The three-dimensional molecular packing structure of collagen

J. Orgel♦, T. Wess♦, A. Miller♦, T. Irving●, and A. Hammersly■

♦*Department of Biological and Molecular Sciences,University of Stirling, Stirling, Scotland, FK9 4LA UK*
●*Department of Biological Chemical and Physical Sciences, Illinois Institute of Technology, 3101 South Dearborn Street, Chicago, IL 60618-3793 USA*
■*ESRF, BP220 F-38043 Grenoble, France*

Introduction

Collagen is the single most abundant protein in the animal kingdom, and is the principal constituent of the extracellular matrix and connective tissues. The definition of the fibril structure of type I collagen is of fundamental importance to understanding such phenomena as crosslinking in fibrils and the packing of microfibrils to form crystallites of supramolecular arrays. This is of particular significance in pinpointing where other molecules may bind to the various collagens within the extracellular matrix and understanding how both healthy and diseased tissues respond to biomechanical stress. This will help to identify the molecular basis of connective tissue diseases.

The principal limiting factor in natural collagen crystallite packing studies so far has been their model-dependent nature, which often fails to take into account the duality of the packing system-crystalline arrays of molecules organized within a liquid-like system. With recent advancements and developments to synchrotron radiation stations such as the Bio-CAT beamline [Advanced Photon Source (APS), IL, USA], it has become feasible to obtain diffraction data of sufficiently high quality from crystallites within collagen fibrils to determine the molecular structure of these crystalline regions for the first time.

Using multiple isomorphus addition, it has been possible to solve the high-angle one-dimensional axial structure of type I collagen [1]. Using similar techniques, an equivalent solution for the structure factors relating to the three-dimensional structure has been sought in a combined effort by the Stirling University x-ray fiber diffraction group and the team at the Bio-CAT (18-ID) beamline.

The major problem in obtaining quantifiable data from the collagen type I x-ray diffraction pattern is the partial overlapping of a large number of Bragg reflections and the nature of the diffuse background [2 and 3]. These problems have been overcome in this study by carefully removing the diffuse background and modelling the x-ray diffraction pattern to obtain the structure function amplitudes. Obtaining data of sufficient quality to be able to even attempt to extract the intensities is a major achievement.

Methods and Materials

Sample preparation and x-ray diffraction experiments were carried out as described in [3]. The staining procedures used to produce isomorphic derivatives are given by [4]. Iodide is believed to bond covalently with tyrosine residues, although iodohistidine can also form under the same reaction conditions [5]. While gold chloride is known to form complexes with cysteine, methionine, and histidine, only the latter two are found in type I collagen.

X-ray diffraction patterns of rat tail tendon were obtained at the APS Bio-CAT beamline (18-ID) using a sample to detector distance of 1029 mm. A CCD detector was used to monitor the quality of diffraction from a particular sample, and a phosphor image plate scanned on a molecular dynamics flatbed scanner was used to collect data. Background subtraction of the sample-derived diffuse scatter was performed utilizing the suite of programs FIT2D [2]. Intensities where extracted via a simulated annealing based algorithm written specifically for the purpose of modelling the collagen equatorial diffraction pattern.

Results

Over 150 Bragg peaks were indexed according to the triclinic unit cell of [2 and 6], and the intensity of these peaks was measured for the native and two-derivative samples. A comparison of the background-subtracted native diffraction pattern with the simulated reconstruction of the diffraction pattern using the intensities obtained from native fiber diagram is shown in Figure 1. The R-factor for the simulation is approximately 0.07.

Discussion

Presented here for the first time is the basis of an unambiguous determination of the three-dimensional molecular packing structure of hydrated fibrillar type I collagen. The structure functions have been determined through a means that is not biased by assumptions of the structure of the protein system itself. The validity of the intensity estimations are currently being tested in the process of solving the phase problem. So far, it has been possible to produce a three-dimensional electron density map of the packing structure localized to the plane of the telopeptide regions. The phase solution is being carefully examined and the corresponding electron density map will be published shortly.

simulated intensites (maxium h=3, k=2, l=12)   observed reflections (background subtracted)
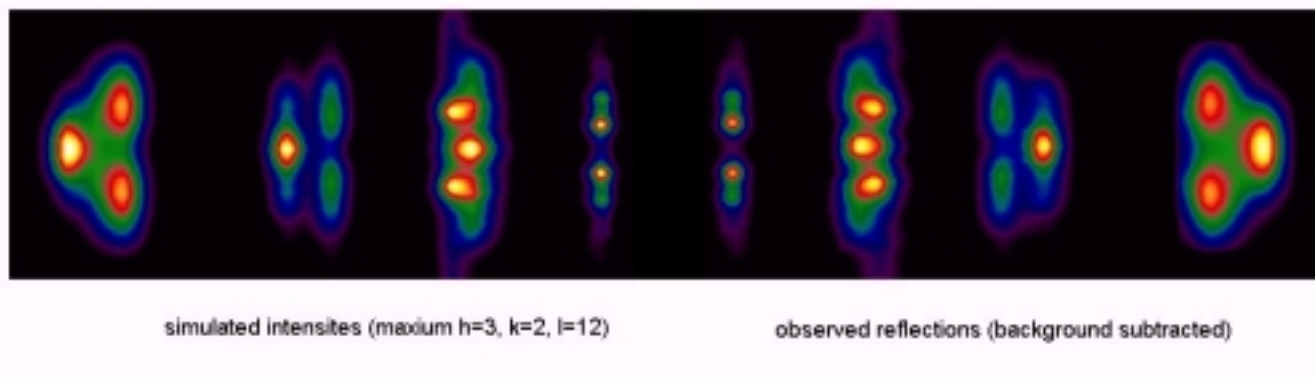
Figure 1: Lateral reflections from native type I collagen diffraction pattern compared with the simulated diffraction pattern used in estimating the intensities of the observed pattern.

References

[1]   J.P. Orgel, T.J. Wess, and A. Miller, *Structure* **8**, 137–142 (2000).

[2]   T.J. Wess, A.P. Hammersley, L. Wess, and A. Miller, *J. Mol. Biol.* **248**, 487–493 (1995).

[3]   T.J. Wess, A.P. Hammersley, L. Wess, and A. Miller, *J. Mol. Biol.* **275**, 255–267 (1998).

[4]   J.P. Bradshaw, A. Miller, and T.J. Wess, *J. Mol. Biol.* **205**, 685–694 (1989).

[5]   G.E. Means and R.E. Feeney, (Holden-Day, San Francisco, 1971) 175–182.
        R.D.B. Fraser, T.P. MacRae, A. Moller, and E. Suzuki, *J. Mol. Biol.* **167**, 497–521 (1983)