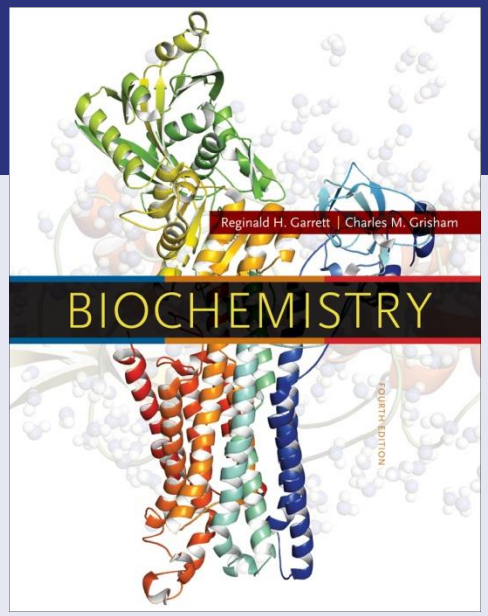
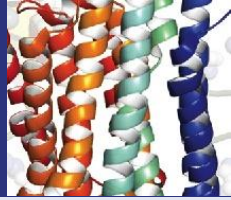


Reginald H. Garrett
Charles M. Grisham

Chapter 6
Proteins: Secondary, Tertiary, and Quaternary Structure
蛋白質二級、三級、四級結構

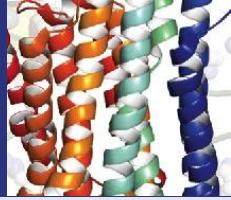
Essential Question



- How do the forces of chemical bonding determine the formation, stability, and myriad functions of proteins?

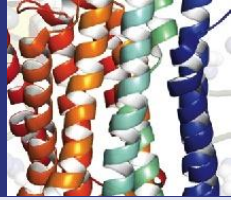
- 化學鍵結力如何決定蛋白質的形成、安定性、與各種功能

Outline



- What noncovalent interactions stabilize protein structure?
- What role does the amino acid sequence play in protein structure?
- What are the elements of secondary structure in proteins, and how are they formed?
- How do polypeptides fold into three-dimensional protein structures?
- How do protein subunits interact at the quaternary level of protein structure?

Protein Structure and Function Are Tightly Linked
蛋白質的(各級)結構與(生理)功能是緊密聯結的關係



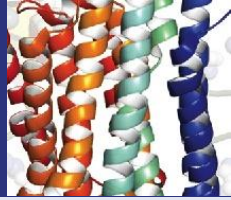
- The three-dimensional structures of proteins and their biological functions are linked by several overarching principles:

- Function depends on structure
- Structure depends on sequence and on weak, noncovalent forces (有那些非共價鍵參與?決定結構, 決定功能)
- The number of protein folding patterns is large but finite

  (蛋白質折疊形式有限)
- Structures of globular proteins are marginally stable

  (球狀蛋白結構不安定)
- Marginal stability facilitates motion (分子運動, 給予功能)
- Motion enables function

6.1 Noncovalent Interactions Stabilize the Higher Levels of Protein Structures



What are these "weak forces"?

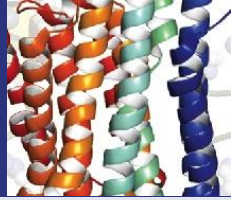
Peptide backbone

Side chains

安定高階結構在骨幹與側鏈的非共價鍵的種類:

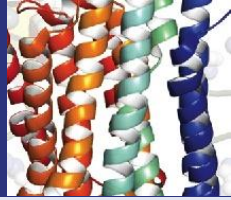
- van der Waals: 0.4 - 4 kJ/mol
- hydrogen bonds: 12-30 kJ/mol
- ionic bonds: 20 kJ/mol
- hydrophobic interactions: <40 kJ/mol

6.1 Noncovalent Interactions in/on Proteins



- Secondary, tertiary, and quaternary structure of proteins is formed and stabilized by weak forces
- Hydrogen bonds are formed wherever possible
- Hydrophobic interactions drive protein folding
- Ionic interactions usually occur on the protein surface
- Van der Waals interactions are ubiquitous (到處存在)

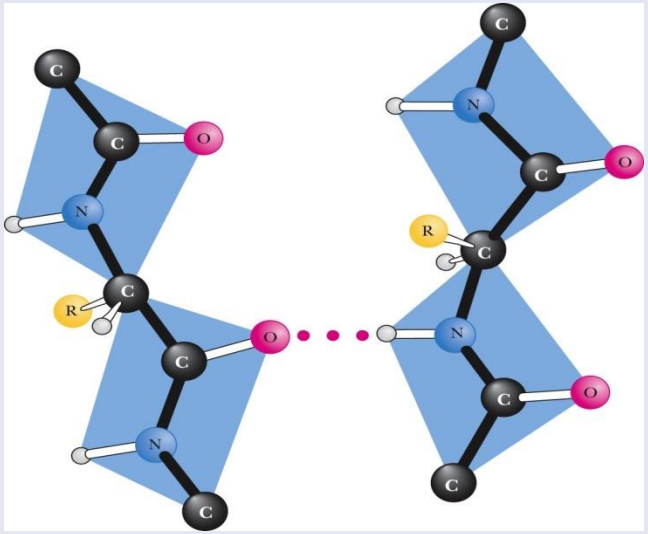
Hydrogen bonds are formed wherever possible
蛋白質結構的氫鍵形成有很多可能的位置



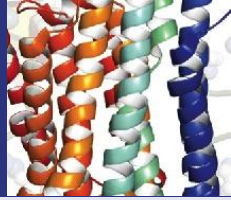
- Component atoms (O, N, S, P etc.) of peptide backbone
- Side chains

  S, T, Y, D, E, N, Q, K, R, H, M, C
- Usually on protein surface
- HB in the protein interior provide stability

Figure 6.5 Schematic drawing of a hydrogen bond between a backbone C=O and a backbone N-H.



Hydrophobic interactions drive protein folding
疏水性作用驅動蛋白質折疊

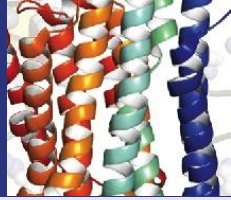


- Clustering between Nonpolar side chains of amino acids and nonpolar solutes

  A, V, L , I, M, W, F, P

- Clustering is entropically driven
- The interior/core of the protein structures are exclusively hydrophobic

Ionic interactions usually occur on the protein surface 離子交互作用常位於蛋白質表面



- Electrostatic attraction between opposite charge
- Electrostatic repulsion between like charges
- Positive: Lys, Arg, His, amino-terminal
- Negative: Glu, Asp, carboxyl-terminal
- Complicated by salt, e.g. NaCl, Na^+ and Cl^- interfere
- e.g.salting-out (鹽析, protein is salted out of the solution, salt compete with water for waters solvation)
  - salting-in (鹽溶)

Electrostatic Interactions in Proteins-1

positively and negatively charged groups

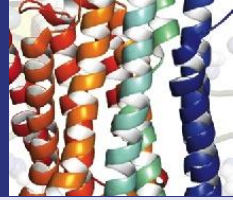
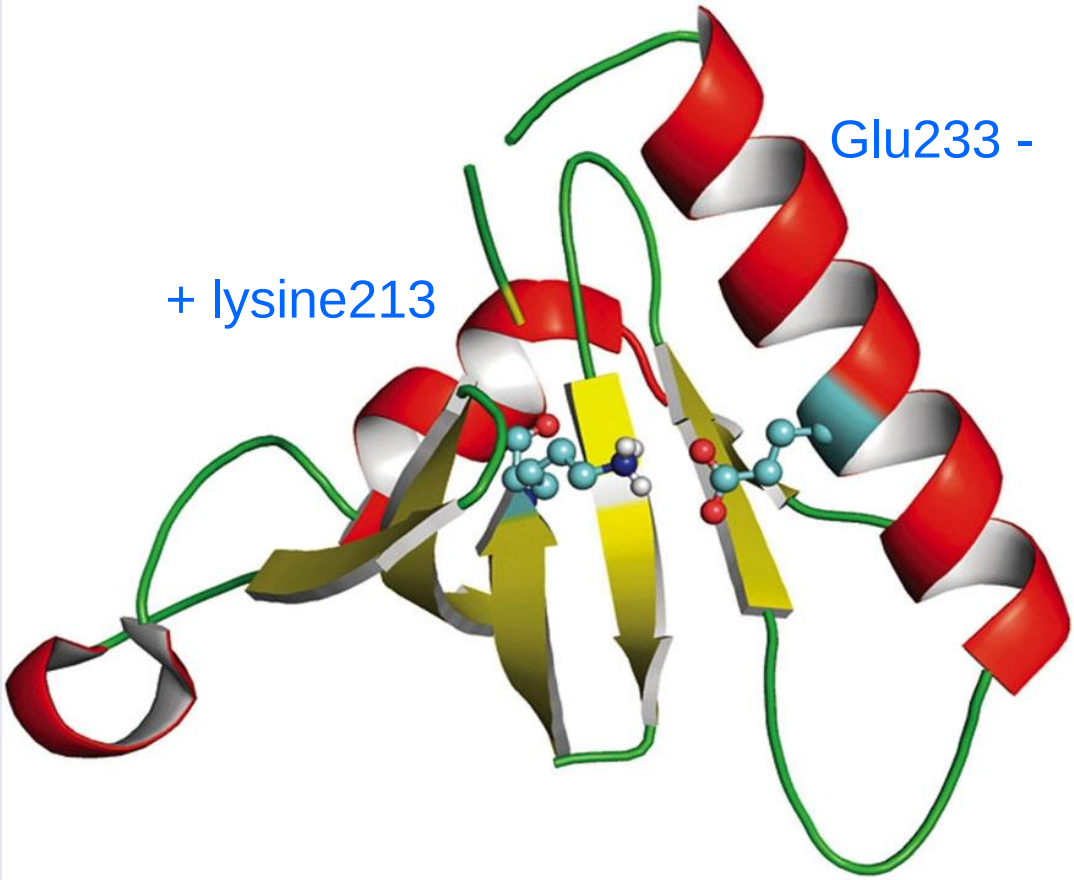


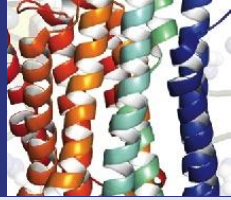
Figure 6.1 An electrostatic interaction between lysine213 and glutamate233 side chains in IRAK-4 (Interleukin-1 (IL-1) receptor-associated kinase-4), an enzyme that phosphorylates other proteins.



Pdb id=2NRY

http://www.ncbi.nlm.nih.gov/Structure/mmdb/mmdbsrv.cgi?ShowOp=VastSum&uid=43499

Electrostatic Interactions in Proteins-2



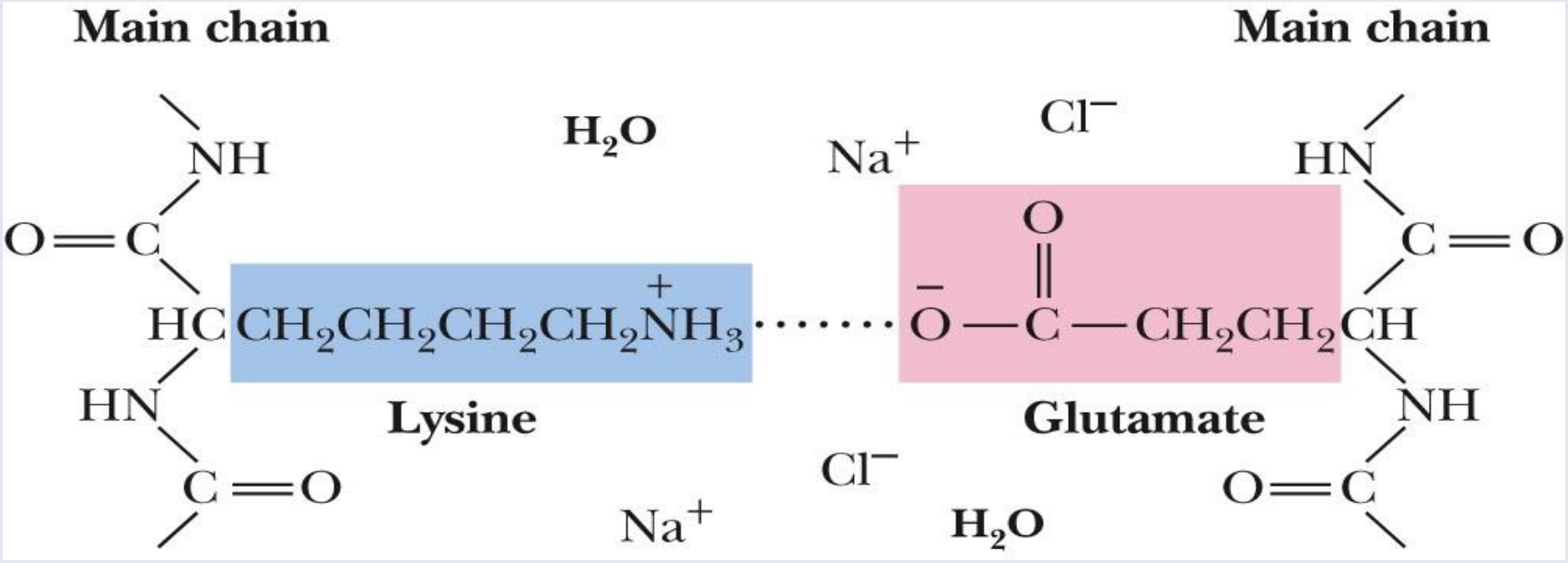
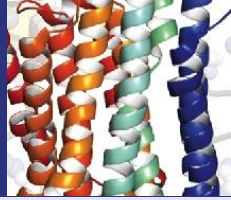


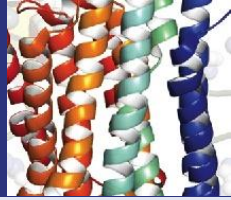
Figure 6.1 An electrostatic interaction between a positively charged lysine ε-amino group (Epsilon) and a negatively charged glutamate γ-carboxyl group.

6.2 What Role Does the Amino Acid Sequence Play in Protein Structure?



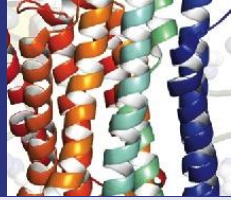
All of the information necessary for folding the peptide chain into its "native" structure is contained in the primary amino acid structure of the peptide.

How do proteins recognize and interpret the folding information?



- Certain loci along the chain may act as nucleation points (折疊核心)
- Protein chain must avoid local energy minima
- Chaperones may help (助折疊的molecular chaperones)
  - 幫助其他蛋白質做出正確折疊產生正確構形的蛋白
  - E.g. the Hsp70 family of proteins, which bind to hydrophobic residues in an unfolded or partially folded protein to prevent incorrect aggregation of these residues.
  - E.g. chaperonins, assist proteins that do not fold spontaneously in their cellular environment.

6.3 What Are the Elements of Secondary Structure in Proteins, and How Are They Formed?



- The atoms of the peptide bond lie in a plane
- All protein structure is based on the amide plane.
- The resonance stabilization energy of the planar structure is 88 kJ/mol.
- A twist about the C-N bond involves a twist energy of 88 kJ/mol times the square of the twist angle.
- Twists can occur about either of the bonds linking the alpha carbon (Cα) to the other atoms of the peptide backbone

  Cα -N Cα-C

The amide or peptide bond planes are joined by the tetrahedral bonds of the α-carbon (Cα).

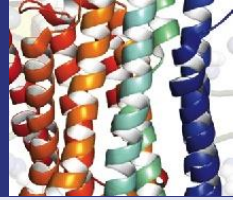
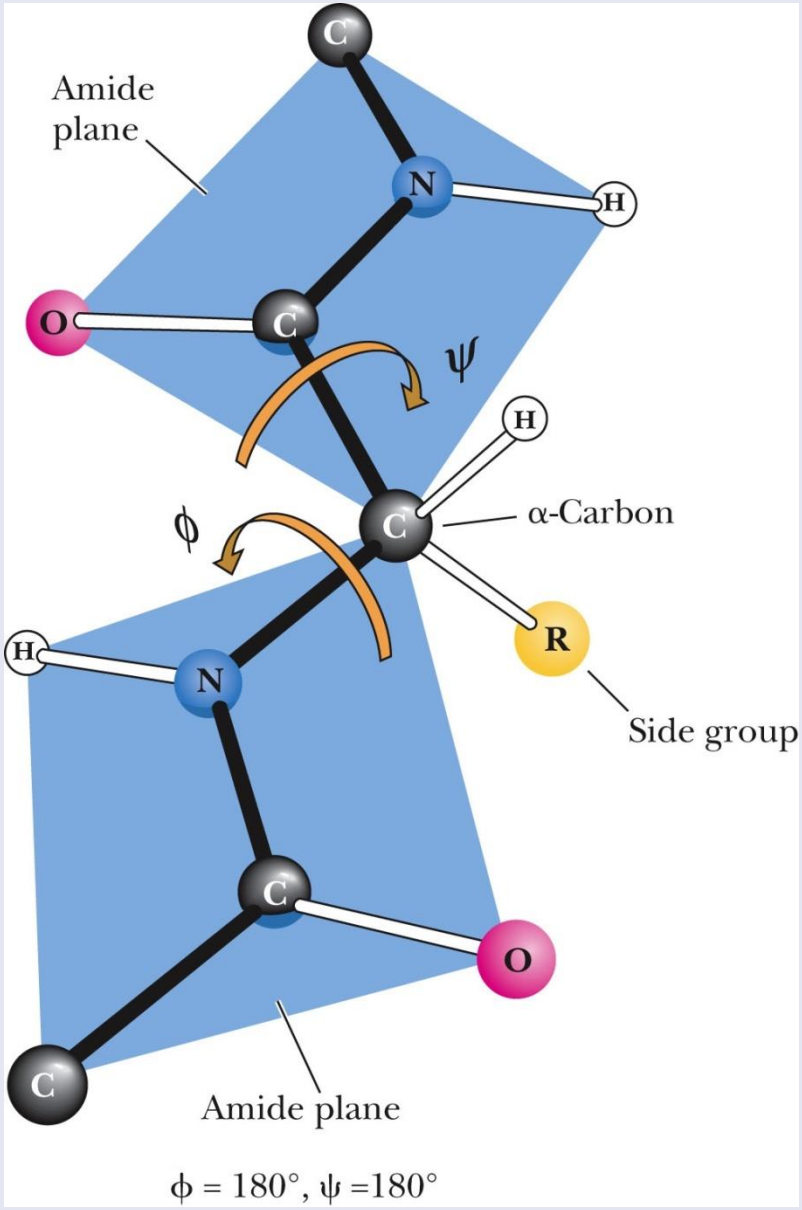


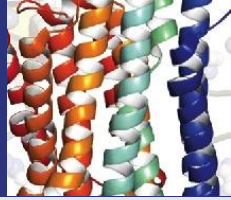
Figure 6.2

The rotation parameters are φ (phi, Cα -N) and ψ (psi, Cα-C).

The conformations shown corresponds to φ= 180° and ψ= 180°.



Consequences of the Amide Plane in the Peptide Bond – specified φ and ψ angles



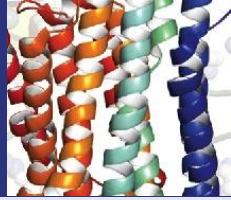
Two degrees of freedom per residue for the peptide chain

Angle about the C_α-N bond is denoted φ (phi)
Angle about the C_α-C bond is denoted ψ (psi)

The entire path of the peptide backbone is known if all φ and ψ angles are specified

Some values of φ and ψ are more likely than others.

Steric crowding cause some values of φ and ψ Are Not Allowed (立體空間擁擠不允許某些轉動角度)



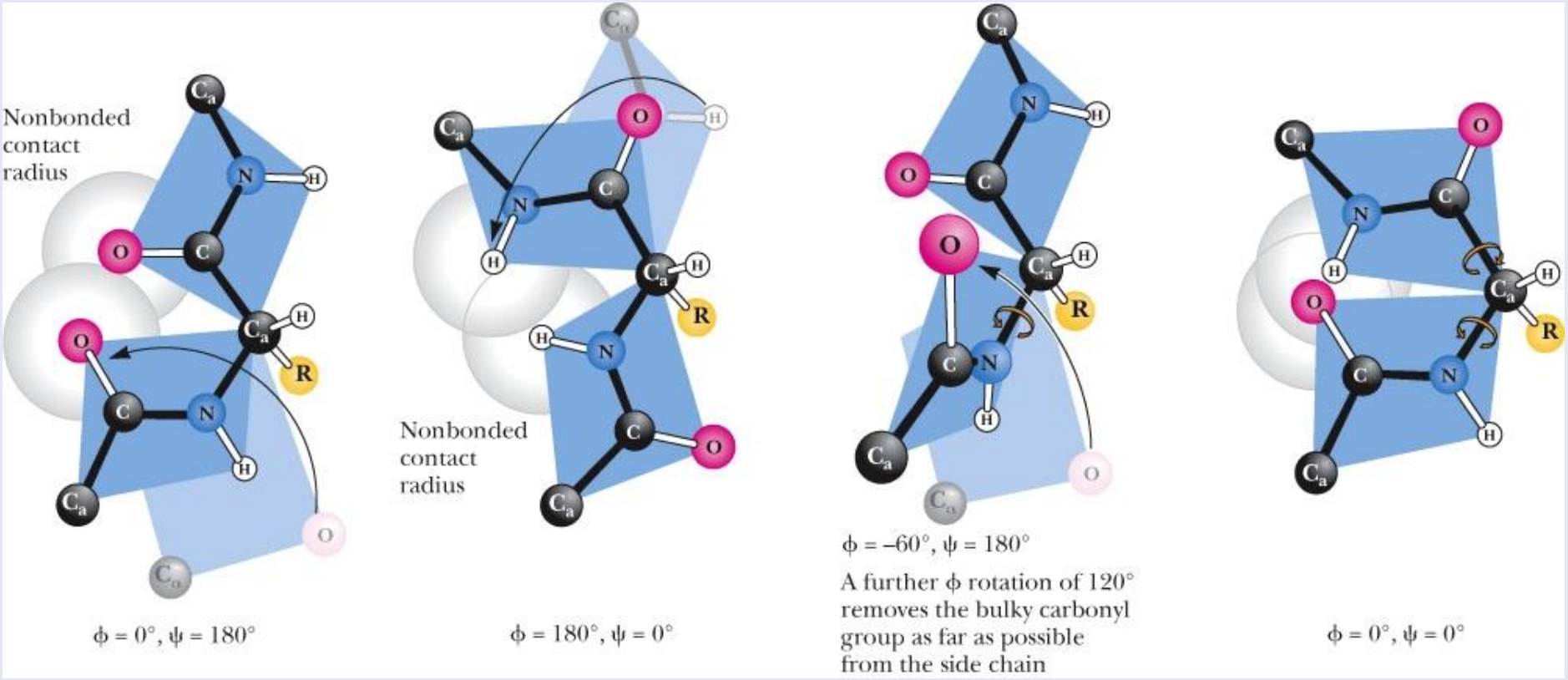
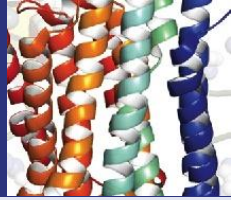


Figure 6.3 Many of the possible conformations about an α-carbon between two peptide planes are forbidden because of steric crowding.

Steric Constraints (*Unfavorable orbital overlap)*

precludes some combinations of φ and ψ



φ = 0°, ψ = 180° is unfavorable
φ = 180°, ψ = 0° is unfavorable
φ = 0°, ψ = 0° is unfavorable

立體空間的約束性排除Cα兩邊C-N φ (phi), C-C ψ (psi)鍵某些旋轉角度的組合

G. N. Ramachandran was the first to demonstrate the convenience of plotting phi,psi combinations from known protein structures

The sterically favorable combinations are the basis for preferred secondary structures

優先形成的二級結構 是基於 有利的立體空間組合

A Ramachandran diagram

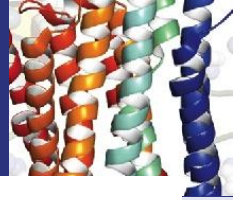
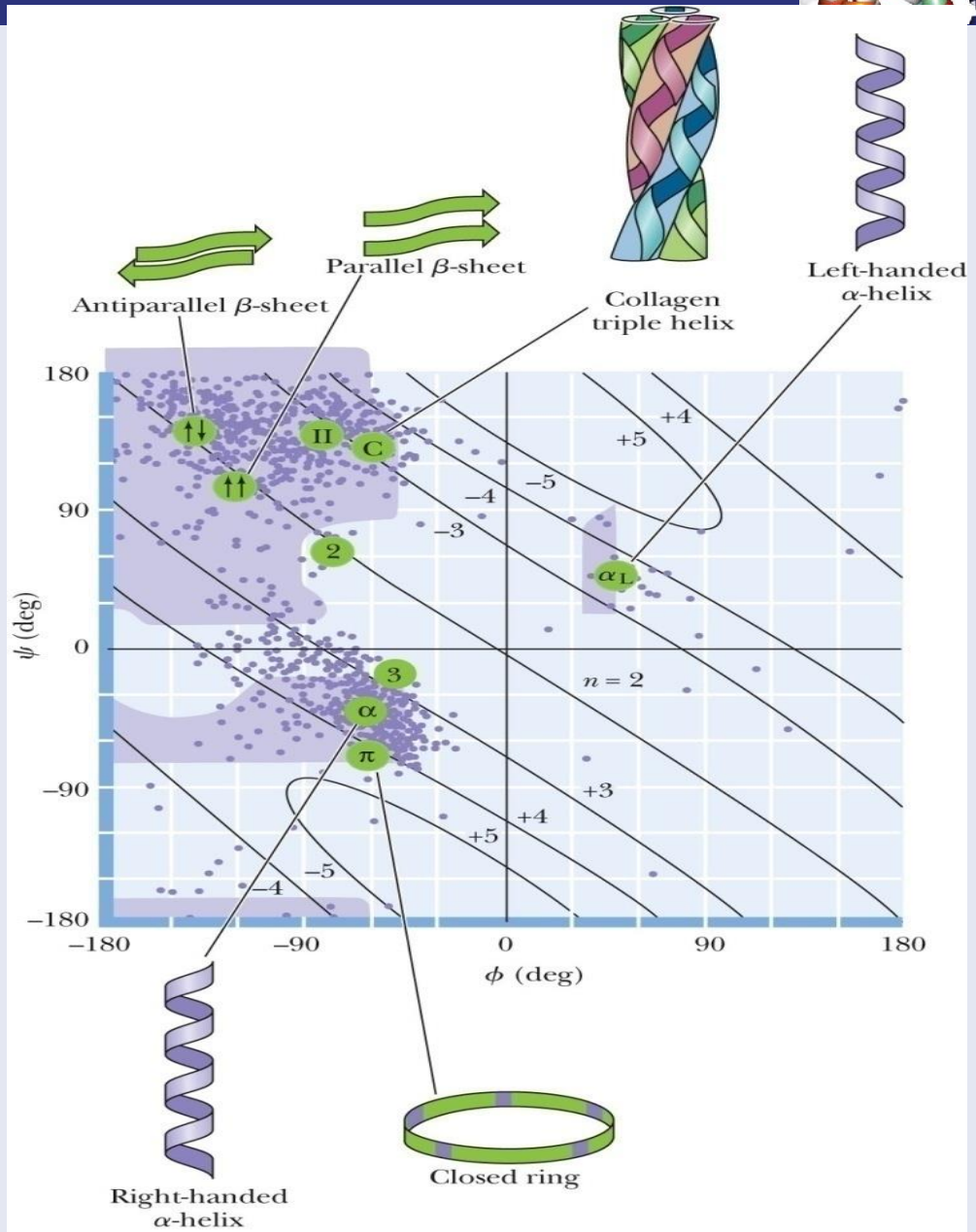


Figure 6.4 showing the sterically reasonable values of the angles φ & ψ.

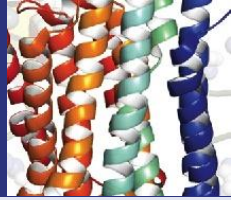
畫出立體性最合理的轉動角度分佈

The shaded regions indicate particularly favorable values of these angles.

Dots in purple indicate actual angles measured for 1000 residues (excluding glycine, for which a wider range of angles is permitted) in eight proteins.



Classes of Secondary Structure



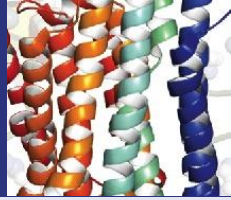
Secondary structures are local structures that are stabilized by hydrogen bonds

二級結構是蛋白質局部結構, 藉H.B. 安定

I. Alpha helices
- Other helices

II. Beta sheet (composed of "beta strands")
- Tight turns (aka beta turns or beta bends)
- Beta bulge

I. The α-Helix (α-螺旋)



- First proposed by Linus Pauling and Robert Corey in 1951
- Identified in keratin by Max Perutz

- A ubiquitous (到處存在) component of proteins
- Stabilized by H bonds

Four different representations of the α-helix

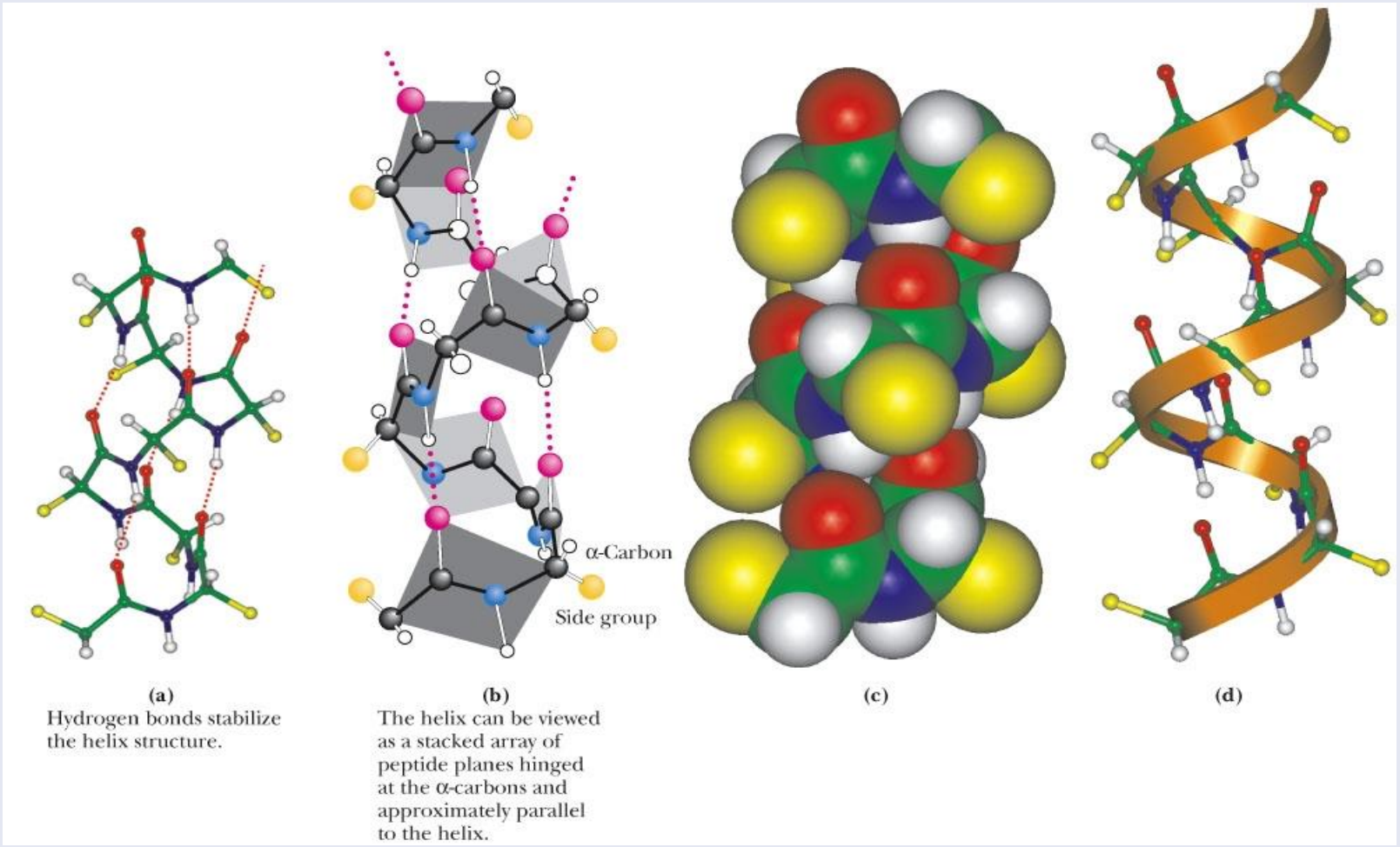
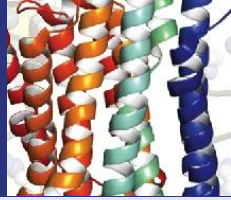


Figure 6.6 (a) ball-stick (b) helix (c) space-filling (d) ribbon

The α-Helix-- *Numbers to Know*



- Residues per turn: 3.6
- Rise per residue: 1.5 Angstroms
- Rise per turn (pitch): 3.6 x 1.5Å = 5.4 Angstroms
- The backbone loop that is closed by any H-bond in an alpha helix contains 13 atoms

(每個氫鍵以13個原子夾著)

- φ = -60 degrees, ψ = -45 degrees (right-handed)
- The non-integral number of residues per turn was a surprise to crystallographers
  - (5轉含18殘基=1轉含3.6殘基)

Proteins contain substantial amounts of α-helix

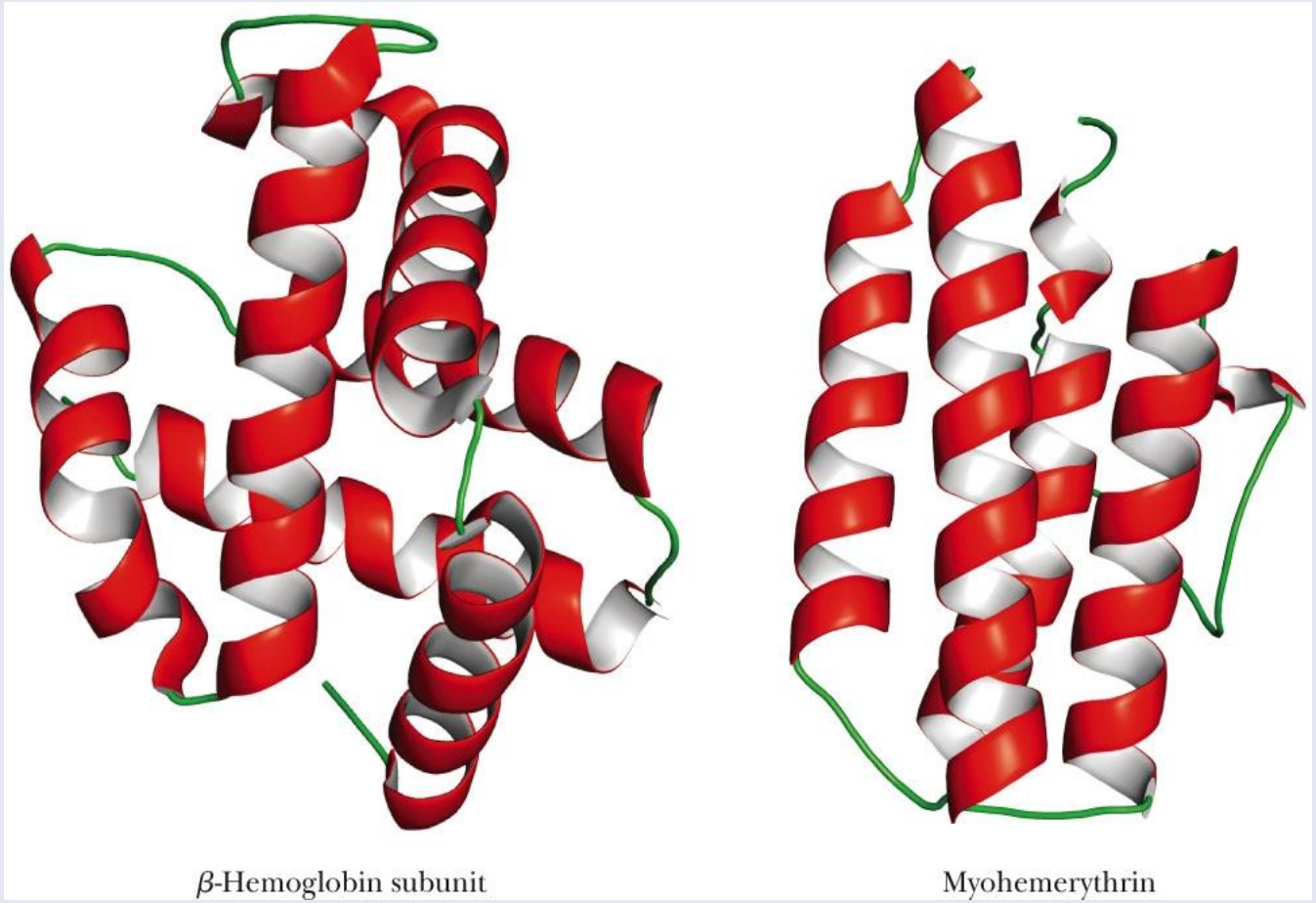


Figure 6.7 Two proteins that contain substantial amounts of α-helix.

Exposed N-H and C=O groups at the N- and C-ends of an α-Helix can be “capped”. 兩尾端有頭蓋

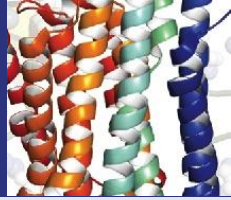


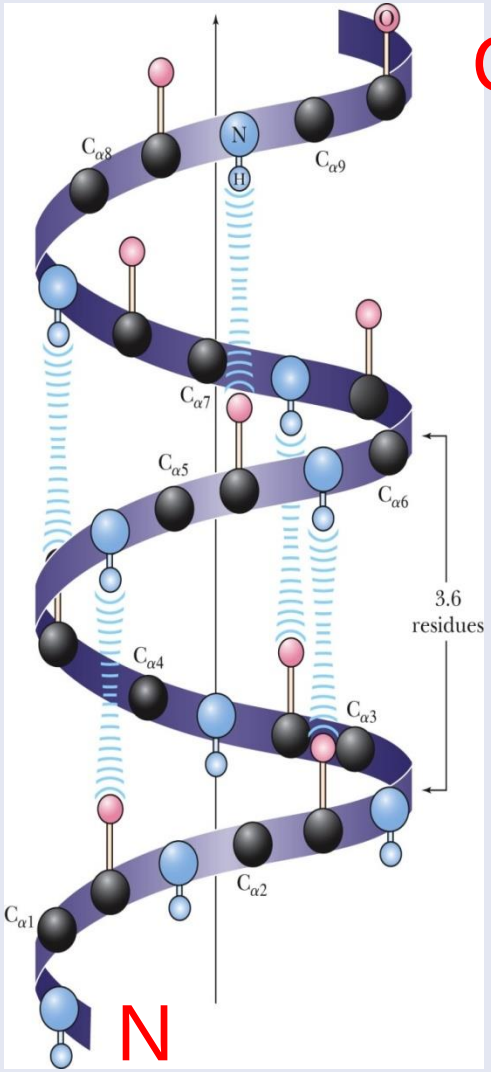
Figure 6.9 Four N-H groups at the N-terminal end of an α-helix and four C=O groups at the C-terminal end lack partners for H-bond formation.

The formation of H bonds with other nearby donor and acceptor groups is referred to as helix capping.

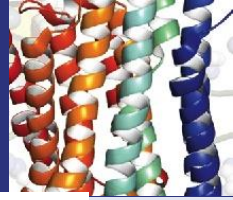
Capping may also involve appropriate hydrophobic interactions that accommodate nonpolar side chains at the ends of helical segments.

參與H.B. 的官能基稱為”helix capping”, 但兩端各有4個N-H 和C=O不參與H.B., “capping”是非極性側鏈參與的疏水性交互作用

Carbonyl(C＝O)與下游的H-N- 生成氫鍵



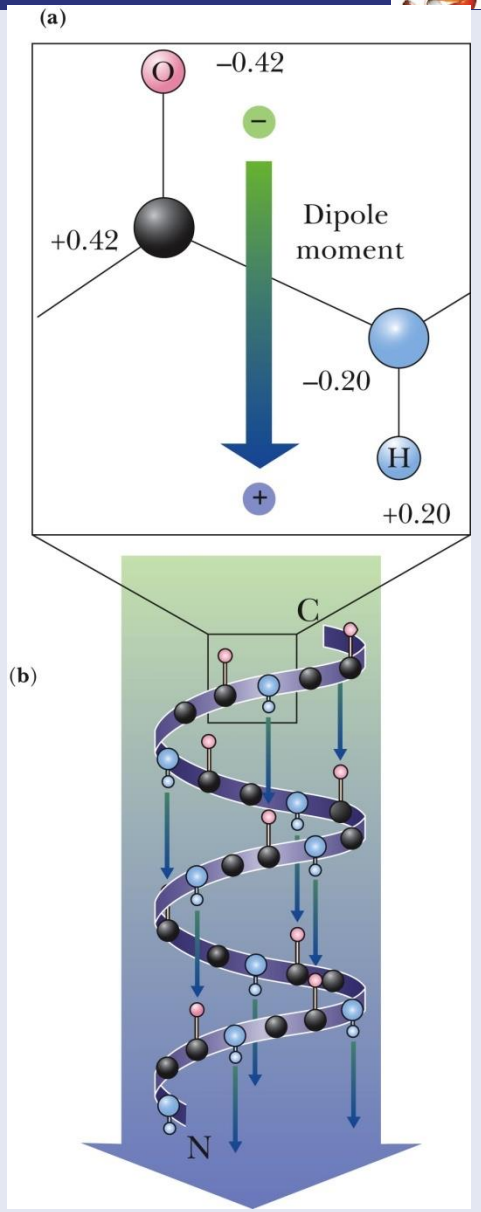
The α-Helix Has a Substantial Net Dipole Moment



整個α helix成圓筒狀，且有偶極性

Figure 6.8 The arrangement of N-H and C=O groups (each with an individual dipole moment) along the helix axis creates a large net dipole moment for the helix.

The numbers indicate fractional charges on respective atoms.



Amino acids can be classified as helix-formers or helix breakers

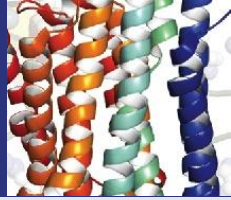


TABLE 6.1 Helix-Forming and Helix-Breaking Behavior of the Amino Acids

Amino Acid		Helix Behavior*	
A	Ala	H	(I)
C	Cys	Variable	
D	Asp	Variable	
E	Glu	H	
F	Phe	H	
G	Gly	I	(B)
H	His	H	(I)
I	Ile	H	(C)
K	Lys	Variable	
L	Leu	H	
M	Met	H	
N	Asn	C	(I)
P	Pro	B	
Q	Gln	H	(I)
R	Arg	H	(I)
S	Ser	C	(B)
T	Thr	Variable	
V	Val	Variable	
W	Trp	H	(C)
Y	Tyr	H	(C)

H: helix former

B: helix breaker

C: random coil

I: indifference

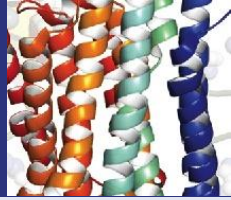
( ): secondary tendency

Variable: depend on pH

Lys, <pH11, Random coil ;
pH12, α-helix

Pro, limit rotation of Cα-N

II. The β-Pleated Sheet (β-摺板, 盾形平面)



- The β-pleated sheet is composed of β-strands
- Also first postulated by Pauling and Corey, 1951
- Strands in a β-sheet may be parallel (同向)or antiparallel (反向)
- Rise per residue:
  - 3.47 Angstroms for antiparallel strands
  - 3.25 Angstroms for parallel strands
  - Each strand of a β-sheet may be pictured as a helix with two residues per turn

parallel (同向)

antiparallel (反向)

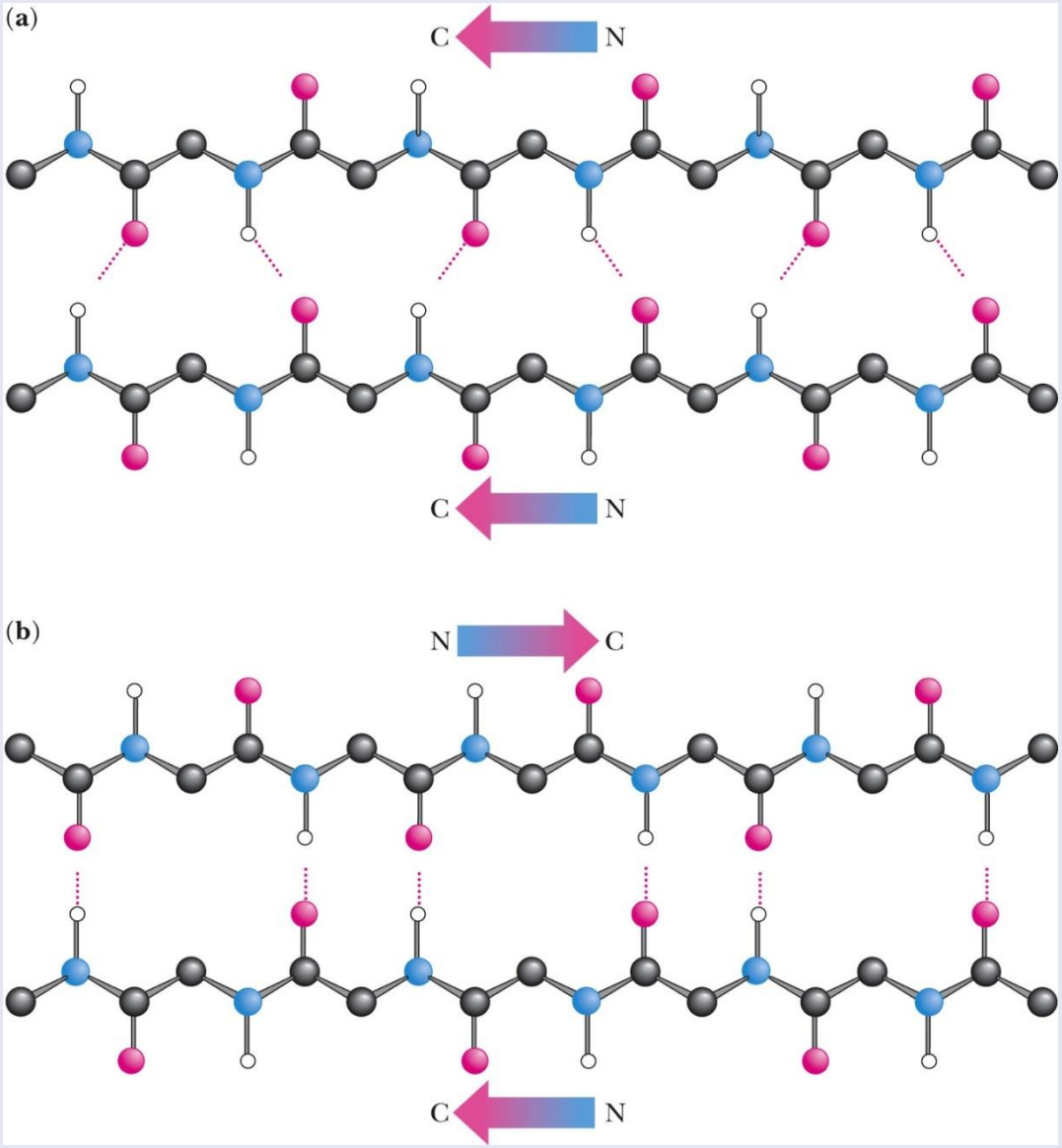
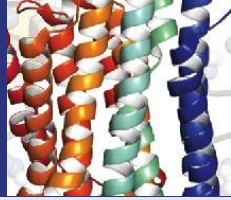


Figure 6-10b p151

An Antiparallel β-Pleated Sheet (silk, flexible, not extendable)



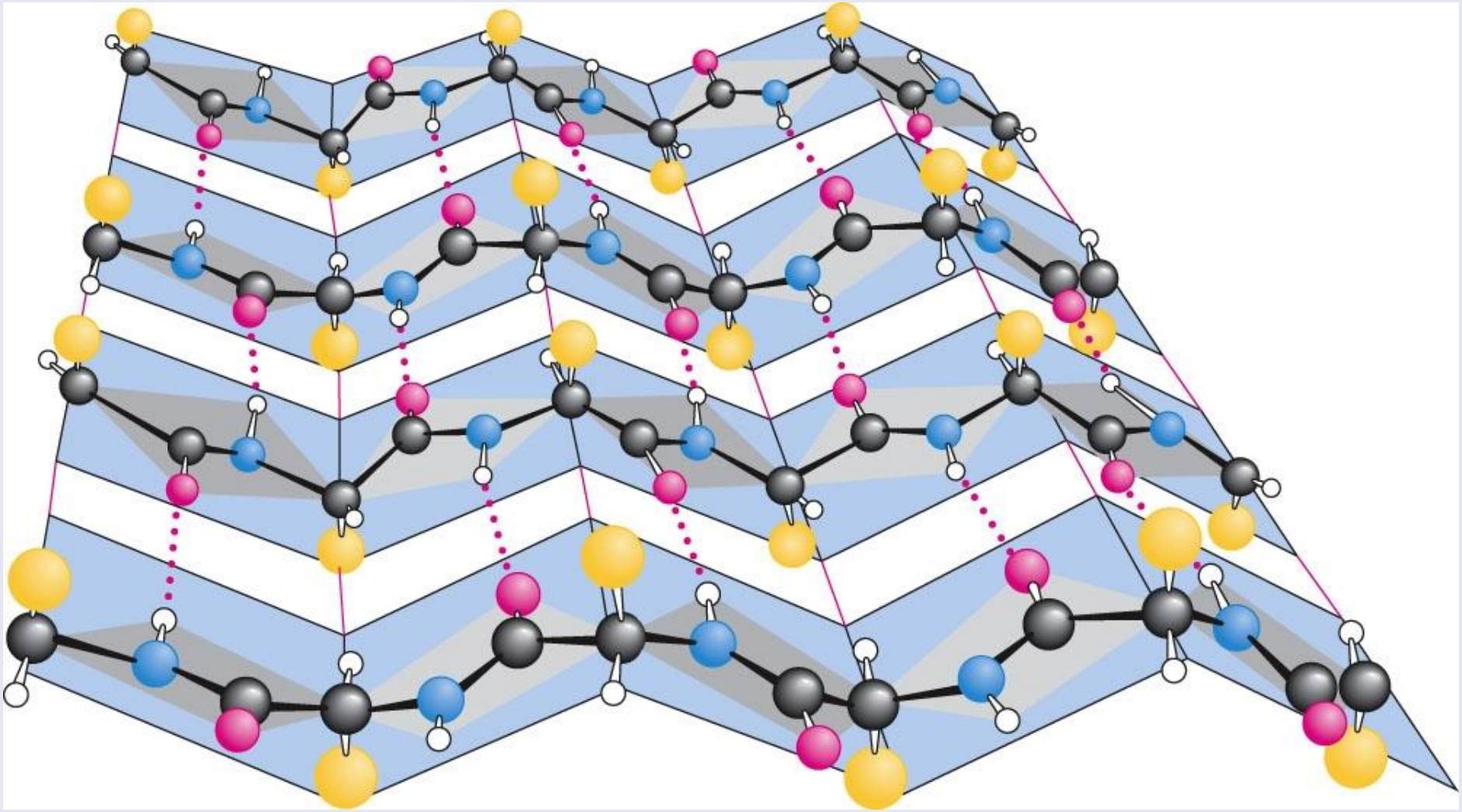
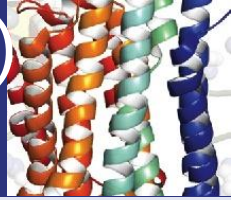


Figure 6.10 多由 R 基團較小的胺基酸 (如 Ala, Gly, Ser) 組成

Helix-Sheet Composites (混合式– strong & elastic) in Spider Silk (蜘蛛網絲, 在地球470 million years)



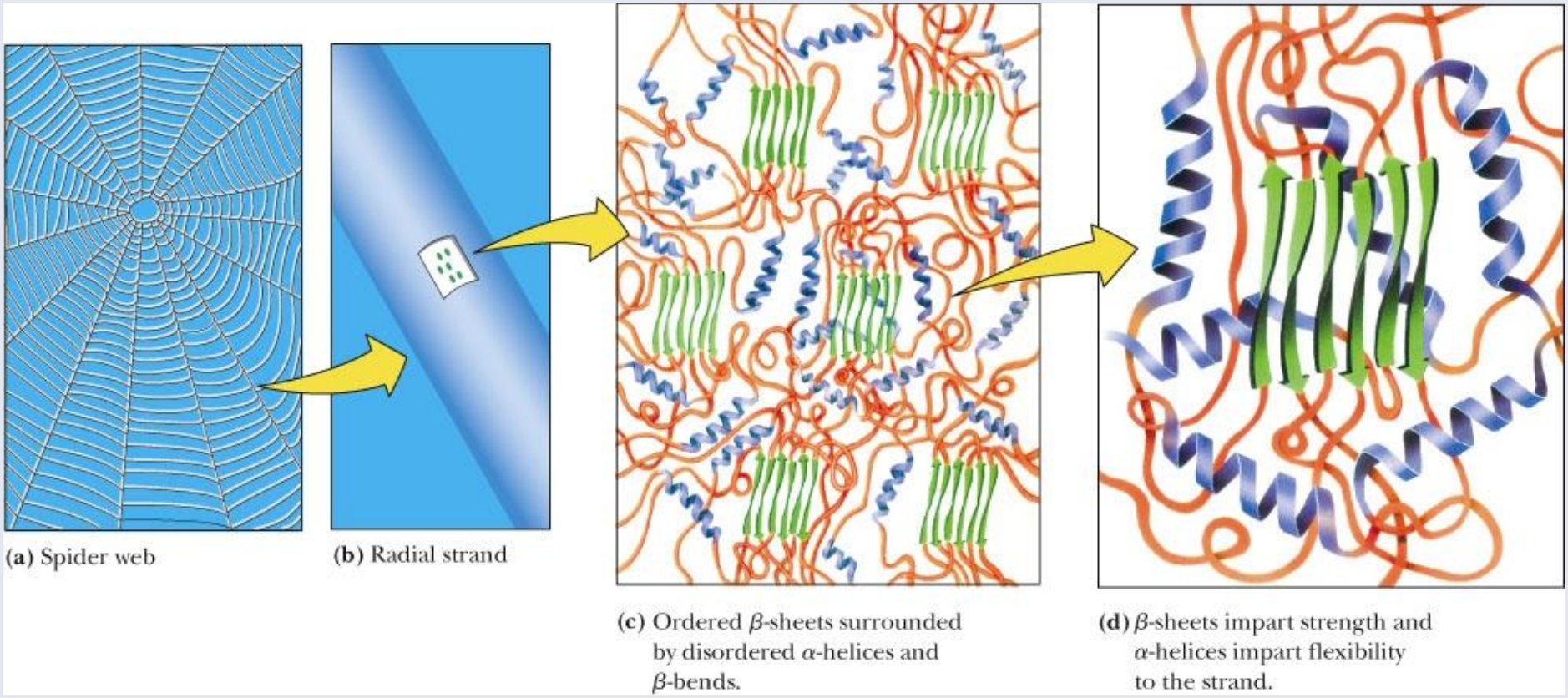
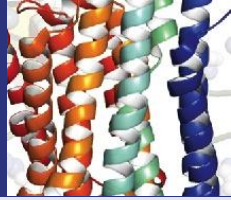


Figure 6.11 Spider web silks are composites of α-helices and β-sheets. The radial strands of webs must be strong and rigid and have a higher percentage of β-sheets. The circumferential strands (termed capture silk) must be flexible and contain a higher percentage of α-helices.

The β-Turn *(aka β-bend, or tight turn)*

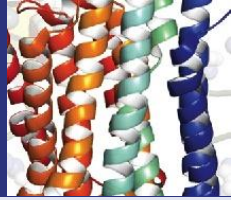


- Allows the peptide chain to reverse direction
  使方向迴轉
- Carbonyl C of one residue is H-bonded to the amide proton of a residue three residues away

- **Proline** (強迫轉向) and **glycine** (小, 提供轉向條件) are prevalent in β-turns

- There are two principle forms of the β-turn
  - tight turns
  - β-bends

Four residues are required to form a β-turn



C=O up the amide plane

C=O below the amide plane

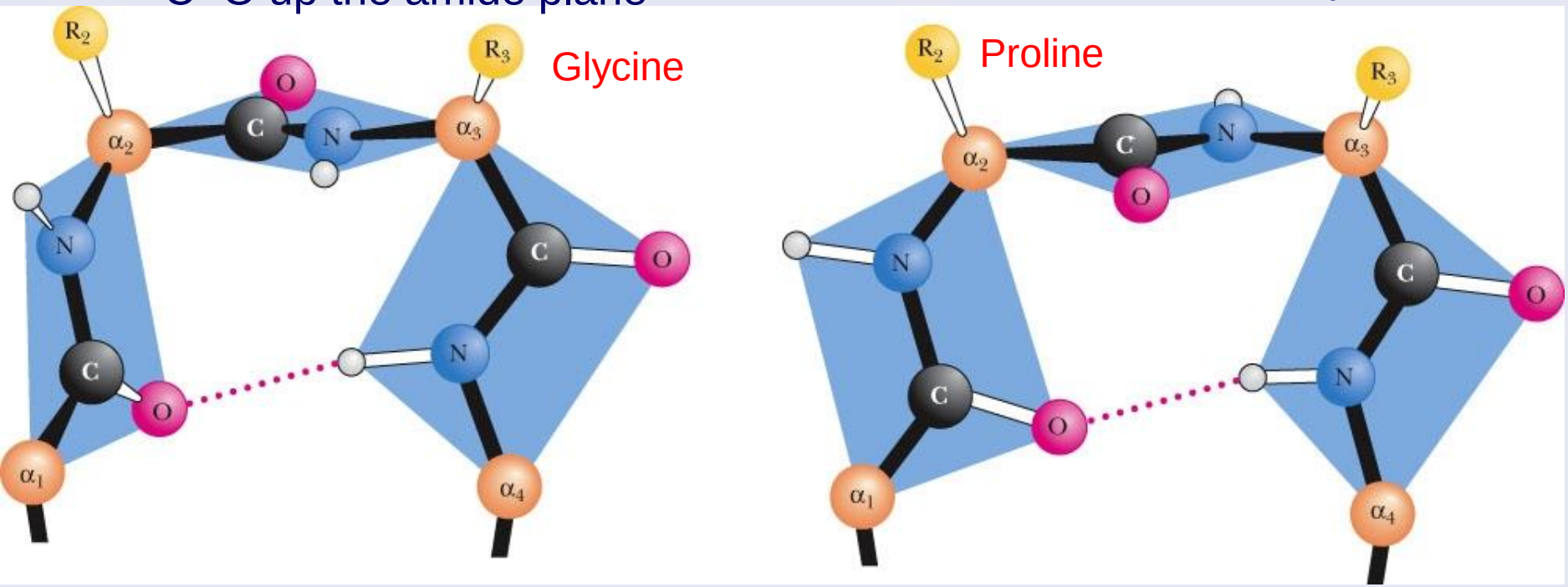
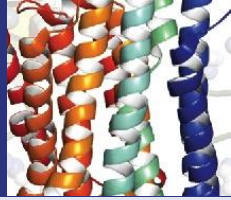


Figure 6.12 The structures of two kinds of β-turns (also called tight turns or β-bends). Four residues are required to form a β-turn. Left: Type I; right: Type II.

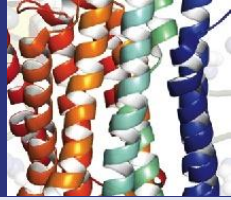
6.4 How Do Polypeptides Fold into Three-Dimensional Protein Structures?



Several important principles:

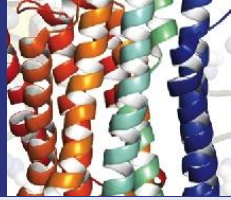
- Secondary structures form wherever possible (due to formation of large numbers of H bonds)
- Helices and sheets often pack close together
- Peptide segments between secondary structures tend to be short and direct
- Proteins fold so as to form the most stable structures.
- Stability arises from: 安定性來自許多分子內氫鍵, 減少折疊後表面積
  - *Formation of large numbers of intramolecular hydrogen bonds*
  - *Reduction in the surface area accessible to solvent that occurs upon folding*

6.4 How Do Polypeptides Fold into Three-Dimensional Protein Structures?



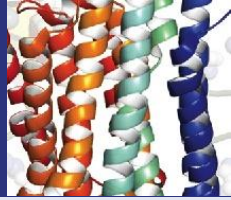
- Two factors lie at the heart of these principles:
  - Proteins are typically a mixture of hydrophilic and hydrophobic amino acids
  - The hydrophobic groups tend to cluster together in the folded interior of the protein
  - 疏水性基團在折疊蛋白內部
  - 親水性基團在外與水環境接觸

Fibrous Proteins (纖維性蛋白)

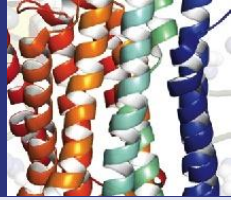


- Much or most of the polypeptide chain is organized approximately parallel to a single axis
- Fibrous proteins are often mechanically strong
- Fibrous proteins are usually insoluble
- Usually play a structural role in nature

- Three examples of fibrous proteins:
  - α-Keratin (角質)
  - β-Keratin
  - Collagen (膠原蛋白)

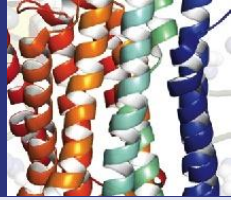


α-Keratin (α-角質)



- A fibrous protein found in hair, fingernails, claws, horns and beaks (指甲,獸爪, 牛角,鳥嘴)
- Sequence consists of 311-314 residue alpha helical rod segments capped with non-helical N- and C-termini
- Primary structure of helical rods consists of 7-residue repeats: $(a\text{-}b\text{-}c\text{-}d\text{-}e\text{-}f\text{-}g)_n$, where a and d are nonpolar.
- This structure promotes association of helices to form coiled coils

Fibroin and β-Keratin: β-Sheet Proteins



Proteins that form extensive beta sheets

- Found in silk fibers and bird feathers (蠶絲, 羽毛)
- Alternating sequence:
  Gly-Ala/Ser-Gly-Ala/Ser....
- Since residues of a β-sheet extend alternately above and below the plane of the sheet, this places all glycines on one side and all alanines and serines on other side!
- This allows Glys on one sheet to mesh with Glys on an adjacent sheet (same for Ala/Sers)

Fibroin and β-Keratin: β-Sheet Proteins

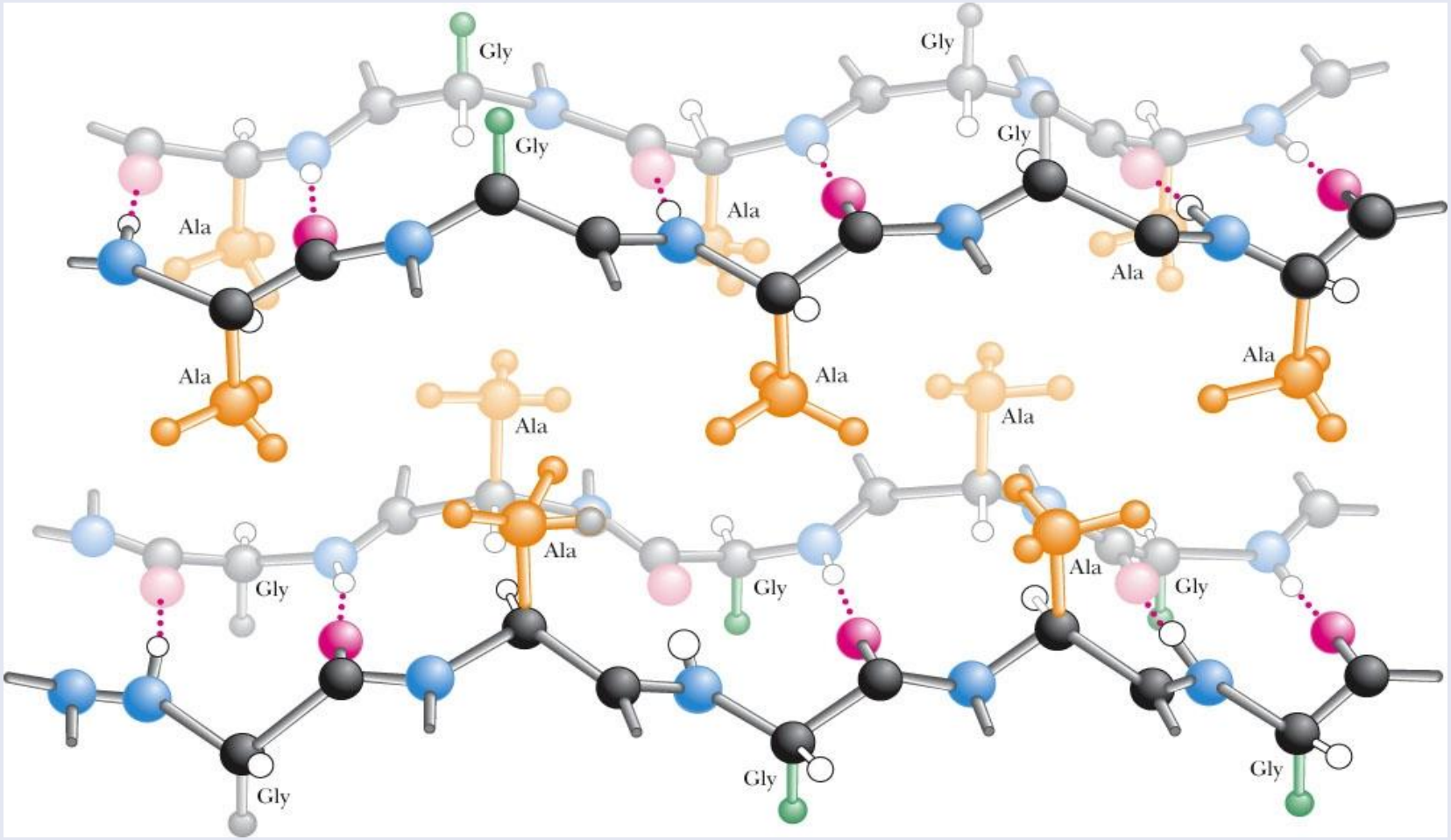
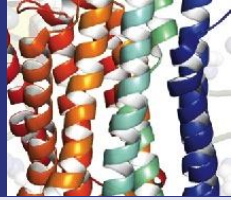


Figure 6.14 Silk fibroin consists of a unique stacked array of β-sheets.



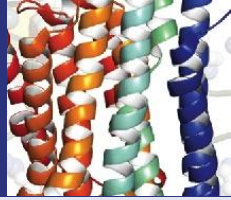
Collagen – A Triple Helix 三股螺旋

Principal component of connective tissue (tendons, cartilage, bones, teeth)

- Basic unit is tropocollagen:
  - Three intertwined polypeptide chains (1000 residues each)
  - MW = 285,000
  - 300 nm long, 1.4 nm diameter

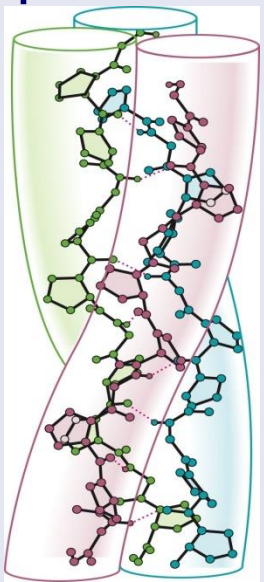
  - Unique amino acid composition, including hydroxylysine and hydroxyproline
  - Hydroxyproline is formed by the vitamin C-dependent prolyl hydroxylase reaction.

The Collagen Triple Helix

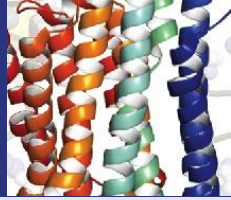


A case of structure following composition

- The unusual amino acid composition of collagen is unsuited for alpha helices or beta sheets
- It is ideally suited for the collagen triple helix: three intertwined helical strands
- Much more extended than alpha helix, with a rise per residue of 2.9 Angstroms
- 3.3 residues per turn
- Long stretches of right-hand Gly-Pro-Pro/HyP

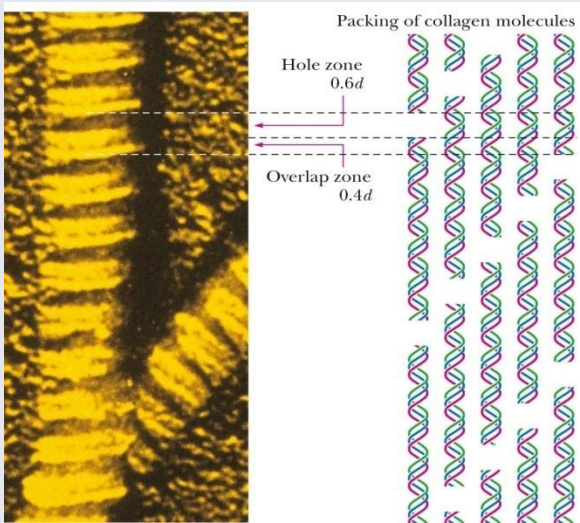


Collagen Fibers are Staggered arrays of tropocollagens

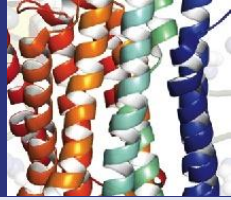


- Banding pattern in EMs with 68 nm repeat
- Since tropocollagens are 300 nm long, there must be 40 nm gaps between adjacent tropocollagens (5 x 68 = 340 nm)
- 40 nm gaps are called "hole regions" - they contain carbohydrate and are thought to be nucleation sites for bone formation

(40 nm gaps是骨骼形成的核心)

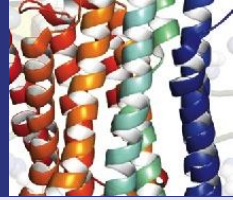


Structural basis of the collagen triple helix 安定三股螺旋的方式



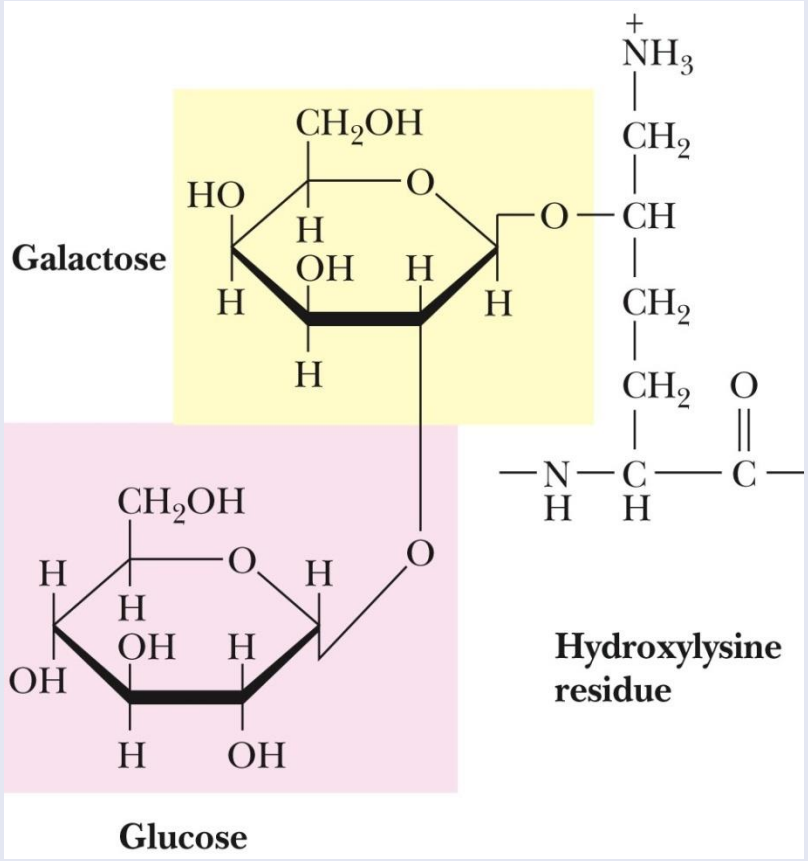
- Every third residue faces the crowded center of the helix - only Gly fits here
- Pro and HyP suit the constraints of φ and ψ
- Interchain H-bonds involving HyP stabilize helix
- Fibrils are further strengthened by intrachain lysine-lysine and interchain hydroxypyridinium crosslinks

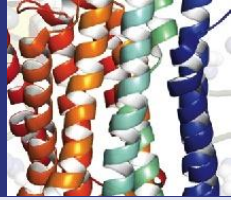
The hole (gaps) regions of collagen fibrils may be the sites of nucleation for bone mineralization



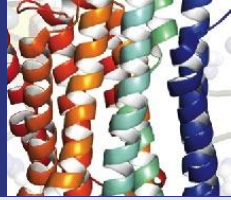
A disaccharide of galactose and glucose is covalently linked to the 5-hydroxyl group of hydroxylysines in collagen by the combined action of galactosyltransferase and glucosyltransferase.

(40 nm gaps是骨骼形成的核心)



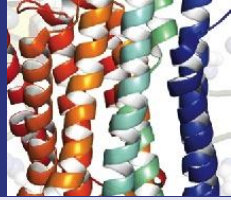


Globular Proteins Mediate Cellular Function
球蛋白 -- 細胞功能



- Globular proteins are more numerous than fibrous proteins
- The diversity of protein structures in nature reflects the remarkable variety of functions they perform
- Functional diversity derives in turn from:
  - The large number of folded structures that polypeptides can adopt
  - The varied chemistry of the side chains of the 20 common amino acids

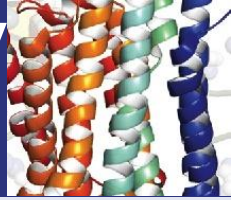
Globular Proteins (球蛋白)



Some design principles

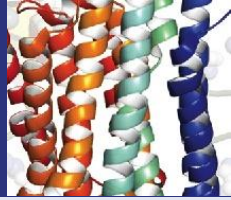
- Helices and sheets make up the core of most globular proteins
- Most polar residues face the outside of the protein and interact with solvent
- Most hydrophobic residues face the interior of the protein and interact with each other
- Packing of residues is close
- However, ratio of vdw (van der Waals) volume to total volume is only 0.72 to 0.77, so empty space exists 28%
- The empty space is in the form of small cavities

Why does the protein core consist primarily of α–helices and β–sheets?



- The protein core is predominantly hydrophobic
- The highly polar N-H and C=O moieties of the peptide backbone must be neutralized in the hydrophobic core (在疏水中心, 極性中和)
- The extensively H-bonded nature of α-helices and β-sheets is ideal for this purpose

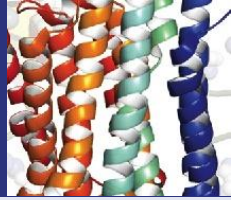
Protein core versus protein surface



- The helices and sheets in the core of a globular protein are typically constant and conserved in sequence and structure
- The protein surface is different in several ways

- Much of the **surface** is composed of loops and tight turns that connect the helices and sheets of the core
- Thus the surface is a complex landscape of different structural elements
- These surface elements can interact with small molecules or with other proteins
- They are the basis for **enzyme-substrate interactions**, cell signaling, and immune responses

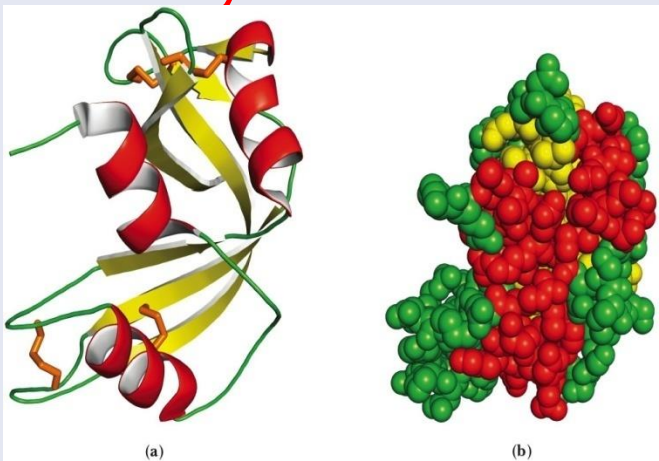
“Random coils” are not random



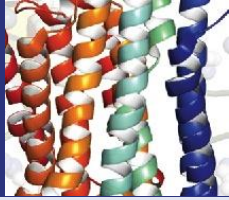
- The segments of a protein that are not helices or sheets are traditionally referred to as “random coil”, although this term is misleading:
- Most of these segments are neither coiled or random

- They are usually organized and stable, but don’t conform to any frequently recurring pattern

- Random coil segments are strongly influenced by side-chain interactions with the rest of the protein



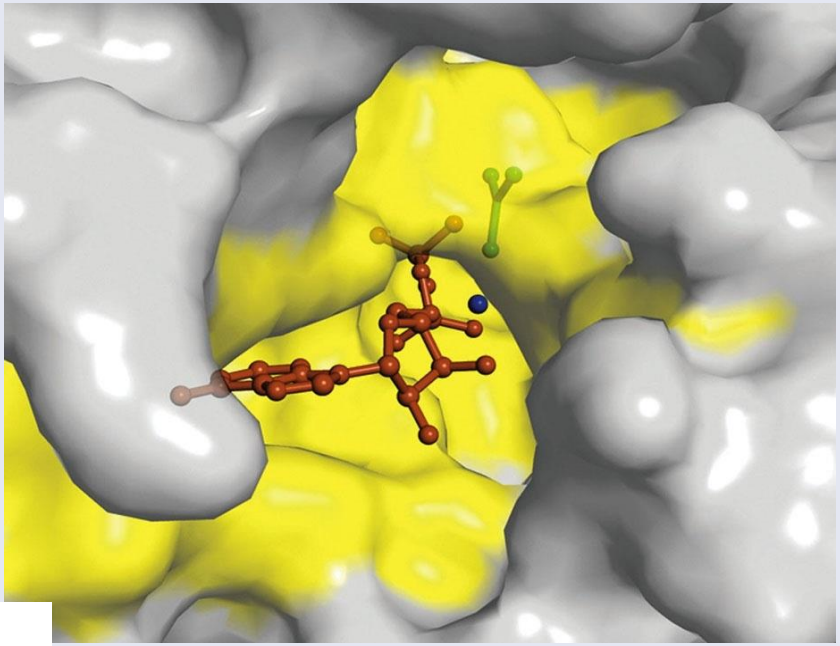
蛋白質表面複雜

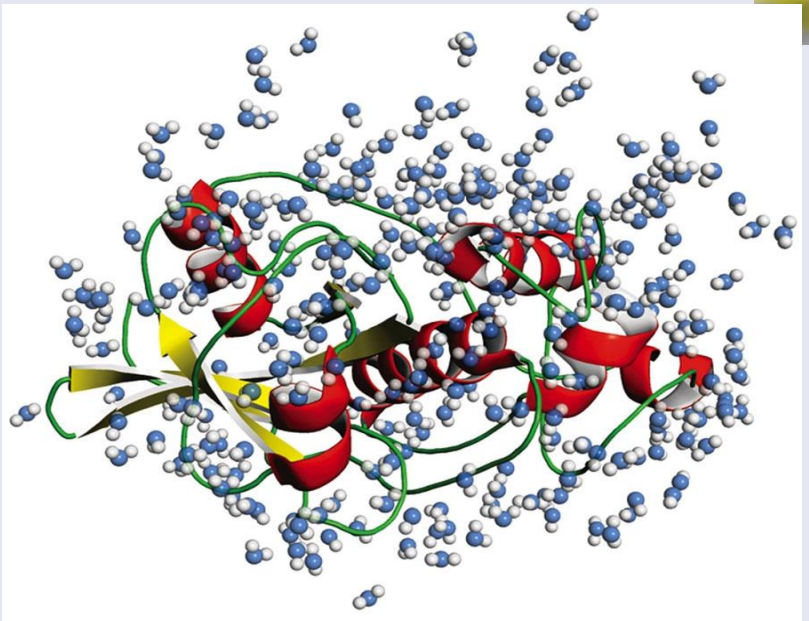


The surfaces of proteins are complementary to the molecules they bind.

H bonds with water molecules

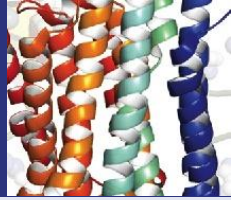
和會結合之分子互補，和水結合





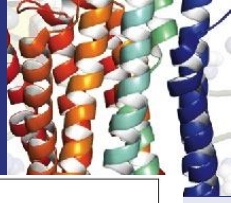
一個side chain會和多個水分子結合 H.B.

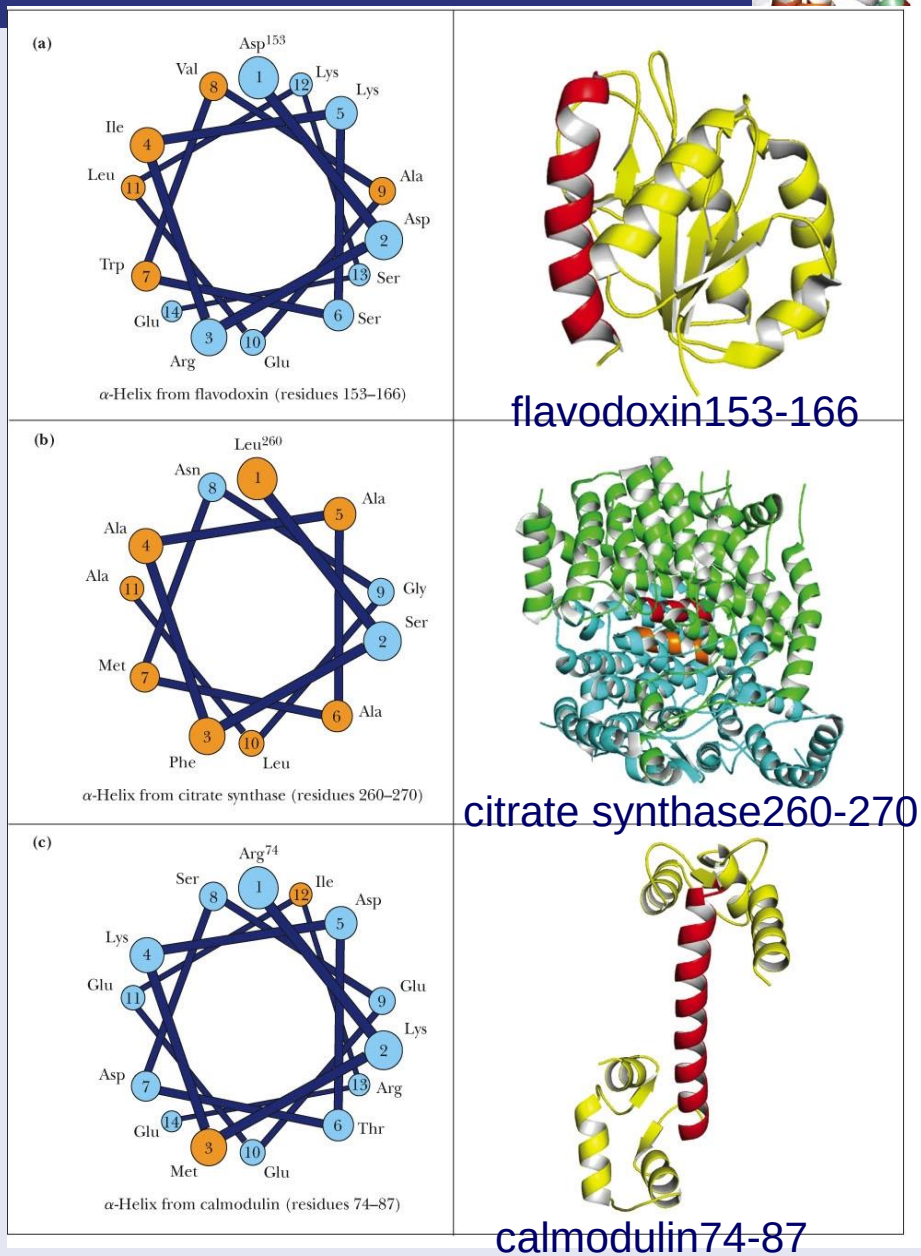
Waters on the Protein Surface Stabilize the Structure



- The surface structure of a globular protein includes water molecules
- The polar backbone and side chain groups on the protein surface make H bonds with solvent water
- α-Helices on a protein surface are usually amphiphilic, with polar and charged residues facing the solvent and nonpolar residues facing the interior
- A helical wheel presentation can reveal the amphiphilic nature of an α-helix
  - Some α-helices are hydrophobic and buried in the protein interior
  - Some helices are polar and entirely solvent-exposed

α-Helices May be Polar, Nonpolar or Amphiphilic (red helix)





flavodoxin153-166

citrate synthase260-270

calmodulin74-87

Figure 6.22 The so-called helical wheel presentation can reveal the polar or nonpolar character of α-helices.

Side chains matter.

(1)多數 “domains ”由單獨連續序列組成

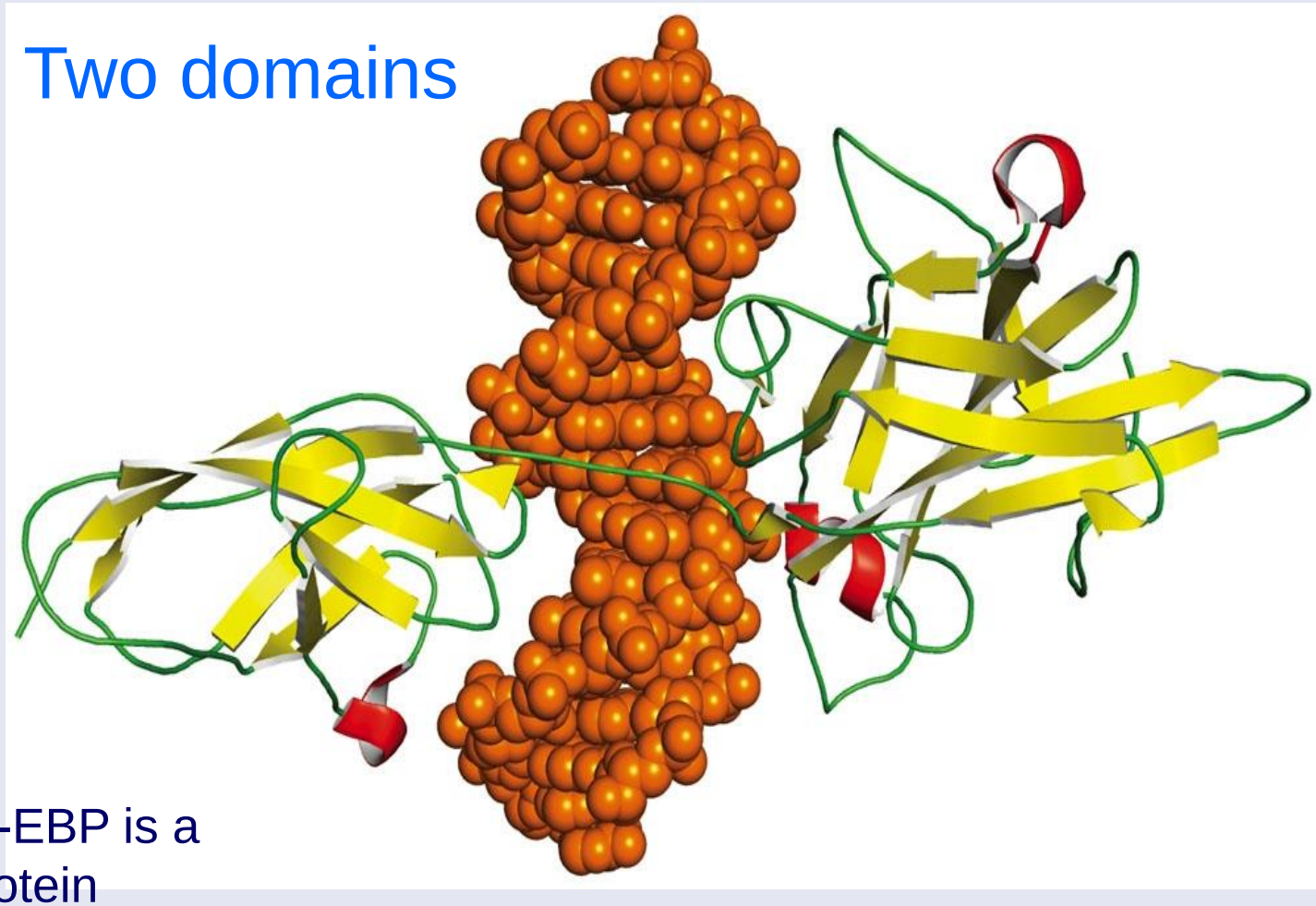
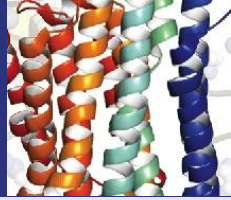


Figure 6.23 Ton-EBP is a DNA-binding protein consisting of two distinct domains.

(2) 大型 “domain ”由中間被另一個domain中斷之 兩段序列組成



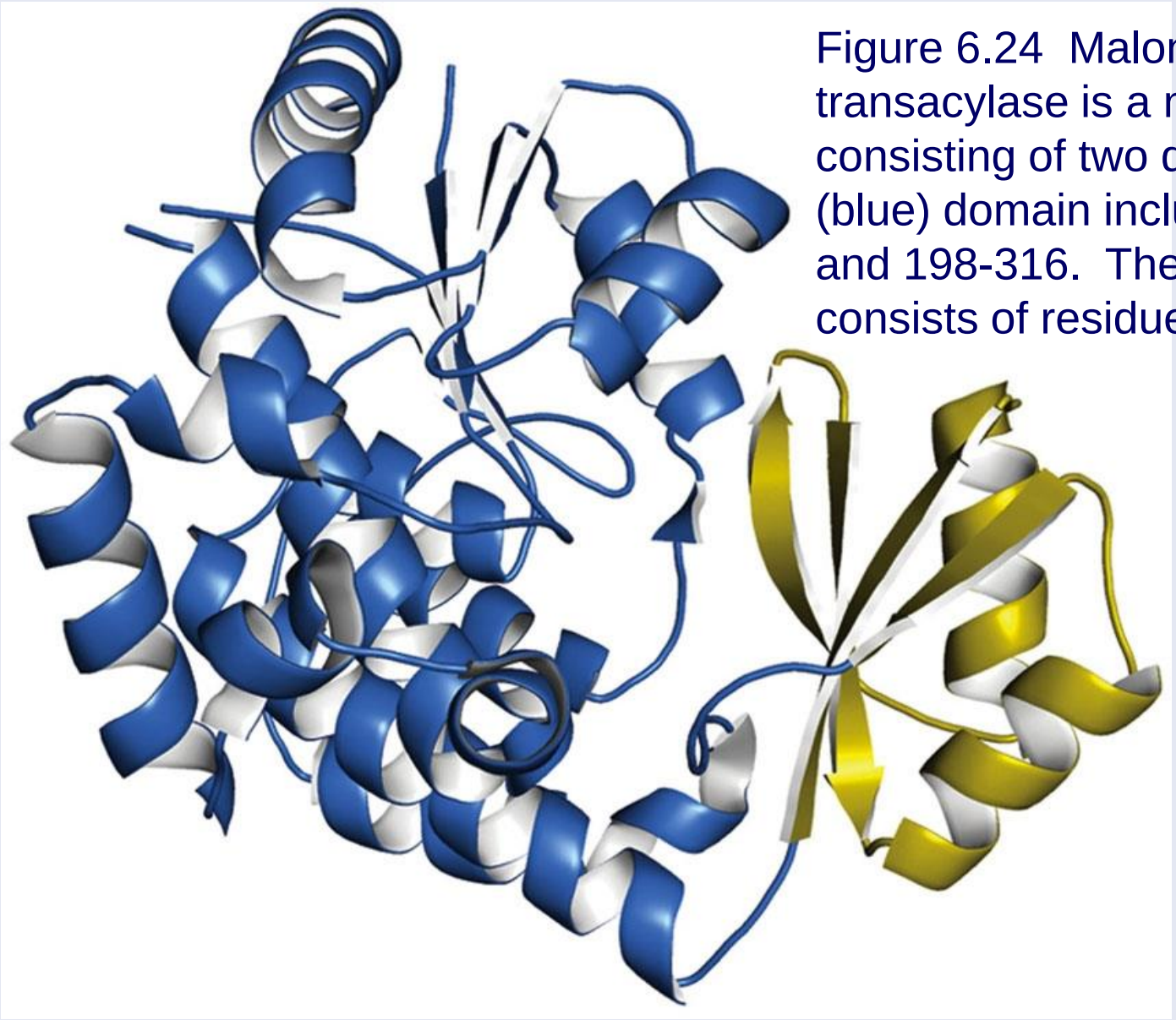
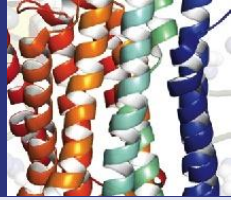


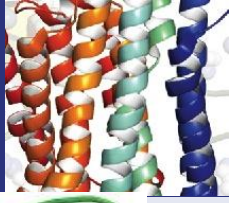
Figure 6.24 Malonyl CoA:ACP transacylase is a metabolic enzyme consisting of two domains. The large (blue) domain includes residues 1-132 and 198-316. The small (gold) domain consists of residues 133-197.

(3) 很多蛋白質 由很多domains組成



- Multidomain proteins typically possess the sum of functional properties and behaviors of their constituent domains
- Proteins consisting of multiple domains probably evolved by the fusion of genes that once coded for separate proteins
  - About 90% of domains in proteins have been duplicated in other proteins
  - Many proteins even contain multiple copies of the same domain
- Some of these often-duplicated domains are shown in Figure 6.25

(4) Several protein modules used in the construction of complex multimodule proteins.



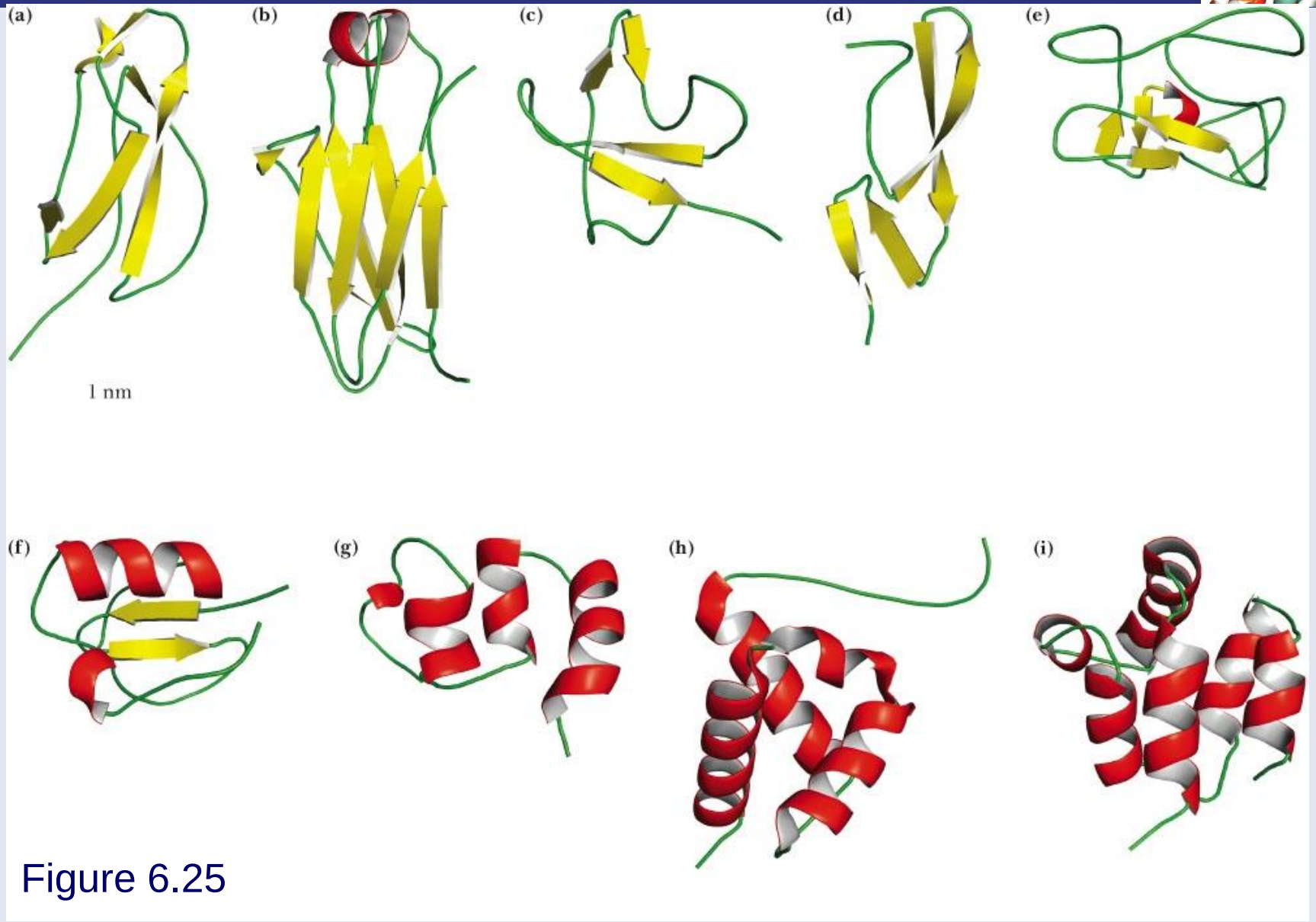


Figure 6.25

Many proteins are composed of several distinct domains (modules, motif, 模組)

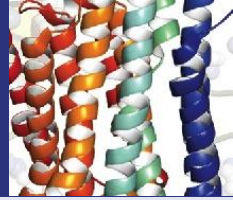
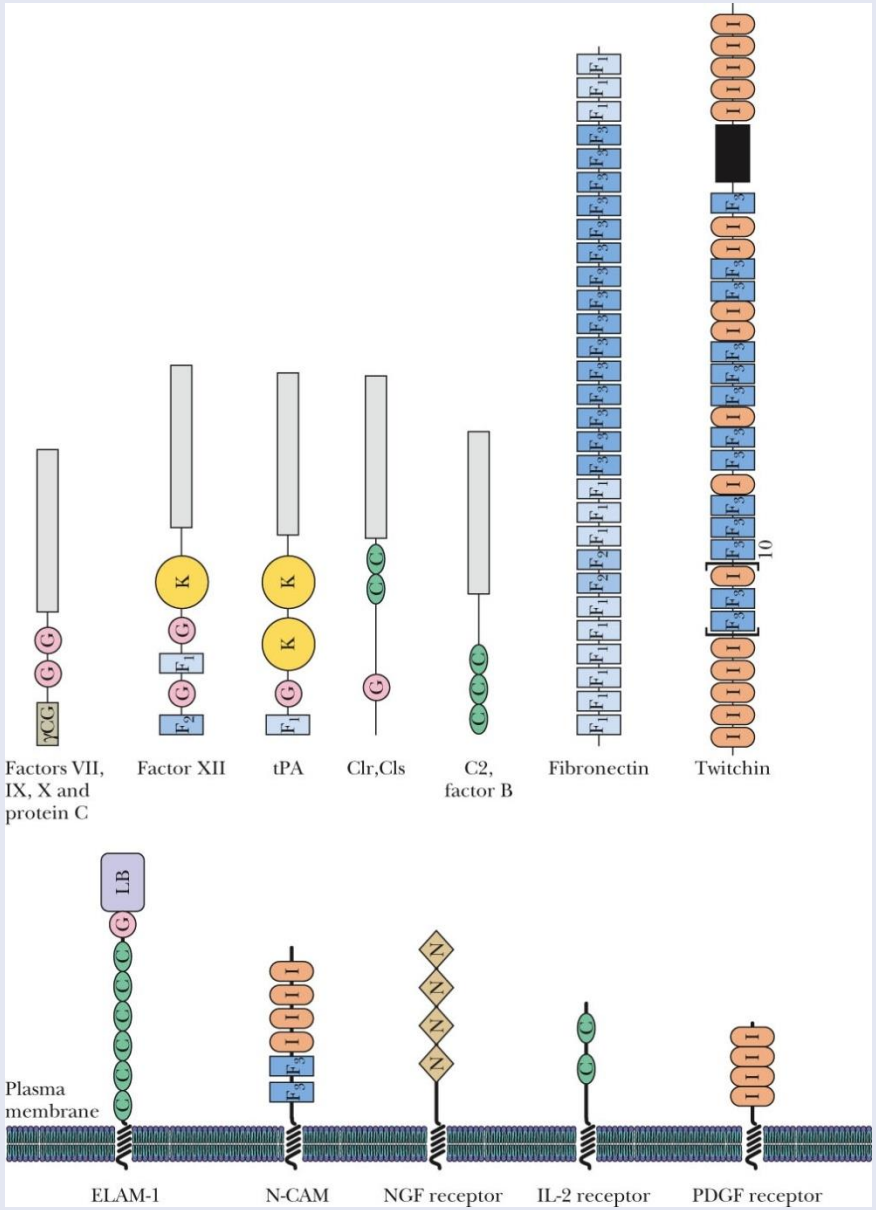
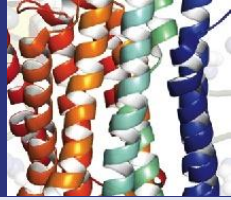


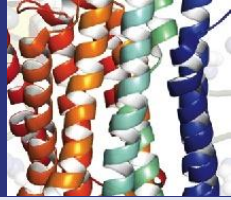
Figure 6.26 A sampling of proteins that consist of mosaics of individual protein modules.

蛋白質由獨立之蛋白質模組的鑲嵌所組成





Denaturation Leads to Loss of Protein Structure and Function (蛋白質變性失去結構與功能)

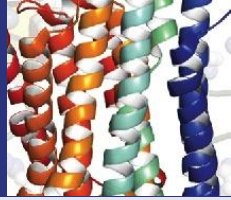


- The cellular environment is suited to maintaining the weak forces that preserve protein structure and function

  * Not including covalent bonds

- External stresses – heat, chemical treatment, etc. – can disrupt these forces in a process termed **denaturation** – the loss of structure and function
  - The cooking of an egg is an everyday example
  - Ovalbumin, the principal protein in egg white, remains in its native structure up to a characteristic melting temperature, T_m
  - $>T_m$, the structure unfolds and function is lost
  - Acid, Base, Organic solvent, Detergent, Particular agent (guanidine-HCl, urea)

Denaturing Agents



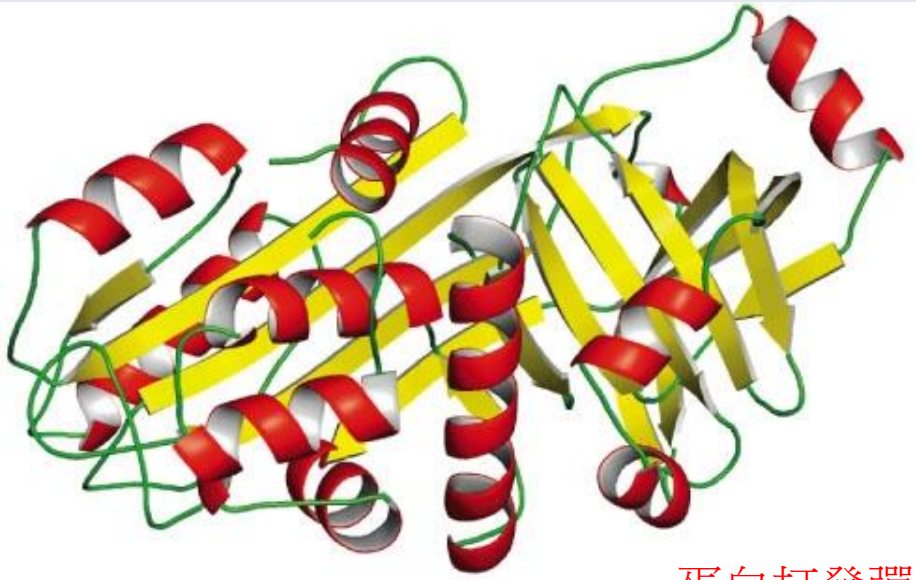
Acid, Base: protonation, deprotonation of dissociable groups, alter inic interaction

Organic solvent: disrupt hydrophobic interactions

Detergent: disrupt both hydrophilic and hydrophobic forces

Particular agent (guanidine-HCl, urea):
Direct: bind to hydrophilic groups
Indirect: altering the structure and dynamics of the water solvents

Denaturation Leads to Loss of Protein Structure and Function



Ovalbumin monomer

蛋白打發彈性與二級結構有關

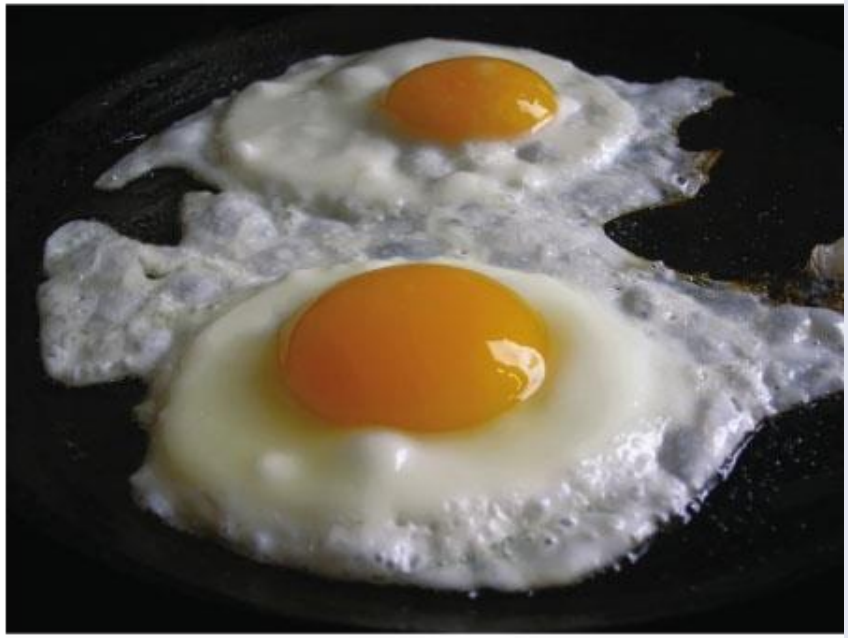
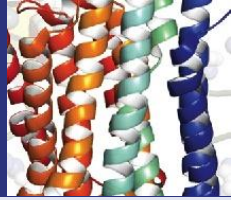


Figure 6.29 The proteins of egg white are denatured during cooking. More than half of the protein in egg white is ovalbumin (10% egg is protein, OV amounts to 54% of it).

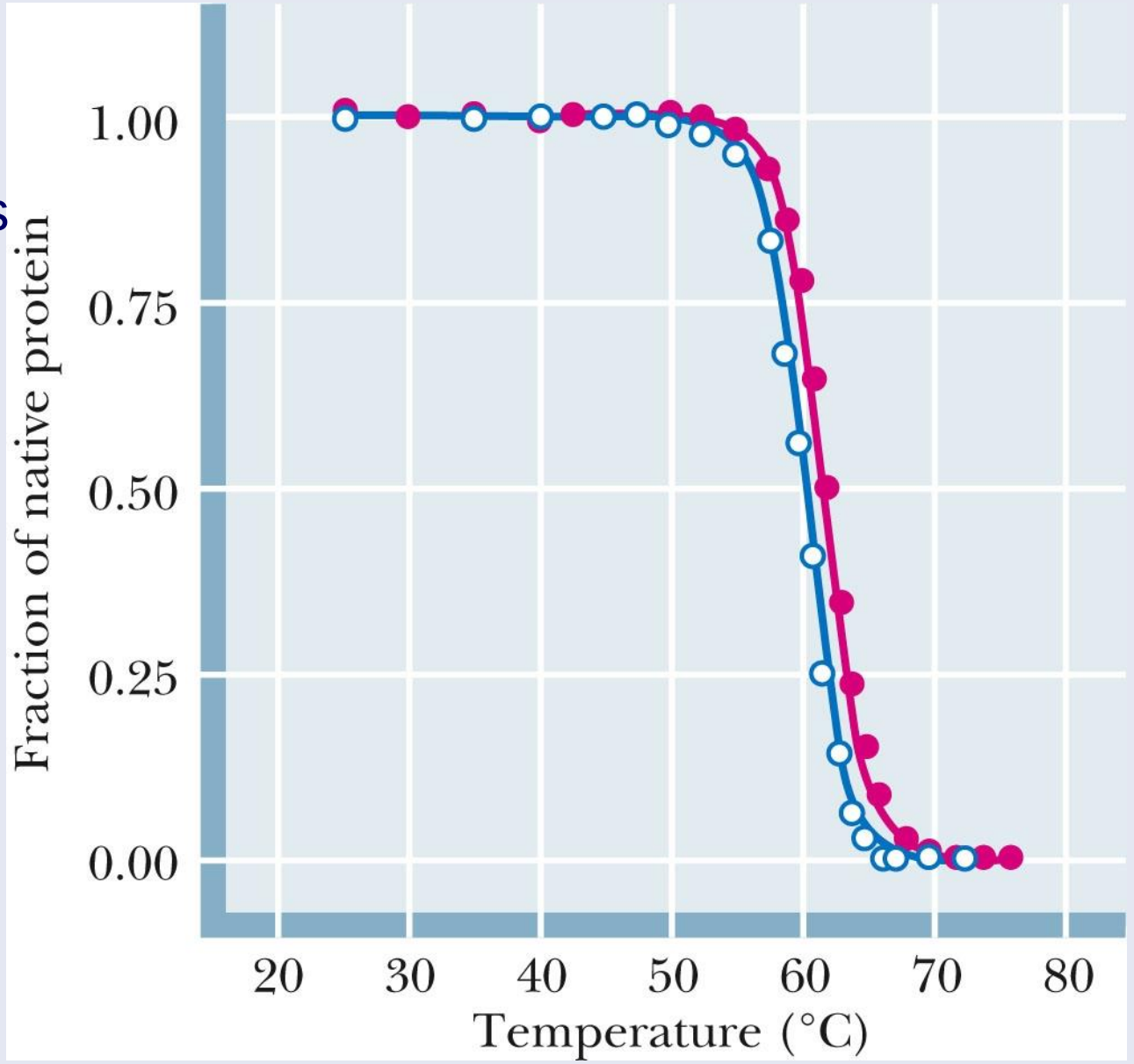
蛋白質變性, unfolding → aggregation (透明, 黏液→ 固態, 白色)

Denaturation Leads to Loss of Protein Structure and Function 高溫造成蛋白質變性



Protein 6.30 Proteins can be denatured by heat, with commensurate loss of function.

"two state transition"



Denaturation Leads to Loss of Protein Structure and Function高濃度鹽酸尿素造成蛋白質變性

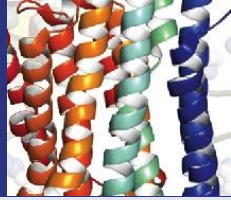
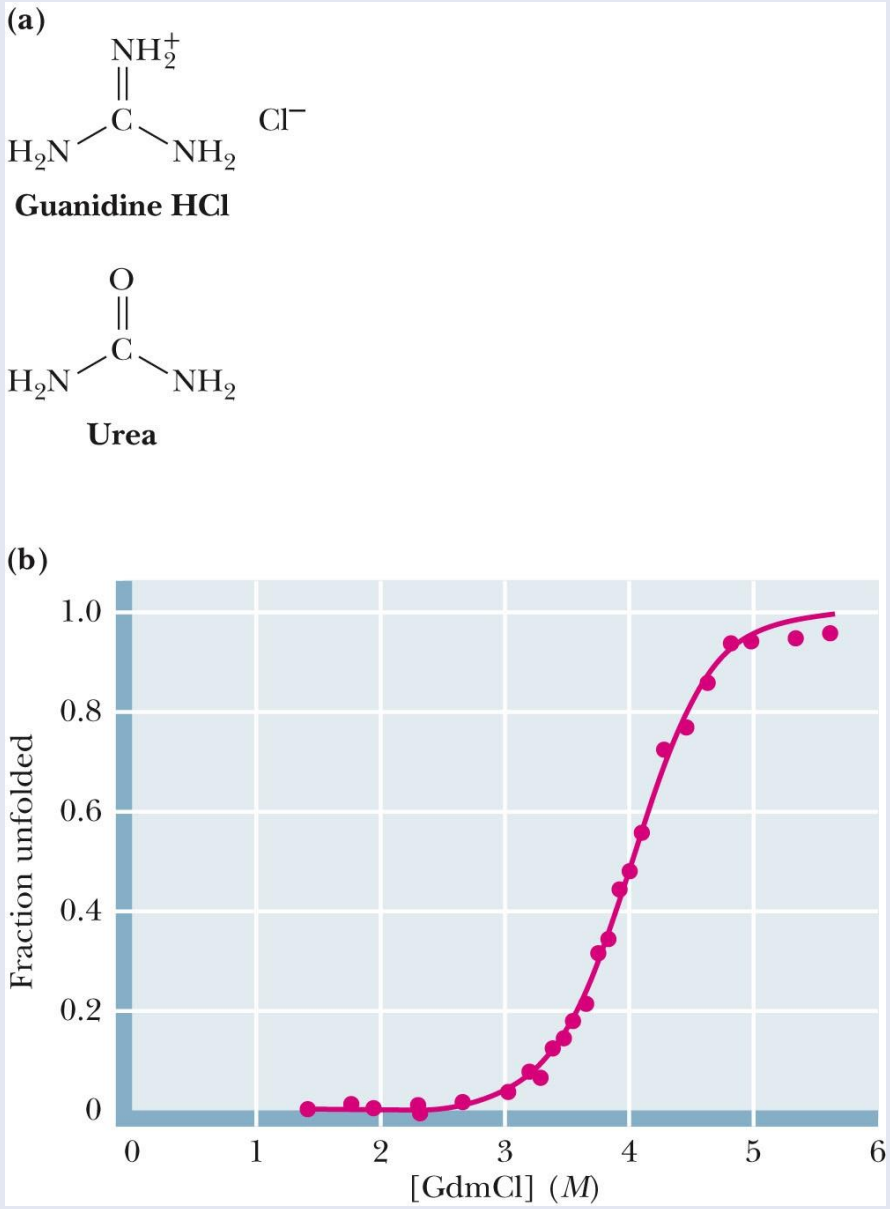


Figure 6.31 Proteins can be denatured (unfolded) by high concentrations of guanidine-HCl or urea.

The denaturation of chymotrypsin is plotted here.



Anfinsen's Classic Experiment Proved that Sequence Determines Structure

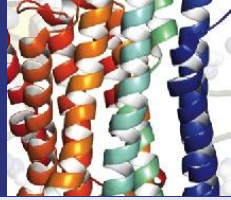
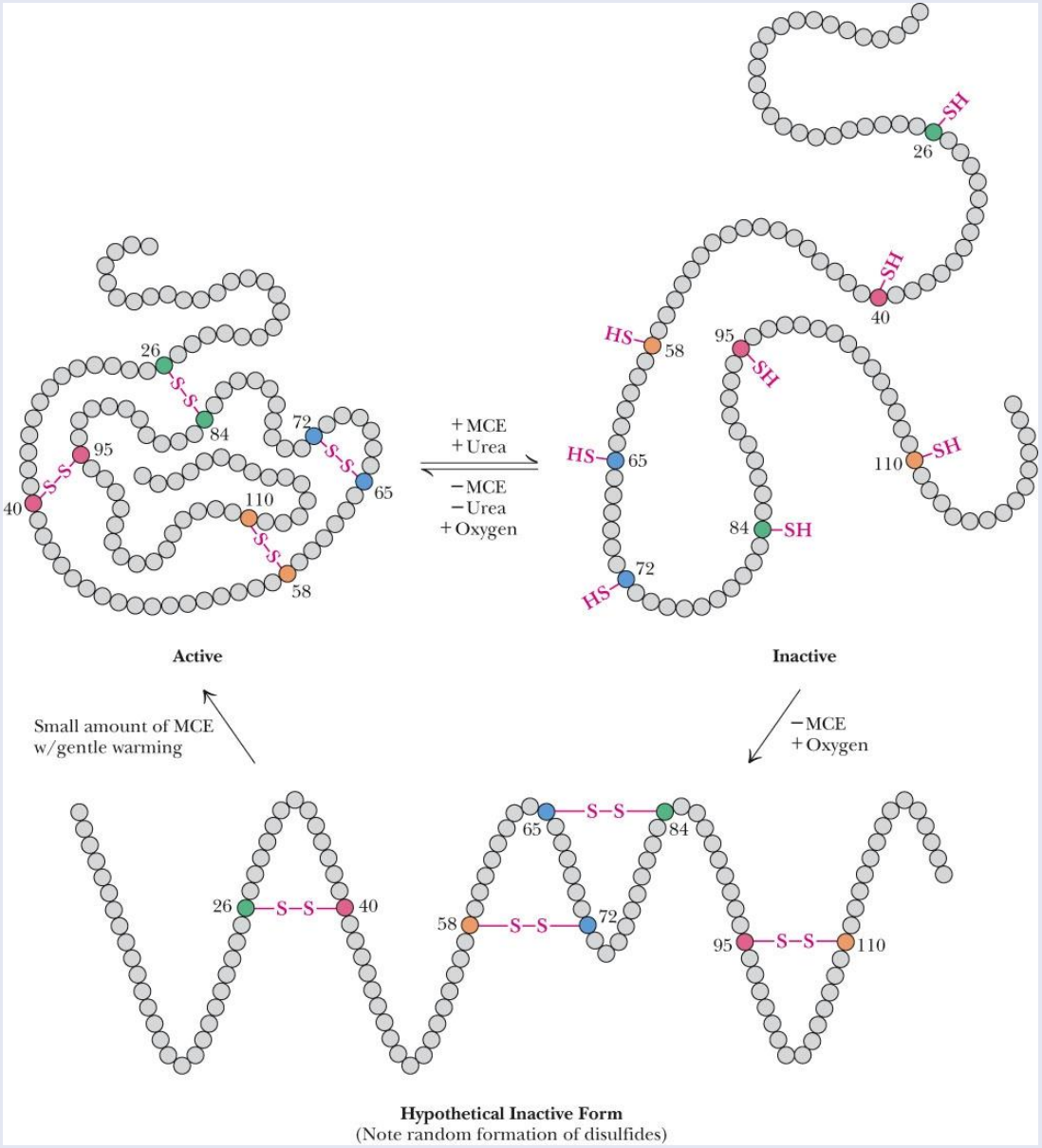
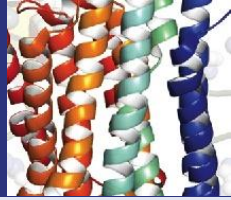


Figure 6.32 Ribonuclease can be unfolded by treatment with urea, and β-mercaptoethanol (MCE) cleaves disulfide bonds.

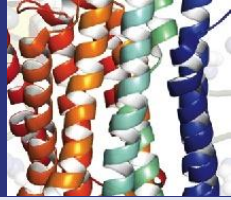
Anfinsen (1951, NIH) showed that ribonuclease structure (and function) could be restored under appropriate conditions.

"information required for folding of globular proteins is contained in the primary structure"



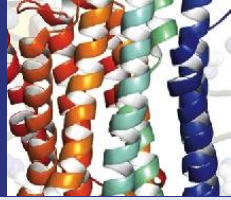


Postulated Themes of Protein Folding
蛋白質摺疊的假設論述



- Secondary structures – helices, sheets, and turns – probably form first
- Nonpolar residues may aggregate or coalesce in a process termed a **hydrophobic collapse**
- Subsequent steps probably involve formation of long-range interactions between secondary structures or involving other hydrophobic interactions
- The folding process may involve one or more intermediate states, including transition states and what have become known as **molten globules**

The Protein Folding Energy Landscape



Ken Dill has suggested that the folding process can be pictured as a funnel of free energies.
自由能的漏斗 –

摺疊而降低成安定型

The rim at the top represents the many unfolded states.

Polypeptides fall down the wall of the funnel to lower energies as they fold.

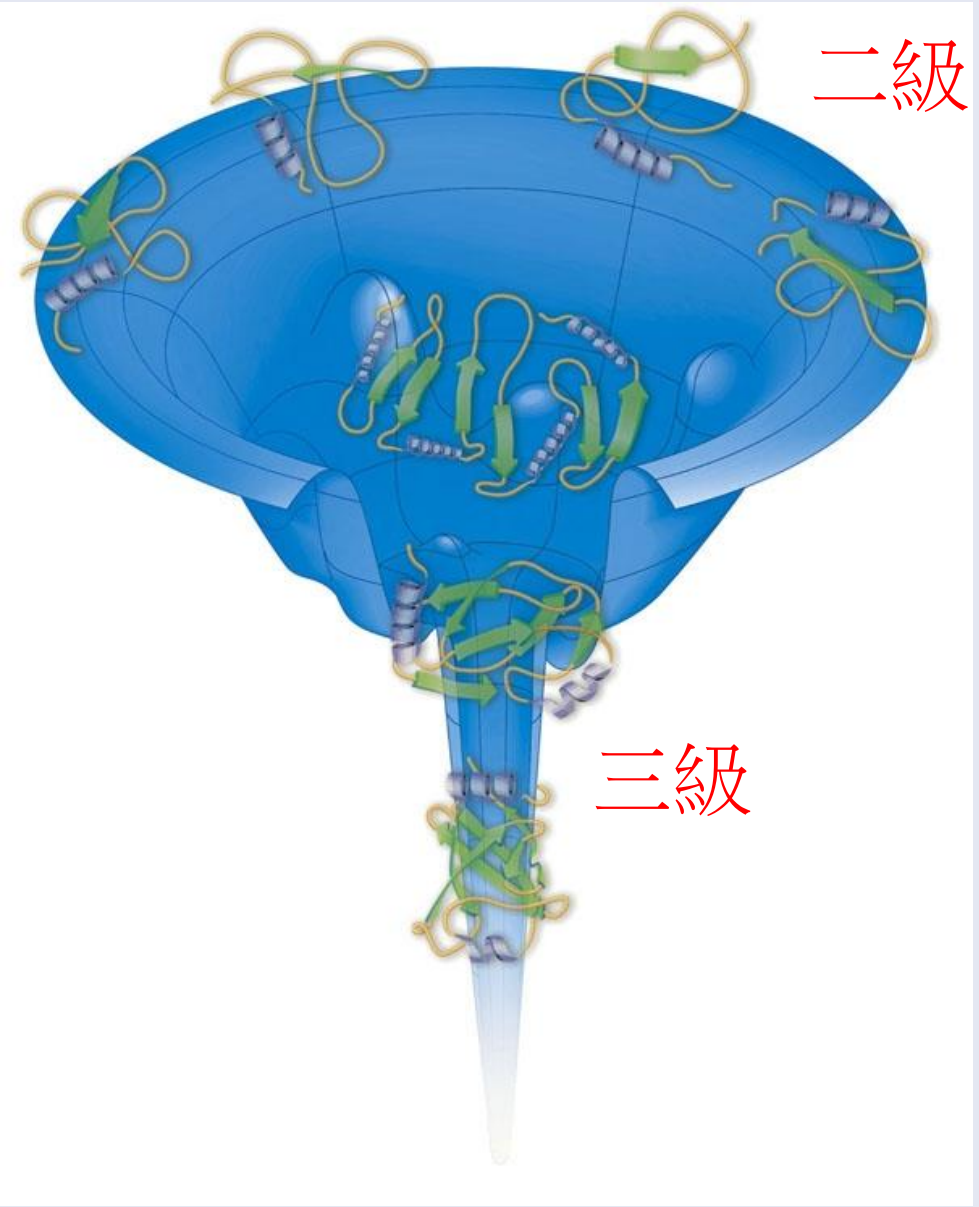
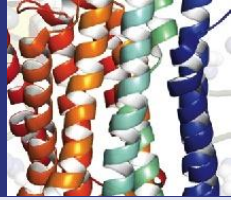


Figure 6.34 Folding model of globular proteins.

Folding is a thermodynamically favored process



$\Delta G = \Delta H - T\Delta S$,
(自由能) (enthalpy) (entropy)

Summing these quantities can make the total free energy of folding negative, so that the folding structure is stable.

折疊使亂度變小
由分子內鍵結的補償作用

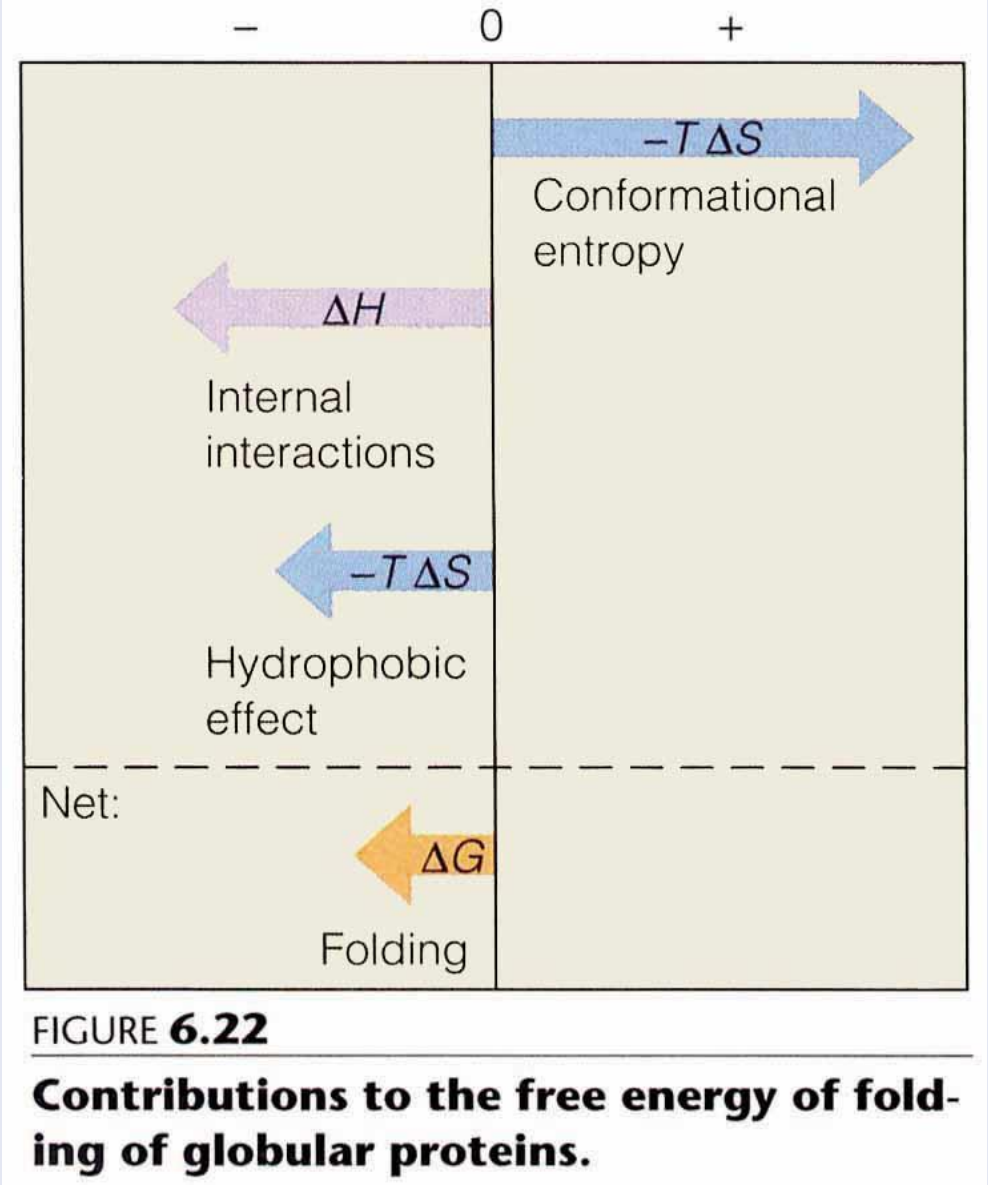
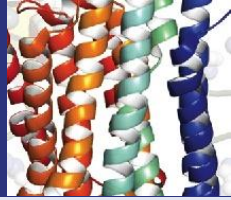


FIGURE **6.22**

Contributions to the free energy of folding of globular proteins.



- **(1) Internal Interactions** –分子內交互作用力 -ΔH
  - Energetically favorable interactions between groups within the folded molecule including charge-charge, internal hydrogen bonding, and van der Waals interactions, are the major source of the negative ΔH of folding.

- **(2) The Hydrophobic Effect** –疏水性聚集使亂度變小
  - Increase entropy by destroying the ordered structures of water around these residues in the unfolded state to become folded.

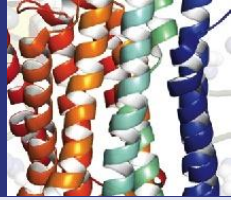


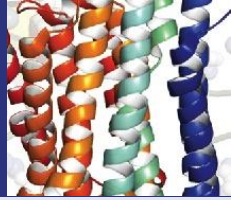
Table 6.3 Thermodynamic parameters for folding of some globular proteins at 25℃ in aqueous solution

Protein	ΔG (kJ/mol)	ΔH (kJ/mol)	ΔS (J/K·mol)
Ribonuclease	−46	−280	−790
Chymotrypsin	−55	−270	−720
Lysozyme	−62	−220	−530
Cytochrome *c*	−44	−52	−27
Myoglobin	−50	0	+170

Stabilized by hydrophobic effect

Note: Data adapted from P. L. Privalov and N. N. Khechinashvili, *J. Mol. Biol.* (1974) 86:665–684. Each data set has been taken at the pH value where the protein is maximally stable; all are near physiological pH. Data are for the folding reaction: Denatured $\rightleftharpoons$ native.

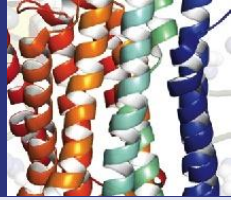
Marginal Stability of the Tertiary Structure Makes Proteins Flexible 邊緣性安定使蛋白質有彈性/運動



- A typical folded protein is only marginally stable
- It is logical to think that stability is important to function, so why are proteins often only marginally stable?
- The answer appears to lie in flexibility and motion
- It is becoming increasingly clear that flexibility and motion are important to protein function

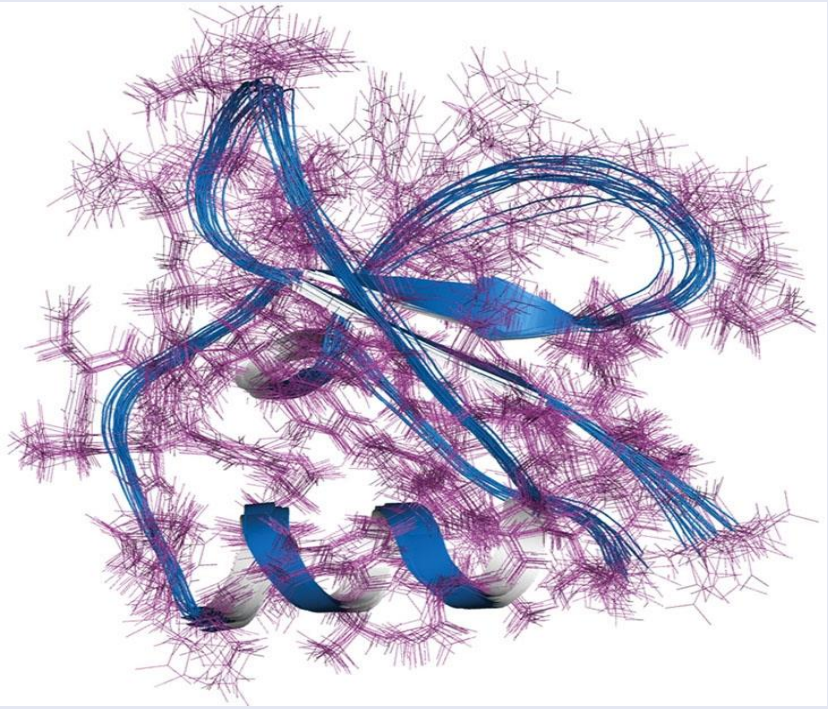
Motion is Important for Globular Proteins

蛋白質運動性 對球蛋白的功能性特別重要

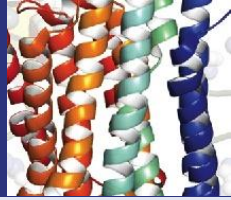


- Protein are dynamic structures – they oscillate (左右振動) and fluctuate (上下波動) continuously about their average or equilibrium structures (達到平均/平衡結構)

- This flexibility is essential for protein functions, including:
  - Ligand binding
  - Enzyme catalysis
  - Enzyme regulation



The Folding Tendencies and Patterns of Globular Proteins (球蛋白摺疊的 趨勢 與 款式)



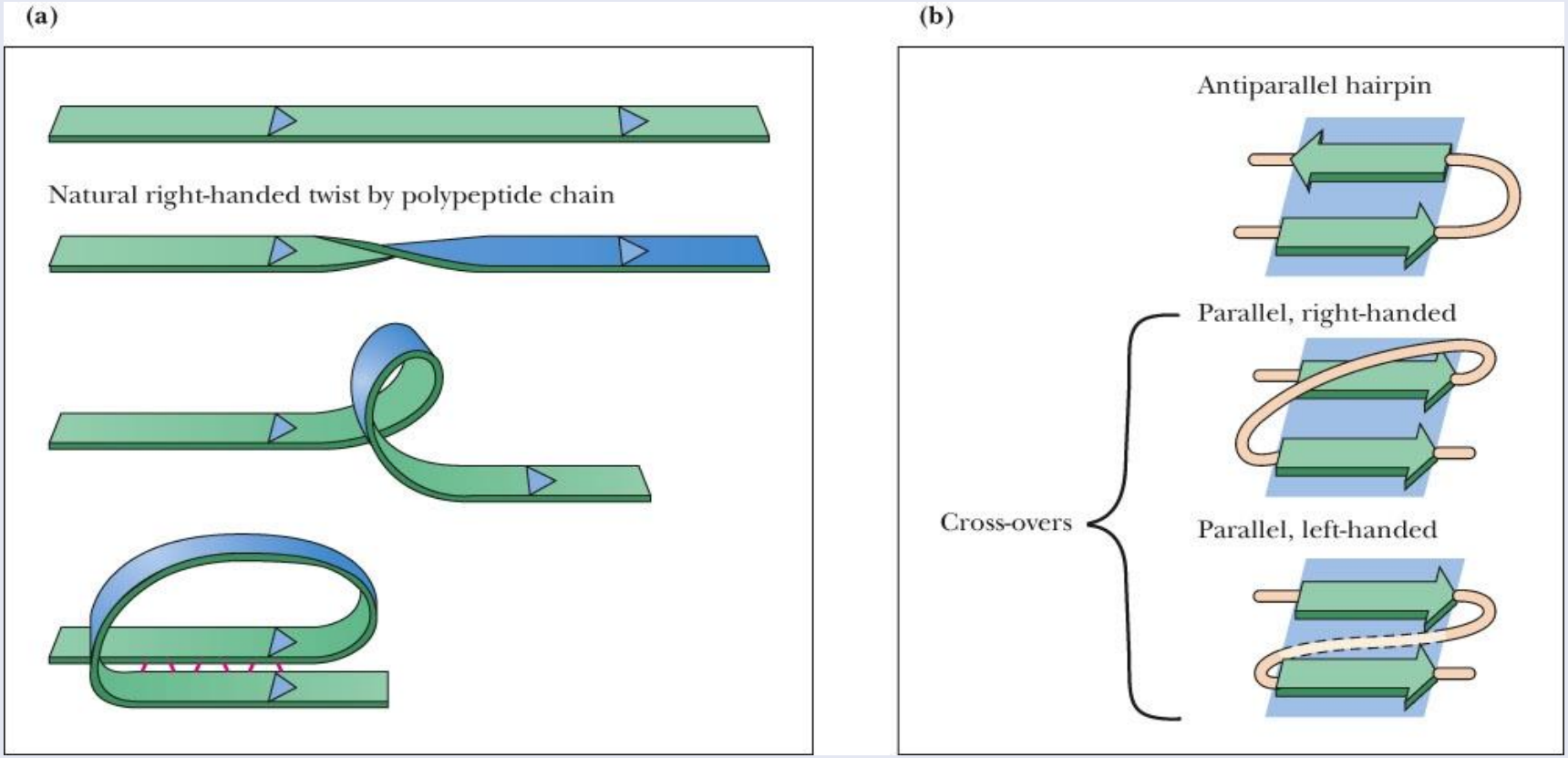


Figure 6.37 (a) The natural right-handed twist of polypeptide chains, and (b) the types of connections between β-strands.

Layer Structures in Globular Proteins 球蛋白的疊層

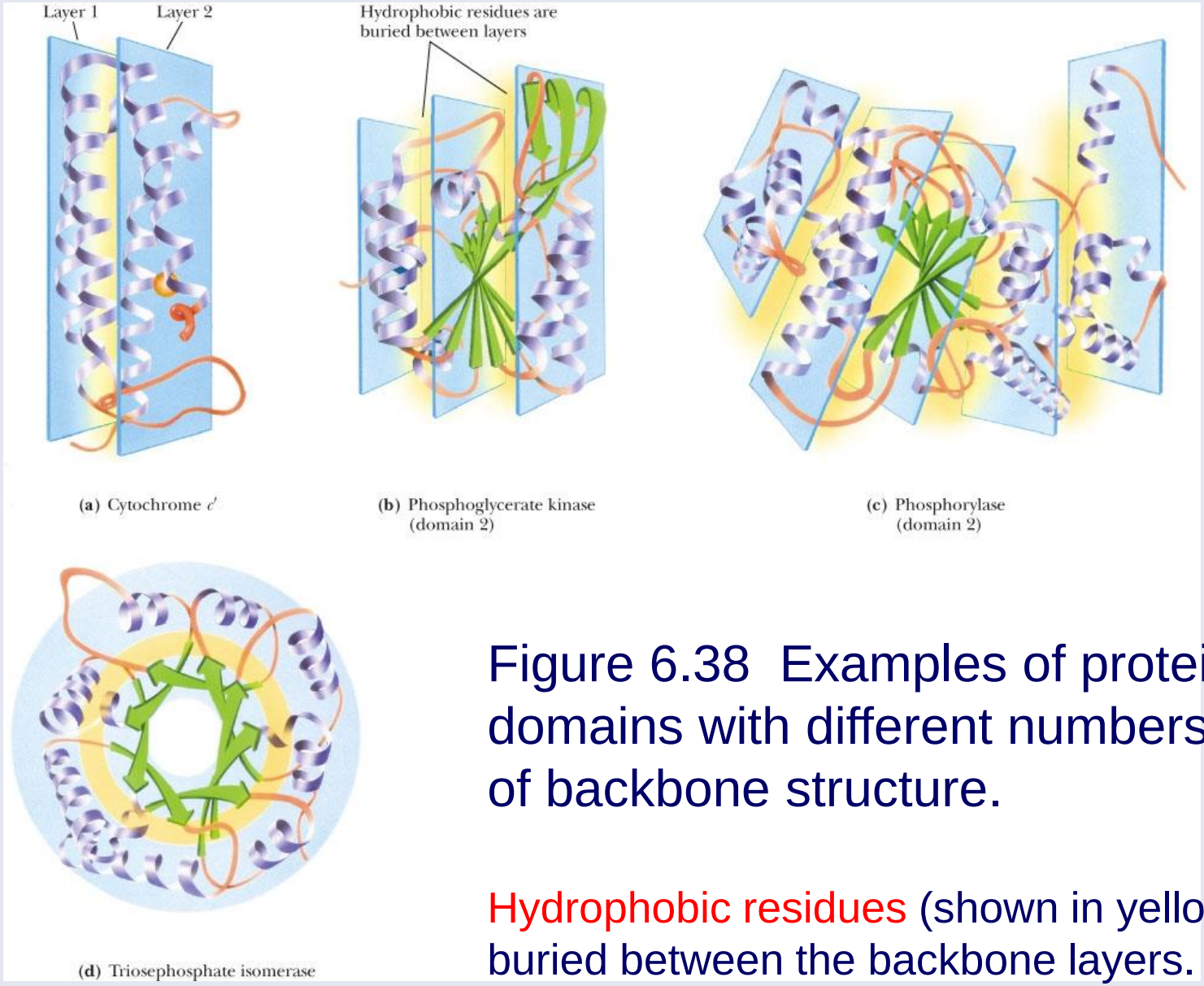
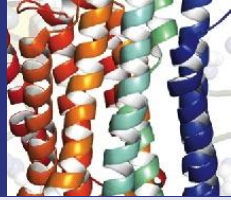


Figure 6.38 Examples of protein domains with different numbers of layers of backbone structure.

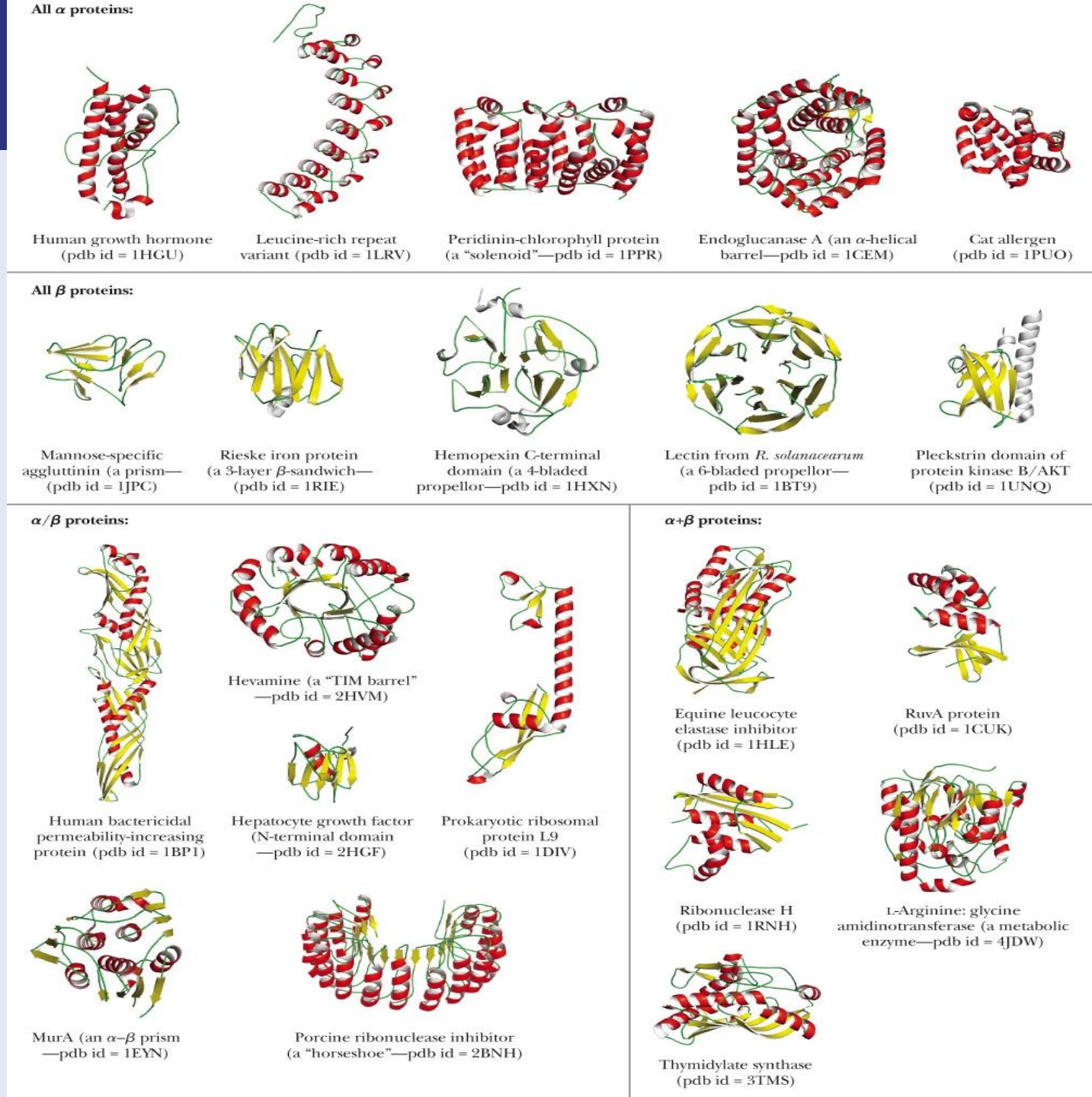
Hydrophobic residues (shown in yellow) are buried between the backbone layers.

Most Globular Proteins Belong to One of Four Structural Classes (多數球蛋白屬於四種二級結構排列的方式)

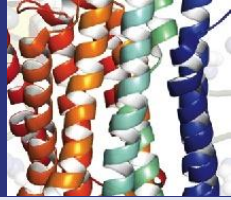


- Proteins can be classified according to the type and arrangement of secondary structure
- There are four classes:
  - All α proteins, in which α helices predominate
  - All β proteins, in which β sheets predominate
  - α/β proteins, in which helices and sheets are intermingled (交錯混雜)
  - α+β proteins, which contain separate α-helical and β-sheet domains (分開)

Figure 6.39 Four major classes of protein structure.



Molecular Chaperones Are Proteins That Help Other Proteins to Fold

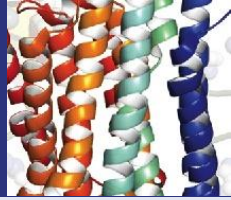


- Why are chaperones needed if the information for folding is inherent in the sequence?
  - to protect nascent proteins from the concentrated protein matrix in the cell and perhaps to accelerate slow steps

- Chaperone proteins were first identified as "heat-shock proteins" (hsp60 and hsp70)

  - 第一群鑑定出的伴護蛋白是熱休克蛋白

Some Proteins Are Intrinsically Unstructured

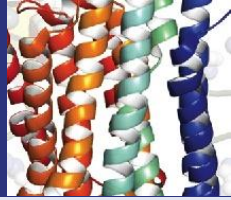


- Many proteins exist and function normally in a partially unfolded state

- These **intrinsically unstructured proteins (IUPs)** do not possess uniform structural properties but are still essential for cellular function 不具規則性的結構, 很多極性殘基, 缺疏水基

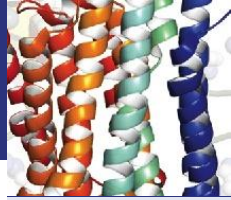
- These proteins are characterized by a nearly complete lack of structure and high intramolecular flexibility
- IUPs adopt well-defined structures in complexes with their target proteins
- IUPs are characterized by an abundance of polar residues and a lack of hydrophobic residues

Diseases of Protein Folding
蛋白質摺疊 的疾病 (組成->結構->功能)



Disease	Affected Protein	Mechanism
Alzheimer's disease	β-Amyloid peptide (derived from amyloid precursor protein)	Misfolded β-amyloid peptide accumulates in human neural tissue, forming deposits known as neuritic plaques.
Familial amyloidotic polyneuropathy	Transthyretin	Aggregation of unfolded proteins. Nerves and other organs are damaged by deposits of insoluble protein products.
Cancer	p53	p53 prevents cells with damaged DNA from dividing. One class of p53 mutations leads to misfolding; the misfolded protein is unstable and is destroyed.
Creutzfeldt-Jakob disease (human equivalent of mad cow disease)	Prion	Prion protein with an altered conformation (PrP^{SC}) may seed conformational transitions in normal PrP (PrP^{C}) molecules.
Hereditary emphysema	α_1-Antitrypsin	Mutated forms of this protein fold slowly, allowing its target, elastase, to destroy lung tissue.
Cystic fibrosis	CFTR (cystic fibrosis transmembrane conductance regulator)	Folding intermediates of mutant CFTR forms don't dissociate freely from chaperones, preventing the CFTR from reaching its destination in the membrane.

6.5 The subunit compositions of several proteins.



Proteins with two or four subunits predominate in nature, and many cases of higher numbers exist.

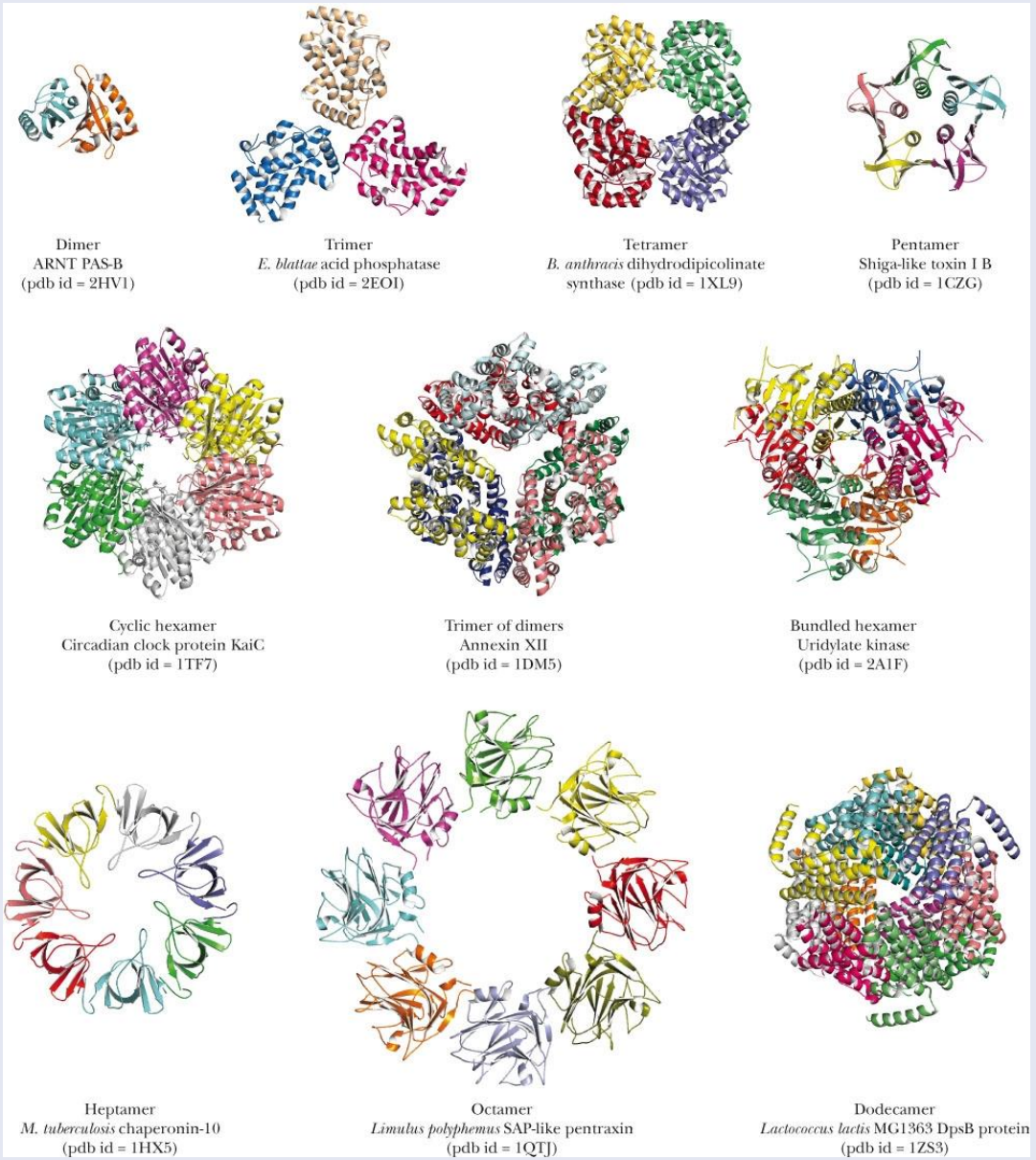
TABLE 6.3 Aggregation Symmetries of Globular Proteins

Protein	Number of Subunits
Alcohol dehydrogenase	2
Malate dehydrogenase	2
Superoxide dismutase	2
Triose phosphate isomerase	2
Glycogen phosphorylase	2
Aldolase	3
Bacteriochlorophyll protein	3
Concanavalin A	4
Glyceraldehyde-3-phosphate dehydrogenase	4
Immunoglobulin	4
Lactate dehydrogenase	4
Prealbumin	4
Pyruvate kinase	4
Phosphoglycerate mutase	4
Hemoglobin	2 + 2
Insulin	6
Aspartate transcarbamoylase	6 + 6
Glutamine synthetase	12
TMV protein disc	17
Apoferritin	24
Coat of tomato bushy stunt virus	180

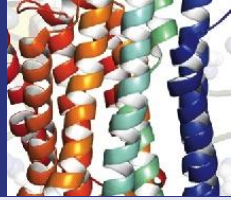
6.5 How Do Protein Subunits Interact at the Quaternary Level of Structure?

Figure 6.44 Multimeric proteins are symmetric arrangements of asymmetric objects.

A variety of symmetries is displayed in these multimeric structures.

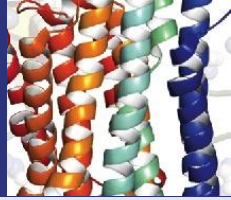


What are the structural and functional advantages driving quaternary association?

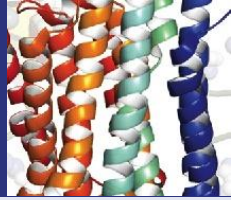


- Stability: reduction of surface to volume ratio
  - shield hydrophobic residues from solvent water
- Genetic economy and efficiency
  - less DNS is required to code for a monomer
- Bringing catalytic sites together
  - associated subunit in an enzyme is active, dissociated monomers are inactive
- Cooperativity
  -

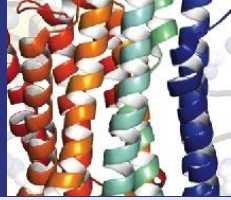
Quaternary structure-四級結構



- multiple protein subunits interact with each other.
  - **Homotypic** quaternary structure occurs when identical or nearly identical polypeptide chains interact.
  - **Heterotypic** quaternary structure occurs when subunits of very different structures interact.
- Interactions between polypeptide subunits that stabilize the multisubunit structure are the same as the ones stabilizing tertiary structure –
  - salt bridges, hydrogen bonds, van der Waals forces, hydrophobic interactions and (occasionally) disulfide bonds.



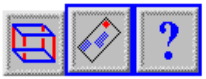
(5) Classification Schemes for the Protein Universe Are Based on Domains

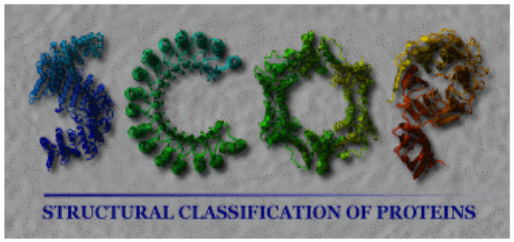


- Several comprehensive projects have organized the available information on protein domains into defined hierarchies or levels of protein structure.
- The Structural Classification of Proteins (SCOP) database recognizes five overarching classes
- SCOP is based on levels that embody the evolutionary and structural relationships among known proteins

  http://scop.mrc-lmb.cam.ac.uk/scop/
- CATH (standing for Class, Architecture, Topology, Homologous superfamily) is another system

  http://www.cathdb.info/
- CATH differs from SCOP in combining manual analysis with quantitative algorithmic analysis





Welcome to **SCOP**: Structural Classification of Proteins.
1.75 release (June 2009)

38221 PDB Entries. 1 Literature Reference. 110800 Domains. (excluding nucleic acids and theoretical models).
Folds, superfamilies, and families statistics here.
New folds superfamilies families.
List of obsolete entries and their replacements.

Authors. Alexey G. Murzin, John-Marc Chandonia, Antonina Andreeva, Dave Howorth, Loredana Lo Conte, Bartlett G. Ailey, Steven E. Brenner, Tim J. P. Hubbard, and Cyrus Chothia. scop@mrc-lmb.cam.ac.uk
Reference: Murzin A. G., Brenner S. E., Hubbard T., Chothia C. (1995). SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247, 536-540. [PDF]
Recent changes are described in: Lo Conte L., Brenner S. E., Hubbard T.J.P., Chothia C., Murzin A. (2002). SCOP database in 2002: refinements accommodate structural genomics. *Nucl. Acid Res.* 30(1), 264-267. [PDF],
Andreeva A., Howorth D., Brenner S.E., Hubbard T.J.P., Chothia C., Murzin A.G. (2004). SCOP database in 2004: refinements integrate structure and sequence family data. *Nucl. Acid Res.* 32:D226-D229. [PDF], and
Andreeva A., Howorth D., Chandonia J.-M., Brenner S.E., Hubbard T.J.P., Chothia C., Murzin A.G. (2007). Data growth and its impact on the SCOP database: new developments. *Nucl. Acids Res.* 2008 36: D419-D425; doi:10.1093/nar/gkm993 [PDF].

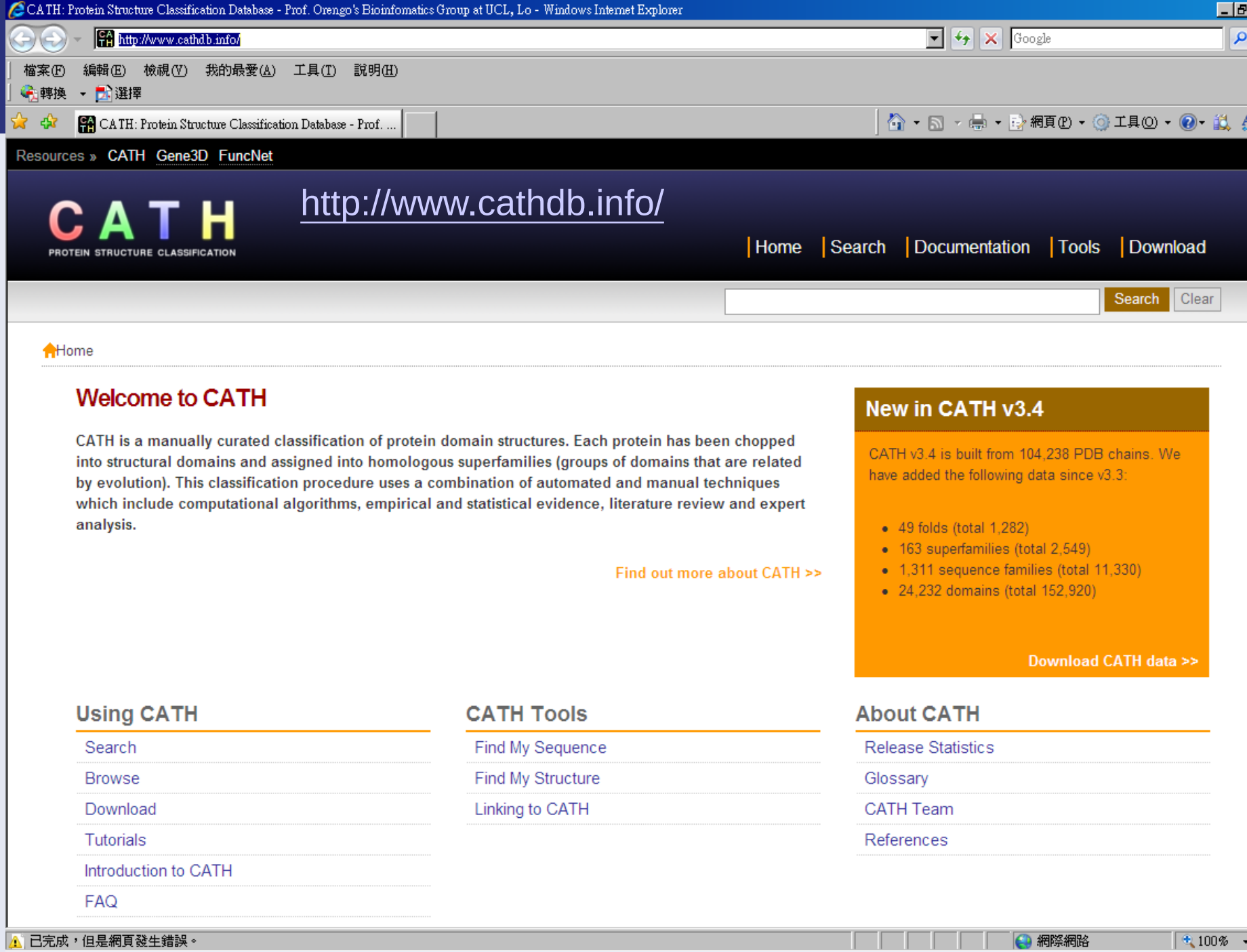
Access methods

- Enter SCOP at the **top of the hierarchy**
- Keyword search of SCOP entries
- SCOP parseable files
- All SCOP releases and reclassified entry history
- **pre-SCOP - preview of the next release**
- SCOP domain sequences and pdb-style coordinate files (ASTRAL)
- Hidden Markov Model library for SCOP superfamilies (SUPERFAMILY)
- Structural alignments for proteins with non-trivial relationships (SISYPHUS)

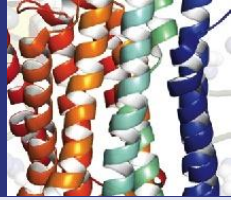
- Online resources of potential interest to SCOP users

SCOP **mirrors** around the world may speed your access.

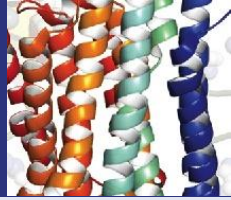
- http://scop.mrc-lmb.cam.ac.uk/scop/


CATH: Protein Structure Classification Database - Prof. Orengo's Bioinformatics Group at UCL, Lo - Windows Internet Explorer
http://www.cathdb.info/
Google
檔案(F) 編輯(E) 檢視(V) 我的最愛(A) 工具(T) 說明(H)
轉換 選擇
CATH: Protein Structure Classification Database - Prof. ...
網頁(P) 工具(O)
Resources » CATH Gene3D FuncNet
CATH
PROTEIN STRUCTURE CLASSIFICATION
http://www.cathdb.info/
Home | Search | Documentation | Tools | Download
Search
Clear
Home
Welcome to CATH
CATH is a manually curated classification of protein domain structures. Each protein has been chopped into structural domains and assigned into homologous superfamilies (groups of domains that are related by evolution). This classification procedure uses a combination of automated and manual techniques which include computational algorithms, empirical and statistical evidence, literature review and expert analysis.
Find out more about CATH >>
New in CATH v3.4
CATH v3.4 is built from 104,238 PDB chains. We have added the following data since v3.3:
49 folds (total 1,282)
163 superfamilies (total 2,549)
1,311 sequence families (total 11,330)
24,232 domains (total 152,920)
Download CATH data >>
Using CATH
Search
Browse
Download
Tutorials
Introduction to CATH
FAQ
CATH Tools
Find My Sequence
Find My Structure
Linking to CATH
About CATH
Release Statistics
Glossary
CATH Team
References
已完成，但是網頁發生錯誤。
網際網路
100%

Classification Schemes for the Protein Universe Are Based on Domains



- Common features of SCOP and CATH:
- **Class** is determined from overall composition of secondary structure elements in a domain
- **Fold** describes the number, arrangement, and connections of these secondary structure elements
- **Superfamily** includes domains of similar folds and usually similar functions
- **Family** usually includes domains with closely related amino acid sequences



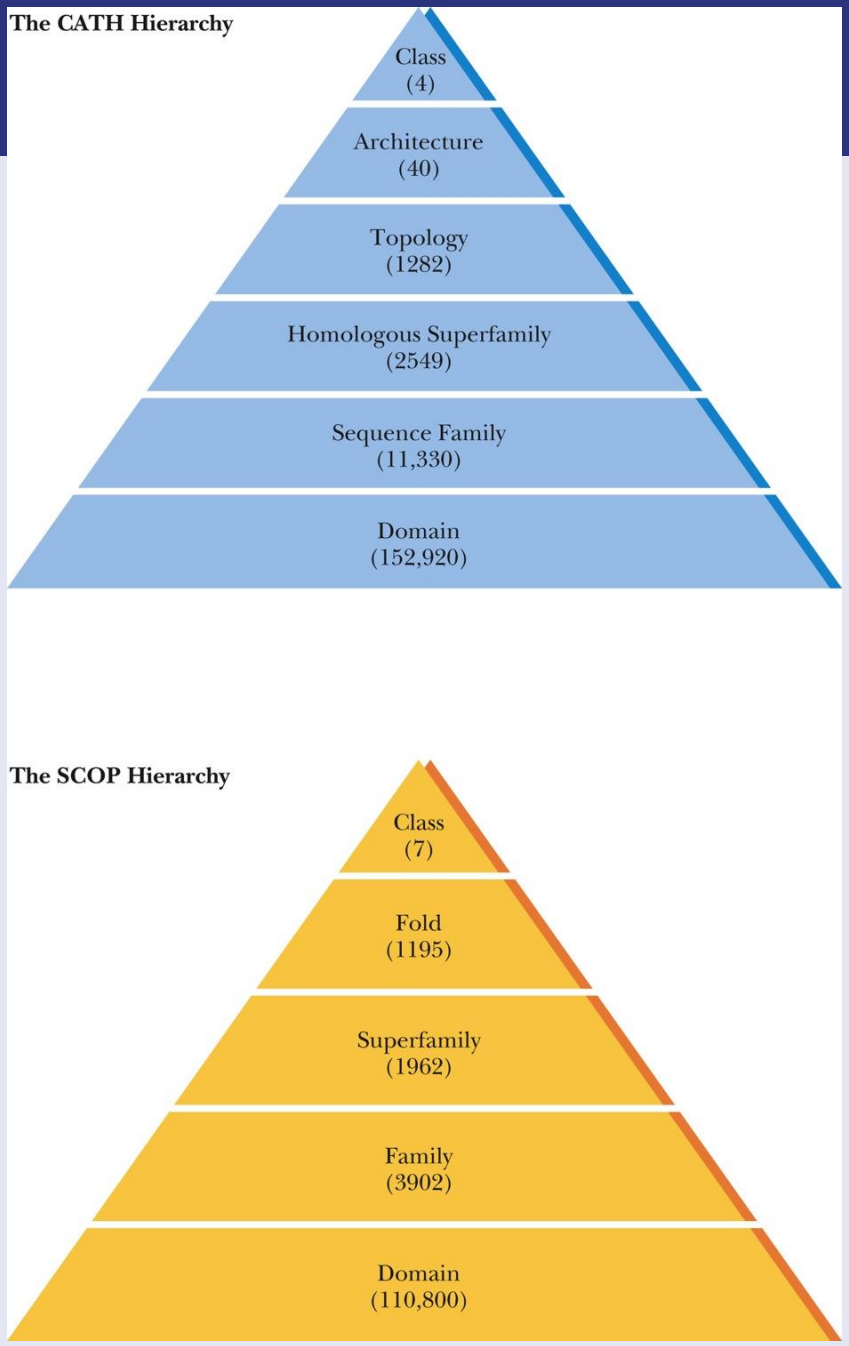


Fig. 6.27. Hierachical classification systems for known proteins.