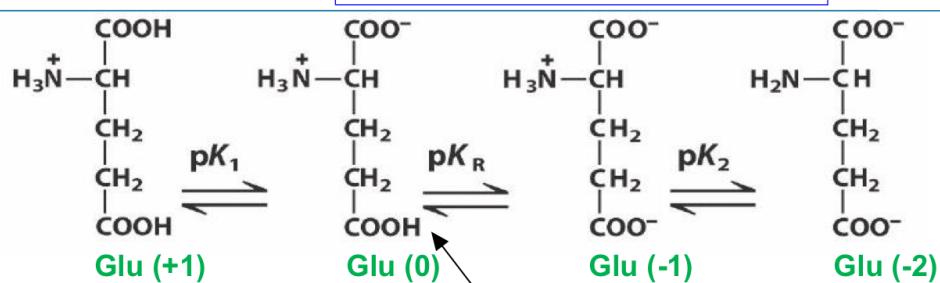
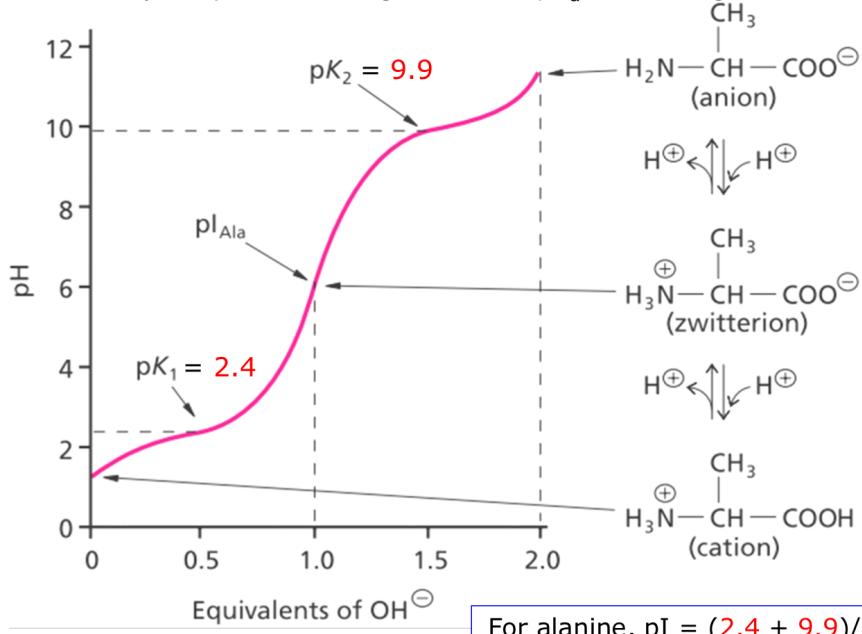
$\text{pH} > \text{p}K_a$ A^- predominates

$\text{pH} < \text{p}K_a$ HA predominates

6° How to calculate pI :

- ① Identify all ionizable groups
- ② Assign $\text{p}K_a$ to each ionizable group
- ③ Start with each ionizable group in protonated form (very low pH — may be 0 or 1) and calculate its net charge.
- ④ Slowly move up in pH to the first ionizable group's $\text{p}K_a$ and deprotonate it (reduce charge by 1)
- ⑤ Do this until each group is deprotonated. Identify all charged forms and at which pH each transition occurs.

Generally, the pI is the average of the two $\text{p}K_a$'s bracketing the isoelectric structure.



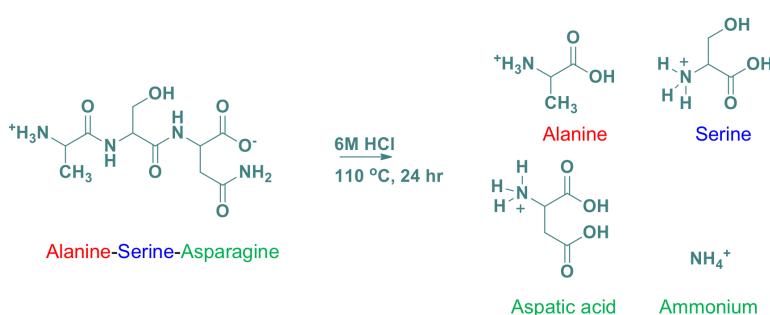
Multiple ionizable groups with both acidic and basic $\text{p}K_a$ values.

Amino acid	$\text{p}K_{\text{a}1}$ (COOH)	$\text{p}K_{\text{a}2}$ (NH ₂)	$\text{p}K_R$ (sidechain)	pI
Glycine	2.34	9.60	—	5.97
Alanine	2.33	9.69	—	6.01
Valine	2.32	9.62	—	5.97
Leucine	2.36	9.60	—	5.98
Isoleucine	2.36	9.68	—	6.02
Proline	2.00	10.96	—	6.48
Phenylalanine	1.83	9.13	—	5.48
Tyrosine	2.21	9.11	10.07	5.66
Tryptophan	2.39	9.39	—	5.89

Amino acid	pK_{a1} (COOH)	pK_{a2} (NH ₂)	pK_R (sidechain)	pI
Serine	2.21	9.15	13.60	5.68
Threonine	2.12	9.62	13.60	5.87
Cysteine	1.96	10.28	8.18	5.07
Methionine	2.27	9.21	—	5.74
Asparagine	2.02	8.80	—	5.41
Glutamine	2.17	9.13	—	5.65
Aspartic acid	1.89	9.60	3.65	2.77
Glutamic acid	2.19	9.67	4.25	3.22
Lysine	2.2	8.95	10.53	9.74
Arginine	1.8	9.04	12.48	10.76
Histidine	1.8	9.17	6.00	7.59

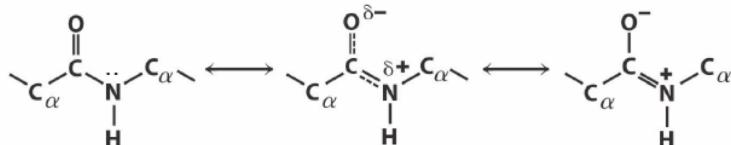
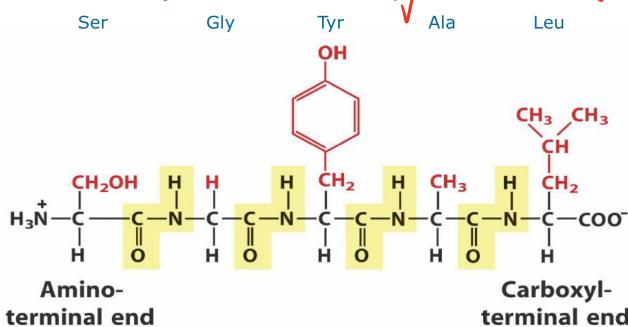
§2 Peptides

1. Peptide can be hydrolyzed to constituent amino acids at a harsh condition



2. Geometry of peptide bond

- 1° The peptide bond (C-N) has partial double bond character due to the resonance (共轭) contribution
- 2° The -CONH- is rigid and planar. This limits orientations available to the polypeptide.
- 3° The carbonyl oxygen has partial negative charge and the amide nitrogen a partial positive charge
- 4° Virtually all peptide bonds in proteins occur in the trans conformation.

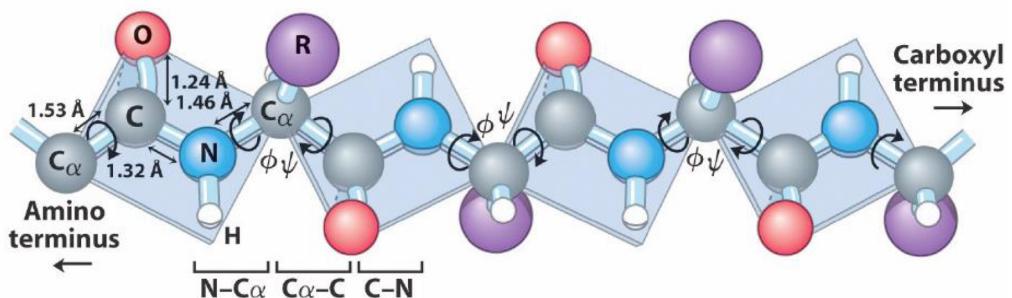
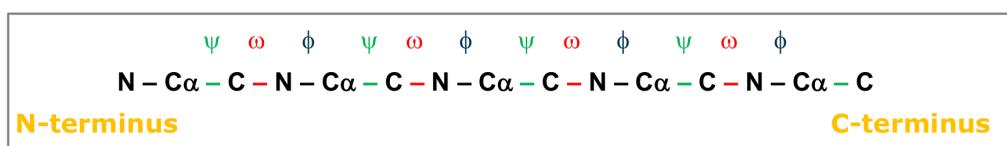


3. Three dihedral angles

$$\phi: \text{C} - \text{N} - \text{C}\alpha - \text{C}$$

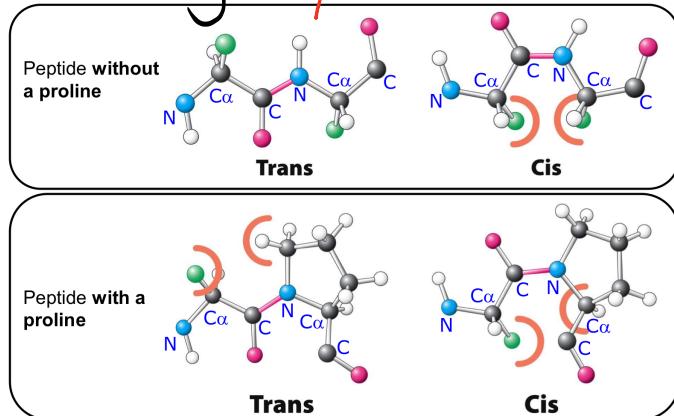
$$\Psi: \text{N} - \text{C}\alpha - \text{C} - \text{N}$$

ω : C α – C – N – C α



For a peptide bond, the *cis* conformation is usually much less favorable than the more extended *trans* conformation because of *steric interference between the side chains* attached to the two α -carbon atoms.

* both cis and trans conformations are populated in a peptide bond involving a proline residue.

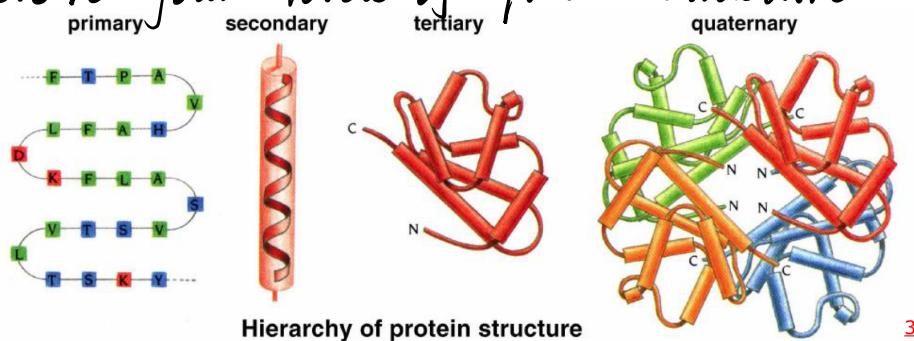


§3 Protein Structures

Protein generally have more than 50 amino acid residues.

There are many possible ways to arrange the residues in 3-dimensional space. Most proteins fold into a single stable shape known as its **native conformation**.

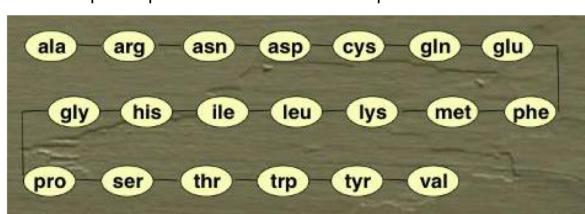
There're four levels of protein structure



1. Primary structure of proteins

- 1^o Primary structure is the **sequence** of amino acids.
- 2^o The amino acid sequence determines the 3-D-structure (fold) of a protein and ultimately specifies its **function**
- 3^o Forces involved in primary structure are very strong covalent bonds: **peptide bonds**

An example sequence of amino acids in a protein:



2. Secondary structure of proteins

- 1^o Primary structure determines the secondary structure.
- 2^o Long chains of amino acids commonly **fold** or **curl** into a regular repeating structure.
- 3^o Structure is the result of **hydrogen bonding** and

other non-covalent interactions among amino acids within the protein.

4^o Common secondary structures are

α -helix (α 螺旋)

β -pleated sheet (β 折叠)

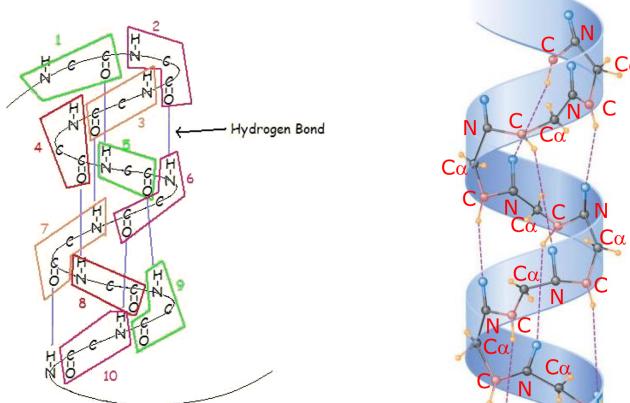
5^o Secondary structures add new properties to a protein
e.g. strength, flexibility

6^o α -helix

① α -helix is rigid, rod-like structure with a right-handed helical conformation

② Each carbonyl oxygen (residue n) of the backbone is hydrogen-bonded to the backbone amide hydrogen (residue $n+4$)

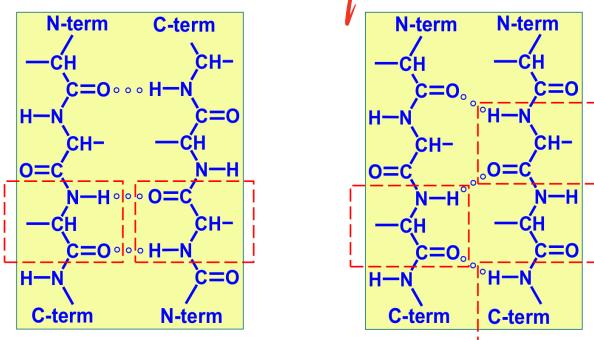
③ There are 3.6 amino acid residues per turn and the pitch is 0.54 nm



7^o β -pleated sheet

β -pleated sheet is held by hydrogen bonding between adjacent sheets of protein.

There're two kinds of β -pleated sheet:
parallel and antiparallel



Antiparallel sheet

Parallel sheet

8° Supersecondary structures

Many proteins contain combinations of α -helix and β -pleated sheets and other secondary structures.

Forces involved in secondary structures are hydrogen bonds and other non-covalent interactions.



3. Tertiary structure of proteins

1° Arrange the secondary structures in relation to each other.

2° Proteins fold into specific structures.

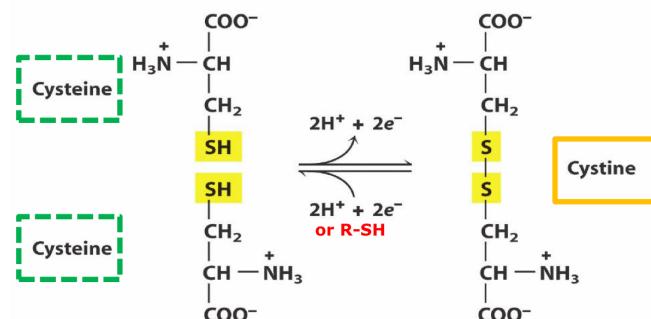
Forces involved in tertiary structures are:

- { Hydrogen bonds
- Electrostatic interactions (静电相互作用)
- Metal ion coordination (金属离子配位键)
- Hydrophobic effect (疏水效应)

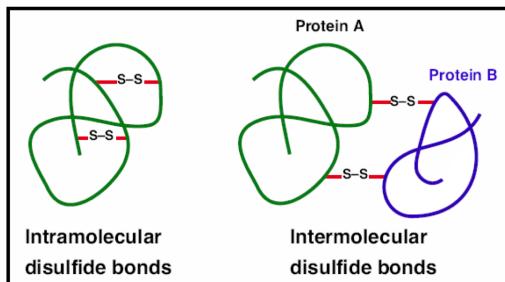
3° Disulfide bonds in proteins

① Disulfide bond formation is an oxidation reaction

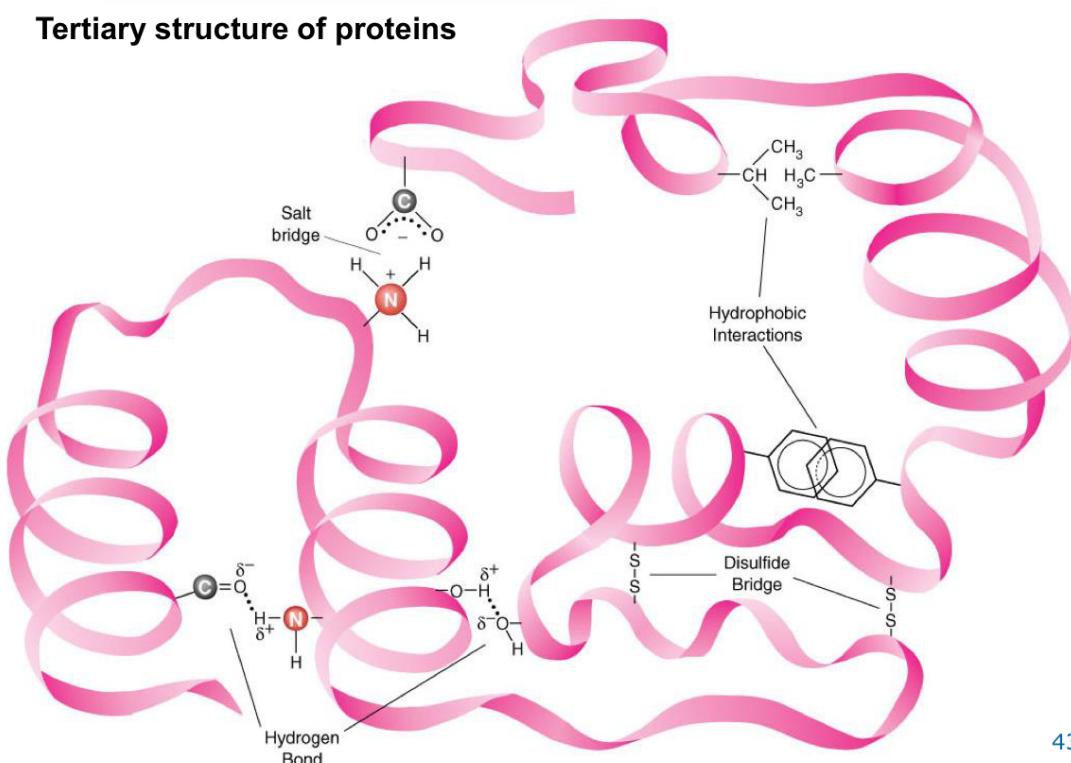
- ② Intracellular conditions are maintained sufficiently reducing to inhibit formation of most disulfide bonds.
- ③ Extracellular conditions are more oxidizing, favoring disulfide formation.
- ④ Thus, **extracellular** proteins containing cysteines often have $-S-S-$ bonds, while **intracellular** proteins rarely have.
- ⑤ $-S-S-$ bonds can be cleaved by **R-SH**



Disulfide bonds in proteins



Tertiary structure of proteins



4. Quaternary structure of proteins

1° Many proteins are not single peptide strands.

Two or more subunits are held together by non-covalent interactions to form quaternary structures.

2° Forces involved in quaternary structures are:


Hydrogen bonds
Electrostatic interactions (静电相互作用)
Metal ion coordination (金属离子配位键)
Hydrophobic effect (疏水效应)
.....

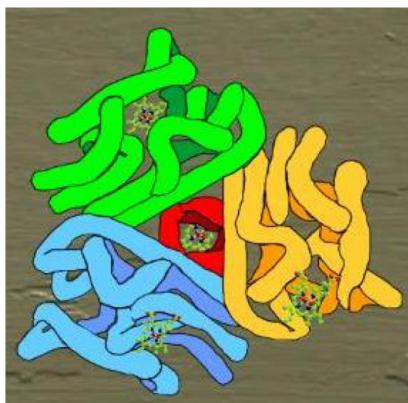
5. The hierarchical protein structure organization

1° Primary structure: Amino acid sequence

2° Secondary structure: Local regions with defined fold
(α -helix, β -sheet, etc)

3° Tertiary structure: How the locally folded structures arrange in 3D

4° Quaternary structure: Packing of several folded polypeptide chains



b. Getting protein structures

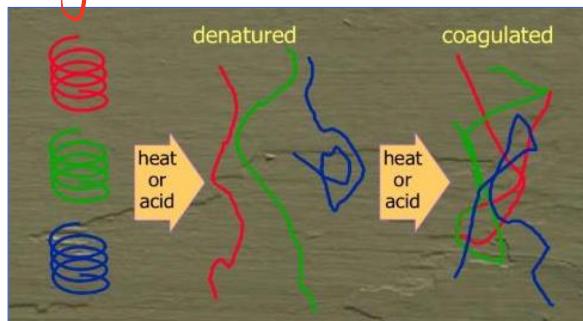
1° Prediction

using { Known structure
Tendencies of amino acids
Energy minimization

2° Experimental determination
{ X-ray crystallography
NMR

7. Effects of temperature and pH on proteins.

If go beyond a "normal" range
{ denatured (变性)
coagulate (凝固)



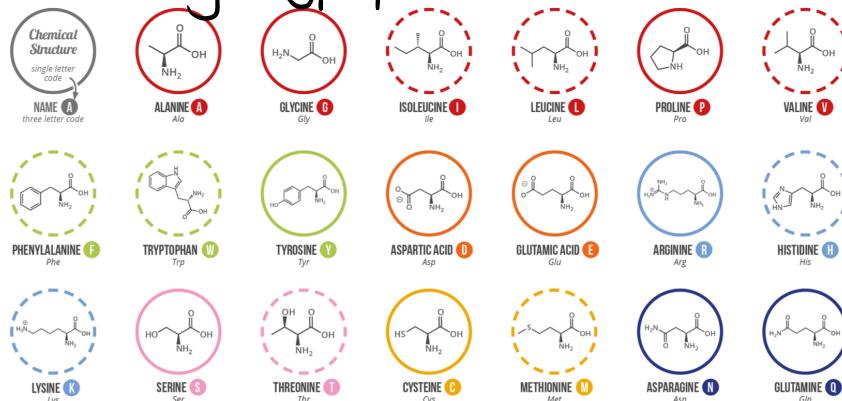
8. Measure/monitor protein unfolding

1° CD (Circular Dichroism)

Measuring content of regular structures such as α -helix and β sheet.

2° Trp fluorescence

Measuring tryptophan local environment changes.

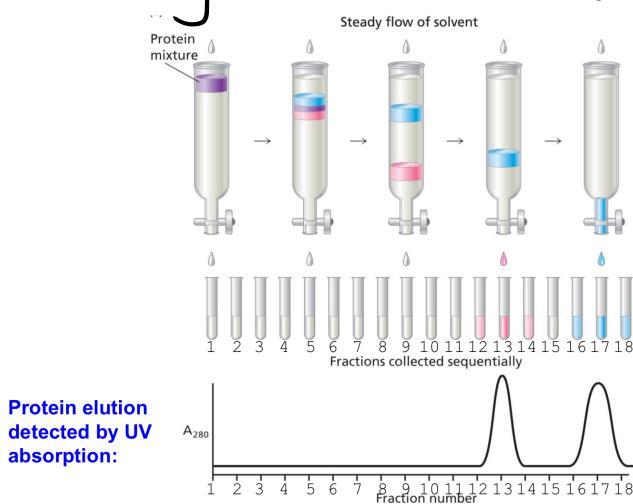


§4 Protein purification and analysis

1. Protein purification using column chromatography (柱层析法)

1° Basic steps

- ① Protein mixture is applied to a cylindrical (圆柱状的) column filled with an **insoluble** materials.
- ② Solvent is added to wash the proteins through the matrix of materials.
- ③ The eluate (洗出液) is collected in **many fractions**.
- ④ The concentration of protein in each fraction can be determined by **UV absorbance** at 280 nm.

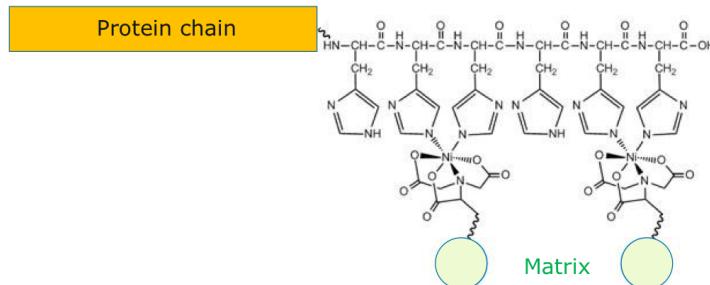


2° Affinity chromatography (亲和色谱法)

- ① This technique takes advantage of the fact that many proteins specifically bind other molecules.
- ② One can use this to construct a column containing the **ligand** (配位键) covalently attached to a matrix.
- ③ The proteins which can bind the ligand will be **retained** on the column.

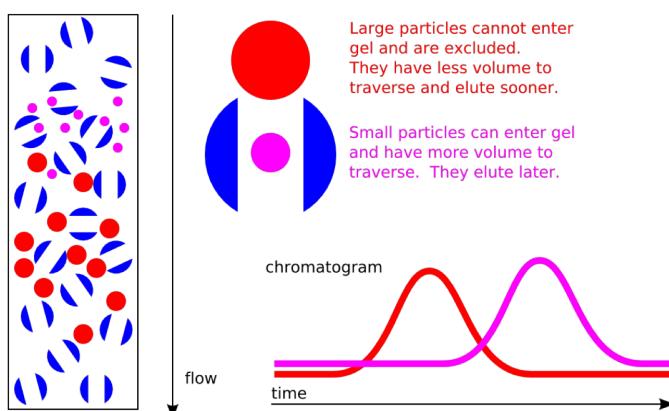
- ④ The protein bound to the ligand will be released by adding competitive binders (e.g. imidazole in Ni columns)

Ni columns for purifying **proteins** labelled with a **histidine tag**:



3° Gel filtration chromatography (Size-exclusion chromatography) (凝胶过滤色谱法)

- ① The protein is applied to the top of column consisting of porous beads.
 - ② The pores of the beads are of a controlled size.
 - ③ The larger proteins are excluded from the beads and flow through faster than smaller ones.
 - ④ Separate proteins by size, the largest are eluted first.



4^o Ion exchange chromatography (离子交换色层析法)

- ① Use the charge on a protein.
 - ② Ion in solution replace that are electrostatically bound to an inert support carrying groups with

opposite charge.

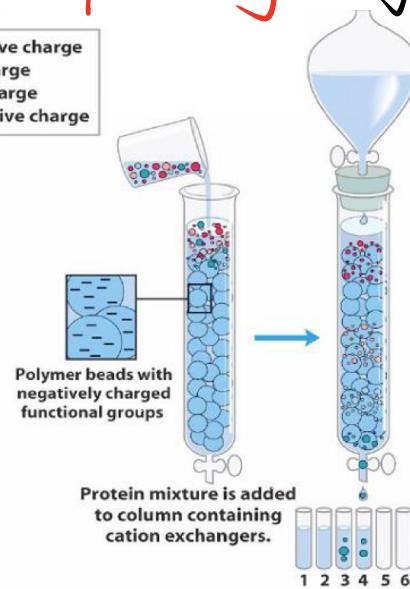
Categories:

Cation exchanger: bear negatively charged groups

Anion exchanger: bear positively charged groups

- Large net positive charge
- Net positive charge
- Net negative charge
- Large net negative charge

- The protein is loaded to the top of column. All proteins with the opposite charge or low charge can be removed.
- The proteins can be eluted from the column by changing the pH or by increasing the salt concentration, which shields the charge and thus decrease their attraction.

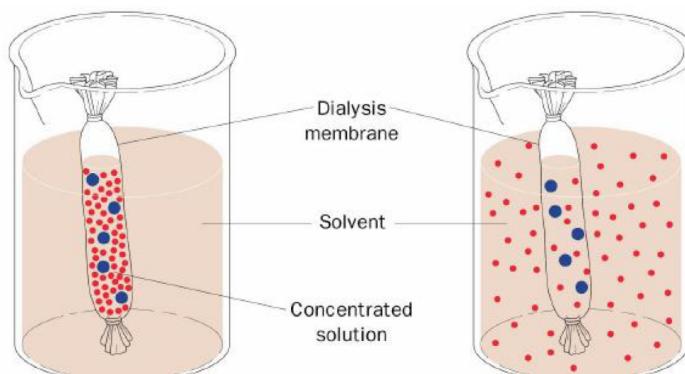


Proteins move through the column at rates determined by their net charge at the pH being used. With cation exchangers, proteins with a more negative net charge move faster and elute earlier.

2. Protein purification by dialysis (透析)

Protein is put within a dialysis membrane with pores of controlled size.

Proteins **bigger** than the pores are **retained**, while **smaller** molecules and salts can **diffuse out**.



3. Other protein purification methods

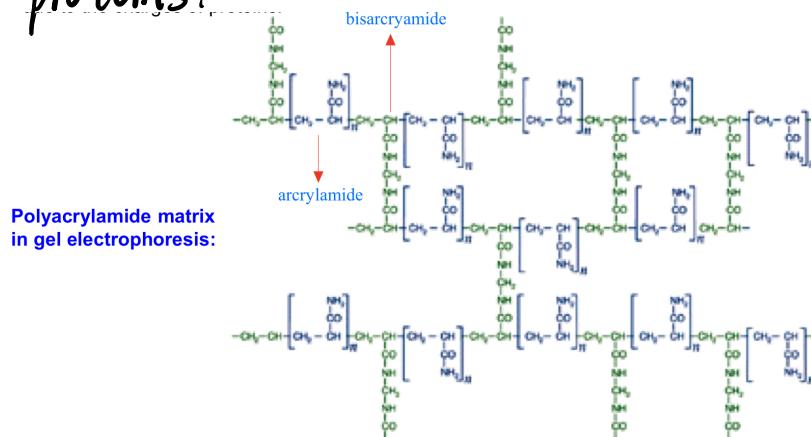
1° **Precipitation** (沉淀) at the pI

With $\text{pH} = \text{pI}$, proteins have the lowest solubility

2° Precipitation by adding high concentration of salt
 At very high concentration ($> 1M$) of certain salts (ammonium sulfate (硫酸铵)), proteins solubility is reduced due to a competition with the protein for interaction with water molecules.

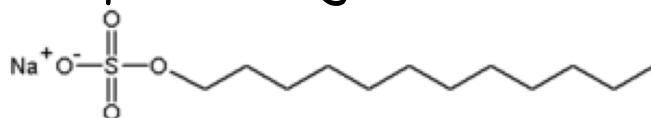
4. Protein analysis by gel electrophoresis (凝胶电泳)

1° All proteins penetrate (渗透) through the pores of gel matrix driven by an electric field, due to the charges of proteins.

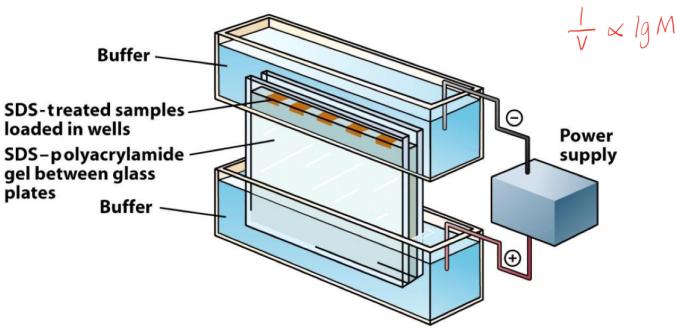


2° SDS - polyacrylamide gel electrophoresis (SDS-PAGE)

- ① Usually for analyzing protein molecule weight
- ② Proteins are denatured and solubilized by the detergent SDS
- ③ Protein and SDS form a roughly rod-like shape with negative charges
- ④ Separation is based on the size (molecule weight)
- ⑤ Smaller proteins migrate faster than larger one.



SDS (12-carbon chain)



$$\frac{1}{V} \propto \lg M$$

