

# Beta Structures

# Beta Structures

-comprise the second largest group of protein domain structures

-functionally, this group is the most diverse

- enzymes
- transport proteins
- antibodies
- cell surface proteins
- virus coat proteins

# Beta Structures

- consist of  $\beta$  strands that can vary in numbers from four or five to over ten
- the strands are arranged in a predominantly antiparallel fashion and usually in such a way that they form two  $\beta$  sheets that are joined together and packed against each other
- the  $\beta$  sheets have the usual twist, and when packed together, they form a barrel-like structure
- antiparallel  $\beta$  structures have a core of hydrophobic side chains inside the barrel provided by residues in the  $\beta$  strands
- the surface is formed by residues from the loop regions and from the strands

# Beta Structures

- the aim of this chapter is to examine a number of antiparallel  $\beta$  structures and demonstrate how these rather complex structures can be separated into smaller comprehensible motifs
- the number of possible ways to form antiparallel  $\beta$  sheet structures rapidly increases as the number of strands increases
- surprisingly, the number of topologies actually observed is small and most  $\beta$  structures fall into a few groups of common topology

# Beta Structures

The three most frequently occurring groups

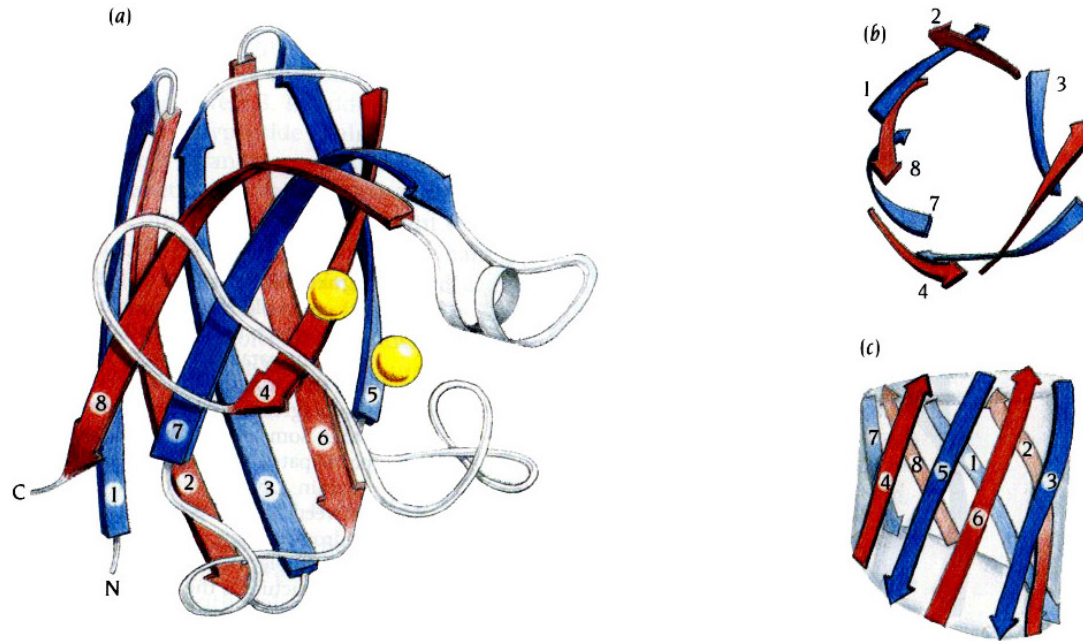
- up-and-down barrels

- Greek key barrels

- jelly roll barrels

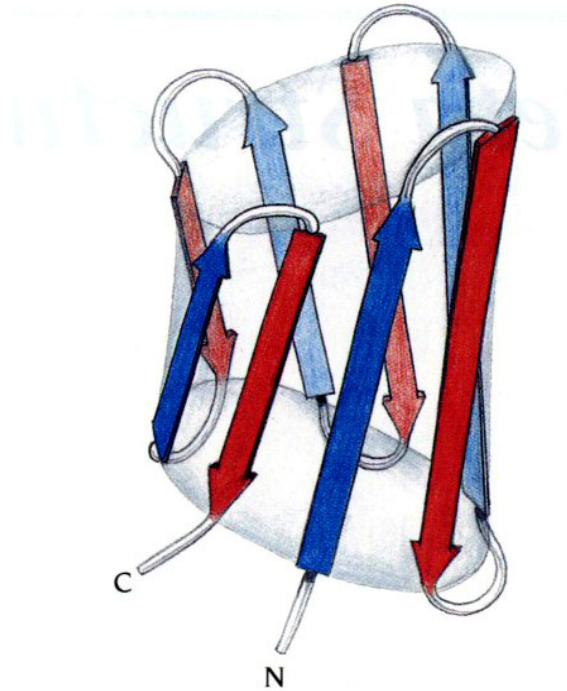
can all be related to simple ways of connecting antiparallel  $\beta$  strands arranged in a barrel structure.

## Two antiparallel $\beta$ sheets form a barrel-like structure when packed together

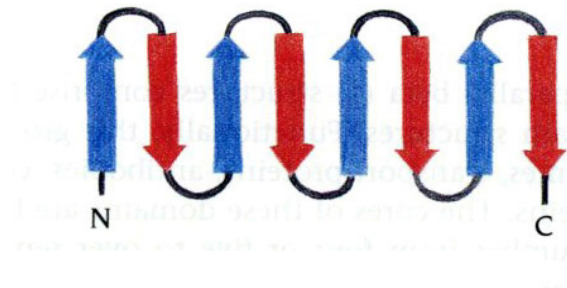


The enzyme superoxide dismutase (SOD) is a  $\beta$  structure comprising 8 antiparallel  $\beta$  strands (a). In addition, SOD has two metal atoms, Cu and Zn, that participate in the catalytic action: conversion of superoxide radical to hydrogen peroxide and oxygen. The eight  $\beta$  strands are arranged in a barrel, which is viewed along the barrel axis in (b) and perpendicular to the axis in (c). In general, antiparallel  $\beta$  structures have a core of hydrophobic side chains inside the barrel provided by residues in the  $\beta$  strands. The surface is formed by the residues from the loop regions and from the strands.

## Topology of up-and-down $\beta$ barrels



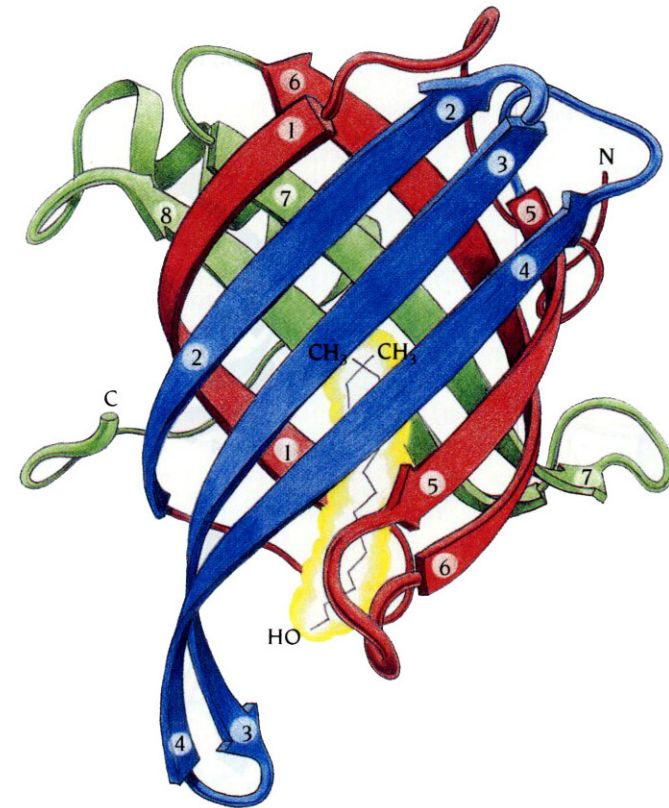
The 8  $\beta$  strands are all antiparallel to each other and are connected by hairpin loops. Beta strands that are adjacent in the amino acid sequence are also adjacent in the three-dimensional structure of up-and-down barrels.



## The retinol-binding protein binds retinol inside an up-and-down barrel

The eight antiparallel  $\beta$  strands twist and curl such that the structure can also be regarded as two  $\beta$  sheets (green and blue) packed against each other. Some of the twisted  $\beta$  strands (red) participate in both  $\beta$  sheets. A retinol molecule, vitamin A (yellow), is bound inside the barrel, between the two  $\beta$  sheets, such that its only hydrophilic part (an OH tail) is at the surface of the molecule. The retinol-binding protein (RBP) is synthesized in hepatocytes, where it picks up one molecule of retinol in the endoplasmic reticulum. In plasma, the RBP-retinol complex is stabilized by interaction with prealbumin. Recognition of this complex by cell-specific receptors causes RBP to release the retinol and, as a result,

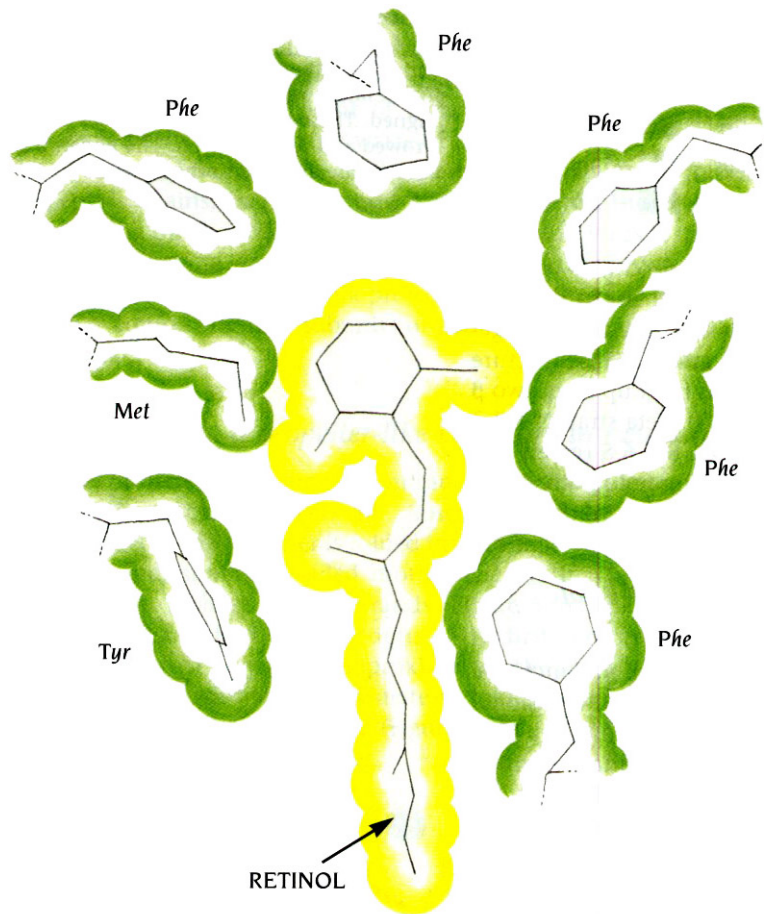
to undergo a conformational change that drastically reduces its affinity to prealbumin. The free RBP molecule is then excreted through the kidney glomerus, reabsorbed in the proximal tubule cells, and degraded.









The hydrophobic retinol molecule is packed against hydrophobic side chains from the  $\beta$  strands in the barrel's core

The binding site for retinol inside the RBP barrel is lined with hydrophobic residues. They provide a hydrophobic surrounding for the hydrophobic part of the retinal molecule.

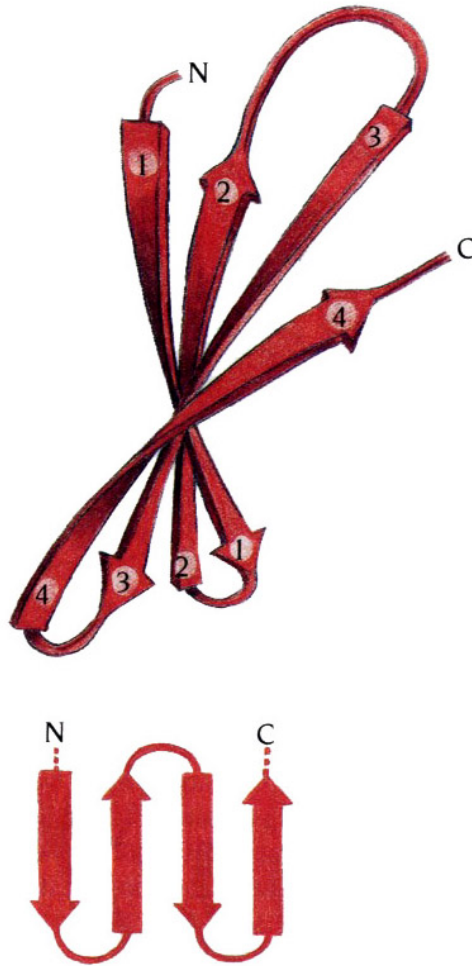


## Amino acid sequence reflects $\beta$ structure

<i>strand no.</i>	<i>residue no.</i>	<i>amino acid sequence</i>
2	41–48	–  <b>Ile</b> – Val –  <b>Ala</b> – Glu –  <b>Phe</b> – Ser –  <b>Val</b> – Asp –
3	53–60	– <b>Met</b> – Ser – <b>Ala</b> – Thr – <b>Ala</b> – Lys – <b>Gly</b> – Arg –
4	71–78	– <b>Ala</b> – Asp – <b>Met</b> – Val – <b>Gly</b> – Thr – <b>Phe</b> – Thr –

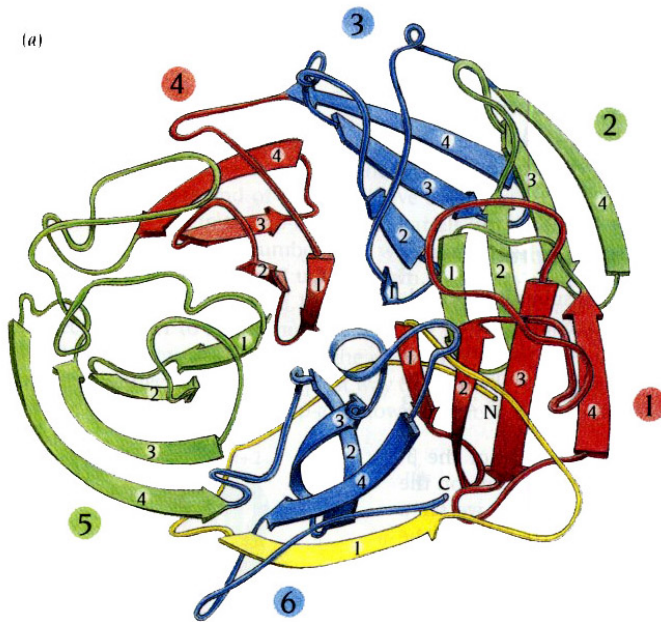
Amino acid sequence of  $\beta$  strands 2 3 4 in human plasma retinol-binding protein. The sequences are listed in such a way that residues which point into the barrel are aligned. These hydrophobic residues are arrowed and colored green. The remaining residues are exposed to the solvent.

## Neuramidase from influenza virus folds into up-and-down $\beta$ sheets

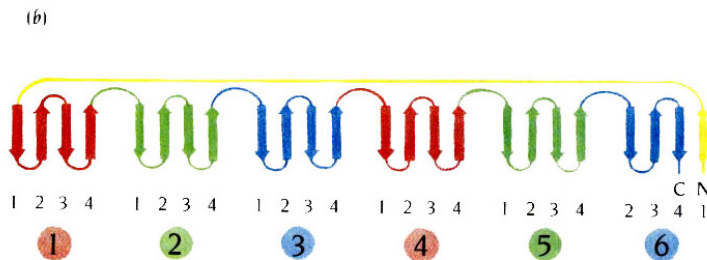


The neuramidase is a homotetramer made up of four identical polypeptide chains, each of around 470 amino acids. Proteolytic treatment releases a soluble head of the neuramidase from the stalk projecting from the viral envelope. The structure of the soluble head, comprising of about 400 amino acids each, was determined. Each of the four subunits of the tetramer is folded into a single domain built up from six closely packed, similar folded motifs. The motif is a simple up-and-down antiparallel  $\beta$  sheet of four strands. The strands have a rather large twist such that the directions of the first and the fourth strands differ by 90 degrees.

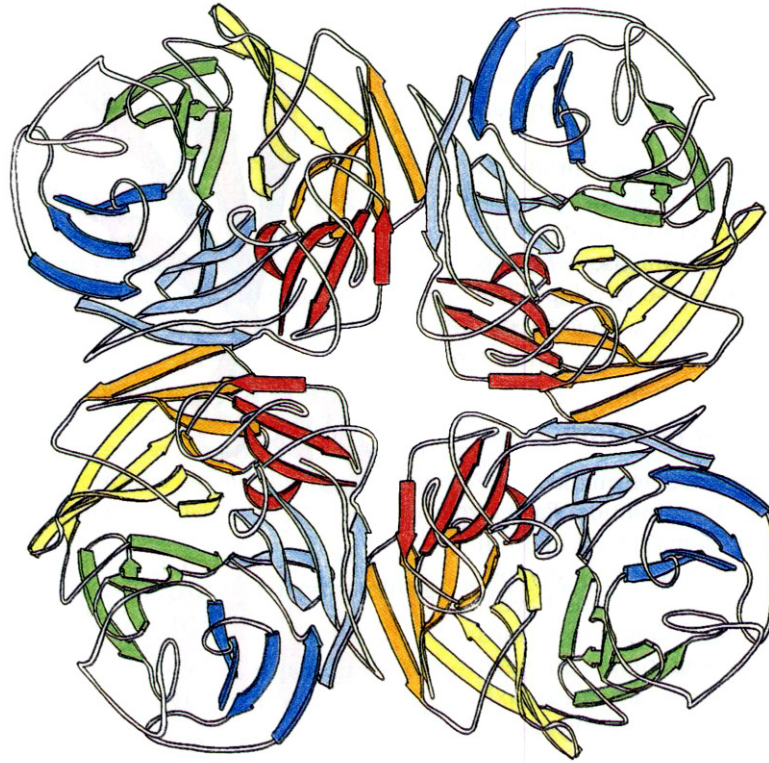
The subunit structure of the neuraminidase headpiece is built up from six similar, consecutive motifs of 4 up-and-down antiparallel  $\beta$  strands



Each such motif has been called a propeller blade and the whole subunit structure a six-blade propeller. The motifs are connected by loop regions from  $\beta$  strand 4 in one motif to  $\beta$  strand 1 in the next motif. The schematic diagram (a) is viewed down an approximate sixfold axis that relates the centers of the motifs.



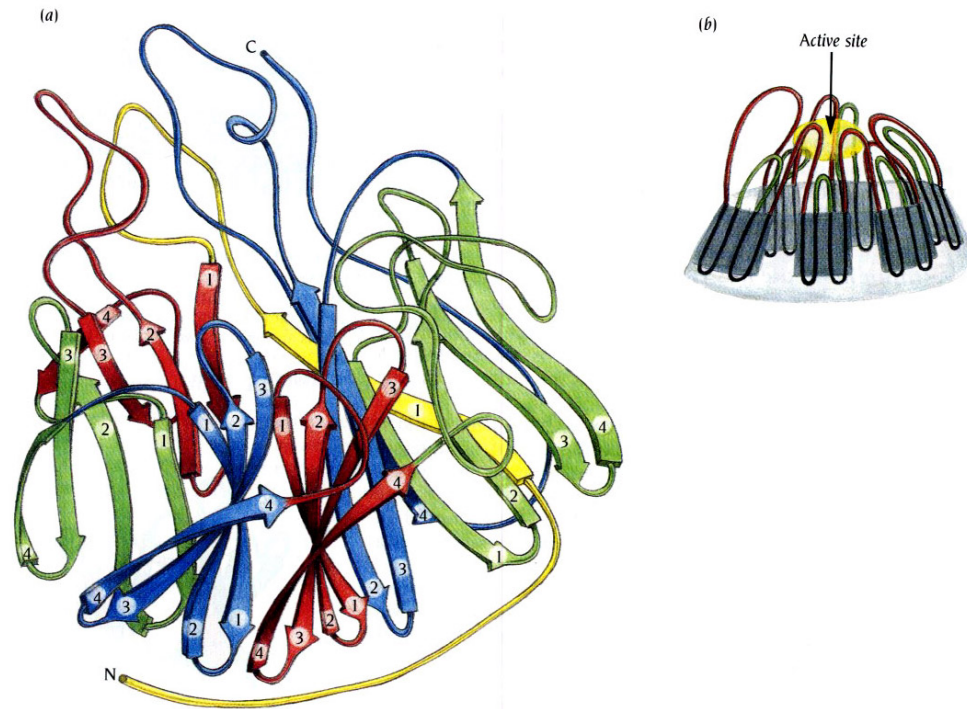
Four six-blade propeller subunits are present in each complete neuraminidase molecule



Schematic view down the fourfold axis of the tetrameric molecule of neuraminidase.

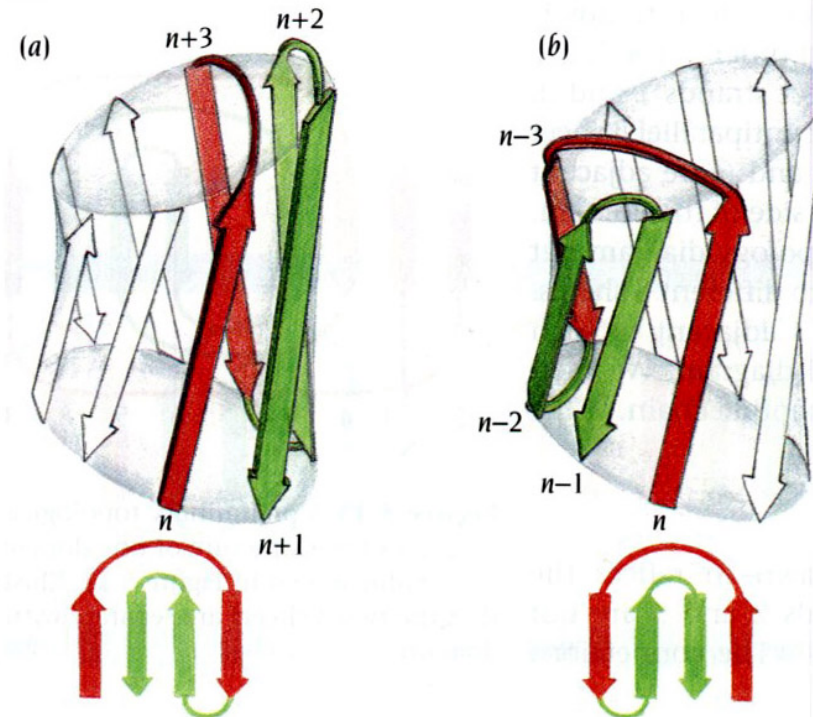


## The neuraminidase active site is in the middle of one side of the propeller



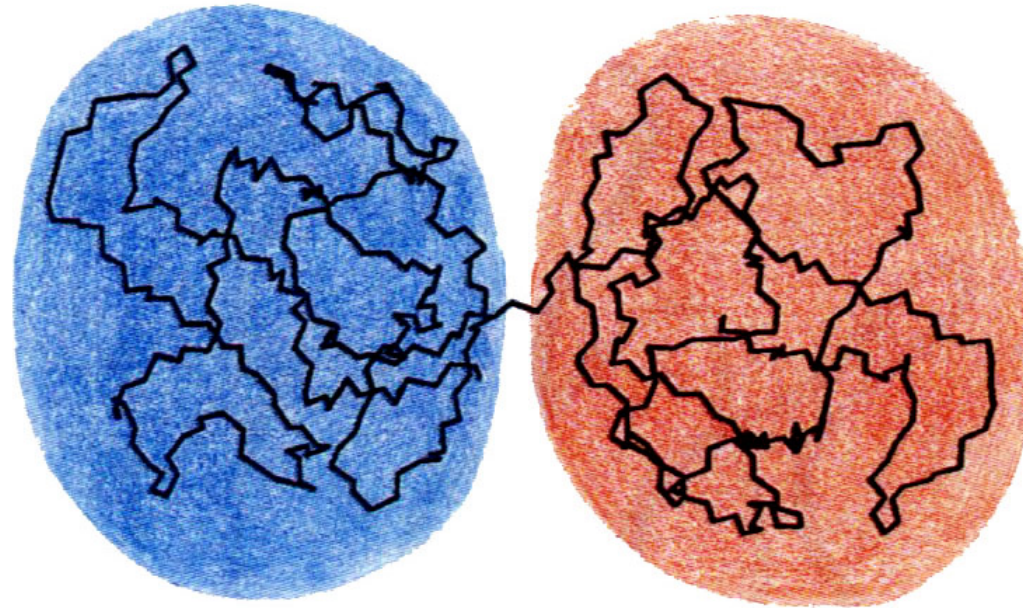
- a) A schematic diagram of the subunit structure shows the propeller viewed from its side.
- b) An idealized propeller structure viewed from the side to highlight the position of the active site. The loop regions that connect the motifs (red in b) in combination with the loops that connect strands 2 and 3 within the motifs (green in b) form a wide funnel-shaped active site pocket.

## Greek key motifs occur frequently in antiparallel $\beta$ structures



Idealized diagram of the Greek key motif is shown. This motif is formed when one of the connections of four antiparallel  $\beta$  strands is not a hairpin connection. The motif occurs when strand number  $n$  is connected to strand  $n + 3$  (a) or  $n - 3$  (b) instead of  $n + 1$  or  $n - 1$  in an eight-stranded antiparallel  $\beta$  sheet or barrel. The two different possible connections give two different hands of the Greek key motif. In all protein structures known so far only the hand shown in (a) has been observed.

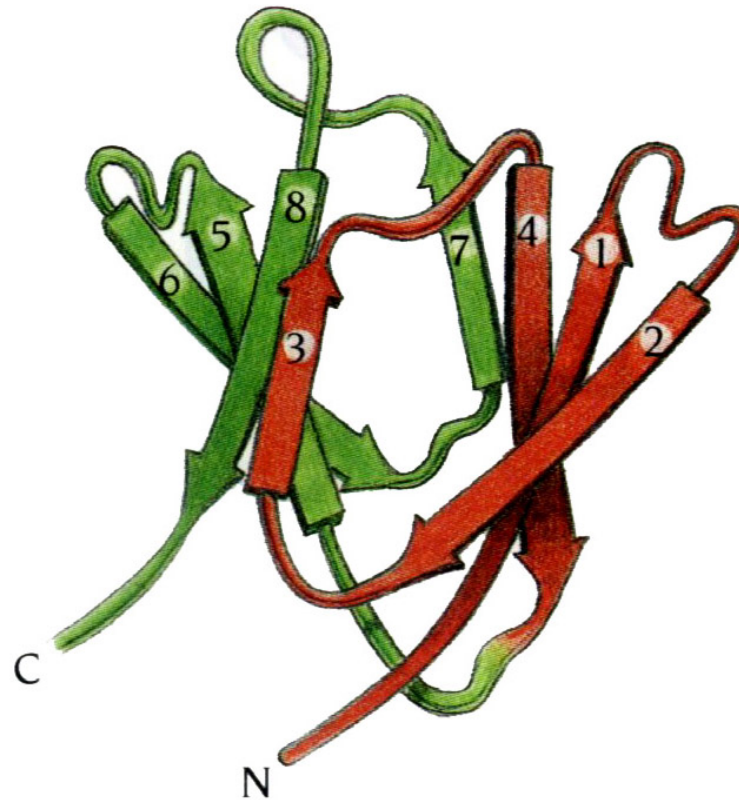
The  $\gamma$ -crystallin molecule has two domains



A computer-generated diagram of the structure of  $\gamma$ -crystallin comprising one polypeptide chain of 170 amino acid residues. Crystallins are lens-specific proteins in our eyes.

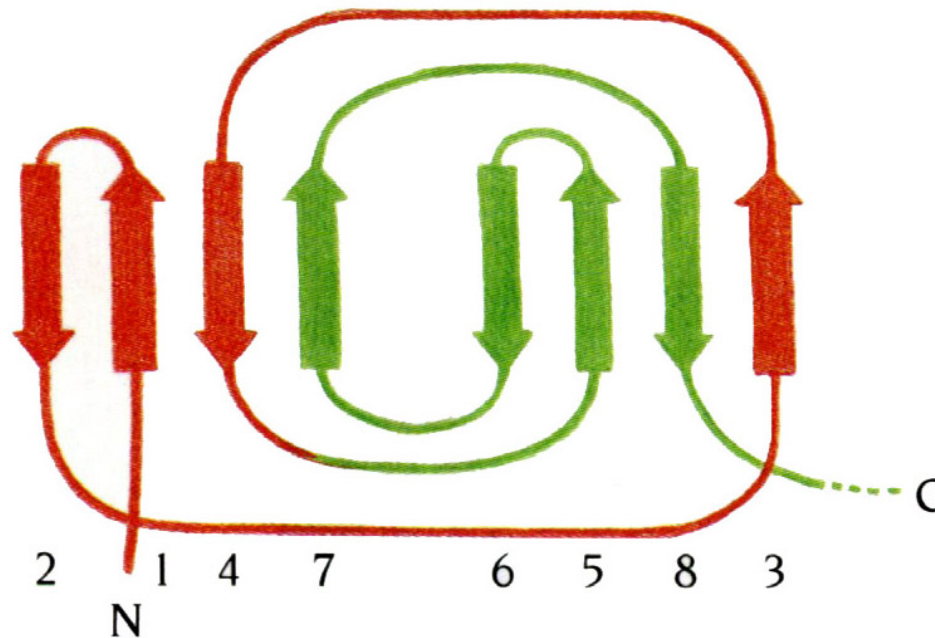


The  $\gamma$ -crystallin domain structure has a simple topology



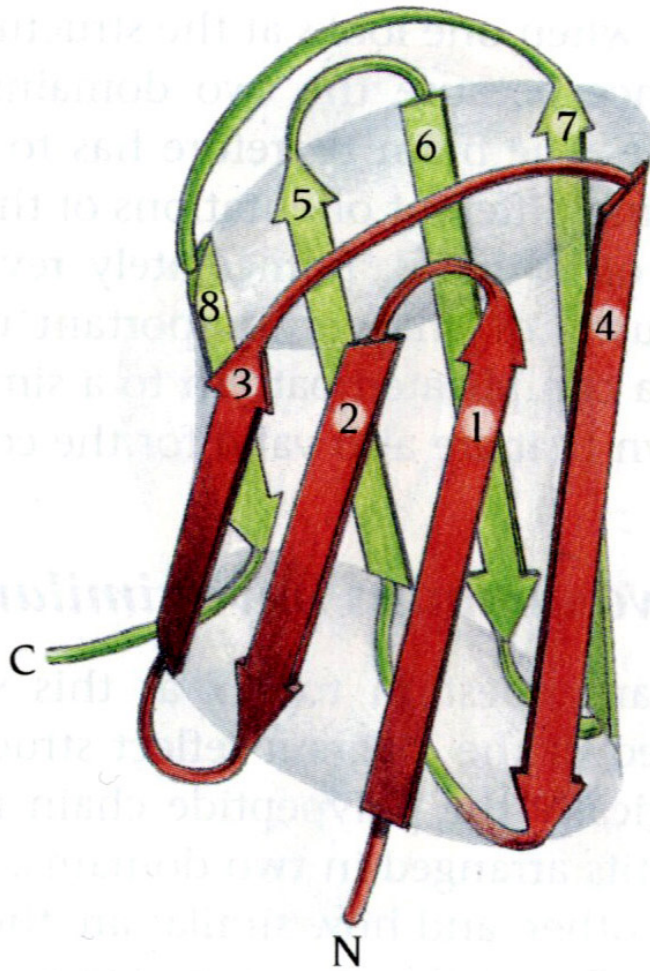
The domain structure is built up from two  $\beta$  sheets of four antiparallel  $\beta$  strands, sheet 1 from  $\beta$  strands 1, 2, 4, and 7 and sheet 2 from strands 3, 5, 6, and 8. The two  $\beta$  sheets are packed against each other so that they form a distorted barrel.

A preliminary topological diagram of one domain  
of  $\gamma$ -crystallin



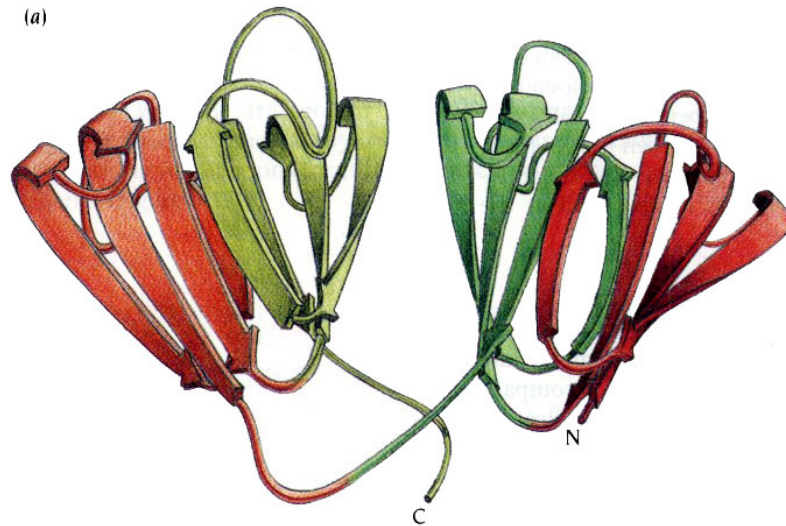
The diagram illustrates that the two  $\beta$  sheets are separated within the domain.

## Two Greek key motifs form the $\gamma$ -crystallin domain

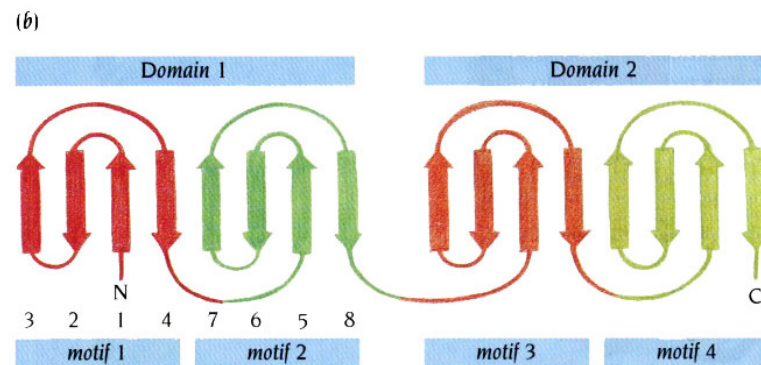


The eight  $\beta$  strands in one domain of the crystalline structure are drawn along the surface of a barrel. From this diagram it is obvious that the  $\beta$  strands are arranged in two Greek key motifs, one (red) formed by strands 1 – 4 and the other (green) by strands 5 – 8. The  $\beta$  strands that form one motif contribute to both  $\beta$  sheets as shown earlier.

## The two $\gamma$ -crystallin domains have identical topology

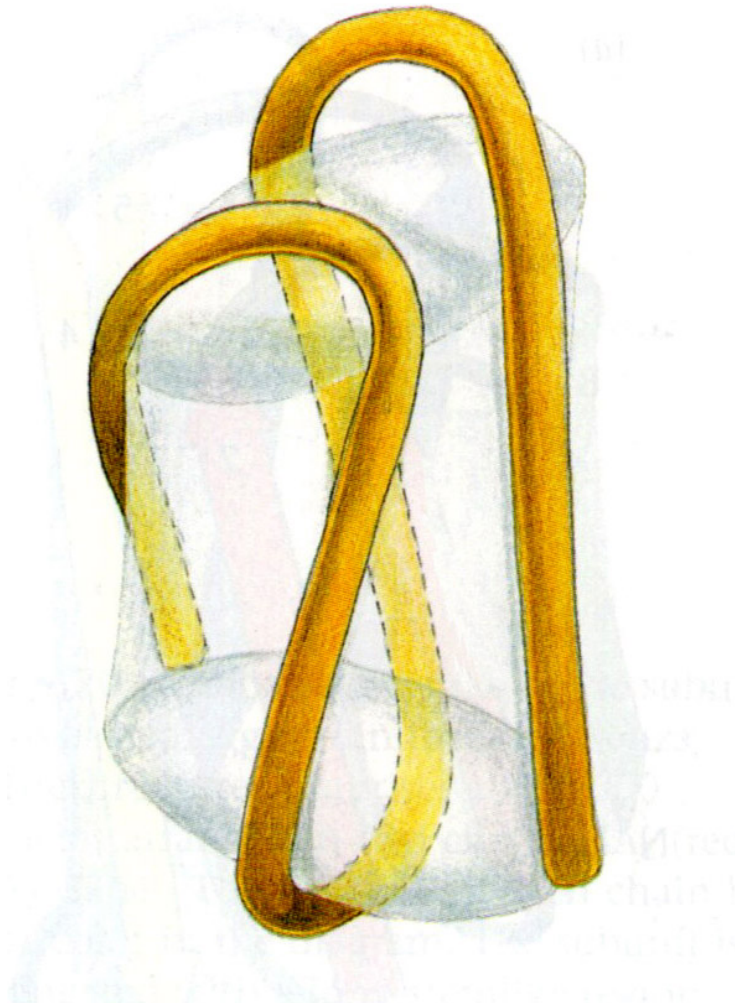


The two domains of the complete molecule have the same topology; each is composed of two Greek key motifs that are joined by a short loop region.



This illustrates one very important use of topology diagrams – namely, to reduce a complicated pattern to a simple one, from which conclusions can be drawn that are valid for the complicated pattern.

## The jelly roll motif is wrapped around a barrel

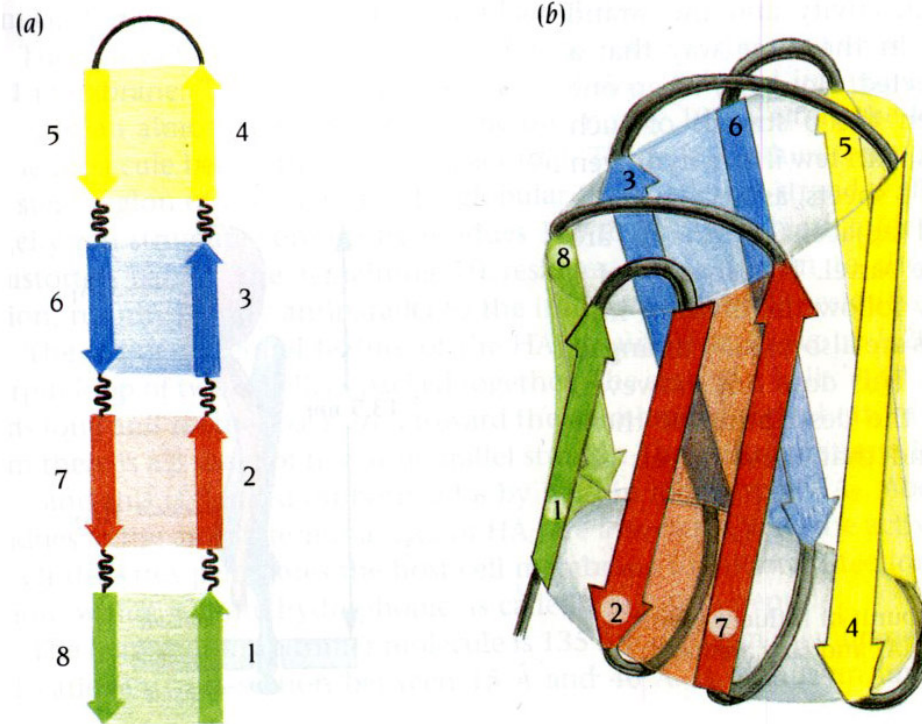


A diagram of a piece of string wrapped around a barrel to illustrate the basic pattern of a jelly roll motif. The string goes up and down the barrel four times, crosses over once at the bottom and twice at the top of the barrel. This configuration is the basic pattern of for the jelly roll motif.

This motif has been found in a variety of different structures including the coat proteins of most spherical viruses examined so far by x-ray crystallography, the plant lectin concanavalin A and the hemagglutinin protein from the influenza virus.

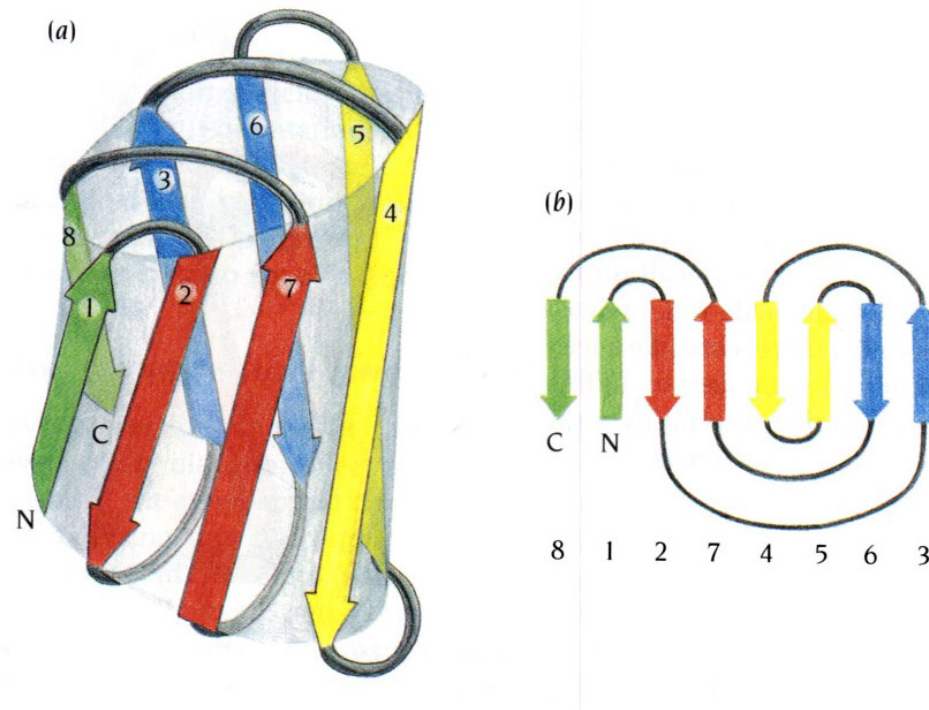


## The jelly roll motif is wrapped around a barrel



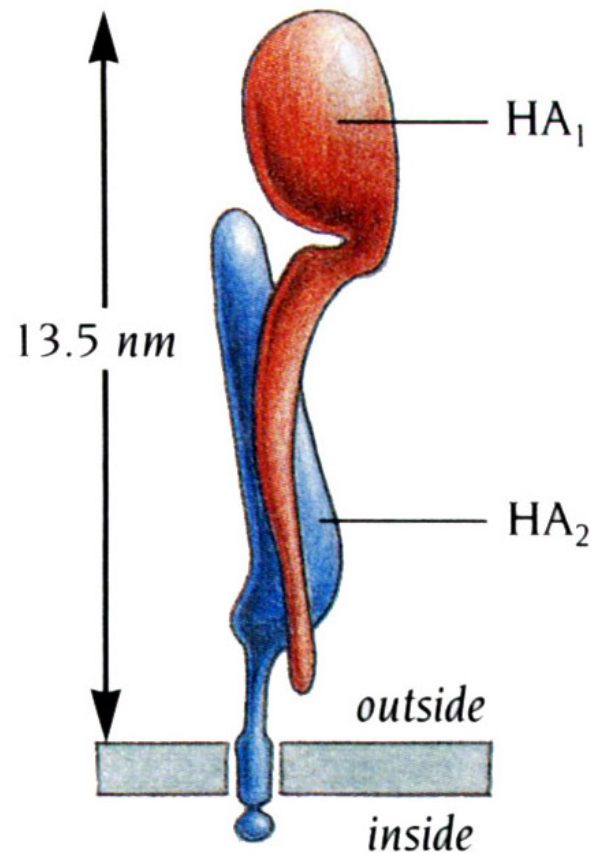
Eight  $\beta$  strands can be arranged in a jelly roll motif. (a) The eight strands are drawn as arrows along two edges of a strip of paper. The  $\beta$  strands are separated by loop regions. (b) The strip of paper is wrapped around a barrel such that the  $\beta$  strands follow the surface of the barrel and the loop regions provide the connections at the bottom ends of the barrel. The  $\beta$  strands are now arranged in a jelly roll motif.

## The topological diagrams of the jelly roll structure



The real-world jelly roll barrels are usually divided into two sheets. When one analyzes the pattern of hydrogen bonds between the  $\beta$  strands of such barrels, one finds that they usually form two sheets with few if any hydrogen bonds between the strands that belong to the different  $\beta$  sheets as we saw in the crystallin structure. The barrel is distorted and adjacent  $\beta$  strands are separated from each other in two places across the barrel. The division of  $\beta$  strands does not necessarily follow the division into topological motifs. Thus, the real-world barrels are distorted and flattened.

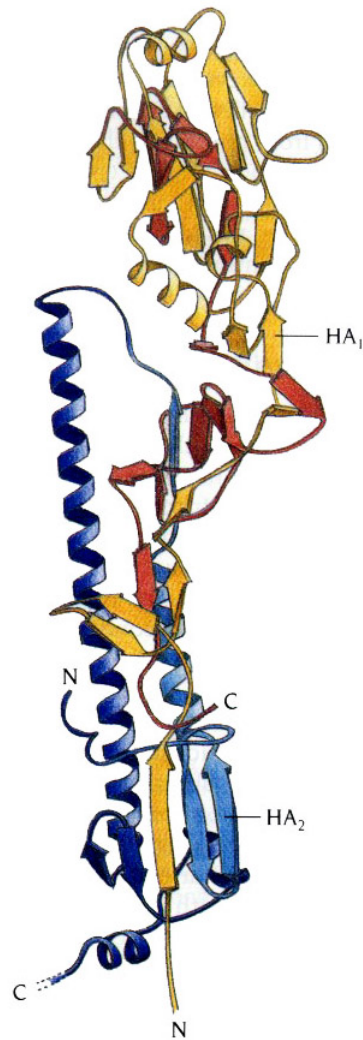
## The functional hemagglutinin subunit has two polypeptide chains



Schematic picture demonstrates that a single subunit of influenza virus consists of two polypeptide chains HA<sub>1</sub> and HA<sub>2</sub>, which are held together by disulfide bonds. The hemagglutinin molecule on the surface of a viral particle is formed from three subunits. Protease treatment of influenza virus particles releases a soluble fragment consisting of three complete HA<sub>1</sub> chains disulfide bonded to three HA<sub>2</sub> chains complete except for their membrane anchor regions. Crystal structure of the fragment has been solved, and will be discussed as an example of a protein consisting partly of jelly roll domains.

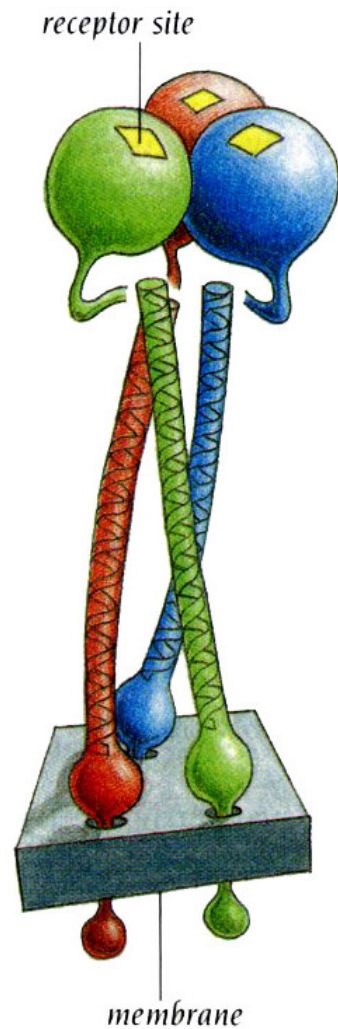


## The hemagglutinin subunit structure is divided into a stem and a tip



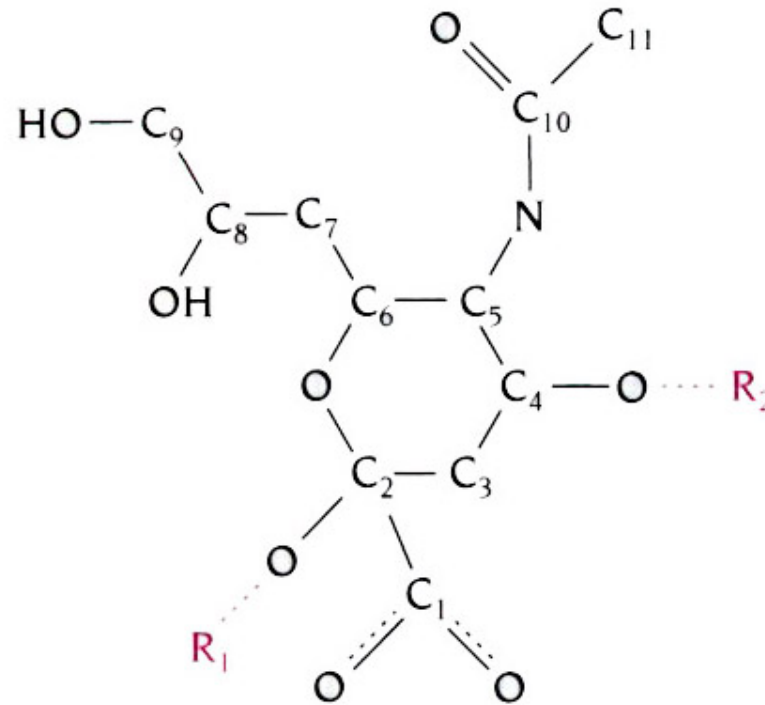
The monomeric subunit is divided into a long, fibrous stem-like region extending outward from the membrane with a globular region at its tip. The structure comprises about 550 amino acids arranged in two chains HA<sub>1</sub> (red) and HA<sub>2</sub> (blue). The first half of each chain has a lighter color in the diagram. The subunit is very elongated with a stem-like region built up by residues from both chains and includes one of the longest  $\alpha$  helices known in globular proteins, about 75 Å long. The globular head is formed by residues only from HA<sub>1</sub>.

## The hemagglutinin molecule is formed from three subunits



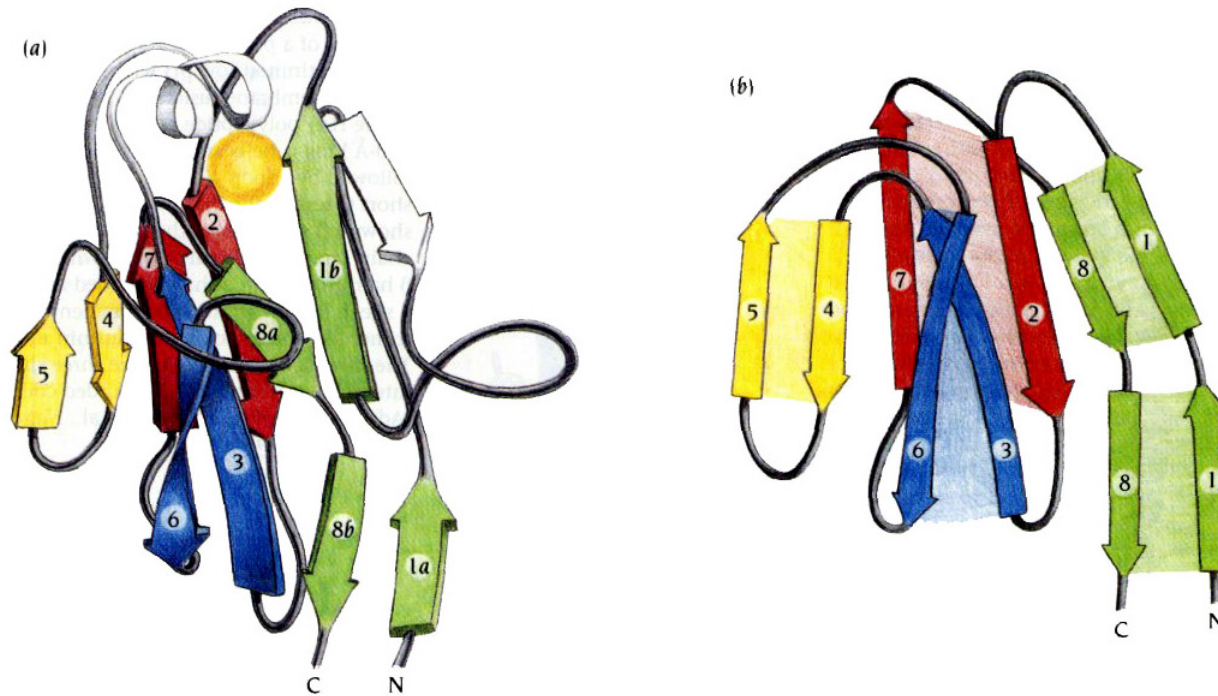
Each of these subunits is anchored in the membrane of the influenza virus. The globular heads contain the receptor sites that bind to sialic acid residues on the surface of eukaryotic cells. A major part of the subunit interface is formed by the three long intertwining helices, one from each subunit.

## Sialic acid – $\alpha$ -5-*n*-acetylneuraminic acid



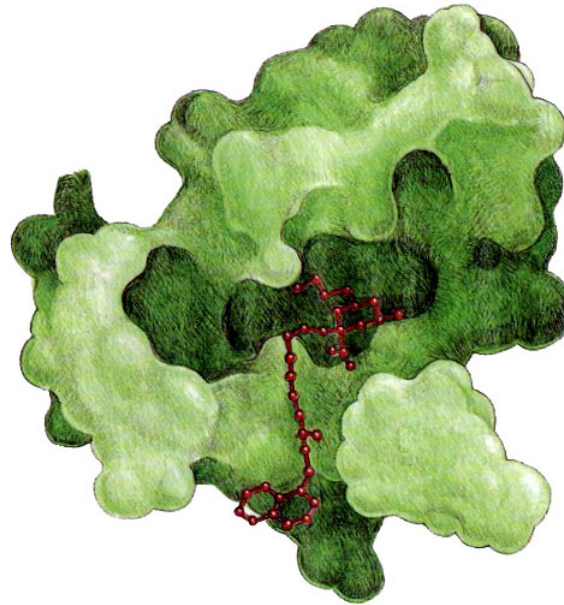
To initiate infection, the virus hemagglutinin binds to the sialic acid residues of glycosylated receptor proteins on the target cell surface. Once bound to the receptor, the virus is then taken into the cell by endocytosis. The receptor binding site on the hemagglutinin molecule has been determined from experiments with mutants of hemagglutinin and from the determination of structures with bond inhibitors (modified sialic acid molecules with large hydrophobic substituents, R<sub>1</sub> and R<sub>2</sub>, at two different positions of the sugar ring).

## The receptor binding site is formed by the jelly roll domain



The globular head of the hemagglutinin subunit is a distorted jelly roll structure (a).  $\beta$  strand 1 contains a long insertion, and  $\beta$  strand 8 contains a bulge in the corresponding position. Each of these two strands is therefore subdivided into shorter  $\beta$  strands. The loop region between  $\beta$  strand 3 and 4 contains a short  $\alpha$  helix, which forms one side of the receptor binding site (yellow circle). A schematic diagram (b) illustrates the organization of the  $\beta$  strands into a jelly roll motif.

## Sialic acid binding domain of hemagglutinin



Space-filling model of the sialic acid binding domain of hemagglutinin with a bound inhibitor illustrates the different binding grooves. The sialic acid moiety of the inhibitor binds in the central groove. A large hydrophobic substituent,  $R_1$ , at the  $C_2$  position of sialic acid binds in a hydrophobic channel that runs from the central groove to the bottom of the domain.

## The receptor binding site is a target for our immune system, and drug design

Antibodies in our immune system bind to the receptor site, so preventing virus from entering a cell. The virus can escape this neutralization through mutations in residues that form this binding site. Such mutations, however, are found only at the rim of the sialic acid binding pocket, presumably because mutation of residues inside this pocket would prevent the virus from binding to the cell surface receptor and consequently prevent viral propagation.

The receptor binding site is therefore an ideal target for drug design, and studies of inhibitor binding to this site have provided valuable clues and ideas for the design of molecules that are candidate drug therapies for influenza virus infection.

## Hemagglutinin acts as a membrane fusogen

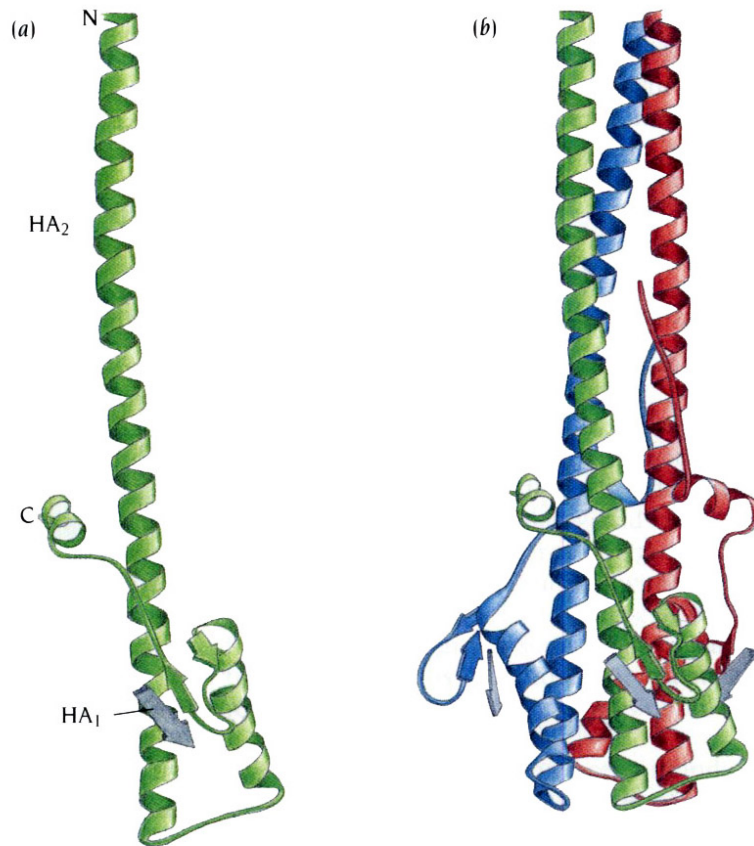
Viruses, bound to the plasma membrane via their membrane receptors, are taken into the cells by endocytosis. Proton pumps in the membrane of endocytic vesicles that now contain the bound viruses cause an accumulation of protons and a consequent lowering of the pH inside the vesicles. The acidic pH (below pH 6) allows hemagglutinin to fulfill its second role namely, to act as a membrane fusogen by inducing the fusion of the viral envelope membrane with the membrane of the endosome. This expels the viral RNA into the cytoplasm, where it can begin to replicate.

This fusogenic activity of influenza hemagglutinin is frequently exploited in laboratory to achieve (i) infection without endocytosis, (ii) delivery of artificial vesicle content into target cells, and (iii) cell fusion.

What are the mechanisms underlying the fusogenic properties of hemagglutinin?



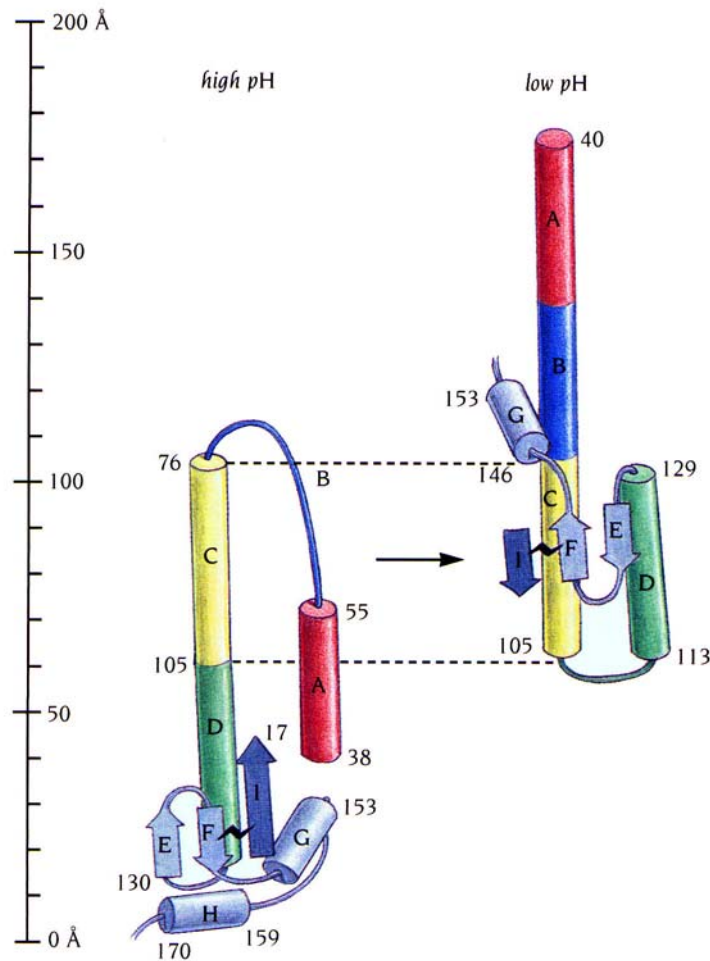
## The structure of hemagglutinin is affected by pH changes



Schematic diagram shows the structure of a proteolytic fragment of hemagglutinin at low pH where the molecule induces membrane fusion. (a) Residues 38 – 175 of the HA<sub>2</sub> polypeptide chain (green) form a 100 Å long  $\alpha$  helix starting at the N-terminus followed by a loop, a  $\beta$  hairpin and finally a short C terminal helix. In addition the diagram shows a  $\beta$  strand from the HA<sub>1</sub> polypeptide chain (gray) which participates with the  $\beta$  hairpin to form a three-stranded antiparallel  $\beta$  sheet. (b) The proteolytic fragment forms a trimer like the intact hemagglutinin molecule. The three  $\alpha$  helices of the three subunits intertwine to form a three-stranded coiled-coil.

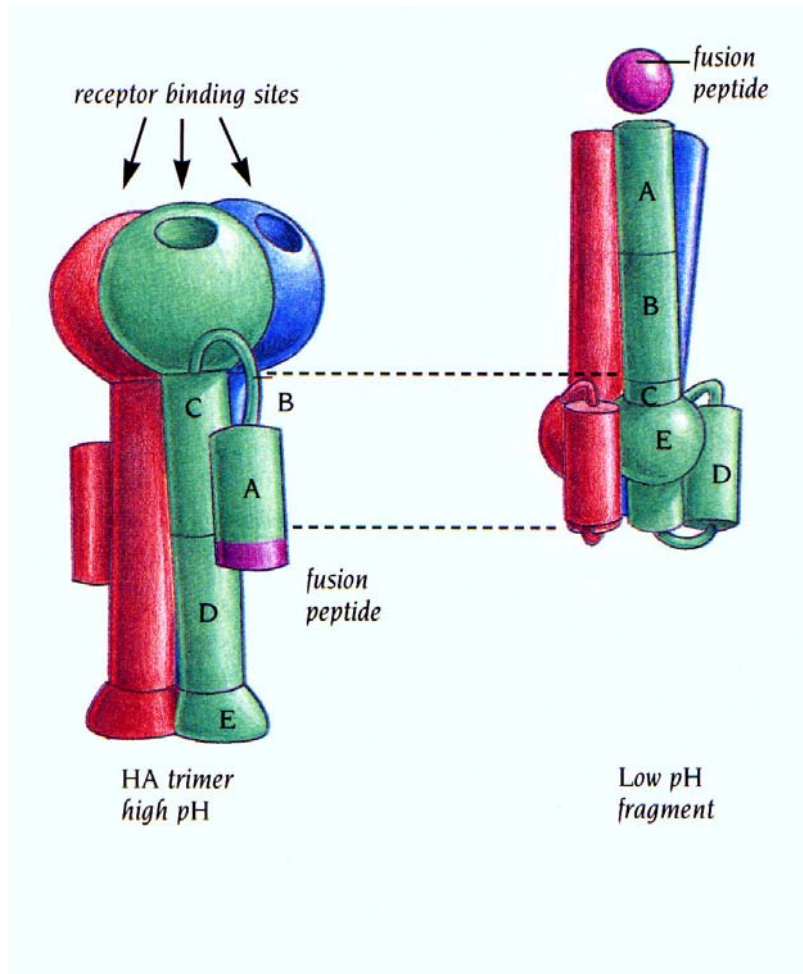


# The structure of hemagglutinin is affected by pH changes



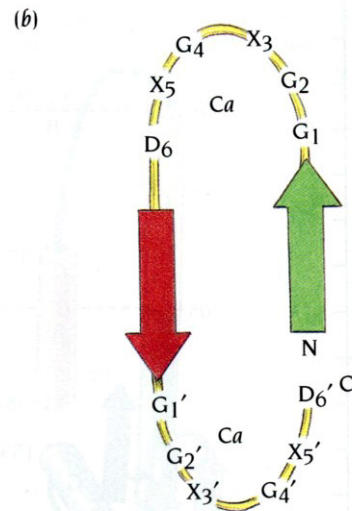
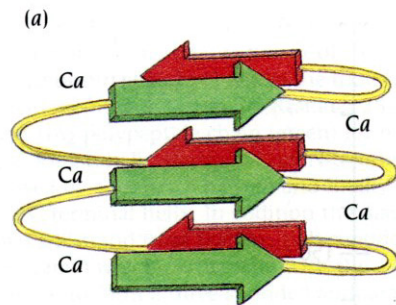
Schematic diagram illustrates the large conformational differences between the high and low pH forms of hemagglutinin. The loop region B in the high pH form has changed into an  $\alpha$  helix producing a continuous 100 Å long helix composed of regions A, B and C at low pH. Furthermore, residues 105-113, which in the high pH form are in the middle of helix C-D, form a loop in the low pH form causing helix D to be at a very different position. Consequently the  $\beta$  hairpin E-F and C-terminal helix G and the  $\beta$  strand 1 from HA<sub>1</sub> occupy very different positions in the two forms even though they have the same internal structure. The squiggle between  $\beta$  strand F and I denotes the S-S bond that joins HA<sub>1</sub> and HA<sub>2</sub>.

## The structure of hemagglutinin is affected by pH changes



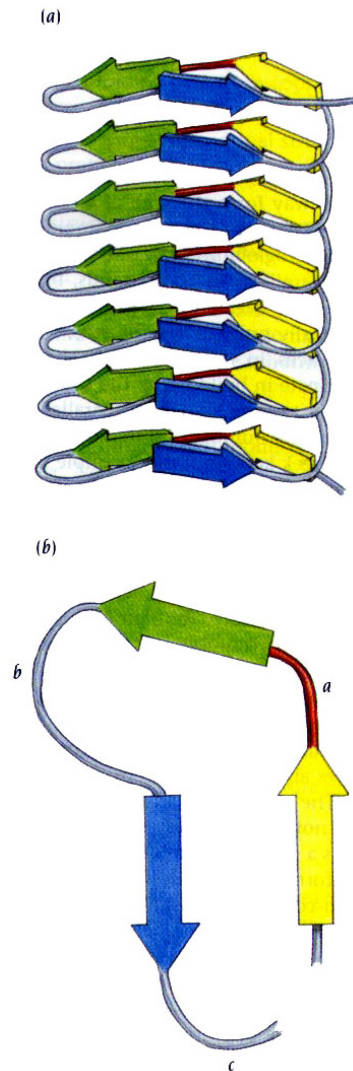
Schematic model represents the conformational change of hemagglutinin that at low pH brings the fusion peptide to the same end of the molecule as the receptor binding site. The fusion peptide (purple) is at the end of  $\alpha$  helix A about 100 Å away from the receptor binding site in the high pH form. In the low pH fragment this region of helix A has moved about 100 Å towards the area where the receptor binding sites are expected to be in the intact hemagglutinin molecule.

## Parallel $\beta$ -helix domains have a novel fold



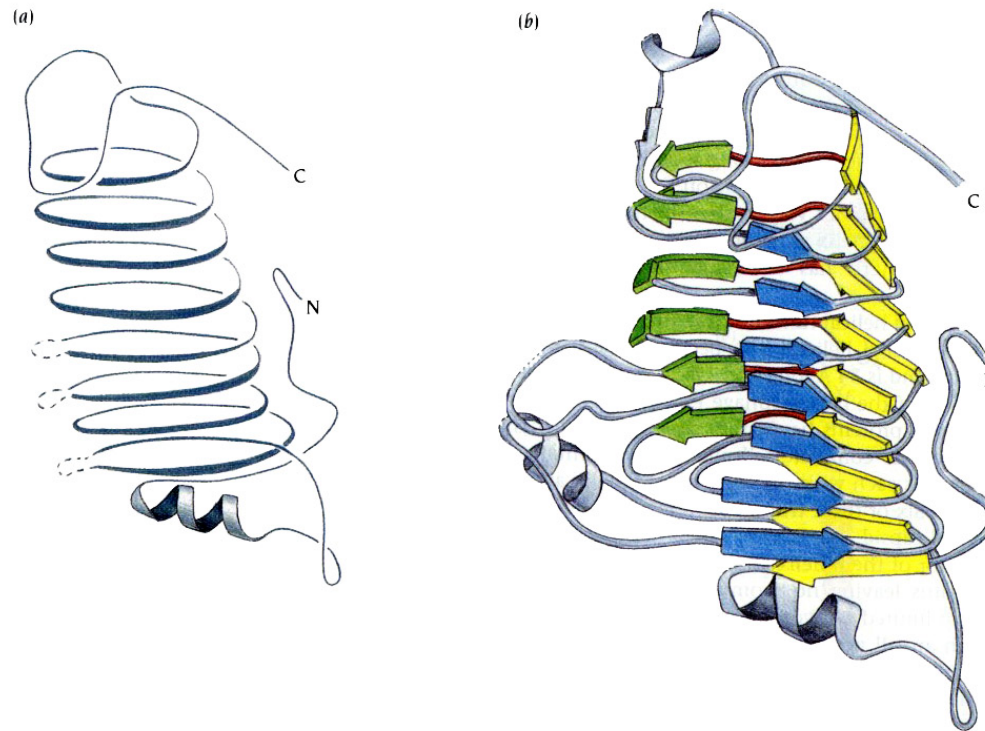
Schematic diagram shows the two-sheet  $\beta$  helix. The complete coils of the helix are shown in (a). The two parallel  $\beta$  sheets are colored green and red, the loop regions that connect the  $\beta$  strands are yellow. (b) Each structural unit is composed of 18 residues forming a  $\beta$ -loop- $\beta$ -loop structure. Each loop region contains six residues of sequence Gly-Gly-X-Gly-X-Asp where X is any residue. Calcium ions are bound to both loop regions.

## Parallel $\beta$ -helix domains have a novel fold



Schematic diagrams show the three-sheet  $\beta$  helix. (a) The three sheet of parallel  $\beta$  strands are colored green, blue and yellow. Seven complete coils are shown in this diagram but the number of coils varies in different structures. Two of the  $\beta$  sheets (blue and yellow) are parallel to each other and are perpendicular to the third (green). (b). Each structural unit is composed of three  $\beta$  strands connected by three loop regions (labeled a, b and c). Loop a (red) is invariably composed of only two residues, whereas the other two loop regions vary in length.

## Parallel $\beta$ -helix domains have a novel fold



Schematic diagrams show the structure of the enzyme [pectate lyase C](#), which has a three-sheet parallel  $\beta$ -helix topology. (a) Idealized diagram highlighting the helical nature of the path of the polypeptide chain which comprises eight helical turns. Dotted regions indicate positions where large external loops have been removed for clarity. (b) Ribbon diagram shows the polypeptide chain. The predominant secondary structural elements are three parallel  $\beta$  sheets. [Each  \$\beta\$  sheet is composed of 7-10 parallel  \$\beta\$  strands with an average length of four to five residues in each strand.](#) The short loop regions of two residues length are shown in red.