LECTURE 06: FIBROUS PROTEINS

Fibrous Proteins Are Adapted for a Structural Function

α-Keratin, collagen, and elastin provide clear examples of the relationship between protein structure and biological function (Table 1).

Table 1 **Secondary Structures and Properties of Some Fibrous Proteins**

Structure	Characteristics	Examples of occurrence
α Helix, cross-linked by disulfide bonds	Tough, insoluble protective structures of varying hardness and flexibility	α-Keratin of hair, feathers, nails
β Conformation	Soft, flexible filaments	Silk fibroin
Collagen triple helix	High tensile strength, without stretch	Collagen of tendons, bone matrix

These proteins share properties that give strength and/or elasticity to structures in which they occur. They have relatively simple structures, and all are insoluble in water, a property conferred by a high concentration of hydrophobic amino acids both in the interior of the protein and on the surface. These proteins represent an exception to the rule that hydrophobic groups must be buried. The hydrophobic core of the molecule therefore contributes less to structural stability, and covalent bonds assume an especially important role.

α-Keratin and **collagen** have evolved for strength.

In vertebrates, α-keratins constitute almost the entire dry weight of hair, wool, feathers, nails, claws, quills, scales, horns, hooves, tortoise shell, and much of the outer layer of skin.

Collagen is found in connective tissue such as tendons, cartilage, the organic matrix of bones, and the cornea of the eye.

The polypeptide chains of both proteins have simple helical structures. The α-keratin helix is the right-handed α helix found in many other proteins (Fig. 2). However, the collagen helix is unique. It is left-handed (see Fig 3) and has three amino acid residues per turn (Fig. 3).

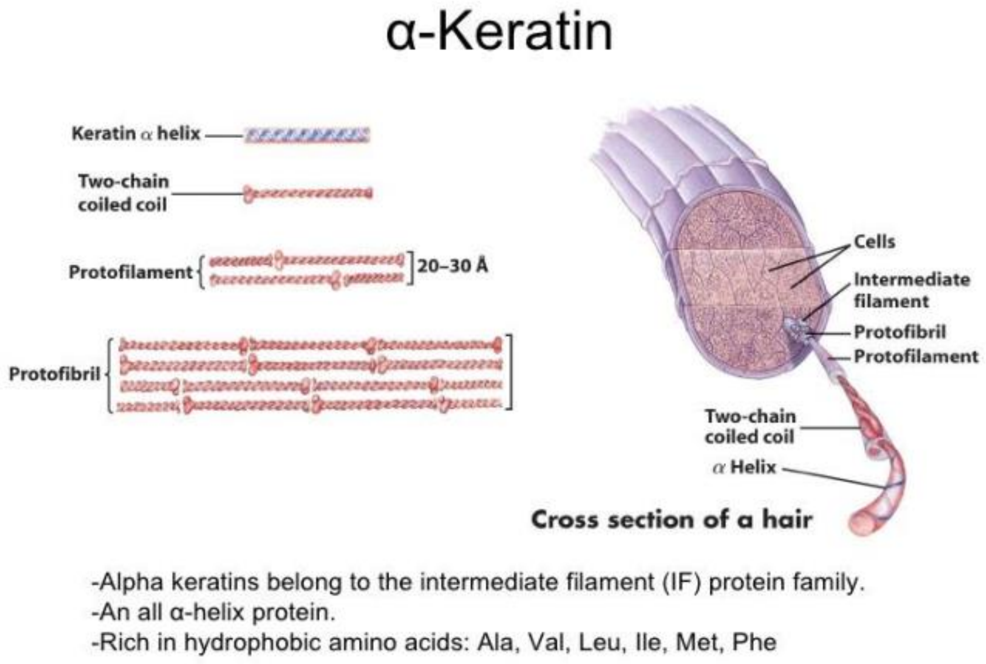
In both α-keratin and collagen, a few amino acids predominate.

α-Keratin is rich in the hydrophobic residues Phe, Ile, Val, Met, and Ala.

Collagen is 35% Gly, 11% Ala, and 21% Pro and Hyp (hydroxyproline; see Fig. 3). The unusual amino acid content of collagen is imposed by structural constraints unique to the collagen helix. The amino acid sequence in collagen is generally a repeating tripeptide unit, Gly-X-Pro or Gly-X-Hyp, where X can be any amino acid. The food product gelatin is derived from collagen. Although it is protein, it has little nutritional value because collagen lacks significant amounts of many amino acids that are essential in the human diet.

In both α-keratin and collagen, strength is amplified by wrapping multiple helical strands together in a superhelix, much the way strings are twisted to make a strong rope (Figs. 2,3). In both proteins the helical path of the supertwists is opposite in sense to the twisting of the individual polypeptide helices, a conformation that permits the closest possible packing of the multiple polypeptide chains.

The superhelical twisting is probably left-handed in α-keratin (Fig.2) and right-handed in collagen (Fig.3). The tight wrapping of the collagen triple helix provides great tensile strength with no capacity to stretch: Collagen fibers can support up to 10,000 times their own weight and are said to have greater tensile strength than a steel wire of equal cross section.



-Alpha keratins belong to the intermediate filament (IF) protein family.
-An all α-helix protein.
-Rich in hydrophobic amino acids: Ala, Val, Leu, Ile, Met, Phe

Fig 2 α-keratin

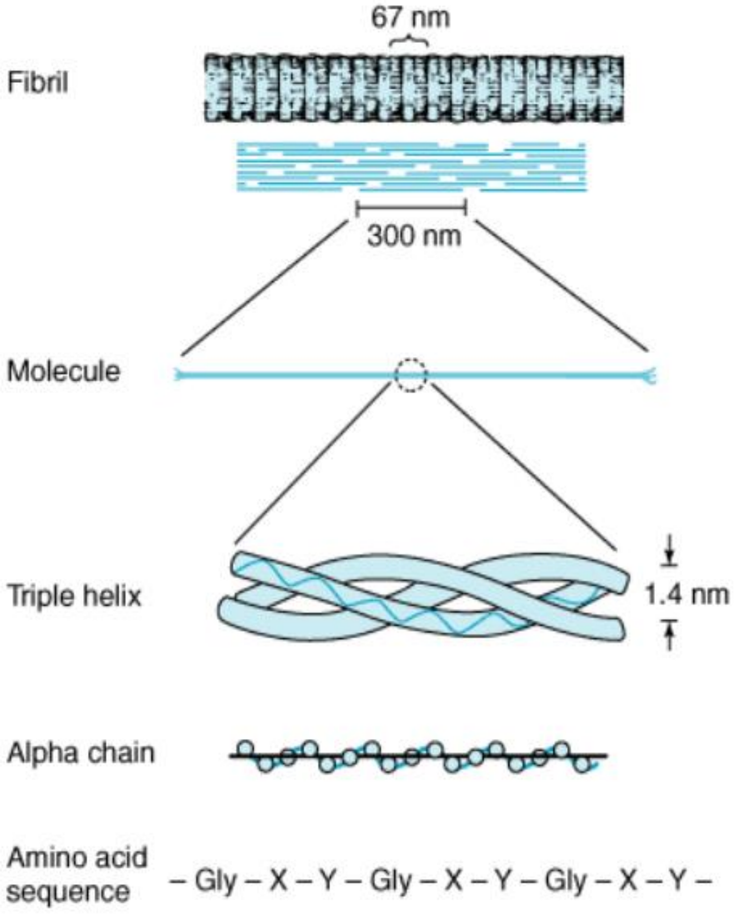


Fig 3 collagen

The fibroin protein consists of layers of antiparallel beta sheets (Fig 4). Its primary structure mainly consists of the recurrent amino acid sequence (Gly-Ser-Gly-Ala-Gly-Ala)$_n$. The high glycine (and, to a lesser extent, alanine) content allows for tight packing of the sheets, which contributes to silk's rigid structure and tensile strength. A combination of stiffness and toughness make it a material with applications in several areas, including biomedicine and textile manufacture.

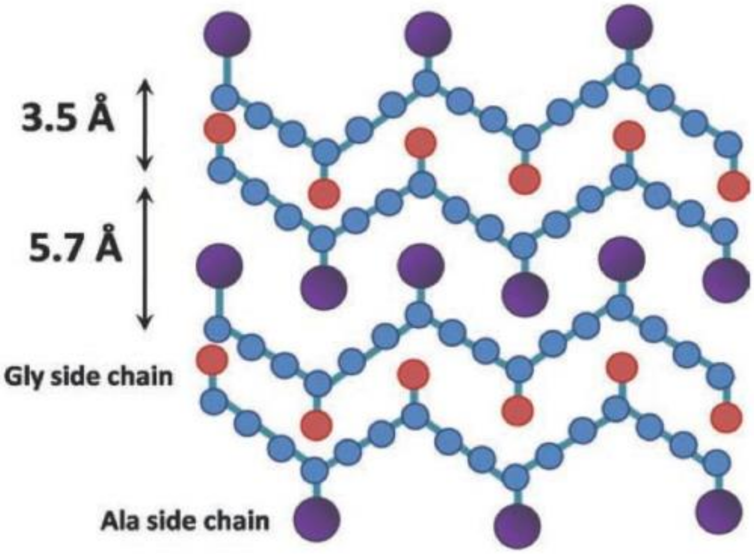
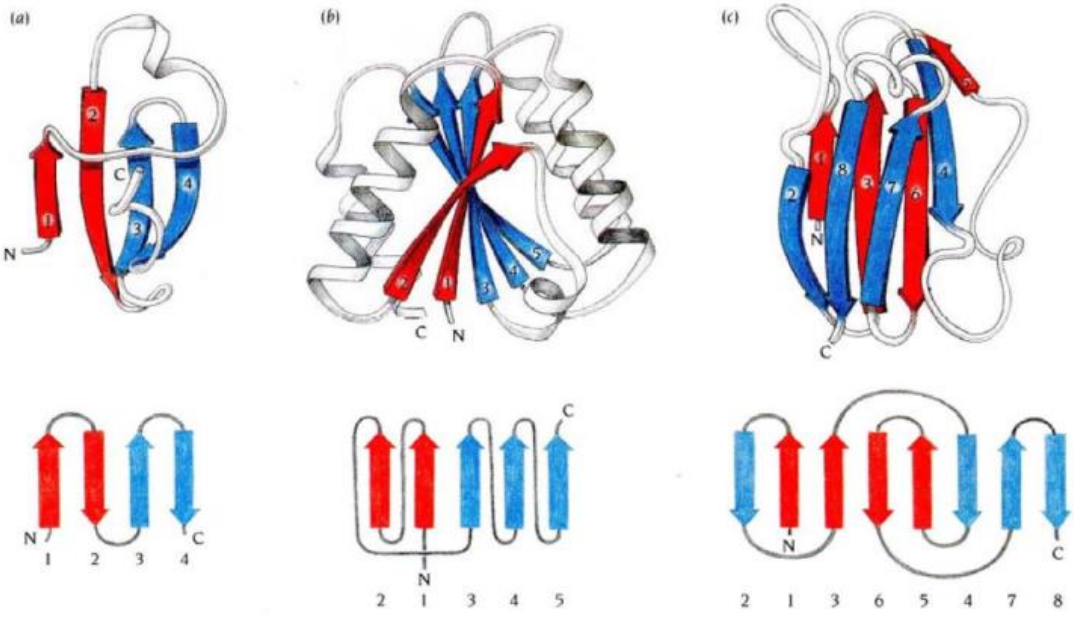


Fig 4 Silk Fibroin

Topology diagrams

The most characteristic features of a β sheet are the number of strands, their relative directions (parallel or antiparallel), and how the strands are connected. This information can be represented by topology diagrams. They are useful to compare β structures.



Protein Tertiary Structure

Tertiary structure refers to the three-dimensional arrangement of all atoms in a protein. Tertiary structure is formed by the folding in three dimensions of the secondary structure elements of a protein. While the α helical secondary structure is held together by interactions between the carbonyl and amide groups within the backbone, tertiary structure is held together by interactions between R-groups of residues brought together by folding. Disulfide bonds are also counted under the category of tertiary structure interactions. Proteins that are compact are known as globular proteins.

Examination of protein structures resolved by X-ray diffraction and NMR has revealed a variety of folding patterns common to many different proteins. However, even within these folds, distinct substructures or structural **motifs**, i.e. distinctive arrangements of elements of secondary structure, have been described. The term **supersecondary structure** has been coined to describe this level of organisation, which is intermediate between secondary and tertiary.

Motifs or folds, are particularly stable arrangements of several elements of the secondary structure. • Supersecondary structures are usually produced by packing side chains from adjacent secondary structural elements close to each other.

Rules for secondary structure.

- Hydrophobic side groups must be buried inside the folds, therefore, layers must be created (β−α−β; α− α).
- α-helix and β-sheet, if occur together, are found in different structural layers.
- Adjacent polypeptide segments are stacked together.
- Connections between secondary structures do not form knots.
- The β-sheet is the most stable.

Motif

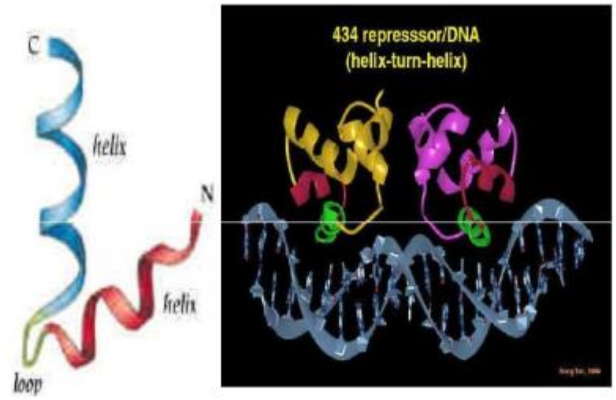
- Secondary structure composition, e.g. all α, all β, segregated α+β, mixed α/β
- Motif = small, specific combinations of secondary structure elements, e.g. β-α-β loop

PROTEIN ENGINEERING

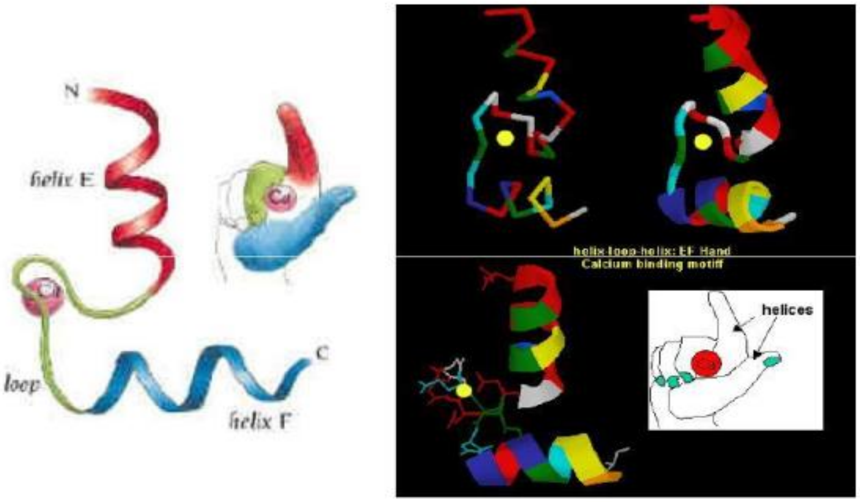
1. Helix super secondary structures

Helix-Turn-Helix Motif

Also called the alpha-alpha type (αα-type). The motif is compromised of two antiparallel helices connected by a turn. The helix-turn-helix is a functional motif and is usually identified in proteins that bind to DNA minor and major grooves, and Calcium-binding proteins.

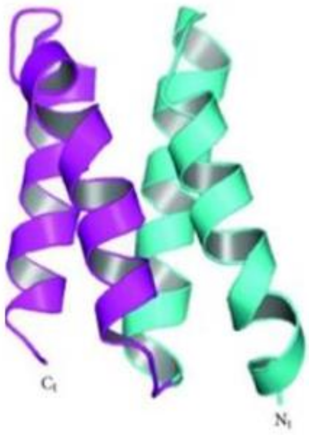


DNA binding Helix-turn-Helix motif



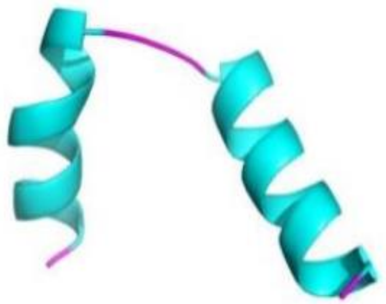
Calcium binding (EF Hand- Calcium binding) motif

Helix-hairpin-helix: Involved in DNA binding



Alpha-alpha corner

Short loop regions connecting helices which are roughly perpendicular to one another



2. Sheet super secondary structures

All beta tertiary structural domains can occur in proteins with one domain (eg. concanavalin A, superoxide dismutase), and occurs at least once in proteins with two domains (eg. chymotrypsin), or three domains (eg. OmpF).

The beta strands making up these domains are all essentially antiparallel and form structures to achieve stable packing arrangements within the protein.

There are presently (as of version 1.39) about 70 subclasses listed in SCOP for this domain, and some examples of these are outlined below.

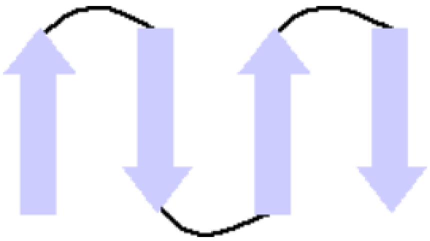
Beta barrels

This is the most abundant beta-domain structure and as the name suggests the domain forms a 'barrel-like' structure. The beta barrels are not geometrically perfect and can be rather distorted.

There are three main types:

1. Up-and-down barrels
2. Greek key barrels
3. Jelly roll (Swiss roll) barrels

Up-and-down beta-sheets or beta-barrels



The simple topology of an up-and-down barrel (named because the beta strands follow each other in sequence in an up-and-down fashion).

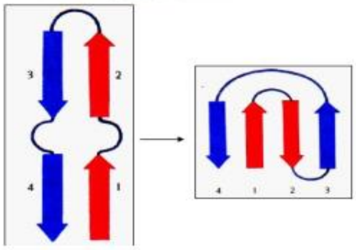
Usually, the loops joining the beta strands do not crossover the 'ends' of the barrel.

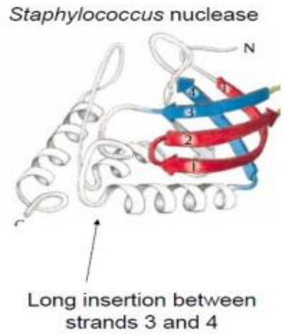
Greek key barrels

These are barrels formed from two, or more, Greek Key motifs.

It is a stable structure

The Greek key barrel consists of four anti-parallel Beta strands where one strand changes the topology direction. Hydrogen bonding occurs between strands 1:4, and strands 2:3. Strand 2 then folds over to form the structural motif.





Jelly roll barrels

These barrels are formed from a 'Greek Key-like' structure called a jelly roll. Supposedly named because the polypeptide chain is wrapped around a barrel core like a jelly roll (swiss roll).

It is a stable structure

This structure is found in coat proteins of spherical viruses, plant lectin concanavalin A, and hemagglutinin protein from influenza virus.

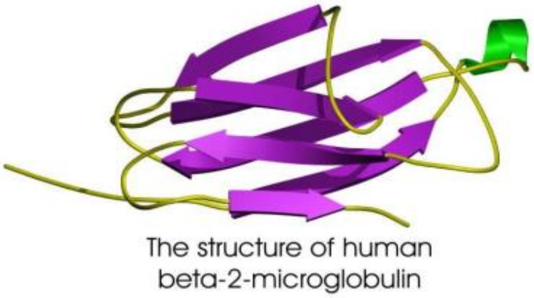
The essential features of a jelly roll barrel are that:

- it is like an inverted 'U' (which is often seen twisted and distorted in proteins)
- it is usually divided into two beta sheets which are packed against each other
- most jelly roll barrels have eight strands although any even number greater than 8 can form a jelly roll barrel
- it folds such that hydrogen bonds exist between strands 1 and 8; 2 and 7; 3 and 6; and 4 and 5

©1999 GARLAND PUBLISHING INC
A member of the Taylor & Francis Group

Beta sandwich

A beta sandwich is essentially a 'flattened' beta barrel with the two sheets packing closely together (like a sandwich!). The first and last strands of the sandwich do not hydrogen bond to each other to complete a 'barrel' structure.

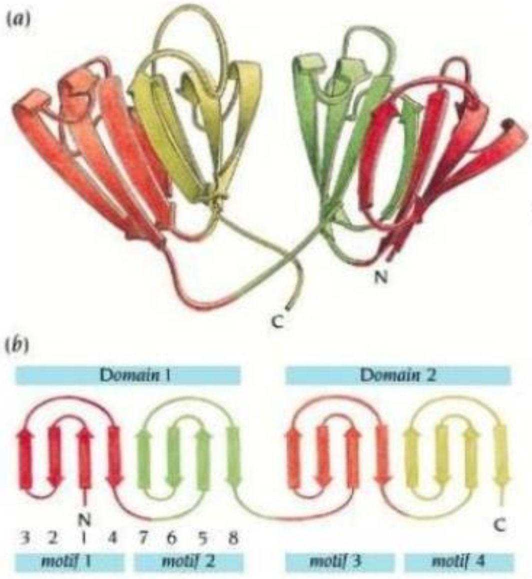


Beta sandwich in beta 2 microglobulin.

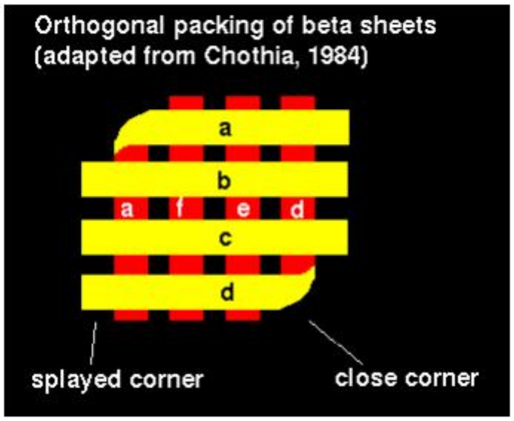
Aligned or Orthogonal beta strands

Beta strands in barrels or sandwich structures can be orientated in two general ways:

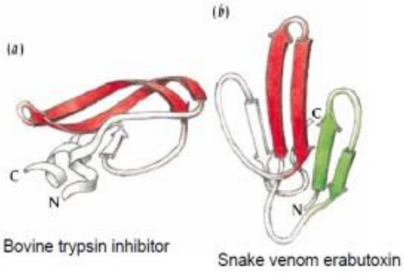
- where the strands in two sheets are almost aligned, and in the same orientation, to each other and form an 'aligned beta' structure (eg. gamma crystallin)



- where the strands, in at least two sheets, are roughly perpendicular to each other and form an 'orthogonal beta' structure.



Beta-hairpin: two antiparallel beta strands connected by a "hairpin" bend, i.e. beta-turn 2 x antiparallel beta-strands + beta-turn = beta hairpin



Bovine trypsin inhibitor

Snake venom erabutoxin

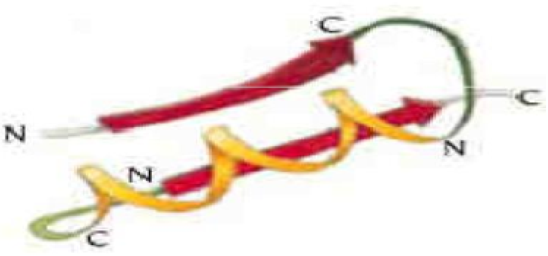
Beta-beta corner

- Two antiparallel beta strands which form a beta hairpin can change direction abruptly. The angle of the change of direction is about 90 degrees and so the structure is known as a 'beta corner'
- The abrupt angle change is achieved by one strand having a glycine residue (so there is no steric hindrance from a side chain) and the other strand having a beta bulge (where the hydrogen bond is broken).
- no known function

α/β Topologies

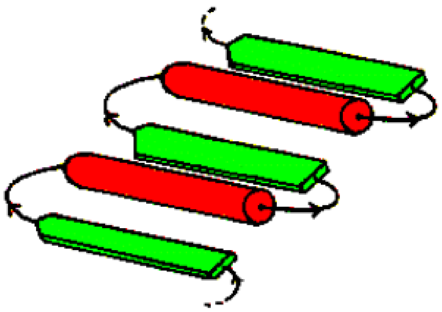
Beta-Helix-Beta Motif

An important and widespread supersecondary structural motif in proteins is known as the β-α-β motif (Beta-Alpha-Beta motif). The motif consists of two parallel Beta strands that is connected via an alpha helix (with two turns). The motif is found in most proteins that contain parallel beta strands, and the axis of the Helix and the Strands are roughly parallel to each other with all three elements forming a hydrophobic core due to shielding. The β-α-β motif may be structurally or functionally involved. The Loop that connects the C-terminal of first Beta strand and N-terminal of Helix is frequently involved in ligand binding functions, and the motif itself is frequently found in ion channels.



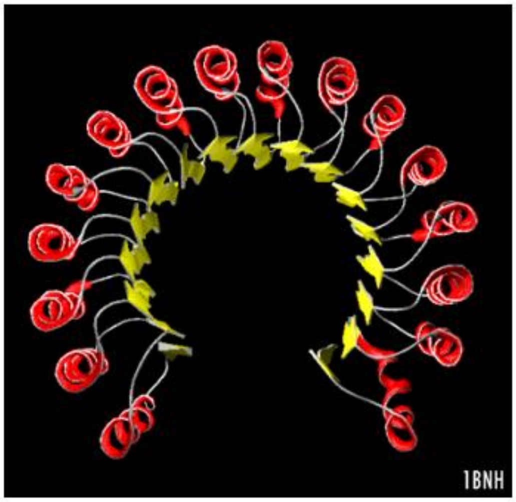
The β - α - β - α - β subunit, often present in nucleotide-binding proteins, is named the **Rossman Fold**, after Michael Rossman

The Rossman fold



α/β horseshoe

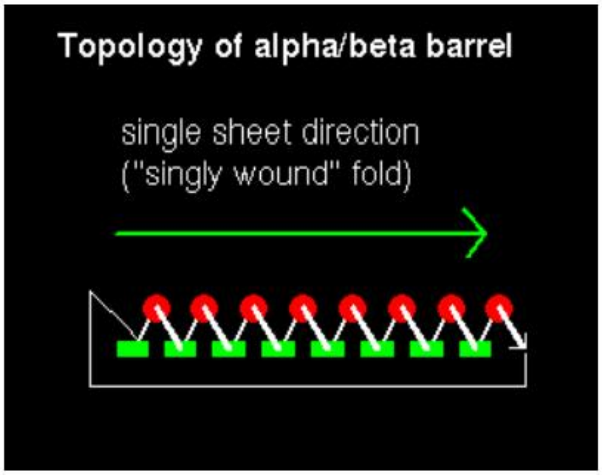
17-stranded parallel b sheet curved into an open horseshoe shape, with 16 a-helices packed against the outer surface. It doesn't form a barrel although it looks as though it should. The strands are only very slightly slanted, being nearly parallel to the central `axis'.



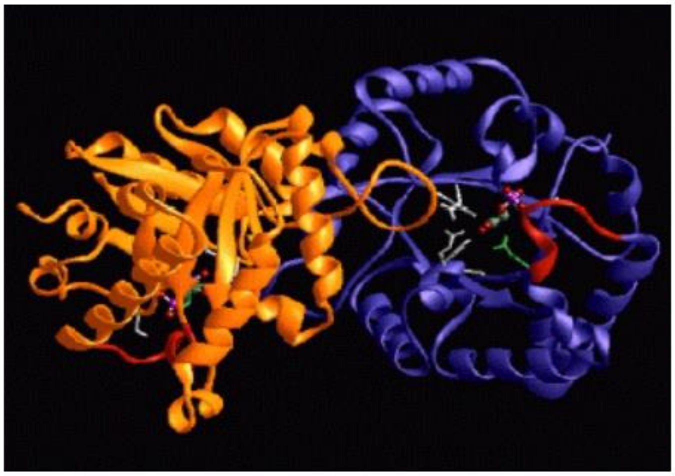
placental ribonuclease inhibitor takes the concept of the repeating α/β unit to extremes.

α/β barrels

Consider a sequence of eight α/β motifs:

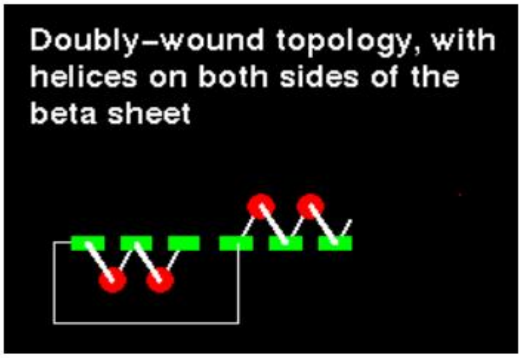


If the first strand hydrogen bonds to the last, then the structure closes on itself forming a barrel-like structure. This is shown in the picture of triose phosphate isomerase.



Note that the "staves" of the barrel are slanted, due to the twist of the b sheet. Also notice that there are effectively four layers to this structure. The direction of the sheet does not change (it is anticlockwise in the diagram). Such a structure may therefore be described as **singly wound**.

In a structure which is open rather than closed like the barrel, helices would be situated on only one side of the b sheet if the sheet direction did not reverse. Therefore open a/b structures must be **doubly wound** to cover both sides of the sheet.

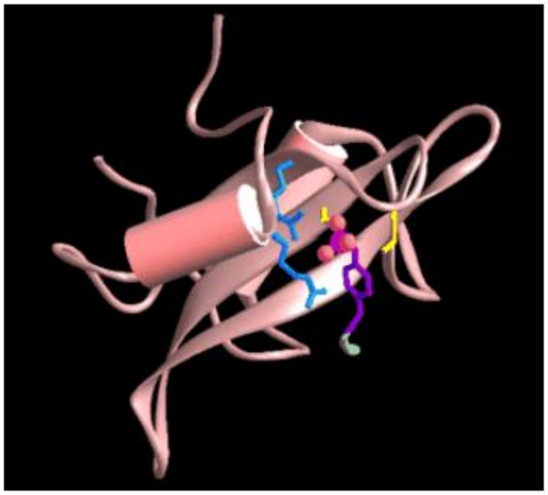


The chain starts in the middle of the sheet and travels outwards, then returns to the centre via a loop and travels outwards to the opposite edge:

Doubly-wound topologies where the sheet begins at the edge and works inwards are rarely observed.

Alpha+Beta Topologies

This is where we collect together all those folds which include significant alpha and beta secondary structural elements, but for which those elements are `**mixed**', in the sense that they do NOT exhibit the wound alpha-beta topology. This class of folds is therefore referred to as **α+ β**



Domains

stable, independently folded, globular units, often consisting of combinations of motifs

vary from 25 to 300 amino acids, average length – 100.

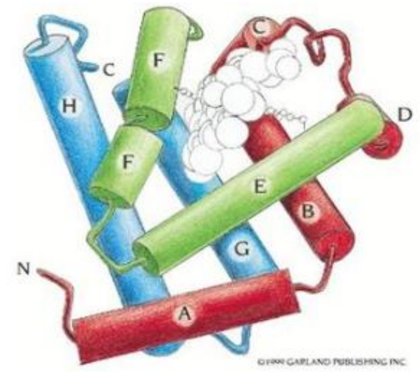
large globular proteins may consist of several domains linked by stretches of polypeptide. Separate domain may have distinct functions (eg G3P dehydrogenase). In many cases binding site formed by cleft between 2 domains

frequently correspond to exon in gene

- Some examples of domains:

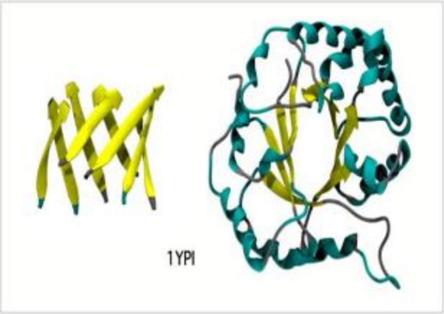
1. in volving α-helix
4-helix bundle
globin fold

PROTEIN ENGINEERING

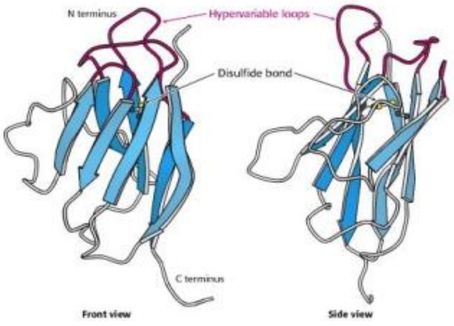


The globin fold is found in its namesake globin protein families: hemoglobins and myoglobins, as well as in phycocyanins. Because myoglobin was the first protein whose structure was solved, the globin fold was thus the first protein fold discovered.

2. parallel β-sheets
   hydrophobic residues on both sides, therefore must be buried.
   □□ barrel: 8 β strands each flanked by an antiparallel α-helix eg triose phosphate isomerase.)



3. antiparallel β -sheet
   hydrophobic residues on one side, one side can be exposed to environment, minimum structure 2 layers
   Sheets arranged in a barrel shape
   More common than parallel β -barrels
   eg. immunoglobulin



PROTEIN ENGINEERING

The **immunoglobulin domain** is a type of protein domain that consists of a 2-layer sandwich of 7-9 antiparallel β-strands arranged in two β-sheets with a Greek keytopology, consisting of about 80 amino acids.

The backbone switches repeatedly between the two β-sheets. Typically, the pattern is (N-terminal β-hairpin in sheet 1)-(β-hairpin in sheet 2)-(β-strand in sheet 1)-(C-terminal β-hairpin in sheet 2). The cross-overs between sheets form an "X", so that the N- and C-terminal hairpins are facing each other.

Members of the immunoglobulin superfamily are found in hundreds of proteins of different functions. Examples include antibodies, the giant muscle kinase titin, andreceptor tyrosine kinases. Immunoglobulin-like domains may be involved in protein–protein and protein–ligand interactions.