Studies of Structure and Dynamics of Biological Macro-molecular Assemblies by Low Angle Neutron Diffraction and Inelastic X-ray Scattering

by

Yun Liu

Submitted to the Department of Nuclear Science and Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Nuclear Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
[September 2005]
August 2005

Author .....
Department of Nuclear Science and Engineering
August 3, 2005

Certified by .....
Sow-Hsin Chen
Professor
Thesis Supervisor

Read by .....
Piero Baglioni
Professor
Thesis Reader

Accepted by .....
Jeffrey A. Coderre
Chairman, Department Committee on Graduate Students

Studies of Structure and Dynamics of Biological Macro-molecular Assemblies by Low Angle Neutron Diffraction and Inelastic X-ray Scattering

by

Yun Liu

Submitted to the Department of Nuclear Science and Engineering
on August 3, 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Nuclear Science and Engineering

Abstract

This thesis is organized into two parts which focus on the studies of the dynamic structure factor and static inter-particle structure factor respectively. In the first part, we have measured and analyzed the dynamic structure factors of aligned 40 $wt\%$ calf-thymus Na-DNA molecules with the inelastic X-ray scattering (IXS). In the second part, we have developed a new efficient method to calculate the inter-particle structure factor in a simple fluid interacting with a two-Yukawa term potential and apply it to study the kinetic phase diagram and analyze the small angle neutron scattering (SANS) intensity distribution of colloidal systems.

By analyzing the dynamic structure factor measured with IXS, the phonon dispersion relations of 40 $wt\%$ calf-thymus Na-DNA with different counterion atmosphere are constructed. It is found that the addition of extra counterions will increase phonon damping at small scattering wave vector, Q. At the intermediate Q range ($12.5\ nm^{-1} < Q < 22.5\ nm^{-1}$), it may even overdamp the phonon so that the phonon feature can not be extracted from the IXS spectra. The measured sound speed is $\sim 3100 m/s$, which is much higher than the sound speed, $\sim 1800 m/s$, obtained by Brillouin light scattering. This difference shows that the atoms of DNA molecules are closely coupled to the surrounding water molecules. Therefore, the different dynamic response of water molecules in different Q range affects the overall dynamic response of the hydrated DNA molecules. By analyzing the IXS spectra, the intermediate scattering function is extracted and shows a clear two step relaxation with the fast relaxation time ranging from 0.1 to 4 ps and the slow relaxation time ranging from 2 to 800 ps.

In order to understand the phase behavior and the interactive potential of a colloidal system, we have developed a new and efficient method to calculate the inter-particle structure factor of a simple fluid interacting with a two-Yukawa term potential. We have applied this method to study the kinetic phase diagram of a system interacting with a short-range attraction and a long-range repulsion. A new

glass phase, *cluster glass*, is determined through the theoretical analysis by the mode coupling theory (MCT). The SANS intensity distribution of cytochrome C protein molecules in solutions is measured and analyzed with our method. A sharp rising intensity at very low Q value has been consistantly observed, which is named zero-Q peak. The existence of the zero-Q peak implies that a weak long-range attraction between protein molecules in solutions exists and has a even longer range than the electrostatic repulsion.

Thesis Supervisor: Sow-Hsin Chen
Title: Professor

Acknowledgments

During my six years at MIT, many people helped me to make this thesis possible. Among them, my special thanks go to my thesis supervisor Professor Sow-Hsin Chen. His tireless teaching has led me into the beautiful field of neutron and X-ray scattering. His deep physical insight has helped me to understand my research problems in the fog of mystery. His research style and working attitude has made me understand how to become a researcher and has become a model that I would like to follow. Without his encouragements and the numerous disscussions with him, I could not finish this thesis on time. My thanks also go to Professor Piero Baglioni, to whom I am truely indebted for providing unconditionally all the experimental materials. His patient teaching in the field of biological systems have also greatly benefited me. I would also like to thank Professor David G. Cory, with whom I have learned the knowledge of the Nuclear Magnetic Resonance, which greatly broadens my eyes in my research. I would also thank Professor Sidney Yip for his suggestions and guidance on both my research and my career life. I would also thank Professor Young S. Lee to join my thesis committee despite his very busy schedule.

I wish to thank Emiliano Fratini, Debora Berti for their support in the sample preparation and help in the experiments. I also wish to thank Ahmet Alatas, Harald Sinn, Ayman Said at APS, Charles J. Glinka, Paul Butler, Steve Kleine at NIST, Pappanan Thiyagarajan, Denis G. Wonzniak at IPNS, for allowing me to use their beam lines and their support during my experiments.

During the years at MIT, I fortunately met a lot of friends, who have made my life more enjoyable and fun. I would like to thank Wei-Ren Chen, Li Liu, Gabriela Leu, Antonio Faraone, Evan M. Fortunato, Zhe Sun, Yong Xiao, Jinhong Chen, Matteo Broccio, Dazhi Liu, Zhiying Chen.

Lastly, I would like to thank my family for their love and encouragements through all these years. Truly it would be impossible to complete this thesis without the love and full support from my wife, Chunguang Xie, and my daughter, Jane Liu. I also wish to thank my grandfather, Gean Liu, and my grandmother, Siyi He.

This thesis is dedicated to my father, Guoping Liu, and my mother, Shanqun Diao.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The understanding of the structure and dynamical behavior of a biological system is the key step to understand many biological functions at molecular level. A well-known example is the discovery of the double helix structure of DNA molecules with the x-ray diffraction in 1950s. This discovery unveil the genetic era since then. Nowadays about 99% of gene-containing part of human sequence has been identified with 99.99% accuracy[1]. The structure and dynamics of a biological system can be roughly classified into two categories: 1) the dynamics and structure within an individual biological molecule; 2) the dynamics and structure of many molecules, i.e., the collective behavior of many particles. The understanding of the former case is important to understand the function of each particle and how it responds to the environmental changes. The understanding of the second one is very important for the phase behavior of the system and for the functions of biological systems.

In order to probe those information, different techniques, such as neutron scattering, x-ray scattering, light scattering, confocal laser, microscope, are applied, depending on the length scale and time scale one would like to investigate. Since the size of biological macromolecules, such as DNA, proteins, ranges from 0.1 nm to 1000 nm, neutron and X-ray scattering become two of the most powerful techniques to probe their static and dynamic information.

This thesis contains two parts. In the first part, we have investigated the dynamic information of 40 $wt\%$ calf-thymus Na-DNA with inelastic X-ray scattering. The

second part is focused at the extraction of the static structure information. To this aim, a new and efficient method is developed to calculate the structure factor in order to understand the collective behavior of a simple liquid interacting with a two-Yukawa term potential. This method has been applied to predict the kinetic phase diagram of a simple fluid system and to investigate the effective inter-protein potential in solutions. In order to study the effective inter-protein potential, we have used small angle neutron scattering technique to measure the scattering intensity distribution.

In the following sections, I will introduce some terminology which is used in the literatures and in this thesis. One can easily find out more complete and detailed description in many books of liquid theory[2, 3]. Considering the large use of inelastic X-ray scatteirng and small angle neutron scattering techniques in our experiments, we report below a brief introduction to these techniques.

1.1 Correlation Functions and Structure Factor

Because the response of a sample due to neutron and X-ray scattering is still in the linear response region, what neutron and x-ray scattering measure is proportional to the fourier transform of density-density correlation functions [3].

The local particle density at a point $\vec{r}$ can be expressed as

$$\rho(\vec{r}) = \sum_{i=1}^{N} \delta(\vec{r} - \vec{r}_i) \tag{1.1}$$

where N is the total number of particles, $\vec{r}_i$ the position of the ith particle.

The correlations between $\rho(\vec{r})$ at two points separated with distance $\vec{r}$ at $t = 0$ is the static density-density correlation function $G(\vec{r})$, which is defined as

$$G(\vec{r}) = \frac{1}{N} \int < \rho(\vec{r}' + \vec{r})\rho(\vec{r}') > \tag{1.2}$$

The *static structure factor*, $S(\vec{Q})$, is defined as the Fourier transform of $G(\vec{r})$ as

$$S(\vec{Q}) = \int e^{-i\vec{Q}\cdot\vec{r}} G(\vec{r}) d\vec{r} = \frac{1}{N} < \rho_{\vec{Q}}\rho_{-\vec{Q}} > \tag{1.3}$$

where

$$\rho_{\vec{Q}} = \int e^{-i\vec{Q}\cdot\vec{r}_i} d\vec{r} \tag{1.4}$$

And the *pair distribution function* $g(\vec{r})$ is defined as

$$g(\vec{r}) = \frac{1}{N} < \sum_{i}^{N} \sum_{i \neq j}^{N} \delta(\vec{r} + \vec{r}_j - \vec{r}_i) > \tag{1.5}$$

In an isotropic sample, we can use the scaler value, r, and Q instead of $\vec{r}$, and $\vec{Q}$ in the functions, $G(\vec{r})$, $g(\vec{r})$, $\rho_{\vec{Q}}$, and $S(\vec{Q})$.

$S(Q)$ is related with $g(r)$ by

$$S(Q) = 1 + \rho \int e^{-i\vec{Q}\cdot\vec{r}} g(r) d\vec{r} \tag{1.6}$$

where, $\rho = \frac{N}{V}$ is the particle number density, V the volume of the sample.

The correlations between $\rho(\vec{r})$ at two points separated with distance $\vec{r}$ with the time difference t is the time dependent density-density correlation function $G(\vec{r}, t)$, which is defined as

$$G(r, t) = \frac{1}{N} \int < \rho(\vec{r}'(t) + \vec{r}) \rho(\vec{r}') > \tag{1.7}$$

$G(r, t)$ is also called *van Hove function.* Its spacial Fourier transform, $F(Q, t)$, is called *intermediate scattering function*, which is expressed as

$$F(Q, t) = \int e^{-i\vec{Q}\cdot\vec{r}} G(r, t) d\vec{r} \tag{1.8}$$

$$= \frac{1}{N} < \sum_{j}^{N} \sum_{l}^{N} e^{i\vec{Q}\cdot\vec{r}_j(t)} e^{-i\vec{Q}\cdot\vec{r}_l} > \tag{1.9}$$

At $t = 0$, $F(Q, t = 0) = S(Q)$ is just the static structure factor.

The *dynamical structure factor* is just the time Fourier transform of the interme-

diate scattering function,

$$S(Q,\omega) = \int_{-\infty}^{\infty} e^{-i\omega t} F(Q,t) dt \tag{1.10}$$

$S(Q,\omega)$ contains the dynamical information and can be measured by inelastic scattering techniques. Since some inelastic scattering methods measure the transferred energy instead of frequency shift, the dynamical structure factor is sometimes expressed in terms of transferred energy, E, as $S(Q,E)$, which is equal to $S(Q,\omega)/\hbar$.

1.2 Introduction to the High Resolution Inelastic X-ray Scattering

The high resolution inelastic X-ray scattering is a relatively new scattering technique. With the combination of the availability of the third generation synchrotron source and the new analyzers of very good quality, the energy resolution of the current IXS instruments can reach about $1.5 \sim 2\ meV$.

The double-differential cross-section of an isotropic one component system that the inelastic X-ray scattering (IXS) measures can be expressed as

$$\frac{d^2\sigma}{dQdE} = N r_0^2 (\varepsilon_i \cdot \varepsilon_f)^2 \frac{k_f}{k_i} f(Q)^2 S(Q,E)$$

where E is the energy transferred in the scattering process, N the total number of atoms, ε_i and ε_f the polarization vectors of X-ray photons before and after the scattering process, k_i and k_f the wave vectors of the incident and scattered X-rays respectively, r_0 the classical radius of an electron, and $f(Q)$ the form factor of the atoms in the system. $k_f \simeq k_i$, $Q = 2k_i sin(\theta/2)$, where θ is the scattering angle. Hence the only energy dependent part in the cross section is the dynamic structure factor, $S(Q,E)$. If one normalizes the measured IXS spectrum at each fixed Q to unity, all the prefactors before $S(Q,E)$ vanish. In the following, when we mention an IXS spectrum, we always mean a normalized IXS spectrum, $S(Q,E)/S(Q)$, where the

structure factor, $S(Q)$, is equal to $\int S(Q,E)dE$. The key to analyze an IXS spectrum is $S(Q,E)$ calculation.

Since IXS is a relatively new technique, it is useful to briefly describe some technical details of the instruments. Our experiments was carried out at the high resolution inelastic X-ray scattering beam line (3-ID) at the Advanced Photon Source (APS). Fig.1 reports the schematic diagram of the inelastic x-ray scattering spectrometer at the Advanced Photon Source, Argonne National laboratory. The 21.657 keV (denoted by E_1) X-ray was produced by a 4.6 m long undulator with a 2.7 cm magnetic period at the storage ring of the APS at Sector 3 [4]. The beam was premonochromated by a water-cooled diamond (1 1 1) double-crystal monochromator. It is further monochromatized by a high resolution in-line monochromator which consists of two nested silicon channel-cut crystals. An asymmetrically cut silicon (4 4 0) crystal is used as the outer channel of the in-line monochromator. The crystal is able to collimate the incoming beam to below one microradian inside the channel cut. After the outer crystal, the beam is reflected twice by the inner channel cut with a Bragg angle at 83.2° at the (15 11 3) reflection. After transmitting through the second monochromator, the beam has a tunability range between 21.50 and 21.70 keV with an energy width of 1.3 meV. By rotating the inner and outer channel-cut of the in-line monochromator, the energy can be tuned at the meV scale. The beam is then focused by a total reflecting mirror to a spot size at the sample of 200 μm $\times$ 100 μm after passing through the high energy resolution monochromator. The photon flux at the sample position is about 6×10^8 photons/s at a current of 100 mA in the storage ring.

The scattered photon was collected by a (18 6 0) reflection of a bent silicon analyzer in extreme back-scattering geometry ($\theta_B = 89.98°$). The analyzer is comprised of a 4 mm thick, 100 mm large focusing silicon disk, with a distance of 6m from the sample. The disk is glued onto a flexible glass wafer and diced with a high-precision diamond saw into small pixels of 1mm size in order to avoid bending stress in the silicon. By etching the crystal in a KOH solution, the stress from distorted areas

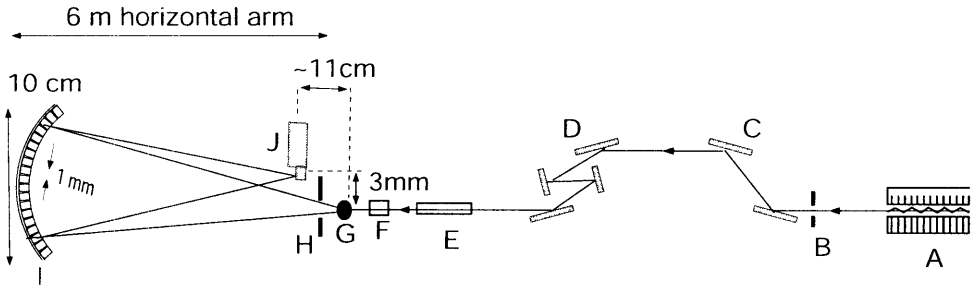

Figure 1-1: Schematic diagram of the inelastic x-ray scattering spectrometer at Beamline 3-ID, Advanced Photon Source, Argonne National laboratory. The beam comes from an undulator (A) and premonochromator (C), then passes through the high-resolution monochromator (D) and focusing mirror (E) before it illuminates the sample (G). The scattering intensity is focused by analyzer (I) into detector (J). (F) is the ionization chamber to monitor the incident flux onto sample. (B) and (H) are the slit systems that determine the source size.

at the cut regions can be removed. A temperature-controlled chamber is used to house the analyser to keep the temperature stabilized near room temperature to $\pm$ 2 mK/24h.

A commercially available Cd-Zn-Te semiconductor detector is used here. Its electronic noise level is below 0.001Hz.

The wavevector transfer, Q, can be varied according to the relation

$$Q = \frac{4\pi}{\lambda} \sin\frac{\theta}{2} = 21.95\mathring{A}^{-1} \sin\frac{\theta}{2} \tag{1.11}$$

where λ is the wavelength of the X-rays and θ the scattering angle at the sample. The maximum scattering angle of the setup is $\theta \leq 15.5°$, which corresponds, according to the above equation, to a maximum momentum transfer of 30 nm^{-1}.

The net energy resolution function was measured by a Plexiglas sample. The shape of the measured resolution function can be reasonably well described by a fit with a Pseudo-Voigt function

$$\begin{aligned} R(E) &= I_o\{\frac{2\eta}{\pi\Gamma}\left[1+4(\frac{E}{\Gamma})^2\right]^{-1} \\ &\quad +(1-\eta)\frac{2}{\Gamma}(\frac{\ln 2}{\pi})^{1/2} \times \exp\left[-4\ln 2(\frac{E}{\Gamma})^2\right]\} \end{aligned} \tag{1.12}$$

where Γ is the FWHM of the curve, η the mixing parameter and I_o a normalization constant.

Notice that, Eq. (1.2) dose not take into account of resolutions function. A measured IXS spectrum is the convolution of the double-differential cross-section with $R(E)$.

1.3 Introduction to the Small Angle Neutron Scattering

Small angle neutron scattering is a well-developed scattering technique. It has been widely applied to study the structures of biological system, micellar solutions, etc.

The small angle neutron scattering measures the spacial differential cross-section per unit volume as

$$I(Q) = \frac{1}{V} < |\sum_{l=1}^{N} b_l e^{i\vec{Q}\cdot\vec{r}_l}|^2 > \tag{1.13}$$

where b_l is the neutron scattering length of lth atom. The neutron scattering length is independent of Q. For a isotropic one component system, $b_l = b$. Therefore,

$$I(Q) = \frac{N}{V} b^2 S(Q) \tag{1.14}$$

In a two phase system, such as protein solutions, micellar solutions, the scattering objects, e.g., proteins, micelles, can be distinguished in appropriate conditions from the continuous solvent. Let's assume that there are total N atoms in the system, N_p macromolecules in solutions, N_m atoms in each macromolecule, and N_s atoms for solvent molecules. Therefore,

$$\begin{aligned} I(Q) &= \frac{1}{V} < |\sum_{j=1}^{N_p}\sum_{l=1}^{N_m} b m_{j,l} e^{i\vec{Q}\cdot\vec{r}_{j,l}} + \sum_{j=1}^{N_s} b_s e^{i\vec{Q}\cdot\vec{r}_j}|^2 > \\ &= \frac{1}{V} < |\sum_{j=1}^{N_p}\sum_{l=1}^{N_m} (b m_{j,l} - b_s) e^{i\vec{Q}\cdot\vec{r}_{j,l}} + \sum_{j=1}^{N} b_s e^{i\vec{Q}\cdot\vec{r}_j}|^2 > \end{aligned} \tag{1.15}$$

$$= \frac{1}{V} < |\sum_{j=1}^{N_p}[\sum_{l=1}^{N_m}(bm_{j,l} - b_s)e^{i\vec{Q}\cdot(\vec{r}_{j,l}-\vec{R}_j)}]e^{i\vec{Q}\cdot\vec{R}_j} + \sum_{j=1}^{N} b_s e^{i\vec{Q}\cdot\vec{r}_j}|^2 >$$

where $bm_{j,l}$ is the scattering length of lth atom in the jth macromolecule, b_s the scattering length of a solvent atom/molecule. Since the solvent molecule is typically very small compared to $2\pi/Q$, it can be treated as a uniform continuous media. Thus, the last term in the above equation, $\sum_{j=1}^{N} b_s e^{i\vec{Q}\cdot\vec{r}_j}$, is equal to $\delta(Q)$, contributing only at $Q = 0$. Therefore, this last term can be considered as zero since SANS only measures the intensity at a finite Q value. If we can further assume that each macromolecule in the solvent is the same, and it is spherically isotropic, a more useful expression is written as

$$\begin{aligned} I(Q) &= \frac{1}{V} < |(\sum_{l=1}^{N_m}(bm_{1,l} - b_s)e^{i\vec{Q}\cdot(\vec{r}_{1,l}-\vec{R}_1)})(\sum_{j=1}^{N_p} e^{i\vec{Q}\cdot\vec{R}_j})|^2 > \\ &= \frac{N_p}{V}P(Q)S_{inter}(Q) \end{aligned} \tag{1.16}$$

where

$$P(Q) = |\sum_{l=1}^{N_m}(bm_{1,l} - b_s)e^{i\vec{Q}\cdot(\vec{r}_{1,l}-\vec{R}_1)}|^2 \tag{1.17}$$

$$S_{inter}(Q) = \frac{1}{N}|\sum_{j=1}^{N_p} e^{i\vec{Q}\cdot\vec{R}_j}|^2 \tag{1.18}$$

$P(Q)$ is called the intra-particle structure factor, or form factor. $S_{inter}(Q)$ is called the inter-particle structure factor. $P(Q)$ is only determined by the structure of each individual macromolecule and $S_{inter}(Q)$ describes the correlations between the centers of macromolecules. In the study of the inter-colloidal correlation in solutions, people usually drop the subscript and use $S(Q)$ as the inter-particle structure factor.

When the volume fraction of the solute is very dilute, $S(Q) \sim 1$. Therefore $I(Q) = \frac{N_p}{V}P(Q)$. The SANS intensity distribution, $I(Q)$, contains the information of the particle structure and shape.

In general, the fitting of $I(Q)$ require the knowledge of $S(Q)$. However, $S(Q)$ is difficult to calculate. In principle, for a simple liquid sample, $S(Q)$ can be calculated

through the Orstein-Zernike (OZ) equation once the inter-particle potential is known.

1.4 Survey of the Thesis

This thesis is outlined as the following.

In the first part, chapter 2 and chapter 3 will focus on the understanding of the dynamical structure factor of the aligned DNA molecules. Chapter 2 will describe the theory used to calculate the dynamical structure factor of $S(Q, E)$ by the generalized five effective eigenmode theory (GFEE) and its special case, the generalized three effective eigenmode theory (GTEE). Chapter 3 will apply the GTEE theory to analyze the IXS spectra of aligned DNA molecules and the effect on $S(Q, E)$ due to different counter-ions is compared.

In the second part, chapter 4, chapter 5, and chapter 6 will focus on the understanding of the static structure information of a colloidal system. Chapter 4 will describe a new and efficient method developed to calculate the inter-particle structure factor, $S(Q)$, of a simple liquid system interacting with a two Yukawa term potential by solving the OZ equation with the mean spherical approximation (MSA) method. This method is called the "two-Yukawa model method". In chapter 5, the "two-Yukawa model method" is applied to calculate $S(Q)$ as the input of the mode coupling theory (MCT) to predict the kinetic phase diagram of a simple liquid system interacting with a short-range attraction and a long-range repulsion. In chapter 6, I have applied this two-Yukawa model method to analyze the SANS intensity distribution of the cytochrome C protein solutions.

Chapter 2

Theory of Dynamic Structure Factor Based on the Generalized Hydrodynamic Model

The knowledge of the dynamical structure factor, $S(Q, E)$, calculated by a reasonable method is fundamental for the analysis of the IXS spectra. In this chapter, based on the generalized hydrodynamic model we will develop a theory, which can be applied to calculate $S(Q, E)$ for a system containing multi-species of atoms.

There are different models used in the literature to calculate $S(Q, E)$. The damped harmonic oscillator (DHO) model was used to fit IXS spectra of water[5, 6] as well as the inelastic neutron scattering (INS) spectra [7]. However, DHO model is only a rough approximation of a disordered system. In a simple liquid, the DHO model works only when the α-relaxation time is very long, which is progressively violated when temperature increases. Therefore, the formula with the memory kernel based on the Zwanzig-Mori expansion is proposed [8] and applied to analyze the IXS spectra of water and liquid neon[9, 10]. At the same time, the three effective eigenmode (TEE) theory[11] based on the generalized hydrodynamic model was also used to analyze the IXS spectra of water[12] to overcome the shortcoming of the DHO model. In order to analyze the IXS spectra of biological systems which contains multi-species of atoms in a system, Chen and Liao generalized the TEE theory to the generalized three effective

eigenmode theory (GTEE) [13] assuming that $S(Q, E)$ can be described as the projection of the complete dynamics onto three microscopic quantities, e.g., atomic form factor weighted number density, longitudinal velocity, and temperature. The GTEE theory was successfully applied to analyze the IXS spectra of lipid bilayers[14, 15, 16]. In the next chapter, we apply the GTEE theory to analyze the liquid crystalline DNA.

The GTEE theory could be improved to the generalized five effective eigenmode theory (GFEE) by introducing two more microscopic quantities, longitudinal momentum flux, and longitudinal heat flux, and projecting the dynamics of the system onto five microscopic quantities. In this chapter, we will present the derivation of the GFEE theory. The derivation of the GTEE theory can be obtained by following the same scheme.

2.1 Generalized Five Effective Eigenmode Theory (GFEE)

In this section, the generalized five effective eigenmode theory (GFEE), will be developed to calculate $S(Q, \omega)$. Traditionally, one usually uses k as the wavevector instead of Q. Therefore, we will use $S(k, \omega)$ in the following. In a scattering experiment of simple liquids or disordered systems, where there is no crystal structure, $Q = k$.

To calculate $S(k, \omega)$, we can equivalently first calculate the intermediate scattering function, $F(k, t)$.

First, we introduce five generalized microscopic quantities as

$$a_j(\vec{k}) = \frac{1}{\sqrt{N}} \sum_{l=1}^{N} A_j^{(l)}(\vec{k}) f_l(k) e^{-i\vec{k}\cdot\vec{r}_l} \tag{2.1}$$

where N is the total number of atoms, $f_l(k)$ the x-ray form factor of lth atom, which can be found in the international tables for x-ray crystallography[17]. $A_j^{(l)}$ is defined as

$$A_1^{(l)} = 1 \tag{2.2}$$

$$A_2^{(l)} = \vec{v}_l \cdot \vec{k}/k \tag{2.3}$$

$$A_3^{(l)} = \frac{1}{2}m\vec{v}_l^2 + \frac{1}{2}\sum_{j=1,j\neq l}^{N}\phi(\vec{r}_{lj}) \tag{2.4}$$

$$A_4^{(l)} = (\vec{v}_l \cdot \vec{k}/k)^2 + \frac{i}{2mk^2}\sum_{j=1,j\neq l}^{N}\vec{k}\cdot\frac{\partial\phi(\vec{r}_{lj})}{\partial\vec{r}_{lj}} \times (e^{i\vec{k}\cdot\vec{r}_{lj}} - 1) \tag{2.5}$$

$$A_5^{(l)} = (\frac{1}{2}m\vec{v}_l^2 + \frac{1}{2}\sum_{j=1,j\neq l}^{N}\phi(\vec{r}_{lj}))\vec{v}_l \cdot \vec{k}/k + \frac{i}{2k}\sum_{j=1,j\neq l}^{N}\vec{v}_l\cdot\frac{\partial\phi(\vec{r}_{lj})}{\partial\vec{r}_{lj}} \times (e^{i\vec{k}\cdot\vec{r}_{lj}} - 1) \tag{2.6}$$

where $\vec{r}_{lj} = \vec{r}_l - \vec{r}_j$, $\phi(\vec{r}_{jl})$ the potential between the atom j and the atom l. $a_1(\vec{k})$ is the microscopic number density, $a_2(\vec{k})$ is the microscopic longitudinal velocity, $a_3(\vec{k})$ is the microscopic energy density, $a_4(\vec{k})$ is the microscopic longitudinal momentum flux, and $a_5(\vec{k})$ is the microscopic longitudinal energy flux. With this five microscopic quantities, we can thus calculate the 5×5 correlation functions, $F_{lj}(k,t)$, as

$$F_{jl} =< a_j(\vec{k})e^{tL}a_l(\vec{k}) > \tag{2.7}$$

where L is the Liouville operator, which is

$$L = \sum_{i=1}^{N}[\vec{v}_i\cdot\frac{\partial}{\partial\vec{r}_i} - \frac{1}{m_i}\sum_{j=1,j\neq l}^{N}\frac{\partial\vec{r}_{ij}}{\vec{r}_{ij}}\cdot\frac{\partial}{\partial\vec{v}_i}] \tag{2.8}$$

These correlation functions have the following properties,

$$F_{jl}(k,t) = F_{lj}(k,t) \tag{2.9}$$

$$F_{2l}(k,t) = \frac{i}{k}\frac{\partial}{\partial t}F_{1l}(k,t) \tag{2.10}$$

$$F_{4l}(k,t) = \frac{i}{k}\frac{\partial}{\partial t}F_{2l}(k,t) \tag{2.11}$$

$$F_{5l}(k,t) = \frac{i}{k}\frac{\partial}{\partial t}F_{3l}(k,t) \tag{2.12}$$

If we project the dynamics to the five independent microscopic quantities defined above, $e^{tL} = e^{t\underline{H}(k)}$, where $\underline{H}(k)$ is a 5×5 matrix with k-dependent elements.

Therefore,

$$F_{jl} = [e^{-t\underline{H}(k)}\underline{V}(k)] \tag{2.13}$$

where

$$V_{jl}(k) = F_{jl}(k) =< [a_j(\vec{k})^* a_l(\vec{k}) > \tag{2.14}$$

Therefore,

$$\mathbf{V}(k) = \begin{pmatrix} V_{11}(k) & 0 & V_{13}(k) & V_{14}(k) & 0 \\ 0 & V_{22}(k) & 0 & 0 & V_{25}(k) \\ V_{13}(k) & 0 & V_{33}(k) & V_{34}(k) & 0 \\ V_{14}(k) & 0 & V_{34}(k) & V_{44}(k) & 0 \\ 0 & V_{25}(k) & 0 & 0 & V_{55}(k) \end{pmatrix} \tag{2.15}$$

It can be shown that

$$V_{11}(k) = S(k) = \frac{1}{N}|\sum_{j=1}^{N} f_j(k)e^{-i\vec{k}\cdot\vec{r}_j}|^2 \tag{2.16}$$

$$V_{22}(k) = \frac{1}{N}\sum_{j=1}^{N} < (\vec{k}\cdot\vec{v}_j)^2 > f_j^2(k) \tag{2.17}$$

$V_{22}(k)$ can thus be considered to be proportional to the weighted average of thermal velocity. Other elements of the matrix (2.15) can be also obtained as the results in Ref[11]. However, due to the existence of the form factor of different atoms, the exact form needs to be modified. The results of V_{11}, and V_{22} are examples of this modification.

After this step, all the derivations can be followed almost exactly the same as the derivation of the five eigenmode theory (FEE) as given in Ref[11]. For completeness, I will outline the derivation here and list the major results.

After obtained the matrix $\underline{V}(k)$, a set of orthonormal quantities, $b_j(\vec{k})$, can be

defined as

$$b_j(\vec{k}) = \sum_{l=1}^{5} U_{jl}(k) a_l(\vec{k}), \tag{2.18}$$

which satisfy

$$< [b_j(\vec{k})]^* b_l(\vec{k}) >= \delta_{jl} \tag{2.19}$$

where U_{jl} is the matrix element of $\underline{U}(k)$, which satisfy

$$\underline{U}^T(k)\underline{U}(k) = V^{-1}(k) \tag{2.20}$$

Thus we define a new set of correlation function as

$$G_{jl}(k,t) =< [b_j(\vec{k})]^* e^{tL} b_l(\vec{k}) > \tag{2.21}$$

The dynamical structure factor $S(k,\omega)$ is thus given as

$$\frac{S(k,\omega)}{S(k)} = \frac{1}{\pi} Re \int_0^\infty e^{-i\omega t} G_{11}(k,t) \tag{2.22}$$

Define $\tilde{G}_{jl}(k,z) = \int_0^\infty dt e^{-zt} G_{jl}(k,t)$. With the projector formalism, it can be proved that

$$z\tilde{G}_{ij}(k,z) = -\sum_{l=1}^{5} H_{il}(k,)\tilde{G}_{lj}(k,z) + \delta_{ij} \tag{2.23}$$

where the matrix $H(k)$ is given as

$$H(k) = \begin{pmatrix} 0 & if_{un}(k) & 0 & 0 & 0 \\ if_{un}(k) & 0 & if_{uT}(k) & if_{u\sigma}(k) & 0 \\ 0 & if_{uT}(k) & 0 & 0 & if_{Tq}(k) \\ 0 & if_{u\sigma}(k) & 0 & z_\sigma(k,z) & iz_{q\sigma}(k,z) \\ 0 & 0 & if_{Tq} & iz_{q\sigma}(k,z) & z_q(k,z) \end{pmatrix} \tag{2.24}$$

$f_{un}(k)$ is the second frequency moment of the dynamic structure factor. It should be noted that no approximation is needed to obtain Eq. (2.24).

Therefore,

$$\begin{aligned} \frac{S(k,\omega)}{S(k)} &= \mathrm{Re}\left\{\tilde{G}_{11}(k,z)\right\}_{z=i\omega} \\ &= \frac{1}{\pi}\mathrm{Re}\left\{\frac{\mathbf{I}}{z+\mathbf{H}(k)}\right\}_{(1,1),\ z=i\omega}, \end{aligned} \tag{2.25}$$

where $\mathbf{I}$ is the 5×5 identity matrix, label (1,1) outside the curly bracket means the (1,1) element of the matrix, i the imaginary unit.

$\tilde{G}_{11}(k,z)$ can be also cast into the form of continued fraction expansion as

$$\tilde{G}_{11}(k,z) = \left(z + \frac{f_{un}^2(k)}{z + z_\phi + \frac{[f_{uT}(k)+\Delta(k,z)]^2}{z+z_T(k,z)}}\right)^{-1} \tag{2.26}$$

where

$$z_\phi(k,z) = \frac{f_{u\sigma}^2(k)}{z + z_\sigma(k,z) + z_{q\sigma}^2(k)/[z+z_q(k)]} \tag{2.27}$$

$$z_T(k,z) = \frac{f_{Tq}^2(k)}{z + z_q(k) + z_{q\sigma}^2(k)/[z+z_\sigma(k)]} \tag{2.28}$$

$$\Delta(k,z) = \frac{-f_{u\sigma}(k)f_{Tq}(k)z_{q\sigma}(k)}{[z+z_q(k)][z+z_\sigma(k)] + z_{q\sigma}^2(k)} \tag{2.29}$$

When analyzing the IXS spectra, $S(k,\omega)$ can be calculated with the Eq. (2.25). And the seven non-zero matrix element can be treated as fitting parameters.

The generalized three effective eigenmode theory is a special case of the GFEE theory by only considering three microscopic quantities, $a_j(\vec{k})$ $(j = 1,2,3)$. The derivations can be obtained by following the derivations of the GFEE theory exactly. A different approach has been given in Ref. [13].

During our analyses of the IXS spectra of aligned DNA samples, the GTEE theory can fit the spectra excellently. Therefore we did not apply the GFEE theory in the analyses to reduce the number of fitting parameters since the GFEE theory needs to have seven fitting parameters instead of four fitting parameters of the GTEE theory.

In the following, the major results of the GTEE theory is given and the results at hydrodynamic limit of the GTEE theory are also discussed

2.2 Generalized Three Effective Eigenmode Theory (GTEE)

In GTEE theory, we can calculate the generalized dynamic structure factor, $S(Q,\omega)$, which may contain multi-species of atoms, with a simple equation bearing the same form as that described by the TEE theory. However, the calculated result is expressed in terms of the generalized dynamic structure factor, which is defined as

$$S(k,\omega) = \frac{1}{2\pi}\frac{1}{N}\int dt e^{i\omega t}\sum_{j,l}^{N}\langle f_j(k)f_l(k)e^{i\mathbf{k}\cdot\mathbf{r}_l(0)}e^{-i\mathbf{k}\cdot\mathbf{r}_j(t)}\rangle. \tag{2.30}$$

where $f_i(k)$ is the form factor of the atom with index i.

The normalized generalized dynamic structure factor can be written as

$$\frac{S(k,\omega)}{S(k)} = \frac{1}{\pi}\,\mathrm{Re}\left\{\frac{\mathbf{I}}{z+\mathbf{H}(k)}\right\}_{(1,1),\ z=i\omega}, \tag{2.31}$$

where $\mathbf{I}$ is the 3×3 identity matrix, label 1,1 outside the curly bracket means the (1,1) element of the matrix, i the imaginary unit.

The generalized hydrodynamic matrix $\mathbf{H}(\mathbf{k})$ is in the form of

$$\mathbf{H}(Q) = \begin{pmatrix} 0 & if_{un}(k) & 0 \\ if_{un}(k) & z_u(k) & if_{uT}(k) \\ 0 & if_{uT}(k) & z_T(k) \end{pmatrix} \tag{2.32}$$

where the four Q dependent matrix elements, z_T, f_{uT}, z_u and f_{un}, are treated as fitting parameters. Among them, $f_{un}(k)$ is related to the second frequency moment of the dynamic structure factor and is given in terms of the structure factor ,$S(k)$, as

$$f_{un}(k) \;=\; kv_0(k)\left[S(k)\right]^{-1/2},$$

$$v_0^2(k) = \sum_\alpha f_\alpha^2(k) w_\alpha^2 v_{0\alpha}^2, \tag{2.33}$$

where $w_\alpha = \sqrt{N_\alpha/N}$ is the square root of the fraction of atomic type α over the total number of atoms, and $v_{0\alpha} = \sqrt{k_B T/m_\alpha}$.

Eq. (2.31) can be written out in an explicit form as

$$\frac{S(k,\omega)}{S(k)} = \frac{1}{\pi} \mathrm{Re} \left\{ \frac{z^2 - (z_u + z_T)z + z_u z_T + f_{uT}^2}{z^3 - (z_u + z_T)z^2 + (z_u z_T + f_{uT}^2 + f_{un}^2)\, z - f_{un}^2 z_T} \right\}_{z=i\omega}. \tag{2.34}$$

Both Eq. (2.31) and Eq. (2.34) have been used to fit the normalized IXS spectra. However, Eq. (2.34) is more computationally efficient.

The denominator of Eq. (2.34) usually has three distinct roots, one real root and two complex conjugate roots, or three different real roots. In principle, it can also have three real roots, among which at least two are equal. However, since we never encounter this case in practice, we will not discuss this situation.

(i) When there are one real root, Γ_h, and two complex conjugate roots, $\Gamma_s \pm i\Omega_s$, Eq. (2.34) can be decomposed into the sum of three Lorentzian terms,

$$\begin{aligned} S(k,\omega)/S(k) &= \frac{1}{\pi}\{A_0 \frac{\Gamma_h}{\omega^2 + \Gamma_h^2} + A_s \frac{\Gamma_s + b(\omega + \Omega_S)}{(\omega + \Omega_S)^2 + \Gamma_s^2} \\ &\quad + A_s \frac{\Gamma_s - b(\omega - \Omega_S)}{(\omega - \Omega_S)^2 + \Gamma_s^2}\} \end{aligned} \tag{2.35}$$

where

$$\begin{aligned} M(x) &= x^2 - (z_u + z_T)x + z_u z_T + f_{uT}^2, \\ N &= \frac{M(\Gamma_s - i\Omega_s)}{2i\Omega_s(\Gamma_h - (\Gamma_s - i\Omega_s))}, \\ A_0 &= \frac{M(\Gamma_h)}{(\Gamma_s - \Gamma_h)^2 + \Omega_s^2}, \\ A_s &= \mathrm{Re}(N), \\ b &= -\mathrm{Im}(N)/\mathrm{Re}(N). \end{aligned} \tag{2.36}$$

This three terms correspond to three peaks in an IXS spectrum, one central Rayleigh peak with the linewidth $2\Gamma_h$, two Brillouin doublet peaks, with Ω_s the phonon excitation energy and Γ_s the phonon damping.

The corresponding ISF is of the form

$$F(k,t) = A_0 e^{-\Gamma_h t} + 2A_s e^{-\Gamma_s t}(\cos\Omega_s t + b\sin\Omega_s t) \tag{2.37}$$

The ISF shows a two-step relaxation. Since Γ_h is typically smaller than Γ_s, the relaxation time of the fast dynamics is $\tau_f = 2\pi/\Gamma_s$, and the relaxation time for the slow dynamics is $\tau_s = 2\pi/\Gamma_h$.

(ii) When there are only three real roots, Γ_1, Γ_2, and Γ_3, there are three peaks all centering at $\omega = 0$ in an IXS spectrum. There are no longer the concept of phonon anymore in this case. Eq. (2.34) can be written as the sum of three terms as

$$S(k,\omega)/S(k) = \frac{1}{\pi}\sum_{i=1}^{3} A_i \frac{\Gamma_i}{\omega^2 + \Gamma_i^2}, \tag{2.38}$$

where

$$\begin{aligned} M(x) &= x^2 - (z_u + z_T)x + z_u z_T + f_{uT}^2, \\ A_1 &= \frac{M(\Gamma_1)}{(\Gamma_1 - \Gamma_(2))(\Gamma_1 - \Gamma_3)}, \\ A_2 &= \frac{M(\Gamma_2)}{(\Gamma_2 - \Gamma_(3))(\Gamma_2 - \Gamma_1)}, \\ A_3 &= \frac{M(\Gamma_3)}{(\Gamma_3 - \Gamma_(1))(\Gamma_3 - \Gamma_2)}. \end{aligned} \tag{2.39}$$

Its correspondent ISF can be expressed as

$$F(k,t) = \sum_{i=1}^{3} A_i e^{-\Gamma_i t} \tag{2.40}$$

Strictly speaking, the ISF is not a two-step relaxation since we have three Lorentizan terms. However, one of the terms usually has a negative amplitude, A, and a much faster relaxation time. Hence if $\Gamma_1 > \Gamma_2 > \Gamma_3$, we assign $\tau_f = 2\pi/\Gamma_2$, and $\tau_s = 2\pi/\Gamma_3$.

We will show the ISF calculated with the fitted parameters in the next chapter. A two-step relaxation can be seen clearly from the figure.

Eq. (2.31) can also be cast into the form of continued fraction expansion as

$$\frac{S(k,\omega)}{S(k)} = \mathrm{Re}\left\{\left(z + \frac{f_{un}^2(k)}{z + z_u(k) + \frac{f_{uT}^2(k)}{z+z_T(k)}}\right)^{-1}\right\}_{z=i\omega} \tag{2.41}$$

By comparing to Eq. (2.26), the relations of the parameters between the GTEE theory and the GFEE theory can be found as

$$z_u(k) = z_\phi(k,z) \tag{2.42}$$

$$f_{uT}(k) = f_{uT}^{(5)}(k) + \Delta(k,z) \tag{2.43}$$

$$z_T(k) = z_T^{(5)}(k,z) \tag{2.44}$$

where the superscript (5) means this term is defined in the GFEE theory.

The GTEE theory can be shown to reduce to some familiar approximate models used in the literature, such as the DHO model, and the viscoelastic model.

(i) When $f_{uT} = 0$, the expression of $S(k,\omega)/S(k)$ is reduced to the DHO model as:

$$S(k,\omega)/S(k) = \frac{1}{\pi}\frac{f_{un}^2(k)z_u(k)}{(\omega^2 - f_{un}^2)^2 + (\omega z_u(k))^2}.$$

In this approximation, there are only two Brillouin peaks without a Rayleigh peak. Usually an emperical Lorentizan term is added to simulate the Rayleigh peak when this model is applied[6, 7].

(ii)When $z_u(k) = 0$, $S(k,\omega)/S(k)$ is reduced to the so-called viscoelastic model as

$$S(k,\omega)/S(k) = \frac{1}{\pi}\frac{z_T f_{un}^2(k) f_{uT}^2(k)}{z_T^2(\omega^2 - f_{un}^2(k))^2 + \omega^2(\omega^2 - f_{un}^2(k) - f_{uT}^2(k))^2}$$

2.3 Results at the Hydrodynamic Limit for the GTEE Theory

In the hydrodynamic limit ($k \to 0$), the matrix element of $\mathbf{H(k)}$ can be shown to reduce to the following forms in the leading order of k expansion as

$$\begin{aligned} f_{un}(k) &= kc_s/\sqrt{\gamma}, \\ z_u(k) &= \phi k^2, \\ f_{uT}(k) &= kc_s\sqrt{(\gamma-1)/\gamma}, \\ z_T(k) &= \gamma D_T k^2, \end{aligned} \tag{2.45}$$

where c_s is the adiabatic speed of sound, ϕ the longitudinal viscosity, $\gamma = c_p/c_v$ the ratio of specific heat at constant pressure and volume, D_T the thermal diffusivity.

Γ_h, Γ_s, and Ω_s can be also solved in the hydrodynamic limit up to the order of $O(k^2)$,

$$\begin{aligned} \Gamma_h &= D_T k^2 \\ \Omega_s &= c_s k \\ \Gamma_s &= (\phi/2 + (\gamma-1)/2D_T)k^2 \end{aligned} \tag{2.46}$$

Chapter 3

Analyses of the IXS spectra of aligned DNA molecules

Due to its biological importance, the structure and dynamics of DNA molecules have been intensively studied both experimentally and theoretically [18, 19, 20, 21, 22, 23, 24, 25, 26].

Evidences from different kind of experiments show that the internal dynamics of DNA molecules is closely coupled to that of the solvent molecules. For example, by studying the average mean square displacement of hydrogen atoms in DNA molecules with inelastic neutron scattering, it is found that DNA molecules undergo a kind of kinetic glass transition in the vicinity of $T \sim 200K$[27, 28], the exact transition temperature of which depends on the hydration level. A similar kind of transition in protein molecules is also observed[29, 30, 31]. The sound speed of hydrated DNA molecules in the hydrodynamic limit ($Q \rightarrow 0$) has been measured with Brillouin light scattering [32, 33, 34]. The longitudinal acoustic ultra sound speed of Na-DNA molecule decreases from about 3500 m/s to about 1800 m/s, when the relative humidity (rh) increases from 23% to 98%[33]. This coupling of the internal dynamics of a DNA molecule to the dynamics of hydration water molecules has been explained in terms of a layered structure of water molecules around DNA [20, 34]. The damping mechanism was attributed to the structural relaxation of water molecules.

It seems that this coupling is also affected by different kind of counterions [21,

26, 33, 35]. The damping of the Brillouin peaks is different in samples of Na-DNA and Li-DNA at the same relative humidity condition [33]. The relaxation of wet DNA in gigahertz range with dielectric measurements shows a two-step relaxation with very strong counterion dependence of the slow relaxation mode [21]. The low frequency mode at $\sim 20\ cm^{-1}$ in the Raman spectra of wet DNA gels were studied at different temperatures. Their temperature dependence is strongly affected by the added counterions[35].

With the availability of the high resolution IXS technique[4] and the high brilliance of the third generation of synchrotron x-ray sources, it is now possible to study the high frequency collective motions of biological samples[13, 14, 15, 16] and molecular liquids[6, 12, 36, 37], whose collective excitation energy has the magnitude of tens of meV. IXS can measure the dynamic structure factor, $S(Q, E)$, as functions of the transferred energy, $E = \hbar\omega$, and the magnitude of the transferred wave vector, Q. The advantage of the IXS technique is that it has a wide energy window and can cover excitations with very large Q (wavevector) range.

Using the IXS technique, we study collective motions of aligned 40 $wt\%$ calf-thymus Na-DNA sample in different solutions with different counterions, and constructed the longitudinal phonon dispersion relation along the axial direction of DNA molecules from $Q = 1.6\ nm^{-1}$ to $Q = 30\ nm^{-1}$ for the first time. Strictly speaking, a DNA rod is not a one dimensional crystal since it does not repeat its structure rigorously along its axial direction. Nevertheless, its IXS spectra show phonon-like excitations. Therefore, we still call these collective excitations phonons. As far as the density fluctuation is concerned, we may neglect the detailed structure of different base pairs and their different orientations, and consider it approximately as a one dimensional crystal. The sound speed along the Na-DNA rod can be extracted. And the damping effect due to different counterions is compared and discussed.

3.1 sample Preparation and Experimental Setup

Highly polymerized (Calf-thymus) DNA has been purchased from Sigma (D-1501, batch number 091K7030) and used without further purification. UV absorption of dilute DNA solution in 0.1 M phosphate buffer (pH7.5) measured with a Perkin Elmer Lambda900 spectrometer, yielded a $A260/A280 = 1.9$. Calf-thymus DNA consists of about 13000 base pairs; its estimated molecular weight is about 8.4×10^6 by considering an average molar mass per nucleotide unit of 324.5. DNA white threads were kept overnight in a dessicator under vacuum, then weighed in a glass vial. The appropriate amount of water or solution was added and the sample vigorously shaken until a homogeneous paste was obtained. This paste was deposited on a $10mm \times 5mm$ quartz slide and shear aligned with the help of a second slide repeatedly rubbed along the 10 mm direction. The samples were then sandwiched between the two slides in order to minimize evaporation during measurement; possible water loss was monitored by weighing the sample before and after measurements.

In order to monitor the DNA alignment, the structure factor $S(Q)$ of different samples at different orientations were measured. Fig. (3-1) shows the measured structure factor $S(Q)$ at different orientations of a DNA sample with respect to the incident X-ray beam. Each cylinder at the top panel represents a single DNA rod. The sample is $40wt\%$ calf-thymus Na-DNA in H_2O. When the incident X-ray beam is perpendicular to the DNA rods, $S(Q)$ shows a very strong scattering peak at $Q = 18.7\ nm^{-1}$ arising from the periodic separation between base pairs along the axis of the double helix, whose distance D is $0.333\ nm^{-1}$, which is a signature of B-DNA. When the incident X-ray beam is parallel to the DNA rods, the prominent Bragg scattering peak at about $Q = 2.4\ nm^{-1}$ clearly indicates the inter-rods correlation effect due to the ordered structure of DNA rods in liquid crystalline phases. By changing the orientations of samples, the alignment of different samples can thus be carefully checked.

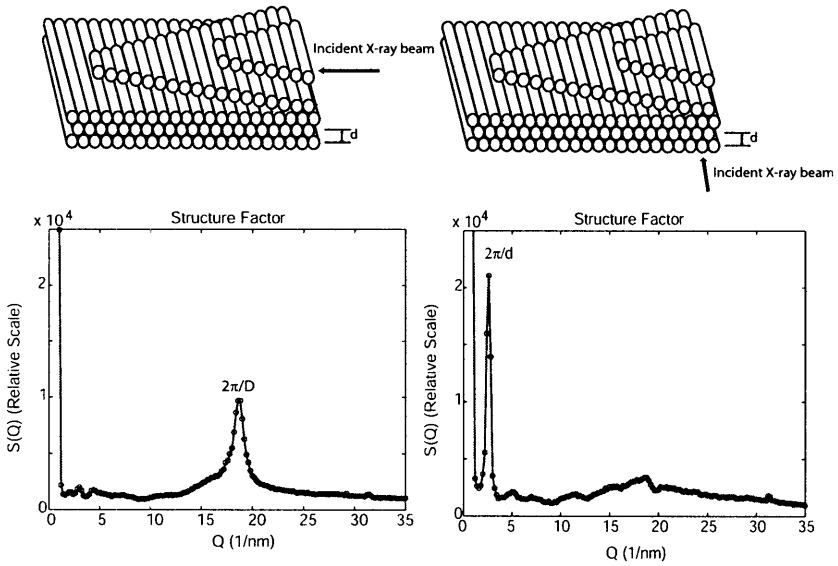

Figure 3-1: The measured $S(Q)$ are shown at different orientations of the $Na - DNA - H_2O$ sample. Each cylinder in the top panel represents a single DNA rod. When the incident X-ray beam is perpendicular to the DNA rods, $S(Q)$ shows a very strong scattering peak at $Q \approx 18.7\ nm^{-1}$, which corresponds to the neighboring base pair distance, $D = 0.333\ nm$. When the incident X-ray beam is parallel to the DNA rods, the prominent Bragg scattering peak at about $Q = 2.4\ nm^{-1}$, which arises from the inter-rod correlation with the inter-rod distance, d, about 2.6 nm

DNA concentration was $40wt\%$ for all samples. We have studied the 40 $wt\%$ calf-thymus Na-DNA samples in H_2O (no additional counterions), and Na-DNA molecules in divalent counterions solutions, 0.083 M $MgCl_2$, 0.083 M $CalCl_2$, 0.083 M $ZnCl_2$, 0.083 M $CuCl_2$, and Na-DNA molecules in trivalent counterions, 0.042 M *spermidine*.

For simplicity, we will call the sample as the $Na-DNA-H_2O$ sample, the $Na-DNA-Mg$ sample, the $Na-DNA-Ca$ sample, the $Na-DNA-Zn$ sample, the $Na-DNA-Cu$ sample, the $Na-DNA-$ *spermidine* sample.

We have also tried the $40wt\%$ Na-DNA in 1 M $MgCl_2$. This sample is called the $Na-DNA-Mg-1M$ sample.

3.2 Method for Data Analysis

When the resolution function is a delta function, the IXS spectra are expressed in Eq. (1.2). After normalization, it becomes $S(Q,E)/S(Q)$.

However, any IXS instrument has a finite energy resolution. A measured IXS spectrum $S_m(Q,E)$ can be expressed as

$$S_m(Q,E) = (S_D(E)S(Q,E)) \otimes R(E)$$

where $S(Q,E)$ is the classical dynamic structure factor of the sample, $S_D(E)$ the detailed balance factor, $R(E)$ the resolution function given by Eq. (1.12), and $\otimes$ means convolution. Typically, $S(Q,E)$ is calculated with a classical model without considering the quantum mechanical effect in the scattering. A classical $S(Q,E)$ is an even function of E. In the thermodynamic state of interest here, we do not expect the quantum mechanic effect dominating so that the classical description is still valid. However, the measured spectra are not balanced at energy gain and energy loss due to the different population at different energy level given by Boltzmann distribution. Therefore the final dynamical structure factor should satisfy the detailed balance condition[2, 3, 38]. This detailed balance condition is commonly fulfilled by empirically multiplying $S(Q,E)$ with a factor $S_D(E)$ [3]. Here we choose $S_D(E) =$

$e^{\frac{E}{2k_BT}}$, where k_B is the Boltzmann constant, and T the temperature in absolute scale.

The energy resolution function of the IXS instrument at Advanced Photon Source in Argonne National Lab is give by Eq. (1.12). $S(Q,E)$ is calculated with the GTEE theory by Eq. (2.34). f_{un}, z_u, z_T, and f_{uT} are treated as the fitting parameters. Among them, f_{un} can be estimated by the second moment of $S(Q,E)$. However, this estimation is not trivial due to the effect of the detailed balance factor.

In the following part of this section, we will describe the method to estimate f_{un}, which can be given as

$$f_{un} = \sqrt{M_2(S(Q,E))/S(Q)} = \sqrt{M_2(Q)/M_0(Q)} \tag{3.1}$$

where

$$M_n(y) = \int_{-\infty}^{\infty} E^n y(E) dE. \tag{3.2}$$

If the resolution function $R(E)$ is an even function, which is the case in our analyses, it is easy to show that

$$M_2(S_D(E)S(Q,E)) = M_2(S_m(Q,E)) - M_2(R(E)) \times M_0(S_D(E)S(Q,E)). \tag{3.3}$$

In the following, for simplicity, we will drop function parameters in equations.

If the absolute value of transferred energy of $S(Q,E)$ is much smaller than $2k_BT$, $S_D(E) \approx 1$ and $M_0(S) = M_0(S_m)$. Therefore,

$$M_2(S) = M_2(S_m) - M_2(R)M_0(S) \tag{3.4}$$

and

$$f_{un}^2 = M_2(S_m)/M_0(S_m) - M_2(R). \tag{3.5}$$

If the absolute value of transferred energy of $S(Q,E)$ is comparable to $2k_BT$, we have to expand the detailed balance factor $S_D(E)$ in the Taylor series to obtain an

approximate result of $S(Q,E)$ as,

$$M_2(S) \approx M_2(S_m) - M_2(R)M_0(S_m) - \frac{\beta}{4}(M_3(S_m) - M_2(R)M_1(S_m)) \tag{3.6}$$

which is a function of different moments determined by experiments. In order to estimate f_{un}, we have to further estimate $M_0(S)$, which can be related to $M_2(S)$ as

$$M_0(S) \approx M_0(S_m) - \frac{\beta^2}{8}M_2(S). \tag{3.7}$$

The estimation of f_{un} is thus expressed as

$$f_{un}^2 \approx \frac{M_2(S)}{M_0(S_m) - \beta^2 M_2(S)/8}. \tag{3.8}$$

Since our experiments were performed at room temperature, $2k_BT$ is about 50 meV. Because the absolute value of transferred energy in $S(Q,E)$ of most experiments is always smaller than 40 meV, we could estimate f_{un} from Eq. (3.8). However, when values of $S(Q,E)$ extends to larger energy range, the estimation become less reliable. Generally the estimated result is smaller than the real result since we always estimate the second moment of a finite spectrum. Nevertheless, this estimation can still give a good initial value for the fitting procedures.

It may be noted that the second moment of the resolution function itself is not convergent, which may bring questions about the validity of the estimation method described above. The divergence of the second moment of the energy resolution function dose not affect our method. Since $S(Q,E)$ calculated with the GTEE theory has finite second moment, we can take a cutoff when the resolution function become the same with measured spectra and calculate the second moments within that finite E range by the method given above.

3.3 Results and Discussions

All the analyses of the IXS spectra have used the Eq. (3.8) to estimate the initial value of f_{un}.

By fitting with the GTEE theory, we have implied that these collective excitation features are only from the longitudinal acoustic (LA) mode. There are several reasons for it. It has been shown that the transverse acoustic (TA) mode disappears at higher hydration level ($rh > 75\%$)[33]. All the samples we studied correspond to the Na-DNA molecule at $rh \simeq 95\%$ studied in Ref[33]. Hence we assume that the TA mode is invisible in our spectra. Although the optical mode was theoretically predicted to be in the order of 80 cm^{-1} [39], which is in the energy range of our IXS spectra, the recent numerical simulation shows that the optical mode may be softened at the room temperature and hence only contribute to such low energy excitations [40] that most of them are masked by the instrumental resolution. Since these works assume the ideal conditions such as neglecting the solvent molecules, we can not estimate how much the modes will change in our system. If we assume the optical mode is similar to the numerical simulation result in Ref[40], the majority of the optical excitation is still within the resolution function. Therefore, at this stage, we assume there is no contribution from the optical mode and the TA mode. From our fitting, the phonon excitation energy behaves linearly at small Q range. Thus it is natural to assign this excitation to the LA mode. However, if the energy resolution of the IXS instrument can be further improved, it is possible to analyze the spectra in a finer way.

Since in our system, there are multi-species of atoms, all of them will contribute to the $S(Q, E)$. Therefore, we tend to think this acoustic wave is due to the collective vibrations of both DNA molecules and water molecules.

3.3.1 Dynamic Structure Factors of Na-DNA Samples

In Fig. (3-2), we show a typical fitting of an IXS spectrum of the $Na - DNA - H_2O$ sample with the GTEE theory at the room temperature. The existence of

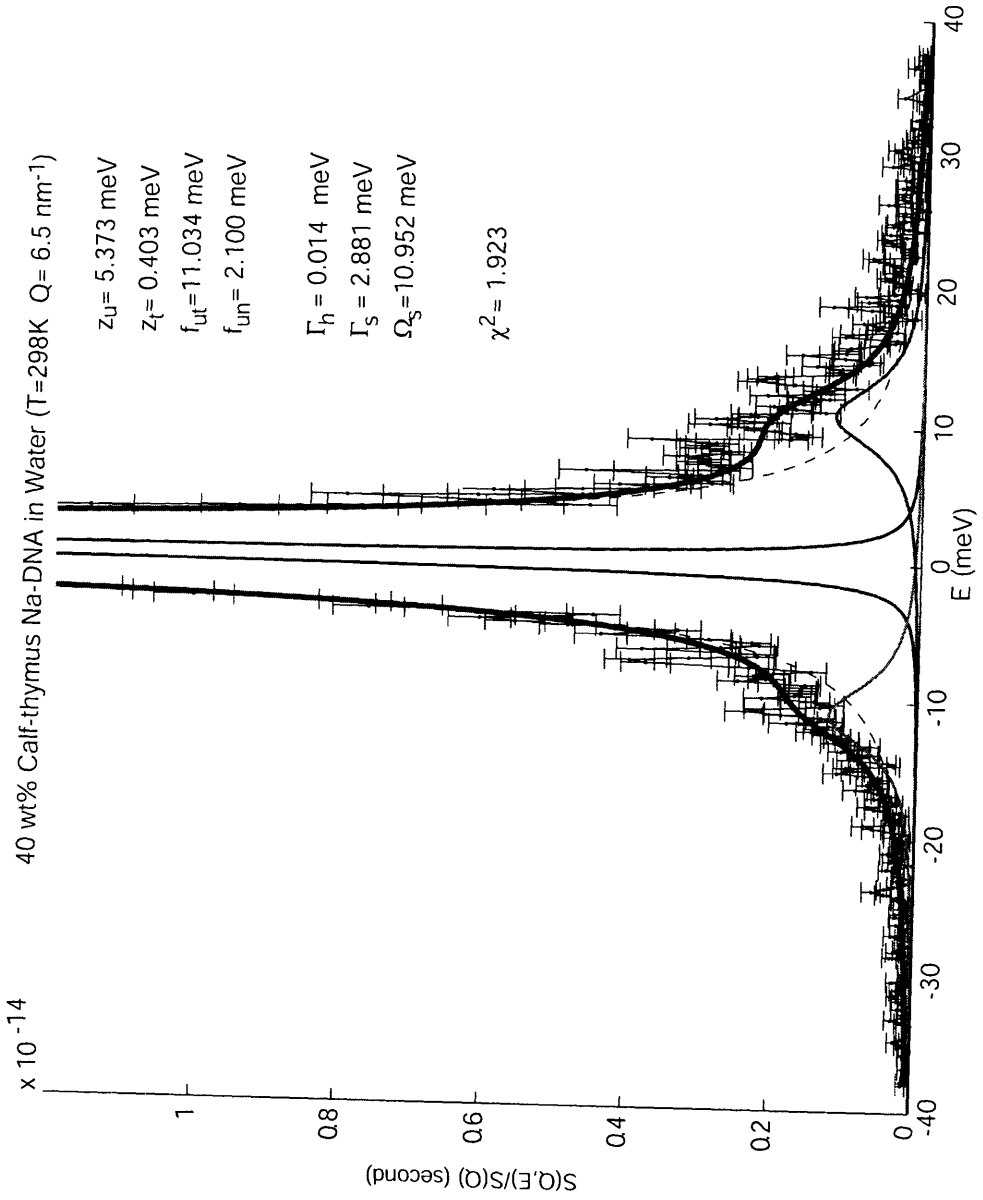

Figure 3-2: Figure illustrates a typical fitting of an IXS spectrum of the $Na-DNA-H_2O$ sample with the GTEE theory. The thin solid lines depicts the three generalized hydrodynamic modes. The dashed line is the resolution function. The thick solid line represents the fitter curves after considering the energy resolution function.

well defined phonon peaks is clearly visible on the right and left hand shoulders of the spectrum. The thin solid line depicts the three generalized hydrodynamic modes, one central and two symmetrically shifted Brillouin modes (phonon). The dashed line is the instrumental resolution function. The thick solid line is the fitted curve, which includes the energy resolution function. The fit is made in an absolute scale with four generalized hydrodynamic parameters: z_u, z_T, f_{uT}, and f_{un}. The fitting results uniquely give the phonon frequency $\Omega_s = 10.95\ meV$, phonon damping $\Gamma_s = 2.88\ meV$, and the width of central peak $\Gamma_h = 0.014\ meV$. The χ^2 of this fitting is about 1.9.

In Fig. (3-3), we show the extracted $S(Q, E)$ by analyzing the IXS spectra of both the $Na - DNA - H_2O$ sample and the $Na - DNA - Mg$ sample. From top to bottom, the Q values are 9.0 nm^{-1}, 12.5 nm^{-1}, and 22.5 nm^{-1} respectively. At the same Q value, the phonon peaks are much more well defined in the spectra of the $Na - DNA - H_2O$ sample than those from the $Na - DNA - Mg$ sample.

In Fig. (3-4), we show the decomposed three Lorentzian peaks calculated by the GTEE theory with the fitted parameters. At $Q = 9.0\ nm^{-1}$, the three peaks consist of one central peak flanked by two shifted Brillouin peaks. For the $Na - DNA - H_2O$ sample, the phonon frequency is $\Omega_s = 11.9\ meV$, the phonon damping 4.5 meV, and the line width of the central peak 0.055 meV. While for the $Na - DNA - Mg$ sample, these results are 10.99 meV, 9.747 meV, and 0.052 meV respectively. It is interesting to note that the phonon damping is stronger in the $Na - DNA - Mg$ sample than that in the $Na - DNA - H_2O$ sample. At $Q = 12.5\ nm^{-1}$ and 22.5 nm^{-1}, the phonon peaks disappear in the $Na - DNA - Mg$ sample. The fittings give three Lorentizan peaks all centered at $E = 0\ meV$. while in the $Na - DNA - H_2O$ sample, Brillouin peaks can still be identified. In the case of $Q = 1.25\ nm^{-1}$, the Q value approaches

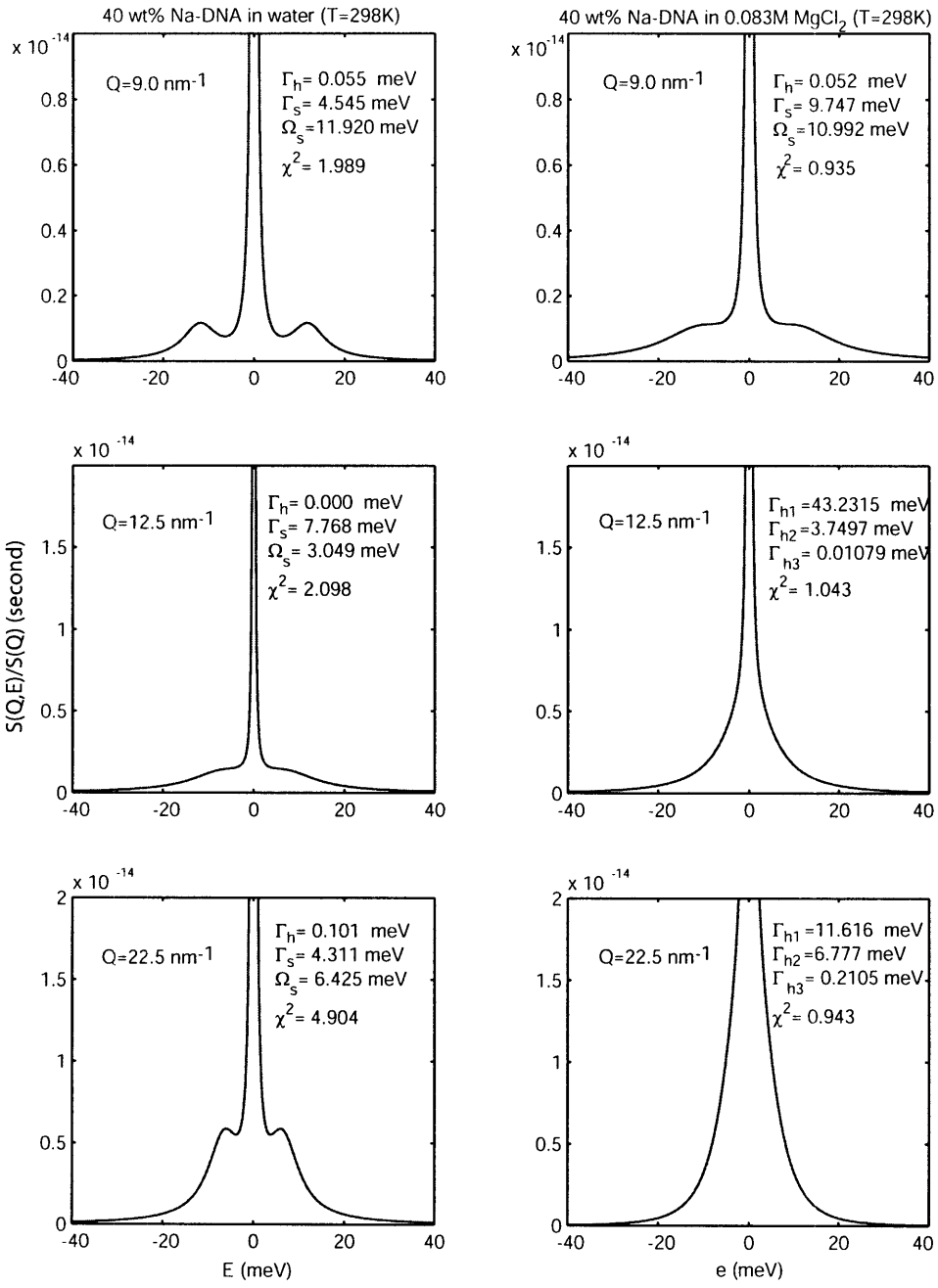

Figure 3-3: Figure shows the extracted dynamic structure factors of two Na-DNA samples with the GTEE theory.

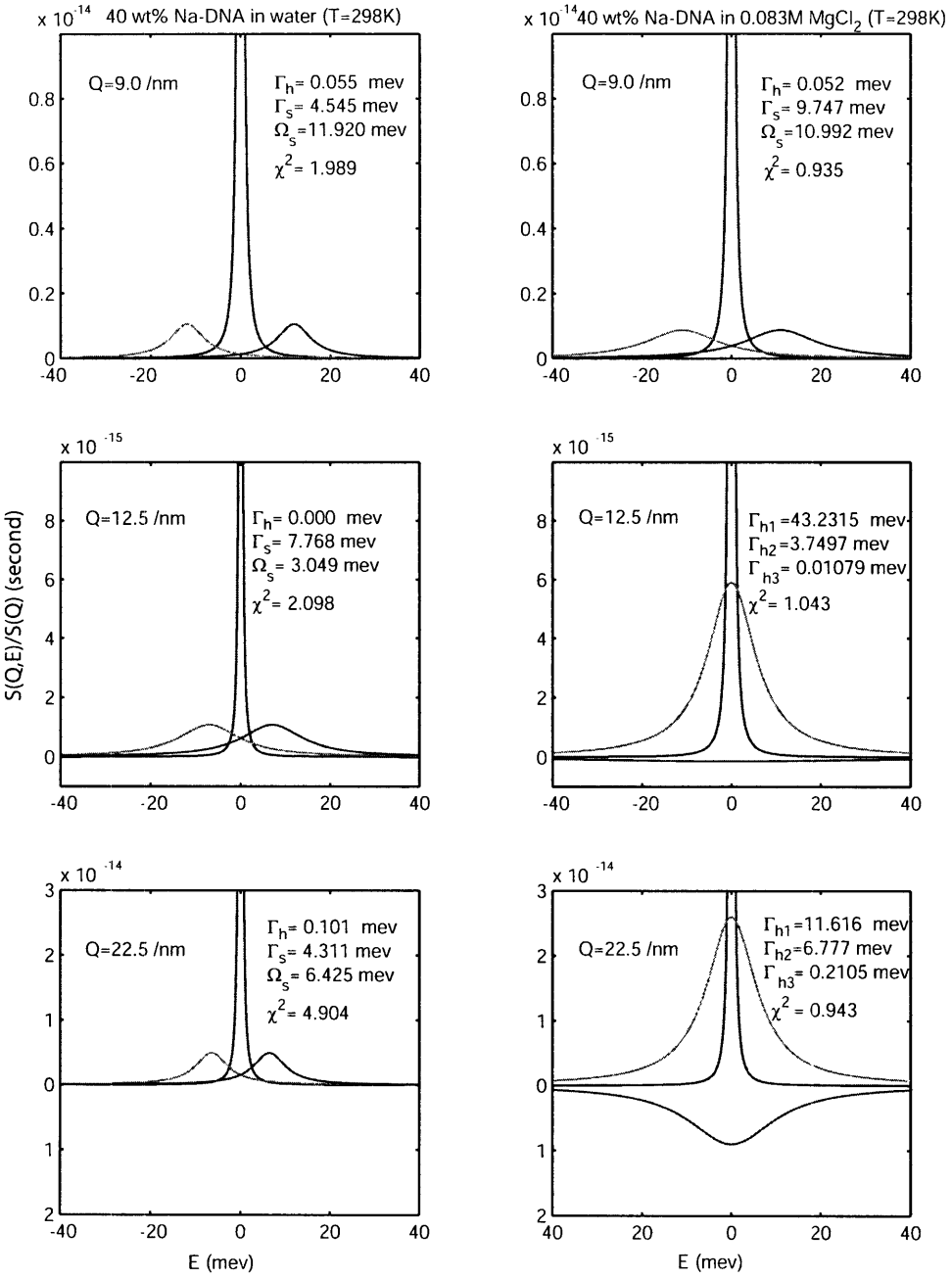

Figure 3-4: Figure shows the decomposed three Lorentzian peaks of $S(Q, E)$ calculated with fitted parameters. The first column shows the results of the $Na - DNA - H_2O$ sample, while the second column shows the results of the $Na - DNA - Mg$ sample.

the diffractoin peak of the structure factor and the effect of De Gennes narrowing is clearly seen.

With the GTEE theory, we thus can extract the phonon energy at different Q values from IXS spectra of different samples. The phonon dispersion of all DNA samples are thus constructed. We also found that the disappearance of phonon features at some Q range seems actually a general feature for the samples with additional counterions.

3.3.2 Phonon Dispersion Relations of Aligned Na-DNA Samples

By analyzing the IXS spectra with the GTEE theory, we are able to construct the phonon dispersion relations for all the DNA samples we measured.

The top panel of Fig. (3-5) shows the phonon dispersion relation, $\Omega_S(Q)$ *vs.* Q, for the $Na-DNA-H_2O$ sample. $\Omega_s(Q)$ is linearly dependent on Q at low Q values and it reaches the first maximum at $Q \approx Q_{max}/2$, where $Q_{max} = 18.7\ nm^{-1}$ is the position of the Bragg diffraction peak in $S(Q)$. The maximum $\Omega_s(Q)$ is about 12 meV. Then $\Omega_s(Q)$ bends down after $Q > Q_{max}/2$ and reaches zero at around $Q = Q_{max}$. For $Q > Q_{max}$, $\Omega_s(Q)$ again becomes nonzero and reaches a secondary maximum at about $Q = 25\ nm^{-1}$, showing phonon propagation at large Q. The sound speed calculated from the initial slope of the phonon dispersion relation is about 3159 m/s. One major distinctive feature of the DNA system is that the dispersion relation extends to a considerably higher Q than that observed in simple liquids. The extended range of Q includes values corresponding to those in the second Brillouin zone of a crystalline system. This feature implies that a DNA rod can be regarded approximately as a one-dimensional crystal, as far as the density oscillations are concerned.

In order to test the consistency of the fitting, we calculated the structure factor, $S(Q)$, from one of the fitting parameters, $f_{un}(Q)$ with Eq. (2.33). The form

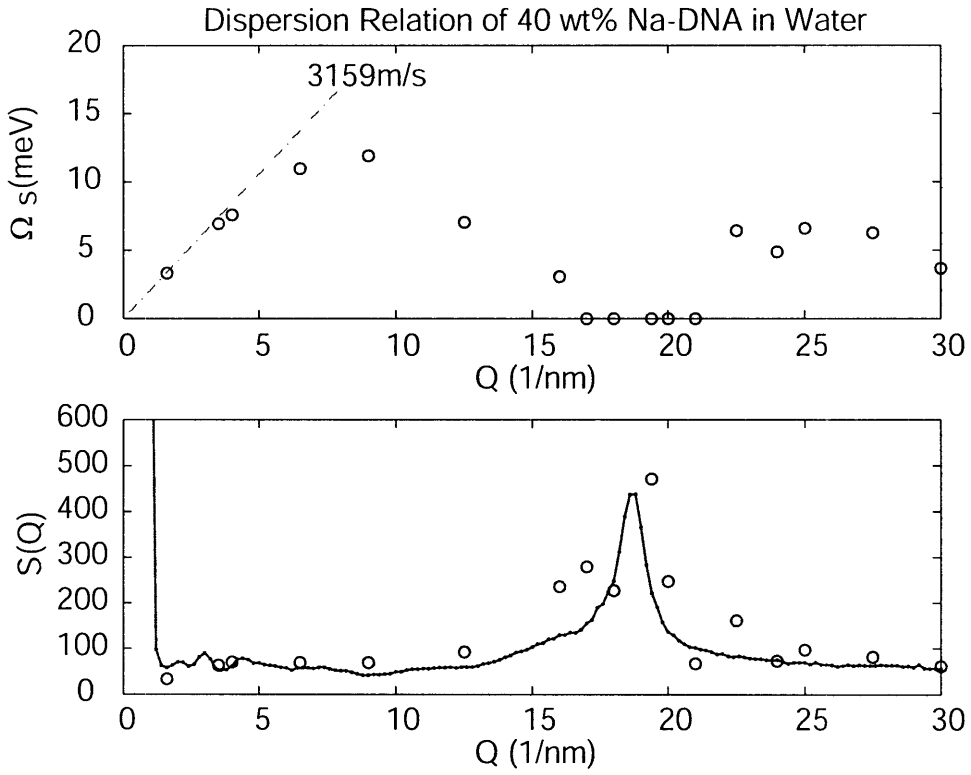

Figure 3-5: Top panel shows the phonon dispersion relation of the $Na - DNA - H_2O$ sample. Bottom panel shows the comparison between the theoretically derived structure factor $S(Q)$ (open circles), in absolute scale, and the measured structure factor (dots with solid line).

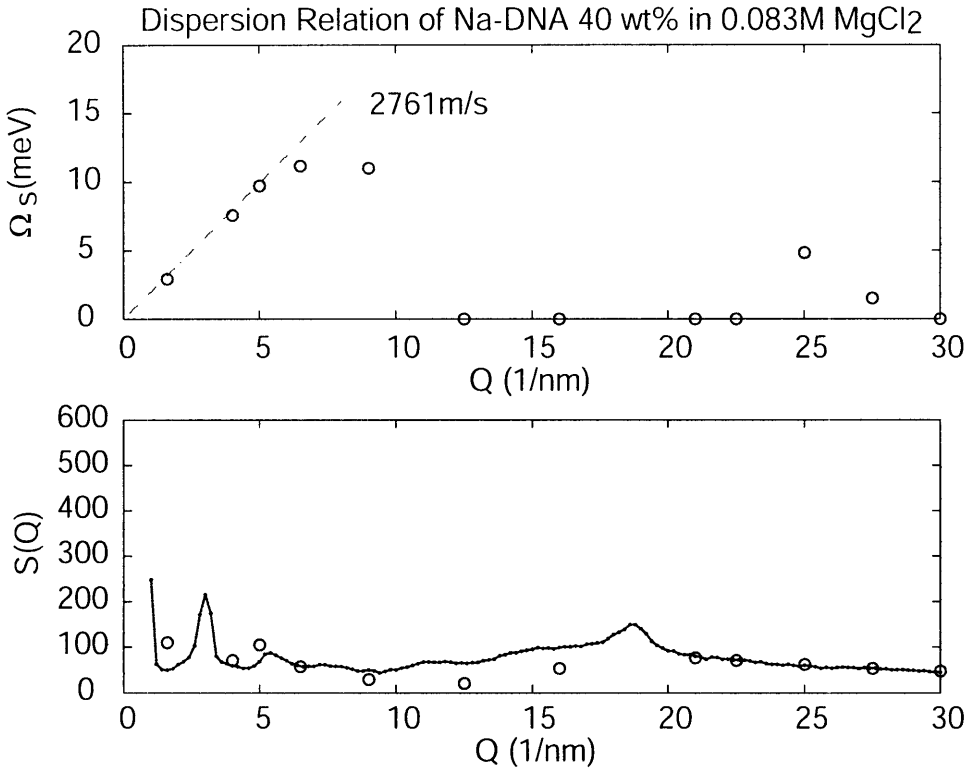

Figure 3-6: Top panel shows the phonon dispersion relation of the $Na - DNA - Mg$ sample. Bottom panel shows the comparison between the theoretically derived structure factor $S(Q)$ (open circles), in absolute scale, and the measured structure factor (dots with solid line).

factor, $f(Q)$, for different atoms can be found in the international tables for X-ray cyrstallography[17]. Therefore, $S(Q) = (Qv_0(Q)/f_{un}(Q))^2$. As an approximation, we have assumed that phosphate have the dominating contribution to $v_0(Q)$. Since the measured structure factor by the instrument is in arbitrary units, the result is scaled by a constant in order to compare with the calculated result. The agreement is fair after considering the simplification of calculation of v_0.

In Fig. (3-6), we show the phonon dispersion relation of the $Na - DNA - Mg$ sample in the top panel. Similarly to the $Na - DNA - H_2O$ sample, the phonon dispersion relation, $\Omega_s(Q)$, is linearly dependent on Q at low Q and it reaches the first maximum at $Q \approx Q_{max}/2$. The maximum of $\Omega_s(Q)$ is about 11 meV. The

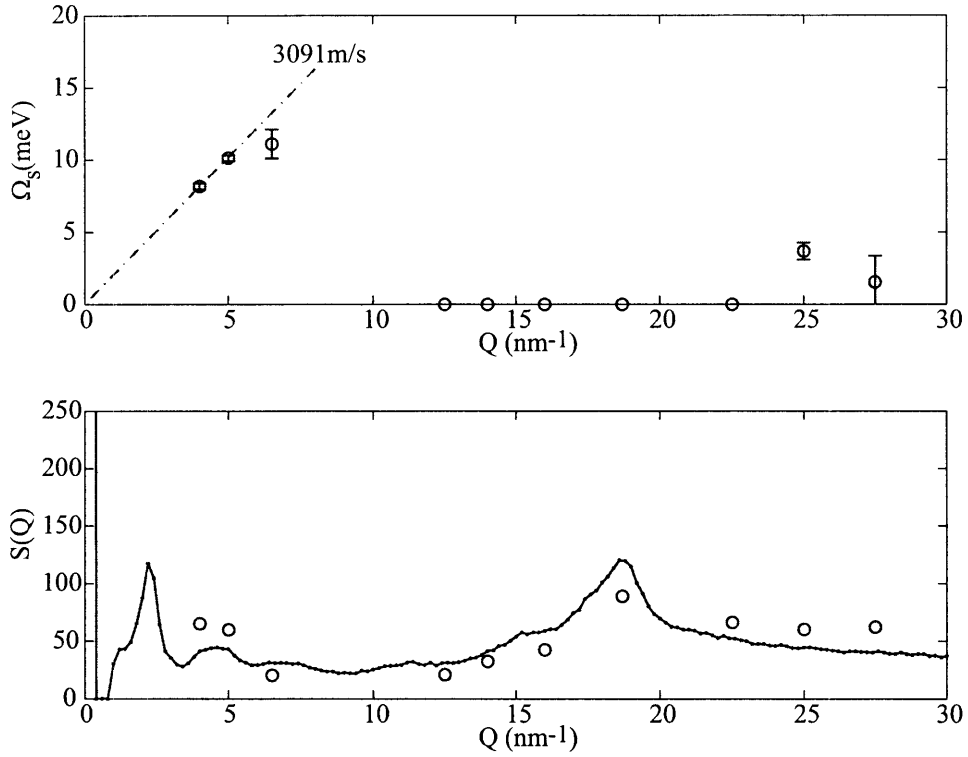

Figure 3-7: Top panel shows the phonon dispersion relation of the $Na - DNA - Zn$ sample. Bottom panel shows the comparison between the theoretically derived structure factor $S(Q)$ (open circles), in absolute scale, and the measured structure factor (dots with solid line).

sound speed obtained from the initial slope of the phonon dispersion relation is about 2761 m/s. When compared with the dispersion relation of the previous sample, the striking difference is related to the Birllouin peaks that completely disappear in the interval from 12.5 nm^{-1} to 22.5 nm^{-1}. This suggests that the different valency of the counterions consistently affects the DNA internal dynamics. This indeed seems to be a general feature for the samples in the presence multivalent counterions.

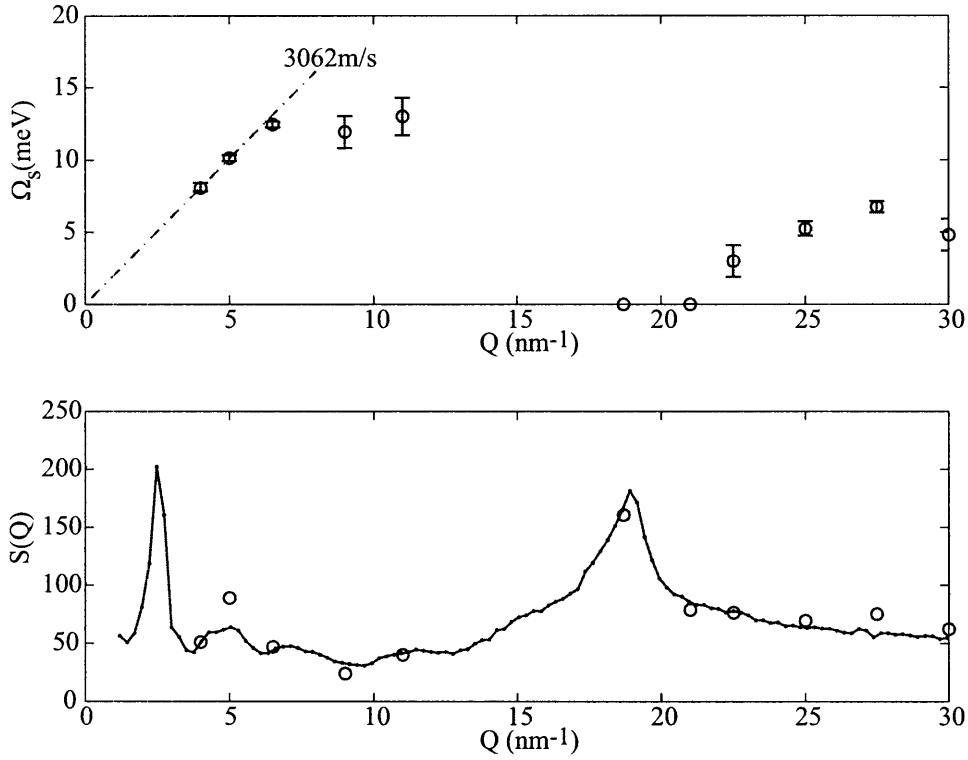

Figure 3-8: Top panel shows the phonon dispersion relation of the $Na - DNA - Ca$ sample. Bottom panel shows the comparison between the theoretically derived structure factor $S(Q)$ (open circles), in absolute scale, and the measured structure factor (dots with solid line).

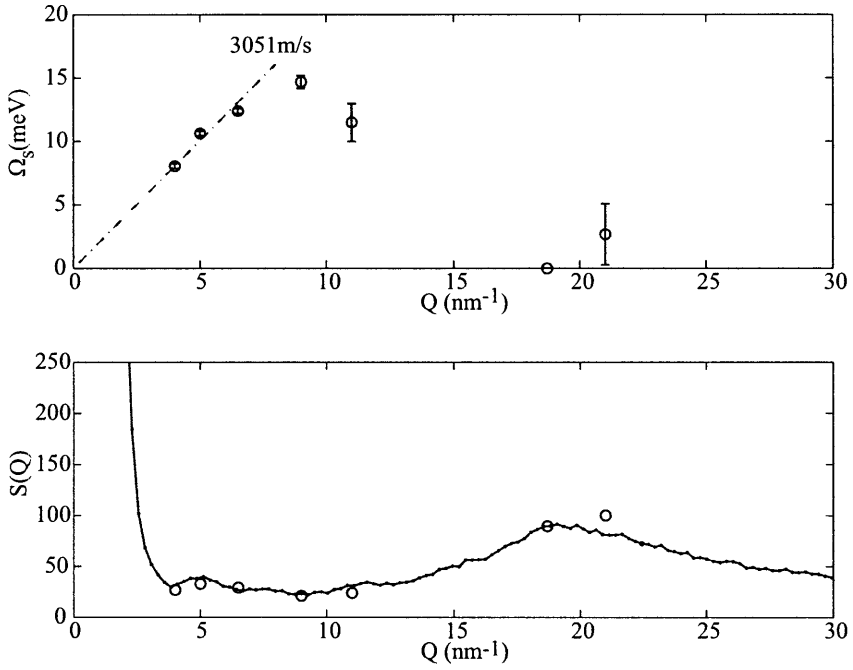

Figure 3-9: Top panel shows the phonon dispersion relation of the $Na - DNA - Cu$ sample. Bottom panel shows the comparison between the theoretically derived structure factor $S(Q)$ (open circles), in absolute scale, and the measured structure factor (dots with solid line).

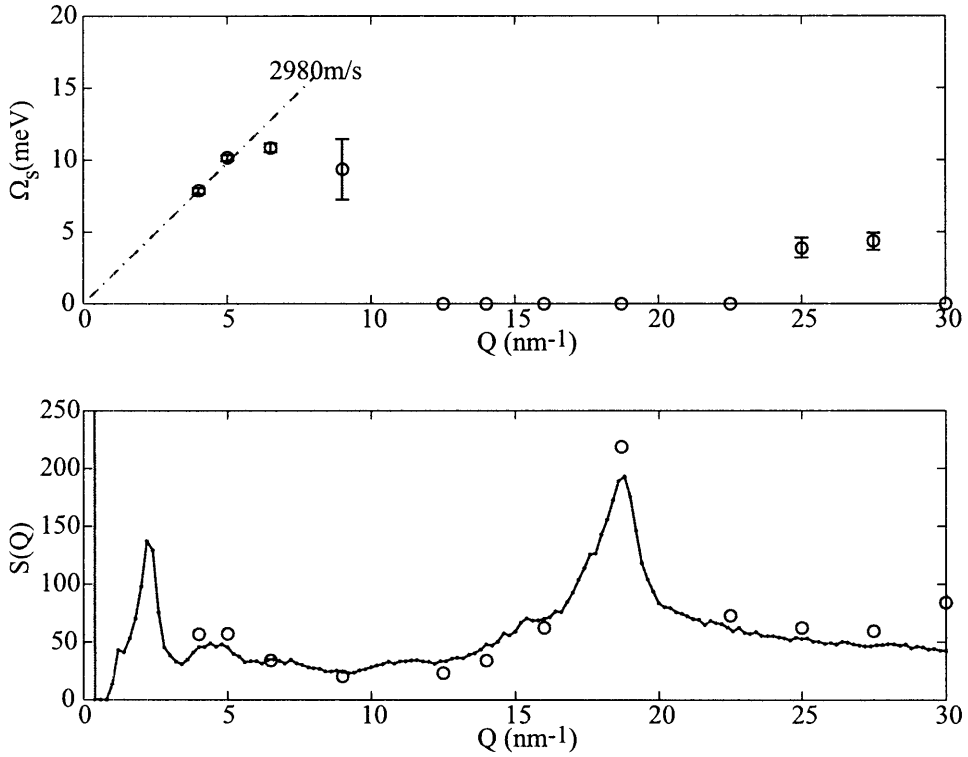

Figure 3-10: Top panel shows the phonon dispersion relation of the $Na - DNA -$ *spermidine* sample. Bottom panel shows the comparison between the theoretically derived structure factor $S(Q)$ (open circles), in absolute scale, and the measured structure factor (dots with solid line).

The dispersion relations for the $Na-DNA-Zn$ sample, $Na-DNA-Ca$ sample, $Na-DNA-Cu$ sample, and $Na-DNA-$ *spermidine* sample, are shown in Fig. (3-7), Fig. (3-8), Fig. (3-9), and Fig. (3-10). The symbols in the bottom panel for these figures have the same meaning as those in Fig. (3-5).

It is interesting to notice between $Q \approx 12.5\ nm^{-1}$ to $Q \approx 22.5\ nm^{-1}$, the phonon disappears for all the samples with added counterions. These results suggest that the addition of positvely charged counterions will increases the damping of the phonons and make them overdamped in some Q range.

The sound speeds obtained from the $Na-DNA-Zn$ sample, $Na-DNA-Ca$ sample, $Na-DNA-Cu$ sample, and $Na-DNA-$ *spermidine* sample, are 3091 m/s, 3062 m/s, 3051 m/s, 2980 m/s respectively. Therefore, the sound speed of aligned calf-thymus Na-DNA molecules is $\sim 3000\ m/s$ at $Q > 1.5\ nm^{-1}$.

This result is about twice the value of the sound speed of Na-DNA molecules, which is $\sim$ 1800 m/s, measured with Brillouin scattering at the similar hydration level[33]. We, therefore, call the sound speed obtained at our Q range as a fast sound speed to distinguish it from the sound speed at very small Q values ($\sim 0.01\ nm^{-1}$), which is the range accessible by Brillouin light scattering.

The reason of having a fast sound speed for a hydrated DNA molecule is attributed to the different dynamic response of water molecular at different Q range. MD computer simulation, the IXS, and INS experimental results have shown that the sound speed of water molecule changes from $\sim 1500\ m/s$[6, 7, 41] at the hydrodynamic limit to $\sim 3000\ m/s$[6, 12] when $Q > 4.0\ nm^{-1}$ at the room temperature. Therefore, we speculate that at hydrodynamic limit, it is the slow dynamics of water molecules that slows down the overall sound speed of hydrated DNA molecules. In the Q range of our experiments, the fast dynamics of water molecules naturally increase the overall sound speed. Hence, not only the hydration level can change the sound speed, but also the dynamic response at different Q range of water molecules will affect the sound speed of DNA molecules. This further proves the close coupling between the DNA molecules and their surrounding water molecules, which have also been seen by

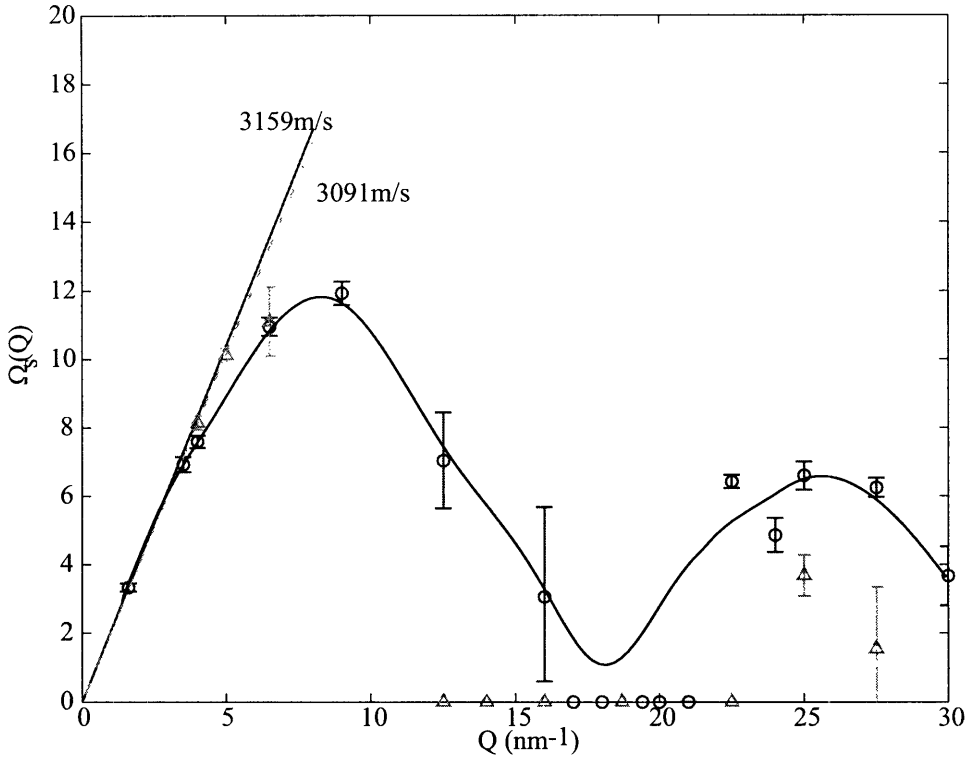

Figure 3-11: The figure shows the comparison of the phonon dispersion relations of the $Na - DNA - Mg$ sample, and the $Na - DNA - Mg - 1M$ sample.

other different techniques[32, 33]. However, it is very difficult with our experiment to know how water molecules interfere with the atoms in a DNA molecule. A computer simulation will be helpful to reveal the relations.

Although the dispersion relation of DNA molecules with different counterions shows consistently linear relation at $Q <\sim 6.5\ nm^{-1}$ with a similar sound speed, the excitation energy differs from each other at the Q range centering around $\sim 9\ nm^{-1}$, and $\sim 25\ nm^{-1}$. The reason for this counterion dependence at these Q ranges is not clear to us at present. The maxima excitation energy of all the phonon dispersion relation is about $\sim 11\ meV$. This therefore suggests that the very broad 80 cm^{-1} mode observed in Raman scattering in most DNA molecule samples may be due to the LA mode according to our observations. This mode has been assigned to the "intrahelical" collective motions of a DNA molecule[19].

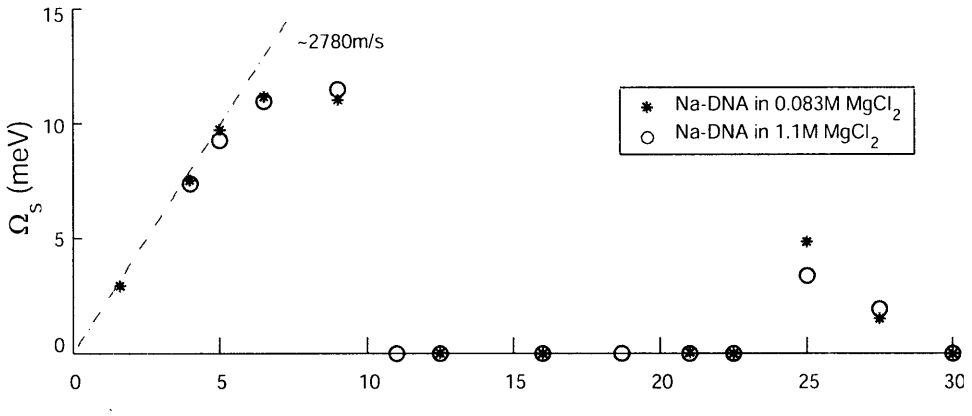

Figure 3-12: The figure shows the comparison of the phonon dispersion relations of the $Na - DNA - Mg$ sample, and the $Na - DNA - Mg - 1M$ sample.

To compare the phonon dispersion relation curves with/without the addition of divalent counterions, we show the results from 40 $wt\%$ calf-thymus DNA molecules in H_2O (circles), and the Zn^{2+} sample (up triangles) together in Fig. (3-11). The sound speed of aligned $40wt\%$ DNA molecules in H_2O is about 3159 m/s estimated from the phonon excitation energy at $Q = 1.6\ nm^{-1}$. It clearly shows the disappearance of phonon at the intermediate range due to stronger damping introduced by additional counter ions. The curve is drawn to guide the eyes.

We have also measured the 40 $wt\%$ Na-DNA sample in $1M\ MgCl_2$. The dispersion relation of this $Na - DNA - Mg - 1M$ sample is compared with that of the $Na - DNA - Mg$ sample in Fig. (3-12). The open circles represent the results of the $Na - DNA - Mg - 1M$ sample, while the star symbols represent the results of the $Na - DNA - Mg$ sample. These two phonon dispersion relation curves are almost the same. The sound speed of the $Na - DNA - Mg$ is about 2761 m/s and the sound speed of the $Na - DNA - Mg - 1M$ is about 2197 m/s. In the figure, the sound speed is indicated as the average of the sound speed of those of two sample. The

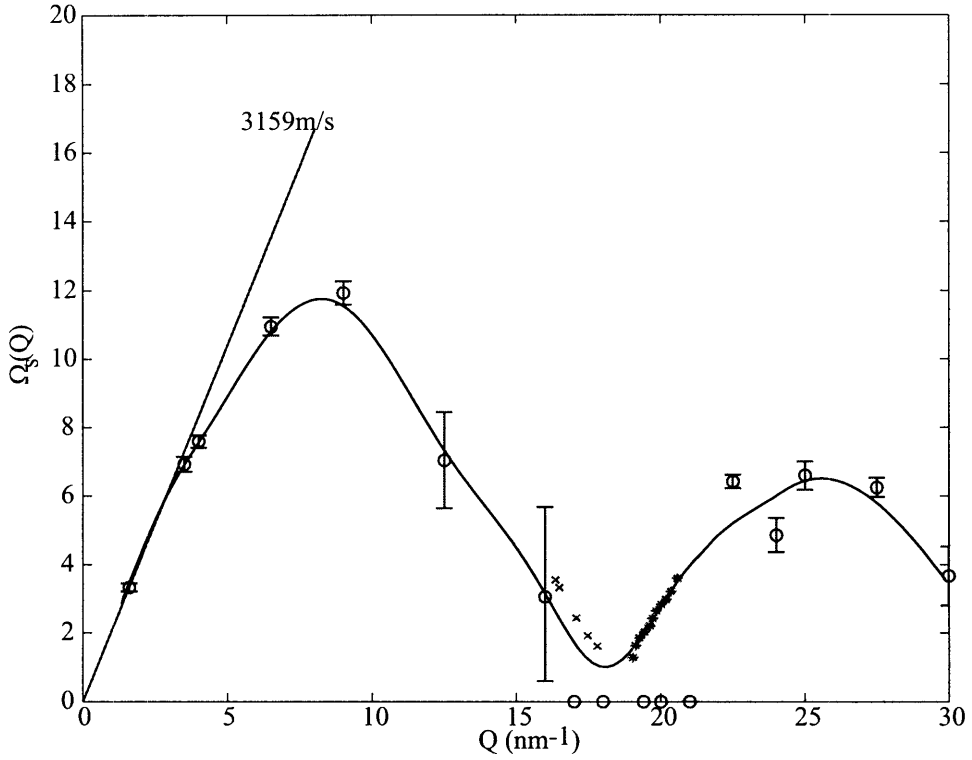

Figure 3-13: The phonon dispersion relation of the $Na - DNA - H_2O$ sample (open circles) is shown together with the results measured by inelastic neutron scattering from Ref[24] (star and cross symbols). The star symbols represent the results by fitting INS spectra with their model, while the cross symbols are calculated results. The solid line is drawn to guide eyes.

sound speed is almost unchanged when the concentration of $MgCl_2$ increases from 0.083 M to 1.1 M, while the increase of $MgCl_2$ concentrations from 0 M to 0.083 M decreases the sound speed from 3100 m/s to 2761 m/s.

The acoustic phonons DNA molecules have been measured by inelastic neutron scattering (INS) by Grimm and co-workers[24]. Grimm's measurements covered the Q range from about 16 nm^{-1} to 21 nm^{-1}. Maret's measurement focused on very small Q values (10^{-2} nm^{-1}). Here we would like to compare our results of the $Na - DNA - H_2O$ sample with their measurements on wet samples, which are hydrated by equilibrating at controlled humidity.

Because Grimm's data is not available for us, we directly extracted the results from a figure of their published paper[24]. Both our results and those of Grimm's are summarized in Fig. (3-13). Open circles show the phonon dispersion relation of the $Na-DNA-H_2O$ sample obtained by us, while the star and cross symbols are Grimm's results. The star symbols were calculated by fitting the INS spectra with one-dimension liquid picture model[42, 43]. The cross symbols were calculated results according to their model. The solid line is drawn to guide eyes. Notice that in our curve, the phonon disappears at a small range centered at $Q = 18.7\ nm^{-1}$, where the Billouin zone center lies if a DNA molecule could be considered as a one-dimensional crystal. We attribute the reason of the failure to identify the phonon peaks in this region to the fact that the energy resolution of IXS spectra is not good enough so that small energy acoustic excitations are masked by the broad resolution functions. According to the results of Fig. (3-13), there are qualitative agreement between these two experimental results.

To extract the sound speed, different models are applied and give different results. The so-called "sound speed" from Grimm's paper is about 2180 m/s. The sound speed we obtain from the acoustic excitations at small Q values is about 3100 m/s. Because of the small Q range in the measurements of Grimm and co-workers due to the limitation of the instrument, they extracted sound speed at relatively large Q values. In our experiments, we could access much lower Q values while keeping the large range of energy scan. Therefore we can directly extract sound speed by assuming the linear relation between excitation energy and Q values at very small Q. This method has been widely used to identify the sound speed of water, liquid metals, and molecular glasses [6, 36, 44].

3.3.3 Analysis of the Phonon Damping

Fig. (3-14) shows the ratio of phonon damping and the phonon frequency as a function of Q for the $Na-DNA-H_2O$ sample, $Na-DNA-Mg$ sample, and $Na-DNA-Mg-1M$ sample, which are shown as solid circle, star, and solid diamond symbols respectively. The solid lines are drawn to guide eyes. It clearly demonstrates that

the increase of $MgCl_2$ concentration consistently increases the phonon damping. This effect can be linked to the increase of viscosity in the presence of $MgCl_2$ concentration. In the hydrodynamic limit, $\Gamma_s/\Omega_s = Q(1/2\phi + 1/2(\gamma - 1)D_T)/c_s$. Therefore, if c_s dose not change too much, the increase of viscosity naturally leads to larger damping results. The increase of stronger damping is thus directly related to the increase of the longitudinal viscosity.

Fig. (3-15) shows the Q dependence of one of the fitted parameters, z_u. The open circle, star, and diamond symbols represent the results for the $Na - DNA - H_2O$ sample, $Na - DNA - Mg$ sample, and $Na - DNA - Mg - 1M$ sample respectively. According to the hydrodynamic limit results, $z_u = \phi Q^2$, where $\phi = (\frac{4}{3}\eta + \zeta)/\rho$ is the longitudinal viscosity. η and ζ are the shear and bulk viscosity, and ρ is the density. The solid lines are best fitting with the function $z_u = \phi Q^2$. From the fitting, ϕ is 0.11, 0.22, and 0.30 $meV\ nm^2$ for the $Na - DNA - H_2O$ sample, $Na - DNA - Mg$ sample, and $Na - DNA - Mg - 1M$ sample respectively. It thus indicates that the increase of counter-ion concentrations increases the viscosity.

Fig. (3-16) shows the phonon damping with error bars for different samples at $Q < 14\ nm^{-1}$. The phonon damping of the $Na - DNA - H_2O$ sample (circles) is much smaller than that of other cases. For the samples with extra added counterions, the spectra of the *spermidine* sample (pentagrams) show largest phonon damping among all the samples presented here. Other three results with divalent counterions are shown as star symbols for $Na - DNA - Ca$ sample, up triangles for $Na - DNA - Zn$ samples, and diamond symbols for $Na - DNA - Cu$ samples. We notice that at $Q = 6.5\ nm^{-1}$, the phonon damping of the $Na - DNA - Zn$ sample is larger than that for the $Na - DNA -$ *spermidine* sample. The mechanism with which the counterions affect the phonon damping is not clear at present. Although some computer simulations studied the motion of counterions around a DNA molecule[45, 46], their effect to the dynamic structure factor has not been checked yet.

It is instructive to compare the ratio between phonon damping and the phonon frequency from different samples, which is shown in Fig. (3-17). The circles with error bars are the result of the $Na - DNA - H_2O$ sample. One of the solid lines is

drawn to cross the results of Na-DNA in H_2O, while another one is drawn to cross the results of the $Na - DNA - spermidine$ samples. These two lines are drawn to guide eyes. The ratio of phonon damping and phonon frequency of samples with divalent counterions falls in between these two lines with the only exception for the case at $Q = 6.5\ nm^{-1}$ for the $Na - DNA - Zn$ sample. At the beginning, the ratio is just a fraction of the phonon energy, which indicates that the IXS spectra contain well defined phonon peaks. With increasing Q values, the ratio becomes larger than one. The damping is so strong that, the phonon peaks becomes not well defined. At the hydrodynamic limit, the ratio between the phonon damping and phonon energy should be linearly dependent on Q. From the results shown in Fig. 3-17, the linear region is only up to about $Q \sim 6.5\ nm^{-1}$, after which the ratio clearly deviate from linear relation with Q.

3.3.4 Relaxation of the Intermediate Scattering Function of Aligned Na-DNA Molecules

With the fitted parameters, we can calculate the ISF with Eq. (2.37) and Eq. (2.40). Fig. (3-18) shows the ISF of $40wt\%$ calf-thymus DNA in H_2O at $Q = 1.6\ nm^{-1}$, $Q = 9.0\ nm^{-1}$, $Q = 18.7\ nm^{-1}$, and $Q = 25.0\ nm^{-1}$. The FWHM of the energy resolution function is about $2\Gamma_R = 2\ meV$, which corresponds to a time window about $0.65\ ps$. In our fitting, $2\Gamma_h$ is the FWHM of the extracted central peak. If we assume that $\Gamma_h > 10\%\Gamma_R$ means that the extracted Γ_h is reliable, the estimated relaxation time from Γ_h is reliable up to $6.5\ ps$. If the relaxation time for some slow dynamics are much larger than $6.5\ ps$, their corresponding energy spectra will be

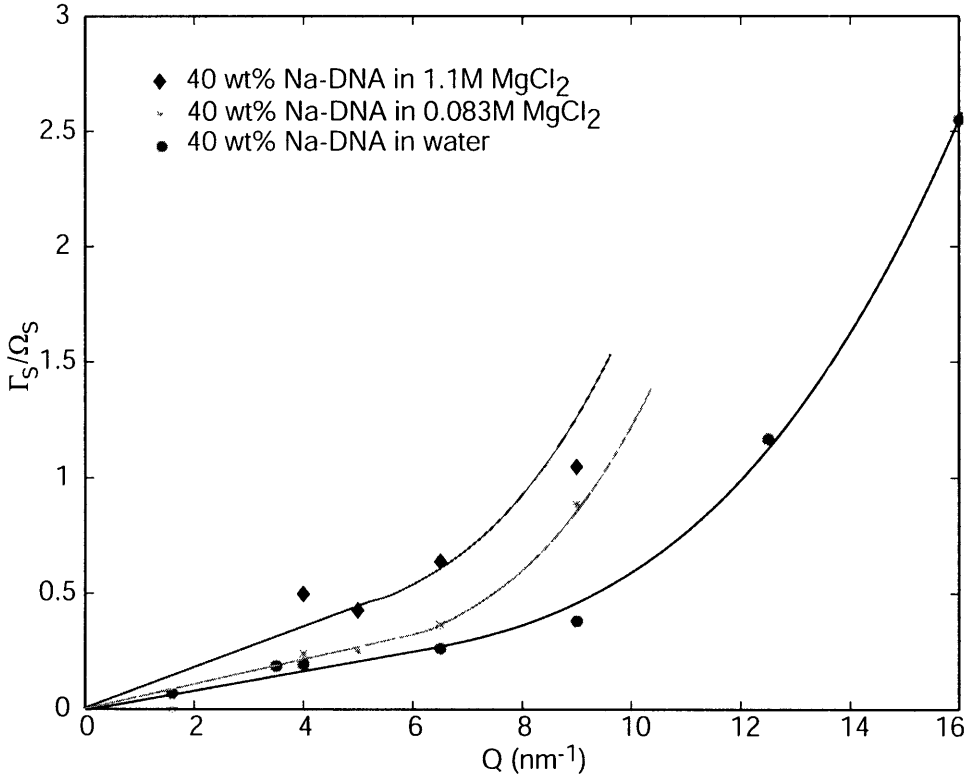

Figure 3-14: Figure shows the Q-dependence of the ratio between the phonon damping, Γ_s, and the phonon frequency, Ω_s at different samples. The solid lines are drawn to guide eyes.

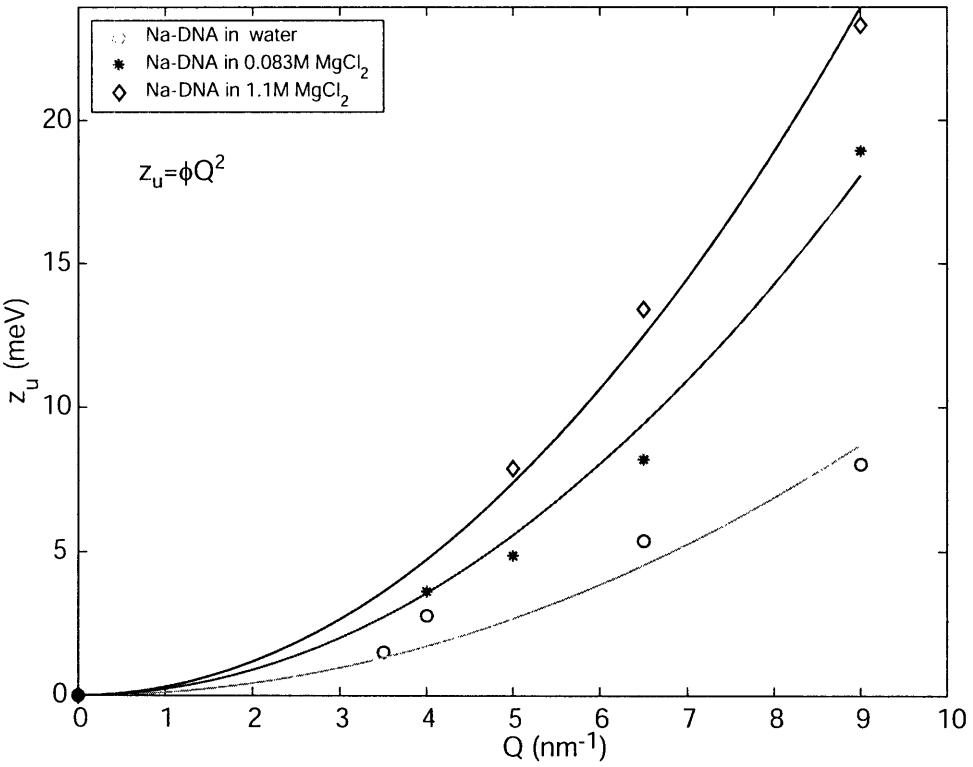

Figure 3-15: Q-dependence of one fitted parameter, z_u, is shown at different conditions. The symbols are results from fitting IXS spectra with the GTEE theory. The solid lines are the best fitting with $z_u = \phi Q^2$ for different samples.

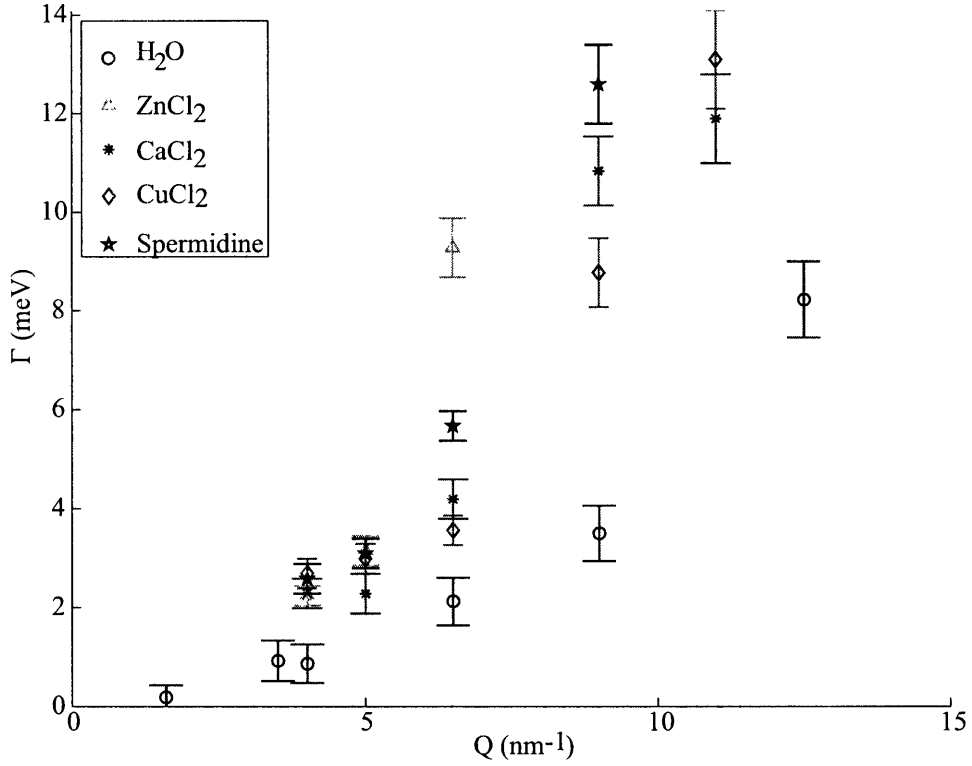

Figure 3-16: Phonon damping of different samples at different Q values is shown. The added counterions in the DNA samples are indicated in the legend. The concentration of the $ZnCl_2$, $CaCl_2$, and $CuCl_2$ are 0.083 M. The spermidine concentration is 0.042 M.

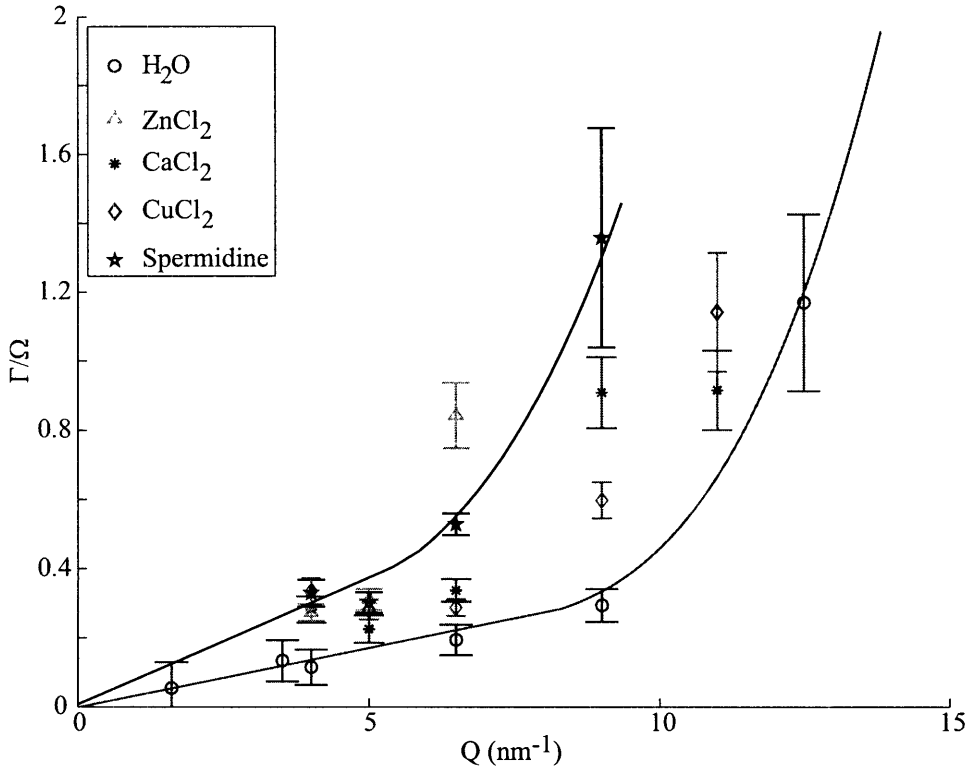

Figure 3-17: Figure shows the ratio between the phonon damping and the phonon frequency of DNA samples with different counterions. Circles with error bars are the results of $Na - DNA - H_2O$ sample. The uptriangle, star, diamond, and pentagram symbols correspond to the results from the $Na - DNA - Zn$, $Na - DNA - Ca$, $Na - DNA - Cu$, and $Na - DNA - spermidine$ samples.

much sharper than the energy resolution function. Therefore, it is very difficult to obtain such kind of slow dynamics information from the IXS spectra. However, it is still possible to extract these information by analyzing the IXS spectra with a model. But the results is highly model dependent. Similar argument holds to the relaxation time estimated with Γ_s. Therefore, a vertical dotted line is thus drawn at $t = 6.5\ ps$ in Fig. (3-18). The extracted results on the left of this line are less model dependent, while the results on the right of this line can be considered as the extrapolation from our model.

In Fig. (3-18), the IXS spectrum at $Q = 1.6\ nm^{-1}$ has a very clear phonon peak with very small damping so that the ISF shows very clear oscillation due to the collective vibration of atoms in the system. The relaxation time for the fast dynamics, τ_f, is about 3.6 ps, while the relaxation time for the slow time dynamics, τ_s, is about 703 ps. When $Q = 9.0\ nm^{-1}$, the decay of the fast dynamics becomes faster due to the much stronger damping with the time domain oscillation barely seen with $\tau_f \sim 0.19\ ps$ and $\tau_s \sim 10\ ps$. At $Q = 18\ nm^{-1}$, close to Brillouin zone center, there is no phonon peak. The ISF also shows a two-step decay. With the definition of τ_f and τ_s described in the previous section, $\tau_f \sim 0.26\ ps$ and $\tau_s \sim 180\ ps$. The increase of the relaxation time of the slow dynamics is due to the De Gennes narrowing effect. When $Q = 25.0\ nm^{-1}$, although there are still phonon peaks in the IXS spectrum, the ISF dose not shown a visible oscillation due to the very strong damping. The estimated relaxation times are $\tau_f \sim 0.1\ ps$, and $\tau_s \sim 7.1\ ps$ respectively. At $Q = 25.0\ nm^{-1}$, the IXS spectrum is so broad that the extracted $S(Q, E)$ should be accurate and independent of the model used as long as the fitting is good. Therefore, the calculated ISF should not be affected by the model used to fit the IXS spectra. Thus the two-step decay observed at this Q value in ISF is a genuine feature of the dynamics of the DNA samples.

Fig. (3-19) and Fig. (3-20) show the calculated ISF with the fitted parameters of the IXS spectra of the $Na - DNA - Ca$ sample, and the $Na - DNA - spermidine$ sample at different Q values indicated in the figure. The symbols and meanings are the same as those in Fig. (3-18). The ISF shows the similar relaxation mode observed

Table 3.1: Estimated fast relaxation time, τ_f, of DNA samples. (unit:ps)

$Q(nm^{-1})$	1.6	4.0	9.0	18	18.7	25.0
H_2O	3.6	0.8	0.19	0.26	-	0.1
$CaCl_2$	-	0.26	0.061	-	0.11	0.065
Spermidine	-	0.25	0.052	-	0.18	0.075

Table 3.2: Estimated slow relaxation time, τ_s, of DNA samples. (unit:ps)

$Q(nm^{-1})$	1.6	4.0	9.0	18	18.7	25.0
H_2O	703	61	10	180	-	7.1
$CaCl_2$	-	720	6.4	-	5.8	1.8
Spermidine	-	780	13	-	33	4.9

in Fig. (3-18). The estimated relaxation time is summarized in the table 3.1 and table 3.2.

From these two tables, the relaxation time in general shows very strong Q dependence. With increasing Q, τ_f and τ_s first decrease. They then seem to increase at $Q \sim 18.7\ nm^{-1}$ and then decrease drastically again. The relaxation time of the fast dynamics of the Na-DNA sample in H_2O is larger than that of the samples with extra added counterions. The typical value of τ_f is about $0.1 - 4\ ps$. The typical value of τ_s ranges from 2 to 800 ps. To further improve our estimations, a better instrument resolution would be necessary. It is noticed that the ISF at $Q = 25.0\ nm^{-1}$ decays very fast so that it is within the window that modern computer simulation can calculate.

The nature of the DNA dynamics is still not clearly understood at present. The relaxation time of the internal dynamics of nucleic acids has been measured by different techniques. The measured relaxation time ranges from $\sim$ 10 to 4000 ps[21, 47, 48, 49]. The relaxation time of water molecules around biological macromolecules has also been measured and estimated with the computer simulation. The estimated relaxation time ranges from subpicosecond to hundreds of picosecond [21, 50, 51]. Since DNA molecules have different modes, such as base oscillations, collective uniform bending and twisting modes, multiple base opening breathing[48], it is very difficult

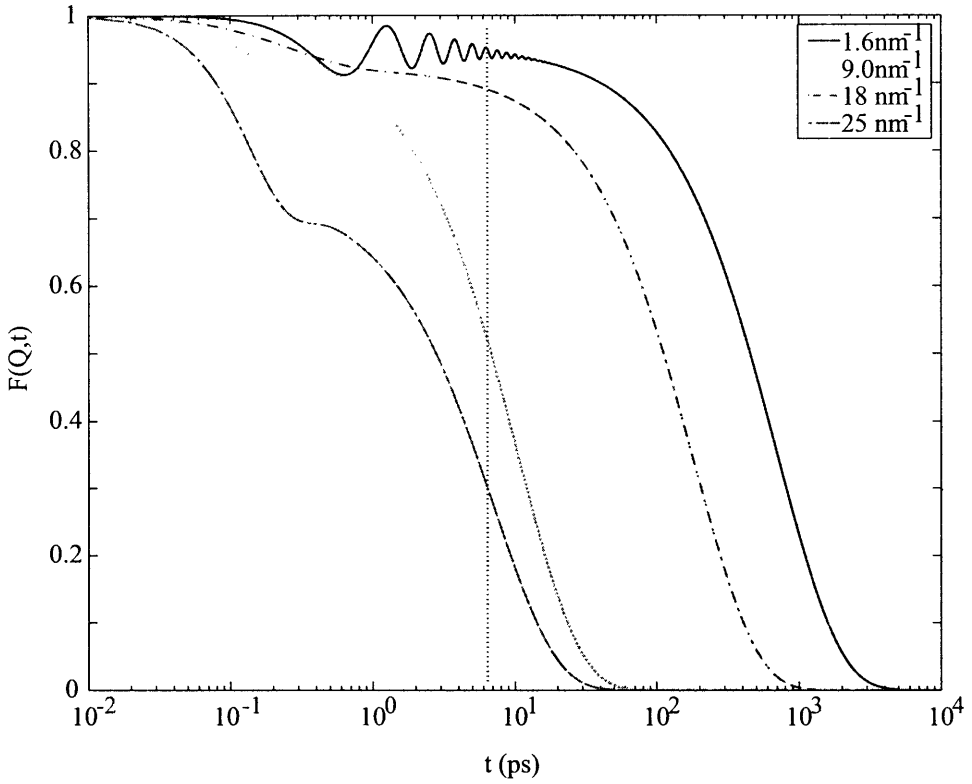

Figure 3-18: The intermediate scattering functions, $F(Q,t)$, are calculated with the fitted parameters of the IXS spectra of 40 $wt\%$ calf-thymus Na-DNA in H_2O at $Q = 1.6\ nm^{-1}$ (solid line), $Q = 9.0\ nm^{-1}$ (dotted line), $Q = 18.0\ nm^{-1}$ (dash dotted line), and $Q = 25.0\ nm^{-1}$ (long dash dotted line). The vertical dotted line drawn at $t \sim 6.5ps$ shows the time window calculated through the energy resolution of the spectrometer (see text). Below this time line, the calculated $F(Q,t)$ is more reliable.

to determine which dynamical modes contribute most to the relaxation at different Q values. The molecular dynamics simulation (MD) may provide more details by comparing the MD results with the experimental results.

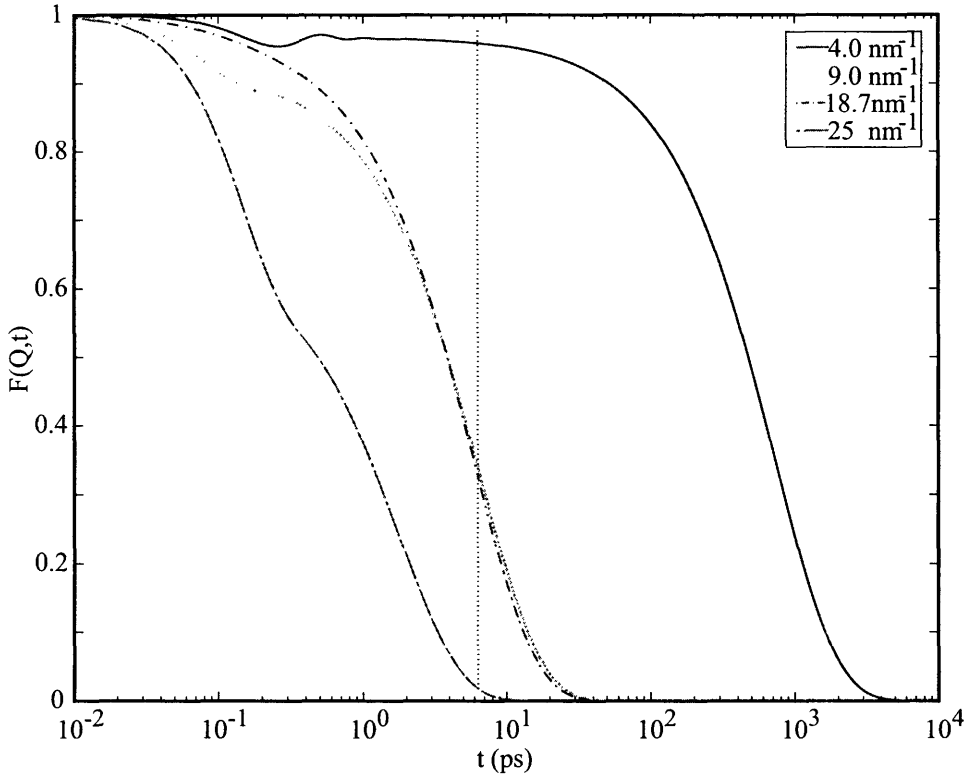

Figure 3-19: The intermediate scattering functions, $F(Q,t)$, are calculated with the fitted parameters of the IXS spectra of 40 $wt\%$ calf-thymus Na-DNA in 0.083 M $CaCl_2$ at $Q = 4\ nm^{-1}$ (solid line), $Q = 9.0\ nm^{-1}$ (dotted line), $Q = 18.0\ nm^{-1}$ (dash dotted line), and $Q = 25.0\ nm^{-1}$ (long dash dotted line). The vertical dotted line drawn at $t \sim 6.5ps$ shows the time window calculated through the energy resolution of the spectrometer. Below this time line, the result is more reliable.

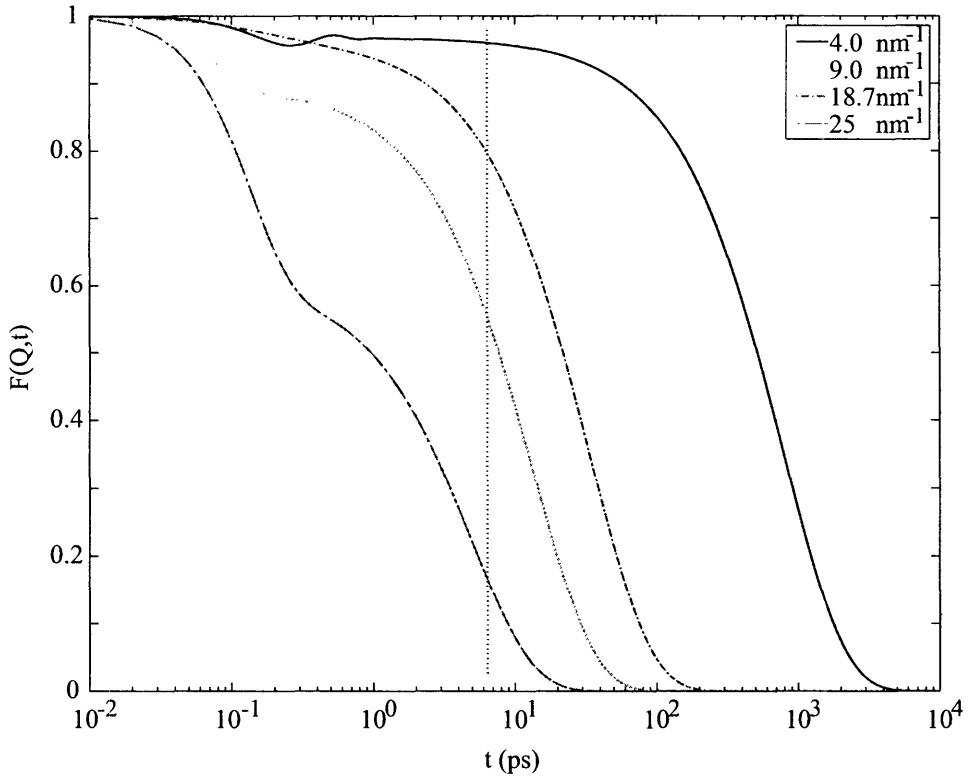

Figure 3-20: The intermediate scattering functions, $F(Q,t)$, are calculated with the fitted parameters of the IXS spectra of 40 $wt\%$ calf-thymus Na-DNA in 0.042 M *Spermidine* at $Q = 4\ nm^{-1}$ (solid line), $Q = 9.0\ nm^{-1}$ (dotted line), $Q = 18.0\ nm^{-1}$ (dash dotted line), and $Q = 25.0\ nm^{-1}$ (long dash dotted line). The vertical dotted line drawn at $t \sim 6.5ps$ shows the time window calculated through the energy resolution of the spectrometer. Below this time line, the result is more reliable.

3.4 Conclusions

We have successfully applied the generalized three effective eigenmode theory (GTEE) to analyze the IXS spectra of aligned $40wt\%$ DNA samples with different counterions. We have constructed the phonon dispersion relation along the DNA axial direction from the calculated eigenmodes. The sound speed obtained from the linear portion of the dispersion relations is about 3100 m/s, which is approximately twice the sound speed of a DNA sample with the similar hydration level observed in the hydrodynamic limit by Brillouin light scattering. Since it is known that the high frequency sound speed of water in this Q-range is also about $3000m/s$, we attribute this difference of the sound speed of aligned DNA rods to the strong coupling of DNA dynamics to that of the hydration water dynamics in the same Q-range. It is observed that the addition of different counterions changes the sound speed at this Q-range only slightly. However, the effect to the phonon damping is very strong. Phonons in the range between $\sim$ 12.5 nm^{-1} and $\sim$ 22.5 nm^{-1} are overdamped by the added counterions according to our model. The phonon damping becomes stronger with the addition of extra counterions. At the same ionic strength, the phonon damping due to the trivalent counterion, *spermidine* , is stronger than that due to divalent counterions. The intermediate scattering function calculated from the fitted parameters of IXS spectra is also discussed. A two-step relaxation of the ISF is observed with the fast relaxation time in the range of $0.1 - 4$ ps and the slow relaxation time in the range of $2 - 800$ ps.

Chapter 4

An Efficient Method to Obtain the Analytical Structure Factor of Two-Yukawa Fluids with the Mean Spherical Approximation

Beginning from this chapter, we will focus on the understanding of the inter-particle structure factor, $S_{inter}(Q)$ defined in Eq. (1.18) for a colloidal system. From now on, when we mention the structure factor, we always mean the inter-particle structure factor. Therefore we will drop the subscript and use $S(Q)$ to represent the inter-particle structure factor.

In a simple fluid, $S(Q)$ is determined by the effective inter-particle potential. Therefore, through the understanding of the inter-particle structure factor, $S(Q)$, the effective inter-particle potential can be extracted. This information is essential to understand the structure and the phase behavior of a colloidal solution, such as micellar solutions and protein solutions.

In this chapter and the chapters those follow, we will focus on the understanding of the structure factor of a simple fluid interacting with a two term Yukawa potential, which will be defined later in this chapter. The motivation to choose this special form

of potential is two fold. First, the analytical structure factor of a system interacting with such potential is available. Second, it can simulate some features of the effective inter-particle potential in a real system, such as a charged colloidal particle solutions.

4.1 Introduction to the Effective Inter-particle Potential in Charge Colloidal Systems

A charged colloidal particle system interacting with a short-range attraction and a long-range electrostatic repulsion has been successfully described by the Derjaguin-Landau-Verwey-Overbeek (DLVO) theory[52]. The stability of such kind of systems has been extensively studied both by perturbation theories and by numerically solving integral equations [53, 54, 55, 56, 57]. A DLVO potential consists of a van der Waals attraction and an electrostatic repulsion. At very low ionic strength, the short-range attraction could be masked by the electrostatic repulsion. For a large particle, for example, the diameter of some particles studied in reference [54] is 350 nm, the electrostatic repulsion could be severely screened by adding salts into solutions. Therefore, besides the *primary minimum* at the very short distance due to the van der Waals attraction, there could be a *secondary minimum* at larger distances. However, within the framework of the DLVO theory, this *secondary minimum* will not happen for a small particle like a Cytochrome C protein or lysozyme protein molecule, whose diameter is about 33 $\mathring{A}$, since the attractive range of the van der Waals potential is a few angstroms for such kind of particles[58, 59, 60]. Therefore, for charge systems, such as protein systems or suspensions with the diameters about a few nanometers, the inter-particle potential according to the DLVO theory is typically made of a short-range attraction and a long-range repulsion. At small volume fractions, the small angle neutron scattering (SANS) and the small angle X-ray scattering (SAXS) spectra can be successfully explained by only considering the screened Coulombic repulsion. At larger volume fractions up to about 20%, Chen and coworkers developed the Generalized One Component Macroion (GOCM)[61, 62] theory to modify the expression

of the screened Columbic repulsion and have successfully explained the SANS and SAXS spectra of Cytochrome C protein solutions at very low ionic strength.

Recent SANS experimental results on protein solutions demonstrate some new interesting results[63]. First, an additional peak, besides the normal diffraction peak, is observed. It appears at a finite wave vector, Q, smaller than the Q value of the first diffraction peak. This peak is called the cluster peak, which signifies the formation of well ordered clusters in the liquid. Secondly, there is always a building up of intensity for Q centered around zero (zero-Q peak) in both liquid-like samples and solid-like samples. These new observations from SANS data indicate that the inter-protein potential is rather complicated, and SANS data could not be explained by a screened Coulombic potential alone[63]. Tardieu and coworkers studied by SAXS the inter-particle potential of lysozyme protein solutions at pH=4.5 with different ionic strengths. They found that at low ionic strength, $I \approx 0.1M$, an electrostatic repulsion plus a short-range (a few $\mathring{A}$), possibly van der Waals, attraction can explain the SAXS data. However, at very large ionic strength, the short-range attraction component also depends on the type of anions added, which can not be explained by the DLVO theory alone[58, 59]. The origin of the short-range attraction dependence on the anions added remains unclear.

Sciortino and coworkers simulated a model system with a short-range Lennard-Jones $2n$ - n potential with a very large n number, plus a long-range Yukawa form of the repulsive potential. Due to the competition of the short-range attraction and the long-range repulsion, the simulation results clearly show that when there is a cluster peak in the structure factor, there are well ordered clusters in the solutions[64]. Therefore they proposed that the gel phase could be stabilized by a weak long-range Coulombic repulsive interaction instead of a short-range attraction. The appearance of the cluster peak in the structure factor generated by the simulations implies that the cluster peak observed in protein SANS data is likely due to the inter-protein potential with a short-range attraction and a long-range repulsion.

There are also experimental observations of macroscopic patterns in both physical and chemical systems[65, 66], which exhibit stripes and circular droplets in two-

dimensional systems and sheets, tubes, and spherical droplets in three-dimensional systems. Sear and coworkers attribute the spontaneous formations of spatial patterns of quantum dots at the air-water interface to the competition of a short-range attraction and a long-range repulsion[66]. A fluid system interacting with a two-Kac-form potential has been studied theoretically with the random phase approximation (RPA) method[67]. All these experiments, simulations, and theoretical works render the systems interacting with a short-range attraction plus a long-range repulsion very interesting.

4.2 An Efficient Method to Calculate the Structure Factor

During their analysis of the experimental data, Tardieu and coworkers have used numerical solutions of the Ornstein-Zernike (OZ) equation[58, 59]. There is still no suitable way to obtain an analytical structure factor to compare with the small angle scattering results to systematically study the properties of fluids with a short-range attraction plus a long-range repulsion. In this chapter, we exploit the known analytical solution for two Yukawa potentials by solving the OZ equation with mean spherical approximation (MSA) given by Waisman, Stell, and Blum[68, 69, 70]. Compared with the Kac form potential, the Yukawa form potential in charge colloidal systems is more appropriate. Moreover comparing to the RPA method, the MSA should be more accurate in the whole Q range of a typical scattering spectra. The screened Coulombic repulsion between charged macroions surrounded by counterions can be expressed as a Yukawa potential, which can be calculated by DLVO[52] and GOCM theories. But the origin of the short-range attraction in inter-protein interaction is still vague. Lennard-Jones potential, Van der Waals potential, and Yukawa potential forms are used in different calculations[58, 59, 64]. Malfois and coworkers show that the van der Waals potential could also be simulated by a short-range attractive Yukawa potential in protein systems[60]. Therefore, we believe that the fluids with two Yukawa

potentials should be able to simulate real systems without losing generality.

In this chapter, we will use this new method to study the properties of two-Yukawa fluids, which further extends the very limited cases discussed by Konior[73]. Although attentions will be focused on discussions of the cluster peak in the structure factor due to the competition of a short-range attraction and a long-range repulsion, we will not restrict ourselves to it. Instead we will also discuss some results obtained for the systems with a relatively short-range repulsion and a long-range attraction. The application of this latter case is immediately found in fitting the SANS experimental data of protein systems in chapter 6.

Although the analytical form of the structure factor of two-Yukawa fluids has been given before, obtaining the coefficients in the analytical form from a given set of control parameters is highly non-trivial, involving works to solve nonlinear algebraic equations. With MSA closure, the analytical form of one Yukawa potential was first given by Waisman[71]. Then Høye, Stell, and Waisman obtained the result of two Yukawa potential[69]. Later, Høye and Blum generalized the analytical form to the case of a multi-Yukawa potential[70]. To obtain the correct coefficients of the analytical form, there are two steps: 1) solving a group of nonlinear equations, the complexity depends on the number of terms in the multi-Yukawa potential; 2) setting criteria to find the right root from multiple roots of the equations. Hayter and Penfold first obtained the coefficients of the analytical form of single term Yukawa potential and calculated the corresponding structure factor for various control parameters[72]. The Hayter-Penfold's method is now widely used to analyze the SANS and SAXS data of polyelectrolyte solutions. To obtain the coefficients for a two-term Yukawa potential is much more difficult. This was first done by Konior and Jedrzejek[73]. Later Arrieta and coworkers worked out the coefficients for a multi-component, multi-Yukawa fluid[74]. However, the approaches of both Konior and Arrieta can not be implemented automatically in computer codes and therefore it will be very difficult to use those methods to fit the experimental data, which usually need tens or hundreds of iterations. Hence, we provide a new and efficient way to obtain the coefficients for a two-term Yukawa potential, which can be implemented into a code to search for the

right solution automatically. This method provides a practical way to fit experimental data with a two-term Yukawa potential model like the Hayter-Penfold's method for a one-term Yukawa potential.

An analytical form of a structure factor is usually obtained by solving the Ornstein-Zernike (OZ) equation,

$$h(r) = c(r) + \rho \int c(|\vec{\mathbf{r}} - \vec{\mathbf{r}}'|)h(r')d\vec{\mathbf{r}}' \tag{4.1}$$

with a suitable closure relation. In Eq. (4.1), ρ is the number density of particles, $c(r)$ the direct correlation function, $h(r) = g(r) - 1$ and $g(r)$ is the pair correlation function.

The OZ equation with a multiple-Yukawa potential has been solved with MSA closure by Blum [70]. Here we will only present equations necessary for solving the case of a two-term Yukawa potential by following Blum's method.

The reduced potential is given as,

$$\frac{V(r)}{k_B T} = \begin{cases} \infty, & (0 < r < 1) \\ -K_1 \frac{e^{-Z_1(r-1)}}{r} - K_2 \frac{e^{-Z_2(r-1)}}{r}, & (r > 1) \end{cases} \tag{4.2}$$

where k_B is the Boltzmann constant, T the absolute temperature, and the diameter of the hardcore is taken to be 1. With the MSA closure,

$$\begin{cases} h(r) = -1, & (0 < r < 1) \\ c(r) = -\frac{V(r)}{k_B T} = K_1 \frac{e^{-Z_1(r-1)}}{r} + K_2 \frac{e^{-Z_2(r-1)}}{r}, & (r > 1), \end{cases} \tag{4.3}$$

Blum solved OZ equation analytically with Baxter's Q-method[75] and obtained the analytical form of $c(r)$ inside the hardcore,

$$-rc(r) = a_0 r + b_0 r^2 + \frac{1}{2}\phi a_0 r^4 +$$

$$\sum_{i=1}^{2} \frac{v_i}{z_i}(1 - e^{-Z_i r}) +$$

$$\sum_{i=1}^{2} \frac{v_i^2}{2K_i e^{Z_i} Z_i^2}(\cosh(Z_i r) - 1), \ (r < 1). \tag{4.4}$$

where ϕ is the volume fraction, $v_i = 24\phi K_i e^{Z_i} \hat{g}(Z_i)$, $a_0 = a^2$, and $b_0 = -12\phi(\frac{1}{2}(a + b)^2 + a\sum_{i=1}^{2} c_i e^{-Z_i})$. And $\hat{g}(s)$ is given as,

$$\hat{g}(s) = \frac{[\frac{1}{s^2}(a(s+1) + bs) - \sum_{i=1}^{2} \frac{Z_i c_i}{Z_i + s} e^{-Z_i}]e^{-s}}{1 - 12\phi q(s)} \tag{4.5}$$

where

$$q(s) = \sigma(s) - \tau(s)e^{-s}$$

$$\sigma(s) = \frac{a}{s^3} + \frac{b}{s^2} - (\frac{a}{2} + b + \sum_{i=1}^{2} c_i e^{-Z_i})\frac{1}{s} + \sum_{i=1}^{2}(c_i + d_i)\frac{1}{Z_i + s} \tag{4.6}$$

$$\tau(s) = a(\frac{1}{s^3} + \frac{1}{s^2}) + \frac{b}{s^2} - \frac{1}{s}\sum_{i=1}^{2} \frac{Z_i}{Z_i + s} c_i e^{-Z_i}. \tag{4.7}$$

It is noted that coefficients in the direct correlation function, Eq. (4.4), is a function of a set of six parameters, $\{a,\ b,\ c_1,\ c_2,\ d_1,\ d2\}$. For a given set of control parameters $\{\phi, Z_i, K_i\}(i = 1, 2)$, $\{a, b, c_i, d_i\}(i = 1, 2)$ can be obtained by solving the following six algebraic equations,

$$b = 12\phi\{-\frac{1}{8}a - \frac{1}{6}b + \sum_{i=1}^{2} c_i[\frac{1}{Z_i^2} - (\frac{1 + Z_i}{Z_i^2} + \frac{1}{2})e^{-Z_i}] + \sum_{i=1}^{2} d_i \frac{1}{Z_i^2}\} \tag{4.8}$$

$$1 - a = 12\phi[-\frac{1}{3}a - \frac{1}{2}b + \sum_{i=1}^{2} c_i(\frac{1}{Z_i} - \frac{1 + Z_i}{Z_i}e^{-Z_i}) + \sum_{i=1}^{2} d_i \frac{1}{Z_i}] \tag{4.9}$$

$$K_i e^{Z_i} = Z_i d_i[1 - 12\phi q(Z_i)], \ (i = 1, 2) \tag{4.10}$$

$$c_i + d_i = 12\phi[\sigma(Z_i)(c_i + d_i) - \tau(Z_i)c_i e^{-Z_i}],\ (i = 1, 2). \tag{4.11}$$

Once $\{a, b, c_i, d_i\}(i = 1, 2)$ are obtained, $c(Q)$ can be calculated by the Fourier transformation of $c(r)$ analytically from Eq. (4.3) and Eq. (4.4) [69], resulting in

$$\begin{aligned}\rho c(Q) = & -24\phi\{\frac{a[\sin Q - Q\cos Q]}{Q^3} + \frac{b[2Q\sin Q - (q^2-2)\cos Q - 2]}{Q^4} + \\ & \frac{a\phi[(Q^2-6)4Q\sin Q - (Q^4 - 12Q^2 + 24)\cos Q + 24]}{2Q^6} + \\ & \sum_{i=1}^{2} h_Q(K_i, Z_i, v_i)\}\end{aligned} \tag{4.12}$$

where

$$\begin{aligned}h_Q(K, Z, v) = & \frac{v}{Z}(1 - \frac{v}{2KZe^Z})(\frac{1-\cos Q}{Q^2} - \frac{1}{Z^2+Q^2}) - \frac{v^2(Q\cos Q - Z\sin Q)}{4KZ^2Q(Z^2+Q^2)} + \\ & \frac{Q\cos Q + Z\sin Q}{Q(Z^2+Q^2)}(\frac{v}{Ze^Z} - \frac{v^2}{4KZ^2e^{2Z}} - K).\end{aligned} \tag{4.13}$$

Therefore, the structure factor, $S(Q)$, can be calculated as $S(Q) = \frac{1}{1-\rho c(Q)}$.

Obtaining $\{a, b, c_i, d_i\}(i = 1, 2)$ from Eq. (4.8)-Eq. (4.11) efficiently is difficult. We will present the method to solve these equations in the following. During the calculations, $\{K_i, Z_i, \phi\}(i = 1, 2)$ are considered known parameters.

Eq. (4.8)-Eq. (4.10) are linear in a, b, c_1, and c_2. Therefore a, b, c_1 and c_2 can be expressed in terms of d_1 and d_2 together with $\{K_i, Z_i, \phi\}$. Substituting the expressions of a, b, c_1 and c_2 into Eq. (4.11), two coupled equations with only d_1 and d_2 are obtained,

$$\begin{aligned}& A_{41}d_1^4 d_2 + A_{32}d_1^3 d_2^2 + A_{31}d_1^3 d_2 + A_{30}d_1^3 + A_{21}d_1^2 d_2 + \\ & A_{12}d_1 d_2^2 + A_{11}d_1 d_2 + A_{10}d_1 + A_{01}d_2 = 0\end{aligned} \tag{4.14}$$

$$B_{14}d_1d_2^4 + B_{23}d_1^2d_2^3 + B_{13}d_1d_2^3 + B_{03}d_2^3 + B_{12}d_1d_2^2 +$$
$$B_{21}d_1^2d_2 + B_{11}d_1d_2 + B_{01}d_2 + B_{10}d_1 = 0 \tag{4.15}$$

which have the same form as Konior[73] obtained. However, β_i in Ref. [73] is equal to $d_i e^{-Z_i}$ in this paper. A_{ij} and B_{ij} are functions of $\{K_i,\ Z_i,\ \phi\}$ $(i = 1, 2)$. Till this step, our method is same as Konior's. Since Eq. (4.15) is a quadratic equation in d_1, Konior expressed d_1 in terms of d_2 and substituted the expression of d_1 back into Eq. (4.14) to obtain a complicated nonlinear equation with only one variable d_2. Instead of directly solving quadratic function of d_1 in Eq. (4.15), we take a different approach, which will finally result in a polynomial equation with only one variable, d_2.

From Eq. (4.15),

$$d_1^2 = -(B_{14}d_1d_2^4 + B_{13}d_1d_2^3 + B_{03}d_2^3 + B_{12}d_1d_2^2 +$$
$$B_{11}d_1d_2 + B_{01}d_2 + B_{10}d_1)/(B_{23}d_2^3 + B_{21}d_2). \tag{4.16}$$

The order of d_1 is one on r.h.s. of Eq. (4.16). Substituting the expression of d_1^2 into Eq. (4.14) for each item, whose order of d_1 is larger than 1. Each substitution will decrease the order of d_1 at least one for the corresponding items. Sort the resultant equation by the order of d_1. If the resultant equation has some items whose order of d_1 is larger than 1, repeat the previous substitution procedure until Eq. (4.14) is reduced to an equation with the highest order of d_1 to be one. Solving this new equation, d_1 could be expressed in terms of d_2 only. Then the expression of d_1 in terms of d_2 can be substituted back into Eq. (4.15) and an equation with only one variable d_2 can be obtained. By letting the numerator of this equation equal to zero, we can obtain a polynomial equation, which has only one variable, d_2, with 27th order,

$$\sum_{i=1}^{27} W_i d_2^i = 0. \tag{4.17}$$

where $\{W_i\}$ $(i = 1, .., 27)$ are coefficients and can be calculated from $\{K_i, Z_i, \phi\}$ $(i = 1, 2)$. For a polynomial equation, such as Eq.(4.17), the complete set of roots can be obtained very fast by numerical method with high accuracy. Once d_2 is available, d_1 can be calculated by solving the quadratic equation of Eq. (4.15). Therefore, $\{a, b, c_i\}$ $(i = 1, 2)$ can be obtained by solving linear equations, Eq. (4.8)-Eq. (4.10). The complete computer codes have been written in Matlab language and could be obtained from the authors[76].

Comparing to the previous methods[73, 74], the advantages of our method are: (1) The complete set of roots for Eq. (4.8)-Eq. (4.11) are obtained. Therefore the right root is guaranteed to be found. By setting proper criteria of root selection in the limit number of roots, our method can ensure the correctness of the results. (2) Our method is more efficient. The search for the correct root can be automatically implemented into codes since the total number of roots is limited. As comparison, the previous methods always need human interventions. The key step in our method is that we reduce the multiple nonlinear equations into a single polynomial equation, Eq.(4.17), which can be solved in a very fast way by the existing known algorithms. In practice, we use *Matlab* language to write the code. There is a function, *roots*(), which can solve roots for polynomial equations. After incorporating the root selection rules, obtaining a structure factor for a given set of $\{K_i, Z_i, \phi\}$ just needs about 10 seconds on a normal desktop computer. Comparing to the method of solving the OZ equation numerically, our method is much faster since a numerical method usually takes a few minutes to solve the OZ equation[54].

Before discussing the root selection rules, we like to point out that obtaining Eq.(4.17) from Eq. (4.14) and Eq. (4.15) is a laborious work, which needs the aid of symbolic calculation softwares, such as *Mathematica*. The expression of the coefficients in Eq.(4.17) can be longer than one page space. Therefore, we have to write a code to directly convert the expressions in *Mathematica* into the expressions *Matlab* can use.

Selecting the correct root is equally important as obtaining the complete set of roots. Pastore has proved that under arbitrary number of potential and number

of components, if a solution of MSA for the OZ equation exists, it is unique and physical[77]. This seems to contradict with the fact that we have generally multiple roots by following Blum's method. Pastore pointed out that this is due to the fact that Baxter's Q-method is not completely equivalent with the OZ equation. Some restrictions need to be added. In our case of two-Yukawa one-component liquids, in order to make Baxter's Q method and the OZ equation equivalent, the Pastore's criterion can be stated as the function

$$\Delta(s) = det(1 - 2\pi\rho\tilde{Q}(s)) \tag{4.18}$$

should have no zero points in the right half of the complex plane, where $\tilde{Q}(s)$ is the Laplace transform of $Q(r)$ and $Q(r)$ is the Q function in Baxter's Q-method. Due to its time cost, this criterion is not widely used. Some alternative methods or simplified criteria were proposed. Before Pastore proposed his criterion, Hayter and Penfold proposed a rule in their method for one Yukawa case. They calculate the pair distribution function, $g(r)$, for each root. The one that gives the correct value of $g(r)$ inside hardcore is the correct root. In Konior and Arrieta's methods, since the complete set of roots is unavailable, they can not use this rule. A simplified version for multi-Yukawa multi-component case for Pastore's criterion is proposed by Arrieta. The version of this simplified criterion in our case is that the correct root should give

$$\Delta(0) = det(1 - 2\pi\rho\tilde{Q}(0)) > 0 \tag{4.19}$$

which is equivalent to require $a > 0$. Although this is just a necessary condition, not sufficient condition, and does not guarantee the equivalence, they found that their results were all correct in their calculations.

In our case, since the complete set of roots is available, we therefore consider both Hayter-Penfold's rule and Arretia's rule to propose our selection rules as following for each set of roots, (1) d_2 and d_1 should be real number; (2) the correct root should have $a > 0$; (3)the correct root should have $v_i/K_i > 0$, which can be equivalently implemented as $\hat{g}(Z_i) > 0$; (4) after the selections of first three conditions, if there

are still more than one root left, calculate $g(r)$ for each root and $g(r)$ of the correct root should have the minimum average value inside the hardcore. The third condition introduced is that the Laplace transform of the pair distribution function $g(r)$ should always be positive.

To check our methods, we first compared the results of our method with the results obtained by Konior[73]. Under the same set of control parameters, $\{K_i, Z_i, \phi\}$ $(i = 1, 2)$, all the calculations give the same results. Secondly, we compared our result with Hayter-Penfold's method for single Yukawa case by setting one of the two Yukawa potential to be a very small value. The result of our method successfully reduce to the result of the Hayter-Penfold's method. These indicate the correctness of our method.

4.3 Analysis of the Structure Factor of Two Yukawa Fluids

In this section, the method developed in the previous section is applied to study the structure factor as a function of different control parameters. We will first discuss the effect on the structure factor due to a short-range attraction plus a long-range repulsion. In particular, the dependence of cluster peaks on the different control parameters are discussed in details. Then a brief discussion of the effect of a relatively short-range repulsion and a long-range attraction will be presented. To be consistent, we will always let the first Yukawa potential to be attractive, i.e., $K_1 > 0$, and let the second Yukawa potential to be repulsive, i.e., $K_2 < 0$. Z_1 is usually set around 10 due to the fact that the range of the short-range attraction between proteins in solutions is about 10% compared with the hardcore diameter[58, 59].

In the analysis of the cluster peak, spinodal lines are always calculated by checking the structure factor intensity at Q=0 to ensure that the cases presented are not macroscopically phase separated. However, due to the lack of dynamical information in a structure factor, this dose not guarantee that systems are always in ergodic

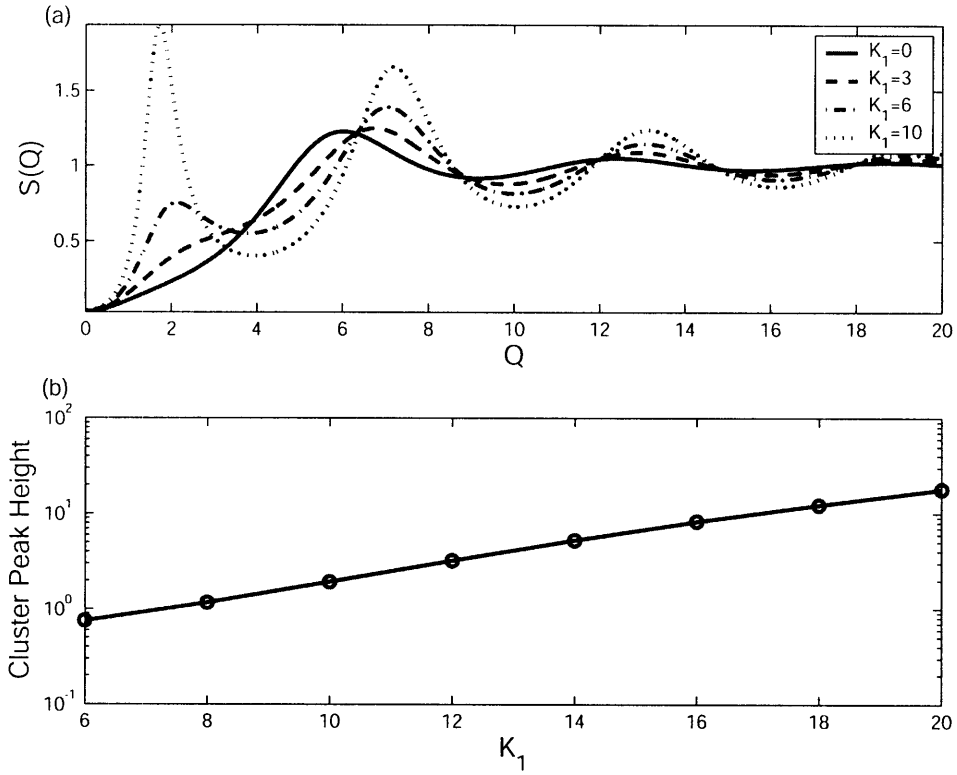

Figure 4-1: This figure illustrates the existence of a novel cluster peak in the structure factor. In panel (a), $\phi = 0.20$, $Z_1 = 10$, $K_1 = -1$, $Z_2 = 0.5$. For $K_1 = 0$ (solid line) and $K_1 = 3$ (dashed line), there is no cluster peak. For $K_1 = 6$ (dash dotted line) and $K_1 = 10$ (dotted line), the attraction is strong enough so that the cluster peak appears. Panel (b) shows the change of cluster peak intensity, $I_{cluster}$, as a function of K_1.

states. The kinetic phase diagrams for the idealized liquid-glass transition can be calculated by the Mode Coupling Theory (MCT)[78, 79, 80] with the structure factor as the input. Such calculations have been done for colloidal systems interacting with a short-range attraction[81, 82]. The kinetic phase diagram predictions of such systems by Dawson et al. have been verified by experiments[83]. The calculations of the kinetic phase diagrams of a fluid system interacting with a two-Yukawa potential are also very interesting and is addressed in chapter 5. In the following calculations, we will ignore the kinetic phase diagrams and focus on how the different control parameters affect the cluster peak.

Figure 4-1 studies the structure factor for a short-range attraction plus a long-

range repulsion with the control parameters, $Z_1 = 10,\ K_2 = -1,\ Z_2 = 0.5,\ \phi = 0.20$ at different values of the attraction depth, K_1. It clearly demonstrates the appearance of the cluster peak due to the competition of the short-range attraction and the long-range repulsion. In panel (a), the solid line, the dashed line, the dash dotted line, and the dotted line correspond to $K_1 = 0,\ 3,\ 6$, and 10 respectively. In the absence of the attraction as shown by the solid line, there is no cluster peak. When the attraction increases, the structure factor first shows an inflection point without a local maximum. When the attraction further increases, the cluster peak begin to appear and the height of the cluster peak increases and the position is shifted to lower Q values. These indicate that in order to observe a cluster peak, the short-range attraction must be strong enough. Once clusters begin to form in the fluids, with the increase of the attractive Yukawa potential, the number of clusters inside the system increases and the size of the cluster grows. Panel (b) shows the change of cluster peak height, $I_{cluster}$, as a function of the attraction depth, K_1. It is interesting to note that there is a linear relation between $log(I_{cluster})$ and K_1 for this set of control parameters. The reasons and implications of this interesting phenomenon need to be studied in future experiments and theories.

Figure 4-2 studies the effect of the attraction range, $\frac{1}{Z_1}$, on the cluster peak height, $I_{cluster}$ with the control parameters, $\{K_1 = 6,\ K_2 = -1,\ Z_2 = 0.5,\ \phi = 0.20\}$ at different values of Z_1. In panel (a), the solid line, the dash dotted line, and the dotted line correspond to $Z_1 = 4,\ 8$, and 14 respectively. With the decrease of Z_1 (the increase of the attraction range), the cluster peak height increase very sensitively and the peak position is also shifted to the lower Q values. Panel (b) indicates the change of the cluster peak, $I_{cluster}$ as a function of the attraction range, $\frac{1}{Z_1}$. When increasing the attraction range, the cluster peak height monotonically increases.

Figure 4-3 studies the effect of the repulsion strength, K_2, on the structure factor with the control parameters, $\{K_1 = 6.9,\ Z_1 = 10,\ Z_2 = 0.5,\ \phi = 0.20\}$ at different values of K_2. In panel (a), the solid line, the dash dotted line, and the dotted line correspond to $K_2 = -1,\ -0.1$, and -0.01 respectively. Without the repulsion, the spinodal line for a pure attractive potential with $Z_1 = 10$ at $\phi = 0.2$ is at $K_1 = 6.96$.

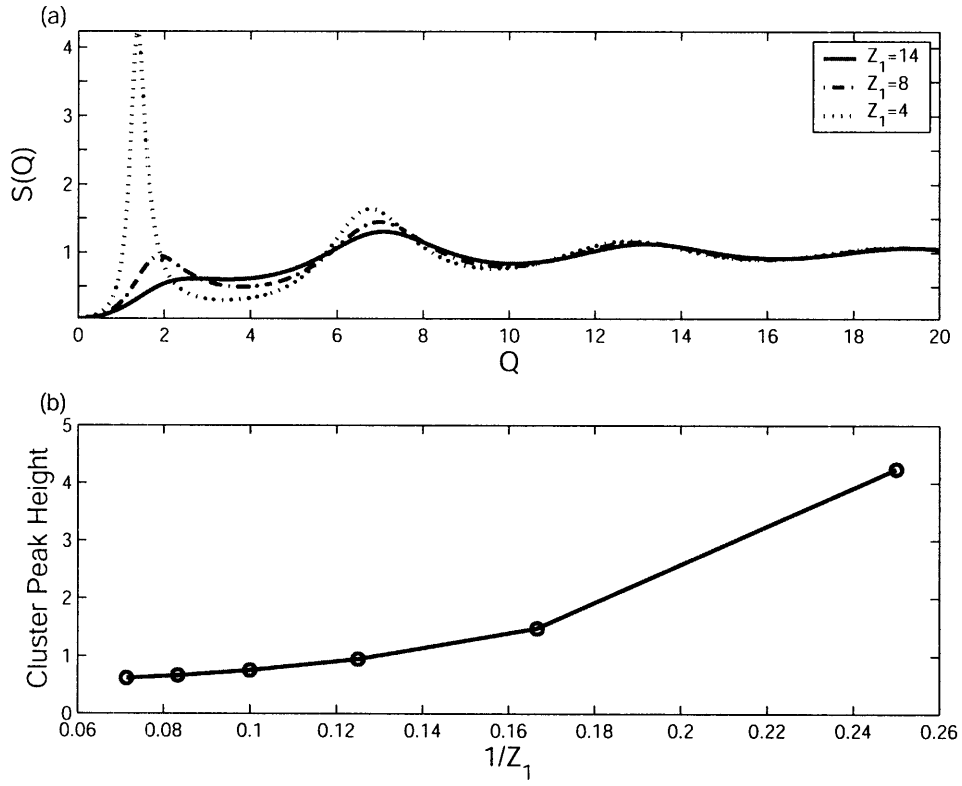

Figure 4-2: In panel (a), $\phi = 0.20$, $K_1 = 6$, $K_2 = -1$, $Z_2 = 0.5$. The solid line, the dash dotted line, and the dotted line correspond to $Z_1 = 14$, 8, and 4 respectively. Panel (b) shows the change of cluster peak intensity, $I_{cluster}$, as a function of attraction range, $\frac{1}{Z_1}$.

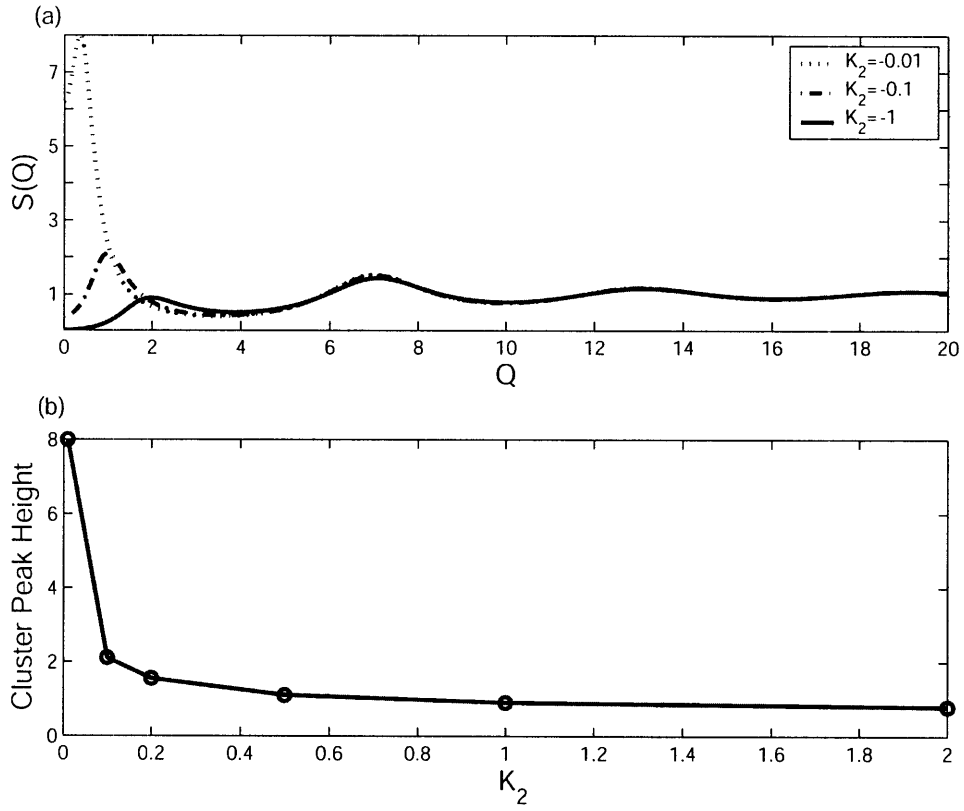

Figure 4-3: In panel (a), $\phi = 0.20$, $K_1 = 6.9$, $Z_1 = 10$, $Z_2 = 0.5$. The solid line, the dash dotted line, and the dotted line correspond to $K_2 = -1$, -0.1, and -0.01 respectively. Panel (b) shows the change of cluster peak intensity, $I_{cluster}$, as a function of the amplitude of repulsive potential, $|K_2|$.

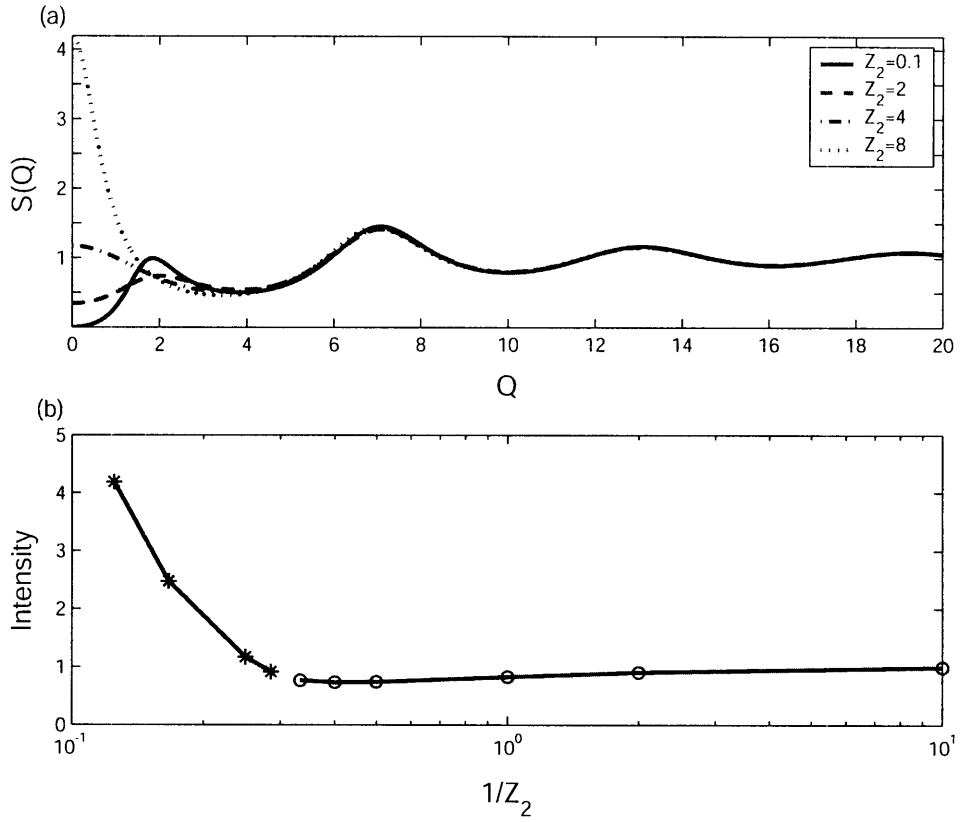

Figure 4-4: Panel (a) demonstrates the effect of repulsion range, $\frac{1}{Z_2}$, on the structure factor, $S(Q)$, at a volume fraction, $\phi = 0.20$. $K_1 = 6.9$, $Z_1 = 10$, $K_2 = -1$. The solid line, the dashed line, the dash dotted line, and the dotted line correspond to $Z_2 = 0.1$, 2, 4, and 8 respectively. Open circles in panel (b) shows the change of cluster peak intensity, $I_{cluster}$, as a function of repulsion range, $\frac{1}{Z_2}$. The $*$ symbol indicates the intensity of $S(Q = 0)$ when there is no finite-Q cluster peak.

For a pure attraction with $\{K_1 = 6.9,\ Z_1 = 10,\ \phi = 0.20\}$, there is no cluster peak in the structure factor and $S(Q = 0)$ is larger than 1000. However, with the addition of a very weak repulsion, $K_2 = -0.01$, the intensity of $S(Q = 0)$ decreases dramatically. Finite size clusters form in the fluids and a finite cluster peak appears in the structure factor. With the further increase of K_1, the cluster peak height decreases and the peak position also shifts to the higher Q values. In panel (b), the change of the cluster peak intensity as a function of the amplitude of the repulsion strength, $|K_2|$, is shown.

Figure 4-4 shows the effect of the repulsion range, $\frac{1}{Z_2}$, on the structure factor with the control parameters, $\{K_1 = 6.9,\ Z_1 = 10,\ K_2 = -1,\ \phi = 0.20\}$ at different values

of Z_2. In panel (a), the solid line, dashed line, dash dotted line, and the dotted line correspond to $Z_2 = 0.1,\ 2,\ 4$ and 8 respectively. With the increase of Z_2 (the decrease of the repulsion range), the cluster peak decreases and the position of the cluster peak shifts to higher Q values. At the same time, the value of $S(Q = 0)$ increases. When the repulsion range become small enough, the cluster peak disappears and there is no local maximum for the range of Q value less than the wave vector of the first diffraction peak, Q_{diff}. During the change of Z_2, the structure factor, $S(Q)$, almost dose not change for $Q > 4$. The reasons are as following. In this set of control parameters, the short-range attractive potential is very strong ($K_1 = 6.9$). The change of the long-range repulsion hardly affects too much the potential form close to the particle, which determines the first diffraction peak of $S(Q)$ and the values for even larger Q values. Therefore, the change of the long-range repulsive potential will mostly affect the system's long-range order. Therefore $S(Q)$ for Q at small values responds very sensitively to the change of the repulsion range. In panel (b), we show the effect of intensity of cluster peak (open circles) as a function of the repulsion range, $\frac{1}{Z_2}$. When there is no cluster peak, we show the intensity of $S(Q = 0)$ as $*$ symbols. If we define the liquid with a cluster peak in the structure factor as cluster liquid, there is a transition of states from cluster liquid to normal liquid.

Figure 4-5 demonstrates the effect of the volume fraction, ϕ, on the structure factor at the control parameters, $\{K_1 = 10,\ Z_1 = 10,\ K_2 = -1,\ Z_2 = 0.5\}$. In panel (a), the solid line, the dashed line, the dotted line, and the dash dotted line correspond to $\phi = 0.05,\ 0.20,\ 0.40$, and 0.55. When $\phi \leq 0.50$, there is a cluster peak in the structure factor and the liquid system shows stable clusters stabilized by the competition of the short-range attraction and the long-range repulsion. When the volume fraction is too large, such as $\phi = 0.55$, the system again is dominating by the single particle motions. For a fixed inter-particle potential, the increase of volume fraction first increases the cluster peak height, $I_{cluster}$. After reaching a maximum at about $\phi = 0.20$, it begins to decrease with the further increase of the volume fraction. Therefore, for a given set of control parameters, if there is a cluster peak, there in general exists an optimal volume fraction value, at which the intensity of

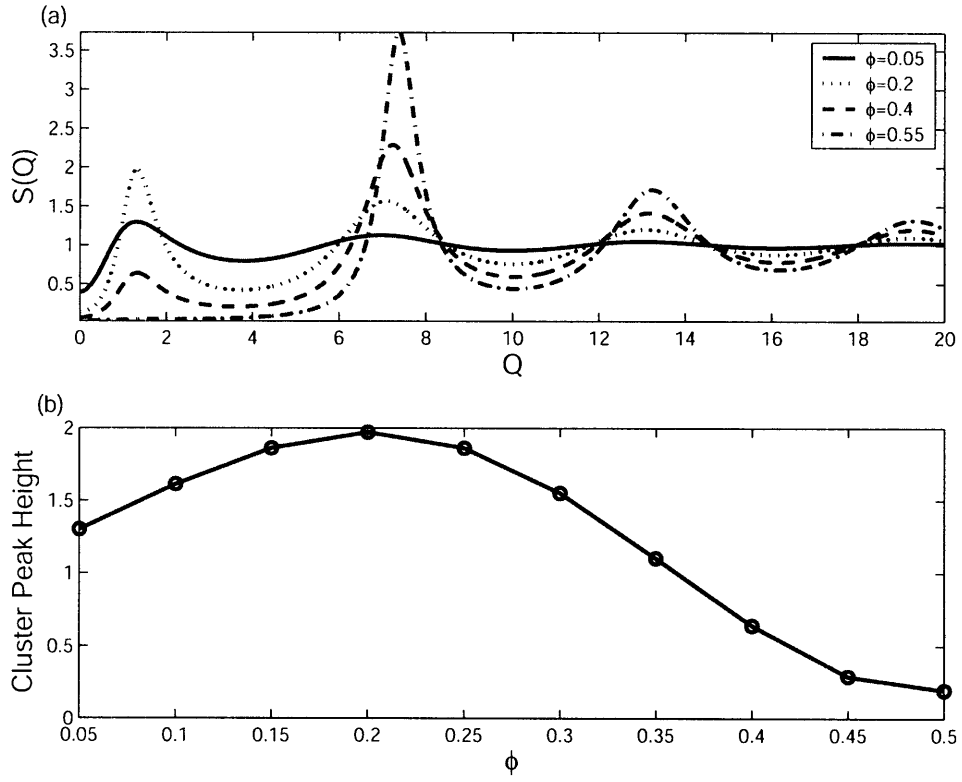

Figure 4-5: Panel (a) shows structure factors at volume fractions, $\phi = 0.05$, 0.2, 0.4, and 0.55. $K_1 = 10$, $Z_1 = 10$, $K_2 = -1$, $Z_2 = 0.5$. Panel (b) shows the change of cluster peak intensity, $I_{cluster}$, up to the volume fraction, $\phi = 0.5$. For a given set of K_i, Z_i, there exists an optimal volume fraction, at which the cluster peak intensity, $I_{cluster}$, is largest. For the case presented in this figure, the optimal volume fraction, ϕ, is about 20%.

the cluster peak is the largest. This is most clearly shown in the panel (b), where the cluster peak intensity, $I_{cluster}$, is shown as a function of the volume fraction, ϕ. Unlike other control parameters, the change of volume fraction dose not change the cluster peak position up to $\phi = 0.40$. This indicates that the inter-cluster distance is not changed with the increase of the volume fraction. Recent results in lysozyme protein solutions[102] show that with the increase of the volume fraction, the cluster size increases while the inter-cluster distance is not changed.

In Figure 4-6, we study the spinodal lines at different control parameters and show them as a function of the effective temperature, $\frac{1}{K_1}$, vs. the volume fraction, ϕ. The solid line in all three panels of Figure 4-6 is the reference spinodal line, which correspond to $\{Z_1 = 10,\ K_2 = -0.3,\ Z_2 = 2\}$. Panel (a) shows the shift of spinodal lines due to the increase of the amplitude of the repulsion, $|K_2|$. The dotted line, the dashed line, and the dash dotted line correspond to $K_2 = -1,\ -2$, and -5 respectively. With the increase of the repulsion strength, the stronger repulsion requires the stronger attraction to have aggregations inside the liquids. Hence the spinodal lines are pushed to lower value of $\frac{1}{K_2}$. In panel (b), the dotted line, the dashed line, the dash dotted line correspond to $Z_2 = 3,\ 4$, and 6 respectively. The increase of Z_2 (the decrease of the repulsion range, $\frac{1}{Z_2}$), weakens the effective repulsion strength. The liquid system needs much less attraction strength to have a phase separation. Therefore, the spinodal lines are shifted to larger value of $\frac{1}{K_1}$. In panel (c), the effect of the attraction range, $\frac{1}{Z_1}$, on the spinodal lines is studied. The dotted line and the dashed line correspond to $Z_1 = 8$, and 6 respectively. At a fixed repulsive potential and the amplitude of the attraction strength, the decrease of Z_1 (the increase of the attraction range $\frac{1}{Z_1}$) increase the effective attraction strength. The spinodal lines are then shifted to larger values of $\frac{1}{K_1}$.

With a long-range repulsive potential and a very weak short-range attraction, the inter-particle potential will be dominated by the repulsion and stable clusters can not form in liquids. On the other hand, if the short-range attraction is too strong, the system may go beyond the formation of finite size clusters and approach the spinodal line. Therefore, there may exist a finite area in the phase plane of $\frac{1}{K_1}$ vs. ϕ that

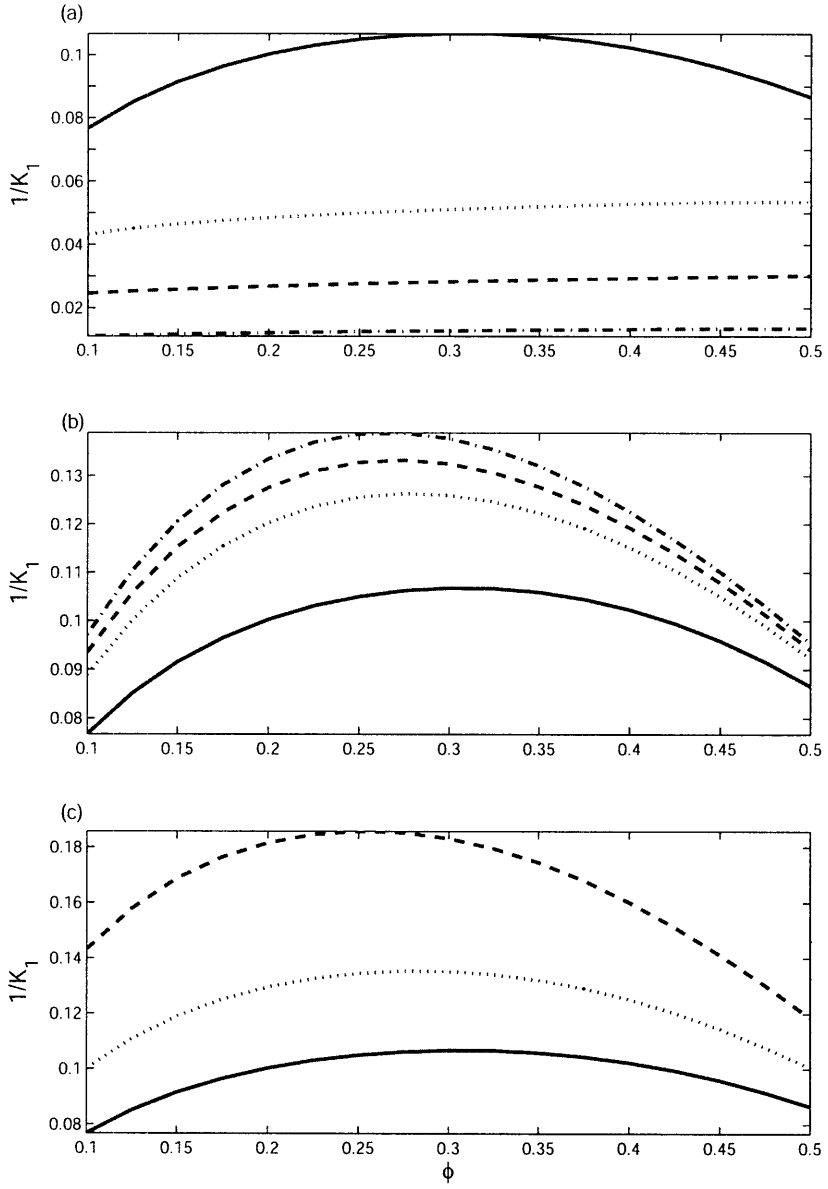

Figure 4-6: The spinodal lines for a short-range attraction and a long-range repulsion are determined as a function of $\frac{1}{K_1}$ vs. ϕ under different control parameters. The solid line in all three panels has the same set of parameters, $Z_1 = 10$, $K_2 = -0.3$, and $Z_2 = 2$. Panel (a) shows the effect on the spinodal line by increasing the repulsion strength, $|K_2|$. The dotted line, the dashed line, and the dash dotted line correspond to $K_2 = -1, \ -2$, and -5 respectively. Panel (b) shows the effect on the spinodal line by increasing Z_2 (decreasing the range of the repulsive potential, $\frac{1}{Z_2}$). The dotted line, the dashed line, and the dash dotted line correspond to $Z_2 = 3$, 4, and 6. Panel (c) shows the effect on the spinodal line by decreasing Z_1 (increasing the range of attractive potential, $\frac{1}{Z_1}$). The dotted line and the dashed line correspond to $Z_1 = 8$, and 6, respectively.

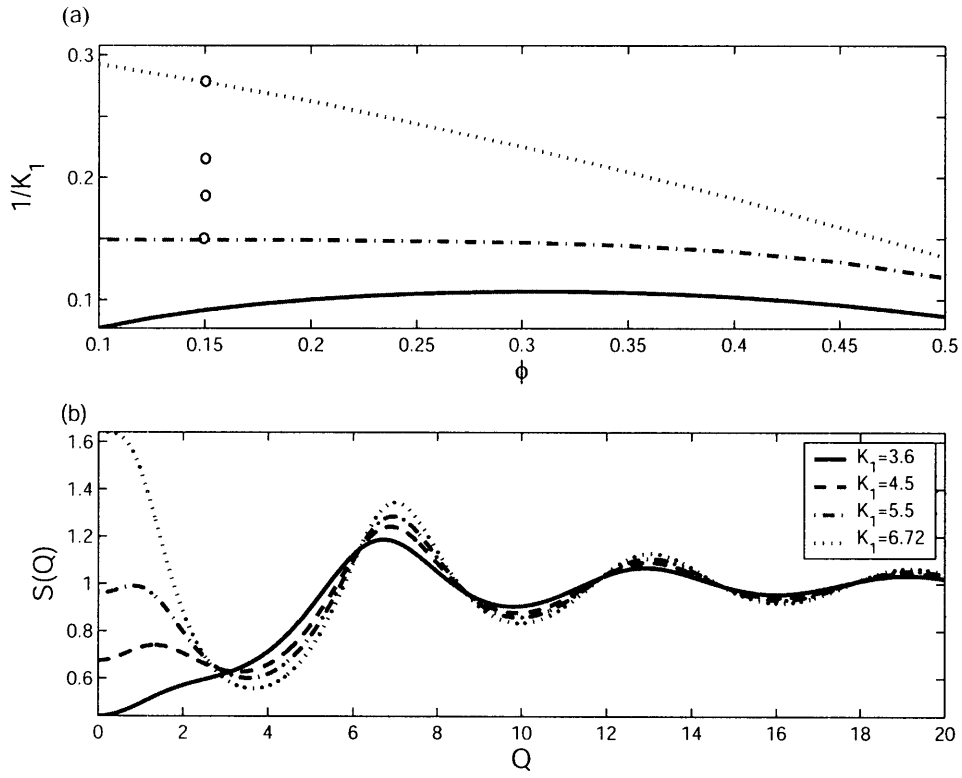

Figure 4-7: In panel (a), the cluster region in the phase diagram is shown as a function of $\frac{1}{K_1}$ (effective temperature), and the volume fraction, ϕ. $Z_1 = 10$, $K_2 = -0.3$ and $Z_2 = 2$. The dotted line is determined as the minimum attraction needed to generate a well defined cluster peak in the structure factor, while the dash dotted line is determined as the maximum attraction needed to have a well defined cluster peak in the structure factor. The region between the dotted line and the dash dotted is named as cluster region. The solid line is the spinodal line. The structure factors are shown in panel (b) with control parameters indicated by the open circles in the phase diagram of panel (a).

finite size clusters are favorably formed and a cluster peak could be observed in the structure factor. Figure 4-7 demonstrates the existence of this **cluster region**. In panel (a) of Figure 4-7, the dotted line is determined as the minimum attraction strength needed to have a cluster peak in the structure factor. And the dash dotted line is determined as the maximum attraction strength needed to have a cluster peak. The control parameters of this system is set as $\{Z_1 = 10,\ K_2 = -0.3,\ Z_2 = 2\}$. The spinodal line is also shown as the solid line in the figure. Figuratively speaking, the dotted line is called **high cluster transition line**, and the dash dotted line **low cluster transition line**. The physical mechanisms of the disappearance of the cluster peak at these two transition lines are different. Above the high cluster transition line, the attraction is weak, clusters can not form due to the strong repulsion and thermal motions. Therefore, the single particle motions dominate the system. Below the low cluster transition line, the attraction is so strong that the system is dominated by the attraction and in favor of forming isolated much bigger islands of dense liquids. In panel (b) of Figure 4-7, the evolution of the structure factor with the increase of the attraction strength, K_1, at the volume fraction $\phi = 0.15$ is shown. The solid line, the dashed line, the dash dotted line, and the dotted line correspond to $K_1 = 3.6,\ 4.5,\ 5.5$ and 6.72, which are indicated as open circles in the phase diagram in panel (a). $K_1 = 3.6$ is at the high cluster transition line. There is only an inflection point instead of a cluster peak. $K_1 = 6.72$ is at the low cluster transition line. $S(Q = 0)$ becomes very large and there is no cluster peak too. Although it is very qualitative to determine cluster transition lines by observing the local maximum (cluster peak) in the structure factor, it should give us a qualitatively correct information. The more quantitative way to determine these transition lines needs computer simulations. However, these two transition lines may not exist in some potentials with a short-range attraction plus a long-range repulsion, which can be seen more clearly in Fig. 4-8.

Fig. 4-8 shows the cluster regions at different long-range repulsions. The high cluster transition line and the low cluster transition line corresponding to the same set of control parameters are drawn in the same symbol. The parameters of the

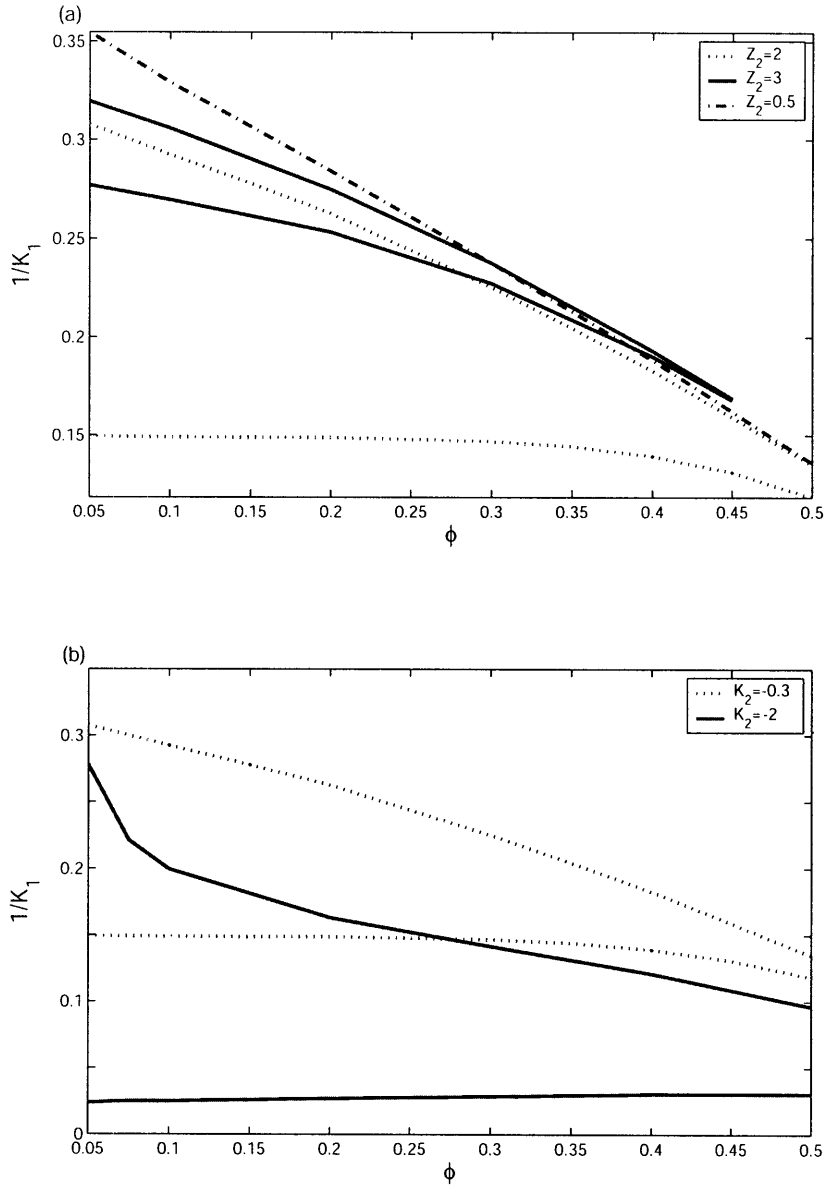

Figure 4-8: The effect of a long-range repulsion on a cluster region is demonstrated. The dotted line in both panels correspond to $\{Z_1 = 10,\ K_2 = -0.3,\ Z_2 = 2\}$. Two lines (high cluster transition line and low cluster transition line) with the same symbol sandwich a cluster region. Panel (a) shows the shift of the cluster region by changing the repulsion range, $\frac{1}{Z_2}$. The solid line and the dash dotted line correspond to $Z_2 = 3$ and $Z_2 = 0.5$ respectively. When $Z_2 = 0.5$, there is no low cluster transition line. Therefore, the area below the high cluster transition line is the cluster region. Panel (b) shows the shift of the cluster region due to the change of the amplitude of the long-range repulsion, $|K_2|$. The solid line corresponds to $K_2 = -2$.

dotted lines in both panels are $\{Z_1 = 10,\ K_2 = -0.3,\ Z_2 = 2\}$, which correspond to the cluster transition lines in panel (a) of Fig. 4-7. Panel (a) of Fig. 4-8 studies the effect on a cluster region by changing Z_2. When Z_2 increases, the cluster region becomes narrower. For $Z_2 = 3$, the cluster region dose not exist for volume fraction larger than $\phi = 0.45$. When Z_2 decreases to 0.5, only the high cluster transition line is observed. In order to find the low cluster transition line, we tested K_1 up to 100 and did not observe it. The increase of K_1 only makes the cluster peak height larger. Although our mathematical method can test even larger values of K_1, we did not see any practical reason to do that. Therefore, for this set of parameters, the whole area below the high cluster transition line belongs to the cluster region. The strong long-range repulsion favors the formation of stable clusters in liquids and prevents the macroscopic phase separation from happening. In panel (b), the effect of K_2 on the cluster region is studied. The solid line corresponds to $K_2 = -2$. The increase of K_2 widens the cluster region and makes both cluster transition lines move to smaller values. From Fig. 4-8 shows that it is easier to observer the cluster peak experimentally if we work with samples at low volume fraction. Adding salt to increase the screening effect will increase Z_2 and make more difficult to observe the cluster peak. Although the increase of K_2 can make the cluster region bigger, it also moves the high cluster transition line to smaller effective temperature. Therefore, if the attraction between two particles in liquids is not strong enough, it will also be difficult to observe the cluster peak.

With the aid of the two Yukawa model, we can also study the fluids with a short-range repulsion and a long-range attraction. Here, the repulsion is said to be short because it is compared with the range of attraction. It may not be necessarily short comparing to the hardcore. The significance of such systems will be seen in the next section.

Figure 4-9 shows the effect of the amplitude of the long-range attraction, K_1, on the structure factor with the control parameters, $\{Z_1 = 0.5,\ K_2 = -2,\ Z_2 = 2\}$ at the volume fraction, $\phi = 0.20$. The solid line, the dotted line, the dashed line, and the dash dotted line correspond to $K_1 = 0,\ 0.2,\ 0.4$, and 0.45 respectively. The

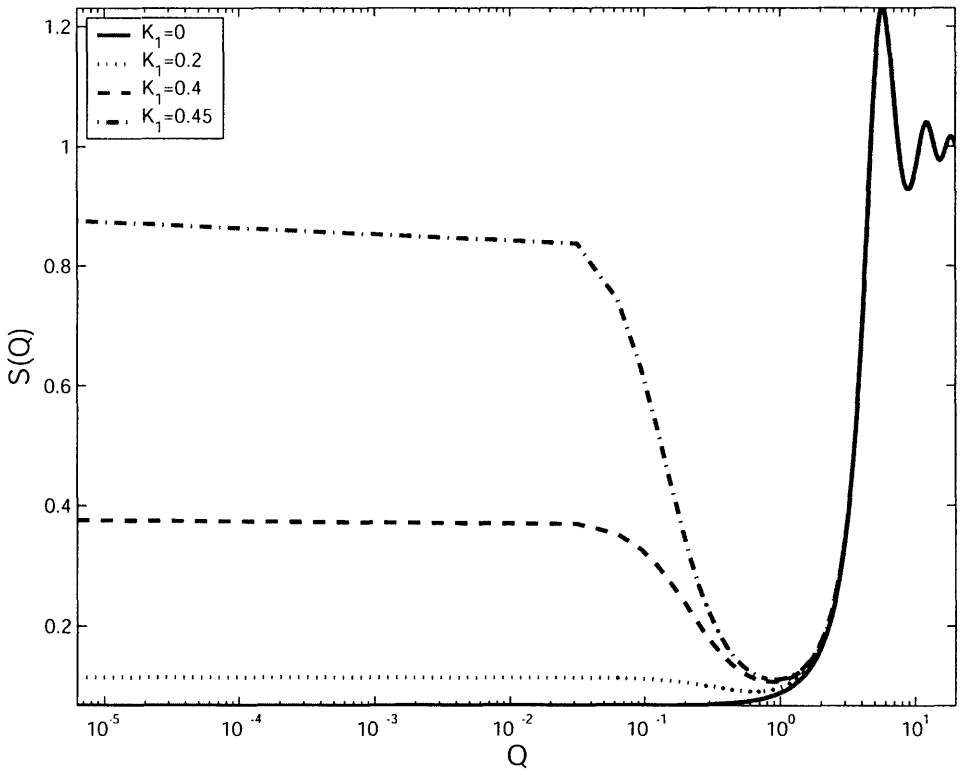

Figure 4-9: The structure factor generated by a relatively short-range repulsion and a long-range attraction is studied. The dependence on the attraction depth, K_1, is presented in this figure. Z_1, K_2, and Z_2 are set as 0.5, -2, and 2, respectively. The solid line, the dotted line, the dashed line, and the dash dotted line correspond to $K_1 = 0,\ 0.2,\ 0.4$, and 0.45 respectively.

Q-axis is plotted in *log* scale to exemplify the change of the structure factor at low Q range due to the increase of the strength of the long-range attraction. The solid line is the structure factor of a single repulsive Yukawa potential, which can also be obtained by the Hayter-Penfold's method. The addition of the long-range attraction dose not affect the first diffraction peak and only show the increase of structure factor intensity for Q close to zero. This is because that due to the long-range property, the attraction can not be very strong. Otherwise, the system is macroscopically phase separated already. Hence the change of the attraction strength dose not affect the potential form and strength close to the hardcore, which dominantly determines the first diffraction peak.

The spinodal lines for a two Yukawa fluids with a short-range repulsion plus a long-range attraction are also studied as a function of the effective temperature, $\frac{1}{K_1}$ vs. the volume fraction, ϕ, in Figure 4-10. The solid line in all three panels correspond to the same control parameters, S$\{Z_1 = 0.5,\ K_2 = -0.3,\ Z_2 = 2\}$. Other spinodal lines in each panel just differ with one control parameter to study the effect of a specific parameter. In panel (a), the effect of the repulsion strength, K_2 is studied. The dotted line, the dashed line, and the dash dotted line correspond to $K_2 = -1,\ -2,$ and -5 respectively. With the increase of the amplitude of the repulsion, a stronger attraction is needed to approach the spinodal line. Therefore the spinodal lines are shifted to smaller effective temperature, $\frac{1}{K_1}$. At the same time, the critical point at spinodal lines is shifted to the lower volume fraction. In panel (b), the effect of the repulsion range, $\frac{1}{Z_2}$, is studied. The dotted line, the dashed line, and the dash dotted line correspond to $Z_2 = 3,\ 4,$ and 6 respectively. The decrease of the repulsion range, $\frac{1}{Z_2}$, shift the spinodal line to higher effective temperature, $\frac{1}{K_1}$. In panel (c), the effect of the attraction range, $\frac{1}{Z_1}$ is studied. The dotted line, and the dashed line correspond to $Z_1 = 1$ and 1.5 respectively. The decrease of the attraction range, $\frac{1}{Z_1}$, shifts the spinodal line to lower effective temperature, $\frac{1}{K_1}$.

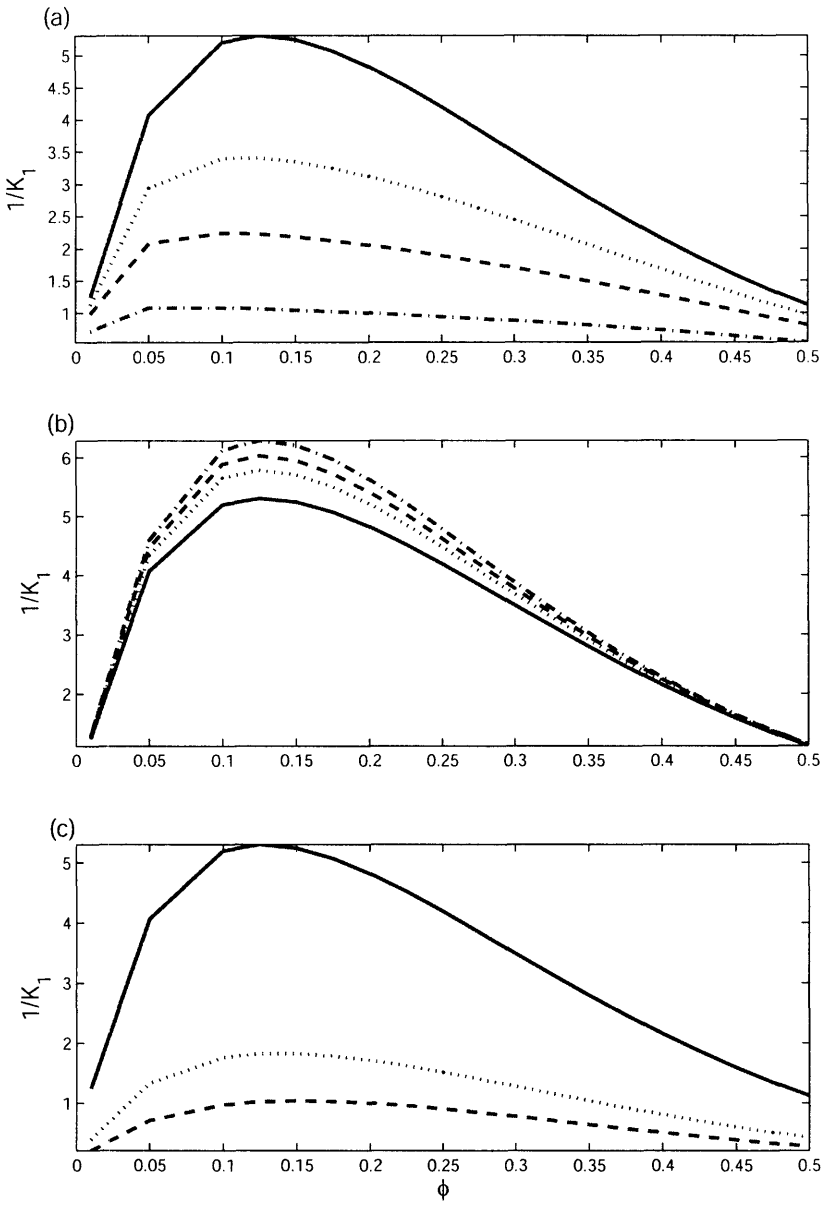

Figure 4-10: The spinodal lines for a long-range attraction plus a short-range repulsion are determined in the plane of $\frac{1}{K_1}$ vs. ϕ. The solid line in all three panels corresponds to the same set of parameters, $Z_1 = 0.5$, $K_2 = -0.3$, and $Z_2 = 2$. Panel (a) shows the effect on the spinodal line by increasing the repulsion strength, $|K_2|$. The dotted line, the dashed line, and the dash dotted line correspond to $K_2 = -1$, -2, and -5, respectively. Panel (b) shows the effect on the spinodal line by increasing Z_2 (decreasing the range of the repulsive potential, $\frac{1}{Z_2}$). The dotted line, the dashed line, and the dash dotted line correspond to $Z_2 = 3$, 4, and 6. Panel (c) shows the effect on the spinodal line by decreasing Z_1 (increasing the range of attractive potential, $\frac{1}{Z_1}$). The dotted line and the dashed line correspond to $Z_1 = 1$, and 1.5, respectively.

4.4 Conclusions

Due to the interest in studying the phenomenon exhibited by fluid systems with a short-range attraction and a long-range repulsion, we present an efficient method to solve the structure factor in a fluid with a two-Yukawa potential from the OZ equation with MSA closure. Compared with the previous methods[73, 74], our method is more complete and efficient. It is easier to implement in computer codes for an automatic search of the correct solution and thus offers an efficient way of fitting experimental data with an analytical structure factor. Using our method, any combinations of a two-Yukawa potential model can be studied besides the interesting case of a short-range attraction plus a long-range repulsion. We have studied the properties of the cluster peak in the structure factor generated by a short-range attraction and a long-range repulsion. We interpret this cluster peak as the result of the cluster formations inside the fluid. First, the case of a short-range attraction plus a long-rang repulsion is studied. Attentions are focused on the cases when a cluster peak appears. The change of the cluster peak as a function of different control parameters is studied. For a given set of control parameters, there is an optimal volume faction, where the intensity of cluster peak, $I_{cluster}$, is the largest. A cluster region, in which a cluster peak can be observed in the structure factor, is identified. The effect of the long-range repulsion on the cluster region is studied. Secondly, the case of a relatively short-range repulsion plus a long-range attraction on the structure factor is also studied. Because of its long range of the attractive potential comparing to the repulsion, the attraction affects the structure factor of very small Q values, while the first diffraction peak is mainly dominated by the short-range repulsion. Due to the presence of the short-range repulsion, the critical points in the spinodal lines are also shifted to a lower volume fraction.

Chapter 5

The Kinetic Phase Diagrams of Two-Yukawa Fluids

As pointed out in the previous chapter, recent experiments and computer simulations on colloidal systems identify a local peak in the structure factor at a wavevector much smaller than that of the first diffraction peak (particle peak)[63, 64]. This observation triggered much interest in the complex liquid community and raised fundamental questions regarding the nature of gelation, aggregation, and the glass transition[64, 84, 85, 86, 87, 88, 89]. Sciortino and coworkers found that the presence of long-range repulsive and short-range attractive potentials[64] leads to the aggregation of colloidal particles and results in a liquid-glass transition at low densities. Experiments on proteins and Laponite systems with this kind of interactions revealed signatures of new cluster regimes[63, 87, 88].

In this chapter, we systematically study a model system with a short-range attractive and a long-range repulsive Yukawa potentials in addition to the hard-core potential. The structure factors are calculated by the mean-spherical approximation (MSA) described in the previous chapter, and the kinetic phase diagrams for the idealized liquid-glass transition are obtained in the framework of mode-coupling theory (MCT)[90, 91, 92]. Liquids and glasses composed of clusters are analyzed in detail. Although the standard MCT is a first order dynamic mean-field approximation[93, 94], its predictions have been extensively verified in colloidal systems. For example,

MCT predicts a liquid-glass transition at a critical volume fraction, $\phi^c \approx 0.516$, in the hard-sphere colloidal system[91], whereas the experimental result is around 0.58[95]. The volume fraction is related to the number density ρ and the colloidal diameter σ by $\phi = \pi\rho\sigma^3/6$. A recent high-order MCT calculation by Wu and Cao[94] improves the prediction to $\phi^c = 0.552$. In attractive colloidal systems, MCT predicts repulsive and attractive glass states[81, 96], which has been confirmed in experiments[83]. MCT is expected to provide reliable predictions, at least in the lowest order, in our model system.

SAXS experiments on lysozyme solutions reveal that the protein-protein potential has a short-range attraction, and a long-range electrostatic repulsion[58, 60] that can be described by a Yukawa form according to DLVO[52] and GOCM[61] theory. With the unclear origin, the attraction has been simulated by different functional forms[60, 64]. To exploit the known analytical form of of S_q, we use a Yukawa form for the attraction part in our system, resulting in

$$\beta V(r) = \begin{cases} \infty, & 0 < r < 1 \\ -\frac{K_1 e^{-Z_1(r-1)}}{r} + \frac{K_2 e^{-Z_2(r-1)}}{r}, & r > 1 \end{cases}, \tag{5.1}$$

where r is the dimensionless inter-particle distance in units of σ, K_1 and Z_1 refer to the attraction, and K_2 and Z_2 refer to the repulsion. As in the screened columbic potential, the inverse of the dimensionless repulsion range is given by the DLVO theory as $Z_2 = \kappa\sigma$[52], where κ is the inverse of the Debye-Hückel screening length. The dependence of K_2 on both the ionic strength I and the particle charge z can be calculated by either DLVO[52] or GOCM[61].

The structure factor can be obtained by the method described in the previous chapter. Compared to other methods[73], our approach generates S_q efficiently over a broad range of wavevectors, which serves as a reliable input for our MCT calculations.

In the remainder of this chapter, we will fix Z_1 and Z_2 and study the kinetic phase diagram as a function of K_1 and ϕ for three values of K_2. Three repulsive heights are employed, $K_2 = 0.3$, 1, and 5, where the corresponding particle charges in the DLVO theory are $z = 1.5$, 2.7, and 6, respectively. We set $Z_2 = 0.5$, which corresponds to a

weak ionic strength of $I = 2.1mM$ and $\sigma = 33$Å, and $Z_1 = 10$, because the attraction range in the protein systems is around 10% of the hard-sphere diameter[58].

Using the analytical structure factor S_q as the input, we employ the standard MCT[90, 91, 92] to calculate the Debye-Waller (nonergodic) factor f_q, defined as the normalized long time limit of the coherent intermediate scattering function, $f_q = \lim_{t\to\infty} F_q(t)/S_q$. The corresponding high-order MCT calculations[94] will be investigated in the future. A central result of MCT is the self-consistent equation for f_q, which predicts the idealized liquid-glass transition[91]. When a system is in the liquid state (ergodic), only the trivial solution, $f_q = 0$, exists but when the system is in the glass state (nonergodic), another nontrivial solution, $f_q > 0$, appears. The transition point in the parameter space is determined by the discontinuity of f_q from zero to a non-vanishing value. To obtain f_q in our system, we follow the numerical method in Ref. [92].

The kinetic phase diagrams are plotted for three K_2 in Fig. 5-1A-C, where the horizontal axis is the volume fraction ϕ and the vertical axis is the inverse of the attraction depth, K_1^{-1}. Similar to attractive colloidal systems[81], the liquid regime (LR) is surrounded by the upper glass regime (UGR) and lower glass regime (LGR). As shown in Fig. 5-1B-C, increasing the repulsion strength extends the liquid phase to larger ϕ and K_1 because the system requires stronger attraction to trap colloids or higher volume fraction to form arrested cages. The glass-glass re-entrance phenomenon occurs at $\phi \geq 0.516$, where the transition line between the LR and UGR appears. When the attraction is weak, the cage effect induced by high ϕ dominates and the system is in the UGR, similar to the repulsive glass in the attractive colloidal suspensions[81]. Because the colloidal particles move closer as attraction increases, the collective motions recover the ergodicity and the system enters the liquid phase. When the attraction is sufficiently strong, colloid particles are randomly trapped by potential wells deeper than thermal fluctuations, and the system enters the LGR. Although sharing a similar mechanism, the LGR for the two Yukawa potential is more complicated than the attractive glass state in the attractive colloidal system[81].

Clusters in the liquid and glass states of our systems are stabilized by the com-

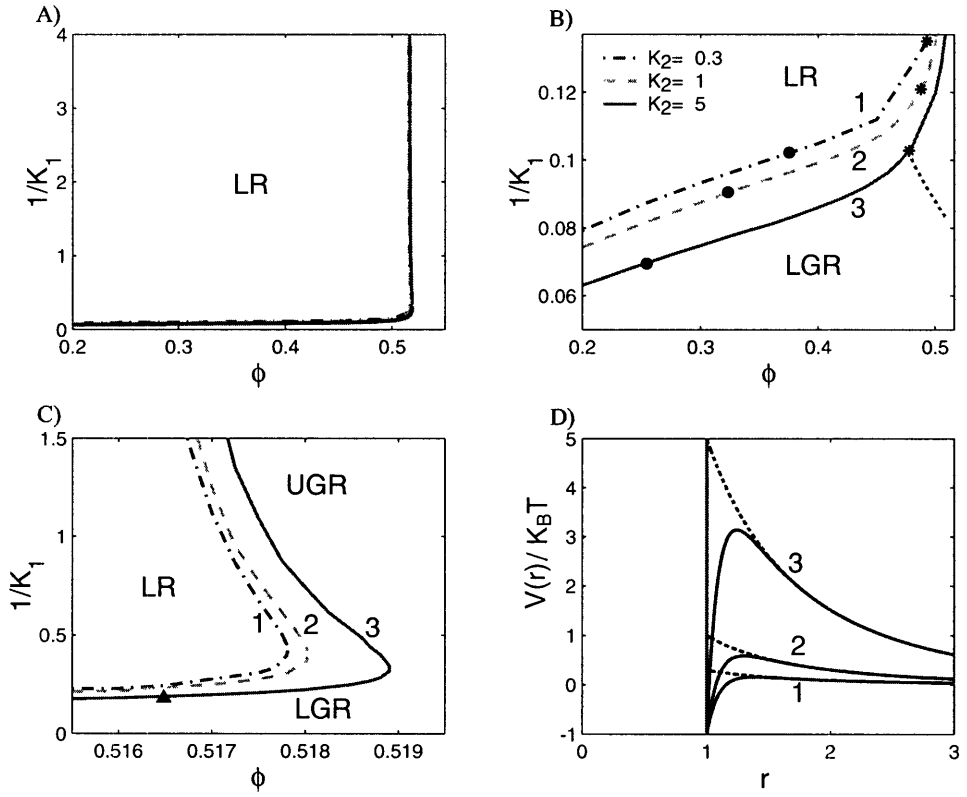

Figure 5-1: Kinetic phase diagrams and potentials for $Z_1 = 10$ and $Z_2 = 0.5$. Curves labelled by 1, 2, and 3 correspond to $K_2 = 0.3$, 1, and 5, respectively. **A**) Kinetic phase diagrams are shown in a wide range of ϕ and K_1^{-1}. **B**) The lower part of A) to show the transition between the LR and LGR. The stars mark the crossover from the static cluster glass to the dynamic cluster glass (see text). The dotted line demonstrates the separation of these two cluster glass states inside LGR for $K_2 = 5$. The solid circles mark ϕ where the cluster peaks are equal to the particle peaks in S_q. **C**) Another part of A) close to the ending point of LR to mark the termination of the dynamic cluster glass. **D**) Potential surfaces for different K_1 and K_2. The dash dotted line is for $K_2 = 0.3$ and $K_1 = 1.3$. The dotted line is for $K_2 = 1$ and $K_1 = 2$. The solid line is for $K_2 = 5$ and $K_1 = 6$. The dotted lines correspond to the pure long-range repulsion.

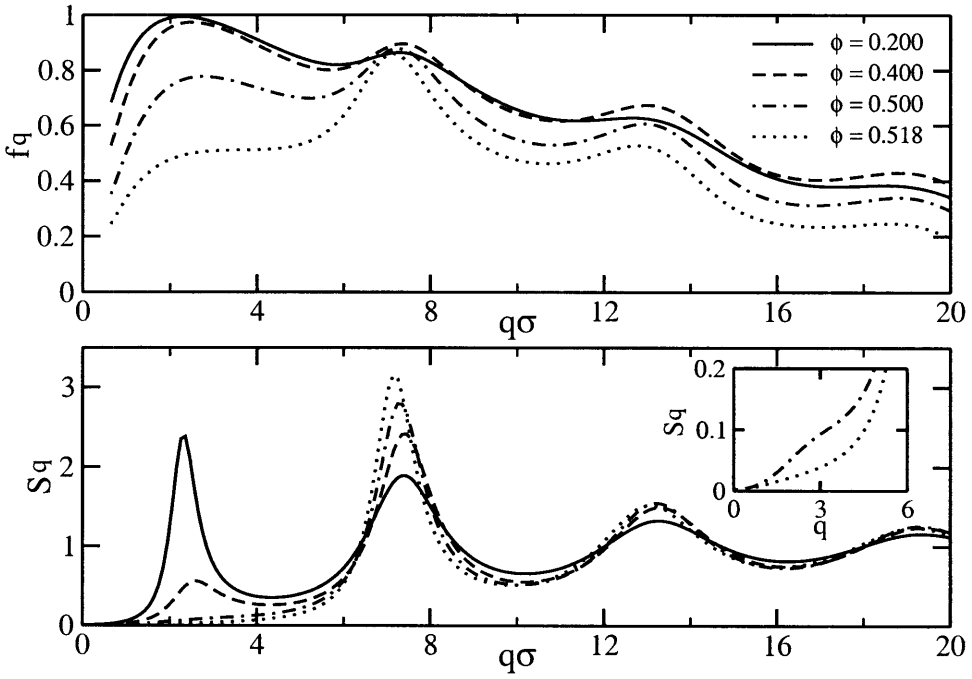

Figure 5-2: **A**) Four Debye-Waller factors f_q along the transition line between the liquid regime (LR) and the lower glass regime (LGR) for $K_2 = 5$. **B**) The structure factor S_q corresponding to the same parameters in A).

petition between the short-range attraction and long-range repulsion[64, 86, 87, 89]. Based on microscopic descriptions with S_q and f_q, we now discuss these cluster states in detail. In Fig. 5-2, the structure factors and the corresponding Debye-Waller factors along the transition line between the LR and LGR are plotted for $K_2 = 5$ and for four values of ϕ ranging from 0.2 to 0.518. In addition to the particle peak at $q_m \approx 7.3\sigma^{-1}$, we observe the cluster peak in several S_q and f_q curves at a much smaller wavevector, $2.3\sigma^{-1} < q'_m < 3.3\sigma^{-1}$. We discuss separately four states in Fig. 5-2. (i) For $\phi = 0.2$, the cluster peaks are higher than the particle peaks in both S_q and f_q, and $f_{q'_m}$ is equal or close to one, indicating that colloidal particles aggregate into stable close packing clusters. The characteristic size of clusters is finite due to the balance of the attraction and repulsion. Because of their thermal and dynamic stability, close packing clusters are the basic units of the system instead of colloidal monomers. As a result, it is possible to improve our predictions for low volume fractions using the effective cluster description proposed recently[89]. Characterized by the presence of the cluster peak in S_q, the ergodic state can be viewed as a cluster liquid, the nonergodic state discussed here is a '**static cluster glass**'. (ii) For $\phi = 0.4$, we observe $S_{q'_m} \ll S_{q_m}$ and $1 \approx f_{q'_m} > f_{q_m}$. Compared to the case of $\phi = 0.2$, the cluster population substantially decreases, suggesting that the cluster structures are less ordered. The large cluster peak in f_q indicates that the less-ordered cluster structure remains dynamically more stable than single colloids. Because the calculation is carried out along the transition line between the LR and LGR in Fig. 5-1B, the attraction strength K_1 decreases as ϕ increases. The disordered cluster structure is related to the decrease of the strength and effective width of the attractive well. Compared to experiments and schematic phase diagrams[84, 87], the glass states at $\phi = 0.4$ and 0.2 may be the precursors of the gel state in Ref.[87], which is an infinite percolating network resulting from the growth of the finite-size clusters in our system. (iii) For $\phi = 0.5$, we observe an inflection point instead of a cluster peak at $q \approx 3.2$ in S_q as shown in the inset of Fig. 5-2B, and the cluster peak in f_q is smaller than the particle peak. Affected by both the increase in ϕ and the decrease in K_1, the cluster structure is highly disordered and does not have a characteristic

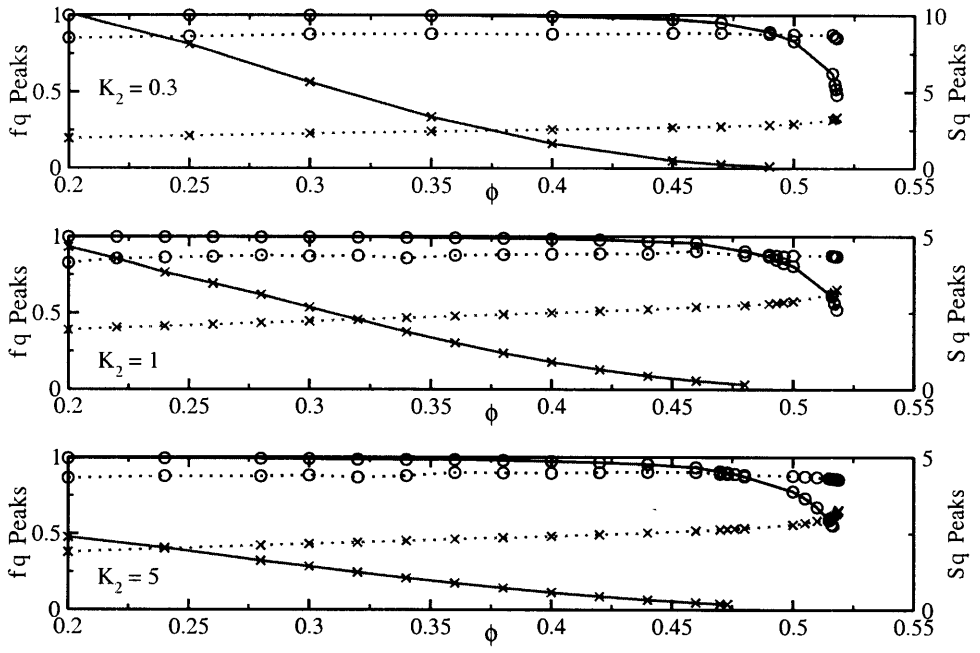

Figure 5-3: This figure summarizes the intensity changes of the cluster peaks and the particle peaks in f_q and S_q along the transition line between the LR and LGR. Lines with circles denote peaks in f_q, referring to the left vertical axes, and lines with crosses denote peaks in S_q, referring to the right vertical axes. The solid lines are the cluster peaks, whereas the dotted lines are the particle peaks.

size. However, the cluster peak in f_q suggests that disordered cluster structures in the glass state are dynamically selected with a characteristic cluster size. The wavevector of the cluster peak in f_q is smaller than the inflection point in S_q, indicating that clusters are more likely to result from dynamics than thermodynamics. These cluster structures may be related to the heterogeneous structures in the activation picture of glass formation[97, 98]. Following these arguments, we consider this glass state as a '**dynamic cluster glass**'. (iv) For $\phi = 0.518$, the cluster peaks in both S_q and f_q disappear. With the attraction depth of $K_1 = 4.5$, we have $\beta V(r = 1^+) > 0$ so the binary bond between two colloidal particles is unstable and will dissociate. As a result, the probability of forming stable clusters is small and the basic dynamic unit in the glass state is a single colloidal particle.

We now extend the above discussion of Fig. 5-2 for one value of K_2 to several values of the repulsion strength. Figure 5-3 is a plot of the cluster and particle peaks

in S_q and f_q along the transition line between the LR and LGR for $K_2 = 0.3$, 1, and 5, respectively. The basic behaviors of the peak values are similar for different K_2. With the rapid decrease of $S_{q'_m}$, the cluster peak in f_q remains close to one for $\phi \leq 0.4$, but quickly decreases around $0.45 < \phi < 0.5$. Interestingly, as $f_{q'_m}$ becomes smaller than f_{q_m}, the cluster peak in S_q turns into an inflection point, indicating a crossover from the static to the dynamic cluster glass regimes. We define the crossover density ϕ_1 between the two regimes by the disappearance of the cluster peak in S_q, giving $\phi_1(K_2 = 0.3) = 0.49$, $\phi_1(K_2 = 1) = 0.48$, and $\phi_1(K_2 = 5) = 0.47$, denoted by stars in Fig. 5-1B. These values of ϕ_1 are close to the freezing point, 0.49, of hard-sphere fluids[95]. This coincidence implies a possible connection between cluster aggregation and crystallization. To demonstrate the tendency for the separation of these two cluster glass states inside LGR, a calculation for $K_2 = 5$ is provided as the dotted line in Fig. 5-1B. The termination of the dynamic cluster glass regime is marked by the disappearance of the cluster peak in f_q, which defines another crossover density ϕ_2. For the weak repulsive strength, e.g., $K_2 = 0.3$ and 1, the attractive well remains negative in the glass regime, excluding the observation of ϕ_2. For the strong repulsive potential $K_2 = 5$, the crossover density is $\phi_2(K_2 = 5) = 0.5165$, denoted by a triangle in Fig. 5-1C. Although the above conclusions drawn from f_q and S_q are crude, future studies will also investigate other measures to explore cluster formation in different regimes.

In summary, we study the structural arrest transitions in a model system with a short-range attractive and a long-range repulsive Yukawa potentials. By computing the Debye-Waller factor f_q from the standard MCT[90, 91, 92], we obtain the kinetic liquid-glass phase diagrams and observe the glass-glass re-entrance phenomenon. The detailed studies of cluster peaks and particle peaks in S_q and f_q along the transition line between the liquid regime and the lower glass regime reveal different cluster regimes in our system, most notably the static cluster glass and the dynamic cluster glass. Our results confirm and significantly extend recent computer simulation results of Sciortino et al.[64] and experiments by Tanaka et al.[87]. More studies are required to investigate detailed mechanisms for forming different cluster regimes as well as their

relationships to the gel phase and attractive glass phase. As shown in the previous chapter, the ranges of attraction and repulsion strongly affect the characteristic size of clusters. Following this line of research, we hope to find the connection between the gelation process, which forms a network with an infinite size, and the aggregation process, which forms clusters with finite sizes.

Chapter 6

Studies of the Effective Potential between Proteins in Solutions

The bottleneck of protein crystallography is the lack of systematic methods to obtain protein crystals. This is partly due to incomplete understanding of the physical chemistry conditions controlling the growth of protein crystals. A full comprehension of the effective protein interactions and phase behavior is therefore essential. It has been shown that the crystallization curves of some globular proteins appear to coincide with the phase diagrams of a hard sphere system interacting with a short range attraction [99, 100, 101]. Small angle neutron and X-ray scattering investigations of proteins suggest the presence of a short-range attractive interaction between protein molecules besides the electrostatic repulsion induced by the residual charges [58, 102, 63]. The DLVO potential has been successfully applied to many colloidal systems and protein solutions [58, 101]. However, it does not seem to fully explain the rich behaviour of proteins [58, 103, 104, 105], and due to the complexity of these systems (anisotropic property, irregular shape, distributed charge patches, etc.), a complete understanding of the properties of the effective interactions between protein molecules in solutions remains a challenge [104].

Recent measurements of small angle neutron scattering (SANS) intensity distribution in protein solutions show interesting results [63, 102, 106]. Beside the normal first diffraction peak, it is present a peak (cluster peak) appearing at a much smaller

scattering wave vector, Q, due to the formation of ordered clusters. The appearance of a cluster peak is explained as due to the competition of a short-range attraction and a long-range electrostatic repulsion [64, 102]. Moreover, a rising intensity as Q approaches zero (zero-Q peak) is observed in both liquid-like and solid-like samples, which implies that the effective potential should have more features in addition to the well known short-range attraction and electrostatic repulsion. The existence of a long-range attraction between protein molecules has already been postulated in theoretical studies. Noro et al. suggested that the presence of a long-range attractive force between protein molecules should shift the metastable fluid-fluid critical point out of gel regime [107]. Thus a protein crystallization may occur without gelation. By employing two Yukawa potential model, Lawlor et al. showed that the introduction of a long-range attraction between protein molecules can enhance crystal growth by avoiding the formation of a disordered state, an attractive glass [81, 83], while preserving the equilibrium features of a system with the short-range attraction [108].

In this chapter, by systematically studying the zero-Q peak, we show that a weak long-range attraction needs to be considered to explain the SANS scattering intensity distributions. The properties of this long-range attraction potential are also investigated.

6.1 Sample Preparation and Experimental Setup

Cytochrome C from horse heart (product no. C7752) was purchased from Sigma and is obtained using a procedure that avoids the trichloroacetic acid (TCA) that is known to promote the dimer formation. Cytochrome C has been dialyzed three times in order to remove any extra salt. Buffers are not used to avoid possible bindings of organic molecules to the protein surface. Lysozyme from chicken egg white was purchased from Fluka (product no. L7651) and used without further purification. The pD of cytochrome solutions at 7.2 and 9.5, has been adjusted by a HCl standard solution. Only samples of Lysozyme have been prepared in 20 mM HEPES buffer. The pD value was sequentially checked by an ISFET pHMeter (KS723) before and after each

performed experiment and found stable within 0.1 units. All samples were prepared a few days before the scheduled experiments to allow the hydrogen/deuterium exchange. For all solutions the final protein concentration has been measured by UV-Visible spectroscopy.

Our experiments were performed at the small angle neutron scattering station, NG7, at the Center of Neutron Research in the National Institute of Standard and Technology. Two configurations have been used to reach a wide range of the wave vector, Q, from $0.004\mathring{A}^{-1}$ to $0.30\mathring{A}^{-1}$, where $Q = \frac{4\pi}{\lambda} sin(\theta/2)$, λ is the neutron wavelength, θ the neutron scattering angle. All the analyses have taken into account the instrumental resolution correction.

SANS intensity distribution, $I(Q)$, can be expressed as

$$I(Q) = A\overline{P}(Q)S(Q) \tag{6.1}$$

where $\overline{P}(Q)$ is the normalized particle structure factor, $S(Q)$, the inter-particle structure factor, A, the known amplitude factor which is proportional to the volume fraction and the square of the neutron scattering length contrast between protein and solvent [61]. Cytochrome C has an ellipsoidal shape with the semi-major and minor axes, $a \times b \times b = 15 \times 17 \times 17\mathring{A}^3$, while a lysozyme molecule has a dimension $a \times b \times b = 22.5 \times 15 \times 15\mathring{A}^3$. $\overline{P}(Q)$ is calculated by considering the ellipsoidal shape of the protein.

In the fitting, $\overline{P}(Q)$ is calculated as an ellipsoidal shape with fixed axis ratio $\frac{b}{a}$. Therefore, $\overline{P}(Q)$ is determined by a single parameter, the length of semi-radius a.

$S(Q)$ is calculated by solving the Ornstein-Zernike (OZ) equation within the mean spherical approximation (MSA) closure involving an effective pair potential, $V(r)$, to be specified later. Denoting the structure factor calculated by the OZ equation $S(Q)_{TH}$, the effect on the inter-particle structure factor, $S(Q)$, due to the non-spherical shape of a particle can be approximately taken into account by the

decoupling approximation as [109, 110]

$$S(Q) = 1 + \beta(Q)(S_{TH}(Q) - 1) \tag{6.2}$$

where

$$\begin{aligned} \beta(Q) &= |< F(Q) >|^2 / < |F(Q)|^2 > \\ < F(Q) > &= \int_{-1}^{1} d\mu \frac{3j_1(u)}{u}, \\ < |F(Q)|^2 > &= \int_{-1}^{1} d\mu (\frac{3j_1(u)}{u})^2 \\ u &= Q\sqrt{a^2\mu^2 + b^2(1-\mu^2)} \end{aligned} \tag{6.3}$$

and $j_1(u)$ is the first order spherical Bessel function.

6.2 The Generalized One Component Macroion Theory (GOCM)

In the fitting, we can calculate the screened Coulomb repulsive potential directly with given charge number of a protein molecule and the ionic strength of the solution. The DLVO theory is widely used to calculate this potential in dilute samples[52]. In moderate concentrations, Wu and Chen successfully applied the GOCM theory to fit the SANS data of Cytochrome C protein solutions up to the volume fraction about 18%[61, 62]. Our SANS data of Cytochrome C protein solutions were obtained with the volume fraction at about 10% and 20%. Therefore GOCM is necessary to explain our experimental results. The brief result of GOCM is given as following[61, 62].

Denote the number of the protein charge z_p and the ionic strength of solution I. The screened Coulombic potential between proteins is expressed as

$$\begin{aligned} V_{GOCM} &= \frac{z_p^2 e^2}{\epsilon \sigma_H} Y^2 \frac{e^{-kx}}{x}, \quad x > 1 \\ Y &= \cosh(k/2) + U[(k/2)\cosh(k/2) - \sinh(k/2)] \end{aligned} \tag{6.4}$$

with

$$
\begin{aligned}
U &= \mu(k/2)^{-3} - \gamma(k/2)^{-1} \\
\mu &= 3\phi/(1-\phi) \\
\gamma &= (G/2+\mu)/(1+G/2+\mu) \qquad (6.5) \\
& \qquad (6.6)
\end{aligned}
$$

and G is the solution of the implicit equations

$$
\begin{aligned}
G^2 &= k^2 + \frac{t_0^2}{(1+G/2+\mu)^2} \qquad (6.7) \\
& \qquad (6.8)
\end{aligned}
$$

where

$$t_0^2 = \frac{24\beta z_p^2 e^2 \phi}{\epsilon \sigma_H}. \qquad (6.9)$$

$\beta = 1/(k_B T)$, $\sigma_H = 2(ab^2)^{1/3}$ is the particle diameter, and k is defined as $k = \kappa\sigma_H$, where the screen constant κ is given by

$$\kappa = (\frac{8\pi e^2 I \beta N_A}{10^3})^{1/2}. \qquad (6.10)$$

6.3 Results and Discussions

Fig. (6-1) shows theoretical calculations together with the SANS result from a cytchrome C solution at $pD = 11$ with 1% volume fraction in $1M$ $NaCl$ in the inset. Conventionally, the effective inter-protein potential is considered to consist of a short-range attractive and a long-range electrostatic repulsive part. The repulsion is screened out by the concentrated salt. Therefore, $S(Q)$ could be obtained by solving OZ equation by considering only a short-range attraction, which can be approximated by an attractive potential of a Yukawa form [60], $V(r) = -K\frac{e^{-Z(r-1)}}{r}$, where K is normalized by $k_B T$, and r is the inter-protein distance normalized by σ

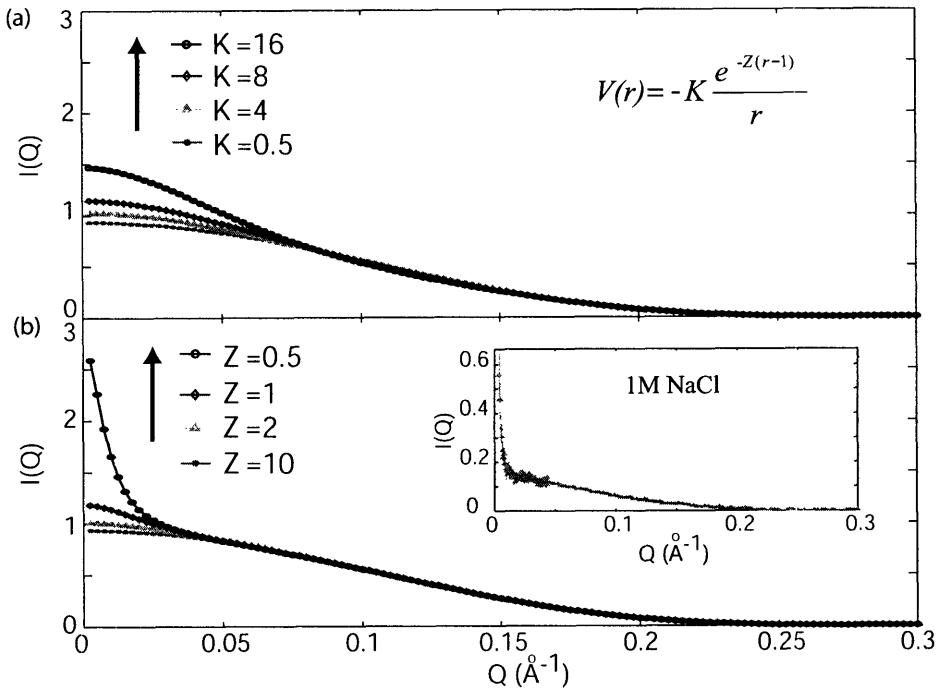

Figure 6-1: Theoretical calculations of $I(Q)$ resulting from one Yukawa attraction at 1% volume fraction. a) the effect due to the variation of the attraction strength, K at $Z = 10$. b) the effect due to the variation of the attraction range, $1/Z$ at $K = 0.5$. The inset shows $I(Q)$ from a cytochrome C sample at 1% volume fraction in $1M$ $NaCl$ at pD=11.

with $\sigma = 2(ab^2)^{1/3}$. The interaction range, $\frac{1}{Z}$, is approximately about 10% of the protein diameter[58]. Fig.(6-1)(a) shows the theoretically calculated SANS intensity distribution with $Z = 10$ by taking the amplitude factor A equal to unity. As K increases from 0.5 to 16, the intensity at $Q = 0$ increases gradually. The results from a short-range attraction always smoothly change the whole scattering curve and can not reproduce the sharp rising up of experimental SANS intensity at very low Q (zero-Q peak) as shown in the inset. Since the typical strength of short-range attraction is smaller than $10k_BT$[58, 111, 112], the short-range attraction can not explain the observed SANS intensity distribution. In Fig. (6-1)(b), K is fixed at 0.5. When Z decreases from 10 to 0.5, i.e. the attraction range increases, the intensity at low Q increases sharply and the theoretical curve exhibits a similar trend to the experimental intensity. The comparison between the experimental result and theoretical curves thus suggests that the zero-Q peak could be induced by a weak long-range attraction. Therefore, the effective potential between protein molecules in solutions should consists of three features: a short-range attraction, an intermediate-range electrostatic repulsion, and a weak long-range attraction. This zero-Q peak has been overlooked in previous experiments due to the limited Q range covered [61]. However, it has been observed before in lysozyme protein solutions [113] and was explained as due to the long-range density fluctuation.

To investigate the properties of this long-range attraction, three different cytochrome C samples at pD=11 with 1% volume fraction added with different salts were measured and their SANS intensity distributions with error bars are plotted in semi-log scale to clearly show zero-Q peaks in Fig. (6-2). The salt concentration, which is indicated in the figure, has been changed to keep the same ionic strength. Very interestingly, the results show that the zero-Q peak depends on the different anions added. $NaCl$ induces a much weaker zero-Q peak compared with that of $NaSCN$ and Na_2SO_4. Since the true form of this long-range attraction is still unknown, in order to fit the experimental results, we assume that the effective inter-protein potential can be simulated by two attractive Yukawa form potential, $V_{TY}(r) = -K_1\frac{e^{-Z_1(r-1)}}{r} - K_2\frac{e^{-Z_2(r-1)}}{r}$. The first term is used to simulate the short-

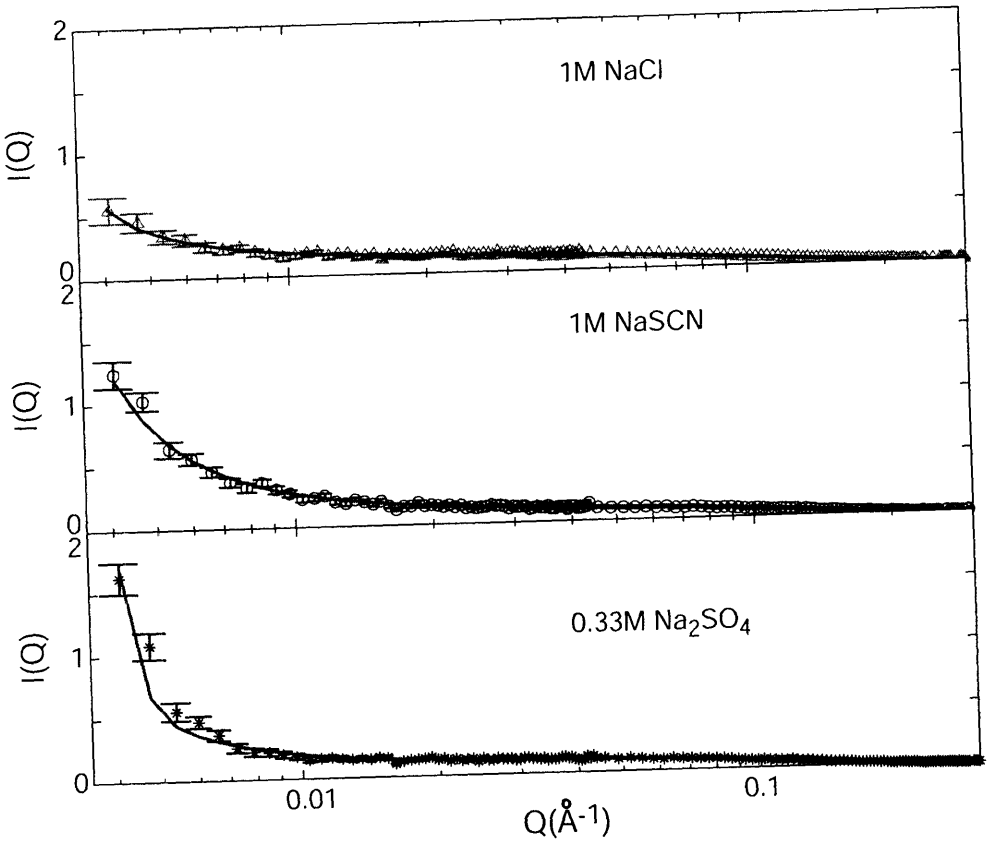

Figure 6-2: SANS intensity distribution of cytochrome C solutions at pD=11 and at 1% volume fraction with different salts added. The solid lines are the theoretical analyses.

Table 6.1: Fitted parameters using the two attractive Yukawa form potential. Results are shown in Figure (6-2).

	$a(\mathring{A})$	K_1	Z_1	K_2	Z_2
$NaCl$	14.7	7 ± 3	7 ± 3	0.11 ± 0.02	0.20 ± 0.01
$NaSCN$	14.6	8 ± 5	10 ± 6	0.33 ± 0.02	0.35 ± 0.01
Na_2SO_4	14.7	4 ± 5	10 ± 10	0.24 ± 0.02	0.27 ± 0.02

range attraction, while the other one is used to simulate the long-range attraction. The fitted results are given in table I.

The fitting is not sensitive at all to the short-range attraction. The long-range attraction has a weak strength (less than 0.5 k_BT) and a range of $3 \sim 5$ times the protein diameter, and it is very sensitive to the anions added to the solutions. This implies that the long-range attraction is at least partly induced by the ion cloud around protein molecules. Depletion forces and van der Waals force can not explain this feature. It is very interesting to notice that the strength and range of the long-range attraction is about the same value as the attractive potential between like-charged particles [114]. However, a charged colloidal particle typically has a uniform charge distribution on the surface, while a protein molecule has both positive and negative charge patches, which make the ion cloud distribution around a protein much more complicated.

Since the long-range attraction seems to be induced by the ion cloud, if we can minimize the ion concentration in solutions, we should expect to greatly suppress the zero-Q peak. Fig. (6-3) shows SANS results of two cytochrome C samples at pD=7.2 and pD=9.5 with 5% volume fraction. Without adding any extra salt in solutions, the ion concentration is determined by the dissociated charges from proteins. The cytochrome C molecule has much smaller charge number at pD=9.5 than that at pD=7.2, being the isoelectric point, pI, about 10.2. This leads to a smaller ion concentration in solutions at pD=9.5. Correspondingly, the SANS intensity distribution shows a zero-Q peak weaker than that at pD=7.2. This implies that the zero-Q peak depends on the ion concentrations, and is not likely due to permanent cluster forma-

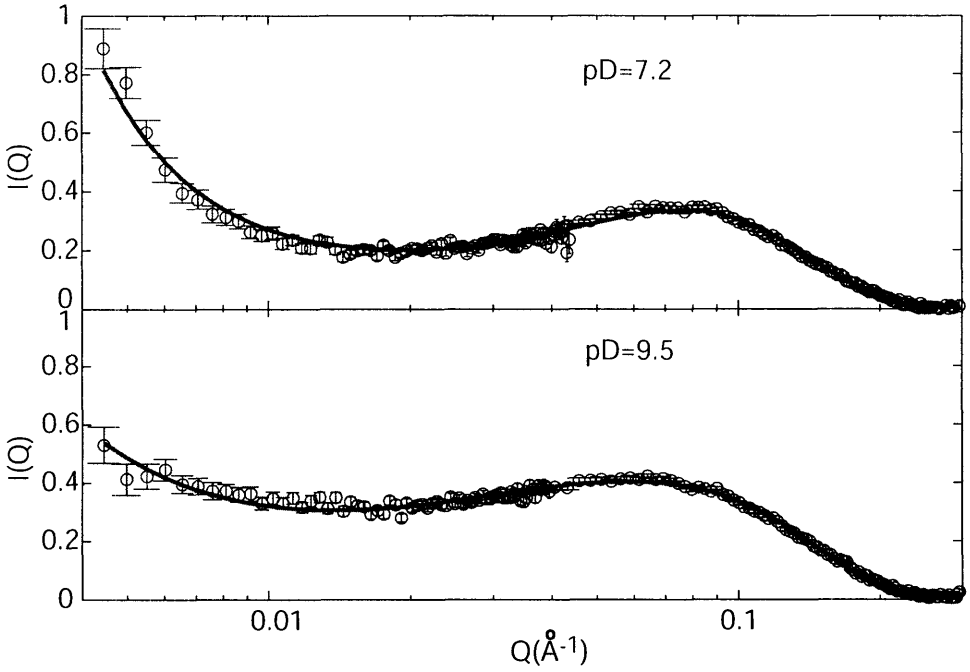

Figure 6-3: SANS intensity distributions of cytochrome C samples at 5% volume fraction. The solid lines are theoretical analyses using two Yukawa potential.

tions, since a weaker repulsion should favor larger cluster formations and thus induce larger zero-Q peak at pD=9.5, which is in contrast with our experimental results. In order to quantitatively analyze the zero-Q peak, we again used the two-Yukawa potential model. However, due to the existence of the weak long-range attraction, we used three Yukawa terms to completely simulate the features of the potential. We argue that in cytochrome C solutions, the short-range attraction is very small so that its effect on the structure factor should be very small at the relatively low volume fraction. Actually, the first diffraction peak of SANS intensity distribution in cytochrome C solutions has been successfully analyzed by only considering the electrostatic repulsion at various pD values [61]. Furthermore, for cytochrome C solutions, we did not observe the cluster peak as it is observed for lysozyme solutions. The lack of cluster peak is attributed to the weak short-range attraction. Therefore, we ignore the short-range attraction and we use the second term of V_{TY} for the long-range attraction contribution and the first term of V_{TY} to simulate the electrostatic

Table 6.2: Fitted parameters obtained using the two Yukawa potential. Results are shown in Figure (6-3).

	$a(\AA)$	charge number	K_2	Z_2
$pD = 7.2$	15.3	3.8 ± 0.1	0.37 ± 0.02	0.34 ± 0.01
$pD = 9.5$	15.4	1.7 ± 0.1	0.08 ± 0.01	0.21 ± 0.01

repulsion. This last can be calculated by the charge number of the protein molecule and the ionic strength [61]. The fitted results given in table II show that the range of the long-range attraction is about $3 \sim 5$ times of protein diameter, and also confirm our direct observation from Fig. (6-3), i.e. the strength of long-range attraction at pD=9.5 is much smaller.

After showing the existence of the long-range attraction in solutions dominated by monomers, it is important to check its existence in protein solutions having equilibrium clusters. Fig. (6-4) shows the results of lysozyme protein solutions at pD=5.1 and at 10% and 20% volume fraction in HEPES buffer. The top panel shows SANS intensity distributions with the fitted results and the bottom panel shows the calculated $S(Q)$ from the fitted parameters. The *main peak* in the top panel is a cluster peak as it is independent of volume fraction. Interestingly, even in the presence of cluster peak, both results have shown the zero-Q peak, which was overlooked by previous experiments [102]. The coexistence of the cluster peak and zero-Q peak clearly indicates the necessity of introducing a third potential feature, a long-range attraction. We used the two Yukawa model previously described to fit the main peak, assuming that the first term is the short range attraction, and the second term the electrostatic repulsion. The fitted curves are shown as dotted line in the figure and the fitting results are given in table III. The attraction strength is about $7k_BT$ with 7% of attraction range, which is consistent with Ref.[111]. In order to fit the complete SANS distribution, we use the two Yukawa model as the reference system and treat the long-range attraction for a perturbation by employing the random phase approximation [115]. The long-range attraction is considered here as a third Yukawa form, $-K_3\frac{e^{-Z_3(r-1)}}{r}$. The results fitted by within this approximation

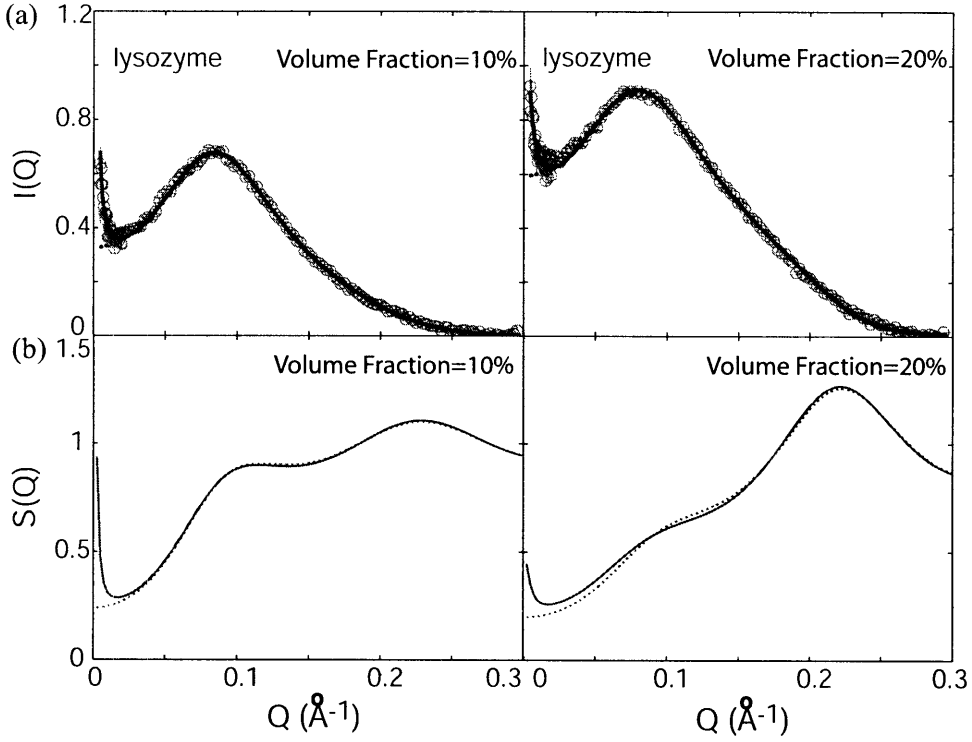

Figure 6-4: SANS intensity distributions of lysozyme samples at pD=5.1 with 20 mM HEPES buffer. Dotted lines (a and b) are fitted by considering only a short-range attraction and electrostatic repulsion. Solid lines (a and b) are the fitted results by using three Yukawa form potential. Lower panel shows the calculated structure factor, $S(Q)$.

Table 6.3: Fitted parameters from fitting the cluster peaks by using only a short-range attraction and the electrostatic repulsion (see Fig. (6-4)).

	$a(\AA)$	charge number	K_1	Z_1
10%	21.1	7.0 ± 1.2	6.9 ± 1.0	14 ± 5
20%	21.2	6.8 ± 1.5	7.7 ± 1.2	15 ± 5

are shown as solid line. The addition of this third potential feature slightly changes the previous fitting parameters. The parameters for the long-range attraction are, $K_3 = 0.038 \pm 0.005$ and $Z_3 = 0.17 \pm 0.01$ for the 10% lysozyme sample, while for 20% sample, $K_3 = 0.019 \pm 0.002$ and $Z_3 = 0.21 \pm 0.03$.

6.4 Conclusions

Judging from SANS results on both cytochrome C and lysozyme protein solutions, we believe that the existence of weak long-range attraction is universal for all protein solutions. Its strength and range depend both on the type of anion and ion concentration in solutions. In other words, it depends on the ion cloud around protein molecules. Our finding has important implications in understanding the protein crystallization process. There are numerous measurements of second virial coefficient in protein solutions by light scattering experiments [116, 117]. Due to the very small Q range the light scattering can access, its intensity is closely related to the height of the zero-Q peak, which depends on the feature of the long-range attraction. Thus we believe that the conclusions derived from those observations should take into account the existence of the weak long-range attraction. The charge dependence of some of these phenomena can then be naturally related to the charge dependence of the long-range attraction [104, 118]. Also the very large cluster formation or gelation in supersaturated protein solutions may be induced by the long-range attraction

6.5 Perspectives for Future Work

Our results shows the existence of a weak long-range attraction between proteins in solutions. In order to have better understanding of the properties of the long-range attraction, we are planning to study how the long-range attraction changes with different control parameters. This information is very helpful to try to formulate a theory for the phenomenon.

The observation of the long-range attraction and its dependence on the ion cloud

may thus help in understanding many phenomena: 1) There are many experimental results on protein solutions depending on the type of ions [103, 104]. For example, it is known that the cloud point temperature of lysozyme protein not only depends on the ionic strength but also on the type of ions added [103]. Since the long-range attraction is apparently dependent on the type of ions, the study of the long-range attraction can shed light in understanding of these puzzles related with the type of ions; 2) In saturated protein solutions prior to the protein crystallization or precipitation, it seems that very large protein clusters may form [120, 121]. Also depending on different solution conditions, the protein precipitation can form either "craggs" (precrystalline aggregates) or "praggs" (precipitating aggregates). The understanding of the role of this long-range attraction during these processes thus may help to have better control of the protein solutions to grow better quality of protein crystals; 3) From our results, the strength and the range of the long-range attraction seems to be about the same value of the long-range attraction between like-charged particles in solutions [114], which has been debated for about two decades. At current stage, we are still not sure if the observed long-range attraction is unique to the protein solutions or it is a general feature for charged colloidal particles. The study of this long-range attraction will help to answer this question and thus provide useful information to the general understanding of the effective inter-particle potential in charged colloidal solutions, which is a fundamental problem in colloidal science.

With this in mind, the long-range attraction between proteins in solutions should be systematically studied. First, we would like to investigate the control parameters of this long-range attraction. From our results, it is obvious that this long-range attraction depends on the type of ions and ion concentrations. Hence it is natural to study its dependence on: 1) the ion concentration, 2) the valency of different anions and cations, 3) ions with different size which directly affect the ion dynamics. Second, it is natural to study the temperature dependence of this long-range attraction. Third, based on the properties extracted from the SANS results, we would like to study the relation between the long-range attraction and the protein behaviors related with different type of ions, such as the Hofmeister series, the ion dependence of the cloud

point temperature[103], and the abnormal behavior of the second virial coefficients in some protein solutions [104, 119]. Fourth, we would like to study the saturated protein solutions to investigate the role of the long-range attraction at protein crystallization conditions and how it affects the path of protein crystallization or gelation. Fifth, it is very appealing to investigate some colloidal systems having similar particle size and charges to see if this long-range attraction is a general feature of all charged colloidal particle or it is just only specific to protein solutions.

Appendix A

A List of Publications

- Y. Liu, S. H. Chen, D. Berti, P. Baglioni, A. Alatas, H. Sinn, E. Alp, A. Said, "Effects of counterion valency on the damping of phonons propagating along the axial direction of liquid crystalline DNA", submitted to *Journal of Chemical Physics*, (2005).

- Y. Liu, E. Fratini, P. Baglioni, W. R. Chen, S. H. Chen, "An effective long-range attraction between protein molecules in solutions studied by small angle neutron scattering", submitted to *Physical Review Letters*, (2005).

- Y. Liu, D. Berti, P. Baglioni, S. H. Chen, A. Alatas, H. Sinn, A. Said, E. Alp, "Inelastic X-ray scattering studies of phonons propagating along the axial direction of a DNA molecule having different counter-ion atmosphere", submitted to *Journal of Physics and Chemistry of Solids*, (2005).

- Y. Liu, W. R. Chen, S. H. Chen, "Cluster Formation in Two Yukawa Fluids", *Journal of Chemical Physics*, **122**, 044507 (2005). (Selected for the January 15, 2005 issue of Virtual Journal of Biological Physics Research.)

- J. Wu, Y. Liu, W. R. Chen, J. Cao, S. H. Chen, "Structural Arrest Transitions in Fluids Described by Two Yukawa Potentials", *Physical Review E* Rapid Communication, **70**, 050401, (2004).

- W. R. Chen, Y. Liu, F. Mallamace, P. Thiyagarajan, S. H. Chen, "Studies of Structural Arrest Transition in $L64/D_2O$ Micellar Solutions", *Journal of Physics: Condensed Matter*, **16**, S4951, (2004).

- Y. Liu, D. Berti, A. Faraone, W. R. Chen, A. Alatas, H. Sinn, E. Alp, P. Baglioni, S. H. Chen, "Inelastic X-ray Scattering Studies of Phonons in Liquid Crystalline DNA", *Physical Chemistry Chemical Physics*, **6**, 1499 (2004).

- P. J. Chen, Y. Liu, T. A. Weiss, H. W. Huang, H. Sinn, E. Alp, A. Said, S. H. Chen, "Studies of Short-Wavelength Collective Molecular Motions in Lipid Bilayers Using High Resolution Inelastic X-ray Scattering", *Biophysical Chemistry*, **105**, 721 (2003).

Appendix B

Manual for TYSQ21 Matlab Package

B.1 Introduction

(Please cite the paper, Yun Liu, Wei-Ren Chen, Sow-Hsin Chen, "Cluster Formation in Two Yukawa Fluids", *Journal of Chemical Physics*, 122, 044507 (2005), if you use the results produced by this code.)

TYSQ21 is the version 2.1 of the matlab package, TYSQ, written to calculate the structure factor of one component liquid systems interacting with a two-term Yukawa potential with the mean spherical approximation (MSA). The structure factor is generated by following Blum's paper (J. Stat. Phys., 16, 399, 1977), in which the structure factor is solved from Ornstein-Zernike equation by Baxter's Q-method with MSA closure.

The potential $V(r)$ is

$$\beta V(r) = \begin{cases} \infty, & 0 < r < 1 \\ -c(r) = -K1\frac{e^{-Z1(r-1)}}{r} - K2\frac{e^{-Z2(r-1)}}{r}, & r > 1 \end{cases} \tag{B.1}$$

where $\beta = \frac{1}{k_B T}$.

The Ornstein-Zernike equation is

$$h(\vec{r}) = c(\vec{r}) + \rho \int d\vec{r}' c(|\vec{r}'|) h(|\vec{r} - \vec{r}'|). \tag{B.2}$$

The closure form is

$$\begin{cases} h(r) = -1, & 0 < r < 1 \\ c(r) = K1 \frac{e^{-Z1(r-1)}}{r} + K2 \frac{e^{-Z2(r-1)}}{r}, & r > 1 \end{cases} \tag{B.3}$$

B.2 Installation

1. Obtain the compressed zip file, TYSQ21.zip

2. Extract the TYSQ21.zip to the directory where you want to install it. Let's assume the directory's name is '**HomeDir**'. You should find **TYSQ21** subdirectory under your directory **HomeDir**. There are also other two files under **HomeDir**: **TwoYukawaSample.m**, **Manual.pdf**. The first file is a sample file you can run to get a flavor how to run this program. The second file is this manual.

B.3 Using the Package

1. Run your Matlab to have your Matlab command window.

2. Choose your working directory.

   Easy way (not recommended): run your own code under the directory **HomeDir/TYSQ21**. The problem of this method is that you can not choose a file which has the same name as that of some file under the directory **HomeDir/TYSQ21/private**.

   Neat way (recommended): write your own program under different directory. Let's assume your working directory name is '**workDir**'. In this case, you have

to let the Matlab to know where your **TYSQ21** package is. The command you should use is 'addpath'.

Windows machine: type the following command in your Matlab command window: addpath HomeDir\TYSQ21 -begin

Linux or Unix machine: type the following command in your matlab command window: addpath HomeDir/TYSQ21 -begin

3. Now you are ready to use the package. The central function file is **CalTYSk.m** under the directory **HomeDir/TYSQ21**. You can copy **TwoYukawaSample.m** from **HomeDir** directory to your **workDir**. Then type in command window TwoYukawaSample, and press ENTER key. It will take about 10 seconds to plot out a structure factor and its pari distribution function $g(r)$ with $Z1 = 10,\ Z2 = 2,\ K1 = 6,\ K2 = -1$, and volume fraction 20%. **CalTYSk.m** function always assumes that the hardcore diameter is one.

B.4 Q&A

1. Can $K1$ and $K2$ be zero?

   No. The codes assume that there must be two Yukawa terms. Therefore, if you want to make one or both of them zero, you can only do it by making them very small values.

2. Can I make $Z1$ and $Z2$ very large number?

   In principle, the answer is Yes. However, this is very subtle. In general, always try to make $Z1 > Z2$, when you want to try a very large number, such as $Z > 20$. This is because that $Z1$ and $Z2$ are treated in an asymmetric way in the computer codes. In order to make the calculation more accurate, $Z1 > Z2$ will be a nice trick. In general, when $Z < 20$, it dose not matter.

   When $Z > 25$, sometimes the intermediate results of this codes may run into the limit of the largest number that a computer can handle. Therefore, results

may potentially become less reliable. Hence, the check of $g(r)$ becomes very important in those situations.

So far, I did not find out any limitation of the value of K except that they can not be zero.

3. Can $Z1$ and $Z2$ be equal?

   $Z1$ and $Z2$ should not be equal. If they are equal, there is essentially only one Yukawa term. Therefore, the algorithm designed for two term Yukawa potential will fail. However, the codes can handle cases that $Z1$ and $Z2$ have only very small differences.

B.5 Improvement of TYSQ21 over TYSQ01

For this version, the major revision is that the input value of Q could have zero value point compared with the previous version, TYSQ01.

Bibliography

[1] The information is obtained from the website of Human Genome Project, whose URL link is http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml

[2] J. P. Boon, S. Yip. 1991. *Molecular Hydrodynamics* (Dover, New York).

[3] S. H. Chen, and M. Kotlarchyk. 1997. *Interaction of Photons and Neutrons with Matters.* (World Scientific, Singapore)

[4] H. Sinn, E. E. Alp, A. Alatas, J. Barraza, G. Bortel, E. Burkel, D. Shu, W. Sturhahn, J. P. Sutter, T. S. Toellner, J. Zhao. Nucl. Instrum. Methods Phys. Res. A **467-468**, 1545 (2001).

[5] F. Sette, G. Ruocco, M. Krisch, C. Masciovecchio, R. Verbeni, U. Bergmann. Phys. Rev. Lett **77**, 83 (1996).

[6] G. Ruocco, F. Sette. J. Phys.: Condens. Matter **11**, R259 (1999).

[7] J. Teixeira, M. C. Bellissent-Funel, S. H. Chen, B. Dorner. 1985. Phys. Rev. Letts **54**, 2681 (1985).

[8] G. Monaco, A. Cunsolo, G. Ruocco, F. Sette. 1999. Phys. Rev. E **60**, 5505 (1999).

[9] C. Masciovecchio, S. C. Santucci, A. Gessini, S. Di Fonzo, G. Ruocco, F. Sette. Phys. Rev. Lett. **92**, 255507 (2004)

[10] A. Cunsolo, G. Pratesi, R. Verbeni, D. Colognesi, C. Masciovecchio, G. Monaco, G. Ruocco, F. Sette. 2001. J. Chem. Phys. **114**, 2259 (2001).

[11] I. M. de Schepper, E. G. D. Cohen, C. Bruin, J. C. van Rijs, W. Montfrooij, L. A. de Graaf. Phys. Rev. A **38**, 271 (1988).

[12] C. Y. Liao, S. H. Chen, F. Sette. 2000. Phys. Rev. E **61**, 1518 (2000).

[13] C. Y. Liao, S. H. Chen. Phys. Rev. E **64**, 21205 (2001).

[14] S. H. Chen, C. Y. Liao, H. W. Huang, T. M. Weiss, M. C. Bellisent-Funel, F. Sette. Phys. Rev. Lett. **86**, 740 (2001).

[15] P. J. Chen, Y. Liu, T. M. Weiss, H. W. Huang, H. Sinn, E. E. alp, A. Alatas, A. Said, S. H. Chen. Biophys. Chem. **105**, 721 (2003).

[16] T. M. Weiss, P. J. Chen, H. Sinn, E. E. Alp, S. H. Chen, H. W. Huang. Biophys. J. **84**(6), 3767 (2003).

[17] *International Tables for x-ray Crystallography* (Reidel, Dordrcht, 1985), Vol. III.

[18] F. Livolant, A. Leforestier. Prog. Polym. Sci. **21**, 1115 (1996).

[19] H. Urabe, H. Hayashi, Y. Tominaga, Y. Nishimura, K. Kubota, M. Tsuboi.J. Chem. Phys. **82**, 531 (1985).

[20] L. L. Van Zandt. Phys. Rev. Lett. **57**, 2085 (1986).

[21] G. Edwards, G. Ying, J. Tribble. Phys. Rev. A **45**, R8344 (1992).

[22] Y. Tominaga, M. Shida, K. Kubota, H. Urabe, Y. Nishimura, M. Tsuboi. J. Chem. Phys. **83**, 5972-5975 (1985).

[23] E. W. Prohofsky. Phys. Rev. A **38**, 1538 (1988).

[24] H. Grimm, H. Stiller, C. F. Majkrzak, A. Rupprecht, U. Dahlborg. Phys. Rev. Lett. **59**, 1780 (1987).

[25] H. Grimm, A. Rupprecht. Physica B **234-236**, 183 (1997).

[26] T. Weidlich, S. M. Lindsay, A. Rupprecht. Phys. Rev. Lett. **61**, 1674 (1988).

[27] A. P. Sokolov, H. Grimm, A. Kisliuk, A. J. Dianoux. J. Biol. Phys. **27**, 313 (2001).

[28] A. P. Sokolov, H. Grimm, R. Kahn. J. Chem. Phys. **110**, 7053 (1999).

[29] W. Doster, S. Cusack, W. Petry. Phys. Rev. Lett. **65**, 1080 (1990).

[30] W. Doster, S. Cusack, W. Petry. Nature **337**, 754 (1989).

[31] A. M. Tsai, D. A. Neumann, L. N. Bell. Biophys. J. **79**, 2728 (2000).

[32] G. Maret, R. Oldenbourg, G. Winterling, K. Dransfeld, A. Rupprecht. Coll. Polym. Sci. **257**, 1017 (1979).

[33] S. A. Lee, S. M. Lindsay, J. W. Powell, T. Weidlich, N. J. Tao, G. D. Lewen, A. Rupprecht. Biopolymers **26**, 1637 (1987).

[34] N. J. Tao, S. M. Lindsay, A. Rupprecht. 1988. Biopolymers **27**, 1655 (1988).

[35] H. Urabe, M. Kato, Y. Tominaga. 1989. J. Chem. Phys. **92**, 768 (1989).

[36] T. Scopigno, G. Ruocco, F. Sette, G. Monaco. Science **302**, 849 (2003).

[37] H. Kriegs, W. Steffen, G. Fytas, G. Monaco, C. Dreyfus, P. Fragouli, M. Pitsikalis, N. Hadjichristidis. 2004. J. Chem. Phys. **121**, 2376 (2004).

[38] M. Bée. 1988. *Quasielastic Neutron Scattering: Principles and Applications in Solid State Chemistry, Biology and Materials Science.* (Adam Hilger, Bristol and Philadelphia)

[39] S. Cocco, R. Monasson. J. Chem. Phys. **112**, 10017 (2000).

[40] N. K. Voulgarakis, G. Kalosakas, K. Ø. Rasmussen, A. R. Bishop. Nano Lett. **4**, 629 (2004).

[41] H. Sampoli, G. Ruocco, F. Sette. Phys. Rev. Lett. **79**, 1678 (1997).

[42] V. J. Emery, J. D. Axe, Phys. Rev. Lett. **40** (1978) 1507.

[43] J. D. Axe, in *Ordering in Strongly Fluctuating Condensed Matter Systems*, edited by T. Riste (Plenum, New York, 1980).

[44] H. Sinn, B. Glorieux, L. Hennet, A. Alatas, M. Hu, E. E. Alp, F. J. Bermejo, D. L. Price, M.-L. Saboungi, Science **299** (2003) 2047.

[45] M. Feig, B. M. Pettitt. Biophys. J. **77**, 1769 (1999).

[46] A. D. MacKerell, Jr.. J. Phys. Chem. B **101**, 646 (1997).

[47] A. Patkowski, W. Eimer, Th. Dorfmüller. Biopolymers **30**, 975 (1990).

[48] R. S. Keyes, A. M. Bobst. Biochemistry **34**, 9265 (1995).

[49] M. E. Hogan, O. Jardetzky. Biochemistry **19**, 3460 (1989).

[50] S. T. Lin, P. K. Maiti, W. A. Goddard, III. J. Phys. Chem. **109**, 8663 (2005).

[51] M. Tarek, D. J. Tobias. Phys. Rev. Lett. **88**, 138001:1-4 (2002).

[52] E. J. W. Verwey and J. TH. G. Overbeek, *Theory of the Stability of Lyophobic Colloids*, (Elsevier, Amsterdam), (1948).

[53] J. M. Victor and J. P. Hansen, J. Chem. Soc., Faraday Trans. 2 **81**, 43 (1985).

[54] V. Morales, J. A. Anta, and S. Lago, Langmuir **19**, 475 (2003).

[55] M. J. Grimson, J. Chem. Soc., Faraday Trans. 2 **79**, 817 (1982).

[56] J. Kaldasch, J. Laven, and H. N. Stein Langmuir **12**, 6197 (1996)

[57] S. K. Lei, W. P. Peng, and G. F. Wang, Phys. Rev. E **63**, 041511 (2001)

[58] A. Tardieu, A.L. Verge, M. Malfois, F. Bonneté, S. Finet, M. Riés-Kautt, and L. Belloni, J. Crystal. Growth **196**, 193, (1999).

[59] A. Tardieu, S. Finet and F. Bonneté, J. Crystal. Growth, **232**, 1, (2001).

[60] M. Malfois, F. Bonneté, L. Belloni, and A. Tardieu, J. Chem. Phys. **105**, 3290, (1996).

[61] C. F. Wu and S. H. Chen, J. Chem. Phys. **87**, 6199, (1987).

[62] C. F. Wu and S. H. Chen, Biopolymers **27**, 1065, (1988).

[63] B. Lonetti, E. Fratini, S. H. Chen, and P. Baglioni, Phys. Chem. Chem. Phys. **6**, 1388, (2004).

[64] F. Sciortino, S. Mossa, E. Zaccarelli, and P. Tartaglia, Phys. Rev. Lett. **93**, 055701 (2004).

[65] M. Seul and D. Andelman, Science **267**, 476 (1995)

[66] R. P. Sear, S. W. Chung, G. Markovich, W. M. Gelbart, and J. R. Heath, Phys. Rev. E. **59**, R6255 (1999).

[67] F. P. Sear and W. M. Gelbart, J. Chem. Phys. **110**, 4582 (1999).

[68] E. Waisman, J. S. Høye, and G. Stell, Chem. Phys. Lett. **40**, 514, (1976).

[69] J. S. Høye, G. Stell, and E. Waisman, Molec. Phys. **32**, 209, (1976).

[70] J. S. Høye and L. Blum, J. Stat. Phys. **16**, 399, (1977).

[71] E. Waisman, Molec. Phys. **25**, 45, (1973).

[72] J. B. Hayter and J. Penfold, Molec. Phys. **46**, 651, (1981).

[73] J. Konior and C. Jedrzejek, Molec. Phys. **48**, 219, (1983).

[74] E. Arrireta, C. Jedrzejek, and K. N. Marsh, J. Chem. Phys. **95**, 6806, (1991).

[75] R. J. Baxter, Austral. J. Phys. **21**, 563, (1968).

[76] The complete codes are written in a Matlab package and could be obtained by contacting Yun Liu (yunliu@mit.edu) or Sow-Hsin Chen (sowhsin@mit.edu).

[77] G. Pastore, Molec. Phys. **63**, 731, (1988).

[78] E. Leutheusser, Phys. Rev. A **29**, 2756 (1984)

[79] W. Götze and L. Sjögren, Prog. Phys. **55**, 241 (1992)

[80] T. Franosch, M. Fuchs, W. Götze, M. R. Mayr, and A. P. Singh, Phys. Rev. E **55**, 7153 (1997)

[81] K. Dawson, G. Foffi, M. Fuchs, W. Götze, F. Sciortino, M. Sperl, P. Tartaglia, Th. Voigtmann, and E. Zaccarelli, Phys. Rev. E **63**, 011401 (2000)

[82] J. Bergenholtz, M. Fuchs, Phys. Rev. E **59**, 5706 (1998)

[83] S. H. Chen, W. R. Chen, F. Mallamace, Science **300**, 619 (2003)

[84] F. Sciortino, Nat. Mater. **1**, 145, (2002).

[85] K. A. Dawson, Curr. Opin. Coll. Int. Sci. **7**, 218, (2002).

[86] J. Groenewold and W. K. Kegel, J. Phys. Chem. B **105**, 11702, (2001).

[87] H. Tanaka et al., Phys. Rev. E **69**, 031404, (2004); R. P. Sear et al., Phys. Rev. E **59**, R6255, (1999).

[88] B. Ruzicka, L. Zulian, and G. Ruocco, to be published in J. Phys. Cond. Matt. **16**(42), S4993 (2004)

[89] M. E. Cates et al., arXiv: cond-mat/0403684.

[90] E. Leutheusser, Phys. Rev. A, **29**, 2765, (1984);

[91] W. Götze and L. Sjögren, Prog. Phys. **55**, 241, (1992).

[92] T. Franosch, M. Fuchs, W. Götze, M. R. Mayr, and A. P. Singh, Phys. Rev. E **55**, 7153, (1997).

[93] J. S. Cao, J. L. Wu, and S. L. Yang, J. Chem. Phys. **116**, 3739, (2002); J. L. Wu and J. S. Cao, Phys. Rev. E **67**, 061116, (2003).

[94] J. L. Wu and J. S. Cao, to be published in J. Phys. Chem.; submitted.

[95] P. N. Pusey and W. van Megen, Nature, **320**, 340 (1986).

[96] J. Bergenholtz and M. Fuchs, Phys. Rev. E **59**, 5706 (1999).

[97] C. A. Angell, J. Phys. Cond. Matt. **12**, 6463 (2000).

[98] X. Xia and P. G. Wolynes, Proc. Natl. Acad. Sci. USA **97**, 2990 (2000); Phys. Rev. Lett. **86**, 5526, (2001).

[99] D. Rosenbaum et al., Phys. Rev. Lett. **76**, 150 (1996).

[100] M. H. J. Hagen et al., J. Chem. Phys. **101**, 4093 (1994).

[101] G. Pellicane et al., J. Phys.: Condens. Matter **16**, S4923 (2004).

[102] A. Stradner, et al., Nature **432**, 492 (2004).

[103] M. L. Broide et al., Phys. Rev. E **53**, 6325 (1996).

[104] R. Piazza, Curr. Opin. Colloidal Interface Sci. **8**, 515 (2004).

[105] A. Striolo et al., J. Chem. Phys. **116**, 7733 (2002).

[106] P. Baglioni et al., J. Phys.: Condens. Matter **16**, S5003, (2004).

[107] M. G. Noro et al., Europhys. Lett. **48**(3), 332 (1999).

[108] A. Lawlor et al., Prog. Colloid Polym. Sci. **123**, 104 (2004).

[109] S. H. Chen, Ann. Rev. Phys. Chem. **37**, 351, (1986).

[110] M. Kotlarchyk et al., J. Chem. Phys. **79**, 2461 (1983).

[111] W. Eberstein et al., J. Cryst. Growth **143**, 71 (1994).

[112] J. Z. Xia et al., Biophys. J. **66**, 861 (1994).

[113] R. Giordano et al., Phys. Rev. A **43**, 6894 (1991).

[114] G.M. Kepler et al., Phys. Rev. Lett. 73, 356 (1994). Y. Han et al., Phys. Rev. Lett. **91**, 038302 (2003).

[115] K. Shukla, R. Rajagopalan, Colloids Surfaces A: Physicochem. Eng. Aspects **81**, 181 (1993).

[116] A. George, W. W. Wilson, Acta Cryst. **50**, 361 (1994).

[117] O. D. Velev et al., Biophys. J. **75**, 2683 (1998).

[118] M. Muschol et al., J. Chem. Phys. **103**, 10424 (1995).

[119] D. N. Petsev, et al., Phys. Rev. Lett. **84**, 1339 (2000).

[120] M. Sokuri et al., FEBS Lett. **295**, 84 (1991)

[121] N. Niimura et al., J. Crystal Growth **154**, 136 (1995)