PROTEIN-PROTEIN INTERACTIONS AND AGGREGATION IN BIOTHERAPEUTICS

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy
in the Faculty of Engineering and Physical Sciences

2015

Mariam Masaud Nuhu

School of Chemical Engineering and Analytical Science, The University of Manchester

CONTENTS

Final Word Count: 49,849

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Arg	arginine
Asn	asparagine
Asp	aspartate
BCM	barycentric mean
BSA	bovine serum albumin
DLS	dynamic light scattering
DLVO	Derjaguin-Landau-Verwey-Overbeek
DS	dextran sulphate
FDA	Food and Drug Administration
Gln	glutamine
His	histidine
Ig	immunoglobulin
Lys	lysine
mAb	monoclonal antibody
MALS	multiangle light scattering
Met	methionine
PCS	photon correlation spectroscopy
PD	1, 2 -propanediol propylene glycol
PEI	polyethyleneimine
pI	isoelectric point
QELS	quasi-elastic light scattering
rhGCSF	recombinant human growth colony stimulating factor
RNase A	ribonuclease A
SLS	static light scattering
Trp	tryptophan

LIST OF SYMBOLS

ΔG_{unf}	free energy of unfolding
B_{22}	osmotic second virial coefficient
B_{33}	osmotic third virial coefficient
π	osmotic pressure
γ_p	activity coefficient
R	universal or ideal gas constant
M_w	molecular weight
c	protein mass concentration
μ	chemical potential
T	absolute temperature
R	intermolecular center-to-center distance
k_B	Boltzmann constant
N_A	Avogadro's constant
Q	net charge
e	elemental charge
d_p	protein diameter
z	valence charge
ε	relative permittivity of water
ε_0	dielectric permittivity of free space
κ	inverse of Debye length
$W(r)$	potential of mean force
R_θ	excess Rayleigh ratio
R_T	Rayleigh ratio of toluene
I_T	reference (toluene) scattering intensity
n_0	refractive indices of solvent

n_T	refractive indices of toluene
A_{cscc}	configuration specific calibration constant
A_{inst}	instrument dependent constant
K	optical constant
λ	wavelength of incident light
dn/dc	differential refractive index increment
$M_{w,app}$	apparent mass-average molecular weight
$M_{w,0}$	infinite dilution molecular weight
$g^{(2)}(\tau)$	intensity autocorrelation function
α	instrument constant
β	background term
D_T	translational diffusion coefficient
D_0	infinite dilution of diffusion coefficient
D_z	mean diffusion coefficient
q	scattering vector length
θ	scattering angle
Pd	polydispersity
k_D	inter-particle interaction parameter
k_f	frictional coefficient
v_{sp}	partial specific volume
μ_E	electrophoretic mobility
$f(\kappa R_h)$	Henry's function
ζ	zeta potential
T_m	melting temperature
T_{agg}	aggregation temperature
η	solution viscosity

Protein-Protein Interactions and Aggregation in Biotherapeutics

Mariam M. Nuhu, The University of Manchester, 2015

Submitted for the degree of Doctor of Philosophy

Abstract

Protein aggregation is a frequently cited problem during the development of liquid protein formulations, which is especially problematic since each protein exhibits different aggregation behaviour. Aggregation can be controlled by judicious choice of solution conditions, such as salt and buffer type and concentration, pH, and small molecule additives. However, finding conditions is still a trial and error process. In order to improve formulation development, a fundamental understanding of how excipients impact upon protein aggregation would significantly contribute to the development of stable protein therapeutics.

The underlying mechanisms that control effects of excipients on protein behaviour are poorly understood. This dissertation is directed at understanding how excipients alter the conformational and colloidal stability of proteins and the link to aggregation. This knowledge can be used for finding novel ways of either predicting or preventing/inhibiting protein aggregation.

Experiments using static and dynamic light scattering, intrinsic fluorescence, turbidity and electrophoretic light scattering were conducted to study the effect of solution conditions such as pH, salt type and concentration on protein aggregation behaviour for three model systems: lysozyme, insulin and a monoclonal antibody. Emphasis is placed on understanding the effects of solution additives on protein-protein interactions and the link to aggregation. This understanding has allowed the rational development of stable formulations with novel additives, such as arginine containing dipeptides and polycations.

DECLARATION

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Signature:

Name of author: Mariam M. Nuhu

Date: 22-07-2015

COPYRIGHT STATEMENT

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the "Copyright") and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the "Intellectual Property") and any reproductions of copyright works in the thesis, for example graphs and tables ("Reproductions"), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property
and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see
http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487), in any relevant Thesis restriction declarations deposited in the University Library, The University

Library's regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University's policy on Presentation of Theses.

ACKNOWLEDGEMENT

First of all, I am most grateful to the Almighty God all through the period of my academic pursuit. I would like to express my deepest gratitude to my doctoral advisor Dr Robin Curtis. I am very grateful to have worked with him on this challenging project. I also admire him for his extensive knowledge in the research field, and excellent teaching skills.

I greatly appreciate the opportunity to work with Dr Dorota Roberts, Dr Abdullatif Alfutimie, and Dr James Austerberry in different phases of this research work. I would like to thank Dr Tom Jowith for facilitating research and access to the Biomolecular analysis lab.

I am grateful to Prof Davor Kovacevic from the University of Zagreb for inviting me to a one-month stay in his group, and for his scientific input in many discussions and meetings. Special thanks also go to Dr Darija Jurasin for her excellent assistance in performing zeta potential measurements, and to the whole team in Zagreb for the nice and supportive atmosphere in the lab. Part of the research work was performed in collaboration with Arecor Ltd and Genzyme. Here, I would like to thank Dr Jan Jezek and Dr Barry Derham for providing the antibody samples and immeasurable help and insight into experimental details and analyses.

I would also like to thank my group members Joe, James and Rose for all the group discussions and lunches both inside and outside of the office or lab. I also thank my friends, particularly Eddy, Efe, Petrus, Blessing, Godwin and Mala for making my life so much more enjoyable at Manchester.

I would very much like to thank my family. Special gratitude goes to my parents, Alhaji and Dr (Mrs) Nuhu, and my brothers, Habib and Masaud Jnr., for their continuous love and support in all my decisions in life. I would also like to thank my

uncle Prof Abdulwaheed Olatinwo and his family, and my aunt Dr Ganiyat Oyeleke for their prayers and support throughout my program. Last but not least, a big thank you to my fiancé Lukman for his never-ending energy, optimism and support, especially towards the finishing of my thesis.

CHAPTER 1

1 INTRODUCTION

This chapter outlines the motivation of the thesis and its structure. The main aim of the work is to gain a better understanding of how excipients affect protein-protein interactions and protein aggregation during the formulation of liquid protein therapeutics. Although the knowledge on protein-protein interactions and aggregation is increasing, a major challenge is optimising formulation conditions to reduce protein aggregation and ensure 18-24 months shelf life. Information on how solution conditions change protein-protein interactions can provide ways of predicting aggregation during drug formulation and storage. This chapter starts with a general overview on challenges faced during development and manufacture of liquid formulations which serves as a build up to the motivation of the work highlighted in section 1.1. Protein aggregation, a commonly encountered problem in each step of development process is presented and its mitigation is discussed along with the link to protein- protein interactions. Finally, the content of the remainder of the report are outlined in section 1.2.

1.1 Motivation

Protein aggregation phenomenon is a major impediment to the production and development of stable protein therapeutics. Aggregation has been observed at all protein concentrations during various bioprocessing steps such as refolding, fermentation, freeze-thawing, reconstitution and also during transport and storage of the protein drugs[1]. For example, the presence of insoluble protein aggregates during recombinant protein expression leads to the formation of inclusion bodies which result in lower protein yields [2, 3]. Moreover, protein aggregation has been implicated in a number of different diseases and disorders arising from changes to

the physical and/or chemical stabilities of proteins. In recent years, the pharmaceutical industry has seen a significant shift towards the development of protein-based therapeutics for the treatment of various diseases and disorders. Subcutaneous or inter-muscular routes of administration are becomingly increasing preferred due to limitations in protein half-life, increasing target coverage, as well as patience convenience which necessitate the need for therapeutics to be formulated at typically high concentrations exceeding 100 mg/mL [4-6]. Development of stable high concentration liquid formulations poses several challenges particularly in relation to solubility, viscosity, phase separation and protein aggregation arising from protein self-interactions and self-association. At high protein concentrations, the volume fraction occupied by the solute is greater than typically present under dilute solution conditions. Deviation from an ideal solution behaviour due to increased protein association can lead to opalescence or phase separation, high viscosities and ultimately aggregation [7, 8]. The presence of aggregates in particular is undesirable as the aggregates often lack bioactivity and may be immunogenic within host patients upon administration thus reducing the safety and efficacy of the protein therapeutic. Therefore, developing a better understanding of how solution conditions change protein-protein interactions and protein aggregation is critical not only in various manufacturing processes but also in the development of stable protein formulations.

Although significant progress has been made to control or inhibit aggregation by manipulating solution conditions, the results are not fully satisfactory for many proteins. This is often due to lack of clear understanding of the process of protein aggregation and underlying mechanisms. As a result this thesis intends to understand how excipients change protein conformational and colloidal stability and the link to

aggregation. This is to be achieved by investigating the effect of the solvent environment on protein-protein interactions which may partly be a pre-requisite to protein aggregation. Detailed analysis and description of various analytical techniques employed to provide in-depth characterisation of protein-protein interactions is discussed later on in the thesis.

1.1.1 Driving force of Aggregation

For many years, protein aggregation has been extensively studied and has continued to gain more light in many areas of scientific and technological research. There is no consistent meaning for the term 'aggregation' as different working definitions are often used. This thesis defines aggregation as *the irreversible formation of high molecular weight aggregates from natively folded proteins that have non-native structures*. Although native proteins have been found to self-associate and precipitate even under physiological, non-denaturing conditions, this process is usually reversible [9].

One of the greatest challenges to understanding and controlling protein aggregation is that there is no single pathway to which aggregates form. Protein aggregates can occur through a number of different mechanisms. Aggregation occurs through at least a two step-process including (1) structural changes to the native state as the native protein partially unfolds and (2) association with other protein molecules to form reversible or irreversible higher-order species [1]. Step (1) is controlled by the conformational stability or the protein's ability to retain the native three-dimensional structure. The second step is however dependent on the nature of protein-protein interactions governed by a range of electrostatic and/or hydrophobic interactions [10] termed colloidal stability (or the protein's ability in solution to resist the formation of aggregates). Figure 1.1 shows a schematic of the aggregation process whereby native

proteins (*N*) form reversible partially unfolded intermediates (*I*) which can lead to totally unfolded (*U*) proteins or irreversible aggregates (*A*). This reflects different aggregation pathways depending on the rate limiting step (*k*) to aggregation which depends on solution conditions and temperature.

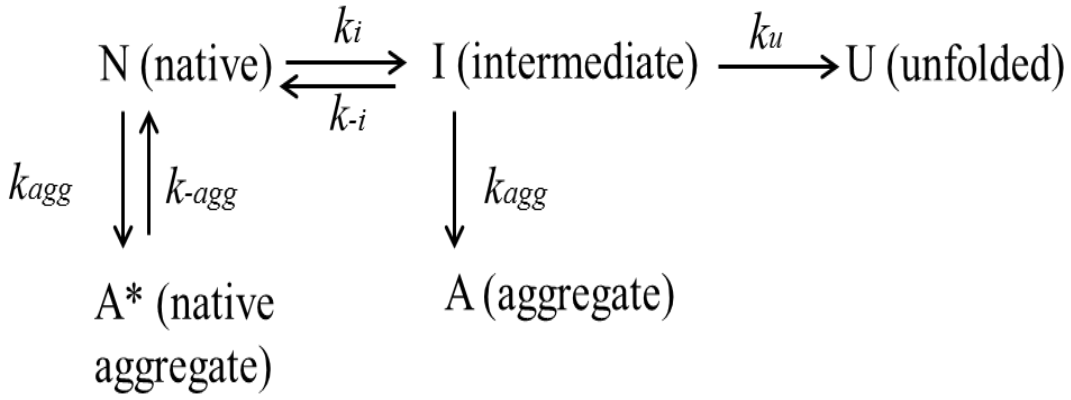


Figure 1.1. Schematic of physical protein aggregation pathways

In most cases, the rate limiting step of the aggregation process occurs through the formation and association steps of the partially folded intermediates. Most methods for inhibiting aggregation target this step by addition of appropriate excipients. For example, arginine improves the refolding yield of some proteins by either stabilizing the native state, solubilising folded intermediates or destabilising incorrectly folded intermediates [11, 12]. The overall outcome of the aggregation process is in part controlled by weak protein-protein interactions therefore, predicting the effect of solution conditions (such as pH, temperature, salt type and concentration) requires elucidating how the solvent environment changes protein-protein interactions.

A number of techniques are used for assessing the conformational and colloidal stability of proteins which can be measured experimentally through the free energy

of unfolding (ΔG_{unf}) and osmotic second virial coefficient (B_{22}) respectively. ΔG_{unf} is defined as the difference in free energy between the native and denatured states as protein stability increases under processes that favour increasing ΔG_{unf} values. B_{22} is a measure of protein-protein interactions in solution where negative B_{22} values reflect attractive protein-protein interactions (or colloidally unstable proteins) while positive values indicate repulsive protein-protein interactions. Figure 1.2 shows a pictorial representation of how protein aggregation process described in Figure 1.1 is controlled by conformational (structural) changes and/ or colloidal changes initiated by assembly processes.

Aggregation proceeds due to structural changes of the native state to an intermediate or transition state species (TS^*), which can be characterized in terms of measured in terms of ΔG_{unf} [13]. Propensity to form further assembly reactions leading to non-native aggregates can be related to values of B_{22}. In another pathway, aggregation occurs due to weak interactions between the native proteins leading to the formation of an aggregation—competent intermediate. This species then undergoes assembly reactions leading to the formation of large aggregates [13].

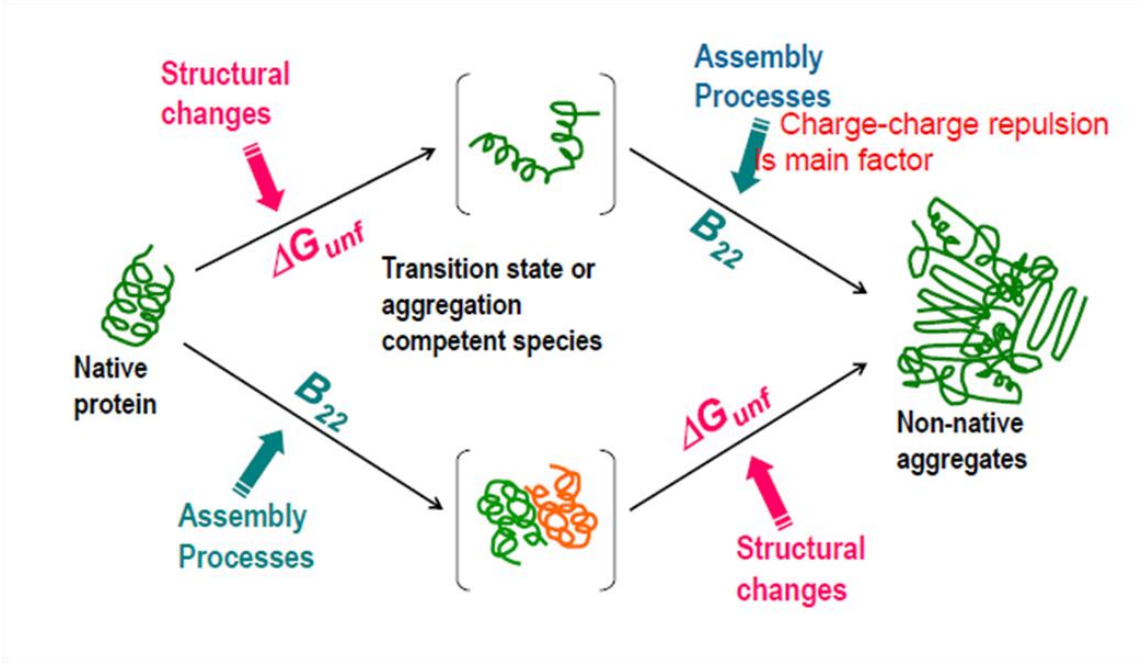


Figure 1.2. Role of conformational and colloidal stability in protein aggregation [13, 14]. An increase in ΔG_{unf} favours increased conformational stability of the native state to prevent formation of aggregation-prone species while decrease in ΔG_{unf} favours reduced conformational stability leading to the formation of structurally altered or aggregation-prone species. A positive B_{22} indicates increased colloidal stability which prevents further assembly reactions that can lead to aggregation whereas a negative B_{22} increases agrgegation propensity due to attractive protein-protein interactions.

It is important to recall that protein aggregation refers to formation of irreversible non-native aggregates which proceeds via aggregation-competent intermediate states. Hence when the sticky hydrophobic residues which are usually buried in the native state become exposed, unfolded or partially unfolded states become prone to aggregation. Thus the most dominant factor underlying protein aggregation is thought to be ‘‘hydrophobic clustering’’. Therefore aggregation is thought to be

predominantly controlled by the initial unfolding of the protein. However, aggregation upon storage is not always linked to conformational stability. Some studies have shown that the colloidal stability or protein-protein interactions can be a controlling factor where attractive protein-protein interactions can lead to aggregation [10, 13]. This indicates that conformational stability is not the only rate controlling step to aggregation. In addition, irreversible protein aggregation formed between partially unfolded proteins is much stronger than the weak self-associations formed between native proteins.

This work focuses on formulation development which is one of the most critical steps in development of protein therapeutics. Realising that it is impossible to create a formulation with perfect stability due to the complexity of the protein structure, the most logical step is to develop predictive ability to control protein aggregation in liquid formulations by understanding how protein aggregation is controlled by protein-protein interactions. This could lead to improved methods for screening solvent conditions against long term storage stability.

1.2 Thesis overview and objectives

This dissertation is presented in the alternative format, as allowed by The University of Manchester thesis submission guidelines. Each of these chapters begins with a providing a general overview of the chapter and has been clearly marked.

The research work was carried out to investigate the role of solvent and co-solvent environment on protein-protein interactions and ultimately protein aggregation. A key aspect is understanding the underlying molecular forces controlling protein solution behaviour. Various analytical techniques were utilised for characterising

protein behaviour at relatively dilute protein concentrations. The thesis is organised into 7 chapters.

Chapter 2 provides relevant background theory to protein stability, protein-protein interactions and explains how link to protein aggregation.

In chapter 3, the analytical characterization methods used in carrying out experiments for studying protein –protein interactions are described in detail. Theoretical descriptions including relevant equations as well as experimental methodology used are discussed.

In the first experimental study presented in Chapter 4, a combination of static and dynamic light scattering and zeta potential analysis is used to investigate protein intermolecular interactions. The main aim of this chapter is to elucidate specific ion effects on protein-protein interactions. Here we examined protein-protein interactions and protein-ion interactions using lysozyme as a model system. Measurements are made for a range of ionic strength, pH values and salt types. In addition, we implement a recently developed rapid screening method for measuring protein-protein interactions. Zeta potential measurements were carried out in collaboration with Professor Davor Kovacevic at the University of Zagreb, Croatia. A manuscript of this work is being prepared for journal publication.

In chapter 5, we establish the effectiveness of a novel set of additives containing arginine (dipeptides) in controlling the association and aggregation behaviour of insulin. The work presented in this chapter was published in the Biotechnology Journal volume 10 (2015) pages 404-416 and titled ‘Arginine dipeptides affect insulin aggregation in a pH- and ionic strength-dependent manner’. The study also shows the applicability of using dipeptides as novel excipients which can be extended to improving stability of other proteins.

Chapter 6 presents a study where a combination of static and dynamic light scattering and intrinsic fluorescence measurements are used to assess the aggregation behaviour and the stability characteristics of a monoclonal antibody (IgG4) under a range of operating conditions in the presence of a novel cationic excipient. The work was performed in collaboration with Arecor Ltd and Genzyme. Measurements are made as a function of pH and ionic strength in order to understand the intermolecular forces controlling the formulation stability.

Chapter 7 summarizes all of the research work carried out in this thesis and brings together the main conclusions from each of the result chapters. Contributions to relevant literature in the field of protein formulation and stability are included. Finally, potential areas for future work are suggested.

1.3 References

[1] Wang, W., Protein aggregation and its inhibition in biopharmaceutics. *International Journal of Pharmaceutics* 2005, *289*, 1-30.

[2] Finke, J. M., Roy, M., Zimm, B. H., Jennings, P. A., Aggregation Events Occur Prior to Stable Intermediate Formation during Refolding of Interleukin 1Î²â€ *Biochemistry* 1999, *39*, 575-583.

[3] De Young, L. R., Fink, A. L., Dill, K. A., Aggregation of globular proteins. *Accounts of Chemical Research* 1993, *26*, 614-620.

[4] Shire, S. J., Shahrokh, Z., Liu, J., Challenges in the development of high protein concentration formulations. *Journal of Pharmaceutical Sciences* 2004, *93*, 1390-1402.

[5] Vineet Kumara, N. D., Liqiang (Lisa) Zhou, Wolfgang Fraunhofer, Impact of short range hydrophobic interactions and long range electrostatic forces on the aggregation kinetics of a monoclonal antibody and a dual-variable domain immunoglobulin at low and high concentrations. *International Journal of Pharmaceutics* 2011, *421*, 82-93.

[6] Samra, H. S., He, F., Advancements in High Throughput Biophysical Technologies: Applications for Characterization and Screening during Early Formulation Development of Monoclonal Antibodies. *Molecular Pharmaceutics* 2012, *9*, 696-707.

[7] Hall, D., Minton, A. P., Macromolecular crowding: qualitative and semiquantitative successes, quantitative challenges. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* 2003, *1649*, 127-139.

[8] Saluja, A., Kalonia, D. S., Nature and consequences of proteinâ€“protein interactions in high protein concentration solutions. *International Journal of Pharmaceutics* 2008, *358*, 1-15.

[9] Minton, A. P., Influence of macromolecular crowding upon the stability and state of association of proteins: predictions and observations. *J Pharm Sci.* 2005, *94*, 1668-1675.

[10] Chi, E. Y., Krishnan, S., Kendrick, B. S., Chang, B. S., *et al.*, Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Science* 2003, *12*, 903-913.

[11] Arakawa, T., and Tsumoto, K., The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation. *Biochemical and Biophysical Research Communications* 2003, *304*, 148-152.

[12] Tsumoto, K., Ejima, D., Kumagaia, I., Arakawa T., Practical considerations in refolding proteins from inclusion bodies. *Protein Expression and Purification* 2003, *28*, 1-8.

[13] 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, *20*, 1325-1336.

[14] Satish, S., BioProcess International European Summit, Czech Republic 2014.

CHAPTER 2

2 PROTEIN AGGREGATION AND PROTEIN-PROTEIN INTERACTIONS

2.1 Introduction

This chapter examines our current understanding of factors influencing protein conformational and colloidal stability (i.e. protein-protein interactions) and the link to protein aggregation. Firstly, we discuss how protein degradation leads to a loss in physical and chemical stability (section 2.2-2.3). We focus on protein aggregation which is a form of physical instability common to most therapeutic proteins and discuss how temperature and solution conditions such as pH, salt type and concentration may affect the aggregation process (section 2.4). Understanding the link between aggregation and protein-protein interactions requires measuring interaction parameters such as the osmotic second virial coefficient which is often a reliable way of interpreting colloidal stability or aggregation propensity, and accounts for exclude volume effects, electrostatic and short ranged interactions such as van der Waal, and hydrophobic interactions (section 2.5).

2.2 Proteins

Proteins are complex macromolecules whose three-dimensional conformation is determined by the primary, secondary, tertiary and quaternary structures. The primary structure is composed of about 20 amino acids (nonpolar, polar, acidic, basic, or aromatic depending on its side chain), containing both amino and carboxylic functional groups covalently linked through peptide bonds in the polypeptide chain [1]. Specific pair residues between the amide protons (N-H) and carbonyl groups (C=O) of the peptide bonds are responsible for the formation of

intra/inter-residue hydrogen bonding, which stabilises the secondary structure. The secondary structure consists of alpha helix, beta sheet, beta turns, or unstructured (random coil) conformations. The tertiary structure refers to the overall 3-dimensional folding of the protein, which is stabilized by interactions between the side chains of the residues located far apart in the linear amino acid sequence. The folding of a protein is stabilised by interactions such as hydrogen bonds, ionic and electrostatic interactions, van der Waals interactions, disulphide bonds, etc. as shown in Figure 2.1. The quaternary structure refers to the association of two or more polypeptide chains or subunits, and is mainly stabilized by weak interactions between the polypeptides. Examples of proteins with quaternary structures include multidomain proteins such as antibodies and haemoglobin. The overall conformational structure of a protein is a critical factor for maintaining its biological activity and long-term storage stability.

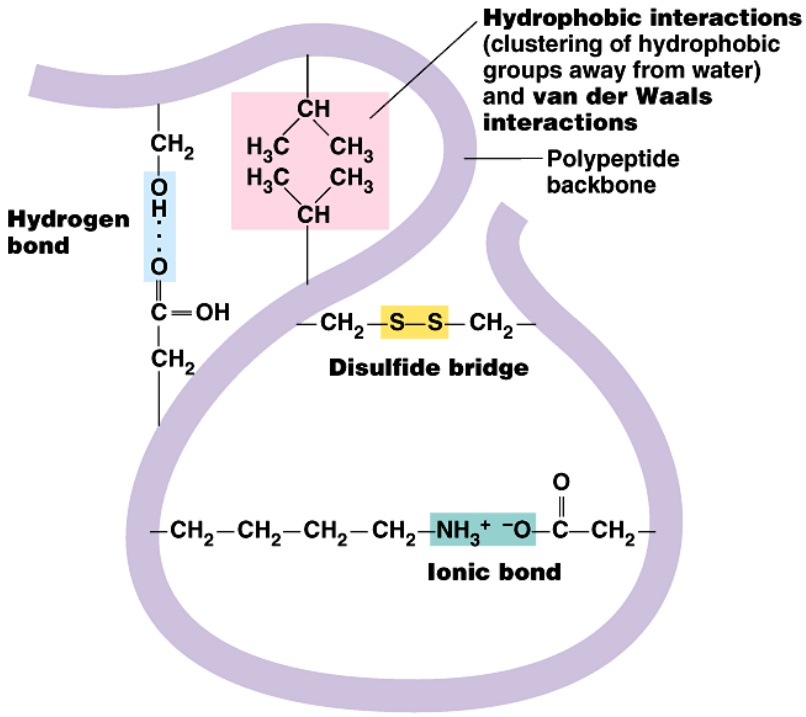


Figure 2.1. Interactions involved in stabilising the tertiary structure

2.3 Protein Stability

In the biotechnology and biopharmaceutical industry, commercial production of proteins is made possible through the advent of recombinant DNA technology. As the number of protein therapeutics on the market increases, the relevance of protein stability will continue to gain more light in the pharmaceutical field. Currently in the United States alone, about 150 protein-based products have received FDA (Food and Drug Administration) approval in addition to 20 antibody products [2]. However, difficulties still exists in commercial viability of proteins due to various instabilities. This poses a challenge during protein purification, separation/isolation, storage and delivery within manufacturing and development. Achieving formulation stability is particularly difficult; proteins are only marginally stable, but formulations need to possess a shelf life of 18-24 months.

Protein degradation arises as a result of either chemical or physical instability which both affects the solubility, immunogenicity and toxicity of the final protein drug. Chemical instability involves processes that modify or alter the chemical structure of proteins via covalent bond formation or cleavage, resulting in a new chemical entity. Examples of reactions linked to chemical instability in therapeutic proteins include asparagine (Asn) or glutamine (Gln) deamidation, aspartate (Asp) isomerisation, methionine (Met) or tryptophan (Trp) oxidation, proteolysis and disulphide bond formation/exchange as shown in Table 2.1. On the other hand, physical instabilities do not involve covalent bond modification as the chemical composition is not affected but the physical state of the proteins is altered. These changes occur as a result of denaturation or protein unfolding, precipitation, aggregation and surface adsorption.

Table 2.1. List of some common protein instabilities

Physical	**Chemical**
• Conformational (thermal, cold, pressure, and chemical denaturation) • Colloidal (aggregation, self-association) • Surface Adsorption	• Asp Isomerization • Asn or Gln Deamidation • Oxidation (metal-catalysed, photo-oxidation or Met /Trp oxidation) • Disulphide bond formation/exchange • Proteolysis

We focus on physical instabilities that occur in proteins often leading to aggregation, which is a severe issue faced during the development of stable protein drug formulation. Physical instability refers to any change in the physical state of proteins in which the chemical composition is retained and not altered in any way. Examples of physical instabilities include denaturation, precipitation and aggregation. Denaturation corresponds to the loss of 3-dimesional structure of natively folded globular proteins into a disordered state. Generally, proteins are stable at temperatures that fall within their maximum free energy change of unfolding, ΔG_{unf} (8-20 kcal/mol). An increase or decrease outside this temperature range can destabilise and denature proteins. Heat-induced denaturation could either be reversible or irreversible. During reversible denaturation, unfolded proteins can regain native state by lowering temperature after heating above the protein melting temperature. Conversely, irreversible denaturation does not allow unfolded proteins to regain their native structure when temperature is lowered. At temperatures below

freezing point of water, proteins have been reported to exhibit cold denaturation [3].Chemical denaturation due to addition of chaotropes, such as urea and guanidine hydrochloride (GdnHCl), also results in loss of protein globular structure. Denaturants function by interacting with unfolded states, but investigations are on-going to understand the intermolecular forces that occur between proteins and denaturants [4, 5].

In this dissertation, aggregation corresponds to the formation of large molecular weight aggregates usually made irreversible through protein conformational changes. Precipitation, on the other hand, corresponds to the formation of an insoluble phase or precipitate. The precipitate could either be reversibly formed as occurs when native proteins self-associate leading to crystallization, or liquid-liquid phase separation. Or alternatively, irreversible aggregation can also lead to the formation or large precipitates.

2.4 Protein Aggregation

Aggregation is the most common manifestation of physical instability. The term ‘aggregation’ involves a number of different types of molecular assemblies, ranging from physical association of native monomers to chemical aggregation resulting in formation of a new chemical bond. Either of these mechanisms can produce soluble or insoluble aggregates depending on environmental conditions and the stage of aggregation process. For instance, insoluble aggregates are formed when the hydrophobic residues of proteins become exposed as occurs in thermal denaturation. Protein aggregation can be affected by external factors (temperature, and solution conditions such as pH, ionic strength, buffer, excipients) or structural factors, which

may destabilize the native state, favour formation of intermediate or aggregate-competent species or stabilise the unfolded states.

2.4.1 Structural factors

Protein aggregation may arise from non-covalent interactions (hydrogen bonding, hydrophobic interactions, van der Waal forces, electrostatic interactions), or interactions between covalently linked species (intrinsic or local /non-local peptide interactions) [6]. The amino acid sequence of proteins is believed to ultimately determine the aggregation propensity and is influenced by the environmental conditions [7]. Calamai et al. [8] showed that the hydrophobicity of proteins influences the likelihood of aggregation as a more hydrophobic protein has greater tendency to aggregate. This view is shared with Fields et al. [9] where increasing the non-polar component of the protein increased the aggregation propensity. Aggregation arises from partially unfolded or structurally-altered proteins that have exposed hydrophobic groups which form intermolecular associations with each other. Further studies indicate that the secondary structures containing substantial amount of β-sheet content are more likely to form aggregates than α-helices. There is an observed transition from the α-helix to β-sheet conformation, followed by aggregation as a result of intermolecular associations [10, 11].

2.4.2 External factors

Temperature

Above certain temperatures, proteins unfold and irreversibly form aggregates when the hydrophobic residues are exposed. The extent of reversibility is dependent on the stage and size of protein aggregates. As the temperature is raised, the frequency of protein diffusion and the strength of hydrophobic interactions are increased thus

affecting the rates of protein aggregation. A number of proteins have been found to aggregate at temperatures slightly above room temperature, for example recombinant human keratinocyte growth factor (rhKGF) at pH 7 [12], tubulin at pH 2.5 [13], and bovine insulin [14]. The thermodynamic folding stability of proteins is characterised by the free energy change of unfolding ΔG_{unf}. High values of ΔG_{unf} indicate a high probability of the protein to be in their native state and vice versa. With increasing temperature, the activation free energy or thermodynamic barrier between the native state (N) and partially folded state is reduced leading to higher concentrations of partially folded or unfolded proteins, which act as aggregation prone intermediates (A_n) before forming irreversible aggregates (A_m) as shown in Figure 2.2.

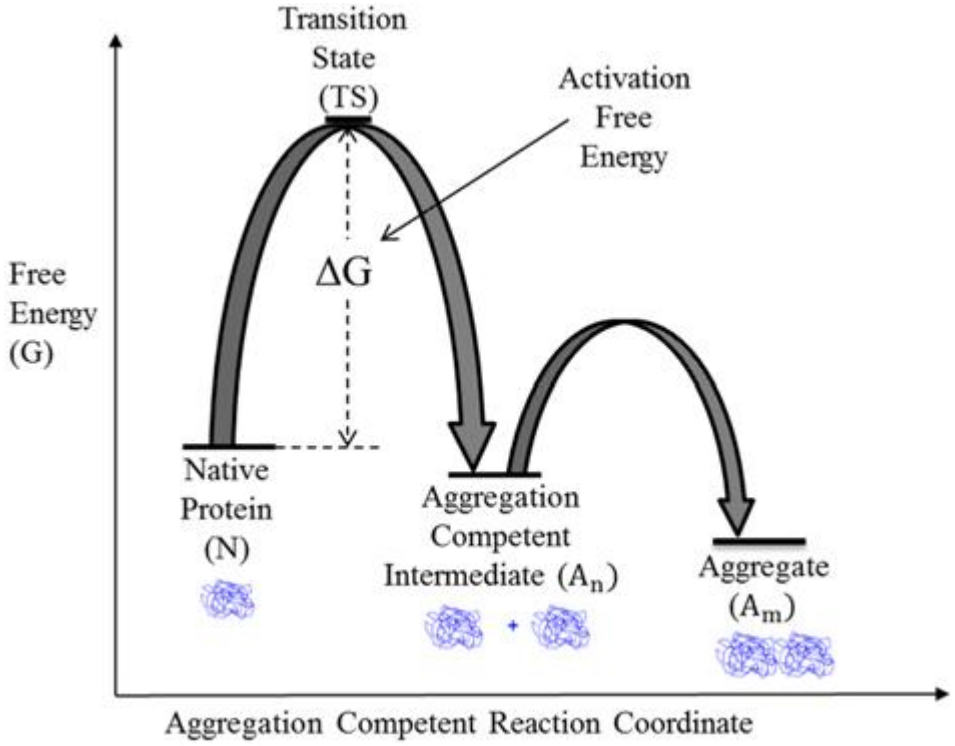


Figure 2.2. Schematic representation of Lumry-Eyring framework [6, 15]. Curved lines represent kinetic energy barrier

Solution pH

The pH of a solution has been shown to affect the rate of aggregation of proteins. pH determines the overall distribution and type of charges on protein surfaces, thus influencing electrostatic interactions. A number of studies have reported the influence of solution pH on aggregation of proteins such as ovalbumin at pH 4.0 [16], recombinant human growth colony stimulating factor (rhGCSF) [17], and bovine serum albumin (BSA) [18]. At pH values far below or far above the isoelectric point (pI), proteins possess highly positive or negative charges, respectively. A large net charge increases nonspecific charge repulsion thereby preventing aggregation. However, a large charge density on the protein causes a lower folding stability as the intramolecular electrostatic repulsion is greater in the compact folded state versus the extended unfolded state. Thus, extreme pH values can also cause proteins to unfold and expose hydrophobic groups leading to aggregation. At pH values close to pI, there is uneven or partial distribution of positive and negative charges on the protein surface which reduces charge-charge repulsion and can even lead to electrostatic association, which in turn, favors aggregation.Giger et al. [19]studied the aggregation behaviour of insulin over a range of pH where maximum aggregation was observed at pH 5.6 close to its pI of 5.5 in low ionic strength solutions.

Buffer Type

Due to the stability of proteins over a narrow range of pH, buffering agents are often added to maintain an optimum pH of the solution. The aggregation behaviour of proteins differ under various buffering conditions, for instance, Katayama et al. [20] studied the effect of different buffer systems on the aggregation behaviour of interferon-tau (IFN-tau) and found that phosphate buffers increased protein

aggregation rates when compared with tris or histidine buffers. Phosphate buffers bind weakly to the native state and increases aggregation by reducing colloidal stability arising from repulsive electrostatic forces between proteins.

Protein Concentration

Protein concentration is another factor that has been shown to affect protein aggregation propensity. Protein denaturation depends weakly on protein concentration. However, increasing protein concentration increases the tendency for protein aggregation and precipitation as a result of greater intermolecular interactions between proteins. Examples include aggregation/fibrillation of insulin [21], apomyoglobin [22], and GCSF [23]. In some cases, aggregation occurs via formation of an initial nucleus only when proteins exceed a critical concentration limit. Kanai et al. [24] and Liu et al. [25] observed an increase in viscosity and self –association with increasing the concentration of proteins.

Aggregation rates typically have a second or higher order dependence on protein concentration. This rate behaviour is quite different from the transition from native state (N) to molten globule (MG) and from MG to the unfolded state (U) in which case the rate of reaction is first order in protein concentration [26]. The higher reaction order of aggregation indicates that interactions between proteins in part control the aggregation behaviour.

Excipients and additives

The aggregation behaviour of proteins is also influenced by the types of co-solvents (often referred to as excipients or additives). Depending on their mechanism of interaction, co-solvents can either stabilise proteins against aggregation or destabilise proteins and promote aggregation. Examples of excipients/additives include salts,

sugars, polyols, amino acids, surfactants, polymers, preservatives and denaturants. The thermodynamic effects of co-solvents on protein stabilisation and destabilisation are determined by a protein- co-solvent preferential interaction parameter. Preferential interaction parameters are equal to the excess or deficiency of the co-solvent within the domain of the solute (protein) molecule relative to the bulk solution, which is referred to as either preferential binding or preferential exclusion, respectively. The stabilising effect of these excipients/additives corresponds with their ability to increase the free energy of unfolding, which is related to the preferential interaction parameter as demonstrated in Figure 2.3. An increase in free energy of unfolding with increasing additive concentration implies that the free energy of the native state is reduced relative to the unfolded state, which, for instance, could result from preferential binding of co-solvent (such as certain ligands) only to the native state (state A in Figure 2.3). Alternatively, the free energy of the unfolded state relative to the folded state is increased through preferential exclusion of co-solvents such as sugars (state B in Figure 2.3). The exclusion is greater about the unfolded state than the folded state due to the larger solvent exposed surface area of the unfolded state.

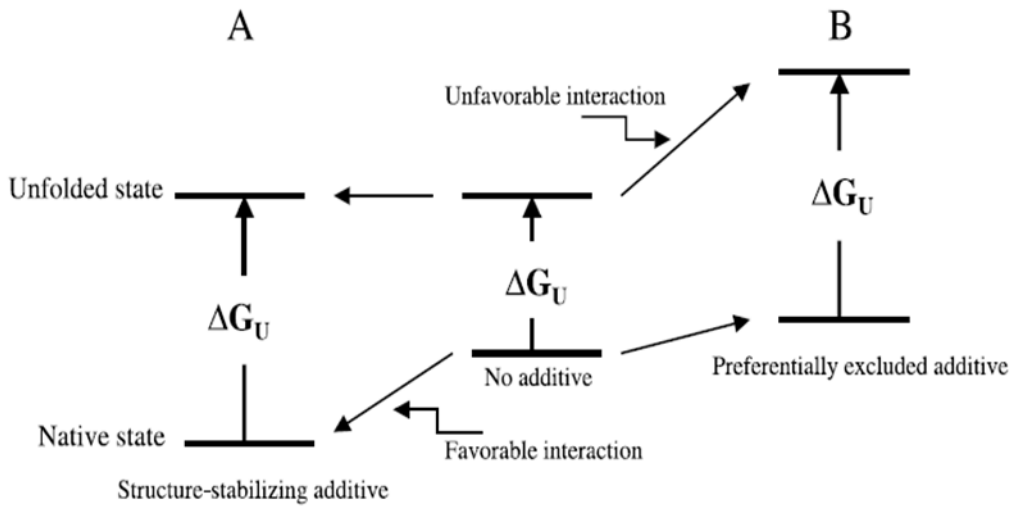


Figure 2.3. Effect of preferential interactions and ΔG_{unf} on folded and unfolded states [27]

Sugars and polyols have been found to stabilise proteins and inhibit aggregation under different conditions [28-30]. For example, the aggregation propensities of IgG1 antibodies at 40 °C were reduced in sucrose solutions by increasing conformational stability of the proteins [28]. The mechanism of stabilisation is most likely due to the exclusion of sucrose from the protein surface such that water molecules are localised around the protein and increases the kinetic barrier between native and aggregation-prone states. Ruzza et al. [31]demonstrated that the addition of 10 mM trehalose stabilised α-synuclein against aggregation by shifting the native state towards a more compact conformation.

Certain ligands such as metal ions can selectively bind to the native state of proteins and confer net stabilization and minimize protein aggregation at high concentrations. For instance, zinc ions are added to hexameric insulin to maintain the conformational stability [32]. At the same time, ligands can cause protein destabilisation by preferentially binding to the unfolded or denatured states of the proteins [2]. The

difference between the mechanism of stabilisation by excluded co-solvents such as sugars and ligands is that co-solvent exclusion is non-specific and occurs with all proteins, whereas ligand binding is protein specific [27]. Cimmperman et al. [33] showed that ligands such as sulphonamide, 1,8-anilinonaphthalene sulfonate (ANS) and zinc ions can selectively bind to native or denatured states to confer either stabilisation or destabilisation to different proteins.

In order to ensure sterility of protein formulations over their shelf life, antimicrobial preservatives such as benzyl alcohol and m-cresol are often added. Effective preservatives are however needed especially in multidose formulations where microbial growth needs to be prevented after a dose has been removed from its vial. Preservatives could be used in injection pens, topical applications and mini pumps. Caution must be taken as the preservatives can induce protein aggregation as not all proteins are compatible with preservatives. Studies suggest that the preservatives bind to the unfolded protein states prone to aggregation, although the mechanism is not well understood. Lam et al. [34] and Zhang et al. [35] reported that benzyl alcohol induced aggregation of recombinant human interferon-γ (rhIFN-γ) by populating aggregation-competent species without significant changes to the secondary structure.

Amino acids such as alanine, proline, lysine, betaine and other amino acid derivatives are often used as additives in inhibiting protein aggregation. Some of these are found in organisms living in extreme environments whereby they stabilise proteins by working as osmolytes or ‘compatible solutes’ in regions of high salinity and extreme temperatures, and do not affect molecular structure and function of proteins [36]. Unlike other stabilising amino acids which are preferentially excluded from the protein surface, arginine preferentially binds to the surface of protein, but

does not destabilise proteins [37].The interactions between arginine and proteins are weak, making arginine effective at high concentrations. Studies have suggested that arginine prevents aggregation by suppressing intermolecular interactions between aggregation –prone species [37, 38].

Salt type and concentration are important solution parameters affecting protein aggregation. Increasing salt concentration weakens the repulsive intermolecular interactions due to favourable interactions between the charges on the protein and salt leading to enhanced protein-protein interactions which results in aggregation. Sahin et al. [39] found that increasing ionic strength of the IgG1 solutions increased aggregation propensity by lowering the repulsive intermolecular interactions. The effect of salts on physical stability is complex as they can alter conformational stability, solubility (salting-in/ out), and aggregate formation. Intramolecular charge-charge interactions affect conformational stability while intermolecular electrostatic interactions affect aggregation and solubility [15].

Protein-salt interactions can be explained in terms of surface tension effects, electrostatic binding (non-specific and specific) and interactions with exposed polar (peptide) groups. Cavity theory relates surface tension and co-solvent exclusion whereby the stabilisation and association of protein molecules is a function of the solvation free energy or work required to create large cavities in a solvent to accommodate proteins. Preferential exclusion is often related to the ability of the co-solvent to increase the surface tension of water indicating the co-solvent exclusion about the protein surface is due to unfavourable interactions between the low dielectric constant of the protein interior and the co-solvents [40]. The extent of interactions between salts and proteins depends on the salt molal surface increment of the salt [41] which is related to the ion position in an empirical series such as the

Hofmeister or lyotropic series. Hofmeister [42] proposed a series for ranking salts in order of salting-out effectiveness for globular proteins now called the Hofmeister series as depicted in Figure 2.4.

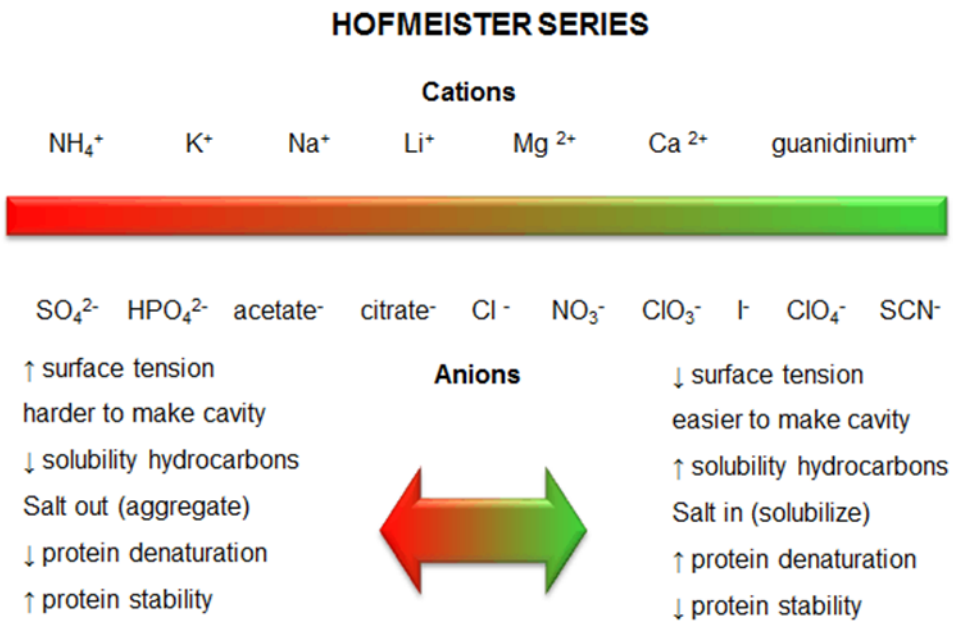


Figure 2.4. Cations and Anions ranking in Hofmeister series [43]

The Hofmeister series is similar to the lyotropic series, which ranks ions according to their hydration. Exceptions occur for multivalent cations which are high on the lyotropic series because divalent cations possess large hydration numbers, but low on the Hofmeister series as divalent cations are weak salting-out ions [44]. For anions, the molal surface tension increment decreases in the order of $SO_4^{2-} > HPO_4^{2-} > F^- > Cl^- > Br^- > I^- > SCN^-$ while for cations follow in the series of $Ca^{2+} > Mg^{2+} > Na^+ > K^+ > NH_3^+$. Ions high on the series interact strongly with water as water molecules are well structured around the ions relative to bulk water. These types of ions are usually referred to kosmotropes, which are water-structure-makers or salting out ions as they stabilise proteins and favour attractive intermolecular interactions,

while ions which interact weakly with water molecules tend to bind to and destabilise proteins and are referred to as chaotropes or water-structure-breakers. Recent studies have found that low concentration of chaotropes in acidic pH solutions, which are preferentially accumulated at the protein surface, do not destabilise proteins but rather salt out proteins more effectively than kosmotropes [45-47].

Generally, ionic (electrostatic) screening interactions occur between the charged groups on the protein surface and all salts, i.e. ionic screening is non-specific. The screening effect depends on pH and ionic strength of the solution. At low ionic strengths, the electrostatic interactions between ions and protein charged groups are always favourable, but are reduced at high ionic strengths as the range of the interaction becomes short-ranged. The electrostatic interactions depend on the sign and number of charges on a protein, which is modulated by pH. Protein-salt interactions are greatest when protein has largest net charge leading to large favourable interactions and salting-in effects. This effect contributes to the low solubility at isoelectric point where protein net charge is equal to zero. Specific electrostatic interactions are dependent on the specific nature of the salt and are usually dominant at high and low ionic strengths. For instance, Yamasaki et al. [48] found that at low ionic strength (10-100 mM), chaotropic salts such as sodium thiocyanate (NaSCN) and sodium perchlorate ($NaClO_4$) stabilised bovine serum albumin (BSA) against thermal unfolding due to specific ion binding to the charged groups on the protein surface and destabilised BSA at high ionic strengths by weakening hydrophobic forces. The latter is consistent with the finding that chaotropes bind non-specifically to non-polar surfaces at high salt concentration, which reduces protein thermodynamic stability of the proteins [49].

The specific binding of monovalent chaotropic anions to net positively charged proteins has been found to follow the reverse Hofmeister trend at concentrations below 200 mM [46, 47, 50]. The monovalent anions have a more pronounced effect on weakening intermolecular interactions and salting out proteins. The strength of interaction of protein-ion interactions can be explained by the Collin's law of matching water affinity or the electroselectivity series. Collins proposed that ions may interact with proteins depending on the hydration properties, size and polarizabilities of the ions and as a result, strong inner sphere ion pairs are formed between ions of opposite signs and similar size having similar water affinities [51, 52]. Large polarizable anions such as thiocyanate (SCN^-) which are weakly hydrated have the strongest interactions with the positively charged side chains of arginine (Arg), lysine (Lys) and histidine (His). The electroselectivity series ranks the affinity of ions for charged groups on the protein surface where ions with higher valency form stronger electrostatic interactions with protein charged groups. Gokarn et al. [45] demonstrated that the binding of divalent sulphate ion (SO_4^{2-}) to positively charged lysozyme is consistent with the electroselectivity series despite being strongly hydrated. However for monovalent anions, protein-ion interactions are consistent with Collins model. This was attributed to specific ion binding of SO_4^{2-} at low salt concentrations resulting in more effective charge neutralisation than the monovalent anions. Salt ions also interact with the large dipole moment of the peptide backbone, which arises from presence of partial positive and negative charges on the amino and carboxyl group of the backbone units respectively. This suggests that cations bind to the negative group on the carbonyl oxygen while anions bind to positive charges on the amino group [41, 53].

Processing conditions

Aggregates can occur in various processing steps such as fermentation, purification, refolding, freeze-thaw process, shaking, lyophilization and during long-term storage. During large-scale protein production, proteins often get expressed in inclusion bodies as insoluble aggregates [54]. During refolding of proteins, aggregates could be formed from high concentrations of misfolded proteins which tend to reduce yields. Proteins stored in either a lyophilised state, liquid state, or solid state have tendency to aggregate and factors such as protein concentration, pH and ionic strength affect the rate of aggregation [55]. Other potential causes of aggregation arise as a result of agitation encountered during processing as well as adsorption to surfaces of containers [56].

2.5 Protein-Protein Interactions

Usually, controlling or inhibiting protein aggregation in biopharmaceutical formulations requires manipulating or screening a wide range of solvent conditions to alter the solution thermodynamics, which are controlled by protein-protein interactions. Due to the complexity in physical and chemical properties of proteins, understanding the nature of protein-protein interactions at a molecular level is yet an unsolved problem. Solution variables that affect protein-protein interactions include temperature and solution conditions such as ionic strength, salt type, and pH. Finding an optimal combination of these variables that control solution properties of interest such as stability, solubility, crystallisation and precipitation requires full understanding of the origin of protein-protein interactions [57]. For instance, knowledge of protein-protein interactions characterized in terms of the osmotic second virial coefficient has been used to define a range of solution conditions that favour crystallisation. Various experimental and theoretical studies have been used

to explain protein-protein interactions in different component systems and to link the observed behaviour with solution properties of interest. Proteins can be described as charged colloids using models such as Derjaguin-Landau-Verwey-Overbeek (DLVO) theory, which can then be utilized for rationalizing colloidal phase diagrams [58, 59]. DLVO theory provides useful starting point in understanding protein-protein interactions. However, the use of this model is limited in that it does not account for specific ion effects, solvation forces and geometric anisotropy, which have been shown to be significant for protein solutions. This section reviews the nature of protein-protein interactions and the link to protein solution behaviour, and examines the effect of solution variables on protein-protein interactions.

One of the earliest measurements for measuring protein-protein interactions was based on using an osmometer [60]. In dilute solutions, the protein activity is directly determined by the actual concentration since the activity coefficient γ_p is approximately 1. However, in concentrated solutions, the effective protein concentration or thermodynamic activity deviates from ideality due to reducing the average protein-protein separation distance and increasing the contribution of protein-protein interactions [61]. This non-ideality can be measured in terms of the difference in pressure (referred to as osmotic pressure, π) of a protein and salt in solution versus the corresponding dialysate, that is the salt solution at the same solvent chemical potentials as the protein solution, in the form of a virial expansion

$$\frac{\pi}{cRT} = \frac{1}{M_w} + B_{22}(T, \mu_w, \mu_s)c + B_{33}(T, \mu_w, \mu_s)c^2 + \ldots \qquad \text{(Equation 2.1)}$$

where R is the universal or ideal gas constant, M_w is the molecular weight of the protein, c is the protein mass concentration, μ_i is the chemical potential of component i, where i equals w for water, and s for salt, T is the absolute temperature,

B_{22} and B_{33} is the second and third osmotic virial coefficient, respectively. From Equation 2.1, measuring osmotic pressure as a function of protein concentration allows for direct measurement of B_{22}. B_{22} reflects behaviour in a non-ideal solution due to two-body protein interactions, where higher order contributions from multi-body protein interactions are accounted for by higher order virial coefficients. When non-idealities due to protein-protein interactions are negligible as occurs at low protein concentration, then the osmotic pressure is given by the ideal van't Hoff equation

$$\pi = \frac{c}{M_w} RT \qquad \text{(Equation 2.2)}$$

The nature of molecular interactions between proteins can be rationalized in terms of B_{22}. A positive B_{22} reflects repulsive protein-protein interactions whereby the osmotic pressure is greater than that of an ideal solution. A negative B_{22} reflects attractive interactions between protein molecules as the osmotic pressure is less than the ideal solution value [62].

2.5.1 Osmotic Second Virial Coefficient

The osmotic second virial coefficient B_{22} provides a measure of the nature and extent of protein-protein interactions. B_{22} can be measured using techniques such as static or dynamic light scattering [63, 64], self-interaction chromatography [65], size exclusion chromatography [66], neutron scattering [67] and sedimentation equilibrium [68]. The general phase behaviour of proteins is ultimately controlled by the protein-protein potential of mean force (or value of B_{22}) [69]. Knowledge of the phase diagram has been used in determining optimal conditions for protein crystallisation. George and co-workers observed a relation between B_{22} and the

conditions favourable for protein crystal growth [70, 71] where high quality crystals are obtained when protein –protein interactions fall within a window of B_{22} values on the phase diagram. The window corresponds a location on the protein phase diagram near to a liquid-liquid equilibrium. B_{22} also provides valuable information on formulation conditions that increase colloidal stability or favour attractive protein-protein interactions.

Within the McMillan Mayer osmotic virial expansion, the second virial coefficient is thermodynamically related to a protein pair potential of mean force. The potential of mean force is related to an interaction averaged over the separation and orientation of a pair of protein molecules in solution [57]. The potential of mean force W(r), is defined as the work required to bring two protein molecules at infinite separation to a fixed configuration [72]. B_{22} is related to the potential of mean force by

$$B_{22} = -\frac{1}{2}\frac{N_A}{M_w^2}\int_0^\infty \left[e^{-w(r)/k_B T} - 1 \right] 4\pi r^2 dr \qquad \text{(Equation 2.3)}$$

where r is the intermolecular center-to-center distance, k_B is the Boltzmann constant and N_A is Avogadro's constant.

Studies have shown that B_{22} values depend on conditions such as pH, salt type, and salt concentration [73, 74]. Narayanan and Liu [67] experimentally measured the B_{22} values of proteins at low and high salt concentrations. The authors found that at low ionic strengths, B_{22} was positive indicating protein repulsion. However at higher ionic strengths, protein-protein interactions became attractive as reflected by the negative B_{22}. The effect of ionic strength on B_{22} observed is consistent with studies conducted by Lehermayr et al. [74] where the propensity of the protein to aggregate increased as the virial coefficients dropped. The pH dependence on protein-protein

interactions of a monoclonal antibody was investigated by Roberts et al. [64]. The authors found that the values of B_{22} decreased as the pH is increased indicating attractive protein-protein interactions and was attributed to a reduction in the double layer forces within the DLVO theory. The pH and ionic strength dependence on protein-protein interactions can be predicted with respect to how the electric double-layer forces changes within the DLVO theory [57].

2.5.2 DLVO theory

Most potential of mean force models use idealised shapes to describe proteins, and in some cases, approaches account for surface roughness/anisotropy of the proteins. The Derjaguin-Landau-Verwey-Overbeek (DLVO) theory is used to describe the behaviour of proteins in dilute electrolyte solutions where the proteins are highly hydrated and the closest distance of approach between protein molecules includes a tightly bound layer of water molecules. In DLVO theory, proteins are treated as hard spheres having uniformly charged surfaces with net charge given by Q, immersed in a dielectric medium (water) ε, containing salt ions behaving as point charges [75]. The potential of mean force W(r) between two protein molecules is expressed as a sum or contribution of three terms, (1) the excluded volume or protein hard sphere term W_{ex}, (2) the electrostatic or electric double layer repulsion W_{elec}, and (3) the dispersion potential of Hamaker W_{disp}. Thus

$$W(r) = W_{ex}(r) + W_{elec}(r) + W_{disp}(r) \qquad \text{(Equation 2.4)}$$

Generally, the DLVO model is applicable to predict protein behaviour at low electrolyte concentration (< 0.1 M) where long-ranged electrostatic forces dominate. However at higher salt concentrations (> 0.1 M), the model does not accurately

describe protein behaviour [76]. It should be noted that ion specific effects are not accounted for by DLVO theory.

Excluded Volume

The contribution of excluded volume to the osmotic second virial coefficient (non-ideal protein solutions) is positive. This volume includes a layer of tightly bound water of thickness δ, which controls the distance of closest approach between protein molecules. Therefore, it is important to accurately estimate the range of δ as it affects B_{22} values used to predict thermodynamic properties of proteins. Vilker et al. [77] showed that using a spheroid approximation to describe the shape of proteins compared to using a sphere provided better fit with experimental data for BSA where difference in excluded volume between the two approximations was greater than 11%.

Van der Waals Interactions

Van der Waal forces are dispersion forces, which are dominant when the protein-protein interaction is short-ranged. The strength of the dispersion force is determined by the nature of solute and solvent as well as the geometry of the molecules [61]. The dispersion potential is proportional to the Hamaker constant A_H, which is determined by the dielectric properties of proteins and the solvent (water). Between two similar bodies such as protein molecules, the Hamaker force is attractive. Because most proteins tend to have similar composition and densities, they have similar Hamaker constants [78]. Care must be taken when fitting the dispersion potential to experimental data because the strength of dispersion attraction is very sensitive to the distance of closest approach between proteins, which is controlled by the parameter describing the hydration layer thickness about proteins, δ.

Electrostatic Interactions

Electrostatic interactions or charge-charge repulsions are greatest in dilute solutions or at low ionic strength where the range of electrostatic forces is greatest; increasing the salt concentration reduces the range of electrostatic interactions between protein molecules. Electrostatic interactions between ions and the protein fixed charges produce an electric double-layer, where the ion density about the protein differs from the bulk solution over a distance given by the Debye length. Within DLVO theory, the ion density is described by treating the salt as point charges within the Poisson-Boltzmann equation. The double layer force arises when double-layers of different proteins overlap with each other. Vilker et al. [77] gives an analytical expression of the double layer force

$$W_{elec}(r) = \frac{z^2 e^2 \left(1/r\right) \exp\left[-\kappa\left(r - d_p\right)\right]}{4\pi\varepsilon_0\varepsilon\left(1 + \kappa d_p/2\right)^2} \qquad \text{(Equation 2.5)}$$

where e is elemental charge, d_p is protein diameter modelled as hard sphere, z is valence charge, ε is relative permittivity of water and ε_0 is dielectric permittivity of free space, and κ is the inverse of the Debye length, which reflects the size of the double layer or the range of the force.

DLVO theory assumes uniform charge distribution which is not always the case for proteins. At ionic strengths greater than 0.1 M, the debye length is short (~ 1 nm) and the equation is not valid as other short-ranged interactions become significant and the assumptions used in the Poisson-Boltzmann equation are no longer value, i.e. ions can no longer be treated as point charges. The equation can only be used at lower ionic strengths (e.g. 0.01 M) where the debye length is larger (~ 3 nm) to predict the behaviour of proteins in solution [41].

The influence of electrostatic interactions on protein-protein interactions can be measured experimentally by observing the effects of pH and ionic strength. pH affects the protonation states of the α–carboxylic and α–amino groups on protein surfaces [79]. At low/acidic pH, proteins are positively charged as the α–amino terminal groups of lysine, histidine and arginine are protonated. The negatively charged aspartate and glutamate α–carboxylic group are protonated and neutralised at acidic pH. Increasing the pH reduces the protein's net charge as carboxylic (–COOH) and amino residues ($-NH_3^+$) become deprotonated. Since the pH of solution determines whether the protein is positively or negatively charged, changes in pH would affect the nature of electrostatic interactions between protein molecules. According to Eq. 2.5, the protein net charge is the main factor contributing to electric double layer forces. At isoelectric point (pI), proteins have no net surface charge due to equal number of positive and negative charges on the protein surface. Hence, when the pH of the protein is far below or above the isoelectric point, the interactions become repulsive due to presence of large net positive or negative charges respectively. Neal et al. [80] reported that at low pH and ionic strength, the virial coefficient of lysozyme becomes highly positive as a result of repulsive interactions and moves from positive to negative B_{22} with increasing pH as the isoelectric point is approached.

2.5.3 Limitations of DLVO

DLVO theory cannot explain many trends of intermolecular protein-protein interactions. Below are given other types of forces that may be important in controlling protein interactions but are not included in DLVO theory.

Specific Ion Effects

Within the DLVO theory, salt ions are treated as point charges and therefore do not include the ion specific interactions that have been observed in a number of studies. In addition, the double layer theory only accounts for electrostatic interactions between ions and proteins, and as such ignores any nonelectrostatic ion-specific interactions at low or high salt concentrations. Specific ion effects are controlled and modulated by presence of salts where the role of salt in determining the extent and magnitude of interactions is still not well understood. Although the Hofmeister or lyotropic series which is based on the molal surface tension increment is often used to explain the effectiveness of ions in salting in/out proteins, not all ions follow this pattern [46, 81]. Within the salting out region, intermolecular interactions become attractive as the salt concentration is raised and the protein solubility decreases depending on the position of the ion in the Hofmeister series. The traditional salting out behaviour is observed when the pH is above the pI. In this case, ions structure surrounding water molecules and prevent them from hydrating the protein surface, thus the proteins form intermolecular interactions with each other. The exact molecular origin of the interaction is still being debated.

On the other hand, a reverse Hofmeister series is noticed at pH values below the pI where chaotropic ions are more effective in salting out proteins than kosmotropes. This anion-induced attraction has been observed especially at low salt concentrations where electrostatic interactions are significant [46]. The order reverts back to the traditional Hofmeister series at a salt concentration where electrostatic interactions are screened. One explanation is that at low salt concentrations, the anions are preferentially accumulated at the protein surface due to strong binding between the anions and positively charged groups. The binding lowers the net positive charge

and neutralizes the double layer forces leading to attractive protein-protein interactions. Specific anion binding increases in the descending order of lyotropic series however divalent anions such as sulphate have been reported to show stronger binding affinity to opposite charges on the proteins due to the higher number of valence [45, 47].

The understanding of specific ion effects requires that contributions from electrostatic interactions, dispersion forces and nonelectrostatic interactions (hydrogen bonding, hydration etc.) are included in protein-protein interaction models.

Solvation Effects

Solvation forces are also not accounted for in DLVO theory. Depending on the chemical nature of the interacting surfaces, solvation forces can either be attractive (between non polar and hydrophobic groups) or repulsive (between polar charged groups). The hydrophobic interaction is an attractive force which occurs when solvent layers around nonpolar groups overlap each other [57]. Also when polar groups approach each other, there is a repulsive solvation force (termed the hydration force), which arises from the free energy required to dehydrate tightly bound water molecules. Repulsive hydration forces play an important role in protein interactions as the force is believed to be responsible for preventing aggregation. These forces are also observed between protein surfaces with bound cations. As the hydration number of the cation increases, the strength and range of the hydration force also increases. This implies that binding of divalent cations such as Mg^{2+} and Ca^{2+} to protein surfaces enhances protein-protein repulsion as they are strongly hydrated [69, 82].

Anisotropic Effects

In solutions at high salt concentration, when the electric double-layer repulsion is screened and protein-protein interactions are short-ranged, the nature of interactions depends on the anisotropic distribution of protein surface groups and the shape of the protein [57]. Anisotropic interactions correspond to the case when the protein-protein interactions depend on the specific orientation of protein molecules [80, 83]. Protein-protein interactions are sensitive to surface properties as slight alterations could result in large changes in B_{22} values. Neal and co-workers examined the effect of protein surface roughness and shape on the contribution of excluded volume and found it the actual excluded volume to be about 1.7 times greater than if the protein molecule was modelled as a smooth surface [80, 84].

2.6 Conclusion

The issue of protein stability is particularly relevant in the field of biopharmaceutics due to the continuous increase in development of therapeutic protein products. Achieving stability is particularly difficult due to the marginal stability of proteins and their susceptibility to degradation. Non-native protein aggregation is problematic as it is encountered in different stages of the development process and is often irreversible. Protein aggregation is controlled by conformational and/or colloidal stability depending on solution conditions. Because proteins aggregate through different mechanisms, it is important to understand how varying solution conditions such as pH, temperature, salt type, salt concentration and other co-solvents affect protein stability and protein-protein interactions and how these link to aggregation. During formulation development, manipulating solution conditions to improve structural stability and protein-protein interactions is commonly employed in rational design of stable therapeutic protein formulations.

2.7 References

[1] Kranz, J., AlAzzam, F., Saluja, A., Svitel, J., Al-Azzam, W., Techniques for Higher-Order Structure Determination, in: Narhi, L. O. (Ed.), *Biophysics for Therapeutic Protein Development*, Springer New York 2013, pp. 33-82.

[2] Manning, M., Chou, D., Murphy, B., Payne, R., Katayama, D., Stability of Protein Pharmaceuticals: An Update. *Pharmaceutical Research* 2010, *27*, 544-575.

[3] Thirumangalathu, R., Krishnan, S., Bondarenko, P., Speed-Ricci, M., *et al.*, Oxidation of Methionine Residues in Recombinant Human Interleukin-1 Receptor Antagonist: Implications of Conformational Stability on Protein Oxidation Kinetics†. *Biochemistry* 2007, *46*, 6213-6224.

[4] Almarza, J., Rincon, L., Bahsas, A., Brito, F., Molecular Mechanism for the Denaturation of Proteins by Urea. *Biochemistry* 2009, *48*, 7608-7613.

[5] Lim, W. K., Rösgen, J., Englander, S. W., Urea, but not guanidinium, destabilizes proteins by forming hydrogen bonds to the peptide group. *Proceedings of the National Academy of Sciences* 2009, *106*, 2595-2600.

[6] Jeong, S., Analytical methods and formulation factors to enhance protein stability in solution. *Arch. Pharm. Res.* 2012, *35*, 1871-1886.

[7] Wang, W., Protein aggregation and its inhibition in biopharmaceutics. *International Journal of Pharmaceutics* 2005, *289*, 1-30.

[8] Calamai, M., Taddei, N., Stefani, M., Ramponi, G., Chiti, F., Relative Influence of Hydrophobicity and Net Charge in the Aggregation of Two Homologous Proteins†. *Biochemistry* 2003, *42*, 15078-15083.

[9] Fields, G. B., Alonso, D. O. V., Stigter, D., Dill, K. A., Theory for the aggregation of proteins and copolymers. *The Journal of Physical Chemistry* 1992, *96*, 3974-3981.

[10] Laurence, J. S., Middaugh, C. R., Fundamental Structures and Behaviors of Proteins, *Aggregation of Therapeutic Proteins*, John Wiley & Sons, Inc. 2010, pp. 1-61.

[11] Mehta, S. B., Bee, J. S., Randolph, T. W., Carpenter, J. F., Partial Unfolding of a Monoclonal Antibody: Role of a Single Domain in Driving Protein Aggregation. *Biochemistry* 2014, *53*, 3367-3377.

[12] Chen, B. L., Arakawa, T., Morris, C. F., Kenney, W. C., *et al.*, Aggregation pathway of recombinant human keratinocyte growth factor and its stabilization. *Pharm Res* 1994, *11*, 1581-1587.

[13] Hall, D., Minton, A. P., Turbidity as a probe of tubulin polymerization kinetics: A theoretical and experimental re-examination. *Analytical Biochemistry* 2005, *345*, 198-213.

[14] Brange, J., Havelund, S., Hougaard, P., Chemical Stability of Insulin. 2. Formation of Higher Molecular Weight Transformation Products During Storage of Pharmaceutical Preparations. *Pharmaceutical Research* 1992, *9*, 727-734.

[15] 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, *20*, 1325-1336.

[16] Bajaj, H., Sharma, V., Badkar, A., Zeng, D., *et al.*, Protein Structural Conformation and Not Second Virial Coefficient Relates to Long-Term Irreversible Aggregation of a Monoclonal Antibody and Ovalbumin in Solution. *Pharmaceutical Research* 2006, *23*, 1382-1394.

[17] Chi, E. Y., Krishnan, S., Kendrick, B. S., Chang, B. S., *et al.*, Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Science* 2003, *12*, 903-913.

[18] Militello, V., Casarino, C., Emanuele, A., Giostra, A., *et al.*, Aggregation kinetics of bovine serum albumin studied by FTIR spectroscopy and light scattering. *Biophysical Chemistry* 2004, *107*, 175-187.

[19] Giger, K., Vanam, R. P., Seyrek, E., and Dubin, P. L., Suppression of Insulin Aggregation by Heparin. *Biomacromolecules* 2008, *9*, 2338-2344.

[20] Katayama, D. S., Nayar, R., Chou, D. K., Valente, J. J., *et al.*, Effect of buffer species on the thermally induced aggregation of interferon-tau. *Journal of Pharmaceutical Sciences* 2006, *95*, 1212-1226.

[21] Brange, J., Andersen, L., Laursen, E. D., Meyn, G., Rasmussen, E., Toward understanding insulin fibrillation. *Journal of pharmaceutical sciences* 1997, *86*, 517-525.

[22] De Young, L. R., Fink, A. L., Dill, K. A., Aggregation of globular proteins. *Accounts of Chemical Research* 1993, *26*, 614-620.

[23] Raso, S. W., Abel, J., Barnes, J. M., Maloney, K. M., *et al.*, Aggregation of granulocyte-colony stimulating factor in vitro involves a conformationally altered monomeric state. *Protein Science* 2005, *14*, 2246-2257.

[24] 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. *Journal of Pharmaceutical Sciences* 2008, *97*, 4219-4227.

[25] 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. *Journal of pharmaceutical sciences* 2005, *94*, 1928-1940.

[26] Hejnaes, K., Matthiesen, F., Skriver, L., Protein Stability in Downstream Processing, *Bioseparation and Bioprocessing*, Wiley-VCH Verlag GmbH 1998, pp. 31-65.

[27] Arakawa, T., Tsumoto, K., Kita, Y., Chang, B., and Ejima, D., Biotechnology applications of amino acids in protein purification and formulations. *Amino Acids* 2007, *33*, 587-605.

[28] Saito, S., Hasegawa, J., Kobayashi, N., Tomitsuka, T., *et al.*, Effects of Ionic Strength and Sugars on the Aggregation Propensity of Monoclonal Antibodies: Influence of Colloidal and Conformational Stabilities. *Pharmaceutical Research* 2013, *30*, 1263-1280.

[29] Han, Y., Jin, B.-S., Lee, S.-B., Sohn, Y., *et al.*, Effects of sugar additives on protein stability of recombinant human serum albumin during lyophilization and storage. *Arch. Pharm. Res.* 2007, *30*, 1124-1131.

[30] Vagenende, V., Yap, M. G. S., Trout, B. L., Mechanisms of Protein Stabilization and Prevention of Protein Aggregation by Glycerol. *Biochemistry* 2009, *48*, 11084-11096.

[31] Ruzza, P., Hussain, R., Biondi, B., Calderan, A., *et al.*, Effects of Trehalose on Thermodynamic Properties of Alpha-synuclein Revealed through Synchrotron Radiation Circular Dichroism. *Biomolecules* 2015, *5*, 724.

[32] Huus, K., Havelund, S., Olsen, H., Weert, M., Frokjaer, S., Chemical and Thermal Stability of Insulin: Effects of Zinc and Ligand Binding to the Insulin Zinc-Hexamer. *Pharmaceutical Research* 2006, *23*, 2611-2620.

[33] Cimmperman, P., Baranauskienė, L., Jachimovičiūtė, S., Jachno, J., *et al.*, A Quantitative Model of Thermal Stabilization and Destabilization of Proteins by Ligands. *Biophysical Journal* 2008, *95*, 3222-3231.

[34] Lam, X. M., Patapoff, T. W., Nguyen, T. H., The Effect of Benzyl Alcohol on Recombinant Human Interferon-γ. *Pharmaceutical Research* 1997, *14*, 725-729.

[35] Zhang, Y., Roy, S., Jones, L. S., Krishnan, S., *et al.*, Mechanism for benzyl alcohol-induced aggregation of recombinant human interleukin-1 receptor antagonist in aqueous solution. *Journal of Pharmaceutical Sciences* 2004, *93*, 3076-3089.

[36] Arakawa, T., Timasheff, S. N., The stabilization of proteins by osmolytes. *Biophysical Journal* 1985, *47*, 411-414.

[37] Arakawa, T., Ejima, D., Tsumoto, K., Obeyama, N., *et al.*, Suppression of protein interactions by arginine: A proposed mechanism of the arginine effects. *Biophysical Chemistry* 2007, *127*, 1-8.

[38] Das, U., Hariprasad, G., Ethayathulla, A. S., Manral, P., *et al.*, Inhibition of Protein Aggregation: Supramolecular Assemblies of Arginine Hold the Key. *PLoS ONE* 2007, *2*, e1176.

[39] Sahin, E., Grillo, A. O., Perkins, M. D., Roberts, C. J., Comparative effects of pH and ionic strength on protein–protein interactions, unfolding, and aggregation for IgG1 antibodies. *Journal of Pharmaceutical Sciences* 2010, *99*, 4830-4848.

[40] Timasheff, S. N., Solvent effects on protein stability: Current Opinion in Structural Biology 1992, 2:35-39. *Current Opinion in Structural Biology* 1992, *2*, 35-39.

[41] Curtis, R. A., Ulrich, J., Montaser, A., Prausnitz, J. M., Blanch, H. W., Protein–protein interactions in concentrated electrolyte solutions. *Biotechnology and Bioengineering* 2002, *79*, 367-380.

[42] Hofmeister, F., Zur Lehre von der Wirkung der Salze. *Archiv f. experiment. Pathol. u. Pharmakol* 1888, *24*, 247-260.

[43] Kunz, W., Specific ion effects in colloidal and biological systems. *Current Opinion in Colloid and Interface Science* 2010, *15*, 34-39.

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

[45] Gokarn, Y. R., Fesinmeyer, R. M., Saluja, A., Razinkov, V., *et al.*, Effective charge measurements reveal selective and preferential accumulation of anions, but not cations, at the protein surface in dilute salt solutions. *Protein Science* 2011, *20*, 580-587.

[46] Zhang, Y., Cremer, P. S., The inverse and direct Hofmeister series for lysozyme. *Proceedings of the National Academy of Sciences* 2009, *106*, 15249-15253.

[47] Zhang-van Enk, J., Mason, B. D., Yu, L., Zhang, L., *et al.*, Perturbation of Thermal Unfolding and Aggregation of Human IgG1 Fc Fragment by Hofmeister Anions. *Molecular Pharmaceutics* 2013, *10*, 619-630.

[48] Yamasaki, M., Yano, H., Aoki, K., Differential scanning calorimetric studies on bovine serum albumin: II. Effects of neutral salts and urea. *International Journal of Biological Macromolecules* 1991, *13*, 322-328.

[49] Majumdar, R., Manikwar, P., Hickey, J. M., Samra, H. S., *et al.*, Effects of Salts from the Hofmeister Series on the Conformational Stability, Aggregation Propensity, and Local Flexibility of an IgG1 Monoclonal Antibody. *Biochemistry* 2013, *52*, 3376-3389.

[50] Zhang, J., Protein-Protein Interactions in Salt Solutions, in: Weibo Cai, H. H. (Ed.), *Protein-Protein Interactions - Computational and Experimental Tools* 2012.

[51] Collins, K. D., Ion hydration: Implications for cellular function, polyelectrolytes, and protein crystallization. *Biophysical Chemistry* 2006, *119*, 271-281.

[52] Collins, K. D., Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Macromolecular Crystallization* 2004, *34*, 300-311.

[53] Nandi, P. K., Robinson, D. R., Effects of salts on the free energy of the peptide group. *Journal of the American Chemical Society* 1972, *94*, 1299-1308.

[54] Finke, J. M., Roy, M., Zimm, B. H., Jennings, P. A., Aggregation Events Occur Prior to Stable Intermediate Formation during Refolding of Interleukin 1Î²â€ *Biochemistry* 1999, *39*, 575-583.

[55] Wang, W., Singh, S., Zeng, D. L., King, K., Nema, S., Antibody structure, instability, and formulation. *Journal of pharmaceutical sciences* 2007, *96*, 1-26.

[56] Wang, W., Li, N., Speaker, S., External Factors Affecting Protein Aggregation, *Aggregation of Therapeutic Proteins*, John Wiley & Sons, Inc. 2010, pp. 119-204.

[57] Curtis, R. A., Lue, L., A molecular approach to bioseparations: Protein–protein and protein–salt interactions. *Chemical Engineering Science* 2006, *61*, 907-923.

[58] Dumetz, A. C., Chockla, A. M., Kaler, E. W., Lenhoff, A. M., Protein Phase Behavior in Aqueous Solutions: Crystallization, Liquid-Liquid Phase Separation, Gels, and Aggregates. *Biophysical journal* 2008, *94*, 570-583.

[59] Kastelic, M., Kalyuzhnyi, Y. V., Hribar-Lee, B., Dill, K. A., Vlachy, V., Protein aggregation in salt solutions. *Proceedings of the National Academy of Sciences* 2015, *112*, 6766-6770.

[60] Starling, E. H., The glomerular functions of the kidney. *The Journal of Physiology* 1899, *24*, 317-330.

[61] Saluja, A., Kalonia, D. S., Nature and consequences of proteinâ€“protein interactions in high protein concentration solutions. *International Journal of Pharmaceutics* 2008, *358*, 1-15.

[62] Park, Y., Choi, G., Effects of pH, salt type, and ionic strength on the second virial coefficients of aqueous bovine serum albumin solutions. *Korean Journal of Chemical Engineering* 2009, *26*, 193-198.

[63] Saluja, A., Fesinmeyer, R. M., Hogan, S., Brems, D. N., Gokarn, Y. R., Diffusion and Sedimentation Interaction Parameters for Measuring the Second Virial Coefficient and Their Utility as Predictors of Protein Aggregation. *Biophysical Journal* 2010, *99*, 2657-2665.

[64] Roberts, D., Keeling, R., Tracka, M., van der Walle, C. F., *et al.*, Specific Ion and Buffer Effects on Protein–Protein Interactions of a Monoclonal Antibody. *Molecular Pharmaceutics* 2014, *12*, 179-193.

[65] Gabrielsen, M., Nagy, L. A., DeLucas, L. J., Cogdell, R. J., Self-interaction chromatography as a tool for optimizing conditions for membrane protein crystallization. *Acta Crystallographica Section D* 2010, *66*, 44-50.

[66] Bloustine, J., Berejnov, V., Fraden, S., Measurements of Protein-Protein Interactions by Size Exclusion Chromatography. *Biophysical Journal* 2003, *85*, 2619-2623.

[67] Narayanan, J., Liu, X. Y., Protein Interactions in Undersaturated and Supersaturated Solutions: A Study Using Light and X-Ray Scattering. *Biophysical Journal* 2003, *84*, 523-532.

[68] Deszczynski, M., Harding, S. E., Winzor, D. J., Negative second virial coefficients as predictors of protein crystal growth: Evidence from sedimentation

equilibrium studies that refutes the designation of those light scattering parameters as osmotic virial coefficients. *Biophysical Chemistry* 2006, *120*, 106-113.

[69] Broide, M. L., Tominc, T. M., Saxowsky, M. D., Using phase transitions to investigate the effect of salts on protein interactions. *Physical Review E* 1996, *53*, 6325-6335.

[70] George, A., Chiang, Y., Guo, B., Arabshahi, A., *et al.*, Second virial coefficient as predictor in protein crystal growth, *Methods in Enzymology*, Academic Press 1997, pp. 100-110.

[71] George, A., Wilson, W. W., Predicting protein crystallization from a dilute solution property. *Acta Crystallographica Section D* 1994, *50*, 361-365.

[72] McMillan Jr, W. G., Mayer, J. E., The statistical thermodynamics of multicomponent systems. *The Journal of Chemical Physics* 1945, *13*, 276-305.

[73] Mehta, C. M., White, E. T. and Litster, J. D. , Correlation of second virial coefficient with solubility for proteins in salt solutions. *Biotechnology Progress* 2012, *28*.

[74] Lehermayr, C., Mahler, H.-C., Mäder, K., Fischer, S., Assessment of net charge and protein–protein interactions of different monoclonal antibodies. *Journal of pharmaceutical sciences* 2011, *100*, 2551-2562.

[75] Verwey, E. J. W., Theory of the Stability of Lyophobic Colloids. *The Journal of Physical and Colloid Chemistry* 1947, *51*, 631-636.

[76] Zhang, Y., Cremer, P. S., Interactions between macromolecules and ions: the Hofmeister series. *Current Opinion in Chemical Biology* 2006, *10*, 658-663.

[77] Vilker, V. L., Colton, C. K., Smith, K. A., The osmotic pressure of concentrated protein solutions: Effect of concentration and ph in saline solutions of bovine serum albumin. *Journal of Colloid and Interface Science* 1981, *79*, 548-566.

[78] Moon, Y. U., Curtis, R. A., Anderson, C. O., Blanch, H. W., Prausnitz, J. M., Protein-Protein Interactions in Aqueous Ammonium Sulfate Solutions. Lysozyme and Bovine Serum Albumin (BSA). *Journal of Solution Chemistry* 2000, *29*, 699-718.

[79] 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. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* 2008, *1784*, 600-610.

[80] Neal, B. L., Asthagiri, D., Velev, O. D., Lenhoff, A. M., Kaler, E. W., Why is the osmotic second virial coefficient related to protein crystallization? *Journal of Crystal Growth* 1999, *196*, 377-387.

[81] Boström, M., Tavares, F. W., Finet, S., Skouri-Panet, F., *et al.*, Why forces between proteins follow different Hofmeister series for pH above and below pI. *Biophysical Chemistry* 2005, *117*, 217-224.

[82] Benas, P., Legrand, L., Ries-Kautt, M., Strong and specific effects of cations on lysozyme chloride solubility. *Acta Crystallographica Section D* 2002, *58*, 1582-1587.

[83] Curtis, R. A., Blanch, H. W., Prausnitz, J. M., Calculation of Phase Diagrams for Aqueous Protein Solutions. *The Journal of Physical Chemistry B* 2001, *105*, 2445-2452.

[84] Neal, B. L., Lenhoff, A. M., Excluded volume contribution to the osmotic second virial coefficient for proteins. *AIChE Journal* 1995, *41*, 1010-1014.

CHAPTER 3

3 EXPERIMENTAL METHODOLOGY

3.1 Introduction

This chapter reviews the analytical methods used in this research work (Chapters 4, 5, and 6). Many of the commonly used methods for characterizing aggregation propensity assess the conformational and/or colloidal stability of the protein. Current available techniques range from basic techniques to high throughput methods. Some of these techniques are either qualitative or quantitative, and differ in detection principles as they possess unique analytical features [1]. Table 3.1 shows the applications and measureable parameters for the methods used in this work. The measurable properties are the parameters used to interpret the results derived from each of the techniques after analysis of the raw data. For example, the osmotic second virial coefficient obtained from static light scattering measurements is used to assess the magnitude of colloidal stability or attractive /repulsive protein-protein interactions under various solution conditions. In order to gain a better understanding of the model systems used in this thesis, several complementary methods were simultaneously used in conducting experiments. Brief theoretical descriptions of the methodology and instrumentation are given in section 3.2 to 3.3.

Table 3.1. Analytical techniques for characterising protein stability

Analytical Technique	Measurable property	Application
Static light Scattering	Molecular weight, osmotic virial coefficient, aggregation onset temperature	Colloidal stability, aggregation propensity, solubility, crystallizability
Dynamic Light Scattering	Hydrodynamic size, diffusion coefficient, protein-protein interaction parameter, particle size distribution, viscosity	Colloidal stability, aggregation propensity, rheology
Intrinsic Fluorescence	Melting temperature, protein conformation and tertiary structure	Conformational and thermal stability
Turbidity	Sample turbidity	Aggregation, particulation
Electrophoretic Light Scattering	Electrophoretic mobility, zeta potential	Colloidal stability, aggregation propensity, ion binding

3.2 Theoretical Descriptions

3.2.1 Static Light Scattering

This technique allows for absolute determination of absolute molecular weight.

In a static light scattering (SLS) experiment, the fraction of incident light scattered by a solution is measured. The intensity of scattered light is recorded and averaged over a time scale. The property of interest is the excess scattered light which corresponds to the difference in scattered light from the protein solution over that of the solvent [2, 3]. Scattering intensity is collected at a single or multiple angles from either a light (multiangle light scattering, MALS) or laser source (multiangle laser light scattering, MALLS) [4]. This technique is also often referred to as static light scattering (SLS).

Consider a sample volume which consists of solution with solute species in the solvent. Once the sample is placed in the path of a light beam with wavelength λ, a fraction of the light is scattered by the solution. Figure 3.1 below shows an illustration of scattered light intensity monitored at a fixed angle by a detector. Light is scattered isotropically and the angular dependence of scattering intensity is lost for particles significantly smaller than $1/20^{th}$ the wavelength of incident light (658 nm). In this case, scattering data is analysed mainly using the angle 90° as other detectors are prone to multiple photon scattering or producing a smaller signal to noise ratio from large impurities such as dust particles as forward scattering is more sensitive to large particles.

The scattering is usually characterized in terms of the excess Rayleigh ratio of the solution, R_θ, which is the difference between the scattering of the solute-containing solution I_θ and scattering from the pure solvent $I_{\theta,\text{ solvent}}$

$$R_\theta = \frac{\left(I_\theta - I_{\theta,solvent}\right)n_0^2}{I_T n_T^2} R_T \qquad \text{(Equation 3.1)}$$

where I_T is the reference (toluene) scattering intensity, R_T is the Rayleigh ratio of toluene and n_0 and n_T are the refractive indices of the solvent and toluene solution respectively. To obtain accurate data, the concentration of the solute needs to be high enough to produce a high signal to noise ratio.

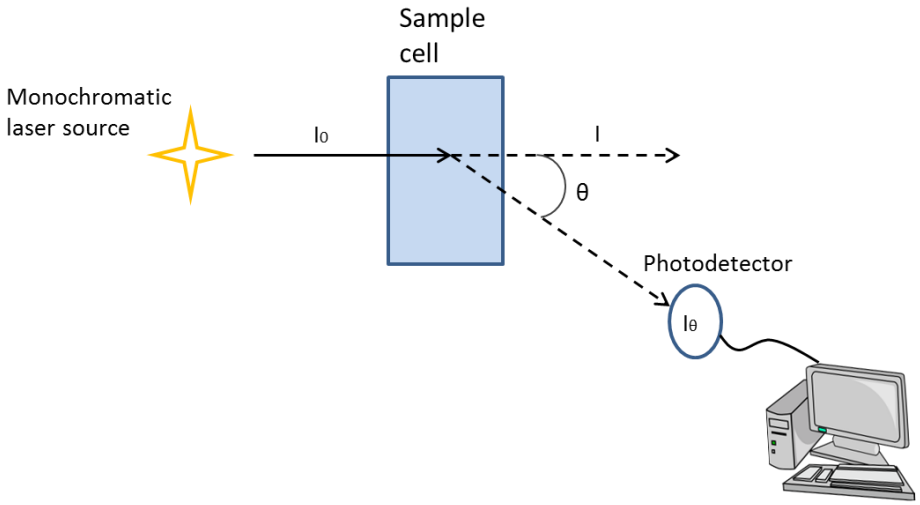


Figure 3.1. Principle operation for Static Light Scattering System. Incident light I_0 passes through the sample and the scattered light intensity, I_θ, is measured at angle θ with the aid of a photodetector.

The light scattering detector measures voltages and not light intensities. The proportionality between the detector voltage and light intensity can be expressed as

$$R_{90} = A_{cscc}\left(\frac{V_{90} - V_{90,solvent}}{V_{laser} - V_{laser,solvent}}\right) \qquad \text{(Equation 3.2)}$$

where A_{cscc} is the configuration specific calibration constant, V_{laser} and $V_{laser,solvent}$ are laser monitor signal for the solute-containing solution and solvent respectively, and V_{90} and $V_{90,solvent}$ are the detector signal voltage for the solute-containing solution and solvent at 90° respectively. The value of A_{cscc} depends on the instrument and solvent properties, such as the refractive index of solvent. Thus, an instrument dependent only constant A_{inst} is introduced to adjust for reflection and geometry corrections

$$A_{cscc} = A_{inst}\frac{n_s n_g}{T_{sg}^2 T_{ga}} \qquad \text{(Equation 3.3)}$$

where n_s and n_g are the refractive index of solvent and glass respectively and T_{sg} and T_{ga} are the transmitted intensity between the glass-solvent and air-glass interfaces of the sample cell. The transmitted intensity between the media 1 and 2 is expressed in terms of the Fresnel equation,

$$T_{12} = \frac{4n_1 n_2}{(n_1 + n_2)^2} \qquad \text{(Equation 3.4)}$$

Calibration is required to calculate the value for $A_{cscc,}$ which is done using toluene with a Rayleigh ratio of 1.406 x 10^{-5} cm^{-1} at a wavelength of 632.8 nm [5]. The basic Rayleigh equation of the scattering intensity of a protein solution is given by the Debye-Zimm relation [6],

$$\frac{Kc}{R_\theta} = \frac{1}{M_{w,0}} + 2B_{22}c \qquad \text{(Equation 3.5)}$$

where c is the protein concentration, $M_{w,0}$ is the infinite dilution molecular weight at zero concentration, B_{22} is the osmotic second virial coefficient (slope) and K is an optical constant given by

$$K = \frac{4\left[\pi n_0 \left(dn/dc\right)\right]^2}{N_A \lambda^4} \qquad \text{(Equation 3.6)}$$

where N_A is the Avogadro's constant (6.022 x 10^{23} mol^{-1}), λ is the wavelength of the incident light, n_0 is the solvent refractive index, and dn/dc is the differential refractive index increment of the solute which can be determined using a differential refractometer. Equation 3.5 is only accurate up to moderate protein concentrations below which contributions from higher order virial coefficients are negligible.

Equation 3.5 is used to analyse the light scattering data to determine the apparent mass-average molecular weight ($M_{w,app}$) and the osmotic second virial coefficient (B_{22}). The $M_{w,app}$ and B_{22} can be obtained directly from the y- intercept and slope of

the Zimm plot respectively when Kc/R_θ (y-axis) is graphed against the protein concentration, *c*, on the x-axis.

3.2.2 Dynamic Light Scattering

Dynamic light scattering (DLS) also known as quasi-elastic light scattering (QELS) or photon correlation spectroscopy (PCS) is a well-established technique for monitoring protein hydrodynamic properties in solution.

DLS is used to measure the time correlation function of fluctuations in scattered intensity of molecules in solution. The decay is determined by the time scale for particles undergoing Brownian motion. Large particles generate intensity fluctuations, which decay slowly while smaller particles move more rapidly leading to smaller time decays in correlations of light fluctuations. The fluctuations in scattered intensity can be mathematically expressed in terms of an intensity autocorrelation function $g^{(2)}(\tau)$ which reflects how quickly correlations in scattered light decay.

$$g^{(2)}(\tau) = \alpha + \beta[g(\tau)]^2 \qquad \text{(Equation 3.7)}$$

where α is an instrument constant and β is a background term in the limit of large delay time τ of a diffusional process.

For a monodispersed sample, the autocorrelation function can be fitted to a single exponential function,

$$g^{(1)}(\tau) = \exp\left(-D_T q^2 \tau\right) \qquad \text{(Equation 3.8)}$$

where D_T is the translational diffusion coefficient and q is the scattering vector length which depends on the scattering angle θ

$$q = \frac{4\pi n_0 \sin\left(\theta/2\right)}{\lambda_0} \qquad \text{(Equation 3.9)}$$

Most protein samples cannot be treated as monodispersed due to the presence of small quantities of aggregates or dust. In this case, the autocorrelation function cannot be represented by a single exponential function. Instead, the cumulants method is used to fit the autocorrelation time to a population distribution of decay rates characterized by a mean diffusion coefficient, D_z [7]

$$\ln g^{(1)}(\tau) = -q^2 D_z \tau + \frac{q^4 \tau^2}{2} (\delta D)_Z^2 \qquad \text{(Equation 3.10)}$$

where z is the z-average of the property and $(\delta D)_z^2$ is the z-average of the distribution width. A polydispersity Pd is defined as

$$Pd = \frac{(\delta D)_z^2}{D_z^2} \qquad \text{(Equation 3.11)}$$

Polydispersity provides information on the width of the particle distribution. A general rule of thumb is that samples with % polydispersity greater than ~ 10 % often indicate presence of higher-order oligomeric species while % Pd less than 10 % contain monodisperse samples. The sensitivity of the technique is limited for large polydispersities and it becomes difficult to obtain accurate measures of the protein size. This technique is used here to measure the hydrodynamic size distribution of the species in solution in response to environmental changes such as pH, ionic strength and temperature [8, 9]. Figure 3.2 shows the basic set up for dynamic light scattering where a monochromatic light beam such as a laser passes through the cell with the solution. The light scattered is detected by photomultiplier, which is then processed to determine the intensity autocorrelation function.

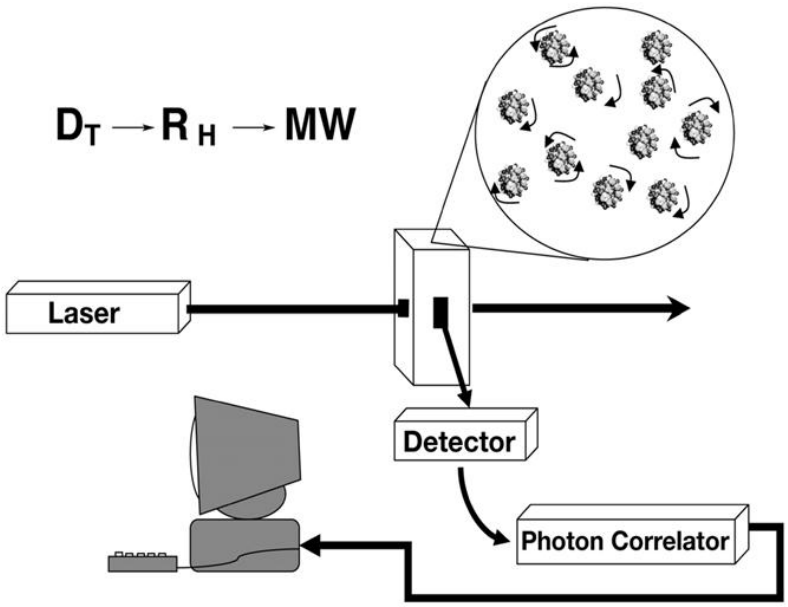


Figure 3.2. Schematic illustration of dynamic light scattering set up [10]. Laser beam illuminates the sample and the detector placed at a fixed scattering angle is used to measure the intensity of scattered light. A correlator is used to determine the correlation function from the time variation of the scattered light intensity.

Protein-protein interactions can be characterized from DLS measurements in terms of an inter-particle interaction parameter (k_D). k_D is obtained from the slope of the concentration dependence of the mutual or translational diffusion coefficient D_T [11] through

$$D_T = D_0(1 + k_D c) \qquad \text{(Equation 3.12)}$$

D_0 is the infinite dilution of the diffusion coefficient at zero concentration c, which is related to an intensity weighted average hydrodynamic radius, $R_{h,0}$, through the Stokes-Einstein relationship

$$R_{h,0} = \frac{k_B T}{6\pi\eta D_0} \qquad \text{(Equation 3.13)}$$

where k_B is Boltzmann constant, η is solvent viscosity and T is absolute temperature.

k_D provides quantitative/qualitative information about protein-protein interactions [12, 13] where larger positive k_D values reflect repulsive interactions and negative k_D values reflect self-association. The interaction parameter can be related to the osmotic second virial coefficient through

$$k_D = 2M_W B_{22} - \left(k_f + 2\nu_{sp}\right) \qquad \text{(Equation 3.14)}$$

where M_wB_{22} and $(k_f + \nu_{sp})$ reflect the thermodynamic and hydrodynamic components of k_D respectively. k_f is the coefficient of the linear term in the virial expansion of the frictional coefficient and ν_{sp} is the partial specific volume of the protein [8, 11].

3.2.3 Intrinsic Fluorescence

Intrinsic protein fluorescence is used to monitor changes in the tertiary structure of proteins under different solution conditions. Aromatic amino acid residues (tryptophan, tyrosine, and phenylalanine) contain chromophores which are usually buried inside native folded proteins. When excited, their electrons absorb UV light at lower wavelengths (blue shift) and jump into an unstable excited state, after which it relaxes back due to some energy loss into an unexcited ground state hence, emitting light at longer wavelengths (red shift) [1]. The excitation and emission profile of the protein is dependent on the polarity of the solvent environment around the aromatic residues. Upon unfolding, the solvent environment around the aromatic residues changes as the buried hydrophobic cores becomes exposed. The fluorescent emission changes leading to observable shifts in peak emission wavelength [3].

Fluorescence parameters such as peak emission wavelength, barycentric mean (BCM) fluorescence, integrated fluorescence intensity and intensity ratio (at 350/330 nm where tryptophan fluoresces) are often used for characterizing protein structures.

The integrated fluorescence intensity is obtained by integrating the area under the fluorescence curve for each spectrum, providing a measure of the total intensity (*I*) given by

$$I = \sum_{\lambda} I(\lambda) \qquad \text{(Equation 3.15)}$$

The peak emission wavelength can be determined by finding the wavelength which corresponds to the maximum intensity. The barycentric mean fluorescence (λ_{bcm}) is the most preferred method for monitoring unfolding as it is sensitive to changes in the average wavelength of florescence emission and takes into account the whole spectrum. λ_{bcm} is defined by the summation of the intensity at a particular wavelength multiplied by the value of the wavelength itself and divided by the integrated fluorescence intensity given by

$$\lambda_{bcm} = \frac{\sum_{\lambda} \lambda I(\lambda)}{\sum_{\lambda} I(\lambda)} \qquad \text{(Equation 3.16)}$$

3.2.4 Turbidity

Turbidity can be used to detect the presence of large protein aggregates under accelerated stress conditions such as pH, high protein concentration, excipients and ionic strength [3] but is unable to provide information about protein shape, size and distribution and should be used in combination with other tools which can detect presence of small oligomeric aggregates [15].

Turbidity is the fraction of incident light that passes through the sample. The reduction in light intensity is due to either light scattering or absorbance. The amount of light scattered depends on the particle size as smaller particles with low turbidities will scatter less light than large particles. The presence of precipitates, large aggregates or particles gives rise to turbid or opalescent solutions which could lead

to multiple scattering and interfere with direct scattering [14]. In Figure 3.3, a source of light beam passes through the sample cell and the scattered light is monitored using a 90° detector while a transmitted light detector measures changes in transmitted light intensity. The backscattering detector measures the amount of light scattered in backward directions at an angle between 90° to 180°.

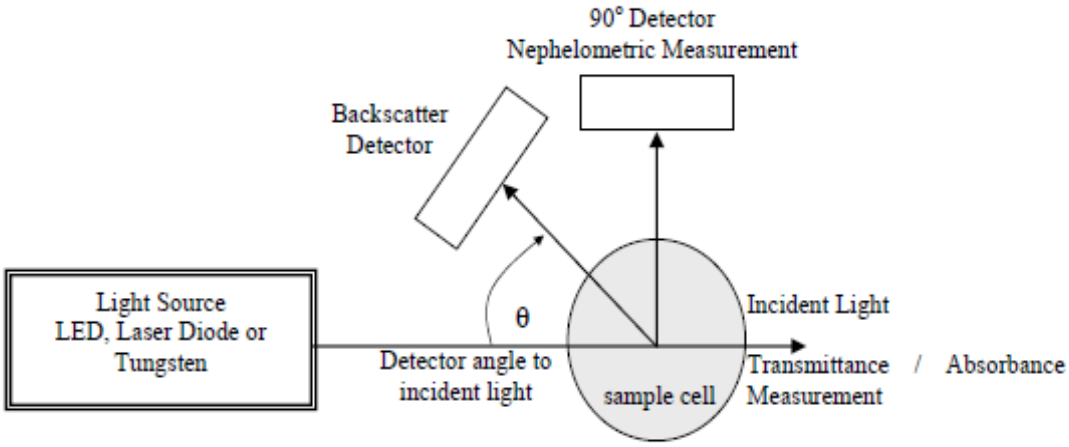


Figure 3.3. Turbidimetry techniques for determination of transmittance and absorbance [14]

Turbidity depends on the path length of the cell, size and concentration of the protein solutions and can be performed using standard UV spectrophotometers where the optical density or absorbance values are recorded at wavelengths (mostly 350 nm to 450 nm) where proteins do not exhibit specific absorbance of visible light [16], or using a colorimeter which provides a direct read out of the percent transmittance (%T) of the solution [17]. Transmittance refers to the amount of light that passes through a sample to the detector and is related to turbidity, τ, by

$$\tau = 100 - \%T \qquad \text{(Equation 3.17)}$$

3.2.5 Electrophoretic Light Scattering

Electrophoretic light scattering is used to measure the rate of migration of charged particles under the influence of an electric field. When an electric field is applied across an electrolyte solution, charged particles move towards the electrodes of opposite charge. This electrokinetic effect is referred to as electrophoresis [18]. At steady-state, the particles migrate at a constant velocity depending on the dielectric constant and viscosity of the medium, and strength of electric field. Once the velocity and applied electric field is known, the electrophoretic mobility μ_E and zeta potential can be calculated. The velocity *v*, is proportional to the electric field, *E*, by

$$v = \mu_E E \qquad \text{(Equation 3.18)}$$

A positive mobility implies that the surface of the particle is positively charged and vice-versa. However the surface charge could be affected by a change in solution conditions leading to a change in charge sign from positive to negative or negative to positive. The presence of a net charge on the surface of a particle in solution affects the distribution of ions surrounding the interfacial region thus, increasing the concentration of counter ions close the particle surface [19]. The region surrounding the particle consists of strongly bound ions referred to as the Stern layer and an outer layer consisting of loosely associated ions known as the diffuse layer. The bound counter ions in the stern layer and charged atmosphere of ions in the diffuse layer make up what is called the electrical double layer. The electrical double layer is formed to neutralise the surface charge.

When an electric field or voltage is applied, particles migrate with the surrounding solvent layer. Somewhere within the diffuse layer, there exists an 'imaginary boundary' beyond which solvent molecules will not move with the particles in

solution. This is called the slipping plane or hydrodynamic shear plane. The electrostatic potential that exists at the slipping plane with respect to the bulk of the solution is known as the zeta potential. Figure 3.4 shows the effect of counter ions on the electrostatic potential of a charged particle. The surface potential drops as the distance away from the particle surface increases, approaching zero as the densities of co-ions and counter ions merge together.

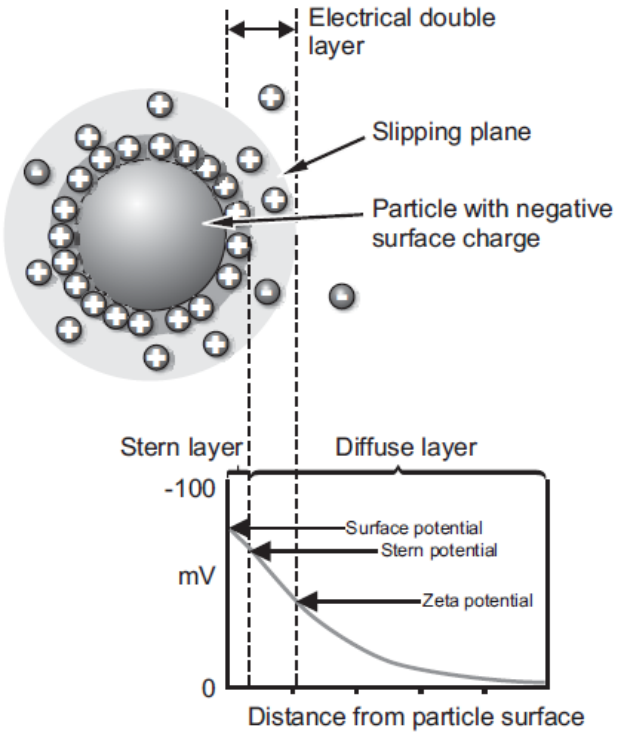


Figure 3.4. Visualization of electrical double layer and zeta potential of a negatively charged particle [19]

The value of the zeta (ζ) potential depends on the electrophoretic mobility and is calculated in two classic limits: the Huckel and Smoluchowski equations. When the particles radius R_h is larger than the debye length κ^{-1} i.e. $\kappa R_h >> 1$, the

Smoluchowski approximation is used and the relationship between ζ-potential and mobility is

$$\zeta = \frac{\mu_E \eta}{\varepsilon \varepsilon_0} \qquad \text{(Equation 3.19)}$$

where ε is relative permittivity of the fluid (water at 25 ° C = 78.3) and ε_0 is electric permittivity of vacuum (8.8547×10^{-12} C^2/ Nm^2).

The Henry- Huckel approximation is valid for particle radius smaller than the debye length ($\kappa R_h << 1$) in which case the ζ-potential is obtained from mobility by

$$\zeta = \frac{3\mu_E \eta}{2\varepsilon \varepsilon_0 f(\kappa R_h)} \qquad \text{(Equation 3.20)}$$

For the determination of Henry's function $f(\kappa R_h)$, two values are used as $f(\kappa R_h)$ varies from 1.0 for low κR_h values to 1.5 as κR_h approaches infinity. A simple analytical expression for the Henry's function derived using the electrostatic double-layer theory is given by [20]

$$f(\kappa R_h) = 1 + \frac{1}{2}\left[1 + \left(\frac{2.5}{\kappa R_h\left[1 + 2\exp(-\kappa R_h)\right]}\right)\right]^{-3} \qquad \text{(Equation 3.21)}$$

The Huckel equation is used in this thesis as measured zeta potential (ζ) values are less than 50 mV indicating that the polarization of ions in the diffuse layer is negligible [18]. The effect of specific ion binding on the surface potential of the protein can be determined by estimating the effective net charge, Z_{eff} from the electrophoretic mobility μ_E, [21]

$$Z_{eff} = \frac{\mu_E 6\pi\eta R_h (1 + \kappa R_h)}{e f(\kappa R_h)} \qquad \text{(Equation 3.22)}$$

where *e* is the unit charge of the proton (1.602 x 10^{-19} C). The debye screening length κ^{-1} can be calculated from the equation

$$\kappa = \sqrt{\frac{2e^2 N_A I}{\varepsilon\varepsilon_0 k_B T}} \qquad \text{(Equation 3.23)}$$

where N_A is Avogadro constant, I is ionic strength, k_B is Boltzmann constant (1.3806 x 10^{-23} J/K) and T is absolute temperature.

The electric double layer force between a pair of proteins is determined by the potential at the beginning of the diffuse layer, which is often approximated by the zeta-potential. As such the zeta-potential provides an indication about the colloidal stability of biological molecules [22]. Zeta potential can be used as a screening tool for selecting optimal solution conditions such as pH, salt type and ionic strength in formulations [23]. For instance, as the concentration or valence of ions in solution is increased, more counter ions will be available to neutralise the particle resulting in a reduction in the double layer thickness. This compression of the double layer leads to a drop in zeta potential resulting in reduced electrostatic repulsions or colloidal stability.

3.3 Experimental methodology

3.3.1 miniDAWN Treos

Static (SLS) and dynamic (DLS) light scattering measurements were performed with the DAWN TREOS light scattering detector. Light scattering measurements are performed in batch-mode for a single protein concentration. Multiple concentration measurements requiring serial dilutions from a stock solution are needed to perform a Zimm analysis. In this approach, individual readings are obtained for each

concentration by manually replacing the protein solution in the syringe used for delivering the sample into the detector. First, the pure solvent is pumped into the Treos detector (Wyatt, Santa Barbara, US) until a stable baseline is obtained. Subsequently, the dilutions made from the stock solution are passed through the detector, starting usually with the most concentrated protein solution. Concentration measurements are usually obtained by connecting an external UV/Vis absorption detector or Optilab-rEX differential refractometer in series with the light scattering flow-cell detector. The intensity of scattered light is collected at three different angles in the detector with an instrument constant of 4.68×10^{-5}. For the experiments conducted in this work, light scattering intensities were collected at detector 2 corresponding to an angle of 90°. Typical raw UV and light scattering data are shown in Figure 3.5. The SLS data are collected on the Wyatt Astra 6 software and exported to an Excel spreadsheet for determination of the virial coefficient and molecular weight. The Wyatt software is used to process the intensity auto-correlation functions using the cumulant analysis to determine the hydrodynamic sizes and diffusion coefficients, which are then exported to excel to determine the protein-protein interaction parameter.

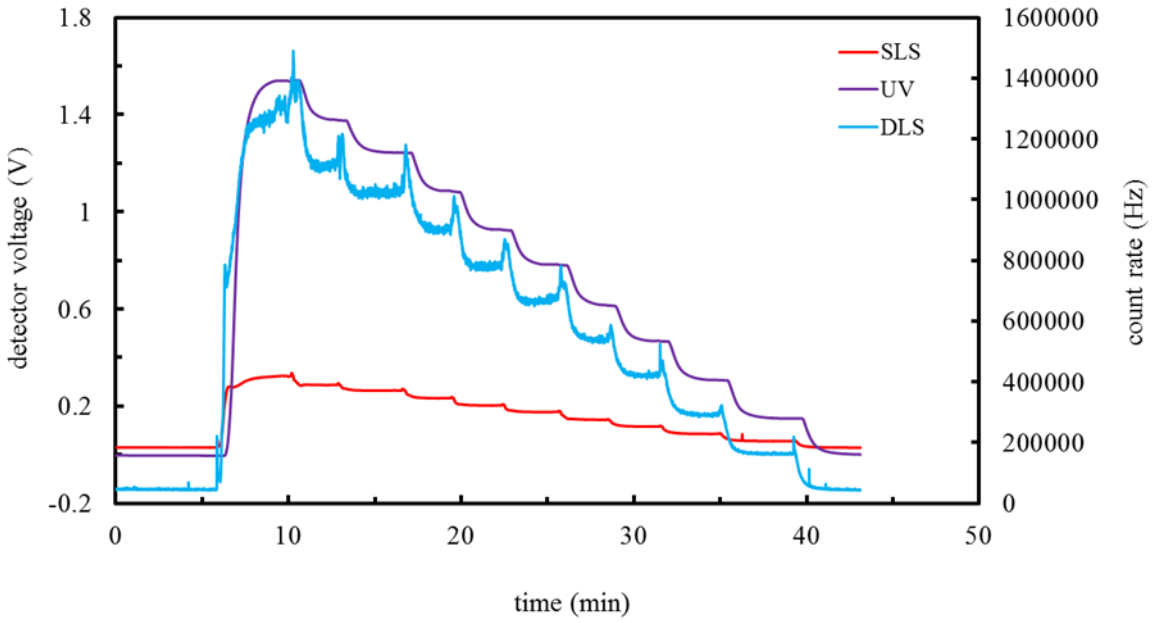


Figure 3.5. Experimental data for lysozyme in a solution of 100 mM KCl pH 7 obtained using the Treos detector. Stock protein samples are manually diluted with buffer to generate 10 different protein concentrations. Measurements of detector voltages for SLS, absorbance at 280 nm for UV, and count rates for DLS are shown.

3.3.2 Calypso

The Calypso instrument provides an automated delivery system which uses programmable syringe pumps to generate stepwise and well-defined concentration gradients from stock solutions of the protein sample and solvent. Figure 3.6 provides a pictorial detail of a typical light scattering arrangement utilizing the Calypso [24, 25]. A typical setup of the Calypso system consists of a multi-angle light scattering detector (mini DAWN-Treos) and UV absorbance detector connected to a computer installed with Calypso software which is used to control the flow rates and analyse scattering data and concentration measurements. The Calypso system generates very precise flowrates allowing for exact serial dilutions. The solutions are usually diluted and injected into detectors by automated-control syringe pumps. The syringe pumps

are programmed to collect the exact aliquot of sample required for each injection into the light scattering detector and UV detector flow cells. A series of concentration gradients is generated as a result of the varying flow ratios of the sample solution and solvent. At the end of each injection, the flow is stopped and allowed to equilibrate for proper mixing. An added advantage to the system is the presence of in-line filters as shown in Figure 3.6 before the mixing point which removes dust contaminants or aggregates.

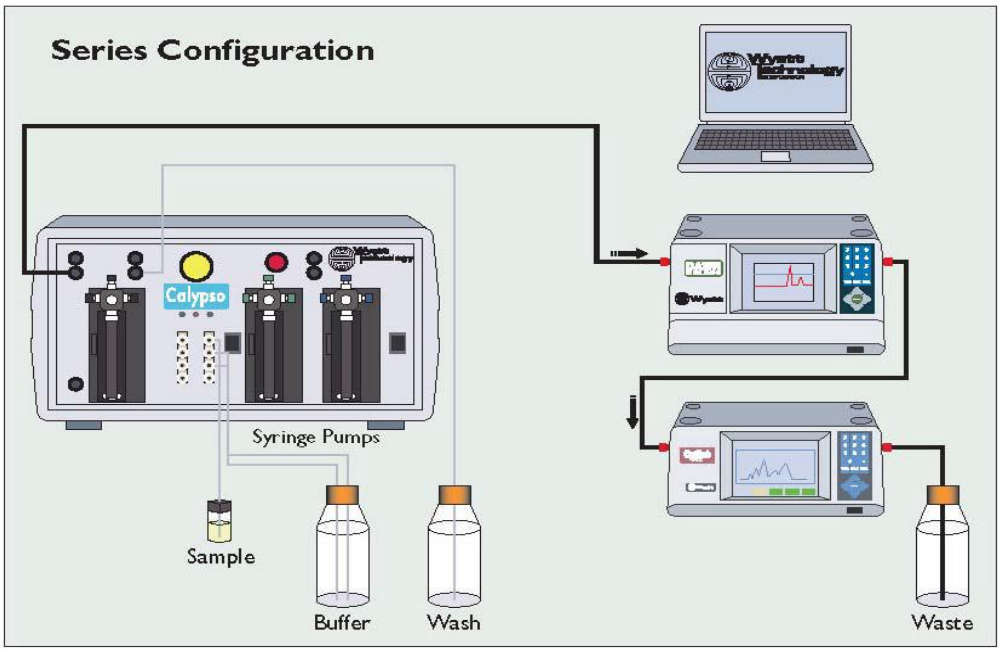


Figure 3.6. Pictorial representation of the Calypso System (CG-MALLS)

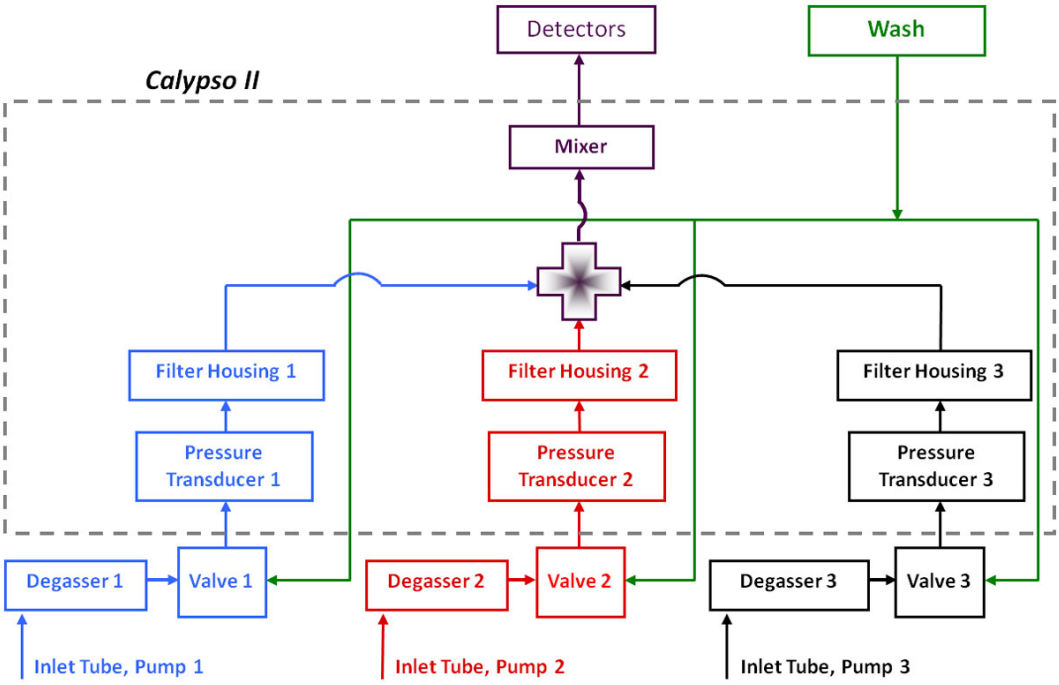


Figure 3.7. Schematic representation of Calypso II CG-MALS [26]

As shown by the schematic of the Calypso in Figure 3.7, each of the computer-controlled syringe pumps is connected to a 3-port distribution valve. The input port is connected to the sample reservoir from where the sample or solvent is taken up into a 0.5 mL syringe. The output port takes the solution from the syringe to the detectors through in-line filters, and a mixer. The wash port is connected to a central reservoir containing wash fluid to flush syringes and detectors. The inlet tubes have integrated holders which fit either 15 mL or 50 mL conical centrifuge tubes. Located adjacent to each syringe pump is the degasser which removes dissolved gases or bubbles that affect the quality of data. At the output port of each syringe pump, pressure transducers are placed to prevent pumps from exceeding the maximum permissible pressure.

Running an experiment in Calypso entails designing a method template in the software, preparing the instrument for experiment and running the experimental

method. Details on preparations for running an experiment which includes pump operations and method design can be found in the manual [26]. Because the Calypso uses 3 pumps, experimental methods can be designed to measure multiple virial coefficients as a function of co-solvent composition. Figure 3.8 shows a typical template for running experiments. In one experiment, the gradient run for lysozyme is generated by setting the flow rates of the 3 syringe pumps in such a way that the flowrate of pump 1 (containing lysozyme solution) increases in steps as the sample concentration gradually increases. Pump 2 contains the salt solution while pump 3 contains the buffer. For each virial coefficient of k_D determination (with a maximum protein concentration of 7 g/L), the salt concentration is fixed by setting the flowrates of pump 2 and 3. In the method shown below, 6 B_{22} and k_D values are obtained in a single run.

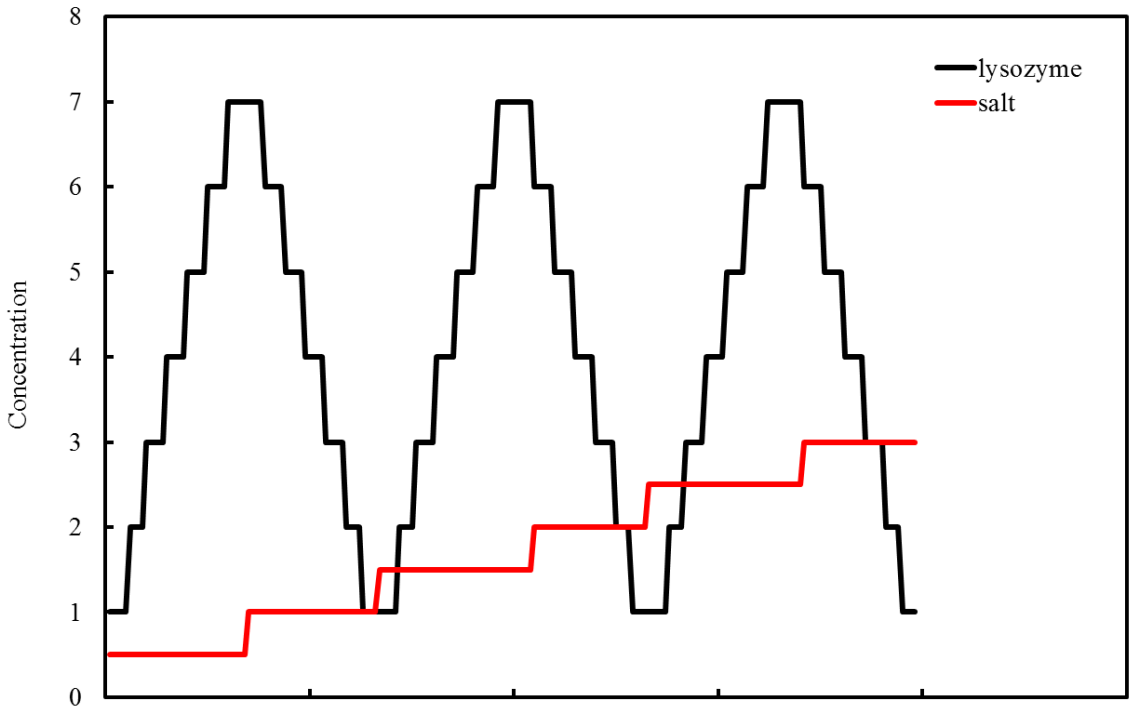


Figure 3.8. Example of a composition-gradient method for lysozyme (black lines) with a maximum concentration of 7 g/L. Measurements are made as a function of

salt concentration (red lines) ranging between 0.5 M and 3 M. Stock protein samples prepared at 14 g/L are automatically diluted with buffer to generate different protein concentrations at a fixed salt concentration.

The stock solution determines the maximum concentration of protein that can be measured on the MALS and concentration detectors. The 7-step lysozyme concentration gradient shown in Figure 3.8 is generated from a 14 mg/mL lysozyme stock solution. The number of steps in the method corresponds to the number of data points on a Zimm plot. The effect of salt concentration on B_{22} and k_D was determined by preparing a 6 M stock salt solution and running experiments within the same method at 6 different salt concentrations (0.5 M to 3 M).

3.3.3 Optim

The Optim (Avacta Analytical) is a low-sample, high-throughput analytical instrument that performs simultaneous measurements of conformational and colloidal stability of proteins by observing changes in intrinsic protein fluorescence and static light scattering respectively. This multi-modal tool is used during pre-formulation, formulation and comparability studies to monitor the effect of different formulations on protein aggregation and conformation [27, 28]. The thermal stability of proteins under different formulations can be assessed by performing temperature-controlled measurements from 15 °C to 95 °C. Under accelerated stress conditions such as temperature, the onset of unfolding and aggregation is usually reported as melting temperature (T_m) and aggregation temperature (T_{agg}) in fluorescence and light scattering measurements respectively [29]. Information on both fluorescence and light scattering measurements are usually obtained by scanning through the entire spectra region i.e. between 200 nm to 500 nm [30]. In addition to being high

throughput, this technique is fully automated, requiring as little as 9 uL of sample for each measurement. Samples are typically placed in specialised microcuvette arrays making it possible to perform 48 measurements in a single experimental run. It is also possible to perform isothermal measurements where samples are held at fixed temperature to determine the rate of unfolding and aggregation, as well as to conduct extrinsic fluorescent measurements in presence of fluorescent dyes.

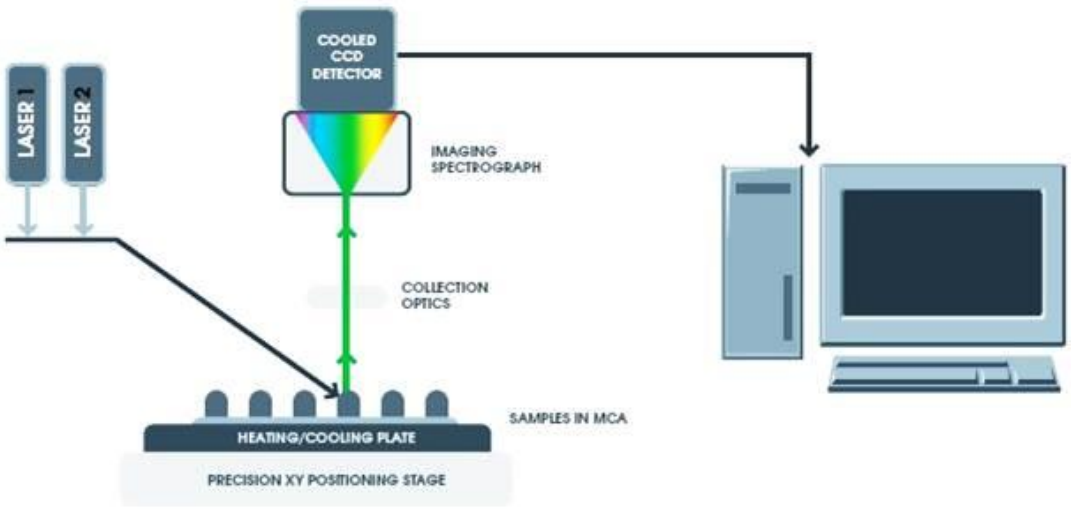


Figure 3.9. Schematic of Optim 1000 system configuration [27]

The Optim uses two lasers at 266 nm and 473 nm for exciting intrinsic fluorescence and light scattering measurements respectively [31]. Figure 3.9 shows an illustration of how the two laser sources pass through the samples and are detected using a Peltier cooled CCD detector over a range of wavelengths to generate a spectrum in a single exposure. In light scattering, the 266 nm laser can be used for greater sensitivity to smaller aggregates or low sample concentrations while the 473 nm illumination is used to detect presence of large aggregates or high sample concentration (greater than 100 mg/mL). For thermal denaturation and aggregation studies, the sample is efficiently heated through a thermoelectrically-controlled

copper heating and cooling plate as the sample is exposed to the laser for a relatively short period (≤ 1second). Measurements are generally taken at a fixed temperature with increasing temperature steps of 1 °C and holding time of 30 seconds to allow for proper sample equilibration.

3.3.4 Zetasizer Nano system

The Malvern Zetasizer Nano system equipped with a built-in Peltier temperature control and an avalanche photodiode detector (Malvern Instruments Ltd., UK) was used to carry out dynamic light scattering and zeta potential measurements. Both measurements are performed at a fixed scattering angle of 173° and measurements are usually fixed for a duration of time to allow for proper equilibration. For DLS measurements, the Nano S series equipped with a 633 nm Helium-Neon (He-Ne) red laser source was used. DLS samples are usually measured in a 45 µL low volume quartz cuvette, DTS2145 (Hellma Analytics, Germany). Zeta potential measurements were carried out the Nano ZS series by phase analysis light scattering method (M3-PALS) which is equipped with a 532 nm green laser source. Samples for zeta potential studies are usually measured with a universal dip cell (ZEN 1002) suitable for high conductivity samples and the dip cell was used in conjunction with a disposable 1 mL polystyrene cuvette (DTS0012).

Data acquisition and analysis for both measurements are carried out using the Dispersion Technology Software (DTS) application (Malvern Instruments). In DLS studies, the acquired correlogram is fitted to the correlation function using the method of cumulants to determine a z-averaged or cumulant-averaged molecular "size" (see Eq. 3.10). In zeta potential studies, the zeta potential is calculated through the measured electrophoretic mobility (μ_E) (see Eq. 3.20). The instrument also has the capability of performing simultaneous static light scattering measurements to

obtain derived count rates which are calculated from mean count rates and corrected attenuation factors which provide a measure of the scattered light intensities. The calculated scattered intensities can then be used to calculate R_θ (see Eq. 3.1).

3.3.5 DynaPro

The DynaPro plate reader II (Wyatt Technology) is a high-throughput temperature-controlled plate reader used to perform DLS measurements. Measurements can be made using 96, 384 or 1,536 well plates requiring sample volumes as low as 4 uL [32]. This makes it possible to achieve rapid and timely detection and characterisation of protein samples. The instrument uses a 830 nm wavelength illumination from a GaAs laser and analyses scattered light at an angle of 158° by using a non-invasive backscattering technique. The high reproducibility, productivity and flexibility of the DynaPro makes it suitable for protein formulation development. For example, thermal-induced aggregation experiments can be performed as a function of time and/or temperature to examine the influence of colloidal and thermal stability in proteins [9]. Thermal experiments can be performed within a temperature range of 4 °C to 85 °C. Usually in these experiments, silicon oil is applied on top of the samples to prevent evaporation. Before running experiments, protein samples and buffer solutions are filtered to remove the presence of unwanted particles which can affect the reliability of the data. The DYNAMICS 7.1.7 data software performs the analysis needed to obtain the diffusion software from the correlation function using the cumulants analysis. The processed results are then exported into an excel spreadsheet to obtain the averaged apparent hydrodynamic sizes and molecular weights, and for the calculation of k_D.

3.3.6 Turbidimetric titration

The Brinkmann PC 950 probe colorimeter with a wavelength of 490 nm was used to perform all turbidity measurements. The experimental set up is illustrated in Figure 3.10 below where a tungsten lamp provides light source which passes into the solution through the fibre optics light path. Light scattered by the proteins in solution is reflected back into the probe by an in-built mirror positioned 2cm away from the end of the fibre optic rod. The light is then returned to the instrument through the return light path where it strikes the photodiode detector which converts the light into transmittance values.

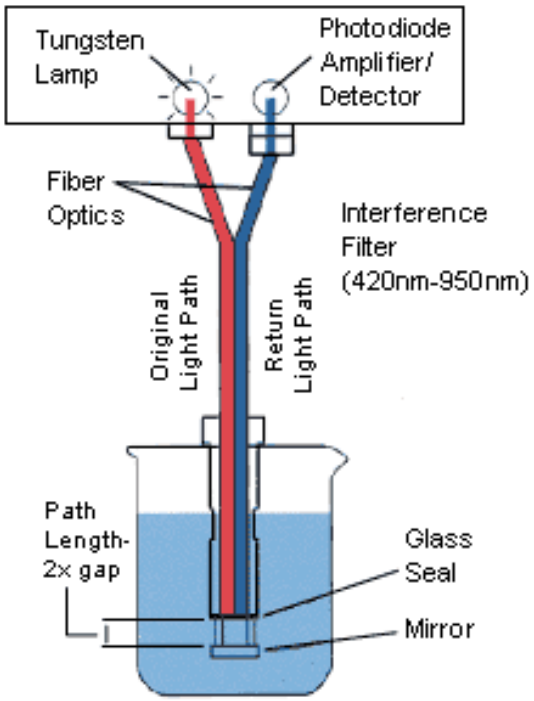


Figure 3.10. Schematic diagram of a PC 950 probe colorimeter [33]

Turbidimetric titration is carried out through incremental or stepwise addition of the additive solvent with a known concentration to a fixed volume of a protein with known concentration. This allows for determination of the effect of the additive

concentration on protein solution turbidity. Before conducting experiments, the reading is usually calibrated to 100% transmittance with a pure solvent like water.

3.4 References

[1] Chaudhuri, R., Cheng, Y., Middaugh, C. R., Volkin, D., High-Throughput Biophysical Analysis of Protein Therapeutics to Examine Interrelationships Between Aggregate Formation and Conformational Stability. *The AAPS journal* 2014, *16*, 48-64.

[2] Ahamed, T., Ottens, M., van Dedem, G. W. K., van der Wielen, L. A. M., Design of self-interaction chromatography as an analytical tool for predicting protein phase behavior. *Journal of Chromatography A* 2005, *1089*, 111-124.

[3] Samra, H. S., He, F., Advancements in High Throughput Biophysical Technologies: Applications for Characterization and Screening during Early Formulation Development of Monoclonal Antibodies. *Molecular Pharmaceutics* 2012, *9*, 696-707.

[4] Sahin, E., Grillo, A. O., Perkins, M. D., Roberts, C. J., Comparative effects of pH and ionic strength on protein–protein interactions, unfolding, and aggregation for IgG1 antibodies. *Journal of Pharmaceutical Sciences* 2010, *99*, 4830-4848.

[5] Harding, S. E., Jumel, K., Light Scattering, *Current Protocols in Protein Science*, John Wiley & Sons, Inc. 2001.

[6] Zimm, B. H., The Scattering of Light and the Radial Distribution Function of High Polymer Solutions. *The Journal of Chemical Physics* 1948, *16*, 1093-1099.

[7] Frisken, B. J., Revisiting the method of cumulants for the analysis of dynamic light-scattering data. *Appl. Opt.* 2001, *40*, 4087-4091.

[8] Kumar, V., Dixit, N., Zhou, L., Fraunhofer, W., Impact of short range hydrophobic interactions and long range electrostatic forces on the aggregation kinetics of a monoclonal antibody and a dual-variable domain immunoglobulin at

low and high concentrations. *International Journal of Pharmaceutics* 2011, *421*, 82-93.

[9] Rubin, J., Sharma, A., Linden, L., Bommarius, A. S., Behrens, S. H., Gauging Colloidal and Thermal Stability in Human IgG1–Sugar Solutions through Diffusivity Measurements. *The Journal of Physical Chemistry B* 2014, *118*, 2803-2809.

[10] Borgstahl, G. O., How to Use Dynamic Light Scattering to Improve the Likelihood of Growing Macromolecular Crystals, in: Walker, J., Doublié, S. (Eds.), *Macromolecular Crystallography Protocols*, Humana Press 2007, pp. 109-130.

[11] Narayanan, J., Liu, X. Y., Protein Interactions in Undersaturated and Supersaturated Solutions: A Study Using Light and X-Ray Scattering. *Biophysical Journal* 2003, *84*, 523-532.

[12] Rubin, J., Linden, L., Coco, W. M., Bommarius, A. S. and Behrens, S. H., Salt-induced aggregation of a monoclonal human immunoglobulin G1. *J. Pharm. Sci.* 2013, 377–386.

[13] Saluja, A., Fesinmeyer, R. M., Hogan, S., Brems, D. N., Gokarn, Y. R., Diffusion and Sedimentation Interaction Parameters for Measuring the Second Virial Coefficient and Their Utility as Predictors of Protein Aggregation. *Biophysical Journal* 2010, *99*, 2657-2665.

[14] Bin Omar, A., Bin MatJafri, M., Turbidimeter Design and Analysis: A Review on Optical Fiber Sensors for the Measurement of Water Turbidity. *Sensors* 2009, *9*, 8311.

[15] Lowe, D., Dudgeon, K., Rouet, R., Schofield, P., *et al.*, Aggregation, stability, and formulation of human antibody therapeutics, *Advances in Protein Chemistry and Structural Biology*, Academic Press 2011, pp. 41-61.

[16] He, F., Razinkov, V., Middaugh, C. R., Becker, G., High-Throughput Biophysical Approaches to Therapeutic Protein Development, in: Narhi, L. O. (Ed.), *Biophysics for Therapeutic Protein Development*, Springer New York 2013, pp. 7-31.

[17] Sharma, V. K., Kalonia, D. S., Experimental Detection and Characterization of Protein Aggregates, *Aggregation of Therapeutic Proteins*, John Wiley & Sons, Inc. 2010, pp. 205-256.

[18] Delgado, A. V., González-Caballero, F., Hunter, R. J., Koopal, L. K., Lyklema, J., Measurement and interpretation of electrokinetic phenomena. *Journal of Colloid and Interface Science* 2007, *309*, 194-224.

[19] Malvern, Zetasizer Nano Series User Manual. *Malvern Instruments Ltd* 2004.

[20] Ohshima, H., A Simple Expression for Henry's Function for the Retardation Effect in Electrophoresis of Spherical Colloidal Particles. *Journal of Colloid and Interface Science* 1994, *168*, 269-271.

[21] Szymański, J., Poboży, E., Trojanowicz, M., Wilk, A., *et al.*, Net Charge and Electrophoretic Mobility of Lysozyme Charge Ladders in Solutions of Nonionic Surfactant. *The Journal of Physical Chemistry B* 2007, *111*, 5503-5510.

[22] Schultz, N., Metreveli, G., Franzreb, M., Frimmel, F. H., Syldatk, C., Zeta potential measurement as a diagnostic tool in enzyme immobilisation. *Colloids and Surfaces B: Biointerfaces* 2008, *66*, 39-44.

[23] Salgın, S., Salgın, U., Bahadır, S., Zeta Potentials and Isoelectric Points of Biomolecules: The Effects of Ion Types and Ionic Strengths. *Int. J. Electrochem. Sci.* 2012, *7*, 12404 - 12414.

[24] Attri, A. K., Minton, A. P., Composition gradient static light scattering: A new technique for rapid detection and quantitative characterization of reversible

macromolecular hetero-associations in solution. *Analytical Biochemistry* 2005, *346*, 132-138.

[25] Attri, A. K., Minton, A. P., New methods for measuring macromolecular interactions in solution via static light scattering: basic methodology and application to nonassociating and self-associating proteins. *Analytical Biochemistry* 2005, *337*, 103-110.

[26] Wyatt, Wyatt Calypso II Syatem User's Guide. *Wyatt Technology Corporation* 2010.

[27] Avacta, www.optim.unchainedlabs.com. *Avacta Analytical Ltd* 2010.

[28] Webster, S., *Pharmaceutical Technology* 2013.

[29] Avacta, Predicting monoclonal antibody stability in different formulations using Optim 2. *Application note*.

[30] Avacta, Optim 2 User Guide v 1.5.4. *Avacta Analytical Ltd* 2012.

[31] Middaugh, C. R., Expediting Early Preformulation Studies : High-Throughput Characterization Systems and Empirical Phase Diagrams Can Speed Process *Genetic Engineering & Biotechnology News* 2012 *32*, 42-43.

[32] Wyatt, DynaPro Pate Reader II. *Wyatt Technology* 2012.

[33] STH, PC 950 Probe Colorimeter. *STH Company* USA, *www.sthcompany.com*.

CHAPTER 4

A manuscript of this work is already being prepared for Journal Submission. Part of the experimental work (zeta potential measurement) was performed in collaboration with Prof Davor Kovacevic and Dr Darija Jurasin.

4 A COMBINED STATIC LIGHT SCATTERING AND ZETA POTENTIAL ANALYSIS FOR SPECIFIC ION EFFECTS ON LYSOZYME ASSOCIATION

4.1 Abstract

Although there has been much progress recently in elucidating the molecular details of specific ion effects, the exact interaction mechanism between proteins and salts remain unclear. A systematic study of specific ion effects on lysozyme solutions using static, dynamic and electrophoretic light scattering was performed as a function of pH and ionic strength for various salts with different anions spanning the range of the Hofmeister series. Experiments reveal that ion binding at low ionic strength can be explained in terms of charge neutralization as sulphate ions are more effective at screening double layer forces. The binding of sulphate ions to positively charged lysozyme through electrostatic interactions is consistent with the electroselectivity theory. At intermediate ionic strength, thiocyanate and nitrate ions are more effective than sulphate at inducing protein-protein attraction. This implies that chaotropic anions induce short range attractive interactions between proteins of non-electrostatic origin. Hence, the salting-out effects of chaotropic anions which leads to the so-called reversal of the Hofmeister series cannot be only rationalized in terms of electrostatic interactions and changes to the double layer potential. The results provide the first estimate of the non-electrostatic contribution to protein-protein attraction in chaotropic anion solutions.

4.2 Introduction

There has been renewed interest in understanding the phenomena in colloid science involving electrolytes that exhibit pronounced ion specificity. Ion specific effects play an important role in modulating many biological processes such as precipitation, aggregation and phase transitions (such as gelation and crystallization) which are partly controlled by protein-protein interactions [1-4]. Studies date back to more than a decade ago where Franz Hofmeister showed a specific ion order according to their ability to precipitate proteins. Hofmeister studied the behaviour of ovalbumin in concentrated salt solutions and noted that the effectiveness of the salt ions in precipitating the protein increases with the hydration size of the ion and was ranked for both cations and anions in the Hofmeister series [5, 6].

Salts have long been used to control protein-protein interactions depending on the pH, salt type and ionic strength of the solution however, the exact mechanism of how salts alter protein-protein interactions remains unclear. Effects arising from electrostatics, hydrophobic forces, van der Waals interactions and hydrogen bonding may contribute to the overall protein-protein interaction, however the exact contribution of each type of interaction is not easily quantifiable. Thus, a useful starting point in understanding salt-mediated protein-protein interactions requires examining the biophysical properties of the proteins (net charge, hydrophobicity) and salt (size, valence, hydration, polarizability) [7]. Ion properties correlate with the Hofmeister series, which was originally developed to characterize the salting-out effectiveness of ions at relatively high concentrations ($\geq$ 0.3 M). Ions to the left side of the series salt-out proteins, and increase protein conformational stability when at high salt concentrations. These ions referred to as kosmotropes are strongly hydrated, which is linked to their ability to dehydrate protein surfaces and reduce

protein solubility. For anions, the order is given by $SO_4^{2-} > F^- > Cl^- > NO_3^- > Br^- > I^- > ClO_4^- > SCN^-$ and for cations $NH_4^+ > K^+ > Na^+ > Li^+ > Mg^{2+} > Ca^{2+}$ for cations. Towards the right side of the series, the anions are chaotropic and weakly hydrated, which is linked to their ability salt-in proteins or lower protein conformational stability. The phenomenon of salting in and salting out have been explained in details in Chapter 2 of this thesis. The effect of anions is more pronounced than cations due to the strong binding affinities of multivalent cations for polar parts of the protein surface, which counteracts their salting-out ability [5].

Measurements from zeta potential [8-10], differential scanning calorimetry [11], cloud-point temperature [12-14], static and dynamic light scattering [15] have revealed ion binding to opposite charged groups on the protein surface. Gokarn and co-workers [8] used effective charge measurements (i.e. zeta potential measurements) to probe specific ion interactions between lysozyme molecules in concentrations up to 0.1 M salt solutions. In solutions below the protein isoelectric point (pI), ion binding was found to follow the reverse Hofmeister series due to preferential accumulation of weakly hydrated monovalent anions at the protein surface. Ion binding here refers to the formation of ion pairs between oppositely charged groups driven by electrostatic interactions [15]. Collins proposed that the affinity for ion binding follows the law of matching water affinity where inner sphere ion pairs are formed between opposite charged ions that have matching absolute free energies of hydration [16, 17]. This implies that stronger ion pairs are formed between kosmotropes or between chaotropes rather than between a kosmotrope and chaotrope ion pair. Based on this model, the positively charged and weakly hydrated amines and amide nitrogen on proteins will interact preferentially with weakly hydrated (chaotropic) anions. Similarly, strongly hydrated cations have

strongest interactions with strongly hydrated and negatively charged carboxylates and carbonyl backbones [16-18]. Contrary to Collins model, strongly hydrated divalent cations have been reported to preferentially interact with weakly hydrated anions [8].

Ion specific effects also arise due to interactions with uncharged protein groups. Ion partitioning studies reveal ion specific salting in interactions due to ion binding around the peptide backbone, while salting out of nonpolar groups on the protein surface is linked to the ion specific surface-tension increment [2]. Depending on the concentration of salt solutions, molecular simulations show that chaotropic anions can display salting in or salting out effects by binding to amide moieties or by increasing the surface tension of nonpolar hydrophobic groups, respectively [19].

The effect of salts on protein-protein interactions is controlled by either specific or nonspecific interactions. With increasing salt concentration, protein-protein interactions become less repulsive due to ionic screening of the double layer forces. Screening is a non-specific effect that can be captured by treating ions as point charges when calculating the repulsive double layer force within DLVO theory, which has previously been used to capture protein behaviour at low salt concentrations ($\leq$ 0.1 M). DLVO theory is not able to account for ion specificity, which can arise when ions bind to within the Stern layer on the protein surface. Double layer forces are reduced due to ion binding as protein charged groups become neutralized. For instance, chaotropes have been observed to salt out positively charged proteins in the low salt concentration regime [13, 20], an effect which has been attributed to neutralizing double layer forces through ion binding. The binding affinity of monovalent anions to the protein generally increases with the anion position in the descending order of the Hofmeister series. However,

multivalent anions such as sulphate, which are high on the series, exhibit strong protein-ion binding as a result of stronger electrostatic interactions [21, 22] but are not as effective at salting-out as chaotropic anions when at intermediate salt concentration [23]. Thus, the effect of ion binding on protein-protein interactions can not only be rationalized in terms of the impact on double layer forces.

Investigating the role of protein-salt interactions on protein-protein interaction requires performing measurements as a function of salt type and ionic strength. Zhang and Cremer [12] used cloud-point temperature measurements to determine the effect of protein-salt interactions on protein-protein interactions and indirectly probed the extent of anion binding in lysozyme solutions as a function of salt type and concentration at pH 9.4. They found that lysozyme precipitation in salt solutions followed 2 different Hofmeister trends with respect to the anion. At low salt concentration (< ~ 0.2 M or 0.3 M), the anion salting-out effectiveness follows the reverse Hofmeister series, whereas the order reverts back to the direct series as the salt concentration is increased. Lysozyme is positively charged at this pH and as such, electrostatic interactions (i.e. double layer forces) are sufficiently repulsive at low ionic strength. Addition of chaotropic counter-ions causes the anions to associate with the positively charged surface leading to electrostatic neutralisation of the protein net charge, which is reflected by an increase in the cloud-point temperature. The finding of a reverse Hofmeister effect on protein solubility at low ionic strength is consistent for monovalent anions for other positively charged proteins such as fusion proteins [24, 25], bovine serum albumin [9, 10, 26] and monoclonal antibodies [23, 27, 28]. The experimental work has been supported by recent computational studies that have linked the preferential adsorption of anions to the protein surface with a reduction in double layer forces [29, 30]. At higher ionic

strengths, the order reverts to the direct Hofmeister series where the competing effect of surface tension is more pronounced than ion binding and electrostatic interactions become completely screened [12]. In that case, the contribution of protein-salt interactions to protein-protein interactions arises from exclusion of ions about the protein surface due to repulsive image forces [31].

Despite the progress in experimental and computational methods to understand the effect of salts on protein-protein interactions, it remains difficult to establish the physical origins of the ion specificity. Due to the heterogeneous nature of the protein surface (comprising of polar, nonpolar and charged side groups), proteins can interact with salt ions via different mechanisms [7, 29]. Recent studies have suggested that effects of ion binding on protein-protein interactions can not only be explained in terms of electrostatic interactions [32]. However, as yet, there have been no systematic studies done where ion binding and protein-protein interactions are measured under the same solution conditions (i.e. ionic strength, salt type, and temperature). The aim of this research work is to perform a detailed study of protein-protein interactions and ion-binding simultaneously using lysozyme as a model protein. Lysozyme, a basic protein with pI of ~ 11 [33] has frequently been used as a model system in understanding specific ion effects [4, 14, 34, 35]. Under acidic conditions below the pI, lysozyme is net positively charged through histidine, arginine and lysine residues. Specific ion effects are probed using static and dynamic light scattering to characterise protein-protein interactions (in terms of the osmotic second virial coefficient, B_{22} and an interaction parameter, k_D). Zeta potential measurements are used to characterise the extent of ion binding and also provide a direct measure of the electrical double layer force between proteins. By measuring the protein-protein interactions and protein-ion binding under the same solution

conditions and temperature, we are able to provide the first estimate of the non-electrostatic contribution to protein-protein attraction in chaotropic anion solutions.

4.3 Materials and Methods

Hen-egg white lysozyme (molecular weight 14.4 kDa) was purchased from Sigma-Aldrich and dissolved in about 30 mL of 10 mM potassium acetate (pH 4 or pH 5.5) or 10 mM Tris (pH 7) buffer solution to a working concentration of 25 mg/mL. The dissolved protein was dialyzed against 2 litres of the corresponding buffer for two hours each to minimise the effects of contamination by small amounts of ions in the lyophilised protein, which may influence protein interactions. Dialysis tubing with molecular weight cut-off of 8 kDa was obtained from Fisher Scientific. The pH of the stock buffers were adjusted with drops of either sodium hydroxide (NaOH) or hydrochloric acid (HCl) where necessary. After dialysis, the protein solution was filtered via a 0.22 um filter (Millipore, UK) and concentration was determined by measuring absorbance at 280 nm with a UV spectrophotometer (VWR International) and calculated using an extinction coefficient of 2.64 mL/mg cm [36] to yield a buffer exchanged stock protein solution at ~ 20 mg/mL. Lysozyme solutions were prepared a day before the experiment and stored at 4°C.

Stock salt solutions were prepared at a concentration of 500 mM by dissolving the salt of interest in 50 mL of desired buffer and filtered using a 0.20 um membrane filter (Millipore, UK). The pH of the solution was checked and adjusted if necessary. Protein sample solutions were prepared by dilution of the stock salt solution with stock lysozyme solution and/ or the corresponding buffer to a final volume of 10 mL. This method allows for the protein solution to contain the same ionic strength as the salt and prevents further dilution of the salts when preparing series of protein

concentrations. The total volume of salt solutions prepared at a fixed ionic strength was 20 mL. In static light scattering each experiment, a series of protein samples with concentrations ranging from 1 g/L to 10 g/L at fixed salt and buffer concentration were then prepared by mass dilutions of the stock protein solution with the corresponding salt solution.

For zeta potential measurements, only one lysozyme concentration prepared at 10 g/L was used since measurements were not conducted as a function of protein concentration. Protein-salt dilutions for the different salt types were made to a final volume of 5 mL.

All salts including potassium chloride (Fluka), potassium nitrate (Sigma-Aldrich), Tris (Formedium) were of analytical grade. Other reagents such as potassium sulphate, potassium thiocyanate, calcium chloride and potassium acetate were purchased from Fischer Scientific.

4.4 Methods

4.4.1 Static and Dynamic Light Scattering

Static and dynamic light scattering (SLS and DLS) measurements were simultaneously carried out using a miniDAWN Treos instrument (Wyatt, Santa Barbara, California) connected to an in-line UV concentration detector. The flow rate for the syringe pump was set to 0.2 mL/ minute. For each salt type and ionic strength, a series of samples with varying protein concentration were injected into the instrument starting with the most concentrated sample at 10 g/L. Buffer only samples containing the salts at a given ionic strength were used as a baseline control. The raw light scattering data was extracted from the software and analysed using an

excel spread sheet to obtain the osmotic second virial coefficient (B_{22}) and infinite dilution molecular weight ($M_{w,0}$). The excess Rayleigh ratios measured at 90° (R_θ) obtained from SLS readings were used to generate the Debye plots according to

$$\frac{Kc}{R_\theta} = \frac{1}{M_{w,0}} + 2B_{22}c \qquad \text{(Equation 4.1)}$$

where c is the protein concentration, and K is an optical constant given by

$$K = \frac{4\left[\pi n_0 \left(dn/dc\right)\right]^2}{N_A \lambda^4} \qquad \text{(Equation 4.2)}$$

where N_A is the Avogadro's constant (6.022×10^{23} mol^{-1}), λ is the wavelength of the incident light, n_0 is the solvent refractive index, and dn/dc is the differential refractive index increment.

For DLS data, ASTRA 6 software was used to determine the intensity auto-correlation function where the diffusion coefficient was extracted using the cumulant analysis method. The translational diffusion coefficient (D_T) was plotted against protein concentration (c) to obtain the diffusion interaction parameter (k_D) derived from the slope according to

$$D_T = D_0\left(1 + k_D c\right) \qquad \text{(Equation 4.3)}$$

where D_0 is the infinite dilution of the diffusion coefficient at zero concentration, D_0 is related to an intensity weighted average hydrodynamic radius, R_h, through the Stokes-Einstein relationship

$$R_h = \frac{k_B T}{6\pi\eta D_0} \qquad \text{(Equation 4.4)}$$

where k_B is Boltzmann constant, η is solvent viscosity and T is absolute temperature.

Errors were calculated based on the standard deviation obtained from linear regression of Equation 4.1 and Equation 4.3.

4.4.2 Zeta Potential Measurements

The zeta potential of lysozyme was measured using the Zetasizer NanoZS instrument by phase analysis light scattering method (M3-PALS) (Malvern Instruments Ltd., UK). The instrument is equipped with a 532 nm green laser source. Scattered light intensities were detected at an angle of 173°. Samples were measured with a universal dip cell (ZEN 1002) suitable for high conductivity samples so as to prevent electrode or sample degradation. The dip cell was used in conjunction with a disposable 1 mL polystyrene cuvette (DTS0012). Measurements were carried out at a temperature of 25.0 ± 0.1 °C in automatic duration mode. For each sample, 10 measurements each of 10 seconds duration were taken and averaged for a single run. The Dispersion Technology Software was used to calculate the zeta potentials (ζ) of lysozyme through the electrophoretic mobility (μ_E) using the Henry equation

$$\zeta = \frac{3\mu_E \eta}{2\varepsilon\varepsilon_0 f(\kappa R_h)} \qquad \text{(Equation 4.5)}$$

where R_h is particle radius, $f(\kappa R_h)$ is Henry's function equal to 1, η is solution viscosity and ε is dielectric constant of water and ε_o is the permittivity of vacuum (8.85 x 10^{-12} C^2/ (Nm^2). The effective charge was calculated from electrophoretic mobility through the relation

$$Z_{eff} = \frac{\mu_E 6\pi\eta R_h (1+\kappa R_h)}{ef(\kappa R_h)} \qquad \text{(Equation 4.6)}$$

where e is unit charge (1.602 x 10^{-19} C), Z_{eff} is effective number of unit charge and κ is the inverse of the Debye Huckel screening length given by

$$\kappa = \sqrt{\frac{2e^2 N_A I}{\varepsilon\varepsilon_0 k_B T}} \qquad \text{(Equation 4.7)}$$

where N_A is Avogadro number, k_B is Boltzmann constant (1.38 x 10^{-23} J/K), T is temperature (in kelvin) and I is ionic strength.

Before running experiments, lysozyme solutions prepared at a final concentration of 10 g/L were filtered using a 0.22 um filter membrane to reduce the presence of small particulates. Between measurements, the cell was thoroughly cleaned with distilled water and gently sonicated for about 10 seconds in order to remove particles adsorbed on the surface of the electrodes. All samples were run in triplicates and error bars represent the standard deviation over three independent runs.

4.5 Results and Discussion

We have investigated the protein-protein interactions in terms of the osmotic second virial coefficient (B_{22}) from static light scattering and an interaction parameter (k_D) derived from dynamic light scattering. Protein-salt interactions have been characterized from zeta potential measurements. The measurements have been carried out for the salts $CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN as a function of ionic strength for pH values equal to 4, 5.5, and 7.

4.5.1 Effect of pH and ionic strength of protein-protein interactions

Figure 4.1 shows the effect of changes in pH and ionic strength on B_{22} and k_D values for lysozyme in potassium chloride solutions. Positive values of B_{22} (SLS) and k_D (DLS) reflect repulsive protein-protein interactions while negative values indicate attractive interactions. B_{22} and k_D follow the same qualitative trends with respect to pH and ionic strength. The behaviour of proteins at low ionic strengths (< 100 mM) can be described using the Derjaguin–Landau–Verwey–Overbeek (DLVO) theory of

colloidal interactions where interactions are controlled by long-range electrostatic interactions [37]. At pH values below the pI, lysozyme is strongly positive charged as the basic residues are ionised while the acidic residues are protonated. DLVO theory can be used to explain the pH and ionic strength dependence of the averaged protein-protein interactions for lysozyme in potassium chloride solutions. As the pH increases at a fixed ionic strength, both interaction parameters decrease due to weaker charge-charge repulsion as the pH is approaching the pI hence B_{22} and k_D are lowest at pH 7. For instance, interactions in 25 mM KCl solutions are less repulsive at pH 7 (where B_{22} is ~ 4 x 10^{-4} mol mL/g^2) compared to pH 4 and 5.5 with B_{22} on the order of 15 x 10^{-4} mol mL/g^2 and 16 x 10^{-4} mol mL/g^2 respectively.

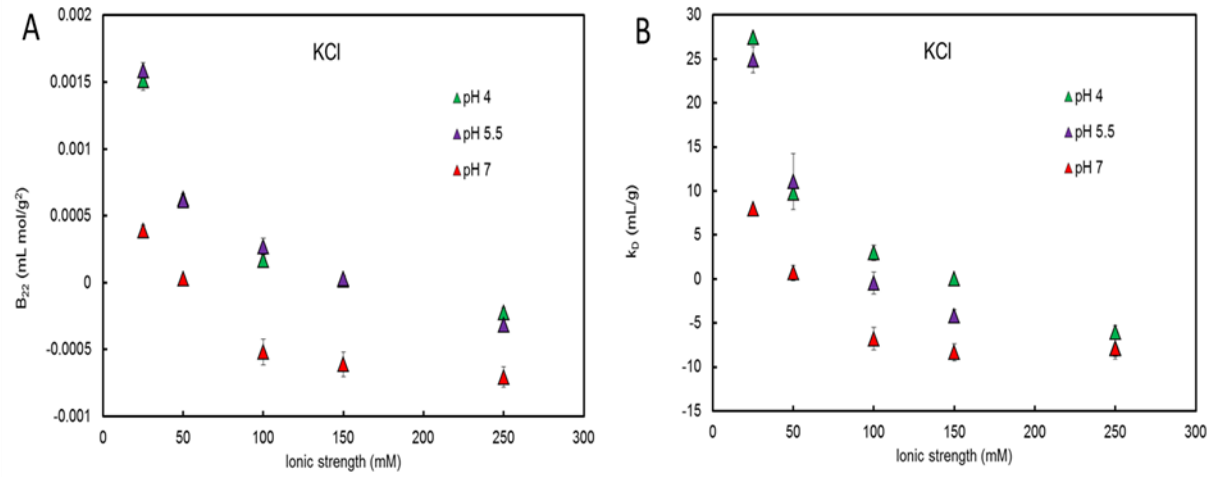


Figure 4.1. (A) Static and (B) dynamic light scattering measurements of lysozyme in potassium chloride solutions as a function of pH and ionic strength in terms of B_{22} and k_D respectively.

As the ionic strength increases, B_{22} and k_D values decrease and change sign from positive to negative. For instance, the value of k_D between 25 mM to 250 mM changes from approximately 26 mL/g to -6 mL/g for pH 4 and 5.5, and at pH 7 changes from 8 mL/g to -8 mL/g. Increasing ionic strength leads to charge neutralisation and ionic screening which both reduce electrical double-layer repulsion. The ionic screening effect is due to compression of the double layer, where the size is given by the debye screening length (κ^{-1}), which is reduced from 3.04 nm to 0.96 nm in 10 mM and 100 mM salt solutions respectively. The values of B_{22} and k_D at pH 4, 5.5 and 7 in 250 mM KCl solutions are all negative indicating the presence of significant short-range attractive (nonelectrostatic) interactions as the double layer becomes compressed. A similar effect of pH and ionic strength on lysozyme-lysozyme interactions have previously been observed by Neal et al. [38] and indirectly through cloud point studies by Grigsby et al. [13] .

4.5.2 Specific ion or Hofmeister effects on protein-protein interactions

The specific salt type contributes significantly towards determining the net protein-protein interactions. In this study, the anion in the salt was systematically changed while the cation was fixed (potassium ion), and the cation was changed while the anion was fixed (chloride ion). Figure 4.2 to 4.4 shows trends for lysozyme in different salt solutions ($CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN) as a function of ionic strength at pH 4, 5.5 and 7 respectively. The overall behaviour of lysozyme molecules characterised from k_D is consistent with B_{22} data. At all pH values, B_{22} and k_D decrease with increasing ionic strength irrespective of the nature of the salt with the most attractive interactions occurring at pH 7. The magnitude of decrease strongly depends on the specific ion type. Salt specific effects at low pH can be

explained in terms of anion binding, which is different from ionic screening caused by a change in the ionic strength of the solution.

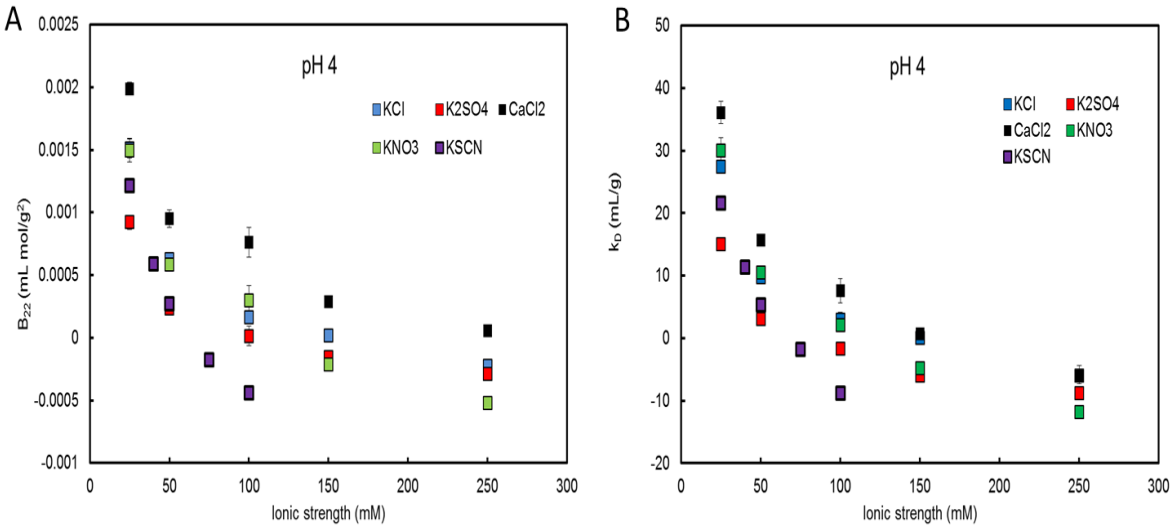


Figure 4.2. Measurements of (A) B_{22} and (B) k_D of lysozyme at pH 4 in different salts solutions ($CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN) represented by colour codes between an ionic strength of 25 mM to 250 mM.

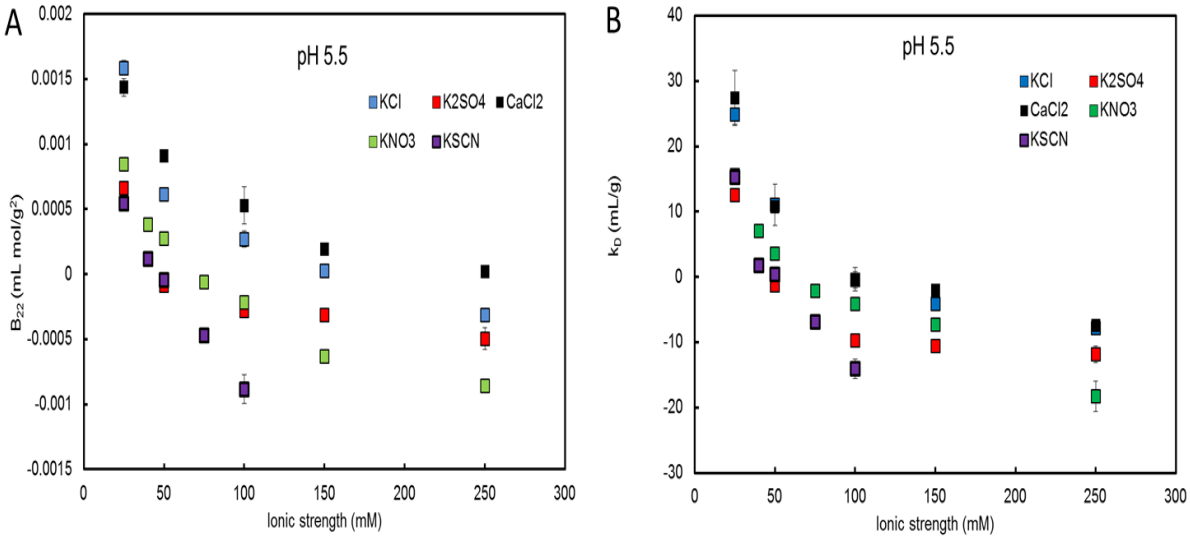


Figure 4.3. Measurements of (A) B_{22} and (B) k_D of lysozyme at pH 5.5 in different salts solutions ($CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN) represented by colour codes between an ionic strength of 25 mM to 250 mM.

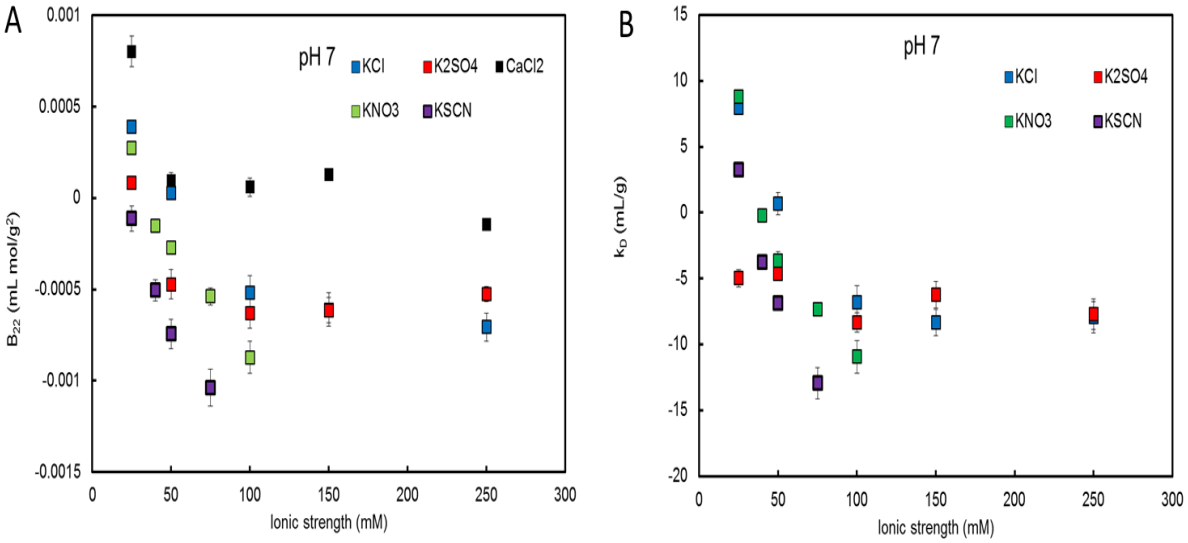


Figure 4.4. Measurements of (A) B_{22} and (B) k_D of lysozyme at pH 7 in different salts solutions ($CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN) represented by colour codes between an ionic strength of 25 mM to 250 mM.

In low ionic strength solutions (< 50 mM), the main contributions to protein-protein interactions are from electric double layer forces. Specific ion effects are apparent when comparing protein-protein interactions at the same ionic strength such that the effect of ionic screening is the same. The specific ion effects are attributable to charge neutralization; ions with higher binding affinity are more effective at reducing double layer forces. In low ionic strength solutions, protein-protein interactions are most repulsive in Cl^- solutions. Cl^- weakly binds to the protein at low ionic strength compared to the other anions (SO_4^{2-}, NO_3^- and SCN^-). At low ionic strength, sulphate ions (SO_4^{2-}) are the most effective at reducing the double layer repulsion. This behaviour is consistent with the ion binding affinity of sulphate to protein charged groups, which is greater than for any monovalent anions [8, 9, 19, 39]. It appears that the only effect of sulphate binding to the protein is to alter the electrostatic interactions because the B_{22} values at high ionic strength are the same for potassium chloride and for potassium sulphate solutions. This behaviour follows what would be expected for a charge neutralization mechanism where the only effect of ion binding is to alter the protein fixed charge distribution. Ramos and Baldwin [22] studied the effect of sulphate and chloride ions on the stabilisation of ribonuclease A (RNase A) and found that sulphate stabilised RNase A via specific ion binding while chloride did not stabilise the protein at low ionic strength.

For monovalent anions, we find the ability of the anion to neutralize lysozyme charge follows the order of the electroselectivity series, which is the inverse of the Hofmeister series in the order $SCN^- > NO_3^- > Cl^-$. At low pH, the chaotropes form strong ion pairs with the positively charge residues on the protein due to their large size and smaller charge density [5, 18]. Thus, weak salting out agents can

preferentially interact with proteins and enhance protein attraction, possibly through neutralizing double-layer forces.

The effect of cation on protein-protein interactions was investigated by substituting potassium (K^+) with calcium (Ca^{2+}) ions in chloride solutions. B_{22} and k_D values are higher in presence of Ca^{2+} indicating increased electrostatic repulsion between lysozyme molecules. Binding of divalent Ca^{2+} to proteins has been reported in a number of studies [15, 40] where increases in protein solubility arise from calcium binding due to repulsive hydration forces. At low ionic strength, the binding of Ca^{2+} will also lead to increased electrostatic repulsion. Conversely, at pH far below the pI, the divalent Ca^{2+} is excluded from around proteins compared to the K^+. Exclusion of the cation will also lower anion binding to protein leading to an increase in electrostatic repulsion [8].

An interesting cross-over effect is seen at intermediate ionic strengths (50 mM) at pH 4 where protein-protein interactions become more attractive in solutions of SCN^- versus those containing SO_4^{2-}. As the ionic strength of the solution is raised further to 250 mM, protein-protein interactions become more attractive in nitrate solutions versus sulphate containing solutions. The reverse Hofmeister series dependence of lysozyme solubility has been previously observed by some other authors [8, 10, 12, 28-30] where the effectiveness of the ions in salting out lysozyme follows the reverse Hofmeister order of $Cl^- < SO_4^{2-} < NO_3^- < SCN^-$ for acidic pH values. Here, 75 mM KSCN solutions are seen to be most effective at salting out lysozyme at pH 7 compared to the other salts. Measurements were not conducted for KSCN at ionic strengths greater than 100 mM due to sample turbidity resulting from lysozyme precipitation.

It is not clear whether or not the cross-over effect can be rationalized in terms of ion binding and charge neutralization. This would require that there are a larger number of chaotropic anion binding sites versus sulphate binding sites, although, the former would be at a weaker affinity. However, the protein-protein interactions in chaotropic anion solutions are more attractive than in sulphate solutions even at higher ionic strength (i.e. 250 mM) where electrostatic interactions are significantly screened, so that the effect of neutralizing charge should be small. Thus, the results suggest that non-electrostatic interactions between proteins are also altered by chaotropic anions.

Link between B_{22} and k_D

A few studies have demonstrated a monotonic correlation between B_{22} and k_D for proteins such as monoclonal antibodies [15, 32, 41-43]. The relationship between B_{22} and k_D has been described by Lehermayr et al. [43] where the correlation line of the plot of B_{22} versus k_D was used to derive an empirical equation ($k_D = 1.06 B_{22}M - 8.9$) for eight different monoclonal antibodies. Connolly et al. [41] showed the linear dependence of k_D with B_{22} obtained in separate measurements where the experimental slope is obtained from the equation $k_D = 1.33\ B_{22}M - 8.2$. Figure 4.5 shows the qualitative rank correlation of B_{22} with k_D for lysozyme solutions in different salt solutions at pH 4. The results demonstrate that the values obtained independently for B_{22} and k_D have a strong linear dependence and can be described by $k_D = 1.31\ B_{22}M - 1.9$ supporting the hypothesis that k_D and B_{22} are proportional measures of protein-protein interactions.

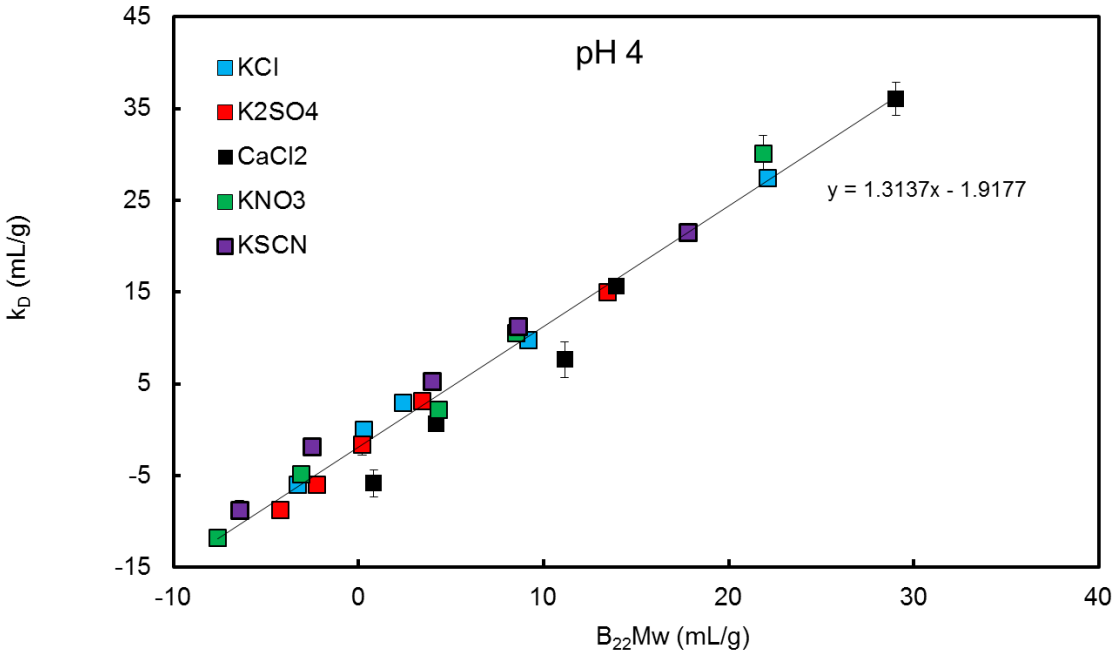


Figure 4.5. Correlation between k_D and B_{22} measurements in lysozyme solutions at 25 mM to 250 mM ionic strength for the different salts ($CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN) at pH 4

In summary, the strong dependence on Hofmeister series indicates DLVO theory cannot be used only to describe the behaviour of lysozyme in very dilute salt solutions. The results observed at ionic strengths above or below 100 mM solutions seem to indicate that at least two different ion specific mechanisms are responsible for controlling protein - protein interactions. At low ionic strength, the ability of the ion to reduce protein-protein repulsion appears to be related to charge neutralization from ion binding which is consistent with previous studies that have demonstrated ion binding affinity follows the electroselectivity series, which agrees with the trends observed at low ionic strength here. However, the strong attractive protein-protein interactions observed in thiocyanate solutions, and to a lesser extent in nitrate solutions, appears to indicate there is an additional effect of ion binding on protein-protein attraction, which cannot be explained by charge neutralization. In order to

establish the link between protein-protein interactions and ion binding, we have carried out zeta potential measurements in the next section, which provide a direct measure of ion binding and the impact on electrostatic interactions between proteins.

4.5.3 Lysozyme Hofmeister effects: zeta potential and net charge

A series of zeta potential experiments were carried out under the same solvent conditions as used for the B_{22} and k_D measurements. The zeta potential (ζ-potential) is the electrostatic potential at the slipping plane located outside the stern layer and provides an estimate for the electrical diffuse layer potential, which in turn, determines the magnitude of the double layer force between proteins. The zeta potential reflects any specific ion binding within the Stern layer. In this study, the measured potentials were less than 25 mV making it valid to use the Henry's equation to relate the zeta-potential to the electrophoretic mobility (see Equation 4.5). Figure 4.6 A to D shows the measured ζ- potentials calculated from electrophoretic mobility (μ_E) for 10 g/L lysozyme solutions over a range of ionic strength (10 mM to 100 mM) at pH 4, 5.5, and 7 for all salts ($CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN).

At such low ionic strengths (10 mM to 100 mM), there were observable differences in the zeta-potential depending on the type of salt. Generally, zeta potentials were positive and slightly decreased as the pH is raised from 4 to 7 due to a reduction in the net positive charges arising from ionisable protein groups.

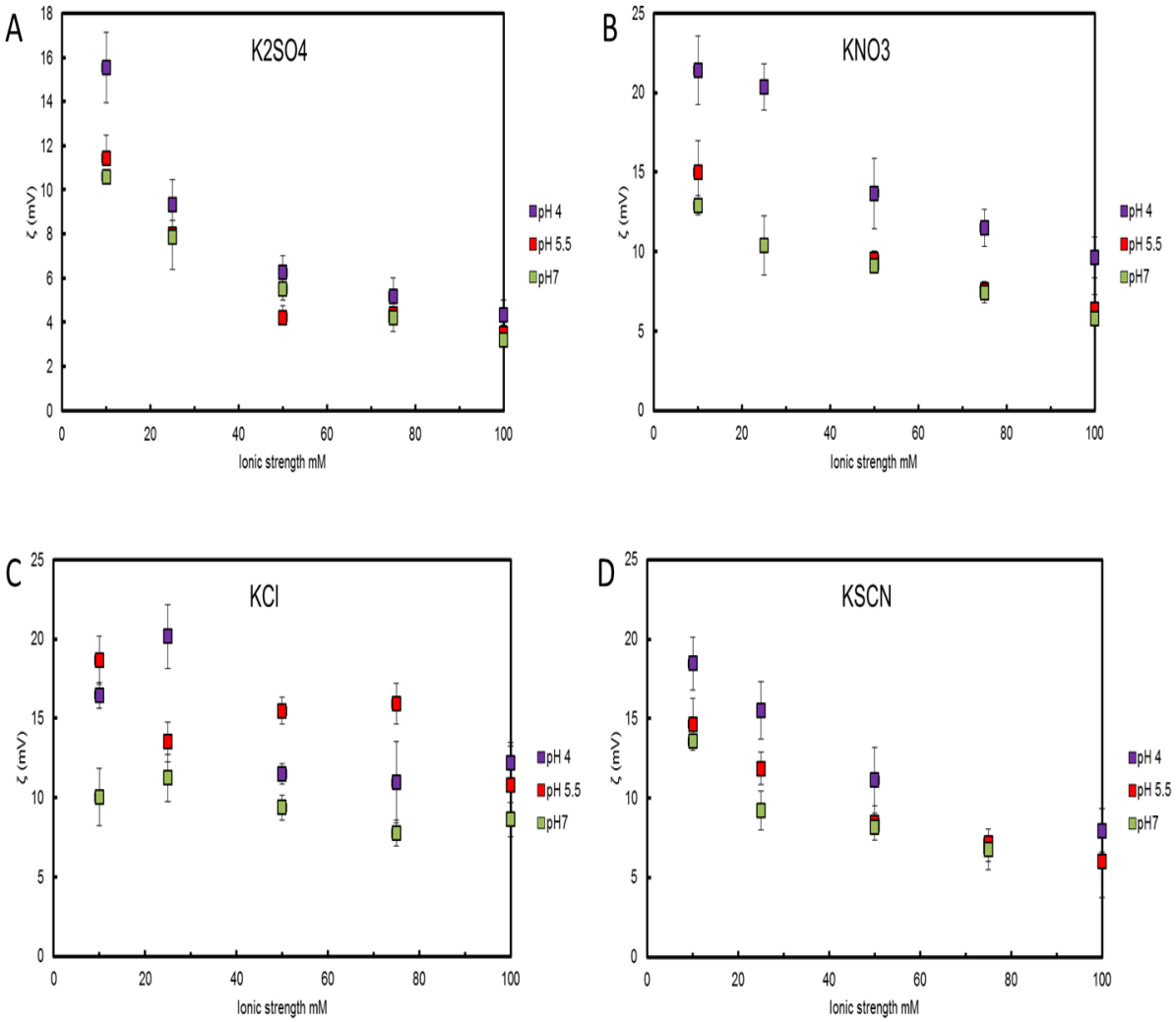


Figure 4.6. Zeta potential of lysozyme as a function of pH and ionic strength in (A) K_2SO_4, (B) KNO_3, (C) KCl and (D) KSCN solutions. Error bars represent standard deviations of 3 independent measurements.

With the exception of chloride ions, the ζ- potential decreased with increasing ionic strength. Values of ζ- potential are seen to depend on the type of salt at a fixed pH and ionic strength. Figure 4.7 shows the ζ- potential of lysozyme in different salt solutions as a function of ionic strength at pH 5.5.

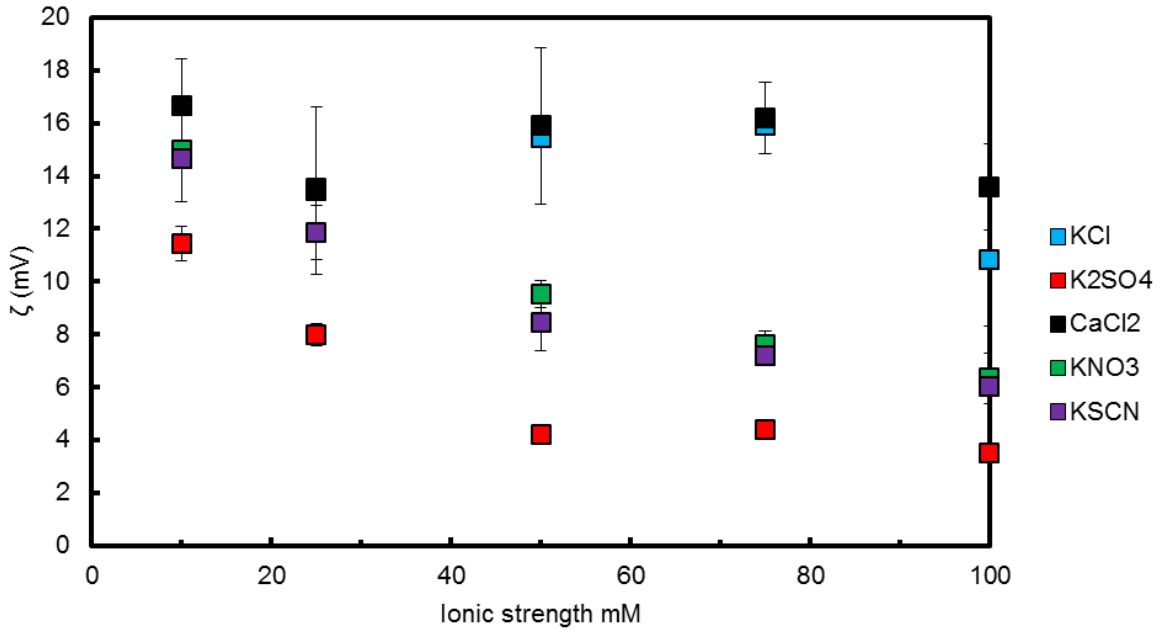


Figure 4.7. Zeta potential of lysozyme as a function of salt type and ionic strength at pH 5.5. Error bars represent standard deviations of 3 independent measurements

Figure 4.8 shows the Z_{eff} of lysozyme at pH 4, 5.5 and 7 for solutions at 100 mM ionic strength containing $CaCl_2$, KCl, K_2SO_4, KNO_3 and KSCN salts. The experimental values of Z_{eff} calculated from the electrophoretic mobility are smaller than that of theoretical calculations from derived pKa values or hydrogen-ion titrations which do not account for charges arising from bound ions in the Stern layer [44]. The finding that the effective charge calculated from zeta potential is less than the theoretical values is consistent with many other studies on different monoclonal antibodies [43, 45, 46]. Since the positively charged lysozyme has affinity for negatively charged ions therefore it is likely that any reduction in Z_{eff} is due to protein-ion binding. It is also possible that approximations used in relating the experimental Z_{eff} to electrophoretic mobility are based on simplifying assumptions that may not describe the true electrostatic potential of proteins. For example, the exact boundary or hydrodynamic shear of plane between the stern layer and diffuse

layer where the zeta potential is measured is not experimentally clear. In addition, Eq. 4.6 used to determine the experimental net charge is made on the assumption that the Stokes-Einstein equation is valid at zero protein concentration and as such assumes a constant diffusion coefficient irrespective of the ionic strength.

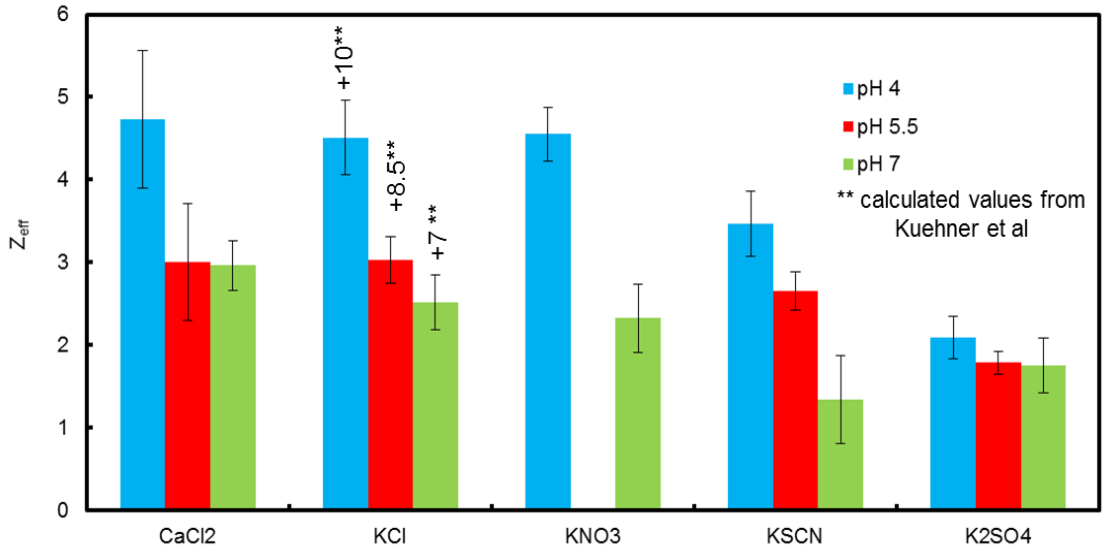


Figure 4.8. Effective charge (Z_{eff}) of lysozyme at an ionic strength of 100 mM as a function of pH and salt type. Error bars represent standard deviations of 3 independent measurements.

The differences in ζ- potentials at the same ionic strength can be attributed to anion specific adsorption. The divalent anion SO_4^{2-} causes a greater reduction in zeta potential compared to monovalent anions (Cl^-, NO_3^-, and SCN^-), which is consistent with its higher binding affinity for protein charged groups [21]. As mentioned previously, the enhanced binding is due to the greater number of valence on SO_4^{2-} and is consistent with the measured B_{22} values at low ionic strength providing support that sulphate binding neutralizes double layer forces. For monovalent anions, the zeta potentials of the ions behaved according to the reverse Hofmeister series

where highly polarisable SCN^- adsorbs to the protein more strongly than Cl^- which is less polarisable. In addition, it seems that Cl^- behaves as a neutral salt and as such does not exhibit changes in ζ-potential even as the ionic strength is increased. Studies have demonstrated that ζ-potential of KCl is independent salt concentration [9, 47, 48]. Thus, the strength of protein-ion binding at low pH in solution between 10 mM and 100 mM ionic strength follows the order $SO_4^{2-} > SCN^- > NO_3^- > Cl^-$, which is the same as the electroselectivity series, which was independently derived to characterize ion binding affinity to protein surfaces [21]. Perhaps, more significantly, no cross-over in the zeta-potential measurements is observed in the ionic strength trends for sulphate versus thiocyanate.

It is interesting to see that the charge of lysozyme is greatest in presence of divalent Ca^{2+} ions compared to the anions. The net charge of lysozyme is slightly higher for $CaCl_2$ solution compared to KCl as shown in Figure 4.9. It is likely that Ca^{2+} reduces the strength of Cl^- binding compared to K^+ because of its divalent nature thereby reducing electrostatic attraction. Alternatively, if Ca^{2+} is excluded from the protein surface, then it is possible that it would draw the chloride ions into the bulk solution. Gokarn et al. [8] suggested that the cations do not interact with the protein surface as there is no strong dependence of the cation identity on the chloride ion and as such, protein–ion binding is restricted to anions.

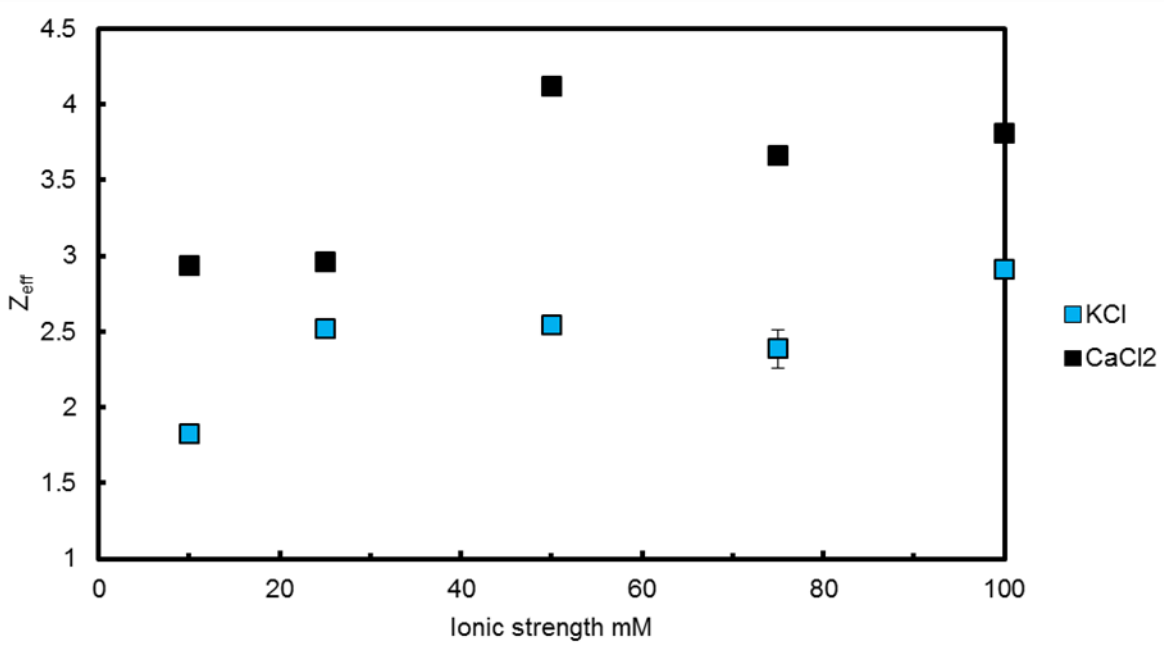


Figure 4.9. Effective charge (Z_{eff}) of lysozyme in $CaCl_2$ and KCl salt solutions at pH 7 as a function of ionic strength. Error bars represent standard deviations of 3 independent measurements.

Relationship between B_{22} and net charge

Protein interaction measurements obtained from B_{22} and k_D data reveal that SO_4^{2-} lowers the electrostatic repulsion of lysozyme compared to the other anions (SCN^-, NO_3^- and Cl^-) in solutions with ionic strengths ≤ 50 mM. This was confirmed by ζ-potential measurements which showed that the divalent anions reduced the effective net charge of lysozyme. The effectiveness of the salts in reducing protein net charge and B_{22} at ionic strengths ≤ 50 mM follows the order $CaCl_2$ < KCl < KNO_3 < KSCN < K_2SO_4, which is the same as the electroselectivity series, which quantifies the anion binding affinity for positively charged resins. Thus, the specific ion effects at low ionic strength (i.e. 25 mM) can be rationalized in terms of ion binding and charge neutralization. Furthermore, the net charge as assessed by zeta-potential is in good agreement with the colloidal stability assessed in terms of B_{22}. This observation

is consistent with measurements performed by Lehermayr et al. [43] where at low ionic strength, B_{22} increased with increasing net charge.

A cross-over is seen at ionic strengths ≥ 50 mM in B_{22} results where chaotropic SCN^- and NO_3^- are more effective in lowering electrostatic repulsion than SO_4^{2-}. This cross-over effect is not seen in zeta potential and charge measurements as SO_4^{2-} still binds more strongly to the positively charged lysozyme and is more effective at lowering the double layer potential or similarly the double layer force. For ionic strengths greater than 50 mM, although the double layer force is lower in SO_4^{2-} solutions, protein-protein interactions are more attractive in SCN^-. In this case, the net charge does not correlate with the B_{22} values and the salting out effectiveness exhibited by chaotropic anions at moderate ionic strength cannot be only attributed to anion binding and neutralization of double layer forces. The salt-induced attraction observed in thiocyanate solutions must arise from a non-DLVO type interaction. We also find the cross-over effect in nitrate solutions occurs at higher ionic strengths where electrostatic interactions are screened. This evidence suggests that chaotropic anion induced attraction is not of an electrostatic nature, but more likely related to changes in short-ranged protein-protein attraction. The molecular origin of protein-protein interactions in moderately concentrated salt solutions is poorly understood, so that it is difficult to interpret the exact mechanism for the chaotropic-anion induced attraction. Curtis et al. [4] explained that upon anion binding the surface chemistry of lysozyme is altered depending on the hydration of the bound anion. Hence, the surface free energy is greater when SCN^- (strongly chaotropic) binds as the water layer becomes less structured, whereas when a kosmotropic anion such as SO_4^{2-} (kosmotropic) binds, the water layer is more structured possibly leading to so-called hydration forces. Another distinct possibility

is that, due to the chaotropic nature of thiocyanate, ion binding causes local protein unfolding leading to exposure of sticky non-polar groups [32]. This hypothesis is consistent with other studies that have found thiocyanate increases local protein flexibility upon binding [11, 49].

4.6 Conclusion

The presented study has shown that the intermolecular interactions between lysozyme molecules at low salt concentrations are controlled by specific ion effects. Although a number of measurements have shown specific ion effects in the low concentration regime, a detailed study on ion specificity has not been reported. With this in mind, the focus of this study was to provide a complete picture of the influence of different salt types on protein interactions in lieu of a better understanding of the mechanisms controlling protein-protein interactions. The combined use of static, dynamic and electrophoretic light scattering measurements allowed for investigating the role of ion binding, debye screening and Hofmeister effect in the pH range studied. The results in this chapter are among the most extensive set of measurements describing the influence of salts on lysozyme interactions.

For each salt type, the measured parameters were found to decrease with increasing ionic strength and pH. The magnitude of attractive or repulsive interactions was also observed to be ion specific. Rationalizing ion specificity requires understanding the effect of the different salt types at the same ionic strength. The origin of the ion specificities seemed to depend on a number of factors such as size, valence and hydration. From our results, we propose that the effectiveness of the salt to affect protein-protein interactions follows two different mechanisms depending on the anion type and ionic strength. First at relatively low ionic strengths where

interactions are of electrostatic origin, the effectiveness of the anion at reducing protein-protein repulsion follows the electroselectivity series. In this scenario, protein-protein interactions are controlled by charge neutralisation arising from ion binding which lowers the electrical double layer. Divalent SO_4^{2-} is the most effective because it displays stronger electrostatic interactions with the positively charged side chains on the surface of the protein compared to the monovalent anions. On the other hand, the cross-over effect observed where interactions are more attractive in thiocyanate and nitrate versus sulphate containing solutions indicates charge neutralization cannot be used to solely describe protein-protein attraction as there is some additional short ranged contribution to the protein-protein attraction. Hence, we observe that the effectiveness of the anions to induce protein-protein association follow the reverse Hofmeister series. This shows that the complexity of specific ion or Hofmeister effects requires better interpretation of ion binding which would account not only for electrostatic theories but ionic dispersion forces.

We hypothesize that the additional protein-protein attraction found in thiocyanate solutions, which cannot be explained by how ion binding alters electrostatic interactions stems from the fact that chaotropic anions can cause proteins to unfold by interacting with polar/non-polar side chains. Thus, the pH dependence appears to indicate the chaotropic anion binding is modulated by the charge on the protein, but when the anions bind, they could cause local protein unfolding and exposure of sticky hydrophobic groups. Therefore the additional attraction could arise because the protein has some local unfolding.

4.7 References

[1] George, A., Wilson, W. W., Predicting protein crystallization from a dilute solution property. *Acta Crystallographica Section D* 1994, *50*, 361-365.

[2] Baldwin, R. L., How Hofmeister ion interactions affect protein stability. *Biophysical journal* 1996, *71*, 2056-2063.

[3] Curtis, R. A., Prausnitz, J. M., Blanch, H. W., Protein-protein and protein-salt interactions in aqueous protein solutions containing concentrated electrolytes. *Biotechnology and Bioengineering* 1998, *57*, 11-21.

[4] Curtis, R. A., Ulrich, J., Montaser, A., Prausnitz, J. M., Blanch, H. W., Protein–protein interactions in concentrated electrolyte solutions. *Biotechnology and Bioengineering* 2002, *79*, 367-380.

[5] Kunz, W., Specific ion effects in colloidal and biological systems. *Current Opinion in Colloid & Interface Science* 2010, *15*, 34-39.

[6] Hofmeister, F., Zur Lehre von der Wirkung der Salze. *Archiv f. experiment. Pathol. u. Pharmakol* 1888, *24*, 247-260.

[7] Zhang, J., Protein-Protein Interactions in Salt Solutions, in: Weibo Cai, H. H. (Ed.), *Protein-Protein Interactions - Computational and Experimental Tools* 2012.

[8] Gokarn, Y. R., Fesinmeyer, R. M., Saluja, A., Razinkov, V., *et al.*, Effective charge measurements reveal selective and preferential accumulation of anions, but not cations, at the protein surface in dilute salt solutions. *Protein Science* 2011, *20*, 580-587.

[9] Salgın, S., Salgın, U., Bahadır, S., Zeta Potentials and Isoelectric Points of Biomolecules: The Effects of Ion Types and Ionic Strengths. *Int. J. Electrochem. Sci.* 2012, *7*, 12404 - 12414.

[10] Medda, L., Barse, B., Cugia, F., Boström, M., *et al.*, Hofmeister Challenges: Ion Binding and Charge of the BSA Protein as Explicit Examples. *Langmuir* 2012, *28*, 16355-16363.

[11] Majumdar, R., Manikwar, P., Hickey, J. M., Samra, H. S., *et al.*, Effects of Salts from the Hofmeister Series on the Conformational Stability, Aggregation Propensity, and Local Flexibility of an IgG1 Monoclonal Antibody. *Biochemistry* 2013, *52*, 3376-3389.

[12] Zhang, Y., Cremer, P. S., The inverse and direct Hofmeister series for lysozyme. *Proceedings of the National Academy of Sciences* 2009, *106*, 15249-15253.

[13] Grigsby, J. J., Blanch, H. W., Prausnitz, J. M., Cloud-point temperatures for lysozyme in electrolyte solutions: effect of salt type, salt concentration and pH. *Biophysical Chemistry* 2001, *91*, 231-243.

[14] Broide, M. L., Tominc, T. M., Saxowsky, M. D., Using phase transitions to investigate the effect of salts on protein interactions. *Physical Review E* 1996, *53*, 6325-6335.

[15] Roberts, D., Keeling, R., Tracka, M., van der Walle, C. F., *et al.*, Specific Ion and Buffer Effects on Protein–Protein Interactions of a Monoclonal Antibody. *Molecular Pharmaceutics* 2014, *12*, 179-193.

[16] Collins, K. D., Ion hydration: Implications for cellular function, polyelectrolytes, and protein crystallization. *Biophysical Chemistry* 2006, *119*, 271-281.

[17] Collins, K. D., Neilson, G. W., Enderby, J. E., Ions in water: Characterizing the forces that control chemical processes and biological structure. *Biophysical Chemistry* 2007, *128*, 95-104.

[18] Collins, K. D., Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Macromolecular Crystallization* 2004, *34*, 300-311.

[19] Rembert, K. B., Paterová, J., Heyda, J., Hilty, C., *et al.*, Molecular Mechanisms of Ion-Specific Effects on Proteins. *Journal of the American Chemical Society* 2012, *134*, 10039-10046.

[20] Ries-Kautt, M. M., Ducruix, A. F., Relative effectiveness of various ions on the solubility and crystal growth of lysozyme. *Journal of Biological Chemistry* 1989, *264*, 745-748.

[21] Goto, Y., Takahashi, N., Fink, A. L., Mechanism of acid-induced folding of proteins. *Biochemistry* 1990, *29*, 3480-3488.

[22] Ramos, C. H. I., Baldwin, R. L., Sulfate anion stabilization of native ribonuclease A both by anion binding and by the Hofmeister effect. *Protein Science* 2002, *11*, 1771-1778.

[23] Zhang-van Enk, J., Mason, B. D., Yu, L., Zhang, L., *et al.*, Perturbation of Thermal Unfolding and Aggregation of Human IgG1 Fc Fragment by Hofmeister Anions. *Molecular Pharmaceutics* 2013, *10*, 619-630.

[24] Saluja, A., Crampton, S., Kras, E., Fesinmeyer, R. M., *et al.*, Anion Binding Mediated Precipitation of a Peptibody. *Pharmaceutical Research* 2009, *26*, 152-160.

[25] Gokarn, Y. R., Fesinmeyer, R. M., Saluja, A., Cao, S., *et al.*, Ion-specific modulation of protein interactions: Anion-induced, reversible oligomerization of a fusion protein. *Protein Science* 2009, *18*, 169-179.

[26] Chen, X., Flores, S. C., Lim, S.-M., Zhang, Y., *et al.*, Specific Anion Effects on Water Structure Adjacent to Protein Monolayers. *Langmuir* 2010, *26*, 16447-16454.

[27] Fesinmeyer, R. M., Hogan, S., Saluja, A., Brych, S., *et al.*, Effect of Ions on Agitation- and Temperature-Induced Aggregation Reactions of Antibodies. *Pharmaceutical Research* 2009, *26*, 903-913.

[28] Zhang, L., Tan, H., Fesinmeyer, R. M., Li, C., *et al.*, Antibody solubility behavior in monovalent salt solutions reveals specific anion effects at low ionic strength. *Journal of pharmaceutical sciences* 2012, *101*, 965-977.

[29] Schwierz, N., Horinek, D., Netz, R. R., Reversed Anionic Hofmeister Series: The Interplay of Surface Charge and Surface Polarity. *Langmuir* 2010, *26*, 7370-7379.

[30] Boström, M., Parsons, D. F., Salis, A., Ninham, B. W., Monduzzi, M., Possible Origin of the Inverse and Direct Hofmeister Series for Lysozyme at Low and High Salt Concentrations. *Langmuir* 2011, *27*, 9504-9511.

[31] Curtis, R. A., Lue, L., A molecular approach to bioseparations: Protein–protein and protein–salt interactions. *Chemical Engineering Science* 2006, *61*, 907-923.

[32] Roberts, D., Keeling, R., Tracka, M., van der Walle, C. F., *et al.*, The Role of Electrostatics in Protein–Protein Interactions of a Monoclonal Antibody. *Molecular Pharmaceutics* 2014, *11*, 2475-2489.

[33] Velev, O. D., Kaler, E. W., Lenhoff, A. M., Protein Interactions in Solution Characterized by Light and Neutron Scattering: Comparison of Lysozyme and Chymotrypsinogen. *Biophysical Journal* 1998, *75*, 2682-2697.

[34] Moon, Y. U., Curtis, R. A., Anderson, C. O., Blanch, H. W., Prausnitz, J. M., Protein-Protein Interactions in Aqueous Ammonium Sulfate Solutions. Lysozyme and Bovine Serum Albumin (BSA). *Journal of Solution Chemistry* 2000, *29*, 699-718.

[35] Neal, B. L., Asthagiri, D., Lenhoff, A. M., Molecular Origins of Osmotic Second Virial Coefficients of Proteins. *Biophysical Journal* 1998, *75*, 2469-2477.

[36] Sophianopoulos, A. J., Rhodes, C. K., Holcomb, D. N., Van Holde, K. E., Physical Studies of Lysozyme: I. CHARACTERIZATION. *Journal of Biological Chemistry* 1962, *237*, 1107-1112.

[37] Verwey, E. J. W., Theory of the Stability of Lyophobic Colloids. *The Journal of Physical and Colloid Chemistry* 1947, *51*, 631-636.

[38] Neal, B. L., Asthagiri, D., Velev, O. D., Lenhoff, A. M., Kaler, E. W., Why is the osmotic second virial coefficient related to protein crystallization? *Journal of Crystal Growth* 1999, *196*, 377-387.

[39] Arosio, P., Jaquet, B., Wu, H., Morbidelli, M., On the role of salt type and concentration on the stability behavior of a monoclonal antibody solution. *Biophysical Chemistry* 2012, *168–169*, 19-27.

[40] López-León, T., Jódar-Reyes, A. B., Ortega-Vinuesa, J. L., Bastos-González, D., Hofmeister effects on the colloidal stability of an IgG-coated polystyrene latex. *Journal of Colloid and Interface Science* 2005, *284*, 139-148.

[41] Connolly, B. D., Petry, C., Yadav, S., Demeule, B., Ciaccio, N., Moore, J. M., Shire, S. J., Gokarn, Y. R., Weak interactions govern the viscosity of concentrated antibody solutions: high-throughput analysis using the diffusion interaction parameter. *Biophys J.* 2012, *103*, 69-78.

[42] Saito, S., Hasegawa, J., Kobayashi, N., Tomitsuka, T., *et al.*, Effects of Ionic Strength and Sugars on the Aggregation Propensity of Monoclonal Antibodies: Influence of Colloidal and Conformational Stabilities. *Pharmaceutical Research* 2013, *30*, 1263-1280.

[43] Lehermayr, C., Mahler, H.-C., Mäder, K., Fischer, S., Assessment of net charge and protein–protein interactions of different monoclonal antibodies. *Journal of pharmaceutical sciences* 2011, *100*, 2551-2562.

[44] Kuehner, D. E., Engmann, J., Fergg, F., Wernick, M., *et al.*, Lysozyme Net Charge and Ion Binding in Concentrated Aqueous Electrolyte Solutions. *The Journal of Physical Chemistry B* 1999, *103*, 1368-1374.

[45] Brummitt, R. K., Nesta, D. P., Chang, L., Kroetsch, A. M., Roberts, C. J., Nonnative aggregation of an IgG1 antibody in acidic conditions, part 2: Nucleation and growth kinetics with competing growth mechanisms. *Journal of Pharmaceutical Sciences* 2011, *100*, 2104-2119.

[46] Winzor, D. J., Jones, S., Harding, S. E., Determination of protein charge by capillary zone electrophoresis. *Analytical Biochemistry* 2004, *333*, 225-229.

[47] Kulmyrzaev, A. A., Schubert, H., Influence of KCl on the physicochemical properties of whey protein stabilized emulsions. *Food Hydrocolloids* 2004, *18*, 13-19.

[48] Johnson, S. B., Scales, P. J., Healy, T. W., The Binding of Monovalent Electrolyte Ions on α-Alumina. I. Electroacoustic Studies at High Electrolyte Concentrations. *Langmuir* 1999, *15*, 2836-2843.

[49] Tobler, S. A., Sherman, N. E., Fernandez, E. J., Tracking lysozyme unfolding during salt-induced precipitation with hydrogen exchange and mass spectrometry. *Biotechnology and Bioengineering* 2000, *71*, 194-207.

CHAPTER 5

This chapter has been submitted and published in the Biotechnology Journal.

5 ARGININE- DIPEPTIDES AFFECT INSULIN AGGREGATION IN A pH- AND IONIC STRENGTH-DEPENDENT MANNER

Nuhu, M. M. and Curtis, R. (2015). Biotechnology Journal, volume 10 (3). pages 404-416.

[School of Chemical Engineering and Analytical Sciences, The University of Manchester, Manchester, United Kingdom]

Correspondence: [Dr Robin Curtis, School of Chemical Engineering and Analytical Sciences, The University of Manchester, Sackville Street, M13 9PL, Manchester, United Kingdom].

E-mail: [r.curtis@manchester.ac.uk]

Keywords: [Arginine; dipeptides; excipients; insulin; protein aggregation]

Abbreviations: [**ADH,** alcohol dehydrogenase; **Arg,** arginine; **ArgEE,** L-arginine ethyl ester; **ArgME,** L-arginine methyl ester; **DLS,** dynamic light scattering; **Glu,** glutamate; **Gly,** glycine; **Leu,** leucine; **MCA,** micro-cuvette array; **Phe,** phenylalanine; **SLS,** static light scattering; **Val,** valine]

5.1 Abstract

Solutions containing arginine or mixtures of arginine and other amino acids are commonly used for protein liquid formulations to overcome problems such as high viscosities, aggregation, and phase separation. The aim of this work is to examine whether the stabilizing properties of arginine can be improved by incorporating the amino acid into a dipeptide. A series of arginine-containing dipeptides have been tested for their ability to suppress insulin aggregation over a range of pH and ionic strength. The aggregation is monitored at room temperature using a combination of turbidimetry and light scattering for solutions at pH 5.5 or 3.7, whereas thermal-induced aggregation is measured at pH 7.5. In addition, intrinsic fluorescence has been used to quantify additive binding to insulin. The dipeptide diArg is the most effective additive in solutions at pH 5.5 and pH 3.7, whereas the dipeptide Arg-Phe almost completely eliminates thermally-induced aggregation of insulin at pH 7.5 up to temperature of 90 °C. Insulin has been chosen as a model system because the molecular forces controlling its aggregation are well known. From this understanding, we are able to provide a molecular basis for how the various dipeptides affect insulin aggregation.

5.2 Introduction

There have been rapid advances in the upstream processing of protein therapeutics, which have shifted the economic bottleneck to downstream purification and formulation. Finding liquid formulations with shelf lives of up to two years is increasingly difficult for some of the newer therapeutics, such as antibodies and antibody fragments[1-3]. These proteins have been engineered for activity, but have a high propensity for irreversible aggregation, which can become exaggerated at higher protein concentrations (150 g/L) needed to meet required doses of therapeutics. Overcoming this problem requires using additives or excipients, which can function at low concentrations and are non-toxic [4, 5].

The effects of excipients on protein aggregation have been well studied for applications in protein refolding. Commonly used additives or excipients to prevent protein aggregation during protein renaturation steps can be classified into two groups, refolding agents and solubility enhancers [4, 6-8]. Refolding agents, which include sugars, polyols, and salts, function by a preferential exclusion mechanism, which is reflected by their ability to raise the protein melting temperature. On the other hand, solubility enhancers do not change the protein melting temperature, but instead, selectively bind to proteins leading to an increase in solubility of partially folded (aggregation-prone) intermediates [7, 9, 10]. Examples of solubility enhancers include various polyelectrolytes, some amino acids, and surfactants [4, 11]. While stabilization agents function at high concentrations, the effectiveness of solubility enhancers is, in part, controlled by their affinity to the protein, which opens up opportunities for engineering or improving excipient function through chemical modifications.

The problem of preventing aggregation in formulation is slightly different than in refolding. Aggregation occurs on longer time scales, because the concentration of aggregate precursors, or partially-folded proteins, is small relative to the native state [12-14]. Nevertheless, a similar classification can be used; excipients either increase the conformational stability (structure stabilizers) or the colloidal stability (solubility enhancers) [15-19].

One of the most universally used excipients is arginine [10, 20], which was originally developed to improve yields during refolding [6, 21, 22], but has since been used as an aid in chromatography [7, 23-25], for reducing viscosity of concentrated protein solutions [26], and for suppressing aggregation in liquid formulations. In general, arginine has little effect on protein melting temperature [6, 10], and as such, is believed to function as a solubility enhancer. This is reflected by its ability to suppress aggregation upon heating and cooling cycles [6, 27-29], or by increasing the solubility of either partially folded protein states formed under mildly denaturing or reducing conditions [30, 31], or chemically modified proteins that cannot refold properly[32-35]. However, the effects of arginine are not universal to all proteins. There are many cases, where arginine solutions have led to increased aggregation propensities of protein solutions [11, 36-40]. In addition, the beneficial effects usually occur over an additive concentration ranging from 100 mM up to 1 M. The use of arginine will be limited in many cases because liquid formulations must be at a physiological osmolarity. As such, there is a need to find alternative additives, especially those that are effective at micromolar concentrations.

A good understanding of the stabilizing mechanism of arginine against aggregation is needed for further developing improved additives. Arginine forms preferential interactions with either acidic or aromatic protein sidechains [9, 20, 27, 39, 41-43].

The interaction with acidic groups is stronger than an ordinary ion pair formation due to the hydrogen bond capabilities between the guanidinium group of arginine and the carboxylates of either glutamate or aspartate [39]. More attention has focused on understanding the interaction with aromatic groups, which are believed to control its effectiveness as an excipient. The preferential interactions formed with aromatics has been inferred from the ability to increase solubility of aromatic amino acids [7], the salting-in ability towards small aromatic compounds [41-43] from the proximity of arginine to aromatic groups in protein crystal structures [9], and the propensity of arginine-aromatic interactions to occur in protein-protein interfaces [44, 45]. What is less clear is how arginine binding alters protein-protein interactions thereby changing solubility. Arginine binding to proteins will alter the protein-protein electrostatic interactions, an effect which can lead to increased aggregation [38, 39]. Because this effect predominantly occurs with acidic proteins, the arginine binding sites have been inferred to be negatively charged groups. Arginine binding to aromatic groups will reduce sticky hydrophobic interactions leading to increased solubility of partially folded states [7, 46]. However, the stabilization mechanism has also been linked to the ability of arginine to self-associate through a head to tail interaction between the guanidinium and carboxylate end group [46-48]. It has been proposed that the self-association leads to cluster formation in bulk, which crowds out the protein-protein interactions [46]. However, other studies have suggested that arginine clusters around hydrophobic groups forming a polar mask pointing out towards the solvent [39, 47]. The latter occurs even when at low pH where the cluster formation does not occur in solution [49].

An improved molecular understanding can be gained from considering how chemical alterations to arginine alter the effects on protein behaviour. Generating a library of

derivatives has the added benefits that more effective excipients can be discovered. In a series of studies, the ability against preventing heat induced aggregation and oxidative refolding was determined for a set of arginine derivatives, and other molecules with similar functional groups [34, 35, 50-54]. Additives that aid in oxidative refolding contain a guanidinium or ureido group, whereas functional groups important for preventing heat-induced aggregation included amino groups, an amino acid backbone, where amino acids with either esterified or amidated carboxyl groups performed better. The thermoinactivation has since been linked to the ability of the additive to prevent chemical modifications, which has led to a third class of excipients being defined as inactivation suppressors under stress conditions [55]. For oxidative refolding, arginine derivatives such as L-Arginine ethyl ester (ArgEE), homo-arginine, and L-argininamide are more effective refolding aids than arginine, whereas L-Arginine methyl ester (ArgME), performed worse [50, 51]. The effectiveness was determined by two factors, the binding affinity for the protein surface (enhanced binding occurs for derivatives with extra methyl sidegroups) and the ability to form large clusters in solution. ArgEE and ArgME lack the carboxy terminal group and formed smaller clusters due to stacking of the guanidinium groups, whereas arginine and homo-arginine form clusters via a head to tail association.

In this work, we extend the studies of arginine derivatives by examining dipeptides containing at least one arginine residue. The advantage is that dipeptides are non-toxic and synthesis strategies are already well established. There have only been a few studies examining the effects of dipeptides on protein aggregation in the context of fibril formation, rather than preventing aggregation of biopharmaceuticals [56-58]. The dipeptide Gly-Phe lowers the aggregation propensity of cytochrome C

through a specific binding interaction confirmed by molecular modelling and consistent with the ineffectiveness of the dipeptide Phe-Gly [56]. However, more recently, dipeptides have been found to exhibit non-specific effects on aggregation; the dipeptide Arg-Phe accelerates the aggregation of acidic proteins at concentrations as low as 0.5 mM but has no effect on the aggregation of basic proteins. However, the same dipeptide inhibits the aggregation of alcohol dehydrogenase (ADH) when under a mild heat shock (at temperatures between 38 to 42 °C). Our study is also partly motivated by the finding that mixtures of arginine and glutamate at equimolar proportions exhibit synergistic effects in increasing the solubility of hydrophobic proteins [59] and inhibiting thermally-induced protein aggregation [60]. We have examined a series of dipeptides containing arginine linked with hydrophobic (Leu, Val), aromatic (Phe), acidic (Glu) and basic (Arg) groups. For selected dipeptides, we have also examined the effect of reversing the sequence and of amidating the carbonyl end-group.

This paper uses insulin, a therapeutic protein, as a model for probing the effects of additives on aggregation. Insulin is a globular protein with molecular weight of 5800 Da made up of two polypeptide chains, the A-chain and B-chain (21 and 30 amino acid residues respectively) held together by 2 disulphide bonds. In the most common therapeutic formulations, insulin forms a hexamer, which is assembled from three dimers in the presence of two zinc ions [61]. Patients also require a zinc-free form of insulin which is faster acting *in vivo*, but is much more aggregation prone [62].

There has been considerable debate over what association model best describes zinc-free insulin solutions. Early studies suggested a similar model to zinc-containing solutions [63-68], where the hexamer and/or tetramer are in equilibrium with the

dimer. However, oligomers greater than hexamer have been observed experimentally, which has led to various isodesmic association models, where the building block is either a dimer [69, 70], hexamer, or a monomer [71, 72], and either in a head-to-tail [70-72] or head-to-head and tail-to-tail manner [73, 74]. The conformation of the dimer in zinc-loaded insulin is similar to that in the zinc-free oligomers [63, 64, 75]. The dimer contains a hydrophobic core formed by burying a number of aromatic and aliphatic groups on the C-terminal end of the B chain, which undergo significant conformational rearrangement. In alkaline solutions at pH values above 7, insulin self-association is in part controlled by the electrostatic repulsion between monomer units. The electrostatic interaction is weakened either by reducing the net charge by lowering the pH or by ionic screening through increases in salt concentration [72, 76]. Insulin self-association follows a similar pattern under acidic conditions with pH below 3 [71, 73], in which case, the electrostatic repulsion between insulin monomers depends only on the net charge, rather than the sign of the charge [69, 72, 77]. In the pH range of 3 to 7, insulin net charge is reduced leading to growth in oligomers and visible solution turbidity at low ionic strength [69, 77-79]. Under these conditions, increasing salt concentration reduces the turbidity reflecting a decrease in insulin aggregation. This behaviour has been attributed to screening of attractive electrostatic interactions, which arise from oppositely charged surfaces on insulin [77]. Insulin association occurs rapidly and remains constant on a time scale of days in solutions with pH ranging from 3 to 10, except for conditions near to the isoelectric point (pH 5.5) at ionic strength values below 30 mM [77]. The latter makes insulin an ideal model for studying effects of additives on protein aggregation by light scattering approaches. Recognizing that additives are known to affect both electrostatic and hydrophobic interactions, we

have chosen to study insulin over a range of pH such that the charge is varied from negative to neutral to positive, with and without sodium chloride to screen the electrostatic interactions.

5.3 Materials and methods

Bovine Zn-insulin was purchased from Sigma-Aldrich (Dorset, UK). L-Glutamate, sodium phosphate (monobasic and dibasic), sodium acetate and benzoylated dialysis tubings were purchased from Sigma-Aldrich. Sodium chloride and acetic acid were obtained from Fischer Scientific. L-Arginine was purchased from VWR International (Poole, UK) while ultra-pure Tris base was gotten from Formedium Ltd (Norfolk, UK). Desalted dipeptides were purchased from Biomatik (Canada) and used without further purification. All chemicals and reagents were of analytical grade and used without further purification. Buffer solutions were filtered using a Millex-GV filter (sterile, 0.22 µm pore size/13 mm diameter) from Millipore. Buffer pH was adjusted with few drops of NaOH or HCl. Deionised distilled water was purified through a Millipore water-purification system with pore-size 0.22 µm before use.

Zn-free insulin was prepared by dissolving insulin powder in 2.5 mL of a 20 mM sodium acetate buffer at pH 3 and stirred gently for 30 minutes at 4 °C. The solution was transferred to dialysis membrane tubing with 2 kDa MWCO and dialysed four times against 1 L of the 20 mM acetate buffer at 4 °C. Each dialysis step lasted at least 3 hours while the fourth step was left to run overnight. The insulin solution was removed from the dialysis tubing and pH was raised from 3 to 7.1 by adding sodium hydroxide. The insulin solution was then transferred into newly prepared dialysis tubing and exhaustively dialysed against 4 L of 5 mM Tris buffer at pH 7.1 using the same dialysis procedure as described above. The solution was filtered through a 0.22

μm Millex filter and concentration determined by UV spectroscopy using an extinction coefficient of 1.0 mL/ (g-cm) at 280 nm [80]. The solution was stored at 4 °C and used within a few days.

Stock solutions of the additives (dipeptides or amino acids) were prepared by dissolving the additive in either a 20 mM sodium acetate buffer at pH 3.7, a 20 mM sodium phosphate buffer at pH 5.5, or a 20 mM sodium phosphate at pH 7.5. The pH of the solution was checked with a pH meter and readjusted using either hydrochloric acid or sodium hydroxide. Stock solutions were then filtered with a 0.22 μm Millex filter and stored at -20 °C.

5.4 Methods

5.4.1 Turbidimetry

Turbidity measurements were performed using a 1 cm path length probe connected to a Brinkmann PC 450 colorimeter. Before taking measurements, Milli-Q water was used to set the transmittance (T) values to 100%. Transmittance is related to turbidity τ by

$$\tau = 100 - \%T \qquad \text{(Equation 5.1)}$$

Filtered Zn-free insulin solution was diluted in 20 mM sodium phosphate buffer at pH 5.5 to a final working concentration of 0.1 g/L in a volume of 20 mL. Turbidimetric titrations were carried out at 25 °C by successive addition of 100 μL of a pH 5.5 solution containing 1 M additive concentration dissolved in 20 mM sodium phosphate buffer. The solution was stirred gently throughout the titration. All titrations were carried out two times to check reproducibility.

5.4.2 Dynamic and static light scattering at room temperature

Dynamic and static light scattering (DLS and SLS) measurements were conducted at a fixed scattering angle using a Malvern Zetasizer Nano-S system equipped with a built-in Peltier temperature control and an avalanche photodiode detector (Malvern Instruments Ltd., UK). Samples were measured in a 45 µL low volume quartz cuvette, DTS2145 (Hellma Analytics, Germany). The Zetasizer Nano S uses a 633 nm Helium-Neon (He-Ne) red laser as a light source and analyses scattered light at an angle of 173°. All measurements were carried out at a temperature of 25.0 ± 0.1 °C in automatic duration mode. For each sample, 11 measurements each of 10 seconds duration were taken and averaged. The measured autocorrelation function was fit to either a regularization analysis or cumulant expansion using the DTS application software (Malvern Instruments Ltd., UK). The results are reported in terms of an averaged hydrodynamic size, d_H which is related to the z-averaged translational diffusion coefficient, D through the Stokes-Einstein relation. The effect of intermolecular interactions on measured value of d_H is negligible for the protein concentrations investigated here. For static light scattering experiments, the scattering intensity of toluene (with a known Rayleigh ratio) was measured to calculate the instrument dependent calibration constant. The intensities of the samples were derived by using the mean count rates after adjusting by the attenuation factors. The Rayleigh ratio was determined from the derived count rate using the DTS software.

In each experiment, the Zn-free insulin solution at the desired pH was diluted volumetrically by a factor of twenty with a stock solution of the additive and immediately placed in the cuvette. The final insulin concentration was 0.2 g/L for samples at pH 5.5 and 2 g/L for samples at pH 3.7. Between measurements the

cuvette was thoroughly cleaned with distilled water (nitric acid, ethanol and surfactant solution) and wiped with lint-free wipes. All samples were run in triplicates, where error bars represent the standard deviation over the three measurements.

5.4.3 Temperature-controlled experiments

Temperature-controlled intrinsic fluorescence and static light scattering experiments were carried out using the Optim (Avacta Analytical Ltd, UK). The Optim contains two lasers that emit light at wavelengths of 266 nm and 473 nm. The light scattering data was recorded at the wavelength of 473 nm. The 266 nm laser source was used for intrinsic fluorescence excitation. Fluorescence spectra were measured over the wavelength interval of 290 to 320 nm using a slit width of 100 µm. The sample temperature was increased from 20 °C to 90 °C at a rate of 2 °C per minute. Light scattering and fluorescence readings were recorded at each temperature point.

Filtered insulin samples were diluted with stock concentrations of additives to a final working concentration of 2 g/L at desired pH conditions. Aliquots of 9 µL of each sample were carefully loaded into the compatible micro-cuvette arrays (MCAs). This was done in such a way to prevent the generation of air bubbles, which could potentially affect the light scattering results. Samples were then sealed at both ends of the cuvette holder in which the wells were placed to avoid leakage at both sides and placed within the instrument. Samples were measured in duplicates and control samples consisting of buffer-only solutions were included.

5.5 Results

5.5.1 Effect of dialysis procedure on insulin preparation

In all experiments we followed the same procedure for preparing a stock solution of insulin at a pH of 7.1 in a 5 mM Tris buffer. Dynamic light scattering measurements on a freshly prepared 3 g/L insulin stock solution indicated two peaks corresponding to a low molecular weight insulin oligomer and a small fraction of a high molecular weight peak reflecting the presence of irreversibly formed aggregates [72]. The small peak corresponding to 5.6 nm reflects an averaged oligomer size formed by the reversible isodesmic self-association. The value is consistent with Kadima et al. [72], which reported a range of 5 to 6 nm for the hydrodynamic diameter when increasing the ionic strength of the solution from 10 to 100 mM at pH 7.5. The insulin size at pH 7 is expected to be slightly greater than the value at pH 7.5 due to a lower net charge of insulin when reducing pH.

5.5.2 Effects of arginine and arginine mixtures on isoelectric precipitation of insulin

We initially examined the effects of using arginine or mixtures of arginine and glutamate on reducing the isoelectric precipitation of insulin. In Figure 5.1 is shown a representative plot of the turbidity versus additive concentration obtained by titrating a solution containing an initial concentration of 0.1 g/L insulin and 20 mM sodium phosphate buffer at pH 5.5. The initial transmittance of each solution remained constant for 5 minutes before the initial titrant was added. The decrease in turbidity with increasing additive concentration reflects a decrease in the average size of insulin oligomers. The solutions containing arginine are more effective at suppressing insulin association as the turbidity goes to 0 at an arginine concentration

of 60 mM, whereas the transmittance is 98.5 % for solutions containing sodium chloride at a concentration of 100 mM. There is no additional benefit of using mixtures of arginine and glutamate indicating glutamate behaves similar to chloride ion.

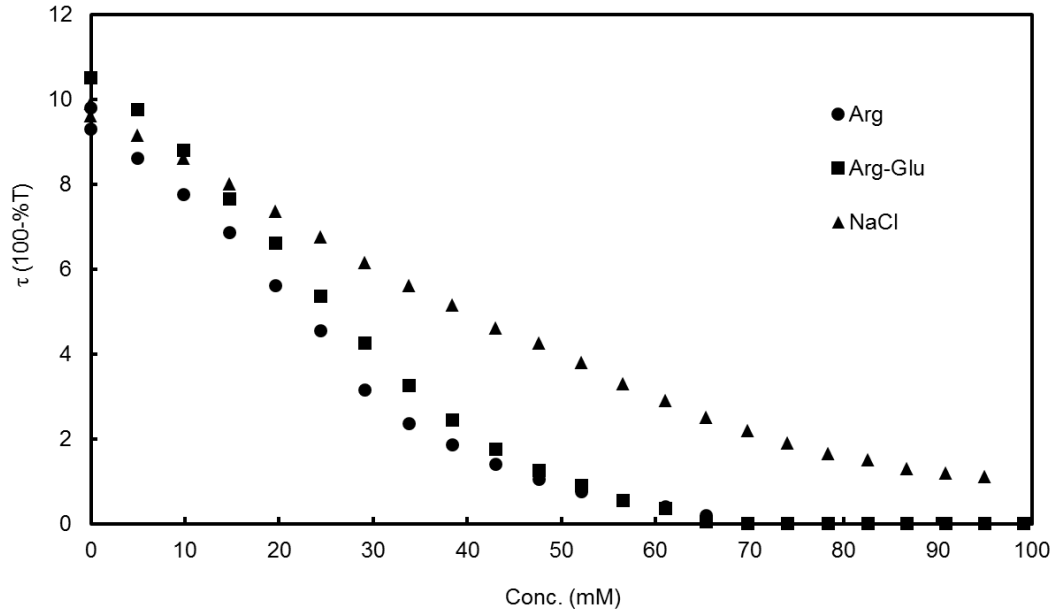


Figure 5.1. Turbidimetric titration of 0.1 g/L insulin against 1M stock concentration of additives in 20 mM phosphate buffer (pH 5.5). 100uL volume of the additives was added until transmittance values reached 100 % and are reported in terms of τ, as 100-%T. The data represent the mean values of two individual experiments. Arg-Glu corresponds to mixtures of Arg and Glu.

5.5.3 Time-evolution of insulin aggregation

Static and dynamic light scattering measurements were used to monitor the effects of various additives on insulin aggregation. In each experiment, the freshly prepared insulin stock solution was diluted volumetrically by a factor of 20 with the corresponding additive solution. Initially, we monitored the time evolution of the

light scattering signal in solutions containing 2 g/L insulin with a pH 3.7 buffer containing 20 mM sodium acetate and 100 mM sodium chloride or with 0.1 g/L insulin in a pH 5.5 buffer solution with 20 mM sodium phosphate. In each case, large oligomers with average sizes much greater than the hexamer are formed within a minute after dilution. However, after the rapid equilibration, there were no observable changes in the static light scattering signal for 60 minutes with either solution. In solutions at pH 5.5, the measured value of R_θ equals 2.00 ± 0.3 x 10^{-3} cm^{-1}. The error corresponds to the standard deviation over three measurements using different freshly prepared insulin stock solutions. In solutions at pH 3.7 and 100 mM NaCl, the measured value of R_θ equals 5.8 ± 0.4 x 10^{-3} cm^{-1}. The results are consistent with Giger et al [77] who found the turbidity readings remained constant after 60 seconds for a 0.1 g/L insulin solution at pH 5.5 for solutions with an ionic strength equal to or greater than 30 mM.

5.5.4 Effects of dipeptides on insulin aggregation state at room temperature

We have examined the effects of various additives on insulin aggregation in solutions at pH 5.5 with 20 mM sodium phosphate buffer (Figure 5.2A) or in solutions at pH 3.7 containing 100 mM sodium chloride and 20 mM sodium acetate buffer (Figure 5.2B). The light scattering readings are reported relative to the corresponding value obtained for insulin in the corresponding additive free solution. The error bars corresponds to the standard deviation over three measurements made using the same insulin stock solution.

The effects of additives are much more pronounced at pH 3.7 than at pH 5.5. As such, results are shown for solutions containing additives at concentrations of 20

mM for pH 3.7, versus 100 mM for pH 5.5. For the solutions at pH 3.7, there is a clear grouping of the additives; diArg, mixtures of arginine and glutamate, or Arg-Phe reduce insulin aggregation, whereas all the other additives shown exhibit little effect. For the measurements at pH 5.5, the light scattering signal is reduced most dramatically in solutions containing diArg relative to all other additives investigated. For the remaining uncapped dipeptides, the effectiveness at reducing aggregation follows the order Arg-Phe (Phe-Arg) > Leu-Arg > Val-Arg (Arg-Val), which correlates roughly with the hydrophobicity of the dipeptide. The amidated dipeptides cause greater light scattering when compared to all other dipeptides except for those containing valine. Interestingly, the light scattering from solutions containing arginine or mixtures of arginine and glutamate is only slightly less than that of sodium chloride, although the amino acid solutions are clearly more effective at reducing the turbidity for additive concentrations much lower than 100 mM (see Figure 5.1). This discrepancy is likely attributable to some irreversible aggregation of insulin that occurs in the solutions at pH 5.5 in the absence of sodium chloride.

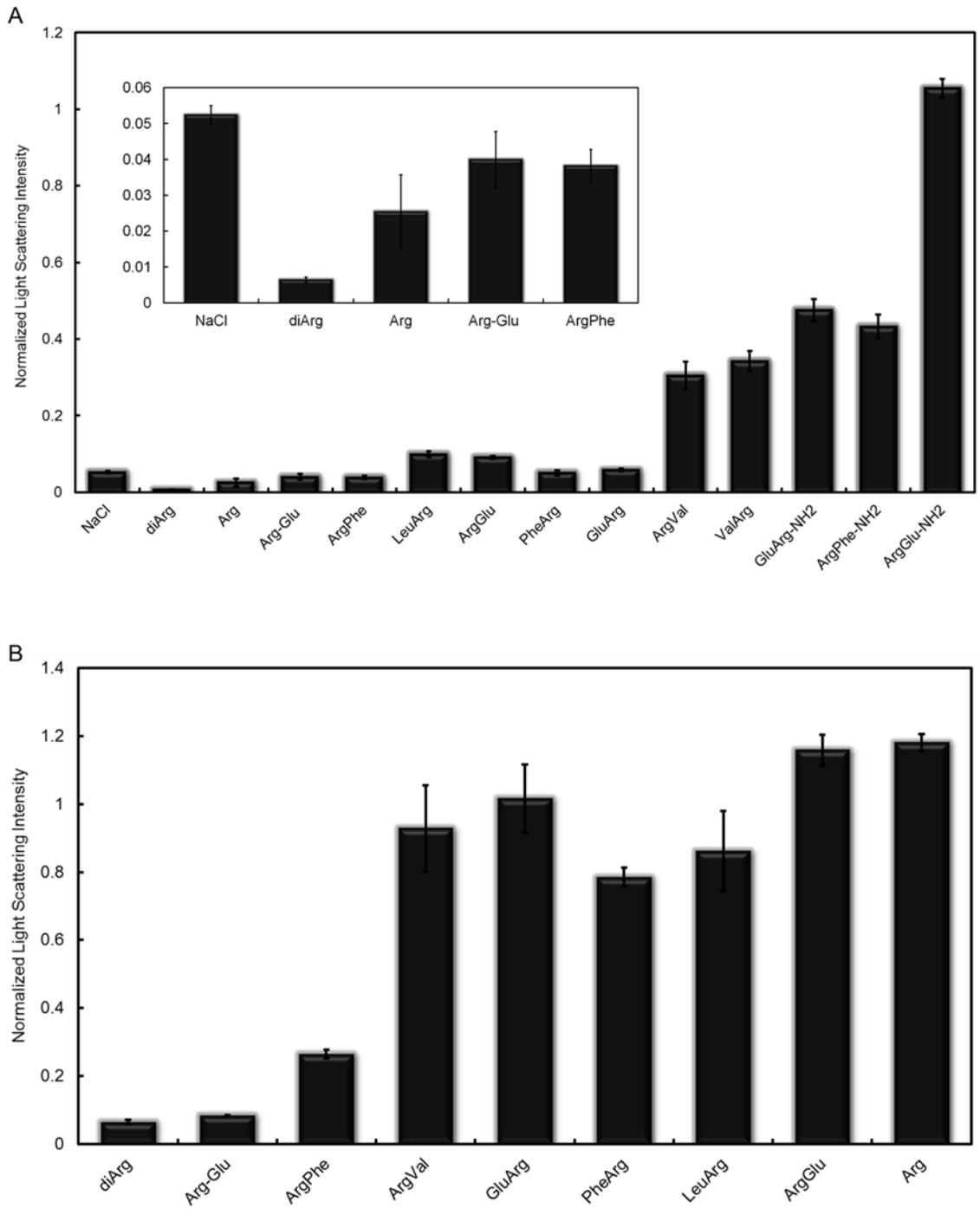


Figure 5.2. (A) Effect of 100 mM additives on scattering intensity of 0.2 g/L Zn-free insulin at 20mM phosphate buffer (pH 5.5). Scattering intensities are normalized with respect to insulin at 20 mM phosphate buffer (pH 5.5) in absence of NaCl. (B) Effect of 100 mM additives on scattering intensity of 2 g/L Zn-free insulin in 20 mM acetate buffer (pH 3.7). Scattering intensities have been normalized with respect

to the scattering intensity of insulin. The derived count rates are calculated from the mean count rates and attenuation factors corrected. Data represents an average of triplicate independent runs ± the standard deviation indicated by the error bars. Amino acids are represented by their three letter abbreviation. Arg-Glu corresponds to mixtures of Arg and Glu

5.5.5 Temperature accelerated aggregation studies of insulin

In solutions at pH 7.5, we used the Optim to measure the effect of temperature on the light scattering intensity for insulin in the presence of various additives. This approach is commonly used for estimating an aggregation temperature, T_{agg}, which corresponds to a rapid increase in the light scattering signal reflecting the onset of aggregation [15, 18]. Temperature induced aggregation occurs when hydrophobic groups are exposed in partially folded or unfolded states. The value of T_{agg} is used as a surrogate for the colloidal stability of partially folded proteins. If conformational stability (as characterized by protein melting temperature) remains constant, an increase in T_{agg} reflects an increase in colloidal stability or reduced self-association.

The light scattering intensities are plotted as a function of temperature for solutions containing 2 g/L insulin at pH 7.5 in the presence of 20 mM additives with and without 100 mM sodium chloride in Figures 5.3A and 5.3B, respectively. The light scattering intensity begins to increase around a temperature of 70 °C, which corresponds to the melting temperature of Zn-free insulin measured at pH 7.4 [81].

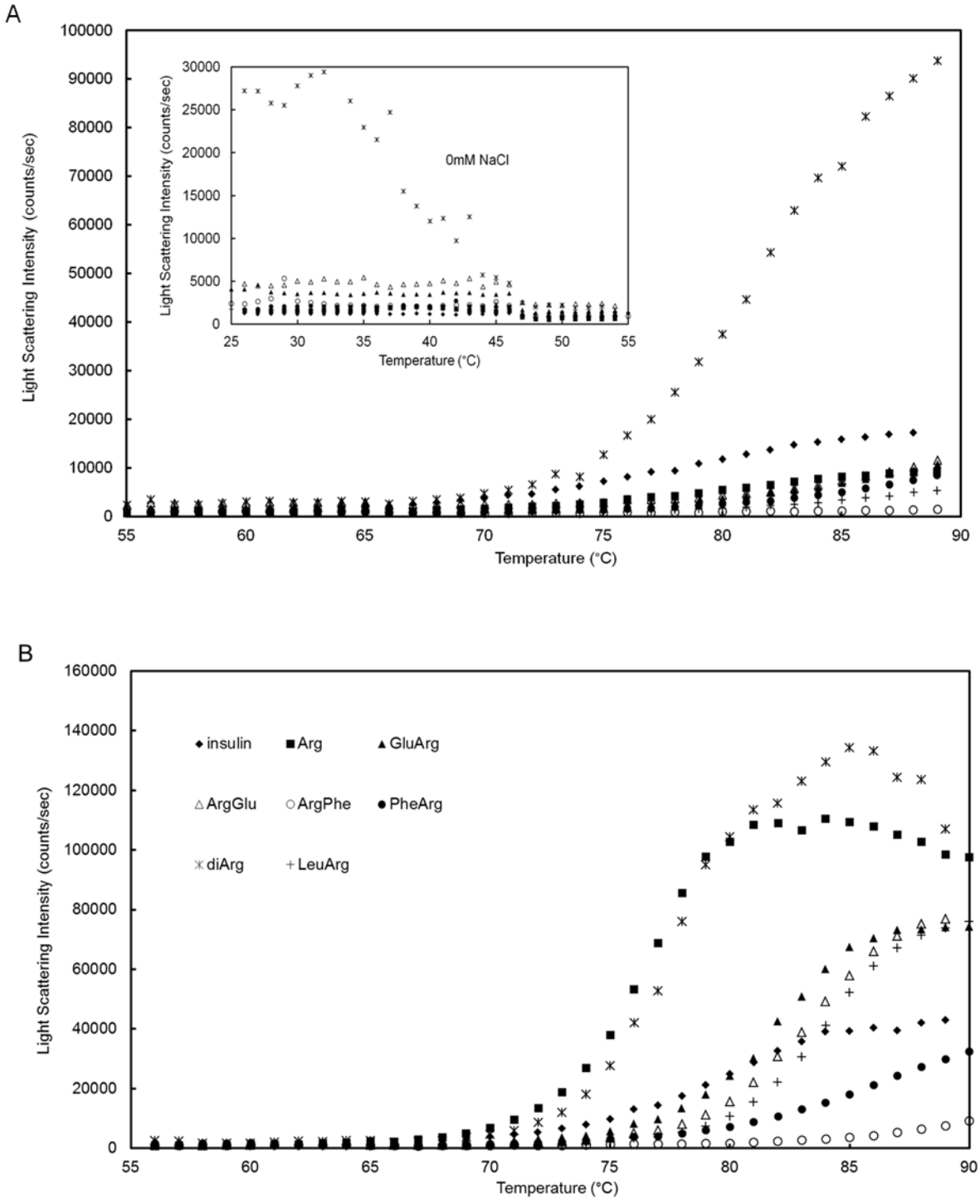


Figure 5.3. Scattering intensities at 473 nm of 2 g/L insulin in 20 mM phosphate buffer (pH 7.5) in the absence (A) and presence of 100 mM NaCl (B) respectively with 20 mM additives. Results shown represent an average of two consecutive measurements. Amino acids are represented by their three letter abbreviation.

Salt addition has a large effect on the measured light scattering intensity at high temperatures. The aggregates grow to a much larger size in solutions containing 100 mM sodium chloride indicating electrostatic repulsion between monomer units in the absence of sodium chloride is significant. The only exception occurs in solutions containing the dipeptide diArg, where the light scattering signals are similar in the presence and absence of salt. When electrostatic repulsion dominates, aggregates grow by monomer addition leading to the formation of linear aggregates, whereas aggregates grow by chain-chain associations leading to more condensed higher molecular weight structures in the absence of an electrostatic barrier [47].

For the salt-free case, all additives except diArg increase the onset temperature and reduce the high temperature light scattering relative to the additive-free solution. In solutions containing 100 mM sodium chloride, the effects of the dipeptides become more pronounced. The dipeptide Arg-Phe suppresses the aggregation significantly at higher temperatures. All other additives except Arg and diArg appear to increase the aggregation onset temperature relative to the additive-free solution.

At room temperature, there is an order of magnitude increase in the light scattering obtained for the solution containing diArg versus any other additive or additive-free solution. Because no visible haze formation occurred for these solutions, it is unlikely that insoluble aggregates are formed at the temperature onset of aggregation, which becomes detectable at light scattering readings similar to those observed for the diArg solutions at room temperature.

5.5.6 Fluorescence studies to probe dipeptide interactions with insulin

In Figure 5.4A to 5.4D are shown fluorescence spectra for 2 g/L insulin solutions at pH 7.5 using an excitation wavelength of 266 nm. Insulin does not contain any

tryptophan residues, so that the fluorescence arises from the tyrosine group with a negligible contribution from phenylalanine or disulfide bonds [82, 83]. The spectra have been shifted vertically such that the intensity at the peak maximum occurs in the same position to emphasize the peak shift.

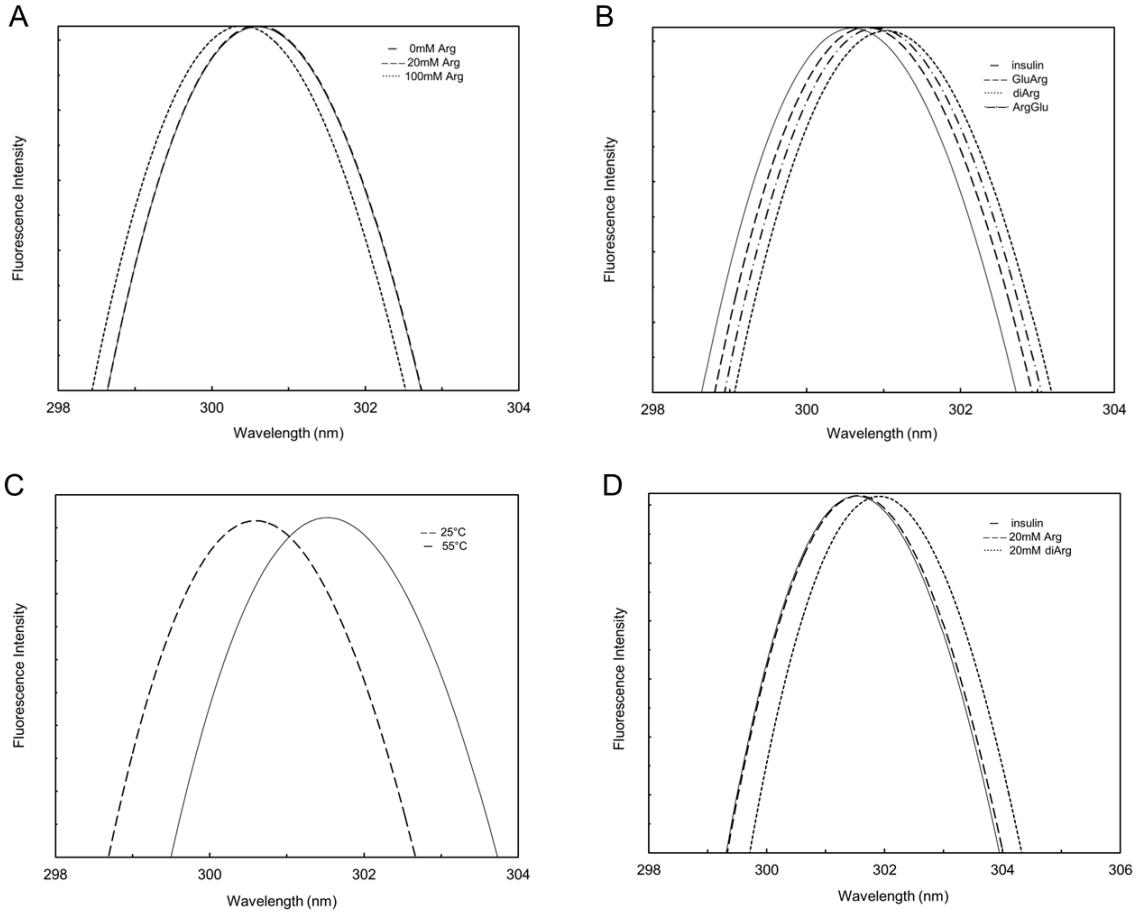


Figure 5.4.Shifted intrinsic fluorescence spectra for 20 mM phosphate solutions (pH 7.5) containing 2 g/L insulin with (A) Arg at concentrations of 0mM, 20 mM or 100 mM at 25 °C; (B) no additive, 20 mM diArg,20 mM GluArg, or 20 mM ArgGlu at 25 °C, (C) no additive at 25°C or 55°C. (D) no additive, 20 mM Arg or 20 mM diArg at 55 °C. Results shown represent an average of two consecutive measurements.

The unshifted fluorescence spectra are shown in Figures 5.5A to 5.5D. The change in the peak maximum is well-correlated with a concomitant change to the peak intensity. In each case, a positive increase in the peak maximum corresponds to a decrease in fluorescence intensity, except for in the presence of 100 mM arginine, where a decrease in the peak maximum was reflected by a decrease in the fluorescence intensity.

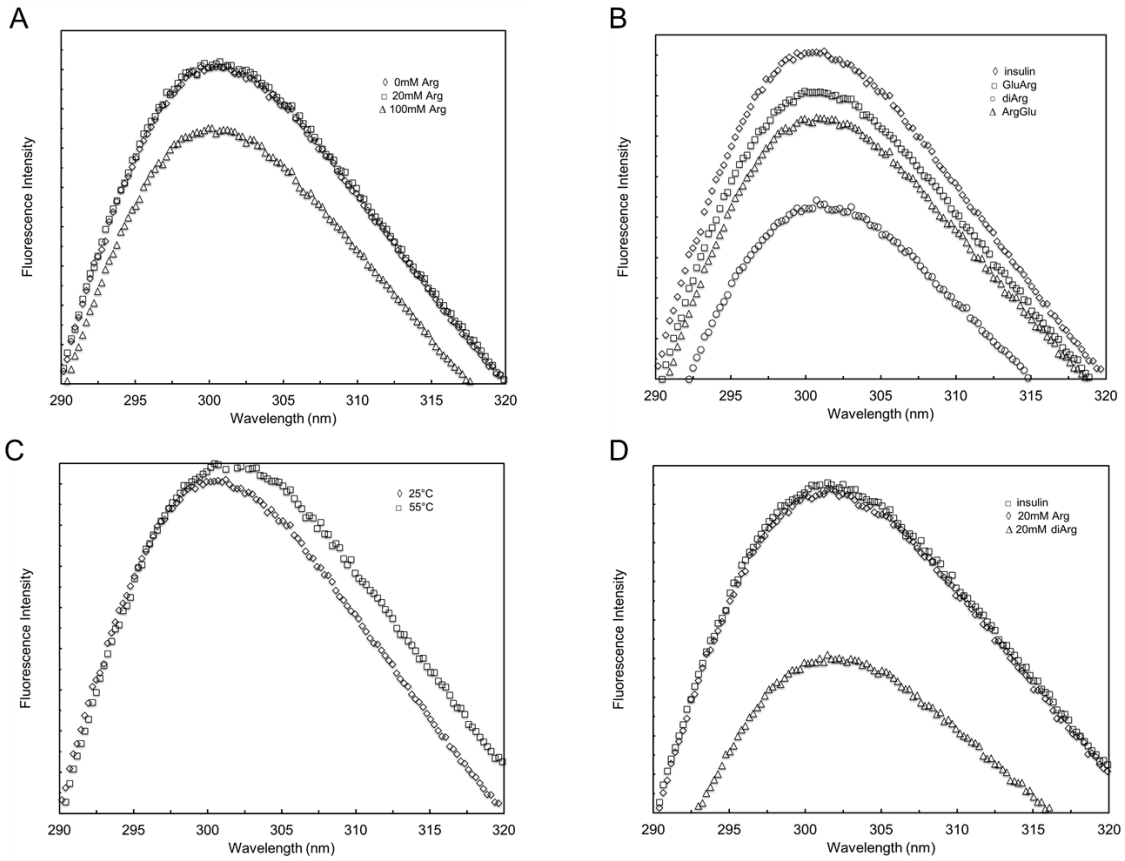


Figure 5.5. Unshifted intrinsic fluorescence spectra for 20 mM phosphate solutions (pH 7.5) containing 2 g/L insulin with (A) Arg at concentrations of 0mM, 20 mM or 100 mM at 25 °C; (B) no additive, 20 mM diArg,20 mM GluArg, or 20 mM ArgGlu at 25 °C, (C) no additive at 25°C or 55°C. (D) no additive, 20 mM Arg or 20 mM diArg at 55 °C. Results shown represent an average of two consecutive measurements.

Figure 5.4A demonstrates that changes to the emission spectrum at 25 °C only occur for insulin solutions at an arginine concentration of 100 mM, as curves corresponding to 20 mM arginine and arginine-free solution overlay with each other. This contrasts with solutions containing 20 mM dipeptides, where the largest shift is found with diArg followed by Arg-Glu and Glu-Arg. The peak intensities for solutions containing either diArg, Arg-Glu, or Glu-Arg are reduced relative to the additive free solution by 24 %, 12 %, or 9 % respectively.

The emission spectrum properties are sensitive to the chemical environment immediately surrounding the tyrosine groups. Changes to the spectrum could be due to changing environment polarity upon exposing tyrosine groups to the solvent or to alterations in the hydrogen bonding patterns from direct or indirect interactions with the dipeptides. In Figure 5.4D is shown the spectra at 55 °C obtained for insulin solutions in the presence and absence of 20 mM diArg. The light scattering intensities shown in the inset to Figure 5.3A reflect a large decrease in insulin aggregation when increasing temperature to 55 °C. Nevertheless, a similar change in the emission spectrum occurs at 55 °C or 25 °C in the presence of diArg suggesting that the intrinsic fluorescence signal is sensitive to additive binding by insulin rather than changes to insulin association state.

5.6 Discussion

5.6.1 Arginine neutralizes negatively charged groups of insulin

There are clear effects of arginine in suppressing the isoelectric precipitation of insulin in solutions at pH 5.5 (see Figure 5.1). The precipitation is driven by attractive electrostatic interactions formed between patches of opposite charge on the surface of insulin [77]. Increasing sodium chloride concentration screens the

electrostatic interactions and reduces insulin aggregation leading to the significant decrease in turbidity. The effectiveness at screening electrostatic interactions is determined by the screening length, which is proportional to the inverse square root of ionic strength. Because the dominant charged state of arginine at pH 5.5 is close to +1*e*, the ionic strength of an arginine chloride solution is equal to that of a sodium chloride solution when compared at the same salt concentration. Thus, the effect of arginine cannot only be rationalized in terms of ionic screening, otherwise arginine would have the same effectiveness as sodium chloride.

Previous studies have established that arginine binds to acidic and aromatic groups on protein surfaces [7, 27, 39, 41-43, 84], where the binding affinity for negatively charged groups is stronger than for aromatics [39]. As such, more likely, arginine binds to acidic insulin groups thereby neutralizing the negatively charged patches and preventing their electrostatic association.

In solutions at pH 7.5 with 100 mM sodium chloride, adding 20 mM arginine reduces the temperature at which insulin molecules aggregate providing evidence that arginine promotes aggregation of the heat denatured states. Drawing this conclusion also requires knowledge of how arginine alters the folding temperature or conformational stability. Protein melting temperatures are generally not sensitive to arginine especially when at such low additive concentration [6, 10]. Similarly, other studies have found that arginine increases aggregation rates of thermally denatured states for negatively charged proteins, such as bovine serum albumin, β-lactoglobulin, α-lactalbumin, and alcohol dehydrogenase [38, 39]. The increased rates are not observed for positively charged proteins, which have led to the suggestion that arginine binds to negatively charged groups and lowers the electrostatic repulsion barrier to protein-protein association. A similar mechanism

can be used to rationalize the data presented here for insulin. Increasing the net negative charge on insulin by increasing pH reduces self-association reflecting the effects of electrostatic repulsion in solutions at room temperature [72]. We expect the same repulsive forces exist at higher temperatures as electrostatic interactions between native proteins also occur between heat-denatured states when compared under the same solvent conditions [12, 85]. Thus, arginine binding to acidic protein groups at pH 7.5 is expected to neutralize repulsive electrostatic interactions leading to enhanced insulin aggregation. This contrasts with what happens at pH 5.5, where neutralization of acidic groups by arginine still occurs, but instead reduces the electrostatic driven aggregation.

5.6.2 Mixtures of arginine and glutamate suppress hydrophobic interactions

In solutions at pH 3.7 and 100 mM sodium chloride, mixtures of arginine and glutamate are much more effective at suppressing aggregation than arginine chloride at concentrations as low as 20 mM. The suppression effects are either due to neutralization of attractive electrostatic interactions, or due to weakening the isodesmic association. There are no synergistic effects of using Arg-Glu mixtures in solutions at pH 5.5, where the electrostatic aggregation is most pronounced. As such, the synergy between the amino acids is not due to impacting upon the electrostatic properties of insulin, rather a more likely explanation is that the mixture is more effective at reducing hydrophobic interactions as previously proposed for other proteins [46, 59]. Hydrophobic forces are especially significant in stabilizing the insulin dimer, which is a building block of insulin oligomers [63].

5.6.3 Dipeptides modulate hydrophobic and electrostatic interactions between insulin molecules

diArg is the most effective dipeptide at suppressing aggregation in solution at low and moderate pH, but enhances insulin association at pH 7.5. In solutions at pH 5.5, additives alter the electrostatic-driven aggregation by neutralizing the charged groups on insulin. Changes to insulin electrostatic properties will be much greater in solutions containing diArg than those with arginine. Because diArg has a greater positive charge than arginine, the electrostatic contribution to the binding affinity with negatively charged groups is larger. In addition, binding to negatively charged groups by diArg will lead to charge inversion, whereas arginine binding only neutralizes the negative charge. It follows that the electrostatic attraction between insulin molecules is suppressed because insulin acidic groups are masked much more effectively in solutions containing diArg versus arginine.

The ability of diArg to alter electrostatic interactions is also reflected by the behaviour at pH 7.5. In solutions without any sodium chloride, in the presence of diArg, aggregation occurs at a lower temperature and aggregates grow to a much larger size than in the additive free solution. Both these trends indicate diArg has neutralized the negatively charged groups on insulin. With increasing the net charge on the protein, the cross-over between these pathways occurs at a higher ionic strength as greater screening is required to reduce the repulsive forces. For most of the additive solutions, increasing sodium chloride concentration from 0 to 100 mM leads to much larger light scattering at high temperatures, reflecting the change in aggregation pathways brought about by screening the electrostatic repulsion. The main exception is the solution containing diArg. The light scattering at 0 mM

sodium chloride is similar to or greater than any of the other additive solutions that contain sodium chloride, indicating the electrostatic repulsion between insulin molecules is significantly reduced in the presence of diArg. The increased light scattering signal at lower temperatures in solutions without sodium chloride when comparing diArg to any of the other additives also provides evidence for increased insulin aggregation due to charge neutralization by additive binding [72].

The dipeptide Arg-Phe and to a lesser extent Phe-Arg prevents heat-induced aggregation at pH 7.5. It is likely that these dipeptides will also affect the aggregation of other proteins. A previous study found solutions of Arg-Phe suppressed the heat-induced aggregation of alcohol dehydrogenase (ADH), but interestingly, enhanced the aggregation at room temperature [58]. Arg-Phe also was found to alter the aggregation of α–lactalbumin in a pH dependent manner, in which case, aggregation was promoted when α–lactalbumin carried a significant negative charge [57]. This finding was taken to indicate a key role of neutralizing acidic groups leading to the enhanced aggregation rates. Because the heat-induced aggregation occurs from partially folded states implies that the suppression effectiveness of Phe-Arg arises from its ability to mask exposed hydrophobic groups. Thus, as with arginine or diArg, Phe-Arg can behave as either an aggregation suppressor or enhancer. The suppression effects are usually linked with additive binding to mask interactions between aromatic or hydrophobic groups [20, 46]. On the other hand, neutralization of protein charged groups by additive binding can either enhance protein aggregation due to reducing protein net charge and lowering the electrostatic repulsion barrier [38, 39], or as with the isoelectric precipitation of insulin, reduce electrostatic driven association by masking charged groups [77].

A key question to address is whether or not the effects of dipeptides on insulin association can be extrapolated to protein therapeutics such as antibodies or antibody-derived proteins. In the latter case, aggregation occurs through a more complicated pathway involving partial unfolding, association of partially folded states, and conformational changes leading to irreversibly formed aggregates [13]. Previous studies have already shown that arginine alone or in mixtures with glutamate is effective at suppressing aggregation of proteins [7,56]. The finding that diArg, and to a lesser extent Arg-Phe, exhibits similar or stronger effects than either of these additive mixtures at effecting electrostatic- or hydrophobic-driven association of insulin provides an initial indication that the dipeptide will also influence aggregation of therapeutic proteins. However, this will be highly dependent on what are the rate controlling steps in the aggregation pathways as well as on the properties of the aggregation prone regions in protein therapeutics.

Acknowledgement

The authors are grateful for useful scientific discussions with Alexander Golovanov and Andrew Doig at the University of Manchester and Jan Jezek and Barry Derham at Arecor. In addition, they thank Avacta for providing in-kind access to the Optim. The project was supported by pump-priming funds from the University of Manchester and a BBSRC-BRIC grant (number BB/I017194/1).

Conflict of interest

The authors declare no commercial or financial conflict of interest.

5.7 References

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

[2] Shire, S. J., Shahrokh, Z., Liu, J., Challenges in the development of high protein concentration formulations. *J. Pharm Sci.* 2004, *93*, 1390-1402.

[3] Shire, S. J., Formulation and manufacturability of biologics. *Curr Opin Biotechnol.* 2009, *20*, 708–714.

[4] Ohtake, S., Kita, Y., Arakawa, T., Interactions of formulation excipients with proteins in solution and in the dried state. *Adv Drug Deliv Rev.* 2011, *63*, 1053-1073.

[5] 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. *Adv Drug Deliv Rev.*2011, *63*, 1118-1159.

[6] Arakawa, T., Tsumoto, K., The effects of arginine on refolding of aggregated proteins: not facilitate refolding, but suppress aggregation. *Biochem Biophys Res Commun.* 2003, *304*, 148-152.

[7] Arakawa, T., Ejima, D., Tsumoto, K., Obeyama, N., *et al.*, Suppression of protein interactions by arginine: A proposed mechanism of the arginine effects. *Biophys Chem.* 2007, *127*, 1-8.

[8] Tsumoto, K., Ejima, D., Kumagaia, I., Arakawa T., Practical considerations in refolding proteins from inclusion bodies. *Protein Expr Purif.* 2003, *28*, 1-8.

[9] Ito, L., Shiraki, K., Matsuura, T., Okumura, M., Hasegawa, K., Baba, S., Yamaguchi, H., Kumasaka, T., High-resolution X-ray analysis reveals binding of

arginine to aromatic residues of lysozyme surface: implication of suppression of protein aggregation by arginine. *Protein Eng Des Sel.* 2011, *24*, 269-274.

[10] Shiraki, K., Kudou, M., Fujiwara, S., Imanaka, T., Takagi, M., Biophysical Effect of Amino Acids on the Prevention of Protein Aggregation. *J Biochem.* 2002, *132*, 591-595.

[11] Liu, Y., Li, J., Wang, F., Chen, J., Li, P., Su, Z., A newly proposed mechanism for arginine-assisted protein refolding - not inhibiting soluble oligomers although promoting a correct structure. *Protein Expr Purif.* 2007, *51*, 235-242.

[12] Chi, E. Y., Krishnan, S., Kendrick, B. S., Chang, B. S., *et al.*, Roles of conformational stability and colloidal stability in the aggregation of recombinant human granulocyte colony-stimulating factor. *Protein Sci.* 2003, *12*, 903-913.

[13] Roberts, C. J., Das, T. K., Sahin, E., Predicting solution aggregation rates for therapeutic proteins: Approaches and challenges. *Int J Pharm.* 2011, *418*, 318-333.

[14] Brummitt, R. K., Nesta, D. P., Roberts, C. J. , Predicting Accelerated Aggregation Rates for Monoclonal Antibody Formulations, and Challenges for Low-Temperature Predictions. *J. Pharm. Sci.* 2011, *100*, 4234-4243.

[15] He, F., Woods, C. E., Becker, G. W., Narhi, L. O., Razinkov, V. I., High-throughput assessment of thermal and colloidal stability parameters for monoclonal antibody formulations. *J. Pharm. Sci.* 2011, *100*, 5126-5141.

[16] Saito, S., Hasegawa, J., Kobayashi, N., Tomitsuka, T., *et al.*, Effects of Ionic Strength and Sugars on the Aggregation Propensity of Monoclonal Antibodies: Influence of Colloidal and Conformational Stabilities. *Pharm Res.* 2013, *30*, 1263-1280.

[17] 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. *Pharm Res.* 2003, *20*, 1325-1336.

[18] 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, *100*, 1306-1315.

[19] Roberts, C. J., Therapeutic protein aggregation: mechanisms, design, and control. *Trends Biotechnol.* 2014, *32*, 372-380.

[20] Arakawa, T., Tsumoto, K., Kita, Y., Chang, B., Ejima, D., Biotechnology applications of amino acids in protein purification and formulations. *Amino Acids* 2007, *33*, 587-605.

[21] Tsumoto, K., Umetsu, M., Kumagai, I., Ejima, D., *et al.*, Role of Arginine in Protein Refolding, Solubilization, and Purification. *Biotechnol Prog.* 2004, *20*, 1301–1308.

[22] Rudolph, R., Fischer, S., 1990, *US4933434 A*.

[23] Falconer, R. J., Chan, C., Hughes, K., Munro, T. P., Stabilization of a monoclonal antibody during purification and formulation by addition of basic amino acid excipients. *Chem. Technol. Biotechnol.* 2011, *86*, 942-948.

[24] Arakawa, T., Philo, J. S., Tsumoto, S., Yumioka, R., Ejima, D., Elution of antibodies from a Protein-A column by aqueous arginine solutions. *Protein Expr Purif.* 2004, *36*, 244-248.

[25] Ejima, D., Yumioka, R., Tsumoto, K., Arakawa, T., Effective elution of antibodies by arginine and arginine derivatives in affinity column chromatography. *Anal Biochem.* 2005, *345*, 250-257.

[26] Inoue, N., Takai, E., Arakawa, T., Shiraki, K., Specific Decrease in Solution Viscosity of Antibodies by Arginine for Therapeutic Formulations. *Mol. Pharm.* 2014, *11*, 1889-1896.

[27] Ghosh, R., Sharma, S., Chattopadhyay, K., Effect of Arginine on Protein Aggregation Studied by Fluorescence Correlation Spectroscopy and Other Biophysical Methods. *Biochemistry* 2009, *48*, 1135-1143.

[28] Tomita, S., Yoshikawa, H., Shiraki, K., Arginine controls heat-induced cluster–cluster aggregation of lysozyme at around the isoelectric point. *Biopolymers* 2011, *95*, 695-701.

[29] Arakawa, T., Kita, Y., Ejima, D., Tsumoto, K., Fukada, H., Aggregation suppression of proteins by arginine during thermal unfolding. *Protein Pept Lett.* 2006, *13*, 921-927.

[30] Ho, J. G., Middleberg, A. P., Estimating the potential refolding yield of recombinant proteins expressed as inclusion bodies. *Biotechnol. Bioeng.* 2004, *87*, 584–592.

[31] Lyutova, E. M., Kasakov, A. S., Gurvits, B. Y., Effects of Arginine on Kinetics of Protein Aggregation Studied by Dynamic Laser Light Scattering and Tubidimetry Techniques. *Biotechnol Prog.* 2007, *23*, 1411-1416.

[32] Reddy, K. R. C., Lilie, H., Rudolph, R., Lange, C., l-Arginine increases the solubility of unfolded species of hen egg white lysozyme. *Protein Sci.* 2005, *14*, 929-935.

[33] Tischer, A., Lilie, H., Rudolph, R., Lange, C., L-Arginine hydrochloride increases the solubility of folded and unfolded recombinant plasminogen activator rPA. *Protein Sci.* 2010, *19*, 1783-1795.

[34] Matsuoka, T., Hamada, H., Matsumoto, K., Shiraki, K., Indispensable structure of solution additives to prevent inactivation of lysozyme for heating and refolding. *Biotechnol Prog.* 2009, *25*, 1515-1524.

[35] Hamada, H., Takahashi, R., Noguchi, T., Shiraki, K., Differences in the Effects of Solution Additives on Heat- and Refolding-Induced Aggregation. *Biotechnol Prog.* 2008, *24*, 436-443.

[36] Cirkovas, A., Sereikaite, J., Different effects of L-arginine on the heat-induced unfolding and aggregation of proteins. *Biologicals* 2011, *39*, 181-188.

[37] Fujimoto, A., Hirano, A., Shiraki, K., Ternary System of Solution Additives with Arginine and Salt for Refolding of Beta-Galactosidase. *Protein J.* 2010, *29*, 161-166.

[38] Smirnova, E., Safenkova, I., Stein-Margolina, B., Shubin, V., Gurvits, B., l-Arginine induces protein aggregation and transformation of supramolecular structures of the aggregates. *Amino Acids* 2013, *45*, 845-855.

[39] Shah, D., Shaikh, A. R., Peng, X., Rajagopalan, R., Effects of arginine on heat-induced aggregation of concentrated protein solutions. *Biotechnol Prog.* 2011, *27*, 513-520.

[40] Fukuda, M., Kameoka, D., Torizawa, T., Saitoh, S., Yasutake, M., Imaeda, Y., Koga, A., Mizutani A., Thermodynamic and fluorescence analyses to determine mechanisms of IgG1 stabilization and destabilization by arginine. *Pharm Res.* 2014, *31*, 992-1001.

[41] Shah, D., Li, J., Shaikh, A. R., Rajagopalan, R., Arginine–aromatic interactions and their effects on arginine-induced solubilization of aromatic solutes and suppression of protein aggregation. *Biotechnol Prog.* 2012, n/a-n/a.

[42] Hirano, A., Arakawa, T., Shiraki, K., Arginine Increases the Solubility of Coumarin: Comparison with Salting-in and Salting-out Additives. *J Biochem.* 2008, *144*, 363-369.

[43] Ariki, R., Hirano, A., Arakawa, T., Shiraki, K., Arginine increases the solubility of alkyl gallates through interaction with the aromatic ring. *J Biochem.* 2011, *149*, 389-394.

[44] Crowley, P. B., Golovin, A., Cation–π interactions in protein–protein interfaces. *Proteins* 2005, *59*, 231–239.

[45] Martis, R. L., Singh, S. K., Gromiha, M. M., Santhosh, C., Role of cation–π interactions in single chain 'all-alpha' proteins. *J Theor Biol.* 2008, *250*, 655-662.

[46] Shukla, D., Trout, B. L., Understanding the Synergistic Effect of Arginine and Glutamic Acid Mixtures on Protein Solubility. *J Phys Chem B*, 2011, *115*, 11831-11839.

[47] Li, J., Garg, M., Shah, D., Rajagopalan, R., Solubilization of aromatic and hydrophobic moieties by arginine in aqueous solutions. *J Chem Phys.* 2010, *133*, 054902-054908.

[48] Vagenende, V., Han, A. X., Mueller, M., Trout, B. L., Protein-Associated Cation Clusters in Aqueous Arginine Solutions and Their Effects on Protein Stability and Size. *ACS Chem Biol.* 2013, *8*, 416-422.

[49] Das, U., Hariprasad, G., Ethayathulla, A. S., Manral, P., *et al.*, Inhibition of Protein Aggregation: Supramolecular Assemblies of Arginine Hold the Key. *PLoS ONE* 2007, *2*, e1176.

[50] Gao, M., Dong, X., Sun, Y., Interactions between l-arginine/l-arginine derivatives and lysozyme and implications to their inhibition effects on protein aggregation. *Biotechnol Prog.* 2013, n/a-n/a.

[51] Shiraki, K., Kudou, M., Nishikori, S., Kitagawa, H., *et al.*, Arginine ethylester prevents thermal inactivation and aggregation of lysozyme. *Eur J Biochem.* 2004, *271*, 3242-3247.

[52] Kudou, M., Shiraki, K., Fujiwara, S., Imanaka, T., Takagi, M., Prevention of thermal inactivation and aggregation of lysozyme by polyamine. *Eur J Biochem.* 2003, *270*, 4547-4554.

[53] Hamada, H., Shiraki, K., l-Argininamide improves the refolding more effectively than l-arginine. *J Biotechnol.* 2007, *130*, 153-160.

[54] Matsuoka, T., Tomita, S., Hamada, H., Shiraki, K., Amidated amino acids are prominent additives for preventing heat-induced aggregation of lysozyme. *J Biosci Bioeng.* 2007, *103*, 440-443.

[55] Tomita, S., Shiraki, K., Why do solution additives suppress the heat-induced inactivation of proteins? Inhibition of chemical modifications. *Biotechnol Prog.* 2011, *27*, 855-862.

[56] La Rosa, C., Milardi, D., Amato, E., Pappalardo, M., Grasso, D., Molecular mechanism of the inhibition of cytochrome c aggregation by Phe-Gly. *Arch Biochem Biophys.* 2005, *435*, 182-189.

[57] Artemova, N. V., Stein-Margolina, V. A., Bumagina, Z. M., Gurvits, B. Y., Acceleration of protein aggregation by amphiphilic peptides: Transformation of supramolecular structure of the aggregates. *Biotechnol Prog.* 2011, *27*, 846-854.

[58] Artemova, N., Stein-Margolina, V., Smirnova, E., Gurvits, B., Formation of supramolecular structures of a native-like protein in the presence of amphiphilic peptides: Variations in aggregate morphology. *FEBS Letters* 2012, *586*, 186-190.

[59] Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., Lian, L., A Simple Method for Improving Protein Solubility and Long-Term Stability. *J. Am. Chem. Soc.* 2004, *126*, 8933-8939.

[60] Kheddo, P., Tracka, M., Armer, J., Dearman, R.J., Uddin, S., van der Walle, C. F., Golovanov, A. P., The effect of arginine glutamate on the stability of monoclonal antibodies in solution. *Int J Pharm.* 2014, *473*, 126-133.

[61] Blundell, T., Dodson, G., Hodgkin, D., Mercola, D., Insulin. The Structure In The Crystal And Its Reflection In Chemistry And Biology. *Adv.Protein Chem. 26* 1972, 279-402.

[62] Brange, J., Ribel, U., Hansen, J. F., Dodson, G., *et al.*, Monomeric insulins obtained by protein engineering and their medical implications. *Nature* 1988, *333*, 679-682.

[63] Pocker, Y., Biswas, S. B., Self-association of insulin and the role of hydrophobic bonding: a thermodynamic model of insulin dimerization. *Biochemistry* 1981, *20*, 4354-4361.

[64] Goldman, J., Carpenter, F. H., Zinc binding, circular dichroism, and equilibrium sedimentation studies on insulin (bovine) and several of its derivatives. *Biochemistry* 1974, *13*, 4566-4574.

[65] Jeffrey, P. D. and Coates, J. H., An Equilibrium Ultracentrifuge Study of the Self-Association of Bovine Insulin*. *Biochemistry* 1966, *5*, 489-498.

[66] Hvidt, S., Insulin Association In Neutral Solutions Studied By Light-scattering. *Biophys Chem.* 1991, *39*, 205-213.

[67] Hansen, J. F., The self-association of zinc-free human insulin and insulin analogue B13-glutamine. *Biophys Chem.* 1991, *39*, 107-110.

[68] Coffman, F. D., Dunn, M. F., Insulin-metal ion interactions: the binding of divalent cations to insulin hexamers and tetramers and the assembly of insulin hexamers. *Biochemistry* 1988, *27*, 6179-6187.

[69] Xu, Y., Yan, Y., Seeman, D., Sun, L., Dubin, P. L., Multimerization and Aggregation of Native-State Insulin: Effect of Zinc. *Langmuir* 2012, *28*, 579-586.

[70] Jeffrey, P. D., Milthorpe, B. K., Nichol, L. W., Polymerization Pattern of Insulin at pH 7.0. *Biochemistry* 1976, *15*, 4660-4665.

[71] Attri, A. K., Fernandez, C., Minton, A. P., pH-dependent self-association of zinc-free insulin characterized by concentration-gradient static light scattering. *Biophys Chem.* 2010, *148*, 28-33.

[72] Kadima, W., Ogendal, L., Bauer, R., Kaarsholm, N., Brodersen, K., Hansen, J. F., Porting, P., The influence of ionic strength and pH on the aggregation properties of zinc-free insulin studied by static and dynamic laser light scattering. *Biopolymers* 1993, *33*, 1643-1657.

[73] Mark, A. E., Nichol, L. W., Jeffrey, P. D., The Self-association of Zinc-free Bovine Insulin - A Single Model Based On Interactions In the Crystal That Describes the Association Pattern In Solution At Ph 2, 7 and 10. *Biophys Chem.* 1987, *27*, 103-117.

[74] Nourse, A., Jeffrey, P. D., A sedimentation equilibrium study of platypus insulin: the HB10D mutant does not associate beyond dimer. *Biophys Chem.* 1998, *71*, 21-34.

[75] Ganim, Z., Jones, K. C., Tokmakoff, A., Insulin dimer dissociation and unfolding revealed by amide I two-dimensional infrared spectroscopy. *Phys. Chem. Chem. Phys.*, 2010, *12*, 3579-3588.

[76] Pedersen, J., Hansen, S., Bauer, R., The aggregation behavior of zinc-free insulin studied by small-angle neutron scattering. *Eur Biophys J.* 1994, *22*, 379-389.

[77] Giger, K., Vanam, R. P., Seyrek, E., Dubin, P. L., Suppression of Insulin Aggregation by Heparin. *Biomacromolecules* 2008, *9*, 2338-2344.

[78] Tanford, C., Epstein, J., The Physical Chemistry of Insulin .1. Hydrogen Ion Titration Curve of Zinc-free Insulin. *J Am. Chem. Soc.* 1954, *76*, 2163-2169.

[79] Kaarsholm, N. C., Havelund, S., Hougaard, P., Ionization behavior of native and mutant insulins: pK perturbation of B13-Glu in aggregated species. *Arch Biochem Biophys.* 1990, *283*, 496-502.

[80] Nielsen, L., Frokjaer, S., Brange, J., Uversky, V. N., Fink, A. L., Probing the Mechanism of Insulin Fibril Formation with Insulin Mutants†. *Biochemistry* 2001, *40*, 8397-8409.

[81] Huus, K., Havelund, S., Olsen, H. B., van de Weert, M., Frokjaer, S., Thermal Dissociation and Unfolding of Insulin. *Biochemistry* 2005, *44*, 11171-11177.

[82] Yong, Z., Yingjie, D., Ming, L., Craig, D. Q. M., Zhengqiang, L., A spectroscopic investigation into the interaction between bile salts and insulin in alkaline aqueous solution. *J Colloid Interface Sci.* 2009, *337*, 322-331.

[83] Bekard, I. B., Dunstan, D. E., Tyrosine Autofluorescence as a Measure of Bovine Insulin Fibrillation. *Biophys J.* 2009, *97*, 2521-2531.

[84] Ito, L., Shiraki, K., Matsuura, T., Okumura, M., *et al.*, High-resolution X-ray analysis reveals binding of arginine to aromatic residues of lysozyme surface: implication of suppression of protein aggregation by arginine. *Protein Eng Des Sel.* 2011, *24*, 269-274.

[85] Li, Y., Ogunnaike, B. A. Roberts, C. J., Multi-variate approach to global protein aggregation behavior and kinetics: effects of pH, NaCl, and temperature for alpha-chymotrypsinogen A. *J Pharm Sci.* 2010, *99*, 645-662.

CHAPTER 6

This work was performed in collaboration with Arecor Ltd and Genzyme.

6 AGGREGATION STABILITY OF AN IgG4 ANTIBODY DURING FORMULATION STUDIES

6.1 Abstract

The aggregation behaviour of an IgG4 antibody was characterised under a range of solution conditions and in the presence of a novel cationic excipient. Conformational and colloidal stability of the antibody were determined from temperature controlled intrinsic fluorescence and light scattering studies. Dynamic and static light scattering were used to measure protein-protein and protein-excipient interactions at room temperature. Lower pH increased net repulsive protein-protein interactions and decreased the thermal stability of the proteins. Fluorescence measurements showed at least two unfolding transitions (or domains) with the lower temperature transition domain as a key driver to the formation of aggregation prone states. Attractive interactions were screened as the ionic strength is increased indicating increased colloidal stability and the magnitude of screening increased with increasing pH. Protein-polycation interactions led to the formation of complexes and are initiated mainly through electrostatic interactions at pH 6 and pH 7. Protein interactions at pH 5 where the protein has more positive charges seem to suggest presence of other short range interactions control protein-polycation behaviour. Both thermal induced unfolding and colloidal interactions showed similar binding of the polycation to the protein. Although changes in intrinsic fluorescence during thermal unfolding suggests that the polycation reduces the conformational stability as aggregation proceeds through partial unfolding, protein-protein interactions are repulsive. Here, partial unfolding is not the rate limiting step to aggregation.

6.2 Introduction

Many important biological processes are modulated by protein-protein interactions and as a result are of particular interest in the area of protein therapeutics development. The main concern of protein therapeutics is the propensity of the proteins to form aggregates during manufacture, transport, storage and delivery. The commercial viability of protein therapeutics depends on its successful formulation especially during long-term storage where attractive protein-protein interactions could lead to reversible self-association and/or aggregation [1-6]. Aggregates which are usually irreversible need to be controlled in order to maintain the stability of the protein therapeutic and minimize the risk of immunogenic responses.

For liquid therapeutic formulations, achieving protein stabilisation requires the addition of excipients (sometimes referred to as additives or co-solvents) such as sugars [7, 8], polyols [9], salts [10], surfactants [11, 12], amino acids [13-16] and other buffers [17]. These excipients interact with proteins thereby affecting protein solubility and folding stability. Unfortunately due to the sheer complexity of proteins, it is difficult to rationalise the exact mechanism through which excipients stabilise proteins when exposed to various stress conditions such as pH, temperature, salt type, ionic strength, and buffer composition.

Protein aggregation is controlled by either conformational or colloidal stability, which in turn are governed by intrinsic or extrinsic factors. Changes in conformational stability lead to exposure of buried nonpolar residues on the protein surface as the native protein unfolds. Colloidal stability depends on the balance between attractive and repulsive intermolecular charge-charge interactions between protein molecules. Thus, the role of conformational and colloidal stability is critical to identifying optimal and stabilising protein formulations.

One of the largest and fastest growing classes of biopharmaceutical products on the market includes antibodies and antibody-derived products. Like other protein therapeutics, antibodies have the tendency to aggregate under a range of formulation conditions. Therefore it is important to generate optimal solution conditions which ultimately control the formulation of the protein drugs.

6.3 Monoclonal Antibodies

Monoclonal antibodies (mAbs) are a rapidly growing class of protein therapeutics used in the treatment of immune and cancer disorders [18]. So far, about 30 mAbs have been approved by the United States Food and Drug Administration (FDA) while over 200 other mAbs are in various stages of development pipeline. mAbs are a highly specific class of immunoglobulins that make up the humoral branch of the immune system. They are produced by the immune system in response to foreign molecules in the body. The basic structures and functions of mAbs are well established. They consist of two identical heavy (H) and two identical light (L) polypeptide chains joined together by disulphide bridges. A schematic of a monoclonal antibody is shown in Figure 6.1. The H-chains are covalently attached to oligosaccharide groups while the L-chains are non-glycosylated. The two H-chains are connected together at the flexible or hinge region of the H-chains. The four polypeptide chains contain constant (C) and variable (V) regions located at the carboxyl and amino terminal of the chains respectively. Both L- and H- chains contain a single variable (V_L and V_H) region. The L-chain has one constant (C_L) region while the H- chain has three constant (C_H1, C_H2, C_H3) regions. The variable regions in both H and L chains join together to form two identical antigen binding (Fab) sites. The constant (Fc) or fragment crystallisable region is responsible for

effector functions and binding such as placental transport. Most mAbs adopt a Y-shaped configuration resulting from the folding of the polypeptide chains.

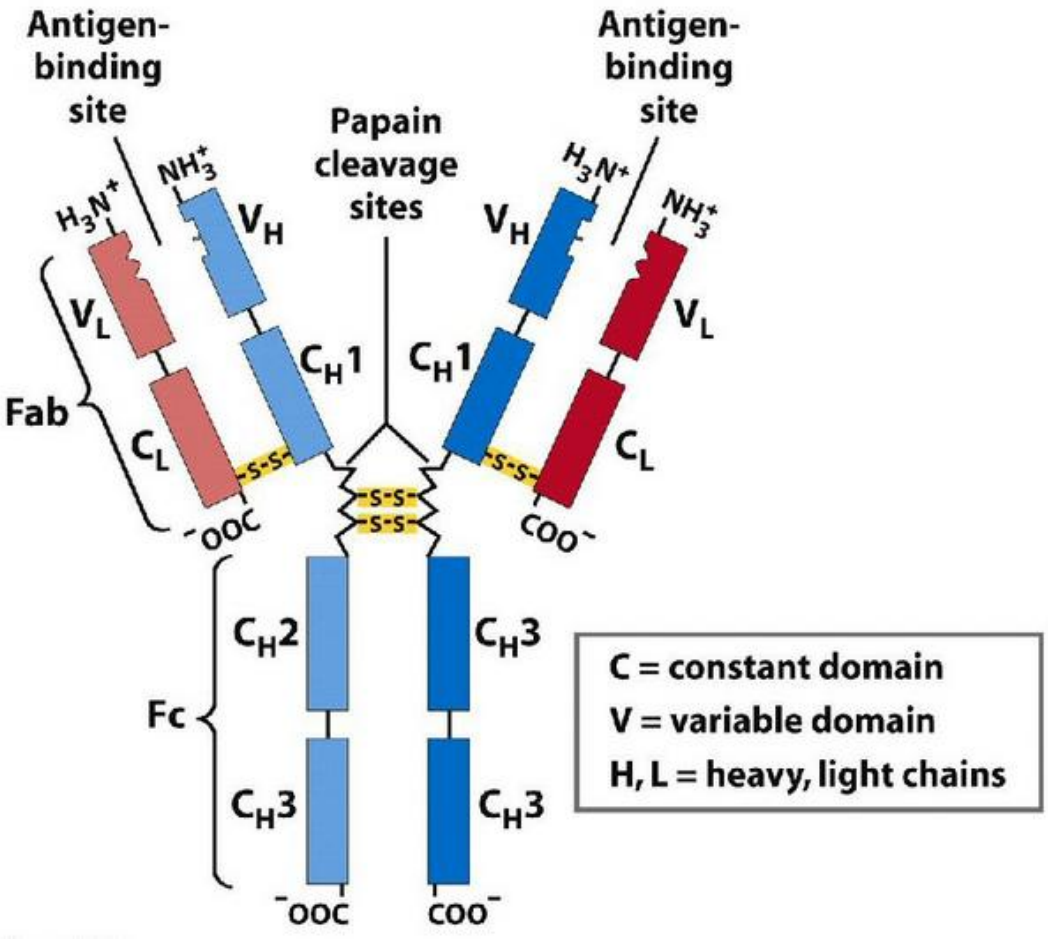


Figure 6.1. Basic structure of an antibody or immunoglobulin Ig molecule. The variable (Fab) domains located at the arm of the Y bind antigens in a lock- and –key mechanism. The fragment constant (Fc) portion at the carboxyl terminus of the H-chains responsible for biological activity [19]

Antibodies are divided into 5 main classes based on their H-chain or C-regions; IgA (α), IgD (δ), IgE (ε), IgM (μ), IgG (γ). They have an approximate molecular weight of 150 kDa (with 50 kDa on the H-chains and 25 kDa on the L-chains) [20]. The IgG class will be discussed in further details as it is the class of mAbs used to conduct

experiments in this chapter. There are four subclasses of the IgG molecule in humans with heavy chains γ1, γ2, γ3 and γ4 giving rise to IgG1, IgG2, IgG3 and IgG4 respectively. The differences arise due to structural differences in the number and location of the disulphide bonds in the heavy chains and length of the hinge region [20].

Like other therapeutic proteins, mAbs are relatively unstable and may have more than one degradation pathway initiated by changes in pH, salt type and concentration, and temperature. Hence it is not possible to generate a single optimal formulation for all mAb-based therapeutics. Understanding antibody aggregation behaviour under different conditions may allow us to understand effects of formulation conditions on protein behaviour to enable predictive methods.

Protein-protein interactions have been strongly correlated to the aggregation propensity of mAbs under various formulation conditions [14, 17, 21-25]. The solution pH controls the ionization properties of charged residues on the surface of proteins and thus, affects electrostatic interactions between protein molecules. Most mAbs are formulated at slightly acidic pH ($\leq$ 6.5) as formulations at neutral and basic pH values enhance chemical instabilities such as deamidation and isomerization [14, 26]. The effect of pH on mAb aggregation can affect either charge-charge interactions and/or destabilisation. The stability of an IgG2 mAb was shown to be sensitive to pH changes as protein-protein attractions increased from pH 4 to 9 due to reducing protein net charge and lowering protein-protein charge repulsion [2]. In addition, pH-induced destabilisation has been observed at low pH values resulting from partial unfolding of the protein which can lead to aggregation even when there is high colloidal stability or increased protein repulsion [27].

Buffer composition also plays an important role in modulating and stabilising protein-protein interactions. Histidine buffers are widely used in preparation of mAb formulations due to their stabilising properties [28-30]. Histidine has a pKa of ~ 6 which makes it suitable for use at slightly acidic conditions. Usually, surfactants are included in mAb formulations to stabilise and prevent proteins form adsorption to interfaces or interfacial stress [31, 32]. Non-ionic surfactants, polysorbate 20 and 80 (Tween 20 and 80) are two most commonly used surfactants in mAb formulations for stabilisation against surface-induced aggregation [11, 12, 33]. Surfactants are surface active amphiphilic agents (containing polar head and hydrophobic tail) that can also interact directly with proteins [26]. Other low molecular weight additives or excipients are usually added to formulations to improve the conformational and colloidal stability of mAbs.

The influence of salts on protein-protein interactions is dependent on a number of factors such as ion hydration, valence, size, net charge, surface charge density and polarizability [34-36]. Salts are known to have a concentration dependent effect on protein-protein interactions according to Hofmeister series [10, 17, 37, 38]. In chapter 4, we have discussed different mechanisms of protein-salt interactions in detail by using lysozyme as a model protein.

Sugars and amino acids have been reported to improve stability of proteins and as such are used in preparation of mAb formulations [7, 8, 14, 15, 39]. In a previous study, we have shown that the use of amino acid derivatives in insulin formulations could either act as an aggregation suppressor or enhancer depending on the solution conditions and rate controlling step of protein-protein interactions [13].

Since there is no ‘universal excipient’ for the stabilisation of all mAbs against aggregation, it is of paramount interest to develop other additives that exhibit

potential to be used as aggregation suppressors. With this in mind, we explored the applicability of polyelectrolytes which have not been fully exploited as additives to improve the solubility and stability of mAbs.

Polyelectrolytes (polyanions or polycations) have been used for modulating protein aggregation behaviour as they bind to and interact with proteins via different mechanisms depending on the pH and the ionic strength of the solution as well as the chemical nature of the protein and polyelectrolyte [40-44]. Inhibition of insulin aggregation by polyanion heparin was attributed to weakening of attractive electrostatic interactions between insulin molecules when the polyanion binds to and masks the exposed positively charged insulin surface [45].The effect of polyelectrolytes on protein-protein interaction is pH dependent. BSA (bovine serum albumin) aggregation was suppressed in presence of dextran sulphate (DS) at pH 5.1 and 6.2 but not at pH 7.5 [46]. Inhibition of association between denatured states was prevented via strong electrostatic interactions between the denatured states of BSA with dextran sulphate which led to formation of DS-BSA complexes. Conversely, the formation of complexes between polyelectrolytes and proteins of opposite charges has been observed to induce protein-protein association due to presence of attractive electrostatic interactions. Goers et al. [42] found that polycations such as polyethyleneimine formed strong complexes with negatively charged α-synuclein through electrostatic interactions which masked charge-charge repulsion to increase oligomerization and fibrillation of the protein. In another study, the addition of a negatively charged polyelectrolyte (poly vinyl sulfonate) to a positively charged protein (bovine pancreatic chymotrypsin) led to the formation of non-soluble complexes which induced precipitation and aggregation [41]. Here in order to understand the stabilizing mechanisms of a novel polycationic excipient, we have

investigated the aggregation behaviour of mixtures of a mAb with the excipient at different pH values and ionic strengths.

The osmotic second virial coefficient (B_{22}) and diffusion interaction parameter (k_D) can be used as indicators to assess protein solution behaviour under different formulation conditions. B_{22} and k_D measurements reflect the averaged protein-protein interactions where an increase in protein-protein association has been linked to an increase in propensity for aggregation suggesting that aggregation is controlled in part by colloidal stability [2, 7, 17, 21-23, 25]. Conformational stability requires measurements of the free energy of unfolding (ΔG_{unf}) where excipients that increase ΔG_{unf} are known to stabilise against aggregation. Conformational stability is also reflected by an increase in unfolding temperature (T_m) values. Because aggregation is controlled by different mechanisms, it is difficult to determine the exact contribution of colloidal and conformational stability and thus, may make interpretation of results challenging.

In this work, we investigated the conformational and colloidal stability characteristics of an IgG4 under different formulation conditions. Colloidal stability was assessed in terms of the osmotic second virial coefficient (B_{22}) from static light scattering and a diffusion interaction parameter (k_D) from dynamic light scattering. Measurements were carried out as function of pH and ionic strength in the presence of sodium chloride (NaCl) or 1, 2 propanediol propylene glycol (PD). The pH was varied to investigate how changes in the protein net charge control electrostatic interactions. The pH dependences of protein-protein interactions were correlated with the aggregation propensity of the mAb using temperature ramped dynamic light scattering and intrinsic fluorescence experiments. We then assessed complex formation of a novel polycation additive (polyethyleneimine, PEI) with the protein

and the impact on the aggregation behaviour. Based on our findings, we were able to establish the contribution of the polycation to colloidal and conformational stability of the mAb under different formulation conditions.

6.4 Materials and Methods

Purified IgG4 (pI between 5 and 6) was provided by Genzyme with an initial protein concentration of 8.13 mg/mL and stored in 24 mM sodium phosphate at pH 7.2 with 28 mM sodium chloride at -20°C.

Frozen samples (~3 mL) were thawed at room temperature and extensively dialyzed (Spectra/Por Float-A-lyser dialysis tubing, molecular weight cut-off 3.5-5 kDa; Spectrum Labs, Cal) against the desired buffer conditions at 4°C. Buffer conditions were 5 mM L- Histidine (Sigma Aldrich) and ~ 0.01 % Tween-80/ litre (Sigma Aldrich) at selected pH values (5, 6, and 7) unless otherwise stated. Buffer solutions were prepared by dissolving histidine in 3 litres of distilled deionized water (ddH_2O) from a Millipore Milli-Q filtration system and titrating to the desired pH with drops of 0.5 M hydrochloric acid or sodium hydroxide. About 4 mL of 100 times diluted Tween was added to the buffer and pH checked before making up the solution to 4 litres with ddH_2O. Protein solutions were dialyzed against a given buffer condition via two 1-L exchanges of 24 hours each after which protein concentration and solution pH were verified. The concentration of protein was determined using an extinction coefficient of 1.356 mL/mg at 280nm. About 500 mL of the second dialysate was filtered using 0.2 µm cut-off sterile syringe filters (Millipore) and used to prepare stock solutions as described below. Table 6.1 shows the final working concentrations of each of the bulk additive solutions that were used in preparing the various formulations. All experiments were carried out at different pH values, high and low ionic strength, and with and without the polycation additive.

Table 6.1. Final additive concentrations (C_2) prepared from initial stock concentrations (C_1) of each additive at different pH values

Conditions	**1.2 M PD**	**600 mM NaCl**	**50% w/v PEI (20 mg/ mL)**	**pH**
M_1	0.3 M			5
M_2	0.3 M			6
M_3	0.3 M			7
M_4		150 mM		5
M_5		150 mM		6
M_6		150 mM		7
M_7	0.3 M		5 mg/mL	5
M_8	0.3 M		5 mg/mL	6
M_9	0.3 M		5 mg/mL	7
M_{10}		150 mM	5 mg/mL	5
M_{11}		150 mM	5 mg/mL	6
M_{12}		150 mM	5 mg/mL	7

1, 2 -propanediol propylene glycol (PD) solution with a molarity of 13.1 M purchased from Sigma Aldrich was diluted with histidine/tween buffer to obtain a working stock concentration of 1.2 M at the desired pH. PD buffer solutions were then diluted to a final concentration of 0.3 M by making a 4 times dilution of 1.2 M PD buffer using the corresponding histidine/tween buffer. For sodium chloride (NaCl) (Fischer Scientific) experiments, 0.6 M NaCl stock solutions were prepared by dissolving the salt in the histidine/tween buffer at the desired pH. The stock salt

solutions were then diluted to a final concentration of 0.15 M NaCl in the corresponding histidine/tween buffers. For polyethyleneimine (PEI, molecular weight ~ 2000 Da, Sigma Aldrich) experiments, stock solutions of PD and PEI were prepared at 1.2 M and 20 mg/mL respectively in desired histidine/tween buffers. The PD/PEI buffer was gently stirred and diluted with the corresponding histidine/tween buffers to a final concentration of 0.3 M PD and 5 mg/mL PEI. For NaCl/PEI experiments, stock solutions of 0.6 M NaCl and 20 mg/mL PEI were prepared in desired histidine/tween buffers. The sample was gently mixed before diluting with corresponding histidine/tween buffers to obtain a final concentration of 0.15 M NaCl and 5 mg/mL PEI.

The pH of all final formulation buffers were checked and adjusted where necessary before filtering with 0.1 um syringe filters. The initial concentration (C_1) of each additive was prepared to a working volume of approximately 50 mL while the after dilution, the final formulation buffers with concentration C_2, was prepared to a working volume of 40 mL. For both light scattering and fluorescence experiments, protein samples were prepared by dilution of the dialysed protein solutions with desired stock concentrations C_1 to a final volume of 5 mL. This method allows for the protein solution to contain the same concentration as the additive. Protein sample solutions were degassed for 1 hour before filtering with 0.1 μm syringe filters (Millipore).

6.5 Methods

6.5.1 Static light scattering

Static light scattering (SLS) experiments were performed to obtain the osmotic second virial coefficient (B_{22}), and/or protein molecular weight values using the

miniDAWN Treos instrument for the various formulations. For each experiment, stock samples (~ 5 g/L) of the mAb were systematically recycled and diluted with the corresponding formulation buffer to generate a series of mAb concentrations in order to conserve sample. The samples were injected into the detectors connected in-line to a UV concentration detector. Buffer only samples were used as a baseline control.

The raw static light scattering data was extracted from the software and analysed using an excel spread sheet. The excess Rayleigh ratios at 90° (R_θ) obtained from SLS readings were used to generate the Debye plots according to

$$\frac{Kc}{R_\theta} = \frac{1}{M_{w,0}} + 2B_{22}c \qquad \text{(Equation 6.1)}$$

where c is protein concentration, K is an optical constant controlled by the solvent properties, and $M_{w,0}$ is the protein infinite dilution molecular weight. Alternatively, we can define an apparent molecular weight $M_{w,app}$ by

$$\frac{R_\theta}{Kc} = M_{w,app} \qquad \text{(Equation 6.2)}$$

where $M_{w,app}$ accounts for attractive and/or repulsive protein-protein interactions. Equation 6.2 is used when protein reversible association occurs at low protein concentrations. When the debye equation was fitted to the light scattering data, the plot of Kc/R_θ versus c yields a slope equal to B_{22} and the inverse of the y-intercept equal to $M_{w,o}$. Errors were calculated based on the standard deviation obtained from linear regression of Equation 6.1.

6.5.2 Dynamic light scattering

Dynamic light scattering measurements were conducted to determine the intensity auto-correlation function and extract the diffusion coefficient D_T using the cumulant

analysis method given in Chapter 3. The measured diffusion coefficient D_T was plotted against protein concentration c to obtain the diffusion coefficient at infinite dilution D_0, which was used in the calculation of the hydrodynamic radius $R_{h,0}$ using the Stokes–Einstein relation

$$R_{h,0} = \frac{k_B T}{6\pi\eta D_0} \qquad \text{(Equation 6.3)}$$

where k_B is Boltzmann constant, η is solvent viscosity and T is absolute temperature.

A plot of D_T/D_0 against c can be used to obtain the interaction parameter k_D or slope to reflect the nature of protein-protein interactions.

$$D_T = D_0(1 + k_D c) \qquad \text{(Equation 6.4)}$$

Equation 6.4 indicates the measured D_T depends on protein-protein interactions through the k_D parameter. We can also define an apparent hydrodynamic size $R_{h,app}$ using an analogy to the Stokes Einstein relation

$$R_{h,app} = \frac{k_B T}{6\pi\eta D_T} \qquad \text{(Equation 6.5)}$$

The advantage of reporting an apparent size versus a diffusion coefficient is that the effect of the solvent viscosity does not impact the hydrodynamic size. DLS at room temperature were performed simultaneously with SLS measurements described above using the miniDAWN Treos and the measured diffusion coefficients obtained using the ASTRA 6 software (Wyatt, Santa Barbara, California) were exported to an excel spread sheet for calculation of the hydrodynamic sizes. Errors were calculated based on the standard deviation obtained from linear regression of Eq. 6.5.

Temperature-controlled DLS studies were carried out using the DynaPro Plate Reader (Wyatt, Santa Barbara, USA) equipped with an 831nm laser. Auto-attenuation was applied to control the signal and strength of the laser beam. The

acquisition time was set to 5 seconds and 60 measurements were performed at each temperature. Temperature was decreased or increased in 5° C increments.

mAb formulations at pH 5 were investigated at five different protein concentrations between 2 and 14 g/L. 40 uL of the samples were added to a 384-well UV-Star Clear Microplate (781801) (Greiner Bio-One GmbH, Germany) in triplicates for each condition. The samples were capped with 20 uL of silicon oil to prevent solvent evaporation. Silicon oil is insoluble in water and has no effect on the measurements. The samples were gently centrifuged at 4000g for 1 minute and lint-free wipes were used to clean the bottom of the plates. The measured diffusion coefficients were analysed by DYNAMICS V6 software (Wyatt, Santa Barbara, USA) and exported to an excel spread sheet for calculation of the interaction parameter and hydrodynamic sizes. The diffusivity readings at the same protein concentration were averaged as measurements were performed in triplicates. Errors were calculated based on the standard deviation obtained from linear regression of Eq. 6.5.

6.5.3 Temperature-controlled light scattering and intrinsic fluorescence experiments

The Optim (Avacta Analytical Ltd, UK) was utilised to perform temperature controlled light scattering and intrinsic fluorescence studies in which the mAb was exposed to temperature scans between 20 °C and 90 °C with a 266 nm laser source at a scan rate of 2 °C per minute. mAb formulations were prepared at a final concentration of 1 g/L and 10 g/L at pH 5 and 7. Protein and buffer solutions were prepared as described above and degassed before running the experiments. Degassing is needed to avoid the introduction of air bubbles which could potentially affect the results. 9 uL of the samples were aliquoted into micro cuvettes and sealed

at the bottom to prevent any leakages. Measurements were performed in duplicates. Colloidal stability was monitored by measuring the light scattering intensities as a function of temperature to determine the aggregation temperature (T_{agg}), which corresponds to the temperature on-set of a light scattering increase. T_{agg} is calculated at the temperature where rate of change of the light scattering signal or differentiated scattering intensity (*dSI/dt*) versus temperature reaches a threshold that is about 10% of the change in signal. Structural stability was examined by recording the fluorescence spectra of the specified temperatures over the wavelength interval of 200 to 500 nm using a slit width of 100 μm. Fluorescence readings were used to determine the unfolding or melting temperature (T_m) of the protein. The T_m values were calculated from a plot of the differentiated fluorescence intensity (*dFI/dt*) versus temperature to give a peak maximum which corresponds to T_m.

6.6 Results

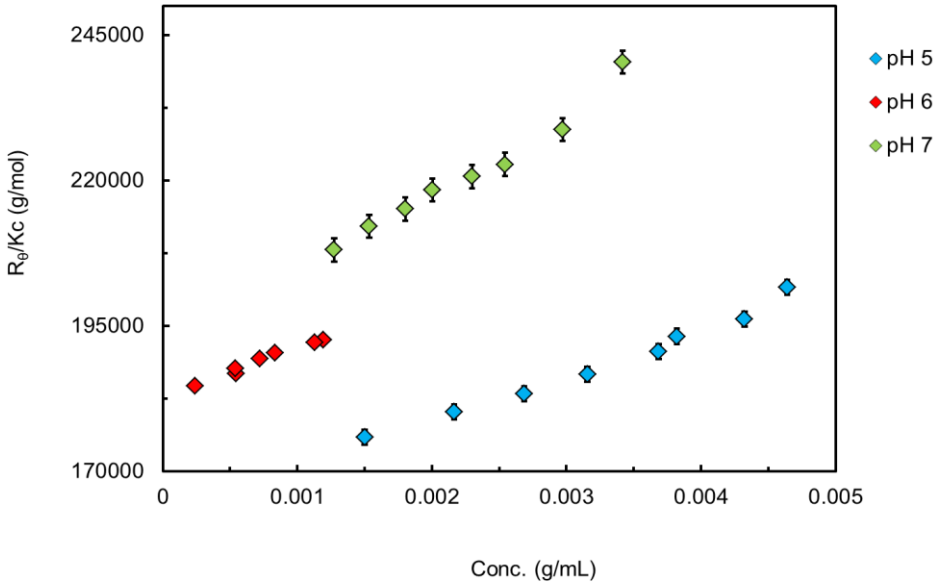
Static and dynamic light scattering were used to determine the colloidal stability in terms of the osmotic second virial coefficient (B_{22}) and diffusion interaction parameter (k_D), respectively, or alternatively, for strongly associating systems, in terms of an apparent weight average molecular weight and hydrodynamic radius. In addition, the infinite dilution values for the hydrodynamic size and weighted average molecular weight were determined to provide information about protein self-association or protein - polycation complexation. Intrinsic fluorescence was applied to investigate the effect of protein-polycation complex formation on the structural stability of the protein and determine whether the polycation inhibits aggregation when bound to the protein. The aggregation temperatures (T_{agg}) were used as a measure of the protein aggregation propensity.

6.6.1 Effect of pH and ionic strength on intermolecular interactions of IgG4

The effect of pH and ionic strength on protein-protein interactions was assessed for the IgG4 antibody with a maximum concentration of 5 g/L at room temperature. In Figure 6.2, values of the excess Rayleigh ratio (R_θ/Kc) was fit to the Debye equation and plotted as a function of protein concentration for solutions containing 300 mM PD at pH 5, 6, and 7. In addition Figure 6.2B is a plot of the apparent hydrodynamic radius obtained from the dynamic light scattering experiments. An upward slope of the static light scattering plot indicates net attractive interactions and a downward slope indicates net repulsion. As the pH of the formulation is increased from 5 to 7, the slope decreases reflecting attractive protein-protein interactions with B_{22} in the order of -11.1 x 10^{-5} mL mol/g^2 (pH 5), -11.8 x 10^{-5} mL mol/g^2 (pH 6), and -13 x 10^{-5} mL mol/g^2 at pH 7. Increasing pH should favour attractive electrostatic interactions or weaker repulsive electrostatics where the net charge of the protein reduces as it approaches its isoelectric point. Sahin et al. [47] studied the aggregation behaviour of four IgG1 antibodies at low ionic strength. They observed that protein-protein interactions are more attractive as the pH is increased from 3.5 to 6.5 at a fixed ionic strength.

However, the weak pH dependence observed is an indication that some other forces are responsible for strong attractive protein-protein interactions. It is possible that the self-attraction follows an oligomerization model, which is not captured with the measured osmotic second virial coefficient. The molecular weight at infinite dilution changes from 166 kDa to 194 kDa as the pH is increased from 5 to 7. This value is greater than the monomer molecular weight of a mAb which is ~ 150 kDa [24]

indicating self-association is occurring at concentrations below the experimentally studied concentration range. The molecular weight dependence on the pH reveals that increased self-association occurs with increasing pH. The static light scattering results are in qualitative agreement with the measured hydrodynamic sizes, which are greatest at pH 7. In addition, a positive slope in the plot of $R_{h,app}$ versus protein concentration reflects the presence of attractive protein-protein interactions.



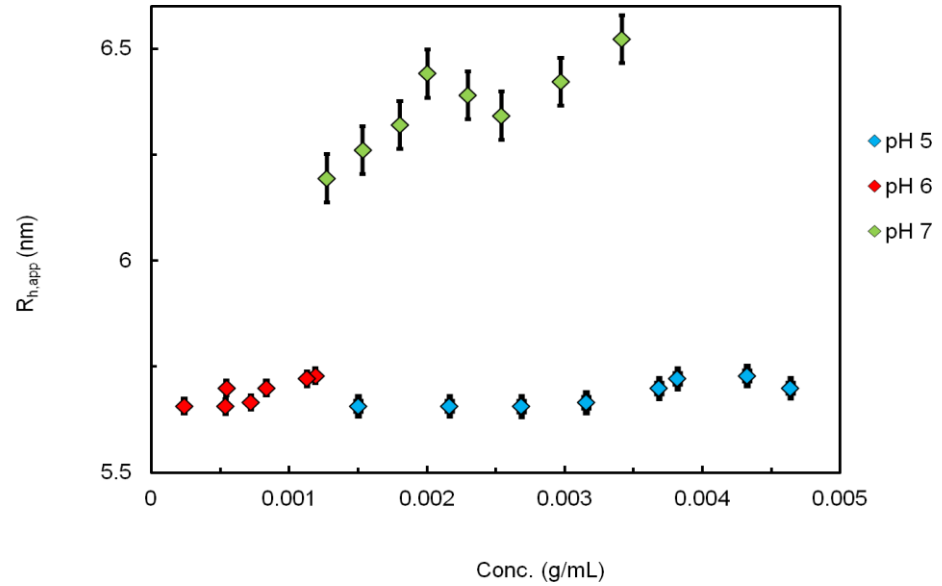


Figure 6.2. (A) Static light scattering measurements for determination of B_{22} values and (B) dynamic light scattering measurements for determination of $R_{h,0}$ of an IgG4 antibody in 300 mM PD solutions at pH 5, 6 and 7. Virial coefficient values decrease with increasing pH: -11.1 x 10^{-5} mL mol/g^2 (pH 5), -11.8 x 10^{-5} mL mol/g^2 (pH 6) and -13 x 10^{-5} mL mol/g^2 (pH 7) and correspond to $M_{w,0}$ values of 166 kDa, 183 kDa and 194 kDa respectively.

The diffusion coefficient at infinite dilution was used to calculate the infinite hydrodynamic sizes ($R_{h,0}$) of the protein. At infinite dilution, the $R_{h,0}$ values are 5.6 nm at pH 5 and 6, and 6.1 nm at pH 7. This value roughly corresponds to the size of a mAb [48]. $R_{h,0}$ reflects self-association below experimental measurements which is why the size at pH 7 is larger than at pH 6 or pH 5.

pH is an important parameter that affects the ionization properties of charges residues on the protein surface and influences intermolecular interactions by altering the electrostatic properties of proteins especially at low ionic strength [49]. However as the ionic strength is increased, long-range electrostatic interactions are screened and the effect of pH on protein-protein interactions is controlled by more subtle mechanisms. Measurements of R_θ/Kc as a function of protein concentration for solutions containing 150 mM NaCl at pH 5, 6, and 7 were carried out to reduce the contribution from electrostatic interactions. However, for these experiments, the light scattering data did not reach a steady-state value at low protein concentration. As a consequence, the data is not reported here.

6.6.2 Effect of polyethyleneimine (PEI) on protein-protein interactions

The effect of PEI on the colloidal stability of the antibody was investigated in formulations containing 300 mM PD and 150 mM NaCl. Polyethyleneimine, a low molecular weight polymer is a highly branched and positively charged polyelectrolyte (or a polycation) [50] and as a result is expected to bind and mask exposed negative charges on the surface of the protein. It is important to note that due to the sheer complexity/ multitude of interactions (hydrogen bonding, van der Waals and electrostatic interactions, salt bridging and ion binding) which affect protein – protein interactions, there may be more than one effect of the additive on

protein interactions. Figure 6.3 shows a plot of excess Rayleigh ratios (R_θ/Kc) and $R_{h,app}$ as a function of protein concentration for solutions containing 300 mM PD and 5 mg/mL PEI at pH 5, 6, and 7. The light scattering data was fit to the debye equation to yield a straight line with a slope equal to the value of B_{22}. There is an observed pH dependence on B_{22} values as protein-protein interactions become more attractive or less repulsive with increasing pH. B_{22} values are positive at pH 5 (5.2 x 10^{-5} mL-mol/g^2) indicating repulsive protein- protein interactions and decrease to negative values at pH 6 (-4.3 x 10^{-5} mL-mol/g^2) and pH 7 (-3.6 x 10^{-5} mL-mol/g^2). Addition of PEI makes protein-protein interactions more repulsive as B_{22} values are higher than in corresponding solutions without PEI. The pH dependence of the protein-protein interactions could be due to the impact of repulsive electrostatic interactions (as interactions are more repulsive at pH 5 than at pH 7) or possibly due to a pH dependence of the PEI binding to the protein.

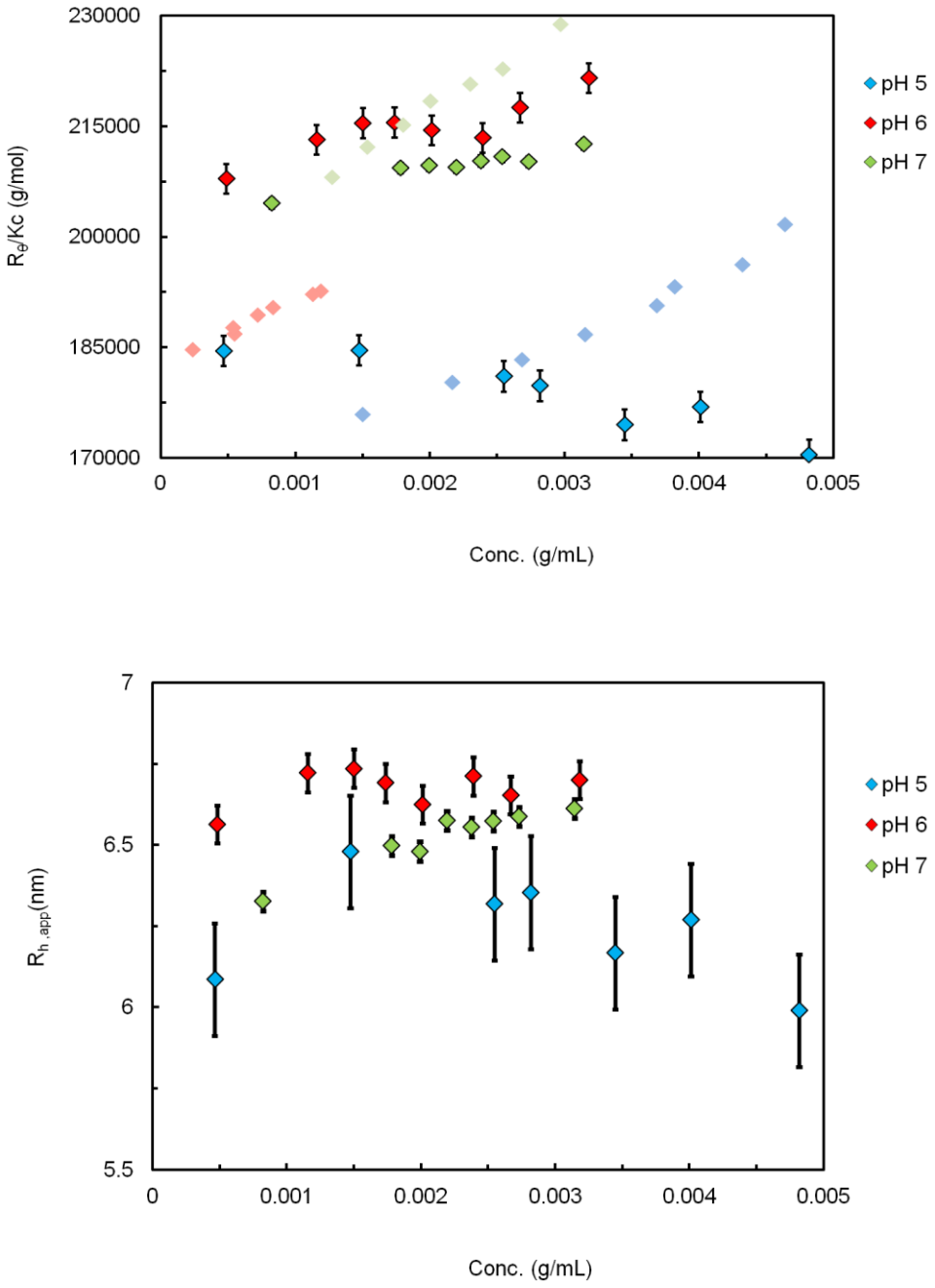


Figure 6.3. (A) Static light scattering measurements for determination of B_{22} values and (B) dynamic light scattering for determination of $R_{h,0}$ of an IgG4 antibody in 300 mM PD and 5 mg/mL PEI solutions (dark symbols) at pH 5, 6 and 7. Light symbols corresponding to solutions without PEI. Virial coefficient values in presence of PEI are greater compared to when there is no PEI and decrease with increasing pH: 5.2 x

10^{-5} mL-mol/g^2 (pH 5), -4.3 x 10^{-5} mL-mol/g^2 (pH 6) and -3.6 x 10^{-5} mL-mol/g^2 (pH 7) and correspond to $M_{w,o}$ values of 188 kDa, 207 kDa and 203 kDa respectively.

We observe that in the presence of PEI there is an increase in $M_{w,0}$ indicating polycation binding to the protein as the measured $M_{w,0}$ is slightly higher than the corresponding value measured in solutions without PEI at pH 5. Alternatively, the high measured value of $M_{w,0}$ could be due to protein self-association at low protein concentration. This cannot be the case, otherwise there would be an increase in $M_{w,app}$ at the experimentally measured concentrations as the association becomes saturated. The strong evidence that polycation binding is occurring at pH 5 also indicates binding is likely at pH 6 and pH 7 as the protein carries a lower positive charge. This is consistent with the measured $M_{w,0}$ values without PEI, which are smaller than the PEI values at the same pH. Furthermore, we notice that the $R_{h,0}$ values are also larger in solutions containing PEI (ranging between 6.3 nm to 6.6 nm) than when there is no PEI providing further evidence of polycation binding.

At pH 5, measured protein-protein interactions are more repulsive in the presence of PEI, which must be due to polycation binding to the protein. However it is unclear as to why the polycation binds to the protein. The observation of protein-polycation binding at pH below the pI indicates presence of negative patches on the protein. In addition to electrostatic interactions, other non-coulombic interactions (such as hydrophobic and hydrogen bonding) have been implicated in enhancing attractive interactions between proteins and polyelectrolytes. For instance, Sedlak and Antalik [51] suggested that ferricytochrome c can interact with the charged and uncharged part of a polyelectrolyte molecule to form complexes via coulombic and non-coulombic interactions respectively. It is likely that PEI binding shields or masks

the opposite charged patches on the protein surface to further enhance protein repulsion or inhibit electrostatic-driven self-association. Dubin and co-workers [44, 52, 53] showed that polycations can bind to proteins in solutions at pH values below the pI i.e. to positively charged proteins that contain negative patches. As the pH approaches the pI, the protein net charge is reduced and has less positive charge hence protein-polycation electrostatic interactions become more attractive at pH 6 and pH 7. From our results, we see that as the pH is raised, PEI binding to the protein increases as reflected by an increase in $M_{w,0}$. However, the measured values of B_{22} indicate the interactions between the protein-polycations are more attractive at the higher pH values.

Figure 6.4 shows a plot of excess Rayleigh ratios (R_θ/Kc) and $R_{h,app}$ as a function of protein concentration for solutions containing 150 mM NaCl and 5 mg/mL PEI at pH 5, 6, and 7. For pH 5 and 6, the static light scattering data was fit to the debye equation to yield a straight line with a slope equal to the value of B_{22} on the order of -7.3 x 10^{-5} mL-mol/g^2 (pH 5) and -10.3 x 10^{-5} mL-mol/g^2 (pH 6) . As the pH was raised, the slope decreases as B_{22} values are negative and protein-protein interactions become more attractive. Light scattering data obtained at pH 7 had a lot of noise and as such were not fit to the debye equation, nevertheless, the values still provide approximate values for the average molecular weight of species in the solution.

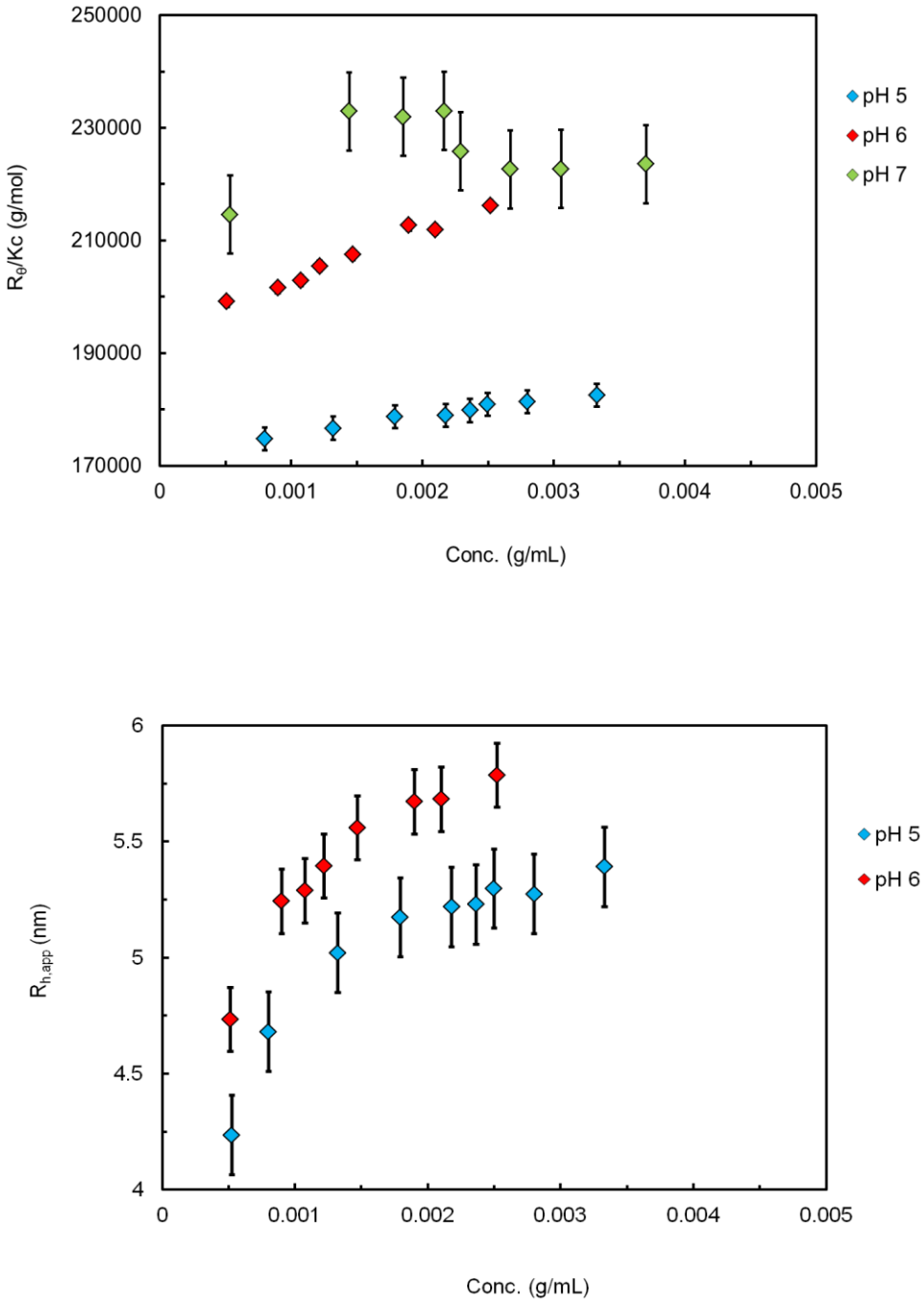


Figure 6.4. (A) Static light scattering measurements for determination of B_{22} values and (B) dynamic light scattering for determination of $R_{h,o}$ of an IgG4 antibody in 150 mM NaCl and 5 mg/mL PEI solutions (dark symbols) at pH 5, 6 and 7. Virial coefficient values are -7.3 x 10^{-5} mL-mol/g^2 (pH 5) and -10.3 x 10^{-5} mL-mol/g^2 (pH 6) and correspond to $M_{w,0}$ values of 169 kDa and 195 kDa respectively.

The infinite dilution molecular weight of the protein was higher at pH 6 (195 kDa) than at pH 5 (169 kDa) and is consistent with a higher hydrodynamic radius at pH 6 (4.7 nm) as compared to pH 5 (4.4 nm). The increase in the infinite dilution molecular weight over monomer value could be due to either polycation binding to the protein or due to self-association occurring at low protein concentrations. At pH 5, $M_{w,0}$ values in PEI-NaCl solutions are less than PEI-PD solutions, and however are slightly larger than the $M_{w,0}$ values obtained in solutions without NaCl and without PEI. The latter experiment is expected to have the greatest self-association, so that there is some evidence to suggest that PEI binding still occurs in NaCl solutions, but to a lesser extent than in the absence of sodium chloride. Similarly, at pH 6, the infinite dilution molecular weights observed with PEI and NaCl are between the values observed with PEI and no NaCl, and for solutions without either additive, also suggesting that PEI binding is occurring in the presence of NaCl. The protein-polyion binding is in part driven by electrostatic interactions, which will be screened by adding salt. The finding that some polyion-protein binding still occurs at higher salt concentrations provides an indication that there are also non-electrostatic, short-ranged interactions that drive the polycation-protein binding [45, 52]. This is consistent with the binding occurring at low pH where non-electrostatic interactions would also help overcome the unfavourable interaction between the positively charged polycation and the net positive charge of the protein.

Interestingly, the hydrodynamic sizes measured at low protein concentration were smaller than the average size of a mAb (i.e. less than 5 nm) indicating the presence of other small colloidal species. Dynamic light scattering measurements on protein-free solutions indicate PEI forms clusters or aggregates in the solution, which lowers the apparent $R_{h,0}$ value measured from the protein-containing solutions. At low

protein concentration, the small PEI aggregates contribute more to the light scattering signal and give an apparent average size less than the protein monomer. Nevertheless the B_{22} and apparent molecular weight values will only reflect the protein complexes since the contributions of small polycation clusters are accounted for by subtracting out the baseline scattering.

Interaction Parameter Determination

The DynaPro plate reader was used to quantify the hydrodynamic properties of the solutions as a function of temperature. At low temperature, measurements provide insight into the sizes of the protein-polycation complexes as well as the interactions between proteins or protein complexes. At higher temperatures, the dynamic light scattering is also used to quantify any aggregation processes. The additional advantage of the technique is that the sample has a much longer time to reach equilibrium. Samples with sodium chloride did not reach equilibrium at low protein concentrations on the flow through system used to generate the results in the previous section.

Measurements are reported in terms of a diffusion interaction parameter (k_D) for various formulations at pH 5 as shown in Figure 6.5. Values of k_D are derived from fitting the diffusion coefficient to Eq. 6.4 where an increase in k_D reflects net protein repulsion and vice-versa.

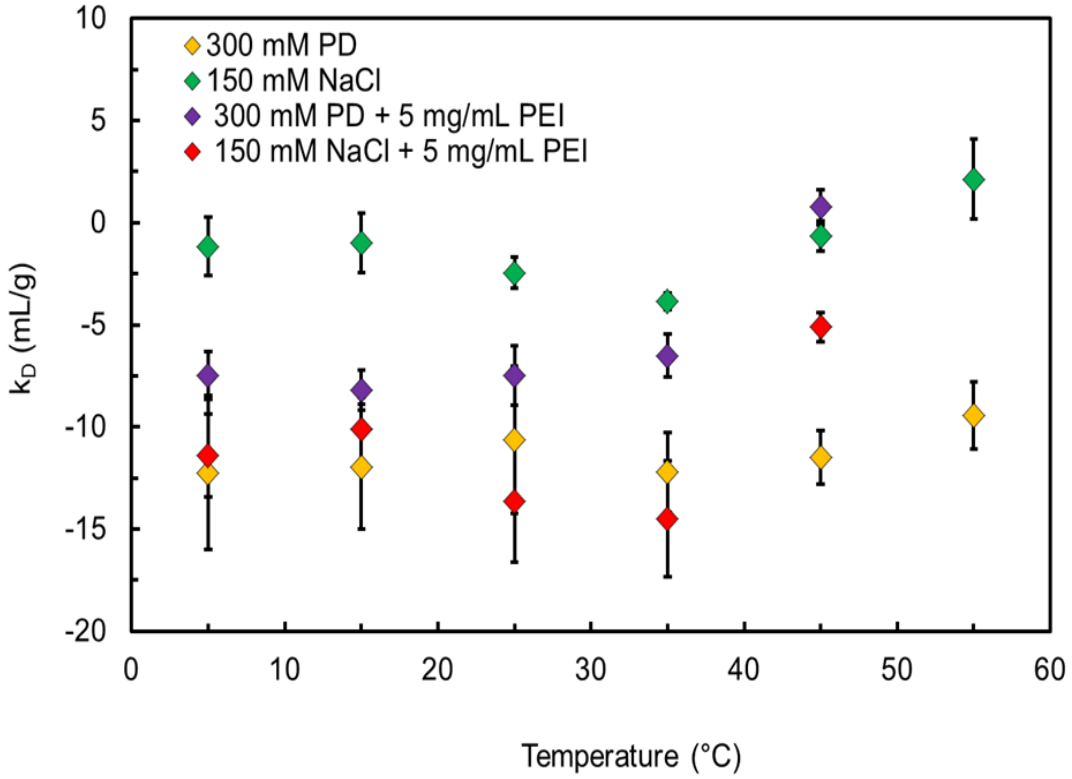


Figure 6.5. Dynamic light scattering measurements at 5 °C showing diffusion interaction parameter (k_D) as a function of temperature for various mAb formulations at pH 5. k_D values are on the order of -12.0 mL/g in PD, -11.4 mL/g in NaCl-PEI, -7.5 mL/g in PD-PEI, and -1.2 mL/g in NaCl solutions. Error bars are calculated from the linear regression calculated by averaging three independent measurements.

At pH 5, solutions containing only 150 mM NaCl reduced protein-protein attraction better than other formulations as reflected by the large k_D values. The main effect of adding NaCl is to reduce the contribution of electrostatic interactions by charge shielding, which is linked to the increase in repulsive protein-protein interactions. This implies that NaCl effectively screens any attractive electrostatic interactions between the protein molecules. Unfortunately this was not established in the static light scattering experiments since the measurements did not reach a steady value.

The apparent negative k_D value obtained on addition of PEI in NaCl solutions implies the PEI is promoting association between proteins. However, the results need to be interpreted with care as the $R_{h,0}$ value in the presence of PEI is smaller than the expected monomer value for the protein. As mentioned previously, PEI forms clusters in solution which give rise to an apparent negative k_D value as the contribution from the clusters to the measured size decreases with increasing protein concentration. However, SLS measurements show that protein-protein interactions are quite attractive in NaCl-PEI solutions as B_{22} values are very negative and almost similar to solutions where there is no NaCl or PEI.

k_D values are observed to increase upon addition of PEI to solutions containing 300 mM PD. It is also possible that there are PEI clusters, which would complicate the analysis further, thus the measured k_D is a lower limit to the actual value; if clusters did occur, the k_D would be higher than the apparent value. The result is consistent with the B_{22} measurements under the same condition where PEI improves the colloidal stability of the protein in solutions containing PD. According to the k_D results, sodium chloride rather than PEI is more effective at screening the protein-protein attraction. However, reaching a definitive conclusion is not possible due to the possibility of PEI cluster formation complicating the interpretation of the diffusion coefficient data in terms of k_D.

6.6.3 Influence of pH, ionic strength and polyethyleneimine on temperature - induced aggregation of an IgG4 antibody

Temperature ramped intrinsic fluorescence and static light scattering were used to assess the conformational stability and the aggregation propensity for various formulations at pH 5 and pH 7 in 1 g/L protein concentrations. The midpoint

temperature at which the protein starts to unfold or melt is reported as the melting temperature or T_m where lower T_m values are indicative of lower conformational stability and vice–versa.

The effect of temperature on the light scattering intensity of the protein formulations is commonly used for estimating an aggregation temperature, T_{agg}. This reflects the temperature onset of aggregation and corresponds to a sharp increase in the light scattering signal. At such temperatures, aggregation is usually caused by hydrophobic groups that are exposed in partially folded or unfolded states. T_{agg} values can also be used as a surrogate to assess the colloidal stability of partially folded proteins as higher T_{agg} values relative to the melting T_m reflect an increase in colloidal stability or reduced self-association of partially folded states.

Figure 6.6 and 6.7 shows a plot of fluorescence intensity as a function of temperature for formulations solutions containing 1 g/L protein at pH 5 and pH 7 respectively. The protein is seen to have two transition temperatures corresponding to the unfolding of different domains on the mAb. Differential scanning calorimetry experiments for an IgG1 antibody revealed multiple unfolding transitions at low pH where the unfolding of the C_H2 domain of the Fc region occurs at a lower temperature (T_{m1}) and the higher temperature transition (T_{m2}) is due to a combination of unfolding of the Fab region and C_H3 domain of the Fc [54]. This suggests that C_H2 domain is the least unstable region as it unfolds first with increasing temperature as has been observed in other studies [55-57]. However, the transition temperatures cannot be correlated based only on the Fab and Fc regions as the Fc region is more sensitive to changes in pH [57, 58]. Brummitt et al. [59] showed that at pH 4.5, the first transition temperature (T_{m1}) of an IgG was around 58.9 °C and the

second transition T_{m2} was observed at around 66 °C. As the pH is raised to 5.5, the conformational stability is increased as only one endotherm peak was observed as the melting temperatures shifted to a higher value of 68 °C.

Fluorescence measurements conducted in this study reveal the presence of two transition temperatures (T_{m1} and T_{m2}) in solutions with only PD at pH 5 corresponding to 62 °C and 74 °C respectively as shown in Table 6.2. As the pH was raised to 7, only one peak was observed as T_{m1} values shifted to 67 °C but second transition T_{m2} was not detected. It is not clear whether T_{m1} corresponds to only the Fab domain or unfolding of both the Fc and the Fab regions at 67 °C.

Table 6.2. Experimentally calculated values for all measured parameters using light scattering and fluorescence techniques.

Additive name	pH	$R_{h,0}$ (nm)	k_D (mL/g)	$M_{w,0}$ (kDa)	B_{22} (x 10^{-5} mL mol/g^2)	Tm$_1$	Tm$_2$	T_{agg}
PD	5	5.6	-13.5	166	-11.1	62	74	63
PD	6	5.6	-19.9	183	-11.8			
PD	7	6.1	-19	194	-13	67		56
NaCl	5					65	75	59
NaCl	6							
NaCl	7					67	75	65
PD-PEI	5	6.4	19.5	188	5.2	52	64.5	61
PD-PEI	6	6.6	-3.6	207	-4.3			
PD-PEI	7	6.3	-18.7	203	-3.6	57	68.9	63
NaCl-PEI	5	4.4	-66	169	-7.3	55	70	59
NaCl-PEI	6	4.7	-79	195	-10.3			
NaCl-PEI	7	5.6		225		61	66.9	65

At pH 5, increasing the ionic strength by addition of NaCl is seen to improve the conformational stability of the protein when there is no PEI in the solution. Melting temperatures increase from 62 °C and 74 °C in solutions of PD to 65 °C and 75 °C in NaCl solutions. This implies that NaCl does not lower conformational stability. However we observe a clear effect of PEI in lowering the conformational stability of the protein as shown in Figure 6.6. The two melting transitions are lowered to 52 °C and 65 °C from 62 °C and 74 °C when adding PEI to PD solutions, and reduced to 55 °C and 70 °C from 65 °C and 75 °C when adding PEI to NaCl solutions. From a thermodynamic analysis, a lowering of conformational stability is due to preferential adsorption of the additive (i.e. PEI) to the protein. This finding provides further support that PEI binds to the protein at pH 5. Previous studies have indicated that polyion binding can lower the conformational stability of proteins [51, 60]. The ability of a polyion to denature proteins has been correlated with the mechanism of binding. Proteins with denaturing ability are classified as strong binders as binding occurs even at high ionic strength when electrostatic interactions are screened [51]. PEI appears to fall into this category as the light scattering studies found binding also does occur in the presence of sodium chloride. It is likely that addition of PEI enhances intramolecular electrostatic repulsions due to the higher charge density upon binding. This leads to protein destabilisation because the charge density is lowered when the protein unfolds.

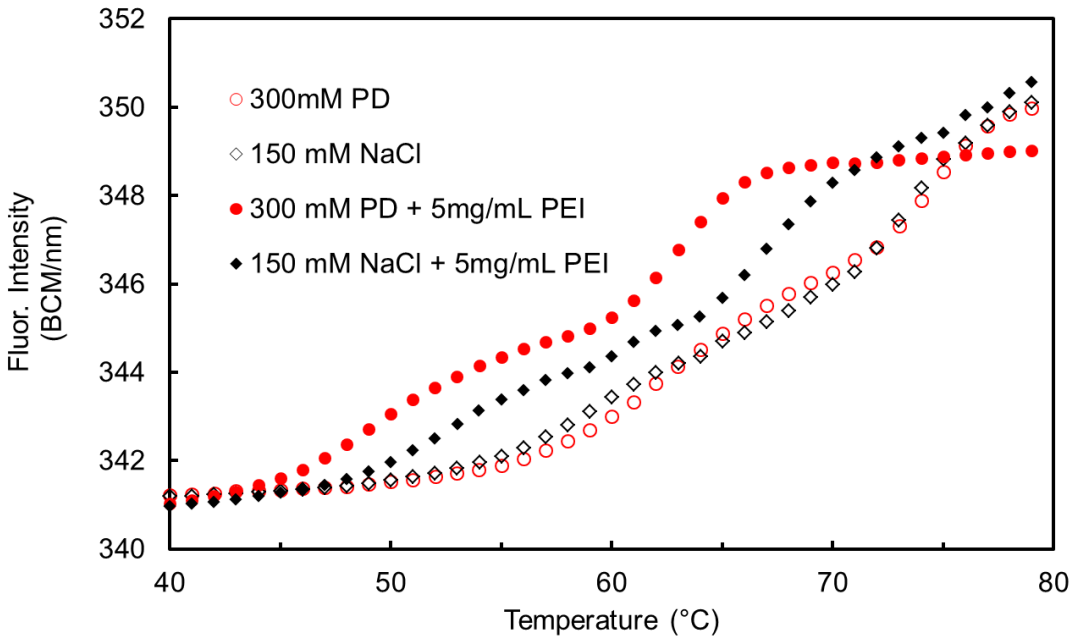


Figure 6.6. Intrinsic fluorescence intensity as a function of temperature at pH 5 in 1 g/L protein formulations. Results represent an average of two consecutive measurements.

At pH 7, although we observe two unfolding transitions, the effect of PEI is more pronounced in solutions containing PD rather than those containing NaCl as shown in Figure 6.7. From Table 6.2 above, a general trend is noticed where the melting temperatures are higher at pH 7 than at pH 5 for all the various formulations. At pH 7 (pH > pI), the protein has more negative charges on its surface and as such should be more susceptible to PEI binding. However the effect of PEI on the protein conformational stability is greater in solutions at low pH, although, the binding to the protein is stronger at high pH. This is in agreement with the general findings that proteins are more stable at pH values close to their isoelectric points [57]. Despite this, adding PEI to pH 7 solutions is seen to marginally lower the melting temperature or conformational stability of the protein.

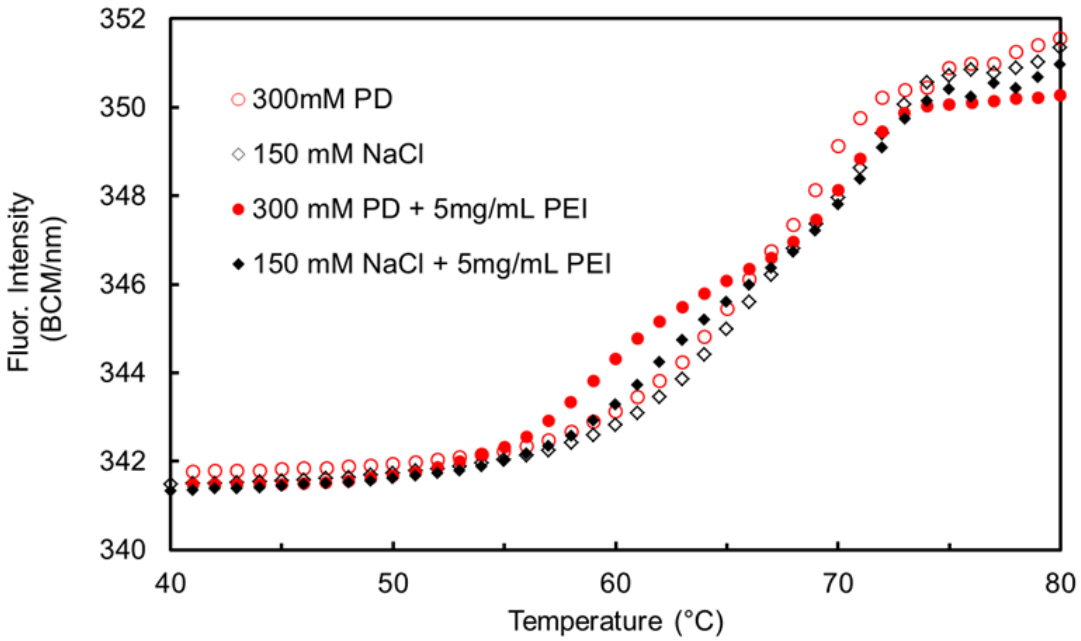


Figure 6.7. Intrinsic fluorescence intensity as a function of temperature at pH 7 in 1 g/L protein formulations. Results represent an average of two consecutive measurements.

Simultaneous acquisition of SLS data was performed to assess the aggregation propensity in terms of the aggregation onset temperatures (T_{agg}). Figure 6.8 and 6.9 shows the plot of the light scattering intensities as a function of temperature for formulations containing 1 g/L of protein at pH 5 and pH 7 respectively. Values for T_{agg} are reported in Table 6.2 above.

At pH 5, in solutions with only PD, the protein aggregates at temperatures (63 °C) lower than the second transition temperature T_{m2} (74 °C) values indicating that the less stable is most likely the aggregation prone state. As the ionic strength is increased, T_{agg} decreases while the low melting transition temperature increases. This implies that addition of NaCl reduces the colloidal stability of the protein. On addition of PEI, the protein starts to aggregate at lower temperatures with the effect

being greater in NaCl solutions versus in PD solutions. T_{agg} are lowered by 2°C when adding PD to PEI solutions, while there is no effect on T_{agg} when adding PEI to NaCl solutions. However in all solutions containing PEI, T_{agg} are higher than their corresponding T_{m1} values indicating the PEI does prevent the partially unfolded proteins from aggregating.

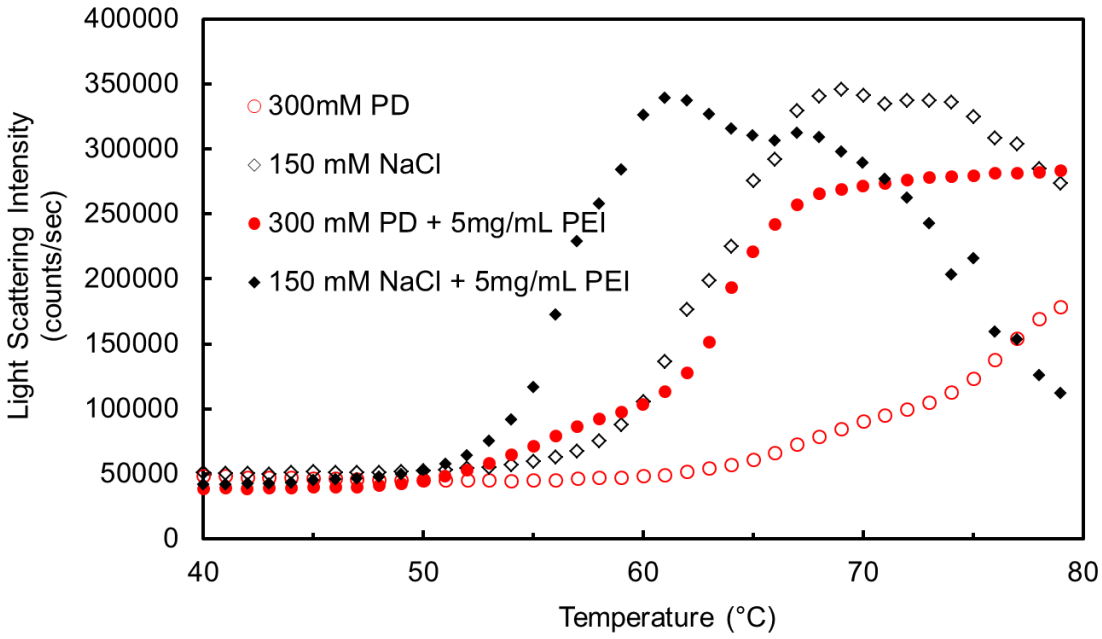


Figure 6.8. Static light scattering intensities as a function of temperature at pH 5 in 1 g/L protein formulations. Results represent an average of two consecutive measurements.

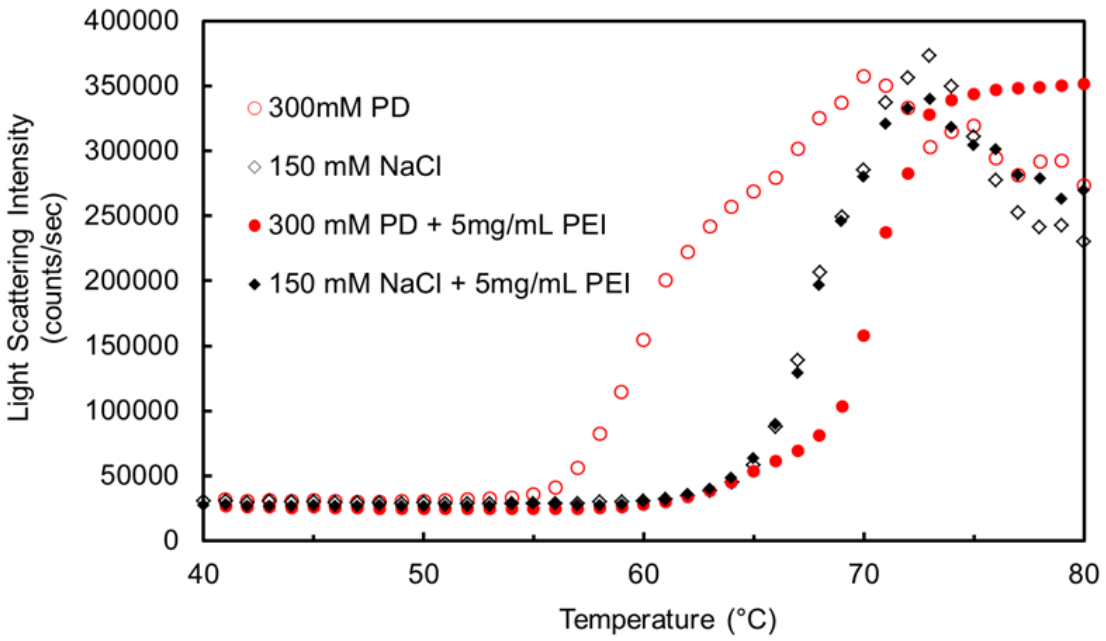


Figure 6.9. Static light scattering intensities as a function of temperature at pH 7 in 1 g/L protein formulations. Results represent an average of two consecutive measurements.

Figure 6.9 shows static light scattering measurements carried out at pH 7. At pH 7, in solutions with only PD, the protein aggregates at a lower temperature at pH 5 than at pH 7, although there is no corresponding decrease in the melting temperature, indicating the colloidal stability decreases with increasing pH. As the ionic strength is increased, T_{agg} increases to 65 °C while the melting temperature is unchanged indicating NaCl is enhancing colloidal stability. This finding is consistent with the k_D measurements where NaCl reduces self-association.

There is a clear effect of PEI on colloidal stability in solutions at pH 7 as the protein starts to aggregate at higher temperatures after PEI addition with the effect being greater in NaCl solutions versus in PD solutions. In all solutions containing PEI, T_{agg} are higher than their corresponding T_{m1} values. This finding indicates the PEI binding is protecting the partially unfolded protein from aggregation. In other words,

aggregation is controlled by the effect of PEI on the colloidal stability and the increased colloidal stability can be attributed to the electrostatic binding of the protein to PEI to induce formation of complexes, which prevents protein-protein attractions. Hence, a decrease in T_m does not result in a decrease in T_{agg}.

6.7 Discussion

The main purpose of this study is to investigate the effect of solution conditions such as pH, ionic strength and a polycationic additive (polyethyleneimine) on the aggregation behaviour of an IgG4 monoclonal antibody. Colloidal interactions were assessed from B_{22} and k_D measurements, while intrinsic fluorescence was used to determine conformational stability. It has been observed that slight changes in solution pH can affect the conformational and colloidal stability of proteins [61]. We observe that the IgG4 shows strong protein-protein association at all pH and the association become stronger with increasing pH to 7 as B_{22} and k_D decrease. The increased association at pH 7 is correlated with an increased aggregation propensity in the absence of salt or the polycation under thermal ramp measurements. This suggests that aggregation propensity (T_{agg}) is increased due to strong protein-protein association as the pH is increased from 5 to 7 despite an increase in conformational stability as the pH is raised. Similarly, an increase in colloidal stability was observed at pH 7 when increasing NaCl concentration. The reduced aggregation propensity measured at high temperatures with either decreasing pH or increasing ionic strength is correlated with a reduction in self-association between natively folded proteins measured at room temperature. This correlation implies the interactions between the natively folded proteins have similar origin to the interactions between partially folded proteins formed at temperatures needed to induce aggregation.

Addition of PEI to the protein solutions revealed a complex relationship as a function of pH between the conformational and colloidal stability of the protein. Polyelectrolyte - protein interactions are mainly driven by electrostatic interactions occurring between oppositely charged side chains leading to the formation of soluble and/or insoluble complexes [62, 63]. Thus increasing the pH above the pI should favour protein-polycation complex formation [64, 65]. Light scattering measurements at room temperature reveal PEI binding as the molecular weights and hydrodynamic sizes increase relative to when there is no PEI. In PD solutions, addition of PEI improves the colloidal stability of proteins as B_{22} and k_D are less negative or more positive. Although the strength of PEI binding increases with increasing pH, protein-protein attractions increase as the pH is raised from 5 to 7. Even at pH 5, the polycation binds to the positively charged protein leading to increased intermolecular electrostatic repulsion. Formation of complexes has also been observed in cases where both the protein and polyelectrolyte bear the same charge sign [52, 53, 63]. Xia et al. [53] found that complexes were formed when BSA and a polyanion had similar net charge due to the heterogeneous nature of the protein surface. In this case, it was postulated that the polyanion binds to positively charged patches on the BSA surface. There is also the possibility that the weak electrostatic repulsion existing between proteins and polyelectrolyte of the same charge sign is overcome by hydrophobic forces [52]. Fluorescence measurements show that PEI reduces the conformational stability of the protein and the effect is greatest at pH 5. Although binding causes the proteins to partially unfold and enhances destabilisation, the aggregation propensity of the proteins is significantly improved at pH 7, but not at pH 5.

On the other hand, colloidal measurements reveal protein-protein interactions are quite attractive on addition of PEI to NaCl solutions and the magnitude increases with increasing pH. PEI is also seen to bind to the proteins even when electrostatic interactions are screened on addition of salt as the apparent molecular weights are greater than when there is no PEI. It is likely that in this case, PEI binding is driven by hydrophobic interactions. We observe that addition of PEI to NaCl solutions also lowers the conformational stability of the protein at pH 5 more than at pH 7 but does not have any significant effect on the aggregation propensity of the protein at both pH values. Previous studies [51, 60] found that polyions that bind through non-electrostatic interactions, also reduce protein conformational stability, which is consistent with the effect of PEI measured in this study.

6.8 Conclusion

From these results, we can see that conformational stability, colloidal interactions and aggregation propensity are sensitive to changes in solution conditions such as pH and ionic strength. PEI lowers the conformational stability of the protein especially at low pH where the protein has a larger positive charge indicating that binding is also driven by, not only electrostatic interactions, but also hydrophobic interactions. At pH 5, protein-protein interactions are attractive as seen in B_{22} measurements and increases with increasing pH. Increasing the ionic strength by addition of NaCl weakens the protein-protein attraction as seen in k_D measurements as attractive electrostatic interactions are screened. Under accelerated conditions (i.e. at high temperatures), PEI increases the aggregation propensity of the protein at pH 5 as the protein aggregates at lower temperatures.

PEI binding lowers the conformational stability more so at pH 5, however, at both pH vales aggregation temperatures are greater than the melting temperature indicating the complexation of the polycation to the protein prevents the aggregation of partially unfolded states. This effect is greatest at pH 7. The stabilizing effect on the protein is also reflected by the decreased self-association probed from the measurements of protein-protein interactions at room temperature. It is likely that at pH 7 the proteins interact with the polycation via electrostatic interactions and reduces the propensity to aggregate therefore aggregation occurs at higher temperatures.

Hence, the overall effect of the polycation cannot be explained solely in terms of electrostatic interactions between PEI and the protein as PEI-protein complex formation occurs on the wrong side of the pI. The binding is driven by strong hydrophobic interactions which also leads to decreased protein conformational stability. Although, the protein conformation is changed, the complexes exhibit reduced aggregation tendency. From the results, we hypothesize that the effect of NaCl on increasing colloidal stability is greater than the polycation (PEI). Furthermore, since NaCl does not alter the conformational stability of the mAb, NaCl is a better aggregation suppressor than PEI.

6.9 References

[1] Neergaard, M. S., Nielsen, A. D., Parshad, H., De Weert, M. V., Stability of monoclonal antibodies at high-concentration: Head-to-head comparison of the IgG1 and IgG4 subclass. *Journal of pharmaceutical sciences* 2014, *103*, 115-127.

[2] Chari, R., Jerath, K., Badkar, A., Kalonia, D., Long- and Short-Range Electrostatic Interactions Affect the Rheology of Highly Concentrated Antibody Solutions. *Pharmaceutical Research* 2009, *26*, 2607-2618.

[3] Harn, N., Allan, C., Oliver, C., Middaugh, C. R., Highly concentrated monoclonal antibody solutions: Direct analysis of physical structure and thermal stability. *Journal of pharmaceutical sciences* 2007, *96*, 532-546.

[4] Inoue, N., Takai, E., Arakawa, T., and, Shiraki, K., Specific Decrease in Solution Viscosity of Antibodies by Arginine for Therapeutic Formulations. *Molecular Pharmaceutics* 2014, *11*, 1889-1896.

[5] Yadav, S., Laue, T. M., Kalonia, D. S., Singh, S. N., Shire, S. J., The Influence of Charge Distribution on Self-Association and Viscosity Behavior of Monoclonal Antibody Solutions. *Molecular Pharmaceutics* 2012, *9*, 791-802.

[6] 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. *Journal of Pharmaceutical Sciences* 2008, *97*, 4219-4227.

[7] Saito, S., Hasegawa, J., Kobayashi, N., Tomitsuka, T., *et al.*, Effects of Ionic Strength and Sugars on the Aggregation Propensity of Monoclonal Antibodies: Influence of Colloidal and Conformational Stabilities. *Pharmaceutical Research* 2013, *30*, 1263-1280.

[8] Rubin, J., Sharma, A., Linden, L., Bommarius, A. S., Behrens, S. H., Gauging Colloidal and Thermal Stability in Human IgG1–Sugar Solutions through Diffusivity Measurements. *The Journal of Physical Chemistry B* 2014, *118*, 2803-2809.

[9] 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, *63*, 1118-1159.

[10] Majumdar, R., Manikwar, P., Hickey, J. M., Samra, H. S., *et al.*, Effects of Salts from the Hofmeister Series on the Conformational Stability, Aggregation Propensity, and Local Flexibility of an IgG1 Monoclonal Antibody. *Biochemistry* 2013, *52*, 3376-3389.

[11] Kerwin, B. A., Polysorbates 20 and 80 used in the formulation of protein biotherapeutics: Structure and degradation pathways. *Journal of pharmaceutical sciences* 2008, *97*, 2924-2935.

[12] Kapp, S. J., Larsson, I., De Weert, M. V., CÁRdenas, M., Jorgensen, L., Competitive Adsorption of Monoclonal Antibodies and Nonionic Surfactants at Solid Hydrophobic Surfaces. *Journal of Pharmaceutical Sciences* 2015, *104*, 593-601.

[13] Nuhu, M. M., Curtis, R., Arginine dipeptides affect insulin aggregation in a pH- and ionic strength-dependent manner. *Biotechnology Journal* 2015, *10*, 404-416.

[14] Kheddo, P., Tracka, M., Armer, J., Dearman, R.J., Uddin, S., van der Walle, C. F., Golovanov, A. P., The effect of arginine glutamate on the stability of monoclonal antibodies in solution. *International Journal of Pharmaceutics* 2014, *473*, 126-133.

[15] Golovanov, A. P., Hautbergue, G. M., Wilson, S. A., and Lian, L., A Simple Method for Improving Protein Solubility and Long-Term Stability. *Journal of the American Chemical Society* 2004, *126*, 8933-8939.

[16] Fukuda, M., Kameoka, D., Torizawa, T., Saitoh, S., Yasutake, M., Imaeda, Y., Koga, A., Mizutani A., Thermodynamic and fluorescence analyses to determine mechanisms of IgG1 stabilization and destabilization by arginine. *Pharmaceutical Research* 2014, *31*, 992-1001.

[17] Roberts, D., Keeling, R., Tracka, M., van der Walle, C. F., *et al.*, Specific Ion and Buffer Effects on Protein–Protein Interactions of a Monoclonal Antibody. *Molecular Pharmaceutics* 2014, *12*, 179-193.

[18] Samra, H. S., He, F., Advancements in High Throughput Biophysical Technologies: Applications for Characterization and Screening during Early Formulation Development of Monoclonal Antibodies. *Molecular Pharmaceutics* 2012, *9*, 696-707.

[19] Nelson, D. L., Cox, M. M., Protein Function, *Lehninger Principles of Biochemistry*, W. H. Freeman & Company 2004.

[20] Wang, W., Singh, S., Zeng, D. L., King, K., Nema, S., Antibody structure, instability, and formulation. *Pharmaceutical Sciences* 2007, *96*, 1–26.

[21] Saluja, A., Fesinmeyer, R. M., Hogan, S., Brems, D. N., Gokarn, Y. R., Diffusion and Sedimentation Interaction Parameters for Measuring the Second Virial Coefficient and Their Utility as Predictors of Protein Aggregation. *Biophysical Journal* 2010, *99*, 2657-2665.

[22] Lehermayr, C., Mahler, H.-C., Mäder, K., Fischer, S., Assessment of net charge and protein–protein interactions of different monoclonal antibodies. *Journal of pharmaceutical sciences* 2011, *100*, 2551-2562.

[23] Connolly, B. D., Petry, C., Yadav, S., Demeule, B., Ciaccio, N., Moore, J. M., Shire, S. J., Gokarn, Y. R., Weak interactions govern the viscosity of concentrated antibody solutions: high-throughput analysis using the diffusion interaction parameter. *Biophys J.* 2012, *103*, 69-78.

[24] Scherer, T. M., Liu, J., Shire, S. J., Minton, A. P., Intermolecular Interactions of IgG1 Monoclonal Antibodies at High Concentrations Characterized by Light Scattering. *The Journal of Physical Chemistry B* 2010, *114*, 12948-12957.

[25] Roberts, D., Keeling, R., Tracka, M., van der Walle, C. F., *et al.*, The Role of Electrostatics in Protein–Protein Interactions of a Monoclonal Antibody. *Molecular Pharmaceutics* 2014, *11*, 2475-2489.

[26] Goswami, S., Wang, W., Arakawa, T., Ohtake, S., Developments and Challenges for mAb-Based Therapeutics. *Antibodies* 2013, *2*, 452-500.

[27] Bajaj, H., Sharma, V., Badkar, A., Zeng, D., *et al.*, Protein Structural Conformation and Not Second Virial Coefficient Relates to Long-Term Irreversible Aggregation of a Monoclonal Antibody and Ovalbumin in Solution. *Pharmaceutical Research* 2006, *23*, 1382-1394.

[28] Chen, B., Bautista, R., Yu, K., Zapata, G., *et al.*, Influence of Histidine on the Stability and Physical Properties of a Fully Human Antibody in Aqueous and Solid Forms. *Pharmaceutical Research* 2003, *20*, 1952-1960.

[29] Falconer, R. J., Chan, C., Hughes, K., Munro, T. P., Stabilization of a monoclonal antibody during purification and formulation by addition of basic amino acid excipients. *Journal of Chemical Technology and Biotechnology* 2011, *86*, 942-948.

[30] Esue, O., Genentech, Inc. 2012.

[31] Daugherty, A. L., Mrsny, R. J., Formulation and delivery issues for monoclonal antibody therapeutics. *Advanced Drug Delivery Reviews* 2006, *58*, 686-706.

[32] Chang, B. S., Yeung, B., Physical Stability of Protein Pharmaceuticals, *Formulation and Process Development Strategies for Manufacturing Biopharmaceuticals*, John Wiley & Sons, Inc. 2010, pp. 69-104.

[33] Cao, X., Fesinmeyer, R. M., Pierini, C. J., Siska, C. C., *et al.*, Free Fatty Acid Particles in Protein Formulations, Part 1: Microspectroscopic Identification. *Journal of Pharmaceutical Sciences* 2015, *104*, 433-446.

[34] Zhang, J., Protein-Protein Interactions in Salt Solutions, in: Weibo Cai, H. H. (Ed.), *Protein-Protein Interactions - Computational and Experimental Tools* 2012.

[35] Kunz, W., Specific ion effects in colloidal and biological systems. *Current Opinion in Colloid & Interface Science* 2010, *15*, 34-39.

[36] Collins, K. D., Ions from the Hofmeister series and osmolytes: effects on proteins in solution and in the crystallization process. *Macromolecular Crystallization* 2004, *34*, 300-311.

[37] Rubin, J., Linden, L., Coco, W. M., Bommarius, A. S. and Behrens, S. H., Salt-induced aggregation of a monoclonal human immunoglobulin G1. *J. Pharm. Sci.* 2013, 377–386.

[38] Zhang, L., Tan, H., Fesinmeyer, R. M., Li, C., *et al.*, Antibody solubility behavior in monovalent salt solutions reveals specific anion effects at low ionic strength. *Journal of pharmaceutical sciences* 2012, *101*, 965-977.

[39] Shukla, D., Trout, B. L., Understanding the Synergistic Effect of Arginine and Glutamic Acid Mixtures on Protein Solubility. *The Journal of Physical Chemistry B* 2011, *115*, 11831-11839.

[40] Assarsson, A., Linse, S., Cabaleiro-Lago, C., Effects of Polyamino Acids and Polyelectrolytes on Amyloid β Fibril Formation. *Langmuir* 2014, *30*, 8812-8818.

[41] Boeris, V., Spelzini, D., Salgado, J. P., Picó, G., *et al.*, Chymotrypsin–poly vinyl sulfonate interaction studied by dynamic light scattering and turbidimetric approaches. *Biochimica et Biophysica Acta (BBA) - General Subjects* 2008, *1780*, 1032-1037.

[42] Goers, J., Uversky, V. N., Fink, A. L., Polycation-induced oligomerization and accelerated fibrillation of human α-synuclein in vitro. *Protein Science* 2003, *12*, 702-707.

[43] Shalova, I. N., Asryants, R. A., Sholukh, M. V., Saso, L., *et al.*, Interaction of Polyanions with Basic Proteins, 2. *Macromolecular Bioscience* 2005, *5*, 1184-1192.

[44] Seyrek, E., Dubin, P. L., Tribet, C., Gamble, E. A., Ionic Strength Dependence of Protein-Polyelectrolyte Interactions. *Biomacromolecules* 2003, *4*, 273-282.

[45] Giger, K., Vanam, R. P., Seyrek, E., and Dubin, P. L., Suppression of Insulin Aggregation by Heparin. *Biomacromolecules* 2008, *9*, 2338-2344.

[46] Chung, K., Kim, J., Cho, B.-K., Ko, B.-J., *et al.*, How does dextran sulfate prevent heat induced aggregation of protein?: The mechanism and its limitation as aggregation inhibitor. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 2007, *1774*, 249-257.

[47] Sahin, E., Grillo, A. O., Perkins, M. D., Roberts, C. J., Comparative effects of pH and ionic strength on protein–protein interactions, unfolding, and aggregation for IgG1 antibodies. *Journal of Pharmaceutical Sciences* 2010, *99*, 4830-4848.

[48] Martin S Neergaard, D. S. K., Henrik Parshad, Anders D Nielsen, Eva H Møller, Marco van de Weert, Viscosity of high concentration protein formulations of monoclonal antibodies of the IgG1 and IgG4 subclass – Prediction of viscosity

through protein–protein interaction measurements. *Eur J Pharm Sci* 2013, *49*, 400-410.

[49] 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. *Biochimica et Biophysica Acta (BBA) - Proteins & Proteomics* 2008, *1784*, 600-610.

[50] Mazzaferro, L., Breccia, J. D., Andersson, M. M., Hitzmann, B., Hatti-Kaul, R., Polyethyleneimine–protein interactions and implications on protein stability. *International Journal of Biological Macromolecules* 2010, *47*, 15-20.

[51] Sedlák, E., Antalík, M., Coulombic and noncoulombic effect of polyanions on cytochrome c structure. *Biopolymers* 1998, *46*, 145-154.

[52] Park, J. M., Muhoberac, B. B., Dubin, P. L., Xia, J., Effects of protein charge heterogeneity in protein-polyelectrolyte complexation. *Macromolecules* 1992, *25*, 290-295.

[53] Xia, J., Dubin, P. L., Dautzenberg, H., Light scattering, electrophoresis, and turbidimetry studies of bovine serum albumin-poly(dimethyldiallylammonium chloride) complex. *Langmuir* 1993, *9*, 2015-2019.

[54] Brummitt, R. K., Nesta, D. P., Chang, L., Kroetsch, A. M., Roberts, C. J., Nonnative aggregation of an IgG1 antibody in acidic conditions, part 2: Nucleation and growth kinetics with competing growth mechanisms. *Journal of Pharmaceutical Sciences* 2011, *100*, 2104-2119.

[55] Souillac, P. O., Biophysical characterization of insoluble aggregates of a multi-domain protein: An insight into the role of the various domains. *Journal of Pharmaceutical Sciences* 2005, *94*, 2069-2083.

[56] Mehta, S. B., Bee, J. S., Randolph, T. W., Carpenter, J. F., Partial Unfolding of a Monoclonal Antibody: Role of a Single Domain in Driving Protein Aggregation. *Biochemistry* 2014, *53*, 3367-3377.

[57] Zhang-van Enk, J., Mason, B. D., Yu, L., Zhang, L., *et al.*, Perturbation of Thermal Unfolding and Aggregation of Human IgG1 Fc Fragment by Hofmeister Anions. *Molecular Pharmaceutics* 2013, *10*, 619-630.

[58] Vermeer, A. W. P., Norde, W., The Thermal Stability of Immunoglobulin: Unfolding and Aggregation of a Multi-Domain Protein. *Biophysical Journal* 2000, *78*, 394-404.

[59] Brummitt, R. K., Nesta, D. P. and Roberts, C. J. , Predicting Accelerated Aggregation Rates for Monoclonal Antibody Formulations, and Challenges for Low-Temperature Predictions. *J. Pharm. Sci.* 2011, *100*, 4234-4243.

[60] Ivinova, O. N., Izumrudov, V. A., Muronetz, V. I., Galaev, I. Y., Mattiasson, B., Influence of Complexing Polyanions on the Thermostability of Basic Proteins. *Macromolecular Bioscience* 2003, *3*, 210-215.

[61] Szenczi, Á., Kardos, J., Medgyesi, G. A., Závodszky, P., The effect of solvent environment on the conformation and stability of human polyclonal IgG in solution. *Biologicals* 2006, *34*, 5-14.

[62] Lankalapalli, S., Kolapalli, V. R. M., Polyelectrolyte Complexes: A Review of their Applicability in Drug Delivery Technology. *Indian Journal of Pharmaceutical Sciences* 2009, *71*, 481-487.

[63] Cooper, C. L., Dubin *, P. L., Kayitmazer, A. B., Turksen, S., Polyelectrolyte–protein complexes. *Current Opinion in Colloid & Interface Science* 2005, *10* 52 – 78.

[64] Sacco, D., Dellacherie, E., Interaction of a macromolecular polyanion, dextran sulfate, with human hemoglobin. *FEBS Letters* 1986, *199*, 254-258.

[65] Andersson, M. M., Hatti-Kaul, R., Brown, W., Dynamic and Static Light Scattering and Fluorescence Studies of the Interactions between Lactate Dehydrogenase and Poly(ethyleneimine). *The Journal of Physical Chemistry B* 2000, *104*, 3660-3667.

CHAPTER 7

7 CONCLUDING REMARKS AND SUGGESTION FOR FUTURE WORK

Previous researches have shown that in order to inhibit or control protein aggregation, one would have to manipulate or screen a wide range of solvent conditions to alter the solution thermodynamics. Understanding protein behaviour by studying protein-protein interactions requires knowledge on how the solution properties are influenced by solution variables such as ionic strength, salt or excipient concentration and type, and pH. This knowledge would help towards understanding how protein aggregation can be prevented or inhibited, which is critical for the rational design of stable protein formulations.

In chapter 4, we investigated protein-protein interactions and protein-ion interactions using lysozyme as a model system to provide insight into specific ion effects and also discuss the role of electrostatic interactions on protein-protein interactions. It was found that ion binding to positively charged proteins has a significant effect on protein-protein interactions. At low ionic strength $\leq$ 50 mM, the primary effect of anion binding is to alter the protein net charge and long-ranged double-layer forces through electrostatic interactions, in which case, sulphate due to the divalent charge is more effective than nitrate or thiocyanate at reducing protein-protein repulsion. This behaviour is consistent with the electroselectivity theory. At intermediate ionic strength, thiocyanate and nitrate ions are more effective than sulphate at inducing protein-protein attraction as lysozyme precipitation follows the reverse Hofmeister series. This implies that chaotropic anions induce short range attractive interactions between proteins of non-electrostatic origin. Therefore, the salting out effect of chaotropic ions or specific ion binding cannot be only rationalized in terms of electrostatic interactions and changes to the double layer potential only. By

combining the measurements of protein-protein interactions with zeta-potential studies under the same solution conditions, our study has provided the first estimate of how chaotropic anions alter non-electrostatic interactions between proteins. This knowledge will be useful for theoreticians and molecular simulation to identify the molecular determinants at the origin of the reverse Hofmeister series effect on protein solubility. For instance, it would be of great interest to rationalise the results using theories that extend the conventional double layer theory to include ion dispersion forces. In addition, performing detailed studies on a different set of proteins above and below the pI at low salt concentrations will shed more light on the propensity of ion specific interactions.

One possible experiment is to check whether or not bovine serum albumin (BSA) follows the reverse Hofmeister series around the isoelectric point (pI) where it has a net neutral charge. The pI of BSA is about 4.7 and is positively charge at low pH (i.e. below 4.7). Therefore one can determine whether or not it is the net charge that relates to chaotropic anion effects, since lysozyme has a large positive charge at pH 4.5. If reverse Hofmeister is observed for BSA then this would imply that ion-protein interactions are short-ranged and insensitive to the net charge of the protein. Static light scattering and zeta potential measurements can be used to probe protein-protein and protein-ion interactions for BSA at various pH values, ionic strengths and salt types. Results from these techniques can be reported in terms of the osmotic second virial coefficient (B_{22}) and zeta potential (ζ- potential).

At relatively low ionic strength, the effect of pH determines the extent of coulombic attraction or repulsion between net charges on the protein surface. This effect can be offset in the presence of anions when the anions bind and neutralise the positive charges on the protein surface. To compare the trends in magnitude between B_{22}

values for each anion type, the ionic strength is systematically varied over a wide range of 10 mM to 2 M. In this proposed experiment, anions such as sulphate, nitrate, thiocyanate, and chloride which span the Hofmeister series will be investigated at both low and high ionic strengths. This is to ascertain whether the reverse Hofmeister series which has been observed for lysozyme solutions follows similar behaviour or correlates with results obtained in BSA solutions depending on pH and ionic strength. Similarly, zeta potential measurements of BSA will be used to determine the contribution of electrostatic interactions when anions bind or accumulate at the protein surface at different ionic strengths. Results obtained from lysozyme measurements performed in this thesis show that sulphate ions do not follow Collins theory of matching water affinities hence, the proposed experiments can show whether or not the effect of anions on BSA molecules can solely be explained in terms of electrostatic /nonelectrostatic interactions or is protein specific. It will also be interesting to see whether the cross-over effect observed for lysozyme molecules between the different anion types follows similar pattern in BSA studies.

In chapter 5, the study aims to understand how dipeptides control the association state of insulin. Previous studies have shown that additives such as amino acids have the ability to suppress protein aggregation however; often the concentrations required for stabilizing formulations are too high for use in formulation. Thus, there is a need for novel additives that stabilize at much lower concentration. Here, we have provided the first investigation into the effects of dipeptides on reducing aggregation for therapeutically relevant proteins. We assess the aggregation behaviour of monomeric or zinc-free insulin under conditions such that its net charge varies from negative to neutral to positive, in the absence and presence of sodium chloride to screen the electrostatic interactions. Turbidimetric titrations showed the

effectiveness of the additives in suppressing insulin aggregation. At the insulin isoelectric pH 5.5, aggregation is greatest at low ionic strength due to strong attractive electrostatic interactions. Increasing the ionic strength of the sample reduces the aggregation due to screening the electrostatic interactions. Light scattering measurements showed that the dipeptide diArg is the most effective at suppressing aggregation in solutions at pH 3.7 and pH 5.5. We hypothesize that the diArg is effective due to its ability to bind to insulin surface and alter the electrostatic properties and interactions between insulin molecules. Most likely the increased binding is due to greater synergistic effects derived when two free arginine molecules are covalently linked together. Intrinsic fluorescence measurements reveal that the dipeptide Arg-Phe reduces thermally-induced aggregation of insulin at pH 7.5 while diArg promotes aggregation of heat denatured states. In this case, the dipeptides modulate hydrophobic and electrostatic interactions between insulin molecules to either inhibit or promote insulin aggregation depending on the rate controlling step to aggregation.

It would be of great interest to extend this experimental study to other protein therapeutics such as antibodies to study whether the different effects of the dipeptides on insulin aggregation are also observed. Here, the effect of promising dipeptides such as diArg, Arg-Phe, Phe-Arg and mixtures of arginine and glutamate can be investigated on a monoclonal antibody (mAb) such as IgG1 in order to develop a strategy for improving mAb formulations. Light scattering and fluorescence techniques can be used to study the effect of the dipeptides on the colloidal and conformational stability of the protein. B_{22} experiments will be conducted as a function of pH and ionic strength to discriminate between the effects of dipeptides on either the electrostatic or hydrophobic interactions between proteins

and the link to colloidal stability. In addition, aggregation propensity and conformational studies will be assessed in terms of aggregation temperatures (T_{agg}) and melting or unfolding temperatures (T_m) using the Optim as a function of temperature, pH, ionic strength and in the absence and presence of the dipeptides. The pH and ionic strength dependence of B_{22} can then be qualitatively correlated with aggregation propensity of the mAb to check that any increase in aggregation can be attributed to the effect of dipeptides on the colloidal stability. If positive B_{22} values correlate with lower aggregation propensity, this will provide an indication the dipeptides stabilize proteins by inducing repulsive protein-protein interactions while a correlation between negative B_{22} and increased aggregation propensity would indicate dipeptides induced attractive protein-protein interactions.

Another extension of this study could be to systematically investigate whether the combination of equimolar mixtures of two dipeptides work better relative to one dipeptide mixture at suppressing insulin aggregation.

In chapter 6, the aggregation behaviour and the stability characteristics of a monoclonal antibody (IgG4) under a range of operating conditions in the presence of a novel cationic excipient was investigated. Results reveal that increasing pH from 5 to 7 increased protein-protein attraction while thermal stability was improved. Increasing the ionic strength by adding of salt indicated attractive electrostatic interactions promote the strong self-association of the IgG4. PEI binding to the IgG4 prevents charged patches on the protein surface from self-association. PEI binding occurs in solutions containing NaCl even when electrostatic interactions are screened implying that binding is driven by hydrophobic interactions. Strong polyion binding to proteins leads to protein conformational destabilization as is also observed here with PEI. Nevertheless, the aggregation onset temperature is greater than the protein

melting temperature indicating the PEI complexation protects the protein against aggregation. The next step is to investigate the effectiveness of the polycation at preventing aggregation at room temperature. This requires understanding how the colloidal and conformational stability change with decreasing temperature.

Another question to address is the correlation between the binding affinity of the polyion and its stabilizing effectiveness against aggregation. In the proposed study, the fluorescence spectra at a fixed protein concentration is recorded at pH 5, 6 and 7 under the same solution conditions at an excitation wavelength of 266 nm. Since the protein contains fluorescent aromatic groups, the effect of the polyion binding can be probed from examining changes to the fluorescence emission spectrum (peak intensity or peak maximum), which would reflect changes in the solvent environment of the buried amino acids due to polycation binding.. Solution conditions in which the polyion exhibits the largest shift will imply that the effectiveness of PEI in stabilizing the protein is as a result of increased binding affinity of PEI to the protein.