strength as steel wires. The three fibrillar collagens are distributed in a tissue-specific manner: type I collagen accounts for most of the protein of dermis, ligaments, tendons, and demineralized bone; type I and type III are the most abundant proteins of large blood vessels; and type II is the most abundant protein of cartilage.

BIOSYNTHESIS AND TURNOVER OF CONNECTIVE TISSUES

Connective tissues are among the most stable components in living organisms, but they are not inert. During embryonic development, connective tissue membranes appear as early as the four-cell blastocyst to provide strength and a structural scaffold for the developing embryo. With the development of blood vessels and skeleton, there is a rapid increase in the synthesis, degradation, and resynthesis of connective tissues. The turnover continues at a slower, but still rapid pace throughout postnatal development and then spikes during the growth spurt of puberty. During adulthood, the metabolic turnover of most connective tissues is slow, but it continues at a moderate pace in bone. With age, malnutrition, physical inactivity, and low gravitational stress, the rate of degradation of most connective tissues, especially in bone and skin, begins to exceed the rate of synthesis and the tissues shrink. In starvation, a large fraction of the collagen in skin and other connective tissues is degraded and provides amino acids for gluconeogenesis (Chap. 97). In both osteoarthritis and rheumatoid arthritis, there is extensive degradation of articular cartilage collagen. Glucocorticoids weaken most tissues by decreasing collagen synthesis. In many pathologic states, however, collagen is deposited in excess. With most injuries to tissues, inflammatory and immune responses stimulate the deposition of collagen fibrils in the form of fibrotic scars. The deposition of the fibrils is largely irreversible and prevents regeneration of normal tissues in hepatic cirrhosis, pulmonary fibrosis, atherosclerosis, and nephrosclerosis.

Structure and Biosynthesis of Fibrillar Collagens The tensile strength of collagen fibers derives primarily from the self-assembly of protein monomers into large fibril structures in a process that resembles crystallization. The self-assembly requires monomers of highly uniform and relatively rigid structure. It also requires a complex series of post-translational processing steps that maintain the solubility of the monomers until they are transported to the appropriate extracellular sites for fibril assembly. Because of the stringent requirements for correct self-assembly, it is not surprising that mutations in genes for fibrillar collagens cause many of the diseases of connective tissues.

The monomers of the three fibrillar collagens are formed from three polypeptide chains, called *α chains*, that are wrapped around each other into a rope-like triple-helical conformation. The triple helix is a unique structure among proteins, and it provides rigidity to the molecule. It also orients the side chains of amino acids in an "inside out" manner relative to most other proteins so that the charged and hydrophobic residues on the surface can direct self-assembly of the monomers into fibrils. The triple-helical conformation of the monomer is generated because each of the α chains has a repetitive amino acid sequence in which glycine (Gly) appears as every third amino acid. Each α chain contains about 1000 amino acids. Therefore, the sequence of each α chain can be designated as $(\text{-Gly-X-Y-})_n$, where X and Y represent amino acids other than glycine and n is >338. The presence of glycine, the smallest amino acid, in every third position in the sequence is critical because this residue must fit in a sterically restricted space in the middle of the helix where the three chains come together. The requirement for a glycine residue at every third position explains the severe effects of mutations that convert any of the glycine residues to an amino acid with a bulkier side chain (see below). Many of the X- and Y-position amino acids are proline and hydroxyproline, which, because of their ring structures, provide additional rigidity to the triple helix. Other X- and Y-positions are occupied by charged or hydrophobic amino acids that precisely direct lateral and longitudinal assembly of the monomers into highly ordered fibrils. Mutations that substitute amino acids in some X- and Y-positions can, in rare instances, also produce genetic diseases.

The fibers formed by the three fibrillar collagens differ in thickness and length, but they have a similar fine structure. As viewed by electron microscopy, they all have a characteristic pattern of cross-striations that are about one-quarter the length of the monomers and reflect the precise packing into fibrils. The three fibrillar collagens, however, differ in sequences found in the X- and Y-positions of the α chains and therefore in some of their physical properties. Type I collagen is composed of two identical α1(I) chains and a third α2(I) chain that differs slightly in its amino acid sequence. Type II collagen is composed of three identical α(II) chains. Type III collagen is composed of three identical α1(III) chains.

To deliver a monomer of the correct structure to the appropriate site of fibril assembly, the biosynthesis of fibrillar collagens involves a large number of unique processing steps (Fig. 427-1). The monomer, first synthesized as a soluble precursor called *procollagen,* contains an additional globular domain at each end. As the proα chains of procollagen are synthesized on ribosomes, the free N-terminal ends move into the cisternae of the rough endoplasmic reticulum. Signal peptides at the N-termini are cleaved, and additional posttranslational reactions begin. Proline residues in the Y-position of the repeating -Gly-X-Y- sequences are converted to hydroxyproline by the enzyme prolyl hydroxylase. The hydroxylation of prolyl residues is essential for the three α chains of the monomer to fold into a triple helix at body temperature. The enzyme requires ascorbic acid as one of its essential cofactors, an observation that explains why wounds fail to heal in scurvy (Chap. 96e). In scurvy, some of the underhydroxylated and unfolded protein accumulates in the cisternae of the rough endoplasmic reticulum and is degraded. Lysine residues in the Y-position are also hydroxylated to hydroxylysine by a separate lysyl hydroxylase. Many of the hydroxylysine residues are glycosylated with galactose or with galactose and glucose. A large mannose-rich oligosaccharide is assembled on the C-terminal propeptide of each chain. The proα chains are assembled by interactions among these C-terminal propeptides that control the selection of the appropriate partner chains to form hetero- or homotrimers and provide the correct chain registration required for subsequent formation of the collagen triple helix. After the C-terminal propeptides assemble the three proα chains, a nucleus of triple helix is formed near the C terminus, and the helical conformation is propagated toward the N terminus in a zipper-like manner that resembles crystallization. The folding into the triple helix is spontaneous in solution, but as discussed below, identification of rare mutations causing OI demonstrated that the folding *in cellulo* is assisted by ancillary proteins. The fully folded protein is then secreted. After secretion, procollagen is processed to collagen by cleavage of the N-propeptides and C-propeptides by two specific proteinases. The release of the propeptides decreases the solubility of the protein about 1000-fold. The entropic energy that is released drives the self-assembly of the collagen into fibrils. Self-assembled collagen fibers have considerable tensile strength, but their strength is increased further by cross-linking reactions that form covalent bonds between α chains in one molecule and α chains in adjacent molecules.

Although the assembly of collagen monomers into fibers is a spontaneous reaction, the process in tissues is modulated by the presence of less abundant collagens (type V with type I, and type XI with type II) and by other components such as a series of small leucine-rich proteins (SLRPs). Some of the less abundant components alter the rate of fibril assembly, whereas others change the morphology of the fibers or their interactions with cells and other molecules.

Collagen fibers are resistant to most proteases, but during degradation of connective tissues, they are cleaved by specific matrix metalloproteinases (collagenases) that cause partial unfolding of the triple helices into gelatin-like structures that are further degraded by less specific proteinases.

OTHER COLLAGENS AND RELATED MOLECULES

The unique properties of the triple helix are used to define a family of at least 28 collagens that contain repetitive -Gly-X-Y- sequences and form triple helices of varying length and complexity. The proteins are