# Biochemistry of connective tissue

*Ass.Proffesor*

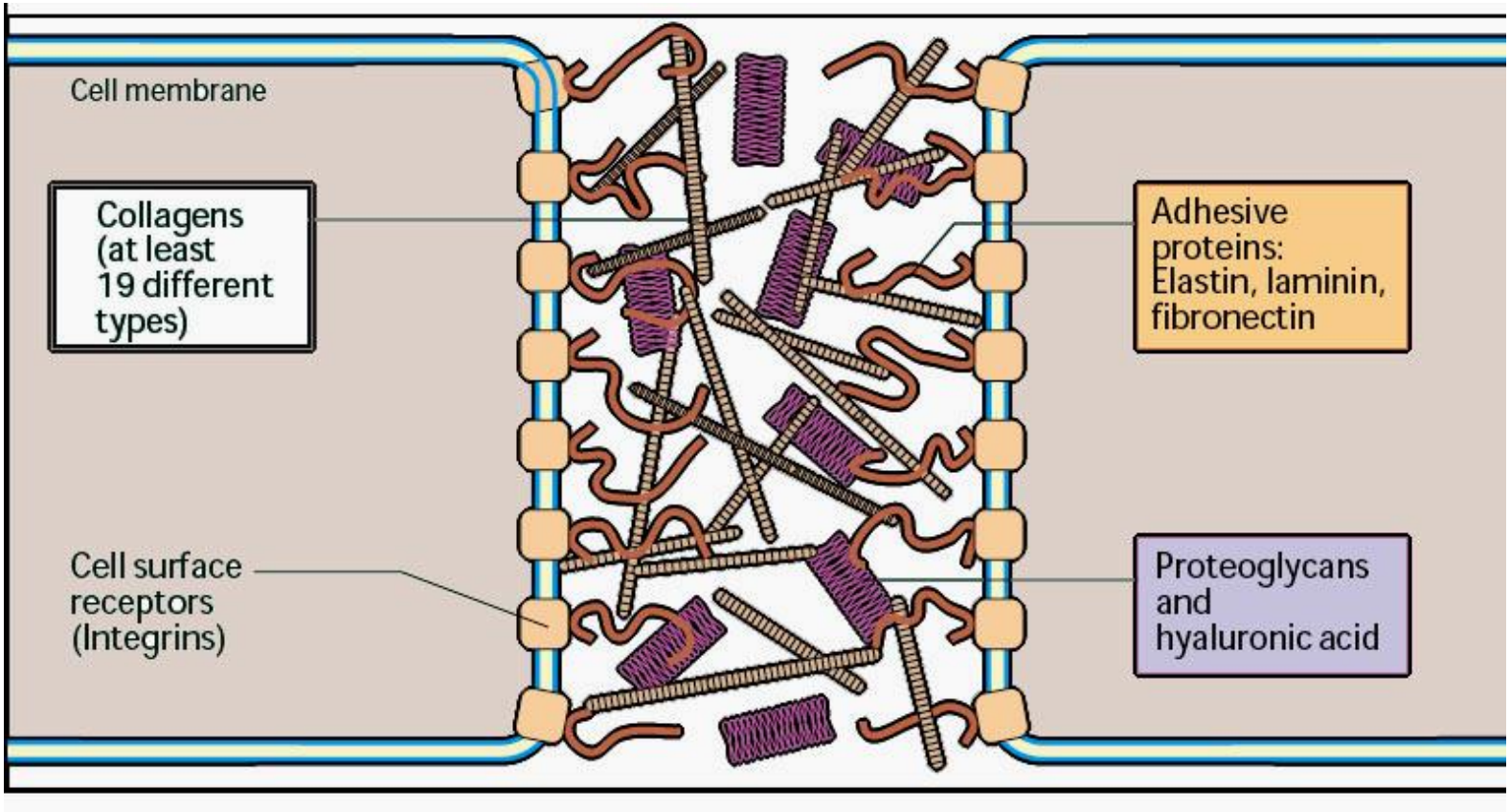
Alexander Naumov

# Extracellular Matrix

# Introduction

- The family of **connective-tissue** cells includes **fibroblasts**, **chondrocytes** (**cartilage cells**), and **osteoblasts** (**bone-forming cells**).
- They are specialized to secrete **extracellular proteins**, particularly ***collagens***, and **mineral substances**, which they use to build up the **extracellular matrix**.
- By contrast, **osteoclasts** dissolve bone matter again by secreting $H^+$ and ***collagenases***.

# Extracellular Matrix




- The **extracellular matrix** (**ECM**) is a complex structural entity surrounding and supporting cells that are found within mammalian tissues.
- The **ECM** is often referred to as the **connective tissue**.

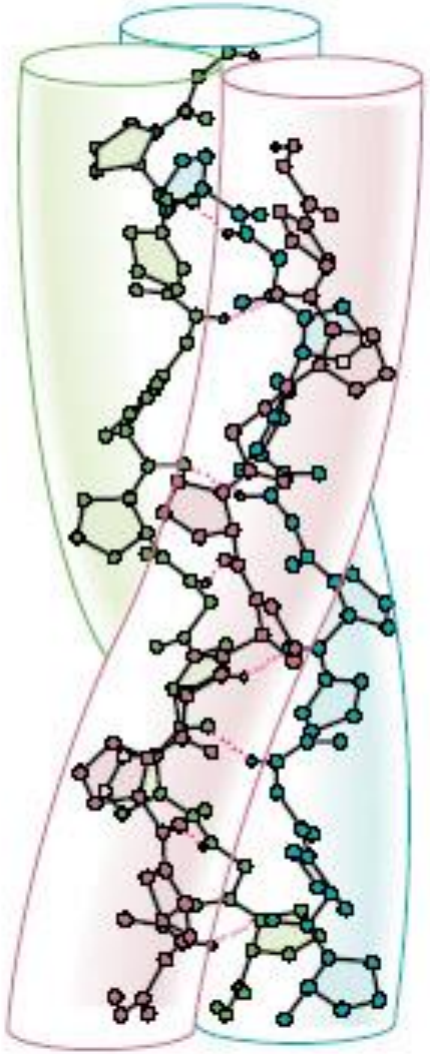
# Extracellular Matrix (cont'd)

- The **ECM** is composed of **3 major classes** of biomolecules:
  1. **Structural proteins**: **collagen** and **elastin**.
  2. **Specialized proteins**: **fibrillin**, **fibronectin**, and **laminin**.
  3. **Proteoglycans**: these are composed of a **protein core** to which is attached long chains of repeating **disaccharide** units termed of **glycosaminoglycans** (**GAGs**) forming extremely complex high molecular weight components of the **ECM**.

# Collagen

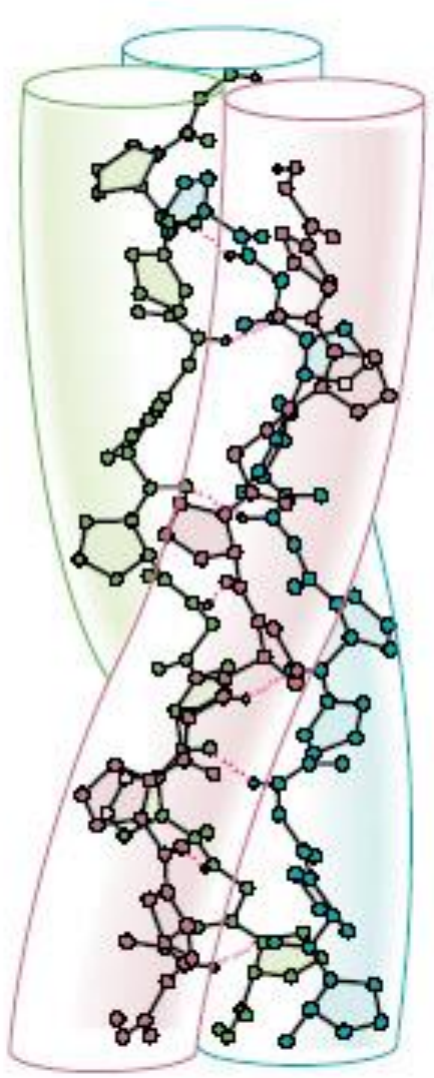
Is the Most Abundant Protein in the Animal World

# Collagens



- **Collagen**, the major component of most connective tissues, constitutes approximately **25%** of the protein of mammals.
- It provides an extracellular framework and exists in virtually **every animal tissue**.
- At least **28** distinct types of **collagen** made up of over **30** distinct polypeptide chains (each encoded by a separate gene) have been identified **in human tissues**.
- A number of proteins - the **C1q component** of the **complement system**, pulmonary **surfactant proteins** SPA and SPD) that are not classified as **collagens** have **collagen-like domains** in their structures.

# Structure of **Collagens**



- The fundamental higher order structure of **collagens** is a long and thin diameter rod-like protein.
  - **Type I collagen** for instance is 300 nm long, 1.5 nm in diameter and consists of 3 coiled subunits composed of two **α1** chains and one **α2** chain.
  - Each chain consists of **1050** AAs wound around each other in a characteristic right-handed triple helix.
    - There are **3** AAs per turn of the helix and **every third amino acid is a Gly**.
  - **Collagens** are also rich in **proline** and **hydroxyproline**.

# Structure of **Collagens**

A striking characteristic of **collagen** is the occurrence of **Gly** residues at every **3** position of the triple helical portion of the **alpha chain**.

This is necessary because **Gly** is the only AA small enough to be accommodated in the limited space available down the central core of the triple helix.

**(Gly-X-Y)n**,

While **X** and **Y** can be any other **AAs**, about **100** of the **X** positions are **Pro** and about **100** of the **Y** positions are **hydroxyproline**.

**Pro** and **hydroxyproline** confer **rigidity** on the **collagen** molecule.

# Structure of **Collagens**

**Hydroxyproline** is formed by the posttranslational hydroxylation of **Pro** residues catalyzed by the enzyme ***prolyl hydroxylase***, whose cofactors are **ascorbic acid** (**vitamin *C***) and **α-ketoglutarate**.

**Lysines** in the **Y** position may also be posttranslationally modified to **hydroxylysine** through the action of ***lysyl hydroxylase***, an enzyme with similar cofactors. Some of these **hydroxylysines** may be further modified by the addition of **galactose** or **galactosyl-glucose** through an **O-glycosidic** linkage, a **glycosylation site** that is **unique** to **collagen**.

# Structure of **Collagens**

**Collagen fibers** are further stabilized by the formation of **covalent cross-links**, both **within and between** the triple helical units.

These **cross-links** form through the action of ***lysyl oxidase***, a **$Cu^{2+}$**-dependent enzyme that oxidatively deaminates the ε-amino groups of certain **Lys** and **hydroxylysine** residues, yielding reactive aldehydes. Such aldehydes can form **aldol condensation products** with other **Lys** or **hydroxylysine**-derived aldehydes or form **Schiff bases** with the ε -amino groups of unoxidized **Lys** or **hydroxylysines**. These reactions, after further chemical rearrangements, result in the stable **covalent cross-links** that are important for the tensile strength of the fibers.

**Histidine** may also be involved in certain cross-links.

# Types of Collagen

| Type | Genes | Tissue |
|---|---|---|
| I | *COL1A1, COL1A2* | **Most connective tissues, including bone** |
| II | COL2A1 | **Cartilage**, **vitreous humor** |
| III | COL3A1 | **Extensible connective tissues such as skin, lung, and the vascular system** |
| IV | COL4A1–COL4A6 | **Basement membranes** |
| V | COL5A1–COL5A3 | Minor component in tissues containing collagen I |
| VI | COL6A1–COL6A3 | **Most connective tissues** |
| VII | COL7A1 | Anchoring fibrils |
| VIII | COL8A1–COL8A2 | **Endothelium**, other tissues |
| IX | COL9A1–COL9A3 | Tissues containing collagen II |

# Types of Collagen

| Type | Genes | Tissue |
|---|---|---|
| X | COL10A1 | Hypertrophic cartilage |
| XI | COL11A1, COL11A2,COL2A1 | Tissues containing collagen II |
| XII | COL12A1 | Tissues containing collagen I |
| XIII | COL13A1 | **Many tissues** |
| XIV | COL14A1 | Tissues containing collagen I |
| XV | COL15A1 | **Many tissues** |
| XVI | COL16A1 | **Many tissues** |
| XVII | COL17A1 | Skin hemidesmosomes |
| XVIII | COL18A1 | **Many tissues** (eg, liver, kidney) |
| XIX | COL19A1 | Rhabdomyosarcoma cells |

# Classification of Collagen

based primarily on the structures that they form

| Class | Type |
|---|---|
| **Fibril-forming** | I, II, III, V, and XI |
| **Network-like** | IV, VIII, X |
| **FACITs1** (fibril-associated collagens with interrupted triple helices) | IX, XII, XIV, XVI, XIX |
| **Beaded filaments** | VI |
| **Anchoring fibrils** | VII |
| **Transmembrane domain** | XIII, XVII |
| Others | XV, XVIII |

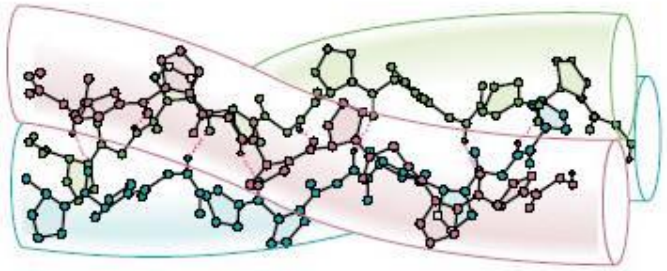
# Formation of **Collagen Fibrils**

In some **collagens**, the entire molecule is triple helical, whereas in others the triple helix may involve only a fraction of the structure.

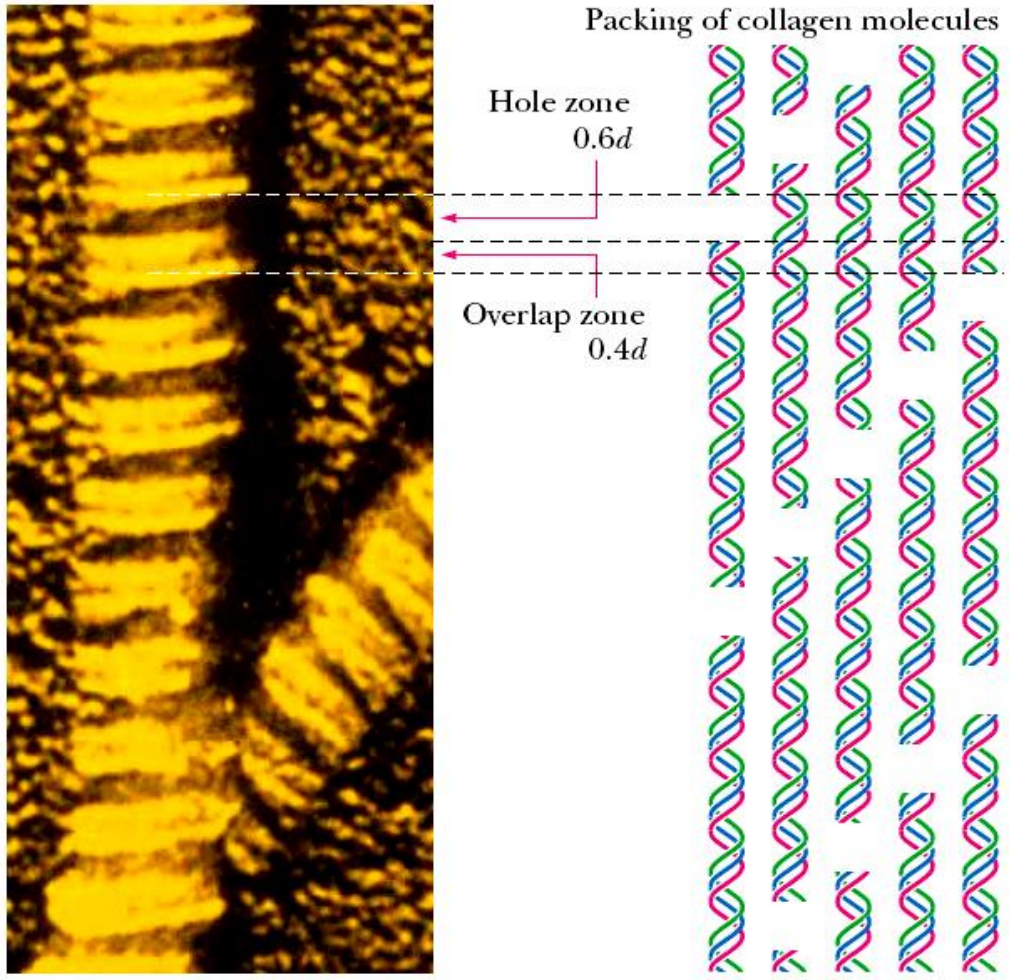
Mature **collagen** **type I**, containing approximately **1000 AAs**;

- in it each polypeptide subunit or **alpha chain** is twisted into a left-handed polyproline helix of three residues per turn.

**Three** of these **alpha chains** are then wound into a right-handed **superhelix**, forming a **rodlike** molecule 1.4 nm in diameter and about 300 nm long.

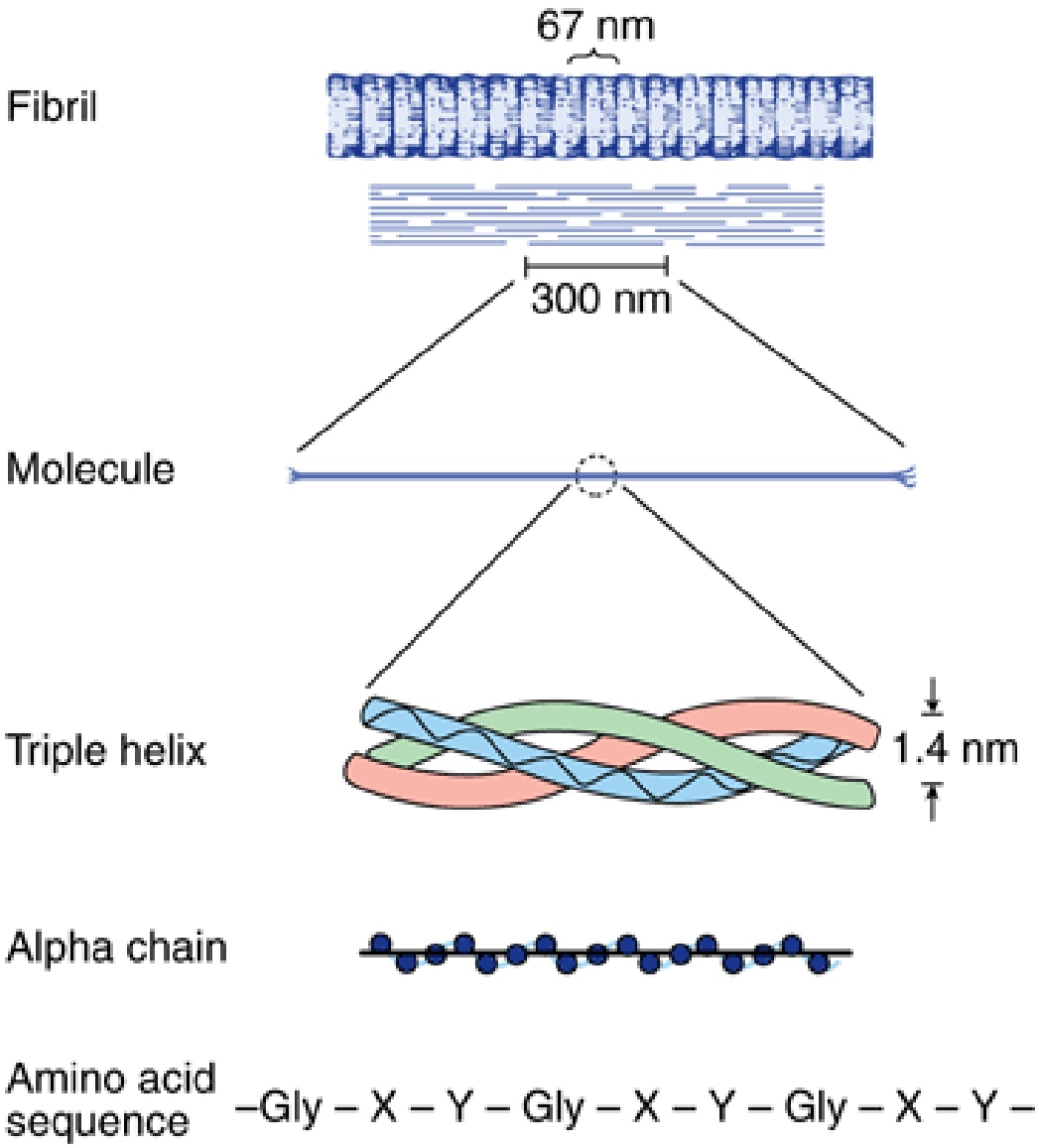


# Formation of **Collagen Fibrils**




- Lateral interactions of triple helices of **collagens** result in the formation of fibrils roughly 50 nm diameter.
- The packing of **collagen** is such that **adjacent** molecules are displaced approximately **1/4 of their length** (67 nm).
  - This **staggered array** produces a **striated effect** that can be seen in the electron microscope.

# Formation of Collagen Fibrils



Collagen
undergoes extensive posttranslational modifications

# **Collagen** processing (1)

- ***Collagens*** are synthesized as longer precursor proteins called ***pre-procollagens***.
- The signal sequence (***pre-***) is removed and numerous modifications take place in the ***procollagen*** chains.
  - Specific **Pro** residues are hydroxylated by ***prolyl 4-hydroxylase*** and ***prolyl 3-hydroxylase***.
  - Specific **Lys** residues also are hydroxylated by ***lysyl hydroxylase***.
  - Both ***prolyl hydroxylases*** are **absolutely dependent upon vit *C*** ($Fe^{2+}$) as co-factor.
  - **Glycosylations** of the O-linked type also occurs during Golgi transit.

# Collagen processing (2)

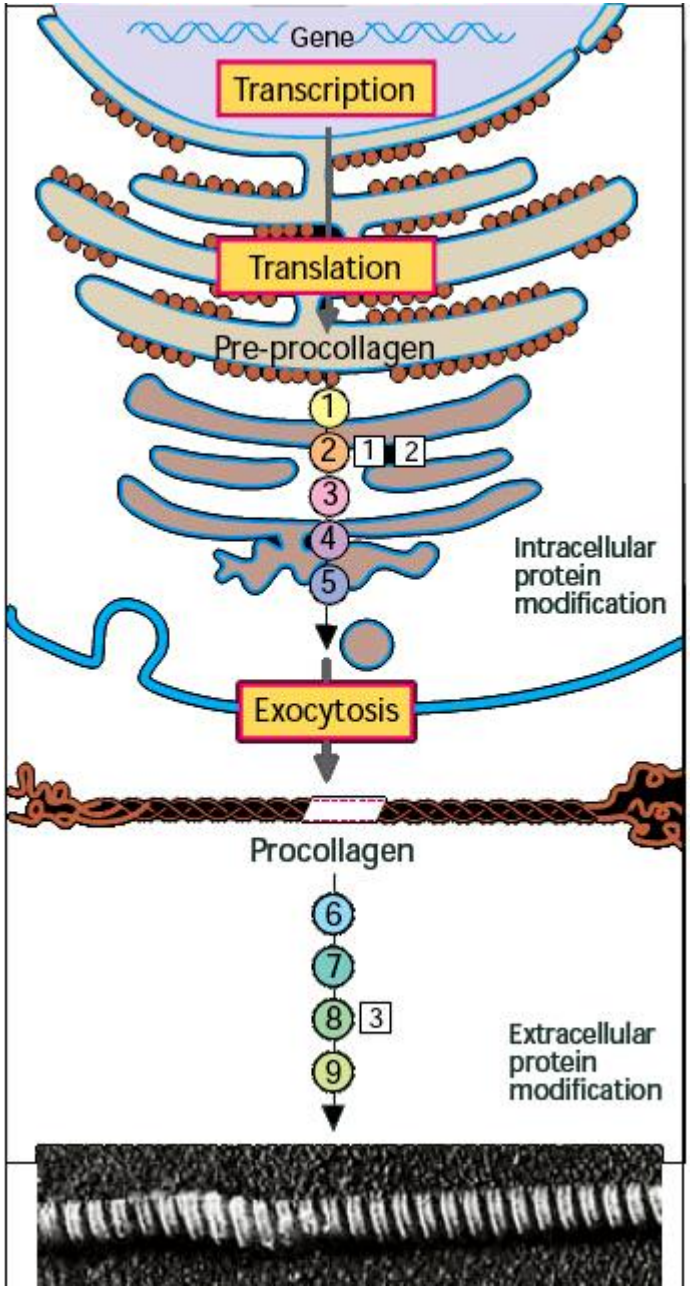
- **Procollagen** contains an additional **150** AAs at the N-terminus and **250** at the C-terminus (**pro-domains**).
  - These **pro-domains** are globular and form multiple intrachain **disulfide bonds**.
  - The disulfides stabilize the **proprotein** allowing the triple helical section to form.

**Collagen fibers** begin to assemble in the **ER** and **Golgi complexes**.

# **Collagen** processing (3)

- Following completion of processing the **procollagens** are secreted into the **extracellular space** where extracellular enzymes remove the **pro**-domains.
  - The **collagen** molecules then polymerize to form **collagen fibrils**.
  - Accompanying **fibril** formation is the oxidation of certain **Lys** residues by the extracellular enzyme ***lysyl oxidase*** foming reactive aldehydes.
  - These **reactive aldehydes** (**allysine**) form specific cross-links between two chains thereby, stabilizing the staggered array of the **collagens** in the fibril.

# Biosynthesis of **Collagen**: Overview




1. Removal of the prepeptide
2. Hydroxylation of Pro and Lys residues
3. Glycosylation of 5Hyl and Asn
4. Oxidation of Cys in propeptides
5. Assemblage to form triple helix
6. Removal of the propeptide
7. Staggered deposition to form fibrils
8. Oxidation of Lys and 5Hyl to aldehydes
9. Cross-linking to form supramolecules

[1] Procollagen-proline 4-dioxygenase 1.14.11.2 [ascorbate, Fe]

[2] Procollagen-lysine 5-dioxygenase 1.14.11.4 [ascorbate, Fe]

[3] Protein-lysine 6-oxidase 1.4.3.13 [Cu]

**Koval A., 2009**

# Biosynthesis of Collagen

**Intracellular**

1. Cleavage of signal peptide

2. Hydroxylation of **Pro** residues and some **Lys** residues; glycosylation of some **hydroxylysyl** residues

3. Formation of intrachain and interchain S–S bonds in extension peptides.

4. Formation of triple helix.

# Biosynthesis of Collagen

**Extracellular**

1. Cleavage of amino and carboxyl terminal **pro-**peptides.

2. Assembly of **collagen** fibers in quarter-staggered alignment.

3. Oxidative deamination of ε-amino groups of **Lys** and **hydroxylysyl** residues to aldehydes.

4. Formation of intra- and interchain cross-links via Schiff bases and aldol condensation products

**Fibronectins**
is a major glycoprotein of the extracellular matrix

# Fibronectins

**Fibronectin is an important glycoprotein involved in cell adhesion & migration**

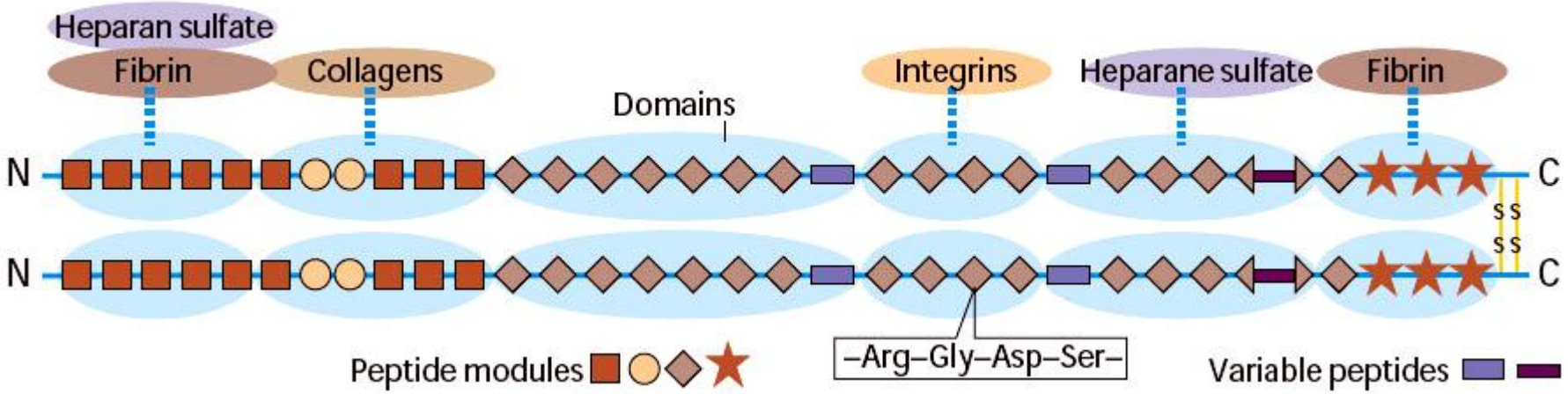
The same cells that secrete **collagen** also secrete **fibronectin**, a large glycoprotein present on cell surfaces, in the extracellular matrix, and in blood.

**Fibronectin** binds to aggregating **precollagen** fibers and alters the kinetics of fiber formation in the pericellular matrix.

Associated with **fibronectin** and **procollagen** in this matrix are the **proteoglycans heparan sulfate** and **chondroitin sulfate**.

**Such interactions may serve to regulate the formation of collagen fibers and to determine their orientation in tissues.**

# Fibronectins




- ***Fibronectins*** contain 6-8 tightly folded **domains** each with a high affinity for a different substrate such as **heparan sulfate**, **collagen** (separate domains for types I, II and III), **fibrin, DNA,** and **cell-surface receptors**.

***Fibronectin*** is a major glycoprotein of the extracellular matrix, also found in a soluble form in plasma. It consists of two identical subunits joined by two disulfide bridges near their carboxyl terminals.

# The Role of **Fibronectins**

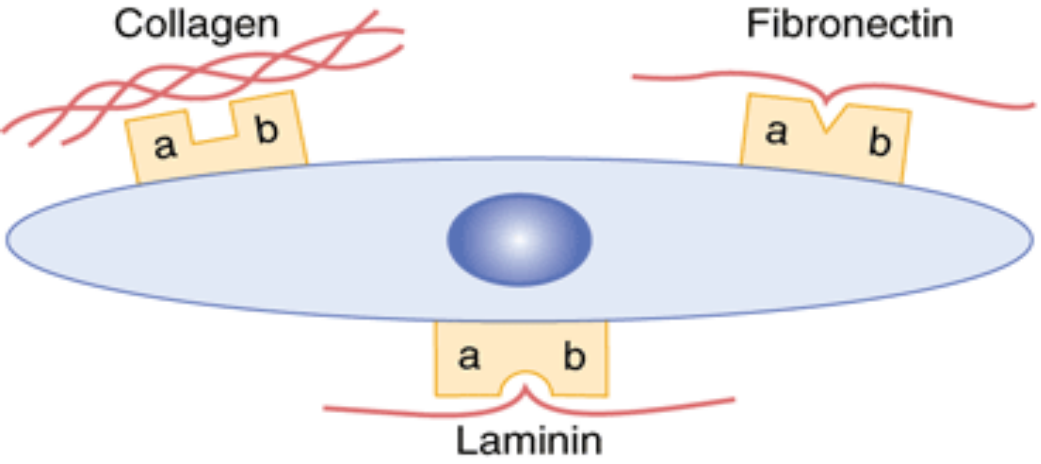
- ***Fibronectin*** attaches cells to **all extracellular matrices** except **collagene type IV** that involves ***laminin*** as the adhesive molecule.
- At least 20 different ***fibronectin*** chains have been identified that arise by alternative RNA splicing of the primary transcript from a single ***fibronectin*** **gene**.

# Integrins

**Fibronectin** contains an **Arg-Gly-Asp** (RGD) sequence that binds to the **receptor**.

This RGD sequence is shared by a number of other proteins present in the ECM that bind to **integrins** present in cell surfaces.

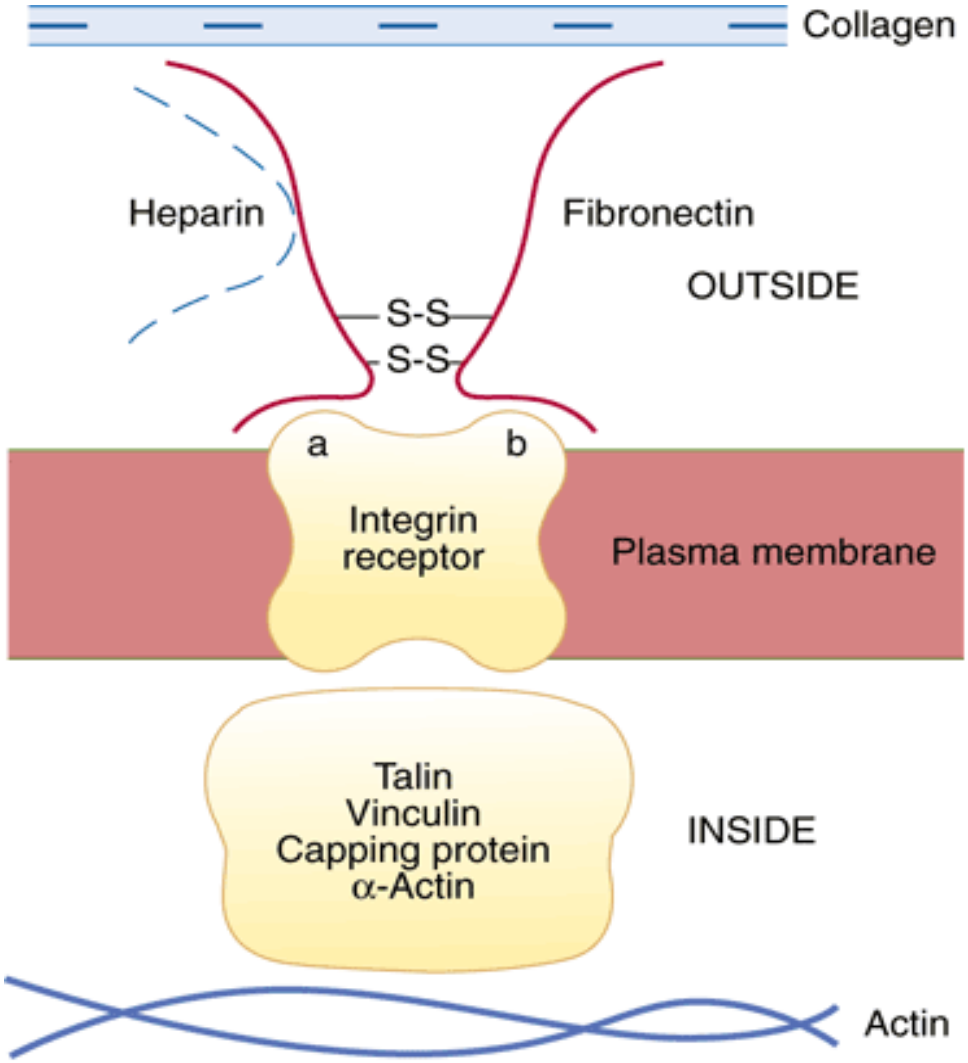
The **integrins** are heterodimers, containing various types of **α-** and **β**-polypeptide chains.




A cell interacting through various **integrin receptors** with **collagen**, **fibronectin**, and **laminin** present in the **ECM**.

# Fibronectin interacting with an integrin fibronectin receptor.

Schematic representation of **fibronectin** interacting with an **integrin fibronectin receptor** in the exterior of the **plasma membrane** of a cell of the **ECM** and of various attachment proteins interacting indirectly or directly with an **actin** microfilament in the cytosol.



# Laminin

is a major protein component of renal glomerular & other basal laminas

# Laminin

**Laminin** has potential binding sites for **type IV collagen**, **heparin**, and **integrins** on cell surfaces.

**Laminin** is a major protein component of renal glomerular & other **basal laminas** - specialized areas of the **ECM** that surround epithelial and some other cells. In that structure, the **basal lamina** is contributed by two separate sheets of cells (one endothelial and one epithelial), each disposed on opposite sides of the lamina; these three layers make up the **glomerular membrane**.

The primary components of the **basal lamina** are three proteins—**laminin**, **entactin**, **type IV collagen** — and the **GAG heparin** or **heparan sulfate**.

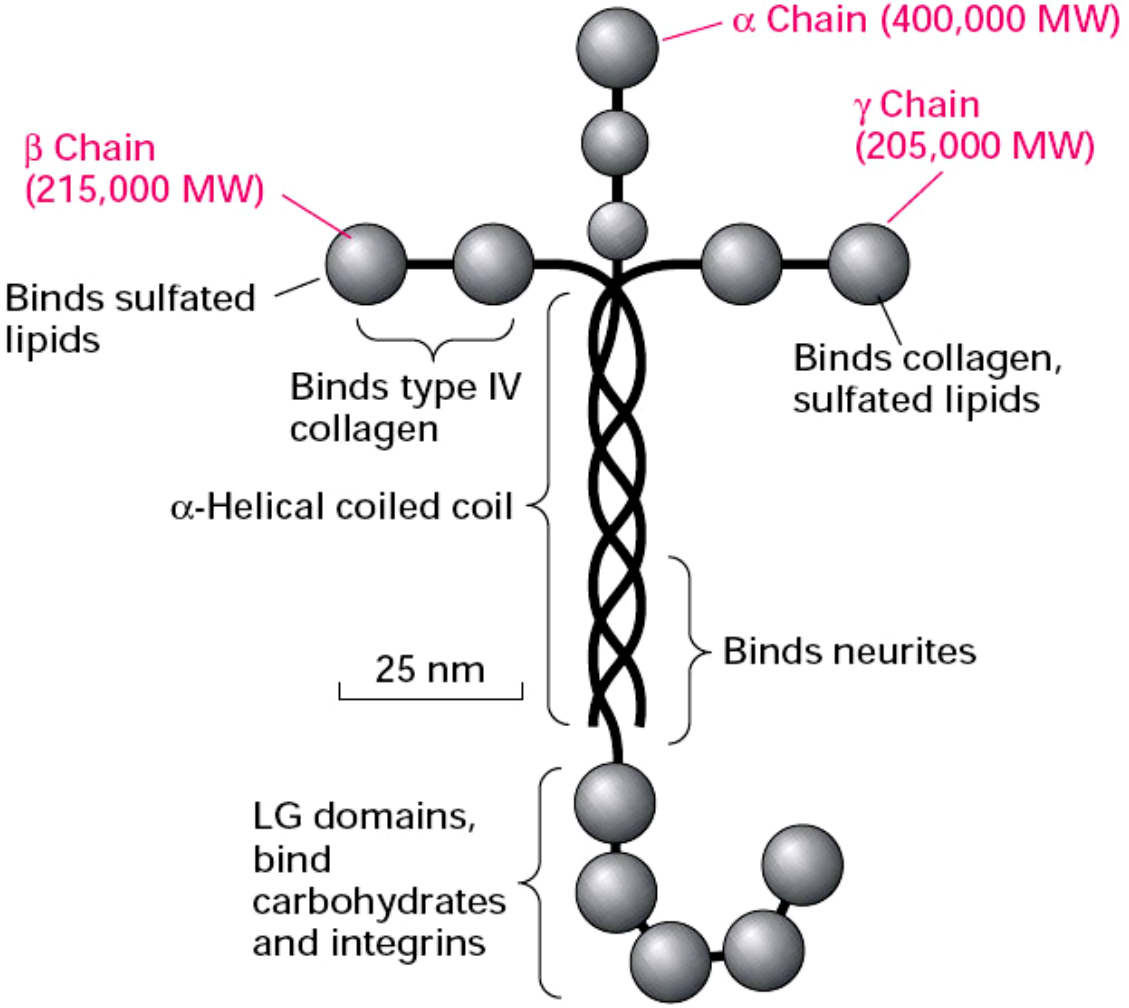
# Basal Lamina Components: Laminin

- All **basal laminae** contain a common set of proteins and GAGs.

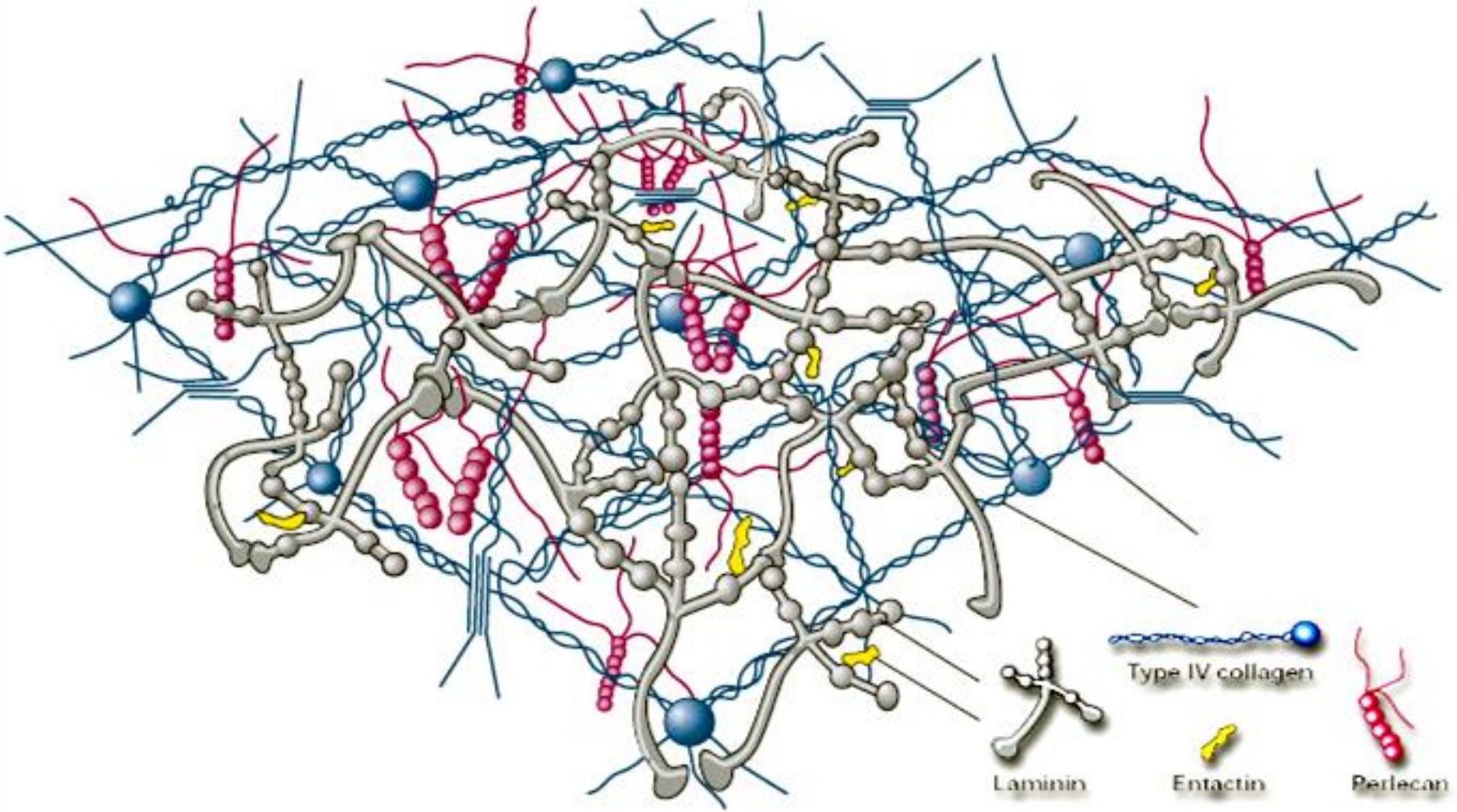
these are

**type IV collagen**, **heparan sulfate proteoglycans**, and **laminin**.

  - **Laminin** anchors cell surfaces to the **basal lamina**.



# Basal Lamina



# Laminin

The **basal lamina** of the **renal glomerulus** has an important role in **glomerular filtration**, regulating the passage of large molecules (most plasma proteins) across the **glomerulus** into the renal tubule.

The normal structure of the glomerulus may be severely damaged in certain types of **glomerulonephritis** (eg, caused by antibodies directed against various components of the glomerular membrane).

This alters the **pores** and the amounts and dispositions of the negatively charged macromolecules, and relatively massive amounts of **albumin** can pass through into the urine, resulting in severe **albuminuria**.

# Elastin

CONFERS EXTENSIBILITY & RECOIL ON LUNG, BLOOD VESSELS & LIGAMENTS

# Elastin

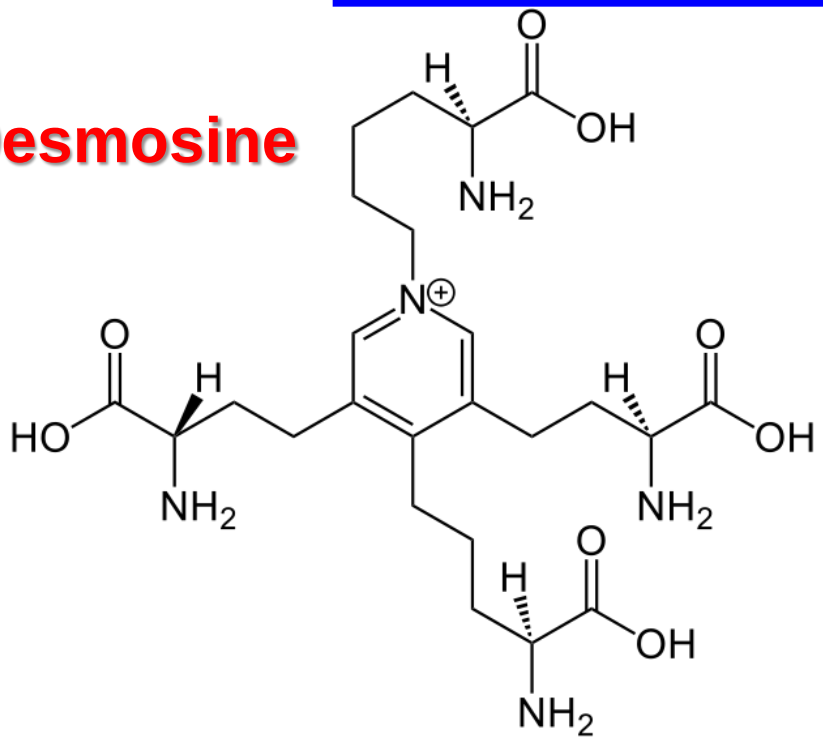
**Elastin** is a highly **elastic protein** in connective tissue and allows many tissues in the body to resume their shape after **stretching** or **contracting**.

Present in large amounts in **lung**, large **arterial blood vessels**, and some elastic **ligaments**. Smaller quantities - are also found in **skin**, **ear cartilage**, and several other tissues.

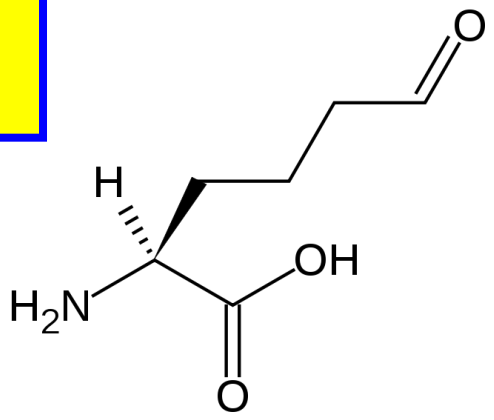
**Elastin** is highly insoluble and extremely stable.

# Elastin

**Desmosine**



**Allysine**



After secretion from the cell, certain **Lys** residues of **tropoelastin** are oxidatively deaminated to aldehydes by **lysyl oxidase**.

The major cross-links formed in **elastin** are the **desmosines**, which result from the condensation of **3** of these **Lys**-derived aldehydes (**allysine**) with an unmodified **Lys** to form a tetrafunctional cross-link unique to **elastin**.

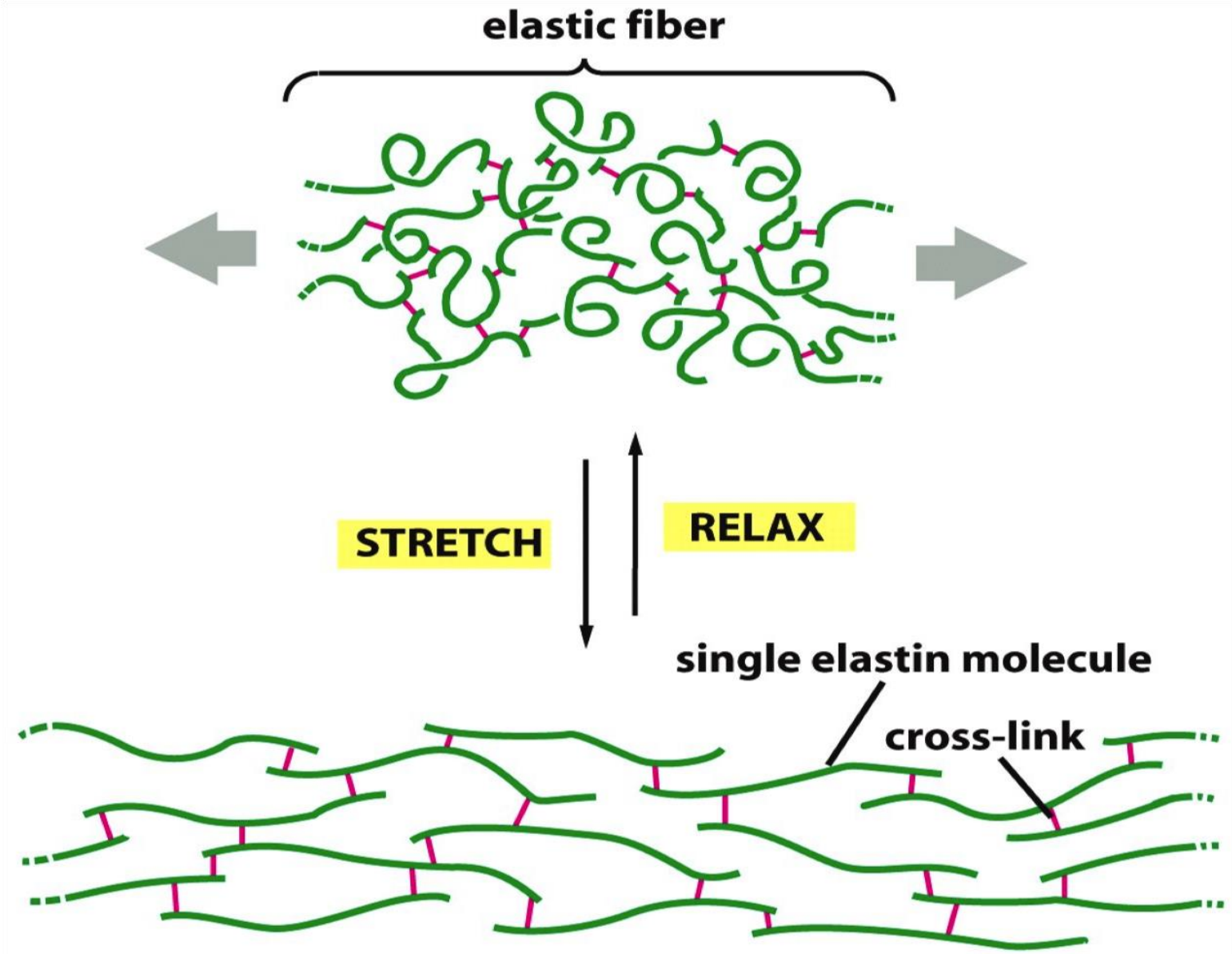



Figure 19-71 Molecular Biology of the Cell 5/e (© Garland Science 2008)

# Elastin

- Deletions in the **elastin** gene have been found in ~ 90% of subjects with the **Williams–Beuren syndrome** - a developmental disorder affecting **connective tissue** and the **central nervous system**.
- The mutations, by affecting synthesis of **elastin** - play a causative role in the **supravalvular aortic stenosis**.
- Fragmentation or a decrease of **elastin** is found in conditions such as:
  - **pulmonary emphysema**,
  - **dermatolysis**,
  - **aging** of the skin.

# PROTEOGLYCANS
# &
# GLYCOSAMINOGLYCANS

# Proteoglycans

**Proteoglycans** are proteins that contain covalently linked **glycosaminoglycans (GAGs).**

At least 30 have been characterized and given names such as:

**syndecan, biglycan, and**
**betaglycan, fibromodulin.**

The proteins bound covalently to **GAGs** are called "**core proteins**".

The amount of carbohydrate in a **proteoglycan** is usually much greater than that found in a **glycoprotein** and may comprise up to **95%** of its weight.

The carbohydrate groups of **proteoglycans** are predominantly GAGs O-linked to **serine** residues.

# Proteoglycans

**Proteoglycans** involve in the **binding of specific proteins** to the **glycosaminoglycan** groups.

The carbohydrate groups of **proteoglycans** are predominantly **glycosaminoglycans** O-linked to **serine** residues.

# Glycosaminoglycans (GAGs)

A **GAG** is an unbranched polysaccharide made up of repeating **disaccharides**, one component of which is always an **amino sugar**, either **D-glucosamine** or **D-galactosamine**.

The other - is a **uronic acid**,

either **L-glucuronic acid** (**GlcUA**)

or **L-iduronic acid** (**IdUA**).

With the exception of **hyaluronic acid**, all the **GAGs** contain **sulfate groups**.

# Glycosaminoglycans (GAGs)

There are at least **7** glycosaminoglycans:

- hyaluronic acid,
- chondroitin sulfate,
- keratan sulfates I and II,
- heparan sulfate, and
- dermatan sulfate.

# The ECM and pathologic processes

The **ECM** has been found to be involved in many processes –

- **in development**,
- **in inflammatory states**,
- **in the spread of cancer cells**.

Involvement of certain components of the **ECM** has been documented in both **rheumatoid arthritis** and **osteoarthritis**.

Several diseases (eg, **osteogenesis imperfecta** and a number of types of the **Ehlers–Danlos syndrome**) are due to genetic disturbances of the synthesis of **collagen**.

Specific components of **proteoglycans** - the **GAGs** - are affected in the group of genetic disorders known as the **mucopolysaccharidoses**.

Changes occur in the **ECM** during the **aging process**.

# Collagen-*Related Diseases*

– The nature and extent of **collagen** cross-linking depends on the age and function of the tissue.

**Collagen** from young animals is predominantly ***uncrosslinked*** and can be extracted in **soluble form**, whereas **collagen** from older animals is highly cross-linked and thus **insoluble**.

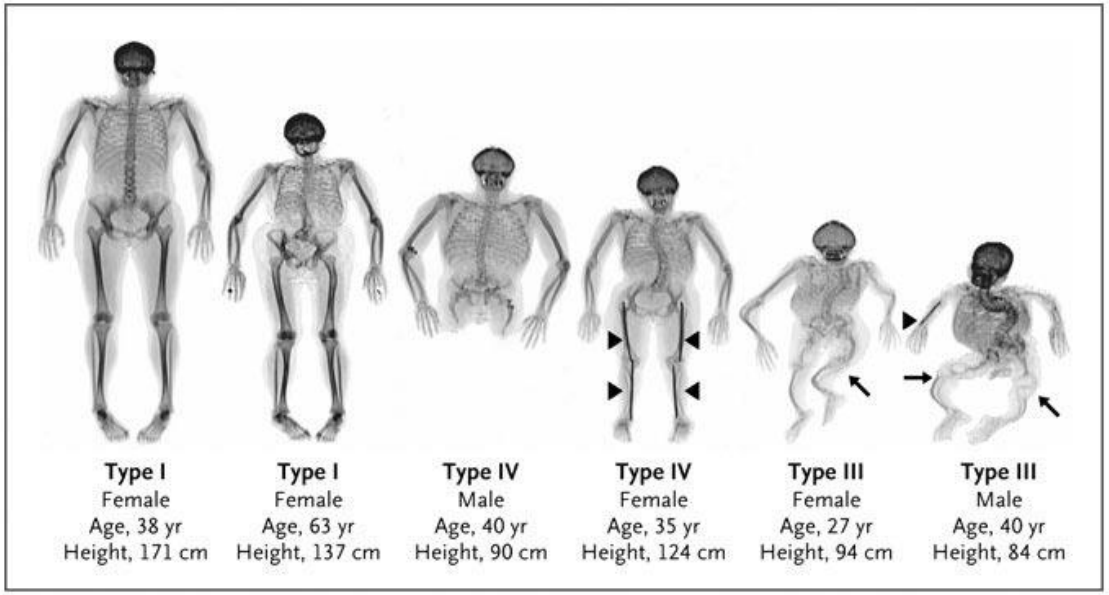
The loss of flexibility of joints with aging is probably due in part to increased cross-linking of **collagen**.

# Collagen Disorders

Alterations in **collagen** structure resulting from abnormal genes or abnormal processing of **collagen** proteins results in numerous diseases:

- **scurvy**,
- **osteogenesis imperfecta**
- **Ehlers-Danlos syndrome**.

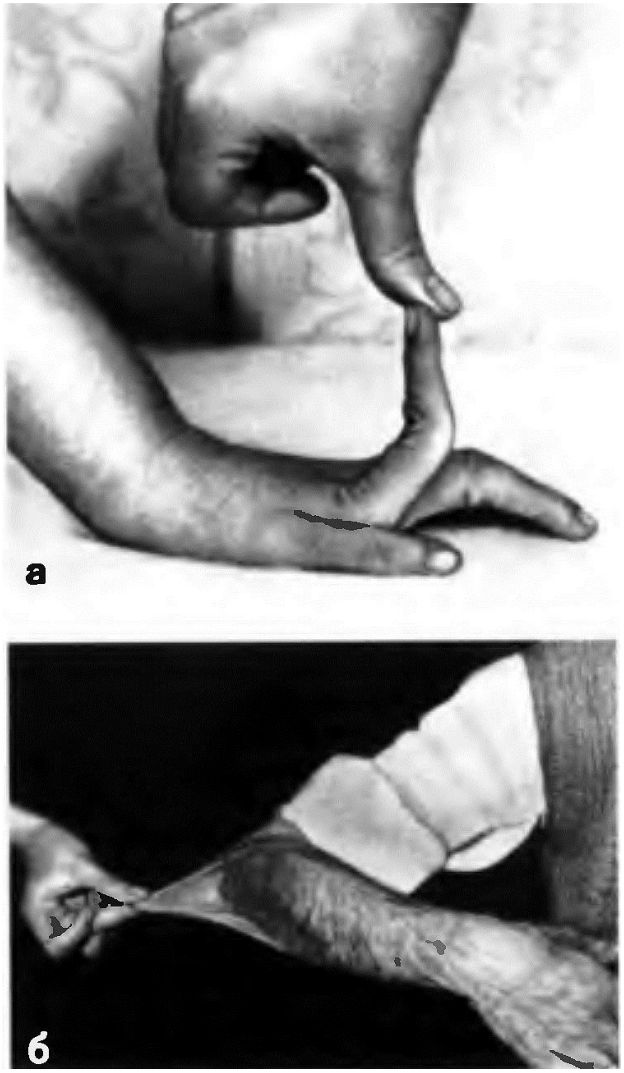
# Osteogenesis Imperfecta


**Osteogenesis imperfecta -** is a congenital bone disorder characterized by **brittle** bones that are prone to fracture ("**the glass man**").

– a deficiency of **Type-I collagen**

# Ehlers-Danlos Syndrome




**Ehlers-Danlos syndrome** is an inherited connective tissue disorder (at least ten distinct disorders ).

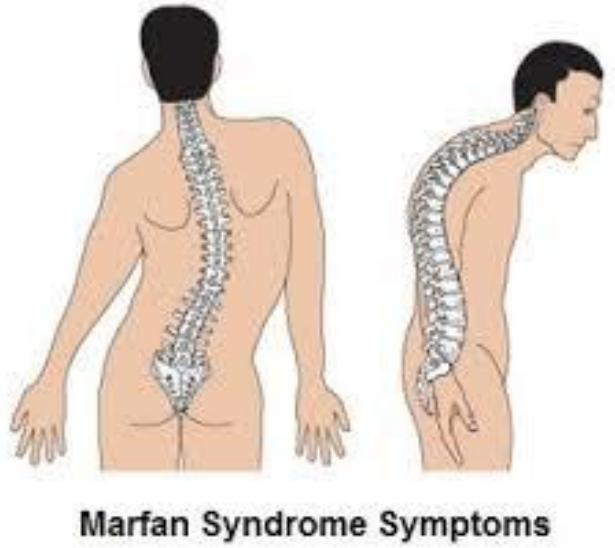
**EDS** is caused by a defect in the

- **structure**,
- **production**, or
- **processing**

of **collagen** or proteins that interact with collagen.

# **Marfan's** syndrome

- **Marfan's syndrome** manifests itself as a disorder of the connective tissue and was believed to be the result of abnormal **collagens**.



Marfan Syndrome Symptoms

However, recent evidence has shown that **Marfan's** results from mutations in the extracellular protein, **fibrillin**, which is an integral constituent of the non-collagenous microfibrils of the **extracellular matrix**.

Thank you for attention