REVIEW

Eur. J. Histochem.
46: 111-124, 2002


# The first and second 'laws' of chemical morphology, exemplified in mammalian extracellular matrices

J.E.Scott

Chemical Morphology, Medical School, Manchester University, M13 9PT, UK

Accepted:30/1/2002

Key words: anionic glycosaminoglycan, collagen, connective tissues, cornea, dermochondan (dermatan) sulphate, extracellular matrix, fibrils, keratan sulphate, tendons, valency

## SUMMARY

Tissues are supramolecular organisations. The permanent and semi-permanent biopolymers therein function collaboratively in specifically bonded frameworks of macromolecules according to the physico-chemical laws that govern the behaviour of all molecules. In this paper two 'laws' or principles are discussed which give insights into the development and function of tissues, particularly the extracellular matrices (ECMs) of connective tissues.

The first 'law' is qualitative;- *The shape of a tissue is implicit in the shapes of the biopolymers from which it is constructed.* The tissue biopolymers are jigsaw pieces, if they don't fit precisely, there is no picture. The second 'law' is quantitative;-*The composition of a tissue is determined by the stable, specific interactions between the macromolecules of which it is constructed.* These basic ideas underlie the discipline of chemical morphology.

The term chemical morphology implies both the chemistry of shape and the shape of chemicals. The first meaning is well exemplified in the ECMs of connective tissues, in which the shape of an organism is defined and maintained. Specific relationships between the fibrillar (collagenous) components and the soluble polymers (proteoglycans) are set in the context of the first law. Tissue electron histochemistry (the morphology of the tissue) and knowledge of secondary and tertiary structures of the participating biopolymers (the shapes of the chemicals) together provide a model susceptible to quantitative testing. Simple calculation shows that the amount of any ligand (e.g. a proteoglycan) specifically bound at a single binding site per unit of collagen fibril length (the D period) increases linearly with the fibril diameter. Given the amount of collagen (measured as hydroxyproline) and its density, the constant of proportionality is ~42. Comparisons of the quantitative relationship between collagen and proteoglycans predicted from the model agree well with those obtained by biochemical analyses of different tendons from three species at all stages of development. Thus, the second 'law' appears to hold in this case.

This article is based on the Sesquicentennial Lecture given at an international symposium sponsored by the Anatomical Society of GB and Ireland, as part of the celebrations of the 150th anniversary of Manchester University and of 25 years of the first Chair in Chemical Morphology, September 20th 2001.

Dedicated to Professor Ulric Welsch, Munich University on his 60th birthday

E-mail: e-mail: jescott@fs1.scg.man.ac.uk

## INTRODUCTION

Until recently chemistry was quite unequal to the problems posed by morphologists. Now, chemists and morphologists look at the same things in different environments with different techniques.The locations of biopolymers of known compositions, shapes and sizes are sought in the tissues and tissue components seen at molecular resolutions await chemical characterisation. Knowledge of what is where poses crucial questions, why do they persist in their locations, and what do they do there? These are fundamental questions in chemical morphology.

With constructive linguistic ambiguity, chemical morphology means both the chemistry of shape and the shape of chemicals. We cannot make real progress with the former without substantial knowledge of the latter. If chemical morphology exists as a discipline in its own right it must be based on principles or 'laws'. In this paper I propose and discuss two such principles. In some contexts they have been in the intellectual atmosphere for many decades, but it is worthwhile to formulate them in terms of chemical morphology.

## PREMISES

This discussion will be restricted to permanent or semi-permanent structural elements. Freely mobile molecules and ions may collaborate in these structures, but at this stage they will be regarded as adjuncts, often functioning on short time scales.

It is assumed that, at least at the outset, most supramolecular complexes of which tissues are made are non-covalent and reversible.

**The first principle** is qualitative. *The shape of a tissue is implicit in the shapes of the biopolymers from which it is constructed.* Naturally, this does not mean that kidneys are made of kidney-shaped molecules, rather that the macromolecules are like jig-saw pieces; if they don't fit there is no picture.

Given the right molecules with the right shape, they must stick together in stable aggregates, to ensure a practical lifetime for the 'tissue'. Because of their large size and complexity most biopolymers can interact with other biopolymers in myriads of ways. If that happened in the tissue the result would be chaos. Specific interactions are needed, - at the expense of all the rest.

*Specificity* is 'a maximum of affinity shown in a minimum number of ways' (Scott 1973). The shapes of the participants control the directions in which this affinity is applied.

**The 2nd principle** is quantitative;-*The composition of a tissue is determined by the stable, specific interactions between the macromolecules of which it is constructed.*

If A interacts specifically with B

$$A + B \leftrightarrow AB$$

a quantitative relationship between the tissue concentrations of A and B exists if all A and B are destined to be in the form of AB. This kind of interaction may be termed 'obligate'. The quantitative relationship between A and B is defined by their valency. If 2 of B combine with one of A then A is divalent, etc.

The concept of 'valency' was proposed by Edward Frankland (1852); the first Professor of Chemistry in Manchester. 150 yrs ago he demonstrated that the composition of compounds was determined by the valency of their constituent elements. His discussion was necessarily limited to relatively simple compounds and I am extending his approach to the chemical morphology of tissues.

These principles grew in importance for me during researches into structure/function relationships in connective tissues (Scott, 1995), which, more than most, demonstrate the importance of shape at all levels from molecules to the whole organism.

Shape is fundamental to all biology. Without permanent reproducible shapes central functions such as digestive, circulatory and nervous systems could not have evolved. Arcadi (1952) quoted his unnamed teacher, who said, 'If by some magic solution one could dissolve all the connective tissues of the body, all that would remain would be a mass of slimy epithelium, quivering muscle and frustrated nerve cells'.

Connective tissues, or more accurately their extracellular matrices (ECMs), are 'systems of insoluble fibres and insoluble polymers which have evolved to take the stresses of movement and the maintenance of shape' (Scott 1975). This definition is illustrated in Fig. 1, a unified view of ECMs. The fibres, which in animal ECMs consist largely of the protein collagen, are inextensible, transmitting and resisting pulling forces, thereby determining the *maximum* size of the tissue. The soluble polymers

are largely carbohydrate and highly negatively charged (i.e. polyanionic) endowing them (the proteoglycans, PGs) with expansile or swelling pressure, like the air in a motor car tyre, therefore resisting compression (Scott 1975).

## SHAPE AND SUPRAMOLECULAR ORGANISATION (THE 1st LAW) IN EXTRACELLULAR MATRICES

The *collagen fibre* exemplifies the elegant ways in which shape and function come together in the ECM.

Collagen fibres must transmit heavy loads without breaking. The strategy underlying the structure of the fibres is to pack the maximum number of polypeptide chains, aligned along the axis, into the fibre, thus increasing the number of chains which must be broken or displaced before the fibre fails, mechanically. The molecules themselves are long, giving fewer natural breaks in a given length of fibril.

Each collagen molecule is a packet of 3 polypeptide chains, twisted round each other to form a triple helix, 200 times longer than it is thick. Each chain is itself a helix, with about 3 aminoacid residues per turn. Since every 3rd residue is glycine (the smallest amino acid), these are exclusively placed along one 'edge' of the approximately 3-fold helix. The glycine edges occupy the centre of the triple helix (Fig. 2), thus permitting the closest possible approach and packing of the 3 chains while the other, bulkier, residues are directed to the outside of the triple helix, where they are free to interact with other molecules. The close proximity of the 3 chains allows the formation of H-bonds and covalent bridges between them, greatly increasing the stability of the triple helix.

The peripheral amino acids are organised along the molecule in bands of polar and non-polar residues, so that when aligned alongside another collagen molecule with about 75% of their lengths overlapping, their summed interactions give optimally stable aggregates (Fig. 2). This 'quarter-staggered' array is consolidated by covalent cross links, effectively converting a large number of molecules into a supermolecule of great tensile strength.

After staining with heavy metals the bands of polar and non-polar aminoacids show up in electron microscopy as a characteristic bar code pattern, the D pattern, labelled a-e, which repeats itself about four times per molecule length, every ~65 nm. It is possible to correlate the position in the banding pattern with the aminoacid sequence of the collagen molecule (Fig. 2), allowing one to identify aminoacid sequences by electron microscopy, in certain circumstances (see below).

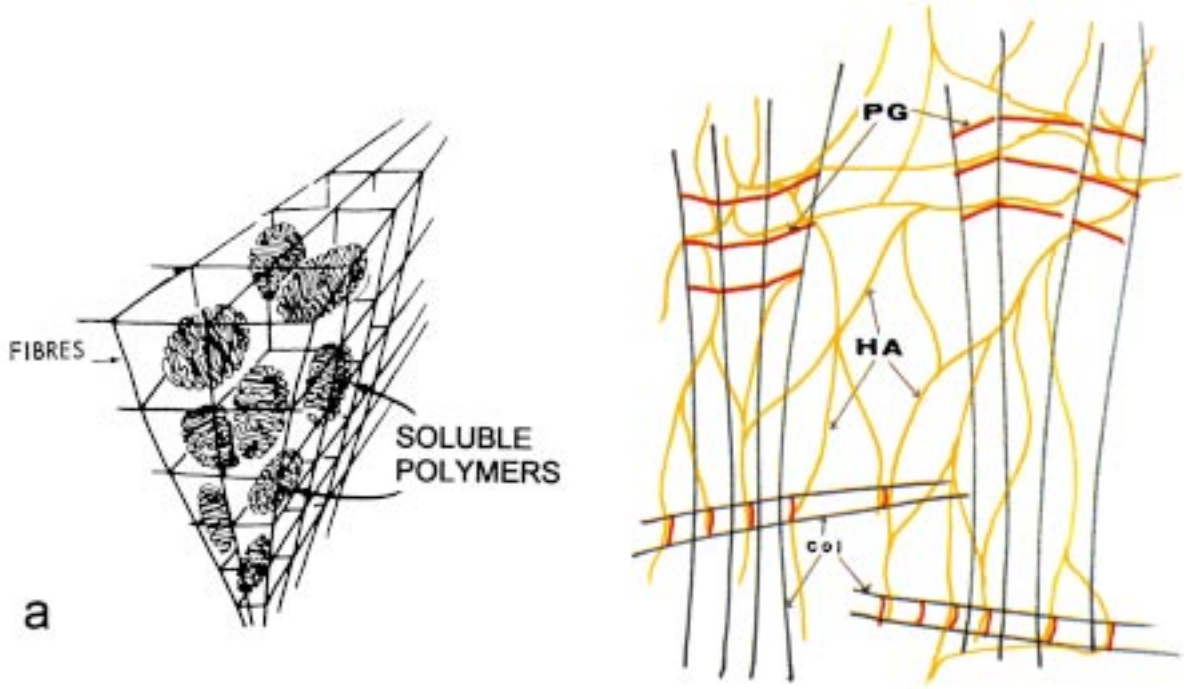
There is no clearer example in all biology in which the fitness for function is traceable back to the primary structure. The interactions of the fibril with other ECM molecules, including the PGs, are crucial in building up the tissue, according to the 1st and 2nd laws.

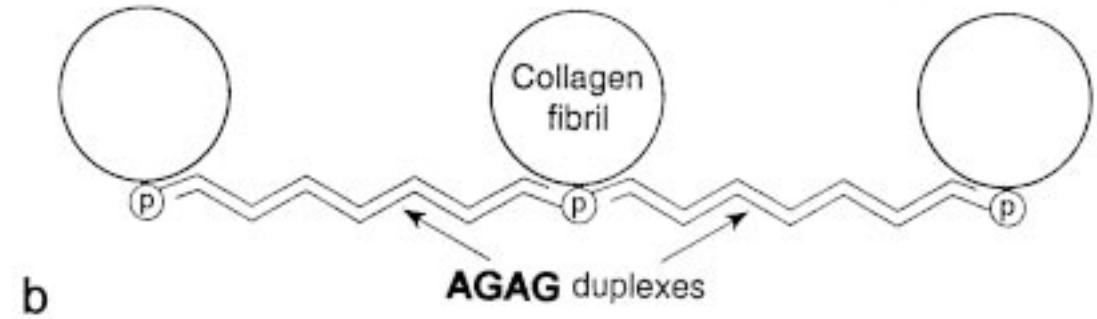
If, as diagrammed in Fig. 1a the PGs were free to move, like fish in a net, they would translocate under pressure, not necessarily returning to their original position, thus seriously hindering maintenance of a permanent shape. A stable relationship between the PGs and the fibrils must exist.

A reagent, Cupromeronic blue, was devised and synthesised (Scott, 1973) with which to label and visualise the polyanions by light and electron microscopy, crucially showing their shape, size and orientation (Scott, 1992). Cupromeronic blue stains the part of the PG carrying the high negative charge, which is located on the polysaccharide chains (anionic glycosaminoglycans, or AGAGs ).

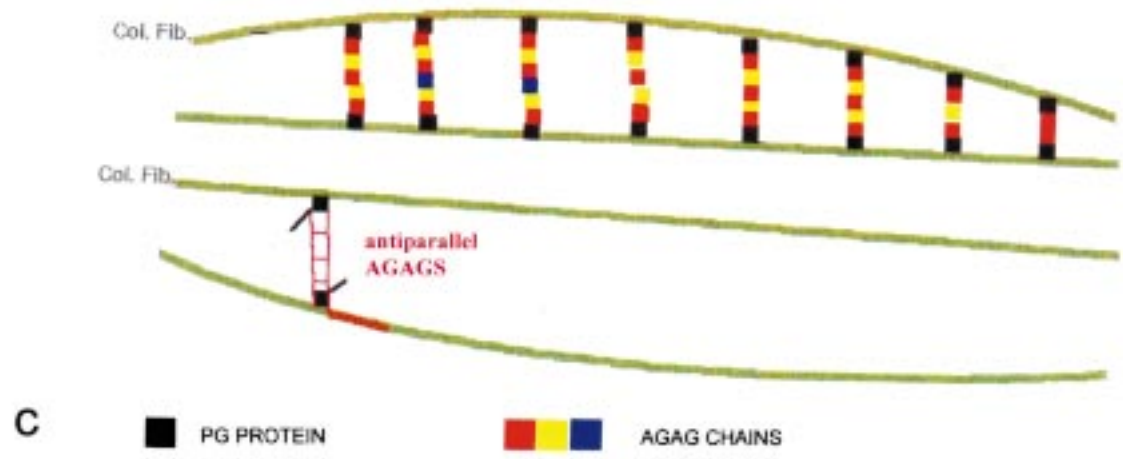
Results on tendon, skin, corneal stroma etc conclusively demonstrated (Fig. 3) regular association of PGs with collagen fibrils, at sites ~62 nm apart, This means there is one binding site in every D period. Struck by the appearance of these pictures, Rhuoslati called the PG '**decorin'** (now **decoran**) because it decorated the collagen fibril. Parallel staining with uranyl acetate located the sites in the d and e bands (Fig. 2). In mammalian corneal stroma, exclusively, the a and c bands are associated with an analogous PG (Scott and Haigh, 1988), but this was not called a decoran. These specific collagen interactions play a dominant role in organising the ECM space (see below).

Homologous sequences of 11 aminoacids (Fig. 4) found at all the association sites in fibrillar collagen - but in no other protein, were proposed (Scott and Glanville, 1993) to be the actual binding sites. Particularly striking is the presence in the first five residues in all the sequences of an invariant charge motif; -,+,0,-. It was therefore intriguing to find the same motif - but with




## The Shape Module



charges reversed, at 2 points in the PG protein, suggesting that ionic attraction between the PG protein and the collagen fibril might be part of the binding mechanism (Scott 1996).

This prospect became exciting when it was found by electron microscopy (Scott 1996) that the PG protein was horseshoe shaped and one copy of the charge motif was present on each arm of the horseshoe, symmetrically placed about the horseshoe axis. The horseshoe is the right size to locate the complementary charge motifs in the PG and collagen fibril over each other, thus linking 2 adjacent collagen molecules together, in quarter stagger (Fig.5a). Preliminary biochemical evidence is compatible with this idea (Scott *et al* 1997).

The consequences of this possibility are far reaching, since every collagen molecule in the quarter stagger can, in principle, participate in 2 of these PG protein bridges, one linking to the molecule to the right and the other to that on the left (Fig. 5b). There is then a very extensive structure in which every collagen molecule is linked to every other via a network of PG-protein bridges. Since this arrangement appears to be restricted to the fibril surface, it may be mportant in controlling the accretion of collagen molecules during fibrillogenesis and maturation of the fibril.

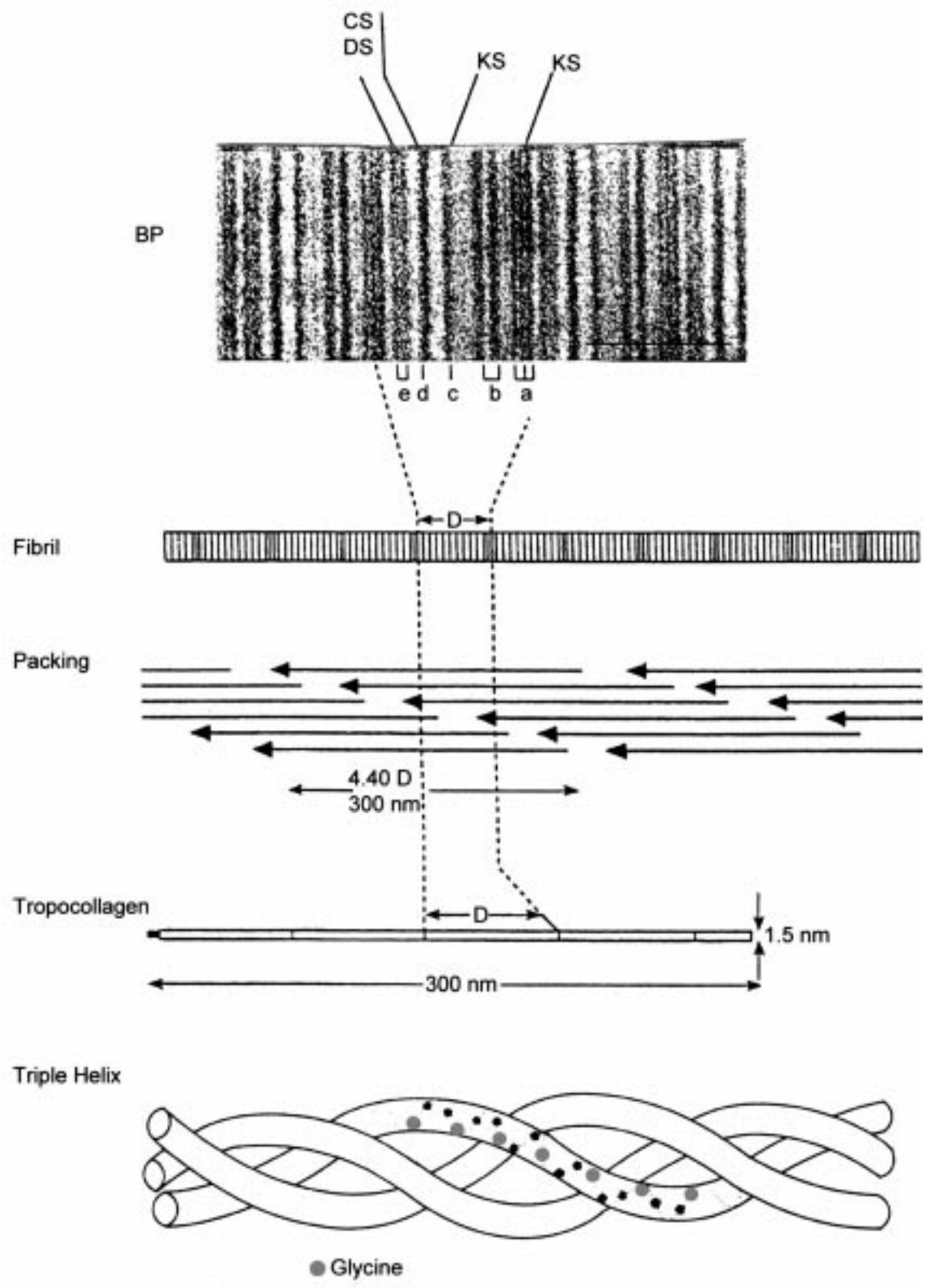
Thus, the shapes of the proteins ensure that the specific valencies are presented in exactly the right way to build up the tissue, as implied by the 1st law.

These bridges are *inter-molecular*. On a larger (>10x) scale the AGAG chains link between *fibrils* (Figs. 1 and 3). They probably orient and maintain orderly parallel arrangements of fibrils (Scott, 1992a). The key to organisation of tissue shapes lies in getting collagen fibrils into the right place. Maintaining them there appears to be a function of the PG bridges. There is no obvious alternative way of doing this. In corneal stroma, young tendons etc. the AGAG bridges are the only visible regular interfibrillar connections.

The bridges are too thick and often too long to consist of single polymer chains (Scott, 1992b). They must consist of aggregates of several AGAG chains, but how do highly charged mutually repelling polyanions form stable aggregates? The answer lies in the shape of the PGs. In fact, they aggregate spontaneously in solution (Ward *et al* 1987) via their AGAG tails, with the AGAG chains lying head to tail, i.e. antiparallel (Fig. 1b). This result translates directly to the tissue, with the PG protein binding to collagen and the AGAG chains forming the bridges. Not only are the soluble polymers, the PGs, tied to the collagen fibrils, they also tie the collagen fibrils to each other. They are bifunctional.

I called this structure the ‘shape module’ (Fig. 1b) because it can orient and space the collagen fibrils. The length of the AGAG chains correlates

**Fig. 1** - Shape/function relationships in extracellular matrices of connective tissues. (a left) Scheme demonstrating the two stress-resisting components of a tissue; fibres (mainly collagen, see Fig. 2) which transmit and resist tensile or pulling forces and soluble polymers which tend to swell in aqueous solution, hence resisting compressive or pushing forces, like the air in a motor car tyre. Direct fibril-fibril connection was originally (1973) invoked in this model to give form to the arrangement, although there was (and still is) no evidence for such connections. No specific associations between the two components had been proven although a requirement was recognised at the time (Scott 1973). (a right) Electron histochemistry of the PGs in many connective tissues (e.g. Fig. 3) provided bases on which to replace the fibril-fibril connections with bridges of pressure-resisting polymers between the fibrils. These PG bridges are regularly distributed along the fibril length at intervals of about 62nm (the D period, see Fig.2), orienting and maintaining the collagen fibrils at separations that vary with the tissue and which correlate with the length of the anionic glycosaminoglycan (AGAG) chain carried by the PG (see (b) and (c) below). Further stabilisation in many tissues is probably provided by AGAG interactions with hyaluronan, as was suggested in the vitreous humor of the eye (Scott 1992c). (b) A cross section orthogonal to the collagen fibrils of an interfibrillar PG bridge. Because these structures are repeated every D period (Fig. 2) they are termed ‘modules’ and because of their role in positioning and ordering the collagen fibrils, which is fundamental to the maintenance of shape (see text), they are ‘shape modules’. In e.g. tendon or cornea there are many millions of these interfibrillar bridges in every $mm^3$ of tissue. The bridges consist of AGAG chains which are of different lengths in different tissues, from about 35nm in skin or tendon to over 600nm in the vitreous humor (see text). p = PG protein. (c) Shape modules at the molecular level. PG proteins non-covalently bound at specific sites along the collagen fibril are linked by AGAG bridges, aligned antiparallel. This type of aggregation occurs spontaneously in aqueous solution (Ward *et al* 1987). Possibly, different lengths of AGAG chains could be used to orient collagen fibrils at varying angles to each other (c, upper) and perhaps to keep fibrils at different distances apart in the same tissue, but no direct evidence is available, although preparations of AGAGs from tissues are heterodisperse to varying extents. The chemical structure of AGAG chains varies along their lengths and this is conveyed by showing blocks within AGAG chains in different colours.


CS
DS
KS
KS
BP
e d c b a
D
Fibril
Packing
4.40 D
300 nm
Tropocollagen
D
1.5 nm
300 nm
Triple Helix
Glycine

with the distance apart of the collagen fibrils in the tissues, being short (molecular mass 15kD) in tendon, longer (50kD) in corneal stroma and very long (300kD) in vitreous humor.

Although different interfibrillar spacings are observed in different tissues, it remains to be seen whether different lengths of AGAGs are used in the same tissue to produce different spacing at different points between 2 fibrils, which would change the shape of the tissue (Fig. 1c).

Obviously the interfibrillar AGAGs must repel each other electrostatically. Equally there must be attractive forces between the AGAGs which outweigh this repulsion, since the aggregates persist and are stable. Non-ionic attractive forces (H-bonding, van der Waals forces and hydrophobic bonding) are very short range and often highly directional. They are therefore dependant on the shapes of the participants, which must be exactly complementary.

All ECM AGAGS prefer to take up the same secondary structure in solution (Scott *et al.* 1995), a 2-fold helix (Fig. 6a). In this configuration there are extended hydrophobic patches of very similar shape in almost identical molecular environments, like chunks of lipid inserted into the centre line of the AGAG chain (Fig. 6b). Hydrophobic bonding between closely juxtaposed patches on neighbouring molecules was an obvious possibility. The two sides of the tapelike molecule, upper and lower, are identical, and when 2 chains are placed antiparallel the patches overlap and convenient H-bonds can form e.g. between acetamido NH and uronate carboxylates (Fig. 6c) There is a close conceptual resemblance to the DNA double helix, in which electrostatic repulsion (between phosphate ester groups) is outweighed by hydrophobic and H-bonding. A fundamental difference is that AGAG aggregates can grow and branch indefinitely, as shown by electron microscopy (Scott *et al* 1991), since the AGAG 2-fold helix is double sided, while DNA stops at a double helix. AGAG bridges can therefore be of variable thickness, as observed in the tissues.

This section has illustrated the 1st law, i.e. the vital part played by shape in supramolecular organ-

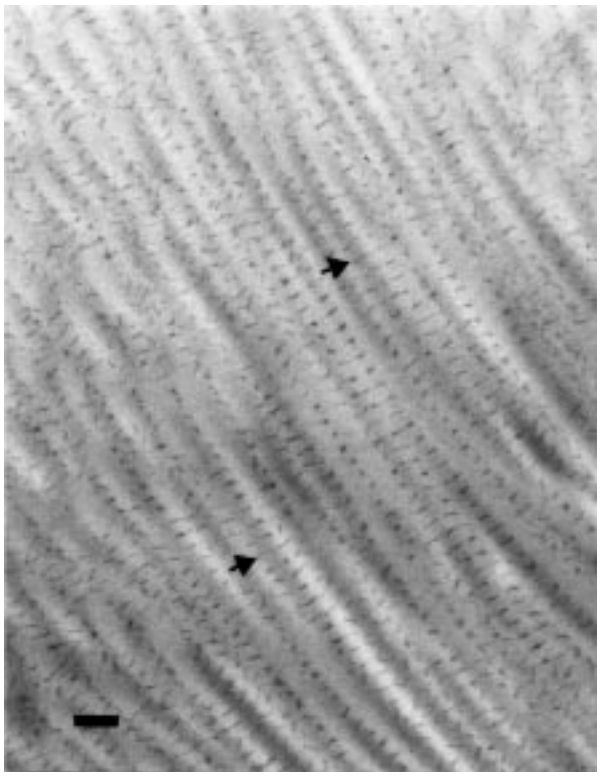


**Fig. 3 -** Electron micrograph of connective tissue from 9 wk rat ear, stained with Cupromeronic blue in 0.1M $MgCl_2$ solution at pH 5.8 (technical details in e.g. Scott 1992 b) Bar =160 nm. Collagen fibrils show as light bands running top to bottom; darker bands in between are the plastic embeddment. PG AGAGs appear as dark filaments orthogonal to the fibrils at separations of about 63nm, often bridging between fibrils (arrowed).

**Fig. 2 -** The collagen fibril hierarchy.The triple helix (bottom) is composed of 3 polypeptide chains each about 1000 aminoacid residues long, consisting of repeating triplets of Gly-X-Y, where X and Y are frequently proline or hydroxyproline. This repeating structure gives rise to a helix with the smallest aminoacid, glycine, at every third position and in the middle of the triple helix, where, consequently,packing is very close (based on a diagram by Nimni and Harkness, 1988). Bulkier amino acids locate to the periphery where they interact with other molecules.The complete triple helix, sometimes called tropocollagen, aligns spontaneously alongside other tropocollagen molecules to give the quarter-staggered packing in the fibril, on the surface of which bands of polar and non-polar aminoacids are present. These show up as a banding pattern (BP, top) bar code labelled a-e in electron microscopy after staining with heavy metal. A complete a-e pattern constitutes a D period (horizontal line in BP) which is about one quarter of the length of a collagen molecule (for further discussion, see text).d and e bands contain the specific binding sites for PGs carrying chondroitin (CS) or dermochondan sulphate (DS) AGAGs (see Fig. 4). a and c bands bind PGs carrying keratan sulphate (KS) AGAGs, in the corneal stroma.

For **Decoron**

| | | | |
|---|---|---|---|
| α1 | d | GDRGEPGPAGP | 886 |
| chain | $e_1$ | GDRGETGPAGP | 856 |
| | $e_1$ | GDRGEPGPPGP | 622 |
| | $e_2$ | GDKGEAGPSGP | 601 |
| | bands | | residue |

For PKS

| | | | |
|---|---|---|---|
| α1 | $a_2$ | GDRGEAGPKGA | 565 |
| chain | $c_2$ | GERGEQGPAGS | 451 |
| | bands | | residue |

Charge motifs

| | |
|---|---|
| In Collagen | - + 0 - |
| In Pg | + - 0 - |

Putative binding sites for collagen fibrils on Decoron

| | |
|---|---|
| RELK | 103 |
| RELH | 242 |
| | residue |

**Fig. 4 -** Probable binding sites on collagen fibrils and associated PGs. Homologous sequences of 11 aminoacids were found in the d and e bands, at which the PG decoran is associated (Fig. 2), in all the non-mineralised mammalian ECMs so far examined (Scott 1995). Similar sequences were found in the binding sites in the a and c bands where proteokeratan sulphates associate in the cornea. Very similar sequences are present in the primary structures of the fibrillar types I, II and III collagens - but in no other protein, and only at the PG binding sites (Scott and Glanville 1993). All the charge is present in the first five residues, (GDRGE, GDKGE, GERGE), as a charge motif. Decoron has a similar motif, but with charges reversed, which led to the postulate that the two complementary charge motifs could bind together as part of the association of PGs with collagen fibrils (Scott 1996, see Fig. 5).

isations which are the basis of tissue ultrastructure. Primarily this is because stabilising forces are cooperative and short range, ineffective except where participants can approach closely, and requiring large numbers of small packets of attractive energy to provide a total energy that resists disaggregation due to physical stresses.

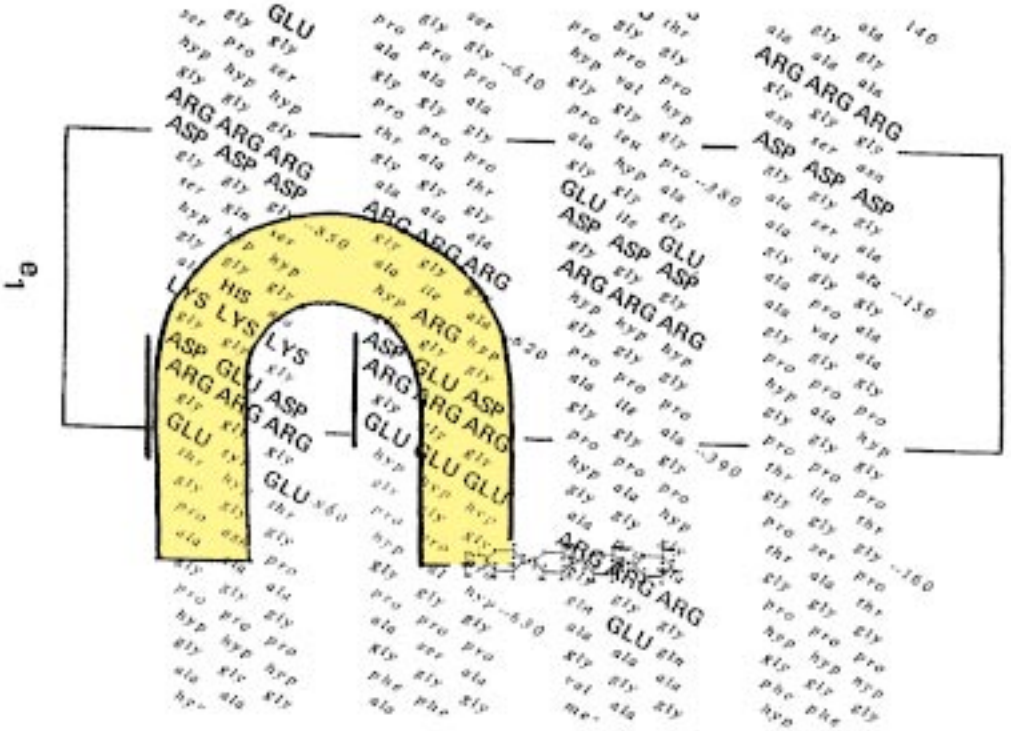
## QUANTIFICATION OF TISSUE SUPRAMOLECULAR ORGANISATION (2nd LAW)

In this context, does the **2nd law** hold? Is the composition of the tissue predictable from our understanding of the specific relationships? In particular does the observed relationship between collagen and decoran permit direct calculation of their weight ratio in the ECM and does this estimate agree with biochemical data?

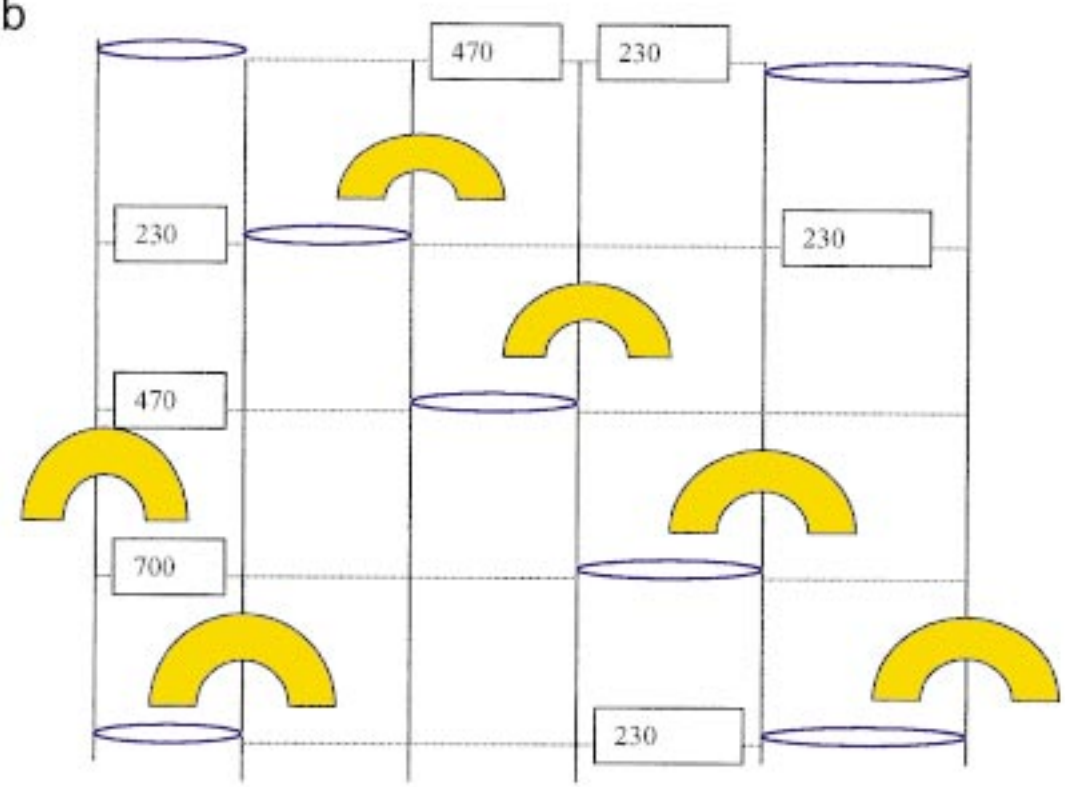
The fibrillar-surface-located d-band binding sites each bind a single PG molecule (Figs. 2 and 3). Each binding site is present at the surface of the fibril in a group of 5 quarter-staggered collagen molecules arranged around the periphery of the fibril (Fig. 7). The number of binding sites therefore increases linearly with the circumference (and hence with the diameter) of the fibril. The amount of collagen in a section of collagen fibril cylinder increases with the square of the diameter. The amounts of binding site per unit cylinder of collagen is therefore in the ratio circumference/(diameter squared).Thus the amount of collagen per molecule, particle or equivalent unit of specifically bound ligand (say, in the ring defined by the d band) increases linearly with the collagen fibril

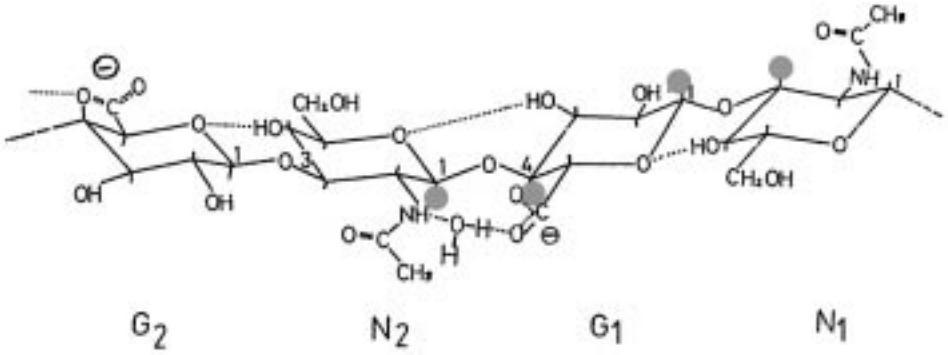
**Fig. 5 -** Putative PG:collagen fibril interactions leading to extended tissue supramolecular organisations. PG proteins (decoron etc) are horseshoe-shaped on which it is possible to locate the probable collagen binding motifs (Scott 1996) (see Fig. 4). (a) Decoron is of the right size to allow juxtaposition of the complementary sites on the collagen fibril, in the quarter staggered array, as found in the tissues. This model, which is to scale,utilises the fact that the binding site charge motifs on the collagen fibril are aligned alongside each other (vertical lines adjacent to residues 855 and 624 in the $e_1$ band) because of the quarter stagger. Hence the positioning of the PG motifs on opposite arms of the horseshoe potentially satisfies the spacial requirements for complementary overlaps of the two sets of binding sequences (Scott 1996). Preliminary biochemical evidence is compatible with this model (Scott *et al* 1997). The aminoacid sequences in the four collagen molecules in quarter-staggered array are as illustrated by Chapman and Hulmes (1984). (b) Proposed scheme showing five adjacent collagen molecules in quarter-staggered array at the fibril surface, as in the tissue fibril, with pairs of putative PG binding sites in 'cascades'. This array is a direct corollary of the PG protein:collagen fibril structure shown in Fig. 5a in which each adjacent pair of binding sites is bridged by a PG protein. Every collagen molecule at the fibril surface is then linked to every other via an extended sheet, as shown.The PG proteins scaffold the aggregated collagen which is the fibril. Numbers refer to the positions in the primary structure of the polypeptide chains, with each cross-line locating the boundary between one D period and the next. The ellipses are the ends of the collagen molecules. The gaps between collagen molecules (see Fig. 2) are not shown.

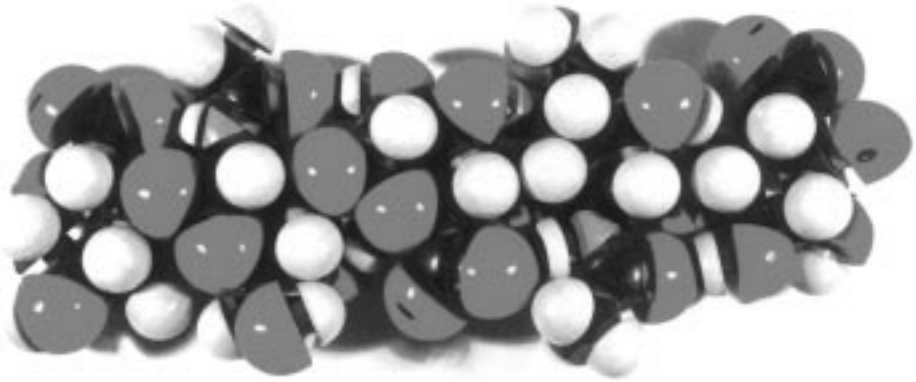
a

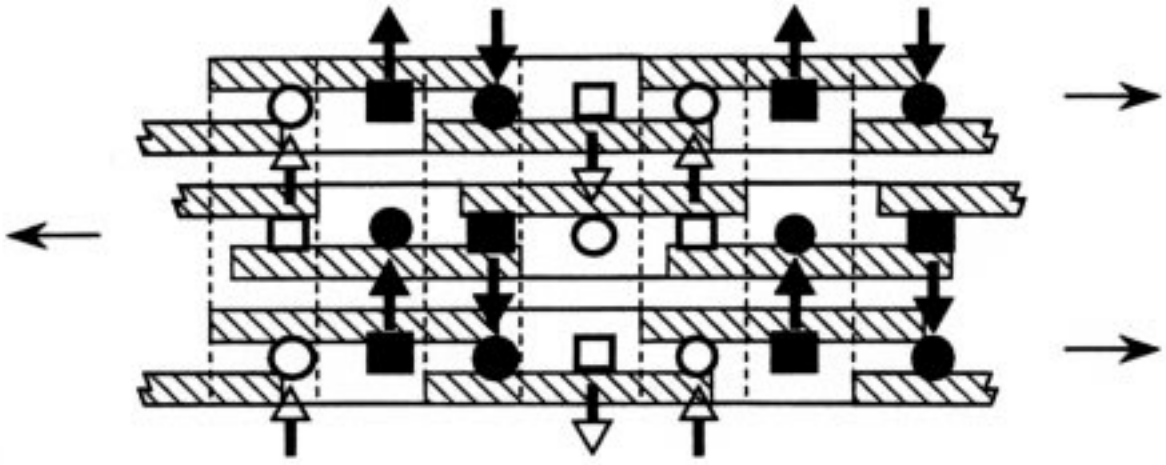


b




a
G2
N2
G1
N1


b


c

diameter. This prediction is supported by data on the amount of dermochondan sulphate associated with collagen fibrils of varying diameters (Fig. 8b).The fibril is made up of a set of cake-slices, defined as to size by the five quarter-staggered collagen molecules (Fig. 7). In a Frankland-like sense, the valency of that piece of collagen fibril is one.

The amount of collagen in the cake-slice of the fibril containing that binding site is easily calculable (Fig. 7). If hydroxyproline (hyp) is assayed as a measure of collagen, the constant relating fibril radius to amount of hyp per binding site is ~42 (Fig. 7). This constant is independant of the nature of the ligand**, which could be any species associating with one type of specific binding site in each D period of collagen fibril as illustrated in Fig. 7.

The quantification argument can be developed further. Given the weight of the decoran AGAG (DS) chain attached at the binding site, the ratio of this weight to that of the collagen section can be calculated, thus allowing direct comparison with the measured wt. ratios in Fig. 8b. The weight of a single decoran DS chain is the weight of one gm. molecule divided by Avogadro's number. The ratio of collagen to DS in this slice of the collagen fibril cake is then 1.28r (Fig. 8a) for a mol. mass of 20kD, (a reasonable estimate for DS from tendons and skin). The ratio based on biochemical analyses is 1.18r, the slope of the regression line in Fig. 8b. Put another way, the calculated value for a fibril of radius 50nm is 64 (Fig. 8a), whereas the biochemical estimate, taken from Fig. 8b and based on measurements from 23 (usually pooled) samples of tendons from 3 different animals is 58. Finally, had a value of 21.4kD been chosen instead of 20kD as the mol. mass of the DS chain the calculated ratio would have been 1.18, in exact agreement with the biochemically measured value.

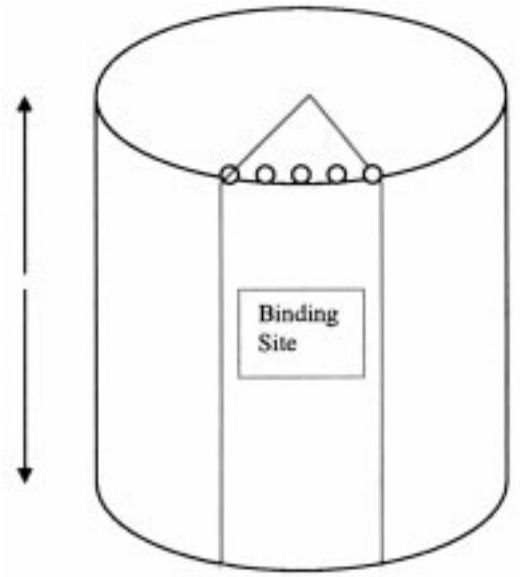
This result suggests that the DS molecular mass (which is an average) was constant throughout the range of tissues at different ages and that the biochemical assays measured these species with equal efficiency.It supports the assumption in the model that one binding site associates with one molecule of DS - and in addition, that all surface binding sites are so occupied. The implications of this finding in fibril growth and maturation remain to be worked out. The assumptions and conclusions appear valid during development and maturation in a range of physiological conditions, since the tendons were from rat tail, or flexor digitorum from cows and chickens, *in utero*, *in ovo*, and *postpartum*. The

**Is this the first experimental evidence for the Adams (1979) conjecture that the answer to Life, the Universe and Everything is 42?

**Fig. 6** - AGAG shapes and supramolecular organisational possibilities. The structures are of hyaluronan, the simplest ECM AGAG, which is without sulphate esters that can modify AGAG behaviour in forming secondary and tertiary structures. Other AGAGs (chondroitin, dermochondan and keratan sulphates) have essentially similar characteristics (Scott *et al.* 1995). (a) The secondary structure of hyaluronan, a 1:4, 1;3 β linked polymer of repeating disaccharides of hexuronic (usually glucuronic) acid (G) and hexosamine (N) typical of the ECM AGAGs. In this two-fold helix the first disaccharide G1N1 is rotated by 180° around the glucosaminide link N2→G1with respect to the G2N2 disacharide, which is in turn rotated similarly about the preceding disaccharide, and so on down the chain. Every alternate disaccharide is then oriented identically, i.e there has been a double or twofold helical rotation to return to the original orientation. In twofold helices of this type the pairs of protons on either side of the glycosidic bonds (indicated in gray in (a)) are very close together. They should therefore give NOEs (nuclear Overhauser enhancements), NMR phenomena that characterise such closely packed atoms. This was demonstrated for chondroitin and dermochondan sulphates and it therefore is very probable that all the ECM AGAGs prefer to take up a two-fold helix (NOT a double helix!) in solution (Scott *et al.* 1995). (b). The skeletal structure in (a) clothed in atomic flesh, realised by modelling with Courtauld spacefilling atoms. Colour coded: gray = oxygen, black = carbon, white = hydrogen. The 2-fold helix is a tapelike molecule with the hydrophilic groups (carboxylate, hydroxyl and amide) at the lateral edges of the tape while the upper and lower surfaces of the tape display hydrophobic groups (CH and C-O-C). The CH units constitute a hydrophobic patch akin to a fatty chain in a lipid. The upper and lower surfaces of the tape are identical, but run antiparallel to each other. (c) 3 flat 2-fold helices stacked vertically and antiparallel above each other with the hydrophobic patches (cross-hatched) juxtaposed. The carboxylate groups (circles) are acceptors in H-bonds from acetamido NH (squares).Because the AGAG tape is double sided, similar interactions can occur in two directions, upwards and downwards, without limit in principle. This allows the AGAG aggregates to form in different thicknesses and also to form branching networks (Scott *et al* 1991, Scott and Heatley, 1999). The arrows at left and right indicate the reducing ends of the molecules. The filled symbols are at the distal edge of the molecular tapes, the open symbols at the proximal edge.

**SCHEME, showing that the weight of any collagen fibril surface-associated species with one specific association site per D period decreases linearly with fibril radius.**

Circumference of d/e band of fibril radius r is $2\pi r$ (measured in nm)
No. of collagen molecule carries a binding site in this ring, **∴** each ring has 4.19r/5=0.84r binding sites




Wt. hyp per D unit;
W = vol. x density x % hyp
**$W = D \times \pi r^2 \times$ collagen density (1.3) x % hyp (14%) in collagen**

**$W = 62 \times \pi r^2 \times 1.3 \times 0.14 = \underline{35.4r^2 \times 10^{-21}\ g}$**

***Wt of hyp per binding site (hyp/Bind.Site)***
**= W**/no. binding sites per D unit

$$\textbf{hyp/Bind.Site} = \frac{35.4r^2 \times 10^{-21}}{0.84r} = \textbf{42.14r x } 10^{-21}\textbf{g}$$

**Fig. 7 -** A section of a collagen fibril cylinder of length one D period, with specific binding sites for ligands (e.g. PG proteins) limited to one band (e.g. the d band) within the D period, circling the collagen cylinder. The cylindrical section consists of a number of slices ('cake slices'), each of which has a periphery of identical length. defined by the 5 collagen molecules (small circles). The total weight of collagen per cake slice is calculable, given the density of collagen (1.3) and the volume of the cylindrical section. The number of cake slices (= number of binding sites) in the section is obtained by dividing the circumference by the length of the 5 laterally aligned collagen molecules. The amount of collagen per binding site is the total collagen in the D section divided by the number of cake slices. If the collagen is assayed via hydroxyproline (hyp) determinations, the constant linking the amount of collagen per binding site to the fibril radius is~42.

PG:collagen fibril interaction, the basis of the shape module, must have been qualitatively and quantitatively important throughout animal evolution.

The good agreement of the calculated with the measured ratio implies that uncertainties in the model and assay methods are not overwhelming. The **2nd law** apparently holds in this case. Adequate data for other ligands are not yet available on which to test further the 2nd law.

## CODA

At a time when macromolecules in tissues are seen as black boxes, identified by an alphabet soup of monoclonal antibodies, we should not forget that they are just molecules (although often rather large) that obey the the same laws governing the behaviour of small molecules and that these laws are fundamental to the understanding of the organisation and stability of tissues.

**a**

**RATIO ($R_{DS}$) of TISSUE CONCENTRATIONS of DERMOCHONDAN SULPHATE TO COLLAGEN HYP in TENDONS**

$$R_{DS} = \frac{\text{wt of hyp/binding site (W, 42.1r x } 10^{-21}\text{ g)}}{\text{wt of DS/binding site (DS/bind. Site)}}$$

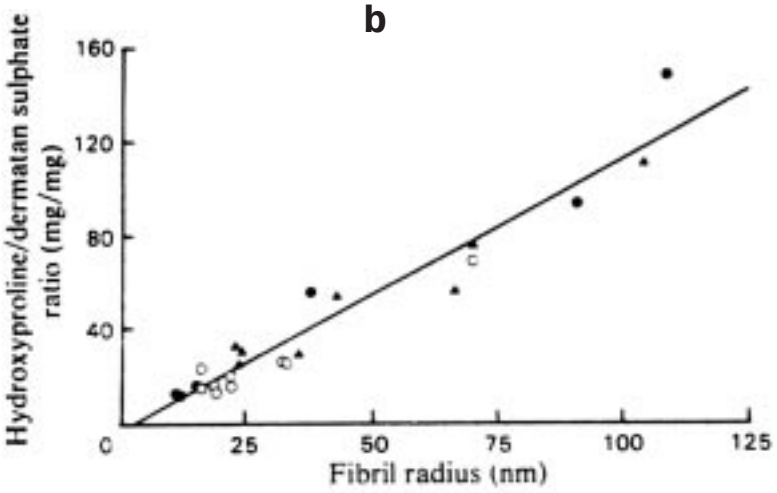
$$\text{DS/bind. Site} = \text{wt of 1 molecule of DS}$$

$$= \frac{\text{mol. wt of DS (~20,000)}}{\text{Avogadro's no (6.0 x } 10^{23})} = 3.3 \times 10^{-20}\text{g}$$

$$R_{DS} = \frac{42.1r \times 10^{-21}\text{g}}{3.3 \times 10^{-20}\text{g}} = 1.28r$$

**For a fibril of radius 50nm RDS = 64**
**Biochemically measured ratio = ~58****

**b**




*Plot of hydroxyproline/dermatan sulphate ratio against collagen-fibril diameter for bovine flexor tendon (○), rat tail tendon (●) and chicken flexor tendon (▲).*

**Fig. 8 -** (a)Assuming that one molecule of DS PG is bound per specific binding site, defined as in Fig. 7, the weight ratio of DS to collagen associated with the binding site is the wt. of one DS AGAG chain per wt. collagen per cake slice (Fig. 7). The wt. of one DS chain is the gm.molecular mass divided by the number of molecules in a gm molecule (Avogadro's number). The weight of collagen hyp in a cake slice as a function of fibril radius is W, derived in Fig. 7. Inserting values of the fibril diameters and DS mol. mass gives the weight ratio of collagen hyp/DS ($R_{DS}$).For a fibril of radius 50nm and DS mol. mass of ~20 kDa the calculated value is 1.28r. (b) Biochemical assays of this wt. ratio, carried out on tendons from 3 species in utero, in ovo and post partum gives a best fit estimate of 1.18r (Scott 1984). **i.e. 1.28 x 50. There is therefore good agreement between calculated and measured $R_{DS}$.

## ACKNOWLEDGEMENTS

I am grateful for Support from Wellcome Trust awards 057601/Z/99 and 063022/Z/0.

## REFERENCES

Adams D.: The hitch hiker's guide to the galaxy. Pan Books, London, 1979.

Arcadi J.A.: Book review of Connective Tissues (Transactions of the first Conference of the Macy Foundation), Bull. Johns Hopkins Hospital *90*, 334-335, 1952.

Chapman J.A., and Hulmes D.J.S.: Electron microscopy of the collagen fibril. In Ultrastructure of the connective tissue matrix (Eds. Ruggeri A, Motta PM). Martinus Nijhoff Publishers, The Hague, pp. 1-33, 1984.

Frankland E.: On a new series of organic bodies containing metals and phosphorus. Phil. Trans. Roy. Soc. London *142*, 417-444, 1852.

Nimni M.E., and Harkness R.D.: Molecular structures and functions of collagen. In Collagen (ed Nimni ME) vol. *1* CRC Press, Boca Raton pp. 1-78, 1988.

Scott J.E.: Affinity, competition and specific interactions in the biochemistry and histochemistry of polyelectrolytes. Biochem. Soc. Trans. *1*, 787-806, 1973.

Scott J.E.: Composition and Structure of the Pericellular Environment: Physiological Function and Chemical Composition of Pericellular Proteoglycan (an evolutionary view). Phil. Trans. Roy. Soc. London B *271*, 235-242, 1975.

Scott J.E.: The periphery of the developing collagen fibril. Quantitative relationships with dermatan sulphate and other surface-associated species. Biochem. J. *218*, 229-233, 1984.

Scott J.E.: Supramolecular organisation of extracellular matrix glycosaminoglycans, *in vitro* and in the tissues. *FASEB J. 6*, 2639-2645, 1992a.

Scott J.E.: Morphometry of Cupromeronic blue-stained proteoglycan molecules in animal corneas, versus that of purified proteoglycans stained *in vitro*, implies that tertiary structures contribute to corneal ultrastructure. J. Anat. *180*, 155-164, 1992b.

Scott J.E.: The chemical morphology of the vitreous. Eye *6*, 553-555, 1992c.Scott J.E.: Extracellular matrix, supramolecular organisation and shape. J. Anat. *187*, 259-269, 1995.

Scott J.E.: Proteodermatan and proteokeratan sulfate (decorin, lumican/fibromodulin) proteins are horseshoe shaped. Implications for their interactions with collagen. Biochemistry *35*, 8795-8799, 1996.

Scott J.E., and Haigh M.: Identification of specific binding sites of keratan sulphate and chondroitin/dermatan sulphate proteoglycans on collagen fibrils in cornea, using Cupromeronic blue in critical electrolyte concentration techniques. Biochem. J. *253*, 607-610, 1988.

Scott J.E., and Glanville R.W.: Homologous sequences in fibrillar collagens may be proteoglycan binding sites. Biochem. Soc. Trans. *21*, 123S, 1993.

Scott J.E., and Heatley F.: Hyaluronan forms specific stable cooperative tertiary structures in solution. A $^{13}C$ n.m.r. study. Proc. Nat. Acad. Sci. USA *96*, 4850-4855, 1999.

Scott J.E., Heatley F., and Wood B.: Comparison of secondary structures in water of chondroitin-4-sulfate and dermatan sulfate: Implications in the formation of tertiary structures. Biochemistry *34*, 15467-15474, 1995.

Scott J.E., Cummings C., Brass A., and Chen Y.: Secondary and tertiary structures of hyaluronan in aqueous solution, investigated by rotary shadowing-electron microscopy and computer simulation. Hyaluronan is a very efficient network-forming polymer Biochem. J. *274*, 699-705, 1991.

Scott J.E., Ritchie M., Glanville R.W., and Cronshaw A.D.: Peptide sequences in glutaraldehyde-linked proteodermatan sulphate:collagen fragments from rat tail tendon locate the proteoglycan binding sites. Biochem. Soc. Trans. *25*, S663, 1997.

Scott J.E., Cummings C., Greiling H., Stuhlsatz H., Gregory J.D., and Damle S.P.: Examination of corneal proteoglycans and glycosaminoglycans by rotary shadowing and electron microscopy, Internat. J. Biol. Macromol. *12*, 180-185, 1990.

Ward N.P., Scott J.E., and Cöster L.: Dermatan sulphate proteoglycans from sclera examined by rotary shadowing and electron microscopy.Biochem. J. *242*, 761-765, 1987.