

## Collagen

by Boyang Zhao

Collagen, a triple helical structure, is a protein is a major component of the connective tissues. It is also the most abundant protein in the human body. So far there have been about 27 types of collagen identified.

Collagen consists of three left-handed polyproline II-like chains supercoiled in a parallel direction with one-residue shift, resulting in a right-handed superhelix. This packing requires glycine (Gly) as the every third residue. As a result, each chain consists of a constant amino acid sequence of Gly-X-Y, where X and Y can be any amino acid residues. About 20% of the amino acids in the collagen structure are the imino acids proline and hydroxyproline.

The substitution of the Gly residues by any other amino acids often lead to diseases such as osteogenesis imperfecta, Ehlers-Danlos syndrome, dystrophic epidermolysis bullosa, Schmid metaphyseal chondrodysplasia, and Alport syndrome as well as others. Extensive research has been done on such mutations through the investigation of short synthetic peptides.

The three peptide chains of collagen are first produced in the endoplasmic reticulum (ER). The C-propeptides then comes together and the three chains start twisting (a process known as propagation) towards the N-terminal ends of the chains. At this stage, the molecule is a procollagen. Cleavage then occurs, where the propeptides are cut off. The resulting collagen then assembles with other collagens, forming fibrils.