

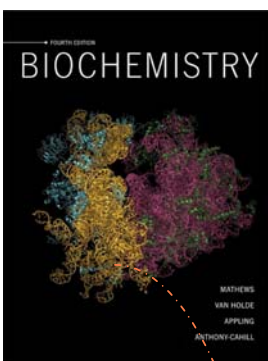
## Chapter 5

# Introduction to Proteins: The **Primary Level** of Protein Structure *Amino acid sequence of a protein*

生化分生科 游佳融  
2014/09

Copyright © 2013 Pearson Canada Inc.

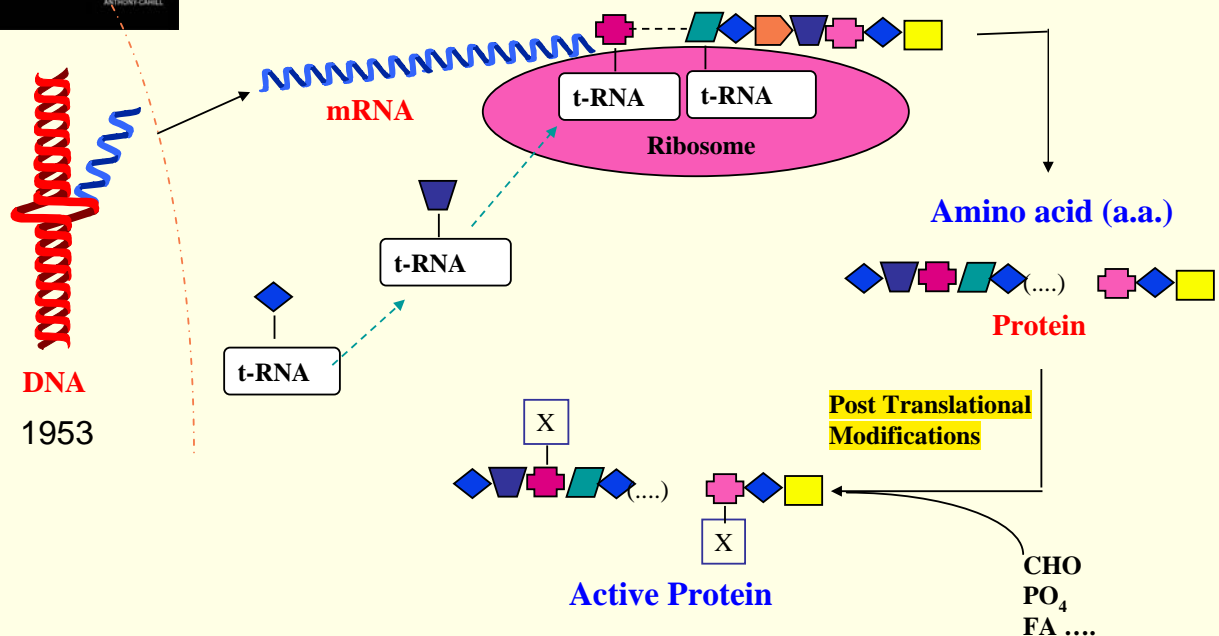
5 - 1



Living cells all have a similar basic chemistry

**Central Dogma 中心法則** –1957, Francis Crick 克里克

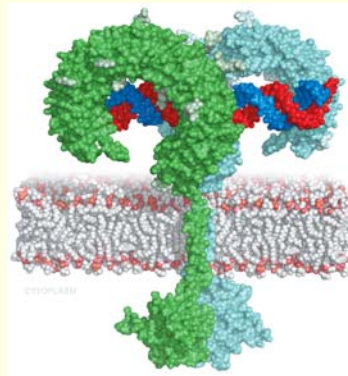
DNA → RNA → protein





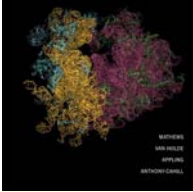
## Proteins play an enormous variety of roles:

- transport and storage of small molecules
- structural framework of cells and tissues
- muscle contraction
- immune responses
- blood clotting
- enzymes-the biological catalysts



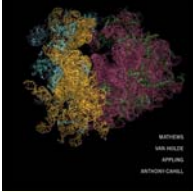
Function:

Structure/conformation  
dependent



***Why** we study the primary level of protein structure (amino acid sequence of a protein)?*

- ◆ The protein structure is able to fulfill a **specific biological function**.
- ◆ The amino acid sequence of a protein **determines its three-dimensional structure**.



## Chapter 5 Outline:

- Structures of the  $\alpha$ -Amino Acids
- Properties of the Amino Acid Side Chains
- Peptides and the Peptide Bond
- Proteins: Polypeptides of Defined Sequence
- From Gene to Protein
- Protein Sequence Homology
- Protein Expression and Purification
- Mass, Sequence, and Amino Acid Analyses of Purified Proteins

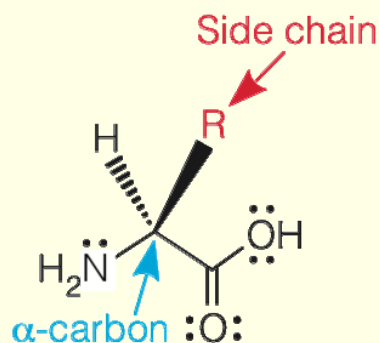


## Structures of the $\alpha$ -Amino Acids

To the  $\alpha$ -carbon of every amino acid are attached:

- an amino group
- a hydrogen atom
- a side chain ("R" group)

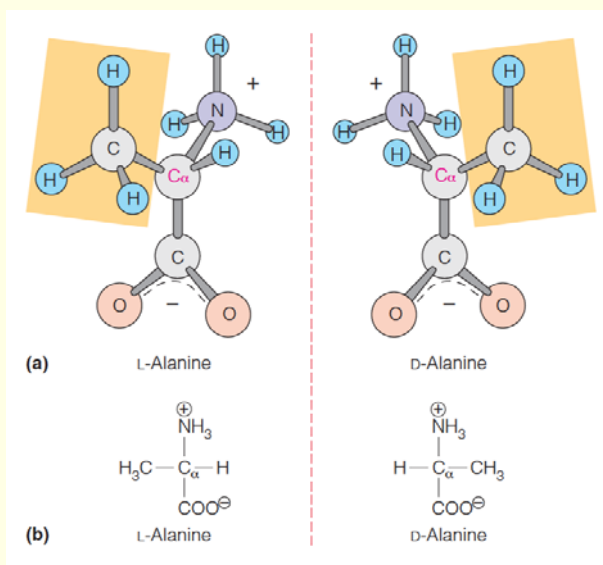
Different  $\alpha$ -amino acids are distinguished by their side chains.



(a)

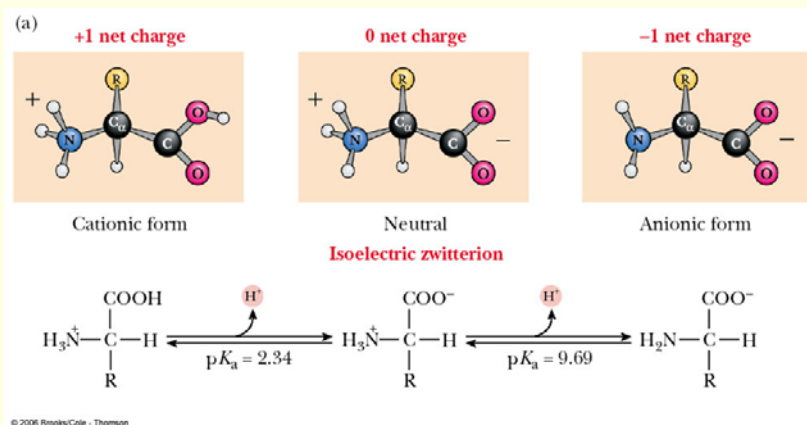
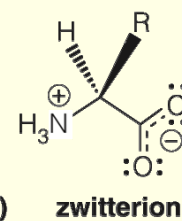
## Stereochemistry of the $\alpha$ -Amino Acids

立體化學



- $\alpha$ -Amino acids' stereochemistry is designated as D- or L-, which is best visualized from its Fischer projection
- There is a preference for L-amino acids in proteins
- Compare this to the preference for D-configured carbohydrates.

- The  $pK_a$  of the carboxylic acid is about 2
- The  $pK_a$  of the  $\alpha$ -amino group is about 10
- Therefore, at physiological pH both the carboxylic acid group and the  $\alpha$ -amino group will be ionized, to yield the **zwitterion form** 兩性離子的
- Amino acids are typically written in their **zwitterionic form**



There are 20 different kinds of amino acids are commonly incorporated into proteins.

TABLE 5.1 Properties of the common amino acids found in proteins

Name	Abbreviations 1- and 3-letter codes	pK <sub>a</sub> of α-COOH Group	pK <sub>a</sub> of α-NH <sub>3</sub> <sup>+</sup> Group	pK <sub>a</sub> of Ionizing Side Chain <sup>a</sup>	Residue <sup>b</sup> Mass (daltons)	Occurrence <sup>c</sup> in Proteins (mol %)
Alanine	A, Ala	2.3	9.7	—	71.08	8.7
Arginine	R, Arg	2.2	9.0	12.5	156.20	5.0
Asparagine	N, Asn	2.0	8.8	—	114.11	4.2
Aspartic acid	D, Asp	2.1	9.8	3.9	115.09	5.9
Cysteine	C, Cys	1.8	10.8	8.3	103.14	1.3
Glutamine	Q, Gln	2.2	9.1	—	128.14	3.7
Glutamic acid	E, Glu	2.2	9.7	4.2	129.12	6.6
Glycine	G, Gly	2.3	9.6	—	57.06	7.9
Histidine	H, His	1.8	9.2	6.0	137.15	2.4
Isoleucine	I, Ile	2.4	9.7	—	113.17	5.5
Leucine	L, Leu	2.4	9.6	—	113.17	8.9
Lysine	K, Lys	2.2	9.0	10.0	128.18	5.5
Methionine	M, Met	2.3	9.2	—	131.21	2.0
Phenylalanine	F, Phe	1.8	9.1	—	147.18	4.0
Proline	P, Pro	2.0	10.6	—	97.12	4.7
Serine	S, Ser	2.2	9.2	—	87.08	5.8
Threonine	T, Thr	2.6	10.4	—	101.11	5.6
Tryptophan	W, Trp	2.4	9.4	—	186.21	1.5
Tyrosine	Y, Tyr	2.2	9.1	10.1	163.18	3.5
Valine	V, Val	2.3	9.6	—	99.14	7.2

Copyright © 2013 Pearson Canada Inc.

5 - 9

Some biologically important amino acids not typically found in protein

TABLE 5.2 Some biologically important amino acids not typically found in proteins

Name	Formula	Biochemical Source, Function
β-Alanine	$\text{H}_3\text{N}^+ - \text{CH}_2 - \text{CH}_2 - \text{COO}^-$	Found in the vitamin pantothenic acid and in some important natural peptides
D-Alanine	$\begin{array}{c} \text{COO}^- \\   \\ \text{H} - \text{C} - \text{NH}_3^+ \\   \\ \text{CH}_3 \end{array}$	In polypeptides in some bacterial cell walls
γ-Aminobutyric acid	$\text{H}_3\text{N}^+ - \text{CH}_2 - \text{CH}_2 - \text{CH}_2 - \text{COO}^-$	Brain, other animal tissues; functions as neurotransmitter
D-Glutamic acid	$\begin{array}{c} \text{COO}^- \\   \\ \text{H} - \text{C} - \text{NH}_3^+ \\   \\ \text{CH}_2 \\   \\ \text{CH}_2 - \text{COO}^- \end{array}$	In polypeptides in some bacterial cell walls
L-Homocysteine	$\begin{array}{c} \text{COO}^- \\   \\ \text{H}_3\text{N}^+ - \text{C} - \text{H} \\   \\ \text{CH}_2 - \text{CH}_2\text{SH} \end{array}$	Many tissues; precursor for methionine biosynthesis
L-Ornithine	$\begin{array}{c} \text{COO}^- \\   \\ \text{H}_3\text{N}^+ - \text{C} - \text{H} \\   \\ \text{CH}_2 - \text{CH}_2 - \text{CH}_2\text{NH}_3^+ \end{array}$	Many tissues; an intermediate in arginine synthesis
Sarcosine	$\begin{array}{c} \text{COO}^- \\   \\ \text{CH}_3 - \text{N} - \text{CH}_2 - \text{COO}^- \\   \\ \text{H} \end{array}$	Many tissues; intermediate in amino acid synthesis
L-Thyroxine	$\begin{array}{c} \text{COO}^- \\   \\ \text{H}_3\text{N}^+ - \text{C} - \text{H} \\   \\ \text{CH}_2 - \text{C}_6\text{H}_3\text{I}_2 - \text{O} - \text{C}_6\text{H}_3\text{I}_2\text{OH} \end{array}$	Thyroid gland; is thyroid hormone (I = iodine)

Copyright © 2013 Pearson Canada Inc.

5 - 10

## Properties of the Amino Acid Side Chains

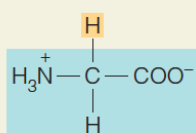
The 20 common amino acids are **classified by their side chains**:

- **Aliphatic (脂肪族的)**
- **Hydroxyl or sulfur-containing (含有 -OH 或 S)**
- **Aromatic (芳香族的)**
- **Basic (鹼性)**
- **Acidic and their amides (酸性以及其氨基化合物)**

## Aliphatic (脂肪族的)

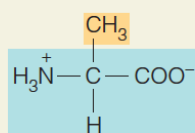
### ALIPHATIC AMINO ACIDS

Smallest



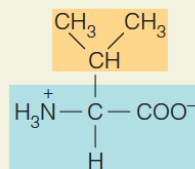
Glycine (Gly) G

甘氨酸



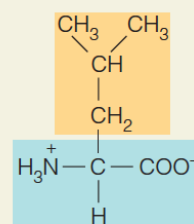
Alanine (Ala) A

丙氨酸



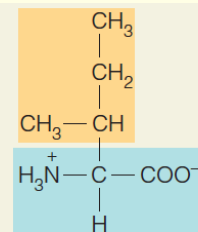
Valine (Val) V

纈氨酸



Leucine (Leu) L

白氨酸



Isoleucine (Ile) I

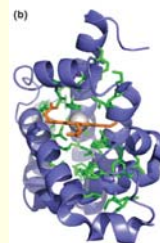
異白氨酸

- The more **hydrophobic** amino acids such as isoleucine are usually found **within the core of a protein molecule**, where they are shielded from water.

(a)

VLSEGEWQLV LHVWAKVEAD VAGHGQDILI RLFKSHPETL EKFD RFKHLK  
TEAEMKASED LKKHGVTVLT ALGAILKKKG HHEAELKPLA QSHATKHKIP  
IKYLEFISEA I IHLVLSRHP GDFGADAQGA MNKALELFRK DIAAKYKELG  
YQG

(b)



(c)

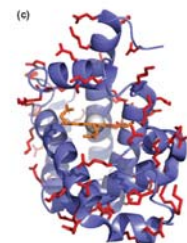
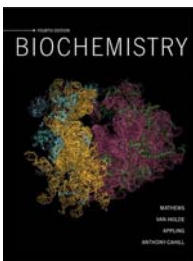


Fig. 6-20



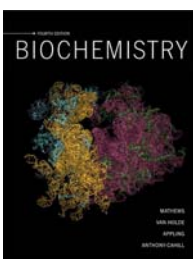
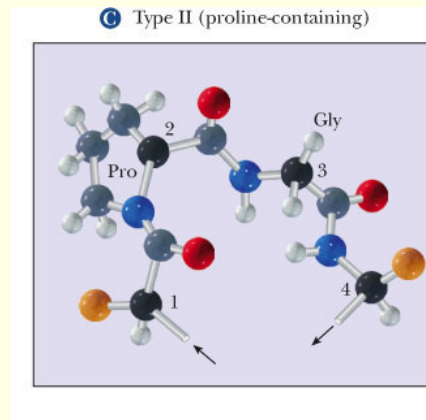
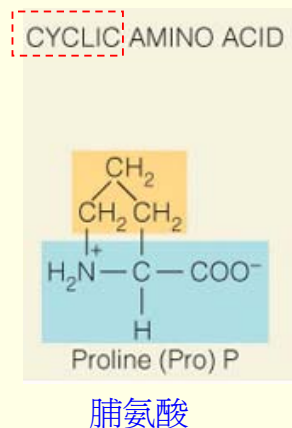


## Biochemistry, 4<sup>th</sup> Edition

● **Proline**, is the only amino acid in this group in which the side chain forms a covalent bond with the  $\alpha$ -amino group.

● The proline side chain has a **primarily aliphatic** character; however, it is frequently **found on the surfaces of proteins** due to **its unique** structural constraints. 獨特

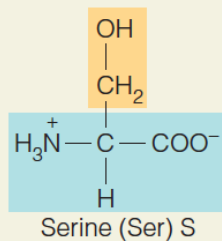
● The rigid ring of proline is well-suited to those sites in a protein structure where the protein must fold back on itself (so-called **"turns"**). 轉彎



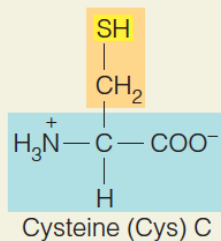
## Biochemistry, 4<sup>th</sup> Edition

### Hydroxyl or sulfur-containing (含有 -OH 或 S)

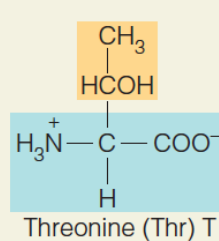
#### AMINO ACIDS WITH HYDROXYL- OR SULFUR-CONTAINING SIDE CHAINS



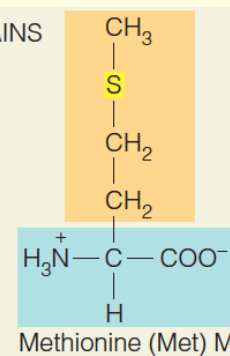
絲胺酸



半胱胺酸



蘇胺酸



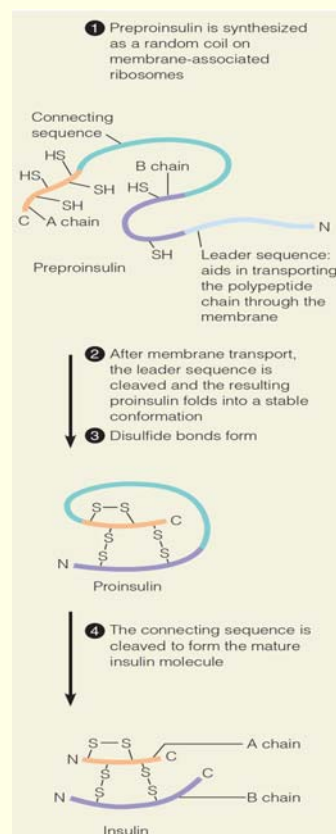
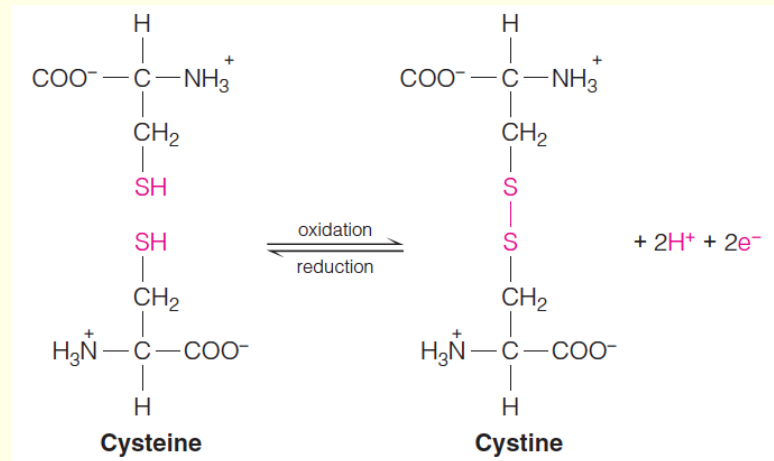
甲硫胺酸

- The **-OH** group of serine and the **-SH** group of cysteine are **good nucleophiles** and often play **key roles in enzyme activity**. 親核性

## Cysteine (半胱氨酸)

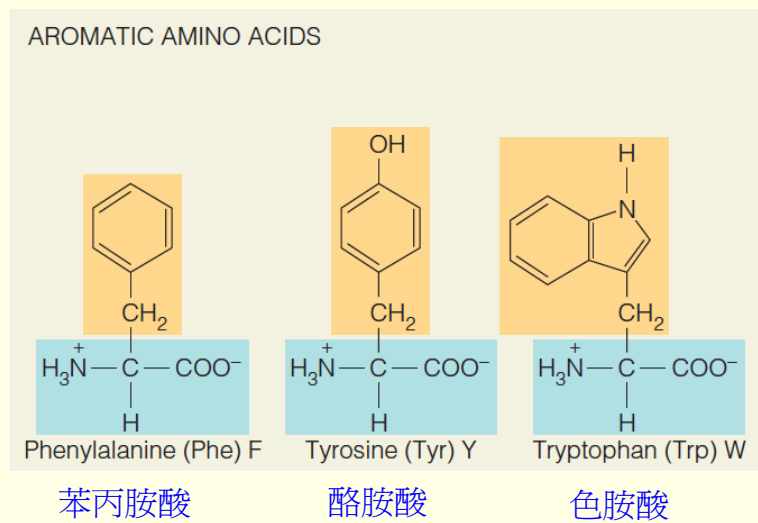
## 雙硫鍵

The oxidation of two cysteine side chains yields a **disulfide bond**. The product of this oxidation is given the name **cystine**. (胱氨酸，双硫丙氨酸)





## Aromatic (芳香族的)



- **Phenylalanine** is one of **the most hydrophobic** amino acids.
- **Tyrosine and tryptophan** have some hydrophobic character as well, but it is tempered by the **polar groups** in their side chains.

Absorption of **ultraviolet light** by aromatic amino acid

應用於偵測與半定量蛋白質

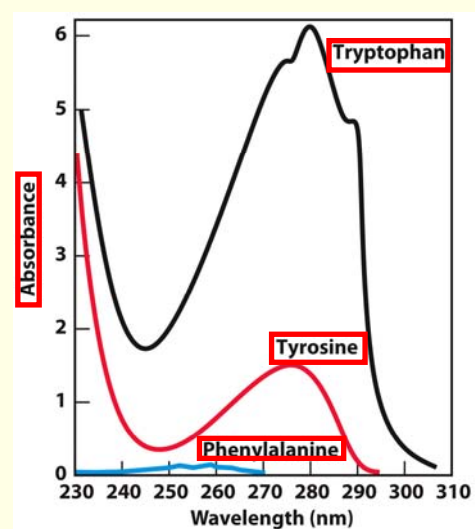
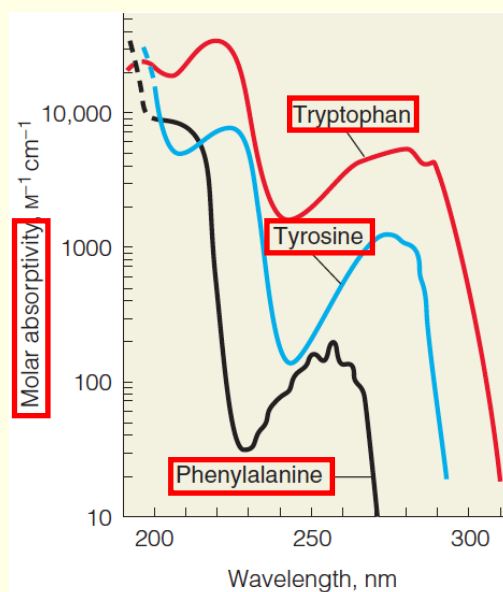
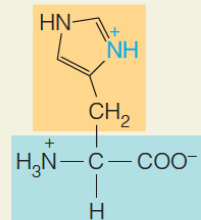


Figure 3-6  
Lehninger Principles of Biochemistry, Sixth Edition  
© 2013 W. H. Freeman and Company

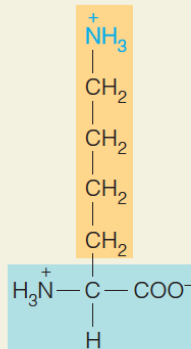
280 nm

## Basic (鹼性)

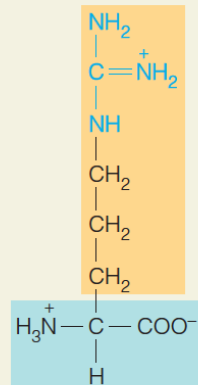
## BASIC AMINO ACIDS



組氨酸



離氨酸



精氨酸

- Lysine ( $pK_a = 10.0$ ) and arginine ( $pK_a = 12.5$ ) are the more basic amino acids.
- Their side chains are **almost always positively charged** under physiological conditions.
- guanidino group ( $\text{HNC}(\text{NH}_2)_2$ )** of arginine, strong base

## Interaction Between A Glycosaminoglycan and Its Binding Protein

Structure of fibroblast growth factor receptor (FGFR) and heparin

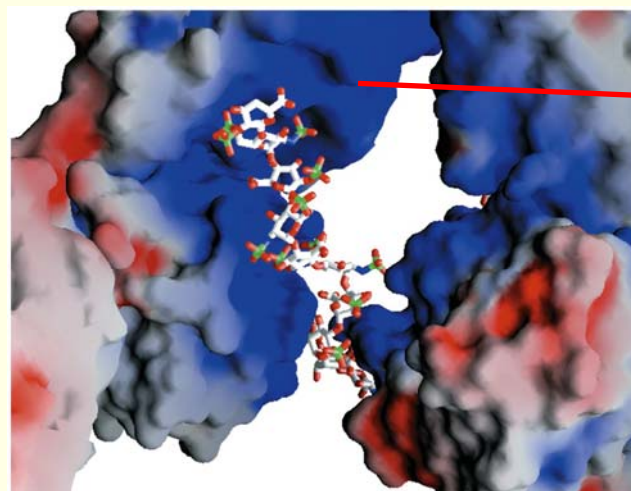
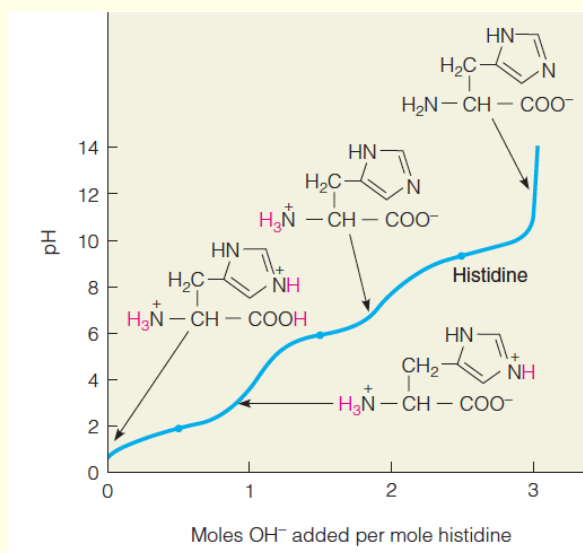
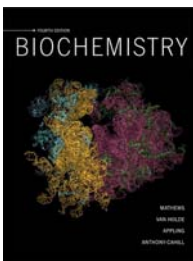
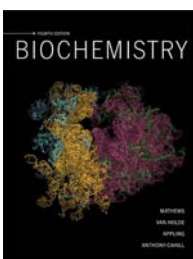


Figure 7-23  
Lehninger Principles of Biochemistry, Fifth Edition  
© 2008 W.H. Freeman and Company

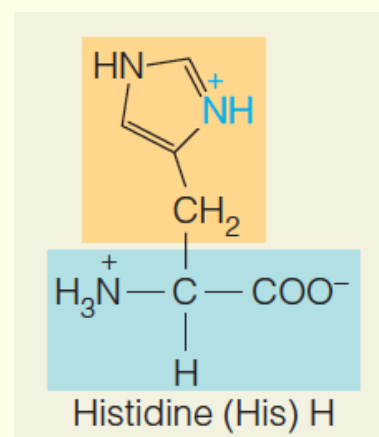
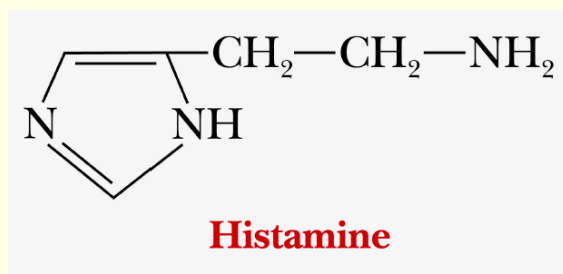
Blue: Basic Red: Acid



- Histidine is the least basic of the three, and as its titration curve shows, the imidazole ring in the side chain of the free amino acid loses its proton at about pH 6.
- Because the histidine side chain has a  $pK_a$  near physiological pH, it often plays a role in enzymatic catalysis involving proton transfer.



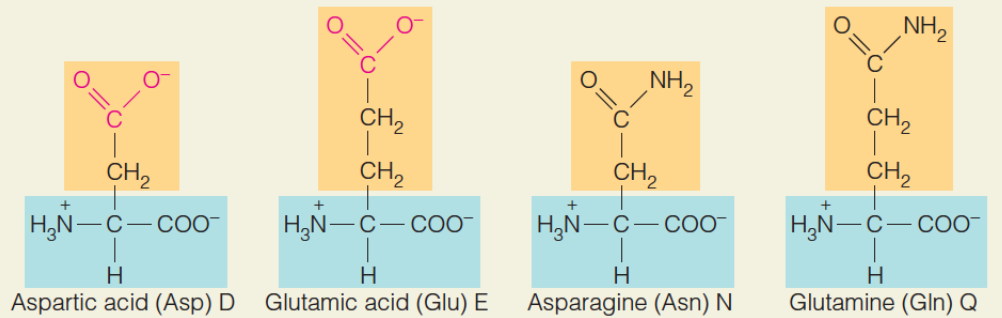
## Amino acid function other than in peptides



組織胺 (組胺酸的衍生物)

## Acidic and their amides (酸性以及其氨基化合物)

## ACIDIC AMINO ACIDS AND THEIR AMIDES



天門冬胺酸

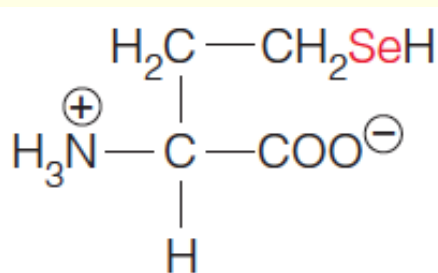
谷胺酸

天門冬醯胺

谷氨醯胺

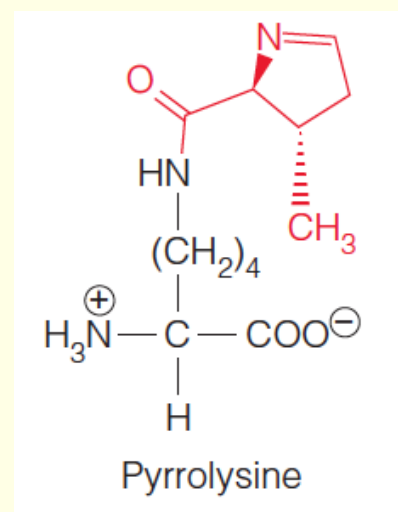
- Aspartic acid ( $pK_a = 3.9$ ) and glutamic acid ( $pK_a = 4.2$ ) typically carry negative charges at pH 7.
- Unlike their acidic analogs, asparagine and glutamine have **uncharged polar side chains**.
- Like the basic and acidic amino acids, Asn and Gln are **hydrophilic** and tend to be on **the surface of a protein molecule**, in contact with the surrounding water.

- Two amino acids, **selenocysteine** and **pyrrolysine**, are encoded genetically and **incorporated into proteins**; however, **they are found in a relatively small number of proteins**.



Selenocysteine

21st aa.

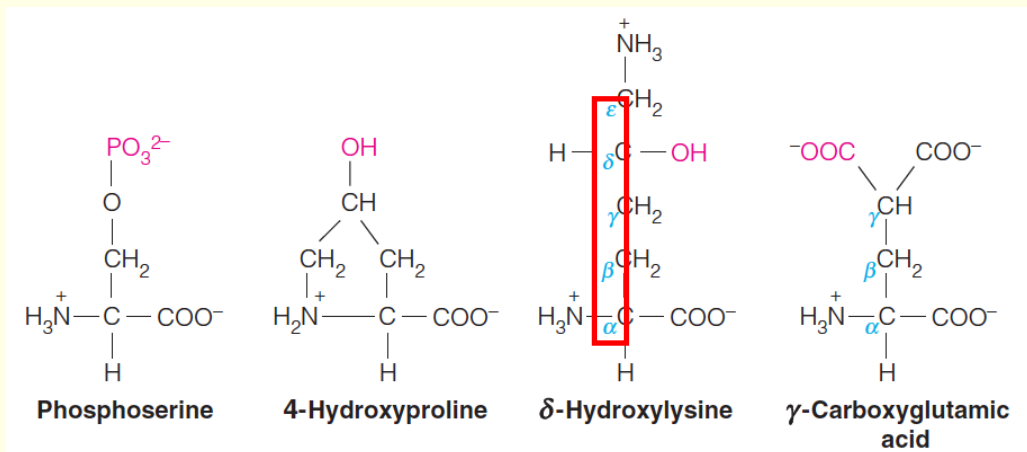


Pyrrolysine

22nd aa.

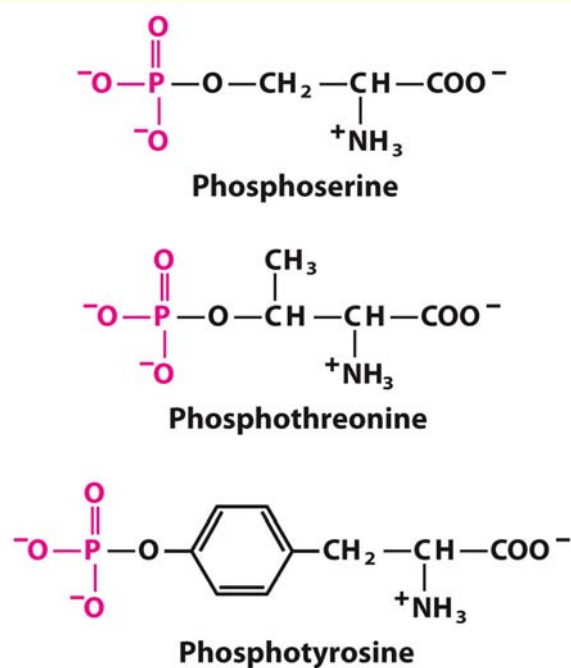
## Modified amino acids

Amino acids can undergo **post-translational modification** (轉譯後修飾) resulting in modified amino acids with unique properties

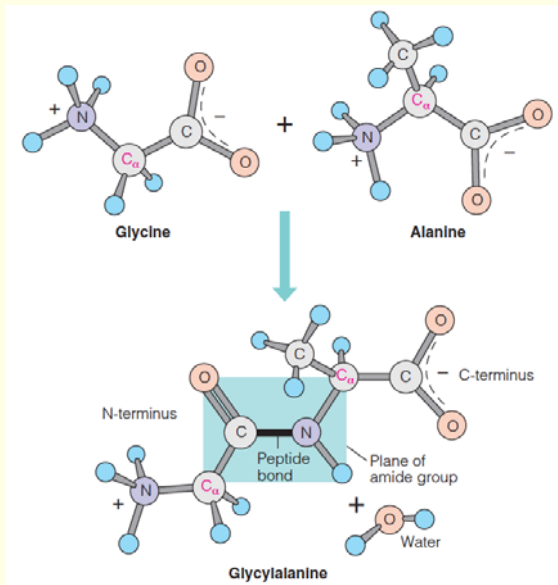


Rich in Collagen

## Protein phosphorylation

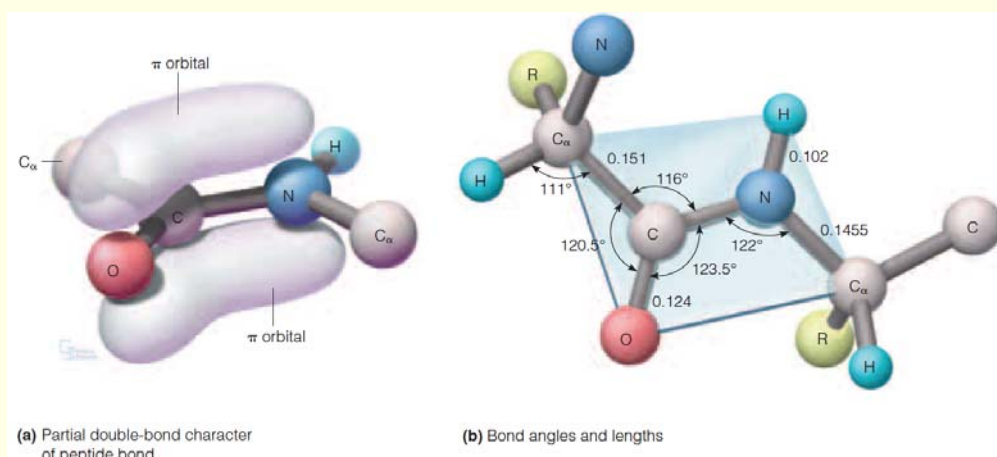


## Peptide Bond (胜肽键)



- Amino acids can be covalently linked together by formation of an **amide bond** between the  **$\alpha$ -carboxylic acid group** on one amino acid and the  **$\alpha$ -amino group** on another.
- Peptide bond**
- Covalent bond**
- Nearly planar**
- A peptide composed of 2 amino acids is called a **dipeptide**.

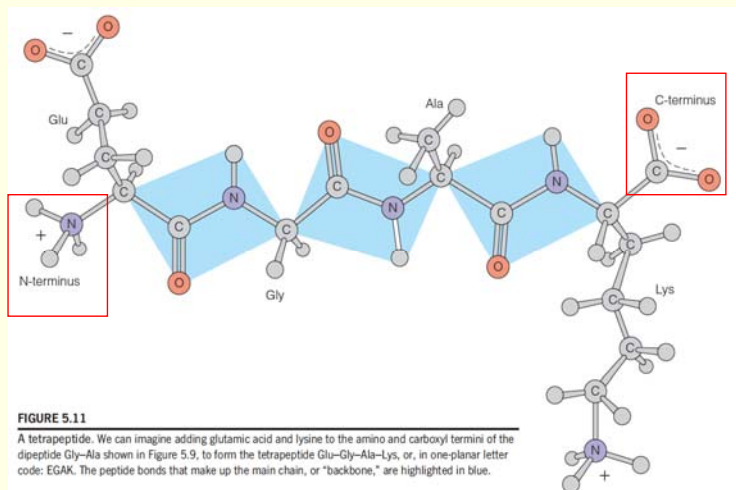
## The property of peptide bond



- The amide carbonyl and amide bonds are nearly parallel.
- The six atoms shown in the blue rectangle in usually **coplanar**.
- There is **little twisting possible** around the peptide bond because the bond has a substantial fraction of **double-bond character**.



## Peptides (胜肽)



Glu-Gly-Ala-Lys

or

EGAK

- **N-terminus** to **C-terminus**.
- **EGAK ≠ KAGE**
- The portion of each amino acid remaining in the chain is called an **amino acid residue**. (残基)

## Peptide chains (胜肽鏈)

- Chains containing only a few amino acid residues (like a tetrapeptide) are collectively referred to as **oligopeptides**.
- If the chain is longer (**>15-20 residues**), it is called a **polypeptide**.
- Polypeptides greater than **~50 residues** are generally referred to as **proteins**.
- Note that **most globular proteins contain 250–600 amino acid residues**.

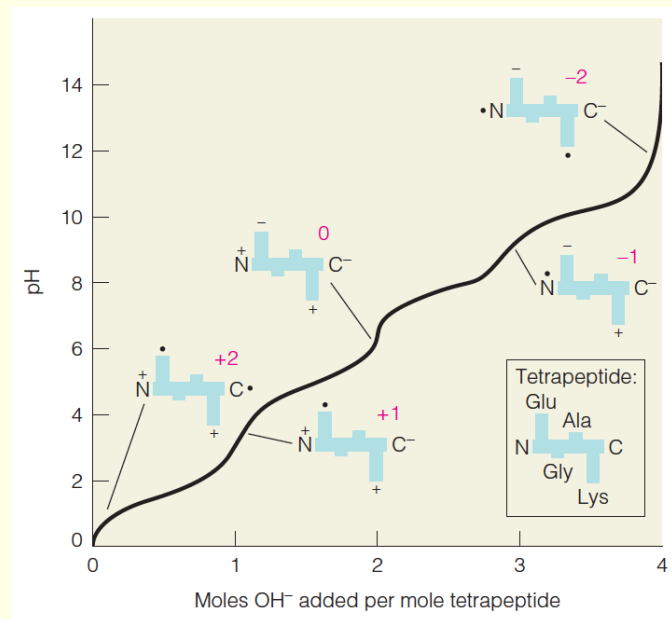
TABLE 5.1 P

Name	Residue <sup>b</sup> Mass (daltons)
Alanine	71.08
Arginine	156.20
Asparagine	114.11
Aspartic acid	115.09
Cysteine	103.14
Glutamine	128.14
Glutamic acid	129.12
Glycine	57.06
Histidine	137.15
Isoleucine	113.17
Leucine	113.17
Lysine	128.18
Methionine	131.21
Phenylalanine	147.18
Proline	97.12
Serine	87.08
Threonine	101.11
Tryptophan	186.21
Tyrosine	163.18
Valine	99.14

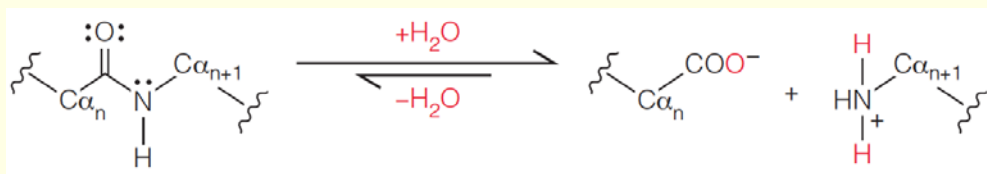
## Polyampholytic behavior of a tetrapeptide

• This titration curve for the tetrapeptide Glu-Gly-Ala-Lys (EGAK) shows the major forms present at several pH values.

• Note the net charges for the different ionization states in red.



## Hydrolysis of peptide bond



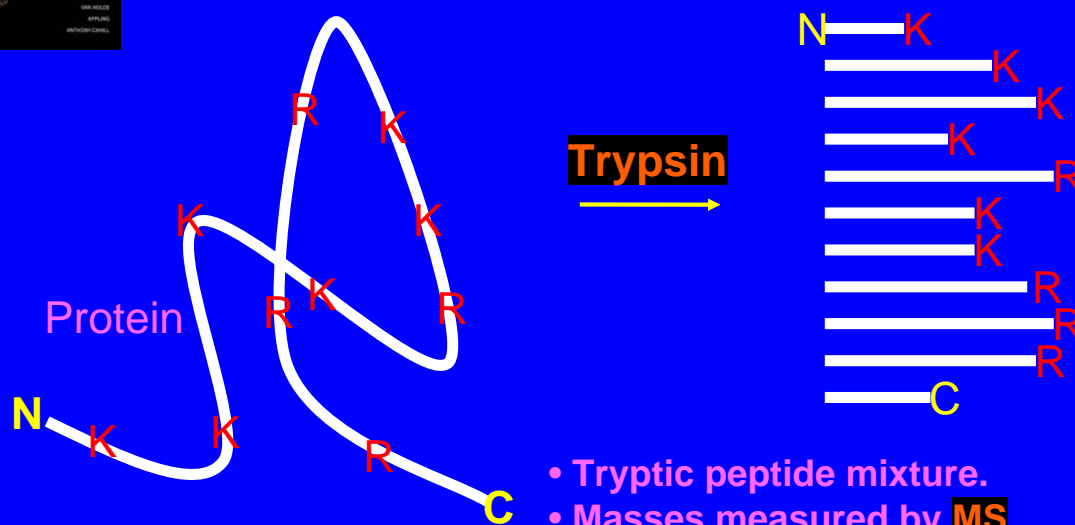
- Peptide bond hydrolysis can be achieved by:
  - Strong mineral acid (e.g., 6 M HCl) cleaves all peptide bonds (including the Asn and Gln amide bonds)
  - Chemicals that cleave at specific sites (e.g., CNBr cleaves at Met)
  - Proteolytic enzymes (proteases) that cleave at specific sites

## Cleavage and proteolysis of peptide chains

TABLE 5.4 The sequence specificities of some proteolytic enzymes

$\text{N-terminal} \cdots \text{---} \underset{\text{H}}{\underset{ }{\text{N}}} \text{---} \underset{\text{H}}{\underset{ }{\text{C}}} \overset{\text{R}_1}{\overset{ }{\text{C}}} \overset{\text{O}}{\parallel} \text{---} \underset{\text{H}}{\underset{ }{\text{N}}} \text{---} \underset{\text{H}}{\underset{ }{\text{C}}} \overset{\text{R}_2}{\overset{ }{\text{C}}} \overset{\text{O}}{\parallel} \text{---} \cdots \text{C-terminal}$		
Enzyme	Preferred Site <sup>a</sup>	Source
<b>Trypsin</b>	R <sub>1</sub> = Lys, Arg	From digestive systems of animals, many other sources
Chymotrypsin	R <sub>1</sub> = Tyr, Trp, Phe, Leu	Same as trypsin
Thrombin	R <sub>1</sub> = Arg	From blood; involved in coagulation
<b>V-8 protease</b>	R <sub>1</sub> = Asp, Glu	From <i>Staphylococcus aureus</i>
Prolyl endopeptidase	R <sub>1</sub> = Pro	Lamb kidney, other tissues
Subtilisin	Very little specificity	From various bacilli
Carboxypeptidase A	R <sub>2</sub> = C-terminal amino acid	From digestive systems of animals
Thermolysin	R <sub>2</sub> = Leu, Val, Ile, Met	From <i>Bacillus thermoproteolyticus</i>

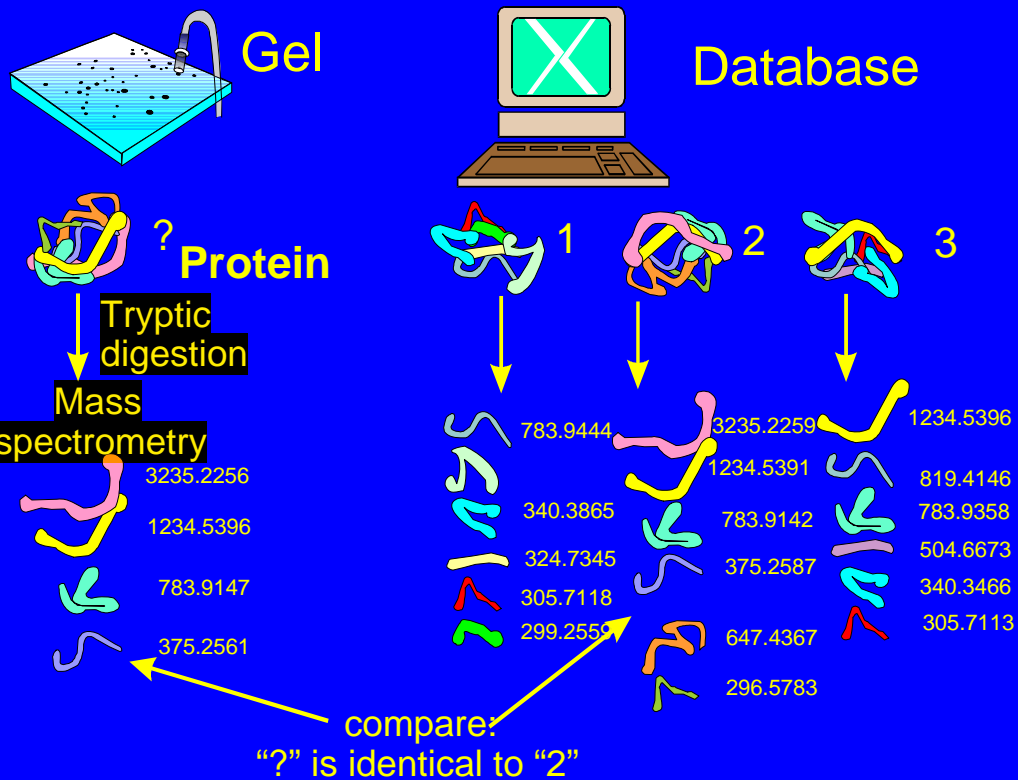
## Protein identification by peptide mass fingerprinting (PMF) 胜肽質量指紋



- Tryptic peptide mixture.
- Masses measured by **MS**.
- Every peptide has a basic C-terminus.

A protein can be identified in a database by **matching masses** of a subset of the tryptic peptides against calculated values.

## Peptide Mass Fingerprinting with Mass spectrometry

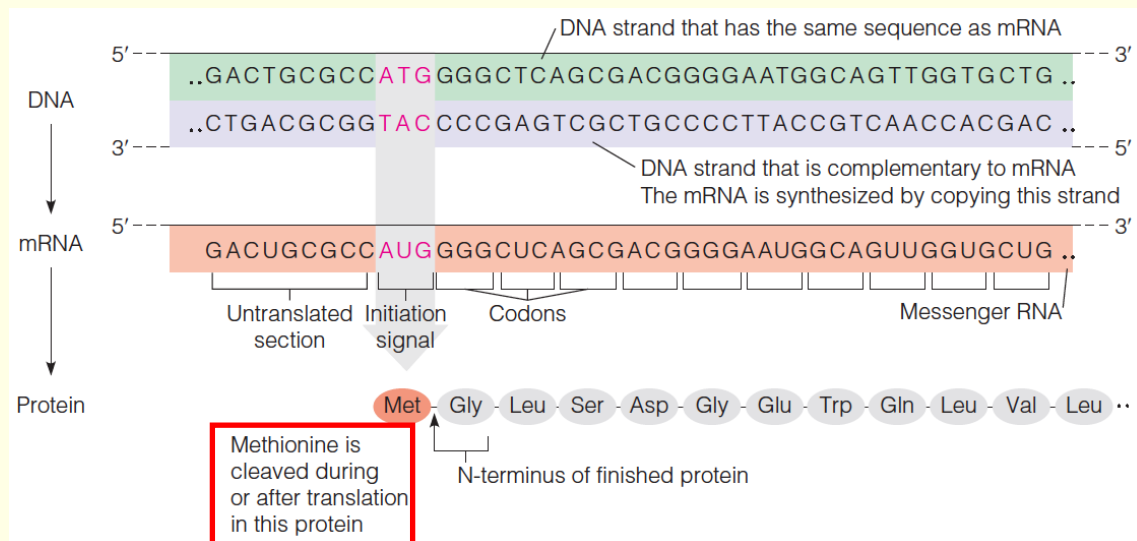


## Protein sequence: from gene to protein

- Remember the **central dogma**:
  - DNA is transcribed into RNA
  - RNA is translated into protein
- Therefore, DNA codes for protein.

		Second position				
		U	C	A	G	
First position U	UUU } Phe	UCU } Ser	UAU } Tyr	UGU } Cys	U	Third position
	UUC } Phe	UCC } Ser	UAC } Tyr	UGC } Cys	C	
	UUA } Leu	UCA } Ser	UAA Stop	UGA Stop	A	
	UUG } Leu	UCG } Ser	UAG Stop	UGG Trp	G	
C	CUU } Leu	CCU } Pro	CAU } His	CGU } Arg	U	
	CUC } Leu	CCC } Pro	CAC } His	CGC } Arg	C	
	CUA } Leu	CCA } Pro	CAA } Gln	CGA } Arg	A	
	CUG } Leu	CCG } Pro	CAG } Gln	CGG } Arg	G	
A	AUU } Ile	ACU } Thr	AAU } Asn	AGU } Ser	U	
	AUC } Ile	ACC } Thr	AAC } Asn	AGC } Ser	C	
	AUA } Ile	ACA } Thr	AAA } Lys	AGA } Arg	A	
	AUG Met/start	ACG } Thr	AAG } Lys	AGG } Arg	G	
G	GUU } Val	GCU } Ala	GAU } Asp	GGU } Gly	U	
	GUC } Val	GCC } Ala	GAC } Asp	GGC } Gly	C	
	GUA } Val	GCA } Ala	GAA } Glu	GGA } Gly	A	
	GUG } Val	GCG } Ala	GAG } Glu	GGG } Gly	G	

## From Gene to Protein



## Proteins: polypeptides of defined sequence

- Every protein has a **defined number and order** of amino acid residues.
- Unique sequence** (獨特序列)
- primary structure** of the protein.

Key:

- Identical amino acids 一致
- Conservative substitutions 保守
- Nonconservative substitutions

Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Human	G	L	S	D	G	E	W	Q	L	V	L	N	V	W	G
Whale	V	L	S	E	G	E	W	Q	L	V	L	H	V	W	A

Number	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Human	K	V	E	A	D	I	P	G	H	G	Q	E	V	L	I
Whale	K	V	E	A	D	V	A	G	H	G	Q	D	I	L	I

Number	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Human	R	L	F	K	G	H	P	E	T	L	E	K	F	D	K
Whale	R	L	F	K	S	H	P	E	T	L	E	K	F	D	R

Number	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60
Human	F	K	H	L	K	S	E	D	E	M	K	A	S	E	D
Whale	F	K	H	L	K	T	E	A	E	M	K	A	S	E	D

Number	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75
Human	L	K	K	H	G	A	T	V	L	T	A	L	G	G	I
Whale	L	K	K	H	G	V	T	V	L	T	A	L	G	A	I

Number	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90
Human	L	K	K	K	G	H	H	E	A	E	I	K	P	L	A
Whale	L	K	K	K	G	H	H	E	A	E	L	K	P	L	A

Number	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105
Human	Q	S	H	A	T	K	H	K	I	P	V	K	Y	L	E
Whale	Q	S	H	A	T	K	H	K	I	P	I	K	Y	L	E

Number	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120
Human	F	I	S	E	C	I	I	Q	V	L	Q	S	K	H	P
Whale	F	I	S	E	A	I	I	H	V	L	H	S	R	H	P

Number	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135
Human	G	D	F	G	A	D	A	Q	G	A	M	N	K	A	L
Whale	G	N	F	G	A	D	A	Q	G	A	M	N	K	A	L

Number	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153
Human	E	L	F	R	K	D	M	A	S	N	Y	K	E	L	G	F	Q	G
Whale	E	L	F	R	K	D	I	A	A	K	Y	K	E	L	G	Y	Q	G

**Sequence homology** (序列相似性) between myoglobin in humans vs. whales.



## Protein Sequence Homology

- **Sequence identity** refers to those parts of the amino acid sequence that are an **exact match**. 一致度
- **Sequence similarity** is based on the classification of the chemical properties of the side chains such as hydrophobicity, polarity, and charge 相似度
- A widely used protocol for conducting database searches is the **Basic Local Alignment Search Tool (BLAST)**.

Score = 30.8 bits (68), **Expect = 6e-06**, Method: Compositional matrix adjust.  
Identities = 32/133 (25%), Positives = 48/133 (37%), Gaps = 40/133 (30%)

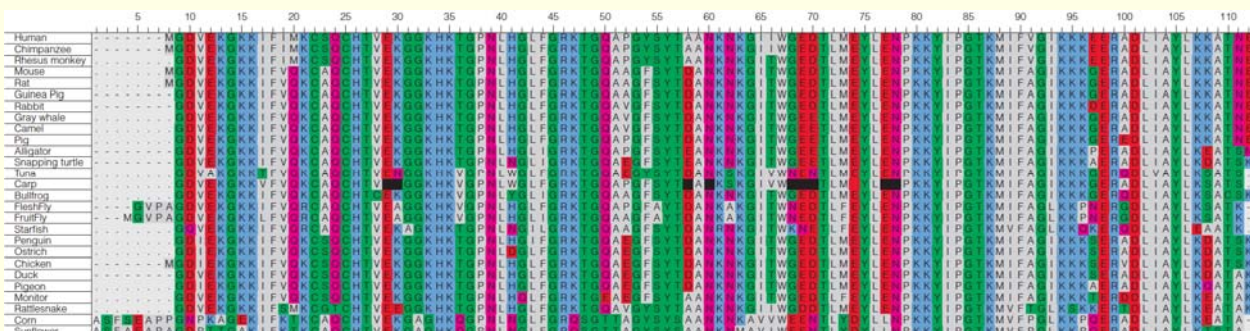
Human Mb	2	LSDGEWQLVNWGKVEADIPGHGOEVLI RLFKGFHETLEKFDKFKHLKSEDEMKASEDL	61
Human α	2	LS + V WGKV A +G EL R+F P T F F	46
Human Mb	62	KKHGATVLTALGG ILKKKGHIHEAEIKPLAQSHATKHKI -PVKYLEFISECI IQVLQSKHP	120
Human α	47	-----ALSALSDI----- HA K ++ PV ++ +S C++ L + P	82
Human Mb	121	GDFGADAQGAMNK 133	
Human α	83	+F ++K AEFTPAVHASL DK 95	

Copyright © 2013 Pearson Canada Inc.

5 - 39

## Protein Sequence Homology

- Protein sequences are classified as "**homologous**" in cases where any sequence similarity is thought to be the result of a common **evolutionary ancestry**. 同源
- **Amino acid conservation** occurs if only one amino acid is found at a given position within an alignment of homologous proteins, that amino acid is said to be absolutely "**conserved**".



Sequence alignment and a phylogenetic tree for cytochrome c from different organisms.

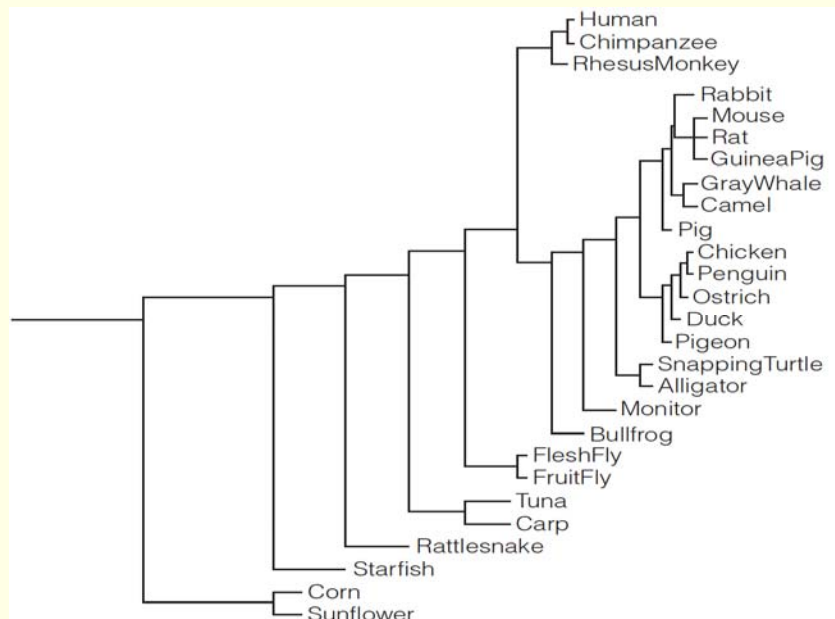
Copyright © 2013 Pearson Canada Inc.

5 - 40



## Sequence homology and its evolutionary relationship

A **greater degree** of protein sequence homology indicates a **closer evolutionary relationship**.

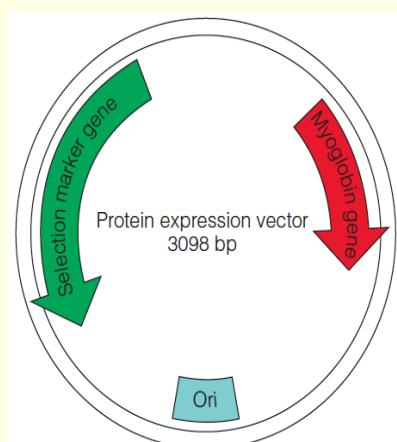


Copyright © 2013 Pearson Canada Inc.

5 - 41

## Recombinant protein expression technology

### 重組蛋白質表現技術

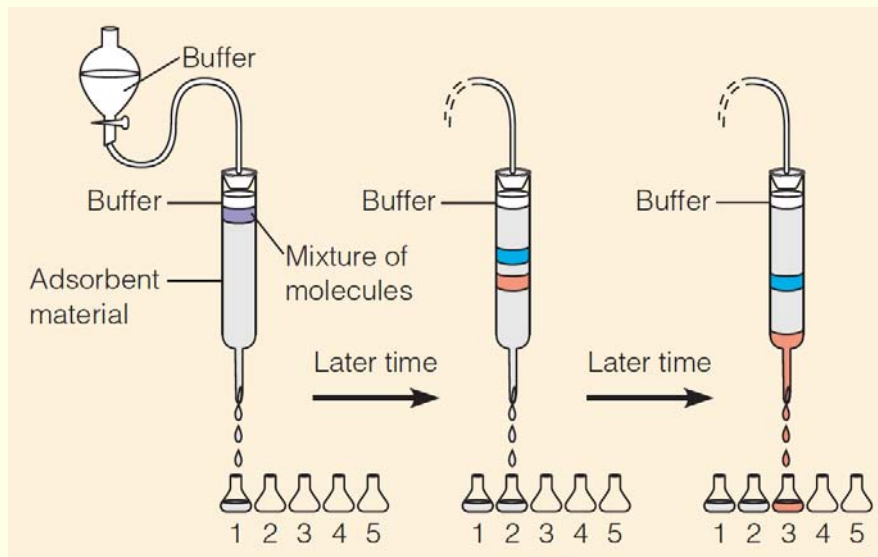


- The **primary structure (sequence)** of a protein corresponds to the DNA (gene) sequence.
- *E. coli* can be made to take up small circular DNA molecules, called "**expression vectors**" (表現載體).
- **Plasmid (質體)**.

Copyright © 2013 Pearson Canada Inc.

5 - 42

## Protein Purification (I)



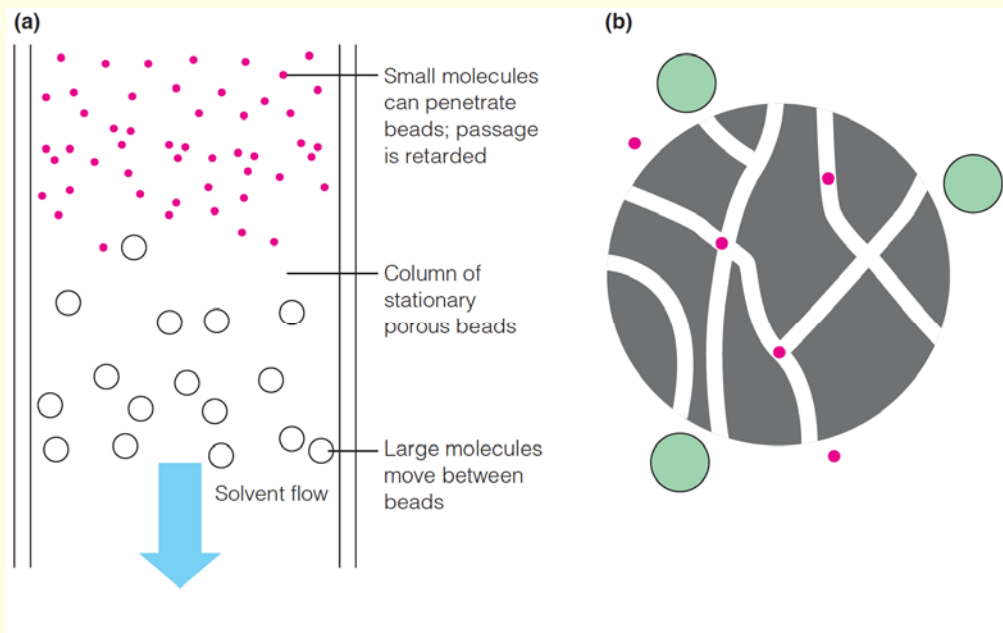
Proteins are generally detected by UV absorbance at 280 nm as they elute from the column.

## Protein Purification (II)

- **Size** (size-exclusion chromatography)  
分子大小，分子篩
- **Ionic charge** (ion-exchange chromatography)  
陰陽離子交換
- **Affinity to ligands** (affinity chromatography)  
親和層析

## Size-exclusion chromatography (aka gel-filtration)

### 分子篩層析



## Ion-exchange chromatography

### 陰陽離子交換層析

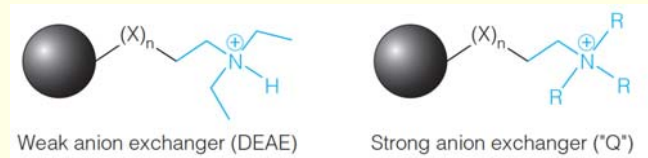
**Ion-exchange chromatography** separates molecules on the basis of their *electrical charge*.

The strength of interaction between a protein molecule and an ion exchange matrix depends on

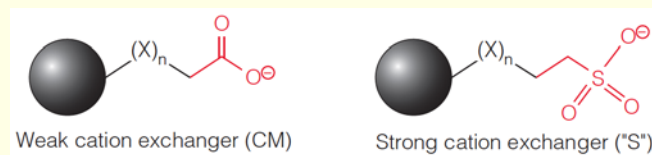
- the **charge density** on the protein
- the **ionic strength** of the mobile phase, which is always a **buffered solution**.

## Two main types of ion exchange matrices

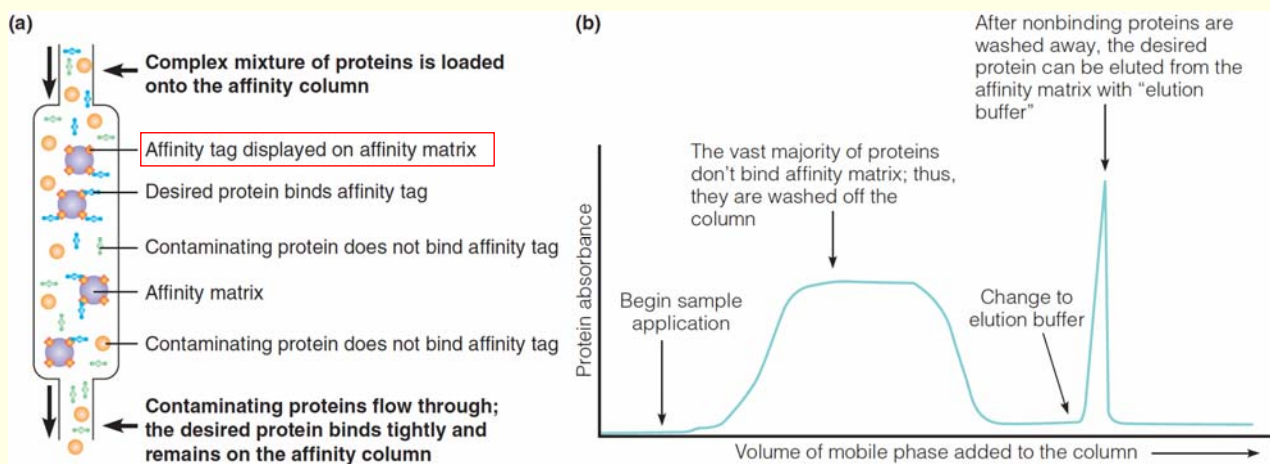
**Anion exchangers** - carry a positive charge and bind to negatively charged proteins 陰離子交換樹脂



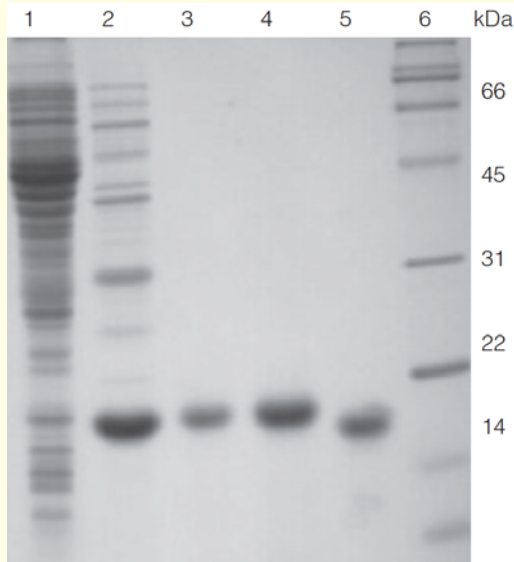
**Cation exchangers** - carry a negative charge and bind to positively charged proteins 陽離子交換樹脂



**Affinity chromatography**  
親和層析法



## Characterization of Purified Proteins (I)



**SDS- PAGE** provides:

- the **purity** of the protein
- an approximate **molecular weight**

測定純度與分子量

Lane 1: *E. coli* lysate (after centrifugation)

Lane 2: IMAC purified proteins

Lane 3: 2 mg of the mutant myoglobin after SEC

Lane 4: 10 mg of the mutant myoglobin after SEC

Lane 5: SEC-purified wild-type myoglobin

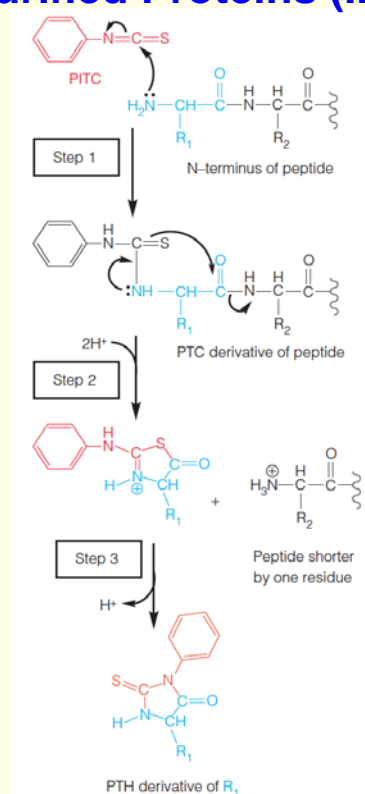
Lane 6: Protein molecular weight markers

## Characterization of Purified Proteins (II)

胺基酸序列分析

Unknown proteins are also characterized by their amino acid sequence.

The Edman method is based on the **stepwise removal of amino acids** from the *N*-terminus of a peptide by a series of chemical reactions called the "**Edman degradation**".



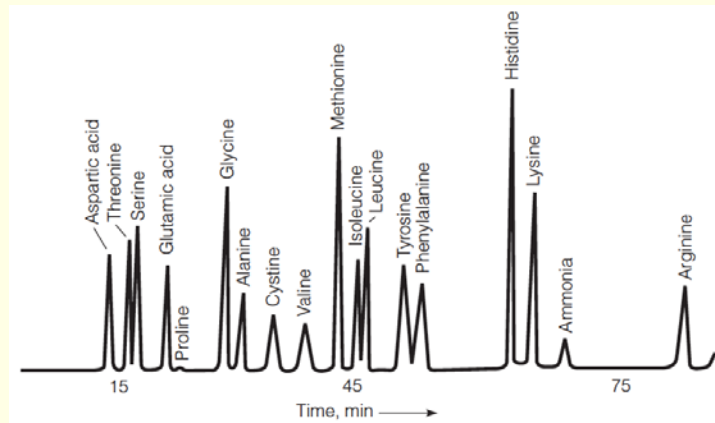
## Characterization of Purified Proteins (III)

**Amino acid analysis** can be used to determine the amino acid composition of a protein. 氨基酸分析

Three steps are involved:

1. **Hydrolysis** of the protein to its constituent amino acids.
2. **Separation** of the amino acids in the mixture by **chromatography**
3. **Quantitation** of the individual amino acids.

On an anionic IEC, acidic amino acids elute first, followed by the hydrophobic, the aromatic, then the basic are last.



Copyright © 2013 Pearson Canada Inc.

5 - 51

## Proteomics (蛋白質體學)

- ◆ The complement of proteins present in a given cell make up the so-called **proteome** of that cell.
- ◆ **Global analysis** of the proteome rather than investigating the properties of purified protein in isolation.
- ◆ A typical proteomics experiment includes the following steps:
  1. Separation and isolation of proteins, or **protein fragments**, from cells or an organism.
  2. Identification by **MS-MS sequencing** of a particular protein within the complex mixture.
  3. **Database searching** to identify the target protein, and its putative function.

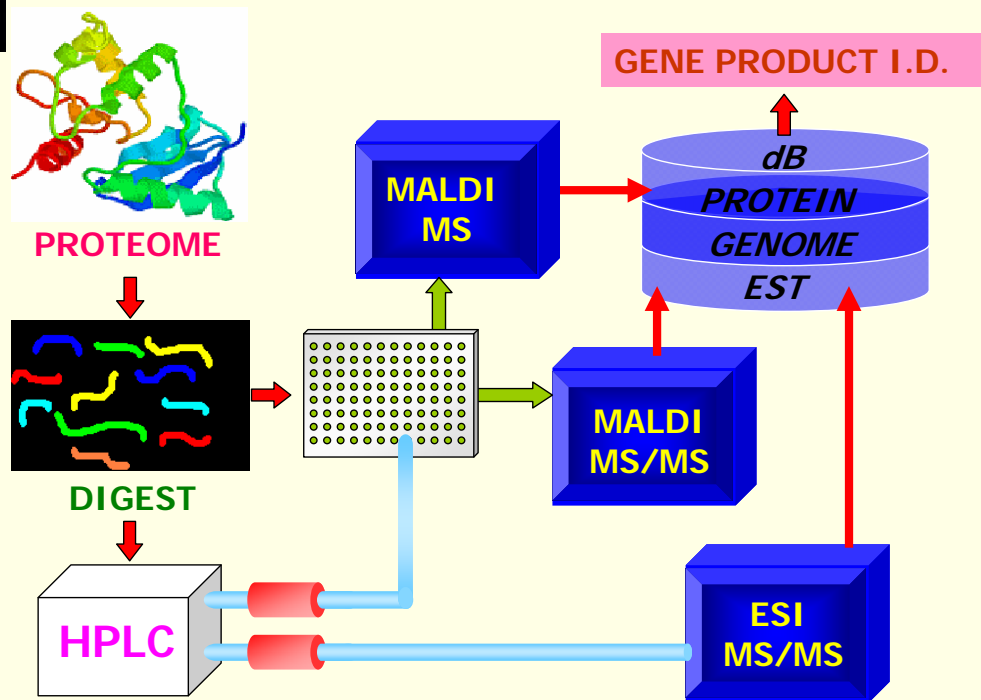
Copyright © 2013 Pearson Canada Inc.

5 - 52



## Basic of protein identification

1. 蛋白質分離 2. 酵素切割 3. 質譜鑑定 4. 比對分析



\*Mass Spectrometry

\*Electrospray ionization

\*Matrix Assisted Laser Desorption and Ionisation

## Identify protein by MS

Mass spectrometry (MS) provides the most accurate mass measurements of large biomolecules.

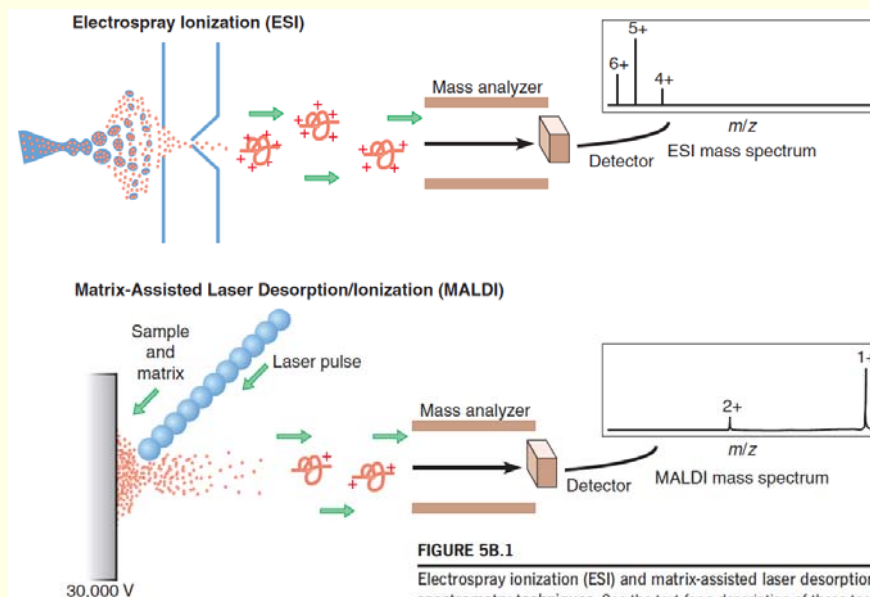
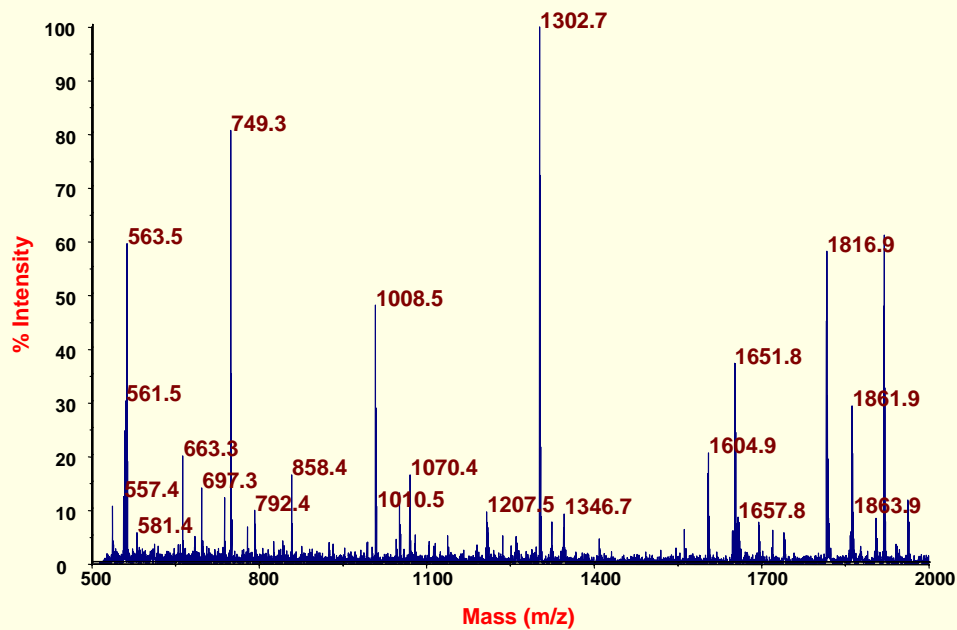


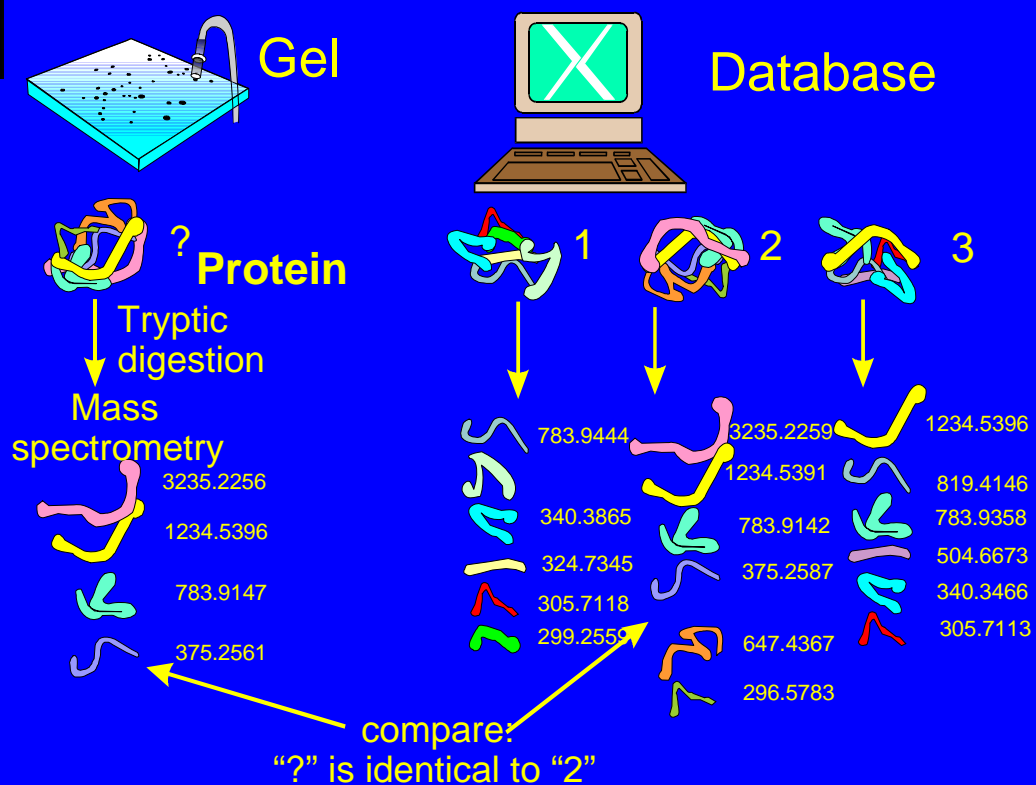
FIGURE 5B.1

Electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI) mass spectrometry techniques. See the text for a description of these techniques.

## MALDI-TOF mass spectrometry profile of tryptic digests of protein 1

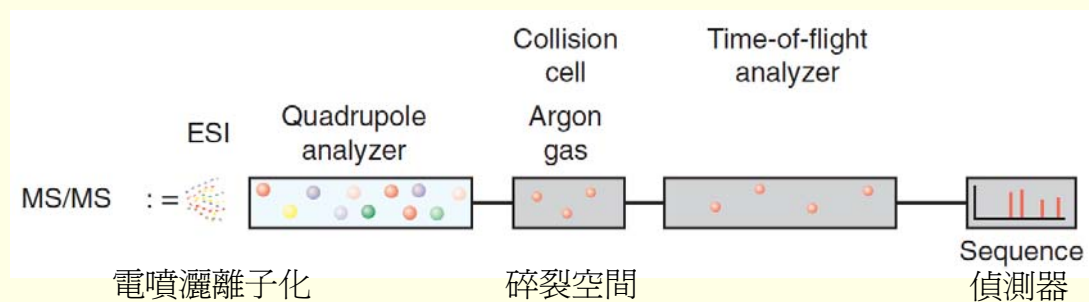


## Peptide Mass Fingerprinting with MS



## Mass spectrometry 質譜分析

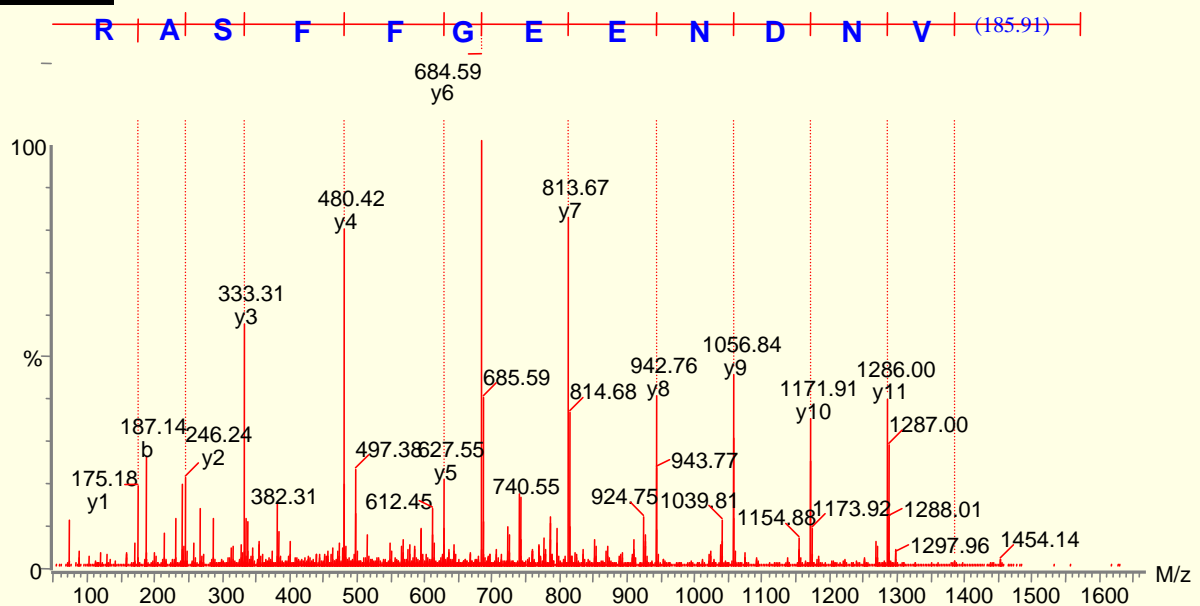
- Mass spectrometry (tandem mass spectrometry, **MS-MS**) can also be used to obtain peptide-sequence information. 串聯式質譜儀分析
- The mass spectrometer must have a collision cell and two mass analyzers: 碎裂
  - quadrupole analyzer 四極桿
  - time-of-flight (TOF) analyzer 飛行時間分析器



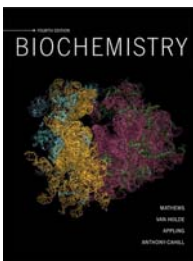
Copyright © 2013 Pearson Canada Inc.

5 - 57

## MS/MS Spectrum of Peptide

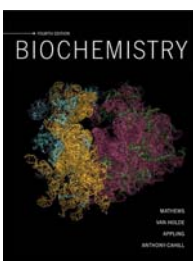


Amino acid sequence read as differences in weight between ions in series.



*Are you understand about*

- Structures of the  $\alpha$ -Amino Acids
- Properties of the Amino Acid Side Chains
- Peptides and the Peptide Bond
- Proteins: Polypeptides of Defined Sequence
- From Gene to Protein
- Protein Sequence Homology
- Protein Expression and Purification
- Mass, Sequence, and Amino Acid Analyses of Purified Proteins



Thank you