Protein-Protein Interactions in Complex Cosolvent Solutions Revealed by Synchrotron Small-Angle X-ray Scattering

DISSERTATION

zur Erlangung des akademischen Grades
Doktor der Naturwissenschaften
(Dr. rer. Nat.)

vorgelegt von
M. Sc. M. Phil. Nadeem Javid
aus Lahore, Pakistan

eingereicht beim
Fachbereich Chemie
der Universität Dortmund

Dortmund 2007

Erster Gutachter:	Prof. Dr. Roland Winter
Zweiter Gutachter:	PD. Dr. Claus Czeslik
Dritter Prüfer:	Dr. Jens Müller
Tag der mündlichen Prüfung:	

To my dear parents

Table of Contents

Chapter 1

Introduction

Protein-protein interactions play an essential role in a series of biological situations, ranging from in vivo problems, where the thermodynamically stable folded state of the proteins and their structural and functional stability in a highly crowded cell environment are directed by intermolecular interactions, to in vitro studies, such as protein crystallization, aggregation and fibrillation. The delicate balance between the marginally stable proteins through van der Waals - forces, electrostatic interactions, hydrophobic interactions, hydrogen bonding and hydration, is the major driving force for such complex processes in nature. Hence, a detailed knowledge of the interactions between protein molecules in solution is instrumental to reach a molecular level understanding of protein function in nature and in technological processes involving proteins [Stradner et al., 2004; Malfois et al., 1996; Svergun and Koch, 2003 and Velve et al., 1998].

The essentiality of acquiring the intricate details about the protein-protein ineractions in order to understand the biological and chemical phenomenon are highlighted in many recent studies. For, instance, the onset of protein precipitation or crystallization as predominantly determined by the mechanism of molecular approach, reorientation and incorporation, which is governed by the strength of various intermolecular interactions, is pointed out by Durbin and Feher [1996]. Controlling and fine tuning of these interactions has been shown to be an effective tool in order to understand the mechanism and describing the conditions required for protein crystallization using a model protein, urate oxidase by Bonnete and coworkers [2001]. It has been proven that crystallization conditions can be obtained by changing the physiochemical parameters in order to induce less repulsive or more attractive interactions. In this study, interaction potentials were studied as a function of the main thermodynamic and chemical parameters: temperature, protein concentration, pH, salt nature and concentration as well as concentration of

polyols. Protein crystallization was observed even at slightly repulsive conditions (with a second virial coefficient B_{22} of about $+10^{-5}$ mol ml g^{-2} instead of – 8 to – 2 $\times 10^{-4}$ mol ml g^{-2}, which is the so-called crystallization window for proteins.

Stradner et al. [2004] have discovered a very different process of small equilibrium clusters formation by the combination of short-range attractive and long-range repulsive interactions in colloids and proteins. This finding is clearly relevant for a range of practically important phenomena including nucleation processes during protein crystallization (possibly the nucleation process in protein fibrillation as well), protein or DNA self-assembly and formation of cluster and gel phases in colloidal suspensions. It was speculated that a deliacte balance between these forces may lead to controlled self-assembly in insulin aggregate formation and the formation of almost mono-disperse mixed aggregates of polyelectrolytes (such as DNA). The driving force for self-assembly into clusters is short range attraction, which effectively acts as a surface tension leading to a decrease in surface energy upon aggregation. On the other hand, cluster growth is limited by the increasing electrostatic energy of the clusters, which counter balances the gain in surface energy. At low ionic strengths, the screeening length is larger or comaparable to the cluster size, so the balance between these forces provides a stabilizing mechanism against gelation and determines a finite aggregation number with a finite size.

The control and fine tuning of protein-protein interactions leads to the formation of protein clusters that eventually evolve into a structural arrested state [Baglioni et al., 2004]. In this work, the effect of pH and different anions of sodium salts on concentrated solutions of cytochrome C protein were investigated by means of small-angle neutron scattering (SANS) and viscocity measurements. The appearance of a low Q (low angle) peak in the SANS intensity distribution is accompanied by a strong increase in the relative viscocity and this phenomenon was considered as the signature of the gelation process promoted by specific co-ion interactions. This structural arrest was induced by salt addition and specifically depends on the nature of the anions, according to the Hofmeister series.

Intermolecular interactions of more or less partially unfolded proteins may undergo non-native self-assembly as a competing pathway to native functional folding, and are the first steps in the nucleation, aggregation and fibrillation process of proteins, which may lead

to a series of diseases, such as Alzheimer, Type-II-Diabetes Mellitus and Parkinson. The rapprochement and spatial reorientation of the insulin-molecules in solution are important steps in the formation of fibrils, which are directly affected by the balance of attractive and repulsive interactions and the resulting close-range order between the protein molecules.

Chi et al. [2003] have commented that in addition to the structural changes that occur during aggregation, protein molecules also assemble to form higher order aggregates. Molecular assembly processes occur as a result of attractive intermolecular interactions. Thus, an understanding of protein aggregation also requires information about the nature and magnitude of these interactions. Moreover, the assembly of protein molecules into non-native aggregates by definition involves the formation of higher molecular weight assemblies from initial lower molecular weight species. Thus, the same intermolecular interactions that govern protein crystallization and salting out are also expected to be important in the formation of non-native protein aggregates, such as fibrils.

Protein aggregation is inherently a nucleation and growth phenomenon where aggregates accumulate, eventually exceeding their solubility limit and precipitate. The existence of a lag phase in the aggregation of some proteins is caused by an energy barrier to assembly (or nucleation). The energy barrier results from the free energy required to create a new solid-liquid interface and depends on the aggregate size [Debenedetti, 1996 and Randolph and Larson, 1998] and so on the intermolecular interactions. When the size of the new phase is above a critical value where the energy barrier is highest, then growth of the nucleus occurs. Nucleation-dependent aggregation behavior is a result of a rate-limiting nucleation step, which is sensitive to the intermolecular interaction potential of the protein molecules. Furthermore, the barrier to assembly may be orientationally specific. If there is an orientation with a lower free energy to assemble, then the growth will occur preferentially in that orientation, resulting in an ordered aggregate morphology.

Intermolecular interactions among the eye lens proteins namely, γ-crystallins and β-crystallins (which in case of aggregation cause pathological disorder in eye called cataract) have been shown qualitatively by using X-ray scattering and osmotic pressure techniques by Tardieu et al. [1992], but the molecular origin and quantitative information regarding the interactions is not shown. Attractive interactions were observed with γ-

crystallins. In this case, the normalized X-ray scattering intensity near the origin increases with increasing protein concentration while the osmotic pressure increases less rapidly than in the ideal case. These attractive interactions were barely dependent on the ionic strength but very sensitive to temperature. However, in the β-crystallins a repulsive Coulombic interactions could be deduced from the observation that the normalized X-ray scattering intensity near the origin decreases with increasing protein concentrations while the osmotic pressure increases much more rapidly than in the ideal case. These interactions depend upon the ionic strength but are hardly affected by temperature.

The physics underlying these processes remains still poorly understood and a growing body of evidence suggests that formation of ordered protein aggregates, the so-called amyloids, is a common, generic feature of proteins as polymers [Chiti and Dobson, 2006; Dzwolak et al., 2004 and Coen et al., 1995].

Cosolvents and cosolutes have a significant influence on the function and stability of proteins. Chaotropic agents such as urea destabilize proteins, while kosmotropic agents such as polyols and sugars stabilize them. Moreover, they also change the interactions between the proteins and in drastic cases lead to effects such as molecular crowding and aggregation. Studies of the influence of cosolvents and cosolutes on the interaction and short-range order in protein solutions play a crucial role in understanding physiological and biochemical processes because biological media such as the cytoplasm are far from being an ideally diluted solution [Back et al., 1979; 10-14; Bennoin and Dagget, 2003; Timasheff and Xie, 2003; Minton, 2001; Munishkina et al., 2004 a; Munishkina et al., 2004 b; Grudzielanek et al., 2005 and Grudzielanek et al., 2006]

Valuable information about the intermolecular interaction potential between proteins in solution can be obtained using the small-angle X-ray scattering technique. In combination with liquid-state theoretical approaches and performing concentration dependent studies, the intermolecular scattering contribution (structure factor, $S(Q)$), which sensitively depends on the interaction potential, can be separated from the intra-particle scattering contribution (the form factor, $P(Q)$) of the molecule. Several works have been carried out to model the pair potentials between protein molecules on the basis of small-angle X-ray scattering (SAXS) data. Using liquid state theories and potential models, it is possible to retrieve the pair potential $V(r)$ between the dissolved protein molecules and to study the

influence of protein concentration, cosolvents and cosolutes. A successful approach is provided by modeling the potential as a sum of a repulsive and an attractive part, in which the former is expressed as a screened Coulomb-term and the latter as attractive Yukawa-potential. Tardieu et al. [1999] demonstrated the usefulness of such potential models for analysing experimental SAXS data as a function of temperature, pH and varying salt concentration, but also highlighted problems at high ionic strength. Niebuhr and Koch [2005] examined the effects of salts as well as of a chaotropic (urea) and kosmotropic (trimethyl-N-oxide) agent. Narayanan and Liu [2003] studied under- and supersaturated lysozyme solutions and applied a Derjaguin-Landau-Verwey-Overbeek (DLVO) potential within the random-phase-approximation (RPA) to fit the measured scattering curves. Using this approach, which accounts for attractive as well as for repulsive forces, analytical solutions for extracting the intermolecular pair potential $V(r)$ are available which are able to successfully describe the measured data for low to medium high ionic strengths [Tardieu et al., 1999; Narayanan and Liu, 2003 and Niebuhr and Koch, 2005].

The aim of our study was to use this approach for studying the intermolecular interaction potential $V(r)$ of a typical monomeric protein, lysozyme, for various kosmotropic and chaotropic cosolvents and cosolutes of different nature as a function of protein concentration and ionic strength. Furthermore, this approach has not been applied to amyloidogenic proteins like insulin. Knowledge of $V(r)$ can contribute to the understanding of the early stages of fibril formation. The probability for the formation of an aggregation seed is directly affected by the interaction potential between the individual insulin molecules. It has been found that the first aggregation seed can determine the structures formed in the following aggregation process – and thus retains information under which circumstances they were formed [Dzwolak, 2005]. This essential finding underlines the importance of the knowledge of interaction potentials among the protein molecules, which will lead to new insights into the forces which govern protein-protein interactions in complex cosolvent mixtures. Also, protein-protein interactions at various concentrations have a drastic effects on the activity coefficients, which are important in thermodynamic calculations of protein solutions already at rather low concentrations.

1.1 Model systems

1.1.1 Lysozyme: a globular protein

Lysozyme is an enzyme (molecular weight 14.6 kD) abundant in egg white and human tears that catalyzes the hydrolytic cleavage of polysaccharides in protective cell walls of some families of bacteria. Lysozyme, because it can lyse, or degrade, bacterial cell walls, serves as a bacterial agent. As in cytochrome *c*, about 40 % of its amino acid residues are in α-helical segments, but the arrangement is different and some ß-sheet structure is also present. Four disulphide bonds contribute to the stability of this structure. The α-helices line a

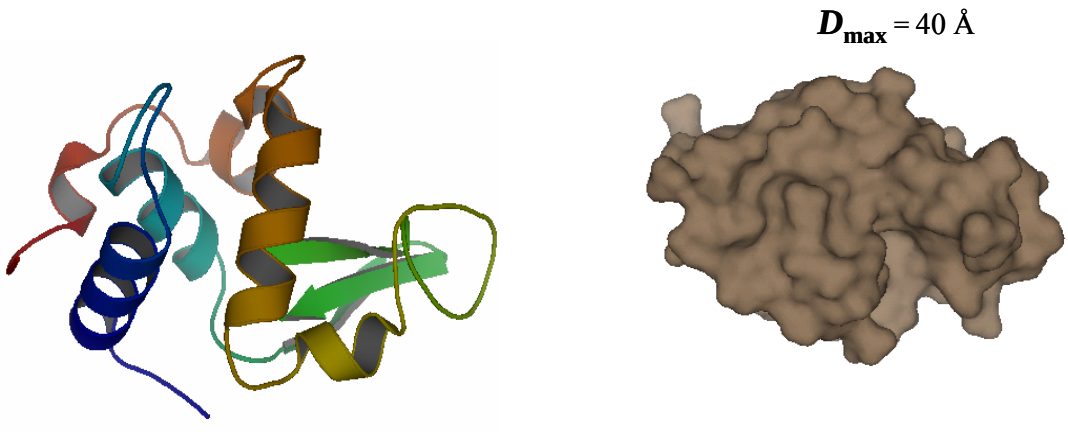


Figure 1.1: Ribbon diagram (left) and surface diagram (right) of the globular lysozyme molecule (PDB entry: 1hel)

long crevice in the side of the molecule, called the active site of substrate binding and catalysis. The bacterial polysaccharide that is the substrate for lysozyme fits into this crevice [Nelson and Koch, 2000].

Lysozyme has radius of gyration of 15.4 ±0.2 Å [Svergun et al., 1998] and has a net positive charge of value ~10 at pH 4.6 [Tanford and Roxby, 1972]. Being amphoteric, individual protein molecules in solution are in equilibrium with H^+ ions and the net surface charge carried by a protein macro-ion is largely determined by the pH of the

solution. In the lysozyme molecule, the charged groups are symmetrically distributed on the surface [Barlow and Thornton, 1986]. Tanford and Roxby [1972] have determined the average charge on a lysozyme macro-ion in aqueous solution as a function of pH from hydrogen ion titration. Lysozyme macro-ions in solution are known to readily complex Cl^- ions [Steinhardt and Renolds, 1969]. The anion adsorption can lower the net charge on the protein which is estimated from hydrogen ion titration for pH < pI.

Among globular proteins, lysozyme has become the prototype for scientific investigations concerning the questions of protein folding, structural stability, aggregation, cluster formation and protein-protein interactions tuning with kosmotropic and chaotropic cosolvents.

Narayanan and Liu [2003] have studied the lysozyme self-interactions in undersaturated and supersaturated salt solutions and related the interaction energies with crystallization behavior of the proteins by the osmotic second virial coefficient, B_{22} and underlined the need for the hydration and other specific forces involved in understanding protein–protein interactions.

Niebuhr and Koch [2005] examined the effect of two physiological cosolutes (urea and trimethylamine-N-oxide) and of KCl on the intermolecular interactions in concentrated lysozyme solutions by synchrotron radiation small-angle x-ray scattering. The interactions were correlated to preferential binding and preferential exclusion of the cosolvent to protein molecules via changing hydration.

Self-interaction chromatography was applied to measure cosolvent induced protein self-interactions by Valente et al. [2005] and calculate second virial coefficients in the presence of polyols and amino acids. All of these cosolvents, alone or in combination, indicated a reduction in intermolecular attractions among the protein molecules (lysozyme).

Ducruix et al. [1996] showed that lysozyme in a low ionic strength buffer presents repulsive protein-protein interactions. Addition of increasing concentrations of salts gradually leads from repulsive to attractive interactions demonstrating the ability of a given protein to change its interactive behavior with additives. While cations (Li^+, Na^+, K^+, NH_4^+, Cs^+) all showed similar effects, large differences were observed between anions in their efficiency to modify the interaction potentials. The order of the anions

(SCN^-, paratoluene sulfonate, NO_3^-, CI^-, $H_2PO_4^-$) was found to be the same as observed for their effectiveness in reducing lysozyme solubility and inducing crystallization.

1.1.2 Insulin: an amyloidogenic protein

Insulin is a small globular protein hormone with a molecular weight of 5.7 kD which is produced in vivo within the pancreatic ß-cells of the islets of Langerhans. It is found not only in all vertebrates including fish but also in starfish, worms and protozoa [Falbe and Regtiz, 1995]. Insulin is the primary hormone responsible for controlling the cellular uptake, utilization and storage of glucose, amino acids and fatty acids while inhibiting the breakdown of glycogen, protein and fat. Together with its antagonist glucagon, insulin keeps the blood sugar level constant within certain physiological limits. Stimulating the glucose intake, especially in muscle-tissue and liver by binding to specific transmembrane receptors, and triggering a complex signalling pathway that leads to a higher glucose cell membrane permeability, insulin is responsible for the blood sugar drop, while in turn glucagon has the opposite effect [Löffler and Petrides, 1990 and Lee and Pilch, 1994].

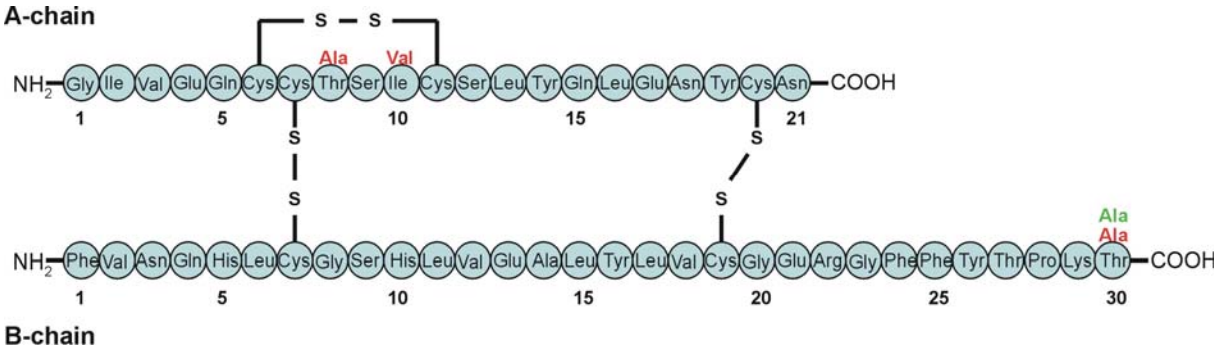


Figure 1.2: Primary structure of human insulin. The colored superscripts indicate those amino acid residues which are different in bovine (red) and porcine (green) insulin. Adapted from [Nielsen et al., 2001b].

The insulin molecule contains 51 amino acids; it is made up of two peptide chains linked by disulphide bonds shown schematically in Fig. 1.2. Although it is active as a monomer, during its biosynthesis and storage it assembles to a dimer and, in the presence of zinc to

hexamers [Derewenda et al., 1989]. The A chain consists of two helical segments (A 2-8 and A 13-20) linked by a turn sequence. The B chain starts with a β-strand-like structure (B 1-6), runs into a long helical strech (B 9-19), then into a turn, and finally into a short strand of β-structure (B 24-29). The two chains are linked by disulphide bonds and stabilized by tight packing of hydrophobic side-chains at the interfaces of structural elements.

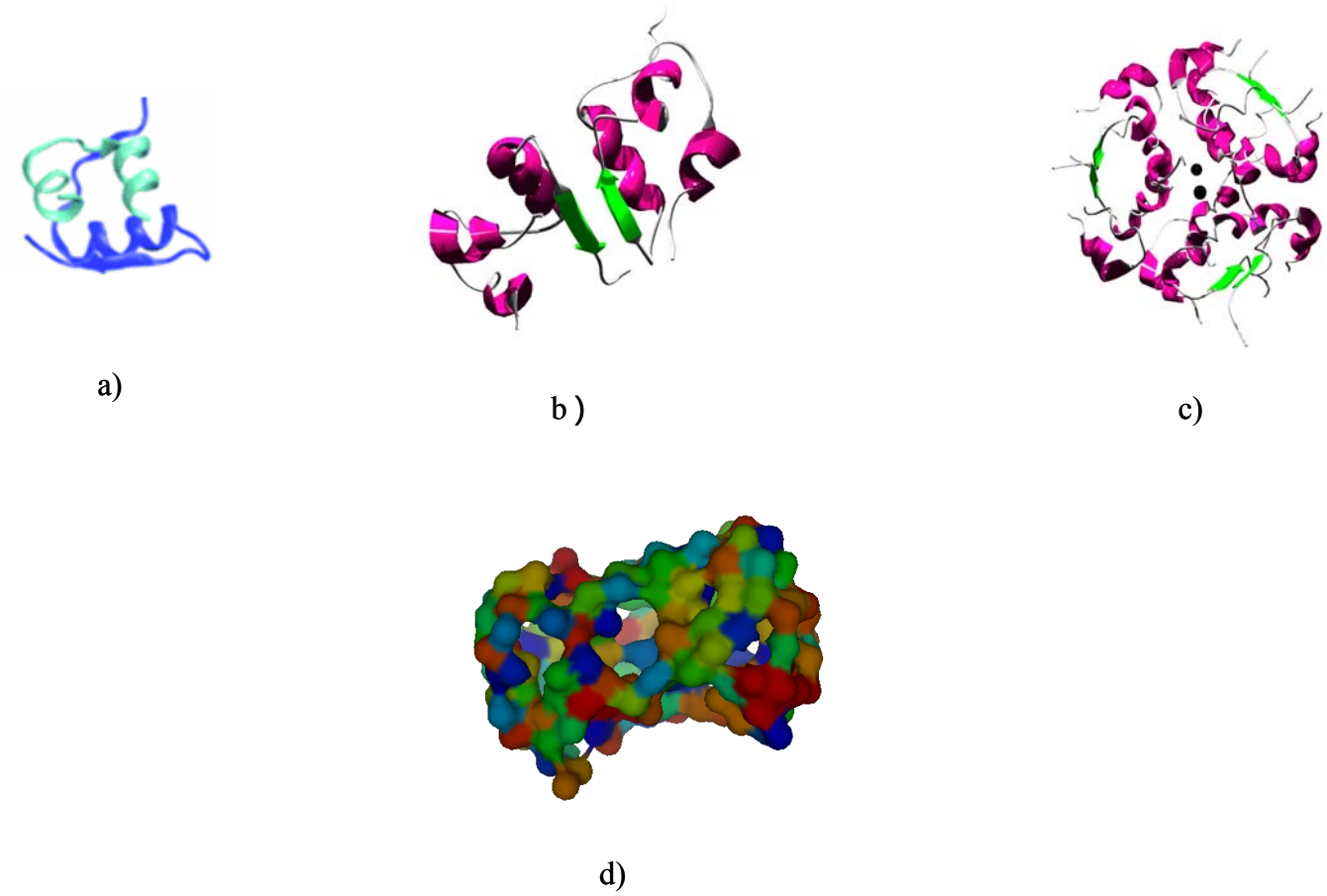


Figure 1.3: Ribbon diagram of an insulin monomer (a), dimer at pH = 2 (b), which is facilitated by antiparallel β-strand formation between residues B24-B28 of the two monomers, and van der Waals contacts between hydrophobic residues of the B chain α-helix and the β-strand [Whittingham et al., 2002]. The hexamer solution structure at pH = 7 (c) contains two zinc ions at the center interacting with the B10 His side chains (R_6 state). This figure was created by POVRAY (version 3.6) [POVRAY], based on the records 1GUJ (dimer) [Whittingham et al., 2002] and 1AIY (hexamer) [Chang et al., 1997], respectively, of the Brookhaven National Laboratory's Protein Data Base [PDB] and tetrameric solution surface structure at pH = 2 (d) in the presence of 100 mM sodium chloride , calculated by GASBOR [Svergun et al., 2001].

Among the proteins which are prone to aggregation and fibrillation, insulin has become a model system due to its medical importance and the wealth of available data. Insulin, which forms hexamers under physiological conditions, aggregates under conditions like low pH, high temperatures or stirring in the presence of hydrophobic surfaces. At pH 2, insulin forms dimers in water, and the addition of particular cosolvents, such as 20 wt% ethanol, does not only render insulin monomeric at pH 2, but also affects the aggregation kinetics. The aggregation and fibril formation of this protein has been investigated with various methods like FT-IR and fluorescence spectroscopy as well as atomic force microscopy (AFM), but studies yielding information on the spatial organisation and intermolecular forces of the molecules in the pre-aggregated state in solution are lacking [Dzwolak, 2006; Whittingham et al., 2002; Sluzky et al., 1991; Ahmad et al., 2004; Podesta et al., 2006; Dzwolak et al., 2005 and Jansen et al., 2005]

1.2 Cosolvents, cosolutes and osmolytes

The biochemical milieu, as well as some solutions used in in vitro experiments involving macromolecules like proteins and nucleic acids, contains additional chemical species such as buffer compounds, simple inorganic salts, amino acids, sugars, urea, polyols, and the like. These species commonly are called cosolvents, cosolutes or osmolytes to indicate that they are present in significant excess to the macromolecule, for instance proteins in solution, as well as to suggest that they do not necessarily bind stoichiometrically to the macromolecule in a mass action sense. Even in the absence of specific stoichiometric binding, however, these cosolvents interact indirectly with macromolecules by virtue of their sharing the same solute environment [Völker and Breslauer, 2005].

Cosolvents, cosolutes and osmolytes influence the proteins in a variety of ways including structural stability, activity, solubility, crowding and aggregation through inter and intra-protein interactions tuning. Cosolvency has been widely used in processes involving proteins, like in vitro crystallization, refolding, dynamically arrested states and as excipients in drug formulations. All the above-mentioned processes rely upon the

mechanism of delicate balancing of protein interactions, leading to provide molecular switches to manipulate the functions in desired directions.

1.2.1 Classification of cosolvents

Generally, cosolvents are classified on the basis of the mechanism of action on the protein molecules or depending upon the ordering/disordering of the water structure. According to their action on the protein, cosolvents are classified as (i) compatible cosolvents, which do not disturb protein functionality, and (ii) non-compatible cosolvents, e.g. denaturants like urea, guanidinium hydrochloride, alcohols, etc. They induce, at sufficiently high concentration, a disruption of the protein structure.

A special class of compatible cosolvents is the so-called compensatory solvents, which stabilize the folded form against denaturation under external stress (examples are sugars, polyols, monomeric amino acids and methylamines).

Osmolytes are compounds which are able to stabilize cells in vivo against dehydrating stress (e.g. salinity) and/or volume changes by maintaining an osmotic equilibrium.

The cosolvents, which are ordering or disordering the water structure are termed as kosmotropes (order - makers) and chaotropes (disorder - makers) respectively [Scharnagal et al., 2005].

The most important kosmotropes are non-reducing sugars like trehalose, the disaccharide sucrose, polyols like glycerol, amino acids like proline and trimethylamine *N*-oxide which vary in their strength of structuring the water shell (hydration shell) around the macromolecules like proteins. For instance, trehalose is particularly effective at stabilizing macromolecules as it has a larger hydrated volume (2.5 times that of sucrose) [Sola-Penna and Meyer-Fernandez, 1998] and sucrose enhances preferential hydration more than glycerol [Mitra et al., 2006]. The well-known chaotropes are large singly charged ions, with low charge density [e.g. SCN^-, $H_2PO_4^-$, HSO_4^-, $(CH_3)_4N^+$ (tetramethylammonium) and $(NH_2)_3C^+$ (guanidinium), urea and alcohols which exhibit a non-monotonic concentration dependence [Mitra et al., 2006].

1.2.2 Mechanisms of cosolvent interactions

As to the mechanism of the cosolvent-protein interaction, there is still considerable discussion going on and the possible mechanisms can be categorized as following [Scharnagal et al., 2005]:

(1) direct contact interaction of the cosolvent molecule with the protein
(2) indirect effects via the perturbation of the hydration layer, including also the perturbation of internal water via some protein specific interaction channels
(3) combination of (1) and (2), a disruption of the water structure in the hydration shell, so that water molecules are released and enable a direct interaction of cosolvent molecules with protein groups.

1.2.3 Preferential hydration

Polar groups on the surface of proteins have tightly bound water molecules in aqueous solutions, generally known as water of hydration and if this hydration is maintained in a concentrated solution of solvent additives (cosolvent, cosolute or osmolyte), a difference in the concentration of the additives develops between the bulk solution and the vicinity of the protein resulting in preferential hydration (also referred as preferential exclusion) such that excess water accumulates near the protein.

The driving force for preferential hydration is the perturbation of the chemical potential of the cosolvent by the protein molecules. The measured change of the amount of water in contact with protein during the course of the reaction modulated by the cosolvent is a change in preferential hydration that is strictly a measure of the cosolvent chemical potential perturbation by the protein in the ternary water-protein-cosolvent system. It is not equal to the change in water of hydration, because water of hydration is a reflection strictly of protein-water forces in a binary system and there is no direct relationship between the water of preferential hydration and water of hydration [Timasheff, 2002].

Preferential hydrations observed for protein stabilizing cosolutes such as certain salts, amino acids, sugars, polyols and methylamines are typically in the range of 0.2-0.4 g per

gram of protein, indicating that water binding is the source of excess water (and cosolute deficiency) near the protein surface in concentrated solutions of these cosolutes [Arakaw, 2002].

1.2.4 Preferential binding

Preferential binding is the adhering of the additives (cosolvent, cosolute or osmolyte) with the protein molecule at a level at which the concentration of the additive in the vicinity of the protein exceeds the concentration of the additive in bulk solution or when the affinity of the additive or ligand towards the protein molecules in aqueous solution is greater than for water. If there is an excess of additive or ligand in the protein domain relative to the bulk solvent composition, it is described as preferential binding [Arakaw, 2002 and Timasheff, 2002].

The principal protein denaturants are urea and guanidine hydrochloride, which induce a random coil state in proteins [Tanford et al., 1967 and Tanford, 1964]. Sodium dodecyl sulfate, alcohols, and some other organic solvents induce a transition into structures rich in α-helices [Tanford et al., 1960 and Reynolds and Tanford, 1970].

More subtle effects are observed in the case of monoalcohols, such as ethanol, and their fluorinated derivatives, such as 2,2,2-trifluoroethanol (TFE), which is assumed to be particularly potent in inducing structural conversions within proteins, likely due to an additive effect of fluorination [Young et al., 1994]. These properties give rise to a series of proposed interaction mechanisms, including perturbation of the protein's water shell, [Buck, 1998 and Walgers et al., 1998] diminution of hydrophobic interactions,[Kentsis and Sosnick, 1998] strengthening of intra-protein hydrogen bonding and less shielding of electrostatic interactions due to the lower polarity of alcohols [Young et al., 1994].

1.3 Small-angle X-ray scattering of biological macromolecules

Small-angle X-ray scattering is a fundamental tool in the study of biological macromolecules providing structural information of particles in near physiological

conditions in monodisperse systems of ideally diluted solutions. Moreover, it provides the information about the structural ordering in solutions along with the knowledge of the spatial distribution of clusters of macromolecules in medium to highly concentrated solutions. The highly crowded environment of cells can be mimicked by employing medium to high concentrations of proteins with added cosolvents, cosolutes and osmolytes by varying the concentration and can be measured with X-ray scattering experiments to deduce a range of information which is useful to explore biologically relevant phenomena like protein- protein interactions, crystallization and folding of proteins.

Scattering particles like proteins can be described as a continuous function of the electron density (ρ) as determined by the chemical composition of the object. If there are no strong variations of electron density over the molecule (e.g., a native protein), a mean value is assumed. In dilute solution, the scattering objects can be considered to be isolated (intraparticle scattering or form factor, $P(Q)$), whereas in a concentrated solution, the interaction between the particles, displaying spatial correlations, also contributes to the scattering intensity (interparticle scattering or structure factor, $S(Q)$).

In the following section, theoretical aspects as relevant for dilute systems are discussed and the nomenclature described by Svergun and Koch [2003] is used mainly.

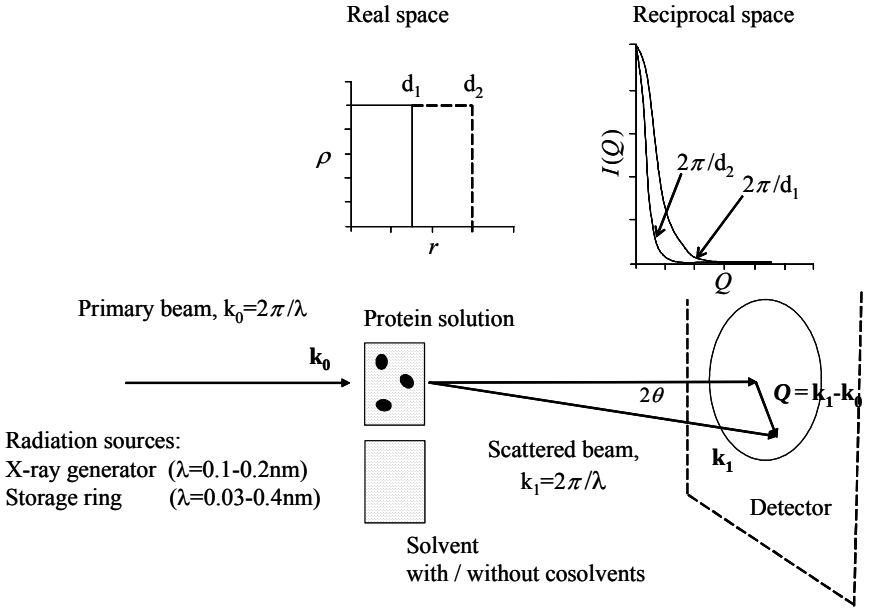


Figure 1.4: Schematic representation of the scattering process modified from Svergun and Koch [2003].

When a protein solution is irradiated by a monochromatic wave with wave vector $k_0 = |\boldsymbol{k}_0| = 2\pi/\lambda$, only its electrons will scatter the incident beam. In case of elastic scattering (i.e., without energy transfer between photons and electrons), the modulus of the scattered wave vector $k_1 = |\boldsymbol{k}_1|$ is equal to k_0 and the scattered vector can be defined as $\boldsymbol{Q} = \boldsymbol{k_1}\text{-}\boldsymbol{k_0}$, with $Q = 4\pi\lambda^{-1}\sin(\theta)$ and 2θ is the scattering angle. The amplitude of the scattered wave is described as

$$A(\boldsymbol{Q}) = \int_v \Delta\rho(\boldsymbol{r}) \exp(i\boldsymbol{Q}\boldsymbol{r}) d\boldsymbol{r} \qquad (1.1)$$

Where, $\Delta\rho(r)$ is the excess scattering length density given by $\Delta\rho(\mathrm{r}) = \rho(r) - \rho_s$, where $\rho(r)$ is the scattering length density of the protein and ρ_s is scattering length density of the solvent (buffer with cosolvents). In SAXS experiments the detector measures only the scattered intensity which is, $I(\boldsymbol{Q}) = A(\boldsymbol{Q})A^*(\boldsymbol{Q})$ and proportional to the number of photons scattered in the given direction $\boldsymbol{Q}$. The intensity from the entire ensemble is a continuous isotropic function and is proportional to the scattering from a single particle averaged over all orientatations Ω, $I(Q) = \langle I(\boldsymbol{Q}) \rangle_\Omega$.

1.3.1 Size and shape parameter evaluation

The scattered intensity can be described as,

$$I(Q) = \langle A(\boldsymbol{Q}) * A(\boldsymbol{Q}) \rangle_\Omega = \left\langle \int_v \int_v \Delta\rho(\boldsymbol{r}) \Delta\rho(r') \exp\{i\boldsymbol{Q}(\boldsymbol{r} - \boldsymbol{r}')\} d\boldsymbol{r} d\boldsymbol{r}' \right\rangle_\Omega \qquad (1.2)$$

Equation 1.2 can be simplified by integrating Eq. 1.2 in spherical coordinates. So,

$$I(Q) = 4\pi \int_0^{D_{\max}} r^2 \gamma(r) \frac{\sin Qr}{Qr} dr \qquad (1.3)$$

Where $\gamma(r)$, is the spherically averaged autocorrelation function of the excess electron density,

$$\gamma(r) = \left\langle \int \Delta\rho(\boldsymbol{u}) \Delta\rho(\boldsymbol{u} + \boldsymbol{r}) d\boldsymbol{u} \right\rangle_\omega \qquad (1.4)$$

The particle shape is evaluated by the function denoted as pair distribution function, $p(r)$. The $p(r)$ function actually gives the distribution of distances between the scattering points (electrons in amino aids) inside the particle (*e.g.*, in one lysozyme molecule). It is described as $p(r) = r^2\gamma(r)$. Inverse Fourier transformation gives the function, $p(r)$

$$p(r) = \frac{r^2}{2\pi^2}\int_0^\infty Q^2 I(Q)\frac{\sin Qr}{Qr}dr \qquad (1.5)$$

The size of the particle can be estimated by using Guinier's approximation,

$$I(Q) = I(0)\left[1-\tfrac{1}{3}R_g{}^2Q^2 + \mathrm{O}(Q^4)\right] \cong I(0)\exp\left(-\tfrac{1}{3}R_g{}^2Q^2\right) \qquad (1.6)$$

where the forward scattering $I(0)$ is proportional to the squared total excess scattering length density of the particle

$$I(0) = \int_v \int_v \Delta\rho(\mathbf{r})\Delta\rho(\mathbf{r}')d\mathbf{r}d\mathbf{r}' = 4\pi\int_0^{D_{\max}} p(r)dr = (\Delta\rho)^2 V^2 \qquad (1.7)$$

and the radius of gyration R_g is the normalized second moment of the distance distribution of the particle around the centre of its scattering length density distribution

$$R_g{}^2 = \int_0^{D_{\max}} r^2 p(r)dr\left[2\int_0^{D_{\max}} p(r)dr\right]^{-1} \qquad (1.8)$$

Equation (1.6), derived by Guinier [Guinier, 1939], is widely used in SAXS experiments to evaluate the size of the particles. For ideal monodisperse systems, the Guinier plot ($\ln(I(Q))$ versus Q^2) is a linear function, whose intercept gives $I(0)$ and the slope yields the radius of gyration R_g. Guinier's approximation is valid for very small-angles only, in the range $Q < 1.3/R_g$.

If the particles in solution are randomly oriented but also interact (non-ideal semi-dilute solutions), local correlations between the neighboring particles must be taken into account. The scattering intensity from the ensemble will be isotropic and for spherical particles can be written as $I(Q) = P(Q)\times S(Q)$, where $P(Q)$ and $S(Q)$ are the form factor (which depends upon the shape and size of the particle) and the structure factor (interference term due to particle interactions), respectively. In biological applications, small-angle X-ray scattering is used to analyze the structure of dissolved macromolecules (based on particle scattering) as well as the interactions based on the interference term.

Separation of the two terms for semi-dilute solutions is possible by using measurements at different concentrations and/or in different solvent conditions (pH, ionic strength, etc.).

1.3.2 Distance distribution function, *p*(*r*)

The scattering intensity curve is a reciprocal representation of the particle shape while the distance distribution function is real space representation of particle's shape which is more intuitive and straightforward to understand. Figure 1.5 presents typical scattering patterns and distance distribution functions of geometrical bodies with the same maximum size. Globular particles display bell-shaped $p(r)$ functions with a maximum at $D_{max}/2$. Elongated particles have skewed distributions with a clear maximum at small distances corresponding to the radius of the cross section. Flattened particles display a rather broad maximum, shifted to distance smaller than $D_{max}/2$. A maximum shifted towards distances larger than $D_{max}/2$ is usually indicative of hollow particle. Particles consisting of well-separated subunits may display multiple maxima, the first corresponding to the intra-subunit distances, the others yielding the separation between the subunits.

Distance distribution functions can be calculated by using the program GNOM [Svergun, 1992]. In a first step, the approximate value of the particle size is given as input, which is calculated from Guinier's analysis and then the program calculates the $p(r)$ function. The particle shape can be simulated by using simulated annealing methods, for example using the program GASBOR [Svergun, 2001].

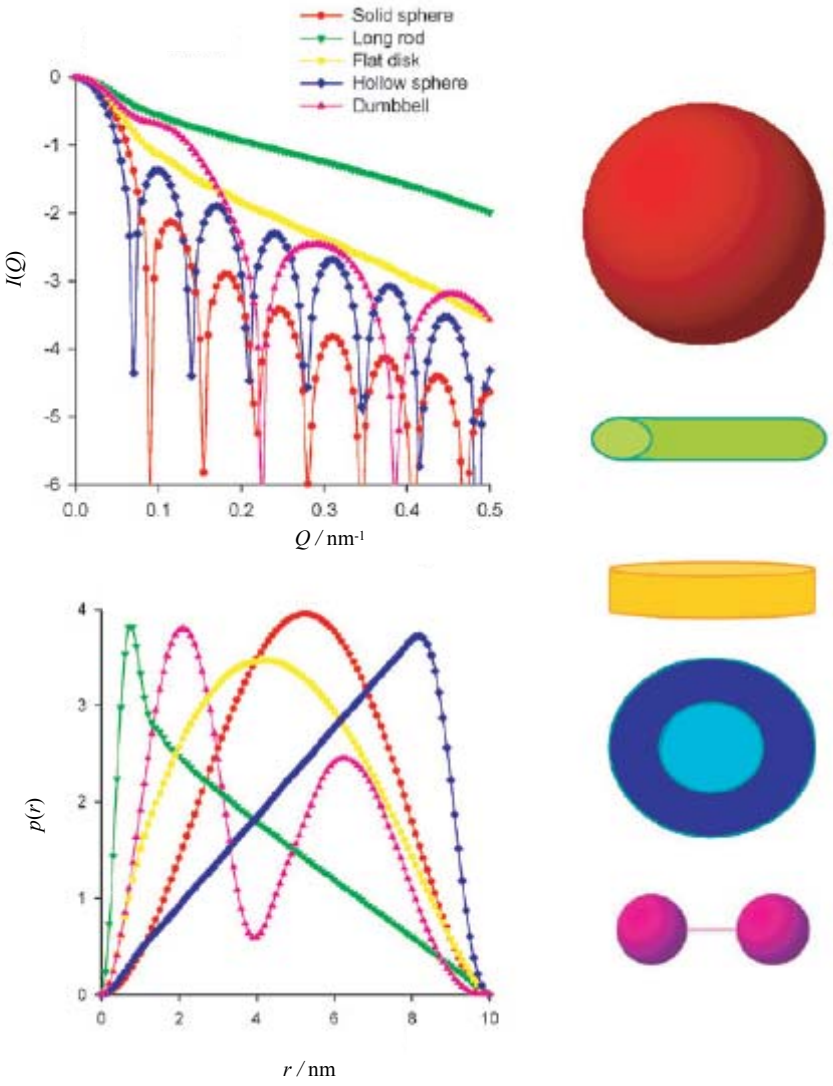


Figure 1.5: Scattering intensities and distance distribution functions of geometrical bodies as adapted from [Svergun and Koch, 2003].

1.4 Protein-protein interaction potentials with small-angle X-ray scattering

A major method for studying interaction potentials of proteins is provided by fitting small-angle neutron or X-ray scattering curves with appropriate potential models within an appropriate liquid-state theoretical approach.

As mentioned in the previous section, the intensity scattered by one particle as a function of momentum transfer Q is usually called particle form factor $P(Q)$, which is the spherically averaged auto-correlation function of the electron density contrast associated with the particle (which in our case is the protein molecule). When the solution is ideal, i.e. in the absence of interactions and where the distribution of proteins within the sample is completely random, the total scattering is the sum of the scattering of the individual particles [Luzzati and Tardieu, 1980]. In reality, the distribution will never be completely random due to the repulsive interactions between the overlapping electron orbitals (hard-sphere interactions). If the relative positions of the proteins are not random due to their interactions, it means that there is a certain order resulting in interference of the scattered waves, which have an impact on the total scattering pattern. This interference term in the presence of interactions, accounting for the departure from ideality, is represented by the so-called structure factor $S(Q)$.

Hence, the scattered intensity of the solution in concentration dependent experiments can be represented as a convolution of the scattered intensity of a single particle averaged over all orientations and an interference term. So, $I(Q,n) = n\ \ P(Q) * S(Q,n)$, where n is the particle number density. The equation can still be considered valid, yet within a smaller Q-range, with quasi-spherical particles and/or polydisperse particles and deviations at high Q-values are expected [Tardieu, 1994].

To extract the experimental structure factor $S_{exp}(Q)$, a form factor is needed. The form factor may be obtained from scattered intensity patterns recorded at low concentrations to avoid the interaction effects and interference term; here the structure factor is 1 for the whole Q-range. Experimental structure factor $S_{exp}(Q)$ is obtained by carrying out concentration dependent experiments and described mathematically in the following Section 1.4.2. The relative distribution of the protein particles under different conditions

of interaction (random distribution, repulsive interactions and attractive interactions), their corresponding scattering intensity pattern and structure factors are shown schematically in Fig 1.6.

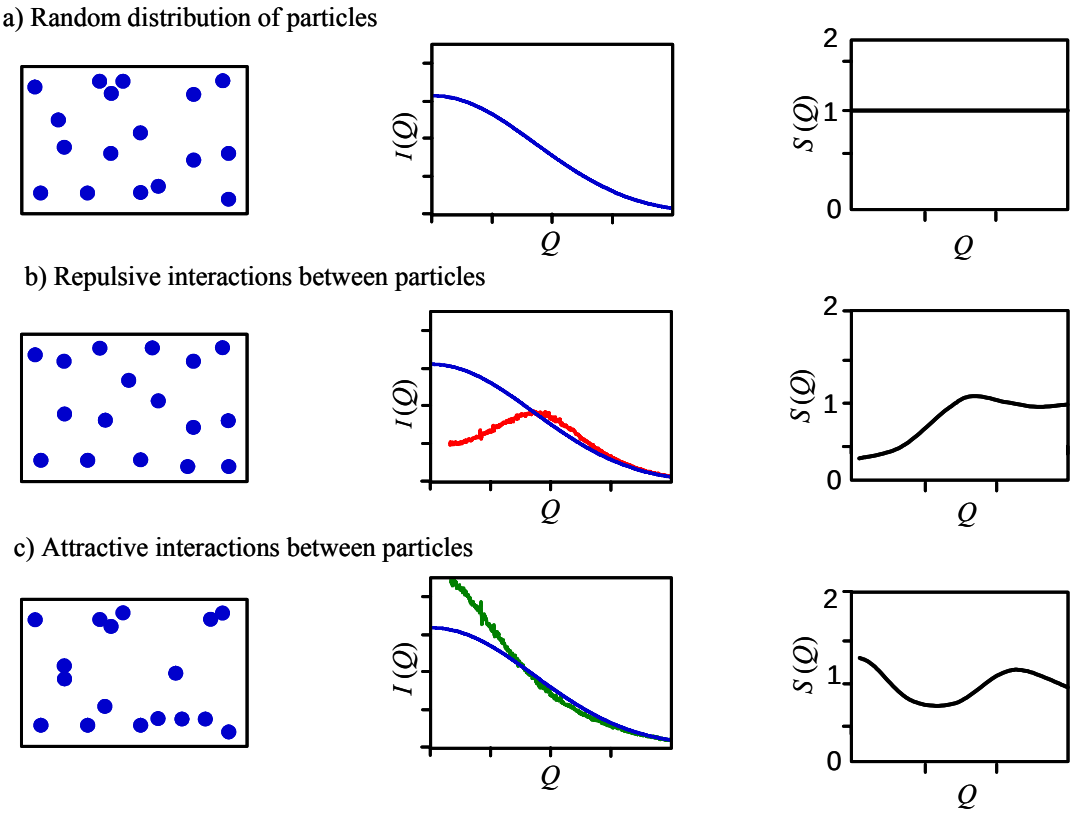


Figure 1.6: Effect of the interactions between proteins on the SAXS intensity pattern $I(Q)$ and structure factors $S(Q)$. a) If the distribution of proteins within the sample is completely random, the scattering pattern will be identical to that of the isolated particle (blue curve). b) In case of repulsive interactions, a strong correlation peak will be observed (red curve) and c) for attractive interactions, a drastic increase in the intensity at $Q \rightarrow 0$ will be observed (green curve).

1.4.1 Numerical treatment and potential models

In aqueous solutions, interactions include inter-particle interactions as well as particle-solvent, particle-ion (or small solute) and solvent-solvent interactions [Hansen and McDonald, 1986 and Israelachvili, 1992]. In a simplified picture, only the interaction

potentials between macromolecules, which interact through the solvent and ions, are explicitly considered [Belloni, 1991]. The solvent is treated as a medium of uniform dielectric constant and the ions as point charges. The solution can be described mathematically as the convolution product of the form factor $P(Q)$ and the structure factor $S(Q)$, which is related to the radial pair distribution function $g(r)$. This function is a measure of the probability of finding another molecule with its centre at distance r from the referrence one, as shown in Fig. 1.7. Mathematically, it is defined as

$$S(Q,n) = 1 + n\int_0^\infty 4\pi^2 r^2 \left(g(r) - 1\right)\frac{\sin(Qr)}{Qr}dr \qquad (1.9)$$

For a hypothetical ideal gas $S(Q)$, or $g(r)$ =1 over the entire Q - or r-range. In a real gas, the molecules cannot interpenetrate and thus the probability of finding two particles of radius R (and diameter σ) within a distance of $r < \sigma$ is zero. In gases and liquids only short range order exists and thus $S(Q)$ or $g(r)$ differs from 1 for small Q - or r-values. For repulsive interactions, the particles are evenly distributed and the structure factor in low Q-range approaching zero is lower than 1. For attractive interactions, fluctuations in the particle distribution are observed, which lead to a structure factor value larger than 1 in the low Q-range [Tardieu et al., 1999] as shown in Figure 1.6.

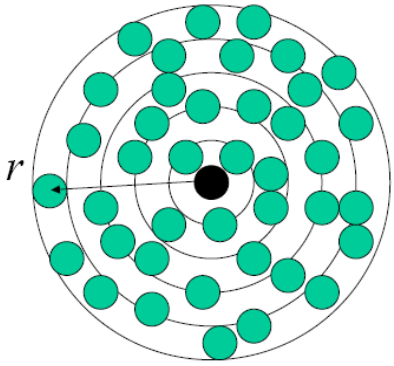


Figure 1.7: The radial pair distribution function $g(r)$ is a measure of the probability to find another molecule with its centre at a distance r from the reference one in the origin.

1.4.2 Structure factor calculation from liquid state theory

One strategy to calculate $S(Q)$ for a given interaction potential is to solve the corresponding Ornstein-Zernike equation with appropriate closure relations in an iterative numerical procedure [Hansen and McDonald, 1986 and Belloni, 1993]. An analytical solution is available by using the so-called random phase approximation (RPA) with a reference system of hard spheres as introduced by Grimson, in which further interactions are treated as a perturbation and given in the form of DLVO potential, i.e., a sum of attractive and repulsive contributions [Grimson, 1983].

The pair wise interaction potentials $V(r)$ are splitted into i) a reference part, $V_0(r)$,

$$V_0(r) = \infty \quad \text{for } r < \sigma$$
$$= 0 \quad \text{for } r > \sigma \qquad (1.10)$$

Where r is the center-to-center separation of the particles and ii) a perturbation part, $V_1(r)$, thus yielding the structure factor $S(Q)$, under RPA by

$$S(Q) = S_0(Q)\left[1 + n\beta S_0(Q)V_1(Q)\right]^{-1}, \qquad (1.11)$$

with β being $(k_B T)^{-1}$, $S_0(Q)$ is the structure factor of the reference system, n the particle number density, and $V_1(Q)$ the Fourier-transform of the perturbation potential $V_1(r)$. $S_0(Q)$ can be calculated according to the empty core model [Kelkar, 1992]:

$$S_0(Q)^{-1} = 1 - 12\eta \frac{(\eta(3-\eta^2)-2)}{(1-\eta)^4} \frac{j_1(Q\sigma)}{Q\sigma}, \qquad (1.12)$$

where η is the volume fraction, σ the hard-sphere diameter of the particles, and $j_1(Q\sigma)$ the first order spherical Bessel-function. The potential $V_1(r)$ is designed as a sum of two potentials, a repulsive Coulomb-potential $V_C(r)$ and an attractive Yukawa-type potential $V_Y(r)$:

$$V_{\mathrm{C}}(r) = \frac{Z^2 \mathrm{e}^2}{\varepsilon(1+0.5\kappa\sigma)^2} \cdot \frac{\exp[-\kappa(r-\sigma)]}{r}, \qquad (1.13)$$

$$V_{\mathrm{Y}}(r) = -J\left(\frac{\sigma}{r}\right)\exp\left[\frac{-(r-\sigma)}{d}\right]. \qquad (1.14)$$

Z is the effective charge of the molecule, e the elementary charge, ε is the dielectric permittivity of the medium, and κ the reciprocal Debye-Hückel screening length defined as :

$$\kappa^2 = \frac{4\pi \mathrm{e}^2}{\varepsilon\, \mathrm{k_B} T} \sum_i n_i Z_i^2 \qquad (1.15)$$

with n_i being the mean density of ions i in solution with charge Z_i. The Yukawa-potential contains the fitting parameters J and d, where J is the modulus of the depth and d the range of the attractive potential. For calculation of $S(Q)$, both potentials have to be Fourier-transformed. For the Coulomb-potential an analytical solution is available [Kelkar, 1992]:

$$V_{\mathrm{C}}(Q) = 4\pi Z^2 \mathrm{e}^2 \frac{[\kappa \sin(Q\sigma) + Q\cos(Q\sigma)]}{\varepsilon(1+0.5\kappa\sigma)^2 Q(Q^2+\kappa^2)}, \qquad (1.16)$$

while the Fourier-transform of the Yukawa-type potential

$$V_{\mathrm{Y}}(Q) = \int_{\sigma}^{\infty} V_{\mathrm{Y}}(r) 4\pi r^2 \frac{\sin(Qr)}{Qr} \mathrm{d}r \qquad (1.17)$$

has to be approximated via numerical integration and the upper limit of integration was restricted to 3σ since the integral value is not sensitive to the values beyond that limit.

Using equations (1.10)-(1.17), it is possible to calculate a theoretical function $S_{\mathrm{th}}(Q,n)$ for a given J- and d-value, which can be compared with the experimentally measured structure factors $S_{\mathrm{exp}}(Q,n)$. The latter can be retrieved from the observed

scattering curves $I(Q,n)$. $I(Q,n)$ is a function of the particle number density n, the form-factor $P(Q)$, the structure factor $S(Q,n)$ and a particle- and instrument-specific constant K:

$$I(Q,n) = n\ K\ P(Q)\ S(Q,n) \qquad (1.18)$$

For low particle concentrations ($n \rightarrow 0$), $S(Q,n)$ equals 1 for all values of Q, so that at a sufficient low reference-concentration n_0, Eq. (1.18) simplifies to:

$$I(Q,n_0) = n_0\ K\ P(Q) \qquad (1.19)$$

Thus, the ratio of the experimental scattering curve $I(Q,n)$ at a given particle concentration n over the scattering curve $I(Q,n_0)$ at low particle concentration n_0 yields the structure factor $S_{exp}(Q,n)$:

$$\frac{I(Q,n)}{I(Q,n_0)} = \frac{n\ K\ P(Q)\ S_{exp}(Q,n)}{n_0\ K\ P(Q)} = \frac{n}{n_0} S_{exp}(Q,n) \qquad (1.20)$$

In this manner, $S_{exp}(Q,n)$ can be determined and compared with the calculated $S_{th}(Q,n)$, and by an efficient fitting procedure, the theoretical structure factor can be refined to the observed one.

The total interaction potential among the lysozyme molecules under charge screening conditions with 250 mM KCl is shown in Fig. 1.8. The interaction potential parameters J and d are indicated as well. The total potential is calculated by the DLVO theory using the hard sphere model and the attractions are highly attractive in nature [Niebuhr and Koch [2005].

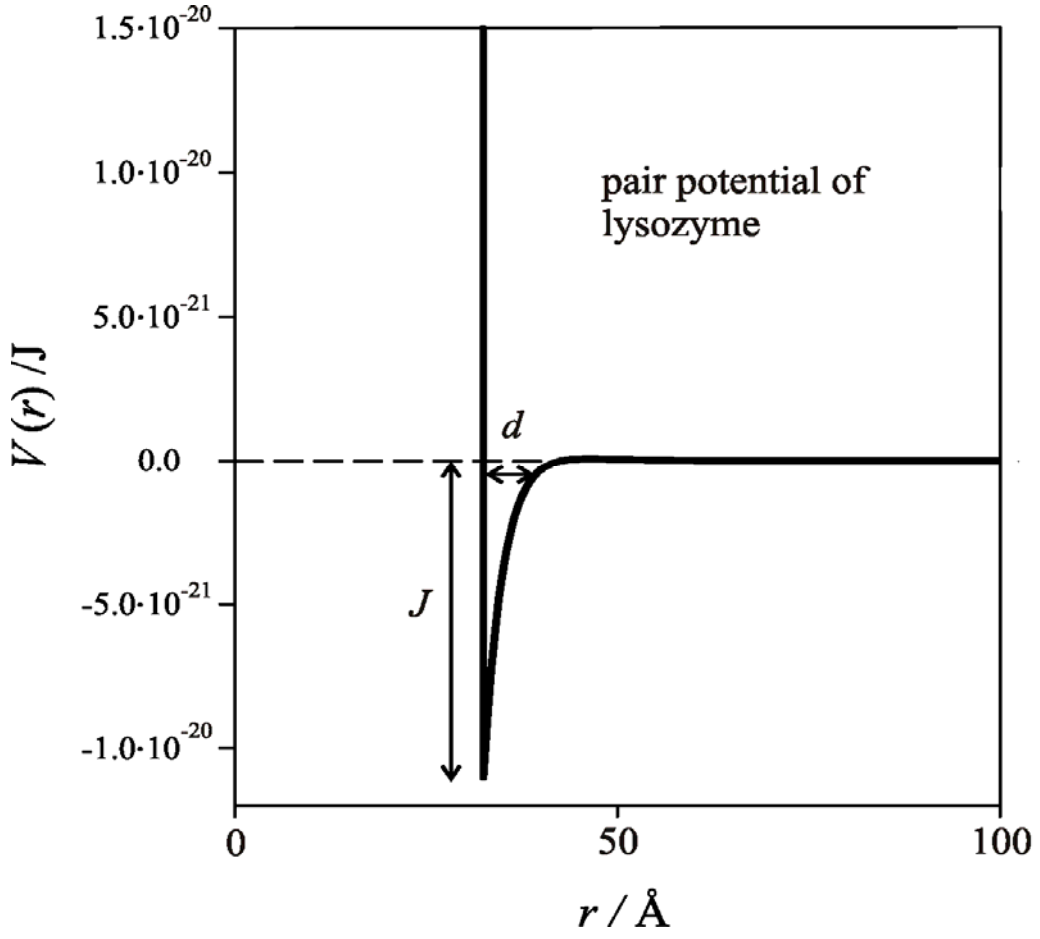


Figure 1.8: Total interaction potential between lysozyme molecules at high salt concentration (250 mM KCl) calculated by the DLVO theory from the given interactions parameters by Niebuhr and Koch [2005].

Chapter 2

Materials and Methods

2.1 Lysozyme sample preparation

Lyophilized and dialyzed lysozyme powder (Sigma-Aldrich, Germany) was dissolved in doubly distilled deionized water to prepare protein stock solutions. At the chosen pH of 4.6, lysozyme has a net charge of +10, and the charged groups are symmetrically distributed on the protein's surface. As cosolvents, glycerol (Fluka, Biochemica), sucrose (Sigma), guanidinium chloride (Sigma), 2, 2, 2-trifluoroethanol (TFE) (Merck), sodium chloride, and potassium sulphate (Sigma) were used as purchased. Solutions for the scattering experiments were prepared from stock solutions. Pure lysozyme solutions were prepared at concentrations of 10, 40 and 100 mg/ml (1, 4, 10 wt %) in 20 mM citrate buffer and the pH was adjusted to 4.6 with 1 M HCl. The solutions were filtered through 0.1 μm anotop (Whatman) filters. The samples for different concentrations of cosolvents with the same concentrations of protein in 20 mM citrate buffer were prepared by adding the appropriate amounts from stock solutions to obtain 500 mM and 2 M glycerol, 500 mM and 1 M sucrose, 10 and 35 % (v/v) 2, 2, 2-trifluoroethanol, 500 mM and 3.5 M guanidinium chloride, 250 mM potassium sulphate and 50, 100 and 200 mM sodium chloride solutions. The cosolvent solutions were also adjusted to pH 4.6 using 1 M HCl or NaOH, respectively, and filtered likewise. The samples were kept at -20 °C after preparation until use to prevent aggregation and degradation of the protein.

2.2 Insulin sample preparation

Bovine insulin was purchased from Sigma and used without further purification. Solutions were prepared by adjusting a pH of 2.0 with 1 M HCl and adding the appropriate amount of

sodium chloride or ethanol from stock solutions (2 M and 99.9 % (v/v), respectively).

Insulin solutions were prepared at concentrations of 5, 10, 40 and 100 mg/ml (0.5, 1, 4 and 10 wt %) in water and the pH was adjusted to 2.0 with 1 M HCl solution. The final concentrations of HCl in these protein solutions required to adjust the pH 2.0 were 50, 50, 130 and 280 mM of HCl, respectively. These concentrations of added HCl were used to calculate the Debye-Hückel screening length. The solutions were filtered through 0.1 μm anotop (Whatman) filters.

Insulin solutions with similar protein concentrations and pH were prepared with a final concentration of 100 mM of sodium chloride solution to probe the charge screening effects. The final concentrations of HCl required to adjust insulin solutions of the different concentrations above mentioned to pH 2.0, were 32.5, 50, 125 and 275 mM respectively, and filtered likewise.

Insulin solutions with similar protein concentrations and pH were prepared with the final concentration of 20 % (v/v) ethanol to probe the solvent hydrophobicity effects. The final concentrations of HCl used, which were required to adjust insulin solutions of the different above mentioned concentrations to pH 2.0 were 32.5, 50, 127 and 250 mM, respectively, and filtered likewise.

All the solutions were prepared freshly prior to the X-ray scattering measurements.

2.3 Beamline X33 at DORIS

The X33 beamline dedicated to small-angle X-ray scattering of biological materials, especially proteins and nucleic acids is located in the D-Fan in the HASYLAB's experimental hall at DESY (Deutsches Elektronen-Synchrotron) in Germany. The experimental set-up is shown in Fig. 2.1. The SAXS set-up consists of the optics hutch and the experimental hutch.

The optical hutch at beamline consists of a monochromator, which is a horizontal focusing triangular Si(111) crystal with asymmetric cut (7°). The monochromator focuses the monochromatic X-rays to a rhodium coated flat mirror on zerodur substrate with gravimetric bending. In the SAXS experiments, the alignment of the beam is very crucial and is carried out by adjusting and overlapping the reflections both from the monochromator and the

mirror. The purpose of this alignment is to reduce the background noise as much as possible and to obtain a smallest possible focal spot of the X-ray incident beam.

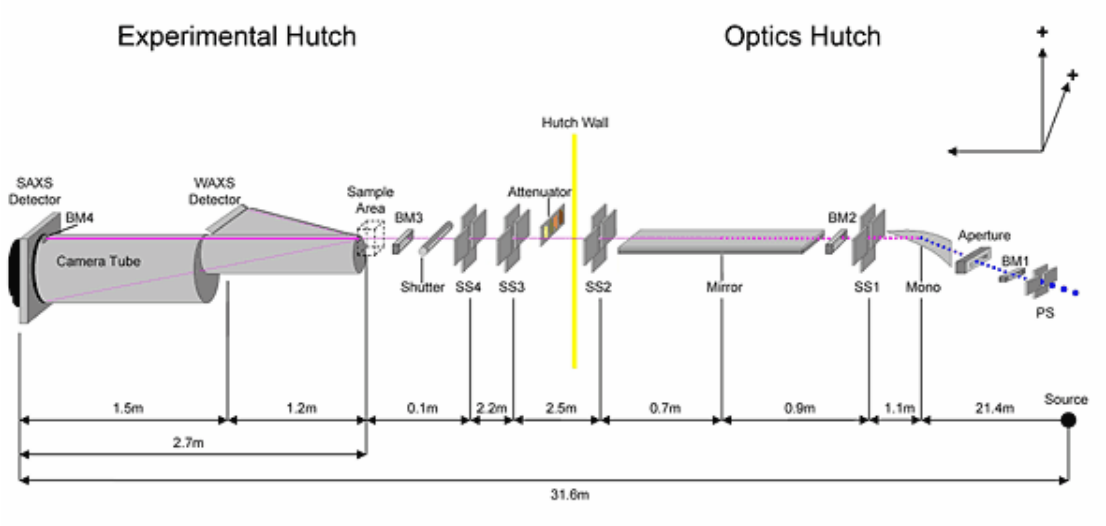


Figure 2.1: The beamline X33 of the EMBL at DORIS (HASYLAB/DESY)

Following the optics hutch is the experimental hutch which contains the ionization chamber, sample holder, camera and the detector. In the DORIS storage ring, the intensity of the beam decays with time and during the measurements the intensity of the incoming beam varies. Therefore, the incident beam intensity should be measured prior to the sample cell for normalizing the data, in order to make different sets of measurements comparable. For this purpose, an ionization chamber having a high voltage capacitor containing air as dielectric is used. The X-rays will ionize the air molecules and resulting current will be measured which is proportional to the intensity of the incident X-ray beam. The aligned and focused incident beam passes through the sample cell which is shown in Fig. 2.2. The windows of the sample container are made from mica and the thickness of the cavity holding the sample is about 1 mm. The incident beam irradiating the sample is scattered at different angles and then enters along with the incident beam a large evacuated metal tube. The large evacuated tube or so called camera is important to avoid parasitic scattering due to air molecules (e.g. CO_2) which interferes to a large extent with the scattering intensity pattern of the protein molecules which are poor scatterers in nature. The incident beam is highly intense as compared to the scattered beam and the direct exposure to the detector can seriously damage the detector and would

intervene the scattering patterns as well. In order to avoid the direct exposure of the incident beam to the detector, a beam stop made of lead is used which absorbs the X-rays. In order to avoid the direct exposure of the incident beam to the detector, a beam stop made of lead is used which absorbs the X-rays.

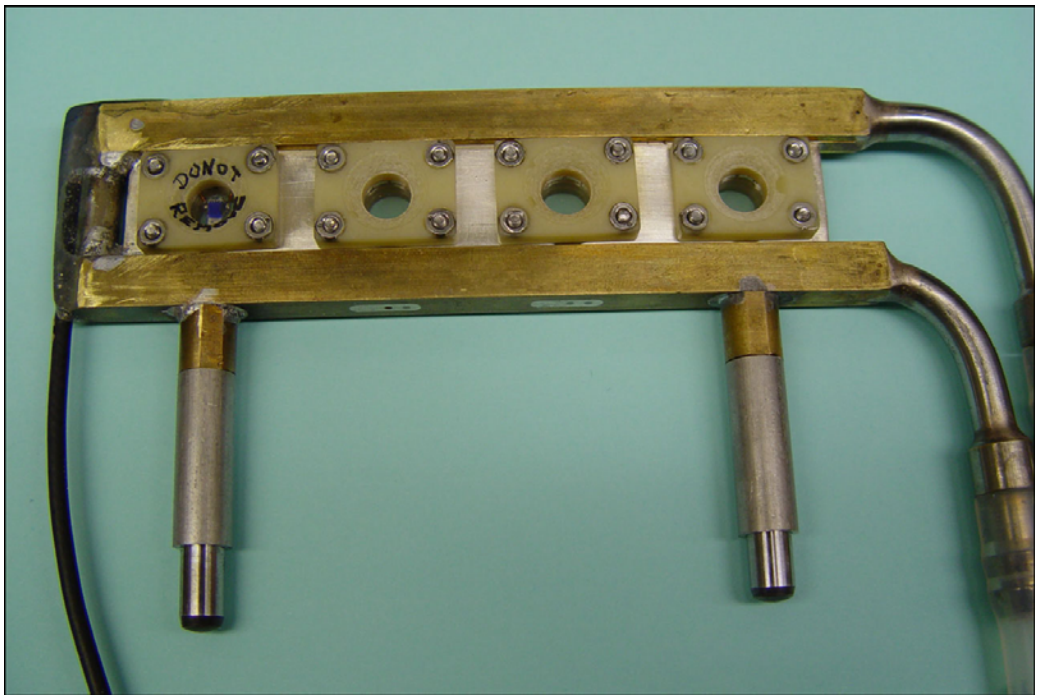


Figure 2.2: Image of a sample holder with three sample cells. The leftmost of the cell contains a temperature sensor. The tubing on the right can be connected to a water bath.

The *Q*-range which can be covered by SAXS depends upon the distance between the sample cell and the detector. In these measurements the sample-to-detector distance was about 2.4 m which covers a *Q*- range from 0.008 $Å^{-1}$ to 0.45 $Å^{-1}$. The measurements carried out at X33 beamline were performed at high concentrations of protein and cosolvents (e.g., 10 wt% lysozyme solution in 3.5 M GdmCl). These samples can also influence the scattering due to absorption effects. Therefore the intensity of the primary beam is measured after the scattering process as well, which is carried out by employing a X-ray sensitive photodiode in the beam stop. The last part of the set-up is a two-dimensional MAR345 image plate detector. The response of the detector is not constant over the entire active area of the image plate, so it is compulsory to measure its response $DR(n)$, which can be done by using the

homogenous irradiation with a radioactive iron source (^{55}Fe). During the processing of scattered data, the experimental data is divided by the detector response as shown in the following section. In order to calculate from the channel numbers of the detector the modulus of the scattering vector, silver behenate is used as calibrant.

In biological samples, especially the proteins where different kinds of buffers with different cosolvents and cosolutes are used, the difference in the scattering intensity is measured which is due to the excess electron density. This is carried out by recording the buffer sample (with/without cosolvent) scattering intensity pattern separately and then later subtracted from the scattering intensity pattern of the protein solution in buffer (with/without cosolvent).

2.3.1 Measurements at beamline X33

A first set of solution scattering measurements of different lysozyme concentrations without and with different concentrations of cosolvents (sucrose, potassium sulphate and guanidinium chloride) was carried out at the beamline X33. The scattering intensities were measured for 300 s each and the buffer/cosolvent background was measured before and after each protein measurement to obtain sufficient statistics and verify that the cell had been properly cleaned. The X-ray scattering measurements are very sensitive to the nature of the solvent (buffers and cosolvents) used. The cosolvents may change the scattering patterns entirely due to different scattering densities and their interactions with proteins. Cosolvents like TFE and GdmCl partially unfold the protein molecules and salts like potassium sulphate precipitate the proteins. These changes drastically affect the scattering measurements, so a very careful approach should be adopted. To avoid all possible interferences, the prerequisite is -as clean as possible the sample cell. Before starting measurements, the sample cell is washed with ethanol and later with deionised water thoroughly. There should be no adsorbed material or scratches on the mica windows. In the second step, the cell is rinsed with the buffer solution to be measured about 2-3 times. After recording the scattering pattern of the buffer, the cell is made empty and dried completely; otherwise the concentration of the following protein sample will be changed. In the third step, after measuring the protein scattering, the sample cell is again washed with water followed by ethanol and water and then the same buffer solution is measured again. Cleaning of the cell is not required after this

step if the sample has a similar buffer but only a higher concentration of protein. In the next step, cell is washed again with water, ethanol, and water and buffer solution. The buffer solution is measured again likewise. In case of protein adsorption, the mica windows needed to be changed.

The final scattering patterns $I(Q)$ were obtained after averaging the frames that were statistically identical, correction of detector response, normalization to the intensity of the transmitted beam and the protein concentration, and subtraction of an averaged buffer background pattern using the program SAPOKO. For dilute solutions, the experimental scattering curve is given by:

$$I(Q) = \frac{1}{c}\left[\frac{I_s(Q)}{I_{s,0}} - \frac{I_b(Q)}{I_{b,0}}\right]\frac{1}{DR(Q)} \qquad (2.1)$$

where $Q = (4\pi/\lambda)\sin\theta$ is the momentum transfer, 2θ is the scattering angle, and λ the wavelength of the incident radiation (0.15 nm), c is the concentration, $I_s(Q)$ is the intensity scattered by protein solution, $I_b(Q)$ is the intensity scattered by the buffer solution including cosolvents/cosolute and salts, $I_{s,0}$ and $I_{b,0}$ are the average intensities of the direct beam transmitted through the sample and buffer along with cosolvent/cosolute and salts, respectively, and $DR(Q)$ is the detector response with respect to Q. For concentrated solutions, the difference in absorption of the solution and solvent must be taken into account. Dividing by the transmitted intensities corrects for the decay of the primary beam intensity and the absorption of the sample. The scattering patterns were scaled in the range 0.18 $Å^{-1} \leq Q \leq 0.3$ $Å^{-1}$ to correct for minor differences in concentration and contrast. As all curves are superimposed in this Q-range, it can be assumed that the interactions no longer influence the scattering pattern at these high momentum transfers.

2.4 Beamline BL9 of DELTA

The beamline BL9 of the Dortmunder Elektronen Spericheringanlage (DELTA) located at the University of Dortmund, Germany [Tolan et al, 2003] is dedicated to material sciences research with multifunctional purposes and focusing on X-ray diffraction, X-ray reflectivity

and small-angle X-ray scattering experiments. A superconducting asymmetric wiggler (SAW), two undulators and several dipole magnets provide synchrotron radiation (maximum energy 1.5 GeV) to the experimental end station of BL9. The experimental set-up at beamline BL9 attached to SAW covers the energy range between 4 keV to 30 keV.

A schematic overview of the experimental set-up for two dimensional SAXS at BL9 of DELTA is depicted in Fig. 2.3.

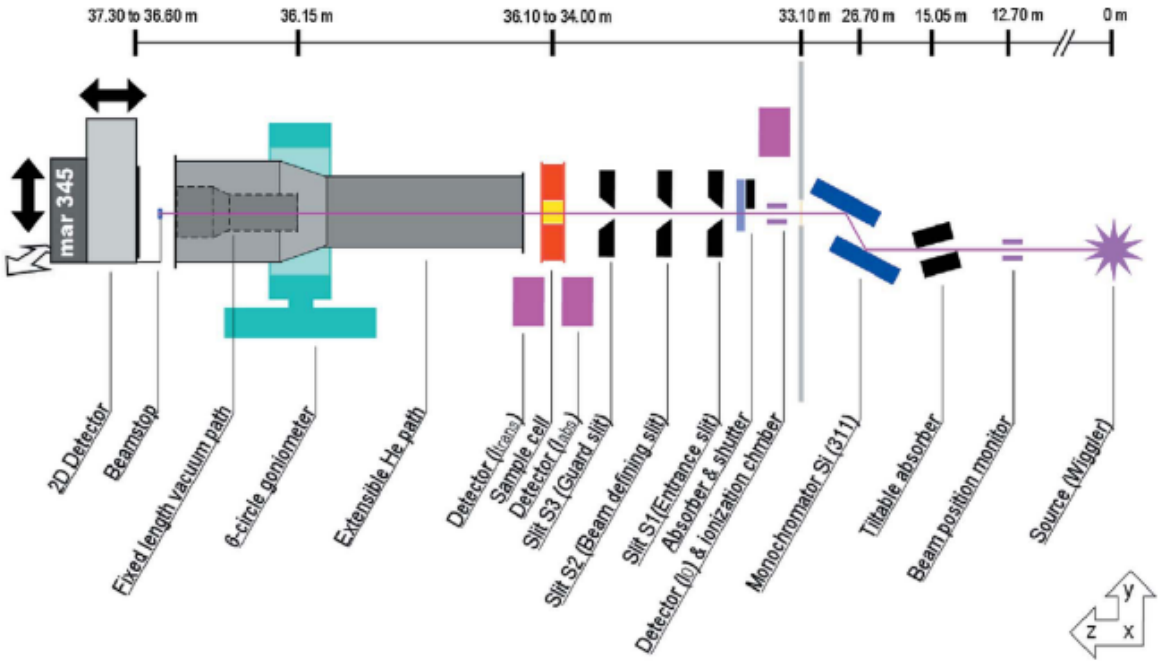


Figure 2.3: Outline of the SAXS/WAXS set-up on BL9 according to Krywka, et al. (2007). The samples can be mounted at different sample-to-detector distances, ranging from 0.45 m to 3.3 m. For the smaller distances, the sample can be placed in the diffractometer and the flexible extensible helium path can be reduced in length.

The monochromatic X-ray beam is vertically focused at the detector position by means of the sagittally bent second monochromator crystal and collimated by using the slit systems S1 and S2, with S3 being a guard slit to reduce slit scattering. The lifetime of X-ray beam at DELTA is about 10 hr. As the current or intensity of the beam decays gradually from 120 mA, the intensity of the incident beam is measured before and behind the sample cell for normalization as discussed in Section 2.3. The samples are filled into a removable easy-to-

change sample cell, which has mica windows. The cell is placed in a sample holder shown in Fig. 2.4.

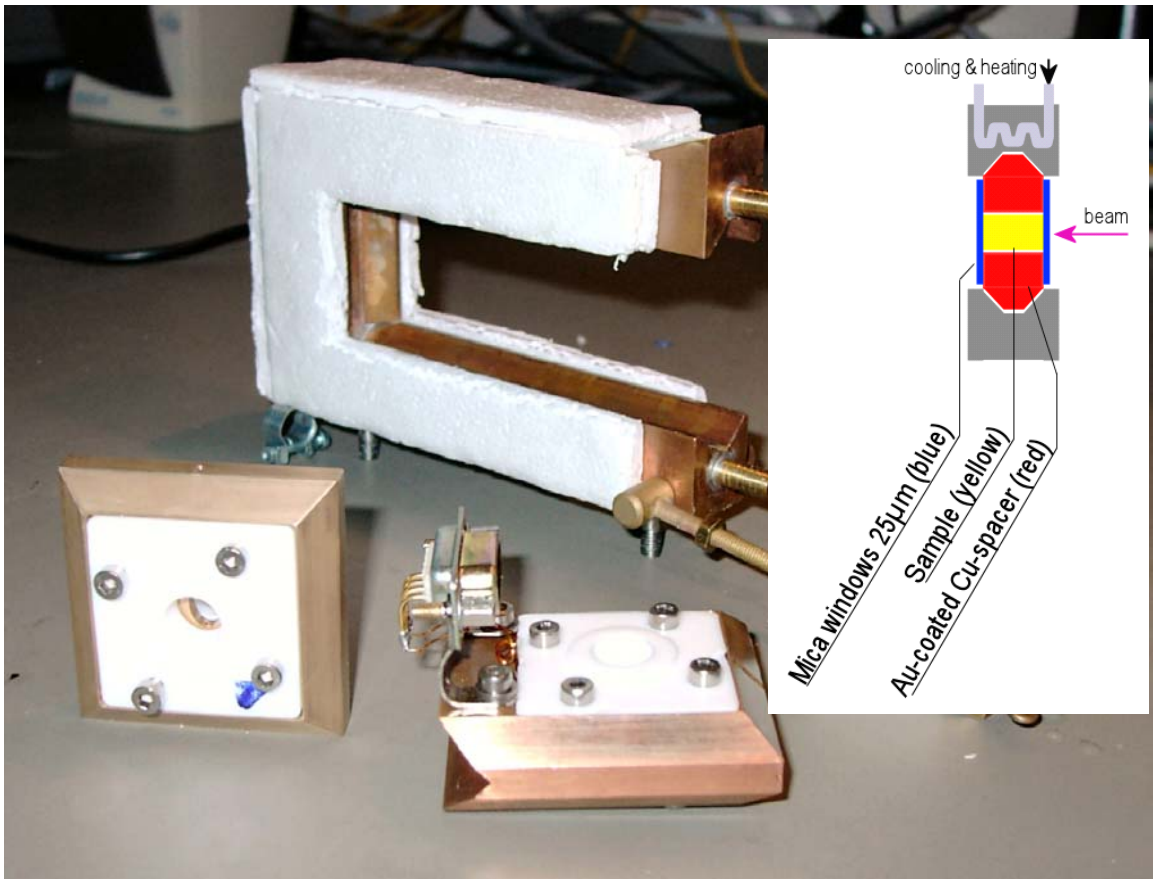


Figure 2.4 Sample environment for SAXS measurements at BL9 adapted from Krywka et al. [2007]. The white U-shaped temperature-controlled sample holder can be loaded with an easy-to-change sample cell and a temperature-readout sensor, the inset shows the schematic profile of the cell.

The samples are filled directly into the cavity of the sample container by a syringe with a thin needle. The sample container is encircled by a gold-coated copper frame which has the mica windows, each of 3 mm thickness. Following the sample holder, is the camera which is a modular extensible tube filled with helium having a large diameter to cover the complete solid angle which can be seen by the detector. Unlike the camera set-up at beamline X33 described in Section 2.3, a large metallic evacuated tube is not used due to its heavy weight and semi-permanent installation. The set-up at BL9 is multipurpose and frequently used for

other experiments which require the beam path to be easily unmounted, so the flexible tubing is a better alternative. The sample-to-detector distance can be varied freely within the range of 45-330 cm. A lead beam stop of diameter 8 mm is positioned in front of the detector. The image plate MAR345 detector is used. The detector-to-sample distance measurements and conversion of detector channel number to modulus of scattering vector is carried out by using silver behenate as calibrant with a *d*-spacing of 58.376 Å. Further information regarding the measurements and processing of data is given in following section.

2.4.1 Measurements at beamline BL9

A second set of the scattering data was recorded at the multi-purpose experimental end station of beamline BL9 at the Dortmund Electron Accelerator DELTA. SAXS images were taken from aqueous solutions of lysozyme at different concentrations (1, 4 and 10 % w/v) in 20 mM citrate buffer at pH 4.6 at room temperature (25 °C), without and with added cosolvents (glycerol, trifluoroethanol and sodium chloride) at different concentrations. Furthermore, insulin solution scattering patterns for different concentrations of insulin (0.5, 1, 4, 10 and 20 wt %) in water at pH 2.0 and with added ethanol and sodium chloride at room temperature (25 °C) were carried out at BL9.

Prior to each measurement, the sample was thawed and vigorously stirred to assure formation of a homogenous mixture of the protein without any aggregation products, which can be easily identified in the SAXS pattern. The scattering intensities were accumulated over different time periods, depending on the concentration of the solution, ranging from 900 to 1800 s. In test measurements, no changes in the scattering intensity were found in 1 % (w/v) lysozyme solution and 0.5 % (w/v) insulin solutions exposed to the X-ray beam over these time intervals, indicating that no radiation damage and no significant radiation-induced aggregation takes place and has to be considered in the evaluation of the data. The measured intensities were integrated radially using the program package FIT2D [Hammersley et al., 1994]. Further normalization of the data was required to account for the variation of the intensity of the incoming beam and the absorption due to the different concentrations of the solutions. Therefore, the intensity of the incoming beam and the transmitted beam was measured and accumulated over the exposure time, thus yielding the normalized intensity

given by $I(Q) = I(Q)_{exp}/I_{trans}$. The scattering profile was then corrected for background scattering by subtraction of the corresponding buffer and sample cell scattering contribution as described in section 2.3.1.

2.5 Measurement of dielectric permittivity

The dielectric permittivity ε of the cosolvent solutions were determined experimentally to be able to use these values as direct input parameters in the screened model potential used, without having to rely on approximate calculations of ε. The electrical conductance of some of the samples did not allow performing routine experiments for measuring the dielectric permittivity. Rather, the determination is based on the analysis of the real part, $\varepsilon'(\nu)$, of the frequency-dependent dielectric permittivity as a function of frequency ν in the range 200 MHz $\leq \nu \leq$ 20 GHz. For this purpose, dielectric spectra were taken employing a network analyzer HP 8720 (Hewlett-Packard) with probe HP 85070B. The spectrum of each solution was measured at least four times and the recorded values were averaged. All measurements were carried out at 25 ^{o}C. The static dielectric permittivity is given by the zero-frequency limit of $\varepsilon'(\nu)$. The measured spectra were fitted to standard expressions such as the Debye function $\varepsilon'(\nu) = \varepsilon_\infty + S/(1+(2\pi\nu\tau)^2)$, where ε_∞ is the high-frequency limit of ε, S the amplitude of the relaxation mode, and τ the relaxation time, followed by zero-frequency extrapolation. The zero-frequency value of interest here is known to be rather insensitive to the choice of the model function provided that the experiments are conducted to low enough frequencies to capture all relaxation modes. For systems of the high fluidity considered here, this was indeed fulfilled in all cases.

2.6 Fitting Strategy

Experimental structure factors $S_{exp}(Q,n)$ were calculated from concentration dependent experiments of the proteins lysozyme and insulin at the previously mentioned conditions by using Eq. 1.20. The experimentally calculated structure factors were then fitted to obtain theoretical structure factors, $S_{th}(Q,n)$, with the help of equations 1.10-1.175, according to the models from liquid state theoretical approaches to find out the best fit

parameters for the depth, *J* and the range, *d* of the attractive potential. All the fittings were carried out by using the program "Mathematica", and the strategy was as followings:

1) All the parameters and constants used in the calculations are read in Mathematica, which include the total number of particles at different concentrations of the lysozyme and insulin, the temperature, charge on the protein molecule, the elementary charge, diameter of the protein molecule, Boltzman's constant, the dielectric constant measured experimentally for different concentrations of added cosolvents, Avogadro's number and the Debye-Hückel screening length which determines the range of the interaction.
2) All the functions/formulas used in calculating the theoretical structure factors in real and *Q*-space are described in separate file as definitions and read by Mathematica. These definitions include the volume/packing fraction of the protein at different concentrations, the radius of the particle, the screened Coulomb potential, sum of the attractive potentials (having *J* and *d* as fitting parameters), first-order spherical Bessel function and the structure factor of the reference system.
3) Global fitting of the experimentally calculated structure factors is carried out through a continuous loop in Mathematica. Different starting values of the fitting parameters *J* and *d* are given. The first theoretical structure factor is calculated after reading all the parameters and the definitions. Then the experimental and first theoretical structure factors are compared for a given set of of *J* and *d* values. Then in the next step these values of fitting parameters are increased and the second theoretical structure factor is calculated. If the differences between experimental and second theoretical structure factors are larger than that obtained in the first one, the first given values of fitting parameters are kept, otherwise these values are again incremented till the best fit with minimum difference is obtained. The best fitting parameters and theoretical structure factors are obtained for all sets of protein concentrations and cosolvents using the same procedure.

The fitting strategy is given schematically in Fig. 2.5 and the details of the above mentioned strategy regarding the data processing in "Mathematica" are described in detail in Appendix 1.

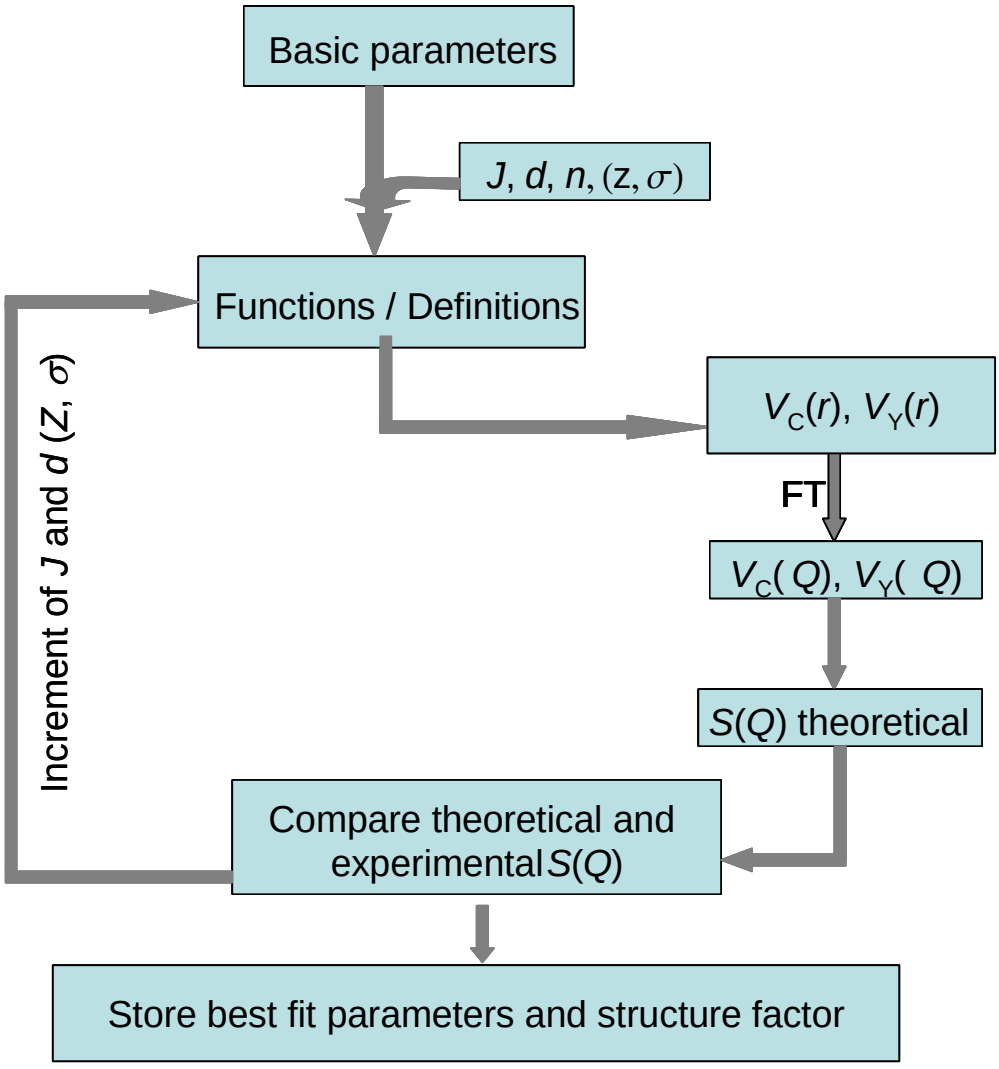


Figure 2.5 Schematic representation of the fitting strategy to determine the potential parameters *J* and *d*.

Chapter 3

Results and Discussion

3.1 Lysozyme intermolecular interactions

3.1.1 Form factors

The scattering intensity curves of 1 wt % lysozyme solutions in 20 mM citrate buffer with all the added cosolvents were measured separately for use as form factors, $P(Q)$, of the protein under the corresponding cosolvent conditions. At this low concentration, interparticle interactions between the protein molecules are essentially negligible and the scattering intensity can still be measured with sufficient statistical accuracy at synchrotron sources. As described previously in section 1.3, when the sample is a macromolecule in solution, like proteins in our case, the scattering due to the excess electron density between the protein molecule and solvent is effectively measured. Thus the scattering of the buffer has to be recorded and subtracted from that of the solution to obtain the scattering pattern owing to the protein molecule only. A typical small-angle X-ray scattering patterns of a 1 wt % lysozyme solution in 20 mM citrate buffer along with the 20 mM citrate buffer at pH 4.6 and at room temperature are shown in Fig. 3.1.

Scattering patterns at higher angles coincide with each other very well in the case of lower protein or cosolute/cosolvent concentrations but vary to a different extent for higher concentrations. In the case, where the scattering intensity at very high angles do not coincide, the scattering patterns were scaled with respect to the scattering patterns of the buffer solutions (in the presence of cosolute/cosolvents) in such a way that the data points match each other and after subtraction of the background scattering pattern, all the

different concentration dependent intensity curves coincide at high angles or *Q*-values respectively.

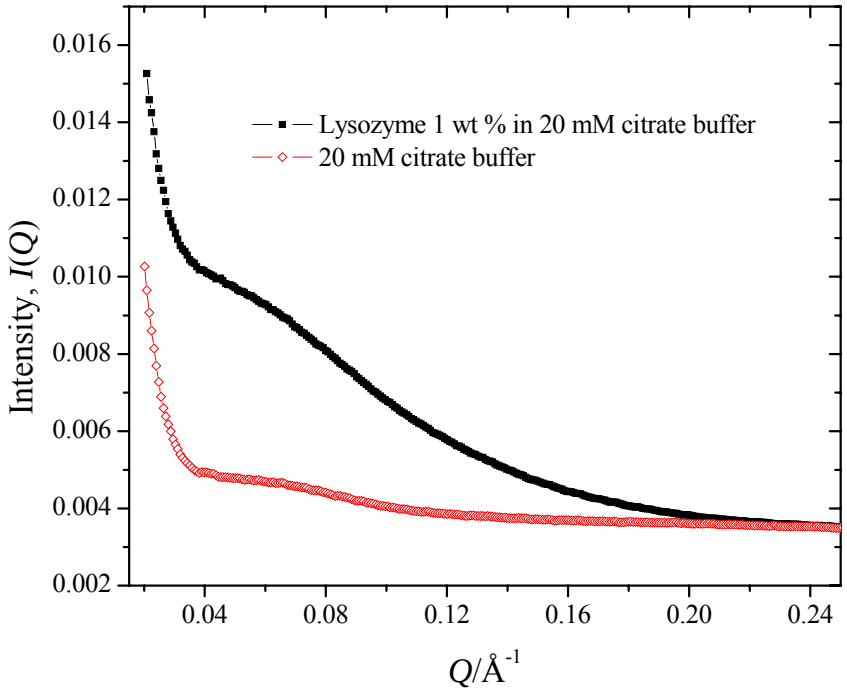


Figure 3.1: Plot of the SAXS intensity *I*(*Q*) (in arbitrary units) of 10 mg/mL of lysozyme solution in 20 mM citrate buffer (in black) and of pure 20 mM citrate buffer (in red) as background.

This approach has been applied to all the experimental data measured for lysozyme and insulin solutions measured at different synchrotron sources. The contribution of attractive or repulsive interactions to the scattered intensity, mostly influence the scattering data at very small angles and at higher angles the structure factor (indicative of intermolecular interactions) approaches $S(Q) = 1$ [Konarev et al., 2003]. The scattering pattern of 1 wt % lysozyme after background subtraction is shown in Fig. 3.2, which is used as form factor *P*(*Q*).

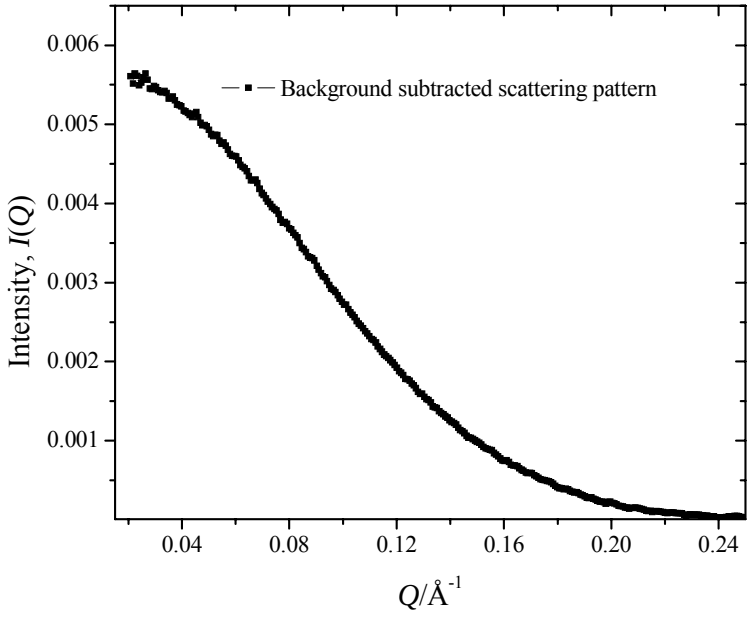


Figure 3.2: Background (buffer) subtracted SAXS intensity $I(Q)$ (in arbitrary units) of 10 mg/ml of lysozyme solution in 20 mM citrate buffer at pH 4.6.

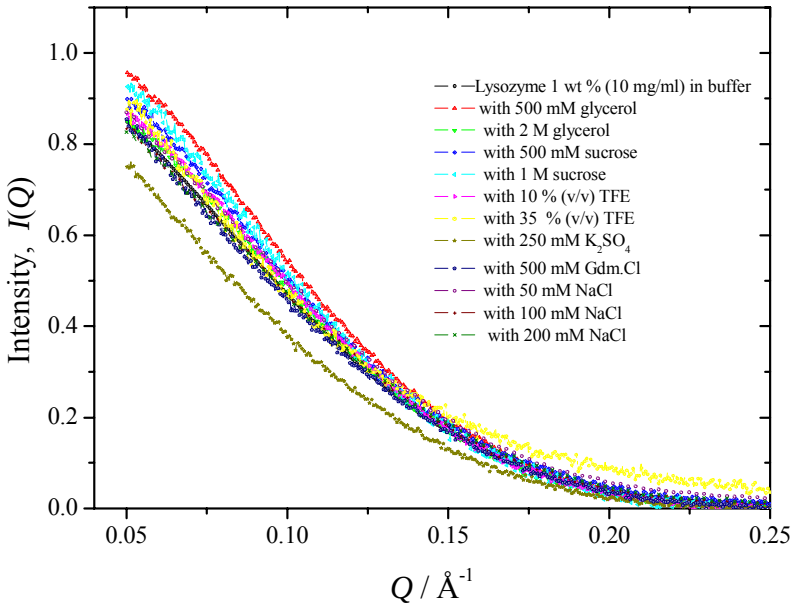


Figure 3.3: Plot of the normalized SAXS intensity $I(Q)$ as a function of momentum transfer Q of 1 wt % lysozyme (used as form factor, $P(Q)$) with different cosolutes and cosolvents at various concentrations.

3.1.2 Guinier analysis

The form factors of all the individual sets of measurements were also used to determine eventually occurring changes in particle dimensions upon addition of the cosolvents. The scattering intensity patterns are depicted in Fig. 3.3. The radii of gyration, R_g, were determined from Guinier plots by using the program PRIMUS [Konarev et al., 2003]. An example of the analysis is shown in Fig. 3.4. The radius of gyration was found to be 15.3 ± 0.15 Å for lysozyme in pure buffer solution, in glycerol, sucrose and potassium sulphate. For 35 % (v/v) trifluoroethanol and 3.5 M guanidinium chloride, R_g was found to be 16.4 ± 0.1 Å, indicating partial unfolding of the protein molecule. Addition of NaCl leads to R_g = 14.6 ± 0.3 Å for 50 mM NaCl and R_g = 15 ± 0.2 Å for the 100 and 200 mM NaCl solutions. These data are in good agreement with previously measured data for lysozyme solutions yielding R_g = 15.4 ±0.2 Å [Svergun et al., 1998].

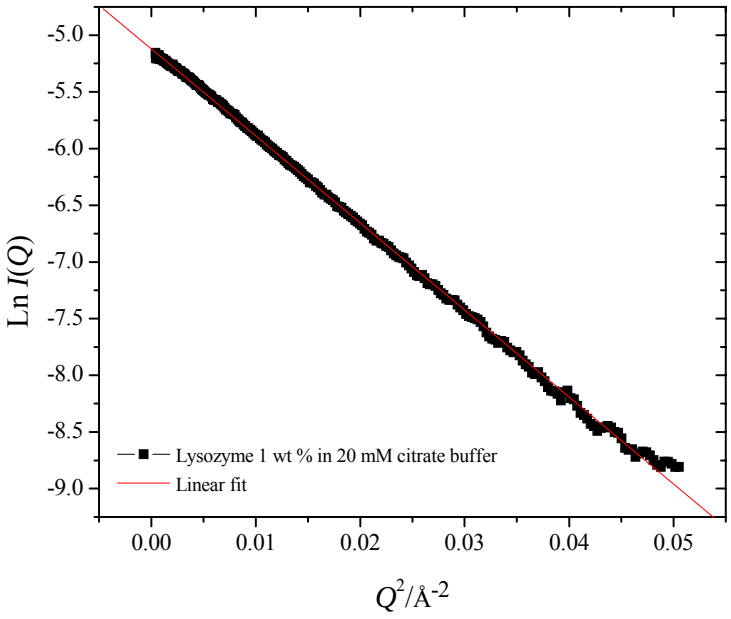


Figure 3.4: Guinier's analysis for a 1 wt % lysozyme solution in 20 mM citrate buffer at pH 4.6.

3.1.3 Shape estimation through distance distribution analysis

The distance distribution function $p(r)$ describes the particle shape as well as particle size as mentioned in Section 1.3.2. The particle shape of the lysozyme molecule in solution was determined by indirect transformation of the scattering intensity pattern of the molecule by using the program GNOM [Svergun, 1992]. The method is based on perceptual criteria of what is the best solution. Initially, an approximate size calculated by Guinier's analysis is given as input and then optimum value of the regularization parameter (Lagrange multiplier) is determined while applying the indirect transformation. The distance distribution function for lysozyme in 20 mM citrate buffer at pH 4.6 is given in Fig. 3.5, which is a typical bell-shaped function and indicative of its globular structure. The maximum of the function gives approximately the radius of the particle which is approximately 20 Å. The particle shape is simulated by using the simulated annealing method using program GASBOR [Svergun, 2001] and a spherical shape of the particle in solution was found. In the case of 35 % (v/v) TFE and 3.5 M guanidinium hydrochloride, an elongated shape, indicating partial unfolding was calculated which is also clear from Guinier's analysis.

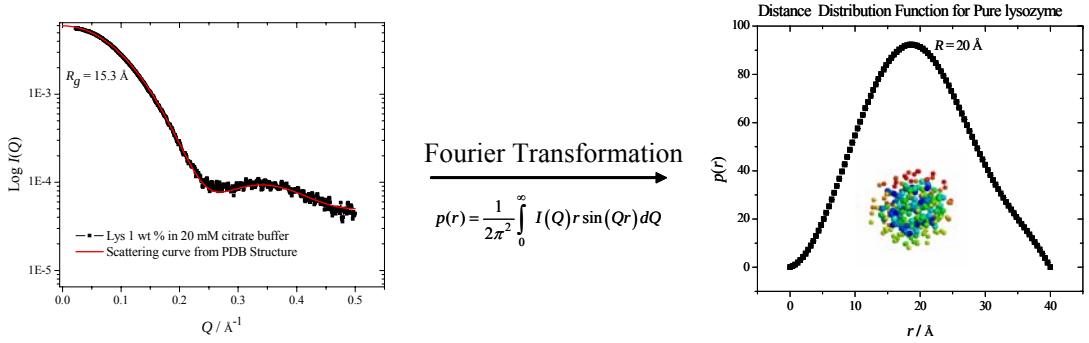


Figure 3.5: Scattering intensity pattern of 1 wt % lysozyme in buffer along with the fit of the lysozyme PDB structure (PDB entry 6LYZ) using the program CRYSOL [Svergun, 1995] shown on the left side and the corresponding distance distribution function calculated by GNOM [Svergun, 1992] along with the estimated shape by the program GASBOR [Svergun, 2001] shown on the right side.

3.2 Interaction potential determination of lysozyme

3.2.1 Lysozyme in pure buffer solution

A first set of experiments was performed with pure lysozyme solutions of concentrations 1, 4 and 10 wt% at pH 4.6 in 20 mM citrate buffer and 25 °C. The data $I(Q)$ are shown in Fig. 3.6. The scattering patterns measured for lysozyme concentrations of 4 and 10 wt% exhibit strong intermolecular correlation peaks and the scattering intensity at low Q-values decreases, indicating repulsive short-range ordering of the dissolved protein molecules. In order to extract the interparticle interaction potentials, $V(r)$, the experimental structure factors $S(Q)$ were calculated by using Eq. 1.20 and then fitted with

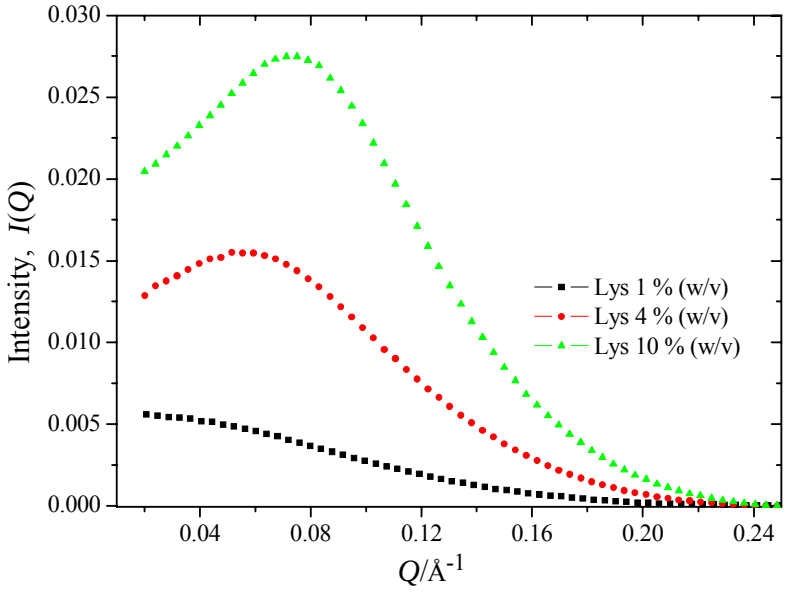


Figure 3.6: Plot of the SAXS intensity $I(Q)$ (in arbitrary units) as a function of momentum transfer Q for different concentrations of lysozyme in wt%: 1 (squares), 4 (circles), 10 (triangles) in 20 mM citrate buffer at pH 4.6.

the model introduced in the section 1.4.2. by using equations 1.10 - 1.17 and are depicted in Fig. 3.11. The Debye–Hückel screening length was calculated by taking into account

the 20 mM citrate buffer at pH 4.6 and the dielectric constant ε was 79.8. Lysozyme has a net positive charge of about 10 [Tanford and Roxby, 1972], but Tardieu et al. [1999] could obtain the best fits for the experimental structure factors by using an effective charge, Z = 6. In other studies, Narayanan and Liu [2003] obtained best fits using the same effective charge of Z = 6, and Niebuhr and Koch [2005] got best results in their analysis using Z = 6.4.

In the presented study, in a first step of simulating the theoretical structure factors, Z = 6 was fixed and then an effective hard sphere diameter (sigma, σ) was determined which could possibly fit the experimental structure factor for lower to higher protein concentrations in reasonable accordance by using the strategy mentioned in Section 2.6. The experimental and calculated (theoretical) structure factors with different σ are shown in Fig. 3.7 for 4 wt % and in Fig. 3.8 for 10 wt % lysozyme solutions. As it can be seen from the fits, the best results could be obtained by using an effective particle diameter of 28 Å. Niebuhr and Koch [2005] used a similar value, σ = 28.4 Å. In a second step, the particle diameter was kept constant at 28 Å and the charge was varied to estimate best fits at this effective particle diameter, which is shown in Fig. 3.9 for 4 wt % and in Fig. 3.10 for 10 wt % lysozyme solutions. Again the best matching fits could be obtained by using Z = 6. This approach of selecting an effective particle diameter and an effective charge was implied in all the fits carried out for different protein concentrations with differing concentrations of various cosolutes, cosolvents and osmolytes.

The attractive part of the pair potential calculated by global fitting of the experimental structure factor data (Fig. 3.11) was found to have a depth of -3.75 k_BT (J = 3.75 k_BT) for the lysozyme concentration of 4 wt%, and of -2.5 k_BT for the 10 wt% solution, in reasonable accordance with the findings of Niebuhr and Koch [2005], who obtained for the depth of the attractive potential -3.6 k_BT for a lysozyme concentration of 6.8 wt% in 20 mM potassium chloride at pH values between 6.5 and 7.0. The smaller J-value at the higher protein concentration of 10 wt% is due to a stronger repulsion between the protein molecules as the protein is positively charged at pH 4.6. Differences of the calculated structure factors from the experimental $S(Q)$ data at small Q-values appear for the highest protein concentration, only, which might be due to limitations of the model used (e.g., non-perfect spherical shape of the protein molecule or transient oligomer formation).

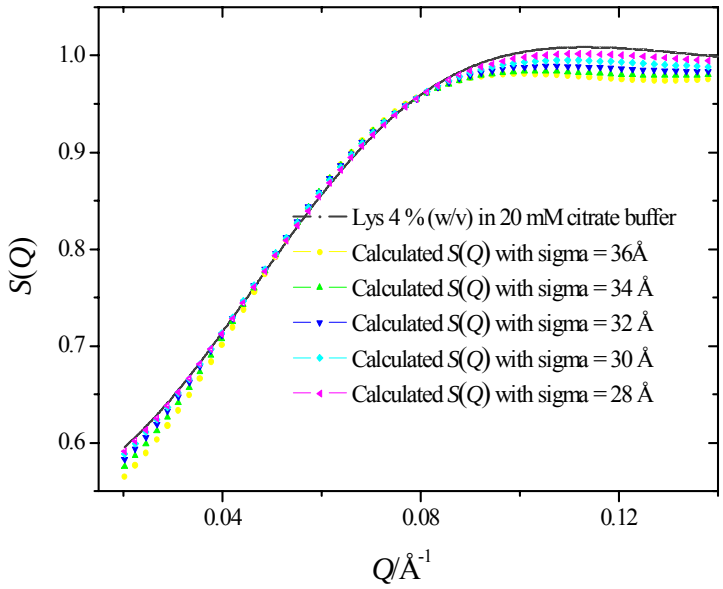


Figure 3.7: Effective sigma selection for a 4 wt% lysozyme solution using $Z = 6$.

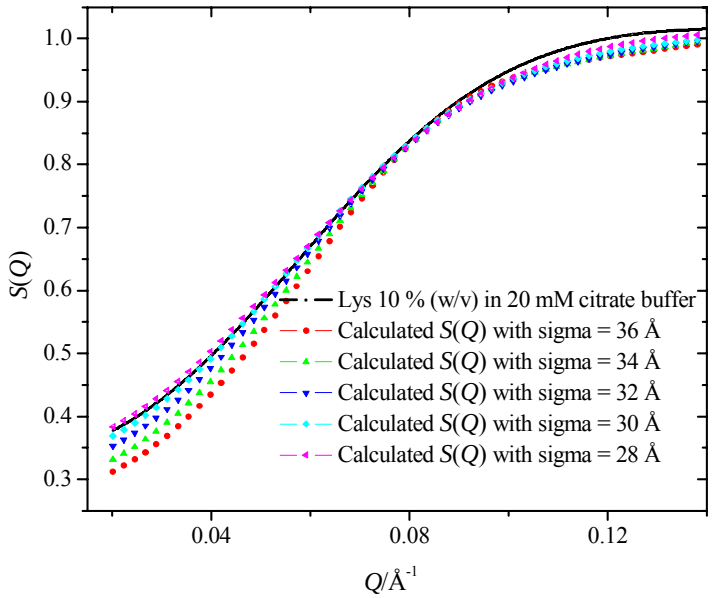


Figure 3.8: Effective sigma selection for a 10 wt% lysozyme solution using $Z = 6$.

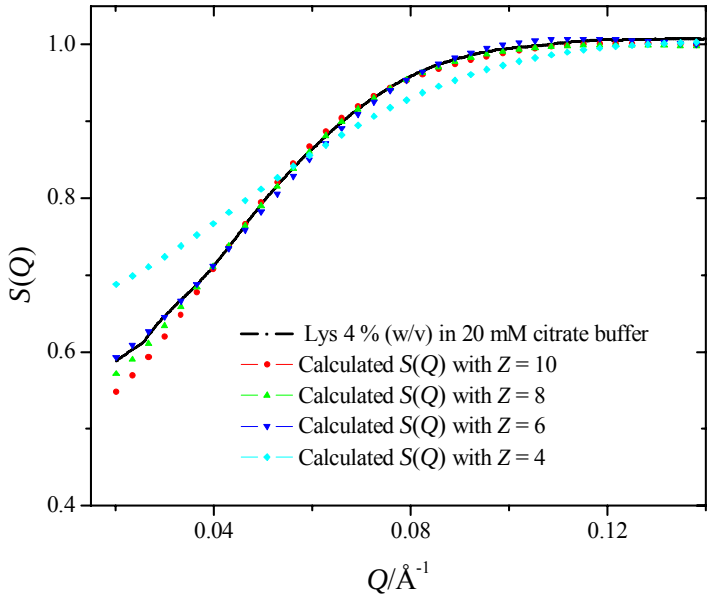


Figure 3.9: Effective charge selection for a 4 wt% lysozyme solution using an effective particle size of 28 Å.

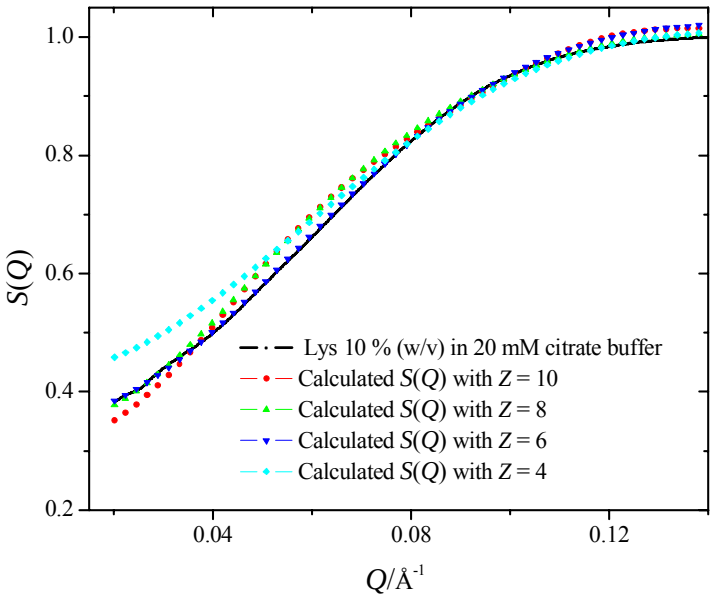


Figure 3.10: Effective charge selection for a 10 wt% lysozyme solution using an effective particle size of 28 Å.

Since the buffer concentration (20 mM) is quite low, the major attractive part of the potential is due to the van der Waals - potential with a small contribution from the osmotic attractive potential, only [Curtis et al., 1998]. As revealed by the fitting procedure, the DLVO potential model based on hard core repulsion, screened Coulomb repulsion and van der Waals - attraction with constant charge and varying depth and range of the attractive part can satisfactorily describe the observed trends of the concentration dependence of $S(Q)$.

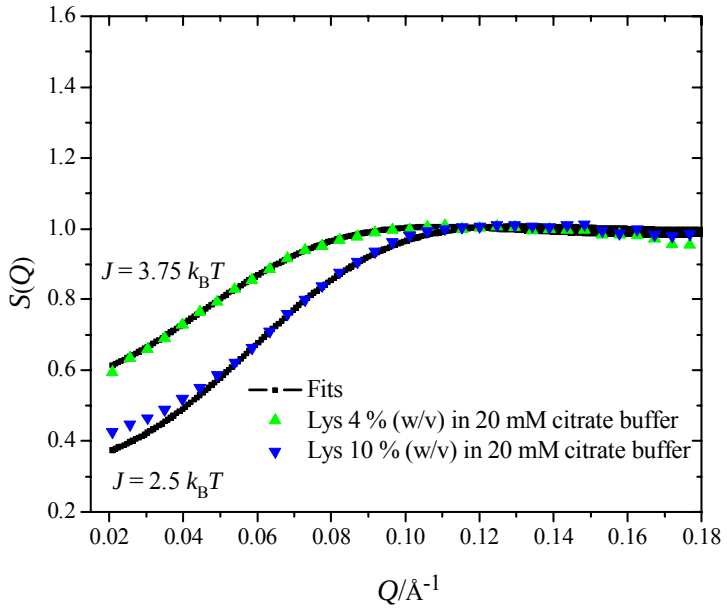


Figure 3.11: Experimental and calculated (full lines in black) structure factors $S(Q)$ for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (downward triangles), in 20 mM citrate buffer at pH 4.6. For clarity, only one experimental point in six is displayed.

3.2.2 Lysozyme in buffer solution with glycerol

Glycerol has been demonstrated to be an effective enhancer of protein structural order and stability [Priev et al., 1996; Gekko and Timasheff, 1981 and Herberhold et al., 2004]. The mechanism proposed for this phenomenon is mainly the preferential hydration of the protein, which means that the cosolvent molecules are thermodynamically favored to be preferentially excluded from the domain of the protein molecule. In terms of

competitive interactions for protein or water, glycerol, being strongly hydrophilic, interacts with water more strongly and its affinity for polar residues on the protein surface is weaker, thus leading to preferential hydration of the protein. Furthermore, glycerol exhibits a stronger steric hindrance as compared to water molecules to interact at closest distances with the protein, so that the volume fraction occupied by the cosolvent at the surface of the protein is less compared to the bulk solvent, which again leads to preferential hydration of the protein. As a result, due to the strong hydration of the protein molecules, a decrease of the intermolecular attractive interaction energies may be expected.

In order to study the effects of glycerol on protein intermolecular interactions, two different glycerol concentrations, 500 mM and 2 M, were selected for each set of lysozyme protein concentration (4 and 10 wt%). The Debye-Hückel screening length is very similar (Table 3.1) as for the protein solution in pure 20 mM citrate buffer solution, and the dielectric permittivity of the medium decreases to 78.13 and 74.90 for 500 mM and 2 M glycerol, respectively. Such decrease of ε itself should lead to an enhancement of the attractive interaction. The SAXS intensity plots for different lysozyme concentrations with 500 mM glycerol concentration are shown in Fig. 3.12. The scattering intensity decreases slightly at small Q-values for both lysozyme concentrations with added glycerol as compared to the protein in pure buffer solution. In fact, the lower intensity values at $Q \rightarrow 0$ indicate a lowering of the attractive part of the interaction potential, which is already largely repulsive in the case of the protein in pure buffer solution, thus leading to enhanced structural short-range ordering. From the measured $I(Q)$ and $P(Q)$ data, the structure factors were calculated and fitted to the theoretical model (Fig. 3.13 , Table 3.1). The modulus of the depth of the attractive interaction potential J decreases slightly from 3.75 k_BT to 3.5 k_BT in the presence of 500 mM glycerol for the 4 wt% lysozyme solution. A more significant decrease was observed for the 10 wt% lysozyme solution: from 2.5 k_BT to 1.5 k_BT in the presence of 500 mM glycerol. These experimental findings are consistent with the concept of preferential hydration. The range of the interaction potential, d, increases from 4.5 Å and 5.75 Å to 5.25 Å and 11.5 Å in the presence of 500 mM glycerol for the 4 and 10 wt% lysozyme solutions, respectively.

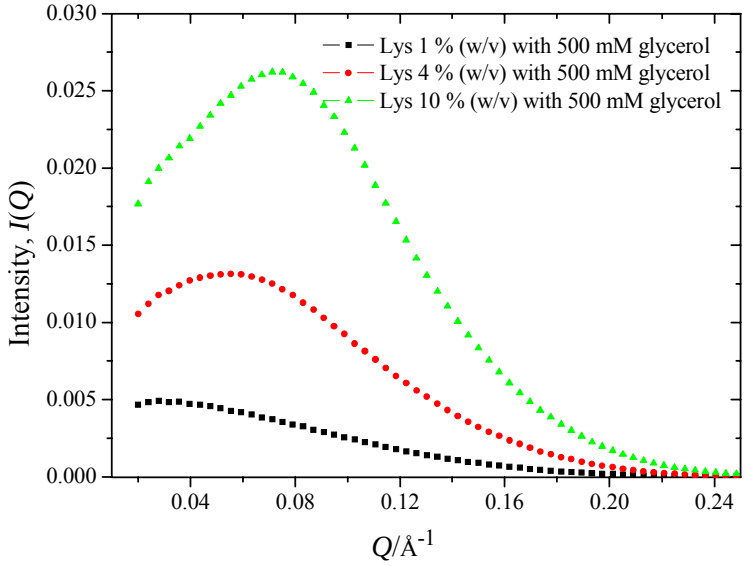


Figure 3.12: Plot of the SAXS intensity $I(Q)$ in arbitrary units against the momentum transfer for different lysozyme concentrations in wt%: 1 (squares), 4 (circles), 10 (triangles) with 500 mM glycerol in 20 mM citrate buffer at pH 4.6, measured at the DELTA synchrotron source.

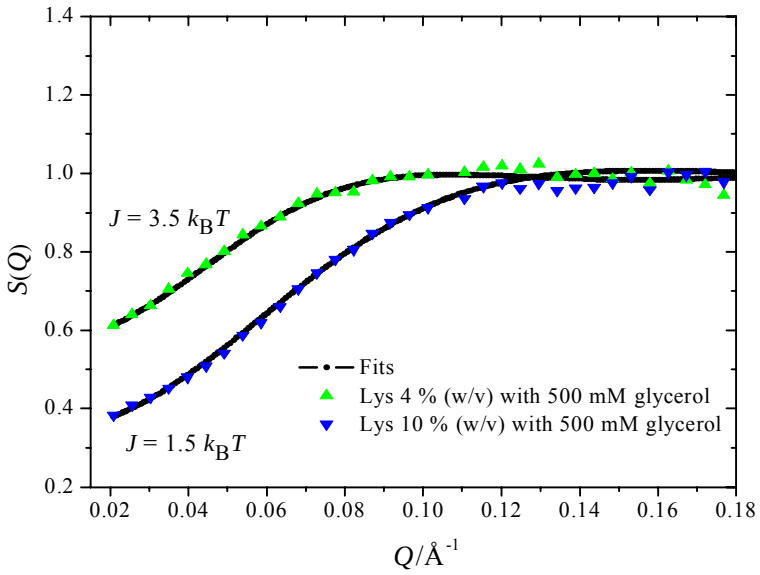


Figure 3.13: Experimental and calculated (black in full lines shown as fits) structure factors for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (downward triangles) with 500 mM glycerol in 20 mM citrate buffer at pH 4.6. For clarity only one experimental point in six is displayed.

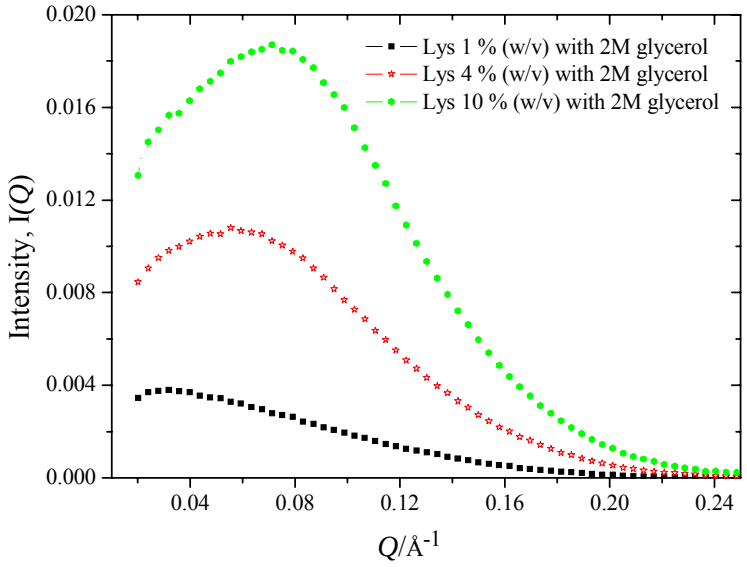


Figure 3.14: Plot of SAXS intensity $I(Q)$ in arbitrary units against the momentum transfer for different lysozyme concentrations in wt%:1 (squares), 4 (stars), 10 (circles) with 2 M glycerol in 20 mM citrate buffer at pH 4.6 measured at the DELTA synchrotron source.

In liquids with pronounced short-range order, the range of the interaction potential, d, is roughly equal to the average nearest neighbor distance between the molecules [Hansen and McDonald, 1986]. These experimental findings are consistent with the concept of preferential hydration. As revealed by pressure perturbation calorimetric (PPC) experiments of Ravindra and Winter [2003], the strength of the hydration layer of the protein increases at higher glycerol concentrations of 1.5 and 3.5 M, and the structural stability is enhanced as observed in the case of SNase [Herberhold et al., 2004] and lysozyme [Back, et al., 1979].

The SAXS intensity plots for different lysozyme concentrations with 2 M glycerol are shown in Fig. 3.14. The further lowering of the scattering intensities at small momentum transfer indicates a further increase of repulsive interaction. Analysis of the structure factor data (Fig. 3.15) yields depth of the interaction potential, J, values of 2.5 k_BT and 1.75 k_BT for lysozyme concentrations of 4 and 10 wt%, respectively. The effect of lowering of the attractive interaction by addition of the osmolyte glycerol hence nullifies a small increase in J expected due to the decrease of the dielectric permittivity of the

medium. The *d*-values and hence average intermolecular distances also increase, yielding corresponding values of 5.75 and 11 Å, respectively.

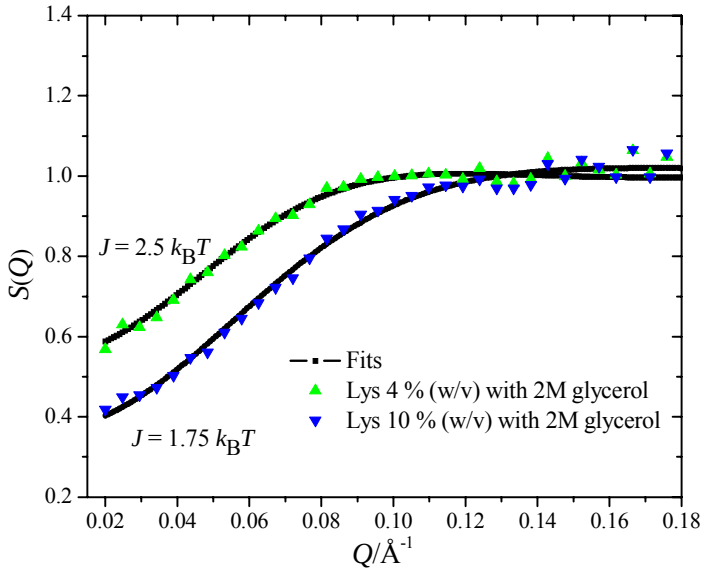


Figure 3.15: Experimental and calculated (black in full lines shown as fits) structure factors for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (downward triangles) with 2 M glycerol in 20 mM citrate buffer at pH 4.6. For clarity only one experimental point in six is displayed.

Farnum and Zukoski [1999] reported similar trends of the effect of glycerol on the intermolecular interaction of bovine pancreatic trypsin inhibitor using static light scattering.

3.2.3 Lysozyme in buffer solution with sucrose

As effective natural osmolyte, addition of sucrose is expected to have similar effects on the intermolecular interaction of proteins, although a different concentration dependence of the hydration properties has been observed in thermodynamic measurements by Ravindra and Winter [2004]. The dielectric permittivities of the solutions with 500 mM and 1 M sucrose solutions were determined to be 75.4 and 73.1, respectively. The

scattering patterns for different lysozyme concentrations in the presence of 500 mM sucrose are depicted in Fig. 3.16. As expected, sucrose exhibits a stronger intermolecular correlation peak with lower intensities at low Q values. Experimental and calculated structure factors are shown in Fig. 3.17 together with the results for the best fits.

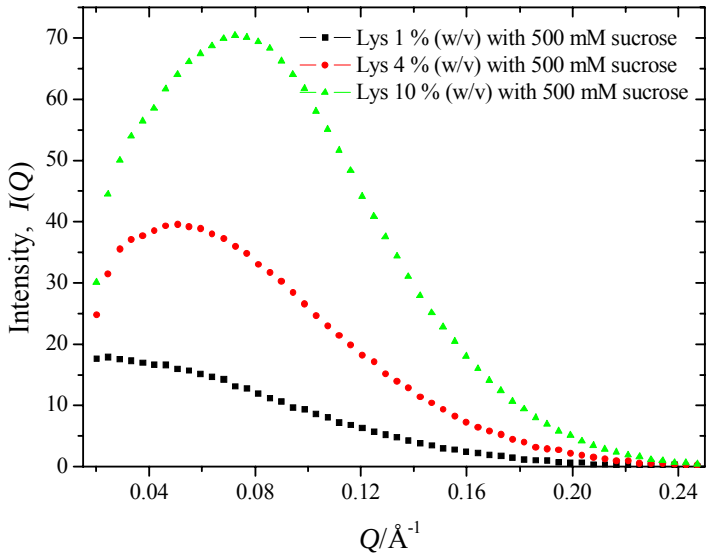


Figure 3.16: Plot of SAXS intensity $I(Q)$ in arbitrary units against the momentum transfer for different lysozyme concentrations in wt%: 1 (squares), 4 (circles), 10 (triangles) with 500 mM sucrose in 20 mM citrate buffer at pH 4.6 measured at the DESY synchrotron source.

The attractive part of the interaction potential increases to a value of -2.25 k_BT for the 4 wt% and to -1.5 k_BT for the 10 wt% protein solution. Like glycerol, the d-values increase to 10.25 and 10.5 Å for the 4 and 10 wt% lysozyme solution, respectively. The slightly stronger repulsive effect of the cosolvent sucrose can be explained in terms of a greater steric exclusion due to the larger size of the sucrose molecule as compared to glycerol, thus resulting in stronger preferential hydration. A similar conclusion has been drawn from thermodynamic measurements on RNase A [Ravindra and Winter, 2004].

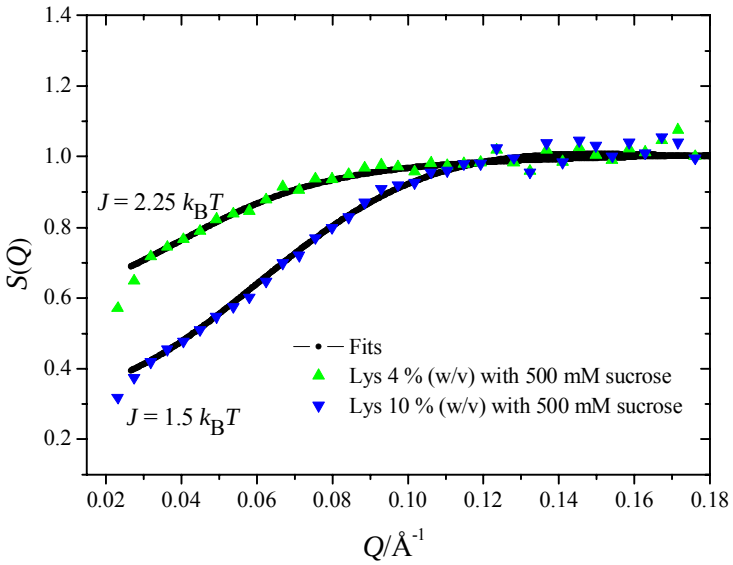


Figure 3.17: Experimental and calculated (black in full lines shown as fits) structure factors for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (downward triangles) with 500 mM sucrose in 20 mM citrate buffer at pH 4.6. For clarity only one experimental point in 20 is displayed.

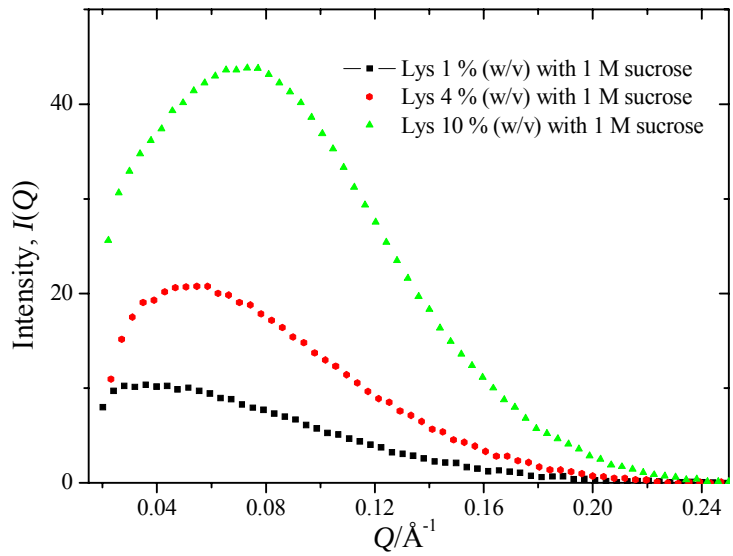


Figure 3.18: Plot of SAXS intensity $I(Q)$ in arbitrary units against the momentum transfer for different lysozyme concentrations in wt%: 1 (squares), 4 (circles), 10 (triangles) with 1 M sucrose in 20 mM citrate buffer at pH 4.6 measured at the DESY synchrotron source.

The scattering intensity patterns for lysozyme solutions of different concentration in 1 M sucrose are shown in Fig. 3.18 and the experimental structure factors with best fitting results are depicted in Fig. 3.19. The *J*-value of the attractive interaction potential is 2.5 k_BT for the 4 wt% solution and 1.25 k_BT for the 10 wt% lysozyme solution, with *d*-values of about 9.25 Å for the 4 wt% and 16.25 Å for the 10 wt% solution. Hence, upon increasing the concentration of sucrose from 500 mM to 1 M, *J* remains essentially constant or increases slightly, which is at first sight unexpected, but may be due to the decrease of the hydration capacity at high sucrose concentrations, as shown by Ravindra and Winter [2004] for sucrose concentrations around 1.5 to 2.5 M.

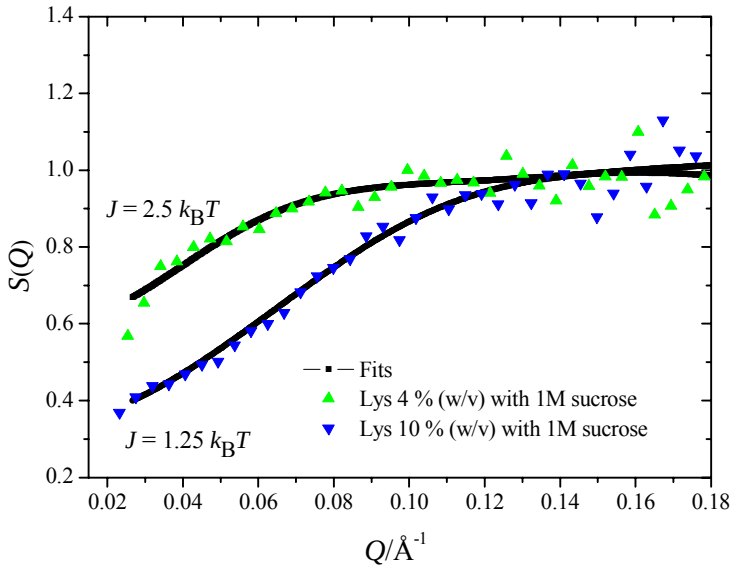


Figure 3.19: Experimental and calculated (black in full lines shown as fits) structure factors for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (downward triangles) with 1 M sucrose in 20 mM citrate buffer at pH 4.6. For clarity only one experimental point in 20 is displayed.

3.2.4 Lysozyme in buffer solution with tri-fluoroethanol (TFE)

It has been generally observed that alcohols promote formation of new secondary structures, in particular formation of α-helices, but at the same time they act as

destabilizers of tertiary and quaternary interactions within the folded protein, occasionally leading to partially unfolded "molten globule"-like conformations [Hirota et al., 1998]. However, their - often more-phasic - modes of action are still a matter of debate. Alcohols exhibit a lower dielectric constant than water and are much weaker hydrogen bond acceptors. They give rise to a series of proposed interaction mechanisms, including perturbation of the protein's water shell [Buck, 1998 and Walgers et al., 1998], diminution of intramolecular hydrophobic interactions, strengthening of intra-protein hydrogen bonding and less shielding of electrostatic interactions due to the low dielectric permittivity of the alcohols. Other scenarios stress preferential binding with simultaneous dehydration in the immediate vicinity of the protein, [Thomas and Dill, 1993 and Bull and Breese, 1978] and the impact of clustering to provide local regions of low polarity. [Hong et al., 1999]. Often the effects of alcohols on protein folding have been shown to depend on the alcohol concentration in a non-monotonic fashion [Cinelli et al., 1997].

The scattering intensity plots and their corresponding structure factors of different concentrations of lysozyme solutions (4 and 10 wt %) in 10 % (v/v) TFE with best fits are presented in Fig. 3.20 and Fig. 3.21, respectively. The experimentally determined dielectric permittivity of the medium is 77, slightly smaller than that of the pure buffer medium (79.8). The depths of the attractive interaction potential J are 2 k_BT and 1.5 k_BT, with d-values of 7 Å and 13.7 Å for the 4 and 10 wt% protein solution, respectively. At low concentrations, TFE behaves like a stabilizer or osmolyte (like glycerol and sucrose) with regards to its impact on the form of $V(r)$.

At higher TFE concentrations, at 35 % (v/v), the scenario changes (Fig. 3.22). The dielectric permittivity of the medium drops drastically to 58.8 and hence the hydrophilicity of the solvent decreases drastically. At these high TFE concentrations, the protein molecules start to change their conformation and more hydrophobic residues of the protein molecules are exposed to the solvent. The radius of gyration, as mentioned in section 3.1.2, increases slightly and the pair distribution function indicates formation of a more elongated structure of the lysozyme molecule under these solvent conditions. The helical content of the protein increases at high TFE concentrations and the spherical shape approximation of the protein molecule is no more valid.

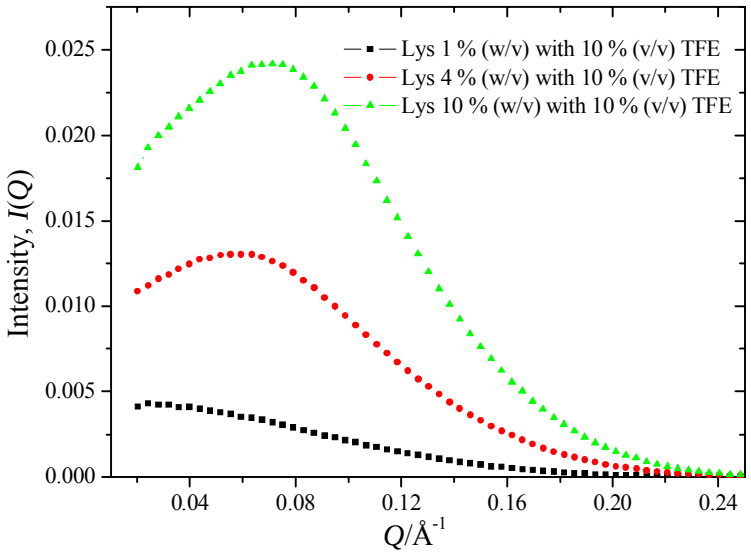


Figure 3.20: Plot of SAXS intensity $I(Q)$ in arbitrary units against the momentum transfer for different lysozyme concentrations in wt%: 1 (squares), 4 (circles), 10 (triangles) with 10 % (v/v) TFE in 20 mM citrate buffer at pH 4.6 measured at the DELTA synchrotron source.

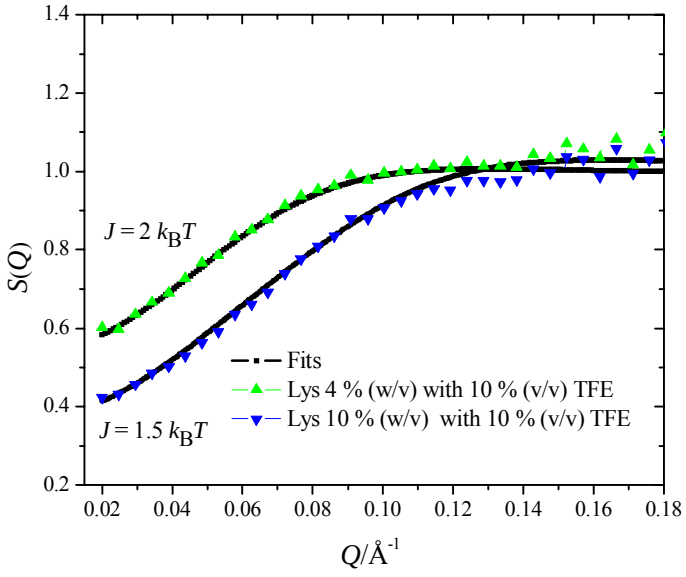


Figure 3.21: Experimental and calculated (black in full lines shown as fits) structure factors for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (downward triangles) with 10 % (v/v) TFE in 20 mM citrate buffer at pH 4.6. For clarity only one experimental point in 6 is displayed.

Hence, as expected, the fits are not as good as for the other cosolvents studied, but at least qualitative trends can be discussed (Fig. 3.23). The depths of $V(r)$ are 7.75 k_BT and 5.25 k_BT for lysozyme at concentrations of 4 and 10 wt%, respectively. The corresponding range of attractive interactions, d, is 2.6 and 4.35 Å, respectively. This observed increase in attractive interaction can be explained in terms of a perturbation of the protein's hydration shell and increasing hydrophobic interactions between the protein molecules by interaction of aromatic residues or other non-polar exposed side chains.

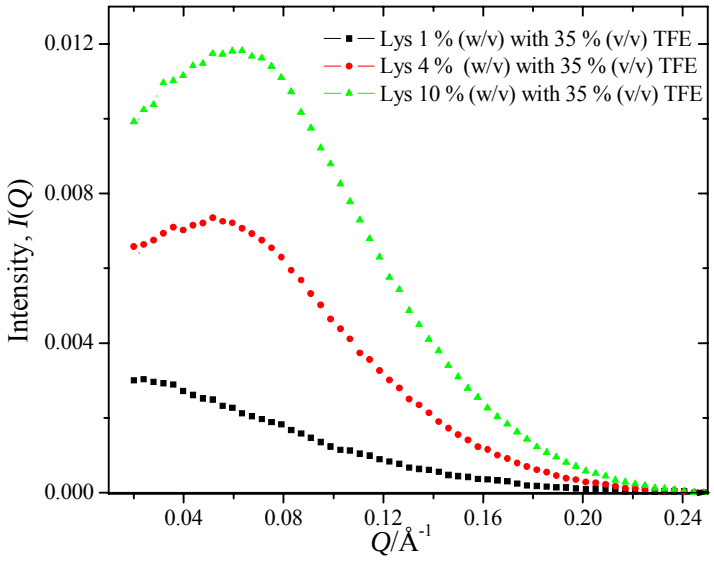


Figure 3.22: Plot of SAXS intensity $I(Q)$ in arbitrary units against the momentum transfer for different lysozyme concentrations in wt%: 1 (squares), 4 (circles), 10 (triangles) with 35 % (v/v) TFE in 20 mM citrate buffer at pH 4.6 measured at the DELTA synchrotron source.

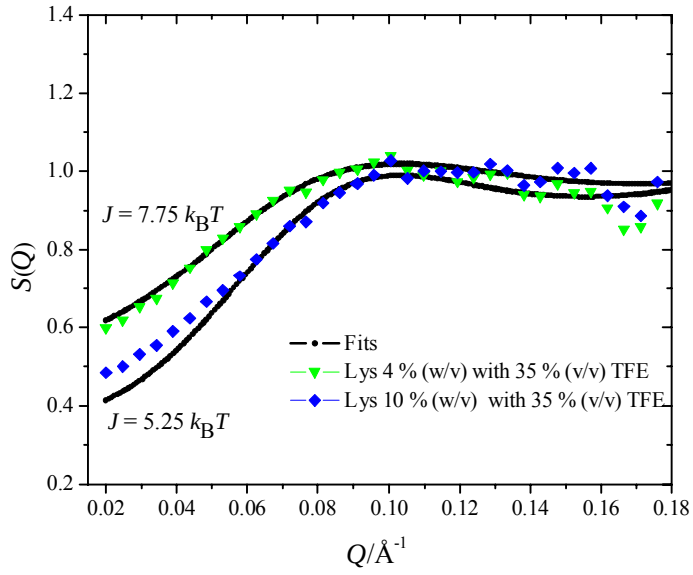


Figure 3.23: Experimental and calculated (black in full lines shown as fits) structure factors for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (diamonds) with 35 % (v/v) TFE in 20 mM citrate buffer at pH 4.6. For clarity only one experimental point in 6 is displayed.

3.2.5 Lysozyme in buffer solution with sodium chloride (charge screening effect)

Charge screening is considered to be a major factor governing the stabilization of proteins [Ravindra and Winter, 2003] and is also of relevance for protein crystallization studies [Rosenberger et al., 1996]. A series of scattering experiments was performed as a function of salt concentration for the 1 wt% and 4 wt% lysozyme concentrations. The measured scattering intensity $I(Q)$ increases drastically at low Q-values with increasing salt concentration as shown in Fig. 3.24 finally leading to the disappearance of the correlation peak. This increase in the scattering intensity at low momentum transfers indicates that repulsive interactions between the protein molecules are gradually lost and attractive interactions dominate, thus leading to a loss of the short-range solution structure.

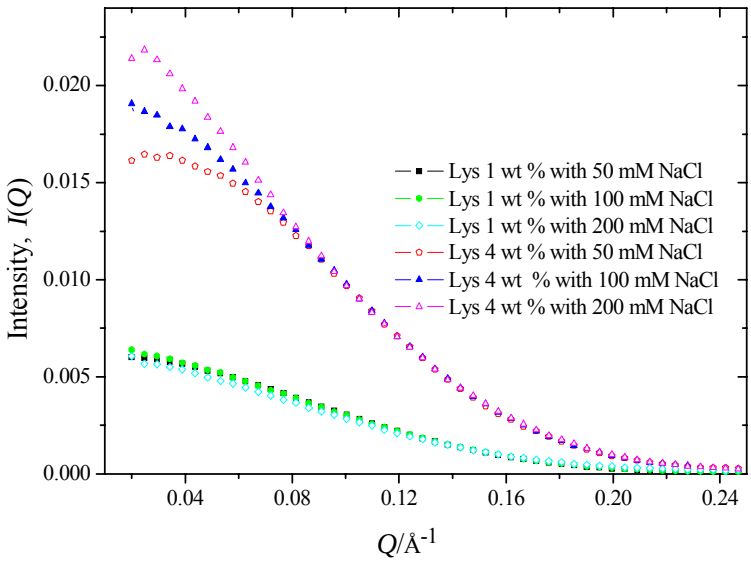


Figure 3.24: Plot of the SAXS intensity $I(Q)$ in arbitrary units as a function of momentum transfer Q for different concentrations of lysozyme in 20 mM citrate buffer at pH 4.6. Lysozyme 1wt %: with 50 mM sodium chloride (squares), with 100 mM sodium chloride (circles), with 200 mM sodium chloride (diamonds). Lysozyme 4 wt%: with 50 mM sodium chloride (pentagons), with 100 mM sodium chloride (filled upward triangles) and with 200 mM sodium chloride (blank upward triangles).

This increase in the scattering intensity at low momentum transfers indicates that repulsive interactions between the protein molecules are gradually lost and attractive interactions dominate, thus leading to a loss of the short-range solution structure.

The experimental and calculated structure factors with their best fits are shown in Fig. 3.25. The dielectric permittivity of the medium was 79.76 and the Debye-Hückel screening lengths were calculated according to the salt concentration used. The experimental $S(Q)$ data were fitted with an effective charge of $Z = 6$ for the protein solution in pure buffer and 50 mM salt concentration, for the higher salt concentrations of 100 and 200 mM, a minor change in charge ($Z = 7$) and varying the depth and range of $V(r)$ resulted in better fits. A continuous increase in the depth of the interaction potential and a concomitant decrease of its range is observed with increasing NaCl concentration

(Table 3.1): The *J*-values of the potential well increases from 3.75 k_BT for 0 mM NaCl to 4.75 k_BT for 50 mM, 8.25 k_BT for 100 mM, to 13 k_BT for 200 mM added NaCl. Conversely, the range of the attractive potential decreases from 4.5 to 3.5, 2.25 and 1.75 Å in the similar order of salt concentration.

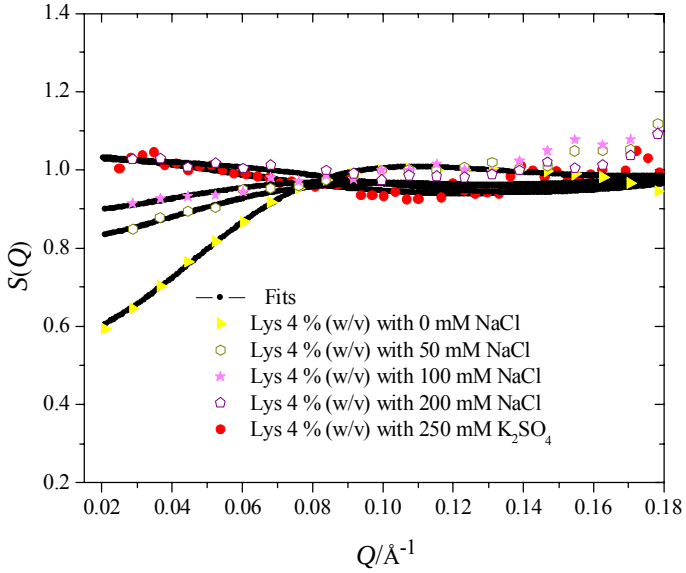


Figure 3.25: Experimental and calculated (full lines in black) structure factors *S*(*Q*) for lysozyme at a concentration of 4 wt % in 20 mM citrate buffer with different concentrations of salts. No salt (triangles), sodium chloride 50 mM (hexagons), sodium chloride 100 mM (stars), sodium chloride 200 mM (pentagons) and potassium sulphate 250 mM (circles).

These systematic trends can be explained by the fact that, as the salt concentration increases, the positive charge on the protein molecule is screened by the salt anions, hence diminishing repulsive protein-protein interactions. Intermolecular distances between the protein molecules start decreasing, leading to the shorter range of the attractive part of the interaction potential. Similar trends were reported by Narayanan and Liu [2003]. At high ionic strengths, an increase in van der Waals attractive - interactions

and an incremental contribution to attraction due to osmotic excluded volume effects of the salt anions explain the increase of the well-depth of $V(r)$.

3.2.6 Lysozyme in buffer solution with potassium sulphate

Hofmeister anions, such as sulphate (SO_4^{2-}), are considered to be strong protein stabilizers and the underlying principle, the reduction of the net positive charge on the

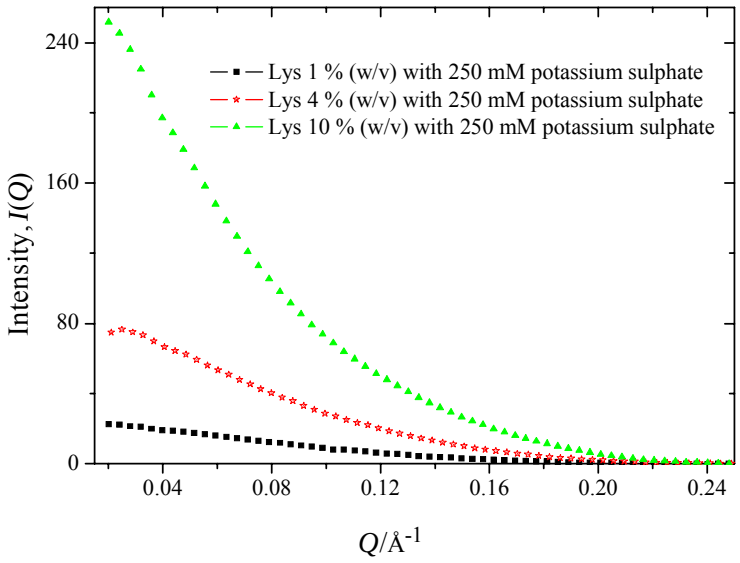


Figure 3.26: Plot of the SAXS intensity $I(Q)$ in arbitrary units as a function of momentum transfer Q for different lysozyme concentrations in wt%: 1 (squares), 4 (stars) and 10 (triangles) with 250 mM potassium sulphate in 20 mM citrate buffer at pH 4.6.

protein surface through charge screening is similar, but more effective owing to the higher negative charge of the sulphate anion. The scattering intensity plots for 1 - 10 wt% lysozyme solutions in 250 mM K_2SO_4 are shown in Fig. 3.26. The strong intensity increase of $I(Q)$ for $Q \rightarrow 0$ suggests a highly attractive nature of $V(r)$. The dielectric constant of the medium decreases to 76.1, which is also favoring attraction between the protein molecules. The experimental structure factor together with the calculated one for the 4 wt % lysozyme solution is shown in Fig. 3.25.It was not possible to fit the structure

factors for the 10 wt % lysozyme solution which may be due to formation of higher order oligomers where the model used is not applicable anymore. The magnitude of the well depth for the 4 wt% lysozyme solution with 250 mM K_2SO_4 is found to be 13 k_BT (Table 3.1), which is similar to the data obtained for 200 mM NaCl. The range of the attractive potential is lowered (d = 1.01 Å), intuitively due to stronger charge screening. In the case of higher protein concentrations (10 wt%, data not shown), $J \approx 18\ k_BT$. The fit is not very satisfactory, however, in particular in the higher Q-range, which may be the result of oligomer formation of the protein molecules in this strongly attractive regime, where the limitation of our model is reached.

3.2.7 Lysozyme in buffer solution with guanidinium chloride

Chemical denaturation with an agent such as urea or guanidinium chloride (GdmCl) is one of the primary ways of assessing protein stability, the effects of mutations on stability and protein un- and re-folding [Fresht, 1999]. It is well known that the destabilizing nature of this agent is mainly caused by its preferential binding to the peptide groups [Bennion and Daggett, 2003 and Makhatadze, 1999]. Recently, the crystal structure of PhoCutA, a heavy metal binding protein, was determined by Tanaka et al. [2004], at 1.6 Å resolution, in the presence of 3 M GdmCl. They found that due to the interaction between the protein and salt, 10 % of the intramolecular hydrogen bonds which are located on the protein surface, especially around the structurally perturbed loops, disappear, again pointing to a strong interaction between the peptide groups and GdmCl. In the crystal structure, few Gdm^+ ions were observed, showing interactions of tryptophan amino acid residues with guanidinium ions by CH-π interactions. Mande and Sobhia [2000] found interactions of Gdm^+ ions with the π-electron cloud of Trp108, His15 and Ile 88 residues as well as water mediated interactions with the Arg14 residue in the crystal structure of hen egg-white lysozyme [Mande and Sobhia, 2000].

In order to study the effect of the chaotropic agent GdmCl on the intermolecular interaction potential of lysozyme, two concentrations of GdmCl (500 mM and 3.5 M) were selected. Intensity plots for lysozyme concentrations at 1, 4 and 10 wt% with the two different GdmCl concentrations are shown in Fig. 3.27. The radius of gyration of

lysozyme in the presence of 500 mM GdmCl as obtained from the Guinier analysis of the $I(Q)$ data is 15.3 ± 0.2 Å, and R_g increases to 16.4 ± 0.5 Å in presence of the 3.5 M GdmCl, indicating partial unfolding of the protein molecule under these cosolvent conditions.

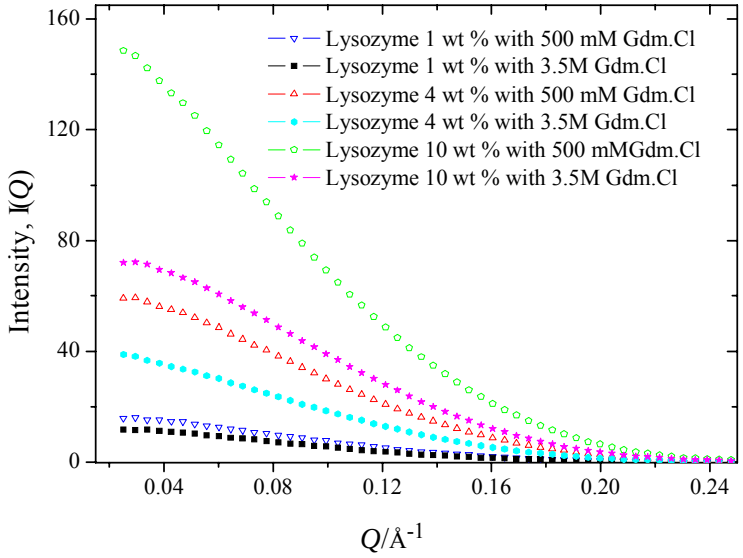


Figure 3.27: Plot of the SAXS intensity $I(Q)$ in arbitrary units as a function of momentum transfer Q for different concentrations of lysozyme in wt% with 0.5 and 3.5 M GdmCl: Lysozyme 1 wt% (downward triangles, squares), 4 wt% (upward triangles, hexagons), 10 wt% (pentagons, stars) in 20 mM citrate buffer at pH 4.6 measured at DESY.

The dielectric permittivity of the medium with 500 mM and 3.5 M GdmCl was found to be 71.3 and 51.7, respectively. The scattering intensities of the GdmCl containing protein solutions at low momentum transfers are lower as compared to the 250 mM salt K_2SO_4 for the respective protein concentrations, indicating a less pronounced salt screening effect. Experimental structure factors were calculated from the data (Fig. 3.28) and the best fits yielded a J value of about 10.5 k_BT for the 4 and 10 wt% lysozyme concentrations in the presence of 500 mM GdmCl, which is slightly smaller as compared to the 250 mM K_2SO_4 solution (13 k_BT for a lysozyme concentration of 4 wt%).

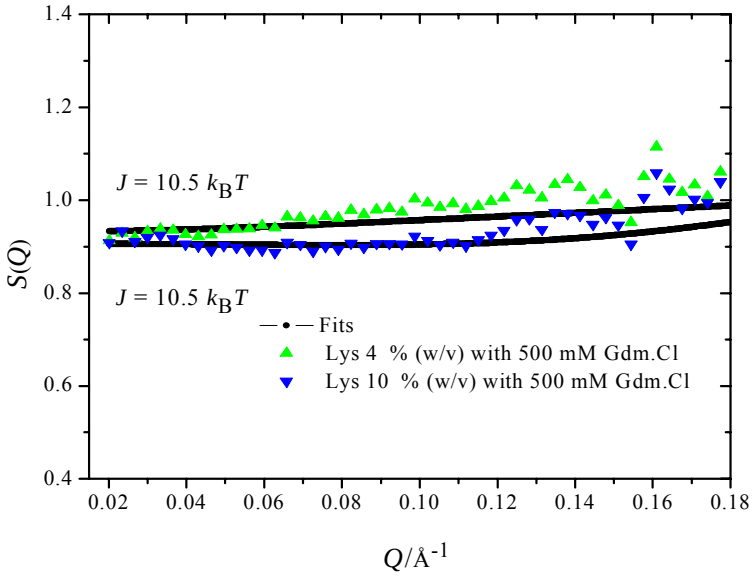


Figure 3.28: Experimental and calculated (full lines in black) structure factors for different lysozyme concentrations in wt%: 4 (upward triangles), 10 (downward triangles) with 0.5 M GdmCl in 20 mM citrate buffer at pH 4.6. For clarity only one experimental point in 20 is displayed.

Recent PPC data show that the hydration level of the protein molecules in the presence of 500 mM GdmCl decreases substantially as compared to the protein in pure buffer solution [Priev et al., 1996, and Ravindra and Winter, 2004]. For the higher GdmCl concentration (3.5 M, Fig. 3.29), the applied potential model does not describe the experimental structure factor quantitatively anymore. The structure factors at low *Q*-values move further to the repulsive regime, which is more pronounced at a protein concentration of 10 wt%. This increased repulsion at the higher cosolvent concentration may arise from the increased concentration of adsorbed positively charged guanidinium ions or/and it may be due to the fact that at higher concentrations of GdmCl, hydration of the protein starts to increase due to partial unfolding of the protein as suggested by Ravindra and Winter [2003], which also favors lowering of attractive interactions.

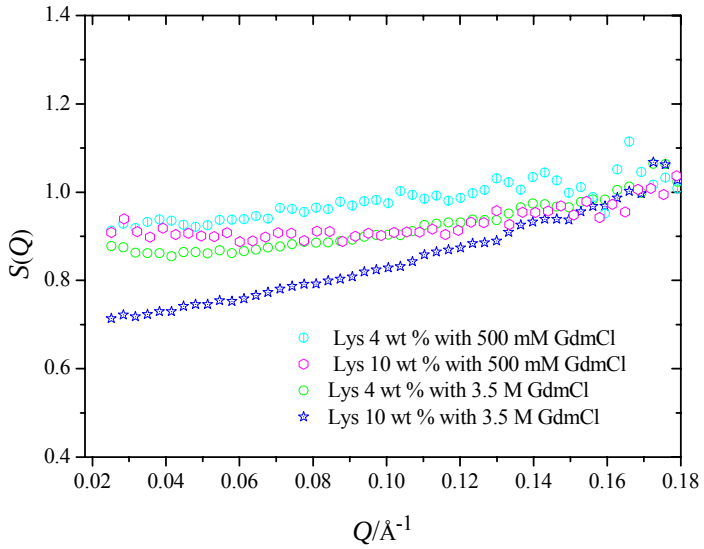


Figure 3.29: Experimental structure factors for different concentrations of lysozyme in with 500 mM and 3.5 M guanidine hydrochloride: 4 wt % (circles with line, circles) and 10 wt % (hexagons and stars).

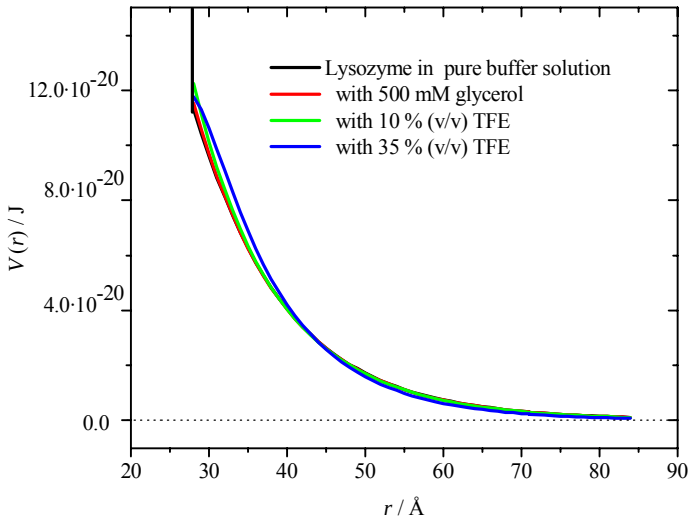
3.2.8 Concluding remarks

The overall parameters of intermolecular interactions among the lysozyme molecules under different cosolvent conditions are given in Table 3.1. Generally, the osmolytes like glycerol and sucrose, in different concentrations decrease the attractive part of the interaction potential, TFE has concentration dependent effects on the attractive interaction potential and salts can change the interactions towards a highly attractive regime by enhancing the salt concentration due to charge screening effects. The total interaction potentials of lysozyme with some representative cosolvents and salts are shown in Fig. 3.30. Different cosolvents and osmolytes like glycerol, sucrose and TFE modify the attractive part of interaction potential, but the total interaction potentials are still largely repulsive in nature, while the salts exhibit concentration dependent effects which can modulate these interactions from the repulsive regime to a highly attractive one.

Samples	Lysozyme 4 % (w/v)		Lysozyme 10 % (w/v)		
	J / k_BT	d / Å	J / k_BT	d / Å	κ^{-1} / Å
Pure lysozyme in 20 mM citrate buffer	3.75	4.5	2.5	5.75	15.4
Lysozyme in 20 mM citrate buffer with 500 mM glycerol (Sucrose 500 mM)	3.5 (2.25)	5.25 (10.25)	1.5 (1.5)	11.5 (10.5)	15.2
Lysozyme in 20 mM citrate buffer with 2 M glycerol (Sucrose 1 M)	2.5 (2.5)	5.75 (9.25)	1.75 (1.25)	11 (16.25)	14.95 --
Lysozyme in 20 mM citrate buffer with 10 % (v/v) TFE	2	7	1.5	13.7	15.16
Lysozyme in 20 mM citrate buffer with 35 % (v/v) TFE	7.75	2.6	5.25	4.35	13.25
Lysozyme in 20 mM citrate buffer with 50 mM NaCl	4.75	3.5	-	-	10.25
Lysozyme in 20 mM citrate buffer with 100 mM NaCl	8.25	2.25	-	-	8.2
Lysozyme in 20 mM citrate buffer with 200 mM NaCl	13	1.75	-	-	6.28
Lysozyme in 20 mM citrate buffer with 250 mM K_2SO_4	13	1.01	18	1.01	3.4
Lysozyme in 20 mM citrate buffer with 500 mM GdmCl	10.5	0.36	10.5	0.51	3.98

Table 3.1: Fitting parameters of $S(Q)$ for lysozyme solutions with and without added cosolvents: J (in k_BT) is the depth of attractive interaction potential $V_Y(r)$ and d (in Å) is the range of the potential. K is the Debye-Hückel screening length.

a)



b)

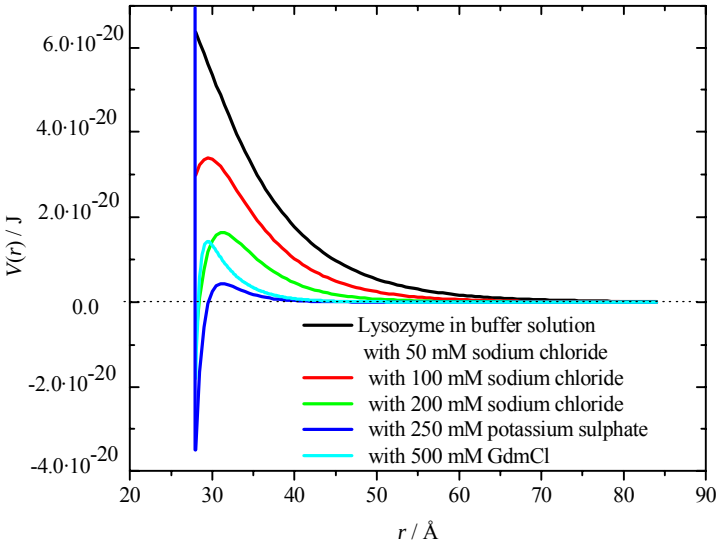


Figure 3.30: Plot of the total intermolecular interaction potential $V(r)$ of 4 wt% lysozyme at pH 4.6 in pure buffer solution as well as in selected cosolvent mixtures (a) as well as salt solutions (b).

3.3 Insulin intermolecular interactions

3.3.1 Form factors

Scattering intensity patterns for 0.5 wt % insulin solutions in water along with added 20 % (v/v) ethanol and 100 mM sodium chloride at pH 2.0 were measured for use as form factor $P(Q)$ for calculating the interaction potentials. Insulin molecules can self assemble into different forms like dimeric, tetrameric and hexamer species in aqueous solutions, depending upon the interactions. These interactions also depend on the concentration of the protein. The minimum possible concentration of insulin was chosen to be a 0.5 wt % in order to avoid interactions and to obtain still statistically accurate scattering data. The scattering intensities after normalization to the incident beam and background subtraction are depicted in Fig. 3.31.

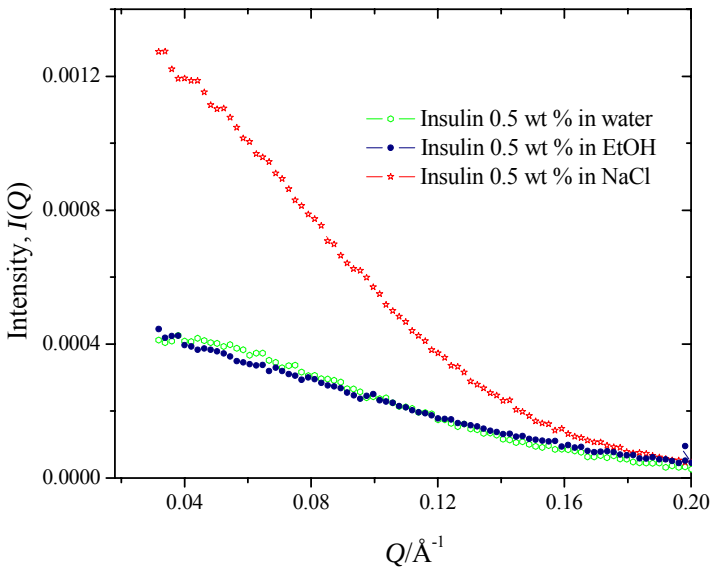


Figure 3.31: Scattering intensities of insulin (0.5 wt%) under various solvent conditions. Insulin in water at pH 2.0 (hexagons), insulin with 20 % (v/v) ethanol at pH 2.0 (circles) and insulin with 100 mM sodium chloride at pH 2.0 (stars).

The form factor analysis of insulin in water at pH 2.0 via Guinier's method yields a radius of gyration, R_g, of 14.3 Å, which matches the value reported for dimeric insulin [Grudzielanek et al., 2005], reflecting the well-known fact that insulin essentially consists of dimers under these experimental conditions.

From the insulin solution of concentration 0.5 wt% in ethanol, a radius of gyration of 11.9 Å is determined, which represents the monomeric form of the insulin, in agreement with data reported in the literature (R_g = 11.6 Å by Nielsen et al. [2001], and R_g = 11.8 Å by Grudzielanek et al. [2005, 2006]).

Guinier's analysis for 0.5 wt% insulin with added 100 mM sodium chloride yielded a radius of gyration of 16 ± 0.2 Å, which indicates that the insulin molecules are predominantly in the tetrameric state under these conditions. Additional light scattering experiments carried out under the same conditions confirmed these findings ($R_h \approx 17.5$ Å) and revealed a small population of dimers, only.

3.4 Interaction potential determination of insulin

3.4.1 Insulin solutions in water and comparison with lysozyme

The scattering intensity pattern $I(Q)$ of insulin at pH 2.0 at different concentrations ranging from 1 to 20 wt% are depicted in Fig. 3.32. The inset shows the scattering intensity for the 0.5 wt% insulin solution, which represents the particle form factor, $P(Q)$. Strong correlation peaks are observed at Q-values of ~0.07 $Å^{-1}$, indicating marked intermolecular protein interactions. The corresponding structure factors $S(Q)$, obtained by using Eq. 1.20, for the 1, 4, and 10 wt% insulin solutions are depicted in Fig. 3.33 along with the best fits calculated theoretically. Structure factors of insulin measured under the conditions which favor amyloidogenesis, differ from the structure factors of a non-aggregation-prone protein, lysozyme, which has been determined for the same concentrations, as displayed in Fig. 3.11. Qualitatively, a peak in $S(Q)$ at Q_{max} can be interpreted in terms of a "Bragg reflection" from planes of particles separated by a mean nearest-neighbor distance l, with $l = 2\pi/Q_{max}$. Typically, charged particles, such as charged proteins, at low ionic strengths maximize their average interparticle distance, and

so l depends on the concentration or volume fraction and one finds $Q_{max} = 2\pi n^{1/3}$, where n is the number density of particles. At the same time, the forward scattering, $I(Q \rightarrow 0)$, becomes suppressed with increasing particle concentration.

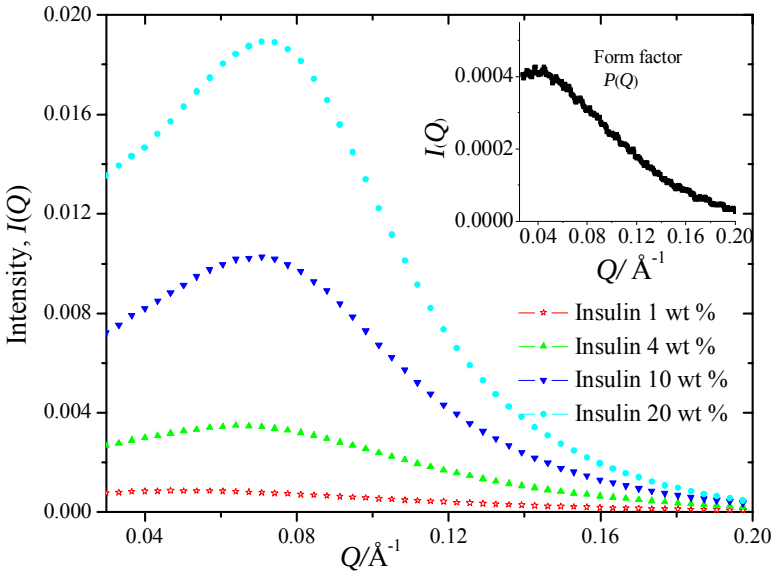


Figure 3.32: Plot of the SAXS intensity $I(Q)$ in arbitrary units as a function of momentum transfer Q for different concentrations of insulin in wt%: 1 (stars), 4 (upward triangles), 10 (downward triangles), 20 (circles) in water at pH 2.0 measured at the DELTA synchrotron source. The inset shows the form factor for dimeric insulin.

Surprisingly, however, the peak position Q_{max} of $S(Q)$ is essentially independent of the insulin concentration. The peak position appears at ~0.08 Å^{-1} for all insulin concentrations (corresponding to a real space dimension of ~80 Å), indicating that insulin particles self-assemble into clusters with a concentration-dependent aggregation number and cluster size distribution, instead of forming a homogeneous spatial distribution of the protein particles. The increase in amplitude of $S(Q)$ reflects an increase in protein cluster-cluster interactions. A similar scenario has been observed for other highly concentrated systems, like colloidal dispersions [Qiu et al., 2006].

The volume fraction or packing density, which determines the hard sphere reference structure factor $S_0(Q)$, is calculated by considering the size and total number of dimeric

insulin particles. The screening lengths are calculated by taking in account the concentrations of the acid required to adjust the pH to 2.0 by 1 M HCl. An effective charge Z of 11 (monomeric insulin: $Z = 6$) with effective hard sphere diameter σ of 23 Å yielded the best fits of the experimental structure factors. As insulin, being partially unfolded under these pH conditions (molten globule kind of state), is less compact, slightly elongated in shape and more flexible as compared to the globular lysozyme molecule [Grudzielanek et al., 2005] , the effective σ for best fits is slightly less than expected for a spherically rigid sphere.

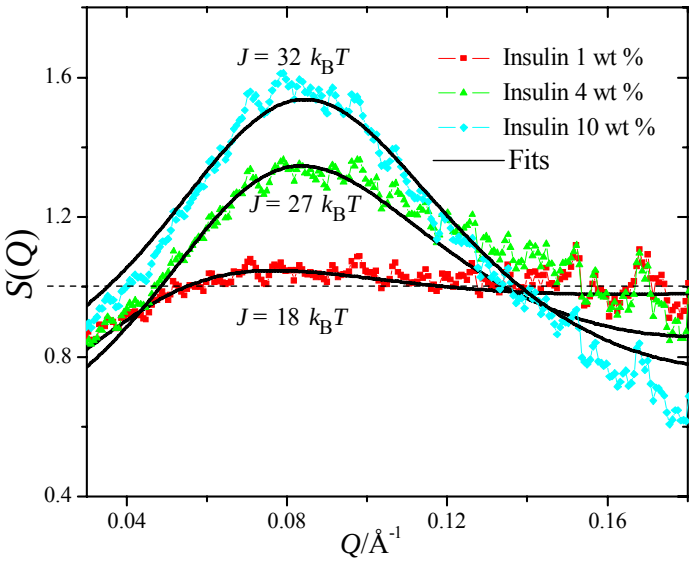


Figure 3.33: Experimental and calculated (full lines in black) structure factors for different concentrations of insulin in wt%: 1 (squares), 4 (upward triangles), 10 (diamonds) in water at pH 2.0 at room temperature.

The depths of the attractive intermolecular interaction potential, J, are 18, 27, and 32 k_BT with ranges d of 11, 4, and 2 Å for insulin concentrations of 1, 4 and 10 wt%, respectively. In contrast, for lysozyme, representing a rather stable protein, the attractive part of $V(r)$ is drastically (a factor of 7 to 13) smaller, and exhibits a different concentration dependence. In the case of the also positively charged lysozyme, the interaction becomes more repulsive with increasing protein concentration. In the case of

insulin, the opposite behavior is observed, indicating pronounced differences in the intermolecular forces between stable and aggregation-prone proteins.

The marked increase in the attractive interaction of insulin molecules, leading to equilibrium cluster formation shows that short-range van der Waals and probably specific hydrophobic attractions between the insulin molecules are strong already at rather low concentrations, and may be perceived as a collective struggle of the polypeptide chains to reduce their surface accessible area and shield hydrophobic groups in clusters, thus leading to a thermodynamically more stable state. However, as revealed by AFM and fluorescence spectroscopic measurements, under these temperature conditions, no stable compact amorphous or fibrillar species are formed [Jansen et al., 2005 and Grudzielanek et al., 2006].

3.4.2 Insulin solutions in water with sodium chloride (charge screening effects)

Charge screening plays a crucial role not only in the stabilization of proteins but also accelerates aggregation or fibrillation processes and is vastly used for protein crystallization studies [Grudzielanek et al., 2006, Stradner et al., 2004, Qiu et al., 2006 and Rosenberger et al., 1996]. In addition, changing charge screening probes the importance of electrostatic forces in the intermolecular interactions of aggregation prone proteins. Hence, an additional series of experiments was performed on insulin solutions of different concentration in the presence of 0.1 M NaCl, an ionic strength which is also of physiological relevance.

The measured scattering intensity $I(Q)$ increases drastically at low Q-values with increasing insulin concentration as shown in Fig. 3.34, and leads to the disappearance of the correlation peak. This monotonic increase in scattering intensity indicates the gradual loss of repulsive interactions and prevalence of short-range attractive interactions. In the presence of 0.1 M NaCl, the positively charged insulin molecules are effectively screened by Cl^- ions.

The corresponding experimental structure factors $S(Q)$ with best fits are shown in Fig. 3.35. The $S(Q)$ data could be fitted only by taking the predominant tetrameric form into account, with an effective charge Z of 24 and an effective σ value of 30 Å. A systematic

continuous increase in the depth of the attractive part of the interaction potential, $V_Y(r)$, and a concomitant decrease of the range of the potential is observed with increasing insulin concentration: the J values increase from 18 to 38 k_BT and from 27 to 50 k_BT for insulin concentrations of 1 and 4 wt%, respectively, as compared to the protein solution in pure acidic water, with ranges d of $V_Y(r)$ of 5 Å and 3 Å, respectively.

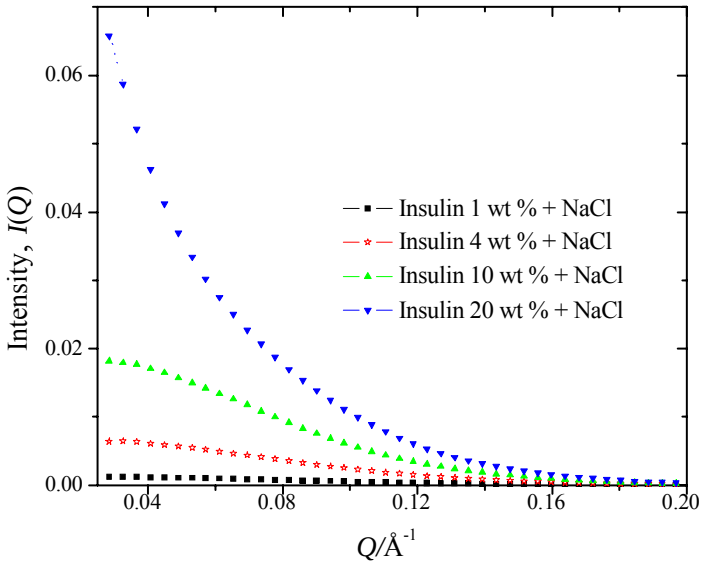


Figure 3.34: Plot of the SAXS intensity $I(Q)$ in arbitrary units as a function of momentum transfer Q for different concentrations of insulin in wt%: 1 (squares), 4 (stars), 10 (upward triangles), 20 (downward triangles) in water with 100 mM sodium chloride at pH 2.0 measured at the DELTA synchrotron source.

The fits for the higher insulin concentrations (10 and 20 wt %) are less good because the model used is not valid anymore as the solutions exhibit a large polydispersity of also larger oligomeric species under these conditions. $S(Q)$ at low Q-values increases drastically for the 4 and 10 wt% solutions, indicating increasing attractive interactions and formation of larger clusters. Owing to the effective screening of ionic changes, repulsive interactions are largely diminished.

The large increase of the attractive well depth, J, is probably due to a further increase in short-range van der Waals attractive interactions and specific hydrophobic interactions, and explains the enhanced rate of protein aggregation and fibrillation under charge-

screening conditions, observed on a shorter time-scale at at higher temperatures. For example, the half time of aggregation of insulin is 470 min in non-agitated buffer solution and 90 min in buffer containing 0.1 M NaCl at 60 ^{o}C [Grudzielanek et al., 2006].

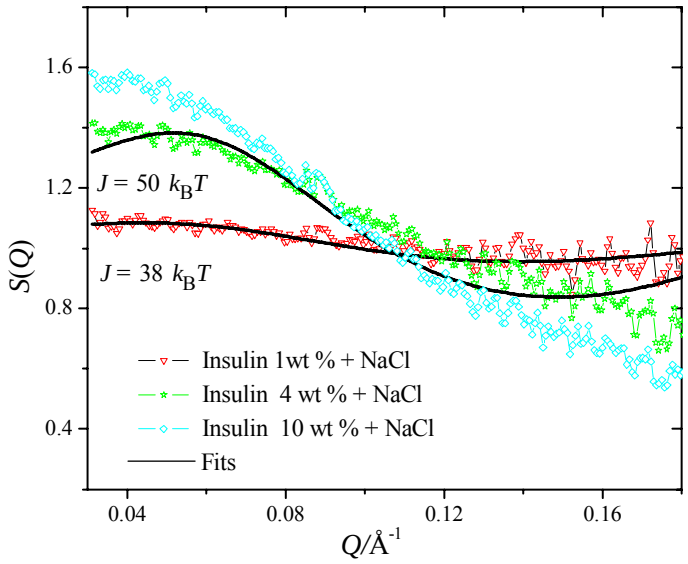


Figure 3.35: Experimental and calculated (full lines in black) structure factors for different concentrations of insulin in wt%: 1 (downwrad triangles), 4 (stars) and 10 (diamonds) in 100 mM sodium chloride at pH 2.0.

3.4.3 Insulin solutions in water with ethanol

Generally, alcohols give rise to a series of proposed interaction mechanisms including perturbation of the protein's water shell, diminution of intermolecular hydrophobic interactions, strengthening of intra-protein hydrogen bonding and less shielding of electrostatic interactions [Buck, 1998 and Walgers et al., 1998]. It has also been observed that alcohols may promote formation of new secondary structures, in particular the formation of α-helices, but at the same time act as destabilizers of tertiary or quaternary interactions within folded proteins [Hirota et al., 1998].

To reveal how the intermolecular interaction depends on the initial aggregation-prone species, additional experiments were carried out at conditions, where insulin exists as

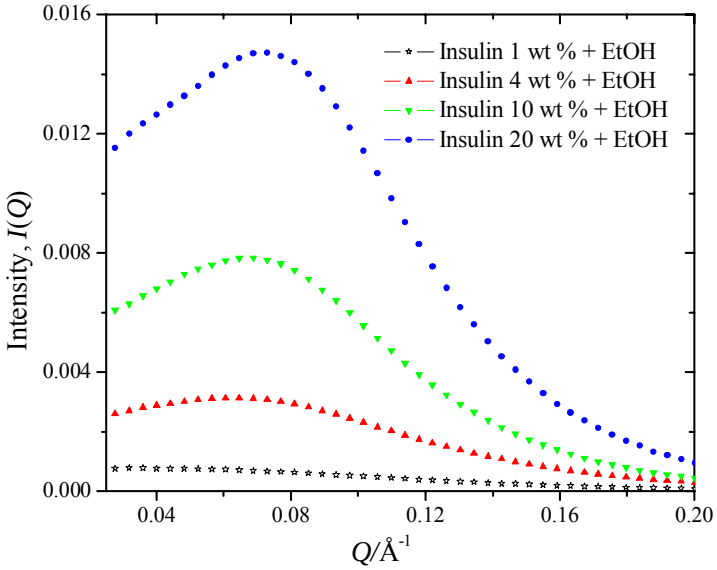


Figure 3.36: Plot of the SAXS intensity $I(Q)$ in arbitrary units as a function of momentum transfer Q for different concentrations of insulin in wt%: 1 (empty stars), 4 (upward triangles), 10 (downward triangles), 20 (circles) in 20 % (v/v) ethanol at pH 2.0 measured at the DELTA synchrotron source.

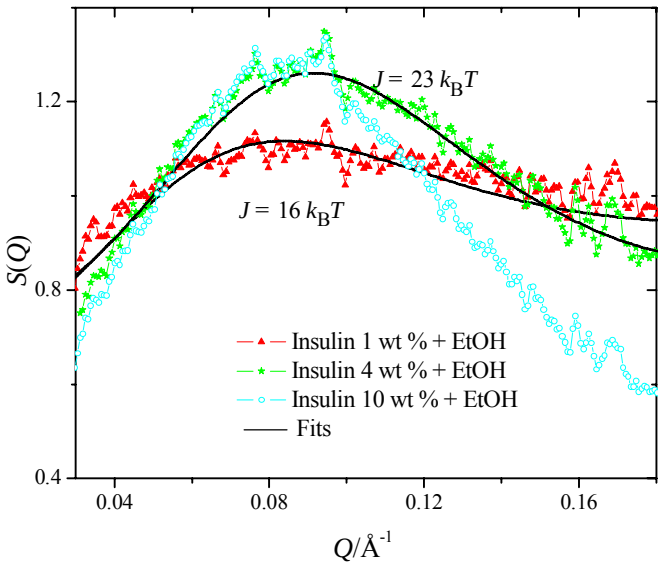


Figure 3.37: Experimental and calculated (full lines in black) structure factors for different concentrations of insulin in wt%: 1 (upwatrd triangles) and 4 (circles) in 20 % (v/v) ethanol at pH 2.0.

monomeric species, only. These conditions are realized at pH 2.0 in 20 wt% ethanol solutions [Grudzielanek et al., 2005 and Grudzielanek et al., 2006]. The scattering intensity plots and their corresponding structure factors for different concentrations of insulin (1, 4, 10 and 20 wt%) in 20% (v/v) ethanol with best fits are presented in Fig. 3.36 and in Fig. 3.37 respectively. The structure factors $S(Q)$ for insulin in the presence of ethanol could be fitted best with a charge $Z = 6$ and an effective hard sphere diameter σ of 20.5 Å. The depths of the attractive interaction potential $V_Y(r)$, J, are 16 and 23 k_BT, with d-values of 5 Å and 2 Å, for insulin concentrations of 1 and 4 wt%, respectively. Thus, the depths and the ranges of $V_Y(r)$ in the presence of 20% (v/v) ethanol are only slightly smaller with respect to the corresponding data of the protein in pure water. The enhanced hydrophobicity of the solvent probably partially balances the hydrophobic part of attractive interaction of the protein molecules, thus slightly disfavoring short range attractive protein-protein interactions. This is in agreement with the observation that the addition of ethanol delays the nucleation process and hence lag-period and aggregation time of insulin fibrillation. For example, at 50 ^oC, the half time for fibrillation is 90 min in 0.1 M HCl, pH 2.0, and 310 min in 20 wt% EtOH, 0.1 M NaCl, pH 2.0, respectively [Grudzielanek et al., 2005]. Remarkably, these drastic changes in aggregation time are based on a small (~10%) change in $V_Y(r)$ of the dimeric and monomeric insulin particles, only.

3.4.4 Concluding remarks

The overall parameters of intermolecular interactions among the insulin molecules under different cosolvent conditions are given in Table 3.2. Insulin as dimeric form at pH 2.0 shows significantly higher magnitudes of the attractive part of the interaction potential as compared to the natively globular stable protein lysozyme. The attractive interactions increase dramatically with increasing concentration of the protein in a converse behavior to the lysozyme interactions. Sodium chloride as charge screener enhances the attractive interactions among insulin molecules (tetrameric form) to approximately two-fold compared to water, only. Addition of ethanol slightly decreases the attractive interactions among the insulin molecules (monomeric form). The total interaction potentials $V(r)$ of

the insulin molecules at a concentration of 4 wt % in acidic water, 20 % (v/v) ethanol and 100 mM sodium chloride at pH 2.0 are shown diagrammatically in Fig. 3.38 along with their corresponding attractive part of the interaction potential, $V_Y(r)$.

Sample	J/k_BT	d /Å	Radius of Gyration / Å	Effective charge	Effective sigma / Å
Insulin 1 wt% in water	18	11	14.3	11	23
Insulin 4 wt% in water	27	4			
Insulin 10 wt% in water	32	2			
Insulin 1 wt% in 100 sodium chloride	38	5	16	24	30
Insulin 4 wt% in 100 sodium chloride	50	3			
Insulin 1 wt% in 20 % (v/v) ethanol	16	5	11.9	6	20.5
Insulin 4 wt% in 20 % (v/v) ethanol	23	2			

Table 3.2: Fitting parameters of $S(Q)$for insulin solutions with and without added cosolvents: J (in k_BT) is the depth of the attractive interaction potential $V_Y(r)$, and d (in Å) is the range of the potential. Sigma is the effective diameter of the insulin molecule.

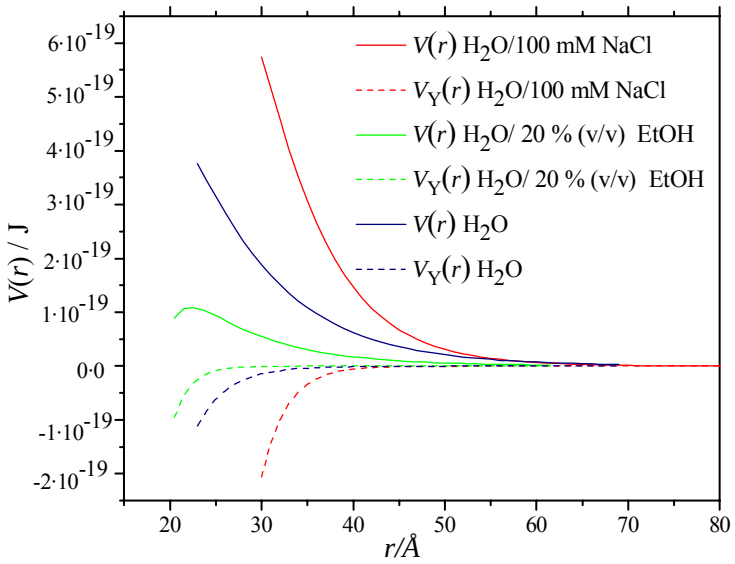


Figure 3.38: Plot of the total intermolecular interaction potential $V(r)$ (in full lines) and the attractive part of the potential $V_Y(r)$ (in dashed lines) for 4 wt% insulin in water (blue), 20% (v/v) ethanol (green) and 100 mM sodium chloride (red) at pH 2.0.

Chapter 4

Conclusions

In this work, intermolecular interactions of solutions of two selected proteins were determined which have transverse implications and crucial importance in vast areas extending from *in-vivo* to *in-vitro* scenarios. Protein-protein interactions have a critical role in a wide variety of situations, like pathological disorders (human eye cataract, Alzheimer's and Parkinson's diseases), production of dynamically arrested states, stabilization and purification of proteins, delivery of protein drugs using cosolvents as interaction modulators and crystallization of proteins for three-dimensional structure determination.

The effects of various kosmotropic and chaotropic cosolvents and cosolutes (salts) on the intermolecular interaction potential $V(r)$ of the proteins lysozyme and insulin and its respective repulsive and attractive contributions were evaluated at various protein concentrations by using synchrotron small-angle X-ray scattering in combination with liquid-state theories. The experimentally derived structure factors, $S(Q)$, obtained from the intensity patterns of different proteins solutions with and without added cosolvents and salts were fitted with a statistical mechanical model based on a DLVO potential which accounts for the repulsive and attractive interactions between the protein molecules. The random phase approximation (RPA) was used for calculation of the theoretical structure factors, which includes the hard sphere potential as reference system and the sum of a repulsive screened Coulomb ($V_C(r)$) interaction and an attractive Yukawa type potential ($V_Y(r)$) as perturbation potential. The attractive Yukawa type potential is taken as the sum of van der Waals-interactions and the osmotic attractive potential due to excluded volume effects of the salt ions.

The scattering data of natively stable globular protein lysozyme in solution indicates that intermolecular interactions of lysozyme solutions are significant above protein

concentrations above ~1 wt%, and the scattering patterns for lysozyme concentrations at and above 4 wt% exhibit strong intermolecular correlation peaks. The repulsive nature of $V(r)$ increases with increasing protein concentration. For example, the potential well of $V_Y(r)$ increases from -3.75 k_BT (modulus of J equals 3.75 k_BT) for the 4 wt% protein solution to -2.5 k_BT for the 10 wt% protein solution.

Addition of kosmotropic cosolvents such as glycerol and sucrose leads to strong concentration dependent effects on the interaction potential. As revealed by complementary pressure perturbation calorimetric experiments, these protein stabilizing agents increase the hydration strength and hence lead to an increase of repulsive forces between the protein molecules. This effect is more pronounced for sucrose, for which a greater steric exclusion is expected due to the larger size of the sucrose molecule, which is in fact in accordance with the increased preferential hydration capacity of this osmolyte. For instance, the potential well of $V_Y(r)$ increases markedly from -3.75 k_BT for the 4 wt% protein solution to -2.25 k_BT already for the protein solution containing 0.5 M sucrose, only. In this concentration range, both glycerol as well as sucrose act in a way that they increase the repulsion between protein molecules, probably due to the enhancement of the strength of the hydration layer. The effect of increased repulsive interactions creates more short-range order between the protein molecules. Only at very high osmolyte concentrations (above about 1 M), sufficient hydration of the protein molecules is no longer possible (in accordance with calorimetric data [Ravindra and Winter, 2003 and Ravindra and Winter, 2004])) and the effect may be reversed.

Trifluoroethanol (TFE) displays a multiphasic effect on $V(r)$ when changing the TFE concentration. By addition of 10 % TFE, intermolecular interactions become more repulsive in nature, similar to the behavior of osmolytes. At higher TFE concentration (35 %) the scenario changes. Due to the drastic decrease in dielectric permittivity of the medium, the attractive as well as the repulsive part of the interaction potential are enhanced. In addition, as a consequence of the drastic decrease in ε, exposure of hydrophobic residues to solvent molecules is increasingly favored, thus leading to partial unfolding (consistent with an increase of R_g) and increased intermolecular hydrophobic interactions of the protein molecules. The potential well of $V_Y(r)$ decreases to -7.75 k_BT for the 4 wt % solution.

Salts like sodium chloride and potassium sulphate exhibit strong concentration dependent changes in the attractive interaction potential due to charge screening of the positively charged protein molecules. With increasing NaCl concentration, the potential well of $V_Y(r)$ decreases from -3.75 k_BT for 0 mM NaCl to -4.75 k_BT for 50 mM, -8.25 k_BT for 100 mM, to -13 k_BT for 200 mM added NaCl and to -13 k_BT for 250 mM K_2SO_4. At high protein concentrations, the limitation of the applied model is reached as a result of oligomer formation of the protein molecules in the strongly attractive regime. With increasing SO_4^{2-} concentration, precipitation of the protein sets in.

Guanidinium chloride, a widely used chaotropic agent, exhibits a similar charge screening effect, resulting in increased attractive interactions between the protein molecules. At higher GdmCl concentrations (3.5 M), however, $V(r)$ becomes more repulsive in nature which might be due to the presence of a large concentration of Gdm^+ ions binding to the protein molecules, hence leading to an enhanced electrostatic repulsion. Such repulsion may also be due to the fact that at higher GdmCl concentrations, expansion and hydration of the protein starts to increase due to partial unfolding of the protein (R_g increases by ≈ 7 %), which also favors an increase of repulsive interactions. At still higher GdmCl concentrations, marked protein unfolding and formation of different populations of protein species sets in, hence the model applied is no longer applicable.

To conclude, this work underlines the need for the solvation and intermolecular interactions of proteins to be well understood and quantified to account for the physico-chemical properties of proteins even at low salt and cosolvent concentrations. In particular, it is essential to point out that these findings also imply that in calculations of thermodynamic properties of proteins, owing to significant intermolecular interactions, activity coefficients are generally not negligible in the concentration range above 1 wt% protein. In most biochemical and biophysical studies, however, they are neglected.

Many studies have shown that protein aggregation is inherently a nucleation and growth phenomenon where aggregates accumulate, eventually exceeding their solubility and precipitate. The existence of a lag phase in the aggregation is caused by an energetic barrier to nucleation or assembly. The barrier to assembly may be orientationally specific. If there is an orientation with a lower free energy to assemble, the growth will occur

preferentially in that orientation, resulting in ordered aggregate morphology, such as fibrils. In order to elucidate the contributions of $V(r)$ responsible for the initiation of protein fibrillar self-assembly, $S(Q)$ for insulin solutions at various concentrations were calculated under charge screening conditions (addition of NaCl) and non charge screening conditions, as well as in 20 % (v/v) ethanol in pre-aggregated states.

Attractive and repulsive interaction potentials were calculated likewise from fitting of the experimental $S(Q)$ to deduce the information about the initiation process of protein aggregation and fibrillation with reference to intermolecular interactions. Surprisingly, the peak maximum of $S(Q)$ is essentially independent of the insulin concentration, indicating that insulin particles self-assemble into equilibrium clusters with a concentration-dependent aggregation number instead of forming a homogeneous spatial distribution of protein particles. The potential well of $V_Y(r)$ decreases drastically to -18, -27, and -32 k_BT for insulin concentrations of 1, 4 and 10 wt%, respectively. In contrast, for lysozyme, representing a rather stable protein, the attractive part of $V(r)$ is much smaller, and exhibits a different concentration dependence. The marked increase in the attractive interaction of insulin shows that short-range van der Waals and probably specific hydrophobic attractions of the partially unfolded insulin molecules (exposing hydrophobic patches) are strong already at rather low concentrations and low temperatures (25 ^oC).

Upon charge screening of the positively charged insulin with 0.1 M NaCl, a drastic decrease in the potential well of $V_Y(r)$ is observed, which is -38 k_BT for 1 wt % and -50 k_BT for the 4 wt % insulin solution. This marked increase in attractive interactions explains the enhanced rate of protein aggregation and fibrillation under charge-screening conditions.

To reveal how the intermolecular interaction depends on solvational conditions and the initial aggregation-prone insulin species, additional experiments were carried out at conditions, where insulin exists as monomeric species, only. The data reveals that the depths and the ranges of $V_Y(r)$ in the presence of 20% (v/v) ethanol are ~10% smaller with respect to the corresponding data of the dimeric protein in pure water. It is very likely that the enhanced hydrophobicity of the solvent partially balances the hydrophobic part of attractive interaction of the protein molecules which is in agreement with the

findings that the addition of ethanol delays the nucleation process of insulin fibrillation Remarkably, these drastic changes in aggregation time are based on a minor change in $V(r)$, only.

To summarise, the combination of a weakly screened long-range electrostatic repulsion and a marked short-range attraction leads to the pronounced formation of equilibrium clusters for the aggregation-prone protein. The partially unfolded state of insulin presents hydrophobic patches which lead to the drastic increase of site-specific short-range hydrophobic interaction in form of clusters in this aggregation-precursor regime. This already occurs far from conditions where the actual aggregation and fibril formation takes place. In other words, it happens before the entropy of releasing the water layers and the attractive short-range H-bonding forces can provide enough driving force to "dry out" the contacting surfaces and ordered fibrillar structures are formed. Such event is facilitated on a short time scale at higher temperatures.

To conclude, the approach employed here is able to distinguish striking differences regarding interaction forces between aggregation-prone proteins such as insulin and natively stable globular proteins such as lysozyme under different hydration, solvation, hydrophobicity and charge screening conditions. This may therefore be used in forthcoming studies on other more disease-related amyloidogenic proteins, such as the Alzheimer peptide or prion protein. Knowledge of these forces may allow to control or fine tune self-assembly of amyloidogenic proteins.

Chapter 5

Zusammenfassung

In dieser Arbeit wurden die intermolekularen Wechselwirkungen zwischen zwei Proteinen in Lösung charakterisiert, welche unterschiedliche Implikationen haben und von essentieller Wichtigkeit - reichend von *in vivo* zu *in vitro* Szenarien - sind. Protein-Protein-Wechselwirkungen spielen eine wichtige Rolle in einer ganzen Bandbreite von Prozessen, wie bei pathologischen Störungen (Katarakt des menschlichen Auges, Alzheimer und Parkinson), der Herstellung dynamisch gefangener Zustände, der Stabilisierung und Reinigung von Proteinen, der Applikation von Protein-Pharmaka mit Cosolventien als Wechselwirkungs-Modulatoren sowie der Kristallisation von Proteinen zur Bestimmung ihrer dreidimensionalen Struktur.

Die Effekte verschiedener kosmotroper sowie chaotroper Cosolventien und Salze auf das intermolekulare Wechselwirkungspotential $V(r)$ der Proteine Lysozym und Insulin und die jeweiligen repulsiven und attraktiven Anteile wurden bei verschiedenen Proteinkonzentrationen mittels Synchrotron Röntgen-Kleinwinkelstreuung in Kombination mit Flüssigkeits-theoretischen Modellen bestimmt. Der experimentell gewonnene Strukturfaktor $S(Q)$, erhalten aus den Streukurven verschiedener Proteinlösung in An- und Abwesenheit von Cosolventien und Salzen, wurde mit einem statistisch-mechanischen Modell angefittet, welches auf dem DLVO-Potential basiert und die repulsiven sowie attraktiven Wechselwirkungen zwischen den Proteinmolekülen berücksichtigt. Die sog. *Random Phase Approxiamtion* (RPA) wurde zur Berechnung des theoretischen Strukturfaktors verwendet, welches aus einem Potential harter, undurchdringlicher Kugeln (“*hard sphere potential*”) als Referenzsystem sowie der Summe aus einer repulsiven, abgeschirmten Coulomb-Wechselwirkung ($V_C(r)$) und einem attraktiven Yukawa-Potential ($V_Y(r)$) als Störpotential besteht. Das attraktive Yukawa-Potential wird als Summe der van-der-Waals Wechselwirkungen und des

attraktiven osmotischen Potentials aufgrund des Ausschlussvolumen-Effekts der Salzionen aufgefasst.

Die Streudaten des nativ stabilen globulären Proteins Lysozym in Lösung zeigen, dass Effekte intermolekularer Wechselwirkungen von Lysozymlösungen oberhalb von ~ 1 % (w/v) signifikant sind, und die Streumuster für Konzentrationen bei und oberhalb von 4 % (w/v) weisen starke intermolekulare Korrelationssignale auf. Die repulsive Natur von $V(r)$ vergrößert sich mit ansteigender Proteinkonzentration. So steigt z.B. die Tiefe des Potentialtopfs von $V_Y(r)$ von –3,75 k_BT für die 4 %ige Lösung auf –2,5 k_BT für die 10 %ige Proteinlösung.

Die Zugabe kosmotroper Cosolventien wie Glycerin und Sucrose führt zu stark konzentrationsabhängigen Effekten auf das Wechselwirkungspotential. Wie durch ergänzende PPC („*Pressure Pertubation Calorimetry*“)-Messungen gezeigt, erhöhen diese proteinstabilisierenden Agenzien die Stärke der Hydratation und führen so zu einer Erhöhung der abstoßenden Kräfte zwischen den Proteinmolekülen. Dieser Effekt fällt für Sucrose stärker aus, da hierfür ein größeres sterisches Ausschlussvolumen aufgrund des größeren Volumens der Sucrosemoleküle erwartet werden kann; und in der Tat ist dies in Übereinstimmung mit der erhöhten Kapazität zur preferentiellen Hydratation („*preferential hydration*“) dieses Osmolyten. So steigt z.B. die Höhe des Potentialtopfes von $V_Y(r)$ deutlich von –3,75 k_BT für die 4 %ige Lösung auf –2,25 k_BT für die Proteinlösung mit nur 0,5 M Sucrose. In diesem Konzentrationsbereich wirken sowohl Glycerin als auch Sucrose in der Art, dass sie die Abstoßung zwischen den Proteinmolekülen vergrößern, wahrscheinlich indem sie die Stärke und Ausdehnung der Hydratationsschicht erhöhen. Dieser Effekt der Erhöhung der repulsiven Kräfte führt zu mehr Ordnung im Nahbereich zwischen den Proteinmolekülen. Nur bei sehr hohen Osmolytkonzentrationen (jenseits von 1 M), ist eine ausreichende Hydratation des Proteins nicht mehr möglich (in Einklang mit kalorimetrischen Daten [Ravindra und Winter, 2003, sowie Ravindra und Winter, 2004]), und der Effekt könnte sich umkehren.

Trifluorethanol (TFE) weist mit sich ändernder Konzentration einen multiphasischen Effekt auf $V(r)$. Bei Zugabe von 10 % TFE werden die intermolekularen Wechselwirkungen repulsiver, ähnlich wie bei den Osmolyten Glycerin und Sucrose. Bei höheren Konzentrationen (35 %) ändert sich das Bild. Aufgrund der drastisch erhöhten

dielektrischen Permeabilität des Mediums werden sowohl der attraktive sowie der repulsive Teil des Wechselwirkungspotentials erhöht. Zusätzlich, als ein Effekt des stark verminderten Werts von ε, wird eine Exposition hydrophober Reste energetisch günstiger, was zu einer teilweisen Entfaltung (in Übereinstimmung mit einem erhöhten Gyrationsradius R_g) und zu erhöhten hydrophoben Wechselwirkungen der Proteinmoleküle führt. Die Tiefe des Potentialtopfes von $V_Y(r)$ sinkt auf –7,75 k_BT für die 4 %ige (w/v) Lösung.

Salze wie Natriumchlorid und Kaliumsulfat bewirken eine stark konzentrationsabhängige Änderung des attraktiven Potentials aufgrund der Abschirmung der positiven Ladungen der Proteinmoleküle. Mit ansteigender Salzkonzentration sinkt die Tiefe des Potentialtopfes von $V_Y(r)$ von –3,75 k_BT für 0 mM NaCl auf –4,75 k_BT für 50 mM, –8,25 k_BT für 100 mM bis auf –13 k_BT für 200 mM NaCl sowie auf –13 k_BT für 250 mM K_2SO_4. Bei hohen Proteinkonzentrationen wird hier die Grenze des verwendeten Modells erreicht, da die Proteine in diesem stark attraktiven Regime Oligomere bilden. Mit ansteigender SO_4^{2-}-Konzentration setzt die Präzipitation des Proteins ein.

Guanidiumchlorid, ein vielfach eingesetztes, chaotropes Agens, weist einen vergleichbaren Abschirmungseffekt auf, der zu erhöhten attraktiven Wechselwirkungen zwischen den Proteinmolekülen führt. Bei hohen GdmCl-Konzentrationen (3,5 M) jedoch wird $V(r)$ repulsiver, was darin begründet sein könnte, dass bei hohen Gdm^+-Konzentrationen dieses Ion an das Protein bindet und so zu einer verstärkten elektrostatischen Abstoßung führt. Diese Abstoßung könnte weiterhin daran liegen, dass bei höheren GdmCl-Konzentrationen die Ausdehnung und Hydratation des Proteins aufgrund der teilweisen Entfaltung des Moleküls zu steigen beginnt (R_g erhöht sich um ~ 7%), was ebenfalls ein Anstieg der repulsiven Wechselwirkungen bewirkt. Bei noch höheren GdmCl-Konzentrationen setzt eine ausgeprägte Entfaltung und die Bildung verschiedener Populationen von Proteinspezies ein, so dass das verwendete theoretische Modell nicht weiter anwendbar ist.

Schließlich unterstreicht dieser Teil der Arbeit die Notwendigkeit, die Solvatation und intermolekularen Wechselwirkungen von Proteinen genau zu verstehen und zu quantifizieren, um den physikalisch-chemischen Eigenschaften von Proteinen - auch bei niedrigen Salz- und Cosolventkonzentrationen - Rechnung zu tragen. Im Besonderen

muss betont werden, dass diese Befunde zeigen, dass bei der Berechnung thermodynamischer Größen von Proteinen Aktivitätskoeffizienten generell oberhalb einer Konzentration von 1 % (w/v) aufgrund der signifikanten intermolekularen Wechselwirkungen nicht vernachlässigbar sind. In den meisten biochemischen und biophysikalischen Studien wird dies jedoch getan.

Viele Studien haben gezeigt, dass die Proteinaggregation und Fibrillbildung von Natur aus ein Nukleations- und Wachstumsprozess ist, bei dem Aggregate akkumulieren, bis sie eventuell ihre Löslichkeit überschreiten und präzipitieren. Die Existenz einer Verzögerungsphase („*lag phase*“) bei der Aggregation wird durch eine energetische Barriere bei der Nukleation verursacht. Diese Barriere beim Zusammenbau der Moleküle kann orientierungsspezifisch sein. Falls eine Orientierung mit niedrigerer freier Energie bei der Clusterbildung existiert, so wird das Wachstum vornehmlich in dieser Orientierung voranschreiten, was zu einer geordneten Struktur des Aggregats führt, wie z.B. zu Fibrillen. Um den Anteil von $V(r)$ aufzuklären, der für die Initiation des protein-fibrillären Self-assembly verantwortlich ist, wurde $S(Q)$ für Insulinlösungen verschiedener Konzentration unter ladungsabschirmenden Bedingungen (Zusatz von NaCl), nicht ladungsabschirmenden Bedingungen sowie in Anwesenheit von 20 % (v/v) Ethanol im prä-aggregierten Zustand bestimmt.
Die attraktiven und repulsiven Wechselwirkungspotentiale wurden durch Fitten des experimentellen Strukturfaktors $S(Q)$ berechnet, um Informationen über den Initiationsprozess der Aggregation und Fibrillbildung bezüglich der intermolekularen Wechselwirkung zu erhalten. Überraschenderweise ist das Maximum von $S(Q)$ im Wesentlichen unabhängig von der Insulinkonzentration, was aufzeigt, dass die Insulinpartikel sich in Gleichgewichts-Clustern mit einer konzentrationsabhängigen Aggregationszahl zusammenlagern, anstatt eine räumlich homogene Verteilung von Proteinpartikeln einzunehmen. Der Potentialtopf von $V_Y(r)$ sinkt drastisch auf –18, –27 und –32 k_BT für Insulinkonzentrationen von 1, 4 bzw. 10 % (w/v). Im Gegensatz dazu ist der attraktive Teil von $V_Y(r)$ für Lysozym, welches ein stabiles Protein darstellt, wesentlich kleiner und weist eine andere Konzentrationsabhängigkeit auf. Der deutliche Anstieg der attraktiven Wechselwirkungen für Insulin zeigt, dass die kurzreichweitigen

van-der-Waals -und wahrscheinlich auch spezifischen hydrophoben Wechselwirkungen des teilentfalteten Insulinmoleküls (welches hydrophobe „Patches“ exponiert) schon bei vergleichsweise niedrigen Konzentrationen und Temperaturen (25 °C) sehr stark sind.

Unter Ladungsabschirmung des positiv geladenen Insulins in Anwesenheit von 0,1 M NaCl wird eine drastische Abnahme der Potentialtopftiefe von $V_Y(r)$ beobachtet, welche $-38\ k_BT$ in der 1 %igen (w/v) und $-50\ k_BT$ in der 4 %igen (w/v) Insulinlösung beträgt. Dieser deutliche Anstieg der attraktiven Wechselwirkungen erklärt die erhöhte Rate der Proteinaggregation und Fibrillbildung unter ladungsabschirmenden Bedingungen.

Um herauszufinden, wie die intermolekularen Wechselwirkungen von den Lösungsbedingungen und den initialen, zur Aggregation neigenden Insulinspezies abhängen, wurden weitere Experimente durchgeführt, unter denen Insulin nur als Monomer vorliegt. Die Daten zeigen, dass die Tiefe und die Reichweite von $V_Y(r)$ in Anwesenheit von 20 % (v/v) Ethanol etwa ~ 10 % kleiner sind als diejenigen des dimeren Proteins in Wasser. Es ist sehr wahrscheinlich, dass die erhöhte Hydrophobizität des Lösungsmittels teilweise den hydrophoben Teil der attraktiven Wechselwirkungen aufhebt, was im Einklang mit dem Befund ist, dass die Zugabe von Ethanol den Nukleationsprozess der Fibrillbildung von Insulin verzögert.

Zusammenfassend lässt sich sagen, dass die Kombination einer schwach abgeschirmten, langreichweitigen Abstoßung und eine signifikante, kurzreichweitige Anziehung zu einer ausgeprägten Bildung von Gleichgewichts-Clustern von zur Aggregation neigenden Proteinen führt. Der teilentfaltete Zustand von Insulin weist hydrophobe „Patches“ auf, die zu einem drastischen Anstieg ortsspezifischer, kurzreichweitiger und hydrophober Wechselwirkungen in Form von Clustern in diesem Aggregations-Vorstadium führen. Dies geschieht bereits unter Bedingungen, die fern von denjenigen sind, unter denen die eigentliche Aggregation und Fibrillbildung einsetzt. In anderen Worten: Es geschieht, bevor die Entropie der freigesetzten Hydratationsschichten und die attraktiven Kräfte der kurzreichweitigen Wasserstoffbrückenbindungen genug treibende Kraft vermitteln können, um die Kontaktflächen zu „trocknen“ und geordnete, fibrilläre Strukturen zu formen. Derlei Vorgänge finden auf kurzer Zeitskala bei höheren Temperaturen statt.

Der hier verwendete Ansatz ist in der Lage, auffallende Unterschiede der Wechselwirkungspotentiale von zur Aggregation neigenden Proteinen wie Insulin, und

nativ stabilen, globulären Proteinen wie Lysozym unter verschiedenen Bedingungen von Hydratation, Solvatation, Hydrophobizität und Ladungsabschirmung aufzuzeigen. Er könnte daher in weiteren Studien über andere, mehr krankheitsbezogene amyloide Proteine, wie das Alzheimer-Peptid oder dem Prion-Protein, Verwendung finden. Die Kenntnis der Wechselwirkungskräfte könnte es erlauben, die Aggregation amyloider Proteine gezielt zu kontrollieren.

References

Ahmad, A., Millett, I. S., Doniach, S., Uversky, V. N. and Fink, A. L. (2004) Stimulation of insulin fibrillation by urea-induced intermediates. *J. Biol. Chem.* **279**, 14999-15013.

Arakawa, A. (2002) Hydration as major factor in preferential solvent-protein interactions. *Cryst. Growth and Desgin*. **2**, 549-551.

Back, J. F., Oakenfull, D. and Smith, M. B. (1979) Increased thermal stability of proteins in the presence of sugars and polyols. *Biochemistry* **18**, 5191-5196.

Baglioni, P., Fratini, E., Lonetti, B. and Chen, S. H. (2004) Structural arrest in concentrated cytochrome C solutions: The effect of pH and salts. *J. Phys. Condens. Matter* **16**, 5003-5022.

Barlow, D. J. and Thornton, J. M. (1986) The distribution of charged groups in proteins. *Biopolymers* **25**, 1717-1733.

Belloni, L. (1991) in: Linder, P. and Zemp, T. (Eds), Neutron, X-ray and light scattering. Elsevier, Amsterdam.

Belloni, L. (1993) Inability of the hypernetted chain integral equation to exhibit a spinodal line. *J. Chem. Phys*. **98** (10), 8080-8095.

Bennion, B. J. and Daggett, V. (2003) The molecular basis for the chemical denaturation of proteins by urea. *Proc. Natl. Acad. Sci. USA.* **100**, 5142-5147.

Bonnete, F., Vivares, D., Robert, C. and Colloc´h, N. (2001) Interactions in solution and crystallization of aspergillus flavus urate oxidase. *J. Cryst. Growth.* **232**, 330-339.

Buck, M. (1998) Trifluoroethanol and colleagues: Cosolvents come of age. Recent studies with peptides and proteins. *Q. Rev. Biophys.* **31**, 297–355.

Bull, H. B. and Breese, K. (1978) Interaction of alcohols with proteins. *Biopolymers* **17**, 2121-2131.

Chang, X., Jorgensen, A. M., Bardrum, P. and Led, J. J. (1997) Solution structures of the R_6 human insulin hexamer. *Biochemistry* **36**, 9409 - 9422.

Chi, E. Y., Krishnan, S., Randolph, T. W. and Carpenter, J. F. (2003) Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. *Pharm. Res.* **20**, 1325-1336.

Chiti, F. and Dobson, C. M. (2006) Protein misfolding, functional amyloid and human disease. *Annu. Rev. Biochem.* **75**, 333-366.

Cinelli, S., Onori, G. and Santucci, A. (1997) Effect of aqueous alcohol solutions on the thermal transition of lysozyme: A calorimetric study. *J. Phys. Chem. B* **101**, 8029-8034.

Coen, C. J., Blanch, H. W. and Prausnitz, J. M. (1995) Salting out of aqueous proteins : Phase equilibria and intermolecular potentials. *AIChE. J.* **41**, 996-1004.

Curtis, R. A., Prausnitz, J. M. and Blanch, H. W. (1998) Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. *Biotechnol. Bioeng.* **57**, 11-21.

Debenedetti, P. G. (1996) Metastable liquids: Concepts and principles. Princeton University Press, Princeton.

Derewenda, U., Derewenda Z., Dodson, G. G., Hubbard R. E. and Korber, F. (1989) Molecular structure of insulin: The insulin monomer and its assembly. *British Medical Bulletin*. **4**, 4-18.

Ducruix, A., Guilloteau, J. P., Ries-Kautt, M. and Tardieu, A. (1996) Protein interactions as seen by solution X-ray scattering prior to crystallogenesis. *J. Cryst. Growth.* **168**, 28-39.

Durbin, S. D. and Feher, G. (1996) Protein crystallization. *Annu. Rev. Phys. Chem*. **47,** 171-204.

Dzwolak, W., Ravindra, R., Nicolini, C., Jansen, R. and Winter, R. (2004) The diastereomeric assembly of polylysine is the low-volume pathway for preferential formation of ß-sheet aggregates. *J. Am. Chem. Soc.* **126**, 3762-3768.

Dzwolak, W., Grudzielanek, S., Smirnovas, V., Ravindra, R., Nicolini, C., Jansen, R., Loksztejn, A., Porowski, S. and Winter, R. (2005) Ethanol-perturbed amyloidogenic self-assembly of insulin: Looking for origins of amyloid strains *Biochemistry* **44**, 8948-8958.

Dzwolak, W. (2006) Tuning amyloidogenic conformations through cosolvents and hydrostatic pressure: When the soft matter becomes even softer. Biochim. *Biophys. Acta, Proteins Proteomics*. **1764**, 470-480.

Falbe, J. and Regtiz, M. (1995) CD-Römpp, Chemielexikon, vers. 1, George Thieme Verlag, Stuttgart.

Farnum, M. and Zukoski, C. (1999) Effect of glycerol on the interactions and solubility of bovine pancreatic trypsin inhibitor. *Biophys. J.* **76**, 2716-2726.

Fersht, A. (1999) Structure and mechanism in protein science: A guide to enzyme catalysis and protein folding. W. H. Freeman company, New York .

Gekko, K. and Timasheff, S. N. (1981) Mechanism of protein stabilization by glycerol: preferential hydration in glycerol-water mixtures. *Biochemistry* **20**, 4667-4676.

Grimson, M. J. (1983) Small-angle scattering from colloidal dispersions. *J. Chem. Soc. Faraday. Trans.* **79**, 817-832.

Grudzielanek, S., Jansen, R. and Winter, R.(2005) Solvational tuning of the unfolding, aggregation and amyloidogenesis of insulin. *J. Mol. Biol*. **351**, 879-894.

Grudzielanek, S., Smirnovas, V. and Winter, R. (2006) Solvation-assisted pressure tuning of insulin fibrillation: From novel aggregation pathways to biotechnological applications. *J. Mol. Biol.* **356**, 497-509.

Guinier A. (1939) La diffraction des rayons X aux tres petits angles; application a l'etude de phenomenes ultramicroscopiques. *Ann. Phys.* **12**, 166–237.

Hammersley, A. P., Svensson, S. O. and Thompson, A. (1994) Calibration and correction of spatial didtortions in 2D detector systems *Nucl. Instr. Meth.* **346**, 312-321.

Hansen, J. P. and McDonald, I.R. (1986) Theory of simple liquids, 2ed., Academic Press, London, Orlando.

Herberhold, H., Royer, C. A. and Winter, R. (2004) Effects of chaotropic and kosmotropic cosolvents on the pressure-induced unfolding and denaturation of proteins: an FT-IR study on staphylococcal nuclease. *Biochemistry* **43**, 3336-3345.

Hirota, N., Mizuno, K. and Goto, Y. (1998) Group additive contributions to the alcohol-induced alpha-helix formation of melittin: implication for the mechanism of the alcohol effects on proteins. *J. Mol. Biol.* **275**, 365-378.

Hong, D. P., Hoshino, M., Kuboi, R. and Goto, Y. (1999) Clustering of fluorine-substituted alcohols as a factor responsible for their marked effects on proteins and peptides. *J. Am. Chem. Soc.* **121**, 8427-8433.

Israelachvili, J. (1992) Intermolecular and surface forces, Academic Press, New York.

Jansen, R., Dzwolak, W. and Winter, R. (2005) Amyloidogenic self-assembly of insulin aggregates probed by high resolution atomic force microscopy *Biophys. J.* **88**, 1344-1353.

Javid, N., Vogtt, K., Krywka, C., Tolan, M., and Winter, R. (2007) Protein-protein interactions in complex cosolvent solutions. *ChemPhysChem*. **8**, 679-689.

Kelkar, V. K., Narayanan, J. and Manohar. C. (1992) Structure factor for colloidal dispersions: Use of exact potentials in random phase approximation. *Langmuir* **8**, 2210-2214.

Kentsis. A. and Sosnick, T. R. (1998) Trifluoroethanol promotes helix formation by destabilizing backbone exposure: Desolvation rather than native hydrogen bonding defines the kinetic pathway of dimeric coiled coil folding. *Biochemistry* **37**, 14613–14622.

Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. and Svergun, D. I. (2003) PRIMUS: A window PC-based system for small angle scattering data analysis. *J. Appl. Crystallogr*. **36**, 1277-1282.

Krywka, C., Paulus, M. Sternemann, C., Volmer, M., Rehmof, A., Nowak, G., Nefedov, A., Pöter, B., Spiegel, M. and Tolan, M. (2006) The new diffractometer for surface X-ray diffraction at beamline BL9 of DELTA *J. Synchrotron Rad.* **13**, 8-13.

Krywka, C., Sternemann, C., Paulus, M., Javid, N., Winter, R., Al-Sawalmih, A., Yi, S., Rabbe, D. and Tolan, M. (2007) The small-angle and wide-angle X-ray scattering set-up at beamline BL9 of DELTA. *J. Synchrotron. Rad*. **14**, 244-251.

Lee, J. and Pilch, P. F. (1994) The insulin receptor: structure, function and signaling. *Am.J. Physio*l. (*Cell Physiol*.) **266**, 319-334.

Luzzati, V. and Tardieu, A. (1980) Recent developments in solution X-ray scattering. *Ann. rev. Biophys. Bioeng*. **9**, 1-30.

Löffler, G. and Petrides, P. E. (1990) Physiologische chemie, 4th ed., Springer-Verlag, Berlin, Heidelberg, New York, London, Paris, Tokyo.

Makhatadze, G. I. (1999) Thermodynamics of protein interactions with urea and guanidinium hydrochloride. *J. Phys. Chem. B.* **103**, 4781-4785.

Malfois, M., Bonnete, F. L. Belloni, L. and Tardieu, A. (1996) A model of attractive interactions to account for fluid–fluid phase separation of protein solutions. *J. Chem. Phys.* **105**, 3290–3300.

Mande, S. C. and Sobhia, M. E. (2000) Structural characterization of protein-denaturant interactions: Crystal structures of hen egg-white lysozyme in complex with DMSO and guanidinium chloride. *Prot. Eng*., *Des. Sel.* **13**, 133-141.

Minton. A. P. (2001) The Influence of macromolecular crowding and macromolecular confinement on biochemical reactions in physiological media. *J. Biol. Chem.* **276**, 10577-10580.

Mitra, L., Smolin, N., Ravindra, R., Royer, C. and Winter, R. (2006) Pressure perturbation calorimetric studies of the solvation properties and thermal unfolding of

proteins in solution-experiments and theoretical interpretation. *Phys. Chem. Chem. Phys.* **8**, 1249-1265.

Munishkina, L. A., Cooper, E. M., Uversky, V. N., and Fink, A. L. (2004 a) The effect of macromolecular crowding on protein aggregation and amyloid fibril formation. *J.Mol. Recognit*. **17**, 456-464.

Munishkina, L. A., Henriques, J., Uversky, V. N., and Fink, A. L. (2004 b) Role of protein-water interactions and electrostatics in alpha-synuclein fibril formation. *Biochemistry* **43**, 3289-3300.

Narayanan, J. and Liu, X. Y. (2003) Protein interactions in undersaturated and supersatureated solutions: A study using light and x-ray scattering. *Biophys. J.* **84**, 523-532.

Nelson, D. L. and Cox, M. M. (2000) Lehninger principles of biochemistry. 3 rd ed., Worth Publisher, New York.

Niebuhr, M. and Koch, M. H. J. (2005) Effects of urea and trimethylamine-N-Oxide (TMAO) on the interactions of lysozyme in solution. *Biophys. J*. **89**, 1978-1983.

Nielsen, L., Khurana, R., Coats, A., Froakjer, S., Brange, J., Vyas, S. Uversky, V. N. and Fink, A. L. (2001) Effect of environmental factors on the kinetics of insulin fibril formation: Elucidation of the molecular mechanism. *Biochemistry* **40**, 6036- 6046.

Nielsen, L., Frokjaer, S., Brange, J., Uversky, V. N. and Fink, A. L. (2001b) Probing the mechanism of insulin fibril formation with insulin mutants. *Biochemistry* **40**, 8397-8409.

Paulus, M., Fendt, R., Sternemann, C., Gutt, C., Hövel, H., Volmer, M., Tolan, M. and Wille, K. (2005) An internet-based synchrotron experiment for students measuring the X-ray magnetic circular dichroism of a Pt-Fe alloy *J. Synchrotron. Rad*. **12**, 246-250.

Podesta, A., Tiana , G., Milani, P. and Manno, M. (2006) Early events in insulin fibrillization studied by time-lapse atomic force microscopy. *Biophys. J*. **90**, 589-597.

Priev, A., Almagor, A., Yedgar, S. and Gavish, B. (1996) Glycerol decreases the volume and compressibility of protein interior. *Biochemistry* **35**, 2061-2066.

Qiu, D., Cosgrove, T., Howe, A. M. and Dreiss, C. A. (2006) A small-angle X-ray scattering study of interactions in concentrated silca colloidal suspensions. *Langmuir* **22**, 546-552.

Randolph, A. D. and Larson, M. A. (1988) Theory of particulate processes: Analysis and techniques of continous crystallization, Academic Press Inc., San Diego, California.

Ravindra, R. and Winter, R. (2003) Pressure perturbation calorimetric studies of the solvation properties and the thermal unfolding of proteins in solution. *Z. Phys. Chem.* **217**, 1221-1243.

Ravindra, R. and Winter, R. (2004) Pressure perturbation calorimetry: A new technique provides surprising results on the effects of co-solvents on protein solvation and unfolding behaviour. *Chemphyschem*. **5**, 566-571.

Reynolds, J. A. and Tanford, C. (1970) The gross conformation of protein-sodium dodecyl sulfate complexes. *J. Biol. Chem.* **245**, 5161-5165.

Rosenberger, F., Vekilov, P. G., Muschol, M. and Thomas, B. R. (1996) Nucleation and crystallization of globular proteins-what we know and what is missing. *J. Cryst. Growth*. **168**, 1-27.

Scharnagal, C., Rief, M. and Friedrich, J. (2005) Stability of proteins: Temperature, pressure and the role of solvent. *Biochim. Biophys. Acta* **1749 (2)**, 187-213.

Sluzky, V., Tamada, J. A., Klibanov, A. M. and Langer, R. (1991) Kinetics of insulin aggregation in aqueous solutions upon agitation in the presence of hydrophobic surfaces. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 9377-9381.

Sola-Penna, M. and Meyer-Fernandez, J. R. (1998) Stabilization against thermal inactivation promoted by sugars on enzyme structure and function: Why is trehalose more effective than other sugars? *Arch. Biochem. Biophys.* **360**, 10-14.

Steinhardt, J. and Reynolds, J. A. (1969) Multiple equilibria in proteins. Academic Press, New York.

Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W. C., Egelhaaf, S. U. and Schurtenberger, P. (2004) Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature* **432**, 492-495.

Svergun, D. I. (1992) Determination of the regularization parameter in indirect transform method using perceptual criteria. *J. Appl. Cryst*. **25**, 495-503.

Svergun, D., Barberato, C. and Koch, M. H. J. (1995) CRYSOL–A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Cryst.* **28**, 768-773.

Svergun, D. I., Richard, S., Koch, M. H. J., Sayers, Z., Kuprin, S. and Zaccai, G. (1998) Protein hydration in solution: Experimental observation by x-ray and neutron scattering *Proc. Natl. Acad. Sci. U.S.A.* **95**, 2267-2272.

Svergun, D. I., Petoukhov, M. V., and Koch, H. J. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* **80**, 2946-2953.

Svergun, D. I. and Koch, M. H. J. (2003) Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.* **66**, 1735-1782.

Tanaka, Y., Tsumoto, K., Umetsu, M., Nakanishi, T., Yasutake, Y., Sakai, N., Yao, M., Tanaka, I. Arakawa, T., Kumagai, I. (2004) Structural evidence for guanidine-protein side chain interactions: Crystal structure of CutA from Pyrococcus horikoshii in 3 M guanidine hydrochloride. *Biochem. Biophys. Res. Commun.* **323**, 185-191.

Tanford, C., Paritosh, K. De. and Taggart, V.G. (1960) The role of the α-helix in the structure of proteins. Optical rotatory dispersion of β-Lactoglobulin. *J. Am. Chem. Soc*. **82**, 6028-6034.

Tanford, C. (1964) Isothermal unfolding of globular proteins in aqueous urea Solutions *J. Am. Chem. Soc*. **86**, 2050-2059.

Tanford, C., Kawahara, K. and Lapanje, S. (1967) Intrinsic viscosities and sedimentation coefficients in concentrated guanidine hydrochloride. *J. Am. Chem. Soc*. **89**, 729-736.

Tanford, C. and Roxby, R. (1972) Interpretation of protein titration curves. Application to lysozyme. *Biochemistry* **11**, 2192-2198.

Tardieu, A., Veretout, F., Krop, B. and Slingsby, C. (1992) Protein interactions in the calf eye lens: Interactions between β-crystallins are repulsive whereas in γ-crystallins they are attractive. *Eur. Biophys.* J. **21**, 1-12.

Tardieu, A. (1994) in: Neutron and synchrotron radiation for condensed matter studies, HERCULES, Les editions de physique, Les Ulis and Springer, Berlin. **3**, 145.

Tardieu, A., Verge, A. Le., Malfois, M., Bonnete, F., Finet, S., Ries-Kautt, M. and Belloni, L. (1999) Proteins in solution : From X-ray scattering intensities to interaction potentials. *J. Cryst. Growth* **196**, 193-203.

Thomas, P. D. and Dill, K. A. (1993) Local and nonlocal interactions in globular proteins and mechanisms of alcohol denaturation. *Protein Sci.* **2**, 2050-2065.

Timasheff, S. N. (2002) Protein-solvent preferential interactions, protein hydration, and the modulation of biochemical reactions by solvent components. *Proc. Natl. acad. Sci.* **99**, 9721-9726.

Timasheff, S. N. and Xie, G. (2003) Preferential interactions of urea with lysozyme and their linkage to protein denaturation. *Biophys Chem*. **105**, 421-448.

Tolan, M., Weis, T., Wille, K. And Westphal, C., (2003) DELTA: Synchrotron light in Nordrhein-Westfalen *Synchrotron Rad. News* **16**, 9-11.

Valente, J. J., Verma, K. S., Manning, M. C., Wilson, W. W. and Henry, C. S. (2005) Second virial coefficient studies of cosolvent-induced protein self-interaction. *Biophys. J.* **89**, 4211-4218.

Velev, O. D., Kaler, E. W., and Lenhoff, A. M. (1998) Protein interactions in solution characterized by light and neutron scattering: Comparison of lysozyme and chymotrypsinogen. *Biophys. J.* **75**, 2682-2697.

Völker, J. and Breslauer, K. J. (2005) Communication between non-acting macromolecules. *Annu. Rev. Biophys. Biomol. Struct.* **34**, 21-42.

Walger, R., Lee, T.C. and Cammers-Goodwin, A. (1998) An indirect chaotropic mechanism for the stabilization of helix conformation of peptides in aqueous trifluoroethanol and hexafluoro-2-propanol. *J. Am. Chem. Soc.* **120**, 5073-5079.

Whittingham, J. L., Scott, D. J., Chance, K., Wilson, A., Finch, J., Brange, J. and Dodson, G. G. (2002) Insulin at pH 2: Structural analysis of the conditions promoting insulin fibre formation. *J. Mol. Biol.* **318**, 479 - 490.

Young, A. C., Tilton, R.F. and Dewan, J. C. (1994) Thermal expansion of hen egg-white lysozyme. Comparison of the 1.9 Å resolution structures of the tetragonal form of the enzyme at 100° K and 298° K. *J. Mol. Biol.* **235**, 302−317.

List of Abbreviations:

Å	Angström
AFM	Atomic force microscopy
$A(Q)$	Scattering amplitude
β	$1/ k_B T$
B_{22}	Second virial coefficient
d	Range of the attractive potential
$\Delta\rho(r)$	Excess scattering length density distribution
DESY	Deutsches Elektronen Synchrotron
DLVO potential	Derjaguin-Landau-Verwey-Overbeek potential
D_{max}	Maximum particle diameter
DNA	Deoxyribonucleic acid
$DR(Q)$	Detector response with respect to Q
e	elementary charge
ε	Dielectric permittivity of the medium
f	Scattering length
FT-IR	Fourier Transform – Infrared Spectroscopy
$\gamma(r)$	Sherically averaged autocorrelation function of the excess scattering density
$g(r)$	Radial pair distribution function
GdmCl	Guanidinium chloride
$I(Q)$	Scattered intensity as a function of momentum transfer
J	modulus of the depth of the attractive potential
j_1	First order spherical Bessel-function
κ	Reciprocal Debye-Hückel screening length
k_B	Boltzmann constant
kD	Kilo Dalton
n	particle number density
N_e	Number of electrons
n_i	Mean density of ions i
$p(r)$	Distance distribution function
$P(Q)$	Form factor
PDB	Protein data base
pI	Isoelectric point
PPC	Pressure perturbation calorimetry
Q	Momentum transfer
r	Distance in real space
r_0	Thomson radius
R_g	Radius of gyration
RPA	Random phase approximation
$\rho(r)$	Scattering length density distribution
σ	Hard sphere diameter
$S(Q)$	Structure factor

$S_0(Q)$	Structure factor of the reference system (empty core model)
SANS	Small angle neutron scattering
SAW	Superconducting asymmetric wiggler
SAXS	Small angle X-ray scattering
$S_{exp}(Q)$	Experimental structure factor
$S_{th}(Q)$	Theoretical structure factor
T	Temperature
TFE	2,2,2-trifluoroethanol
$V(Q)$	Total pair potential in Q-space (reciprocal space)
$V(r)$	Total pair potential in real space
$V_0(r)$	Reference potential (Hard sphere potential)
$V_1(Q)$	Fourier-transform of the perturbation potential
$V_1(r)$	Perturbation potential
$V_C(Q)$	Fourier-transform of the Coulomb-potential
$V_C(r)$	Repulsive Coulomb-potential in real space
$V_Y(Q)$	Fourier-transform of the Yukawa-type potential
$V_Y(r)$	Attractive Yukawa-type potential in real space
Z	Effective charge on the protein molecule

Appendix 1:

Mathematical Scripts for "Mathematica"

a) List of parameters

(**********)
(**********Parameters-list for fitting $S_{exp}(Q)$ **********)

(* units in cgs-system (Narayanan, J. and Liu, X .Y., Biophysical Journal 84, 523-532) *)

(**********)

n1 = 4.211 10^17 (* *absolute number of particles for lysozyme 1 wt % (10 mg/ml) per cubic centimeter* *)

N_A = 6.022 10^23 (* *Avogadro's number* *)

A = 1 10^-8 (* *1 Angström* *)

kb = 1.3806*10^-16 (* *Boltzmann-constant in cgs-units (g cm^2 / (s^2 K))* *)

e = 4.8033 * 10^-10 (* *elementary charge in cgs-units* *)

es = 78 (* *relative dielectric constant of water in C^2/(J cm)* *)

T = 303 (* *Temperature in K* *)

sigma = 28 A (* *effective diameter of lysozyme molecule* *)

Z = 6 (* *effective charge of one Lysozyme molecule* *)

kappa = ((4 Pi e^2)/(e kb T)(0.561*20*10^-6 N_A 1^2 + 0.381*20*10^-6 N_A 2^2 +0.031*20*10^-6 N_A 3^2 + 31.7*10^-6 N_A))^0.5

(* *reciprocal Debye-Hückel-screening length; Ionic strength calculated with absolute number of particles per cubic centimeter (based on the dissociation equilibrium of 20 mM citrate at pH 4.6* *)

(* *kappa^-1 [nm] = 0.304/I^0.5 ; (Glaser, R., "Biophysik"; Gustav Fischer Verlag Jena; ISBN 3-334-60967-7; S. 72* *)

b) List of definitions/formulas

```
(*************************************************************************

                Formulas for the potentials, form factor
                structure factor and scattering curve

*************************************************************************)

 (************************************
Volume/packing fraction of lysozyme
************************************)

eta = Pi sigma^3 n/6

(************************************
Radius of the particle
************************************)

R = sigma / 2

(*****************************************
Screened Coulomb potential in real space "r"
*****************************************)

 VCr = (( Z^2 e ^2) / (es(1 + 0.5 kappa sigma)^2)) * Exp[-(kappa(r - sigma))] / r

(*************************************************************************
*****
Screened coulomb Potential in Q-Space.Reference: Langmuir 1992, 8, 2210-2214
*************************************************************************
*****)

VCQ = 4 Pi Z^2 e^2 (kappa Sin[Q sigma] + Q Cos[Q sigma]) / (es(1 + 0.5 kappa
sigma)^2 Q (Q^2 + kappa^2))

        (*** i.e.  VCQi = Integrate[(VCr Sin[Q r] 4 Pi r^2) / (Q r), {r, sigma(1+1*10^-
13), 3 sigma}] *)


(*****************************************
sum of attractive potentials in real space
*****************************************)

VYr = -J (sigma/r) Exp[-(r - sigma)/d]
```

```
(*************************************************************************
Fourier transform of sum of attractive potentials in Q- space, Reference: Langmuir 1992,
8, 2210-2214
*************************************************************************)

VYQ = Integrate[(VYr Sin[Q r] 4 Pi r^2) / (Q r), {r, sigma(1+1*10^-13), 3 sigma}]

(*******************************************
First order spherical Bessel function for S0(Q)
*******************************************)

j1 = (Sin[Q sigma] - Q sigma Cos[Q sigma]) / (Q sigma)^2


(*************************************
Structure factor for reference system
*************************************)

S0 = ((1 - (12 eta (eta(3 - eta^2) - 2)*j1/(Q sigma))/(1-eta)^4))^-1

(**************
Structure factor calculated from theory, Sth(Q)
***************)

SQ = S0*(1 + (1/(kb*T))*n*S0*(VYQ+VCQ))^-1
```

c) Comparing list for fitting experimental structure factors (e.g. lysozyme 4 wt %, 40 mg/ml)

```
ReadList["PATH /definitions.txt"];

(*** "PATH" is path of directory, where the lists or experimental structure factors are saved***)

fit = Table[Re[SQ], {Q, 0.1*10^7, 2.0*10^7, 2.2*10^4}] (****SQ is liquid state theoretical structure factor, S_th(Q)*****)

temp = 0

For[i = 1, i < 800, i = i + 10, {temp = temp + ((fit[[i]] - Sexp4[[i]][[2]])^2)^0.5}]
(****Sexp4 is  experimental structure factor****)

If[temp < Abwalt, {Abwalt = temp, d >>" PATH /d_4", J/(kb T) >> " PATH /J_4"}]
```

d) List for starting several data-sets fitting experimental structure factors

```
(*************************************************************************
                start_all.txt: List to start fit for several data-sets
*************************************************************************)

(***********************************
   Read non-fitting parameters
***********************************)

ReadList["PATH /parameters.txt"];
ReadList["PATH /definitions.txt"];

(*******************************
  Read measured data
*******************************)

Sexp4 = ReadList["PATH /S4_0.DAT", Number, RecordLists->True]

Sexp10 = ReadList["PATH /S10_0.DAT", Number, RecordLists->True]


(*******************************
   Fitting 4 % solution
*******************************)

Print["Fitting 4% solution"]

n = 4 n1

For[J = 1 kb T, J < 50 kb T, J = J +0.1 kb T, {Print[J/(kb T)],For[d = 0.25 A, d < 50 A, d
= d +0.25 A , {ReadList["PATH /compare4.txt"]}]}]

Abwalt = Infinity

(******* Plot result ***********)

d = << " PATH /d_4"

J = << " PATH /J_4"*kb T

n = 4 n1

ReadList["PATH /definitions.txt"]
```

```
calc4 = Plot[SQ, {Q, 0, 2.0*10^7}]

SexpPlot4 = ListPlot[Semp4, PlotStyle->{RGBColor[1,0,0]}, PlotRange->{{0,
2*10^7},{0, 1.5}}]

Show[SexpPlot4, calc4]

Print["J = ", J/(kb T), " k T", "   d = ", d/A, " A", "   sigma = ", sigma/A, " A" ]
Print["n = ", n ,"   Z = ", Z,]

(*****************************
    Fitting 10 % solution
*****************************)

Print["Fitting 10% solution"]

n = 10 n1

For[J = 1 kb T, J < 50 kb T, J = J +0.1 kb T, {Print[J/(kb T)],For[d = 0.25 A, d < 50 A, d
= d +0.25 A , {ReadList["PATH /compare10.txt"]}]}]

Abwalt = Infinity

(******* Plot result ***********)

d = << " PATH /d_10"

J = << " PATH /J_10"*kb T

n = 10 n1

ReadList["PATH /definitions.txt"]

calc10 = Plot[SQ, {Q, 0, 2.0*10^7}]

SexpPlot10 = ListPlot[Semp10, PlotStyle->{RGBColor[0,0,1]},PlotRange->{{0,
2.0*10^7},{0, 1.5}}]

Show[SexpPlot10, calc10]

Print["J = ", J/(kb T), " k T", "   d = ", d/A, " A", "   sigma = ", sigma/A, " A" ]
Print["n = ", n ,"   Z = ", Z, " ]

Show[SexpPlot 4,  SexpPlot 10, calc10,  calc10, PlotRange->{{0, 2.0*10^7},{0, 1.5}}]

Print["Fit pure finished"]
```

e) List for retrieving data for best fits experimental/theoretical structure factors and potentials

```
ReadList["PATH /parameters.txt"]

Sexp4 = ReadList["PATH /S4_0.DAT", Number, RecordLists->True]

Sexp10 = ReadList["PATH /S10_0.DAT", Number, RecordLists->True]

SexpPlot4 = ListPlot[Semp4,  PlotStyle->{RGBColor[0,0,1]}]

SexpPlot 10 = ListPlot[Semp10,  PlotStyle->{RGBColor[1,0,0]}]

Table[Q, {Q, 0.1*10^7, 2.0*10^7, 2.2*10^4}] >> " PATH /Q_val"

n = 4 n1

d = << " PATH /d_4"

J = << " PATH "*kb T

ReadList["PATH /definitions.txt"]

calc4 = Plot[SQ, {Q, 0, 2.0*10^7}]

Table[Re[SQ], {Q, 0.1*10^7, 2.0*10^7, 2.2*10^4}] >>" PATH /table_4_0"

Table[VCr + VYr, {r, 36 A, 3 sigma, 1 A}] >> " PATH /table_pot_4"

n = 10 n1
```

```
d = << " PATH /d_10"

J = << " PATH /J_10"*kb T

ReadList["PATH /definitions.txt"]

calc10 = Plot[SQ, {Q, 0, 2.0*10^7}]

Table[Re[SQ], {Q, 0.1*10^7, 2.0*10^7, 2.2*10^4}] >> " PATH /table_10_0"

Table[VCr + VYr, {r, sigma, 3 sigma, 1 A}] >> " PATH /table_pot_10"

Show[SexpPlot4, calc4, SexpPlot 10, calc10]
```

Acknowledgements

The work carried out in the last three and half years by me is a minor contribution in scientific research, which is getting enriched by many scientists in various disciplines.

Fore mostly, my heartiest gratitude is indebted to Prof. Dr. Roland Winter for highly intellectual and skillfully valuable guidance throughout this research project till its accomplishment. Multidisciplinary research area including Chemistry, Physics and Biology with well-planned and thoughtfully-designed approach made this work highly interesting and demanding as well. I pay most cordial thanks to my Doctor Father for his endless efforts and immense help in arranging all kind of experimental and theoretical resources.

Dr. Karsten Vogtt made my first introduction to Small angle X-ray scattering with a Kratky camera and then this lead to further higher levels of scientific discussions. I am obliged to him for his always kind and well-informative suggestions.

I thank PD Dr. Claus Czeslik for worthful discussions, which improved my ideas to fully understand the basics and finally writing articles as well.

Dr. Manfred Roessle and Dr. Dmitri Svergun showed me measuring procedures at X33 beamline in DESY-synchrotron source, Germany. I would like to thank them for their professional help and making the first measurements feasible.

I thank Christina Krywka and Prof. Dr. Metin Tolan for providing me a lot of experimental time at the beamline BL9 at DELTA in Dortmund, Germany. Knowledgeable theoretical discussions with Professor Tolan for analyzing the scattering data made me more confident to clarify some doubts.

My stay in Dortmund was made joyful and cheerful by fantastic and wonderful company of many friends. I thank all of them especially, Dr. Karsten Vogtt with whom I was discussing so many and diverse topics. Evolution and theism were among the favourite topics which cleared lot of my personal concepts and made me more flexible and rational in my approach.

I thank Lally Mitra for her loving, caring and nice company and also for introducing me to PPC and DSC measurements.

I am very happy that I met Gurpreet Singh, who is a really wonderful person and amazing friend.

I thank o Michael Sulc, Nagarajan Perisamy, Vytautas Smirnovas, Suman Jha, Diana Radovan, Andrea Gohlke, Dr. Rajesh Mishra, Dr. Katrin Weise, Shuang Zhao and Maximilian Andrews for providing co-operative and diversely cultural working environment. I thank Maximilian Andrews also for his assistance to correct my article in American English.

I thank a lot and appreciate the contributions of Andrea Kreusel, Kirsten Skodzik, Milan Saskovic and Dr. Werner Horstmann. I am thankful to Bertina Schuppan for her help in making diagrams and posters.

I would like to thank the International Max-Planck Research School – Chemical Biology (IMPRS-CB) for providing financial support and to Dr. Jutta Roetter for arranging very excellent lecture series in the field of Chemical Biology.

I thank PD Dr. Claus Czeslik and Dr. Jens Müller for being members of the examination committee.

In the last but not least, I would like to thank and acknowledge my parents for their endless love and all efforts throughout my life, which is always the supreme motivating power for me. I thank my brothers and sisters as well for their love which keeps me always happy.

Nadeem Javid, Dortmund, May 2007