Salt Effects in Protein Solutions

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Felix Roosen-Runge

aus Göttingen

Tübingen

2013

Tag der mündlichen Qualifikation:	19.12.2013
Dekan:	Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter:	Prof. Dr. Frank Schreiber
2. Berichterstatter:	Prof. Dr. Roland Roth

Contents

Chapter 1

Introduction

Solutions of globular proteins are ubiquitous in nature and represent a complex system with a rich phase behavior governed by various environmental parameters such as pH, temperature and salt content. The understanding and control of this phase behavior is an essential cornerstone for both bio-related applications and better description of cellular processes (Gunton et al., 2007). Using the presence of salts, interesting questions on the statistical physics of protein solutions can be addressed, since salt potentially acts as a second control parameter beside temperature. The salt–protein interaction varies the protein–protein interaction, which in turn controls phase behavior and nucleation processes. This introduction provides an overview on the topics under investigation and embeds the thesis into the scientific literature.

1.1 Proteins as an Example of Biological Soft Matter

Proteins are essential in all processes of life. Proteins promote DNA translation and methylation, act as enzymes in metabolic pathways, and are involved in structure formation via fibrillation in the cytoskeleton and control of membrane shapes (Alberts et al., 2008). Each protein is characterized by its specific cellular function, which is related to the individual protein structure and dynamics. From a physical perspective, proteins combine different generic features of soft matter.

First, globular proteins in solutions have been used as a monodisperse and natural model system for colloids. Phase behavior and diffusional dynamics of protein solutions can be successfully described at least qualitatively in terms of colloidal physics. Second, the hierarchical structure of proteins is the key to a wide variety of structures and functions in the protein domain: from the primary structure—a linear chain of amino acids encoded by the DNA sequence—the secondary structure of the chain emerges following generic templates of α-helical, β-sheet or random coil formation. These secondary structures assemble to compact protein domains (tertiary structure) that can finally contribute to a quarternary structure of the protein. Proteins thus cannot be seen only as macromolecular colloids, but also include features of a polymer such as folding and denaturation as well as

internal dynamics and spectroscopic properties of the proteins. Third, protein structures as well as their interactions with the environment are governed by water-mediated effects. Hydrophobic and hydrophilic patterns on the surface are important factors governing protein association in the case of molecular recognition and embedding into cellular structures such as biomembranes.

Proteins thus represent an interesting research topic, in particular when both biological (mal)function and physical description can be complemented, leading to a more complete understanding. Important challenging questions include the formation of amyloid fibrils in Alzheimer's disease (Sipe and Cohen, 2000; Hardy and Selkoe, 2002) and protein condensation in the case of cataract (Benedek, 1997; Gunton et al., 2007). In this context, the view point of soft matter cannot only contribute a more detailed understanding of proteins and cellular processes, but can also help to advance applications such as protein crystallization (McPherson, 2004; George and Wilson, 1994; Vliegenthart and Lekkerkerker, 2000) and self-assembly of nanostructures from protein building blocks (Knowles et al., 2010; Ninham and Nostro, 2010).

Proteins are complex solutes with a non-spherical, rough, porous and flexible shape and inhomogeneous surface patterns of charge and hydrophobicity. Furthermore, proteins are subject to a variety of interactions, including Coulomb repulsion and van der Waals interaction as well as solvent-mediated interactions such as depletion and hydrophobic interaction. Because of this complexity, theoretical results for colloidal model systems such as spheres or rods can only be applied with care. The modeling and understanding of proteins in solution is a challenging but promising research field for statistical physics (cf. Ch. 10).

1.2 Protein–Ion Interaction

Information on interactions between proteins and ions is essential to the understanding of the resulting phase behavior of protein solutions. In the following subsections, we discuss the basic case for unspecific protein–ion interaction—the Poisson-Boltzmann theory—and then move on with three deviations from this behavior: the Hofmeister series, counterion condensation and binding of ions to the protein surface.

1.2.1 Ion Distribution and Charge Screening: Poisson-Boltzmann and DLVO Theory

Considering a charged object in a solution with charged ions, the Poisson-Boltzmann theory provides a basic mean-field approach to describe the ion distribution. This approach combines the exact Poisson equation with a mean-field relation between the electrostatic potential and the potential of mean force on the ions (Gouy, 1910; Chapman, 1913; Messina, 2009). The resulting ion distribution around charged objects forms the so-called electrostatic double-layer that causes a screening of charges in electrolyte solutions. The

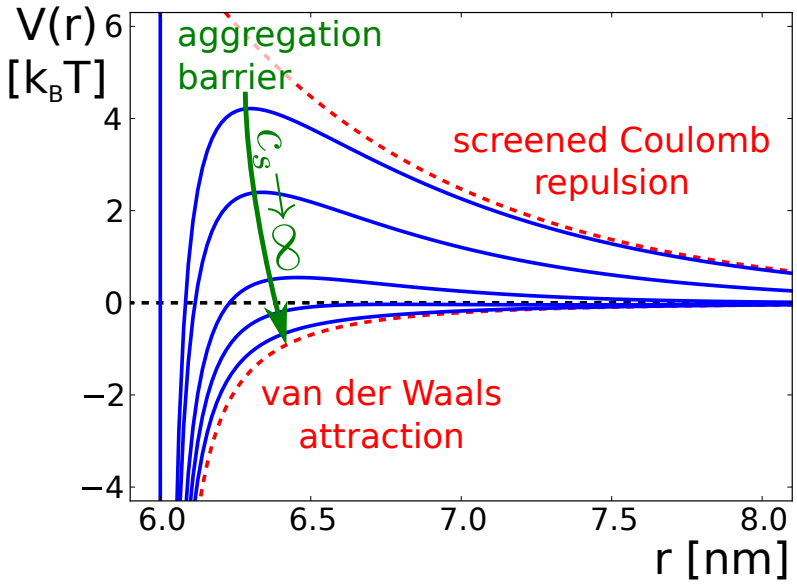

Figure 1.1: DLVO potential for varying salt concentration c_s. With increasing c_s, the potential changes from repulsive to attractive. The aggregation barrier reflects the charge stabilization behavior that becomes weaker due to charge screening.

Poisson-Boltzmann theory and its linearized version, the Debye-Hückel theory (Debye and Hückel, 1923), provides a very useful and important framework for the understanding of electrostatic phenomena in soft matter. Experimentally, the applicability of the Poisson-Boltzmann theory has been supported e.g. for distributions of mono- and divalent ions around DNA (Andresen et al., 2004; Das et al., 2003; Burak et al., 2004).

One very important consequence of the Poisson-Boltzmann approach is the DLVO theory named after Derjaguin and Landau (1941) and Verwey and Overbeek (1948). In the DLVO potential, screened Coulomb interaction and van der Waals attraction are combined to explain the charge stabilization of solutions with charged solutes (Fig. 1.1). With increasing ionic strength, the charge screening decreases the electrostatic repulsion more effectively and on shorter ranges, and finally the van der Waals attraction causes aggregation and precipitation.

The Poisson-Boltzmann theory is based on strong assumptions. In particular it ignores any ion–ion correlations, arising e.g. due to excluded volume and electrostatics. Furthermore, other ion properties such as polarizability and hydration effects are not included, but can actually cause significant effects. On the one hand, it is thus not surprising that multiple ion effects have been observed that cannot be explained by Poisson-Boltzmann theory and DLVO theory (Petsev and Vekilov, 2000; Boström et al., 2001; Messina, 2009). On the other hand, it is even more interesting that the Poisson-Boltzmann theory and the DLVO theory performs so well in many cases, providing approximate theoretical predictions when a full description of the system is out of reach even with elaborate and costly computer simulations (cf. Ch. 5).

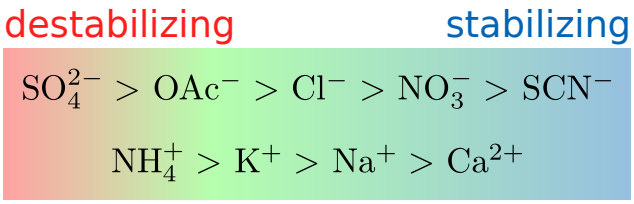

Figure 1.2: Part of the Hofmeister series for anions and cations. Ions on the left hand side of the series destabilize solutions and "salt out" solutes, whereas ions on the right stabilize ("salt in") solutions.

1.2.2 Ion Effects on the Protein Solubility: The Hofmeister Series

Salt-induced phase behavior in protein solutions ranges back to the Hofmeister series on protein solubility in the presence of different salts (Hofmeister, 1888) and the related salting-in and salting-out behavior (Arakawa and Timasheff, 1984). Figure 1.2 shows part of the Hofmeister series for anions and cations. The combination of the cation and anion strongly affects the ability of the salt to precipitate ("salt out") or stabilize ("salt in") colloidal solutions (Collins and Washabaugh, 1985). These variations of the phase behavior cannot be explained by the DLVO theory and imply that the protein–ion interactions vary considerably beyond Poisson-Boltzmann theory due to ion-specific effects.

A comprehensive molecular understanding of the ion-specific effects is a challenge for theory (Kunz, 2010), although it is clear that water-mediated effects are a key ingredient. Baldwin (1996) argues that the Hofmeister effect can be understood considering the two competing abilities of ions to "salt out" nonpolar functional groups and "salt in" the polar peptide group.

A prominent theme for the molecular origin of the Hofmeister effect is the propensity of ions to change the water structure, i.e. the ion hydration (Collins, 2004, 1997). If certain ions—so-called "kosmotropes"—interact with water strongly, the surrounding water is aligned relative to the ion, thus additional water structure is formed. Ions with weak interaction and inappropriate size—so-called "chaotropes"—are not able to induce any water structure and even distort the bulk structure. Another theme involves the change of the dielectric constant at the protein–water interface, which allows non-localized adsorption of polarizable ions at non-polar, hydrophobic areas of the protein surface (Jungwirth and Winter, 2008; Levin, 2009; Dahirel and Jardat, 2010), representing another possible mechanism for the Hofmeister effect via dispersion forces (Ninham and Nostro, 2010; Schwierz et al., 2010; Heyda et al., 2009). Finally, the interfacial tension of the protein–water interface has been linked to the Hofmeister series (Dér et al., 2007). Thus, although of practical importance and known for over a century, the Hofmeister effect, and salt effects on protein solutions in general, remain an interesting and challenging field of research (cf. Ch. 4).

1.2.3 Condensation of Counterions

If the charge density in the system is strong, significantly varied ion distribution are obtained compared to the mean-field Poisson-Boltzmann approach. Manning (1969) found a condensation of counterions on surfaces as long as the surface charge density is higher than a critical value dependent on the surface geometry and counterion valence.

More recently, statistically more advanced approaches accounting for ion–ion correlations due to strong Coulomb coupling have found counterion condensation at strongly charged surfaces as well as ion distributions that depend on the valence and size of the counterions (Messina, 2009; Moreira and Netz, 2000; Deserno et al., 2000; Grosberg et al., 2002; Terao, 2006). In this case, ion–ion correlations cause an inversion of the surface charge (Shklovskii, 1999; Grosberg et al., 2002; Besteman et al., 2004, 2007). Theoretical approaches even predict a so-called "giant overcharging" (Nguyen et al., 2000a) due to increasing monovalent salt content, while simulations suggest a lower reversed charge for these conditions (Lenz and Holm, 2008).

The effects of ion–ion correlations are, in general, expected to be small compared to specific interactions between ions and surfaces (Besteman et al., 2004). While ion–ion correlations might add a finite contribution to the protein–ion interaction, other more specific effects appear to be more relevant (Zhang et al., 2010) (cf. Ch. 6).

1.2.4 Functional Groups at the Protein Surface: Charge Regulation and Ion Binding

On the protein surface, numerous side chains with different chemical properties are exposed to the solvent. As a first consequence, charge regulation of the protein surface is occuring in aqueous solution. Functional surface groups—basic and acidic amino acid side chains (Lys, Arg, Glu, Asp, His) and the protein carboxy and amino terminus—are (de)protonated depending on the pH of the solution (Tanford et al., 1955; Tanford and Kirkwood, 1957), thus coupling pH to protein surface charge.

As a second effect, also other ions than the hydronium ion interact with functional groups. Numerous studies report equilibrium constants for the binding of multivalent counterions to proteins specialized for metal storage and transport (Permyakov, 2009; Williams and Peacocke, 1967; Aasa et al., 1963; Trapp, 1983; Martin, 1986). Models for the ion binding have been developed in order to understand interactions of proteins with ions and ligands (Scatchard, 1949; Feldman, 1972; Klotz, 1982; Tipping, 1998). On the molecular level, amino acids with carboxylate, thiol, thioether, and imidazole groups bind transition metal ions coordinatively (Tainer et al., 1991; Harding, 2001; Pastore et al., 2007; Berthon, 1995). The binding of ions is enhanced at hydrophilic sites surrounded by hydrophobic surface areas (Yamashita et al., 1990). The overall ubiquity of these surface groups suggests that the association of salt counterions with inversely charged side chains at the protein surface is the basic model for the interaction of ions and proteins (Jungwirth and Winter, 2008) (cf. Ch. 5 and 9).

This notion has been explicitly supported for a study on the oligopeptide tetra-aspartate. Kubíčková et al. (2012) observed a charge inversion both experimentally and by molecular dynamics simulations for tetra-aspartate with trivalent cations. Mono- and divalent ions also decreased the overall charge, but did not overcome the initial negative protein charge. As the basic mechanism, the ion binding to carboxylic acids is evidenced by radial distribution functions that also explain the different behavior of multi- and monovalent cations (Kubíčková et al., 2012).

1.3 Ion-Induced Phase Behavior

A rich phase behavior has been found in protein solutions, including liquid–liquid phase separation (LLPS), protein clusters, and crystallization as well as other aggregates such as fibers. The nucleation kinetics differ considerably for different phases, which enables metastable phases such as LLPS or clusters as precursor structures during crystallization as well as arrested phases such as gels (Anderson and Lekkerkerker, 2002; Gibaud and Schurtenberger, 2009). In this section, we provide an overview on the different phenomena that also play an important role in the present thesis. In order to embed the special case of protein solutions into a broader context of biological and soft matter, also comparable systems such as DNA or colloids are discussed.

1.3.1 Protein Surface Charge and Ion-Induced Charge Inversion

Charges on the protein surface are an important property that ensures stability and function of proteins (Warshel and Levitt, 1976; Pace et al., 2000; Lund and Jönsson, 2005). Charge patterns represent anisotropic interaction patches that affects the phase behavior of protein solutions (Sear, 1999; Kern and Frenkel, 2003; Gögelein et al., 2008; Bianchi et al., 2011) as well as pathways for aggregation and crystallization (Whitelam, 2010a; Buell et al., 2009; Fusco and Charbonneau, 2013).

Protein–protein interactions are linked to the charge regulation that in turn is a complex process depending on system geometry and ion specific effects like binding or condensation. A comprehensive understanding of charge regulation at the protein surface has to also account for ion specific effects such as binding and condensation as well as for the system geometry, e.g. the proximity of a wall (Lund and Jönsson, 2005; Kurut and Lund, 2013).

A special case is the charge inversion of surfaces in the presence of counterions. A comprehensive understanding of the charge inversion has to account for both local ion binding and non-local contributions such as ion–ion correlations and hydrophobic effects (Lyklema, 2006; Travesset and Vangaveti, 2009). Charge inversion has been observed for a broad range of systems such as silica spheres (Besteman et al., 2004), insoluble oxides (Lyklema, 2006) and also biological systems such as DNA (Nguyen et al., 2000a).

In particular, a charge inversion has been observed in solutions of globular proteins with multivalent cations (Zhang et al., 2008, 2010). The lower charge density of the protein surface disable ion–ion correlations to be the main cause of charge inversion. Instead, Zhang et al. (2011) support the notion of a charge inversion due to ion binding to acidic residues on the protein surface, based on information from crystal structures (cf. Ch. 5 and 9).

1.3.2 Reentrant Condensation

The inversion of the surface charge is related to a specific phase behavior called reentrant condensation. While the solutions are charge-stabilized at low and high salt concentration, a condensed regime is observed for intermediate salt concentration. Although induced by salts, this phenomenon is clearly beyond the usual salting in and salting out behavior and needs an individual explanation.

The reentrant condensation of DNA macromolecules has been studied intensively and can be explained by an charge inversion and like-charge attractions induced by ion–ion correlations of multivalent counterions (Saminathan et al., 1999; Nguyen et al., 2000b; Burak et al., 2004).

A macroscopically similar observation has been observed in protein solutions with trivalent cations (Zhang et al., 2008, 2010; Ianeselli et al., 2010). As discussed in the previous subsection, ion binding dominates the protein case. A more detailed discussion of the reentrant condensation in protein solutions is subject of the present thesis (Ch. 5 and 6).

1.3.3 Liquid–Liquid Phase Separation (LLPS)

Liquid–Liquid Phase Separation (LLPS) in protein solutions was first found for natural systems, namely in calf, rat and human eye lenses (Tanaka and Benedek, 1975; Tanaka et al., 1977). The LLPS of different γ-crystallins from the eye lens and lysozyme has been studied in order to understand the formation of cold cataract (Siezen et al., 1985; Broide et al., 1991) as well as to investigate critical phenomena in a biological model system (Ishimoto and Tanaka, 1977; Thomson et al., 1987; Schurtenberger et al., 1989; Fine et al., 1995). A metastable LLPS in hemoglobin solutions has been found to be the primary event of sickle cell anemia (Galkin et al., 2002). Experimental results confirmed the metastability of the LLPS with respect to the solid phase (Berland et al., 1992) as theoretically expected for an attraction that is short-ranged compared to the protein size (Asherie et al., 1996; Rosenbaum et al., 1996; Noro and Frenkel, 2000).

Metastable LLPS in protein solutions is of specific interest due to its connection to protein crystallization that will be discussed in Sec. 1.3.5. In this context, control of the phase behavior is essential to optimize nucleation conditions for high-quality protein crystals. Additives in protein solutions such as PEG, glycerol, monovalent salts or a second protein species have been found to shift the coexistence curve of protein solutions in

temperature (Muschol and Rosenberger, 1997; Galkin and Vekilov, 2000b; Annunziata et al., 2002, 2003; Chen et al., 2004; Mason et al., 2010; Wang et al., 2010, 2011; Gögelein et al., 2012). Theoretical studies have reproduced these shifts based on colloidal models and nonspecific interactions between protein and additives (Sear, 1999; Kern and Frenkel, 2003; Wentzel and Gunton, 2007, 2008; Gögelein et al., 2008; Dorsaz et al., 2011; Gögelein et al., 2012). In the present thesis, the effect of multivalent cations with specific protein–ion interaction on the LLPS of protein solutions is demonstrated and intensively discussed (Ch. 7 and 9).

1.3.4 Cluster Formation

The formation of equilibrium clusters in solutions of charged particles has been predicted by a simple argument (Groenewold and Kegel, 2001, 2004; Sciortino et al., 2004): If particles exhibit a short-ranged attraction and a long-ranged (Coulomb) repulsion, monomers attach due to the attraction until the repulsion of the entire cluster grows strong enough to destabilize further attachment. Computer simulations support this notion of cluster formation due to a balance between competing interactions (Sciortino et al., 2004; Mossa et al., 2004; Kowalczyk et al., 2011; Spohr et al., 2002; Jiang and Wu, 2009; Fierro et al., 2011). Colloidal clusters have been observed experimentally in colloidal solutions (Segrè et al., 2001).

Experiments on lysozyme indicate the formation of equilibrium clusters in concentrated protein solutions (Stradner et al., 2004). Further studies by combining small-angle scattering techniques (SAXS, SANS), nuclear magnetic resonance (NMR), neutron spin echo (NSE), and dynamic light scattering (DLS) suggest that the protein clusters are rather dynamic and transient instead of static and irreversible (Stradner et al., 2006; Liu et al., 2011; Porcar et al., 2010; Barhoum and Yethiraj, 2010; Shukla et al., 2008; Cardinaux et al., 2011; Li et al., 2011). In β-lactoglobulin solutions, Piazza and Iacopini (2002) observed the spontaneous formation of transient clusters with a very short life time. Large metastable clusters of proteins have been studied experimentally by light scattering and Brownian microscopy for proteins such as hemoglobin, lumazine synthase, and lysozyme (Pan et al., 2007, 2010; Gliko et al., 2005, 2007; Li et al., 2011).

Reversible cluster formation in protein solutions is of practical interest, since it would be promising for drug delivery at high antibody volume fraction and moderate viscosity (Johnston et al., 2012). The presence of flexibly bound clusters might affect pathways of protein crystallization, as discussed in the next section (cf. also Ch. 8).

1.3.5 Protein Crystallization and Nucleation Pathways

The lack of a systematic and general procedure to obtain high-quality protein crystals has inspired numerous studies on the connection of phase behavior and crystal nucleation as well as the control of optimum conditions for protein crystallization.

Based on experimental observation, George and Wilson (1994) suggested the so-called crystallization slot for the second virial coefficient. While for too weak attractive interaction the nucleation is very slow, too strong attraction causes multiple nucleation events and irregular and arrested assembly of proteins. This line of thought has been put forward in several studies. Vliegenthart and Lekkerkerker (2000) explain optimum crystallization conditions by two effects: First, the approximity to the critical point of a metastable liquid–liquid phase separation has been found to enhance nucleation rates in simulations considerably due to critical fluctuations (Wolde and Frenkel, 1997). Second, along the binodal of the liquid–liquid phase separation, small dense droplets might act as regions with increased nucleation rates (Haas and Drenth, 1999; Vekilov, 2004, 2010a). Indeed, the nucleation rate of protein crystals has been found to be considerably enhanced close to the LLPS boundary (Galkin and Vekilov, 2000b,a) and close to the critical point (Jion et al., 2006).

Both proposed conditions represent two-step nucleation pathways in the sense that the two order parameters density and structure, which are coupled in classical nucleation theory, are separated and established subsequently. In a first step, the solution forms a dense precursor which then reorders to a structured crystal nucleus (Lutsko and Nicolis, 2006). The exact nature of the precursor in protein solutions remains unclear and is, most likely, not a general feature (Vekilov, 2010b). Besides the LLPS related precursors, also protein clusters might represent precursor structures (Hutchens and Wang, 2007) (cf. Ch. 8). Notably, these findings on two-step pathways not only apply to proteins and colloidal systems, but also to small molecular systems (Vekilov, 2010b; Erdemir et al., 2009; Bonnett et al., 2003; Navrotsky, 2004).

1.3.6 Arrested Phases: Gels and Amorphous Aggregates

Besides the equilibrium properties, also kinetic pathways matter for the observed phase behavior. The LLPS in protein solutions is metastable with respect to the solid phase, and thus is only observable due to differing kinetics of phase transition. However, also arrested phases such as gels and glasses (Segrè et al., 2001; Sztucki et al., 2006; Sciortino et al., 2004) are observed in colloidal systems with short-ranged attractions. For the case of proteins, the gel formation has been related to an arrested metastable LLPS (Gibaud and Schurtenberger, 2009) and the formation of clusters (Baglioni et al., 2004; Cardinaux et al., 2011; Gibaud et al., 2011). Poon (1997) argues that the arrested LLPS might represent the reason that crystals cannot grow at high attraction strength (cf. Ch. 8 and 9).

1.4 Patchy Particle Models for Proteins

The non-spherical shape of proteins and inhomogeneous surface pattern of charge and hydrophobicity suggest that protein interactions cannot be understood by isotropic interaction potentials only. Further indirect evidence for anisotropic interactions stem from protein crystals with low packing fraction and a low critical volume fraction of the liquid–liquid phase separation. Isotropic potentials cannot reproduce these features correctly. A suitable reference frame for anisotropic interaction is provided by patchy particles (Bianchi et al., 2011), i.e. particles interacting with other particles through specific surface areas, the so-called patches. The notion of patches is also supported by the effects of some point mutations in the protein sequence that do not alter the protein shape, but significantly vary the protein solubility (McManus et al., 2007; Buell et al., 2009). Numerical calculations of the second virial coefficient based on real protein structures show that few highly attractive relative orientations dominate the protein interaction (Neal et al., 1999).

During the last two decades, several patchy particle models have been investigated to explain equilibrium and non-equilibrium features of protein solutions. At roughly the same time, Sear (1999) and Lomakin et al. (1999) suggested two different protein models with patchy interactions. Lomakin et al. (1999) used a sphere with homogeneously distributed surface sites to calculate an effective isotropic interaction strength for the crystalline phase via averaging over several neighbor configurations. By this means, the solid-liquid and liquid–liquid phase separation can be rationalized by effective isotropic models (Lomakin et al., 1999). Sear (1999) employed a theoretical model system for associating fluids, which was first explored by Wertheim (1984a,b, 1986a,b), and reformulated by Jackson et al. (1988) and Chapman et al. (1988). Complemented with a cell model for the crystal phase, a full phase diagram for protein solutions could be reproduced (Sear, 1999; Vega and Monson, 1998). The Wertheim theory with a related cell model has been exploited to explicitly model the solubility of lysozyme (Curtis et al., 2001) also in the presence of NaCl (Warren, 2002). Based on the Wertheim theory, various studies on patchy colloids found features comparable to protein solutions (Bianchi et al., 2006, 2008; Liu et al., 2009). In particular, depending on the model parameters such as number, size and strength of the patches, low density crystal phases and low critical volume fractions for the LLPS are obtained that can induce both a stable and metastable LLPS with respect to the solid phase. Complementary simulations on patchy colloids support this picture (Vega and Monson, 1998; Kern and Frenkel, 2003; Bianchi et al., 2006; Liu et al., 2007; Foffi and Sciortino, 2007).

A different theoretical approach was used by Gögelein et al. (2008) to describe the effects of NaCl on the LLPS in lysozyme solutions. Using the thermodynamic perturbation theory of Barker-Henderson, attractive patches as well as an isotropic electrostatic repulsion has been accounted for simultaneously, allowing for an excellent reproduction of experimental results on the LLPS (Gögelein et al., 2008).

Recent studies focused on the question how geometry, specificity and variable strengths of patches affect the phase behavior and crystal nucleation. Strong effects have been found

for dissimilar patches (Dorsaz et al., 2012; Fusco and Charbonneau, 2013) that can shift the critical temperature strongly, even above the solubility line. Using real crystal structures for three related mutants of one protein, a theoretical study parametrized a patchy model, obtaining good agreement with experiments and providing an explanation for the effects of point mutations (Fusco et al., 2013). In simulations on a patchy model geometrically designed according to real proteins, one and two-step crystallization processes have been observed depending on the patch strength and isotropic interactions (Whitelam, 2010a).

Summarizing these findings, patchy particles have been proven to represent a useful and versatile tool to describe and understand equilibrium and non-equilibrium features in protein solutions (cf. Ch. 9).

1.5 Outline

The thesis comprises several studies on phenomena in aqueous solutions of proteins in the presence of salt. The guiding idea of the thesis is that salt ions vary the protein interactions that in turn determine the equilibrium phase behavior and non-equilibrium kinetic pathways:

Guiding research idea						
protein–ion interaction	$\longrightarrow$	protein–protein interaction	$\longrightarrow$	phase behavior of protein solutions	$\longrightarrow$	nucleation behavior in protein solutions

Considerable parts of the present thesis are based on publications in scientific journals. In the following, a brief overview of the thesis is given, summarizing the results of the studies and listing the publications along with a full list of coauthors and their individual contributions, which are of course not to be understood as absolutely clear-cut assignments. In addition, the reference is specified at the beginning of the respective chapter.

1.5.1 Methodological Aspects

After the introductory part, two main experimental methods are mentioned. Following a brief outline of small-angle scattering (Chapter 2), the basics of electrophoretics are summarized (Chapter 3).

Throughout the thesis, proteins as well as the interaction of protein and salt are modeled in various ways. For the protein solutions, colloidal solutions represent a suitable reference system that allows a theoretical sound modeling of and predictions for several aspects of protein solutions. The protein–ion interaction is modeled by the binding of cations to surface groups at the protein surface, thereby introducing effective interactions between the proteins. The applicability of these concepts is discussed in Chapter 10.

1.5.2 Studies Contained in the Thesis

Chapter 4

Hydration and interactions in protein solutions containing concentrated electrolytes studied by small-angle scattering

Fajun Zhang, Felix Roosen-Runge, Maximilian W. A. Skoda, Robert M. J. Jacobs, Marcell Wolf, Philip Callow, Henrich Frielinghaus, Vitaliy Pipich, Sylvain Prevost and Frank Schreiber

Phys. Chem. Chem. Phys., 14 (2012) 2483–2493

Contributions:

Research design	FZ, FS
Experiments	FZ, FRR, MWAS, RMJJ, MW
Technical Assistence	PC, HF, VP, SP
Data analysis and interpretation	FZ, FRR
Paper writing	FZ, FRR, FS

(Zhang et al., 2012b)

Protein–protein interaction is studied by small-angle scattering in the presence of different monovalent salts, addressing the so-called Hofmeister effect. The protein–protein interaction and, in particular, the second virial coefficient is found to follow the Hofmeister series. A consistent colloidal interpretation with effective spheres can be drawn, if effects of hydration and non-spherical shape of the protein molecules are accounted for.

Chapter 5

Interplay of pH and Binding of Multivalent Metal Ions: Charge Inversion and Reentrant Condensation in Protein Solutions

Felix Roosen-Runge*, Benjamin S. Heck, Fajun Zhang, Oliver Kohlbacher and Frank Schreiber*

* corresponding authors

J. Phys. Chem. B, 117 (2013) 5777–5787

Contributions:

Research design	FRR,FZ
Experiments	BSH, FZ
Model development and calculation	FRR
Data analysis and interpretation	FRR
Paper writing	FRR, OK, FS

Roosen-Runge et al. (2013a)

The interaction of protein and multivalent counterions is modeled using association reactions, thereby accounting for counterion binding and charge regulation of protein surface groups and pH effects due to salt hydrolysis. The aggregation barrier from a DLVO potential is successfully used to estimate the reentrant condensation of the protein solutions. The results agree qualitatively and even to a reasonable degree quantitatively with experimental results for the protein charge, determined by the electrophoretic mobility, and for the pH of the solution as well as with experimental reentrant phase diagrams.

Chapter 6

Competing Salt Effects on Reentrant Phase Behavior of Protein Solutions: Tailoring of Interaction Strength and Range by the Binding of Multivalent Ions and Charge Screening

Elena Jordan[†], Felix Roosen-Runge[†*], Sara Leibfarth, Fajun Zhang[*], Michael Sztucki, Andreas Hildebrandt, Oliver Kohlbacher and Frank Schreiber

[†] E. J. and F. R.-R. contributed equally to this paper. [*] corresponding authors

submitted and under review

Contributions:

Research design	FZ, FS
Experiments	EJ, FZ
Technical assistance	MS
Simulations and model calculations	FRR, SL
Data analysis and interpretation	EJ, FRR, SL
Paper writing	EJ, FRR, FZ, FS

Jordan et al. (2013)

The influence of monovalent salt on the reentrant condensation is used to shed light on the mechanisms underlying the phase behavior. Both salt concentrations at the first and second transition are found to increase with increasing monovalent salt concentration. These trends are also reproduced in SAXS measurements. Using results for the protein charge from an analytical association model and Monte Carlo sampling of the detailed binding configurations, theoretical estimates for the transition concentrations are calculated using the aggregation barrier of the DLVO potential or a criterion of a critical charge. The trends at the second transition are consistent with screened electrostatic repulsion of protein charges. The trends at the first transition suggest that the binding of ions is not only reducing the protein charge, but also including additional attraction.

Chapter 7

Charge–controlled metastable liquid–liquid phase separation in protein solutions as a universal pathway towards crystallization

Fajun Zhang, Roland Roth, Marcell Wolf, Felix Roosen-Runge, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Sztucki and Frank Schreiber

Soft Matter, 8 (2012) 1313–1316

Contributions:

Research design	FZ, RR, FS
Experiments	FZ, MW, FRR, MWAS, RMJJ
Technical Assistence	MS
Data analysis and interpretation	FZ, RR, FRR
Paper writing	FZ, RR, FRR, FS
Illustration	FRR

(Zhang et al., 2012c)

The meta-stable liquid–liquid phase separation (LLPS) occuring in the condensation regime of the reentrant condensation is discussed in terms of colloidal systems with short-ranged attraction. The closed area of the LLPS is explained by a varying protein–protein interaction due to binding of multivalent ions. The reduced second virial coefficient measured by SAXS supports this notion. The area showing protein crystallization coincides with suitable nucleation conditions for colloids close to a LLPS.

Chapter 8

The Role of Cluster Formation and Metastable Liquid–Liquid Phase Separation in Protein Crystallization

Fajun Zhang, Felix Roosen-Runge, Andrea Sauter, Roland Roth, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Sztucki and Frank Schreiber

Faraday Disc., 159 (2012) 313–325

Contributions:

Research design	FZ, FS
Experiments	FZ, FRR, AS, MWAS, RMJJ
Technical Assistence	MS
Data analysis and interpretation	FZ, FRR,RR
Paper writing	FZ, FRR, FS

(Zhang et al., 2012a)

The occurrence of two-step nucleation pathways for the crystallization process in protein solutions with multivalent counterions close to the LLPS is discussed. In contrast to other protein systems with LLPS, for which two-step nucleation is observed from the dense protein phase, crystals seem to grow from the dilute phase. Using real-time SAXS monitoring crystallization, the formation of protein clusters as precursors for crystals is suggested.

Chapter 9

Proteins as particles with ion-activated attractive patches

Felix Roosen-Runge, Fajun Zhang, Frank Schreiber and Roland Roth

submitted and under review

Contributions:

Research design	FRR, RR
Model development and calculations	FRR, RR
Interpretation	FRR, FZ, FS, RR
Paper writing	FRR, FS, RR

Roosen-Runge et al. (2013b)

The overall reentrant phase behavior of protein solutions is modeled using particles with attractive patches that depend on the binding state of multivalent counterions. The closed shape of the LLPS is reproduced as well as cluster formation throughout the phase diagram. The reentrant condensation is reproduced by the onset of scattering due to clustering and percolation.

Part I

Experimental Methods

Chapter 2

Small-Angle Scattering

In this thesis, both small-angle X-ray scattering (SAXS) and small-angle neutron scattering (SANS) is used in multiple studies (Ch. 4, 6, 7, 8). Here we provide the conceptual and theoretical background for these techniques. The specific experimental details are summarized in the studies, since they differ between the techniques SAXS/SANS, between the beamlines used for the experiments, and with respect to the sample preparation.

2.1 Geometry of Small-Angle Scattering Experiments

The geometry of a SAS experiment is depicted in Fig. 2.1. A sample is illuminated by a collimated beam of X-rays or neutrons, and the scattering is collected by a 2D detector. For the case of isotropic samples, such as protein solutions, the scattering is radially symmetric and depends on the scattering angle 2θ. The scattering angle is related to the scattering vector

$$q = \frac{4\pi}{\lambda} \sin\left(\frac{2\theta}{2}\right) , \tag{2.1}$$

where λ is the wavelength of the scattered beam. Importantly, q sets the length scale $\Delta r = 2\pi/q$ that is accessed by the scattering.

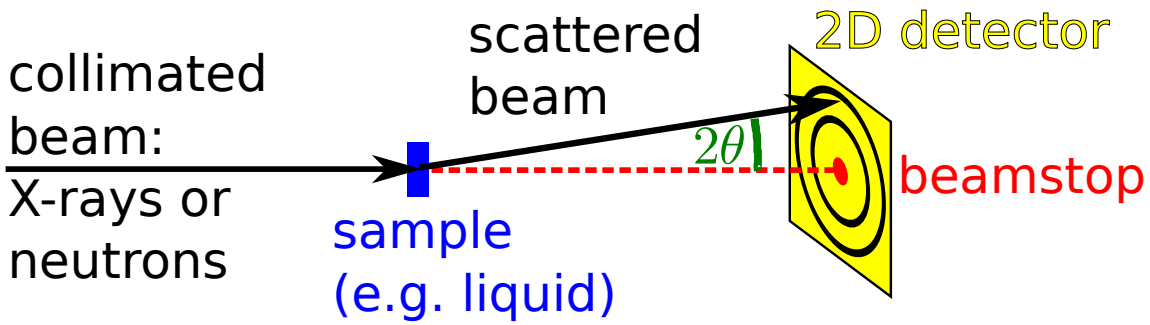

Figure 2.1: Geometry of a small-angle scattering (SAS) experiment. The scattering from a sample is collected under the scattering angle 2θ.

2.2 Theory of Small-Angle Scattering

In this section, the scattering intensity for SAS is calculated for the case of isotropic solutions of particles. We start with the general expression that is then evaluated for the special case of a monodisperse solution of spherical-symmetric particles. We discuss briefly which information can be obtained from SAS. Finally, we discuss the decoupling approximation as a way to analyze samples with polydisperse and non-spherical particles.

2.2.1 General Expression of the Small-Angle Scattering

Considering that N particles are dispersed in a solution, one can divide the full space into several compartments around the particles and thus rewrite the full scattering density $\rho(\vec{r})$ as a sum of partial scattering densities. Writing the partial scattering densities relative to the scattering centers $\vec{X}_j$, one directly obtains

$$\rho(\vec{r}) \;=\; \sum_{j=1}^{N} \rho_j(\vec{x}) \qquad \text{and} \qquad \rho_j(\vec{x}) = \rho(\vec{x} + \vec{X}_j) \quad . \tag{2.2}$$

Note that the system is not assumed to be periodic, and the individual compartments, indicated by j, do not necessarily have a similar size and shape.

While the general theory of SAXS and SANS is similar, the central theoretical difference lies in the origin of the scattering density. For X-ray scattering, the scattering density corresponds to the electron density in the sample, whereas the scattering density is more complicated for neutrons. In general, the scattering cross section of neutrons is the sum of an incoherent and a coherent part. Since the former is constant over the q range probed by SANS, incoherent scattering contributes to a flat background. Thus, the scattering density for SANS is given by the average of the coherent cross sections in the respective area (Lindner and Zemb, 2002).

The scattering intensity for the given system can be obtained directly from the scattering density $\rho(\vec{r})$ (Lindner and Zemb, 2002):

$$I(\vec{q}) \;=\; \left\langle \left| \int \mathrm{d}^3 r e^{i\vec{q}\vec{r}} \rho(\vec{r}) \right|^2 \right\rangle \;=\; \left\langle \left| \sum_j e^{i\vec{q}\vec{X}_j} \int \mathrm{d}^3 x e^{i\vec{q}\vec{x}} \rho_j(\vec{x}) \right|^2 \right\rangle \tag{2.3}$$

$$\;=\; \left\langle \sum_{k,j} e^{i\vec{q}(\vec{X}_j - \vec{X}_k)} F_j(\vec{q}) F_k^*(\vec{q}) \right\rangle \tag{2.4}$$

$$\;=\; \sum_j \left\langle F_j(\vec{q}) F_j^*(\vec{q}) \right\rangle + \left\langle \sum_{k \neq j} e^{i\vec{q}(\vec{X}_j - \vec{X}_k)} F_j(\vec{q}) F_k^*(\vec{q}) \right\rangle \;, \tag{2.5}$$

where $F_j(\vec{q}) = \int \mathrm{d}^3 x e^{i\vec{q}\vec{x}} \rho_j(\vec{x})$ is the scattering amplitude of the jth particle.

2.2.2 Scattering for a Monodisperse Solution of Spherically Symmetric Particles

In the special case of a monodisperse solution of spherically symmetric particles, the scattering amplitude $F_j(\vec{q})$ of all particles is the same for all particles and only dependent on $q = |\vec{q}|$ due to radial symmetry. Eq. (2.5) can be simplified considerably:

$$I(\vec{q}) \quad = \quad I(q) \quad = \quad N\Delta\rho^2 P(q)S(q) \quad , \tag{2.6}$$

where the contrast $\Delta\rho = \rho_p - \rho_s$ is the difference between the scattering densities of particle and solvent. The form factor $P(q)$ represents the scattering of a virtual particle corresponding to the angular averaged scattering density of all particles (Lindner and Zemb, 2002):

$$P(q) \quad = \quad \langle\, F_1(\vec{q})F_1^*(\vec{q})\,\rangle \;/\; \Delta\rho^2 \quad . \tag{2.7}$$

From the form factor, information on the size and shape of the particles can be extracted by direct Fourier inversion (Lindner and Zemb, 2002; Feigin and Svergun, 1987; Glatter and Kratky, 1982), analytical fit functions (Pedersen, 1997; Kline, 2006) as well as by more involved computational approaches for more complex shapes (Svergun, 1999).

The structure factor $S(q)$ encodes the structure of the solution, determined by the interactions between all particles (Hansen and McDonald, 2006; Lindner and Zemb, 2002):

$$S(q) \quad = \quad \frac{1}{N}\left\langle \sum_{k,j} e^{i\vec{q}(\vec{X}_j - \vec{X}_k)} \right\rangle = 1 + \frac{N}{V}\int_0^\infty (g(r)-1)\frac{\sin(qr)}{qr}\,4\pi r^2 \mathrm{d}r \quad . \tag{2.8}$$

Importantly, $S(q)$ is related to the pair correlation function $g(r)$, which is dependent on the interaction potential $u(r)$ between the particles. $S(q)$ can be fitted with model functions developed from liquid theory to provide information on the interaction (Hansen and McDonald, 2006; Pedersen, 1997; Kline, 2006).

2.2.3 Polydisperse Solutions and Non-Spherical Particles: Decoupling Approximation and Effective Structure Factor

The previous section analyzed the special case of a monodisperse solutions of spherical particles. Real solutions, however, often contain polydisperse and non-spherical particles. This section outlines the theory how to account for these effects, obtaining an equation analogous to Eq. (2.6), but with an effective structure factor.

The ensemble average $\langle\bullet\rangle$ in the general expression for the scattering (Eq. (2.5)) runs over several variables that characterize the scattering density. In particular, two kinds of variables can be distinguished. First, the absolute positions of the particles, $\vec{X}_j$, reflect the solution structure and the interactions. Ensemble averages related to the positions will be denoted as $\langle\bullet\rangle_{\mathrm{pos}}$.

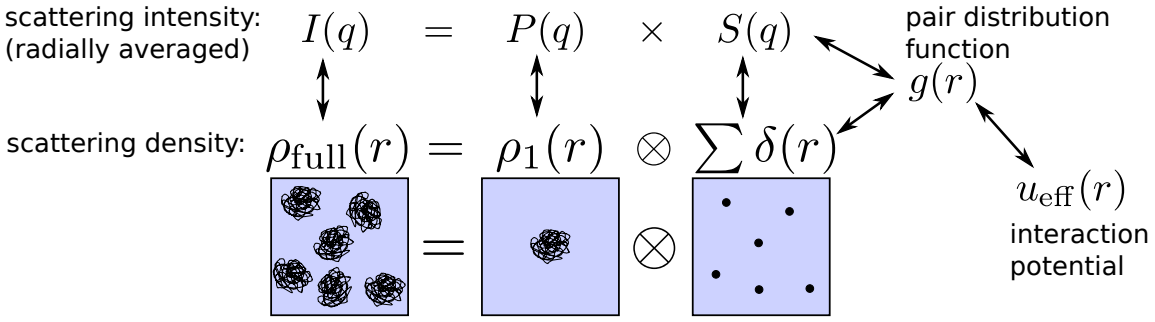

Figure 2.2: Schematic overview of information accessible by small-angle scattering from isotropic solutions. While the form factor represents the shape of the single particle, the structure factor provides access to the interaction potential via fitting methods and liquid theory.

Second, properties of the individual particles have to be characterized. For a monodisperse solution, the only relevant variable for the ensemble average is the orientation of the individual particles. For polydisperse solutions, orientation has to be complemented with variables such as size or shape. Ensemble averages related to the individual features of the jth particle will be denoted as $\langle\bullet\rangle_j$.

In general, all degrees of freedoms are statistically correlated, rendering the full ensemble average $\langle\bullet\rangle_{(\mathrm{pos},(j)_{j=1\ldots N})}$ a complicated quantity. In isotropic systems—i.e. systems without superstructure or orientational order—it is reasonable to use the so-called decoupling approximation to simplify the expression (Kotlarchyk and Chen, 1983; Hayter and Penfold, 1983; Bendedouch and Chen, 1983). The underlying assumption is that internal variables of a given particle, e.g. orientation, are independent of all positions and internal variables of the other particles. With the decoupling approximation (DA), the ensemble average reads:

$$\langle\bullet\rangle_{(\mathrm{pos},(j)_{j=1\ldots N})} \overset{(\mathrm{DA})}{\approx} \left\langle\langle\ldots\langle\langle\bullet\rangle_1\rangle_2\ldots\rangle_N\right\rangle_{\mathrm{pos}} \quad . \tag{2.9}$$

The statistical independence of the different variables also allows a factorization of the ensemble average if the function is separable with respect to position and internal variables, i.e. $\bullet = \alpha((\vec{X}_j)_{j=1\ldots N})\beta_1\beta_2\ldots\beta_N$:

$$\langle\bullet\rangle_{(\mathrm{pos},(j)_{j=1\ldots N})} \overset{(\mathrm{DA})}{\approx} \left\langle\alpha((\vec{X}_j)_{j=1\ldots N})\right\rangle_{\mathrm{pos}}\langle\beta_1\rangle_1\langle\beta_2\rangle_2\ldots\langle\beta_N\rangle_N \quad . \tag{2.10}$$

Importantly, this separation is also possible for the scattering intensity which in the decoupling approximation is given by

$$I(q) = \sum_j \left\langle F_j(\vec{q}) F_j^*(\vec{q}) \right\rangle_{(\mathrm{pos},(k)_k)} + \left\langle \sum_{k\neq j} e^{i\vec{q}(\vec{X}_j - \vec{X}_k)} F_j(\vec{q}) F_k^*(\vec{q}) \right\rangle_{(\mathrm{pos},(k)_k)} \tag{2.11}$$

$$\overset{(\mathrm{DA})}{\approx} N\left\langle F_1(\vec{q}) F_1^*(\vec{q}) \right\rangle_{j=1} + \left| \left\langle F_1(\vec{q}) \right\rangle_{j=1} \right|^2 \left\langle \sum_{k\neq j} e^{i\vec{q}(\vec{X}_j - \vec{X}_k)} \right\rangle_{\mathrm{pos}} \tag{2.12}$$

$$= N\Delta\rho^2 P(q) + N|\mu(q)|^2 (S(q) - 1) \tag{2.13}$$

$$= N\Delta\rho^2 P(q) S_{\mathrm{eff}}(q) \quad . \tag{2.14}$$

Note that $\langle \bullet \rangle_{j=1}$ is equal to all $\langle \bullet \rangle_{j\neq 1}$, because the average runs over all orientations and sizes of particles. $\mu(q) = \langle F_1(\vec{q}) \rangle_1$ can be referred to as the coherent scattering amplitude, based on the analogy of Eq. (2.13) to the full scattering cross section with incoherent and coherent part known for neutron scattering.

The decoupling approximation separates inter-particle correlations from particle properties. In order to derive an expression analogous to Eq. (2.6), the so-called effective structure factor $S_{\mathrm{eff}}(q)$ is introduced in Eq. (2.14), and accounts for both structure of the solution and effects of variations of the particle scattering density due to polydispersity and orientation:

$$S_{\mathrm{eff}}(q) = \left(1 + \frac{|\mu(q)|^2}{\Delta\rho^2 P(q)} (S(q) - 1)\right) \quad . \tag{2.15}$$

In the special case of monodisperse and spherical particles, $S_{\mathrm{eff}}(q) = S(q)$. In general, the effective structure factor allows an analysis for non-spherical and polydisperse particles that is similar to the analysis for spherical and monodisperse particles. In practice, the effective structure factor $S_{\mathrm{eff}}(q)$ is often considered as a real structure factor $S(q)$ with effective parameters. This approach is denoted the "average structure factor" approximation and is reasonable as long as the particles are rather compact and not too extended. In this thesis, we use this approach in Ch. 4, 6, 7 and 8 to extract information on protein interaction from SAS intensities.

Chapter 3

Electrophoretic Mobility and Surface Charge

Measurements of the electrophoretic mobility are used in Ch. 5 to monitor the surface charge of proteins in solutions of multivalent cations. In this chapter we provide the background for this technique and summarize the relations that are used to determine the protein surface charge via the zeta potential from the electrophoretic mobility.

Charged surfaces induce a spatial distribution of ions in the surrounding solution. In the common picture of this so-called electrical double layer, the ion distribution can be separated into the Stern layer of specifically adsorbed counterions close to the surface and the diffuse layer, divided by the outer Helmholtz plane (Hunter, 2000; Delgado et al., 2007). More relevant for electrophoretic measurements, the shear plane separates the hydrodynamically stagnant layer, consisting of an inner part of the diffusive layer and the Stern layer, from the outer part of the diffusive layer. While the latter flows around the particles, the stagnant layer follows the motion of the particle upon shear. The potential at the shear plane is called the zeta potential ζ. In general, ζ reflects the total charge within the stagnant layer, but the analytical relation is not trivial in the case of soft, rough and non-spherical objects (Delgado et al., 2007), and only an effective protein charge can be measured using the electrophoretic mobility (Menon and Zydney, 1998).

3.1 Electrophoretic mobility and Zeta Potential

The electrophoretic mobility can be measured by different methods. In this thesis, we used a Zetasizer Nano (Malvern Instruments) that performs electrophoretic light scattering (ELS) via phase analysis light scattering (PALS). In this technique, the motion of charged particles in an external alternating electrical field is probed, which allows a reliable determination of the electrophoretic mobility of biomolecules.

The electrophoretic mobility μ_E is related to the zeta potential ζ for a spherical particle with radius a by the Henry function $f(\kappa a)$ (Delgado et al., 2007):

$$\mu_E = \frac{2\varepsilon}{3\eta} f(\kappa a)\, \zeta \quad . \tag{3.1}$$

Here, ε is the total dielectric permittivity and η is the viscosity of the medium. The inverse Debye screening length κ represents the relevant length scale for the electrostatic interaction:

$$\kappa^2 = 4\pi\lambda_B \sum_i n_i Z_i^2 \quad . \tag{3.2}$$

In the sum, all present ionic species i are accounted for, with the respective valency Z_i and number concentration n_i. e denotes the elementary charge and $k_B T$ the Boltzmann constant multiplied by the temperature.

The Henry function $f(\kappa a)$ depends on the geometry of the particle and is given for spheres by (Deshiikan and Papadopoulos, 1998)

$$f(x) = \begin{cases} 1 + \frac{x^2}{16} - \frac{5x^3}{48} - \frac{x^4}{96} + \frac{x^5}{96} - \left[\frac{x^4}{8} + \frac{x^6}{96}\right] \exp(x) \int_\infty^x \frac{e^{-t}}{t} & \text{for } x < 1 \\ \frac{3}{2} - \frac{9}{2x} + \frac{75}{2x^2} - \frac{330}{x^3} & \text{for } x > 1\ . \end{cases} \tag{3.3}$$

The Henry function $f(\kappa a)$ interpolates between the Helmholtz-Smoluchowski limit ($\kappa a \gg 1$: $f = 1.5$) and the Hückel-Onsager limit ($\kappa a < 1$: $f = 1$). A more practicable approximation for $f(\kappa a)$ is provided by Ohshima's relation (Delgado et al., 2007; Ohshima, 1994):

$$f(\kappa a) = 1 + \frac{1}{2}\left[1 + \left(\frac{2.5}{\kappa a[1 + 2\exp(-\kappa a)]}\right)\right]^{-3} \quad . \tag{3.4}$$

While retardation—i.e. the opposite movement of the particle and the counterion cloud—is included in this equation, relaxation—i.e. ions have to rearrange themselves constantly due to the movement of the particle—is not included and can cause further strong changes in the range between the two limiting cases (Deshiikan and Papadopoulos, 1998).

3.2 Zeta Potential and Surface Charge Density

In first approximation, the zeta potential corresponds to the potential at the shear plane (Kirby and Hasselbrink, 2004). The zeta potential is affected by the initial surface charge Q, charges of adsorbed ions in the Stern layer, and charges within the hydrodynamically stagnant part of the diffuse layer.

The Poisson-Boltzmann equation

$$\nabla^2 \Psi = -\frac{1}{\epsilon} \sum_i n_{0,i} e z_i \exp\left(-\frac{e z_i}{k_B T} \Psi\right) \tag{3.5}$$

describes the resulting potential in the solution on a mean-field level, i.e. ignoring ion–ion correlations, and can be solved in special cases or under special approximations. Doing so, expressions can be derived that relate the zeta potential to the surface charge. The final approximate expression in Eq. (3.14) is used in Ch. 5 to determine the total charge of the complex of proteins with bound ions.

3.2.1 Infinite Flat Surface

As a first step, it is instructive to consider a symmetric binary electrolyte ($z^+ = -z^- = z$) and an idealized and infinite flat surface. Under these conditions, the nonlinear Poisson-Boltzmann equation reads

$$\nabla^2\Psi = -\frac{2n_0ez}{\epsilon}\sinh\left(-\frac{ez}{k_BT}\Psi\right) \tag{3.6}$$

and can be solved analytically (Kirby and Hasselbrink, 2004):

$$\Psi^*(y) = 2\ln\left[\frac{1+\tanh(\zeta^*/4)\exp(-\kappa y)}{1-\tanh(\zeta^*/4)\exp(-\kappa y)}\right] \tag{3.7}$$

with the distance y perpendicular to the surface, and the reduced potentials $\Psi^* = \Psi ze/k_BT$ and $\zeta^* = \zeta ze/k_BT$.

3.2.2 Spherical Particle

For other geometries than a flat surface, there is in general no exact solution. However, the case of spherical particles reduces to the planar solution in the two limiting cases $\kappa a \ll 1$ and $\kappa a \gg 1$ (Ohshima et al., 1982). While in the first limit the particle can be considered as point-like compared to the double-layer—thus allowing linearization—, the second limit corresponds to a very thin double layer relative the the particle size—thus allowing for neglecting the surface curvature.

For the general case, the Poisson-Boltzmann equation in spherical coordinates reads (Ohshima et al., 1982)

$$\frac{\mathrm{d}^2\Psi}{\mathrm{d}r^2} + \frac{2}{r}\frac{\mathrm{d}\Psi}{\mathrm{d}r} = \frac{2}{\epsilon_0\epsilon_r}\sinh\left(\frac{ze\Psi}{k_BT}\right) \tag{3.8}$$

with the boundary conditions

$$\Psi(a) = \Psi_0 \quad , \qquad \Psi(\infty) = 0 \quad , \qquad \left.\frac{\mathrm{d}\Psi}{\mathrm{d}r}\right|_a = -\frac{\sigma}{\epsilon_0\epsilon_r}\,. \tag{3.9}$$

Debye-Hückel approximation

An exact solution for the surface potential is available with the Debye-Hückel approximation, i.e. the linearization of the Poisson-Boltzmann equation:

$$ze\Psi \ll k_B T \qquad \Rightarrow \qquad \sinh\left(-\frac{ez_i}{k_B T}\Psi\right) \approx -\frac{ez_i}{k_B T}\Psi \quad , \tag{3.10}$$

which yields a simple exponential decay for the potential on a flat plane with symmetrical electrolyte (White, 1977):

$$\Psi_{\mathrm{lin}} = \Psi_0 \kappa a \exp(\kappa a)\frac{\exp(-\kappa y)}{\kappa r} \quad . \tag{3.11}$$

Using this linearization the surface charge density σ returns (White, 1977)

$$\sigma_{\mathrm{lin}} = \epsilon_0 \epsilon_r \kappa \left(\frac{1+\kappa a}{\kappa a}\right)\Psi_0 \quad . \tag{3.12}$$

Approximate Expressions

As an improvement beyond the linearization, different approximate analytic expressions have been derived for the relation of surface charge and surface potential. For comparison and validation, the electrophoretic mobility and its relation to the zetapotential have been calculated numerically (Wiersema et al., 1966).

Ohshima et al. (1982) derived an accurate analytic expression for symmetric electrolytes. By replacing the non-planar part of the spherical equation by the values from the limit $\kappa a \gg 1$ one obtains

$$\sigma = \frac{\epsilon_0 \epsilon_r \kappa k_B T}{ze}\left(2\sinh(\zeta^*/2) + \frac{4}{\kappa a}\tanh(\zeta^*/4)\right) \quad , \tag{3.13}$$

where $\zeta^* = ze\Psi_0/(k_B T)$. For further improvement, not the limit value, but the limiting form has been used, yielding (Ohshima et al., 1982)

$$\sigma = \frac{\epsilon_0 \epsilon_r \kappa k_B T}{ze} 2\sinh(\zeta^*/2)\left[1 + \frac{2}{\kappa a \cosh^2(\zeta^*/4)} + \frac{8\ln(\cosh(\zeta^*/4))}{(\kappa a)^2 \sinh^2(\zeta^*/2)}\right]^{1/2} \quad . \tag{3.14}$$

Compared to the values computed by Wiersema et al. (1966) this formula is accurate with errors smaller than 5 % for $\kappa a > 0.1$ (Ohshima et al., 1982).

In this thesis in Ch. 5, Eq. (3.14) is used in combination with Eq. (3.1) and (3.4) to calculate the surface charge density from the measured electrophoretic mobilities.

Part II

Results and Discussion

Chapter 4

Hydration and Interactions in Protein Solutions Containing Concentrated Electrolytes Studied by Small-Angle Scattering

Chapter 4 is based on the following publication:

Hydration and interactions in protein solutions containing concentrated electrolytes studied by small-angle scattering
Fajun Zhang, Felix Roosen-Runge, Maximilian W. A. Skoda, Robert M. J. Jacobs, Marcell Wolf, Philip Callow, Henrich Frielinghaus, Vitaliy Pipich, Sylvain Prevost and Frank Schreiber

Phys. Chem. Chem. Phys., 14 (2012) 2483–2493

Contributions:

Research design	FZ, FS
Experiments	FZ, FRR, MWAS, RMJJ, MW
Technical Assistance	PC, HF, VP, SP
Data analysis and interpretation	FZ, FRR
Paper writing	FZ, FRR, FS

(Zhang et al., 2012b)

During protein crystallization and purification, proteins are commonly found in concentrated salt solutions. The exact interplay of the hydration shell, the salt ions, and protein–protein interactions under these conditions is far from being understood on a fundamental level, despite the obvious practical relevance. We have studied a model globular protein (bovine serum albumin, BSA) in concentrated salt solutions by small-angle neutron scattering (SANS). The data are also compared to previous studies using

SAXS. The SANS results for dilute protein solutions give an averaged volume of BSA of 91 700 Å^3, which is about 37 % smaller than that determined by SAXS. The difference in volume corresponds to the contribution of a hydration shell with a hydration level of 0.30 g/g protein. The forward intensity $I(0)$ determined from Guinier analysis is used to determine the second virial coefficient, A_2, which describes the overall protein interactions in solution. It is found that A_2 follows the reverse order of the Hofmeister series, i.e. $(NH_4)_2SO_4 < Na_2SO_4 <$ NaOAc $<$ NaCl $<$ $NaNO_3 <$ NaSCN. The dimensionless second virial coefficient B_2, corrected for the particle volume and molecular weight, has been calculated using different approaches, and shows that B_2 with corrections for hydration and the non-spherical shape of the protein describes the interactions better than those determined from the bare protein. SANS data are further analyzed in the full q-range using liquid theoretical approaches, which gives results consistent with the A_2 analysis and the experimental structure factor.

4.1 Introduction

Protein interactions and phase behavior in solutions containing concentrated electrolytes are crucial for understanding the mechanism of protein crystallization or the specificity (Curtis et al., 2001, 1998, 2002; Durbin and Feher, 1996; Piazza, 2000, 2004). For example, salt induced precipitation has been extensively used as an initial step for protein purification (Curtis et al., 1998; Collins, 2004; Collins and Washabaugh, 1985; Collins, 1997). However, protein solubility is not well understood on a molecular level, and selecting the optimum conditions to precipitate a target protein is difficult because the solubility is governed by many factors including pH, surface hydrophobicity, surface charge distribution, salt type and concentration. Proteins are commonly found in concentrated salt solutions during protein crystallization which is essential for the most efficient way of determining the protein structure, namely X-ray crystallography (Durbin and Feher, 1996). The optimization of conditions for preparation of protein single crystals is still largely a trial and error process. Theories that can reliably predict protein solubility and crystallization conditions in a complex solution are currently not available, but understanding the factors that affect the protein interactions as well as the phase behavior of protein solution is the only way towards developing a theoretical framework that can be used to optimize or predict the desired conditions.

Small angle X-ray and neutron scattering (SAXS and SANS) as low resolution diffraction methods have been widely used for structural determination as well as understanding protein interactions in solutions (Jacrot, 1976; Jacrot and Zaccai, 1981; Lipfert and Doniach, 2007; Perkins, 2001; Perkins et al., 2008; Petoukhov and Svergun, 2006; Svergun et al., 1998; Stradner et al., 2006, 2004). Using SANS combined with SAXS, Sinibaldi et al. studied the solvation properties of BSA and lysozyme in urea solution and water/glycerol mixtures, respectively (Sinibaldi et al., 2007, 2008). SANS and SAXS provide complementary

information due to the different responses to the hydration shell surrounding proteins. The interpretation of SANS data requires knowledge of the hydration level and the H–D exchange ratio of proteins. Zaccai and Jacrot discussed the volume determined by SANS measurement from the forward intensity, $I(0)$ (Zaccai and Jacrot, 1983). They pointed out that the volume term derived from $I(0)$ includes not only the "dry" volume of the protein, but also the volume change of the water molecules from the bulk to the hydrated state. This perturbation mainly depends on the density of the hydrated water compared to bulk water. Svergun et al. studied the density of hydration shell using SAXS and SANS (Svergun et al., 1998). They found that the density of the hydration shell is about 10 % higher than that of bulk water. This result was supported by molecular dynamics simulations (Merzel and Smith, 2002). Given the negative contrast of the high density hydration shell for SANS measurements in both H_2O and D_2O, this will give a reduced size of the target protein. Indeed, the radius of gyration (R_g) determined by SAXS was found to be larger than that calculated from the atomic structure, while R_g values determined by SANS in H_2O and D_2O are similar to and slightly smaller compared to the calculated value. From these results, it seems that SANS is not sensitive to the hydration shell. Perkins discussed the effect of hydration shell in both SAXS and SANS measurements (Perkins, 2001): For SAXS, the electron density difference between the hydrated water (0.408 e $Å^{-3}$) and bulk water (0.334 e $Å^{-3}$) is similar to that of proteins (0.425 e $Å^{-3}$), and it therefore measures the hydrated protein molecules as a complete object. For SANS, proteins in H_2O have a SLD of $0.19 \cdot 10^{-6} Å^{-2}$, +134 % compared to H_2O ($-0.562 \cdot 10^{-6} Å^{-2}$), while the hydration shell ($-0.69 \cdot 10^{-6} Å^{-2}$) has a contrast of about –23 %, which is similar to the 31 %variation in SLD within the common 20 amino acids with proteins, the hydration shell is therefore not detectable in a SANS experiment with H_2O buffer. Proteins in D_2O have a SLD of $3.2 \cdot 10^{-6} Å^{-2}$, –50 % compared to D_2O ($6.35 \cdot 10^{-6} Å^{-2}$). The hydration shell (SLD of $7.82 \cdot 10^{-6} Å^{-2}$) has a contrast of about +22 %. The contribution from the hydration shell is still weak, but may not be negligible. As shown by Svergun et al. (1998), R_g values determined by SANS in D_2O are slightly smaller than the calculated value (within systematic error bar). It was concluded that to a good approximation, the hydration shell is not detectable by SANS in either H_2O or D_2O (Perkins, 2001). These studies indicate that there is a discrepancy between the theoretical considerations and SANS experiments and more detailed experimental studies are required to clarify this issue; and the comparison between SAXS and SANS is necessary for a comprehensive picture of the hydration shell surrounding proteins. It has long been known that the solubility of proteins in solution depends not only on the salt concentration, but also on the type of the salt. This important phenomenological concept is known as "Hofmeister effect" (Hofmeister, 1888). The classification of ions was initially performed on their ability to precipitate a protein, and could be extended to many other systems (Collins and Washabaugh, 1985). The order is related to many key parameters, in particular the propensity of ions to affect water structure (Collins and Washabaugh, 1985). It has been shown that anions have a stronger effect. However, the combination of cation and anion

in a salt does not result in a trivial sum of ion effects. A representative Hofmeister anion series can be given as follows:

$$\begin{aligned}\mathrm{Citrate}^{3-} > \mathrm{SO}_4^{2-} > \mathrm{PO_4H}^{2-} > \mathrm{F}^- > \mathrm{CH_3COO}^- > \ldots \\ \cdots > \mathrm{Cl}^- > \mathrm{Br}^- > \mathrm{I}^- > \mathrm{NO}_3^- > \mathrm{ClO}_4^- > \mathrm{SCN}^- \quad .\end{aligned}$$

Ions on the left hand side of the series will precipitate ("salt out") solutes, whereas ions on the right will dissolve or denature ("salt in") solutes. Hofmeister effects have been explained in terms of structure-making ("kosmotrope") and structure-breaking ("chaotrope") abilities of these ions with water (Collins, 2004, 1997). Kosmotropes interact with water strongly and water molecules surrounding the salt ions are more structured relative to bulk water. Chaotropes break up the structure of the surrounding water molecules because of the large size of the ion and its weak interaction with water. For example, the specific ion effects on protein solubility were described in terms of the ability of ions to "salt in" the polar peptide group and "salt out" the nonpolar side chains (Baldwin, 1996). Hofmeister effects on the solubility of proteins have also been related to the protein–water interfacial tension (Dér et al., 2007). Although the molecular basis of Hofmeister effects, i.e. the specific ion effects, is not yet fully understood, significant progress has been made by inclusion of dispersion forces and hydration. Simplified models of electrolytes with proper inclusion of dispersion forces can capture the essentials of Hofmeister effects qualitatively (Ninham and Nostro, 2010; Levin, 2009; Schwierz et al., 2010).

Serum albumin is the most abundant protein in mammalian blood with a concentration around 30 g/l. Bovine and human serum albumin (BSA and HSA) as model globular proteins have been involved in many studies due to their applications as a carrier protein and as a stabilizing agent in enzymatic reactions. BSA in solution has been studied by SAXS and SANS with various protein concentration, ionic strength, and pH (Barbosa et al., 2010; Nossal et al., 1986; Kotlarchyk and Chen, 1983; Bendedouch and Chen, 1983; Bendedouch et al., 1983). BSA is a good model for protein interactions in concentrated salt solutions because of its stability in a wide salt concentration range. In this context, we have studied the protein–protein interactions in aqueous solution using SAXS as a function of protein and ionic strength (Roosen-Runge et al., 2010, 2011; Zhang et al., 2007, 2008). The protein–protein interactions strongly depend on the nature of the added salts. When a trivalent salt was added into protein solution, a novel phase behavior, named "reentrant condensation", was observed recently in our laboratory (Zhang et al., 2008; Ianeselli et al., 2010; Zhang et al., 2010).

In this paper, we present a systematic study of protein (BSA) in concentrated salt solutions with different types of salt by using SANS, which provides high contrast and no radiation damage. By comparison with our previous SAXS study (Zhang et al., 2007), the contribution from the hydration shell will be specifically addressed. By using several typical salting-in and salting-out anions selected from the Hofmeister series, we address the following questions: first, how sensitive is SANS in this case to the hydration shell? By

calculating the protein volume and comparing with that determined from SAXS and protein crystallography, we determine the contribution of the hydration shell. Second, how to determine the interactions in such concentrated salt solutions? The second virial coefficient is evaluated and discussed along the predictions for the Hofmeister series. Furthermore, modeling based on liquid state theory will be used to address these questions.

4.2 Experimental Section

4.2.1 Materials

Bovine serum albumin (BSA) (99 % purity) was purchased from Sigma-Aldrich. It is a lyophilized powder with a molar molecular weight of ≈ 66 kDa and was used as received. The following salts were used as received: sodium acetate (NaOAc), $NaNO_3$ (Sigma-Aldrich), Na_2SO_4, NaCl, $(NH_4)_2SO_4$ (Merck), NaSCN (Alfa Aesar). BSA solutions were prepared by diluting a stock solution of 180 ± 1 mg /mL determined by UV absorption at 280 nm with an extinction coefficient of 39 020 M^{-1} cm^{-1}, or 0.5912/(mg/mL) cm^{-1}, calculated from the amino acid sequence (Bendedouch and Chen, 1983). In order to avoid the effect of other ions, no buffer was used. The pH of protein solutions was 6.8 ± 0.1 for most salts used except $(NH_4)_2SO_4$ and NaOAc with pH values of about 6.2 ± 0.1 and 8.0 ± 0.1, respectively. All the SANS measurements were performed at 20 °C using D_2O (99.9 %, Aldrich or ILL) as solvent.

Caution: It should be noted that concentrated sodium salt solutions can be radioactive after SANS measurements.

4.2.2 Small-Angle Neutron Scattering

Small-angle neutron scattering (SANS) measurements were performed at three different stations, D22 (ILL, Grenoble, France), KWS2 (JCNS, FRM2, Munich, Germany) and V4 (Helmholtz Center, Berlin, Germany).

D22 diffractometer at the high flux reactor of the Institut Laue-Langevin (ILL) in Grenoble, France: Two configurations were used with sample-to-detector distances of 2 m and 8 m and collimation length of 2 and 8 m, respectively, in order to cover the q-range from 0.007 to 0.35 $Å^{-1}$ at a wavelength l of 6.0 Å ($\Delta\lambda/\lambda = 10\,\%$). Protein solutions in D_2O were filled in quartz cells with a path-length of 1 or 2 mm. Pure H_2O was used for the absolute intensity calibration. Data correction was performed using GRASP (Dewhurst, 2012).

SANS instrument V4 at the Helmholtz Center, Berlin, Germany (Keiderling and Wiedenmann, 1995): Three configurations were used with sampleto- detector (SD) distances of 1 m, 4 m and 12 m and collimation lengths of 8 m, 4 m, and 12 m, respectively, in order to cover the q-range from 0.005 to 0.40 $Å^{-1}$ at a wavelength of 6 Å ($\Delta\lambda/\lambda = 10\,\%$). The data were recorded on a 64 × 64 cm^2 two-dimensional detector and radially averaged. The

data reduction was performed using the software BerSANS (Keiderling, 2002). Protein solutions in D_2O were filled in quartz cells with a path-length of 5 or 10 mm. Pure H_2O was used for the absolute intensity calibration.

KWS2 instrument located at the FRM2, Munich, Germany: Two configurations were used with SD distances of 2 and 8 m and a collimation length of 8 m in order to cover the q-range from 0.005 to 0.35 Å^{-1} at a wavelength λ of 4.5 Å ($\Delta\lambda/\lambda = 20\,\%$). A two-dimensional array detector was used to detect neutrons scattered from sample solutions. Protein solutions were filled in rectangular quartz cells with path-length of 1 or 2 mm. Plexiglas was used as secondary standard to calibrate the absolute scattering intensity. The data correction and absolute intensity calibration were obtained using the software QtiKWS (Pipich, 2012).

It is worthy to note that the absolute scale of SANS is known to be difficult to obtain with a high precision, and different set-ups and methods from different facilities might lead to slightly different absolute intensity values (Grillo, 2008; Lindner, 2000).

4.2.3 Data Analysis

SANS intensity from protein solutions can be described as:

$$I(q) = N_p(\Delta\rho)^2 V_P^2 P(q) S(q) + B \quad , \tag{4.1}$$

where $q = 4\pi/l \cdot \sin\theta$, 2θ denotes the scattering angle, N_P is the number of protein molecules per unit volume in the solution, V_P is the volume of a single protein and $\Delta\rho = (\rho_P - \rho_S)$, is the difference of scattering length density between protein and solvent. $P(q)$ is the form factor of a given protein, i.e., the scattering from a single protein molecule after orientation averaging. $S(q)$ is the structure factor describing the time-averaged interaction and distribution of proteins in solution. B is the protein concentration dependent incoherent background.

Determination of the Radius of Gyration, R_g. In a dilute protein solution, where protein molecules are well dispersed, the interaction between them is negligible. In this case, the scattering intensity is the summation of the scattering intensities of all the proteins within the illuminated volume. At sufficiently low q ($qR_g < 1$), the scattering intensity can be approximated by the Guinier law (Glatter and Kratky, 1982; Guinier and Fournet, 1955):

$$\ln[I(q) - B] = \ln I(0) - \frac{1}{3} R_g^2 q^2 \quad , \tag{4.2}$$

where $I(0)$ is the forward scattering at zero angle, R_g the radius of gyration. Eqn (4.2) provides a direct method to determine the R_g. In practice, eqn (4.2) is valid in the range of qR_g up to 1.5 (Perkins et al., 2008).

Determination of the Second Virial Coefficient, A_2. The forward intensity determined from the Guinier analysis can be used to determine A_2. Tardieu and co-workers have demonstrated that under weak interactions and low protein concentrations, the second virial coefficient can be determined from SAXS measurements (Bonnet et al., 1999; Bonneté and Vivarès, 2002; Tardieu et al., 1999; Finet et al., 2004). Eqn (4.1) can be re-written as:

$$I(q) = c\alpha P(q)S(q) + B \quad , \tag{4.3}$$

where c is the protein concentration in unit of mg/ml, and a a pre-factor. The structure factor at $q = 0$ is related to the osmotic pressure, Π, by:

$$S(0) = \frac{RT}{M_W}\left(\frac{\partial \Pi}{\partial c}\right)^{-1} \quad , \tag{4.4}$$

where R is the gas constant, and Π can be expressed via a virial expansion:

$$\frac{\Pi}{cRT} = \frac{1}{M_W} + A_2 c + A_3 c^2 \ldots \tag{4.5}$$

For low protein concentration and weak interactions, A_3 and higher order terms can be neglected, which then yields with eqn (4.3):

$$\frac{1}{S(0)} = 1 + (2M_W A_2)\, c = \frac{c}{I(0) - B}\alpha P(0) \quad . \tag{4.6}$$

Model Fitting using Liquid State Methods. The scattering intensity, $I(q)$, for a polydisperse or a non-spherical system, is calculated based on approximation approaches. Most often used are the "decoupling approximation" and "average structure factor" approximation (Chen and Lin, 1987; Pedersen, 1997; Hayter and Penfold, 1983; Chen, 1986). Both approaches assume that the particle orientation is independent of the particle position, and the form factor results consistently in both approaches from the angular average of a single particle.

In the decoupling approximation, an explicit angular average for the assumed particle shape is performed and returns a scaling factor $\beta(q)$ for the structure factor (Chen and Lin, 1987; Pedersen, 1997; Hayter and Penfold, 1983). The average structure factor approximation accounts implicitly for the angular average by introducing an effective monodispersed particle (Chen, 1986). While this thus technically neglects polydispersity and non-sphericity, the decoupling approximation can induce more delicate errors, since it is sensitive to the actual choice of the particle shape, which is not unique if based on the experimental form factor. For the case of moderately non-spherical but monodispersed particles, such as globular proteins in solution, both assumptions were proven to give comparable results. Therefore, in this work, the scattering intensity is calculated using the average structure factor approximation (Glatter and Kratky, 1982; Chen and Lin, 1987; Chen, 1986; Lindner and Zemb, 2002).

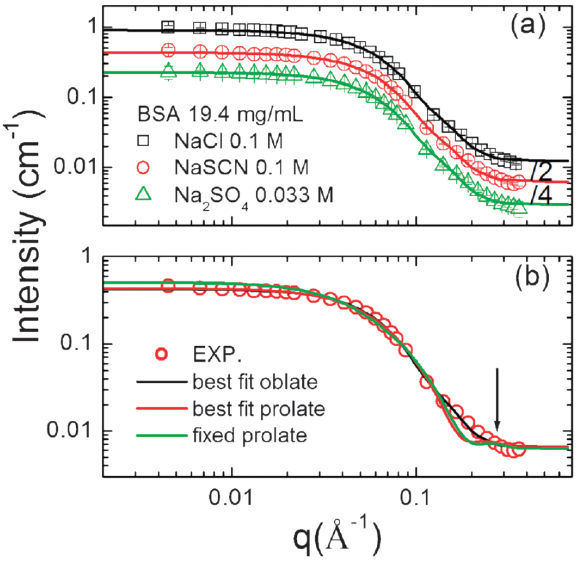

Figure 4.1: Determination of the form factor from SANS profiles for BSA solutions at low ionic strength. Data were taken at station V4, Helmholtz Zentrum Berlin. Only 20 % of data were plotted for clarifying (a) best fit using an oblate form factor, data are shifted for clarity, and (b) comparison of best fits from oblate, prolate and fixed prolate ($70 \times 20 \times 20\text{Å}^3$) as form factors.

In our case, the protein molecules have an overall ellipsoidal geometry. An effective sphere radius is calculated by equating the second virial coefficient, A_2, of the ellipsoid to a sphere having the same A_2. In the later discussion, this effective radius is called "Isihara radius", and will be used to calculate $S(q)$ (Zhang et al., 2007, 2008; Kline, 2006; Isihara, 1950). BSA is negatively charged at physiological pH. At low ionic strength (< 100 mM), the charge induced interaction can be described using a screened Coulombic potential (Hayter and Penfold, 1981; Hansen and Hayter, 1982). At moderate and high ionic strength, the surface charge of proteins is sufficiently screened. The overall interaction is rather weak, and the protein molecules interact with each other mainly through hard sphere (excluded volume effect) interactions. Detailed description can be found in our previous publications (Zhang et al., 2007, 2008; Ianeselli et al., 2010). The data analysis was carried out using macros developed by the NIST center for neutron scattering research (Kline, 2006).

4.3 Results and Discussion

4.3.1 Form Factor and Hydration Shell

We first present the results on the shape and size of proteins in dilute solutions (below 20 mg/ml) with an ionic strength sufficient to screen the surface charge. Since the global shape of a particle is mainly reflected in the intermediate q-range where the scattering intensity decays quickly as a function of q, a model with a simplified particle shape (sphere

or ellipsoid etc.) has the advantage of using a larger q-range compared to the Guinier analysis. It thus provides more precise information on the size and shape of the particle. The information on the form factor is very sensitive to the fit in the q-range of 0.01 to 0.2 Å^{-1}. Data points at larger q have a lower statistical certainty and are affected by the incoherent background correction, while the data points at lower q are sensitive to the effects of beam stop and aggregation.

Data fitting using an ellipsoid form factor was performed on SANS data from all instruments. Fig. 4.1 shows the SANS data collected on V4. The SANS curves are almost identical for BSA 19.4 mg/ml with three different salts at an ionic strength of 100 mM (Fig. 4.1a). The best fits using an oblate ellipsoid gives the dimension $R_a = 12.5 \pm 0.2$ Å and $R_b = 41.9 \pm 1.1$ Å, with a constant volume fraction of (1.1 ± 0.1) %, $\Delta\rho = -(3.3 \pm 0.2) \cdot 10^{-6}$ Å^{-2} and incoherent background 0.01 cm^{-1}. In the literature (Nossal et al., 1986; Bendedouch and Chen, 1983), BSA has been proposed to possess a cigar-like shape and modeled using a prolate ellipsoid form factor with $R_a = 70$ Å and $R_b = 20$ Å. Due to the limited q-range in Nossal et al. (1986) and Bendedouch and Chen (1983) ($< 0.16\text{Å}^{-1}$), the model is not sensitive to the difference between prolate and oblate ellipsoid form factor. As seen in Fig. 4.1b, the differences for these two types of form factor are clearly seen at $q > 0.14\text{Å}^{-1}$. First of all, the prolate form factor decays much faster than that of the oblate. Second, a second scattering maximum would have to be visible at $q = 0.28$ Å^{-1} (indicated by an arrow for the red and green curves in Fig. 4.1b), which is not observable from the experimental data and the oblate form factor. The best fit using prolate form factor gives the dimension of $R_a = 53.9$ Å and $R_b = 21.6$ Å in our case. Based on these, we conclude that an oblate ellipsoid form factor is more suitable in our data analysis. Note that we do not intend to describe the details of the shape of protein since a high resolution crystal structure for BSA is not available, but estimate only qualitatively the overall protein geometry. Despite the simplicity, the form factor should give a very good estimate of the volume to be compared to the hydrated volume or dry volume of proteins in solution, in particular since the focus is on relative changes of the volume instead of the aspect ratio of the axis.

Fig. 4.2a shows SANS data from KWS2 for BSA 9.5 mg/ml with NaCl from low to high ionic strength. The three curves almost overlap indicating the negligible effects of ionic strength on the globular structure of proteins and negligible interactions for these sample solutions. Data from KWS2 give $R_a = 13.5 \pm 0.6$ Å and $R_b = 41.0 \pm 0.9$ Å. Data from D22 give average dimensions of $R_a = 12.3 \pm 0.8$ Å and $R_b = 42.2 \pm 0.7$ Å. Fig. 4.2b presents the Kratky plots ($I(q)q^2$ versus q) of SANS data for samples with 9.5 mg/ml BSA and various added salts. The Kratky plot is useful to clarify the folding state of proteins under various conditions (Doniach, 2001). For compact proteins in solution with their native folding state, as described by the Porod law, the scattering intensity decays as $I(q) \propto q^{-4}$ (Glatter and Kratky, 1982), which in the Kratky plot gives a pronounced peak with a "bell" shape. On the other hand, for completely unfolded proteins, the scattering intensity of biopolymers in a random coil state follows the Debye law, i.e. $I(q) \propto q^{-1}$, which in the Kratky plot shows

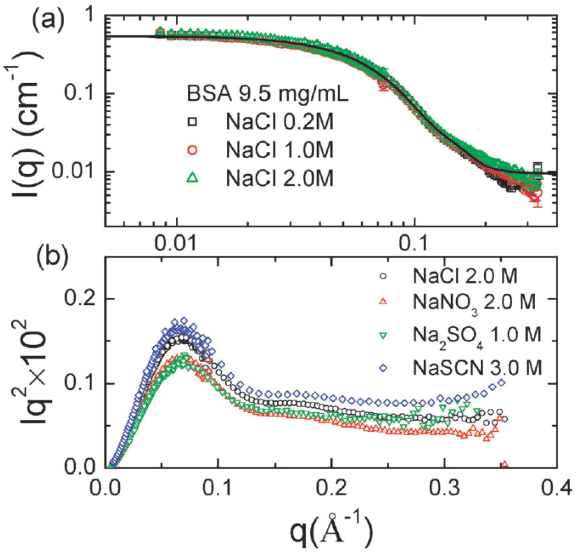

Figure 4.2: (a) SANS profiles for BSA solutions with different NaCl concentration and model fitting using ellipsoid as form factor (data from KWS2). Only 30 % of data points were plotted. (b) Kratky plots of typical SANS data. A well developed bell shape indicates the globular shape of protein without unfolding or denaturation under the concentrated electrolyte solutions

a steady increase at high q values (Doniach, 2001). The resulting plots shown in Fig. 4.2b indicate that the addition of salts from both sides of the Hofmeister series does not reduce the compactness of protein significantly, consistent with other studies (Curtis et al., 2002; Doniach, 2001; Baglioni et al., 2004; Dumetz et al., 2007).

SANS results collected from different instruments (V4, KWS2, D22) give an averaged dimension of BSA molecule of $(12.5 \pm 0.8) \times (41.9 \pm 0.9) \times (41.9 \pm 0.9)$ Å^3. Using the maximum error of the axes, we estimate the upper limit for the error of the volume as $\approx 6\,\%$, which corresponds to a volume of 91700 ± 5500 Å^3. This value is consistent to that reported in the literature (Sinibaldi et al., 2008), and close to the volume of HSA determined from the crystal structure (85 280 Å^3) using the PDB file of 1N5U and calculated by CRYSOL (Svergun et al., 1995). The dry volume of BSA can also be calculated from the specific volume (0.74 ± 0.02 cm^3/g) (Mylonas and Svergun, 2007) divided by N_A and M_W (66 100 Da), which gives a value of 81 200 Å^3. SANS measurements thus give a volume about 7.5 % larger than the values calculated from crystal structure of HSA and 13 % larger than that from the specific volume, respectively. This slight enlargement of the resulting volume may be partly due to the existence of dimer in BSA solutions. Hunter and Carta analyzed the fraction of dimer and other oligomers in commercial BSA products and found an overall fraction about 6.4 % for the product used in this work.66 Dimerization thus cannot be the major contribution to the observed volume increase. In contrast, this dimension is significantly smaller than that determined by SAXS as $17 \times 42 \times 42$ Å^3 with a volume of 125 300 Å^3. Assuming the density increase of the hydrated water is $\approx 10\,\%$ (Svergun et al., 1998), the hydration level can be calculated by $0.74 \cdot 36.6\,\% \cdot 1.1 \approx 0.30$ g/g protein

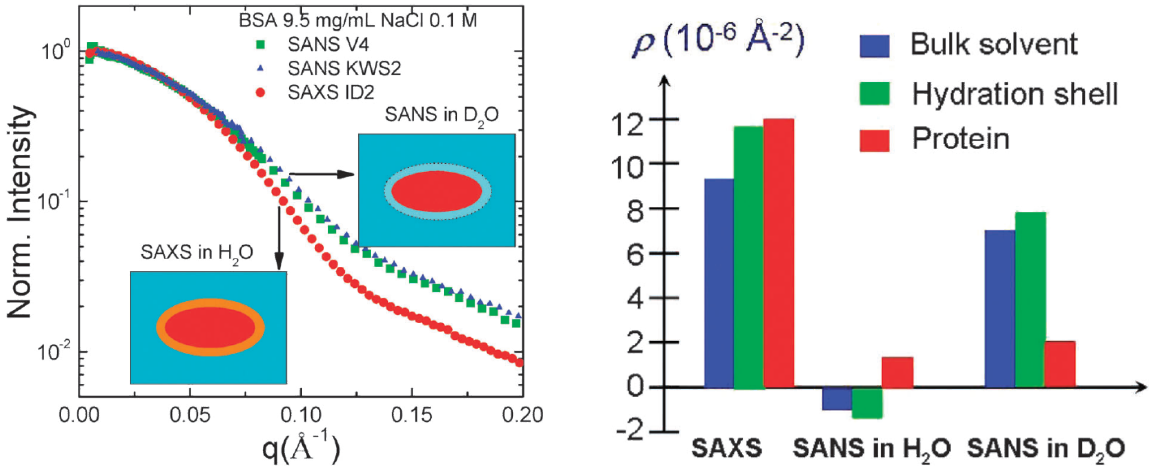

Figure 4.3: (a) Comparison of SANS and SAXS data reveals the effect of the hydration shell. Data are normalized to the forward scattering $I(0) = 1$. SAXS data are collected at ESRF, Grenoble. Details can be found in Roosen-Runge et al. (2011). (b) A diagram of SLD explains the contribution of hydration in SANS and SAXS measurements (Svergun et al., 1998).

for BSA. This hydration level is in good agreement with the values (from 0.3 to 0.4 g/g BSA) reported in the literature (Perkins, 2001; Nossal et al., 1986; Bendedouch and Chen, 1983; Kuntz and Kauzmann, 1974). This hydration effect is clearly seen by comparing the SAXS and SANS profiles on BSA in H_2O and D_2O (Fig. 4.3). A diagram of SLD (Fig. 4.3b) describes the difference of hydration in SANS and SAXS measurement, which is consistent with literature reports (Perkins, 2001; Svergun et al., 1998). It is clearly visible that the SANS data shift towards higher q values compared to the SAXS data at $q > 0.05\,\text{Å}^{-1}$, indicating a shrinkage of the apparent volume of proteins measured by SANS in D_2O. It is worth noting that the shift of SANS data cannot be due to the smearing effect. As discussed in the literature, the major smearing effect of SANS measurement is due to the wavelength spread, $\Delta\lambda/\lambda$ (Grillo, 2008; Pedersen, 1997). The SANS data from V4 and KWS2 with $\Delta\lambda/\lambda$=10 % and 20 %, respectively, show very small differences compared to the shift with respect to the SAXS data. Similar behavior has been observed by Svergun et al. (1998).

4.3.2 Protein–Protein Interaction from the Second Virial Coefficient

Fig. 4.4a presents typical SANS profiles of BSA with 2.0 M $NaNO_3$. A Guinier analysis is used to determine the forward intensity $I(0)$ and radius of gyration, R_g, in the q range from 0.005 to 0.045Å^{-1}. Typical Guinier plots (Fig. 4.4b) show a linear relationship for samples with protein concentration up to 89.4 mg/ml. Note that the scattering intensity near the beam stop has a higher uncertainty, and the low q upturn (e.g. 4.5 mg/ml) does not happen systematically. Therefore, it is impossible to judge the overall interaction by just looking at the scattering profile in the low q region. Guinier analysis from all samples with low protein concentration (< 20 mg/ml) gives similar values of $R_g = 28.7 \pm 1.5$ Å, which is

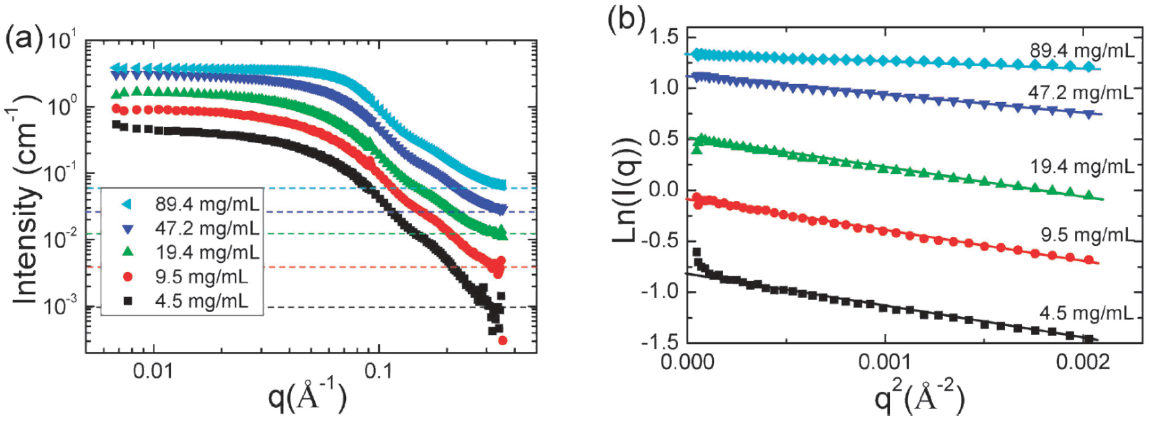

Figure 4.4: (a) Representative SANS data (from D22, ILL) for BSA solutions with $NaNO_3$ 2.0 M. The dashed lines correspond to the estimated incoherent background (see text). (b) Guinier plots of SANS data in (a), a good linear relationship was obtained for all samples. R_g and $I(0)$ are obtained from the linear fitting in the q^2 range of $5 \cdot 10^{-5}$ to $1.5 \cdot 10^{-3}$ Å^{-2}.

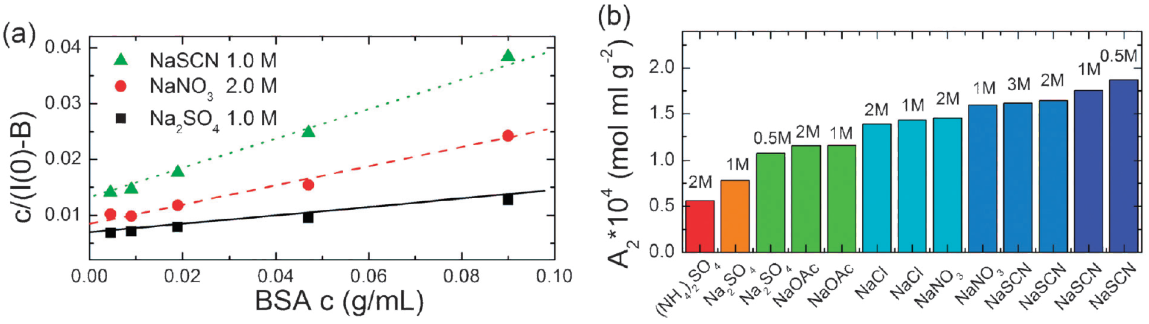

Figure 4.5: (a) Plots of $c/(I(0)-B)$ as a function of protein concentration. The intercept and the slope of linear fitting are used to determine the second virial coefficient, A_2, according to eqn (4.6). (b) Presentation of A_2 with various salts in the increase trend. The effect of the type of salts follows the Hofmeister series.

consistent with the values of 28.7 and 30.5 Å reported in the literature (Bendedouch and Chen, 1983; Kozak, 2005). The molecular weight has been estimated for samples with protein concentration less than 20 mg/ml. Using an average contrast of $2.73 \cdot 10^{-6}$ Å^{-2}, we have 168.8 ± 26.7 kDa (D22, ILL), 83.1 ± 14.5 kDa (KWS2, Munich) and 151.9 ± 15.2 kDa (V4, Berlin). Compared to the theoretical value of the BSA monomer (66.1 kDa), a factor of 2 may prevail for values from different instruments. Note that the samples have been prepared always in the same way. However, the average error is larger than 10 % in any case, therefore, we did not use these values for further analysis.

The forward intensity, $I(0)$, is used to calculate A_2 according to eqn (4.6). In the case of the SANS data analysis, the incoherent background has to be subtracted from the forward intensity. The background is determined using the Porod law.46 By plotting $I(q)q^4$ vs. q^4, the slope of the linear relationship gives the incoherent background (Glatter and Kratky, 1982). As indicated in Fig. 4.4a, the constant backgrounds for BSA 4.5, 9.5, 19.5, 47.2 and 89.4 mg/ml with 2.0 M NaNO3 are 0.0004, 0.002, 0.010, 0.023 and 0.055

cm^{-1}, respectively. Compared to the forward intensity, the background contributes about 0.5–2 %. It is important to correct for it. For example, without background correction, A_2 for this condition would be $1.41 \cdot 10^4$ mol ml/g^2, however after background correction, $A_2 = 1.45 \cdot 10^4$ mol ml/g^2. Fig. 4.5a displays the typical plots of $c/(I(0) - B)$ versus c for protein solutions with various salts and ionic strength. In all of these, a good linear relationship is obtained. The factor a is constant for each series of samples where only protein concentrations are varied and the salt concentrations are kept constant. From the intercept and the slope of the linear fit, A_2 values are determined and listed in Table 4.1 and presented in Fig. 4.5b. In all cases, $A_2 > 0$, indicating the repulsive nature of the effective interactions under the present experimental conditions. This is consistent with the fact that all solutions are stable during the experiments. As displayed in Fig. 4.5b, the effect of added salt on the protein interactions is clearly visible. The A_2 follows the inverse Hofmeister series, $(NH_4)_2SO_4 < Na_2SO_4 <$ NaOAc $<$ NaCl $<$ $NaNO_3 <$ NaSCN. This effect, namely that protein interaction in solution follows the inverse order of the Hofmeister series, has been observed in several protein systems, such as lysozyme, α- and γ-crystallins (Bonnet et al., 1999; Tardieu et al., 1999; Finet et al., 2004; Finet and Tardieu, 2001). Using SAXS, Bonneté et al. studied the effective protein–protein interactions in lysozyme solutions as a function of temperature, ionic strength and salt type (Bonnet et al., 1999; Bonneté and Vivarès, 2002; Finet et al., 2004). A_2 for various conditions from under-saturated to supersaturated solutions were determined from the structure factor. The addition of salt can induce an additional attractive potential that is a function of the anion type and follows the order of the Hofmeister series. The order is direct or reverse when the pH values of the solutions are above or below the pI of proteins, respectively (Finet et al., 2004). A value of $A_2 = 1 \cdot 10^{-4}$ mol ml/g^2 has been reported by Tessier et al. for BSA solutions with NaCl at neutral pH (Tessier et al., 2002). A similar value is reported by Vilker et al. as $A_2 = 1.34 \cdot 10^{-4}$ mol ml/g^2 for BSA with 150 mM NaCl (Vilker et al., 1981). These values are in good agreement with our measurements by SANS (Table 4.1). With 1.0 and 2.0 M NaCl, the A_2 are 1.40 and 1.37 $\cdot 10^{-4}$ mol ml/g^2, respectively.

A_2 has been proposed to be the indicator for optimizing the conditions of protein crystallization. George and Wilson proposed that the best condition for protein crystallization corresponds to a narrow range of A_2, from -1.0 to $-8.0 \cdot 10^{-4}$ mol ml/g^2 (George and Wilson, 1994). Bonneté and Vivarés studied the crystallization conditions of different biological macromolecules with molecular weight from 14 kDa to 4600 kDa (Bonneté and Vivarès, 2002). It was found that for large macromolecules, the A_2 values are slightly positive, but smaller than those of the corresponding hard-spheres, indicating an attraction in addition to the hard-sphere repulsion. The dimensionless second virial coefficient $B_2 = M_w A_2/v = A_2 M_w^2/V/N_A$, which is rescaled from experimental to theoretical units by the molecular weight and corrected for the excluded volume V of a single particle, is suggested to be a better indicator for good protein crystallization conditions, because it depends only on the nature and the relative strength of the pair interaction potential and

Salt condition (M)		Station	A_2 (10^{-4} mol ml/g)	B_2 v	B_2 SANS	B_2 SANS + Isih.	B_2 v + hydr.	B_2 SAXS	B_2 SAXS + Isih.
				81.2[a]	91.7[a]	134.7[a]	114.5[a]	125.3[a]	157.5[a]
$(NH_4)_2SO_4$	1.5	KWS2	0.55 ± 0.16	4.9	4.4	3.0	3.5	3.2	2.5
Na_2SO_4	0.5	D22	1.06 ± 0.05	9.5	8.4	5.7	6.7	6.1	4.9
Na_2SO_4	1.0	D22	0.77 ± 0.03	6.9	6.1	4.2	4.9	4.5	3.6
NaOAc	1.0	D22	1.14 ± 0.11	10.2	9.0	6.1	7.2	6.6	5.3
NaOAc	2.0	D22	1.12 ± 0.09	10.0	8.9	6.1	7.1	6.5	5.2
NaCl	1.0	D22	1.40 ± 0.07	12.5	11.1	7.6	8.9	8.1	6.5
NaCl	2.0	D22	1.37 ± 0.09	12.2	10.8	7.4	8.7	7.9	6.3
$NaNO_3$	1.0	D22	1.56 ± 0.12	14.0	12.4	8.4	9.9	9.0	7.2
$NaNO_3$	2.0	D22	1.45 ± 0.06	12.9	11.4	7.8	9.2	8.4	6.7
NaSCN	0.5	KWS2	1.91 ± 0.05	17.0	15.1	10.3	12.1	11.0	8.8
NaSCN	1.0	KWS2	1.79 ± 0.08	16.0	14.2	9.7	11.4	10.4	8.3
NaSCN	2.0	KWS2	1.67 ± 0.04	15.0	13.2	9.0	10.6	9.7	7.7
NaSCN	3.0	KWS2	1.66 ± 0.09	14.9	13.2	9.0	10.5	9.6	7.7

Table 4.1: The second virial coefficient, A_2 and B_2, of BSA in various salt solutions, B_2 is calculated with different approaches as described in the text.
[a] The volume used for each method with the unit of $1000\,\text{Å}^3$.

not on the molecular size (Asthagiri et al., 2005). From the definition of B_2, the effective interaction between protein molecules becomes attractive for $B_2 < 4$ and repulsive for $B_2 > 4$.

However, the calculation of $B_2 = M_w A_2/v = A_2 M_w^2/(V/NA)$ depends on the choice of the volume, in particular whether the hydration shell and the non-spherical shape are taken into account. In principle, the volume can be calculated from the well-known specific volume of the protein, $V_{\text{bare}}^{(\text{th})} = 0.74\ \text{cm}^3/\text{g}\ M_w/N_A$, also including the hydration shell (assuming a 0.3 g/g hydration level), $V_{\text{hydr}}^{(\text{th})} = (0.74 + 0.3)\ \text{cm}^3/\text{g}\ M_w/N_A$. The ellipsoidal shape determined from SANS and SAXS experiments can be used to calculate the bare volume (SANS), $V_{\text{bare}}^{(\text{exp})} = 4\pi/3 \cdot 12.5 \times 41.9 \times 41.9\,\text{Å}^3$, and hydrated volume (SAXS), $V_{\text{hydr}}^{(\text{exp})} = 4\pi/3 \cdot 17 \times 42 \times 42\,\text{Å}^3$. Importantly, for a hard ellipsoid, the second virial coefficient B_2 does not only depend on the size but also on the aspect ratio. This effect of non-spherical shape can be accounted for with an effective sphere with the Isihara radius (see Methods section). Thus, we can calculate the volumes of these Isihara spheres for the experimentally determined ellipsoids, obtaining $V_{\text{bare}}^{(\text{Isihara})} = 4\pi/3 \cdot (31.8\,\text{Å})^3$ and $V_{\text{hydr}}^{(\text{Isihara})} = 4\pi/3 \cdot (33.5\,\text{Å})^3$. We summarize the different volumes and the obtained B_2 values in Table 4.1.

Obviously, different volumes simply correspond to a scaling factor for the virial coefficient. This scaling behavior, however, can induce mis-interpretation of the qualitative nature of the underlying interaction. Reconsidering that the excluded volume contribution is represented with $B_2 = 4$, a small factor can change the interpretation from a hard sphere

with additional repulsion to one with additional attraction. Indeed, this case is observed for samples with 1.5 M $(NH_4)_2SO_4$, where the virial coefficient changes from $B_2 = 4.9$ for the bare protein volume, i.e. additional repulsion, to $B_2 = 3.5$ for the hydrated protein volume, i.e. additional attraction. When considering the non-spherical shape, the values get even smaller, suggesting a stronger attraction.

For our measurements, the virial coefficients generally follow the expected relative order from the Hofmeister series for each of the six approaches. In particular, for salting-in salts, all calculated absolute values for B_2 correspond to a stronger repulsion than only the excluded volume. For salting-out conditions like $(NH_4)_2SO_4$ 1.5 M, Na_2SO_4 1.0 M or NaOAc 2.0 M, we expect an additional attraction or at least an interaction close to the excluded volume case. This expectation corresponds to $B_2 < 4$ and is only found if hydration, non-spherical shape or both are taken into account. Unfortunately, these two contributions are not as well-defined as the bare protein volume. The real protein shape differs from the ellipsoid of revolution and is hard to predict. Surface roughness has been shown to increase the excluded volume considerably (Frauenfelder et al., 2009; Makarov et al., 2002; Paliwal et al., 2005). For example, Neal and Lenhoff found that the excluded volume is about 6.7 times the molecular volume from the crystallographic structure instead of 4 times for a smooth sphere (Neal and Lenhoff, 1995). This result is comparable with our calculation using the volume from SAXS. However, the hydration shell is not a uniform layer of increased water density, and the hydration level can change due to salt conditions and temperature (Ortore et al., 2008). A precise description of the excluded volume contribution to the second virial coefficient is difficult and the use of A_2 or B_2 to describe the effective interaction between particles should be interpreted with care, in particular when the interactions are weak.

4.3.3 Experimental Structure Factor and Effective Protein–Protein Interactions from Model Fitting

For BSA in D_2O, assuming no H–D exchange, the scattering length density (SLD) would be about $1.84 \cdot 10^{-6}$ Å^{-2}. However, there are 1018 exchangeable hydrogens in one BSA molecule and up to 90 % of them can be exchanged after 24 h (Lu, 1999), which corresponds to a SLD of $3.04 \cdot 10^{-6}$ Å^{-2}. In practice, the H–D exchange depends on protein concentration (Nossal et al., 1986). By considering 70 % exchange, the SLD is $2.77 \cdot 10^{-6}$ Å^{-2}, which is close to the fitted value. For example, from the data with 4.9 and 89.4 mg/ml BSA with 0.1 M NaCl collected at D22 (Fig. 4.6a), when fixing the solvent SLD (D_2O, $6.34 \cdot 10^{-6}$ Å^{-2}), we get SLD of BSA of 2.71 and $2.87 \cdot 10^{-6}$ Å^{-2}, respectively. These values are consistent to those reported in the literature (Nossal et al., 1986). Since the difference of SLD between protein and solvent contributes to a constant factor of data fitting, for the remaining data fitting, we fix the SLD of the protein to $2.70 \cdot 10^{-6}$ Å^{-2}.

Fig. 4.6a presents the full q-range SANS data with simulated data derived from the model. With low ionic strength, SANS data for higher protein concentrations, the inter-

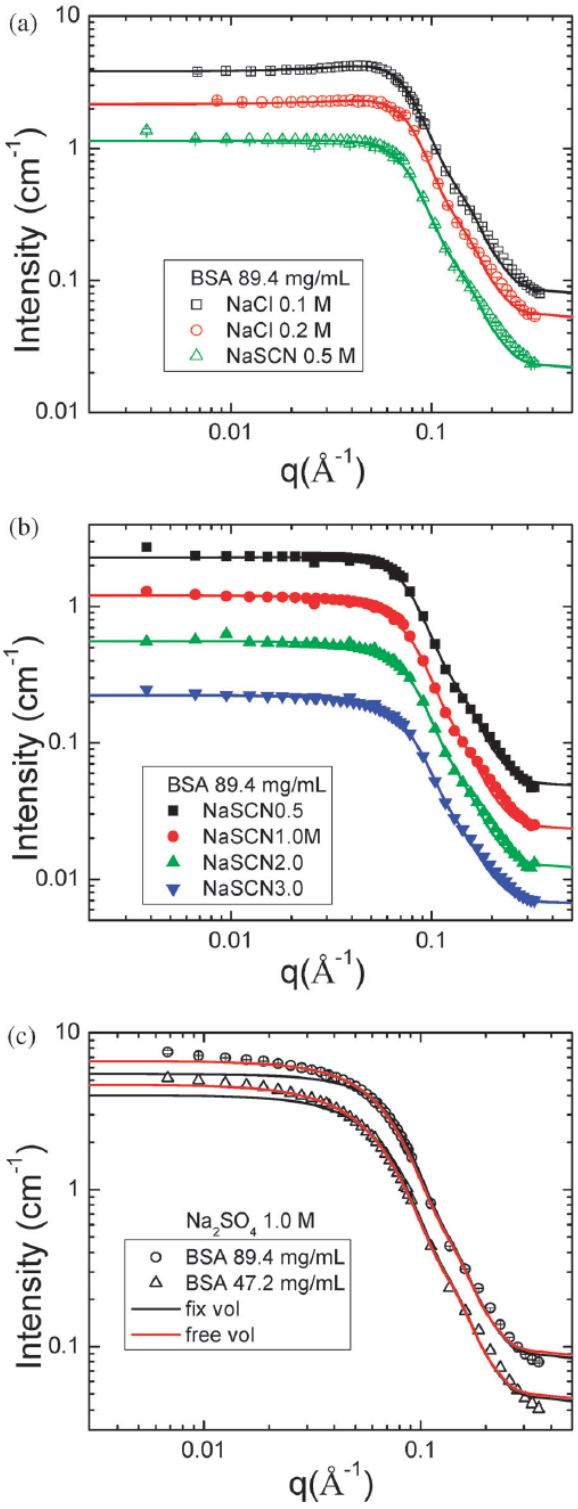

Figure 4.6: SANS data with model fitting: (a) BSA with NaCl (0.1 and 0.2M) and NaSCN (0.5 M). (b) Typical SANS data with model fitting for BSA 89.4 mg/ml with concentrated NaSCN (data from KWS2). (c) SANS data with model fitting for BSA with Na_2SO_4 indicate significant deviation from the hard sphere potential (see text). In all plots, only 30 % of data points were plotted.

Samples				Fitted	Fixed ionic	SLD solvent	Effective
protein	(g/l)	salt	(M)	volume (%)	strength (M)	(10^{-6} Å^{-2})	Charge (e)
BSA	47.2	NaCl	0.1	5.75	0.1	6.32	17.0
BSA	89.4	NaCl	0.1	7.82	0.1	6.32	16.6
BSA	47.2	NaCl	0.2	4.37	0.2	5.35	18.6
BSA	89.4	NaCl	0.2	9.22	0.2	5.27	21.5
BSA	89.4	NaSCN	0.5	7.08	0.5	5.42	36.6

Table 4.2: Fitting parameters for sample solutions with a low ionic strength. The errors of the parameters pertaining to the fitting procedure are better than $5\,\%$, but the systematic errors, including sample preparation, raw data reduction, and calibration, are estimated to $\approx 10\,\%$.

BSA		Fitted	SLD solvent	SLD protein
47.2/89.4 mg/ml		volume (%)	(10^{-6} Å^{-2})	(10^{-6} Å^{-2})
NaCl	1.0 M	4.78/9.97	5.23/5.12	2.70
NaCl	2.0 M	4.68/10.1	5.22/5.14	2.70
$NaNO_3$	1.0 M	3.80/8.84	6.40/6.06	2.70
$NaNO_3$	2.0 M	3.53/8.86	6.16/5.73	2.70
NaOAc	1.0 M	3.28/7.87	6.29/5.82	2.70
NaOAc	2.0 M	3.58/8.07	5.24/4.96	2.70
NaSCN	0.5 M	3.48	5.43	2.70
NaSCN	1.0 M	3.63/8.54	5.39/5.13	2.70
NaSCN	2.0 M	3.72/9.47	5.18/5.00	2.70
NaSCN	3.0 M	3.79/9.42	4.95/4.74	2.70
Na_2SO_4	1.0 M	1.49/4.73	6.34 fix	1.78/1.67
Na_2SO_4	1.0 M	4.00/8.00 fix	6.34 fix	2.49/2.63

Table 4.3: Fitting parameters for sample solutions with a high ionic strength. The errors of the parameters pertaining to the fitting procedure are better than $5\,\%$, but the systematic errors, including sample preparation, raw data reduction, and calibration, are estimated to $\approx 10\,\%$.

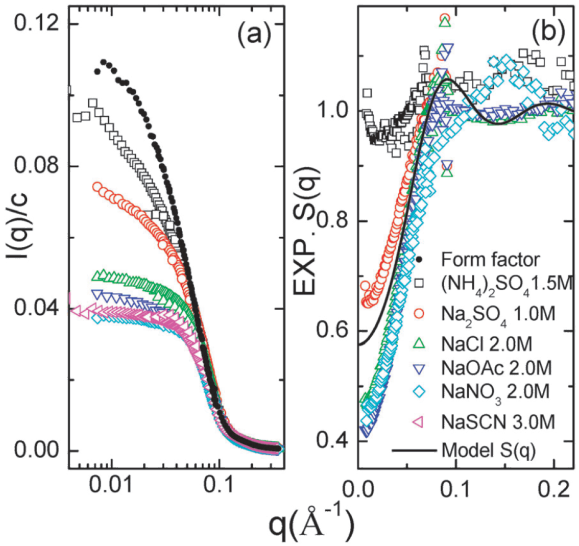

Figure 4.7: (a) A comparison of SANS data for BSA 89.4 mg/ml with various salts with the experimental form factor from BSA 9.5 mg/ml with NaCl 1.0 M. All data have been corrected for protein concentration. (b) Experimental structure factor from (a). The solid line is the theoretical curve of $S(q)$ of a hard sphere potential with an effective diameter of 63.5 Å and a volume fraction of 7.0 % for comparison.

action can be well described using a screened Coulomb potential in the structure factor. The fit parameters are listed in Table 4.2. For BSA 47.2 and 89.4 mg/ml, the experimental volume fractions are 3.49 % and 6.62 %, respectively. The fitted volume fraction values are systematically higher and the surface charges are 16 to 20 e. These results are in good agreement with our previous study using SAXS as well as those reported in the literature (Sinibaldi et al., 2008; Zhang et al., 2007). It is interesting to see that for 0.5 M NaSCN, the data with 89.4 mg/ml BSA are better described using a screened Coulomb potential, indicating that the surface charge still plays a role. The higher surface charge (36.6 e) can be explained by the strong salting-in effect of the thiocyanate anion. Studies have shown that SCN^- can strongly adsorb onto serum albumin. As many as 40 SCN^- ions can be bound to each albumin molecule (Arakawa and Timasheff, 1982; Scatchard et al., 1950). With high salt concentration (> 0.5 M), the SANS data are best described using form factor plus a hard sphere structure factor. Fig. 4.6b shows typical SANS data along with model fitting using a hard sphere potential in the presence of NaSCN. The only fitting parameter is the volume fraction of protein in the solutions. As listed in Table 4.3, the obtained volume fractions are generally higher than the experimental values. This is consistent since the fitted volume fraction is based on effective spheres with the Isihara radius.

When salting-out salts were added, such as Na_2SO_4 and $(NH_4)_2SO_4$, the SANS data of course cannot be understood using a hard sphere potential (Fig. 4.6c). Using the free volume and the SLD of protein as the fit parameters, the resulting volume fraction and the SLD values are too low (Table 4.3), i.e. SLD $< 1.84 \cdot 10^{-6}$ Å^{-2}. When fixing the volume fraction, one finds that the model calculated profiles deviate significantly from the SANS

data in the low q region (Fig. 4.6c). This deviation is also observed for samples with $(NH_4)_2SO_4$. The steady increase of scattering intensity in the low q region might be due to the appearance of an attractive potential. These results SANS data by the form factor directly. The normalized SANS data were plotted together with an experimental form factor (with BSA 9.5 mg/ml and NaCl 1.0 M) (Fig. 4.7a). It is obvious that all the samples in the low q region ($< 0.05\,\text{Å}^{-1}$) have lower intensity compared to that of the form factor, indicating a weak repulsion. Experimental structure factors are shown in Fig. 4.7b. For comparison, a hard sphere structure factor was calculated using the Isihara radius of 31.8 Å and a volume fraction of 7 %. The hard sphere potential is used here to monitor the subtle changes on potentials upon addition of different salts. Comparing with the model curve in the low q region, one sees that with Na_2SO_4 and $(NH_4)_2SO_4$, the low q values are higher than the model value, which indicates attraction, but the overall values of $S(q)$ in the low q region are still smaller than unity. Comparing with the previous section, we note that the second virial coefficient with correction for the non-spherical shape of proteins using the Isihara radius describes the interactions better than that determined from the bare protein. By comparing different salts, one can observe the consistent results of the Hofmeister effect as observed in other protein systems (Bonnet et al., 1999; Finet et al., 2004).

4.4 Conclusions

In this work, we have reported a SANS study on the hydration and protein interactions in solution containing concentrated electrolytes. The data are also compared with a previous study using SAXS (Zhang et al., 2007). SANS measurements from three stations (D22, KWS2 and V4) give very consistent results: first, the form factors $P(q)$ extracted from dilute protein solutions give an averaged dimension of BSA of $12.5 \times 42 \times 42\,\text{Å}^3$, corresponding to a volume of 91 700 Å^3 which is about 7.5 % and 13 % larger than the values calculated from crystal structure of HSA and the specific volume of BSA, respectively. This dimension is significantly smaller than that determined by SAXS as $17 \times 42 \times 42\,\text{Å}^3$ with a volume of 125 300 Å^3. The difference in volume (37 %) supports the idea of the contribution of a hydration shell with a hydration level of ≈ 0.3 g/g protein. These results indicate that compared to SAXS, the SANS measurements are not sensitive to the $\approx$ 10 % density difference between the hydrated and bulk water due to the small thickness of this hydration shell ($\approx$ 3 Å), which makes a combination of SANS and SAXS an ideal tool to study the hydration effect in protein solutions under various conditions.

Second, the effective protein–protein interactions in solutions containing concentrated salts can be described using the second virial coefficient, A_2, and its dependence on the nature of the added salts is clearly visible. It is found that A_2 follows the reverse order of the Hofmeister series, i.e. $(NH_4)_2SO_4 < Na_2SO_4 < NaOAc < NaCl < NaNO_3 < NaSCN$. The calculation of the dimensionless second virial coefficient B_2 reveals that the hydration and the non-spherical shape of proteins have to be considered for a better description of

interactions in protein solutions with concentrated electrolytes. SANS data are further analyzed using the full q-range based on liquid theoretical approaches, confirming the results of experimental structure factor and the A_2 analysis. Consistently, the additional interaction on top of the hard sphere repulsion is found to change from repulsive for salting-in conditions to attractive for salting-out conditions.

Acknowledgements

We gratefully acknowledge financial support from Deutsche Forschungsgemeinschaft (DFG) and the beam time allocation from ESRF, ILL, JCNS and Helmholtz-Center Berlin (BENSC). The beam time on V4 at the Helmholtz Zentrum Berlin has been supported by the European Commission under the 6th Framework Program through the Key Action: Strengthening the European Research Area, Research Infrastructures. Contract No. RII3-CT-2003-505925 (NMI3).

Chapter 5

Interplay of pH and Binding of Multivalent Metal Ions: Charge Inversion and Reentrant Condensation in Protein Solutions

Chapter 5 is based on the following publication:

Interplay of pH and Binding of Multivalent Metal Ions: Charge Inversion and Reentrant Condensation in Protein Solutions

Felix Roosen-Runge*, Benjamin S. Heck, Fajun Zhang, Oliver Kohlbacher and Frank Schreiber*

* corresponding authors

J. Phys. Chem. B, 117 (2013) 5777–5787

Contributions:

Research design	FRR,FZ
Experiments	BSH, FZ
Model development and calculation	FRR
Data analysis and interpretation	FRR
Paper writing	FRR, OK, FS

Roosen-Runge et al. (2013a)

Tuning of protein surface charge is a fundamental mechanism in biological systems. Protein charge is regulated in a physiological context by pH and interaction with counterions. We report on charge inversion and the related reentrant condensation in solutions of globular proteins with different multivalent metal cations. In particular, we focus on the changes in phase behavior and charge regulation due to pH effects caused by hydrolysis of metal ions. For several proteins and metal salts, charge inversion as measured by elec-

trophoretic light scattering is found to be a universal phenomenon, the extent of which is dependent on the specific protein-salt combination. Reentrant phase diagrams show a much narrower phase-separated regime for acidic salts such as $AlCl_3$ and $FeCl_3$ compared to neutral salts such as YCl_3 or $LaCl_3$. The differences between acidic and neutral salts can be explained by the interplay of pH effects and binding of the multivalent counterions. The experimental findings are reproduced with good agreement by an analytical model for protein charging taking into account ion condensation, metal ion hydrolysis and interaction with charged amino acid side chains on the protein surface. Finally, the relationship of charge inversion and reentrant condensation is discussed, suggesting that pH variation in combination with multivalent cations provides control over both attractive and repulsive interactions between proteins.

5.1 Introduction

Tuning of protein surface charge is a fundamental mechanism ensuring protein stability and function in aqueous solutions (Warshel and Levitt, 1976; Pace et al., 2000; Lund and Jönsson, 2005). The protein surface charge stabilizes protein solutions (Derjaguin and Landau, 1941; Verwey and Overbeek, 1948) and cluster phases (Groenewold and Kegel, 2001; Stradner et al., 2004). Non-homogeneous surface charge distributions represent anisotropic interaction patches which crucially affect phase behavior of solutions (Bianchi et al., 2006; Gögelein et al., 2008) as well as pathways for aggregation and crystallization (Whitelam, 2010a; Buell et al., 2009).

In electrolyte solutions, control on the protein charge is realized mainly by two processes: (de)protonation of functional surface groups and counterion condensation on the protein surface. Protonation and deprotonation, as investigated in titration experiments (Tanford et al., 1955; Tanford and Kirkwood, 1957), modulates the charge state of basic and acidic amino acid side chains (Lys, Arg, Glu, Asp, His) as well of the protein termini (carboxy and amino terminus). As in every proteolytic equilibrium, the charge state of each side chain – and thus the overall protein charge – depends on the pH of the surrounding medium.

Upon adding salt ions, charge regulation becomes more complex due to the ion-protein interaction. Many phenomena such as ion condensation (Manning, 1978) and ion binding (Scatchard, 1949) have been found to play a role in protein-ion interaction and charge regulation (Kurut and Lund, 2013). The discovery of the Hofmeister series (Hofmeister, 1888) inspired an enduring effort to understand the protein-salt interaction, focusing on surface hydrophobicity (Schwierz et al., 2010), ion hydration (Collins, 1997) and ion polarizability (Heyda et al., 2009). In fact, ion specific effects on the surface charge of proteins in electrolyte solutions are still a challenge for theory (Kunz, 2010). The association of salt counterions with inversely charged side chains at the protein surface seems to be the dominant effect and is suggested as the basic model for ion-protein interaction (Jungwirth

and Winter, 2008). Ion association has been identified to occur mainly at hydrophilic sites surrounded by hydrophobic surface areas (Yamashita et al., 1990). Amino acids with carboxylate, thiol, thioether, and imidazole groups can coordinatively bind transition metal ions (Tainer et al., 1991; Harding, 2001). Taking into account also the charge of the molecule, binding models have provided the basis for the understanding of interactions of proteins with ions and ligands (Scatchard, 1949; Feldman, 1972; Klotz, 1982; Tipping, 1998).

Importantly, also non-local effects, i.e. not only dependent on one single functional group, affect ion association. First, since the protein/water interface exhibits a change of dielectric constant, also non-localized adsorption of polarizable, i.e. large, ions could occur at non-polar, hydrophobic areas of the protein surface (Jungwirth and Winter, 2008; Dahirel and Jardat, 2010). Second, ion-ion correlations and finite size effects close to surfaces with high charge density have been shown to cause overcharging of the surface, attraction between like-charged surfaces, and charge inversion of colloidal particles (Messina, 2009; Moreira and Netz, 2000; Deserno et al., 2000; Grosberg et al., 2002; Allahyarov et al., 1998; Besteman et al., 2004).

A complete model for the protein charge in any system has to include both local and non-local contributions, although many studies focus on only one approach (Lyklema, 2006; Travesset and Vangaveti, 2009).

The driving force of ion-induced surface charge is the change in free energy upon ion condensation on the surface, which can be described qualitatively – and to a reasonable degree also quantitatively – with an effective short-ranged interaction between protein and ions, even for long-ranged and non-localized interactions (Travesset and Vangaveti, 2009).

In this article, we present findings on the charge inversion of globular proteins, in particular bovine and human serum albumin (BSA, HSA), ovalbumin (OVA) and β-lactoglobulin (BLG), in the presence of multivalent metal cations. We focus on the effects of the trivalent metal ions Al^{3+}, Fe^{3+} and Y^{3+} and, in particular, the effects of pH variations arising from the metal hydrolysis. The charge inversion is reflected in a reentrant condensation of the protein solution (Zhang et al., 2008, 2010; Ianeselli et al., 2010), i.e. the protein solution is homogeneous and stable for low and high metal ion concentrations, while for intermediate metal salt concentrations a phase-separated state is found, corresponding to either a liquid-liquid phase separation or formation of amorphous clusters and crystals (Zhang et al., 2012c,a). These phenomena do not occur in solutions of monovalent cations and cannot be understood by conventional salting-in or salting-out effects (Arakawa and Timasheff, 1984; Zhang et al., 2007, 2012b).

After an outline of the experimental methods, we present experimental results on reentrant condensation in protein solutions with trivalent salts and the related inversion of the protein charge. We introduce a model for metal ion association and hydrolysis, which conceptually explains the observed phenomena. The theoretical methods are summarized in more detail in the Appendix.

5.2 Experimental Methods

5.2.1 Sample Preparation and Optical Characterization

All proteins were purchased from Sigma-Aldrich (BSA: 99 % purity, A3059; HSA: 97–99 %, A9511; BLG: 90 %, L3908; OVA: 98 %, A5503) and used as received. $FeCl_3$, $AlCl_3$ and YCl_3 were purchased from Sigma Aldrich at high purity grade (>99 %). Deionized water was used to prepare all solutions.

Stock solutions for proteins (200 mg/ml) and salts (200 mM) were prepared and then diluted in order to obtain the desired concentrations. The protein stock solution was filtered through 0.22 μm membrane filters (Millipore). No buffer was used in order to study solely the effect of the salt in the protein-protein interactions. pH of freshly prepared protein-salt solutions was measured using the pH meter Mettler Toledo InLab 413 SG/2n IP67, which was calibrated before measurement to standard buffer solutions. The phase behavior of protein solutions was determined using the optical transmission of the solutions (Zhang et al., 2010) or by visual inspection. All preparations and measurements were performed at room temperature.

In order to ensure structural stability of the proteins at the given metal ion concentration and pH, the secondary structure content was monitored using Fourier Transform Infrared spectroscopy. No significant changes have been found in the amide I ($\approx 1650\,\mathrm{cm}^{-1}$) and amide II bands ($\approx 1450\,\mathrm{cm}^{-1}$) at different metal ion concentrations (for details and data see Supplementary Material), implying that the secondary structure of the protein remains invariant for the range of salt concentration studied.

5.2.2 Electrophoretic Light Scattering

All measurements of electrophoretic mobility were performed with a Zetasizer Nano (Malvern Instruments GmbH, Germany), using electrophoretic light scattering (ELS) via phase analysis light scattering (PALS). From the electrophoretic mobility, the so-called zeta potential ζ can be calculated, which is the potential at the outer border of the hydrodynamic stagnant layer around the protein. In general, ζ reflects the total charge within the stagnant layer, but the analytical relation is not trivial in the case of soft, rough and non-spherical objects (Delgado et al., 2007). The next paragraphs outline the calculation of the protein charge from ELS measurements. Note that this approach is based on several assumptions, most importantly a mean-field theoretical approach for the ion distribution and an effective sphere as representation of the protein. Since both assumptions are not fully fulfilled in our system, the resulting charges should be considered as effective charges (Menon and Zydney, 1998) which are not absolutely precise, but suitable for relative comparison and further discussion.

ELS measures the particle mobility μ_E in an external electrical field which is related to the zeta potential for a spherical particle with radius a (Delgado et al., 2007):

$$\mu_E = \frac{2\varepsilon}{3\eta} f(\kappa a)\, \zeta \; . \tag{5.1}$$

ε is the total dielectric permittivity and η the viscosity of the medium. The inverse Debye screening length κ and the Bjerrum length λ_B represent the relevant length scales for the electrostatic interaction:

$$\kappa^2 = 4\pi\lambda_B \sum_i n_i Z_i^2 \qquad , \quad \lambda_B = e^2/(\varepsilon k_B T) \; . \tag{5.2}$$

Here, all present ionic species i are accounted for, with the respective valency Z_i and number concentration n_i. e denotes the elementary charge and $k_B T$ the Boltzmann constant multiplied with the temperature.

The Henry function $f(\kappa a)$ in Eq. (5.1) interpolates between the Helmholtz-Smoluchowski limit ($\kappa a \gg 1$: $f = 1.5$) and the Hückel-Onsager limit ($\kappa a < 1$: $f = 1$). $f(\kappa a)$ can be approximated by Ohshima's relation (Delgado et al., 2007; Ohshima, 1994):

$$f(\kappa a) = 1 + \frac{1}{2}\left[1 + \left(\frac{2.5}{\kappa a[1 + 2\exp(-\kappa a)]}\right)\right]^{-3} \; . \tag{5.3}$$

The reduced zeta potential $\tilde{\zeta} = \frac{e\zeta}{2k_B T}$ is related to the surface charge density σ (Ohshima et al., 1982):

$$\begin{aligned} \frac{\sigma e}{k_B T} \;=\; 2\epsilon\kappa \Bigg[& \sinh^2\left(\frac{\tilde{\zeta}}{2}\right) + \frac{2}{\kappa a}\tanh^2\left(\frac{\tilde{\zeta}}{4}\right) \\ & + \frac{8}{(\kappa a)^2}\ln\left(\cosh\left(\frac{\tilde{\zeta}}{4}\right)\right) \Bigg]^{1/2} \; . \end{aligned} \tag{5.4}$$

Based on σ we calculate the total charge of the protein-salt complex from

$$Q = 4\pi a^2 \sigma \quad . \tag{5.5}$$

For the different proteins we estimated the following effective radii (Zhang et al., 2012b): $a_{\mathrm{BSA}} = a_{\mathrm{HSA}} = 3.3\,\mathrm{nm}$, $a_{\mathrm{BLG}} = 2.7\,\mathrm{nm}$, $a_{\mathrm{OVA}} = 3.6\,\mathrm{nm}$. For BLG and OVA, we considered dimers.

5.3 Experimental Results

5.3.1 Reentrant Condensation

Fig. 5.1 summarizes the phase diagrams determined by optical transmission and visual inspection showing reentrant condensation behavior for BSA with three different metal

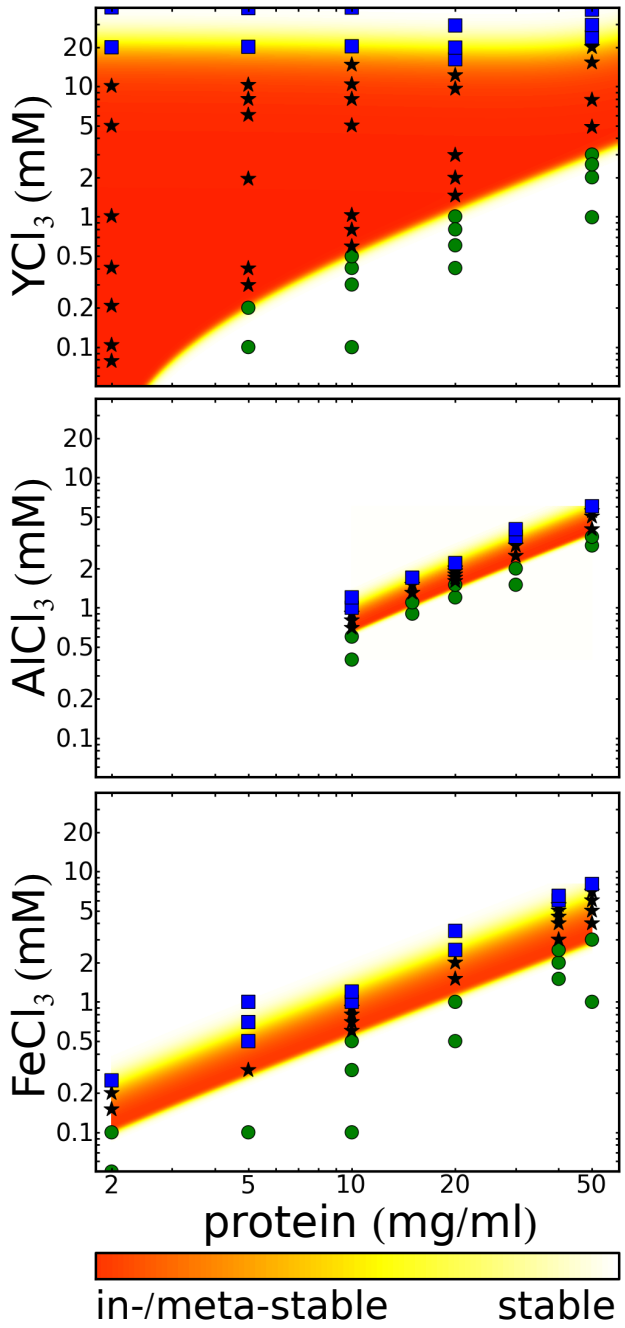

Figure 5.1: Reentrant phase behavior for BSA and different multivalent metal ions (Al^{3+}, Fe^{3+}, Y^{3+}). The protein solution is stable for low (green circles) and high (blue squares) metal ion concentrations (white color-coding). For intermediate metal ion concentrations, an unstable or metastable protein solution exists (black stars and red color-coding). Clear differences can be observed between the different metal ions. In particular, the condensed regime occurs on a much narrower metal concentration range for Al^{3+} and Fe^{3+} than for Y^{3+}, which is an indication of relevant pH effects due to metal ion hydrolysis. Part of the data are taken from Zhang et al. (2008, 2010)

ions. In earlier studies, reentrant condensation has been observed as a universal phase behavior for several globular proteins with negative native charge (e.g. BSA, HSA, OVA, BLG) in the presence of multivalent metal cations (e.g. Al^{3+}, Fe^{3+}, Y^{3+}, La^{3+}) (Zhang et al., 2008, 2010). At low salt concentrations, solutions are clear; above a critical salt concentration c^* protein solutions enter a phase-separated state. Further increase of the salt concentration results in a clear solution again.

In the context of this study, it is important to emphasize that there are significant differences concerning the phase-separated regime (Fig. 5.1). Quantitatively, the concentration range for the phase-separated regime is clearly narrower for Al^{3+} and Fe^{3+} than for Y^{3+}. Qualitatively, the nature of the phase separation differs: while for $AlCl_3$, gels and aggregates are formed, only amorphous aggregates are found in the case of $FeCl_3$. For YCl_3, clustering, amorphous aggregates and liquid-liquid phase separation are observed (Zhang et al., 2012c).

5.3.2 Inversion of Surface Charge

Fig. 5.2 presents the inversion of surface charge for several typical sets of protein-metal ion solutions. The surface charge is calculated by Eq. (5.4, 5.5) from ELS measurements. For all studied combinations of proteins with acidic pI (here: BSA, HSA, BLG, OVA) and multivalent metal salt (here: YCl_3, $AlCl_3$, $FeCl_3$), a clear inversion of the surface charge is observed upon increasing the metal ion concentration. In Fig. 5.2(top), we observe a clear dependence of the point of zero charge c_0 on the protein concentration. This effect is expected, since the higher number density of protein requires a higher concentration of metal ions to compensate for the native protein surface charge.

Fig. 5.2(bottom) illustrates the deviations of the quantitative extent of charge inversion with respect to the combination of protein and metal salt. First, the maximum surface charge after charge inversion depends on the protein. This finding is expected, since the number of association sites for metal ions will differ with protein size and structure. While BSA and HSA have comparable size and show comparable absolute numbers, the OVA dimer is significantly larger and apparently obtains higher surface charges after charge inversion. Second, the nature of the metal salt has an effect. Different association behavior is possible, and pH effects due to metal salt hydrolysis are expected to contribute to the charge inversion. The strength of the pH effect is reflected in the large shift of c_0 between rather neutral salts (YCl_3) and acidic salts ($AlCl_3$, $FeCl_3$).

5.3.3 Semiempirical Fit of Surface Charge

The solid lines in Fig. 5.2 represent fits with a Langmuir-like equation for the protein charge upon variation of salt concentration. The starting point for fitting these charge profiles is the well-known Langmuir isotherm for the association of a ligand at N independent sites with association constant K. In our case, however, two mechanisms are expected to

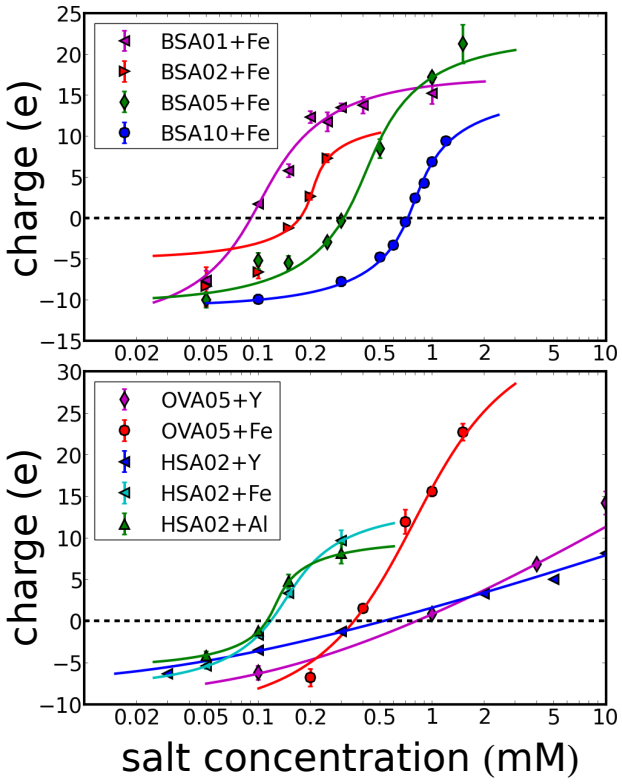

Figure 5.2: Inversion of protein surface charge induced by multivalent metal ions as calculated from electrophoretic measurements. All solid lines correspond to Langmuir isotherms with an additional cooperativity parameter (Eq. (5.6)). Top: Charge inversion for protein concentration series (1, 2, 5, 10 mg/ml BSA) with $FeCl_3$: With increasing protein concentrations, the point of zero charge shifts to higher salt concentrations. The higher concentration of proteins requires a higher amount of metal ions to compensate the native surface charge. Bottom: Comparison between different salts for OVA (5 mg/ml) and HSA (2 mg/ml) with $FeCl_3$, $AlCl_3$ and YCl_3: The charge inversion is universally present in all measured samples of BSA, HSA, BLG and OVA in presence of $FeCl_3$, $AlCl_3$ and YCl_3. In general, the charge inversion occurs at lower salt concentration for acidic salts like $FeCl_3$ and $AlCl_3$ compared to YCl_3, indicating a relevant pH effect due to metal salt hydrolysis.

change this profile. First, each counterion on the surface will change the protein charge and thus suppress further ion association. Second, the hydrolysis of the metal salts lowers the pH in the experimental concentration range, causing an additional positive charge due to protonation of acidic residues.

In order to semiempirically allow for these two mechanisms, i.e. competitive binding and pH effects, the association constant K is rewritten as a product of a intrinsic association constant K' and an exponential, representing the first-order correction of the free energy with respect to Q: $K(Q) = K' \cdot \exp(pQ)$. The cooperativity parameter p corresponds to the free energy per protein charge. Introducing this modification into the Langmuir isotherm, the fit function reads:

$$Q = Q_0 + \frac{ZNc}{c + K' \cdot \exp(pQ)} \,. \tag{5.6}$$

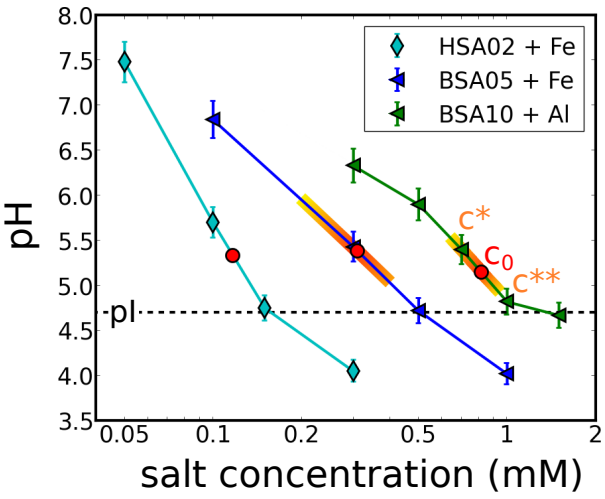

Figure 5.3: pH of protein solutions (HSA 2 mg/ml, BSA 5 mg/ml and 10 mg/ml) with $FeCl_3$ and $AlCl_3$. Both charge inversion, indicated by the estimated point of zero charge c_0 (red circle), and reentrant condensation, indicated by the critical salt concentrations c^* and c^{**} (colored area), occur at pH values above the isoelectric point pI (black dashed horizontal line). Thus, binding of counterions has a non-negligible effect on the charge regulation.

Z is the valency of the metal ions, c is the metal ion concentration, Q_0 is the protein charge without metal ions and N is the number of accessible association sites. The exponential factor renders the equation implicit in the desired variable Q; the solution is obtained from a numerical bisection with respect to Q using the routine supplied by the python package *scipy.optimize* (Jones et al., 2001; Oliphant, 2007) (version 0.7.0).

Using Eq. 5.6, the charge profiles from ELS can be fitted with reasonable accuracy (Fig. 5.2). The fit parameters for all sample series are given in the supplementary material. Except from some unphysical fits due to insufficient range of metal concentration, the fit parameters show a consistent behavior: the protein charge Q_0 lies between -5 and -12. The number of association sites N varies between 4 and 11. Interestingly, the parameter p is mainly negative. The binding of counterions to the protein, however, is a competitive binding due to the surface charge variation, implying a positive p. The negative value of p thus implies a non-negligible pH effect on the charge inversion.

5.3.4 pH of Protein Solutions with Multivalent Salts

Fig. 5.3 summarizes some results on the pH values of protein solutions with acidic multivalent salts. As expected, the pH decreases with increased salt concentration down to pH 4. At higher protein concentration, the pH decreases more slowly, presumably due to self-buffering of the protein and binding of multivalent salt. Along with the measured pH values, the point of zero charge c_0 (red circles), as taken from the fits, and the phase boundaries of the phase-separated regime, c^* and c^{**} (yellow stars), are shown. Importantly, both

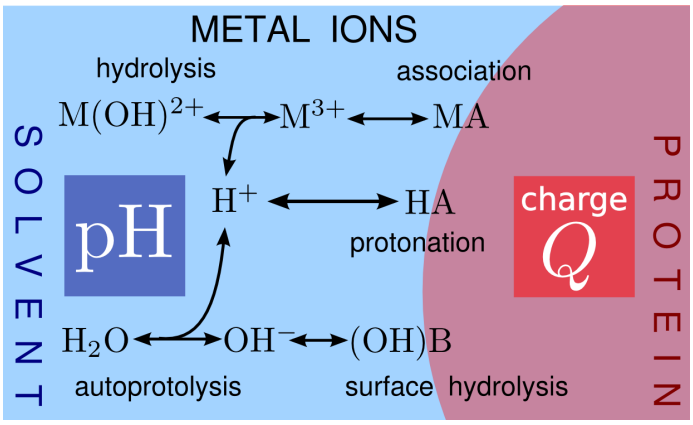

Figure 5.4: Schematic representation of the reactions incorporated in the suggested model. The surface equilibria of metal association, protonation, and hydrolysis link the total particle charge to the solution pH. Metal ions with pronounced pH effect due to hydrolysis thus have a non-trivial effect on the surface charge.

charge inversion and reentrant condensation take place at pH values above the isoelectric point pI $\approx$ 4.7 of BSA (Peters, 1996). The charge inversion thus cannot be driven solely by pH variation of the solution (Zhang et al., 2010); binding of multivalent counterions has also a non-negligible effect.

5.4 Analytical Model for the Interplay of Ion Association and pH Effects

The experimental results in the previous section indicate that both pH and binding of multivalent ions play an important role for the protein surface charge. Inspired by these experimental findings, we present an analytical model which incorporates several association reactions (for a schematic representation, see Fig. 5.4). (De)Protonation and association of metal counterions take place at the same functional groups at the protein surface and are thus in competition. Due to the geometrical compactness of the protein, the functional groups are not independent, but coupled to each other since the protein total charge contributes non-negligibly to the association free energy. Furthermore, the solution chemistry adds another level of complexity due to the hydrolysis of metal salts and the related changes of the pH in the protein solution. In essence, there is a interplay of protein surface charge, solution pH and salt concentration.

Using a model of a spherical particle in a solution, we aim to extract some essential features by solving a system of association reactions while simultaneously correcting for charge and pH effects on the association equilibria. Note that the model should not be regarded as a detailed quantitative description, which would obviously fail, amongst other factors, for the reasons of missing anisotropy effects from non-spherical particle geometry and mean-field approximations for the ion distributions. However, general trends of ion

association, interwoven with pH effects, should be accessible from this model which we believe to capture the essential points.

Concerning the surface association, the calculation scheme is inspired by adsorption models for metal complexes at surfaces (James and Healy, 1972; Agashe and Regalbuto, 1997) as well as ideas from well-known titration calculations for proteins (Tanford et al., 1955; Tanford and Kirkwood, 1957; Bashford and Karplus, 1991) and models for ion binding (Scatchard, 1949; Feldman, 1972; Klotz, 1982; Tipping, 1998; Munson and Rodbard, 1980). The hydrolysis of the trivalent metal salts used here is well-characterized (Byrne et al., 2000; Flynn, 1984; Lydersen, 1990; Wood, 1990). To our knowledge, so far neither simulation nor theoretical models have addressed the coupling of metal salt hydrolysis and surface association as present in unbuffered solutions of proteins with salts.

5.4.1 Association Reactions

When considering a protein in aqueous solution, first and importantly, the charge regulation via acidic and basic residues has to be accounted for. The basic residues are denoted as B^+. Not all acidic residues are expected to act as metal association site, e.g. for steric reasons. D^- represents acidic residues acting only as protonation site while A^- denotes a site with both protonation and metal association.

$$\begin{aligned} \mathrm{HA} &\overset{K_a}{\rightleftharpoons} \mathrm{A}^- + \mathrm{H}^+ \\ \mathrm{HD} &\overset{K_a}{\rightleftharpoons} \mathrm{D}^- + \mathrm{H}^+ \\ \mathrm{BH}^+ &\overset{K_b}{\rightleftharpoons} \mathrm{B} + \mathrm{H}^+ \end{aligned} \tag{5.7}$$

This reaction has to be complemented by the autoprotolysis of water

$$\mathrm{H_2O} \overset{K_w}{\rightleftharpoons} \mathrm{OH}^- + \mathrm{H}^+ \quad . \tag{5.8}$$

Adding multivalent metal ions to the solution adds further association reactions. The hydrolysis of metal ions varies the solution pH, which in turn affects the charge regulation of the protein.

$$\begin{aligned} \mathrm{M}^{3+} + \mathrm{H_2O} &\overset{K_0}{\rightleftharpoons} \mathrm{M(OH)}^{2+} + \mathrm{H}^+ \\ \mathrm{M(OH)}^{2+} + \mathrm{H_2O} &\overset{K_1}{\rightleftharpoons} \mathrm{M(OH)}_2^+ + \mathrm{H}^+ \\ \mathrm{M(OH)}_2^+ + \mathrm{H_2O} &\overset{K_2}{\rightleftharpoons} \mathrm{M(OH)}_3 + \mathrm{H}^+ \end{aligned} \tag{5.9}$$

We use a simplified notation where M^{3+} represents possible metal aquo complexes with the given charge, for example the octahedral hexaaquaaluminium complex ion $\mathrm{Al(H_2O)}_6^{3+}$, for other valencies analogously (Martin, 1986). The limited solubility of $\mathrm{M(OH)}_3$ complexes will be discussed in the next subsection. We have neglected further species like $\mathrm{M(OH)}_4^-$ since these will play a role in basic solutions only (Martin, 1986).

	pK_s'	pK_0	pK_1	pK_2	pK_{m0}
Fe	4.9 [a]	2.2 [a]	3.5 [a]	6.0 [a]	8.6
Al	10.4 [b]	5.0 [b]	5.1 [b]	6.6 [b]	5.7
Y	18.05 [c]	7.7 [d]	10.0	15.0	4.2

Table 5.1: Values for the equilibrium constants and solubility products defined in the previous section. All values are given as decadic logarithm. Note that superscripts do not denote exponents, but specify the references. Values without reference have been estimated, since no precise literature value is given to our knowledge (see discussion in text). [a] Flynn (1984); [b] Lydersen (1990); [c] Diakonov et al. (1998); [d] Wood (1990)

Finally, binding of metal ions to acidic sites at the protein surface (each specimen, respectively) modulates the surface charge:

$$\begin{aligned} \mathrm{M}^{3+} + \mathrm{A}^- &\overset{K_{m0}}{\rightleftharpoons} \mathrm{MA}^{2+} \\ \mathrm{M(OH)}^{2+} + \mathrm{A}^- &\overset{K_{m1}}{\rightleftharpoons} \mathrm{M(OH)A}^{+} \\ \mathrm{M(OH)}_2^{+} + \mathrm{A}^- &\overset{K_{m2}}{\rightleftharpoons} \mathrm{M(OH)}_2\mathrm{A} \\ \mathrm{M(OH)}_3 + \mathrm{A}^- &\overset{K_{m3}}{\rightleftharpoons} \mathrm{M(OH)}_3\mathrm{A}^- \quad . \end{aligned} \tag{5.10}$$

We do not incorporate other association reactions of monovalent ions (like Na^+ and Cl^-) with both protein and metal salt, since these reactions would unnecessarily complicate the model. For simplicity we assume complete solvation of salt and protein. Note that monovalent ions are implicitly accounted for by the ionic strength which screens the electrostatic contributions (see subsections on free energy contributions and on phase behavior from a DLVO picture). Although assumptions on the ionic strength might quantitatively change the model parameters, the qualitative behavior from the binding of trivalent salt ions and their coupling with the pH is expected to be similar.

5.4.2 Equilibrium Constants and Solubility

The association reactions mentioned in the previous subsection incorporate several equilibrium constants K_i which have to be chosen reasonably. While the equilibrium constants for the aqueous chemistry can be obtained from literature values, an estimation of equilibrium constants for association reactions on the protein surface is more involved. In principle, the association constant for metal ions and protonation at the protein surface is locally varying with a rather broad distribution. The local shifts arise from inhomogeneities in the electrostatic surface potential and solvation energy due to the local surface morphology as well as possible multidentate binding configurations. Since we are aiming for a qualitative understanding of the pH effects on the charge inversion in a simplified picture, we use a single value to represent a hypothetical homogeneous association process of each species, respectively. For the autoprotolysis of water, we used $pK_w = 14$.

Hydrolysis of Metal Ions: The used values are summarized in Table 5.1. For the hydrolysis constants, systematic variations are expected due to temperature and total ionic strength (Byrne et al., 2000). The general trend and relation between the metal salts, however, is expected to be conserved. The values for $\mathrm{p}K_2$ have additional uncertainties since the values will depend on the choice how to distinguish between $M(OH)_3$ and the corresponding insoluble solid state which might be still present in a cluster state (Byrne et al., 2000).

For the case of Y^{3+}, the literature values for $\mathrm{p}K_1$ and $\mathrm{p}K_2$ have been estimated since no published values exist to our knowledge. However, the pH below 7 in the examined solutions makes these reactions negligible.

Solubility of Metal Hydroxides: Using the solubility product K_s, we obtain $K'_s = K_s/K_w^3 = [\mathrm{M}^{3+}]_{\max}/[\mathrm{H}^+]^3$, which can be used to calculate the maximum concentration of free, unbound metal ions at a given pH: $[\mathrm{M}^{3+}]_{\max} = 10^{-\mathrm{p}K'_s - 3\mathrm{pH}}$.

Solubilities vary by several orders of magnitude between solid phases with aging times over years (e.g., gibbsite $Al(OH)_3$ or different hydrous ferric oxides as hematite and geothite) and the amorphous state aged only for hours (Martin, 1986; Lydersen, 1990; Flynn, 1984; Diakonov et al., 1998). The experiments discussed in this study were performed within hours after preparation, suggesting the amorphous state after several hours as the suitable reference system. Accordingly, these solubility products K'_s are listed in Table 5.1.

During our experiments, no precipitation of metal hydroxide $M(OH)_3$ has been observed. Metal ions are introduced via a chloride salt. Most metal atoms occur in hydroxy complexes and thus lower the pH considerably. Furthermore, a considerable portion of the metal ions is removed from the solubility equilibrium due to association to the protein surface. During the calculation, the concentration of free, unbound ions was checked and found generally smaller than the corresponding maximum concentration $[\mathrm{M}^{3+}]_{\max}$. Thus, no metal hydroxide precipitation has to be taken into account.

Protonation and Dissociation Constants of Amino Acid Functional Groups: Protonation at the protein surface occurs mainly at the acidic amino acids, i.e. aspartic and glutamic acid. The related intrinsic equilibrium constants of the side chain in an isolated amino acid are $\mathrm{p}K'_a = 3.71$ and 4.15, respectively (Lide, 2009). Protonation of basic amino acid residues arginine and lysine has $\mathrm{p}K'_b = 12.10$ and 10.67 (Lide, 2009). In both cases, however, it has to be kept in mind that solvation energies as well as the local electrostatic environment shift the equilibrium constants for individual sites considerably. To account for these variations is the central task of more detailed titration calculations (Tanford and Kirkwood, 1957; Bashford and Karplus, 1991).

Estimating shifts in $\mathrm{p}K$ and aiming for representative binding constants for the pH range between 3 and 7, we use the following equilibrium constants: $\mathrm{p}K_a = 5$ and $\mathrm{p}K_b = 9$. The number of basic residues is set to 40. The number of acidic residues is set to 52, out of which 10 also can bind metal ions. The number of metal binding sites is chosen

consistent with the semiempirical fit from the experimental section. The total numbers of functional groups are reasonable choices based on reported values for functional groups. BSA and HSA as well as dimers of BLG and OVA have approximately 80 basic and 90-100 acidic residues (Zhang et al., 2010). Note, though, that considerable part of these might not dissociate in the given pH range due to solvation shifts in the dissociation constant.

Association of Metal Ions to the Protein: Several studies report equilibrium constants for the association of Y(III) to bone sialoprotein (Williams and Peacocke, 1967) as well as Fe(III) and Al(III) to transferrin (Aasa et al., 1963; Trapp, 1983; Martin, 1986). These values have to be taken with care, since a single equilibrium constant subsumes all hydrolysis species of metal ions in one constant which thus presumably will be pH-dependent (Martin, 1986). Furthermore, these studies mainly focused on metal storage and transport proteins with specialized cavities. However, also solvent-exposed glutamates and aspartates have been shown to represent binding sites for multivalent cations with low affinity and low cation specificity (Pastore et al., 2007). Since no values for the association constants of metal ions to side chains are given in the literature, we used those reported for carboxylate groups of single amino acids in solution. The equilibrium constants vary with the actual amino acid and solution conditions like buffer salts as well as temperature. For Fe^{3+}, values for $\mathrm{p}K_m$ range from 8.6 to 9.2; for Al^{3+}, values between 5.7 and 6.7 have been reported (Berthon, 1995). Even lower values are given for Y^{3+}, ranging between 4.4 and 5.4 (Sandhu, 1977; Berthon, 1995).

These values do not include any solvation effects which are relevant for binding towards a carboxylic acid at the protein surface and in general shift all equilibrium constants to lower values. Furthermore, the values do not correspond directly to $\mathrm{p}K_{m0}$, but incorporate binding of all metal hydroxy complexes simultaneously. $\mathrm{p}K_{m1}$ and $\mathrm{p}K_{m2}$ should be smaller due to a decreased electrostatic contribution to the binding. Since our approach is not aimed at describing the charging in full quantitative accuracy, but rather at qualitatively rationalizing the main effects, we choose the equilibrium constants $\mathrm{p}K_{m0}$, $\mathrm{p}K_{m1}$ and $\mathrm{p}K_{m2}$ with the following estimation procedure: First, we choose $\mathrm{p}K_{m0}$ as the lowest value from the published range as listed also in Table 5.1. These values account already for the loss of solvation due to the binding site, i.e. one amino acid. Neighboring amino acids of the proteins will cause further loss of solvation; the effect, however, is not easy to estimate. Thus, in a second step, we decrease the $\mathrm{p}K_{m0}$ by 2 to roughly account for the additional solvation effect in our coarse-grained model. Third, we estimate the pure electrostatic contribution to $\mathrm{p}K_{m0}$ from the binding distance $r_b \approx 2\,\text{Å}$ found in protein crystal structures (Harding, 2001). This gives $\Delta\mathrm{p}K_m \approx e^2/(4\pi\epsilon r_b k_B T \ln(10)) \approx 1.5$ per ligand charge. We thus obtain $\mathrm{p}K_{m1}$ and $\mathrm{p}K_{m2}$ by subtraction of 1.5 and 3 from $\mathrm{p}K_{m0}$, respectively.

5.4.3 Contributions to the Free Energy and Surface Charge Effects

The equilibrium constants given in the previous sections can be rewritten as

$$K = \exp(\beta \Delta G) = \exp(\beta(\Delta G_{\text{intr}} + \Delta G_{\text{es}})) \quad . \tag{5.11}$$

Here we subsume all contributions under only two terms: first, the intrinsic term ΔG_{intr} accounts for short-range, non-coulombic contributions. In particular, this term includes localized contributions like hydrogen bonding and coordinative bonds, but also accounts for non-localized free energy contributions arising from hydrophobic binding, solvation effects and other water-mediated interactions (Lyklema, 2006). Furthermore, entropic contributions and ion activities as well as free energy contributions due to ion correlations are included in the intrinsic part, since they show a comparable effect (Travesset and Vangaveti, 2009).

Second, ΔG_{es} accounts for the electrostatic long-range interaction of the global surface charge Q with the ligand. Using the surface potential Ψ_0, the terms reads

$$\Delta G_{\text{es}} = Ze\Psi_0 \quad , \tag{5.12}$$

where Z is the formal charge of the ligand. In other words, the local density of ions varies according to the electrostatic potential at the reaction center and, thus, the association reaction will also depend on the protein charge. Eq. (5.4) defines the relation between the zeta potential Ψ_0 and the global surface charge Q. With this in mind, all binding constants have to be considered as depending on Q:

$$K_a = K'_a\, \tau(Q)^{-1}\,,\; K_b = K'_b\, \tau(Q)\,,\; K_{mi} = K'_{mi}\, \tau(Q)^{2-i} \quad ,$$

where $K' = \exp(\beta \Delta G_{\text{intr}})$ denotes the intrinsic binding constant, respectively, and $\tau(Q) = \exp(\,\beta\, e\Psi_0(Q)\,)$ represents the Boltzmann factor from the surface potential.

The surface charge Q can be expressed in terms of the dissociated surface groups and bound metal ions (see Appendix, Eq. (5.13)). Since the amount of binding in turn depends on the association constants $K(Q)$ and thus the surface charge Q, rendering the full determination an implicit problem. The method to solve the system of equations applied for this work is explained in more detail in the Appendix.

5.5 Model Results

5.5.1 Charge Inversion and Solution pH

Fig. 5.5 summarizes the model results for a fixed protein concentration (BSA or HSA 2 mg/ml) with different salts. Both solution pH and protein charge vary faster for acidic salts - as intuitively expected and observed in the experiments (cf. Fig. 5.2, Fig. 5.3). In

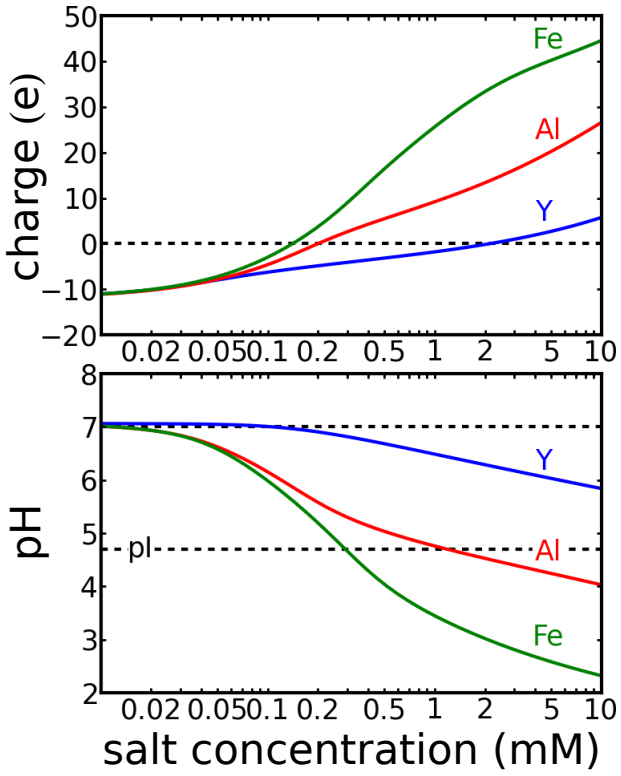

Figure 5.5: Results from the binding model for BSA / HSA (2 mg/ml) with $FeAl_3$, $AlCl_3$ and YCl_3. Top: The inversion of protein surface charge occurs at much lower salt concentrations for acidic salts. Bottom: pH in general decrease upon addition of salt. While for YCl_3 the solutions stays nearly neutral, the other two salts cause the pH to decrease below the isoelectric point pI. Note that the model results show good agreement with the experimental results (Fig. 5.2, Fig. 5.3).

particular, for $FeCl_3$, the extent of charge inversion is much larger and the point of zero charge occurs at much lower salt concentration compared to YCl_3. While for YCl_3, the pH stays rather close to neutral within the experimental concentration range, it decreases below the isoelectric point, pI, for both other salts at rather low salt concentrations. Note that in both cases charge inversion occurs clearly before the pH value reaches pI.

When compared to the experiments, the qualitative agreement is very good. Even quantitatively, several quantities such as the point of zero charge and the salt concentration at the pI do not differ too much, indicating an appropriate choice for the model parameters.

5.5.2 Phase Behavior: Reentrant Condensation from a DLVO Picture

The phase behavior can be estimated by calculating the stability from a DLVO-like picture (Derjaguin and Landau, 1941; Verwey and Overbeek, 1948) by using the surface charge as evaluated from the association model for the repulsive screened Coulomb interaction. We choose a Hamaker constant of $A = 3k_BT$ for the attractive van der Waals interaction (Heinen et al., 2012). The energy barrier is calculated as the difference between the

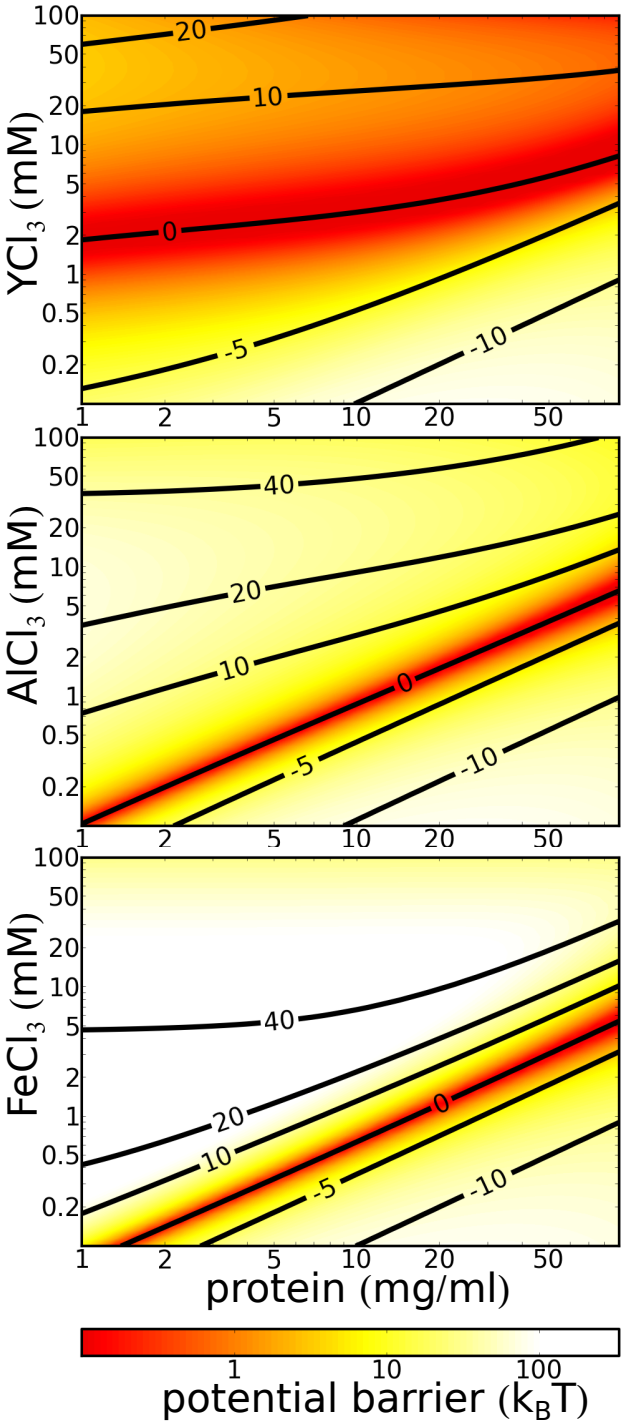

Figure 5.6: Stability barriers (color-coded in units of k_BT) and protein charges (profile lines with indicated absolute charges in e) for the full parameter space of the different metal ions. In general, a reentrant phase behavior is observed with very similar topology as the experimental results (cf. Fig. 5.1). For acidic salts like $FeCl_3$, the charge inversion occurs at much lower salt concentration and within a smaller concentration range than for rather neutral salts like YCl_3.

maximum of the potential and the potential at equilibrium distance (Roosen-Runge et al., 2010), $r_{\mathrm{eq}} = 2\left(\frac{\pi}{\sqrt{18}}\frac{3}{4\pi n}\right)^{1/3} = \left(\frac{\sqrt{2}}{n}\right)^{1/3}$ with the protein number density n. If no maximum exists, the barrier is set to zero, representing phase regimes without stabilization against aggregation.

Fig. 5.6 presents the results in color coded form (in units of k_BT). In addition, profile lines for the surface charge are given (in units of elementary charges e). The general shape of the phase diagram (Fig. 5.6) resembles obviously the experimental phase diagram (Fig. 5.1). For all salts and protein concentrations, a charge inversion is found, which causes a reentrant charge stabilization and thus reentrant condensation. Furthermore, the phase diagram of acidic salts like $AlCl_3$ and $FeCl_3$ differs clearly from those found for YCl_3. As in the experimental findings, the phase-separated regime extends over a much larger salt concentration range for the case of YCl_3.

5.6 Discussion

5.6.1 Mechanism of Protein Charge Inversion

The experimental data for charge inversion have been explained via a physicochemical model. Interaction of multivalent metal ions with amino acid side chains occurs mostly in a very specific manner. Primarily responsible for this specificity is the metal ion chelation by carboxylate groups of Asp and Glu. There is clear evidence for this interaction mechanism in one of the systems studied here. A recently published (Zhang et al., 2011) crystal structure of BLG in the presence of Y^{3+} reveals four well-defined binding sites for Y^{3+} cations. The cations coordinate in multidentate geometries with up to four carboxylate groups, thereby acting as cross-linkers between the protein molecules.

The reentrant condensation of proteins could be qualitatively described using Monte Carlo simulations for the binding of cations to acidic amino acids exposed to the solvent (Zhang et al., 2008, 2010). A recent study investigated the molecular mechanism of charge inversion in solutions of tetra-aspartate, representing a model "protein" in the sense of exposing carboxylic acids to the solvent (Kubíčková et al., 2012). Mono- and bidentate association geometries of cations at carboxylate groups were found, indicating ion-carboxyl pairing to be the underlying mechanism. Ion-protein interaction was found to be dominated by electrostatic attraction between cations and the carboxyl group, while ion specific chemical contributions seemed less important (Kubíčková et al., 2012).

Protonation equilibria of charged amino acid side chains are the second factor contributing to charge inversion. Variations of the pH cause considerable contributions to the surface charge. This effect is particularly pronounced in the case of metal ions like Fe^{3+} and Al^{3+}, which cause a significant pH shift by hydrolysis themselves. Although a generic effect of proteins, the extent is dependent on the specific protein and its structure, in particular the number of acidic and basic residues. For unbuffered solutions, the specific combination of protein and salt thus can have an important effect on the results. A recent

simulation study based on a coarse-grained protein model found a dependence of the isoelectric point on monovalent ion concentration (Kurut and Lund, 2013), supporting the notion that salt binding in combination with charge regulation governs the protein surface charge.

5.6.2 Relation of Charge Inversion to Reentrant Condensation

Reentrant condensation has been suggested to be a universal phenomenon in solutions of several globular proteins with negative native charge in the presence of multivalent cations (Zhang et al., 2010). From a naïve point of view, charge inversion represents the essential mechanism allowing residual attraction to cause phase separation, once the Coulombic repulsion vanishes close to zero protein net charge.

Recent results inspire a more refined relation of reentrant condensation and charge inversion.

For BLG with divalent salts the solutions do not clear up completely for high salt concentrations, although a significant charge inversion has occurred. Similarly, reentrant condensation is found to be incompletely established for solutions of glucose isomerase in the presence of Y^{3+}, whereas a charge inversion is present. Combining these finding with the presented results of this study, charge inversion is found to be a universal phenomenon for globular proteins with negative native charge in the presence of multivalent cations, while the occurrence of the reentrant regime seems to depend on other more specific conditions.

This finding is also in accordance with the observation that charge inversion caused by titration does not induce reentrant condensation. For both HSA and BSA, solutions remain stable when reducing the pH way below the isoelectric point, even when passing a conformational transition (normal to fast form) (Peters, 1996; Dockal et al., 2000; Barbosa et al., 2010).

The protein charge controls the Coulombic repulsion between the protein molecules. For the case of salt-induced charge inversion, the molecules repel each other at low and high salt concentrations, whereas no Coulombic repulsion is present at the point of zero charge. The difference in phase behavior, i.e. the occurrence of the reentrant regime, mainly arises from the strength and nature of attraction in the system. Since several proteins without multivalent ions do not phase separate at the isoelectric point, the binding of multivalent cations does not only cause charge inversion, but also causes additional attraction between the proteins. Possible mechanism for this attraction could be cation-cross-links between two protein molecules (Zhang et al., 2011) or attractive patches due to charge and hydrophobicity patterns.

These findings also imply another hypothetical mechanism besides the pH effect for the narrowing of the condensation regime for $AlCl_3$ and $FeCl_3$ compared to YCl_3. For acidic salts, the variation of pH causes a significant part of the charge inversion due to side chain protonation. The number of bound cations at the point of zero overall charge

is decreased, implying that ion-induced attractions like cross-linking or patch formation are less established. Less attraction in turn means a narrower range around the point of zero charge, where repulsion is small enough to allow sensitive phase separation and aggregation. Thus, variation of the pH might not only influence the charge regulation, but could also be a control parameter to further fine tune the protein interactions, both attractive and repulsive, in order to allow control of phase behavior and nucleation conditions (Zhang et al., 2012c).

5.7 Conclusions

Using electrophoretic light scattering, charge inversion of globular proteins with negative native charge in the presence of multivalent metal cations has been shown to be a common phenomenon. Effects of the pH arising from the hydrolysis of different metal salts induce significant changes for charge inversion and reentrant phase behavior. An analytical model based on chemical association reactions and metal hydrolysis provides good qualitative and even reasonable quantitative agreement with the experimental data, supporting the view of general features governing the charge inversion in protein solutions. Both local pairing of ions and carboxyl groups and protonation represent likely mechanisms dominating the charge inversion. Thus, in contrast to charge inversion in other biological molecules such as DNA, protein charge inversion should mainly be considered as a phenomenon driven by ion-protein pairing. The actual extent of charge inversion is determined by the specific combination of the protein and added salt, rendering a detailed prediction rather difficult.

For the phenomenon of reentrant condensation, the picture becomes even more involved. Multivalent cations do not only compensate the protein charge, but also provide attractive interaction patches between proteins. These attractive sites can be represented by cross-links of multivalent cations as found in protein crystals as well as by varied surface patterns of charges and hydrophobicity. The interplay of cation binding and pH, thus, not only governs the protein surface charge, but also governs at least part of the protein interactions. The control of pH in combination with multivalent cations is suggested as a methodological improvement to control phase and nucleation behavior of protein solutions. In a broader context, the interplay of pH and ion association is relevant for a comprehensive understanding of biological and soft matter systems in aqueous electrolyte solutions.

5.8 Appendix: Model for Ion Association and pH Effects: Calculation Method

In this appendix we present the model combining protein-ion interactions and protolysis equilibria. Starting from conservation laws for the number densities of all involved components, we transform the system of non-linear equation to a polynomial in $[H^+]$, which has been found to have only one physical solution. The charge effect is implemented by a numeric bisection approach with respect to the surface charge.

5.8.1 Conservation Laws

The reactions in Eq. (5.7–5.10) imply conservation laws for the total number density of all components:

$$\begin{aligned} m &= \sum_n [\mathrm{M(OH)}_n^{3-n}] + \sum_n [\mathrm{M(OH)}_n\mathrm{A}^{2-n}] \\ a &= [\mathrm{A}^-] + [\mathrm{HA}] + \sum_n [\mathrm{M(OH)}_n\mathrm{A}^{2-n}] \\ d &= [\mathrm{D}^-] + [\mathrm{HD}] \\ b &= [\mathrm{B}^+] + [\mathrm{B(OH)}] \,. \end{aligned}$$

For simplicity of notation, we use m for the total metal ion concentration (bound and free), b for the total basic residue concentration and a and d for the acidic residue concentrations (ion binding and inert). The charge neutrality of the system implies a fifth equation incorporating all charged species:

$$\begin{aligned} 0 &= \sum_n \left((3-n)[\mathrm{M(OH)}_n^{3-n}] + (2-n)[\mathrm{M(OH)}_n\mathrm{A}^{2-n}]\right) \cdots \\ &\quad -3m + a - [\mathrm{A}^-] + [\mathrm{B}^+] - b + d - [\mathrm{D}^-] + [\mathrm{H}^+] - [\mathrm{OH}^-] \,, \end{aligned}$$

including monovalent ions from the dissociation of the salt ($3m$) and the amino acids of the protein (a, d, b), which are assumed not to have a significant influence on the initial pH at zero salt concentration.

From mass action we obtain the following polynomials in $[H^+]$, which summarize the total concentrations of different species occuring in the reactions Eq. (5.7–5.10):

$$\begin{aligned} \alpha([\mathrm{H}^+], K_i) &= \sum_{j=0}^{3} [\mathrm{H}^+]^j \prod_{n=3-j}^{2} K_n \\ \beta([\mathrm{H}^+], K_i, L_i) &= \sum_{j=0}^{3} L_{3-j} [\mathrm{H}^+]^j \prod_{n=3-j}^{2} K_n \\ \gamma([\mathrm{H}^+], K_a) &= 1 + [\mathrm{H}^+] K_a \end{aligned}$$

$$\begin{aligned}
\delta([\mathrm{H}^+], K_b) &= 1 + [\mathrm{H}^+]K_b \\
\epsilon([\mathrm{H}^+], K_i, L_i) &= \sum_{j=0}^{3}(j-1)L_{3-j}[\mathrm{H}^+]^j \prod_{n=3-j}^{2} K_n \\
\phi([\mathrm{H}^+], K_i) &= \sum_{j=0}^{3} j\,[\mathrm{H}^+]^j \prod_{n=3-j}^{2} K_n \\
\eta(v[\mathrm{H}^+], K_0) &= (K_0[\mathrm{H}^+])
\end{aligned}$$

Using these polynomials, the conservation laws can be rewritten in matrix form and the equation system can be transformed into a set of four linear and one non-linear equations:

$$\begin{pmatrix} m \\ a \\ d \\ b \\ h' \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 & 0 & \beta \\ 0 & \gamma & 0 & 0 & \beta \\ 0 & 0 & \gamma & 0 & 0 \\ 0 & 0 & 0 & \delta & 0 \\ -\eta\phi & \eta & \eta & (1-\delta)\eta & -\epsilon\eta \end{pmatrix} \cdot \begin{pmatrix} [\mathrm{M(OH)_3}] \\ [\mathrm{A^-}] \\ [\mathrm{D^-}] \\ [\mathrm{B(OH)}] \\ [\mathrm{A^-}][\mathrm{M(OH)_3}] \end{pmatrix},$$

$$\chi \begin{pmatrix} [\mathrm{M(OH)_3}] \\ [\mathrm{A^-}] \\ [\mathrm{D^-}] \\ [\mathrm{B(OH)}] \\ [\mathrm{A^-}][\mathrm{M(OH)_3}] \end{pmatrix} =$$

$$\begin{pmatrix} \gamma\delta\eta(\beta+\epsilon\gamma) & -\beta\gamma\delta\eta & -\beta\gamma\delta\eta & -\beta(1-\delta)\eta\gamma^2 & \beta\delta\gamma^2 \\ \beta\gamma\delta\eta\phi & \gamma\delta\eta(\alpha\epsilon-\beta\phi) & -\alpha\beta\delta\eta & -\alpha\beta\gamma(1-\delta)\eta & \alpha\beta\gamma\delta \\ 0 & 0 & \delta\eta(\alpha(\beta+\epsilon\gamma)-\beta\gamma\phi) & 0 & 0 \\ 0 & 0 & 0 & \gamma\eta(\alpha(\beta+\epsilon\gamma)-\beta\gamma\phi) & 0 \\ -\delta\eta\gamma^2\phi & \alpha\gamma\delta\eta & \alpha\delta\eta\gamma & \alpha(1-\delta)\eta\gamma^2 & -\alpha\delta\gamma^2 \end{pmatrix} \cdot \begin{pmatrix} m \\ a \\ d \\ b \\ h' \end{pmatrix}$$

where we use the short-hand notations $h' := -1 + ([\mathrm{H}^+] - 3m + (a+d) - b)\eta$ and $\chi = \gamma\delta\eta(\alpha\beta + \alpha\gamma\epsilon - \beta\gamma\phi)$. We used *Wolfram Mathematica® 8* to perform this task.

5.8.2 Derivation of Polynomial in $[\mathrm{H}^+]$ and Physical Solution

Equating the product of the first and second row with the fifth row, we obtain a polynomial in $[\mathrm{H}^+]$:

$$\begin{aligned}
0 \;&=\; \left(\chi\,[\mathrm{A^-}]\right)_{\mathrm{row2}} \cdot \left(\chi\,[\mathrm{M(OH)_3}]\right)_{\mathrm{row1}} \\
&\qquad -\chi\,\left(\chi\,[\mathrm{A^-}][\mathrm{M(OH)_3}]\right)_{\mathrm{row5}} \\
&=:\; f_{(m,a,b,d)}\left([\mathrm{H}^+]\right)\,.
\end{aligned}$$

Any solution of the system of equations has to be a root of $f_{(m,a,b,d)}\left([\mathrm{H}^+]\right)$. Given a value of $[\mathrm{H}^+]$, all other concentrations of reaction species can be calculated. The physical solution is obtained by testing straight-forward physical conditions for the mathematical roots $[\mathrm{H}^+]$, returning under all studied conditions a unique solution:

(i) $[\mathrm{H}^+] \in \mathbb{R}^+$

(ii) $m > [\mathrm{M(OH)_3}]_{[\mathrm{H}^+],m,a,b,d} > 0$

(iii) $a > [\mathrm{A}^-]_{[\mathrm{H}^+],m,a,b,d} > 0$

(iv) $b > [\mathrm{B(OH)}]_{[\mathrm{H}^+],m,a,b,d} > 0$

(v) $d > [\mathrm{D}^-]_{[\mathrm{H}^+],m,a,b,d} > 0$

5.8.3 Surface Charge Effect on Association and Implicit Relation

As pointed out in the main text, surface charge influences the association constants and is in turn determined by the amount of association. Using the solution for $[\mathrm{H}^+]$ from the previous subsection, the surface charge can be calculated for a given set of association constants K_a, K_b, K_w, K_i, L_i:

$$\begin{aligned} Q &= [\mathrm{B}^+] - [\mathrm{A}^-] - [\mathrm{D}^-] + \sum_n (2-n) \cdot [\mathrm{M(OH)}_n\mathrm{A}^{2-n}] \\ &=: g([\mathrm{H}^+], m, a, b, d) \quad . \end{aligned} \tag{5.13}$$

In order to solve this implicit problem, we performed a numerical bisection with respect to Q for the function $g\left([\mathrm{H}^+], m, a, b, d\right)\big|_Q - Q$ in the interval $\left[0,\, g\left([\mathrm{H}^+], m, a, b, d\right)\big|_{Q=0}\right]$ using the routine supplied by the Python package *scipy.optimize* (Jones et al., 2001; Oliphant, 2007) (version 0.7.0).

Acknowledgements

The authors acknowledge financial support by the Deutsche Forschungsgemeinschaft (DFG). FRR acknowledges a fellowship by the Studienstiftung des Deutschen Volkes. The authors benefited from discussions with M. Hennig (ILL, Grenoble, France), O. Lenz, C. Holm (University of Stuttgart, Germany), M. Antalik (Slovak Academy of Sciences, Kosice), and P. Jungwirth (Academy of Sciences of the Czech Republic, Prague).

Chapter 6

Competing Salt Effects on Reentrant Phase Behavior of Protein Solutions: Tailoring of Interaction Strength and Range by the Binding of Multivalent Ions and Charge Screening

Chapter 6 is based on the following publication:

Competing Salt Effects on Reentrant Phase Behavior of Protein Solutions: Tailoring of Interaction Strength and Range by the Binding of Multivalent Ions and Charge Screening

Elena Jordan[†], Felix Roosen-Runge[†*], Sara Leibfarth, Fajun Zhang[*], Michael Sztucki, Andreas Hildebrandt, Oliver Kohlbacher and Frank Schreiber

[†] E. J. and F. R.-R. contributed equally to this paper. [*] corresponding authors

submitted and under review

Contributions:

Research design	FZ, FS
Experiments	EJ, FZ
Technical assistance	MS
Simulations and model calculations	FRR, SL
Data analysis and interpretation	EJ, FRR, SL
Paper writing	EJ, FRR, FZ, FS

Jordan et al. (2013)

The phase behavior of protein solutions is affected by additives such as crowder molecules or salts. In particular, upon addition of multivalent counterions, a reentrant

condensation can occur, i.e. protein solutions are stable for low and high multivalent ion concentrations, but aggregating at intermediate salt concentrations. The addition of monovalent ions shifts the phase boundaries to higher multivalent ion concentrations. This effect is found to be reflected in the protein interactions, as accessed via small-angle X-ray scattering. Two simulation schemes—a Monte Carlo sampling of the counterion binding configurations using the detailed protein structure and an analytical coarse-grained binding model—reproduce the shifts of the experimental phase boundaries. The results support a consistent picture of the protein interactions responsible for the phase behavior. The repulsive Coulomb interaction is varied by the binding of multivalent counterions and additionally screened by any increase of the ionic strength. The attractive interaction is induced by the binding of multivalent ions, most likely due to ion bridging between protein molecules. The overall picture of these competing interactions provides interesting insight into possible mechanisms for tailoring interactions in solutions via salt effects.

6.1 Introduction

Salts are ubiquitous in nature. Under realistic physiological conditions, multiple types of salts are present and have profound effects on cellular processes. The interaction of salt ions with macromolecules such as proteins or DNA is a long-studied field with important open challenges. Starting with the Hofmeister series for ion-dependent effects of protein solubility (Hofmeister, 1888; Collins, 2004; Curtis et al., 1998), the interaction of ions with aqueous surfaces shows interesting and manifold manifestations. The most basic one is the formation of the electrostatic double-layer as predicted by the Poisson-Boltzmann theory (Das et al., 2003), causing a screening of charges in electrolyte solutions. Ion condensation at the surface has been found for highly charged surfaces (Manning, 1978). For multivalent salts, even a surface charge inversion due to condensed ions has been found caused by ion-ion correlations (Deserno et al., 2000; Moreira and Netz, 2000; Grosberg et al., 2002; Besteman et al., 2004; Messina, 2009). A similar charge inversion can also be observed in systems where ions bind to specific surface sites (Lyklema, 2006; Jungwirth and Winter, 2008; Lund et al., 2008; Kalcher et al., 2009; Kubíčková et al., 2012), such as metal ions at functional groups of protein side chains (Harding, 2001; Permyakov, 2009).

For mixtures of mono- and multivalent salts, the picture of the ion–surface interaction becomes still more complicated. For the distributions of mono- and divalent counterions around DNA, competing effects between the two types of salt were found, at the same time supporting and challenging the Poisson-Boltzmann theory for mixtures of ion types (Andresen et al., 2004; Das et al., 2003; Burak et al., 2004). Regarding the effective charge inversion due to ion-ion correlations, theoretical approaches predict a so-called "giant overcharging" (Nguyen et al., 2000a), while simulations suggest that monovalent salt is lowering the effective reversed charge (Lenz and Holm, 2008). For the binding of

multivalent ions to functional groups such as carboxylates in amino acids, the equilibrium constants are decreased slightly upon addition of monovalent salt (Sandhu, 1977). In real systems such as protein solutions, these effects all occur simultaneously, rendering a complete description impossible.

A better description of the interaction between ions and macromolecules is important for the detailed understanding of macromolecular interactions and the related phase behavior. For DNA, the ion condensation and a related reentrant condensation of DNA macromolecules has been studied intensively (Saminathan et al., 1999; Nguyen et al., 2000b; Burak et al., 2004). Reentrant condensation means that solutions are clear for low and high salt concentrations, indicating stable solutions. At intermediate multivalent salt concentration between the transition concentrations c^* and c^{**}, the solutions become turbid. It was found experimentally (Saminathan et al., 1999) as well as theoretically (Nguyen et al., 2000b; Burak et al., 2004) that c^* increases with the addition of monovalent salt, suggesting that condensation of multivalent ions is weakened with increased ionic strength.

In the context of phase behavior of biological solutions, proteins represent a challenging system. Control on the phase behavior of protein solutions is a major goal of protein science, since it allows more efficient biotechnological methods for e.g. protein crystallization and opens possibilities for biomedical applications and food industry including issues related to, e.g., shelf life. Proteins have a non-spherical and rough shape with inhomogeneous patterns of surface charge and hydrophobicity. Thus, a variety of anisotropic forces can act between proteins, and theoretical results for model systems like spheres or rods can only be applied with care.

In this article, we study the role of monovalent (NaCl) and multivalent (YCl_3) salts in solutions for the phase behavior of the globular protein human serum albumin (HSA). Recently it was found that multivalent ions induce a reentrant condensation in protein solutions (Zhang et al., 2008, 2010; Roosen-Runge et al., 2013a). This reentrant behavior can be understood by an inversion of the protein surface charge due to binding of Y^{3+} ions to surface groups. This charge inversion has been found as a universal behavior in protein solutions in the presence of multivalent metal counterions (Zhang et al., 2010; Roosen-Runge et al., 2013a). Both Monte Carlo sampling of the binding configurations (Zhang et al., 2008) and an analytical coarse-grained model (Roosen-Runge et al., 2013a) reproduce the charge inversion with good qualitative and even reasonable quantitative agreement. Here we study the yet unexplored effect of monovalent salt on the charge inversion and reentrant phase behavior induced by multivalent salt to elucidate the underlying protein interactions.

6.2 Experimental

6.2.1 Materials

De-fatted Human Serum Albumin (HSA) was purchased from Sigma-Aldrich (A 9511, essentially globulin-free). YCl_3 was purchased from Sigma Aldrich as anhydrous powder with a high-purity grade (99.99%, no. 451363). NaCl (>99.5%) was purchased from Merck. Degassed Milli-Q water was used to prepare all solutions. All solutions were prepared and studied at room temperature (295 ± 2 K).

Stock solutions for the protein (HSA 200 mg/mL) and the salts (salt concentration 100 mM) were prepared and then diluted in order to obtain the desired concentrations. For SAXS measurements, the samples were centrifuged at 10,000 rpm. The protein concentrations of the samples were determined by UV-visible spectroscopy using the absorption at 280 nm with an extinction coefficient for HSA of 0.531 L/g. The pH of the solution was found to be close to 7 and did not vary significantly between different protein and salt concentrations. Note that this pH stability observed for the HSA/YCl_3/NaCl system does not necessarily apply to other proteins and salts. Strong pH shifts are observed in systems where the metal ions undergo hydrolysis (e.g. $FeCl_3$, $AlCl_3$) (Roosen-Runge et al., 2013a).

6.2.2 The Protein Phase Diagram as a Function of Protein and Multivalent Salt Concentration

In order to determine changes in the phase behavior qualitatively, the turbidity of the samples was determined by visual inspection 3–5 min after the preparation, when the opacity was constant. Simultaneously, the absorption was determined with UV-visible spectroscopy. Spectra were recorded in the wavelength range from 400–800 nm, where the protein shows no characteristic absorption. The transmitted intensity was integrated over the spectral range. The absorption increases significantly once the samples become turbid. The results from both methods are consistent.

In order to determine the transition concentrations, series of samples with the same protein concentration and various YCl_3 concentrations close to c_Y^* and c_Y^{**} were prepared. The highest YCl_3 concentration at which the sample is still clear, c_{Y1}, and the lowest YCl_3 concentration at which the sample is turbid, c_{Y2}, were determined. The first transition concentration c_Y^* with an inaccuracy of δc_Y^* was calculated as

$$c_Y^* = \frac{c_{\mathrm{Y1}} + c_{\mathrm{Y2}}}{2} \quad , \quad \delta c_Y^* = \frac{c_{\mathrm{Y1}} - c_{\mathrm{Y2}}}{2} \quad . \tag{6.1}$$

c_Y^{**} is determined analogously.

6.2.3 Small-Angle X-ray Scattering (SAXS)

Experiments

SAXS measurements were performed at the European Synchrotron Radiation Facility (ESRF) in Grenoble, France, at the beamline ID02 with a sample-to-detector distance of 2 m. The energy of the incoming beam was 16.062 keV (wavelength 0.078 nm). A q-range from 0.072 nm^{-1} to 4.5 nm^{-1} was covered. The SAXS detector was a fiber optically coupled fast-readout low-noise (FReLoN) CCD based on a Kodak KAF-4320 image sensor in an evacuated flight tube. About 100 μl sample solution was filled into a flow-through quartz capillary. The sample in the scattering volume was exchanged for every exposure. For each sample, 10 exposures of 0.3 s each were measured. The 2D data were normalized to an absolute scale and azimuthally averaged to obtain the intensity profiles, and the solvent background was subtracted. More detailed information on data reduction and q-resolution calibration can be found in the literature (Narayanan, 2008).

Data Analysis

Small-angle X-ray scattering data can be used to obtain information on the pair interaction potential (Glatter and Kratky, 1982; Svergun and Koch, 2003; Hansen and McDonald, 2006). The scattering intensity, $I(q)$, for a polydisperse or a non-spherical system, can be calculated on the basis of approximation approaches such as the "decoupling approximation" and the "average structure factor" approximation (Pedersen, 1997; Chen and Lin, 1987). Both approaches assume that the particle position is not correlated with its orientation. For the case of non-spherical but monodisperse solutes at a low to intermediate concentration, such as the studied protein solutions, both approaches give comparable results (Zhang et al., 2007). Therefore, in this work, the scattering intensity is calculated using the average structure factor approximation, which can be expressed by

$$I(q) = N(\Delta\rho)^2 V^2 P(q)\bar{S}(q) \tag{6.2}$$

where $q = \frac{4\pi}{\lambda}\sin(2\Theta/2)$ is the scattering vector, 2Θ is the scattering angle, N is the number of protein molecules per unit volume in the solution, $\Delta\rho = \rho_\text{p} - \rho_\text{s}$ is the scattering length density difference between the solvent and the solute, and V is the volume of a single protein. $P(q)$ is the form factor of a given protein, i.e., the scattering from a single protein molecule after orientation averaging. An ellipsoid form factor with semi-axes a and b is used to model HSA (Feigin and Svergun, 1987):

$$P(q) = \int_0^1 dx \left| \frac{3(\sin u - u\cos u)}{u^3} \right|^2, \tag{6.3}$$

$$u = qb[(a/b)^2 x^2 + (1-x)^2]^{1/2}. \tag{6.4}$$

Using the average structure factor approximation, $\bar{S}(q)$ is calculated from a monodisperse structure factor with an effective sphere diameter. In our case, the protein solution is a monodisperse but non-spherical (ellipsoidal) system. The effective sphere diameter is calculated from a virtual sphere with the same second virial coefficient as the ellipsoid (Isihara, 1950; Zhang et al., 2012b). In the following parts and for simplicity, we use $S(q)$ to denote $\bar{S}(q)$.

In solutions with monovalent salt and at smaller salt concentrations, the inter-particle potential can be modeled reasonably well by the repulsive part of a full DLVO potential, *i.e.* a screened Coulomb potential (Nägele, 1996; Heinen et al., 2012):

$$U_{\mathrm{SC}}(r) \quad = \quad \frac{Z^2 e^2 \exp(\kappa\sigma)}{\varepsilon(1+\kappa\sigma/2)^2} \frac{\exp(-\kappa r)}{r} \qquad (\text{for } r > \sigma) \quad . \tag{6.5}$$

r is the center-to-center distance between two particles with diameter σ and charge $Q = Ze$. ε is the dielectric constant of the continuous medium.

The inverse Debye screening length κ depends on the ionic strength I,

$$\kappa = \sqrt{\frac{\varepsilon k_B T}{2 N_A e^2 I}} \quad \text{and} \quad I = \frac{1}{2} \sum_{\alpha} n_\alpha Q_\alpha^2 \, , \tag{6.6}$$

where n_α and Q_α denote the number density and charge of the ions of species α in the solution.

A solution for the structure factor using the rescaled mean-spherical approximation (RMSA) (Hayter and Penfold, 1981; Hansen and Hayter, 1982) was used for the fits (Kline, 2006).

6.3 Theoretical modeling

6.3.1 Detailed Protein Structure: Monte Carlo Sampling of Ion Binding Configurations

Using a high-resolution structure of human serum albumin (HSA) from the Protein Data Bank (ID: 1N5U) (Wardell et al., 2002), the binding probabilities of multivalent ions to solvent-exposed carboxylate groups were calculated with a Monte Carlo approach. Results and further description of this method have been published previously (Zhang et al., 2008, 2010). An established approach for the calculation of protonation states of proteins (Bashford and Karplus, 1990) has been modified for the case of the binding of multivalent metal ions (Y^{3+}). The protonation state of an individual protein was determined with the H++ webserver (Gordon et al., 2005). All solvent-exposed carboxylate groups were regarded as potential binding sites for Y^{3+}. The electrostatic partition function of Y^{3+} binding was sampled in a Monte Carlo scheme (Zhang et al., 2008) to determine the Y^{3+} binding probabilities as a function of the Y^{3+} concentration. The simulation approach uses classical

Poisson-Boltzmann theory to predict the electrostatic potential at the Y^{3+} binding sites in the presence of water. All parameters involved in the simulation were taken from literature values (Sandhu, 1977; Wardell et al., 2002; Zhang et al., 2008). Note that the Monte Carlo sampling includes the full ion-ion correlation between bound Y^{3+} ions and accounts for the full atomic detail of HSA. These effects are expected to dominate the less important ion-ion correlations in the bulk solution, which are not accounted for in the Poisson-Boltzmann approach.

The average number of the bound Y^{3+} ions was determined at several Y^{3+} concentrations. The net charge of the protein is calculated for different fixed concentrations of monovalent salt from the bare charge of the protein and the Y^{3+} ions bound to the protein. The Y^{3+} concentration used in this approach corresponds to the concentration in the surrounding bulk solvent, thus only the free Y^{3+} ions. In order to obtain the full Y^{3+} concentration, both free ions and bound ions have to be added up:

$$c_{\mathrm{Y}} = c_{\mathrm{Y,free}} + n_{\mathrm{Y,bound}}\, c_{\mathrm{HSA}}/M_{\mathrm{HSA}}\ . \tag{6.7}$$

$n_{\mathrm{Y,bound}}$ is the average number of Y^{3+} ions bound to one protein, c_{HSA} is the protein concentration in mg/ml and $M_{\mathrm{HSA}} = 66500\,\mathrm{Da}$ is the molecular weight of HSA.

The addition of monovalent salt (NaCl) is included by a change of the Debye screening length as well as a small shift of the equilibrium constant for the binding of multivalent ions to carboxylate groups (Sandhu, 1977). The full simulation setup has been repeated for each monovalent salt concentration, respectively.

6.3.2 Coarse-Grained Model: Analytical Calculation Scheme for Protein Total Charge

An analytical model for the binding of multivalent ions to a spherical particle, taking into account charge regulation of surface groups, counterion binding as well as pH effects due to hydrolysis of multivalent metal ions, has been developed recently (Roosen-Runge et al., 2013a). The scheme is based on association reactions between the present ions and surface sites, the binding to which is further dependent on the total particle charge. A detailed description of the method as well as the choice of suitable model parameters solely based on literature values and theoretical estimates can be found in Ref. (Roosen-Runge et al., 2013a).

In this coarse-grained model, the addition of monovalent salt varies the surface potential via the Debye screening length κ, and thus changes the binding equilibria of multivalent ions. The analytical scheme allows an investigation of the effects of monovalent ions on the binding of multivalent ions and the protein total charge at low computational cost.

6.3.3 Estimation of Phase Boundaries in the Reentrant Phase Diagram

Two methods have been used to estimate the transition Y^{3+} concentrations, c_Y^* and c_Y^{**}, from the calculated protein total charges and the related full DLVO potential.

Charge Estimation: The transition concentrations can be estimated directly under the assumption that the transition salt concentrations, c_Y^* and c_Y^{**}, correspond to transition protein charges Q^* and Q^{**}. Ignoring charge regulation of the protein and using the HSA charge without YCl_3 of $Q_0 \approx -11e$, the transition charges can be related to specific numbers of bound Y^{3+} ions $N \approx (Q^* - Q_0)/3$. We chose $Q^* = -5e$ and $Q^{**} = 5e$, corresponding to roughly 2 and 5 bound ions. Note that the exact choice of Q^* and Q^{**} does not change the obtained trends, since the charge profile is a monotonous function in c_Y and c_{NaCl}.

Barrier Estimation: The full DLVO potential between two spherical particles with diameter σ is given by a screened Coulomb repulsion and a van der Waals term (Nägele, 1996; Likos, 2001):

$$\begin{aligned} U_{\mathrm{DLVO}}(r) &= U_{\mathrm{SC}}(r) + U_{\mathrm{vdW}}(r) \quad , \\ U_{\mathrm{vdW}}(r) &= -\frac{A}{12}\left(\frac{\sigma^2}{r^2-\sigma^2} + \frac{\sigma^2}{r^2} + 2\log\frac{r^2-\sigma^2}{r^2}\right) \quad . \end{aligned} \tag{6.8}$$

where A is the system-dependent Hamaker constant. At low ionic strength and large charges, the DLVO potential shows a maximum $U_{\mathrm{DLVO}}(r_{\max})$, representing a repulsive barrier. $U_{\mathrm{DLVO}}(r_{\max})$ can be used to estimate the propensity for aggregation. In our case, we used $A = 3\,k_BT$ as employed for a similar protein (bovine serum albumin) with realistic results (Heinen et al., 2012; Roosen-Runge et al., 2013a). As the criterion for the transition concentrations we used $U_{\mathrm{DLVO}}(r_{\max}) - U_{\mathrm{DLVO}}(r_{\mathrm{eq}}) \stackrel{!}{=} 1\,k_BT$, where $r_{\mathrm{eq}} = (\sqrt{2}/n)^{1/3}$ with the number density n is the maximized distance between the particles in the solution (Roosen-Runge et al., 2010, 2013a). This approach provided realistic results for a similar system and related issues (Roosen-Runge et al., 2013a).

6.4 Results and Discussion

6.4.1 Experimental Phase Diagram: Reentrant Condensation

Experimental phase diagrams for HSA in the presence of both monovalent and multivalent ions were determined as described in Sec. 6.2.2. Several HSA concentrations c_{HSA} between $5-50$ mg/mL and NaCl concentrations c_{NaCl} between $0-100$ mM were chosen to determine the transition concentration c_Y^* between the clear regime I and the turbid regime II and the second transition concentration c_Y^{**} between regime II and the clear reentrant regime III. Fig. 6.1 shows an example of a phase diagram for HSA with YCl_3 and a constant NaCl concentration of 20 mM. In the phase diagram the added YCl_3 concentration is plotted

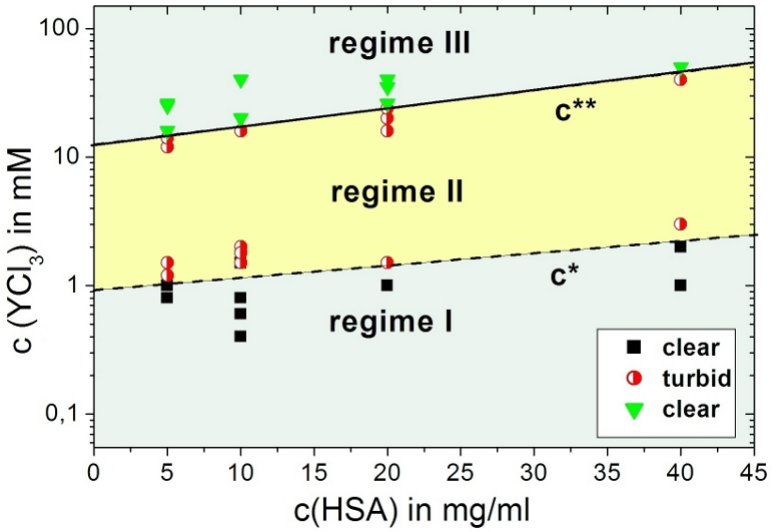

Figure 6.1: Phase diagram as determined by visual inspection for HSA with YCl_3 with a constant NaCl concentration of 20 mM. The black squares stand for samples in the solution regime I, the red circles for samples in the aggregation regime II, and the green triangles for samples in the reentrant regime III. The samples in regime I and III are clear, whereas the samples in regime II are turbid due to aggregation.

against the protein concentration in the sample. For YCl_3 concentrations lower than c_Y^* and higher than c_Y^{**} the samples are clear (solution regime I and reentrant regime III). Samples with YCl_3 concentrations between c_Y^* and c_Y^{**} are turbid (aggregation regime II). Similar diagrams were generated for several different NaCl concentrations. Both transition concentrations c_Y^* and c_Y^{**} are found to increase with increasing c_{NaCl} as depicted in Fig. 6.5(a). This finding is discussed in more detail in Sec. 6.4.4.

The results from optical observation are also supported by measurements on the amount of precipitates. A set of samples was prepared with the same protein concentration of about 45 mg/ml, NaCl concentrations of 0.0 mM, 10 mM, 20 mM, 50 mM and 80 mM and various YCl_3 concentrations between 2 mM and 40 mM. Aggregates formed after the addition of YCl_3, as predicted by the phase diagram. We centrifuged the samples to accelerate the precipitation. The protein concentration of the supernatant was monitored after centrifugation by UVvis absorption. The results for NaCl 10 mM, 50 mM and 80 mM are shown in Fig. 6.2, indicating that the amount of precipitation also follows the reentrant behavior.

The increasing trend of the transition concentrations c_Y^* and c_Y^{**} with increasing c_{NaCl} is recovered. At higher NaCl concentrations, higher YCl_3 concentrations are required to cause precipitation. A similar trend is observed for the second transition, although no complete redissolution is obtained within the probed YCl_3 concentration range for $c_{\mathrm{NaCl}} = 50$ mM and 80 mM.

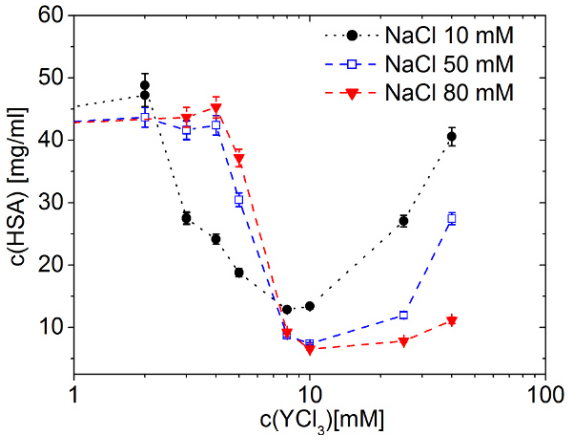

Figure 6.2: Protein concentration in the supernatant after the precipitation induced by YCl_3. The different colors refer to different NaCl concentrations (10 mM, 50 mM, 80 mM). With increasing NaCl concentration, both the increase of precipitation and redissolution occurs at higher YCl_3 concentration.

For samples with $c_{\mathrm{HSA}} \gtrsim 30$ mg/ml with YCl_3 and without NaCl, a liquid-liquid phase separation (LLPS) is found in regime II (Zhang et al., 2012c). The investigation of NaCl effects on the LLPS was beyond the scope of this study, but we speculate that the underlying protein interaction causes qualitatively similar effects and trends.

6.4.2 Variations of Protein Interaction Throughout the Reentrant Condensation: SAXS

SAXS data provide direct access to the protein interaction via liquid state theory fitting methods. Samples containing only NaCl and no YCl_3 can be fitted very well with a form factor of an oblate ellipsoid with fixed semi-axes ($a = 1.6$ nm and ($b = 4.2$ nm) and a screened Coulomb structure factor. The results for a sample series with fixed HSA concentration $c_{\mathrm{HSA}} = 62$ mg/ml, no YCl_3, and increasing NaCl concentration (10 mM $< c_{\mathrm{NaCl}} <$ 200 mM) are summarized in Fig. 6.3(a). For low NaCl concentrations, a pronounced correlation peak is observed, the height of which decreases with increasing NaCl concentration. This behavior can be explained by the screening of the protein charge and the reduction of the repulsive electrostatic interaction. The fit parameters at low salt concentration take realistic values (Tab. 6.1). With increasing NaCl concentration, the fitted volume fraction $\varphi^{(\mathrm{fit})}$ shifts to higher values, probably because the screened Coulomb structure factor becomes less accurate. The general results are consistent with results on protein interactions in charge-screened solutions of bovine serum albumin (BSA) at low to high protein concentrations (Zhang et al., 2007; Roosen-Runge et al., 2010) and results on HSA in water and buffer (Fukasawa and Sato, 2011).

Upon addition of YCl_3, the phase behavior changes qualitatively (cf. Fig. 6.1). In order to monitor the protein interactions throughout the phase diagram and their dependence on

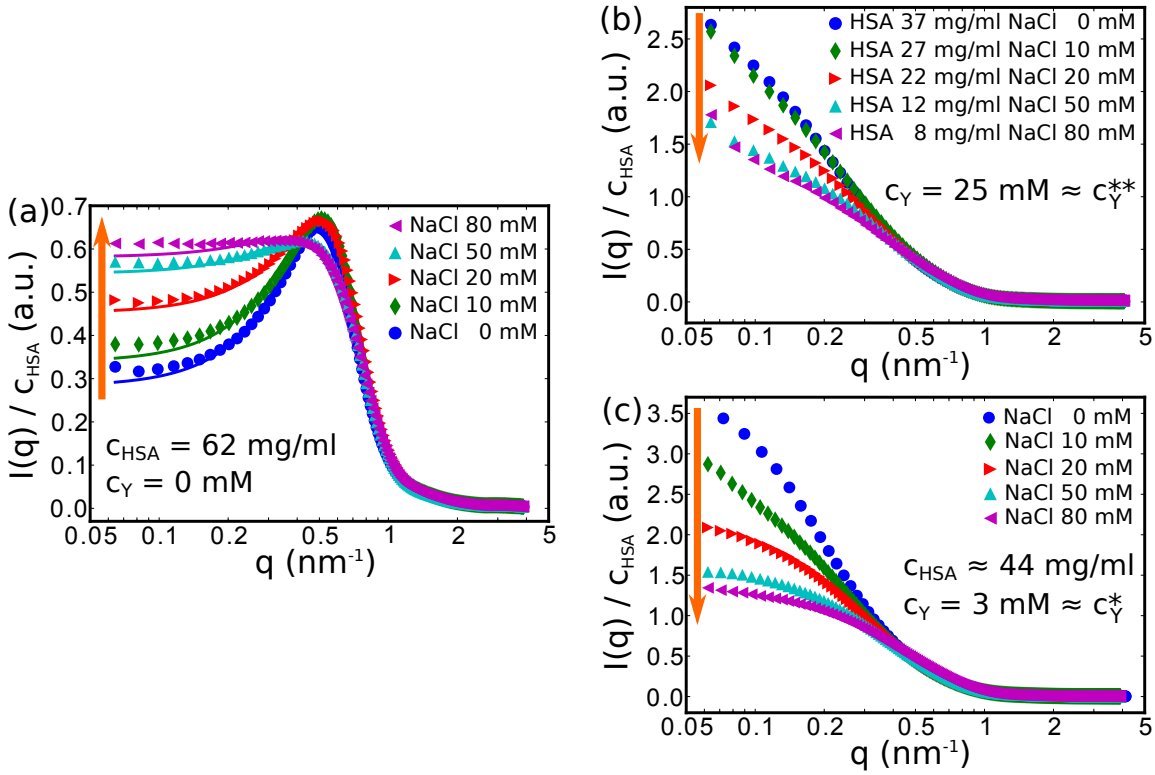

Figure 6.3: SAXS curves for HSA with different salt conditions. The orange arrows indicates the trend for increasing c_{NaCl}. For clarity, only every second data point is shown. (a) no YCl_3: With increasing c_{NaCl} the low-q intensity increases due to the reduced screened Coulomb repulsion. The intensity is fitted using an ellipsoid form factor and a screened Coulomb structure factor (fit parameters: Tab.6.1). (b) $c_Y = 25\,\mathrm{mM} \approx c_Y^{**}$: The protein concentration in the measured supernatant decreases strongly with increasing c_{NaCl}, reflecting the increasing of c_Y^{**} with c_{NaCl}. The scattering intensity in the limit of low q does not level off but increases steeply, thereby indicating the presence of larger aggregates. (c) $c_Y = 3\,\mathrm{mM} \approx c_Y^{*}$: The low-$q$ intensity decreases with increasing c_{NaCl}, corresponding to weaker attractive interaction.

NaCl concentration, SAXS intensities were recorded for HSA concentrations c_{HSA} between 5 and 50 mg/ml at YCl_3 concentrations c_{Y} of 1 to 50 mM and with NaCl concentrations c_{NaCl} of 0, 10, 20, 50 and 80 mM. Protein interactions deep in the solution regime I, i.e. at low c_{Y}, can still be modeled reasonably well with a repulsive screened Coulomb interaction. In solutions close to the aggregation regime II or in the reentrant regime III, the protein interactions cannot be described by a repulsive interaction anymore, indicating the increasing relevance of attractive interactions. In the context of this study, we are interested in the changes of attraction and repulsion. Since information on this overall interaction can be obtained by model-free means from the low-q intensity, no detailed data fitting is required.

In Fig. 6.3(c), a series with fixed HSA concentration $c_{\mathrm{HSA}} \approx 44\,\mathrm{mg/ml}$ and fixed YCl_3 concentration $c_{\mathrm{Y}} = 3\,\mathrm{mM}$ is shown, corresponding to conditions close to c_Y^*. Around c_Y^* the effect of NaCl is most pronounced, since the overall interaction is expected to change

c_{NaCl} [mM]	$\varphi^{(\mathrm{fit})}$ [%]	$Q^{(\mathrm{fit})}[e]$	$c_{\mathrm{NaCl}}^{(\mathrm{fit})}$ [mM]
10	6.6	-10.9	16
20	7.6	-10.5	21
50	7.8	-14.4	51
80	7.5	-14.1	81
200	8.1	-10.5	200

Table 6.1: Fit parameters for the curves in Fig. 6.3(a). The intensity from HSA 62 mg/ml in NaCl solutions is fitted with a screened Coulomb structure factor. The free fit parameters are shown in the table, the others were fixed. The parameters kept fixed were the semi-axes $a = 1.6$ nm and $b = 4.6$ nm of the oblate ellipsoid, and the temperature $T = 295$ K. The fit parameter for the salt concentration c_s(NaCl) was first held on the prepared value and then let free for a second fit. The relative errors of the parameters were less than 1%.

from a charge-screened repulsion to attraction. The low-q intensity increases for decreasing c_{NaCl}, indicating an increasing strength of the overall attraction. Note that the opposite behavior would be expected if only charge screening was relevant, since the screening reduces the repulsion and thus increases the overall attraction with increasing c_{NaCl}.

Fig. 6.3(b) depicts HSA solutions at fixed $c_{\mathrm{Y}} = 25\,\mathrm{mM}$, corresponding to conditions around c_Y^{**}. The solutions were prepared at a HSA concentration $c_{\mathrm{HSA}} \approx 44\,\mathrm{mg/ml}$ for a series of NaCl concentrations c_{NaCl}. Before measurement, the samples were centrifuged to remove large aggregates. With increasing c_{NaCl}, the amount of precipitated aggregates increases strongly, reflecting the trend of increasing c_Y^{**} and leaving a smaller protein concentrations in the measured supernatant. The scattering intensity around c_Y^{**} does not level off on the given q range, suggesting that larger aggregates are still present in the supernatant, or forming again. The decreasing scattering intensity with increasing c_{NaCl} is thus related to the internal structure of larger aggregates and no good indicator for varying protein interactions. The results will be exploited and discussed in more detail in Sec. 6.5.

6.4.3 Protein Charge Inversion as a Function of Ionic Strength

Measurements of the electrophoretic mobility as well as simulations show an inversion of the protein net charge, i.e. the total charge of the protein-ion complex, in the presence of multivalent metal ions (Zhang et al., 2008, 2010; Roosen-Runge et al., 2013a). The results from both simulation schemes—the Monte Carlo sampling of binding configurations and the analytical coarse-grained model—for the protein net charge Q as a function of the Y^{3+} concentration are shown in Fig. 6.4. All charge curves show an inversion from negative to positive net charge in the simulated range of Y^{3+} concentration, which is in qualitative agreement with experiments. Moreover, the simulated curves all meet at one point close to zero net charge for both approaches, respectively. This simple behavior originates from the vanishing monopole moment of the protein in this regime.

Both simulation schemes show similar trends with increasing monovalent salt concentration. In the regime of negative net charge a higher NaCl concentration leads to a lower

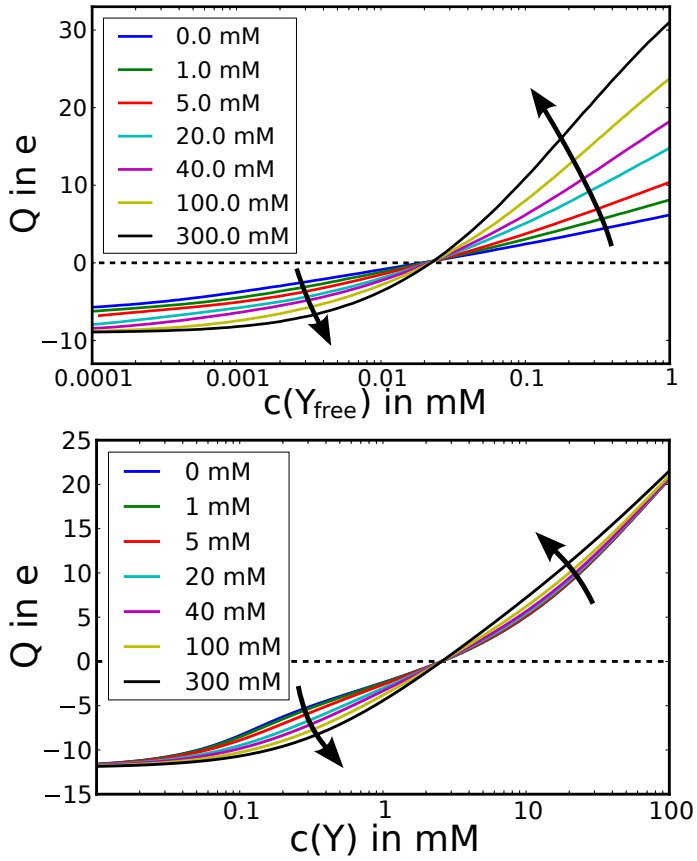

Figure 6.4: Charges of the protein-ion complex as a function of Y^{3+} concentration and for increasing NaCl concentration: (top) Net charges from Monte Carlo sampling of binding configurations. (bottom) Net charges from analytical coarse-grained model for a protein concentration of 5 mg/ml. The black arrows indicate the trends for increasing monovalent salt concentration. Note that the MC sampling method (top) is based on a single protein molecule, thus no absolute protein concentration is provided and the salt concentration $c_{\mathrm{Y,free}}$ corresponds to the concentration of free ions in the solvent. The coarse-grained method uses a real salt concentration c_{Y} incorporating free and bound ions.

number of bound Y^{3+} ions. This binding behavior is reasonable because the favorable electrostatic interaction between the Y^{3+} ions and the negatively charged protein is weakened by the screening effect of the salt. In the regime of positive net charge the number of bound Y^{3+} ions increases with increasing NaCl concentration, since the electrostatic surface potential due to the positive protein net charge is decreased and thus allows more Y^{3+} ions to bind to the surface driven by the NaCl-independent binding enthalpy.

The differences between the simulations can be explained by the different nature of the schemes. The differing Q spread at low salt concentration might be due to the fact that the coarse-grained model takes into account only one effective binding site, while the Monte Carlo scheme calculates the binding equilibrium for each site individually and, thus, incorporates a broader distribution of equilibrium constants. The qualitative difference at high concentrations—divergence for Monte Carlo scheme and convergence

for the coarse-grained model—arises from the significantly differing absolute values of the salt concentrations in the charge-inverted regime. This effect might be caused by the fact that the two schemes account for the multivalent ions in fundamentally different ways and the binding parameters are determined differently. Furthermore, the Monte Carlo scheme accounts explicitly also for interactions between the bound ions, while the coarse-grained model treats all bound ions in a mean-field picture. Finally, pH effects and charge regulation varying with multivalent salt concentration are only accounted for in the coarse-grained model.

The general consistency of the observed trends despite the difference in the simulation schemes suggests that these trends are rather robust properties of a protein in a solution containing multivalent ions. Based on these charge curves, effects on the phase behavior are estimated in the following section.

6.4.4 Dependence of Transition Concentrations c_Y^* and c_Y^{**} on Ionic Strength

The results on the dependence of the transition concentrations c_Y^* and c_Y^{**} of HSA solutions on monovalent salt are summarized in Fig. 6.5. From the experimental results, a clear increase of the two transition concentrations is observed with increasing monovalent salt concentration. This trend is observed for all protein concentrations probed in the range between 5–50 mg/ml (Fig. 6.5(a)). For samples with $c_{\mathrm{NaCl}} \geq 100\,\mathrm{mM}$, no reentrant phase behavior is observed, and even samples at very high $c_{\mathrm{Y}} \geq 100$ mM are still turbid.

For the theoretical methods, the results differ between the estimation methods and simulation schemes. While c_Y^* from the charge estimation shows a clear increase for the full monovalent salt range (Fig. 6.5(b+c))—and thus agrees with the experimental observations (Fig. 6.5(a))—, c_Y^* from the barrier estimation decreases for higher ionic strengths (Fig. 6.5(d+e)). In contrast, for c_Y^{**} only the results from the barrier estimation show the observed general increase, while the charge estimation misses the trends at least for smaller protein concentrations.

6.5 Discussion: Nature of the Protein Interactions Driving the Reentrant Phase Behavior

From the comparison of the different results from experimental phase diagrams, SAXS, and simulations, a consistent picture of the underlying protein interactions can be drawn for both the first transition c_Y^* and the second transition c_Y^{**}. The general picture involves charge-stabilized solutions at low and high c_{Y}, i.e. repulsive Coulomb interaction due to considerable protein net charges, while in the aggregation regime II an overall attractive interaction dominates. The competition of the attractive and repulsive interaction can be exploited by the addition of NaCl, thereby fine-tuning the interactions in both interaction strength and repulsion range. The occuring changes in protein interaction are schematically depicted in Fig. 6.6 and will be discussed in the following for the two transitions. Starting

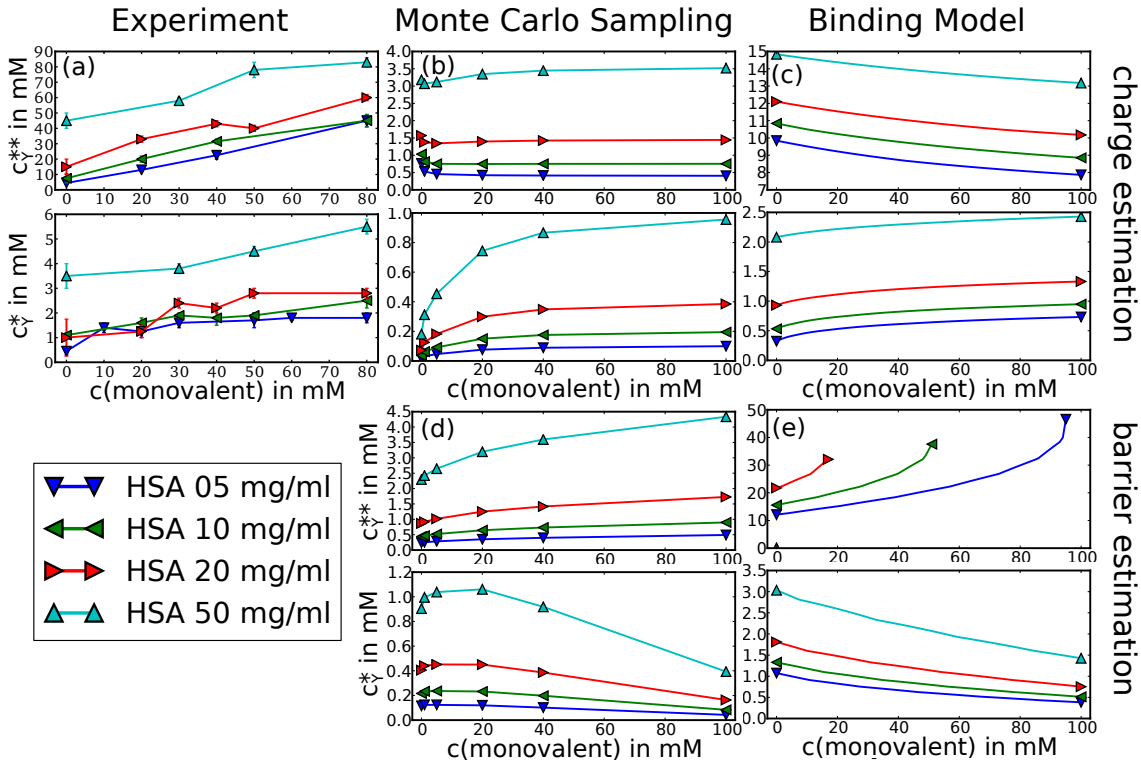

Figure 6.5: Transition concentrations c_Y^* and c_Y^{**} for the reentrant phase boundaries. The left column (a) summarizes the experimental results for the phase boundaries as determined by visual inspection at room temperature, showing a clear increase of both c_Y^* and c_Y^{**} with increasing monovalent salt. For $c_{\mathrm{NaCl}} \geq 100\,\mathrm{mM}$, no second transition is observed – the samples remain turbid up to very high c_Y. The center and right column show the results from the Monte Carlo sampling of binding configurations (b+d) using the detailed protein structure and the analytical coarse-grained model (c+e). Two estimations for c_Y^* and c_Y^{**} are used: The top row uses a critical charge of $Q^* = -5e$ and $Q^{**} = 5e$ as indicator (b+c). The bottom row estimates the transition concentration using the height of the repulsive barrier in the DLVO potential (d+e). For the analytical coarse-grained model (c+e), the curve corresponds to 100 data points that were not explicitly drawn for clarity. For the data points missing in (e) compared to (c), no repulsive barrier above $1k_BT$ was found.

from the experimental evidence by visual inspection and SAXS, the simulation results are used to provide insight into the underlying mechanisms of the ion–protein interaction.

A recent study investigated protein interactions in the presence of a set of mono-, di- and trivalent counterions, indicating stronger changes of attraction and repulsion at similar ionic strength for multivalent ions compared to monovalent ions (Kundu et al., 2013). These findings clearly suggest an effect of multivalent ions beyond charge screening and are consistent with the interpretation for HSA and YCl_3 concentrations below c_Y^*.

In protein solutions with little or no YCl_3, the notion of a charge stabilized solution is supported clearly by the SAXS results and the excellent agreement with model fits involving a screened Coulomb repulsion (Fig. 6.3(a)). The correlation peak evidences the overall repulsion that becomes weaker once the ionic strength increases.

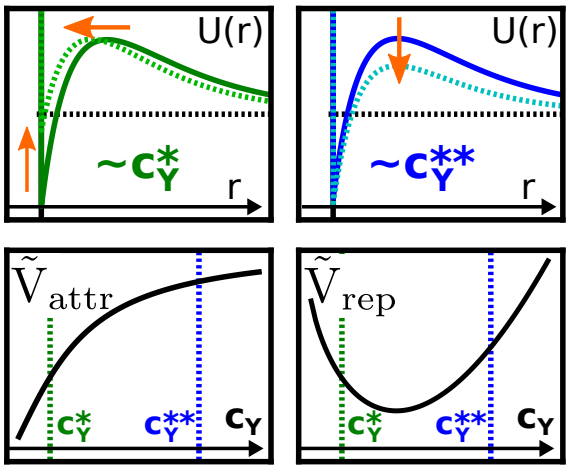

Figure 6.6: Schematic representation of the change of protein interaction throughout the phase diagram. The orange arrows indicate trends with increasing c_{NaCl}. Around c_Y^*, the addition of NaCl results in a weaker overall attraction, although the Coulomb repulsion is screened (top left). Around c_Y^{**}, the repulsion is screened with no evidence for varied attraction (top right). The addition of YCl_3 has several simultaneous effects on the strength of attraction, $\tilde{V}_{\mathrm{attr}}$ (bottom left), and repulsion, $\tilde{V}_{\mathrm{rep}}$ (bottom right). First, the increase of the ionic strength shortens the repulsion range. Second, ion binding to the protein surface causes a charge inversion and, thus, changes the repulsion strength non-monotonously (bottom right). Third, the presence of multivalent ions causes an additional attractive interaction, presumably mediated by ion bridges (bottom left).

Increasing the YCl_3 concentration, the first transition c_Y^* occurs. The experimental observation of c_Y^* for increasing NaCl concentration (Fig. 6.5(a)) agrees well with the trend observed by SAXS (Fig. 6.3(c)). With increasing c_{NaCl}, the sample conditions move away from the phase boundary into the stable solution regime I (Fig. 6.1), and the overall interaction has to become less attractive. Importantly, this trend cannot be caused by charge screening, since for net-negative proteins around c_Y^* charge screening would increase the overall attraction by a diminished repulsion for higher ionic strengths. Using the simulation results, the mechanism behind the first transition can be elucidated. The increase of c_Y^* seems to be related rather to a critical charge Q^* (Fig. 6.5(b+c)) than to a critical repulsive barrier (Fig. 6.5(d+e)). While counterintuitive at first glance, this finding can be rationalized by a consideration of the nature of the attractive interaction throughout the phase diagram. Multivalent metal ions can form ion-mediated intermolecular salt bridges between globular proteins (Zhang et al., 2011) and by these contribute considerably to the attraction. Ion-induced attraction is also suggested by a related experimental finding: HSA solutions without salt are stable even at the isoelectric point around pH 4.7 (Dockal et al., 2000). Since the isoelectric point corresponds to zero net charge and vanishing Coulomb monopole repulsion, the attraction without salt is not strong enough to cause aggregation. Thus, the reentrant condensation in the presence of multivalent ions implies that the ions indeed play an essential role for the attractive interaction (Roosen-Runge et al., 2013a), most likely via

ion bridging (Zhang et al., 2011; Roosen-Runge et al., 2013b). Since the protein net charge is reflecting ion binding, the critical charge corresponds to a critical number of bound ions that allow enough ion-induced attraction to destabilize the protein solution and at the same time neutralize the protein net charge. By contrast, the barrier estimation (see Sec. 6.3.3) uses a constant attraction represented by the Hamaker constant. The increase of c_Y^* with monovalent salt is thus most likely due to the effect of monovalent salt on the binding equilibrium of multivalent salts to the protein surface, and, interestingly, not due to a stronger charge screening of Coulomb repulsion due to higher ionic strengths.

For the second transition (c_Y^{**}), a different picture is found. The increasing trend of the experimental transition concentrations c_Y^{**} (Fig. 6.5(a)) suggests stronger attraction and weaker repulsion with increasing c_{NaCl}, since the phase boundary from the reentrant regime III to the phase-separated regime II (Fig. 6.1) is approached with increasing c_{NaCl}. This behavior agrees with the picture of a screened Coulomb repulsion. Using the simulation results, the importance of Coulomb repulsion for the solution stability of the reentrant regime III is supported. The increase of c_Y^{**} for all protein concentrations is reproduced by the barrier estimation and not by the charge estimation (Fig. 6.5(b–e)). This finding implies that charge is not enough to stabilize suspensions once the ionic strength is too high: Although the protein charge increases with monovalent salt concentration at fixed multivalent salt concentration, this gain in positive charge is not enough to overcome the effects of charge screening. Experimentally, this argument is further confirmed by the missing second transition c_Y^{**} for $c_{\mathrm{NaCl}} > 100\,\mathrm{mM}$ and the strong increase of precipitating aggregates with increasing c_{NaCl} (cf. protein concentrations in the measured supernatant in Fig. 6.3(b)). Considering the arguments on ion bridging from the previous paragraph on c_Y^*, a constant attraction, as used in the barrier estimation, is a reasonable assumption at high multivalent salt concentration: Once several ions have bound to the protein surface, further ions cannot contribute additional attraction via ion bridges for steric reasons. We therefore infer that charge screening plays the primary role for the reentrant transition c_Y^{**}, while ion binding is only a secondary effect.

The SAXS results around c_Y^{**} demonstrate that the underlying mechanism of the second transition is far from trivial. The increasing amount of precipitated aggregates with increasing c_{NaCl} (cf. c_{HSA} in the supernatant in Fig. 6.3(b)) clearly evidences an overall increased attraction, consistent with the arguments of the previous paragraph. The low q behavior of the scattering intensity suggests the presence of larger aggregates in the supernatant measured after centrifugation (Fig. 6.3(b)). From these results, aggregates seems to play an important role around c_Y^{**}.

At this point it is important to also note the limitations of the two simulation schemes. Both schemes are designed for the calculation of ion binding to a single protein molecule. Real protein solutions, however, might form clusters and aggregates that affect both charge regulation and binding equilibria considerably. In particular in the reentrant and aggregated regime, corresponding to higher multivalent salt concentrations, the stability of

the protein solutions can be determined by other properties beyond those of the individual molecule. Furthermore, both schemes are based on theoretical estimations and literature values and involve no fitting of experimental observables. In particular, the large deviations in the absolute values for the transition concentrations between the Monte Carlo scheme and the experiments might be caused by a non-optimized choice for the binding constant.

Despite these limitations, the comparison of both schemes helps to prevent conclusions from potential artifacts. The results of both simplified schemes reproduce the experimental observations semi-quantitatively and lead to a consistent interpretation with an attractive interaction caused by ion bridges and repulsion due to Coulomb interaction.

It is worth noting that both the reentrant phase behavior and the mechanism of ion condensation differ considerably from the reentrant condensation in DNA systems (Saminathan et al., 1999; Nguyen et al., 2000b; Burak et al., 2004). In the case of DNA, the ions condense due to ion-ion correlations, while the counterion–protein interaction is dominated by binding energy. The trend with increasing monovalent salt concentration for the second transition is opposite between the two systems. For proteins, the transition concentration c_Y^{**} increases, while it stays constant or decays in the case of DNA (Saminathan et al., 1999).

6.6 Conclusions

We have presented a comprehensive study of the impact of monovalent ions (NaCl) on the reentrant phase behavior of the globular protein HSA induced by YCl_3. Experimental data on the phase diagram and from small angle X-ray scattering on the protein interaction have been compared to results from two simulation methods—a Monte Carlo sampling of binding configurations and an analytical coarse-grained model—on the protein net charges and the related solution stability. All results support a consistent picture of the protein interactions causing the reentrant phase diagram, as summarized schematically in Fig. 6.6. Using the effect of NaCl, effects of charge screening, i.e. diminishing Coulomb repulsion, could be separated from effects of increased attraction, providing insight into the nature of the underlying protein interaction.

At low YCl_3 concentrations c_Y, screened Coulomb interaction dominates. Approaching the first transition concentration c_Y^*, the binding of multivalent counterions not only reduces the net protein charge and Coulomb repulsion, but importantly also contributes to the attractive interaction between the proteins, probably by ion bridging between molecules(Zhang et al., 2011; Roosen-Runge et al., 2013b). Under these conditions, a higher ionic strength, although screening the Coulomb repulsion, decreases the overall attraction, since less counterions bind to the surface (Fig. 6.4). As observed in the experiments, the transition concentration c_Y^* increases with increasing c_{NaCl}. The finding that the first transition is driven by the ion-induced increase of attraction (Fig. 6.6) is supported by the SAXS results (Fig. 6.3(c)) and the transition concentrations based on the charge estimation (Fig. 6.5).

For c_Y^{**}, a different mechanism appears to dominate. An increasing ionic strength decreases the Coulomb repulsion and renders charge stabilization of the protein solution only possible for higher protein charges, i.e. more bound counterions. Although more multivalent counterions are bound for higher c_{NaCl} (Fig. 6.4), the effect of charge screening overcompensates the gain in charge, as seen from the barrier estimation in Fig. 6.5. Consistent with the experiments, c_Y^{**} thus increases with increasing c_{NaCl}. These results support that the change of the second transition with increasing c_{NaCl} is driven by the reduction of repulsion due to charge screening (Fig. 6.6). The related SAXS results suggest a relevant role of aggregates around the second transition.

In summary, the present study provides a consistent picture for the underlying interactions driving the reentrant condensation in protein solutions in the presence of YCl_3. Monovalent salt tunes the interaction strength and the range of interaction. The competition of binding of multivalent ions and nonspecific effects of monovalent salts provides interesting opportunities to tune interactions in solutions.

Acknowledgements

We are grateful for the valuable discussions and comments from T. Narayanan (ESRF, Grenoble, France). We gratefully acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG). F. R.-R. acknowledges a fellowship of the Studienstiftung des deutschen Volkes. E. J. acknowledges a studentship of the ILL (Grenoble, France).

Chapter 7

Charge–Controlled Metastable Liquid–Liquid Phase Separation in Protein Solutions as a Universal Pathway Towards Crystallization

Chapter 7 is based on following publication:

Charge–controlled metastable liquid-liquid phase separation in protein solutions as a universal pathway towards crystallization
Fajun Zhang, Roland Roth, Marcell Wolf, Felix Roosen-Runge, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Sztucki and Frank Schreiber

Soft Matter, 8 (2012) 1313–1316

Contributions:

Research design	FZ, RR, FS
Experiments	FZ, MW, FRR, MWAS, RMJJ
Technical Assistance	MS
Data analysis and interpretation	FZ, RR, FRR
Paper writing	FZ, RR, FRR, FS
Illustration	FRR

Zhang et al. (2012c)

We demonstrate that a metastable liquid-liquid phase separation (LLPS) in protein aqueous solutions can be induced by multivalent metal ions at room temperature. We determine the salt and protein partitioning in the two coexisting phases. The structure factor obtained by small angle X-ray scattering provides direct evidence for a short-ranged attraction, which leads to the metastability of the LLPS. An extended phase diagram with three control parameters (temperature, protein and salt concentration) provides a

conclusive physical picture consistent with a criterion for the second virial coefficient. The presented isothermal control mechanism of the phase behavior opens new perspectives for the understanding of controlled phase behavior in nature. Furthermore, we discuss the application of this framework in predicting and optimizing conditions for protein crystallization.

7.1 Introduction

Metastable liquid-liquid phase separation (LLPS) in protein solution is a fundamental biophysical phenomenon and provides a mechanism for biological structure formation (Gunton et al., 2007; Annunziata et al., 2003; Pande et al., 2001; Wang et al., 2010; Galkin et al., 2002; Asherie, 2004; Vekilov, 2004; Muschol and Rosenberger, 1997) such as a prerequisite for the formation of crystals in cataracts (Annunziata et al., 2003; Pande et al., 2001; Wang et al., 2010) and fibers responsible for sickle cell anemia and Alzheimer's disease (Annunziata et al., 2003; Galkin et al., 2002), and changes on the pathways of protein crystallization (Wolde and Frenkel, 1997; Vekilov, 2004; Lutsko and Nicolis, 2006). Although claimed to be universal, a metastable LLPS has so far been observed only in a few protein systems due to the rather small accessible temperature window in conventional approaches (Gunton et al., 2007; Annunziata et al., 2003; Asherie, 2004). Proteins, like colloids, in solution often interact via effective interactions caused by other components of the solution, i.e. the solvent, salt, *etc.* Changing the solvent conditions alters the resulting interactions to a large extent. Usually the range of the effective attraction for colloids or proteins in solutions is shorter than the size of the particles, which leads to a metastable LLPS (corresponding to the gas-liquid phase transition in one-component systems) (Rosenbaum et al., 1996; Anderson and Lekkerkerker, 2002; Poon, 2002; Asherie et al., 1996; Wolde and Frenkel, 1997; Gunton et al., 2007). Theory, simulations, and experiments have predicted that density fluctuations near LLPS can significantly reduce the energy barrier of protein condensation including crystallization, i.e. in addition to the classical nucleation mechanism, a two-step nucleation mechanism has been proposed (Rosenbaum et al., 1996; Asherie et al., 1996; Wolde and Frenkel, 1997; Vekilov, 2004; Lutsko and Nicolis, 2006). Experimental and theoretical studies of colloid-polymer systems provide insight into the phase behavior of a system with short-ranged attractions, where the range and the strength of attraction can be tuned by the molecular weight and concentration of the polymer (Poon, 2002; Belloni, 2000). While for conventional colloids it is relatively easy to induce attractions with polymer, in protein solutions, where charge effects are ubiquitous, it seems natural to employ charges to tune the interactions. It is not obvious *a priori*, though, whether the predictions for conventional colloids apply to proteins with their discrete charge pattern, non-spherical shape and softness.

7.2 Results

In this Letter, we demonstrate that trivalent salts can be used to optimize crystallization conditions for globular proteins along general physical arguments. We first present the phase diagram of human serum albumin (HSA), a medium size globular protein very abundant in blood with a molecular weight of 67 kDa, as a function of the concentration of protein, c_p, and the concentration of the trivalent salt YCl_3, c_s, at 20° C (FIG. 7.1a). Trivalent counterions have been demonstrated to be an efficient method for tuning interactions in protein solutions by the binding of cations to acidic residues (Asp and Glu) on the protein surface, causing a charge-inversion of the protein (Zhang et al., 2008, 2010), whereas mono- and divalent ions do not cause such phase behavior in general. In the experimental range (3 to 80 mg/mL), the (c_p, c_s) plane is divided into three regimes, with regime I and III corresponding to a single-phase solution and regime II a two-phase state (Zhang et al., 2008). Here, we focus on the condensed phase in regime II, which is initially turbid (FIG. 7.1a). Closer examination by optical microscopy (FIG. 7.1b) reveals tiny droplets of the protein-rich phase suspended in solution. These can merge and grow, indicating a LLPS. A similar LLPS induced by YCl_3 is observed in bovine serum albumin (BSA) and β-lactoglobulin solutions (Zhang et al., 2011). Thus, LLPS is not specific for HSA, but a universal phenomenon for negatively charged proteins in the presence of trivalent cations, which makes the approach discussed here applicable to, in principle, 46 % of the entire protein family (Zhang et al., 2010).

For 31 mg/mL HSA solutions at 20° C, the LLPS occurs at a salt concentration c_s between 3 and 20 mM (FIG. 7.1a). The partitioning of both protein and salt into two coexisting phases is determined by X-ray and ultra-violet light absorption (see supplementary material in Sec. 7.5). In FIG. 7.1a, the pairs of coexisting phases ($c_p^{(1)}$, $c_s^{(1)}$) and ($c_p^{(2)}$, $c_s^{(2)}$) (connected by dotted tie-lines) define the coexistence region (red ellipsoidal area). The LLPS occurs within a closed area in the $c_p \sim c_s$ plane within Regime II. While inside the coexistence we find LLPS, the condensed phase of Regime II outside the area is populated with amorphous clusters.

Also the stability of the L-L coexistence phase, which is controlled by the effective attraction range, supports this picture as follows. The freshly prepared protein-rich phase is transparent with light yellow color. After several days up to weeks of storage at 20° C, the dense liquid eventually becomes gel-like (FIG. 7.1e). For the protein-poor phase, crystals can grow from these solutions (FIG. 7.1c). These observations indicate that the LLPS is ultimately metastable: the free energy of the dense liquid phase is lower than that of the initial solution, whereas it is higher than that of the solid state (aggregates or gel) (Vekilov, 2004). The metastability of the LLPS induced by multivalent counterions is similar to the thermally-driven LLPS in concentrated protein solutions, such as lysozyme, where the protein-rich phase decays into a solid phase over time (Muschol and Rosenberger, 1997; Shah et al., 2004). These observations are consistent with the theoretical predictions that short-ranged attractions in protein solutions cause the L-L coexistence curve to move below

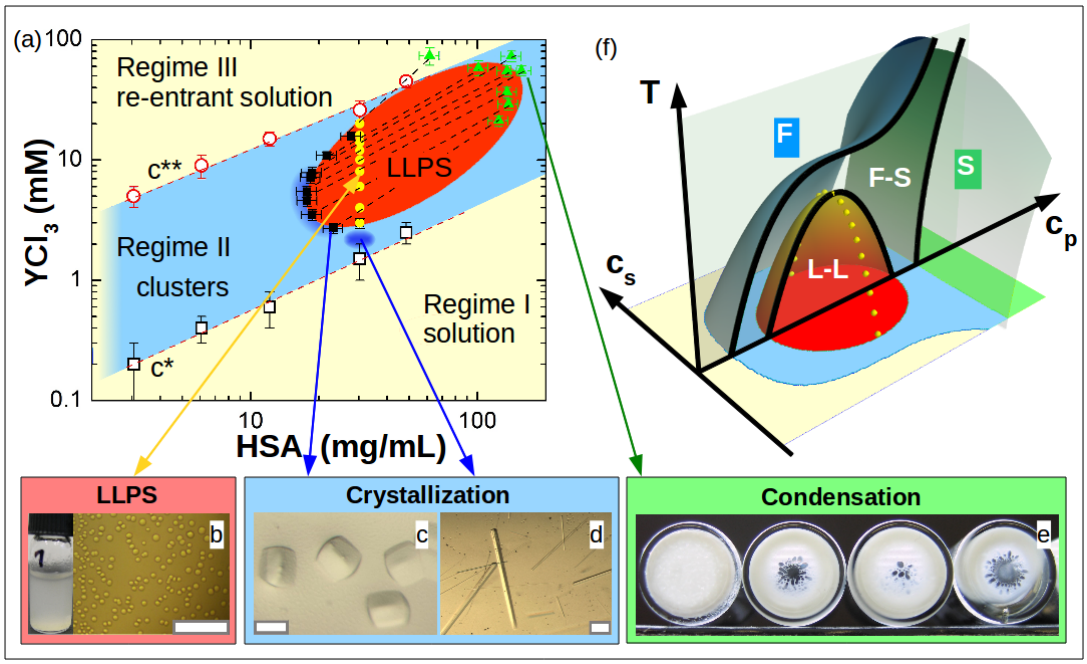

Figure 7.1: Tuning interactions in protein solutions towards controlled protein crystallization (all scale bars correspond to 0.1 mm): (a) Phase diagram of HSA controlled by YCl_3 in the (c_p, c_s) plane. The open symbols with error bar represent the boundary between the regimes as determined by optical transmission, while the solid symbols (red area) denote coexisting liquid phases. (b) A typical optical microscopy image of a freshly prepared sample with 31.0 mg/mL HSA and 4.0 mM YCl_3 shows small droplets of protein-rich phases, which coalesce proving that a non-arrested LLPS occurs. (c+d) Crystallization with different growth mechanisms is observed in the dilute coexisting phase (c) and in the region slightly below the LLPS boundary (c_s=2.0 mM) (d). (e) Amorphous aggregation in the protein-rich coexisting phase after storage at 20° C for two weeks indicates the general metastability of the regime II to aggregation (pictures for initial HSA concentration 31.0 mg/mL with 4, 6, 10 and 15 mM YCl_3). (f) Sketch of a phase diagram with three control parameters: temperature, T, protein and salt concentration, c_p and c_s. The phase behavior can be connected to the phase diagram established in colloidal systems with a short-range attraction. Importantly, the two crystallization areas (c and d) at the same temperature are consistent with predictions for optimized crystallization conditions around the LLPS critical point (d) and at the dilute coexisting phase at a suitable attraction strength (c) (Vliegenthart and Lekkerkerker, 2000). The presented phase diagram with T and c_s as control parameter in addition to c_p thus proves the metastable LLPS to be a universal pathway towards crystallization.

the gas-solid curve and the LLPS becomes metastable (Wolde and Frenkel, 1997; Petsev et al., 2003; Muschol and Rosenberger, 1997).

From the perspective of colloid theory, a metastable LLPS is caused by a strong attractive potential with a range much smaller than the effective hard sphere particle diameter (Poon, 2002; Asherie et al., 1996; Wolde and Frenkel, 1997). Regardless of the precise origin of the short-ranged attraction between the proteins, its presence is essential for the LLPS: the loss of entropy in the high density phase, compared to the corresponding entropy in the low density phase, has to be compensated by the increase in internal energy due to the attraction.

Using the virial expansion of the osmotic pressure, one obtains the second viral coefficient B_2 as a measure of the integrated strength of the interaction ($\beta = 1/(k_B T)$):

$$\beta P(\rho) \approx \rho + B_2\rho^2 + B_3\rho^3 + \dots \tag{7.1}$$

$$B_2 = 2\pi \int_0^\infty \mathrm{d}r\, r^2 [1 - \exp(-\beta V_{\text{eff}}(r))] \tag{7.2}$$

Mechanical equilibrium at coexistence implies that the osmotic pressure in the high density phase is equally low as in the low density phase. This can only be achieved by a sufficiently negative value of B_2. In fact it was observed that a reduced second virial coefficient of $B_2/B_2^{HS} < -1.5$ is required for the occurrence of a LLPS (see also supplementary material in Sec. 7.5)(Vliegenthart and Lekkerkerker, 2000; Noro and Frenkel, 2000).

In order to further understand the multivalent cation induced LLPS on the molecular level, the effective protein-protein interaction has been investigated by SAXS (for details see supplementary material in Sec. 7.5) (Zhang et al., 2008, 2007; Kline, 2006; Ianeselli et al., 2010). Typical SAXS curves of the protein-poor phase after the LLPS are shown in FIG. 7.2. Data fitting of a series of sample solutions with $c_p = 6.0$ mg/mL gives a form factor of $1.7 \times 5.3 \times 5.3$ nm^3 (see supplementary material in Sec. 7.5) resulting in an effective sphere diameter of $\sigma = 8.1$ nm. To further quantify the attractive potential, a short-ranged square well structure factor was used to fit the data (FIG. 7.2) (Kline, 2006; Menon et al., 1991), accounting for steric and short-range interaction:

$$\beta V_{\text{eff}}(r) = \begin{cases} \infty & 0 < r < \sigma \quad , \\ \ln(12\,\tau\,\frac{\Delta}{\sigma+\Delta}) & \sigma < r < \sigma + \Delta \quad , \\ 0 & \sigma + \Delta < r \quad , \end{cases} \tag{7.3}$$

where Δ is the width of the attractive well. In the limit $\Delta \to 0$ the well-known case of sticky hard spheres (SHS) is recovered and only the stickiness parameter τ is required to account for the attraction and relates directly to the virial coefficient

$$\lim_{\Delta\to 0} \frac{B_2(\tau)}{B_2^{HS}} = 1 - \frac{1}{4\tau} \quad . \tag{7.4}$$

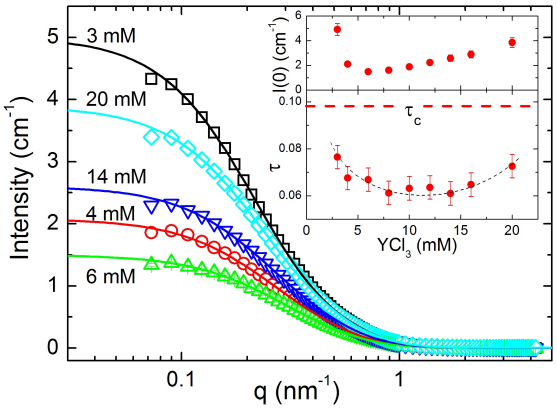

Figure 7.2: SAXS data with model fitting for the protein-poor phases from sample solutions with initial c_p of 31.0 mg/mL after LLPS. Only every second data point is plotted for clarity. The insets show τ and $I(0)$ as a function of c_s.

Within the SHS model, the critical point of the LLPS occurs at a stickiness parameter of $\tau_c = (2-\sqrt{2})/6 \approx 0.0976$ (Baxter, 1968), which corresponds to $B_2(\tau = \tau_c)/B_2^{HS} \approx -1.56$ and thus satisfies the criterion for the second virial coefficient (Vliegenthart and Lekkerkerker, 2000; Noro and Frenkel, 2000).

For the SAXS data fitting, Δ was fixed to $0.02\,\sigma$ to prevent artificial coupling with τ. The obtained τ values (see inset of FIG. 7.2) are generally below τ_c for YCl_3 concentrations between 3 and 20 mM, which is consistent with the observed LLPS. The related interaction potential has a depth of $\sim 4\,k_BT$. Approaching the upper and lower LLPS boundary for the salt concentration, τ increases towards τ_c, reflecting the vicinity to a critical point as expected from FIG.1f.

The absolute scattering at low q, $I(0)$, adds to this consistent physical picture in a model-free way. $I(0)$ is determined by the compressibility χ_T, since $S(q \to 0) = k_BT\rho\,\chi_T$ (Hansen and McDonald, 2006). The compressibility χ_T diverges at the spinodal line. Thus, the closer the coexisting densities are, the closer binodal and spinodal lines are, and the larger is χ_T and hence $S(q \to 0)$ in the coexisting phases. The experimental results on the microscopic interactions thus reflect the phase behavior, as can be seen easily by comparing the SAXS intensity for small values of q (FIG. 7.2, Inset) with the LLPS phase behavior shown in FIG. 7.1. As the YCl_3 concentration, for a fixed c_p of 31.0 mg/mL, is increased up to 3 mM, the system phase-separates. For this c_s the system is close to the lower critical point of the LLPS, which results in large values of χ_T and $S(q \to 0)$. As the YCl_3 concentration is increased further up to 10 mM, the coexistence region broadens, causing χ_T and $S(q \to 0)$ to decrease. In the range between 10 to 30 mM of YCl_3 the trend is reversed: χ_T and $S(q \to 0)$ increase again, until the system mixes for YCl_3 concentrations above 30 mM.

7.3 Discussion

The phase behavior shown in FIG. 7.1a can now be understood along the usual phase diagram for colloids with short-range attraction, when taking into account that protein and bound Y^{3+} ions build effective complexes with a tunable (short-range) attraction. Importantly, this binding mechanism is generally present in protein systems, rendering the presented approach a universal pathway towards controlling phase behavior in protein solutions.

In FIG. 7.1f, the black lines present a typical phase diagram in the (T, c_p) plane for such a protein-ion complex. By changing the ratio between protein and bound Y^{3+} ions, the interaction strength varies, which is reflected in a shift of the critical point of the LLPS (visualized for selected complexes by the yellow points). Regarding an isothermal plane (c_s, c_p) leads back to the phase diagram presented in FIG. 7.1a. Note that the two coexistence lines for the solid-fluid transition at the dilute limit should merge and for high protein concentration, the phase behavior becomes more complex (Zaccarelli, 2007).

The critical points of LLPS can be nicely described using the thermodynamic criterion based on B_2, which has been used as a predictor in protein crystallization. George and Wilson observed that B_2 falls in a narrow range for protein crystallization (George and Wilson, 1994). Subsequent theoretical work by Vliegenthart and Lekkerkerker demonstrated that B_2 has a nearly constant value at the critical point and indeed can be used as a predictor for protein crystallization, i.e. the optimal conditions for crystal growth are either near the critical point where the density fluctuation enhances the nucleation rate (Wolde and Frenkel, 1997), or below the critical point but near the protein-poor phase boundary where crystals grow via a two-step procedure (Vliegenthart and Lekkerkerker, 2000).

The growth of high quality protein single crystals (FIG. 7.1c+d) supports these theoretical predictions. Importantly, by using multivalent ions, both mechanisms can be observed at the same temperature: FIG. 7.1c corresponds to the case of a two-step nucleation and FIG. 7.1d represents nucleation close to the LLPS critical point. By varying temperature, crystallization conditions can be further optimized. While a detailed crystallographic study is beyond the scope of this Letter, several proteins have been previously crystallized using YCl_3 as a crucial additive without clarifying its exact role (Bouyain et al., 2005; Xu et al., 2001). FIG. 7.1 now provides a context for the physical understanding.

The exact physical origin of the short-range attraction remains to be identified – van der Waals, solvation and structural forces, which have a range in the order of $1.1\,\sigma$ (Rosenbaum et al., 1996), can contribute to the short-range attraction. Effective spheres as representation of proteins has proved successful for data modeling and the description of the dynamics of protein solutions (Jennings and Parslow, 1988; Roosen-Runge et al., 2011; Heinen et al., 2011). However, protein-Y^{3+} complexes with a non-spherical shape, nonhomogeneous surface charge pattern and complex hydration properties are expected to show highly directional, short-ranged interactions. Thus, a more detailed theoretical analysis of this system could be achieved within the framework of patchy particles (Wertheim, 1984a; Foffi

and Sciortino, 2007; Pawar and Kretzschmar, 2010), which could also explain the rather low volume fractions at the LLPS critical point below 10 % (Bianchi et al., 2006).

However, even without the precise knowledge of the microscopic mechanism, multivalent metal ions can be successfully used to control the macroscopic phase behavior of globular proteins. The observed phase behavior can be understood and described even quantitatively based on fundamental thermodynamic principles and colloid theory.

7.4 Conclusions

In summary, we have presented a theoretically consistent and comprehensive understanding of the controlled optimization of protein crystallization around the metastable LLPS in protein solutions solely induced by multivalent ions. The results suggest a universal applicability of the approach and open the field for further systematic studies of protein phase behavior.

The authors profited from discussions with C.P. Royall (Bristol, UK), S. Dietrich (Stuttgart, Germany) and T. Narayanan (ESRF, Grenoble, France). We acknowledge financial support from the Deutsche Forschungsgemeinschaft (DFG).

7.5 Appendix: Supplementary Material

Here we provide supplementary material for the paper Zhang et al., Metastable Liquid-Liquid Phase Separation in Protein Solutions as a Universal Pathway Towards Crystallization, 2012. First, we assemble supporting experimental information. Second, we give a theoretical consideration on the criterion of the second virial coefficient related to the LLPS.

7.5.1 Supporting Theoretical Information

Effective One-Component Systems

We briefly elaborate on the phenomenon of a liquid-liquid phase separation in a potentially complex mixture of several components like colloids or proteins in solution, which consists of the colloids/proteins, the solvent, added salt, etc. We provide a simple and intuitively transparent argument how it relates to the second virial coefficient. Often it is natural to focus on the behavior of the largest component, the colloids or proteins, and treating the remaining parts of the mixture as a background medium. This can be done in a rigorous way by mapping the Hamiltonian of the mixture onto an effective one-component system (McMillan and Mayer, 1945; Roth et al., 2001) by integrating out the degrees of freedom of those components that should no longer be considered. This mapping changes the interactions in the system from the *bare* interactions between all possible species combination, i.e. protein-water, protein-ion, etc., to *effective* interactions between particles of the remaining species.

For such a system to undergo a phase separation into a low density fluid phase (gas) and a high density fluid phase (liquid), it is important that the effective interaction possesses in addition to the repulsion at short distances an attractive tail. The reason is simple to understand: When a low density phase is transformed into a high density phase, then the system loses entropy, which has to be compensated by the gain in interaction energy due to the attractive tail. This argument holds independent of the precise nature of the effective interaction.

Clearly, the attraction has to be sufficiently strong for a phase separation to occur. A convenient measure for the strength of the attraction is the *effective* second virial coefficient, which is defined by

$$B_2 = 2\pi \int_0^\infty \mathrm{d}r\, r^2 \left\{1 - \exp(-\beta V_{\mathrm{eff}}(r))\right\}. \tag{7.5}$$

If B_2 is positive, the interaction is mainly repulsive and the system behaves hard-sphere-like, if it is negative, then the attraction is strong.

The question is how strong the *effective* interaction has to be in order to drive a phase separation. To this end Vliegenthart and Lekkerkerker made the interesting observation that for various systems the value of the second virial coefficient B_2 at the critical point seems universal (Vliegenthart and Lekkerkerker, 2000):

$$\frac{B_2}{B_2^{HS}} \approx -1.5, \tag{7.6}$$

with $B_2^{HS} = 16\pi R^3/3$ the second virial coefficient of a hard sphere of radius R. This observation was also confirmed by Noro and Frenkel (2000) and was tested for hard-sphere mixtures (Dijkstra et al., 1999) using the depletion potential (Roth et al., 2001).

The sticky hard-sphere model was introduced by Baxter (1968) as an example of a system with hard-core repulsion and additional attraction, which can undergo fluid-vapor phase separation and which can be treated analytically. The interaction potential is given by

$$\beta V_{\mathrm{eff}}(r) = \begin{cases} \infty & r < \sigma = 2R, \\ -\beta u_0 = \ln\left(\frac{12\tau\Delta}{\sigma+\Delta}\right) & \sigma < r < \sigma + \Delta, \\ 0 & r > \sigma + \Delta, \end{cases} \tag{7.7}$$

where usually the limit $\Delta \to 0$ is taken. In this limit one finds for the reduced second virial coefficient

$$\lim_{\Delta\to 0} \frac{B_2}{B_2^{HS}} = 1 - \frac{1}{4\tau}. \tag{7.8}$$

Baxter (1968) found that within the Percus-Yevick closure relation at the critical point the stickiness parameter is given by

$$\tau_c = \frac{2-\sqrt{2}}{6} \approx 0.0976. \tag{7.9}$$

If the value of the stickiness parameter at the critical point, τ_c is inputted into the expression for the reduced second virial coefficient, Eq. (7.8), one finds

$$\frac{B_2(\tau = \tau_c)}{B_2^{HS}} = 1 - \frac{1}{4\tau_c} \approx -1.56, \tag{7.10}$$

which agrees nicely with the aforementioned criterion for the second virial coefficient.

In order to better understand the observation by Vliegenthart and Lekkerkerker we consider phase coexistence between a low and a high density fluid in more detail. A low density fluid phase with density ρ_I can coexist with a high density phase with density ρ_{II}, if they are in mechanical and in chemical equilibrium, i.e.

$$p(\rho_I) = p(\rho_{II}), \quad \text{and} \quad \mu(\rho_I) = \mu(\rho_{II}), \tag{7.11}$$

where p is the pressure and μ the chemical potential. For mechanical equilibrium to be possible, the pressure at coexistence has to be low, because the pressure of a low density phase will always be low. Therefore, at coexistence, it is possible to expand the pressure into a virial expansion and truncate the series after a few terms with the most prominent contribution being the second virial term:

$$\beta p(\rho) \approx \rho + B_2\rho^2 + B_3\rho^3 + \dots. \tag{7.12}$$

Only, if the attraction is sufficiently strong can the pressure of a high density phase be equally low as that of the low density phase. The onset of a fluid-fluid phase separation is a critical point.

The location of the critical point follows from

$$\left.\frac{\partial p(\rho)}{\partial \rho}\right|_{\rho=\rho_c} = 0 = \left.\frac{\partial^2 p(\rho)}{\partial \rho^2}\right|_{\rho=\rho_c}. \tag{7.13}$$

Note that the vanishing of the first derivative of the pressure p w.r.t. density ρ expresses the condition of vanishing inverse compressibility or the bulk modulus. This is important because the structure factor $S(q)$ in the limit of $q \to 0$ reduces to the compressibility χ_T, i.e. $S(q \to 0) = k_B T \rho \chi_T$. This means that at the critical point or at the spinodal line, when the derivative of the pressure w.r.t. density, or the inverse compressibility, vanishes, the structure factor diverges for $q \to 0$. If one considers a system at a coexisting density, i.e. on or close to the binodal line, then the state is close to the spinodal, which implies that the compressibility is large, but does not diverge, which can be seen by a large value of $S(q \to 0)$. The closer the system is to the critical point, the closer the binodal and spinodal lines are. This means that close to the critical point, on the binodal (at either the low or the high coexisting density), the compressibility and hence $S(q \to 0)$ are large, while further away from the critical point, the difference in density between the binodal and spinodal increases, which means that the compressibility at a coexisting density and the structure factor $S(q \to 0)$ decrease.

By combining Eqs. (7.12) and (7.13) we obtain a set of equations which can be solved e.g. for the second and third virial coefficient at the critical point. Here only the result for B_2 is of interest. One finds that

$$\frac{B_2}{B_2^{HS}} \approx -\frac{1}{4\eta_c} \approx -2.06, \tag{7.14}$$

with the packing fraction at the critical point due to Baxter (1968)

$$\eta_c = \frac{3\sqrt{2}-4}{2} \approx 0.1213. \tag{7.15}$$

While the agreement between Eqs. (7.10) and (7.14) is clearly far from perfect, this simple argument helps to rationalize the origin of the B_2 criterion based solely on the idea of mechanical equilibrium.

Note that the value of the second virial coefficient at the critical point based either on Eq. (7.10) or on Eq. (7.14) change somewhat, if other estimates for the location of the critical point are employed, however, the magnitude remains of the right order.

The conclusion is that for $\tau < \tau_c$ or $B_2/B_2^{HS} < -1.5$ the interaction potential is sufficiently strong to drive a phase separation into a low density and a high density phase.

7.5.2 Supporting Experimental Information

Sample Preparation

Human serum albumin, HSA (A9511) and yttrium chloride were purchased from Sigma-Aldrich and used as received. All samples were prepared at 20° C. A series of protein solutions with various salt concentrations were prepared by mixing stock solutions. No buffer was used to avoid the effect of other co-ions. The phase transitions were determined by optical transmission. Protein concentrations (c_p) in each phase were determined using UV absorption at a wavelength of 280 nm with a coefficient of 0.51 mL/mg (Carter and Ho, 1994). LLPS was determined by optical microscopy. The protein-poor (phase-1) and protein-rich phase (phase-2) are separated by centrifugation. $c_p^{(1)}$ was determined directly by UV absorption, and $c_p^{(2)}$ was calculated from the volume of each phase and the initial protein concentration.

Small-Angle X-ray Scattering

Small-angle X-ray scattering (SAXS) was performed at beamline ID2, ESRF (Grenoble, France) at a wavelength of the incident beam of 0.08 nm (16038 eV). A q range of 0.07 to 4.2 nm^{-1} was covered using a sample to detector distance of 2.0 m. The 2D data were normalized and azimuthally averaged to obtain the intensity profiles on an absolute scale; the solvent background was subtracted. For more detailed information on data reduction and q-resolution calibration see Narayanan (2008) and Ianeselli et al. (2010). In order to determine the salt partitioning after LLPS, sample solutions of phase-1 were measured by

SAXS at three incoming energies well below ($E1 = 16038$ eV), near ($E2 = 17032$ eV) and on the absorption edge of yttrium ($E3 = 17038$ eV). The measured data were automatically corrected for the varying detector efficiency and normalized to absolute units, by using water as intensity calibration standard. The remaining difference of intensity integrated over a q range of 2.2 - 4.1 nm^{-1} gives the fluorescence contributed from Y^{3+}, which is proportional to salt concentration (c_s). Using a series of pure salt solutions, the fluorescence intensity $\Delta I(E3 - E1)$ as a function of c_s was calibrated (Figure 7.4). From $\Delta I(E3 - E1)$ of phase-1 $c_s^{(1)}$ was determined, and $c_s^{(2)}$ was calculated from the volume of each phase and the initial salt concentration.

SAXS Form Factor Determined from Dilute Protein Solutions

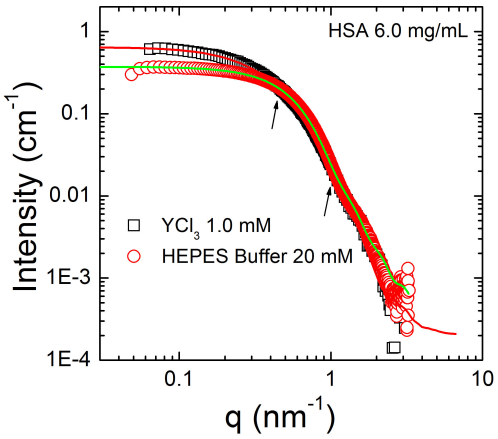

Figure 7.3: SAXS data with model fitting for HSA 6.0 mg/mL in Regime II, for comparison, SAXS data of HSA with HEPES buffer was also shown.

Calibration of Concentration Measurements of Y^{3+} from X-ray Absorption

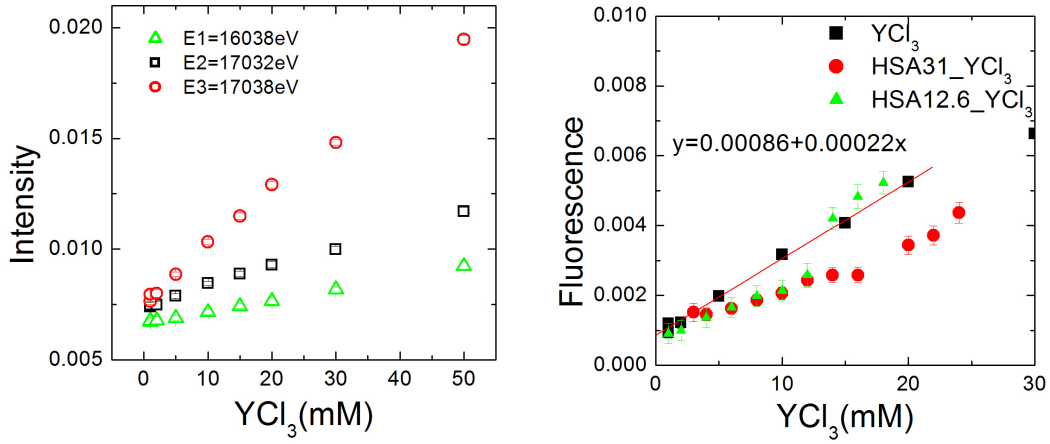

Figure 7.4: (left): Integrated intensity in the q range of 2.2 to 4.1 nm^{-1} for pure YCl_3 solutions from 1 to 50 mM at three energies. (right): Plots of fluorescence intensity, $\Delta I(E3 - E1)$, as a function of YCl_3 concentration.

Chapter 8

The Role of Cluster Formation and Metastable Liquid–Liquid Phase Separation in Protein Crystallization

Chapter 8 is based on the following publication:

The Role of Cluster Formation and Metastable Liquid–Liquid Phase Separation in Protein Crystallization

Fajun Zhang, Felix Roosen-Runge, Andrea Sauter, Roland Roth, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Sztucki and Frank Schreiber

Faraday Disc., 159 (2012) 313–325

Contributions:

Research design	FZ, FS
Experiments	FZ, FRR, AS, MWAS, RMJJ
Technical Assistance	MS
Data analysis and interpretation	FZ, FRR,RR
Paper writing	FZ, FRR, FS

Zhang et al. (2012a)

We discuss the phase behavior and in particular crystallization of a model globular protein (β-lactoglobulin) in solution in the presence of multivalent electrolytes. It has been shown previously that negatively charged globular proteins at neutral pH in the presence of multivalent counterions undergo a "re-entrant condensation (RC)" phase behavior (Zhang et al., 2008), i.e. a phase-separated regime occurs in between two critical salt concentrations, $c^* < c^{**}$, giving a metastable liquid–liquid phase separation (LLPS). Crystallization from the condensed regime has been observed to follow different mechanisms. Near c^*, crystals grow following a classic nucleation and growth mechanism; near c^{**}, the crystallization follows a two-step crystallization mechanism, i.e, crystal growth follows a metastable

LLPS. In this paper, we focus on the two-step crystal growth near c^{**}. SAXS measurements indicate that proteins form clusters in this regime and the cluster size increases approaching c^{**}. Upon lowering the temperature, in situ SAXS studies indicate that the clusters can directly form both a dense liquid phase and protein crystals. During the crystal growth, the metastable dense liquid phase is dissolved. Based on our observations, we discuss a nucleation mechanism starting from clusters in the dilute phase from a metastable LLPS. These protein clusters behave as the building blocks for nucleation, while the dense phase acts as a reservoir ensuring constant protein concentration in the dilute phase during crystal growth.

8.1 Introduction

Crystallization plays a decisive role in many processes in nature and industry. For example, the pharmaceutical industry requires production of the desired crystal form of the drug molecules which is important for their biofunction and stability (Gunton et al., 2007). However, it has long been known that the unique features of crystallization, including crystal lattice and polymorphism, particle size, and its distribution, are defined in the nucleation stage. In spite of the huge efforts over the last decades, our understanding of the early stage of crystallization is still limited. A breakthrough over the last decade has revealed new insight into this step: studies have shown that in colloid and protein solutions, the attractive potential is short-ranged as compared to their size, which is crucial for their phase behavior, and a "two-step" mechanism has been proposed to explain the crystallization behavior in these systems under suitable conditions, i.e. nucleation events follow a metastable liquid–liquid phase separation (LLPS) (Gunton et al., 2007; Vekilov, 2004, 2010b).

In the region of LLPS, however, colloidal systems with spherical isotropic short-range potential will become a gel or arrested phase instead of a dense liquid phase. In other words, the gelation line coincides with the phase separation boundary (Lu et al., 2008). Proteins, which differ from conventional colloids by having a non-spherical shape and inhomogeneous charge pattern, have shown liquid–liquid coexistence under suitable conditions (Delaye et al., 1981; Broide et al., 1991; Muschol and Rosenberger, 1997; Galkin and Vekilov, 2000b; Annunziata et al., 2002; Galkin et al., 2002; Grouazel et al., 2002; Annunziata et al., 2003; Chen et al., 2004; Dorsaz et al., 2011; Wang et al., 2011; Zhang et al., 2011, 2012c) and gelation at higher volume fraction or lower temperature (Cardinaux et al., 2007; Gibaud et al., 2011; Gibaud and Schurtenberger, 2009; Sciortino and Zaccarelli, 2011). Near LLPS, recent studies reveal the formation of dense protein clusters. The formation and physical origin of clusters in colloidal and protein solutions have attracted significant attention in soft matter studies during the last few years. For instance, it is a crucial step for understanding the mineralization process (Navrotsky, 2004). The cluster phenomenon is closely related to the interactions and phase behavior of these systems. For example, counter-balanced interactions have been reported to lead to equilibrium as well as transient

clusters in concentrated protein solution (Liu et al., 2005; Shukla et al., 2008; Stradner et al., 2004). Recent studies have revealed another type of cluster phase existing in concentrated protein solutions, such as lysozyme, hemoglobin, etc (Gliko et al., 2007; Pan et al., 2007, 2010). The clusters can be very big with 105 to 106 molecules. A theoretical model has been proposed to explain the origin of such a long-living cluster phase (Pan et al., 2010).

The consequence of the formation of protein clusters (transient or equilibrium) and metastable LLPS is that it changes the kinetic pathway of crystal nucleation significantly. There is increasing evidence that clusters, nanoscale amorphous precipitates, and other more complex precursors in the aqueous phase play an important role in crystallization (Navrotsky, 2004; Banfield et al., 2000; Furrer et al., 2002; Mintova et al., 1999). Computer simulations by ten Wolde and Frenkel show that far from the metastable LLPS critical point, the nucleation follows the classical nucleation mechanism. However, when approaching the critical point, the critical nucleus becomes highly disordered, liquid-like droplets which further follow a structural change to eventually become crystalline (Wolde and Frenkel, 1997). This finding inspires a two-step model, i.e. nucleation occurs within the dense liquid phase, which corresponds to the separation of order parameters (density and structure) during crystallization. It applies not only to proteins and colloidal systems, but also to small molecular systems (Erdemir et al., 2009; Bonnett et al., 2003). However, the role of the protein cluster as well as the dense liquid phase during nucleation and protein crystallization is still not entirely clear (Vekilov, 2010b).

We have recently studied the phase behavior of globular proteins in solution in the presence of multivalent metal ions. It has been shown that solutions of negatively charged globular proteins at neutral pH in the presence of multivalent counterions undergo a "re-entrant condensation (RC)" phase behavior (Zhang et al., 2008; Ianeselli et al., 2010; Zhang et al., 2010, 2011, 2012c), i.e. a phase-separated regime occurs in between two critical salt concentrations, $c^* < c^{**}$, including a metastable liquid–liquid phase separation (LLPS) (Zhang et al., 2012c). Crystallization from the condensed regime follows different mechanisms. Near c^*, crystals grow following a classic nucleation and growth mechanism; near c^{**}, the crystallization occurs from phase-separated protein solutions, indicating a two-step mechanism, i.e. crystal growth follows a metastable LLPS (Zhang et al., 2011).

In this work, we aim to achieve new insights into the role of LLPS and clustering in protein crystallization by presenting a study of the structural evolution during a two-step process. The questions we are interested in are the following: (1) what is the structure of proteins in solution near c^{**}? Are they still in their dimeric state or forming clusters? (2) If clusters are formed, how do the size and structure of the protein clusters depend on the location in the phase diagram? (3) Can the nuclei of crystals be formed via cluster–cluster aggregation? (4) What is the relationship between clusters and the LLPS? With these questions in mind, we have performed systematic SAXS measurements on a series of solutions, and the crystal growth has been followed using in situ measurements as a function of temperature.

8.2 Experimental

8.2.1 Materials

Globular β-lactoglobulin (BLG) from bovine milk (L3908) and yttrium chloride (YCl_3) were purchased from Sigma-Aldrich. Solutions were prepared by mixing stock solutions of BLG (67 mg/ml) and YCl_3 (100 mM). The phase diagram (protein concentration c_p vs. salt concentration c_s) was determined at room temperature ($\approx 22\,^{\circ}$C) by monitoring the optical transmission of a series of protein solutions containing different salt concentrations (Zhang et al., 2008). The c_p values were determined by UV absorption using an extinction coefficient of 0.96 ml/mg at a wavelength of 278 nm (Sober, 1970). Note that the presence of high concentration of buffer (such as HEPES and Tris buffer) can affect the phase behavior and the solubility of yttrium salts. However, with lower buffer concentration (about 5 mM), the effect on the solubility of yttrium salts is negligible. To avoid the effect of other ions, no buffer was used in this work for sample preparation except for specific reference data set in Fig. 8.2 (Zhang et al., 2010).

8.2.2 Small-Angle X-ray Scattering (SAXS) Measurements

The SAXS measurements were performed at the ESRF (Grenoble, France) on the beamline ID02 with a sample-to-detector distance of 2 m or a combination of 0.85 and 5 m in order to cover a larger q range of 0.04 to 7.8 nm^{-1} (Sztucki et al., 2006). The data were collected by a high sensitivity fiber-optic coupled CCD (FReLoN) detector placed in an evacuated flight tube. The protein solutions were loaded using a flow-through capillary cell (diameter ≈ 2 mm; wall thickness ≈ 10 mm). No variation that would indicate radiation damage of SAXS profiles was observed comparing successive short exposures of cumulated time up to 3 s. For the temperature scan, each temperature was held for 15 min to allow equilibration. The incident and transmitted beam intensities were simultaneously recorded for each SAXS pattern with an exposure of 0.3 s. The 2D data were normalized to an absolute scale and azimuthally averaged to obtain the intensity profiles, and the solvent background was subtracted. For more detailed information on data reduction and q-resolution calibration, see Narayanan (2008).

8.3 Results and Discussions

8.3.1 Phase Behavior of Protein Solutions in the Presence of Multivalent Counterions

We first briefly summarize the phase diagram of BLG (c_p) as a function of the YCl_3 concentration (c_s) at room temperature (Fig. 8.1) with an extended range of c_p. This diagram provides a guide for optimizing the conditions for protein crystallization. For a given protein concentration c_p, an increase of the salt concentration c_s above a certain threshold

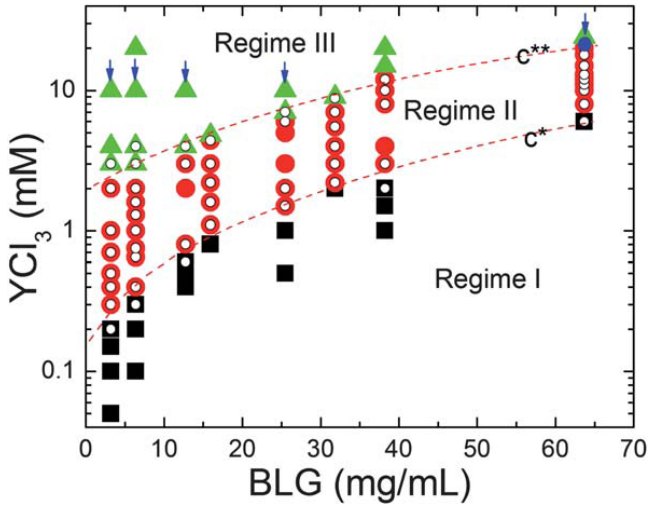

Figure 8.1: Extended phase diagram of protein BLG as a function of protein (c_p) and salt, YCl_3 (c_s) (Zhang et al., 2011). (a) Phase diagram at room temperature (22 °C). Solid symbols present the sample solutions at different regimes (see text). Small open symbols present the samples solutions where crystallization was observed at 4 °C . The small arrows in Regime III indicate the samples measured by SAXS.

(c^*) results in the protein solution becoming turbid and entering a two-phase state. When c_s is increased further (above c^{**}), the protein solution turns clear again. Thus, the two salt concentrations c^* and c^{**} divide the phase diagram into three regimes (Fig. 8.1). Regimes I and III contain clear protein solution, whereas the protein condenses (or aggregates) in Regime II. The addition of YCl_3 leads to a charge inversion on the protein surface, which has been proved using zeta potential measurements (Zhang et al., 2011, 2010). The experimental observations on the crystal growth from different regions indicate different growth mechanisms (Zhang et al., 2011): near c^*, crystal growth follows the classical nucleation theory, i.e. nucleation occurs from homogeneous supersaturated solutions. Near c^{**}, the solution has a transition temperature, Th. Below Th, a metastable LLPS occurs before crystallization, apparently resembling the so-called "two-step" nucleation mechanism. In the remaining part of paper, we will focus on the structure of protein clusters and crystallization in Regime III.

8.3.2 Protein Clusters and their Structure in Regime III Studied by SAXS

We now examine the solution structure of proteins in Regime III using SAXS. By fixing the salt concentration and varying the protein concentration, we approach the phase transition boundary, c^{**}. The selected solutions are shown in Fig. 8.1 as indicated by vertical arrows. The additional sample with higher protein concentration (67.0 mg/ml) with 15 mM YCl_3 near c^{**} is chosen for a temperature dependent measurement. Fig. 8.2 shows a series of SAXS profiles of BLG in solution with 10 mM YCl_3 and different protein concentrations. The scattering curves have been scaled by their protein concentrations in terms of the absolute intensity.

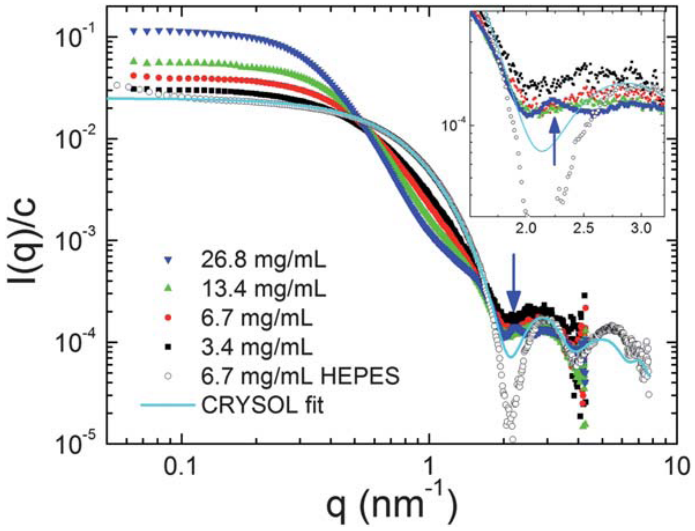

Figure 8.2: SAXS profiles of BLG solutions with YCl3 in the re-entrant regime. For a given value of c_s, a clear transition from dimer to cluster was observed with increasing protein concentration. For comparison, the SAXS curve of BLG in HEPES buffer is also shown, which can be well described by a dimer using the crystal structure (PDB code of 1BEB). Inset shows the expanding of the region with the arrow.

Sample	$I(0)/c$/cm^{-1}	R_g/nm	Number of dimer in cluster[a]
BLG 6.7 HEPES	0.0248	2.33	Monodisperse dimer
BLG 3.4 YCl3 10 mM	0.0321	3.08	1.3
BLG 6.7 YCl3 10 mM	0.0432	3.36	1.7
BLG 13.4 YCl3 10 mM	0.0663	3.67	2.7
BLG 27.0 YCl3 10 mM	0.1126	4.90	4.5

Table 8.1: Structural parameters obtained by Guinier analysis on the SAXS data.
[a] These values are calculated by dividing $I(0)$ by 0.0248, $I(0)$ of the pure dimer solution.

For comparison, a SAXS profile for BLG in 20 mM HEPES buffer is also plotted. These data were collected at two sample-to-detector distances (0.85 and 5 m) covering a large q range from 0.03 to 7.8 nm^{-1}. Two scattering minima are clearly visible at q = 2.2 and 4.2 nm^{-1}, respectively. In HEPES buffer (pH 7.0), BLG is dissolved as dimers (Elofsson et al., 1997; Gottschalk et al., 2003; Vogtt et al., 2011), which is further confirmed by fitting the SAXS data using CRYSOL (Svergun et al., 1995) with its crystal structure (PDB code 1BEB). The radius of gyration, R_g, of the BLG dimer obtained by the fitting procedure is 2.33 nm, which is in good agreement with those reported in the literature (Moitzi et al., 2011). For solutions with 10 mM YCl_3 and increasing BLG concentration, the scattering curves change smoothly. At 3.4 mg/ml protein, well below the phase boundary, the scattering intensity in the low q region ($q < 0.5$nm^{-1}) is higher as compared to the dimeric state, whereas in the intermediate q region (0.5nm$^{-1} < q < 1.4$nm^{-1}), the scattering intensity is lower. With increasing the protein concentration, this trend becomes more significant and two crossing points are observed at $q = 0.5$ and 1.4 nm^{-1}, respectively. With 26.8 mg/ml

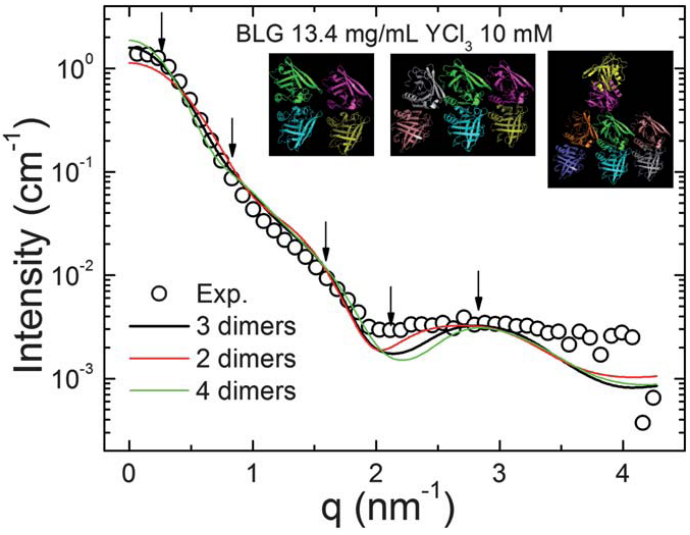

Figure 8.3: Experimental SAXS data of protein clusters compared to those created using the crystal structure (PDB code of 3PH5). Only 10 % of experimental data are plotted for clarity. The inset shows the structure of protein clusters created by the crystal structure with 2, 3 and 4 dimers.

protein, a kink is well-developed in the q region between the crossing points, indicating the formation of clusters. Basic structural parameters of these clusters, such as R_g, and the normalized forward intensity, $I(0)$, are obtained from a Guinier analysis (Glatter and Kratky, 1982). As shown in Table 8.1, R_g increases when approaching the phase boundary. $I(0)$ is related to the molecular weight of the cluster, which can be used to estimate the number of dimers within the clusters. Simply dividing the normalized $I(0)$ by 0.0248 (the value obtained from the form factor), the numbers are estimated as 1.3, 1.7, 2.7, 4.5, respectively. These results are listed in Table 1. A new maximum appears at $q = 2.2\,\mathrm{nm}^{-1}$ for $c_p > 20$ mg/ml as indicated by the arrow in Fig. 8.2 (expanded in the figure). This maximum, which corresponds to a center-to-center distance of $d = 2p/q = 2.82$ nm, is very close to the monomer–monomer distance within a dimer (3.08 nm) (calculated from the crystal structure: (1BEB). Thus, the appearance of this peak is attributed to the nearest neighbor correlation within clusters, indicating the liquid-like structure within the clusters.

The cluster phase observed in our system is due to the balance of electrostatic repulsion and the attraction (Groenewold and Kegel, 2001). We have demonstrated that in the re-entrant regime, the effective surface charge of proteins is inverted from negative to positive, which provides the long range repulsion (Zhang et al., 2008, 2010). On the other hand, the bridging effect of the yttrium cations provides the short range attraction. Although other interactions, such as van der Waals attraction may also contribute to the overall attractive potential, we assume that the bridging effect dominates, which is consistent with the crystallography study (Zhang et al., 2011). Clustering for qualitatively comparable interactions has been observed also for lysozyme and colloids in solution (Stradner et al., 2004).

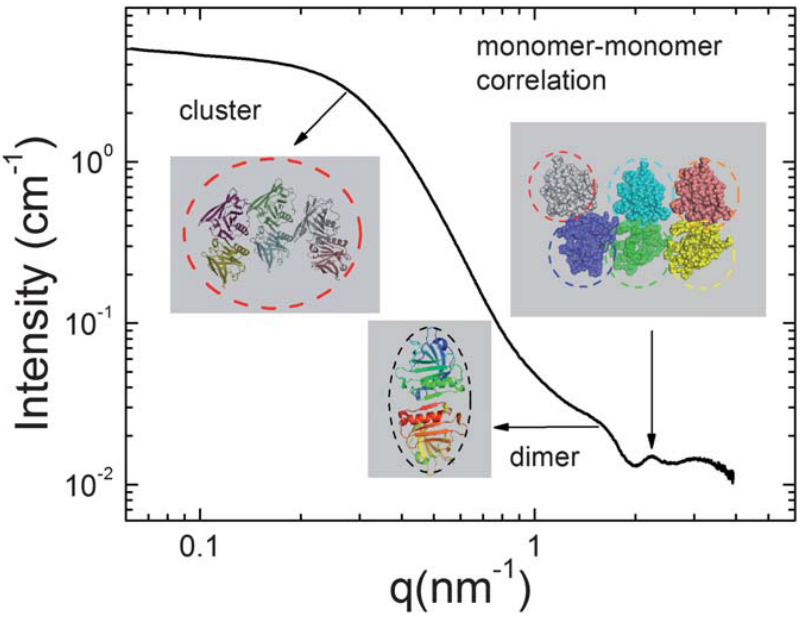

Figure 8.4: Structural hierarchy of BLG solutions in re-entrant regime revealed by SAXS. The peak and shoulders in the SAXS curve at $q = 2.2$, 1.8 and 0.3 nm^{-1} correspond to the monomer–monomer correlation, the form factor of a dimer, and the cluster, respectively.

Since the bridging effect seems to contribute anisotropic attractions, not only the crystal structure, but also the structure of these clusters should reflect the anisotropic interaction between the proteins being bridged by counterions. Indeed, the SAXS profiles of the cluster phases can be reasonably fitted by typical cluster structures created based on the crystal structure determined in our previous work (PDB 3PH5) (Zhang et al., 2011) with 2, 3 and 4 dimers, as shown in Fig. 8.3. The clusters are created as follows: we first choose one dimer from the lattice and set its center of mass to be the origin. Then the positions of the other dimers within the clusters are chosen based on the shortest center-to-center distance to the origin. The experimental SAXS profile of BLG 13.4 mg/ml with 10 mM YCl_3 is used to compare with these cluster structures using CRYSOL (Svergun et al., 1995). The best fitting comes from the cluster with 3 dimers (consistent with Table 1) which reproduces the R_g and $I(0)$ quantitatively, and other features of the SAXS profile qualitatively, such as the position of the kink, the shoulder for dimer, the minimum and maximum from the form factor as indicated by arrows in Fig. 8.3. Of course, this simple structural model of the cluster cannot reproduce all the details since the size and structure of the clusters will have a distribution. In spite of these difficulties, it is plausible to assume that the protein clusters have a "pre-crystalline" structure (Fig. 8.3) similar to those created from the crystal structure.

Having elaborated on the formation of small clusters with dimers as the principle building blocks from Fig. 8.2 and 8.3, we now review the full hierarchic structure of protein assembly in solution. Fig. 8.4 shows the SAXS profile of a sample close to the phase boundary with BLG 67 mg/ml with 15 mM YCl_3. A pronounced maximum at $q = 2.2\,\mathrm{nm}^{-1}$

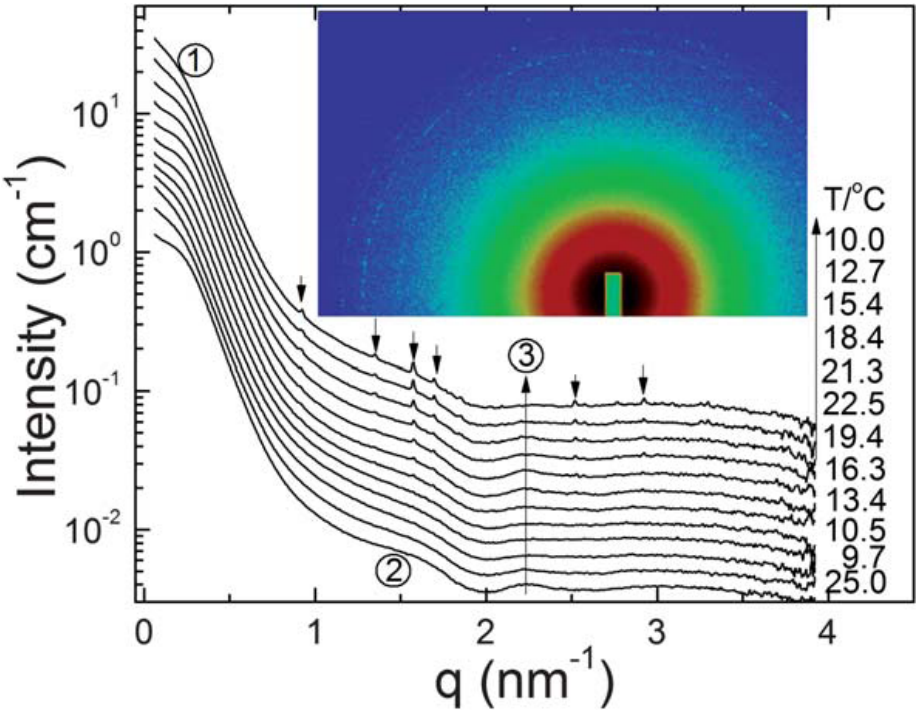

Figure 8.5: SAXS curves at different temperatures during cooling. Crystallization occurs below 25 °C . The intensity of the maximum at $q = 2.2$ nm^{-1} decreases with lowering temperature, and the low q intensity increases steadily. Bragg peaks appearing in the intermediate q range have been indexed using the crystal structure (see text). The curves are shifted upward for clarify. The inset shows the 2D scattering pattern at 10 °C .

is clearly visible. As mentioned above, the scattering peak at $q = 2.2\,\mathrm{nm}^{-1}$ is attributed to the monomer–monomer (M–M) correlation within clusters, which cannot be reproduced from the crystal structures in Fig. 8.3. The shoulder at $q \approx 1.8\,\mathrm{nm}^{-1}$ corresponds to the dimer scattering, and the shoulder at $q \approx 0.3\,\mathrm{nm}^{-1}$ corresponds to the form factor of the protein cluster. The number of dimers within the clusters is estimated to be around 3–4 for this sample since the intensities continuously increase in the low q region and both the forward intensity and the radius of gyration cannot be precisely determined by the Guinier analysis. The continuous increase indicates that the clusters may further build up higher level structures with a larger length scale, which makes the M–M correlation more significant.

8.3.3 Crystal Growth Followed by In Situ SAXS

We now present the results of crystal growth directly from the cluster solution by an in situ SAXS observation. Fig. 8.5 shows a series of SAXS profiles for BLG 67 mg/ml with YCl_3 15 mM. The transition temperature, T_h, is about 22 °C . At 25 °C (after preparation), the SAXS curve is the same as shown in Fig. 8.4. The sample is firstly quenched to 10 °C , then heated up stepwise to 22 °C , and then cooled down stepwise to 10 °C . The increment of heating and cooling is 1 °C and the time for equilibration is 15 min. After quenching to 10 °C , the

scattering intensity in the low q region (region (1) in Fig. 8.5) increases as compared to that at 25 °C , indicating the formation of larger clusters. The solution becomes turbid. In the meantime, the M–M correlation peak (region (3) in Fig. 8.5) becomes weaker. In the intermediate q region (region (2) in Fig. 8.5), where the scattering is mainly contributed by the form factor of a dimer, no changes could be observed. The subsequent stepwise heating shows the inverse trend, i.e. with increasing temperature, the intensity in the low q region decreases, and the M–M correlation peak becomes stronger. The solution becomes clear again. This observation suggests that the formation of larger clusters is a reversible process, which is consistent with our previous measurement on T_h (Zhang et al., 2011). So far, no Bragg peaks could be observed. Upon further lowering the temperature, one observes the following significant changes: (i) Bragg peaks appear, indicating the crystallization of proteins, indicating that protein crystals can grow directly from the clusters. The inset 2D pattern of the final state shows sharp circles indicating the isotropic distribution of the orientation in the solution. We have recently determined the structure of crystals grown under similar conditions in the phase diagram. All crystals have an orthorhombic structure with space group P212121 and contain one dimer in their asymmetric unit (PDB code 3PH5 and 3PH6), which can be used to index the Bragg peaks in the SAXS profiles (Zhang et al., 2011). For example, the first three Bragg peaks at $q = 0.924\,\mathrm{nm}^{-1}$, $q = 1.348\,\mathrm{nm}^{-1}$ and $q = 1.571\,\mathrm{nm}^{-1}$ can be assigned as (002), (012) and (100) of the crystal structure, respectively. This means that although M–M correlation is visible in the cluster phase, the dimer represents the building block of the crystals. (ii) In the low q region, the scattering intensity increases with decreasing temperature, indicating the formation of larger objects, i.e. clusters and crystals. (iii) No change is observed in region (2) except the appearance of Bragg peaks. (iv) the M–M correlation peak at q = 2.2 nm^{-1} continuously decreases in intensity as the temperature decreases. The temperature dependence of the M–M correlation peak may be related to the flexibility of the monomers within the cluster.

The crystal growth of similar samples has been followed using an optical microscope (Zhang et al., 2011). It shows that after quenching the solution below T_h, the solution becomes turbid instantly. After some time (hours) the dense liquid phase appears, then crystal growth appears to start from the dilute phase. Over time, the crystal growth consumes the material and leads to the dissolution of the droplets. Our observations of crystal growth in the presence of YCl_3 suggest that the most common way is the growth from the dilute phase directly leading to the coexistence of a dense liquid phase and crystals. From the results of the SAXS measurements, we now know that even in the dilute phase, proteins form clusters (Fig. 8.2, 8.3) and crystals grow directly from such protein clusters (Fig. 8.5). These observations suggest that the protein clusters undergo two competing pathways during cooling, i.e. either forming a dense liquid phase or reordering into a crystal, which can modify the pathways of crystal growth as to be discussed below.

8.4 Discussion

8.4.1 Pathway of the Two-Step Crystallization

We first discuss the possible pathway of protein crystallization during a two-step growth procedure, which leads to the discussion of the role of protein clusters and a metastable LLPS. A two-step mechanism following a metastable LLPS suggests that the nucleation occurs within the dense liquid phase instead of the dilute phase (Vekilov, 2010b; Wolde and Frenkel, 1997; Talanquer and Oxtoby, 1998). Because the surface free energy at the interface between the crystal and the solution is significantly higher than at the interface between the crystal and the dense liquid, the barrier for nucleation of crystals from the solution would be much higher. This would lead to much slower nucleation of crystals directly from the solution than inside the clusters. While indeed some experimental observations are consistent with this prediction, e.g. for glucose isomerase and hemoglobin (Galkin et al., 2002; Vivarès et al., 2005), other observations however, indicate that crystallization prefers to start in the dilute phase (Liu et al., 2010). While this has often been attributed to the high viscosity or gelation of the dense liquid phase, a clear understanding of their exact role is still missing.

In our system, we observe in addition to the metastable LLPS a non-negligible effect of clustering. Although the mechanism of clustering observed here may not be the same as in the concentrated protein solutions (Gliko et al., 2007; Pan et al., 2007, 2010), it can be assumed to play a similar role on protein crystallization. The protein clusters observed show interesting features with respect to crystallization, which make them suitable as building blocks for crystal formation. Firstly, the clusters have a pre-crystalline structure, since the cation binding sites represent specific interaction patches. These imprint a favorable structure, as implied by the SAXS scattering curves of the clusters being simulated well using the clusters created from the crystal structure. Secondly, clusters seem to have an internal flexibility, as evidenced by the pronounced M–M correlation peak in larger clusters. This flexibility enables local reorientation within the clusters, rendering the pre-crystalline structure of clusters to be an ideal building block for crystallization, since the enthalpy cost of nucleation via local reorientation is much lower than for a hypothetical nucleation directly from a dimer in solution. Thirdly, the cluster size and its distribution vary throughout the phase diagram; in particular, the clusters grow when the solution conditions approach the coexistence region. Large pre-crystalline clusters can serve as the precursor (stabilized nuclei) for further crystal growth, as recently observed in various systems (Navrotsky, 2004; Banfield et al., 2000; Furrer et al., 2002; Mintova et al., 1999).

Inspired by our observations, we propose an alternative pathway of protein crystallization in Fig. 8.6. The monodisperse dimeric proteins in solution form clusters upon adding YCl_3 due to the bridging effect of cations (Fig. 8.6a), as indicated by SAXS measurements (Fig. 8.2). Upon lowering the temperature across the phase boundary (below Th), LLPS takes place, leaving a dense phase and a dilute phase with clusters (Fig. 8.6b).

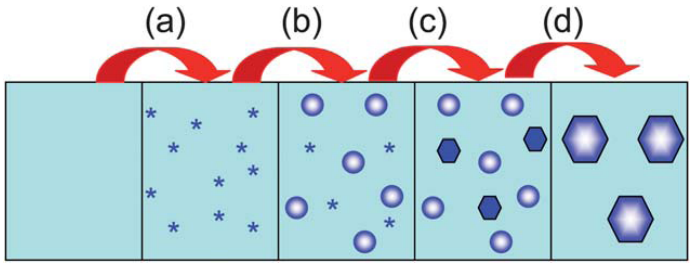

Figure 8.6: Schematic illustration of the phase transitions (LLPS and crystallization) from protein clusters (see text for details). (a) Upon addition of YCl_3, protein clusters formed via ion bridging. (b) Lowering temperature to $T < T_h$ and crossing the phase boundary, the clusters first aggregate with a liquid-like structure. (c) After the induction time protein crystals grow directly from the clusters. (d) With time increasing, while the crystals grow steadily, the dense liquid phase dissolves and disappears.

In solution, clusters can move freely and restructure themselves internally. Furthermore, clusters can assemble to larger clusters with a pre-crystalline structure. As a next step, internal reorientation turns the clusters to stable precursor nuclei (Fig. 8.6c). Because the dense liquid phase is metastable with respect to the crystalline phase, the crystals grow and consume the protein molecules from the dense liquid phase which dissolves and disappears (Fig. 8.6d). This pathway differs from the previous view of the two-step mechanism involving a metastable LLPS by the phases where nucleation starts. While the previous view predicts the crystals to be nucleated within the dense liquid phase, our observations suggest that both condensed phases can arise from the protein clusters in the protein-poor phase.

8.4.2 The Role of Clusters and Dense Liquid Phase in Protein Crystallization

The possible pathway of protein crystallization discussed above (Fig. 8.6) leads to a further discussion of the role of the protein clusters and the metastable dense liquid phase during the two-step crystallization procedure. In our system, both the dense liquid phase and crystals grow directly from the protein clusters (Fig. 8.6): we thus speculate that while the protein clusters enhance the nucleation, the dense liquid phase may play a role in optimizing the conditions for crystallization. Some clues arising from the detailed studies on lysozyme in solution support this explanation: Pan et al. have shown that the protein clusters exist between the solubility line and the LLPS in the phase diagram (Liu et al., 2005), which coincides with the optimum conditions for protein crystallization predicted using the second virial coefficient as predictor by Vliegenthart and Lekkerkerker (2000). Our observations from in situ SAXS (Fig. 8.5) indicate that nucleation is possible from protein clusters, and microscope observations indicate that crystals grow from the dilute solution instead of the dense liquid phase. Employing this picture, the metastable LLPS occurs as a helpful side effect at these thermodynamic conditions, since the LLPS indicates suitable interaction conditions and ensures that the dilute protein solutions stay at a defined

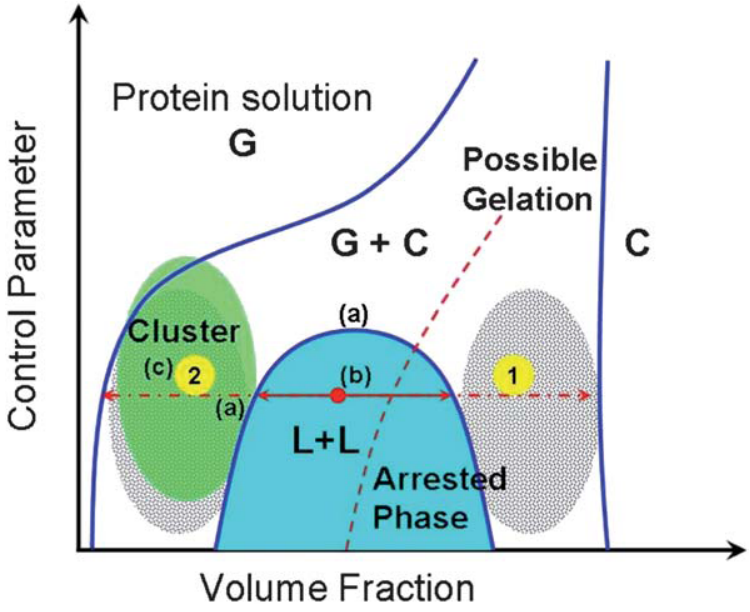

Figure 8.7: Schematic phase diagram of protein solutions. The region of clustering (Zhang et al., 2008) is marked in green and the gray shaded area corresponds to the optimum condition for crystallization (Glatter and Kratky, 1982). Two-step nucleation mechanism predicts that protein solutions located within LLPS region (red point) undergo a LLPS in the first step and then nucleation occurs within the dense liquid phase (along pathway (1)). An alternative pathway proposed in this work follows pathway (2): after LLPS, protein clusters in the dilute phase with the optimized thermodynamic conditions for crystallization can initialize crystal growth with a reduced energy barrier. Note that the exact conditions or the location of the gelation line will vary with the specific system.

concentration. Furthermore, the dense liquid phase potentially acts as a reservoir feeding the dilute phase with protein molecules and thus stabilizing the thermodynamic conditions during the ongoing crystal growth.

8.4.3 Connection of Two-Step Crystallization to the Phase Diagram

In addition to the two-step nucleation mechanism (Vekilov, 2010b; Wolde and Frenkel, 1997; Talanquer and Oxtoby, 1998), we have discussed an alternative pathway of two-step crystallization via cluster precursors. Considering the complexity of protein phase behavior, we expect both to be relevant under suitable conditions, which we discuss with the schematic phase diagram in Fig. 8.7. Vliegenthart and Lekkerkerker (2000) elaborated on earlier findings of George and Wilson (1994) to provide a criterion for optimum crystallization conditions based on the reduced second virial coefficient. Besides the metastable critical point enhancing nucleation via critical density fluctuations (Wolde and Frenkel, 1997), the identified optimum conditions in terms of the reduced second virial coefficient correspond to a temperature window, i.e. both the dilute and the dense phase (indicated with gray shading). Considering the lower nucleation barrier in the dense phase, the optimum conditions for protein crystallization should be represented by the dense solution

as predicted in the two-step nucleation mechanism (Vekilov, 2010b; Wolde and Frenkel, 1997; Talanquer and Oxtoby, 1998), although in practice, such condition often leads to poly-crystallites instead of single crystals, which are needed for the crystallographic study. These conditions, however, are not necessarily accessible for all systems, since arrested phase behavior may render the high-concentration coexistent phase a gel instead of an equilibrium solution (Lu et al., 2008; Gibaud et al., 2011; Gibaud and Schurtenberger, 2009; Sciortino and Zaccarelli, 2011; Bianchi et al., 2006). While for isotropic interaction potentials, a metastable LLPS is generally arrested (Lu et al., 2008; Gibaud and Schurtenberger, 2009), anisotropic and patchy spheres have been shown to phase separate completely into a metastable "gas" and "empty liquid" phase (Sciortino and Zaccarelli, 2011; Bianchi et al., 2006). At least for the arrested systems, however, protein clusters in the dilute phase could step in as precursors, providing an alternative (even dominate) way of crystallization. Interestingly, the cluster phase is expected to overlap with the optimum crystallization conditions based on the second virial coefficient. The proposed mechanism via cluster precursors could thus explain the frequent observation of crystal growing in the dilute phase.

Finally, since there appear to be different views in the literature regarding the terminology of the "two-step" crystallization process, we would like to comment on the different scenarios. In protein solutions, depending on the position in the phase diagram (labeled in Fig. 8.7 as a, b and c), at least three scenarios for the intermediate state could be realized: (a) a transient density fluctuation or cluster near the critical point or near the G–L binodal, this scenario is supported by simulations and theoretical studies (Wolde and Frenkel, 1997; Oxtoby, 1998; Lutsko, 2012); (b) a metastable co-existence of liquid phases, which has been observed in several protein systems (Galkin et al., 2002; Vivarès et al., 2005); and (c) protein clusters as shown in our system and other proteins in solution (Pan et al., 2007, 2010). In real protein solutions, these mechanisms might also occur at the same time, rendering a deeper understanding of protein crystallization an interesting challenge for future research.

8.5 Conclusions and Outlook

In summary, our SAXS study of the structure and crystallization from bovine β-lactoglobulin (BLG) solutions in the presence of YCl_3 leads to the following conclusions. First, near the phase boundary, the balance between the bridging effect of yttrium cations and effective charge inversion of proteins leads to the formation of small protein clusters which have a pre-crystalline structure. Second, in situ SAXS measurements on the crystallization together with previous observation by optical microscopy indicate that these protein clusters can form both a dense liquid phase and crystals. The crystals occur later compared to the dense liquid phase due to the higher energy barrier of nucleation. This procedure leads to the competing coexistence of these two structures. Third, due to the metastable character of the dense liquid phase, it re-dissolves during the ongoing crystal growth and disappears

in the post-crystal growth stage. The observed two-step crystal growth procedure suggests that clusters play the main mechanistic role for nucleation, while LLPS allows the system to access and stabilize suitable conditions.

Acknowledgements

We acknowledge the useful discussion with G. Zocher (IFIB, Universität Tübingen) and T. Narayanan (ESRF, Grenoble, France). We acknowledge financial support from Deutsche Forschungsgemeinschaft (DFG) and beamtime allocation at ESRF, Grenoble, France.

Chapter 9

Proteins as Particles with Ion-Activated Attractive Patches

Chapter 9 is based on the following publication:

Proteins as particles with ion-activated attractive patches
Felix Roosen-Runge, Fajun Zhang, Frank Schreiber and Roland Roth

submitted and under review

Contributions:

Research design	FRR, RR
Model development and calculations	FRR, RR
Interpretation	FRR, FZ, FS, RR
Paper writing	FRR, FS, RR

Roosen-Runge et al. (2013b)

We introduce an explicit mechanism of patchy interaction between proteins, thereby demonstrating that proteins can be modeled by patchy particles. Experimentally, ion bridges of multivalent cations between protein molecules have been observed in protein crystals. Modeling this mechanism via particles with ion-activated attractive patches, we evaluate the resulting phase behavior. A broad variety of experimental results for protein solutions with multivalent cation is reproduced and explained very naturally by our model, including charge inversion, reentrant condensation, metastable liquid-liquid phase separation, cluster formation and different pathways of crystallization. The good agreement indicates that the proposed mechanism of ion bridging can be understood with attractive patches, and vice versa protein-cation solutions represent a natural model system for patchy particles.

9.1 Introduction

Patchy particles represent elegant models to explore the statistical physics of soft matter systems with directional interactions, in particular via the Wertheim theory for associating fluids (Jackson et al., 1988; Chapman et al., 1988; Wertheim, 1984a,b, 1986a,b). Exploiting different choices of the patch-patch interaction, novel phenomena such as empty liquids (Bianchi et al., 2006), reentrant network formation (Russo et al., 2011) and control of crystallization pathways (Whitelam, 2010a; Fusco and Charbonneau, 2013) have been found in calculations and simulations. In several theoretical studies, patchy particles have been suggested to represent a model system for proteins (Sear, 1999; Kern and Frenkel, 2003; Gögelein et al., 2008; Fusco and Charbonneau, 2013; Whitelam, 2010a). However, the connection between patchy particles and proteins has been based on rather general considerations concerning the non-spherical shape and inhomogeneous charge and hydrophobicity surface patterns, as well as indirect indications such as low density crystals and low critical volume fractions (Bianchi et al., 2011).

In this Letter, we present a mechanism for ion-activated attractive patches between proteins which is directly supported by experimental evidence. Protein crystals grown from solutions with multivalent cations show ion bridges between different protein molecules (Fig. 9.1B) (Zhang et al., 2011). These ion bridges are formed due to coordinative bonds of trivalent metal ions such as Yttrium(III) with solvent-exposed carboxylic side chains on the protein surface. Thus, by introducing multivalent metal cations, attractive patches on the protein surface are activated, as explained below, and can be used to tune the interactions and the associated phase behavior of globular proteins. Indeed, trivalent cations have been found to induce a rich phase behavior and non-classical nucleation processes in solutions of globular proteins, including inversion of the protein surface charge (Fig. 9.1A), reentrant condensation, liquid-liquid phase separation (LLPS) (Fig. 9.1C), cluster phases, and one-step as well as two-step pathways of protein crystallization (Zhang et al., 2008, 2011, 2012c). Importantly, the ubiquity of solvent-exposed carboxylic side chains in globular proteins implies that the presented mechanism applies to a large part of the protein family, statistically almost 50 % (Zhang et al., 2010).

The phase behavior of protein solutions has been studied experimentally under different conditions of salt and other additives. In particular, LLPS (Thomson et al., 1987; Asherie et al., 1996; Muschol and Rosenberger, 1997; Annunziata et al., 2003; Wentzel and Gunton, 2007; Gögelein et al., 2008), cluster phases (Stradner et al., 2004) and solvent-controlled crystallization pathways (Wolde and Frenkel, 1997; Galkin and Vekilov, 2000b) have been found and are usually discussed in terms of general arguments involving screened Coulomb interaction or depletion effects. For our system, we propose a model based on the formation of ion bridges that accounts for all these features in a very natural way. Although charge effects are relevant in the experimental system, the effective interactions introduced by our model reproduce the observed phenomena without explicitly accounting for electrostatics.

This simplification allows an analytical evaluation of the ion-bridging mechanism as a reference for further experimental and theoretical studies as well as simulations, which are potentially computationally very expensive, for this system.

9.2 Particles with Ion-Activated Attractive Patches

Within our approach proteins are modeled as particles with m patches per particle. Multivalent cations are modeled as bridge particles, which can bind to a patch and thereby activate it. Other ions are not considered here. The interaction between proteins is given by the hard sphere repulsion V_{HS} between the particles and attractive square-well attractions between the patches: (Jackson et al., 1988)

$$V(1,2) = V_{HS}(R_{12}) + \sum_{i=1}^{m}\sum_{j=1}^{m} V_{pp}(r_{12}^{ij}), \tag{9.1}$$

$$V_{pp}(r_{12}^{ij}) = \begin{cases} -\varepsilon_{pp} & \text{for } r_{12}^{ij} < r_c \\ 0 & \text{for } r_c \leq r_{12}^{ij}. \end{cases} \tag{9.2}$$

r_{12}^{ij} is the distance between the centers of patch i of particle 1 and patch j of particle 2.

A key ingredient in our model is the variable square-well depth ε_{pp} which is generated by the activation of patches through the binding of multivalent cations to the surface of the proteins with binding energy ε_b. We consider the binding sites to be independent from each other. Since a binding site can either be unoccupied (u), or occupied (o) by at most one particle, the average occupancy Θ of each binding site is given by the statistics of a two-state system in the grand canonical ensemble:

$$\Theta = \frac{1}{\exp(\beta(\varepsilon_b - \mu_s)) + 1}. \tag{9.3}$$

The chemical potential μ_s of the ions is related to its reservoir concentration via $c_s^r \approx \exp(\beta\mu_s)$ with $\beta = (k_B T)^{-1}$.

The patch-patch interaction energy ε_{pp} is determined from three contributions (see Fig. 9.1D,E) depending on the average occupancy Θ of two interacting patches:

$$\varepsilon_{pp} = \varepsilon_{uu}(1-\Theta)^2 + 2\,\varepsilon_{uo}\Theta(1-\Theta) + \varepsilon_{oo}\Theta^2. \tag{9.4}$$

The first term accounts for the interaction between two unoccupied patches, which meet with a probability of $(1-\Theta)^2$. The second term represents the ion-bridge attraction, when an unoccupied and an occupied patch become cross-linked, with a probability of $2\Theta(1-\Theta)$. The third term describes the interaction between two occupied patches, with a probability of Θ^2. In order to recover a purely repulsive system for fully occupied or unoccupied

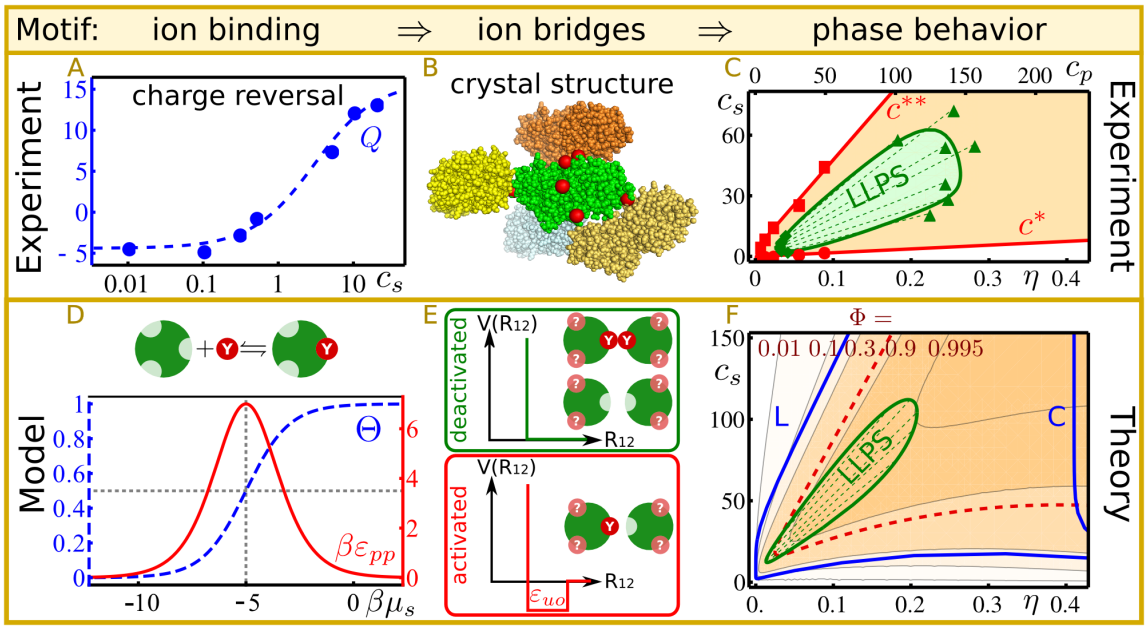

Figure 9.1: Multivalent ions have been experimentally found to bind to specific sites on the surface of globular proteins, causing a charge reversal (A) and forming ion bridges in crystal structures (red spheres in B). The isothermal phase diagram shows a reentrant condensation with salt concentration c_s in the range $c^* < c_s < c^{**}$ and a homogeneous solution otherwise. In the condensed regime, a metastable liquid-liquid phase separation (LLPS) is found at low volume fractions (C). The results on the binding and bridging of ions inspire a model involving particles with ion-activated attractive patches. The average occupancy Θ [Eq. (9.3)] reflects the charge reversal (D). From the ion-bridge energy ε_{uo} and the probability that an occupied and unoccupied patch meet (E), the ion-induced patch-patch attraction strength ε_{pp} [Eq. (9.5)] results as a non-monotonous function of the chemical potential μ_s that represents the ion concentration in the reservoir (D). The resulting phase diagram reproduces all experimental observations (F). The LLPS occurs at similar volume fractions and salt concentrations (green closed area) and is metastable with respect to the liquid-crystalline transition (blue solid line). Ion bridges build clusters, as indicated via the cluster fraction Φ [Eq. (9.8), orange contour plot] and the percolation line (red dashed line), beyond which dynamics might by slowed down. The connection to the reentrant condensation and physical implications of the percolation are discussed in the text. Details: The concentrations of salt, c_s, and protein, c_p, are given in units of mM and g/l, respectively. A: Human serum albumin (5 g/l) with YCl_3. Surface charges were obtained as described in Ref. (Roosen-Runge et al., 2013a). The dashed line corresponds to a Langmuir isotherm $Q = Q_0 + 3N\Theta$ with $N = 6.8$, $Q_0 = -4.4$ and $\beta\varepsilon_b = -5.9$. B: Crystal structure for dimers of β-lactoglobulin with Yttrium(III) (PDB 3PH6) (Zhang et al., 2011). For clarity, only 4 out of 8 neighboring dimers are shown. C: Human serum albumin with YCl_3. Data are replotted from Zhang et al. (2012c). The volume fraction η is calculated via an effective sphere diameter (Roosen-Runge et al., 2011).

patches, we choose $\varepsilon_{uu} = \varepsilon_{oo} = 0$. Then ε_{pp} reads

$$\beta\varepsilon_{pp} = \frac{2\,\beta\varepsilon_{uo}\exp(\beta\varepsilon_b - \beta\mu_s)}{(\exp(\beta\varepsilon_b - \beta\mu_s) + 1)^2}. \tag{9.5}$$

The experimental control parameter is the total salt concentration c_s, which is given in our model by

$$c_s = m\Theta\rho + \exp(\beta\mu_s)\left(1 - \eta(1 + R_s/R_p)^3\right), \tag{9.6}$$

where R_s and R_p are the (effective) radii of salt ions and proteins, respectively. The first term accounts for the ions bound to proteins. The second term originates from the free ions in the solution, corrected for the volume excluded to ions (Roth et al., 2001). The number density ρ is related to the protein volume fraction $\eta = \frac{4}{3}\pi R_p^3\rho$.

For our calculations we set the number of patches $m = 4$, according to the number of ion bridges per monomer in the protein crystal. The centers of the patches are located within the particle at a distance $d = 0.9\,R_p$ from the particle center. The range of the square-well is chosen to be $r_c = 0.33\,R_p$, which is sufficiently small to ensure that a patch can form at most one bond. The ratio between the ion and protein radii is $R_s/R_p = 2\,\text{Å}/36\,\text{Å} = 1/18$. The bridge energy $\beta\varepsilon_{uo} = 14$ and binding energy $\beta\varepsilon_b = -5$ are chosen inspired by the experimental behavior of HSA and YCl_3 (cf. Langmuir isotherm in Fig. 9.1A).

9.3 Liquid-Liquid and Solid-Liquid Phase Coexistence

The liquid-liquid and solid-liquid phase coexistence at temperature T follows from chemical and mechanical equilibrium, $\mu_1 = \mu_2$ and $P_1 = P_2$, respectively, between phases 1 and 2. The chemical potential $\mu = \partial f/\partial\rho|_{VT}$, pressure $P = \rho\mu - f$, and isothermal compressibility $\chi_T = 1/\rho(\partial P/\partial\rho)^{-1}$ are calculated analytically from the free energy density f that is given for the fluid by the Wertheim theory (Jackson et al., 1988) and for the solid by a related cell model assuming the formation of all m ion bridges (Sear, 1999).

The main feature of our model is the formation of ion bridges: If an occupied patch interacts with an unoccupied patch, an ion bridge forms and links the participating proteins (Fig. 9.1E). Based on the attractive interaction induced by the ion bridges, several phenomena in the system become intuitively clear. At low ion concentrations ($\Theta \to 0$) the proteins repel each other – in the model by the hard-sphere interaction and in the experimental system through electrostatic repulsion. As more ions are added to the system more binding sites become occupied, i.e. Θ increases, which in the experimental system results in a reduced net charge (Fig. 9.1A,D). For $\Theta < 0.5$ the addition of ions increases the attraction between proteins, since the probability for an occupied and an unoccupied patch to meet increases. Further increasing the ion concentration ($\Theta > 0.5$) decreases the attraction since too many binding sites are already occupied, thereby reflecting the charge reversal (Fig. 9.1A,D). At high enough ion concentration, the probability for ion-bridge formation is low and proteins mainly repel each other, corresponding to the reentrant homogeneous phase.

If the attraction is sufficiently strong and the protein density ρ is in the right range, a LLPS occurs in a closed area (green line in Fig. 9.1C,F). Importantly, the low volume fractions of the coexistent liquid phases can be reproduced with patchy particles, which is not possible using only isotropic potentials (Noro and Frenkel, 2000). As expected for protein solutions, the LLPS is found to be metastable with respect to a crystal phase (blue line in Fig. 9.1F). We find excellent agreement between theory and experiment (Fig. 9.1C,F).

9.4 Cluster Formation

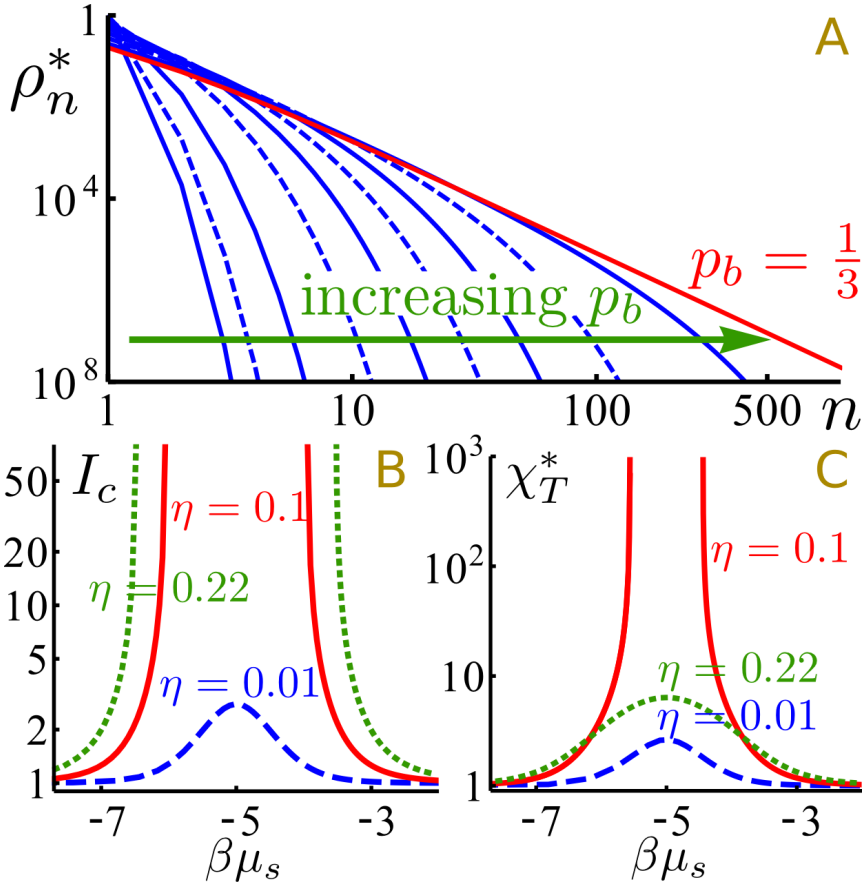

Figure 9.2: A: The cluster size distribution $\rho_n^* = \rho_n/\rho$ broadens with increasing bond probability (blue lines: p_b = 0.0001, 0.001, 0.01, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3; red line: $p_b = p_b^* = 1/3$), implying the growth of larger clusters, in particular when approaching the percolation limit p_b^*. B: The integrated scattering power of all clusters, normalized to the monomer solution, shows a steep increase at intermediate chemical potentials μ_s, i.e. intermediate salt concentrations. This behavior represents one reason for the turbidity in the solution during the reentrant condensation. C: Another reason for turbidity is opalescence close to the LLPS boundary and the LLPS itself. The forward scattering intensity is proportional to the isothermal compressibility $\chi_T^* = \chi_T/\chi_T^{HS}$, normalized to the hard sphere value, that outlines the reentrant effect. The calculations in B, C are performed for three volume fractions below, through and above the LLPS.

A natural consequence of the ion bridges is the formation of clusters. Using the Flory-Stockmeyer theory (Flory, 1941; Bianchi et al., 2008) the number density ρ_n of n-clusters and the cluster fraction Φ are given by

$$\rho_n = \rho(1-p_b)^m[p_b(1-p_b)^{m-2}]^{n-1}\frac{m(mn-n)!}{(mn-2n+2)!n!} \quad , \tag{9.7}$$

$$\Phi = \frac{\rho-\rho_1}{\rho} = 1-(1-p_b)^m \quad . \tag{9.8}$$

With increasing average ion-bridge probability p_b, which is provided by the Wertheim theory (Jackson et al., 1988), larger cluster become more frequent (Fig. 9.2A). Once p_b exceeds the percolation value $p_b^* = 1/(m-1)$, a system-spanning cluster can form, implying that dynamics in the solution might be slowed down or even arrested.

9.5 Concluding Discussion

The formation of clusters provides an explanation for the experimentally observed turbidity in the condensed regime between c^* and c^{**} (red lines in Fig. 9.1C). Using the approximation of Rayleigh scattering, the scattering of n-clusters grows with $V^2 \propto n^2$, which causes a steep increase of the integrated scattering power $I_c = \sum_{n=1}^{\infty} n^2\rho_n$ with increasing p_b, in particular when approaching the percolation threshold p_b^* (Fig. 9.2B). In addition, as the liquid-liquid phase boundary is approached, opalescence causes turbidity in the solution, being manifested in the increase of the isothermal compressibility of the patchy particle solution (Fig. 9.2C).

The LLPS in combination with cluster formation and, in particular, percolation helps to rationalize the experimental observations that crystals grow around the dilute boundary of the LLPS, but not in the dense solutions, although the latter should be more suitable for crystallization according to classical nucleation theory (Zhang et al., 2012c,a). Two theoretical results can be used to explain this behavior: First, the dense phase occurs beyond the percolation line, possibly hindering motions necessary to reorganize from disordered clusters to ordered crystals. Second, a considerable amount of clusters is present around the dilute coexistence boundary. These clusters might act as precursors in two-step nucleation pathways that occur much faster than the classical one-step nucleation from homogeneous solution (Zhang et al., 2012a).

To conclude, we have presented a comprehensive interpretation of the experimental observations in protein solutions that is based on an experimentally supported mechanism for ion-activated attractive patches. The nature of the patchy interaction, namely ion bridges between neighboring molecules, allows validation and application of the patchy particle model in real protein solutions. The cation concentration represents an independent control parameter. These findings promise control of the phase behavior of protein solutions using multivalent cations and the statistical physics of particles with ion-activated patches. Beside proteins solution, the model presented in this Letter might also prove useful for other soft matter systems that allow for bridge formation.

Acknowledgements

We thank Martin Oettel (Tübingen) for stimulating discussions on the manuscript. F.R.-R. acknowledges a fellowship of the Studienstiftung des deutschen Volkes. The work was supported by the Deutsche Forschungsgemeinschaft (DFG).

Appendix: Details of the Calculation

Wertheim Theory for Patchy Particles

The Wertheim theory for associating liquids (Jackson et al., 1988; Chapman et al., 1988; Wertheim, 1984a,b, 1986a,b) provides an elegant perturbation approach to analytically calculate properties for particles with m attractive patches.

The Wertheim theory is limited to the case of only one bond per patch. This condition is fulfilled due to the steric hindrance of hard core repulsion, if the attraction range r_c is smaller than $r_c^{(\max)} = \sqrt{d^2 - 2\sqrt{3}dR_p + 4R_p^2} - d$, which for the parameter choice in this study returns $r_c^{(\max)} = 0.40\, R_p$. Since the chosen interaction range is $r_c = 0.33\, R_p < r_c^{(\max)}$, the condition of only one bond is fulfilled.

The basic result of the Wertheim theory is the bonding probability from which the free energy and other properties can be calculated. The bonding probabilities for m different patches follow from the mass-action equation (Jackson et al., 1988)

$$(1 - p_i) = \left(1 + \sum_{j=1}^{m} \rho (1 - p_i) \Delta_{ij} \right)^{-1} \quad , \tag{9.9}$$

which considerably simplifies under the assumption that all patches are similar: (Bianchi et al., 2008)

$$\frac{p_b}{(1 - p_b)^2} = \rho m \Delta_{ij} \quad , \tag{9.10}$$

where an overall bonding probability $p_b \equiv p_i$ is introduced. The perturbation integral from the patches is given by

$$\Delta_{ij} \;=\; 4\pi \int_0^\infty g_{\mathrm{HS}}(R_{12}) \, \langle f_{ij}(1,2) \rangle R_{12}^2 \, dR_{12} \quad . \tag{9.11}$$

$\langle f_{ij}(1,2) \rangle$ denotes the average of the Mayer function $f_{ij}(1,2) = \exp(-\beta V(1,2)) - 1$ for all relative orientations of the two particles at fixed distance R_{12}. Using the Carnahan-Starling contact value for the pair correlation function of the hard sphere reference system with volume fraction η and $\sigma = 2R_p$,

$$g_{\mathrm{HS}}(\sigma) = \frac{1 - \frac{1}{2}\eta}{(1 - \eta)^3} \quad , \tag{9.12}$$

Eq. (9.11) can be approximated (Jackson et al., 1988):

$$\Delta_{ij} \approx 4\pi g_{\mathrm{HS}}(\sigma)\, K_{AB}(\sigma, r_c, d)\, F_{AB}(\beta\varepsilon_{\mathrm{pp}}) \tag{9.13}$$

$$F_{AB} = \exp(\beta\varepsilon_{\mathrm{pp}}) - 1 \quad , \tag{9.14}$$

$$K_{AB} = \frac{\sigma^2}{72d^2}\Bigg((2d + r_c - \sigma)(22d^2 - 5dr_c - 8r_c^2 - 7d\sigma + r_c\sigma + \sigma^2) + 6(d - r_c)(2d + r_c)^2 \log\left[\frac{\sigma}{2d + r_c}\right]\Bigg) \quad . \tag{9.15}$$

Evaluation of Phase Behavior

Given the bond probability p_b, the phase behavior can be calculated by evaluating thermodynamic functions. A coexistence of two phases is found under three conditions:

$$T_1 = T_2 \quad , \qquad \mu_1 = \mu_2 \quad , \qquad P_1 = P_2 \quad . \tag{9.16}$$

Free Energy of Liquid Phase The free energy density of the patch bonds between the particles is given by: (Jackson et al., 1988; Bianchi et al., 2008)

$$\frac{\beta F_{bond}}{V} = \rho \sum_{i=1}^{m} \left(\ln(1 - p_i) + \frac{1}{2}p_i\right) \tag{9.17}$$

$$= \rho m \left(\ln(1 - p_b) + \frac{1}{2}p_b\right) \tag{9.18}$$

with the particle number density $\rho = N/V$ and $\beta = (k_B T)^{-1}$. Both the chemical potential μ and the pressure P consists of the reference part from the hard sphere system and a part arising from the patch attraction:

$$\beta\mu = \frac{\eta(8 + 3(\eta - 3)\eta)}{(1 - \eta)^3} + \ln[\eta] + \beta\mu_{\mathrm{bond}} \quad , \tag{9.19}$$

$$\beta P = \rho\frac{1 + \eta + \eta^2 - \eta^3}{(1 - \eta)^3} + \rho\beta\mu_{\mathrm{bond}} - \frac{\beta F_{\mathrm{bond}}}{V} \quad , \tag{9.20}$$

$$\beta\mu_{\mathrm{bond}} = \frac{d}{d\rho}\left(\frac{\beta F_{\mathrm{bond}}}{V}\right) \tag{9.21}$$

$$= \frac{3\eta^2(2\eta - 5)F_{AB}K_{AB}f^2}{(\eta - 1)^4 z^2 R^3} + m\ln\left(\frac{2}{z}\right) \quad ,$$

$$z = 1 + \sqrt{1 + \frac{6(\eta - 2)\eta F_{AB}K_{AB}m}{(\eta - 1)^3 R^3}}$$

Free Energy of Solid Phase For the solid phase, we used a cell model assuming that all m bonds are formed in a crystal (Sear, 1999). The free energy density $f_c(\rho)$ is calculated from the restricted volume for free motion V_{free} in radial and angular direction at particle distance r:

$$f_c(\rho) = -\rho \log[V_{\text{free}}(R_p(\rho_c/\rho)^{1/3})] - \rho\frac{m}{2}\varepsilon_{pp} \tag{9.22}$$

$$V_{\text{free}}(r) = -4\frac{(r-R_p)(r-d-r_c)}{(d+r_c-R_p)R_p}\,\frac{r_c^2-(r-d)^2}{R_p^2} \quad . \tag{9.23}$$

The density ρ_c is given by the volume fraction of a simple cubic crystal $\eta_c = \pi/6 = \rho_c R_p^3 4\pi/3$.

Chapter 10

Remarks on Modeling of Protein–Salt Solutions

This chapter is designated to an overview of the modeling of protein–salt solutions that is employed and developed throughout the present thesis. First, we briefly outline the challenges and potential difficulties for the modeling of proteins. In particular, protein properties such as non-sphericity, roughness and conformational flexibility have to be evaluated carefully, before further modeling can be performed. Second, we discuss the modeling of proteins in terms of colloidal models. Colloidal models represent an interesting reference system, since they allow a fruitful connection of observed phenomena in protein solutions to predictions based on statistical physics and colloid theory. Third, the modeling of the interaction between proteins and salt is reviewed. The specific interaction of cations with the protein surface can be very specific and complex, making the extraction of the essential features for a conceptual understanding very interesting and educative.

This chapter complements modeling aspects directly related to studies from the present thesis with information from other studies that are not included in the thesis, but were performed during the time of the thesis (Roosen-Runge et al., 2011; Heinen et al., 2012).

10.1 Proteins as a Challenge for Modeling

Proteins exhibit a complex variety of interactions. Starting with the excluded volume, globular proteins are non-spherical objects which are to a certain extent soft, flexible and permeable. Short-range interactions arise from van der Waals interactions as well as water-mediated interactions due to hydrophobic and hydrophilic surface patterns. Long-range Coulomb interactions are based on a non-homogeneous surface charge distribution, implying that besides the monopole term also dipole and higher order multipoles can have an effect. Protein–protein and protein–ligand interactions are highly specific in many cases, as evidenced in situations of molecular docking and key-lock mechanisms (Permyakov, 2009).

Different computational and theoretical approaches have been used to represent the non-spherical protein with highly anisotropic interactions in theoretical terms. Using simulations such as Monte Carlo approaches and molecular dynamics, properties of single proteins have been studied in full atomic detail. This detailed approach fails for reasons of computational time as soon as many proteins are involved, which is inevitable for the study of phase behavior and protein aggregation. Possible schemes with frozen and coarse-grained protein structures as well as implicit water models enable the study of protein ensembles, relying on the assumption that the coarse-grained models are chosen such that the essential features are reproduced.

Different from these "bottom-up" approaches which build suitable models starting from the details of the system, also a "top-down" approach can be used. Given a theoretical well-characterized system, the experimental system can be mapped on theoretical parameters in a way that the features of interest are correctly reproduced. In this context, colloids provide a suitable model system that has been studied intensively. Most analytical approaches from colloid theory are based on hard spheres with additional isotropic interactions. A comparison of protein solutions with predictions from colloid theory, thus, starts with the non-trivial question how to represent proteins with a non-spherical, rough and flexible shape and anisotropic interactions in colloidal models.

10.2 Modeling of Proteins

10.2.1 Conformational Stability of Proteins

Proteins are biopolymers and, thus, can react on environmental factors such as temperature, salts or other additives by the variation of the geometrical conformation and even a complete unfolding of the native globular structure. If a model in colloidal terms is desired, further information on the conformational flexibility is necessary. Ideally, the protein conformation should be stable and invariant for all experimental conditions investigated.

In general, the globular proteins used in this work have a conformation robust against changes of salt concentration and pH. For BSA and HSA, transitions between globular states—the "normal", "fast" and "extended" conformation—has been found at acidic pHs below 4 (Leggio et al., 2008; Dockal et al., 2000; Barbosa et al., 2010; Peters, 1996). Although this transition might change details, a global account of BSA and HSA with colloidal particles as performed in Ch. 5 should be sensible for two reasons. First, the transition is at very low pH which is outside the salt concentration range in which charge inversion and reentrant condensation occur. Second, both conformations are globular shapes, which might be not too different in effective colloidal terms. For even lower pH values, another transition to an extended conformation takes place and the molecules become partly unfolded at low concentrations, while they seem to stay globular for higher concentrations due to crowding effects (Barbosa et al., 2010).

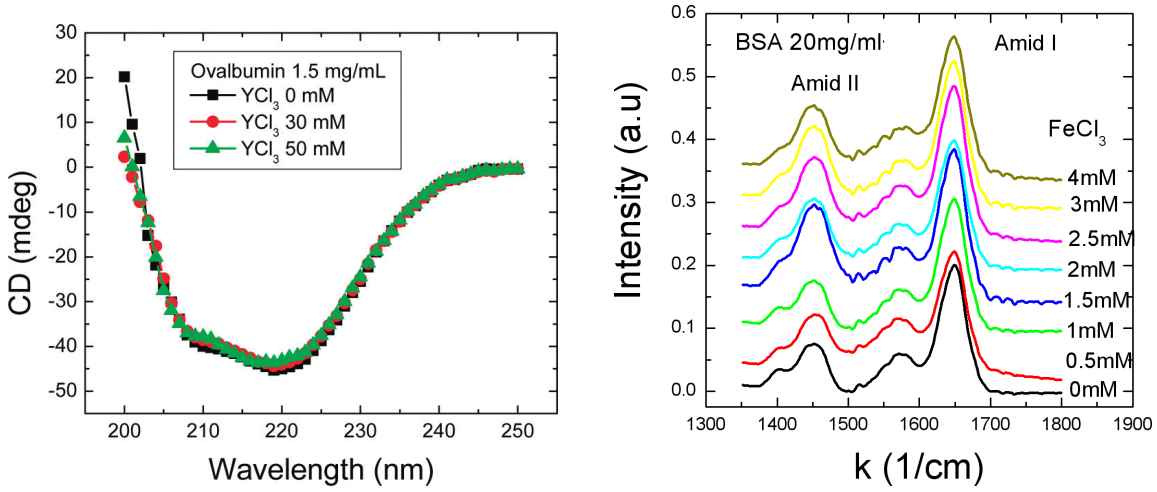

Figure 10.1: Optical spectroscopy evidences the invariance of the protein secondary structure throughout the phase diagram. Left: Spectra from circular dicroism (CD) for Ovalbumin in the presence of YCl_3 (figure taken from Ianeselli et al. (2010)). Right: Fourier transform infrared (FTIR) spectra for BSA in the presence of $FeCl_3$ (data measured by Z. Öztürk).

In order to monitor the protein structure on the level of the secondary structure, spectroscopic techniques at different salt conditions have been used. Both circular dichroism (CD) and Fourier transform infrared (FTIR) spectroscopy evidence an invariance of the protein secondary structure throughout the reentrant phase diagram (Fig. 10.1). Note that the FTIR spectra are taken with $FeCl_3$ in which case not only salt, but also pH variations could vary the protein structure.

The SAXS results via the Kratky plots $I(q)q^2$ (e.g. Fig. 4.2) show no sign of a significant increase for high q, thereby evidencing that the geometry of the protein shape is not altered from a globular state and thus no unfolding and no random coil formation is present in the protein–salt solutions.

The overall results suggest that the proteins can indeed be described with a stable globular conformation under the investigated conditions. The next subsection outlines by which means proteins can be represented in terms of effective spheres.

10.2.2 Accounting for Non-Sphericity: Effective Spheres

A conventional way for a representation of a protein is to map the non-spherical shape to an effective sphere (Jennings and Parslow, 1988). The exact method how to choose an effective sphere cannot be universal since it, in principle, depends on the phenomenon of interest. Here we summarize arguments which have been used in previous publications (Roosen-Runge et al., 2011; Heinen et al., 2012) and studies included in this thesis (Chapter 4).

In order to account for effects of non-sphericity with only one parameter, it is convenient to consider ellipsoids of revolution. Dividing the polar axis a by the (double) equatorial axis b, one obtains the aspect ratio p as a suitable measure of non-sphericity:

$$p = \frac{a}{b} \, .$$

For hard ellipsoids of revolution, the mapping to effective hard spheres can be calculated for different observables based on well-established relations from statistical physics. Naturally, similar relations can also be obtained, in principle, for triaxial ellipsoids, rods, cylinders or other shapes – here, we restrict the overview to the case of ellipsoids of revolutions outlining the general strategy to obtain effective spheres.

Importantly, the usage of ellipsoids of revolution also opens an experimental way to extract appropriate dimensions and aspect ratios via the fitting of the protein form factor observed by small-angle scattering. This method has performed well in comparing results from protein solutions to theoretical predictions of hard spheres for topics ranging from statics like protein interaction (Zhang et al., 2007, 2012b) to dynamics like diffusion and viscosity (Roosen-Runge et al., 2011; Heinen et al., 2012).

Particle Volume

The simplest way to an effective radius is to determine the radius of a sphere that has the same volume as the ellipsoid. The derived geometric effective radius reads

$$R_{\mathrm{geo}} = (ab^2)^{1/3} = bp^{1/3} \quad . \tag{10.1}$$

Since the simple geometric volume does not account for thermo- and hydrodynamic effects of non-sphericity (Svergun et al., 1998; Zhang et al., 2012b), it should be considered as a lower boundary to the effective sphere diameter (Heinen et al., 2012).

Excluded Volume Interaction: Isihara Sphere

From the viewpoint of interactions, the most natural comparison of hard ellipsoids with hard spheres is the excluded-volume interaction. In particular, the second virial coefficient of hard ellipsoids has been calculated by Isihara (1950):

$$B_2 = 4\left(\frac{4\pi}{3}ab^2\right) f \quad , \tag{10.2}$$

$$f = \frac{1}{4} + \frac{3}{16}\left(1 + \frac{\arcsin(\epsilon)}{\epsilon\sqrt{1-\epsilon^2}}\right)\left(1 + \frac{1-\epsilon^2}{2\epsilon}\log\frac{1+\epsilon}{1-\epsilon}\right) \quad , \tag{10.3}$$

$$\epsilon = \begin{cases} 1 - p^2 & \text{for } p < 1 \quad , \\ 1 - p^{-2} & \text{for } p > 1 \quad . \end{cases} \tag{10.4}$$

Equating B_2 with the second virial coefficient of a hard sphere $B_2^{(\mathrm{HS})} = 4\left(\frac{4\pi}{3}R^3\right)$, the effective radius is derived directly as

$$R_{B_2} = b(fp)^{1/3} \quad . \tag{10.5}$$

This effective radius is also frequently used in the data analysis of small-angle scattering to represent a sensible effective sphere based on the form factor for the fitting of the structure factor (Kline, 2006; Pedersen, 1997).

Note at this point that we restricted ourselves to analytical estimations that allow a straight-forward and experimentally grounded analysis via the ellipsoid model. The excluded volume interaction between two macromolecules can also be evaluated numerically based on the crystal structure, yielding a comparable increase of the effective radius compared to the geometric radius (Neal and Lenhoff, 1995).

Self-Diffusion: Perrin Factors for Hydrodynamic Radius

The effective sphere for a hard ellipsoid with respect to self-diffusion has been derived by Perrin (1934), with corrections pointed out by Koenig (1975). The Perrin factor is used to calculate the hydrodynamic radius R_h of an ellipsoid of revolution. This radius is inversely proportional to the diffusion constant D_0 as specified in the Stokes-Einstein relation

$$D_0 = \frac{kT}{6\pi\eta R_h} \quad , \tag{10.6}$$

where kT is the thermal energy, and η is the solvent viscosity.

The diffusion coefficient of non-spherical particles depends on the orientation of the particle with respect to the direction of motion. Averaging over all orientations, one obtains for an ellipsoid of revolution (Jennings and Parslow, 1988; Ferrer et al., 2001; Roosen-Runge et al., 2011):

$$R_h = \begin{cases} b\sqrt{1-p^2} \,/\, \arctan\left[\frac{1}{p}\sqrt{1-p^2}\right] & \text{for } p < 1 \\ b\sqrt{p^2-1} \,/\, \log\left[p+\sqrt{p^2-1}\right] & \text{for } p > 1 \end{cases} \quad . \tag{10.7}$$

Importantly, the hydrodynamic radius can be directly measured by dynamic light scattering in the limit of dilute concentrations that, in combination with the rotational diffusion, indirectly provides experimental access to the shape of the protein (Ferrer et al., 2001).

Intrinsic Viscosity

The intrinsic viscosity

$$[\eta] = \lim_{\varphi \to 0} \frac{\eta(\varphi) - \eta(0)}{\eta(0)\varphi} \tag{10.8}$$

is another dynamic quantity that allows the estimation of an effective radius. For hard spheres, the Einstein result $[\eta]^{(\mathrm{HS})} = 2.5$ holds, whereas the intrinsic viscosity of ellipsoids

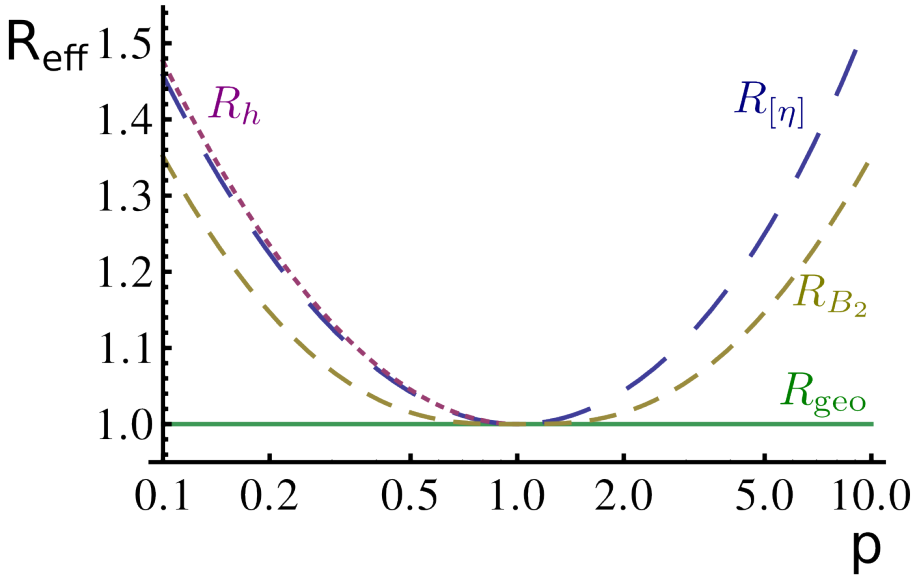

Figure 10.2: Comparison of different effective radii obtained for ellipsoid of revolutions with aspect ratio p and constant particle volume $V = 4\pi/3$. The effective radii are calculated via the hydrodynamic radius (R_h), the intrinsic viscosity ($R_{[\eta]}$), the second virial coefficient (R_{B_2}) and the geometric radius ($R_{\mathrm{geo}} \equiv 1$).

$[\eta]^{(ell)}$ is a complicated quantity derived from the average over a fourth rank tensor (Douglas and Garboczi, 1995). For oblate ellipsoids, the intrinsic viscosity can be represented in a more simple form by (van der Kooij et al., 2001; Heinen et al., 2012)

$$[\eta]^{(ell)} = \frac{5}{2} + \frac{32}{15\pi}\left(\frac{1-p}{p}\right) - 0.628\left(\frac{1-p}{1-0.075p}\right) \quad . \tag{10.9}$$

Equating the linear term in the expansion of the viscosity $\eta = \eta(0) + \eta(0)[\eta]\varphi + \mathcal{O}[\varphi^2]$, one obtains an effective radius

$$R_{[\eta]} = b\left(p\frac{[\eta]^{(ell)}}{2.5}\right)^{1/3} \tag{10.10}$$

that accounts for the non-sphericity of the solute.

Comparison and Discussion of the Different Methods

Having presented four different methods to calculate an effective radius for a hard ellipsoid of revolution, it is interesting to compare the different means. Fig. 10.2 shows the different effective radii for ellipsoids with constant volume but different aspect ratio p. The geometric radius R_{geo} does not change, while the other effective radii increase with increasing non-sphericity, i.e. an aspect ratio $p \neq 1$. The effective radii from dynamic properties—i.e. the hydrodynamic radius R_h and $R_{[\eta]}$ from the intrinsic viscosity—vary more strongly than the effective radius from the static interaction, i.e. R_{B_2}. Generally, a higher non-sphericity implies an effective sphere with a larger particle volume.

10.2.3 Accounting for Softness and Permeability of the Shape

The effects of a soft and partially permeable shape represent more complicated tasks for the modeling, depending on the specific conditions and observables. For the case of softness, hard sphere models of soft spheres have been proposed, however, with modifications such as a temperature-dependent effective radius (Protopapas et al., 1973). For other phenomena such as melting and crystallization, the picture of soft spheres differs considerably from hard spheres, thus obstructing models of effective hard spheres (Senff and Richtering, 1999). For the case of permeability, the short-time collective diffusion and sedimentation has been found to be significantly fastened with increasing particle permeability (Abade et al., 2010). Since this fastening is dependent on the particle volume fraction, effective hard spheres are not suited to reproduce the behavior and, in general, more detailed modeling of permeability is required if this feature seems important for the given phenomenon.

For the present thesis, softness and permeability have not been taken into account. Models using hard particles produced consistent results, suggesting that softness and permeability do not play an essential role for the observed phenomena.

10.3 Modeling of Protein–Ion Interaction

For several combinations of multivalent ions and proteins such as hemoglobin and Fe^{2+} or calmodulin and Ca^{2+}, specific binding configurations are known and their properties have been investigated intensively (Permyakov, 2009). However, for the globular proteins in combination with the multivalent ions investigated in this work, no such specific binding sites are known. The following paragraphs summarize the indications and assumptions used to model the interaction of protein and multivalent cations.

10.3.1 Molecular Dynamics of Detailed Protein Structure

As a first approach, molecular dynamics simulations have been performed on one β-lactoglobulin dimer in the presence of YCl_3, thereby accounting explicitly for water and counterions (Leibfarth, 2011). Since the standard force fields implemented via GROMACS do not contain suitable interaction parameter for Yttrium(III), quantum chemical calculations have been mapped to effective Lennard-Jones parameters and implemented in the simulation (Leibfarth, 2011). Fig. 10.3 depicts the final state of one simulation run, showing multidentate binding of Y^{3+} to solvent-exposed carboxylic acids. The binding of the Y^{3+} to the protein surface groups was even so strong that no detaching could be observed during the simulation runs. Although the simulations thus do not provide information on equilibrium properties, this finding clearly suggests binding of Y^{3+} to specific groups on the protein surface.

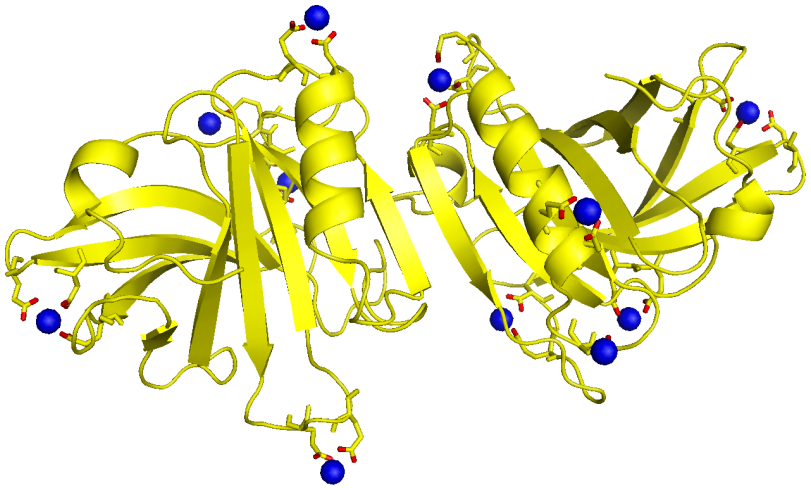

Figure 10.3: Final state of a molecular dynamics simulation for one β-lactoglobulin dimer in the presence of YCl_3. The Y^{3+} ions (blue spheres and with slightly enlarged relative size) bind with multidentate geometries to solvent-exposed carboxylic acids on the protein surface (sketched as sticks where the oxygens are red). The simulation has been performed within the diploma thesis of Leibfarth (2011).

10.3.2 Monte Carlo Sampling of Binding States

In a second step, the knowledge that surface-exposed acidic residues represent possible binding sites for Y^{3+} is used for the simulation of the overall protein charge. Using a frozen protein structure with a protonation state adjusted to pH 7, preferable binding sites are identified based on the solvation effect of the ions and a binding energy from the literature, yielding around 20 binding sites with overall positive binding constants (Zhang et al., 2008, 2010). Combining the binding constant with the electrostatic coupling between the different binding sites, binding configurations were sampled for different multivalent cations and additional monovalent salt concentrations (Ch. 6). Indeed, a charge inversion is found that can be used to rationalize the reentrant condensation due to reentrant charge stabilization (Ch. 6).

10.3.3 Coarse-Grained Analytical Binding Model

As a third approach, the ion binding mechanism evidenced from the atomic-detailed protein simulations is transfered to a coarse-grained analytical model of a spherical particle with N binding sites that are self-consistently coupled via the protein charge (Ch. 5). A qualitatively similar behavior as for the Monte Carlo sampling is found. However, the analytical tractability allows the introduction of pH effects that on the one hand change the protonation state and on the other hand couple to the hydrolysis of the multivalent metal salt. Furthermore, the trends with additional monovalent salt in the system with proteins and multivalent salt can be explained reasonably (Ch. 6).

10.3.4 Brownian Dynamics of Colloids with Ion Binding

In a forth approach, the proteins are represented by spherical colloids, and the protein–cation binding interaction is modeled with a Lennard-Jones potential. Using this model in combination with Brownian dynamics simulations, the basic mechanisms of ion binding and ion bridging between colloids can be studied and provides information on charge inversion, ion-induced colloid interaction and, potentially, phase behavior. Since this study is work in progress and no comprehensive picture has reached so far, it is not included in the thesis.

10.3.5 Ion-Activated Attractive Patches

Finally, the protein–ion interaction is reduced to the binding of ions to independent binding sites on the surface of an effective sphere representing the protein. Allowing the ions to simultaneously bind to patches of two different proteins, the effect of ion bridging on the phase behavior can be explored (Ch. 9). This simple model for the protein–ion interaction in combination with an ion-bridge model for particles with ion-activated attractive patches explains the full ion-induced phase behavior and nucleation pathways in a very natural way (Ch. 9).

10.4 Concluding Remarks on Modeling of Proteins

Although proteins are complex macromolecules with various non-trivial properties, simple colloidal models can provide suitable reference systems to compare, describe and explain phenomena in protein solutions. It should, however, be kept in mind that models are designed with a focus on specific phenomena and thus cannot be expected to represent the full reality of the protein. While the protein by itself is just one entity, numerous models are required to reproduce the protein properties.

In this thesis, different colloidal perspectives on proteins are presented and provide a consistent physical picture of the equilibrium phase behavior of protein solutions in the presence of salt.

Chapter 11

Concluding Discussion and Outlook

We have presented several studies on the ion-induced phase behavior of protein solutions. In this chapter, these results are discussed also in connection with results from the literature to achieve a comprehensive picture of the phenomena observed in protein solutions with multivalent cations.

11.1 Well-Established Salt Effects in Protein Solutions

Salt-induced effects in protein solutions have been studied for over a century. Salt effects established in the literature include the unfolding of proteins during cold denaturation, the screening of charge stabilization and the Hofmeister series on the solubility of protein solutions in the presence of different salts. These effects have also been studied for the globular proteins in the present thesis, focusing on the qualitative salt effect on the phase behavior of protein solutions.

Using different experimental techniques monitoring both the secondary structure and the overall geometrical shape and compactness, no evidence for unfolding or strong conformational changes under the salt and pH conditions investigated has been found (Sec. 10.2.1). This conformational stability allows the modeling along the lines of colloidal models as applied successfully in several studies on the dynamics of serum albumins in solution (Ferrer et al., 2001; Roosen-Runge et al., 2010, 2011; Heinen et al., 2012)(Ch. 10).

The protein interactions under different salt conditions along the Hofmeister series have been studied and also used to test the concept of effective spheres for the modeling of protein interactions (Ch. 4). The Hofmeister series is reflected in the second virial coefficients B_2 measured by the low q intensity from SAXS. In particular, B_2 decays from salting-in to salting-out conditions. After normalization of B_2 to the excluded volume contribution, salting-in conditions should correspond to additional repulsion on top of hard sphere behavior, while salting-out conditions mean additional attraction. Indeed, this behavior is found if the effective sphere is chosen such that it accounts for both the hydration shell and non-spherical shape. Using this effective sphere, the crossover from additional repulsion to additional attraction along the Hofmeister series is observed

in a model-free manner via the experimental structure factors measured by SAXS. The protein interactions along the Hofmeister series can be understood both quantitatively and qualitatively using SAXS measurements and a suitable protein model including hydration and non-sphericity (Ch. 4).

A similar concept has been used for BSA solutions with NaCl (Zhang et al., 2007), finding good qualitative agreement with expectations based on the DLVO theory for the protein interaction (cf. Fig. 1.1). For low salt concentration, the screened Coulomb repulsion dominates, while for high salt concentration, a short-ranged attraction can be used to model the protein interactions.

Salt effects known from the literature thus show the expected behavior in these protein systems and no anomalies have been observed. In particular, the well-known salt effects do not provide any explanation for the ion-induced reentrant phase behavior observed in protein solutions in the presence of multivalent cations. Thus, another more specific mechanism of ion-protein interaction has to be relevant.

11.2 Ion Binding to the Protein: Charge Inversion

In the present thesis, the binding of multivalent cations to the protein surface is identified as the origin of the ion-induced effects in protein solution. In this section, we discuss the binding of multivalent cations to the surface of individual proteins before we move on to properties involving many protein molecules such as interactions and phase behavior.

The binding of ions to protein surface sites is a well-known fact in metall-binding proteins (Permyakov, 2009). For the case of multivalent cations, solvent-exposed carboxylic acids represents binding sites that are very ubiquitous in protein systems. The ion binding has been found in several simulation studies (Zhang et al., 2008, 2010; Kubíčková et al., 2012; Leibfarth, 2011) (cf. also Fig. 10.3) as well as in experimental studies determining the association constants of multivalent cations to individual amino acids (Berthon, 1995; Sandhu, 1977).

We developed a coarse-grained model incorporating multivalent cation binding to the protein surface, hydrolysis of the salt, charge regulation of the protein surface, and their mutual coupling (Ch. 5). Importantly, the model parameters are taken from literature values and physical estimations so that no artificial agreement is caused by free parameters with fitted, but unphysical values. Charge effects are corrected in a mean-field approach and also allow a direct way to account for charge screening due to changes of ionic strength (Ch. 5 and 6).

The charge inversion is reproduced with excellent qualitative and reasonable quantitative agreement for different globular proteins with different trivalent salts, i.e. YCl_3, $AlCl_3$, and $FeCl_3$. The three salts have an increasing pH effect due to hydrolysis, causing a more rapid charge inversion. Comparing the point of zero charge with the isoelectric point, we can conclude that both pH and ion binding have a non-negligible effect for the charge

inversion. The experimental results and the good agreement with the model calculations suggest that charge inversion is a universal phenomenon in solutions of proteins with native negative charge and multivalent cations (Ch. 5).

11.3 Ion-Induced Attraction Causing the Reentrant Condensation

For solutions of different globular proteins in the presence of multivalent cation, reentrant condensation has been found (Zhang et al., 2008, 2010). Reentrant condensation means that the solutions are clear and thermodynamically stable at low and high cation concentration, but condense at cation concentration between c^* and c^{**} (Zhang et al., 2008, 2010). In this thesis, the mechanism of the reentrant condensation behavior is investigated in more detail, thereby focusing on the underlying ion-induced protein interactions.

The condensed regime of the reentrant condensation in BSA solutions is narrowed significantly for salts with strong pH effects ($FeCl_3$, $AlCl_3$) compared to more neutral ones (YCl_3). Using the protein net charge from the model calculation, the screened Coulomb repulsion can be added to a fixed, effective van der Waals attraction to calculate an aggregation barrier in a DLVO-like spirit (cf. Fig. 1.1). Indeed, the aggregation barrier reflects the reentrant condensation with very good agreement (Ch. 5).

This model suggests that a charge inversion is enough to cause a reentrant condensation, since charge stabilization is only possible at high and low cation concentration. However, titration of BSA solutions through the isoelectric point induces no reentrant condensation (Dockal et al., 2000; Barbosa et al., 2010). Thus, the presence of multivalent cations seems to be essential for the attractive interaction between the proteins. This argument also implies that reentrant condensation is less universal than charge inversion, but also more interesting (Ch. 5).

Using results from experiments as well as two simulation schemes, the effect of monovalent salt on the charge inversion and reentrant condensation induced by multivalent salt is investigated, thereby providing information on the underlying mechanism for the variation of protein interaction. Experimentally, the addition of NaCl shifts the two transition concentrations c_Y^* and c_Y^{**} of the reentrant condensation to higher YCl_3 concentrations. For high enough NaCl concentrations, no second transition to the reentrant dissolution is observed anymore. The two simulation schemes—a Monte Carlo sampling of binding configurations and a coarse-grained model—evidence that the addition of NaCl increases the absolute charge before and after charge inversion (Ch. 6).

The combination of experiment and simulations provides a consistent picture on the different mechanisms at the transitions. The effect of monovalent salt on the second transition c_Y^{**} is consistent with a DLVO-like mechanism: with increasing NaCl concentration, the repulsion becomes screened and, thus, weaker. More trivalent cations are needed to reach the protein net charge required to charge stabilize the protein solution. The mechanism around the first transition concentration c_Y^* is clearly different. If charge screening was

the only effect, c_Y^* would be expected to decrease with increasing NaCl concentration, since less cations would be needed for a low enough charge. Since the opposite trend is observed, ion-induced effects beyond charge screening play an important role. The first transition c_Y^* seems to be related to a critical charge, i.e. a critical number of bound cations. This finding suggests that the attraction between the proteins is induced by the binding of cations (Ch. 6).

Using SAXS, the picture is further solidified, supporting the notion of an ion-induced attraction. Around the second transition, the protein concentration in the measured supernatant is decreased strongly, suggesting precipitation of aggregates. Also the increase of the SAXS intensity towards low q indicates the presence of aggregates in the samples. While the first transition thus seems to be caused essentially by the increased ion-induced attraction in an equilibrium liquid state, the second transition seems to depend rather on the charge stabilization against the formation of larger aggregates (Ch. 6).

11.4 Comparison to Reentrant Phase Behavior of Colloidal Particles with High Surface Charge

At this point, it is important to note the analogues—both similarities and differences—between the reentrant phase behavior observed in protein solutions and that for other solutes with high surface charge such as DNA.

First, a reentrant condensation is observed in all cases upon addition of multivalent counterions, with very similar topology and shape of the phase boundaries (Zhang et al., 2008; Grosberg et al., 2002). Second, the reentrant phase behavior in all systems is linked back to a gain in free energy when ions approach the surface of the macromolecule (Travesset and Vangaveti, 2009; Lyklema, 2006). Third, the reentrant condensation is driven by the interplay of Coulomb repulsion and an ion-induced attraction.

The molecular origin of the phenomena differ by the source of the free energy gain. For the case of DNA, ion–ion correlations cause the condensation of ions on the surface of DNA and induce a like-charge attraction between the colloids. Proteins can also carry high surface charges, so contributions from ion–ion correlations might also be present. However, the binding of cations to acidic residues on the protein surface represents a specific interaction not accounted for in electrostatic calculations. In general, the free energy contribution from these specific interactions is expected to be stronger than the contribution of ion–ion correlations (Besteman et al., 2004; Zhang et al., 2010). Furthermore, the non-spherical shape, inhomogeneous surface charge of proteins and discrete binding sites for the counterions might weaken the contribution of ion–ion correlations further.

When investigating the phase behavior in more detail, qualitative differences are observed upon addition of monovalent salt (Ch. 6). For protein solutions, both transition concentrations of multivalent cations increase with increasing ionic strength, whereas the second transition remains constant or even decays in the case of DNA (Saminathan

et al., 1999). Thus, although, at first glance, the phase behavior looks comparable for the protein and DNA case, the interaction mechanisms governing the solutions differ clearly. Protein solutions are an independent system of biological soft matter with interesting phase behavior governed also by electrostatic effects.

11.5 Liquid–Liquid Phase Separation

In colloidal solutions with attractive interaction, a liquid–liquid phase separation (LLPS) is expected once the attractive interaction becomes strong enough. As a measure for the required strength for isotropic interaction potentials, the reduced second virial coefficient B_2^*, which is normalized by the excluded volume contribution, should be smaller than roughly -1.5 based on simulations and theoretical predictions (Vliegenthart and Lekkerkerker, 2000; Noro and Frenkel, 2000) as well as on a very general thermodynamic argument from the virial expansion (Ch. 7).

LLPS is observed for different proteins in a closed area in the isothermal (c_p, c_s) plane (Ch. 7). From SAXS measurements on the dilute phase after LLPS, both the scattering intensity at low q, as a rough measure of interaction, and the reduced second virial coefficients B_2^* are found to vary non-monotonously with increasing YCl_3 concentration, evidencing first an increase and subsequently a decrease of attraction. Importantly, the reduced second virial coefficient consistently is below the critical value of -1.5, approaching it at the lower and higher YCl_3 boundary of the LLPS regime (Ch. 7). These experimental findings on protein solutions are in accordance with predictions for the LLPS in colloidal solutions, when the salt concentration is considered as a second control parameter in addition to the temperature.

The observed phase behavior is qualitatively in accordance with the LLPS phase diagrams modeled for lysozyme in the presence of different mono- and divalent salts (Wentzel and Gunton, 2007) using experimental results on the clouding point (Grigsby et al., 2001; Park and Bae, 2004). For several salts, the clouding point temperature increases with increasing salt concentration and then decreases again, indicating a reentrant phase behavior (Grigsby et al., 2001; Park and Bae, 2004). Mapping these results to a square well system with varying attraction strength, Wentzel and Gunton (2007) derived an isothermal phase diagram with a closed LLPS region similar to the one discussed in Ch. 7.

Interestingly, two different mechanisms seem to cause the variation of the attraction. In the case of YCl_3 and globular proteins with negative surface charge, the charge inversion due to ion binding is expected to drive the reentrant behavior. Consistently, for lysozyme with a positive surface charge no reentrant behavior is observed with YCl_3 (Zhang et al., 2010). The reentrant effect in the case of lysozyme is not related to a charge inversion, but can be attributed to hydration effects, consistent with the observed trends upon changing the mono- and divalent salts along the Hofmeister series (Grigsby et al., 2001).

Two further aspects related to the LLPS deserve discussion. First, the thermodynamic status of a LLPS in colloidal solutions, whether stable or metastable, is an interesting question also related to applications. Decreasing the range of the attraction between the colloids, the coexisting phases after a LLPS become metastable with respect to crystallization (Anderson and Lekkerkerker, 2002). A metastable LLPS is thus not a real equilibrium phase transition, but is observed due to kinetic reasons: The kinetics of spinodal decomposition or nucleation of dense droplets occur on much shorter time scales than the nucleation of the solid phase. This time scale separation renders the LLPS experimentally observable and even apparently stable over some time, although the coexisting phases do not represent equilibrium. For isotropic interactions, the attraction range related to the onset of metastability is estimated to around 15–25 % of the colloidal diameter (Noro and Frenkel, 2000). In solutions of proteins, the attractions are short-ranged compared to the protein size, and the LLPS should be metastable, which is consistent with the observation that crystals and aggregates form from the coexisting liquid phases. The metastable LLPS in protein solutions with multivalent cations is inspiring from the viewpoint of crystallization processes as discussed in Sec. 11.6.

Second, the critical volume fraction η_c of the LLPS caused by isotropic interactions is expected to vary between the short-range sticky-sphere limit of $\eta_c \approx 0.3$ (Noro and Frenkel, 2000) and the long-range mean-field limit of $\eta_c \approx 0.13$ (Hansen and McDonald, 2006). The observed critical volume fraction for the protein solutions with multivalent cations lies below this range around $\eta_c \approx 0.08$ (Ch. 7). This finding suggests that anisotropic interactions are present, which are discussed in more detail in Sec. 11.7.

11.6 Pathways of Protein Crystallization: Two-Step Nucleation Process with Cluster Precursor

The growth of protein crystals is observed within the condensed regime of the reentrant condensation close to the dilute boundary of the metastable LLPS and not from the dense phases where aggregates form (Ch. 7). This observation needs more discussion since it is neither in accordance with the classical nucleation theory nor with the conventional notion of a two-step nucleation process involving a LLPS (Ch. 8).

Classical nucleation theory (CNT) starts from the picture of a homogeneous solutions which is metastable with respect to another phase, e,g. crystal. In this case, a nucleation event corresponds to a local variation of structure and density, and nucleus growth is not favorable until a critical nucleus size is reached. Importantly, within the framework of CNT, the nucleus is conceptualized to form directly with the final crystalline structure, implying that structure and density have to develop simultaneously from a homogeneous solution. Thus, CNT is not prepared to deal with inhomogeneous solutions and possible intermediate precursor states induced e.g. by a LLPS.

The alternative picture is a multi-step nucleation process during which different intermediate states occur since the formation of intermediates is much faster than the nucleation of the equilibrium crystal phase. The characteristic feature of a two-step nucleation process is that two order parameters—density and local structure—develop subsequently and not coupled as in the classical nucleation theory. The conventional notion in the case of protein solutions involves in a first step a liquid–liquid phase separation with only the density as order parameter. In the second step, the local structure emerges by rearrangements in the dense phase, representing a nucleation without strong changes of density but local structure as the second order parameter. Thus, the nucleation of a crystal nucleus should occur in the dense phase due to the higher prevalent volume fraction, as it has been observed for several protein solutions (Galkin et al., 2002; Vivarès et al., 2005).

Two factors could rationalize the finding that crystallization starts from the dilute phase as observed in the present thesis: First, the salt ions could be distributed into the two coexisting phases in a way that the effective interactions differ considerably. However, the cation concentration determined in the experiment seems to be proportional to the protein concentration (cf. Fig. 7.1(a)). Based on the mechanisms of ion binding and ion-induced attraction, it seems very unlikely that the protein interactions differ considerably in the two coexisting phases.

Second, the dense phase might be an arrested phase in which aggregates form that hinder local rearrangements to build crystal nuclei. In the dilute phase, another two-step nucleation process seems probable for protein crystallization, i.e. via cluster precursor as observed experimentally during crystallization (Ch. 8). Importantly, the clusters seem to be rather loosely bound, allowing for structural reorganization. The good agreement of calculated scattering curves of cluster structures, which are build from an experimental crystal structure, with the experimental scattering data suggest that the cluster might already have a semi-crystalline structure. In this picture, clusters grow as precursors with higher local density and intermediate local order, from which then in a second step the pure crystal nucleus can rearrange. The dense phase represents a protein reservoir that is consumed during the growth of the crystals in the dilute phase (Ch. 8).

11.7 Proteins as Particles with Ion-Activated Attractive Patches

While the molecular origin for the charge inversion can be identified as the binding of cations to protein surface groups (cf. Sec. 11.2), the molecular origin of the ion-induced attraction causing the reentrant condensation is still to be discussed (cf. Sec. 11.3). In crystal structures of proteins, the formation of ion bridges has been observed. Yttrium ions bind simultaneously to several carboxylic acids on the surface of two neighboring protein molecules. For β-lactoglobulin, dimers are connected by 8 ion bridges, suggesting these cross-links to represent an orientation-dependent attraction between the proteins (Zhang et al., 2011). For monomers of human serum albumin, 4 ion bridges are formed with neighboring molecules.

These ion bridges can be described by surface patches of a special kind: The patches do not only act as potential attractive patches between proteins but also as binding sites for cations. The patchy interaction becomes attractive only when an occupied and unoccupied patch meet. Otherwise they remain inert or even repulsive. Modeling the proteins by particles with these ion-activated attractive patches, the phase behavior is calculated within the framework of the Wertheim theory for associating fluids (Ch. 9).

With increasing cation concentration, an increasing number of cations bind to the patches. The bound ions correspond not only to the charge inversion, but also induce a non-monotonously varying attraction strength with a maximum at intermediate cation concentration, reflecting the reentrant condensation behavior. The experimentally observed turbidity of the samples can be explained by the emergence of clusters as well as the divergence of the isothermal compressibility with increasing attraction strength. The liquid–liquid phase separation is also predicted within a similar range of protein volume fraction and salt concentration compared to the experimental results (Ch. 9). Furthermore, the differing pathways of crystallization can be rationalized using the formation of larger aggregates close to the percolation of ion bridges, as predicted from the patchy particle model (Ch. 9). Finally, the formation of clusters of only several proteins is expected theoretically also at lower cation concentration (Ch. 9). This prediction couples to experimental results for these systems on the protein diffusion and cluster formation, by light scattering as well as quasi-elastic neutron scattering (Soraruf et al., 2013; Grimaldo, 2013).

The concept of patchy particles has been suggested and successfully applied as a model for proteins with non-spherical shape and inhomogeneous surface patterns of charge and hydrophobicity (Sear, 1999; Kern and Frenkel, 2003; Gögelein et al., 2008; Whitelam, 2010a; Fusco and Charbonneau, 2013). Independent experimental results such as a low critical volume fraction of LLPS in protein solutions and a low density of protein crystals point toward anisotropic interactions (Bianchi et al., 2011). Recently, a patchy particle model has been parametrized using the crystal contacts from real protein crystals and performed well to explain differences in the phase behavior of a protein in its wild-type and mutated form (Fusco et al., 2013).

Based on this studies, the study presented in this thesis provides a good test case whether the patchy model is really suitable to model the protein interaction, since the cation concentration as a second control parameter determines the attraction strength of the patches in a model-intrinsic way. The excellent agreement and consistent explanatory power of this simple picture of proteins as particles with ion-activated attractive patches promises an interesting connection of statistical physics to dynamic as well as static properties of protein solutions (Ch. 9).

11.8 Implications and Outlook: I. Effect of Temperature

The full thesis focuses on isothermal effects in protein solutions in which the concentration of multivalent cations represents the control parameter. Reconsidering the reentrant phase behavior and the closed area of the LLPS, the cation concentration can be understood as an effective temperature that varies non-monotonously with a minimum value at intermediate cation concentration, analogously to the varying attraction strength in the experiments (Wolf et al., 2013)(Ch. 7) and the patchy particle model (Ch. 9). The potential interplay of this effective temperature with the system temperature represents an interesting topic for further investigations.

This question becomes more relevant due to the fact that for several protein solutions with multivalent salts, a lower critical solution temperature (LCST) has been observed, i.e. the coexisting densities of the LLPS broaden with increasing temperature. This LCST behavior suggests a dominating entropic contribution to the free energy, and thus, a relevant role of water-mediated effects related to hydration (Vekilov et al., 2002; Shiryayev et al., 2005; Wentzel and Gunton, 2007, 2008). In this picture, the role of the cations is not clear at all. Possible ion effects due to hydration range from solvation effects to changes of the hydrophobicity patterns on the protein surface. Thus, an experimental phase diagram for both control parameters—temperature and salt concentration—could prove very helpful to elucidate the molecular mechanisms of protein–ion and protein–protein interactions further. The understanding of temperature effects is essential to describe the full phase behavior of protein solutions in more detail, and would also allow a systematic optimization of protein crystallization with respect to temperatures during storage and preparation.

11.9 Implications and Outlook: II. Pathways of Structure Formation and Crystallization in Protein Solutions

During the thesis, we outlined a consistent picture how to understand the equilibrium phase behavior, protein–protein and protein–ion interactions. The implementation of this information into a predictive framework for protein crystallization is, however, an unsolved challenge. In protein solutions, several features and levels of structure formation—cross-linking, clustering, LLPS, precipitation and crystallization—are observed. The hierarchy and temporal order between the formation of different structures is far from trivial and is most probably linked to the specific interaction mechanisms, in particular the relative strengths of isotropic and anisotropic interactions.

As an example of hierarchical structure formation in solutions of bovine serum albumin, Wang and Annunziata (2008) observed the formation of microspheres resembling a LLPS and a subsequent cross-linking between the microspheres upon the addition of PEG as depletion agent and glutaraldehyde as cross-linker. Both the cross-linking of microspheres

and the depletion interaction are rather isotropic and the ubiquity of depletion agent compared to the cross-linker seems to be the reason that the formation of microspheres is established first and followed by cross-linking.

A very different realization of hierarchical structure formation is observed for the protein–salt system discussed in this thesis. Ion bridges—i.e. cross-links between the proteins—are formed first, represent highly anisotropic interactions, and induce further phenomena such as clustering and LLPS (Ch. 7 and 9). These clusters seem to have a rather liquid-like structure and act as precursors for crystal nuclei (Ch. 8).

A comparable behavior has been observed in simulations of the crystallization process for a patchy model system for S-layer proteins (Whitelam, 2010a). For weak isotropic attraction compared to directional attraction, crystals nucleate directly from the solution. Increasing the ratio between the isotropic and the directional interaction strength, liquid-like clusters form first, from which then crystals nucleate. For even higher ratios, large aggregates are observed and nucleation is arrested (Whitelam, 2010a). These results have been further complemented with theoretical calculations for the evolution of mean-field order parameters—density and local orientation—in solutions of particles with anisotropic and isotropic interactions (Whitelam, 2010b). The pathways from a homogeneous solution to the crystal pass through intermediate, amorphous or semi-ordered phases and depend not only on the isotropic and anisotropic interaction strengths, but also on dynamical properties of the order parameters (Whitelam, 2010b). The classical view of two-step nucleation via critical fluctuations (Wolde and Frenkel, 1997) or the formation of dense liquids as precursor (Vekilov, 2004) thus have to be reconsidered once anisotropic interactions are relevant and could potentially represent templates promoting the formation of crystalline structure.

This finding also affects the classical approach by George and Wilson (1994) as well as the consideration of Vliegenthart and Lekkerkerker (2000), both of which focus on the second virial coefficient and the related phase behavior, disregarding the potential anisotropic nature of the protein interactions. Haxton and Whitelam (2012) performed a simulation study to explore the question how to optimize conditions for protein crystallization, allowing also for anisotropic interactions. In this study, best assembly is not related to a crystallization slot of the second virial coefficient, but found for slight supersaturation when staying away from the LLPS boundary. For this purpose, a sufficient metastability gap between solubility line and LLPS boundary has to be present, which can be ensured by a suitable choice of sufficiently anisotropic and very weak isotropic interaction (Haxton and Whitelam, 2012). Further insight has been added by a simulation study investigating the effect of different patch geometries and asymmetries of attraction strengths (Fusco and Charbonneau, 2013). While the patch geometry mainly altered the density of the crystal phase, an increasing asymmetry of attraction strengths changes the LLPS from a metastable phase transition with respect to crystallization into a stable state, thereby causing substantial effects on the thermodynamics and kinetics of crystallization (Fusco

and Charbonneau, 2013). Patchy interaction thus changes the picture for the protein crystallization considerably.

The calculation of realistic phase diagrams for explicit proteins via patchy models (Fusco et al., 2013; McManus et al., 2007) promises new theoretical concepts towards protein crystallization based on equilibrium and kinetic arguments. To realize the full power of these theoretical concepts, experimental techniques have to be developed that allow a tuning of patchy interactions in protein solutions. In this context, the protein–salt systems studied in this thesis become very interesting, since the binding of multivalent cations varies at the same time the surface charge, interaction strength and the patch geometry on the protein surface. Established knowledge on ion binding sites in proteins could be implemented into patchy models to predict phase diagrams. Furthermore, the interplay of ion bridging and electrostatics might not only govern the phase behavior, but also kinetics of nucleation and clustering, since charge patterns could represent association funnels between proteins. For a reliable descriptive or predictive framework of protein crystallization, the characterization of the equilibrium properties is not enough, but has to be complemented by real-time monitoring of structure formation (cf. Ch. 8).

Solutions of globular proteins with multivalent cations thus represent challenging, but also promising experimental systems, since they provide unparalleled opportunities to tune anisotropic interactions, and a variety of experimental techniques is ready to access structural as well as dynamical features.

11.10 Final Conclusions

Solutions of globular proteins in the presence of multivalent cations exhibit a rich phase behavior including reentrant condensation, liquid–liquid phase separation, cluster formation and crystallization. The protein–ion interaction is dominated by the binding of the multivalent cations to acidic residues on the protein surface, inducing an inversion of the protein net charge. The reentrant phase behavior is caused by the competition of the Coulomb repulsion, which varies in strength throughout charge inversion, and an ion-induced attraction. Cross-linking between protein molecules by bridges of multivalent cations can be used to reproduce the full phase behavior with excellent agreement. The related formation of clusters might act as precursors in a two-step nucleation pathway for protein crystallization.

Throughout the thesis, experimental results on protein solutions with salt are complemented by an understanding based on colloid-like models. First, the concept of effective spheres proves suitable to describe protein interactions measured in salt solutions along the Hofmeister series. Second, the interplay of ion binding, charge regulation and pH effects due to salt hydrolysis is combined in a coarse-grained model that successfully predicts the experimentally observed charge inversion and reentrant condensation. Third, a model of particles with ion-activated attractive patches reproduces the ion-induced phenomena observed in protein solutions with multivalent cations.

In addition to temperature, the concentration of multivalent cations is a second control parameter for the phase behavior of protein solutions, and allows the simultaneous variation of interaction strength and patch geometry. These features are promising for future research on optimum equilibrium and non-equilibrium conditions for protein crystallization. The combination of experiments and the statistical physics of patchy models based on real protein structures might be an important next step.

Bibliography

Aasa, R., Malmström, B. G., Saltman, P., and Vänngård, T. (1963). The specific binding of iron(III) and copper(II) to transferrin and conalbumin. *Biochim. Biophys. Acta*, 75(0):203 – 222. [5, 64]

Abade, G. C., Cichocki, B., Ekiel-Jeyewska, M. L., Nägele, G., and Wajnryb, E. (2010). Short-time dynamics of permeable particles in concentrated suspensions. *J. Chem. Phys.*, 132(1):014503. [139]

Agashe, K. and Regalbuto, J. (1997). A revised physical theory for adsorption of metal complexes at oxide surfaces. *J. Colloid Interface Sci.*, 185(1):174–189. [61]

Alberts, B., Johnson, A., Lewis, J., Rafi, M., Roberts, K., and Walter, P. (2008). *Molecular Biology of the Cell*. Garland Science, Taylor & Francis Group, New York, 5th revised edition. [1]

Allahyarov, E., D'Amico, I., and Löwen, H. (1998). Attraction between like-charged macroions by coulomb depletion. *Phys. Rev. Lett.*, 81(6):1334–1337. [53]

Anderson, V. and Lekkerkerker, H. (2002). Insights into phase transition kinetics from colloid science. *Nature*, 416(6883):811–815. [6, 96, 148]

Andresen, K., Das, R., Park, H. Y., Smith, H., Kwok, L. W., Lamb, J. S., Kirkland, E. J., Herschlag, D., Finkelstein, K. D., and Pollack, L. (2004). Spatial distribution of competing ions around DNA in solution. *Phys. Rev. Lett.*, 93:248103. [3, 76]

Annunziata, O., Asherie, N., Lomakin, A., Pande, J., Ogun, O., and Benedek, G. B. (2002). Effect of polyethylene glycol on the liquid–liquid phase transition in aqueous protein solutions. *Proc. Natl. Acad. Sci. USA*, 99(22):14165–14170. [8, 108]

Annunziata, O., Ogun, O., and Benedek, G. B. (2003). Observation of liquid–liquid phase separation for eye lens γS-crystallin. *Proc. Natl. Acad. Sci. USA*, 100(3):970–974. [8, 96, 108, 124]

Arakawa, T. and Timasheff, S. N. (1982). Preferential interactions of proteins with salts in concentrated solutions. *Biochemistry*, 21(25):6545–6552. [48]

Arakawa, T. and Timasheff, S. N. (1984). Mechanism of protein salting in and salting out by divalent cation salts: balance between hydration and salt binding. *Biochemistry*, 23(25):5912–5923. [4, 53]

Asherie, N. R. (2004). Protein crystallization and phase diagrams. *Methods*, 34(3):266–272. [96]

Asherie, N. R., Lomakin, A., and Benedek, G. B. (1996). Phase diagram of colloidal solutions. *Phys. Rev. Lett.*, 77(23):4832–4835. [7, 96, 99, 124]

Asthagiri, D., Paliwal, A., Abras, D., Lenhoff, A., and Paulaitis, M. (2005). A consistent experimental and modeling approach to light-scattering studies of protein-protein interactions in solution. *Biophys. J.*, 88(5):3300 – 3309. [44]

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(42):S5003. [9, 40]

Baldwin, R. (1996). How Hofmeister ion interactions affect protein stability. *Biophys. J.*, 71(4):2056 – 2063. [4, 34]

Banfield, J. F., Welch, S. A., Zhang, H., Ebert, T. T., and Penn, R. L. (2000). Aggregation-based crystal growth and microstructure development in natural iron oxyhydroxide biomineralization products. *Science*, 289(5480):751–754. [109, 117]

Barbosa, L. R., Ortore, M. G., Spinozzi, F., Mariani, P., Bernstorff, S., and Itri, R. (2010). The Importance of Protein-Protein Interactions on the pH-Induced Conformational Changes of Bovine Serum Albumin: A Small-Angle X-Ray Scattering Study. *Biophys. J.*, 98(1):147 – 157. [34, 69, 134, 145]

Barhoum, S. and Yethiraj, A. (2010). NMR detection of an equilibrium phase consisting of monomers and clusters in concentrated lysozyme solutions. *J. Phys. Chem. B*, 114:17062–17067. [8]

Bashford, D. and Karplus, M. (1990). pKa's of ionizable groups in proteins: atomic detail from a continuum electrostatic model. *Biochemistry*, 29(44):10219–10225. [80]

Bashford, D. and Karplus, M. (1991). Multiple-site titration curves of proteins: an analysis of exact and approximate methods for their calculation. *J. Phys. Chem.*, 95(23):9556–9561. [61, 63]

Baxter, R. J. (1968). Percus–Yevick equation for hard spheres with surface adhesion. *J. Chem. Phys.*, 49(6):2770–2774. [100, 103, 105]

Belloni, L. (2000). Colloidal interactions. *J. Phys.: Condens. Matter*, 12(46):R549. [96]

Bendedouch, D. and Chen, S.-H. (1983). Structure and interparticle interactions of bovine serum albumin in solution studied by small-angle neutron scattering. *J. Chem. Phys.*, 87(9):1473–1477. [22, 34, 35, 39, 41, 42]

Bendedouch, D., Chen, S. H., and Koehler, W. C. (1983). Determination of interparticle structure factors in ionic micellar solutions by small angle neutron scattering. *J. Phys. Chem.*, 87(14):2621–2628. [34]

Benedek, G. B. (1997). Cataract as a protein condensation disease: the Proctor Lecture. *Invest. Ophth. Vis. Sci.*, 38(10):1911–1921. [2]

Berland, C. R., Thurston, G. M., Kondo, M., Broide, M. L., Pande, J., Ogun, O., and Benedek, G. B. (1992). Solid-liquid phase boundaries of lens protein solutions. *Proc. Natl. Acad. Sci. USA*, 89(4):1214–1218. [7]

Berthon, G. (1995). The stability constants of metal complexes of amino acids with polar side chains. *Pure Appl. Chem.*, 67:1117–1240. [5, 64, 144]

Besteman, K., Van Eijk, K., and Lemay, S. G. (2007). Charge inversion accompanies dna condensation by multivalent ions. *Nat. Phys.*, 3(9):641–644. [5]

Besteman, K., Zevenbergen, M. A. G., Heering, H. A., and Lemay, S. G. (2004). Direct observation of charge inversion by multivalent ions as a universal electrostatic phenomenon. *Phys. Rev. Lett.*, 93(17):170802. [5, 6, 53, 76, 146]

Bianchi, E., Blaak, R., and Likos, C. N. (2011). Patchy colloids: state of the art and perspectives. *Phys. Chem. Chem. Phys.*, 13:6397–6410. [6, 10, 124, 150]

Bianchi, E., Largo, J., Tartaglia, P., Zaccarelli, E., and Sciortino, F. (2006). Phase diagram of patchy colloids: Towards empty liquids. *Phys. Rev. Lett.*, 97:168301. [10, 52, 102, 120, 124]

Bianchi, E., Tartaglia, P., Zaccarelli, E., and Sciortino, F. (2008). Theoretical and numerical study of the phase diagram of patchy colloids: Ordered and disordered patch arrangements. *J. Chem. Phys.*, 128(14):144504. [10, 129, 130, 131]

Bonneté, F. and Vivarès, D. (2002). Interest of the normalized second virial coefficient and interaction potentials for crystallizing large macromolecules. *Acta Crystallogr., Sect. D*, 58(10):1571–1575. [37, 43]

Bonnett, P. E., Carpenter, K. J., Dawson, S., and Davey, R. J. (2003). Solution crystallisation via a submerged liquid-liquid phase boundary: oiling out. *Chem. Commun.*, 0:698–699. [9, 109]

Bonnet, F., Finet, S., and Tardieu, A. (1999). Second virial coefficient: variations with lysozyme crystallization conditions. *J. Cryst. Growth*, 196(24):403 – 414. [37, 43, 49]

Boström, M., Williams, D. R. M., and Ninham, B. W. (2001). Specific ion effects: Why DLVO theory fails for biology and colloid systems. *Phys. Rev. Lett.*, 87(16):168103. [3]

Bouyain, S., Longo, P. A., Li, S., Ferguson, K. M., and Leahy, D. J. (2005). The extracellular region of ErbB4 adopts a tethered conformation in the absence of ligand. *Proc. Natl. Acad. Sci. USA*, 102(42):15024–15029. [101]

Broide, M. L., Berland, C. R., Pande, J., Ogun, O. O., and Benedek, G. B. (1991). Binary-liquid phase separation of lens protein solutions. *Proc. Natl. Acad. Sci. USA*, 88(13):5660–5664. [7, 108]

Buell, A. K., Tartaglia, G. G., Birkett, N. R., Waudby, C. A., Vendruscolo, M., Salvatella, X., Welland, M. E., Dobson, C. M., and Knowles, T. P. J. (2009). Position-dependent electrostatic protection against protein aggregation. *ChemBioChem*, 10(8):1309–1312. [6, 10, 52]

Burak, Y., Ariel, G., and Andelman, D. (2004). Competition between condensation of monovalent and multivalent ions in DNA aggregation. *Curr. Opin. Colloid Interface Sci.*, 9(12):53 – 58. [3, 7, 76, 77, 92]

Byrne, R., Luo, Y.-R., and Young, R. (2000). Iron hydrolysis and solubility revisited: observations and comments on iron hydrolysis characterizations. *Mar. Chem.*, 70(13):23 – 35. [61, 63]

Cardinaux, F., Gibaud, T., Stradner, A., and Schurtenberger, P. (2007). Interplay between spinodal decomposition and glass formation in proteins exhibiting short-range attractions. *Phys. Rev. Lett.*, 99:118301. [108]

Cardinaux, F., Zaccarelli, E., Stradner, A., Bucciarelli, S., Farago, B., Egelhaaf, S. U., Sciortino, F., and Schurtenberger, P. (2011). Cluster-driven dynamical arrest in concentrated lysozyme solutions. *J. Phys. Chem. B*, 115(22):7227–7237. [8, 9]

Carter, D. C. and Ho, J. X. (1994). Structure of serum albumin. *Adv. Protein Chem.*, 45:153–203. [105]

Chapman, D. L. (1913). A contribution to the theory of electrocapillarity. *Phil. Mag.*, 25(148):475–481. [2]

Chapman, W. G., Jackson, G., and Gubbins, K. E. (1988). Phase equilibria of associating fluids. *Mol. Phys.*, 65(5):1057–1079. [10, 124, 130]

Chen, Q., Vekilov, P. G., Nagel, R. L., and Hirsch, R. E. (2004). Liquid-liquid phase separation in hemoglobins: Distinct aggregation mechanisms of the β6 mutants. *Biophys. J.*, 86(3):1702 – 1712. [8, 108]

Chen, S.-H. (1986). Small angle neutron scattering studies of the structure and interaction in micellar and microemulsion systems. *Annu. Rev. Phys. Chem.*, 37(1):351–399. [37]

Chen, S. H. and Lin, T. L. (1987). *Neutron Scattering*, pages 489–543. Academic Press, London. [37, 79]

Collins, K. (1997). Charge density-dependent strength of hydration and biological structure. *Biophys. J.*, 72(1):65 – 76. [4, 32, 34, 52]

Collins, K. D. (2004). Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Methods*, 34(3):300 – 311. [4, 32, 34, 76]

Collins, K. D. and Washabaugh, M. W. (1985). The Hofmeister effect and the behaviour of water at interfaces. *Q. Rev. Biophys.*, 18:323–422. [4, 32, 33]

Curtis, R. A., Blanch, H. W., and Prausnitz, J. M. (2001). Calculation of phase diagrams for aqueous protein solutions. *J. Phys. Chem. B*, 105(12):2445–2452. [10, 32]

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(1):11–21. [32, 76]

Curtis, R. A., Ulrich, J., Montaser, A., Prausnitz, J. M., and Blanch, H. W. (2002). Protein–protein interactions in concentrated electrolyte solutions. *Biotechnol. Bioeng.*, 79(4):367–380. [32, 40]

Dahirel, V. and Jardat, M. (2010). Effective interactions between charged nanoparticles in water: What is left from the DLVO theory? *Curr. Opin. Colloid Interface Sci.*, 15(12):2 – 7. [4, 53]

Das, R., Mills, T. T., Kwok, L. W., Maskel, G. S., Millett, I. S., Doniach, S., Finkelstein, K. D., Herschlag, D., and Pollack, L. (2003). Counterion distribution around DNA probed by solution X-ray scattering. *Phys. Rev. Lett.*, 90:188103. [3, 76]

Debye, P. and Hückel, E. (1923). De la theorie des electrolytes. i. abaissement du point de congelation et phenomenes associes. *Phys. Z.*, 24(9):185–206. [3]

Delaye, M., Clark, J., and Benedek, G. (1981). Coexistence curves for phase separation in the calf lens cytoplasm. *Biochem. Biophys. Res. Commun.*, 100(2):908 – 914. [108]

Delgado, A., González-Caballero, F., Hunter, R., Koopal, L., and Lyklema, J. (2007). Measurement and interpretation of electrokinetic phenomena. *J. Colloid Interface Sci.*, 309(2):194 – 224. [25, 26, 54, 55]

Dér, A., Kelemen, L., Fbin, L., Taneva, S. G., Fodor, E., Pli, T., Cupane, A., Cacace, M. G., and Ramsden, J. J. (2007). Interfacial water structure controls protein conformation. *J. Phys. Chem. B*, 111(19):5344–5350. [4, 34]

Derjaguin, B. V. and Landau, L. (1941). Theory of the stability of strongly charged lyophobic sols and of the adhesion of strongly charged particles in solutions of electrolytes. *Acta Physicochim. USSR*, 14:633–662. [3, 52, 66]

Deserno, M., Holm, C., and May, S. (2000). Fraction of condensed counterions around a charged rod: Comparison of Poisson–Boltzmann theory and computer simulations. *Macromolecules*, 33(1):199–206. [5, 53, 76]

Deshiikan, S. R. and Papadopoulos, K. D. (1998). Modified booth equation for the calculation of zeta potential. *Colloid Polym. Sci.*, 276:117–124. [26]

Dewhurst, C. (2012). Grasp: Graphical reduction and analysis SANS program for Matlab. http://www.ill.eu/?id=6300. [35]

Diakonov, I., Ragnarsdottir, K., and Tagirov, B. (1998). Standard thermodynamic properties and heat capacity equations of rare earth hydroxides:: II. Ce(III)-, Pr-, Sm-, Eu(III)-, Gd-, Tb-, Dy-, Ho-, Er-, Tm-, Yb-, and Y-hydroxides. Comparison of thermochemical and solubility data. *Chem. Geol.*, 151(14):327 – 347. [62, 63]

Dijkstra, M., van Roij, R., and Evans, R. (1999). Phase diagram of highly asymmetric binary hard-sphere mixtures. *Phys. Rev. E*, 59:5744–5771. [103]

Dockal, M., Carter, D. C., and Rüker, F. (2000). Conformational transitions of the three recombinant domains of human serum albumin depending on pH. *J. Biol. Chem.*, 275(5):3042–3050. [69, 90, 134, 145]

Doniach, S. (2001). Changes in biomolecular conformation seen by small angle X-ray scattering. *Chem. Rev.*, 101(6):1763–1778. [39, 40]

Dorsaz, N., Filion, L., Smallenburg, F., and Frenkel, D. (2012). Spiers memorial lecture: Effect of interaction specificity on the phase behaviour of patchy particles. *Faraday Discuss.*, 159:9–21. [11]

Dorsaz, N., Thurston, G. M., Stradner, A., Schurtenberger, P., and Foffi, G. (2011). Phase separation in binary eye lens protein mixtures. *Soft Matter*, 7(5):1763–1776. [8, 108]

Douglas, J. F. and Garboczi, E. J. (1995). Intrinsic viscosity and the polarizability of particles having a wide range of shapes. *Adv. Chem. Phys.*, 91:85–154. [138]

Dumetz, A. C., Snellinger-O'Brien, A. M., Kaler, E. W., and Lenhoff, A. M. (2007). Patterns of protein–protein interactions in salt solutions and implications for protein crystallization. *Protein Sci.*, 16(9):1867–1877. [40]

Durbin, S. and Feher, G. (1996). Protein crystallization. *Annu. Rev. Phys. Chem.*, 47(1):171–204. [32]

Elofsson, U. M., Paulsson, M. A., and Arnebrant, T. (1997). Adsorption of β-lactoglobulin A and B in relation to self-association:effect of concentration and pH. *Langmuir*, 13(6):1695–1700. [112]

Erdemir, D., Lee, A. Y., and Myerson, A. S. (2009). Nucleation of crystals from solution: Classical and two-step models. *Acc. Chem. Res.*, 42(5):621–629. [9, 109]

Feigin, L. A. and Svergun, D. I. (1987). *Structure Analysis by Small-Angle X-Ray and Neutron Scattering*. Plenum Press, New York. [21, 79]

Feldman, H. A. (1972). Mathematical theory of complex ligand-binding systems at equilibrium: Some methods for parameter fitting. *Anal. Biochem.*, 48(2):317 – 338. [5, 53, 61]

Ferrer, M. L., Duchowicz, R., Carrasco, B., de la Torre, J. G., and Acuna, A. U. (2001). The conformation of serum albumin in solution: A combined phosphorescence depolarization-hydrodynamic modeling study. *Biophys. J.*, 80(5):2422 – 2430. [137, 143]

Fierro, A., Abete, T., Coniglio, A., and de Candia, A. (2011). Clusters in colloidal systems. *J. Phys. Chem. B*, 115(22):7281–7287. [8]

Fine, B. M., Pande, J., Lomakin, A., Ogun, O. O., and Benedek, G. B. (1995). Dynamic critical phenomena in aqueous protein solutions. *Phys. Rev. Lett.*, 74(1):198–201. [7]

Finet, S., Skouri-Panet, F., Casselyn, M., Bonnet, F., and Tardieu, A. (2004). The Hofmeister effect as seen by SAXS in protein solutions. *Curr. Opin. Colloid Interface Sci.*, 9(12):112 – 116. [37, 43, 49]

Finet, S. and Tardieu, A. (2001). α-crystallin interaction forces studied by small angle X-ray scattering and numerical simulations. *J. Cryst. Growth*, 232(14):40 – 49. [43]

Flory, P. J. (1941). Molecular size distribution in three dimensional polymers. I. gelation. *J. Am. Chem. Soc.*, 63(11):3083–3090. [129]

Flynn, C. M. (1984). Hydrolysis of inorganic iron(III) salts. *Chem. Rev.*, 84(1):31–41. [61, 62, 63]

Foffi, G. and Sciortino, F. (2007). On the possibility of extending the Noro–Frenkel generalized law of correspondent states to nonisotropic patchy interactions. *J. Phys. Chem. B*, 111(33):9702–9705. [10, 101]

Frauenfelder, H., Chen, G., Berendzen, J., Fenimore, P. W., Jansson, H., McMahon, B. H., Stroe, I. R., Swenson, J., and Young, R. D. (2009). A unified model of protein dynamics. *Proc. Natl. Acad. Sci. USA*, 106(13):5129–5134. [45]

Fukasawa, T. and Sato, T. (2011). Versatile application of indirect Fourier transformation to structure factor analysis: from X-ray diffraction of molecular liquids to small angle scattering of protein solutions. *Phys. Chem. Chem. Phys.*, 13:3187–3196. [84]

Furrer, G., Phillips, B. L., Ulrich, K.-U., Pöthig, R., and Casey, W. H. (2002). The origin of aluminum flocs in polluted streams. *Science*, 297(5590):2245–2247. [109, 117]

Fusco, D. and Charbonneau, P. (2013). Crystallization of asymmetric patchy models for globular proteins in solution. *Phys. Rev. E*, 88:012721. [6, 11, 124, 150, 152]

Fusco, D., Headd, J. J., Simone, A. D., and Charbonneau, P. (2013). Characterizing protein crystal contacts and their role in protein crystallization. *arXiv preprint arXiv:1301.3349*. [11, 150, 153]

Galkin, O., Chen, K., Nagel, R. L., Hirsch, R. E., and Vekilov, P. G. (2002). Liquid–liquid separation in solutions of normal and sickle cell hemoglobin. *Proc. Natl. Acad. Sci. USA*, 99(13):8479–8483. [7, 96, 108, 117, 120, 149]

Galkin, O. and Vekilov, P. G. (2000a). Are nucleation kinetics of protein crystals similar to those of liquid droplets? *J. Am. Chem. Soc.*, 122(1):156–163. [9]

Galkin, O. and Vekilov, P. G. (2000b). Control of protein crystal nucleation around the metastable liquid–liquid phase boundary. *Proc. Natl. Acad. Sci. USA*, 97(12):6277–6281. [8, 9, 108, 124]

George, A. and Wilson, W. W. (1994). Predicting protein crystallization from a dilute solution property. *Acta Crystallogr., Sect. D*, 50(4):361–365. [2, 9, 43, 101, 119, 152]

Gibaud, T., Cardinaux, F., Bergenholtz, J., Stradner, A., and Schurtenberger, P. (2011). Phase separation and dynamical arrest for particles interacting with mixed potentials-the case of globular proteins revisited. *Soft Matter*, 7:857–860. [9, 108, 120]

Gibaud, T. and Schurtenberger, P. (2009). A closer look at arrested spinodal decomposition in protein solutions. *J. Phys.: Condens. Matter*, 21(32):322201. [6, 9, 108, 120]

Glatter, O. and Kratky, O. (1982). *Small angle X-ray scattering*. Academic Press. [21, 36, 37, 39, 42, 79, 113, 119]

Gliko, O., Neumaier, N., Pan, W., Haase, I., Fischer, M., Bacher, A., Weinkauf, S., and Vekilov, P. G. (2005). A metastable prerequisite for the growth of lumazine synthase crystals. *J. Am. Chem. Soc.*, 127(10):3433–3438. [8]

Gliko, O., Pan, W., Katsonis, P., Neumaier, N., Galkin, O., Weinkauf, S., and Vekilov, P. G. (2007). Metastable liquid clusters in super- and undersaturated protein solutions. *J. Phys. Chem. B*, 111(12):3106–3114. [8, 109, 117]

Gögelein, C., Nägele, G., Tuinier, R., Gibaud, T., Stradner, A., and Schurtenberger, P. (2008). A simple patchy colloid model for the phase behavior of lysozyme dispersions. *J. Chem. Phys.*, 129(8):085102. [6, 8, 10, 52, 124, 150]

Gögelein, C., Wagner, D., Cardinaux, F., Nägele, G., and Egelhaaf, S. U. (2012). Effect of glycerol and dimethyl sulfoxide on the phase behavior of lysozyme: Theory and experiments. *J. Chem. Phys.*, 136(1):015102. [8]

Gordon, J. C., Myers, J. B., Folta, T., Shoja, V., Heath, L. S., and Onufriev, A. (2005). H++: a server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.*, 33(suppl 2):W368–W371. [80]

Gottschalk, M., Nilsson, H., Roos, H., and Halle, B. (2003). Protein self-association in solution: The bovine β-lactoglobulin dimer and octamer. *Protein Sci.*, 12(11):2404–2411. [112]

Gouy, M. (1910). Sur la constitution de la charge electrique a la surface d'un electrolyte. *J. Phys. Theor. Appl.*, 9(1):457–468. [2]

Grigsby, J., Blanch, H., and Prausnitz, J. (2001). Cloud-point temperatures for lysozyme in electrolyte solutions: effect of salt type, salt concentration and pH. *Biophys. Chem.*, 91(3):231 – 243. [147]

Grillo, I. (2008). *Soft-Matter Characterization*, chapter Small-angle neutron scattering and applications in soft condensed matter, pages 723–782. Springer. [36, 41]

Grimaldo, M. (2013). Protein self-diffusion studied by quasi-elastic neutron scattering: The effect of charge tuning using a trivalent salt. Master's thesis, Universität Tübingen, Germany. [150]

Groenewold, J. and Kegel, W. K. (2001). Anomalously large equilibrium clusters of colloids. *J. Phys. Chem. B*, 105(47):11702–11709. [8, 52, 113]

Groenewold, J. and Kegel, W. K. (2004). Colloidal cluster phases, gelation and nuclear matter. *J. Phys.: Condens. Matter*, 16(42):S4877. [8]

Grosberg, A. Y., Nguyen, T. T., and Shklovskii, B. I. (2002). Colloquium: The physics of charge inversion in chemical and biological systems. *Rev. Mod. Phys.*, 74(2):329–345. [5, 53, 76, 146]

Grouazel, S., Perez, J., Astier, J.-P., Bonneté, F., and Veesler, S. (2002). BPTI liquid-liquid phase separation monitored by light and small angle X-ray scattering. *Acta Crystallogr., Sect. D*, 58(10 Part 1):1560–1563. [108]

Guinier, A. and Fournet, G. (1955). *Small-angle scattering of X-rays*, volume 14. Wiley New York. [36]

Gunton, J., Shiryayev, A., and Pagan, D. (2007). *Protein Condensation: Kinetic Pathways to Crystallization and Disease*. Cambridge University Press. [1, 2, 96, 108]

Haas, C. and Drenth, J. (1999). Understanding protein crystallization on the basis of the phase diagram. *J. Cryst. Growth*, 196(24):388 – 394. [9]

Hansen, J.-P. and Hayter, J. B. (1982). A rescaled MSA structure factor for dilute charged colloidal dispersions. *Mol. Phys.*, 46(3):651–656. [38, 80]

Hansen, J.-P. and McDonald, I. R. (2006). *Theory of simple liquids*. Academic press, 3rd edition. [21, 79, 100, 148]

Harding, M. M. (2001). Geometry of metal–ligand interactions in proteins. *Acta Crystallogr., Sect. D*, 57(3):401–411. [5, 53, 64, 76]

Hardy, J. and Selkoe, D. J. (2002). The amyloid hypothesis of alzheimer's disease: Progress and problems on the road to therapeutics. *Science*, 297(5580):353–356. [2]

Haxton, T. K. and Whitelam, S. (2012). Design rules for the self-assembly of a protein crystal. *Soft Matter*, 8:3558–3562. [152]

Hayter, J. B. and Penfold, J. (1981). An analytic structure factor for macroion solutions. *Mol. Phys.*, 42(1):109–118. [38, 80]

Hayter, J. B. and Penfold, J. (1983). Determination of micelle structure and charge by neutron small-angle scattering. *Colloid Polym. Sci.*, 261:1022–1030. [22, 37]

Heinen, M., Holmqvist, P., Banchio, A. J., and Nägele, G. (2011). Pair structure of the hard-sphere Yukawa fluid: An improved analytic method versus simulations, Rogers-Young scheme, and experiment. *J. Chem. Phys.*, 134(4):044532. [101]

Heinen, M., Zanini, F., Roosen-Runge, F., Fedunova, D., Zhang, F., Hennig, M., Seydel, T., Schweins, R., Sztucki, M., Antalik, M., Schreiber, F., and Nägele, G. (2012). Viscosity and diffusion: crowding and salt effects in protein solutions. *Soft Matter*, 8:1404–1419. [66, 80, 82, 133, 135, 136, 138, 143]

Heyda, J., Hrobárik, T., and Jungwirth, P. (2009). Ion-specific interactions between halides and basic amino acids in water. *J. Phys. Chem. A*, 113(10):1969–1975. [4, 52]

Hofmeister, F. (1888). Zur Lehre von der Wirkung der Salze. *N-S Arch. Pharmacol.*, 24:247–260. [4, 33, 52, 76]

Hunter, R. J. (2000). *Foundations of Colloid Science*. Oxford University Press, USA, New York. [25]

Hutchens, S. B. and Wang, Z.-G. (2007). Metastable cluster intermediates in the condensation of charged macromolecule solutions. *J. Chem. Phys.*, 127(8):084912. [9]

Ianeselli, L., Zhang, F., Skoda, M. W. A., Jacobs, R. M. J., Martin, R. A., Callow, S., Prévost, S., and Schreiber, F. (2010). Protein-protein interactions in ovalbumin solutions studied by small-angle scattering: effect of ionic strength and the chemical nature of cations. *J. Phys. Chem. B*, 114(11):3776–83. [7, 34, 38, 53, 99, 105, 109, 135]

Ishimoto, C. and Tanaka, T. (1977). Critical behavior of a binary mixture of protein and salt water. *Phys. Rev. Lett.*, 39(8):474–477. [7]

Isihara, A. (1950). Determination of molecular shape by osmotic measurement. *J. Chem. Phys.*, 18(11):1446–1449. [38, 80, 136]

Jackson, G., Chapman, W. G., and Gubbins, K. E. (1988). Phase equilibria of associating fluids. *Mol. Phys.*, 65(1):1–31. [10, 124, 125, 127, 129, 130, 131]

Jacrot, B. (1976). The study of biological structures by neutron scattering from solution. *Rep. Prog. Phys.*, 39(10):911. [32]

Jacrot, B. and Zaccai, G. (1981). Determination of molecular weight by neutron scattering. *Biopolymers*, 20(11):2413–2426. [32]

James, R. O. and Healy, T. W. (1972). Adsorption of hydrolyzable metal ions at the oxide–water interface. III. A thermodynamic model of adsorption. *J. Colloid Interface Sci.*, 40(1):65 – 81. [61]

Jennings, B. R. and Parslow, K. (1988). Particle Size Measurement: The Equivalent Spherical Diameter. *Proc. R. Soc. Lond. A*, 419(1856):137–149. [101, 135, 137]

Jiang, T. and Wu, J. (2009). Cluster formation and bulk phase behavior of colloidal dispersions. *Phys. Rev. E*, 80:021401. [8]

Jion, A. I., Goh, L.-T., and Oh, S. K. (2006). Crystallization of IgG1 by mapping its liquid–liquid phase separation curves. *Biotechnol. Bioeng.*, 95(5):911–918. [9]

Johnston, K. P., Maynard, J. A., Truskett, T. M., Borwankar, A. U., Miller, M. A., Wilson, B. K., Dinin, A. K., Khan, T. A., and Kaczorowski, K. J. (2012). Concentrated dispersions of equilibrium protein nanoclusters that reversibly dissociate into active monomers. *ACS Nano*, 6(2):1357–1369. [8]

Jones, E., Oliphant, T., Peterson, P., et al. (2001). SciPy: Open source scientific tools for Python. [59, 73]

Jordan, E., Roosen-Runge, F., Leibfarth, S., Zhang, F., Sztucki, M., Kohlbacher, O., Hildebrandt, A., and Schreiber, F. (2013). Competing salt effects on reentrant phase behavior of protein solutions. *submitted and under review*. [13, 75]

Jungwirth, P. and Winter, B. (2008). Ions at aqueous interfaces: From water surface to hydrated proteins. *Annu. Rev. Phys. Chem.*, 59(1):343–366. [4, 5, 52, 53, 76]

Kalcher, I., Horinek, D., Netz, R. R., and Dzubiella, J. (2009). Ion specific correlations in bulk and at biointerfaces. *J. Phys.: Condens. Matter*, 21(42):424108. [76]

Keiderling, U. (2002). The new BerSANS-PC software for reduction and treatment of small angle neutron scattering data. *Appl. Phys. A*, 74(1):s1455–s1457. [36]

Keiderling, U. and Wiedenmann, A. (1995). New SANS instrument at the BER II reactor in Berlin, Germany. *Physica B*, 213214(0):895 – 897. [35]

Kern, N. and Frenkel, D. (2003). Fluid–fluid coexistence in colloidal systems with short-ranged strongly directional attraction. *J. Chem. Phys.*, 118(21):9882–9889. [6, 8, 10, 124, 150]

Kirby, B. J. and Hasselbrink, E. F. (2004). Zeta potential of microfluidic substrates: 1. theory, experimental techniques, and effects on separations. *Electrophoresis*, 25(2):187–202. [26, 27]

Kline, S. R. (2006). Reduction and analysis of SANS and USANS data using IGOR Pro. *J. Appl. Cryst.*, 39(6):895–900. [21, 38, 80, 99, 137]

Klotz, I. (1982). Numbers of receptor sites from scatchard graphs: facts and fantasies. *Science*, 217(4566):1247–1249. [5, 53, 61]

Knowles, T., Oppenheim, T., Buell, A., Chirgadze, D., and Welland, M. (2010). Nanostructured films from hierarchical self-assembly of amyloidogenic proteins. *Nat. Nanotechnol.*, 5(3):204–207. [2]

Koenig, S. H. (1975). Brownian motion of an ellipsoid. a correction to Perrin's results. *Biopolymers*, 14(11):2421–2423. [137]

Kotlarchyk, M. and Chen, S.-H. (1983). Analysis of small angle neutron scattering spectra from polydisperse interacting colloids. *J. Chem. Phys.*, 79(5):2461–2469. [22, 34]

Kowalczyk, P., Ciach, A., Gauden, P., and Terzyk, A. (2011). Equilibrium clusters in concentrated lysozyme protein solutions. *J. Colloid Interface Sci.*, 363(2):579 – 584. [8]

Kozak, M. (2005). Glucose isomerase from *Streptomyces rubiginosus* – potential molecular weight standard for small-angle X-ray scattering. *J. Appl. Cryst.*, 38(3):555–558. [42]

Kubíčková, A., Křížek, T., Coufal, P., Vazdar, M., Wernersson, E., Heyda, J., and Jungwirth, P. (2012). Overcharging in biological systems: Reversal of electrophoretic mobility of aqueous polyaspartate by multivalent cations. *Phys. Rev. Lett.*, 108:186101. [6, 68, 76, 144]

Kundu, S., Das, K., and Aswal, V. (2013). Modification of attractive and repulsive interactions among proteins in solution due to the presence of mono-, di- and tri-valent ions. *Chem. Phys. Lett.*, 578(0):115–119. [89]

Kuntz, Jr., I. and Kauzmann, W. (1974). Hydration of proteins and polypeptides. *Adv. Protein Chem.*, 28:239 – 345. [41]

Kunz, W., editor (2010). *Specific ion effects*. World Scientific Publishing Company. [4, 52]

Kurut, A. and Lund, M. (2013). Solution electrostatics beyond pH: a coarse grained approach to ion specific interactions between macromolecules. *Faraday Discuss.*, 160:271–278. [6, 52, 69]

Leggio, C., Galantini, L., and Pavel, N. V. (2008). About the albumin structure in solution: cigar expanded form versus heart normal shape. *Phys. Chem. Chem. Phys.*, 10:6741–6750. [134]

Leibfarth, S. (2011). Simulation of protein charge inversion by trivalent metal ion binding. Master's thesis, Universität Tübingen, Germany. [139, 140, 144]

Lenz, O. and Holm, C. (2008). Simulation of charge reversal in salty environments: Giant overcharging? *Eur. Phys. J. E*, 26:191–195. [5, 76]

Levin, Y. (2009). Polarizable ions at interfaces. *Phys. Rev. Lett.*, 102:147803. [4, 34]

Li, Y., Lubchenko, V., and Vekilov, P. G. (2011). The use of dynamic light scattering and brownian microscopy to characterize protein aggregation. *Rev. Sci. Instrum.*, 82(5):053106. [8]

Lide, D. R. (2009). *CRC handbook of chemistry and physics, 90th edition*. CRC press Boca Raton, Florida, USA. [63]

Likos, C. N. (2001). Effective interactions in soft condensed matter physics. *Phys. Rep.*, 348(45):267 – 439. [82]

Lindner, P. (2000). Water calibration at D11 verified with polymer samples. *J. Appl. Cryst.*, 33:807–811. [36]

Lindner, P. and Zemb, T. (2002). *Neutrons, X-rays, and Light: Scattering Methods Applied to Soft Condensed Matter*. Elsevier North-Holland. [20, 21, 37]

Lipfert, J. and Doniach, S. (2007). Small-angle X-ray scattering from RNA, proteins, and protein complexes. *Annu. Rev. Biophys. Biomol. Struct.*, 36:307–327. [32]

Liu, H., Kumar, S. K., and Sciortino, F. (2007). Vapor-liquid coexistence of patchy models: Relevance to protein phase behavior. *J. Chem. Phys.*, 127(8):084902. [10]

Liu, H., Kumar, S. K., Sciortino, F., and Evans, G. T. (2009). Vapor-liquid coexistence of fluids with attractive patches: An application of Wertheim's theory of association. *J. Chem. Phys.*, 130(4):044902. [10]

Liu, Y., Fratini, E., Baglioni, P., Chen, W.-R., and Chen, S.-H. (2005). Effective long-range attraction between protein molecules in solutions studied by small angle neutron scattering. *Phys. Rev. Lett.*, 95:118102. [109, 118]

Liu, Y., Porcar, L., Chen, J., Chen, W.-R., Falus, P., Faraone, A., Fratini, E., Hong, K., and Baglioni, P. (2011). Lysozyme protein solution with an intermediate range order structure. *J. Phys. Chem. B*, 115(22):7238–7247. [8]

Liu, Y., Wang, X., and Ching, C. B. (2010). Toward further understanding of lysozyme crystallization: Phase diagram, proteinprotein interaction, nucleation kinetics, and growth kinetics. *Cryst. Growth Des.*, 10(2):548–558. [117]

Lomakin, A., Asherie, N., and Benedek, G. B. (1999). Aeolotopic interactions of globular proteins. *Proc. Natl. Acad. Sci. USA*, 96(17):9465–9468. [10]

Lu, J. (1999). Neutron reflection study of globular protein adsorption at planar interfaces. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.*, 95:3–46. [45]

Lu, P. J., Zaccarelli, E., Ciulla, F., Schofield, A. B., Sciortino, F., and Weitz, D. A. (2008). Gelation of particles with short-range attraction. *Nature*, 453(7194):499–503. [108, 120]

Lund, M. and Jönsson, B. (2005). On the charge regulation of proteins. *Biochemistry*, 44(15):5722–5727. [6, 52]

Lund, M., Vrbka, L., and Jungwirth, P. (2008). Specific ion binding to nonpolar surface patches of proteins. *J. Am. Chem. Soc.*, 130(35):11582–11583. [76]

Lutsko, J. F. (2012). On the role of metastable intermediate states in the homogeneous nucleation of solids from solution. *Adv. Chem. Phys.*, 151:137. [120]

Lutsko, J. F. and Nicolis, G. (2006). Theoretical evidence for a dense fluid precursor to crystallization. *Phys. Rev. Lett.*, 96:046102. [9, 96]

Lydersen, E. (1990). The solubility and hydrolysis of aqueous aluminium hydroxides in dilute fresh waters at different temperatures. *Nord. Hydrol.*, 21(3):195–204. [61, 62, 63]

Lyklema, J. (2006). Overcharging, charge reversal: Chemistry or physics? *Colloid Surface A*, 291(1-3):3 – 12. [6, 53, 65, 76, 146]

Makarov, V., Pettitt, B. M., and Feig, M. (2002). Solvation and hydration of proteins and nucleic acids: A theoretical view of simulation and experiment. *Acc. Chem. Res.*, 35(6):376–384. [45]

Manning, G. S. (1969). Limiting laws and counterion condensation in polyelectrolyte solutions: I. colligative properties. *J. Chem. Phys.*, 51(3):924–933. [5]

Manning, G. S. (1978). The molecular theory of polyelectrolyte solutions with applications to the electrostatic properties of polynucleotides. *Q. Rev. Biophys.*, 11(02):179–246. [52, 76]

Martin, R. B. (1986). The chemistry of aluminum as related to biology and medicine. *Clin. Chem.*, 32(10):1797–806. [5, 61, 63, 64]

Mason, B. D., van Enk, J. Z., Zhang, L., Remmele, Jr., R. L., and Zhang, J. (2010). Liquid-liquid phase separation of a monoclonal antibody and nonmonotonic influence of Hofmeister anions. *Biophys. J.*, 99(11):3792 – 3800. [8]

McManus, J. J., Lomakin, A., Ogun, O., Pande, A., Basan, M., Pande, J., and Benedek, G. B. (2007). Altered phase diagram due to a single point mutation in human γD-crystallin. *Proc. Natl. Acad. Sci. USA*, 104(43):16856–16861. [10, 153]

McMillan, Jr., W. G. and Mayer, J. E. (1945). The statistical thermodynamics of multicomponent systems. *J. Chem. Phys.*, 13(7):276–305. [102]

McPherson, A. (2004). Introduction to protein crystallization. *Methods*, 34(3):254 – 265. Macromolecular Crystallization. [2]

Menon, M. K. and Zydney, A. L. (1998). Measurement of protein charge and ion binding using capillary electrophoresis. *Anal. Chem.*, 70(8):1581–1584. [25, 54]

Menon, S. V. G., Manohar, C., and Rao, K. S. (1991). A new interpretation of the sticky hard sphere model. *J. Chem. Phys.*, 95(12):9186–9190. [99]

Merzel, F. and Smith, J. C. (2002). Is the first hydration shell of lysozyme of higher density than bulk water? *Proc. Natl. Acad. Sci. USA*, 99(8):5378–5383. [33]

Messina, R. (2009). Electrostatics in soft matter. *J. Phys.: Condens. Matter*, 21(11):113102. [2, 3, 5, 53, 76]

Mintova, S., Olson, N. H., Valtchev, V., and Bein, T. (1999). Mechanism of zeolite a nanocrystal growth from colloids at room temperature. *Science*, 283(5404):958–960. [109, 117]

Moitzi, C., Donato, L., Schmitt, C., Bovetto, L., Gillies, G., and Stradner, A. (2011). Structure of β-lactoglobulin microgels formed during heating as revealed by small-angle X-ray scattering and light scattering. *Food Hydrocolloids*, 25(7):1766 – 1774. [112]

Moreira, A. G. and Netz, R. R. (2000). Strong-coupling theory for counter-ion distributions. *Europhys. Lett.*, 52(6):705. [5, 53, 76]

Mossa, S., Sciortino, F., Tartaglia, P., and Zaccarelli, E. (2004). Ground-state clusters for short-range attractive and long-range repulsive potentials. *Langmuir*, 20(24):10756–10763. [8]

Munson, P. J. and Rodbard, D. (1980). Ligand: A versatile computerized approach for characterization of ligand-binding systems. *Anal. Biochem.*, 107(1):220 – 239. [61]

Muschol, M. and Rosenberger, F. (1997). Liquid–liquid phase separation in supersaturated lysozyme solutions and associated precipitate formation/crystallization. *J. Chem. Phys.*, 107(6):1953–1962. [8, 96, 97, 99, 108, 124]

Mylonas, E. and Svergun, D. I. (2007). Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J. Appl. Cryst.*, 40(s1):s245–s249. [40]

Nägele, G. (1996). On the dynamics and structure of charge-stabilized suspensions. *Phys. Rep.*, 272(5-6):215–372. [80, 82]

Narayanan, T. (2008). *Soft Matter Characterization*, chapter Synchroton small-angle X-ray Scattering, page 899ff. Springer. [79, 105, 110]

Navrotsky, A. (2004). Energetic clues to pathways to biomineralization: Precursors, clusters, and nanoparticles. *Proc. Natl. Acad. Sci. USA*, 101(33):12096–12101. [9, 108, 109, 117]

Neal, B. and Lenhoff, A. (1995). Excluded volume contribution to the osmotic second virial coefficient for proteins. *AIChE J.*, 41(4):1010–1014. [45, 137]

Neal, B. L., Asthagiri, D., Velev, O. D., Lenhoff, A. M., and Kaler, E. W. (1999). Why is the osmotic second virial coefficient related to protein crystallization? *J. Cryst. Growth*, 196(2-4):377 – 387. [10]

Nguyen, T. T., Grosberg, A. Y., and Shklovskii, B. I. (2000a). Macroions in salty water with multivalent ions: Giant inversion of charge. *Phys. Rev. Lett.*, 85:1568–1571. [5, 6, 76]

Nguyen, T. T., Rouzina, I., and Shklovskii, B. I. (2000b). Reentrant condensation of DNA induced by multivalent counterions. *J. Chem. Phys.*, 112(5):2562–2568. [7, 77, 92]

Ninham, B. W. and Nostro, P. L. (2010). *Molecular Forces and Self Assembly: In Colloid, Nano Sciences and Biology*. Cambridge University Press Cambridge. [2, 4, 34]

Noro, M. G. and Frenkel, D. (2000). Extended corresponding-states behavior for particles with variable range attractions. *J. Chem. Phys.*, 113(8):2941–2944. [7, 99, 100, 103, 128, 147, 148]

Nossal, R., Glinka, C. J., and Chen, S.-H. (1986). SANS studies of concentrated protein solutions. I. Bovine serum albumin. *Biopolymers*, 25(6):1157–1175. [34, 39, 41, 45]

Ohshima, H. (1994). A simple expression for henry's function for the retardation effect in electrophoresis of spherical colloidal particles. *J. Colloid Interface Sci.*, 168(1):269 – 271. [26, 55]

Ohshima, H., Healy, T. W., and White, L. R. (1982). Accurate analytic expressions for the surface charge density/surface potential relationship and double-layer potential distribution for a spherical colloidal particle. *J. Colloid Interface Sci.*, 90(1):17 – 26. [27, 28, 55]

Oliphant, T. E. (2007). Python for scientific computing. *Comput. Sci. Eng.*, 9(3):10–20. [59, 73]

Ortore, M. G., Sinibaldi, R., Spinozzi, F., Carsughi, F., Clemens, D., Bonincontro, A., and Mariani, P. (2008). New insights into urea action on proteins: A SANS study of the lysozyme case. *J. Phys. Chem. B*, 112(41):12881–12887. [45]

Oxtoby, D. W. (1998). Nucleation of first-order phase transitions. *Acc. Chem. Res.*, 31(2):91–97. [120]

Pace, C. N., Alston, R. W., and Shaw, K. L. (2000). Charge–charge interactions influence the denatured state ensemble and contribute to protein stability. *Protein Sci.*, 9(7):1395–1398. [6, 52]

Paliwal, A., Asthagiri, D., Abras, D., Lenhoff, A., and Paulaitis, M. (2005). Light-scattering studies of protein solutions: Role of hydration in weak protein-protein interactions. *Biophys. J.*, 89(3):1564 – 1573. [45]

Pan, W., Galkin, O., Filobelo, L., Nagel, R. L., and Vekilov, P. G. (2007). Metastable mesoscopic clusters in solutions of sickle-cell hemoglobin. *Biophys. J.*, 92(1):267 – 277. [8, 109, 117, 120]

Pan, W., Vekilov, P. G., and Lubchenko, V. (2010). Origin of anomalous mesoscopic phases in protein solutions. *J. Phys. Chem. B*, 114(22):7620–30. [8, 109, 117, 120]

Pande, A., Pande, J., Asherie, N., Lomakin, A., Ogun, O., King, J., and Benedek, G. B. (2001). Crystal cataracts: Human genetic cataract caused by protein crystallization. *Proc. Natl. Acad. Sci. USA*, 98(11):6116–6120. [96]

Park, E. J. and Bae, Y. C. (2004). Cloud-point temperatures of lysozyme in electrolyte solutions by thermooptical analysis technique. *Biophys. Chem.*, 109(1):169 – 188. [147]

Pastore, C., Franzese, M., Sica, F., Temussi, P., and Pastore, A. (2007). Understanding the binding properties of an unusual metal-binding protein – a study of bacterial frataxin. *FEBS J.*, 274(16):4199–4210. [5, 64]

Pawar, A. B. and Kretzschmar, I. (2010). Fabrication, assembly, and application of patchy particles. *Macromol. Rapid Commun.*, 31(2):150–168. [101]

Pedersen, J. S. (1997). Analysis of small-angle scattering data from colloids and polymer solutions: modeling and least-squares fitting. *Adv. Colloid Interface Sci.*, 70(0):171 – 210. [21, 37, 41, 79, 137]

Perkins, S. J. (2001). X-ray and neutron scattering analyses of hydration shells: a molecular interpretation based on sequence predictions and modelling fits. *Biophys. Chem.*, 93(23):129–139. [32, 33, 41]

Perkins, S. J., Okemefuna, A. I., Fernando, A. N., Bonner, A., Gilbert, H. E., and Furtado, P. B. (2008). Xray and neutron scattering data and their constrained molecular modeling. In Correia, D. J. J. and Dr. H. William Detrich, I., editors, *Biophysical Tools for Biologists, Volume One: In Vitro Techniques*, volume 84 of *Methods in Cell Biology*, pages 375 – 423. Academic Press. [32, 36]

Permyakov, E. Y. (2009). *Metalloproteomics*. John Wiley & Sons, Hoboken, New Jersey. [5, 76, 133, 139, 144]

Perrin, F. (1934). Mouvement Brownien d'un ellipsoide (I). Dispersion dielectrique pour des molecules ellipsoidales. *J. Phys.-Paris*, 7(1):497–511. [137]

Peters, T. (1996). *All about albumin: biochemistry, genetics, and medical applications*. Academic Press San Diego, CA. [60, 69, 134]

Petoukhov, M. and Svergun, D. (2006). Joint use of small-angle X-ray and neutron scattering to study biological macromolecules in solution. *Eur. Biophys. J.*, 35(7):567–576. [32]

Petsev, D. N. and Vekilov, P. G. (2000). Evidence for non-DLVO hydration interactions in solutions of the protein apoferritin. *Phys. Rev. Lett.*, 84(6):1339–1342. [3]

Petsev, D. N., Wu, X., Galkin, O., and Vekilov, P. G. (2003). Thermodynamic functions of concentrated protein solutions from phase equilibria. *J. Phys. Chem. B*, 107(16):3921–3926. [99]

Piazza, R. (2000). Interactions and phase transitions in protein solutions. *Curr. Opin. Colloid Interface Sci.*, 5(1-2):38 – 43. [32]

Piazza, R. (2004). Protein interactions and association: an open challenge for colloid science. *Curr. Opin. Colloid Interface Sci.*, 8(6):515 – 522. [32]

Piazza, R. and Iacopini, S. (2002). Transient clustering in a protein solution. *Eur. Phys. J. E*, 7(1):45–48. [8]

Pipich, V. (2012). QtiKWS: user-friendly program for reduction, visualization, analysis and fit of SA(N)S data. http://www.qtikws.de. [36]

Poon, W. (2002). The physics of a model colloidal-polymer mixture. *J. Phys.: Condens. Matter*, 14:R859. [96, 99]

Poon, W. C. K. (1997). Crystallization of globular proteins. *Phys. Rev. E*, 55:3762–3764. [9]

Porcar, L., Falus, P., Chen, W.-R., Faraone, A., Fratini, E., Hong, K., Baglioni, P., and Liu, Y. (2010). Formation of the dynamic clusters in concentrated lysozyme protein solutions. *J. Phys. Chem. Lett.*, 1(1):126–129. [8]

Protopapas, P., Andersen, H. C., and Parlee, N. A. D. (1973). Theory of transport in liquid metals. I. Calculation of self-diffusion coefficients. *J. Chem. Phys.*, 59(1):15–25. [139]

Roosen-Runge, F., Heck, B. S., Zhang, F., Kohlbacher, O., and Schreiber, F. (2013a). Interplay of pH and binding of multivalent metal ions: Charge inversion and reentrant condensation in protein solutions. *J. Phys. Chem. B*, 117(18):5777–5787. [12, 51, 77, 78, 81, 82, 86, 90, 126]

Roosen-Runge, F., Hennig, M., Seydel, T., Zhang, F., Skoda, M. W., Zorn, S., Jacobs, R. M., Maccarini, M., Fouquet, P., and Schreiber, F. (2010). Protein diffusion in crowded electrolyte solutions. *BBA-Proteins Proteom*, 1804(1):68 – 75. [34, 68, 82, 84, 143]

Roosen-Runge, F., Hennig, M., Zhang, F., Jacobs, R. M. J., Sztucki, M., Schober, H., Seydel, T., and Schreiber, F. (2011). Protein self-diffusion in crowded solutions. *Proc. Natl. Acad. Sci. USA*, 108(29):11815–11820. [34, 41, 101, 126, 133, 135, 136, 137, 143]

Roosen-Runge, F., Zhang, F., Schreiber, F., and Roth, R. (2013b). Proteins as particles with attractive patches. *submitted and under review*. [15, 91, 92, 123]

Rosenbaum, D., Zamora, P., and Zukoski, C. (1996). Phase behavior of small attractive colloidal particles. *Phys. Rev. Lett.*, 76(1):150–153. [7, 96, 101]

Roth, R., Evans, R., and Louis, A. A. (2001). Theory of asymmetric nonadditive binary hard-sphere mixtures. *Phys. Rev. E*, 64(5):051202. [102, 103, 127]

Russo, J., Tavares, J. M., Teixeira, P. I. C., Telo da Gama, M. M., and Sciortino, F. (2011). Reentrant phase diagram of network fluids. *Phys. Rev. Lett.*, 106:085703. [124]

Saminathan, M., Antony, T., Shirahata, A., Sigal, L. H., Thomas, T., and Thomas, T. J. (1999). Ionic and structural specificity effects of natural and synthetic polyamines on the aggregation and resolubilization of single-, double-, and triple-stranded DNA. *Biochemistry*, 38(12):3821–3830. [7, 77, 92, 146]

Sandhu, R. S. (1977). A thermodynamic study of complexation reaction of yttrium(III), lanthanum(III) and cerium(III) with tyrosine. *Monatsh. Chem.*, 108:51–55. [64, 77, 81, 144]

Scatchard, G. (1949). The attractions of proteins for small molecules and ions. *Ann. NY. Acad. Sci.*, 51(4):660–672. [5, 52, 53, 61]

Scatchard, G., Scheinberg, I. H., and Armstrong, S. H. (1950). Physical chemistry of protein solutions. v. the combination of human serum albumin with thiocyanate ion1a. *J. Am. Chem. Soc.*, 72(1):540–546. [48]

Schurtenberger, P., Chamberlin, R. A., Thurston, G. M., Thomson, J. A., and Benedek, G. B. (1989). Observation of critical phenomena in a protein-water solution. *Phys. Rev. Lett.*, 63(19):2064–2067. [7]

Schwierz, N., Horinek, D., and Netz, R. R. (2010). Reversed anionic Hofmeister series: the interplay of surface charge and surface polarity. *Langmuir*, 26(10):7370–7379. [4, 34, 52]

Sciortino, F., Mossa, S., Zaccarelli, E., and Tartaglia, P. (2004). Equilibrium cluster phases and low-density arrested disordered states: The role of short-range attraction and long-range repulsion. *Phys. Rev. Lett.*, 93:055701. [8, 9]

Sciortino, F. and Zaccarelli, E. (2011). Reversible gels of patchy particles. *Curr. Opin. Solid. St. M*, 15(6):246 – 253. [108, 120]

Sear, R. P. (1999). Phase behavior of a simple model of globular proteins. *J. Chem. Phys.*, 111(10):4800–4806. [6, 8, 10, 124, 127, 132, 150]

Segrè, P. N., Prasad, V., Schofield, A. B., and Weitz, D. A. (2001). Glasslike kinetic arrest at the colloidal-gelation transition. *Phys. Rev. Lett.*, 86:6042–6045. [8, 9]

Senff, H. and Richtering, W. (1999). Temperature sensitive microgel suspensions: Colloidal phase behavior and rheology of soft spheres. *J. Chem. Phys.*, 111(4):1705–1711. [139]

Shah, M., Galkin, O., and Vekilov, P. G. (2004). Smooth transition from metastability to instability in phase separating protein solutions. *J. Chem. Phys.*, 121(15):7505–7512. [97]

Shiryayev, A., Pagan, D. L., Gunton, J. D., Rhen, D. S., Saxena, A., and Lookman, T. (2005). Role of solvent for globular proteins in solution. *J. Chem. Phys.*, 122(23):234911. [151]

Shklovskii, B. I. (1999). Screening of a macroion by multivalent ions: Correlation-induced inversion of charge. *Phys. Rev. E*, 60(5):5802–5811. [5]

Shukla, A., Mylonas, E., Di Cola, E., Finet, S., Timmins, P., Narayanan, T., and Svergun, D. I. (2008). Absence of equilibrium cluster phase in concentrated lysozyme solutions. *Proc. Natl. Acad. Sci. USA*, 105(13):5075–5080. [8, 109]

Siezen, R. J., Fisch, M. R., Slingsby, C., and Benedek, G. B. (1985). Opacification of γ-crystallin solutions from calf lens in relation to cold cataract formation. *Proc. Natl. Acad. Sci. USA*, 82(6):1701–1705. [7]

Sinibaldi, R., Ortore, M., Spinozzi, F., de Souza Funari, S., Teixeira, J., and Mariani, P. (2008). SANS/SAXS study of the BSA solvation properties in aqueous urea solutions via a global fit approach. *Eur. Biophys. J.*, 37:673–681. [32, 40, 48]

Sinibaldi, R., Ortore, M. G., Spinozzi, F., Carsughi, F., Frielinghaus, H., Cinelli, S., Onori, G., and Mariani, P. (2007). Preferential hydration of lysozyme in water/glycerol mixtures: A small-angle neutron scattering study. *J. Chem. Phys.*, 126(23):235101. [32]

Sipe, J. D. and Cohen, A. S. (2000). Review: History of the amyloid fibril. *J. Struct. Biol.*, 130(23):88 – 98. [2]

Sober, H. A. (1970). *CRC Handbook of Biochemistry: Selected data for molecular biology*. CRC. [110]

Soraruf, D., Roosen-Runge, F., Grimaldo, M., Zanini, F., Schweins, R., Seydel, T., Zhang, F., Oettel, M., and Schreiber, F. (2013). Protein cluster formation in aqueous solution in the presence of multivalent metal ions – a light scattering study. *submitted and under review*. [150]

Spohr, E., Hribar, B., and Vlachy, V. (2002). Mechanism of macroion-macroion clustering induced by the presence of trivalent counterions. *J. Phys. Chem. B*, 106(9):2343–2348. [8]

Stradner, A., Cardinaux, F., and Schurtenberger, P. (2006). A small-angle scattering study on equilibrium clusters in lysozyme solutions. *J. Phys. Chem. B*, 110(42):21222–31. [8, 32]

Stradner, A., Sedgwick, H., Cardinaux, F., Poon, W. C. K., Egelhaaf, S. U., and Schurtenberger, P. (2004). Equilibrium cluster formation in concentrated protein solutions and colloids. *Nature*, 432:492–495. [8, 32, 52, 109, 113, 124]

Svergun, D. (1999). Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, 76(6):2879 – 2886. [21]

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(6):768–773. [40, 112, 114]

Svergun, D. I. and Koch, M. H. J. (2003). Small-angle scattering studies of biological macromolecules in solution. *Rep. Prog. Phys.*, 66(10):1735. [79]

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. USA*, 95(5):2267–2272. [32, 33, 40, 41, 136]

Sztucki, M., Narayanan, T., Belina, G., Moussaïd, A., Pignon, F., and Hoekstra, H. (2006). Kinetic arrest and glass-glass transition in short-ranged attractive colloids. *Phys. Rev. E*, 74(5):051504. [9, 110]

Tainer, J. A., Roberts, V. A., and Getzoff, E. D. (1991). Metal-binding sites in proteins. *Curr. Opin. Biotechnol.*, 2(4):582 – 591. [5, 53]

Talanquer, V. and Oxtoby, D. W. (1998). Crystal nucleation in the presence of a metastable critical point. *J. Chem. Phys.*, 109(1):223–227. [117, 119, 120]

Tanaka, T. and Benedek, G. B. (1975). Observation of protein diffusivity in intact human and bovine lenses with application to cataract. *Invest. Ophth. Vis. Sci.*, 14(6):449–56. [7]

Tanaka, T., Ishimoto, C., and Chylack, L. (1977). Phase separation of a protein-water mixture in cold cataract in the young rat lens. *Science*, 197(4307):1010–1012. [7]

Tanford, C. and Kirkwood, J. G. (1957). Theory of protein titration curves. I. General equations for impenetrable spheres. *J. Am. Chem. Soc.*, 79(20):5333–5339. [5, 52, 61, 63]

Tanford, C., Swanson, S. A., and Shore, W. S. (1955). Hydrogen ion equilibria of bovine serum albumin. *J. Am. Chem. Soc.*, 77(24):6414–6421. [5, 52, 61]

Tardieu, A., Verge, A. L., Malfois, M., Bonnet, F., Finet, S., Ris-Kautt, M., and Belloni, L. (1999). Proteins in solution: from X-ray scattering intensities to interaction potentials. *J. Cryst. Growth*, 196(2–4):193–203. [37, 43]

Terao, T. (2006). Nonlinear screening on the charged surfaces with trivalent and tetravalent salt ions: Monte carlo simulations. *Colloid Surface A*, 273(1-3):141 – 146. [5]

Tessier, P. M., Vandrey, S. D., Berger, B. W., Pazhianur, R., Sandler, S. I., and Lenhoff, A. M. (2002). Self-interaction chromatography: a novel screening method for rational protein crystallization. *Acta Crystallogr., Sect. D*, 58(10 Part 1):1531–1535. [43]

Thomson, J. A., Schurtenberger, P., Thurston, G. M., and Benedek, G. B. (1987). Binary liquid phase separation and critical phenomena in a protein/water solution. *Proc. Natl. Acad. Sci. USA*, 84(20):7079–7083. [7, 124]

Tipping, E. (1998). Humic ion-binding model VI: An improved description of the interactions of protons and metal ions with humic substances. *Aquat. Geochem.*, 4:3–47. [5, 53, 61]

Trapp, G. A. (1983). Plasma aluminum is bound to transferrin. *Life Sci.*, 33(4):311 – 316. [5, 64]

Travesset, A. and Vangaveti, S. (2009). Electrostatic correlations at the stern layer: Physics or chemistry? *J. Chem. Phys.*, 131(18):185102. [6, 53, 65, 146]

van der Kooij, F. M., Boek, E. S., and Philipse, A. P. (2001). Rheology of dilute suspensions of hard platelike colloids. *J. Colloid Interface Sci.*, 235(2):344 – 349. [138]

Vega, C. and Monson, P. A. (1998). Solid–fluid equilibrium for a molecular model with short ranged directional forces. *The Journal of Chemical Physics*, 109(22):9938–9949. [10]

Vekilov, P. G. (2004). Dense liquid precursor for the nucleation of ordered solid phases from solution. *Cryst. Growth Des.*, 4(4):671–685. [9, 96, 97, 108, 152]

Vekilov, P. G. (2010a). Nucleation. *Cryst. Growth Des.*, 10(12):5007–5019. [9]

Vekilov, P. G. (2010b). The two-step mechanism of nucleation of crystals in solution. *Nanoscale*, 2:2346–2357. [9, 108, 109, 117, 119, 120]

Vekilov, P. G., Feeling-Taylor, A. R., Yau, S.-T., and Petsev, D. (2002). Solvent entropy contribution to the free energy of protein crystallization. *Acta Crystallogr., Sect. D*, 58(10 Part 1):1611–1616. [151]

Verwey, E. and Overbeek, J. (1948). *Theory of the Stability of Lyophobic Colloids*. Elsevier, Amsterdam. [3, 52, 66]

Vilker, V. L., Colton, C. K., and Smith, K. A. (1981). The osmotic pressure of concentrated protein solutions: Effect of concentration and pH in saline solutions of bovine serum albumin. *J. Colloid Interface Sci.*, 79(2):548 – 566. [43]

Vivarès, D., Kaler, E. W., and Lenhoff, A. M. (2005). Quantitative imaging by confocal scanning fluorescence microscopy of protein crystallization *via* liquid–liquid phase separation. *Acta Crystallogr., Sect. D*, 61(6):819–825. [117, 120, 149]

Vliegenthart, G. A. and Lekkerkerker, H. N. W. (2000). Predicting the gas–liquid critical point from the second virial coefficient. *J. Chem. Phys.*, 112(12):5364–5369. [2, 9, 98, 99, 100, 101, 103, 118, 119, 147, 152]

Vogtt, K., Javid, N., Alvarez, E., Sefcik, J., and Bellissent-Funel, M.-C. (2011). Tracing nucleation pathways in protein aggregation by using small angle scattering methods. *Soft Matter*, 7:3906–3914. [112]

Wang, Y. and Annunziata, O. (2008). Liquid–liquid phase transition of protein aqueous solutions isothermally induced by protein cross-linking. *Langmuir*, 24(6):2799–2807. [151]

Wang, Y., Lomakin, A., Latypov, R. F., and Benedek, G. B. (2011). Phase separation in solutions of monoclonal antibodies and the effect of human serum albumin. *Proc. Natl. Acad. Sci. USA*, 108(40):16606–16611. [8, 108]

Wang, Y., Lomakin, A., McManus, J. J., Ogun, O., and Benedek, G. B. (2010). Phase behavior of mixtures of human lens proteins Gamma D and Beta B1. *Proc. Natl. Acad. Sci. USA*, 107(30):13282–7. [8, 96]

Wardell, M., Wang, Z., Ho, J. X., Robert, J., Ruker, F., Ruble, J., and Carter, D. C. (2002). The atomic structure of human methemalbumin at 1.9 Å. *Biochem. Biophys. Res. Commun.*, 291(4):813 – 819. [80, 81]

Warren, P. B. (2002). Simple models for charge and salt effects in protein crystallization. *J. Phys.: Condens. Matter*, 14(33):7617. [10]

Warshel, A. and Levitt, M. (1976). Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.*, 103(2):227 – 249. [6, 52]

Wentzel, N. and Gunton, J. D. (2007). Liquid–liquid coexistence surface for lysozyme: role of salt type and salt concentration. *J. Phys. Chem. B*, 111(6):1478–1481. [8, 124, 147, 151]

Wentzel, N. and Gunton, J. D. (2008). Effect of solvent on the phase diagram of a simple anisotropic model of globular proteins. *J. Phys. Chem. B*, 112(26):7803–7809. [8, 151]

Wertheim, M. (1984a). Fluids with highly directional attractive forces. I. Statistical thermodynamics. *J. Stat. Phys.*, 35:19–34. [10, 101, 124, 130]

Wertheim, M. (1984b). Fluids with highly directional attractive forces. II. Thermodynamic perturbation theory and integral equations. *J. Stat. Phys.*, 35:35–47. [10, 124, 130]

Wertheim, M. (1986a). Fluids with highly directional attractive forces. III. Multiple attraction sites. *J. Stat. Phys.*, 42:459–476. [10, 124, 130]

Wertheim, M. (1986b). Fluids with highly directional attractive forces. IV. Equilibrium polymerization. *J. Stat. Phys.*, 42:477–492. [10, 124, 130]

White, L. R. (1977). Approximate analytic solution of the Poisson-Boltzmann equation for a spherical colloidal particle. *J. Chem. Soc., Faraday Trans. 2*, 73(5):577–596. [28]

Whitelam, S. (2010a). Control of pathways and yields of protein crystallization through the interplay of nonspecific and specific attractions. *Phys. Rev. Lett.*, 105(8):088102. [6, 11, 52, 124, 150, 152]

Whitelam, S. (2010b). Nonclassical assembly pathways of anisotropic particles. *J. Chem. Phys.*, 132(19):194901. [152]

Wiersema, P. H., Loeb, A. L., and Overbeek, J. T. G. (1966). Calculation of the electrophoretic mobility of a spherical colloid particle. *J. Colloid Interface Sci.*, 22(1):78 – 99. [28]

Williams, P. and Peacocke, A. (1967). The binding of calcium and yttrium ions to a glycoprotein from bovine cortical bon. *Biochem. J.*, 105(3):1177–85. [5, 64]

Wolde, P. R. t. and Frenkel, D. (1997). Enhancement of protein crystal nucleation by critical density fluctuations. *Science*, 277(5334):1975–1978. [9, 96, 99, 101, 109, 117, 119, 120, 124, 152]

Wolf, M., Roosen-Runge, F., Zhang, F., Roth, R., Skoda, M. W. A., Jacobs, R. M. J., Sztucki, M., and Schreiber, F. (2013). Effective interactions in protein-salt solutions approaching liquid-liquid phase separation. *in preparation*. [151]

Wood, S. A. (1990). The aqueous geochemistry of the rare-earth elements and yttrium: 1. Review of available low-temperature data for inorganic complexes and the inorganic ree speciation of natural waters. *Chem. Geol.*, 82(0):159 – 186. [61, 62]

Xu, H. E., Lambert, M. H., Montana, V. G., Plunket, K. D., Moore, L. B., Collins, J. L., Oplinger, J. A., Kliewer, S. A., Gampe, R. T., McKee, D. D., Moore, J. T., and Willson, T. M. (2001). Structural determinants of ligand binding selectivity between the peroxisome proliferator-activated receptors. *Proc. Natl. Acad. Sci. USA*, 98(24):13919–13924. [101]

Yamashita, M. M., Wesson, L., Eisenman, G., and Eisenberg, D. (1990). Where metal ions bind in proteins. *Proc. Natl. Acad. Sci. USA*, 87(15):5648–5652. [5, 53]

Zaccai, G. and Jacrot, B. (1983). Small angle neutron scattering. *Annu. Rev. Biophys. Bio.*, 12(1):139–157. [33]

Zaccarelli, E. (2007). Colloidal gels: equilibrium and non-equilibrium routes. *J. Phys.: Condens. Matter*, 19(32):323101. [101]

Zhang, F., Roosen-Runge, F., Sauter, A., Roth, R., Skoda, M. W. A., Jacobs, R., Sztucki, M., and Schreiber, F. (2012a). The role of cluster formation and metastable liquid-liquid phase separation in protein crystallization. *Faraday Discuss.*, 159:313–325. [14, 53, 107, 129]

Zhang, F., Roosen-Runge, F., Skoda, M. W. A., Jacobs, R. M. J., Wolf, M., Callow, P., Frielinghaus, H., Pipich, V., Prevost, S., and Schreiber, F. (2012b). Hydration and interactions in protein solutions containing concentrated electrolytes studied by small-angle scattering. *Phys. Chem. Chem. Phys.*, 14:2483–2493. [12, 31, 53, 55, 80, 136]

Zhang, F., Roth, R., Wolf, M., Roosen-Runge, F., Skoda, M. W. A., Jacobs, R. M. J., Sztucki, M., and Schreiber, F. (2012c). Charge-controlled metastable liquid-liquid phase separation in protein solutions as a universal pathway towards crystallization. *Soft Matter*, 8:1313–1316. [14, 53, 57, 70, 84, 95, 108, 109, 124, 126, 129]

Zhang, F., Skoda, M. W. A., Jacobs, R. M. J., Martin, R. A., Martin, C. M., and Schreiber, F. (2007). Protein Interactions Studied by SAXS: Effect of Ionic Strength and Protein Concentration for BSA in Aqueous Solutions. *J. Phys. Chem. B*, 111(1):251–259. [34, 38, 48, 49, 53, 79, 84, 99, 136, 144]

Zhang, F., Skoda, M. W. A., Jacobs, R. M. J., Zorn, S., Martin, R. A., Martin, C. M., Clark, G. F., Weggler, S., Hildebrandt, A., Kohlbacher, O., and Schreiber, F. (2008). Reentrant condensation of proteins in solution induced by multivalent counterions. *Phys. Rev. Lett.*, 101(14):148101. [7, 34, 38, 53, 56, 57, 68, 77, 80, 81, 86, 97, 99, 107, 109, 110, 113, 119, 124, 140, 144, 145, 146]

Zhang, F., Weggler, S., Ziller, M. J., Ianeselli, L., Heck, B. S., Hildebrandt, A., Kohlbacher, O., Skoda, M. W. A., Jacobs, R. M. J., and Schreiber, F. (2010). Universality of protein reentrant condensation in solution induced by multivalent metal ions. *Proteins*, 78(16):3450–3457. [5, 7, 34, 53, 54, 56, 57, 60, 64, 68, 69, 77, 80, 86, 97, 109, 110, 111, 113, 124, 140, 144, 145, 146, 147]

Zhang, F., Zocher, G., Sauter, A., Stehle, T., and Schreiber, F. (2011). Novel approach to controlled protein crystallization through ligandation of yttrium cations. *J. Appl. Cryst.*, 44(4):755–762. [7, 68, 69, 90, 91, 92, 97, 108, 109, 111, 113, 114, 116, 124, 126, 149]

List of Own Publications

1. Felix Roosen-Runge, Marcus Hennig, Tilo Seydel, Fajun Zhang, Maximilian W.A. Skoda, Stefan Zorn, Robert M.J. Jacobs, Marco Maccarini, Peter Fouquet, and Frank Schreiber. Protein diffusion in crowded electrolyte solutions. *BBA-Proteins Proteom*, 1804(1):68 – 75, 2010.

2. Felix Roosen-Runge, Marcus Hennig, Fajun Zhang, Robert M. J. Jacobs, Michael Sztucki, Helmut Schober, Tilo Seydel, and Frank Schreiber. Protein self-diffusion in crowded solutions. *Proc. Natl. Acad. Sci. USA*, 108(29):11815–11820, 2011.

3. Fajun Zhang, Roland Roth, Marcell Wolf, Felix Roosen-Runge, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Sztucki, and Frank Schreiber. Charge-controlled metastable liquid-liquid phase separation in protein solutions as a universal pathway towards crystallization. *Soft Matter*, 8:1313–1316, 2012.

4. Marco Heinen, Fabio Zanini, Felix Roosen-Runge, Diana Fedunova, Fajun Zhang, Marcus Hennig, Tilo Seydel, Ralf Schweins, Michael Sztucki, Marian Antalik, Frank Schreiber, and Gerhard Nägele. Viscosity and diffusion: crowding and salt effects in protein solutions. *Soft Matter*, 8:1404–1419, 2012.

5. Marcus Hennig, Felix Roosen-Runge, Fajun Zhang, Stefan Zorn, Maximilian W. A. Skoda, Robert M. J. Jacobs, Tilo Seydel, and Frank Schreiber. Dynamics of highly concentrated protein solutions around the denaturing transition. *Soft Matter*, 8:1628–1633, 2012.

6. Fajun Zhang, Felix Roosen-Runge, Maximilian W. A. Skoda, Robert M. J. Jacobs, Marcell Wolf, Philip Callow, Henrich Frielinghaus, Vitaliy Pipich, Sylvain Prevost, and Frank Schreiber. Hydration and interactions in protein solutions containing concentrated electrolytes studied by small-angle scattering. *Phys. Chem. Chem. Phys.*, 14:2483–2493, 2012.

7. Fajun Zhang, Felix Roosen-Runge, Andrea Sauter, Roland Roth, Maximilian W. A. Skoda, Robert Jacobs, Michael Sztucki, and Frank Schreiber. The role of cluster formation and metastable liquid-liquid phase separation in protein crystallization. *Faraday Discuss.*, 159:313–325, 2012.

8. Felix Roosen-Runge, Benjamin S. Heck, Fajun Zhang, Oliver Kohlbacher, and Frank Schreiber. Interplay of pH and binding of multivalent metal ions: Charge inversion and reentrant condensation in protein solutions. *J. Phys. Chem. B*, 117(18):5777–5787, 2013.

9. Fajun Zhang, Felix Roosen-Runge, Andrea Sauter, Marcell Wolf, Robert M. J. Jacobs, and Frank Schreiber. Reentrant condensation, liquid-liquid phase separation and crystallization in protein solutions induced by multivalent metal ions. *Pure Appl. Chem.*, accepted for publication, 2013.

10. Elena Jordan†, Felix Roosen-Runge†, Sara Leibfarth, Fajun Zhang, Michael Sztucki, Oliver Kohlbacher, Andreas Hildebrandt, and Frank Schreiber. Competing salt effects on reentrant phase behavior of protein solutions. *submitted*, 2013.

11. Felix Roosen-Runge, Fajun Zhang, Frank Schreiber, and Roland Roth. Proteins as particles with attractive patches. *submitted*, 2013.

12. Moritz Schollbach, Fajun Zhang, Felix Roosen-Runge, Maxmilian W. A. Skoda, Robert M. J. Jacobs, and Frank Schreiber. Gold nanoparticles decorated with oligo(ethylene glycol) thiols: Surface charges, and interactions with proteins in solution. *submitted*, 2013.

13. Daniel Soraruf, Felix Roosen-Runge, Marco Grimaldo, Fabio Zanini, Ralf Schweins, Tilo Seydel, Fajun Zhang, Martin Oettel, and Frank Schreiber. Protein cluster formation in aqueous solution in the presence of multivalent metal ions a light scattering study. *submitted*, 2013.

14. Felix Roosen-Runge, Baohu Wu, Marcell Wolf, Fajun Zhang, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Sztucki, and Frank Schreiber. Multivalent ions at protein surface: an ASAXS study in protein solutions with attractive interactions. *in preparation*.

15. Marcell Wolf, Felix Roosen-Runge, Fajun Zhang, Roland Roth, Maximilian W. A. Skoda, Robert M. J. Jacobs, Michael Sztucki, and Frank Schreiber. Effective interactions in protein-salt solutions approaching liquid-liquid phase separation. *in preparation*, 2013.

Acknowledgements

The scientific work included in this thesis has been enabled by a living and inspiring work environment. Numerous people have contributed to it and I am pleased to acknowledge their support.

First of all, I would like to thank my first supervisor Frank Schreiber for giving me the opportunity and freedom to work in this interesting scientific field as independently as I wished to. I benefited greatly from discussions on data interpretation, general aspects of soft condensed matter, academic life and scientific writing.

I am indebted to Fajun Zhang for his enduring support and patience as well as for the scientific inspiration and innumerable discussions. During experiments and the development of the conceptual understanding, Fajun's experience how to tackle scientific problems and his thoughtful scientific guidance for the projects made it a pleasure to work in his subgroup.

I express my deepest gratitude to my second supervisor Roland Roth for the very constructive and pleasant collaboration. In particular, I have always appreciated the calm and concentrated atmosphere of our discussions, which deepened my theoretical understanding of statistical physics of soft matter.

I would like to thank Oliver Kohlbacher for his thoughtful advice on the modeling of protein–ion systems and the general information on proteins and biochemistry that emerged during several discussions.

I am grateful to all the coworkers from the protein subgroup that accompanied my research. In particular, I would like to thank Andrea Sauter and Marcell Wolf for the smooth team work during numerous beamtimes at the FRM II as well as for the good cooperation regarding our connected research projects. I am indebted to Benjamin Heck, Elena Jordan and Sara Leibfarth, whose results from their diploma work provided a solid basis for projects in my doctoral thesis.

It is a pleasure to thank Martin Oettel for several discussions that not only clarified the specific topics, but also broadened my knowledge of general statistical physics.

I would like to thank Tilo Seydel and Marcus Hennig who have shaped my thinking about scattering, soft matter, physics in general and, in particular, dynamical features of protein solutions. Although the studies on protein dynamics did not enter the present thesis, the ongoing collaboration has been a steady source of ideas, inspiration and motivation while working on my thesis. I would also like to thank Daniel Soraruf and Marco Grimaldo

for inspiring discussions on protein dynamics.

I enjoyed the joint work with Fabio Zanini very much and I do appreciate his sharp and incorruptible thinking about scientific questions as well as his broad horizon and open mind for new ideas.

I would like to express my sincere gratitude to Marco Heinen und Gerhard Nägele for the fruitful collaboration on protein diffusion and the related insights on colloidal modeling in connection with proteins.

While working on the thesis, I participated in numerous experiments at large-scale facilities. I would like to thank the coworkers—Robert Jacobs, Maximilian Skoda, and all members of the subgroup—as well as the scientific staff at the beamlines—in particular Aurel Radulescu, Henrich Frielinghaus, Michael Sztucki, Noemi Szekely, Ralf Schweins, Theyencheri Narayanan and Vitaly Pipich—for their help in good humor and the willingness to share their knowledge on scattering techniques and data analysis.

During conferences and visits, the research topics of the thesis were discussed with many scientists. I would like to thank Paddy Royall for his interest in our work and his comments as well as Pavel Jungwirth for comments on ion effects at the protein surface.

I express my deepest appreciation to Oliver Lenz and Christian Holm for interesting and helpful discussions on salt effects in colloid solutions, in particular in relation to simulations.

I would like to thank Edward Goldobin, Katharina Broch, Oliver Eibl, Reinhold Kleiner and Frank Schreiber for the good team work during the tutorials accompanying the lecture "Condensed Matter".

Many thanks to present and former members of the work group in Tübingen for enjoyable conversations and the friendly atmosphere, in particular during the group excursions to Oberjoch and traveling to conferences. Alexander Gerlach has my special admiration for his steady and excellent organization and unpretentious help within the work group.

I am happy to acknowledge the prompt, concise and very helpful proofreading by Andrea Sauter, Elena Jordan, Marcell Wolf, Marco Grimaldo, Michal Braun, Olga Matsarskaia, Saliba Barsaume and Tilo Seydel.

I acknowledge a doctoral scholarship of the Studienstiftung des deutschen Volkes.

Finally, numerous influences from the outside of the academic world of physics supported and enabled this thesis in such important ways that few lines cannot do justice to my parents and siblings, my friends and, in particular, my wife. I am very grateful for the time spent together, and I am looking forward to more.