View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Elsevier - Publisher Connector

Structure, Vol. 9, 1061–1069, November, 2001, ©2001 Elsevier Science Ltd. All rights reserved. PII S0969-2126(01)00669-4

The In Situ Supermolecular Structure of Type I Collagen

Joseph P.R.O. Orgel,[1,4,5] Andrew Miller,[1] Thomas C. Irving,[2] Robert F. Fischetti,[2] Andrew P. Hammersley,[3] and Tim J. Wess[1]

[1] Centre for Extracellular Matrix Biology
Department of Biological Sciences
University of Stirling
Stirling FK9 4LA
United Kingdom

[2] Center for Synchrotron Radiation Research and Instrumentation
Department of Biological, Chemical, and Physical Sciences
Illinois Institute of Technology
3101 South Dearborn Street
Chicago, Illinois 60616

[3] European Synchrotron Radiation Facility
BP220
Grenoble F-38043
France

Summary

Background: The proteins belonging to the collagen family are ubiquitous throughout the animal kingdom. The most abundant collagen, type I, readily forms fibrils that convey the principal mechanical support and structural organization in the extracellular matrix of connective tissues such as bone, skin, tendon, and vasculature. An understanding of the molecular arrangement of collagen in fibrils is essential since it relates molecular interactions to the mechanical strength of fibrous tissues and may reveal the underlying molecular pathology of numerous connective tissue diseases.

Results: Using synchrotron radiation, we have conducted a study of the native fibril structure at anisotropic resolution (5.4 Å axial and 10 Å lateral). The intensities of the tendon X-ray diffraction pattern that arise from the lateral packing (three-dimensional arrangement) of collagen molecules were measured by using a method analogous to Rietveld methods in powder crystallography and to the separation of closely spaced peaks in Laue diffraction patterns. These were then used to determine the packing structure of collagen by MIR.

Conclusions: Our electron density map is the first obtained from a natural fiber using these techniques (more commonly applied to single crystal crystallography). It reveals the three-dimensional molecular packing arrangement of type I collagen and conclusively proves that the molecules are arranged on a quasihexagonal lattice. The molecular segments that contain the telopeptides (central to the function of collagen fibrils in health and disease) have been identified, revealing that they form a corrugated arrangement of crosslinked molecules that strengthen and stabilize the native fibril.

Introduction

Collagen is one of the most abundant proteins within the animal kingdom, with diverse locations and specific roles concerned principally with the maintenance of tissue and cellular shape, strength, and structural integrity. As one of the most crucial components of the extracellular matrix—the arena of many important cellular functions—it is perhaps surprising that relatively little is known about the lateral supermolecular organization of collagen. Relating the structure of connective tissue matrices at the molecular level to their function is not just an academic problem; it is a vital step in understanding the pathology and basis for possible treatment of many human diseases. In the context of diseased tissues, abnormalities of collagen molecular structure may well adversely affect the packing of collagen molecules in fibrillar-forming types (in bone, skin, tendon, and so on), which in turn reduces the stability of these fibers (characterized by reduced fibril diameter) and produces weaker and structurally inferior connective tissues. Clearly, then, understanding the supermolecular (packing) structure of healthy connective tissue is a prerequisite to elucidating the pathology of diseases such as Osteogensis inperfecta and rheumatoid arthritis, to name but two.

It was over 30 years ago that Crick [1] observed, "The superlattice of collagen is a neglected problem and it is time somebody took it up again." This statement seems to be just as true today. While many models attempting to combine X-ray diffraction data, electron microscopic results, and biochemical data have been proposed over the years [2–21], no consensus structure has emerged. We have circumvented this impasse by taking the model independent approach of solving the phases for the reflections in the native X-ray diffraction patterns by multiple isomorophous replacement (MIR) and by calculating the electron density map directly. By using diffraction data from native and isomorphous fibers, we ensure that the unresolved questions of the supermolecular organization of the tissue are addressed, since the fiber diffraction patterns obtained from tendons arise due to the superlattice formed by the packing arrangement of the collagen molecules themselves, rather then from isolated collagen chains.

Biological fibers have a much higher degree of order in the direction parallel to the fiber axis than perpendicular to it. Hence, the resolution in our electron density map of collagen, in contrast to that from a crystal, is anisotropic, being of higher resolution in the direction parallel to the fiber axis and less in the lateral direction. However, the lateral resolution is sufficient to resolve

[4] Correspondence: jorgel@tigger.cc.uic.edu
[5] Present address: Laboratory for Molecular Biology (MC567), Department of Biological Sciences, University of Illinois, 900 S. Ashland Avenue, Chicago, Illinois 60607.

Key words: collagen; extracellular matrix; structure; X-ray diffraction; packing; fibril

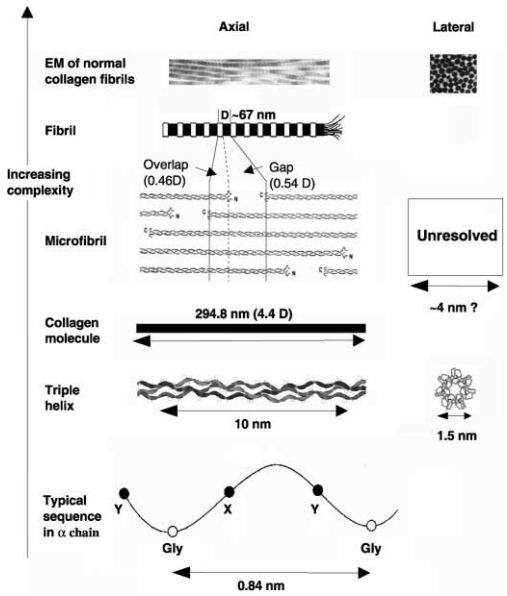

Figure 1. Known Structural Hierarchy of Collagen

The principal component of fibrillar tissues is the collagen fibril that is made of collagen triple helices that are axially staggered by D (67 nm) and regularly organized in the lateral direction (unresolved structure). The topmost element in the diagram shows normal positively stained collagen fibrils in mammalian tissue, just below which is a schematic representation of a single fibril that shows the dark/light banding pattern of negatively stained isolated fibrils. The structural hierarchy of polypeptide to fibril is shown though the bottom to top figure elements.

the large-scale supermolecular structural question—how the collagen molecules are arranged in the fibrils. Other studies [22] have focused on the structure of short soluble collagen-like peptides ("short" meaning that these peptides are approximately 30 amino acids in length, in contrast to the over 1000 residues that make up the length of the native collagen molecule). Although such studies have been very useful, they cannot comment on the native packing structure of the five molecular segments found within the naturally occurring unit cell and, hence, the need for the approach presented here.

The fundamental units in the animal extracellular matrix are collagen fibrils which, in tendons, are parallel to each other and to the long axes of the tendons and are up to ~1000 nm in diameter. The type I collagen molecules which make up most of tendons are some 300 nm long (L). In electron micrographs, the fibrils show a periodicity along the axis (the D period), which is shown by X-ray diffraction to be 67 nm (234 amino acid residues) in the native, hydrated state. The periodicity is due to the molecules being staggered in the axial direction by 67 nm with respect to each other (see Figure 1). Since L = 4.46 D, this means that there are gaps of 0.54 D between the molecular ends distributed, D apart, on a regular array within the fibril in the direction parallel to the fibril axis. The X-ray fiber diffraction pattern from tendon shows over 140 reflections on a central linear line (the meridian) through reciprocal space parallel to the fiber axis. These meridional reflections index on the repeat of 67 nm. Our previous MIR study [23] had produced a profile of the electron density to 0.54 nm resolution of collagen fibrils projected onto the fibril axis, sufficient to ascertain the conformation of the telopeptides. Each collagen molecule consists of four segments of length D and a fifth segment of length 0.46 D. Segments 1 and 5 contain the nontriple helical telopeptides that are sites of the intermolecular crosslinks and provide the structural integrity and strength of the fibrils (see Figures 1 and 2).

In the X-ray fiber diffraction pattern, the strong near-equatorial reflection at a spacing close to that expected from near-neighbor molecules (1.3 nm) was shown to be split into three strong components and to suggest a quasihexagonal packing of collagen molecules [8]. An improved set of unit cell parameters was deduced by more detailed measurements on the X-ray pattern ([11, 13]; see Table 1). Further improvements in the quality of the X-ray diffraction patterns obtained by using synchrotron radiation and image treatment techniques have allowed the measurement of the intensities of the Bragg reflections from native tissue and from isomorphous derivatives. Derivatives used were either iodine (which labels the tyrosine amino acids that only occur in the telopeptides) and gold chloride (which labels histidine and selected methionine residues) [23, 24]. Building on these advances, we have solved the structure of collagen packing in situ by MIR. The packing of the collagen triple helices is immediately revealed as quasihexagonal, the location of the crosslinking helices have been determined, and a pattern of intermolecular and possibly interfibrillar crosslinking is inferred on the basis of this new data.

Results and Discussion

Direct Determination of the Electron Density Map

The tendon specimens from which we have obtained the X-ray patterns are not, of course, single crystals. They are made up of millions of fibrils which, while closely parallel to each other, are in fact misaligned from true parallelism by 1°–2°. The Bragg reflections in the X-ray patterns are therefore drawn out into arcs which often overlap with the arcs from other reflections. The fibrils are randomly oriented azimuthally. Finally, the crystallographic unit cell is triclinic (see Table 1) with its long axis inclined at 5° to the axis of the fibril. Hence, the X-ray pattern is a 360° rotation pattern with some crystal (fibril) misalignment and with the long *c* axis not coincident with the axis of azimuthal rotation. In addition, the X-ray patterns always show diffuse scattering as well as the Bragg reflections. This diffuse scatter originates from the specimen (rather than the X-ray optics) and is specifically localized along the principal layer lines in the collagen molecular transform. There are, therefore, significant technical challenges to be met in estimating the intensities of the Bragg reflections. These are primarily acquiring patterns of sufficiently high quality to allow accurate separation of the diffuse background from the Bragg reflections, determination of the intensities from these closely spaced reflections, and then solving the phases. These problems were largely

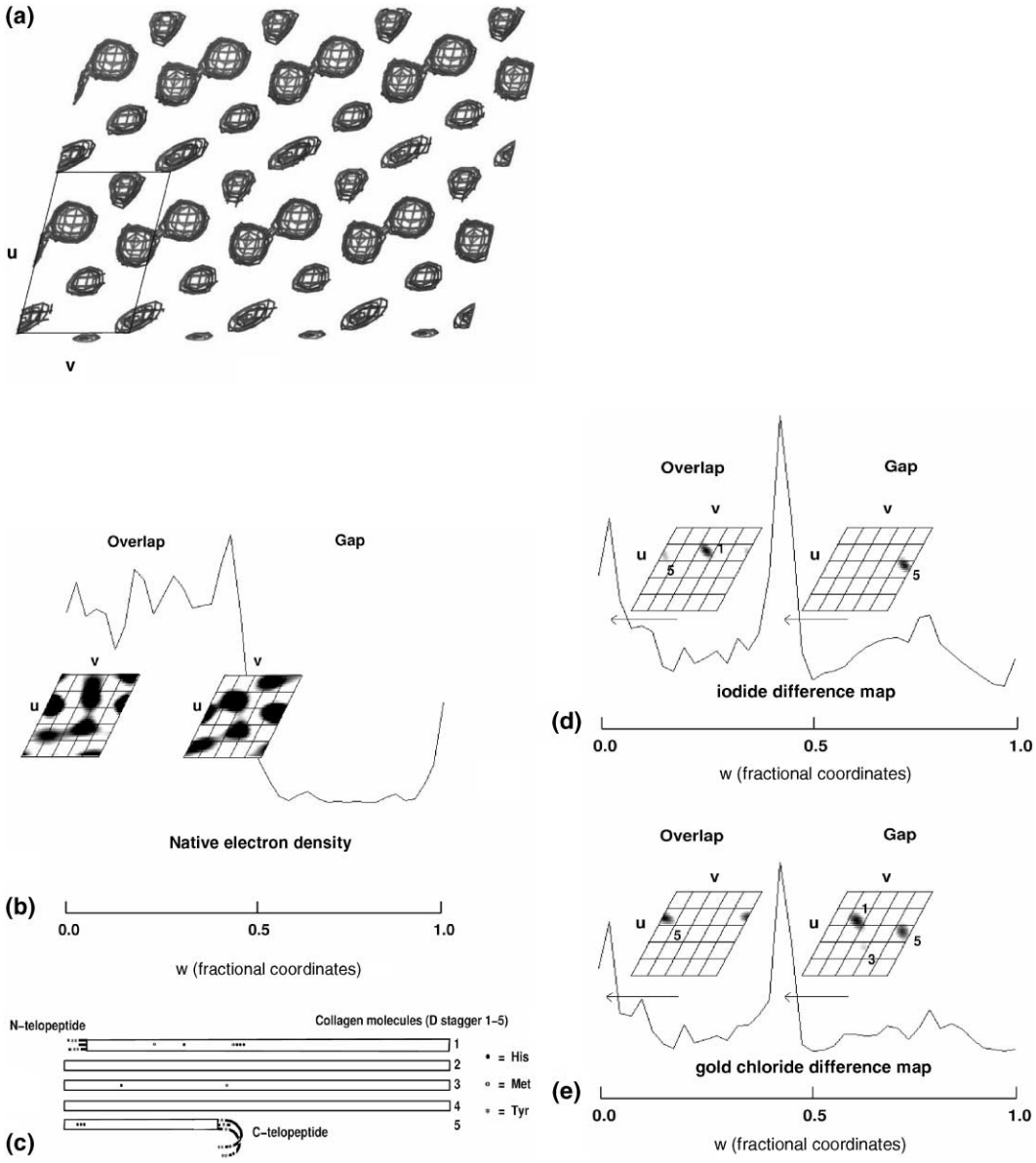

Figure 2. Quasihexagonal Packing and Identification of the Telopeptide-Containing Segments

(a) The quasihexagonal packing of the collagen molecules seen in a transverse section through the electron density map generated by plotting 2×4 unit cells with an axial thickness of 0.1 units of the unit cell *c* axis in the region of the C-terminal telopeptide. One unit cell is outlined (bottom left).

(b) The native axial and lateral electron density profiles. Here, the axially projected electron density is shown as a line profile that exhibits the characteristic density differences in the gap and overlap regions obtained previously [23]. Superimposed on this at the axial levels of the N- and C-terminal telopeptides are the corresponding two-dimensional crosssections through the electron density map showing the lateral electron density profile obtained in this study.

(c) Our knowledge from [23] of the axial positions of the heavy-atom binding sites on each molecular D-periodic segment (numbered 1–5) is shown here. These were predicted from the amino acid sequence and confirmed in previous studies [23, 24]. This information was used in deducing the three-dimensional positions of the heavy atoms from our three-dimensional Patterson difference maps.

(d) The iodide derivative difference map.

(e) The gold chloride derivative difference map.

The iodide and gold chloride derivative difference maps confirm the location of segments 1 and 5, while only the gold chloride derivative shows the possible location of segment 3. The segment numbers are indicated on the diagrams. The arrangement of segments 1 and 5 in particular imply that the nonidealized [13] quasihexagonal packing arrangement of collagen molecules is due in part to the intermolecular crosslinks between these segments. The thickness of the slices shown in (b), (d), and (e) is the same as shown in (a).

overcome by improving sample preparation and by optimizing the experimental conditions to enhance the quality of the X-ray patterns as well as by using software developed from well-established principles for analysis and interpretation (see Experimental Procedures; Tables 1–3).

Table 1. Experimental Details

Source	ID18 BioCAT undulator beamline (APS)
Beam size (unfocused)	0.8 mm × 3.6 mm (at sample)
Flux	1.5×10^{13}–5×10^{13} ph/s @ 100 mA
Path length	1026 mm
Wavelength	1.003 Å
Detector	Fuji BAS V image plates with a Fuji BAS2500
Data sets	One native and two isomorphous derivative fiber diagrams
Derivatives	Iodide and gold chloride
Unit cell dimensions	Triclinic a = 39.97, b = 26.95, c = 677.9 Å
Unit cell angles	$\alpha = 89.24°$ $\beta = 94.59°$ $\gamma = 105.58°$
Space group	P1
Resolution limits (Å)	5.4 meridional, 10.0 equatorial
Observed unique reflections	410 (124 meridional, 286 equatorial/off-meridional)
Figure of merit	0.4

Table 2. Rms Derivatives for Determination of Closely Spaced Bragg Reflections

	Group				
Data Set	1	2	3	4	Mean
Native	3.23	4.21	1.57	2.99	3.02
Iodide derivative	0.97	1.69	1.48	2.08	1.56
Gold derivative	1.64	1.55	0.64	1.52	1.34

Interpretation of the Electron Density Map

The electron density map shown in Figures 2 and 3 is the first direct visualization of the molecular packing in collagen fibrils. Figure 2a shows transverse sections $0.0 < z < 0.1$ of the fractional coordinate *z* through the native fibril at right angles to the unit cell *c* axis and at the level of the C-terminal telopeptide. It immediately reveals that the molecules are packed on a hexagonal lattice. Since the molecular segments are not all identical, the packing is better described as quasihexagonal. The relative positions of the five molecular segments can be seen to deviate from the expected ideal quasihexagonal coordinates in such a way as to indicate bunching of the molecular segments. This is most probably due to the connectivity between the five different molecular segments in one D period, as suggested by Fraser et al. [13].

Figure 2b shows the native, axially projected electron density profiles and the *z* levels of the N and C telopeptides. The collagen molecules are tilted to the *c* axis in the unit cell. As a consequence, the distribution of the molecular segments is best shown by local projections of thin slices through the unit cell rather than a single projection down the *c* axis of the whole unit cell. A schematic diagram of the positions of the amino acids that act as binding sites for heavy atom labeling is shown in Figure 2c. The electron density difference maps in Figures 2d and 2e show corresponding axial projections and lateral sections through the (d) iodinated and (e) gold chloride difference electron density maps at the level of the telopeptides. Since iodine principally labels the telopeptides, this makes it clear that segments 1 and 5 of the collagen molecules are located next to each other and that they form a continuous corrugation of crosslinked nearest neighbors along a line approximately parallel to the *a* axis of the unit cell. This fits well with biochemical studies on the intermolecular crosslinks [25].

The molecules are ordered to different degrees within the D-repeating unit. Figure 3a, which is a view of the electron density perpendicular to the fibril axis, shows that they are best ordered at the axial level of the telopeptides (the *c* axis of the electron density map has been compressed by 5 times in all axial views). In the rest of the overlap region, they are ordered sufficiently well to allow recognition of the paths of the individual segments of the collagen molecules, but they are poorly ordered in the gap region where no individual molecular paths can be discerned. The fibrils, therefore, contain a D-periodic (67 nm repeat) pattern of ordered and disordered domains along the fibril axis. This finding is consistent with the observation that the X-ray pattern contains both Bragg reflections and specifically localized diffuse scatter. In our electron density map, the highest degree of order is at the axial level of the N and C telopeptides. The folded nature of the C-terminal telopeptide and the contracted nature of the N telopeptide have been established by our one-dimensional study of collagen electron density [23]. The packing density in the plane of the telopeptides would be the highest in the unit cell and would possibly restrict lateral and azimuthal movement of chains and, hence, account for the higher degree of order in this region. Lateral and azimuthal motions are thought to be appreciable in the collagen fibril as demonstrated by NMR studies [26] and by analysis of the X-ray pattern [15]. Our results show that lateral disorder is highest in the gap region. This difference in the degree of molecular order, with the highest order at the level of the telopeptides and the lowest in the gap region with an intermediate degree of order in the overlap region, fits with the classification proposed by Jones and Miller [27] and Hulmes and coworkers [19]. Differences in the lateral ordering of long molecules were also reported by Phillips and coworkers [28] in studies of highly hydrated crystals of the fibrous protein tropomyosin.

Table 3. Definition of Row-Line Groups

Group	Reciprocal Space Range	(h,k,l) Index of Row-Lines
1	3.9 nm	(1,0,/),($\bar{1}$,0,/)
2	2.7 nm	(0,1,/),(0,$\bar{1}$,/),(1,$\bar{1}$,/),($\bar{1}$,1,/)
3	1.8 nm	($\bar{1}$,$\bar{1}$,/),(1,1,/),(2,0,/),($\bar{2}$,0,/),($\bar{2}$,1,/),(2,$\bar{1}$,/)
4	1.3 nm	(2,1,/),($\bar{2}$,$\bar{1}$,/),($\bar{1}$,2,/),(1,$\bar{2}$,/),(0,2,/),(0,$\bar{2}$,/),($\bar{3}$,1,/),(3,1,/),(3,0,/),($\bar{3}$,0,/)

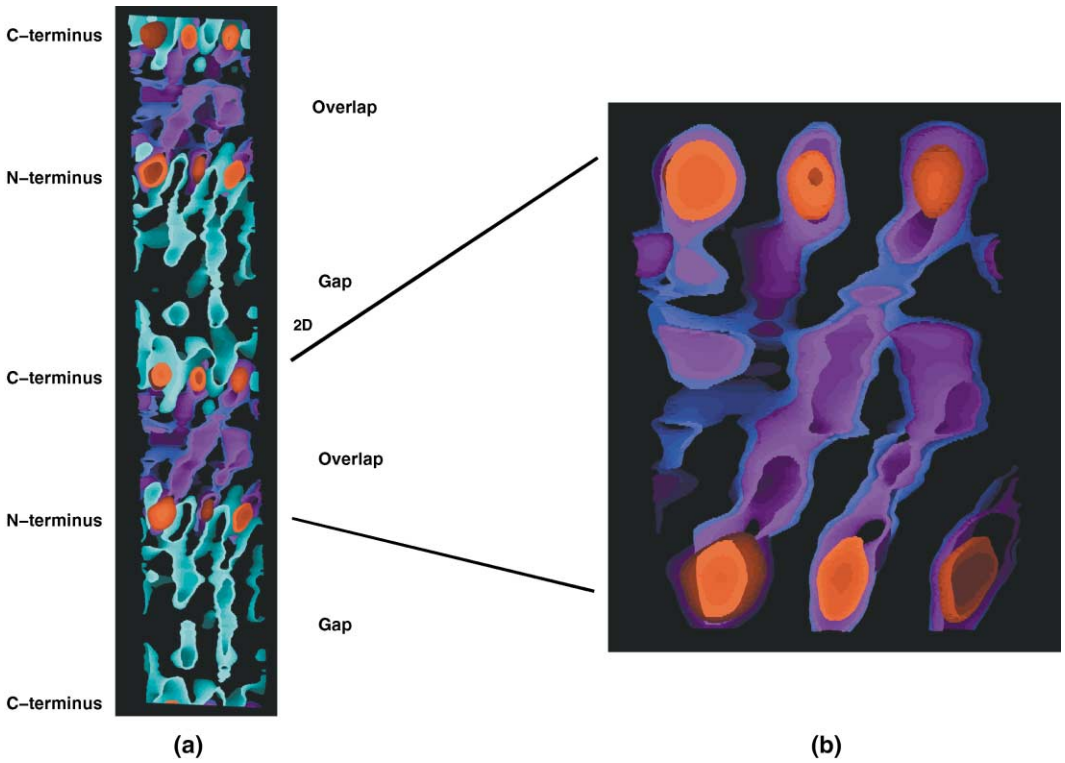

Figure 3. Electron Density Map of Type I Collagen Molecular Packing

(a) A view perpendicular to the fibril axis of two-unit cells (D periods). The molecular packing arrangement is most clearly discernible at the axial level of the telopeptides (shown as the C-terminal and N-terminal regions in the figure) but is also discernible within the rest of the overlap region. There are no complete but only partially discernible molecular paths in the gap region. This implies a large degree of lateral disorder of the molecules in the lower packing density gap region, as predicted in several model studies [6, 15, 19]. Low electron density is commonly encountered in macromolecular crystallography in regions of crystal structures subject to thermal motion (such as chain loops, which disappear into the background). The electron density is viewed here as a 20 Å slice, perpendicular to the fibril axis (which has been compressed 5 times).

(b) The molecular tilt of the collagen molecular segments of the overlap region are observed to follow a vector approximately parallel to the line (0,0,0) to (0,2,1) (u,v,w, as in Fraser et al. [15]). This corresponds to a tilt of about 5° relative to the *c* axis of the unit cell. The molecular segments in the overlap region follow parallel paths. The formation of intermolecular crosslinks and the higher packing density at the interfaces of the overlap/gap regions (at the telopeptides, Figure 2b) ensures that the overlap region is well ordered, particularly in the plane of the telopeptides in contrast to the less-ordered state in the gap. The overlap region is illustrated here with the *c* axis compressed by 5 times to show clearly the tilt of the individual molecular segments. The view, slab, and axial compression are as in (a).

Figure 3b is a view through the electron density map of the overlap region and parallel to one of the principal planes of the quasihexagonal lattice. It shows that the molecules are tilted with respect to the unit cell *c* axis by about 5° as proposed by Fraser et al. [15]. The image we have obtained is therefore consistent with findings well established by a variety of experimental techniques. However, it also reveals important new features about the molecular packing, the sites of intermolecular crosslinks, and the different degrees of molecular disorder.

Implications for Molecular Arrangement and Connectivity

The distribution of the molecular segments in the unit cell is of paramount importance since it can allow the distinction to be made between the different topologies of molecular packing that are theoretically possible. The packing topology, in turn, has implications for molecular assembly and the mechanical function of the collagen fibril. In the electron density maps presented here, the segments that contain the telopeptide regions are bunched together, away from the ideal positions in the quasihexagonal packing model. This relationship is not unexpected since the intermolecular connectivity is through telopeptide-derived crosslinks. This information can be used to reduce the number of possible packing arrangements, although by itself it is insufficient to distinguish between the sheet and compressed microfibril topologies proposed previously [8, 12] since this would require at least three segments to be defined. We can make a tentative positional assignment to segment number 3 (see Figures 2d and 2e), and if correct, this means that the packing topology is microfibrillar. Figure 4 uses our arrangement to show the chains of crosslinked segments 1 and 5 traversing the fibrils together with a possible scheme of intermolecular connectivity.

The First "Crystal" Structure for a Natural Fiber

The methods employed here to determine the three-dimensional structure of type I collagen have proven

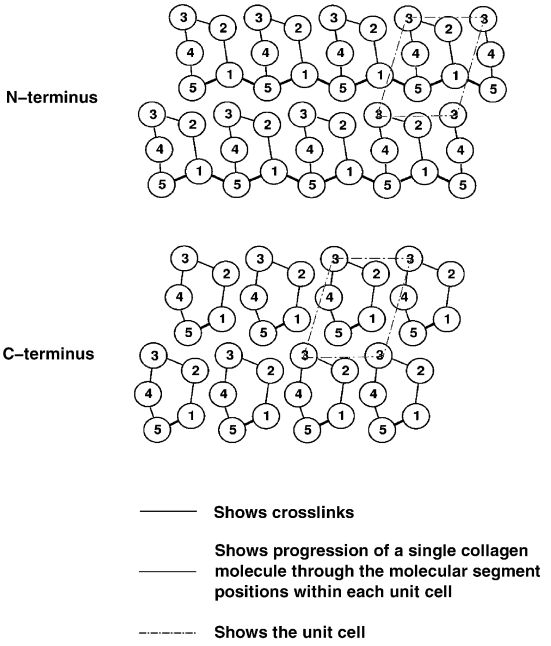

Figure 4. Possible Molecular Topologies

The assignment of positions in the unit cell to segments 1 and 5 (which contain the telopeptides) show that there are continuous chains of crosslinked molecules running across the fibril. This is illustrated by the bold lines between segments 1 and 5. More speculatively, we have assigned positions to the other segments based on our tentative identification of segment 3 in Figures 2d and 2e. If correct, this shows that microfibrils could exist and that there would be both inter- and intramicrofibrillar crosslinks, making microfibrils difficult to isolate.

successful. The three-dimensional molecular packing arrangement of type I collagen has been revealed, and, of the plethora of models proposed over half a century, the quasihexagonal model has been established. The molecular segments that contain the telopeptides have been identified and their coordinates in the unit cell determined, revealing that the molecular segments that are crosslinked are packed together more closely than the remaining three segments, particularly in the C-terminal region. Given the resolution of the X-ray diffraction patterns, we can only discern the arrangement of the molecules. No information is available on the atomic arrangement within molecules, although broad features such as intermolecular crosslinks can be recognized.

A tentative pattern of interconnectivity of collagen molecules has been suggested, and this involves the formation of both intermolecular and intermicrofibrillar crosslinks. This arrangement of intermicrofibrillar crosslinking would add a degree of integrity and strength to the fibril, as does the fact that the packing arrangement of segments 1 and 5 produces a corrugated arrangement of crosslinked molecules. These must contribute a significant degree of stability and resistance to shearing forces within and outside of the fibril and therefore throughout the connective tissue generally.

Biological Implications

Over 20 collagen types have now been discovered and are found widespread throughout the animal kingdom. They convey the principal mechanical support and structural organization in the extracellular matrix to connective tissues such as bone, skin, tendon, blood vessels, and other tissue types, as well as maintaining cell shape and probably a whole array of yet unidentified roles. Collagen type I is the most abundant and diversely located member of the collagen family, found principally in fibril form.

Collagen fibrils form in a self-assembly process where the correct axial registration between collagen molecules and telopeptide-mediated crosslink formation is crucial to the formation of normal fibrils. Specific axial and lateral organization is somehow derived in this process with the characteristic 67 nm axial staggering of molecular ends visualized by repeating dark and light bands in electron micrographs of negatively stained fibrillar tissues. Until now, the lateral organization has been little understood, with no clear and unambiguous evidence linking the structure of individual collagen triple helices with how they pack and how they are organized laterally to form the much larger supermolecular arrays, i.e., fibrils.

In studying the in situ structure of type I collagen by using synchrotron radiation, it has been possible to identify the packing mode of collagen molecules and the probable basis for this quasihexagonal packing as arising from the corrugated pattern of crosslinked molecular segments. This corrugated pattern must impart a degree of resistance to shearing forces within and outside of the fibril, and this, combined with the intermolecular and possibly interfibrillar crosslinking mediated by the telopeptide-containing molecular segments, explains something of the considerable tensile strength displayed by normal fibrillar connective tissue. Inhibition of the normal crosslinking process or even alteration of the (otherwise-normally crosslinked) telopeptide structure could result in a significant weakening of the connective tissues integrity, thereby rendering it more brittle and prone to damage.

If our tentative assignment of the position of segment 3 is correct, then the basis of fibrillar tissue construction is that of fibrils made from corrugated sheets of interlinked microfibrils, rather than the (perhaps) over-simplified concept of sheets of collagen chains. The existence of such interfibrillar crosslinking is appealing in that it shows a high degree of stability and continuity within the structural hierarchy of the fibril and connective tissue generally, as well as providing an explanation as to why such microfibrils have proved as yet difficult to isolate.

Experimental Procedures

Sample preparation, derivative labeling, and X-ray diffraction experiments were carried out as described previously [21, 23, 24]. Data sets for the native, iodide, and gold chloride derivatives were collected on the ID 18 BioCAT undulator beamline (Advanced Photon Source, Argonne National Laboratory, IL); for details see Table 1.

Prior to data extraction, background scatter was removed according to the procedures described in Wess et al. [18, 21]. Spatially overlapped reflections were separated by using in-house software that made use of the Metropolis algorithm [29]. The coordinate position of the intensities in the diffraction pattern were fixed according to the unit cell parameters previously described [18], which marked the center of a two-dimensional Gaussian function used to determine each Bragg reflection intensity in the off-meridional diffraction pattern. The optimum set of separated X-ray diffraction

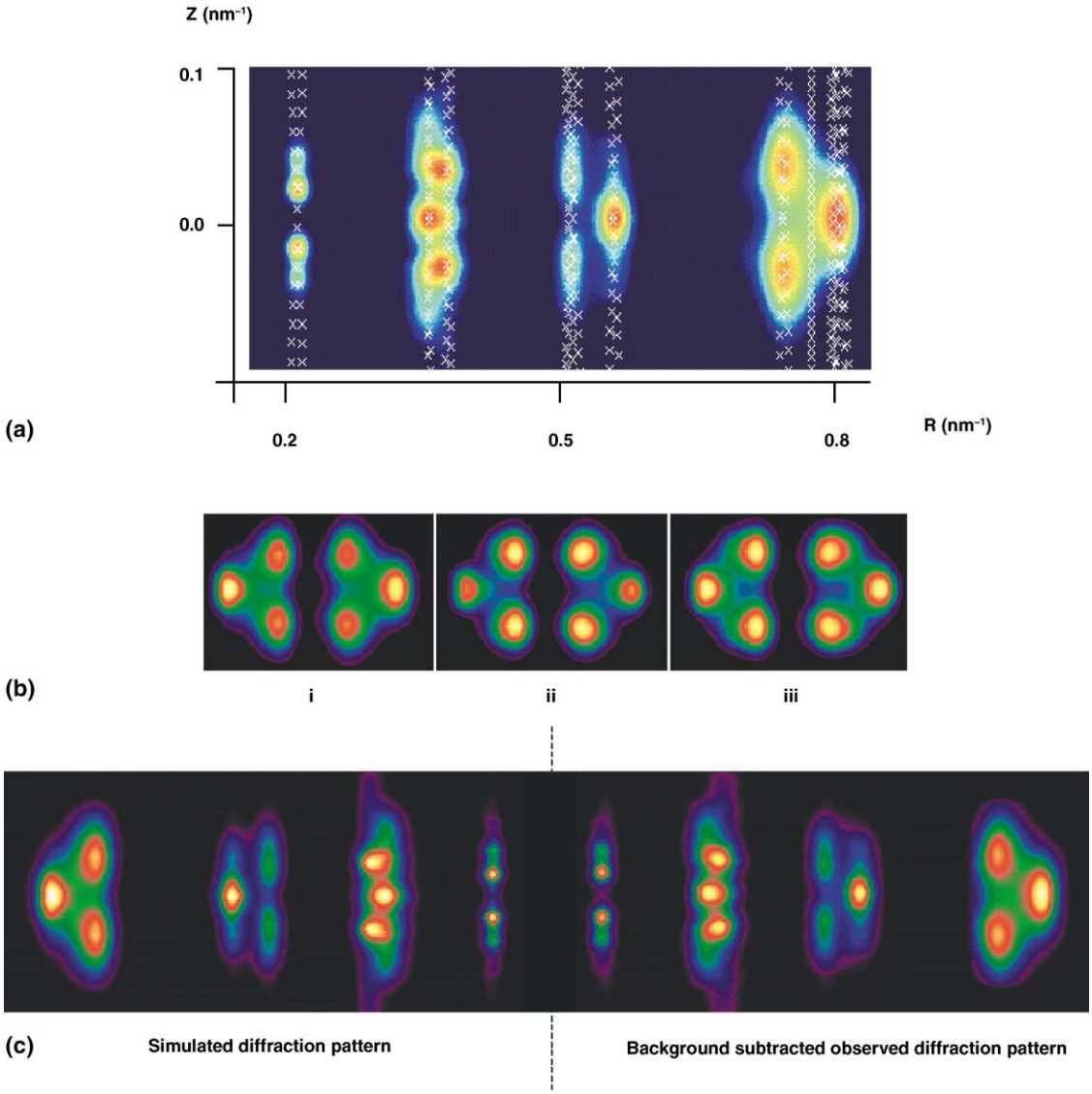

Figure 5. Background Subtraction and Intensity Determination

(a) Diffuse background subtracted equatorial diffraction pattern (low-angle section) from native rat tail tendon. The reciprocal lattice points are marked by crosses and were calculated from the triclinic unit cell [13, 18]. The indices of reflections indicated a high degree of overlap, particularly in the area corresponding to R = 0.8 nm^{-1} and Z = 0.018 nm^{-1}, where R and Z are cylindrical polar coordinates of the cylindrically projected central section of the fiber diffraction pattern. Note that this shows only the most closely spaced reflections. Other Bragg reflections with greater spatial separation appear above this region in the diffraction pattern and were also used in the phase calculations.

(b) The area in the region of R = 0.8 nm^{-1}. This particular area from the observed and simulated diffraction patterns of the native and two derivatives are, from left to right, simulated and observed native, simulated and observed iodide derivative, and simulated and observed gold chloride derivative. The region corresponds to the "triplet" of intensity and derives from the interplanar spacing of molecules in the quasihexagonal packing scheme. Ten row-lines of closely spaced Bragg peaks contribute directly to the triplet. Clear differences can be seen between the intensities of the reflections with the same (h,k,l) indices in the observed native and derivative X-ray patterns, as well as good agreement between each of the simulated and observed diffraction patterns.

(c) The quality of the intensity determination is made more evident in the full low-angle equatorial native diffraction pattern after background subtraction. The right side is the observed pattern; the left side is the simulated pattern used to obtain the Bragg intensities.

intensities obtained in this manner was used to calculate the structure factor phases (see Figures 5 and 6 and Table 2).

Putative information about the relative distances between the 1st and 5th molecular segments was obtained from the iodide and gold chloride difference Patterson functions with reference to the known axial distribution of the heavy atom vectors [23, 24] and the amino acid sequence. This information was then employed as a starting point to solve the phase problem for the three-dimensional unit cell with the Xtalview crystallographic software suite [30] (specifically Xfft and Xpatpred) to produce a visualization of the three-dimensional molecular packing arrangement of collagen molecules. The axial positions of the heavy atoms as determined from the previous one-dimensional study were combined with initial difference Patterson function data relating to the lateral structure to obtain an estimate of lateral heavy atom groupings. The three-dimensional coordinates of the heavy-atom binding sites were estimated and refined using Xpatpred and Xheavy of the XtalView crystallographic software suite [30].

Calculation of the electron density map was performed by using 410 unique reflections that were obtained from each of the native

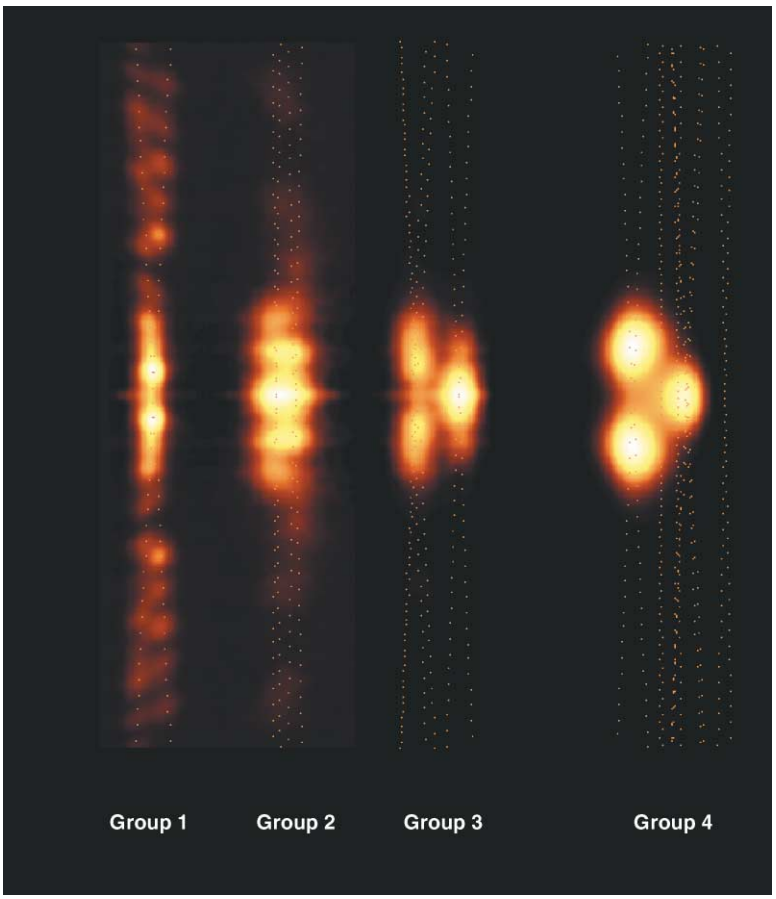

Figure 6. Off-Meridional Diffraction Pattern of Iodinated Derivative Tendon

Background subtracted false color image showing the indexed Bragg reflection positions as spots and the row-line group definitions listed in Table 3.

and two derivative patterns. The reflections used lay in the (h,k,l) range $(\bar{3},2,0)$ to (3,2,12) and meridional intensities (0,0,13) to (0,0,124) determined previously [23]. Xheavy was applied to the native and derivative data sets to calculate the native phases, the electron density map, and the difference Fourier maps. Table 1 summarizes the experimental details.

Acknowledgments

We wish to acknowledge the staff at the European Synchrotron Research Facility and Daresbury Synchrotrons for help and assistance while we developed our X-ray techniques and the Carnegie Trust (Scottish Universities) for travel and subsistence support for J.P.R.O.O. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Basic Energy Sciences, Office of Science (under contract number W 31 109 ENG 38). BioCAT is a National Institutes of Health supported Research Center (RR-08630). This work was also supported by the Biotechnology and Biological Sciences Research Council (98/B13991). We are grateful to Malcolm Walkinshaw and Constance Jeffery for comments on the manuscript.

Received: June 1, 2001
Revised: September 6, 2001
Accepted: September 10, 2001

References

1. Crick, F.H.C. (1966). Principles of biomolecular organization. In CIBA Foundation Symposium. (London: J. and A. Churchill) p. 132.
2. Bear, R.S. (1952). The structure of collagen fibrils. Adv. Prot. Chem. *7*, 69–160.
3. Smith, J.W. (1968). Molecular packing in native collagen. Nature *219*, 157–158.
4. Miller, A., and Wray, J.S. (1971). Molecular packing in collagen. Nature *230*, 437–439.
5. Miller, A., and Parry, D.A. (1973). Structure and packing of microfibrils in collagen. J. Mol. Biol. *75*, 441–447.
6. Woodhead-Galloway, J., and Machin, P. (1976). Modern theories of liquids and the diffuse equatorial x-ray scattering from collagen. Acta Crystallogr. A *32*, 368–372.
7. Woodhead-Galloway, J., and Young, H. (1978). Probabilistic aspects of the structure of the collagen fibril. Acta Crystallogr. A *34*, 12–18.
8. Hulmes, D.J.S., and Miller, A. (1979). Quasi-hexagonal molecular packing in collagen fibrils. Nature *282*, 878–880.
9. Ruggeri, A., Benazzo, F., and Reale, E. (1979). Collagen fibrils with straight and helicoid microfibrils: a freeze fracture and thin section study. J. Ultrastruct. Res. *68*, 101–108.
10. Fraser, R.D.B., Macrae, T.P., and Suzuki, E. (1979). Chain conformation in the collagen molecule. J. Mol. Biol. *129*, 463–481.
11. Miller, A., and Tochetti, D. (1981). Calculated X-ray diffraction from a quasi-hexagonal model for the molecular arrangement in collagen. Int. J. Macromol *3*, 9–18.
12. Piez, K.A., and Trus, B.L. (1981). A new model for packing of type I collagen molecules in the native fibril. Biosci. Rep. *1*, 801–810.
13. Fraser, R.D.B., MacRae, T.P., Miller, A., and Suzuki, E. (1983). Molecular conformation and packing in collagen fibrils. J. Mol. Biol. *167*, 497–521.
14. Hulmes, D.J.S., Holmes, D.F., and Cummings, C. (1985). Crystalline regions in collagen fibrils. J. Mol. Biol. *184*, 473–477.
15. Fraser, R.D.B., Macrae, T.B., and Miller, A. (1987). Molecular packing in type I collagen fibrils. J. Mol. Biol. *193*, 115–125.
16. Raspanti, M., Ottani, V., and Ruggeri, A. (1989). Different archi-

tectures of the collagen fibril: morphological aspects and functional implications. Int. J. Biol. Macromol. *11*, 367–371.
17. Fratzl, P., Fratzl-Zelman, N., and Klaushofer, K. (1993). Collagen packing and mineralization. Biophys. J. *64*, 260–266.
18. Wess, T.J., Hammersley, A., Wess, L., and Miller, A. (1995). Type I collagen packing, conformation of the triclinic unit cell. J. Mol. Biol. *248*, 487–493.
19. Hulmes, D.J.S., Wess, T.J., Prockop, D.J., and Fratzl, P. (1995). Radial packing, order and disorder in collagen fibrils. Biophys. J. *68*, 1661–1670.
20. Lee, J., and Scheraga, H.A., and Rackovsky, S. (1996). Computational study of packing a collagen-like molecule: quasi-hexagonal vs "Smith" collagen microfibril model. Biopolymers *40*, 595–607.
21. Wess, T.J., Hammersley, A., Wess, L., and Miller, A. (1998). Molecular packing of type I collagen in tendon. J. Mol. Biol. *275*, 255–267.
22. Bella, J., Eaton, M., Brodsky, B., and Berman, H.M. (1994). Crystal-structure and molecular-structure of a collagen-like peptide at 1.9-angstrom resolution. Science *266*, 75–81.
23. Orgel, J.P., Wess, T.J., and Miller, A. (2000). The *in situ* conformation and axial location of the intermolecular crosslinked non-helical telopeptides of type I collagen. Structure *8*, 137–142.
24. Bradshaw, J.P., Miller, A., and Wess, T.J. (1989). Phasing the meridional diffraction pattern of type I collagen using isomorphous derivatives. J. Mol. Biol. *205*, 685–694.
25. Bailey, A.J., Light, N.D., and Atkins, E.D.T. (1980). Chemical cross-linking restrictions models for the molecular organization of the collagen fibre. Nature *288*, 408–410.
26. Torchia, D.A. (1982) Solid state NMR studies of molecular motion in collagen fibrils. In Methods of Enzymology, Volume 82, L. Cunningham and D. Fredriksen, eds. (New York: Academic Press), pp. 174–186.
27. Jones, E.Y. and Miller, A. (1991). Analysis of structural design-features in collagen. J. Mol. Bio *218*, 209–219.
28. Phillips, G.N., Fillers, J.P., and Cohen, C. (1980). Motions of tropomyosin. Crystals as metaphor. Biophys. J. *32*, 485–502.
29. Metropolis, N., Rosenbluth, M., Teller, A., and Teller, E. (1953). Equation-of-state calculations by fast computing machines. J. Chem. Phys *21*, 1087.
30. McRee, D.E. (1993). Practical Protein Crystallography (San Diego, CA: Academic Press).