

主題：CH5 後段+CH6	
教師：游佳融	日期：2014/03/03
撰稿組：林佩萱 關好珊 何佩萱 孫慶語	審稿組：黃尹珊 李佳樺 王韻婷

## CH5.Introduction to Protein: the Primary Level of Protein Structure

### 蛋白質：具有固定序列的多肽鏈

1. 每個蛋白質都具有一獨特、定義數量以及定義次序的胺基酸鏈。而此序列在核酸指的即是蛋白質的初級結構(primary structure)。

2. 右圖為抹香鯨以及人類的肌紅蛋白(myoglobin)初級結構的胺基酸序列圖，可以發現有高達 84% 的序列相似度(sequence similarity)，這是因為它們具有相同功能的緣故。然而隨著物種演化，蛋白質也會以改變胺基酸序列的形式進化(evolve)，進化的方式有兩種：以保存其固有化學性質以及/或者支鏈大小的進化方式稱作保守取代(conservative changes)，而若是具有較大性質上變動的進化方式則稱作非保守取代

(nonconservative changes)。注意即使是再微小的序列改變都有可能大大影響蛋白質的功能性；排除突變的因素，每一隻抹香鯨的肌紅蛋白胺基酸序列應為相同。

圖示：藍色代表一致的胺基酸、綠色代表保守取代的胺基酸、黃色代表非保守取代的胺基酸。

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

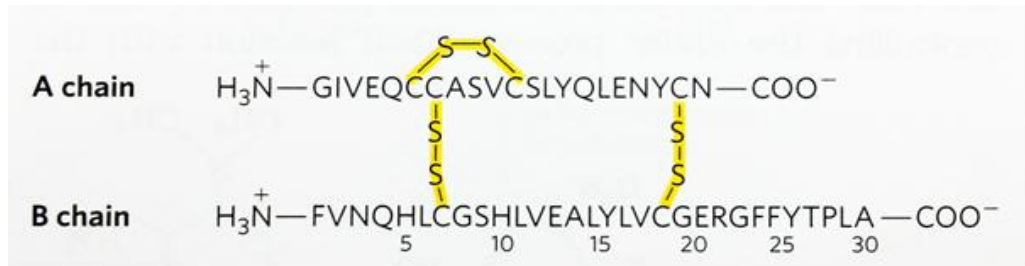
Number	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135
Human	G	D	F	G	A	D	A	Q	G	A	M	N	K	A	L
Whale	G	N	F	G	A	D	A	Q	G	A	M	N	K	A	L

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

3. 部分蛋白質含有兩條或兩條以上以共價鍵/非共價鍵連接的多肽鏈組成。

(1) 世界上第一個序列被解出來多肽鏈的是胰島素(insulin)。胰島素具有 A、B 兩個多肽鏈，黃色的即為雙硫鍵(disulfide bonds)，如下圖所示。由此可知雙硫鍵可以在分子間或在跟分子內形成。



(2) 關於蛋白質的分子量大小，最小已經有定義約 10Kda，但蛋白質的分子量變異性很高，例如角質(Titin)為目前人類最大的蛋白質，雖然它只有一條多肽鏈但是它的分子量卻是其他蛋白質的三百倍，常出現在角質、皮膚、髮削，因此常造成鑑定上的誤差。

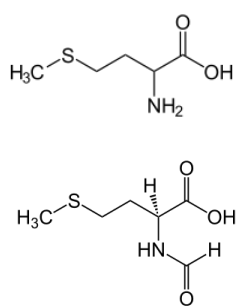
(3) 有功能的蛋白質，可能由兩條以上的多肽鏈形成，涉及到蛋白質的四級結構。例如血紅蛋白(hemoglobin)就由 4 條肽鏈(peptide)組成，一般稱一條肽鏈為一個次單元(subunit)，次單元具有調節功能，一條肽鏈可折疊成三級構造。若一個完整具有功能的蛋白質是由兩條或以上肽鏈所組成，則稱之具有四級結構。

(4) 註：HbA(hemoglobin adult)：成人的紅血球，含有 2 個  $\alpha$  及 2 個  $\beta$  次單元；HbF(hemoglobin fetal)：幼兒的紅血球，含有 2 個  $\alpha$  及 2 個  $\gamma$  次單元。

## 從基因到蛋白質

1. 勿忘中心法則(Central Dogma)：DNA 轉錄(transcribe)成 RNA，而 RNA 轉譯(translate)成蛋白質。因此 DNA 的序列決定其產生的蛋白質。

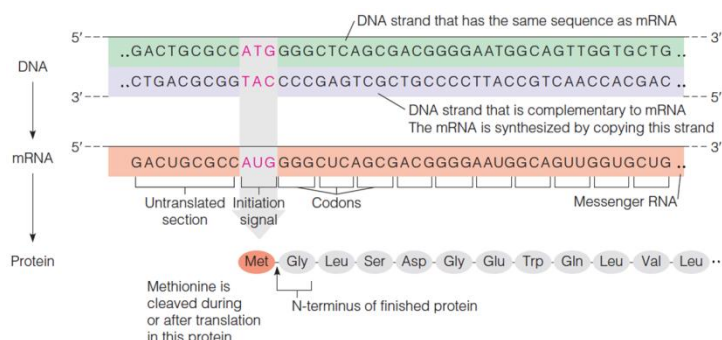
2. 右圖為基因密碼子圖表，可以得知 UAA、UAG 和 UGA 三個密碼子通常不會產生任何胺基酸(極少數的狀況，UGA 會轉譯出硒半胱胺酸<selenocysteine>，UAG 會轉譯出吡咯賴氨酸<pyrrolysine>)，而是作為終止密碼子，在 C



端(C-terminus)停止轉譯的進行。而 AUG 則是作為起始密碼子，在真核生物中轉譯出甲硫胺酸(methionine，縮寫 Met，左上圖)，在原核生物以及真核生物器官中則是轉譯出 N-甲酰甲硫氨酸(N-Formylmethionine，縮寫 fMet，左下圖)。理論上來說此兩種胺基酸應作為 N 端(N-terminus)的開頭，不過它們常

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

常會在修飾的過程中遭到切除。下圖則是描述人類血紅蛋白在 N 端的 DNA、mRNA 以及多肽鏈序列三者的關係，可以發現 N 端的甲硫胺酸被移除了。



## 蛋白質序列同源度

1. 一種稱作 Basic Local Alignment Search Tool(簡稱 BLAST)的工具被廣泛使用在研究蛋白質序列上。簡單來說，當發現一串新的基因序列，為了猜測其功能，我們會使用 BLAST 的蛋白質基因序列資料庫進行比對，來得到我們想要的結果。序列的相似度與其功能的相似度是相關的。根據規則，當序列有至少 25% 的相似度，它們可能就有相似的功能存在。下圖是 BLAST 使用的實例，比較人類肌紅蛋白(human myoglobin)以及人類  $\alpha$  球蛋白(human  $\alpha$  globin)之間的相似度。

圖示：藍色代表一致的胺基酸、綠色代表相似的胺基酸、紅色代表間隔。

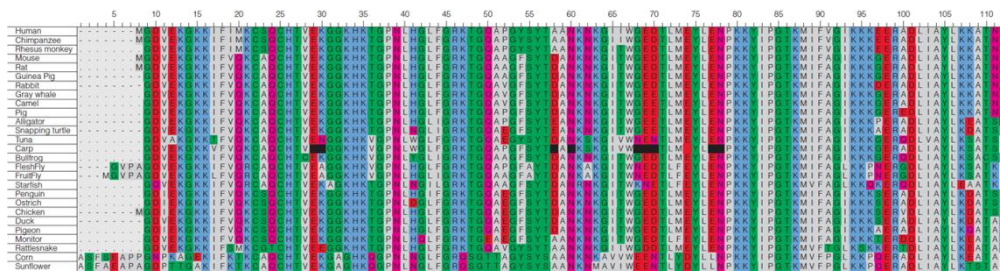
Score = 30.8 bits (68), Expect = 6e-06, Method: Compositional matrix adjust.  
Identities = 32/133 (25%), Positives = 48/133 (37%), Gaps = 40/133 (30%)

Human Mb	2	LSDGEWQLVLNVWGKVEADIPGHGQEVLI RLFKGHPETLEKFDKFKHLKSEDEMKASEDL	61
Human $\alpha$	2	LSPADKTNVKAAWGKVGAAHAGEYGAELERMFLSFPTTKTYFPHF	46
Human Mb	62	KKHGATVLTALGG ILKKKGHHEAEIKPLAQSHATKHKI-PVKYLEFISECI IQVLQSKHP	120
Human $\alpha$	47	----- ALSALSDI----- HAHKLRVDPVNF-KLLSHCLLVTLAAHL P	82
Human Mb	121	GDFGADAQGAMNK	133
Human $\alpha$	83	AEFTPAVHASLDDK	95

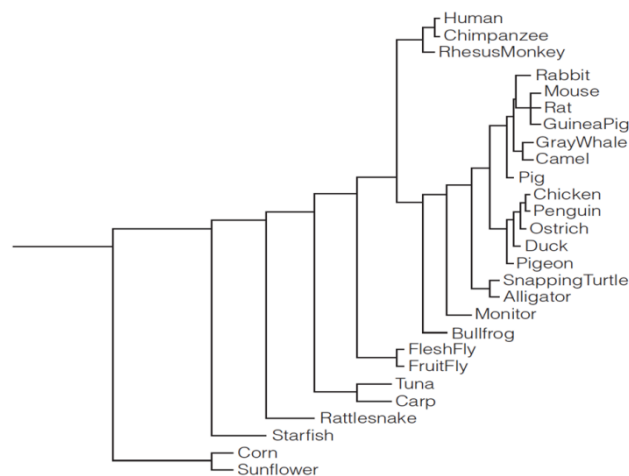
2. 同源與相似的比較：若某兩物種是基於相同演化祖先，而擁有相關的蛋白質序列的話，那麼它們是同源的(homologous)；若某兩物種並不是基於相同祖先，但擁有相似的蛋白質序列，那麼它們是相似的(similar)。因此，序列相似度(sequence similarity)是基於序列的相似程度，而序列同源度(sequence homology)是基於兩個序列在演化上的相關程度。

3. 序列一致度(sequence identity)是指胺基酸序列中完全相同的部分；序列相似度(sequence similarity)則是基於支鏈的化學性質，如疏水性(hydrophobicity)、極性(polarity)、帶電(charge)等等，進行分類而得。

4. 下圖為 27 不同物種的細胞色素 c(cytochrome c)的序列圖，疏水性胺基酸標示為灰色、普通胺基酸標示為藍色、酸性胺基酸標示為紅色、極性不帶電的胺基酸標示為綠色(除了 Asn 與 Gln 是標示為洋紅色以外)



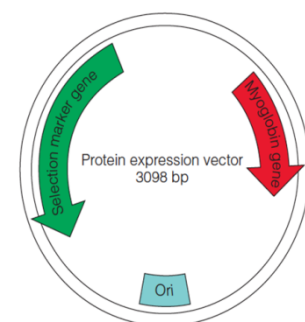
5. 下圖為接續上頁末圖、根據胺基酸序列而判斷出在演化上的差異樹狀圖。可以從此圖了解到蛋白質序列同源度代表著在演化上的密切相關性。



6. 胺基酸保存(amino acid conservation)只有在當一個胺基酸皆在同源蛋白質中的一特定位置被發現時，該胺基酸就稱為「保存的(conserved)」。

## 蛋白質表現與純化

1. 重組蛋白質表現技術(Recombinant Protein Expression Technology)：將所需的蛋白質的基因注入載體，利用 E. coli 攜帶質體(plasmid)作為表現載體(expression vector)，可以大量複製所需之蛋白質。右圖為其形式，右方紅色箭頭代表重組的蛋白質基因(recombinant protein of interest)，而左方綠色箭頭代表選擇的標示基因(selection marker gene)。



2. 純化過程(Purification Process)：層析法(column chromatography)可用於純化蛋白質溶液，主要方法有以下三種：

(1) By charge：使用離子交換層析法(ion-exchange chromatography)，左下圖由不同蛋白質和固定相(stationary phase)間吸引力的不同而分離。已知



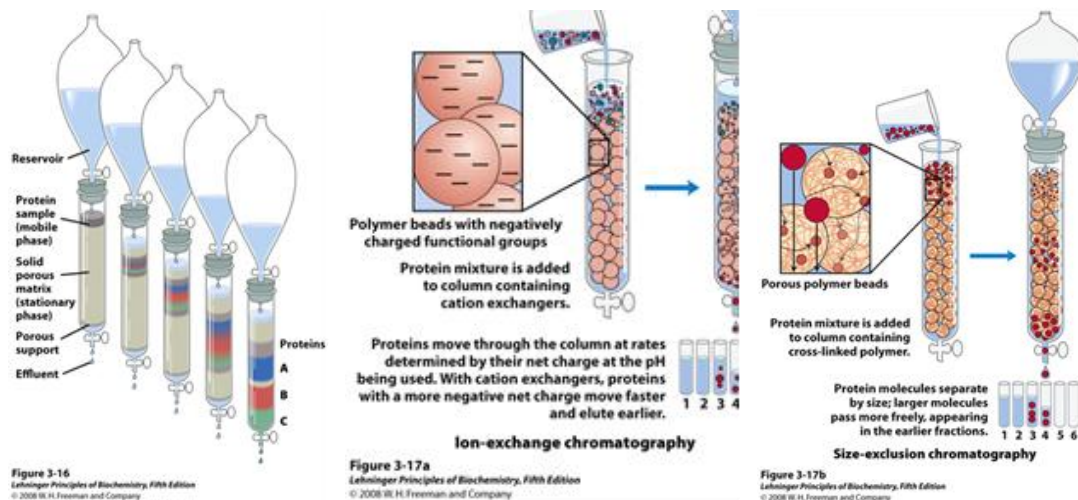
蛋白質的 PI 值，再選擇適合的 buffer(需加減等電點一個 pH 值，因在 PI 值時溶解度最低)。

(2) By size：樹脂過濾層析法(gel filtration chromatography)，下圖

分子量較大的蛋白質會先分離出來，因為其不會經過樹脂中的孔洞而從旁邊經過，而分子量小的物質會經過樹脂中較細小的孔徑而較分子量大的慢沖提出來。

(3) By affinity：右下圖

利用配位體(ligand，e.g. DNA binding protein、抗體)與欲分離的蛋白質間的親和度來純化，純化效益最高(可達 1000 倍)，因配位體和欲分離物有專一性。e.g. 管內填充物(圖中圓球)球上結合特殊序列的 DNA，而目標蛋白質可以辨識並與特殊序列 DNA 相吸，而達分離效果。



## 純化蛋白質的質量、序列、與胺基酸分析

### 1. 蛋白質分析—SDS-PAGE(SDS – sodium dodecyl sulfate – a detergent)

(1) Standard(markers)：具有固定且已知的蛋白質所組成

(2) 功能：可知蛋白質的分子量與純度

(3) 步驟：使蛋白質結構及電性解除

加熱約 95°C、5~10min 使蛋白質展開，解開二、三級結構

(4) SDS 是一種介面活性劑(清潔劑)，SDS 帶負電，使所有的 protein 包裹成帶負電(統一電性)

(5) 低分子量的蛋白質跑較快，較起始點較遠

(6) 可用來解讀蛋白質的純度，以及含量。由圖可知，右邊四個物質相較於左邊較為單純(純度高)。

### 2. 蛋白質分析—Mass Spectrometry

如下圖所示，ESI(electrospray ionization)技術是將蛋白質溶液經過加速，在抵

達分析器之前絕大部分的溶液已經蒸發，殘存下來的蛋白質分子就能經儀器分析繪製出 ESI-MS mass spectrum；MALDI(matrix-assisted laser desorption/ionization)技術則是將蛋白質嵌在可以吸收 UV 光線的 matrix 上，當雷射光打到 matrix 上時，matrix 會蒸發並釋出蛋白質氣體分子，再透過儀器分析繪製出 MALDI-MS spectrum。

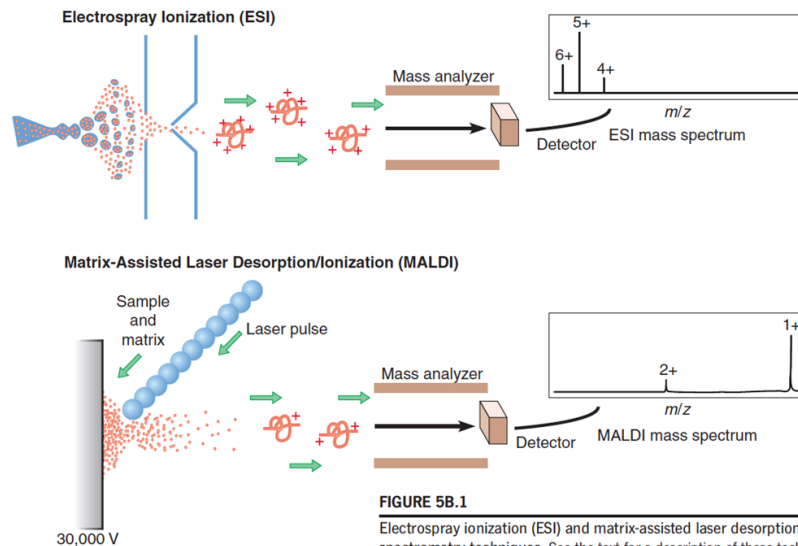


FIGURE 5B.1

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

## CH6. The Three-Dimensional Structure of Protein

### 一、蛋白質的四級結構：

一級結構(primary structure) – 胺基酸序列

二級結構(secondary structure) – 一級結構局部摺疊為重複單元

三級結構(tertiary structure) – 整條胺基酸序列摺疊成單體蛋白質(monomeric protein)或蛋白質次單位(subunit)

四級結構(quaternary structure) - subunit association 次單位連結

下圖：一級：AA. (amino acid)sequence；二級： $\alpha$ -helix,  $\beta$ -sheet, turns and loops

三級：一整條胺基酸序列之空間配置；四級：多個 subunit

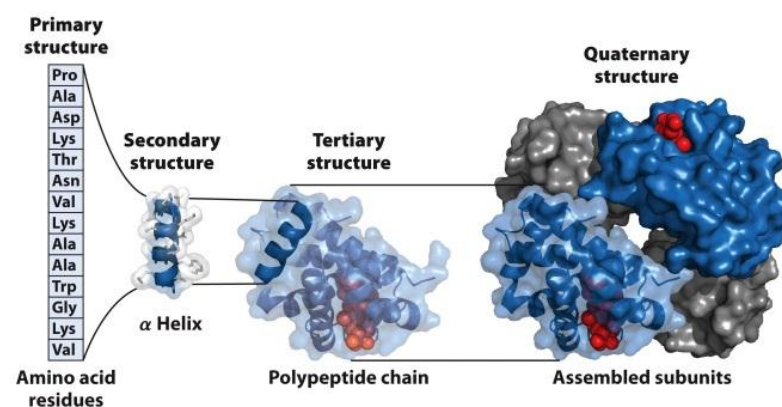
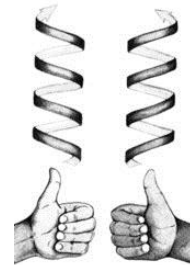


Figure 3-23  
Lehninger Principles of Biochemistry, Sixth Edition  
© 2013 W. H. Freeman and Company

## 二、 Secondary Structure ( $\alpha$ -helix, $\beta$ -sheet, turns and loops)

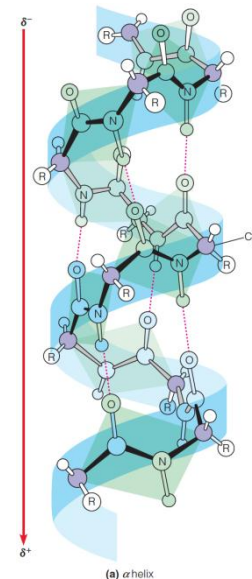
鮑林整理出多肽鍊摺疊結構準則：

1. 鍵長與鍵角盡可能保持與未摺疊多肽鍊相似
2. 原子間距離不可小於二者之凡得瓦半徑(van der waals radius)
3. amide group 要維持一平面且為trans configuration(fig.5.14(b))，僅與  $C\alpha$  連接的鍵結可以旋轉。
4. 使產生最多 amide proton 和 carbonyl oxygen 之 H-bond.



Regular ways to fold the polypeptide chain

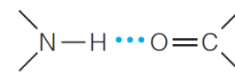
1. Peptide bond 無法旋轉(共振使其可視為平面，see p.148 fig5.14)  
→ polypeptide 由一系列平面組成，而僅與  $C\alpha$  連接的鍵結可以旋轉。
3.  $\Phi$ (phi): angle around the  $\alpha$ -carbon—amide nitrogen bond  
 $\Psi$ (psi): angle around the  $\alpha$ -carbon—carbonyl carbon bond  
由  $C\alpha$  看去，順時針旋轉為正。
4. 在一個完全伸展的多肽鍊中，  
 $\phi$  和  $\psi$  皆為 $+180^\circ$ (如右圖)。



最常見二級結構： $\alpha$ -helices &  $\beta$ -sheets

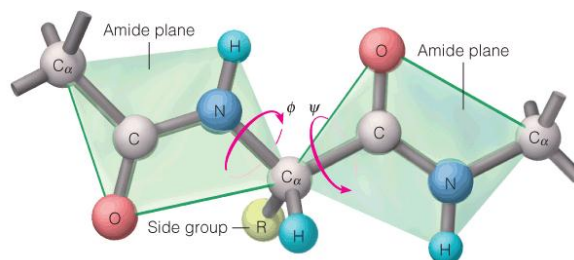
### Amphiphilic helices and sheets

因相同極性的側鏈易團聚在一塊， $\alpha$ -helices &  $\beta$ -sheets 常形成兩性(amphiphilic)分子：同時具疏水面與親水面。此構造利於兩個至多個二級結構在水溶液中以疏水面相連結，親水面向外形成穩定結構。



### $\alpha$ -helix

1. 所有 R group 向外突出，而一個 helix 轉一圈為 3.6 個 residues(殘基)。
2. 蛋白質中  $\alpha$ -helix 為 right-handed。(如右示意)
3. H bond 大致平行於 helix axis，使得 N-H 與 C=O 大致平行於 helix axis，產生(N-terminus→C-terminus)為 $(\delta+ \rightarrow \delta-)$ 的 helical dipole moment(又稱 macrodipole)
4.  $i$  殘基上之 carbonyl oxygen 與向



C 端移動四個殘基( $i + 4$  殘基)上之 **amido proton** 形成 hydrogen bond

→ A hydrogen-bonded loop of **13 atoms** is formed.

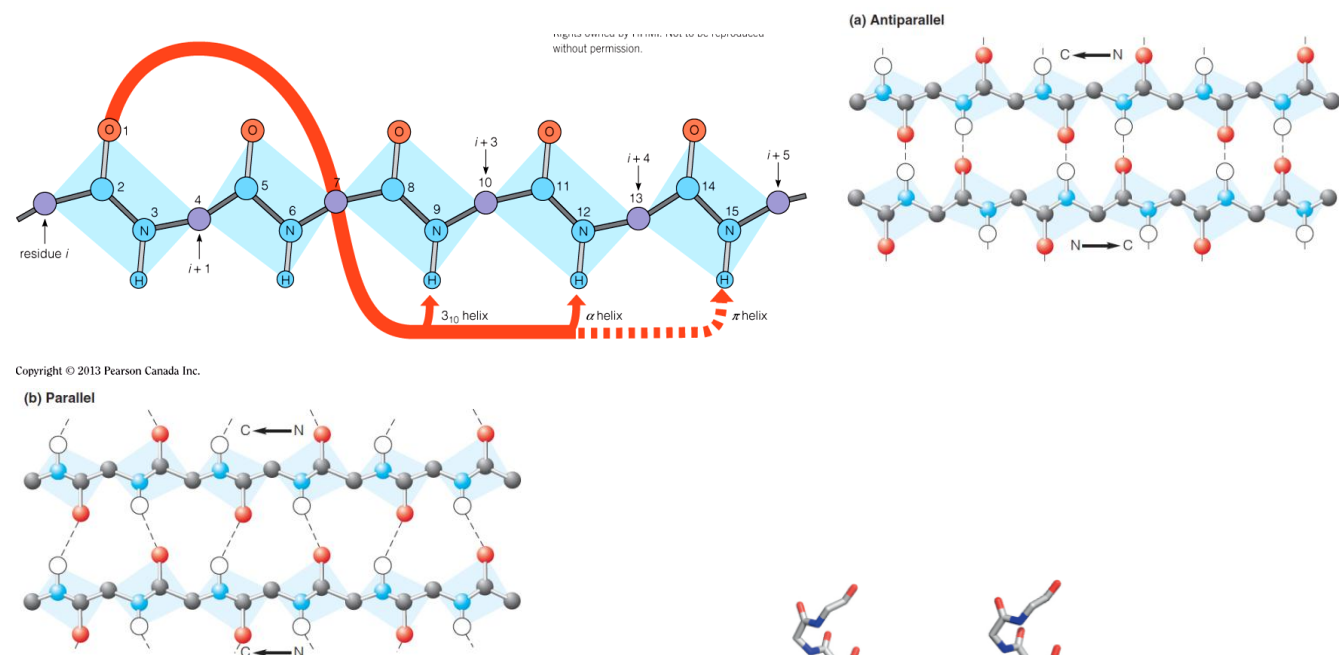
→  $\alpha$ -helix 又可被稱為 3.6<sub>13</sub> helix.

※3<sub>10</sub> helix：亦是組成蛋白質之二級結構之一，但不如  $\alpha$ -helix 常見。轉一圈為 3 個殘基，而一個 hydrogen-bonded loop 則含有 10 個原子，故名 “3<sub>10</sub>” helix。

※  $\pi$  helix：由於突變使  $\alpha$ -helix 結構插入一 amino acid，使得一般  $i$  與  $i+4$  之間之 hydrogen bond 變為  $i$  與  $i+5$ ，形成  $\pi$ -bulge(腫脹、凸塊，又稱 “ $\alpha$ -bulge” 或 “ $\alpha$ -aneurism”)一個序列中通常只有一個 insertion。

$\pi$  helix 並非重複性結構，但可作為追蹤蛋白質功能演化的 marker。

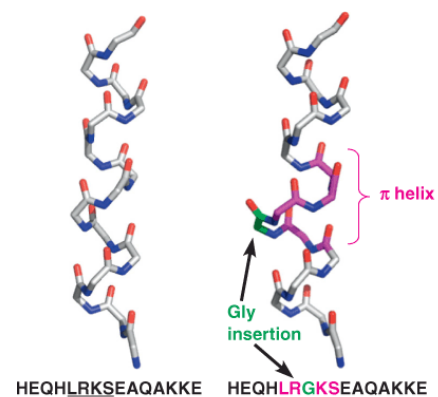
## $\beta$ -sheet



1.  $\beta$  strand：每一個 residue 與前後 residue 角度旋轉 180 度(亦即 2 residues/turn)，許多  $\beta$  strand 並排並產生分子間氫鍵。

2. 依 N(amino)到 C(carboxyl)方向，分為 antiparallel 和 parallel。

相異處：(1)重複片段長度：parallel=6.5Å；  
antiparallel=7 Å (節自 M107 共筆)



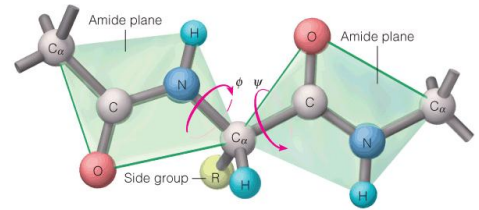
Copyright © 2013 Pearson Canada Inc.



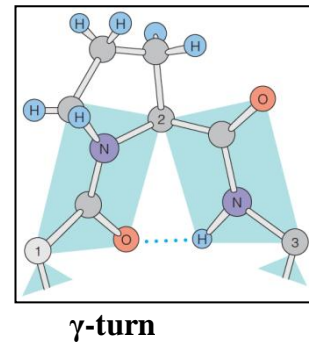
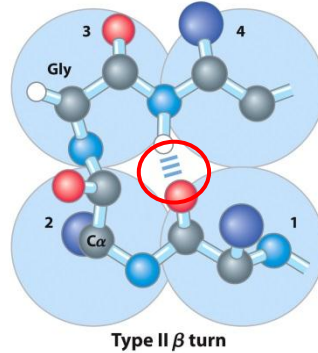
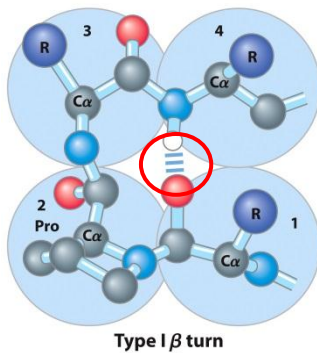
(2) 氫鍵方向：parallel 整齊(linear)；antiparallel 歪曲(distorted)

**Polypeptide II helix**(p.181 下半以及 fig.6.8，較不重要，看看就好)

1. 並不符合鮑林的準則：不含穩定結構之氫鍵
2. Left-handed helix
3. 大約 1/3 的 residue 為 proline，因此也常被稱為“polyproline II helix”  
而 glycine 也是常見的 residue。



## Turn 轉折

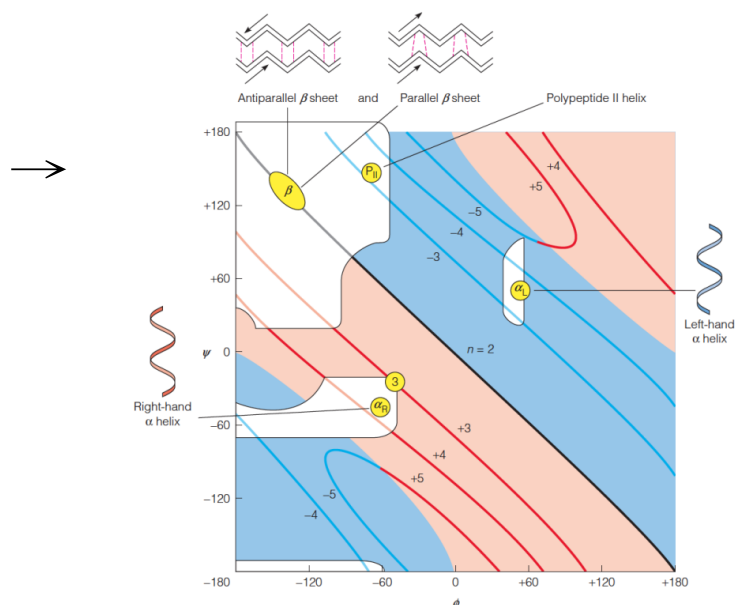


**β-turns**：β-sheet 中片段改變方向時即需 β-turn 連接。Antiparallel β-sheet 中尤其常見。

1. **Type I**：proline 常於 2 號位置
2. **Type II**：glycine 常於 3 號位置

**γ-turns**：proline 常於 3 號位置

## 三、Ramachandran plot(拉氏圖)



※一種使蛋白質結構中，主鏈胺基酸殘基的二面角  $\psi$  和  $\phi$  可視化的方法。對主鏈胺基酸構型系統化描述：橫坐標為  $\phi$  值，縱坐標為  $\psi$  值。

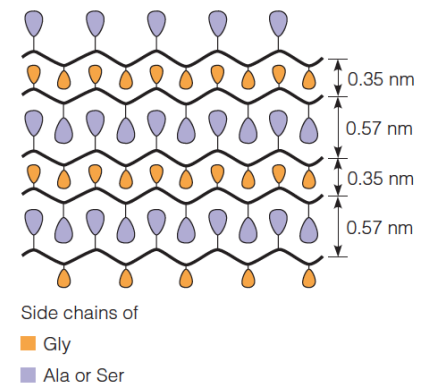
※每一種不同的二級結構都有不同的  $\psi$  和  $\phi$  值組合，詳見 p.184 table 6.2.

右圖：有標碼之曲線代表 **residue per turn** 的數值，此為 **poly-L-alanine** 之拉氏圖，若使用 **R group** 較 **bulky** 之胺基酸，圖中白色區塊(sterically allowed conformation)面積會更小。

→glycine 具有最多可能 ( $\phi, \psi$ )，proline 具有最少可能 ( $\phi, \psi$ )

#### 四、Fibrous proteins 纖維狀蛋白

纖維狀蛋白具延展的、纖維狀結構，在動物體內扮演維持結構的角色(例如：表皮、結締組織、毛髮、絲)。每一種特定纖維蛋白都會有 3-4 種含量較多的胺基酸，用以穩定其延展的二級結構。(見 p.185 下方 table 6.3)



##### 1. $\alpha$ -keratin :

- 角質以及中間絲(intermediate filament protein)之組成成分，構成動物毛髮、指甲以及大部分表皮，也是細胞核、細胞質、細胞表面維持結構的成分。
- 以毛髮為例，其組成由 **Monomer**(約 300residues 含  $\alpha$ -helical domain 和 Globular domain 之單體，right-handed  $\alpha$ -helix)→**Dimer**(因為  $\alpha$ -keratin 每 3-4 個 residue 便有一個非極性 R group 之 residue，而  $\alpha$ -helix 是 3.6residue/turn，故每一 monomer 都有一 hydrophobic surface area(雖然不完全吻合 3.6residue/turn)→在水溶性環境中 hydrophobic surface 會相互連結→兩個 Monomer 得以 bonded by hydrophobic interactions→left-handed coiled-coil)→**Protofilament**→**Protofibril**(intermediate filament 中會產生後兩者結構)(見課本 p.186 fig.6-14)
- 可以藉由 cysteine 架 disulfide cross-links 使  $\alpha$ -keratin 結構硬化。指甲中便含有許多 disulfide cross-links，而頭髮則較少(燙髮：打斷 disulfide cross-links→改變形狀→reoxidation)

##### 2. $\beta$ -keratin：含有較多的 $\beta$ -sheet，大多存在於鳥類(羽毛)與爬蟲類(鱗片)

##### 3. Fibroin(纖維蛋白)：存在於蠶絲與蜘蛛絲中，以蠶絲為例(如右下圖)

- Antiparallel  $\beta$ -sheet，在其組成中大部分為 **Glycine**，其次為 **Alanine** 與 **Serine**(不同種類的 Fibroin 有不同的組成，但 glycine 總是最多)
- 因為共價鍵結幾乎都伸展到最大長度→較強韌而不具延展性
- 因為鏈與鏈之間是用凡得瓦力連結→flexible
- 並非所有 fibroin 結構都是  $\beta$ -sheet，有時 **bulky amino acid** 會阻斷  $\beta$ -sheet 片段(eg. valine, tyrosine)，這樣的成分會增加 **fibroin** 的延展性。

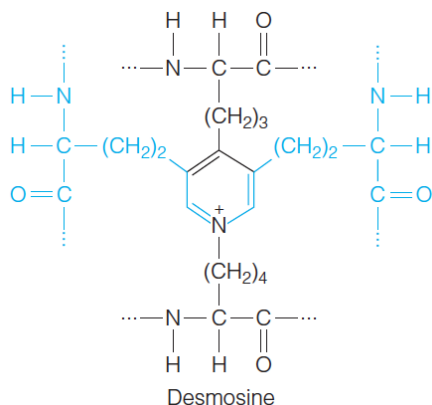
##### 4. Collagen(膠原蛋白)：在動物體中最常見的蛋白質，構成骨質(bone matrix)、

肌腱(tendon)、以及皮膚構造。

- 以 tropocollagen 為基本單位構成(由三個多肽鏈構成一 right-handed triple helix，每一個多肽鏈約 1000 residues，left-handed，3.3 residues/turn，每三個 residue 就會出現一個 glycine 且位於靠近 triple helix 中心的位置→基本 motif 為 **Gly-X-Y**，其中 X 通常是 **proline**，而 Y 通常是 **proline 或 hydroxyproline**，不過 X、Y 位 residue 可以是其他種類胺基酸，越「高等」有機體(higher organism)會有越多種的變異)
- Hydroxyproline 及 hydroxylysine(含量較少)→其-OH 可產生氫鍵穩定結構
- 缺乏 vitamin C(L-ascorbic acid)→催化 proline 和 lysine hydroxylation 之酵素無法運作→**scurvy(壞血病)**：**H-bond 減少，skin, gums, blood vessels weaken.**
- 相鄰 Tropocollagen 約相距 64nm→具條紋的外觀，也因 overlap 以及因為與 lysine side chain 反應相關的 cross-link(p.189 反應式)→強韌
- Collagen synthesis(如右圖)**

## 5. Elastin(彈性蛋白)：

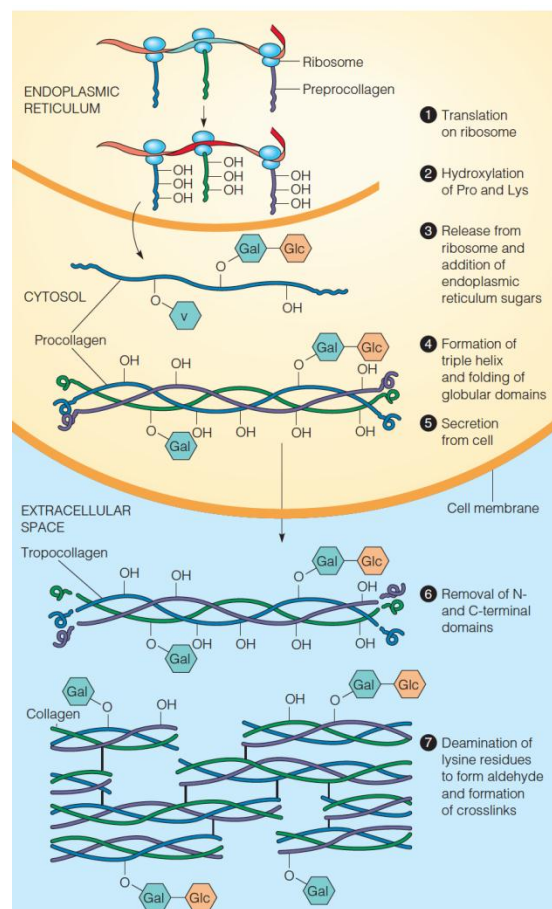
- 含許多 glycine、alanine 和 valine
- random coil→幾乎沒有所謂二級結構，不過亦會有 lysine cross-links.
- 四個 lysine side chain 可形成 desmosine(鎖鏈素) cross-link.(如下圖)



Protein Secondary Structure (Motif &

domain & fold)

超二級結構(Super-secondary structure)



## 1. Motifs (folds)



數個二級結構蛋白質所組成，可能為單純  $\alpha$ 、 $\beta$ (ex:  $\beta$ - $\alpha$ - $\beta$  loop)，或是以不同比例混和。物種間有可能有相似的序列，例如植物和動物的 protein kinase c 有 50 個胺基酸序列很相似，那形成的 folds 也有同樣的功用，所以說若兩者的 motifs 相似度越高，那表示其物種相似性就越高。

用以分類的話，protein kinase c 有四種，為何分成四種(family)也是依據 motifs 些微的差異。

簡單來說，就是相似且重要的氨基酸序列造成相似的 motifs，而這些有相似的 motifs 的蛋白質就會被歸在同一個 family，這使不同的物種間有類似的生理機轉。

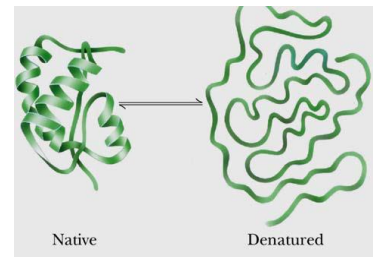
## 2. 同源蛋白

- I. Identity: 相同度。(Identity=100%即兩蛋白質上的胺基酸序列結構完全相同)
- II. Homologous: 具有相同 domain 的蛋白質，即有相同的功能。
- III. Similarity: 相似度，同一屬性的蛋白質分類。(由同族的胺基酸所構成的蛋白質即相似，如: Ser=Thr)

$$\text{Similarity} \geq \text{Identity}$$

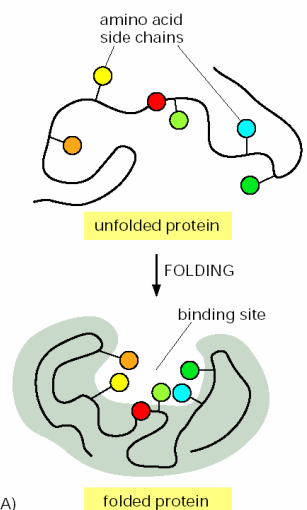
## 3. Protein stability and folding

- I. 蛋白質在特定細胞環境下才能行使功能。當受到高溫、低溫、強酸、強鹼、有機溶劑、chaotropic agent 的刺激後，蛋白質有可能會解構(denaturation)而當這些因子去除後，蛋白質有可能會再回復(refolding)。



- II. Folding 是一個高度協同作用途徑的結果，有三種非共價鍵結會參與摺疊：

- (1) 正負電相吸(例如 glutamic acid 和 lysine)
- (2) 氫鍵
- (3) 凡德瓦力(aromatic、疏水性)



III. ribonuclease A 會被高濃度的 urea 和還原劑作用產生 denature



→還原劑切斷雙硫鍵，形成八個 Cys

→尿素破壞 hydrophobic interaction 使之失去活性

移除尿素和還原劑，則蛋白質有機會摺疊回去用活性測試或使用精密儀器測定三級結構可知道蛋白質有無 refold。

IV. 有些蛋白質在摺疊時要有特殊蛋白質(chaperones 意指僕人)幫助，Chaperone 會與 peptide 結合，提供摺疊所必需的 microenvironment。

#### a. 熱休克蛋白(Hsp70)

救援蛋白，細胞遭受危難的時候會大量分泌，主要出現在 ER。會和未折疊的 peptide 上的疏水性殘基結合，以防止不正確的聚合。

b. GroES 和 GroEL 是一個複合體 chaperonin，長得像圓筒狀，未折疊的蛋白質和未被 GroES 塞住的 GroEL 結合

→ATP 結合在 GroEL 七圓環上的每一個次單元

→ATP 的水解使 14 個 ADP 和 GroES 離開

→7 個 ATP 和 GroES 結合在口袋被塞滿的 GroEL 上

→蛋白質在內部摺疊

→蛋白質離開時已經是完全摺疊或部分折疊

Chaperonin 可重複使用，因為細胞很需要他，又不能耗時一直重做(因需要能量)，若蛋白質經其他蛋白質協助後還是無法摺疊成功，細胞會將之移除或排出(亦可能沉澱)。

(課本上好像沒有)

Protein tertiary structure

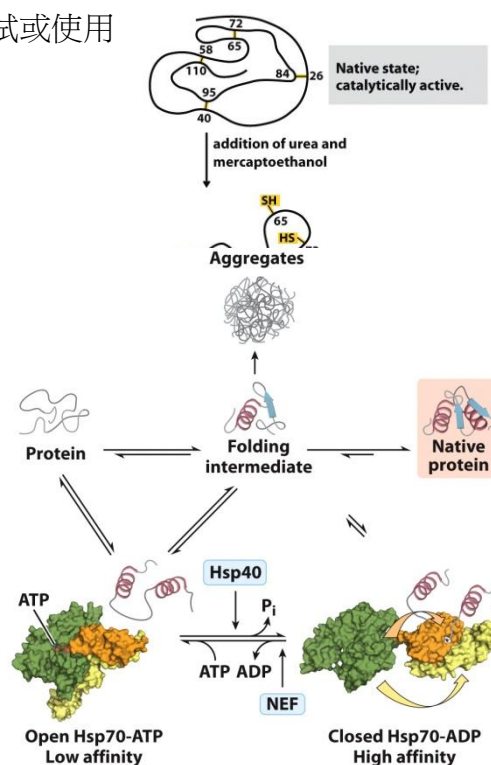


Figure 4-30  
Lehninger Principles of Biochemistry, Sixth Edition  
© 2013 W. H. Freeman and Company

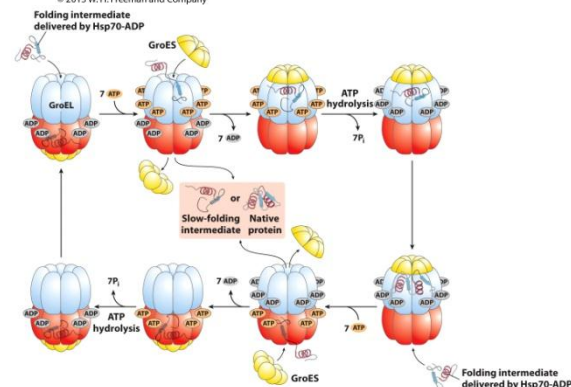


Figure 4-31a  
Lehninger Principles of Biochemistry, Sixth Edition  
© 2013 W. H. Freeman and Company

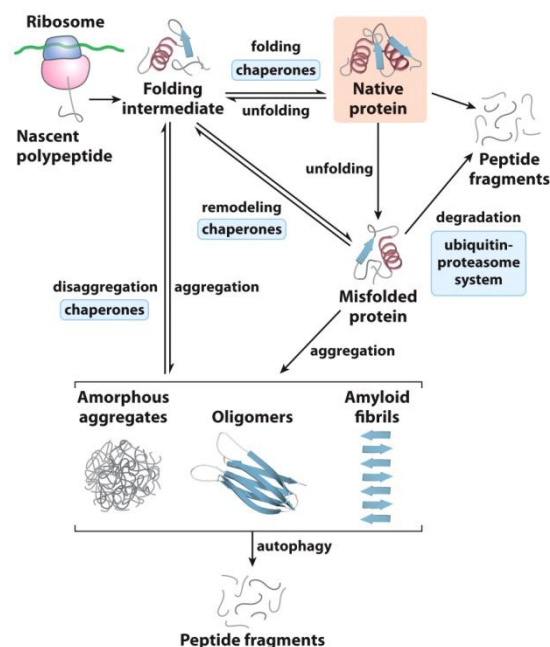


Figure 4-25  
Lehninger Principles of Biochemistry, Sixth Edition  
© 2013 W. H. Freeman and Company

## 一、蛋白質的三級結構

### 1. Fibrous protein(纖維蛋白)

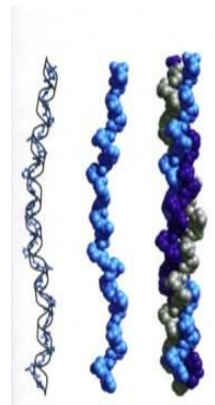
不溶於水，由單一二級結構組成。負責結構部份，提供脊椎動物支持、形狀和外在保護。

### 2. Globular protein(球蛋白)

- I. 佔蛋白質的 80%，常包含多種類型的二級蛋白(如酵素、調控蛋白)，再組合成三級結構。多為外層是 hydrophilic aa. 內層是 hydrophobic aa.(不可套用再穿膜蛋白)
- II. 執行細胞中的多種功能，如合成、運輸、新陳代謝
- III. Many carry **prosthetic groups** (small molecules that may be noncovalently or covalently bonded to the protein (*ex.* heme in myoglobin).)

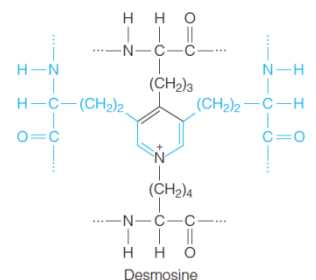
### 3. collagen(膠原蛋白)

- I. 纖維狀，張力高(二集結構以  $\alpha$  helix 為主)
- II. 結締組織：肌腱、軟骨、骨頭、眼角膜
- III. A repetitive motif in the sequence is of the form Gly-X-Y, where X is often proline and Y is proline or hydroxyproline.( 故 Proline 和 4-hydroxyproline 佔了 30%)
- IV. 每條 collagen 為 left-handed，三條 collagen 纏繞形成 right-handed super helical triple helix
- V. 許多 triple-helices 組成 Collagen fibril(膠原纖維)
- VI. 裡面是疏水性，外面的 side chain 是極性的，所以具有水溶性
- VII. 4-hydroxyproline 因有 OH 基，所以不會破壞  $\alpha$ -helix
- VIII. Scurvy(敗血病): Collagen fibers are weakened(less H-bonding between the chains of tropocollagen), caused by failure to hydroxylate prolines and lysines in collagen.



### 4. Elastin 彈性蛋白

- I. Rich in Gly, Ala, Val.
- II. Very flexible and easily extended.
- III. 少有二級結構
- IV. Contains lysine side chains.( Four lysine side chains can be combined to yield a **desmosine** cross-link.)



## 二、 Domain

Domain 是蛋白質三級結構摺疊而成的反應作用區，Domain 比 motifs 來到具有功能的特性，motifs、folds、domain 因其有特異性，可以用來分類蛋白質，有些小分子蛋白只有一個 domain，而有些蛋白質則可能有不只一個 domain。

### 三、Factors determining Secondary and Tertiary Structure.(熱力學)

三級結構蛋白質的形成是 *thermodynamically favorable process* (熱力學趨向低能量的過程)， $\Delta G$  必小於零。

以下幾種情況會使蛋白質 denaturation(變性)，失去功能：

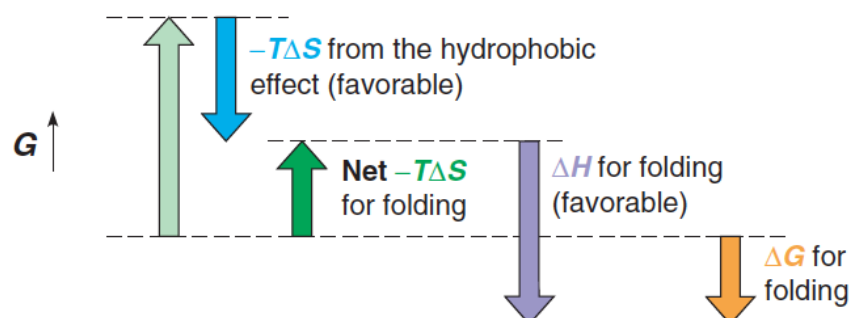
1. 溫度上升
2. 環境酸鹼度劇烈變化
3. 有機溶劑

受以下熱力學因素影響：

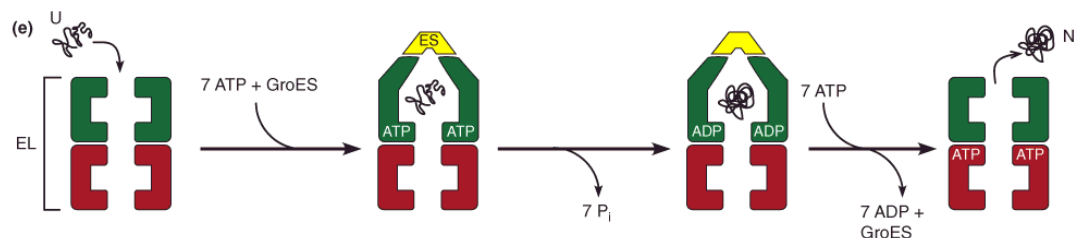
1. Conformational Entropy(從高能量構型到低能量構型)
2. Charge-Charge Interactions(發生在正負電不同的 side chain 上，有時又叫 salt bridges)
3. Internal Hydrogen Bonds
4. van der Waals Interactions
5. the Hydrophobic Effect (雖然原本沒有折疊的 protein 的熵 entropy 較低，但將疏水基團包入蛋白質中之後反而可以形成較穩定但熵較高的情況)

$-T\Delta S$  from decrease in conformational entropy (unfavorable)

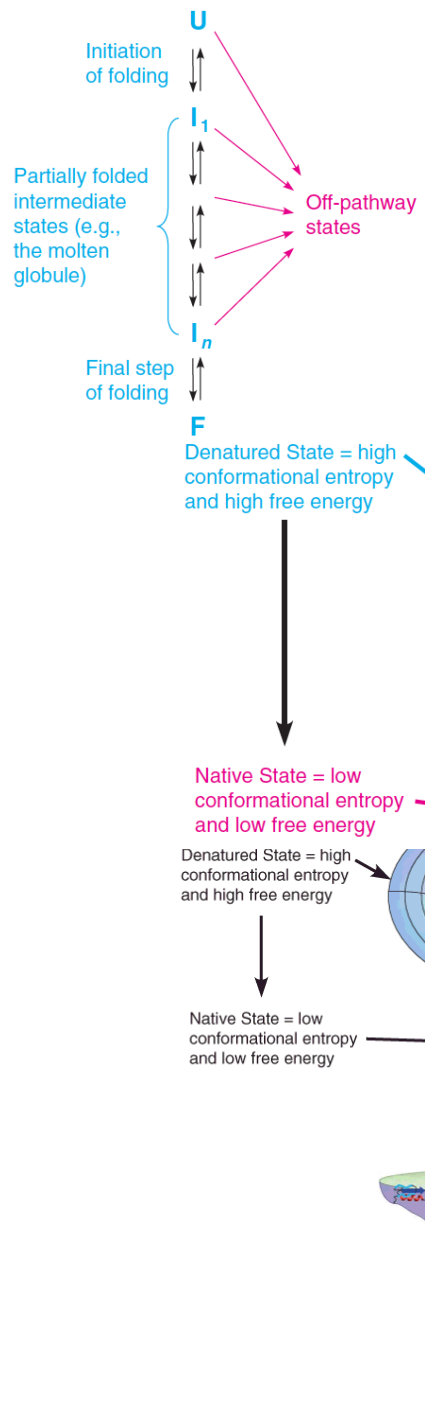
$$\Delta G = \Delta H - T\Delta S$$



6. 雙硫鍵的功能：可以使蛋白質三級結構較不易變性(主要是在高溫下)ex. BPTI protein 有三個雙硫鍵，為最穩定的蛋白質之一
7. **Molecular Chaperones** 分子伴侶：幫助蛋白質摺疊(需要提供能量)，使他不會折疊錯誤或和其他蛋白質聚合的蛋白質結構 ex. GroEL-GroES chaperonin



### 四、Tertiary Structure 的動力學



1. 簡單蛋白質摺疊模型：U 表示尚未摺疊的蛋白質，F 表示摺疊好的有功能蛋白質，I 是(許許多多的)中間產物，基本上蛋白質摺疊時會趨向最低能量，Off-pathway states include aggregates and other non-native states that may be kinetic or thermodynamic “dead-ends”，會使摺疊速度變慢(類似錯誤的嘗試)，此模式並不能完全解釋蛋白質的摺疊(Levinthal's paradox)
2. energy landscape/folding funnel 顯示了各種不同的蛋白質能量和他們之間的互換關係，解釋蛋白質的摺疊時間很短，通常只有幾秒而已
3. the molten globule 是指三級結構相似但二級結構改變的輕微變性蛋白質

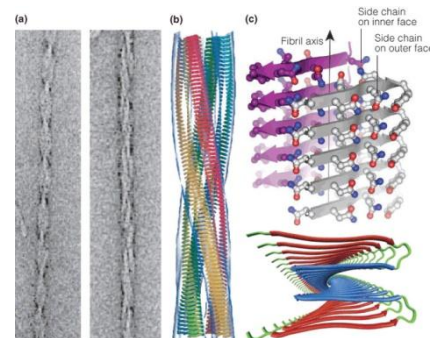
## 五、folding 錯誤造成的 disease

蛋白質的摺疊錯誤會使蛋白質功能無法正常運作而造成疾病，例如 Alzheimer's disease and Parkinson's disease.

(狂牛症的普恩蛋白降至比一般蛋白質更低的能量，故高溫也無法摧毀他)

## 六、Prediction of Secondary and Tertiary Structure.

1. 二級結構的預測：(課本 p.209 table6.8)
2. 三級結構的預測：Schematic of *de novo* structure prediction





using Rosetta:

- I. Assembly of fragments of local secondary structure.
- II. Final low-energy conformation produced by fragment packing.
- III. All-atom model produced after high-resolution refinement.

#### Quaternary Structure of Proteins

三、蛋白質的四級結構：兩個以上的 subunits

1. 血紅蛋白 (adult) :  $2\alpha 2\beta$
2. 血紅蛋白 (fetus) :  $2\alpha 2\gamma$
3.  $\beta$  與  $\gamma$  之間只有一個 amino acid 不同
4.  $\gamma$  攜帶氧氣的能力大於  $\beta$
5. 一個血紅蛋白可攜帶 4 個分子的氧
6. 一個肌紅蛋白只有一條 peptide，所以只攜帶一分子的氧

#### 蛋白質的鍵結力量(國考愛考)

- 甲、一級：肽鍵
- 乙、二級：氫鍵
- 丙、三級：支鏈的交互作用
- 丁、四級：非共價鍵的交互作用

Spectroscopic methods for studying macromolecular conformation in solution

一、X 光繞射(X-ray crystallography or X-ray diffraction)

- 1、X 光繞射是用來鑑定球狀蛋白或生物聚合物三級結構的良好工具。
- 2、使用限制：分子必須事先結晶化(crystallized)。
- 3、鑑定步驟：
  - (1)純化蛋白質
  - (2)將其結晶化
  - (3)蒐集繞射數據
  - (4)計算並獲得晶體中電子密度的分布情況
  - (5)從中分析獲得原子的位置信息，鑑定出蛋白質的三級結構
- 4、優點：
  - (1)無分子大小(size)限制
  - (2)容易鑑定
- 5、缺點：
  - (1)不易應用於膜蛋白
  - (2)無法得知氫原子訊息
  - (3)無法得知環境改變後，蛋白質如何改變其構型(conformational changes)

二、核磁共振(Nuclear Magnetic Resonance, NMR)

1、原理：由 NMR spectroscopy 紀錄待測物各處原子的 chemical shift(單位 ppm)，再與已知數據做對照，藉以得知結構。(不懂的同學請自己翻閱上學期有機化學課本\^O^/)

2、步驟：(1)純化蛋白質

(2)將其溶解

(3)蒐集 NMR 數據

(4)分析數據

(5)計算得知結構

3、優點：(1)不需要將蛋白質結晶化

(2)可得知氫原子訊息

(3)可用來測量蛋白質的 dynamic motions

4、缺點：(1)不適用於無法溶解的(insoluble)蛋白質

(2)小分子蛋白質應用效果較好

### 三、圓偏光二色光譜(Circular Dichroism, CD)

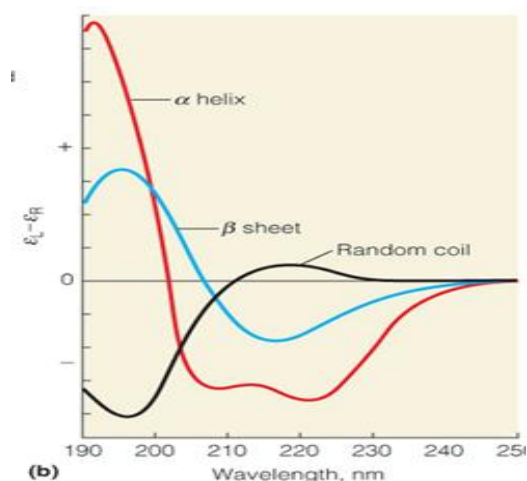
1、CD 是測量蛋白質二級結構的良好工具。

2、原理：

(1)圓偏振光(circularly polarized light)：光波的電場方向稱為「偏振方向(Polarized direction)」。如果一道光波前進的時候，其電場變換方向一直在改變，稱為「非偏振光(Non-polarized light)」；如果一道光波前進的時候，其電場變換方向固定不變，則稱為「偏振光(Polarized light)」。而這個偏振方向又可分為順時針旋轉或逆時針旋轉，分別稱為 right circularly polarized light 與 left circularly polarized light。

(2)大部分分子並不對稱(asymmetric)，會對右圓偏光與左圓偏光有不同的吸收度，而這個差異即被定義為 circular dichroism。

其公式為 $\Delta A = (A_L - A_D)/A$ ，其中  $A_L$  是對左圓偏光的吸收度， $A_D$  是對右圓偏光的吸收度， $A$  是對非偏振光的吸收度。故 $\Delta A$  可為正值或負值，且可做為判定為哪種二級結構的標竿，可說 circular dichroism 也是分子的特性之一，如下圖所示。



- 3、優點：容易觀察蛋白質在溶液中的構型變化(conformational changes)
- 4、缺點：無法剖析原子層級的結構