CHAPTER VIII

X-Ray Diffraction Studies on the Lens esculenta Protein Fractions

Theoretical:

1. The Problem:

The structural aspects of macromolecules, like proteins, have three distinctive features -

(i) Precise knowledge of bond lengths and bond angles
(ii) Molecular configuration
(iii) Overall size and shape of the molecule.

Determination of these structural features is of definite help in understanding and interpreting the various physico-chemical properties of the molecule. The experiment and discussion in this chapter mainly concentrates in deriving general picture of the macromolecular structure by X-Ray scattering in the lower range of interatomic distances, say around 10 Å.

Here an attempt has been made to derive, with a good degree of explicitness, the molecular intensity function as an average value overall orientations. Thus the problem of setting up an ideal diffraction experiment did not arise. Debye (1915, 1930)\(^1\), Pauling and Brockway (1934)\(^3\), Warren and Gingrich (1934)\(^4\) were among the pioneer workers who have investigated the structure of some relatively simple vapors and simple amorphous solids.

Since the effects of intermolecular interference are relatively non-confusing with large molecules, and are confined
Debye (1930) expressed the average intensity per molecule scattered coherently by an assembly of independent molecules in random orientation as:

\[ I(s) = \sum_i \sum_j f_i(s) f_j(s) \frac{\sin S \ell_{ij}}{S \ell_{ij}} \]  

(1)

In our discussion, the angle variable has been defined as,

\[ S = 4 \pi \frac{\sin \theta}{\lambda} \]  

(2)

where, \(2 \theta = \) angle of scattering

\(\lambda = \) wavelength of the monochromatic X-Ray employed

\(\ell_{ij} = \) Distance between two atoms, \(i\) and \(j\), constituting a pair

and \(f_i(s)f_j(s) = \) product of the scattering factors of \(i\) and \(j\).

Separating out the portion of the scattered intensity that is due to interatomic interferences,

\[ I(s) = A(s) + B(s) \]  

(3)

when, \(i = j\)

\[ A(s) = \sum_i f_i^2(s) \]  

(4)

when, \(i \neq j\)

\[ B(s) = \sum_i \sum_{j \neq i} f_i(s)f_j(s) \frac{\sin S \ell_{ij}}{S \ell_{ij}} \]  

(5)

Assuming that all the atomic scattering factors have the same dependence upon the angle variable \(s\), and differ only in magnitude, and defined by the atomic number \(Z\),

\[ f_Z(s) = z f(s) \]  

(6)

It is now possible to bring out \(f^2(s)\) outside the summation in Eqn. (5). Thus,
\[ i(s) = \text{Intensity function} = \frac{I(s)}{A(S)} = \]

**Interference Scattering**

Independent scattering

\[ \text{Hence, } \quad i'(s) = \frac{1}{\gamma} \sum_i \sum_j \frac{\delta_i^j}{S_l^{ij}} W_{ij} \frac{S_i L_{ij}}{S_{ij}} \]  

\[ \text{Where, } \quad \gamma = \sum \mu^2 \]

\[ W_{ij} = Z_i Z_j = \text{product of the atomic numbers in the pair of atoms } i,j. \]

Here, \( i(s) \) has been termed in the "Interference Intensity Function".

Since Eqn. (8) is rather complicated for routine calculation, in the case of a large molecule like a protein molecule, the Eqn. (8) may be rewritten as the equivalent Fourier integral,

\[ \psi S i'(s) = \int_0^\infty \frac{\phi(r)}{r} \delta_i^j S r \, dr \]  

in which the smoothed continuous radial distribution function \( \phi(r) \) replaces the discontinuous point function \( W_{ij}(l_{ij}) \) in Eqn. (8)

The radial distribution function is defined as:

\[ \phi(r) = \int_0^\infty \nu \sigma(l) \, dl \]  

where, \( \sigma(x) \) is a smoothing function folded into the discontinuous histogram \( H(l) \) representative of the distribution of \( W_{ij} \) in terms of \( l_{ij} \).
In the Eqn. (9), \( i'(s) \) may be replaced by the Fourier transform, \( \psi'(s) \), or \( \psi(x) \):

\[
i(s) = \frac{i'(s)}{\psi'(s)} \quad - (12)
\]

The exact form of the smoothing function is therefore merely a matter of convenience. The deviation in \( i(s) \) between Eqns. (8) and (12) can be made negligible by choosing \( \Delta = 0.1 \AA \).

4. The Use of Radial Distribution Curves:

The radial distribution function \( \phi(r) \) can be calculated from the atomic positions, as evidenced by the Eqns. (10) and (11), and is characteristic of the molecule, whether large or small. Radial distribution function may be derived from experimental intensity curve as follows:

Fourier transformation of Eqn. (9) gives,

\[
\phi(r) = \frac{2\lambda}{\pi} \int_0^{\infty} S_i'(s) \psi_0 \psi_s \, ds \quad - (13)
\]

As \( i'(S) \) may be easily obtained from experimentally observed values of \( i(S) \) by the Eqn. (12), it is possible to derive by experiment a radial distribution function with the same significance as \( \phi(r) \) computed for a hypothetical molecule.

Experimental:

1. Experimental Procedure: Sample Preparation:

(i) X-Ray Source: The general practice in the X-Ray study of proteins is to use on Cu - K\( \lambda \) radiation (\( \lambda = 1.542 \AA \)). The usual Nickel filter (K\( \beta \) filter)
helps to achieve monochromatization ca. 99% in terms of energy (Arndt and Riley, 1952). The voltage and current of X-Ray machine were 40 KV and 10 mA respectively.

(ii) The Sample: The proteins to be studied were taken as amorphous powder, finally ground in a marble stone mortar and pestle and packed in the sample holder. Packing in the sample holder has to be well compressed so that it is almost completely homogeneous in density.

(iii) The Camera and the Sample Holder: In order to measure, in a single continuous experiment, the intensity scattered over a certain angular range, the usual Debye-Scherrer method poses some difficulties. In the Debye-Scherrer method a cylindrical specimen is employed, and the diffraction pattern is recorded in a photographic film. The measured intensity ordinates must be corrected for absorption in the sample; although methods for determining this correction have been developed, an accurate knowledge of the atomic composition of the specimen is needed, since the absorption factor cannot be measured directly but must be calculated. In case of polyelectrolytes, containing mainly light atoms, the absorption is often determined by small amounts of heavier atoms whose exact proportion is not known with certainty.

Because of this difficulty it is customary to make accurate intensity measurements using parallel sided powder blocks and employing a side reflection technique. The camera used for this purpose is known as "Brindlay Camera". If the specimen is sufficiently thick to be considered "infinite" from the standpoint of absorption, the correction to be applied becomes invariant with the scattering angle.
In these experiments with Brindleys camera, the radius of the camera was 10 cm. The scattering angle was chosen to be 5°, and an exposure was given continuously for two hours.

The sample holder was made from lead sheet, with a rectangular slot in the middle. The dimensions of the slot were 12 mm. x 5 mm. and the thickness was 2 mm. The back of the slot was covered by fixing a thin cellophane sheet. The powdered sample was packed into the slot with proper compression.

(iv) Recording: In our experiments, the photographic recording was preferred to counters. In all types of counters, a certain degree of background count is always present. While taking the measurements, the error due to background count has to be corrected for. With photographic recording, this trouble can be avoided. We have used Ilford Ilfex Industrial X-Ray films for this recording.

(v) Microphotometry: Intensity curves were drawn from the scattering photographed with the help of Microphotometric recording of the film strips in the low and medium angle region. Radial distribution function were calculated from the intensity curve and the radial distribution curves were drawn.

Results and Discussions:

(I) Sensitivity and Reliability of the Method:

The intensity curves obtained in this way refer to the sample only and have been corrected in the usual way for experimental factors. These include the incoherent Compton Scattering, but the effect of fluorescent radiation is usually negligible.
The degree of reproducibility of results for a given sample, when in the form of fine powder, was very high, even when conditions were deliberately varied in order to constitute a test. The accuracy of measurement of scattered intensity, over a range of S say from 0.1 to 6.5, was within 1% in the most favourable cases, although for general study, a somewhat lower accuracy was accepted.

In the symmetrical transmission setting, almost grazing angles of incidence (about 5°) was necessitated. Therefore the geometrical volume correction factor has to be very efficiently noted. If the specimen deviates from the assumed shape of a block with perfectly plane and parallel sides, the correction factor will not operate satisfactorily. In particular, any bowing of the specimen into a concave or convex shape will seriously affect the extreme high angle point of the intensity curve, and a satisfactory correction procedure will be impossible. However, the low angle region, which is of interest in these studies, is not greatly affected as the change of irradiated volume is there much less marked.

After summing up, the measurement of corrected intensity curve is sufficiently accurate for the low and medium angle region.
Microphotometric Figure of *L. leucothrix* BR 25
Protein Fractions after X-Ray Diffraction Expt.

Sample LEW-IP 6.45

Sample LEW-IP 6.2

Sample LEW-IP 5.5
INTERFERENCE INTENSITY CURVE FOR SAMPLE LEW - I.R. 6'45

\((s), I\)
Interpretation for the X-Ray scattering data for amorphous substances is difficult due to the absence of complete order. The calculation of the radial distribution function can give all the thermodynamic properties correctly. The structure studies from the radial distribution function is a rare one, but considerable research work is going on along this line. Riley and Arndt³ have carried out the calculations for the following helical models:

\[ \alpha \quad \beta \quad \gamma \quad 4_{13} \quad 3_{10} \quad 3_{8} \quad 2_{76} \]

(1) Curves 1-3 are the results of the microphotometric analysis of the X-Ray Diffraction photographs of three *Lens esculenta* protein samples.

(2) Curves 4-6 are the $S$ versus $i'(s)$ curves, where

\[ S = \frac{4 \pi}{\lambda} \int s \sin \theta \]

(3) Curves 7-9 are the radial distribution functions $\Phi(r)$ versus $r$. The radial distribution functions were calculated by means of the equation No. (13). Table 1 tabulates the position of the peaks and peak heights of the radial distribution curves for different protein samples.

Table is given in the next page.
Table - 1

Comparison of Radial Distribution Curves for the three Protein Samples

<table>
<thead>
<tr>
<th>Sample No.</th>
<th>Distance of peaks (Å)</th>
<th>Peak Heights</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEW-IP 6.45</td>
<td>2.25</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>4.00</td>
<td>1.10</td>
</tr>
<tr>
<td></td>
<td>6.25</td>
<td>0.20</td>
</tr>
<tr>
<td></td>
<td>7.80</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>8.70</td>
<td>0.75</td>
</tr>
<tr>
<td></td>
<td>9.85</td>
<td>0.45</td>
</tr>
<tr>
<td>LEW-IP 6.2</td>
<td>1.15</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>4.05</td>
<td>2.50</td>
</tr>
<tr>
<td></td>
<td>7.05</td>
<td>2.05</td>
</tr>
<tr>
<td></td>
<td>8.20</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td>9.50</td>
<td>0.85</td>
</tr>
<tr>
<td>LEW-IP 5.5</td>
<td>2.00</td>
<td>1.05</td>
</tr>
<tr>
<td></td>
<td>4.40</td>
<td>2.20</td>
</tr>
<tr>
<td></td>
<td>5.65</td>
<td>0.50</td>
</tr>
<tr>
<td></td>
<td>6.40</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>7.00</td>
<td>0.70</td>
</tr>
<tr>
<td></td>
<td>8.30</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>9.40</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Table 2 shows the diameter of the helix chain.

Table 2
Calculation of the Diameter of Helix Chain from the Radial Distribution Curve

<table>
<thead>
<tr>
<th>Sample No.</th>
<th>Diameter of Helix Chain, (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEW-IP 6.45</td>
<td>4.00</td>
</tr>
<tr>
<td>LEW-IP 6.2</td>
<td>4.05</td>
</tr>
<tr>
<td>LEW-IP 5.5</td>
<td>4.40</td>
</tr>
</tbody>
</table>

Analysis of Results:

The basic structure considered for a protein is the α-helix of Pauling, Corey and Branson (1951). A protein is made up of 20 odd amino acid residue in a particular sequence, specific for specific protein. A typical single amino acid residue with the stereochemistry around the asymmetric Cα-atom is the following:

\[
\begin{align*}
\text{C}^\prime & \quad \text{O} \\
\cdots \text{N} & \quad \text{C}_\alpha \\
\cdots \text{H} & \\
\text{C}_\beta & \\
\text{R} & 
\end{align*}
\]
Fig 10 Radiol distribution curves \( \phi(r) \) derived various helical configurations of a polypeptide chain. The number of amino acid residues included in the calculations is shown in parentheses.
In this the amino acid residue is in the L-configuration usually found in natural proteins drawn according to the usual conventions. $C_\alpha$ is the plane of the paper; the bonds of $\text{C}^\prime\text{O}$ and $C_\beta$ are below and those of $\text{NH}$ and $H$ above this plane.

There are possible different variables of helix, or indeed of any molecule containing optically asymmetric atoms as an integral part of a helical chain.

In order to interpret the experimental data we ought to know behaviours of various experimentally determinable functions for standard configurations. We have appended the following figures of model polypeptides from literature for immediate use.

The curve 10 corresponds to the hypothetical model polypeptides drawn from the calculations of Riley (1960). While comparing our experimental $\phi(r)$ versus $r$ get the following:

(i) The general trend of the curves depend on the length of the helix.

(ii) An infinitely long chain gives rise to a radial distribution function which at large values of $r$ oscillates asymptotically about a constant finite value.

(iii) Chain of finite length gives after initial region, a curve which descends steadily until it reaches zero at a value of $r$ approximately equal to the folded chain length and diameter.
In this the amino acid residue is in the L-configuration usually found in natural proteins drawn according to the usual conventions. $C_\alpha$ is the plane of the paper; the bonds of $C^\prime 0$ and $C_\beta$ are below and those of NH and H above this plane.

There are possible different variables of helix, or indeed of any molecule containing optically asymmetric atoms as an integral part of a helical chain.

In order to interpret the experimental data we ought to know behaviours of various experimentally determinable functions for standard configurations. We have appended the following figures of model polypeptides from literature for immediate use.

The curve 10 corresponds to the hypothetical model polypeptides drawn from the calculations of Riley (1960). While comparing our experimental $\Phi(r)$ versus $r$ get the following:

(i) The general trend of the curves depend on the length of the helix.

(ii) An infinitely long chain gives rise to a radial distribution function which at large values of $r$ oscillates asymptotically about a constant finite value.

(iii) Chain of finite length gives after initial region, a curve which descends steadily until it reaches zero at a value of $r$ approximately equal to the folded chain length and diameter.
(iv) An idea of the shape function $\phi (r)$ may be obtained depending on the form of the molecule, which is clearly demonstrated by these model $\phi (r)$ curves.

(v) Initial part of all the curves are virtually the same, corresponding to the nearest and next nearest atom neighbours (Peaks at 1.5 Å and 2.6 Å).

(vi) The region most sensitive to the chain configuration lies between about $r = 4\,\text{Å}$ and $r = 10\,\text{Å}$.

(vii) A prominent peak between 4 and 5 Å occurs in most cases and is related to the average diameter of the helix.

(viii) There exists a general similarity between the radial distribution curves for the two variants of the $\alpha$, $\Pi$ and $\frac{4}{13}$ helices. This makes their differentiation, by means of radial distribution functions, a matter of uncertainty.
REFERENCES

1. Debye, P.

2. Debye, P.
   Physik. Z. (1930), 21, 419.

3. Pauling, L. and Brockway, L.O.

4. Gingrich, N.S. and Warman, E.B.

5. Riley, D.P. and Arndt, U.W.
   Nature, (1953), 171, 144.

6. Riley, D.P. and Arndt, U.W.

7. Pauling, L., Corey, R.B. and Branson, H.R.

8. Riley, D.P.
   In "Non-crystalline Solids"
DISCUSSION

Summary of the Main Results:

We may summarise now the concrete results we have achieved during our investigations.

1) In Chapter II, we have studied the extractability of proteins of *Lens esculenta* BR-25 pulse seeds. We have observed that the extractability of proteins is lowest in distilled water and maximum in a phosphate buffer of pH 8.6. We have observed that the amount of protein extracted from different genetic variants of *Lens esculenta* seeds, using distilled water as extractant, differed from each other, though not to a large extent. This differential solubility corresponds to the differences among the characteristics of the proteins of the different genetic variants of the same *Lens esculenta* pulse seeds.

2) We have described in Chapter III, the preparation of eleven homogenous fractions by column chromatography of aqueous extracts of the *Lens esculenta* BR-25 pulse seeds. Out of these eleven fractions, we worked in the eight fractions. The eight fractions which we have studied, were obtained from the column chromatography on Carboxymethyl cellulose, using the step-wise elution technique. We tried several other resin like DEAE cellulose, hydroxyapatite, Amberlite IRC-50 (H) etc. but the CMC column alone provided an efficient fractionation. The step-wise elution technique was also very useful as is evident from the chromatogram.

3) In the Chapter IV the Poly Acrylamide gel Electrophoresis is described. This method is one of the most widely accepted methods for testing the homogeneity of a protein preparation. We have shown that all the eight fractions of *Lens esculenta* BR-25 seed protein prepared from
the CMC column chromatography exhibit sharp single bands in the gels. We have also calculated their individual mobilities and have concluded that they differ from each other in mobility values. We also studied electrophoresis in starch but have found that starch is not as good a carrier as the PAE, and so we have not incorporated our starch electrophoresis results.

(4) In Chapter V, we have measured the isoelectric points of the above protein fractions, by the Reverse Dye Partition technique. Since this technique involves direct measurement of charge of a protein molecule, we preferred this method after comparing with the commonly employed indirect methods of pH determination which involve determination of minima of certain physico-chemical properties.

In the end of this Chapter, we have named our protein fractions based on the isoelectric points of the individual sample. Initially "LE" is put to signify *Lens esculenta*, followed by the symbol of its solvent, say "W" for water etc. The last term in the nomenclature is the isoelectric point denoted as IP - say 5.5. Thus, the name of a particular sample becomes "LEW-IP 5.5". A detailed description regarding the procedure for nomenclature and its basis has been narrated in this Chapter.

(5) From our viscosity measurements, we made calculations of various viscosity parameters. First of all we determined the intrinsic viscosity, from the reduced viscosity vs. concentration curve.

(6) We calculated the viscosity increment and axial ratio of some of our samples in the Newtonian line of approach.
(7) We calculated the translational and rotary frictional ratios of some of our samples.

(8) A description along the Scheraga-Mandelkern line of approach for viscosity theories has been made followed by calculations of $\beta$ and $\delta$ functions of the protein samples.

(9) We made a critical discussion comparing the Einstein-Peterlin-Oncley treatment and Scheraga-Mandelkern treatment of viscosity.

(10) Amino acids of several of our fractions have been performed and we found that lysine is present in all of the three samples in a good amount. The amino acid composition of the three samples do not vary much.

(11) We have taken powder photographs of three of our samples from X-Ray scattering experiments and obtained the intensity curves for each of them. The intensity curves were helpful for the calculation of radial distribution functions, and the radial distribution curves have also been drawn. The interference intensity curves have been drawn too.

(12) The radial distribution curves furnished the value of the diameter of the helical structures.