LUND UNIVERSITY

Attractive Patchy Protein Interactions

Understanding of protein dimerization

Li, Weimin

2016

Link to publication

Citation for published version (APA):
Li, W. (2016). *Attractive Patchy Protein Interactions: Understanding of protein dimerization*. Division of Physical Chemistry, Faculty of Science, Lund University.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Attractive Patchy Protein Interactions

Understanding of Protein Dimerization

by Weimin Li

LUND
UNIVERSITY

Thesis for the Degree of Doctoral of Philosophy in Physical Chemistry

Faculty opponent: Abraham M. Lenhoff

To be presented, with the permission of the Faculty of Science of Lund University, for public criticism in lecture hall B at the Division of Physical Chemistry on Friday, the 16th of September 2016 at 09:30.

Organization
LUND UNIVERSITY

Division of Physical Chemistry
P.O. Box 124, SE–221 00, Lund
Sweden

Document name
DOCTORAL DISSERTATION

Date of disputation
2016-09-16

Sponsoring organization OMM Linneaus center; Swedish Research Council; Crafoord Foundation; Royal Swedish Academy of Science

Author(s)
Weimin Li

Title and subtitle
Attractive Patchy Protein Interactions: Understanding of Protein Dimerization

Abstract
Protein interactions are important in understanding various protein relevant cellular and technological processes. This is a challenging task due to the intrinsically irregular shape and inhomogeneous surface distribution of neutral, charged, hydrophobic amino acid residues. This results in more or less anisotropic interactions depending on external parameters such as salt and pH. However, a deeper understanding of these complex interactions is essential in order to understand the behavior of proteins in solution and protein phase diagrams. The aim of this work was to gain a deeper understanding of protein attractions, in particular, anisotropic protein attractions. We have for this reason studied the protein lactoferrin. We have found and characterized a highly directional attraction which combined with Monte Carlo (MC) simulations show the mechanism as a highly directional short-ranged electrostatic attraction, originated from few ionizable amino acids. This gives rise to a non-monotonic dependence of the second virial coefficient, B_2, with ionic strength both in scattering experiments and MC simulations, in qualitative agreement. Further, we show how the directional attraction is effectively behaving as an attractive patch that leads to dimerization of lactoferrin under patch attractive conditions. The phase diagram is also determined under patch conditions where new phases are discovered using cryo-transmission electron microscope. Finally, we explored protein capacitance obtained both from titration experiments and MC computer calculations which is related to charge regulation attraction.

Key words
Patchy attractions, protein-protein interactions, second virial coefficients

Classification system and/or index terms (if any)

Supplementary bibliographical information

Language
English

ISSN and key title

ISBN
978-91-7422-466-5 (print)

Recipient's notes

Number of pages
106

Price

Security classification

I, the undersigned, being the copyright owner of the abstract of the above-mentioned dissertation, hereby grant to all reference sources the permission to publish and disseminate the abstract of the above-mentioned dissertation.

Signature _____ Date _____

DOKUMENTDATABLAD enl SIS 61 41 21

Attractive Patchy Protein Interactions

Understanding of Protein Dimerization

by Weimin Li

LUND
UNIVERSITY

Cover illustration front: Taiji, one ultimate that is composed of yin and yang and reconciled by yin's yang and yang's yin. Lactoferrin dimer, one unit that is composed of two monomers through complementary patchy interaction.

Cover illustration back: To my family, my wife, my relatives and my friends.

Funding information: The OMM Linneaus center in Lund, the Swedish Research Council, the Crafoord Foundation, and the Royal Swedish Academy of Science

© Weimin Li 2016

Faculty of Science, Division of Physical Chemistry

ISBN: 978-91-7422-466-5 (print)

Printed in Sweden by Media-Tryck, Lund University, Lund 2016

"Life is like a box of chocolates. You never know what you're gonna get",
but "All is well".

From
— Forrest Gump (R. Zemeckis, W. Finerman, S. Tisch and S. Starkey)
— 3 Idiots (R. Hirani and V.V. Chopra)

Contents

Popular Science Summary iii

1 Introduction **1**

2 Materials **5**
2.1 Buffer preparation 5
2.2 Purify and concentrate the protein solution 6
2.3 Protein characterization 7

3 Theory and Methods **13**
3.1 Small-angle scattering 13
3.2 Second virial coefficient 16
3.3 Dynamic light scattering 19
3.4 Integral equation theory 21
3.5 Monte Carlo simulations 24
3.6 Protein titration 25
3.7 Cryogenic transmission electron microscope 26
3.8 Confocal laser scanning microscopy 26

4 Results **27**
4.1 Charge-induced patchy attractions between proteins 27
4.2 Concentration-induced protein association from a directional patch attraction 32
4.3 Solution structures formed by proteins with patch-patch attractions 38
4.4 Protein charge capacitance obtained from experiments and MC simulations 42

5 Conclusions and Future Prospects **47**

References **49**

Acknowledgements **55**

Papers **57**
List of Papers 57
My contribution to the papers 59

Popular Science Summary

If DNA is the genetic code, proteins are the resulting products from this code. They are all unique with a certain structure which is intimately connected to their biological function. This is encrypted in the sequence. The sequence is the ordering of how individual building blocks, the amino acids, are arranged. This ends up in a complex, mosaic surface with regions of different kinds. These regions can be either attractive, repulsive or neutral giving rise to "finger-print"-like identity of each protein. The forces which are operating between proteins originates from the "finger-print"-like surface distribution of the amino acids. This is highly important to understand since most of protein functions, solution behavior such as aggregation, self-association and crystallization, are direct consequences of the forces.

In this work, we have investigated and determined what is the hidden mechanism behind intrinsic protein attractions, i.e. attractive forces which acts to push proteins closer, such as in formation of a protein complex or in an aggregates found in neurodegenerative diseases.

In particular we wanted to understand asymmetric, patchy, attractions where the forces are unevenly distributed. We have found, from studies of a model protein, lactoferrin, a new mechanism behind such directional patchy interactions. We have nice agreement between computer model calculations and experiments showing how the protein in a delicate balance uses its amino acids to exactly regulate the patch which also determines the solution behavior and new intriguing structures have been found.

Chapter 1

Introduction

Protein interactions

Proteins are large biomolecules of fundamental importance involved in the biologic machinery in all living organisms. Their function is broad including the construction of cytoskeleton[1], catalyzing metabolic reactions[2], conveying signals between cells or transferring molecules[3], to mention a few. The differences in activities of proteins are primarily originating from their constituents, which are amino acids. Proteins are built of amino acids that are linked through peptide bonds between amine and carboxylic groups forming the backbone chain. The side chains or residues of these natural amino acids are unique and used in categorizing the amino acids. The side chains are either hydrophilic which includes chargeable and neutral residues, or hydrophobic. Proteins fold into specific tertiary structures determined from the sequence of amino acids, and they function by interacting in a complex fashion with other proteins or bio-molecules.

Repulsive interactions between proteins are essential to maintain stability of protein solutions, for example in food applications where emulsions are stabilized by adding polysaccharide chains[4]. Attractive interactions between proteins are also important, one example is in protein crystallization because in a protein crystal monomers are arranged in a tight, periodic fashion. Directional attractions are also behind the formation of large multi-protein complexes , such as the major histocompatibility complex, MHC-I, which are responsible for the specific immune system[5]. The ability of proteins to target specific chemicals due to attractive forces is widely used in formulation of drugs[6]. However, it is a challenging task to understand the mechanisms behind protein interactions[7–9]. This is further complicated by the fact that protein interactions can be sensitive on the level of specific amino acids[10]. The level

of complexity originates from each protein having a unique tertiary structure accompanied by an irregular distribution of different types of amino acids which can be charged, neutral or hydrophobic depending on the solution conditions. The irregular surface distribution of amino acids results in an interaction landscape consisting of charged, uncharged and hydrophobic regions or patches. These mosaic patterns of amino acid residues together with the protein shape contribute in a complex way to the overall protein interactions.

Understanding of interactions in general is crucial since they are responsible for the overall phase behavior of a system including phenomena such as crystallization, aggregation, and phase separation[7,11,12].

Exactly how the regions or patches correlate with simpler colloidal models of spheres decorated with attractive regions is hard to see. Again, in the case of protein crystallization, previous studies showed that the nature of contact amino acids is a key factor in growth of protein crystals[13] and that the protein interactions involved in crystallization should be described as patchy contact sites[12].

In the past, most of the understanding on protein phase or solution behavior comes from isotropic interaction models such as the hard sphere potential, used to predict crystal-nucleation rates[14] and the adhesive hard sphere (AHS) potential from Baxter[15]. These are colloidal models which have been successfully employed in the understanding of phase diagrams of some globular proteins[16–18].

Introducing patchy attractions is expected to give rise to a very different phase or solution behavior compared to isotropically interacting colloids. They predict a vast array of new phases depending on the number of patch and size[19–23]. Attempts have been made to describe protein interactions with an anisotropic attraction models[24]. However, how to go from a protein to a more coarse-grained representation is non-trivial and a framework is lacking.

Furthermore, dissecting the overall interactions to grasp the main contributor is difficult since the intermolecular interactions are combined in a complex manner. One experimentally accessible quantity probing the molecular interactions is the osmotic second virial coefficient, B_2 [25–27]. To distinguish each contribution to the overall interactions is not feasible from singular B_2. This however, do not prevent it from being an important quantity capable of discriminating between overall attractions and overall repulsions. It can be obtained from static light scattering experiments[28], and from measurements of the osmotic pressure of the system[29]. Combining experiments, colloidal theories and computer simulations together is needed to understand protein interactions and predict protein phase diagrams [30–33].

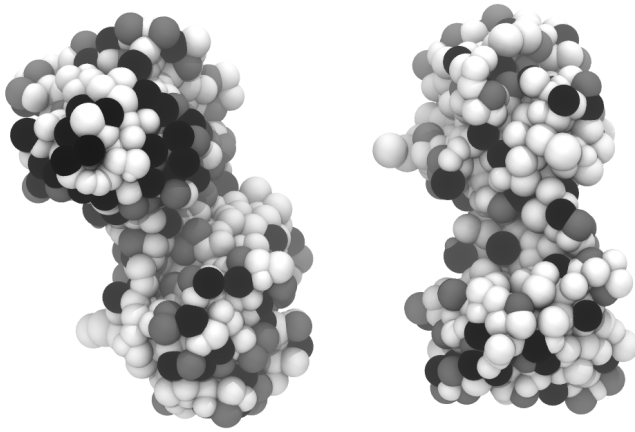


Figure 1.1: Monte Carlo model of lactoferrin where amino acids are represented as a collection of beads. Neutral amino acids are depicted as white spheres, cationic as black spheres and anionic as grey spheres. The left protein has the deeper groove pointing towards us and the right protein shows the other side.

Model protein lactoferrin

In this study, lactoferrin was chosen as a model system based on previous Monte-Carlo (MC) simulations which had discovered a highly directional attraction between lactoferrin molecules[34]. It is a water-soluble globular glycoprotein with a molecular weight of 80 kDa that is abundant in various secretory fluids such as tears, saliva, and in milk with excellent antimicrobial properties[35]. Colostrum, the first milk given to newborn babies of mammals, contains a particularly high concentration of lactoferrin serving as the front line of defense for the newborn. It is also present in neutrophils, the white blood cells involved in the innate immune system[36]. Human colostrum contains the highest lactoferrin concentration in mammals followed by cow milk[37], which motivates studying lactoferrin from a biological side.

We have chosen to work with bovine holo-lactoferrin whose structure is more conserved than apo-lactoferrin[38]. The resolved crystal structure of lactoferrin presents a peanut shape composed of two domains with the long axis around 9 nm and short axis of around 4 nm, molecular weight of 80 kDa, as figure 1.1 shows[39].

Outline and purpose of thesis

The overall goal of this thesis is to improve our understanding of protein attractions based on the model protein lactoferrin. Here experimental observations are combined with Monte Carlo simulations and integral equation theory. I start with a description of the preparation procedures of the protein, given in more detail than are found in the papers. Next, I describe the techniques which are used in the experiments and related theory.

One particular aim of the thesis was to elucidate which mechanism dominates protein attractions. Guided by Monte Carlo simulations of lactoferrin a directional short-range attraction was investigated in paper I[34]. Experimentally the attraction was shown to be a highly directional short-range electrostatic attraction manifested as a nonmonotonic behavior of B_2 as a function of ionic strength. Similar nonmonotonic behavior of B_2 has also been observed in other protein systems pointing to some level of generality[8,33].

Further we show in paper II how the isothermal compressibility, $S(0)$ measured by SAXS and SLS, displays a maximum as a function of volume fraction when the patch attraction is operating. Intentionally, we wrongfully apply a conventional AHS model to explain how the behavior is signaled as an unphysical adaption of the attraction strength as a function of volume fraction. Instead the Wertheim integral equation theory[40–42] provides a dimerization mechanism explaining the behavior of $S(0)$ without any unphysical adjustment.

In paper III we investigate the phase diagram under patchy conditions in terms of protein concentration. Patchy short-range attractions in spherical models produce rich phase diagrams which is also found in the lactoferrin system when the patch directional attraction is active. A new stripe-like phase was investigated and imaged using mainly cryo-TEM.

At the end we turn our attention to titration experiments and charge capacitance, which are related to the charge regulation mechanism[43]. Charge regulation is directly linked to the protein's charge capacitance and in paper IV we determined capacitance from experiments and compared with simulations on lactoferrin.

Chapter 2

Materials

Bovine lactoferrin was purchased in powder from Morinaga Milk Industry Co. Ltd., Japan with a purity above 96%. A systematic procedure of dissolving, purifying, and changing the buffers was established. A detailed description is given here on the purification and stabilization intended to serve as a future reference.

We use exclusively monovalent buffers to modulate the buffer pH. In the acidic pH range from pH 5-7, NaOAc was used. At a higher range from pH 7-9, Tris was chosen. The contribution of ionic strength from buffer is taken into account by using the Henderson-Hasselbalch equation. The buffer chemicals used in the experiments were purchased from Sigma company in powder form. Prepared buffer solutions were filtered with 0.45 μm and stored in a refrigerator at 4 °C prior to use.

2.1 Buffer preparation

Buffer solution was composed of 3 mM target buffer (Tris or NaOAc in most cases in our study) which was the smallest amount of buffer needed to provide stable pH values which ensured that the ionic strength was mainly set by added salt and not the buffer components. Furthermore, 1 mM NaN_3 was added to prevent the growth of bacteria. This recipe was applied to prepare all buffers. The target pH was adjusted by adding HCl or NaOH. The contribution from each species in the buffer solution to the total ionic strength, [I], was taken into account according to the equation 2.1

$$I = \frac{1}{2}\sum_{i=1}^{n} k_i z_i^2, \tag{2.1}$$

where i is the index of ion species, n is the total number of ion species, k_i is the molar concentration of i-th species, z_i is the number of charges (valency) of i-th species. The desired ionic strength is achieved by adding NaCl. Adjusting the pH by adding acid or base also contributes to the total ionic strength which thus restricts the available pH range at low ionic strength, for example, adding HCl to reach a low pH may exceed the desired ionic strength because of adding Cl^-.

The contribution to the total ionic strength from buffer components is taken into account by using the acidic dissociation constant and calculating the concentration of the ionized buffer or conjugate base as following

$$HA + H_2O \rightleftharpoons A^- + H_3O^+ \tag{2.2}$$

where HA is the weak acid and A^- is the conjugate base which contributes to the total ionic strength. The way of calculating the concentration of A^- is described by the Henderson-Hasselbalch equation

$$pH = pK_a + log_{10}\left(\frac{[A^-]}{[HA]}\right) \tag{2.3}$$

pK_a is the acidic dissociation constant of the corresponding weak acid, $[A^-]$ and [HA] are the concentration of the conjugate base and the weak acid. In order to calculate $[A^-]$, equation 2.3 gives

$$[A^-] = 10^{pH-pK_a}[HA] \tag{2.4}$$

2.2 Purify and concentrate the protein solution

Powder lactoferrin was dissolved in NaOAc buffer at pH 5.5, ionic strength 5 mM, with the initial concentration of 2 mg/mL. The protein solution was equilibrated for at least 48 hours to have an adequate amount of monomers dissolved. This original protein solution is referred to as the mother stock solution. When the mother stock solution is ready, the first step is to filter it by using AMICON centrifugal filters with a cutoff of 100 kDa in a centrifuge (4000 rpm, 15 minutes). The following step is to immediately transfer the filtrates out from 100 kDa filter tubes into 50 kDa filter tubes (4000 rpm, 15 minutes) in order to concentrate the purified protein solution.

Buffer exchange was undertaken when needed by repeatedly adding target buffer in 50 kDa filters until the filtrates reached the same pH as the target buffer. An example is shown in figure 2.1. The purity of the final protein solution was examined by dynamic light scattering (DLS). Samples were stored in 4 °C for further research work.

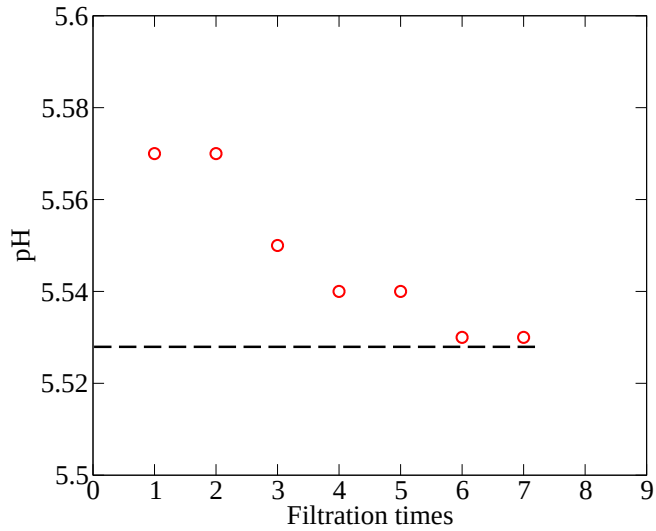


Figure 2.1: pH of filtrates after each washing step, red circle. Dash line is the pH of washing buffer.

2.3 Protein characterization

In order to know the protein concentration accurately, the extinction coefficient of lactoferrin was measured on a spectrophotometer. Also, the protein density was measured on a densitometer in order to convert mass concentration to volume fraction. The stability of lactoferrin in terms of structural conformation was characterized by using circular dichroism. The stability of protein solution, i.e. aggregation or degradation over time, was inspected by both circular dichroism and DLS. The aim of testing the stability was to locate the pH where protein starts to denature.

2.3.1 UV-Vis spectrophotometer

The extinction coefficient of a protein, ε, measured on a UV-Vis spectrophotometer satisfies a linear relation between protein concentration and the amount of light absorbed according to the Beer-Lambert's law

$$A = \varepsilon c l \tag{2.5}$$

where A is the absorbance, c is the protein concentration which depends on the unit of ε, and l is the beam path-length which is 1 cm in our work.

The protein concentration is typically calculated by measuring absorbance A at

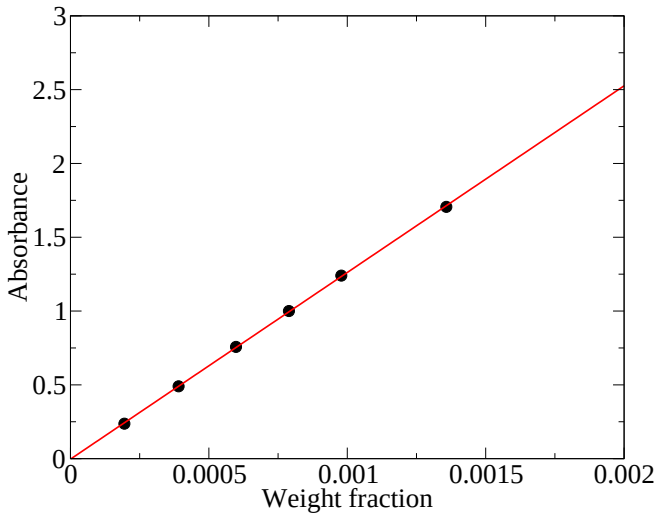


Figure 2.2: Abbsorbance, at wavelength 280 nm, as a function of protein concentration. Symbols are date, line is linear fit. Samples are in buffer NaOAc, pH 5.5 and ionic strength 5 mM .

wavelength of 280 nm with a known ε. Thus, the accuracy of determination of protein concentration depends on the accuracy of extinction coefficient.

We performed a careful dry-mass concentration determination on a large volume of protein sample to obtain the initial protein concentration accurately in weight fraction. A concentration series was prepared from this sample based on mass ratio. By measuring the absorbance on the series, the extinction coefficient was extracted from the slope of the absorbance as a function of protein concentration as it is shown in figure 2.2. In this figure, the linear fitting of the data gives an extinction coefficient of 1264.5 in the unit of cm^{-1} on the basis of monomeric weight fraction.

The instrument used was a Cary 300 Bio spectrophotometer, from Agilent Technologies, Inc. An external thermocontrol was implemented, and all the experiments were performed at 25 °C. Before the measurements start, spectrophotometer is switched on for one hour to stabilize the light emission. Sample is measured in a quartz cuvette with a light pass length of 1 cm. The ultra-violet (UV) light is selected to start from 340 nm and end at 200 nm. The detector of this instrument will be saturated when the absorbance is above 2.5 and the lower limit is at 0.1.

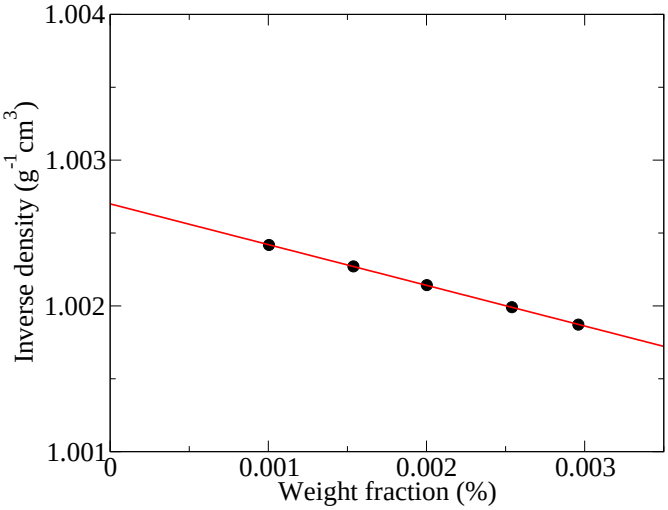


Figure 2.3: The inverse total density plotted as a function of protein concentration in weight fraction (black dots) with linear fit (red line). Samples are in NaOAc buffer, pH 5.5 and ionic strength 5 mM.

2.3.2 Protein density

Protein density is needed in order to convert the protein concentration in weight faction to weight per volume, or volume faction. The Density-meter used to measure the density of the protein solution was a DMA 5000 from Anton Paar. The temperature was kept at 25 °C. The following equation was utilized to extract the protein density from the density of protein solution

$$\rho_{tot}^{-1} = (\rho_{pro}^{-1} - \rho_{sol}^{-1}) \cdot c + \rho_{sol}^{-1}, \tag{2.6}$$

where ρ_{tot}^{-1}, ρ_{pro}^{-1} and ρ_{sol}^{-1} are the inverse density of the protein solution, the inverse protein density and the inverse solvent density, respectively, and c is the protein concentration given in weight fraction. Equation 2.6 contains the protein density in the slope of ρ_{tot}^{-1} as a function of c as it is shown in figure 2.3.2. The protein density of lactoferrin obtained from our experiment was 1.3824 g/cm^3.

2.3.3 Stability of protein conformation

Proteins have their native conformation within a certain pH range[44,45]. The stability of lactoferrin with respect to denaturation or partial unfolding was investigated

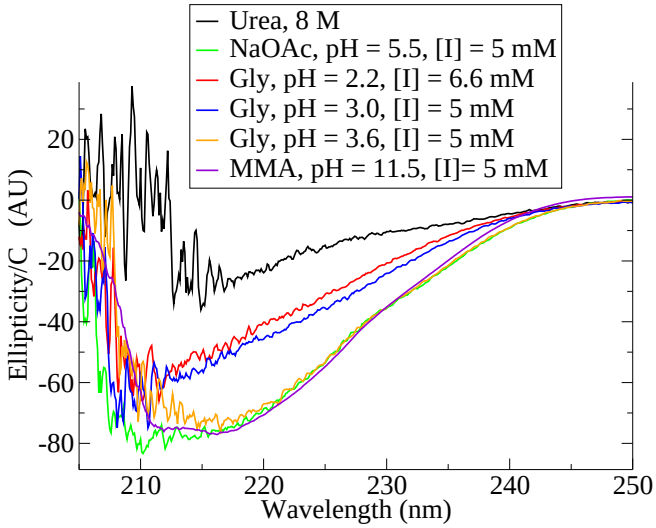


Figure 2.4: Circular dichroism spectra normalized with concentration. Conditions are shown in the legend.

using circular dichroism (CD). Lactoferrin is stable at various conditions *in vivo* with an isoelectric point of about 8[37], therefore the lactoferrin denaturation point is expected at extreme pH. Glycine (Gly) buffer was prepared for pH below 4 and methylamine (MMA) buffer for pH above 10. Protein concentration was around 1.0 mg/mL giving good signal to noise ratio. All samples were measured in a 0.1 cm quartz cuvette at 25 °C.

Two positive control experiments were included for comparison. One is lactoferrin in NaOAc solution at pH 5.5 and ionic strength 5 mM, where the protein structure is fully intact. The other is lactoferrin in 8 M of urea, where the protein is fully untangled. A pH point where protein starts denaturing should give a spectrum in between the spectra of the two controls as it is shown in figure 2.4.

The CD spectra of protein solutions at pH 11.5 buffered by methylamine and at pH 3.6 buffered by glycine do not differ from the spectra of protein solution buffered by NaOAc, the control experiment. This was interpreted as that in the pH range from 3.6 to 11.5, the protein keeps its native conformation. The spectra at pH 2.2 and pH 3.0 are referring to a certain degree of protein conformational loss because they are close to the control of 8 M urea. Overall this experiment suggests that the denaturation process of lactoferrin starts between pH 3.0 and pH 3.6.

Regarding the stability of protein conformation, it should be noted that it can be permanently altered by long time X-ray exposure which imposes a physical dam-

age on the protein interior structure. Although this issue is commonly known in literature it is often ignored.

2.3.4 Stability of protein solution

We examined the protein solution stability over time searching for signs of any possible aggregation or degradation process. DLS provides the hydrodynamic radius of proteins (section 3.3), which is used to monitor the changes of apparent protein size. Hydrodynamic radius does not necessarily represent the true size of a protein but well servers the purpose of indicating any size changes because of aggregation or degradation over time.

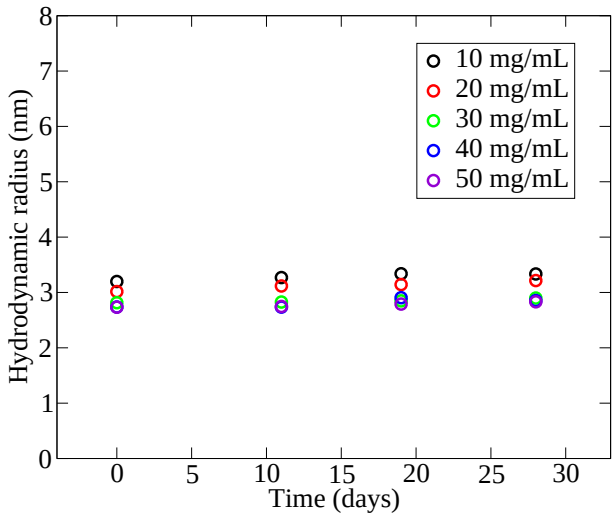


Figure 2.5: Hydrodynamic radius at five protein concentrations as a function of time in days. The protein solutions are in NaOAc buffer with a pH of 5.5 and ionic strength of 5 mM.

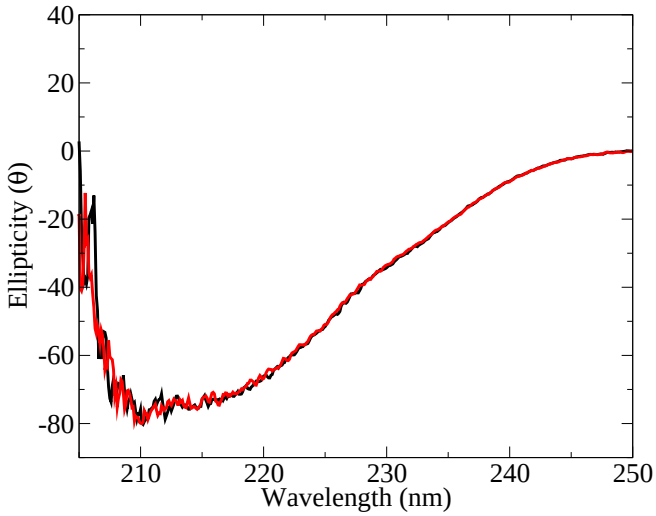


Figure 2.6: Circular dichroism measurements performed on the same sample twice (red line and black line) with a time gap of a month. Protein is in NaOAc buffer, pH 5.5 and ionic strength 5 mM.

The protein solutions scrutinized by DLS on a time scale of days as it is shown in figure 2.5 demonstrate a constant hydrodynamic radius around 3 nm. This shows that within the time scale of one month, protein solutions in NaOAc buffer with a pH of 5.5 and ionic strength of 5 mM remain stable. This is satisfactory for long term experiments.

Another indicator of protein solution being stable over time is the CD measurements. It is expected that aggregation or degradation modifies protein conformation which in turn changes the CD spectra. As the figure 2.6 shows, spectra recorded with a time difference of a month on the same sample are close to identical proving that the protein has a conserved structure during this time in NaOAc buffer at pH of 5.5 and ionic strength of 5 mM.

The circular dichroism (CD) instrument is type PTC-348WI from JASCO Corporation. β-sheet, α-helix and random coil structure have distinct absorption patterns in the wavelength region between 190 nm and 250 nm, which is used to quantify the abundance of the proteins having the above mentioned secondary structures. Peltier cell holder which allows for temperature control of 25 °C.

Chapter 3

Theory and Methods

This chapter introduces a number of techniques that have been used throughout this work. It briefly describes the integral equation theory used, and gives an introduction to Monte Carlo (MC) simulations.

3.1 Small-angle scattering

In small-angle light and X-ray scattering experiments the scattered intensity, $I(q)$, is measured as a function of scattering vector q, defined as

$$q = \frac{4\pi n}{\lambda} \sin\frac{\theta}{2}, \tag{3.1}$$

where λ is the beam wavelength, n the index of refraction in light scattering and θ is the scattering angle. A suspension composed of monodisperse particles has the scattered intensity expressed as

$$I(q) = nP(q)S(q), \tag{3.2}$$

here n is the number density of the particles defined as number of particles per volume, N/V, $P(q)$ is the form factor and $S(q)$ is the structure factor.

The form factor relates to the square of the form amplitude $|f(q)|^2$ which reads for a spherical particle with a radius of a as

$$f(q) = 4\pi \int_0^a r^2(\varrho(r) - \varrho_{solv}) \frac{\sin qr}{qr}\, \mathrm{d}r, \tag{3.3}$$

where ϱ stands for scattering length density, $\varrho(r)$ refers to particle and ϱ_{solv} refers to solvent. The ability of a particle to scatter radiation is determined by its scattering length density which is in turn determined by its internal structure. Hence $P(q)$ is determined by the shape, size and internal density profile of the particle. In X-ray scattering, due to the short wavelength, $P(q)$ resolves the conformation of a nano-scaled object.

The structure factor $S(q)$ is the Fourier transform related to the radial distribution function $g(r)$ [46] as

$$S(q) = 1 + 4\pi\frac{N}{V}\int_0^\infty [g(r) - 1]r^2\frac{\sin qr}{qr}\,\mathrm{d}r \tag{3.4}$$

The radial distribution function $g(r)$ defines how the number density, $\frac{N}{V}g(r)$, at distance r is influenced by a particle at $r = 0$. Thus $S(q)$ contains the information about interactions and correlations between particles.

$P(q)$ is determined experimentally in a dilute condition where particles are not correlated, $I(q) \approx nP(q)$. $S(q)$ is obtained by dividing the scattered intensity measured on a concentrated system by scattered intensity of a dilute system as

$$S(q) = \frac{I(q)_{conc}}{I(q)_{dilu}}\frac{N_{dilu}}{N_{conc}}, \tag{3.5}$$

where $I(q)$ is the measured intensity, N_{dilu}/N_{conc} is the protein concentration ratio. The limit of $S(q)$, $q \to 0$, probes the whole ensemble and is related to the isothermal compressibility. $S(0)$ grows when the system is attractive indicating an increased compressibility. For isotropic repulsions, $S(0)$ is monotonically decreased indicating a decreased compressibility.

Such a valuable information as the molecular weight is contained in the absolute scattered intensity. For light scattering it is expressed via the Rayleigh ratio. The method of calculating the absolute intensity is the same for X-ray scattering and light scattering, which is

$$I_{par}^{abs} = (I_{sam} - I_{bac})\frac{I_{ref}}{I_{ref}^{abs}} \tag{3.6}$$

All measured intensities are transmission corrected. In order to obtain the absolute intensity of the particles, the measured intensity of a sample with the contribution from background subtracted is corrected by an absolute scaling factor which is calculated by dividing the measured intensity of a reference by the known absolute

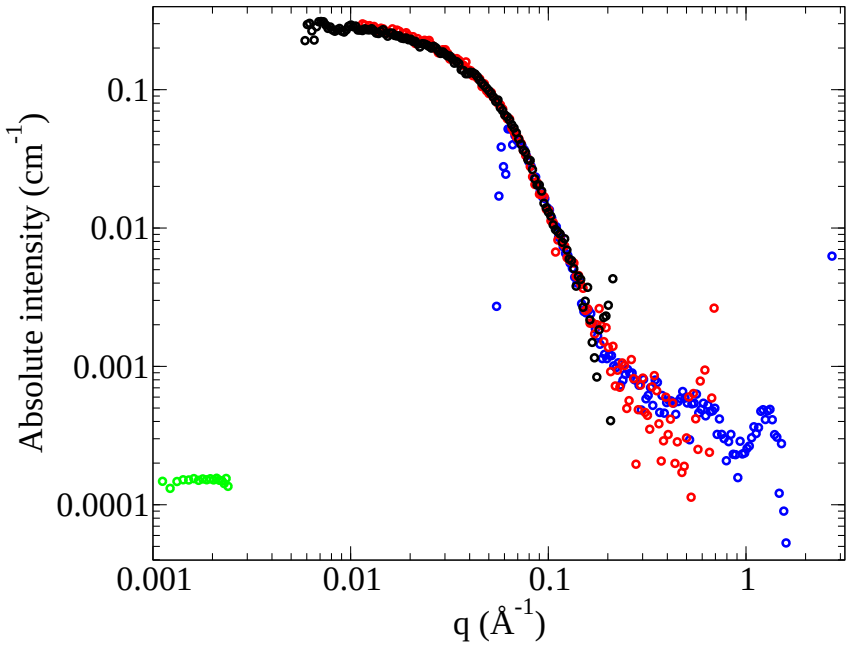


Figure 3.1: The absolute scattered intensity of lactoferrin as a function of q. Black, red and blue circles correspond to SAXS measurements. Green circles correspond to SLS measurements. Samples are in Tris buffer, ionic strength 35 mM and pH 7.

intensity of the reference. For SLS and SAXS, toluene and water are the commonly used references, respectively.

Absolute scaling of SLS and SAXS intensities does not imply that they have similar order of magnitude as shown in figure 3.1. The scattering ability depends on the scattering length density of the particles, which is physically different between light scattering and X-ray scattering experiments. In the light scattering experiments, the light propagation is altered by the refractive index while in the X-ray scattering experiments, electron density determines the scattering ability[47].

SAXS was measured on a Ganesha 300XL (JJ X-ray Systems). Temperate was controlled at 25 °C. SLS was performed on a ALV with ALV/CGS-8F goniometer, ALV-5000 correlator, from Germany (ALV GmbH). Temperature was controlled to have 25 °C for both SLS and SAXS.

3.2 Second virial coefficient

A small-q virial expansion of $S(q)$, employing the power series of $\sin qr/qr$ in equation 3.4, results in

$$S(q) = 1 - 2\frac{N}{V}B_2 + \frac{2}{3}\pi\frac{N}{V}q^2C + \cdots \tag{3.7}$$

when it is combined with an expansion in number density. Here B_2 is a widely used property for assessing particle interactions, termed the second virial coefficient. It is defined as

$$B_2 = 2\pi\int_0^\infty (1 - e^{-u(r)/(k_BT)})r^2\mathrm{d}r, \tag{3.8}$$

where N and V are the number of particles and the volume of the particle, $u(r)$ is a pairwise potential, k_BT is the thermal energy at temperature T, r is the particles center-to-center distance. B_2 indicates attraction when it is negative and repulsion when it is positive.

The second virial coefficient, B_2, is experimentally accessible from SLS by constructing a Zimm-plot using the following relation [48]

$$\frac{Kc}{R_\theta} = \frac{1}{M_w} + \frac{2N_AB_2}{M_w^2}c, \tag{3.9}$$

where K is the optical constant, c is the concentration of the particles, R_θ is the Rayleigh ratio, N_A is Avogadro's number and M_w is the molecular weight.

The optical constant K is defined as

$$K = \frac{4\pi^2n_0^2(\mathrm{d}n/\mathrm{d}c)^2}{N_A\lambda^4}, \tag{3.10}$$

where n_0 is the refractive index of the solvent, $\mathrm{d}n/\mathrm{d}c$ is the increment of refractive index of the protein solution as a function of protein concentration determined by using refractometer (Abbe 60/ED). The linear dependence corresponding to 3.10 is shown in figure 3.2. The measured $\mathrm{d}n/\mathrm{d}c$ gives a value of $1.73\times10^{-4}\mathrm{m}^3\mathrm{kg}^{-1}$, which agrees with literature [49].

R_θ is the Rayleigh ratio at the angle θ. Lactoferrin system has a weakly pronounced q-dependence on the conventional Zimm-plot due to the fact that the scattered intensity does not depend on the scattering angle as it is shown in figure 3.3, which is also referred to as the Debye plot.

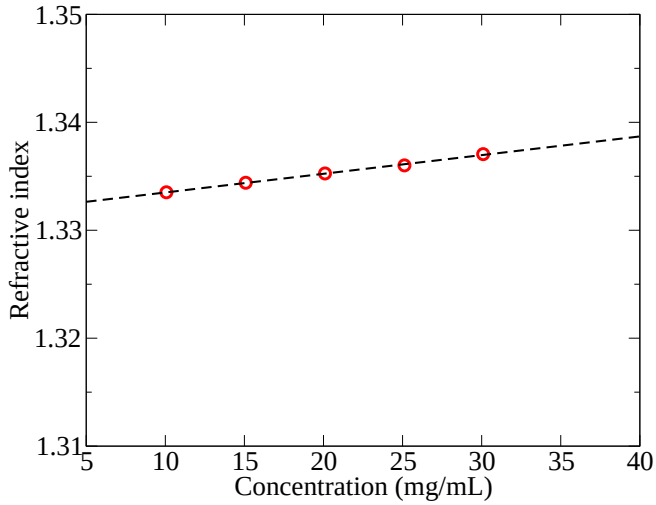


Figure 3.2: Refractive index as a function of protein concentration. Sample are in Tris buffer, pH 7 and ionic strength 35 mM.

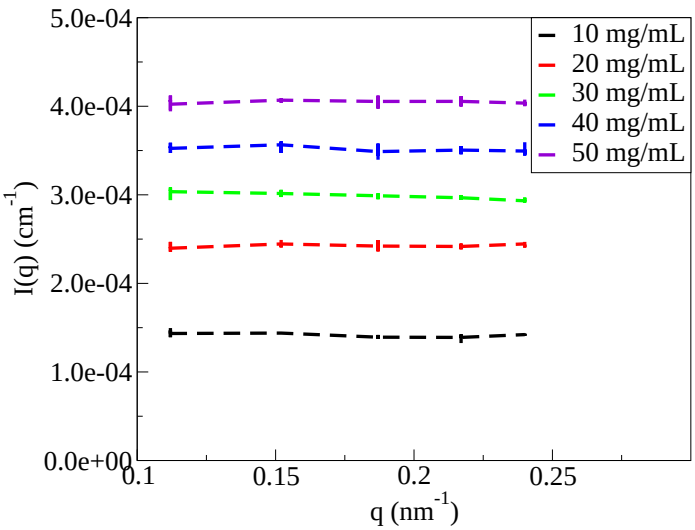


Figure 3.3: Scattered intensity of five concentrations as a function of q, as labeled. Error bars are shown as vertical lines. Samples are in NaOAc buffer, pH 5.5 and ionic strength 5 mM.

An example of applying equation 3.9 for calculating molecular weight is shown in

figure 3.8 by plotting $Kc/R_{\theta=90^\circ}$ as a function of c. The molecular weight M_w can be obtained from the intercept at zero concentration, second virial coefficient B_2 can be extracted from the slope.

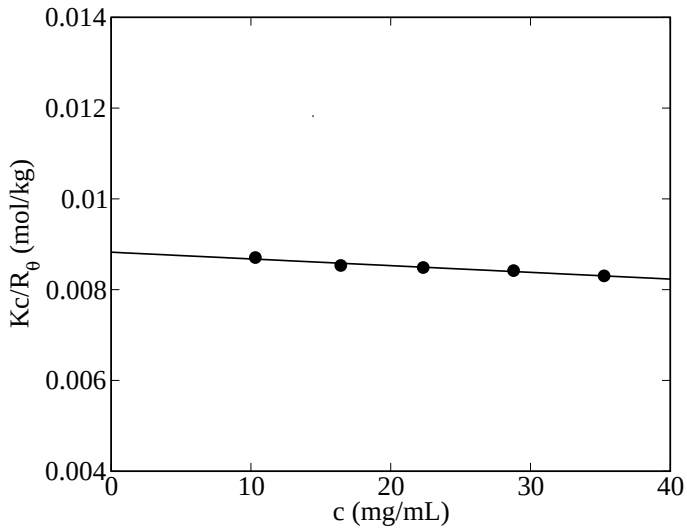


Figure 3.4: Kc/R_{90} as a function of c. Line is the linear fit. Sample is in Tris buffer, pH 7 and ionic strength 35 mM.

The necessary conditions of applying the equation 3.9 are 1) q is in the Guinier range, and 2) low enough particle concentration so that pairwise interactions dominate in the solution[48].

3.3 Dynamic light scattering

Dynamic light scattering (DLS) records how the scattered intensity $I(q,\tau)$ evolves as a function of time by correlating the scattered intensity at time 0 with the scattered intensity at time τ[47], which is defined as the intensity auto-correlation function (ACF)

$$g^{(2)}(q,\tau) \equiv \frac{\langle I(q,0)I(q,\tau)\rangle}{\langle I(q)\rangle^2}, \tag{3.11}$$

where $\langle I(q,0)I(q,\tau)\rangle$ is the ACF defined as

$$\lim_{\mathcal{T}\to\infty} \frac{1}{\mathcal{T}} \int_0^{\mathcal{T}} I(q,t)\, I(q,t+\tau)\mathrm{d}t \tag{3.12}$$

This implies the ACF contains information about how much the scattered intensity at time t correlates with the scattered intensity at time $t+\tau$, where τ is a delay time.

The intensity ACF is related to the field ACF $g^{(1)}(q,\tau)$ through Siegert relation[50]

$$g^{(2)}(q,\tau)1 + \beta|g^{(1)}(q,\tau)|^2, \tag{3.13}$$

where $\beta \leq 1$ is the deviation from ideal correlation function depending on the geometry of instrument. The information about Brownian motion which introduces fluctuations of particles' positions known as diffusion is contained in $g^{(1)}(q,\tau)$ when it is evaluated during a short time period under the condition that the system represents monodisperse spherical particles

$$g^{(1)}(q,\tau) = e^{-q^2 D(q)\tau}, \tag{3.14}$$

where $D(q)$ is the diffusion constant. In the limit of infinite dilute concentration, D_0 is applied to calculate hydrodynamic radius via the Stokes-Einstein equation

$$D_0 = \frac{k_B T}{6\pi\eta r_h}, \tag{3.15}$$

where k_B is the Boltzmann constant, T is the temperature, η is the viscosity of the solvent and r_h is the hydrodynamic radius which is calculated from measuring D_0 and η.

Figure 3.5 shows the raw scattered intensity $I(q,\tau)$ in the upper panel. $I(q,\tau)$ fluctuates around its mean value $< I(q,\tau) >$. The bottom panel shows that $I(t)$ loses its correlation with $I(t+\tau)$ with time. The correlation function starts at $< I^2 >$ and decays to $< I >^2$.

DLS is performed on the same instrument ALV as SLS experiments. Temperature was controlled to be at 25 °C.

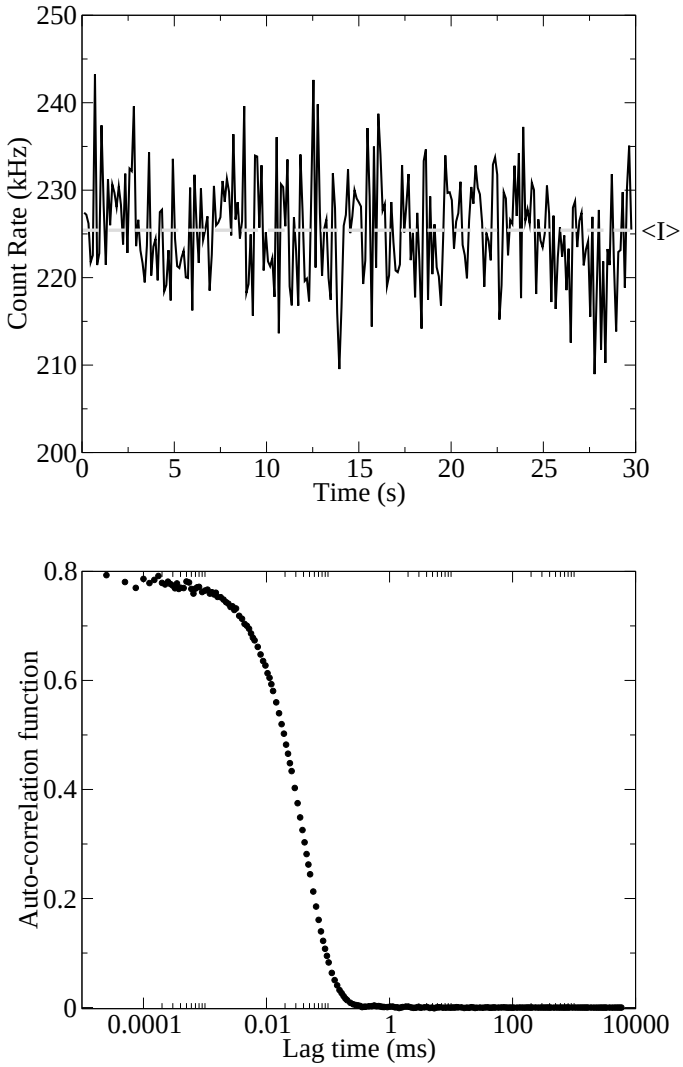


Figure 3.5: Upper panel represents scattering intensity fluctuating around it mean value <I> as a function of time. Bottom panel is the intensity auto-correlation function as a function of time.

3.4 Integral equation theory

In an equilibrium system the distribution of particles can be calculated approximately using integral theory. Interactions between particles lead to correlations between particle positions. Modeling scattering requires only correlations depending on two position coordinates contained in the radial distribution function g(r) (see equation 3.4). The correlation can be divided into direct and indirect parts. This division results in the Ornstein-Zernike (OZ) equation [51]

$$h(\mathbf{r}_1, \mathbf{r}_2) = c(\mathbf{r}_1, \mathbf{r}_2) + \int c(\mathbf{r}_1, \mathbf{r}_3) n(\mathbf{r}_3) h(\mathbf{r}_3, \mathbf{r}_2) d\mathbf{r}_3 \tag{3.16}$$

The total correlation between a particle at $\mathbf{r}_1$ and another particle at $\mathbf{r}_2$ is given by $h(\mathbf{r}_1, \mathbf{r}_2) = g(\mathbf{r}_1, \mathbf{r}_2) - 1$. The first term on the right-hand side of equation 3.16 represents the direct influence and the second term comprises the indirect influence, which is transmitted via a third particle. This third particle exerts a direct influence via $c(\mathbf{r}_1, \mathbf{r}_3$. In addition, the number of particles at position $\mathbf{r}_3$ is altered by the presence of the particle at $\mathbf{r}_2$, which accounts for the remainder of the indirect correlation. Finally, the third particle is allowed to be anywhere in the available volume, which explains the integration.

To solve the OZ equation, one needs an additional relation between the total and direct correlation functions, a so-called closure relation. Two common closure functions are introduced here. The *Hypernetted Chain* (HNC) approximation is a well established closure that usually works well for long-range repulsion. It is given by

$$g(r) \approx e^{-u(r)/kT + h(r) - c(r)} \tag{3.17}$$

The *Percus-Yevick* (PY) approximation is usually preferred when dealing with short-range interactions. It is given by

$$g(r) \approx e^{-u(r)/kT}[1 + h(r) - c(r)] \tag{3.18}$$

Note that these closures introduce the pair potential that is absent from the OZ equation.

3.4.1 Numerical solution of the integral equation theory

For isotropic, homogeneous systems the OZ function can be simplified as

$$h(r) = c(r) + \int c(r') n h(|\mathbf{r} - \mathbf{r}'|) d\mathbf{r}', \tag{3.19}$$

where r and r' are the different correlation distances, and n is the number density.

This equation can be solved numerically by iteration. In practice, the following procedures can be adopted, here exemplified for the PY closure.

1. the continuous function $\gamma(r) = h(r) - c(r)$ is used to reformulate the OZ equation and the closure relation as

$$c(r) = (e^{-u(r)/kT} - 1)(\gamma(r) + 1)$$

2. with an initial guessing for $\gamma(r)$, $c(r)$ is determined and Fourier transformed numerically.

3. numerically Fourier transform the OZ function, which can be written as

$$\tilde{\gamma}(k) = (n\tilde{c}^{\,2}(k))/(1 - n\tilde{c}(k))$$

   The variable k is here the Fourier transform variable conjugate to r.

4. back Fourier transform the results from step 3 to obtain a new $\gamma(r)$ to restart the iteration from step 2 until the $\gamma(r)$ function converges.

3.4.2 Modeling patchiness

To model the patchiness of the attractive interaction between lactoferrin molecules, a reformulated integral equation theory due to Wertheim has been implemented[40–42]. It considers dimers as separate species of the system in addition to monomers. This is schematically shown in figure 3.6.

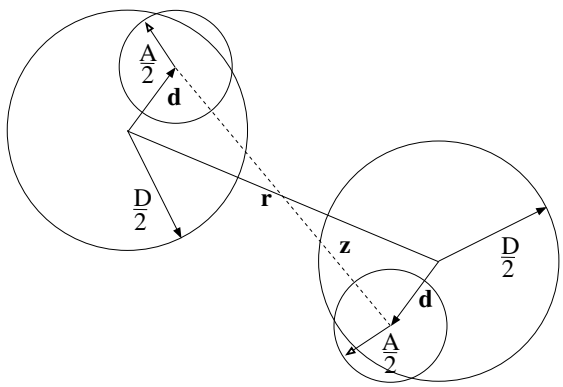


Figure 3.6: Spherical particles of diameter D have been given an attractive site position at $d = |\mathbf{d}|$ from the center with a range of attraction characterized by A along the site-site separation vector separated by center-center distance r.

Dimers are considered to exist whenever the small off-center spheres in figure 3.6 overlap. The Wertheim theory requires that A, d and D satisfy the following conditions

$$0 < \mathrm{A} + 2d - \mathrm{D} < (2 - \sqrt{3})d \leq 0.134\mathrm{D} \tag{3.20}$$

in order to ensure that only dimers and no larger aggregates can form due to the patch attraction. In this case the OZ equation has been derived by Wertheim as

$$h_{ij}(r) = c_{ij}(r) + \sum_{\substack{k=0,1 \\ l=0,1}} \int c_{ik}(r')\rho_{kl}h_{lj}(|\mathbf{r} - \mathbf{r}'|)d\mathbf{r}', \tag{3.21}$$

where $h_{00}(r)$ and $c_{00}(r)$ are (orientationally averaged) monomer-monomer correlation functions. Similarly, $h_{01}(r)$, $h_{10}(r)$, $c_{01}(r)$ and $c_{10}(r)$ are monomer-dimer correlation functions and $h_{11}(r)$ and $c_{11}(r)$ are dimer-dimer correlation functions. Also, in equation 3.21, n_{00} is the total number density, $n_{01} = n_{10}$ is the monomer number density, and $n_{11} = 0$.

In addition, Wertheim has derived HNC- and PY-like closures, which have been found to produce almost identical results, at least when the volume fraction is below 0.15. Another model, the Baxter's adhesive hard sphere[15], which just incorporates an isotropic short-range attraction, is also used in this work.

3.5 Monte Carlo simulations

The Metropolis Monte Carlo (MC) simulations use lactoferrin crystal structure (PDB 1BLF) to model the protein as a rigid body. Simulations were used to calculate the Helmholtz interaction free energy in aqueous salt solution. The atomistic amino acids were coarse-grained to spheres, centered on the center of mass of fully resolved residues. The canonical (NVT) ensemble was computed having translational, rotational MC moves and proton swap moves on ionizable sites.

Salt is treated explicitly in the titration simulations[52], while in the two body simulations it is treated implicitly. The effective interaction potential is described as

$$\beta u = \sum_{i \neq j}^{N} \lambda_B z_i z_j \exp\left(-\kappa r_{ij}\right)/r_{ij} + 4\beta\epsilon\left[\left(\sigma_{ij}/r_{ij}\right)^{12} - \left(\sigma_{ij}/r_{ij}\right)^{6}\right] + \sum_{i}^{N_p} (\mathrm{pH} - \mathrm{p}K_{a,i}) \ln 10, \quad (3.22)$$

where N and N_p represent all the residues and residues that can be protonated, respectively. κ^{-1} and λ_B are the Debye length and Bjerrum length, respectively. Debye length describes the length scale of the exponential decay of the electrostatic interactions. Bjerrum length defines a length scale where the electrostatic interaction is comparable with thermal energy. z is the valencies of the particles, $\beta\epsilon$ is the strength of Lennard-Jones potential, which has a well established value 0.05 for lactoferrin system[34,53], σ_{ij} is the mean particle diameter, r_{ij} is the inter-residue distances, $\beta^{-1} = k_B T$ is the thermal energy, and $\mathrm{pK}_{a,i}$ are the acid dissociation constants of titratable amino acid i. The pKa values of chargeable amino acid residues are listed in table 3.1[54].

Table 3.1: pH values at the isoelectric point of the amino acids in water at 25 °C

Amino acid	Arginine	Histidine	Lysine	Aspartic acid	Glutamic acid
pI	10.76	7.64	9.47	2.98	3.08

The angular averaged pair distribution function is calculated from sampling the protein-protein center of mass separations, $g(r) = exp(-\beta w(r))$. The long-range part of the potential of mean force is fitted to the Debye-Hückel expression to extend the $g(r)$ and suppress the statistical noise. Integration of $g(r)$ gives the osmotic second virial coefficient.

$$B_2 = -2\pi \int_0^\infty \left(g(r) - 1\right) r^2 dr. \quad (3.23)$$

3.6 Protein titration

Protein titration is composed of an acidic titration from pH=7 to pH=2 and a basic titration from pH=4 to pH=11. A complete titration is not feasible in one experiment due to the difficulty of preparing samples at extreme pH. Protein denaturation occurs in between pH 3.0 and 3.6 (see section 2.3.3). High pH is hard to maintain due to the absorption of CO_2.

The consumption of titrant by the bulk is removed by solving the following equation

$$\begin{cases} \pm n_t &= [H^+]_f V_f - [H^+]_i V_i + [OH^-]_i V_i - [OH^-]_f V_f \,, \\ n_t &= (V_f - V_i) c_t \,. \end{cases}$$

where n_t is the molar number of titrant, $\pm$ depends on the type of titrant, $[H^+]_i$ and $[H^+]_f$ are the concentrations of protons before and after adding the titrant. Proton concentration $[H^+]$ is calculated from pH via pH = $-log_{10}[H^+]$. The hydroxide ions $[OH^-]$ are obtained by using the water dissociate constant, $[H^+][OH^-] = 13.998$ at 25°C[55].

Protein is fully protonated at low pH manifested as a titration plateau. However in our measurements the experimental titration curves do not show a satisfactory plateau. Theoretical fit at low pH to the experimental titration curves provides the shift to the absolute scale where titration curves start at fully protonated state. This is done by applying[56]

$$n_{H^+} = N_{Asp} \frac{e^{n_{Asp}(\mathrm{pH}-\mathrm{pKa}_{Asp})}}{1 + e^{n_{Asp}(\mathrm{pH}-\mathrm{pKa}_{Asp})}} + N_{Glu} \frac{e^{n_{Glu}(\mathrm{pH}-\mathrm{pKa}_{Glu})}}{1 + e^{n_{Glu}(\mathrm{pH}-\mathrm{pKa}_{Glu})}} \tag{3.24}$$

where N_{Asp} and N_{Glu} are numbers of aspartic acid (Asp) and glutamic acid (Glu) in lactoferrin, respectively, pKa_{Asp} and pKa_{Glu} are acidic dissociation constants of Asp and Glu, respectively. We consider fitting for only Asp and Glu at low pH because they have the lowest pKa among the chargeable amino acid residues[54].

It is experimentally challenging to obtain the titration plateau. This is due to the disturbance from protein denaturation and background noise from titrating water at low pH.

Taking the derivative of the protein titration curve provides the capacitance[57]

$$C = -\frac{1}{\ln 10} \frac{\partial \langle Z \rangle}{\partial \mathrm{pH}} \tag{3.25}$$

The charge capacitance determines the strength of an attraction introduced by the fluctuation of charges on the ionizable amino-acid residues when proteins move towards each other[43].

Titration was conducted in *Probe Drum*. pH-meter is calibrated from pH=1 to pH=12 with an accuracy on the third digit by using pH standards from ThermoFisher. Electrode is maintained clean before each measurements. Protein solution has a volume of 1 mL and a concentration in the range from 5 mg/mL to 10 mg/mL to have sufficient proton response. HCl and NaOH are the acidic and basic titrant separately with a concentration of 1 M. The step size of adding titrant into protein solution is adjusted to have a reasonable distribution of pH points, After each titrant injection, a stirring process and a period of equilibrating time are given to the system. The pH-meter is calibrated before the titration experiments. Two different ionic strengths, 35 mM and 200 mM are measured to compare the protein titrating ability at different salt concentration.

3.7 Cryogenic transmission electron microscope

Cryogenic transmission electron microscope (cryo-TEM) was used for imaging using Philips CM120 BioTWIN Cryo with images recorded on a CCD camera (Gatan 791). Samples were prepared by using $2 \sim 3$ μL of protein solution on the grid and plugging into liquid ethane immediately after removing excess sample volume by carefully tapping filter paper on the back side of the grid. Stand-by samples were kept in liquid nitrogen. Cryogenic method is to keep the solution structure as natural as possible. The picture analysis was done by using the software Fiji[58].

3.8 Confocal laser scanning microscopy

Confocal laser scanning microscopy (CLSM) was performed on an inverted confocal laser scanning microscope (Leica DMI6000) equipped with SP5 tandem scanner with an oil immersion 100 X objective lens. The sample loading glass was pre-coated with poly-L-lycine for stabilization, and samples were sandwiched between cover glass and hermetically sealed. Lactoferrin was labeled by Alexa 647 fluorescence (Molecular Probes), and dye to protein ratio was 1:156 in order to minimize the dye influence on protein-protein interactions.

Chapter 4

Results

The major results obtained experimentally verify predictions of lactoferrin patch attraction from computer simulations. The consequence of this protein attraction is dimerization at an intermediate protein concentration. A preliminary phase diagram has been determined that is interpreted based on the path-patch interaction mechanism. At the end we explore the relation between titration curves and charge capacitance, obtained both experimentally and from computer simulations, which is related to the charge regulation mechanism.

4.1 Charge-induced patchy attractions between proteins

Work from Persson et al. demonstrated an anisotropic attraction derived from a regio-specific attraction of lactoferrin, which implies that the proteins can dimerize, specifically in terms of orientation[34]. As computer simulations showed there are a few chargeable amino acids located on protein surface. They compose a patch that is complementary to itself, matching electrostatically with the same patch on the other protein surface. This patch results in a short-range electrostatic attraction between lactoferrin molecules.

We investigated this patchy attraction via tuning the Debye screening length by adding salt. This demonstrated the balance between two opposing electrostatic contributions, a short-range electrostatic attraction described above, and the Coulomb repulsion.

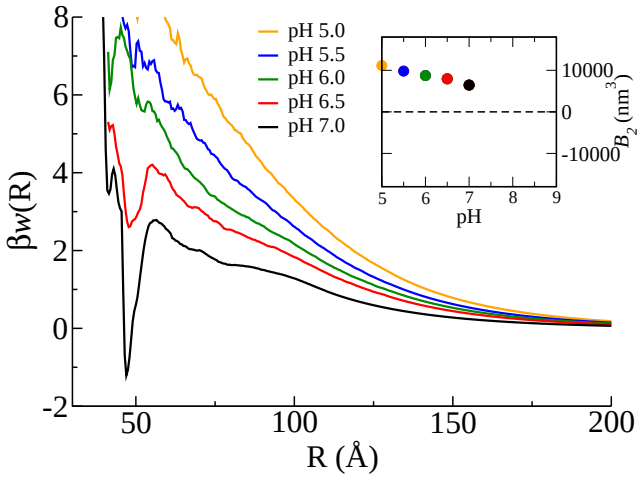


Figure 4.1: Angularly averaged protein interaction potential of mean force, $\beta w(R)$, is calculated as a function of separation between protein centers of mass, R. The pH values are calculated at a constant ionic strength of 5 mM. The inset figure shows the osmotic second virial coefficient, B_2, of corresponding pH.

4.1.1 Locating the condition for short-range attraction

The electrostatic interaction of a protein is different at different pH values. This is because the chargeable amino acid side chains on the surface of protein have different protonation state. The overall net charge of the protein is higher when the pH is far from the protein's isoelectric point (pI) than when the pH is close to pI. Higher overall net charge grants a dominating long-range Coulomb repulsion. In order to diminish the overwhelming repulsion and find a condition where attraction starts to appear, we studied the protein interaction potential of mean force at different pH by MC simulations as shown in figure 4.1.

The interaction potential of mean force (PMF), $\beta w(R)$, stands for an attraction when it is negative and repulsion when it is positive. When the pI is approached, the repulsion decreases from pH=5.0 to pH=7.0 and PMF shows a negative minimum at separation of 45 Å. The inset in the figure shows how B_2 values decline as a function of pH due to the repulsion originating from the overall net charge is reduced gradually when the pH is moving towards the pI. At pH 7 the short-range attraction appears at a short interacting distance.

In order to explain this attraction, experiments on lactoferrin solutions with various pH were performed. From SLS the B_2 was extracted to compare with B_2 obtained from simulations.

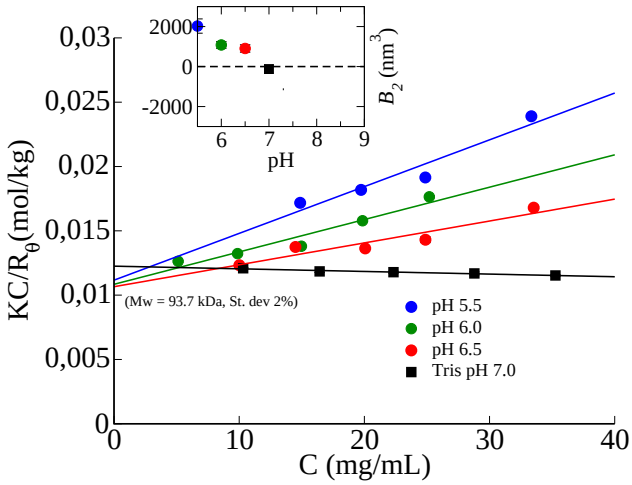


Figure 4.2: Debye plot at different pHs and constant ionic strength (5 mM). Buffer is acetate except at pH 7 (Tris). Lines are weight linear least-square fits to the data as labeled. The inset shows corresponding virial coefficients with error bars.

It is observed from the Debye plot shown in figure 4.2 that slope decreases with increasing pH values. This corresponds to decreasing B_2 values shown in the inset of figure 4.2. At pH=7, B_2 is close to zero, which represents a balance between attraction and repulsion. Based on the interaction energy in the simulations, pH=7 was chosen as the condition of studying the directional attraction in order to test if it was measurable experimentally.

4.1.2 Influences of ionic strength

We next investigate how ionic strength influences the interactions, starting at 5 mM and pH=7. The experimental results from SLS are shown in figure 4.3. At 5 mM the slope is slightly negative, and is decreasing as salt is added corresponding to a more negative B_2. However, as we continue adding salt, the slope instead is increasing and B_2 becomes more positive. The nonmonotonic dependence of the slope as a function of ionic strength results in a B_2 minimum.

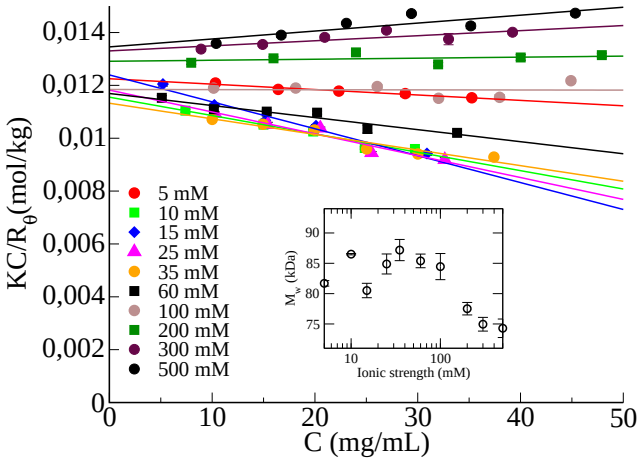


Figure 4.3: Debye plot of different ionic strengths as a function of protein concentration. The measurements were performed in Tris buffer and pH=7. The lines represent linear regression fit of each data set. The inset shows the molecular weight extracted from each intercept as a function of ionic strength.

From the experimental data, it is still unclear which mechanism is responsible for the minimum of B_2. Nevertheless, B_2 calculated from MC simulations at matching conditions demonstrates a similar attraction minimum as in the experimental curve, figure 4.4. The minimum in B_2 has a direct connection to the minimum of angular averaged PMF due to the effect of salt. Figure 4.5 demonstrates that increasing salt concentration from 5 mM to 35 mM leads to a more pronounced energy minimum, while further addition of salt eliminates both the repulsive barrier and the energy minimum.

The observed phenomena comes from two competing electrostatic contributions, a short-range patch-patch attraction and a long-range Coulomb repulsion. Adding salt initially screens the Coulomb repulsion thereby enhancing the effect of the attractive patch. Further grow of salt concentration leading to a decreased Debye length screens the short-range attraction as well. This balance is manifested as a nonmonotonic B_2 as a function of ionic strength observed in both SLS experiments and MC simulations.

We have not intentionally attempted to adjust the only free parameter in MC simulations, the strength of Lennard-Jones potential, ϵ, in equation 3.22, therefore the B_2 values agree qualitatively but not quantitatively.

The inset in figure 4.5 shows an analysis of the angular average PMF, $\beta w(R)$, as

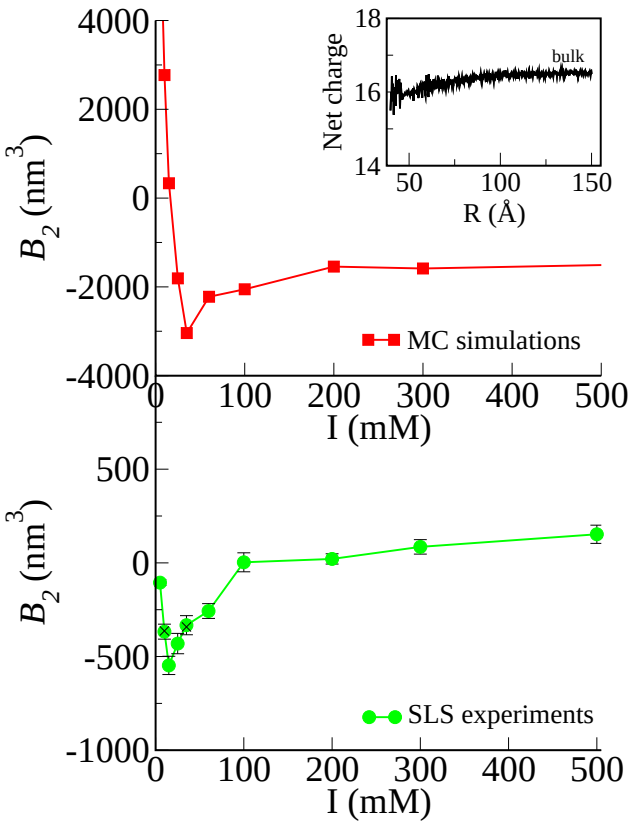


Figure 4.4: B_2 calculated from MC simulations is on the top, and B_2 obtained from experiments is on the bottom, both are plotted as a function of ionic strength. The inset shows the overall net charge as a function of protein interacting distance. In the experiments, 10 mM and 35 mM are repeated on different batches shown as crosses.

a function of protein-protein separation. $\beta w(R)$ contains a Debye-Hückel term accounting for electrostatic repulsion and the Lennard-Jones potential accounting for van der Waals short-range attraction. They both were found to be necessary for the distinct and narrow minimum, corresponding to the tightly bound stereospecific configuration[34]. Artificially switching off either of them eliminates the energy minimum, showing that both contributions are important.

Another type of attraction which could contribute to the protein behavior is the charge regulations. That implies that ionization states of two proteins become correlated at low ionic strength when they come close to each other[43,57]. To confirm the contribution of this interaction, the variation in protein charge as a function

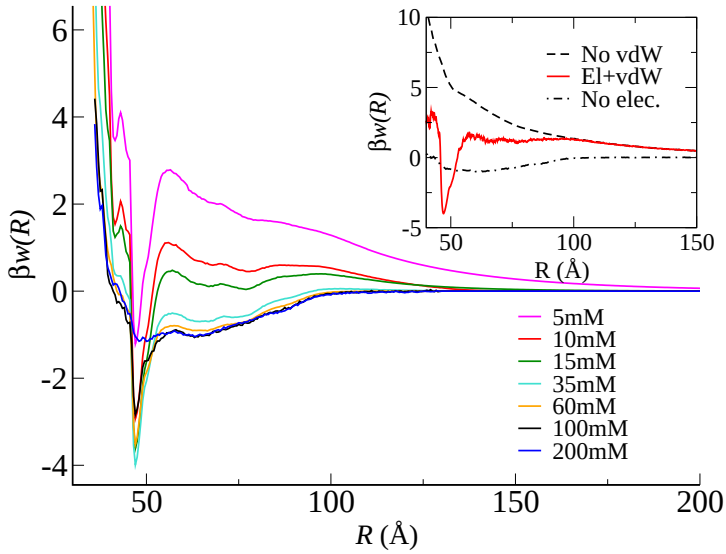


Figure 4.5: Angular averaged protein-protein interaction potential of mean force, $\beta w(R)$ as a function of the distance between centers of mass of interacting proteins, R, at different ionic strengths, pH=7. The inset shows $\beta w(R)$ with either van der Waals attractions or electrostatic repulsion artificially disabled in similar conditions.

of separation was analyzed. As shown in the inset of figure 4.4, the fluctuation of charges is almost negligible when the proteins approach each other. Therefore the fluctuation force seems to be less important at least under these conditions.

4.2 Concentration-induced protein association from a directional patch attraction

Lactoferrin interacts via a directional patch-patch attraction at certain conditions. This attraction was investigated at relatively low protein concentrations where the proteins interact via attractive forces without an indication of association. In this section lactoferrin is examined using SLS and SAXS at higher protein concentration. The aim is to investigate how directional attraction influences the microstructure and how distinct the patchy interaction is compared to the well-known centrosymmetric case. The measurements were performed at two salt concentrations: 35 mM and 200 mM. Initially, the results are analyzed by short-range isotropic attractions[16,59,60]. Alternatively, they are analyzed by integral equation theory for associating systems under the influence of directional attraction from Wertheim[40–42].

4.2.1 Baxter model

Figure 4.6 is a reduced Zimm plot based on SLS data for lactoferrin solutions at ionic strengths 35 and 200 mM. The patch attraction is active at 35 mM and inactive at 200 mM. This is observed from the difference of slopes at low-concentration region, where B_2 is provided. The inset of figure 4.6 demonstrates the nonmonotonic dependence of B_2 on salt concentration (see section 4.1.2). The arrows in the inset point to 35 mM salt where B_2 is negative, and 200 mM salt where B_2 is close to zero, respectively. These two conditions were studied for protein concentrations of up to 150 mg/mL by using SLS and SAXS. Here we focus on the behavior of the effective structure factor, $S(q)$, instead of scattering intensity. By removing the effect of the form factor, $P(q)$, the concentration-induced changes in the structure factor are investigated based on protein interactions.

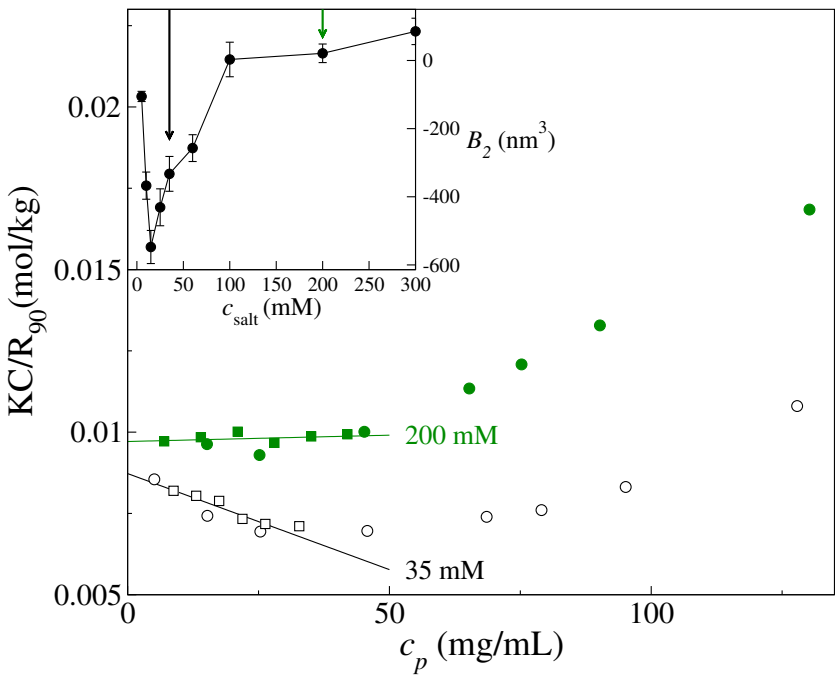


Figure 4.6: Reduced Zimm plot at two different ionic strengths, 35 and 200 mM of salt, as a function of protein concentration. Two sets of data (squares and circles) are measured at both salt concentration. The inset shows measured B_2 in the unit of nm^3 as a function of ionic strength. Two dilute series provide B_2 values from the slopes, and the corresponding B_2 values are indicated by the arrows in the inset.

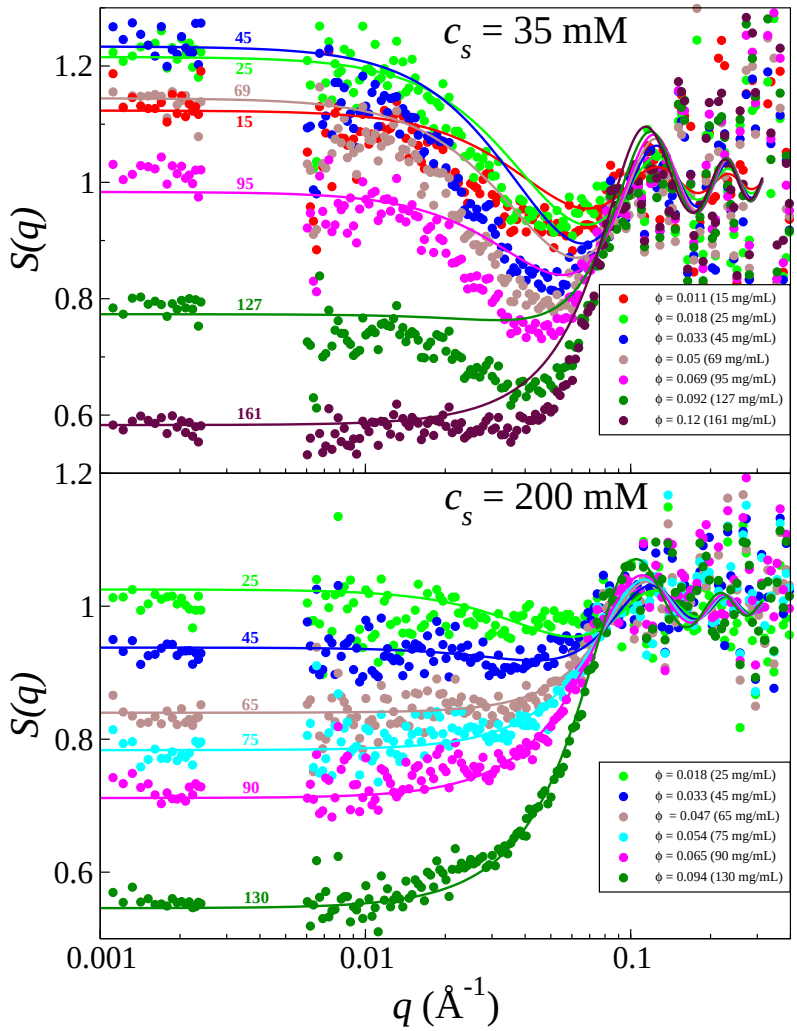


Figure 4.7: Baxter adhesive sphere structure factors (lines) with effective structure factors obtained from SLS (left of gap) and SAXS (right of gap) at 35 and 200 mM of salt, pH=7. The sphere diameter in the analysis was set to 6 nm to acquire a reasonable agreement between experiments and model at high q.

Figure 4.7 shows lowest-q data obtained from SLS, and data to the right of the gap from SAXS. In the presence of the patch attraction, the low-q scattering exhibits a nonmonotonic concentration dependence. The $S(q)$ at low q shows a characteristic ascending of an attractive system as the protein concentration increases initially. As the protein concentration is increased further, the system appears to become more repulsive, presenting a descending $S(q)$ at low q. This suggests that the patch attraction leads to excluded-volume-like interactions as the protein solution becomes

more concentrated. The behavior is quite different for the patch inactive case at 200 mM salt concentration. The scattering at low-q descends constantly as the protein concentration is increased, as expected for particles interacting via repulsive forces.

Using spherically symmetric adhesive model[16,60] as given by Baxter's adhesive sphere[15], the experimental structure factors are fitted as shown in figure 4.7. However, in Baxter model the stickiness parameter, τ, must be varied freely to obtain a reasonable agreement with the experimental data at any salt concentration. The resulting value of τ as a function of volume fraction, ϕ, shown in figure 4.8 demonstrates a non-physical variation of adhesion while changing the particle concentration, indicating a limitation of applying isotropic interaction to patch-patch interaction.

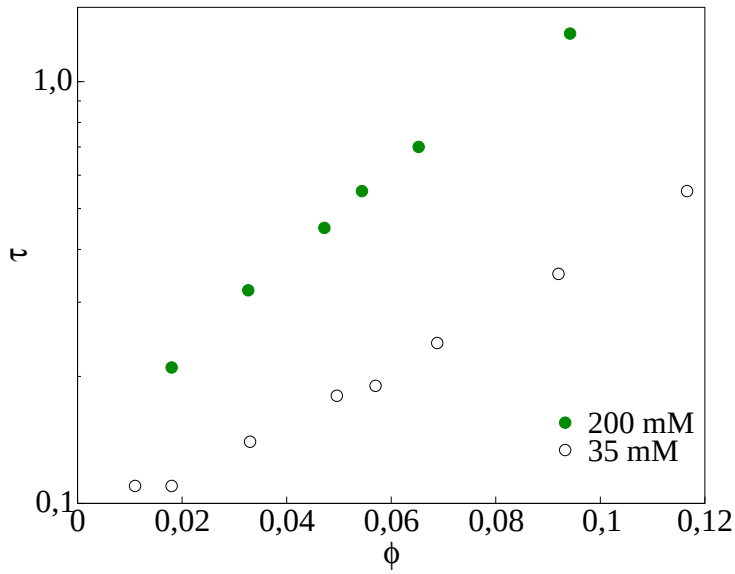


Figure 4.8: Baxter adhesion τ parameter as a function of protein volume fraction ϕ at two different ionic strengths, as extracted from the fitting of the adhesive sphere structure factors in figure 4.7

The work from Scherer et al.[61], by using a constant τ in the adhesive shpere model, shows a significant deviation from monoclonal antibody system where B_2 was negative. They assign this deviation to self-association. Following this reasoning, the apparent concentration dependence of interaction disappears by using a model with directional interaction.

4.2.2 Wertheim model

By considering a model of spheres interacting via a single off-center attractive site, the particles spontaneously associate into dimers to varying extents, depending on patch strength and particle concentration. The steric blocking of the attractive patch when dimer is formed contributes to excluded-volume interactions which govern the system when sufficient amount of dimers are formed as the volume fraction is increased. Sciortino et al. have studied this patch system over rather broad range of parameters[21–23,62]. For dilute-to-semidilute solutions, Wertheim model appropriately handles the highly directional attractions.

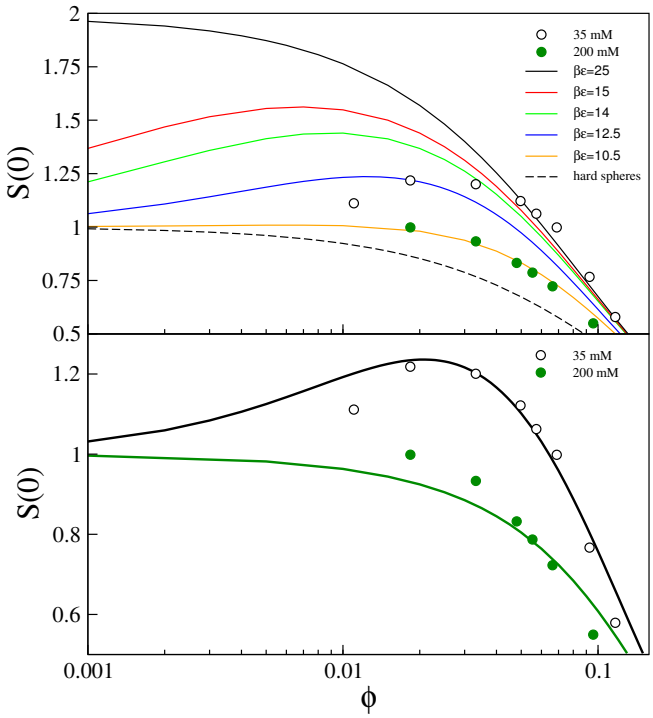


Figure 4.9: The structure factor at long wavelength limit as a function of volume fraction, ϕ. The upper panel is the result from Wertheim's integral equation theory for single-patch hard spheres shown as solid lines for various patch strengths as labeled. Experimental $S(0)$ is obtained by extrapolating the measured structure factor to low-q region at two salt concentrations as labeled. The bottom panel is the result from Wertheim model with addition of isotropic Yukawa and Coulomb potential, at patch strength of 11.2 (top) and 0 (bottom), as solid lines.

Proteins are initially treated as hard spheres with addition of an attractive patch

controlled in a way such that trimer formation is not allowed. Figure 4.9 shows the prediction for large wavelength limit of the structure factor, $S(0)$, from Wertheim's integral equation theory for one-patch spheres. The system is determined by the volume fraction and patch strength. For a very strong patch attraction, such as $\beta\epsilon$ = 25 in the upper panel of figure 4.9, $S(0)$ equals to the isothermal compressibility of hard-sphere dimers with the result for an ideal gas in the dilute limit, $S(0)$ = 2. Because of the excluded-volume interactions, $S(0)$ decreases with increasing volume fraction, ϕ. Similar monotonic decrease of $S(0)$ is found for weak patch attractions, $\beta\epsilon \lesssim 10$ in the upper panel of figure 4.9, where the excluded-volume interactions dominate. In contrast, for intermediate patch strengths, $S(0)$ behaves monotonically as a function of ϕ. In this range of patch strength, there is an equilibrium between monomers and dimers. Initially, dominating patch attraction from monomers increases $S(0)$. Once the dimer concentration becomes sufficiently high and the excluded-volume interactions become dominant, $S(0)$ starts to decrease.

Comparing the model with the experimental data in the upper panel of figure 4.9, the intermediate patch strength $\beta\epsilon = 12.5$ provides qualitatively similar $S(0)$ trend for the experimental data at 35 mM salt where the patch attraction dominates. The $S(0)$ falls monotonically with volume fraction when the patch strength is strongly reduced for experiments at 200 mM salt. For experiments, a sharp maximum of $S(0)$ appears at $\phi \approx 0.02 - 0.03$.

The agreement between Wertheim theory for $S(0)$ and the experimental data at 35 mM salt can be improved by adding isotropic interactions that is closer to reality than the hard spheres. The following isotropic potentials are superimposed on the model

$$\beta\phi_{\mathrm{iso}}(r) = \begin{cases} \infty & r < D \\ -\dfrac{Ke^{-\kappa_d(r-D)}}{r/D} + \dfrac{L_B Z_{\mathrm{eff}}^2 e^{-\kappa_d(r-D)}}{(1+\kappa_d D/2)^2 r} & r > D \end{cases} \tag{4.1}$$

where the repulsive part is a screened Coulomb interaction, and the attractive part has the Yukawa form. With these two terms considered, the monomer-dimer equilibrium can be shifted to give a maximum close to that seen experimentally at 35 mM salt, as shown in the bottom panel of figure 4.9. In contrast to the concentration-dependent τ used in the absence of the patch attraction, a constant patch strength was applied for all ϕ to capture the behavior of 35 mM salt condition.

Directional attraction leading to dimer formation at intermediate protein concentrations is also supported by the DLS measurements. The diffusion coefficient behaves differently between the situation when the patch attraction is switched-on, 35 mM salt, and switched-off, 200 mM salt, as shown in figure 4.10. The diffusion coefficient at zero concentration limit, D_0, obtained for protein concentration in

200 mM salt coincides with D_0 obtained for the dilute protein solution in 35 mM salt, as labeled in blue and green line respectively in figure 4.10. The intermediate concentrations provide a D_0 close to half of the above D_0. Form Stokes-Einstein equation, this gives two hydrodynamic radius that differ by a factor of close to two. This suggests the monomeric form of the protein dominates at 200 mM salt and also in the dilute solution at 35 mM salt, however, the intermediate concentrations at 35 mM could have dimers dominating. The boundary for monomers and dimers in 35 mM salt solution is at a volume fraction of $\phi \approx 0.04$.

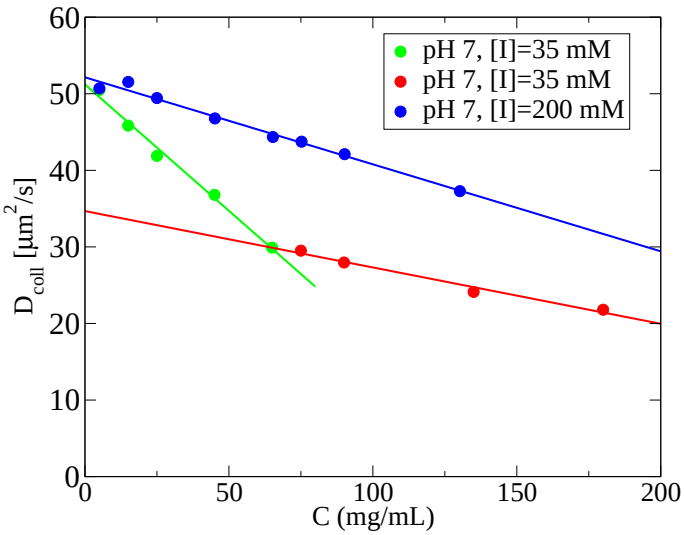


Figure 4.10: Collective diffusion measured as a function of protein concentration at 35 mM and 200 mM ionic strength separately. 35 mM is fitted by two lines crossing at a protein concentration around 60 mg/mL ($\phi \approx 0.04$).

4.3 Solution structures formed by proteins with patch-patch attractions

In this section, the self-association and phase diagram of lactoferrin at pH and ionic strength where proteins are interacting via patch-patch attraction is discussed. Previously it has been shown that with the grow of protein concentration, protein dimers and monomers can be in equilibrium due to the patch attractions. As the concentration is increased further, the patch attraction induces formation of highly organized, monodisperse macroscopic stripe-like structures.

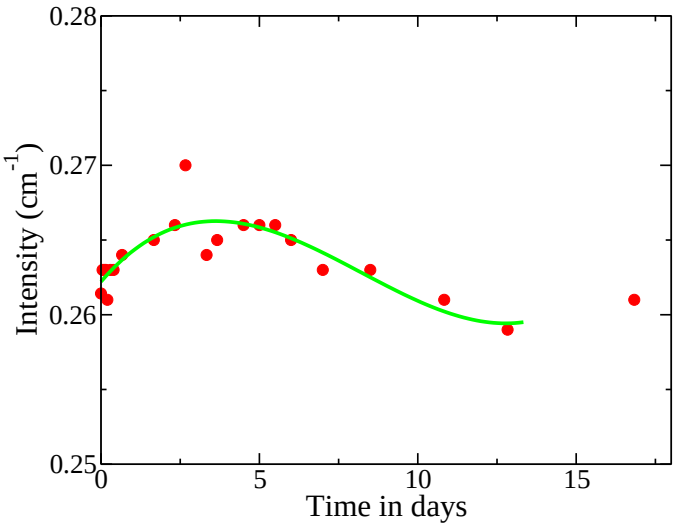


Figure 4.11: Time study from SLS for protein concentration of 100 mg/mL, ionic strength of 35 mM and pH=7. Data points are angular averaged intensity with polynomial fit as guided by eye.

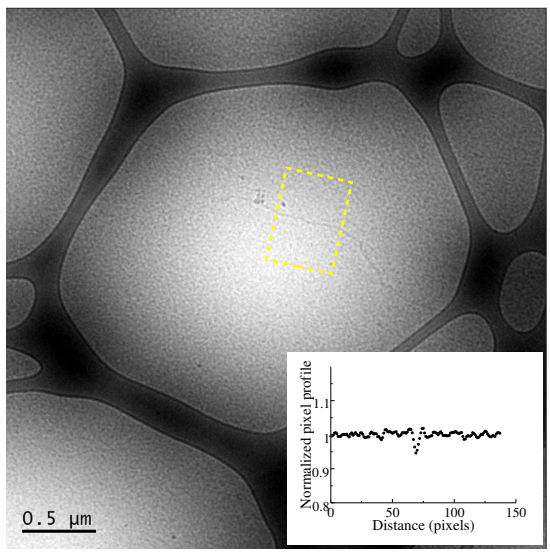


Figure 4.12: Cryo-TEM image of a thread-like structure. Protein solution used here has a concentration of 80 mg/mL, 35 mM salt and pH=7. Equilibrating time is 7 days. Inset shows the normalized pixel profile where average intensity was taken along the direction of the stiring in the marked yellow dash box.

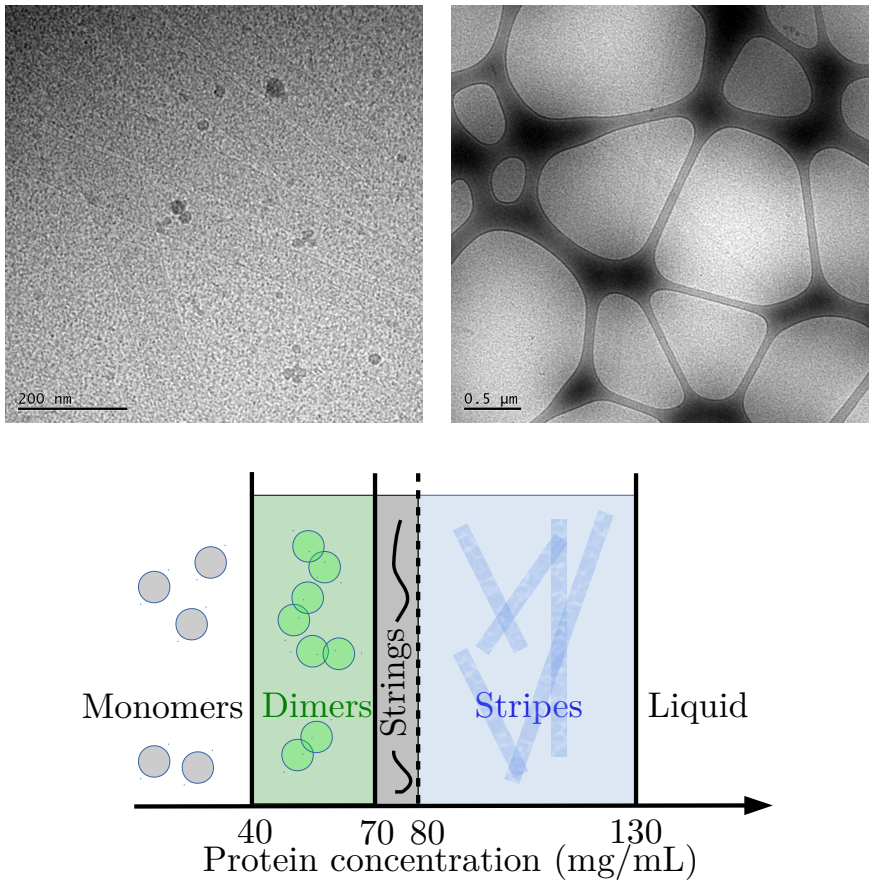


Figure 4.13: Schetch of lactoferrin's phase diagram under a pair-wised attraction. Bottom left is the Wertheim model on single patch-patch attractive spheres. Bottom middle is a representation of lactoferrin dimer. Top two are electron microscopic picture with different magnification connected to the bottom right which is the overall phase diagram on protein concentration with simplified carton demonstrating the corresponding phase.

The stripe phase was difficult to detect by SAXS because only a small portion of the proteins participates in the formation of stripes and they are susceptible to sedimentation due to micron-sized entities. Instead we detected the self-association of lactoferrin monomers by monitoring the overall scattered intensity by means of SLS, since the formation of structures influences the static and dynamic properties of the solution. At the initial stage, dimerization of the monomers decreases the number density but increases the particle size, leading to an overall effect of slight increment of scattered intensity due to the fact that scattered intensity depends more on particle size than the number density[47].

As shown in figure 4.11, the scattered intensity increment originates from multiple molecular species within the first four days. After four days, due to the sedimentation of macro-sized structures the scattered intensity decreases. Since it is experimentally challenging to investigate this structure by SAXS, we studied it by using cryo-transmission electron microscope (cryo-TEM) and confocal laser scanning microscopy (CLSM).

The upper part of figure 4.13 shows the cryo-TEM images of the new stripe-like phase acquired at different magnifications. The frosts appear as big dark speckle, which are not protein structures. A thin thread-like pre-structure was also discovered at concentrations slightly smaller than 80 mg/mL, as shown in figure 4.12. This could be a separate structure or precursors before the formation of stripes. The stripe are constant on the width which is approximately 100 nm and micron-sized in length. Experiments were validated by repeating the cryo-TEM measurements on different batches of protein solutions. In general, the structures were able to be detected after few weeks of equilibrating, images taken on freshly prepared sample showed neither stripes nor wires. Poor contrast in cryo-TEM images is coming from high concentrations of dimers and possible higher order oligomers in the background. This is reflected in the pixel profile shown in the inset of figure 4.12, i.e. it is shown that the difference in contrast between string and background is only about 5%.

To strengthen the observation that the macro-sized protein solution structure is not an artificial effect induced by cryo freezing, CLSM was applied. The resulting image is shown in figure 4.14.

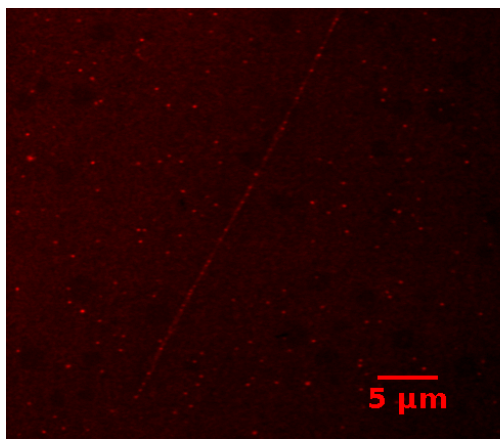


Figure 4.14: Confocal microscope image of the stripe structure form by lactoferrin at 100 mg/mL, 35 mM salt and pH=7. Alexa 647 was the fluorescent label.

4.4 Protein charge capacitance obtained from experiments and MC simulations

Titrations performed at pH 10 on two different ionic strengths, 35 mM and 200 mM, are shown in figure 4.15. Proton association per protein molecule is the difference between protons associated to the protein solution and protons associated to water. It shows difficulty to experimentally titrate water repetitively, so the water titration calculated from theory is used as the background. Data is presented in terms of associated protons from the acidic end, so negative values meaning number of protons dissociated from positively charged amino acid residues. Experimental repetitions were spline interpolated, then acidic titration (full circles) and basic titration (full squares) are merged based on minimizing the standard deviations of overlapping points, and the averages (full lines) are calculated afterwards.

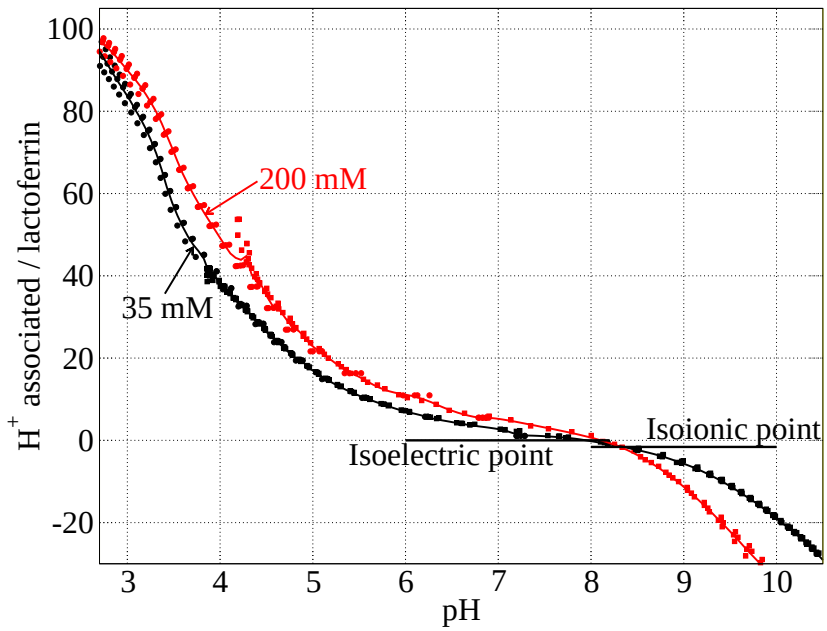


Figure 4.15: Lactoferrin titration at two ionic strengths as labeled. Symbols are experimental data, repetitions are included. Two lines are the average of spline interpolation of each salt repetitions.

The titration plateau is used to shift titration curve to absolute scale, which shows the correct number of protein overall charges. The difficulties we have encountered in experiments of obtaining a titration plateau are protein denaturation and large proton consumption by water, at low pH. Instead, a theoretical titration plateau

that is calculated from fitting the low pH part of titration curve by using Asp and Glu residues scales the experimental data to absolute values. The discrepancy of experimental repetitions at low pH reduces the accuracy of theoretical plateau's position. Preliminary theoretical plateau gives an isoelectric point (zero charge) and isoionic point (zero ion adsorption) around 8 as shown in figure 4.15. Similar values are observed for lactoferrin in literature[63].

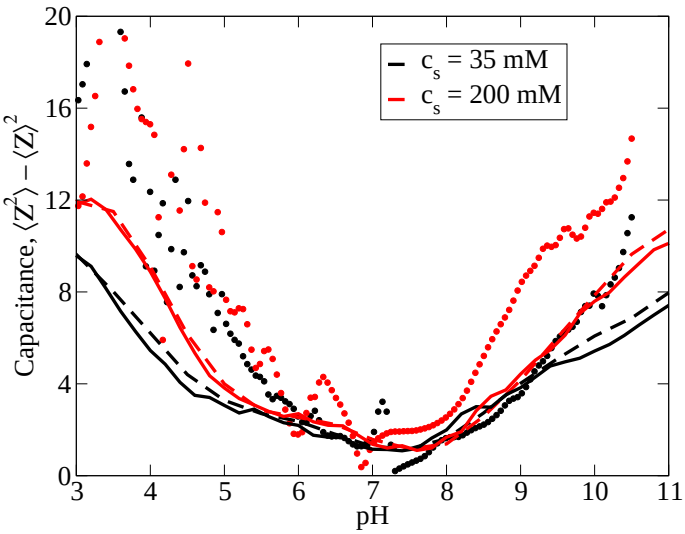


Figure 4.16: Capacitance (left) and protein charge (right) of lactoferrin as a function of pH at 35 mM (red) and 200 mM (black) salt concentrations. From MC simulations of lactoferrin monomers (dashed), dimers (fully drawn), and measured (symbols).

Comparison of capacitance by differentiating titration curve obtained from MC simulations and experiments is revealed in Fig. 4.16. Both show a generally increased capacitance as salt is increased. This is expected due to decreased internal repulsion due to screening. Lack of agreement in the very high and low pH is likely to be understood from the fact that not only protein but also water titrates significantly in these regimes.

At pH 7, lowest point of capacitance as seen in figure 4.16, has a zeta potential of 4.55 mV found by the measures of the electrophoretic mobility, at 5 mM salt. By variation of pH we searched for the corresponding mobility above the isoelectric point and found the zeta potential at pH 10 to be -5.23 mV. The values are listed in table 4.1

From this one would expect the absolute charge of the protein to be similar at these

Table 4.1: Measured properties at 5 and 20 mM salt, (I), at pH 7 and pH 10.

	B_2 (nm^3)		ZP (mV)		D_0 (μm^2/s)		M_w (kg/mol)	
I (mM)	5	20	5	20	5	20	5	20
pH 7	-105.58	-490*	4.55	|	57.078	56.705	81	83*
pH 10	-68.76	-251.72	-5.23	|	35.684	40.157	373	268

* value is the average of corresponding 15 and 25 mM salt results.

two pH values. Simulations put the valency of the protein at 15.0 and -5.4 at pH 7 and 10 respectively. This suggests that ion binding could have occurred[65].

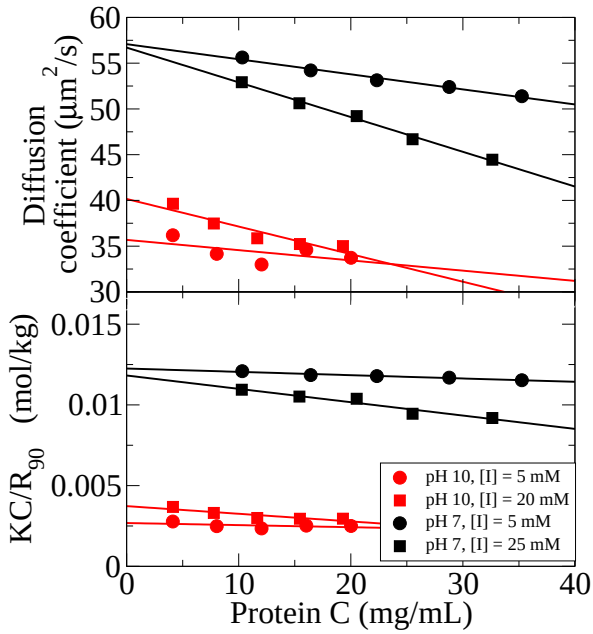


Figure 4.17: Top panel shows collective diffusion as a function of protein concentration. Bottom panel show the reduced Zimm-plot. Conditions are labeled.

Figure 4.17 shows collective diffusion as a function of protein concentration and reduced Zimm-plot at 5 and 20 mM salt, each at pH 7 and pH 10, as shown in the legend. The D_0 from each salt and pH conditions, listed in table 4.1, are indicating two sized species. By applying the Stoke-Einstein equation, one small sized species given by the D_0 measured at pH 7 is about 4.3 nm corresponding to the radius of a monomer[37,39], and one big sized species at pH 10 is about 6.5 nm. Table 4.1 also provides M_w, calculated from the intercept of reduced Zimm-plot. Values of M_w from pH 7 and pH 10 can be again considered as two species, and they differs about four time. DLS measurements on human lactoferrin suggests that

diffusion can be influenced by changing the pH[66]. Titration on lactoferrin dimer was studied by MC simulation shown in figure 4.16. Dimerization, if driven by charge regulation, would be manifested in the stabilization of either protonated or deprotonated states of certain residues, resulting in a decreased capacitance. What is found, is a capacitance that is independent of aggregation state.

B_2 in table 4.1, extracted from slopes of reduced Zimm-plot, is showing increased attraction by adding salt from 5 to 20 mM, at both pH 7 and pH 10. Reduced Coulomb repulsion because of adding salt could be the main reason. At pH 7, the identified patch attraction[67,68] is contributing to the total interactions. At pH 10, attraction induced by charge regulation is speculated to be involved in the interaction between the protein aggregates (trimers or tetramers) due to its prominent capacitance.

Chapter 5

Conclusions and Future Prospects

Study on lactoferrin has revealed an patch-patch attraction at pH close to the isoelectric point. This attraction is short-range electrostatic, originating from few ionizable amino acid residues. The electrostatic origin is understood and proved by tuning the ionic strength in Monte Carlo (MC) simulations and in static light scattering (SLS) experiments. A qualitative agreement was found on the second virial coefficient B_2 between simulations and experiments, both showing a nonmonotonic behavior of B_2 as a function of salt concentration. This is the consequence of screening two opposing interactions, Coulomb repulsion and short-range electrostatic attraction.

Under attractive condition, using small-angle X-ray scattering (SAXS), increasing the protein concentration results in a maximum on isothermal compressibility. Studying this results by applying isotropic Baxter model needs a nonphysical adaption on the interaction strength as protein concentration increases. This limitation is eliminated by modeling a patch-patch interaction via Wertheim's integral equation theory. The patch-patch interaction leads to dimerization when protein concentration increases, which contributes excluded-volume repulsion to the total interaction, manifested as decreasing compressibility. Dimerization is supported by other observations. The hydrodynamic radius obtained from dynamic light scattering (DLS) is increased about two times from dilute to semi-dilute protein concentrations. The concentration boundary between dilute to semi-dilute matches the position of compressibility maximum.

Further increase of protein concentration when patchy interaction plays, macro-sized structures are discovery by using cryo-transmission microscopy (cryo-TEM). A stripe-like structure is seen from cryo-TEM images, with a dimension of 100 nm wide and few micrometer long. Similar phase was predicted by using a comparable

patchy colloids[21]. We have shown a preliminary lactoferrin phase diagram based on what we have studied.

At the end protein capacitance was calculated from experimental titration and MC simulated titration, qualitatively good agreement was found between these two methods. From electrophoretic mobility experiment, two pH conditions having opposing zeta potential but similar absolute charges are located, pH 7 and pH 10. The dynamic light scattering (DLS) has revealed that at pH 10, protein solution provides a radius that is two time larger than at pH 7, which provides monomers. Possible protein aggregation or association is also observed by using static light scattering (SLS) where the molecular weight calculated from Debye-plot shows a four-fold increase from pH 7 to pH 10. The effect of associated proteins on protein capacitance seems minor by comparing monomer and dimer titrations in MC simulations. Capacitance is related to the charge regulation attraction, but further investigations are needed on higher salt concentration to ensure the role of charge regulation.

Overall, this work has studied the patch-patch interaction on lactoferrin, and show different phases due to this patchy interaction as protein concentration increases.

References

[1] Bill Wickstead and Keith Gull. The evolution of the cytoskeleton. *J. Cell Biol.*, 194:513–525, 2011.

[2] Louise N. Johnson and Petsko Gregory A. David philips and the origin of structural enzymology. *Trends Biochem. Sci.*, 24:287–289, 1999.

[3] Carlos Bustamante, Chemla Yann R, Nancy R. Forde, and David Izhaky. Mechanical processes in biochemistry. *Annu. Rev. Biochem.*, 73:705–748, 2004.

[4] Eric Dickinson. Interfacial structure and stability of food emulsions as affected by protein–polysaccharide interactions. *Soft Matter*, 4:932–942, 2008.

[5] Gustav Røder, Linda Geironson, Iain Bressendorff, and Kajsa Paulsson. Viral proteins interfering with antigen presentation target the major histocompatibility complex class i peptide-loading complex. *J. Virology*, 82(17), 2008.

[6] David A. Nagib and David W. C. MacMillan. Trifluoromethylation of arenes and heteroarenes by means of photoredox catalysis. *Nature*, 480:224–228, 2011.

[7] Raffaele Mezzenga and Peter Fischer. The self-assembly, aggregation and phase transitions of food protein systems in one, two and three dimensions. *Rep. Prog. Phys.*, 76(046601), 2013.

[8] Roberto Piazza. Protein interactions and association: an open challenge for colloid science. *Curr. Opin. Colloid Interface Sci.*, 8:515–522, 2004.

[9] Roberto Piazza. Interactions and phase transitions in protein solutions. *Curr. Opin. Colloid Interface Sci.*, 5:38–43, 2000.

[10] Jennifer J. McManus, Alekey Lomakin, Olutayo Ogun, Ajay Pande, Markus Basan, and Jayanti Pande. Altered phase diagram due to a single point mutation in humam γd-crystallin. *Proc. Natl. Acad. Sci. U. S.*, 104:16856–16861, 2007.

[11] A. C. Dumetz, M. Snellinger-O'Brien, E.W. Kaler, and A. M. Lenhoff. Patterns of protein-protein interactions in salt solutions and implications for protein crystallization. *Protein Sci.*, 16:1867–1877, 2007.

[12] Diana Fusco, Jeffrey J. Headd, Alfonso De Simone, Jun Wang, and Patrick Charbonneau. Characterizing protein crystal contacts and their role in crystallization: rubredoxin as a case study. *Soft Matter*, 10:290–302, 2014.

[13] W Nicholson Price II, Yang Chen, Samuel K Handelman, Helen Neely, Philip Manor, Richard Karlin, Rajesh Nair, Jinfeng Liu, Michael Baran, John Everett, Saichiu N Tong, Farhad Forouhar, Swarup S Swaminathan, Thomas Acton, Rong Xiao, Joseph R Luft, Angela Lauricella, George T DeTitta, Burkhard Rost, Gaetano T Montelione, and John F Hunt. Understanding the physical properties that control protein crystallization by analysis of large-scale experimental data. *Nat. Biotechnol.*, 27:51–57, 2009.

[14] Stefan Auer and Daan Frenkel. Prediction of absolute crystal-nucleation rate in hard-sphere colloids. *Nature*, 409:1020–1023, 2000.

[15] R.J. Baxter. Percus-yevick equation for hard spheres with surface adhesion. *J. Chem. Phys.*, 49(2770), 1968.

[16] Roberto Piazza, Véronique Peyre, and Vittorio Degiorgio. "sticky hard spheres" model of proteins near crystallization: A test based on the osmotic compressibility of lysozyme silutions. *Phys. Rev. E*, 48:R2733–R2736, 1998.

[17] Stefano Buzzaccaro, Roberto Rusconi, and Roberto Piazza. "sticky" hard spheres: Equation of state, phase diagram, and metastable gels. *Phys. Rev. Lett.*, 99:098301, 2007.

[18] Frédéric Cardinaux, Thomas Gibaud, Anna Stradner, and Peter Schurtenberger. Interplay between sponodal decomposition and glass formation in proteins exhibiting short-range attractions. *Phys. Rev. Lett*, 99(118301), 2007.

[19] E. Bianchi, J. Largo, P. Tartaglia, E. Zaccarelli, and Francesco Sciortino. Phase diagram of patchy colloids: Towards empty liquids. *Phys. Rev. Lett.*, 97(168301), 2006.

[20] N. Kern and D. Frenkel. Fluid-fluid coexistence in colloidal systems with short-ranged strongly directional attraction. *J. Chem. Phys.*, 118:9882–9889, 2003.

[21] Gianmarco Munaò, Zdeněk Preisler, Teun Vissers, Frank Smallenburg, and Francesco Sciortino. Cluster formation in one-patch colloids: low coverage results. *Soft Matter*, 9:2652–2661, 2013.

[22] Achille Giacometti, Fred Lado, Julio Largo, Giorgio Pastre, and Francesco Sciortino. Effects of patch size and number within a simple model of patchy colloids. *J. Chem. Phys.*, 132(174110), 2010.

[23] Zdeněk Preisler, Teun Vissers, Gianmarco Munaò, Frank Smallenburg, and Francesco Sciortino. Equilibrium phase of one-patch colloids with short-range attractions. *Soft Matter*, 10:5121–5128, 2014.

[24] Christoph Gögelein, Gerhard Nögele, Remco Tuinier, Thomas Gibaud, Anna Stradner, and Peter Schurtenberger. A simple patchy colloid model for the phase behavior of lysozyme dispersions. *J. Chem. Phys.*, 129(085102), 2008.

[25] B.L. Neal, D. Asthagiri, O.D. Velev, A.M. Lenhoff, and E.W. Kaler. Why is the osmotic second virial coefficient related to protein crystallization? *J. Cryst. Growth*, 196:377–387, 1999.

[26] Abraham George and W. William Wilson. Predicting protein crystallization from a dilute solution property. *Acta Cryst.*, 50:361–365, 1994.

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

[28] Bruno H. Zimm. Application of the methods of molecular distribution to solutions of large molecules. *J. Chem. Phys.*, 14(3):164–179, 1946.

[29] Y. U. Moon, C. O. Anderson, H. W. Blanch, and J. M. Prausnitz. Osmotic pressures and second virial coefficients for aqueous saline solutions of lusozyme. *Fluid Phase Equilib.*, 168:229–239, 2000.

[30] Alexander Grünberger, Pin-Kuang Lai, Marco A. Blanco, and Christopher J. Roberts. Coarse-grained modeling of protein second osmotic virial coefficients: Steric and short-ranged attractions. *J. Phys. Chem. B*, 117:763–770, 2012.

[31] Marcin Deszczynski, Stephen E. Harding, and Donald J. Winzor. Negative second virial coefficients as predictors of protein crystal growth: Evidence from sedimentation equilibrium studies that refutes the designation of those light scattering parameters as osmotic virial coefficients. *Biophys. Chem.*, 120:106–113, 2005.

[32] Patrick B. Warren. Simple models for charge and salt effects in protein crystallization. *J. Phys.: Condens. Matter*, 14:7617–7629, 2002.

[33] A. C. Dumetz, A. M. Chockla, E.W. Kaler, and A. M. Lenhoff. Effects of ph on protein-protein interactions and implications for protein phase behavior. *Biochim. Biophys. Acta*, 1784:600–610, 2008.

[34] Björn A. Persson, Mikael Lund, Jan Forsman, Dereck E.W. Chatterton, and Torbjörn Åkesson. Molecular evidence of stero-specific lactoferrin dimers in solution. *Biophys. Chem.*, 151:187–189, 2010.

[35] Håvard Jenssen and Robert E. W. Hancock. Antimicrobial properties of lactoferrin. *Biochimie*, 91(19–29), 2009.

[36] B. Lönnerdal and S. Iyer. *Lactoferrin: Molecular Structure And Biological Bunction*. Annual Reviews Inc. a, P.O. Box 10139, 4139 EI Camino Way, Palo Alto, California 94306, USA, Dep. Nutr., Univ. Calif., Davis, CA 95616, USA, 1995.

[37] Jeremy H. Brock. The physiology of lactoferrin. *Minireview, Biochem. Cell Biol.*, 80:1–6, 2002.

[38] Hum Bokkhim, Nidhi Bansal, Lisbeth Grøndahl, and Bhesh Bhandari. Physico-chemical properties of different forms of bovine lactoferrin. *Food Chem.*, 141:3007–3013, 2013.

[39] Stanley A. Moore, Bryan F. Anderson, Colin R. Groom, M. Haridas, and Edward N. Baker. Three-dimensional structure of diferric bovine lactoferrin at 2.8 øA resolution. *J. Mol. Biol.*, 274:222–236, 1997.

[40] M.S. Wertheim. Fluids with highly directional attractive forces. ii. thermodynamic perturbation theory and integral equations. *J. Stat. Phys.*, 35:35–47, 1984.

[41] M.S. Wertheim. Fluids with highly directional attractive forces. i. statistical thermodynamics. *J. Stat. Phys.*, 35:19–34, 1984.

[42] M.S. Wertheim. Fluids of dimerizing hard spheres, and fluid mixtures of hard spheres and dispheres. *J. Chem. Phys.*, 85:2929–2936, 1986.

[43] John G. Kirkwood and John B. Shumaker. Forces between protein molecules in solution arising from fluctuations in proton charge and configuration. *Proc. Natl. Acad. Sci*, 38:863–871, 1952.

[44] Melanie R. Nilsson and Christopher M. Dobson. *In vitro* characterization of lactoferrin aggregation and amyloid formation. *Biochem.*, 42:375–382, 2003.

[45] Ashoka Sreedhara, Ragnar Flengsrud, Vishweshwaraiah Prakash, Daniel Krowarsch, Thor Langsrud, Purnima Kaul, Tove G. Devold, and Gerd E. Vegarud. A comparison of effects of ph on the thermal stability and conformation of caprine and bovine lactoferrin. *Int. Dairy J.*, 20:487–494, 2010.

[46] J. K. G. Dhont. *An introduction to Dynamics of Colloids*. Elsevier, Amsterdam, 1996.

[47] Peter Lindner and Thomas Zemb. *neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter*. ELSEVIER SCIENCE B.V. Sara Burgerhartstraat 25 P.O. Box 211, 1000 AE Amsterdam, The Netherlands, Elsevier Science Global Rights Departent, PO Box 800, Oxford OX5 1DX, UK, 2002.

[48] Bruno H. Zimm. The scattering of light and the radial distribution function of high polymer solutions. *J. Chem. Phys.*, 16(12):1093–1099, 1948.

[49] Huaying Zhao, Patrick H. Brown, and Peter Schuck. On the distribution of protein refactive index increments. *Biophys. J.*, 100:2039–2317, 2011.

[50] A. J. F. Siegert. *Radiation Laboratory report No. 465*. Cambridge, Mass. Radiation Laboratory, Massachusetts Institute of Technology, 1943.

[51] Roland Kjellander. *The basis of statistical thermodynamics or My favourite path to thermodynamics and beyond*. R. Kjellander, Compendium, 1991.

[52] Christophe Labbez and Bo Jönsson. *A new Monte Carlo method for the titration of molecules and minerals*, chapter 8, pages 66–72. Computational Science & Engineering, 2007.

[53] Mikael Lund and Bo Jönsson. A mesoscopic model for protein-protein interactions in solution. *Biophys. J.*, 85:2940–2947, 2003.

[54] Robert C. Weast and Malvin J. Astle. *CRC Handbook of Chemistry and Physics*. CRC Press, Inc. Boca Raton, Floride, 1982.

[55] H. S. Harned and B. B. Owen. *The Physical Chemistry of Electrolytic Solutions (3rd ed.)*. New York: Reinhold, 1958.

[56] Thom Leiding, Kamil Górecki, Tomas Kjellman, Sergei A. Vinogradov, Cecilia Hägerhäll, and Sindra Peterson Årsköld. Precise detection of ph inside large unilamellar vesicles using membrane-impermeable dendritic porphyrin-based nanoprobes. *Anal. Biochem*, 338(296–305), 2009.

[57] Mikael Lund and Bo Jönsson. Charge regulation in biomolecular solution. *Q. Rev. Biophys.*, 46:256–281, 2013.

[58] Schindelin Johannes, Arganda Carreras Ignacio, Frise Erwin, Kaynig Verena, Longair Mark, Pietzsch Tobias, Preibisch Stephan, Rueden Curtis, Saalfeld Stephan, Schmid Benjamin, Tinevez Jean Yves, White Daniel James, Hartenstein Volker, Eliceiri Kevin, Tomancak Pavel, and Cardona Albert. Fiji: an open-source platform for biological-image analysis. *Nature Methods*, 9:676–682, 2012.

[59] D. Roberts, R. Keeling, M. Tracka, C. F. van der Walle, S. Uddin, J. Warwicker, and R. Curtis. The role of electrostatics in protein-protein interactions of a monoclonal antibody. *Mol. Pharmaceutics*, 11:2475–2489, 2014.

[60] Bernard M. Fine, Aleksey Lomakin, Olutayo O. Ogun, and George B. Benedek. Static structure factor and collective diffusion of globular proteins in concentrated aqueous solution. *J. Chem. Phys*, 104:326–335, 1996.

[61] T. M. Scherer, J. Liu, S. J. Shire, and A. P. Minton. Intermolecular interactions of igg1 monoclonal antibodies at high concentrations characterized by light scattering. *J. Phys. Chem. B*, 114:12948–12957, 2010.

[62] Achille Giacometti, Fred Lado, Julio Largo, Giorgio Pastre, and Francesco Sciortino. Phase diagram and structural properties of a simple model for one-patch particles. *J. Chem. Phys.*, 131(174114), 2009.

[63] Fabiana Superti, Rosa Siciliano, Barbara Rega, Francesco Giansanti, Piera Valenti, and Giovanni Antonini. Involvement of bovine lactoferrin metal saturation, sialic acid and protein fragments in the inhibition of rotavirus infection. *Biochim. Biophys. Acta*, 1528:107–115, 2001.

[64] W. Richard Bowen, Xiaowei Cao, and Paul M. Williams. Use and elucidation of biochemical data in the prediction of the membrane separation of biocolloids. *Proc: Math. Phys. Eng. Sci.*, 455(2933–2955), 1999.

[65] I. Mela, E. Aumaitre, A.M. Williamson, and G.E. Yakubov. Charge reversal by salt-induced aggregation in aqueous lactoferrin solutions. *Colloids Surf. B*, 78(53–66), 2010.

[66] Yuncheng Liang and W. Richard Bowen. Correlation of the gradient diffusion coefficients of human lactoferrin with interparticle interactions validated by photon correlation spectroscopy. *J. Colloid Interface Sci.*, 284(157–166), 2004.

[67] Weimin Li, Björn A. Persson, Maxim Morin, Manja A. Behrens, Mikael Lund, and Malin Zackrisson Oskolkova. Charge-induced patchy attractions between proteins. *J. Phys. Chem. B*, 119:503–508, 2014.

[68] Weimin Li, Björn A. Persson, , Mikael Lund, Johan Bergenholtz, and Malin Zackrisson Oskolkova. Concentration-induced association in a protein system caused by a highly directional attraction. *X*, X:X, 2016.

Acknowledgements

Fifty two months of working in Physical Chemistry Department in Lund has been a great experience in my life. All the people that I have encountered are interesting like waves with different wavelength. Now when I look back on all the waves, a beautiful rainbow is over there. Here I want to put my sincere acknowledgments to the most special people.

First of all I would like to thank my main supervisor Malin Zackrisson Oskolkova, her passion on the projects, curiosity of new discoveries and rigorousness on fine details are the indispensable factors for all the progresses we have accomplished. I also want to thank my co-supervisor Johan Bergenholtz for patiently explaining the complicated integral equation theory to me and my second co-supervisor Ulf Olsson for the suggestions on my study plan and the influence of dealing challenges with happy altitude. On my study plan I also want to give tusen tack to Olle Sörderman, his ability of grasping the core of a subject and saying it with a very clam tone always impresses me.

I would also like to thank Björn Persson and Mikael Lund for showing me how to perform Monte Carlo simulations, giving me the money in pocket metaphor to explain the protein capacitance and correcting me on the calculations of ionic strength.

Thanks to the people who give me all the technical and administrative supports. Busy but always nice and helpful Christopher Hirst, very kind and thoughtful Maria Södergren, lovely and very efficient Ingird, and our secretary Helena Persson and Ganesha responsible Peter Holmqvist. Thanks to Paula Leckius helping me on organizing and printing my thesis.

I have to thank the master students whom I have worked with me during my PhD life. Maxim Morin, thank you for being so supportive on the ALV measurements with me in the late nights, having interesting conversations about almost everything and sharing a lot of fun time with me. Jasenko Gavran and Emil Gustafsson, it has

been a pleasure to work with you and I also got inspired and motivated by you when I was asked the basic work related questions. Paula da Silva, thank you for joining me at the very beginning of my PhD time and working together on the lactoferrin project.

Thanks to my great office mates, Marta Gubitosi, Axel Rüter, Joel Hagman, Luigi Gentile, Christian M. Jessen, Manja A. Behrens, Celen C. Cenker and Sebastian Lages. There are a lot of fun of sharing the office with you, and I appreciate for all the help and supports. Thanks to my best colleagues, here I want mention Thiago Ita, João Martins, Ilaria Idini, Janne-Mieke Meijer, Julien Schmitt, Ricardo Gaspar, Ruiyu Lin, Emelie Nilsson, Dat Pham, Jenny Algotsson, Jerome Crassous, Aleksandra Dabkowska, Najet Mahmoudi, Stefan Kuczera. Thanks for sharing great time with we.

At the end, I want to thank my mom, my dad, my brother and also my tiny Tiny. Thank you for providing a safe place for me to anchor when it is storming outside.

Papers

List of Papers

I **Charge-Induced Patchy Attractions between Proteins**

Weimin Li, Björn A. Persson, Maxim Morin, Manja A. Behrens, Mikael Lund and Malin Zackrisson Oskolkova
The Journal of Physical Chemistry B 2015, 119, 503-508

II **Concentration-Induced Association in a Protein System Caused by a Highly Directional Patch Attraction**

Weimin Li, Björn A. Persson, Mikael Lund, Johan Bergenholtz, and Malin Zackrisson Oskolkova
Journal Physical Chemistry B 2016,

III **Solution Structures Formed by a Protein System Interacting via a Directional Patch Attraction**

Weimin Li, Maxim Morin, Emil Gustafsson, Björn A. Persson, Mikael Lund, and Malin Zackrisson Oskolkova
Submitted

IV **Charge Fluctuations Calculated from Experimental Titration and Monte Carlo Simulations on Large Protein Molecules**

Weimin Li, Björn A. Persson, Mikael Lund, and Malin Zackrisson Oskolkova
Manuscript

All papers are reproduced with permission of their respective publishers.

My contribution to the papers

I I prepared most of the samples and performed most of the experiments and most of the experimental data analysis. I contributed to the writing of the paper.

II I prepared most of the samples and performed all the experiments and all the experimental data analysis. I contributed to the writing of the paper.

III I prepared all the sample and performed part of the experiments. I wrote the first draft.

IV I prepared all the samples, involved in the experiments and data analysis. I wrote the first draft.

Paper 1

This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

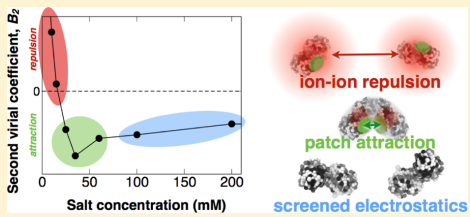
pubs.acs.org/JPCB

Charge-Induced Patchy Attractions between Proteins

Weimin Li,[†] Björn A. Persson,[‡] Maxim Morin,[†] Manja A. Behrens,[†] Mikael Lund,*[,‡] and Malin Zackrisson Oskolkova*[,†]

[†]Division of Physical Chemistry and [‡]Division of Theoretical Chemistry, Lund University, POB 124, 22100 Lund, Sweden

ABSTRACT: Static light scattering (SLS) combined with structure-based Monte Carlo (MC) simulations provide new insights into mechanisms behind anisotropic, attractive protein interactions. A nonmonotonic behavior of the osmotic second virial coefficient as a function of ionic strength is here shown to originate from a few charged amino acids forming an electrostatic attractive patch, highly directional and complementary. Together with Coulombic repulsion, this attractive patch results in two counteracting electrostatic contributions to the interaction free energy which, by operating over different length scales, is manifested in a subtle, salt-induced minimum in the second virial coefficient as observed in both experiment and simulations.



INTRODUCTION

Understanding how protein−protein interactions originate from the level of specific amino acids is of great importance to discern protein function and solution behavior.[1] Due to the inherent complexity, this is a considerable challenge[2−4] because both protein shape and interactions are anisotropic where surface-localized amino acid residues create irregular patterns of neutral, charged, and hydrophobic regions. Overall, these regions or patches contribute to the anisotropic protein interaction energy in a complex manner that ultimately governs if the protein undergoes crystallization, phase separation, and aggregation.[2,3,5,6]

A recent study shows that protein crystallization is often dominated by a limited number of amino acid contacts,[7] giving rise to patchy protein−protein interactions.[5] Describing proteins as patchy, spherical colloids is expected to change the overall phase diagram as compared to that of particles interacting via a centrosymmetric potential.[8−12] It is, however, nontrivial to map the effect of specific amino acid sequences onto such models and more granular approaches seem warranted to study the effect of, i.e., point mutations, solution pH, and salt concentration.

Dissecting which noncovalent, intermolecular interactions that dominate under certain solution conditions is difficult. Experiments probing, i.e., the osmotic second virial coefficient, B_2, does not allow for a separation of the different contributions to the overall interaction. The virial coefficient nontheless remains an important, experimentally accessible quantity that provides a thermodynamic measure of protein−protein interactions.

In this work, we combine static light scattering measurements of B_2 with Metropolis Monte Carlo (MC) simulations from which B_2 is calculated by taking into account the detailed protein structure as well as solution conditions. We show how protein interactions operate in a delicate balance of several both attractive and repulsive contributions, resulting in an electrostatic patchy attraction. The mechanism is recognized both in SLS experiments and in the MC simulations via a nonmonotonous ionic strength dependence of the second virial coefficient at low to moderate ionic strengths, where salt first lowers B_2 and then increases it. To determine which mechanism is responsible for the observed nonmonotonic behavior, we consider also the possibility of Kirkwood fluctuation forces, which are attractive protein−protein interactions due to correlated protonation states in the two proteins.[13,14] This was, however, dismissed by allowing the amino acids to titrate, following a proton titration scheme implemented in the simulations.[14] Left to consider are two electrostatic contributions counteracting each other. The first is a generic, screened Coloumb repulsion due to the net charge of the protein at the studied conditions; the second is a local patch in the charge distribution resulting in a directional, attractive patch.[15] Further, the simulations show that electrostatic interactions alone are insufficient and that there is a nonadditive coupling with van der Waals interactions. Similar salt-screened attractions appearing at low-to-intermediate ionic strengths have previously been observed in several experimental studies of proteins.[6,16,17] Also, more recently, a nonmonotonic behavior of B_2 for monoclonal antibodies was found[18] and the authors indeed argued for a similar mechanism. This body of data points to the possibility of electrostatic attractive patchiness, with a high complementarity, which if present, would appear at low ionic strength and in the neighborhood of the isoelectric point.

METHODS

Sample Preparation. We use bovine lactoferrin[19,20] (>96% purity, Morinaga Milk Industry Co., Ltd., Japan), a globular

Received: December 2, 2014
Revised: December 10, 2014
Published: December 10, 2014

ACS Publications © 2014 American Chemical Society

DOI: 10.1021/jp512027j
J. Phys. Chem. B 2015, 119, 503−508

milk protein with a molecular weight of 80 kDa.[21] Lactoferrin resembles the shape of a dumbbell, as can be seen in Figure 1,

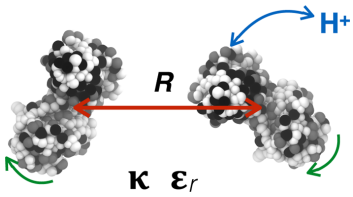


Figure 1. Monte Carlo model of two lactoferrin molecules built from collections of amino acid beads that can be neutral (white spheres), cationic (black spheres), or anionic (gray spheres). Solvent and salt particles are treated implicitly by the relative dielectric constant, ϵ_r, and the inverse Debye screening length, κ. During thermal averaging, the proteins translate (red), rotate (green), and fluctuate (blue) according to solution pH and intermolecular interactions.[14]

with the half-axis equal to 4.7 and 2.6 nm.[22] The isoelectric point (pI) is estimated to be 9.4 from titration simulations,[15] slightly higher than the experimental value of 9.[23] In this study we perform our experiments at pH $\leq$ 7. Stock solutions of lactoferrin were prepared at a concentration of 2 mg/mL by dissolving in NaOAc buffer, I = 5 mM, at pH 5.5 $\pm$ 0.02. We use the Henderson–Hasselbalch equation to account for the contribution to the ionic strength due to changes in dissociation of the buffer species with pH. Also, 1 mM NaN_3 was included to prevent microbial growth; this was also accounted for in the ionic strength.

The chemical nature of salt and buffer components were specifically chosen to be monovalent to avoid ion adsorption, which can dramatically change protein interactions.[24] A stock solution was equilibrated for at least 48 h at room temperature to allow for sufficient solubilization, after which an extensive filtration (centrifugal filters, Millipore) procedure was performed to remove possible contaminants and aggregates. Satisfactory buffer exchange was reached when the filtrate reached the desired target pH ($\pm$0.02). We stress the importance to remove contaminants and insoluble aggregates, which are present in the commercial powder, to obtain experimental B_2 values in close agreement with values from simulations. At low pH we used NaOAc buffer, and at higher pH, Tris buffer, both chosen to produce only monovalent buffer components. Overall, this procedure results in monomeric protein solutions that were monitored by dynamic light scattering (DLS) prior to and after each measurement. DLS allows for extraction of the hydrodynamic radius r_H of the solute species from the cumulant expansion of the autocorrelation function and by the CONTIN analysis[25,26] where a convolution of the distribution was consistent with the presence of a single protein species of same size. Throughout this study, samples contained a hydrodynamic radius of approximately 4 nm in good agreement with literature.[21] A constant temperature of 25 $\pm$ 0.1 °C was used throughout all experimental measurements. Determination of the protein concentration was done by UV absorption spectroscopy; the absorption coefficient was determined to be 1.224 $\pm$ 0.004 cm^2/mg using only monomeric samples, monitored by DLS, to avoid incorrect concentration determination due to contributions from light scattering of aggregates. The static and dynamic light scattering experiments were performed using an ALV 5000F CGS-8F goniometer (ALV, Germany) and correlator equipped with a He–Ne laser diode (Spectra Physics, 127 V/50 mW), operating at a wavelength λ = 632.8 nm.

Experimental Determination of Second Virial Coefficients. Static light scattering allows for determination of the molecular weight M_w and the second virial coefficient B_2 of solutions, which is the property of interest here. B_2 is an important property of the overall interactions where a positive value indicates overall repulsive protein interactions and a negative value signals overall attractive interactions. It is determined by using the relationship between the Rayleigh ratio, R_θ (m^{-1}), and the mass concentration of the protein C (kg/m^3), referred to as a Debye plot

$$\frac{KC}{R_\theta} = \frac{1}{M_w} + \frac{2N_A B_2}{M_w^{\,2}} C \tag{1}$$

where

$$K = \frac{4\pi^2 n_0^{\,2} (dn/dC)^2}{N_A \lambda^4} \tag{2}$$

In eq 1 and eq 2 N_A is Avogadro's number, n_0 is the refractive index of the buffer solution, λ (m) is the wavelength, and dn/dC (m^3/kg) is the increment in the refractive index with respect to increasing protein concentration. In all measurements, static and dynamic light scattering experiments were performed in parallel, at a fixed angle of 90°. However, angular scans were performed initially to ensure no angular dependence. The refractive indices of the buffer solutions, n_0, were measured on an Abbe refractometer. For the refractive index increment, dn/dC, a standard literature value for globular proteins (0.000 186 m^3/kg)[27] was used, which returned molecular weights in close agreement with the literature value of 80 kDa.[21] The same molecular weight was used to determine the B_2 values from the slope in the Debye plots.

Monte Carlo Simulations. Metropolis Monte Carlo (MC) computer simulations[28] were used to compute the Helmholtz interaction free energy between two rigid lactoferrin molecules in an aqueous salt solution. Amino acids are represented by spheres located at the residue mass center according to the crystal structure (PDB 1BLF), Figure 1. The canonical (NVT) ensemble is sampled using molecular translational and rotational MC moves as well as proton swap moves on titratable (acid and basic) sites to account for charge fluctuations; i.e., the simulations are performed at constant pH.[14] Production runs consist of at least 10^8 configurations, preceded by 10 times shorter equilibration runs. Solvent and salt are treated at the Debye–Hückel level whereas exchange repulsion and short-ranged attractions, such as van der Waals (vdW), are described by a Lennard-Jones potential, yielding the effective system energy

$$\beta u = \sum_{i\neq j}^{N} \lambda_B z_i z_j \exp(-\kappa r_{ij})/r_{ij} + 4\beta\epsilon[(\sigma_{ij}/r_{ij})^{12} - (\sigma_{ij}/r_{ij})^6] + \sum_{i}^{N_p} (\mathrm{pH} - \mathrm{p}K_{a,i}) \ln 10 \tag{3}$$

where N and N_p run over all residues and protonated sites, respectively. κ^{-1} is the Debye length, λ_B = 0.7 nm (water, 298 K) is the Bjerrum length, z is particle valency (−1, 0, +1), $\beta\epsilon$ = 0.05 is the vdW strength,[15,29] σ_{ij} is the arithmetic mean particle diameter, r_{ij} is inter-residue distances, $\beta^{-1} = k_B T$ is the thermal

DOI: 10.1021/jp512027j
J. Phys. Chem. B 2015, 119, 503–508

energy, and $pK_{a,i}$ are the unperturbed acid dissociation constants for titratable amino acids. The magnitude of the short-ranged interaction, ϵ, has been chosen such that its effective contribution is the same as in a previously investigated model,[29] and we have made no attempts to fit it to the current case. The angularly averaged pair distribution function, $g(R) = \exp(-\beta w(R))$, is obtained by sampling the histogram of protein−protein mass center separations, R, and is subsequently integrated to give the osmotic second virial coefficient

$$B_2 = -2\pi \int_0^\infty (g(R) - 1)R^2 \, dR \tag{4}$$

The precision of the interaction free energy, $w(R)$, is $\pm 0.05\ k_BT$ or better for all relevant protein separations. To minimize noise amplification for large R when eq 4 is integrated, the tail of the sampled $g(R)$ is substituted with a smooth function at long separations where only Coulombic repulsion persists. The functional form is the linearized Poisson−Boltzmann result for two charged, macro-ions with charges smeared over their surfaces, $w(R)^{ion} = \lambda_B Z^2 \sinh^2(\kappa a) e^{-\kappa R}/R(\kappa a)^2$. Here Z is the average protein net charge and a an approximate protein radius, obtained by fitting to the $g(R)$ tail.

Finally, the inverse Debye length is calculated according to $\kappa = (8\pi\lambda_B I)^{1/2}$ where the ionic strength, $I = {}^1/_2 \sum_i \rho_i z_i^2$, is summed over all ion types of density ρ_i and valency z_i. Thus, for a 1:1 salt such as NaCl, I simply equals the salt concentration.

RESULTS AND DISCUSSION

This work was prompted by previous structure-based MC simulations on lactoferrin, which predicted the presence of an electrostatic anisotropic attraction.[15] The attraction came from a few localized amino acid residues that gave rise to an unusually distinct and narrow minimum in the interaction free energy as a function of protein−protein separation, $\beta w(R) = -\ln g(R)$. This is shown in Figure 2, and the minimum was found to deepen upon increasing pH toward the isoelectric point.[15] The involved amino acids identified in the simulations[15] are charged, pointing to a mechanism of electrostatic origin. Indeed, electrostatics was found to lock the two proteins into a few orientations, i.e., a regio-specific interaction. Using the angularly averaged protein−protein potential of mean force, we calculate B_2 as a function of pH, as shown in the inset to Figure 2. The minimum in $w(R)$ gives a negative contribution to B_2, which decreases with increasing pH. The simulations thus guide to the narrow set of conditions where to expect the interaction energy to display the minimum and electrostatic anisotropic attractions. They are expected at low ionic strength, near the isoelectric point which corresponds to B_2 values close to zero, as shown in the inset to Figure 2.

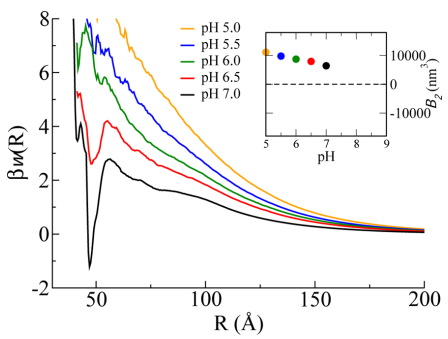


Figure 2. Simulated angularly averaged protein−protein potential of mean force, $\beta w(R)$, as a function of protein−protein mass center separation, R, and at different pHs. The salt concentration is 5 mM for all pHs. The inset shows the corresponding virial coefficients, B_2, cf. eq 4.

The experimental conditions corresponding to the simulations were investigated by performing static light scattering experiments as a function of protein concentration. Second virial coefficients were obtained from eq 1 (Methods) in a Debye representation of the scattered intensity, which is shown in Figure 3. Here K is an optical constant, C is the protein concentration and R_θ is the scattered light expressed as the excess Rayleigh ratio. The ionic strength is fixed at 5 mM while pH is varied. As shown in Figure 3 at pH 5.5, the charged proteins repel each other, resulting in a positive slope. This corresponds to a positive second virial coefficient, shown in the inset of Figure 3. By gradually increasing pH, while keeping the ionic strength fixed at 5 mM, we find the attractive and repulsive contributions to B_2 to balance out, at pH 7, resulting in a slightly negative B_2. Note that at pH 7 we switch to Tris buffer to maintain stable buffering capacity.

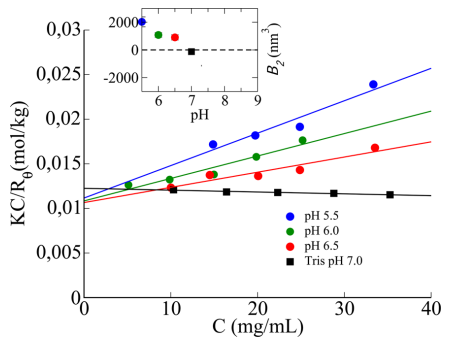


Figure 3. Debye plot at different pHs and constant ionic strength (5 mM) in acetate and Tris buffer (pH 7). The lines are weighted linear least-squares fits to the data. Error bars are included but lay within the symbols. The inset shows corresponding virial coefficients with error bars.

We next investigate how ionic strength influences the interactions, starting at 5 mM and pH 7. These are the ionic strength and pH conditions where the slope in Figure 3 turned negative, and where to expect the electrostatic patch attraction to be at its strongest, providing it is a real, measurable effect. The experimental results from static light scattering are shown in Figure 4. We observe how the slope, and correspondingly B_2, first decreases when the ionic strength is increased. This behavior is expected because electrostatic screening weakens the repulsion between proteins, resulting in a lowering of the slope and thus B_2. However, as more salt is added, surprisingly, the slope starts to *increase* at around 25 mM.

The corresponding second virial coefficients extracted from the experiments in Figure 4 are shown in the lower panel in Figure 5 as a function of ionic strength. The nonmonotonic dependence of the slope results in a B_2 minimum as a function of ionic strength.

DOI: 10.1021/jp512027j
J. Phys. Chem. B 2015, 119, 503−508

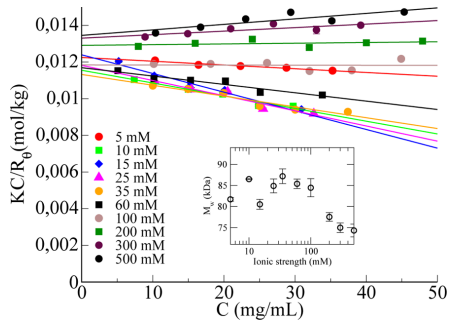


Figure 4. Debye plot at different ionic strengths at pH 7 in Tris buffer. The lines are weighted linear least-squares fits to the data. Error bars are included but lay within the symbols. The inset shows corresponding molecular weights as a function of ionic strength with error bars.

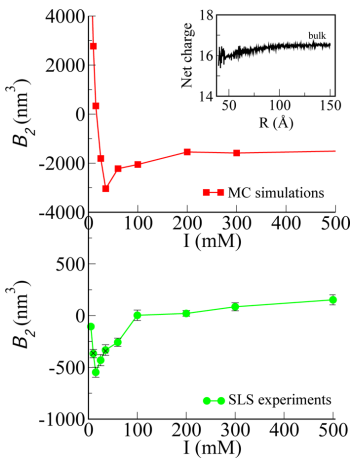


Figure 5. Calculated (top) and measured (bottom) second virial coefficient as a function of ionic strength, at pH 7. In the measurements, the ionic strengths 10 and 35 mM (shown as crosses) were repeated from additional concentration series and batches. The inset shows the calculated overall protein charge as a function of protein−protein mass center separation.

From the isolated experimental data, it would be difficult to determine which mechanism is responsible for the minimum. We also have B_2 values calculated from simulations at matching conditions; see the upper panel of Figure 5. Here, we observe exactly the same minimum in the B_2 values as a function of salt concentration. Importantly, we can directly link the minimum in B_2 to the minimum in the angularly averaged free energy, shown in Figure 2. The effect of salt on the angularly averaged potential of mean force is shown in Figure 6. It can be seen that addition of salt initially acts to expose the attractive minimum as the repulsive barrier at longer separations decreases. Additional salt screens both the repulsive barrier and the free energy minimum. The observed mechanism comes from two opposing, electrostatic contributions: a short-ranged, attractive patch and a long-ranged, screened Coulomb repulsion. Salt

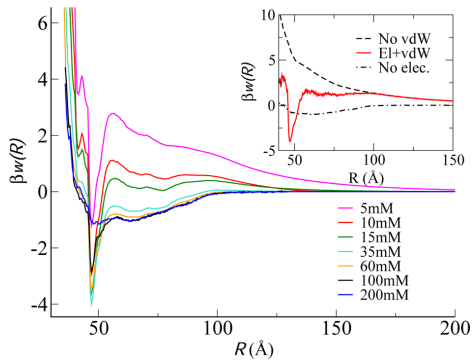


Figure 6. Angularly averaged protein−protein potential of mean force, $\beta w(R)$, at different salt concentrations, pH 7. The inset shows $\beta w(R)$ when either electrostatic or short-ranged, attractive ("vdW") interactions are artificially disabled at 5 mM and comparable conditions, pH 8.

initially screens the Coloumb repulsion, thus strengthening the effect of the attractive patch. As the salt concentration is further increased (decreasing Debye length), the attractive interaction is ultimately screened as well. This balance is manifested as a minimum in the second virial coefficient as a function of ionic strength, as observed in both SLS experiments and MC simulations.

Although the qualitative agreement between measured and simulated B_2 is excellent, the absolute values differ. This is expected from a coarse-grained model and we have deliberately made no attempts to fit the only free parameter in our model, the Lennard-Jones strength, ϵ, to the measurements. As also discussed in the Methodology, this single value is taken from a different protein system and essentially encompassed the net effect of all short-ranged, nonelectrostatic interactions. For this reason it is unlikely to be universal and as hinted at in a recent combined SAXS/simulation study of protein solution structure,[30] our choice of ϵ may be slightly too attractive but nonetheless close enough to capture complex, qualitative trends in arbitrary protein systems.

The inset to Figure 6 shows a study of the free energy, βw, as a function of the protein−protein separation. The βw, shown in red, where both electrostatic and short-ranged attractions are included, has a distinct and narrow minimum, corresponding to the tightly bound and stereospecific configuration.[15] Artificially turning off either electrostatic or short-ranged interactions suppresses the free energy minimum, showing that both interactions are needed due to nonlinear coupling of the Boltzmann weight.

An alternative explanation for the observed salt behavior of B_2 may stem from proton fluctuations. That is, a mechanism whereby the ionization states of the two proteins close to each other become correlated at low ionic strengths.[13,14] To assert if this mechanism is operating, we have analyzed the variation in protein charge as a function of separation. As shown in the inset of Figure 5, the net charge varies only weakly as the proteins approach. The slight decrease at shorter distance is a mutual response of the two approaching charge distributions. Hence, we conclude that fluctuation forces are unimportant for the present system as the average net charge of the proteins are only weakly perturbed.

DOI: 10.1021/jp512027j
J. Phys. Chem. B 2015, 119, 503−508

Other attractive protein interactions, signaled by a minimum in the second virial coefficient, have been reported[31,32] albeit at significantly higher salt concentrations. We argue that these observations cannot originate from the same electrostatic mechanism found here, because the ionic strength at which the attraction is located is too high to survive the electrostatic screening.

CONCLUSIONS

We have used static light scattering and computer simulations to obtain osmotic second virial coefficients, B_2, for lactoferrin as a function of pH and salt concentration. The simulations were used to guide the experiments to a narrow set of conditions where B_2 first decreases with added salt, then increases to reach a plateau. We show that this nonmonotonous behavior—observed unambiguously in experiment and simulations—originates from a high charge complementarity on the binding interface between two proteins. This creates a short-ranged attraction that is competing with a long-ranged repulsion due to the protein net charges. Addition of salt modulates the electrostatic screening length, whereby the balance between long- and short-ranged electrostatic interactions can be finely tuned. This study provides insight into how anisotropic protein attractions come about in a nonlinear combination of several electrostatic as well as other short-ranged forces.

AUTHOR INFORMATION

Corresponding Authors

*M. Lund. E-mail: mikael.lund@teokem.lu.se. Phone: +46 (0) 46 222 1428. Fax: +46 (0)46 222 8648.

*M. Zackrisson Oskolkova. E-mail: malin.zackrisson@fkem1.lu.se. Phone: +46 (0)46 222 8185. Fax: +46 (0)46 222 4413.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The authors wish to thank the OMM Linneaus center in Lund; the Swedish Research Council; the Swedish Foundation for Strategic Research; the Crafoord Foundation; the Royal Swedish Academy of Sciences; LUNARC in Lund for computational resources; and Morinaga Milk Industry Co., Ltd, Japan for kindly providing us with the protein.

REFERENCES

(1) McManus, J.; Lomakin, A.; Ogun, O.; Pande, A.; Basan, M.; Pande, J.; Benedek, G. B. Altered Phase Diagram Due to a Single Point Mutation in Human γD-Crystallin. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 16856−16861.

(2) Mezzenga, R.; Fischer, P. The Self-Assembly, Aggregation and Phase Transition of Food Protein Systems in One, Two and Three Dimensions. *Rep. Prog. Phys.* **2013**, *76*, 046601.

(3) Piazza, R. Protein Interactions and Association: an Open Challenge for Colloid Science. *Curr. Opin. Colloid Interface Sci.* **2004**, *8*, 515−522.

(4) Piazza, R. Interactions and Phase Transitions in Protein Solutions. *Curr. Opin. Colloid Interface Sci.* **2000**, *5*, 38−43.

(5) Fusco, D.; Headd, J. J.; de Simone, A.; Wang, J.; Charbonneau, P. Characterizing Protein Crystal Contacts and Their Role in Crystallization: Rubredoxin as a Case Study. *Soft Matter* **2014**, *10*, 290−302.

(6) Dumetz, A. C.; Snellinger-O'Brien, A. M.; Kaler, E. W.; Lenhoff, A. M. Patterns of Protein-Protein Interactions in Salt Solutions and Implications for Protein Crystallization. *Protein Sci.* **2007**, *16*, 1867−1877.

(7) Price, W. N.; Chen, Y.; Handelman, S. K.; Neely, H.; Manor, P.; Karlin, R.; Nair, R.; Liu, J.; Baran, M.; Everett, J.; et al. Understanding the Physical Properties that Control Protein Crystallization by Analysis of Large-Scale Experimental Data. *Nat. Biotechnol.* **2009**, *27*, 51−57.

(8) Wertheim, M. S. Fluids with Highly Directional Forces. I. Statistical Thermodynamics. *J. Stat. Phys.* **1984**, *35*, 19−34.

(9) Kern, N.; Frenkel, D. Fluid-Fluid Coexistence in Colloidal Systems with Short-Ranged Strongly Directional Attraction. *J. Chem. Phys.* **2003**, *118*, 9882−9889.

(10) Bianchi, E.; Largo, J.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Phase Diagram of Patchy Colloids: Towards Empty Liquids. *Phys. Rev. Lett.* **2006**, *97*, 168301.

(11) Bianchi, E.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Theoretical and Numerical Study of the Phase Diagram of Patchy Colloids: Ordered and Disordered Patch Arrangements. *J. Chem. Phys.* **2008**, *128*, 144504.

(12) Munao, G.; Preisler, Z.; Vissers, T.; Smallenberg, F.; Sciortino, F. Cluster Formation in One-Patch Colloids: Low Coverage Results. *Soft Matter* **2013**, *9*, 2652−2661.

(13) Kirkwood, J. G.; Shumaker, J. B. Forces Between Protein Molecules in Solution Arising from Fluctuations in Proton Charge and Configuration. *Proc. Natl. Acad. Sci. U. S. A.* **1952**, *38*, 863−871.

(14) Lund, M.; Jönsson, B. Charge Regulation in Biomolecular Solution. *Q. Rev. Biophys.* **2013**, *46*, 265−281.

(15) Persson, B. A.; Lund, M.; Forsman, J.; Chatterton, D. E. W.; Åkesson, T. Molecular Evidence of Stereo-Specific Lactoferrin Dimers in Solution. *J. Biophys. Chem.* **2010**, *151*, 187−189.

(16) Neal, B. L.; Asthagiri, D.; Velev, O. D.; Lenhoff, A. M.; Kaler, E. W. Why Is the Osmotic Second Virial Coefficient Related to Protein Crystallization? *J. Cryst. Growth* **1999**, *196*, 377−387.

(17) Dumetz, A. C.; Chockla, A. M.; Kaler, E. W.; Lenhoff, A. M. Effects of pH on Protein-Protein Interactions and Implications for Protein Phase Behavior. *Biochim. Biophys. Acta* **2008**, *1784*, 600−610.

(18) Roberts, R.; Keeling, R.; Tracka, M.; van der Walle, C. F.; Uddin, S.; Warwicker, J.; Curtis, R. The Role of Electrostatics in Protein-Protein Interactions of a Monoclonal Antibody. *Mol. Pharmaceutics* **2014**, *11*, 2475−2489.

(19) Baker, E. N.; Baker, H. M. A structural Framework for Understanding the Multifunctional Character of Lactoferrin. *Biochimie* **2009**, *91*, 3−10.

(20) Brock, J. H. The Physiology of Lactoferrin. *Biochem. Cell Biol.* **2002**, *80*, 1−6.

(21) Lönnerdal, B.; Iyer, S. Lactoferrin Molecular Structure and Biological Function. *Annu. Rev. Nutr.* **1995**, *15*, 93−110.

(22) Babina, S. E.; Tuzikov, F. V.; Tuzikova, N. A.; Buneva, V. N.; Nevinskii, G. A. Effect of Nucleotides on the Oligomeric State of Human Lactoferrin. *Mol. Biol. (Moscow, Russ. Fed., Engl. Ed.)* **2006**, *40*, 121−131.

(23) Superti, F.; Siciliano, R.; Rega, B.; Giansanti, F.; Valenti, P.; Antonini, G. Involvement of Bovine Lactoferrin Metal Saturation, Silica Acid and Protein Fragments in the Inhibition of Rotavirus Infection. *Biochim. Biophys. Acta* **2001**, *1528*, 107−115.

(24) Roosen-Runge, F.; Heck, B. S.; Zhang, F.; Kohlbacher, O.; Schreiber, F. Interplay of pH and Binding of Multivalent Metal Ions: Charge Inversion and Reentrant Condensation in Protein Solutions. *J. Phys. Chem. B* **2013**, *117*, 5777−5787.

(25) Provencher, S. W. A. Constrained Regularization Method for Inverting Data Represented by Linear Algebraic or Integral Equations. *Comput. Phys. Commun.* **1982**, *27*, 213−227.

(26) Provencher, S. W. CONTIN: A General Purpose Constrained Regularization Program for Inverting Noisy Linear Algebraic and Integral Equations. *Comput. Phys. Commun.* **1982**, *27*, 229−242.

(27) Zhao, H.; Brown, P. H.; Schuck, P. On the Distribution of Protein Refractive Index Increments. *Biophys. J.* **2011**, *100*, 2309−2317.

(28) Metropolis, N. A.; Rosenbluth, A. W.; Teller, M. N. R. A.; Teller, E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.* **1953**, *21*, 1087−1092.

DOI: 10.1021/jp512027j
J. Phys. Chem. B 2015, 119, 503−508

(29) Lund, M.; Jönsson, B. A Mesoscopic Model for Protein-Protein Interactions in Solution. *Biophys. J.* **2003**, *85*, 2940−2947.
(30) Kaieda, S.; Lund, M.; Plivelic, T. S.; Halle, B. Weak Self-Interactions of Globular Proteins Studied by Small-Angle X-ray Scattering and Structure-Based Modeling. *J. Phys. Chem. B* **2014**, *118*, 10111−10119.
(31) Petsev, D. N.; Vekilov, P. G. Evidence for Non-DLVO Hydration Interactions in Solutions of the Protein Apoferritin. *Phys. Rev. Lett.* **2000**, *84*, 1339.
(32) Piazza, R.; Iacopini, S.; Galliano, M. BLGA Protein Solutions at High Ionic Strength: Vanishing Attractive Interactions and "Frustrated" Aggregation. *Europhys. Lett.* **2002**, *59*, 149−154.

DOI: 10.1021/jp512027j
J. Phys. Chem. B 2015, 119, 503−508

Paper II

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

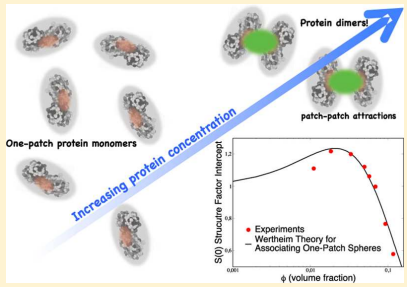
Concentration-Induced Association in a Protein System Caused by a Highly Directional Patch Attraction

Weimin Li,[†] Björn A. Persson,[‡] Mikael Lund,[‡] Johan Bergenholtz,[§] and Malin Zackrisson Oskolkova*[,†]

[†]Division of Physical Chemistry and [‡]Division of Theoretical Chemistry, Lund University, P.O. Box 124, SE-22100 Lund, Sweden

[§]Department of Chemistry and Molecular Biology, University of Gothenburg, SE-412 96 Göteborg, Sweden

ABSTRACT: Self-association of the protein lactoferrin is studied in solution using small-angle X-ray scattering techniques. Effective static structure factors have been shown to exhibit either a monotonic or a nonmonotonic dependence on protein concentration in the small wavevector limit, depending on salt concentration. The behavior correlates with a nonmonotonic dependence of the second virial coefficient on salt concentration, such that a maximum appears in the structure factor at a low protein concentration when the second virial coefficient is negative and close to a minimum. The results are interpreted in terms of an integral equation theory with explicit dimers, formulated by Wertheim, which provides a consistent framework able to explain the behavior in terms of a monomer−dimer equilibrium that appears because of a highly directional patch attraction. Short attraction ranges preclude trimer formation, which explains why the protein system behaves as if it were subject to a concentration-dependent isotropic protein−protein attraction. Superimposing an isotropic interaction, comprising screened Coulomb repulsion and van der Waals attraction, on the patch attraction allows for a semiquantitative modeling of the complete transition pathway from monomers in the dilute limit to monomer−dimer systems at somewhat higher protein concentrations.



INTRODUCTION

Understanding how proteins behave in solution requires detailed knowledge about the structural arrangement of their amino acids, particularly how they are organized near the protein surface. In addition, one needs a way to translate how this structure affects protein interactions. Protein interactions are generally complex because of surface chemical heterogeneities associated with different surface-located amino acid residues. This is expected to impart an orientational dependence to the overall interaction, in addition to the orientational dependence due to the nonspherical shape of the proteins. Simplified models of proteins and their interactions are attractive from computational considerations and may suffice for gaining qualitative insight into solution behavior. This type of approach has resulted in a great improvement in the understanding of how patchy or anisotropic attractions modify quite drastically the phase behavior and other properties as compared to those of systems governed by isotropic attractive potentials. For example, there is a shift in the liquid−liquid binodal and critical point toward lower volume fractions depending on the number of attractive sites.[1−3] In the case of one-patch spheres, computer simulations suggest that new types of structures should appear, such as wires and lamellar phases.[4] Protein interactions are often used as the basis for examining such simplified models with patchy or anisotropic interactions, but little is known about how models of patchiness should be designed to capture protein interactions in a more realistic manner. There is also a lack of guidelines when it comes to how one can recognize the effect of patchiness and interaction anisotropy on measurements of physicochemical properties of actual protein solutions.

Lactoferrin is a protein that, under the right conditions in terms of pH and ionic strength, exhibits a particularly strong patch−patch attraction, which may lock nearby proteins into a tightly bound and stereospecific configuration.[5] This effect was initially predicted in structure-based, two-protein Monte Carlo simulations of the potential of the mean force of lactoferrin,[5] the results of which have recently been corroborated experimentally by static light scattering (SLS) experiments.[6] Because of the presence of a few localized residues, the interaction free energy reaches a distinct and narrow minimum upon increasing the pH on the acidic side of the isoelectric point. These amino acids are charged, which results in an attractive electrostatic patch that is modulated not only by ionic strength but also requires van der Waals attractions. This interaction manifests itself as a negative-valued second virial coefficient, B_2, at NaCl concentrations in the range 10−60 mM.[6] However, if the ionic strength is either increased or decreased, then B_2 increases to reach positive values, indicating that the protein interaction is dominated by excluded-volume or repulsive electrostatic interactions away from the narrow range of ionic strength centered at around 15 mM.[6]

Received: July 10, 2016
Revised: July 20, 2016

ACS Publications © XXXX American Chemical Society

DOI: 10.1021/acs.jpcb.6b06873
J. Phys. Chem. B XXXX, XXX, XXX−XXX

This minimum in B_2 is not observed only for lactoferrin. Other proteins have been found to show a qualitatively similar nonmonotonic dependence of B_2 on salt concentration.[7,8] In these cases, it is observed at much higher salt concentrations, at which electrostatic effects are far less pronounced, which points to other mechanisms at play. However, some monoclonal antibodies exhibit this effect at reduced salt concentrations,[9] which has similarly been attributed to anisotropic interactions of electrostatic origin. In addition, given that self-association of proteins into dimers and high-order oligomers is so widespread, at least in nature,[10] it is possible that many more proteins may be subject to this type of mechanism. However, studies will have to be conducted in the future to examine whether self-association of this sort can also occur under physiologically relevant conditions.

In this work, lactoferrin solutions are examined using SLS and small-angle X-ray scattering (SAXS) as the protein concentration is increased away from the dilute limit studied previously.[6] The aim is to investigate how directional attraction impacts the microstructure and to see more generally how one can distinguish this patch interaction from the well-known centrosymmetric case. To this end, the lactoferrin system is studied both under conditions (35 mM) in which the patch attraction is active and dominant and under conditions (200 mM) in which it is screened, leading to B_2 values close to zero. On the one hand, results are analyzed in terms of a well-known model for short-range isotropic attractions, widely employed to model the behavior of protein solutions,[9,11−14] and on the other hand, they are analyzed in terms of an integral equation theory for associating systems under the influence of directional attractions, formulated by Wertheim.[15−17]

MATERIALS AND METHODS

Bovine lactoferrin (>96% purity), with a molecular weight of 80 kDa,[18] was purchased in the powder form from Morinaga Milk Industry Co. Ltd., Japan. Monomeric solutions were prepared by dissolving the protein powder in NaOAc buffer of 5 mM ionic strength and pH of 5.5 for a minimum of 48 h at 25 °C. Purification and adjustment of buffer conditions to those of tris buffer at pH 7 proceeded by centrifugal filtration (100 kDa molecular weight cut-off (MWCO); Millipore Amicon), as did concentration of dilute solutions (50 kDa MWCO; Millipore Amicon). The buffers used contained monovalent electrolyte to avoid ion adsorption. Conversion from mass concentration (mg/mL) to volume fractions was done using a protein density of 1.3824 g/mL. A constant temperature of 25 ± 0.1 °C was maintained in all experimental measurements.

The purified solutions were monitored by dynamic light scattering (DLS) to ensure that the samples exhibited only a single narrow size distribution centered at 8 nm (diameter), in agreement with the literature value.[18] Protein concentrations were determined spectrophotometrically using an extinction coefficient of 1.224 cm^2/mg. SLS and DLS data were collected using an ALV 5000F CGS-8F goniometer (ALV GmbH, Langen, Germany) at a wavelength of 632.8 nm. Angular scans were performed from 50 to 130° in 5° incremental steps. All measurements were performed at 25 °C, and toluene was used as a reference to obtain data on an absolute scale.

SAXS spectra were recorded on an automated pinhole system (Ganesha, JJ X-ray System ApS) equipped with a motorized two-dimensional Pilatus detector (Dectris Ltd., Switzerland), at a wavelength of 1.5408 Å. Data were collected at three different sample-to-detector distances (1491, 480, and 180 mm) to give a scattering vector range of 0.003 $Å^{-1} \leq q \leq$ 1.4 $Å^{-1}$. Raw SAXS images were processed using the SAXSGUI software for radial averaging and background subtraction. All SAXS measurements were performed at 25 °C. A measurement of water was used to bring data onto an absolute scale.[19] To obtain effective structure factors, the scattering values from low-concentration samples ($C \approx 5$ mg/mL) were taken as form factors.

INTEGRAL EQUATION THEORY

We follow Wertheim[17] in adopting a model consisting of an isotropic interaction, $\phi_{iso}(r)$, complemented by a site−site attraction of the square-well form, as illustrated in Figure 1. The

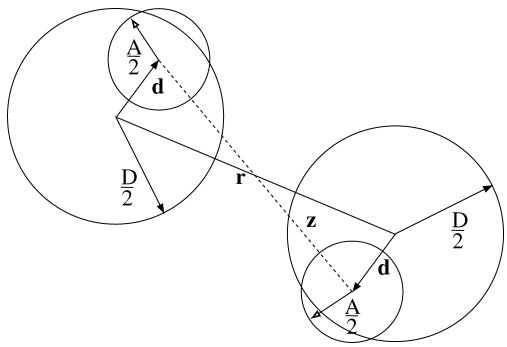


Figure 1. Two like-sized spheres of diameter D separated by center−center distance r. Each sphere carries an attraction site a distance of $d = |\mathbf{d}|$ from the center, with the attraction range, A, measured along the site−site separation vector, $\mathbf{z}$. When the smaller spheres centered on the attractive sites overlap, an attraction of magnitude ϵ is generated.

orientationally averaged Mayer function involving the site−site attraction is

$$\bar{f}_A(r) = (e^{\beta\epsilon} - 1)\frac{(A + 2d - r)^2(2A + 2d + r)}{24d^2 r} \quad (1)$$

where A and ϵ are the range and depth, respectively, of the square-well attraction, d is the distance from the sphere center to the attraction site, r is the center−center separation, and $\beta = (k_B T)^{-1}$ is the inverse thermal energy. As noted by Wertheim,[17] for site−site attraction ranges that satisfy (with D being the sphere diameter)

$$0 < A + 2d - D < (2 - \sqrt{3})d \leq 0.134D \quad (2)$$

the attraction site can only accommodate at most one other particle. As a consequence of the restriction to dimer formation, the closure and Ornstein−Zernike equations can be orientationally averaged to read

$$h_{ij}(r) = c_{ij} + \sum_{\substack{k=0,1 \\ l=0,1}} \int c_{ik}(r')\rho_{kl}h_{lj}(|\mathbf{r} - \mathbf{r}'|)\, d\mathbf{r}' \quad (3)$$

where $\rho_{00} = \rho$, $\rho_{01} = \rho_{10} = \rho_0$, and $\rho_{11} = 0$ in terms of the total number density, ρ, and the monomer number density, ρ_0. The monomer number density is obtained by enforcing the following relation

DOI: 10.1021/acs.jpcb.6b06873
J. Phys. Chem. B XXXX, XXX, XXX−XXX

$$\rho_0 = -\frac{1}{2X} + \sqrt{\frac{1}{4X^2} - \frac{\rho}{X}} \quad (4)$$

$$X = 4\pi \int_D^{2d+A} (h_{00}(r) + 1)\overline{f}_A(r) r^2 \, dr \quad (5)$$

A Percus–Yevick-like closure (with $g_{ij}(r) = h_{ij}(r) + \delta_{i0}\delta_{j0}$)

$$e^{-\beta\phi_{iso}(r)} c_{ij}(r) = (e^{-\beta\phi_{iso}(r)} - 1) g_{ij}(r) + g_{00}(r)\overline{f}_A(r)\delta_{i1}\delta_{j1} \quad (6)$$

and a reference hypernetted chain closure, both devised by Wertheim,[16] were used. These were found to yield results that were indistinguishable in the range of volume fractions investigated, $0 < \phi < 0.15$.

The above equations were reformulated in terms of the continuous function, $\gamma_{ij}(r) = h_{ij}(r) - c_{ij}(r)$, and solved by iteration. Briefly, the iteration was initiated by taking $\gamma_{ij}(r) = 0$ and $g_{00}(r) = 1$, based on which eq 4 was solved to yield an estimate for ρ_0. The $c_{ij}(r)$ functions were then determined from the closure and numerically Fourier transformed. The Ornstein–Zernike equations were solved for the Fourier-transformed $\gamma_{ij}(r)$ functions, which were Fourier back-transformed to serve as input in the second iteration step. An algorithm to improve the convergence was also employed,[20] and discontinuities in the $c_{ij}(r)$ functions were handled in the Fourier transformation step as described by Klein and D'Aguanno.[21] The model structure factors were determined from the Fourier transform of the complete radial distribution function, given as $g(r) = g_{00}(r) + (\rho_0/\rho)(g_{01}(r) + g_{10}(r)) + (\rho_0/\rho)^2 g_{11}(r)$.[22] In typical calculations, the center–center separation was uniformly discretized in 2^{15} increments, with $\Delta r = 0.002D$. In addition, for the sake of comparison, Baxter's adhesive sphere model[23] was also employed.

RESULTS AND DISCUSSION

Figure 2 shows a partial Zimm plot based on SLS data for lactoferrin solutions collected for ionic strengths of 35 and 200 mM. For the former case, the patch attraction is active and dominant, whereas for the latter case, it is inactive. This can be seen from the difference in low-concentration slopes, which provide a measure of the second virial coefficient. The inset of Figure 2 shows B_2 values for a range of ionic strengths, illustrating the nonmonotonic dependence of B_2 on salt concentration studied previously.[6] The arrows in the inset point to the negative-valued B_2 for 35 mM added salt, and the close-to-zero value of B_2 results when 200 mM salt is added. In the following sections, lactoferrin solutions under these two conditions have been studied for concentrations of up to about 150 mg/mL using both SLS and SAXS. Rather than analyzing the scattered intensity directly, we focus on the behavior of the effective structure factor, $S(q) = (\rho P(q))^{-1} I(q)$, which removes some of the effect of the form factor, $P(q)$. Importantly, this manipulation removes the effect of increasing the number density of scatterers, which, together with excluded-volume interactions, causes the intensity at low q to show a maximum as a function of concentration. In other words, the concentration-induced changes in the structure factors investigated here are due to the protein interactions.

Figure 3 shows effective structure factors, obtained from absolute intensities divided by a dilution-limiting measurement in the absence of correlation effects. The lowest-q data were obtained from SLS measurements, and the data to the right of

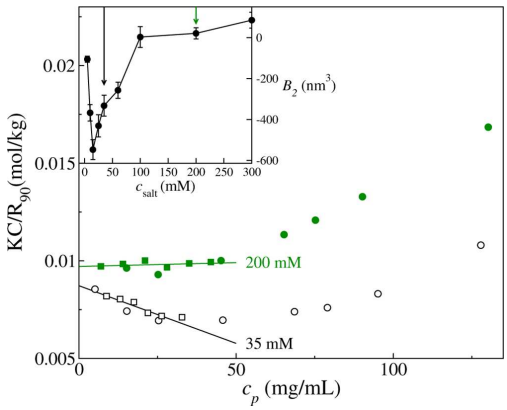


Figure 2. Partial Zimm plot using SLS data for the inverse concentration-normalized Rayleigh ratio, R_{90}, as a function of protein concentration for added salt concentrations of 200 and 35 mM, as labeled. Two sets of data (squares and circles) have been recorded at both salt concentrations. The dilution-limiting slopes determine the second virial coefficient, B_2, given in units of nm^3 as a function of salt concentration in the inset. The arrows in the inset indicate the B_2 values for 35 and 200 mM.

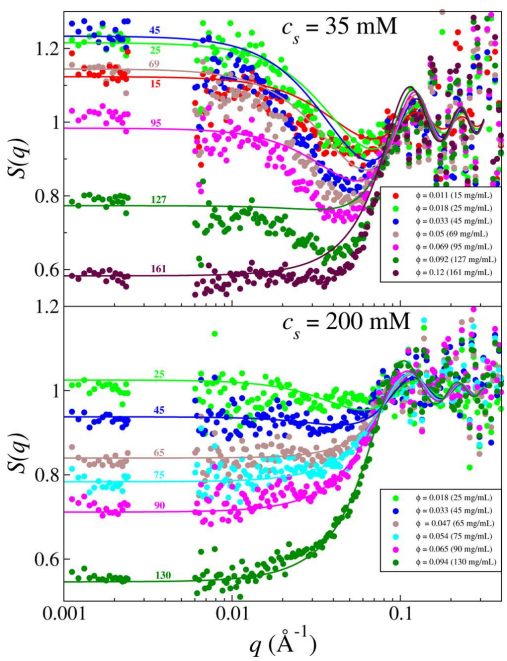


Figure 3. Comparison of Baxter adhesive sphere structure factors (lines) with effective structure factors obtained from SAXS (right of gap) and SLS (left of gap) of lactoferrin at 35 and 200 mM salt, pH 7 (T = 25 °C). The effective protein diameter in the analysis was set to 6 nm to obtain reasonable agreement between the data and model at higher q.

DOI: 10.1021/acs.jpcb.6b06873
J. Phys. Chem. B XXXX, XXX, XXX−XXX

the slight gap in q were from SAXS measurements. These structure factors behave very differently, depending on whether the patch is turned on (35 mM salt) or off (200 mM salt). In the presence of the patch attraction, not only does $S(q)$ shows a characteristic low-q upturn for all except the most concentrated sample, but the low-q scattering also exhibits a nonmonotonic concentration dependence (cf. Figure 3). As the protein concentration is increased away from the dilute limit, the low-q upturn of the structure factor initially becomes more pronounced. As the protein concentration is increased further, the low-q limit of the structure factor begins to decrease, suggesting that the patch attraction gives way to excluded-volume-like interactions as the protein solution becomes more concentrated. The behavior is quite different at the 200 mM salt concentration. As the protein concentration is increased, the scattering at low-q values becomes increasingly more suppressed, as expected for particles interacting via repulsion.

Proceeding along the lines of Fine et al.[12] and Piazza et al.,[13] who adopted a spherically symmetric adhesive interaction, as given by Baxter's adhesive sphere model,[23] to model γ_{II}-crystallin and lysozyme solutions, we attempt to fit the experimental structure factors using this model. In the present case, however, the stickiness parameter, τ, in Baxter's model must be varied freely to get reasonable agreement with the scattering data at either salt concentration. As shown in Figure 3, the adhesive sphere model applied in this way captures the dependence on q and volume fraction reasonably well for 35 mM NaCl and quantitatively for 200 mM NaCl. However, the resulting values for τ in Figure 4 show that the system behaves

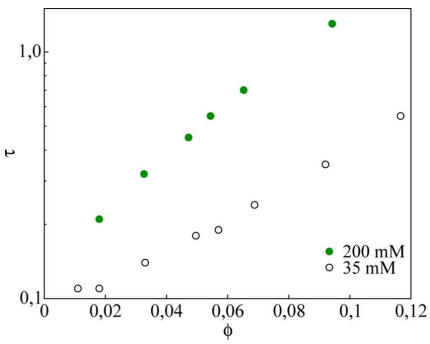


Figure 4. Baxter τ parameter as a function of protein volume fraction and added salt concentration, as extracted from the fitting of the adhesive sphere structure factors in Figure 3.

as if the particles were subject to a strong attraction (low values of τ) at low protein concentrations and a much weaker attraction (higher values of τ) at higher protein concentrations at both salt concentrations studied. In fact, τ values near 0.1, as obtained at the lowest protein concentrations in the presence of a strong patch attraction, cause a fluid−fluid phase transition at about 10-fold higher concentrations for the adhesive sphere model.[24] This apparent variation in the interaction with concentration is in complete contrast to the findings for γ_{II}-crystallin and lysozyme solutions.[12,13] However, Scherer and co-workers,[14] using a constant τ in the adhesive sphere model, found significant deviations from data for a monoclonal antibody when the second virial coefficient was negative. They attribute the deviation to self-association. Following this line of reasoning, we now demonstrate that the apparent concentration dependence of the interaction can indeed be removed simply by treating the attraction as highly directional instead.

We adopt a model of spherical particles interacting via a single attractive site placed off-center such that the overall interaction becomes patchy. The tendency for the particles to appear less attractive at higher concentrations now comes about far more naturally. By constraining the off-center attraction to short ranges, the system spontaneously associates into dimers to varying extents, depending on patch strength and protein concentration. Because of steric blocking of the attractive patch, the system becomes increasingly governed by excluded-volume interactions, once a sufficient amount of dimers have formed, as the volume fraction is increased. Such single-patch systems have been studied in the past by conventional integral equation theory and computer simulations over rather broad ranges of parameters.[25] Here, we are interested in dilute-to-semidilute solutions and turn to the integral equation theory by Wertheim,[15−17] which is particularly well suited for dealing with highly directional attractions. Proteins are initially treated as hard spheres interacting in addition via a short-range square-well attraction placed at the sphere surface, such that trimer formation is completely suppressed. For this simple model, the two-density integral equations to be solved involve only orientationally averaged quantities,[17] and these have been solved numerically.

Figure 5 shows predictions for the long wavelength limit of the structure factor, $S(0)$, from Wertheim's integral equation theory for one-patch spheres, with the patch located at the surface with an attractive range extending 10% of the sphere diameter. The system is controlled by the volume fraction of the spheres and the patch strength. For very strong patch attractions, such as $\beta\epsilon = 25$ in Figure 5, the structure factor of the spheres tends to the isothermal compressibility of hard-sphere dimers with the ideal gas result, $S(0) = \left(\frac{\partial \rho_0}{\partial \beta p}\right)_T = 2$, in the dilute limit.[26] Because of excluded-volume interactions, $S(0)$ decreases monotonically from 2 as a function of the volume fraction. A similar monotonic decrease with increasing ϕ occurs for weak patch strengths, $\beta\epsilon \lesssim 10$ in Figure 5, such that excluded-volume interactions again dominate. In contrast, for intermediate patch strengths, a nonmonotonic behavior of $S(0)$ as a function of volume fraction is predicted. For this range of patch strengths, the attraction leads to a buildup of dimers as a function of volume fraction and a decrease of monomers. Once the dimer concentration becomes sufficiently high, the effect of the patch attraction is diminished, and again, excluded-volume interactions become dominant because the short range of the patch attraction prohibits bonding of multiple spheres.

Compared to the experimental results in Figure 5, the model calculations capture the behavior qualitatively for the case of an active, dominant patch, that is, at 35 mM. As previously shown by computer simulations,[5,6] the patch originates from a subtle interplay between electrostatic and dispersion interactions. Adding an electrolyte screens the electrostatics, which eliminates the effect on measurements of the second virial coefficient in Figure 2. Consistent with these observations, the data for $S(0)$ for an ionic strength of 200 mM in Figure 5 exhibit a close-to-zero slope in the dilute limit, corresponding to $B_2 \approx 0$ (see Figure 2), and no maximum as a function of ϕ. Lowering the ionic strength to 35 mM, on the other hand,

DOI: 10.1021/acs.jpcb.6b06873
J. Phys. Chem. B XXXX, XXX, XXX−XXX

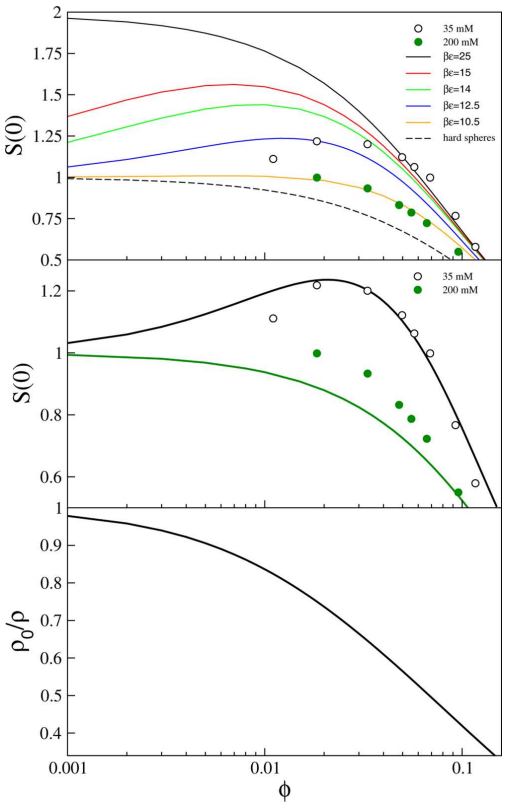


Figure 5. Long wavelength limit of the structure factor and monomer−dimer equilibrium as functions of the volume fraction. Top panel: results from Wertheim's integral equation theory for single-patch hard spheres are shown as solid lines for various patch strengths, as labeled. For the experiments, the data have been obtained by extrapolation of low-q data for the measured structure factor at the two NaCl concentrations, as labeled. Middle panel: results from Wertheim's integral equation theory for single-patch spheres are shown as solid lines for patch strengths of 11.2 (top) and 0 (bottom), using in addition an isotropic double Yukawa interaction. Bottom panel: mole fraction of monomers, produced by Wertheim theory, corresponding to the 35 mM case above.

produces an initial increase in $S(0)$ with volume fraction, corresponding to $B_2 < 0$ in Figure 2, and a sharp maximum at about $\phi \approx 0.02-0.03$.

Although in qualitative accord with the experiments, the integral equation theory for hard spheres with a single patch produces maxima in $S(0)$ at $\phi \lesssim 0.01$. The agreement between Wertheim theory for $S(0)$ and the experimental data for 35 mM NaCl can be improved by adding an isotropic interaction that is closer to reality than the hard-sphere one. Adding isotropic repulsions beyond the hard-core interaction, in addition to the patch attraction, results in moving of the maximum in $S(0)$ to lower volume fractions. This shift in the monomer−dimer equilibrium presumably occurs because more particles are repelled into bonded patch−patch configurations. Conversely, isotropic attractions in addition to the patch attraction lead to the opposite effect, whereby the maximum is shifted to higher volume fractions. In this case, the patch−patch attraction competes with the isotropic attraction, and somewhat higher concentrations are required before a sufficient number of dimers are formed to bring about a behavior dominated by excluded volume.

In Figure 5, we have superimposed the following isotropic potential[27] on the patch attraction

$$\beta\phi_{\mathrm{iso}}(r) = \begin{cases} \infty & r < D \\ -\dfrac{K\, \mathrm{e}^{-\kappa_{\mathrm{a}}(r-D)}}{r/D} + \dfrac{L_{\mathrm{B}} Z_{\mathrm{eff}}^2\, \mathrm{e}^{-\kappa_{\mathrm{d}}(r-D)}}{(1+\kappa_{\mathrm{d}}D/2)^2 r} & r > D \end{cases} \quad (7)$$

where the repulsive part is a screened Coulomb interaction, involving Bjerrum length L_{B}, Debye length κ_{d}^{-1}, and effective protein charge Z_{eff}. The additional isotropic attraction in eq 7, intended to model van der Waals attraction, has for the sake of simplicity been given the Yukawa form, with decay length κ_{a}^{-1} and maximum well depth K. In an attempt to assign reasonable values to these parameters, the diameter has been taken as 6 nm, as was done in the modeling of the structure factor using the adhesive sphere model in Figure 3. To assign a value to the effective protein charge, we follow the method of Palberg et al.[28] and estimate it from the Debye−Hückel expression, $Z_{\mathrm{eff}} = \frac{D}{2L_{\mathrm{B}}}(1+\kappa D/2)\frac{e|\zeta|}{k_{\mathrm{B}}T}$, with the surface potential exchanged for zeta potential ζ. This results in a low charge of $Z_{\mathrm{eff}} = 2$, which reflects a measured zeta potential of +4.5 mV at pH 7 and 35 mM. This low value of zeta potential agrees well with results from other experiments.[29] The Debye length was determined from the 35 mM bulk electrolyte concentration, and the parameters governing the isotropic attraction were set to $\kappa_{\mathrm{a}} = 20D^{-1}$ and $K = 1k_{\mathrm{B}}T$ to mimic van der Waals attraction.

With these parameters, the monomer−dimer equilibrium can be shifted so as to give a maximum close to that seen experimentally at 35 mM, as shown in Figure 5. In contrast to the concentration-dependent τ used in the absence of the patch attraction, a constant patch strength of $11.2k_{\mathrm{B}}T$ was used for all ϕ to capture the behavior at 35 mM in Figure 5. With the same parameter values for the isotropic potential, aside from the much shorter Debye length, the monotonic behavior of the data for 200 mM can be qualitatively reproduced if the patch strength is reduced or removed altogether. In Figure 5, a patch strength of $0k_{\mathrm{B}}T$ has been chosen.

The lower panel of Figure 5 shows the resulting monomer mole fraction for the 35 mM case as a function of volume fraction. As the overall concentration is increased, the fraction of dimers gradually increases at the expense of the monomer fraction. However, at volume fractions well past the maximum in $S(0)$, there remains an appreciable monomer fraction in equilibrium with the dimers formed because of the patch attraction.

A final remark concerns the one-patch model, which leaves particles free to rotate about the site−site axis. In the Monte Carlo simulations[5] of protein pairs, such rotations were never observed. Rather, for proteins in a dimer, only a few configurations were observed, with protein molecules severely restricted orientationally relative to one another. A two-patch model could be used to constrain dimer configurations, and although it would be more physically appealing, it would also lead to a more complicated model, which we leave for future studies.

DOI: 10.1021/acs.jpcb.6b06873
J. Phys. Chem. B XXXX, XXX, XXX−XXX

CONCLUSIONS

Effective structure factors have been determined for lactoferrin solutions. These reveal a drastically different behavior at low q, depending on the salt concentration. At 35 mM, a non-monotonic dependence of $S(0)$ on volume fraction results, whereas a monotonic decrease is observed at 200 mM. This behavior correlates with a negative-valued and near-zero-valued second virial coefficient, respectively. Such maxima in $S(0)$ generally appear for particles subject to sufficiently strong attractions. However, in the present case of lactoferrin solutions, the low volume fraction at the maximum points to a patch−patch attraction, such that association into dimers prevents further association into higher-order oligomers.

A simple model based on a patch attraction, restricted to inhibit trimers and higher-order oligomers, explains how a low-concentration maximum in the low-q-limiting structure factor develops because of dimerization and subsequent blocking of the attraction site. In contrast, assuming that the system is governed by isotropic interactions only leads to a strongly concentration-dependent interaction, which is, if not unphysical, certainly highly questionable. Furthermore, Wertheim's integral equation theory with explicit dimers can be brought into semiquantitative accord with data by superposing an isotropic interaction on the patch attraction. In this way, the pathway from monomers under very dilute conditions to mixed monomer−dimer states at higher concentrations can be described within a consistent statistical mechanical framework, with the capacity of predicting both structure and solution properties.

AUTHOR INFORMATION

Corresponding Author

*E-mail: malin.zackrisson@fkem1.lu.se. Phone: +46 (0)46 222 8185. Fax: +46 (0)46 222 4413.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank the OMM Linneaus Center, the Swedish Research Council, the Crafoord Foundation, The Royal Swedish Academy of Sciences (KVA), and LUNARC in Lund for computational resources and Morinaga Milk Industry Co., Ltd., Japan, for kindly providing us with the protein.

REFERENCES

(1) Kern, N.; Frenkel, D. Fluid−Fluid Coexistence in Colloidal Systems with Short-Ranged Strongly Directional Attraction. *J. Chem. Phys.* **2003**, *118*, 9882.

(2) Bianchi, E.; Largo, J.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Phase Diagram of Patchy Colloids: Towards Empty Liquids. *Phys. Rev. Lett.* **2006**, *97*, No. 168301.

(3) Bianchi, E.; Tartaglia, P.; Zaccarelli, E.; Sciortino, F. Theoretical and Numerical Study of the Phase Diagram of Patchy Colloids: Ordered and Disordered Patch Arrangements. *J. Chem. Phys.* **2008**, *128*, No. 144504.

(4) Munaò, G.; Preisler, Z.; Vissers, T.; Smallenburg, F.; Sciortino, F. Cluster Formation in One-Patch Colloids: Low Coverage Results. *Soft Matter* **2013**, *9*, 2652−2661.

(5) Persson, B. A.; Lund, M.; Forsman, J.; Chatterton, D. E. W.; Åkesson, T. Molecular Evidence of Stereo-Specific Lactoferrin Dimers in Solution. *Biophys. Chem.* **2010**, *151*, 187−189.

(6) Li, W.; Persson, B. A.; Morin, M.; Behrens, M. A.; Lund, M.; Oskolkova, M. Z. Charge-Induced Patchy Attractions between Proteins. *J. Phys. Chem. B* **2015**, *119*, 503−508.

(7) Piazza, R.; Iacopini, S.; Galliano, M. BLGA Protein Solutions at High Ionic Strength: Vanishing Attractive Interactions and "Frustrated" Aggregation. *Europhys. Lett.* **2002**, *59*, 149−154.

(8) Petsev, D. N.; Vekilov, P. G. Evidence for Non-DLVO Hydration Interactions in Solutions of the Protein Apoferritin. *Phys. Rev. Lett.* **2000**, *84*, 1339−1342.

(9) Roberts, D.; Keeling, R.; Tracka, M.; van der Walle, C. F.; Uddin, S.; Warwicker, J.; Curtis, R. The Role of Electrostatics in Protein-Protein Interactions of a Monoclonal Antibody. *Mol. Pharmaceutics* **2014**, *11*, 2475−2489.

(10) Marianayagam, N. J.; Sunde, M.; Matthews, J. M. The Power of Two: Protein Dimerization in Biology. *Trends Biochem. Sci.* **2004**, *29*, 618−625.

(11) Rosenbaum, D. F.; Zamora, P. C.; Zukoski, C. F. Phase Behavior of Small Attractive Colloidal Particles. *Phys. Rev. Lett.* **1996**, *76*, 150−153.

(12) Fine, B. M.; Lomakin, A.; Ogun, O. O.; Benedek, G. B. Static Structure Factor and Collective Diffusion of Globular Proteins in Concentrated Aqueous Solution. *J. Chem. Phys.* **1996**, *104*, 326−335.

(13) Piazza, R.; Peyre, V.; Degiorgio, V. "Sticky Hard Spheres" Model of Proteins Near Crystallization: A Test Based on the Osmotic Compressibility of Lysozyme Solutions. *Phys. Rev. E* **1998**, *58*, R2733−R2736.

(14) Scherer, T. M.; Liu, J.; Shire, S. J.; Minton, A. P. Intermolecular Interactions of IgG1 Monoclonal Antibodies at High Concentrations Characterized by Light Scattering. *J. Phys. Chem. B* **2010**, *114*, 12948−12957.

(15) Wertheim, M. S. Fluids with Highly Directional Attractive Forces. I. Statistical Thermodynamics. *J. Stat. Phys.* **1984**, *35*, 19−34.

(16) Wertheim, M. S. Fluids with Highly Directional Attractive Forces. II. Thermodynamic Perturbation Theory and Integral Equations. *J. Stat. Phys.* **1984**, *35*, 35−47.

(17) Wertheim, M. S. Fluids of Dimerizing Hard Spheres, and Fluid Mixtures of Hard Spheres and Dispheres. *J. Chem. Phys.* **1986**, *85*, 2929−2936.

(18) Lönnerdal, B.; Iyer, S. Lactoferrin - Molecular Structure and Biological Function. *Annu. Rev. Nutr.* **1995**, *15*, 93−110.

(19) Orthaber, D.; Bergmann, A.; Glatter, O. SAXS Experiments on Absolute Scale with Kratky Systems Using Water as a Secondary Standard. *J. Appl. Crystallogr.* **2000**, *33*, 218−225.

(20) Ng, K. C. Hypernetted Chain Solutions for Classical One-Component Plasma up to Γ = 7000. *J. Chem. Phys.* **1974**, *61*, 2680−2689.

(21) Klein, R.; D'Aguanno, B. Static Properties of Colloidal Suspensions. In *Light Scattering, Principles and development*; Brown, W., Ed.; Oxford University Press Inc.: New York, 1996.

(22) Kimura, Y.; Yoshimura, Y. Solvophobic Interaction in a Simple Associated Fluid. A Study on the Basis of Wertheim's Theory on Associated Fluids. *Mol. Phys.* **1992**, *76*, 737−755.

(23) Baxter, R. J. Percus-Yevick Equation for Hard Spheres with Surface Adhesion. *J. Chem. Phys.* **1968**, *49*, 2770−2774.

(24) Miller, M. A.; Frenkel, D. Competition of Percolation and Phase Separation in a Fluid of Adhesive Hard Spheres. *Phys. Rev. Lett.* **2003**, *90*, No. 135702.

(25) Giacometti, A.; Lado, F.; Largo, J.; Pastore, G.; Sciortino, F. Phase Diagram and Structural Properties of a Simple Model for One-Patch Particles. *J. Chem. Phys.* **2009**, *131*, No. 174114.

(26) Kalyuzhnyi, Y. V.; Holovko, M. F.; Haymet, A. D. J. Integral Equation Theory for Associating Liquids: Weakly Associating 2-2 Electrolytes. *J. Chem. Phys.* **1991**, *95*, 9151−9164.

(27) Gibaud, T.; Cardinaux, F.; Bergenholtz, J.; Stradner, A.; Schurtenberger, P. Phase Separation and Dynamical Arrest for Particles Interacting with Mixed Potentials - the Case of Globular Proteins Revisited. *Soft Matter* **2011**, *7*, 857−860.

(28) Palberg, T.; Mönch, W.; Bitzer, F.; Piazza, R.; Bellini, T. Freezing Transition for Colloids with Adjustable Charge: a Test of Charge Renormalization. *Phys. Rev. Lett.* **1995**, *74*, 4555−4558.

(29) Valiño, V.; San Roman, M. F.; Ibanez, R.; Benito, J. M.; Escudero, I.; Ortiz, I. Accurate Determination of Key Surface

DOI: 10.1021/acs.jpcb.6b06873
J. Phys. Chem. B XXXX, XXX, XXX−XXX

529 Properties that Determine the Efficient Separation of Bovine Milk BSA
530 and LF Proteins. *Sep. Purif. Technol.* **2014**, *135*, 145−157.

DOI: 10.1021/acs.jpcb.6b06873
J. Phys. Chem. B XXXX, XXX, XXX−XXX

Paper III

ARTICLE TYPE www. rsc.org/xxxxxx | XXXXXXXX

Solution Structures Formed by a Protein System Interacting via a Directional Patch Attraction†

Weimin Li,[a] Maxim Morin,[a] Emil Gustafsson,[a] Björn A. Persson,[b] Mikael Lund,[b] and Malin Zackrisson Oskolkova*[a]

Received Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX
First published on the web Xth XXXXXXXXXX 200X
DOI: 10.1039/b000000x

The phase behavior of lactoferrin has been studied as a function of concentration at a pH and ionic strength where lactoferrin is known to interact effectively via a patch-patch attraction. In contrast to isotropic attractive potentials, the directional attraction gives rise to a different phase or solution behavior. Initially, at low concentrations, the protein starts to dimerize. As the concentration is increased, the protein self-assembles into elongated, stripe-like structures at intermediate protein concentrations, a behavior which has been predicted for the case of attractive one-patch colloids. The stripe phase is surprisingly difficult to detect using conventional techniques, i.e. small-angle X-ray scattering, due to only a small fraction of the protein are participating in the stripes combined with susceptible sedimentation due to micron-sized entities. This is circumvented by monitoring the change in the overall protein concentration by static light scattering and the stripe formation can be followed. For visualization of the structures Cryo-TEM is used.

1 Introduction

Protein dimerization, oligomerization and association into higher ordered structures are fundamental processes in biology[1,2]. For some proteins it is their natural behavior and part of their biological function[3]. For others it is a part of *progressio morbi* of a number of neurodegenerative diseases where partial protein unfolding may lead to fibrillation and amyloid formation[4]. From a biophysical point of view, to understand various pathways of protein folding/unfolding and self-assembly and identifying their key mechanisms is of high importance. This is severely hampered by the intrinsic complexity of the vast diversity in protein shape and the distribution of the amino acids. The resulting overall protein-protein interactions are highly complex, and it quickly becomes difficult to determine which kind of interactions is not only present but also dominates for a particular protein.

Patchy interactions is a special kind of intermolecular interactions that implies the presence of special groups of residues or sites on the interacting molecules that induce anisotropic potentials, i.e. intermolecular interactions become highly directional and spatially asymmetric. In the case of proteins, the origin of patch-patch attractions was recently found to be of electrostatic origin which explains the strong impact of the solution pH and ionic strength[5–8]. Also the nature of the surface residues[9] and the directional nature of the patch[7] are important factors in alternating the protein interaction potentials.

The idea to view proteins as patchy colloids is not new[10], but for this particular protein studied here, lactoferrin, recent orthogonal data from which lactoferrin has been shown to interact in a highly stereospecific way show the interactions to follow the behavior of spheres with a one-site attraction[5,6]. Metropolis Monto Carlo simulation of lactoferrin predicted a highly stereo-specific attraction stemming from an electrostatic attractive patch composed of a few amino-acid residues located at the protein surface[7]. This highly directional interaction between lactoferrin molecules manifests as a non-monotonic behavior of the second virial coefficient, B_2, as a function of ionic strength, and leads to a concentration-induced self-association into dimers as the precursor for further phase variations[5,6]. To calculate a phase diagram for globular proteins based on spherical models with anisotropic interactions one needs to take directional patchy attractions into account[11,12]. Computer simulations on beads with variable number, size and strength of patches display rich phase diagrams[13,14]. In particular, the formation of elongated structures, wires, occurs in simulations under relatively small patch-to-sphere coverage ratio[14] which has until now not been seen for one-patch proteins or colloids.

In this work, we investigate the self-association and phase

† Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/b000000x/

[a] *Division of Physical Chemistry, Lund University, POB 124, 22100 Lund, Sweden.*
[b] *Division of Theoretical Chemistry, Lund University, POB 124, 22100 Lund, Sweden.*
* *Fax: +46 (0)46 222 4413; Tel: +46 (0)46 222 8185; E-mail: malin.zackrisson@fkem1.lu.se*

diagram of the milk protein lactoferrin in the range of pH and ionic strength where it is known to be interacting effectively as an attractive one-patch colloid[6]. We show that with the grow of protein concentration, protein dimers in equilibrium with monomers[6] appear due to the patch attractions. As the concentration is increased further, the patch attractions lead to formation of highly ordered, monodisperse macroscopic stripe-like structures. These structures are experimentally challenging to detect and investigate due to the low concentration and the structures being large enough to sediment due to gravity. We here show how the formation of the structures can be followed indirectly by monitoring the overall scattered intensity by means of static light scattering (SLS) and investigated and visualized using cryo-TEM. Furthermore, we hypothesize that the structures are formed in equilibrium as a consequence of the attractive patch. This is corroborated on the prediction of such wires in simulations on one-patch attractive spheres[14].

2 Materials and Methods

Bovine lactoferrin (>96% purity) with a molecular weight of 80 KDa[15] was purchased from Morinaga Milk Industry Co., Ltd., Japan. Monomeric solutions were prepared by dissolving the protein powder at a concentration of 2 g/L in NaOAc buffer of 5 mM ionic strength and pH 5.5 for a minimum of 48 hours, at room temperature. Purification and buffer exchange to tris buffer at pH 7 followed by concentrating dilute solutions were achieved by centrifugal filtration (Millipore Amicon, 50 and 100 kDa MWCO). The buffers used contained only monovalent electrolyte to minimize ion adsorption. Conversion from mass concentration (mg/mL) to volume fractions was done using a protein density of 1.3824 g/mL. A constant temperature of 25 ± 0.1°C was maintained in all experimental measurements. Further details of the characterization of the protein can be found elsewhere[5].

Purified protein solutions were monitored by dynamic light scattering (DLS) to ensure a narrow size distribution, ($\geq$ 95%) centered at 4 nm in agreement with the literature value[15]. Protein concentrations were determined spectrophotometrically with an extinction coefficient of 1.2673 cm^2/mg.

The static and dynamic light scattering experiments were performed using an ALV 5000F CGS-8F goniometer (ALV, Germany) and correlator equipped with a He-Ne laser diode (Spectra Physics, 127V/50 mW), operating at a wavelength $\lambda = 632.8$ nm.

Cryo-transmission electron microscopy (cryo-TEM) was conducted at the National Center for High Resolution Electron Microscopy at Lund University, Sweden. Images were recorded on a Gatan 791 CCD camera equipped with an Gatan GIF 100 imaging filtering system. Samples were prepared and transferred using an Oxford CT 3500 Cryoholder. Images were analyzed by using the ImageJ software[16].

Confocal laser scanning microscopy (CLSM) was carried out using an inverted confocal laser scanning microscope (Leica DMI6000) equipped with an SP5 tandem scanner using a 100 X times oil immersion objective. The protein solutions were sandwiched between coverslips and hermetically sealed. The glass surface were pre-coated with poly-L-lycine for stabilization. Fluorescence labeling of lactoferrin was done with Alexa 647 purchased from Molecular Probes. The dye to protein molar ratio was kept at 1:156, in order to minimize dye molecules influencing the protein-protein interactions.

3 Results

We here bring lactoferrin up to high protein concentrations (200 mg/mL) at conditions where it self-associates and forms concentration-induced dimers exactly in the attractive minima of the second virial coefficient in order to investigate how the presence of highly directional patch attractions[5,17] affects the phase diagram at higher concentrations. The self-association of lactoferrin monomers influences the static and dynamic properties of the solution. At the initial stage, one expects dimerization of the monomers to decrease the number density while increasing the particle size. The overall effect is a slight increment of scattered intensity due to the fact that scattered intensity has a stronger dependence on the particle size than on the number density of particles[18]. This trend is depicted as a function of time in figure 1.

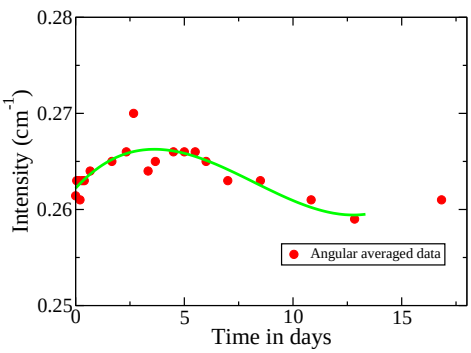


Fig. 1 Time evolution of angularly averaged intensity obtained from static light scattering on a wire-forming 100 mg/mL lactoferrin sample, at pH 7 and 35 mM. Green line is a polynomial fit to the data.

Within the first four days, the scattered intensity originates from multiple molecular species of different size. After four days incubation the intensity profile starts to decrease. This is

This journal is © The Royal Society of Chemistry [year]

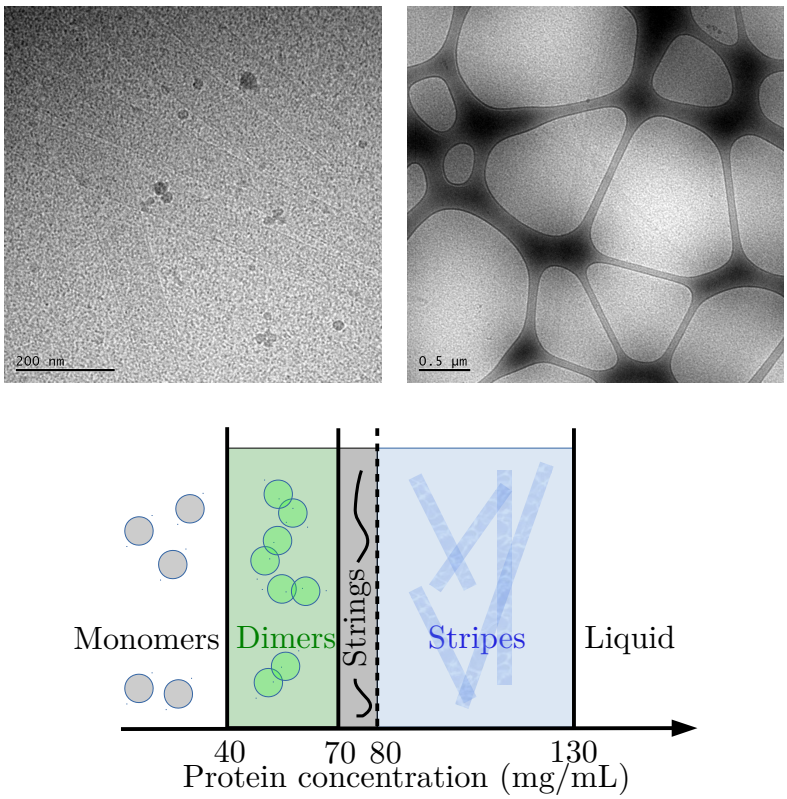


Fig. 2 Upper panel show two typical cryo-TEM images, of different magnification, of the new stripe structures of lactoferrin at patch attraction conditions corresponding to pH 7 and 35 mM salt. Lower panel show the schematic phase diagram of lactoferrin.

due to the formation of macro-sized objects which sediment to the bottom of the light scattering cuvette. This leads to a small but detectable decrease of the protein concentration in the bulk which manifests by a lowering of the total scattered intensity. The characterization of the new structures was inherently difficult because the amount of macro-sized objects is too small, and could not be detected using small-angle X-ray scattering. Instead, they were investigated using different microscopy techniques (cryo-TEM and CSLM). Cryo-TEM images of the new stripe phase are shown in the upper part of figure 2 at different magnifications. Ice is seen as dark spherical points in one of the micrographs and should not be misinterpreted as protein structures.

Systematic cryo-TEM measurements were undertaken at different protein concentrations, at patch conditions corresponding to pH 7 and 35 mM salt. Stripes formed at a protein concentration range between 80 mg/mL and 135 mg/mL after allowing for equilibration which roughly for this system took up to 14 days. We note (but do not show) at concentrations slightly smaller than 80 mg/mL, thin thread-like pre-structures had formed which may be unripe stripes or a separate structure. The stripes are stiff structures approximately 100 nm wide and micron-sized in length, ranging from one micrometer to few micrometers. The experiments were validated by repeating the cryo-TEM experiments on separate series of concentrations, which confirmed the stripe and string structures. In general, the structures were only formed after a few weeks of equilibration, images taken immediately after preparation showed no signs of stripes or wires. The low contrast found in the cryo-TEM images is a consequence of the high concentrations of single, dimer and possible higher order oligomeric protein molecules sitting in the background solvent. This is also reflected in the pixel profile densities. The difference in contrast between strings and background is very low, around

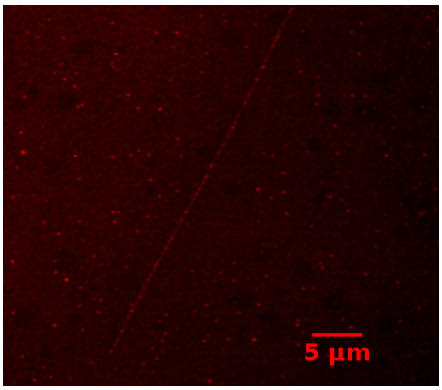


Fig. 3 Confocal microscope image of the stripe structures formed by lactoferrin at pH 7 and 35 mM salt. Alexa 647 was used as fluorescent label.

5% intensity. Nevertheless, the width of the strings can be extracted from the pixel profile and is determined to be around 12 nm which is comparable to twice the monomer size, obtained from DLS[6]. From this we conclude that the strings are likely composed of dimeric building blocks. The pixel density of the stripes are much higher with a difference in the intensity of about 15% compared to the background. Interestingly, the stripes have similar thickness as the string structures.

We conclude our observations by a simple test of the fragility of the structures. The stripes break into short fragments when a gentle stirring is applied before loading on the cryo-grid, which implies that the interactions between the building blocks of the stripes are not strong.

To further strengthen our observations and dismiss the possibility of the stripes and strings as structural artifacts induced by the cryo freezing, confocal microscopy was also conducted on samples where stripe structures had been observed in Cryo-TEM. The resulting imaging from CSLM shows also the stripes which can be seen in figure 3.

4 Conclusions

A schematic phase diagram of protein concentration under conditions when patch attraction dominates in the system is here presented. Increasing the protein concentration leads the stripe formation which is agreement with computer simulations on one-patch colloids with small patch coverage. We hypothesize that the structures are formed in equilibrium as a consequence of the attractive patch.

References

1 J. J. McManus, P. Charbonneau, E. Zaccarelli and N. Asherie, *Curr. Opin. Colloid Interf. Sci.*, 2016, **22**, 73–79.

2 N. J. Marianayagam, M. Sunde and J. M. Matthews, *Trends Biochem. Sci.*, 2004, **29**, 618–625.

3 R. Dominquez and K. C. Holmes, *Annu. Rev. Biophys.*, 2011, **40**, 169186.

4 V. N. Uversky and D. Eliezer, *Curr. Protein Pept. Sci.*, 2009, **10**, 483–489.

5 W. Li, B. A. Persson, M. Morin, M. A. Behrens, M. Lund and M. Z. Oskolkova, *J. Phys. Chem. B*, 2014, **119**, 503–508.

6 W. Li, B. A. Persson, M. Lund, J. Bergenholtz and M. Z. Oskolkova, *J. Phys. Chem. B.*, 2016, DOI: 10.1021/acs.jpcb.6b06873.

7 B. A. Persson, M. Lund, J. Forsman, D. E. W. Chatterton and T. Åkesson, *Biophys. Chem.*, 2010, **151**, 187.

8 D. Roberts, R. Keeling, M. Tracka, C. F. van der Walle, S. Uddin, J. Warwicker and R. Curtis, *Mol. Pharmaceut.*, 2014, **11**, 2475–2489.

9 W. N. P. II, Y. Chen, S. K. Handelman, H. Neely, P. Manor, R. Karlin, R. Nair, J. Liu, M. Baran, J. Everett, S. N. Tong, F. Forouhar, S. S. Swaminathan, T. Acton, R. Xiao, J. R. Luft, A. Lauricella, G. T. DeTitta, B. Rost, G. T. Montelione and J. F. Hunt, *Nat. Biotechnol.*, 2009, **27**, 51–57.

10 C. Gögelein, G. Nägele, R. Tuinier, T. Gibaud, A. Stradner and P. Schurtenberger, *J. Chem. Phys.*, 2008, **129**, 085102.

11 P. B. Warren, *J. Phys.: Condens. Matter*, 2002, **14**, 7617–7629.

12 R. P. Sear, *J. Chem. Phys.*, 1999, **111**, 4800–4806.

13 A. Giacometti, F. Lado, J. Largo, G. Pastre and F. Sciortino, *J. Chem. Phys.*, 2010, **132**, 174110.

14 Z. Preisler, T. Vissers, G. Munaò, F. Smallenburg and F. Sciortino, *Soft Matter*, 2014, **10**, 5121–5128.

15 B. Lönnerdal and S. Iyer, *Lactoferrin: Molecular Structure And Biological Bunction*, Annual Reviews Inc. a, P.O. Box 10139, 4139 EI Camino Way, Palo Alto, California 94306, USA, Dep. Nutr., Univ. Calif., Davis, CA 95616, USA, 1995.

16 S. Johannes, A. C. Ignacio, F. Erwin, K. Verena, L. Mark, P. Tobias, P. Stephan, R. Curtis, S. Stephan, S. Benjamin, T. J. Yves, W. D. James, H. Volker, E. Kevin, T. Pavel and C. Albert, *Nature Methods*, 2012, **9**, 676–682.

17 W. Li, B. A. Persson, , M. Lund, J. Bergenholtz and M. Z. Oskolkova, *X*, 2016, **X**, X.

18 P. Lindner and Th.Zemb, *neutrons, X-rays and Light: Scattering Methods Applied to Soft Condensed Matter*, ELSEVIER SCIENCE B.V. Sara Burgerhartstraat 25 P.O. Box 211, 1000 AE Amsterdam, The Netherlands, Elsevier Science Global Rights Departent, PO Box 800, Oxford OX5 1DX, UK, 2002.

This journal is © The Royal Society of Chemistry [year]

Paper IV

Charge Fluctuations Calculated from Experimental Titration and Monte Carlo Simulations on Large Protein Molecules

Weimin Li,[†] Björn Persson,[‡] Mikael Lund,[‡] and Malin Zackrisson Oskolkova[†]

Division of Physical Chemistry, Lund University, P. O. Box 124, SE 22100, Lund, Sweden, and Division of Theoretical Chemistry, Lund University, P. O. Box 124, SE 22100, Lund, Sweden

E-mail:

Abstract

Charge capacitance of lactoferrin, which is a large protein molecule, was revealed from titration experiments and Monte Carlo (MC) simulations. The results from both methods were found to be in a qualitatively good agreement. Dimerization of lactoferrin at pH 10 was found to have a minor effect on the capacitance as a function of pH compared to the monomer titration shown by MC simulations. Dynamic and static light scattering (DLS, SLS) experiments show that lactoferrin molecules are associated into trimers or tetramers, at pH 10. The interaction mechanism is discussed.

Introduction

Electrostatic interactions play an important role for in the solution behavior of proteins. This is manifested by the fact that a change in pH or ionic strength have a significant effect on the properties of protein solutions. For pH values far from the isoelectric point of the protein in question, Coulomb interactions originating from the overall net charges are predominant. When pH is close to the isoelectric point of the protein, electrostatic interactions sensitive to the protein charge distribution become important. This type of interaction was described

*To whom correspondence should be addressed

[†]Division of Physical Chemistry, Lund University, P. O. Box 124, SE 22100, Lund, Sweden

[‡]Division of Theoretical Chemistry, Lund University, P. O. Box 124, SE 22100, Lund, Sweden

by Kirkwood to originate from charge fluctuations.[1] Proteins composed of large number of neutral and ionizable amino acid residues where protons can attach, have numerous possible electric charge configurations. Fluctuations between the possible configurations induce an electrostatic correlation which has an attractive nature.

Previous studies of lactoferrin solution by both computer simulations and experiments[2–4] have suggested an attractive mechanism of interactions in the system at 35 mM of salt and pH 7 by showing that absolute charges do not fluctuate when proteins approach each other, instead lactoferrin shows a patch attraction due to a few charged amino acids under these conditions. Here, we are interested in conditions where charge fluctuations are expected to play a dominant role.

Interactions induced by charge fluctuations have a direct connection with the charge capacitance of a protein.[5,6] It has been shown previously that the protein capacitance is the derivative of the protein titration curve.[5,7] Experimentally the protein charge capacitance was calculated from titration measurements on a small-sized protein, calbindin D_{9k}.[8,9] Thus to determine the conditions for charge regulation in a solution of large protein, lactoferrin, a study of charge capacitance is needed.

Materials and Methods

Experiments

Lactoferrin is a globular protein composed of two domains, having a peanut shape with long axis about 10 nm and short axis 6 nm.[10,11] The theoretical isoelectric point (pI) of lactoferrin is 9.4.[2] Experiments provide a pI ranged from 8 to 9.[12,13] Raw lactoferrin was purchased from Morinaga Milk Industry Co. Ltd., Japan with a purity of >96%. Mature and pure protein solution was prepared as described in the previous work.[3]

A titrating spectrometer (Probe Drum) was used for titration experiments. The pH-meter was calibrated on pH standards (Thermo Science) from pH 1–12 before the titration experiments. The protein solution was controlled to have a concentration in the range from 5 mg/mL to 10 mg/mL with a volume of 1 mL, in order to have measurable proton association/dissociation from the proteins. Buffer chemicals were not included in the solvent. The protein solution was prepared at two ionic strengths, 35 mM and 200 mM, for the titration experiments.

One complete titration is composed of two parts: an acidic titration roughly from pH 7–1.5, and a basic titration roughly from pH 4–12. Titrating from both pH ends is not feasible because of the difficulty in sample preparation. Another level of complexity comes from the

fact that lactoferrin denatures at low pH, and high pH solution is hard to be maintained due to the adsorption of CO_2.

HCl and NaOH are the acidic and basic titrants with a concentration of 1 M. The step size of adding a titrant is adjusted to have a reasonable distribution of pH points. After each titrant injection a stirring process and equilibrating time are imposed. The consumption of titrant by the bulk water is removed by using the following relation

$$\begin{cases} \pm n_t &= [H^+]_f V_f - [H^+]_i V_i + [OH^-]_i V_i - [OH^-]_f V_f \ , \\ n_t &= (V_f - V_i) c_t, \end{cases} \tag{1}$$

where n_t is the molar number of titrant, $\pm$ depends on if it is acidic or basic titration. $[H^+]_i$ and $[H^+]_f$ are the concentration of protons before and after adding a titrant, calculated from pH via $pH = -log_{10}[H^+]$. $[OH^-]_i$ and $[OH^-]_f$ are obtained by using water dissociate constant, $[H^+][OH^-] = 13.998$ for 35 mM and 14.003 for 200 mM at 25°C.[14]

Theoretically a protein is expected to be fully protonated at low pH resulting in a titration plateau. By fitting the acidic part of the titration curve at low pH using [15]

$$n_{H^+} = N_{Asp} \frac{e^{n_{Asp}(pH - pKa_{Asp})}}{1 + e^{n_{Asp}(pH - pKa_{Asp})}} + N_{Glu} \frac{e^{n_{Glu}(pH - pKa_{Glu})}}{1 + e^{n_{Glu}(pH - pKa_{Glu})}} \tag{2}$$

provides an absolute scaling for the titration curve of the overall protein charges. n_{H^+} is the number of dissociated protons, n_{Asp} and n_{Glu} are the numbers of dissociated aspartic (Asp) and glutamic acids (Glu). Asp and Glu have the lowest acid dissociation constants,[16] and for this reason they are used to calculate the theoretical fittings at low pH.

Taking the derivative of the protein titration curve provides the capacitance, C, resulting in the equation [5]

$$C = -\frac{1}{\ln 10} \frac{\partial \langle Z \rangle}{\partial \mathrm{pH}}, \tag{3}$$

where Z is the number of protons dissociated per protein.

Collective diffusion as a function of protein concentration is measured by dynamic light scattering (DLS). Extrapolation of diffusion to a zero-concentration limit gives D_0, which can be used to calculate the hydrodynamic radius of protein, r_h, by using Stokes-Einstein equation

$$D_0 = \frac{k_B T}{6\pi\eta r_h}, \tag{4}$$

where k_B is the Boltzmann constant, T is the temperature, η is the viscosity of the solvent.

Static light scattering can be used to determine the molecular weight of a protein, M_w, and the second virial coefficient, B_2. B_2 is a quantity describing the overall interaction

of the system, positive second virial coefficient indicates overall repulsion and negative B_2 corresponds to overall attraction. B_2 can be extracted from a reduced Zimm plot, for samples where there is no angular variation using the equation

$$\frac{KC}{R_{90}} = \frac{1}{M_w} + \frac{2N_A B_2}{M_w^2} C, \tag{5}$$

where

$$K = \frac{4\pi^2 n_0^2 (dn/dC)^2}{N_A \lambda^4} \tag{6}$$

C is the protein concentration, R_{90} is the Rayleigh ratio measured at angle of 90°, N_A is Avogadro's number, n_0 is the refractive index of solvent, λ is the wavelength, and dn/dC is the increment refractive index.

Static and dynamic light scattering (SLS and DLS) are performed simultaneously on ALV instrument 5000F CGS-8F goniometer from Gernamy.

The zeta potential (ZP) was calculated from the measured electrophoretic mobility according to the Henry equation

$$U_E = \frac{2\varepsilon z f(Ka)}{3\eta}, \tag{7}$$

where U_E is the electrophoretic mobility, z is the Zeta potential, ε is the dielectric constant, η is the viscosity, $f(Ka)$ is Henry's function which is approximated (Huckel approximation) to be 1.0 in a small particles system. The instrument used was Malvern Zetasizer (Nano-ZS), and the diffusion barrier technique was applied to minimize the probability of protein being in contact with electrodes.

Metropolis Monte Carlo Simulations

The theoretical results were obtained from Metropolis Monte Carlo (MC) simulations, performed using the Faunus framework.[17] We use a rigid coarse-grained protein model, derived from experimental protein structures, where amino acids were represented by soft spheres. The protein models were based on crystal structures from the PDB repository of bovine (1BLF).[10] This method was applied previously to several different systems,[2,18] including studies of large proteins and thermodynamic properties of protein solutions.[19] pH dependent charges of titratable residues were placed in the center of the amino acid spheres. Charges were obtained from simulations with single proteins in precence of explicit salt as described previously.[2,6] The protonation states were allowed to fluctuate and salt particles were treated explicitly in a Grand Canonical MC scheme, i.e. under constant chemical section.

Results and Discussion

The titration results at two different ionic strengths, 35 mM and 200 mM, at pH 10 are shown in figure 1. To obtain proton association per protein, titration background needs to be subtracted. It is highly challenging to titrate water repetitively, therefore the theoretical titration of water is used as a background. The data is presented in terms of associated protons from the acidic end. It is calculated from experimental number of added titrants at each pH step. Negative values indicate protons dissociating from positively charged amino acid residues. Using spline interpolation of the experimental repetitions, acidic titration (full circles) and basic titration (full squares) are merged based on minimizing the standard deviations of overlapping points, and the averages (full lines) are calculated afterwards.

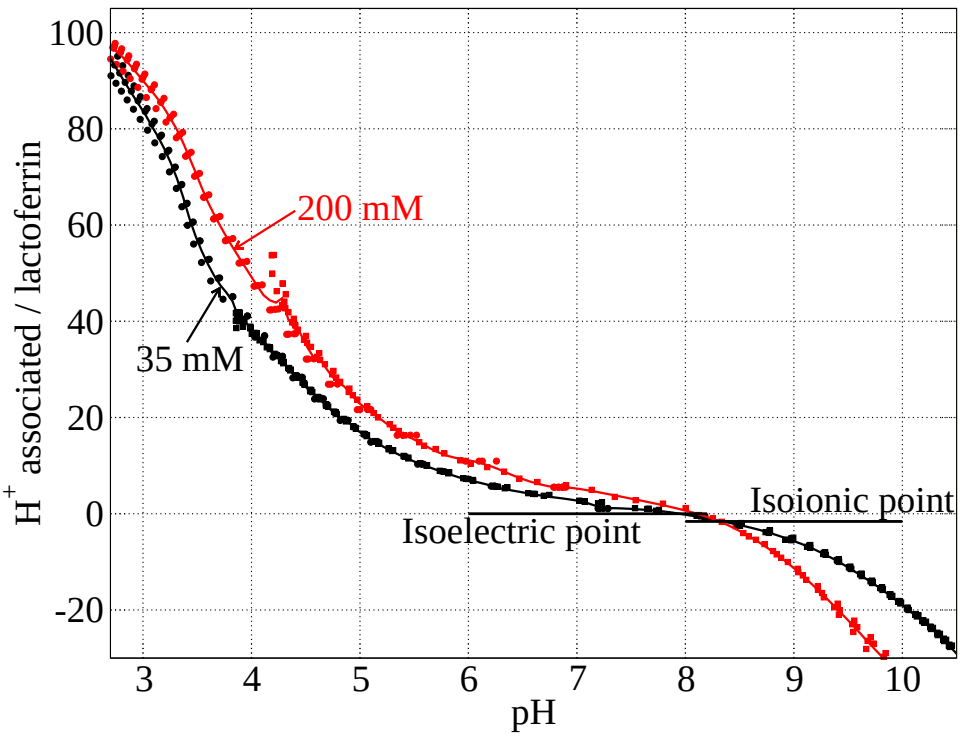


Figure 1: Lactoferrin titration at two ionic strengths (35 mM in black, and 200 mM in red). Symbols are experimental data, repetitions are included. Two lines are the average of spline interpolation of each salt repetitions.

A titration plateau is necessary to put the titration curve on an absolute scale, which

represents the correct state of protein overall charges. Experiments show difficulty obtaining this titration plateau due to protein denaturation and large proton consumption by water at low pH. Instead, a theoretical titration plateau was calculated by using Asp and Glu residues to fit the low pH part of titration curve. Theoretical plateau guides to scale the experimental data to absolute values. The discrepancy of experimental repetitions at low pH decreases the accuracy of theoretical plateau. Preliminary theoretical fits provide an isoelectric point (zero charge) and isoionic point (zero ion adsorption) around 8 as shown in figure 1, which is reasonable when comparing with the corresponding values in literature.[12]

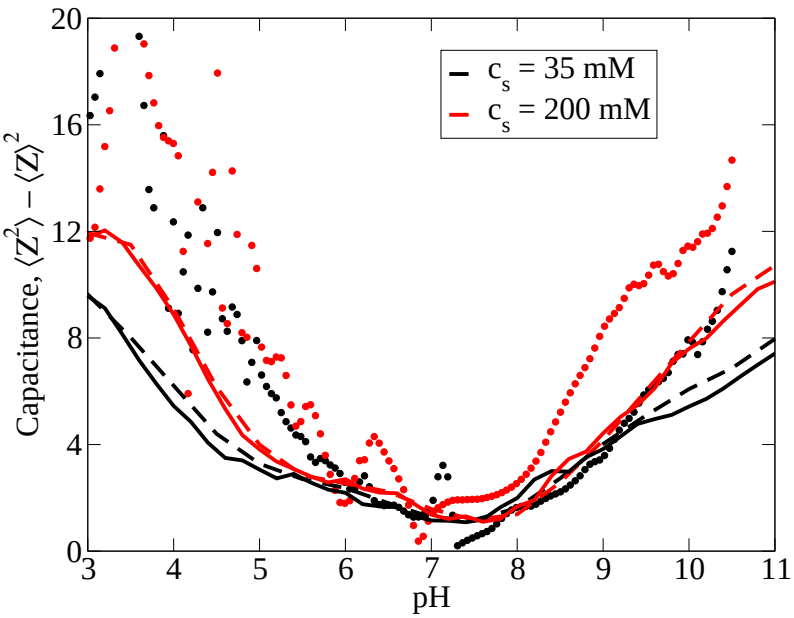


Figure 2: Capacitance (left) and protein charge (right) of lactoferrin as a function of pH at 35 mM (red) and 200 mM (black) salt concentrations. From MC simulations of lactoferrin monomers (dashed), dimers (fully drawn), and measured (symbols).

Comparison of the computed from MC simulations and experimentally measured capacitance which was obtained by differentiating the titration curve is present in figure. 2. Both show a generally increased capacitance as the salt concentration is increased. This is expected due to decreased internal repulsion because of screening. The lack of agreement at high and low pH likely originates from the fact that not only protein but also water titrates significantly in these regimes.

In studies of lactoferrin it has previously been suggested that binding of chloride needs to

be taken into account to describe the electrophoretic mobility of the proteins.[20] The mobility was measured in terms of zeta potential which was found to be 4.55 mV at pH 7 and 5 mM of salt, see table 1. By varying pH we explored the corresponding mobility above the isoelectric point and found the zeta potential at pH 10 to be -5.23 mV.

Table 1: Measured properties at 5 and 20 mM of salt, [I], at pH 7 and pH 10.

	B_2 (nm^3)		ZP (mV)		D_0 (μm^2/s)		M_w (kg/mol)	
I (mM)	5	20	5	20	5	20	5	20
pH 7	-105.58	-490*	4.55	—	57.078	56.705	81	83*
pH 10	-68.76	-251.72	-5.23	—	35.684	40.157	373	268

* value is the average of corresponding 15 and 25 mM salt results.

The results imply the absolute charge of the protein to be similar at these pH values. However, the MC simulations set the valency of the protein at 15.0 and -5.4 at pH 7 and pH 10, respectively. This discrepancy is in accordance with the previous observations and suggests that ion binding may occur here.[21]

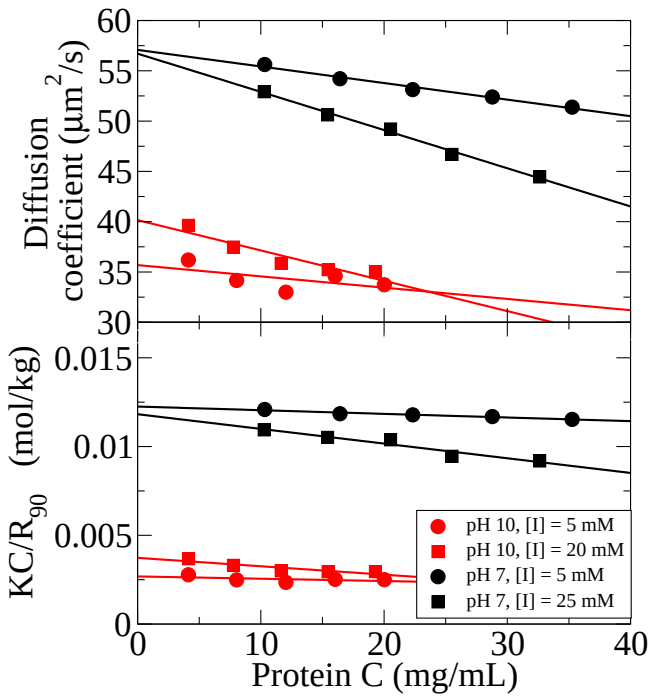


Figure 3: Top panel shows collective diffusion as a function of protein concentration. Bottom panel shows the reduced Zimm-plot. Conditions are labeled.

In fact, the picture is more complicated when comparing these solution conditions at pH 7 and 10. Figure 3 shows in the top graph the collective diffusion as a function of protein concentration and the lower graph show a reduced Zimm plot at 5 and 20 mM of salt, at pH 7 and pH 10, as shown in the legend. The D_0 and M_w from each of the salt and pH conditions, listed in the table 1, indicates different species at hand. At pH 7 the apparent size is around 4 nm corresponding to a monomer,[10,22] also in agreement with our previous result at these concentrations[4] and the SLS derived molecular weight listed in table 1. Larger oligomers are present at pH 10 with an apparent size of 6.5 nm and a molecular weight suggesting trimers or possibly tetramers.

The influence of association on the titration of the protein was studied by MC simulation and is shown in figure 2. This is a comparison of capacitance between lactoferrin dimer and monomer. If dimerization was driven by charge regulation it would be manifested in the stabilization of either protonated or deprotonated states of certain residues and result in a decreased capacitance. Instead we find the capacitance independent of the aggregation state.

B_2 in table 1, extracted from slopes of the reduced Zimm plot, shows increased attraction by increasing salt concentration from 5 to 20 mM, at both pH 7 and pH 10. This could be due to reduced Coulomb repulsion because of salt screening. At pH 7, the identified patch attraction[3,4] contributes to the total interactions which is manifested by decreasing B_2. At pH 10, the charge regulation induces attraction which is believed to be involved in the interaction between the trimers or tetramers due to its prominent capacitance.

Conclusions

We have examined the capacitance of lactoferrin by experimental titration and Monte Carlo simulated titration. A qualitatively good agreement between them has been found. Two pH conditions were identified, pH 7 and pH 10, where the proteins have opposing zeta potential but similar absolute charges. The hydrodynamic radius and molecular weight calculated from dynamic and static light scattering, respectively, indicates protein association. The protein capacitance has been shown to be independent of aggregation state, as least for dimers. Charge regulation may be a significant contribution to the total interaction at pH 10, but requires more investigation at higher salt concentration in order to draw a more convincing conclusion.

Acknowledgement

The thanks from authors are given to OMM Linneaus center in Lund; the Swedish Research Council; the Swedish Foundation for Strategic Research; the Crafoord Foundation; the Royal Swedish Academy of Sciences; LUNARC computational source in Lund; Morinaga Milk Industry Co., Ltd, Japan.

References

1. Kirkwood, J. G., and Shumaker, J. B. (1952) Forces between protein molecules in solution arising from fluctuations in proton charge and configuration. *Proc. Natl. Acad. Sci 38*, 863–871.

2. Persson, B. A., Lund, M., Forsman, J., Chatterton, D. E., and Åkesson, T. (2010) Molecular evidence of stero-specific lactoferrin dimers in solution. *Biophys. Chem. 151*, 187–189.

3. Li, W., Persson, B. A., Morin, M., Behrens, M. A., Lund, M., and Oskolkova, M. Z. (2014) Charge-Induced Patchy Attractions between Proteins. *J. Phys. Chem. B 119*, 503–508.

4. Li, W., Persson, B. A., , Lund, M., Bergenholtz, J., and Oskolkova, M. Z. (2016) Concentration-induced association in a protein system caused by a highly directional attraction. *J. Phys. Chem. B* X.

5. Lund, M., and Jönsson, B. (2013) Charge regulation in biomolecular solution. *Q. Rev. Biophys. 46*, 256–281.

6. Lund, M., and Jönsson, B. (2005) On the charge regulation of proteins. *Biochemistry 44*, 5722–5727.

7. Ullner, M., and Jönsson, B. (1993) Conformational properties and apparent dissociation constants of titrating polyelectrolytes: Monte Carlo simulation and scaling arguments. *J. Chem. Phys. 100*.

8. Kesvatera, T., Jönsson, B., Thulin, E., and Linse, S. (1999) Ionization behaviour of acidic residues in Calbindin D_{9k}. *Proteins 37*, 106–115.

9. Kesvatera, T., Jönsson, B., Thulin, E., and Linse, S. (2001) Focusing of the electrostatic potential at ef-hands of Calbindin D_{9k}: Titration of acidic residues. *Proteins 45*, 129–135.

10. Moore, S. A., Anderson, B. F., Groom, C. R., Haridas, M., and Baker, E. N. (1997) Three-dimensional structure of diferric bovine lactoferrin at 2.8 øA resolution. *J. Mol. Biol. 274*, 222–236.

11. Lönnerdal, B., and Iyer, S. *Lactoferrin: Molecular Structure And Biological Bunction*; Annual Reviews Inc. a, P.O. Box 10139, 4139 EI Camino Way, Palo Alto, California 94306, USA: Dep. Nutr., Univ. Calif., Davis, CA 95616, USA, 1995.

12. Superti, F., Siciliano, R., Rega, B., Giansanti, F., Valenti, P., and Antonini, G. (2001) Involvement of bovine lactoferrin metal saturation, sialic acid and protein fragments in the inhibition of rotavirus infection. *Biochim. Biophys. Acta 1528*, 107–115.

13. Groves, M. L. (1960) The isolation of a red protein from milk. *J. Am. Chem. Soc. 82*, 3345–3350.

14. Harned, H. S., and Owen, B. B. *The Physical Chemistry of Electrolytic Solutions (3rd ed.)*; New York: Reinhold, 1958; pp 639–643.

15. Leiding, T., Górecki, K., Kjellman, T., Vinogradov, S. A., Hägerhäll, C., and Årsköld, S. P. (2009) Precise detection of pH inside large unilamellar vesicles using membrane-impermeable dendritic porphyrin-based nanoprobes. *Anal. Biochem 338*.

16. Weast, R. C., and Astle, M. J. *CRC Handbook of Chemistry and Physics*; CRC Press, Inc. Boca Raton, Floride, 1982; pp C756–C759.

17. Lund, M., Trulsson, M., and Persson, B. (2008) Faunus: An object oriented framework for molecular simulation. *Source Code for Biology and Medicine 3*.

18. Persson, B., and Lund, M. (2009) Association and electrostatic steering of α-lactalbumin and lysozyme heterodimers. *Phys. Chem. Chem. Phys. 11*.

19. Kurut, A., Persson, B. A., Åkesson, T., Forsman, J., and Lund, M. (2012) Anisotropic Interactions in Protein Mixtures: Self Assembly and Phase Behavior in Aqueous Solution. *J. Phys. Chem. B 3*.

20. Bowen, W. R., Cao, X., and Williams, P. M. (1999) Use and elucidation of biochemical data in the prediction of the membrane separation of biocolloids. *Proc: Math. Phys. Eng. Sci. 455*.

21. Mela, I., Aumaitre, E., Williamson, A., and Yakubov, G. (2010) Charge reversal by salt-induced aggregation in aqueous lactoferrin solutions. *Colloids Surf. B 78*.

22. Brock, J. H. (2002) The physiology of lactoferrin. *Minireview, Biochem. Cell Biol. 80*, 1–6.