Acta Cryst. (1976). A32, 368

Modern Theories of Liquids and the Diffuse Equatorial X-ray Scattering from Collagen

BY JOHN WOODHEAD-GALLOWAY*

Laboratory of Molecular Biophysics, Zoology Department, University of Oxford, England and Department of Rheumatology, Stopford Building, Oxford Road, Manchester M13 9PT, England

AND PENELOPE A. MACHIN

Atlas Computing Laboratory, Chilton, Didcot, Berks., England

(*Received* 22 *July* 1975; *accepted* 27 *September* 1975)

The idea that the lateral arrangement of a proportion of molecules in the collagen fibril displays a short-range two-dimensional liquid-like order is investigated and shown to be true. Theoretical calculations of a dense disordered assembly of hard discs yield X-ray scattering curves which possess the features of the near equatorial diffuse scattering from rat-tail tendon collagen. The theoretical model used is the integral equation formulation of Percus & Yevick [*Phys. Rev.*(1958). **110**, 1] known to be almost exact for hard-disc potentials. The molecular diameter and number density of molecules obtained from the comparison with experiment are close to those suggested by Woodhead-Galloway, Hukins & Wray [*Biochem. Biophys. Res. Commun.* (1975). **64**, 1237–1244] to explain the ordered part of the collagen and to those suggested by Katz & Li [*J. Mol. Biol.* (1973). **73**, 351–369] on the basis of an investigation of the density of wet tendon. A brief discussion of elastoidin is included. Discrepancies between the observed and predicted scattering are discussed.

1. Introduction

A short-range order similar to that supposed to be characteristic of liquids seems to be a feature of the distribution of structural units, molecules, microfibrils, crystallites, *etc.*, in a number of animal and plant connective and skeletal tissues. Two differences may be noticed, however, between the local order displayed in the two sorts of situation; liquids are fluid – a consequence of *their* irregularity – whereas it seems that the tissues are static, at least in this respect. Second, structural tissues are usually fibrous, being relatively well ordered (although not always) in the axial direction – the short-range order refers to the lateral arrangement of molecules or groups of molecules, [the analogy with liquid crystals has been drawn in some examples (*e.g.* Neville & Luke, 1971)] so that one has in essence 'planar liquids'. Blaisie & Worthington (1969) suggested the idea of a planar liquid to account for X-ray observations of the distribution of retinal pigment in the eye of the frog, although Fraser, MacRae, Miller & Suzuki (1964) had considered it in a more structural context for the distribution of α-keratin microfibrils in the matrix, and more recently a similar point has been made for the distribution of chitin crystallites in insect skeletons (Neville, Parry & Woodhead-Galloway, 1976). In both the last two situations the actual distribution is visualizable directly using the electron microscope. Situations where short-range order has been proposed but the evidence for it is again X-ray diffraction and not so readily interpretable, are the collagens of rat-tail tendon (Burge, 1965; Hoseman, Dreissig & Nemetschek, 1974; Nemetschek & Hoseman, 1973) and dogfish-fin elastoidin (Woodhead-Galloway & Knight, 1975).

Since good quantitative theories of liquids have been proposed over the last decade or so, it is tempting to borrow these to account for the near-equatorial diffuse scattering that provides the basis of the suggestion in the collagens. There are, of course, the two drawbacks alluded to above. The two situations are not quite the same. The first, though the more serious in principle, is easily met by the ergodic theorem with its assertion that time and space averages are the same, and, of course, Bernal (1964) has proposed excellent static models of liquids. The second is a trivial point theoretically, but presents some technical difficulties. The two-dimensional problem of liquid-like disorder is more formidable that either the three- or one-dimensional problems (Rowlinson, 1964).

2. The theoretical approach

Modern theories of liquids presume that they are like dense gases rather than disordered crystals in the way that they scatter X-rays. That is not to say that the scattering they predict does not have some crystal-like features, but the power of the methods is that these features arise in the course of calculation rather than being introduced in an *ad hoc* way. Indeed, it might be claimed that such theories are the one area where structure has been succesfully predicted on the basis of potential-energy functions.

For the purposes of accounting for X-ray diffraction results, we may take it that the radial distribution function is a sufficient description of a liquid (although it does not necessarily define a statistical structure

* Present address: Medical Research Council, 20 Park Crescent, London W1N 4AL, England.

unambiguously which requires a complete set of multiparticle correlation functions). It is customary to write the pair distribution function, F_{12}, which is the probability density function for finding two molecules simultaneously a distance r apart in the liquid as

$$F_{12}=n_0^2 g(r)\,, \tag{2·1}$$

where n_0 is the molecular number density and $g(r)$ the radial distribution function, and Zernike & Prins obtained the formula for the X-ray interference function $S(k)$, usually called the structure factor in work on liquids, describing the scattering by such a distribution of molecules as

$$S(k)=1+n_0\int_V [g(r)-1]\exp(ik\,.\,r)\mathrm{d}r\,. \tag{2·2}$$

The term in brackets is often called the 'total correlation function' and represents deviations from the average density of the liquid. The total scattering from the liquid $I(k)$ is written, as is well known

$$I(k)/NI_0=f^2(k)S(k)\,, \tag{2·3}$$

where $f(k)$ is the Fourier transform of the electron distribution of the molecule, I_0 is the incident intensity and N the total number of scatterers.

Notice that crystallographers use the term 'structure factor' differently as the square root of a reflexion's intensity and that they use $f(k)$ as the form factor for an atom.

Present-day theoretical work on liquids has often attempted to calculate $g(r)$ using as parameters the intermolecular potential-energy function $V(r)$, temperature K_BT and density n_0. They are fundamental theories in that they begin explicitly or implicitly with the partition function following the work of Mayer. (For a review of Mayer's work see Salpetre, 1958.) From the radial distribution function the X-ray scattering can then be calculated. Since, however, the experimental quantity is $I(k)$, or, if sufficient is known about $f(k)$, it may be $S(k)$, and since it is inordinately difficult to obtain accurate estimates of $g(r)$, for example, because the observable data is limited, there seems on the whole little point in calculating $g(r)$ and then calculating the diffraction intensity. It is more elegant and economical to find a theory of $S(k)$ and use it directly for comparison, and, indeed, there are many real advantages in doing so (Woodhead-Galloway, 1968; Woodhead-Galloway, Gaskell & March, 1968).

A simple theoretical model, conceptually, which lends itself to the above approach and has provided the basis of understanding diffraction results for simple inert gases, and to some extent for liquid metals, is the integral equation formulation of Percus & Yevick (1958) (see also Lebowitz & Percus, 1966) based on Ornstein & Zernike's (1914) direct correlation function and employing the hard-sphere potential

$$\begin{aligned} V(r)&=+\infty \quad r<2R, \text{ the molecular diameter}\\ &=0 \qquad r>2R\,. \end{aligned} \tag{2·4}$$

This model has the attractive feature that it allows immediately, though not easily, in three dimensions (and one) a calculation of $S(k)$ in closed form in terms of simple trigonometric functions (Thiele, 1963; Wertheim, 1963). It has the added advantage that at a realistic density it provides an almost exact description of the scattering for the potential given in (2·4). A further advantage for calculation using such a potential is that only one real parameter, the packing fraction $\eta=n_0\pi R^3/6$, is involved. The molecular radius, R, enters calculation explicitly only as a trivial scaling factor. The temperature does not enter at all. The success of such a simple model in accounting for diffraction observations in real liquids seems to be that at the high densities encountered in liquids the structure is dominated by the packing of the impenetrable molecular cores, the 'hard spheres' of the model, and any long-range attractive parts in the potential, such as dispersion forces, serve only as a minor perturbation on the structure (Woodhead-Galloway *et al.*, 1968). In view of the success and manifest advantages for calculation of the approach, it is the one we shall employ here. However, since the problem is one of irregularity in two dimensions, the model is, of course, one for an assembly of 'hard discs' rather than spheres. Both in collagen and elastoidin the molecules are arranged almost parallel to the fibre axis. There seems to be some tilt, but it is small – a few degrees at most. In a first approximation the tilt will be ignored. In projection down the axis the molecules may, with some reservations, be regarded as discs, and the problem of the lateral packing one of the local order of an assembly of 'hard discs'. There is, of course, in rat-tail tendon, though not in elastoidin, X-ray evidence that part of the collagen exists with long-range lateral order (see Woodhead-Galloway, Hukins & Wray, 1975, for a possible interpretation), but we shall also ignore this feature.

3. Calculations

As was remarked earlier, two-dimensional problems of disorder are troublesome, and no closed-form solution has yet been shown to exist for the Percus–Yevick equation even for the simple hard-disc potential. However, recently, the structure factor $S(k)$ has been calculated for the model by Machin & Woodhead-Galloway (1975) with an eye to the sort of problem discussed here, and those results will be employed. The method used was to find a series solution to the Percus–Yevick equation using the packing fraction as an expansion parameter (a customary approach). Relatively good convergence of the series which finally employed eight terms was found up to a density of about $\frac{2}{3}$ that of closest packing ($\eta=0{\cdot}9069$). It was felt that at the moment further calculations along these particular lines are not warranted; the returns for the effort are small. Some further information was elicited by interpolating between the maximum value calculated and

that of closest packing. The calculations were of structure factors which are the quantities readily available from the Percus–Yevick (and direct correlation function appraoches generally). Total intensities $I(k)$ were calculated by assuming the molecule to have uniform electron density so that

$$f(k)=2\pi RJ_1(kR)/k\,, \qquad (3\cdot1)$$

where $k=2\pi/d$, d being measurements in reciprocal space converted to distance in real space.

4. Results and comparison with experiment

Fig. 1 shows structure factors $S(k)$ calculated for a number of values of the packing fraction η. Fig. 2 shows the total intensity $I(k)$ for a system with a packing fraction of 0·236 using 2·3 and 3·1. A first maximum can just be resolved, and its position corresponds to a value of $d/2R$ of about 2·5. Fig. 3 shows that at higher densities the peak is better resolved and moves towards the value expected for the (1,0) reflexion of a hexagonally close-packed lattice where $d/2R=\sqrt{3}/2=0\cdot866$. Table 1 shows the calculated variation in position of the first peak with packing fraction. The value at $\eta=0\cdot7854$ (equivalent to tetragonal packing) was found by interpolation since our calculated series does not contain enough terms to ensure good convergence much above $\eta=0\cdot6$.

Table 1. *Position of the first peak in $S(k)$ and $I(k)$ given as the ratio $d/2R$ for a number of values of the packing fraction η and also measured in* nm *based on a value of* $2R=1\cdot18$ nm (*see* § 5)

The value for $\eta=0\cdot9069$ (hexagonal close packing) is $\sqrt{\frac{3}{2}}$. The value for $\eta=0\cdot7854$ (tetragonal packing of hard discs) is found by interpolation.

	$d/2R$ for position of the first peak in $S(k)$	$d/2R$ for position of the first peak in $I(k)$	d in nm for position of first peak in $S(k)$ $2R=1\cdot18$ nm	d in nm for position of first peak in $I(k)$ $2R=1\cdot18$ nm
0·236	1·17	2·5	1·38	2·95
0·392	1·13	1·43	1·33	1·69
0·608	1·05	1·10	1·24	1·30
0·785	~0·9	~0·9	~1·09	~1·09
0·907	0·866	0·866	1·02	1·02

In Fig. 4 a comparison is made of the theoretical $I(k)$ with the diffuse intensity distribution on the equator of the collagen diffraction pattern.

Matching of the patterns was performed by fitting the height and position of the first peak in the intensities and finding a value of η that gave a reasonable fit with the shape of the experimental curve. The best fit yields a value of η of 0·608 and a value of $2R$ of 1·18 nm (the experimental peak is at $d=1\cdot3$ nm and $d/2R$ theoretically at $\eta=0\cdot608$ is 1·1). The theoretical curve seems to give a good account of experiment except near the origin (see below), and the values of the parameters derived from experiment fit in quite well with other observations (also see below).

The differences between theory and experiment are of two sorts. First, there are peaks superposed on the

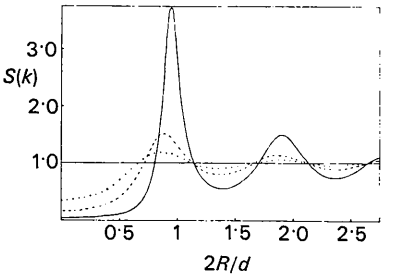

Fig. 1. The structure factor or interference function $S(k)$ calculated for different values of the packing fraction 1 ····· $\eta=0\cdot236$, 2 ------- $\eta=0\cdot392$, 3 ——— $\eta=0\cdot608$.

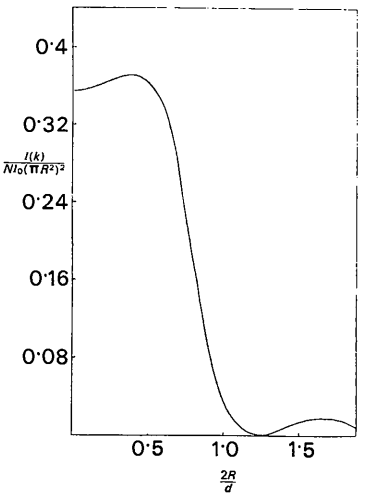

Fig. 2. Total scattering curve $I(k)\propto4(\pi R)^2\ J_1^2(kR)\ S(k)/k^2$ for packing fraction $\eta=0\cdot236$. The first peak in the function is just resolved and is at a value in reciprocal space given by $2\pi R/d=\pi/2\cdot5$, *i.e.* $d/2R=2\cdot5$.

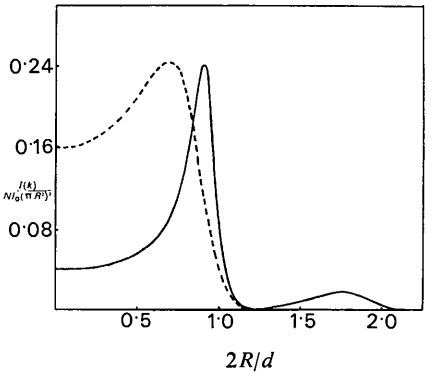

Fig. 3. As Fig. 2, but at higher values of the packing fraction ------- $\eta=0\cdot392$, ——— $\eta=0\cdot608$. The first peak in $I(k)$ is becoming more pronounced as η becomes larger and is moving outwards (towards smaller values of d).

experimental diffuse scatter at 1·26 and 1·75 nm (and perhaps near 2·5 nm). These refer to the ordered phase of collagen (see Woodhead-Galloway, Hukins & Wray, 1975). Second, near the origin, that is as $k \rightarrow 0$ ($d \rightarrow \infty$) the experimental intensity is much greater than that predicted by the model. This is a common feature of observations in simple liquids (see, for example, Mikaloj & Pings, 1967) and is a consequence of the long-range attractive part of the intermolecular potential (Woodhead-Galloway *et al.*, 1968; Machin & Woodhead-Galloway, 1970). An understanding of the feature here would require a more complicated potential-energy function than (2·4) involving an attractive part to describe cohesion between the molecules and it is proposed to leave the reporting of such calculations to a later paper. However, the point has been investigated to some extent. We have attempted to find an experimental $S(k)$ from the observed scattering using expression (2·3). Since a value of $2R=$ 1·18 nm fits some of the scattering curve, we divided the experimental intensity by $[J_1(2\pi R/d)/(2\pi/d)]^2$ using $2R=1{\cdot}18$ nm and the resulting structure factor is given in Fig. 5. This shows the deviation in intensity near the origin from that expected for a hard-disc model, and a comparison of $S(k)$ with those obtained experimentally for the inert gases (Mikaloj & Pings, 1967) shows a qualitative similarity.

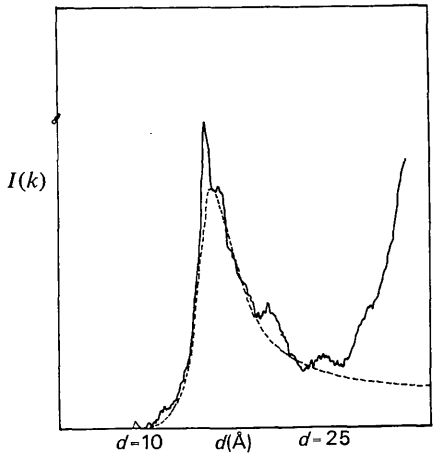

Fig. 4. Fit between equatorial trace of diffuse scatter for wet rat-tail tendon and calculated $I(k)$. The position of the first peak has been fitted to the experimental results, and the best fit between the shapes of the two curves is with $\eta \simeq 0{\cdot}608$. The position of the peak implies that $2R$ (the diameter of the molecular hard core) $\simeq 1{\cdot}18$ nm corresponding to a number density of molecules per square nm of a section perpendicular to the axis of 0·55. As in the case of simple liquids, a hard-disc model gives no account of the scattering near the origin.

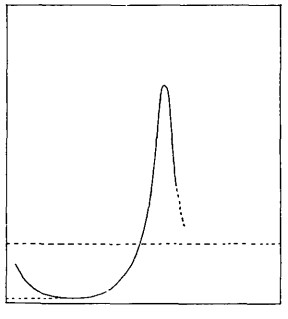

Fig. 5. ——— $S(k)$ experimental for wet rat-tail tendon. The curve was obtained by dividing the experimental $I(k)$ by $(J_1(k)/k)^2$ where $k=2\pi R/d$ and $2R=1{\cdot}18$ nm. The curve near the origin can be seen to be turning upwards – a feature usually observed in structure factors for simple liquids and reflecting some tendency to cluster (due to long-range forces) above that predicted by packing the hard discs -------.

It is worth mentioning that production of experimental structure factors is done routinely for atomic liquids. There is a serious drawback to doing it here. An atomic form factor (see *International Tables for X-ray Crystallography*) is a monotonically decreasing function in reciprocal space, whereas a molecular form factor such as is being proposed here passes periodically through zero making the division of $I(k)$ by $f^2(k)$ impossible in certain regions.

Further confirmation of the appropriateness of the theoretical treatment reported here is obtained by a consideration of the diffuse scattering in dry tendon and that observed in the other well-studied from of collagen, elastoidin.

On drying, the first peak in the scattering form tendon moves out to a value of d of 1·06 nm. Since removal of water results in closer packing, it is interesting that this is the sort of number predicted at high values of the packing fraction in Table 1.

In wet elastoidin the diffuse equatorial scattering is different from that observed in wet tendon. The first peak is not so well resolved (Wray, 1973; Woodhead-Galloway & Knight, 1975) and appears at $d \simeq 2{\cdot}1$ nm [McGavin (1962) suggested that it is not resolved at all]. This is consistent with the calculations reported here if the density in wet elastoidin is much lower than in wet tendon, say, $\eta \simeq 0{\cdot}3$, a point confirmed by the differences in negative staining observed in the electron microscopy of tendon and elastoidin (Knight & Woodhead-Galloway, 1975). No estimate of the density of elastoidin seems to exist so the point cannot be checked, but the large value of d specifying the first peak in the intensity is consistent with the difficulty encountered in its resolution.

Are the molecular diameter and packing fraction obtained on the basis of the analysis of wet tendon consistent with other observations? The value of $2R=$ 1·18 nm is a little lower than that suggested by Woodhead-Galloway, Hukins & Wray to explain the ordered phase in collagen. On the basis of the indexing of the discrete reflexions in the diffraction pattern they argued that the intermolecular spacing must be close to 1·23 nm. The agreement is good. More refined calculations will be required to decide how good the true agreement

is. Using the value of $2R = 1{\cdot}18$ nm and $\eta = 0{\cdot}608$ we may obtain an estimate of the number density of molecules in tendon which is the only useful measure of density for a system having both protein and water. The predicted number density is 0·55 molecules nm^{-2} of cross section. Katz & Li (1973) suggested that the density of collagen in wet tendon is 1·14 molecules gm^{-1} of collagen which may be converted into a number density if the molecular weight of the collagen molecule and the axial arrangement of molecules are known. The molecular weight is about 300 000, and the suggestion of Hodge & Petruska (1963) is widely accepted as a model for the axially projected structure. The calculated experimental number density n_0 turns out to be 0·6 molecules nm^{-2}, again in good agreement with that predicted here. The number density is much higher and the intermolcular distance much lower than suggested by the microfibril model of collagen (Smith, 1968; Miller & Parry, 1973).

Little more can be achieved with the calculations as they stand. To go any further two things are needed: first, some sort of attractive potential function must be introduced to account for the deviation from the hard-disc models (see Woodhead-Galloway, 1968, for an account of the problem in liquids), and second, to cover situations such as that in dry tendon where the packing fraction is much higher than it is in wet, an extended series (or a better method) must be obtained.

J. W. G. gratefully acknowledges the award of the Guinness Fellowship at New College, Oxford and the Sir Henry Royce Fellowship in the Rheumatology Department, University of Manchester, during the tenures of which this research was carried out.

References

BERNAL, J. D. (1964). *Proc. Roy. Soc.* A**280**, 299–322.

BLAISIE, J. K. & WORTHINGTON, C. R. (1969). *J. Mol. Biol.* **39**, 417–439.

BURGE, R. E. (1965). In *Structure and Function of Connective and Skeletal Tissue*, edited by S. FITTON JACKSON, R. D. HARKNESS, S. M. PARTRIDGE & G. R. TRISTAM, pp. 2–7. London: Butterworth.

FRASER, R. D. B., MACRAE, T. P., MILLER, A. & SUZUKI, E. (1964). *J. Mol. Biol.* **9**, 250–252.

HODGE, A. J. & PETRUSKA, J. A. (1963). *Aspects of Protein Structure*. edited by G. N. RAMACHANDRAN, pp. 289–300. New York: Academic Press.

HOSEMAN, R., DREISSIG, W. & NEMETSCHEK, TH. (1974). *J. Mol. Biol.* **83**, 275–280.

KATZ, E. P. & LI, S. T. (1973). *J. Mol. Biol.* **73**, 351–369.

KNIGHT, D. P. & WOODHEAD-GALLOWAY, J. (1975). In preparation.

LEBOWITZ, J. L. & PERCUS, J. (1966). *Phys. Rev.* **144**, 151.

MACHIN, P. A. & WOODHEAD-GALLOWAY, J. (1970). *J. Phys.* C**3**, 2216–2222.

MACHIN, P. A. & WOODHEAD-GALLOWAY, J. (1975). *Mol. Phys.* Submitted.

MCGAVIN, S. (1962). *J. Mol. Biol.* **5**, 275.

MIKALOJ, M. & PINGS, C. J. (1967). *J. Chem. Phys.* **46**, 1401.

MILLER, A. & PARRY, D. A. D. (1973). *J. Mol. Biol.* pp. 437–439.

NEMETSCHEK, TH. & HOSEMAN, R. (1973). *Kolloid Z. Pol.* **251**, 1044.

NEVILLE, A. C. & LUKE, B. M. (1971). *J. Cell. Sci.* **8**, 93–109.

NEVILLE, A. C., PARRY, D. A. D. & WOODHEAD-GALLOWAY, J. (1976). *J. Cell. Sci.* In the press.

ORNSTEIN, L. S. & ZERNIKE, F. (1914). *Proc. Acad. Sci. Amsterdam*, **17**, 793.

PERCUS, J. K. & YEVICK, G. J. (1958). *Phys. Rev.* **110**, 1.

ROWLINSON, J. S. (1964). *Mol. Phys.* **7**, 593–595.

SALPETRE, E. E. (1958). *Ann. Phys. New York*, **5**, 183–223.

SMITH, J. W. (1968). *Nature, Lond.*, **219**, 157–158.

THIELE, E. (1963). *J. Chem. Phys.* **39**, 474.

WERTHEIM, M. S. (1963). *Phys. Rev. Lett.* **10**, 321.

WOODHEAD-GALLOWAY, J. (1968). Ph.D. Thesis. Univ. of Sheffield.

WOODHEAD-GALLOWAY, J., GASKELL, T. & MARCH, N. H. (1968). *J. Phys. C*, **1**, 271.

WOODHEAD-GALLOWAY, J., HUKINS, D. W. L. & WRAY, J. S. (1975). *Biochem. Biophys. Res. Commun.* **64**, 1237–1244.

WOODHEAD-GALLOWAY, J. & KNIGHT, D. P. (1975). *Proc. Roy. Soc.* B submitted.

WRAY, J. S. (1973). D.Phil. Thesis. Univ. of Oxford.