



中山大學
SUN YAT-SEN UNIVERSITY

第2章 氨基酸、多肽和蛋白质

(Amino Acids, Polypeptides and Protein)

生物化学 I

生命科学学院

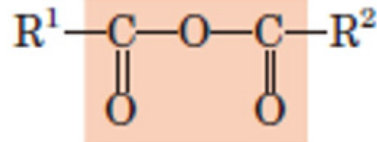
刘峰

liufeng23@mail.sysu.edu.cn

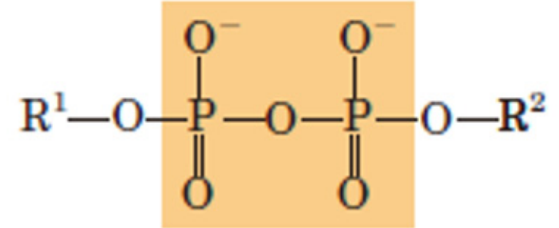
生命科学楼1号楼324室

绪论作业

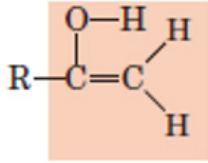
Anhydride
(two car-
boxylic acids)



Phosphoanhydride



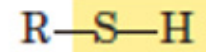
Enol



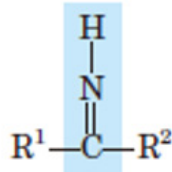
Carbonyl
(aldehyde)



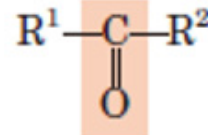
Sulfhydryl



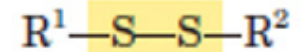
Imine



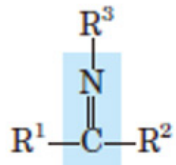
Carbonyl
(ketone)



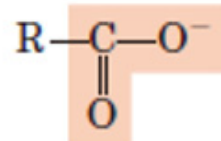
Disulfide



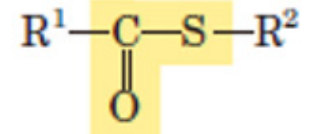
N-Substituted
imine (Schiff
base)



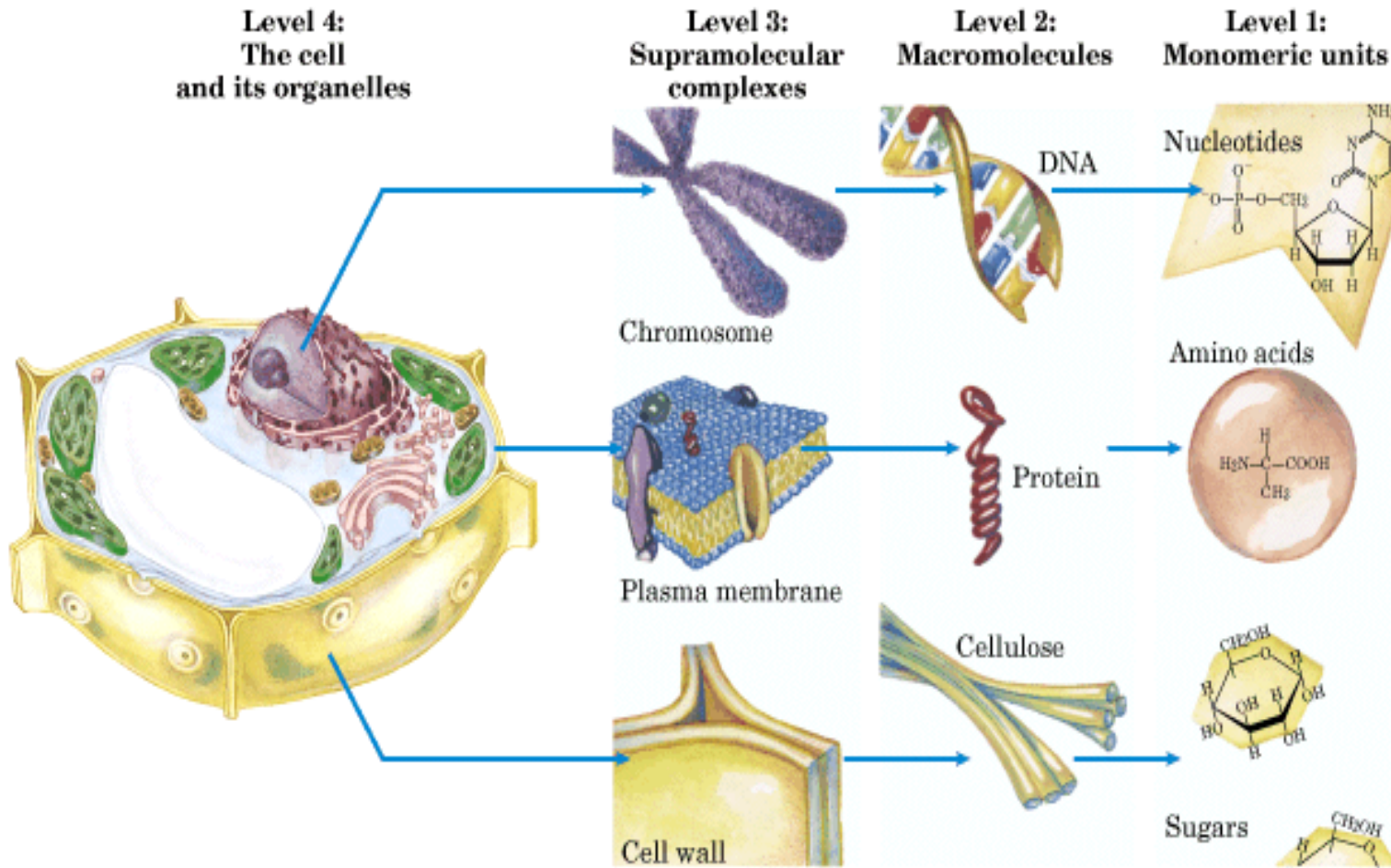
Carboxyl



Thioester



蛋白质是“行使生物学功能的生物大分子”

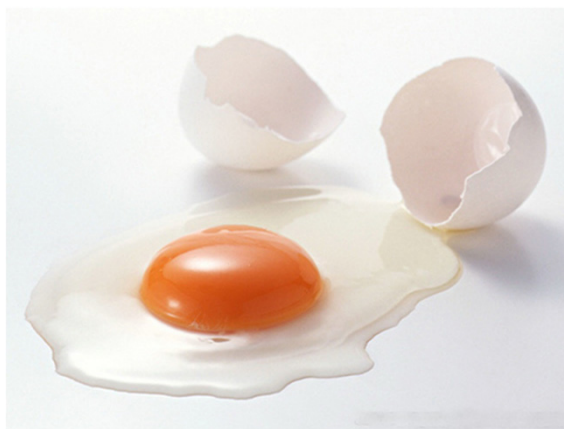


稳定

多样的化学基团

多样的组合方式

蛋白质在生活中以各种形式存在

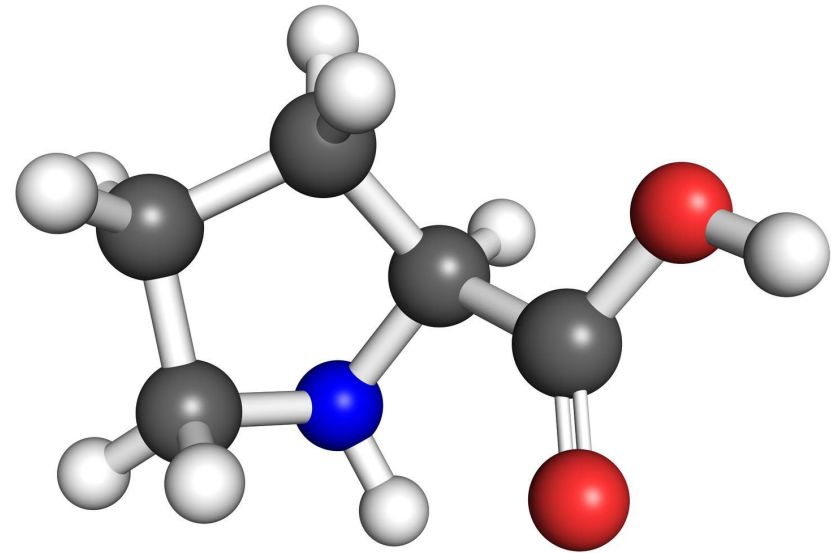


野生动物真的有营养么？



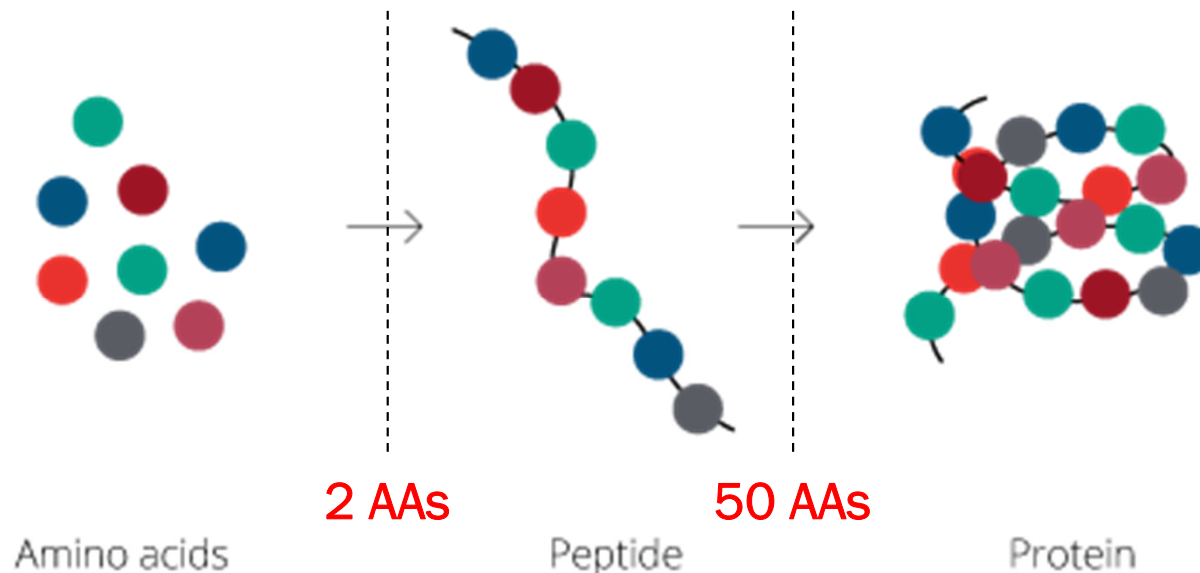
第2章 氨基酸、多肽与蛋白质

- ▶ 氨基酸的基本定义与特征
- ▶ 氨基酸的种类和结构
- ▶ 氨基酸的化学性质
- ▶ 多肽：氨基酸低聚物
- ▶ 蛋白质：氨基酸高聚物
- ▶ 蛋白质的研究方法



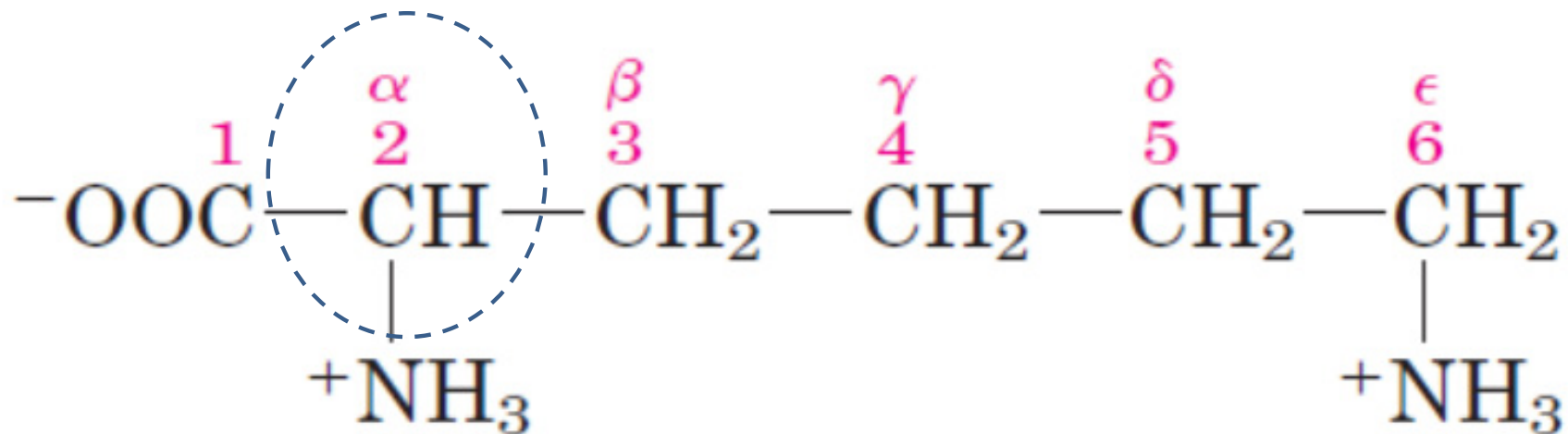
蛋白质、(多)肽、氨基酸

- ▶ **蛋白质**可以被酸、碱或蛋白酶催化水解。
- ▶ 在水解过程中，逐渐降解成相对分子质量越来越小的**肽段**，直到最后成为**氨基酸**的混合物。

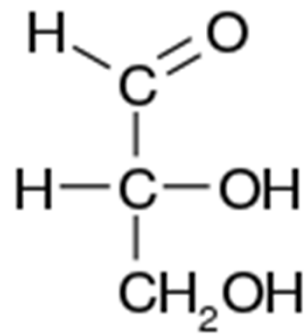


氨基酸 (amino acids) 是构成蛋白质的基础

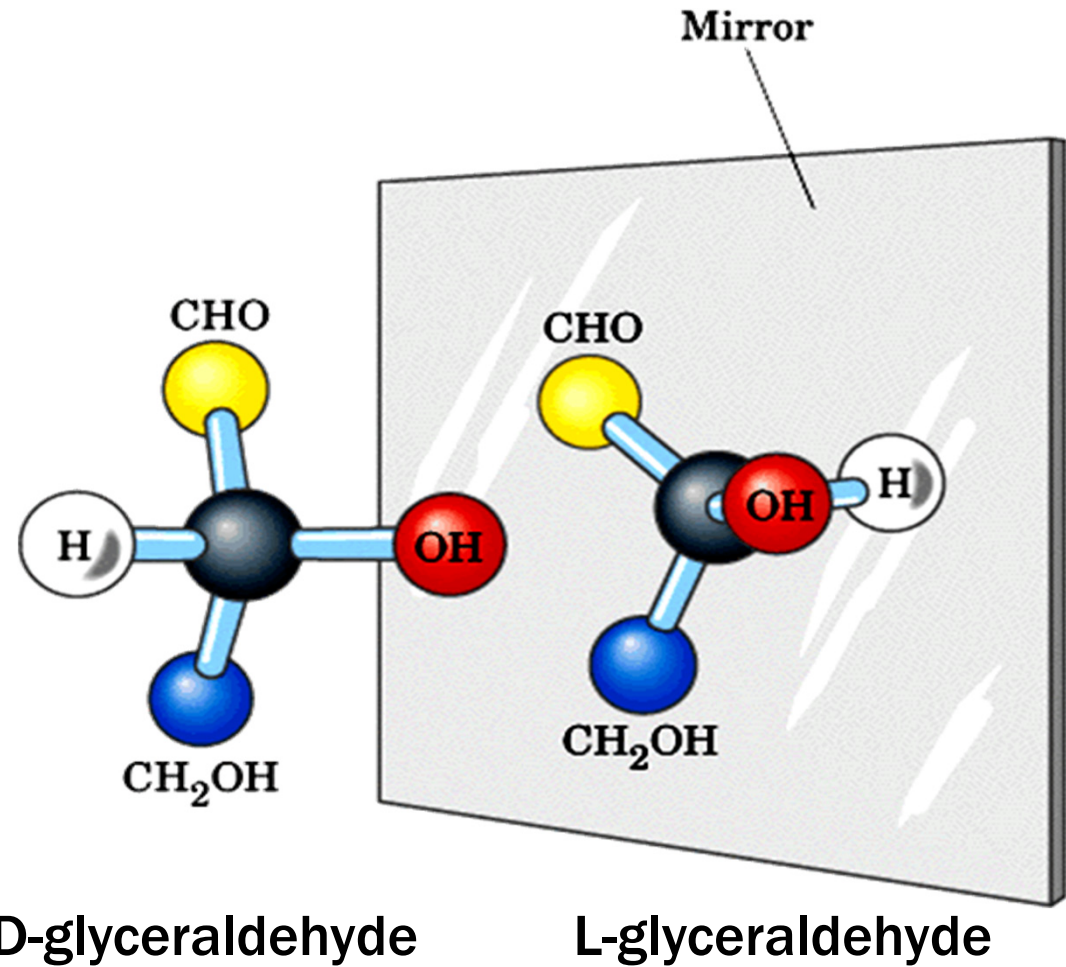
- ▶ 各种生物体中发现的氨基酸已有300多种
- ▶ From N to T (1806-1938) : 发现构成蛋白质的20种常见天然氨基酸 (常见蛋白质氨基酸)
- ▶ 天然氨基酸具有共同特征, 氨基 (或亚氨基) 连接在 α -碳原子 (α -氨基酸)



D (*dextro*)、L (*levo*) 构型

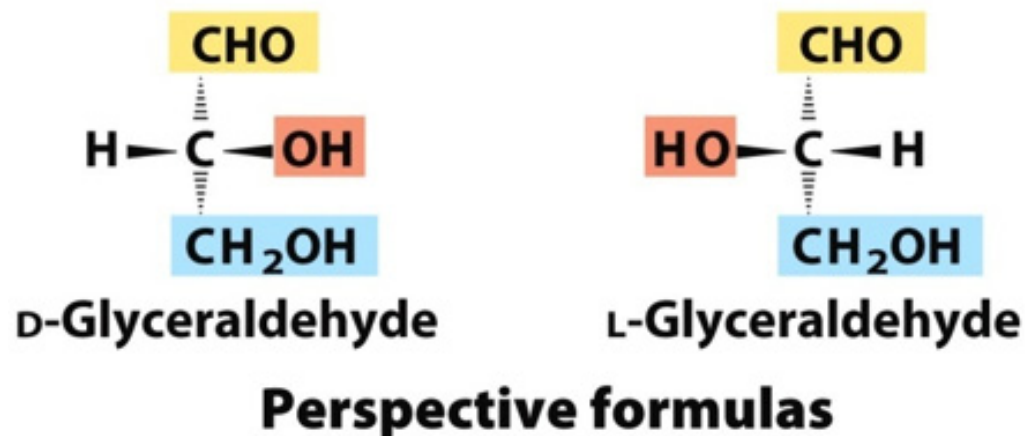


甘油醛
Glyceraldehyde

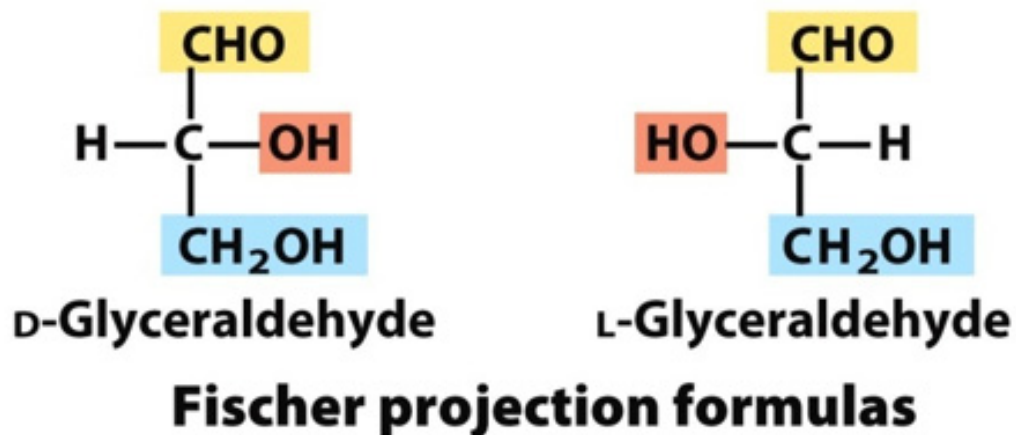


对映异构体 Enantiomers

表示形式 formulas



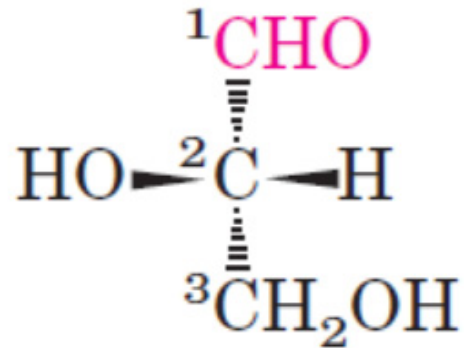
透视式



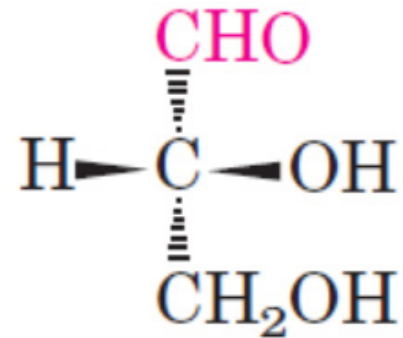
费歇尔投影式



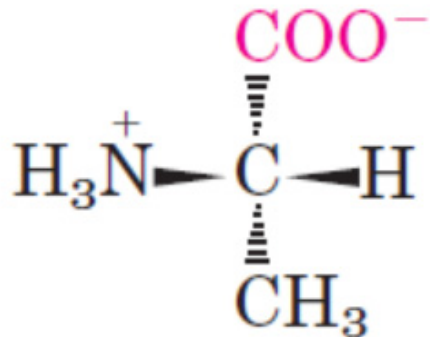
氨基酸的相对构型根据甘油醛决定



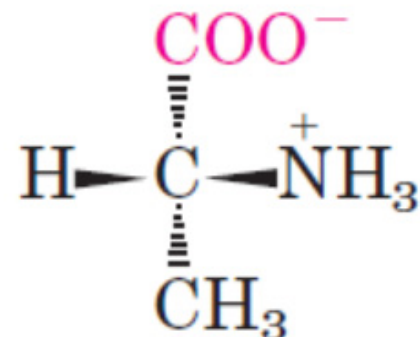
L-Glyceraldehyde



D-Glyceraldehyde



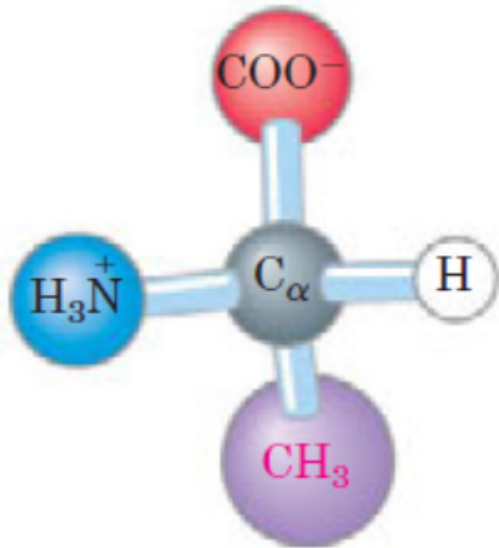
L-Alanine



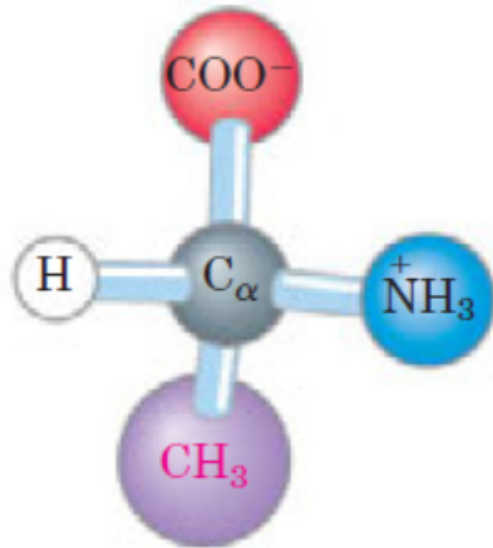
D-Alanine



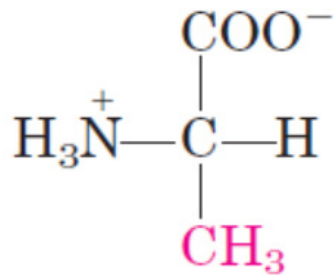
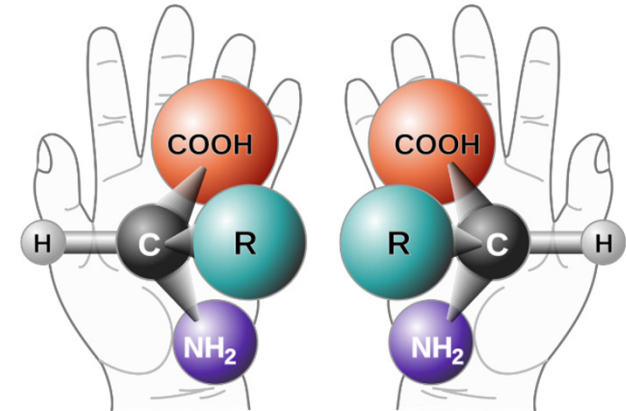
丙氨酸 (Alanine) 的分子式



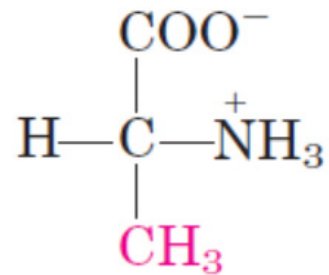
L-Alanine



D-Alanine



L-Alanine

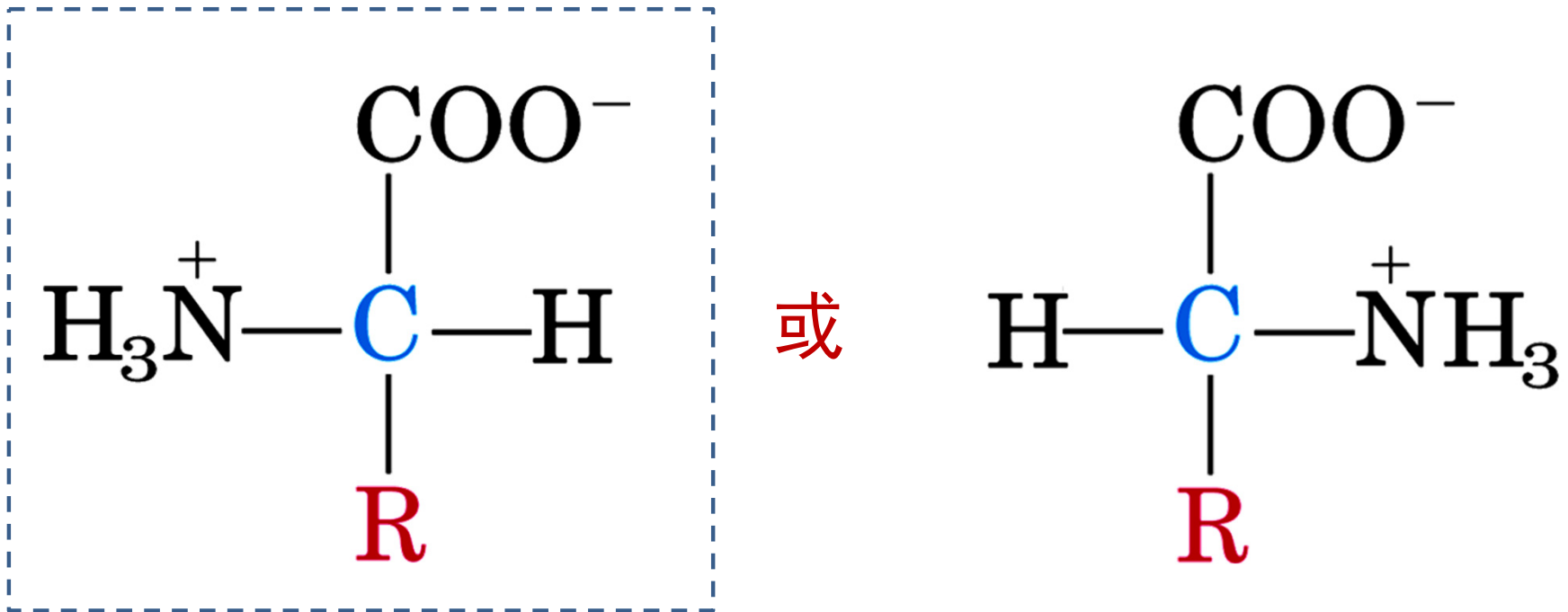


D-Alanine



氨基酸的通式

目前发现的参与形成蛋白质的氨基酸（蛋白质氨基酸）都是L型



1. 这两个式子有什么不同？
2. 为什么羧基和氨基如此写？

为什么目前蛋白质氨基酸都是L型？

nature
biotechnology

ARTICLES

<https://doi.org/10.1038/s41587-021-00969-6>

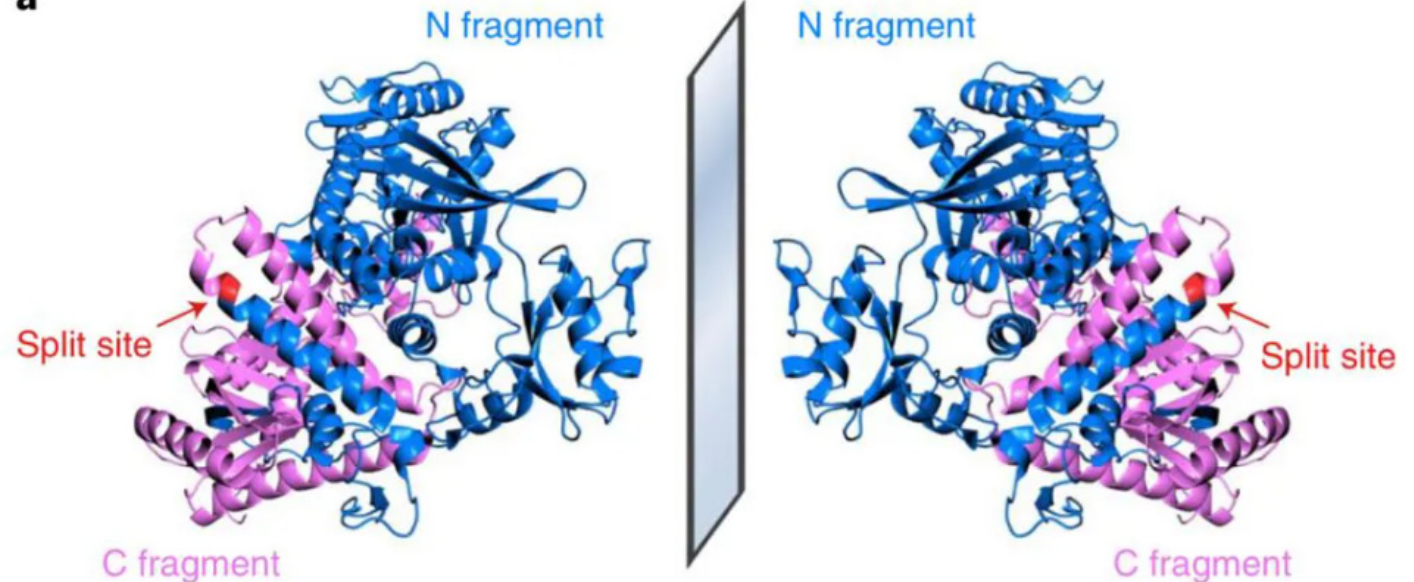
Check for updates

Bioorthogonal information storage in L-DNA with a high-fidelity mirror-image *Pfu* DNA polymerase

Chuyao Fan^{1,2}, Qiang Deng^{1,2} and Ting F. Zhu¹✉

Natural DNA is exquisitely evolved to store genetic information. The chirally inverted L-DNA, possessing the same informational capacity but resistant to biodegradation, may serve as a robust, bioorthogonal information repository. Here we chemically synthesize a 90-kDa high-fidelity mirror-image *Pfu* DNA polymerase that enables accurate assembly of a kilobase-sized mirror-image gene. We use the polymerase to encode in L-DNA an 1860 paragraph by Louis Pasteur that first proposed a mirror-image world of biology. We realize chiral steganography by embedding a chimeric D-DNA/L-DNA key molecule in a D-DNA storage library, which conveys a false or secret message depending on the chirality of reading. Furthermore, we show that a trace amount of an L-DNA barcode preserved in water from a local pond remains amplifiable and sequenceable for 1 year, whereas a D-DNA barcode under the same conditions could not. Our findings suggest that molecular tools may transform the development of advanced technologies for the exploration of the mirror-image central dogma and exploration of their

a



Natural *Pfu* DNA polymerase

Mirror-image *Pfu* DNA polymerase

D型氨基酸一样能够构成有功能的蛋白

使用L型真的只是一种巧合么？



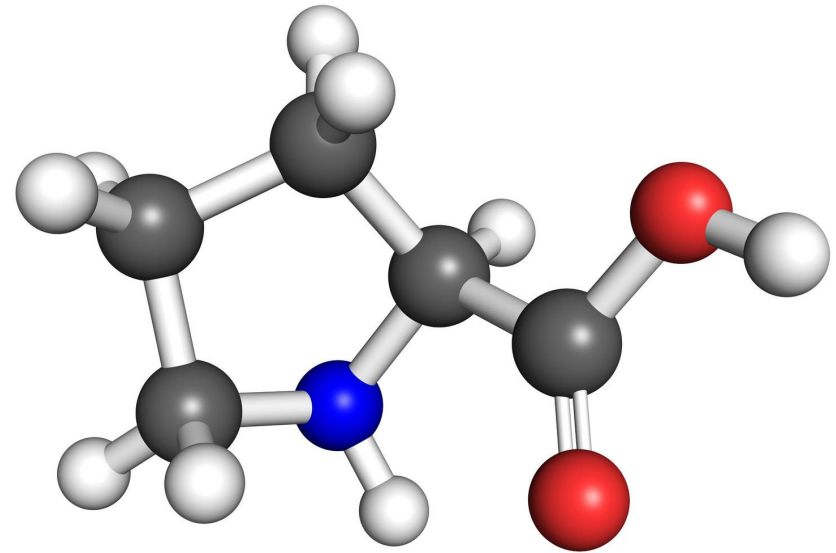
Q: 人体内有D型氨基酸吗?

- ▶ 蛋白质分子中没有D-氨基酸，但有一些小肽是有D-氨基酸的，如细菌细胞壁的主要成分肽聚糖中的小肽和很多细菌产生的抗菌肽。
 - ▶ 人体至少已经发现两种D-氨基酸：一种是D-丝氨酸，另外一种是N-甲基-D-天冬氨酸 (NMDA)
 - ▶ 细胞先合成L-氨基酸，在由特异的消旋酶将L-氨基酸异构成D-氨基酸。
 - ▶ 在人体内还没有发现有**任何肽含有D-氨基酸**。
-



第2章 氨基酸、多肽与蛋白质

- ▶ 氨基酸的基本定义与特征
- ▶ 氨基酸的种类和结构
- ▶ 氨基酸的化学性质
- ▶ 多肽：氨基酸低聚物
- ▶ 蛋白质：氨基酸高聚物
- ▶ 蛋白质的研究方法

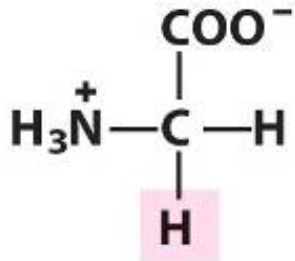


常见的蛋白质氨基酸

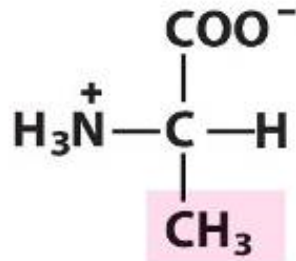
- ▶ R为非极性（7种）
- ▶ R为芳香族（3种）
- ▶ R为极性（10种）
 - ▶ 不带电5种
 - ▶ 带正电3种
 - ▶ 带负电2种



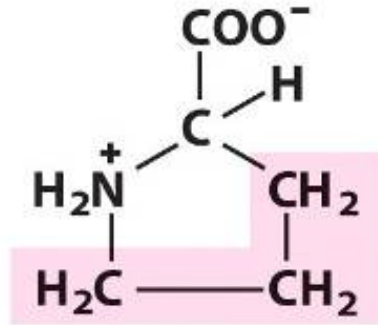
非极性脂肪族 (Nonpolar aliphatic)



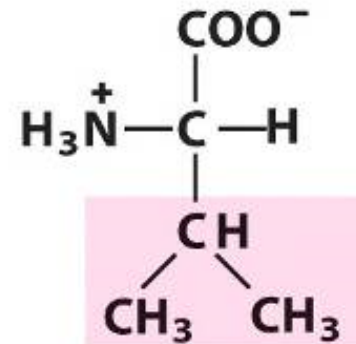
甘氨酸
Glycine



丙氨酸
Alanine

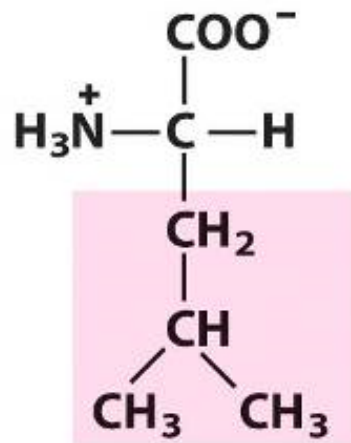


Proline 脯氨酸

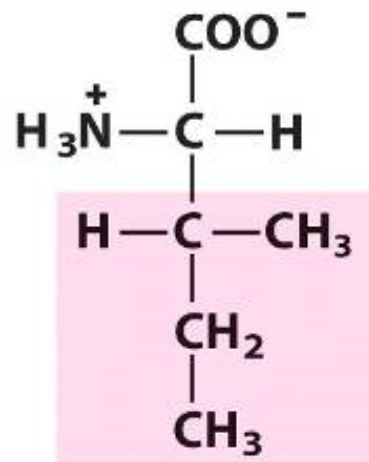


Valine 缬氨酸

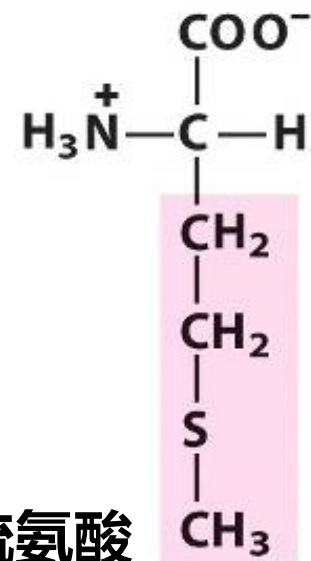
最好记的缩写



亮氨酸
Leucine



异亮氨酸
Isoleucine

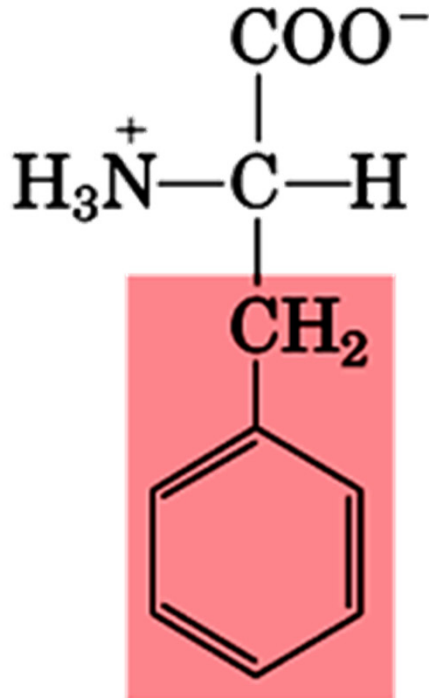


甲硫氨酸
Methionine

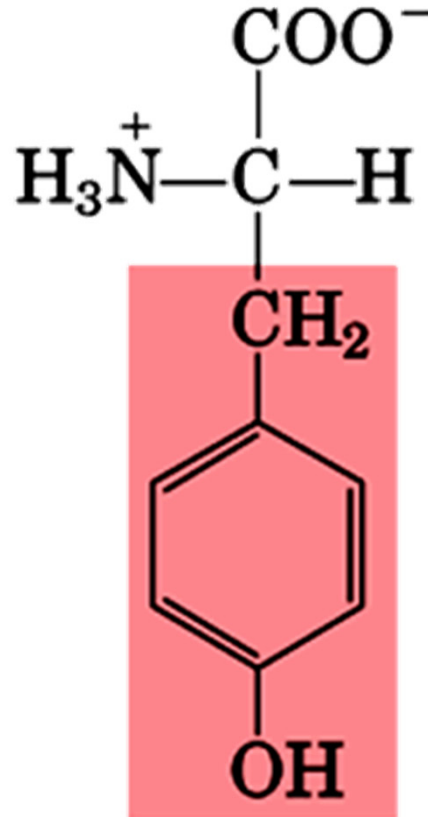
芳香族 (Aromatic)

最难记的缩写

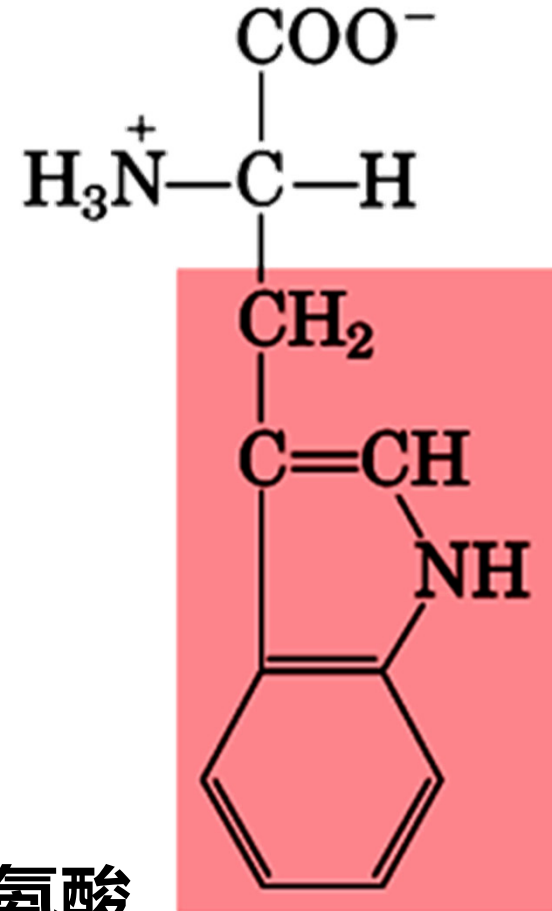
3



苯丙氨酸 **F**
Phenylalanine



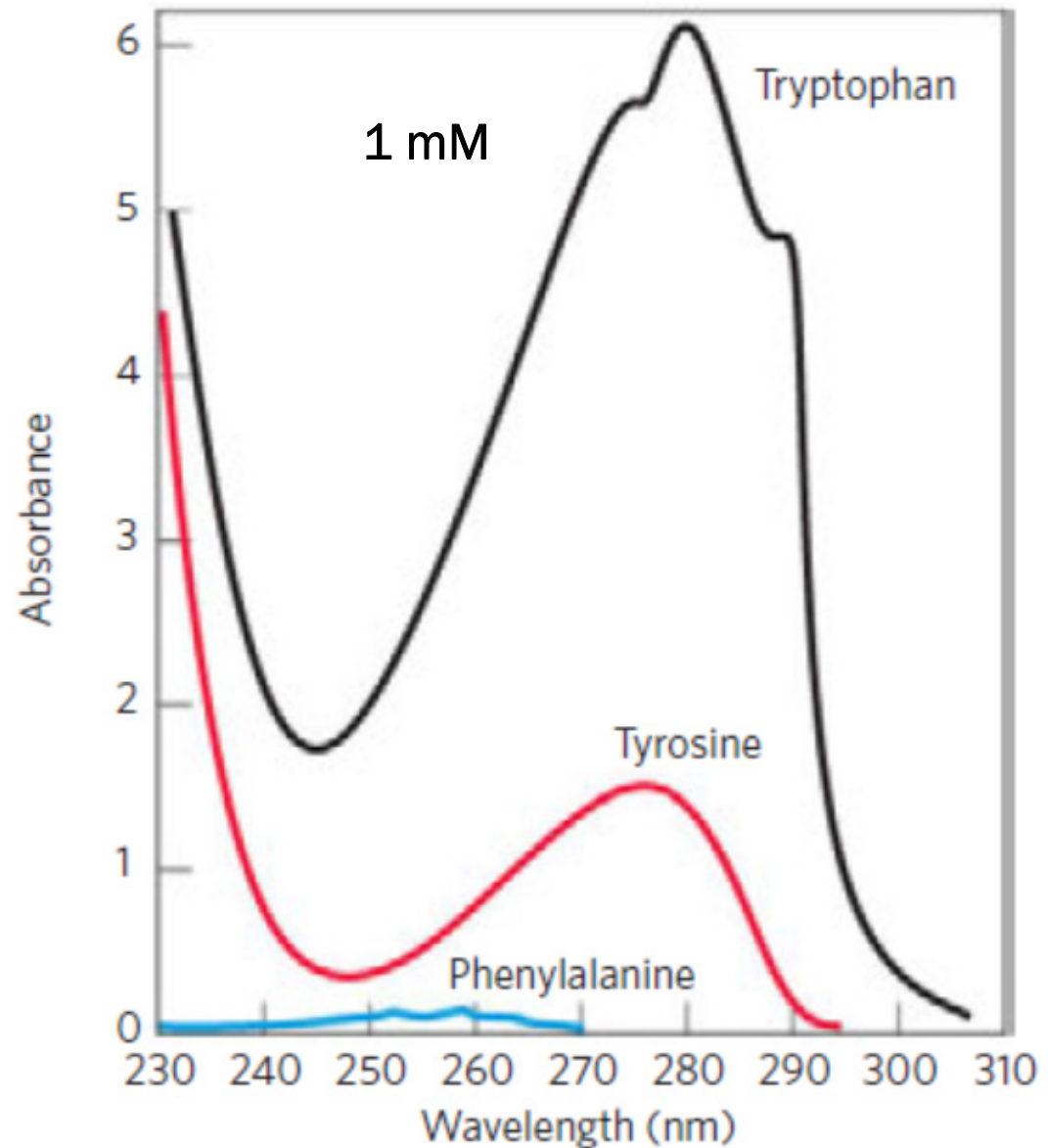
酪氨酸
Tyrosine



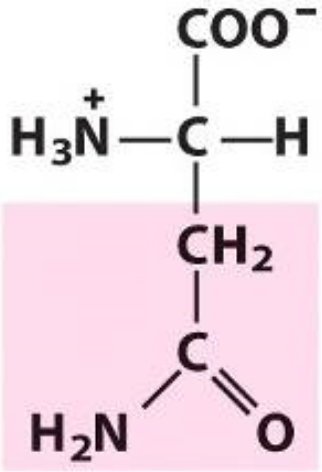
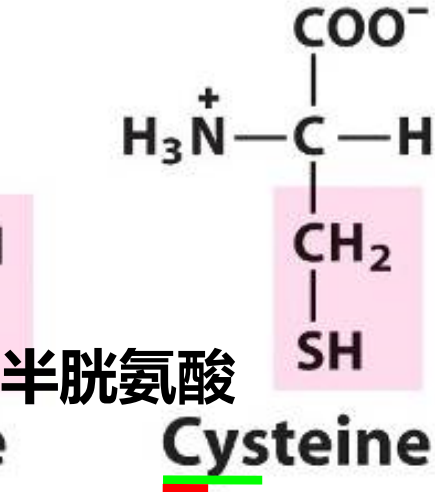
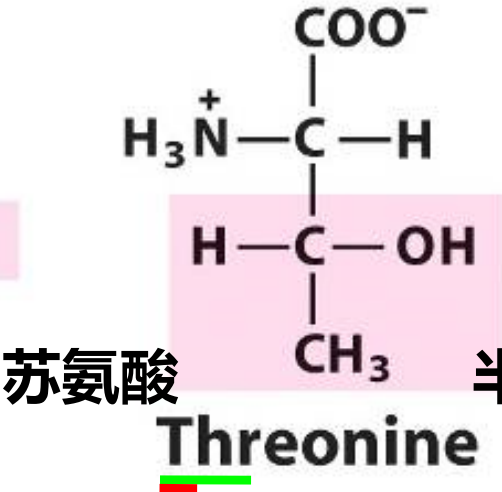
色氨酸
W Tryptophan

Q: 能否用紫外分光光度计测量蛋白质浓度?

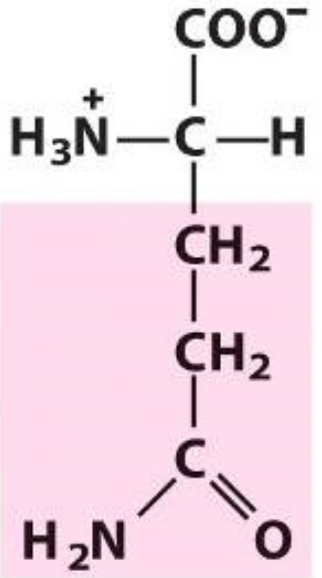
- ▶ 依据朗伯比尔定律
(The Lambert-Beer Law)
- ▶ 色氨酸和酪氨酸在约280 nm处有强吸收, 苯丙氨酸也有一定的紫外吸收



极性不带电 (Polar, uncharged)



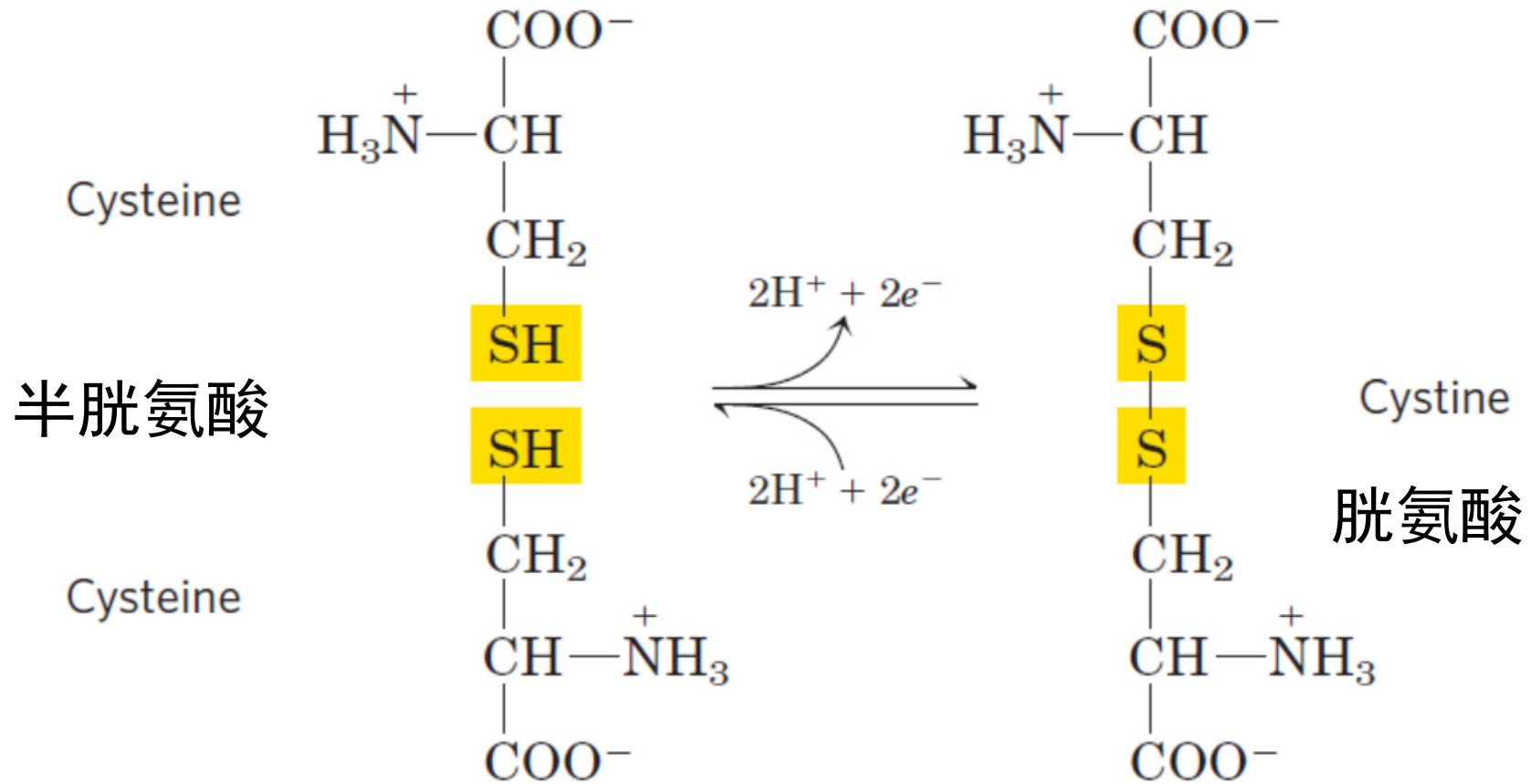
天冬酰胺 Asparagine



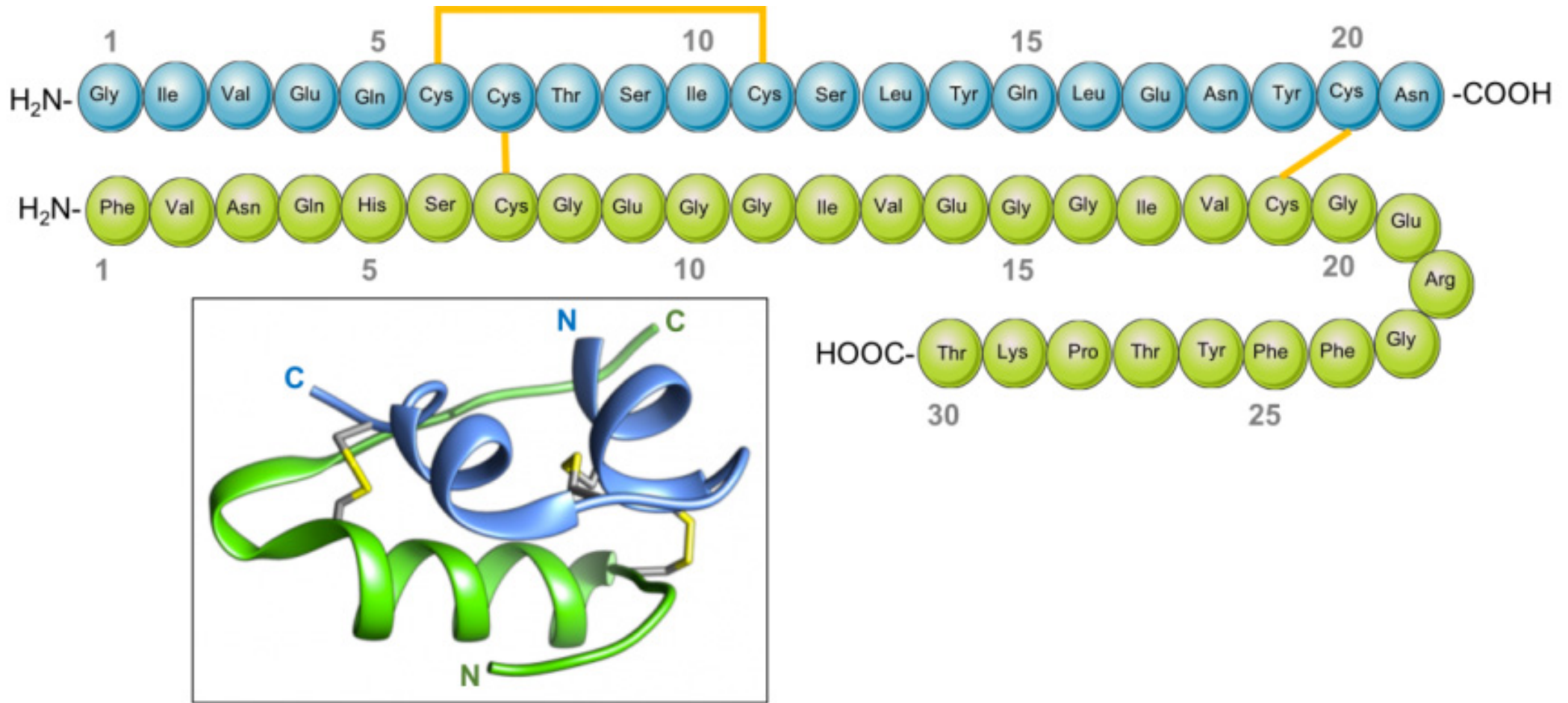
Glutamine 谷氨酰胺

Q

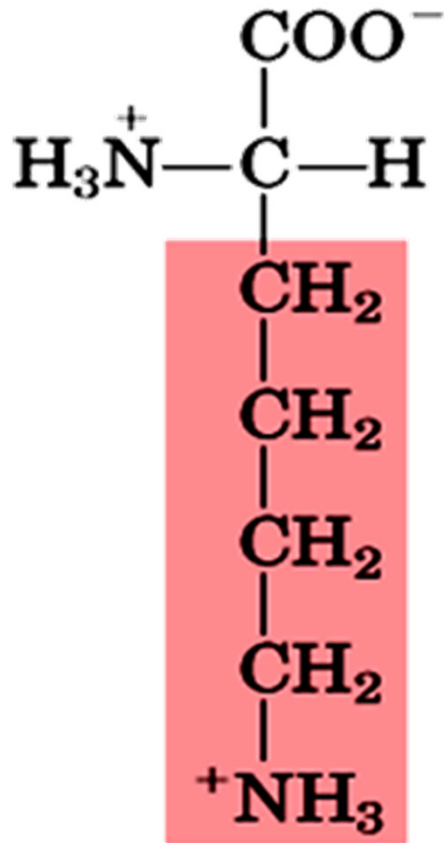
两个半胱氨酸被氧化后可以形成二硫键



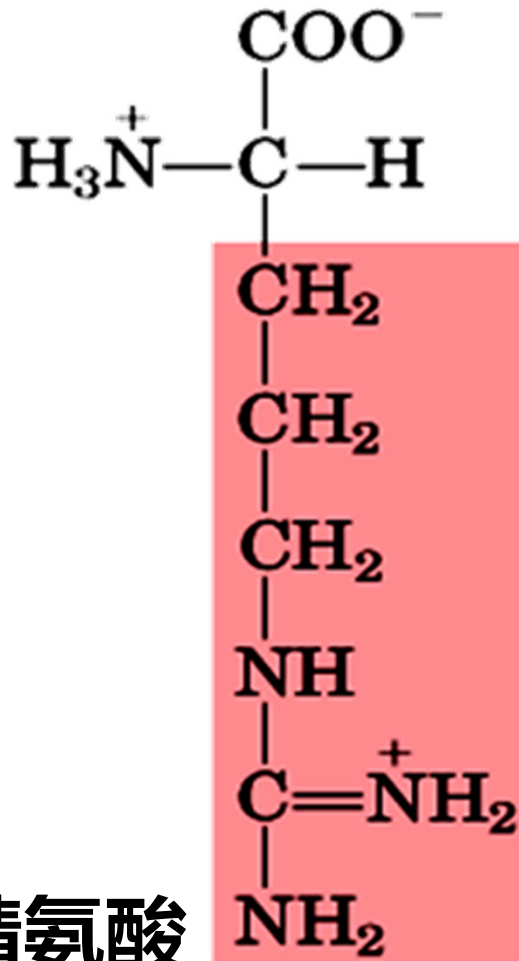
二硫键对于胰岛素形成正确的结构至关重要



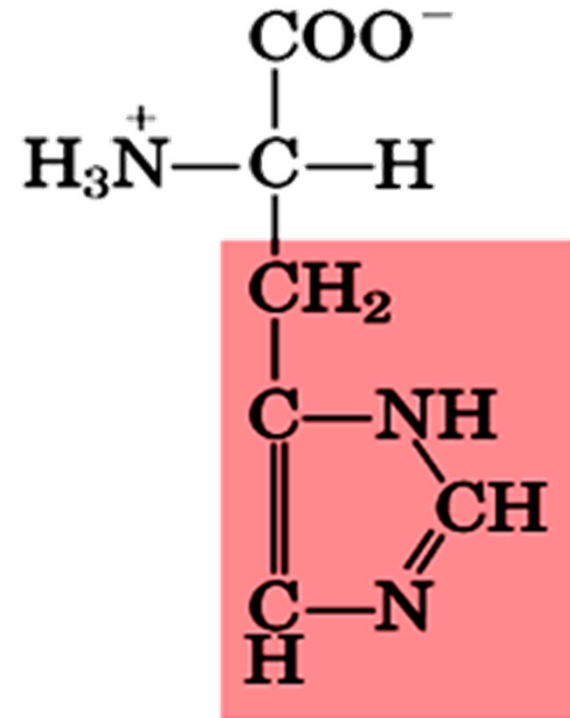
带正电 (Positively charged)



K 赖氨酸
Lysine

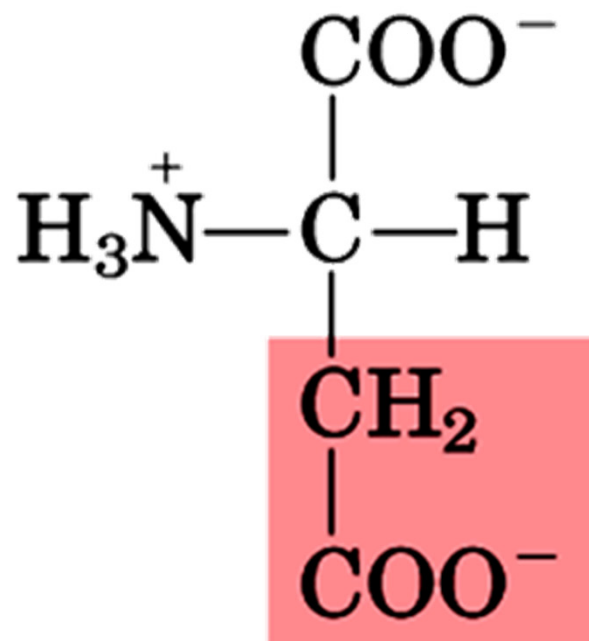


精氨酸
Arginine



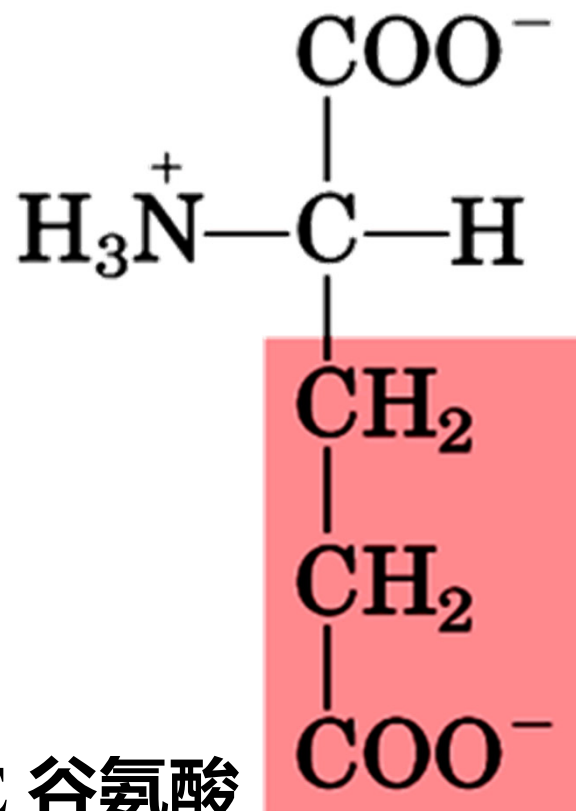
组氨酸
Histidine

带负电 (Negatively charged)



D 天冬氨酸

Aspartate



E 谷氨酸

Glutamate



氨基酸周期表

■ 碱性氨基酸

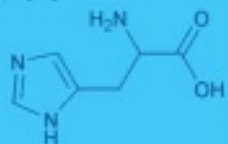
■ 极性氨基酸(不电离)

■ 非极性氨基酸(疏水性)

■ 酸性氨基酸

H His

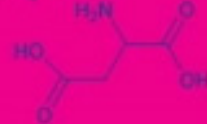
155.16
137.14
C₆H₉N₃O₂



组氨酸 Histidine

D Asp

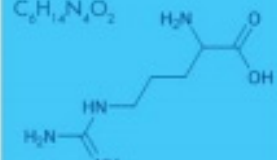
133.10
115.09
C₄H₇NO₄



天门冬氨酸 Aspartic Acid

R Arg

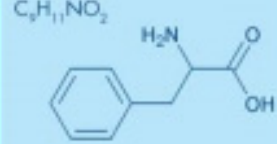
174.20
156.19
C₆H₁₂N₄O₂



精氨酸 Arginine

F Phe

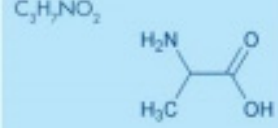
165.19
147.18
C₉H₉NO₂



苯丙氨酸 Phenylalanine

A Ala

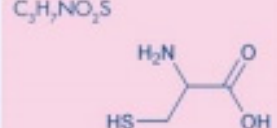
89.09
71.08
C₃H₇NO₂



丙氨酸 Alanine

C Cys

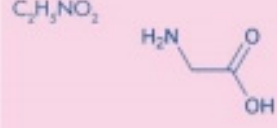
121.16
103.14
C₃H₇NO₂S



半胱氨酸 Cysteine

G Gly

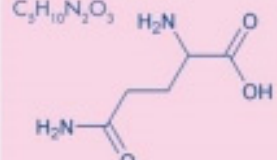
75.07
57.05
C₂H₃NO₂



甘氨酸 Glycine

Q Gln

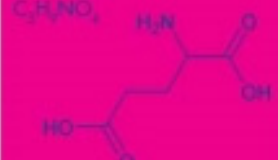
146.15
128.13
C₅H₁₀N₂O₃



谷氨酰胺 Glutamine

E Glu

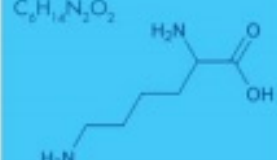
147.13
129.11
C₅H₉NO₄



谷氨酸 Glutamic Acid

K Lys

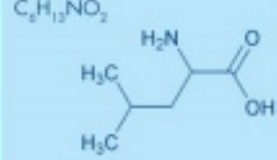
146.19
128.17
C₆H₁₂N₂O₂



赖氨酸 Lysine

L Leu

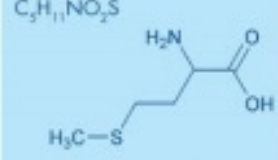
131.18
113.16
C₆H₁₃NO₂



亮氨酸 Leucine

M Met

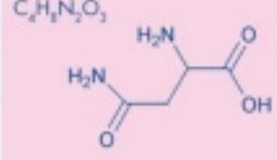
149.21
131.20
C₃H₇NO₂S



甲硫氨酸 Methionine

N Asn

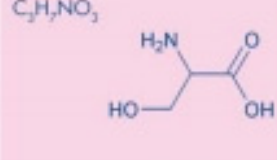
132.12
114.10
C₄H₈N₂O₃



天门冬酰胺 Asparagine

S Ser

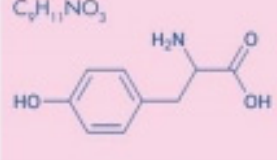
105.09
87.08
C₃H₇NO₃



丝氨酸 Serine

Y Tyr

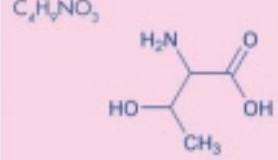
181.19
163.17
C₉H₉NO₃



酪氨酸 Tyrosine

T Thr

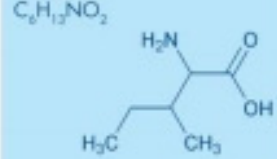
119.12
101.10
C₄H₉NO₃



苏氨酸 Threonine

I Ile

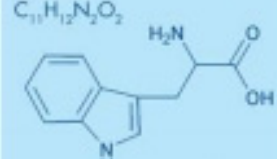
131.18
113.16
C₆H₁₃NO₂



异亮氨酸 Isoleucine

W Trp

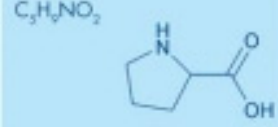
204.23
186.21
C₁₁H₁₂N₂O₂



色氨酸 Tryptophan

P Pro

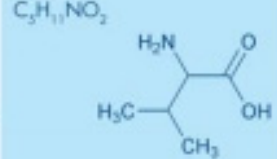
115.13
97.12
C₃H₇NO₂



脯氨酸 Proline

V Val

117.15
99.13
C₃H₇NO₂



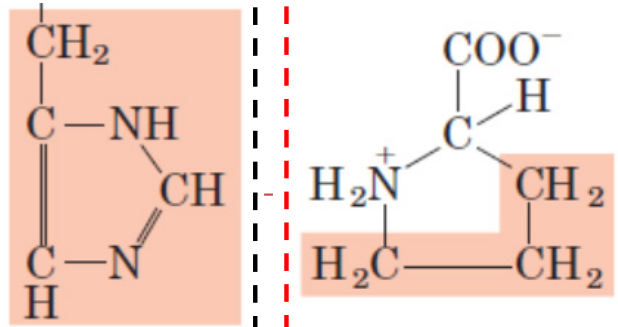
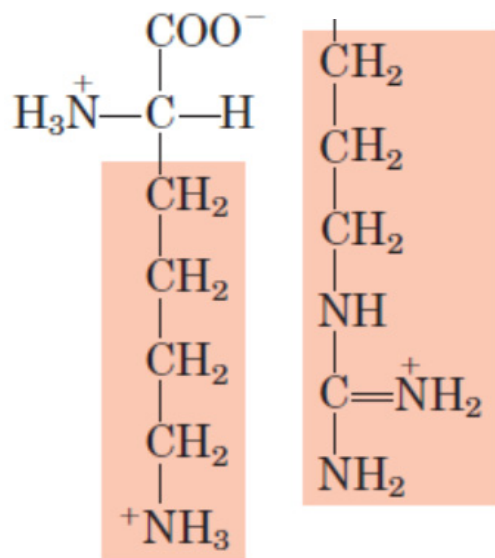
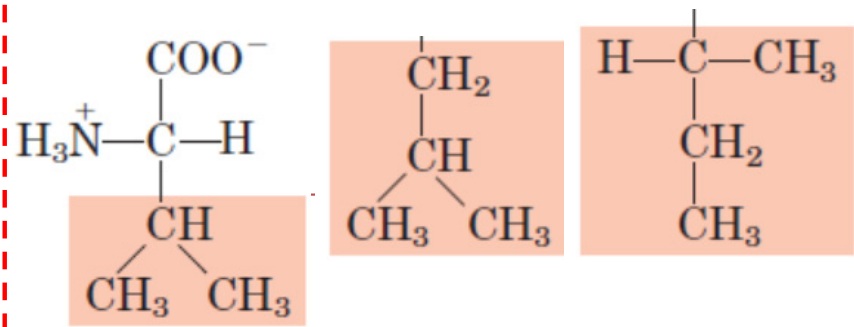
缬氨酸 Valine

单字母缩写 — **S** — 三字母缩写

分子量 — 105.09
残基分子量 — 87.08
分子式 — C₃H₇NO₃

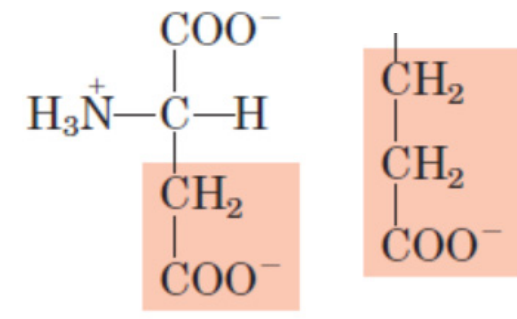
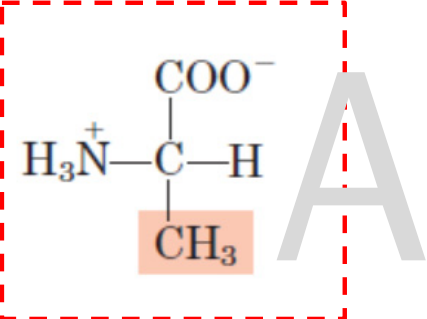
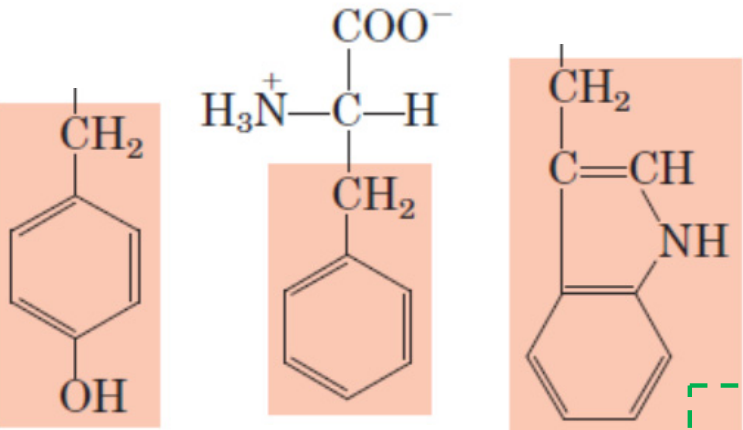
化学结构式

丝氨酸 Serine — 化学名



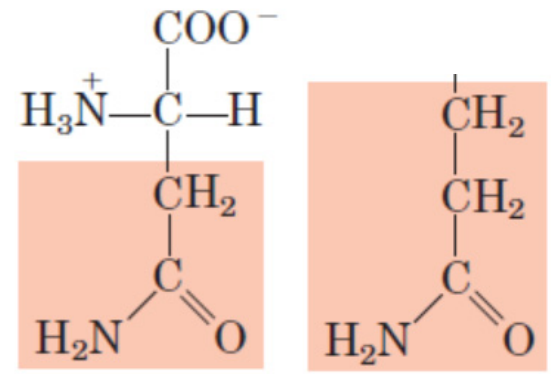
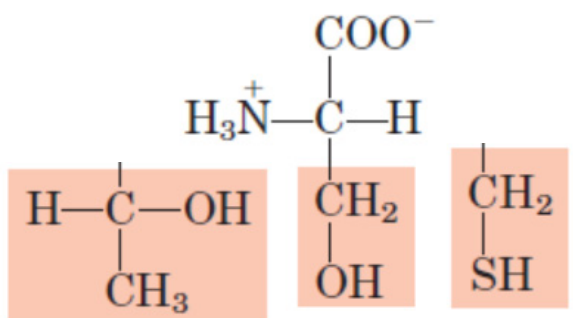
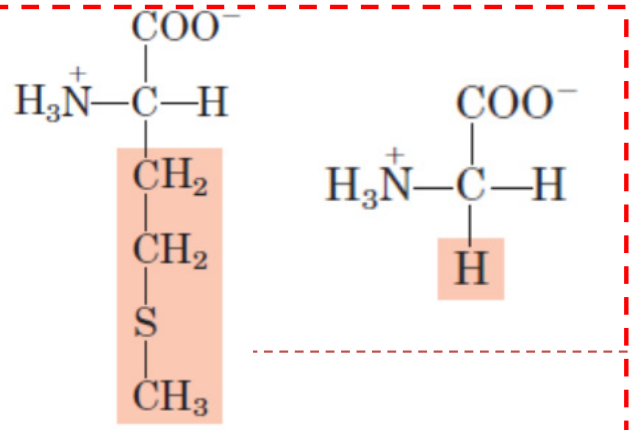
H→芳环 (苯丙氨酸)

H→羧酸 (天冬氨酸)



H→羟 (丝氨酸)

H→酰胺 (天冬酰胺)



氨基酸的单字母编码的创建

MTTC SRQFTSSSS MKGSCGIGGGIGAGSSRISSVLAGGSCRAPNTYGGGLSVSSSRFSSGGAYGLGG
GYGGGFSSSSSSSFGSGFGGGYGGGLGAGLGGGFAGGDGLLVGSEKVTMQNLNDR LASYLDK
VRALEEANADLEV KIRDWYQRQRPAEIKDYSPYFKTIEDLRNKILTATVDNANVLLQIDNARLAADDFRTK
YETELNLRMSVEADINGLRRVLDLTLARADLEMQIESLKEELAYLKKNHEEEMNALRGQVGGDVNVEM
DAAPGVDL SRILNEMRDQYEKMAEKNRKDAEEWFFTKTEELNREVATNSELVQSGKSEISELRRTMQN
LEIELQS QLSMKASLENSLEETKGRYCMQLAQIQEMIGSVEEQLAQLRCEMEQQNQEYKILLDVKTRLE
QEIATYRRLLGEDAHLSSSQFSSGSQSSRDVTSSSRQIRTKVMDVHDGKVVSTHEQVLRTKN



- ▶ Margaret Oakley Dayhoff (1925 - 1983), 美国物理化学家, “生物信息学之母”。
- ▶ 1965年出版《The Atlas of Protein Sequence and Structure》奠基了生物信息学。
- ▶ 促成1971年成立蛋白质资源信息数据库 (Protein Information Resource, PIR)。

氨基酸的英文缩写

1 独特的首个字母（6个）

中文名	英文名	三字母	单字母	性质
半胱氨酸	Cysteine	Cys	C	极性
组氨酸	Histidine	His	H	正电
异亮氨酸	Isoleucine	Ile	I	非极
甲硫氨酸	Methionine	Met	M	非极
丝氨酸	Serine	Ser	S	极性
缬氨酸	Valine	Val	V	非极

氨基酸的英文缩写

2 首字母有重复，更常见氨基酸优先命名（5个）

中文名	英文名	三字母	单字母	性质
丙氨酸	Alanine	Ala	A	非极
甘氨酸	Glycine	Gly	G	非极
亮氨酸	Leucine	Leu	L	非极
脯氨酸	Proline	Pro	P	非极
苏氨酸	Threonine	Thr	T	极性

氨基酸的英文缩写

最难记忆的一页

3 发音类似的名称 (8个)

中文名	英文名	三字母	单字母	性质
精氨酸	Arginine	Arg	R	正电
天冬酰胺	Asparagine	Asn	N	极性
天冬氨酸	Aspartate	Asp	D	负电
谷氨酸	Glutamate	Glu	E	负电
谷氨酰胺	Glutamine	Gln	Q (Q-tamin)	极性
苯丙氨酸	Phenylalanine	Phe	F	芳香
酪氨酸	Tyrosine	Tyr	Y	芳香
色氨酸	Tryptophan	Trp	W (双环)	芳香

氨基酸俗名的来历

- ▶ 天冬酰胺（Asparagine）发现于芦笋（asparagus）（首个发现的氨基酸，1806）
- ▶ 谷氨酸（Glutamate）表示来自面筋（wheat gluten）的酸性氨基酸
- ▶ 酪氨酸（Tyrosine）来自cheese（希腊语“tyros”）



氨基酸的英文缩写

4 临近首字母（1个）

中文名	英文名	三字母	单字母	性质
赖氨酸	Lysine	Lys	K	正电

A	B	C	D	E	F
G	H	I	J	K	L
M	N	O	P	Q	R
S	T	U	V	W	X
Y	Z				

Which letters are still available for other amino acids?



氨基酸的英文缩写

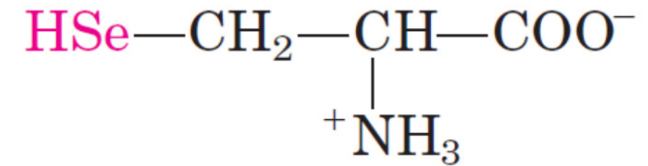
5 剩余字母（6个）

中文名	英文名	三字母	单字母	性质
吡咯赖氨酸	Pyrrolysine		O	非极
硒代半胱氨酸	Selenocysteine		U	负电
天冬氨酸/酰胺	/	Asx	B	
谷氨酸/酰胺	/	Glx	Z	
亮/异亮氨酸	/	Xle	J	
未知氨基酸	/	Xaa	X	

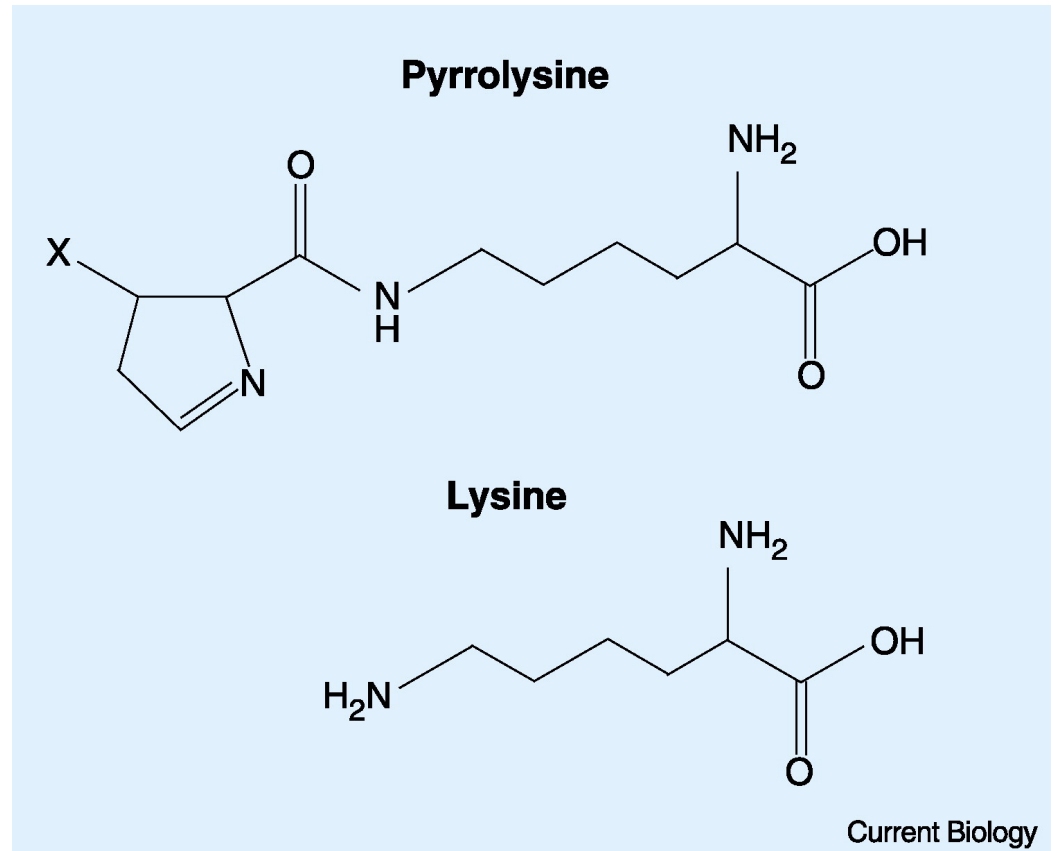
Try to “translate” any sentence into AMINO ACIDS.

重要但不常见的蛋白质氨基酸

存在于古细菌、细菌和动植物中



- ▶ 硒代半胱氨酸
(selenocysteine)
- ▶ 吡咯赖氨酸
(Pyrrolysine)

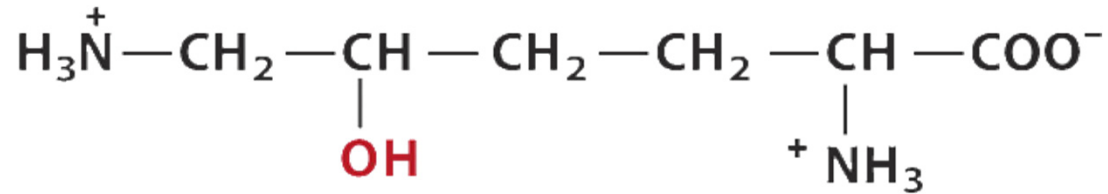


✦ *Selenocysteine: UGA*

✦ *Pyrrolysine: UAG*

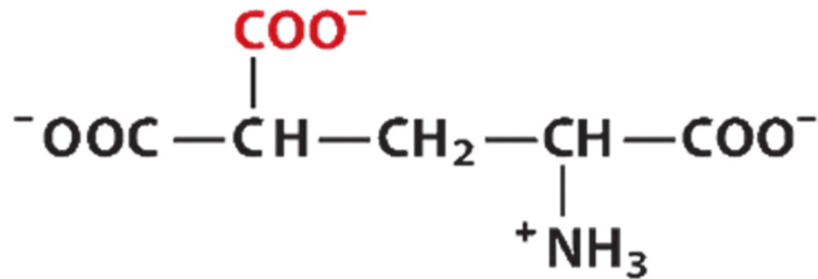
✦ *Science (2002) vol. 296, 1409; 1459; 1462*

Q: 这些氨基酸是蛋白质氨基酸么?



5-Hydroxylysine

5-羟基赖氨酸（胶原蛋白中）



γ -Carboxyglutamate

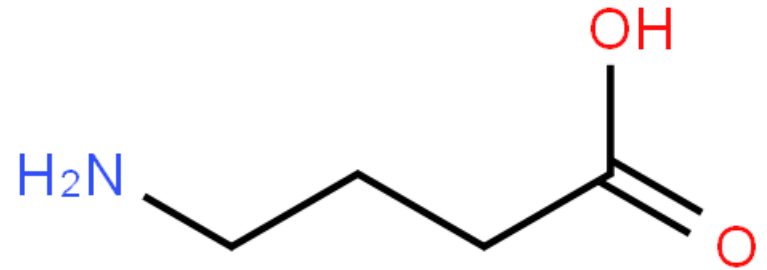
γ -羧基谷氨酸（凝血酶原中）

- ▶ 不是。合成后再修饰的产物，没有对应的密码子。
-



非蛋白质氨基酸

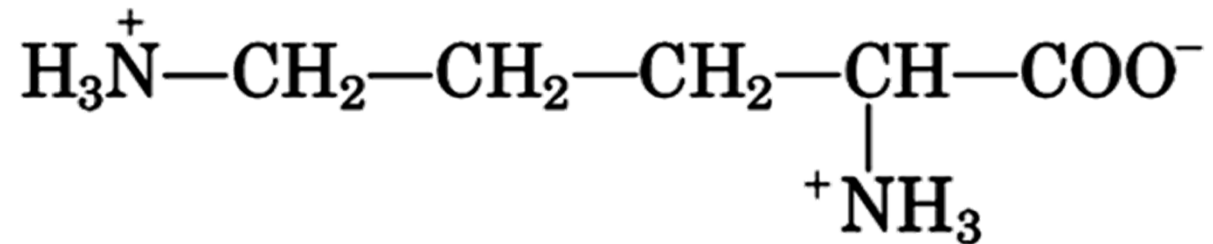
▶ 神经递质：γ-氨基丁酸



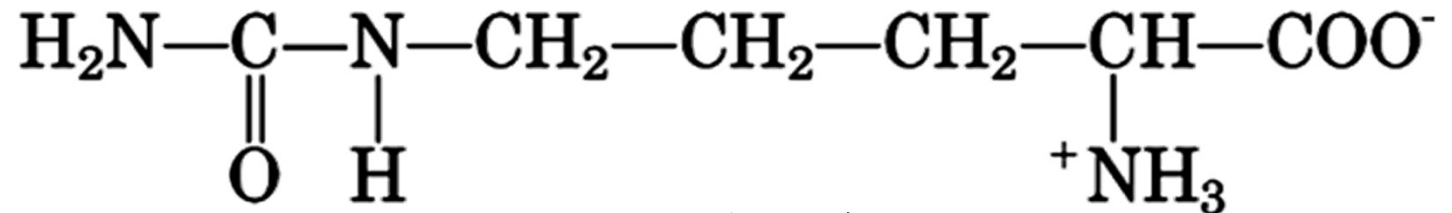
▶ 精氨酸合成和尿素循环中间产物

▶ L-瓜氨酸

▶ L-鸟氨酸



鸟氨酸



瓜氨酸



必需氨基酸 (Essential Amino Acids)

- ▶ 人体生命活动需要但**自身无法合成或合成量不足**
- ▶ 含有这8种氨基酸且比例合适的蛋白质是（成人）优质蛋白
- ▶ 野生动物与家畜在蛋白组成上无特别之处
- ▶ 笨蛋来宿舍凉一凉鞋（FMKTWLIV）
- ▶ 婴儿也需要组氨酸（H）

- ▶ 植物和微生物**不存在必需氨基酸**



Q: 人体中有多少种蛋白质氨基酸?

- ▶ 人体中有21种蛋白质氨基酸
- ▶ 生物中有22种蛋白质氨基酸
- ▶ 生物（或人体）中有20种常见的蛋白质氨基酸

- ▶ 成人中有8种必需氨基酸
- ▶ 婴儿中有9种必需氨基酸



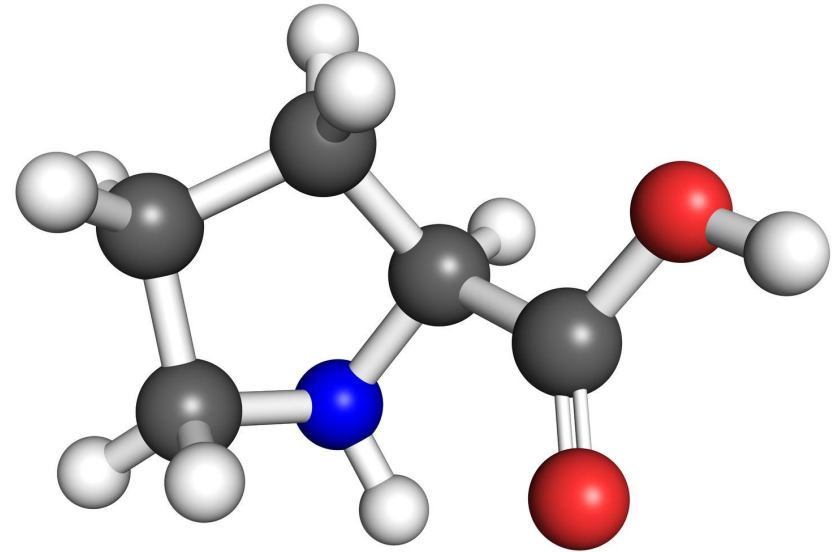
▶ I LOVE SUN YAT-SEN UNIVERSITY

▶ WE LIKE BIOCHEMISTRY

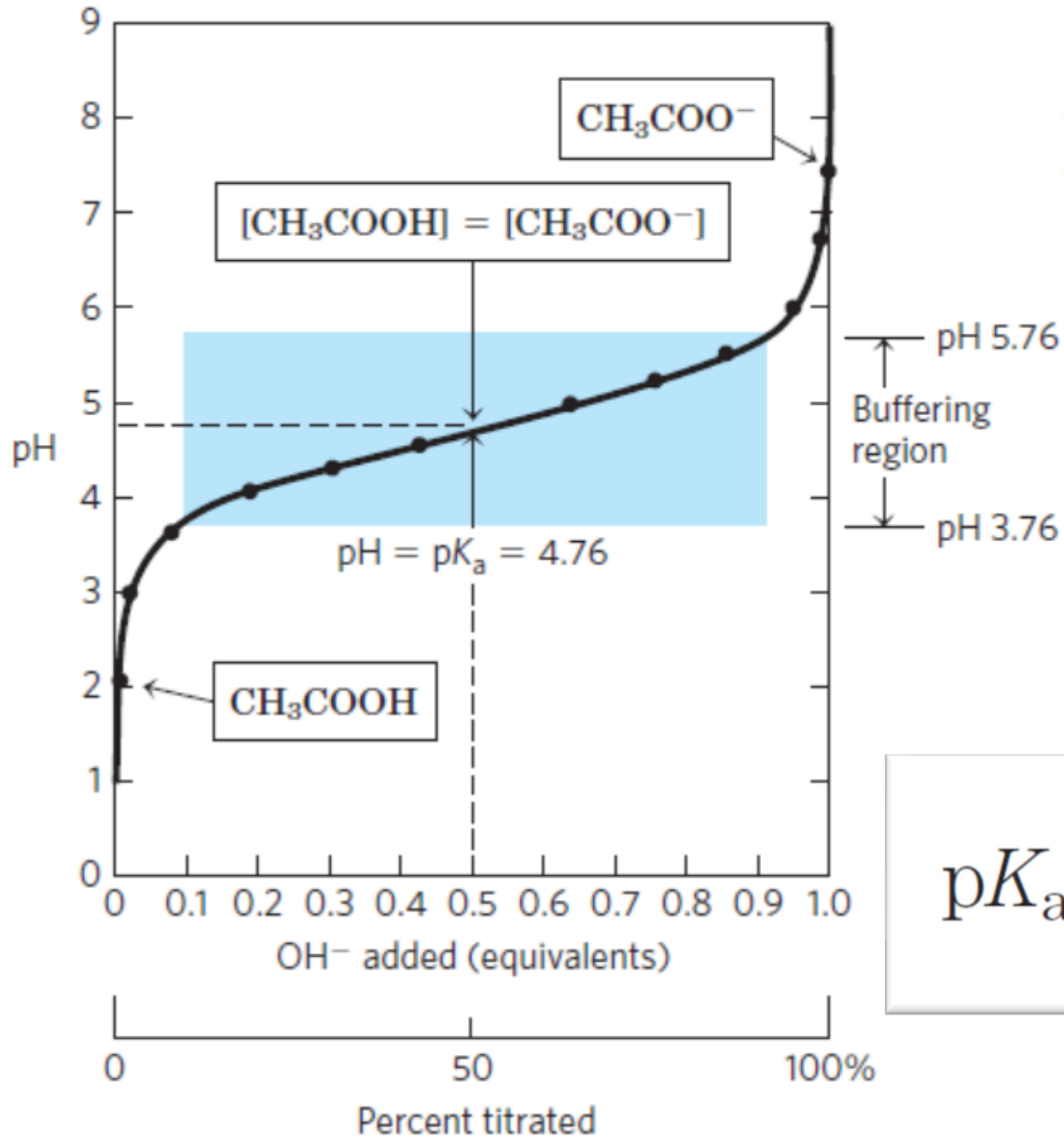


第2章 氨基酸、多肽与蛋白质

- ▶ 氨基酸的基本定义与特征
- ▶ 氨基酸的种类和结构
- ▶ 氨基酸的化学性质
- ▶ 多肽：氨基酸低聚物
- ▶ 蛋白质：氨基酸高聚物
- ▶ 蛋白质的研究方法



酸解离常数 (acid dissociation constant)



$$\frac{[\text{H}^+][\text{A}^-]}{[\text{HA}]} = K_a$$

$$\text{p}K_a = \lg \frac{1}{K_a} = -\lg K_a$$

酸碱平衡方程

- ▶ 亨德森-哈塞尔巴尔赫方程 (Henderson-Hasselbalch equation)
- ▶ 用于计算 pK_a 、pH、估算电离程度等

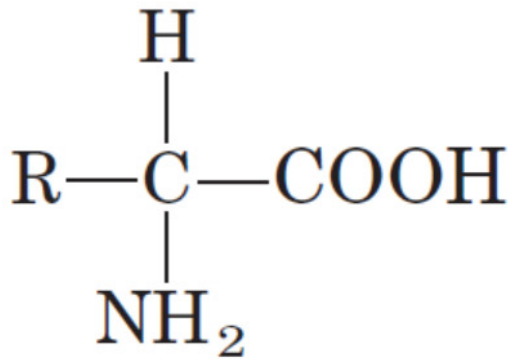
$$K_a = \frac{[H^+][A^-]}{[HA]} \longrightarrow [H^+] = K_a \frac{[HA]}{[A^-]}$$

$$-\lg [H^+] = -\lg K_a - \lg \frac{[HA]}{[A^-]}$$

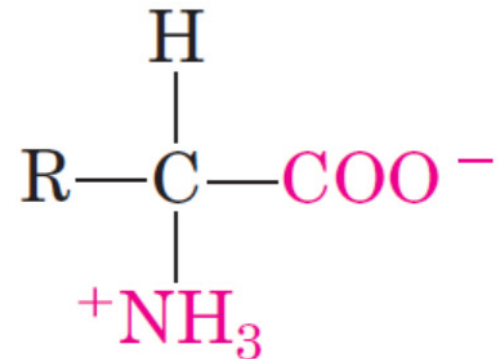
$$pH = pK_a + \lg \frac{[A^-]}{[HA]}$$

氨基酸的兼性离子 (zwitterion) 形式

- ▶ 氨基酸在特定pH值条件下形成的**偶极离子**形式。
- ▶ 净电荷为零，在**电场中不移动**。
- ▶ 既起**酸**的作用，又起**碱**的作用，是一类两性电解质。



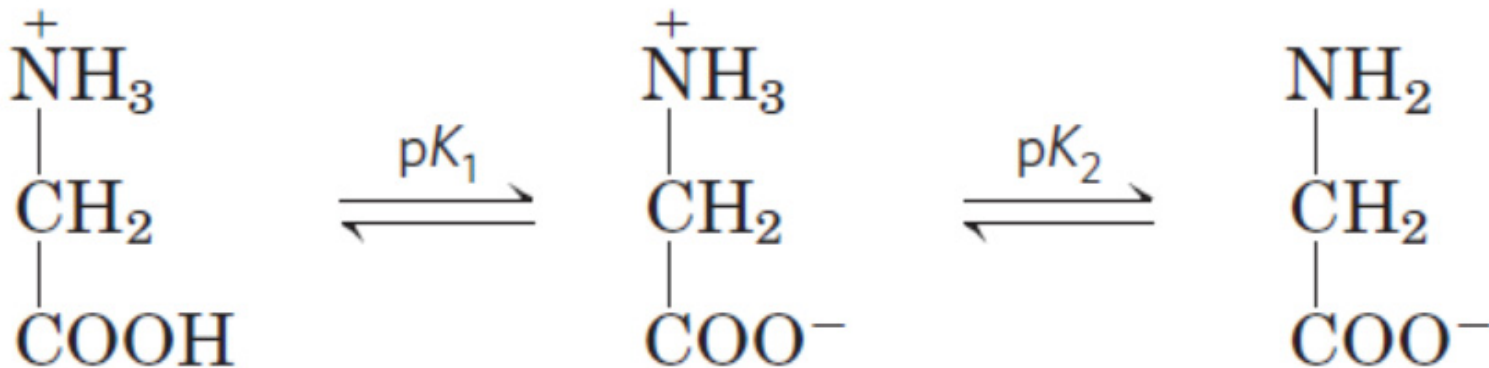
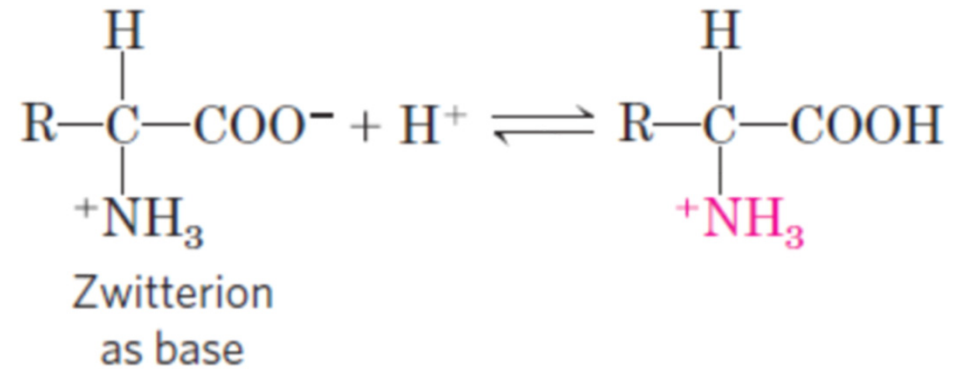
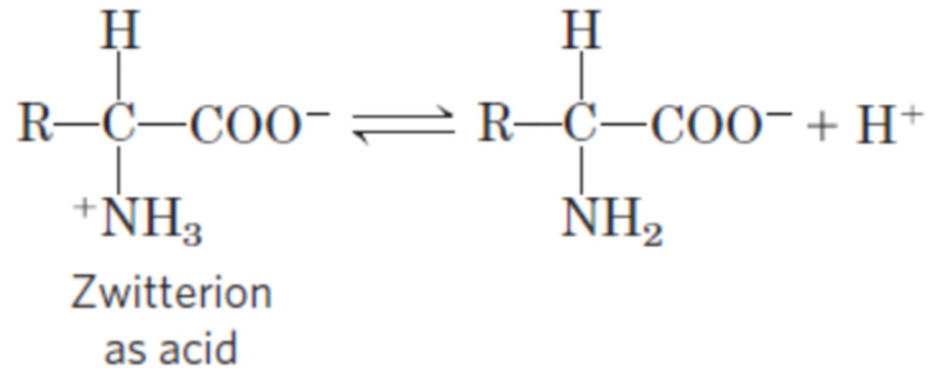
Nonionic form



Zwitterionic form



氨基酸的解离 (以甘氨酸为例)



氨基酸至少有两个 $\text{p}K_a$ ，记作 $\text{p}K_1$ 和 $\text{p}K_2$

氨基酸的酸性

甘氨酸
酸性强于乙酸，
碱性弱于甲胺

pK_a

2

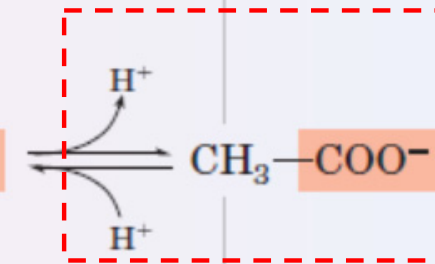
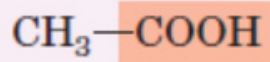
4

6

8

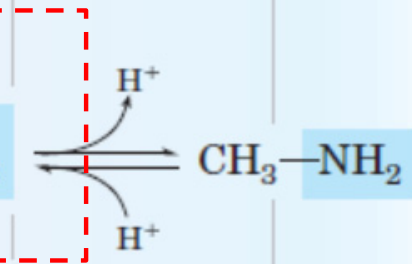
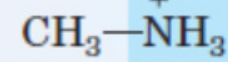
10

12



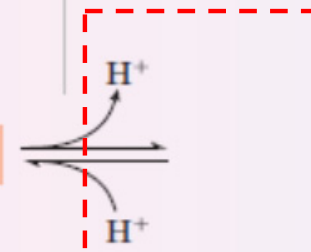
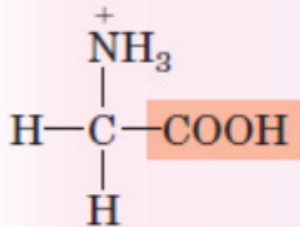
Acetic acid

The normal pK_a for a carboxyl group is about 4.8.



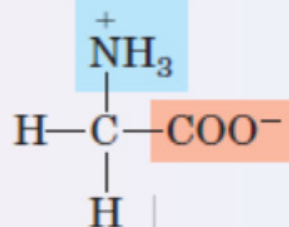
Methylamine

The normal pK_a for an amino group is about 10.6.

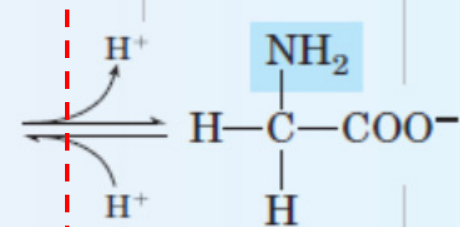


α-Amino acid (glycine)
pK_a = 2.34

Oppositely charged groups lower the pK_a by stabilizing the zwitterion.



不带电
(等电点)



α-Amino acid (glycine)
pK_a = 9.60

Electronegative oxygen atoms in the carboxyl group pull electrons away from the amino group, lowering its pK_a.

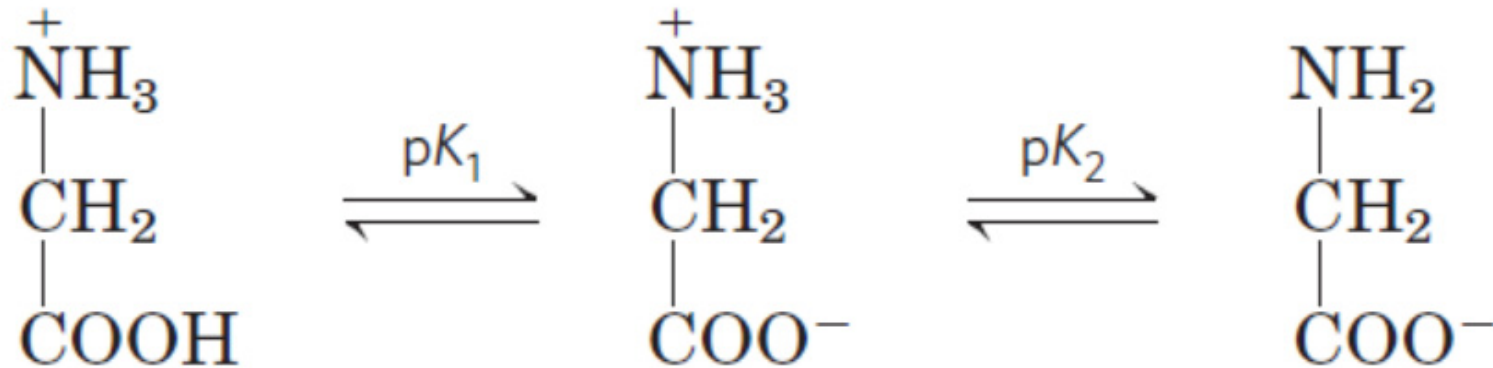
带正电

带负电

pH值影响解离和带电情况

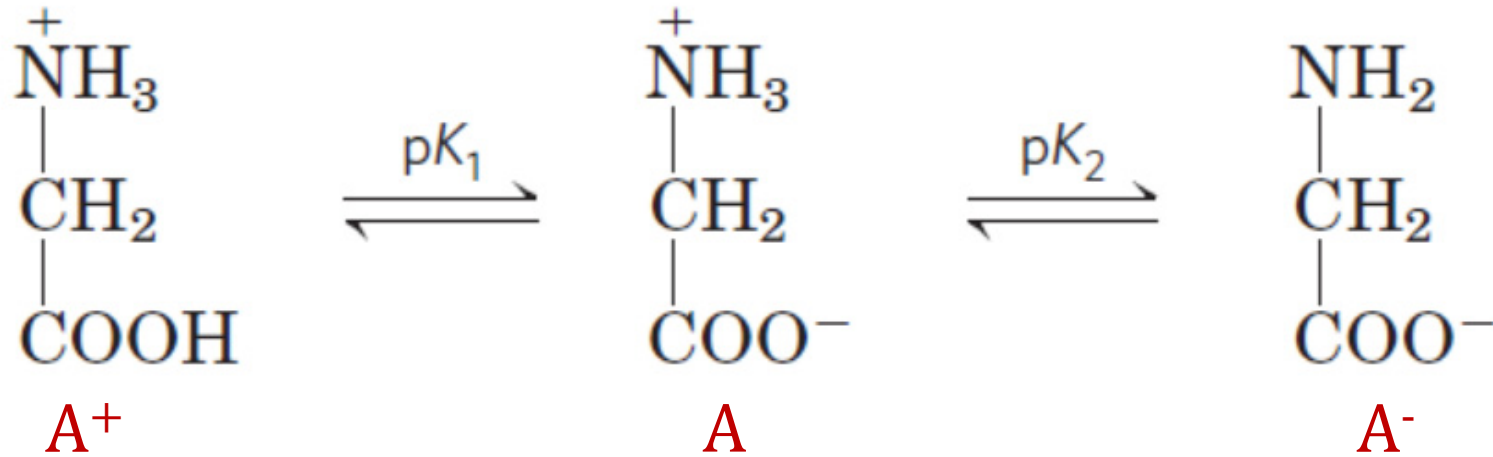
氨基酸的等电点 (重点)

- ▶ 调节氨基酸溶液的pH，使氨基酸分子所带净电荷为零。该pH值即为**氨基酸的等电点**（isoelectric point, pI）。
- ▶ 在等电点时，氨基酸处于兼性离子状态，在电场中不发生移动。



$$\text{pI} = \frac{1}{2} (\text{p}K_1 + \text{p}K_2)$$

pi的推导过程



$$K_1 = \frac{[\text{A}][\text{H}]}{[\text{A}^+]}$$

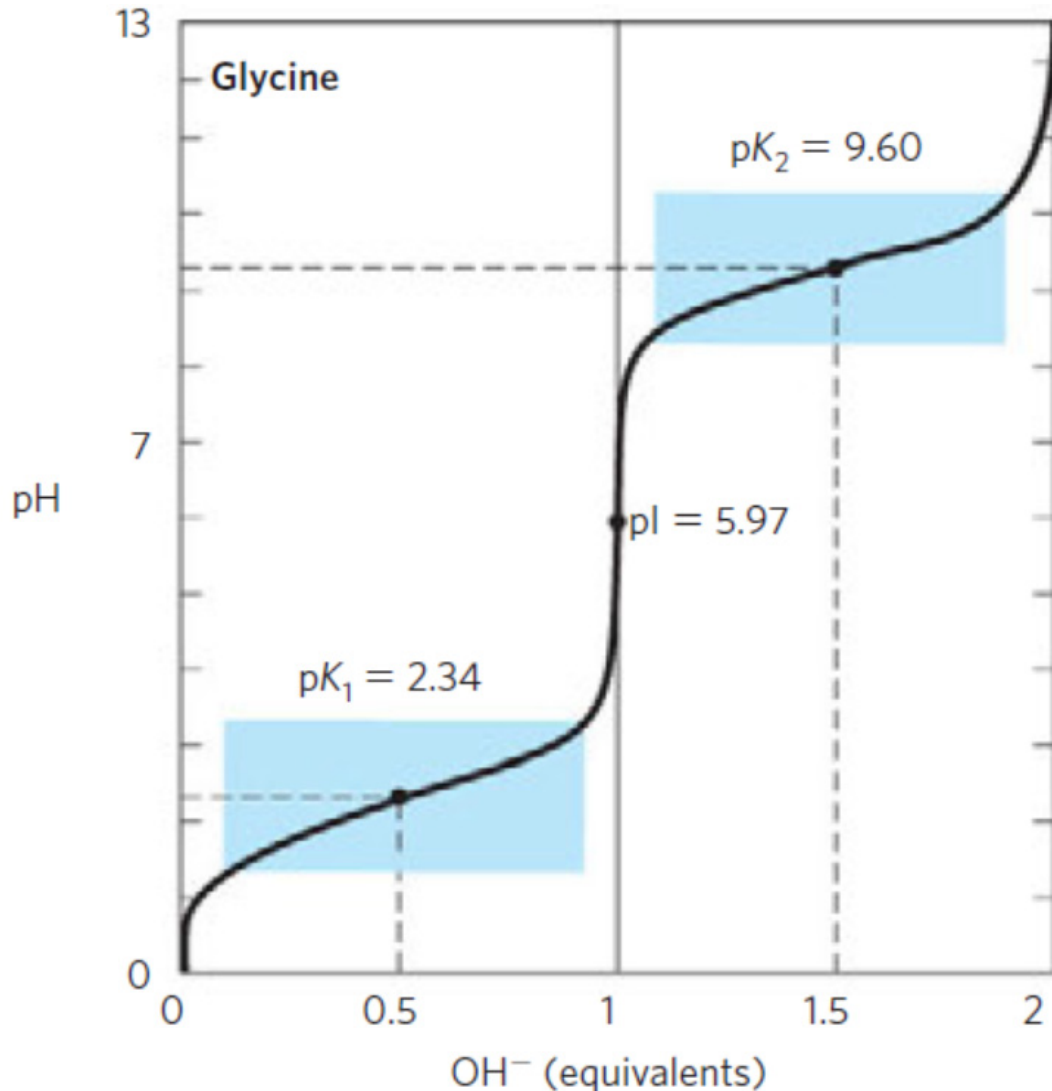
$$K_2 = \frac{[\text{A}^-][\text{H}]}{[\text{A}]}$$

$$K_1 K_2 = \frac{[\text{A}][\text{A}^-][\text{H}]^2}{[\text{A}^+][\text{A}]} = [\text{H}]^2$$

$$\text{pI} = \frac{1}{2} (\text{p}K_1 + \text{p}K_2)$$

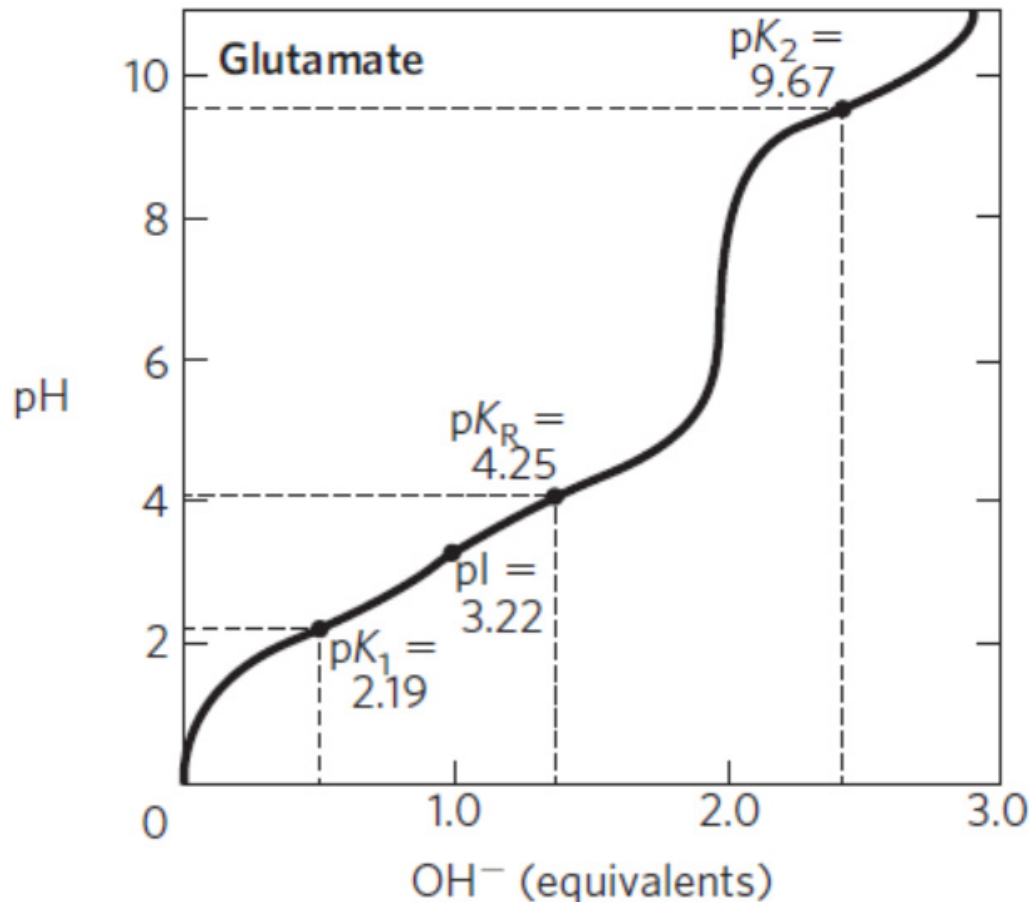
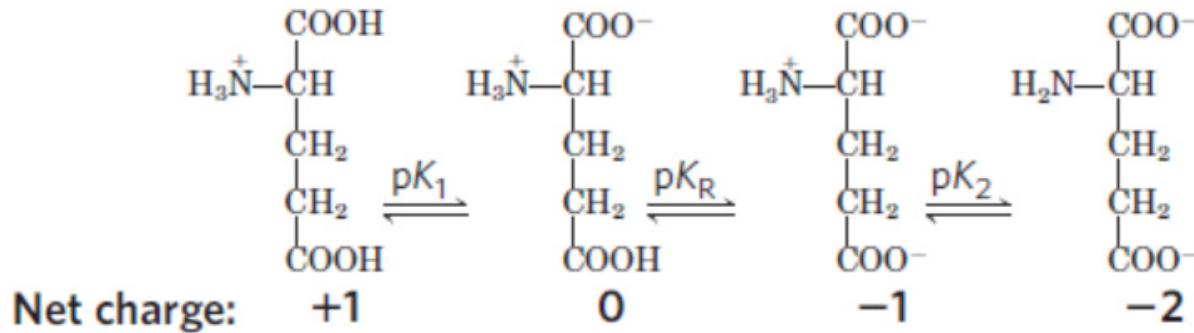
等电点时 $[\text{A}^+] = [\text{A}^-]$ ，两边同取对数

如何确定氨基酸的pK和pI值？ 氨基酸滴定



- ▶ 从左向右依次加入NaOH进行滴定
- ▶ 匀速加入NaOH
 - ▶ 在pK值附近，溶液pH变化最小（缓冲区）
 - ▶ 在pI值附近，溶液pH变化最大
- ▶ pH等于pI时氨基酸的溶解度最低

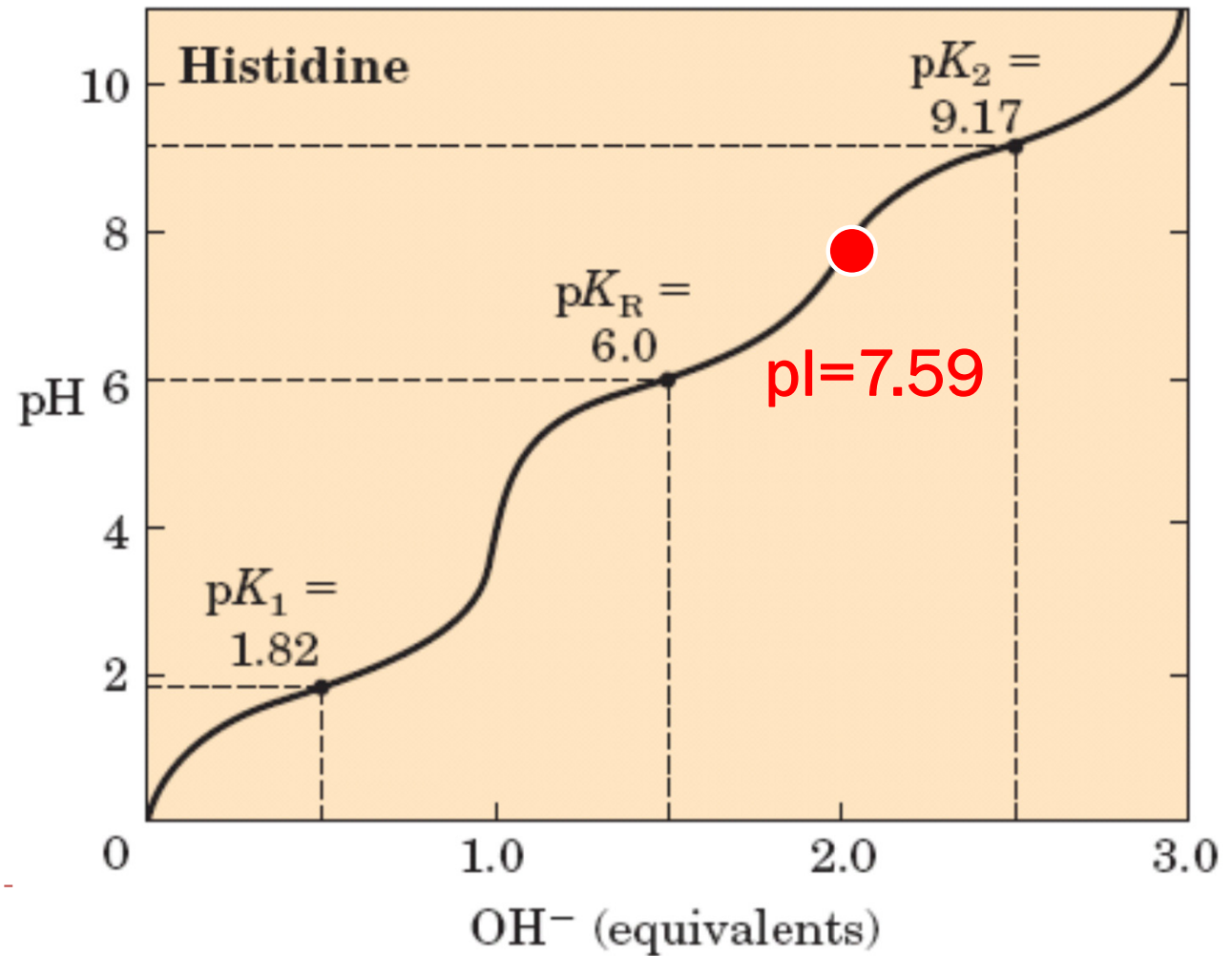
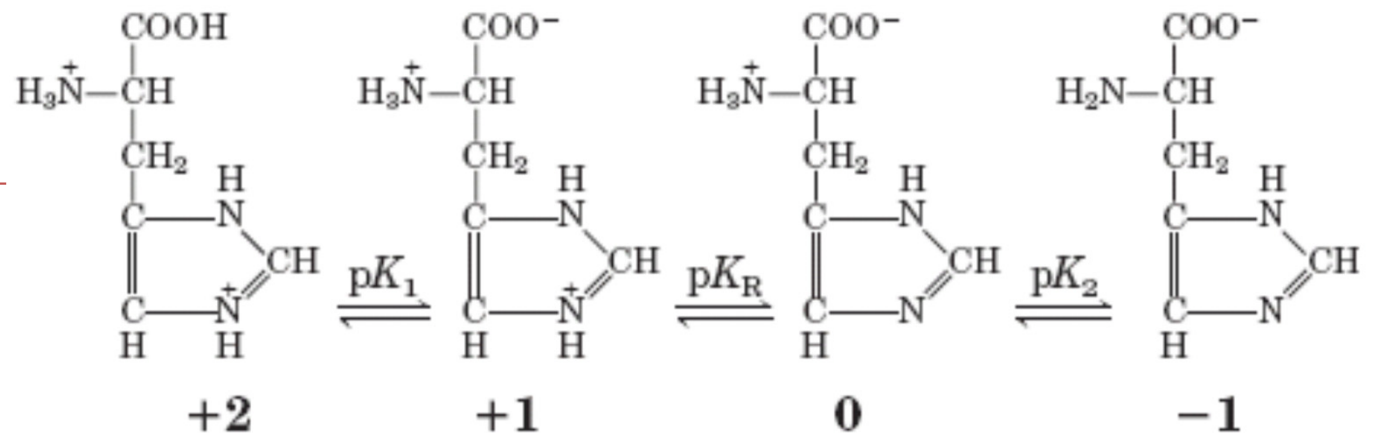
Q: 如何确定含有多个pK值氨基酸的pI?



▶ 找到净电荷为零的兼性离子，将其两侧的pK求和除2即为等电点。

▶ 对于三级以上的电离平衡，远处的pK值与兼性离子两侧的pK值相比差别较小时，不再适用，需要精确计算。

组氨酸的pI



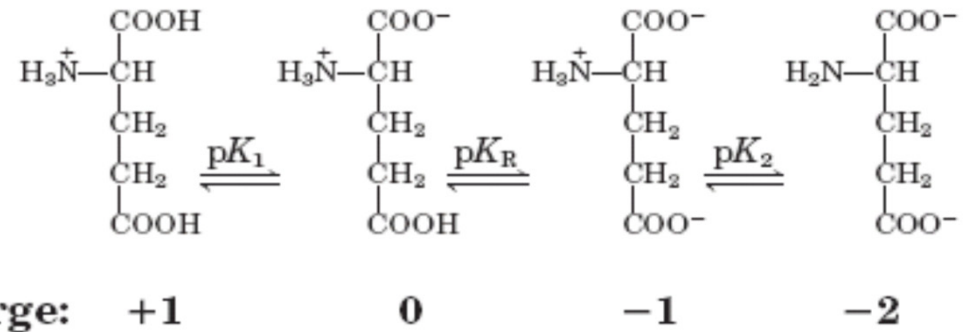
分子等电点在生物化学中的利用

- ▶ 在等电点时，由于静电作用，分子的**溶解度最小**，容易沉淀，进行分离纯化和鉴定
- ▶ 通过比较pH与pI，确定分子带电性质
- ▶ 调节溶液pH值，改变分子带电性质，据此对分子进行**分离纯化、电泳鉴定**等操作

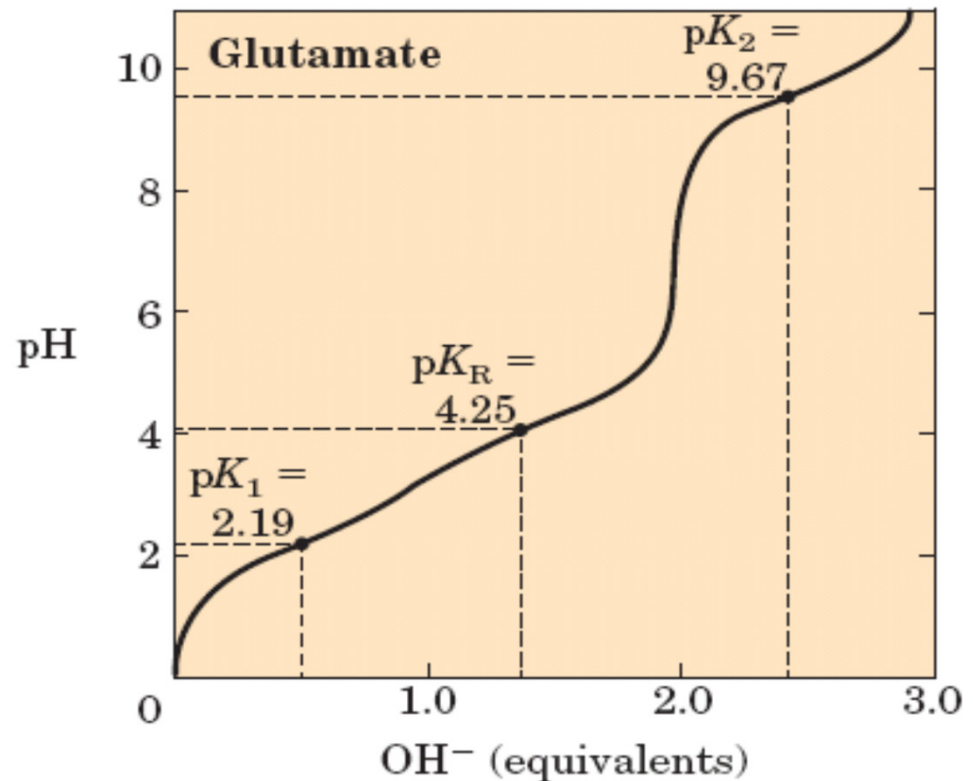


Q: 在pH10的Glu溶液中，下列哪一种结构占优？

- ▶ A. α -羧基和 γ -羧基都解离
- ▶ B. 羧基氨基都不解离
- ▶ C. 只 α -羧基解离
- ▶ D. 只 γ -羧基解离

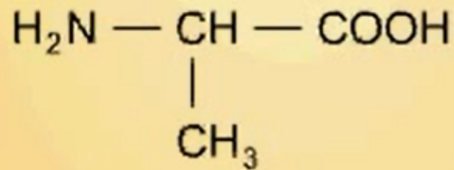


▶ A

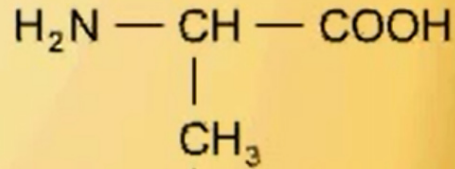


(a)

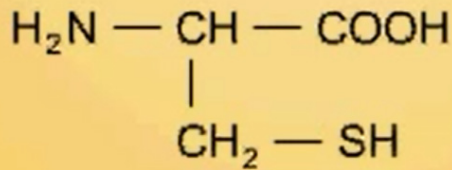
Q: 如何利用电荷分离Ala、Cys和His?



alanine (Ala)
pI = 6.11



histidine (His)
pI = 7.64

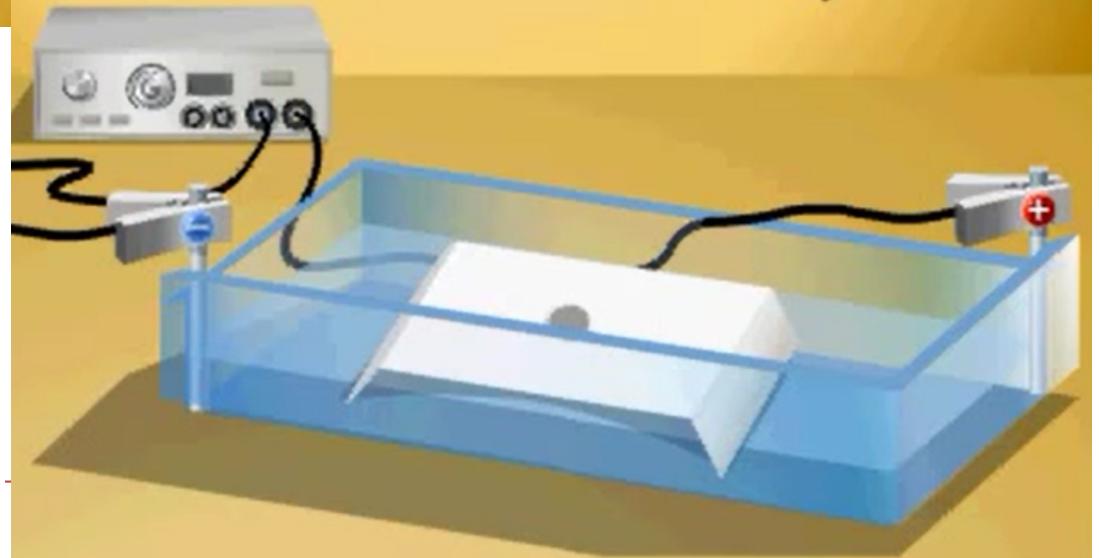
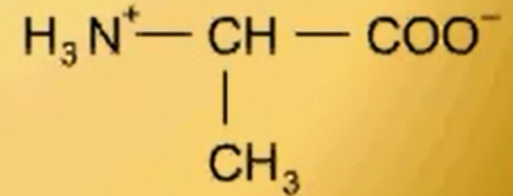


cysteine (Cys)
pI = 5.02

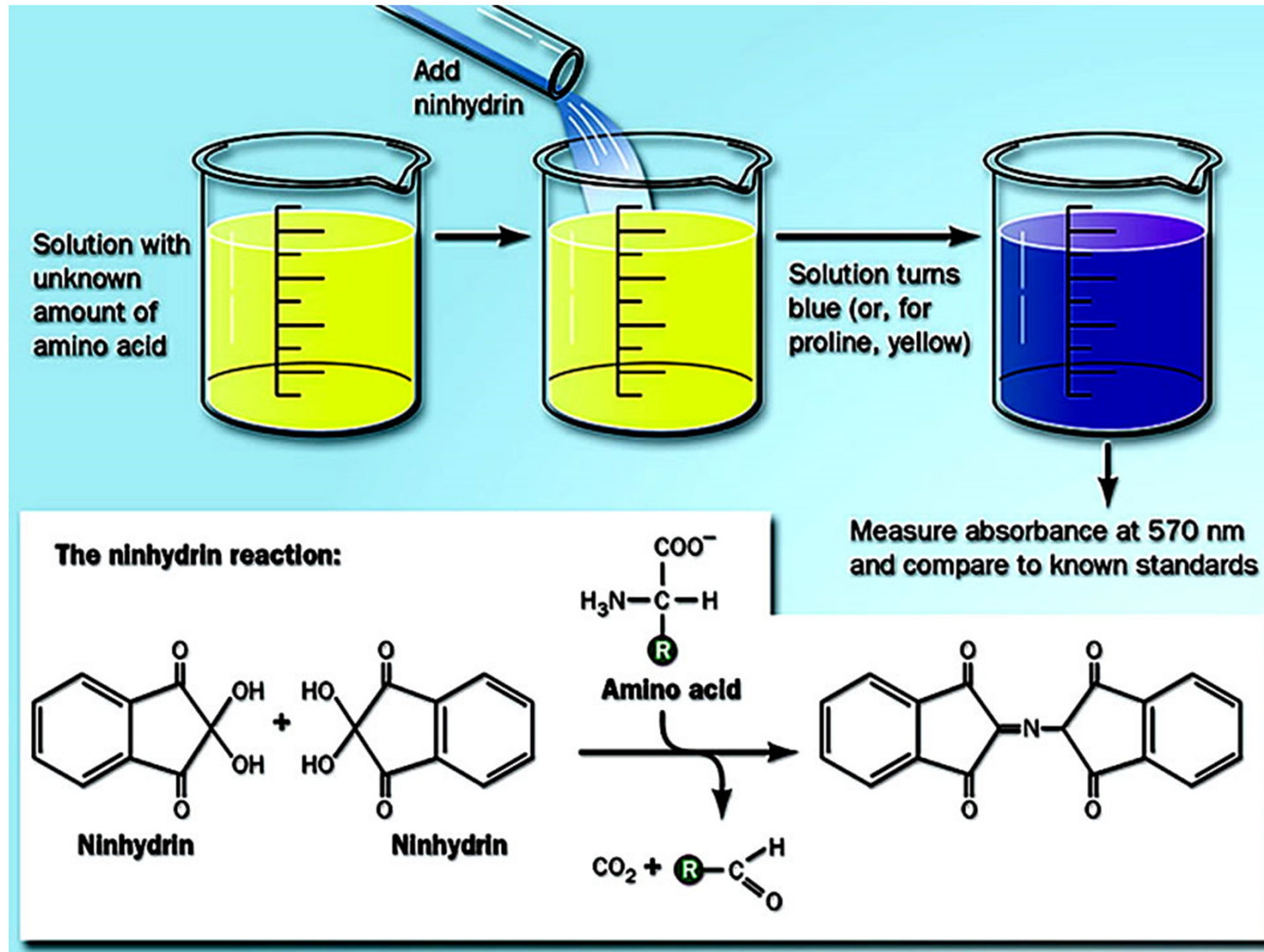
← His Cys →

pI of alanine \approx pH of the buffer solution

溶液pH 6.11

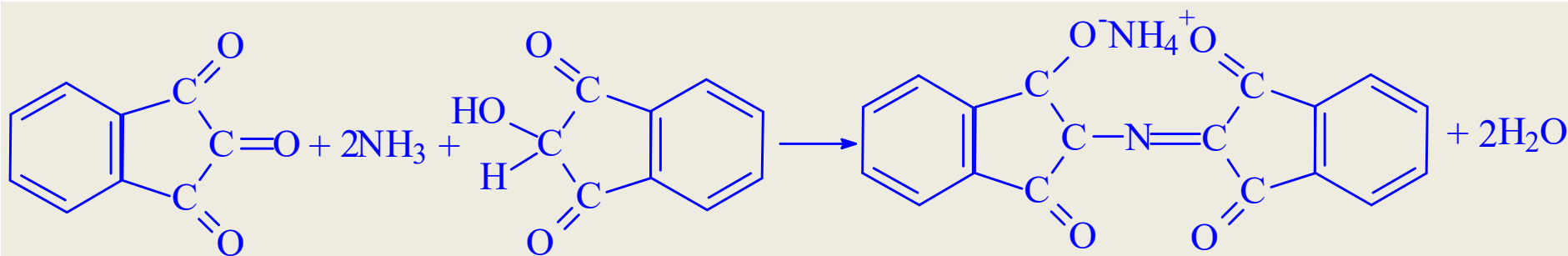
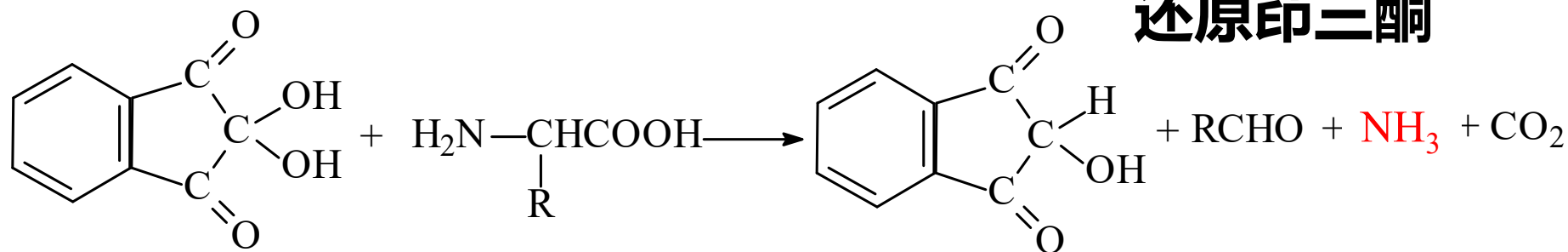
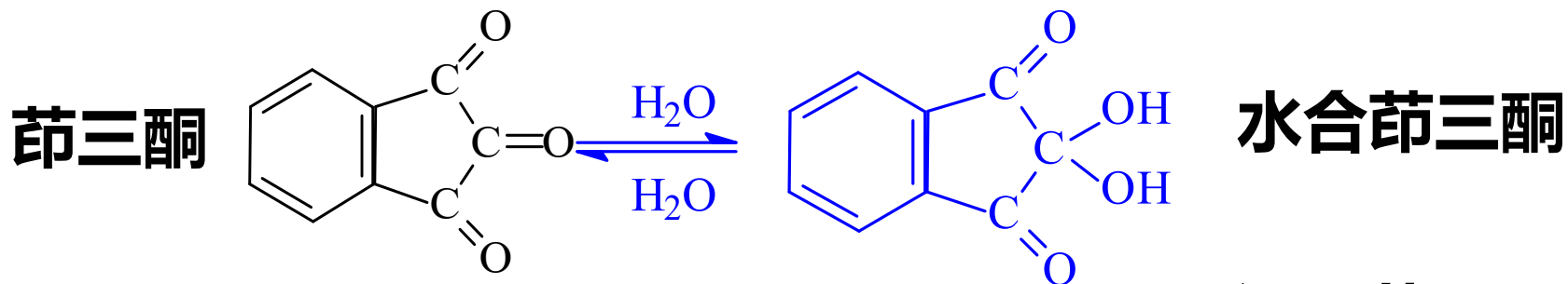


氨基酸的化学反应：水合茚三酮反应



α -氨基酸能与
水合茚三酮
(ninhydrin)
反应生成兰紫色物质

氨基酸的化学反应：水合茚三酮反应



水合茚三酮反应：氨基酸的定性与定量分析

- ▶ 大多数氨基酸的 α -氨基与茚三酮反应生成蓝紫色产物
- ▶ 用纸层析或柱层析把各种氨基酸分开，利用水合茚三酮显色可以定性鉴定并用分光光度法在570 nm定量测定各种氨基酸含量。
- ▶ 亚氨基酸——脯氨酸和羟脯氨酸，与水合茚三酮反应并不释放 NH_3 ，而直接生成亮黄色化合物，最大光吸收在440 nm。



氨基酸的其他化学反应

- ▶ 与亚硝酸反应生成 N_2 （Pro除外）
- ▶ 与2,4-二硝基氟苯（DNFB, Sanger试剂）
- ▶ 与异硫氰酸苯酯（PITC, Edman试剂）



氨基酸侧链的反应性

- ▶ S、T、Y 含有羟基，可以被磷酸化修饰
- ▶ H、K、R 碱性氨基酸
 - ▶ K和R的侧链氨基可被修饰，如酰基化、甲基化等
 - ▶ H的咪唑基pKa在生理pH值附近，可以作为质子供体或质子受体，常见于酶活中心。
- ▶ C的巯基可以生成二硫键

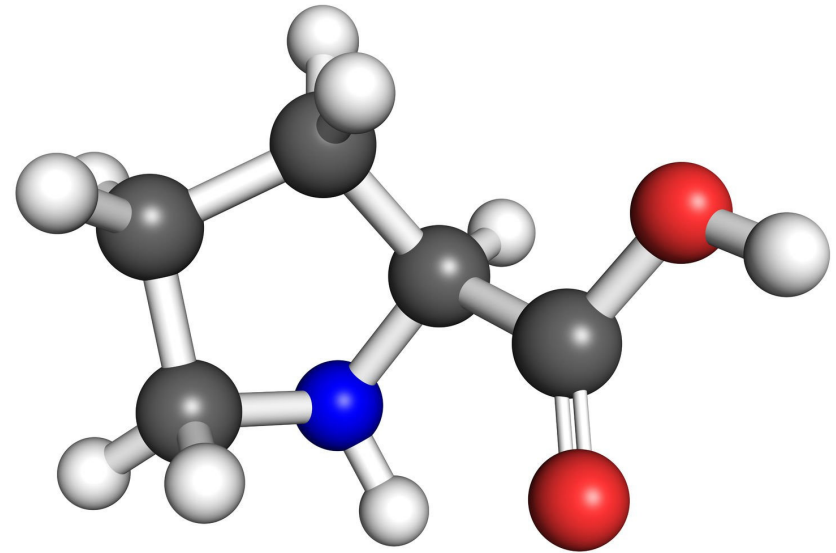


Properties and Conventions Associated with the Common Amino Acids Found in Proteins

Amino acid	Abbreviation/ symbol	M_r	pK_a values			pI	Hydropathy index*	Occurrence in proteins (%) [†]
			pK_1 (—COOH)	pK_2 (—NH ₃ ⁺)	pK_R (R group)			
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

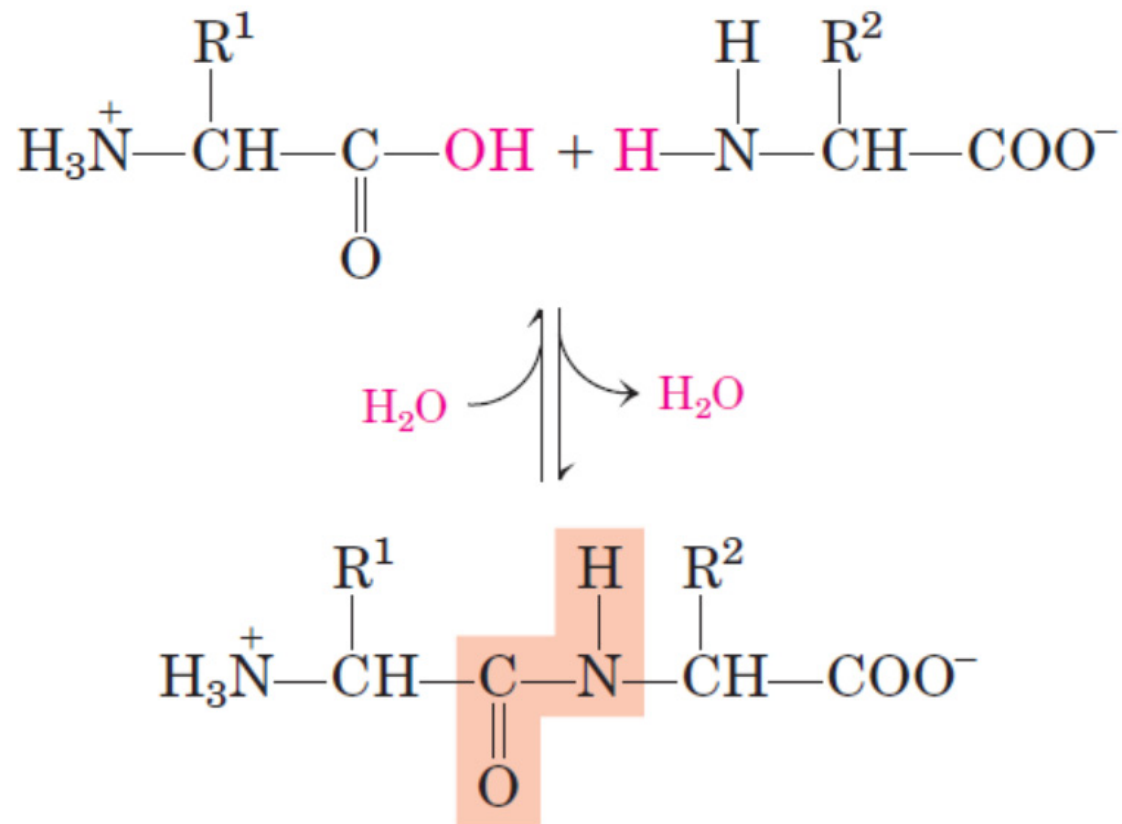
第2章 氨基酸、多肽与蛋白质

- ▶ 氨基酸的基本定义与特征
- ▶ 氨基酸的种类和结构
- ▶ 氨基酸的化学性质
- ▶ **多肽：氨基酸低聚物**
- ▶ **蛋白质：氨基酸高聚物**
- ▶ 蛋白质的研究方法



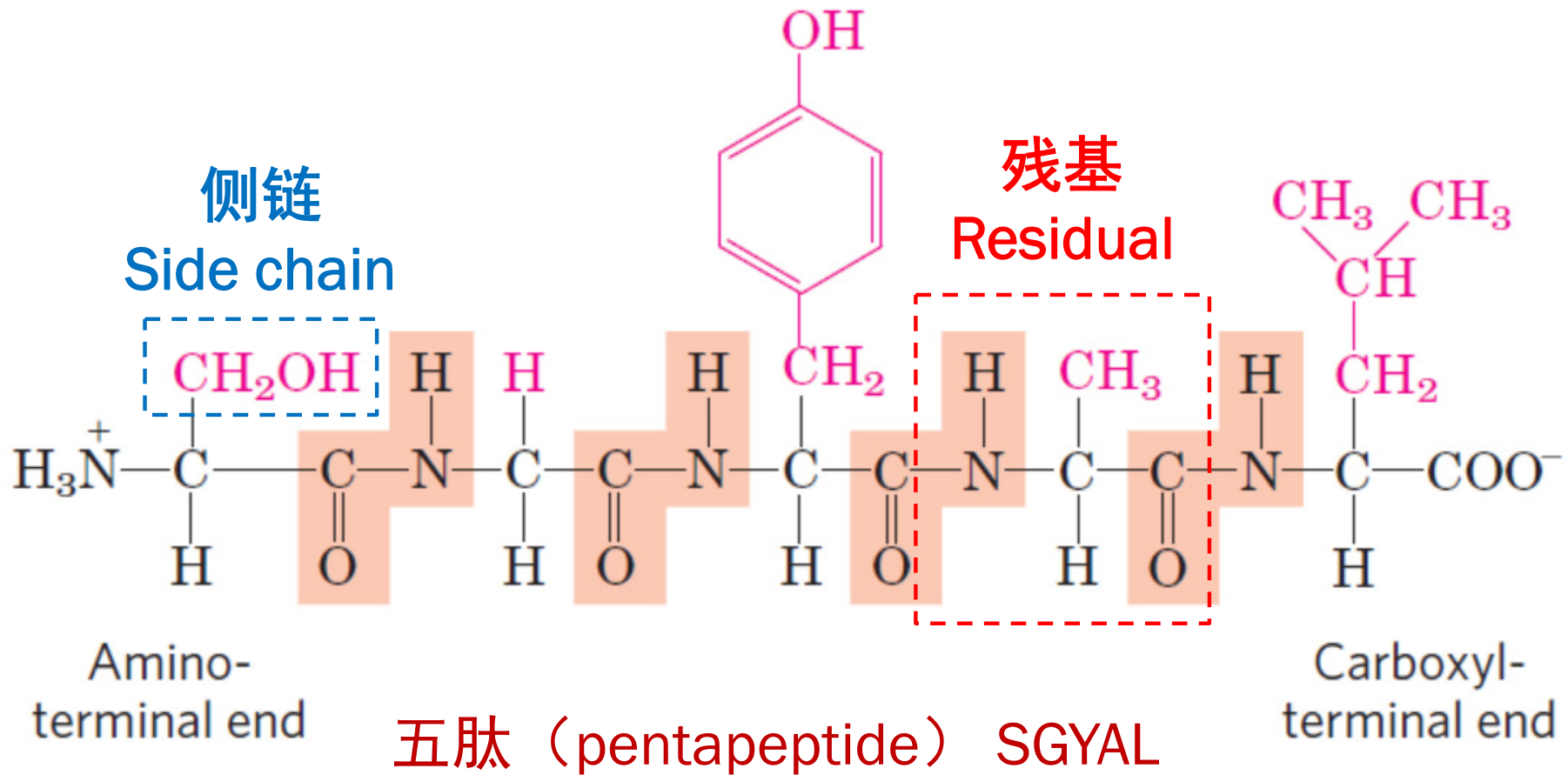
肽键 (peptide bond)

- ▶ α -氨基与 α -羧基缩合脱水生成
- ▶ 标准生化条件下不利于生成肽键
- ▶ 氨基酸需要活化后生成肽键



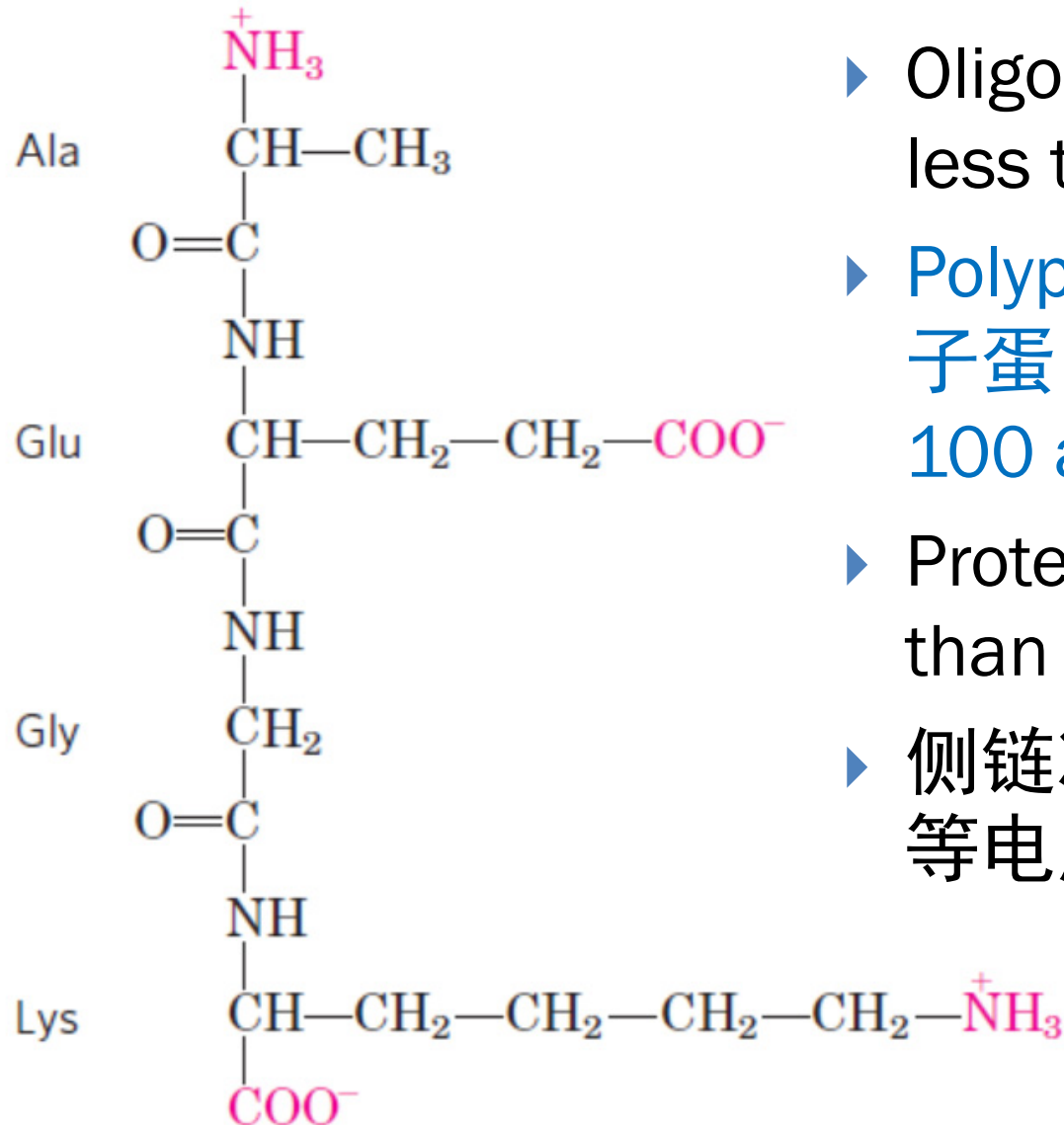
肽 (peptide)

- ▶ 长度、序列、方向、残基、侧链



肽分子量估算: $110 \times$ 残基数

肽和蛋白质的等电点



- ▶ Oligopeptides/寡聚肽: less than 50 aa residues
- ▶ Polypeptides/多肽（小分子蛋白质）: between 50-100 aa residues
- ▶ Protein/蛋白质: more than 100 aa residues
- ▶ 侧链决定了肽和蛋白质的等电点等诸多性质。

肽和蛋白质的等电点和分子量计算

▶ https://web.expasy.org/compute_pi/

Expasy

Compute pI/Mw

Compute pI/Mw tool

Compute pI/Mw is a tool which allows the computation of the theoretical pI (isoelectric point) and Mw (molecular weight) for a for user entered sequences [\[reference\]](#).

[Documentation](#) is available.

Compute pI/Mw for Swiss-Prot/TrEMBL entries or a user-entered sequence

Please enter one or more UniProtKB/Swiss-Prot protein identifiers (ID) (e.g. *ALBU_HUMAN*) or UniProt Knowledgebase accession numbers. Alternatively, enter a protein sequence in single letter code. The theoretical pI and Mw (molecular weight) will then be calculated.

Or upload a file from your computer, containing one or more sequences.

Resolution: Average or Monoisotopic

[Click here to compute pI/Mw](#) [Reset](#)

Compute pI/Mw

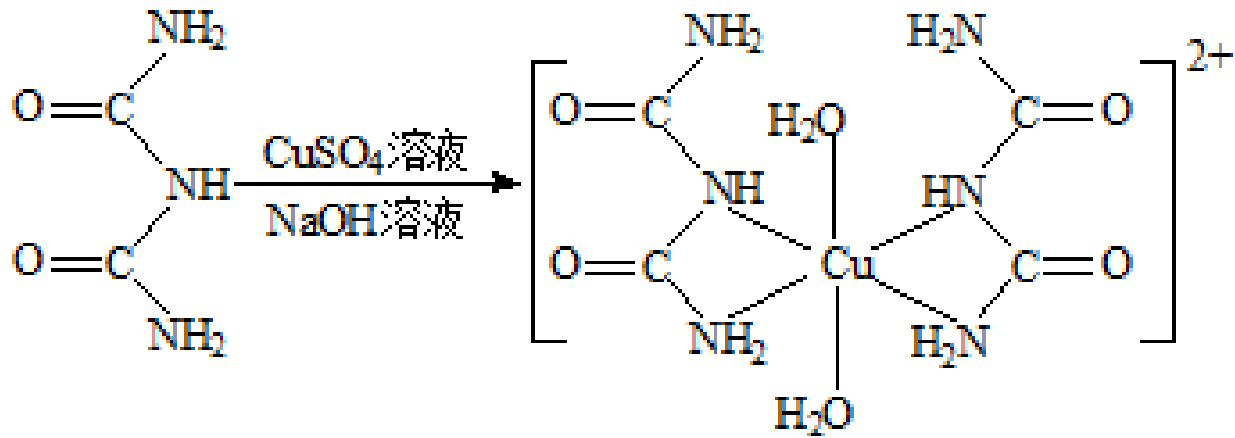
Theoretical pI/Mw (average) for the user-entered sequence:

SGYAL

Theoretical pI/Mw: 5.24 / 509.56

肽的鉴定：双缩脲反应

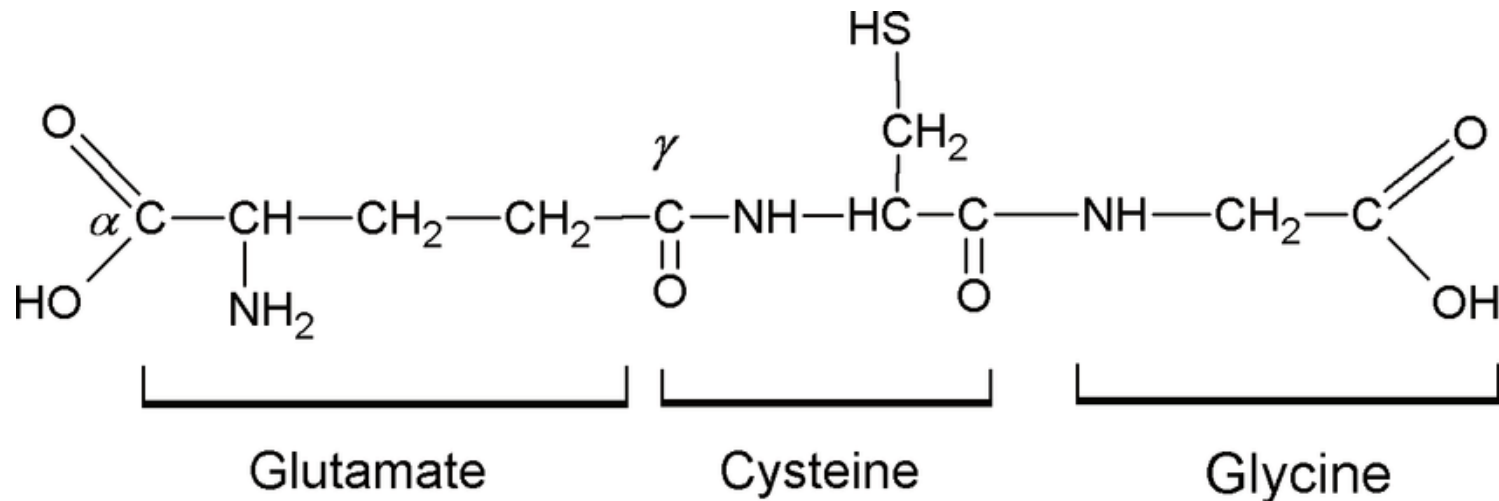
- ▶ 多肽和蛋白质的定性反应
 - ▶ 碱性条件下，反应物与铜离子生成紫红色络合物
 - ▶ 氨基酸和二肽不能发生此反应



多肽的分类

▶ 天然肽

- ▶ 核糖体合成肽（催产素 oxytocin）
- ▶ 酶合成肽（谷胱甘肽 glutathione, GSH）

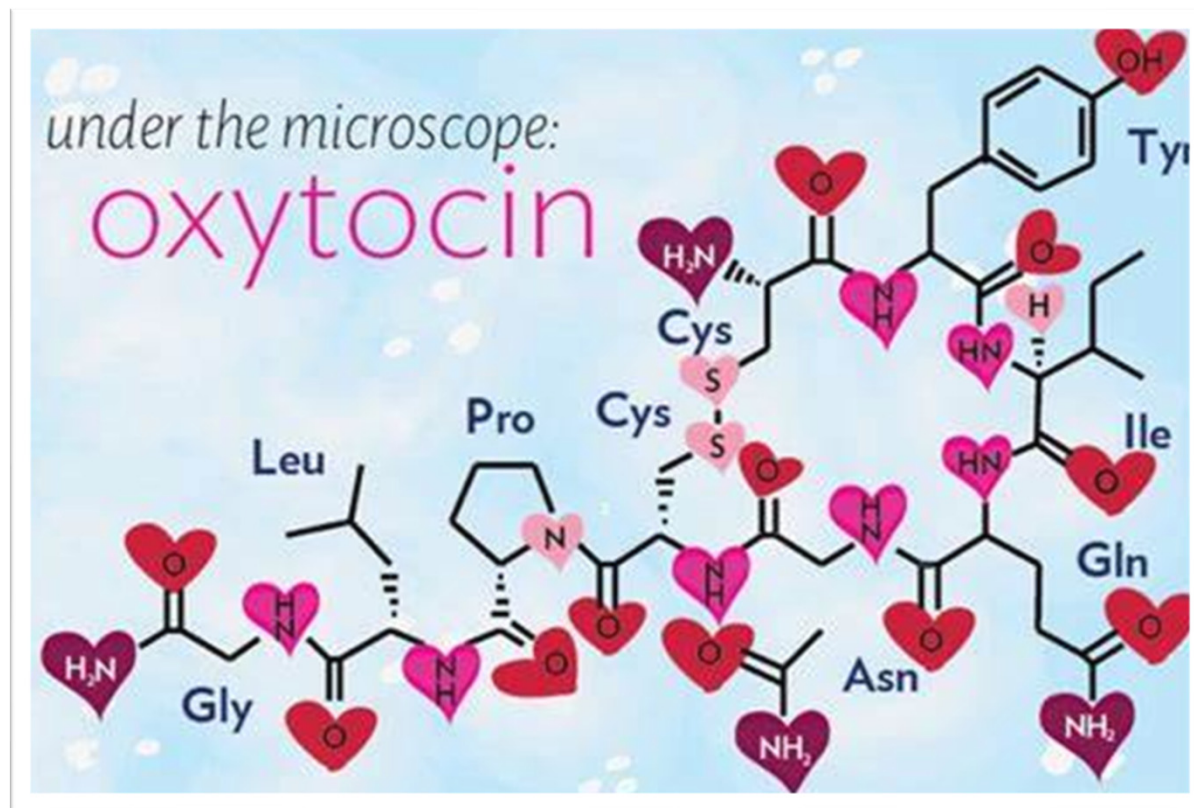


▶ 人工肽（阿斯巴甜 aspartame）



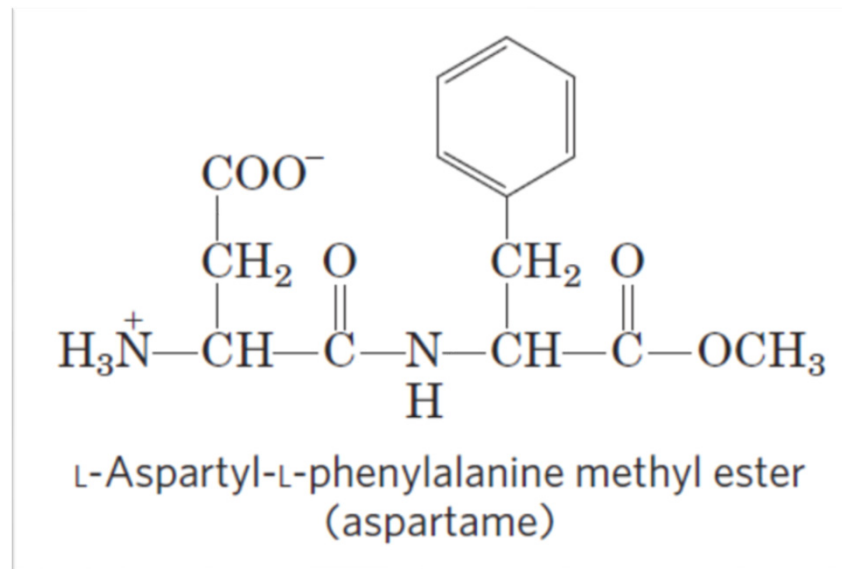
生物活性肽 (Biologically active peptides)

- ▶ 催产素是大脑产生的一种激素，男女都有。
- ▶ 催产素还是诱发人与人之间亲密关系的因素之一。
- ▶ 九肽 Cys-Tyr-Ile-Gln-Asn-Cys-Pro-Leu-Gly (含二硫键)



生物活性肽 (Biologically active peptides)

- ▶ 阿斯巴甜(Aspartame)是一种人造的糖替代品，是现在世界上6000多种食品和饮料的主要成分。又称甜味素、蛋白糖、天冬甜母、天冬甜精、天苯糖等。
- ▶ 由必需氨基酸之一的L-苯丙氨酸先与甲醇酯化后再和氨基酸L-天冬氨酸(Asp, D)缩合酰胺化所产生的一种经甲基酯化后的二肽类化合物。



各种蛋白大小差异巨大

TABLE 3-2 Molecular Data on Some Proteins

	Molecular weight	Number of residues	Number of polypeptide chains
Cytochrome <i>c</i> (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) 血红蛋白	64,500	574	4
Serum albumin (human)	66,000	609	1
Hexokinase (yeast)	107,900	972	2
RNA polymerase (<i>E. coli</i>)	450,000	4,158	5
Apolipoprotein B (human)	513,000	4,536	1
Glutamine synthetase (<i>E. coli</i>)	619,000	5,628	12
Titin (human)	2,993,000	26,926	1

缀合蛋白质 (conjugated proteins)

- ▶ 简单蛋白质 (simple proteins)
- ▶ 缀合蛋白质含有非氨基酸成分的辅因子 (cofactors)

TABLE 3-4 Conjugated Proteins

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

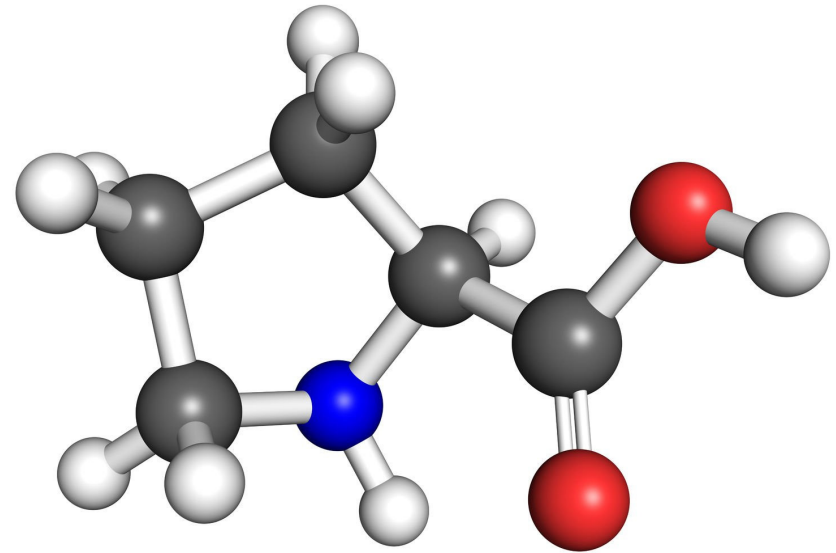
辅因子 (cofactors)

- ▶ 辅基 (prosthetic group)
 - ▶ 结合非常紧密，有时用共价结合，用剧烈的方法才能除去，往往引起蛋白结构改变
- ▶ 辅酶 (coenzyme)
 - ▶ 非共价结合，透析和超滤可以除去



第2章 氨基酸、多肽与蛋白质

- ▶ 氨基酸的基本定义与特征
- ▶ 氨基酸的种类和结构
- ▶ 氨基酸的化学性质
- ▶ 多肽：氨基酸低聚物
- ▶ 蛋白质：氨基酸高聚物
- ▶ 蛋白质的研究方法



利用蛋白质的物理化学性质对其进行研究

蛋白大小

蛋白电荷

亲疏水性

分子互作

▶ 蛋白质的分离与纯化

▶ 沉淀与离心、透析与超滤、层析、.....

▶ 蛋白质的鉴定

▶ 电泳、免疫印迹、化学测序、质谱分析、.....

▶ 蛋白质序列分析

▶ 序列预测、序列比对、.....

▶

蛋白质的分离与纯化

- ▶ 一个中心：保护蛋白活性（稳定性）
 - ▶ 低温、浓度、pH、缓冲液、蛋白酶抑制剂、避免反复冻融、EDTA螯合重金属、灭菌溶液、.....

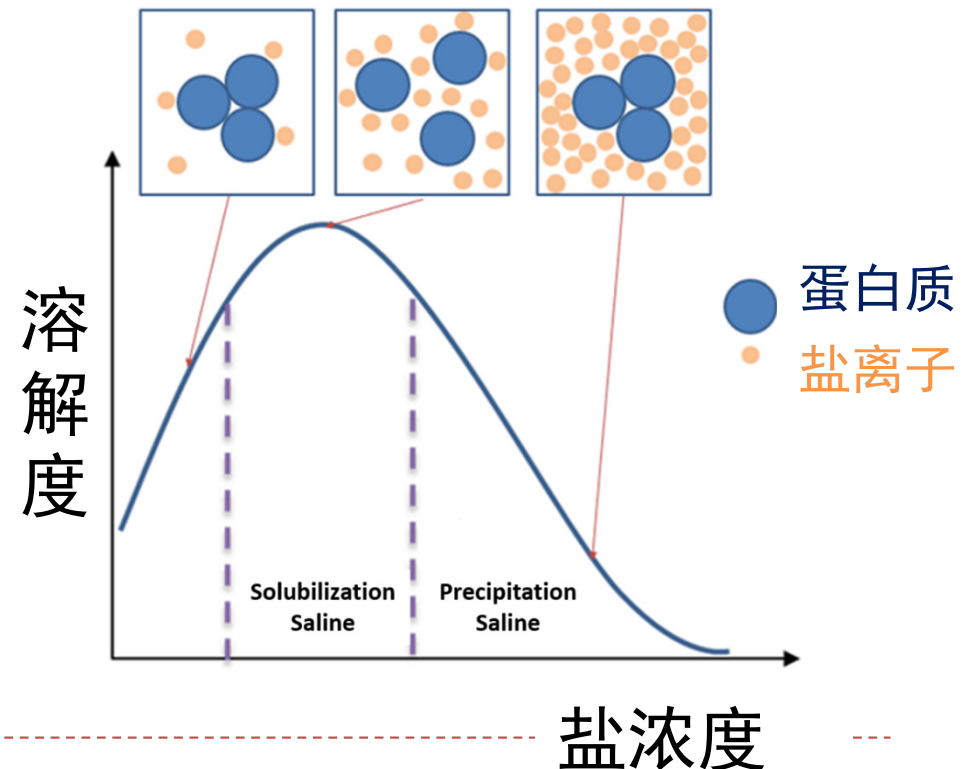
▶ 三个问题

- ▶ 纯化蛋白的目的是什么？
- ▶ 测定活性或鉴定蛋白质的方法是什么？
- ▶ 初始原料是什么？



蛋白质分离纯化方法1：沉淀与离心

- ▶ 盐析 (salting out)
 - ▶ 硫酸铵等中性盐对蛋白质的溶解度有显著影响。
 - ▶ 一般在低盐浓度下随着盐浓度升高，蛋白质分子表面的电荷会增加，增强蛋白质分子与水分子的作用，蛋白质的溶解度增加，此称盐溶。
 - ▶ 当盐浓度继续升高时，蛋白质的溶解度不同程度下降并先后析出，这种现象称盐析。



离心 (centrifuge)

- ▶ 水平转子离心
- ▶ 角转子离心



蛋白质分离纯化方法1：沉淀与离心

▶ 沉降离心

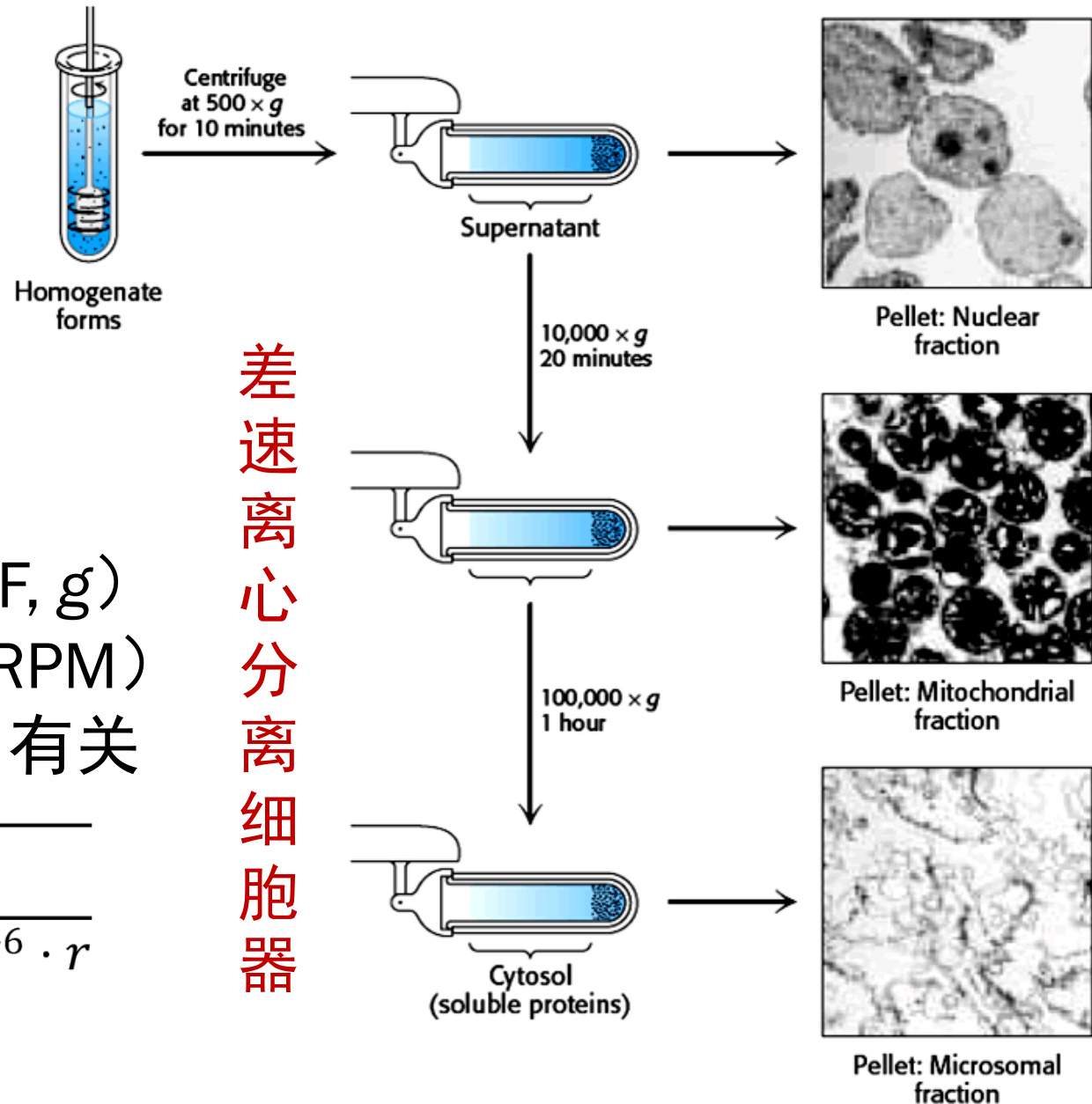
- ▶ 常规离心
- ▶ 差速离心

▶ 梯度离心

- ▶ 密度梯度离心
- ▶ 平衡梯度离心

- ▶ 相对离心力 (RCF, g) 与每分钟转速 (RPM) 和转子半径 (r) 有关

$$RPM = \sqrt{\frac{RCF}{1,118 \cdot 10^{-6} \cdot r}}$$



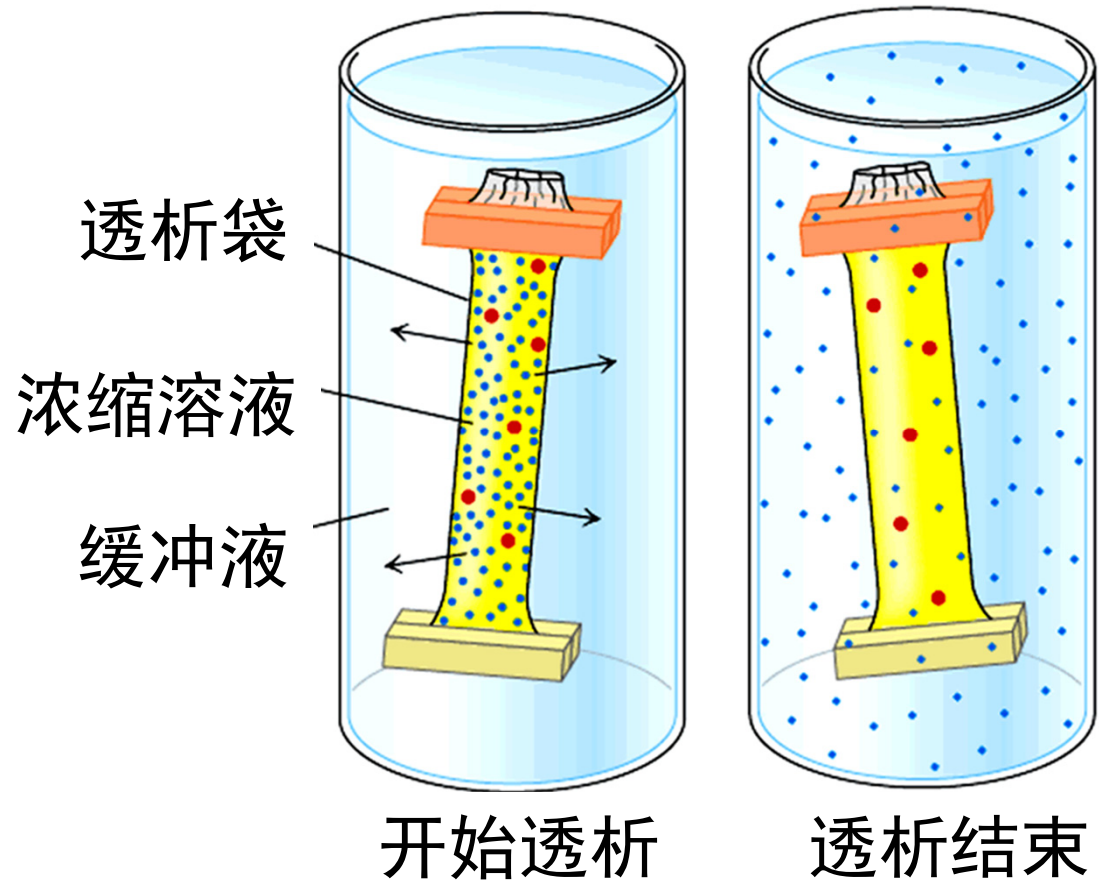
蛋白质分离纯化方法2：透析与超滤

▶ 透析 (dialysis)

- ▶ 小分子经过半透膜扩散到水（或缓冲液），与生物大分子分开。

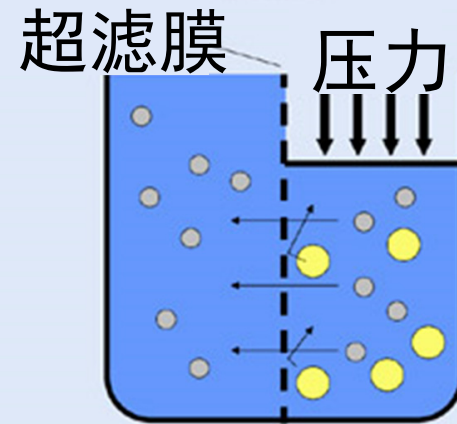
▶ 反透析

- ▶ 装有蛋白溶液的透析袋放入聚乙二醇，被吸水浓缩。



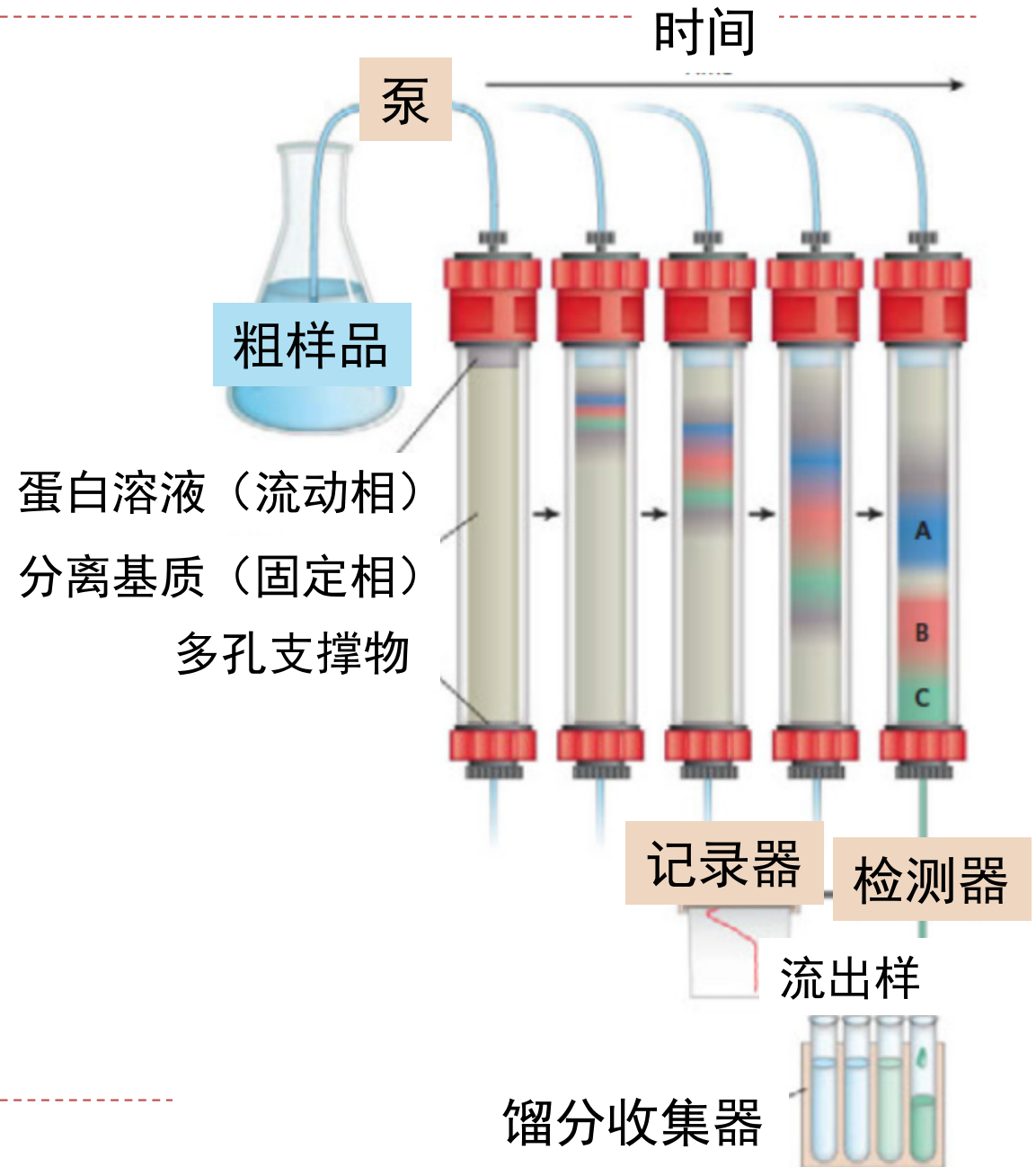
蛋白质分离纯化方法2：透析与超滤

- ▶ 超滤（ultrafiltration）
 - ▶ 可在加压情况下，迫使大分子留在膜上的溶液，水和小分子通过膜，达到浓缩蛋白质、脱盐或只换缓冲液的目的。
 - ▶ 可以离心代替加压。



蛋白质分离纯化方法3：层析

- ▶ 层析（又称色谱，Chromatography）
 - ▶ 利用不同物质在两相介质（固定相和流动相）中具有不同的分配系数达到分离目的。
 - ▶ 影响因素包括：分子**电荷**不同、**大小**不同、**亲疏水性**不同、是否能与特定的相具有**特异性结合**（亲和力）等。
- ▶ 柱层析

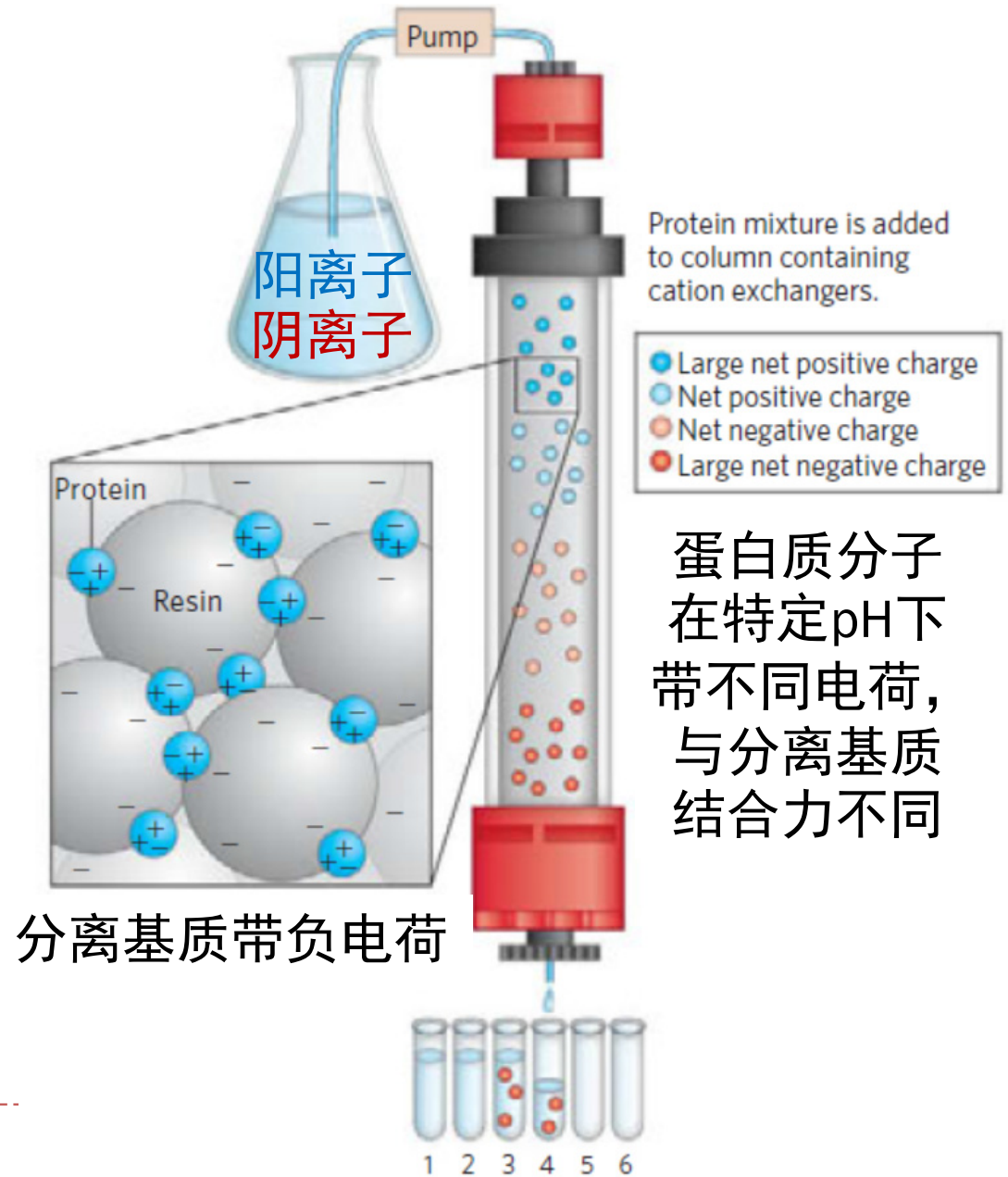


1. 离子交换层析 (Ion-exchange chromatography)

- ▶ 以**离子交换剂**为固定相，依据流动相中的组分离子与分离基质上的平衡离子进行可逆交换时的结合力大小的差别进行分离。

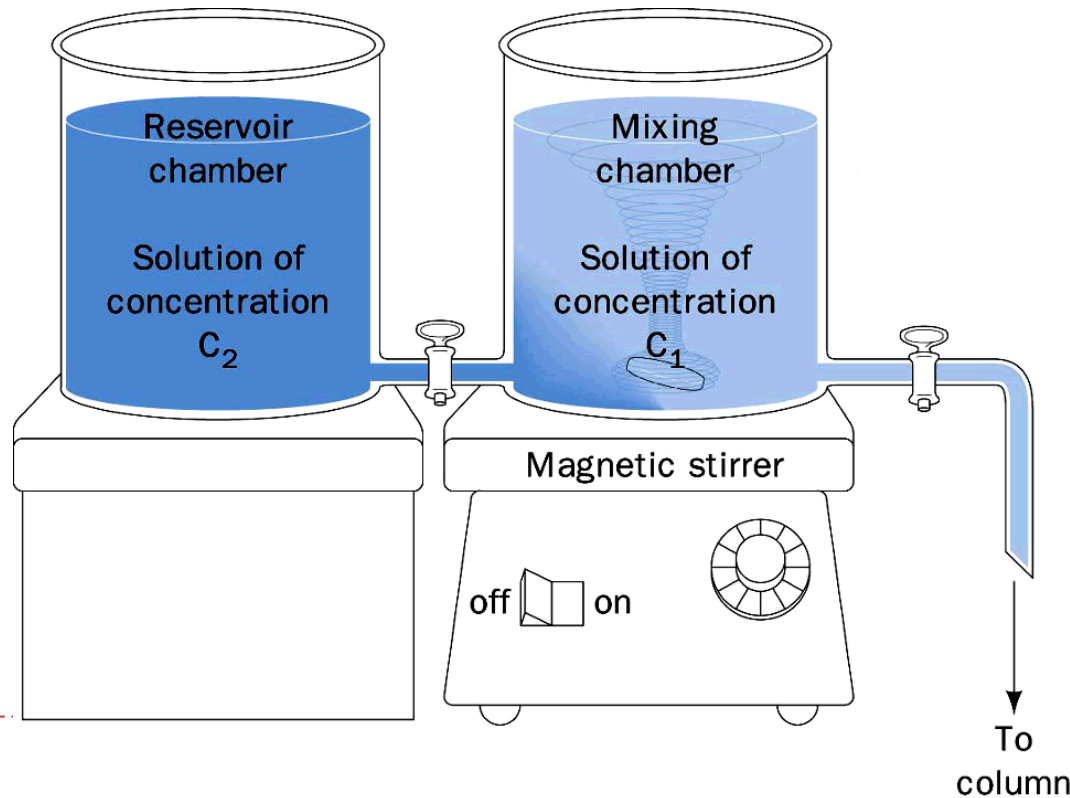
- ▶ 阳离子交换层析（与基质交换阳离子，阴离子先出）
- ▶ 阴离子交换层析（与基质交换阴离子，阳离子先出）

- ▶ 右图为阳离子交换



洗脱：逐步提高pH和盐浓度

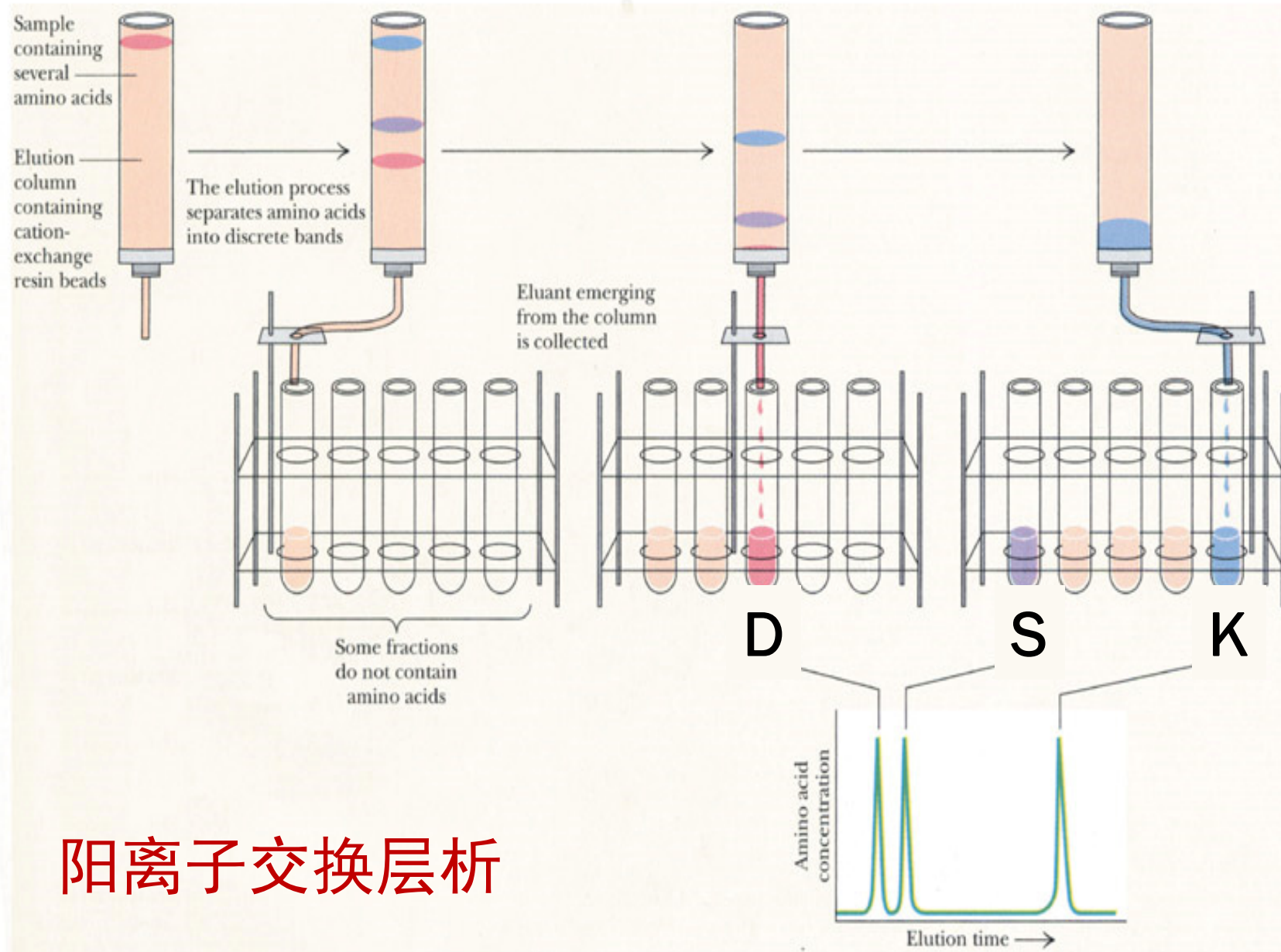
- ▶ 为了使氨基酸从树脂柱上洗脱下，需要降低它们之间的亲和力，有效的方法是逐步提高洗脱剂的pH和盐浓度，这样各种氨基酸将以不同的速度被洗脱下来。



用于制备线性pH
或盐浓度梯度

$$\text{pH} = \text{p}K_a + \lg \frac{[\text{A}^-]}{[\text{HA}]}$$

Q: 这是什么离子交换层析?



阳离子交换层析

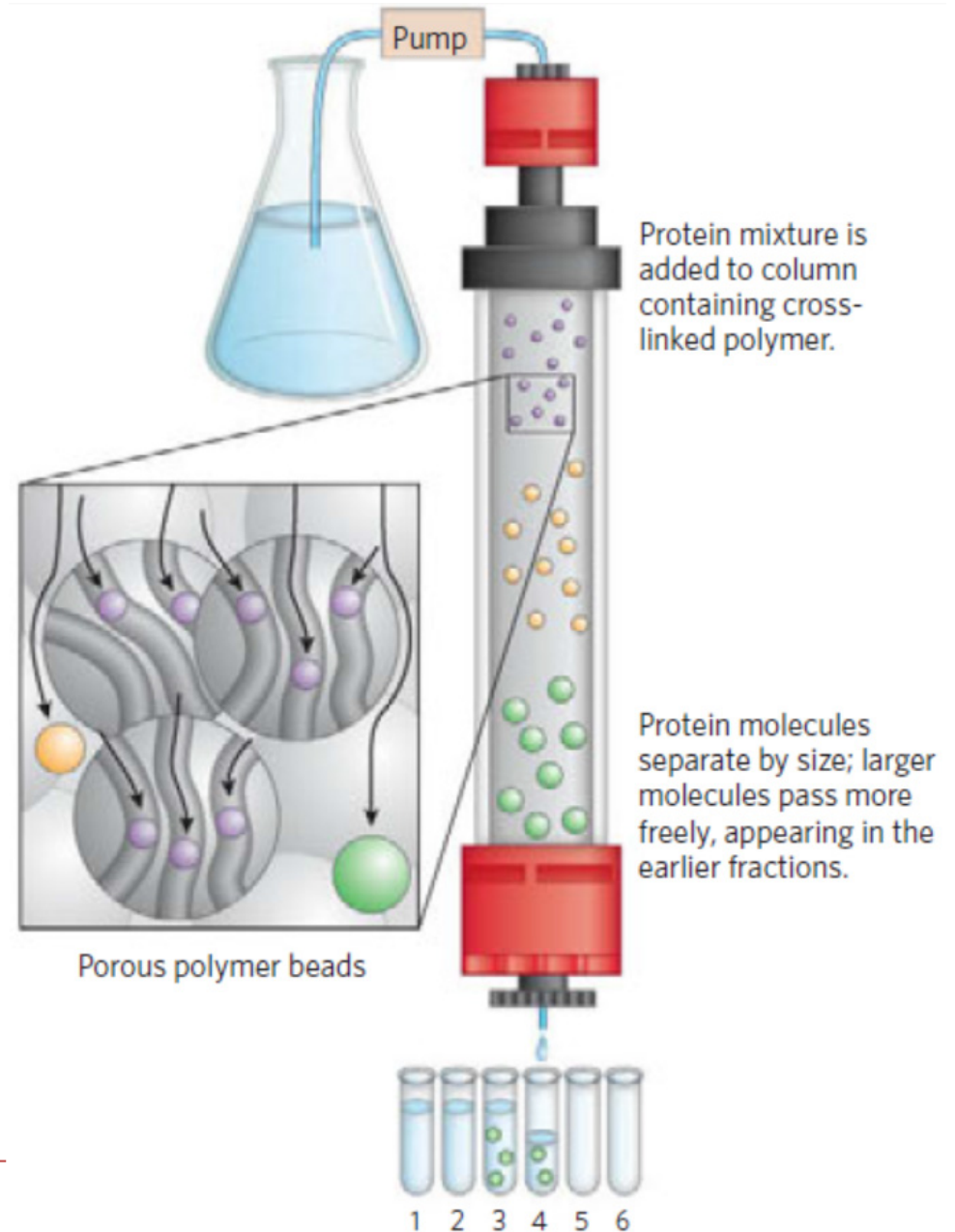
Q: 推测洗脱顺序

- ▶ 某氨基酸溶液中，有四种氨基酸A、B、C、D，其等电点分别为10、4、7、5，如不考虑次要因素：
 - ▶ 1) 采用阳离子交换柱并用盐浓度或pH梯度进行洗脱，将按何种顺序被洗脱？
 - ▶ 2) 采用阴离子交换柱并用盐浓度或pH梯度进行洗脱，将按何种顺序被洗脱？
- ▶ 1) BDCA
- ▶ 2) ACDB

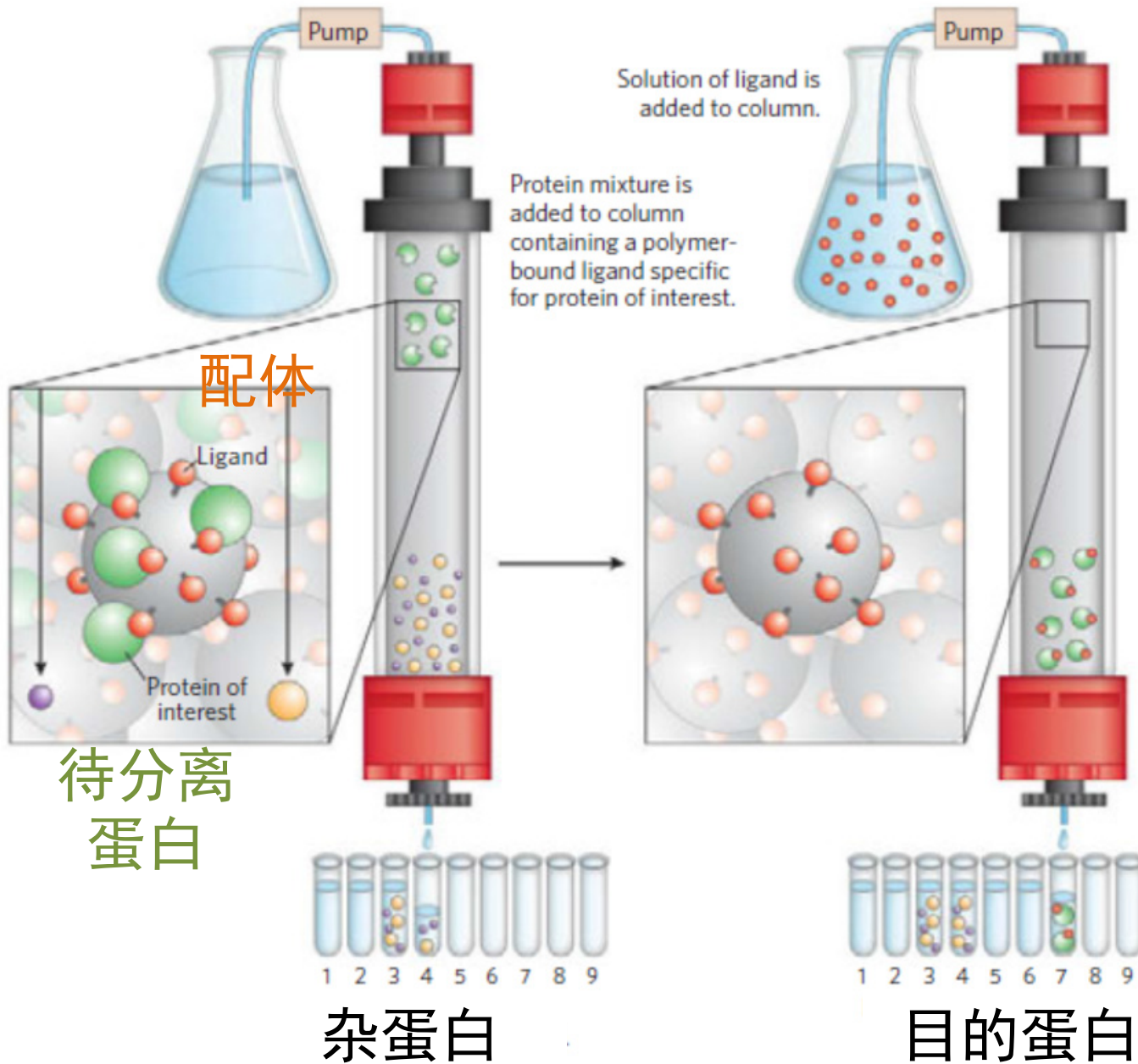


2. 排阻层析 (Size-exclusion chromatography)

- ▶ 又称为凝胶过滤 (Gel filtration)
- ▶ 根据样品分子的大小尺寸进行分离的层析技术。
- ▶ 分离基质中含有微孔
 - ▶ 大分子路径短，先流出
 - ▶ 小分子路径长，后流出



3. 亲和层析 (Affinity chromatography)



▶ 分离基质上的分子或基团（配体）可以**特异结合**待纯化的蛋白

▶ 洗脱方式

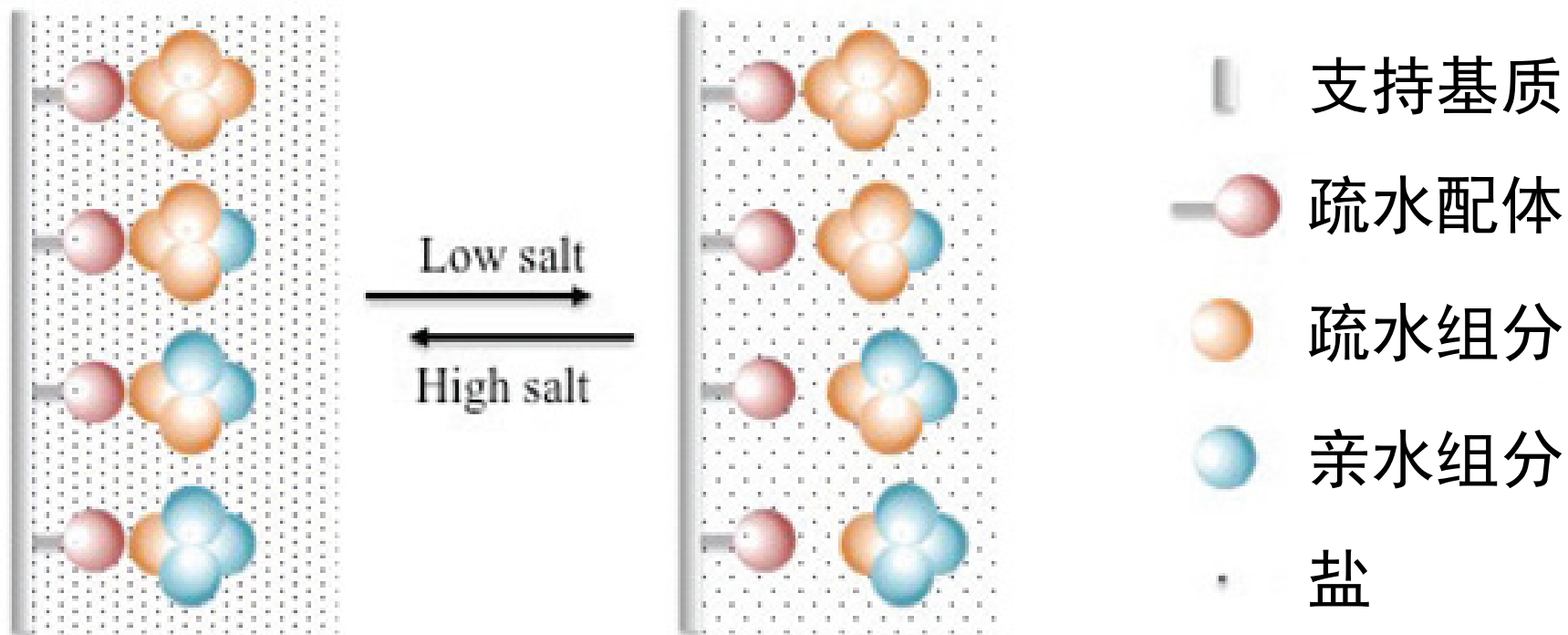
▶ 游离配体

▶ 高盐

▶ 酸

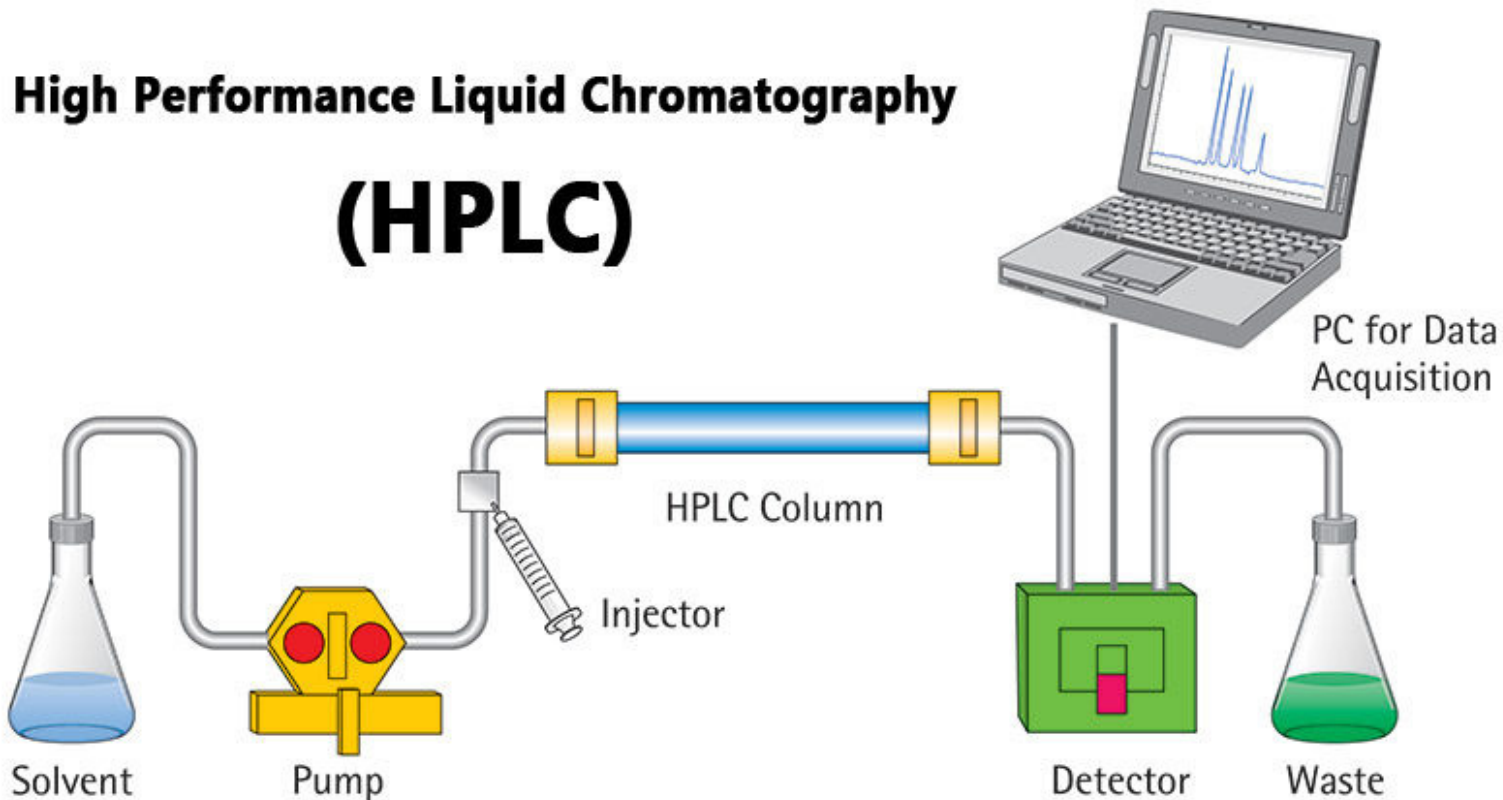
4. 疏水层析 (Hydrophobic interaction chromatography)

- ▶ 疏水基团（如辛烷基、苯环等）作为固定相
- ▶ 高盐促进蛋白暴露疏水面，与基质结合
- ▶ 洗脱方法：低盐、去垢剂、pH值等



5. 高效液相色谱 (high-performance liquid chromatography)

- ▶ 高效液相色谱从原理上与经典的液相色谱没有本质的差别
- ▶ 特点是采用了高压输液泵、高灵敏度检测器和高效微粒固定相
- ▶ 速度快、分辨率高、灵敏度高、柱可反复使用、样品用量少等



分离纯化蛋白通常需要多种手段联合使用

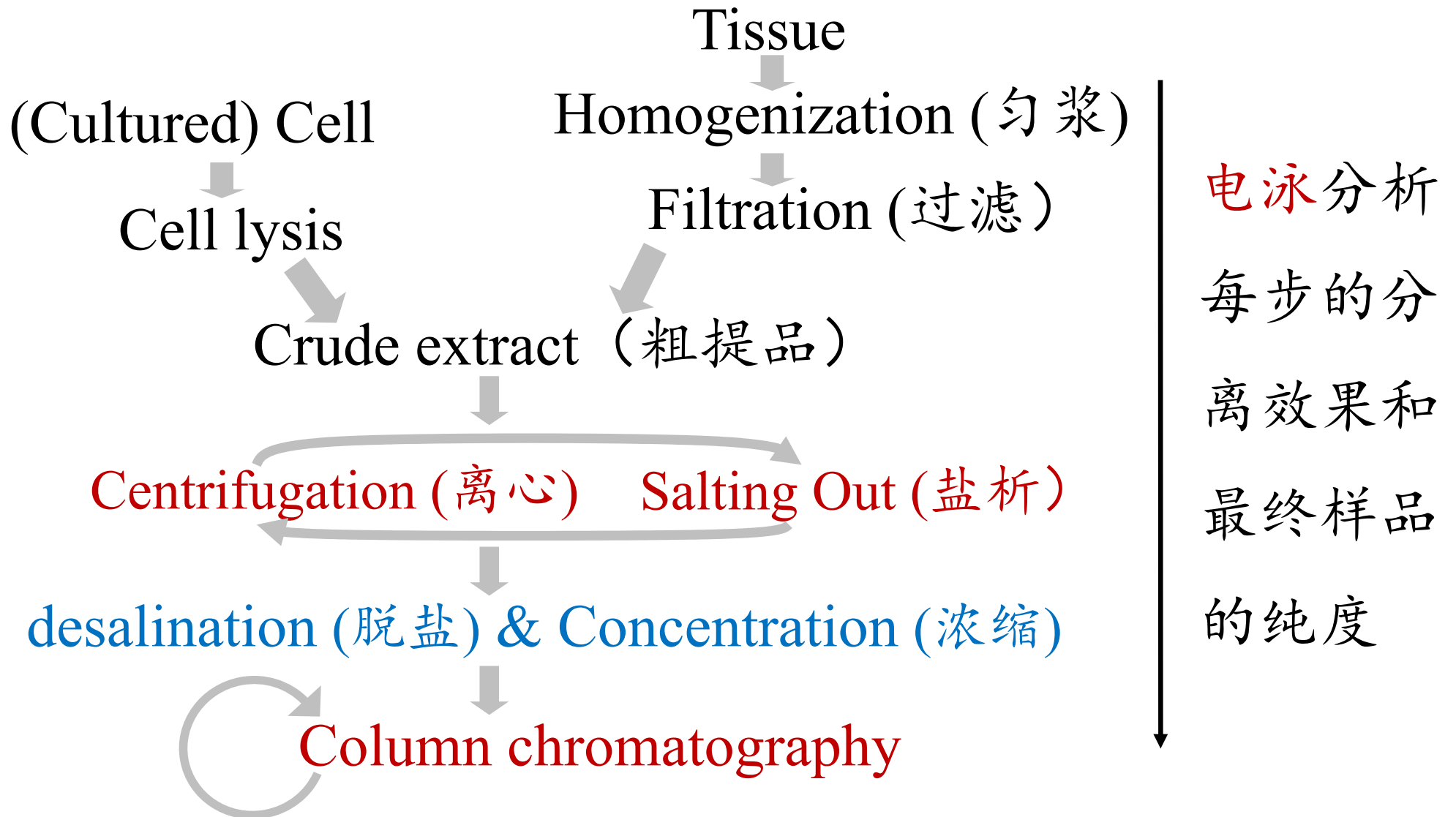
- ▶ 样品总量减少、活力降低
- ▶ 目的蛋白的**比活力**（单位质量的蛋白活性，specific activity）大大提高



TABLE 3-5 A Purification Table for a Hypothetical Enzyme

Procedure or step	Fraction volume (mL)	Total protein (mg)	Activity (units)	Specific activity (units/mg)
1. Crude cellular extract	1,400	10,000	100,000	10
2. Precipitation with ammonium sulfate	280	3,000	96,000	32
3. Ion-exchange chromatography	90	400	80,000	200
4. Size-exclusion chromatography	80	100	60,000	600
5. Affinity chromatography	6	3	45,000	15,000

总结：蛋白质纯化的基本流程



利用蛋白质的物理化学性质对其进行研究

蛋白大小

蛋白电荷

亲疏水性

分子互作

▶ 蛋白质的分离与纯化

▶ 沉淀与离心、透析与超滤、层析、.....

▶ 蛋白质的鉴定

▶ 电泳、免疫印迹、化学测序、质谱分析、.....

▶ 蛋白质序列分析

▶ 序列预测、序列比对、.....

▶

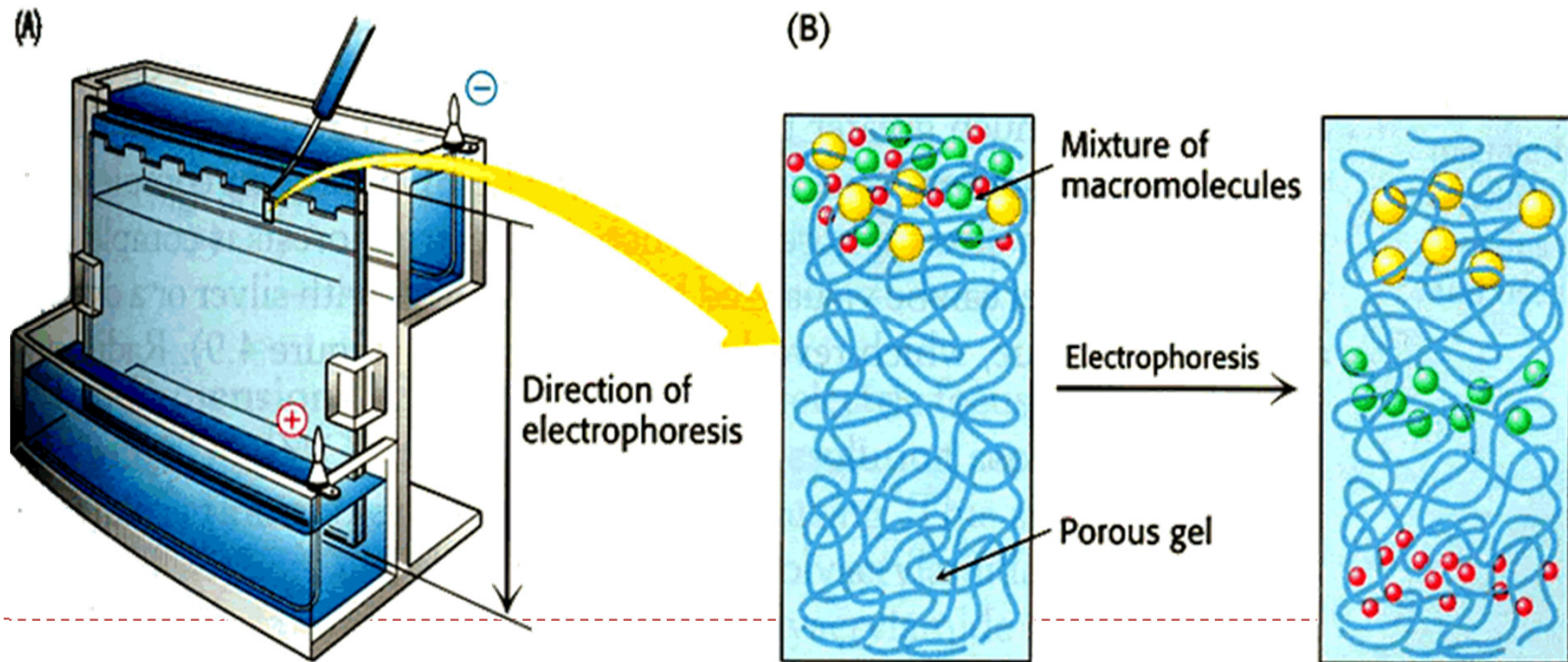
分析假想蛋白X的流程

- ▶ 研究大小和纯度
- ▶ 电泳
- ▶ 研究在那些组织或细胞中表达
- ▶ 免疫印迹
- ▶ 蛋白质的序列
- ▶ 化学测序、质谱分析

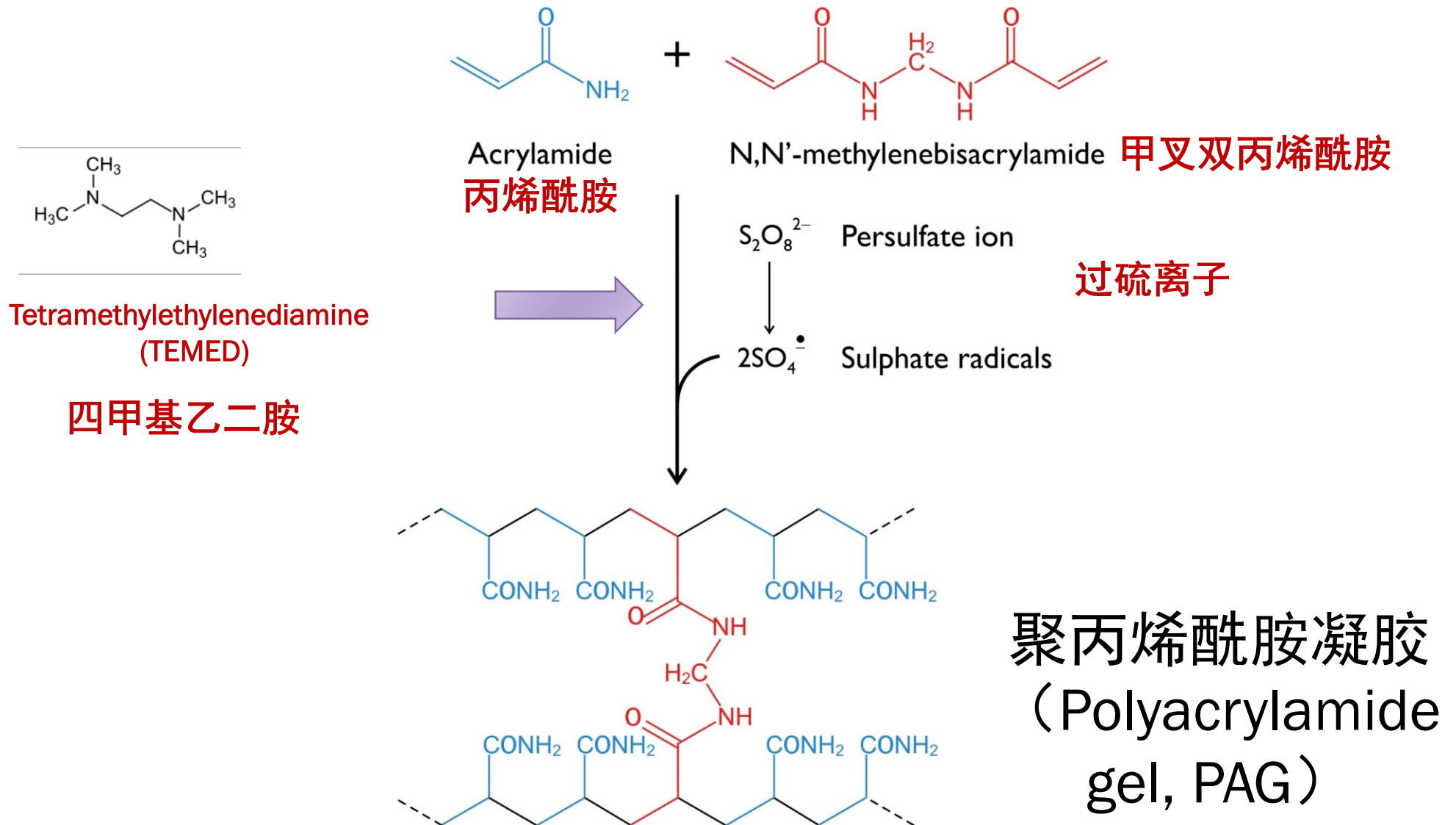


蛋白质鉴定方法1：电泳 (Electrophoresis)

- ▶ 带电颗粒在**电场**作用下，在**多孔电泳介质**中向着与其**电性相反**的电极移动。
- ▶ 泳动快慢与带电**电荷量**成正比，与**自身半径**和**电泳溶液的黏度**成反比。

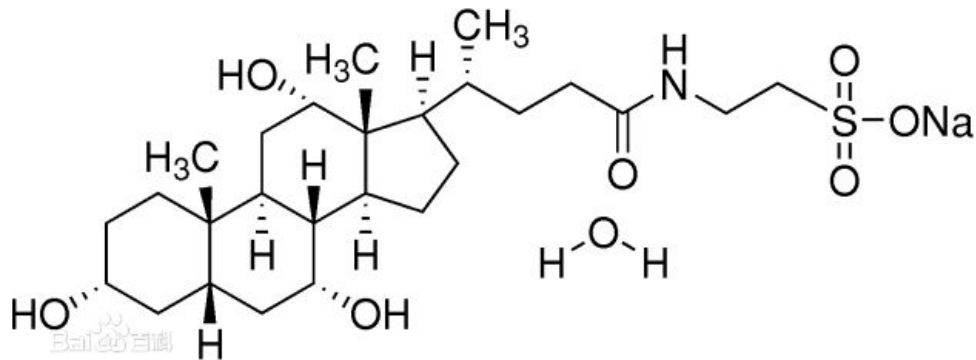


常用聚丙烯酰胺凝胶来分离蛋白质



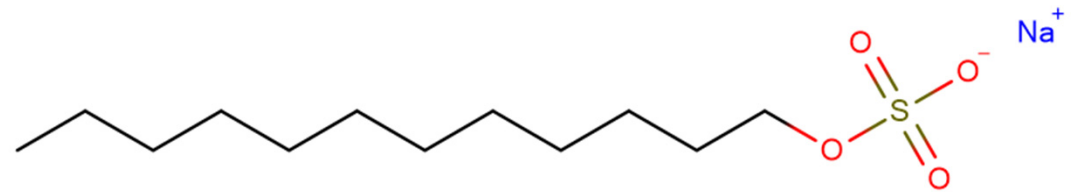
聚丙烯酰胺凝胶电泳 (PAGE)

种类	非变性 (Native)	变性 (Denatured)
典型目的	蛋白的多聚化	蛋白质的常用分离鉴定方式
影响因素	分子量、分子构象、分子电荷等	分子量
去垢剂	脱氧胆酸钠 (弱)	SDS (强)



脱氧胆酸钠

► (Sodium deoxycholate)

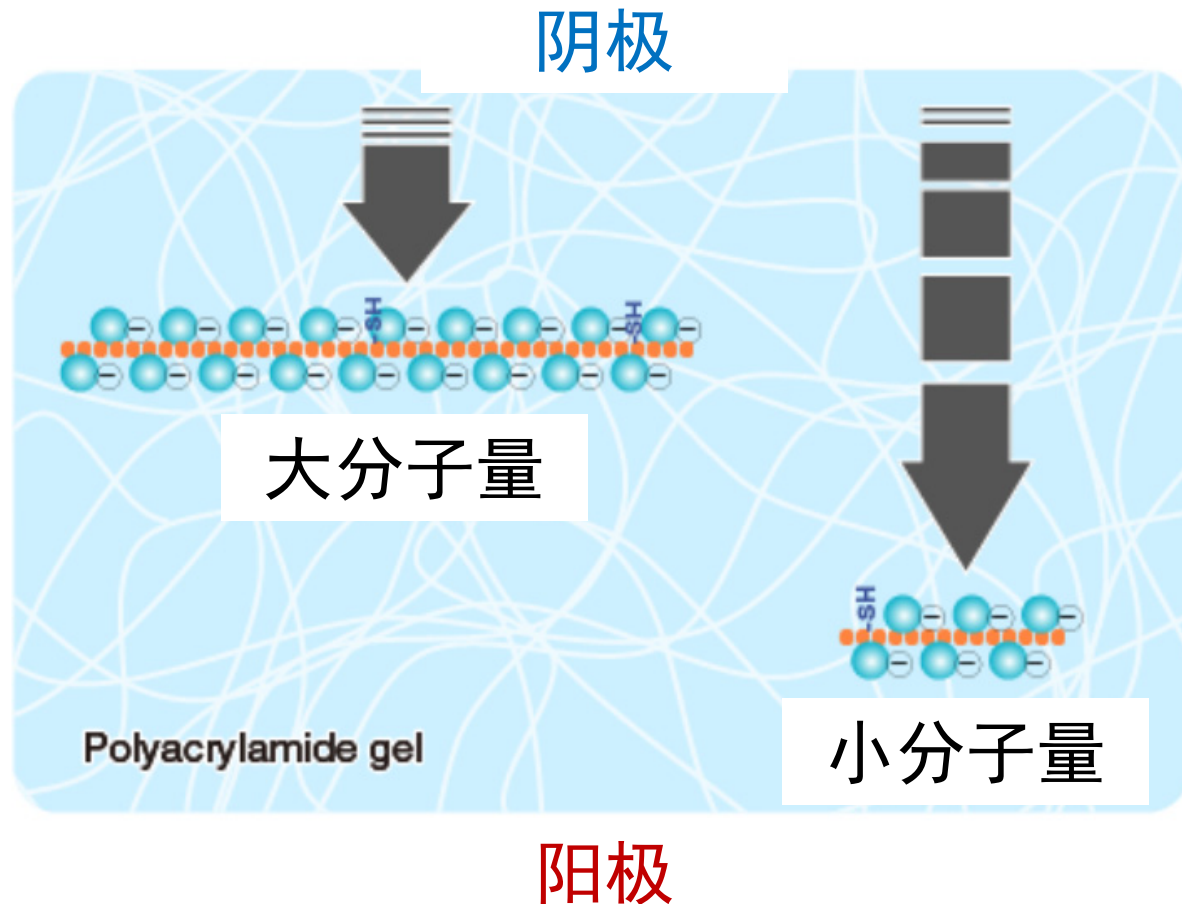


十二烷基硫酸钠

(Sodium dodecyl sulfate)

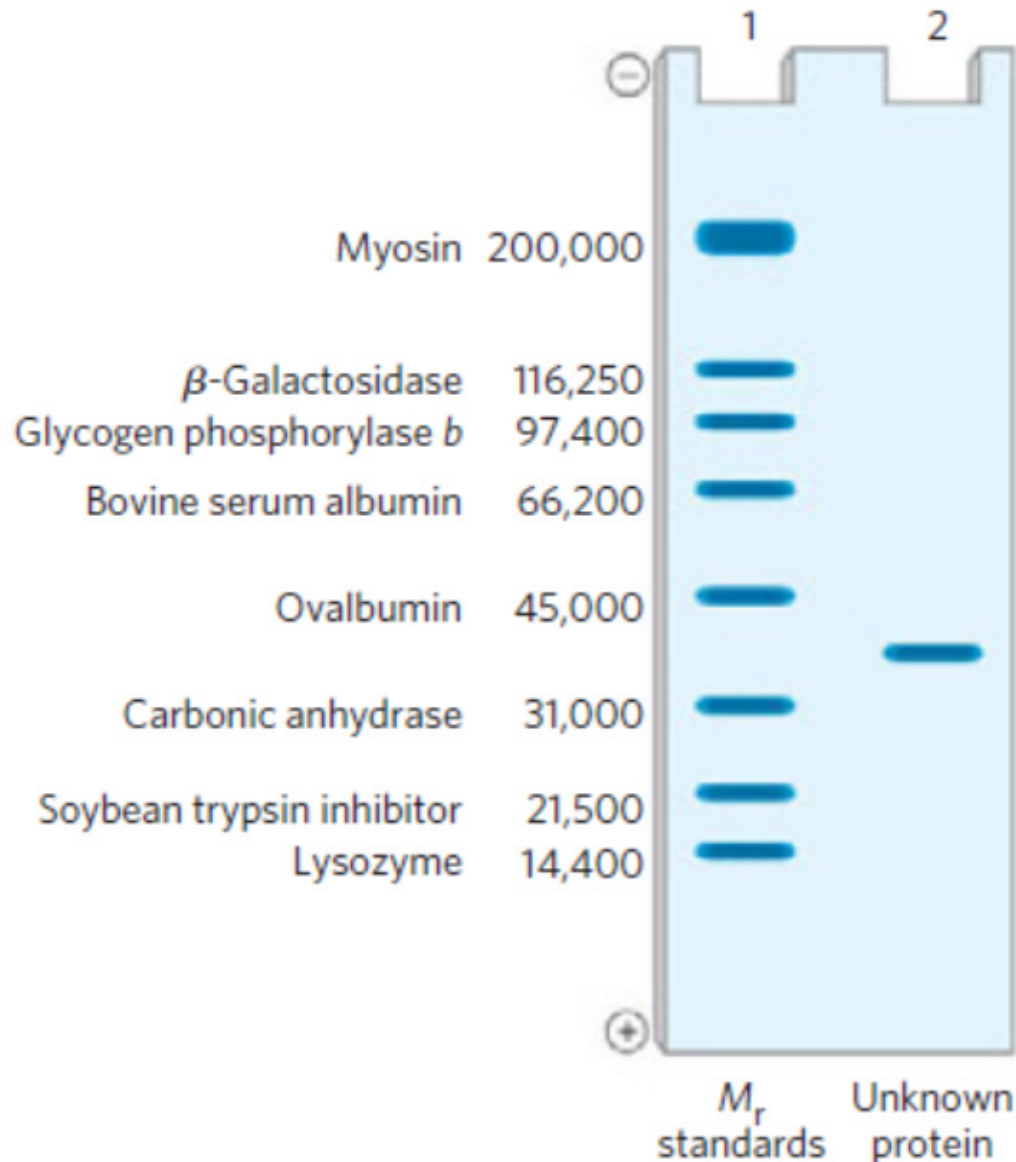
按蛋白的分子量分离：SDS-PAGE

- ▶ 如果在PAGE中加适量**SDS**和**还原剂**，可用于测定蛋白质的分子质量。

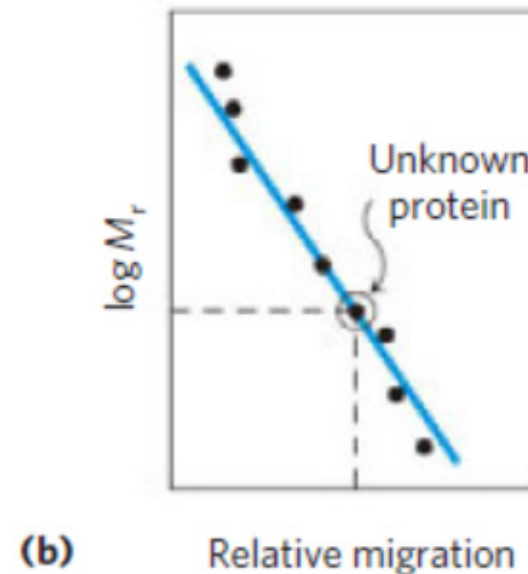


- ▶ 胶束在SDS-PAGE系统中的电泳迁移率不再受蛋白质原有电荷的影响，而主要取决于椭圆棒的长轴长度，即蛋白质或亚基分子量的大小。

SDS-PAGE用来估测蛋白质的分子量



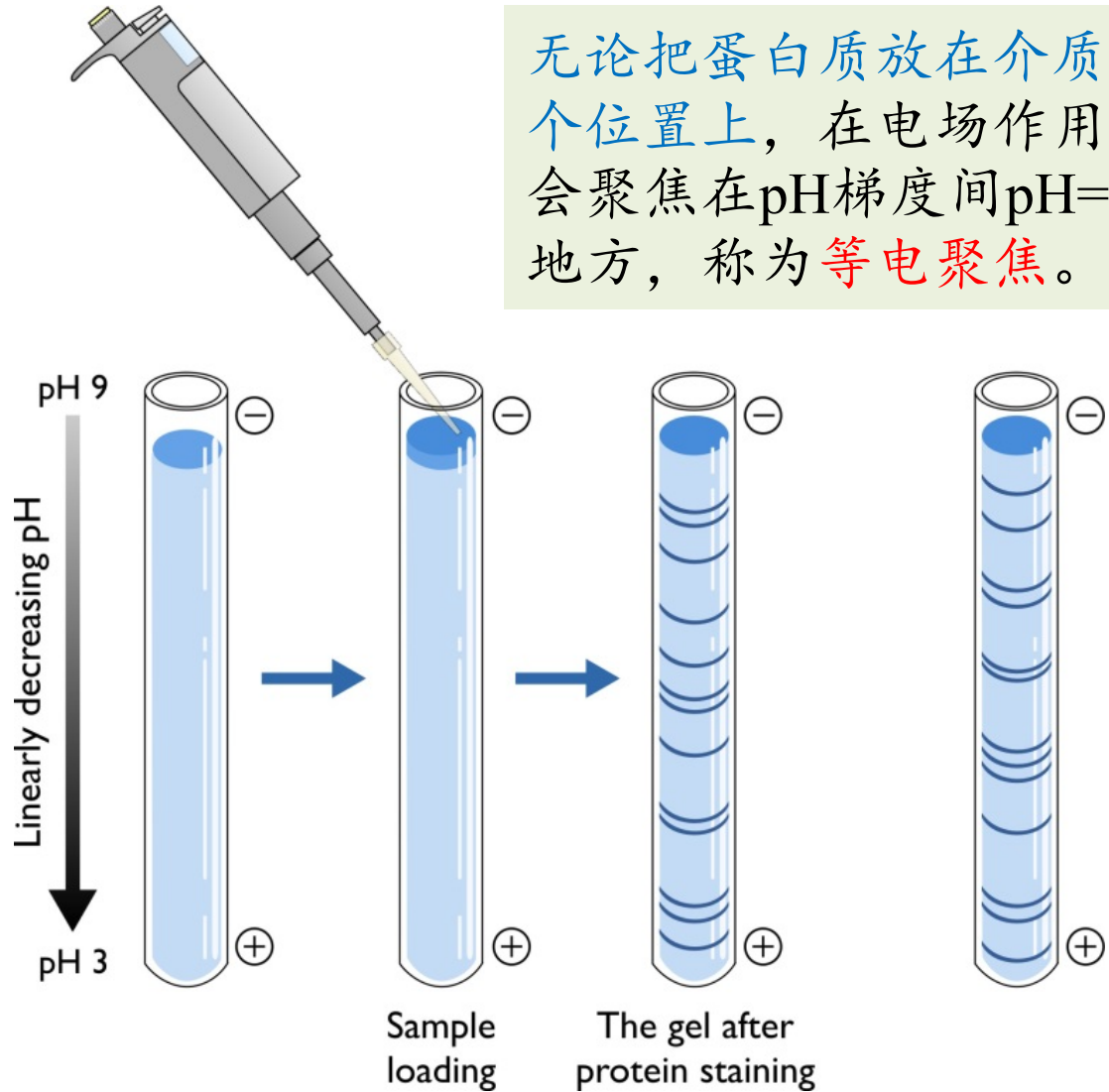
- ▶ 当蛋白质的**分子量**在15 kDa–200 kDa之间时，电泳迁移率与分子量的对数呈线性关系。



按蛋白的pH值分离：等电聚焦 (Isoelectric focusing)

加入两性电解质形成自然pH梯度

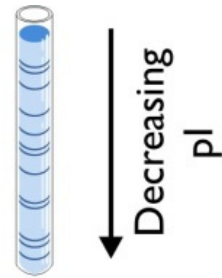
载体两性电解质
(商品名Ampholine)
是一种脂肪族多胺
基多羧基的混合物，
用于形成pH梯度。



无论把蛋白质放在介质的哪个位置上，在电场作用下都会聚焦在pH梯度间 $pH=pI$ 的地方，称为**等电聚焦**。

双向电泳 (two-dimensional electrophoresis)

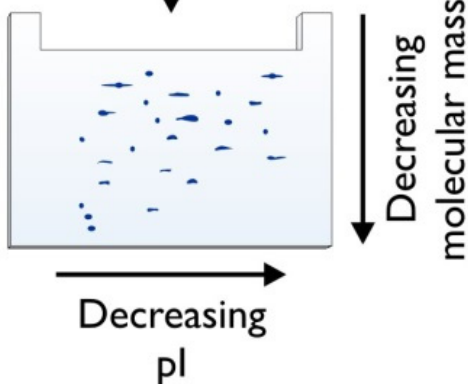
1. Separation of proteins by pI value



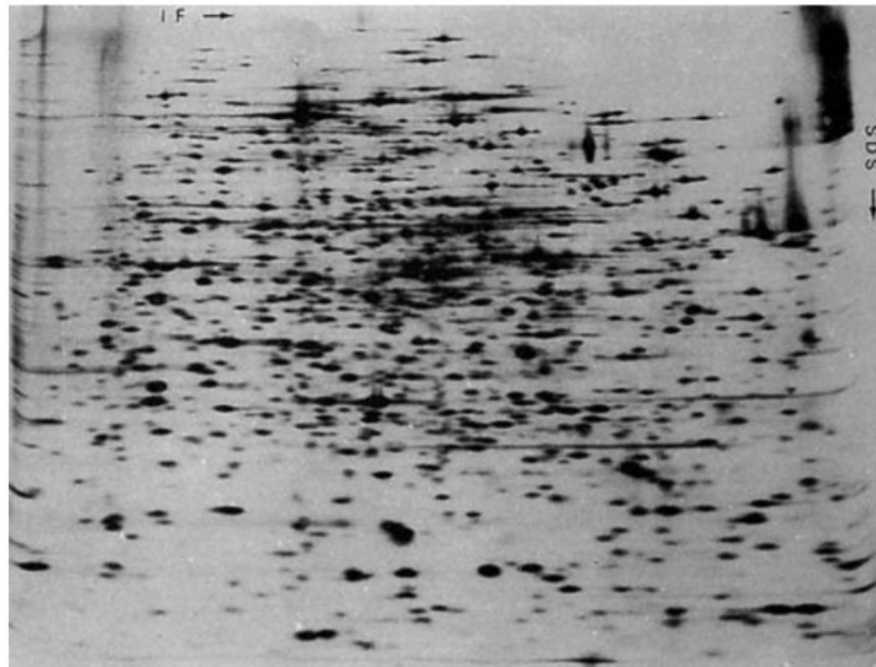
2. Soaking the gel in SDS solution and fitting it on an SDS PA gel



3. Separating the proteins by molecular mass with SDS PAGE

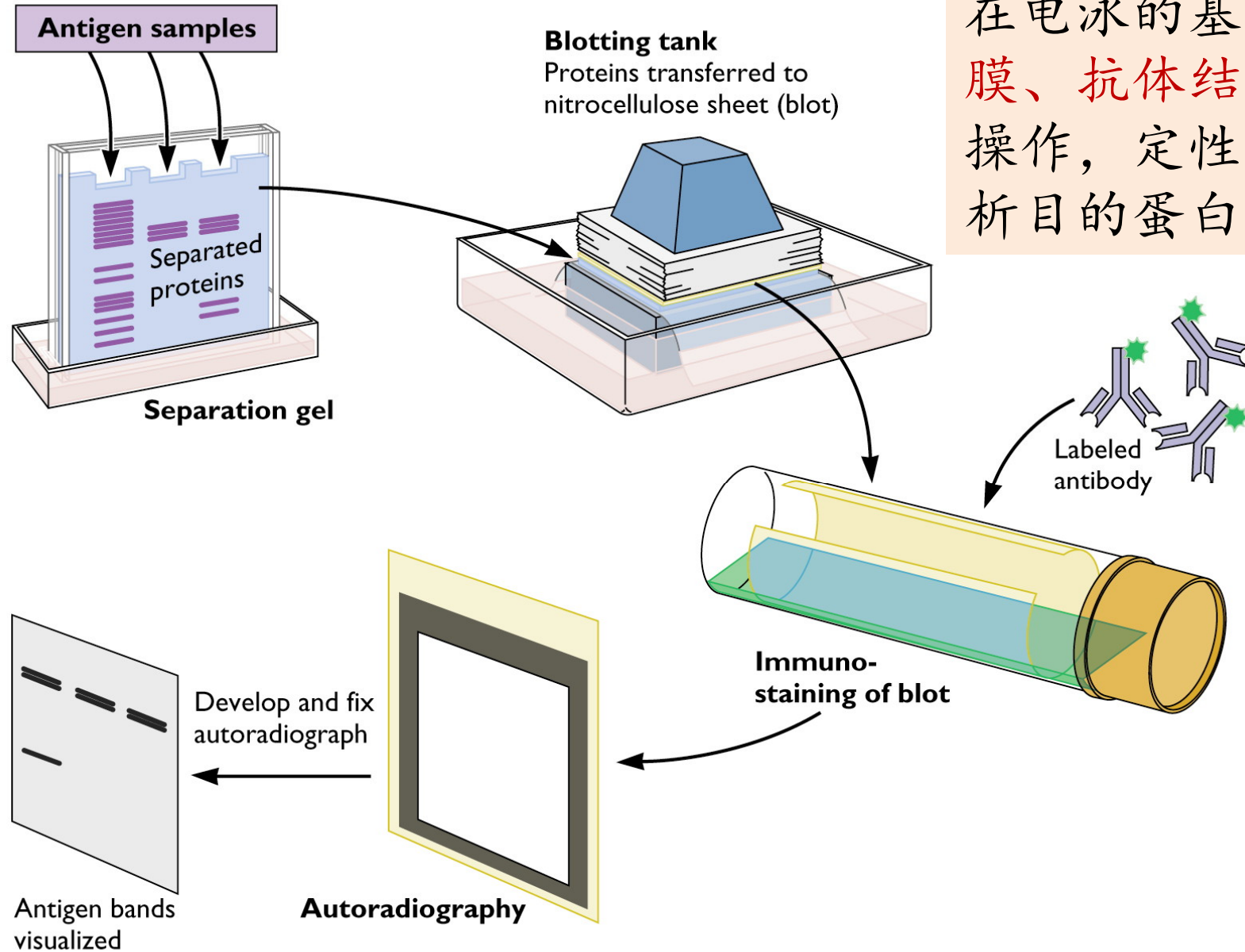


在两个维度更精密的分离各种蛋白质



Approximately 1000 *E. coli* proteins on a single 2D gel

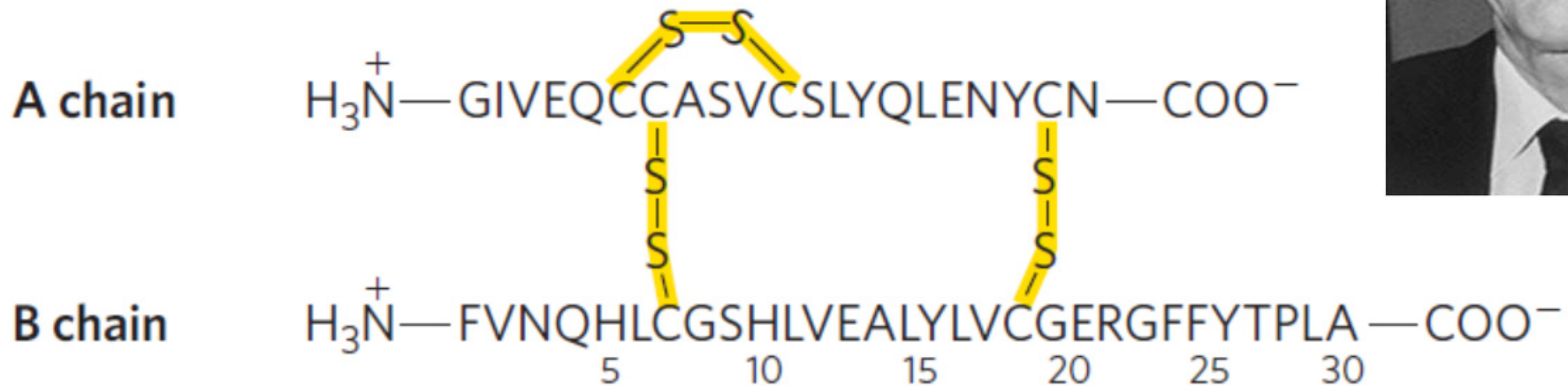
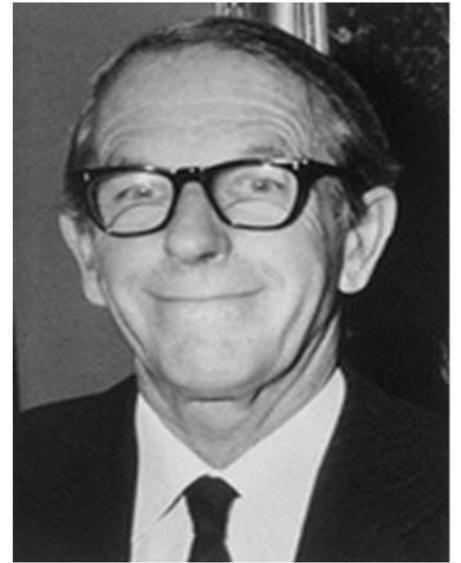
蛋白质鉴定方法2：免疫印迹



在电泳的基础上进行**转膜、抗体结合、显影**等操作，定性或半定量分析目的蛋白。

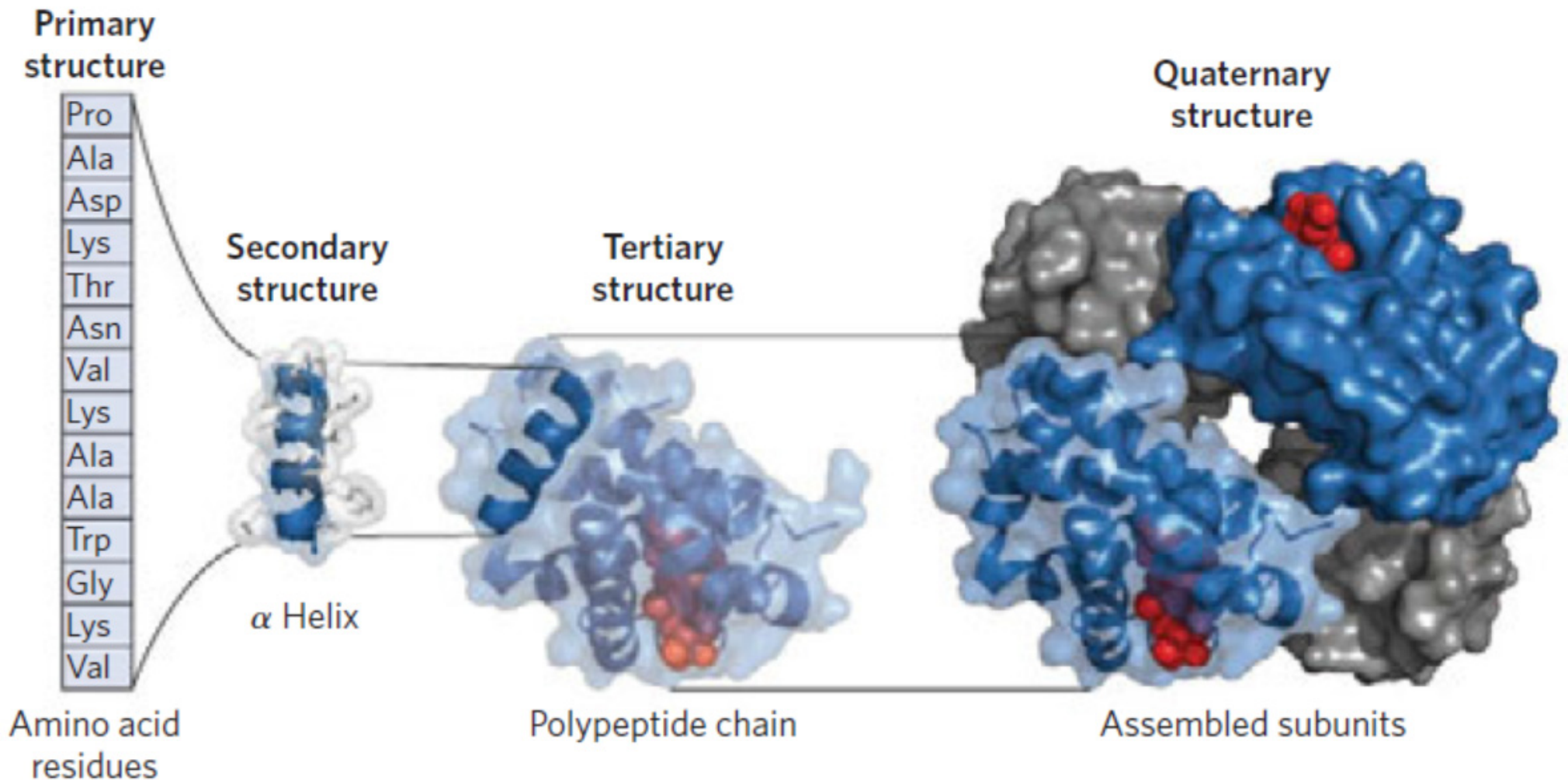
蛋白质鉴定方法3：化学测序

- ▶ F. Sanger经过10多年的努力，使用超过100g牛胰岛素（bovine insulin），于1953年公布测序结果；
- ▶ 因该工作获得1958年的诺贝尔化学奖。



蛋白质的各级结构

- ▶ 不是所有的蛋白都有四级结构



蛋白质测序策略

- ▶ 纯化目标蛋白（纯度大于97%）
- ▶ 拆分肽链（8 M尿素、8 M盐酸胍、高盐、极端pH等）
- ▶ 分析单链氨基酸组成（充分水解后分离分析氨基酸）
- ▶ 打开二硫键（巯基还原后封闭，或氧化成磺酸基）
- ▶ 不同酶将肽链切成小片段
- ▶ 测定每个小片段的序列
- ▶ 序列拼接和定位二硫键



蛋白质水解

▶ 酸水解

- ▶ 6 M盐酸或4 M硫酸煮20小时，可以得到L-氨基酸的混合物
- ▶ 被破坏的氨基酸：**色氨酸**、丝氨酸、苏氨酸、天冬酰胺、谷氨酰胺

▶ 碱水解

- ▶ 5 M烧碱煮20小时，得到**消旋**的氨基酸（racemization）
- ▶ 被破坏的氨基酸：包括精氨酸在内的大多氨基酸；获得稳定的**色氨酸**

▶ 酶解

- ▶ 不消旋，不破坏氨基酸，多种酶混合才能完全水解，序列特异性
 - ▶ 胰蛋白酶（trypsin）、胃蛋白酶（pepsin）等
-



不同酶将肽链切成小片段

*Reagent (biological source)**

Cleavage points[†]

Trypsin **胰蛋白酶**
(bovine pancreas)

Lys, Arg (C) **N/C表示切点**

Submaxillaris protease
(mouse submaxillary gland)

Arg (C)

Chymotrypsin **糜蛋白酶**
(bovine pancreas)

Phe, Trp, Tyr (C)

Staphylococcus aureus V8 protease
(bacterium *S. aureus*)

Asp, Glu (C)

Asp-N-protease
(bacterium *Pseudomonas fragi*)

Asp, Glu (N)

Pepsin **胃蛋白酶**
(porcine stomach)

Phe, Trp, Tyr (N)

Endoproteinase Lys C
(bacterium *Lysobacter
enzymogenes*)

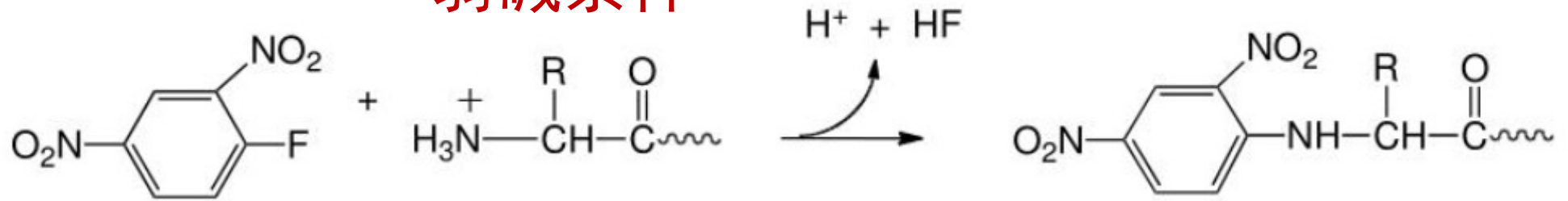
Lys (C)

Cyanogen bromide

Met (C)

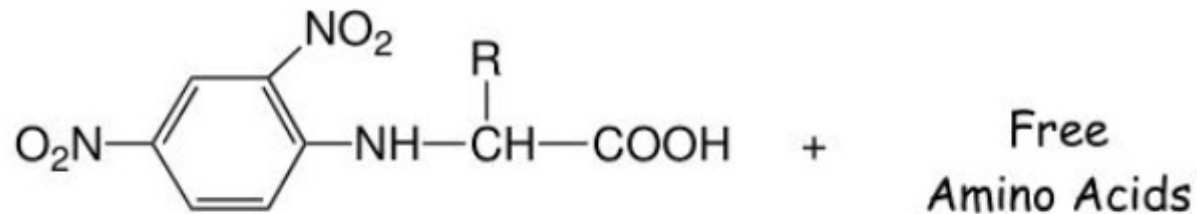
测定每个小片段的N端序列：Sanger法 (1945)

弱碱条件



2,4-二硝基氟苯
(DNFB, Sanger试剂)

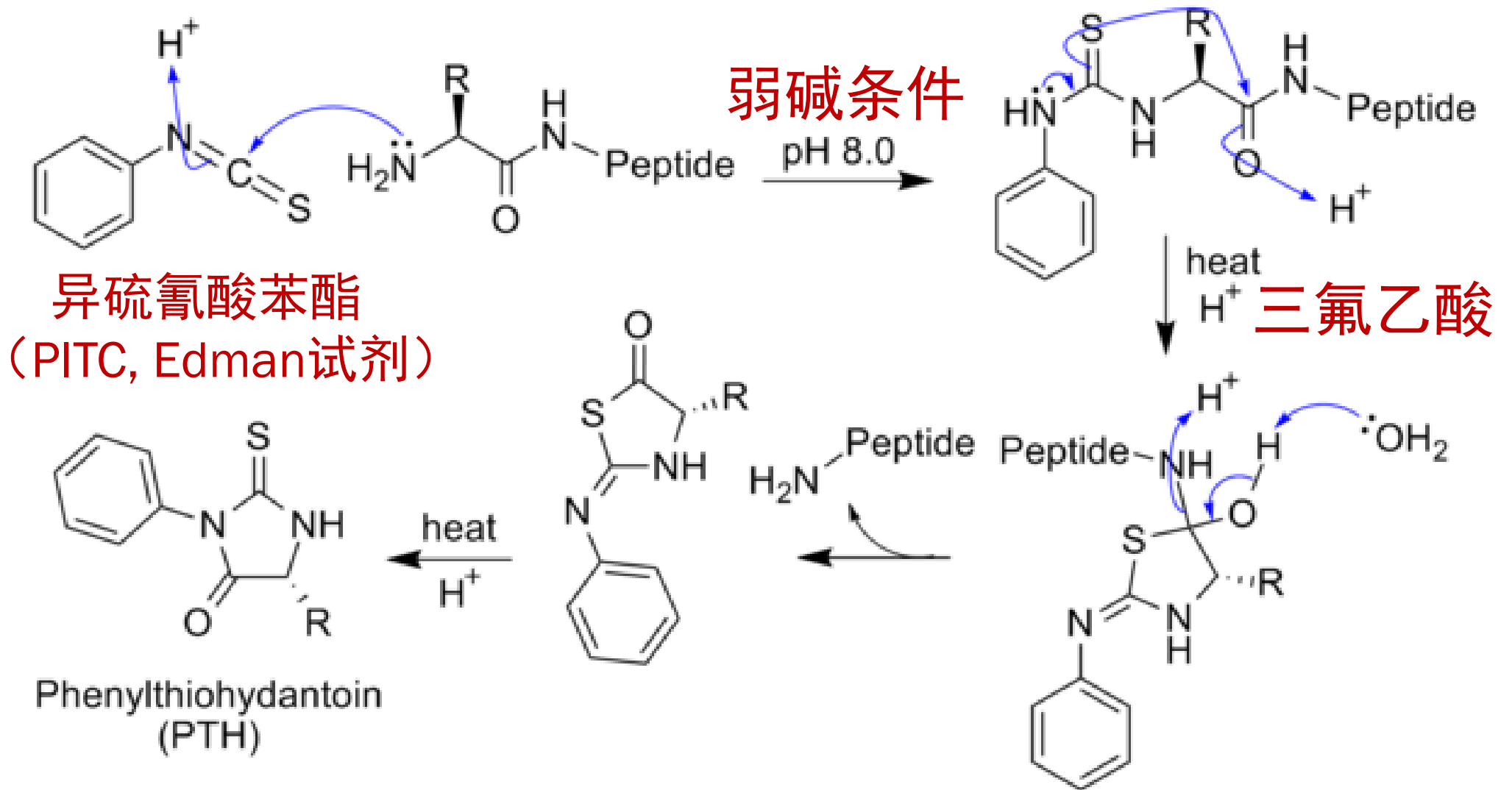
强酸水解



经乙醚抽提后
纸层析鉴定

2,4-二硝基苯氨基酸
(DNP-氨基酸) (黄色)

测定每个小片段的N端序列：Edman法 (1950)



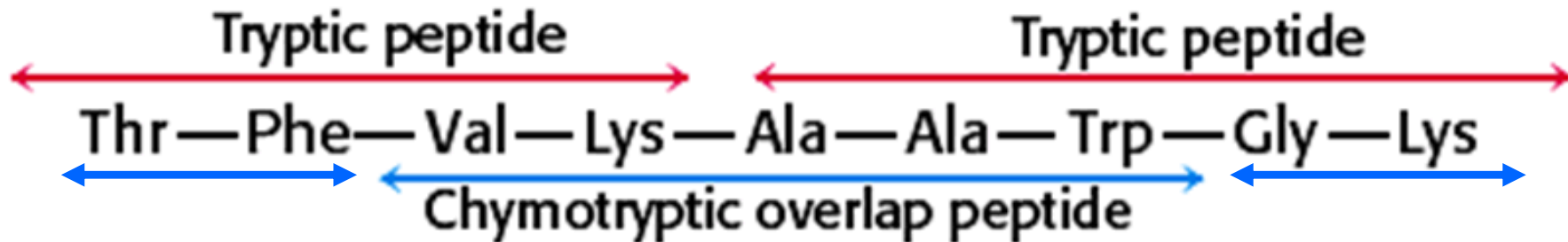
Edman降解法一次可连续准确测出20-30个残基序列

序列拼接

- ▶ 至少有两种切割方式
- ▶ 片段必须重叠
- ▶ 对可能出现二硫键的位置进行酶切电泳等实验验证

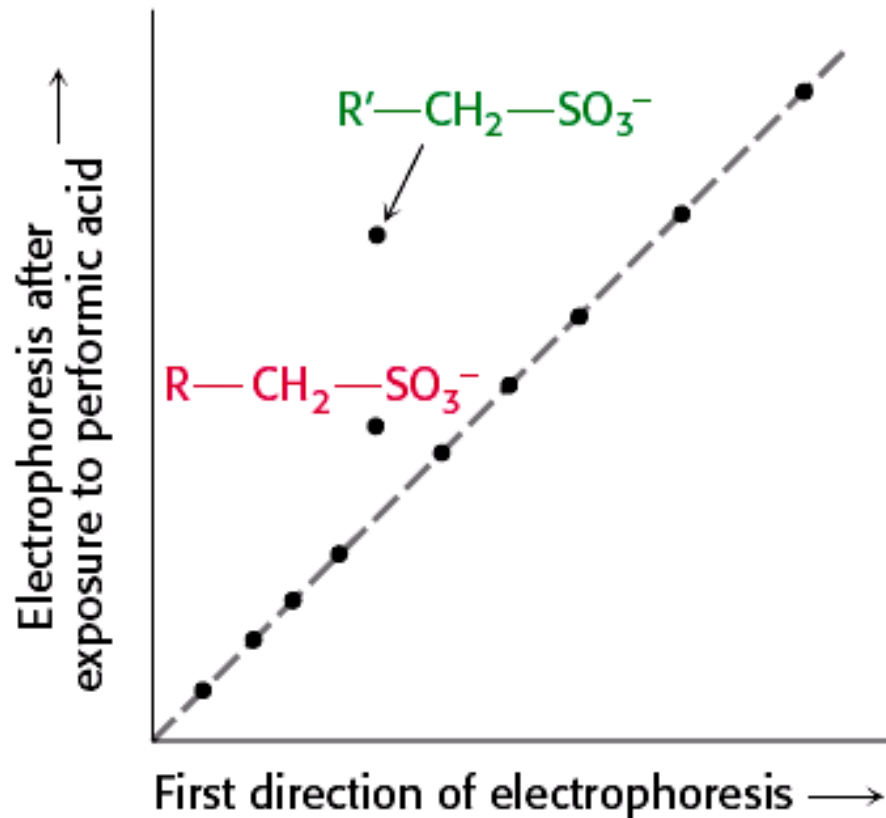
Tryptic peptides
Ala—Ala—Trp—Gly—Lys
Thr—Phe—Val—Lys

Chymotryptic peptide
Val—Lys—Ala—Ala—Trp
Thr—Phe Gly—Lys



定位二硫键

Diagonal electrophoresis
对角线电泳



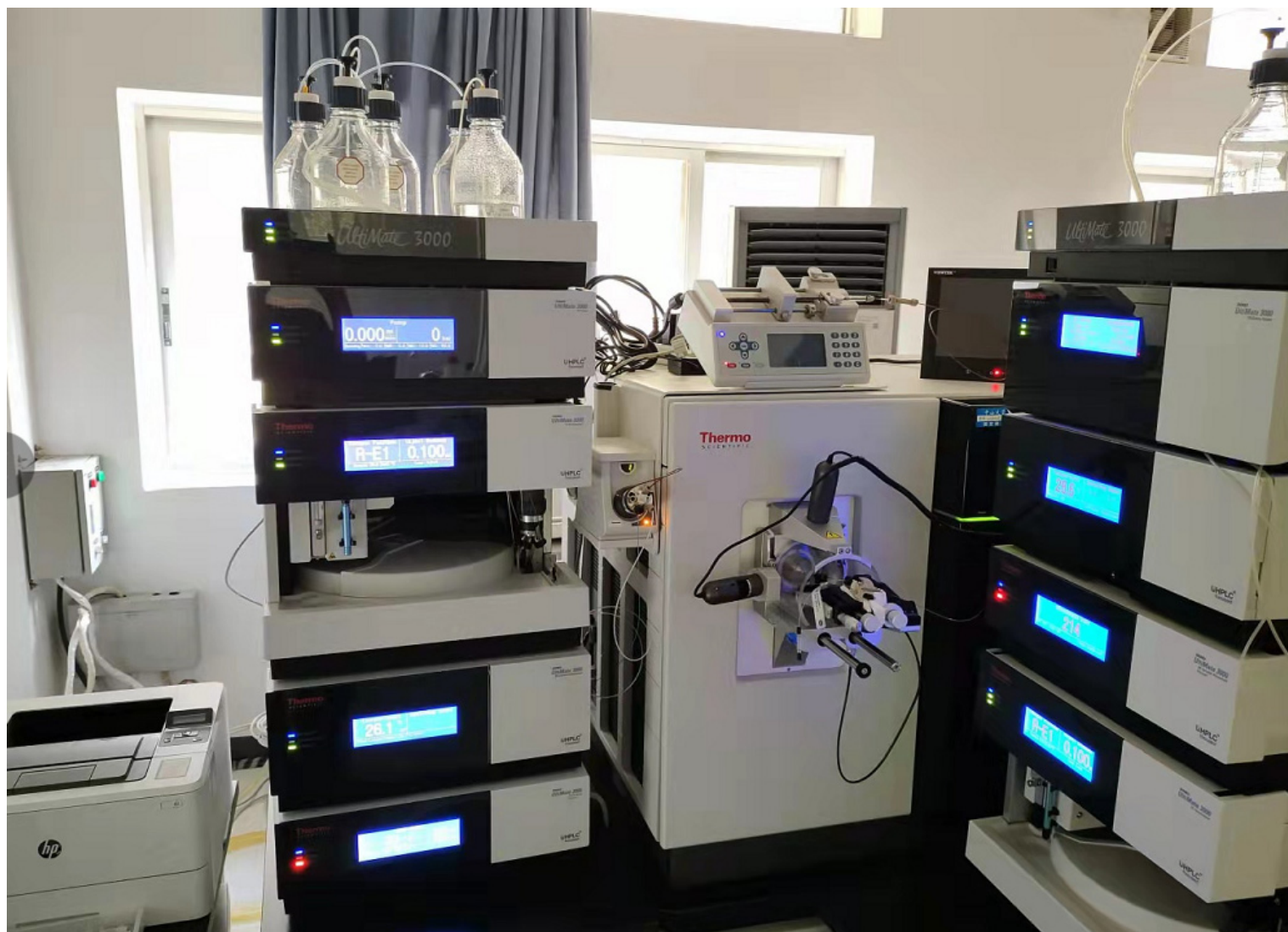
- ▶ 选用胃蛋白酶水解，对角线电泳分离
- ▶ 第一向电泳：大小及电荷的不同
- ▶ 滤纸暴露在过甲酸蒸气中
- ▶ 第二向电泳：含磺酸基的成对肽段比原来含二硫键的肽小而负电荷增加
- ▶ 肽斑用茚三酮显色
- ▶ 未用茚三酮显色的肽段进行氨基酸分析，与多肽链序列比较

蛋白质鉴定方法4：质谱分析 (Mass spectrometry)

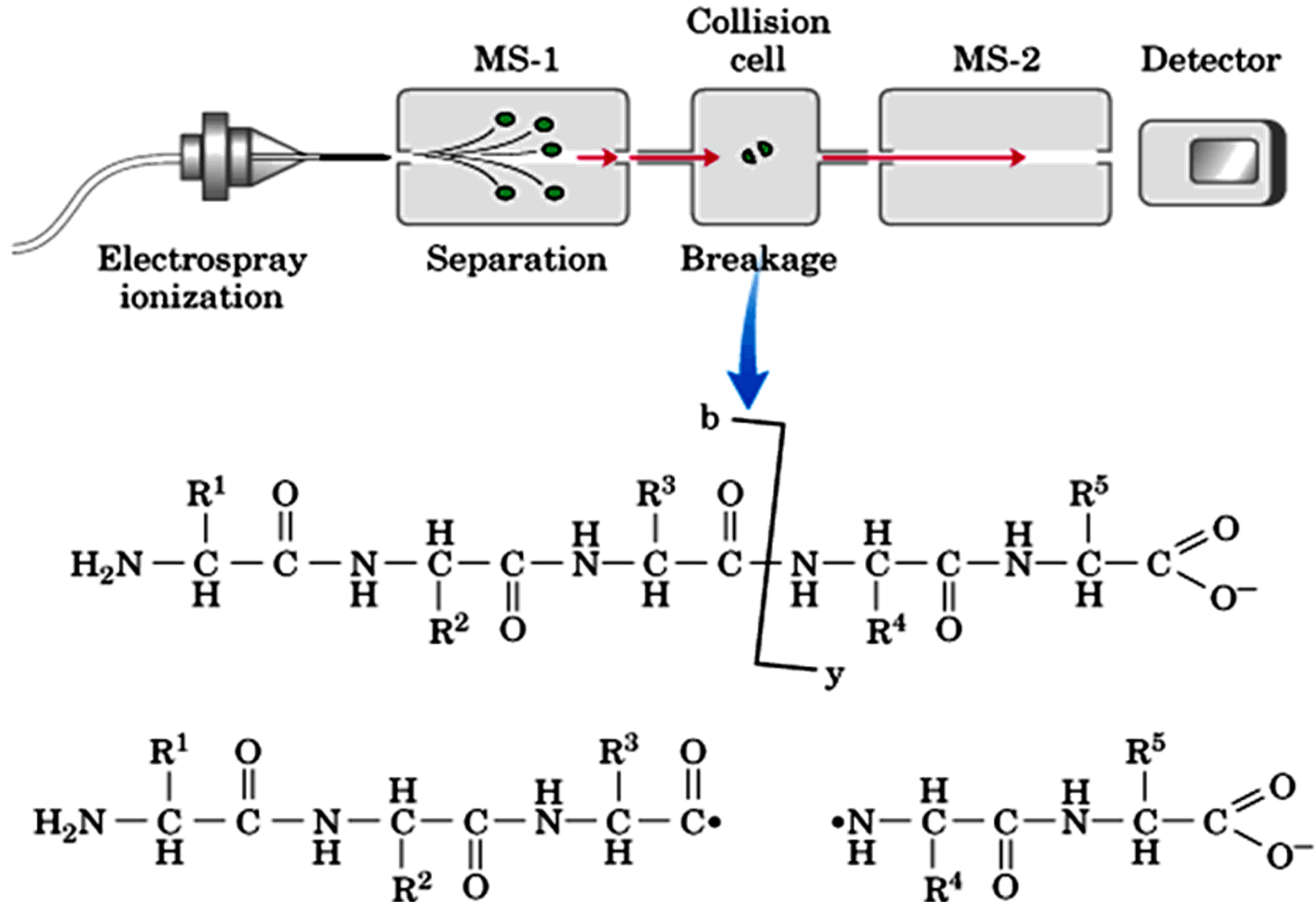
- ▶ 带电粒子在电场/磁场中的运动行为取决于其**荷质比** (mass-to-charge ratio)，荷质比不同的带电粒子其运动行为不同，利用这种差异可以对分子进行测量分析。
- ▶ 质谱是目前**测定分子量**最准确的方法，并且可用于**测定分子结构、蛋白质核酸的测序**等。



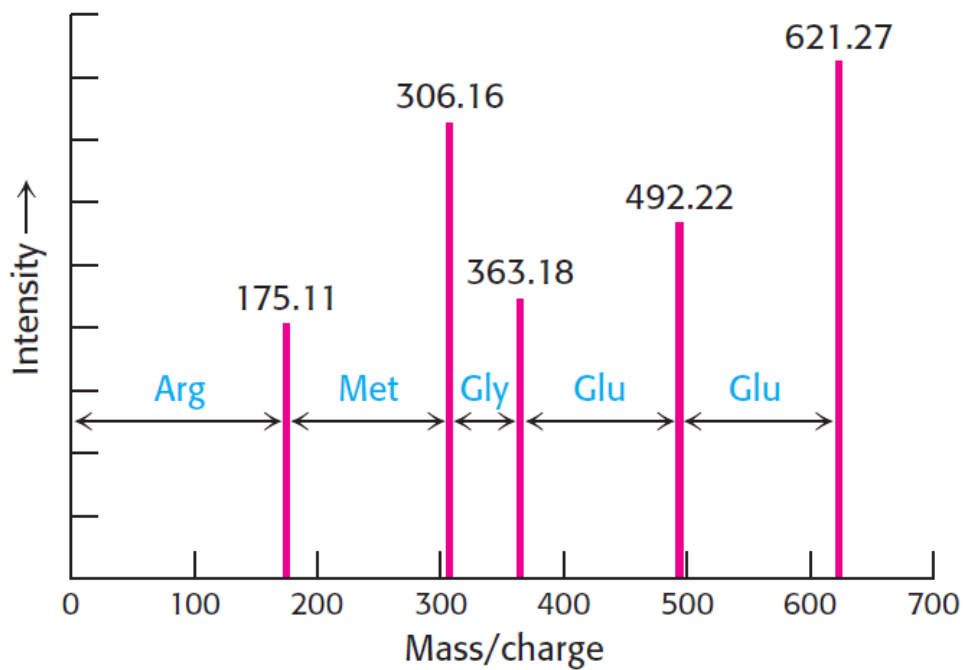
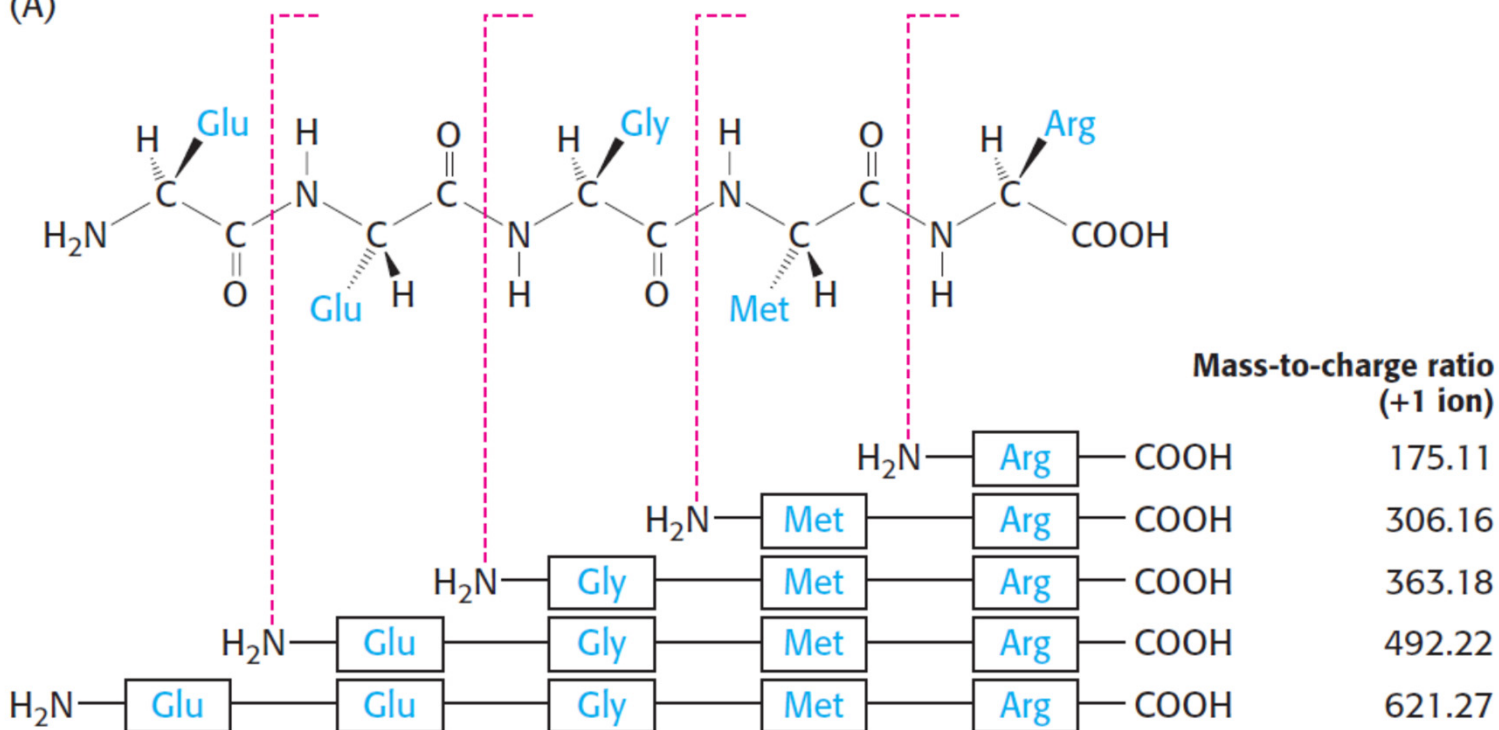
Orbitrap Fusion Lumos 高分辨质谱仪



串联质谱分析可以给出肽段的序列信息



(A)



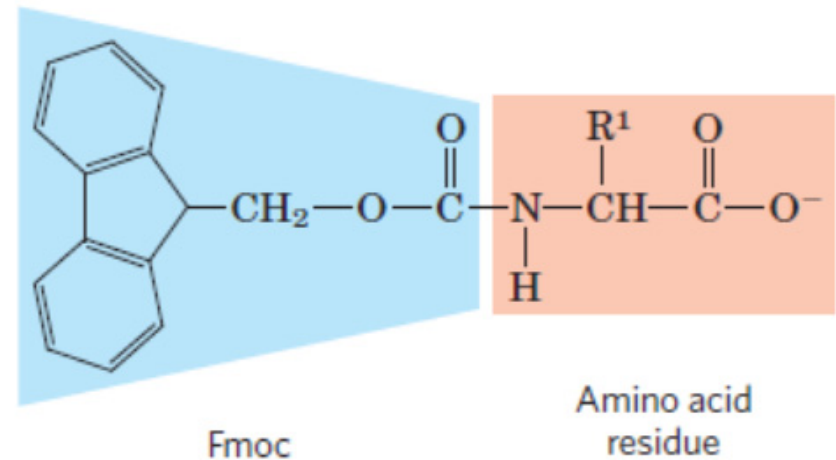
通过碎片的分子量计算，拼接出片段的序列

多肽的固相合成

- ▶ Bruce Merrifield于1963年创立，1984年获诺贝尔化学奖
- ▶ 将肽链的C端氨基酸的羧基共价固定在不溶性高分子树脂上
- ▶ 重复下列步骤，从C端向N端合成多肽

- ▶ 去保护
- ▶ 活化
- ▶ 耦联
- ▶ 洗脱和过滤

N端保护基团



- ▶ 去除保护基团，从树脂上裂解释放多肽
- ▶ 特点：边合成边纯化



不完全反应的积累会导致产物纯度大大下降

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

利用蛋白质的物理化学性质对其进行研究

蛋白大小

蛋白电荷

亲疏水性

分子互作

▶ 蛋白质的分离与纯化

▶ 沉淀与离心、透析与超滤、层析、.....

▶ 蛋白质的鉴定

▶ 电泳、免疫印迹、化学测序、质谱分析、.....

▶ 蛋白质序列分析

▶ 序列预测、序列比对、.....



蛋白质序列分析方法1：序列预测

- ▶ 利用中心法则，从核酸序列推测出氨基酸序列
- ▶ 寻找开放阅读框 ORFfinder
 - ▶ <https://www.ncbi.nlm.nih.gov/orffinder/>
- ▶ mRNA翻译成蛋白质 Translate Tool
 - ▶ <https://web.expasy.org/translate/>

ORF16 (813 aa)

Display ORF as...

Mark

```
>lcl|ORF16
MKAKVSRGGGFRGALNYVFDVGKEATHTKNAERVGGNMAGNDPRELSREF
SAVRQLRPDIDGKPVWHCSLSLPPGERLSAEKWEAVAADFMQRMGFDQTNT
PWVAVRHQDTRDKDHIHIVASRVGLDGKVVWLGQWEARRAIEATQELEHTHG
LTLTPGLGDARAERRKLTDKKEINMAVRTGDEPPRQRLQRLLEAVKDKPT
ALELAERLQAAGVGVANLASTGRMNGFSFEVAGVVPFKGSDLGKGYTWAG
LQKAGVTYDEARDRAGLERFRPTVADRGERQDVAAVREPARGLEAPTGR
SLDRDGDALGTAGPTPAGRDAGSGSLRQGDGHSAQDAGRADAADERERGA
GLRAEGREAGRDHLRPVAQPVRAENEPQQHGADRAAGGDLAQAGERTAG
HDESRRPTDRGSEERDAPAPLAAGAGADSGRGRGRDAGSDWASRFKQASAA
KRRAADGRLGQRDLEQGHAGARVAETDRQSARELDPTAYLEASGYTVKR
EGRHLSVRAGGDEAYRVTRQQDGRWLWCDRYGNDGGDNIDLVREIEPGTG
YAEAVYRLSGAPTVRQQPRPSEPKRQPPQLPAQGLAAREHGRDYLGKRG I
SQDTIEHAEKAGMVRVYADGGVLFVGYDRAGTAQNATRRAIAPADPVQKRD
LRGSDKSYPPILPGDPAKVWIVEGGPDALALHDIKRSGQQPPPTVIVSGG
ANVRSFLERADVQAILKRAERTVAGENEKNPEAQAKADAGHQKQQRVA
KITGREVRQWTPKPEHGKDLADMNARQVAEIERKRQAEIEAERARNRELS
RKSRRYDGPSTFGR
```

Mark subset...

Marked: 0

Download marked set as

Protein FASTA ▾

Label	Strand	Frame	Start	Stop	Length (nt aa)
ORF16	+	3	3153	5594	2442 813
ORF7	+	1	9907	10908	1002 333
ORF17	+	3	5841	6767	927 308
ORF5	+	1	7273	7995	723 240
ORF10	+	2	2456	3169	714 237
ORF3	+	1	3925	4608	684 227
ORF13	+	2	8078	8737	660 219
ORF6	+	1	8995	9624	630 209
ORF2	+	1	766	1392	627 208

蛋白质序列分析方法2：序列比对

▶ 序列相似性 (Similarity)

- ▶ 多个蛋白质的氨基酸残基序列的相似程度，以%表示。

▶ 序列同源性 (Homology)

- ▶ 多个蛋白质 (的基因) 具有共同的起源，称为同源蛋白 (Homologous proteins)，它们往往具有相似或相同的结构和功能，以是或否来表示。



同源蛋白中的氨基酸残基可分为三类

- ▶ 不变氨基酸残基 (Invariant residues) (identities)
 - ▶ 在序列的某些位置上种类不变的氨基酸残基，用一致 (identities) 来表示
- ▶ 可变氨基酸残基 (Variable residues)
 - ▶ 在序列的某些位置上种类可变的氨基酸残基
- ▶ 保守氨基酸残基 (Conservative residues)
 - ▶ 在序列的某些位置上，可替换的相似氨基酸残基，比如R变为K，用Positives来表示

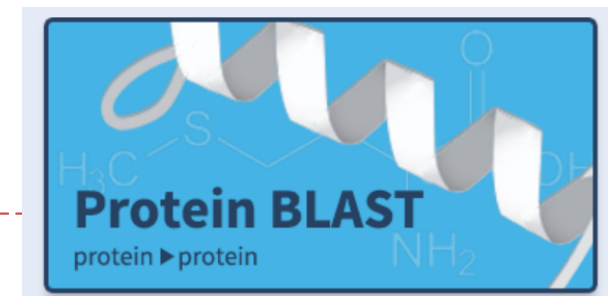


利用生物信息学工具对比蛋白质序列

▶ 序列比对工具 Basic Local Alignment Search Tool

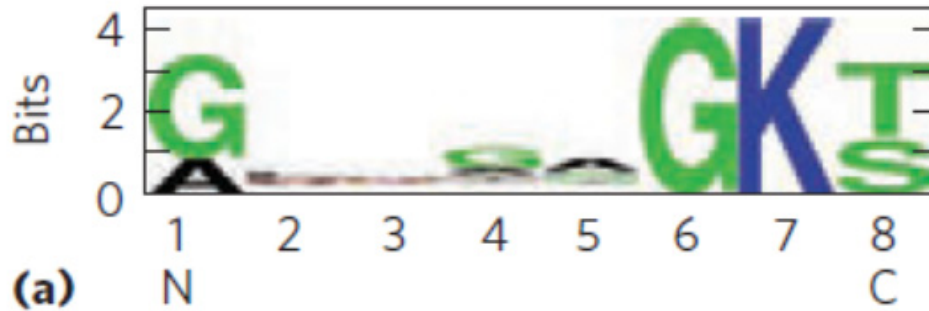
▶ <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

NW Score	Identities	Positives	Gaps
213	79/330(24%)	139/330(42%)	14/330(4%)
Query 1	MKDNTVPLKLIALLANGEFHSGEQLGESLGMSRAAINKHQTLRDWGVDVFTVPGKGYSL		60
Sbjct 1	M + + + + SG+ + ESL +SR A+ K I L+ G + +V KG+ L		60
MSKYSQDVLQLLYKNKPNYISGQSIAESLNISRTAVKKVIDQLKLEGCKIDSVNHKGHLL			60
Query 61	PEPIQLLNAEQILGQLDGGSVT-----VLPVIDSTNQYLLDRIGELKSGDACVAEYQQAG		115
Sbjct 61	+ + + + I+ Q S V IDST + +S +++ Q G		119
QQ-LPDIWYQGIIDQYTKSSALFDFSEVYDSIDSTQLAAKSLVGNQSSFFILSDEQTKG			119
Query 116	RGRRGRKWFSPFGANLYLSMFWRLEQGPAAAIGLSLVIGIVMAEVLRLKLGADKVRVKWPN		175
Sbjct 120	RGR R W S G L++S+ R + +L I + + + ++ D+V+VKWPN		179
RGRENHRWSSSKGQGLWMSVVLRPNVAFSMISKFNLFIALGIRDAlQHFSQDEVKVKWPN			179

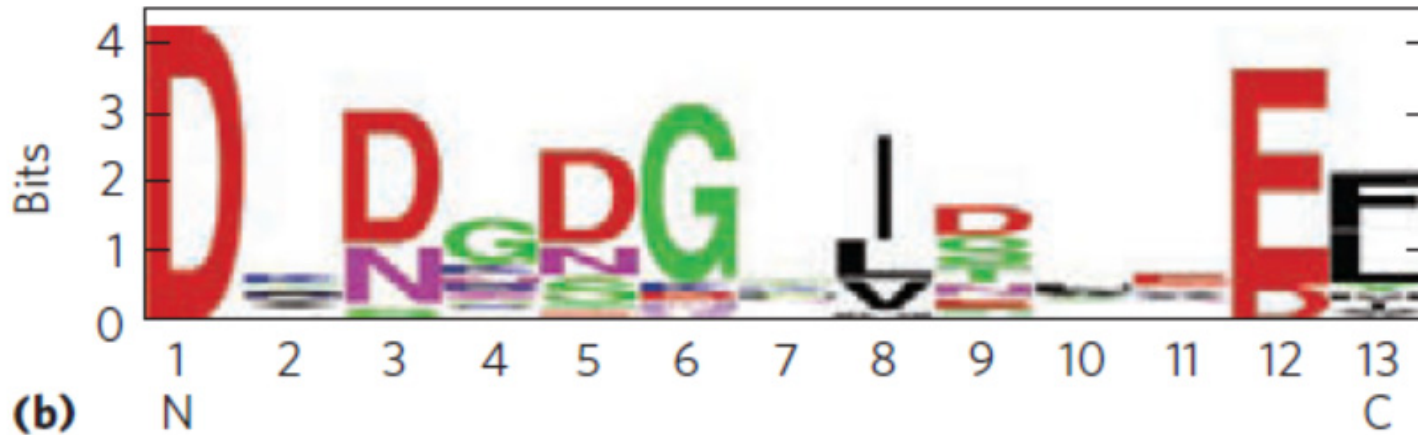


作图描述序列的保守性

[AG]-x(4)-G-K-[ST].

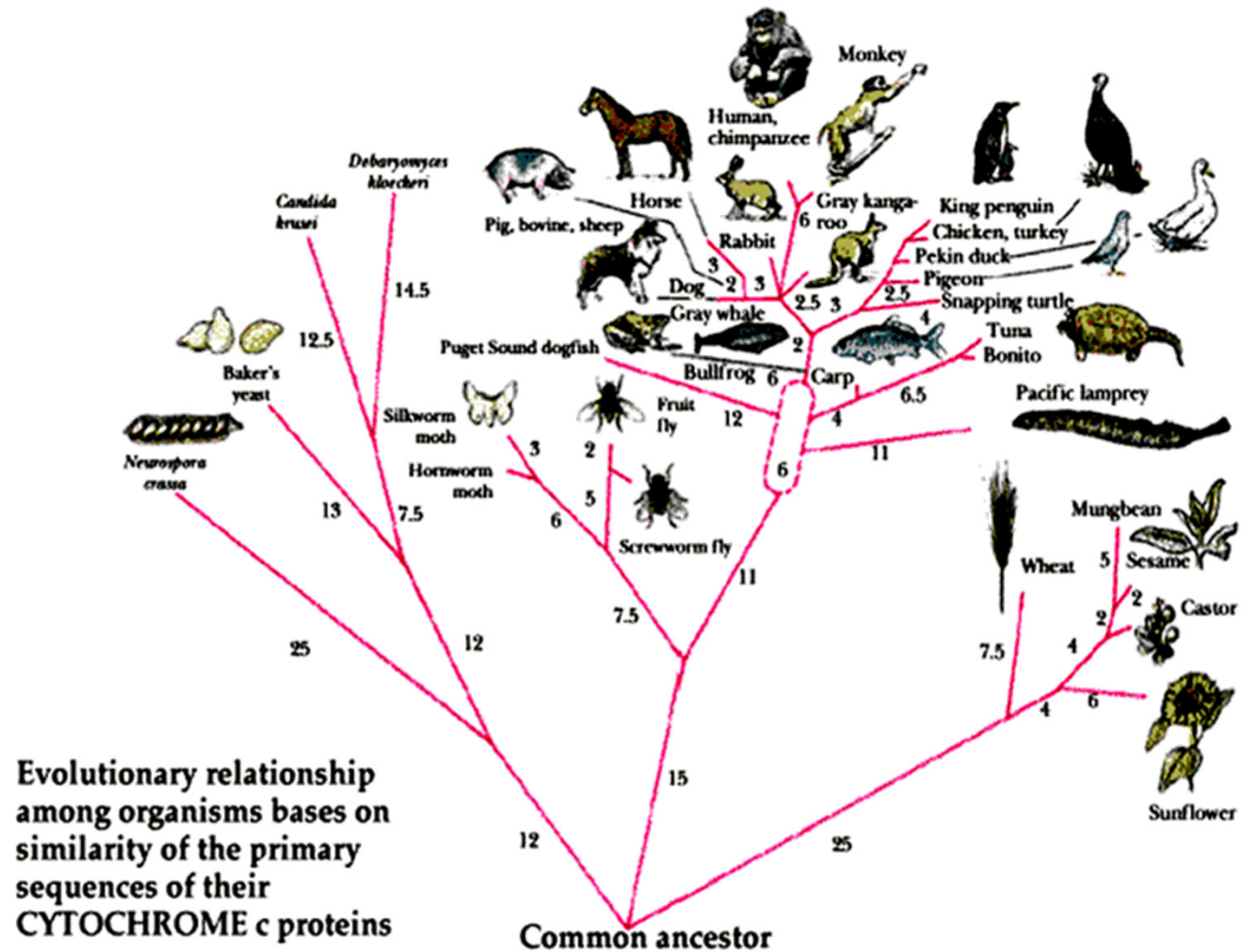


D-{W}-[DNS]-{ILVFYW}-[DENSTG]-[DNQGHRK]-{GP}-
[LIVMC]-[DENQSTAGC]-x(2)-[DE]-[LIVMFYW].



蛋白质序列保守性与物种进化的关系

- 来自任两个物种的同源蛋白质，其序列间的氨基酸差异数目与这些物种间的系统发生差异是成比例的，也即在进化位置上相差愈远，其氨基酸序列之间的差别愈大。



学习目标

- ▶ 氨基酸的通式与立体化学
- ▶ 氨基酸的中英文全称、英文三字母和单字母缩写
- ▶ 氨基酸的等电点、简单计算及应用
- ▶ 蛋白质的分离纯化鉴定的思路、常见方法及基本原理
 - ▶ 分离纯化
 - ▶ 鉴定
 - ▶ 序列分析



第2章 作业

- ▶ 1. 名词解释（英文作答）（手写拍照上传） Isoelectric point; Peptide bond; Salting out; Electrophoresis; Homologous proteins; Conjugated protein。
 - ▶ 2. 书写参与蛋白质组成的20种氨基酸的结构式、简写和缩写（手写拍照上传）；英文教材课后习题：P111第7题和P112第15题（中文或英文作答均可）。
 - ▶ 3. 自出题目环节：第2章内容中，自拟6道试题（填空、单选、判断各2题）和答案，并给出简单解析。
-

