Collagen structure and evolution of multicellular life

Collagens are major extracellular structural proteins accounting for around 25% of proteins in humans. They are found in connective tissues and are a key ingredient of skin, tendons, bones and cartilage. The role of collagen fibres in the extracellular matrix of tissues could be compared to that of cellulose in plant cell walls and the steel bars of reinforced concrete. They act as a tissue scaffolding e. g. in basement membranes, provide tensile strength, facilitate cell adhesion as well as having roles in tissue repair, cell migration and protection. The conservation of certain types of collagen within all animals, from the simplest sponge to ultra complex human beings is indicative of their importance to life. Mutations in collagens may be lethal or result in many severe diseases including, osteoarthritis, osteoarthritis, arterial aneurysms and Alports Syndrome. Inhibition of collagens in sponges has been found to result in tissue regeneration failure.

Collagen Classification and Structure

Different types of vertebrae collagen are distinguished with Roman numerals that are given in the order of their discovery. They are classified as either: fibrillar (I, II, III, V, XI), basement membrane (IV), short chain, facit or other, depending on their function. Some homologies have been identified between collagens found in sponges and fibrillar and basement membrane collagens in animals.

The structure of a collagen subunit (tropocollagen) is analogous to that of a rope; each subunit consists of three polypeptide chains (α chains) wound around each other to form a triple helix. Hydrogen bonding holds the three chains together. Each tropocollagen is approximately 300 nm in length and 1. 5 nm in diameter, they aggregate to form larger structures such as fibrils.

Vertebrates have over 40 genes that code for collagen α chains, which combine to form 29 different types of collagen. The majority of collagens are homotrimers, containing three identical α chains, but heterotrimers also exist which contain a combination of two or three different chains. Each polypeptide chain will spontaneously form a left-handed helix (therefore not to be confused with right-handed α-helices), and combine with two other chains to produce a right-handed coiled coil or superhelix. This triple-helical region is a key characteristic of collagen and is referred to as the collagenous domain. Each α chain has a repeating Gly-X-Y sequence, where the X and Y residues can be any of the 20 amino acids but are frequently proline and hydroxyproline. Approximately 50% of a collagen molecule is made up of glycine, proline and hydroxyproline. This sequence is vital for the formation of the collagenous domain. A Glycine residue is required at every third position because it is the only amino acid small enough to be compacted inside the helix, since its side chain consists of single hydrogen. Other amino acids in the X and Y positions project their side chains outwards. Hydrogen bonding between Proline and hydroxyproline residues help to stabilise the triple helix; the concentration of theses residues in the collagens of animals adapted to low temperatures e. g. fish and sponges is less than those used to warmer conditions. Imperfections (1-3 residues) and interruptions (4+ residues) of the triplet sequence result in kinks and flexible regions within the helix. Fibrillar collagens have the least tolerance for imperfections where the substitution of a single glycine can result in pathological conditions. Interruptions in other collagens are fairly common place and may have evolved to fulfil certain structural functions. Some alterations to the sequence and their role as integrin-binding sites will be

considered later. There are other proteins that contain these collagen-like triple helical domains but are not considered to be collagens. Therefore, having a similar domain alone is not considered sufficient; a protein must also have a structural function in the extracellular matrix to be considered as a true collagen.

Collagen Synthesis

The synthesis of fibrillar collagen has been studied extensively due to its presence and abundance in all animals; therefore, it is a relatively well understood process. Translation of fibrillar collagen mRNAs by ribosomes attached to the endoplasmic reticulum results in the production of a precursor molecule called preprocollagen. The preprocollagen consists of a collagenous domain with two adjacent non-collagenous domains known as the C- and N- propeptides along with a RER signal sequence. Once inside the ER lumen preprocollagen undergoes several modifications before being secreted from the cell. Firstly, the RER signal is cleaved; the chains are now know as procollagen. Next hydroxylation of proline and lysine residues is carried out by different enzymes in the presence of vitamin C. This stage may be followed by O-linked glycosylation of specific hydroxylysine residues. The function of these glycosylated-hydroxylysines is not fully understood, however studies suggest that their varying concentrations in fibrillar collagens may somehow regulate fibril thickness; a high concentration results in thin fibrils, while low concentrations results in the formation of thick fibrils. Assembly of the triple helix begins with the association of appropriate α chains at their C- propeptide domain, three α chains combine to nucleate the triple helix which then undergoes elongation in the C- to N-

direction. At this stage the fibrillar procollagen is exported to the golgi where it is packaged before being exocytosed from the cell. Extracellular proteases cleave the non-collagenous domains from the procollagen producing mature tropocollagen. These mature units aggregate laterally to form striated fibrils. The final step is the formation of intermolecular cross links between lysine and hydroxylysine residues of adjacent tropocollagens. This step provides the tensile strength of fibrils. Cross linking is catalysed by the enzyme lysyl oxidase.

Role of collagen in evolution

Multicellularity is one of natures greatest innovations, the leap from unicellular to multicellular life forms has been facilitated by collagen based solutions to three key problems; cell adhesion, mechanical support and protection. This report will proceed to explain how collagen has enable life to overcome the first two of these problems, giving rise to complex multicellular life forms such as ourselves.

Adhesion

The evolution of the multicellular life has been dependent upon the development of an extracellular matrix that acts to cement the cells together and enable cell-cell communication. One very primitive way to achieve this would be via secreted intracellular cement, although this alone would provide little possibility for coordination. A better approach would be to combine an intracellular matrix with cell-cell juctions. This arrangement has been identified in the sponges. Sponges have two covering layers of cells which are connected via cell junctions and an inner jelly-like mesohyl region that is separated from the external environment and contains a collection of

other specialised cells. Analysis of the jelly of sponge mesohyl regions has shown it to contain collagenous substances. Phylogenic studies suggest that sponges are the oldest surviving anscestors of all metazoan life, with some of the oldest fossils dating back 580 million years. Therefore, their collagen-like molecules have gained much interest. Subsequently a lot of research has been conducted on the sponges with particular attention being paid to the demosponnge taxa, which encompasses around 90% of sponge species. Sponges in this taxa support their collagenous mesohyl with spongin, a short-chain collagen that aggregates to form a skeletal structure. Significant structural similarities have been identified between spongin and verterbrae IV collagen, strongly suggesting that they are related and share a common anscestor. Spongin may be the closest descendent of a collagen that appeared aroung the time of milticelularity.

As the complexity of life has increased so too has the complexity of the extracellular matrix. The ECM of vertebrates contains numerous types of collagenous, elastic and reticular fibres. Fibrillar collagens are by far the most abundant; they provide the structural support, help cell to withstand pulling forces and enable cell movement. Vertebrates cells are able to mediate with collagens directly via integrin receptors on their surface or indirectly via glycoproteins found withing the matrix, such as fibronectin a protein which facilitates cell movement. Integrins are a group of cell-surface receptors that mediate interactions between a cells cytoskeleton and the extracellular matrix. They are crucial for cell attachment to the ECM and also for transducing external signals to the cell. Integrins have been identified in all animal species including sponges. Integrins bind collagen at specific

sequences within the collagenous and non-collagenous domains, there are four methods for the attachment.

They appear to have evolved later than the collegens themselves. Suggesting that collagens original role was purely structural protein only … specialized and taken on other roles.

The evolution of the multicellular life has been dependent upon the development of an extracellular matrix that acts to cement the cells together and enable cell-cell communication. One way primitive way to achieve this would be via a secreted intracellular cement, although this alone would provide little possibility for coordination. A better approach would be to combine an intracellular matrix with cell-cell juctions. This arrangement has been identified in the sponges, which have two covering layers of cell connected by junctions and an inner jelly-like mesohyl region, that is separated from the external environment and contains a collection of other specialised cells. The mesohyl of sponges contains collagenous substances. Phylogenic studies suggest that sponges are the oldest surviving anscestors of all metazoan life, with fossils dating back to 580 million years being recovered. Therefore, their collagen-like molecules have gained much interest. Subsequently a lot of research has been conducted on the sponges with particular attention being paid to the demosponnge taxa, which encompasses aroung 90% of sponge species. Sponges in this taxa support their collagenous mesohyl with a skeleton of spongin. Spongin is a short-chain collagen that aggregates to form the skeletal structure. Significant structural similarities have been identified between spongin and verterbrae IV collagen, strongly suggesting that they are related and share a common

anscestor. Spongin may be the closest descendent of a collagen that appeared aroung the time of milticelularity. As the complexity of life has increased so to has the complexity of the extracellular matrix. The ECM of vertebrates contains numerous types of collagenous, elastic and reticular fibres. Fibrillar collagens are by far the most abundant, they provide the structural support and enabling cell movement. Vertebrates cells are able to mediate with collagens directly via integrin receptors on their surface or indirectly via glycoproteins found withing the matrix, such as fibronectin a protein which facilitates cell movement. Integrins have been identified in all animal species, they are receptor that mediate cell-cell interactions and interactions with the extracellular matrix, they appear to have evolved later than the collegens themselves. Suggesting that collagens original role was purely structural protein only … specialized and taken on other roles.