INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.
RICE UNIVERSITY

PROTEIN DYNAMICS AT VARIOUS HYDRATION LEVELS USING THE INCOHERENT QUASIELASTIC NEUTRON SCATTERING TECHNIQUE.

by

Hung Cao

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Huey W. Huang, Chairman
Professor of Physics

Stanley A. Dodds
Associate Professor of Physics

George N. Phillips, Jr.
Professor of Biochemistry and Cell Biology

Carlton F. Hazlewood
Adjunct Professor of Biophysics

Houston, Texas
May, 1995
Abstract

Protein Dynamics at Various Hydration Levels Using the Incoherent Quasielastic Neutron Scattering Technique.

The incoherent quasi-elastic neutron scattering (IQNS) method is a useful technique to study biomolecular dynamics. The versatility of the method makes possible motional studies of biomolecules in different forms: powder, crystal, and solution; and at different temperatures. Thus, it allows for the investigation of biomolecular dynamics over a wide-range of physical conditions. We have used the IQNS method to study the motions of side chains in trypsin and myoglobin at various D2O hydration levels. The scattering spectra $S(Q,\omega)$ were measured in constant-Q mode. The protein in powder form exhibits vibrational high-frequency motions, while the protein in solution and in crystals are characterized by diffusive jumps, and high-frequency vibrations. At temperatures below 200K, the $S(Q,\omega)$ for these proteins in solution is similar to an harmonic solid. As temperature increases, a transition is seen at 200K, above which the protein becomes more liquid-like with rapid transitions between conformational substates. The diffusion constant $D$ for the side chains is on the order of $10^{-6}$ cm$^2$/sec.
Acknowledgements.

This thesis is dedicated to Dr. H. E. Rorschach. I would like to express my gratitude and respect to him for his guidance, and support.

I also would like to express my appreciation and respect to Dr. C. F. Hazlewood for his encouragement, guidance, and support.

I would like to thank Dr. S. Shapiro for all his help. The experiment could not be done without him.

Dr. T. Thurston has been very helpful all along. I appreciate his friendship.

Dr. G. N. Phillips and his students are appreciated for their help in providing the facility to prepare the sample and molecular dynamics simulation.

I would like to thank Dr. H. W. Huang for his willingness to serve as the chairman of my thesis committee. Our many discussions have helped to improve my thesis.

Dr. S. Dodds also helped in improving the thesis. I appreciate his fairness.

My fellow graduate students, especially Kevin Minard, will always be remembered for their support and enlightening discussions.

I am very thankful to have a supportive and wonderful family.

This work was funded by the Department of Energy and the R. A. Welch foundation.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>1.1. The Objective of This Thesis</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Biomolecules' Related Motions</td>
<td>1</td>
</tr>
<tr>
<td>1.3. Techniques for Probing Protein Motions</td>
<td>3</td>
</tr>
<tr>
<td>1.4. Previous Related Experiments</td>
<td>3</td>
</tr>
<tr>
<td>1.5. Preview of the Thesis Contents</td>
<td>3</td>
</tr>
<tr>
<td>II. Experimental Methods</td>
<td>5</td>
</tr>
<tr>
<td>2.1. Sample</td>
<td>5</td>
</tr>
<tr>
<td>2.2. Spectrometer</td>
<td>8</td>
</tr>
<tr>
<td>III. Scattering Laws of Diffusive Motion</td>
<td>11</td>
</tr>
<tr>
<td>3.1. General Formulation of Scattering Law</td>
<td>13</td>
</tr>
<tr>
<td>3.2. Inelastic, Elastic, and Quasielastic Scatterings</td>
<td>16</td>
</tr>
<tr>
<td>3.3. Vibrational-Jump Diffusional Model</td>
<td>20</td>
</tr>
<tr>
<td>3.4. Scattering Law of Vibrational-Jump Diffusion Motion</td>
<td>23</td>
</tr>
<tr>
<td>IV. Spectrometer Resolution, Background and Multiple Scattering</td>
<td>28</td>
</tr>
<tr>
<td>4.1. Spectrometer Resolution Function</td>
<td>28</td>
</tr>
<tr>
<td>4.2. Inelastic Scattering</td>
<td>33</td>
</tr>
<tr>
<td>4.3. D$_2$O Background</td>
<td>34</td>
</tr>
<tr>
<td>4.4. Multiple Scattering</td>
<td>34</td>
</tr>
<tr>
<td>V. Data Analysis</td>
<td>36</td>
</tr>
<tr>
<td>5.1. Data Correction</td>
<td>36</td>
</tr>
<tr>
<td>5.2. Experimental Results</td>
<td>44</td>
</tr>
<tr>
<td>5.3. Discussion</td>
<td>64</td>
</tr>
<tr>
<td>5.4. Conclusion</td>
<td>76</td>
</tr>
</tbody>
</table>
References.
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE II-1a</th>
<th>Aluminum chamber for holding dry sample.</th>
</tr>
</thead>
<tbody>
<tr>
<td>FIGURE II-1b</td>
<td>Aluminum chamber for holding liquid sample.</td>
</tr>
<tr>
<td>FIGURE II-2</td>
<td>Sketch of the BNL triple-axes spectrometer at the cold neutron source H9.</td>
</tr>
<tr>
<td>FIGURE III-1</td>
<td>Layout of a typical triple-axes spectrometer.</td>
</tr>
<tr>
<td>FIGURE IV-1</td>
<td>Sketch of trajectories for vibrational-jump-diffusive motion.</td>
</tr>
<tr>
<td>FIGURE IV-2</td>
<td>Schematic diagram of experimental incoherent neutron scattering spectra: elastic, inelastic and quasielastic.</td>
</tr>
<tr>
<td>FIGURE IV-3</td>
<td>Restricted diffusion between two impermeable walls.</td>
</tr>
<tr>
<td>FIGURE IV-4</td>
<td>Restricted diffusion on the surface of a sphere.</td>
</tr>
<tr>
<td>FIGURE IV-5</td>
<td>Restricted jump diffusion between two impermeable walls.</td>
</tr>
<tr>
<td>FIGURE V-1a</td>
<td>Intergrated intensity S(Q) for D\textsubscript{2}O solution.</td>
</tr>
<tr>
<td>FIGURE V-1b</td>
<td>D\textsubscript{2}O background spectra for 300K.</td>
</tr>
<tr>
<td>FIGURE V-2a</td>
<td>Results of fitting simulated data with only a Lorentzian.</td>
</tr>
<tr>
<td>FIGURE V-2b</td>
<td>Results of fitting simulated data with only a Lorentzian.</td>
</tr>
<tr>
<td>FIGURE V-3a</td>
<td>Spectra of powder protein for various Q at 300K (raw data).</td>
</tr>
<tr>
<td>FIGURE V-3b</td>
<td>Elastic scattering spectrum of powder protein at 320K.</td>
</tr>
<tr>
<td>FIGURE V-4a</td>
<td>Spectra of trypsin-solution for various Q at 200K (raw data).</td>
</tr>
<tr>
<td>FIGURE V-4b</td>
<td>Elastic scattering spectrum of protein-solution at 200K.</td>
</tr>
<tr>
<td>FIGURE V-5a</td>
<td>Spectra of myoglobin-solution for various Q at 280K (raw data).</td>
</tr>
<tr>
<td>FIGURE V-5b</td>
<td>Spectra of myoglobin-solution for various Q at 300K (raw data).</td>
</tr>
<tr>
<td>FIGURE V-6a</td>
<td>Linewidth vs. Q for myoglobin-solution at 280K-single-jump length fit.</td>
</tr>
<tr>
<td>FIGURE V-6b</td>
<td>Linewidth vs. Q for myoglobin-solution at 280K-gaussian jump length distribution fit.</td>
</tr>
</tbody>
</table>
FIGURE V-7. Intergrated intensity vs. Q for myoglobin-solution at 280K-harmonic model.

FIGURE V-8a. Linewidth vs. Q for myoglobin-solution at 300K-single-jump length fit.

FIGURE V-8a. Linewidth vs. Q for myoglobin-solution at 300K-gaussian jump length distribution fit.

FIGURE V-9. Intergrated intensity vs. Q for myoglobin-solution at 300K-harmonic model.

FIGURE V-10a. Spectra of protein crystals for various Q at 300K.

FIGURE V-10b. Intergrated intensity vs. Q for myoglobin crystals at 300K-harmonic model.

FIGURE V-10c. Linewidth vs.Q for trypsin crystals-gaussian jump length distribution fit.

FIGURE V-11. Model potentials for protein motions. Representations of harmonic potential (left), and multi-minimum potential (right).

FIGURE V-12a. Thermal vibrations at different temperature for myoglobin in solution.

FIGURE V-12b. Different studies of thermal vibrations of proteins in solution.

FIGURE V-13a. MD simulation of myoglobin for 17 ps-jump motion.

FIGURE V-13b. MD simulation of myoglobin for 17 ps-jump motion.
1. Chapter I
   Introduction.

1.1. The Objective of This Thesis

We have used the incoherent quasi-elastic neutron scattering (IQNS) technique to explore the dynamics of chain segments of biomolecules, trypsin and myoglobin. Our main goal is to characterize the motions of protein hydrogens to further our understanding of the basic mechanism of diffusion that gives rise to relaxation times in nuclear magnetic resonance studies, and to compare these motions in different states: powder, solution, and crystals\(^\text{1}\). Previous work related to ours will be discussed. Some relevant results from molecular dynamics simulations will also be mentioned\(^\text{2}\).

We were able to perform a detailed study of the quasi-elastic linewidth and the Debye-Waller factor for trypsin and myoglobin proteins in powder, solution and crystals at different temperatures. For the data in solution and crystals about room temperature, we model the protein's hydrogens motions using the vibrational-jump diffusion picture. For data in powder and low-temperature solution samples, we model these motions using harmonic vibration picture\(^\text{3}\). Before details of our study are presented, a brief introductions of neutron techniques, biomolecules and techniques used in the study of protein motions are presented.

1.2. Biomolecules' Related Motions

The motions of proteins cover a wide range of amplitudes (0.01 to 100 Å) and time scales (10\(^{-15}\) to 10\(^3\) s)\(^\text{4}\). Among this richness of motional phenomena, we are mainly interested in the thermal vibrations and diffusive jump motions (≈ 10\(^{-11}\) s) of protein side
chains. The high frequency localized thermal motion of each atom in the molecule can be
determined from the measurements of the coherent scattering of X-rays and neutrons
scattering from single crystals of biological materials. The diffusive jumps motions are
obtained by slight shifts in positions of various segments of the molecules. Each of these
positions is often called a micro-conformation. The transitions among these different
"micro-conformations" are of a diffusive nature that are thought to be important to the
biomolecule's functions<2-9>. We will show that a convenient way of measuring such
diffusive motion is IQNS.

1.3. **Techniques for Probing Protein Motions**

The wide range of motions of biomolecules requires more than one technique to
fully explore their nature. Several techniques have been used: light scattering, X-ray and
neutron diffraction, diffuse X-ray scattering, small-angle neutron scattering, nuclear
magnetic resonance (NMR), Mössbauer spectroscopy, flash photolysis, inelastic neutron
scattering, computer simulation, quasi-elastic neutron scattering (QNS), etc.<6,10,11>.

Among scattering techniques, X-ray and neutron are the two most powerful
methods used to study biomolecules at the atomic level. In many respects, these two
methods are comparable in both their experimental methodologies, and in the resulting
informational content. Unlike X-ray, however, neutron scattering factors do not increase
with heavier elements. It also has an unusual advantage, that is being able to identify light
element such as hydrogen due to its large incoherent scattering cross section. This special
feature is very useful in the study of proteins whose atoms consist of a large number of
hydrogens – nearly one half the total number of atoms – and these hydrogens dominate
much of the chemical and physical structure of proteins.
The main advantage of neutron scattering over nuclear magnetic resonance (NMR) lies in the fact that very short correlation times can be investigated (10^{-9} sec to 10^{-12} sec). For NMR, this time region is beyond the minimum measurable relaxation rate 1/T_1.

One other important and powerful technique is molecular dynamics simulation. It has provided insights in bridging the gap between theory and experiments in structural and dynamical properties of biomolecules. It also provides a wealth of information which cannot be obtained by other techniques individually. Furthermore, it is an excellent way to visualize some of the dynamical concepts that could be difficult to grasp otherwise.<12-17>

1.4. **Previous Related Experiments**

QNS studies of polymers, trypsin and muscle tissues were done by C. Trantham, D. Bearden and C. Lin<18-20>. These studies focus on the high frequency thermal vibrational motion of hydrogen in trypsin by measuring S(Q) – the integrated intensity of the quasi-elastic peak. From these data, the mean-squared local vibrational amplitude can be extracted from the Debye-Waller (DW) factor. However, the details of the atomic diffusive motion can not be obtained in this way because no information on the line shape was obtained. Corrections for background, spectrometer-resolution, and inelastic scattering proved to be very difficult using this low-resolution IQNS technique.

Several other studies<5,21-25> also concentrate on protein in slightly hydrated powder forms. Very few experimental studies of protein motion in picosecond timescale used proteins in solution.

1.5. **Preview of the Thesis Content**
After the introduction, the chapters are arranged in the following order. First, chapter II briefly introduces the proteins under consideration: trypsin and myoglobin; and then details the sample preparation methods that include D$_2$O exchanged, crystallization and also describes the sample chambers and the triple-axes spectrometer. In chapter III, we present the space-time correlation functions and scattering laws and show how these functions can be used to extract dynamical properties of proteins using quasielastic neutron scattering technique. In addition, we discuss the physical origin of the elastic peak and relate it to various physical pictures under which it can arise. The later part of chapter III shows the derivation of the scattering law for the vibrational-jump diffusion model that was used in our data analysis. Much of the theoretical derivation on the vibrational-jump diffusion picture was done recently by Dr. H. E. Rorschach. The idea of vibrational-jump diffusion model is not original, but what is interesting is the general formulation that allows for inclusion of anharmonicity if it arises. Next, chapter IV discusses the data analysis procedures. This chapter includes spectrometer resolution function: the approximated form (a gaussian); the effects of multiple scattering and inelastic background; and the D$_2$O background correction. Chapter V is the central part that includes the findings, discussion and summary of our experiments.
Chapter II
Experimental Methods

2.1. Sample

2.1.1. Myoglobin

Myoglobin is a globular protein that plays an important role in the respiratory cycle in vertebrates. It is located in the muscle and acts as the oxygen-buffer within the muscle. The myoglobin's role of binding oxygen depends on a heme group, which consists of an organic part and an iron atom. Myoglobin was the first protein to have its three-dimensional structure seen by X-ray at the atomic resolution. It consists of 153 amino acids. Of these amino acids, 1247 atoms are hydrogens, 769 are carbons, 220 are nitrogens and 215 are oxygens. Here, the hydrogens are slightly more than 50% of the total number of atoms. Myoglobin's structure consists of eight α-helices folded together\(^{26}\). It is one of the most widely studied proteins because it can be obtained in large quantities, relatively small in size, and easily crystallized.

2.1.2. Trypsin

Trypsin is an enzyme released from the pancreas. It is responsible for the hydrolysis of peptide bonds in proteins. Trypsin consists of 223 amino acids. Of these amino acids, 1604 atoms are hydrogens, 1003 are carbons, 284 are nitrogens and 320 are oxygens. Here, the hydrogens are about 50% of the total number of atoms. The structure type of trypsin is mostly \(\beta\) form. \(^{27-30}\)
Trypsin has the ability to digest itself, so in order to prevent this autodigestion process during the experiment we adjusted the pH of trypsin-D$_2$O solution to between 4 and 5 using DCl. We also used benzamidine to inhibited the trypsin from auto digestion in the trypsin crystals. <27>

2.1.3. **Sample Preparation**

Since our experiment was designed to look at the motions of the hydrogens in the proteins, D$_2$O is used as the solvent. Initially, the loosely bound hydrogens were exchanged with deuterons by the following procedure. First, the protein powders were mixed with D$_2$O in a closed container at about 6% of protein concentration. The sample was then left in a refrigerator at about 30°C for about 20 hours. Next, we freeze-dried the sample by placing it in an ethyl alcohol + dry-ice bath; and then in a vacuum chamber for about 30 hours. The proteins that are prepared in this way are called treated proteins.

In the solution experiment, the treated proteins were mixed with D$_2$O to form a protein solution. The sample was placed in a cryostat filled with He gas during the neutron scattering experiment for temperature-control purpose and to prevent the condensation of water on the sample container surface. There are two different designs of the aluminum sample chambers. One is made for holding dry samples (see Fig. II-1a). The other is made for holding liquid samples to minimize the liquid loss during the experiment (see Fig. II-1b).
Fig. II-1a. Aluminum chamber for holding dry sample.

holes for injecting liquid sample

groove for Indium wire seal

Fig. II-1b. Aluminum chamber for holding liquid sample.
We prepared the protein solution for crystallization using the standard procedures as described in the references. This protein solution contains about 60 mg trypsin, 10 mg benzamidine and 1 mg CaCl per ml (between pH 5 and 7). The benzamidine was used to prevent the trypsin from autodigestion. This type of trypsin is often called as benzamidine-inhibited trypsin. In order to obtain a large quantity of crystals (about 2g), we placed several drops of prepared protein-solution on each of several glass plates inside large containers with 1.4 to 1.6 M-ammonium sulfate. Within a few days, long orthorhombic crystals of trypsin, up to 0.5 mm x 0.5 mm x 2.0 mm, grew at room temperature. This standard technique of growing crystals is called the vapor diffusion technique.

Horse myoglobin crystals were grown using the standard batch method. First, the proteins were dissolved (8mg/ml) in a solution containing 3.0 M of (NH₄)₂SO₄ and 0.1 M of sodium acetate adjusted to pH 7.0. This solution was then filtered through a 0.22m low pressure filter-funnel. The batch crystallizations were carried out in varies quantities 1.0 ml to 5.0 ml samples, which were allowed to stand in disposable glass tubes. The crystals appeared within a week. The sperm whale myoglobin crystals were also grown by a similar method.

2.2. **Spectrometer**

Our experiment was carried out at Brookhaven National Laboratory (BNL) on the H9-A triple-axis spectrometer, which is located at the cold neutron facility (see Fig. II-2). To provide intense beams of very low energy neutrons (≤ 5 meV), a liquid-hydrogen moderator system has been installed in the beam thimble of H9. The moderator chamber contains 1.4 liters of liquid hydrogen and is located near the beginning of the beam lines. Circulating cold helium gas from an external refrigerator maintains the hydrogen at a
temperature of 19K. Neutrons from the surrounding heavy water blanket are moderated by scattering processes within the cold hydrogen, thus shifting the spectral distribution of the emerging beam towards lower energies. In operation, the system serves as a cold neutron source for three separate beams A, B, and C as shown in Fig. II.2.

H9-A is primarily used for high resolution inelastic neutron scattering studies of single crystal, polycrystalline, or amorphous sample. However, as in our case, it can also easily be adjusted to quasielastic scattering experiments. It is equipped with a double monochromator housing that allows for large monochromator scattering angles. A Beryllium filter is placed before the monochromator for elimination of neutrons with energy above the cutoff energy (≈ 5 meV for Be). The incident energy can be varied, so scans with a fixed final energy can be obtained, or a fixed incident energy with varied final energies can also be used. The sample table is equipped with a manual two-axis goniometer with a ±10° range and is compatible with a variety of available cryostats and furnaces. External collimation is provided by Soller slits with 5x5 cm² cross section<sup>31</sup>.

**H9-A Station Parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monochromator scattering angle</td>
<td>60°&lt;θ&lt;134°</td>
</tr>
<tr>
<td>E₀ or λ₀ (PG 002)</td>
<td>2.15&lt;E₀&lt;7.3meV</td>
</tr>
<tr>
<td></td>
<td>3.3&lt;λ₀&lt;6.2Å</td>
</tr>
<tr>
<td>Beam size at sample</td>
<td>7.5 (h) x 5 (w) cm²</td>
</tr>
<tr>
<td>Sample scattering angle</td>
<td>-50°&lt;θ&lt;150°</td>
</tr>
<tr>
<td>Analyzer scattering angle</td>
<td>-125°&lt;θ&lt;125°</td>
</tr>
<tr>
<td>In-pile collimation</td>
<td>15', 30', or 60'</td>
</tr>
<tr>
<td>Detector</td>
<td>BF₃</td>
</tr>
</tbody>
</table>
Fig. II-2. Sketch of the BNL triple-axis spectrometer at the cold neutron source H9.
Chapter III

Scattering Law of Diffusive Motion

In this chapter, the general forms of neutron scattering law are discussed using the space-time correlation functions, or the van Hove formalism. The experimental neutron scattering spectrum – the elastic and the quasi-elastic parts – are also considered. In particular, we will present some examples under which the elastic peak plays a major role. The last part of this chapter presents the formulation that leads to the vibrational-jump diffusion model, which we used to analyze our data.

Before presenting the vibrational-jump diffusion model, we first show how the vibrational motion can be related to the single-particle distribution function, which takes into account the anharmonicity. Anharmonicity, in our terminology, is any motion(s) that causes a deviation from the often-used harmonic Debye-Waller factor (see discussion of eq. III-7). We then arrive at the general expression for the scattering law of jump diffusion. In the later part of this section, we discuss the inclusion of vibrations and of jump-diffusion to make the vibrational-jump diffusion picture.

Fig. III-1. Sketch of trajectories for vibrational-jump-diffusive motion.
We seek to explain the hydrogens' motion associated with the biomolecules, trypsin and myoglobin, using the vibrational-jump diffusion model, which fits the data better than other related models. Previous experiments in our group on polymers have shown a behavior similar to that predicted by vibrational-jump diffusion theory. We later attempted to use the molecular dynamic (MD) simulation results only to elucidate qualitatively the vibrational-jump-diffusion mechanism for the sidechains' hydrogens and for visualization purpose.

Using the general formulation for the vibrational motion, i.e. the DWF (Debye-Waller Factor), and the jump-diffusion, we put together a general formulation of vibrational-jump-diffusion picture that allows both harmonic and anharmonic descriptions of the DWF, in which the harmonic approximation is a special case.

Before going into the details of jump-diffusion theory, we first look at the inadequacy of continuous diffusion at large values of momentum transfer $Q$, or small space-time values. The continuous diffusion predicts a rms velocity $\sqrt{\langle r^2 \rangle / t^2} \propto t^{-0.5}$, so for $t$ small the rms velocity becomes physically unreasonable. This model is indeed only applicable to the continuum; the motion of particles under the detailed influence of their neighbors requires a different picture.

In general, comparisons between continuous and jump-diffusion models can be summarized in a qualitative way. For small $Q$ values, the scattering neutrons interact with motion over large length scale or many diffusive steps. In this $Q$ range, the linewidth varies linearly with $Q$ for both continuous and jump-diffusion models. For large $Q$ values, the neutron scattering wave can interact with single jump steps (as in jump-diffusion) or small length scale. Here ($> 1.5 \text{ Å}^{-1}$), the linewidth versus $Q$ is non-linear for jump diffusion processes; while the continuous diffusion model predicts that the linewidth varies linearly with $Q$. 
3.1. **General Formulation of Scattering Law**

3.1.1. **Van Hove Correlation Functions**

Information about both structure and dynamics of molecules can be obtained from the scattering law through the following van Hove's formalism. The van Hove neutron-scattering law can be written as

$$S(Q,\omega) = \frac{N\sigma k_1}{\hbar k_0} \int_{-\infty}^{\infty} G(r,t) e^{iQ \cdot r} e^{-i\omega t} \, d^3r \, dt,$$  \hspace{1cm} (III-1)

where

- $S(Q,\omega)$ is the scattering cross section,
- $\hbar\omega = E_0 - E_1$ is the neutron's energy lost,
- $k_1, k_0$ are the neutron's incident and scattered wave vector respectively,
- $Q = k_1 - k_0$ is the scattering vector,
- $N$ is the number of nuclei in the sample,
- $\sigma$ is related to the scattering length $b$ which will be discussed later.

Here $S(Q,\omega)$ is proportional to the space and time Fourier transforms of the time-dependent pair-correlation function or the van Hove correlation function $G(r,t)$. This general result gave a unified description for all neutron scattering experiments. The fact that $S(Q,\omega)$ is simply the Fourier transform of a function that gives the probability of finding two atoms a certain distance apart is responsible for the power of neutron scattering. The van Hove correlation can be written as

$$G(r,t) = \left< \frac{1}{N} \sum_{jk} \delta(r - [r_k(0) - r_j(t)]) \right>$$  \hspace{1cm} (III-2)
where $< \ldots >$ is called the thermal average. The delta function in the definition of $G(\mathbf{r}, t)$ is zero except when the position of an atom $k$ at time zero and the position of atom $j$ at time $t$ are separated by the vector $\mathbf{r}$. Because the delta functions are summed over all possible pairs of atoms to obtain $G(\mathbf{r}, t)$, this function is equal to the probability of an atom being at the origin of a coordinate system at time zero and an atom being at position $\mathbf{r}$ at time $t$. $G(\mathbf{r}, t)$ is also often referred to as the time-dependent pair-correlation function because it describes how the correlation between two particles evolves with time.

3.1.2. Coherent and Incoherent Scatterings

In a system of a given atomic species $i$, with many isotopes having a nuclear spin, the interaction depends on the properties of the nucleus and on the total spin state of the nucleus-neutron system, and most isotopes have several spin states. In general, however, there is no correlation between the spin of a nucleus and its position in a sample. Thus, the scattering lengths $b$ can be averaged over the nuclear spin states without affecting the thermodynamic average $<\ldots>$.

As we average the scattering law $S(Q, \omega)$ over the nuclear spin states, there are two quantities to be considered: the average value of $b$ ($<b>$) and the average value of $b^2$ ($<b^2>$). The average over the nuclear spin states of $S(Q, \omega)$ gives

$$\sum_{jk} b_j b_k \Omega_{jk} \Omega_{jk} = \sum_{jk} b_j^2 \Omega_{jk} + \sum_{j} (b_j^2 - <b>^2) \Omega_{jj} = S_c(Q, \omega) + S_i(Q, \omega) \quad (III-3)$$

where

$$\Omega_{jk} = \int_{-\infty}^{\infty} <\exp(iQ \cdot r_k(0)) \exp(iQ \cdot r_j(t))> e^{-i\omega t} \, dt$$

$$= \int_{-\infty}^{\infty} \delta(r - [r_k(0) - r_j(t)]) e^{iQ \cdot r} e^{-i\omega t} \, dt \, d^3r$$
The first term on the right side of above equation represents the so-called coherent scattering, whereas the second represents the incoherent scattering.

Based on van Hove's formulation above, the distinct and self correlation functions are defined respectively as,

$$ G(r,t) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} S_c(Q,\omega) e^{iQ\cdot r} e^{i\omega t} dQ d\omega, \quad (\text{III-4}) $$

$$ G_s(r,t) = \frac{1}{(2\pi)^3} \int_{-\infty}^{\infty} S_i(Q,\omega) e^{iQ\cdot r} e^{i\omega t} dQ d\omega. \quad (\text{III-5}) $$

By taking inverse Fourier transform of the above equations, we can obtain $S_c(Q,\omega)$ and $S_i(Q,\omega)$.

Furthermore, we can define the coherent and incoherent scattering lengths

$$ b_{\text{coh}} = <b> \quad (\text{III-6}) $$

$$ b_{\text{inc}} = [ <b^2> - <b>^2 ]^{\frac{1}{2}} \quad (\text{III-7}) $$

The coherent scattering law is a sum over both $j$ and $k$ (see Eq. III-3) and thus involves correlations between the position of an atom $j$ at $t=0$ and the position of a second atom $k$ at time $t$. Thus the coherent scattering represents interference between waves produced by the scattering of all the nuclei in the sample.

The incoherent scattering law, on the other hand, describes the correlations between the position of an atom $j$ at $t=0$ and the position of the same atom at time $t$. Thus, in the incoherent scattering, there is no interference effect between scattered waves from different nuclei. For this reason, incoherent scattering provides a good way of examining processes in which atoms diffuse.
3.2. **Inelastic, Elastic and Quasielastic Scatterings**

The incoherent scattering spectra can be divided into three regions: inelastic, elastic, and quasielastic (see Fig. III-2). The inelastic spectra contain regions of non-zero values of

![Schematic diagram of experimental incoherent neutron scattering spectra: elastic, inelastic and quasielastic.](image)
energy transfer. It gives details on the exchange of energy with vibrational modes in the molecule, and thus is an indispensable tool of vibrational spectroscopy. The elastic peak has information on the average spatial distribution of the restricted nuclei. The origin of the elastic peak will be discussed in a later section. The quasieelastic profile includes region around \( \omega=0 \). It provides details of stochastic motions such as diffusive motions and high frequency vibrations. It is often modeled as a Lorentzian(s). In our experiment, we are only interested in the elastic and quasielastic region.

3.2.1. Physical Origin of Elastic Peak

In general, the elastic scattering arises from scatterers being confined in space. The following examples discuss circumstances under which the elastic peak presents in the scattering law. There are other examples under which the elastic peak can arise\(^{<43>}\), but will not be discussed here.

3.2.1.1. Restricted Diffusion between Two Impermeable Walls

![Fig. III-3. Restricted diffusion between two impermeable walls.](image-url)
The simplest example of restricted diffusion is the case of diffusion in one dimension between two impermeable walls separated by distance L<sup>47</sup>.

The probability \( G(x,x_0,t) \) of finding a particle at \((x,t)\) if it was at \((x_0,t_0)\) follows the one-dimensional diffusion equation,

\[
\frac{\partial}{\partial t} G(x,x_0,t) = D_x \frac{\partial^2}{\partial x^2} G(x,x_0,t) \tag{III-8}
\]

with the boundary conditions

(i) \( G(x,x_0,t=0) = \delta(x-x_0) \)

(ii) \( \left( \frac{\partial G}{\partial x} \right)_{x=0,L} = 0 \) for all \( t > 0 \)

where \( D_x \) is the one-dimensional diffusion coefficient. Here the boundary condition (ii) requires zero concentration gradient inside the walls.

By Fourier transforming the solution of the above equation, we can obtained the incoherent scattering law<sup>43</sup>

\[
S_{\text{inc}}(Q,\omega) = A_0(Q_xL)\delta(\omega) + \sum_{n=1}^{\infty} A_n(Q_xL)L_n(\omega) \tag{III-9}
\]

where the widths of the Lorentzian functions \( L_n(\omega) \) are

\[
\tau_n^{-1} = \frac{n^2\pi^2D_x}{L^2},
\]

and the elastic and quasielastic structure factors are defined as,
\[ A_0(QxL) = \frac{2(1 - \cos(QxL))}{(QxL)^2} = J_0^2 \left( \frac{L}{2Q} \right) \]  \hspace{1cm} (III-10)

and

\[ A_n(QxL) = \frac{4(QxL)^2[1 - (-1)^n \cos(QxL)]}{[(QxL)^2 - (n\pi)^2]^2} \]  \hspace{1cm} (III-11)

respectively. The emphasized feature here is the elastic peak that originates from the particle being restricted in space.

3.2.1.2. **Restricted Diffusion on an Impermeable Sphere**

For the case of diffusion on a sphere of radius \( \rho \), the probability distribution \( G_s \) of its coordinate \( \Omega \) is governed by

\[ \frac{\partial G_s}{\partial t} = D_r \nabla^2 G_s \]  \hspace{1cm} (III-12)

Fig. III-4. Restricted diffusion on the surface of a sphere.

where \( G_s (\Omega, \Omega_0, t) \) is the probability of finding the particle at \( \Omega \) at time \( t \) if it was at \( \Omega_0 \) at \( t = 0 \). \( D_r \) is the rotational diffusion coefficient.
The scattering law is\textsuperscript{<40,47,48>}

\[ S (Q, \omega) = J_0^2 (Q \rho) \delta(\omega) + \frac{1}{\pi} \sum_{l=1}^{\infty} (2l + 1) J_l^2 (Q \rho) \frac{D_r}{[D_r (1+1)]^2 + \omega^2} \]  \hspace{1cm} (III-13)

where all the functions above carry their usual mathematical definitions. It is seen that the scattering law is composed of a sharp peak superimposed on a broadened component (a sum of Lorenztians) whose width is of the order of a few \( D_r \) and whose intensity depends on \( Q \). The elastic term will vanish quickly as soon as \( Q \rho > \pi \).

Another example worth mentioning is the diffusion inside an anisotropic shape of an impermeable cylinder surface. The presence of elastic scattering is again due to confinement by an impermeable wall.

3.3. **Vibrational-Jump Diffusion Model**

3.3.1. **Vibrational Motion (Debye-Waller Factor DWF)**

In general the single-particle DWF can be expressed as\textsuperscript{<49>},

\[ \text{DWF} = \exp (-2W) = |\exp(-iQ \cdot r)|^2 \]  \hspace{1cm} (III-14)

Let

\[ W(Q) = -\ln<\exp(-iQ \cdot r)>. \]  \hspace{1cm} (III-15)

Using Taylor series expansion and assuming \( -iQ \cdot r = 0 \), we obtained
\[ W(Q) = \frac{1}{2} \langle (Q_r)^2 \rangle - \frac{1}{6} i \langle (Q_r)^3 \rangle - \frac{1}{24} \{ \langle (Q_r)^4 \rangle - 3 \langle (Q_r)^2 \rangle^2 \} + \cdots \quad \text{(III-16)} \]

The first term in the expression is the harmonic approximation of the DWF. The second term is zero if the potential is even in \( r \). The third term contribution is also zero if the vibration motion distributions are Gaussian caused by harmonic forces. In the presence of the anharmonic force, this term will not vanish.

In the quasi-elastic region, the vibrational motion can be approximated by the harmonic DWF \( \exp(-Q^2 <u^2>) \), \( <u^2> \) is the mean square vibration amplitude; showing the intensity decreased with increasing \( Q \). This Debye-Waller factor plays an important role in protein dynamics, especially in the conformational substates concept pursued by Frauenfelder and others (see discussion in chapter V)<sup>36,40</sup>.

### 3.3.2. Scattering Law of Jump Diffusion

The theory of continuous diffusion is satisfactory when spatial and temporal values \( (r,t) \) are large compared to the elementary values \( (l,\tau) \) of diffusion steps. When the experimental values of \( (Q,\omega) \) are equivalent to small values of \( (r,t) \) or close to \( (l,\tau) \), the continuous diffusion theory breaks down. In this case, it is necessary to take into account the details of a diffusive step in order to calculate the scattering law<sup>50</sup>.

A picture of a jump diffusive particle is as following: the particle spends some time \( \tau_0 \) at a given site. In this time \( \tau_0 \), it may vibrate about the center of equilibrium. After a time \( \tau_0 \), the particle then moves quickly, in a time much smaller compared to \( \tau_0 \), to a new site located at distance \( l \) relative to the original site. The distance \( l \) is larger than the dimensions of the thermal cloud (see Fig. III-1). This diffusive picture is repeated indefinitely. Let \( G_s(r,t) \) be the probability of finding an atom at \( r \), the time rate of change of \( G_s(r,t) \) is given by a rate equation, i.e.
\[ \frac{\partial G_s(r,t)}{\partial t} = \frac{1}{n\tau_0} \sum_{i=1}^{n} [G_s(r+I_i,t) - G_s(r,t)] \]  

(III-17)

where \( n \) is the number of available sites for jumping. In the limit \(|r| \gg \lll\), i.e. macroscopic limit, the term \( G_s(r+I_i,t) \) can be expanded about \( r \) and continuous diffusive theory is recovered.

Equation (III-17) can be Fourier transformed (\( r \) to \( Q \)) to give

\[ \frac{\partial I_s(Q,t)}{\partial t} = \frac{-1}{n\tau_0} \sum_{i=1}^{n} (1 - \exp(-iQ \cdot I)) \ I_s(Q,t) \]  

(III-18)

using

\[ \int G_s(r+I_i,t) \exp(-iQ \cdot r) \ dr = \exp(-iQ \cdot I) \ I_s(Q,t) \]  

(III-19)

Applying the appropriate initial condition \( G_s(r,0) = \delta(r) \), we obtain

\[ G_s(r,t) = \exp[-(t/r) f(Q)]. \]  

(III-20)

The Fourier transformation with respect to time finally gives,

\[ S_s(Q,\omega) = \frac{1}{\pi} \frac{f(Q)}{\omega^2 + [f(Q)]^2}, \]  

(III-21)

where

\[ f(Q) = \frac{1}{n\tau_0} \sum_{i=1}^{n} (1 - \exp(-iQ \cdot I)). \]

For continuous diffusion equation, it can be shown that
\[ G_s(r,t) \propto \exp[-r^2/4Dt], \quad (\text{III-22}) \]

and

\[ S_s(Q,\omega) = \frac{1}{\pi} \frac{DQ^2}{\omega^2 + [DQ^2]^2}. \quad (\text{III-23}) \]

Here, we have the usual linewidth = $DQ^2$ for continuous diffusion. Notice that the general form of the scattering function $S_s(Q,\omega)$ is a Lorentzian. Assume that the vectors $l$ are oriented at random (i.e. jumping takes place in random directions) and that the distribution of lengths $l$ is a continuous function of $l$ (let this length distribution function be $a(l)$). Thus $f(Q)$ is averaged over angles and the sum changed to an integral to give

\[ f(Q) = \frac{1}{\tau_0 A} \int \left( 1 - \frac{\sin Ql}{Ql} \right) a(l) \, dl \quad (\text{III-24}) \]

To illustrate the technique, let us assume that the length distribution $a(l)$ has a shape of a random distribution, e.g. $a(l) = l \exp(-l/l_0)$, then $f(Q)$ becomes

\[ f(Q) = \frac{1}{\tau_0} \left[ 1 - \frac{1}{1 + [Ql_0]^2} \right]. \quad (\text{III-25}) \]

### 3.4. Scattering Law of Vibrational-Jump-Diffusion Motion

#### 3.4.1. Theoretical background

In our quasi-elastic scattering experiments, the integrated intensity is determined from the area of the quasi-elastic peak, after subtracting the inelastic and $D_2O$ background:
It will be shown below that $S_m(Q)$ is roughly equal to the DWF in certain approximation. In the process of measuring this quantity, we only accepted the scattering intensity in the range $-\omega_q < \omega < \omega_q$. This process is equivalent to the above integral Eq. (III-26) over the range $-\omega_q < \omega < \omega_q$, where $\omega_q$ is chosen so as to include the quasi-elastic line but not the inelastic part of the spectrum (see Fig. III-2). Substituting Eq. (III-26) in Eq. (III-1), and integrating over $\omega$, we obtain:

\[
S_m(Q) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 2 \frac{\sin \omega_q t}{t} \exp(iQ \cdot r) G_s(r,t) \, dr \, dt \quad \text{(III-27)}
\]

We now suppose that the self-correlation function $G_s(r,t)$ is due to two uncorrelated types of motion, one slow on the scale of $\omega_q$ and the other fast on this scale:

\[
G_s(r,t) = G_{\text{slow}}(r,t) \otimes G_{\text{fast}}(r,t) \quad \text{(III-28)}
\]

where $\otimes$ represents a convolution of the two correlation functions. Substituting Eq. (III-28) in Eq. (III-27), the resulting function contains the factor $\frac{\sin \omega_q t}{t} G_{\text{slow}}(r,t)$. This function is nearly constant for $\omega_q t \ll 1$, and it is slowly-varying compared with $G_{\text{fast}}$, whose component frequencies are large compared to $\omega_q$. In the integral in Eq. (III-27), $G_{\text{fast}}$ may therefore be replaced by its time-average:

\[
<G_{\text{fast}}(r,t)>_t = \int \rho(r - r') \rho(r') \, dr' = \rho(r) \otimes \rho(r) \quad \text{(III-29)}
\]
where \( \rho(\mathbf{r}) \) is the time-averaged single-particle density function, and \( \int \rho(\mathbf{r}) \, d\mathbf{r} = 1 \).<sup>53</sup>

The Fourier transform over \( \mathbf{r} \) in Eq. (III-27) can then be performed, making use of the convolution theorem:

\[
S_m(Q) = \left[ \int_{-\infty}^{\infty} <G_{\text{fast}}(\mathbf{r},t) \exp(iQ \cdot \mathbf{r}) \, d\mathbf{r} \right] \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} 2 \frac{\sin \omega_q t}{t} G_{\text{slow}}(\mathbf{r},t) \exp(iQ \cdot \mathbf{r}) \, d\mathbf{r} \, dt
\]

(III-30)

Substituting Eq. (III-29) in Eq. (III-30) and making use of the convolution theorem in the first integral,

\[
\int_{-\infty}^{\infty} <G_{\text{fast}}(\mathbf{r},t) \exp(iQ \cdot \mathbf{r}) \, d\mathbf{r} = \int_{-\infty}^{\infty} \rho(\mathbf{r}) \exp(iQ \cdot \mathbf{r}) \, d\mathbf{r} \]

(III-31)

The term in brackets in Eq. (III-30) is the square of the "Debye-Waller factor" (DWF = \( \exp(-W_D) \)). This term is independent of \( Q \) and \( \omega_q \) if \( \omega_q \) is large compared to the line-width of the quasi-elastic spectrum associated with the diffusive motion. The second term in Eq. (III-30) gives the quasi-elastic line.

In the above form, the DWF is seen to be due to diffraction from the time-averaged density distribution of the high-frequency motion of the scattering particle. If the particle oscillates about an equilibrium position with a \( \rho(\mathbf{r}) \) that is Gaussian, then the Fourier transform will be Gaussian, and we obtain the usual form for the DWF:

\[
\text{DWF} = \exp(-Q^2 \langle u^2 \rangle / 6)
\]

(III-32)
If, however, the \( \rho(r) \) for example involving rapid jumps between two or more quasi-equilibrium positions, then the DWF will contain \( Q \)-dependent diffraction-like terms. In any case, the DWF is the Fourier transform of the density distribution of the time-averaged high-frequency motion.

Returning to Eq. (III-30), we see that the quasi-elastic contribution can now be written as:

\[
S_q(Q,\omega) = (\text{DWF})^2 \times \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(i\omega t) \exp(iQ \cdot r) \ G_{\text{slow}}(r,t) \ dr \ dt \quad (\text{III-33})
\]

This term will produce a quasi-elastic line that depends on the character of \( G_{\text{slow}}(r,t) \).

For Brownian diffusion,

\[
G_{\text{slow}}(r,t) = (4\pi Dt)^{-3/2} \exp(-r^2/4Dt) \quad (\text{III-34})
\]

and the quasi elastic spectrum becomes:

\[
S_q(Q,\omega) = \frac{1}{\pi} \frac{\Gamma}{\omega^2 + (\Gamma)^2} \quad (\text{III-35})
\]

with \( \Gamma(Q) = DQ^2 \).

Other types of diffusive motion (e.g. jump diffusion) give similar Lorentzian scattering laws with different forms for \( \Gamma(Q) \), which depend on the assumed distribution of jump lengths\(<52-56>\).

3.4.2. Different jump-length distributions

We will consider some possible forms for the jump distribution \( \rho(r) \)
3.4.2.1. Exponential

\[ \rho(r) = \frac{1}{4\pi r_0^2} \exp\left(-\frac{r}{r_0}\right), \]  
\[ (III-36) \]

\[ F(Q) = \text{FT} \, \rho(r) = \frac{1}{1 + Q^2 r_0^2}, \]

\[ \frac{\Gamma}{\tau} = \frac{1}{\tau} \left( 1 - \frac{1}{1 + Q^2 r_0^2} \right) = \frac{1}{\tau} \left( \frac{Q^2 r_0^2}{1 + Q^2 r_0^2} \right). \]  
\[ (III-37) \]

3.4.2.2. Dirac-delta

\[ \rho(r) = \delta(r - r_0), \]  
\[ (III-38) \]

\[ F(Q) = \frac{1}{Q r_0} \sin(Q r_0), \]

\[ \frac{\Gamma}{\tau} = \frac{1}{\tau} \left( 1 - \frac{1}{Q r_0} \sin(Q r_0) \right). \]  
\[ (III-39) \]

In all cases, the width is \( \sim Q^2 \) (continuous diffusion) for \( Q r_0 \ll 1 \), and \( \Gamma \) approaches constant for \( Q r_0 >> 1 \).
Chapter IV

Spectrometer Resolution, Background, and Multiple Scattering

Before the data is analyzed, a few corrections must be done on the raw data. These corrections include the effects of the spectrometer resolution, the inelastic background and the D$_2$O background whenever it is appropriate to do so. The multiple scattering effect is also mentioned, we however did not make correction but only minimize this effect.

4.1. Spectrometer Resolution Function

One of the data corrections is the effect of the spectrometer resolution. This problem arises from the fact that the signal being detected is averaged over the geometry of the monochromator, collimator, and detector. Ideally, the instrument should have a source-monochromator-collimator combination which delivers perfectly monochromatic, infinitely collimated radiation which is directed through an infinitesimally thin specimen. The scattered neutrons then are registered by a detector with an infinitesimally small acceptance aperture. In reality, the beam has finite size, and even an ideal detector will be seeing scattered radiation over a range of angles. A less-than-perfect monochromatization of the beam has a similar effect. In all, the signal measured at scattering angle $\theta$ will not be $I(\theta)$ but some kind of weighted average over a range of angles $\theta \pm \Delta$. The signal measured is an integrated version of the desired signals. Mathematically, the effect of smearing the signal can be thought of as a convolution$^{57-59}$.

$I_0 (Q, \omega)$, the measured signal, is given as follows:$^{60}$

\[
I_0 (Q, \omega) = \int_{-Q'}^{Q'} dQ_0 \int_{-\omega'}^{\omega'} d\omega_0 \ S_i (Q-Q_0, \omega-\omega_0) \ R (Q_0, \omega_0), \quad \text{(IV-1)}
\]
where \( R(Q, \omega) \) is the resolution function in \( (Q, \omega) \) space, and \( S_1(Q, \omega) \) is the ideal signal.

4.1.1. General Spectrometer Resolution Function

![Diagram of a neutron spectrometer](image)

Fig. IV-1. Layout of a typical triple-axes spectrometer.
The sketch of a typical triple-axes spectrometer is shown in Fig. IV-1. The scattering angles 2M, 2S, and 2A are the those of the monochromator, sample, and analyzer respectively. <61,62>

4.1.2. Approximated Form of the Spectrometer Resolution

In our spectrometer, the width of the Q-resolution function is much smaller than any feature in a constant \(\omega\) scan, except in the case of a Bragg peak. Thus for constant Q, the scattering law is averaged over a narrow range of Q-space in which \(S_i (Q, \omega)\) is almost constant. So in this approximation, it can be approximated as a function of Q

\[
R (Q, \omega) \approx \delta (Q) R (\omega). \quad (IV-3)
\]

Then the equation (IV-1) becomes

\[
I_0 (Q,\omega) = \int_{-\infty}^{\infty} d\omega_0 \ S_i (Q, \omega - \omega_0) \ R (\omega_0) . \quad (IV-4)
\]

For convenience, the limits in (IV-4) are usually taken to be \(\pm \infty\).

Our resolution function was measured by scanning the elastic incoherent peak from vanadium. It should be noted that the resolution is not symmetric about \(\omega = 0\), due to the fact that the final wave vector \(k_f\) is varied<61,62>. Taking the variation of \(k_f\) into account, the \(R(\omega)\) has approximately a Gaussian form in \(\omega\) multiplied by a skew factor, <63>

\[
R (\omega) \propto (A - B\omega) \exp \left( -\frac{\omega^2}{K^2} \right) , \text{ where } K = \frac{\Gamma r}{2\sqrt{\ln 2}} \quad (IV-5)
\]
\( \Gamma_f \) is the resolution full-width-half-maximum, A and B are arbitrary constants.

### 4.1.2.1 The Voigt Function

Using the (IV-5) resolution function, we can write \( I_0(Q, \omega) \) explicitly as,

\[
I_0(Q, \omega) = \frac{1}{K \sqrt{\pi}} \int_{-\omega}^{\omega} d\omega_0 \ S_i(Q, \omega - \omega_0)(A - B\omega_0) \exp\left(-\frac{\omega_0^2}{K^2}\right). \tag{IV-6}
\]

In some models used to describe the dynamics of protons in solution, the shape of \( S_i(Q, \omega) \) is Lorentzian, including the model we used to describe our system. So \( S_i(Q, \omega) \) takes the form

\[
S_i(Q, \omega) = \frac{1}{\pi \hbar} \exp\left(-\frac{Q^2 <u^2>}{3}\right) \frac{\Gamma/2}{(\Gamma/2)^2+(\omega-\omega_0)^2}. \tag{IV-7}
\]

Then the experimental spectrum is:

\[
I_0(Q, \omega) = \frac{1}{K \hbar \pi^{3/2}} \int_{-\omega}^{\omega} d\omega_0 (A - B\omega_0) \exp\left(-\frac{\omega_0^2}{K^2}\right) \exp\left(-\frac{Q^2 <u^2>}{3}\right) \frac{\Gamma/2}{(\Gamma/2)^2+(\omega-\omega_0)^2}. \tag{IV-8}
\]

Using the definition of the complex Voigt function,

\[
V(z) = \exp(-z^2) \left(1 - \text{erf}(-iz)\right), \tag{IV-9}
\]

\[
\text{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z dt \exp(-t^2).
\]

where \( \text{erf} \) is the error function, we have
\[
\int_{-\infty}^{\infty} \delta t \frac{y \exp \left( -t^2 \right)}{(x - t)^2 + y^2} = \pi \, \Re \{ V(x + iy) \} \quad (x \text{ is real}, \, y > 0 ). \quad \text{(IV-10)}
\]

It can be shown that \( I_0(Q,\omega) \) can be approximately written as,

\[
I_0(Q,\omega) = \frac{1}{K\hbar \nu \pi} \exp \left( -Q^2 \frac{\langle u^2 \rangle}{3} \right) [1 - C\omega] \, \Re \left\{ V\left( \frac{\omega}{K} + \frac{\Gamma(Q)}{2K} \right) \right\}, \quad \text{(IV-11)}
\]

where \( C \approx 0.02 \), a small correction, so it is permissible to assume the resolution function is an ideal Gaussian in which \( I_0(Q,\omega) \) has the form

\[
I_0(Q,\omega) = \frac{1}{K\hbar \nu \pi} \exp \left( -Q^2 \frac{\langle u^2 \rangle}{3} \right) \Re \left\{ V\left( \frac{\omega}{K} + \frac{\Gamma(Q)}{2K} \right) \right\}. \quad \text{(IV-12)}
\]

As shown in my MA thesis\(^{<68>}\), the fits done with both forms of resolution functions (eq. IV-5 and a Gaussian) are approximately of equal quality, within the error bars. The numerical values of the Voigt function can be found by Taylor series expansion, or by asymptotic series expansion, or by numerical integration.

4.1.2.2. Fitting Techniques to Extract the Width and Intensity

There are more than one way to extract the width and the intensity from our experimental data. We have chosen the convolution technique, but will briefly discuss other possible techniques.

The first technique involves the deconvolution of the experimental spectrum to obtain the scattering law. In principle when one of the two contributions to the experimental spectrum is known, e.g. the resolution function, then the scattering law can be obtained
through Fourier deconvolution. This procedure can go wrong if the transform of the response function is zero for some values of $S_\eta$ - the values for the experimental spectrum. This procedure is also quite sensitive to noise in the input data, and to the accuracy of the known response function. It requires excellent statistics in the data set for success. So a perfectly reasonable deconvolution procedure sometimes produces nonsense.

The second approach involves the approximation of the Voigt function using a linear combination of Gaussian and Lorentzian curves. The approximation can be written as

$$ F(\Gamma, \eta, x) = \frac{\eta}{1 + \left( \frac{x - x_0}{0.5\Gamma} \right)^2} + (1 - \eta) \exp\left( -\frac{x - x_0}{b\Gamma} \right)^2. $$

(IV-13)

In the study of hydrated proteins, \textsuperscript{69} this technique has been applied to extract the linewidth from a Voigt function. In general, this method is less exact except in a certain special case and is rarely used.\textsuperscript{64}

The third method involves the comparison the experimental spectrum with the convolution of the assumed scattering law and the spectrometer resolution function (a gaussian in our case)\textsuperscript{64}. The scattering laws being fitted are varied depending on whether the sample is powder, frozen solution, or solution. These scattering laws are mentioned in chapter V. This method is more advantageous, because the problem of deconvolution can be avoid as in the first method. We had chosen this method to analyze our data.

4.2. Inelastic Scattering

In addition to quasi-elastic scattering near $\omega=0$ in which we are interested, there is a contribution to the experimental spectrum near $\omega=0$ from the inelastic scattering from the sample. We call this the inelastic background. In our experiment with the sample, we hold
Q constant and scan the quasi-elastic peak by varying $\omega$. In this case we find that the inelastic background is well represented by a simple linear function,

$$I_{\text{inel. bg.}} = A + B\omega.$$  \hfill (IV-14)

In practice, when second term $B\omega$ was left out, the change on the fitting parameters were negligible; thus we decide to use only $A$ for the inelastic background. This indicates that the inelastic background in our region of interest, i.e. the quasielastic region, can be approximated by a constant background.

4.3. **D$_2$O Background**

In order to observe the dynamics of protons associated with the biomolecules, we placed them in a D$_2$O solution. The scattering from the D$_2$O in the solution also requires some corrections to the data. For small $Q$, it has been shown that the scattering function for coherent scattering (D$_2$O) can be approximated as a Lorenztian with $Q$-dependent width.<sup>71</sup> For D$_2$O, the coherent cross section is 80% of the total scattering cross section, so in the fitting, we require a Lorentzian for the D$_2$O contribution. The fitting parameters were taken from a separate fit on the D$_2$O-only scans and these parameters were held fixed in the protein + D$_2$O scans fitting.

4.4. **Multiple Scattering**

There is also another effect that makes the data analysis more difficult, which is the effect of multiple scattering. Realistically, the scattering law should include second and higher-order scattering-terms. However, the multiple scattering effect is complicated. It is
advantageous if the effect can be avoided altogether; as in our case we minimize the effect of multiple scattering by employing samples with high transmission, i.e. the thickness of the sample is small (≈1.5 mm). The transmission rates for the samples were measured. These values range from .87 to .90 for solution and powder respectively.
Chapter V  
Data Analysis

After the data correction and fitting, the analysis of the neutron scattering spectra focuses on the integrated intensity (or the Debye-Waller factor) which reveals the characteristic of vibrational motion; and the linewidth which describes the diffusive behavior of the system under study. In the case of protein powder and protein in solution at low temperature, the only quantity of interest is the integrated intensity. In the case of protein in solution and in crystals at room temperature, both the integrated intensity and the linewidth are analyzed.

5.1. **Data Correction**

5.1.1. **Inelastic and D$_2$O Background Correction**

The background scans on the D$_2$O were done to allow proper extraction of the real lineshape of the scattering by the protein molecule. We fit the D$_2$O background with a Lorentzian as discussed in chapter IV. The parameters obtained from the D$_2$O-background fitting were used as fixed parameters in the analysis of protein-D$_2$O spectra. The exact form of the total scattering cross section of protein-D$_2$O spectra is as follows,

\[
S(Q,\omega) = A + L_{D_2O}(Q,\omega) \otimes R(Q,\Gamma_{res}) + L_{protein}(Q,\omega) \otimes R(Q,\Gamma_{res}) \quad (V-1)
\]

where $L_{D_2O}(Q,\omega)$ is the Lorentzian representing the dynamical behavior of the D$_2$O; $L_{protein}(Q,\omega)$ is the Lorentzian representing the dynamical behavior of the protein; $R(Q,\omega)$ is the general form of the resolution function of the spectrometer; and the inelastic
background is well approximated by the constant $A$. Higher order terms could be included for the inelastic background, but for the limited range that we measured, this approximation with $A$ is adequate.

The $S(Q)$ spectra for $D_2O$ show a $Q$-dependence behavior. It decreases between $0.5-1.5 \, \text{Å}^{-1}$ and increases again around $2.0 \, \text{Å}^{-1}$ (see Fig. V-1a). The physical origin of this peak at $\approx 2.0 \, \text{Å}^{-1}$ is the coherent effect due to the atomic spacing $r_0$ ($Q=2\pi/r_0$) of neighboring molecules of the liquid. The raw $S(Q,\omega)$ for $D_2O$ spectra are shown in Fig. V-1b. These spectra are Lorentzians and their amplitudes are much weaker compared to protein-solution spectra. The $D_2O$ spectra is more pronounced at small $Q$ (=0.5 \, \text{Å}^{-1}) and become flaten at larger $Q$, except for $Q$ around $2.0 \, \text{Å}^{-1}$.

5.1.2. The Elastic Peak.

As indicated in chapter III, the elastic incoherent structure factor (EISF) arises due to diffusion in restricted geometry. It has been shown that the scattering laws for diffusion in these cases can be separated into two parts:

$$S_{\text{inc}}(Q,\omega) = I_{\text{inc}}(Q,\infty)\delta(\omega) + S_{\text{inc}}^Q(Q,\omega). \quad (V-2)$$

The $S_{\text{inc}}(Q,\omega)$ has an elastic part $I_{\text{inc}}(Q,\infty)\delta(\omega)$, and a broadening quasielastic part $S_{\text{inc}}^Q(Q,\omega)$. The linewidth of this quasielastic term reveals information about the characteristic times and space of the motion. On the other hand, the EISF reflects the time-averaged spatial distribution of the moving molecules in restricted geometry at $t = \infty <72>$. The EISF is a significant and a measurable parameter with powder proteins, protein with small hydration levels, low-temperature protein-solution, and hydrated membranes. These systems indicate that there are hydrogens whose motions are restricted in space. In systems
Fig. V-1a. Integrated intensity $S(Q)$ for $\text{D}_2\text{O}$ solution.
Fig. V-1b. D$_2$O background spectra for 300 K.
with dynamical translational disorder, the elastic peak does not appear. Such systems include hydrogens in metals, liquids etc. \textsuperscript{72}. A similar study of protein myoglobin in solution by M. Karplus, et. al.\textsuperscript{73} also observed no EISF component. For protein in solution at a comparable concentration, they found that the scattering cross section to be purely quasielastic\textsuperscript{73}.

Another factor that plays a crucial role in the measurement of the EISF is the effect of the spectrometer resolution width. In the scattering processes, the elastic peak \(I_{\text{el}}(Q, \omega)\delta(\omega)\) is convoluted with the spectrometer's resolution function, which in our cases can be approximated by gaussians.\textsuperscript{74,75}

5.1.3. Simulated Results on the Effects of Elastic Peak

Since in all our experiments, we could not resolve the elastic peak in protein-solution at 280K and 300K nor in protein crystals, we try to estimate the errors caused by the assumed presence of the elastic peak in the quasi-elastic spectrum through simulated data. This will give us a rough estimate on the error if we know the amplitude of the elastic peak.

The estimation of the errors introduced by an elastic peak (or the elastic incoherent structure factor – EISF) to a quasi-elastic spectrum is a simple-minded approach. First, the experimental scattering intensity was simulated with a Lorentzian and a delta-function. The Lorentzian represents the quasi-elastic intensity, which provides both time constants and spatial information; and the delta-function or the elastic scattering intensity \(\delta(\omega)\) represents the time-averaged spatial distribution of the moving molecule. The parameters used in the simulation are comparable to those obtained from the experimental data.
This simulated spectra were then convoluted with the spectrometer resolution function to produce the "real" experimental scattering cross sections and in a certain approximation have the form,

\[ S(Q,\omega) = \delta(\omega) \otimes R(Q,\Gamma_{\text{res}}) + L(Q,\omega) \otimes R(Q,\Gamma_{\text{res}}), \quad (V-3) \]

where \( \otimes \) is the usual notation for the convolution; \( R(Q,\Gamma_{\text{res}}) \) is the resolution function with resolution width \( \Gamma_{\text{res}} \); \( \delta(\omega) \) is the EISF which describes the spatial distribution of restricted molecules; and \( L(Q,\omega) \) represents the jump-diffusive mechanism. To test the reliability of the fitting functions, we fitted the simulated data with the same function in (V-3) and obtained the original parameters within 0.1% error.

The next step is to fit the simulated data (= a Lorentzian + a EISF component) with only a Lorentzian (i.e. \( L(Q,\omega) \)) and leave out the EISF component \( \delta(\omega) \). The results are shown in Fig. V-2a & Fig. V-2b. The abscissa denotes the ratio of the effective areas of elastic component \( \delta(\omega) \otimes R(Q,\Gamma_{\text{res}}) \approx \) a gaussian) to that of the quasi-elastic component \( L(Q,\omega) \otimes R(Q,\Gamma_{\text{res}}) \approx \) a Lorentzian). The ordinates in Fig. V-2a & Fig. V-2b represent the fractional difference in the fitted amplitude and the width respectively, as compared to the simulated amplitude and width. The results show that if the area of the EISF is comparable or less than the Lorentzian, then the error in the amplitude is equal or less than 10%, while the error in the width is also about 10% or less.

Before we proceed to analyze the protein-solution data at ambient temperature, we need to assume that the EISF is small compared to the quasi-elastic component; thus the error that we make, by modeling the scattering cross section with only the Lorentzian or the quasi-elastic component, would also be small as shown in the simulated-fitting results.
Fig. V-2a. Results of fitting simulated data with only a lorentzian.
Fig. V-2b. Results of fitting simulated data with only a lorentzian.
5.2. **Experimental Results**

In general, the vibrational motion of the proteins under study agrees with the harmonic model – as indicated by the study of the integrated intensity. The harmonic model assumes that the probability of finding an atom about its equilibrium position is a gaussian. Furthermore, the diffusional behavior of the protein can be described by the jump-diffusion model with a gaussian jump-length distribution – as indicated by the study of linewidth vs. Q. We also were able to extract the self-diffusion constant $D_s$ which quantitatively describes the diffusional behavior of the side chains' hydrogens.

5.2.1. **Protein Powder**

The trypsin and myoglobin powders were lyophilized, freeze-dried in D$_2$O to exchange loosely bound hydrogens on the protein with about .086 g of D$_2$O/g of protein. This method has been described in chapter II. The scans of $S(Q,\omega)$ were taken at various Q's around the quasielastic region. Fig. V-3a shows a typical scattering spectra that consists only detectable elastic peak for protein powder. The fitting of these gaussian spectra (see Fig. V-3a) indicates that the scattering is purely elastic with no detectable quasi-elastic line – consisting only of a delta-function convoluted with the spectrometer resolution.

Between 200K-320K, there are generally two models to be considered for the elastic intensity. The first model is the non-gaussian two-states model, i.e. the molecule is hopping between two allowed states $<76>$:

$$S(Q) = e^{-Q^2 \Delta x^2} \left[ 1 - 2p_1 p_2 \left( 1 - \frac{\sin(Qd)}{Qd} \right) \right]$$  \hspace{1cm} (V-4)
Fig. V-3a. Spectra of powder protein for various Q at 300K (raw data).
The term $e^{-Q^2<\Delta X^2>}$ represents the well-known harmonic approximation of the Debye-Waller factor. In terms of the mean-square displacement $<\Delta X^2>$, it has a gaussian form. The second term in bracket is the elastic incoherent structure factor of the two-state model. The probability of finding a hydrogen in the ground or excited state are denoted by variables $p_1$ and $p_2$ respectively. However, due to limit $Q^2$-range in our triple-axes spectrometer from 0.25 to 6.25 Å$^2$, our data can not differentiate between the non-gaussian behavior and gaussian behavior. Also as indicated by others studies, it is sufficient to invoke the harmonic model to extract the mean-square displacement $<\Delta X^2>$<76,77>.

We proceeded with the harmonic DWF model,

$$S(Q) = e^{-Q^2<\Delta X^2>}.$$ (V-5)

The results are shown in Fig. V-3b for protein powder at 320 K. The measurement in Fig. V-3b was done with the spectrometer's mode set at $\omega=0$ (elastic scan) for various momentum transfer $Q$'s. It is then fitted with the harmonic model. The result show that for protein powder at 320K the mean-square displacement is $<\Delta X^2>=0.139 \pm .007$ Å$^2$.

5.2.2. Protein in Solution

5.2.2.1. Results for Low Temperatures

The scattering cross sections, $S(Q,\omega)$ of trypsin and myoglobin in solution between 100-200K, exhibit characteristics similar to powder protein (see Fig. V-3a), i.e. the scattering consists of only an elastic peak. The widths of these spectra are independent of Q (equal to the resolution width of the spectrometer $\approx .08$ meV). Furthermore, the lineshape of these spectra can be approximated by gaussians. Fig. V-4a shows the raw
Fig. V-3b. Elastic scattering spectrum of powder protein at 320K.

\[
S(Q) = \exp(-\langle \Delta x^2 \rangle Q^2)
\]

<table>
<thead>
<tr>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.139</td>
<td>0.007</td>
</tr>
</tbody>
</table>
Fig. V-4a. Spectra of trypsin-solution for various Q at 200K (raw data).
scattering cross sections of the protein-solution sample between 0.5–2.5 Å⁻¹ at 200K. These spectra characteristics similar to protein powder, i.e. purely elastic scattering with gaussian lineshapes.

The elastic scattering cross sections S(Q) were measured at 100K, 150K and 200K. Here, we assume that at these low temperatures the harmonic description of the molecule is adequate, since in the frozen solution the motion of molecules are mainly high-frequency vibrations. The elastic scattering cross section S(Q) is related to the mean-square displacement <ΔX²> of the molecules through the harmonic approximation above. Here <ΔX²> is proportional to the temperature.

The fitting results are shown in Fig. V-4b for S(Q) at 200 K. The mean-square displacements <ΔX²> are 0.055 Å², 0.06 Å², and 0.072 Å², for 100 K, 150 K and 200 K respectively.

5.2.2.2. Results at 280K and 300K

The raw spectra of protein in D₂O solution at various Q's are shown in Fig. V-5a & Fig. V-5b for 280K and 300K respectively. Compared to powder and frozen-solution, these solution spectra exhibit more broadening at the wings. These are indications of quasi-elastic scattering, which are fitted with Lorentzian.

At temperature 280K, the linewidth vs. Q and the fitting models are shown in Fig. V-6a & Fig. V-6b. The linewidths increase and levels off at large values of Q. This characteristic is common for jump diffusion models. We used, in Fig. V-6b, the jump diffusion model described in section 3.4.2.1 with the gaussian jump-length distribution

\[ p(r) = \frac{1}{4\pi r_0^2} \exp\left(-\frac{r}{r_0}\right). \]
Fig. V-4b. Elastic scattering spectrum of protein-solution at 200K.

\[ S(Q) = \exp(-<\Delta x^2>Q^2) \]

<table>
<thead>
<tr>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.072</td>
<td>0.004</td>
</tr>
</tbody>
</table>
Fig. V-5a. Spectra of myoglobin-solution for various Q at 280K (raw data).
Fig. V-5b. Spectra of myoglobin-solution at various Q at 300K (raw data).
Fig. V-6a. Myoglobin-solution at 280K.

![Graph showing linewidth versus Q(Å⁻¹) with data points and error bars.]

<table>
<thead>
<tr>
<th>Width = τ⁻¹(1-sin(Qr₀)/(Qr₀)</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>τ⁻¹</td>
<td>0.050</td>
<td>0.004</td>
</tr>
<tr>
<td>r₀</td>
<td>2.086</td>
<td>0.157</td>
</tr>
</tbody>
</table>

single jump length
Fig. V-6b. Myoglobin-solution at 280K.

\[ \text{Width} = 2Q^2D(1+Q^2Dt) \]

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.023</td>
<td>0.004</td>
</tr>
<tr>
<td>Dt</td>
<td>0.572</td>
<td>0.154</td>
</tr>
</tbody>
</table>

gaussian jump length distribution
\[ \Gamma(Q) = \frac{2Q^2D}{1 + Q^2Dt_0} \quad (V-6) \]

The model in Fig. V-6a is also a jump diffusion model but with a discrete jump length distribution \( \rho(r) = \delta(r - r_0) \) as discussed in section 3.4.2.2,

\[ \Gamma(Q) = \left( 1 - \frac{1}{Qr_0} \sin(Qr_0) \right) \quad (V-7) \]

which often is used to describe the motions in an order systems such as hydrogens diffusion in metals. The fitting results and functions are shown in the boxes on the figures.

Also at 280K, the integrated intensity vs. Q is shown in Fig. V-7. The integrated intensity varies inversely with Q. The fitting using the harmonic Debye-Waller factor is satisfactory.

At 300K, the linewidth vs. Q and the fitting are shown in Fig. V-8a & Fig. V-8b. The models used at 300K is similar to those in 280K.

The integrated intensity vs. Q for 300K, shown in Fig. V-9, is also fitted with the harmonic model.

### 5.2.3. Protein in Crystals

The analysis of crystal data is more difficult than we had originally anticipated. The first problem encountered is the absence of an elastic peak whose origin is discussed earlier in chapter III. The second problem involves the measurement of the exact amount of the exchangable D_2O in the crystals. These problems make a meaningful interpretation of the crystals data difficult. We proceeded to extract the linewidth and integrated intensity but did not analyze the data with the jump-diffusion model.
Fig. V-7. Myoglobin-solution at 280K.

\[ S(Q) = A \exp(-<\Delta X^2>_Q^2) \]

<table>
<thead>
<tr>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>154.27</td>
</tr>
<tr>
<td>( &lt;\Delta X^2&gt; )</td>
<td>0.15</td>
</tr>
</tbody>
</table>
Fig. V-8a. Myoglobin-solution at 300K.

Line Width (meV)

Q (Å⁻¹)

Width = t⁻¹(1 - sin(Qr₀) / (Qr₀))

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>t⁻¹</td>
<td>0.057</td>
<td>0.002</td>
</tr>
<tr>
<td>r₀</td>
<td>2.175</td>
<td>0.081</td>
</tr>
</tbody>
</table>

single jump length
Fig. V-8b. Myoglobin-solution at 300K.

Width = $2Q^2D/(1+Q^2Dt)$

<table>
<thead>
<tr>
<th></th>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>D</td>
<td>0.0322</td>
<td>0.002</td>
</tr>
<tr>
<td>Dt</td>
<td>0.794</td>
<td>0.078</td>
</tr>
</tbody>
</table>

gaussian jump length distribution
Fig. V-9. Myoglobin-solution at 300K.

\[ S(Q) = A \exp(-\langle \Delta x^2 \rangle Q^2) \]

<table>
<thead>
<tr>
<th>Value</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>142.058</td>
</tr>
<tr>
<td>\langle \Delta x^2 \rangle</td>
<td>0.182</td>
</tr>
</tbody>
</table>

harmonic model
The raw data for protein in crystals is shown in Fig. V-10a. The integrated intensity vs. Q in Fig. V-10b. exhibits the characteristic as seen in the solution data, i.e. the appearance of the $D_2O$ peak at about $Q \approx 2\text{Å}^{-1}$. The determination of $D_2O$ contribution in the crystals data is less certain. Thus proper extraction of $D_2O$ background cannot be done without some guess work. In Fig. V-10c, the width vs. Q plot also indicates the leveling off at high Q's.
Fig. V-10a. Spectra of protein crystals for various Q at 300K.
Fig. V-10b. Trypsin crystals at 300K.
Fig. V-10c. Width vs. Q for trypsin crystals.
5.3. **Discussion**

The theoretical picture of motions are discussed in an earlier section. Basically, the self-correlation function $G_2(r,t)$ is due to two types of motion: the vibrational fluctuation and the jump-diffusion. These two motions are assumed to be uncorrelated.

5.3.1. **Dynamical Information from the Integrated Intensity**

5.3.1.1. **Protein Powder.**

The vibrational motion is proportional to the Debye-Waller factor. To be more specific, the DW factor gives the average mean-square vibration amplitude for the hydrogens associated with the protein. The basic assumption is contained in the following equation:

$$S(Q) \approx e^{-Q^2<\Delta x^2>}. \quad (V-8)$$

Here, the probability of finding an atom a distance $x$ from its equilibrium position due to its high-frequency thermal motion is isotropic and has a gaussian distribution.

In general, two models for the Debye-Waller factors are being considered for the fast internal motions of proteins (see Fig. V-11). In one model (left), the fluctuations are assumed to occur within a single well that can be approximate as harmonic. The other model (right) assumes multi-minima potential, with the internal motions corresponding to a superposition of oscillations within wells and transitions between them <78>.
Fig. V-11. Model potentials for protein motions. Representations of harmonic potential (left), and multi-minimum potential (right).

Experimental data of powder protein have been interpreted using both pictures, and it is difficult to differentiate the two models in our limited Q-range. However, it also has been pointed out that the use of non-harmonic dynamics to interpret the data is unnecessary, at least for neutron scattering studies. In addition, the non-harmonic models, e.g. the two-state model shown in eq. V-5, generally consists of more free parameters, so it is less appealing to use than the harmonic model. Thus, it is adequate to use the harmonic approximation in our data.

The data for temperature dependence of the mean-square displacement $\langle \Delta X^2 \rangle$ of protein in solution and in dry powder protein are shown in Fig. V-12a & b. For the result of dry protein, it is not clear if there is a change in the mean-square displacement $\langle \Delta X^2 \rangle$ as the temperature increases to above 200K as suggested by other studies shown in Fig. V-12b. One of the reasons is that our protein powder is much drier with 0.086 g solvent / 1.0 g of protein compared to others' 0.33 g solvent / 1.0 g of protein. On a separate study using neutron scattering technique, the onset of $\langle \Delta X^2 \rangle$ at 200K was also not clearly observed due to low hydration level.
Fig. V-12a. Thermal vibrations at different temperature for Mb in solution.
Fig. V-12b. Different studies of thermal vibrations of proteins.

- ○ 0.38 g H₂O/g protein [others]
- □ powder 0.08 g D₂O/g protein [ours]
- ◇ MD [others]
- ✗ solution 5 g D₂O/g protein [ours]
5.3.1.2. Protein in Solution.

At temperature below 200 K, the elastic scattering intensity of the protein-solution has the form as expected for a system with scatterers in a harmonic potential whose vibrational motion can be described by a harmonic Debye-Waller function \( S(Q) \approx e^{-Q^2\langle \Delta x^2 \rangle} \). These features of the Debye-Waller factor are due to motions of protein's localized thermal oscillations of hydrogens confined by D\(_2\)O ice. Here, the protein molecules are trapped in localized regions of conformational space and their diffusive degree of freedom is also "frozen out". In these temperatures' ranges (below 200 K), the mean-square displacement \( \langle \Delta x^2 \rangle \) is directly proportional to the temperature as shown in Fig. V-12a.

The temperature dependence of \( \langle \Delta x^2 \rangle \) can provide us with additional interesting information if we assume that each atom moves in a conformational potential, \( V(x) \). Supposed \( V_\nu(x) \) has this form,

\[
V_\nu(x) \propto x^{1/\nu}, \quad \text{for } x > 0. \hspace{1cm} (V-9)
\]

Based on a Boltzmann distribution, substates at distance \( x \) and temperature \( T \) are populated according to,

\[
P_\nu(x, T) \propto x^2 \exp(-\text{const. } x^{1/\nu}/T). \hspace{1cm} (V-10)
\]

The mean-square displacement is given by

\[
\langle \Delta x^2 \rangle_\nu = \frac{\int dx \ x^2 \ P_\nu(x, T)}{\int dx \ P_\nu(x, T)} = \text{const. } T^{2\nu}, \hspace{1cm} (V-11)
\]
where the constant does not depend on \( T \).

The temperature dependence of \( \langle \Delta X^2 \rangle \), above 200K (the so-called transition temperature \( T_t \)) and below, indicates that \( v = 0.5 \) is satisfactory, i.e. \( \langle \Delta X^2 \rangle \) is linearly depended on \( T \). This result is consistent with our previous assumption that the potential is approximately harmonic. Furthermore, using the energy equipartition principle,

\[
\frac{1}{2} K_{\text{eff}} \langle u^2 \rangle = \frac{3}{2} k_B T \quad (V-12)
\]

(note: \( \langle \Delta X^2 \rangle = \frac{\langle u^2 \rangle}{3} \)).

In this simplified picture, \( K_{\text{eff}} \) is thought as the average effective spring constant of the interactions which keep the protein's molecules in their equilibrium positions. The \( K_{\text{eff}} \) for frozen, powder, and solution are 0.071, 0.043, and 0.008 (10^5 \text{ dyne cm}^{-1}) respectively. Here, due to the weaker spring constant \( K_{\text{eff}} \), the molecules in solution can have more freedom to move rapidly from one substate to another at room temperature above the \( T_t \). \(^{81}\)

At higher temperatures, specifically 280 K and 300 K, a presumably new degree of motional freedom emerges that caused the detectable quasi-elastic scattering intensities – an indication of transition between various conformational substates – (the hydrogen bonds can occupy various positions; the side chains of the amino acids can be in a number of different positions; and consequently a given protein can take up slightly different overall structures, which we call conformational substates). For a protein system of \( N \) components (amino acids) and if each of the \( N \) amino acids can assume roughly 2 different configurations of equal energy, then there are approximately \( 2^N \) states with energies close to that of ground states, for example a trypsin with 223 amino acids can have states of \( 2^{223} \approx 10^{67} \).

We recall the often used definition of \( \langle \Delta X^2 \rangle \) in X-ray methods,
\[ <\Delta X^2>_{\text{X-ray}} = <\Delta X^2>_{\nu} + <\Delta X^2>_{\text{cs}} + <\Delta X^2>_{\text{ld}}, \quad (V-13) \]

where \(<\Delta X^2>_{\nu}\) represents the contribution from local atomic vibration; \(<\Delta X^2>_{\text{cs}}\) gives the contribution of the transitions between conformational substates; \(<\Delta X^2>_{\text{ld}}\) is due to lattice disorder\(^{80}\). In protein-solution study, we can neglect the \(<\Delta X^2>_{\text{ld}}\), which represents the smearing of the atomic position due to lattice disorders for protein in crystalline form; thus is not relevant for incoherent neutron scattering since each scatterer contributes to the signal independently of the others. Hence the QNS mean-square displacement has the form,

\[ <\Delta X^2>_{\text{neutron}} = <\Delta X^2>_{\nu} + <\Delta X^2>_{\text{cs}}. \quad (V-14) \]

In Fig. V-12b, there is good agreement between our experiment of protein-solution and other's neutron experiment and MD simulation at temperatures above freezing. In the temperature-range of 100 K–200 K, our \(<\Delta X^2>_{\text{neutron}}\) is slightly larger than that of other techniques.

Closer inspection of our data on \(<\Delta X^2>\) between solution and powder, we can estimate the value of \(<\Delta X^2>_{\text{cs}}\), which gives rise to the hopping between substates, by finding the difference,

\[ <\Delta X^2>_{\text{cs}} = <\Delta X^2>_{\text{solution}} - <\Delta X^2>_{\text{powder}}. \quad (V-15) \]

Here we assume the \(<\Delta X^2>_{\text{powder}}\) is mainly contributed by vibrational motions.

The below table shows the value of \(<\Delta X^2>_{\text{cs}}\) using neutron scattering method and X-ray scattering,
Results of mean-square displacement $<\Delta X^2>_{cs}$ of protein-solution at 300K. \(^8\)

<table>
<thead>
<tr>
<th>Types of atoms</th>
<th>X-ray scattering</th>
<th>Our QNS experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Side chains</td>
<td>0.04 to 0.25 Å²</td>
<td>-</td>
</tr>
<tr>
<td>Hydrogens</td>
<td>-</td>
<td>0.10 Å²</td>
</tr>
</tbody>
</table>

The difference between X-ray and our QNS results is partly due to the locations of hydrogens and non-hydrogens atoms in the protein. The non-hydrogens are generally less mobile since they are often located on the backbone, and their movements usually require other small group of atoms to reposition. While the hydrogens are found in the sidechains that largely are in direct contact with the solvent; thus they have larger motional fluctuations.

5.3.1.3. Protein in Crystals.

We have extract the linewidth and the integrated intensity but did not analyze the crystals data according to protein solution picture due to difficulties as discussed earlier.

5.3.2. Dynamical Information from the Linewidth.

The dynamical information from the linewidth is only obtainable from protein-solution data. One of the useful obtained characteristic is the jump-length distribution of individual molecules, which is the measure of the system randomness. Another useful quantity is the self-diffusion constant $D_s$. 
5.3.2.1. **Protein in Solution.**

One significant difference between powder and solution protein is the presence of elastic peak broadening in solution due to transition between substates of the protein. This "jump diffusive" effect is represented by the Lorentzian term, as discussed earlier. The data indicate that a gaussian distribution of jump-lengths \( \rho(r) = \frac{1}{4\pi r_0^2} \exp(-\frac{r}{r_0}) \) is more satisfactory than a discrete distribution \( \rho(r) = \delta(r - r_0) \) with single jump-length \( r_0 \).

With regard to the "jump diffusive" mechanism of protein sidechains' hydrogens, several studies have indicated that these motions do exist and occur frequently <78-80>. Our own molecular dynamics simulation for 18 ps on hydrated sperm-whale myoglobin also indicated these jumps between different conformations of the sidechains (see Fig. V-13a, & Fig. V-13b) on the order of picosecond timescale. The magnitude of these jumps sometimes can be as large as few Å as compared to the gaussian jump-length distribution with the mean value of 0.8 Å at 300 K. The simulation is by no mean a quantitative comparision with the experimental results. We were curious to see if in fact the simulation would show these jumps.

The linewidth under discussion is based on the jump-diffusion model:

\[
\Gamma(Q) = \frac{2Q^2D}{1 + Q^2D\tau_0}.
\]  

(V-16)

The slope of the curve at small Q's gives \( \Gamma(Q) = 2Q^2D \), and the leveling off at large Q's gives the residence time \( \Gamma(Q) = 2/\tau_0 \).

At 300 K and 280 K, the effective diffusive constants D for protein's hydrogens are 0.49 and 0.34 \( (10^5 \text{ cm}^2/\text{sec, <10\% error}) \) respectively. The reduced diffusive coefficient
Fig. V-13a. MD simulation of myoglobin for 17ps.

terminal H atom of residue Lys. 34
Fig. V-13b. MD simulation of myoglobin for 17ps.

terminal H atom of residue Asp.44
indicates that the molecules are less mobile as temperature decreases. While the residence times $\tau_0$ of the hydrogens are similar 1.54 and 1.68 ($10^{-11}$ sec, $<10\%$ error) for 300 K and 280 K respectively. For pure water, $D$ is about 2.5 ($10^5$ cm$^2$/sec) and $\tau_0$ 2.0 ($10^{-12}$ sec).

Our results are consistent with other studies, with slightly smaller $D$ at lower temperature, and with smaller $D$ and longer $\tau_0$ values compared to the values for bulk water.
5.4. **Summary.**

The neutron scattering technique can be used to study motions of biomolecules at different temperatures; and in different states: powder, solution and may be crystals. Hence, it allows us to probe dynamical properties related to these states. As in our case, there are dynamical properties deduced from two important experimental quantities: the elastic peak and quasi-elastic bands.

From the elastic peak in powder and solution, the proteins' hydrogens exhibit high frequency vibrations, which can be modeled by the harmonic DW factor $e^{-q^2\langle\Delta x^2\rangle}$. The dependence of mean-square displacement $\langle\Delta x^2\rangle$ on temperature indicates that $\langle\Delta x^2\rangle$ increase with increasing temperature. Furthermore, there is an increase in the mean-square displacement $\langle\Delta x^2\rangle$ at about 200K. Below the temperature 200K, the molecules only vibrate harmonically in local potential wells. Above 200K, there is a steep increase in the values of $\langle\Delta x^2\rangle$, indicating an extra degree of freedom in the motions of protein. This anomalous behavior is likely an attribute of the transition of molecules between conformational substates, due to the loosen up of the surface water as the temperature increases, which in turn allows the protein to be more flexible to hop between substates.

In crystals, the deviation from the harmonic DW factor is caused by the presence of D$_2$O in crystal cells. The usual background subtraction technique does not apply here due to the possibility of a small unexchangable amount of H$_2$O.

The globular protein's motion of hydrogens in solution and in crystals can be modeled by jump-diffusion picture as indicated by the linewidths. The self-correlation function $G_6(r,t)$ of the related proteins' motion in liquid is known from the diffusion equation in the theory section. There are two limiting cases, depending on the scales of the observation times. First, for the times much longer than the time between two diffusion
steps (i.e. small Q), the translational motion here is described by simple continuous diffusion, where the self-diffusion coefficient $D$ obeyed the rule $\Delta E \propto DQ^2$. For short time (i.e. large Q), the characteristics of elementary jump-diffusion steps dominate. Thus, the width approaches an asymptotic value $\Delta E \propto 1/\tau$. The diffusive motions in crystals are significantly reduced as compared to that in solution. While in crystals the molecules spend longer time at their equilibrium sites.

Another important result is that the observed motion here does not depend on the secondary structures of the proteins — myoglobin consisted of mostly alpha-helical structure, whereas trypsin is beta-sheet, yet their observed motions are similar.
References.


