Collagens: the major ECM component
it constitutes about 25% of the proteins in mammals

Type	*Chain(s)*	*Molecular forms*	*Main Distribution*	*Key Features*
I	α1(I) α2(I)	$[\alpha1(I)]_2$ α2(I)	Very widespread: dermis, bone, ligament, tendon, etc	Most abundant collagen type. Main constituent of major fibre bundles that gives strength to connective tissues
		$[\alpha1(I)]_3$	Dermis, dentin	Apparently a minor form
II	α1(II)	$[\alpha1(II)]_3$	Cartilage, intervertebral disc	Main collagen of cartilage: Forms the main fibrils of this tissue
III	α1(III)	$[\alpha1(III)]_3$	Blood vessels, dermis, intestine, etc.	Fibrillar, frequently associated with type I in extensible tissues. More abundant in foetal tissues
IV	α1(IV) α2(IV) α3(IV) α4(IV) α5(IV) α6(IV)	$[\alpha1(IV)]_2$ α2(IV) $[\alpha3(IV)]_2$ α4(IV) (?) $[\alpha5(IV)]_2$ α6(IV) (?) (?)	Basement membranes	A non-fibril forming collagen. Forms a two-dimensional network
V	α1(V) α2(V) α3(V)	$[\alpha1(V)]_3$ $[\alpha1(V)]_2$ α2(V) α1(V) α2(V) α3(V)	Widespread in low quantity; appears associated with collagen I fibrils	Forms fibrils; may form fibril core with collagen I
VI	α1(VI) α2(VI) α3(VI)	α1(VI) α2(VI) α3(VI)	Widespread	Forms beaded filaments. Alternative spliced forms
VII	α1(VII)	$[\alpha1(VII)]_3$	Skin, oral mucosa, cervix	Forms anchoring structure linking epithelial basement membrane to underlying tissue
VIII	α1(VIII) α2(VIII)	$[\alpha1(VIII)]_3$ $[\alpha2(VIII)]_3$ $[\alpha1(VIII)]_2$ α2(VIII)	Associated with endothelial cell layers, e.g. Descemet's membrane	May form a hexagonal lattice in some tissues. Short chain length
IX	α1(IX) α2(IX) α3(IX)	α1(IX) α2(IX) α3(IX)	Cartilage, vitreous body	A collagen II fibril-associated collagen with an interrupted triple helix (FACIT); can contain a glycosaminoglycan chain
X	α1(X)	$[\alpha1(X)]_3$	Hypertrophic mineralising cartilage	A short-chain collagen; has similar structure to collagen VIII
XI	α1(XI) α2(XI) α3(XI)	α1(XI) α2(XI) α3(XI)	Cartilage, intervertebral disc	Forms fibrils which are associated with collagen II
XII	α1(XII)	$[\alpha1(XII)]_3$	Ligament, tendon	A collagen I fibril-associated collagen with an interrupted triple helix (FACIT)
XIII	α1(XIII)	(?)	Widespread in low quantity	Transmembrane collagen with an hydrophobic domain
XIV	α1(XIV)	$[\alpha1(XIV)]_3$	Skin, tendon	A fibril-associated collagen with an interrupted triple helix (FACIT)
XV	α1(XV)	(?)	Expressed in fibroblasts, smooth muscle cells	Triple helix is interrupted in several places
XVI	α1(XVI)	(?)	Expressed in fibroblasts, keratinocytes	A fibril-associated collagen with an interrupted triple helix (FACIT)

Type	*Chain(s)*	*Molecular forms*	*Main Distribution*	*Key Features*
XVII	α1(XVII)	(?)	Bullous pemphigoid antigen, expressed at dermal-epidermal junction	Triple helix is interrupted in several places. Transmembrane collagen with an hydrophobic domain
XVIII	α1(XVIII)	(?)	Expressed in highly vascularised tissues	Triple helix is interrupted in several places
XIX	α1(XIX)	(?)	Expressed in very small amounts by cultured skin fibroblasts and tumour cells	A fibril-associated collagen with an interrupted triple helix (FACIT) with five triple helical subdomains
XX	α1(XX)	(?)	Minor component of several connective tissues, such as sternal cartilage, cornea and tendon and embryonic tissue as corneal epithelium	A fibril-associated collagen with an interrupted triple helix (FACIT)
XXI	α1(XXI)	(?)	Expressed in tissues with a muscle phenotype (heart, skeletal muscle, smooth muscle and placenta)	A fibril-associated collagen with an interrupted triple helix (FACIT)
XXII	α1(XXII)		Hair follicle	Unspecified
XXIII	α1(XXIII)	$[\alpha1(XXIII)]_3$ (?)	Cornea, overexpressed in adenocarcinoma cells; cellular localization	Transmembrane collagen with an hydrophobic domain
XXIV	α1(XXIV)	(?)	Cartilage, retina, cornea, skin	Fibrillar collagen close to collagens type V and XI
XXV	α1(XXV)	(?)	Overexpressed in neurons	Transmembrane collagen with an hydrophobic domain
XXVI	α1(XXVI)	(?)	Specifically expressed in testis and ovary	Unspecified, maybe related to collagen type XIII and XXV
XXVII	α1(XXVII)	(?)	Cartilage, eye, ear, lung and colon	Fibrillar collagen

Vertebrate s have at least 28 collagen Types with 42 distinct alpha chains!!!!

General Collagen Structure

✓ Collagen is a trimeric protein constituted by three left-handed α chains

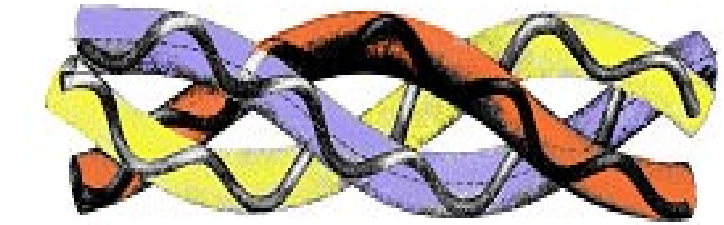
✓The three α chains are supercoiled in a right-handed superhelix

✓Each chain is characterized by the repeating of a tripeptide unit

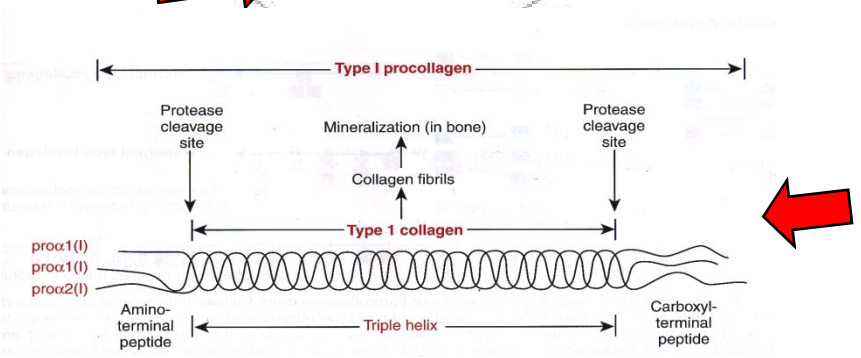
$(Gly\text{-}X\text{-}Y)_n$

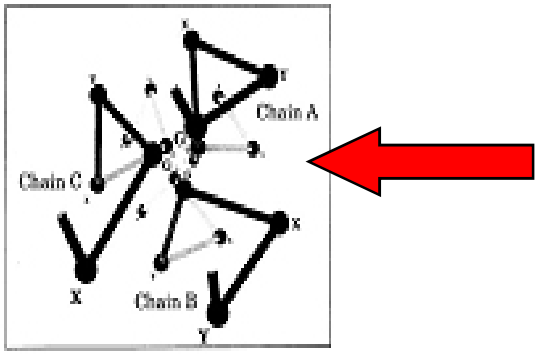
General Collagen Structure

- ✓ **The three α chains can be the same or different (homotrimers or heterotrimers)**
- ✓ Glycine residue every three aminoacidids is necessary for proper folding
- ✓ Regions not folded at triple helix can be removed after synthesis and secretion of the collagen molecules



$(Gly\text{-}X\text{-}Y)_n$





General Collagen Structure

- ✓ The three α chains can be the same or different (homotrimers or heterotrimers)
- ✓ **Glycine residue every three aminoacidids is necessary for proper folding**
- ✓ Regions not folded at triple helix can be removed after synthesis and secretion of the collagen molecules

```
31
-Gly-Pro-Met-Gly-Pro-Ser-Gly-Pro-Arg-
22
-Gly-Leu-Hyp-Gly-Pro-Hyp-Gly-Ala-Hyp-
31
-Gly-Pro-Gln-Gly-Phe-Gln-Gly-Pro-Hyp-
40
-Gly-Glu-Hyp-Gly-Glu-Hyp-Gly-Ala-Ser-
49
-Gly-Pro-Met-Gly-Pro-Arg-Gly-Pro-Hyp—
58
-Gly-Pro-Hyp-Gly-Lys-Asn-Gly-Asp-Asp-
```

Figure 2-45
Biochemistry, Sixth Edition
© 2007 W. H. Freeman and Company

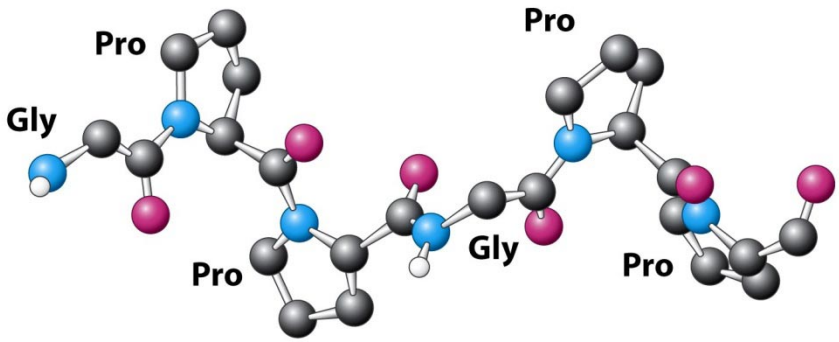
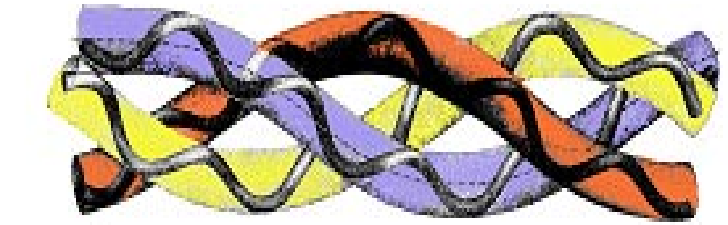


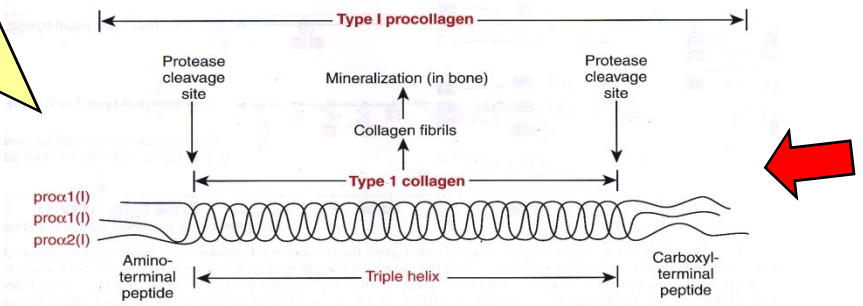
Figure 2-46
Biochemistry, Sixth Edition
© 2007 W. H. Freeman and Company

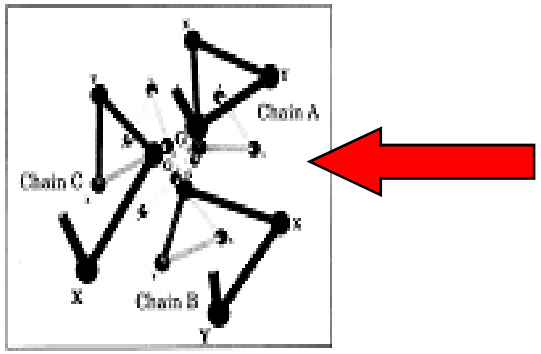
General Collagen Structure

- ✓ The three α chains can be the same or different (homotrimers or heterotrimers)
- ✓ Glycine residue every three aminoacidids is necessary for proper folding
- ✓ **Regions not folded at triple helix can be part of the collagenic molecules or can be removed after synthesis and secretion of the collagen molecules**



$(Gly\text{-}X\text{-}Y)_{II}$





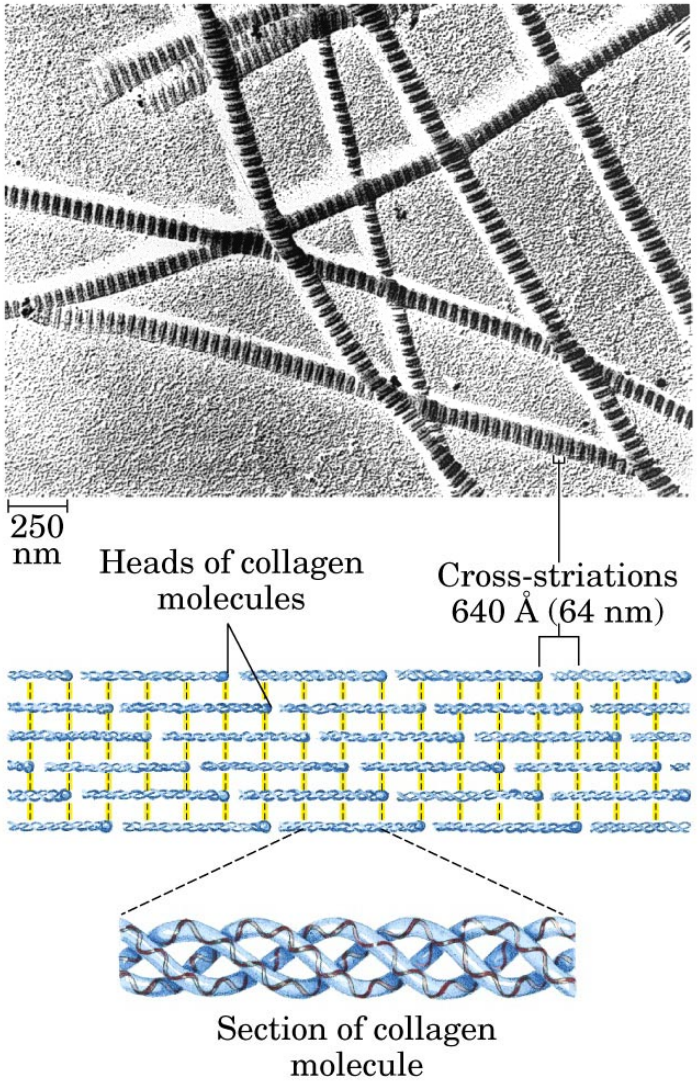
Collagen self assembly

i.e. Collagen Type I that:

✓ generates small fibrils forming a network

✓At TEM reveals a typical banding pattern

✓The regula patter is determined by head to tail assembly of collagen molecules



Collegen molecules stabilization

The α chains of single collagen molecules are stabilized by weak bonds such as H bond, both direct or mediated by water molecules

$(Pro\text{-}Hyp\text{-}Gly)_3$-Ile-Thr-Gly-Ala-Arg-Gly-Leu-Ala-Gly-Pro-Hyp-Gly-$(Pro\text{-}Hyp\text{-}Gly)_3$

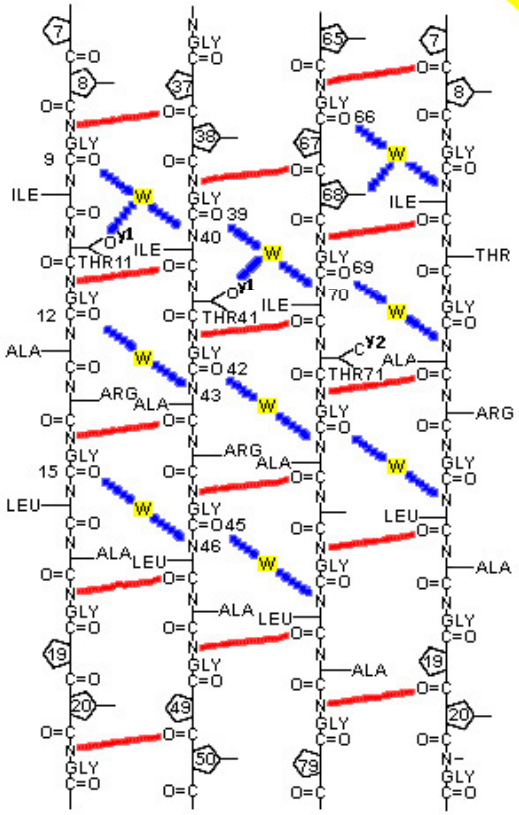
Residues 785-796 of the human sequence of type III collagen

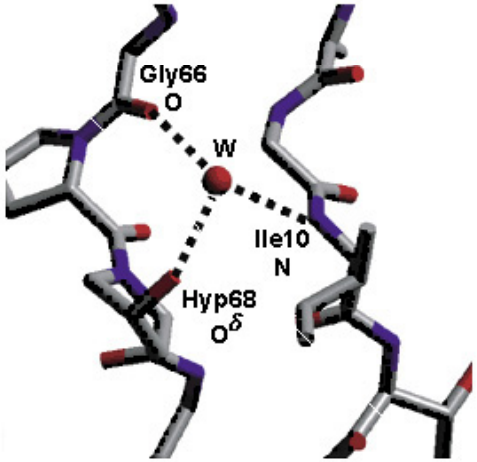
Direct interchain H bonds

(Gly) N-H ⋮ C=O (X)

Water-mediated interchain H bonds

(X) N-H ⋮ W ⋮ C=O (Gly)



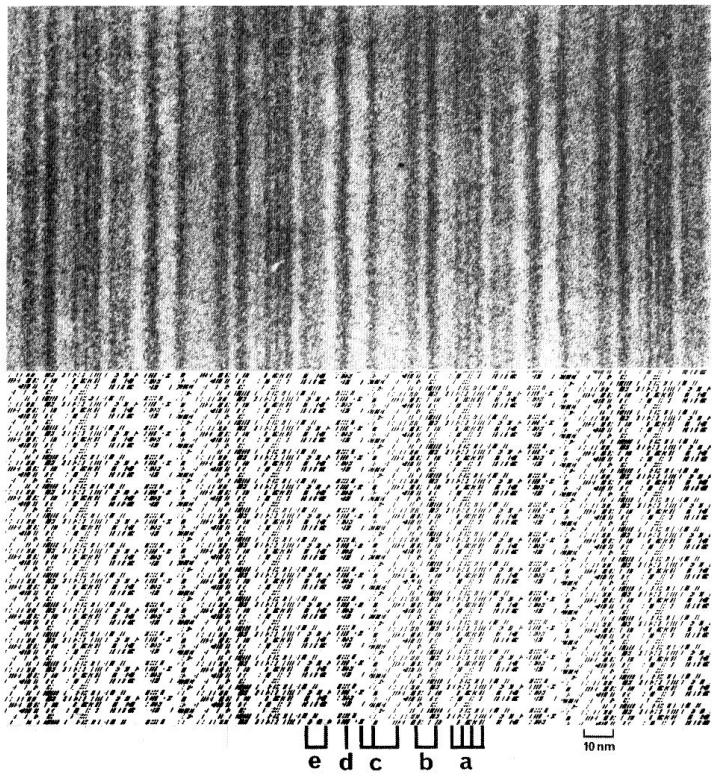


Collagen fibrils satbilization

Collagen fibrils are stabilized by weak bonds such as
✓ interchain ionic interactions
✓Interchain hydrophobic interactions

These bonds allow a regular packing of collagen molecules that can be observed as specific pattern when the fibrils are analyzed by TEM

TEM banding of a collagen molecules after positive staining



From Chapman JA and Hulmes DJS "Electron microscopy of the collagen fibril" in "Ultrastructure of the Connective Tissue Matrix", M. Nijhoff Publ.

Computer digitalization of collagen fibrils in which each ionic residue was marked as a little sign.
The digital reconstruction reveals the banding of collagen fibrils

Collagen thermal stability

Collagen triple helix denaturation happens at ~4-5° above physiological temperature

Collagen proteins and genes nomenclature

Collagen proteins

$\alpha y(Z)$

Z is a roman number indicating the collagen type

The collagen chains are indicated with the greek symbol alpha

Y is a number (from 1 to 6) indicating the particular chain type.It does not have any particular meaning beside that of underlying the similar or different composition of the collagen chains.

Human Collagen genes

COLZAy

Z is an arabic number indicating the collagen type.

Y is an arabic number indicationg the collagen chain type.

i.e. Homotrimeric collagen: $[\alpha 1(II)]_3$, COL2A1, collagen type II
Heterotrimeric collagen $[\alpha 1(I)]_2\alpha 2(I)$, COL1A1 e COL1A2, collagen type I