IIT Guwahati

Lecture 37

Course BT 631

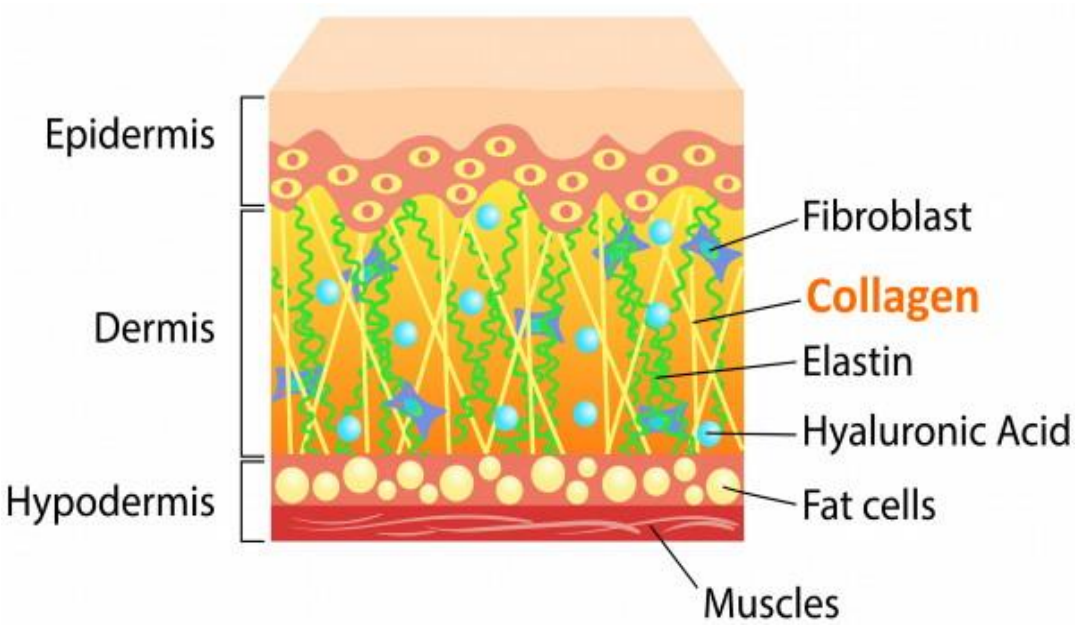
Protein Structure, Function and Crystallography

Prof. Arun Goyal

Dept. of Biosciences and Bioengineering

Collagen

- Collagen is a major component of connective tissue present in skin, tendons, ligaments, teeth and bones where it performs a wide variety of structural roles.
- Collagens have the structure of a triple helix assembled from three polypeptide chains.
- In mature adult, the collagen fibers are extremely robust and insoluble. The insolubility of collagen was for many years a barrier to its chemical characterization, until it was realized that the tissues of younger animals contained a greater proportion of collagen with higher solubility.



Structure of skin

Types of Collagen

Four major classes of collagen are known as summarized in the Table below.

Type I collagen consists of two identical chains called α_1(I) chains and a third chain called α_2.

In contrast Type II collagen contains three identical α_1 chains.

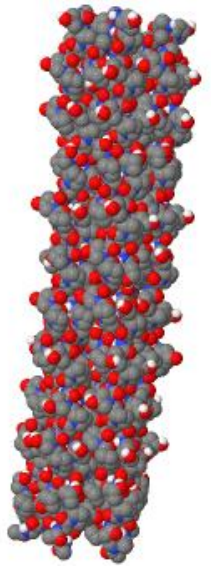
Type	Function and location
Type I	The chief component of tendons, ligaments and bones
Type II	Represents over 50% of the protein in cartilage. It is also used to build the notochord of vertebrate embryos
Type III	Strengthens the walls of hollow structures like arteries, the intestine and the uterus
Type IV	Forms the basal lamina (sometimes called a basement membrane) of epithelia. For example, a network of Type IV collagen provides the filter for the blood capillaries and the glomeruli of the kidneys

The structure of Collagen

- **Tropocollagen is the structural unit of collagen fiber. It is regarded as the subunit from which, the polymeric collagen materials are made.**
- **Tropocollagen is a triple helix of three similarly sized polypeptide chains each of average about 1000 amino acid residues in length. This leads to an approximate M_r of 285 kDa, an average length of 300 nm and a diameter of 1.4 nm.**
- **The tropocollagens are unusual in their amino acid composition as they have high proportions of glycine and proline residues. Collagen has a repetitive primary sequence in which every third residue is glycine. The sequence of the polypeptide chain can therefore be written as**

  -Gly-Xaa-Yaa-Gly-Xaa-Yaa-Gly-Xaa-Yaa-

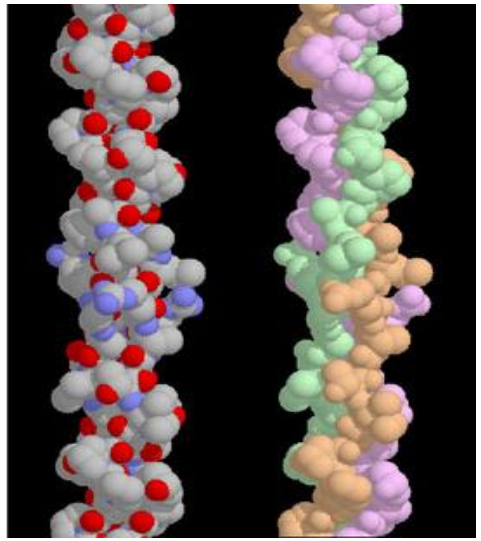
  where Xaa and Yaa are any other amino acid residue. However, analysis of collagen sequences reveals Xaa and Yaa as proline or lysine.



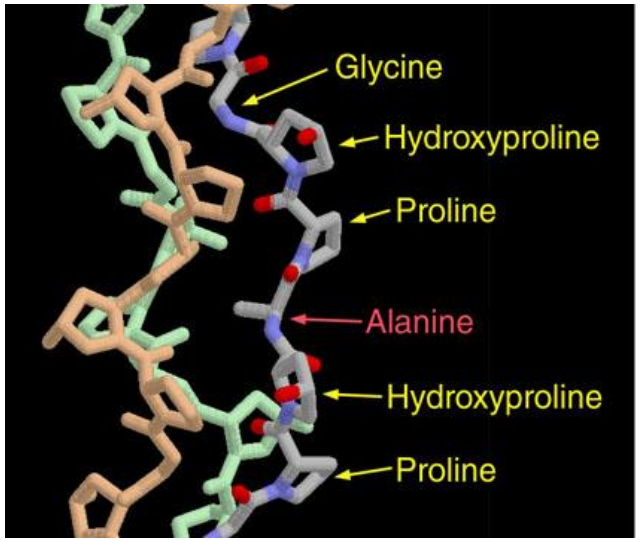
Structure of collagen

The structure of Collagen

- Many of proline and lysine residues are hydroxylated via post-translational enzymatic modification to yield hydroxyproline (Hyp) and hydroxylysine (Hyl), respectively.
- The sequence Gly-Pro-Hyp occurs frequently in collagen. The existence of repetitive sequence is a feature of collagen, keratin and fibroin proteins and is in marked contrast to globular proteins where repetitive sequences are the exception.



Space filling model of Triple helix



Wireframe view of three chains

Fig. The basic structure of the triple helix of collagen. The wireframe view reveals regular repeating Proline residues in all three chains

The structure of Collagen

- Each polypeptide chain intertwines with the remaining two chains to form a triple helix. Each chain has the sequence Gly-Xaa-Yaa and forms a left-handed super helix with the other two chains.
- The rise or translation distance per residue for each chain in the triple helix is 0.286 nm whilst the number of residues per turn is 3.3. Combining these two figures yields a value of ~0.95 nm for the helix pitch (rise per turn, 0.286 x 3.3 = 0.95 nm).
- Glycine lacking a chiral center and possessing considerable conformational flexibility presents a significant contrast to proline.
- In proline conformational restraint exists as a result of the limited variation in the torsional angle (ϕ) permitted by a cyclic ring.

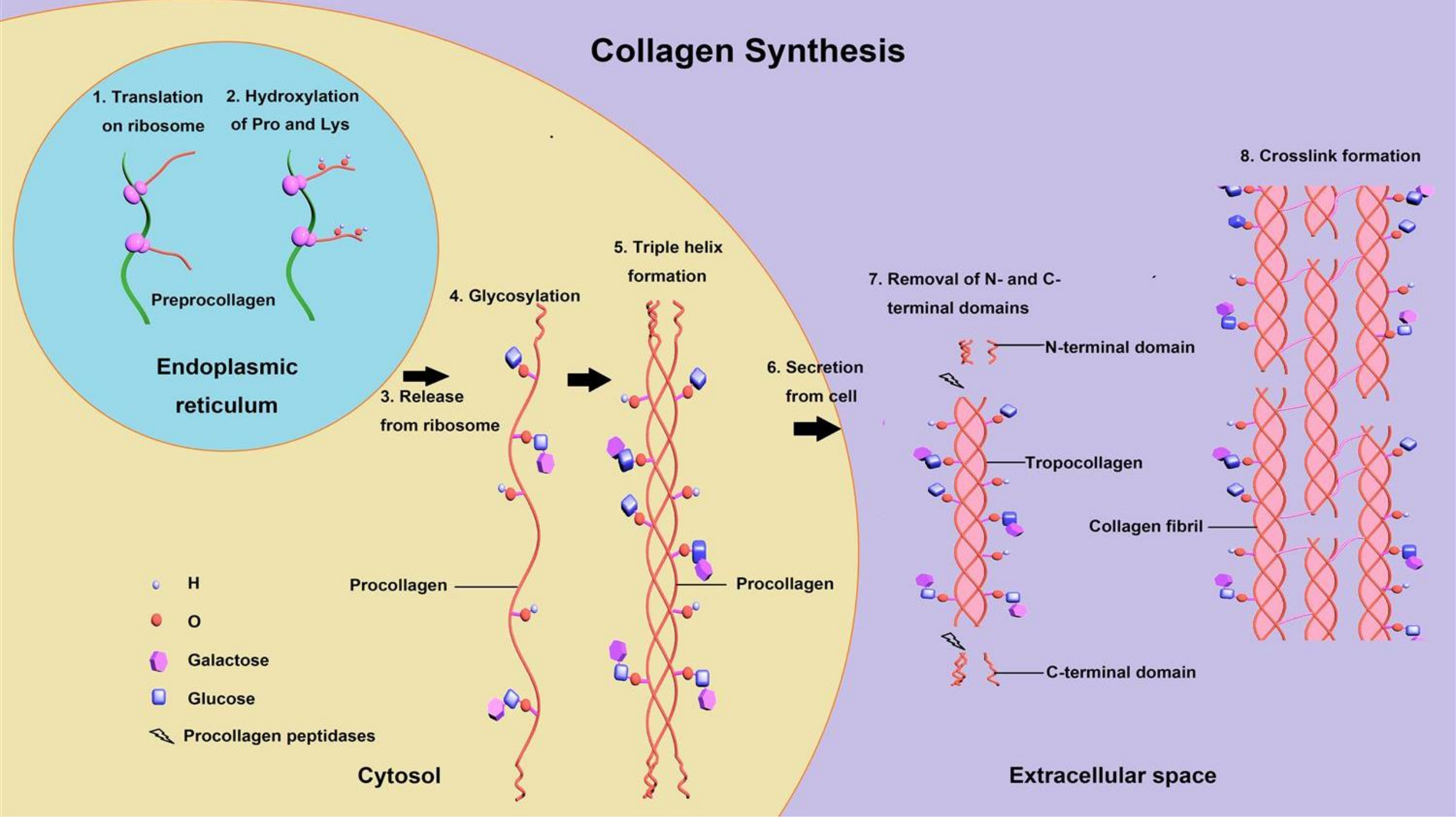
The structure of Collagen

- Since the glycine residues are located at every third position and make contact with the two remaining polypeptide chains, it is clear that only a very small side chain (i.e., glycine) can be accommodated at this position. Any side chain bigger than hydrogen would disrupt the conformation of the triple helix.
- As a result, there is very little space along the helix axis of collagen and glycine is always the residue closest to the helix axis. The side chains of proline residues along with lysine and other residues are on the outside of the helix.
- The close packing of chains clearly stabilizes the triple helix through van der Waals interactions, but in addition extensive hydrogen bonding occurs between polypeptide chains. The hydrogen bonds form between the amide (NH) group of one glycine residue and the carbonyl (C=O) group of residues Xaa on adjacent chains. The direction of hydrogen bonds is transverse or across the long axis of the helix. Interactions within the triple helix are further enhanced by hydrogen bonding between (CO-NH) amide groups and the hydroxyl groups of Hyp residues.

Synthesis of Collagen

- Collagen is synthesized as a precursor termed procollagen in which additional domains at the N and C terminal specifically modulate the folding process. However mature collagen lacks these domains so any unfolding that occurs is difficult to reverse.
- The tropocollagen triple helix, as a structural component associate together as a part of collagen fibre.
- Each tropocollagen molecule is approximately 300 nm in length and packs together with neighbouring molecules to produce a characteristic banded appearance of fibres.
- The banded appearance arises from the overlapping of each triple helix by approximately 64 nm thereby producing the striated appearance of collagen fibrils.
- This pattern of association relies on further cross-linking both within individual helices, known as *intra*molecular cross-links, as well as bonds between helices where they are called *inter*molecular cross-links. Both cross-links are the result of covalent bond formation.
- The covalent cross-links among collagen molecules are derived from lysine or hydroxylysine and involve the action of an enzyme called lysyl oxidase.

Collagen Synthesis



Functions of Collagen

- Cross-linking of collagen is a progressive process but does not occur in all tissues to the same extent.
- In general, younger cells have less cross-linking of their collagen than older cells, as in wrinkled skin in the elderly, as compared with that of a newborn baby.
- It is also the reason why meat from older animals is tougher than that derived from younger individuals.