INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
34100 TRIESTE (ITALY) - P.O.B. 586 - MIRAMARE - STRADA COSTIERA 11 - TELEPHONES: 224281/2/3/4/5/6
CABLE: CENTRATOM - TELEX 460392 - I

SMR/111 - 6

SECOND SUMMER COLLEGE IN BIOPHYSICS

30 July - 7 September 1984

The X-ray Fiber Diffraction Studies of DNA and RNA

K. TOMITA
Faculty of Pharmaceutical Sciences
Osaka University
Yamadaoka 1 - 6
Suita
Osaka 565
Japan

The X-ray fiber diffraction studies of DNA and RNA

Three lectures by Prof. Ken-ichi Tomita (Osaka, Japan)

Lecture 1.

1) INTRODUCTION

I will introduce briefly the historical backgrounds before and on the era of discovery of double-stranded DNA helix by Watson & Crick. Accumulations of biochemical and physico-chemical studies[1-4] on DNA structure led to the Watson-Crick helical model[5] which was confirmed by new X-ray diffraction data obtained in Wilkin's laboratory. "The Double Helix" written by J.D.Watson[6] is a very fresh and impressive story on the discovery of the DNA structure.

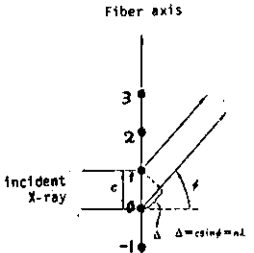
1) W.T.Astbury(1947).Symp.Soc.Exp.Biol.I.Nucleic Acid, p.66.
2) S.Furberg(1950).Acta Cryst.,3,325.
3) J.M.Gulland(1947).Cold Spr.Harb.Symp.Quant.Biol.,XII,p.95.
4) S.Zamenhof, G.Brawenmann & E.Chargaff(1952).Biochim.Biophys. Acta.,9,402.
5) J.D.Watson & F.H.C.Crick(1953).Nature,171,964.
6) J.D.Watson(1968). "The Double Helix", Atheneum, New York.

2) FUNDAMENTALS OF X-RAY FIBER DIFFRACTION THEORY

(a) X-RAY SCATTERING FROM FIBERS.

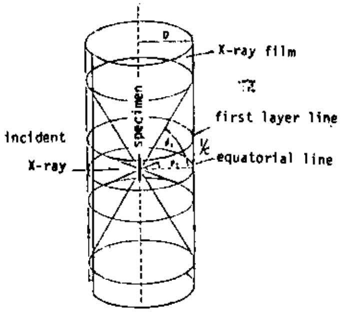
Fibrous materials such as polypeptides and polynucleotides cannot be obtained as single crystals. When such molecules occur in ordered states a high degree of orientation normally occurs only in the direction of the fiber axis. The lateral orientation of the crystallites is usually imperfect.

As an example of the simplest fibrous molecule a one-dimensional lattice consisting of identical atoms equally spaced at intervals of c along the z-axis(fiber axis) is considered.



Diffraction by one-dimensional lattice

Fig.1



Diffraction photograph

Fig.2

The numbering of the atoms is given in Fig.1. The coordinate of the nth atom is $x_n=y_n=0$, $z_n=nc$. Therefore, the molecular scattering factor, F_M, is given by

$$F_M = \sum_n f_n \exp[2\pi i(Xx_n+Yy_n+Zz_n)] \\ = f\sum_{n=-N}^{N}\exp(2\pi iZnc) \quad . \qquad (1)$$

Since this equation is a series of geometrical progressions and $\exp(i\alpha) - \exp(-i\alpha) = 2i\sin\alpha$, Eq.(1) becomes

$$F_M = f\,\frac{\sin[\pi Z(2N+1)c]}{\sin(\pi Zc)} \quad . \qquad (2)$$

Now for simplicity, 2N+1 is replaced by N. The square of the sine part of Eq.(2) is

$$G\cdot G^* = \frac{\sin^2(\pi NZc)}{\sin^2(\pi Zc)} \quad (\text{ Laue function }) \qquad (3)$$

If N is large, $G\cdot G^*$ has maximum value, N^2, on the planes of $Z=\ell/c$ (ℓ is an integer) and nearly zero elsewhere.

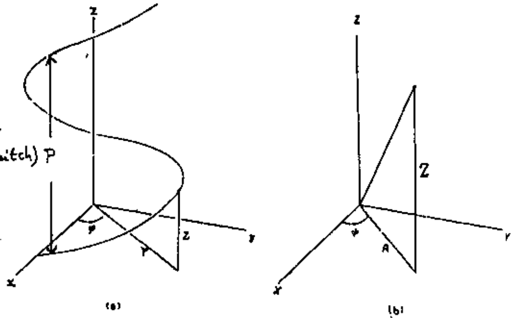
Accordingly, Eq.(1), shown in reciprocal space(X-ray film) in Fig.2, has a value Nf only on the planes normal to the Z-axis with equal separations 1/c and whose intersections with the sphere of reflection are coaxial circles. This result shows that when there is a periodicity only along the z direction(fiber axis) in real space, a periodicity exists only along the Z direction in reciprocal space, that is, Fourier transform of one-dimensional lattice is the parallel layer lines with the spacing of 1/c in reciprocal space(X-ray film).

The fiber repeat distance is readily obtainable directly, if strong layer line spacings are present, from the relationship, $c\sin\phi = n\lambda$, where c is the fiber repeat distance, λ is X-ray wave length and n is an integer. From the spacing of equatorial reflections the distance of reflecting planes parallel to the fiber axis may often obtained.

(b) X-RAY SCATTERING FROM A HELIX.

Fig.3

Cylindrical coordinates



(a) real space (b) reciprocal space

$$\begin{cases} x = r\cos 2\pi \frac{z}{P} \\ y = r\sin 2\pi \frac{z}{P} \\ z = z \end{cases} \qquad \begin{cases} X = R\cos\psi \\ Y = R\sin\psi \\ Z = Z \end{cases} \qquad (4)$$

It is often convenient to describe the structure of a helical molecule in cylindrical coordinates instead of Cartesian ones as shown in Fig.3 and Eq.(4). The Fourier transform F (or the intensity) will also be described by cylindrical coordinates. Such coordinates are most suitable if the object has some circular symmetry, i.e., if the principal axis is a rotation or screw axis.

Cochran et al.[7)] have discussed diffraction by helical molecules.

Consider first the Fourier transform for a continuous helix of radius r and pitch(repeat distance) P (Fig.3). $\varphi = 2\pi z/P$, and the equation of the helix in cylindrical coordinates is

$$r = r, \text{ and } \varphi = 2\pi z/P . \qquad (5)$$

We assume that the density along the helix is unity. The Fourier transform at a point (X,Y,Z) in reciprocal space is given by

$$T(X,Y,Z) = \int \exp[2\pi i(xX + yY + zZ)]dV \qquad (6)$$

when dV = dxdydz is a volume element of the helix and is directly proportional to dz. Thus,

$$T(X,Y,Z) = \int_0^P \exp[2\pi i(rX\cos 2\pi z/P + rY\sin 2\pi z/P + zZ)]dz \qquad (7)$$

apart from unimportant constant of proportionality ($2\pi r_0$). Using Eq.(4), this result can be written as

$$T(R,\psi,Z) = \int_0^P \exp[2\pi i\{Rr\cos(2\pi z/P - \psi) + zZ\}]dz \qquad (8)$$

where $R^2 = X^2 + Y^2$, and $\tan\psi = Y/X$.

The right-hand side of Eq.(8) is equal to zero except when Z = n/P where n is an integer. This is in accord with the fact that the diffraction pattern of a helix with an exact vertical repeat distance P is confined to layer lines at heights Z = n/P in reciprocal space. Hence,

$$T(R,\psi, n/P) = \int_0^P \exp[2\pi i\{Rr\cos(2\pi z/P - \psi) + nz/P\}]dz \qquad (9).$$

The above integral may be evaluated by means of the identity:

$$\exp(iX\cos\varphi)\exp(in\varphi)d\varphi = 2\pi i^n J_n(X)$$
$$= 2\pi J_n(X)\exp in\pi/2 \qquad (10)$$

where $X = 2\pi Rr$ and $\varphi = 2\pi z/P - \psi$, and J_n is an n-th order Bessel function (Fig.4).

From Eqs.(9) and (10) we have

$$T(R,\psi,n/P) = J_n(2\pi Rr)\exp[in(\psi + \pi/2)] \qquad (11).$$

The function T has several notable features;

1)the function T gives directly the amplitude and phase of the X-ray scattering on the n-th layer line, that is, the function T is structure factor F of a continuous helix.

Bessel functions.

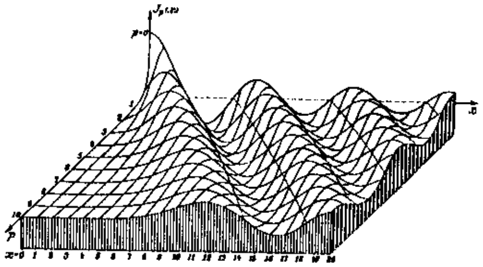


Fig.4

2) the modulus $|T| = |J_n(2\ Rr)|$ is independent of φ and hence possesses cylindrical symmetry: the distribution of $|T|^2 = |F|^2$ on the layer line with $\ell = n$ is defined by the square of J_n.

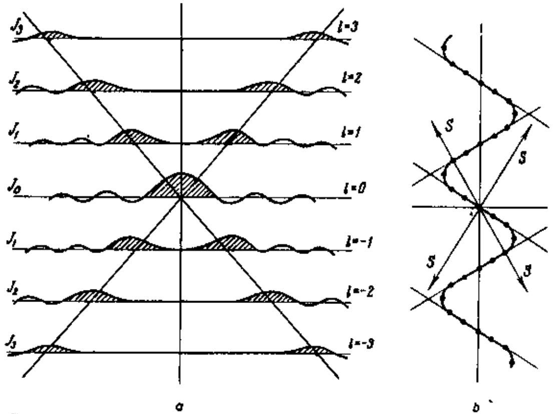


Fig.5. *a*) Distribution of the Bessel functions with $n = l$ which determine the intensities along the layer lines; *b*) scheme illustrating the correlation between the array of main peaks in the oblique form of a cross and the normals to the densest rows of atoms in the helix

Since X(the radius of the first peak in J_n) increases with n (Fig.4), the intensity distribution has the characteristic form of a cross(Fig.5 a). This form is well explained by the array of the most densely populated "rows" of atoms in the helix(Fig.5 b), normal to which the intensity is highest in reciprocal space (the arrow ; S in Fig.5 b).

The continuous wire model is, of course, somewhat unrealistic. If the helix is broken up into a series of points , the central part of the pattern is unchanges but further patterns are superimposed upon it and the resultant pattern has empty diamond-shaped regions above and below the center.

We define a discontinuous helix as a set of points occurring with a vertical spacing p on a continuous helix.

Consider a function H which is zero everywhere except on a continuous helix, where it assumes the value unity, that is,

$$H(x,y,z) = \delta(x - r\cos 2\pi z/P)\,\delta(y - r\sin 2\pi z/P) \qquad (12).$$

The Fourier transform of the function H is given by Eq.(11). Now, we consider another function K which is zero everywhere

except on a set of horizontal planes of spacing p, where it assumes the value unity, that is,

$$K(x,y,z) = \sum_{k=-\infty}^{\infty} \delta(z - kp) \qquad k=0,\pm1,\pm2,\cdots \qquad (13).$$

The Fourier transform of the product H·K of these two functions gives a structure factor of a discontinuous helix, $F_d(X,Y,Z)$.

$$F_d(X,Y,Z) = \iiint_{-\infty}^{\infty} H(x,y,z)K(x,y,z)\exp[2\pi i(xX + yY + zZ)]\,dxdydz \qquad (14)$$

According to the convolution theory, the Fourier transform of the product H·K is given by the convolution of T and S which are the Fourier transform of H and K, respectively, that is,

$$F_d(X,Y,Z) = \iiint_{-\infty}^{\infty} S(\xi,\eta,\zeta)\,T(X-\xi,Y-\eta,Z-\zeta)\,d\xi d\eta d\zeta \qquad (15).$$

Using Eq.(13), Eq.(15) can be written as

$$F_d(R,\psi,Z) = \sum_{m=-\infty}^{\infty}\sum_{n=-\infty}^{\infty} \delta(Z-n/P-m/p)\,J_n(2\pi Rr)\exp[in(\psi+\pi/2)] \qquad (16).$$

Therefore, $F_d(R,\psi,Z)$ is to be finite only in the planes at height

$$Z = n/P + m/p \qquad n,m= 0,\ \pm1,\ \pm2\ \cdots \qquad (17)$$

on which it assumes the value $J_n(2\pi Rr)\exp[in(\psi+\pi/2)]$.

Eq.(17) is a selection rule, and if P/p can be expressed as a ratio of whole numbers, the transform is confined to a set of planes, that is, if $P/p = M/N_0$, $Z = \ell/c$ and $N_0P = c$, Eqs.(16) and (17) are

$$F_d(R,\psi,\ell/c) = \sum_n J_n(2\pi Rr)\exp[in(\psi+\pi/2)] \qquad (18)$$

$$nN_0 + mM = \ell \qquad (19)$$
$$n,m,\ell = 0,\ \pm1,\ +2\ \cdots\cdots$$

(c) THE STRUCTURE FACTOR CALCULATION.

The calculated intensity of diffraction by a proposed fiber structure can be obtained either by computation, using the structure factor expression (Eq.(18)) or it can be obtained optically with the optical diffraction spectrometer[8]. The latter method provides only approximate structure factors for comparison with the observed X-ray diffraction photograph, because of the difficulty of dealing with overlapping atoms and with the differences in scattering factors.

We shall now consider how numerical calculation can be made when exact coordinates are assumed for all the atoms in a helical structure. A real helical molecule is a system of atoms whose centers may have different r; such a molecule can be described as a combination of several discontinuous helices differing in their r_j. The positions of the atoms in a unit group may be specified in terms of their initial coordinates (r_j, φ_j, z_j); in addition, we have the atomic scattering amplitude f_j for each.

Thus, in place of Eq.(18) we have the general relation

$$F(R,\psi,\ell/c) = \sum_j \sum_n f_j J_n(2\pi Rr) \exp[i\{n(\psi-\varphi_j+\pi/2)+2\pi\ell z_j/c\ ]$$
$$= \sum_n G_{n,\ell}(R) \exp[in(\psi+\pi/2)] \qquad (19)$$

where

$$G_{n,\ell}(R) = \sum_j f_j J_n(2\pi Rr) \exp[i(-n\varphi_j + 2\pi\ell z_j/c\ )] \qquad (20)$$

but with the selection rule

$$\ell/c = n/P + m/p.$$

Eq.(19), being a function of ψ, is not cylindrically symmetrical. In a fiber diagram, where the molecules are parallel, but oriented at random about the fiber axis, the relevant intensity is cylindrically averaged: that is, averaged over all values of ψ 9).

$$\langle F^2(R,\psi,\ell/c)\rangle_\psi = \langle F\cdot F^*\rangle_\psi = \frac{1}{2\pi}\int_0^{2\pi} F\cdot F^*\, d\psi$$
$$= \frac{1}{2\pi}\sum_n\sum_{n'} G_{n,\ell}(R)\, G^*_{n,\ell}(R)\int_0^{2\pi}\exp[i(n-n')(\psi+\pi/2)]\,d\psi$$
$$(21)$$

Since n and n' are integers and F is a periodic function of the period 2π with respect to ψ,

$$\int_0^{2\pi}\exp[i(n-n')\psi]\,d\psi = \begin{cases} 0 & n \neq n' \\ 2\pi & n = n' \end{cases} \qquad (22)$$

Therefore,

$$\langle F^2(R,\psi,\ell/c)\rangle_\psi = \sum G_{n,\ell}(R)\cdot G^*_{n,\ell}(R) = \sum_n (\ A_n^2 + B_n^2\ ) \qquad (23)$$

where

$$A_n = \sum_j f_j J_n(2\pi Rr)\cos(n\varphi_j + 2\pi\ell z_j/c)$$
$$B_n = \sum_j f_j J_n(2\pi Rr)\sin(n\varphi_j + 2\pi\ell z_j/c) \qquad (24).$$

References

7) W.Cochran, F.H.C.Crick & V.Vand(1952).Acta Cryst.,5,581.

8) C.A.Taylor & H.Lipson(1960). "Optical Transforms",Bell, London.

9) D.R.Davies & A.Rich(1959).Acta Cryst.,12,97.

Lecture 2.

3) EXPERIMENTAL TECHNIQUES FOR X-RAY FIBER DIFFRACTION

I will explain in some detail by using slides, and only a brief outline of the contents is given here.

(a) X-RAY SOURCES

For fiber diffraction studies, Cu Kα radiation(wavelength 1.542 Å) is usually employed as X-ray source. The Kβ line (λ=1.392 Å) is removed (or more accurately the intensity is reduced) by a nickel filter. Monochromatic beams are also obtained by diffraction through particular single crystal (crystal monochromator). Large structures diffract relatively weakly, and a fundamental practical problem is the generation of a sufficiently intense X-ray beam.

Commercially available X-ray generators fall into three classes: (i) sealed off tubes (ii) microfocus tubes and (iii) rotating anode tubes.

(b) X-RAY CAMERAS

Flat cameras are used for taking the X-ray photographs of the oriented samples. However, the flat camera is not suitable for measuring the fiber period(helical pitch:P, axial raise per residue:p etc.), since the layer lines appear as hyperbolas instead of straight lines as in those photographs obtained with cylindrical camera. Very clear photographs can be obtained by using a vacuum camera with a large distance between film and specimen. Helium is bubbled through appropriate saturated salt solutions and then through the cameras to control the relative humidity of the sample (calcium chloride,33% r.h., sodium nitrite,66% r.h., sodium chlorate, 75% r.h.,and sodium tartrate, 92% r.h.).

(c) DEVICES FOR MAKING THE ORIENTED FIBERS

In mounting a fiber, it is very often necessary to establish or to improve the degree of order. The method used depends on the physical properties of tne material being studied. In the case of polynucleotide, the method of pulling out into thin threads is commonly used. Langridge et al.[10] used fibers 50 to 100μ in diameter, made by wetting with a droplet of glass-distilled water a few mg of DNA on a glass slide, dipping a pointed glass rod into the sticky and very stiff mass, and slowly withdrawing it. Well oriented samples show good uniform extinction under the polarizing microscope.

(d) INTENSITY MEASUREMENT AND HELICAL PARAMETERS

In order to measure the diffraction spots over a wide range of optical densities, it is usual to use a multiple film technique, where a stack of two or more films absorbing about 75% of incident energy and transmitting 25% to the next film. Measurements of the reflection intensity are made using a microphotometer and by visual comparison with a standard intensity scale. From the distances between the equator and the layer lines of X-ray fiber diagram, the fiber period (c-value) is obtained using Polanyi's formula, $c\sin\phi = n\lambda$ (see Fig.1). Indexing of reflections and space group determination are possible when very well oriented fiber diagram is obtained. If the X-ray fiber diagram is as simple as those shown in Fig.5, a measurement of the position of the first maximum of each Bessel function on each layer line would give the helix radius. For example, if we know that there is a J_4 term on the 4th layer line having a maximum value at R_{max}, we can refer to the tables to find that the first maximum of $J_4(x)$ is at x = 5.3 and put $2\pi r_0 R_{max} = 5.3$, thereby obtaining r_0.

(e) MODEL BUILDINGS AND REFINEMENTS

Model molecules made of metal or plastics are very useful for setting up molecular models of polynucleotides. We can use the model molecules (usually Kendrew's wire model) made on a scale, for example, of 20 mm or 50 mm corresponding to 1 Å. The molecular model is mounted with the helix axis perpendicular to a drawing board and coordinates are measured directly with ruler or using the shadow of the model in a beam of parallel light. Fourier refinement technique has been applied to DNA, RNA and synthetic polynucleotides with considerable success.

Reference

10) R.Langridge, H.R.Wilson, C.W.Hooper, M.H.F.Wilkins & L.D.Hamilton(1960). J.Mol.Biol.,2,19.

Lecture 3.

4) APPLICATION TO POLYNUCLEOTIDES

Using slides, I will explain the details of the following polynucleotide structures;

a) SYNTHETIC POLYNUCLEOTIDES

(i) poly(I)·poly(C^+)[11]

(ii) poly(I_f)·poly(C)[12]

(iii) poly(m^2A)·poly(U)[13]

b) DNA

(i) A-DNA[14] (ii) B-DNA[15-16] (iii) C-DNA[17]

(iv) D-DNA[18] (v) Z-DNA[19]

c) Double-stranded RNA[20]

(i) reovirus RNA[21] (ii) wound tumor virus RNA[22]

(iii) rice dwarf virus RNA[23]

d) DNA-RNA hybrid[24]

References.

11) K.Tomita & A.Rich(1964). unpublished result.

12) K.Tomita, T.Hakoshima, T.Fukui & M.Ikehara(1982). Nucleic Acids Res., Symp.Series,11,177.

13) T.Hakoshima & K.Tomita(1980). in "Biomolecular structure, conformation, function and evolution" Vol.I.p.313, Pergamon Press, Oxford & New York.
T.Hakoshima, T.Fukui, M.Ikehara & K.Tomita(1981). Proc.Nat. Acad.Sci.USA.,78,7309.

14) S.Arnott, D.W.L.Hukins(1972).Biochem.Biophys.Res.Comm., 47,1504.

15) R.Langridge, D.A.Marvin, W.E.Seeds, H.R.Wilson, C.W. Hooper, M.H.F.Wilkins & L.D.Hamilton(1960). J.Mol. Biol.,2,38.

16) A.G.W.Leslie, S.Arnott, R.Chandrasekaran, R.L.Ratliff (1980). J.Mol.Biol.,143,49.

17) D.A.Marvin, M.Spencer, MH.F.Wilkins & L.D.Hamilton(1961). J.Mol.Biol.,3,547.

18) S.Arnott, R.Chandrasekaran, D.W.L.Hukins, P.J.C.Smith & L.Watts(1974). J.Mol.Biol.,88,523.

19) S.Arnott, R.Chandrasekaran, D.L.Birdsall, A.G.Leslie & R.L.Ratliff(1980). Nature,283,743.

20) S.Arnott, F.Hutchinson, M.Spencer, M.H.F.Wilkins, W. Fuller & R.Langridge(1966). Nature,211,227.

21) R.Langridge & P.J.Gomatos(1963). Science,141,694.

22) K.Tomita & A.Rich(1964). Nature,201,1160.

23) T.Sato, Y.Kyo-oku, S.Higuchi, Y.Mitsui, Y.Iitaka & M.Tsuboi(1966). J.Mol.Biol.,16,180.

24) G.Milman, R.Langridge & M.J.Chamberlin(1967). Proc.Nat. Acad.Sci.USA.,57,1804.