Structure of the Organic Matrix: Collagen Structure (Chemical)

D.J. Prockop and C.J. Williams
Dept. of Biochemistry, College of Medicine and Dentistry of New Jersey – Rutgers Medical School
Piscataway, NJ 08854, USA

INTRODUCTION
Collagen is the major constituent of the organic matrix of bone in that it accounts for about 30% of its dry weight and 90 to 95% of its nonmineral content (see (14)). Of the proteins found in nature, collagen is one of the largest. In terms of structure and function, it is among the simplest. In contrast, its biosynthesis is complex in that the initial RNA transcripts of the genes must be extensively processed in order to generate the functional mRNAs which are used for translation. Also, the polypeptides first synthesized on ribosomes must undergo extensive post-translational processing before the protein can be assembled into an extracellular collagen fiber. (For reviews see (2,11,13); additional references cited here are to a few recent papers not included in these reviews.)

THE COLLAGEN MOLECULE
The function of collagen fibers is to hold together and give form to the tissues of the body. Collagen fibers have about the same tensile strength as steel wires and they consist of highly ordered polymers of a rod-like molecule. The molecule itself contains three polypeptide chains, called α chains (Fig. 1). Each of these α chains is coiled into a left-handed helix with about 3 amino acids per turn. The three helical
chains are then twisted around each other into a right-handed superhelix to form a rigid structure similar to a long, thin segment of a rope. Each $\alpha$ chain contains about 1000 amino acid residues and, with the exception of short sequences at the ends of the chains, every third amino acid in each chain is glycine. The molecular formula of an $\alpha$ chain can thus be approximated as $(X-Y-Gly)_{333}$ or $(Gly-X-Y)_{333}$, where $X$ and $Y$ represent amino acids other than glycine. The presence of glycine, the smallest amino acid, in every third position is crucial; the amino acid in this position occupies the restricted space in which the three helical $\alpha$ chains come together in the center of the triple helix. The triple-helical conformation also depends on the presence of proline and hydroxyproline in