Dynamics Of Biopolymers On Nanomaterials Studied By Quasielastic Neutron Scattering And Mdsimulations

Gurpreet Dhindsa
Wayne State University,
DEDICATION

‘To my Grand Parents, Parents and loving family’
ACKNOWLEDGEMENTS

It is my great pleasure to work under the guidance and supervision of my advisor Dr. Xiang-Qiang Chu throughout my research period leading to my Ph.D. degree as her first student. She played a major role in helping me to learn a new field of research. This research work would not have been possible without her support. I am fortunate to have her as my dissertation advisor.

I would like to thank our post-doctoral fellow Dr. Debsindhu Bhowmik who helped me learn another technique (MD Simulation) in addition to experimental work. Simulation project would not be possible without his help. I appreciate the time he spent with me while teaching the basics and having other meaningful discussions about our projects.

I would also like to thank Dr. Ratna Naik, not only for being my dissertation committee member but also for all her support and encouragement during my career. I would also like to thank Dr. Zhi-Feng Huang and Dr. Arun Anantharam for being in my dissertation committee.

I would also like to thank Dr. Eugene Mamontov, Dr. Hugh O'Neill and Dr. Madhu Tyagi for helping me in performing the experiments at Oak Ridge National lab and NIST respectively.

I would also like to thank our collaborators Dr. Panchapakesan Ganesh, Dr. Monojoy Goswami who helped us analyzing the simulation data.

I am grateful to my teachers at Wayne State University and all other institutes I have attended for teaching and training in the subject areas of physics.

I am thankful to my husband Kulwinder S. Dhindsa for being my second half, and always stand with me. I am fortunate to have him in my life. During my PhD career, we were also blessed with our little princess, Ramneek Dhindsa, who always refreshes me with her innocent and invaluable smile and small hugs.
I am also thankful to my late grandfather, late father, grandmother and mother for their belief, support and love throughout my life. Their blessings made me to reach at this position. I am also grateful to my younger brother for all his support.
# TABLE OF CONTENTS

Dedication ................................................................................................................................. ii  
Acknowledgements .................................................................................................................. iii  
List of Tables ............................................................................................................................... viii  
List of Figures ............................................................................................................................. ix  

Chapter 1: Introduction ........................................................................................................... 1  
1.1 Neutron Scattering in Biology ......................................................................................... 1  
1.2 Scope of Thesis .................................................................................................................. 2  

Chapter 2: Neutron Scattering ............................................................................................... 3  
2.1 Introduction ....................................................................................................................... 3  
2.2 Coherent and Incoherent Scattering ................................................................................ 4  
2.3 Structure Factor $S(Q,ω)$ and Intermediate Scattering Function (ISF) ....................... 5  
2.4 Atomic Motion .................................................................................................................. 6  
2.5 Neutron Scattering Spectrometers .................................................................................. 8  
2.5.1 Backscattering Spectrometer ...................................................................................... 8  
2.5.1.1 BASIS at ORNL ...................................................................................................... 9  
2.5.1.2 High Flux Backscattering Spectrometer (HFBS) ..................................................... 11  
2.5.2 Neutron Spin Echo Spectroscopy (NSE) .................................................................... 11  
2.5.3 Small Angle Neutron Scattering (SANS) ................................................................... 14  

Chapter 3: Molecular Dynamics Simulations ........................................................................ 16  
3.1 Introduction ....................................................................................................................... 16  
3.2 Classical Equations of motion ......................................................................................... 17
3.3 Force Fields

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.3.1 Bonded Interactions</td>
<td>20</td>
</tr>
<tr>
<td>3.3.2 Non-bonded Interactions</td>
<td>21</td>
</tr>
</tbody>
</table>

3.4 Periodic Boundary Conditions

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
</table>

3.5 Cut off Radius and Ewald Sum

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
</table>

3.6 Statistical Mechanics and Ensembles

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
</table>

3.7 Transport Properties

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7.1 Mean Square Displacement (MSD)</td>
<td>26</td>
</tr>
</tbody>
</table>

Chapter 4: Effect of Nanodiamond on RNA dynamics

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Introduction</td>
<td>27</td>
</tr>
<tr>
<td>4.1.1 Nanoparticle-Biomolecules Complex</td>
<td>27</td>
</tr>
<tr>
<td>4.1.2 Nanomaterials</td>
<td>28</td>
</tr>
<tr>
<td>4.2 Materials and Methods</td>
<td>32</td>
</tr>
<tr>
<td>4.2.1 Neutron Scattering Data Analysis</td>
<td>33</td>
</tr>
<tr>
<td>4.2.2 Simulation Methodology</td>
<td>33</td>
</tr>
<tr>
<td>4.3 Results and Discussion</td>
<td>36</td>
</tr>
<tr>
<td>4.4 Conclusion</td>
<td>52</td>
</tr>
</tbody>
</table>

Chapter 5 Dynamics of an Intrinsically Disordered Protein: β-Casein

<table>
<thead>
<tr>
<th>Subsection</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.1 Introduction</td>
<td>54</td>
</tr>
<tr>
<td>5.2 Materials and Methods</td>
<td>57</td>
</tr>
<tr>
<td>5.3 Results and discussion</td>
<td>59</td>
</tr>
<tr>
<td>5.3.1 Mean Square Displacement (MSD) and Protein Flexibility</td>
<td>59</td>
</tr>
<tr>
<td>5.3.2 Analysis of QENS Data in Energy Domain</td>
<td>61</td>
</tr>
</tbody>
</table>
5.3.3 Time Domain Analysis .................................................................66
5.4 Conclusion ..................................................................................70

Chapter 6: Domain Motion of an Oligomeric Protein IPPase studied by

Neutron Spin Echo (NSE) .................................................................72
6.1 Introduction ................................................................................72
6.2 Materials and Methods ...............................................................74
6.3 Results and Discussion ...............................................................75
6.4 Conclusion ..................................................................................79

Chapter 7 Summary ........................................................................80
References .......................................................................................83
Abstract ..........................................................................................103
Autobiographical Statement .............................................................105
LIST OF TABLES

Table 4.1 Lennard Jones parameters, epsilon (Kcal/mol) and Rmin/2 (Angstrom) are calculated for the Nanodiamond carbon ..........................................................35

Table 5.1 The fitting parameter \( a \) at different temperatures for both dry and hydrated \( \beta \)-Casein ........................................................................................................66
LIST OF FIGURES

Figure 2.1 Scattering triangle in a neutron scattering. $k_i$ and $k_f$ are the incident and emergent wave vectors; $q$ is the scattering vector and $\theta$ is the scattering angle ........................................................................................................4

Figure 2.2 Three types of scattering are used to study the atomic motion. Elastic peak (red color) is represented by delta function, quasi-elastic contains (blue color) translational and rotational motion, and inelastic scattering is used to study very fast motion of atoms such as vibrational motion........................................................................................................7

Figure 2.3 Schematic picture of BASIS tank, white arrows show the path of neutrons. Neutrons scatter from the sample, reflected from the top and bottom analyzers and then detected by the top and bottom Detectors .......................10

Figure 2.4 Schematic set up of NSE. Upper part shows the instrument set up where the neutrons enter the magnetic field coil ($B_1$) having length ($l_1$) perpendicular to the direction of magnetic field. Before neutrons scattered by the sample, they flipped by \(\pi\)-flipper and then pass through the second magnetic field coil ($B_2$). At the end, scattered neutrons flipped by the $\pi/2$ flipper so that they can pass through the analyzers and detected by the detectors. Lower portion represents the spin orientation of neutrons .................................................................................................12

Figure 2.5 Schematic diagram of Small Angle Neutron Scattering (SANS). $k_i$ and $k_s$ are the incident and scattered wave vectors along with the scattering wave vector ($q$) ..............................................................................................................15

Figure 3.1 (a) Covalent bond, (b) Bond angle, (c) Dihedral angle, (d) Improper dihedral angle, (e) Long range Vander Waals, and (f) Coulomb electrostatic interactions ........................................................................................................21

Figure 3.2 Periodic Boundary Conditions are shown in 2-D. The primary cell is in the middle in grey color. One atom is exiting the cell, and its position is replaced by the other one. ........................................................................................................23

Figure 4.1 Root Mean Square Deviation (RMSD) values of the D$_2$O hydrated RNA with and without ND at three different temperatures, 260, 280 and 300 K.................................................................36

Figure 4.2 Comparison of normalized $S(Q,\omega)$ of three different samples ND+tRNA+D$_2$O, ND+tRNA and tRNA+D$_2$O at 300 K and at one $q$-value 1.1Å$^{-1}$ ........................................................................................................38
Figure 4.3 Intermediate Scattering Function (ISF) of ND+tRNA+D2O, tRNA+D2O and tRNA+ND sample is plotted at 300 K and at five different q-values .................................................................39

Figure 4.4 a and b, H₁(Q,T) and B(T) are the fitted parameters of logarithmic model (Eq. 4.6) respectively. H₁(Q,T) is plotted as a function of Q at three different temperatures while B(T) is plotted as a function of temperature ..........................................................................................................................39

Figure 4.5 Mean Square Displacement (MSD) of hydrated and dry tRNA on nanodiamond (ND) surface is plotted as a function of temperature. Dynamical transition (Tₐ) is observed at 230 K only in hydrated sample.......40

Figure 4.6 a and b Experiment and MD simulation Intermediate Scattering Function (ISF) of ND+tRNA+D₂O, ND+tRNA and tRNA+D₂O at one q-value 1.1Å⁻¹ and at three different temperatures 260, 280 and 300 K. Higher decay in the slope of hydrated tRNA on ND surface compared to other samples implies faster dynamics of tRNA in the presence of ND. .........................................................41

Figure 4.7 a and b Experimental relaxation time τ of hydrogen of RNA is plotted as a function of inverse temperature in a system of ND+tRNA+D₂O and tRNA+D₂O system at six different q-values respectively. The cusp shape of the curve shows that ND+tRNA+D₂O follow Vogel Fulcher Tammann law. The straight line shape of the curve shows that tRNA+D₂O system follow Arrhenius law .........................................................................................................43

Figure 4.8 a and b Experimental relaxation time τ of hydrogen of tRNA in ND+tRNA+D₂O and tRNA+D₂O system are plotted as a function of q at eight different temperatures, T = 240 to 310 K with 10 K interval ........................................................................................................................................43

Figure 4.9 a-c Comparison of Stokes-Einstein Relation (SER) D*τ/T Hydrogen of RNA (H_RNA) and Deuterium of Water (D₂O) at three q-values 0.5, 1.1 and 1.9 Å⁻¹Black circles and red squares represent hydrogen of RNA of RNA+ND+D₂O and RNA+D₂O samples; green diamonds and blue up-triangles represent D₂O of water for RNA+ND+D₂O and RNA+D₂O samples respectively. ........................................................................................................43

Figure 4.10 Diffusion coefficient ‘D’ is plotted as a function of τ for Hydrogen of RNA at three different Q values, 0.5 (circle), 1.1 (square) and 1.9 (up triangle) respectively. The black and red symbols represent with
Figure 4.11 Diffusion coefficient ‘D’ is plotted as a function of τ for deuterium of D₂O at three different Q values, 0.5 (circle), 1.1 (square) and 1.9 (up triangle) respectively. The black and red symbols represent with (RNA+D₂O) and without ND (RNA+ND+D₂O) sample respectively ........................................46

Figure 4.12 Scaling exponent, α as a function of Q representing the values at different length scales. The α values are derived from Fig. 4.10 and 4.11 using a non-linear least-squares Marquardt-Lavengberg algorithm. The dashed line at 1.0 represents normal Fickian diffusion ........................................47

Figure 4.13 Left and right images are the simulated snapshots of RNA+D₂O and ND+RNA+D₂O system from MD simulation respectively .........................48

Figure 4.14 Radial Pair distribution function, g(r) (left axis) and number of water molecules (D₂O) from the surface of ND are plotted as a function of distance in ND+RNA+D₂O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules (right axis) for the respective temperatures .................................................................50

Figure 4.15 Radial Pair distribution function, g(r) (left axis) and number of water molecules (D₂O) (right axis) around RNA are plotted as a function of distance in ND+RNA+D₂O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules for the respective temperatures .................................................................50

Figure 4.16 Radial Pair distribution function, g(r) (left axis) and number of hydrogen of RNA around ND are plotted as a function of distance in ND+RNA+D₂O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules (right axis) for the respective temperatures .................................................................51

Figure 4.17 Radial Pair distribution function, g(r) (left axis) and number of water
molecules around RNA are plotted as a function of distance in RNA+D_{2}O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules (right axis) for the respective temperatures.

Figure 5.1 The mean square displacement (MSD), \( \langle x^2(T) \rangle \) of dry β-CN and D_{2}O hydrated β-CN plotted as functions of temperature. The dynamic transition is observed at \( T \approx 225 \) K in hydrated β-CN. Error bars are within the size of symbols.

Figure 5.2 (a) and (b) Normalized QENS spectra of dry and hydrated β-CN at different temperatures from 220 K to 340 K and 260K to 340K respectively at \( Q = 0.9907 \) Å\(^{-1}\). (c) and (d) Analysis of the QENS spectra of dry and hydrated β-CN. The circles are the experimental data, the blue lines represent the fitted curves. The orange and magenta lines represent quasi-elastic scattering component and elastic scattering component respectively. Error bars throughout the text represent one standard deviation.

Figure 5.3 The EISF of (a) hydrated and (b) dry β-Casein at different temperatures fitted according to eq. (5.8). The solid lines are the fitted curves. The error bars of the data points are within the symbols if not shown.

Figure 5.4 The fitting parameter \( p_0 \), representing the fraction of immobile protons, plotted as function of temperature for both dry and hydrated β-Casein.

Figure 5.5 Analysis of the ISF of hydrated β-CN in the β-relaxation region at 10 different Q-values. The upper three panels show the results at lower temperatures \( T = 220, 260, \) and \( 280K \); the lower three panels show at higher temperatures \( T = 300, 320, \) and \( 340 K \). The solid lines represent the fitted curves.

Figure 5.6 (a) Analysis of the fitting parameter \( H_1(Q,T) \) of both samples as a function of Q at different temperatures. For the hydrated sample (solid symbols), \( H_1(Q,T) \) is fitted by a power law, \( H_1(Q,T)= B_1(T)Q^\beta \) at temperatures \( T = 220, 240, 260, 280, 300, 320, 340 K \). For the dry β-CN (open symbols), \( H_1(Q,T) \) does not show much temperature or Q dependence; (b) Comparison of the ISF(Q,t) of hydrated and dry β-CN at three different temperatures, \( T = 300, 320, 340 K \). The open symbols represent data belongs to dry β-CN, while solid symbols belong to hydrated β-CN. The error bars of the data points are within...
the symbols if not shown ........................................................................................................68

Figure 5.7 (a) Linear fitting of the parameter $B_1(T)$ as a function of temperature for hydrating $\beta$-CN. (b) Arrhenius plot of the relaxation time $\tau_\beta(T)$ as a function of inverse of temperature in log scale.................................................................70

Figure 6.1 (a) IPPASE structure with three domains in different colors (green, royal blue and yellow); (b) protein in D$_2$O buffer matched with the number of protein and water during NSE experiment (one IPPASE protein is surrounded by 20833 water molecules) molecules in our system ...............75

Figure 6.2 SANS experimental results on IPPase solutions (a) and (b) Intensity versus $q$ at two temperatures 300 and 320 K and three different concentrations 50, 10 and 5 mg/ml; (c)-(f) fitted data with Guinier law and calculated the $R_g$ at two different temperatures and two concentrations ........................................77

Figure 6.3 NSE data for IPPase (a),(b) $I(Q,t)/I(Q,0)$ on logarithmic scale versus $q$ at 320 and 346 K at eighteen different $q$-values from 0.05-1.6 Å$^{-1}$; Straight lines fit with single exponential function (c) and (d) Diffusion Coefficient (DC) versus $q$ at two temperatures 320 and 346 K .........................................................78
CHAPTER 1 INTRODUCTION

1.1 Neutron Scattering in Biology

Proteins are the essential components of all living organisms. These molecules are responsible for performing a wide range of biological functions [1]. Complete understanding of biomolecular system is not only from their structures, but from their dynamics too. Biomolecules are flexible in nature, and their flexibility is required for them to function [2]. It is widely accepted that proteins comprise a large number of conformational sub-states. Transitions between these states are related to the biological functions of proteins, such as enzyme catalysis [3]. To understand the biological function of proteins, it is essential to study their motions at different conditions, such as hydration, temperature and pressure.

At present, a number of techniques are used to probe soft-matter at atomic level, such as X-ray diffraction, light scattering, Nuclear Magnetic Resonance (NMR), and neutron scattering [4-7]. Among them, neutron scattering has the unique advantage in studying biological systems, which are consisting of large amount of hydrogen atoms, since the hydrogen atoms have much larger incoherent neutron cross-section than other atoms in the biological systems, such as oxygen, carbon and nitrogen. Hence, incoherent neutron scattering spectra reflect only the dynamics of hydrogen atoms in the measured biological systems [8, 9]. Additionally, neutron scattering technique has the ability to cover length scales in the range of few angstroms to hundreds of nanometers. Neutrons exhibit energies in the range of meV (milli-electronvolt), and these low energies are comparable with the excitation energies of the molecular motion. This makes neutron scattering techniques well-suited for the study of the dynamics of biomolecules [10, 11].
In this thesis, when studying the motion of biomolecules, we use D$_2$O as a solvent or in the hydration layer of the biomolecules, instead of H$_2$O, since the incoherent neutron cross section of hydrogen (80 barns) is much larger than deuterium (2 barns). Thus, the incoherent neutron scattering signal will come only from hydrogen atoms within the biomolecules. This enables us to study the dynamics of the biomolecules without removing water from the system.

1.2 Scope of Thesis

The goal of this thesis is to study the dynamics of nucleic acid on nanomaterial and proteins under various conditions. In addition to neutron scattering experiments, we used Molecular Dynamics (MD) simulations to further interpret the experimental results, and to extend data to the time range which cannot be reached by experiments. Neutron scattering technique has special advantage to study biological systems at various length and time scales. The theoretical background and importance of neutron scattering is discussed in Chapter 2. The details of MD simulation are explained in Chapter 3. In our first project (Chapter 4), we approach to understand the effect of nanomaterial (ND) on the dynamics of RNA by quasielastic neutron scattering (QENS) and MD simulations. Studying of this system will provide fundamental knowledge of surface interactions of nanomaterials with biomolecules for their future use in drug delivery. In Chapter 5, our second project studied the role of hydration for the functioning of a natively disordered protein β-casein by quasielastic neutron scattering. To understand the importance of water, we compared the dynamics of hydrated and dry protein samples. In our third project (Chapter 6), we study the domain motion of an oligomeric protein Inorganic Pyrophosphatase (IPPase) at high temperatures by neutron spin echo technique (NSE). Chapter 7 is the summary of the thesis.
CHAPTER 2 NEUTRON SCATTERING

2.1 Introduction

Neutron scattering theory was established in 1954 by Van Hove [12] after neutrons were discovered by James Chadwick in 1932 [13]. This technique can be applied to many fields, including condensed matter physics, material sciences, polymers, chemistry and biology. Neutrons can penetrate into the matter without any destruction and are especially sensitive to hydrogen atoms. These properties make neutron scattering technique one of the most highly demanded techniques in biology, which can study the structure and motion of proteins, nucleic acids, membranes etc. Neutron scattering technique has advantage to explore the motion of atoms and molecules at the nano-scale in length and time domain centered on 1nm and 1ns respectively. This length and time scale occurs in the middle of microscopic and macroscopic scale, which lies in the arena of molecular biology. With neutron spectroscopy, atomic motion can be explored from the $10^{-14}$ to $10^{-6}$ s [14]. This chapter starts with the basic theory of neutron scattering.

When scattered, neutrons exchange energy and momentum with the nuclei of the samples, as shown in Fig. 2.1. The energy and momentum transfers in a neutron scattering event can be written as

$$\hbar \omega = \frac{\hbar^2}{2m} (k_f^2 - k_i^2)$$

$$\hbar q = \hbar (k_f - k_i)$$

(2.1)  (2.2)
Fig. 2.1 Scattering triangle in a neutron scattering event. \( k_i \) and \( k_f \) are the incident and emergent wave vectors; \( q \) is the scattering vector and \( \theta \) is the scattering angle.

The double differential cross-section with respect to change in solid angle \((\Omega, \Omega+\Delta \Omega)\) and energy transfer \((E, E+\Delta E)\) is measured during the neutron scattering experiments. It is written as

\[
\frac{d^2\sigma}{d\Omega dE} = \frac{\text{Flux of neutrons in solid angle } d\Omega \text{ per unit Energy}}{\text{flux of incoming neutrons}} \tag{2.3}
\]

Since neutrons directly interact with the nuclei of a sample, their interaction is very short range. This interaction is characterized by the parameter called scattering length \((b)\). It gives the strength of interactions between neutrons and nuclei. Scattering length is expressed as a real and imaginary quantity \([15]\). It has a different value for different elements and isotopes. Its value varies with spin of atom. The scattering cross section from a single stationary atom is given by \(4\pi b^2\) \([16]\).

### 2.2 Coherent and Incoherent Scattering

There are two types of scattering, one due to the average scattering length \((\bar{b})\) called coherent scattering, and the other due to the fluctuations from the average \((b - \bar{b})\) called incoherent scattering. Coherent scattering occurs when the neutron waves interact with the whole sample as a unit and hence the scattered waves from different nuclei interfere with each other showing interference effects. Its cross-section is proportional to \((\bar{b}^2)\). Incoherent scattering occurs when
the neutron waves interact independently with each nucleus, and the scattered waves do not show any interference effects. Its cross-section is proportional to the \((b - \bar{b})^2\) [8, 15]. The overall scattering is the sum of both coherent and incoherent scattering. It comes as

\[
\frac{d^2 \sigma}{d \Omega dE} = \frac{1}{2 \pi \hbar k_i} \sum_{l,l'} b_{coh,l} b_{coh,l'} \int_{-\infty}^{\infty} dt e^{-i\omega t} < e^{-iQR_l} e^{iQR_l'(t)} + \\
\sum_{l=l'} \left(b_{inc,l} \right)^2 \int_{-\infty}^{\infty} dt e^{-i\omega t} < e^{-iQR_l} e^{iQR_l(t)}
\]

(2.4)

\[
= \frac{k_f \sigma_c}{k_i} S_{coh}(Q,\omega) + \frac{k_f \sigma_i}{k_i} S_{inc}(Q,\omega)
\]

(2.5)

where \(\sigma_c = 4\pi b_{coh,l} b_{coh,l'}\) and \(\sigma_{inc} = 4\pi \left(b_{inc,l}\right)^2\)

\[
S_{coh}(Q,\omega) = \frac{1}{2 \pi \hbar} \sum_{l,l'} \int_{-\infty}^{\infty} dt e^{-i\omega t} < e^{-iQR_l} e^{iQR_l(t)} >
\]

(2.6)

\[
S_{inc}(Q,\omega) = \frac{1}{2 \pi \hbar} \sum_{l} \int_{-\infty}^{\infty} dt e^{-i\omega t} < e^{-iQR_l} e^{iQR_l(t)} >
\]

(2.7)

where \(R_l\) and \(R_l'(t)\) are the positions of different atoms in coherent scattering while \(R_l\) and \(R_l(t)\) represent the position of same atom but at different times in incoherent scattering. Studying the motion of biological samples, the main contribution in the intensity comes from incoherent scattering.

### 2.3 Structure Factor S(Q,ω) and Intermediate Scattering Function (ISF)

\(S(Q,\omega)\) and ISF\((Q,t)\) are interlinked to each other. \(S(Q,\omega)\) represents the motion in energy domain and ISF\((Q,t)\) in time space [6, 16]. These represent the atomic correlation between atoms
positions at different times in real space and time. By taking the inverse Fourier transform of 
$S(Q,\omega)$ (equation 2.6 and 2.7) give ISF(Q,t).

$$I_{\text{coh}}(Q,t)=\sum_{l,l'} <e^{-iQRL}e^{iQRl'(t)}>$$  \hspace{1cm} (2.8)

$$I_{\text{inc}}(Q,t)=\sum_{l} <e^{-iQRl}e^{iQRl(t)}>$$ \hspace{1cm} (2.9)

All of the above equations (2.6-2.9) showed that coherent scattering occurs due to the 
correlation between the positions of different nuclei at $t = 0$ and time 't' and gives interference 
while incoherent scattering occurs due to the correlation between the positions of same nuclei at 
different times [6].

**2.4 Atomic Motion**

Generally, incoherent neutron scattering contains three types of scatterings, which includes 
elastic, quasielastic and inelastic scattering shown in Fig. 2.2. Elastic peak is represented by delta 
function when there is no transfer of energy. In this thesis, work is done in quasielastic region, 
broadening of the peak around the elastic peak. This broadening incorporates the rotational, 
translational and vibrational motion of atoms in energy domain is given by

$$S(Q,\omega) = S_{\text{inc rot}}(Q,\omega) \otimes S_{\text{inc vib}}(Q,\omega) \otimes S_{\text{inc tr}}(Q,\omega)$$ \hspace{1cm} (2.10)

The vibrational motion is high (meV) to come in quasielastic window (μeV); therefore, it is 
measured by inelastic scattering. The above expression becomes like this:

$$S(Q,\omega) = S_{\text{inc rot}}(Q,\omega) \otimes S_{\text{inc tr}}(Q,\omega)\exp(-\frac{1}{3} <r^2>Q^2)$$ \hspace{1cm} (2.11)
The vibrational term takes the form of Gaussian distribution in quasi elastic region. This factor is known as the Debye-Waller factor.

Analyzing the data in time domain, the convolution sign in expression (eq. 2.10 -2.11) becomes the multiplication sign, and it can be written as

\[
I(Q,t) = I_{inc}^{rot}(Q,t) I_{inc}^{vib}(Q,t) I_{inc}^{tr}(Q,t)
\]  

(2.12)

\[
I(Q,t) = I_{inc}^{rot}(Q,t) I_{inc}^{tr}(Q,t) \exp\left(-\frac{1}{3} r^2 Q^2 \right)
\]  

(2.13)

---

**Fig. 2.2** Three types of scattering are used to study the atomic motion. Elastic peak (red color) is represented by delta function, quasi-elastic contains (blue color) translational and rotational motion, and inelastic scattering is used to study very fast motion of atoms such as vibrational motion.
2.5. Neutron scattering spectrometers

Neutron scattering is a powerful tool for the study of structure and dynamics of biological materials. Here we list some spectrometers below that we mostly use them to probe the dynamics and structure of various biomolecules.

2.5.1 Backscattering Spectrometer

This spectrometer is conceived based on the 'inverted scattering geometry' phenomenon. According to the phenomenon, after cooling down the neutrons, their initial energy, $E_i$, is selected by the monochromators or choppers while the final energy, $E_f$, is selected by the analyzer crystals. This technique utilizes energy analyzer crystals with a fixed Bragg reflection angle at or near 90° when $\theta \approx 90°$ the angular term vanishes. It is shown by this equation

$$ \frac{\Delta \lambda}{\lambda} = \frac{\Delta d}{d} + \frac{\Delta \theta}{\tan \theta} \quad (2.14) $$

where $\lambda$ is the wavelength of neutron, $d$ represents spacing between the crystals. Hence, the resolution depends only on the spread of the analyzer lattice constant, and it helps in minimizing the energy resolution of analyzer. This technique is characterized for the high energy resolution on the scale of $\mu$eV [17]. In neutron spectroscopy, backscattering spectrometer probes dynamics in the range of pico to nano second scale. This spectrometer comes in the middle of time of flight spectrometer that probes to the pico second scale and the spin echo spectrometer that probes to the scale of tens and hundreds of nano-seconds. The neutron energy transfer is the difference in the initial and final energies, $E_i - E_f$. At the reactor based sources, initial energy $E_i$ is varied by varying the temperature of the monochromators with respect to the analyzer. At some reactors,
Doppler-driven monochromator method is also used. In the former case, the lattice constants of monochromator are varied. On the contrary, at the spallation neutron source $E_i$ is determined from the time of flight between the neutron source and the detectors. Building the spectrometer at the reactor based or at the spallation neutron source has its own limitations and advantages. Spectrometer at the spallation source suffers from the energy resolution while the spectrometers at reactors have some limitations with the dynamic range. In this thesis, dynamics are studied by using two backscattering spectrometers BASIS at the Oak Ridge National laboratory (ORNL) [18] and the HFBS at National Institute of Standards and Technology (NIST) Center for Neutron Research (NCNR) [19].

2.5.1.1 BASIS at ORNL

BASIS is a near-backscattering, crystal-analyzer spectrometer that provides very fine energy resolution [18]. It uses Si (111) as analyzer crystals and has a long primary flight path of 84m, which is used to control the energy resolution. Its resolution is estimated by using the initial and final time of flights (initial measures from the moderator to the sample, and final measures from the sample to analyzer to detectors). The overall resolution of this spectrometer can be estimated by using these equations:

$$\delta E = \sqrt{\delta E_p^2 + \delta E_s^2}$$

(2.16)

$$\delta E_p = 2E_i \sqrt{\left(\frac{\delta L_i}{L_i}\right)^2 + \left(\frac{\delta t_0}{t_i}\right)^2}$$

(2.17)

$$\delta E_s = 2 \sqrt{\left(\frac{E_i^2}{E_f^2} \left(\frac{\delta t_f}{t_f}\right)^2 + E_f^2 \left(\cot \theta \delta \theta^2 + \left(\frac{\delta d}{d}\right)^2\right)\right)}$$

(2.18)
where $\delta E_p$ includes the contribution from the incident flight path to the sample while $\delta E_s$ is the contribution both from the sample geometry and the final flight path.

The neutrons travelled with an energy of 2082 $\mu$eV (wavelength 6.267 Å) passed only once through the sample to reduce the absorption of neutrons by the sample. After scattering from the sample, they are Bragg reflected by the analyzers and then detected by the detectors. Detectors at BASIS allow generating the reduced data sets at true constant q-values. The large q (0.2 to 2Å$^{-1}$) and energy transfers (± 200 $\mu$eV) range are covered at BASIS [18]. The layout of BASIS spectrometer is shown in Fig. 2.3.

![Image of BASIS tank with neutrons path](image)

**Fig. 2.3** Schematic picture of BASIS tank, white arrows show the path of neutrons. Neutrons scatter from the sample, reflected from the top and bottom analyzers and then detected by the top and bottom Detectors [18].

### 2.5.1.2 High Flux Backscattering Spectrometer (HFBS) at NCNR

Backscattering spectrometers have flux limited due to the narrow energy resolution. At HFBS, neutrons travelled with wavelength $\lambda_0 = 2d = 6.2712$ Å, $k_0 = 2\pi/\lambda_0 = 1.00$ Å$^{-1}$, $v_0 = 630.8$
ms\(^{-1}\) and energy \(E_0 = 2.08\) meV. The maximum accessible momentum transfer at HFBS is \(Q = 4\pi/\lambda_0 = 2.0\) Å\(^{-1}\). HFBS became the first spectrometer where Phase space transformation chopper (PST) was installed to boost the count rates without affecting the performance of the instrument. Comparing to the other spectrometers, it varies the incident energy with the cam-based Doppler-drive system [19].

### 2.5.2 Neutron Spin Echo Spectroscopy (NSE)

NSE technique was developed by Mezei in 1972 [20]. The NSE spectroscopy has the advantage that it is able to probe the slow inter-domain motions in the protein directly in the time-domain. This technique is related to spin Larmor precession angle to achieve the good resolution. The length and timescale of Neutron spin-echo (NSE) exist in the range of sub-Angstrom and picoseconds to nanometers and several tens of nanoseconds and beyond. Comparing to the QENS experiments, NSE can provide information about the dynamics at a similar q-range but at longer times.

NSE spectrometers principle is based on the spin Larmor precession \(\omega_L = \gamma B\). Spin Echo instrument details are described in Fig. 2.4. Initially neutrons enter the \(\pi/2\) flipper that flips the neutrons perpendicular to the magnetic field direction. Now, they enter the first homogeneous magnetic field coil where the neutrons start precession. After travelling 't' time through the magnetic field \(B_1\) coil of arm length \(l_1\) neutrons perform a \(N_1\) number of precessions correspond to precession angle

\[
N_1 = \frac{\alpha_1}{2\pi} = \frac{\omega_L t}{2\pi} \tag{2.19}
\]
Fig. 2.4 Schematic set up of NSE. Upper part shows the instrument set up where the neutrons enter the magnetic field coil \((B_1)\) having length \((l_1)\) perpendicular to the direction of magnetic field. Before neutrons scattered by the sample, they flipped by \(\pi\)-flipper and then pass through the second magnetic field coil \((B_2)\). At the end, scattered neutrons flipped by the \(\pi/2\) flipper so that they can pass through the analyzers and detected by the detectors. Lower portion represents the spin orientation of neutrons.

where travelling time depends on the incident velocity of neutrons \(t = l_1/v_i\). Equation 2.19 can be rewritten as:

\[
N_1 = \frac{\gamma B_1 \lambda l_1 m}{2\pi h} = \frac{\gamma B_1 l_1}{2\pi v_i} \tag{2.20}
\]

There is a \(\pi\)-flipper to rotate the spins at an angle of 180º. It has an effect to reverse the direction of beam in the second solenoid so that all the precessions are undertaken in the opposite direction. Now, neutrons are scattered by the sample, and again travelled through the second magnetic field \((B_2)\) coil where they continue to precesses \((N_2)\) given as:

\[
N_1 = \frac{\alpha_2}{2\pi} = \frac{\gamma B_2 \lambda f l_2 m}{2\pi h} = \frac{\gamma B_2 l_2}{2\pi v_f} \tag{2.21}
\]
If both magnetic coils exhibit the same length and magnetic field, then the final precession angle for a neutron is given by the difference in the initial and final wavelengths

$$\Omega = \alpha_2 - \alpha_1 = 2\pi(N_2 - N_1)$$  \hspace{1cm} (2.20)

$$= \frac{\gamma B l m}{\hbar} (\Delta \lambda)$$  \hspace{1cm} (2.21)

The second magnetic coil is followed by $\pi/2$ flipper that works as a "stop watch" to stop the precession of neutrons. This flipper also converts the final precession angle to a longitudinal polarization component because these analyzers only allowed the longitudinal component to pass through. At the end, beam polarization $(P)$ is detected by the detectors:

$$P = \langle \cos(\Omega) \rangle$$  \hspace{1cm} (2.22)

where $\langle .... \rangle$ denotes the average over all the scattered neutrons. When there is elastically coherent scattered neutrons the number of precessions are same in both of the coils, and it gives the maximum polarization $P = 1$. On the contrary, inelastically scattered neutrons exhibit small transfer of energy and the polarization can be expressed as:

$$P(Q,t) = \frac{\int_{-\infty}^{\infty} S(Q,\omega) \cos(\omega t) d\omega}{\int_{-\infty}^{\infty} S(Q,\omega) d\omega} = \frac{I(Q,t)}{I(Q,0)}$$  \hspace{1cm} (2.23)

NSE spectrometers are particularly used to measure the polarization in time domain as a function of scattering vector $q$. The scattering vector is measured by the angle $2\theta$ formed by the arm of second coil with the arm of first coil.
2.5.3 Small Angle Neutron Scattering (SANS)

SANS enables us to study the structure of biomolecules in solution, without crystalizing them. It measures the interference of scattered neutrons from all atoms of a sample in the direction of scattering vector, \( q = 4\pi \sin \theta / \lambda \) is shown in Fig. 2.5. The measured intensity reflects the cross section scattered by a protein and its solution. It is given by

\[
I_{\text{meas}}(q) = I_{\text{coh}}(q) + I_{\text{inc}} + \text{bgr}
\]

\[
I_{\text{coh}} = \frac{N}{V} S(q,c)P(q)
\]

A set of 'N' proteins in a volume 'V' at the center of mass positions. Background (bgr) includes the coherent and incoherent scattering of the solvent, which is independently measured and subtracted. The remaining quantity \( I_{\text{coh}}(q) \) measures the spatial distribution of protein atoms.

The structure factor \( S(q,c) \) of protein depends on the protein interactions and concentration, \( c = N/V \). It is the Fourier transform of radial distribution function \( g(r) \). \( P(q) \) is called the form factor that related to the spatial arrangement of atoms inside the protein. It provides the information about the size, shape and internal structure [21]. It is concentration independent and given as

\[
P(Q) = \sum_{j,k} < b_j b_k \exp (iq(r_j - r_k)) >
\]

where 'b' is the scattering length and 'r' is the position vectors of atoms.
Fig. 2.5 Schematic diagram of Small Angle Neutron Scattering (SANS); $k_i$ and $k_s$ are the incident and scattered wave vectors along with the scattering wave vector ($q$).
CHAPTER 3 MOLECULAR DYNAMICS SIMULATIONS

3.1 Introduction

Proteins belong to the large and complex system. They perform wide variety of functions in our body, for instance, antibodies, enzymes, messengers etc. Different amino acids are used for the formation of a chain of protein. To this date 105732 number of proteins, nucleic acids and their complex structure is confirmed in the protein data bank [22]. There are various experimental techniques used to study the structure and dynamics such as NMR (Nuclear Magnetic Resonance [7]), X-ray crystallography [4] and neutron scattering [6] etc. However, these experimental techniques have their limitations and can only detect the proteins motion at very limited length and time scales. Bio-molecular motion exist over wide range of time scale, for instance, local motion in proteins contains atomic fluctuations, side chain motions and loop motions exist in the length scale of 0.01 to 5 Å and time scale $10^{-15}$ to $10^{-12}$ s (femto second to pico second). The domain and helix motion comes under the rigid body motion that occurs in the length scale of 1 to 10 Å and time scale is $10^9$ to $10^6$ s [23]. There is also a folding unfolding transition that occurs in the range of 5 Å or more and time involved is $10^6$ to $10^1$ s [24]. Hence, MD Simulation is one of the important tools to investigate the bimolecular motion at all time and length scales.

MD Simulation was first introduced by Alder and Wainwright in 1959 on hard sphere model [25]. Nowadays, it is quite common to study the local and global motion as well as folding and unfolding of proteins, membranes, lipids, proteins adsorbed on nanomaterials such as Graphene, Carbon nanotubes, Gold nanoparticles, Nanodiamonds. Simulations of proteins on nanomaterials will be helpful in understanding their complex system. MD Simulations are based on classical equations of motion and statistical mechanics.
3.2 Classical Equations of motion

The total energy of the system is the sum of the kinetic and potential energy

\[ H = K + U(r) \]

K represents the kinetic energy and U(r) is the potential energy calculated from the potential energy function. Kinetic energy is represented as \( K = \sum_{i=1}^{N} m_i v_i^2 / 2 \) where \( m_i \) and \( v_i \) are the masses and velocities of the \( i \)th particle. During the simulation, instantaneous temperature can be calculated from the kinetic energy, \( E_K = \frac{3}{2} Nk_B T \) where \( k_B \) is the Boltzmann’s constant \[26\].

While running the simulation, atoms continuously change their positions and velocities after each collision. Atoms follow the Newton's laws, thus the new positions and velocities can be calculated by integrating Newton's equation of motion

\[ F = m^*a_i = m \frac{d^2 x}{d t^2} \quad (3.1) \]

If we know the potential energy of each atom then the force can be calculated from the negative gradient of potential energy function

\[ F = -\frac{\partial U}{\partial r} \quad (3.2) \]

With the given coordinates new forces, positions and velocities are generated. It generates a new trajectory by repeating the calculations at each interval of time. The more realistic calculation comes from the breaking of integration of equations of motion into a very short time steps. At each step, the forces act between the atoms are computed and combined with the new values of positions and velocities for all atoms. It is assumed that the force is constant during that time step. There are many algorithms which are used to solve the Newton's equations of motion in molecular dynamics for instance verlet, leap frog, velocity verlet etc. Velocity verlet is very common in NAMD and other programs to run the simulation \[27\]. Position and dynamic
coordinates (velocities, accelerations etc.) in all of these algorithms are expanded by the Taylor's series. In verlet it comes as

\[ r(t + \delta t) = r(t) + v(t) \delta t + \frac{1}{2} a(t) \delta t^2 \]  
\[ r(t - \delta t) = r(t) - v(t) \delta t + \frac{1}{2} a(t) \delta t^2 \]  

(3.3)

(3.4)

Summation of 3.3 and 3.4 equations gives

\[ r(t + \delta t) = 2r(t) - r(t - \delta t) + a(t) \delta t^2 \]  

(3.5)

The velocities can't calculate directly in this algorithm. There are several ways to calculate the velocities. The simplest one is to take the difference of velocities at time \( t+\delta t \) and \( t-\delta t \) and divided by \( 2\delta t \).

\[ v(t) = \frac{r(t + \delta t) - r(t - \delta t)}{2\delta t} \]  

(3.6)

Alternatively, velocities can also be calculated at the half step, \( t + \frac{1}{2} \delta t \)

\[ v(t + \frac{1}{2} \delta t) = \frac{r(t + \delta t) - r(t - \delta t)}{\delta t} \]  

(3.7)

The advantage of this algorithm is that the implementation is very easy. The storage requirements are also modest, only three parameters are required to store. Two sets of positions \( r(t), r(t-\delta t) \) and the accelerations, \( a(t) \). It has many disadvantages too. The position coordinates are calculated by the addition of this small term \( a(t) \delta t^2 \). In addition, the lack of velocity term makes it difficult to get the velocity value, and you can't get it until the positions have been computed at the next step. The next position is calculated from the current position and previous position. At the beginning of the simulation, \( t = 0 \), only one set of positions are available. Therefore, it is must to provide the values at time \( t-\delta t \). One method is to use the Taylor's expansion.
There are several variations employed on the Verlet algorithm. Then, it has been improved by coupling with the leap frog algorithm. The leap frog algorithm comes in this was

\[ r(t+\Delta t) = r(t) + v(t+\frac{1}{2}\Delta t)\Delta t \]  \hspace{1cm} (3.8)

\[ v(t+\frac{1}{2}\Delta t) = v(t-\frac{1}{2}\Delta t) + a(t)\Delta t \] \hspace{1cm} (3.9)

In leap frog algorithm, velocities at this time, \((t+\frac{1}{2}\Delta t)\) are first calculated from the velocities at \((t-\frac{1}{2}\Delta t)\) and the acceleration at time \(t\). The velocities can be calculated from the

\[ v(t) = \frac{1}{2}[v(t+\frac{1}{2}\Delta t) + v(t-\frac{1}{2}\Delta t)] \]  \hspace{1cm} (3.10)

The velocities first leap frog over the positions to give their values at time \(t+\frac{1}{2}\Delta t\). The positions then leap frog over velocities at time \(t+\Delta t\) and so on. Although velocities are explicitly included in it and do not need to consider the difference of two equations but the positions and velocities are not synchronized. Consequently, it is difficult to calculate the contribution of kinetic and potential energy at the same time (at which positions are defined).

Velocity verlet algorithm has advantage over leap frog that positions, velocities and accelerations are computed at the same time. Its expression is given as:

\[ r(t+\Delta t) = r(t) + \Delta t v(t) + \frac{1}{2}\Delta t^2 a(t) \]  \hspace{1cm} (3.11)

\[ v(t+\Delta t) = v(t) + \frac{1}{2}\Delta t[a(t) + a(t+ \Delta t)] \] \hspace{1cm} (3.12)

Using the acceleration at time \('t'\) and \('t+\Delta t'\) can calculate the new velocities. The velocities are determined at half time step as follows:

\[ v(t+\frac{1}{2}\Delta t) = v(t) + \frac{1}{2}a(t)\Delta t \] \hspace{1cm} (3.13)

The final positions and velocities comes this way

\[ r(t+\Delta t) = r(t) + \Delta t v(t+\frac{1}{2}\Delta t) \]  \hspace{1cm} (3.14)
In the next step, using the new positions the updated values of forces $f(t+\delta t)$ are obtained, and the full step velocities are calculated using this expression

$$v(t+\delta t) = v(t+\frac{1}{2}\delta t) + \frac{1}{2}\delta t*a(t+\delta t)$$

Hence all of the co-ordinates (positions, velocities and accelerations) are synchronized [28].

### 3.3 Force fields

The most important calculation in the MD simulation is the Force fields and its parameters. There are various force fields that used in MD simulation such as Charmm [29], Amber [30], Gromacs [31] etc. Charmm force fields were used to run the simulation on NAMD software [32].

$$U(r) = U_{bond} + U_{angle} + U_{dihedral} + U_{improper} + U_{LJ} + U_{Coulomb}$$

(3.16)

The potential energy function is separated into two terms: covalent bonded and non-covalent bonded interactions between atoms. In our MD simulation, both of these interactions were calculated.

#### 3.3.1 Bonded Interactions:

$$U_{bond} = \sum_{bonds} K_b(b - b_0)^2$$

(3.17)

$$U_{angles} = \sum_{angles} K_\theta(\theta - \theta_0)^2$$

(3.18)

$$U_{dihedral} = \sum_{dihedral} K_\psi(1 + \cos(n_\psi - \delta))$$

(3.19)

$$U_{improper} = \sum_{improper} K(\phi - \phi_0)^2$$

(3.20)

In all of the above cases summation goes from $i = 1$ to $N$. The bond stretching exists between the two covalent bonds where $b$ is the actual bond length, $b_0$ is the reference length, and $K_b$ is the bond stretching force constant shown in Fig. 3.1(a). The reference bond length and force constant are specifically defined for only that pair. Their values are extracted from experiment or calculated from quantum mechanics. The bond angle exists between triplets of consecutive
atoms where $K_\theta$ is the angle bending force constant, $\theta$ is the actual angle and $\theta_0$ is the reference angle. Dihedral angles are formed between the planes formed by the set of three atoms where $n$ is the multiplicity, which implies the number of minima exists as the bond is rotated through $360^\circ$. The multiplicity is a positive non-zero integer. $K_\chi$ and $\delta$ are the force constant and phase angle. The atoms in a plane are chosen according to the IUPAC convention. The improper dihedral potential is used to maintain the planarity of the molecular structure. The same atoms are used in the formation of planes as in the dihedral potential but at different positions shown in Fig. 3.1(d). In the improper dihedral potential, three atoms are centered on the fourth atom.

![Fig. 3.1](image) (a) Covalent bond, (b) Bond angle, (c) Dihedral angle, (d) Improper dihedral angle, (e) Long range Vander-Waals and (f) Coulomb electrostatic interactions.

### 3.3.2 Non-bonded interactions

There are two types of non-bonded interactions occur between the same atoms or with the other atoms: Vander Waals Fig. 3.1(e) and Electrostatic interactions Fig. 3.1(f).

**a) Van der Waals interactions:** These referred to the combination of repulsive and attractive forces between atoms. In MD simulations, Lennard-Jones (LJ) potential (equation 3.21) is used to describe these forces. The first term represents the repulsive, and the second term
represents the attractive interactions. The energy is zero when the atoms are far apart from each other. As they start coming towards each other, the potential energy becomes negative. After passing through the minima, the energy rises up. It is written as

\[ U_{\text{LJ}} = 4 \varepsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^{6} \right] \]  

(3.21)

\( \varepsilon \) is the depth of the potential, \( \sigma \) is the distance at which the potential is zero and \( r \) is the distance between the atoms. In our simulation, the Lorentz-Berthelot mixing rules are applied to determine the values of LJ parameters of carbon of Nanodiamond with these equations

\[ \sigma_{ij} = \frac{1}{2} (\sigma_{ii} + \sigma_{jj}) \text{ and } \varepsilon_{ij} = \sqrt{\varepsilon_{ii} \varepsilon_{jj}} \]  

(3.22)

**Electrostatic Interactions:** These interactions are modeled by the coulomb potential

\[ U_{\text{coulomb}} = \frac{q_i q_j}{4\pi \varepsilon_0 r} \]  

(3.23)

\( q_i \) and \( q_j \) are the atomic charges, \( r \) is the distance between the atoms and \( \varepsilon_0 \) is the permittivity of space [23].

### 3.4 Periodic Boundary Conditions

MD Simulations are in high demand in smaller as well as in larger systems, which consists of up to millions of atoms. In a system containing millions of atoms, they do not stay in the finite box size, holes are created, broken molecules appear etc. Moreover, the large number of atoms exists on the simulation box surface than those in the bulk. Hence, the surface atoms experience different forces than the atoms in the bulk. It affects the properties of system [33]. Therefore, Periodic Boundary Conditions (PBC) is used in MD Simulation to avoid these problems, and the boundary effects due to the finite box size. The PBC mainly used for the bulk systems such as proteins, nucleic acids, glasses, gases and mixtures. In PBC, a primary cell is created in the middle, and the exact replicas of this cell are surrounded around the primary cell. During the
simulation, if one atom exits the primary cell then the other atom enters from the opposite side of the cell. The atoms in the replicas also moved with the same way as they are moving in the primary cell shown in the 2-D Fig. 3.2. PBC has open boundaries and atoms can easily move from one cell to another. Consequently, it keeps the density of the cell conserve by exiting the atom from one side and enters from the other side. In this way, the entire system is conserved and shown in the Fig. 3.2 [34]. Although there are various cases where PBC not good to use because they require the presence of large amount of water molecules in the system.

![Fig. 3.2](image)

Fig. 3.2 Periodic Boundary Conditions are shown in 2-D. The primary cell is present in the middle in grey color. Eight replicas surround the primary cell. One atom is exiting the cell, and its position is replaced by the other one [34].

### 3.5 Cut off radius and Ewald Sum

The non-bonded interactions make the simulation very time consuming and expensive due to the large number of atoms interact with each other. Ewald Sum and Cut off radius are used to cut the cost of simulation. The short range interactions are calculated with the cut off radius. It is
defined as the small sphere is made around the primary atom. In this way, a primary atom interacts only with those atoms which come inside that defined sphere. This rule is good to calculate the LJ interactions. The electrostatic interactions are calculated with the implicit methods such as continuum dielectrics and static dipolar fields as well as explicit methods such as Ewald sum, truncation and interaction shifting. Among them Ewald sum is the most accurate and efficient with the periodic boundary conditions. With the Ewald sum, the convergent sum is converted into the sum of two convergent summations: (a) short-range real space, and (b) a long range reciprocal-space summations [35].

3.6 Statistical Mechanics and Ensembles

MD simulation deals with the microscopic states at atomic level such as atomic position, atoms velocities etc. while the physical science studies the relation among macroscopic observables such as temperature (T), pressure (P) and volume (V). Statistical mechanics make a link between these microscopic and macroscopic observables. Different types of ensembles are used in MD simulations.

**Microcanonical ensemble (NVE):** In Newton's integration of equations of motion, the total energy of the system is conserved for fixed total number of particles and volume. This system is referred to as micro-canonical ensemble (NVE).

**Canonical ensemble (NVT):** The number of particles, volume and the temperature of the system are constant. This system is coupled with thermostat to keep the temperature constant.

**Isothermal-isobaric ensemble (NPT):** In MD Simulations, Isobaric-isothermal ensemble (NPT) is most common ensemble used especially for the bio-molecules. It is characterized by fixing the total number of particles, pressure and temperature of the system.
In NPT and NVT ensemble, the thermostat is required to keep the temperature constant and acts as a source of thermal energy, which supplies and removes the heat from the system. The temperature constant simulation may be required if someone wants to see the temperature dependent behavior of the system. Canonical (NVT) and isothermal-isobaric ensembles (NPT) are mainly used to keep the temperature constant. The temperature is maintained by coupling the system to an external bath, which is fixed at the desired temperature. There are several thermostats that are used in the NVT and NPT ensembles for instance Nose Hoover, Berendsen, Gaussian (velocity-rescaling), Anderson and Langevin thermostats. Among them NAMD uses the Langevin thermostat. This thermostat is based on the langevin equation, which considers the addition of frictional force and noise terms. It can be written as

\[ m\ddot{a}_i = F_i - \frac{1}{\tau_d} m \dot{v}_i + \sqrt{\frac{2kTm_i}{\tau_d}} \, R_i(t) \]  

where \( kT \) denotes the thermal energy, \( \tau_d \) is the characteristic viscous damping time and \( R_i(t) \) is the delta-correlated stationary Gaussian process with zero-mean.

### 3.7 Transport Properties

Neutron scattering and molecular dynamics have a great combination to study the structure and dynamics of complex molecular systems. Neutron scattering is sensitive to atomic correlations in length at Angstrom scale and time to ps scale. Molecular Dynamics (MD) simulations covered exactly the same range and beyond that too [36]. MD simulations provide us a trajectory of atomic position coordinates as a function of time while neutron scattering measurements are presented in reciprocal space. The Fourier transform of dynamic structure factor gives intermediate scattering function (ISF) in time domain. The same quantities such as ISF(Q,t), S(Q,ω) etc can be obtained in MD simulation by using the corresponding equations.
In addition to these quantities, some other properties can also be calculated, which are not measured by quasielastic neutron scattering technique. nMoldyn software is used to measure all of the quantities required according to the project requirements [36].

### 3.7.1 Mean Square Displacement (MSD)

MSD is another important quantity describing the atomic or molecular dynamics. It is defined as change in the atomic position

\[ d_\alpha(t) = R_\alpha(t) - R_\alpha(0) \]  

(3.25)

MSD of \( \alpha \) particle is given as

\[ \Delta^2_\alpha(t) = <d_\alpha(t)^2> \]  

(3.26)

Diffusion coefficient is obtained from MSD by using the standard Einstein-diffusion law

\[ D_\alpha = \lim_{t \to \infty} \frac{1}{6t} \Delta^2_\alpha(t) \]  

(3.27)
CHAPTER 4 EFFECT OF NANODIAMOND ON RNA DYNAMICS

4.1 Introduction

4.1.1 Nanoparticle-Biomolecules Complex

Genetic disorder in human beings causes complex diseases, such as Cancer, Alzheimer’s disease, and Parkinson's disease. Many strategies have been employed for their treatment, but it is still an area of interest in the research, for instance surgical removal of the infected parts [37]. The bio-nanotechnology is emerging as a new interdisciplinary research field for the medical purposes. Nanoparticles serve as a basic building block for the bio-nanotechnology research including metals, semiconductors, polymers, carbon-based nanomaterials etc. These nanoparticles have got a great attention due to their optical, electrical and structural properties [38], and they exist in the same size domain (1-100 nm) of intercellular structures and biomolecules [39]. The coupling between nanoparticles and biomolecules provide the basis for understanding the surface interactions, there by their future use in drug delivery. Incorporating the biomolecules on the surface of nanomaterials increase their stability [40]. Moreover, the knowledge of surface interaction between amino acid side chains and nanoparticles pave the way to generate novel materials such as biosensors [41]. Bio-nanomaterials also have applications for the cure of genetically based diseases, these nanomaterials target the area of interest such as variation in pH value [42] and inhibit the overdevelopment of gene expressions or proteins, for instance RNA nanotechnology [43, 44]. The composite of biomolecules and nanomaterials have strong potential for the treatment of these diseases. Hence, it becomes important to study the surface interaction between nanomaterials and biomolecules.

Biomolecules, such as nucleic acids and proteins have been used for long time in drug delivery [41]. The discovery of RNAi (RNA interference) opens the way for the cure of many
incurable diseases [45, 46]. They have advantage to target undrug-able diseases while therapeutic based drugs target only the receptors present on the cells surface [47]. Currently, main focus comes on the siRNA (short interfering or silencing RNA) due to its smaller size, and hence easy to get absorbed into the body through the cell's membrane [48]. However, siRNA-based therapeutics faced major challenge in reaching the target inside the cell [49]. siRNA gets reflected from its path because of the electrostatic repulsions caused by negative charge on the backbone of siRNA and on the cell membrane [50]. Therefore, use of siRNA needs a carrier to get into the target cells. Nucleic acid and nanomaterials complex has a great potential for the treatment of uncurable diseases, but first we need to do fundamental research on their complex such as understanding the behavior of nucleic acids with nanomaterials. In this thesis, we study the effect of nanomaterial on the dynamics of RNA.

4.1.2 Nanomaterials

Nanomaterials have received a lot of attention in the area of drug delivery due to their high stability, low toxicity and biocompatibility. To this date, many nanomaterials have been developed as drug carriers to deliver drug at the target inside the cells [51]. The first nanotechnology drug delivery system was developed by using 'lipid vesicles' in 1960 which are later known as 'Liposomes' [52]. Liposomes are the spherical vesicles with phospholipid bilayer membrane structure that makes them an attractive drug vehicle for delivering therapeutic cargos such as nucleotides, proteins etc. In 1973, they were reported to use them in vivo delivery, but suffered from the premature delivery of cargo. After the liposomes, polymer-drug conjugates became the second most important nanoparticle-based therapeutic. There has been various polymers developed over decades [53]. In addition, the recent developments in nanoscience and nanotechnology have showed the potential in carbon based nanomaterials for biomedical
applications. Carbon nanotubes [54], fullerene C60 and fluorescent carbon dots have also been developed for biomedical purposes [55]. All of these have some advantages or disadvantages over each other, and there is another class of carbon material introduced at the end of 1980s called ‘Nanodiamonds’ (NDs). These nanoscale diamond paricles were first discovered in the Union of Soviet Socialist Republics (USSR) in 1960s [56], where they were produced by detonation technique, but they remained unknown to the rest of the world until the end of 1980s. By then, people sought interest in these nanoparticles. By definition, NDs are carbon nanoparticles, and exhibit truncated octahedral geometry. As the name 'nanodiamond' infers, they exhibit the diamond like characteristics such as chemical stability, hardness, stiffness and strength at nanoscale. In addition to these properties, they are available at nanoscale having large surface area and high adsorption capacity. They have been synthesized by various techniques, for instance, detonation technique, laser ablation, high-energy ball milling of high-pressure high temperature (HPHT) diamond microcrystals, plasma-assisted chemical vapor deposition (CVD) etc. Among them first three methods are commercially used [57].

Versatile properties of NDs enable them to be used in various aspects, for example, drug delivery and biosensors. Recently, they attracted much attention due to its luminescence in the visible range, which can be used as a tracker in biological applications [58]. Production of ND is environment friendly and high purity nanodiamonds (powder form) are produced in large amounts at a low cost [59]. They are found to be less toxic compared to other carbon nanomaterials [51] and, consequently, well suited for the drug delivery, tissue engineering [44], tribology [38] etc. The effectiveness of drug delivery depends upon the following factors: (1) a large binding surface area, (2) the ability to reach the targeted cells, and (3) the ability to deliver the active molecules to their target [60]. NDs exist in the size domain of intercellular structures
and large biomolecules. They are low toxic and biocompatible, and they can be quickly removed from the body [39]. NDs have attracted great attention due to the above all of its properties. This thesis deals with the study of the NDs effect on the dynamics of RNA with neutron scattering and Molecular Dynamics (MD) simulation.

Billions of years of natural evolution has created a set of functional building-block molecules like RNA and DNA [61] that enable various fundamental biological functions. Humankind, on the other hand, is no less creative in rivaling the biological complexity while producing materials like nanodiamonds that can have ramifications for biomedical applications [62]. DNA based nanomaterials are now well established [63] in nanomechanical devices and drug delivery applications [64]. Similarly, there is a surge in interest in RNA nanotechnology in recent years due to its potential applications in the treatment of cancer and genetic disorders [61]. On the carbon-based nanomaterials front, nanodiamond (ND) is considered a viable candidate for drug delivery in lung carcinoma cells [65] and bactericidal applications [66] due to its low cytotoxicity and higher biocompatibility. ND is a good platform because of the simple geometry, consistency in shape and uniform surface chemistry. ND is also safer compared to tubular shaped nanomaterials [51]. Functionalized ND complexed with biodegradable, biocompatible polymers has recently been demonstrated for biomedical applications [67]. Since the surface of ND can be easily tailored with ionogenic groups (ether-C-O-C, peroxide –C-O-O-, carbonyl –C=O and hydroxyl type C-O-H bonding, etc.) and hydrocarbon fragments to adsorb large number of biologically active molecules [68], it is prudent to pursue RNA-ND nanocomposites as possible materials for effective biomedicines. The non-toxic ND nanocomposites possess excellent mechanical and optical properties with high surface areas and are of the size of intercellular structures and large biomolecules [69]. ND complexes have been shown to be well suited for
applications in drug-delivery [61, 62], tissue engineering [61, 70], tribology [62] and bio-
ingaging [62, 71], however there is very little understanding of the structure as well as dynamics of these composite materials. Since the presence of a ND leads to a corona [72] between ND, water and RNA, a basic understanding of the structure/dynamics and its tenability of a biomolecules(RNA/ND) composite is essential for further development/refinement in designing novel materials.

In this work, we investigate the dynamics of hydrated and de-hydrated RNA molecules on ND surfaces relative to the freestanding RNA by using quasi-elastic neutron scattering (QENS) and atomistic molecular dynamics (MD) simulations. Through these methods, we are able to distinguish the water dynamics from that of the RNA dynamics at atomic and molecular levels. Contrary to the generally held notion that nanoscale spatial heterogeneity at interfaces lead to correlated particle motions [73], there by slowing down the dynamics, our results show faster RNA dynamics on ND surfaces compared to dry or freestanding cases within the temperature range investigated. Also, these results are different in comparison to previous experimental observations on another set of biomolecules on silica surface [74], where it was argued that the adsorption of biomolecules not only decreased the flexibility but simultaneously modified the mobility of residues and dynamics upon surface interaction. We attribute the faster dynamics on ND surfaces to the de-confinement of both the RNA and water molecules, a result of the triple phase coexistence of water+RNA+ND at the biomolecular corona on the surfaces of the ND. This is purely an entropic effect and can as such be fine-tuned by chemical functionalization of the ND surface, fundamentally altering the properties of the biomolecules on nanomaterial surfaces.
4.2 Materials and Methods

NDs (radius ~2.5nm) were prepared with detonation technique [69] and were placed inside a oven to remove most of water molecules adsorbed on ND surfaces. tRNA was bought from Sigma Aldrich and was used without further purification. The hydrated sample was prepared by adsorbing 0.03 g of tRNA on 0.33g of dry ND surface, and then hydrated with 0.06g of D₂O. The dry sample was prepared by directly adsorbing tRNA on ND surfaces without hydration process. For D₂O hydrated tRNA sample without ND, we used data from our previous experiment where tRNA hydration level was \( h = 0.5 \) (g of D₂O/g of tRNA) [75].

The QENS experiment was performed on backscattering spectrometer (BASIS) at SNS, Oak Ridge National Laboratory. The elastic scans were measured by cooling down from 300 to 20K at a decay rate of 5 K/min. In the quasi-elastic mode, the energy resolution is 3.4 μeV (full-width at half maximum) and the used dynamic range was ±119.5μeV. The hydrated sample was measured from 220 K to 300 K with 10 K interval, and the dry sample was measured from 220 K to 320 K with 20 K intervals. The resolution for both samples was measured at 4 K. The measured QENS spectra \( S_m(Q,\omega) \) provides the self-dynamic structure factor. \( S_H(Q,\omega) \) of hydrogen atoms of tRNA convoluted with instrumental resolution \( R(Q,\omega) \),

\[
S_m(Q,\omega) = S_H(Q,\omega) \otimes R(Q,\omega)
\]  

(4.1)

The measured dynamic structure factor is modeled as

\[
S_m(Q,\omega) = [A_0 \delta(\omega) + (1 - A_0) S(Q,\omega) + B(Q,\omega)] \otimes R(Q,\omega)
\]

(4.2)

where \( \delta(\omega) \) is a delta function centered at zero energy transfer, \( A_0 \) is fractional intensity of elastic scattering, \( R(Q,\omega) \) is the resolution function, \( B(Q,\omega) \) is linear background and \( S(Q,\omega) \) is the model scattering function. The dynamics of tRNA are also studied in time domain. Before
performing the Fourier transform of BASIS data from energy into time domain, we subtracted the linear background from the QENS spectra. Fourier Transform converts equation (4.1) to

\[ F_m(Q,t) = F_H(Q,t) \times R(Q,t) \] (4.3)

where \( F_H(Q,t) \) is the ISF of the hydrogen atom of tRNA and can be calculated

\[ F_H(Q,t) = \frac{F_m(Q,t)}{R(Q,t)} \] (4.4)

where \( F_m(Q,t) \) is the Fourier transform of measured QENS spectra, \( S_m(Q,\omega) : F_m(Q,t) = \text{FT}[S_m(Q,\omega)] \), and \( R(Q,t) \) is Fourier transformed resolution.

### 4.3 Neutron Scattering Data Analysis

The Intermediate Scattering Function (ISF) of the hydrogen atoms, \( F_H(Q,t) \) can be fitted by an asymptotic expression derived from Mode Coupling Theory (MCT) [76-78]:

\[ F_H(Q,t) \sim [f(Q,T) - H_1(Q,T) \ln(t/\tau_\beta(T)) + H_2(Q,T) \ln^2(t/\tau_\beta(T))] \exp[t/\tau_\alpha(Q,T)] \] (4.5)

where \( \tau_\beta(T) \) and \( \tau_\alpha(Q,T) \) are the characteristic \( \beta \)- and \( \alpha \)-relaxation times, \( f(Q,T) \) is a temperature-dependent pre-factor, which is proportional to the Debye-Waller factor for small \( Q \)s, i.e. \( f(Q,T) = \exp[-A(T)Q^2] \). The \( Q \)-dependent parameters, \( H_1(Q,T) \) and \( H_2(Q,T) \), can be written as \( H_1(Q,T) = h_1(Q)B_1(T) \) and \( H_2(Q,T) = h_1(Q)B_2(Q,T) \), representing the first and second order logarithmic decay parameter [78, 79]. In our experiment, the time range (up to 1 ns) is much shorter than \( \alpha \)-relaxation time range (\( \mu \)s to ms), hence the value of the last exponential factor can be approximated to unity. Equation (4.5) can be simplified as

\[ F_H(Q,t) \sim f(Q,T) - H_1(Q,T) \ln(t/\tau_\beta(T)) + H_2(Q,T) \ln^2(t/\tau_\beta(T)) \] (4.6)

### 4.4 Simulation Methodology

MD Simulations are performed on hydrated RNA systems both with and without the presence of ND. The initial coordinates for hammerhead (RNA) were taken from the protein data bank (PDB: 299D). This RNA has already been used to explain and match the neutron scattering
results from tRNA [80]. We extended this idea by using MD simulations of D$_2$O hydrated RNA in the presence and absence of ND same as the experimental conditions. The hydration level in the D$_2$O hydrated RNA without ND, is kept same with the experiments i.e., h=0.5 (g D$_2$O/g of RNA). Initially for the simpler of the two systems i.e., without ND, a single RNA is placed into a pre-equilibrated box of water eliminating the overlapping water molecules and subsequently replicating into 8 clones. Furthermore, the system was neutralized by the addition of sodium ions. Also each of the eight RNA molecules, the ions and water are rotated by a random angle around a randomly chosen principal axis. In case of the system with ND, eight 2.5 nm (radius) ND spheres are prepared to mimic the experimental conditions. 8 RNAs and 10464 water molecules are surrounded around the eight 2.5 nm radius ND balls (containing a total of 92320 carbon atoms) matched with the experimental RNA, water and ND ratio. Simulations are performed on both systems using the CHARMM-27 protein nucleic acid force field [81] and TIP3P [82] water model using NAMD [32]. We used a ND ball fully made of carbon atoms and available Lennard-Jones (LJ) parameters from literature to keep it hydrophilic. In our work, the $\sigma$ and $\varepsilon$ parameters of the LJ potential, representing the interactions between the non-bonded carbon atoms are calculated from the parameters of graphite and carbon nanotubes [83] by applying the Lorentz-Berthelot mixing rule ($\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2}$, $\varepsilon_{ij} = \sqrt{\varepsilon_{ii} \varepsilon_{jj}}$). The values of the parameters used in simulation are shown in Table 4.1.

<table>
<thead>
<tr>
<th>Bonds</th>
<th>$R_{\text{min}}/2$</th>
<th>Epsilon($\varepsilon$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C-O</td>
<td>1.845</td>
<td>0.138</td>
</tr>
<tr>
<td>O-O</td>
<td>1.7682</td>
<td>-0.15210</td>
</tr>
<tr>
<td>C-C</td>
<td>1.92</td>
<td>-0.13</td>
</tr>
</tbody>
</table>

**Table 4.1** Lennard Jones parameters, epsilon (Kcal/mol) and Rmin/2 (Angstrom) are calculated for the Nanodiamond carbon.
The surface of the system satisfies the condition of hydrophilicity, which requires the size of the solvent particles to be smaller than that of solute particles and the strength of solvent-solvent intermolecular interactions to be weaker than that of solute-solvent interactions [84]. Here, the solute is nanodiamond and the solvent is water. The deuteration of water is done during analysis by using nMoldyn software [85]. This is a standard procedure used for simulated data in order to match neutron experiments data [85, 86]. The only difference is that deuteration increases the viscosity by 1.23 (at 298 K) compared to water. So, we see \( R_{\text{min}/2 \ \text{CC}} = 1.92 \ \text{Å} > R_{\text{min}/2 \ \text{HH}} = 0.2245 \ \text{Å} \) and \( R_{\text{min}/2 \ \text{CC}} = 1.92 \ \text{Å} > R_{\text{min}/2 \ \text{OO}} = 1.7682 \ \text{Å} \) (\( R_{\text{min}} \) is the distance where the potential attains the minimum). In other relevant work, Sendner et al. [87] showed how hydrophobic-hydrophilicity varies with the interaction energy while keeping the size unchanged. In that work the interaction energy (\( \varepsilon_{\text{CO}} \)) between carbon and oxygen atoms was tuned between 0.026 Kcal/mol and 0.171 Kcal/mol, with greater hydrophobic behavior corresponding to the decreased interaction energy. In our simulation, \( \varepsilon_{\text{CO}} = 0.138 \ \text{Kcal/mol} \), \( \varepsilon_{\text{CH}} = 0.077 \ \text{Kcal/mol} \) and \( \varepsilon_{\text{OO}} = -0.15210 \ \text{Kcal/mol} \), \( \varepsilon_{\text{HH}} = -0.04600 \ \text{Kcal/mol} \). This made the ND surface more hydrophilic in nature. Periodic boundary conditions are used to determine the long-range electrostatic interactions. The short-range interactions are calculated within the cut off 12Å and Particle Mesh Ewald (PME) is used to evaluate the long-range interactions. Prior to data collection runs, energy of the systems was minimized and the systems were equilibrated for 6ns in the NPT ensemble at five different temperatures, \( T = 260K, 270K, 280K, 290K \) and 300 K. The Root Mean Square Deviation (RMSD) plots of two different Systems ND+RNA+D₂O and RNA+D₂O at different temperatures are presented in Fig. 4.1. After equilibration, simulations are carried out for another 5ns at each temperature with time step of 1fs. The subsequent analyses on the dynamics were performed by nMoldyn [85].
Fig. 4.1 Root Mean Square Deviation (RMSD) values of the D₂O hydrated RNA without and with ND at three different temperatures 260, 280 and 300 K.

4.5 Results and Discussion

The incoherent neutron scattering cross-section for hydrogen is at least 20 times larger than that of other atoms in the bio-ND system, such as carbon, oxygen, nitrogen and sulfur atoms. This unique characteristic is exploited to capture the dynamics of the tRNA (transfer RNA) within the tRNA+ND+water system by replacing the hydration water H₂O with heavy water (D₂O). Here we compare three different samples, tRNA hydrated with D₂O (tRNA+D₂O), dehydrated tRNA on ND (tRNA+ND) and hydrated tRNA on a ND surface (ND+tRNA+D₂O). In all these three samples, the majority of hydrogen atoms only occur in tRNA molecules, indicating that the measured QENS spectra solely represent the dynamics from the tRNA component of the three systems respectively. The measured normalized QENS spectra, namely, the self-dynamic structure factor $S(Q,\omega)$, is plotted on a log scale as function of energy transfer in Fig. 4.2. The broadening of the central peaks from the resolution function is the result of quasi-elastic scattering of neutrons from hydrogen atoms in the sample. The broader the central
peaks, the faster dynamics of hydrogen atoms in the samples. The QENS spectra at \( Q = 1.1\text{Å}^{-1} \) (\( Q \) being the magnitude of the wave vector transfer in the scattering) are compared for three different samples at temperature \( T = 300 \text{ K} \). One can observe that apparently the central peak of the ND+tRNA+D\textsubscript{2}O sample is broader than the other two samples implying faster tRNA dynamics in the presence of the ND and D\textsubscript{2}O.

The Fourier transform of \( S(Q,\omega) \), the intermediate scattering function (ISF), \( F_H(Q,t) \), is plotted in Fig. 4.3 at three temperatures 300K for the three different samples. The ISF \( F_H(Q,t) \), known as the particle-particle correlation function, is the key function to connect theoretical prediction, neutron scattering experimental data and MD simulation results. Here, we use an asymptotic expression (Equation 4.6) [75] derived from the most popular glass transition theory, mode coupling theory (MCT) [76, 77, 79, 88], to fit our experimental \( F_H(Q,t) \) in the time range from ps to ns. The solid lines in Fig. 4.3 represent the fitted curves. Here, we show again in the time domain that the tRNA within the ND+tRNA+D\textsubscript{2}O (blue symbols) sample has the fastest relaxation dynamics.

![Fig. 4.2 Comparison of normalized \( S(Q,\omega) \) of three different samples ND+tRNA+D\textsubscript{2}O, ND+tRNA and tRNA+D\textsubscript{2}O at 300 K and at one q-value 1.1Å\(^{-1}\).](image)
The Fourier transform of $S(Q,\omega)$, the intermediate scattering function (ISF), $F_H(Q,t)$, is plotted in Fig. 4.3 at 300K for the three different samples. The ISF $F_H(Q,t)$, known as the density-density correlation function or particle-particle correlation function, is the key function to connect theoretical prediction, neutron scattering experimental data and MD simulation results. Here, we use an asymptotic expression [76] (see equation 4.6) derived from the most popular glass transition theory, mode coupling theory (MCT), to fit our experimental $F_H(Q,t)$ in the time range from ps to ns. The four parameters in this model, $A(T)$, $\tau_\beta(T)$, $H_1(Q,T)$ and $B(T)$ are obtained by fitting different Q curves simultaneously at each temperature. $A(T)$ is fixed to zero because the pre-factor $f(Q,T)$ goes to 1 at all the Q-values at a specific short time $\tau_\beta(T) \sim 10$ ps which is much shorter than our measured time range. The relaxation of tRNA+ND sample is the slowest (open symbols) followed by tRNA+D$_2$O (half-filled symbols). ND+tRNA+D$_2$O (solid symbols) sample shows the fastest relaxation. The $H_1(Q,T)$ and $B(T)$ factor also showed the faster dynamics exhibit by the tRNA+D$_2$O+ND system than tRNA+D$_2$O and tRNA+ND shown in Fig. 4.4 a and b. The solid lines in Fig. 4.3 represent the fitted curves. Here, we show again in the time domain that the tRNA within the ND+tRNA+D$_2$O (blue symbols) sample has the fastest relaxation dynamics.
**Fig. 4.3** Intermediate Scattering Function (ISF) of ND+tRNA+D2O, tRNA+D2O and tRNA+ND sample is plotted at 300 K and at five different q-values.

**Fig. 4.4** a and b, $H_1(Q,T)$ and $B(T)$ are the fitted parameters of logarithmic model (Eq. 4.6) respectively. $H_1(Q,T)$ is plotted as a function of Q at three different temperatures while $B(T)$ is plotted as a function of temperature.

The mean square displacement (MSD) plotted as a function of temperature T is traditionally used as an indicator of the flexibility (or ‘softness’) of the measured biomolecules, in our case, tRNAs. At a given temperature, the steeper the slope of MSD vs T curve, the softer the biomolecule is [89]. The MSD is calculated through analysis of incoherent elastic scattering using the Debye Waller Factor, $S(Q,\omega=0) = \exp[-Q^2<x^2(T)>]$, where $<x^2(T)>$ is the MSD at temperature T. The MSD of hydrated and dehydrated tRNA with ND are plotted in Fig. 4.5. A
sudden change in the slope of MSD at around 230K is observed in D$_2$O hydrated tRNA on the
surface of ND, but not noticeable in the other dehydrated sample. This inflection in the slope of
MSD is often referred to as the dynamic transition, of biomolecules, and has been observed in
hydrated proteins [90, 91] and hydrated tRNA [90, 92], but not in dry biomolecules. It is
considered to be a transition from harmonic to anharmonic behaviors that is responsible for the
performance of biological functions. Our observation confirms that the activation of tRNA is
hydration dependent and this inflection is absent in dry sample [93, 94].

![Fig. 4.5 Mean Square Displacement (MSD) of hydrated and dry tRNA on nanodiamond surface is plotted as a function of temperature. Dynamical transition ($T_d$) is observed at 230 K only in hydrated sample.](image)
**Fig. 4.6** a and b Experiment and MD simulation Intermediate Scattering Function (ISF) of ND+RNA+D$_2$O, ND+RNA and tRNA+D$_2$O samples at one q-value 1.1Å$^{-1}$ and at three different temperatures 260, 280 and 300 K respectively. Higher decay in the slope of hydrated tRNA on ND surface compared to other samples implies faster dynamics of tRNA in the presence of ND.

Fig. 4.6 a shows the experimental ISF at Q = 1.1 Å$^{-1}$ and T = 260, 280 and 300K. Hydrated tRNA experiences faster dynamics in the presence of ND at 260K, 280 and 300K. We estimated the relaxation, $\tau$, from an intuitive way, by taking the relaxation time at ISF=0.92. The results are shown in Fig. 4.7-4.8 for different temperatures and Q. Figure 4.7 a represent $\tau$ as a function of inverse T and Q respectively for RNA+ND+D$_2$O and figure 4.7 b represent the same for RNA+D$_2$O (i.e absence of ND) samples. At all different Qs and temperatures, the relaxation time $\tau$ is 2-4 times faster in the presence of the ND surface (Fig. 4.7 a and 4.8 a). While $\tau$ shows Vogel-Fulcher-Tammann (VFT) type decay (curves) in the presence of ND, a Arrhenius type decay (straight line) is observed in the absence of ND, resembling a ‘fragile’ (Fig. 4.7 a) and strong (Fig. 4.7 b) glassy behavior respectively. These resemble a ‘weak’ (Fig. 4.7 a) and ‘strong’ (Fig. 4.7 b) glassy behavior respectively. We refer these as a ‘jammed’ state, as the
system is not ‘naturally glassy’ but the large macromolecular motion in the presence of D₂O hinders the dynamics. The tRNA dynamics is a consequence of ‘weak’ jamming, which is considered to be responsible for the faster dynamics in the presence of ND. In Figure 4.9, the scaling shows a stronger Q-dependence of \( \tau \) compared to the smaller macromolecular counterparts [95-98]. In macromolecules [95-98], the Q-dependence in the Q-range < 1 Å⁻¹, follows a power law, \( \tau(Q) \sim Q^{2/\beta} \), where \( \beta \) is the stretching exponent ranging between 0.4-0.65. Fig. 4.9 shows a much slower decay (fit not shown here) with \( \beta \) ranging from 0.32-38 and 0.26-32 for ND+RNA+D₂O and RNA+D₂O respectively. Therefore, the tRNA dynamics becomes slower compared to typical polymers in both cases. As a general tendency, the presence of ND typically should slow down the motion of RNA as has been observed on a silica surface [74]. However, counterintuitive an ‘enhancement’ of the dynamics of the highly interactive hydrated RNA is observed on ND surface in the present study. This phenomenon is critical to the motion of large biomolecules in the presence of a functionalizable nanoparticle, in this case ND. In Fig. 4.9 b, we also noticed a step jump of \( \tau \) within the Q-range 1.1-1.3 that relates to 4.8 to 5.7 Angstrom length scale. tRNA motion is observed faster at > 5.7Å length scales and slower at the shorter length scales. The same is observed until \( T = 260K \) in ND+RNA+D₂O case (Figure 4.9 a) but with a kink-jump at \( Q = 1.1\text{Å} \). Therefore, the molecular motions of the hydrated only tRNA are not homogeneous at all the length scales; however the same gives a molecular scale intrinsic non-exponential relaxation in the presence of ND. This critical feature of tRNA dynamics can be related to heterogeneous dynamics of the large biomolecules on two different length scales in the absence of ND [99, 100].
Fig. 4.7 a and b Experimental relaxation time $\tau$ of hydrogen of RNA is plotted as a function of inverse temperature in a ND+tRNA+D$_2$O and tRNA+D$_2$O system at six different $q$-values respectively. The cusp shape of the curve shows that ND+tRNA+D$_2$O follow Vogel-Fulcher-Tammann law. The straight line shape of the curve shows that tRNA+D$_2$O follow Arrhenius law.

Fig. 4.8 a and b Experimental relaxation time $\tau$ of hydrogen of tRNA in ND+tRNA+D$_2$O and tRNA+D$_2$O system are plotted as a function of $q$ at eight different temperatures, $T = 240$ to $310$ K with $10$ K interval.

To further investigate the underlying mechanism of the tRNA motion, we carried out all-atom MD simulations of the D$_2$O hydrated RNA in the presence or absence of ND. It has already been established [101] that tRNA can be represented by hammerhead RNA to accurately explain the neutron scattering results from tRNA [101]. We also used hammerhead RNA in the presence
and absence of ND in our simulations [102] with experimental parameters. The simulation results are shown in Fig. 4.6 b. The simulated ISFs also show a faster dynamics of the D$_2$O hydrated RNA on ND surface and hence in qualitative agreement with experimental results.

In Fig. 4.9, we present the scaling analysis of the simulated transport coefficients, which provides a better understanding of the critical dynamics of hydrated RNA on ND. To answer “why does RNA exhibit faster dynamics on a ND surface?”, as observed in our QENS experiments, we plotted the standard numerical quantity for the validation of Stokes-Einstein Relation (SER), Dτ/T, as a function of 1000/T and D (where D is diffusion constant) as a function of τ shown in Fig. 4.9 at short, medium and long length scales, Q = 0.5, 1.1 and 1.9 Å$^{-1}$ respectively. Regions of highly mobile molecules move through less mobile molecules, which has consequences in the validity of SER[103, 104] given by, D ~ τ$^{\alpha}$, where ‘α’ is the exponent. For α = 1.0, the dynamics follow normal Fickian diffusion i.e., no violation of SER. To follow SER, Dτ/T, must be constant with 1/T. Fig. 4.9, the normalized Dτ/T (normalized to the highest temperature value) shows a large deviation of the RNA from unity (plotted with a black line at 1.0) without ND; however, the presence of ND reduces the deviation. Therefore, the RNA shows a strong violation of SER without ND and a weaker violation with ND. Water, on the other hand, shows a smaller deviation representing the weakest violation of SER. The RNA Dτ/T always exhibits strong violation of SER. The violation of SER is related to the heterogeneous dynamics (HD) of the molecule that is related to ‘highly mobile’ and ‘immobile’ motion relative to the average motion. The stronger SER violation represents stronger heterogeneity in the dynamics leading to a lesser mobile molecular relaxation as in the case of “freestanding” RNA. On the other hand, a weaker SER violation leads to a faster dynamics as is observed in the presence of ND. The same was confirmed by the experiments in Fig. 4.8 a and b. As the length
scale decreases (from Fig. 4.9 a to c) the absolute value of $D\tau/T$ reduces, indicating a reduction of heterogeneity in molecular motion. This kind of heterogeneous dynamics is well known in biomolecules in living systems [100, 105, 106]; therefore, the observed HD falls under the generalized characteristic motion of bio-macromolecules. The importance of the MD simulation is that it allows critical analysis of the D$_2$O dynamics that cannot be detected by neutron experiments.

![Fig. 4.9 a - c Comparison of Stokes-Einstein Relation (SER) $D^*\tau/T$ of Hydrogen of RNA (H$_x$RNA) and Deuterium of Water (D$_2$O) at three q-values 0.5, 1.1 and 1.9 Å$^{-1}$. Black circles and red squares represent hydrogen of RNA of ND+RNA+D$_2$O and RNA+D$_2$O samples; green diamonds and blue up-triangles represent D$_2$O of water for ND+RNA+D$_2$O and RNA+D$_2$O samples respectively.](image)

To show the breakdown of SER breakdown, $D \sim \tau^\alpha$ in a rigorous way, we plotted $D$ versus $\tau$ in Fig. 4.10 and 4.11. From Fig. 4.10, RNA dynamics in the presence of ND (black symbols) shows faster dynamics than without ND (red symbols). The plot shows two different scaling behaviors. The red symbols scales with $\alpha = 0.1$-0.2 whereas the black symbols scales with $\alpha = 0.27$-0.52. Fig. 4.11 shows scaling behavior of water varying from 0.6-0.66 without ND (red symbols) to 0.58-0.78 with ND (black symbols). Therefore, the hydrated RNA shows a stronger ‘glass-type’ behavior without ND. The master curve shown in Fig. 4.10 represents the universal scaling dynamics of D$_2$O, similar with our without ND although the diffusion coefficient is
higher in the former case. Such a violation is almost absent or weaker in strong liquids consistent with the idea that decoupling is related to heterogeneous dynamics that are absolutely important for fragile liquids [107, 108]. The D$_2$O dynamics, therefore, is always decoupled from RNA dynamics and shows a ‘fragile’ nature compared to the RNA. The $\alpha$ values are plotted in Fig. 4.12. It shows small differences for water and is always ‘fragile’, falling within the dashed region (close to $\alpha = 1.0$). However, there is a large difference in RNA $\alpha$ with or without ND. The $\alpha = 0.1$-0.2 range represents a much stronger ‘glassy’ (or jammed) behavior without the ND that fundamentally restricts the motion of RNA leading to a slow decay in the ISF (Fig. 4.6 b).

**Fig. 4.10** Diffusion coefficient ‘D’ is plotted as a function of $\tau$ for Hydrogen of RNA at three different Q values, 0.5 (circle), 1.1 (square) and 1.9 (up triangle) respectively. The black and red symbols represent with (ND+ RNA+D$_2$O) and without (RNA+D$_2$O) ND samples respectively.
Fig. 4.11 Diffusion coefficient ‘D’ is plotted as a function of τ for deuterium of D₂O at three different Q values, 0.5 (circle), 1.1 (square) and 1.9 (up triangle) respectively. The black and red symbols represent with (ND+ RNA+D₂O) and without (RNA+D₂O) ND samples respectively.

Fig. 4.12 Scaling exponent, α as a function of Q representing the values at different length scales. The α values are derived from fig. 4.10 and 4.11 using a non-linear least-squares Marquardt-Lavenberg algorithm. The dashed-line at 1.0 represents normal Fickian diffusion.
Fig. 4.13 Left and right images are the simulated snapshots of RNA+D$_2$O and ND+RNA+D$_2$O system from MD simulations respectively.

Fig. 4.13 left and right show simulation snapshots of the RNA+D$_2$O and RNA+ND+D$_2$O samples, respectively from MD simulation data. The radial distribution function, g(r), of water hydrogen (H) and RNA hydrogen (H) is plotted in Fig. 4.14-4.17 and extends to 10 Å from the surface (25Å) of the ND in Fig. 4.14 and 4.16, and 10 Å from the RNA surface (0Å) in Figure 4.15 and 4.17. The corresponding number of water molecules is shown on the right. In Fig. 4.14 and 4.16, water and RNA both are observed on the ND surface in RNA/ND/D$_2$O sample at 260 and 300K. However, water profusely hydrates the ND surface, (~3000 molecules Fig. 4.14) compared to RNA (~400 molecules Fig. 4.16) on the ND surface. While the large number of water molecules (~3000) surrounds the ND surface, the number of water molecule around the RNA is negligibly low, around ~120 molecules around, as shown in Figure 4.15. Also, the g(r) peak, probability of finding a water H on ND surface is twice as high as the probability of finding a RNA H on ND surface or a water H around RNA H. Fig. 4.15 shows similar number (~around 120) of water H around RNA H in the absence of ND (as Fig. 4.17); however with a weaker g(r) peak. Considering the natural water H distribution around RNA H without any
surface interactions, the presence of hydrophilic ND surface shows ten times (8.0 in Fig. 4.15 compared to 0.8 in Fig. 4.17) higher probability of finding water H around RNA. The presence of large number of water on ND surface and higher peak in g(r) of RNA H on the surface is possible only if there is an “interfacial layer” between the ND and the RNA that makes water accessible to both the RNA and the ND surfaces there by forming a ND-corona complex[72]. In Fig. 4.14, g(r) of water around a ND center shows three peaks at 27Å, at 30Å and 32.5Å. The peaks for the RNA are at 26.9, 29 and 31.5 Å respectively. While the first peak of RNA and water H are located at roughly the same position, the second and third peaks differ by ~1Å. This represents de-confinement of RNA on the far side of the ND. The first peak can be looked on as a near-monolayer of RNA molecules ("hard" corona) which binds tightly to the ND surfaces; while the second peak represents a "soft" corona on top of the "hard” corona, which is more loosely associated to the "interfacial layer”. Because of the loose structure and rapid exchanging feature of the "soft" corona, the otherwise attractive interaction between RNA and ND is weakened due to a high-density interfacial water layer. Moreover, it was observed in Fig. 14.9 that the RNAH violates SER with a weaker heterogeneity on ND surface compared to a freestanding RNA leading to a faster RNA motion on ND surface. The weaker heterogeneous dynamics of the RNA leading to their faster dynamics, while having the same number of water molecule around RNA with or without ND, can only be explained if the interfacial water layer is also ‘de-confined’ from the RNA. The existence of large number of water on the ND surface and the interfacial water layer is a result of the combination of de-confinement of RNA and a strong hydrophilic interaction of water with the ND surface.
**Fig. 14.14** Radial Pair distribution function, $g(r)$ (left axis) and number of water molecules (D$_2$O) from the surface of ND are plotted as a function of distance in ND+RNA+D$_2$O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules (right axis) for the respective temperatures.

**Fig. 14.15** Radial Pair distribution function, $g(r)$ (left axis) and number of water molecules (D$_2$O) (right axis) around RNA are plotted as a function of distance in ND+RNA+D$_2$O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules for the respective temperatures.
Fig. 14.16 Radial Pair distribution function, $g(r)$ (left axis) and number of hydrogen of RNA around ND are plotted as a function of distance in ND+RNA+D$_2$O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules (right axis) for the respective temperatures.

Fig. 14.17 Radial Pair distribution function, $g(r)$ (left axis) and number of water molecules around RNA are plotted as a function of distance in RNA+D$_2$O system. The magenta and blue colors represent temperature at 260 and 300 K respectively. The dashed lines represent number of water molecules (right axis) for the respective temperatures.
4.6 Conclusion

We conclude that the underlying mechanism for RNA dynamics on the ND surface is twofold: structural and dynamical. RNA exhibits a weaker heterogeneous dynamics on a ND surface than a freestanding hydrated RNA. The weaker heterogeneity is facilitated by a weaker SER violation leading to a ‘fragile’ jammed stated compared to freestanding RNA that show a ‘strong’ jammed states facilitated by a strong SER violation. Structurally, water molecules form an interfacial layer between RNA and ND with a high probability of water molecule around RNA on ND surface. This gives rise to a de-confinement of RNA molecule with weaker heterogeneous dynamics. A de-confined RNA molecule leads to a faster relaxation dynamics consistent with SER violation. The water molecule, on the other hand, shows ‘weak’ SER violation and follows a universal scaling law.

Hydrated RNA exhibits faster dynamics on a ND surface compared to a freestanding case by using QENS and MD simulation. The fragility of the RNA is suppressed in the absence of ND leading to a higher order breakdown of SER. The SER breakdown is weaker in the presence of ND. At the same time, the water molecule from the hydrated RNA is released on hydrophilic ND surface to form an interfacial water layer that leads to a de-confinement of RNA giving rise to a faster RNA. The simulations are in ‘qualitative’ agreement with the experimental results. The techniques used here allow a precise determination and the inference of a detailed explanation of RNA dynamics and structures in bio-macromolecules nanocomposites. The combination of faster dynamics and de-confinement of RNA on ND surface provides unique opportunities to enhance the drug-delivery mechanism(s) in RNA nanotechnology by introducing a small portion of non-toxic, highly functionalizable ND. ND can also be integrated with other hydrophilic
biomolecules and enhance the properties of materials for bactericidal activity, drug-delivery and even for treating viral diseases.
CHAPTER 5 DYNAMICS OF AN INTRINSICALLY DISORDERED PROTEIN: β-CASEIN

5.1 Introduction

Bovine milk approximately contains 3.0-3.5% protein where the concentration and composition can change during lactation. The properties of milk mainly depend on the proteins; although fat, lactose and salts are also play an important role. The main function of milk is to provide the essential nutrients, for instance, calcium that is required for the growth and development of infants. Initially, it was believed that milk is a simple homogeneous liquid, but later on, it was found that milk contains colloidal particles that are composed of complex associated proteins and calcium phosphates. Milk proteins were divided into two categories: Casein and Serum proteins. These heterogeneous associated colloids are named as "casein micelles [109]. Casein proteins play an important role on keeping the milk to highly stable at the high heat treatments for various dairy products. This makes the casein micelles to be highly studied for long time. It occupied specific position among biological systems because different models were proposed to explain its structure such as coat-core, sub-unit and internal structure models [110].

Casein is a group of unique - specific proteins which constitutes approximately 80% of the total protein in the cow milk and other dairying species. Basically, they exist in the form of stable colloidal particles in the spherical shape of diameter in the range of 50-300 nm, and having mass 2.2 \times 10^{-15} g [109, 111, 112]. It is composed of four main types of proteins: α_{S1}-casein (CN), α_{S2}-CN, β-CN and κ-CN where they present in milk in the molar ratio of 4: 1: 4: 1 respectively [113]. Among them β-CN is the major source to provide protein to the infants growth. It is found that the growth of pups are reduced on giving them β-CN absent in milk
All of them differ in structure, type and degree of post-translational modification [115]. αs-CN proteins are the major casein proteins that contain the 8-10 seryl phosphate groups while β-CN contains about 5 phosphoserine residues. αs-CN and β-CN are highly phosphorylated proteins so they become high calcium ion binding proteins [115]. Out of them κ-CN has only one phosphoseryl residue and κ-CN is also glycosylated. In addition, casein proteins contain large amount of propyl residues especially in β-CN. The presence of large amount of propyl residues is responsible to affect the structure of casein because these proline residues interrupt the formation of α-helical and β-sheet. Moreover, the casein proteins have different hydrophobic and hydrophilic regions along the protein chain, and among them β-CN is the most hydrophobic in nature.

β-CN has been the subject of research since a long time because of its versatile properties such as formation and stabilization of food emulsions [116], binding sites of retinol and retinoic acid in milk [117], amphiphilic and self-assemble into stable micellar structure in aqueous solution [115]. As β-CN is a natural milk protein whose micellar structure is analogous to di-block copolymers. It is expected that β-CN micelles are more stable compared to low molecular weight surfactant micelles. Moreover, hydrophobic nature of β-CN attracts great attention to be used as a drug carrier for oral delivery system in cancer treatments. More stability in its structure decreases the chances of early release of drug than low molecular weight surfactants [118]. Therefore, the efficiency, capacity and stability of β-CN enable it as a good entrapment in using as a drug carrier for oral delivery system [119-121]. Furthermore, its open tertiary structure makes it to be easily accessible for gastric proteases. Hence, it may be used as an oral drug delivery carrier for releasing the drug in stomach for the treatment of gastric carcinoma. β-CN
hydrophobic property makes it to use for many hydrophobic chemotherapies such as mitoxantrone, vinblastine, irinotecan, docetaxel and paclitaxel [122].

Although caseins have important biological functions, they lack secondary structures and are recognized as "natively disordered" proteins [113]. There are various models proposed by different research groups to explain the internal structure of casein micelle and its components, however no general agreement has been met [123]. Small angle neutron scattering (SANS), small angle X-ray scattering (SAXS) and static light scattering (SLS) techniques were used to compare the different models to explain the internal structure of casein micelles [124]. Their SAXS study on casein micelles report “Yet their structure is still a puzzle and is continuously the subject of furious debates among the scientific community” [124, 125]. A hypothesis about the function of "natively disordered" proteins like casein [123, 126] is that it might need a high structural flexibility to perform their function [113, 127]. The flexibility of protein, or protein "softness" [128], is generally known to be essential for their biological activities and is believed to be influenced by its hydration water and other solvents [90, 129-134]. Some recent studies suggest that there are both solvent dependent and solvent independent fluctuations in the protein dynamics. Some aspects of protein dynamics are considered to be "slaved" to the solvent fluctuations (α-fluctuations), while some additional local protein motions (β-fluctuations) are independent of the solvent fluctuations [135-139].

In this study, we use both elastic neutron scattering (ENS) and quasielastic neutron scattering (QENS) to investigate the temperature-dependent relaxation dynamics of both hydrated and dry β-CN. ENS provides information about mean square displacement (MSD) of hydrogen atoms within protein molecules, which is traditionally used as a measure of the protein flexibility [128]. QENS has been proved to be a prevailing tool to study protein dynamics in the time scale of
The hydrogen atom has large incoherent scattering cross-section compared to other atoms present in proteins. In QENS study of protein dynamics, the incoherent scattering signal dominates only from non-exchangeable hydrogen atoms [9] which belongs to following groups: backbone hydrogen (H), methyl hydrogen (CH₃), methylene hydrogen (CH₂) and hydrogen atom from aromatic rings [139]. Therefore QENS intensity selectively characterizes hydrogen atoms' motions in this protein. In this paper, QENS results are discussed in both energy domain and time domain. In energy domain, a model-independent analysis is employed to study the QENS spectra at different temperatures. In time domain, the Fourier transform of the QENS spectra, intermediate scattering functions (ISF) of hydrogen atoms in both dry and hydrated β-CN are analyzed by an asymptotic formula developed from mode coupling theory (MCT) [76, 88, 140], approving that β-CN undergoes a logarithmic decay in its dynamics which has been observed in other well-folded proteins [76, 141]. The temperature dependence of the fitting parameters is consistent for both analyses. These results provide evidence that a "natively disordered" protein β-CN has a higher structural flexibility in its dynamics, which may result in an easier micellization [142] and thermal unfolding (or so-called reversible thermal denaturation) at lower temperatures compare to other well-folded proteins and rigid biomolecules [76, 143].

5.2 Materials and Methods

Lyophilized β-casein (β-CN) was purchased from Sigma-Aldrich without further purification. For comparison purpose, both dry and hydrated protein samples were measured. The lyophilized protein powder was hydrated with D₂O vapor inside a glove box for 24 hours. The final hydration level of casein was 0.35 g of D₂O per gram of casein. This hydration level is chosen to
make sure the protein has a monolayer of D$_2$O covering its surface to maintain its activity [144].

The hydration level was measured by directly measuring the weight of absorbed D$_2$O.

QENS experiment was performed on high flux back-scattering spectrometer (HFBS) at National Institute of Standards and Technology (NIST) Center for Neutron Research (NCNR) [19]. In elastic mode, elastic intensity was recorded from 330K down to 4K with a ramp rate of 1K/min. In the quasi-elastic mode, the energy resolution is 0.8 μeV (full- width at half maximum, for the Q- averaged resolution value) and accessible dynamic range is ±17μeV. The hydrated sample was measured at seven temperatures from 220 K to 340 K with 20 K interval, and the dry casein sample was measured from 260 K to 340 K, 20 K interval. The resolution function was measured at 4 K where the entire signal is expected to be completely elastic in nature.

The QENS spectra $S_m(Q,\omega)$ is measured by using equation 4.1 and modeled with equation 4.2. We found a single Lorentzian was appropriate to fit the data in entire temperature range and for all Q values which is given by:

$$S(Q,\omega) = \frac{1}{\pi} \frac{\Gamma(Q)}{\omega^2 + \Gamma^2(Q)}$$

(5.1)

The data fitting was done by the peak analysis software PAN in package DAVE developed by NCNR [145].

To study the dynamics of β-CN in time domain, ISF is calculated by using equation 4.4. ISF is fitted with logarithmic model (equation 4.6).
5.3 Results and Discussion

5.3.1 Mean Square Displacement (MSD) and Protein Flexibility

The "softness", or the flexibility of the protein, can be quantified by the mean square displacement (MSD) $\langle x^2(T) \rangle$ of hydrogen atoms in protein molecules [128, 146, 147]. The MSD is calculated from the measured incoherent elastic neutron scattering (IENS) intensity using Debye-Waller factor derived from Gaussian approximation

$$I_{el}(Q,T,\omega = 0) / I_{el}(Q,T = 4 K, \omega = 0) = \exp[-Q^2\langle x^2 \rangle / 3]$$

(5.2)

where $I_{el}(Q,T,\omega = 0)$ is the temperature dependent elastic scattering intensity.

![Graph showing mean square displacement](image)

**Fig. 5.1** The mean square displacement (MSD), $\langle x^2(T) \rangle$ of dry $\beta$-CN and D$_2$O hydrated $\beta$-CN plotted as functions of temperature. The dynamic transition is observed at T~ 225 K in hydrated $\beta$-CN. Error bars are within the size of symbols.

The measured MSDs $\langle x^2(T) \rangle$ for both dry and hydrated $\beta$-CN are plotted in Fig. 5.1 as functions of temperature. Neither of the curves show sharp drop at the freezing point (273 K),
indicating no ice formation in our measured temperature range for either dry or hydrated samples. Two onsets inflected in the MSD slopes are observed in the hydrated sample. The first onset occurs at T~100 K and is common for both samples. This an-harmonic behavior has been observed in a number of proteins [90, 148] and ascribes to onset of methyl group rotations. Interestingly, both samples show exact same MSDs until about 200 K indicating methyl group dynamics of β-CN is independent of hydration which has been found to be the case for other proteins as well [91, 149-151]. The second increase in the slope at T~ 225 K is only observed in hydrated protein. Although a very recent study [152] demonstrate that the MSD of dry bovine serum albumin (BSA) powder exhibits the change in the slope in the vicinity of 240-260 K, our results do not show the same transition in dry β-CN as T-dependence of MSD slope remains unchanged in the above T-range. A possible reason is due to their much different energy window (~meV for time-of-flight spectrometer) comparing to our backscattering spectrometer with μeV energy window. In meV range, instrument is more sensitive to localized fast motions rather than conformational fluctuations. Furthermore, BSA is a well-folded protein and is much more rigid than the natively disordered β-CN. Thus hydration effect can be more prominent on β-CN molecules.

The second change in MSD slope is highly hydration dependent [90] and is generally recognized as the dynamic transition. This transition involves backbone fluctuations [139, 153], and activation of hydrophilic side chains which are strongly coupled to the relaxation rates of the hydrogen bonds that they form with hydration water [154, 155]. In some recent studies [156-158], it is suggested that the protein dynamic transition, shown as the an-harmonic onset in the MSDs depends on the energy resolution of the spectrometer. Here, our results clearly demonstrate that the local and global dynamics of proteins are affected by the hydration. While β-CN has both
hydrophobic and hydrophilic regions [115], the sharp increase in MSD of hydrated at 225 K represents the hydrophilic region. No such transition or an-harmonic behavior is observed in MSD of dry β-CN at 225 K, indicating that hydrophilic residues in β-CN protein are highly hydration dependent.

5.3.2 Analysis of QENS Data in Energy Domain

The normalized QENS spectra for both dry and hydrated β-CN at different temperatures are plotted in Fig. 5.2 (a) and (b) in log scale. The broadening of central peaks from the resolution indicates the quasi-elastic scattering of hydrogen atoms in the protein molecules. The central peaks become broader as the temperature increases, implying faster dynamics of hydrogen atoms in the β-CN at higher temperatures. Comparing Fig. 5.2 (a) and (b), the central peaks for hydrated sample have more broadenings than that of the dry sample which implies more dynamics in the hydrated sample. Fig. 5.2 (c) and (d) illustrate the analysis of QENS data by DAVE [145]. A single Lorentzian fitting agrees with the measured QENS data satisfactorily. The fitting parameters for dry and hydrated samples at the same temperatures (shown in Fig. 5.3, Fig. 5.4 and Table 5.1) demonstrate that hydrated β-CN has much larger quasi-elastic scattering component than the dry one, while their elastic scattering components are comparable. This fact is consistent with the observation of the raw data shown in Fig. 5.2 (a) and (b).
Fig. 5.2 (a) and (b) Normalized QENS spectra of dry and hydrated β-CN at different temperatures from 220 K to 340 K and 260K to 340K respectively at $Q = 0.9907 \text{ Å}^{-1}$. (c) and (d) Analysis of the QENS spectra of dry and hydrated β-CN. The circles are the experimental data; the blue lines represent the fitted curves. The orange and magenta lines represent quasi-elastic scattering component and elastic scattering component respectively. Error bars throughout the text represent one standard deviation.

The elastic incoherent structure factor (EISF), defined as the ratio of the elastic intensity to the total intensity, i.e. $A_0$ in Eq. (4.2), is plotted as a function of $Q$ in Fig. 5.3. EISF provides the information of the fraction of atoms that are mobile and immobile in the range of spectrometer. It contains the information on geometry of motion of atoms exhibits translational or rotational motion and represents the probability that a particle can be found in the same volume of space at some subsequent time [159, 160]. As the QENS data is fitted by using a single Lorentzian, we expect the EISF curves to be fitted well with the expression of single diffusive motion of atoms. As this diffusion is expected to be isotropic, we started with diffusive motion within a sphere.
expression, which is given by \((3j_1(Qa)/(Qa))^2\), where \(j_1\) denotes the spherical Bessel function of the first kind of order 1. In the above expression ‘\(a\)’ is the radius of the diffusive sphere, in which atoms diffuse freely. The complete model can be written as [161]

\[
EISF = p_0 + p_1((3j_1(Qa)/(Qa))^2)
\]

(5.3)

where \(p_0\) denotes the immobile fraction of protons and \(p_1\) denotes the mobile fraction of atoms and \(p_1 + p_0 = 1\). We use the above model to fit EISF of both dry and hydrated samples, to better compare the two cases with the same set of parameters. Although in recent literature [6, 90, 162] it is suggested to fit the EISF of the dry sample with a model for methyl rotations, the difference between two analyses are indistinguishable in our measured Q and energy range for the low hydration and dry samples [90]. Moreover, our current data show a much smaller mobile fraction of hydrogen atoms participates when only methyl group contribute to the dynamics which justifies the approach that we have used in this study. The fitted curves of EISF are shown as solid lines in Fig. 5.3.

![Fig. 5.3 The EISF of (a) hydrated and (b) dry β-Casein at different temperatures fitted according to eq. (5.8). The solid lines are the fitted curves. The error bars of the data points are within the symbols if not shown.](image)
The mobile population of atoms undergoes the diffusive motions in spheres as the EISF is reasonably fit with corresponding model (eq. 5.3) as can be seen in Fig. 5.3. The deviations on low Q side could be due to some small multiple scattering contributions to the observed data. The fitting parameter $p_0$ is plotted in Fig. 5.4 as a function of temperature for both dry and hydrated samples. Apparently the population of the immobile hydrogen atoms, $p_0$, is much lower for hydrated sample than the dry one at all temperatures, indicating that when hydrated, more hydrogen atoms are observable within our measurement dynamic window. Moreover, the decrease in $p_0$ value with increase in temperature indicates that more hydrogen atoms are participating in motion as temperature goes up. The parameter $a$, representing the radius of the diffusive sphere is shown in Table 5.1 for both samples. It is clear that for hydrated sample with increasing temperature, both the fraction of mobile atoms ($p_1$) and the radius of sphere ($a$) increase significantly, indicating that more hydrogen atoms undergo diffusive motions with larger amplitudes with temperature increases. Such temperature dependence of parameters is not
clearly evident for the dry sample. The stronger temperature dependence of mobile fraction and radius of sphere together with the fact that mobile fraction is much smaller in dry sample clearly show that motions in hydrogen atoms in methyl groups and side chains together are highly hydration dependent. Additionally, in Table 5.1, the sphere radius 'a' has a value of ~3 Å for both dry and hydrated samples, which is much larger than the radius of a methyl group (~1 Å). Thus the measured dynamics in β-casein demonstrates primarily slow relaxation process. Considering the bovine β-casein lacks tertiary structure and its primary structure suggested being amphiphilic in nature [141], our results reveal that the increase in hydration of β-casein leads to an increase in the flexible regions in the protein, which is crucial for β-caseins to form micelle-like structures.

Interestingly, when temperature goes even higher, say at 320 and 340 K, a decrease in volume of diffusion is observed (decrease in the radius a). It might be due to the partially thermal unfolding (or so-called thermal denaturation) of the proteins at these higher temperatures [163, 164]. The partially thermal unfolding is suggested to correlate with the relaxation in the hydrogen bond network formed between protein and its hydration water [143, 154, 163, 165]. Upon thermal unfolding, the hydrogen bond network collapses and the flexibility of the protein is reduced. The number of hydrogen atoms with diffusive motion (p₁) increases markedly, given that more hydrogen atoms are able to diffuse than in the native state; while the radius of the sphere in which the hydrogen atoms move is reduced, because of less flexibility in the protein molecule. These results are consistent with the previous observation of thermal denaturation (heat-induced unfolding) of protein in solution [163] and suggest the thermal unfolding of hydrated β-casein at 320K and 340K.
Table 5.1 The fitting parameter \( a \) at different temperatures for both dry and hydrated \( \beta \)-Casein

<table>
<thead>
<tr>
<th>Temperature (K)</th>
<th>( a_{dry}(\text{Å}) )</th>
<th>( a_{hydrated}(\text{Å}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>220</td>
<td>N/A</td>
<td>2.39 (± 0.25)</td>
</tr>
<tr>
<td>240</td>
<td>N/A</td>
<td>2.41 (± 0.21)</td>
</tr>
<tr>
<td>260</td>
<td>3.30 (± 0.25)</td>
<td>2.83 (± 0.20)</td>
</tr>
<tr>
<td>280</td>
<td>3.21 (± 0.35)</td>
<td>3.08 (± 0.19)</td>
</tr>
<tr>
<td>300</td>
<td>2.05 (± 0.21)</td>
<td>3.37 (± 0.38)</td>
</tr>
<tr>
<td>320</td>
<td>3.67 (± 0.23)</td>
<td>2.56 (± 0.29)</td>
</tr>
<tr>
<td>340</td>
<td>2.44 (± 0.33)</td>
<td>2.89 (± 0.17)</td>
</tr>
</tbody>
</table>

5.3.2 Time Domain Analysis

The intermediate scattering function (ISF) \( F_H(Q,t) \), referred to as density-density correlation function, is a primary tool to unravel the relaxation dynamics in the protein molecules. In this paper, the ISF of both samples are analyzed by an asymptotic expression derived from Mode Coupling Theory (MCT) (Eq. 4.6). The MCT is a theory originally developed to describe glass-forming liquids [77, 145]. Here it is employed to describe the dynamics of proteins because proteins and glass forming liquids share common dynamical behaviors in many aspects [140, 148, 166-170]. The MCT has predicted a logarithmic decay in the \( \beta \)-relaxation range of protein dynamics and has been proven to be efficient in fitting protein dynamics in time domain both by molecular dynamics (MD) simulations [88] and experiments [75]. In Fig. 5.5, ISF of the hydrated \( \beta \)-CN is analyzed according to Eq. (4.4) at six temperatures and ten \( Q \)-values. Four fitting parameters, \( A(T) \), \( \tau_\beta(T) \), \( H_1(Q,T) \) and \( H_2(Q,T) \) are obtained by fitting curves at all ten \( Q \)-values together, where \( A(T) \) and \( \tau_\beta(T) \) are \( Q \)-independent parameters, \( H_1(Q,T) \) and \( H_2([76]Q,T) \) \( Q \)-dependent parameters. The fitting results show that the value of \( A(T) \) is very close to 0. This is because the pre-factor \( f(Q,T) \) goes to 1 at all \( Q \)-values at a specific short time \( \tau_\beta(T) \sim 10 \text{ ps} \), which is much shorter than our measured time range (above 100 ps).
Fig. 5.5 Analysis of the ISF of hydrated β-CN in the β-relaxation region at 10 different Q-values. The upper three panels show the results at lower temperatures $T = 220$, 260, and 280K; the lower three panels show at higher temperatures $T = 300$, 320, and 340 K. The solid lines represent the fitted curves.

Fig. 5.6(a) compares the Q-dependent parameter $H_1(Q,T)$ of both dry and hydrated β-CN. $H_1(Q,T)$ represents quantitatively the slope of the decay, which can be observed qualitatively in Fig. 5.5 and Fig. 5.6 (b). For the dry sample, $H_1(Q,T)$ hardly show any Q or T dependence, due to lack of hydrogen motions detected in our measurement time window. For the hydrated sample, $H_1(Q,T)$ increases with Q and follows a power law in Q at small Q-values, i.e. $H_1(Q,T) = B_1(T)Q^\beta$, where $\beta$ is a value between 1 and 2 and $B_1(T)$ is a temperature dependent parameter, which is plotted in Fig. 5.7 (a). Fig. 5.6(b) shows the comparison of relaxational dynamics of hydrated and dry β-CN at $Q = 0.99$ Å$^{-1}$ at three different temperatures in time domain. It is clearly shown that the slope of the decay $H_1(Q,T)$ of the dry sample is much smaller than the
hydrated sample, which means the relaxation dynamics of dry β-CN is much slower than hydrated one and is much less Q or temperature dependent. This provides another evidence that the dynamics of β-CN is highly hydration dependent and solvent plays an important role in the activation of proteins [137].

![Figure 5.6](image)

**Fig. 5.6** (a) Analysis of the fitting parameter $H_1(Q,T)$ of both samples as a function of Q at different temperatures. For the hydrated sample (solid symbols), $H_1(Q,T)$ is fitted by a power law, $H_1(Q,T) = B_1(T)Q^\beta$ at temperatures $T = 220, 240, 260, 280, 300, 320, 340$ K. For the dry β-CN (open symbols), $H_1(Q,T)$ does not show much temperature or Q dependence; (b) Comparison of the ISF(Q,t) of hydrated and dry β-CN at three different temperatures, $T = 300, 320, 340$ K. The open symbols represent data belongs to dry β-CN, while solid symbols belong to hydrated β-CN. The error bars of the data points are within the symbols if not shown.

The temperature dependent $B_1(T)$ is fitted linearly in Fig. 5.7(a). The characteristic relaxation time $\tau_\beta(T)$ is plotted versus $1000/T$ (the so called Arrhenius plot) in Fig. 5.7(b). Below 300K, $\tau_\beta(T)$ can be fitted with the Vogel-Fulcher-Tammann (VFT) law, $\tau_\beta(T) = \tau_1(T)\exp[DT_0/(T-T_0)]$, where D is the dimensionless parameter providing measurement of fragility and $T_0$ is the ideal glass transition temperature. According to the glass transition theory, the VFT behavior typically represents α-process. Such apparent non-Arrhenius behavior of relaxation time is a common
feature in the dynamics of biomolecules and has been observed in many previous works [75, 141, 158, 171]. Above T~300K, the temperature dependence of $\tau_\beta(T)$ does not follow the VFT law anymore. It switches to an Arrhenius-like behavior. Since it is hard to determine whether this behavior is an Arrhenius behavior due to limited data points (3 points here), we use a straight line to fit the data points as guide of the eyes (the green dashed line in figure 5.7(b)). Notice that this switch in $\tau_\beta(T)$ at 300K is consistent with our analysis in energy domain shown in Fig. 5.4 and Table 5.1. This high-temperature transition in hydrated $\beta$-Casein is possibly related to the reversible thermal denaturation or thermal unfolding of the protein, which has been observed in hydrated lysozyme and its hydration water in previous experiments [141, 143] and has also been observed in hydrated tRNA [75]. It is highly hydration dependent and is suggested to correlate with the relaxation in the hydrogen bond network formed between protein and hydration water [143, 154, 165]. The lower transition temperature in $\beta$-Casein (300K) comparing to other biomolecules (320K for tRNA [75] and 345K for lysozyme [143]) indicate that lack of secondary structure will result in thermal unfolding or thermal denaturation of $\beta$-Casein at relatively lower temperatures compare to other well-folded proteins and rigid biomolecules. This observation confirms that natively disordered $\beta$-Casein has relatively higher structural flexibility compare to well-folded proteins.
Fig. 5.7 (a) Linear fitting of the parameter $B_1(T)$ as a function of temperature for hydrated $\beta$-CN. (b) Arrhenius plot of the relaxation time $\tau_\beta(T)$ as a function of inverse of temperature in log scale.

5.4 Conclusion

In summary, the relaxation dynamics of a natively disordered protein $\beta$-Casein has been studied by quasielastic neutron scattering (QENS) both in energy and time domain to study the effect of hydration and temperature. The temperature dependent mean square displacement (MSD) of hydrogen atoms within $\beta$-Casein molecules demonstrates two onsets (~100 K and ~225 K) inflected in its slope in the hydrated sample while only one onset (~100 K) is observed in the dry sample. These results indicate that the activation of natively disordered $\beta$-Casein relies on both methyl group rotations, which is independent of hydration, and the backbone fluctuations and activation of hydrophilic side chains, which are highly hydration dependent. This is consistent with the previous literature that $\beta$-Casein has higher flexibility in its dynamics compared to other well folded proteins [113].

From the analysis of EISF in energy domain, it is shown that for hydrated $\beta$-Casein, as temperature increases, both the fraction of mobile fraction of hydrogen atoms and the radius of free diffusive sphere increase significantly. However, such temperature dependence of
parameters is not evident for the dry sample. This indicates that such motion in hydrogen atoms are highly hydration dependent. At even higher temperatures of 320 and 340 K, a partial unfolding of hydrated β-CN causes a decrease in volume of diffusion.

In time domain, we found that the relaxation dynamics of dry β-CN is much slower than the hydrated one and is much less Q or temperature dependent. This provides another evidence that solvent plays an important role in the activation of proteins [137]. The intermediate scattering functions (ISF) of the hydrated β-CN demonstrate a logarithmic-like decay which is predicted by mode coupling theory (MCT) and has been observed in well-folded proteins [76, 141], and other biopolymers such as tRNA [75]. The high temperature transition in the characteristic β-relaxation time \( \tau_\beta(T) \) reveals that lack of secondary structure in β-CN will result in easier unfolding at lower temperatures compare to other more rigid biomolecules.
CHAPTER 6  DOMAIN MOTION OF AN OLIGOMERIC PROTEIN

IPPASE STUDIED BY NEUTRON SPIN ECHO

6.1 Introduction

Proteins are formed from the long chains of amino acids bonded together by peptide bonds [172]. X-ray studies showed that the polypeptide chain folds into several globular units in all proteins, and they are called 'domains' [173]. Domains are characterized that they form compact three dimensional structures, which combine in different arrangements to form proteins with different functions. The domain and helix motion of proteins exist at the length scale of 1 to 10 Å and time scale $10^{-9}$ to $10^{-6}$ s whereas atomic fluctuation, side chain and loop motion exist over the wide range of time scale and length scale of $10^{-15}$ to $10^{-12}$ s (femto second to pico second) and 0.01 to 5 Å respectively [23]. Domain motions contribute to many important protein functions, such as catalysis, transport of metabolites, and formation of protein assemblies. In addition, domain motions play an important role in the biological functioning of proteins, for instance, coupled domain motion occurs in motor proteins, signaling and structural proteins [174]. Domain movements are coupled with the hinge and shear motions [175], and the rigid domain movements are connected by the flexible joints or soft hinges that constrain the domain movements [176]. Domain motions are also responsible for the configurational change in the proteins [172].

Inorganic Pyrophosphatase (IPPase) enzyme has been of great interest because it helps in catalyzing the hydrolysis of inorganic phosphate to form orthophosphate. It plays an important role in the synthesis of nucleic acids and the production of enzymes and proteins [177]. Virtually, it is present in any cell- animal, vegetable or microbial and mainly located in the cytosol (cytosol is within the cell membrane and part of cytoplasm) [178]. This enzyme has been
studied by preparing it from a variety of sources, but the enzyme which was prepared from Saccharomyces cerevisiae and Escherichia coli, are the most highly characterized both structurally and biochemically [179]. In this study, we used IPPase prepared from Thermococcus thioreducens, which has ability to resist the denaturation above 348 K [177]. Therefore, the thermal property of IPPase makes it interesting to study its motion at high temperatures. In our previous work, the dynamics of IPPase were studied at 10 ps to 0.5 ns scale with the quasi-elastic neutron scattering technique. Its dynamics were compared with the small monomer model protein named lysozyme. Our results reveal that the ISF(Q,t) of IPPase follows a similar logarithmic decay as lysozyme; however, it has slower activity in midtime β-relaxation region. The slower relaxation dynamics of IPPase can be linked to the high temperature catalytic activity of IPPase than lysozyme [141]. IPPase exhibits three domains (pdb 3Q9M) [22] which exist and function independently to the rest of the protein chain.

The protein domain motions are controlled by various factors such as solvent contributions, internal protein friction, ligand binding, salt or buffer pH etc. Generally, the structure of folded protein is determined by X-ray crystallography that needs crystal size of several hundred micrometers. Studying the subdomains need to determine some multidomain structures; however, the spatial arrangement of proteins in crystals decreases the configurational freedom of protein. Proteins also might select different configuration than in solution. The lack of preferred configuration hinders crystallization such as in Intrinsically disordered proteins (IDPs). On the other hand, proteins have more freedom in solution to modify their configuration, and domains will also have unrestricted motion [172]. Neutron scattering has shown the potential to enlighten the structure and dynamics of proteins as well as biopolymers [21]. Neutron Spin Echo (NSE) is a powerful tool to study the motion of proteins, membranes etc at longer time scales [180].
Therefore, here we focus on studying the domain motion of IPPase at ns timescale with Neutron Spin Echo Spectroscopy (NSE).

NSE technique has been initially demonstrated by Alpert et al [181] to study the internal protein motion. NSE spectroscopy determines the time-space correlation function on timescale from nanoseconds to microsecond and at length scales from several Angstrom to hundreds of Angstrom [182]. It is used to observe the long-range relaxation motion in macromolecules. Moreover, it also has the potential to determine the global shape fluctuations and protein domain motion. In this study, we focus on exploring the slow inter-domain motions of IPPase at high temperatures by NSE technique. We also used SANS to monitor and analyze the structural changes of the protein before and after NSE measurements.

6.2 Materials and Methods

50 mg/ml IPPase protein in D$_2$O buffer (also includes 25 mM Tris-Hcl pH 9.0 and 50 mM NaCl) were purchased from iXpressGenes. The expression and purification processes of IPPase can be found in reference [141]. The IPPase structure are shown in Fig. 6.1(a) and (b), by taking snap shots in the software VMD (Visual Molecular Dynamics). Fig. 6.1(a) represents IPPase structure with three domains and Fig. 6.1(b) represents this IPPase molecule within D$_2$O solution, mimicking the experimental conditions.

Small angle neutron scattering (SANS) and Neutron Spin Echo (NSE) experiments were performed at Oak Ridge National Laboratory (ORNL). During the SANS experiment, we measure protein solution samples with three different concentrations 50mg/ml, 10mg/ml and 5mg/ml, at two temperatures 300 and 320 K. One detector was used at a distance of 6 m (distance from sample to detector) to cover the required q range from 0.00851878 to 0.152857 Å$^{-1}$. During the NSE experiment, 8 Å neutrons wavelength was used, and the data was collected at
the q range from 0.051894 to 0.163215 Å⁻¹. NSE experiments were performed at two temperatures 320 and 346 K.

Fig. 6.1 (a) IPPASE structure with three domains in different colors (green, blue and yellow); (b) protein in D₂O buffer matched with the number of protein and water during NSE experiment (one IPPASE protein is surrounded by 20833 water molecules) molecules in our system.

In the NSE experiments, correlation time \( t = \gamma J m_n^2 \lambda^3 / h^2 \) is dependent on the neutron mass, the third power of wavelength, path integral of magnetic field, gyro magnetic ratio and Planck constant. The neutron wavelength was chosen in a way to get the expected relaxation time in ns. NSE measurements were done at two high temperatures (320 and 346 K) and 50 ns time range were used to see the domain motion and some other internal modes in IPPase protein.

6.3 Results and Discussion

From SANS experiments, the coherent intensity of IPPase protein versus q is plotted at different proteins concentration and two temperatures, 300 and 320 K, as shown in Fig. 6.2 (a) and (b). There is no apparent difference between the intensity curves at two different temperatures for all concentrations, implying that the structure of IPPase does not change at these temperatures. The SANS intensity can be represented by the form factor \( P(Q) \) and the
Structure factor $S(Q)$, as shown in Chapter 2, $I_{coh} = \frac{N}{V} S(Q)P(Q)$. When the intensities from low concentration solutions are extrapolated to infinite dilution, the structure factor $S(Q)$ is equal to one and the shape of the intensity curves follows that of the form factor $P(Q)$, which determines the particle shape. Here we didn’t observe any peak at lower concentrations, but observed a peak at 50 mg/ml concentration, which means that the peak is generated from the structure factor, representing the protein-protein interactions in the high concentration solutions.

Therefore, we can safely apply the 'Guinier approximation' to fit the data measured at two lower concentrations, but cannot fit the data at the highest concentration (50 mg/ml) due to the presence of the structure factor. The Guinier analysis of small angle scattering data was developed by Andre Guinier in 1930 [183]. It was first introduced to the study of proteins in 1950s [184]. The data is fitted with the Guinier approximation, $I(q) = \exp(-Q^2R_g^2/3)$ within its valid regime of $Q*R_g < 1.73$. Despite of some limitations, it is widely used in the small angle scattering technique in neutrons and X-rays. Here, the radius of gyration $R_g$ calculated from the Guinier approximation is shown in Fig. 6.2 (c)-(f). We found that $R_g$ increases slightly with temperature, while does not change much with concentration. Increase in $R_g$ value implies that volume of proteins is expanding slightly with increase in temperature, which is consistent with our intuition. Such behavior is observed at both concentrations 5mg/mL and 10mg/mL.
Fig. 6.2 SANS experimental results on IPPase solutions. (a) and (b) Intensity versus $q$ at two temperatures 300 and 320 K and three different concentrations 50, 10 and 5 mg/ml; (c)–(f) fitted data with Guinier law and calculated $R_g$ at two different temperatures and two concentrations.

The NSE measurements provide the normalized intermediate scattering function $I(Q,t)/I(Q,0)$. The background noise includes the elastic signal from the instrument and the faster dynamics of the solvent are subtracted from the data. In Fig. 6.3 (a) and (b), ISF is plotted on a logarithmic scale as functions of time and fitted at eighteen different $q$-values at two temperatures $T = 320$ and 346 K. ISF was fitted with single exponential function $\exp(-t/\tau)$ where $1/\tau = D^*Q^2$ and $D$ is the diffusion coefficient. Fig. 6.3 (c) and (d) represent the temperature dependence of the diffusion coefficient 'D' of hydrogen atoms within IPPase molecules. From the two panels, we can clearly see that the values of $D$ are much larger at higher temperature 346 K than that at 320 K, while it follows the same trend at each temperature. Increase in the value of $D$ implies faster
motion in protein with increase in temperature, which is consistent with our previous observations in QENS experiments [141]. At low q-values, the diffusive coefficient D is almost independent of q until it reaches q = 0.1 Å⁻¹. This fact implies that at low q-values below 0.1 Å⁻¹, only the rotational motions of proteins, such as methyl group rotations or side chain motions, which are q-independent, can be detected. Above q = 0.1 Å⁻¹, there is a sharp increase in the diffusive coefficient D until q = 0.122 Å⁻¹. This sharp increase in D might be related to the translational motion of hydrogen atoms within IPPase. But above q = 0.122 Å⁻¹, we observe a decrease in the diffusion coefficient D, which might be the contribution from both translational and rotational motions [185]. Such q-values correspond to the length scale of the domains within IPPase (30~50 Å), implying this decrease in D is related to domain motion of IPPase at these q-values.

Fig. 6.3 NSE data for IPPase. (a), (b) I(Q,t)/I(Q,0) on logarithmic scale versus time at 320 and 346 K at eighteen different q-values ranging from 0.05-1.6 Å⁻¹. Straight lines represent fit with single exponential function (c) and (d) Diffusion Coefficient (D) versus Q at two temperatures 320 and 346 K.
6.4 Conclusion

SANS results revealed that the volume of IPPase expands slightly with increase in temperature up to 320K, but the conformational structure of IPPase does not change with temperature. It confirms the stability of IPPase at very high temperatures. The ISF obtained in NSE experiments was fitted with a single exponential function. Fitting of intensity with one exponential function reveals one type of motion in our examined time scale and Q range. The Q dependence and temperature dependence of D reveals the domain motions of IPPase within our measured temperature and length scale. Decrease in diffusive coefficient at q-values between 0.122 Å⁻¹ and 0.163 Å⁻¹ is related to the slow motions of domains within IPPase molecules in our measured time window and q-range.
CHAPTER 7 SUMMARY

In this thesis, we investigated dynamics of β-Casein, Inorganic Pyrophosphatase (IPPase) protein and tRNA on Nanodiamond surface (ND). These interesting biological systems were studied by using neutron scattering and MD simulations. The importance of neutron scattering in biology was explained in chapter 1. The principle of various neutron scattering techniques such as Backscattering spectrometry, Spin Echo spectroscopy and Small Angle Neutron scattering schematic and principles were explained in chapter 2, as well as the details of MD simulations was described in chapter 3. In chapter 4, we have demonstrated the effect of Nanodiamond (ND) on tRNA dynamics by quasielastic neutron scattering and MD simulations. Experimental and simulation results revealed that D$_2$O-hydrated tRNA shows a faster dynamics in the presence of ND, compared to freestanding tRNA and dehydrated tRNA. To answer this question "why does tRNA exhibit faster motion in the presence of ND?" We used simulation results to check whether that hydrogen of RNA and deuterium of water (D$_2$O) on ND surface as well as freestanding RNA follows the Stokes Einstein relation (SER) or not. We found that the SER breakdown is higher in freestanding RNA due to the presence of heterogeneities. SER breakdown follows a simple scaling law between D and $\tau$, $D \sim \tau^\alpha$. We found the large difference in the values of scaling parameter 'α' of RNA with and without ND. The farther away value of 'α' from 1 in freestanding RNA represents the jammed behavior of RNA without ND. This jammed behavior restricts the motion of freestanding RNA and hence slows down its motion.

We also described faster dynamics of RNA in the presence of ND on the structural basis, water molecules form an interfacial layer between RNA and ND with a high probability of water molecules on ND surface. This gives rise to a de-confinement of RNA molecule from ND surface with weaker heterogeneous dynamics. A de-confined RNA molecule leads to a faster
relaxation dynamics consistent with SER violation. The water molecule, on the other hand, shows ‘weak’ SER violation and follows a universal scaling law.

In Chapter 5, we have demonstrated the importance of water for the activity of milk protein β-Casein with quasielastic neutron scattering in energy and time domain. Mean Square Displacement (MSD) of hydrogen atoms of hydrated β-Casein show sharp increase in its slope around 225 K. This transition is called dynamic transition and related to the anharmonic behavior of protein while this transition is not observed in dry sample. In energy domain, Elastic Incoherent Structure Factor (EISF) results show that the increase in the fraction of hydrogen atoms and radius of diffusive sphere on increasing the temperature. However, such temperature dependent behavior is not observed in dry sample. It implies the motion of hydrogen atoms are highly hydration and temperature dependent. In time domain, we found the similar results as energy domain that dynamics of dry β-CN is much slower than the hydrated one. The Intermediate Scattering Function (ISF) of hydrated and dry β-CN demonstrates a logarithmic decay, which is predicted by Mode Coupling Theory (MCT).

In chapter 6, we explored the slow domain motions of Inorganic Pyrophosphatase Protein (IPPase) at very high temperatures with Neutron Spin Echo (NSE). We also used Small Angle Neutron Scattering (SANS) to monitor and analyze the structural change of IPPase at different temperatures. We observed that the conformational structure of IPPase does not change with temperature, due to its high temperature stability, although its radius of gyration increased slightly with temperature. From the NSE experiments, we were able to use a single exponential function to fit the ISF data and obtained diffusion coefficient D. The decrease of D with Q at higher q values reveals the slow domain motions in the oligomeric protein IPPase.
In future, we will investigate the behavior of other proteins such as Green Fluorescent protein (GFP), lysozyme etc on ND surface. This way, we might be able to universalize the faster dynamics of proteins on ND surface. In addition, we will study the dynamics of proteins with other carbon based materials such as 3D Graphene foam.
REFERENCES


ABSTRACT

DYNAMICS OF BIOPOLYMERS ON NANOMATERIALS STUDIED BY QUASIELASTIC NEUTRON SCATTERING AND MD SIMULATIONS

by

GURPREET K. DHINDSA

December 2015

Advisor: Dr. Xiang-Qiang Chu

Major: Physics

Degree: Doctor of Philosophy

Neutron scattering has been proved to be a powerful tool to study the dynamics of biological systems under various conditions. This thesis intends to utilize neutron scattering techniques, combining with molecular dynamics (MD) simulations, to develop fundamental understanding of several biologically interesting systems. Our systems include a drug delivery system containing Nanodiamonds with nucleic acid (RNA), and two specific model proteins, β-Casein and Inorganic Pyrophosphatase (IPPase).

RNA and nanodiamond (ND) both are suitable for drug-delivery applications in nanobiotechnology. The architecturally flexible RNA with catalytic functionality forms nanocomposites that can treat life-threatening diseases. The non-toxic ND has excellent mechanical and optical properties, as well as functionalized high surface area, which is actively considered for biomedical applications. In this thesis, we utilized two tools, quasielastic neutron scattering (QENS) and MD Simulations to probe the effect of ND on RNA dynamics. Our work provides fundamental understanding of how hydrated RNA motions are affected in the RNA-ND nanocomposites. From the experimental and MD simulation results, we found that the motion of hydrated RNA is faster on ND surface than that of a freestanding RNA. MD
Simulation results showed that the failure of Stokes-Einstein relation resulting in the presence of dynamic heterogeneities in the biomacromolecules. Radial pair distribution function from MD Simulations confirmed that the hydrophilic nature of ND attracts more water than RNA, resulting in the de-confinement of RNA on ND. Therefore, RNA exhibits faster motions in the presence of ND than freestanding RNA.

In the second project, we studied the dynamics of a natively disordered protein β-Casein which lacks secondary structures. In this study, the temperature and hydration effects on the dynamics of β-Casein are explored by Quasielastic Neutron Scattering (QENS). We investigated the mean square displacement (MSD) of hydrated and dry β-Casein as a function of temperature, to study the effect of hydration on their flexibility. The Elastic Incoherent Structure Factor (EISF) in the energy domain reveals the fraction of hydrogen atoms participating in motion in a sphere of diffusion. In the time domain analysis, a logarithmic-like decay is observed in the range of picosecond to nanosecond (β-relaxation time range) in the dynamics of hydrated β-Casein. Our temperature dependent QENS experiments provide evidence that lack of secondary structure in β-Casein results in higher flexibility in its dynamics and easier reversible thermal unfolding compared to other rigid biomolecules.

Lastly, we studied the domain motion of IPPase protein by Neutron Spin Echo Spectroscopy (NSE). We found that decrease in diffusion coefficient belongs to domain motion of IPPase. Moreover, the conformational structure of IPPase does not change with temperature, due to its high temperature stability, while the radius of gyration of the protein molecule is varied by temperature.
AUTOBIOGRAPHICAL STATEMENT

Education

1. PhD (Physics) Wayne State University, Detroit, Michigan, US - 2015 (expected)
2. M.S. (Physics), Wayne State University, Detroit, Michigan, US - 2013
4. B.Ed, Panjab University, Chandigarh, India, 2005
5. B.Sc (Physics, Chemistry, Mathematics), Punjabi University Patiala, Punjab, India - 2004

Publications

2. “Nanoscale de-confinement of hydrated RNA on a Nanodiamond Surface”, Gurpreet K. Dhindsa¹, Debsindhu Bhowmik¹, Monojoy Goswami²,³, Panchapakesan Ganesh², Vadym N. Mochalin⁴, Hugh O’Neill⁵, Yury Gogotsi⁴, Eugene Mamontov⁶, Bobby G. Sumpter²,³ Liang Hong⁷ and Xiang-qiang Chu¹ - to be submitted, 2015

Conferences and meetings attended

1. Oral presentation at Midwest Graduate Research Symposium (MGRS), University of Toledo, Ohio, March 21, 2015
3. Poster presentation at the Michigan Microscopy and Microanalysis Society, Detroit, MI, Nov 6, 2014
4. Poster presentation at the American Conference on Neutron Scattering (ACNS), Knoxville, Tennessee, June 2, 2014
5. Poster presentation at 8th Annual Midwest Conference on Protein Folding, Assembly and Molecular Motions, University of Notre Dame, Notre Dame, Indiana, USA, May 3, 2014
6. Poster presentation at American Physical Society (APS) March Meeting, Denver, Colorado, March 5, 2014
7. Poster presentation at the Physics Department Graduate Research Day, Department of Physics and Astronomy, Wayne State University, April 18, 2013

Awards

1. Travel Award, American Council of Neutron Scattering, June, 2014