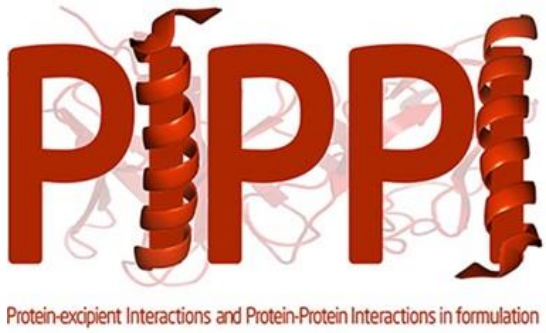
Protein-protein interactions in high protein concentrations

Mahapatra, Sujata

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record



Protein-protein interactions in high protein concentrations

By

Sujata Mahapatra

A thesis submitted in partial fulfilment for the
degree of Doctor of Philosophy at the
Department of Chemistry,
Technical University of Denmark

October 2019

Preface

The dissertation has been submitted to the Department of Chemistry, Technical University of Denmark as partial fulfillment of the requirements to obtain a PhD at DTU. The presented work was carried out in the Department of Protein Biochemistry and Stability, Novozymes A/S and Department of Chemistry, Technical University of Denmark, under the supervision of Science manager Werner W. Streicher and Professor Pernille Harris and co-supervision from Science manager Allan Nørgaard and Professor Günther H. J. Peters. Part of the thesis work were carried out during my external stays, which presents static light scattering measurements at Wyatt Technologies Europe under the guidance of Managing Director Dierk Roessner, theoretical modeling at University of Manchester under the guidance of Professor Robin Curtis and coarse-grained modeling at Lund University under the guidance of Professor Mikael Lund. A part of the project was to purify and characterize lipase for PIPPI using state-of-the-art techniques from Novozymes (not included in the thesis).

Small-angle X-ray scattering experiments were carried out at the EMBL, P12, BioSAXS beamline at DESY (Hamburg, Germany) and at the EMBL, BM29, BioSAXS beamline at ESRF (Grenoble, France). This project was funded by a part of the EU Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie grant agreement No 675074 – "Protein excipient Interactions and Protein-Protein Interactions in formulation" (PIPPI).

The thesis contains six chapters in total and the work presented in chapters 3, 4 and 5 are intended to publish in three manuscripts.

Contribution to other publications and manuscripts

Gentiluomo, L., Roessner, D., Augustijn, D., Svilenov, H., Kulakova, A., **Mahapatra, S.**, Winter, G., Streicher, W., Rinnan, Å., Peters, G.H.J., Harris, P., Frieß, W. Application of Interpretable Artificial Neural Networks to Early Monoclonal Antibodies Development. *Eur J Pharm Biopharm*. **2019**, 141, 81-89.

Augustijn, D., **Mahapatra, S.**, Streicher, W., Svilenov, H., Kulakova, A., Pohl, C., Rinnan, Å. Novel Non-linear Curve Fitting to Resolve Protein Unfolding Transitions in Intrinsic Fluorescence Differential Scanning Fluorimetry. *Eur J Pharm Biopharm.* **2019**, 142, 506-517.

Gentiluomo, L., Roessner, D., Streicher, W., **Mahapatra, S.**, Harris, P., Frieß, W. Characterization of native reversible self-association of a monoclonal antibody mediated by Fab-Fab interaction. *J Pharm Sci*. **2019**, doi:10.1016/j.xphs.2019.09.021.

Gentiluomo, L., Svilenov, H., Augustijn., D., El Bialy, I., Greco, M.L., Kulakova, A., Indrakumar, S., **Mahapatra, S**., Morales, M., Pohl, C., Roche, A., Tosstorff, A., Curtis, R., Nørgaard, A., Khan, T.A., Peters, G.H.J., Rinnan, Å., Streicher, W., van der Walle, C., Uddin, S., Winter, G., Roessner, D., Harris, P., Frieß, W. Boosting therapeutic protein development by publicly available datasets including comprehensive computational and biophysical characterization. **2019**. (Manuscript submitted).

Kulakova, A., Indrakumar, S., Sønderby, P., **Mahapatra, S.**, Streicher W., Peters G.H.J., Harris, P. Albumin-Neprilysin Fusion Protein: Understanding Stability using Small-Angle X-ray Scattering and Molecular Dynamics Simulations. 2019. (Manuscript submitted).

Augustijn, D., Kulakova, A., **Mahapatra, S**., Harris, P., Rinnan, Å. Isothermal Chemical Denaturation: Data Analysis, Error Detection and Correction by PARAFAC2. **2019**. (Manuscript in preparation).

Acknowledgments

I am immensely grateful for the guidance and support of my supervisors Science Manager Werner W. Streicher and Professor Pernille Harris and of my co-supervisors Science Manager Allan Nørgaard and Professor Günther H. J. Peters at both Novozymes and DTU, all of whom I would like to thank for including me in the PIPPI project and giving me the opportunity to work on an exciting project. My special thanks to Werner Streicher and Pernille Harris for their invaluable feedback, helping me to stay on track in difficult moments, and for allowing me to participate in conferences around the world and correction of the thesis.

I would like to thank the European Union's Horizon 2020 research and innovation program for funding the PIPPI project. Everyone in the Marie Skłodowska-Curie PIPPI project has contributed to the teamwork necessary to successfully complete the many parts of our combined achievements. Indeed, it has been inspirational to be a member of this group of talented people who are not only strong scientists but who are also great colleagues and so much fun to be with. I would happily work with all of you again.

DANSCATT provided the financial support making the synchrotron trips possible, for which I am grateful. I wish to thank the great beamline scientists from P12 BioSAXS at DESY in Hamburg, Germany and from B29 BioSAXS at ESRF in Grenoble, France for all the help and assistance during our measurements.

My acknowledgment also goes to my collaborators for being gracious hosts during my research visits. Professor Robin Curtis at the Manchester Institute of Biotechnology, thank you for taking time out of your busy schedule and for all the inspiring research discussions about protein-protein interactions. I am very thankful to Professor Mikael Lund and my colleague Marco Polimeni at Lund University for introducing me to the world of coarse-grained modeling and for all our scientific discussions. I would also like to thank Dr. Dierk Roessner and my colleague Lorenzo Gentiluomo for hosting me at Wyatt Technologies Europe in Dernbach, Germany, to perform light scattering measurements and helping with the data analysis.

I would like to give my special thanks to Senior Advisor Jens T. Bukrinski for our interesting scientific discussions about antibodies and for generously offering his thoughts regarding my future opportunities.

A special thanks to Gitte J. Funch at Novozymes who patiently taught me how to find my way around the lab and never thought that I asked too many questions about how to operate the equipment. I am very thankful to Pernille Sønderby for teaching the SAXS course and helping me with SAXS data interpretation and planning my samples. Many thanks to my colleagues from the PIPPI project in Denmark, Alina Kulakova, Christin Pohl, Sowmya Indrakumar and Dillen Augustijn for all their scientific input. You are amazing persons and I had such a great time working with you.

The IT staffs at DTU deserve my gratitude who did their superhero stuff when on very short notice I had to transfer all of my documents and data to a new laptop during the end of my project.

Heartfelt gratitude goes to all my colleagues at DTU and Novozymes, Alina Kulakova, Christin Pohl, Bartłomiej Kołaczkowski, Sowmya Indrakumar, Line A. Ryberg, Ulf Molich, Tine M. Frederiksen, Natalia Skawinska, Suk Kyu Ko, Ida M. Vedel and Kasper D. Tidemand who made the days at the office enjoyable. I am thankful for all the good times together, the numerous social activities, Christmas lunches, Friday breakfasts and cakes. Many thanks to Line, Christin, Alina, Natalia, Marco, Aisling, Pernille Harris, and Günther Peters for a great company and help during our synchrotron trips. Thanks again to Ulf for helping out with Matlab scripts and IT-related issues. Thanks to Maria Blanner Bang for helping with practical matters in the lab at DTU.

My family and friends have made me realize that no other success can compensate for failure at home. I send my love to my grandparents, Uchhav and Saraswati, to my parents, Rabindra and Kishori, to my sisters Sunita, Pooja and Liza and to my brother Om all of whom constantly have been providing me with uplifting comments and much-needed encouragement. I thank my close friends, Sandeep and Anurag for all the fun trips and putting up with me during my good and bad days. I am especially grateful to Carsten Bober for reading the thesis and providing helpful comments. I also thank my other friends in Denmark, Ankita, Shajeel, Hariom, Minkle, Maulik, Krishna, Omkar and Gaurav whom I love to spend time with whenever possible. I am truly lucky to have all of you in my life.

In the loving memory of Umacharan and Satyabhama.

Sujata Mahapatra

Abstract

Protein-based therapeutics are increasingly important due to their high specificity, potency and low toxicity. In the drug product, the protein is expected to remain chemically and physically stable over time, as degradation and aggregation can decrease the efficacy of the protein-drug, or be regarded as potentially toxic. High protein concentrations may lead to self-association, phase separation, high viscosity, opalescence, gel formation or increased propensity for protein particle formation. In the PIPPI Horizon2020 ETN project, different scientific fields are combined to systematically investigate different formulation conditions and map the proteins´ physicochemical properties, the colloidal and conformational protein stability, to assess the intermolecular interactions for the investigated proteins and also, where possible to generate molecular models. To achieve this, all the data from the PhD students in the PIPPI project is collated in a large database to make it publicly accessible for the scientific community.

This PhD project is one of the PhD projects in the PIPPI consortium. The thesis is divided into three parts: In the first part, seven proteins were chosen from PIPPI protein library to perform stability studies, from which there were five monoclonal antibodies (mAbs), a human serum albumin and one lipase. Each protein was extensively studied under different physicochemical conditions using high-throughput techniques like nano differential scanning fluorimetry (nanoDSF) and isothermal chemical denaturation (ICD). These studies were used to choose conditions to perform structural and interaction studies using small angle X-ray scattering (SAXS). In the continuation of the first part, the molecular basis for protein-protein interactions at high concentration were studied by measuring the static structure factor in the second part. Four mAbs that could be obtained at high concentrations were chosen. SAXS and static light scattering (SLS) were performed to determine the structure factor, the second virial coefficient and thereby the nature of the intermolecular interactions. Further, the effect of NaCl was also studied for one of the mAbs. According to experimental data, all four mAbs were repulsive in nature and presence of NaCl screened the repulsion. The experimental behavior could be captured by two simplified theoretical hard sphere models containing long-range repulsion and short range attraction potentials. In the third part, interaction and ion binding studies were performed for two mAbs using SAXS, SLS and coarse-grained modeling. The SAXS and SLS data indicated that the antibodies behave differently in presence of NaCl. One mAb (PPI03) has repulsive interaction although increasing ionic strength screens part of the electrostatic repulsion, while the other mAb (PPI13) has additional attractive interactions even at low NaCl concentration. To study the Cl^- ion binding, metropolis monte carlo (MC) simulations were performed on the SAXS derived rigid body models. The MC simulation was used in an attempt to explain the different behavior in terms of charges and ion distributions around the mAbs by creating iso-density maps of Cl^- ions. The study showed that PPI03 has more positive patches compared to PPI13 due to which Cl^- ions accumulate more on PPI03. Moreover, this mechanism can be responsible for repulsion in case of PPI03 in presence of NaCl.

Dansk resumé

Proteinbaseret medicin bliver i stigende grad benyttet på grund af deres høje specificitet og potens og lave toksisitet. Produkterne skal være kemisk og fysisk stabile, fordi nedbrydning og aggregering kan formindske effekten eller fremkalde toksisitet. Høj proteinkoncentration kan føre til selvassociering, faseseperation, høj viskositet, gelatering eller til dannelse af proteinpartikler. "PIPPI Horizon2020 ETN"-projeketet kombinerer systematiske undersøgelser til at kortlægge proteinernes fysisk-kemiske egenskaber, stabilitet, og intermolekylære vekselvirkninger og, hvor det har været muligt, til at lave molekylære modeller. For at opnå dette er de samlede data fra alle de PhD studerende in PIPPI-projektet blevet samlet i en database, som vil blive offentligt tilgængelig.

Dette PhD-projekt er et af PhD-projekterne i PIPPI-konsortiet. Afhandlingen er opdelt i tre dele: I første del blev syv proteiner udvalgt fra PIPPI's proteinbibliotek: fem monoklonale antistoffer (mAbs), human serum albumin og en lipase. Hvert protein blev undersøgt ved brug af højeffektive ”high throughput” metoder som nano differential scanning fluorimetry (nanoDSF) og isotermisk kemisk denaturering (ICD). Resultaterne herfra blev brugt til at udvælge betingelser for små-vinkel røntgenspredning (SAXS). I anden del af afhandlingen undersøges den statiske strukturfaktor ved høje proteinkoncentrationer. Fire mAbs, som kunne fremskaffes ved høje koncentrationer, blev valgt. SAXS og static lysspredning (SLS) blev benyttet til at bestemme strukturfaktoren og den anden virialkoefficient og derved bestemme de intermolekylære interaktioner. Ydermere blev effekten af NaCl undersøgt for en enkelt mAb. Det eksperimentelle data viste, at alle fire mAbs var repulsive, og at NaCl skærmer denne frastødning. Eksperimenterne kunne forklares ved to simplificerede ”hard sphere” modeller med både repulsion og attraktion. I afhandlingens tredje del undersøgtes vekselvirkninger og ionbinding på to mAbs ved SAXS, SLS og course grain modellering. Data fra SAXS og SLS viser at antistofferne opfører sig forskelligt ved tilsætning af NaCl. Et mAb er repulsivt selv med stigende ionstyrke, mens det andet mAb viser øget attraktion selv ved lave NaCl-koncentrationer. Der blev lavet metropolis monte carlo (MC) simulering for at forklare den forskellige opførsel af ladninger og ionfordeling omkring disse mAbs Disse undersøgelser viste, Cl^- ioner ophober sig omkring de positive områder på proteinet.

Abbreviations

A(q)	amplitude of X-ray
Ace	acetate
B_{22}	osmotic second-virial coefficient
C	antibody constant region
$c_{½}$	concentration of denaturant required to unfold 50% of the protein
CDR	complementarity determining regions
CG-MC	Coarse Grained Monte Carlo
C_H	antibody heavy constant region
C_L	antibody light constant region
CMC	canonical Monte Carlo simulations
DLS	Dynamic Light Scattering
DLVO	Derjaguin−Landau−Verwey−Overbeek
D_{max}	maximum dimension
DSF	Differential Scanning Fluorimetry
F	fluorescence
Fab	fragment antigen-binding
Fc	fragment crystallizable region
FDA	Food and Drug Administration
FT	Fourier Transform
Fv	variable region
GuHCl	guanidine hydrochloride
His	histidine
HSA	Human Serum Albumin
$I(q)$	intensity of X-ray at q
ICD	Isothermal Chemical Denaturation
IgG	Immunoglobulin G
K	optical constant
M	molar (mol/L)
mAb	monoclonal antibody
MALS	Multi-Angle Light Scattering
M_W	molecular weight
N_A	avogadro number
$P(q)$	form factor
$p(r)$	pair-distance distance distribution
PD	polydispersity
PDB	Protein Data Bank
Phos	phosphate
pI	isoelectric point
P_i	penalty term
PMF	potential of the mean force
q	scattering vector (momentum transfer)
R	gas constant
$\bar{R}_\theta$	the excess Rayleigh ratio
R_G	radius of gyration
R_H	hydrodynamic radius
$S(q)$	structure factor
SAXS	small angle X-ray scattering

scFv	single chain variable region
SLS	Static Light Scattering
T	temperature
$T_{½}$	temperature of unfolding
T_{agg}	onset temperature of aggregation
T_m	melting temperature
Tris	trizma
V	antibody variable region
V_H	antibody heavy variable region
V_L	antibody light variable region
V_p	Porod volume
ΔG	variation of Gibbs free energy
$\Delta\rho$	electron density contrast
ε	extinction coefficient
$\partial n/\partial c$	refractive index increment
λ	wavelength
χ^2	minimization of discrepancy
υ	partial specific volume
τ	strength of the adhesive force between the particles
σ	Effective hard sphere radius
Z	Total charge

Table of Contents

1 Introduction

1.1 Introduction to protein therapeutics

Biopharmaceuticals, alternatively known as biotherapeutics or biologics are alternative pharmaceutical products that are derived from cell-based therapy, gene therapy and recombinant techniques[1]. Today, recombinant protein based therapeutics are an important class of medicines that have become an integral part of current practice for medical and health treatment. Since the approval of the first therapeutic protein, human insulin, in the 1980s, the pharmaceutical industry has been increasing its focus on recombinant DNA technology, which has led to a large number of approved biopharmaceuticals[2,3]. As of 2019, nearly 380 protein therapeutics have been approved by the US Food and Drug Administration (FDA), seeing protein based drugs making up approximately 10% of the entire therapeutic market[4]. Recently approved protein therapeutics have been used to treat several human disorders, such as cancer, autoimmunity/inflammation, diabetes, hemophilia and genetic disorders[5]. The molecular weight of protein therapeutics can be in a range of 6 kDa, insulin, to exceeding over 100 kDa which includes antibodies and large fusion proteins. Further, they exhibit complex conformations known as secondary, tertiary and quaternary structures, which must be maintained. The manufacturing and production of protein therapeutics are highly complex and challenging compared to small molecule drugs[6].

Protein formulation is an essential part of the development of therapeutic protein-based drug products. These often face several challenges, such as protein stability, which affects behavior during manufacturing, storage, conformation of the protein, solubility and the possibility of specific and non-specific self-association[7]. Protein stability is generally related to its physical stability (e.g. thermal, conformational, colloidal, aggregation, unfolding and others) and its chemical stability (e.g. oxidation, deamination, isomerization, chemical modification of amino acids and others)[8]. Maintaining the structure of the protein is imperative to maintaining proper function in the solution. There are several factors that play an important role for protein stability, such as buffer, pH, temperature and excipients (e.g. sugars, salts, amino acids, and denaturants)[9]. There is no specific criteria for choosing the above mentioned factors for formulation has been reported and yet remains as the expensive process in experimental drug development.

The typical volume for injectable drug products for subcutaneous administration is limited to a maximum volume of between 1 and 1.5 ml[10]. This limitation seems partly from the risk that the injected drug might be expelled by back pressure created in the subcutaneous tissue and partly due to

creation of excessive injection pain. This is why concentrated protein solutions are necessary to meet patients' dose requirements; however, protein solutions at high concentrations tend to self-associate, potentially leading to phase separation, high viscosity, opalescence, gel formation and the increased propensity for protein particle formation which makes it difficult/impossible to administer[11–14].

To address the issues involved in formulation by better understanding protein stability at the molecular level, a consortium named PIPPI (Protein-excipient Interactions and Protein-Protein Interactions in formulation) was designed[15]. The PIPPI protein library consists of 21 different proteins, representing different classes of biopharmaceuticals. All the proteins have been extensively studied using various biophysical techniques to investigate the physicochemical properties with a comprehensive understanding of conformational changes at the molecular level complemented by *in-silico* approach.

1.2 Properties influence solution behavior

Therapeutics can be administered subcutaneously, intravenously or orally, the focus here being on liquid formulations. During the development process, proteins are often screened using high-throughput techniques over a wider range of formulation conditions[16]. The physical stability of the protein depends on the physicochemical properties and on how the interdependent factors contribute to solution behavior the structural properties which are critical for stable formulations. The general factors are pH, charge, effect of high concentrations, temperature and co-solvents. These factors are described more extensively in the following sections.

1.2.1 pH and charge

Proteins' charges have an important role to play in determining their roles in protein-protein interactions. It may influence electrostatic interactions or may affect the charges of the neighboring atoms. The protonation state of a protein is changed by the association and dissociation of H^+ ions which is dependent on the overall pH of the solution, on the presence of salts, buffer and ionisable residues[9,17]. It is essential to have an accurate measurement of pH during the formulation process endeavoring to achieve protein stability. At a pH near to the isoelectric point (pI) of the protein, there will be no net charge across the protein, but patches of charges that are balanced by opposing charges elsewhere on the surface still exist. If pH is significantly higher or lower than pI, the protein will exhibit a net charge that favors intermolecular repulsion, normally, resulting in increased colloidal stability[18,19].

1.2.2 Effect of high concentrations

In high concentration protein formulation, concentration is typically in the ranges varying between 50 and 150 mg/ml[20]. Compared to other types of proteins, antibodies have been available for some time in the market at high concentrations[21]. Previous studies show that an antibody solution at 125 g/L is usually 60 times more viscous than the solution in the absence of protein[22].

Proteins at high concentrations have a larger probability of either native or non-native self-association. Native self-association leading to oligomerisation or clustering may cause phase separation and high viscosity[13]. These phenomena are mostly reversible but risk destabilizing the highly concentrated formulations, which contribute to protein stability in formulation. Non-native protein self-association

of the protein arising from non-specific protein-protein interaction leads to irreversible aggregation. The presence of non-native proteins causes loss of activity and may also cause immunogenicity[23].

There still are significant issues in applying the experimental techniques and insights into protein formulation development processes. Several techniques have been used traditionally to investigate protein solution behavior and stability[24]. However, when it comes to high concentration protein solutions, the choices are limited. We cannot simply extrapolate from a protein's low concentration behavior to its high concentration behavior due to the very complex behavior of proteins[25]. In this work, using small angle X-ray scattering (SAXS), static light scattering (SLS) and two simplified theoretical hard sphere models containing interaction potentials, we are able to investigate the physical properties of the proteins at high concentrations.

1.2.3 Temperature

The protein's structure and solution behavior are interrelated and strongly influenced by temperature[26–28]. In therapeutics, the effects of temperature, such as thermal denaturation caused by extreme temperature, have been described previously[27,28]. Moreover, even a moderate temperature increase will influence the protein, i.e., can influence the partial unfolding of a protein or the kinetic energy of the solution leading to aggregation. Generally, the thermal denaturation process is irreversible as a result of aggregation of unfolded proteins[29,30]. Accordingly, finding an accurate storage temperature for long term stability of biopharmaceuticals is a challenge[8]. For liquid formulations, storage temperatures between -10ºC and +50ºC currently are the commonly used storage temperature range[31].

Thermal denaturation of proteins has been well studied using several techniques (e.g. differential scanning calorimetry, differential scanning fluorescence and differential scanning light scattering and others)[32,33]. Multidomain proteins like antibodies show complex thermal denaturation processes which can be domain dependent due to specific stabilizing interactions between the domains[34]. The thermal stability of a protein can be affected by varying pH and ionic strength of a solution[35]. It is important, therefore, to find a stable pH and ionic strength where the protein's thermally stable and has long-term storage stability.

1.2.4 Co-solvents

Protein stability can be altered by addition of salt to the formulation. There are a wide range of different salts available and these can have different effects depending on the types and concentrations of the salts in solution[36]. Mainly, it affects the solubility, ion binding and protein charge[37]. The first protein stability study with specific ions was conducted by Hofmeister and the effects of salts were ranked in a series for their ability to precipitate proteins[37]:

Cations

$$(CH_3)_4N^+ > NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+} > Ba^{2+} > GdnH^+$$

Anions

$$CO_3^{2-} > SO_4^- > S_2O_3^{2-} > H_2PO_4^- > OAc^- > Cl^- > Br^- > NO_3^- > I^- > ClO_4^- > SCN^-$$

Depending on the interaction with water, ions are divided into two categories, i.e., kosmotropes and chaotropes. Here, chaotropes are the ions listed to the right of Cl^- and kosmotropes are listed to the left of Cl^- ion. Kosmotropes are highly hydrated in solution, are small in size and have a high charge density which leads to a reduction in the ability to solubilize the protein and enhances hydrophobic interactions[38]. On the contrary, chaotropes are poorly hydrated breaking the water structure and leading to an increased protein solubility[38].

At low ionic strength, the addition of lower concentrations of salts leads to an increase in protein stability, known as salting in. Generally occurring with chaotropes, the interaction between protein and salt leads to preferential adsorption of ions which causes repulsive interaction among the proteins resulting in increased solubility. As the salt concentration increases, known as salting-out, the solubility decreases, generally occurring with kosmotropes. Here, the interaction between ion and solvent is more favorable compared to that of protein-ion interactions leading to preferential exclusion of salt ions causing attractive non-specific interactions and thereby decreasing solubility. The salting-in and salting-out effects are highly dependent on the salt concentrations used and further affect protein stability. The Hofmeister series have a strong correlation between both salting-in and salting-out effects[39,40].

1.3 PIPPI project

The main goal of the PIPPI project is to achieve a better molecular understanding of protein stability in formulation. The PIPPI protein library consists of a diverse set of proteins of different folds and molecular weights. Figure 1.1 shows the composition of the PIPPI protein library. We applied a systematic approach to map the physicochemical properties of and we studied their stability as a function of the solution conditions. Stability studies of all the proteins have been performed under different formulation conditions by varying pH, NaCl, buffer species and excipients.

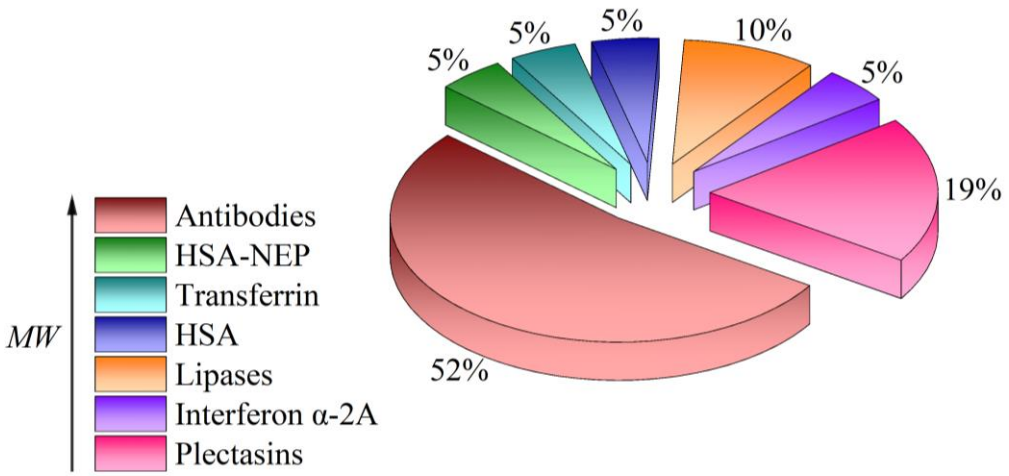


Figure 1.1 Protein library: Pie chart showing different classes of proteins with increasing molecular weight.

In therapeutics, stability studies for proteins may be approached from different aspects generally related to the physical and chemical stability[8,41]. Physical stability can be seen as conformational and colloidal, aggregation, while chemical stability is related mainly to oxidation, deamination and others[8,42]. A number of biophysical techniques are used to characterize different properties of proteins, represented schematically in Figure 1.2.

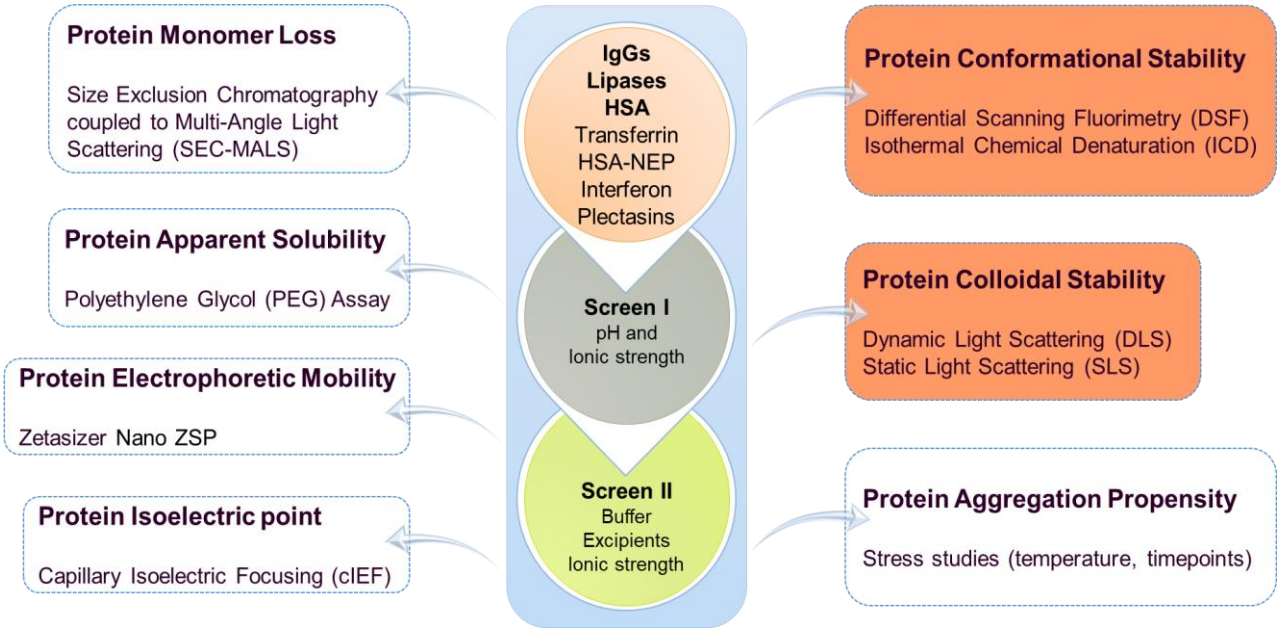


Figure 1.2 Schematic representation of stability studies and methods used.
Proteins and methods I have used are highlighted in bold and orange boxes, respectively.

Table 1.1 Extracted parameters from the instruments used in stability studies.

Methods	Extracted parameters
DSF	$T_{½}$ (apparent melting temperature)
ICD	ΔG (Gibbs free energy of unfolding), $c_{½}$ (concentration of denaturant required to unfold 50% of the protein), *m*-value
SLS	M_W (molecular weight) and B_{22} (second virial coefficient)
DLS	R_h (radius of hydration), *PD* (polydispersity), T_{agg} (onset temperature of aggregation), *kD* (interaction diffusion parameter)
SEC-MALS	Retention time, monomer loss, M_W, *PD*
PEG-assay	Turbidity midpoint
Zetasizer	zeta potential, particle size
cIEF	Isoelectric point

Various studies have been performed by different methods: conformational stability studies using differential scanning fluorimetry (DSF) and isothermal chemical denaturation (ICD); colloidal stability studies using static and dynamic light scattering techniques (SLS and DLS); aggregation propensity

using DLS by undertaking stress studies at different temperatures and time points; monomer loss of the protein using size exclusion chromatography coupled to multi-angle light scattering (SEC-MALS); apparent solubility using polyethylene glycol (PEG) assay, electrophoretic mobility by using zetasizer and isoelectric point calculated using capillary isoelectric focusing (cIEF). To complement the experimental data, *in-silico* approaches such as monte carlo (MC) and molecular dynamics (MD) simulations were incorporated. In Table 1.1, the extracted parameters from all the instruments are shown.

The protein stability studies have been divided into two parts: Screen I and Screen II (see Table 1.2). This work is a group effort that is distributed among the 15 PIPPI students. My contribution to the screening is to study the conformational and colloidal stability for five antibodies, one lipase and human serum albumin (HSA). For Screen I, measurements were performed in different chemical conditions by varying pH from 5.0 to 9.0 with 0.5 increments and in 3 different ionic strengths (0mM, 70 mM and 140 mM of NaCl). From pH 5.0 to pH 7.5, histidine buffer was selected and from pH 8.0 to pH 9.0, tris buffer was selected. In all, a total 27 formulation conditions have been chosen for Screen I. For Screen II, we study the effect of buffers and excipients, two pH values were selected for each protein, which are pH 5.0 and pH 6.5 for antibodies and human serum albumin, while pH 5.5 and pH 7.5 for lipase. For pH 5.0 and pH 5.5, acetate and histidine buffers were chosen and for pH 6.5 and pH 7.5, phosphate and histidine buffers were chosen. For the excipients study, three excipients were selected: sucrose, arginine, and proline. All the proteins were extensively characterized by using the above mentioned techniques. For structural studies, small angle X-ray scattering (SAXS) and nuclear magnetic resonance (NMR) measurements were performed on selected conditions from Screen I and Screen II. The totality of the collected data has been stored in the PIPPI database.

Table 1.2 Formulation conditions chosen for Screen I and Screen II measurements for antibodies, HSA and lipase.

Protein	**Screen I**	**Screen II**
Antibodies and human serum albumin	10 mM histidine for pH 5.0 to pH 7.5 10 mM tris for pH 8.0 to pH 9.0 NaCl: (0, 70, 140 mM)	10 mM histidine for pH 5.0 10 mM histidine for pH 6.5 10 mM acetate for pH 5.0 10 mM phosphate for pH 6.5 excipients: 140 mM NaCl with histidine, 280 mM sucrose, 140 mM arginine, 280 mM proline
Lipase	10 mM histidine for pH 5.0 to pH 7.5 10 mM tris for pH 8.0 to pH 9.0 NaCl: (0, 70, 140 mM)	10 mM histidine for pH 5.5 10 mM histidine for pH 7.5 10 mM acetate for pH 5.5 10 mM phosphate for pH 7.5 excipients: 140 mM NaCl with histidine, 280 mM sucrose, 140 mM arginine, 280 mM proline

My PhD project focused on the stability studies of the proteins in combination with structure and interaction studies at high protein concentrations. Chapter 3 contains the stability studies of monoclonal antibodies, lipase and human serum albumin using DSF and ICD and complemented by SAXS. Chapter 4 covers the structure and interaction studies of four monoclonal antibodies at high concentrations using SAXS and SLS measuring the static structure factor complemented by theoretical models with interaction potentials for structure factor fitting. In chapter 5, we performed SAXS and SLS on two monoclonal antibodies in combination with coarse-grained modeling to see the effect of ion binding. Chapter 6 contains the overall conclusion. Further investigation of macromolecular crowding in protein formulation were performed in presence of three polysaccharides (dextran sulfate 10, ficoll 70 and PEG 35000) as crowding agent. This investigation is not conclusive yet and have been added in the appendix to document it.

References

1. Halpern, W. & Hutto, D. *Biopharmaceuticals*. *Haschek and Rousseaux's Handbook of Toxicologic Pathology* (Elsevier, 2013). doi:10.1016/B978-0-12-415759-0.00025-X

2. Shah, D. K. Pharmacokinetic and pharmacodynamic considerations for the next generation protein therapeutics. *J. Pharmacokinet. Pharmacodyn.* **42**, 553–571 (2015).

3. Carton, J. M. & Strohl, W. R. *Protein therapeutics (introduction to biopharmaceuticals)*. *Introduction to Biological and Small Molecule Drug Research and Development: Theory and Case Studies* (Elsevier Ltd, 2013). doi:10.1016/B978-0-12-397176-0.00004-2

4. Usmani, S. S. *et al.* THPdb: Database of FDA-approved peptide and protein therapeutics. *PLoS One* (2017). doi:10.1371/journal.pone.0181748

5. Leader, B., Baca, Q. J. & Golan, D. E. Protein therapeutics: A summary and pharmacological classification. *Nature Reviews Drug Discovery* (2008). doi:10.1038/nrd2399

6. Carter, P. J. Introduction to current and future protein therapeutics: A protein engineering perspective. *Exp. Cell Res.* **317**, 1261–1269 (2011).

7. Otvos, L. & Wade, J. D. Current challenges in peptide-based drug discovery. *Front. Chem.* (2014). doi:10.3389/fchem.2014.00062

8. Manning, M. C., Chou, D. K., Murphy, B. M., Payne, R. W. & Katayama, D. S. Stability of protein pharmaceuticals: An update. *Pharm. Res.* **27**, 544–575 (2010).

9. Wang, W. Instability, stabilization, and formulation of liquid protein pharmaceuticals. *International Journal of Pharmaceutics* (1999). doi:10.1016/S0378-5173(99)00152-0

10. Narasimhan, C., MacH, H. & Shameem, M. High-dose monoclonal antibodies via the subcutaneous route: Challenges and technical solutions, an industry perspective. *Therapeutic Delivery* (2012). doi:10.4155/tde.12.68

11. Garidel, P., Kuhn, A. B., Schäfer, L. V., Karow-Zwick, A. R. & Blech, M. High-concentration protein formulations: How high is high? *Eur. J. Pharm. Biopharm.* **119**, 353–360 (2017).

12. Warne, N. W. Development of high concentration protein biopharmaceuticals: The use of platform approaches in formulation development. *Eur. J. Pharm. Biopharm.* **78**, 208–212

(2011).

13. Salinas, B. A. *et al.* Understanding and modulating opalescence and viscosity in a monoclonal antibody formulation. *J. Pharm. Sci.* (2010). doi:10.1002/jps.21797

14. Raut, A. S. & Kalonia, D. S. Pharmaceutical Perspective on Opalescence and Liquid-Liquid Phase Separation in Protein Solutions. *Mol. Pharm.* (2016). doi:10.1021/acs.molpharmaceut.5b00937

15. PIPPI (Protein-excipient Interactions and Protein-Protein Interactions in formulation). (2016). Available at: http://www.pippi.kemi.dtu.dk/.

16. Cheng, W. *et al.* Comparison of high-throughput biophysical methods to identify stabilizing excipients for a model IgG2 monoclonal antibody: Conformational stability and kinetic aggregation measurements. *J. Pharm. Sci.* (2012). doi:10.1002/jps.23076

17. Zhang, J. Protein-Protein Interactions in Salt Solutions. in *Protein-Protein Interactions - Computational and Experimental Tools* (2012). doi:10.5772/38056

18. Olsen, S. N., Andersen, K. B., Randolph, T. W., Carpenter, J. F. & Westh, P. Role of electrostatic repulsion on colloidal stability of Bacillus halmapalus alpha-amylase. *Biochim. Biophys. Acta - Proteins Proteomics* (2009). doi:10.1016/j.bbapap.2009.02.010

19. Saito, S. *et al.* Behavior of Monoclonal Antibodies: Relation between the Second Virial Coefficient (B 2) at Low Concentrations and Aggregation Propensity and Viscosity at High Concentrations. *Pharm. Res.* (2012). doi:10.1007/s11095-011-0563-x

20. Garidel, P., Kuhn, A. B., Schäfer, L. V., Karow-Zwick, A. R. & Blech, M. High-concentration protein formulations: How high is high? *Eur. J. Pharm. Biopharm.* **119**, 353–360 (2017).

21. Das, N. Commercializing high-concentration mAbs. *BioPharm International* (2016).

22. Liu, J., Nguyen, M. D. H., Andya, J. D. & Shire, S. J. Reversible self-association increases the viscosity of a concentrated monoclonal antibody in aqueous solution. *J. Pharm. Sci.* **94**, 1928–1940 (2005).

23. Mitragotri, S., Burke, P. A. & Langer, R. Overcoming the challenges in administering biopharmaceuticals: Formulation and delivery strategies. *Nature Reviews Drug Discovery* (2014). doi:10.1038/nrd4363

24. Filipe, V., Hawe, A., Carpenter, J. F. & Jiskoot, W. Analytical approaches to assess the degradation of therapeutic proteins. *TrAC - Trends in Analytical Chemistry* (2013). doi:10.1016/j.trac.2013.05.005

25. Saluja, A. & Kalonia, D. S. Nature and consequences of protein-protein interactions in high protein concentration solutions. *International Journal of Pharmaceutics* (2008). doi:10.1016/j.ijpharm.2008.03.041

26. Mirsky, A. E. & Pauling, L. On the Structure of Native, Denatured, and Coagulated Proteins. *Proc. Natl. Acad. Sci.* (1936). doi:10.1073/pnas.22.7.439

27. deWit, J. N. & Klarenbeek, G. Effects of Various Heat Treatments on Structure and Solubility of Whey Proteins. *J. Dairy Sci.* (1984). doi:10.3168/jds.S0022-0302(84)81628-8

28. German, B., Damodaran, S. & Kinsella, J. E. Thermal Dissociation and Association Behavior of Soy Proteins. *J. Agric. Food Chem.* (1982). doi:10.1021/jf00113a002

29. Sánchez-ruiz, J. M., López-lacomba, J. L., Cortijo, M. & Mateo, P. L. Differential Scanning Calorimetry of the Irreversible Thermal Denaturation of Thermolysin. *Biochemistry* (1988). doi:10.1021/bi00405a039

30. Milardi, D., La Rosa, C. & Grasso, D. Extended theoretical analysis of irreversible protein thermal unfolding. *Biophys. Chem.* (1994). doi:10.1016/0301-4622(94)00033-G

31. Porter, W. R. Thermally Accelerated Degradation and Storage Temperature Design Space for Liquid Products. (2012).

32. Goldberg, D. S., Bishop, S. M., Shah, A. U. & Sathish, H. A. Formulation development of therapeutic monoclonal antibodies using high-throughput fluorescence and static light scattering techniques: Role of conformational and colloidal stability. *J. Pharm. Sci.* (2011). doi:10.1002/jps.22371

33. Menzen, T. & Friess, W. Temperature-ramped studies on the aggregation, unfolding, and interaction of a therapeutic monoclonal antibody. *J. Pharm. Sci.* (2014). doi:10.1002/jps.23827

34. Vermeer, A. W. P. & Norde, W. The thermal stability of immunoglobulin: Unfolding and aggregation of a multi-domain protein. *Biophys. J.* (2000). doi:10.1016/S0006-3495(00)76602-1

35. He, F., Hogan, S., Latypov, R. F., Narhi, L. O. & Razinkov, V. I. High throughput thermostability screening of monoclonal antibody formulations. *J. Pharm. Sci.* (2010). doi:10.1002/jps.21955

36. Kohn, W. D., Kay, C. M. & Hodges, R. S. Salt effects on protein stability: Two-stranded α-helical coiled-coils containing inter- or intrahelical ion pairs. *J. Mol. Biol.* (1997). doi:10.1006/jmbi.1997.0930

37. Baldwin, R. L. How Hofmeister ion interactions affect protein stability. *Biophys. J.* (1996). doi:10.1016/S0006-3495(96)79404-3

38. Collins, K. D. Ions from the Hofmeister series and osmolytes: Effects on proteins in solution and in the crystallization process. *Methods* (2004). doi:10.1016/j.ymeth.2004.03.021

39. Curtis, R. A., Prausnitz, J. M. & Blanch, H. W. Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. *Biotechnol. Bioeng.* (1998). doi:10.1002/(SICI)1097-0290(19980105)57:1<11::AID-BIT2>3.0.CO;2-Y

40. Curtis, R. A., Ulrich, J., Montaser, A., Prausnitz, J. M. & Blanch, H. W. Protein-protein interactions in concentrated electrolyte solutions: Hofmeister-series effects. *Biotechnol. Bioeng.* (2002). doi:10.1002/bit.10342

41. Manning, M. C., Liu, J., Li, T. & Holcomb, R. E. Rational Design of Liquid Formulations of Proteins. *Adv. Protein Chem. Struct. Biol.* (2018). doi:10.1016/bs.apcsb.2018.01.005

42. Chi, E. Y., Krishnan, S., Randolph, T. W. & Carpenter, J. F. Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. *Pharmaceutical Research* (2003). doi:10.1023/A:1025771421906

2 Methods and Theory

2.1 Differential Scanning Fluorimetry

Differential scanning fluorimetry (DSF) is a biophysical technique employing thermal denaturation to assess protein thermal stability[1–3]. DSF applies a linear temperature gradient from 20ºC to 95ºC to cause protein unfolding. Most of the therapeutic proteins aggregate during thermal unfolding and the process is irreversible[4]. Due to the irreversibility of the unfolding process, the thermal denaturation methods provide apparent protein melting temperature (denoted as $T_{½}$), an approximation of the true protein melting temperature.

At ambient conditions, the proteins exist in a folded state, where the hydrophobic regions are buried in the core of the protein. When the protein unfolds, hydrophobic regions are exposed to the solvent. This is utilized in DSF, which was developed as an extrinsic dye-based method, which monitors the change in the fluorescence intensity of a fluorescent dye upon interactions with hydrophobic patches exposed during protein unfolding[5]. The dye used in this method is highly fluorescent in a nonpolar environment. During the thermal unfolding process, an increase in exposed hydrophobic residues leads to an increase in fluorescence.

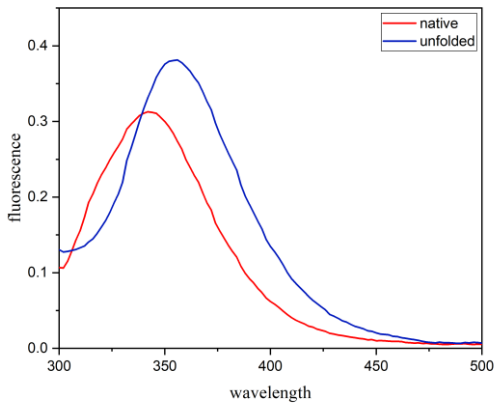


Figure 2.1 Fluorescence spectra of a native and unfolded protein.

DSF can be also used as a label-free technique monitoring the change in intrinsic fluorescence from Trp, Tyr and Phe when these are exposed to solvent during unfolding. NanoDSF is a label-free, native DSF technique that utilizes the native protein's intrinsic fluorescence to measure protein unfolding.

NanoDSF monitors the change in the intrinsic protein fluorescence intensity or peak maximum due to change in the environment of fluorescent amino acids (phenylalanine, tyrosine, and tryptophan) upon unfolding[6]. Generally proteins only show fluorescence in the tryptophan (290–350 nm) or tyrosine (280–340 nm) wavelength range. However, tyrosine has lower absorption coefficient, a lower quantum yield, and the emission at a lower wavelength, due to which tyrosine fluorescence can only be studied when there is very few or no tryptophan residuespresent[7]. In nanoDSF, the changes in fluorescence peak are monitored at two different emission wavelengths, normally 330 and 350 nm, where these wavelengths are in the range for the peak maxima for folded (330) or unfolded (350) proteins[8] (see Figure 2.1).

The analysis of thermal protein unfolding is based on the ratio between the intensities at the two emission wavelengths. The apparent melting temperature, $T_{½}$, is at the maximum in the first derivative of the ratio between fluorescence at 350nm and 330 nm (F350/F330), where approximately 50% of the protein is unfolded[2]. An example of nanoDSF unfolding curves is shown in Figure 2.2. The F350/F330 vs. temperature results in a sigmoidal curve in most cases. Moreover, multidomain proteins show multiple unfolding transitions due to multi-state unfolding processes exemplified by different thermal stabilities of individual domains within a protein.

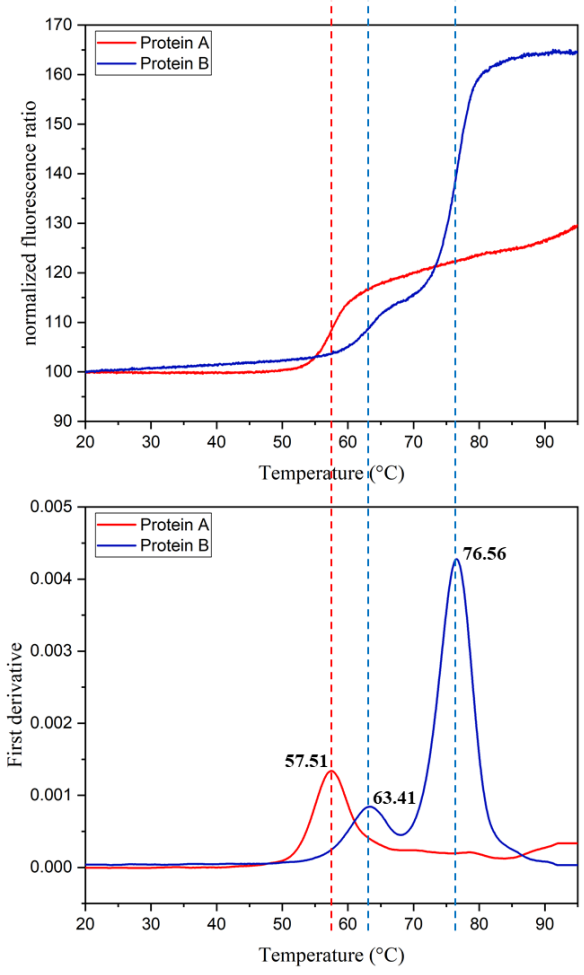


Figure 2.2 $T_{½}$ measurement. Upper panel, F350/330 fluorescence ratio intensity of intrinsic tryptophan plotted against temperature. Lower panel, $T_{½}$ calculation by first derivate analysis.

A Prometheus NT.48 instrument from NanoTemper Technologies GmbH, Munich, Germany was used to determine the thermal stability of therapeutic proteins in a variety of different formulations[9]. Prometheus NT.48 can measure 48 different experimental conditions in one run, with low sample consumption (< 10 μg) and low sample volume (~10 μL) regardless of choice of buffers and excipients[9]. Thermal unfolding was performed at a heating rate of 1°C/minute, resulting in a data point density of 10 points/°C. Due to high data point density, high-resolution measurements are obtained allowing precise fitting of the folded-unfolded transition by mathematical models[10].

2.2 Isothermal Chemical Denaturation

Isothermal chemical denaturation (ICD) is also used to study the physical/conformational stability of the protein[4,11]. Chemical denaturation is often times more reversible than thermal denaturation, due to which this technique can provide the thermodynamic parameter, Gibbs free energy of unfolding (ΔG), describing the protein conformational stability and the protein unfolding process[12]. In general, the folded protein is must be in equilibrium with a fraction of unfolded protein in a solution, described by an equilibrium constant, K_{eq}. The equilibrium constant is related to ΔG of unfolding:

$$\Delta G = -RTln(K_{eq}) \tag{2.1}$$

In ICD, the protein is formulated in the formulation of interest and mixed with increasing concentration of denaturant (guanidine hydrochloride or urea). Further, the mixture is incubated at a constant temperature until it reaches equilibrium and the intrinsic fluorescence is used to monitor unfolding. The maxima for folded and unfolded protein are measured and used to monitor unfolding (from these, the ratio can also be calculated). The ICD data is fitted to a suitable model[13,14] to extract the parameters, ΔG, m, $c_{½}$, see Figure 2.3. The parameters $c_{½}$ corresponds to concentration of denaturant needed to unfold 50% of the protein and m-values indicates the cooperativity of the unfolding and is correlated with the change in solvent accessible area of the protein during unfolding[15]. ΔG follows a linear dependence with the denaturant concentration:

$$\Delta G = \Delta G^{\circ} - mc_{denaturant} \tag{2.2}$$

where, ΔG° is the Gibbs free energy of protein unfolding in the absence of denaturant and the m-value is the slope of the line. The relation between the extracted parameters from the analyzed ICD data is:

$$c_{½} = \Delta G^{\circ}/m \tag{2.3}$$

In relation to formulation development, high ΔG represents the presence of a small fraction of unfolded protein, while low delta G represents the presence of a large fraction of unfolded protein, which are likely to form aggregates[16,17]. ICD can also be performed in the presence of extrinsic dye to follow the unfolding process by monitoring change in fluorescence. For multidomain proteins, the chemical denaturation process also shows multiple unfolding transitions, giving some insight into the stability of individual domains.

HUNK, AVIA ICD 2304 from Unchained Labs was used to perform chemical denaturation studies in a high throughput way. It is a fully automated instrument, which prepares denaturation curves from protein and formulation stocks, incubates the samples and measure the fluorescence.

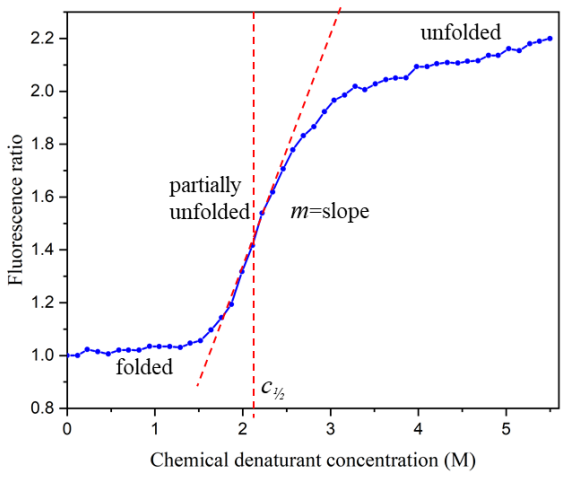


Figure 2.3 Protein unfolding curve obtained by isothermal chemical denaturation and the parameters extracted from the analysis.

2.3 Small Angle X-ray Scattering

Small angle X-ray scattering (SAXS) is a contrast method used to study the structural properties and interactions of biological molecules in solution[18]. Schematic representation of a SAXS experiment is shown in Figure 2.4. SAXS is performed using a monochromatic X-ray source, which is directed to the sample solution leading to scattering of the X-rays. The scattering intensity is recorded by the a detector as a function of momentum transfer vector q, i.e., $I(q)$. Due to the random orientations of the molecules in the solution, SAXS intensities is an average over all the positions in the particle. The relation between the scattering angle (2θ), momentum transfer (q) and the wavelength of the incident beam (λ) is defined as[18]:

$$q = \frac{4\pi \sin\theta}{\lambda} \tag{2.4}$$

For an isotropic solution, $I(q)$ is related to an ensemble average of the instantaneous scattering amplitudes of the protein particles $A(q)$ by:

$$I(q) = \frac{\langle A(\boldsymbol{q})A^*(\boldsymbol{q})\rangle}{V} \tag{2.5}$$

Where V is the scattering volume. The scattering amplitude is given by a sum over particles between the scattering length of individual scatters (b_i) and a phase factor related to the center of mass position $\boldsymbol{r}_i$ is defined by:

$$A(\boldsymbol{q}) = \sum_{\boldsymbol{i}} b_i \exp(-i\boldsymbol{q}\cdot\boldsymbol{r}_i) \tag{2.6}$$

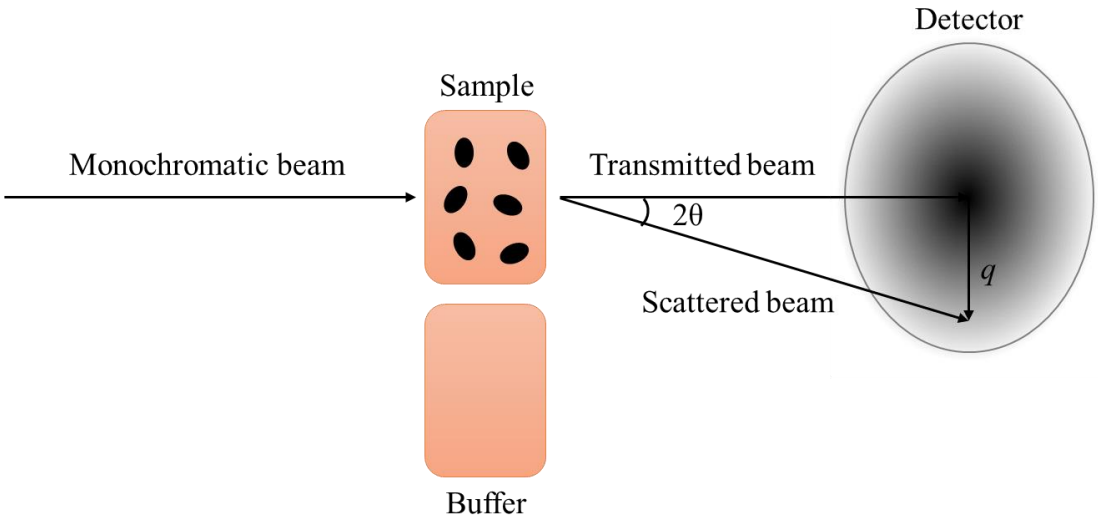


Figure 2.4 Schematic representation of SAXS: The incident X-ray beam is scattered by the particles in the sample by a scattering angle 2θ and the intensity is recorded on a detector. The change in the direction of scattered beam is compared to the incident beam by a momentum transfer vector, *q*. Buffer measurement were performed in between sample measurements.

From the illustration of Figure 2.4, we see that SAXS measurements includes separate sample and buffer measurements. The scattering profile for the protein is obtained by subtracting the buffer from the protein sample measurement. Since the difference in electron density, and therefore the difference in scattering power, between the buffer and the protein is quite small, it is extremely important that the buffer match the buffer of the protein sample[19]. Therefore, the buffers are always the last dialysis buffer from preparation of the sample.

The scattering intensity of a monodisperse solution without particle interference with the maximum particle diameter, D_{max} is related to the pair distance distribution function, *p(r)* by Fourier transform given by[20]:

$$I(q) = 4\pi \int_0^{Dmax} p(r) \frac{4\pi \sin(qr)}{qr} \mathrm{d}r \tag{2.7}$$

Likewise, *p(r)* can be determined from the inverse Fourier transformation of the intensity, *I(q)*:

$$p(r) = \frac{r^2}{2\pi^2} \int_0^{\infty} q^2 I(q) \frac{\sin(qr)}{qr} \mathrm{d}q \tag{2.8}$$

Preliminary SAXS analysis is often performed by looking at the Guinier approximation or Pair-distance distribution, *p(r)*, function (using Equation 2.9). From these several parameters can be calculated: molecular weight (*MW*), radius of gyration (R_G) and D_{max}. The Guinier approximation is only valid at low angles and is true up to $qR_G < 1.3$[18]. It is derived by Taylor series expansion of *I(q)*, and can be extrapolated to $q = 0$ by Equation 4.

$$I(q) \cong I(0)\, exp\left[\frac{-(q^2 * R_G^2)}{3}\right] \tag{2.9}$$

After calculating *I(0)*, M_W can be calculated in absolute scale using Equation 2.10:

$$M_W = \frac{N_A I(0)/c}{\Delta\rho_M^2} \tag{2.10}$$

where N_A is Avogadro constant, $\Delta\rho_M$ is the scattering contrast per mass. $\Delta\rho_M$ was calculated using proteins average partial specific volume of $0.7425\ cm^3 g^{-1}$ by Mylonas and Svergun[21]. For Guinier approximation, only the first part of the scattering curve is used to calculate the M_W and R_G, while the full scattering curve is used to calculate D_{max}, M_W and R_G from *p(r)* function.

Another way of analyzing SAXS data is by Kratky plot (see Figure 2.5), which tells about the conformation and flexibility of the protein.

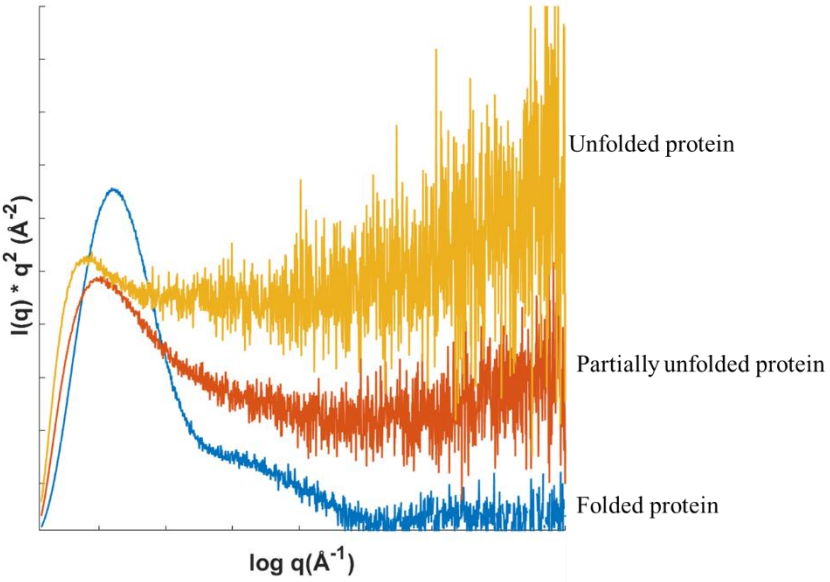


Figure 2.5 Kratky plot for folded globular protein (blue), Partially unfolded protein (orange) and unfolded protein (yellow). Curvature is depending on molecular shape, degree of flexibility.

SAXS data also provide information about the particle distribution in terms of the structure factor, *S(q)*[22]. The total scattering intensity, $I(q)$, is the results of $P(q)$ multiplied by from $S(q)$, where *P(q)* is the form factor that depends on the particle structure. The $I(q)$ can be written as:

$$I(q) \propto P(q) \cdot S(q) \tag{2.11}$$

There are two types of effects which can describe the molecular interactions; one is intra-particle effects describes by the form factor, *P(q)*, is averaged over size and orientation of the scattering particle, while another one is inter-particle interaction describes by the structure factor, *S(q)*, such as repulsion and attraction. At infinite dilution, if *S(q)* = 1, and no inter-particle interaction is present[23], we get the *P(q)* by merging the low protein concentration data of a particular sample, which is the form factor of the sample. This was done in order to obtain the structure factor, *S(q)*, with varying concentrations which can be obtained by dividing *I(q)* by *P(q)*. If the scattering intensity is decreasing with increasing protein concentration and *S(q)* < 1, it is an indication of repulsion in the system. Moreover, if the intensity is increasing with increasing protein concentration and *S(q)* > 1, it is an indication of attraction in the system. The schematic illustration is presented in Figure 2.6.

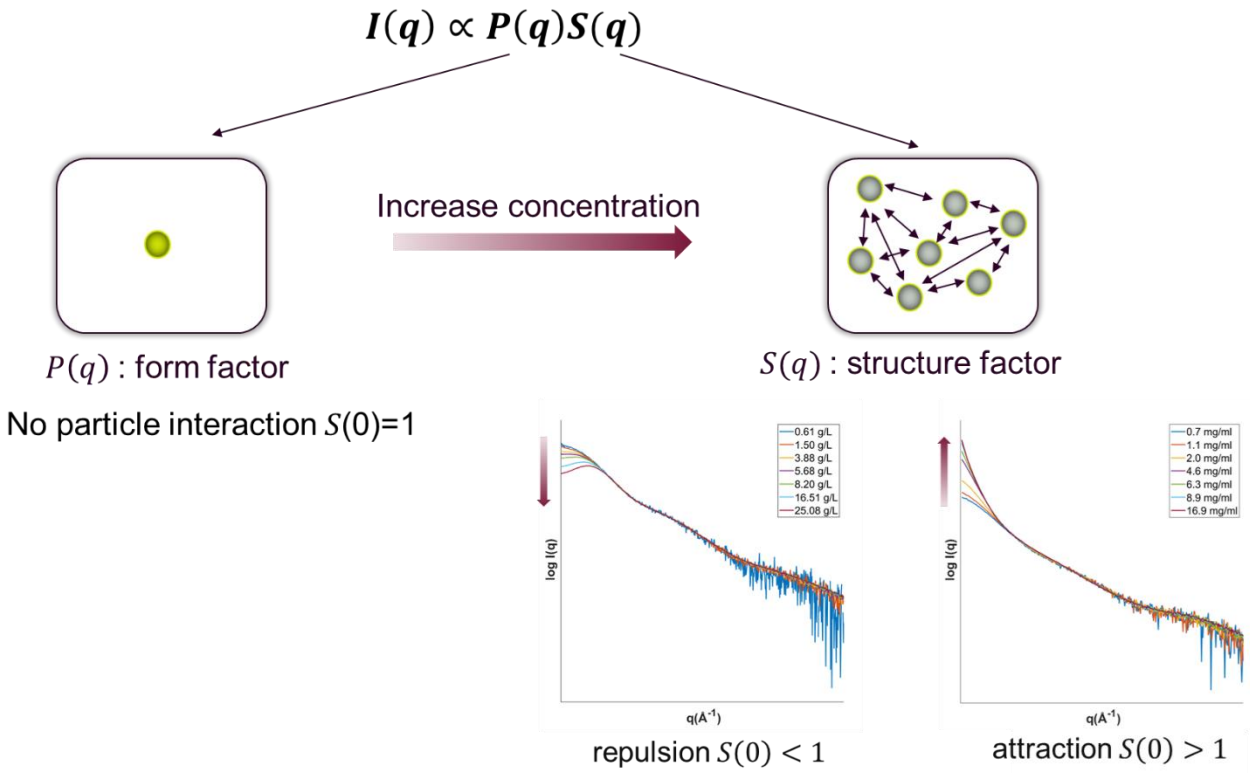


Figure 2.6 The solution intensity is the product of form factor (***P*(*q*)**) and structure factor (***S*(*q*)**). *S(q=0)* is not directly accessible from SAXS due to the beamstop, so it can be calculated by extrapolation of *S(q)* curves.

SAXS modelling

SAXS data can be used to perform rigid body modelling or *ab* initio modelling, when high quality data are available[24]. One challenge for SAXS modelling is, we can get similar scattering profiles for different models[20,25]. The ambiguity can be reduced if we have the information about protein sequence, high resolution model or homology model and include this information while modelling. In SAXS based modelling, we optimize a set of parameters, which describe the model by minimization of discrepancy (χ^2) between the experimental and modelled data. χ^2 can be calculated by using Equation 2.12, where I_{exp} is the experimental intensities, I_{calc} is the calculated intensities form the model, N is the number of data points, σ is the experimental errors and c is the scaling factor[26].

$$\chi^2 = \frac{1}{N-1}\sum_{i=1}^{N}\left[\frac{I_{exp}(q_i) - c\, I_{calc}(q_i)}{\sigma(q_i)}\right]^2 \tag{2.12}$$

The acceptable value for χ^2 is around 1, but it highly depends the quality of the data. If the scattering data are noisy, then the χ^2 value will be smaller. To reduce the ambiguity, it is important that penalties are employed in SAXS based modelling, This means that a target function, E is minimized, where E is the sum of χ^2 and P_i (penalty term) that are weighted by α:

$$E = \chi^2 + \sum \alpha_i P_i \tag{2.13}$$

Ab-initio modelling

Ab-initio modelling can be used to for low resolution protein model reconstruction without any knowledge of the structure. It generates an approximate three dimensional shape of the protein. There are two common approach to do this modelling: bead modelling and dummy residues modelling[26]. DAMMIF[27] and DAMMIN[28] are two bead modelling program and GASBOR is the dummy residue modelling program from ATSAS[29] software package. In the bead modelling approach, the initial search volume for solute and solvent consist of beads, the simulated annealing tries to minimize E (see Equation 2.13). In dummy residues modelling, each C_α atom of the protein as dummy residue and also during the simulated annealing search, the position of the residues changes to minimize E.

Rigid body modelling

Rigid body modelling can be performed for multidomain proteins, complexes. if the high resolution structure of the individual subunits are available[26]. In this thesis CORAL[30] is used to perform the rigid body modelling. CORAL stands for complexes with random loops. In this case flexibility of the protein is taken into account. The program can translate and rotate the individual domains of a multidomain protein. If there is any missing fragments, CORAL employs the pre-generated library of loops. The program generates 20 structures considering every possible end to end distances and provides the best fitted model with the experimental data. To validate the model, CRYSOL[31] program is used to generate the theoretical scattering curve of the model, which is then compared with our experimental scattering curve. An example of CORAL model is shown in Figure 2.7.

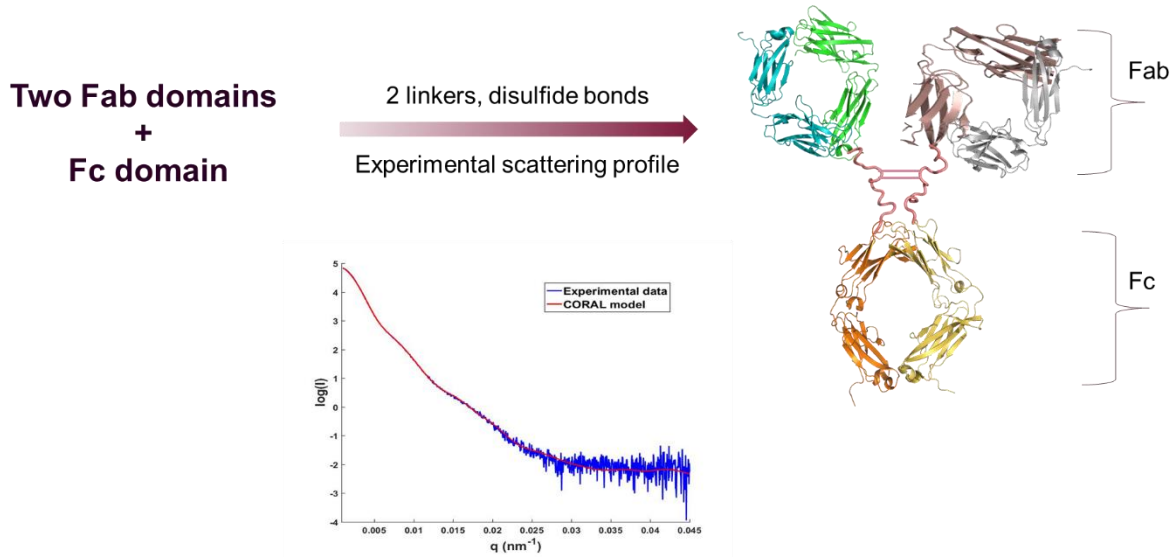


Figure 2.7 Representation of a CORAL modelling of an antibody by using the crystal structure of domains. A fitting plot of SAXS scattering curve and CORAL model using CRYSOL was shown in bottom panel.

2.4 Static Light Scattering

Static light scattering (SLS) can be used to calculate the molecular weight and size of the protein. It can also be used to measure interactions in terms of second virial coefficients, B_{22}, to observe the protein solution properties[32]. SLS measurements relates to the intensity of the scattered light, the

excess Rayleigh ratio, $\bar{R}_{\theta,}$ is the measured quantity of an SLS experiment. $\bar{R}_{\theta}$ is related to the light scattered by the sample at a fixed protein concentration, scattering from the solvent and osmotic compressibility of the solution is given by Equation 2.14 [33,34].

$$\frac{Kc(\partial n/\partial c)^2}{\bar{R}_{\theta}} = S(0)^{-1} = \frac{1}{M_{\mathrm{w}}RT}\left(\frac{\partial \Pi}{\partial \rho}\right) \tag{2.14}$$

Here M_{w} is the protein average molecular weight, *S(0* is the structure factor estimated at q = 0, *R* is the gas constant, *T* is the temperature, $\partial n/\partial c$ is the refractive index increment of the protein solution, c is protein concentration, $\partial \Pi/\partial \rho$ is the osmotic compressibility which is related to the structure factor and *K* is the optical constant which is equal to $2\pi^2 n_0^2/(N_{\mathrm{A}}\lambda^4)$, where n_0 is the refractive index of the solvent and N_A is Avogadro's number.

The osmotic pressure, Π, of the solution can be expanded in a virial expansion, where the second order term is the osmotic second virial coefficient denoted by B_{22}. The osmotic compressibility can be related to B_{22} by taking the derivative of the virial expansion with respect to protein concentration, and the limit lies until low protein concentration.

$$\frac{\Pi}{RT} = \rho + B_{22}\rho^2 + higher\ order\ terms \tag{2.15}$$

The Rayleigh ratio can also be related to B_{22} according to:

$$\frac{Kc(\partial n/\partial c)^2}{\bar{R}_{\theta,}} = \frac{1}{M_{\mathrm{w}}} + 2B_{22}c \tag{2.16}$$

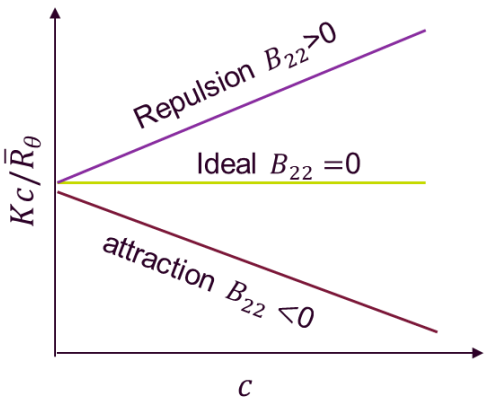


Figure 2.8 Determination of the interaction parameter B_{22} from static light scattering.

B_{22} can be determined from the slope of the linear fit to Equation 2.16 by measuring at different protein concentration and M_W is the inverse of the *y*-intercept. B_{22} is a measure of non-ideality of the solution or solute-solute interaction in a specific solvent. A positive B_{22} value indicates that the interaction of the macromolecules is repulsive while a negative B_{22} value indicates that the interaction is attractive (see Figure 2.8). SLS measurements can be also used as an alternative approach for measuring the structure factor at q = 0 (see Equation 2.14).

2.5 Dynamic Light Scattering

Dynamic light scattering (DLS) is a technique for particle size analysis based on Brownian motion of particles in solution. In a solution, particles are constantly colliding with the solvent molecules due to which there is constant movement of the particles. DLS measures the total light scattering intensity over time to monitor the particle movement. Due to the particle movements, the scattered intensity fluctuate and not constant over time. Larger particles moves slowly compared to smaller particle in the solution and shows slower fluctuations. While measuring the intensity fluctuations, we can get the a correlation function (see Figure 2.9), which can be used to determine the translational diffusion coefficient[35].

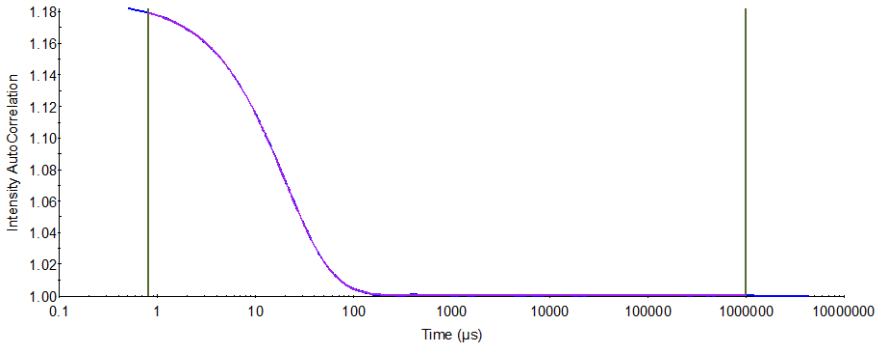


Figure 2.9 Autocorrelation function derived from the change in measured intensity

The diffusion coefficient can be interpreted as the hydrodynamic radius, R_h, of the particle and is directly proportional to the exponential decay rate of the autocorrelation function via the Stokes-Einstein equation:

$$R_h = \frac{K_B T}{6\pi\eta D} \tag{2.17}$$

Here D is the diffusion coefficient, K_B is the Boltzmann constant, T is the absolute temperature in kelvin and η is the viscosity of the solvent. Here, the particle is assumed to be a hard sphere, the hard sphere is diffusing at the same rate as the particle of interest.

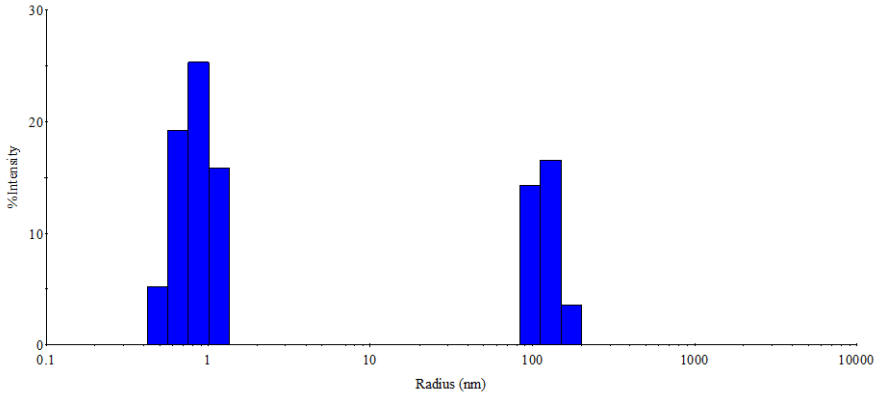


Figure 2.10 Size distribution profile of a multiple decay data.

From DLS measurements, we can also get the information size distribution profile of the particle of interest and give information whether the solution is mono or multimodal, by fitting the multiple exponential time constants to the autocorrelation function[36]. The size distribution histogram gives the information about the number of species present in the solution in combination with their mean size by measuring the relative intensity scattered (% intensity) by each population (see Figure 2.10). The DLS data can be also used to describe the polydispersity of the species, estimated relative amount of mass (% mass) or number of particles (% number)[37].

2.6 Protein-Protein Interaction Model

In chapter 4, two simplified protein-protein interaction model has been used to capture the behavior at high protein concentrations. The model is a hard sphere where the protein pair potential mean force, *w(r)*, is used to provide the input to predict the thermodynamic properties, where *r* is the center to center separation between proteins. The protein pair potential mean force represents an interaction free energy averaged over the orientation between the solvents degree of freedom and pair of proteins. In the potential mean force models, protein-protein interaction is described using an excluded volume potential and the electric double-layer force from Derjaguin−Landau−Verwey−Overbeek (DLVO) theory for electrostatic contributions[38–40]. In this approach, protein is treated as a uniformly charged sphere immersed in a dielectric continuum containing point charges. An approximation for two body interaction free energy is shown in equation 4.5:

$$\beta w(r) = \begin{cases} \infty & r < \sigma \\ \dfrac{Z^2 \lambda_B}{(1+\kappa\sigma/2)^2} \dfrac{\exp[-\kappa(r-\sigma)]}{r} & r > \sigma \end{cases} \tag{2.18}$$

Z corresponds to the protein valency, σ is an effective hard sphere radius, β is the inverse temperature $1/K_bT$ where K_b is Boltzmann's constant, λ_B is the Bjerrum length $\lambda_B = e^2/(4\pi\beta\varepsilon\varepsilon_0)$. Bjerrum length corresponds to the distance between a pair of ions under the condition when coulomb energy is equal to the thermal energy. The inverse Debye− Hückel screening length κ controls the range of the electrostatic interactions:

$$\kappa = \sqrt{2e^2 N_A I/(\beta\varepsilon\varepsilon_0)} \tag{2.19}$$

I is the ionic strength of the solution, *e* is the electronic charge, ε_0 is the vacuum permittivity, ε is the dielectric constant of water, and N_A is Avogadro's number. λ_B is the Bjerrum length - the separation between a pair of ions for Coulomb energy equal to thermal energy.

In this study, firstly we used the hard sphere model with repulsive yukawa potential, which we call as Yukawa model in this chapter. This model accounts the repulsion using Yukawa interaction potential, where the Yukawa potential has the same mathematical form as the electric double layer potential mentioned in Equation 4.5[39]. Yukawa model uses a combination of a hard sphere repulsion (for radii below the hard sphere radius, σ) in combination with an electrostatic repulsion (for radii above σ). This means that in the Yukawa model no attractive terms are present.

To account the short range attraction term, here we use an adjustable parameters to reflect the adhesive forces between the surfaces, by using Baxter adhesive potential, $w(r)^{sr}$. The adhesive potential

corresponds to a square well potential taken in the limit when the range or width of square well goes to zero. In this model, only one parameter is needed to describe the interaction, which is tau, τ, corresponds to the strength of the adhesive force between the particles. In this limit, the contribution of the attractive interaction to the virial coefficient becomes $B_{22}^{hs}/4\tau$, where the adhesion is proportional to 1/tau.

$$B_{22}^{sr} = -\frac{B_{22}^{hs}}{4\tau} \tag{2.20}$$

In the second model, which we call as Baxter model in this chapter, there we have included the repulsive yukawa potential in combination with baxter adhesive potential to account for both repulsion and attraction for protein-protein interaction. It has also been previously studied to describe the compressibility curves of protein[41].

We can calculate the structure factor from SAXS for an isotropic system, the *S(q)* is related to the Fourier transform of the pair distribution function *g(r)* by

$$S(q) = 1 + 4\pi\rho \int \mathrm{d}r \frac{r \sin(qr)}{q}[g(r) - 1] \tag{2.21}$$

Here ρ is the protein density and q is the momentum transfer. The pair distribution function corresponds to the normalized density for the centers of protein molecules in a spherical shell located at *r* with volume $4\pi r^2$ about a protein molecule fixed at the origin. By using the Ornstein-Zernike equation[42], we can calculate the *S(q)* in terms of interaction potential.

References

1. Malik, K., Matejtschuk, P., Thelwell, C. & Burns, C. J. Differential scanning fluorimetry: Rapid screening of formulations that promote the stability of reference preparations. *J. Pharm. Biomed. Anal.* (2013). doi:10.1016/j.jpba.2013.01.006

2. Senisterra, G., Chau, I. & Vedadi, M. Thermal denaturation assays in chemical biology. *Assay and Drug Development Technologies* (2012). doi:10.1089/adt.2011.0390

3. He, F., Hogan, S., Latypov, R. F., Narhi, L. O. & Razinkov, V. I. High throughput thermostability screening of monoclonal antibody formulations. *J. Pharm. Sci.* (2010). doi:10.1002/jps.21955

4. Freire, E., Schön, A., Hutchins, B. M. & Brown, R. K. Chemical denaturation as a tool in the formulation optimization of biologics. *Drug Discov. Today* **18**, 1007–1013 (2013).

5. Menzen, T. & Friess, W. High-throughput melting-temperature analysis of a monoclonal antibody by differential scanning fluorimetry in the presence of surfactants. *J. Pharm. Sci.* (2013). doi:10.1002/jps.23405

6. Garidel, P., Hegyi, M., Bassarab, S. & Weichel, M. A rapid, sensitive and economical assessment of monoclonal antibody conformational stability by intrinsic tryptophan fluorescence spectroscopy. *Biotechnol. J.* (2008). doi:10.1002/biot.200800091

7. Lakowicz, J. R. *Principles of fluorescence spectroscopy*. *Principles of Fluorescence*

Spectroscopy (2006). doi:10.1007/978-0-387-46312-4

8. Duy, C. & Fitter, J. How aggregation and conformational scrambling of unfolded states govern fluorescence emission spectra. *Biophys. J.* (2006). doi:10.1529/biophysj.105.078980

9. Magnusson, A. O. *et al.* nanoDSF as screening tool for enzyme libraries and biotechnology development. *FEBS J.* (2019). doi:10.1111/febs.14696

10. Augustijn, D. *et al.* Novel Non-linear Curve Fitting to Resolve Protein Unfolding Transitions in Intrinsic Fluorescence Differential Scanning Fluorimetry. *Eur. J. Pharm. Biopharm.* (2019). doi:10.1016/j.ejpb.2019.06.001

11. Pace, C. N., Shirley, B. A. & Thomson, J. A. Measuring the conformational stability of a protein. *Protein Struct. A Pract. approach* (1997).

12. Schellman, J. A. Temperature, stability, and the hydrophobic interaction. *Biophys. J.* (1997). doi:10.1016/S0006-3495(97)78324-3

13. Pace, C. N. & Shaw, K. L. Linear extrapolation method of analyzing solvent denaturation curves. *Proteins Struct. Funct. Genet.* (2000). doi:10.1002/1097-0134(2000)41:4+<1::aid-prot10>3.3.co;2-u

14. Niklasson, M. *et al.* Robust and convenient analysis of protein thermal and chemical stability. *Protein Science* (2015). doi:10.1002/pro.2809

15. Myers, J. K., Nick Pace, C. & Martin Scholtz, J. Denaturant m values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding. *Protein Sci.* (1995). doi:10.1002/pro.5560041020

16. Chi, E. Y., Krishnan, S., Randolph, T. W. & Carpenter, J. F. Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. *Pharmaceutical Research* (2003). doi:10.1023/A:1025771421906

17. Clarkson, B. R., Schön, A. & Freire, E. Conformational stability and self-association equilibrium in biologics. *Drug Discov. Today* **21**, 342–347 (2016).

18. Svergun, D. I. & Koch, M. H. J. Small-angle scattering studies of biological macromolecules in solution. *Reports Prog. Phys.* (2003). doi:10.1088/0034-4885/66/10/R05

19. Chaudhuri, B., Muñoz, I. G., Qian, S. & Urban, V. S. *Biological Small Angle Scattering: Techniques, Strategies and Tips*. *Biological Small Angle Scattering: Techniques, Strategies and Tips* **1009**, (Springer Singapore, 2017).

20. Putnam, C. D., Hammel, M., Hura, G. L. & Tainer, J. A. X-ray solution scattering (SAXS) combined with crystallography and computation: Defining accurate macromolecular structures, conformations and assemblies in solution. *Quarterly Reviews of Biophysics* (2007). doi:10.1017/S0033583507004635

21. Mylonas, E. & Svergun, D. I. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. in *Journal of Applied Crystallography* (2007). doi:10.1107/S002188980700252X

22. Svergun, D. I., Koch, M. H. J., Timmins, P. A. & May, R. P. *Small Angle X-Ray and Neutron Scattering from Solutions of Biological Macromolecules*. (2013).

doi:10.1093/acprof:oso/9780199639533.001.0001

23. Fritz, G. & Glatter, O. Structure and interaction in dense colloidal systems: Evaluation of scattering data by the generalized indirect Fourier transformation method. *J. Phys. Condens. Matter* (2006). doi:10.1088/0953-8984/18/36/S14

24. Trewhella, J. *et al.* 2017 publication guidelines for structural modelling of small-angle scattering data from biomolecules in solution: An update. *Acta Crystallogr. Sect. D Struct. Biol.* (2017). doi:10.1107/S2059798317011597

25. Vestergaard, B. Analysis of biostructural changes, dynamics, and interactions - Small-angle X-ray scattering to the rescue. *Arch. Biochem. Biophys.* (2016). doi:10.1016/j.abb.2016.02.029

26. Tuukkanen, Maxim V. Petoukhov, A. SAS-Based Structural Modelling and Model Validation. in *Biological Small Angle Scattering: Techniques, Strategies and Tips* 87–105 (2017).

27. Franke, D. & Svergun, D. I. DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* (2009). doi:10.1107/S0021889809000338

28. Svergun, D. I. Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.* (1999). doi:10.1016/S0006-3495(99)77443-6

29. Franke, D. *et al.* ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* (2017). doi:10.1107/S1600576717007786

30. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* (2012). doi:10.1107/S0021889812007662

31. Svergun, D., Barberato, C. & Koch, M. H. CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* (1995). doi:10.1107/S0021889895007047

32. Murphy, R. M. Static and dynamic light scattering of biological macromolecules: What can we learn? *Curr. Opin. Biotechnol.* (1997). doi:10.1016/S0958-1669(97)80153-X

33. Casassa, E. F. & Eisenberg, H. Thermodynamic Analysis of Multicomponent Solutions. *Adv. Protein Chem.* (1964). doi:10.1016/S0065-3233(08)60191-6

34. Vrij, A. & Overbeek, J. T. G. Scattering of light by charged colloidal particles in salt solutions. *J. Colloid Sci.* (1962). doi:10.1016/0095-8522(62)90023-5

35. Minton, A. P. Recent applications of light scattering measurement in the biological and biopharmaceutical sciences. *Analytical Biochemistry* (2016). doi:10.1016/j.ab.2016.02.007

36. Kuebler, S. Characterizing stable protein formulations. *Genet. Eng. Biotechnol. News* (2007).

37. Wyatt Technology Corporation. *Dynamics User's Guide*. (2010).

38. Zhang, F. *et al.* Protein interactions studied by SAXS: Effect of ionic strength and protein concentration for BSA in aqueous solutions. *J. Phys. Chem. B* (2007). doi:10.1021/jp0649955

39. Liu, Y., Chen, W. R. & Chen, S. H. Cluster formation in two-Yukawa fluids. *J. Chem. Phys.* (2005). doi:10.1063/1.1830433

40. Curtis, R. A., Prausnitz, J. M. & Blanch, H. W. Protein-protein and protein-salt interactions in

aqueous protein solutions containing concentrated electrolytes. *Biotechnol. Bioeng.* (1998). doi:10.1002/(SICI)1097-0290(19980105)57:1<11::AID-BIT2>3.0.CO;2-Y

41. 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). doi:10.1021/jp1028646

42. Høye, J. S. & Blum, L. Solution of the Yukawa closure of the Ornstein-Zernike equation. *J. Stat. Phys.* (1977). doi:10.1007/BF01013184

3 Stability and self-interactions of monoclonal antibodies, albumin and lipase

3.1 Introduction

Biotherapeutics are believed to represent the next generation of pharmaceutical products[1,2]. In the 1980s, therapeutic recombinant insulin[3] was developed and since then, the importance of biotherapeutics has been significantly increasing. Although many different proteins are used as therapeutics, this chapter focuses on monoclonal antibodies, human serum albumin and lipase. Protein stability can be studied from using various aspects[4,5], i.e., physical stability (e.g. colloidal, conformational, aggregation, unfolding and others) and chemical stability (e.g. oxidation, deamination, isomerization and others). Formulation conditions (pH, ionic strength, excipients) that mainly have an impact on physical stability might adversely effect of chemical stability. In this study, we are focusing on the physical stability of proteins, which was investigated by thermal and chemical denaturation studies in different physicochemical conditions, complemented by SAXS.

3.1.1 Monoclonal antibodies

Muronomab was the first monoclonal antibody approved by FDA in 1985 under the name of Orhtoclone OKT3[6] treating cardiac, hepatic and combined kidney-pancreas transplants. Monoclonal antibodies have been used for the previous 30 years[7] and their clinical application spectrum is growing constantly[8,9].

Monoclonal antibodies (mAbs) play an important role in the human immune system. They originate from a single primordial gene[10], has a specific antigen-binding site, which helps to recognize and remove foreign objects. Antibodies are glycoproteins, which lie within the larger immunoglobulin superfamily. The five classes of antibody superfamilies' found in serum are IgA, IgM, IgD, IgE and IgG[11]. IgA produced in mucosal membranes represents 15% of total immunoglobulins produced throughout the body; IgM produced by vertebrates; IgD is found in blood serum, representing 0.25% of immunoglobulins in serum; IgE found in mammals, utilized during immune defense against parasites; IgG is the most common type of antibody found in human blood circulation, represents 75% of human serum antibodies.

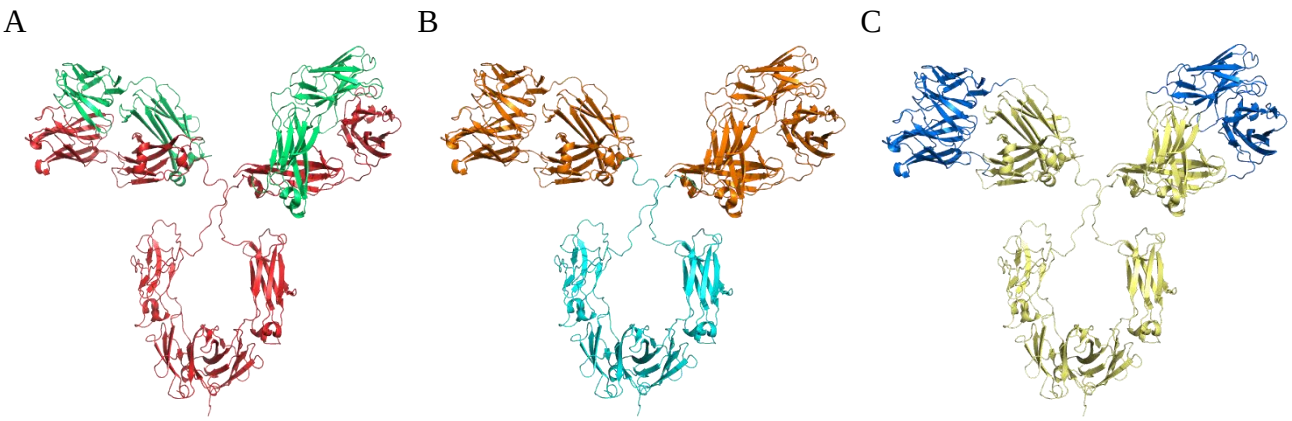


Figure 3.1 IgG structure representation: Panel A shows the four chains where the two heavy chains are shown in red and the two light chains in green. Panel B shows the functional subunits namely fragment antigen-binding (Fab) domains (orange) and fragment crystallizable (Fc) domain (sky blue) where Fab domains contain the antigen-binding sites (shown in blue in Panel C). Structures are based on the PDB entry 1HZH[12] and made in PyMOL[13].

This study focuses on the IgG type antibody. The IgG antibody is a multidomain protein (~150 kDa) (see Figure 3.1(A)), consisting of four polypeptide chains: two heavy (~50 kDa) and two light chains (~25 kDa)[14]. It can be divided into functional subunits two fragments antigen-binding (Fab) domains and one fragment crystallizable Fc domain (see Figure 3.1(B, C)). The three individual domains contain two folded β sheets. There are four disulfide bonds present in an IgG antibody molecule: two of them are connecting the heavy and light chains and two are within the hinge region (see Figure 3.2). The hinge region provides flexibility to the antibody and is necessary for performing immunological functions.

IgG can be divided into four subclasses, i.e., IgG1, IgG2, IgG3 and IgG4[15] depending on their biological properties, immunological function and class of heavy chain. Additionally, there has been further division based on the light chain class, namely kappa (κ) or lambda (λ). These different classes and sub-classes take into account their size, flexibility and immune response of the antibody[15] and therefore while considering the solution behavior, these properties should be taken into consideration.

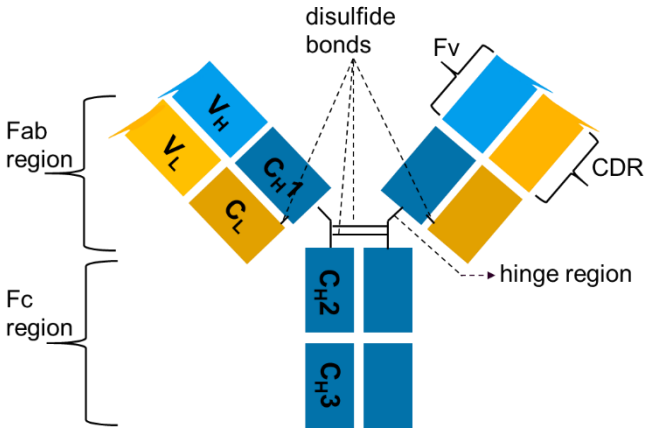


Figure 3.2 Structure of IgG representing constant (C_H and C_L) and variable (V_H and V_L) regions.

It can be further divided into variable (V) and constant (C) regions (see Figure 3.2). Within antibodies of the same class, constant regions (C_H and C_L) are similar and variable regions (V_H and V_L), which

can be also denoted as complementarity determining regions (CDR) are different in their amino acid sequences. CDR regions determine the antigen specificity within the variable regions[16].

Additionally, single chain Fv (scFv) and single domain V_H are also other possible specifically engineered substructures for particular biotherapeutics reasons (e.g. bispecific antibodies). Bispecific antibodies[17] contains two different antigen binding sites in one molecule, which can be generated by attachment of scFv fragment, single domain V_H or other genetic engineering structures carboxyl ends of IgG molecules[18].

In this study, five mAbs have been used. Details of each mAb are mentioned in Table 3.1.

Table 3.1 Overview of proteins used for stability and structural studies

PIPPI Code	Type	M_W (kDa)	Extinction coefficient[19], ε (L g^{-1} cm^{-1})	pI	Sequence	Notes
Monoclonal antibodies						
PPI01	IgG1λ	144.8	1.56	7.96	available	
PPI02	Human IgG1κ	148.2	1.47	8.53	available	
PPI03	Human IgG1κ	144.8	1.435	8.44	available	Wild-type IgG
PPI08	IgG1κ + scFv	204.4	1.57	8.9-9.2	not available	Bispecific
PPI10	Human IgG1	144.2	1.533	8.95	available	
Human serum albumin						
PPI49	HSA	66.4	0.486	5.67	available	
Lipase						
PPI45	Lipozyme RM	29.5	1.2	4.7	available	

3.1.2 Human serum albumin (HSA)

Human serum albumin (66.5 kDa) is the most abundant transport protein in human blood plasma with a concentration up to 50 g/L[20]. It helps in transportation of fatty acids, ions and small molecules in blood. The very first reference to albumin was found back in 400 years BC by Hippocrates of Cos[21] in urine of patients with chronic liver disease.

HSA is a single peptide chain of 585 amino acids and has three major domains (I, II and III), where each domain consists of two subunits (see Figure 3.3). It has 17 intra-subdomain disulfide bridges with one free thiol at position 34[22].

HSA is often used as a standard protein, but it has also been utilized in the pharmaceutical industries as a drug carrier to design drug conjugates, treatment of shock, hypoalbuminemia and cirrhosis of the liver[23,24]. HSA has a long half-life of 19-22 days[21], a characteristic which is utilized in drug delivery. Due to its unique properties and high stability, HSA has been used as an excipient in liquid and lyophilized biopharmaceuticals[25]. Further research is ongoing to increase the half-life, to harvest the high stability and solubility to stabilize other proteins[23,26]. Details of HSA used in this study are shown in Table 3.1.

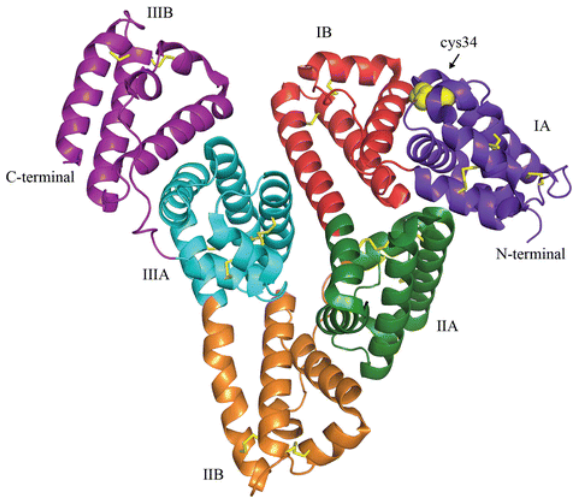


Figure 3.3 Structure of human serum albumin with three domains shown in purple (IA), red (IB), green (IIA), orange (IIB), blue (IIIA), and violet (IIIB). Yellow sticks depicture disulfide bridges. Structure is based on PDB:1E7E[27].

3.1.3 Lipase

Lipase belongs to the family of triglycerides lipases. Triglycerides lipases are a member of the *α/β* hydrolase superfamily that hydrolase the ester linkages of triglycerides. In biotechnology and biocatalysts, lipases are one of the most important molecules[28]. They have a number of attractive application in cosmetics, organic chemistry and pharmaceuticals[29]. The use of this enzyme has been well established since 1980s[30] and is currently used in pharmaceutical industry routinely for the synthesis enantiomerically pure active pharmaceutical ingredients and their intermediates for the production of single-isomer chiral drugs[31,32].

In this study, we have used lipase isolated from Rhizomucor meihei, Lipozyme RM (see Figure 3.4) consist of 272 amino acids with a molecular weight 29.5 kDa. It is a single polypeptide chain. The polypeptide chain is folded into coiled β-sheet domain with predominantly parallel strands which are connected by a variety of loops and helical segments[33]. Details of the lipase used in this study are shown in Table 3.1.

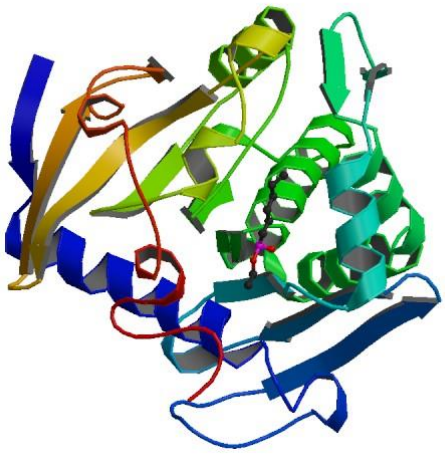


Figure 3.4 Structure of Lipozyme RM, PDB: 5TGL[34].

3.1.4 Overview of experiments

The aim of this study is to provide better understanding of the stability of mAbs, HSA and lipase in different formulation conditions in combination with pH, ionic strength, buffers and excipients. In general, mAbs stability increases with increasing pH and addition of NaCl decreases the thermal stability, that has also been observed in previous studies[35–37]. In case of HSA and lipase, stability gradually increases with increasing pH and then decreases from neutral to high pH. The buffers used (histidine, phosphate and acetate) have nearly similar effects with some exception (see discussion section). Excipients (sucrose, arginine and proline) have different effects on the stability of all the measured proteins. Sucrose and proline in general increase the stability, while arginine decreases the stability of mAbs (consistent with previous studies[38]).

3.2 Materials and Methods

3.2.1 Sample preparation

Monoclonal antibodies were provided by AstraZeneca, Grant Park, UK; Human serum albumin was provided by Albumedix Ltd, Nottingham, UK; and lipase was provided by Novozymes A/S, Denmark. Stock concentrations of each protein: PPI01 – 51.3 g/L, PPI02 – 100g/L, PPI03 – 46.4 g/L, PPI08 – 48.8 g/L, PPI10 – 49.3 g/L, PPI45 – 24.5 g/L and PPI49 – 200 g/L. Schematic flow chart of dialysis procedure is shown in Figure 3.5. For pH and NaCl screening, the stock solutions were dialyzed in 10 mM histidine pH 5.5, 7.0 and 10 mM tris pH 8.5.

For the buffer and excipients screening, all mAbs and HSA were dialyzed into 10 mM histidine pH 5.0 and 6.5, 10 mM acetate pH 5.0 and 10 mM phosphate pH 6.5 and lipase was dialyzed in 10 mM histidine pH 5.5 and 7.5, 10 mM acetate pH 5.5 and 10 mM phosphate pH 7.5. The buffer of the proteins was exchanged three times by extensive dialysis and final samples were prepared as previously described[39]. Protein concentrations were measured using Nanodrop™ 8000 (Thermo Fisher Scientific, Waltham, USA) using the respective extinction coefficients, ε, at 280 nm (see Table 3.1).

3.2.2 Differential Scanning Fluorimetry with Intrinsic Protein Fluorescence Detection (nanoDSF)

Samples containing 1 g/L protein in the respective formulations were filled in standard nanoDSF capillaries (NanoTemper Technologies, Munich, Germany). Measurements were performed using the Prometheus NT.48 (NanoTemper Technologies, Munich, Germany) system that measures the intrinsic protein fluorescence intensity at 330 nm and 350 nm after excitation at 280 nm. A temperature ramp of 1°C/min was applied from 20 to 95°C. The fluorescence intensity ratio (F350/F330) was plotted against the temperature, the inflection points of the unfolding transitions were determined of the first derivative of each measurement using the PR Control software V1.12 (NanoTemper Technologies, Munich, Germany). All measurements were performed in triplicates and within-triplicate outliers were removed[40].

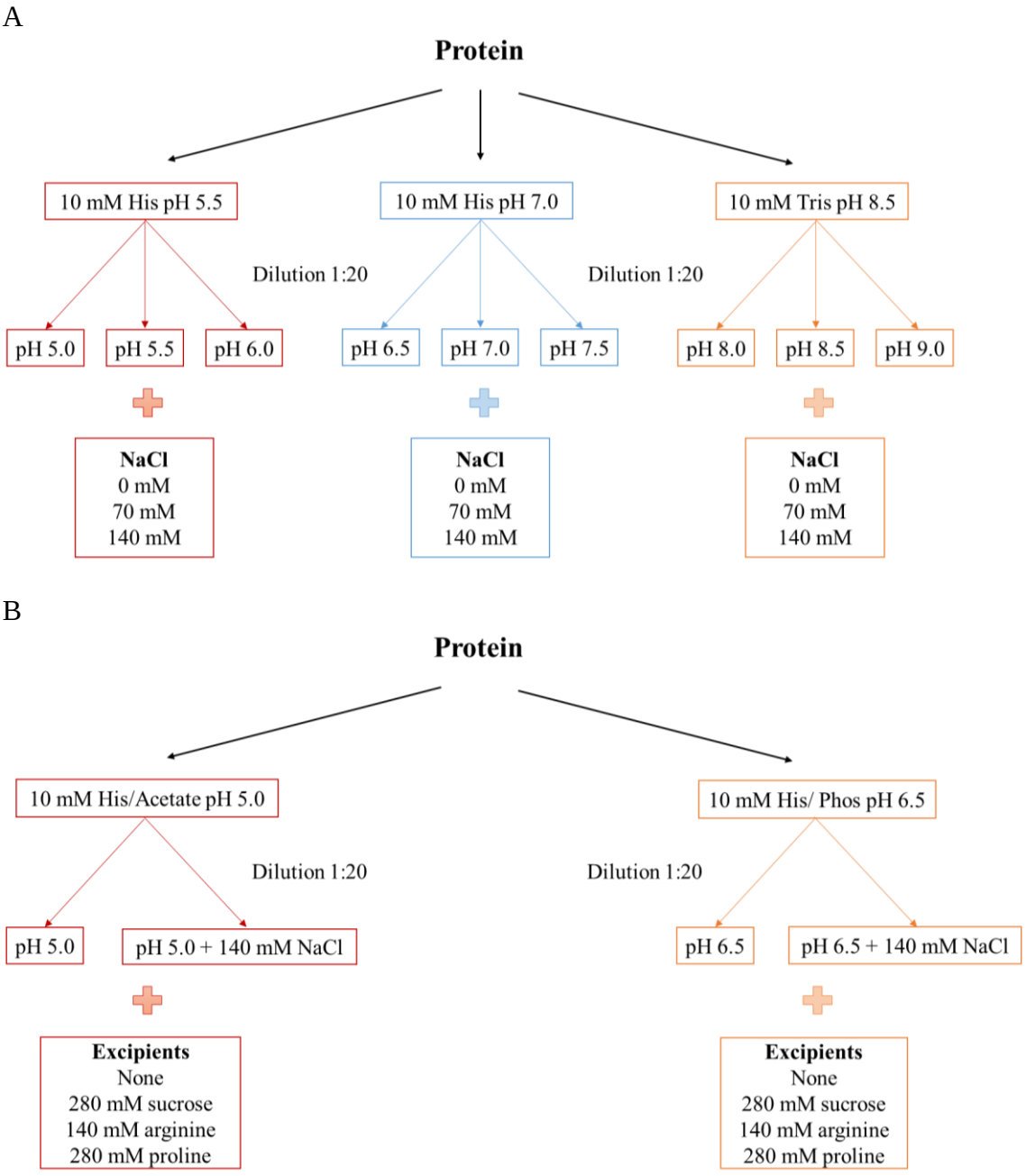


Figure 3.5 Schematic representation of dialysis for (A) pH/NaCl effect and (B) buffer/excipients effect

3.2.3 Isothermal Chemical Denaturation (ICD)

Chemical denaturation studies for all the proteins were performed using Unchained Labs HUNK system (AVIA ICD 2304). The excitation wavelength was 285 nm and emission intensities were recorded from 300 nm to 450 nm. The gain setting was set 10 for antibodies and albumin and 100 for lipases, based on a previously performed gain test. From the incubation test, 18.9 h of additional incubation time were set for antibodies, 5.24 h for lipase and no incubation needed for albumin. 48-point linear gradient was automatically generated for each condition. Guanidine hydrochloride (GuHCl) was used as denaturant. 6M GuHCl stock solutions were prepared in each condition. Protein stock solutions were prepared at 1 g/L and diluted 12.5 times to the final condition. The data analysis was performed using the Unchained Labs software (Formulator 3.02). For the native protein, the

fluorescence emission maximum $\lambda_{max\ (native)}$ was selected from the spectrum of the sample containing no denaturant. For the unfolded state, the fluorescence emission maximum $\lambda_{max\ (den)}$ was chosen from the fluorescence emission spectrum of the sample containing 5.5 M GuHCl. The ratio $\lambda_{\max\ (den)}/\lambda_{\max(native)}$ was plotted against denaturant concentration to monitor the unfolding process. The secondary fit was performed for each pH combining different NaCl concentrations in order to minimize the error. Free energy of unfolding (ΔG), $c_{½}$ and m-values were calculated for using different state model for different proteins; for antibodies (two state transitions) with 3 state models and for lipases and albumin with two state models (one state transition).

3.2.4 Small Angle X-ray Scattering (SAXS)

SAXS experiments were performed at the ESRF synchrotron, BM29 bioSAXS[41] beamline at Grenoble, France and P12 beamline at the Petra III[42] storage ring, DESY, Hamburg, Germany. Measurements on pure water were used to get the data on an absolute scale. Buffers were measured both before and after each sample and averaged before subtraction. Data collection parameters are listed in Table S3.1 for BM29, ESRF and P12, DESY.

Buffer averaging and subsequent subtraction prior to data analyses were performed in Primus[43]. The ATSAS program package version 2.8.4[44] was used for further data analysis. Primus was also used to perform Guinier region analysis and GNOM[45] was used for pair distribution, $p(r)$, analysis. The intensity, $I(q)$ is measured as a function of scattering vector $q = 4\pi\frac{sin\theta}{\lambda}$ here λ is the wavelength, and 2θ the scattering angle. M_w calculations were performed using, $M_W = [N_A I(0)/c]/\Delta\rho_M^2$ where N_A is Avogadro constant, $I(0)/c$ is concentration normalized forward scattering and $\Delta\rho_M$ is the scattering contrast per mass. $\Delta\rho_M$ was calculated using proteins average partial specific volume[46], $0.7425\ cm^3g^{-1}$.

3.3 Results

3.3.1 Monoclonal antibodies (mAbs)

Thermal and chemical unfolding studies of five mAbs in different formulation condition complemented by SAXS

Thermal and chemical unfolding studies have been performed on five mAbs using differential scanning fluorimetry (DSF) and isothermal chemical denaturation (ICD) in different conditions as a function of pH from 5.0 to 9.0 with 0.5 intervals and ionic strengths (0, 70, 140 mM NaCl) in different buffers (e.g. histidine, etc. all at a concentration of 10 mM) and excipients.

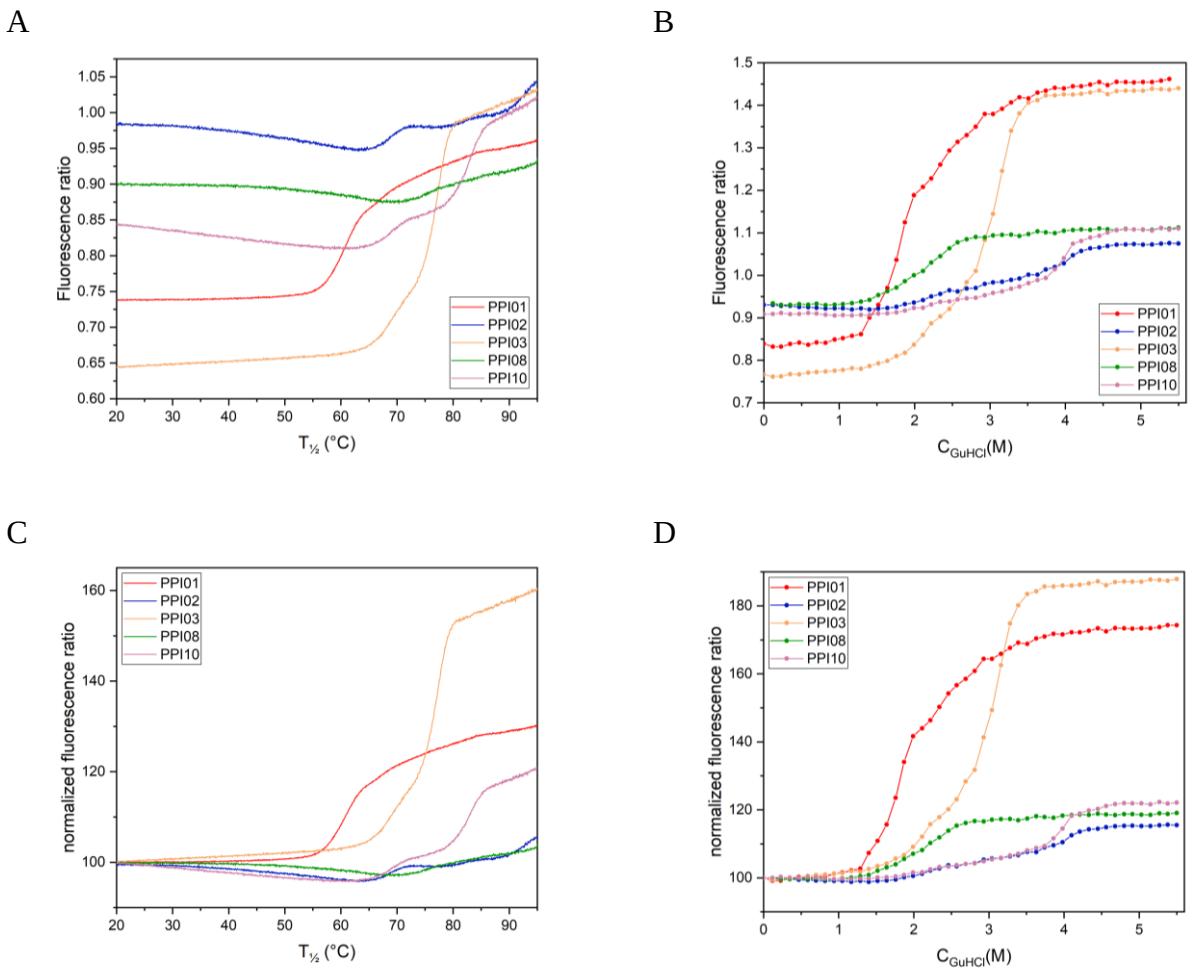


Figure 3.6 Thermal (A) and chemical (B) unfolding curves of PPI01 (red), PPI02 (blue), PPI03 (yellow), PPI08 (green) and PPI10 (purple) in 10 mM histidine pH 6.0. In Panel C and D, normalized thermal and chemical unfolding curves are also shown for the all five mAbs in similar formulation condition.

Shown above in Figure 3.6 is an example of the thermal and chemical unfolding curves for all five mAbs at 10 mM histidine pH 6.0, 0 mM NaCl. Measurements of the thermal and chemical unfolding reveal typical sigmoidal curves with a plateau at low temperature representing the fully native state (pretransition) and a plateau at higher temperatures (posttransition), representing the denatured state. All the mAbs showed two unfolding transitions from both methods but have different unfolding steps. The distinct unfolding process can be attributed to the different conformational stabilities of Fab and Fc domains of the mAbs. Thermal stability studies were examined by calculating the apparent melting temperature ($T_{½}$) for each protein and chemical denaturation studies were examined by calculating the concentration of denaturant ($c_{½}$) to unfold 50% the protein.

Structural studies were performed using SAXS on selected well behaved conditions from screen I and screen II and mainly based on ICD and nanoDSF data. All selected conditions were combinations of pH, NaCl, buffers and excipients. Measurement conditions are listed in Table 3.2.

Table 3.2 Selected conditions for SAXS measurements from thermal and chemical unfolding studies

Protein	**Buffer**	**Additive**
PPI01 (IgG1λ)	10 mM His pH 5.0	-
	10 mM His pH 5.5	-
	10 mM His pH 6.0	-
	10 mM His pH 6.5	-
PPI03 (Human IgG1κ)	10 mM His pH 6.0	NaCl (0, 35, 70, 140 and 250 mM)
	10 mM His pH 6.5	-
		Arginine (0, 35, 70 and 140 mM)
	10 mM Phos pH 6.5	-
		Arginine (0, 35, 70 and 140 mM)
PPI02 (Human IgG1κ)	10 mM His pH 6.5	-
		NaCl (140 mM)
		Arginine (0, 35, 70 and 140 mM)
	10 mM His pH 7.0	NaCl (140 mM)
PPI08 (IgG1κ + scFv)	10 mM His pH 6.5	-
PPI10 (Human IgG1)	10 mM His pH 5.0	-
		Arginine (0, 35, 70 and 140 mM)
PPI49 (HSA)	10 mM His pH 6.0	-
	10 mM His pH 7.0	-
PPI45 (Lipozyme RM)	10 mM Acetate pH 5.5	-
		Arginine (0, 35, 70 and 140 mM)
	10 mM His pH 5.5	-
	10 mM His pH 7.5	-
	10 mM Phos pH 7.5	-
		Arginine (0, 35, 70 and 140 mM)
	10 mM Tris pH 9.0	-

Sequence alignment were performed for mAbs by using Clustal Omega program[47] and shown in section 3.6.1 in supplementary materials. All the mAbs have more than 80% sequence similarity in heavy and light chain except PPI01, which has 40% sequence similarity in light chain.

Comparison between PPI03 - Human IgG1κ (Wild-type IgG) and PPI02 - Human IgG1κ

pH and NaCl effect

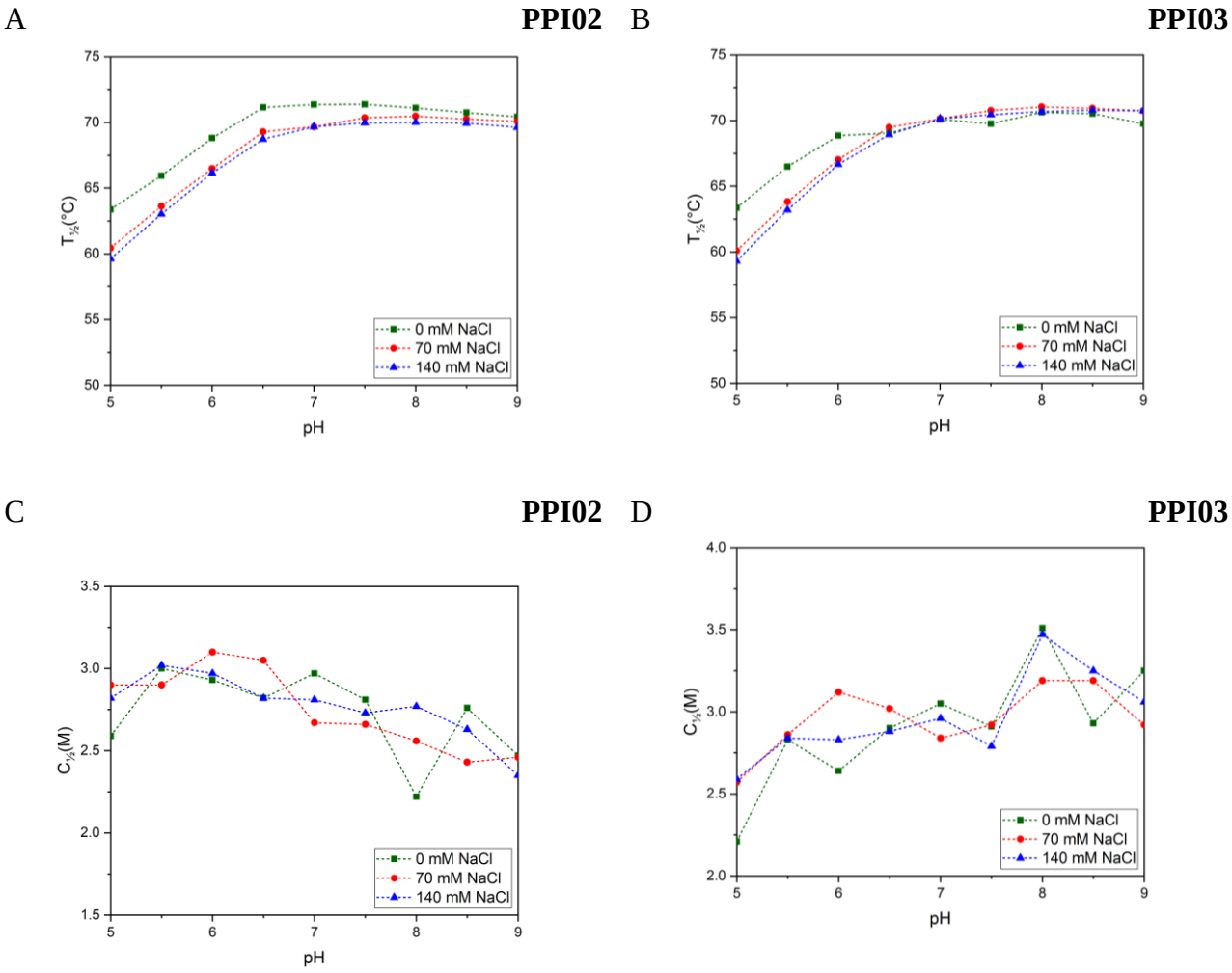


Figure 3.7 Apparent melting temperature of PPI02 (A) and PPI03 (B) measured by using nanoDSF at different pH (5.0-9.0) and different NaCl concentrations (0, 70, and 140 mM). The protein concentration in the measured samples in (A) and (B) is 1 g/L; Denaturant needed to unfold 50% of the protein using guanidine hydrochloride of PPI02 (C) and PPI03 (D). The protein concentration in the measured samples in (C) and (D) is 0.08 g/L. In both methods, stability parameters are calculated determined by measuring the change in intrinsic protein fluorescence intensity ratio and fitting the data.

PPI02 and PPI03 belong to human IgG1k type, have a molecular weight of 148.2 kDa and 144.8 kDa, respectively. They have a 86.4% sequence similarity in the heavy chain and differ in the Fv region while having a 92.1% sequence similarity in the light chain being different in the CDR region (see section 3.6.1).

NanoDSF studies show an increase in $T_{½}$ with increasing pH, reflecting an increase in thermal stability as shown in Figure 3.7(A, B). However, for PPI02, the addition of NaCl decreases the thermal stability in all the pH's tested, but in the case of PPI03, NaCl has an effect below pH 6.5. Visual inspection of the ICD plots shows that a higher concentration of GuHCl is needed to unfold PPI03 compared to PPI02, which indicates higher resistance to GuHCl unfolding of PPI03. In chemical denaturation

process, with increasing pH, $c_{½}$ is decreasing for PPI02, while it increases for PPI03 with increasing pH. However, a significant difference cannot be proved in case of NaCl due to no distinctive trend.

Buffer and excipient effect

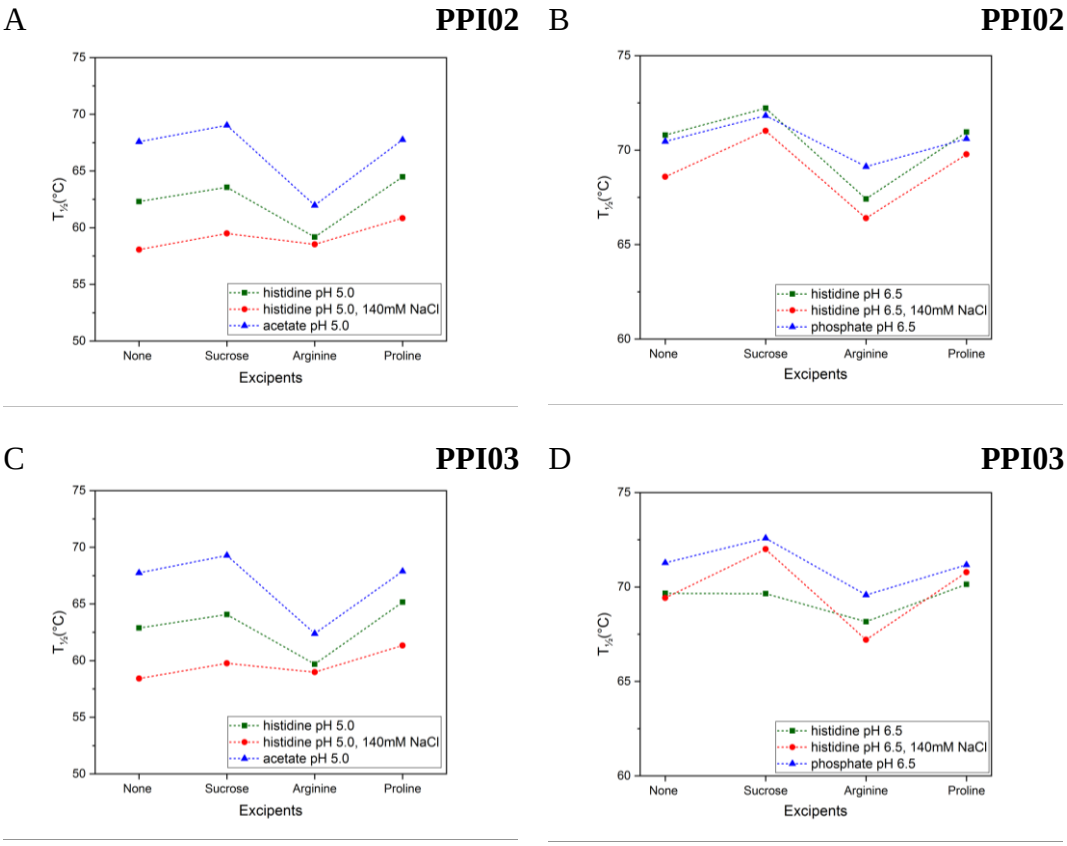


Figure 3.8 nanoDSF measurements of PPI02 and PPI03 at pH 5.0 and pH 6.5 in different buffers, NaCl and excipients. A and C: thermal denaturation studies of PPI02 and PPI03 respectively at pH 5.0, blue: 10 mM acetate pH 5.0, green: 10 mM histidine pH 5.0, red: 10 mM histidine pH 5.0 with 140 mM NaCl. B and D: thermal denaturation studies of PPI02 and PPI03 respectively at pH 6.5, blue: 10 mM phosphate pH 6.5, green: 10 mM histidine pH 6.5, red: 10 mM histidine pH 6.5 with 140 mM NaCl.

NanoDSF and ICD analyses for PPI02 and PPI03 are shown in Figure 3.8. At pH 5.0, the apparent melting temperature is higher in acetate than in histidine and lowest in histidine with NaCl added. In acetate and histidine buffers addition of arginine lowers $T_{½}$ while addition of sucrose increases $T_{½}$. Addition of proline seem to increase $T_{½}$ in histidine and histidine with NaCl added, but has no effect in acetate. The two proteins behave somewhat differently at pH 6.5. PPI02 is more stable in histidine and acetate buffer while addition of NaCl destabilizes about 2ºC. PPI03 is more stable in phosphate than in the two histidine buffers. Generally, sucrose seem to stabilize, except for PPI03 in histidine buffer and addition of proline has no significant effect. Addition of arginine seem to lower $T_{½}$. An overview of these results are shown in Table 3.5.

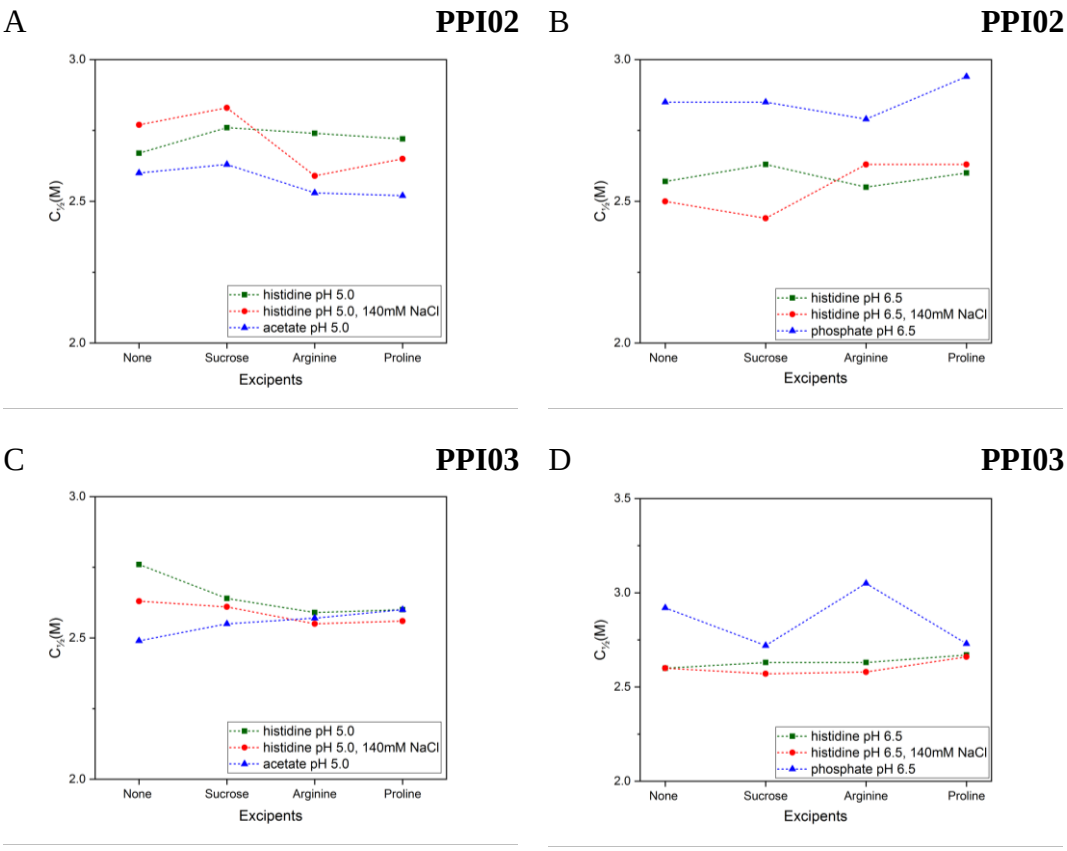


Figure 3.9 ICD measurements of PPI02 and PPI03 at pH 5.0 and pH 6.5 in different buffers, NaCl and excipients. A and C: chemical denaturation studies of PPI02 and PPI03 respectively at pH 5.0, blue: 10 mM acetate pH 5.0, green: 10 mM histidine pH 5.0, red: 10 mM histidine pH 5.0 with 140 mM NaCl. B and D: chemical denaturation studies of PPI02 and PPI03 respectively at pH 6.5, blue: 10 mM phosphate pH 6.5, green: 10 mM histidine pH 6.5, red: 10 mM histidine pH 6.5 with 140 mM NaCl.

In Figure 3.9, we are comparing the results in histidine and acetate buffers at pH 5.0 for both PPI02 and PPI03, $c_{½}$ is decreasing in presence of acetate buffer, which points to decreases in stability, however, PPI03 is stabilized by arginine and proline. At pH 6.5 it appears that, phosphate stabilizes both the mAbs under all tested conditions.

SAXS studies

For structure and interaction studies, SAXS concentrations series measurement were performed at pH 6.5 for both PPI02 and PPI03 in combination with different buffers, NaCl and arginine as an excipient based on ICD and nanoDSF data. Selected pH was chosen from PIPPI screen II.

SAXS scattering intensities of PPI02 and PPI03 are shown in Figure 3.10(A, C). The intensity curves superimpose for q-values greater than 0.08 $Å^{-1}$, which indicates that no conformational changes are occurring at higher protein concentration. For q-values less than 0.08 $Å^{-1}$, the scattering intensities decreases with increasing protein concentration (shown in Figure 3.10(A, C)), indicating repulsive interactions in 10 mM histidine at pH 6.5. Kratky plots are shown in Figure 3.10(B, D) for PPI02 and PPI03. They represent characteristics of folded protein and the shape of the plots represents the behavior of multidomain proteins.

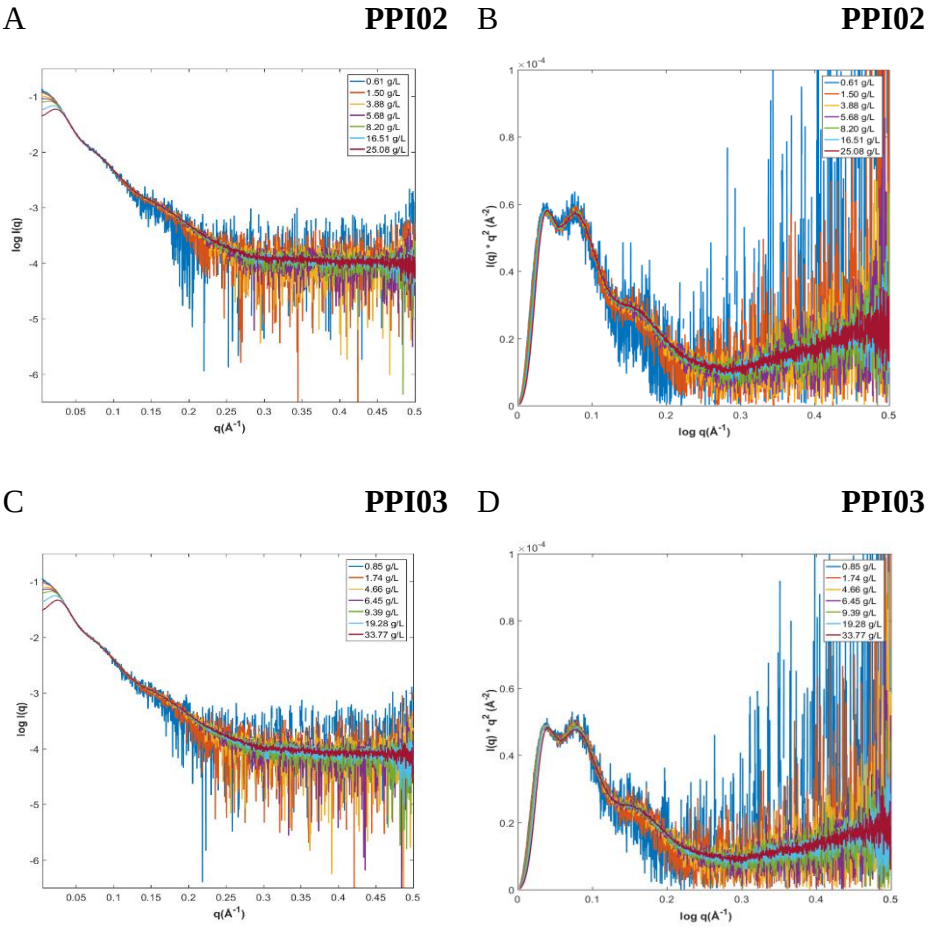


Figure 3.10 SAXS measurements of PPI02 and PPI03 in 10 mM histidine at pH 6.5. A and C: scattering curves, normalized for concentration, of PPI02 and PPI03 respectively. B and D: kratky plots of PPI02 and PPI03 respectively.

In Figure 3.11, the shown plot is the pair distribution function of PPI02 and PPI03 at 0.61 g/L and 0.85 g/L respectively in 10 mM histidine at pH 6.5. Each of the P(r) curves contains two peaks, which has been observed previously for antibodies across different subtypes including IgG1, IgG2, and IgG4, indicates characteristics of a multidomain protein[48–50].

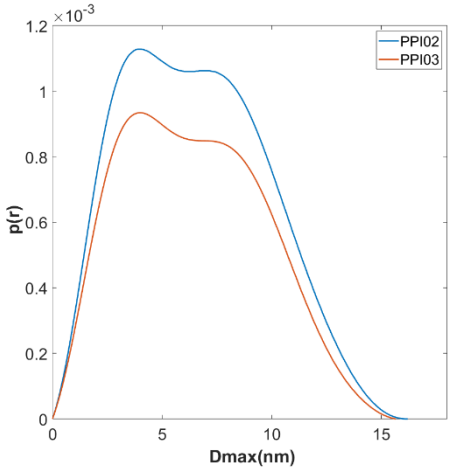


Figure 3.11 P(r) curves obtained from the SAXS profile for PPI02 and PPI03.

From Figure 3.12(A), it can be seen that PPI02 is repulsive in nature. Calculated R_G and D_{max} seems to be consistent over the concentration range, while the apparent M_W is decreasing with increasing concentration (see Table S 3.2). In order see the effect of NaCl, SAXS measurements were performed in presence of 140 mM NaCl in 10 mM histidine pH 6.5 and pH 7.0. At pH 6.5, the addition of NaCl decreases the repulsion, while at pH 7.0 NaCl introduce attraction (see Figure 3.12(D, E)). However, at pH 6.5, an increasing D_{max} in combination with a decreasing R_G and M_W, indicates both repulsive and attractive interactions, while at pH 7.0, all the calculated SAXS parameters are increasing leads to the presence of aggregates (see Table S 3.2). From the Kratky plot in Figure 3.12(E) shows a different shape for concentration above 1.21 g/L at pH 7.0, which indicates conformational changes at higher concentration. However, at pH 6.5, PPI02 does not have any conformational changes in the presence of NaCl.

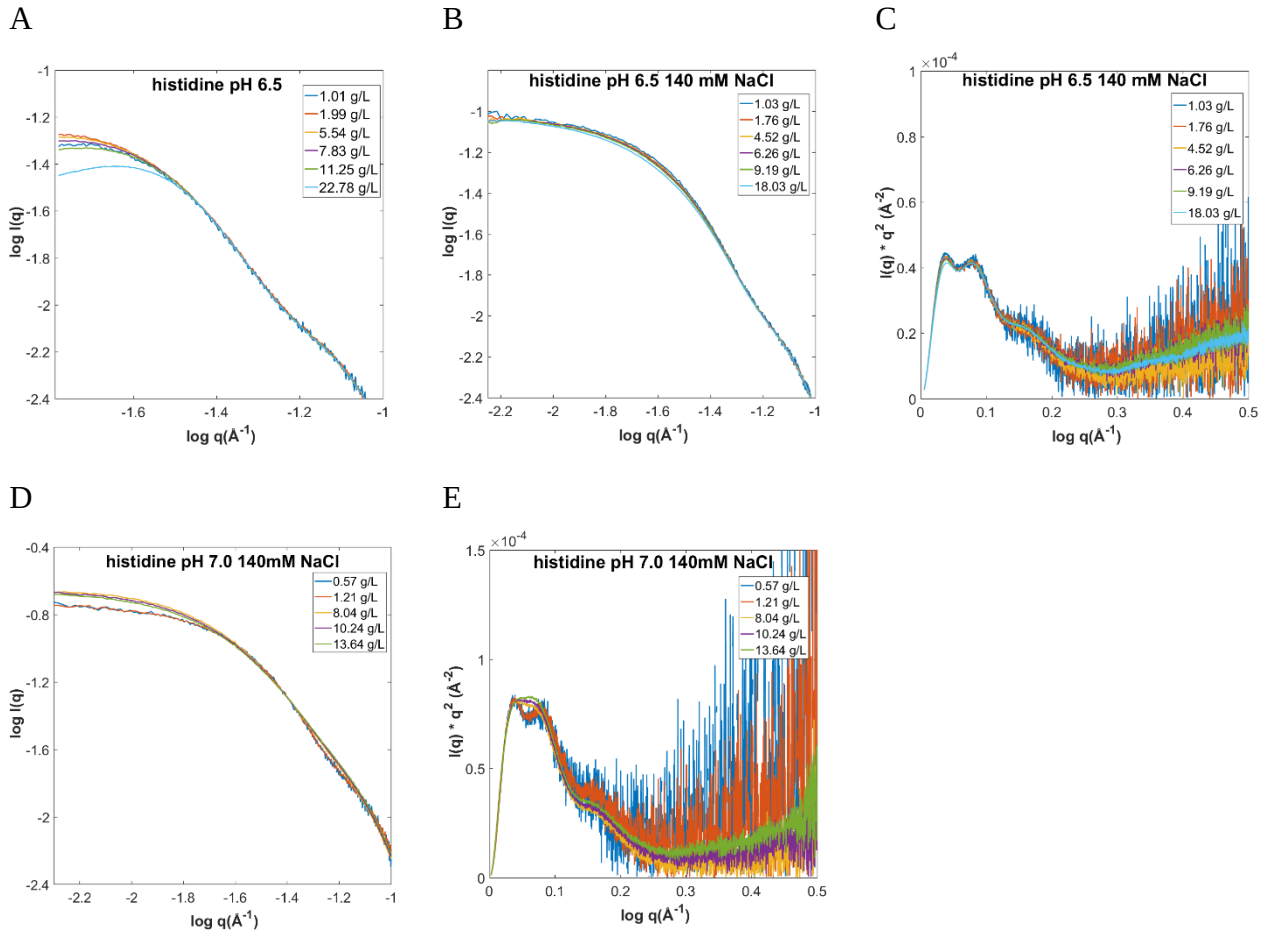


Figure 3.12 SAXS scattering profiles for PPI02: (A) 10 mM histidine pH 6.5, X-axis scale is different due to radiation damage during the measurements; (B) 10 mM histidine pH 6.5 and 140 mM NaCl; (D) 10 mM histidine pH 7.0 and 140 mM NaCl. Kratky plot for PPI02: (C) 10 mM histidine pH 6.5 and 140 mM NaCl; (E) 10 mM histidine pH 7.0 and 140 mM NaCl.

The effect of arginine was studied for PPI02 under two different protein concentrations of 1 g/L and 5 g/L in 10 mM histidine at pH 6.5 shown in below. In this case, addition of arginine introduces attraction as there is increase in scattering intensities, indicating that arginine destabilizes PPI02. Additionally, an increase in M_W, R_G and D_{max} points to the presence of aggregates (see Table S 3.2).

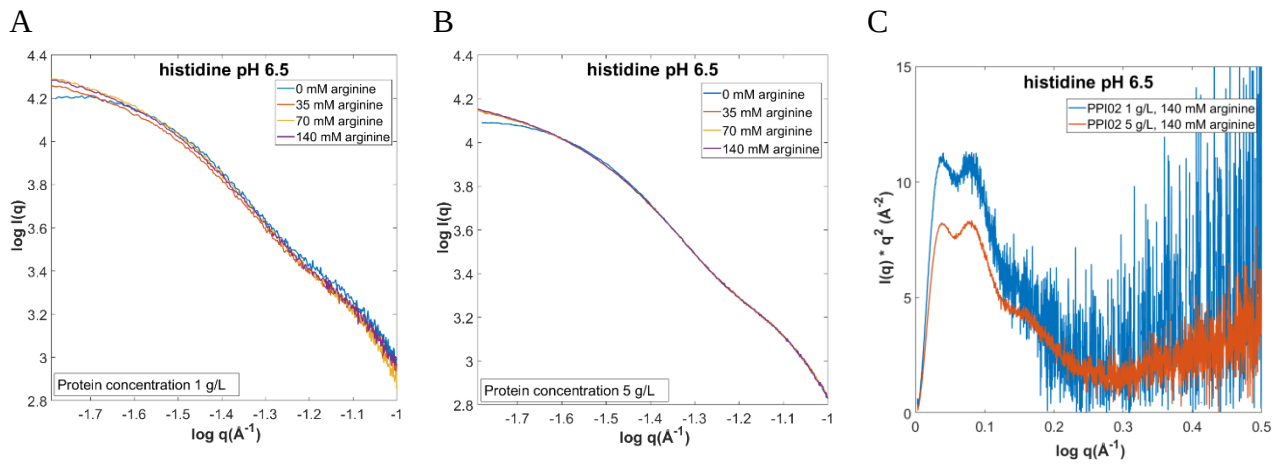


Figure 3.13 Scattering profiles for PPI02 at 10 mM histidine pH 6.5 with increasing concentration of arginine.

Unlike PPI02, PPI03 is repulsive at pH 6.5 (see Figure 3.14(A)) in histidine buffer as the intensity at low q-values decreases with increasing concentrations. Moreover, decrease of M_W, R_G and D_{max} indicates increase in repulsion (see Table S 3.3). Calculated SAXS parameters in phosphate buffer shows decrease of M_W, R_G and D_{max} as well, while the parameters are slightly higher than the monomeric IgG1 (see Table S 3.3). However, phosphate reduces the repulsion of PPI03 at higher concentration compared to histidine as the intensities are higher in phosphate (see Figure 3.14 (A, B) and Table S 3.3).

To observe how NaCl affects the structural features, SAXS measurements were performed at 10 mM histidine at pH 6.0 with the addition of 0, 35, 70 and 140 mM NaCl (see Figure 3.14(C)), data indicates that addition of NaCl screens the repulsive effect, which is completely gone above 70 mM NaCl.

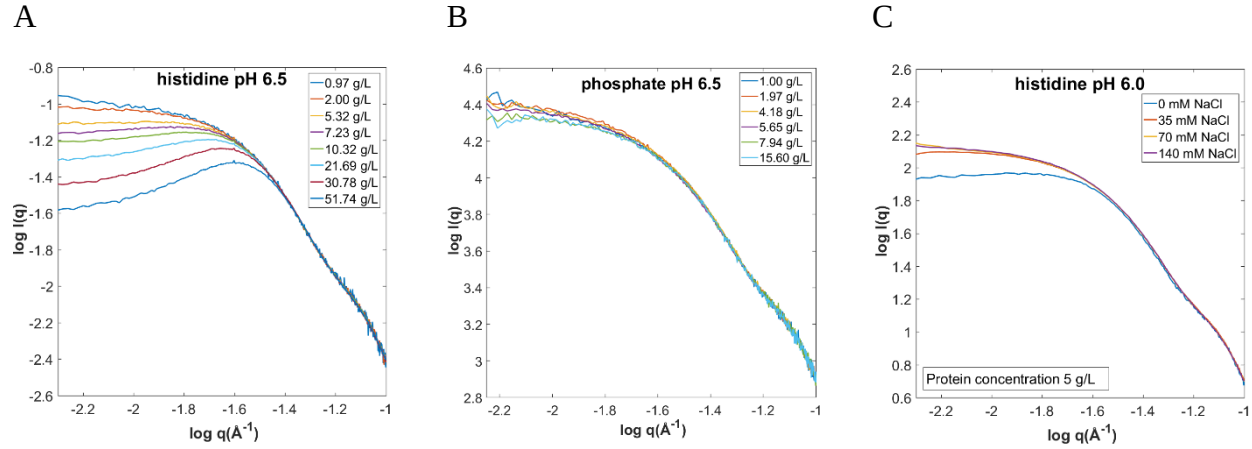


Figure 3.14 SAXS scattering profiles for PPI03 in correlation with pH, ionic strength: (A) 10mM histidine pH 6.5; (B) 10 mM phosphate pH 6.5; (C) 10 mM histidine pH 6.0 at 0, 35, 70, 140 mM NaCl.

In Figure 3.15, SAXS data to study buffer and excipient effects are shown for PPI03. The measurements have been performed in the presence of arginine in two different buffer systems: 10 mM histidine and 10 mM phosphate at pH 6.5. In Figure 3.15(B, E), increasing arginine concentration does not show any significant effect in histidine and phosphate buffer at 5 g/L protein concentration. However, in case of histidine at pH 6.5, M_W, R_G and D_{max} is increasing with increasing arginine concentration (see Table S 3.3) but does not lead to aggregation, while in case of phosphate M_W, R_G

and D_{max} are decreasing indicates arginine is reducing repulsion. Moreover, Kratky plots of PPI03 show an increase in scattering at a higher angle at 1 g/L in case of histidine (see Figure 3.15(C, F)), which is a characteristic for a flexible system.

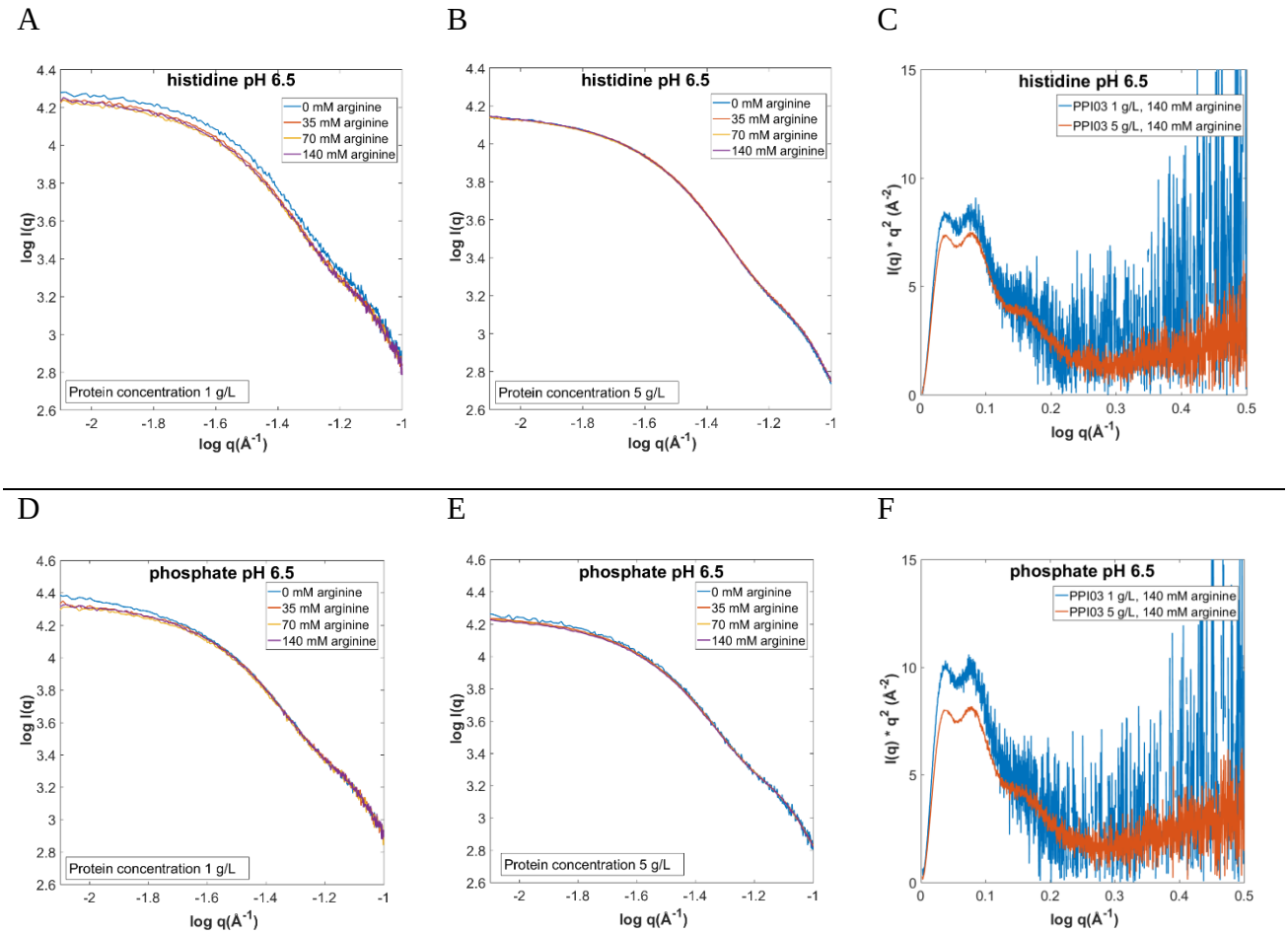


Figure 3.15 SAXS scattering profiles for PPI03 in combination with pH with histidine and phosphate buffers (10 mM buffer concentrations) and arginine (0, 35, 70, 140 mM) as an excipients at two different protein concentration to see the effect on protein concentration. A and B: 10 mM histidine pH 6.5 at 1 g/L and 5 g/L respectively;
D and E: 10 mM phosphate pH 6.5 at 1 g/L and 5 g/L respectively;
C and F: kratky plots in 10 mM histidine and 10 mM phosphate at pH 6.5.

PPI01 - IgG1λ type

pH and NaCl effect

Data acquired from nanoDSF and ICD are shown in Figure 3.16 (A, B). Thermal stability of PPI01 increases with increasing pH, with a plateau at pH 7.5. The thermal unfolding process is different from PPI02 and PPI03, as the shape of the obtained $T_{½}$ curve is different across the pH (Figure 3.7 (A, B) and Figure 3.16 (A)). In chemical denaturation process, we see a general trend of increasing $c_{½}$ from pH 6.0 to pH 8.0, except the extreme low and high pH. No effect of NaCl is observed in thermal denaturation process, while a clear dependence of NaCl is observed in chemical unfolding process. Also addition of 70 mM NaCl has a stabilizing effect as the $c_{½}$ is high.

The thermal and chemical unfolding curves for low, mid and high pH (pH 5.0, 7.5, 9.0) in presence of 70 mM NaCl are shown in Figure S 3.2. The shape of the thermal denaturation curves and the unfolding transition process appears to be fairly similar in all three pHs, while the shape of the chemical denaturation curves are slightly different from each other. Due to different chemical unfolding transition process, the effect of NaCl is pronounced for PPI01 in chemical denaturation.

A B

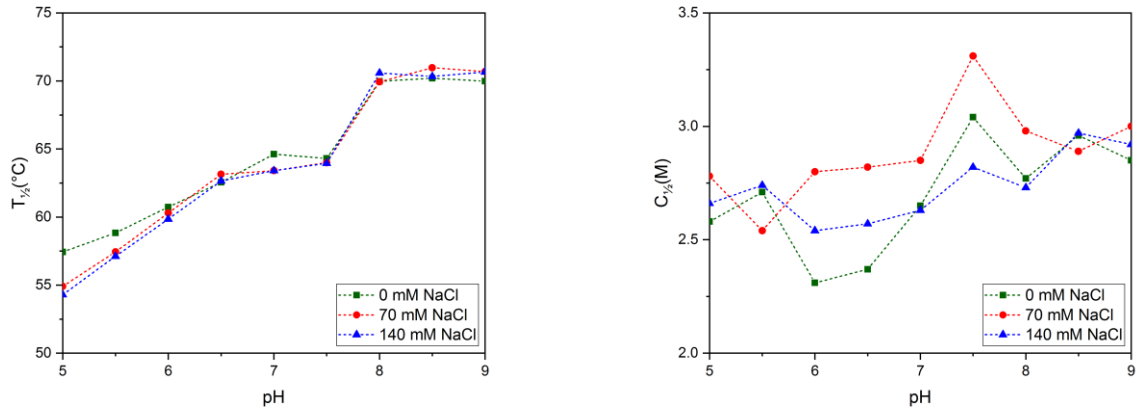


Figure 3.16 (A) $T_{½}$ measurements and (B) $c_{½}$ measurements of PPI01 at different pH (5.0-9.0) and different NaCl concentrations (0, 70, and 140 mM) using nanoDSF and ICD at 1 g/L and 0.08 g/L protein concentration respectively. Color code: 0 mM NaCl in green, 70 mM NaCl in red, and 140 mM NaCl in blue.

Buffer and excipient effect

To study the effect of buffer and excipient, pH 5.0 and pH 6.5 were chosen. From Figure 3.17(A, B), we observe that PPI01 in both acetate and phosphate buffer have higher $T_{½}$ relative to histidine at pH 5.0 and pH 6.5 by 2˚C. Among the excipients, arginine has destabilizing effect on thermal stability at pH 5.0 as arginine seem to lower $T_{½}$. However, addition of sucrose and proline has stabilizing effect as the $T_{½}$ is increasing in all the measured condition at pH 5.0. At pH 6.5, the effect on $T_{½}$ is greater in the presence arginine indicates increase in stability in comparison to sucrose and proline as an excipient. However, arginine in the combination of histidine and NaCl decreases the thermal stability compared to sucrose and proline at pH 6.5.

When denatured by GuHCl, PPI01 behaves equally in the presence of acetate and histidine at pH 5.0 (see Figure 3.17(C)). Arginine also has a negative effect on chemical denaturation process at pH 5.0 in histidine and acetate buffer, while sucrose has a positive effect. At pH 6.5, $c_{½}$ does not show any significant effect from excipients. ICD studies could not be obtained at phosphate pH 6.5 due to precipitation during dialysis. The supernatant of dialyzed protein at phosphate pH 6.5 was filtered using 0.22 µm cellulose acetate filters from VWR International and concentration obtained was 0.48 g/L. Thermal stability studies were performed due to the low sample concentration requirement of nanoDSF, as the concentration obtained after filtration was 0.48 g/L.

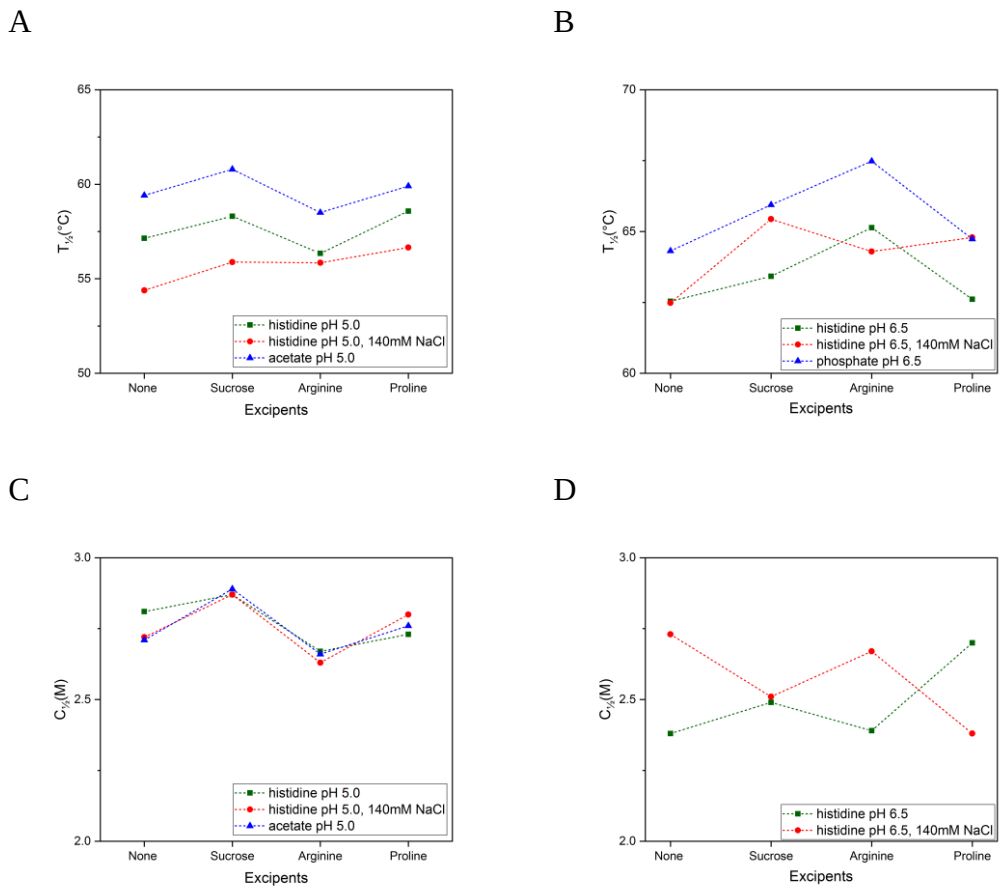


Figure 3.17 nanoDSF and ICD measurements of PPI01 at 1 g/L and 0.08 g/L protein concentration respectively. A and C: thermal and chemical denaturation studies at pH 5.0 respectively, blue: 10 mM acetate pH 5.0, green: 10 mM histidine pH 5.0, red: 10 mM histidine pH 5.0 with 140 mM NaCl. B and D: thermal and chemical denaturation studies at pH 6.5 respectively, blue: 10 mM phosphate pH 6.5, green: 10 mM histidine pH 6.5, red: 10 mM histidine pH 6.5 with 140 mM NaCl.

SAXS studies

As a part of my collaboration work for PPI01, characterization of the mAb association as a function of multiple factors: pH, temperature, salt concentration and protein concentration has been performed on PPI01 together with Lorenzo Gentiluomo at Wyatt Technology, Europe. In the study, reversibility of self-association has been studied. This association occurs mainly due to Fab-Fab hydrophobic interaction, while irreversible association occurs once the mAb is unfolded. One of the results is increased amount of oligomers detected as the pH is increased and at only pH 5.0 the monomer peak is detected[51]. These observations provided the motivation for using SAXS to study the effect of pH on structure and oligomer formation.

SAXS measurements were performed at 10mM histidine buffer at four different pH (5.0, 5.5, 6.0, 6.5) in concentration range from 0.5g/L – 17g/L. For all the curves in Figure 3.18, there are no changes observed for $q \geq 0.04$ Å^{-1} with increasing concentrations, but for the lower q region, $q < 0.04$ Å^{-1}, there is a significant increase in normalized forward scattering intensity, indicating aggregation. All the calculated parameters from the scattering curves are provided in Table 3.3. From the parameters, we observed the presence of larger oligomers/aggregates in all measured conditions with increasing

concentrations, which is also shown in Figure 3.19. At pH 5.0 and pH 5.5, PPI01 is monodisperse below 1 g/L as the apparent M_W is almost equal to the actual M_W. For pH 6.0 and above, PPI01 starts self-associating even at lower concentrations (see Table 3.3).

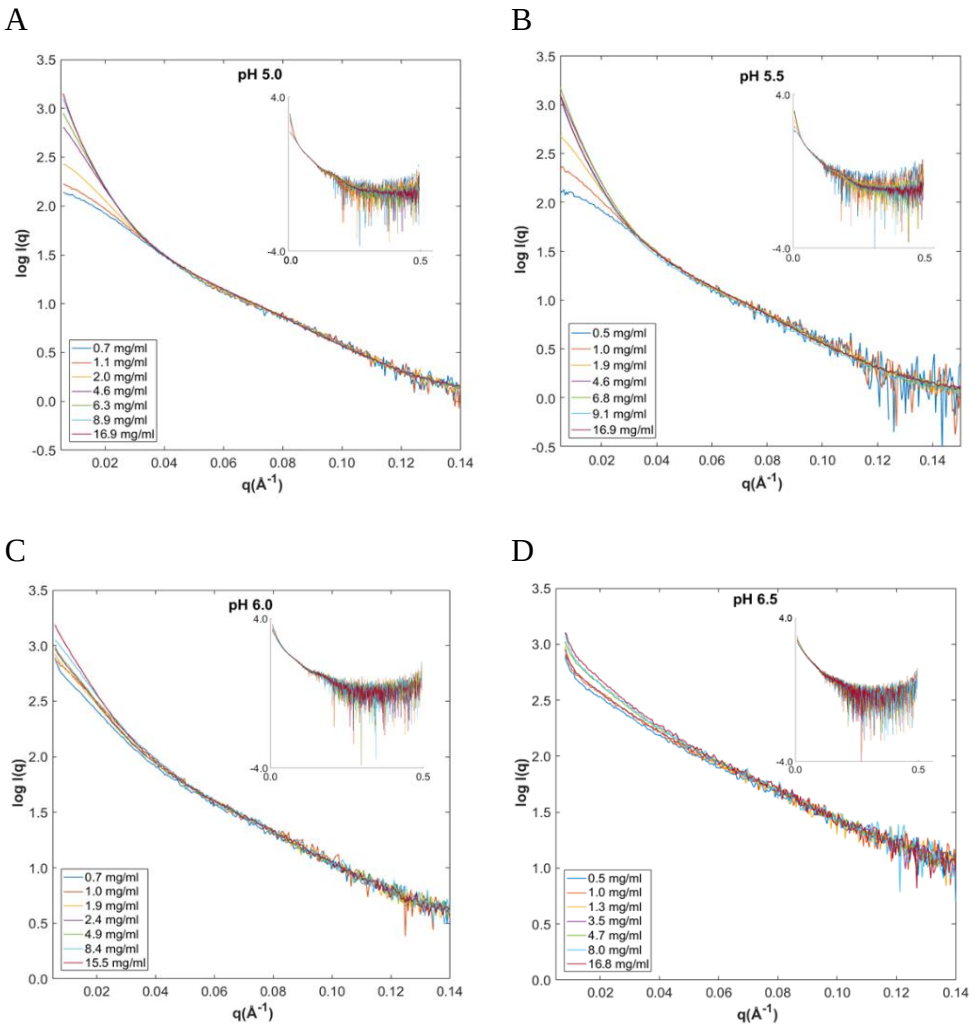


Figure 3.18 SAXS scattering curves: (A) 10 mM histidine pH 5.0, (B) 10 mM histidine pH 5.5, (C) 10 mM histidine pH 6.0, (D) 10 mM histidine pH 6.5. Data are shown for different formulation conditions with increasing concentrations.

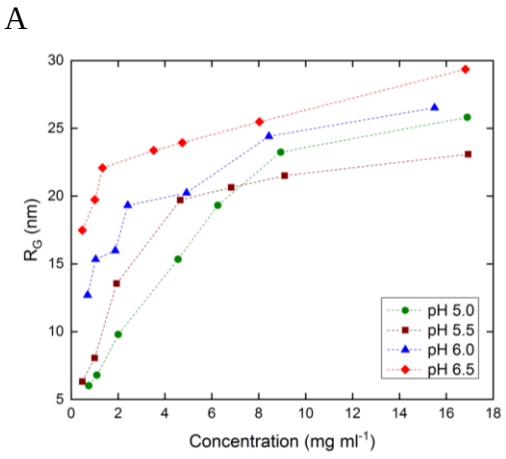


Figure 3.19 Calculated R_G for PPI01 in formulation conditions with increasing concentrations.

Table 3.3 An overview of the samples measured by SAXS and data treatment parameters for PPI01

Formulation condition	**Protein (g/L)**	**NaCl (mM)**	**p(r)**			**Apparent M_W (kDa)**
			I(0)/c	R_G (nm)	D_{max} (nm)	P(r)
10 mM histidine pH 5.0	0.74	-	0.11	6.01	15.11	156
	1.10	-	0.14	6.79	19.47	188
	2.01	-	0.24	9.80	40.70	333
	4.56	-	0.66	15.34	61.68	916
	6.26	-	1.05	19.32	75.49	1451
	8.94	-	1.74	23.24	80.00	2412
	16.89	-	2.24	25.82	85.00	3109
10 mM histidine pH 5.5	0.47	-	0.11	6.32	16.29	163
	1.00	-	0.14	8.06	24.51	249
	1.93	-	0.24	13.55	60.80	610
	4.65	-	0.66	19.71	83.00	1931
	6.82	-	1.05	20.64	62.00	1828
	9.11	-	1.74	21.51	81.52	1966
	16.93	-	2.24	23.09	81.30	2456
10 mM histidine pH 6.0	0.70	-	0.48	12.70	29.56	663
	1.04	-	0.56	15.34	38.31	774
	1.88	-	0.69	15.98	40.00	949
	2.41	-	1.05	19.32	34.57	1451
	4.92	-	1.58	20.24	32.50	2190
	8.43	-	1.70	24.42	38.18	2351
	15.50	-	2.16	26.52	43.04	2998
10 mM histidine pH 6.5	0.48	-	0.81	17.48	40.00	1119
	1.01	-	1.26	19.74	45.00	1747
	1.34	-	1.54	22.09	48.00	2129
	3.52	-	1.84	23.38	50.00	2555
	4.74	-	2.38	23.94	20.96	3298
	8.03	-	2.50	25.48	56.95	3459
	16.82	-	2.96	29.36	60.51	4106

PPI08 - IgG1κ + scFv (Bispecific)

pH and NaCl effect

PPI08 is a bispecific mAb which has an extra scFv fragment and a molecular weight of 204.4 kDa. Thermal stability is low at pH 5.0 and 5.5 but increases at pH 6.0 by ~10ºC, however, from pH 6.0 to pH 9.0 there are no significant changes (see Figure 3.20 (A)). For chemical denaturation process, in Figure 3.20 (B), there is no distinctive trend of $c_{½}$ as a function of pH and NaCl.

A B

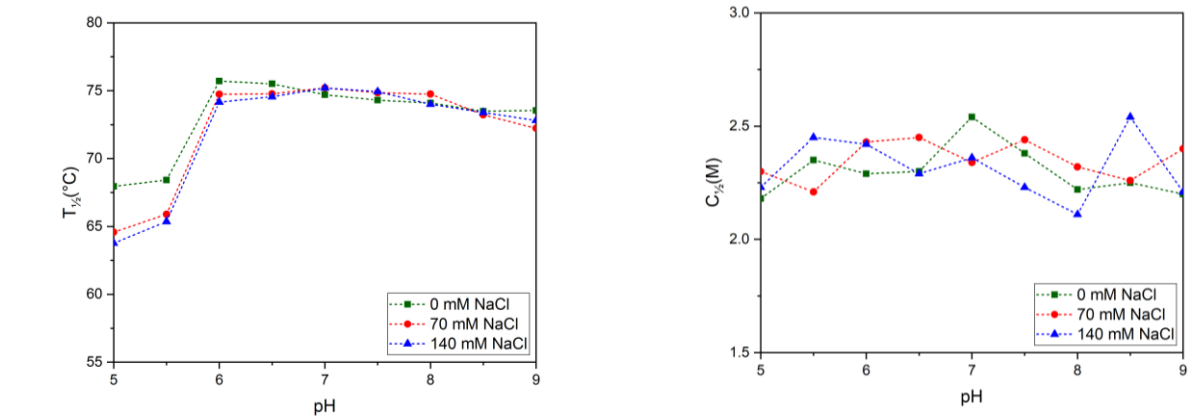


Figure 3.20 (A) $T_{½}$ measurements and (B) $c_{½}$ measurements of PPI08 at different pH (5.0-9.0) and different NaCl concentrations (0, 70, and 140 mM) using nanoDSF and ICD at 1 g/L and 0.08 g/L protein concentration respectively. Color code: 0 mM NaCl in green, 70 mM NaCl in red, and 140 mM NaCl in blue.

Buffer and excipient effect

From Figure 3.21(A, C), at we observe that pH 5.0 acetate stabilizes PPI08 in comparison to histidine, as the $T_{½}$ and $c_{½}$ is higher compared to histidine in almost all the tested conditions, while at pH 6.5, phosphate stabilizes PPI08 except in the absence of excipients. Moreover, presence of NaCl decreases the thermal stability of the protein at both pH 5.0 and pH 6.5 as the $T_{½}$ is decreasing.

By comparing the excipients at pH 5.0, arginine has a destabilizing effect on thermal stability and chemical denaturation process, while sucrose and proline is stabilizing the protein in all the tested conditions. However at pH 6.5, arginine is destabilizing PPI08 except in the presence of phosphate, while presence of sucrose and proline stabilizes PPI08 except in histidine pH 6.5, 140 mM NaCl.

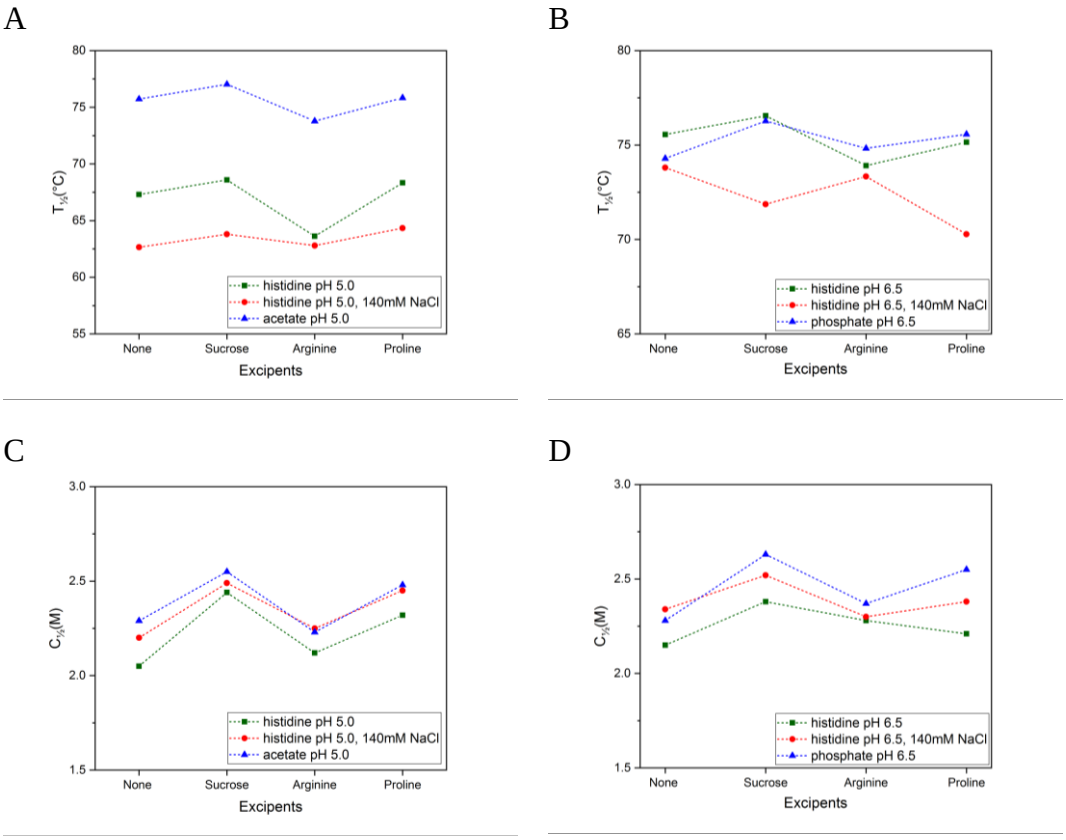


Figure 3.21 nanoDSF and ICD measurements of PPI08 at 1 g/L and 0.08 g/L protein concentration respectively. A and C: thermal and chemical denaturation studies at pH 5.0 respectively, blue: 10 mM acetate pH 5.0, green: 10 mM histidine pH 5.0, red: 10 mM histidine pH 5.0 with 140 mM NaCl. B and D: thermal and chemical denaturation studies at pH 6.5 respectively, blue: 10 mM phosphate pH 6.5, green: 10 mM histidine pH 6.5, red: 10 mM histidine pH 6.5 with 140 mM NaCl.

SAXS studies

SAXS measurements were performed at 10 mM histidine at pH 6.5. Figure 3.22 reflects PPI08 has repulsion, as the scattering intensities decrease with increasing protein concentration at pH 6.5. and shows similar repulsive tendency. Calculated M_W, R_G and D_{max} is decreasing with increasing concentration also indicates repulsion (see Table S 3.4). However, the apparent molecular weight is slightly lower than the actual molecular weight while considering R_G and D_{max}, PPI08 seems to be monomeric at pH 6.5.

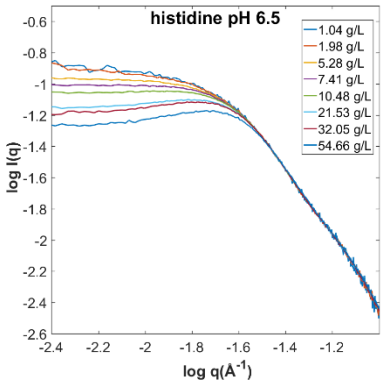


Figure 3.22 SAXS concentration series of PPI08 in 10 mM histidine pH 6.5.

PPI10 - Human IgG1

pH and NaCl effect

The pH and NaCl dependency of PPI10 shows similar thermal and chemical unfolding trend as PPI02 (see Figure 3.7(A, C) and Figure 3.23). Thermal stability increases with increasing pH 5.0 to pH 9.0. However, the addition of NaCl decreases the thermal stability for all pH values (see Figure 3.23(A)). In chemical denaturation process, $c_{½}$ decreases with increasing pH (see Figure 3.23(B)), although the effect of NaCl does not show a distinctive trend, except $c_{½}$ is maximum around pH 6.5.

A B

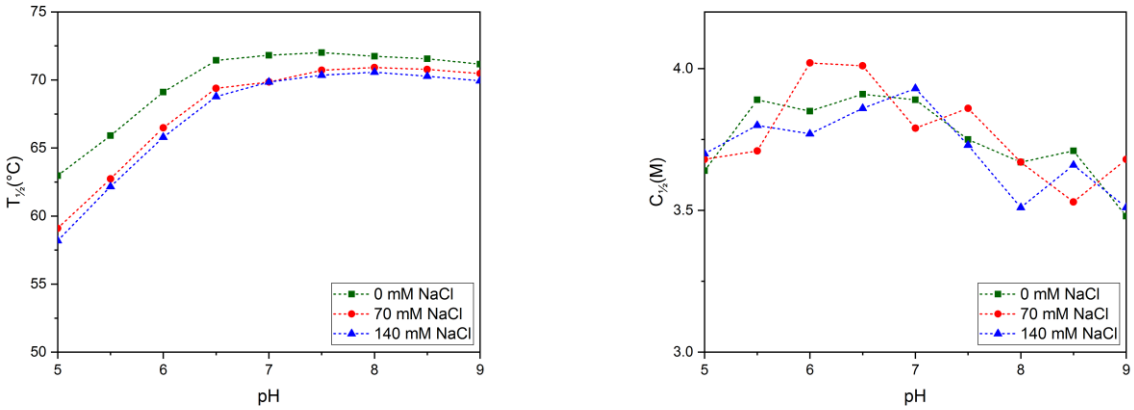


Figure 3.23 $T_{½}$ measurements and (B) $c_{½}$ measurements of PPI08 at different pH (5.0-9.0) and different NaCl concentrations (0, 70, and 140 mM) using nanoDSF and ICD at 1 g/L and 0.08 g/L protein concentration respectively. Color code: 0 mM NaCl in green, 70 mM NaCl in red, and 140 mM NaCl in blue.

Buffer and excipient effect

NanoDSF and ICD data are shown in Figure 3.24 (A, C). At pH 5.0, PPI10 is more thermally stable in acetate versus histidine as the $T_{½}$ is higher in all the tested conditions. Presence of NaCl destabilizes the PPI10. Moreover at pH 6.5, PPI10 is equally thermally stable in presence of phosphate compared to histidine (see Figure 3.24 (B)). Adding NaCl to histidine makes it less stable except $c_{½}$ is lower at pH 6.5 in the presence and absence of sucrose.

By looking at excipients effect, in presence of sucrose and proline, PPI10 has higher $T_{½}$ at pH 5.0 and pH 6.5, while arginine shows negative effect for thermal stability (see Figure 3.24(A, B)). If we look at ICD, at pH 6.5, arginine effect is enormous and stabilizes PPI10, while at pH 5.0, arginine destabilizes PPI10 except in acetate buffer. However, sucrose and proline do not any distinguishable trend in chemical denaturation process (see Figure 3.24 (C, D).

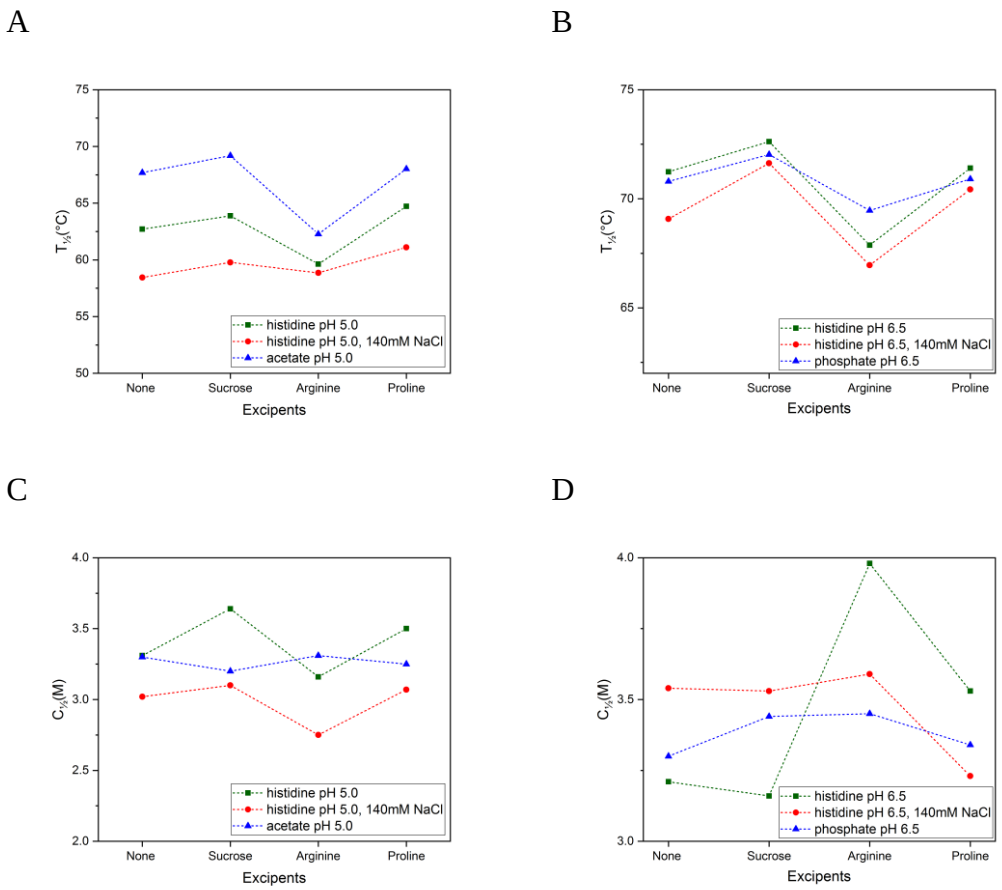


Figure 3.24 nanoDSF and ICD measurements of PPI10 at 1 g/L and 0.08 g/L protein concentration respectively. A and C: thermal and chemical denaturation studies at pH 5.0 respectively, blue: 10 mM acetate pH 5.0, green: 10 mM histidine pH 5.0, red: 10 mM histidine pH 5.0 with 140 mM NaCl. B and D: thermal and chemical denaturation studies at pH 6.5 respectively, blue: 10 mM phosphate pH 6.5, green: 10 mM histidine pH 6.5, red: 10 mM histidine pH 6.5 with 140 mM NaCl.

SAXS studies

SAXS concentration series were performed at pH 5.0 and pH 6.0 in 10 mM histidine. PPI10 shows repulsive behavior as the intensities are decreasing with increasing protein concentration and repulsion is decreasing from pH 5.0 to pH 6.0 (see Figure 3.25). Additionally, decrease in apparent M_W, R_G and D_{max} at pH 5.0 and pH 6.0 indicates that PPI10 is less repulsive at pH 6.0 (see Table 3.4).

A B

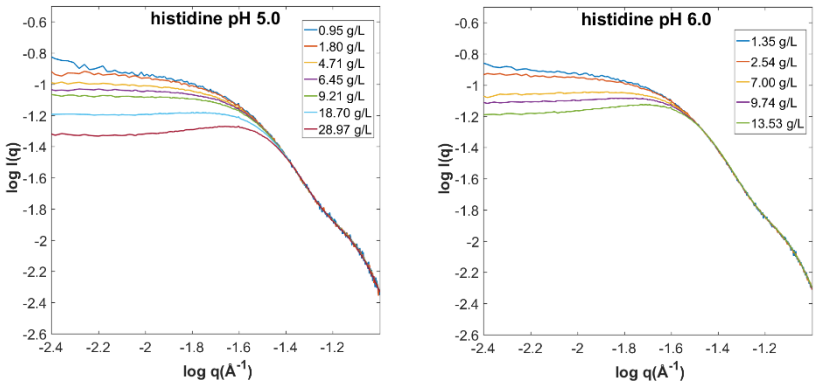


Figure 3.25 SAXS concentration series measurements of PPI10: (A) 10 mM histidine pH 5.0 and (B) 10 mM histidine pH 6.0.

Table 3.4 An overview of the samples measured by SAXS and data treatment parameters for PPI10.

Formulation condition	**Protein (g/L)**	**NaCl (mM)**	**Arginine (mM)**	**Guinier**		**p(r)**			**Apparent M_W (kDa)**	
				I(0)/c	R_G (nm)	*I(0)/c*	R_G (nm)	D_{max} (nm)	Guinier	P(r)
10 mM histidine pH 5.0	0.95	-	-	0.12	5.12	0.12	5.17	16.60	166	166
	1.80	-	-	0.12	4.91	0.12	5.03	14.11	166	166
	4.71	-	-	0.11	4.47	0.11	4.93	15.22	152	152
	6.45	-	-	0.10	4.44	0.11	4.83	14.73	138	152
	9.21	-	-	0.10	4.17	0.11	4.79	14.68	133	152
	18.70	-	-	0.09	4.10	0.10	4.66	14.00	120	138
	28.97	-	-	0.08	3.78	0.09	4.60	14.00	104	125
	4.71	-	0	0.11	4.47	0.11	4.93	15.22	152	152
	4.89	-	35	0.13	5.10	0.13	5.18	18.01	180	180
	4.83	-	70	0.15	5.13	0.15	5.21	18.49	208	208
	4.79	-	140	0.15	5.18	0.15	5.26	18.48	208	208
10 mM histidine pH 6.0	1.35	-	-	0.12	4.80	0.12	5.00	15.26	166	166
	2.54	-	-	0.12	4.41	0.12	4.90	15.02	166	166
	7.00	-	-	0.10	3.75	0.12	4.82	15.04	138	166
	9.74	-	-	0.09	3.38	0.11	4.71	14.44	130	152
	13.53	-	-	0.09	3.32	0.10	4.45	13.50	152	180

SAXS scattering curves for PPI10 in presence of arginine in 10 mM histidine at pH 5.0 are shown in Figure 3.26. It shows that increasing arginine introduces attraction at pH 5.0 as the scattering intensities increase with increasing arginine concentration from 0 mM to 140 mM. At 1 g/L protein concentration, the observed interaction is low, while at protein concentration 5 g/L, we see that at 70 mM arginine, the repulsion has been cancelled. Additionally, increases in M_W, R_G and D_{max} with increasing concentration of arginine indicates presence of aggregates at pH 5.0 (see Table 3.4).

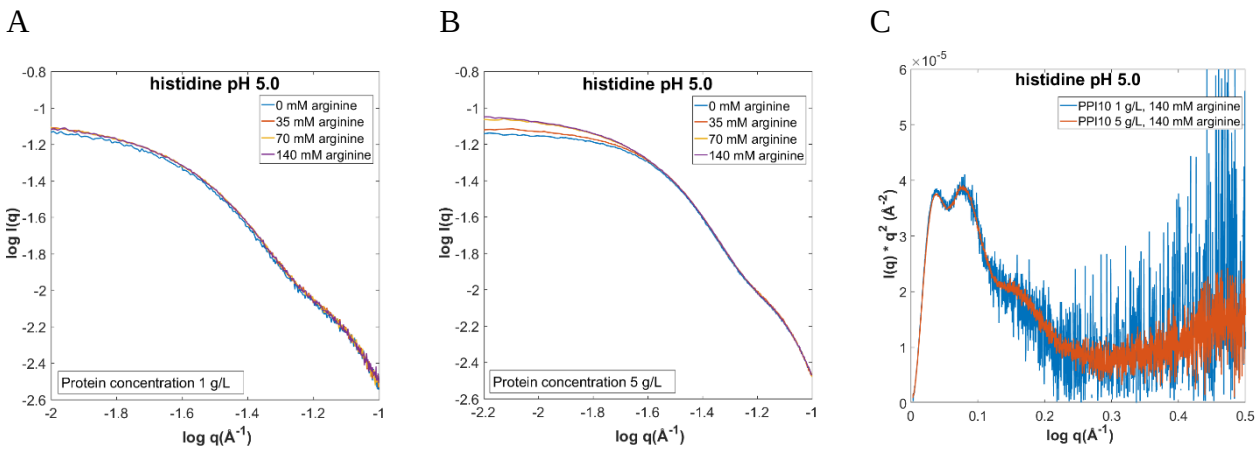


Figure 3.26 Scattering profiles for PPI10 at 10 mM histidine pH 5.0 with increasing concentration of arginine in at 1 g/L (A) and 5 g/L (B) protein concentrations.

3.3.2 Human Serum Albumin

Thermal and chemical unfolding studies have been performed on HSA using nanoDSF and ICD in different conditions as a function of pH from 5.0 to 9.0 and ionic strengths (0, 70, 140 mM NaCl) in different buffers (e.g. histidine, etc. all at a concentration of 10 mM) and excipients.

In Figure 3.28, an example of thermal unfolding curve for HSA at 10 mM histidine pH 6.5, 0 mM NaCl is shown. In thermal unfolding process, the fluorescence ratio of F350/F330 is decreasing first and increasing after around 63˚C. The decrease in ratio could be due to tryptophan is reburied during the unfolding process.

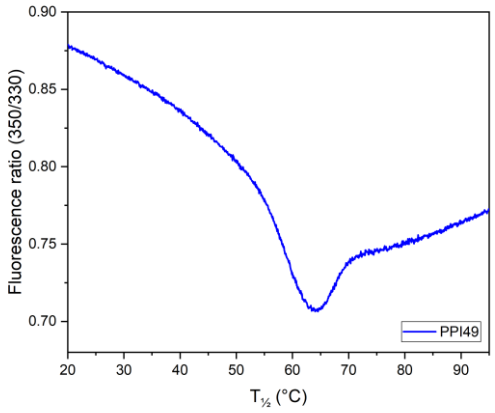


Figure 3.27 Thermal unfolding curve of HSA (blue) in 10 mM histidine pH 6.5 from nanoDSF.

In Figure 3.28, an example of chemical unfolding curve for HSA at 10 mM histidine pH 6.5, 0 mM NaCl is shown. In ICD, we measure the intrinsic fluorescence from 300 to 500 nm along the wavelength, which allows to follow a different ratio to get clear overview. At 5.5 M GuHCl, there are two observable peaks, the broad one for tryptophan and also a tyrosine peak at 304 nm (see Figure 3.28 (A)). ICD data analysis were performed using two state model by taking the fluorescence ratio of F304/F346, where a clear sigmoidal unfolding curve is observed (see Figure 3.28 (B)).

A B

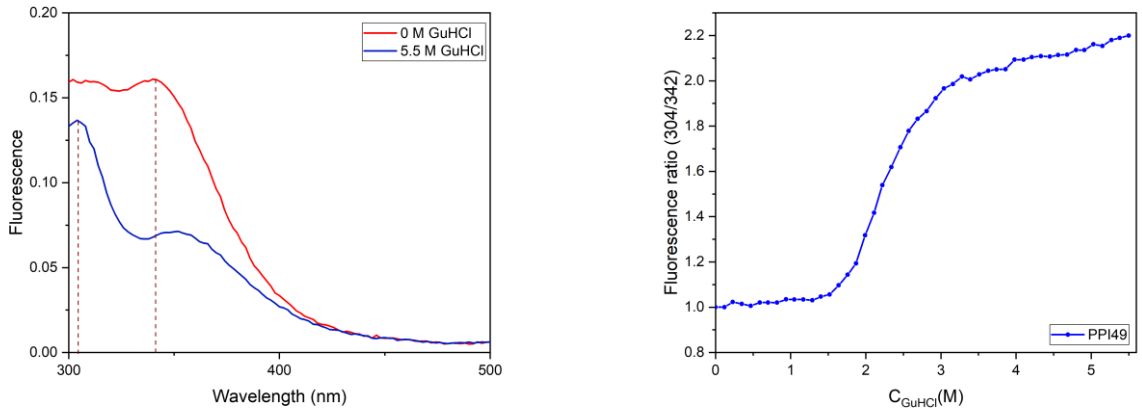


Figure 3.28 Chemical unfolding curves of HSA (blue) in 10 mM histidine pH 6.5. A: intrinsic fluorescence changes from 300 – 500 nm wavelength at 0 M GuHCl (red) and 5.5 M GuHCl (blue); B: changes in fluorescence ratio (304/342) vs GuHCl concentration.

pH and NaCl effect

Thermal and chemical denaturation studies of HSA in different formulation conditions are shown in Figure 3.29. From the thermal denaturation studies, it can be seen that, $T_{½}$ gradually increases from pH 5.0 to pH 7.5 and then decreases between pH 7.5 to pH 9.0, this indicates that Trizma buffer and the pH shift has destabilizing effect and HSA is less thermally stable at high pH compared to mid-range pH (see Figure 3.29(A)). NaCl has a stabilizing effect in thermal stability in all the tested conditions as the $T_{½}$ is increased by at least 5˚C in all the measured conditions.

A B

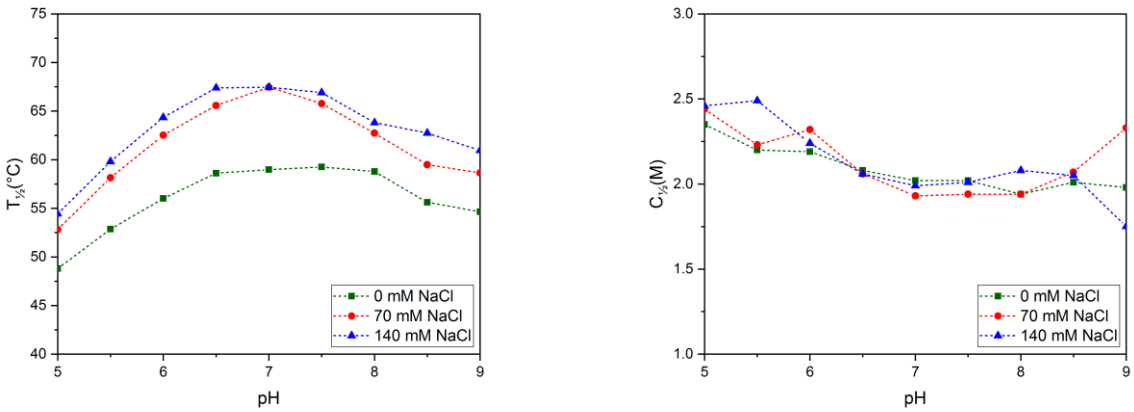


Figure 3.29 $T_{½}$ measurements and (B) $c_{½}$ measurements of HSA (PPI49) at different pH (5.0-9.0) and different NaCl concentrations (0, 70, and 140 mM) using nanoDSF and ICD at 1 g/L and 0.08 g/L protein concentration respectively. Color code: 0 mM NaCl in green, 70 mM NaCl in red, and 140 mM NaCl in blue.

HSA has one unfolding transition state (see Figure 3.28 (B)). Figure 3.29(B) indicates decrease in $c_{½}$ with increasing pH up to pH ~7.0 and where it reaches a plateau, while the effect of NaCl doesn't show a distinctive trend.

Buffer and excipient effect

To determine buffer and excipient effect, pH 5.0 and pH 6.5 were chosen from the above mentioned studies. Acetate buffer has a stabilizing effect on PPI49, compared to histidine, in the absence and presence of excipients as both $T_{½}$ is increasing at pH 5.0 with one exception in presence of proline. At pH 6.5, phosphate also has a stabilizing effect on both thermal and chemical unfolding process in all the tested conditions compared to histidine. Histidine in combination with NaCl has stabilizing effect on thermal stability (see Figure 3.30) compared to only histidine as a buffer component except in presence arginine at pH 6.5. At pH 5.0, sucrose, proline and arginine has positive effect on the thermal stability with an exception in presence of acetate, while at pH 6.5, there is no distinctive effect observed.

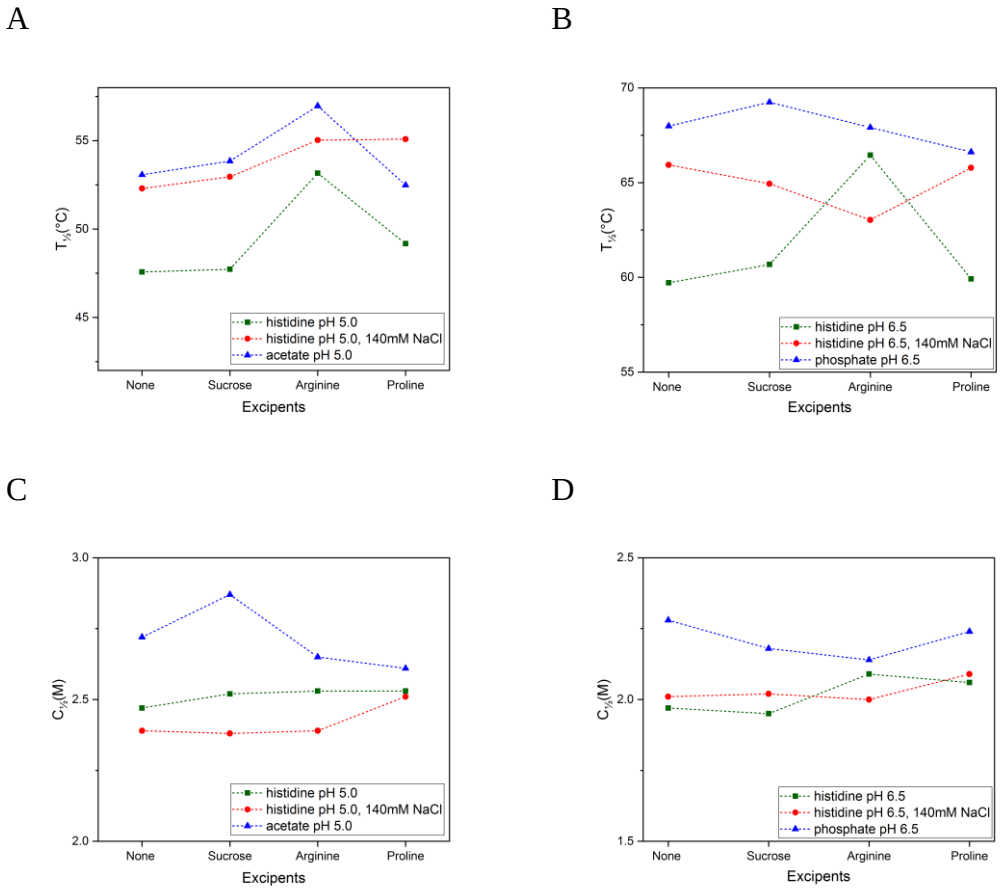


Figure 3.30 A and C: thermal and chemical denaturation studies of HSA at pH 5.0 respectively, blue: 10 mM acetate pH 5.0, green: 10 mM histidine pH 5.0, red: 10 mM histidine pH 5.0 with 140 mM NaCl. B and D: thermal and chemical denaturation studies of HSA at pH 6.5 respectively, blue: 10 mM phosphate pH 6.5, green: 10 mM histidine pH 6.5, red: 10 mM histidine pH 6.5 with 140 mM NaCl.

SAXS studies

Two near neutral pH conditions were chosen to perform SAXS concentration series measurements, i.e., pH 6.0 and pH 7.0. Data were shown in Figure 3.31, HSA is repulsive in nature in pH 6.0 and pH 7.0. There is no conformational changes with increasing protein concentrations, as the scattering curves superimpose at high *q*-values. Kratky plots are shown in Figure 3.31(B, D), indicate the protein is in fully folded state in both measured conditions.

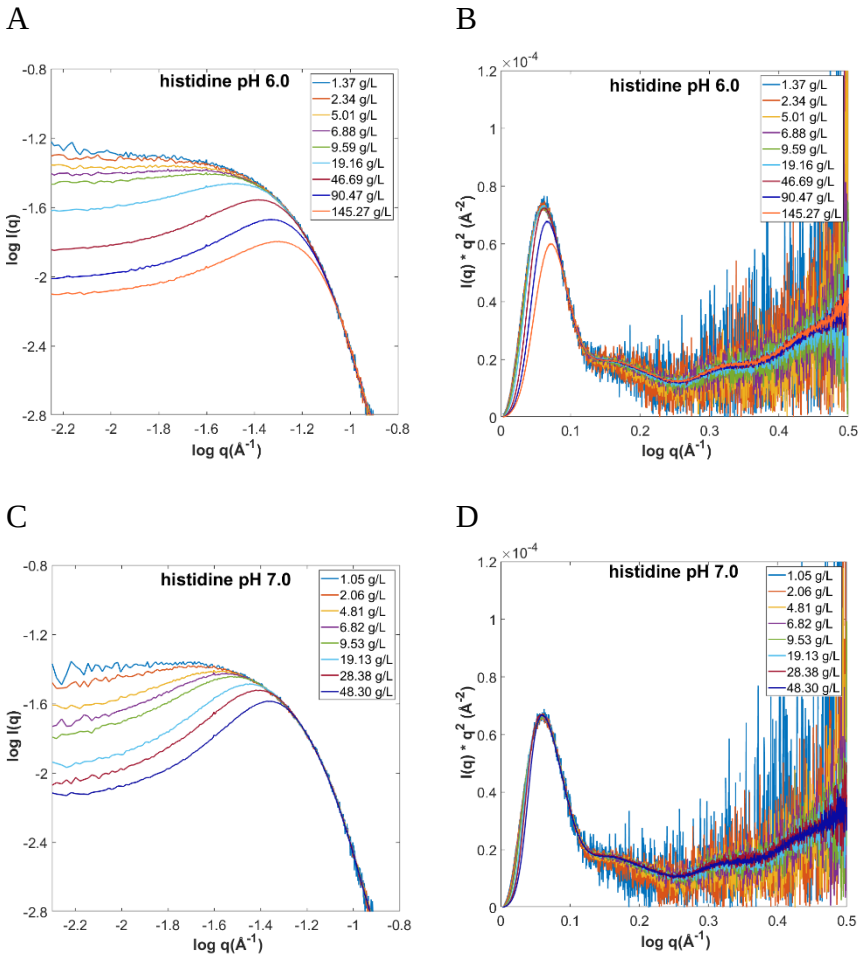


Figure 3.31 SAXS measurements of HSA in 10 mM histidine at pH 6.0 (A, B) and 10 mM histidine at pH 7.0 (C, D). A and C: scattering curves, normalized with concentration; B and D: kratky plots of HSA.

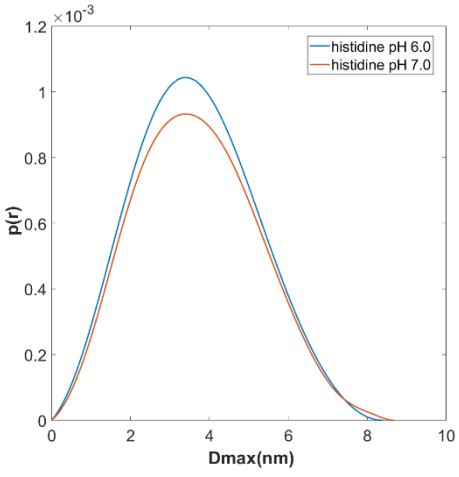


Figure 3.32 P(r) curves obtained from SAXS profile for HSA at 1 g/L.

In Figure 3.32, pair distribution function of HSA is shown where shape of curve indicates the behavior of a globular protein[52]. D_{max} at pH 7.0 is slightly higher than pH 6.0 due the presence of a tiny tail, which indicates presence of a smaller amount of aggregates. Additionally, at pH 6.0, calculated M_W,

R_G and D_{max} is decreasing with increasing protein concentration indicating repulsion. At pH 7.0, the calculated M_W, R_G and D_{max} shows less repulsion compared to pH 6.0 (see Table S 3.5).

3.3.3 Lipases

In Figure 3.33, thermal and chemical unfolding curves are shown. In thermal unfolding process, we see the similar pattern as HSA. The decrease in fluorescence intensity ratio (350:330) could be due to tryptophan is reburied during the unfolding (see Figure 3.33 (A)). However, we see a sigmoidal transition in case of chemical denaturation process for fluorescence intensity ratio of 354:342 (see Figure 3.33 (B)). Two state model has been used to analyze the chemical denaturation data.

A B

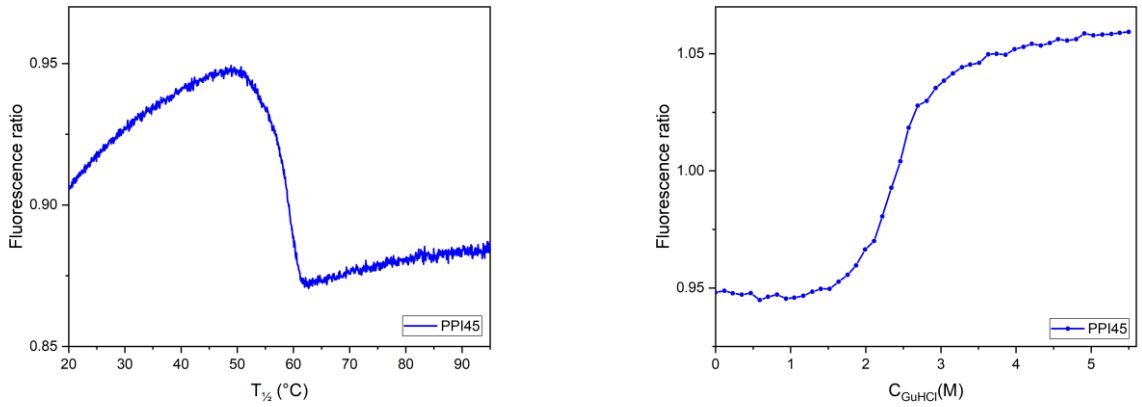
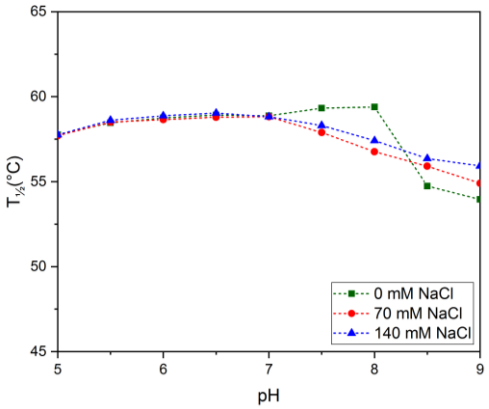


Figure 3.33 Thermal (A) and chemical (B) unfolding curves of lipase (blue) in 10 mM histidine pH 7.5.

pH and NaCl effect

Thermal and chemical denaturation studies were performed in different conditions as a function of pH from 5.0 to 9.0 and ionic strengths (0, 70, 140 mM NaCl). In Figure 3.34 (A), NanoDSF studies shows an increase in $T_{½}$ from pH 5.0 to pH 8.0, reflecting an increase in thermal stability in the absence of NaCl. However, $T_{½}$ is decreasing from pH 8.0 to pH 9.0 reflecting a decrease in thermal stability at higher pH. Measured $T_{½}$ is similar in the presence of NaCl until pH 7.0. Moreover, NaCl has a destabilizing effect in thermal stability from pH 7.0 and above with exceptions at pH 8.0 and pH 9.0. Figure 3.34 (B) shows, increase in pH does not have any significant effect on the chemical denaturation process. However, presence of 70 mM NaCl has a stabilizing effect from pH 6.0 and above as the $c_{½}$ is higher.

A



B

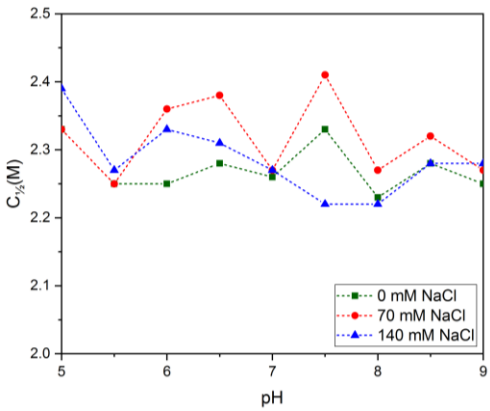
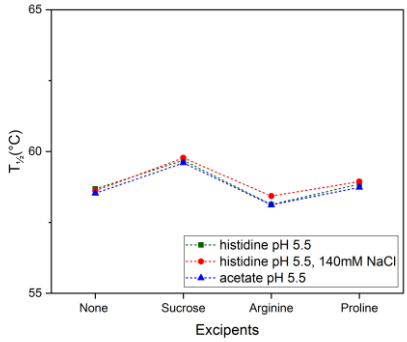


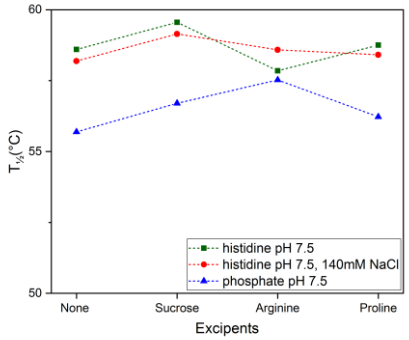
Figure 3.34 $T_{½}$ measurements and (B) $c_{½}$ measurements of lipase (PPI45) at different pH (5.0-9.0) and different NaCl concentrations (0, 70, and 140 mM) using nanoDSF and ICD at 1 g/L and 0.08 g/L protein concentration respectively. Color code: 0 mM NaCl in green, 70 mM NaCl in red, and 140 mM NaCl in blue.

Buffer and excipient effect

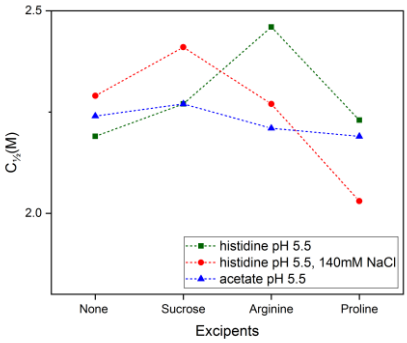
A



B



C



D

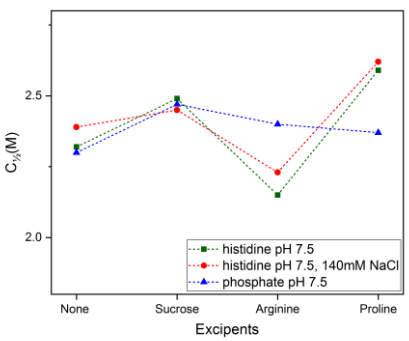


Figure 3.35 A and C: thermal and chemical denaturation studies of lipase at pH 5.5 respectively, blue: 10 mM acetate pH 5.5, green: 10 mM histidine pH 5.5, red: 10 mM histidine pH 5.5 with 140 mM NaCl. B and D: thermal and chemical denaturation studies of lipase at pH 7.5 respectively, blue: 10 mM phosphate pH 7.5, green: 10 mM histidine pH 7.5, red: 10 mM histidine pH 7.5 with 140 mM NaCl.

To study the buffer and excipient effect, pH 5.5 and pH 7.5 were chosen. NanoDSF and ICD ,measurements for lipase at pH 5.5 and pH 7.5 are shown in Figure 3.35. At pH 5.5, the apparent melting temperature is similar in all the measured buffers. However at pH 7.5, phosphate has a destabilizing effect on lipase. Sucrose as an excipient has stabilizing effect as the $T_{½}$ and $c_{½}$ is increasing at both pH 5.5 and pH 7.5 in all the tested conditions. Addition of arginine seem to lower $T_{½}$, in all the tested buffers with an exception in presence of phosphate. Addition of proline has no significant effect. For chemical unfolding process, at pH 5.5, $c_{½}$ is higher in presence of arginine in histidine buffer, indicates arginine stabilizes lipases. At pH 7.5, proline and sucrose stabilizes the lipase, as the $c_{½}$ is higher compared to lipase in the absence of excipients. Furthermore addition of NaCl has no significant effect in the chemical unfolding process

SAXS studies

In order to study the interactions and conformational stability, SAXS measurement were performed at pH 5.5 and pH 7.5, selected conditions from PIPPI screen and also at pH 9.0 as lipase is thermally less stable at higher pH.

Shown below is an example of scattering curves and kratky plot of SAXS concentrating series measurements of lipase in 10 mM histidine at pH 5.5. In Figure 3.36(A), the scattering intensities superimpose for q-values greater than 0.08 $Å^{-1}$ indicating no conformational changes with increasing protein concentration. The kratky plot represents characteristics of folded protein and the shape of the plots represents the behavior of a globular protein.

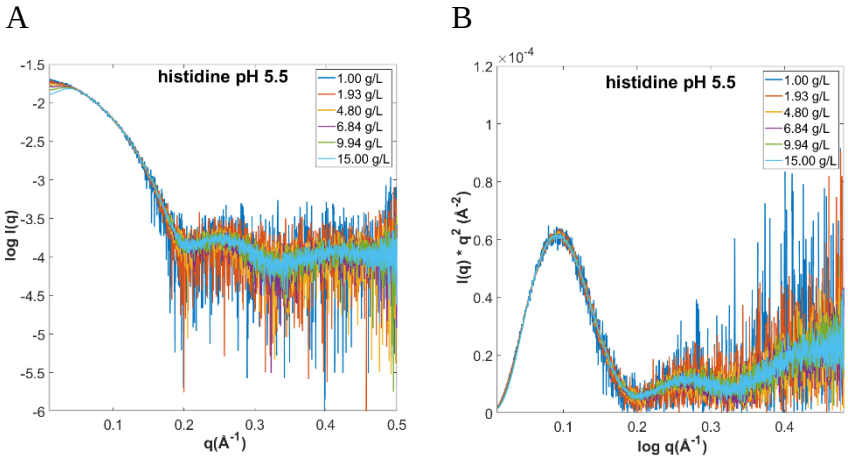


Figure 3.36 SAXS measurements of lipase in 10 mM histidine at pH 5.5. A: scattering curves, normalized with concentration; B: kratky plots of lipase.

Figure 3.37 shows that lipase with increasing protein concentration. However, D_{max} is slightly higher in acetate buffer in comparison to histidine and phosphate buffer (see Table S 3.6). This reflects a screening effect of acetate and lipase is less stable in acetate. Additionally, 10 mM histidine pH 7.5 is the best behaving formulation condition as the calculated M_W, D_{max} and R_G seems to be consistent and appears to be close to the exact parameters for a monomeric lipase throughout the concentration series measurements.

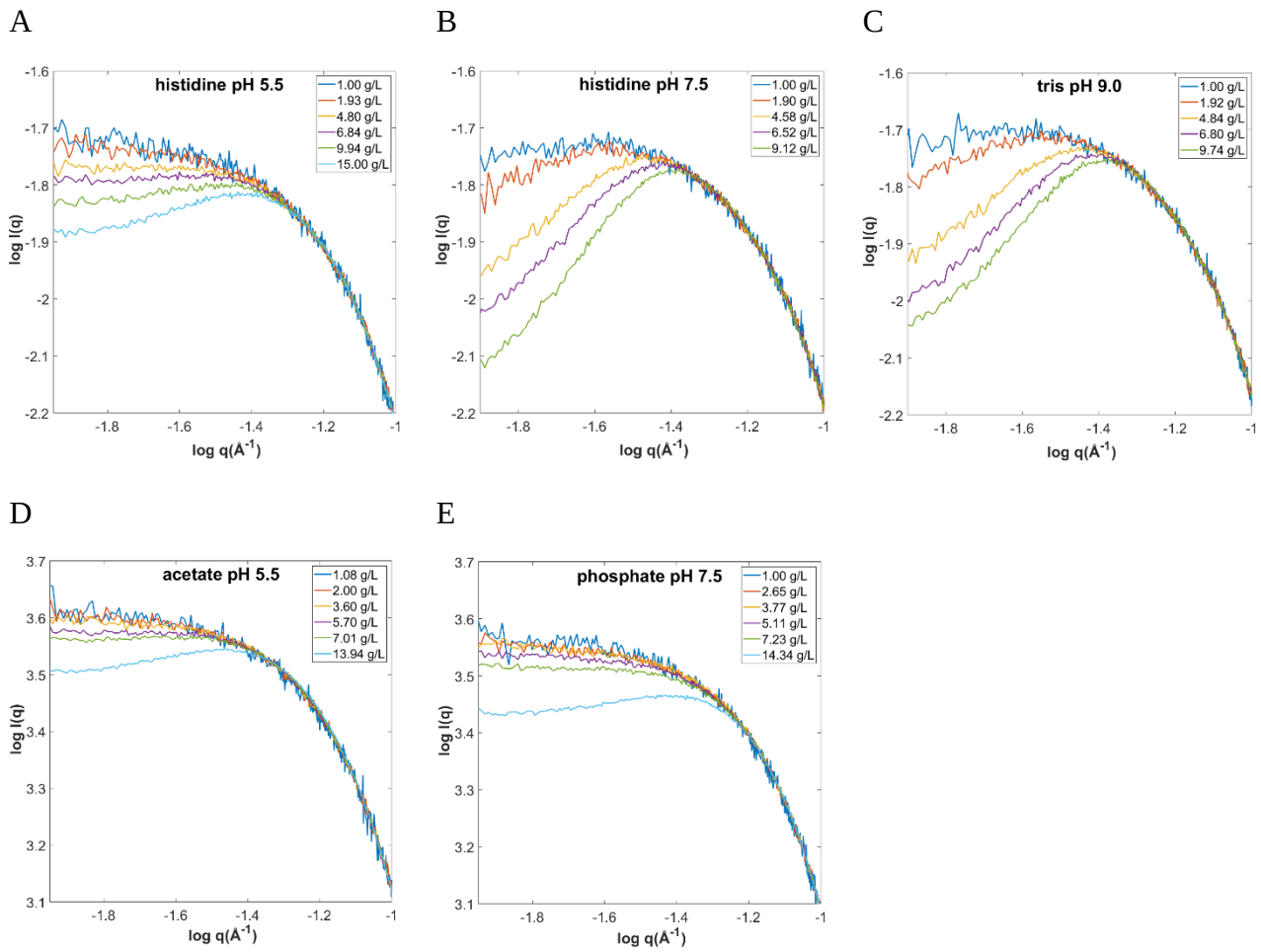


Figure 3.37 SAXS concentration series measurements of lipase: (A) 10 mM histidine pH 5.5, (B) 10 mM histidine pH 7.5, (C) 10 mM tris pH 9.0, (D) 10 mM acetate pH 5.5, (E) 10 mM phosphate pH 7.5.

To study the effect of excipients, SAXS measurements have been performed at pH 5.5 and pH 7.5 in presence of arginine. The scattering intensities increases with increasing arginine concentration at both pH 5.5 and pH 7.5, indicating attraction. Calculated parameters, M_W, D_{max} and R_G (see Table S 3.6) increases with increasing arginine concentration at 5 g/L, however the apparent M_W at pH 5.5 is slightly higher than the monomeric molecular weight, indicating presence of aggregates, while the apparent M_W at pH 7.5 does not exceed the 30 kDa, indicating decrease in repulsion in presence of arginine (see Table S 3.6).

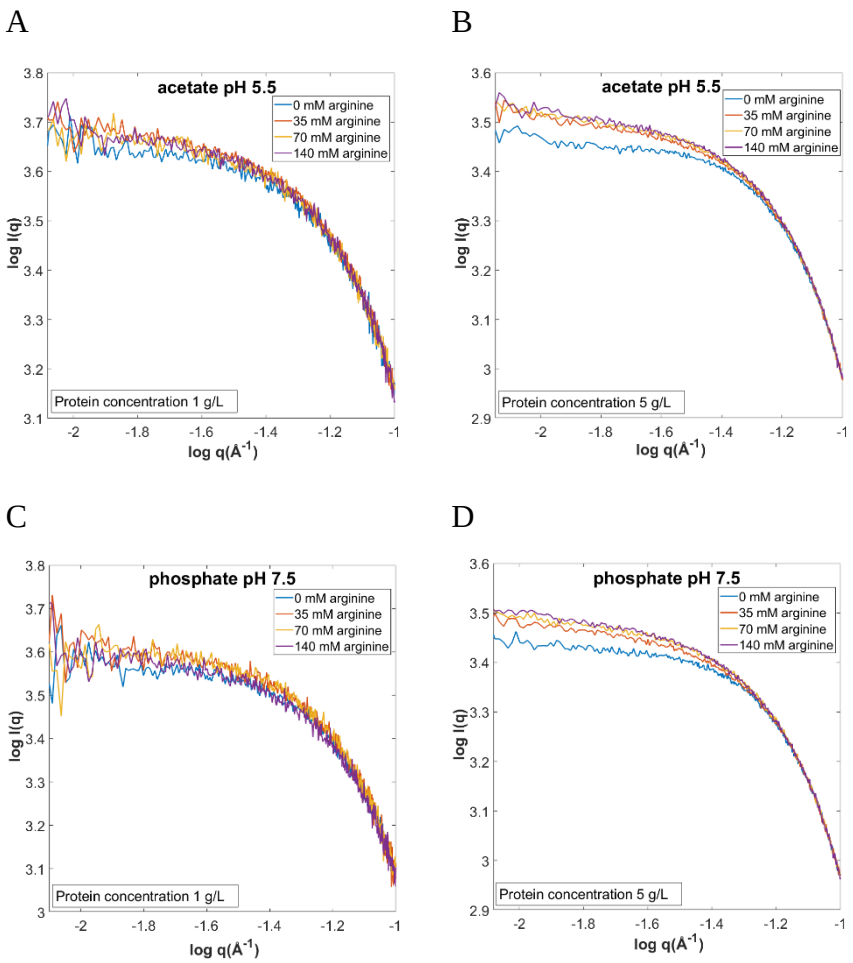


Figure 3.38 Scattering profiles for lipase at 10 mM histidine pH 5.0 with increasing concentration of arginine in at 1 g/L (A) and 5 g/L (B) protein concentrations.

3.4 Overall Discussion

3.4.1 mAbs

pH and NaCl effect

The overall results of the stability experiments for all the mAbs, albumin and lipase are shown in Table 3.5 and Table 3.6. In general, thermal and chemical unfolding studies show different unfolding processes. Thermal unfolding studies of all the mAbs in all tested conditions indicate that, increasing pH leads to increase in stability. $T_{½}$ increases with increasing pH up to 6.5, where it reaches a plateau of maximum stability for PPI02, PPI03, PPI08 and PPI10. For PPI01 however, $T_{½}$ has an upward slope until pH 9.0. All the tested mAbs are less stable at low pH, especially below pH 6.0, which has also been previously observed by *Feng et.al*[35]. In case of HSA and lipase, thermal stability increases from pH 5.0 to around neutral pH and then gradually decreases form pH 7.5 to pH 9.0. Comparing the apparent melting temperature over the pH range, HSA is least thermally stable at low pH while, lipase at high pH. Addition of NaCl, has stabilizing effect on HSA, while it has destabilizing effect on all the mAbs. In case of lipase, NaCl increases the thermal stability ay high pH (from pH 7.0 to pH 9.0), while it has no effect at lower pH.

According to SAXS studies, lipase, HSA and all the mAbs are repulsive in nature (see Figure 3.14, Figure 3.12, Figure 3.22, Figure 3.31, Figure 3.37), except for PPI01 which is attractive in nature at pH 6.5 and below (see Figure 3.18). Moreover, addition of NaCl is screening the charges responsible for the repulsion, leading to a decrease in stability. In case of PPI01, increasing pH from pH 5.0 to pH 6.5 leads to increase in larger aggregates (see Figure 3.19). Based on the chemical denaturation data, $c_{½}$ increases with increasing pH for PPI01 and PPI03, while it moderately decreases for PPI02, PPI03 and PPI10. Addition of NaCl doesn't have any effect on the denaturant chemical stability.

Buffer and excipients effect

In the current study, three different buffers were investigated: acetate, phosphate, and histidine in both pH 5.0 and 6.5 for HSA and mAbs, pH 5.5 and pH 7.5 for lipase. To study the effect of excipients, three different excipients (sucrose, arginine, and proline) were studied in the above mentioned buffers. Differences in $T_{½}$ and $c_{½}$ ($\Delta T_{½}$ and $\Delta c_{½}$) in the presence of acetate and phosphate with respect to histidine, and in the presence of excipients with respect to no excipients were calculated and represented in Table 3.5 and Table 3.6.

Among all the tested buffers, acetate in general increases the thermal stability for HSA and all the mAbs compared to histidine, while it has negligible effect on lipase. In chemical unfolding process, Acetate has a destabilizing effect for the mAbs, while it stabilizes HSA and lipase (see Table 3.5and Table 3.6). At pH 6.5, phosphate increases the overall stability of all the mAbs, (except for PPI01 which was precipitated) in comparison to histidine which is in good agreement with previously performed studies for IgG1 mAbs[38]. SAXS studies for wild type IgG (PPI03) also show less repulsion in presence of phosphate compared to histidine (see Figure 3.14(C, D)) indicating also that phosphate might be screening the charges.

From Table 3.5, the largest excipient effect was observed in presence of arginine. The destabilizing effect of arginine was more pronounced at pH 5.0 for all the mAbs, while at pH 6.5, arginine still has the destabilizing effect mAbs except for PPI01. Also in combination with NaCl, Arginine has minimal positive effect at histidine pH 5.0 for all the mAbs. Moreover, arginine has stabilizing effect on HSA, except in presence of phosphate, while the effect is opposite in case of lipase, i.e., minimal destabilizing effect except in presence of phosphate buffer. Additionally, from SAXS measurements for mAbs, in presence arginine at pH 5.0 a tiny tail has been observed in pair distribution function (see Figure S3.1 in supplementary material) indicates the presence of aggregates, which disappears at pH 6.5. SAXS studies in case of lipase shows that, presence of arginine reduces the repulsion in phosphate buffer and small amount of aggregates are seen in acetate buffer (see Figure 3.38 and Table S 3.6).

The most significant stabilization was observed in presence of sucrose for all the mAbs and lipase in all conditions tested. Proline had a stabilizing effect on the protein stability in presence of histidine, except at pH 6.5 in presence of NaCl, while it had a minimal effect in acetate and phosphate buffers.

The effect of the excipients has been previously studied for mAbs[36,53]. The stabilizing effects of sucrose previously shown for other proteins appeared to be independent of pH[54,55] which is in good agreement with our data. It has been previously observed that the stabilizing effect of sucrose is due to during protein solvent interaction, sucrose is preferentially excluded from the protein domain leads to increase in the free energy of the system which thermodynamically leads to protein stabilization[54]. In contrast,

the combined effect in case of NaCl and arginine is might be due to their interaction with proteins via a charge dependent mechanism which is mediated by pH[35].

Table 3.5 Overview of thermal and chemical unfolding studies of mAbs and HSA at different formulation conditions in combination with pH, NaCl, buffers and excipients.

	Add	**nanoDSF**						**ICD**					
		$\Delta T_{½}$						**$\Delta c_{½}$**					
		PPI01	**PPI02**	**PPI03**	**PPI08**	**PPI10**	**PPI49**	**PPI01**	**PPI02**	**PPI03**	**PPI08**	**PPI10**	**PPI49**
↑ pH		++++	+++	+++	+++	+++	+++	++	-	+++	-	-	-
	NaCl	--	--	--	--	--	+++	0	-	+	0	+	+
Ace pH 5.0		++	+++	++	+++	++	+++	-	0	--	-	-	+
	Sucrose	+	+	+	+	+	+	0	+	+	+	+	+
	Arginine	-	---	--	-	---	++	0	-	0	-	0	0
	Proline	0	0	0	0	0	-	0	0	+	-	+	-
His pH 5.0	Sucrose	+	+	+	+	+	0	+	+	-	--	+	0
	Arginine	-	--	--	--	--	+++	0	+	0	--	0	0
	Proline	+	++	++	+	++	+	0	0	0	-	+	0
+ NaCl	Sucrose	+	+	+	+	+	+	+	+	0	+	+	0
	Arginine	+	0	+	0	0	++	-	0	-	-	0	0
	Proline	++	++	++	+	++	++	0	+	0	+	+	+
Phos pH 6.5		+	--	+	-	0	+++	x	0	-	0	+	+
	Sucrose	+	+	+	+	+	-	x	0	-	-	+	-
	Arginine	++	-	-	+	-	--	x	0	0	-	0	-
	Proline	0	0	0	+	0	0	x	+	0	-	+	0
His pH 6.5	Sucrose	+	+	0	+	+	+	+	0	--	--	+	0
	Arginine	++	--	-	-	--	+++	0	0	0	--	+	+
	Proline	0	0	0	0	0	0	+	+	--	0	+	0
+ NaCl	Sucrose	++	++	0	-	++	+	+	+	0	+	-	0
	Arginine	+	--	--	0	--	0	0	+	--	0	-	0
	Proline	++	+	+	--	+	-	+	+	0	0	--	0

Reference point for buffers: pH 5.0 – histidine pH 5.0; pH 6.5 – histidine pH 6.5
Reference point for excipients: respective buffer without excipient
+ stabilizes
- destabilizes
0 no trend for pH/NaCl study
x data not acquired

++++/----	$\Delta T_{½}$> 10ºC	$\Delta c_{½}$> 1.5 M
+++/---	5ºC <$\Delta T_{½}$< 10ºC	1 M <$\Delta c_{½}$< 1.5 M
++/--	2ºC <$\Delta T_{½}$< 5ºC	0.5 M <$\Delta c_{½}$< 1.0 M
+/-	0.5ºC <$\Delta T_{½}$< 2ºC	0.10 M <$\Delta c_{½}$< 0.5 M
0	$\Delta T_{½}$< 0.5ºC	$\Delta c_{½}$< 0.10 M

Table 3.6 Overview of thermal and chemical unfolding studies of lipase at different formulation conditions in combination with pH, NaCl, buffers and excipients.

	Add	**nanoDSF**	**ICD**
		$\Delta T_{½}$	**$\Delta c_{½}$**
		PPI45	**PPI45**
↑ pH		---	-
	NaCl	0	+
Acetate pH 5.5		0	+
	Sucrose	++	0
	Arginine	-	0
	Proline	+	-
Histidine pH 5.5	Sucrose	++	+
	Arginine	--	++
	Proline	0	0
+ NaCl	Sucrose	++	+
	Arginine	0	0
	Proline	+	--
Phosphate pH 7.5		--	0
	Sucrose	++	+
	Arginine	++	+
	Proline	+	+
Histidine pH 7.5	Sucrose	+	+
	Arginine	-	-
	Proline	0	++
+ NaCl	Sucrose	+	+
	Arginine	+	-
	Proline	+	++

Reference point for buffers: pH 5.5 – histidine pH 5.5; pH 7.5 – histidine pH 7.5
Reference point for excipients: respective buffer without excipient
\+ stabilizes
\- destabilizes
0 no trend for pH/NaCl study
x data not acquired

++++/----	$\Delta T_{½}$> 5ºC	$\Delta c_{½}$> 1.0 M
+++/---	3ºC <$\Delta T_{½}$< 5ºC	0.5 M <$\Delta c_{½}$< 1.0 M
++/--	1ºC <$\Delta T_{½}$< 3ºC	0.2 M <$\Delta c_{½}$< 0.5 M
+/-	0.2ºC <$\Delta T_{½}$< 1ºC	0.05 M <$\Delta c_{½}$< 0.2 M
0	$\Delta T_{½}$<0.2ºC	$\Delta c_{½}$< 0.05 M

3.5 Conclusion

The overall stability study of five mAbs, HSA and one lipase are reported in this chapter with the effect of pH, buffers and excipients on conformational stability. The mAbs, despite having more than 80% sequence similarity in both heavy and light chain, show differences in stability. Four of the mAbs (PPI02, PPI03, PPI08, PPI10) show similar behavior with pH, NaCl, acetate, sucrose and arginine in most of the cases, while PPI01 shows opposite behavior with pH and arginine which has 40% sequence similarity in light chain. HSA and lipase have unique behaviors in the tested conditions. Further, a combination of excipients can result in the opposite effect compared to their individual effects, such as arginine and NaCl alone destabilizes the measured protein, while the combination of arginine and NaCl have stabilizing effect on the proteins. We also observe that thermal and chemical denaturation of the proteins not always complement each other, indicating that the unfolding processes are different. Additionally, SAXS in combination with ICD and nanoDSF gives a better overview of the protein conformational changes and interactions with respect to protein concentrations. The stabilizing and destabilizing effects of buffer components and excipients are difficult to explain at a molecular level due to lack of information about the distinct binding site of the proteins and binding affinities of the different excipients to the proteins. In order to understand these effects, we need to perform some additional measurements in combination with *in-silico* approaches.

3.6 Supplementary Materials

3.6.1 Sequence alignment of mAbs by Clustal2.1

Heavy chain - percent identity matrix and multiple sequence alignment

1: PPI02	100.00	82.77	86.35	85.39
2: PPI01	82.77	100.00	86.25	86.19
3: PPI03	86.35	86.25	100.00	96.21
4: PPI10	85.39	86.19	96.21	100.00

```
sp|PPI02      QVTLRESGPALVKPTQTLTLTCTFSGFSLSTAGMSVGWIRQPPGKALEWLADIWWDDK-K  59
sp|PPI01      EVQLVQSGAEVKKPGATVKISCKVYGYIFTD--YNIYWVRQAPGKGLEWMGLIDPDNGET  58
sp|PPI03      QVNLRESGGGLVQPGGSLRLSCAASGFTFGS--YAMSWVRQAPGKGLEWVSAISGSGGST  58
sp|PPI10      EVQLLESGGGLVQPGGSLRLSCAASGFTFGN--SWMSWVRQAPGKGLEWVSAISGSGGST  58
              :* * :**  : :*  :: ::*   *: :      : *:** ***.***:. *  ..  .

sp|PPI02      HYNPSLKDRLTISKDTSKNQVVLKVTNMDPADTATYYCARDMIF--NFYFDVWGQGTTVT  117
sp|PPI01      FYAEKFQGRATMTADTSSDRAYMELSSLRFEDTAVYYCATVMGKWIKGGYDYWGRGTLVT  118
sp|PPI03      YYADSVKGRFTISRDNSKNSLYLQMNSLRAEDTAVYYCARRSIYGGNYYFDYWGRGTLVT  118
sp|PPI10      YYADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCTRDL--PGIAVAGYWGQGTLVT  116
              .*  ..:.* *:: *.*.:   ::::.:   ***.***:           . **:** **

sp|PPI02      VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVL  177
sp|PPI01      VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVL  178
sp|PPI03      VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVL  178
sp|PPI10      VSSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVL  176
              ************************************************************

sp|PPI02      QSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCPPCPAPEL  237
sp|PPI01      QSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPEL  238
sp|PPI03      QSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPEL  238
sp|PPI10      QSSGLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPEL  236
              ************************************:***********************

sp|PPI02      LGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE  297
sp|PPI01      LGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE  298
sp|PPI03      LGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE  298
sp|PPI10      LGGPSVFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREE  296
              ************************************************************

sp|PPI02      QYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS  357
sp|PPI01      QYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS  358
sp|PPI03      QYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS  358
sp|PPI10      QYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPS  356
              ************************************************************

sp|PPI02      REEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK  417
sp|PPI01      RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK  418
sp|PPI03      RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK  418
sp|PPI10      RDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDK  416
              *:*:********************************************************

sp|PPI02      SRWQQGNVFSCSVMHEALHNHYTQKSLSLSPG-      449
sp|PPI01      SRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK      451
sp|PPI03      SRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK      451
sp|PPI10      SRWQQGNVFSCSVMHEALHNHYTQKSLSLSPGK      449
              ********************************
```

Light chain - percent identity matrix and multiple sequence alignment

1: PPI01	100.00	43.06	42.86	41.90
2: PPI02	43.06	100.00	90.14	92.02
3: PPI03	42.86	90.14	100.00	91.59
4: PPI10	41.90	92.02	91.59	100.00

```
sp|PPI01      -QSVLTQPPSVSGAPGQRVTISCTGSSSNIGAGYDVHWYQQLPGTAPKLLIYDNFNRPSG  59
sp|PPI02      DIQMTQSPSTLSASVGDRVTITCSASSR-VG---YMHWYQQKPGKAPKLLIYDTSKLASG  56
sp|PPI03      DIQMTQSPSSLSASVGDRVTITCRASQSISS---YLNWYQQKPGKAPKLLIYAASSLQSG  57
sp|PPI10      DTQMTQSPSTLSASVGDRVTITCRASEGIYH---WLAWYQQKPGKAPKLLIYKASSLASG  57
                .:  .* ::*.: *:****:* .*.        : **** **.*******   .  **

sp|PPI01      VPPRFSGSKSGTSASLAITGLQAEDEADYYCQSYDSPTLTSPFGTGTLTVLGQPKAAPSV  119
sp|PPI02      VPSRFSGSGSGTEFTLTISSLQPDDFATYYCFQGSGY--PFTFGGGTKVEIKRTVAAPSV  114
sp|PPI03      VPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYST--PLTFGGGSKVEIKRTVAAPSV  115
sp|PPI10      VPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQYSNY--PLTFGGGTKLEIKRTVAAPSV  115
              ** ***** ***. :*:*:.** :* * *** .  .      ** *:   : :  *****

sp|PPI01      TLFPPSSEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGV-ETTTPSKQSNNKYAA  178
sp|PPI02      FIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSL  174
sp|PPI03      FIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSL  175
sp|PPI10      FIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSL  175
               :****.*:*::..*::***:.:***  ..* **.*.:  ...  *:.* ....:..*:

sp|PPI01      SSYLSLTPEQWKSHRSYSCQVTHEGSTV--EKTVAPTECS       216
sp|PPI02      SSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC-       213
sp|PPI03      SSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC-       214
sp|PPI10      SSTLTLSKADYEKHKVYACEVTHQGLSSPVTKSFNRGEC-       214
              ** *:*:  :::.*: *:*:***:* :    *:.   **
```

3.6.2 SAXS additional information's

Table S3.1 Experimental set up of SAXS measurements.

Instrument	**BM29, ESRF**	**P12, DESY**
Wavelength (Å)	0.99	1.24402
q-range ($Å^{-1}$)	0.004 – 0.49	0.00261 – 0.7263
Sample-to-detector distance ($\boldsymbol{m}$)	2.864	3.000
Detector	PILATUS 1M	PILATUS 6M
Exposure time (s)	10 x 1.00	30 x 0.095
Beam size	700 x 700 μm^2	0.2 × 0.12 mm^2
Sample configuration	1.8 mm quartz glass capillary	Quartz glass capillary
Absolute scaling method	Comparison to water in sample capillary	Comparison with scattering from pure H2O
Normalization	To transmitted intensity by beam-stop counter	To transmitted intensity by beam-stop counter
Monitoring for radiation damage	Control of un-subtracted and scaled subtracted data for systematic changes typical for radiation damage	Frame-by-frame comparison

Table S 3.2 An overview of the samples measured by SAXS and data treatment parameters for PPI02.

Formulation condition	**Protein (g/L)**	**NaCl (mM)**	**Arginine (mM)**	**Guinier**		**p(r)**			**Apparent M_W (kDa)**	
				I(0)/c	R_G (nm)	*I(0)/c*	R_G (nm)	D_{max} (nm)	Guinier	P(r)
10 mM histidine pH 6.5	1.00	-	-	0.13	4.98	0.13	5.14	14.90	180	180
	1.99	-	-	0.12	4.71	0.12	4.96	15.20	166	166
	5.54	-	-	0.11	4.32	0.12	4.91	15.02	152	166
	7.83	-	-	0.11	4.41	0.12	4.87	15.01	152	166
	11.25	-	-	0.11	4.60	0.12	5.04	15.50	152	166
	22.78	-	-	0.10	4.37	0.12	4.86	15.00	138	166
	1.00	140	-	0.15	4.97	0.15	5.04	16.34	208	208
	1.99	140	-	0.16	4.91	0.16	4.73	16.87	222	222
	5.54	140	-	0.16	4.87	0.16	5.01	16.80	222	222
	7.83	140	-	0.15	4.91	0.15	4.97	17.02	208	208
	11.25	140	-	0.14	4.82	0.15	4.99	17.04	194	208
	22.78	140	-	0.14	4.94	0.15	4.78	17.01	194	208
	1.24	-	0	20539*	4.27	22320*	4.75	14.10	137	149
	1.21	-	35	22736*	5.02	22790*	5.08	16.66	152	152
	1.19	-	70	23336*	5.03	24540*	5.07	16.90	156	164
	1.22	-	140	24120*	5.06	24260*	5.16	17.15	161	162
10 mM histidine pH 7.0	0.57	140	-	0.18	5.05	0.18	5.18	18.03	249	249
	1.21	140	-	0.18	5.12	0.18	5.24	19.07	249	249
	8.04	140	-	0.22	6.26	0.23	6.58	27.40	305	319
	10.24	140	-	0.22	6.20	0.22	6.67	30.48	305	305
	13.64	140	-	0.21	6.36	0.22	6.61	27.58	291	305

Note: *DESY – BSA calibration; ¤ - data collected in ESRF

Table S 3.3 An overview of the samples measured by SAXS and data treatment parameters for PPI03.

Formulation condition	**Protein (g/L)**	**NaCl (mM)**	**Arginine (mM)**	**Guinier**		**p(r)**			**Apparent M_W (kDa)**	
				I(0)/c	R_G (nm)	*I(0)/c*	R_G (nm)	D_{max} (nm)	Guinier	P(r)
10 mM histidine pH 6.0	0.85	-	-	0.11	4.92	0.11	5.05	15.65	152	152
	1.74	-	-	0.10	4.60	0.10	4.98	15.50	137	138
	4.66	-	-	0.09	4.53	0.10	4.89	15.01	130	138
	6.45	-	-	0.09	4.52	0.10	4.95	15.26	127	138
	9.39	-	-	0.09	4.60	0.10	4.95	15.49	130	138
	19.28	-	-	0.09	4.38	0.10	4.89	15.00	119	138
	33.77	-	-	0.08	4.20	0.08	4.61	14.50	107	111
	5.00	0	-	102.55¤	4.7	117.00¤	4.78	14.38	114	130
	5.00	35	-	131.68¤	4.96	132.20¤	5.03	15.60	146	146
	4.81	70	-	133.91¤	5.07	133.70¤	5.10	16.10	148	148
	5.00	140	-	136.05¤	5.11	136.50¤	5.20	16.14	151	151
10 mM histidine pH 6.5	0.97	-	-	0.10	4.85	0.11	4.98	15.30	138	152
	2.00	-	-	0.10	4.59	0.10	4.91	14.92	138	138
	5.32	-	-	0.10	4.67	0.10	4.89	15.12	136	138
	7.23	-	-	0.10	4.66	0.10	4.98	15.00	136	138
	10.32	-	-	0.10	4.65	0.11	4.99	15.00	134	152
	21.69	-	-	0.09	4.53	0.10	4.95	15.50	129	138
	30.78	-	-	0.09	4.39	0.10	4.86	15.00	122	138
	51.74	-	-	0.08	4.13	0.08	4.57	14.00	105	111
	5.32	-	-	0.10	4.67	0.10	4.89	15.12	136	138
	4.85	-	35	20841*	4.93	20977*	5.01	15.51	140	142
	4.89	-	70	21226*	5.03	21276*	5.10	16.63	142	143
	4.94	-	140	21049*	5.04	21126*	5.12	16.90	141	143
10 mM phosphate pH 6.5	1.00	-	-	26123*	5.78	25980*	5.85	20.22	174	173
	1.97	-	-	23800*	5.37	23540*	5.35	17.17	159	157
	4.18	-	-	24575*	5.47	24400*	5.50	18.34	164	163
	5.65	-	-	24639*	5.71	24410*	5.72	18.87	164	163
	7.94	-	-	22557*	5.21	21970*	5.07	14.39	151	147
	15.60	-	-	22304*	5.10	21900*	5.03	20.22	149	146
	4.18	-	-	24575*	5.47	24400*	5.50	18.34	164	163
	4.32	-	35	17391	5.08	17430	5.15	17.44	116	116
	4.40	-	70	16980	5.07	17020	5.13	17.08	113	114
	4.36	-	140	16134	5.06	16180	5.13	16.00	108	108

Table S 3.4 An overview of the samples measured by SAXS and data treatment parameters for PPI08.

Formulation condition	**Protein (g/L)**	**Additives (mM)**	**Guinier**		**p(r)**			**Apparent M_W (kDa)**	
			I(0)/c	R_G (nm)	*I(0)/c*	R_G (nm)	D_{max} (nm)	Guinier	P(r)
10 mM histidine pH 6.0	0.99	-	0.14	5.71	0.14	5.86	19.06	194	194
	2.06	-	0.13	5.48	0.13	5.80	18.59	180	180
	5.37	-	0.11	4.71	0.12	5.58	17.40	152	166
	7.52	-	0.11	4.97	0.12	5.59	17.56	152	166
	11.02	-	0.10	4.09	0.12	5.41	16.99	137	166
	22.10	-	0.10	4.92	0.12	5.65	18.00	138	166
	30.81	-	0.10	4.77	0.11	5.42	17.00	136	152
10 mM histidine pH 6.5	1.04	-	0.13	5.72	0.13	5.88	18.98	180	180
	1.98	-	0.13	5.50	0.13	5.73	16.51	180	180
	5.28	-	0.12	5.24	0.13	5.69	17.94	166	180
	7.41	-	0.11	4.89	0.12	5.64	17.99	152	166
	10.48	-	0.11	4.76	0.12	5.59	17.90	152	166
	21.53	-	0.11	5.12	0.12	5.63	18.00	152	166
	32.05	-	0.11	5.20	0.12	5.61	18.00	152	166
	54.66	-	0.10	4.77	0.11	5.44	18.00	133	152

Table S 3.5 An overview of the samples measured by SAXS and data treatment parameters for HSA.

Formulation condition	**Protein (g/L)**	**Additives (mM)**	**Guinier**		**p(r)**			**Apparent M_W (kDa)**	
			I(0)/c	R_G (nm)	*I(0)/c*	R_G (nm)	D_{max} (nm)	Guinier	P(r)
10 mM histidine pH 6.0	1.37	-	0.05	2.76	0.05	2.79	8.40	73	69
	2.34	-	0.05	2.70	0.05	2.76	7.97	71	69
	5.01	-	0.05	2.64	0.05	2.75	7.98	69	69
	6.88	-	0.05	2.65	0.05	2.77	8.15	69	69
	9.59	-	0.05	2.55	0.05	2.76	8.15	66	69
	19.16	-	0.04	2.34	0.05	2.77	8.00	61	69
	46.69	-	0.04	2.24	0.05	2.73	7.80	54	69
	90.47	-	0.03	2.19	0.04	2.64	7.50	44	55
	145.27	-	0.02	1.74	0.04	2.51	6.90	29	55
10 mM histidine pH 7.0	1.05	-	0.05	2.81	0.05	2.83	8.68	68	69
	2.06	-	0.05	2.81	0.05	2.84	8.70	66	69
	4.81	-	0.05	2.82	0.05	2.87	8.80	68	69
	6.82	-	0.05	2.79	0.05	2.83	8.50	66	69
	9.53		0.05	2.82	0.05	2.85	8.90	68	69
	19.13		0.05	2.80	0.05	2.85	9.00	68	69
	28.38		0.05	2.80	0.05	2.87	9.00	66	69
	48.30	-	0.05	2.75	0.05	2.82	8.50	64	69

Table S 3.6 An overview of the samples measured by SAXS and data treatment parameters for lipase.

Formulation condition	**Protein (g/L)**	**NaCl (mM)**	**Arginine (mM)**	**Guinier**		**p(r)**			**Apparent M_W (kDa)**	
				I(0)/c	R_G (nm)	*I(0)/c*	R_G (nm)	D_{max} (nm)	Guinier	P(r)
10 mM histidine pH 5.5	1.00	-	-	0.02	1.88	0.02	1.83	5.53	28	28
	1.93	-	-	0.02	1.83	0.02	1.81	5.76	26	28
	4.80	-	-	0.02	1.83	0.02	1.81	5.74	26	28
	6.84	-	-	0.02	1.82	0.02	1.82	6.41	26	28
	9.94	-	-	0.02	1.82	0.02	1.83	6.58	26	28
	15.00	-	-	0.02	1.69	0.02	1.80	6.97	25	28
10 mM acetate pH 5.5	1.08	-	-	4189*	1.84	4173*	1.82	5.99	28	28
	2.00	-	-	4158*	1.80	4163*	1.80	5.56	28	28
	3.60	-	-	4218*	1.83	4199*	1.81	6.06	28	28
	5.70	-	-	4209*	1.83	4220*	1.83	6.06	28	28
	7.01	-	-	4236*	1.83	4258*	1.85	6.09	28	28
	13.94	-	-	3704*	1.18	4204*	1.82	6.08	25	28
	1.02	-	0	4523*	1.84	4506*	1.82	5.99	30	30
	1.08	-	35	4828*	1.92	4789*	1.88	6.03	32	32
	1.05	-	70	4725*	1.91	4695*	1.88	6.13	32	31
	1.12	-	140	4731*	1.90	4703*	1.87	6.02	32	31
	4.89	-	0	4037*	1.83	4045*	1.82	5.74	27	27
	4.98	-	35	4157*	1.89	4149*	1.88	6.39	28	28
	4.65	-	70	4220*	1.92	4211*	1.91	6.50	28	28
	4.79	-	140	4245*	1.93	4241*	1.92	6.97	28	28
10 mM histidine pH 7.5	1.00	-	-	0.02	1.84	0.02	1.80	5.20	29	28
	1.90	-	-	0.02	1.82	0.02	1.80	5.41	28	28
	4.58	-	-	0.02	1.87	0.02	1.83	5.84	29	28
	6.52	-	-	0.02	1.88	0.02	1.83	5.58	29	28
	9.12	-	-	0.02	1.89	0.02	1.86	6.14	29	28
10 mM phosphate pH 7.5	1.00	-	-	3807*	1.80	3809*	1.79	5.36	25	25
	2.65	-	-	3721*	1.73	3760*	1.76	4.81	25	25
	3.77	-	-	3808*	1.78	3841*	1.79	5.47	25	26
	5.11	-	-	3782*	1.78	3814*	1.78	5.30	25	25
	7.23	-	-	3680*	1.71	3791*	1.78	5.17	25	25
	14.34	-	-	3410*	1.54	3727*	1.76	5.01	23	25
	1.03	-	0	3807*	1.80	3809*	1.79	5.36	25	25
	1.22	-	35	4065*	1.89	4023*	1.84	5.47	27	27
	1.15	-	70	4083*	1.90	4054*	1.86	5.79	27	27
	1.07	-	140	3876*	1.91	3853*	1.86	5.86	26	26
	5.01	-	0	3873*	1.78	3899*	1.79	5.47	26	26
	4.92	-	35	3979*	1.84	3973*	1.82	5.86	27	27

	5.11	-	70	4069*	1.89	4056*	1.86	5.90	27	27
	5.09	-	140	4095*	1.91	4077*	1.88	5.99	27	27
10 mM tris pH 9.0	1.00	-	-	0.02	1.84	0.02	1.82	5.76	30	28
	1.92	-	-	0.02	1.86	0.02	1.82	5.54	30	28
	4.84	-	-	0.02	1.88	0.02	1.83	5.60	30	28
	6.80	-	-	0.02	1.89	0.02	1.82	5.32	30	28
	8.74	-	-	0.02	1.85	0.02	1.91	7.43	30	28

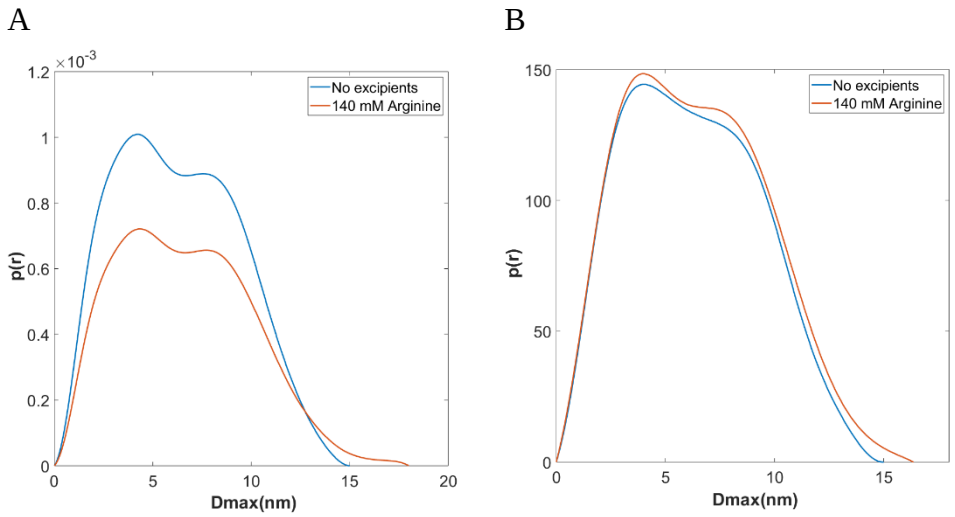


Figure S3.1 Effect of arginine on changes in p(r) functions at (A): PPI 10 at pH 5.0 and (B): PPI02 at pH 6.5. Experiments performed at a concentration around 5 g/L.

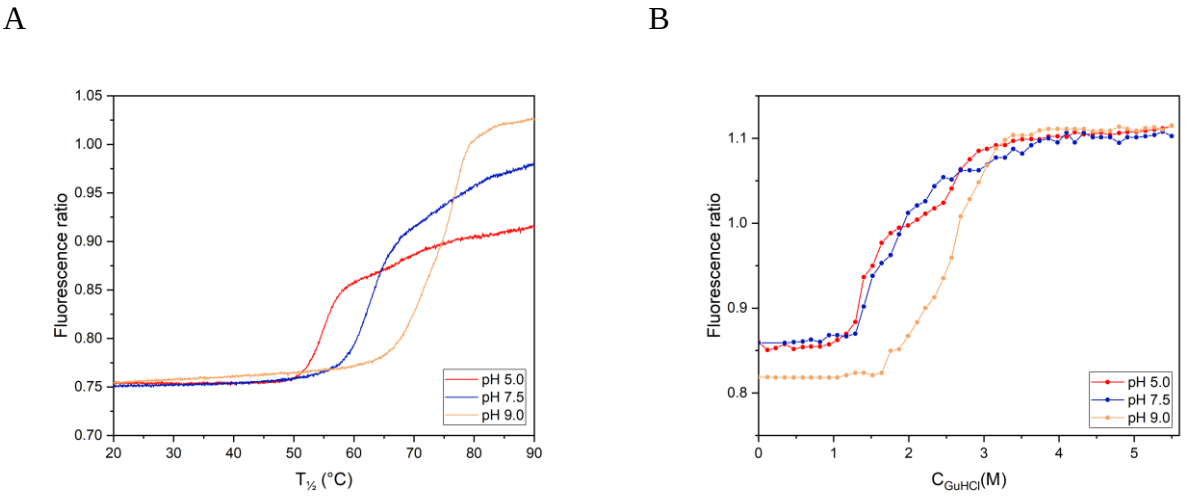


Figure S 3.2 (A) Thermal and (B) chemical unfolding curves of PPI01 at three different pH. Color code: 10 mM histidine pH 5.0 (red), 10 mM histidine pH 7.7 (blue), 10 mM tris pH 9.0 (yellow).

3.6.3 Sequences

Human IgG1κ (PPI-03)

Heavy chain

QVNLRESGGGLVQPGGSLRLSCAASGFTFGSYAMSWVRQAPGKGLEWVSAISGSGGSTYY
ADSVKGRFTISRDNSKNSLYLQMNSLRAEDTAVYYCARRSIYGGNYYFDYWGRGTLVTVS
SASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSG
LYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVF
LFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRV
VSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQV
SLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFS
CSVMHEALHNHYTQKSLSLSPGK

N-glycosylation site

Light chain

DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYAASSLQSGVPSRFS
GSGSGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGGGSKVEIKRTVAAPSVFIFPPSDEQLKS
GTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
HKVYACEVTHQGLSSPVTKSFNRGEC

Human IgG1 (PPI-10)

Heavy chain

EVQLLESGGGLVQPGGSLRLSCAASGFTFGNSWMSWVRQAPGKGLEWVSAISGSGGSTYY
ADSVKGRFTISRDNSKNTLYLQMNSLRAEDTAVYYCTRDLPGIAVAGYWGQGTLVTVSSAS
TKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYS
LSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVFLFP
PKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVVS
VLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKNQVSL
TCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCS
VMHEALHNHYTQKSLSLSPGK

N-glycosylation site

Light chain

DTQMTQSPSTLSASVGDRVTITCRASEGIYHWLAWYQQKPGKAPKLLIYKASSLASGVPSRF
SGSGSGTEFTLTISSLQPDDFATYYCQQYSNYPLTFGGGTKLEIKRTVAAPSVFIFPPSDEQLK
SGTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYE
KHKVYACEVTHQGLSSPVTKSFNRGEC

Human IgG1λ (PPI-01)

Heavy chain

EVQLVQSGAEVKKPGATVKISCKVYGYIFTDYNIYWVRQAPGKGLEWMGLIDPDNGETFY
AEKFQGRATMTADTSSDRAYMELSSLRFEDTAVYYCATVMGKWIKGGYDYWGRGTLVTV
SSASTKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSS
GLYSLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPS
VFLFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTY
RVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSRDELTKN
QVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNV
FSCSVMHEALHNHYTQKSLSLSPGK

N-glycosylation site

Light chain

QSVLTQPPSVSGAPGQRVTISCTGSSSNIGAGYDVHWYQQLPGTAPKLLIYDNFNRPSGVPP
RFSGSKSGTSASLAITGLQAEDEADYYCQSYDSPTLTSPFGTGTLTVLGQPKAAPSVTLFPPS
SEELQANKATLVCLISDFYPGAVTVAWKADSSPVKAGVETTTPSKQSNNKYAASSYLSLTPE
QWKSHRSYSCQVTHEGSTVEKTVAPTECS

IgG1κ (PPI-02)

Heavy chain (by peptide digest)

QVTLRESGPALVKPTQTLTLTCTFSGFSLSTAGMSVGWIRQPPGKALEWLADIWWDDKKHY
NPSLKDRLTISKDTSKNQVVLKVTNMDPADTATYYCARDMIFNFYFDVWGQGTTVTVSSA
STKGPSVFPLAPSSKSTSGGTAALGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLY
SLSSVVTVPSSSLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCPPCPAPELLGGPSVFLF
PPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNSTYRVV
SVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKGQPREPQVYTLPPSREEMTKNQVS
LTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSC
SVMHEALHNHYTQKSLSLSPG

N-glycosylation site

Light chain (by peptide digest)

DIQMTQSPSTLSASVGDRVTITCSASSRVGYMHWYQQKPGKAPKLLIYDTSKLASGVPSRFS
GSGSGTEFTLTISSLQPDDFATYYCFQGSGYPFTFGGGTKVEIKRTVAAPSVFIFPPSDEQLKS
GTASVVCLLNNFYPREAKVQWKVDNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEK
HKVYACEVTHQGLSSPVTKSFNRGEC

Human Serum albumin (PPI49)

MKWVTFISLLFLFSSAYSRGVFRRDAHKSEVAHRFKDLGEENFKALVLIAFAQYLQQCPFED
HVKLVNEVTEFAKTCVADESAENCDKSLHTLFGDKLCTVATLRETYGEMADCCAKQEPER
NECFLQHKDDNPNLPRLVRPEVDVMCTAFHDNEETFLKKYLYEIARRHPYFYAPELLFFAK
RYKAAFTECCQAADKAACLLPKLDELRDEGKASSAKQRLKCASLQKFGERAFKAWAVARL
SQRFPKAEFAEVSKLVTDLTKVHTECCHGDLLECADDRADLAKYICENQDSISSKLKECCEK
PLLEKSHCIAEVENDEMPADLPSLAADFVESKDVCKNYAEAKDVFLGMFLYEYARRHPDYS
VVLLLRLAKTYETTLEKCCAAADPHECYAKVFDEFKPLVEEPQNLIKQNCELFEQLGEYKF
QNALLVRYTKKVPQVSTPTLVEVSRNLGKVGSKCCKHPEAKRMPCAEDYLSVVLNQLCVL
HEKTPVSDRVTKCCTESLVNRRPCFSALEVDETYVPKEFNAETFTFHADICTLSEKERQIKKQ
TALVELVKHKPKATKEQLKAVMDDFAAFVEKCCKADDKETCFAEEGKKLVAASQAALGL

Lipase (PPI-45)

SIDGGIRAATSQEINELTYYTTLSANSYCRTVIPGATWDCIHCDATEDLKIIKTWSTLIYDTNA
MVARGDSEKTIYIVFRGSSSIRNWIADLTFVPVSYPPVSGTKVHKGFLDSYGEVQNELVATV
LDQFKQYPSYKVAVTGHSLGGATALLCALDLYQREEGLSSSNLFLYTQGQPRVGNPAFANY
VVSTGIPYRRTVNERDIVPHLPPAAFGFLHAGSEYWITDNSPETVQVCTSDLETSDCSNSIVPF
TSVLDHLSYFGINTGLCT

References

1. Carter, P. J. Introduction to current and future protein therapeutics: A protein engineering perspective. *Exp. Cell Res.* **317**, 1261–1269 (2011).

2. Sutradhar, K. B., Khatun, S., Al Mamun, A. & Begum, M. Distribution and elimination of protein therapeutics: A review. *Stamford J. Pharm. Sci.* **4**, 1–12 (2011).

3. Goeddel, D. V., Kleid, D. G. & Bolivar, F. Expression in Escherichia coli of chemically synthesized genes for human insulin. *Proc. Natl. Acad. Sci. U. S. A.* (1979). doi:10.1073/pnas.76.1.106

4. Chi, E. Y., Krishnan, S., Randolph, T. W. & Carpenter, J. F. Physical stability of proteins in aqueous solution: Mechanism and driving forces in nonnative protein aggregation. *Pharmaceutical Research* (2003). doi:10.1023/A:1025771421906

5. Manning, M. C., Chou, D. K., Murphy, B. M., Payne, R. W. & Katayama, D. S. Stability of protein pharmaceuticals: An update. *Pharm. Res.* **27**, 544–575 (2010).

6. Smith, S. L. Ten years of Orthoclone OKT3 (muromonab-CD3): A review. *J. Transpl. Coord.* (1996).

7. Weiner, G. J. Building better monoclonal antibody-based therapeutics. *Nature Reviews Cancer* (2015). doi:10.1038/nrc3930

8. Kaplon, H. & Reichert, J. M. Antibodies to watch in 2019. *MAbs* (2019). doi:10.1080/19420862.2018.1556465

9. Beck, A., Wurch, T., Bailly, C. & Corvaia, N. Strategies and challenges for the next generation

of therapeutic antibodies. *Nature Reviews Immunology* (2010). doi:10.1038/nri2747

10. Hill, R. L., Delaney, R., Fellows, R. E. & Lebovitz, H. E. THE EVOLUTIONARY ORIGINS OF THE IMMUNOGLOBULINS. *Proc. Natl. Acad. Sci.* (1966). doi:10.1073/pnas.56.6.1762

11. Harlow, E. & Lane, D. *Using antibodies: a laboratory manual, CSHL Press; 1998. Antibodies A Laboratory Manual Illustrated* (1999).

12. Saphire, E. O. *et al.* Crystal structure of a neutralizing human IgG against HIV-1: A template for vaccine design. *Science (80-.).* (2001). doi:10.1126/science.1061692

13. Schrödinger, L. The PyMol Molecular Graphics System, Versión 1.8. *Thomas Holder* (2015). doi:10.1007/s13398-014-0173-7.2

14. Poljak, R. J., Amzel, L. M. & Phizackerley, R. P. Studies on the three-dimensional structure of immunoglobulins. *Progress in Biophysics and Molecular Biology* (1978). doi:10.1016/0079-6107(78)90005-6

15. Vidarsson, G., Dekkers, G. & Rispens, T. IgG subclasses and allotypes: From structure to effector functions. *Front. Immunol.* **5**, 1–17 (2014).

16. Capra, J. D. & Kehoe, J. M. Variable region sequences of five human immunoglobulin heavy chains of the V(H)III subgroup: Definitive identification of four heavy chain hypervariable regions. *Proc. Natl. Acad. Sci. U. S. A.* (1974). doi:10.1073/pnas.71.3.845

17. Sedykh, S. E., Prinz, V. V., Buneva, V. N. & Nevinsky, G. A. Bispecific antibodies: Design, therapy, perspectives. *Drug Design, Development and Therapy* (2018). doi:10.2147/DDDT.S151282

18. Spiess, C., Zhai, Q. & Carter, P. J. Alternative molecular formats and therapeutic applications for bispecific antibodies. *Molecular Immunology* (2015). doi:10.1016/j.molimm.2015.01.003

19. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* (2003). doi:10.1093/nar/gkg563

20. Kratz, F. & Elsadek, B. Clinical impact of serum proteins on drug delivery. *Journal of Controlled Release* (2012). doi:10.1016/j.jconrel.2011.11.028

21. Peters, T. & Peters, T. 2 – The Albumin Molecule: Its Structure and Chemical Properties. *All About Albumin* (1995). doi:10.1016/B978-012552110-9/50004-0

22. Curry, S., Mandelkow, H., Brick, P. & Franks, N. Crystal structure of human serum albumin complexed with fatty acid reveals an asymmetric distribution of binding sites. *Nat. Struct. Biol.* (1998). doi:10.1038/1869

23. Kratz, F. Albumin as a drug carrier: Design of prodrugs, drug conjugates and nanoparticles. *J. Control. Release* (2008). doi:10.1016/j.jconrel.2008.05.010

24. Fanali, G. *et al.* Human serum albumin: From bench to bedside. *Molecular Aspects of Medicine* (2012). doi:10.1016/j.mam.2011.12.002

25. Hawe, A. & Frieß, W. Formulation development for hydrophobic therapeutic proteins. *Pharmaceutical Development and Technology* (2007). doi:10.1080/10837450701247350

26. Sleep, D., Cameron, J. & Evans, L. R. Albumin as a versatile platform for drug half-life

extension. *Biochimica et Biophysica Acta - General Subjects* (2013). doi:10.1016/j.bbagen.2013.04.023

27. Larsen, M. T., Kuhlmann, M., Hvam, M. L. & Howard, K. A. Albumin-based drug delivery: harnessing nature to cure disease. *Mol. Cell. Ther.* (2016). doi:10.1186/s40591-016-0048-8

28. Jaeger, K. E. & Eggert, T. Lipases for biotechnology. *Current Opinion in Biotechnology* (2002). doi:10.1016/S0958-1669(02)00341-5

29. Schmid, R. D. & Verger, R. Lipases: Interfacial Enzymes with Attractive Applications. *Angew. Chemie Int. Ed.* (1998). doi:10.1002/(sici)1521-3773(19980703)37:12<1608::aid-anie1608>3.3.co;2-m

30. Margolin, A. L. Enzymes in the synthesis of chiral drugs. *Enzyme and Microbial Technology* (1993). doi:10.1016/0141-0229(93)90149-V

31. Gotor-Fernández, V., Brieva, R. & Gotor, V. Lipases: Useful biocatalysts for the preparation of pharmaceuticals. *J. Mol. Catal. B Enzym.* (2006). doi:10.1016/j.molcatb.2006.02.010

32. Carvalho, A. C. L. D. M. *et al.* Recent advances in lipase-mediated preparation of pharmaceuticals and their intermediates. *International Journal of Molecular Sciences* (2015). doi:10.3390/ijms161226191

33. Derewenda, Z. S., Derewenda, U. & Dodson, G. G. The crystal and molecular structure of the Rhizomucor miehei triacylglyceride lipase at 1.9 Å resolution. *J. Mol. Biol.* (1992). doi:10.1016/0022-2836(92)90225-9

34. Brzozowski, A. M. *et al.* A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. *Nature* (1991). doi:10.1038/351491a0

35. He, F., Hogan, S., Latypov, R. F., Narhi, L. O. & Razinkov, V. I. High throughput thermostability screening of monoclonal antibody formulations. *J. Pharm. Sci.* (2010). doi:10.1002/jps.21955

36. Tian, F. *et al.* Spectroscopic evaluation of the stabilization of humanized monoclonal antibodies in amino acid formulations. *Int. J. Pharm.* (2007). doi:10.1016/j.ijpharm.2006.10.037

37. Wang, W., Singh, S., Zeng, D. L., King, K. & Nema, S. Antibody structure, instability, and formulation. *J. Pharm. Sci.* **96**, 1–26 (2007).

38. He, F. *et al.* High-throughput dynamic light scattering method for measuring viscosity of concentrated protein solutions. *Anal. Biochem.* **399**, 141–143 (2010).

39. Augustijn, D. *et al.* Novel Non-linear Curve Fitting to Resolve Protein Unfolding Transitions in Intrinsic Fluorescence Differential Scanning Fluorimetry. *Eur. J. Pharm. Biopharm.* (2019). doi:10.1016/J.EJPB.2019.06.001

40. Gentiluomo, L., Svilenov, H., Augustijn., D., El Bialy, I., Greco, M.L., Kulakova, A., Indrakumar, S., Mahapatra, S., Morales, M., Pohl, C., ..., Harris, P., Frieß, W. Boosting therapeutic protein development by publicly available datasets including comprehensive computational and biophysical characterization. *Manuscr. Submitt.* (2019).

41. Pernot, P. *et al.* New beamline dedicated to solution scattering from biological macromolecules at the ESRF. in *Journal of Physics: Conference Series* (2010). doi:10.1088/1742-

6596/247/1/012009

42. Blanchet, C. E. *et al.* Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY). *J. Appl. Crystallogr.* (2015). doi:10.1107/S160057671500254X

43. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* (2003). doi:10.1107/S0021889803012779

44. Franke, D. *et al.* ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* (2017). doi:10.1107/S1600576717007786

45. Semenyuk, A. V. & Svergun, D. I. GNOM. A program package for small-angle scattering data processing. *J. Appl. Crystallogr.* (1991). doi:10.1107/S002188989100081X

46. Mylonas, E. & Svergun, D. I. Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. in *Journal of Applied Crystallography* (2007). doi:10.1107/S002188980700252X

47. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* (2011). doi:10.1038/msb.2011.75

48. Rayner, L. E. *et al.* The solution structures of two human IgG1 antibodies show conformational stability and accommodate their C1q and FcγR ligands. *J. Biol. Chem.* (2015). doi:10.1074/jbc.M114.631002

49. Abe, Y., Gor, J., Bracewell, D. G., Perkins, S. J. & Dalby, P. A. Masking of the Fc region in human IgG4 by constrained X-ray scattering modelling: Implications for antibody function and therapy. *Biochem. J.* (2010). doi:10.1042/BJ20100641

50. Mosbaek, C. R., Konarev, P. V., Svergun, D. I., Rischel, C. & Vestergaard, B. High concentration formulation studies of an IgG2 antibody using small angle x-ray scattering. *Pharm. Res.* (2012). doi:10.1007/s11095-012-0751-3

51. Gentiluomo, L. *et al.* Characterization of native reversible self-association of a monoclonal antibody mediated by Fab-Fab interaction. *J. Pharm. Sci.* (2019). doi:10.1016/j.xphs.2019.09.021

52. Skou, S., Gillilan, R. E. & Ando, N. Synchrotron-based small-angle X-ray scattering of proteins in solution. *Nat. Protoc.* (2014). doi:10.1038/nprot.2014.116

53. Arakawa, T. & Timasheff, S. N. The stabilization of proteins by osmolytes. *Biophys. J.* (1985). doi:10.1016/S0006-3495(85)83932-1

54. Arakawa, T. & Timasheff, S. N. Stabilization of Protein Structure by Sugars. *Biochemistry* **21**, 6536–6544 (1982).

55. Jones, L. S. *et al.* The effects of Tween 20 and sucrose on the stability of anti-L-selectin during lyophilization and reconstitution. *J. Pharm. Sci.* (2001). doi:10.1002/jps.1098

4 Monoclonal antibodies at high concentrations

4.1 Introduction

Monoclonal antibody (mAb) derived products are an important factor in the treatment of immunological disorders and have a major contribution in biotherapeutics, with many still are in the process of development[1–3]. As of 2019, more than 80 mAb therapeutics have been granted market approval and are one of the rapidly growing biotherapeutics during the past several years[4]. In the administration, often they are injected at very high protein concentrations (100-200 g/L) for subcutaneous delivery and predicting a stable solution behavior makes the formulation development challenging[5].

Proteins at high concentrations tend to self-associate, which can cause instability potentially leading to opalescence, liquid-liquid phase separation, high viscosity solutions, where concentrated protein solutions are required to meet patient dose requirements[6–8]. Protein aggregation is another important issue, where aggregated proteins are a risk factor for protein drug immunogenicity due to conformational changes in the structure and self-association which leads to a formation of large particles[9,10]. Absorption of any chemical modification such as oxidation in the vial are additional concerns that could lead to degradation and loss of activity if the three-dimensional structure of a protein is not conserved[11]. The interplay between formulation attributes of a protein, their molecular structure and aggregation are not thoroughly understood at a molecular level, which implies that there is a lack of knowledge to bridge the formulation attributes of proteins with physical stability and their three-dimensional structure. In order to overcome these challenges, several factors need to be considered, such as pH, ionic strength, excipients, etc[12,13]. To minimize the sample consumption and their wastage, a stable formulation condition should be determined at an early stage during the drug discovery process.

The molecular basis of protein-protein interactions at high concentrations is poorly understood due to their transient nature and inability to capture the experimental behavior. Various biophysical techniques are available to predict and extract the information about the protein-protein interactions between mAbs, such as static or dynamic light scattering[14,15], rheology[16,17], molecular modeling[18], coarse-grained modeling[19] and X-ray and neutron scattering[19–21]. Protein intermolecular interactions are not limited to stability or structural studies but also molecular crowding, long-range and short-range interactions and cluster formation which triggered by electrostatic and hydrophobic interactions[22–24].

Previous studies have been performed on α-crystalline found in the eye lens using small-angle X-ray scattering (SAXS) showed how this method could be used for studying the concentrated solution behaviors of proteins[25,26]. Here, a theoretical model has been used to study the interactions by using electric double-layer potential where the system is interacting through electrostatic repulsive interactions[27]. Small-angle neutron scattering studies performed by Sjöberg and Mortensen on human serum albumin (HSA) in D_2O in the presence and absence of NaCl and fitted to an ellipsoid model by Monte Carlo simulation using a square well repulsive potential at low ionic strength and a short-ranged repulsive Yukawa at high ionic strength[28,29]. Recently, SAXS and static light scattering studies have been performed on HSA at high concentrations (up to 450g/L) to study the solution properties as a function of ionic strength by measuring the structure factors[30]. The authors fitted the structure factors using a hard sphere and ellipsoid model with repulsive Yukawa potential to capture the behavior over the concentration range, where the ellipsoid model shows a better fit for the thermodynamic data.

In the present work, we focus on the interactions study of four mAbs (PPI02, PPI03, PPI08 and PPI10) at high concentrations under specific formulation conditions where they remain soluble. The formulation conditions used for these studies have been determined by a large screening effort of an EU funded project called PIPPI which aims to study a diverse set of proteins and deals with different physicochemical conditions by varying pH, ionic strength and excipients to find stable formulation attributes.

As the intermolecular solution structure of proteins is the controlling factor in determining the formulation attributes. This has been investigated using SAXS and static light scattering (SLS) under specific well-behaved formulation conditions to determine the solution structure and self-interaction of four mAbs and the effect of ionic strength. SAXS experiments provide information about the spatial protein density distribution in terms of the structure factor, $S(q)$, which is a measure of the inter-particle interaction, because of attraction or repulsion. SLS measurement provides the weight-averaged molar mass and the osmotic second virial coefficient, B_{22}[30–32]. The mean spherical approximation has been used to fit the structure factors from SAXS using two hard sphere theoretical models (Yukawa and Baxter) to capture the behavior at high protein concentration. The first model accounts for repulsion using Yukawa interaction potential and the second model accounts for both repulsion using Yukawa potential and short-range attraction using Baxter adhesive potential. The objective is to see whether or not we can capture the behavior of the concentrated protein solutions using a simplified model. The motivation for using the simplified model was that it would be easier to make the connection to the properties like viscosities, etc. As this is an analytical model, the properties can be calculated quickly compared to a coarse-grained model.

4.2 Materials and Methods

Four monoclonal antibodies belong to IgG subclass type-1 (PPI02, PPI03, PPI08 and PPI10) were provided by AstraZeneca (AZ), Grant Park, UK, details of the proteins are mentioned below. Sequence alignment has been done for PPI02, PPI03 and PPI10 by using Clustal Omega program[33] and shown in section 3.6.1 in supplementary materials of chapter 3. PPI02, PPI03 and PPI08 have more than 80% sequence similarity in both heavy and light chain. PPI08 is a bispecific antibody and the sequence is not provided by the AZ. Additionally, PPI08 have an additional single chain Fv fragment of molecular weight ~50 kDa.

Table 4.1 Protein overview.

Code	Type	M_W (kDa)	Extinction coefficient[34], ε (L g^{-1} cm^{-1})	pI	Sequence	Notes
PPI02	Human IgG1κ	148.2	1.47	9.275	available	
PPI03	Human IgG1κ	144.8	1.435	9.410	available	Wild-type IgG
PPI08	IgG1κ + scFv	204.4	1.57	9.165	not available	Bispecific
PPI10	Human IgG1	144.2	1.533	9.160	available	

4.2.1 Sample preparation

Slide-A-Lyzer™ dialysis cassettes (Thermo Fisher Scientific®) were used for dialysis of each protein in 10 mM histidine buffer at pH 6.0. The buffer of the proteins was exchanged three times by extensive dialysis and final samples were prepared as previously described[35]. Each dialysis was continued overnight at a cold temperature (4 °C) with gentle stirring. The individual samples were prepared by diluting or concentrating (Vivaspin® centrifugal concentrators with polyethersulfone membrane, 50 kDa cutoff) the dialyzed sample. The protein concentration after dialysis was measured using NanoDrop™ 8000 Spectrophotometer using the respective extinction coefficients (see Table 4.1) at 280 nm. Samples with varying salt concentrations were prepared from buffer stock solutions with a 1M NaCl content by adding an appropriate amount of buffer to the individual samples. Importantly, buffers for buffer subtraction were also prepared to match the difference in NaCl content.

4.2.2 Small angle X-ray scattering

SAXS experiments were performed at the ESRF synchrotron, BM29 bioSAXS[36] beamline at Grenoble, France. It is used to measure the scattering intensities, *I(q)*, over an accessible range of q from 0.005−0.4 Å^{-1}. Measurements on pure water were used to get the data on an absolute scale. In order to compare the scattering profiles of the measurements, the data has to be normalized by the protein concentration of each sample, where the intensity profiles are scaled against each other for 0.15 Å^{-1} < q < 0.3 Å^{-1}. Buffers were measured both before and after each sample and averaged before subtraction from the sample's scattering profile. Data collection parameters for BM29 are listed in Table S4.1 in supplementary materials. Measurements were performed on a series of samples at various concentrations from 1 g/L to 130 g/L in different formulation conditions and an overview of the samples measured by SAXS are provided in Table S4.2, supplementary information. Samples above 30 g/L were injected manually in the flow cell for measurement due to high viscosity, where the lower concentrated samples were injected automatically by the sample changer.

The ATSAS program package version 2.8.4[37] was used for data analysis. Evaluation of the Guinier region was performed within Primus[38]. The atomic pair distribution function, $p(r)$, was determined by using the program GNOM[39]. The effective structure factors, $S_{eff}(q)$, were determined by dividing the normalizing scattering intensities by the form factor, $P(q)$, for the corresponding protein solution. To reduce the noise at low q values, we used GNOM program and fitted the form factor. The purpose to derive the effective structure factors is to get the information about the solution structure and inter

particle interactions. The effective structure factor for q values less than 0.06 $Å^{-1}$ is equal to the true structure factor $S(q)$[40].

4.2.3 Static light scattering

A Wyatt DynaPro NanoStar cuvette-based system was used for the static light scattering experiments. The molecular weight and second virial coefficient of the samples were measured using this instrument. It is a cuvette based system with low sample consumption. It uses a 60 mW GaAs diode laser at a wavelength of 658 nm with vertically polarized light. Samples should be filtered before measuring to eliminate dust particles. Dynamic Software (Wyatt Technology Corporation, Santa Barbara, CA) was used to acquire and process the SLS data. In this study, SLS measurements have also been used as an alternative approach for measuring $S(q)$ (theory for obtaining the B_{22} and $S(0)$ is described in chapter 2 in section 2.4).

4.2.4 Protein-protein interaction models

Here, we have used two simplified theoretical hard sphere model is used to capture the behavior of the concentrated protein solutions by fitting the structure factor profiles obtained from SAXS. The first model called as Yukawa model in this chapter, accounts for repulsion using Yukawa interaction potential, where we have used two fitting parameters, σ (hard sphere radius) and Z (net charge). Moreover, the second model called as Baxter model in this chapter, accounts for both attraction and repulsion. It accounts for repulsion using Yukawa potential and short-range attraction using Baxter adhesive potential. In this model, there were three fitting parameters, σ and Z and τ (strength of the adhesive force between the particles). Theory behind the models and interaction potentials are discussed in chapter 2 section 2.6.

4.3 Results and Discussion

4.3.1 Solution condition chosen to characterize protein – protein interactions and conformation

The protein-protein interaction for PPI02, PPI03, PPI08 and PPI10 used in this study have been screened extensively as a function of pH and ionic strength and for a series of different buffers and salts[41]. In chapter 3, nanoDSF and ICD studies were shown for these four mAbs for a wide range of pH and ionic strength, where we observed that all the mAbs are stable around the mid pH range (pH 6.0 to pH 7.0). We have chosen a well-behaved condition from the large screening to study the protein-protein interactions in solutions at an ionic strength of 10mM with histidine buffer at pH 6.0 at high concentrations. The effect of NaCl (0 mM, 35 mM, 70 mM and 140 mM) have also been studied for PPI03 at high protein concentrations. Due to time restriction and availability of the proteins, measurements has not been done in the presence of NaCl for PPI02, PPI08 and PPI10. Measurements cover protein concentrations ranging from 1 g/L to a maximum of approximately 150 g/L for mAbs.

4.3.2 Interaction and conformation of mAbs with changes in protein concentration

Monoclonal antibodies are y-shaped molecule, containing to Fab domains and one Fc domain with a hinge region, which allows for the movement. In SAXS, the formfactor reflects an average

conformations of various spatial arrangements of the Fc and Fab domains. Scaled SAXS intensities as a function of q of all the measured conditions for all four mAbs are shown in Figure 4.2. For an IgG1 mAb, the relative spatial rearrangements of the fragments about the hinge region are expected to cause changes to the SAXS profile in the q range 0.02 $Å^{-1} < q < 0.2$ $Å^{-1}$ [20]. Furthermore, for $q > 0.1$ $Å^{-1}$, the effective structure factor is close to 1. The changes in the SAXS profile in the region 0.1 $Å^{-1} < q < 0.2$ $Å^{-1}$ correspond to changes in the form factor and can be used to investigate the effects of conformational changes. In Figure 4.2, the intensity curves overlay with each other for q-values greater than 0.1 $Å^{-1}$, which indicates no conformational changes are occurring at high protein concentration. There are observable deviation intensities with increasing protein concentration for q values less than 0.1 $Å^{-1}$, which reflects repulsive protein-protein interaction for all the mAbs (see Figure 4.2). The scattering profiles for PPI03 in presence of NaCl in Figure 4.2 (A, B, C, D) shows less repulsion compared to in the absence of NaCl.

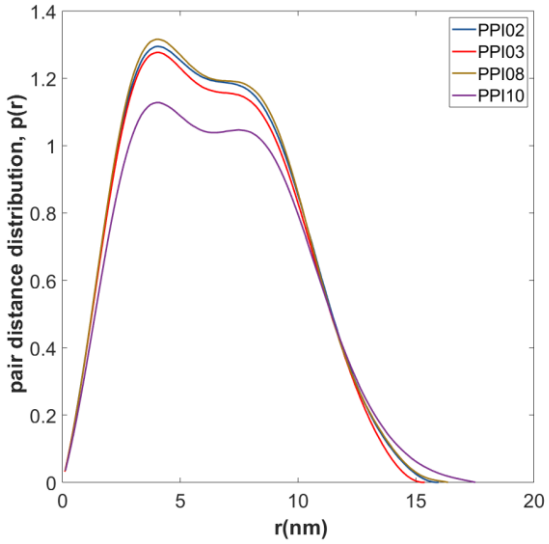


Figure 4.1 Pair distribution function obtained from the SAXS profiles for all the four mAbs. Color code: PPI02 (blue), PPI03 (red), PPI08 (dark yellow), PPI10 (violet).

In Figure 4.1, we show a plot of pair distribution function, *P(r)*, obtained from the X-ray scattering data at the concentration of around 1 g/L of all four mAbs. Each of the *P(r)* curves contains two peaks, which has been observed previously for antibodies across different subtypes including IgG1, IgG2, and IgG4[42–44]. The peak locations are similar to those observed in SAXS studies of other IgG1 antibodies.

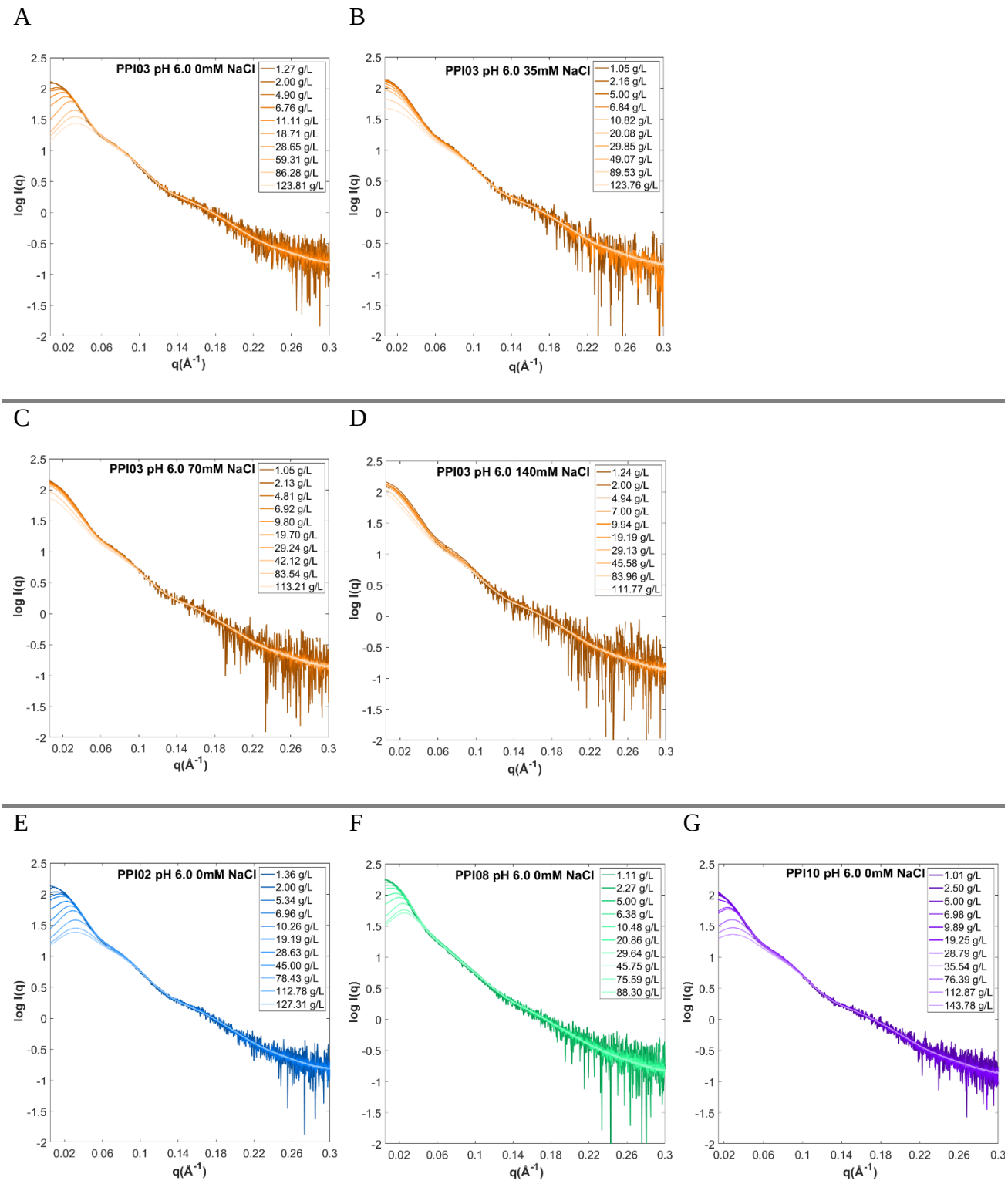


Figure 4.2 SAXS scattering curves. (A): PPI03 in 10 mM histidine pH 6.0, (B): PPI03 in 10 mM histidine pH 6.0, 35 mM NaCl, (C): PPI03 in 10 mM histidine pH 6.0, 70 mM NaCl, (D): PPI03 in 10 mM histidine pH 6.0, 140 mM NaCl, (E): PPI02 in 10 mM histidine pH 6.0, (F): PPI08 in 10 mM histidine pH 6.0, (G): PPI10 in 10 mM histidine pH 6.0. Data are shown for different concentrations where increasing concentration is indicated by decreasing color intensity.

4.3.3 Experimental structure factors and second virial coefficients

The SAXS data used to derive the experimental structure factors in all the measured conditions for all four mAbs and the SLS data were used to determine the osmotic second virial coefficient, B_{22}. SLS data were also used to calculate the structure factor at $q = 0$, as an alternative approach. SLS and SAXS provide complementary techniques for estimating *S(0)* as both approaches rely on assumptions about the angle dependence of the scattered light. For the SAXS data, fitting is required to extrapolate the *S(q)* profile to the long wavelength limit from data obtained for q values greater than 0.01 Å^{-1}, whereas the SLS experiment assumes no q-dependence of scattered light for q less than approximately 0.001 Å^{-1}. As a check of the assumptions, Figure 4.3(A) shows a comparison of the measured values of *S(0)* by SLS to those obtained from the fits to the PPI03 SAXS data in 10 mM histidine at pH 6.0. The two methods are shown to give results that are in very good agreement.

A B

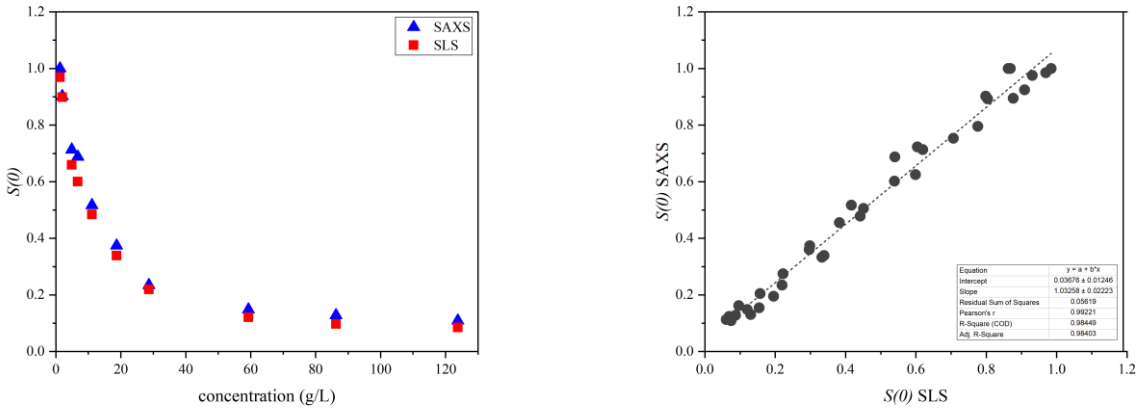


Figure 4.3 Extrapolated values of *S(0)* from SAXS and calculated values of *S(0)* from SLS as a function of concentration. A: *S(0)* values of PPI03 at 10 mM histidine pH 6.0; B: Linear fit of *S(0)s* obtained from SAXS and SLS.

A decrease in structure factor at low q region with increasing protein concentration reflects the effects of repulsive protein-protein interactions. In Figure 4.3 (B), data for all the measured mAbs are shown. We have plotted $S(0)$ obtained by SAXS as a function of $S(0)$ obtained by SLS. A very good correspondence is seen.

The reciprocal of *S(0)* obtained in 10 mM histidine pH 6.0 for the different mAbs are plotted in Figure 4.4 to get a better overview of the protein behavior. All the mAbs shows repulsive behavior in 10mM histidine at pH 6 as $S(0) < 1$. From the calculated SAXS parameters (see Table S4.3), R_G and D_{max} is seen to be consistent over the concentration range, while M_W is decreasing with increasing concentration, showing that all the measured mAbs are monodisperse and repulsive in the formulation condition.

A B

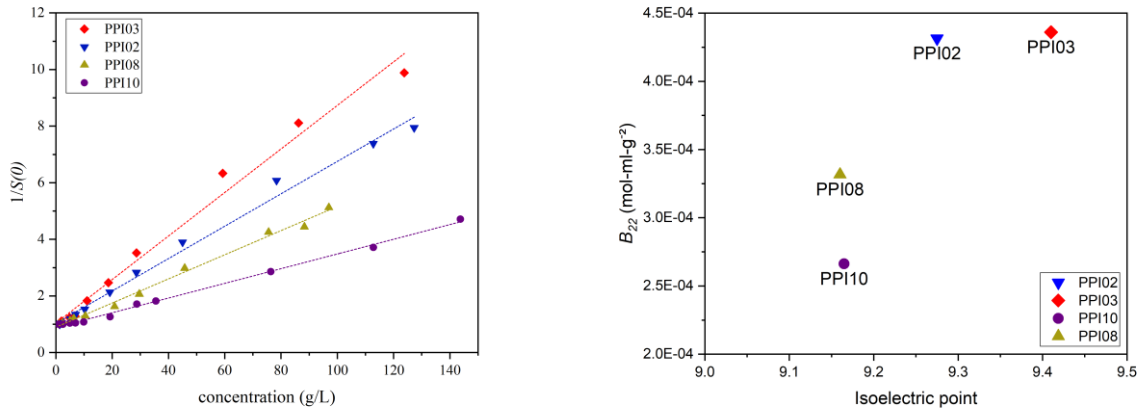


Figure 4.4 (A): 1/*S(0)* values from SAXS experiments and (B): Osmotic second virial coefficients, B_{22}, measured using SLS at 10 mM histidine, 0mM NaCl at pH 6.0. for all the four mAbs. Color code: PPI02 (blue), PPI03 (red), PPI08 (dark yellow) and PPI10 (violet).

PPI03 is the most repulsive system and PPI10 is the least repulsive system compared to other mAbs as the osmotic second virial coefficients, B_{22}, of PPI03 is higher and lower for PPI10 (see Figure 4.4). To see if this is connected with the iso-electric point we see in Figure 4.4(B) that the more far pI is from the measured solution pH, the more repulsive the protein behavior. The colloidal stability order of the mAbs is PPI03 > PPI02 > PPI08 > PPI10.

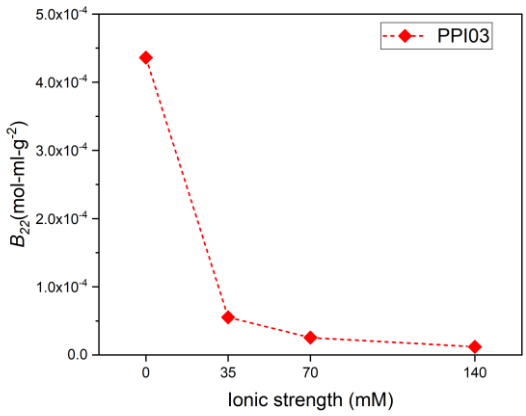


Figure 4.5 B_{22} values plotted vs ionic strength PPI03 in histidine pH 6.0.

To study the effect of the ionic strength on the osmotic second virial coefficient, B_{22}, SLS measurements have been performed over a range of ionic strengths (0 mM, 35 mM, 70 mM and 140 mM NaCl) in 10 mM histidine buffer at pH 6.0.

All the measured B_{22} values for PPI03 are positive (see Figure 4.5), reflecting that PPI03 is repulsive in all the tested conditions. Moreover, with increasing ionic strength up to 140 mM, B_{22} is decreasing, indicating that the ionic strength screens the electrostatic repulsion. Measured molecular weight for PPI03 in all the tested conditions was on the order of 146.0 ± 5 kDa (see Table S4.4), indicating monodisperse samples.

4.3.4 Protein-protein interaction model of the SAXS data

SAXS derived structure factors profiles for four mAbs has been fitted with the two simplified theoretical models, Yukawa and Baxter model as a function of protein concentration. In Table 4.2, the fitting parameters from Yukawa and Baxter models are shown for structure factors in 10 mM histidine at pH 6.0 solution condition. The ionic strength of the solution is 0.005 M. For fitting the structure factors, we have included the uncertainties into the fit parameters. The fitting with repulsive Yukawa model were done using two fitting parameters, i.e., σ (hard sphere radius) and Z (net charge). The model do not work well for all the antibodies probably due to the estimated uncertainties of the fit parameters were orders of magnitude larger than the parameters themselves.

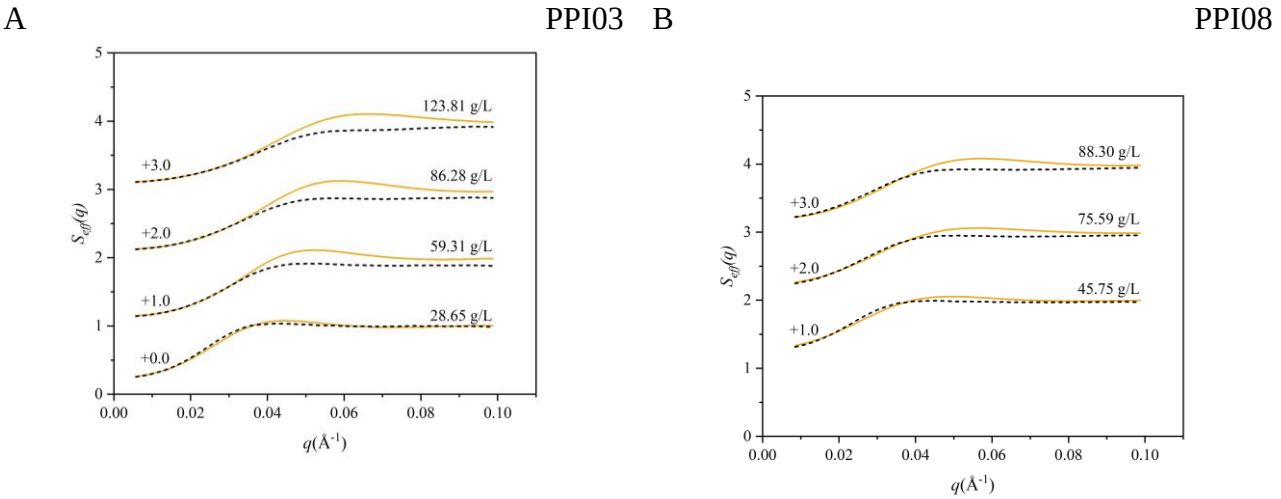


Figure 4.6 Selected structure factor profiles $S_{eff}(q)$ obtained from SAXS measurements all four mAbs in 10 mM histidine at pH 6.0, ~28 to ~145 g/L protein concentration. The yellow lines correspond to the hard sphere fitting with repulsive yukawa potential for $q < 0.03$ Å^{-1} and the black dashed lines corresponds to experimental structure factor profiles.

Using Yukawa model, the fitting of $S(q)$ profiles for PPI03 and PPI08 were possible and showed reasonable fit, while for PPI02 and PPI10 the fitting did not converge at any ionic strength. For PPI03, the fitting only converged at an ionic strength of 0.002 M, so this ionic strength was used for both PPI03 and PPI08. In Figure 4.6, shown is the measured structure factors profiles for PPI03 and PPI08 as a function of protein concentration in 10 mM histidine pH 6.0, with a minimum protein concentration of ~30 g/L. Here, the yellow lines correspond to the theoretical model and black dashed line to experimental structure factors. These selected structure factors are fitted with repulsive Yukawa potential for q-values, $q < 0.03$ Å^{-1}. Small q ranges for structure factors have been fitted to avoid any artifacts from orientational correlations.

Table 4.2 Parameters obtained from fitting to $S_{eff}(q)$ of all four mAbs in 10 mM histidine, 0 mM NaCl at pH 6.0, Representation: [a]Yukawa and [b]Baxter model.

Protein	Conc. (g/L)	σ (Å) [a]	Z (e) [a]	χ^2 [a]	σ (Å) [b]	Z (e) [b]	tau [b]	χ^2 [b]	S(0) [b]
PPI02	127.31	-	-	-	100.0 ± 1.9	12.4 ± 0.16	0.248	1.13	0.110
	112.78	-	-	-	97.0 ± 1.5	13.2 ± 0.08	0.194	1.22	0.118
	78.43	-	-	-	96.5 ± 1.2	14.2 ± 0.05	0.161	1.77	0.143
	45.00	-	-	-	99.7 ± 1.8	15.8 ± 0.05	0.178	3.60	0.203
	28.63	-	-	-	91.9 ± 4.0	14.9 ± 0.12	0.160	12.3	0.271
PPI03	123.81	84.7 ± 0.5	5.8 ± 0.10	28.3	111.0 ± 0.7	15.0 ± 0.17	0.115	1.01	0.105
	86.28	90.7 ± 0.7	6.5 ± 0.13	57.3	116.8 ± 0.6	15.9 ± 0.12	0.178	0.77	0.126
	59.31	65.5 ± 3.2	8.5 ± 0.09	141	123.1 ± 0.7	17.6 ± 0.10	0.232	1.01	0.145
	28.65	87.3 ± 5.6	9.0 ± 0.11	295	132.0 ± 0.6	18.6 ± 0.04	0.389	0.67	0.232
PPI08	88.30	101.2 ± 0.2	5.5 ± 0.08	8.4	102.4 ± 0.7	13.1 ± 0.07	0.193	0.74	0.190
	75.59	94.6 ± 0.7	6.5 ± 0.14	58.9	105.0 ± 0.7	13.8 ± 0.07	0.183	1.08	0.208
	45.75	90.5 ± 4.3	7.6 ± 0.25	184	107.5 ± 0.9	15.8 ± 0.05	0.193	1.49	0.263
	29.64	-	-	-	102.3 ± 1.2	14.6 ± 0.05	0.192	2.29	0.412
PPI10	143.78	-	-	-	91.7 ± 2.3	10.9 ± 0.09	0.202	3.88	0.130
	112.87	-	-	-	91.5 ± 1.7	11.8 ± 0.08	0.168	3.92	0.154
	76.39	-	-	-	88.7 ± 1.9	12.2 ± 0.10	0.131	13.9	0.195
	35.54	-	-	-	105.1 ± 1.8	15.9 ± 0.08	0.176	8.80	0.338

The fitting provides a bit lower net charge for PPI03 than expected which was studied by Roberts et al[24], but the charge does not change much with increasing protein concentration. It could be lower to compensate for fitting with a lower ionic strength than the experimental value. In addition the size is slightly lower than the excluded volume diameter equal to 10.4 nm (the excluded volume diameter is equal to the hydrodynamic diameter, it corresponds to the diameter of a sphere with the same excluded volume as the protein), but in the study by Hung et al[45], they found similar sizes, hydrodynamic radius of 4.5 nm from fitting to *S(q)* curves as a function of protein concentration. The diameters remain relatively constant except at 59.31 g/L.

PPI08 fitting parameters are pretty similar to PPI03, except the size of the hard sphere is a bit larger, but similar charges are obtained. It is interesting because PPI08 has a lower B_{22} value compared to PPI03, but this is not reflected in the fitting parameters, which would require either a smaller size or lower charges.

In the second approach by using the Baxter model, which have an additional short ranged attractive term work well for all the antibodies. Here we consider three fitting parameters, σ and Z and τ with uncertainties into the fit parameters. A comparison between the experimental and fitted S(q) profiles are shown in Figure 4.7, data have been fitted for low q-values, $q < 0.025$ Å^{-1}. All the measured S(q) profiles have a good fit with the Baxter model for low q-values. From the fitting data, S(0) values can be extracted by extrapolating the S(q) profiles to zero and the contribution of the attraction. An ionic strength 0.01 M was used as the fitting only converged at an ionic strength of 0.01 M. Further, this ionic strength was used for fitting of *S(q)* profiles of all four mAbs. As there are uncertainties in the

screening length for describing the electrostatic interactions arising from the assumption that the antibody is a sphere interacting through centro-symmetric potentials, this might be the reason for a higher ionic strength provides slightly better fits, because it is reflecting the effects of anisotropic interactions or shape.

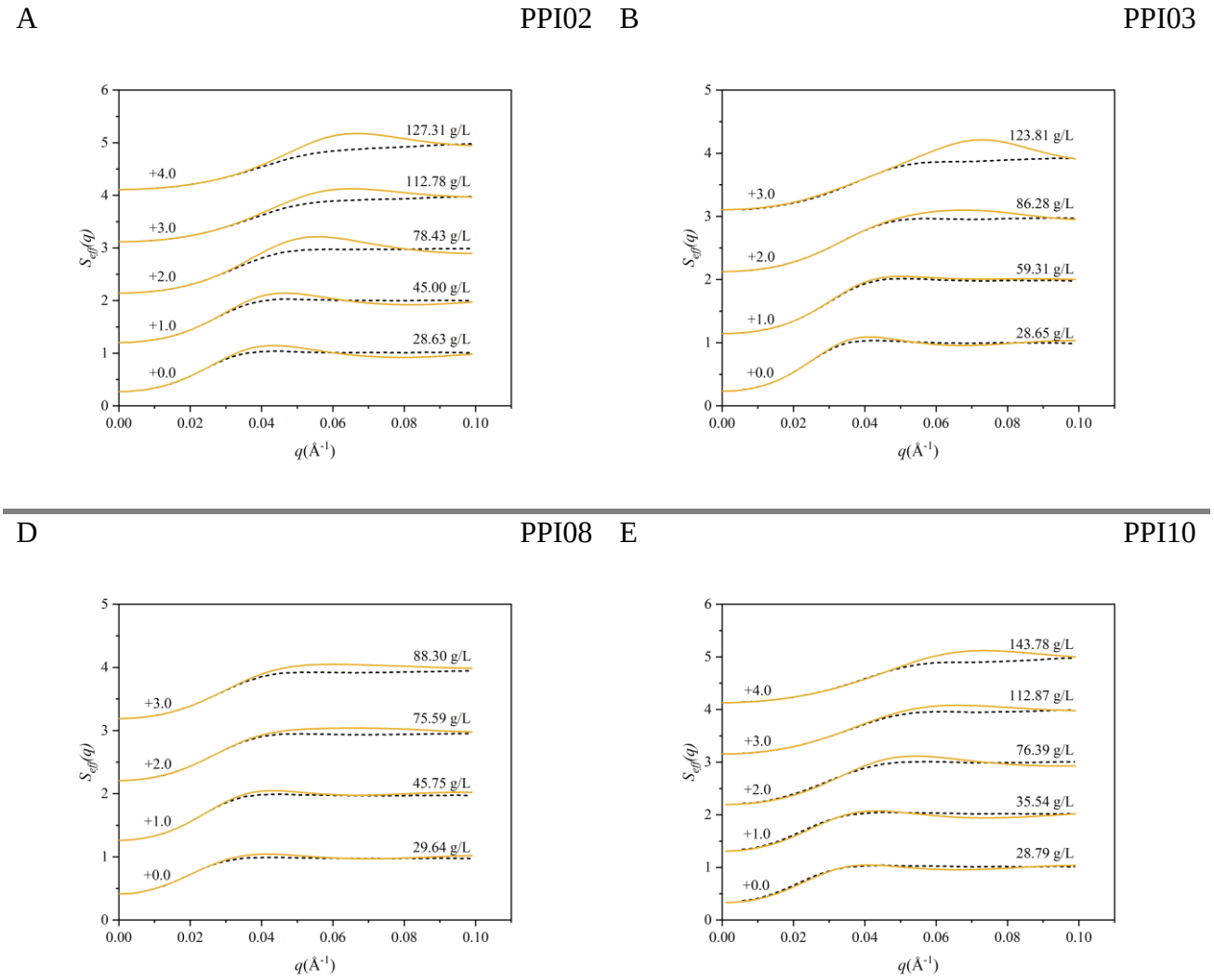


Figure 4.7 Selected structure factor profiles $S_{eff}(q)$ obtained from SAXS measurements all four mAbs in 10 mM histidine at pH 6.0, ~28 to ~145 g/L protein concentration. The yellow lines correspond to the fitting with Baxter model for $q < 0.025$ Å^{-1} and the black dashed lines corresponds to experimental structure factor profiles.

The results of the fitting are shown in Table 4.2 along with the reduced χ^2 values. By comparing the χ^2 values from Baxter fitting to yukawa fitting, we get much lower χ^2 values and better fits from Baxter fitting. This is expected since there should be a missing attraction in the Baxter model, even at low ionic strength. For PPI03, the charges are relatively constant and agree well with the expected charges from the Roberts et al[24], where the charges were obtained from fitting the DLVO potential to B_{22} data. The obtained charges from Baxter model are greater than the values obtained from fitting with the yukawa potential due to the difference in ionic strengths used in the fitting. Both the σ values and τ values follow a systematic trend with changing protein concentration. With increasing protein concentration, σ and τ tau values decrease which indicates that there is less attraction in the model. As

the values for both parameters changing quite a lot, so the behavior with respect to protein concentration is difficult to explain.

For each of the other proteins, PPI02, PPI08 and PPI10, σ remains relatively constant with protein concentration, which is encouraging, and follow the order PPI03 > PPI08 > PPI02 > PPI10. Moreover, the net charges are also consistent with the isoelectric points of the mAbs which indicate PPI10 and PPI08 might have slightly lower net charges than PPI03. For PPI08, the value of τ remains constant, while there is slight variation in the values for PPI02 and PPI10. This indicates that the model could be used to fit the whole *S(q)* profiles and capture the behavior.

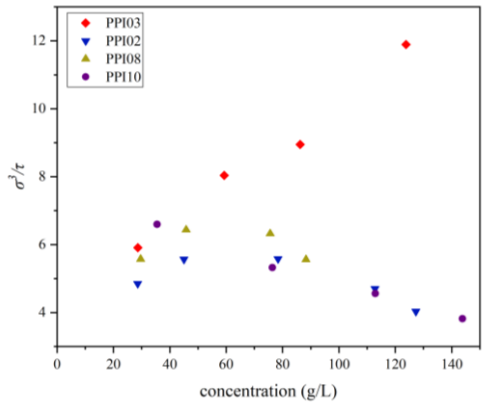


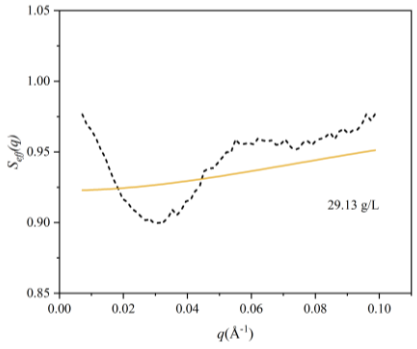
Figure 4.8 Relative contribution of short range attraction shown by plotting σ^3/τ vs. protein concentration.

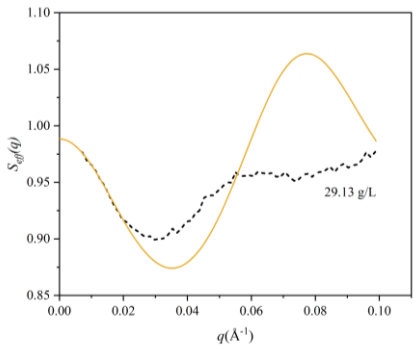
The contribution of the short ranged attraction from the Baxter fitting results was given by $B_{22}^{hs}/4\tau$. In Figure 4.8, the relative contribution of short range attraction for the four mAbs was compared. It is the same as a function of protein concentration for PPI02, PPI08, PPI10 except PPI03, so it seems like they have all the same approximate attraction. In addition, PPI03 which should have the least attraction, does have the least attraction, but the behavior is opposite to what happens with the other mAbs.

For PPI03, fitting has also been done for 35 mM, 70 mM and 140 mM NaCl and the fitting parameters are shown in Table 4.3. In Figure 4.9, fitting of both Yukawa and Baxter model are shown for PPI03 in 140 mM NaCl. We see that the Yukawa model does not fit properly and gives high χ^2 values. However, Baxter model fits well with the all structure factor profiles, as there is upward movement of the curve at low q indicates presence of attraction (see Figure 4.9). The upturn in the structure factor appears due to the presence of weak protein-protein associaltion[20,46]. In Figure 4.10, the fittings for PPI03 in presence of 35 mM , 70 mM and 140 mM NaCl are shown using Baxter model, where the model has good fit for all the measured structure factors in presence of NaCl. .

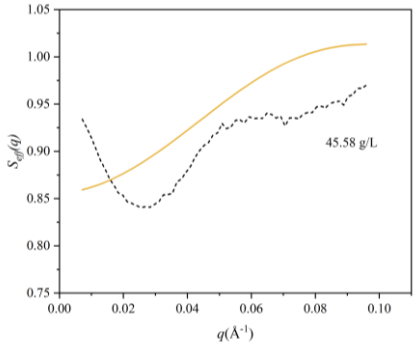
Fitting using Yukawa model **Fitting using Baxter model**

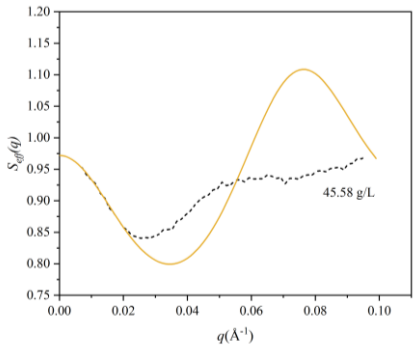
Protein concentration: 29.31 g/L



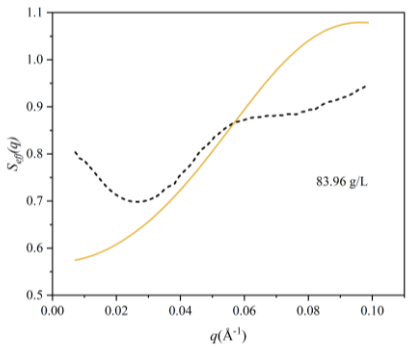


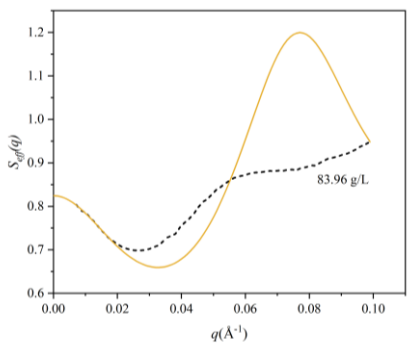
Protein concentration: 45.58 g/L



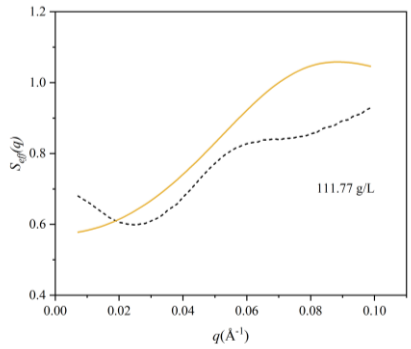


Protein concentration: 83.96 g/L





Protein concentration: 111.77 g/L



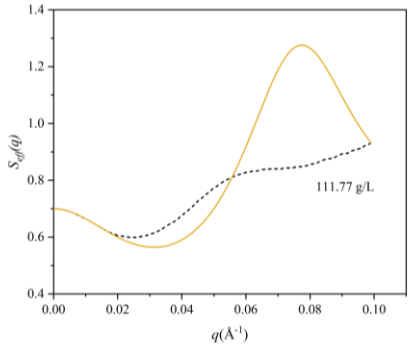


Figure 4.9 Structure factor profiles $S_{eff}(q)$ obtained from SAXS measurements for PPI03 in 10 mM histidine in 140 mM NaCl in different protein concentration. Left panel shows the Yukawa fitting and right panel shows the Baxter fitting. The yellow lines correspond to the fitting model and the black dashed lines corresponds to experimental structure factor profiles. Here we could see perfectly that there is attraction due to upward curvature of the structure factor profiles

For ionic strength 75 and 145 mM, both the σ values and τ is remaining constant with increasing protein concentration, which indicates that we can use these parameters to extrapolate across protein concentration. In addition, the *S(0)* is slightly higher for higher for 145 mM ionic strength which indicates that with increasing ionic strength from 75 to 145 mM, it is screening the electrostatics.

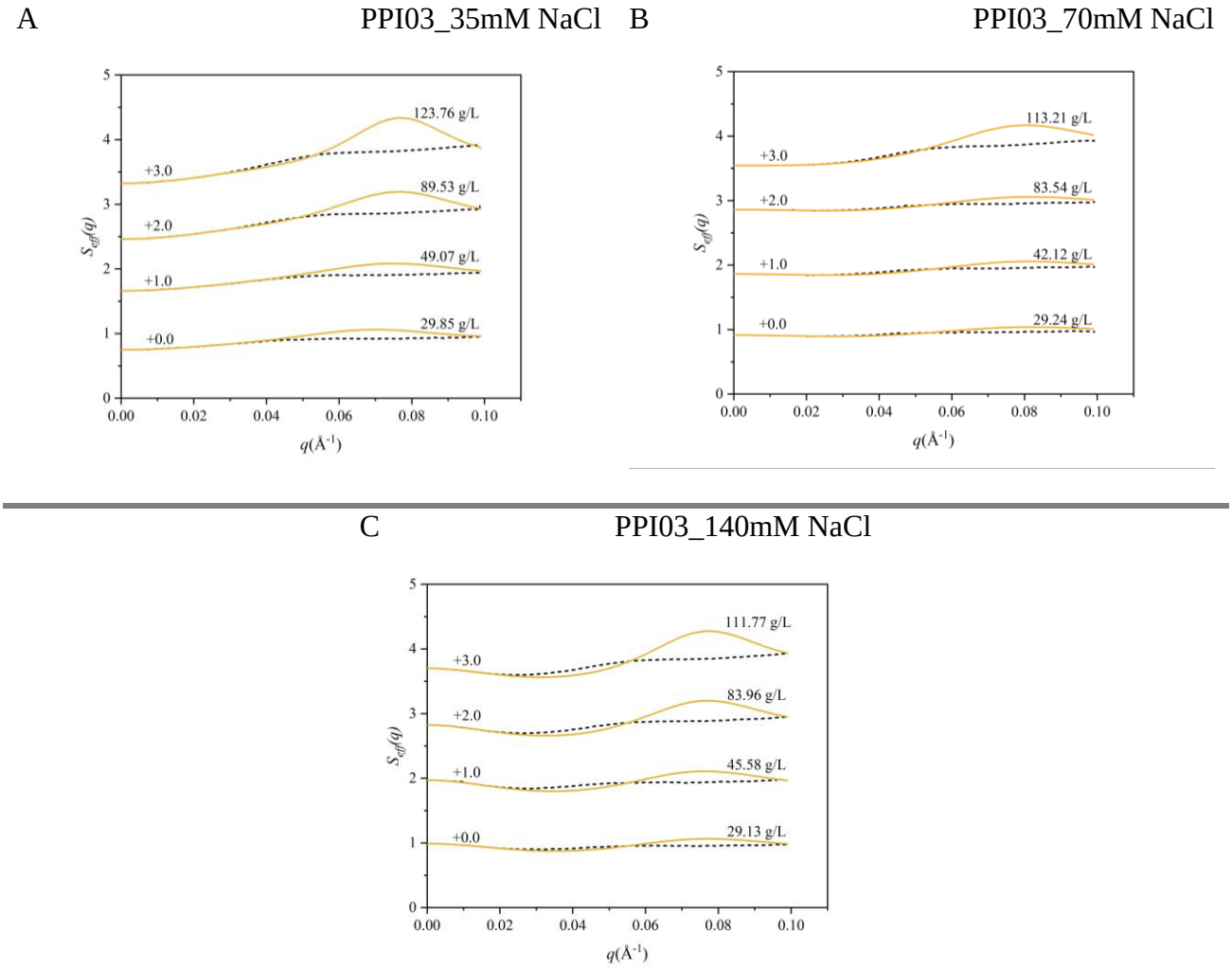


Figure 4.10 Structure factor profiles $S_{eff}(q)$ of PPI03 in presence of NaCl obtained from SAXS measurements and fitted with Baxter model for $q < 0.025$ Å^{-1}. The yellow lines correspond to the fitting and the black dashed lines corresponds to experimental structure factor profiles.

Table 4.3 Parameters obtained from fitting to $S_{eff}(q)$ for PPI03 in presence of NaCl, Representation: [a]Yukawa and [b]Baxter model.

Protein	Ionic strength (M)	Conc (g/L)	σ (Å) [a]	Z (e) [a]	χ^2 [a]	σ (Å) [b]	Z (e) [b]	tau [b]	χ^2 [b]	S(0) [b]
PPI03	0.040	123.76	-	-	-	99.70	19.43	0.15	9.74	0.32
	0.040	89.53	78.23	0.39	132.35	98.33	15.83	0.21	8.62	0.46
	0.040	49.07	42.97	12.04	5.57	97.42	12.25	0.29	2.32	0.66
	0.040	29.85	82.87	0.02	27.80	104.21	13.21	0.29	2.26	0.75
PPI03	0.075	113.21	-	-	-	88.12	5.44E-04	0.39	10.98	0.54
	0.075	83.54	-	-	-	88.74	7.11E-04	0.35	13.25	0.69
	0.075	42.12	65.58	0.48	174.88	87.51	7.53E-05	0.34	7.32	0.86
	0.075	29.24	80.11	9.18	30.45	87.03	5.54E-04	0.33	3.09	0.91
PPI03	0.145	111.77	-	-	-	93.00	5.92E-03	0.26	4.35	0.69
	0.145	83.96	-	-	-	93.41	3.39E-04	0.25	4.73	0.81
	0.145	45.58	-	-	-	93.90	1.05E-02	0.24	1.38	0.97
	0.145	29.13	74.68	10.94	65.93	92.68	4.01E-03	0.25	0.43	0.98

In this study, Baxter model fits well compared to Yukawa due to additional attractive term for different solution with large range of ionic strengths. Fitting can be improved if we have less fitting parameters. For example: if we keep the sigma fixed for both the models and redoing the fitting, then we could improve the fitting. Model can be improvised by restraining the inter dependable parameters. Also as we only used hard sphere model and antibody is not really spherical, changing the shape of the model (ellipsoid) could be another option to improvise the fitting and getting a better understanding of different interaction parameters contributions.

4.4 Conclusion

In this study, four IgG1 mAbs have been studied in concentrated solutions. SAXS data shows all the mAbs are monodisperse and there were no conformational changes with increasing concentration. Here, SAXS and SLS techniques complement each other very well. Experimentally derived structure factors for mAbs contain features that give the information about the protein density distribution and inter-particle interactions. SAXS and SLS study shows all the mAbs are repulsive in nature in the measured solution conditions. From SLS, we fit a scalar number and only say if the net interaction is repulsive or attractive. However, with SAXS data curves $S(q)$ and modelling here of, we are able to resolve attractive and repulsive forces. The spherical model study provides a more accurate extrapolation of the structure factors. Model that only accounts for repulsive interaction do not provided a good fit for all the mAbs ,while model that accounts for both repulsion and attraction shown much better fit to the experimentally derived structure factors. Even we see the net interaction is repulsive form SAXS and SLS measurements, there was evidence of presence of some short ranged protein-protein attraction in the system. From SLS, we can only say if the net interaction is repulsive or attractive, but with SAXS data and modelling, we could get a more clear idea and say about the relative contribution of repulsion and attraction.

Supplementary Materials

Table S4.1 Experimental set up of SAXS measurements

Instrument	**BM29, ESRF**
Wavelength (Å)	0.99
q-range ($Å^{-1}$)	0.004 – 0.49
Sample-to-detector distance ($\boldsymbol{m}$)	2.864
Detector	PILATUS 1M
Exposure time (s)	10 x 1.00
Beam size	700 x 700 μm^2
Sample configuration	1.8 mm quartz glass capillary
Absolute scaling method	Comparison to water in sample capillary
Normalization	To transmitted intensity by beam-stop counter
Monitoring for radiation damage	Control of un-subtracted and scaled subtracted data for systematic changes typical for radiation damage

Table S4.2 Samples for SAXS experiments performed at varying protein concentrations.

Protein/Code	**Buffer**	**Additive**	**Protein concentration (g/L)**
PPI02 (Human IgG1κ)	10 mM histidine pH 6.0	-	1.36 – 127.31
PPI03 (Human IgG1κ)	10 mM histidine pH 6.0	-	1.27 – 123.81
		35 mM NaCl	1.05 – 123.76
		70 mM NaCl	1.05 – 113.21
		140 mM NaCl	1.24 – 111.77
		250 mM NaCl	0.98 – 129.42
PPI08 (IgG1κ + scFv)	10 mM histidine pH 6.0	-	1.11 – 88.30
PPI10 (Human IgG1)	10 mM histidine pH 6.0	-	1.01 – 143.78

Table S4.3 Overview over all calculated SAXS parameters for samples 1-11 g/L. All other samples show large deviations from ideality due to repulsion.

Protein Code	**Formulation condition**	**Protein (g/L)**	**NaCl (mM)**	**Guinier**		***P(r)***			**Apparent M_W (kDa)**	
				I(0)/c	R_G (nm)	*I(0)/c*	R_G (nm)	D_{max} (nm)	Guinier	*P(r)*
PPI02	10 mM histidine pH 6.0	1.36	-	141.1	4.91	142.9	5.05	15.96	156	158
		2.00	-	142.9	4.79	143.7	5.00	15.40	158	159
		5.34	-	128.5	4.32	139.2	4.89	15.23	142	154
		6.96	-	124.7	4.27	134.1	4.77	14.63	138	149
		10.26	-	117.6	4.10	129.4	4.68	14.35	130	143
PPI03	10 mM histidine pH 6.0	1.27	-	136.2	4.83	141.7	5.00	15.37	151	157
		2.00	-	134.0	4.64	136.1	4.97	15.30	148	151
		4.90	-	117.2	4.09	133.1	4.80	14.70	130	147
		6.76	-	116.6	3.98	132.2	4.83	14.45	129	146
		11.11	-	114.7	3.70	131.8	4.71	14.40	127	146
		1.05	35	144.2	4.95	145.2	5.05	16.11	160	161
		2.16	35	143.7	4.95	144.7	5.04	16.36	159	160
		5.00	35	143.4	4.94	143.4	5.03	16.30	159	159
		6.84	35	142.2	4.91	142.7	5.01	16.31	158	158
		10.82	35	133.2	4.92	133.8	5.00	16.06	148	148
		1.05	70	139.3	5.08	138.9	5.11	15.37	154	154
		2.13	70	136.9	5.03	137.5	5.12	14.98	152	152
		4.81	70	135.8	5.07	135.2	5.07	14.99	150	150
		6.92	70	133.5	5.02	134.2	5.06	14.87	148	149
		9.80	70	133.9	5.00	134.5	5.04	14.50	148	149
		1.24	140	136.0	5.07	134.1	5.08	15.80	151	149
		2.00	140	133.9	5.06	133.3	5.09	15.01	148	148
		4.94	140	133.4	5.05	132.3	5.09	14.90	148	147
		7.00	140	132.2	5.05	131.3	5.09	14.71	146	145
		9.94	140	130.4	5.05	130.9	5.09	14.60	144	145
		0.98	250	123.2	5.11	124.1	5.26	19.66	136	137
		2.00	250	121.9	5.08	123.2	5.26	20.75	135	137
		5.15	250	121.1	5.13	121.9	5.30	20.08	134	135
		7.11	250	121.5	5.16	122.8	5.38	21.60	135	136
		9.94	250	122.6	5.31	123.1	5.44	22.11	136	136
PPI08	10 mM histidine pH 6.0	1.11	-	202.1	5.74	206.0	5.97	20.73	224	228
		2.27	-	201.3	5.57	207.3	5.87	18.89	223	230
		5.00	-	180.0	5.06	192.8	5.63	17.75	199	214
		6.38	-	172.9	4.88	190.6	5.62	17.43	192	211
		10.48	-	163.9	4.79	182.3	5.48	17.23	182	202
PPI10	10 mM histidine pH 6.0	1.01	-	133.6	5.16	134.0	5.23	16.52	148	148
		2.50	-	131.2	5.00	133.2	5.14	16.27	145	148
		5.00	-	122.7	4.74	127.7	5.01	15.42	136	141
		6.98	-	119.0	4.56	125.4	4.96	15.52	132	139
		9.89	-	114.0	4.39	123.8	4.92	15.00	126	137

Table S4.4 Molecular weight *(M_W)* and B_{22} values calculated from SLS measurements.

PPI03 10 mM histidine pH 6.0	B_{22} **x** 10^{-4} (mol-mL/g²)	M_W (kDa)
0 mM NaCl	4.36	143.2
35mM NaCl	0.553	141.4
70 mM NaCl	0.255	138.2
140 mM NaCl	0.121	151.3

References

1. Chames, P., Van Regenmortel, M., Weiss, E. & Baty, D. Therapeutic antibodies: Successes, limitations and hopes for the future. *British Journal of Pharmacology* (2009). doi:10.1111/j.1476-5381.2009.00190.x

2. Weiner, L. M., Surana, R. & Wang, S. Monoclonal antibodies: Versatile platforms for cancer immunotherapy. *Nature Reviews Immunology* (2010). doi:10.1038/nri2744

3. Sharkey, R. M. & Goldenberg, D. M. Use of antibodies and immunoconjugates for the therapy of more accessible cancers. *Advanced Drug Delivery Reviews* (2008). doi:10.1016/j.addr.2008.04.011

4. Kaplon, H. & Reichert, J. M. Antibodies to watch in 2019. *MAbs* (2019). doi:10.1080/19420862.2018.1556465

5. Shire, S. J., Shahrokh, Z. & Liu, J. Challenges in the development of high protein concentration formulations. *Journal of Pharmaceutical Sciences* (2004). doi:10.1002/jps.20079

6. Mason, B. D., Zhang, L., Remmele, R. L. & Zhang, J. Opalescence of an IgG2 monoclonal antibody solution as it relates to liquid-liquid phase separation. *J. Pharm. Sci.* (2011). doi:10.1002/jps.22650

7. Raut, A. S. & Kalonia, D. S. Pharmaceutical Perspective on Opalescence and Liquid-Liquid Phase Separation in Protein Solutions. *Mol. Pharm.* (2016). doi:10.1021/acs.molpharmaceut.5b00937

8. Salinas, B. A. *et al.* Understanding and modulating opalescence and viscosity in a monoclonal antibody formulation. *J. Pharm. Sci.* (2010). doi:10.1002/jps.21797

9. Saluja, A. & Kalonia, D. S. Nature and consequences of protein-protein interactions in high protein concentration solutions. *International Journal of Pharmaceutics* (2008). doi:10.1016/j.ijpharm.2008.03.041

10. Kayser, V. *et al.* Evaluation of a non-arrhenius model for therapeutic monoclonal antibody aggregation. *J. Pharm. Sci.* (2011). doi:10.1002/jps.22493

11. Manning, M. C., Chou, D. K., Murphy, B. M., Payne, R. W. & Katayama, D. S. Stability of protein pharmaceuticals: An update. *Pharm. Res.* **27**, 544–575 (2010).

12. Kamerzell, T. J., Esfandiary, R., Joshi, S. B., Middaugh, C. R. & Volkin, D. B. Protein-excipient interactions: Mechanisms and biophysical characterization applied to protein formulation development. *Advanced Drug Delivery Reviews* (2011). doi:10.1016/j.addr.2011.07.006

13. Ohtake, S., Kita, Y. & Arakawa, T. Interactions of formulation excipients with proteins in solution and in the dried state. *Advanced Drug Delivery Reviews* (2011). doi:10.1016/j.addr.2011.06.011

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). doi:10.1021/jp1028646

15. He, F. *et al.* High-throughput dynamic light scattering method for measuring viscosity of concentrated protein solutions. *Anal. Biochem.* **399**, 141–143 (2010).

16. Saluja, A., Badkar, A. V., Zeng, D. L., Nema, S. & Kalonia, D. S. Application of high-frequency rheology measurements for analyzing protein-protein interactions in high protein concentration solutions using a model monoclonal antibody (IgG2). *J. Pharm. Sci.* (2006). doi:10.1002/jps.20663

17. Burckbuchler, V. *et al.* Rheological and syringeability properties of highly concentrated human polyclonal immunoglobulin solutions. *Eur. J. Pharm. Biopharm.* (2010). doi:10.1016/j.ejpb.2010.08.002

18. Brandt, J. P., Patapoff, T. W. & Aragon, S. R. Construction, MD simulation, and hydrodynamic validation of an all-atom model of a monoclonal IgG antibody. *Biophys. J.* (2010). doi:10.1016/j.bpj.2010.05.003

19. Dear, B. J. *et al.* X-ray Scattering and Coarse-Grained Simulations for Clustering and Interactions of Monoclonal Antibodies at High Concentrations. *J. Phys. Chem. B* (2019). doi:10.1021/acs.jpcb.9b04478

20. Yearley, E. J. *et al.* Small-angle neutron scattering characterization of monoclonal antibody conformations and interactions at high concentrations. *Biophys. J.* (2013). doi:10.1016/j.bpj.2013.06.043

21. Corbett, D. *et al.* Coarse-Grained Modeling of Antibodies from Small-Angle Scattering Profiles. *J. Phys. Chem. B* **121**, 8276–8290 (2017).

22. Chari, R., Jerath, K., Badkar, A. V. & Kalonia, D. S. Long- and short-range electrostatic interactions affect the rheology of highly concentrated antibody solutions. *Pharm. Res.* (2009). doi:10.1007/s11095-009-9975-2

23. Ralston, G. B. Effects of ‘crowding’ in protein solutions. *J. Chem. Educ.* (1990). doi:10.1021/ed067p857

24. Roberts, D. *et al.* The role of electrostatics in protein-protein interactions of a monoclonal antibody. *Mol. Pharm.* (2014). doi:10.1021/mp5002334

25. Tardieu, A. Eye lens proteins and transparency: from light transmission theory to solution X-ray structural analysis. *Annual review of biophysics and biophysical chemistry* (1988). doi:10.1146/annurev.bb.17.060188.000403

26. Vérétout, F., Delaye, M. & Tardieu, A. Molecular basis of eye lens transparency. Osmotic

pressure and X-ray analysis of α-crystallin solutions. *J. Mol. Biol.* (1989). doi:10.1016/0022-2836(89)90316-1

27. Tardieu, A., Laporte, D. & Delaye, M. Colloidal dispersions of α-crystallin proteins. - I. Small angle X-ray analysis of the dispersion structure. *J. Phys.* (1987). doi:10.1051/jphys:019870048070120700

28. Sjöberg, B. & Mortensen, K. Interparticle interactions and structure in nonideal solutions of human serum albumin studied by small-angle neutron scattering and Monte Carlo simulation. *Biophys. Chem.* (1994). doi:10.1016/0301-4622(94)00089-1

29. Sjöberg, B. & Mortensen, K. Structure and thermodynamics of nonideal solutions of colloidal particles: Investigation of salt-free solutions of human serum albumin by using small-angle neutron scattering and Monte Carlo simulation. *Biophys. Chem.* (1997). doi:10.1016/S0301-4622(96)02251-X

30. Sønderby, P. *et al.* Self-Interaction of Human Serum Albumin: A Formulation Perspective. *ACS Omega* **3**, (2018).

31. Curtis, R. A., Ulrich, J., Montaser, A., Prausnitz, J. M. & Blanch, H. W. Protein-protein interactions in concentrated electrolyte solutions: Hofmeister-series effects. *Biotechnol. Bioeng.* (2002). doi:10.1002/bit.10342

32. Voges, M., Herhut, M., Held, C. & Brandenbusch, C. Light-scattering data of protein and polymer solutions: A new approach for model validation and parameter estimation. *Fluid Phase Equilib.* (2018). doi:10.1016/j.fluid.2018.02.022

33. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* (2011). doi:10.1038/msb.2011.75

34. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* (2003). doi:10.1093/nar/gkg563

35. Augustijn, D. *et al.* Novel Non-linear Curve Fitting to Resolve Protein Unfolding Transitions in Intrinsic Fluorescence Differential Scanning Fluorimetry. *Eur. J. Pharm. Biopharm.* (2019). doi:10.1016/J.EJPB.2019.06.001

36. Pernot, P. *et al.* New beamline dedicated to solution scattering from biological macromolecules at the ESRF. in *Journal of Physics: Conference Series* (2010). doi:10.1088/1742-6596/247/1/012009

37. Franke, D. *et al.* ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* (2017). doi:10.1107/S1600576717007786

38. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* (2003). doi:10.1107/S0021889803012779

39. Semenyuk, A. V. & Svergun, D. I. GNOM. A program package for small-angle scattering data processing. *J. Appl. Crystallogr.* (1991). doi:10.1107/S002188989100081X

40. Greene, D. G., Ferraro, D. V., Lenhoff, A. M. & Wagner, N. J. A critical examination of the decoupling approximation for small-angle scattering from hard ellipsoids of revolution. *J. Appl. Crystallogr.* (2016). doi:10.1107/S1600576716012929

41. Gentiluomo, L., Svilenov, H., Augustijn., D., El Bialy, I., Greco, M.L., Kulakova, A., Indrakumar, S., Mahapatra, S., Morales, M., Pohl, C., ..., Harris, P., Frieß, W. Boosting therapeutic protein development by publicly available datasets including comprehensive computational and biophysical characterization. *Manuscr. Submitt.* (2019).

42. Rayner, L. E. *et al.* The solution structures of two human IgG1 antibodies show conformational stability and accommodate their C1q and FcγR ligands. *J. Biol. Chem.* (2015). doi:10.1074/jbc.M114.631002

43. Abe, Y., Gor, J., Bracewell, D. G., Perkins, S. J. & Dalby, P. A. Masking of the Fc region in human IgG4 by constrained X-ray scattering modelling: Implications for antibody function and therapy. *Biochem. J.* (2010). doi:10.1042/BJ20100641

44. Mosbaek, C. R., Konarev, P. V., Svergun, D. I., Rischel, C. & Vestergaard, B. High concentration formulation studies of an IgG2 antibody using small angle x-ray scattering. *Pharm. Res.* (2012). doi:10.1007/s11095-012-0751-3

45. Hung, J. J. *et al.* Protein-Protein Interactions of Highly Concentrated Monoclonal Antibody Solutions via Static Light Scattering and Influence on the Viscosity. *J. Phys. Chem. B* (2019). doi:10.1021/acs.jpcb.8b09527

46. Abramo, M. C. *et al.* Effective interactions in lysozyme aqueous solutions: A small-angle neutron scattering and computer simulation study. *J. Chem. Phys.* (2012). doi:10.1063/1.3677186

5 Self-interactions and ion binding of two monoclonal antibodies: small-angle X-ray scattering, static light scattering and coarse-grained modeling

Monoclonal antibodies are one of the largest classes of biotherapeutics due to their high binding specificity. However, the formulation still remains a challenge, especially at high concentrations, to meet patient dosage requirements. Parameters like pH, ionic strength, excipients can be optimized to achieve a stable formulation. The solution properties are highly dependent on the colloidal and conformational stability of the proteins which is governed by intermolecular interactions.

In this study, we have selected two monoclonal antibodies (PPI03 and PPI13) form the PIPPI protein library and chosen a well behaving solution condition from the large screening of these mAbs. In the presence of 10 mM histidine at pH 6.0 as a function of NaCl concentration, self-interactions and ion binding studies were performed using SAXS, SLS and coarse-grained modeling at high protein concentration. SAXS provided information about the conformation and self-interaction presence in terms of structure factors, while SLS studies provided the information for net interaction in terms of osmotic second virial coefficient. SAXS based models have been used for coarse-grained modeling. Here, we construct a single protein model, where each amino acid is represented by a spherical bead. Using this model, we could visualize the ions accumulations on the surface, which could explain the types of interactions present in the system. Also, we have constructed a two-body model and performing monte carlo simulation to reproduce the osmotic second virial coefficient obtained from SLS using *in-silico* approach.

Manuscript in preparation

Self-interactions and ion binding of two monoclonal antibodies: small-angle X-ray scattering, static light scattering and coarse-grained modeling

Sujata Mahapatra[a,b], Marco Polimeni[c], Mikael Lund[b], Günther H.J. Peters[b], Pernille Harris[b], Werner W. Streicher[a]

[a]Novozymes A/S, Biologiens Vej 2, 2800 Kgs. Lyngby, Denmark
[b]Department of Chemistry, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
[c]Division of Theoretical Chemistry, Naturvetarvägen 14, 223 62 Lund, Sweden

Abstract

In the present study, small angle X-ray scattering small-angle X-ray scattering (SAXS) and static light scattering (SLS) and coarse-grained modeling have been used to study the self-interactions and Cl^- ion binding of two monoclonal antibodies, PPI03 and PPI13. SAXS concentration series measurements were performed at high protein concentration and as a function of NaCl concentration to study the behavior at high concentration, while SLS measurement was performed to obtain the osmotic second virial coefficient. The experiments are complemented with metropolis monte carlo simulations using one body model. In the one body model, each amino acids were considered as a bead. The SAXS derived rigid body models were used to carry out the monte carlo simulation in the presence of NaCl to examine the Cl^- binding on mAbs. We presented the ion density maps of the mAbs in the presence of NaCl that showed the Cl^- ions accumulate on particular regions or patches on mAbs surface, subsequently explain the type of interactions present.

5.1 Introduction

Monoclonal antibody (mAb) derived products have become a major contribution in biotherapeutics for their ability to bind specifically to target antigens[1,2]. The stability of mAbs is governed by intermolecular self-interactions, crucial to the formulation properties of antibody variants. The formulation remains challenging at high protein concentrations due to undesirable solution properties, such as high viscosity and/or opalescence[3]. Additionally, aggregation of mAbs could also lead to harmful immune reactions[4]. In concentrated protein solution, weak protein-protein interactions can be used to determine the phase behavior and to the behavior of the protein in a crowded environment[5,6].

Solution pH and ionic strength are two widely used parameters for optimizing the formulation stability. The impact of pH and ionic strength of mAbs are poorly understand in comparison to globular proteins[7–9].The solution properties of mAbs variants depend on the colloidal stabilities, which is governed by the intermolecular self-interaction. Mainly, the amino acid residues exposed in the complementarity determining regions (CDRs) are involved in these intermolecular self-interaction[10,11]. These interactions have been studied by using small X-ray scattering, light scattering, by measuring the osmotic second virial coefficient, B_{22}, and the structure factor, $S(q)$[12–14].

The experimental data obtained from SAXS has been computed by molecular simulations using analytical models (e.g. coarse grained modeling), based on Derjaguin–Landau–Verwey–Overbeek (DLVO) potential[15,16]. In this model, the approximation is, protein is described as a uniformly surface-charged in a dielectric continuum. Coarse-grained modeling has been performed on mAbs at high concentration by Dear et al[17] using a 12 bead model of the mAb shape to fit the SAXS derived structure factors, where the model provides an insight into short-ranged anisotropic attraction and cluster size distribution properties[17]. The osmotic second virial coefficient has been previously modeled using monte carlo simulations for many protein systems by Prytkova et al[18], where the method has been validated using the experimental data including B_{22} and $S(q)$ of hen egg-white lysozyme in solution.

Here, we have used two monoclonal antibodies, PPI03 and PPI13 that have more than 90.8% sequence similarity, for interactions and ion binding studies using SAXS, SLS and coarse-grained modeling in solution pH 6.0 in the absence and presence of NaCl at high protein concentration. We report the B_{22} and $S(q)$ obtained from SAXS and SLS to extract the information about protein-protein interactions. Further, we construct a single protein coarse-grained (CG) model using SAXS derived rigid body model. In this model, each amino acid is represented by a spherical bead that can be neutral or charged according to the amino acid species and the solution pH. This model could predict the Cl^- ion accumulation on the surface of the mAbs, which could help to gain further insight into the protein-protein interaction in presence of NaCl.

5.2 Theory

5.2.1 Small-Angle X-ray Scattering (SAXS)

Small angle X-ray scattering (SAXS) is a biophysical method used to study the structural properties and interactions of biological molecules in solution[19]. The scattering intensity is recorded by the a detector as a function of momentum transfer vector q, i.e., $I(q)$. Due to the random orientations of the molecules in the solution SAXS intensities is an average over all the positions in the particle.

The intensity, $I(q)$, in SAXS is measured as a function of the momentum transfer vector q[19]:

$$q = \frac{4\pi \sin\theta}{\lambda} \tag{5.1}$$

λ is the wavelength and 2θ the scattering angle. The intensity can be written as:

$$I(q) \propto P(q) * S(q) \tag{5.2}$$

The total scattering intensity, $I(q)$, thereby results from $P(q)$ and the contribution from $S(q)$. Here, $P(q)$ is the form factor represents averaged over all sizes and orientations of the scattering particle, while *S(q)* describes the inter-particle interaction present in the solution, such as repulsion and attraction. At infinite dilution, *S(q)* = 1 and no inter-particle interaction is presents[20]. We get the $P(q)$ by merging the low to high concentration data of a particular sample, which is the form factor of the sample. This was done in order to obtain the effective structure factor, *S(q)* , with varying concentrations which was found by dividing $I(q)$ by $P(q)$.

5.2.2 Static Light Scattering (SLS)

Static light scattering (SLS) is used to measure interactions in terms of second virial coefficients, B_{22}, to observe the protein solution propertiess[21]. At fixed protein concentration, the static light scattering measures the excess Rayleigh ratio, $\bar{R}_{\theta,}$ is given by[22]:

$$\frac{Kc(\partial n/\partial c)^2}{\bar{R}_\theta} = \frac{1}{M_{\mathrm{w}}RT}\left(\frac{\partial \Pi}{\partial \rho}\right) \tag{5.3}$$

Here M_{w} is the protein average molecular weight, R is the gas constant, T is the temperature, $\partial n/\partial c$ is the refractive index increment of the protein solution, c is protein concentration, $\partial \Pi/\partial \rho$ is the osmotic compressibility and K is the optical constant which is equal to $2\pi^2 n_0^2/(N_{\mathrm{A}}\lambda^4)$, where n_0 is the refractive index of the solvent and N_A is Avogadro's number.

The osmotic compressibility of the solution is a virial expansion measured second virial coefficient denoted by B_{22}, and the limit lies until low protein concentration.

The Rayleigh ratio can be related to B_{22} according to:

$$\frac{Kc(\partial n/\partial c)^2}{\bar{R}_{\theta,}} = \frac{1}{M_{\mathrm{w}}} + 2B_{22}c \tag{5.4}$$

where B_{22} can be calculated from the linear fit to Equation 5.4 and molecular weight by the inverse of y-intercept.

5.2.3 Monte Carlo Simulations

One–body model

We construct a coarse-grained (CG) model using SAXS derived rigid body model where each amino acid is represented by a spherical bead that can be neutral or charged according to the amino acid species and the solution pH. Each bead has a diameter equal to $\sigma_{ii} = (6M_w/\pi\rho)^{1/3}$ as specified in[23]. Here, M_W is the molecular weight of the amino acid[24] (in g mol^{-1}) in the protein chain (after neglecting all the hydrogens) and $\rho = 1$ (in g mol^{-1} Å^{-3}) is an average amino acid density. It has been shown[23] that varying the amino acid density by $\pm 20\%$ does not influence considerably structure factor calculations. N- and C- protein terminals are represented by independently beads. M_W and σ_i used in our simulations are reported in Table 5.1.

Our simulation model is composed of a single antibody fixed in the center of a spherical cell of radius R=100 Å containing mobile salt ions (NaCl) and counter ions to ensure electroneutrality, see Figure 5.1. Here, smaller beads represent Na^+ (blue) and Cl^- (red) salt particles, while bigger beads represent amino acids neutral (grey), positively (blue), and negatively (red) charged. The solvent is treated implicitly by a dielectric continuum. The salt is treated explicitly (see salt parametrization for more details) with a grand-canonical scheme, or implicitly using the Debye-Hückel approximation with a Debye length, λ, matching the ionic strength, I, via $\lambda = \sqrt{\varepsilon_r \varepsilon_0 k_B T / 2 N_A e^2 I}$. Here, ε_0 is the vacuum permittivity, $\varepsilon_r = 78.7$ is the water relative permittivity at the temperature $T = 298.15$ K, k_B is Boltzmann's constant, N_A is Avogadro's number, and e is the positive electron unit charge.

Table 5.1 M_W and σ_i of amino acid residues and N and C terminals used in the simulations

Amino Acid	**Abbreviation**	M_w	σ_{ii} (Å)
C- terminal	CTR	16	3.1
N-terminal	NTR	14	3.0
Alanine	ALA	66	5.0
Isoleucine	ILE	102	5.8
Leucine	LEU	102	5.1
Methionine	MET	122	6.2
Phenylaniline	PHE	138	6.4
Valine	VAL	90	5.6
Proline	PRO	90	5.6
Arginine	ARG	144	6.5
Lysine	LYS	116	6.1
Aspartic Acid	ASP	110	5.9
Glutamic Acid	GLU	122	6.2
Glutamine	GLN	120	6.1
Asparagine	ASN	108	5.9
Histidine	HIS	130	6.3
Serine	SER	82	5.4
Threonine	THR	94	5.6
Tyrosine	TYR	154	6.7
Cysteine	CYS	103	5.8
Tryptophan	TRP	176	7.0

Interactions in solution take place accordingly to the following potential energy where n is the total number of interaction sites,

$$U = \underbrace{\sum_{i}^{n-i} \sum_{j=i+1}^{n} \frac{e^2 z_i z_j}{4\pi\varepsilon_0 \varepsilon_r r_{ij}} e^{-\frac{r_{ij}}{\lambda}}}_{Electrostatic\ energy} + \underbrace{4\varepsilon_{ij} \sum_{i}^{n-i} \sum_{j=i+1}^{n} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6} \right]}_{vdW\ \&\ excluded\ volume\ energy} + \underbrace{k_B T \sum_{i}^{n_p} (\mathrm{p}K_{a,i} - \mathrm{pH}) \ln 10}_{Titration\ energy} \tag{5.5}$$

The first term accounts for electrostatic interactions where z_i and z_j are the charges of the i^{th} and j^{th} interaction site (amino acid or a salt/counter-ion particle bead); r_{ij} is the centre distance between them. For explicit salt simulations, the Debye length is set to infinity. The second term is the Lennard-Jones potential, accounting for van der Waals and excluded volume interactions. σ_{ij} and ε_{ij}, are obtained using the Lorentz-Berthelot mixing rules, $\sigma_{ij} = (\sigma_{ii} + \sigma_{jj})/2$; $\varepsilon_{ij} = \sqrt{\varepsilon_{ii}\varepsilon_{jj}}$, where σ_{ii} and σ_{jj} are the diameters and ε_{ii} and ε_{jj} are the self-interaction energy of the i^{th} and j^{th} bead. Finally, the last term accounts for the energy due the protonation/deprotonation of titratable amino acids, n_p, and p$K_{a,i}$ are model acid dissociation constants for titratable amino acid side-chains[25].

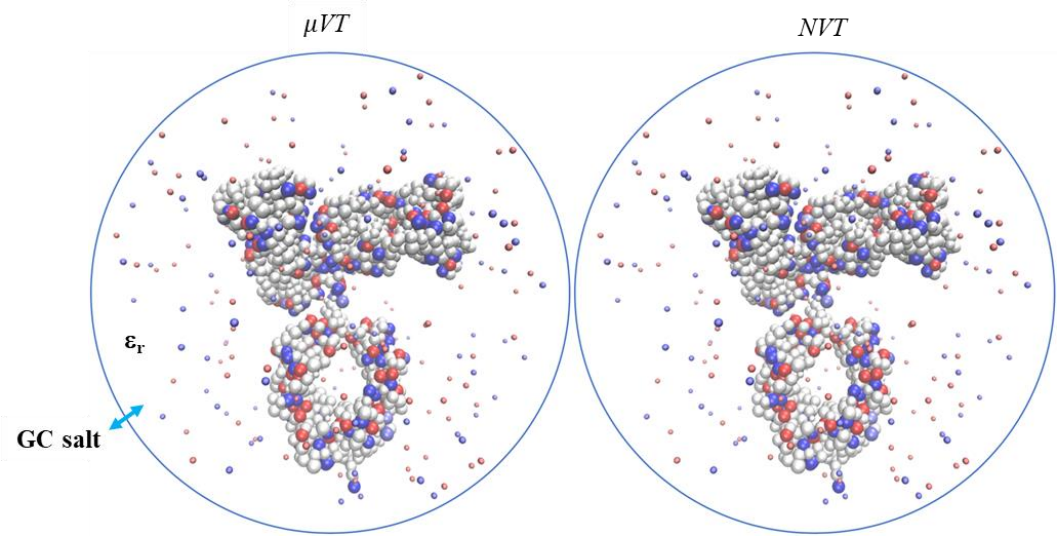


Figure 5.1 Monte Carlo simulation models for a single protein in an aqueous salt solution. Smaller beads represent Na^+ (blue) and Cl^- (red) ions, while bigger beads represent amino acid neutral (grey), positively (blue) and negatively (red) charged. Left: salt particles are inserted and deleted using a grand-canonical scheme, μVT, according to their activity or chemical potential corresponding to experimental salt concentration. Right: a *fixed* average number of salt particles is inserted in the cell according to the preceding μVT simulations.

Salt Parametrization

To have a realistic force field for ions in implicit solvent simulations, we find ionic radii that best match the experimental activity coefficients of Na^+ and Cl^-. Further, a number of different salt species are used as proxy for interactions between ions and amino acids, see Table 5.2.

Table 5.2 Ionic radii matching the experimental activity coefficient for NaCl and other salt species.

Salt	σ_{ii} **(Å)**	**Proxy for**
NaCl	4.6	
NaAc	5.25	$Na^+ \leftrightarrow$Asp/Glu/CTR
NH_4Cl	4.7	Lys$\leftrightarrow Cl^-$
GndCl	3.5	Arg$\leftrightarrow Cl^-$

We perform monte carlo (MC) simulations where salt particles are inserted into a cubic box using a grand-canonical scheme accordingly to a given activity. Varying the σ_{ij} value for a give salt species

and maintaining all the other parameter fixed in the Equation 5.5 (only the first two terms of the equation are used for the salt parametrization and the Debye length is set to infinity) we found the values that best match the experimental activity coefficient in the concentration range 0 to 1 M (see Figure 5.2).

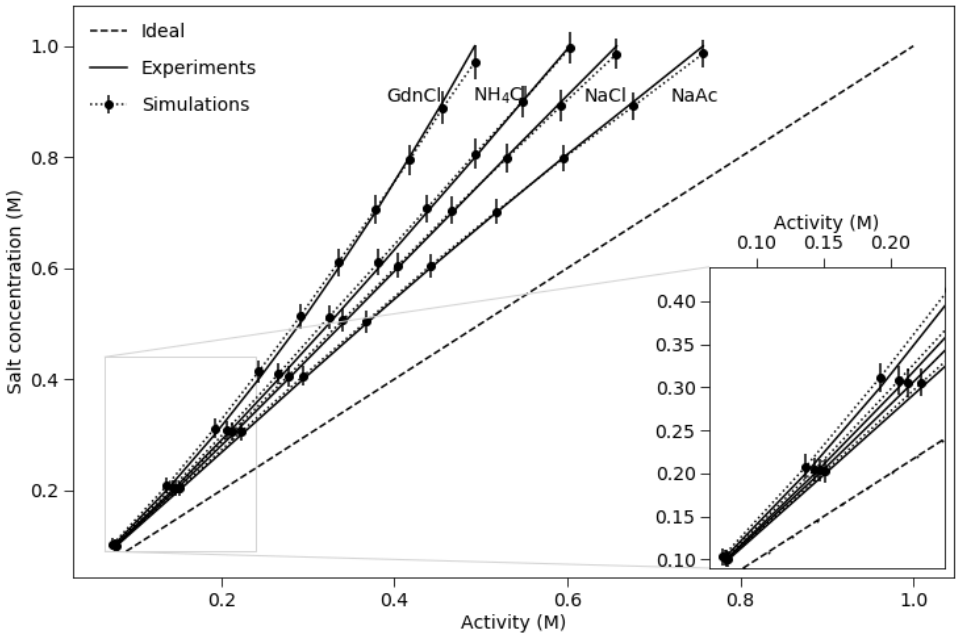


Figure 5.2 Simulated and experimental activity coefficients for different salt concentrations. The experimental data is used to fit the Lennard-Jones parameters for salts and amino-acid ion interactions.

5.3 Materials and Methods

Two monoclonal antibodies belong to IgG subclass type-1 (PPI13 and PPI03) were provided by AstraZeneca, Grant Park, UK, details of the proteins are mentioned below. They have a 90.8% sequence similarity in the heavy chain (see Table S5.1) and differ in the Fv region while having a 96.7% sequence similarity in the light chain (see Table S5.2) being different in the CDR region. Experimentally, we investigated ionic strength dependency by using Small-Angle X-ray Scattering (SAXS) and Static Light Scattering (SLS). Coarse Grained Monte Carlo (CG-MC) simulation was performed on SAXS based rigid body modeling structure.

Table 5.3 Protein overview.

Code	**Type**	M_W **(kDa)**	**Extinction coefficient**[26]**, ε** (L g^{-1} cm^{-1})	**pI**	**Sequence**	**Notes**
PPI03	Human IgG1κ	144.8	1.435	8.44	available	Wild-type IgG
PPI13	Human IgG1κ TM	148.9	1.66	9.04	available	

5.3.1 Sample preparation

Slide-A-Lyzer™ dialysis cassettes (Thermo Fisher Scientific®) were used for dialysis of each protein in 10 mM histidine buffer at pH 6.0. Buffer exchange was performed after two and four hours, ensuring a dilution of at least 500 times sample volume. Each dialysis was continued overnight at a cold temperature (4 °C) with gentle stirring. The individual samples were prepared by diluting or concentrating (Vivaspin® centrifugal concentrators with polyethersulfone membrane, 50 kDa cutoff) the dialyzed sample. The protein concentration after dialysis was measured using NanoDrop™ 8000 Spectrophotometer using the respective extinction coefficients (see Table 5.3) at 280 nm. Samples with varying salt concentrations were prepared from buffer stock solutions with a high NaCl content by adding an appropriate amount of buffer to the individual samples.

5.3.2 SLS data collection

A Wyatt DynaPro NanoStar was used for the static light scattering experiments. The molecular weight and second virial coefficient of the samples were measured using this instrument. . It is a cuvette based system with low sample consumption. Samples were filtered before measuring to eliminate dust particles. Dynamic Software (Wyatt Technology Corporation, Santa Barbara, CA) was used to acquire and process the SLS data. Measurements were performed on a series of samples at various concentrations from 1 g/L to 10 g/L in different formulation conditions (see Table S5.4).

5.3.3 SAXS data collection

SAXS experiments were performed at the ESRF synchrotron, BM29 bioSAXS[27] beamline at Grenoble, France. Measurements on pure water were used to get the data on an absolute scale. Buffers measurements were done both before and after every sample and averaged before subtraction from the sample's scattering profile. Data collection parameters for BM29 are listed in Table S5.3 in supplementary materials. Measurements were performed on a series of samples at various concentrations from 1 g/L to 130 g/L in different formulation conditions (see Table S5.4). The highly concentrated samples (more than 30 g/L) were injected manually in the flow cell due to high viscosity, where the lower concentrated samples were injected automatically by the sample changer.

The ATSAS program package version 2.8.4[28] was used for data analysis. In order to compare the scattering profiles of the measurements, the data has to be normalized by the protein concentration of each sample. Using Primus[29], subsequent buffer averaging and subtraction were done prior to the data analysis. Evaluation of the Guinier region was performed within Primus. The atomic pair distribution function, $p(r)$, was determined by using the program GNOM[30]. Rigid body modelling was performed for mAbs using CORAL[31] using the homology model high resolution structure of the individual subunits, which takes flexibility of the protein into account.

5.4.3 Metropolis Monte Carlo Simulations

Mimicking the experimental conditions at pH 6.0 and 35 mM, 70 mM, and 140 mM NaCl concentration and 10 mM histidine buffer, we use a one-body protein model with the purpose of predicting ion distributions around the antibodies. We approximate the buffer effects adding 10 mM of NaCl on each of the previous experimental condition.

Metropolis Monte Carlo (MC) simulations are performed with Faunus[32,33] which incorporates different trial moves: translation and rotation (salt particles) and amino acid titration. At each simulation step, *N* random moves are attempted. First, we use a grand canonical Monte Carlo (GCMC) scheme, where salt particles are inserted accordingly to their activity or chemical potential, corresponding to the above salt concentrations. During these simulations, the amino acid protonation states are allowed to fluctuate around their equilibrium value accordingly to the solution conditions and local environment. Equilibration runs are performed (10^4 steps) before performing the production runs (10^5 steps) and starting to sample the properties of interest. The average number of salt particles and the mean values of amino acid charges are then fixed and used in Canonical Monte Carlo simulations (CMC) in order to spatially sample ion distributions in the simulation cell. Using these data, ion density maps are created through the VolMap Plugin of the VMD software[34]. Identical iso-density values are chosen for each condition in order to highlight differences in ion densities between the two antibodies.

5.4 Results and Discussion

5.4.1 Protein-protein interactions by SLS

The SLS data were used for determination of the osmotic second virial coefficient, B_{22}, and the molecular weight, M_W of PPI03 and PPI13 over a range of ionic strengths (0 mM, 35 mM, 70 mM and 140 mM NaCl) in 10 mM histidine buffer at pH 6.0.

All the measured B_{22} for PPI03 are positive (see Figure 5.3), reflects PPI03 is repulsive in all the tested conditions. Moreover, with increasing ionic strength up to 140 mM, B_{22} is decreasing, indicates ionic strength screens the electrostatic repulsion. Measured molecular weight for PPI03 in all the tested conditions was on the order of 146.0 ± 5 kDa (see Table S5.5), indicating monodisperse samples.

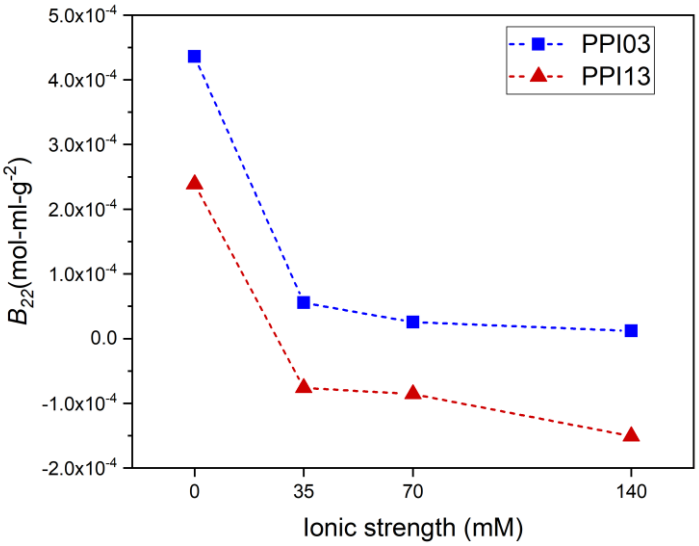


Figure 5.3 B_{22} values plotted vs ionic strength for PPI03 (blue) and PPI13 (red).

In Figure 5.3, calculated B_{22} for PPI13 are negative even at low ionic strengths, reflects attraction in the system. However, in the absence of NaCl, PPI13 has positive B_{22}. Calculated M_W for PPI13 is 146.9

kDa at 0 mM NaCl, while the M_W in the presence of NaCl is above 150 kDa (see Table S5.5), indicates the presence of aggregates.

5.4.2 Protein-protein interactions by SAXS

To study the interaction and conformational changes, SAXS concentration series data were collected at 10 mM histidine buffer at pH 6.0. In Figure 5.4, we show plots of pair distribution function, *p(r)*, obtained from SAXS measurements at 1 g/L protein concentration for PPI03 and PPI13. Each of the *p(r)* curves contains two peaks, which has been previously observed for antibodies[35]. The first peak at ~4.2 nm reflects interatomic distances occurring within the same domain, whereas the second peak at ~7.4 nm corresponds to the distance between atoms in different domains[36].

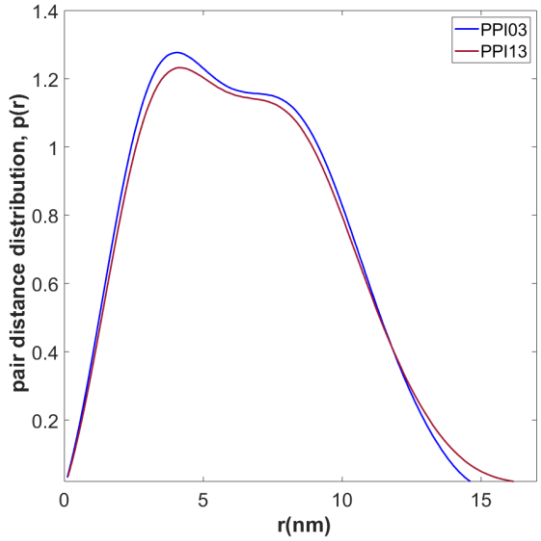


Figure 5.4 Pair distribution function obtained from the SAXS profiles for PPI03 and PPI13 at 1 g/L protein concentrations at 10 mM histidine pH 6.0.

For interactions studies, structure factor analysis was performed and shown in Figure 5.5, where $S(q) > 1$ indicates attraction and $S(q) < 1$ indicates repulsion. Structure factor curves for PPI03 lies below 1 in y-axis scale, indicates repulsion and with increasing ionic strength repulsion decreases as the *S(q)*s are increasing but all of them lies below 1. Moreover, *S(q)* plots lies above 1 in x-axis scale in presence of NaCl indicates aggregation, while lies below 1 for PPI13 in at 0 mM NaCl.

SAXS data shows similar self-association interaction patterns for both mAbs as previously observed from SLS (see Figure 5.3 and Figure 5.5). From structure factor analysis, it has been observed that PPI13 shows attractive interactions even at lower concentrations of proteins, while PPI03 is stable around 100 g/L concentration.

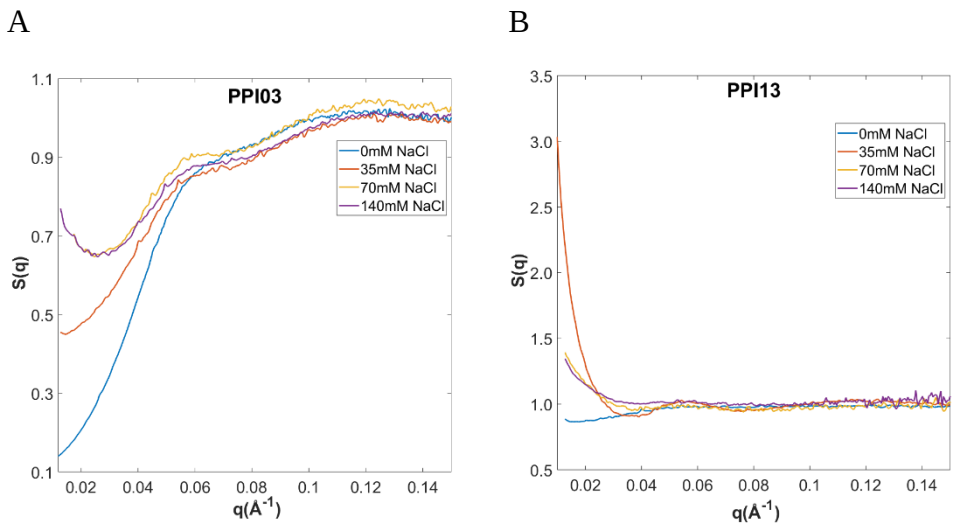


Figure 5.5 SAXS structure factor profiles for (A) PPI03 at around 125 g/L and (B) PPI13 at around 25 g/L.

5.4.3 Rigid body modeling: SAXS

To further study the overall conformation of both the mAbs, rigid body modeling for form factors of all measured conditions was performed using CORAL[31], which allows for domain flexibility and generates the best fitted model for the experimental scattering curve. CORAL derived models for PPI03 and PPI13 are shown in Figure 5.6. The CORAL models for all the tested conditions are later used for coarse grained modeling.

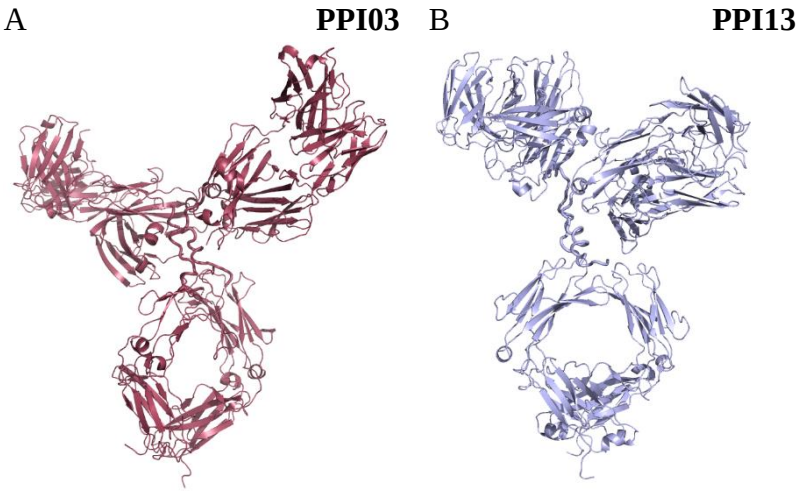


Figure 5.6 SAXS derived CORAL based rigid body model of PPI03 and PPI13 at 10 mM histidine pH 6.0 and the figures are made using PyMOL[37].

The formfactor of each measurement was fitted with the theoretical scattering curve from the CORAL model using CRYSOL[38] are shown in Figure 5.7. All the models have good fit with the experimental data, which can be compared by the values obtained for χ^2, which describes the discrepancy between calculated and experimental data.

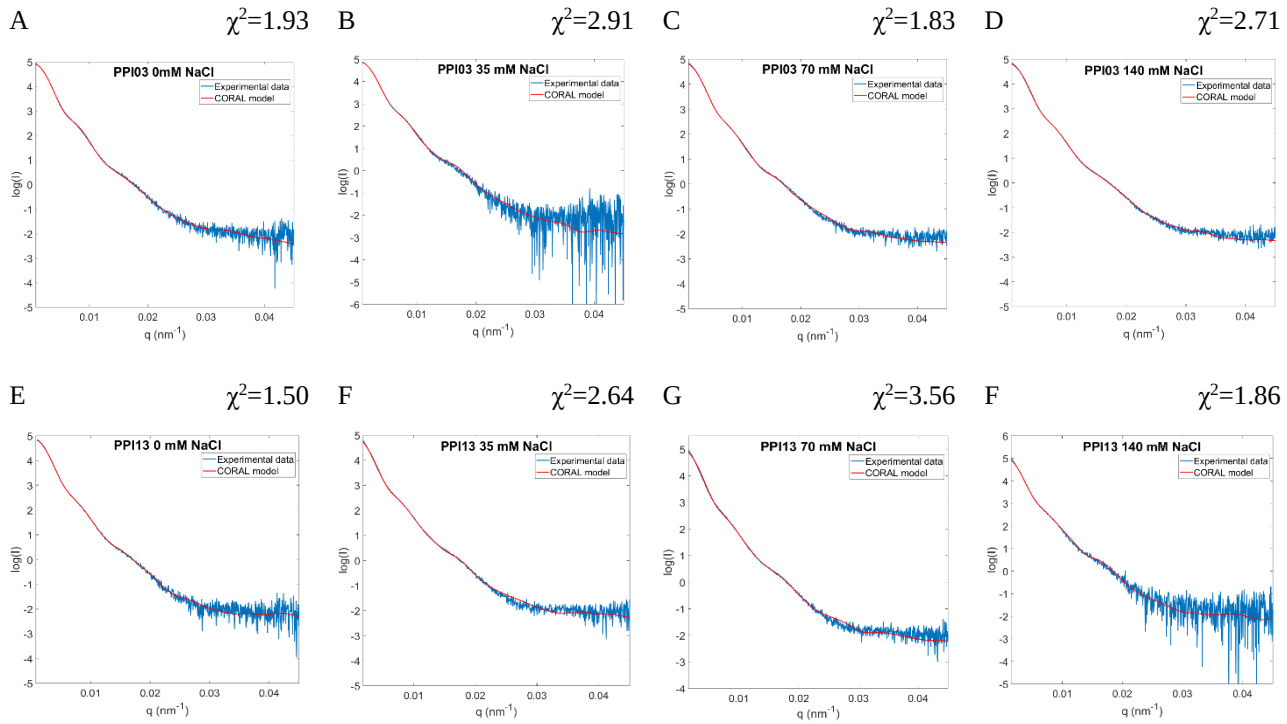


Figure 5.7 Fit plots for rigid body models from CORAL to the experimental data, CORAL models for PPI03 and PPI13 in (A, E)10 mM histidine 0 mM NaCl, (B, F) 10 mM histidine 35 mM NaCl, (C,G) 10 mM histidine 70 mM NaCl and (D,F) 10 mM histidine 140 mM NaCl.

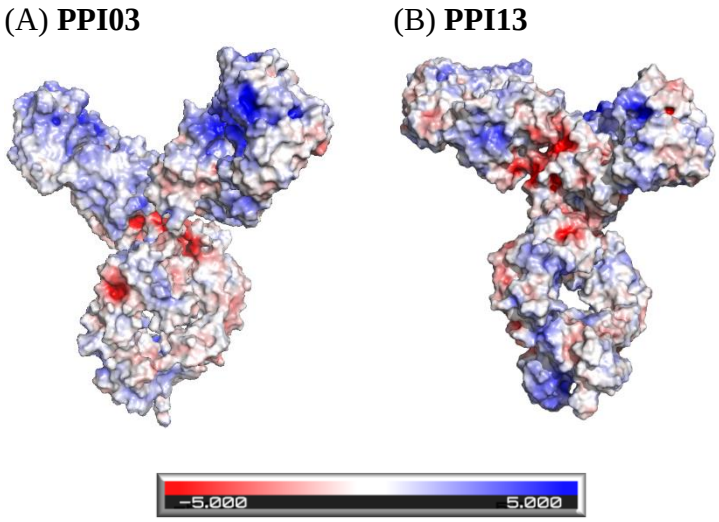


Figure 5.8 Surface electrostatics representation of SAXS derived structures for (A) PPI03 and (B) PPI13 using APBS tools in PyMOL[37].

As presence of NaCl, effect both the mAbs in a unique way and influences the electrostatic interactions, in order to understand these behavior, surface electrostatic potentials were investigated qualitatively using SAXS derived models at 10 mM histidine at pH 6.0 using APBS tools in PyMOL. In Figure 5.8, the electrostatic surface for PPI03 shows less prominent negative and more positive patches around the Fab domains compared to PPI13. This could describe the more repulsive characteristics of PPI03 due to presence of more positive charge. Moreover, for PPI13, due to presence of more hydrophobic patches, it is easier to screen the electrostatics in presence of NaCl and leads to attraction. To get a

better understanding of the Cl ion binding on the surface, Monte Carlo simulations has been performed using the SAXS derived models.

5.4.4 Monte Carlo Simulations

Static light scattering (SLS) measurements show that PPI03 and PPI13 monodisperse salt solutions behave differently in terms of the osmotic second virial coefficient, B_{22}. Considering that both antibodies have a net positive charge and considering the degree of similarity of them in both heavy and light chains, we would expect similar behaviors in solution. Instead, while for PPI03 B_{22} is positive for all salt conditions, indicating repulsion in solution, for PPI13 it becomes negative even for small amount of salts, indicating attraction. We tried to explain this different behavior in terms of charges and ions distribution around the antibodies. In Table 5.4, the average antibody charge calculated using our GCMC scheme as a function of the solution conditions are reported.

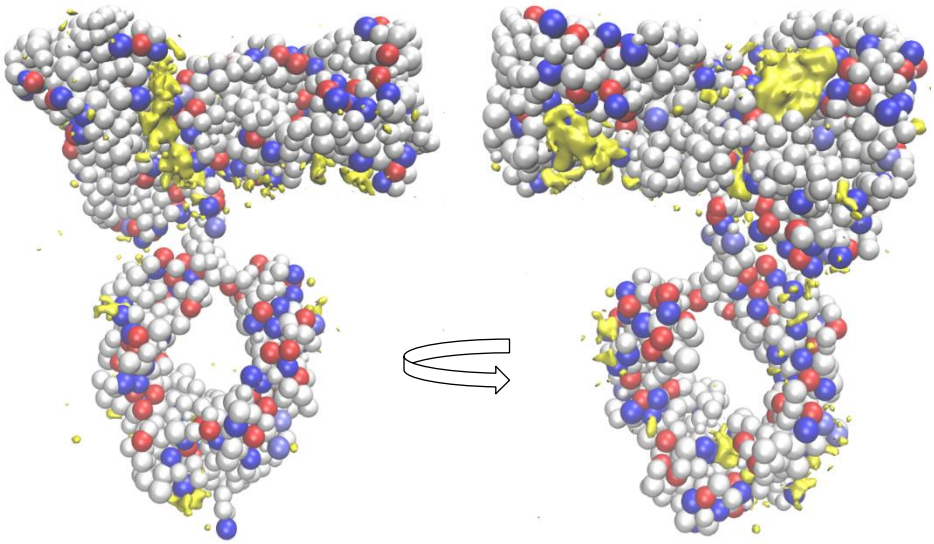
Table 5.4 Average antibodies charges obtained from GCMC simulations for different ionic strength conditions

Protein – Ionic Strength	**Average charge**
PPI03 - 35 mM	35.4
PPI03 - 70 mM	36.3
PPI03 - 140 mM	37.4
PPI13 - 35 mM	29.2
PPI13 - 70 mM	30.4
PPI13 - 140 mM	31.1

Charges increase with salt concentration for both antibodies and PPI03 is more positively charged than PPI13 by roughly +5*e*. Also, as shown in Figure 5.8, the electrostatic surface for PPI03 have more positive patches compared to PPI13. These, differences could result in a larger chloride ions accumulation around PPI03.

To study this, we create iso-density maps of Cl^- ions for both PPI03 and PPI13. In Figure 5.9, Figure 5.10 and Figure 5.11, the chloride density maps (yellow surfaces) are shown for different salt concentrations. It is easy to note that for all conditions explored, Cl^- ions accumulate more on PPI03 compared to PPI13. Also, it seems that Cl^- ions accumulate on particular regions or patches on PPI03 surface while in PPI13 ions accumulation is more spread out. We can hypothesize that if positive patches are "blocked" by ions accumulation, only negative patches remain involved in protein-protein interactions. This mechanism can be responsible for leading to a poorly repulsion in solution and then in a slightly positive B_{22} value for PPI03.

PPI03 35 mM NaCl (isovalue = 0.0022)



PPI13 35 mM NaCl (isovalue = 0.0022)

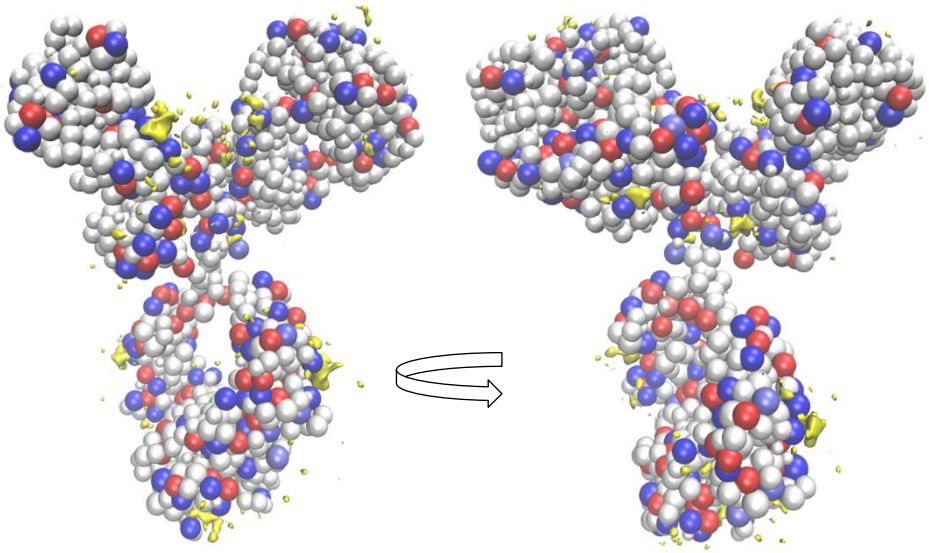
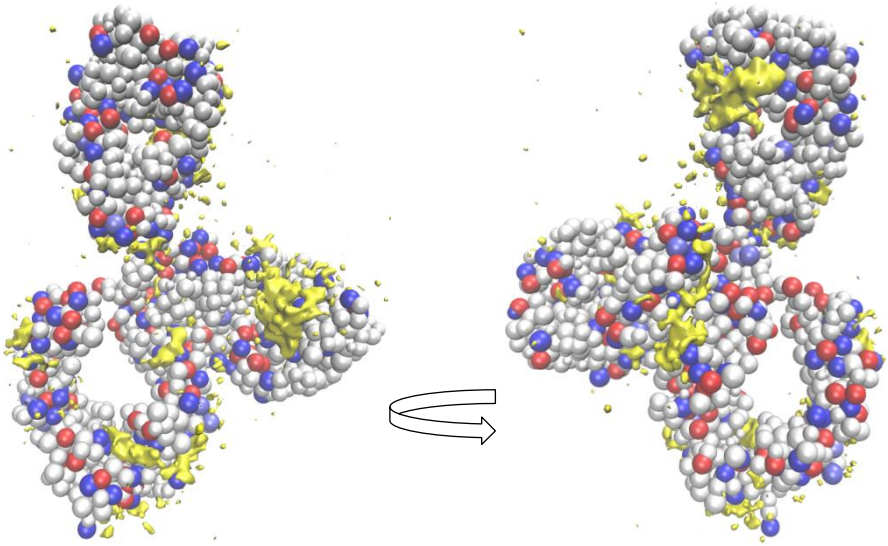


Figure 5.9: Iso-density maps of Cl^- for PPI03 (upper panel) and PPI13 (lower panel) at 35 mM NaCl at pH 6.0 with iso values = 0.0022.

PPI03 70 mM NaCl (isovalue = 0.0027)



PPI13 70 mM NaCl (isovalue = 0.0027)

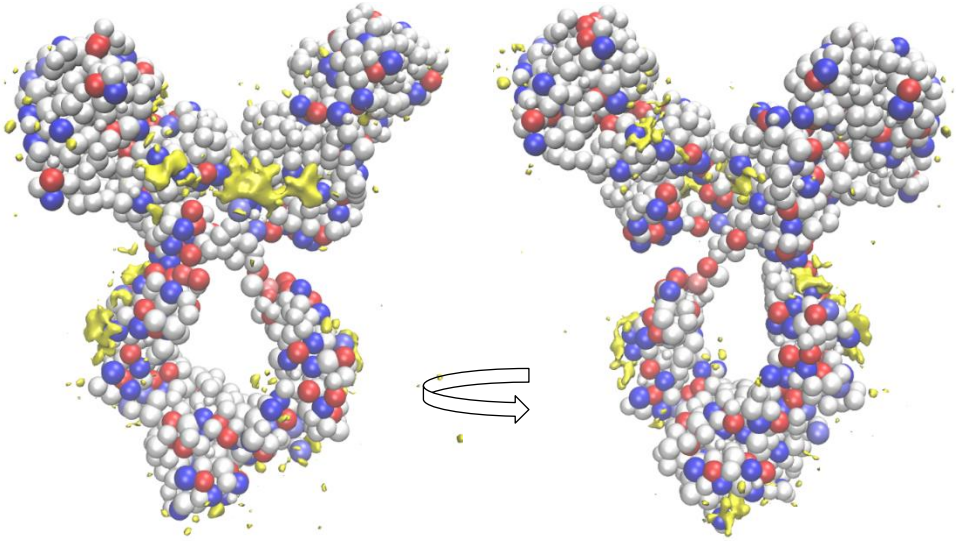
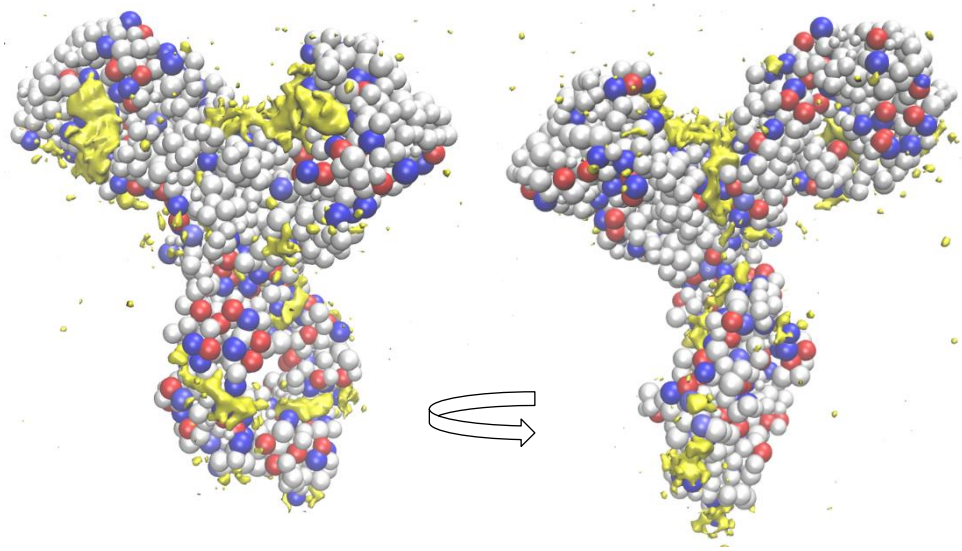


Figure 5.10 Iso-density maps of Cl^- for PPI03 (upper panel) and PPI13 (lower panel) at 70 mM NaCl at pH 6.0 with iso values = 0.0027.

PPI03 140 mM NaCl (isovalue = 0.0036)



PPI13 140 mM NaCl (isovalue = 0.0036)

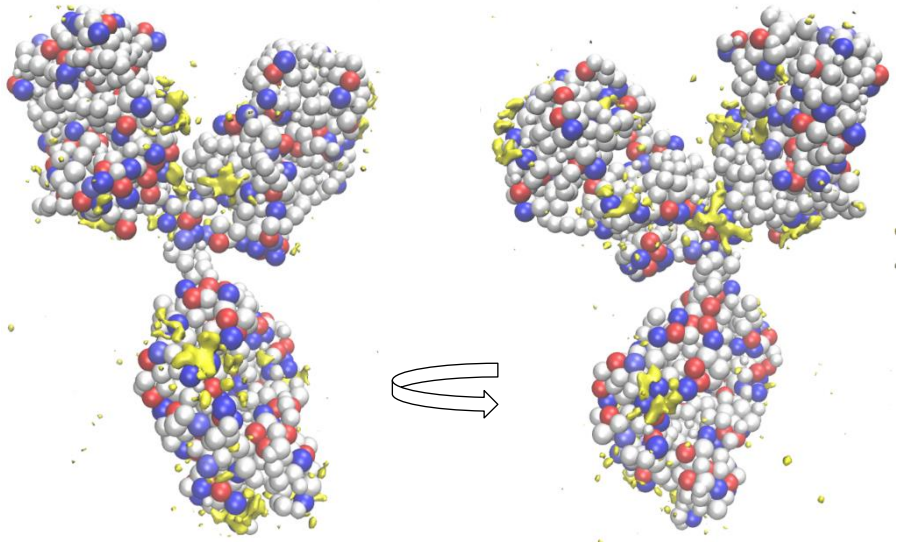


Figure 5.11 Iso-density maps of Cl^- for PPI03 (upper panel) and PPI13 (lower panel) at 140 mM NaCl at pH 6.0 with iso values = 0.0036.

Ongoing work: Two-body Simulation Model

Two identical proteins are placed in a cylindrical box with radius *R*=150 Å and length *L*=500 Å. They can translate and rotate along *z*-axis and amino acid charges are allowed to titrate as in the one-body model. The energy function is identical to the single protein model and due to the added computational complexity, we use implicit and explicit salt particles for the simulation.

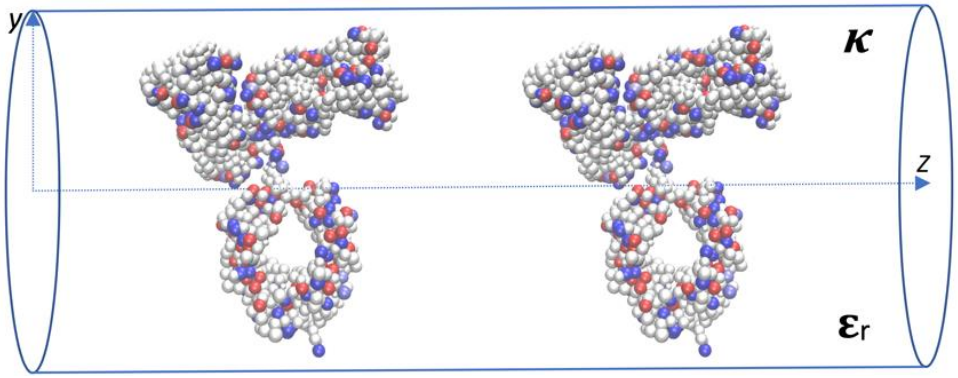


Figure 5.12 Monte Carlo simulation models for a two protein in an aqueous salt solution.

The angularly averaged protein-protein radial distribution function, $g(r)$, is sampled as a function of protein mass center separation, *r*, using the histogram method. From $g(r)$, using Boltzmann inversion, we get the potential of the mean force (PMF) with the purpose of calculate the osmotic second viral coefficient, $B_{22} = -2\pi \int_0^\infty [g(r) - 1] r^2 dr$. We are performing this study to develop a model which could complement the second virial coefficient obtain from light scattering technique.

5.5 Conclusion and Perspective

The multidomain mAbs show that their self-interaction behavior would be unique from one another. In the NaCl screening of the two mAbs, with increasing ionic strength we see repulsion until 140 mM NaCl for PPI03, whereas PPI13 shows attraction even at low ionic strength. Here SAXS derived CORAL models have acceptable χ^2 for CG simulations. From the monte carlo simulations from one-body model, we could find the Cl^- ion binding to the surface of the proteins. The simulation results shows more Cl ion is binding to PPI03 compared to PPI13 and the positive patches in PPI03 are blocked.

Using these models, we could also find out which amino acids are involved in the binding. Further to characterize the protein-protein interactions, two body model can be used, where this model becomes important for higher protein concentration for determining the second virial coefficients.

5.6 Supplementary Materials

Sequence alignment

Differences, Gaps

Table S5.1 Heavy chain alignment of PPI03 and PPI13.

Heavy Chain: Identity: 395/456 (86.6%), Similarity: 414/456 (90.8%), Gaps: 6/456 (1.3%)

```
PPI03 1   QVNLRESGGGLVQPGGSLRLSCAASGFTFGS--YAMSWVRQAPGKGLEWV  48
          ||:|:|||:|||:|:::|:|:|::||:::::  |::||:||:|||||||:
PPI13 1   QVQLQESGPGLVKPSETLSLTCTVSGGSISADGYYWSWIRQPPGKGLEWI  50

PPI03 49  SAISGSGGSTYYADSVKGRFTISRDNSKNSLYLQMNSLRAEDTAVYYCAR  98
          :::::| |||||::|:|||:|||:|:|||:::|:::|::|:|||||||||
PPI13 51  GSLYYS-GSTYYNPSLKGRVTISGDTSKNQFSLKLSSVTAADTAVYYCAR  99

PPI03 99   RSIYGG---NYYFDYWGRGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAA 145
           ...|.|   ..:||:||||||||||||||||||||||||||||||||||
PPI13 100  TPAYFGQDRTDFFDVWGRGTLVTVSSASTKGPSVFPLAPSSKSTSGGTAA 149

PPI03 146  LGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSS 195
           ||||||||||||||||||||||||||||||||||||||||||||||||||
PPI13 150  LGCLVKDYFPEPVTVSWNSGALTSGVHTFPAVLQSSGLYSLSSVVTVPSS 199

PPI03 196  SLGTQTYICNVNHKPSNTKVDKKVEPKSCDKTHTCPPCPAPELLGGPSVF 245
           |||||||||||||||||||||:||||||||||||||||::||||||
PPI13 200  SLGTQTYICNVNHKPSNTKVDKRVEPKSCDKTHTCPPCPAPEFEGGPSVF 249

PPI03 246  LFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 295
           ||||||||||||||||||||||||||||||||||||||||||||||||||
PPI13 250  LFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKP 299

PPI03 296  REEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTISKAKG 345
           |||||||||||||||||||||||||||||||||||:||||||||||
PPI13 300  REEQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPASIEKTISKAKG 349

PPI03 346  QPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 395
           |||||||||||||:|:||||||||||||||||||||||||||||||||||
PPI13 350  QPREPQVYTLPPSREEMTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNY 399

PPI03 396  KTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||
PPI13 400  KTTPPVLDSDGSFFLYSKLTVDKSRWQQGNVFSCSVMHEALHNHYTQKSL 449

PPI03 446  SLSPGK 451
           ||||||
PPI13 450  SLSPGK 455
```

Table S5.2 Light chain alignment of PPI03 and PPI13.

Light Chain: Identity: 197/214 (92.1%), Similarity: 207/214 (96.7%), Gaps: 0/214 (0.0%)

```
PPI03 1 DIQMTQSPSSLSASVGDRVTITCRASQSISSYLNWYQQKPGKAPKLLIYA 50
        |||||||||:||||||||||||||||:|||:|:||||||||||:|||:
PPI13 1 DIQMTQSPSTLSASVGDRVTITCRASQGISSWLAWYQQKPGKAPKVLIYK 50

PPI03 51 ASSLQSGVPSRFSGSGSGTDFTLTISSLQPEDFATYYCQQSYSTPLTFGG 100
         ||:|:|||||||||||||:||||||||||:||||||:::|:|||:
PPI13 51 ASTLESGVPSRFSGSGSGTEFTLTISSLQPDDFATYYCQQSHHPPWTFGQ 100

PPI03 101 GSKVEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKV 150
          |:|:||||||||||||||||||||||||||||||||||||||||||||||
PPI13 101 GTKLEIKRTVAAPSVFIFPPSDEQLKSGTASVVCLLNNFYPREAKVQWKV 150

PPI03 151 DNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQG 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||
PPI13 151 DNALQSGNSQESVTEQDSKDSTYSLSSTLTLSKADYEKHKVYACEVTHQG 200

PPI03 201 LSSPVTKSFNRGEC 214
          ||||||||||||||
PPI13 201 LSSPVTKSFNRGEC 214
```

Table S5.3 Experimental set up of SAXS measurements

Instrument	**BM29, ESRF**
Wavelength (Å)	0.99
q-range ($Å^{-1}$)	0.004 – 0.49
Sample-to-detector distance (m)	2.864
Detector	PILATUS 1M
Exposure time (s)	10 x 1.00
Beam size	700 x 700 μm^2
Sample configuration	1.8 mm quartz glass capillary
Absolute scaling method	Comparison to water in sample capillary
Normalization	To transmitted intensity by beam-stop counter
Monitoring for radiation damage	Control of un-subtracted and scaled subtracted data for systematic changes typical for radiation damage

Table S5.4 Samples for SAXS and SLS experiments performed at varying protein concentrations

Protein	**Buffer**	**Additive**	**SAXS measured concentrations (g/L)**	**SLS measured concentrations (g/L)**
PPI03 (Human IgG1κ)	10 mM histidine pH 6.0	-	0.5 - 130	1 - 10
		NaCl (35, 70 and 140 mM)		
PPI13 (Human IgG1κTM)	10 mM histidine pH 6.0	-	0.5 - 120	1 - 10
		NaCl (35, 70 and 140 mM)		

Table S5.5 Molecular weight *(M_W)* and B_{22} values calculated from SLS measurements.

PPI03 10 mM histidine pH 6.0	B_{22} **x** 10^{-4} (mol-mL/g²)	M_W (kDa)
0 mM NaCl	4.36	143.2
35mM NaCl	0.553	141.4
70 mM NaCl	0.255	138.2
140 mM NaCl	0.121	151.3

PPI13 10 mM histidine pH 6.0	B_{22} **x** 10^{-4} (mol-mL/g²)	M_W (kDa)
0 mM NaCl	2.39	146.9
35mM NaCl	-1.51	165.8
70 mM NaCl	-0.762	181.2
140 mM NaCl	-0.856	190.3

Acknowledgement

We thank the European Union's Horizon 2020 research and innovation program for funding (grant agreement No 675074).

Danish Agency for Science, Technology, and Innovation for funding the instrument center DanScatt.

BM29 ESRF, Grenoble for providing beam time for performing the SAXS experiments.

Static light scattering measurements were performed at Wyatt Technology Europe, Dernbach, Germany. Thank you Lorenzo Gentiluomo for helping with the measurements.

AstraZeneca for kindly providing monoclonal antibodies.

Monte carlo simulations were performed at department of theoretical chemistry at Lund university.

References

1. Chames, P., Van Regenmortel, M., Weiss, E. & Baty, D. Therapeutic antibodies: Successes, limitations and hopes for the future. *British Journal of Pharmacology* (2009). doi:10.1111/j.1476-5381.2009.00190.x

2. Weiner, L. M., Surana, R. & Wang, S. Monoclonal antibodies: Versatile platforms for cancer immunotherapy. *Nature Reviews Immunology* (2010). doi:10.1038/nri2744

3. Harn, N., Allan, C., Oliver, C. & Middaugh, C. R. Highly concentrated monoclonal antibody solutions: Direct analysis of physical structure and thermal stability. *J. Pharm. Sci.* (2007). doi:10.1002/jps.20753

4. Fradkin, A. H., Carpenter, J. F. & Randolph, T. W. Immunogenicity of aggregates of recombinant human growth hormone in mouse models. *J. Pharm. Sci.* (2009). doi:10.1002/jps.21834

5. Vekilov, P. G. Phase transitions of folded proteins. *Soft Matter* (2010). doi:10.1039/c0sm00215a

6. Zhou, H.-X., Rivas, G. & Minton, A. P. Macromolecular Crowding and Confinement: Biochemical, Biophysical, and Potential Physiological Consequences. *Annu. Rev. Biophys.* (2008). doi:10.1146/annurev.biophys.37.032807.125817

7. 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 - Proteins Proteomics* (2008). doi:10.1016/j.bbapap.2007.12.016

8. 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). doi:10.1110/ps.072957907

9. Haynes, C. A., Tamura, K., Körfer, H. R., Blanch, H. W. & Prausnitz, J. M. Thermodynamic properties of aqueous ??-chymotrypsin solutions from membrane osmometry measurements. *J. Phys. Chem.* (1992). doi:10.1021/j100181a069

10. Kanai, S., Liu, J., Patapoff, T. W. & Shire, S. J. Reversible self-association of a concentrated monoclonal antibody solution mediated by fab-fab interaction that impacts solution viscosity. *J. Pharm. Sci.* (2008). doi:10.1002/jps.21322

11. Wu, S. J. *et al.* Structure-based engineering of a monoclonal antibody for improved solubility. *Protein Eng. Des. Sel.* (2010). doi:10.1093/protein/gzq037

12. Narayanan, J. & Liu, X. Y. Protein interactions in undersaturated and supersaturated solutions: A study using light and X-ray scattering. *Biophys. J.* (2003). doi:10.1016/S0006-3495(03)74871-1

13. Zhang, F. *et al.* Protein interactions studied by SAXS: Effect of ionic strength and protein concentration for BSA in aqueous solutions. *J. Phys. Chem. B* (2007). doi:10.1021/jp0649955

14. Ianeselli, L. *et al.* Protein-protein interactions in ovalbumin solutions studied by small-angle scattering: Effect of ionic strength and the chemical nature of cations. *J. Phys. Chem. B* (2010). doi:10.1021/jp9112156

15. Seung, J. K., Dumont, C. & Gruebele, M. Simulation-based fitting of protein-protein interaction potentials to SAXS experiments. *Biophys. J.* (2008). doi:10.1529/biophysj.107.123240

16. Abramo, M. C. *et al.* Effective interactions in lysozyme aqueous solutions: A small-angle neutron scattering and computer simulation study. *J. Chem. Phys.* (2012). doi:10.1063/1.3677186

17. Dear, B. J. *et al.* X-ray Scattering and Coarse-Grained Simulations for Clustering and Interactions of Monoclonal Antibodies at High Concentrations. *J. Phys. Chem. B* (2019). doi:10.1021/acs.jpcb.9b04478

18. Prytkova, V. *et al.* Multi-Conformation Monte Carlo: A Method for Introducing Flexibility in Efficient Simulations of Many-Protein Systems. *J. Phys. Chem. B* (2016). doi:10.1021/acs.jpcb.6b00827

19. Svergun, D. I. & Koch, M. H. J. Small-angle scattering studies of biological macromolecules in solution. *Reports Prog. Phys.* (2003). doi:10.1088/0034-4885/66/10/R05

20. Fritz, G. & Glatter, O. Structure and interaction in dense colloidal systems: Evaluation of scattering data by the generalized indirect Fourier transformation method. *J. Phys. Condens. Matter* (2006). doi:10.1088/0953-8984/18/36/S14

21. Murphy, R. M. Static and dynamic light scattering of biological macromolecules: What can we learn? *Curr. Opin. Biotechnol.* (1997). doi:10.1016/S0958-1669(97)80153-X

22. Vrij, A. & Overbeek, J. T. G. Scattering of light by charged colloidal particles in salt solutions. *J. Colloid Sci.* (1962). doi:10.1016/0095-8522(62)90023-5

23. 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). doi:10.1021/jp505809v

24. Lide, D. R. *CRC Handbook of Chemistry and Physics, 89th Edition*. *CRC Handbook of Chemistry and Physics* (2009). doi:10.1016/0165-9936(91)85111-4

25. Thurlkill, R. L., Grimsley, G. R., Scholtz, J. M. & Pace, C. N. pK values of the ionizable groups of proteins. *Protein Sci.* (2006). doi:10.1110/ps.051840806

26. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* (2003). doi:10.1093/nar/gkg563

27. Pernot, P. *et al.* New beamline dedicated to solution scattering from biological macromolecules at the ESRF. in *Journal of Physics: Conference Series* (2010). doi:10.1088/1742-6596/247/1/012009

28. Franke, D. *et al.* ATSAS 2.8: A comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* (2017). doi:10.1107/S1600576717007786

29. Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. & Svergun, D. I. PRIMUS: A Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* (2003). doi:10.1107/S0021889803012779

30. Semenyuk, A. V. & Svergun, D. I. GNOM. A program package for small-angle scattering data processing. *J. Appl. Crystallogr.* (1991). doi:10.1107/S002188989100081X

31. Petoukhov, M. V. *et al.* New developments in the ATSAS program package for small-angle scattering data analysis. *J. Appl. Crystallogr.* (2012). doi:10.1107/S0021889812007662

32. Stenqvist, B., Thuresson, A., Kurut, A., Vácha, R. & Lund, M. Faunus-A flexible framework for Monte Carlo simulation. *Mol. Simul.* (2013). doi:10.1080/08927022.2013.828207

33. Lund, M., Trulsson, M. & Persson, B. Faunus: An object oriented framework for molecular simulation. *Source Code Biol. Med.* **3**, 1–8 (2008).

34. Humphrey, W., Dalke, A. & Schulten, K. VMD: Visual molecular dynamics. *J. Mol. Graph.* (1996). doi:10.1016/0263-7855(96)00018-5

35. Mosbaek, C. R., Konarev, P. V., Svergun, D. I., Rischel, C. & Vestergaard, B. High concentration formulation studies of an IgG2 antibody using small angle x-ray scattering. *Pharm. Res.* (2012). doi:10.1007/s11095-012-0751-3

36. Inouye, H., Houde, D., Temel, D. B. & Makowski, L. Utility of Solution X-Ray Scattering for the Development of Antibody Biopharmaceuticals. *J. Pharm. Sci.* (2016). doi:10.1016/j.xphs.2016.07.021

37. Schrödinger, L. The PyMol Molecular Graphics System, Versión 1.8. *Thomas Holder* (2015). doi:10.1007/s13398-014-0173-7.2

38. Svergun, D., Barberato, C. & Koch, M. H. CRYSOL - A program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.* (1995). doi:10.1107/S0021889895007047

6 Conclusion

In this work, we have used high throughput techniques for biophysical characterization of the therapeutic proteins in combination with SAXS and SLS in order to study the protein-protein interactions at high concentration in various physicochemical conditions. SAXS and SLS were complemented with *in-silico* approach, such as an analytical model with interaction potentials and coarse-grained modeling.

The initial stability and interaction studies of mAbs, albumin and lipase were performed at low protein concentration. The study shows different behavior in in various physicochemical conditions and the behavior depends on their size, structure, excipients present in the solution and function of the proteins. Further, to study the solution structure and characterize the interactions present at high concentration, proteins have been selected which were available at high concentration such as mAbs.

The observation from the form factor of the mAbs at high protein concentrations suggested no changes in mAb conformations. It was possible to observe the inter particle interaction by looking at the structure factors obtained at different protein concentration, i.e., the mAbs show repulsive interaction in the absence of NaCl. SLS data provided the information of the net interaction of the mAbs, i.e., repulsive interaction. SLS and SAXS complement each other well. The simplified spherical model with interaction potentials study provides a more accurate extrapolation of the structure factors. It was concluded that the model containing both a repulsive and an attractive part had the best fit to the experimental SAXS data at low q values, which contain the important information about protein-protein interactions. The model explained that even though the net interaction is repulsive, there might be some weak attractive interaction is present.

The effect of NaCl was also studied on mAbs at high concentration for two mAbs using SAXS and SLS and complemented using monte carlo simulations. SAXS based models can be used to perform coarse-grained modelling. The self-interaction behavior of multidomain mAbs are unique from one another. Using the single protein coarse grained model, we found the information about accumulation of Cl^- ion on the surface on particular region on the mAbs, which is the reason for the difference in self-interaction behavior among the mAbs. In the ongoing work using a two-body model we could determine the net protein-protein interactions in terms of osmotic second virial coefficient.

Appendix

Macromolecular crowding in lipase

Introduction

Macromolecular "crowding" is the term used to describe the excluded-volume effect of macromolecules in biological systems[1,2]. Volume exclusion is a phenomenon resulting from the physical occupation of a finite volume by a population of a given macromolecular species making that space unavailable to other molecules[2,3]. The degree of volume excluded depends on the size of the crowding molecules that are occupying the available volume. In a protein solution, a protein molecule diffuses and distributes itself in the available volume, however, the presence of crowding molecules reduces the volume available. This reduction of the available volume is a consequence of the steric repulsion between the crowding and test macromolecules depending on the chemical nature of the crowder[4].

Effects of molecular crowding has been previously studied in biological systems[5–8]. Thermal studies have been performed creatine kinase in presence of ficoll 70 by Wang et al[7], where they observed that increase in protein concentration increases the $T_{½}$ in the presence of crowding agent. Moreover, the key finding here is macromolecular crowding enhances protein thermal stability. Protein stability study have been also performed in presence of dextran, where high concentrations of dextran in the solutions stabilized cytochrome c by changing conformation of from fully unfolded polypeptide to more compact state in crowded environments[8].

This study will focus on molecular crowding effects in protein formulation. It is not clear what the general effect is of crowding, which can be attributed to the size of the protein (versus the size of the crowding agent) and the intrinsic solubility of the molecule (surface charge distribution, net charge, etc.). To explore these effects, we have chosen one medium-sized protein (Lipase). By increasing the concentrations of protein in a solution, we go from ideality to non-ideality, and if we continue increasing concentration, then they may aggregate or precipitate (depending on the conditions). Excluded volume is one effect on the behavior of the molecule of interest, but another factor is the size of the interstitial spaces. To address these questions, we have studied the behavior of the molecule at a higher concentration by adding large inert molecules (Dextran sulfate 10, Ficoll 70 and PEG 35000) in the protein solution (artificially increasing the protein concentration by excluding the available volume). By using molecules of different sizes, we will be able to address the question of crowder size

versus the size of the protein. One thing is the excluded volume, which can be the same for small, medium or large crowders, depending on the concentration, but what is different and cannot be changed, is the size of the interstitial spaces, and it's the effect of these that we are trying to determine. Dynamic light scattering, nanoDSF were performed to determine the thermal stability and size distribution of the particles.

Materials and Methods

Lipase was chosen as a medium sized protein from the PIPPI library and three crowders (Dextran sulfate 10, Ficoll 70 and PEG 35000) were chosen to perform the crowding experiment. Lipase was provided by Novozymes A/S, Denmark and the three crowders, dextran sulfate 10, ficoll 70 and PEG 35000 were used from Sigma Aldrich.

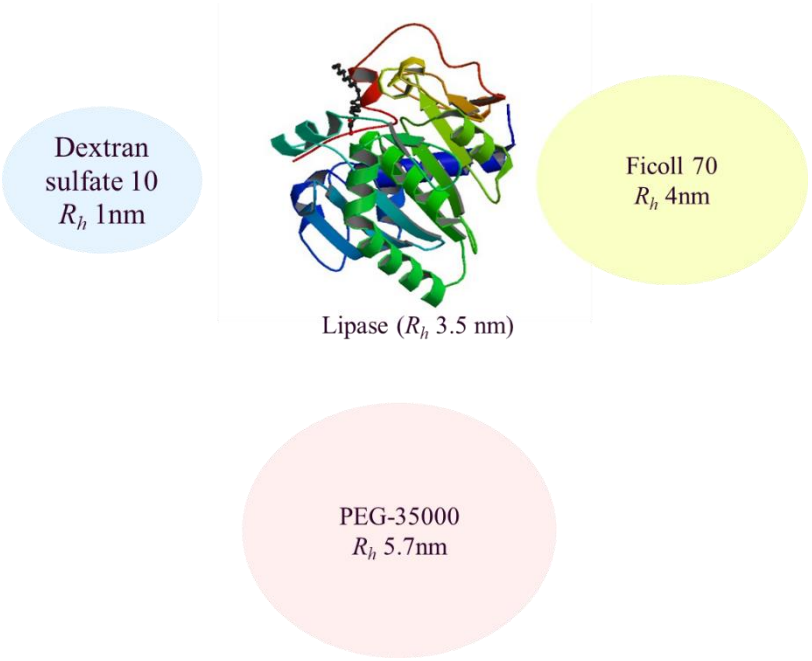


Figure 1 Schematic representation of three crowders and a lipase molecule.

Sample preparation

Slide-A-Lyzer™ dialysis cassettes (Thermo Fisher Scientific®) were used for dialysis of lipase in 10 mM histidine buffer at pH 5.5, 70 mM NaCl. Dialysis buffer was exchanged after a minimum of two hours with the final dialysis step running for 16 hours 4 °C with gentle stirring. The individual samples were prepared by diluting or concentrating (Vivaspin® centrifugal concentrators with polyethersulfone membrane, 50 kDa cutoff) the dialyzed sample. The protein concentration after dialysis was measured using NanoDrop™ 8000 Spectrophotometer using the respective extinction coefficient[9]e, 1.2, at 280 nm. Samples with crowder concentrations were prepared from stock solutions of each crowder in 10 mM histidine buffer at pH 5.5, 70 mM NaCl by adding an appropriate amount for the individual samples.

Dynamic light scattering (DLS) data collection

High throughput dynamic light scattering was conducted on a DynaPro® II Plate Reader (Wyatt Technology, Santa Barbara, USA) to obtain the hydrodynamic radius, R_h. The samples were filtered using a Millex® 0.22 µm filter (Merck Millipore, Darmstadt, Germany) and equilibrated at 25 °C for 10 minutes in the Plate reader. 20 µL per well of each sample were pipetted in triplicates into Aurora 384 Lobase Assay Plates (Aurora Microplates, Whitefish, USA). The plates were sealed with a few µL of silicone oil and then centrifuged at 2000 rpm for 1 minute. Data were processed by the Dynamics software V7.8 (Wyatt Technology, Santa Barbara, USA). DLS data were analyzed by using the DYNAMICS software V7.8.

Differential Scanning Fluorimetry with Intrinsic Protein Fluorescence Detection (nanoDSF)

Samples were filled in standard nanoDSF capillaries (NanoTemper Technologies, Munich, Germany). Measurements were performed using the Prometheus NT.48 (NanoTemper Technologies, Munich, Germany) system that measures the intrinsic protein fluorescence intensity at 330 nm and 350 nm after excitation at 280 nm. A temperature ramp of 1°C/min was applied from 20 to 95°C. The fluorescence intensity ratio (F350/F330) was plotted against the temperature, the inflection points of the unfolding transitions were determined of the first derivative of each measurement using the PR Control software V1.12 (NanoTemper Technologies, Munich, Germany).

Results and Discussions

Selection of solution condition and experimental criteria

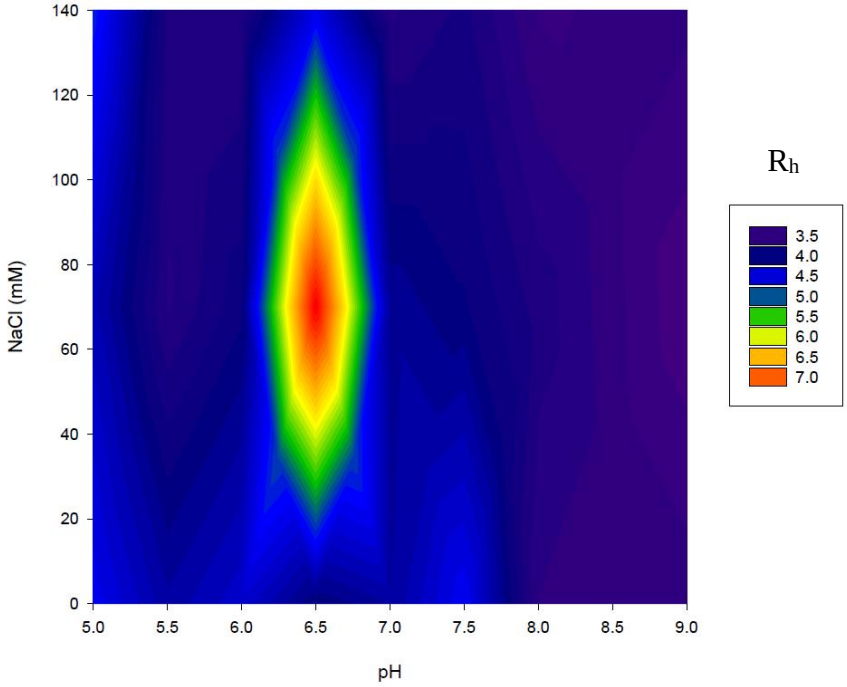


Figure 2 Contour plot of the measured hydrodynamic radius of lipase in 10 mM histidine from pH 5.0 to pH 9.0 in the presence of NaCl.

Initially, DLS measurement were performed to find a solution condition, where the lipase is monodispersed, at least at the lower protein concentration rang by observing the hydrodynamic radius of the lipase. The measurement were performed as a function of pH and NaCl concentrations (0, 70 and 140 mM) in 10 mM histidine buffer.

In Figure 2, we show a contour plot of the measured hydrodynamic radius of lipase, where we see that R_h of the lipase below pH 6.0 and above pH 7.0 lies between 3.5 nm to 4.0 nm, which we were expecting for a monodispersed lipase and from pH 6.0 to pH 7.0, R_h is gradually increasing and have a two fold increase at pH 6.5, in 70 mM NaCl. By observing the overall R_h, we decided to choose a low pH condition with the presence of NaCl, 10 mM histidine pH 5.5, 70 mM NaCl to perform further studies with the crowders.

Table 1 Effective concentration of the crowders[3,6].

Macromolecular crowders	Effective concentration (mg/ml)
Dextran sulfate 10	45.64
Ficoll 70	37.5
PEG 35000	5.21

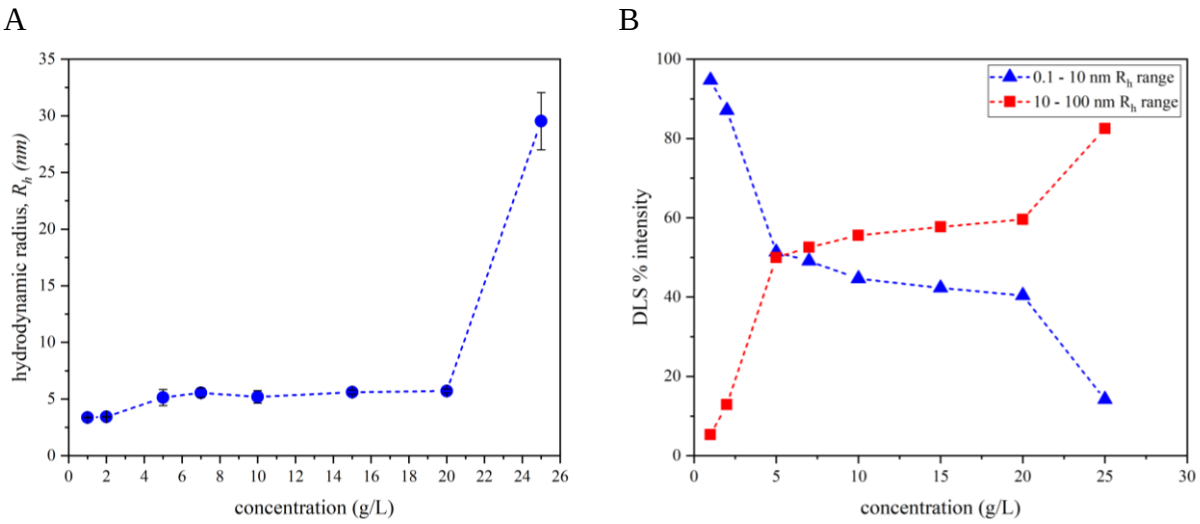


Figure 3 Concentration dependent increase in hydrodynamic radius (Panel A) and % population based on intensity of the concentration series measurements of lipase showing large particle formation at 5 mg/ml and above (Panel B).

A concentration series measurements were performed for lipase. In Figure 3, DLS measurements show with increasing lipase concentration, the R_h is increasing and larger particles of size 10 – 100 nm range are forming. Moreover, the thermal stability of lipase is considerably decreasing as the $T_{½}$ is decreasing with increasing concentration (see Figure 4). From the R_h values for the concentration range above 2 mg/ml there is an increase in R_h if we compared to lower concentrations where the Rh is similar to the expected hydrodynamic radius of the lipase. 2.5 mg/ml was chosen to do further studies with the

crowders to study the effect of molecular crowding. The aim for the measurements is to add differently sized crowders to a solution containing 2.5 mg/ml lipase whilst keeping the volume constant resulting in a concentration series of lipase due to the excluded volume by the crowders. These data will be compared to lipase only data, i.e. without crowder, in order to determine the effect of the crowder on the stability/solubility of the lipase. The effective concentration of each crowder has been calculated and shown in Table 1[6]. Knowing the effective concentration would help to avoid precipitation or phase separation issues of the crowders during the sample preparation steps.

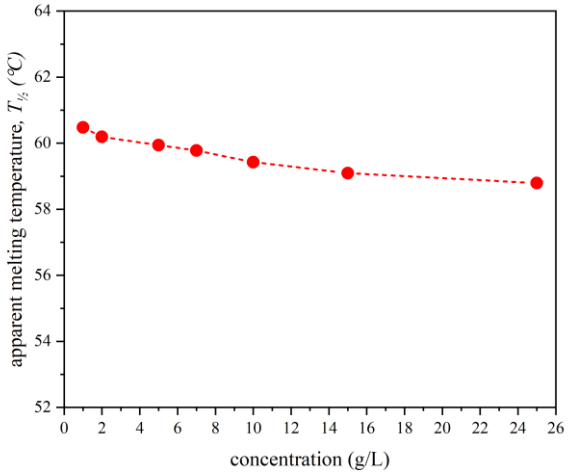


Figure 4 nanoDSF measurements of lipase concentration series.

fractional volume occupancy and crowders concentrations

Table 2 Calculations for estimating the fractional volume occupancy of lipase.

Volume of each lipase molecule considered as a sphere with a radius of 3.5 nm (3.5×10^{-9} m) Volume of a sphere = $4/3\ \pi r^3$ = $4/3\ \pi\ (3.5 \times 10^{-9})^3$ =179.5×10^{-27}cubic meter → 1
Number of lipase molecules in 1 mg ▪ We know that the MW is ~ 29500 Da. ▪ That means, 29500 grams of lipase contains Avogadro number of molecules. ▪ From this, we then calculate the number of molecules of lipase present in 1 mg as follows: $(1 \times 10^{-3}) \times (6.023 \times 10^{23}) \div 29500$ = 2.04×10^{16} → 2
Fraction Volume Occupancy in 1 ml Now, we calculate the volume occupied by these number of lipase molecules in 1 ml by multiplying 1 & 2 and then express in percentage to estimate the fraction volume occupancy: $179.5 \times 10^{-27} \times 2.04 \times 10^{16}$ = 0.37% (v/v)

Macromolecular crowding functions by way of the excluded volume effect and is often referred to as the volume of a solution that is excluded by the presence of one or more background particles in the solution. To evaluate it, the fraction of the total volume occupied by macromolecules was calculated for lipase and each crowder can be calculated from the R_h.

The volume of a single crowding molecule is calculated by assuming that each crowder are spherical. Multiplied by the amount of crowding molecules (to be calculated via molecular weight and concentration used), the total volume that the crowders occupy can be calculated[10]. Calculated fraction volume occupancy of each crowder are shown in Table 3.

Table 3 Fractional volume occupancy exerted by crowders at 1 mg/ml(determined empirically)

Crowders	Molecular Mass (g/mol)	Hydro dynamic Radius (nm)	Volume of each crowder (cubic meter) 4/3 πr3	Effective Conc. (mg)	No. of crowders molecules [(Conc) × (6.023 × 10^{23}) ÷ MM]	Fractional vol. Occupancy in 1ml (column 4 × Column 6)
Dextran sulfate 10	8000	1	4.19×10^{-27}	1	7.53×10^{16}	0.03%
Ficoll 70	70000	4	268×10^{-27}	1	0.86×10^{16}	0.23%
PEG 35000	35000	5.7	776×10^{-27}	1	1.72×10^{16}	1.33%

In Table 4, we calculated the % volume occupied by lipase at 5 mg/ml and above after the subtraction of 2.5 mg/ml of lipase to fill the volume occupancy by adding the crowders. Further, the amount of crowders calculated to fill up the above % volume occupancy were calculated and shown in Table 4.

Table 4 Calculations of concentration required of individual crowders for making the crowding environment same at the concentration series measurements of lipase.

lipase (mg/ml)	Subtracted lipase concentration after 2.5 mg/ml lipase	%Volume occupied by Lipase (from column 2)	Amount of crowders needed to fill the %volume form column 3 (in mg/ml)		
			Dextran sulfate 10	Ficoll 70	PEG 35000
5	2.50	0.93	30.83	4.02	0.70
7	4.50	1.67	55.50	7.24	1.25
10	7.50	2.78	92.50	12.07	2.09
15	12.50	4.63	154.17	20.11	3.48
20	17.50	6.48	215.83	28.15	4.87
25	22.50	8.33	277.50	36.20	6.26

The calculated concentration of the dextran and PEG marked in red crosses the effective concentration limits (see Table 1), so those measurements were not performed in combination with lipase.

Hydrodynamic radius and thermal stability measurements

DLS measurements on the lipase in the presence of crowder and crowders alone were performed to see the effect of the crowders.

In Figure 5(A), shown plots are the DLS measurements of lipase and the crowders alone in 10 mM histidine, 70 mM NaCl at pH 5.0. In panel A, a size distribution profile for a concentration series measurement of lipase from 1 mg/ml to 25 mg/ml was shown over logarithmic R_h in the x-axis. Here, we see the relative scattered intensity by two different sized populations, the % intensity of R_h below 10 nm is higher for lipase concentration below 5 mg/ml. The % intensity for R_h above 10 nm is increasing with increasing concentration indicates presence of larger particles at higher lipase concentration.

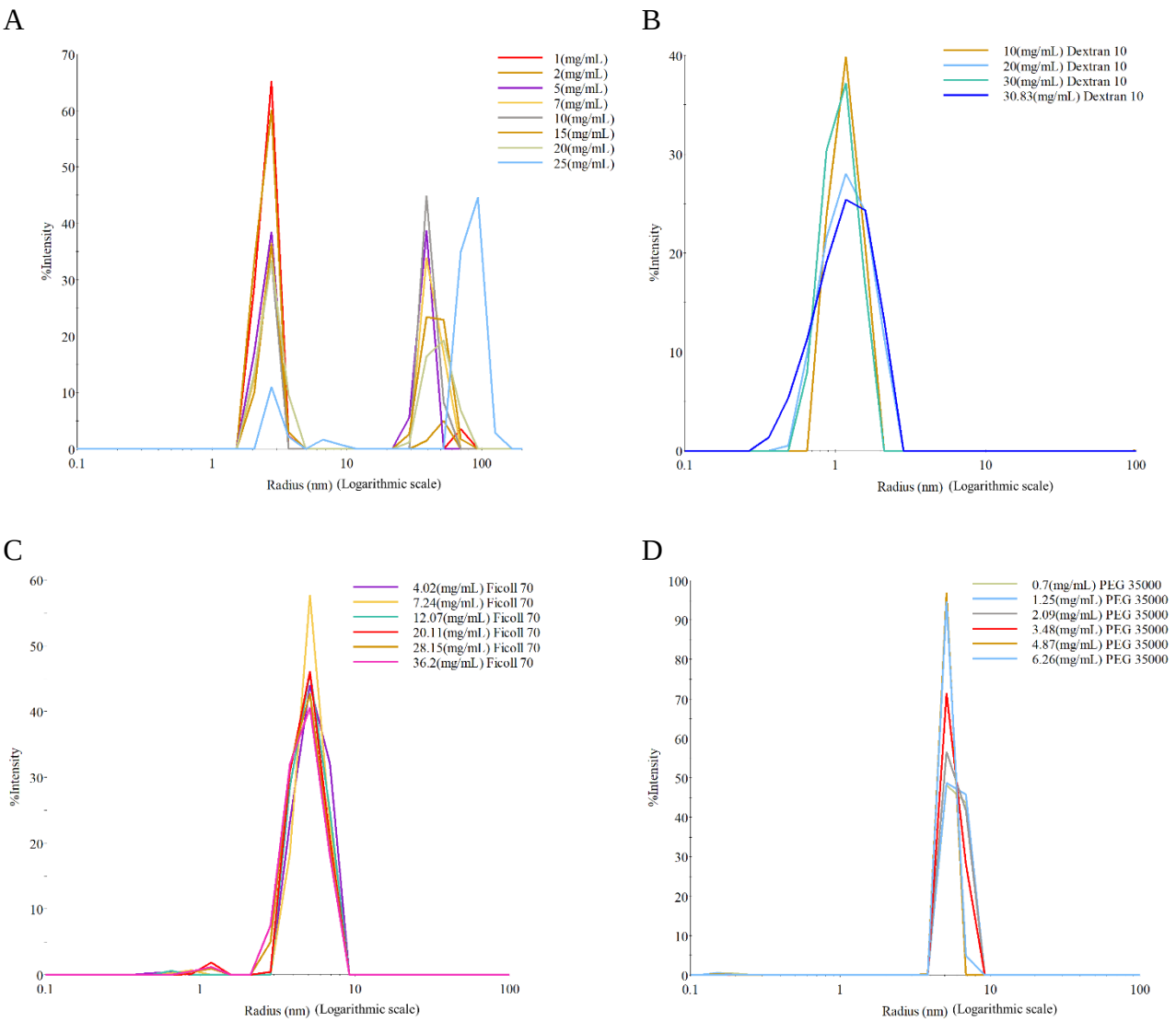


Figure 5 DLS measurements of concentration series of Lipase (panel A) and three crowders, Dextran, Ficoll, PEG in panel B, C and D respectively.

In Figure 5 (B, C and D), size distribution profiles for three individual crowders are shown at the calculated concentrations from Table 4. As the calculated concentration for dextran above the effective concentration limit except for 5 mg/ml condition, we perform an additional three measurements with 10 mg, 20 mg and 30 mg of dextran in the presence of lipase. We see that each crowder has one population and there were no larger particle formation occurs. Measured R_h of three crowders are shown in Table 5.

Table 5 Hydrodynamic radius, R_h, of crowders measured using DLS.

Dextran conc. (mg/ml)	R_h (nm)	Ficoll conc. (mg/ml)	R_h (nm)	PEG conc. (mg/ml)	R_h (nm)
10	1.56	4.02	4.76	0.70	5.60
20	1.42	7.24	4.73	1.25	5.57
30	1.41	12.07	4.62	2.09	5.51
30.83	1.21	20.11	4.41	3.48	5.27
		28.15	4.31	4.87	5.03
		36.20	4.14	6.26	4.80

Table 6 $T_{½}$ value at each given concentration of crowders in presence of 2.5 mg/ml of lipase.

Dextran (mg/ml)	Lipase (mg/ml)	$T_{½}$ (ºC)	Ficoll (mg/ml)	Lipase (mg/ml)	$T_{½}$ (ºC)	PEG (mg/ml)	Lipase (mg/ml)	$T_{½}$ (ºC)
0	2.5	60.2	0	2.5	60.2	0	2.5	60.2
10	2.5	59.7	4.02	2.5	60.1	0.70	2.5	60.4
20	2.5	59.6	7.24	2.5	60.1	1.25	2.5	60.1
30	2.5	59.6	12.07	2.5	60.2	2.09	2.5	60.2
30.83	2.5	59.6	20.11	2.5	60.2	3.48	2.5	60.3
			28.15	2.5	60.1	4.87	2.5	60.1
			36.20	2.5	60.2	6.26	2.5	60.3

In Table 6, obtained apparent melting temperature of lipase with crowders are shown. Data indicates that increase in concentrations of ficoll and PEG do not affect lipase's thermal stability as the $T_{½}$ remains constant. However, $T_{½}$ is decreasing by 0.5 ºC in presence of dextran, which is negligible. This means, the thermal stability was not affected by the presence of crowders.

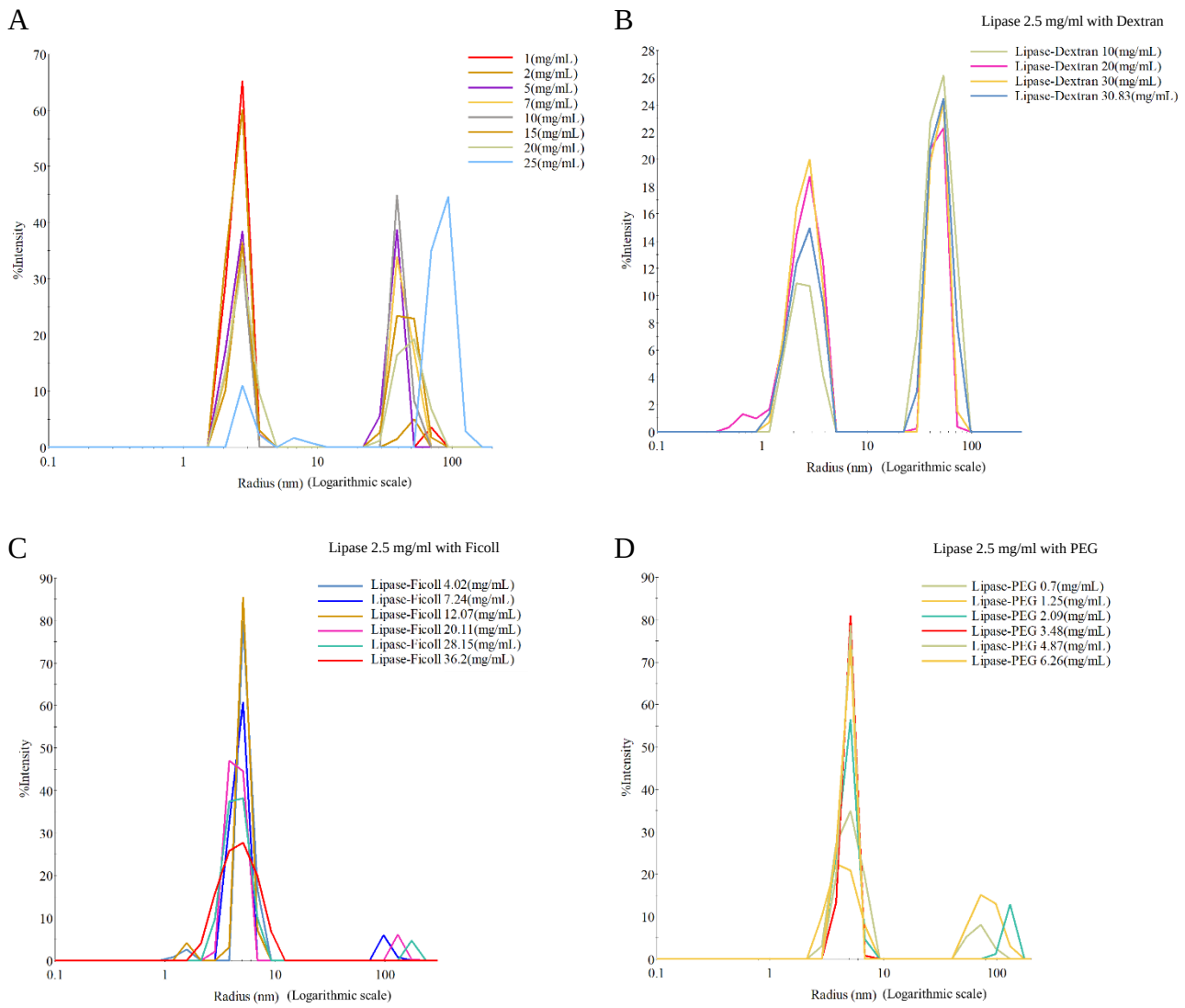


Figure 6 DLS measurements of three crowders with lipase are shown. Panel A: lipase in 10 mM histidine pH 5.5, 70 mM NaCl, Panel B: lipase with dextran in 10 mM histidine pH 5.5, 70 mM NaCl, Panel C: lipase with ficoll in 10 mM histidine pH 5.5, 70 mM NaCl, Panel D: lipase with PEG in 10 mM histidine pH 5.5, 70 mM NaCl.

In Figure 6, size distribution profiles of three crowders with lipase are shown. In the presence of dextran, we still see two populations, one of size smaller than 10 nm and size above 50 nm, similar to what we see with lipase alone in the solution, this could explain that crowders smaller than the test molecule do not prevent the large particle formation. However, in presence of ficoll and PEG, we see a higher percentage of intensity for the distribution profile below 10 nm and a very small amount of larger particles compared to the percentage we see with lipase alone in the crowded environment at concentration 5 mg/ml and above. This suggests that crowders larger than the test molecule, such as ficoll and PEG in presence of lipase may prevents the formation of larger particle (prevent molecules from self-associating in order to minimize the size of the oligomer). Moreover, this study could be complemented by circular dichroism spectroscopy study in order to explain the conformational changes of the lipase.

References

1. Chebotareva, N. A., Kurganov, B. I. & Livanova, N. B. Biochemical effects of molecular crowding. *Biochemistry (Moscow)* (2004). doi:10.1007/s10541-005-0070-y

2. HARVE, K. S., RAGHUNATH, M., LAREU, R. R. & RAJAGOPALAN, R. MACROMOLECULAR CROWDING IN BIOLOGICAL SYSTEMS: DYNAMIC LIGHT SCATTERING (DLS) TO QUANTIFY THE EXCLUDED VOLUME EFFECT (EVE). *Biophys. Rev. Lett.* (2006). doi:10.1142/s1793048006000215

3. Kuznetsova, I. M., Turoverov, K. K. & Uversky, V. N. What macromolecular crowding can do to a protein. *International Journal of Molecular Sciences* (2014). doi:10.3390/ijms151223090

4. Minton, A. P. The Influence of Macromolecular Crowding and Macromolecular Confinement on Biochemical Reactions in Physiological Media. *Journal of Biological Chemistry* (2001). doi:10.1074/jbc.R100005200

5. van den Berg, B. Effects of macromolecular crowding on protein folding and aggregation. *EMBO J.* (1999). doi:10.1093/emboj/18.24.6927

6. Armstrong, J. K., Wenby, R. B., Meiselman, H. J. & Fisher, T. C. The hydrodynamic radii of macromolecules and their effect on red blood cell aggregation. *Biophys. J.* (2004). doi:10.1529/biophysj.104.047746

7. Wang, Y., He, H. & Li, S. Effect of Ficoll 70 on thermal stability and structure of creatine kinase. *Biochem.* (2010). doi:10.1134/S0006297910050160

8. Sasahara, K., McPhie, P. & Minton, A. P. Effect of dextran on protein stability and conformation attributed to macromolecular crowding. *J. Mol. Biol.* (2003). doi:10.1016/S0022-2836(02)01443-2

9. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* (2003). doi:10.1093/nar/gkg563

10. Chen, C., Loe, F., Blocki, A., Peng, Y. & Raghunath, M. Applying macromolecular crowding to enhance extracellular matrix deposition and its remodeling in vitro for tissue engineering and cell-based therapies. *Advanced Drug Delivery Reviews* (2011). doi:10.1016/j.addr.2011.03.003