

A Critical Evaluation of Screening versus Sample-by-Sample Analysis of Protein Aggregation using a Variety of Chromatographic and Spectroscopic Techniques



A thesis submitted to the National University of Ireland in fulfilment of the requirements for the degree of

Doctor of Philosophy

By

Ruth McNamara, M.Sc.

Department of Chemistry
Maynooth University

October 2014

Research Supervisor: Dr. Jennifer McManus

Head of Department: Dr. John Stephens

*Dedicated to my parents, Maureen and Matthew.
Without their encouragement none of my success would be
possible.*

Declaration

I hereby certify that this thesis has not been submitted before, in whole or in part, to this or any university for any degree and is, except where otherwise stated, the original work of the author.

Signed: _____

Date: _____

Maynooth University

Acknowledgements

First and foremost, I would like to sincerely thank my supervisor and thesis advisor, Dr. Jennifer McManus, for her help, support and friendship. I have learned so much throughout my four years under her guidance. Without her advice and patience this PhD thesis and course of study would never have amounted to what it is today.

I would also like to thank all those who have helped me accomplish this research. Both Prof. John Lowry and Dr. John Stephens for allowing me the opportunity to carry out my research in the Department of Chemistry. A special thank you to John and Pat Hume Scholarship and to Science Foundation Ireland for supporting this work financially. Thank you to the staff at NIBRT, especially to Prof. Pauline Rudd for allowing me to spend some time at their research facility. Particularly to Tharmala Tharmalingam for her help and guidance during my time there. I would also like to thank Prof. Ian Marison and his lab group, including Dr. Pirkko Muhonen and Karen Byrne for the IgG1 cell culture medium. Thank you to Mr. Brenton Cavanagh from the RSCI for his assistance with TEM images. Finally to all the fourth year chemistry students who in their time within our lab group, provided help to my research; Ciara Glynn, Neil Conlon and Ciara Higgins.

My sincere thanks to all other members of the Chemistry department for their support and assistance. To the administration staff; Niamh, Carol and Donna for their help whenever it was called upon. To the technical staff and lab co-ordinators for their advice, especially Barbra who was always willing to help with HPLC troubleshooting. To Noel Williams – I'm glad I'm gone before you, you are a fundamental part of the department and I don't know how it'll survive after you! Thank you for always being there to answer any questions and providing help when needed.

To all members of our research group present, (can't say past as I'm the first), it has been a privilege working with all of you. Susan James, my first lab buddy! Thank you for the head rubs! Even when my hair resembled a (ginger) bush, you were still willing. Wherever you end up, keep in touch - I'll send you on some kit-kats. Urszula-Ula-Migas, if I ever need something pickled, black rubber wellies or the shoppings I know where to go! Thanks Gods for always to be there for breaks when I needed them most, even if it was to keep you company out-front. Remember the 20th of December, that's my name day!!! Alice B-L-U-M (for Mary)-L-E-I-N (for Nancy), the lab never was the same after

you came, nor is it ever the same when you're there. I shall miss our fishing trips! Thanks for all the dinners and giggles- now stay out of the rain or you'll start to rust. Michelle K. Quinn my sista from another country (yes Mayo is THAT far away!!) your friendship since your arrival has been most pleasant and rewarding- mostly at Alices expense ☺ the numerous times we scared Alice are without doubt my funniest memory. Don't visit Stephen without me (we have to get writing that book). Don't be too sad, we'll always have Denver. I hope that Jack can fill the very, very, very, very large void I've left in your life. Mark Grace, our very first lab boy to play with. If you are ever looking to cuddle up on the couch and watch a Disney movie –DON'T, go drink a beer like a man!! I hope the girls go easy on you in my absence but if not, do as Queen Elsa does and just "Let it go, let it go!" I will miss you all. I know that I may have left the lab but I'm sure I have not left any of your lives – you won't get rid of me that quick!

To the rest of the chemistry department, past and present, thank you all for the good times and the bad (generally 1.45pm for me). To the Laser Lads, thank you for providing so much entertainment. Barry and his partner in crime, Christopher (Chiggy), a constant source of tea time chats, although I'm sure they were to your advantage not mine, for the GREEN tea – my favourite, the Halloween jokes and all the puns- I don't have thyme to go into them all now but they were greatly appreciated! To the 1 o'clock lunch time crew - those gone (Lynn, Joey, John, Colin and Conor) and still there - I've learned a lifetime of information some of which I hope to quickly forget, most of which I will never use, some of which I'll not pass down for fear of judgement and even more I still don't understand and am afraid to google. Justine for being a constant supply of meat and insurance claim forms, Ross for always holding the door open (fridge or otherwise), Andrew always lending a hand and Jack- look after Michelle for me. Thank you all for your friendship, laughs, nights out, nights in, Halloween, Summer, Christmas, and it's a (insert day)day parties.

To my writing buddies for the past few months Thank you for the help and support to get through this writing process – WE MADE IT!!! Especially Ursula Sheridan for putting up with my constant interruptions and questions and distractions and nuisances. It's not like a Cavan woman to share anything so your guidance towards putting this thesis together is very much appreciated. We started in this process together four years ago so it's nice that we can finish up together too!

To my home-birds- Aoife, Eimear, Jen (+1), Niamh, Sinead, Sharon and Tara, you're many years of friendship have gotten me through so much, especially lately when I needed a rant! I cannot wait to celebrate with a glass of wine (and not just to get drunk).

To the constant source of happiness in my life – Tommy. I came to Maynooth to pursue my love of research but little did I know my love for you would also pursue. The journey of my last four years led to you and it's a path I'm glad I had you to travel with. I cannot put into words how much you mean to me (the tears in my eyes say a million words). You have given me so much support and love that I will be forever grateful. You are always there for me when I need a cuddle or provide a home to put a randomly placed finger. Thank you for having me in your life (and currently your house too)! You “always” listened to my rants, and knew the words to calm me down (generally something to do with food). We pushed, supported and inspired each other during the last few months to both achieve our goals, for that I'm forever grateful- but for now its holiday time.

Finally, I would like to thank my family; Alan, Michelle and Peter, and my parents, Maureen and Matthew. Without your support and love I could not have got through these years (oh the car helped too). Thank you all for providing me with continuous motivation and encouragement during this important time in my life. Now you can have me back 😊

Abstract

While protein aggregation has been widely studied, it is still poorly understood. Protein aggregation is a complex process, resulting in heterogeneous populations of particles from a variety of mechanisms, which provides both intellectual and technical challenges. This thesis probes the mechanisms leading to protein aggregation by characterising the aggregation of two model proteins, lysozyme and bovine serum albumin, using a variety of chemical and biophysical techniques.

High-throughput assays are commonly utilised for screening protein formulations. Two high-throughput assays were critically examined to assess the effectiveness and robustness of this approach for detecting protein aggregation after forced degradation (or accelerated aggregation testing). While these high-throughput assays do play a role in formulation and screening, careful consideration must be given to the interpretation of the analysis and the results must be viewed in context and as part of a wider characterization process.

Chemical modification of protein structure via oxidation was also examined as a mechanism leading to aggregation. Oxidation of lysozyme and bovine serum albumin was induced with hydrogen peroxide. The secondary structure of both proteins were vulnerable to oxidative attack, resulting in a loss of native structure and the formation of both protein aggregates and protein fragments, the extent of which depended on both the specific protein chemistry and solution conditions.

The mechanism of chemical unfolding was probed for both lysozyme and BSA with the addition of two denaturants; urea and guanidine hydrochloride. The mechanism by which each denaturant unfolds both proteins has been thoroughly studied in the past. Lysozyme is known to follow a two-state model, while BSA conforms to a two-step unfolding process via three transition states. We have demonstrated the sensitivity and reliability of second derivative UV analysis to characterise the unfolding states of both proteins.

Finally, a monoclonal antibody, IgG1, was purified and characterised. Initial studies were carried out to assess the stability of this protein. Glyco-analysis was carried out to determine the glycan profile of the IgG1 protein and to ensure for consistency in product glycosylation. The stability of the glycoform was also examined in terms of length of storage under, different storage conditions.

List of Abbreviations

| | |
|---|--|
| % <i>T</i> – percentage transmission | NP-HPLC – normal phase HPLC |
| 2-AB – 2-aminobenzamide | NTA – nanoparticle tracking analysis |
| ABS – <i>Arthrobacter ureafaciens</i> sialidase | O ₂ • ⁻ – superoxide anion radical |
| AF4 –asymmetric-Flow Field Flow Fractionation | OS – oxidative stress |
| AFM – atomic force microscopy | PAGE – polyacrylamide gel electrophoresis |
| APS – ammonium persulphate | Phen – phenylalanine |
| Asn – Asparagine | <i>pI</i> – isoelectric point |
| AUC – analytical ultracentrifugation | RON – reactive nitrogen species |
| <i>B</i> ₂₂ – osmotic second virial coefficient | ROS – reactive oxygen species |
| BFK – Bovine kidney alpha-fucosidase | S/N – signal to noise ratio |
| BSA – Bovine serum albumin | SANS – small-angle neutron scattering |
| BTG – Bovine testes β-galactosidase | SAXS – small-angle X-ray scattering |
| Ca ²⁺ – calcium ion | SDS – sodium dodecyl sulphate |
| CCD – charged coupled device | SE-HPLC – size exclusion HPLC |
| CHO – Chinese hamster ovaries | SOD – superoxide-dismutase |
| Cys - Cysteine | TEM – transmission electron microscopy |
| DAD – diode array detector | TEMED – Tetramethylethylenediamine |
| DLS – dynamic light scattering | ThT – Thioflavin T |
| DSC – differential scanning calorimetry | <i>T_m</i> – melt transmission temperature |
| DTT – dithiothreitol | <i>t_R</i> – retention time |
| <i>g</i> – gravitational force | Tris-HCl – |
| GdmHCl – guanidine hydrochloride | tris(hydroxymethyl)aminomethane hydrochloride |
| GlcNAc – N-acetylglucosamine | Trp – Tryptophan |
| GU – glucose unit | Tyr – Tyrosine |
| GUH B – N-acetylglucosaminidase | U – urea |
| <i>h</i> – Planck’s constant | UPLC – ultra performance liquid chromatography |
| H ₂ C ₂ O ₄ – oxalic acid | UV – ultraviolet |
| H ₂ O ₂ – hydrogen peroxide | UV/Vis – ultraviolet/visible |
| H ₂ SO ₄ – sulphuric acid | v/v – volume per volume |
| HAS – Human serum albumin | w/v – weight/volume |
| HCl – hydrochloric acid | Δ <i>G</i> [‡] – activation energy |
| HEWL – Hen egg white lysozyme | ε – extinction co-efficient |
| HILIC - hydrophilic interaction liquid chromatography | λ – lambda |
| HO• – hydroxyl radical | λ _{max} – lambda maximum |
| HPLC – high performance liquid chromatography | λ _{min} – lambda minimum |
| HTS – high-throughput screening | |
| IgG – immunoglobulin G | |
| JBM – Jack Bean α(1-2,3,6)-Mannosidase | |
| <i>k_D</i> – dissociation constant | |
| KMnO ₄ – potassium permanganate | |
| mAbs – monoclonal antibodies | |
| MCO – metal catalysed oxidation | |
| MTX – methotrexate hydrate | |
| <i>M_w</i> – molecular weight | |
| MWCO – molecular weight cut off | |
| MWM – molecular weight marker | |
| NaH ₂ PO ₄ – monobasic sodium phosphate | |
| NaHCO ₃ – sodium bicarbonate | |
| NH ₄ HCO ₂ – ammonium formate | |

Table of Contents

| | |
|--|----------|
| Declaration | i |
| Acknowledgements | ii |
| Abstract..... | v |
| List of Abbreviations | vi |
| Chapter 1 Introduction | 7 |
| 1.1 Protein aggregation..... | 8 |
| 1.1.1 The importance of understanding protein aggregation | 8 |
| 1.1.2 Aggregation models and pathways | 9 |
| 1.1.2.1 Mechanism 1 – Reversible association of the native monomer | 9 |
| 1.1.2.2 Mechanism 2 – Aggregation of conformationally - altered monomer .. | 11 |
| 1.1.2.3 Mechanism 3 – Aggregates by chemically-modified product..... | 13 |
| 1.1.2.4 Mechanism 4 – Nucleation-controlled aggregation | 14 |
| 1.1.2.5 Mechanism 5 – Surface induced | 15 |
| 1.2 Solution conditions..... | 16 |
| 1.2.1 Solution pH | 16 |
| 1.3 Characterising protein aggregates | 17 |
| 1.3.1.1 Fibrillar aggregates | 18 |
| 1.3.1.2 Amorphous aggregates | 19 |
| 1.4 Methods to measure protein aggregation | 20 |
| 1.5 Unfolding pathways | 23 |
| 1.6 Forced degradation and aggregation | 25 |
| 1.6.1.1 Thermal stress – Mechanism 2 | 25 |
| 1.6.1.2 Chemical denaturation induced unfolding – Mechanism 2..... | 27 |
| 1.6.1.3 Chemical alteration – Mechanism 3 | 28 |
| 1.6.1.4 Oxidation | 28 |
| 1.7 Model proteins..... | 31 |
| 1.7.1 Lysozyme | 31 |
| 1.7.2 Bovine serum albumin | 33 |
| 1.8 Therapeutic Proteins | 35 |
| 1.8.1 Production of therapeutic proteins – monoclonal antibodies..... | 35 |
| 1.8.2 Antibodies | 36 |
| 1.8.3 Immunoglobulin G | 36 |
| 1.8.4 Glycosylation | 37 |
| 1.8.4.1 Glycosylation of immunoglobulins | 39 |
| 1.9 Thesis motivation | 41 |

| | | |
|------------------|---|-----------|
| Chapter 2 | Experimental techniques and data analysis | 42 |
| 2.1 | Preparation of Reagents and Materials | 43 |
| 2.1.1 | Preparation of buffers | 43 |
| 2.1.1.1 | Sodium phosphate | 43 |
| 2.1.1.2 | Tris-HCl | 43 |
| 2.1.1.3 | 10% SDS | 44 |
| 2.1.2 | Preparation of reagents | 44 |
| 2.1.2.1 | Preparation of denaturants | 44 |
| 2.1.2.2 | Standardizing hydrogen peroxide | 44 |
| 2.1.3 | Preparation of protein stock solutions | 45 |
| 2.2 | Mechanisms of aggregation | 45 |
| 2.2.1 | Thermal denaturation | 45 |
| 2.2.1.1 | Temperature induced unfolding | 45 |
| 2.2.2 | Oxidation | 46 |
| 2.2.2.1 | Oxidant induced de-stabilisation of protein | 46 |
| 2.2.3 | Chemical unfolding | 46 |
| 2.3 | SE-High Performance Liquid Chromatography | 46 |
| 2.3.1 | Description of the technique | 46 |
| 2.3.2 | Instrumentation | 47 |
| 2.3.3 | Running a sample | 48 |
| 2.3.4 | Calibration of the column | 48 |
| 2.3.5 | Determination of protein concentration by SE-HPLC | 49 |
| 2.4 | Electrophoresis | 50 |
| 2.4.1 | Description of the technique | 50 |
| 2.4.2 | Types of PAGE electrophoresis | 51 |
| 2.4.2.1 | Native polyacrylamide gels | 51 |
| 2.4.2.2 | SDS polyacrylamide gels | 52 |
| 2.4.3 | Protocol | 52 |
| 2.4.3.1 | PAGE | 52 |
| 2.4.4 | Modifications to the PAGE protocol | 54 |
| 2.4.4.1 | Native PAGE | 54 |
| 2.4.4.2 | Reducing SDS-PAGE | 54 |
| 2.4.4.3 | Gel concentration | 54 |
| 2.5 | Spectroscopy | 55 |
| 2.5.1 | Absorption spectroscopy | 55 |
| 2.5.1.1 | UV-Vis measurement | 58 |
| 2.5.2 | Fluorescence Spectroscopy | 58 |

| | | |
|---------------------------|---|-----------|
| 2.5.2.1 | Fluorescence Measurements..... | 60 |
| 2.5.3 | Derivative spectroscopy..... | 60 |
| 2.5.3.1 | Second derivative absorbance spectroscopy measurement..... | 62 |
| 2.5.3.2 | Second derivative fluorescence spectroscopy..... | 62 |
| 2.6 | High throughput assays..... | 63 |
| 2.6.1 | Turbidity assay..... | 63 |
| 2.6.1.1 | Protocol..... | 64 |
| 2.6.2 | Thioflavin T assay..... | 64 |
| 2.6.2.1 | Protocol..... | 65 |
| 2.7 | Phase separation..... | 66 |
| 2.7.1 | Removal of large (precipitated) aggregates..... | 66 |
| 2.7.2 | Characterisation of supernatant..... | 66 |
| 2.8 | Microscopy..... | 66 |
| 2.8.1 | Light microscopy imaging protocol..... | 67 |
| 2.8.2 | Polarised microscopy imaging protocol..... | 67 |
| 2.8.3 | Transmission electron microscopy..... | 68 |
| 2.8.3.1 | TEM Protocol..... | 68 |
| 2.9 | Isolation of monoclonal antibody IgG1..... | 69 |
| 2.9.1 | Expression and purification of IgG1 from CHO DP-12..... | 69 |
| 2.9.2 | Identifying the glycoform of IgG1..... | 70 |
| 2.9.2.1 | Extraction of glycan from IgG1..... | 70 |
| 2.9.2.2 | Labelling of glycan..... | 71 |
| 2.9.2.3 | Label clean up – paper chromatography..... | 71 |
| 2.9.2.4 | Separation of glycoforms..... | 72 |
| <i>Undigested Samples</i> | | 72 |
| <i>Digested Samples</i> | | 72 |
| 2.9.2.5 | Peak assignment..... | 74 |
| Chapter 3 | High-Throughput Approach to Measure Protein Aggregation..... | 75 |
| 3.1 | Introduction..... | 76 |
| 3.1.1 | Background..... | 76 |
| 3.1.2 | The problem..... | 76 |
| 3.1.3 | High-throughput as a solution..... | 77 |
| 3.1.4 | Advancements in high-throughput technologies..... | 77 |
| 3.1.5 | HTS methodology..... | 78 |
| 3.1.6 | Scepticism of HTS..... | 79 |
| 3.1.7 | The end goal..... | 80 |
| 3.2 | Aims of the study..... | 81 |

| | | |
|------------------|---|------------|
| 3.3 | Results | 83 |
| 3.3.1 | High-throughput assay - Turbidity..... | 83 |
| 3.3.1.1 | Lysozyme | 83 |
| 3.3.1.2 | BSA | 84 |
| 3.3.2 | Solution turbidity over time | 85 |
| 3.3.3 | Analysis of the supernatant | 87 |
| 3.3.4 | Further Screening | 89 |
| 3.3.5 | Aggregate morphology..... | 93 |
| 3.3.5.1 | Appearance of the protein aggregates | 93 |
| 3.3.5.2 | Polarised light microscopy | 94 |
| 3.3.6 | High-throughput assay – Thioflavin T..... | 95 |
| 3.3.7 | SE-HPLC | 96 |
| 3.3.7.1 | SE-HPLC of lysozyme | 97 |
| 3.3.7.2 | SE-HPLC of BSA..... | 100 |
| 3.4 | Discussion | 104 |
| 3.5 | Conclusions | 108 |
| Chapter 4 | Oxidation as a Pathway to Protein Aggregation | 109 |
| 4.1 | Protein Oxidation | 110 |
| 4.2 | Aims of the study | 115 |
| 4.3 | Results | 116 |
| 4.3.1 | Oxidation of Lysozyme..... | 116 |
| 4.3.1.1 | Oxidative modification of lysozyme secondary structure | 116 |
| 4.3.1.2 | Macroscopic changes following oxidation..... | 118 |
| 4.3.1.3 | Lysozyme turbidity after oxidation | 118 |
| 4.3.1.4 | Characterisation of oxidised lysozyme precipitates | 120 |
| 4.3.1.5 | Analysis of small lysozyme aggregates..... | 123 |
| 4.3.1.6 | Fragmentation of oxidised lysozyme..... | 127 |
| 4.3.2 | Oxidation of Bovine serum albumin..... | 129 |
| 4.3.2.1 | Oxidative modification of BSA secondary structure..... | 129 |
| 4.3.2.2 | Analysis of BSA aggregation and fragmentation..... | 130 |
| 4.3.2.3 | Quantification of oxidation products..... | 135 |
| 4.4 | Discussion | 136 |
| 4.5 | Conclusions | 141 |
| Chapter 5 | Second Derivative Spectroscopy to Monitor Protein Unfolding..... | 142 |
| 5.1 | Introduction | 143 |
| 5.1.1 | Protein unfolding with chemical denaturants..... | 143 |
| 5.1.2 | Spectroscopy and protein unfolding..... | 143 |

| | | |
|--|---|------------|
| 5.1.3 | Derivative spectroscopy to monitor protein unfolding | 144 |
| 5.1.4 | Unfolding of BSA | 145 |
| 5.1.5 | Unfolding of lysozyme..... | 147 |
| 5.2 | Aims of the study | 148 |
| 5.3 | Results | 150 |
| 5.3.1 | Unfolding of BSA | 150 |
| 5.3.1.1 | Fluorescence spectroscopy of BSA unfolding | 150 |
| 5.3.1.2 | Second derivative fluorescence spectroscopy of BSA unfolding | 153 |
| 5.3.1.3 | Zero order UV absorbance of BSA unfolding..... | 156 |
| 5.3.1.4 | Second derivative UV absorbance during BSA unfolding..... | 157 |
| 5.3.2 | Unfolding of lysozyme..... | 161 |
| 5.3.2.1 | Fluorescence spectroscopy of lysozyme unfolding..... | 161 |
| 5.3.2.2 | Second derivative fluorescence spectroscopy of lysozyme unfolding | 163 |
| 5.3.2.3 | Zero order UV absorbance during lysozyme unfolding | 164 |
| 5.3.2.4 | Second derivative UV absorbance of lysozyme unfolding | 165 |
| 5.4 | Discussion | 169 |
| 5.4.1 | Unfolding of BSA | 169 |
| 5.4.2 | Unfolding of lysozyme..... | 172 |
| 5.5 | Conclusions | 175 |
| Chapter 6 Purification and Characterisation of IgG1 | | 177 |
| 6.1 | Introduction | 178 |
| 6.2 | Aims of the study | 181 |
| 6.3 | Results | 182 |
| 6.3.1 | Purification of monoclonal antibody..... | 182 |
| 6.3.2 | SE-HPLC | 183 |
| 6.3.3 | Gel electrophoresis..... | 184 |
| 6.3.3.1 | Identifying intact monomeric IgG1 | 184 |
| 6.3.3.2 | Identifying heavy and light chains of IgG1 | 186 |
| 6.3.4 | Batch to batch variation | 187 |
| 6.3.4.1 | Protein yield | 187 |
| 6.3.4.2 | Gel Electrophoresis | 188 |
| 6.3.4.3 | SE-HPLC..... | 189 |
| 6.3.5 | Stability of purified IgG1 | 190 |
| 6.3.6 | Identifying the glycoform of IgG1 | 193 |
| 6.3.6.1 | Undigested samples | 193 |
| 6.3.6.2 | Digested samples | 198 |
| 6.4 | Discussion | 203 |

| | | |
|-----|---|------------|
| 6.5 | Conclusions and future work..... | 205 |
| | Summary and Concluding Remarks | 206 |
| | References..... | 208 |

Chapter 1 Introduction

1.1 Protein aggregation

Proteins are synthesized on ribosomes as a linear chain composed of 20 naturally occurring amino acids. The linear arrangement of these amino acids through peptide bonds makes up the primary structure of a protein; the basis of the polypeptide chain. The use of the term “protein folding” refers to the process whereby the newly synthesised polypeptide chain adopts a unique three-dimensional shape; the native conformation of the protein. This is a spontaneous process that gives the protein the shape necessary for it to function. This is achieved without the need for any other information beyond what is contained within the amino acid sequence and its surroundings (Anfinsen 1973; Dobson, Sali *et al.* 1998).

Protein solutions are thermodynamically unstable and prone to the formation of protein aggregates, due to net-attraction between protein molecules. Environmental changes can lead to partial unfolding of the protein, and by different mechanisms to aggregation. Almost all proteins can aggregate in certain experimental conditions; however the degree to which they aggregate and the types of aggregates formed will vary depending on such conditions. The overall process of aggregation depends on several factors including temperature, solution pH, protein concentration and the presence of co-solutes. The process of aggregation results from different interconnected mechanisms (Fink 1998; Manno, Emanuele *et al.* 1999; San Biagio, Martorana *et al.* 1999; Bondos and Bicknell 2003; Wang, Nema *et al.* 2010). Changes in protein structure, formation of intermolecular bonds, protein self-association and nucleation processes all contribute to the formation of protein aggregates. Despite extensive research both within academia and medicine, and in the biopharmaceutical industry, the complex processes that govern protein aggregation are still poorly understood.

1.1.1 The importance of understanding protein aggregation

Protein aggregation is important in many different fields from biomedicine to biotechnology, and also many diseases associated with the deposition of amyloid aggregates (Dobson 2004; Ami, Natalello *et al.* 2013). Over 40 human diseases are commonly associated with protein aggregation, in particular neurodegenerative/ageing diseases such as Alzheimer’s disease, Parkinson’s disease, Huntington’s disease, and other human diseases such as type II diabetes (DeToma, Salamekh *et al.* 2012). Protein aggregation in biopharmaceutical protein products is of growing concern, related to the

potential impact of aggregated protein particles on product quality, efficacy and safety. The increase in safety concerns stems from the suggestion that protein particles have been the cause of adverse events and immunogenic responses in patients (Rosenberg 2006; Zolls, Tantipolphan *et al.* 2012), yet a full understanding of the risk still remains unclear. How exactly particle size, morphology or molecular structure affects immunogenicity has not yet been established (Ripple and Dimitrova 2012). Consequently, manufacturers of therapeutic proteins are implementing various control strategies and continued improvements to minimise particle formation (Ripple and Dimitrova 2012). A better understanding of the aggregation risks associated with the manufacturing process, transportation, storage and clinical administration is required. Recognising any potential for manufacturing processes to induce aggregation, along with improving the shelf-life of the product to ensure better stability of the formulations are some of the ways of managing and addressing the issues of protein aggregation (Cromwell, Hilario *et al.* 2006).

1.1.2 Aggregation models and pathways

The term “protein aggregation” has many definitions. “Protein aggregates” are here defined as any non-native protein species with a molecular weight higher than a protein monomer (*i.e.* dimers, oligomers, multimers...) that are formed by either physical or chemical changes and/or by covalent or non-covalent interactions (Mahler, Friess *et al.* 2009).

Protein aggregation can occur via numerous mechanisms, some that involve the interplay between different structural conformations of the protein (non-native), while others are not related to changes in protein structure (Fink 1998). The aggregation pathways of a protein are not exclusive, in that they can occur simultaneously within the same solution (Philo and Arakawa 2009). Protein aggregation depends strongly on the solution environment, as well as on the relative intrinsic thermodynamic stability of the protein native state. Different stress conditions can favour various aggregation pathways, leading to a number of mechanisms by which protein aggregation can occur leading to either the formation of amorphous aggregates or amyloid-type fibrils.

1.1.2.1 Mechanism 1 – Reversible association of the native monomer

The surface of protein monomers are self-complementary, so that they readily self-associate without the formation of an intermediate state, to form small oligomers which

are generally reversible. No structural change is required for mechanism 1. The number and variability of the surface patches has the potential to form multiple conformations and different patterns of oligomer growth. With time these oligomers can grow (dependent on protein concentration). Irreversible aggregates may occur through the formation of covalent bonds such as disulphide linkages (Philo and Arakawa 2009) within the non-covalent aggregate.

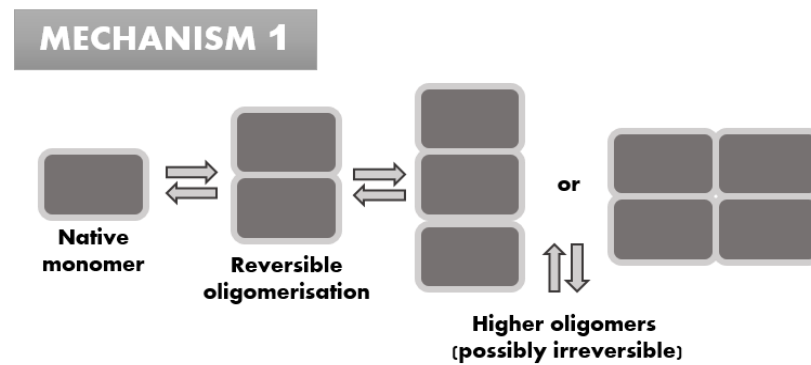
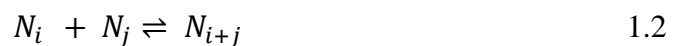


Figure 1.1: Schematic of protein aggregation by Mechanism 1 modified from (Philo and Arakawa 2009).

In 1970, Eisenberg suggested a reversible association mechanism for the aggregation of glutamate dehydrogenase (Reisler, Pouyet *et al.* 1970);



where N is the native monomeric state, and N_n , is the polymerised species of length n . This was disputed by Thusius who claimed Eisenberg’s reversible association mechanism alone is not able to account for all the aggregation data in varying concentration ranges (Thusius 1975). They termed their mechanism the “random association mechanism” where by any two native protein units (N) of any size (mono- or polymeric) can associate together where i and $j = 1, 2, \dots, \infty$;



Self-association of native protein is understood to be a result of non-covalent protein-protein interactions (Wang, Nema *et al.* 2010). The primary step in protein condensation diseases stems from a loss in protein solubility due to the net attraction between the proteins (Benedek 1997; Annunziata, Ogun *et al.* 2003). An understanding of the molecular interactions that govern the conformation of proteins in solution goes some way to understanding the drive for a protein to self-associate and in determining its colloidal stability. In some solutions, for protein in their native state the repulsive forces

(table 1.1) are not sufficiently strong to overcome the attractive forces, and self-association can occur through intermolecular interactions.

| <i>Type</i> | <i>Sign</i> | <i>Strength</i> | <i>Range</i> |
|-------------------------|-------------------------|----------------------------|---------------------------|
| <i>Hydrophobic</i> | Attractive | Strong | Long |
| <i>Electrostatic</i> | Repulsive or attractive | Weak → Strong ^a | Short → Long ^a |
| <i>Hydrogen bonding</i> | Attractive | Weak | Short |
| <i>Hydration</i> | Repulsive | Strong | Short |
| <i>Van der Waals</i> | Attractive | Weak | Short |
| <i>Steric repulsion</i> | Repulsive | Strong | Short |

Table 1.1: Types of molecular interactions between two similar proteins in aqueous solution. [^a depends on pH and ionic strength] (Bryant and McClements 1998).

A structural change is not required, for this mechanism of aggregation, as shown in the case of human eye lens proteins (crystallins), where aggregation and phase transitions are thought to be driven by attractive inter-protein interactions (McManus, Lomakin *et al.* 2007; Wang, Lomakin *et al.* 2010). Measuring the net interaction potential is a reliable way of predicting aggregation potential. George and Wilson (George and Wilson 1994) have shown that the thermodynamic parameter, the osmotic second virial coefficient (B_{22}) can be used to assess the tendency of a protein to self-associate. B_{22} represents a Boltzmann-weighted average measure of protein-protein interactions; a positive B_{22} value indicates protein-protein net repulsion and favours protein-solvent interactions (proteins are colloiddally stable), while a negative value indicates protein-protein net attraction. Traditional methods to measure B_{22} values include static light scattering, dynamic light scattering, membrane osmometry (Tessier and Lenhoff 2003). B_{22} has been successfully used to directly quantify protein-protein interactions (Saluja, Fesinmeyer *et al.* 2010; James and McManus 2012).

1.1.2.2 Mechanism 2 – Aggregation of conformationally - altered monomer

The key difference between mechanism 1 and 2 is that a conformational change or partial unfolding of the native state is required to form a non-native aggregate. Without this structural change/unfolding, the native monomer has a low propensity to aggregate through reversible interactions. Once it undergoes conformational changes or partial unfolding (forming an intermediate state) associations typical of mechanism 1 are formed (Philo and Arakawa 2009).

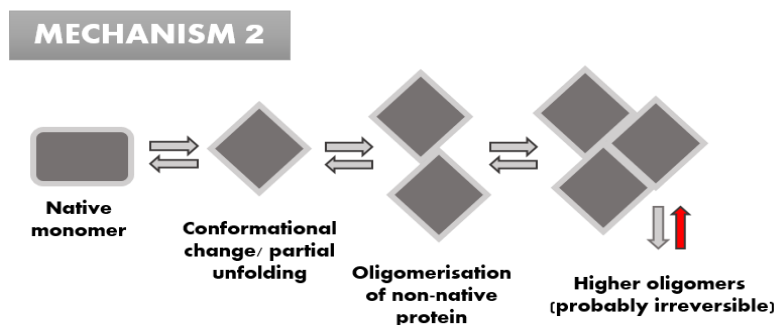


Figure 1.2: Schematic of protein aggregation by Mechanism 2 modified from (Philo and Arakawa 2009).

In order for a natively folded monomer to form structurally perturbed, higher order aggregates the native state needs to first undergo both a structural change and an assembly process which is generally irreversible once the aggregates get to a sufficient size (Chi, Krishnan *et al.* 2003). This model of protein aggregation was first presented by Lumry and Eyring in 1954 known as the Lumry-Eyring framework model (Lumry and Eyring 1954), proposing that the native protein (N) undergoes an intramolecular transformation into a transition state (TS^*) preceding the formation of an aggregation-competent intermediate (A_I), followed by the irreversible assembly into aggregates. A_m and A_{m+I} are aggregates containing m and $m + I$ protein molecules;



A representative reaction diagram for the irreversible association of non-native protein is shown in Figure 1.3. In this reaction the formation of A_m is favoured thermodynamically and therefore has the lowest free energy. Each reaction proceeds through energy barriers. The maximum energy state is the formation of the TS^* state and the free energy between this state and that of the native state (N) is the activation free energy (ΔG^\ddagger) (Chi, Krishnan *et al.* 2003).

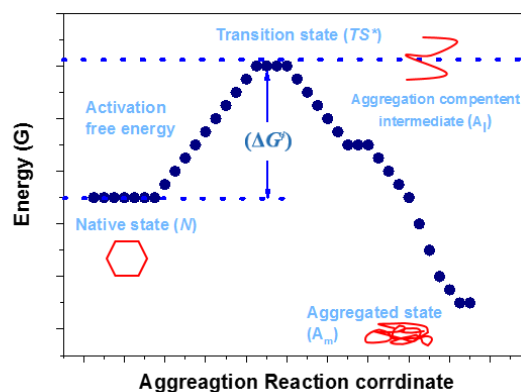


Figure 1.3: Schematic reaction coordinate diagram of a protein aggregation mechanism of a structurally altered protein (mechanism 2). Modified from (Chi, Krishnan *et al.* 2003).

Much of the research in this area suggests that this is probably the most dominant mechanism of protein aggregation and the most problematic because it is encountered so routinely in protein formulation (Wang, Singh *et al.* 2012; Elvin, Couston *et al.* 2013). Many conformational diseases, also known as misfolding diseases, result from proteins with a mutated gene sequence whereby the protein cannot adopt a native folded conformation, structural modification or by an incorrect folding process leading to self-association and subsequent aggregation *e.g.* prion, tau, alpha-synuclein and amyloid-beta peptide (Moreno-Gonzalez and Soto 2011). Experimental examples and theoretical models suggest that fundamentally any protein can form non-native aggregates in solution if incubated for sufficient time without the intervention of other process such as chemical or proteolytic degradation (Roberts 2007). The aggregation of protein via this mechanism is promoted by numerous stresses such as heat, chemical unfolding or shear.

1.1.2.3 Mechanism 3 – Aggregates by chemically-modified product

This mechanism is similar to mechanism 2, except that the aggregation proceeds from a covalent structural alteration (specific interaction) to the native state through a method of chemical degradation *i.e.* oxidation, deamidation, proteolysis, photo decomposition or disulphide bond shuffling. The non-native monomers are capable of inducing aggregation of native monomers by altering their physical properties such as their hydrophobicity or association tendency through the generation of sticky patches on their surface, or a reduction or change in the electrical charge in a way that reduces the electrostatic repulsion between monomers. They also induce changes to the secondary/tertiary structures of a protein. Aggregates formed through this pathway can be composed of

either chemically modified monomers, but can also have native monomers engaged in the aggregates (Philo and Arakawa 2009; Amani and Naeem 2013).

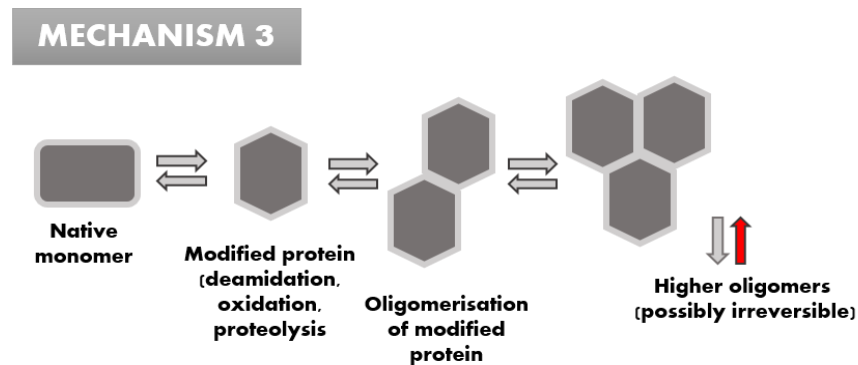


Figure 1.4: Schematic of protein aggregation by Mechanism 3 modified from (Philo and Arakawa 2009).

Chemical modification can result in cleavage of protein to yield lower-molecular weight products or cross-linkage of protein to yield higher molecular weight products.

1.1.2.4 Mechanism 4 – Nucleation-controlled aggregation

Nucleation driven aggregation involves free monomers coming together to form a critical sized nucleus, which is followed by elongation steps via the irreversible addition of monomers. Once the nucleus is formed, it acts as a seed for further growth of the aggregate. Nucleation driven aggregation is generally preceded by some structural alteration to the protein (Moreno-Gonzalez and Soto 2011). Aggregation proceeds in two phases, nucleation and then growth. The formation of a critical nucleus is stochastic in nature and lag times before the appearance of aggregates can vary from days to years (Wu and Shea 2011; Cohen, Vendruscolo *et al.* 2012).

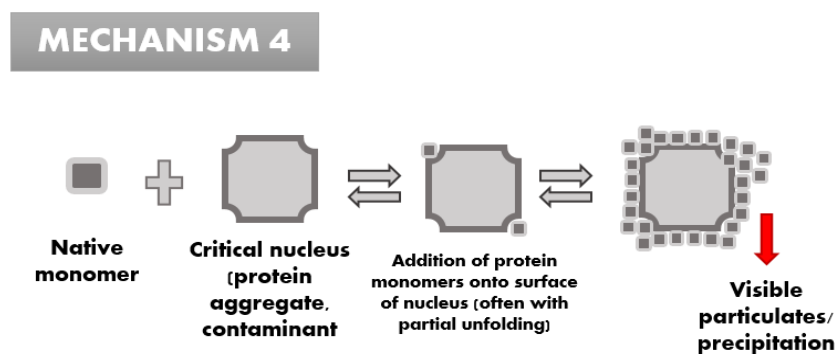


Figure 1.5: Schematic of protein aggregation by Mechanism 4 modified from (Philo and Arakawa 2009).

The nucleation process is the rate limiting step for aggregates of this type *i.e.* the energy barrier results from the free energy required to create a new solid-liquid interface and depends on the size of the aggregate, once the size of the new phase is above a critical size (energy barrier is at its highest) the nucleus grows in size (Chi, Krishnan *et al.* 2003). Seeding, known as seeded polymerization, of protein solutions with preformed aggregates can also encourage nucleation driven aggregation. These can be made from homologous seeds (same as monomeric protein) or heterologous seeds (different to monomer *i.e.* non-native). Because the addition of monomer to aggregates is strongly favoured by the thermodynamic stability of the aggregates, aggregates will dominate and recruit monomeric protein (Kopito 2000; Morales, Moreno-Gonzalez *et al.* 2013). Heterogeneous nucleation can also occur due to contaminants or impurities (foreign matter) within protein solutions acting as seeds for aggregation. Sources of contaminants can be manufacturing impurities such as silica particles shed by glass vials, or steel particles shed by piston filling pumps (Roberts and Wang 2010).

1.1.2.5 Mechanism 5 – Surface induced

Protein aggregation can be mediated by undesirable adsorption to surfaces during purification, filtration, transportation and storage. This mechanism is closely related to, and can sometimes be considered alongside mechanism 4 where the surface acts as the critical nucleus. This aggregation pathway is initiated by reversible native monomer binding to surfaces inducing a conformational change in the protein structure. Aggregation will then propagate either on the surface or from the altered monomer released back into the solution. Partially unfolded protein can adsorb at interfaces directly (Philo and Arakawa 2009).

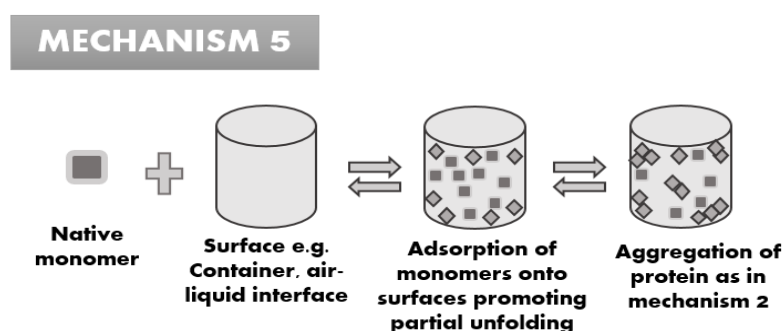


Figure 1.6: Schematic of protein aggregation by Mechanism 5 modified from (Philo and Arakawa 2009).

This mechanism does not always result in structural change in the protein. The reversible absorption of rvPAF-AH was observed without detectable changes in the secondary or tertiary structure but resulted in extensive aggregation (Chi, Weickmann *et al.* 2005). The driving force behind the interactions will depend on the surface and on the solution conditions *i.e.* solution pH, salt concentration, temperature (Wang, Nema *et al.* 2010). Surface induced aggregation can also occur at air-water interfaces found in vials, syringes and tubing. This is generally driven by hydrophobic interactions (Amani and Naeem 2013). The most reliable ways to measure protein adsorption and surface induced protein aggregation is to use fluorescent labelling to monitor changes in protein hydrophobicity (Yano 2012) or by monitoring loss of protein in the bulk solution (Roberts and Wang 2010).

1.2 Solution conditions

The aggregation behaviour of a protein relies on many factors, both intrinsic (protein structure) and extrinsic. The most significant extrinsic factor is the solution environment surrounding the protein, including pH, ionic strength, buffer species, excipients and contact materials (Mahler, Friess *et al.* 2009). Of these, pH is one of the most critical (Wang, Nema *et al.* 2010). Solution conditions for protein formulations need to be tightly controlled as proteins are expressed in an aqueous environment and many therapeutic proteins are in liquid form as they are typically administered by intravenous infusion or subcutaneous injection (Jiskoot, Randolph *et al.* 2012).

1.2.1 Solution pH

Throughout this thesis, solution pH was varied as each mechanism has been systematically studied. The pH of a protein solution dictates the distribution of charge on the protein surface and is governed by differences in pKa values of the ionisable groups (Garcia-Moreno 2009). These charges can influence intramolecular folding interactions and also intermolecular protein-protein interactions. Proteins are generally only stable in a narrow pH range due to a delicate balance between these interactions. By simply altering the solution pH, the charge-charge interactions are altered (Wang, Nema *et al.* 2010). In general, protein aggregation is most likely to occur at the pI of a protein, as its solubility should be at its lowest due to the loss in protein charge-charge repulsive interactions. At

extreme pH values, proteins are strongly charged. In some cases these dense charges on the protein surface greatly increase repulsive intramolecular and/or intermolecular interactions inducing partial unfolding of the protein exposing hydrophobic patches potentially increasing protein aggregation (Roberts and Wang 2010). However this is not the case for all proteins, some proteins can aggregate at a lower rate around their *pI* value (Majhi, Ganta *et al.* 2006). Solution pH can also play a role in the morphology of aggregates since the solution pH affects the type and density of charge on the surface and the degree of structural disruption (Krebs, Domike *et al.* 2009).

1.3 Characterising protein aggregates

Protein aggregates can be categorised or classified into several different types (usually governed by the mechanisms mentioned in section 1.1.2), by the interactions involved in their formation, and their size (Ripple and Dimitrova 2012). Currently there is no universal method of nomenclature. The following system is a useful way of categorising aggregates (modified from (Mahler, Friess *et al.* 2009; Roberts and Wang 2010).

1. **Protein Conformation:** predominantly native structure or non-native (significantly altered structure compared to native state/ partially unfolded)
2. **Linkage/type of bond:** covalent (held through covalent links *e.g.* disulphide links, will not typically dissociate with chemical denaturant) or physical (held through non-covalent physical interactions *e.g.* electrostatic, hydrogen bonds, will typically dissociated with chemical denaturant)
3. **Reversibility: reversible** (dissociates upon reversal of solution condition that initiated their formation) or **irreversible** (remain associated in higher order state upon reversal of solution conditions that initiated their formation)
4. **Size:** soluble (typically dimers, trimers, oligomers that are not visible under a microscope), subvisible (between 0.1 – 100 μm , can induce visible turbidity if present in high concentrations) or visible (larger than 100 μm , can be visualised with naked eye)

The aggregation process can lead to aggregates ranging in size from nanometres to millimetres. For this reason, it is useful to categorise the particle types produced. Again there is no universal method, nor is there a consistent definition of what “soluble” or

“insoluble” aggregates refers to. Soluble aggregates are generally considered as those that are not visible as discrete particles, while insoluble aggregates are referred to aggregates that are often visible to the unaided eye (Cromwell, Hilario *et al.* 2006). The terms subvisible and visible are also used to discriminate between particle sizes (Ripple and Dimitrova 2012). Figure 1.7 is a representation of the size range of aggregates and their relevant terms of characterisation.



Figure 1.7: Representation of the approximate size and characterisation terms for protein aggregates modified from (Mahler, Friess *et al.* 2009).

The macroscopic attributes (morphology) of protein aggregates is highly variable but can be generalised into amorphous or fibrillar material depending on the protein and its environment (Mahler, Friess *et al.* 2009; Roberts and Wang 2010).

1.3.1.1 Fibrillar aggregates

Fibrillar aggregates are frequently associated with amyloid plaque formation. Amyloidogenesis is a condition defined by the process of soluble protein turning into insoluble aggregates known as amyloid fibrils generally driven by a nucleation dependent pathway (Bhak, Choe *et al.* 2009). Some of the most-well known amyloid related diseases are neurodegenerative disorders such as Alzheimer’s diseases, Parkinson’s disease and Huntington’s diseases associated with the accumulation of these fibrils (Dobson 1999). The proteins and peptides associated with these diseases do not share primary amino acid sequences or structural homologies. There is no apparent structural link between them that may account for their tendency to form fibrils, yet regardless of the proteins initial conformation, they all share common fibril characteristics (Roberts and Wang 2010; Amani and Naeem 2013). Proteins not associated with these diseases states have also been shown to form fibril structures with the same properties as amyloid fibrils (Krebs, Bromley *et al.* 2005).

Fibrils are so called after their long, thin, unbranched fibrous shape (figure 1.8). They can be quite diverse in their diameter, length and interconnection between amyloid stands.

They generally exhibit optical birefringence and increased fluorescence upon association of dyes such as Congo red and Thioflavin T, yet the mechanism of their association to such dyes is still largely unknown (Biancalana, Makabe *et al.* 2009; Biancalana and Koide 2010). The pathways to fibril formation usually consists of a conformational change from random coil (no structure) or α -helix into β -strand. They lack alpha content but have considerable beta-sheet formation, with parallel strands orientated perpendicular to the fibre axis. Antiparallel sheet orientations have also been observed. Their formation is stabilised by backbone hydrogen bonding and hydrophobic interactions, rather than through specific interactions of different side chains (Chiti, Webster *et al.* 1999; Hou and Zagorski 2004).

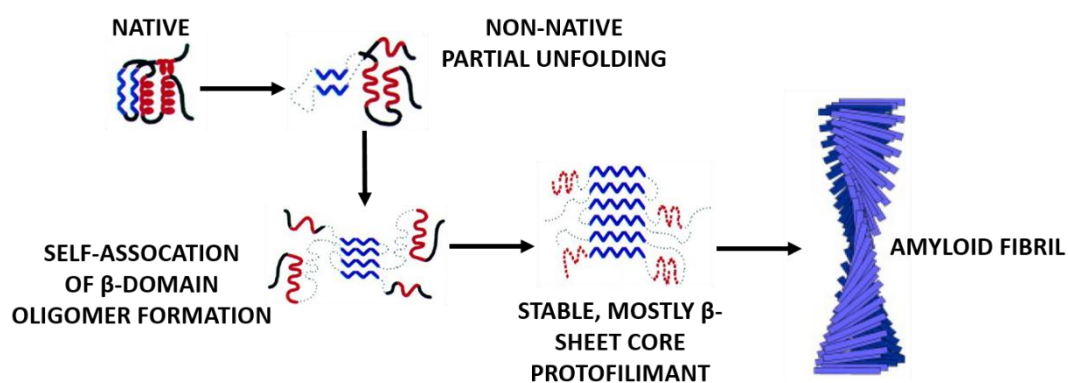


Figure 1.8: Schematic of fibrillar aggregate formation modified from (Booth, Sunde *et al.* 1997).

1.3.1.2 Amorphous aggregates

Amorphous aggregates (figure 1.9) lack long range order. Originally it was assumed they were composed of completely unfolded material held together by random associations between hydrophobic residues, since harsh denaturants (SDS, urea) were required to resolubilise them. Today amorphous aggregates have been shown to consist of various types of assemblies without ordered intermolecular interactions (Yoshimura, Lin *et al.* 2012) and also by fully folded proteins without covalent interactions *e.g.* α -crystallin in patients with cataracts (Sugiyama, Fujii *et al.* 2010). They can be thought of as heterogeneous populations of aggregates, from subpopulations of reversible, native, unmodified aggregates and irreversible, unfolded disulphide cross-linked aggregates (Narhi, Schmit *et al.* 2012). Detailed information about the structure of amorphous aggregates is lacking as they tend to scatter light interfering with spectroscopic techniques

typically used to characterise structure, and also because of their heterogeneity they are hard to study (Roberts and Wang 2010).

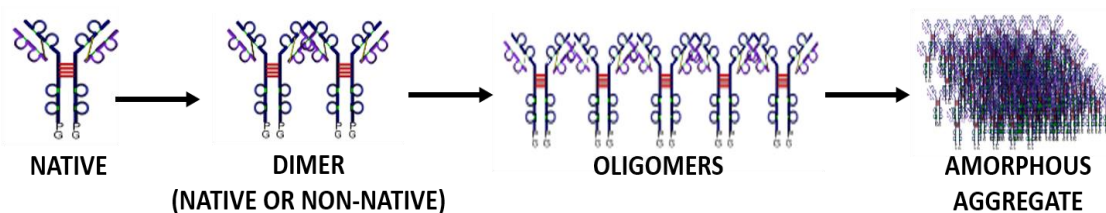


Figure 1.9: Schematic of amorphous aggregate formation.

Many amorphous aggregates have increased beta-sheet structure and diminished alpha helicity compared to the native state. Some amorphous aggregates have been shown to associate with dyes such as Congo red and Thioflavin T, suggesting at least some level of ordered structural elements common with fibrillar aggregates (Biancalana and Koide 2010). Insulin initially converted into reversible amorphous aggregates, undergoes structural reorganisation from α -helix to β -sheet structures capable of associating with ThT (Owczarz and Arosio 2014). It was also shown to associate with the hydrophobic pocket of human serum albumin devoid of cross β -structure characteristic of fibril structure (Sen, Fatima *et al.* 2009). ThT has been shown to associate with non-amyloid amorphous β -aggregates devoid of fibril morphology (Chang, Liao *et al.* 2009).

1.4 Methods to measure protein aggregation

Aggregation of protein samples during formulation and production is a serious safety concern due to their adverse effects and immune response in patients (Cromwell, Hilario *et al.* 2006). Therefore there is a need to establish enhanced methods to detect and monitor protein aggregation during manufacturing during both upstream and downstream processing. Many forced degradation studies have been conducted to induce aggregation in protein solutions, to study the aggregation process and to identify the appropriate techniques to monitor aggregation (Hawe, Wiggenhorn *et al.* 2012). While these studies are useful, care must be taken when interpreting the results. Since there are several mechanisms by which aggregation can proceed, it must be clear from the stress applied, which particular mechanism is driving aggregation. Techniques or approaches suitable for studying the fibrillation process of proteins as a result of structural change are not necessarily applicable to the study of reversible aggregates, while techniques useful for measuring protein unfolding will not hold for detecting oxidation related modifications.

In the absence of any prior knowledge about the characteristic of the aggregated material, it is necessary to employ multiple techniques to obtain as much information as possible (Zolls, Tantipolphan *et al.* 2012). As discussed in section 1.2, aggregates across a range of sizes (nm to μm) may be present, meaning no single technique can reliably characterize the full spectrum of aggregates present. A combination of several techniques is necessary (Narhi, Jiang *et al.* 2009). A range of the major analytical techniques used for aggregation analysis is shown in Figure 1.10 spanning the wide variety of particle sizes that may arise as a result of aggregation.

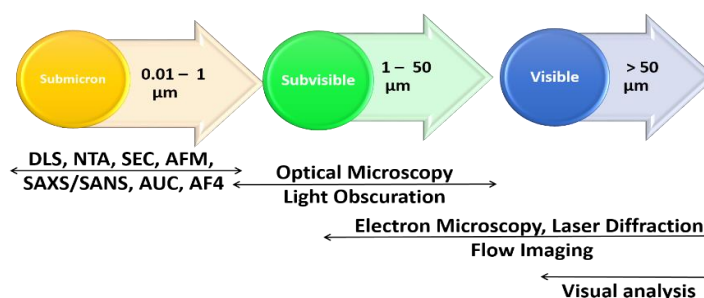


Figure 1.10: Variety of analytical techniques used to measure protein aggregation over a range of length scales.

The kinetics and products of protein aggregation have been measured using at least 20 different analytical techniques, each having pros and cons, which should be carefully considered in designing or selecting a suite of techniques to monitor aggregation of protein. Table 1.2 lists such techniques modified from (Mahler, Friess *et al.* 2009; Roberts and Wang 2010; den Engelsman, Garidel *et al.* 2011).

| Method | Application | Size range | Advantages | Disadvantages |
|--------------------------|--|-------------------------|---|---|
| SE-HPLC | Size estimation and quantification (soluble aggregates) | 50 –1000 kDa | Robust, sensitive, precise | Sample dilution, limited resolution, limited particle size range |
| Mass spectroscopy | Size/charge estimation and quantification | Atomic resolution - MDa | Very high resolution, high accuracy, precision, detailed structural information | Gas phase analysis, requires exchange into volatile buffer, costly and expertise training |
| Electrophoresis | | | | |
| - Native | Size estimation and to distinguish reducible covalent from non-covalent aggregates | 5 – 1000 kDa | - Easy to perform, detect native aggregates | - Does not detect non-covalent aggregates |
| - SDS-PAGE | | | - Differentiation disulphide linked, non-covalent | - Mobility depends on charge, size, size resolution |

| | | | | |
|--|--|--|---|--|
| Capillary electrophoresis | Size estimation and quantification (soluble aggregates) | 5 – 1000 kDa | High resolution, quantification, fast separation | Possible protein reactions with capillary, doesn't detect non-covalent |
| Microscopy - Light - Electron - Atomic force | Size and shape estimation | - > 1 μm - mm - nm – mm - nm (resolution 0.01 nm) | - Easy, size & shape- - Large size range, high resolution - Molecular resolution | - Limited resolution- - Possible artifacts, not quantitative - Time consuming, expensive, limited representativeness |
| Static light scattering | Size and shape estimation | 20 nm – 2 μm | Single particles detection, counting, characterisation | Requires sample dilution |
| Dynamic light scattering | Size distribution | 1 nm – 10 μm | Easy to perform, non-destructive, high sensitivity | Complicated data analysis, low resolution, sensitivity to artifacts |
| Analytical ultracentrifugation | Size, shape estimation and quantification | 1 nm – 0.1 μm | Absolute method, measurement of molecule size, shape, high resolution | Strongly dependent on instrument quality, complex data analysis |
| Light obscuration | Size and quantification (insoluble aggregates) | 2 – 100 μm | Rapid analysis, counting and clustering | Large sample volume, no morphological information |
| Coulter counter | Size and number quantification (insoluble aggregates) | 0.4 – 1.2 μm | Single particle detection, counting and characterisation | Dilution of low conductivity samples into electrolyte solution |
| Visible inspection | Absence or presence of visible aggregates | > 50 μm - mm | Easy to perform, information on size and shape | Low resolution, limited particle discrimination, subjective |
| Turbidimetry, nephelometry | Particle size/concentration | 100 nm – several μm | Easy to perform, various designs and methodologies | Observed signal depends on both size and concentration, no particle specific detail |
| UV-Vis spectroscopy | Soluble and insoluble aggregates; solution property (no quantification possible) | N/A | Easy to perform, non-destructive, detection of particles outside absorption bands | Complicated data interpretation, limited information on particle properties |
| Circular dichroism | Structural analysis | N/A | Non-destructive, easy to perform, low protein amount required, undiluted sample analysis possible | Interference by light scattering, limited resolution, complicated data analysis |
| Fluorescence spectroscopy | Structural analysis | nm - μm | Non-destructive, comprises multiple fluorescent techniques | Destructive when using extrinsic dyes, interference from excipients or impurities |

| | | | | |
|-----------------------------------|---------------------|-----|---|---|
| | | | combined with microscopic tools | |
| (FT-)infrared spectroscopy | Structural analysis | N/A | Non-destructive, solid state analysis possible, little inference from scattering, microscopic tools | Low sensitivity, protein concentration >1 mg/ml, possible artifacts |
| Raman spectroscopy | Structural analysis | N/A | Non-destructive, solid state analysis possible, little inference from scattering, microscopic tools | Particle isolation can lead to artifacts |

Table 1.2: Overview of analytical methods used for detecting, measuring and characterising protein aggregates

1.5 Unfolding pathways

Most of the knowledge we have about the forces that stabilise protein conformation comes from thermodynamic studies of folding/unfolding (Gomez, Hilser *et al.* 1995) (Pace 1986). Protein conformational stability can be described in term of Gibbs free energy, ΔG , between the folded and unfolded state;

$$\Delta G_{UN} = G_U - G_F \quad 1.5$$

where ΔG_{UN} is the difference in Gibbs free energy upon unfolding, and G_U and G_F are the Gibbs free energies of the unfolded and folded state respectively (Sancho 2013). In order for proteins to maintain their folded native conformation against more stable aggregated conformations, a high energy barrier must be overcome. This energy barrier is related to the mechanical stability of a protein *i.e.* the resistance of native proteins to be unfolded by external forces (Sancho 2013).

Two-state behaviour systems are the simplest models for protein unfolding (Pace 1986; Eftink 1994); the unfolded state (U) and the folded, native state (F) are the only two conformations and the equilibrium constant, K_{un} , is the ratio of unfolded protein to the folded native state;

$$U \rightleftharpoons F \quad 1.6$$

$$K_{UN} = \frac{[F]}{[U]} \quad 1.7$$

which can be used to calculate the ΔG_{UN} of unfolding:

$$\Delta G_{UN} = -RT \ln K_{UN} = \Delta H_{UN} - T \Delta S_{UN} \quad 1.8$$

where ΔH_{UN} and ΔS_{UN} are the enthalpy and entropy associated with unfolding.

A two-state model of protein unfolding is characterised typically by a linear plot of the natural logarithm of the rate constants for unfolding, $\ln K_{UN}$, and refolding, $\ln K_F$, versus denaturant concentration [D]. A non-linear plot of such parameters can result from the presence of a meta-stable intermediate on or off the folding pathway (Jackson 1998). For some proteins the initially unfolded protein rapidly adopts a limited set of conformations under refolding conditions before refolding is complete. These states are termed “intermediates” (*I*) (Creighton, Darby *et al.* 1996);



where there is evidence of one or more intermediates in the folding process, then the thermodynamic models become more complex;

$$K_{F \rightarrow I} = \frac{[I]}{[F]}, \quad K_{F \rightarrow U} = \frac{[U]}{[F]} \quad 1.10$$

The corresponding rate constants need to be calculated for each intermediate state transition; $\Delta G_{F \rightarrow I}$, $\Delta H_{F \rightarrow I}$, $\Delta S_{F \rightarrow I}$, $\Delta G_{I \rightarrow U}$, $\Delta H_{I \rightarrow U}$, $\Delta S_{I \rightarrow U}$.

Intermediate states are generally less folded than the native state but more folded than the denatured state (Sancho 2013). Identifying the existence of intermediate states and characterising its spectral properties is a difficult challenge as it will depend on how extensively the intermediate state is populated and requires a significant difference in fluorescence properties of the three states (Eftink 1994). The kinetic importance of these states is still largely unknown, some believing they are off-pathway conformations that are kinetically trapped (Gruebele 1999). In order for the intermediate state to be deemed a productive and obligatory state through which the molecule must pass, a lag phase before the appearance of the native state is formed is required, corresponding to the time it takes for *I* to be generated. However this lag phase is generally undetectable (Creighton, Darby *et al.* 1996). Conventional free energy diagrams as in Figure 1.11 are commonly used to illustrate the folding/unfolding pathway of a protein. The wells portray the sequence and stability of intermediates and end states.

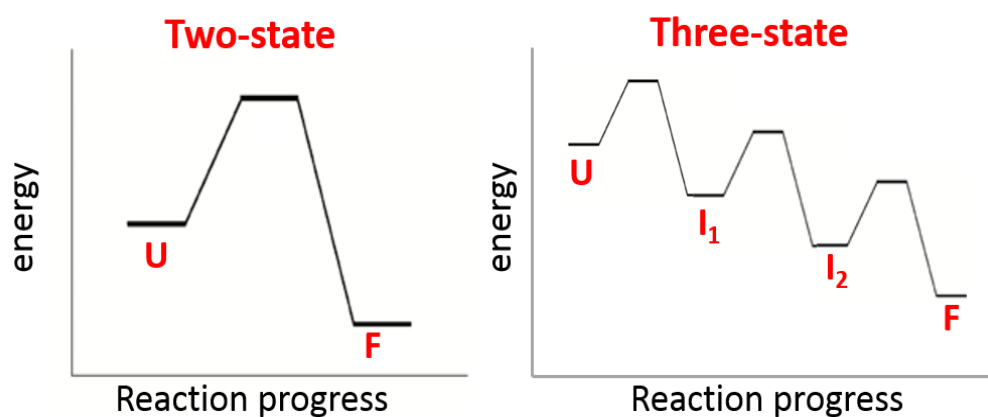


Figure 1.11: Free energy diagram illustrating the folding/unfolding of protein molecules. Change in the relative free energies for two-state or three-state (intermediate) folding. Thermodynamic states (U- unfolded, I- intermediate, F- folded).

1.6 Forced degradation and aggregation

Tuning environmental conditions and stress mechanisms can result in the destabilisation of the native conformation of protein and also encourage protein-protein interactions leading to protein aggregation as discussed in section 1.1.2. This type of stress-testing/forced-degradation studies (*i.e.* accelerated aggregation) are commonly used during the development of protein formulations (Hawe, Wiggenhorn *et al.* 2012). These practises are essential for obtaining information on potential degradation pathways, ensuring analytical methods are satisfactory and identifying potential stabilisers and optimal conditions are achieved (Hasija, Li *et al.* 2013).

1.6.1.1 Thermal stress – Mechanism 2

Proteins have a limited temperature range where their structural integrity is maintained, outside of this range the protein is likely to denature. Thermal denaturation, is one of the most common stress mechanisms used for accelerated aggregation testing (Bischof and He 2005; Hawe, Wiggenhorn *et al.* 2012). Exposure of proteins to elevated temperatures perturbs their native conformation to such a degree that intermediate (partially unfolded) states are formed that can promote aggregation. The melting temperature (T_m) of a protein is often used to assess its thermal stability. T_m is expressed as the peak temperature where a protein molecule is half way through the unfolding process (temperature where half the protein is denatured) and therefore proteins often rapidly aggregate near their T_m (Cooper 1999; Wang, Nema *et al.* 2010). The prediction of protein aggregation on T_m is not wholly reliable, a higher T_m does not ensure a lower tendency of a protein to aggregate by other

mechanisms, and instead it is the actual solution conditions that control protein aggregation (Blumlein and McManus 2013). The thermal unfolding of IL-Ira mutant ($T_m = 63$ °C) did not lead to immediate aggregation, and was reversible measured by CD (Myers, Pace *et al.* 1995). To accurately interpret thermal scans, a prior knowledge or assumptions of the relative enthalpy and heat capacity values of the aggregates being formed are required. T_m values should only be used for qualitatively ranking of solution conditions where aggregation is likely to occur (Weiss, Young *et al.* 2009).

After thermal denaturation, during cooling and refolding of the protein there are competing pathways leading to either the native state or an aggregated state. These are governed by attractive and repulsive forces brought about by the solution environment surrounding the protein. At pH values close to the pI of a protein, there is a shift from net repulsion between the protein molecules to attraction. Therefore repulsive interactions are essentially screened and there is an overall attraction between the molecules increasing the chance of aggregation as the protein is thermally unfolded at elevated temperatures. The screening of the repulsive charges, is coupled with the exposure of hydrophobic residues on the protein surface thus increasing the probability of protein aggregation through either specific or non-specific interactions between hydrophobic regions of different molecules (Goldberg, Rudolph *et al.* 1991; Calamai, Canale *et al.* 2005).

Several methods can be used to measure thermal denaturation of protein. Differential scanning calorimetry (DSC) is commonly applied to monitor the thermal behaviour of protein solutions as it is the only technique that can measure the calorimetric effect associated with the protein structural change. It can also directly quantitate protein unfolding. It measures the specific heat of a sample as temperature is increased and can detect the heat absorbed during transition from a native to a denatured state, thus determining the relative amounts of each state (Freeman, Borrelli *et al.* 1999). Numerous groups have followed the calorimetric events associated with protein structural change as a result of their structural environment, altering parameters such as protein concentration, pH, ionic strength, presence of co-solutes and ion type (Garbett, Mekmaysy *et al.* 2009; Luo, Wu *et al.* 2011; Blumlein and McManus 2013)

1.6.1.2 Chemical denaturation induced unfolding – Mechanism 2

The native state of a protein is characterised by its tightly folded and highly ordered conformation. With the use of chemical denaturants (urea, guanidine hydrochloride), this folded native state can be induced to unfold the polypeptide chain, resulting in the protein adopting a largely disordered structure. The formation of these partially unfolded states can induce protein aggregation through interactions detailed in Table 1.1. (Roberts and Wang 2010).

The chemical unfolding of a protein using denaturants has been carried out in countless studies (Dill and Shortle 1991; Street, Courtemanche *et al.* 2008). Despite their widespread use, the molecular basis for urea and guanidine hydrochloride induced protein denaturation remains unknown and is under intense investigation (Cao and Li 2008; Constatinescu, Herrmann *et al.* 2010; Gao, She *et al.* 2010; Buchner, Murphy *et al.* 2011). Both “indirect” and “direct” mechanisms are proposed. The ‘indirect’ mechanism suggests that the denaturant alters the solvent environment by disruption of the water structure, weakening the hydrogen bonds between water and the protein to facilitate the exposure of residues in the hydrophobic core. One such model was proposed by Frank and Franks whereby urea only interacted with the disordered state of water, pulling it away from the more ordered state and thereby indirectly “breaks” the water structure (Frank 1968).

In recent years, the general consensus falls to the ‘direct’ mechanism which suggests that the denaturant unfolds the protein through direct interactions with the protein by displacement of water molecules from the protein. It directly binds to and stabilises the denatured state favouring unfolding. Kumaran and Ramamurthy used several urea derivatives to show direct hydrogen-bonding interaction of the N-H moiety of urea with the hydrogen bonding functional groups of BSA to induce unfolding (Kumaran and Ramamurthy 2011). Schellman (Schellman 1994) reported that one urea replaces one water molecule, a view not supported by Lim *et al.* who proposed the idea that since urea is approximately three times larger than a water molecule, that a single urea molecule should replace three water molecules (Lim, Rosgen *et al.* 2009). A third view believes that both mechanisms might play a role in chemical unfolding of protein. Daggett and co-workers (Bennion and Daggett 2003) showed both direct urea interactions consisting of

hydrogen bonding to the polar moieties of the protein, while also promoting indirect unfolding by altering the water structure and dynamics.

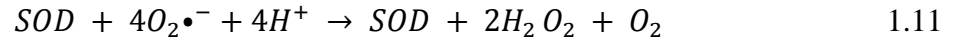
1.6.1.3 Chemical alteration – Mechanism 3

Chemical modifications of proteins are complex and do not appear to be solely dependent on the amino acid affected but also on sequential and structural neighbours (Reubsaet, Beijnen *et al.* 1998). Several spontaneous and induced chemical alterations such as oxidation, deamidation, hydrolysis, β -elimination and racemisation can affect protein degradation and aggregation resulting in fragmentation and aggregation of the protein (Roberts and Wang 2010). The oxidation of proteins will be discussed in more detail below.

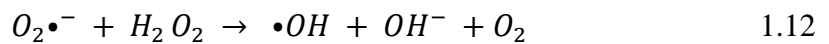
1.6.1.4 Oxidation

Aerobic organisms constantly produce small amounts of free radicals. Free radicals are a reactive chemical species that have an odd number of electrons rendering them highly active with relatively short lifetimes (Davies and Pryor 2005). Oxidative stress (OS) can occur within organisms when their defence mechanisms designed to protect themselves are unable to adequately remove these free radicals, or they are generated in such excess beyond the ability of the defence to cope with. OS can result in several metabolic dysfunctions; peroxidation of membrane lipids, depletion of nicotinamide nucleotides, rises in intracellular free Ca^{2+} ions, cytoskeleton disruption and DNA damage (Jomova, Vondrakova *et al.* 2010).

Free radicals are formed via biochemical processes that involve the generation of reactive oxygen and nitrogen species (ROS and RON), as a by-product from normal cellular metabolism (Kaur, Kaur *et al.* 2014). ROS are molecules that contain one or more unpaired electrons and singlet oxygen, rendering them highly reactive and damaging to cells. They are formed by the consecutive intracellular reduction of molecular oxygen to produce the superoxide anion radical ($\text{O}_2\bullet^-$), hydrogen peroxide (H_2O_2) and hydroxyl radical ($\bullet\text{OH}$) (Liu, Fiskum *et al.* 2002). $\text{O}_2\bullet^-$ is the primary free radical and is produced as a side product in the mitochondria during energy transduction reactions. $\text{O}_2\bullet^-$ is a short lived radical and the major proportion of it is converted into H_2O_2 and molecular oxygen by superoxide-dismutase (SOD), an enzyme that catalyses the partitioning of the toxic $\text{O}_2\bullet^-$ radical (Henkler, Brinkmann *et al.* 2010);



Hydrogen peroxide is a harmful by-product of many normal metabolic processes. To prevent damage, it must be quickly converted into other, less dangerous substances. H_2O_2 is a strong oxidant and can be further reduced into the hydroxyl radical $\bullet OH$, the neutral form of hydroxide ion and is a highly reactive radical with a half-life of nanoseconds (Sies 1997). Among the many types of free radicals generated, the hydroxyl radical is one of the most potent. It is formed via the Haber-Weiss reaction by oxidation of superoxide anion radical;



The oxidation of proteins is believed to be a relatively rapid process as the changes in conformation exposes more susceptible residues. For example, the nitration of tyrosine to form peroxynitrite ($ONOO^-$) from the reaction of NO^- with superoxide, $NO\bullet$ and $O_2\bullet^-$ occurs in $k = 6.7 \times 10^9 \text{ M}^{-1} \text{ s}^{-1}$ (Huie and Padmaja 1993), while peroxiredoxin has been shown to reaction with H_2O_2 with a rate constant of $k = 1.3 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ (Peskin, Low *et al.* 2007). In excessive cases, massive protein damage can result in the formation of large aggregates that cannot be degraded themselves. In excessive cases, massive protein damage can result in the formation of large aggregates that cannot be degraded themselves. There are numerous mechanisms for the induction of protein oxidation as highlighted in Table 1.3. All of the amino side chains can become oxidatively modified making them more susceptible to oxidation (Shacter 2000). The rate of protein oxidation can be affected by both intrinsic (flexibility of the peptide backbone and overall structure) and extrinsic factors (pH and buffer type). It can also be broken down into two categories; site specific (metal catalysed oxidation (MCO)) or non-site specific (photooxidation and free radical cascades) (Manning, Chou *et al.* 2010).

| <i>Modification</i> | <i>Amino acids involved</i> | <i>Oxidising source*</i> |
|---|-----------------------------|---|
| Disulfides, glutathiolation | Cys | All, ONOO ⁻ |
| Methionine sulfoxide | Met | All, ONOO ⁻ |
| Carbonyls (aldehydes, ketones) | All (Lys, Arg, Pro, Thr) | All |
| Oxo-histidine | His | γ -Ray, MCO, ¹ O ₂ |
| Dityrosine Tyr | Tyr | γ -Ray, MCO, ¹ O ₂ |
| Chlorotyrosine | Tyr | HOCl |
| Nitrotyrosine | Tyr | ONOO ⁻ |
| Tryptophanyl modifications | | |
| (N-formyl)kynurenine | Trp | γ -Ray |
| Hydro(pero)xy derivatives | Val, Leu, Tyr, Trp | γ -Ray |
| Chloramines, deamination | Lys | HOCl |
| Lipid peroxidation adducts | | |
| (MDA, HNE, acrolein) | Lys, Cys, His | γ -Ray, MCO (not HOCl) |
| Amino acid oxidation adducts | Lys, Cys, His | HOCl |
| Glycooxidation adducts | Lys | Glucose |
| Cross-links, aggregates, fragments | Several | All |

Table 1.3: Oxidative modifications of proteins (modified from (Shacter 2000))

*MCO-metal catalysed oxidation; All = γ -ray, HOCl, ozone, ¹O₂.

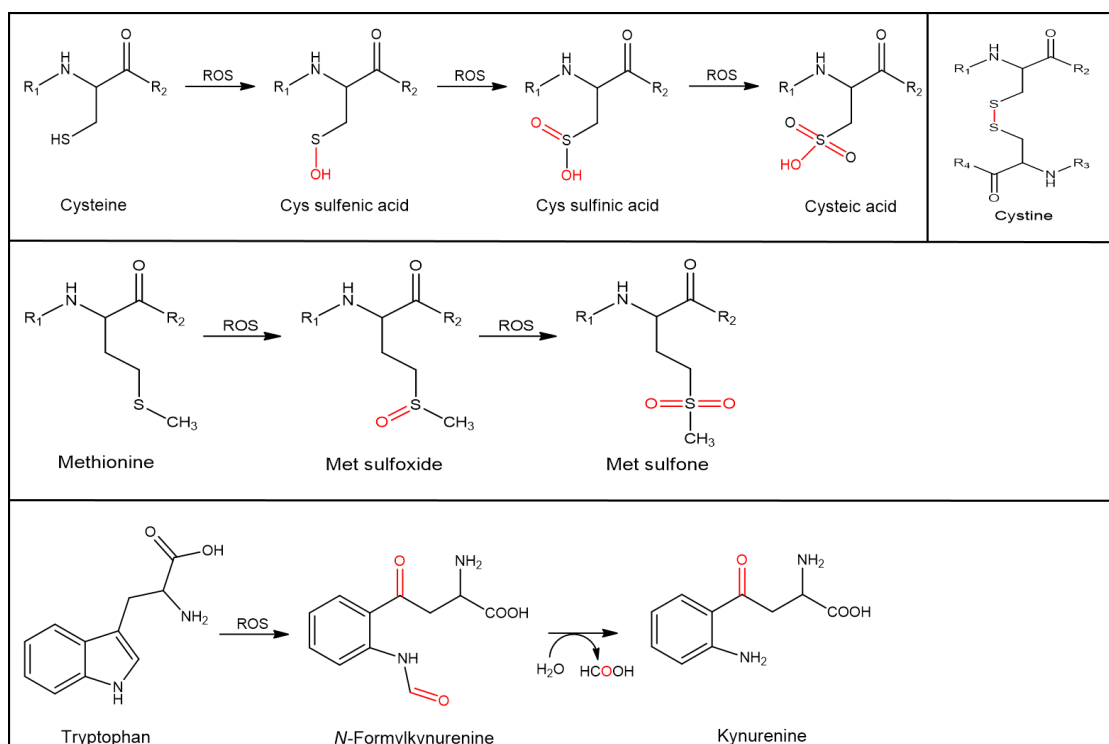


Figure 1.12: Commonly observed oxidative modifications of protein amino acids caused by reactive oxygen species (ROS) modified from (Moller, Jensen et al. 2007).

The most common oxidative protein modifications of amino acids residues are shown in Figure 1.12. These include cysteine, methionine and tryptophan. Peroxide molecules can be introduced into a formulation by trace amounts of impurities in the added excipients such as surfactants. As a result, oxidation of proteins in a forced condition of excess H₂O₂ is one of the most commonly used models in order to understand protein oxidation in solutions. The oxidation of protein can be limited with more efficient packing and storage, lessening the likelihood for photooxidation or less exposure to oxygen by decreasing the deal volume in vials. The addition of additives into the formulation has shown to improve oxidation levels (Manning, Chou *et al.* 2010)

1.7 Model proteins

Throughout this thesis, lysozyme and bovine serum albumin have been used as model systems to examine how protein aggregation proceeds for each mechanism. These two proteins were selected since they are available in high purity, at a reasonable cost and the properties of these proteins have been well established through many studies. Some of their structural features are listed in Table 1.4

| <i>Protein</i> | <i>Mw</i> (Da) | <i>Secondary</i> <i>structure</i> | <i>pI</i> | <i>S-S</i> <i>bridges</i> | <i>Chromophores</i> | <i>T_m</i> |
|-----------------|-------------------|--|-----------|------------------------------|----------------------------------|----------------------|
| <i>Lysozyme</i> | 14,307 | 4 α -helices 5 β -sheets random coil | 11 | 4 | Phen - 3 Tyr - 3 Trp - 6 | >75 °C |
| <i>BSA</i> | 66,000 | 67% α -helices 16 % β -sheets 29% random coil | 5.1 | 17 | Phen - 27 Tyr - 20 Trp - 2 | >65 – 70 °C |

Table 1.4: Properties of lysozyme and BSA.

1.7.1 Lysozyme

Lysozyme is an enzyme found in bodily secretions (nasal, tears, saliva, serum). Hen egg white lysozyme (HEWL), is a natively monomeric protein made up of 129 amino acids that fold into a compact globular structure with a molecular weight of 14, 307 Da. This structure was first sequenced in 1963 by enzymatic digestion of reduced

carboxymethylated HEWL (Canfield 1963; Jolles, Jauregui Adell *et al.* 1963). It is composed of two domains, alpha (green features) and beta (red features) (figure 1. 13).

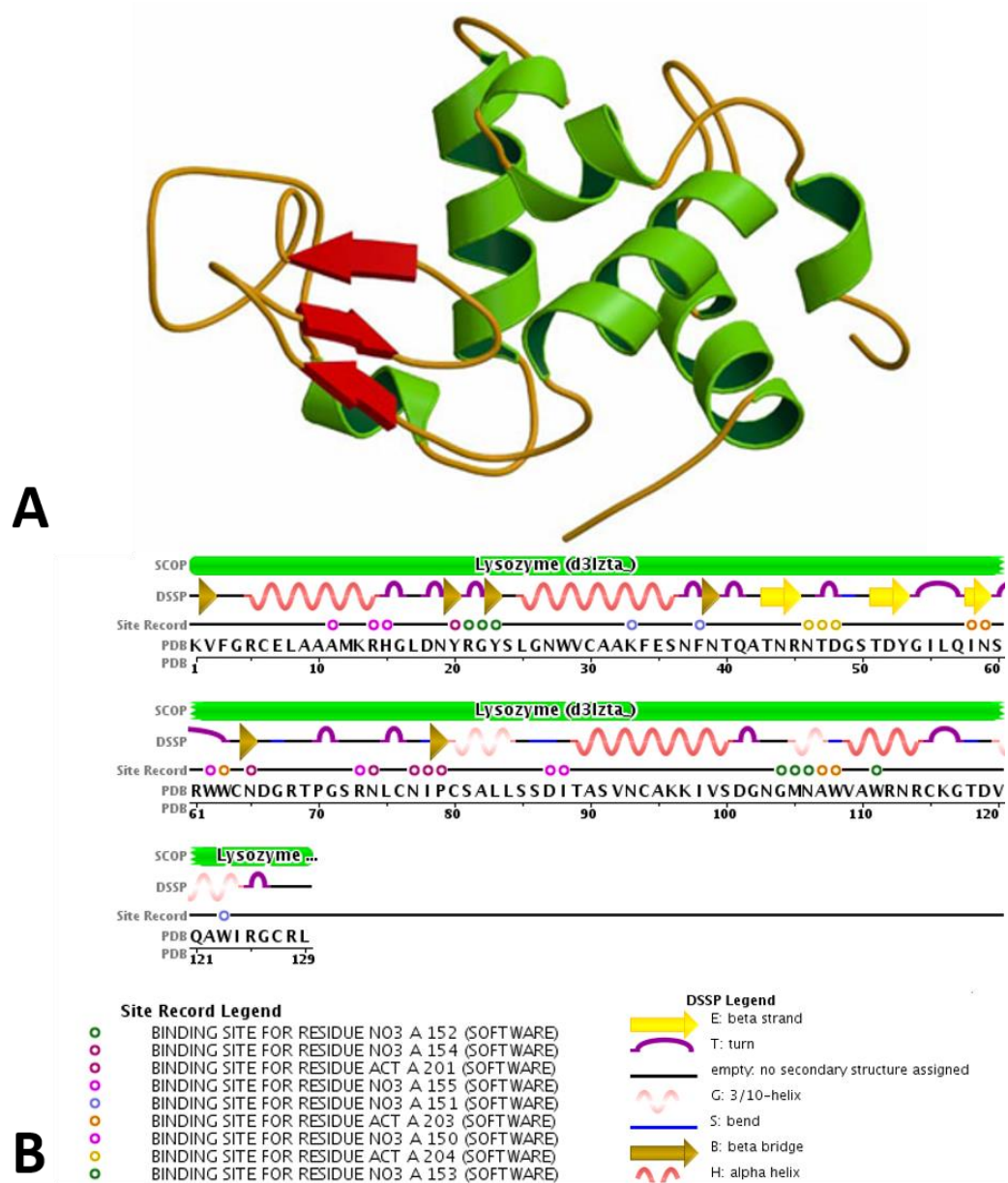


Figure 1.13: A) Ribbon representation of the crystal structure of and B) the amino acid sequence (PDB code: 3LZT) (ProteinDataBank).

The α -domain (residues 1 – 35, 85 – 129) comprises four α -helices, and a single short 3_{10} -helix. Within the α -domain there is a core of hydrophobic side-chains that are packed closely together. Its β -domain (residues 36 – 84) is made up of a triple-stranded antiparallel β -sheet, a long loop and a 3_{10} -helix. There is no core hydrophobic region, instead hydrogen bonds and a number of small hydrophobic clusters are believed to be responsible for maintaining the tertiary structure. There is also a long exposed loop region

(Schwalbe, Grimshaw *et al.* 2001). Four disulphide bonds formed between eight cysteine residues (6:127, 30:115, 64:80 and 76:94) also contribute to the folded conformation. The disulphide bond between cysteine residues 76:94 supports both domains. The ratio of tryptophan to tyrosine is 2:1. There are 6 tryptophan residues from which the intrinsic fluorescence dominates (Trp-28, -62, -63, -108, -111 and -123). Of these Trp-62, -63 and -108 are located in the active site, from which the bulk fluorescence comes from Trp-62 and -108 (Imoto, Forster *et al.* 1972; Formoso and Forster 1975).

1.7.2 Bovine serum albumin

Serum albumin is one of the most abundant proteins in plasma and aids in maintenance of circulating fluid within the vascular system, acting to transport a variety of endogenous and exogenous compounds (Zilg, Schneider *et al.* 1980). Bovine and human serum albumin are commonly used in experimental studies. The structure of serum albumins have been extensively studied (Peters 1985). Both are very similar globular proteins performing the same functions and sharing high sequence homology with the major difference being their number of amino acids; HSA is made up of 585 with BSA having 582.

BSA has a molecular weight of 66,000 Da. It is a multi-domain monomer protein made up of three homologous domains (I, II, III) which are divided into nine loops giving it an overall heart-like shape. The secondary structure consists of 55% α -helix, 16% β -sheet and 29% disordered structure. Each domain displays specific structural functional characteristics and is composed of two further subdomains; subdomain A containing six helices and subdomain B with four helices. It contains 35 cysteine residues of which 17 are involved in disulphide bonds stabilising its tertiary structure, with one free thiol group at position 34 (Reed, Feldhoff *et al.* 1975; Peters 1985; Carter, He *et al.* 1989; Bujacz 2012). It's believed that this free thiol group may not be essential in the formation of large aggregates, but is likely to play a role in the intermolecular cross-links between aggregates (Hillier, Lyster *et al.* 1980; Matsudomi, Rector *et al.* 1991; Boye, Alli *et al.* 1996) and in the thermal aggregation pathways of BSA (Militello, Casarino *et al.* 2004). Despite numerous methods to determine its tertiary structure, its three dimensional molecular structure has remained largely under debate (Ferrer, Duchowicz *et al.* 2001; Huang, Kim *et al.* 2004). Based on hydrodynamic experiments (Bloomfield 1966; Wright

and Thompson 1975) and low angle X-ray scattering (Bloomfield 1966), serum albumin was postulated to be an oblate ellipsoid. NMR studies have indicated it's a heart shaped structure (Bos, Labro *et al.* 1989). Bujacz recently determined the crystal structure of BSA at 2.47 Å resolution (figure 1.14) exhibiting only 75.8% homology to HSA with differences in surface structure, the binding pockets and charge distribution (Bujacz 2012).

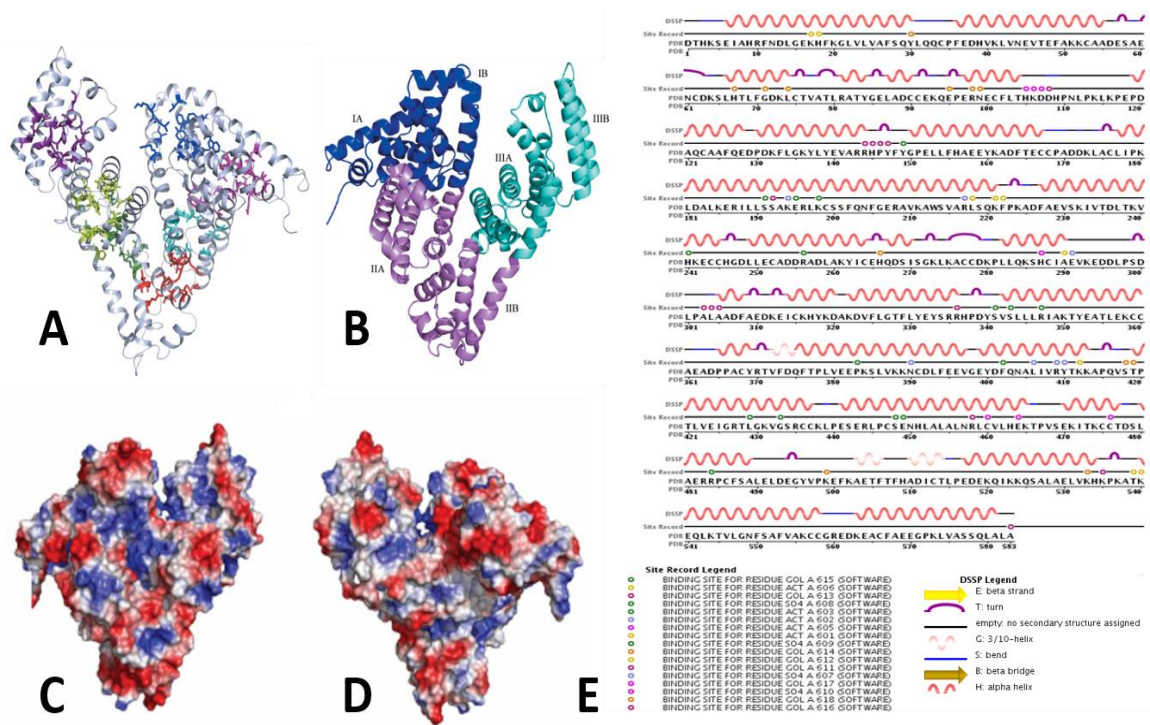


Figure 1.14: A) Crystal structure of BSA. B) Domain division of BSA (I – blue, II – purple, III – cyan). C) Charge distribution on the surface of BSA; molecules in the orientation with domain I on the left and further domains arranged anticlockwise. D) Charge distribution on the surface of BSA; domain I is on the right and the domains are arranged clockwise. E) The amino acid sequence (PDB code: 1A06) (ProteinDataBank ; Bujacz 2012).

BSA has several conformational transformations induced by pH changes (figure 1.15). BSA is natively folded between pH 4.5 to 7. Lowering the pH to 3.5 – 2 the fast form (F, fast migration in gel electrophoresis) and the expanded form (E). Raising the pH from 7 to 9 introduces the basic form (B) and finally the expanded form at pH 11.5 (Michnik, Michalik *et al.* 2005)

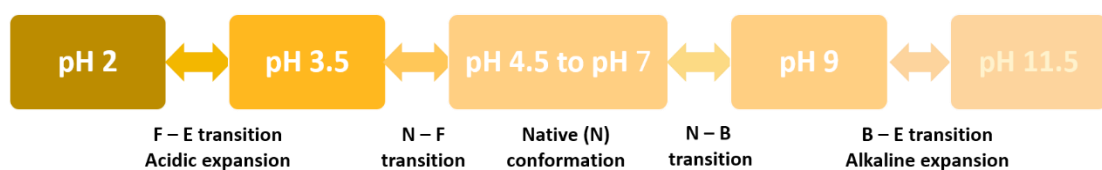


Figure 1.15: BSA structural conformations related to pH changes

1.8 Therapeutic Proteins

Aggregation of bio-therapeutics is a major issue for the advancement of products to market. With numerous monoclonal antibodies already on the market and many more in development, understanding the factors that affect their efficiency and safety is greatly significant (Schellekens 2003; Bhogal 2010; Tovey and Lallemand 2011; Arvinte, Palais *et al.* 2013). Development of stable formulations is of utmost importance for the success of protein drugs. Various factors contribute to the aggregation of therapeutic proteins including temperature, light, solution conditions (pH, buffer type, concentration, and ionic strength), protein concentration, process and purification steps, and storage conditions meaning that control of protein stability is a challenging task for many pharmaceutical companies. Protein aggregation can result in reduced biological activity, but most importantly it can result in immunogenicity and/or cellular toxicities (Demeule, Gurny *et al.* 2006). Compromised immunoglobulin formulations have been shown to cause serious renal failure, and anaphylactoid reactions such as headache, fever and chills (Nydegger and Sturzenegger 1999; Aghamohammadi, Farhoudi *et al.* 2003). Despite the numerous advancements in improving therapeutic formulations, novel drug modifications are also helping to overcome these issues. These can either be genetic or chemical modifications; site-specific mutagenesis, PEGylation, sialylation, lipid based vehicles, hyperglycosylation and protein fusion (Pisal, Kosloski *et al.* 2010). Key steps in all such modifications is preservation of protein activity and stability as all can result in reduction and/or complete loss in protein activity (Roberts and Wang 2010).

1.8.1 Production of therapeutic proteins – monoclonal antibodies

Antibodies (also known as immunoglobulins) are glycoproteins made up of 3 – 12% carbohydrates at conserved *N*-glycosylation sites (Leibiger, Wustner *et al.* 1999), produced naturally in the body by plasma cells in response to an invasion of foreign substances (antigens). They are the body's first natural defence system (Paul 1993; Pathak and Benita 2012). The advent of hybridoma technology by Georges Köhler and Cèsar Milstein in 1975 (Kohler and Milstein 1975) has enabled the production of monoclonal antibodies. Monoclonal antibodies (mAbs) are the fastest growing category of therapeutics entering clinical study, primarily in development for treatment of cancer and immunological disorders (Nelson, Dhimolea *et al.* 2010; Casi and Neri 2012). The use of glycoengineering to produce antibodies with specific glyco forms can increase the

therapeutic efficiency of mAbs (Zheng, Bantog *et al.* 2011). Transfected cell lines are used to synthesize a wide range of compounds that can be extracted and purified for use as biopharmaceuticals. It also makes it possible to generate cells lines that can produce molecules to be easily engineered to optimise their therapeutic properties (Paul 1993). The predominant cell line used for human recombinant protein production is the Chinese Hamster Ovary (CHO) cell line (Chusainow, Yang *et al.* 2009). This cell line has been approved and licensed for the synthesis of therapeutic products including glycoproteins (Hossler, Khattak *et al.* 2009). It is the most commonly used cell line for antibody expression and has been shown to produce glycosylation patterns compatible with the human immune system (Shukla and Thommes 2010).

1.8.2 Antibodies

Antibodies are divided into 5 classes based on their differing structure IgA, IgD, IgE, IgM and IgG (figure 1.16). Generally antibodies are Y-shaped molecules, either singular (monomeric) or in a combination (dimers, pentamers). They can be divided into two regions; the variable (V) region has the defining antigen binding properties, while the constant (C) region interacts with effector cells and molecules. They are divided into classes based on their C regions (Beale and Feinstein 1976).

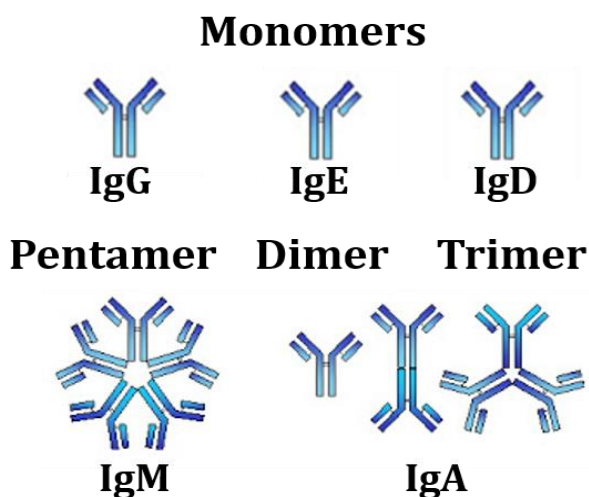


Figure 1.16: Structure of 5 classes of antibodies and their molecular formation.

1.8.3 Immunoglobulin G

Immunoglobulin G (IgG) is the most abundant serum immunoglobulin with concentrations of approximately 10–15 mg/ml and also the majority of marketed mAbs belong to the IgG class (Weiner, Surana *et al.* 2010). They have two identical heavy

chains (H, 50 kDa) and two identical light chains (L, 25 kDa) and can be further subdivided into four different subclasses (IgG1, IgG2, IgG3 and IgG4) depending on the number and location of interchain disulphide bonds and the length of the hinge region of the heavy chain, named $\gamma 1$, $\gamma 2$, $\gamma 3$ and $\gamma 4$. There is also variation in the light chain giving rise to two types- lambda (λ) and kappa (κ). Each IgG molecule is bi-functional. The variable region of both chains mostly consists of the first 110 amino acids and forms the antigen-binding fragment (Fab) that recognizes the respective antigen targets. These can be further sub-divided into three hypervariable sequences (HV1, HV2 and HV3) on both the heavy and light chains and provides the structural basis for the enormous immunological diversity in antibodies. The rest of the sequence consists of the Fc region that works to interact with receptors on effectors cells (Amzel and Poljak 1979; Schroeder and Cavacini 2010; Weiner, Surana *et al.* 2010).

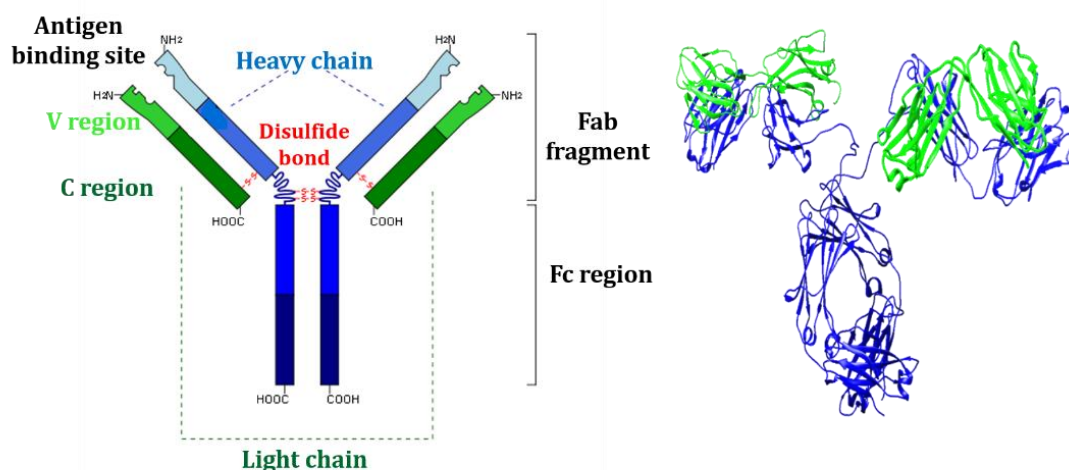


Figure 1.17: Line structure of immunoglobulin G along with crystal structure representation of IgG2a monoclonal antibody as per the protein data bank (light), created using Chimera 1.8.1.

1.8.4 Glycosylation

A glycoprotein is one that contains an oligosaccharide chain covalently attached to the polypeptide side chain. This process, known as glycosylation, is one of the most complex post-translational modifications. There are two major types; the covalent attachment of oligosaccharides through the amide side-chain nitrogen of an asparagine (Asn) residue (*N*-linked), or through the oxygen side chain of serine or threonine (*O*-linked) of a polypeptide chain (figure 1.18).

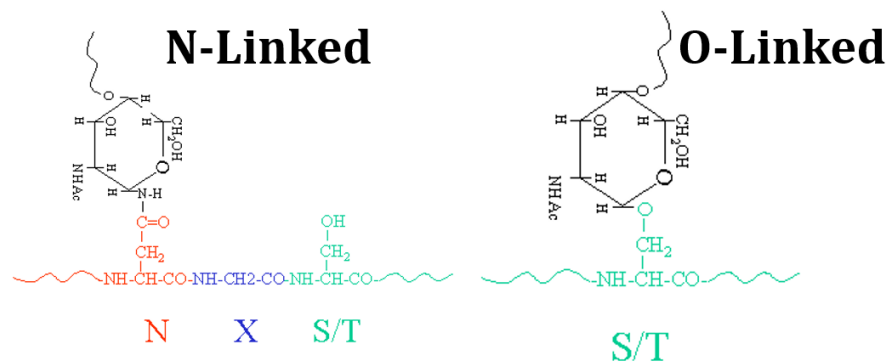


Figure 1.18: Structure of *N*-linked and *O*-linked glycans.

Glycosylation is carried out in a step-wise, highly ordered, site specific manner in the Golgi apparatus, where the activity of each enzyme is dependent upon the completion of the previous enzymatic reaction. The structure of the oligosaccharides (the glycan chain) is determined in part by the conformation of the polypeptide chain in the immediate vicinity of the amino acid residue being processed and the availability of different glycoenzymes to complete the glycan structure (Schwarz and Aebi 2011).

There are five different *N*-glycosidic linkages, with *N*-acetylglucosamine to asparagine (GlcNAc-Asn) being the most common. This linkage was discovered by biochemical analyses of immunoglobulins (Stanley, Schachter *et al.* 2009). Not all asparagine residues can accept an *N*-glycan, the minimal sequence must contain the sequon (Asn-Xaa-Ser/Thr) where Xaa is any amino acid except proline. All *N*-glycans share a mutual core sugar sequence, $\text{Man}\alpha 1-6(\text{Man}\alpha 1-3)\text{Man}\beta 1-4\text{GlcNac}\beta 1-4\text{GlcNac}\beta 1-\text{Asn-Xaa-Ser/Thr}$ and can be classified into three types as shown in Figure 1.19. 1) Oligomannose - only mannose residues are attached at the core, 2) Complex - antennae initiated by *N*-acetylglucosaminyltransferases (GlcNAcTs) are attached at the core and 3) Hybrid - only mannose residues are attached to the $\text{Man}\alpha 1-6$ arm of the core and one or two antennae are on the $\text{Man}\alpha 1-3$ arm. Researchers in this field use pictorial diagrams such as those in Figure 1.17, composed of small geometric symbols to represent individual constituent monosaccharides to aid in describing the structures of many carbohydrates.

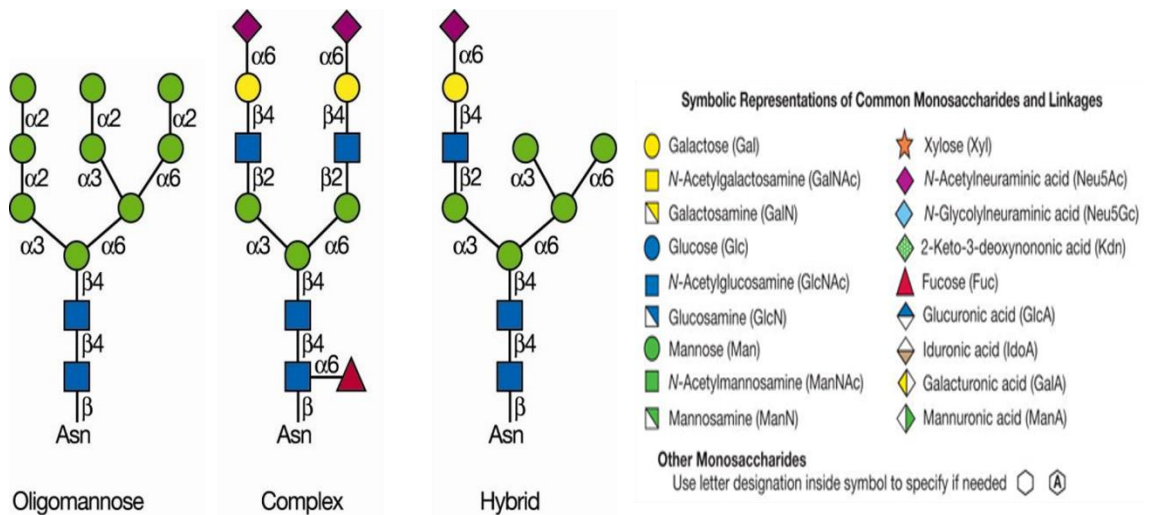


Figure 1.19: Structural diversity in the three classes of *N*-linked glycans (Stanley, Schachter *et al.* 2009). *N*-glycans added to protein at Asn-Xaa-Ser/Thr sequons and each contains the common core $\text{Man}_3\text{GlcNAc}_2\text{Asn}$ which can be then elongated with the addition of other monosaccharides.

1.8.4.1 Glycosylation of immunoglobulins

All immunoglobulins are glycosylated with an inherent set of glycoforms which can differ in number, type and glycosylated site. IgG is the least glycosylated with just two glycan attachment sites, *N*-linked to the conserved Asn297 on both heavy chains (Arnold, Wormald *et al.* 2007). These attachments are ~2 kDa each and help to maintain the structure and stability of the Fc fragment of IgG. This associated sugar domain is not exposed on the surface of IgG, but buried within the hydrophobic core between the two heavy chains, impacting the Fc structure. Each of the Asn-297 sites contains one family of glycans that can be assigned to three subsets, IgG -G0, -G1 and -G2 depending on the number of galactoses present. Figure 1.20 outlines the basic glycan structure found in IgG. The oligosaccharide found in the Fc region is a biantennary complex made up of a core block of sugars (blue boxes) (Asn 297 -GlcNAc -GlcNAc -mannose ($\alpha 1$ -6/ $\alpha 1$ -3) - [mannose -GlcNAc]₂), where GlcNAc is *N*-acetylglucosamine. Variations to this pattern include (red boxes) attachment of terminal sialic acid, a third GlcNAc arm (a bisecting GlcNAc), terminal galactosylation and core fucosylation (Teillaud 2005).

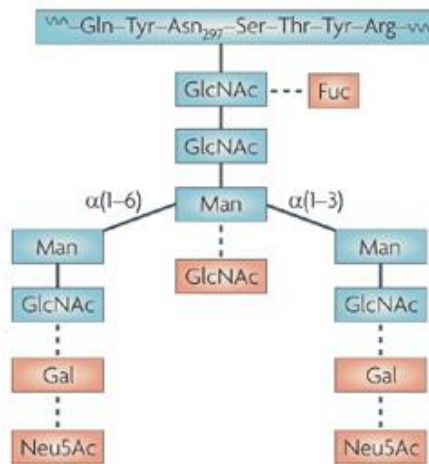


Figure 1.20: The oligosaccharides found in immunoglobulin G (IgG) attached to the Fc fragment of the heavy chain through N-linkage to the Asn297 (Jefferis 2009).

Engineering glycoforms to increase the biological activity of mAbs allows optimisation of their roles and enhances the functional diversity of the proteins. Glycosylation achieves this as every aspect of glycosylation can be modified from the glycosidic linkage, the glycan composition, structure and shape. Figure 1.21 outlines just a few of the glycan variations of the glycoforms of IgG. Each heavy chain may bear one of a total of 32 unique oligosaccharides and random pairing of heavy side glycoforms could generate ~500 glycoforms.

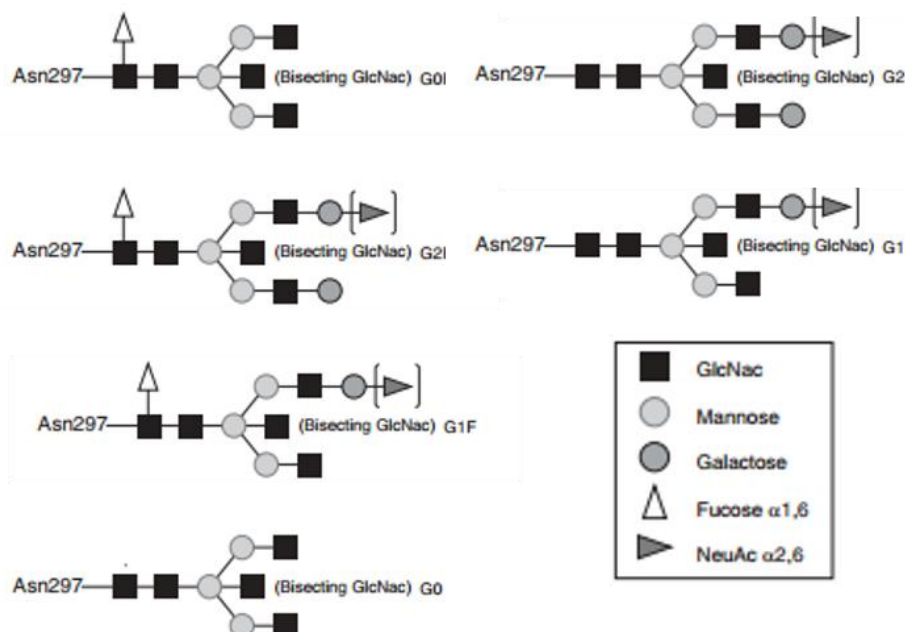


Figure 1.21: Structural variations of the glycoforms of IgG (Teillaud 2005).

The activity, solubility and immunogenicity of monoclonal antibodies can be influenced by glycosylation (Yamada 2011; Costa, Rodrigues *et al.* 2013). Glycan heterogeneity can

be altered by many cell cultural factors including the choice of cell line. The function of a particular glycan can be as important and complex as the function of an amino acid. Thus there is a requirement to monitor the impact cell cultural variables will have in order to ensure consistent glycosylation of monoclonal antibodies (Hossler, Khattak *et al.* 2009).

1.9 Thesis motivation

The motivation for this work was to further our understanding of the mechanisms by which proteins aggregate. Most of the previous studies in this area have been conducted in isolation *e.g.* a specific protein under a specific condition with a specific stress. Here, an inclusive, broad ranging comparative study, using two different proteins, lysozyme and BSA, for a fixed range of solution conditions was conducted. After forced degradation we measured and characterised their response to three specific stresses (thermal unfolding, chemical unfolding and chemical modification) with a collection of high-throughput and orthogonal techniques to meet the following objectives;

1. To understand the molecular interactions governing the aggregation process for each mechanism.
2. To measure the degree of structural change associated with each aggregation pathway.
3. To relate the aggregation behaviour to specific solution conditions.
4. To assess if certain types of aggregate forms (amorphous/fibril) are associated with a specific mechanism.

The second aspect of this work was the isolation and characterisation of a monoclonal antibody, IgG. Specifically, the stability of the glycosylated form of this protein was assessed. The effects of short and long term storage conditions of the glycosylation pattern were also explored.

Chapter 2 Experimental techniques and data analysis

2.1 Preparation of Reagents and Materials

2.1.1 Preparation of buffers

The buffer salts used were analytical grade and were prepared in Milli-Q water (ultra-pure). The pH of the buffer solutions was adjusted using concentrated sodium hydroxide or hydrochloric acid (HCl) solutions as appropriate. Stock buffer solutions were filtered through 0.45 μm , 47 mm nylon membranes using a Millipore filter assembly under vacuum (Merck Millipore). All buffers for SE-HPLC use were of HPLC grade, filtered using the above method and degassed using an online system degasser. All buffers for gel electrophoresis were also filtered as per above and degassed using a vacuum pump.

2.1.1.1 Sodium phosphate

To prepare sodium phosphate buffer at the desired pH, firstly stock solutions of monobasic 1 M NaH_2PO_4 and dibasic 1 M Na_2HPO_4 were prepared and mixed in appropriate volumes as per Table 2.1 to achieve the desired pH. To prepare the stock solutions, 11.998 g of NaH_2PO_4 (Mw = 119.98) and 14.196 g of Na_2HPO_4 (Mw = 141.96) were each dissolved in Milli-Q water and made up to 100 ml.

| <i>pH</i> | <i>Vol (ml) of 1 M</i> | | <i>Final vol (ml)</i> |
|-----------|--------------------------------------|--------------------------------------|-----------------------|
| | <i>NaH₂PO₄</i> | <i>Na₂HPO₄</i> | |
| 5.8 | 8.0 | 92.0 | 100 |
| 6.0 | 12.3 | 87.7 | 100 |
| 7.0 | 61.0 | 3.9 | 100 |
| 7.8 | 91.5 | 8.5 | 100 |

Table 2.1: Sodium phosphate pH guide.

The 1 M solutions were then diluted as necessary *i.e.* to make a 50 mM, 1L solution, 50 ml of the 1M stock was diluted into 950 ml of Milli-Q water.

2.1.1.2 Tris-HCl

To prepare a 1 M Tris-HCl buffer stock solution, appropriate weight of Tris solid (Fisher) was dissolved in Milli-Q water. The pH was adjusted using HCl.

2.1.1.3 10% SDS

To prepare a 10% SDS stock solution, 1 g of SDS solid (Fisher) was dissolved in 10 ml Milli-Q water in a volumetric flask.

2.1.2 Preparation of reagents.

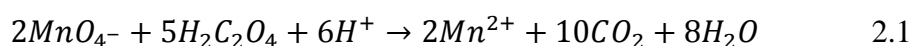
2.1.2.1 Preparation of denaturants

Urea (U) and Guanidine hydrochloride (GdmHCl) [Calbiochem] were made up as stock solutions to a maximum concentration of 11 M and 8 M respectively. Denaturants were fully dissolved in appropriate buffer in volumetric flasks and prepared directly before use.

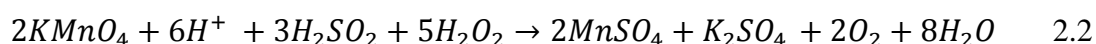
2.1.2.2 Standardizing hydrogen peroxide

Hydrogen peroxide (Calbiochem) was purchased as a 30% (v/v) solution. Hydrogen peroxide decomposes over time into water and oxygen. Before each use, H₂O₂ was standardised against potassium permanganate (KMnO₄) to obtain the absolute concentration by titration against a known concentration of KMnO₄.

Firstly, a 2 mM solution KMnO₄ was made up in Milli-Q water (0.03 g in 100 ml). Solid KMnO₄ contains traces of other impurities, so to determine the exact concentration of the KMnO₄, it was titrated against oxalic acid (H₂C₂O₄) in the presence of excess 6 M sulphuric acid (H₂SO₄). By using the stoichiometry of the standardization reaction, the exact concentration of the KMnO₄ solution can be determined;



The standardised solution of KMnO₄ was then titrated against the unknown concentration of H₂O₂, in the presence of excess 6 M sulphuric acid (H₂SO₂). In acidic solutions, potassium permanganate reacts with hydrogen peroxide as follows;



Once a permeant pink colour change was detected the titration was stopped and the molar concentration of H₂O₂ was calculated. Hydrogen peroxide was standardised in this way directly before each use.

2.1.3 Preparation of protein stock solutions

Chicken egg white lysozyme [Calbiochem, Lot # D00140217] and bovine serum albumin [Sigma Aldrich, Lot # SLBL2871V] were used without further purification. Protein was dissolved in sodium phosphate buffer, 50 mM pH 5 – 8 or sodium acetate buffer pH 5 – 8, then washed of co-precipitated salts by repeated ultrafiltration using Amicon Ultra-4 MWCO 10 kDa (lysozyme) or 30 kDa (BSA) centrifugal devices at $4000 \times g$. Both lysozyme and BSA concentrations were determined spectrometrically by UV/Vis absorbance using the extinction coefficients, $2.64 \text{ ml mg}^{-1} \text{ cm}^{-1}$ and $0.66 \text{ ml mg}^{-1} \text{ cm}^{-1}$ respectively. [See section 2.5]

2.2 Mechanisms of aggregation

2.2.1 Thermal denaturation

Thermal denaturation is achieved by heating the protein solution beyond its melt transition temperature T_m . Upon heating and cooling of the protein, it begins to unfold and refold. The overall structural stability of the protein solution during this process is dependent on many factors such as protein concentration, pH, ionic strength, presence of co-solutes and ion type. The degree of structural change caused by thermal denaturation can very much depend on the solution conditions (Wang, Nema *et al.* 2010).

2.2.1.1 Temperature induced unfolding

A thermal denaturation assay was developed to suit a 96 well plate format. Prepared protein samples were aliquoted in triplicate into a 96 well plate. Solution conditions were varied including, protein concentration (0.5 – 20 mg/ml), buffer type (50 mM sodium phosphate) and pH of buffer solution (range from pH 5 – 8). Stock solutions of protein were prepared and diluted to the desired concentration using buffer to a total volume of 200 μl , below the maximum well volume of 300 μl . The multi-well plate was covered using sealing tape (Thermo Scientific) and a plate lid. The proteins were thermally denatured by incubation at 80°C , (beyond the T_m of the protein) for 20 minutes to allow sufficient time for the protein to fully denature and then cooled on the bench at room temperature (20 – 24°C).

2.2.2 Oxidation

The hydroxyl radical ($\bullet\text{OH}$) is the most powerful oxidizing species among several reactive oxygen radicals (ROS), and is known to oxidize proteins (Shacter 2000). Therefore hydrogen peroxide (H_2O_2) was chosen to act as a ROS to induce oxidation and the associated conformational rearrangements within the protein structure.

2.2.2.1 Oxidant induced de-stabilisation of protein

Stock protein solutions were prepared and diluted in 50 mM sodium phosphate buffer and hydrogen peroxide to give the final protein concentration of 10 mg/ml. (BSA was examined in pH 7 buffer, while lysozyme was examined over the pH range 5 – 8.) H_2O_2 ranging from 0.01 mM to 5 M was added to the protein solutions to a total volume of 300 μl . H_2O_2 is sensitive to light and can readily decompose to water and oxygen. Therefore protein solutions after addition of H_2O_2 were stored in amber glass vials sealed with parafilm and wrapped in aluminium foil. Glass syringes were used to transfer H_2O_2 solutions to avoid leaching from plastic pipette tips which may have impacted on the aggregation study. Identical protein solutions not treated with H_2O_2 were also prepared under the same conditions to be used as a reference.

2.2.3 Chemical unfolding

Chemical denaturants (urea and guanidine hydrochloride) were used to induce protein unfolding. Stock protein solutions of lysozyme (0.5 mg/ml) and BSA (1 mg/ml) were prepared in 50 mM sodium phosphate buffer, pH 8. In microcentrifuge tubes, the stock protein solutions were diluted with appropriate amounts of buffer and denaturant. The total volume was fixed at 1 ml in every sample. The concentration of denaturant ranged from 0.5 – 10.5 M for urea, and 0.5 – 7.5 M for guanidine hydrochloride. Protein samples without denaturant at the same protein concentration and final sample volume were used as a control.

2.3 SE-High Performance Liquid Chromatography

2.3.1 Description of the technique

High performance liquid chromatography (HPLC) is an analytical technique and physical separation method in which the components to be separated are selectively distributed

between two immiscible phases. The liquid mobile phase carrying the molecules to be separated is forced under pressure through a stationary phase bed, held in a column. The size exclusion HPLC (SE-HPLC) principle of separation is based on molecular size (figure 2.1). Larger molecules will move more quickly through the packing material of the column and elute first, while the smaller molecules will take longer as they are slowed down by passing through the pores in the packing material. The time taken for a molecule to travel through the column and reach the detector is known as the retention time (t_R) (Hamilton and Sewell 1982).

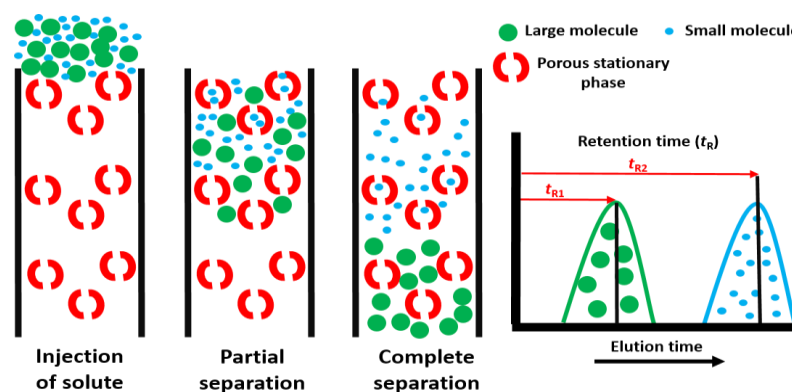


Figure 2.1: Schematic diagram of the three phases of SE-HPLC. A representative chromatogram is shown indicating two elution peaks (one for each particle size) and their corresponding retention times (t_R).

In SE-HPLC, the stationary phase generally consists of gel particles having a porous matrix with a closely controlled pore size. A schematic of the main components of required for an effective SE-HPLC system is shown in Figure 2.2.

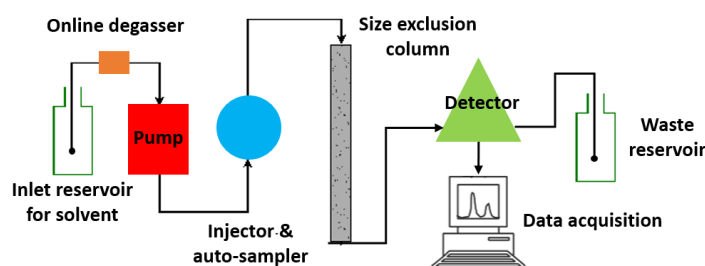


Figure 2.2: Schematic of components of SE- HPLC system

2.3.2 Instrumentation

SE-HPLC was carried out using a Shimadzu SPD HPLC system. Shimadzu LC solution software was used to run the instrument. Before each use the HPLC system has all its lines purged of trapped air or gas bubbles to ensure no damage to the system or column. A steady baseline was achieved by running mobile phase at 0.75 ml/min for 30 – 60

minutes before a sample injection. The column used was a Superdex 200 10/300 GL SE-HPLC column (GE Healthcare). This column allows for high resolution analytical separations of protein molecules with molecular weights between 10 and 600 kDa. Protein was detected using a diode array detector (DAD) at 190 – 800 nm for each measurement.

2.3.3 Running a sample

For HPLC, samples were placed in Chromacol glass vials (Lennox) with a crimp cap with a pre-fitted seal. Samples were introduced to the column via the instrument autosampler. Typical conditions used to run a SE-HPLC sample are given in Table 2.3. All SE-HPLC measurements were carried out at room temperature.

SE-HPLC sample run

| | |
|--------------------------------|----------------------------------|
| <i>Mobile phase</i> | 50 mM sodium phosphate, pH 5 - 8 |
| <i>Pump flow rate</i> | 0.75 ml/min |
| <i>Analysis run time</i> | 45 minutes |
| <i>Sample injection volume</i> | 20 – 30 μ l |
| <i>Sample concentration</i> | 0.5 – 20 mg/ml |
| <i>Pressure</i> | ~ 11 bar |

Table 2.3: Typical conditions used to run a protein sample using Superdex 200 10/300 GL column.

2.3.4 Calibration of the column

A mixture of proteins with different molecular weights, within the resolution range of the column were prepared. Each protein was dissolved in 50 mM sodium phosphate buffer and filtered through 0.22 μ m syringe driven filters (Milli-pore Durapore) to remove particulates. All proteins were run individually (1 mg/ml, sample injection volume 20 μ l, flow rate 0.75 ml/minute) to obtain a characteristic chromatogram for that protein and assign a retention time as per Table 2.4.

| <i>Peak #</i> | <i>Protein</i> | <i>tr</i> | <i>M W</i> | <i>Conc. mg/ml</i> | <i>Injection vol. (μl)</i> |
|---------------|----------------------------|-----------|------------|--------------------|---|
| 1 | Gama Globulin dimer | 10.40 | 300 | | |
| 2 | Gamma globulin | 13.20 | 150 | 1 | 20 |
| 3 | Bovine serum albumin dimer | 14.09 | 133.4 | | |
| 4 | Bovine serum albumin | 16.30 | 66.7 | 1 | 20 |
| 5 | Chicken serum albumin | 17.53 | 44.2 | 1 | 20 |
| 6 | Myglobulin | 21.06 | 16.7 | 1 | 20 |
| 7 | Lysozyme | 23.29 | 14.7 | 1 | 20 |

Table 2.4: Proteins used for the calibration of Superdex 200 10/300 GL column.

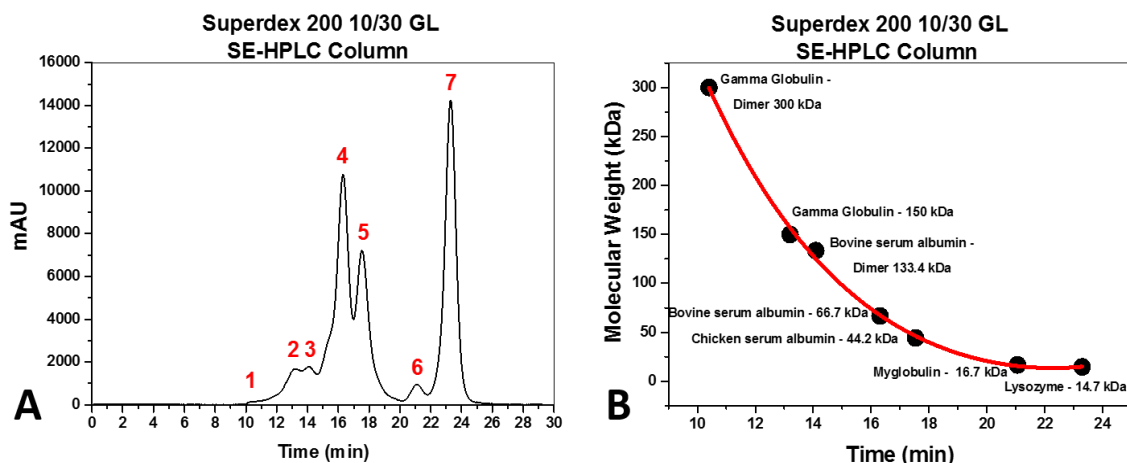


Figure 2.3: Calibration of Superdex 200 column using a range of different molecular weight proteins.

A mixture of all proteins was then prepared keeping each individual protein concentration at 1 mg/ml; 200 μ l of each protein sample at 5 mg/ml, made up to a total mixture volume of 1 ml. 20 μ l of this protein mixture was then loaded on to the column (figure 2.3A). A calibration curve was then created (figure 2.3B).

2.3.5 Determination of protein concentration by SE-HPLC

Calibration curves of protein over a range of concentrations were prepared allowing conversion of the area under a chromatographic peak to a protein concentration (figure 2.4). Protein samples were prepared at concentrations from 0.5 – 20 mg/ml and the concentration was confirmed by UV absorbance. The sample volume was 20 μ l in each case and detection was at 280 nm. Using OriginPro 9.1 software, the chromatogram at each concentration was baseline subtracted and then the area under all peaks were integrated to give a total peak area per mg/ml of protein.

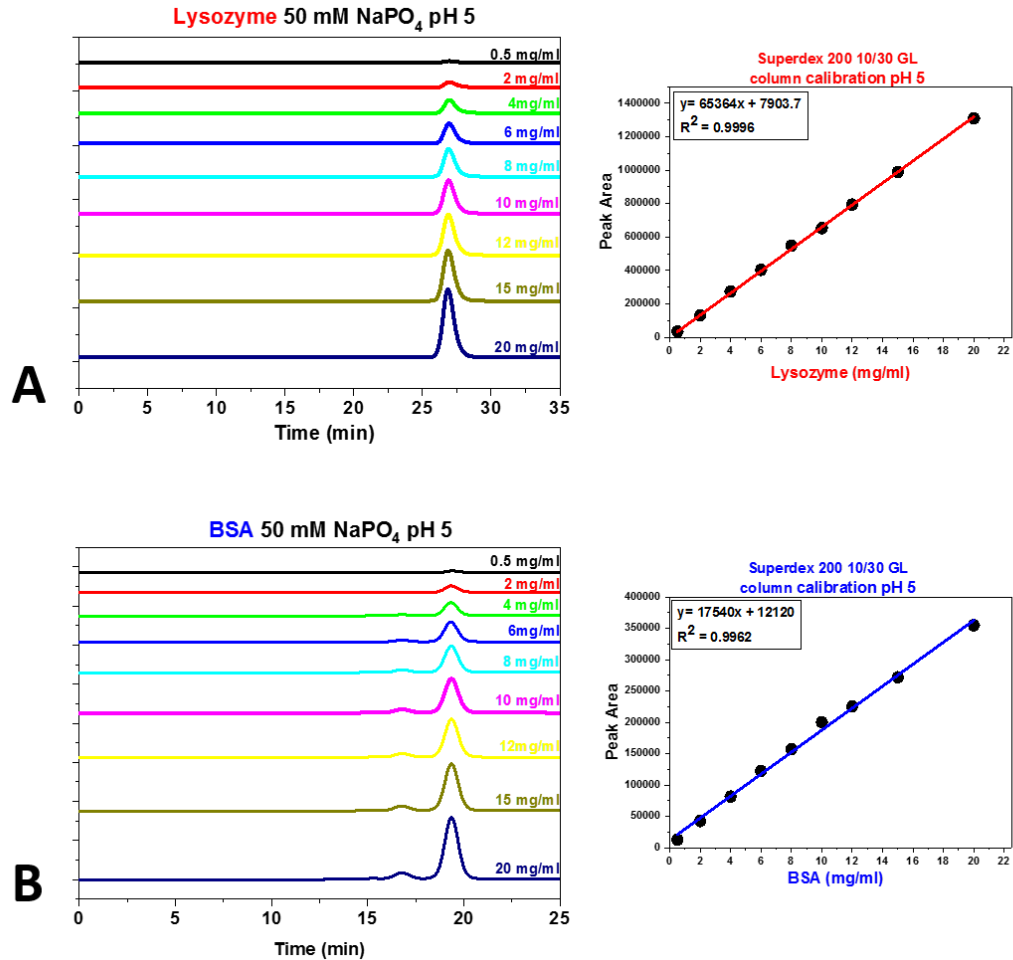


Figure 2.4: SE-HPLC for A) lysozyme and B) BSA at increasing concentrations. Relationship between peak area and protein concentration of each is also shown.

2.4 Electrophoresis

2.4.1 Description of the technique

Electrophoresis is an analytical method used to separate, identify and purify biomolecules. Under the influence of an electric field, charged molecules will migrate in the direction of the electrode bearing an opposite charge. Due to the varying charges and masses on the molecules, they will migrate at different rates and thus be separated into single fractions. This technique is typically used for proteins, peptides, sugars and nucleic acids. It can be used for qualitative characterisation of samples, quantitative determination and assessment of purity, and preparative purposes (Hames and Rickwood 1981; Westermeier 2001). The speed at which a protein will migrate in an electric field will be characteristic of that protein under defined conditions (pH, ionic strength,

temperature and voltage). The electrophoretic mobility (μ) of a protein molecule can be defined as;

$$\mu = \frac{d}{Et} \quad 2.3$$

where d is the distance travelled from the origin (cm), E is the strength of the electrical field (V/cm) and t is the duration of electrophoresis (seconds) (Keren 2003).

Polyacrylamide gel was first used for electrophoresis by Raymond and Weintraub (Raymond and Weintraub 1959). PAGE gels consist of an electrically neutral matrix of pores to separate proteins according to size (molecular weight) and conformation. They can resolve proteins between 5 and 400 kDa. The clear gel is formed by chemical co-polymerisation of acrylamide monomers with a cross-linking reagent, usually bisacrylamide. The reaction is started with ammonium persulphate (APS) acting as a polymerising agent with TEMED (N,N,N,N'-tetramethylethylenediamine) providing the tertiary amino groups to release the free radicals by APS (Westermeier 2001). The pore size of the gels is inversely related to the amount of acrylamide used and can be controlled by the amount of acrylamide (T) and the degree of cross-linking (C)

$$T = \frac{(a+b) \times 100}{V} [\%], \quad C = \frac{b \times 100}{a+b} [\%] \quad 2.4$$

where a is the mass of acrylamide, b is the mass of cross-linker and V is the volume (Westermeier 2001).

2.4.2 Types of PAGE electrophoresis

2.4.2.1 Native polyacrylamide gels

Native or non-denaturing polyacrylamide gels are used for separation and characterisation of protein depending on the net charge, size and shape of the protein in its native state. Samples are prepared in a non-reducing, non-denaturing sample buffer, which maintains the protein secondary structure and native charge density. Proteins migrate through the gel matrix as the majority will have a net negative charge in the alkaline buffer used. The higher the negative charge density, the faster the protein will migrate. The frictional force of the gel matrix also acts like a sieve, restricting the mobility of proteins by their size and shape (Westermeier 2001).

2.4.2.2 SDS polyacrylamide gels

SDS-PAGE is a method of separating proteins exclusively on the basis of molecular weight. The structural complexity of the protein is reduced and a uniform charge to mass ratio is applied. Sodium dodecyl sulphate (SDS) is an anionic detergent that binds to most proteins in constant weight ratio, giving it an overall negative charge density. The tertiary and secondary structures are also disrupted through the breakage of hydrogen bonds causing the protein to unfold. This type of gel is commonly referred to as a “non-reducing” SDS-PAGE. “Reducing” SDS-PAGE involves breakage of the disulphide bonds between the cysteine residues of the protein. This is achieved by a thiol reducing agent such as 2-mercaptoethanol or dithiothreitol. The protein migrates in the electric field towards the positive anode, separating according to their size with the smaller proteins migrating faster (further) than the larger protein (Westermeier 2001).

2.4.3 Protocol

2.4.3.1 PAGE

A 4-gel electrophoresis tank (Bio-Rad) was used for all separations. Gel plates rinsed with Milli-Q water, followed by ethanol and air-dried. Once dried, plates were assembled as per manual. Gel solutions were prepared with the following components:

Resolving Gel (12%)

Milli-Q water

Acryl amide: Bis solution (37.5:1) (Bio-Rad)

1 M Tris-HCL pH 8.8

10% SDS (w/v)

Mixed and degased for 15 minutes

10% (w/v) ammonium persulfate (prepared fresh, added just before use)

TEMED (added just before use)

Stacking Gel (4%)

Milli-Q water

Acryl amide: Bis solution (37.5:1)

0.5 M Tris-HCL pH 6.8

10% SDS

Mixed and degased for 15 minutes

10% (w/v) ammonium persulfate (prepared fresh, added just before use)

TEMED (added just before use)

Sample buffer

Milli-Q water
0.5 M Tris-HCL (pH depending on sample)
Glycerol (0.8 ml)
10% SDS (w/v)
Bromophenol Blue (0.5 mg)

Running buffer

10X Tris-glycine (Bio-Rad, diluted 1/10)

Protein Molecular weight markers

Molecular weight markers were purchased from Bio-Rad (figure 2.5). Markers were diluted in SDS reducing sample buffer and stored at -20°C . Markers were thawed and boiled at 95°C just before use.

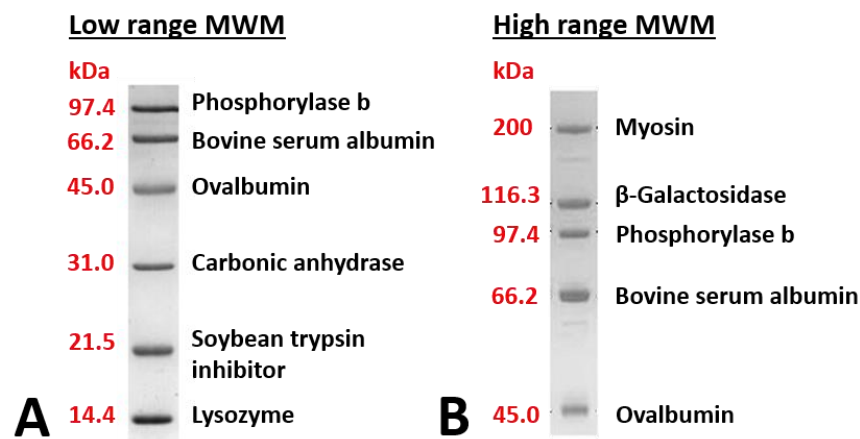


Figure 2.5: Bio-Rad molecular weight markers. A) Low range and B) High range.

Gel preparation

Appropriate amounts of both separating gel and stacking gel were prepared as described above. Just before casting the gel, the ammonium persulfate and TEMED were gently mixed into the gel mixture. Using Pasteur pipettes the separating gel solution was slowly pipetted into the gap between the glass casting plates, to about 2/3 full. A small amount of tert-amyl alcohol was used to cover the gel solution while the gel was setting. The gel was left for 40 – 45 minutes to set. Once set, the tert-amyl alcohol was washed off the gel with Milli-Q water and residual water was wiped away. The stacking gel was then added and the well combs inserted. A further 40 – 45 minutes was allowed for the stacking gel to set before running samples.

Sample preparation

Samples were prepared by mixing appropriate amounts of sample and sample buffer. The maximum loading capacity for each well was 30 µl. Samples were heated to 90 - 95°C for 5 minutes (for SDS PAGE only).

Running the gel

Gels were run using an external power source at 200 V (400 mA) until the visual running band was 1 cm above the end of the glass plates.

Gel staining

Gels were stained for 1+ hour in Coomassie Brilliant Blue R-250 staining solution (Biorad). Gels were then de-stained in a mixture of Milli-Q water, methanol and acetic acid in a ratio 50/40/10 (v/v/v). The protein molecular weight was determined by comparing the migration of the protein band to the molecular weight standards.

2.4.4 Modifications to the PAGE protocol

The protocol described above is an example of a non-reducing SDS 12% gel. This type of gel is ideal for identifying the molecular weight of a protein. Modifications to this protocol were needed for a number of different reasons.

2.4.4.1 Native PAGE

For native PAGE the protein samples were prepared without the addition of reducing or denaturing sample buffer and without heat treatment. The electrophoresis was also run without SDS in the running buffer.

2.4.4.2 Reducing SDS-PAGE

Reducing gels were prepared by adding 2-mercaptoethanol to SDS sample buffer just before use allowing for identification of disulphide linked aggregates or reduction of proteins that were disulphide inked.

2.4.4.3 Gel concentration

By varying the percentage of cross-linker, the pore size of native and SDS-PAGE gels were optimised to yield the best separation and resolution of the protein of interest. The gel concentration was adjusted according to Table 2.5 by varying the amount of cross-linker in the resolving gel solution.

| <i>Acrylamide %</i> | <i>MW Range</i> | <i>Vol Acrylamide (ml)</i> | <i>Vol Milli Q water (ml)</i> |
|---------------------|-----------------|----------------------------|-------------------------------|
| <i>7%</i> | 50 – 500 kDa | 2.3 | 5.1 |
| <i>10%</i> | 20 – 300 kDa | 3.3 | 4.1 |
| <i>12%</i> | 10 – 200 kDa | 4 | 3.4 |
| <i>15%</i> | 3 – 100 kDa | 5 | 2.4 |

Table 2.5: *Optimum acrylamide gel composition for the resolution of different Mw proteins. Volume of acrylamide and Milli-Q water needed to achieve each gel composition (keeping the total resolving gel solution at 10 ml).*

2.5 Spectroscopy

Spectroscopy deals with the production, measurement, and interpretation of spectra arising from the interaction of radiation with matter. When radiation interacts with matter a number of processes can occur; absorbance, fluorescence/phosphorescence, reflection and scattering. Radiation is characterised by energy, E , that is related to the frequency, ν , or wavelength, λ , of the radiation by the Planck relationship;

$$E = h\nu = hc/\lambda \quad 2.5$$

where c is the speed of light (2.998×10^8 cm/s) and h is Planck's constant (6.625×10^{-27} J-s).

2.5.1 Absorption spectroscopy

The absorbance of light is due to the interaction of an oscillating electromagnetic field of radiation with the charged particles in an atom (Hammes 2005). The principle is based on the ability of a molecule to absorb photons of light. When UV-visible light is absorbed by a molecule, energy from the light can promote an electron from a bonding or non-bonding orbital into one of the empty anti-bonding orbitals. The process of absorption is extremely fast occurring in about 10^{-15} s. Electrons in the excited state equilibrate with the surroundings and returns to the ground electronic state in a decay process. This occurs with either the production of heat (non-radiative decay) or the emission of light (radiative decay) with an energy equal to or less than that of the absorbed radiation (figure 2.6). This process is known as fluorescence and will be discussed in more detail in section 2.5.2.

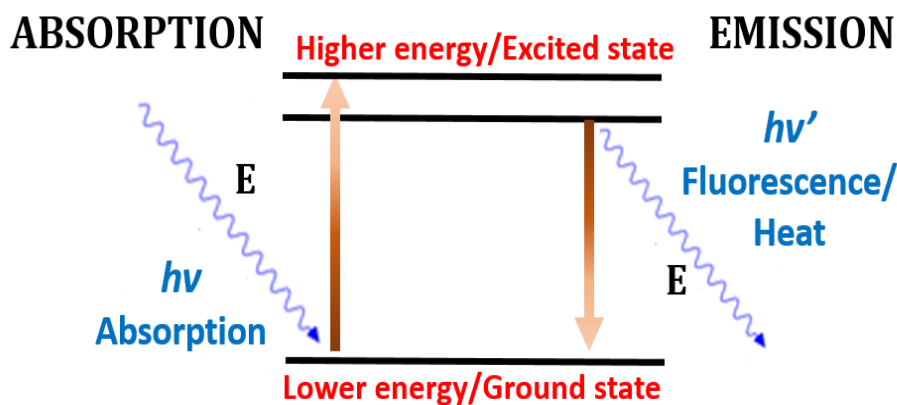


Figure 2.6: Absorption and emission of light.

The potential energy of a molecule can be represented as the sum of its electronic, vibrational and rotational energies;

$$E_{total} = E_{electronic} + E_{vibrational} + E_{rotational} \quad 2.6$$

The amino acids that make up the backbone of protein are connected by peptide bonds. This gives rise to two characteristic absorption bands in the far-UV region ($\lambda = 190 - 250$ nm): $\pi \rightarrow \pi^*$ transition at 190 nm and a weaker, broader $n \rightarrow \pi^*$ transition at 210 nm. A π orbital is formed between two p atomic orbitals overlapping laterally, resulting in a π bond. A π electron can be promoted to an antibonding orbital, π^* , with the absorption of sufficient energy. This is known as a $\pi \rightarrow \pi^*$ transition. Non-bonding electrons, can be promoted to an antibonding orbital and this transition is denoted by $n \rightarrow \pi^*$ transition. A σ orbital can also be formed between two s atomic orbitals, between two p atomic orbitals or from one s and one p atomic orbital, forming a σ bond (Valeur 2002). Figure 2.7 shows the different types of electronic transitions and their energy is generally in the order;

$$n \rightarrow \pi^* < \pi \rightarrow \pi^* < n \rightarrow \sigma^* < \pi \rightarrow \sigma^* < \sigma \rightarrow \pi^* < \sigma \rightarrow \sigma^* \quad 2.7$$

Of these six transitions, only the two of lowest energy occur in absorbance and fluorescence spectroscopy (200 – 800 nm spectrum) for small molecules.

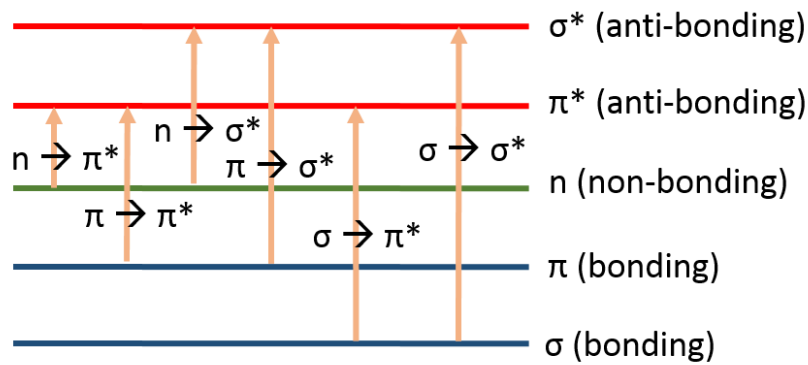


Figure 2.7: Electronic transitions and their corresponding orbitals.

The ability of a protein to absorb UV-visible light is based on the presence of chromophores such as amide bonds in the peptide backbone and the aromatic amino acids. For proteins, there are three amino acid residues that absorb light in the UV region: phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) due to their conjugated π bonds in the non-polar side chains (figure 2.8).

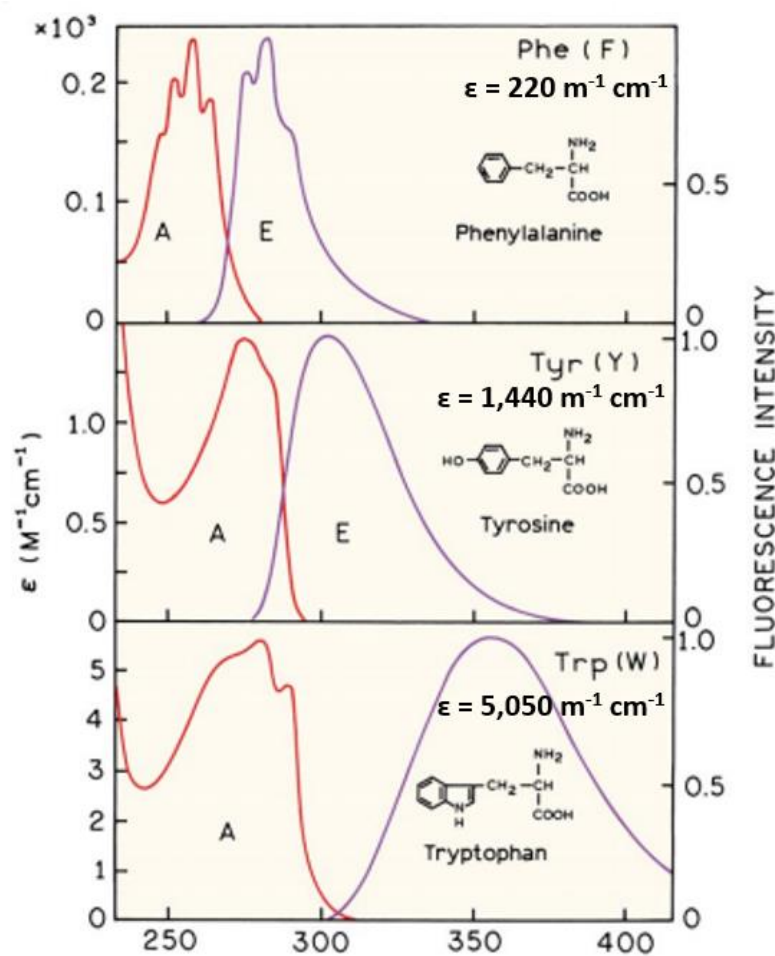


Figure 2.8: Aromatic residues found in protein: Phenylalanine, tryptophan and tyrosine contain aromatic side chains, which allow them to absorb UV light (A) and emit fluorescence (E) from (Lakowicz 2006).

Absorbance spectroscopy is a commonly used tool to estimate protein concentration, according to the Beer-Lambert equation;

$$A = \epsilon cl \quad 2.8$$

where A is absorbance, ϵ is the extinction co-efficient of the protein ($1 \text{ dm}^3 \text{ mol}^{-1} \text{ cm}^{-1}$), c is the concentration of the protein (mol dm^{-3}) and l is the path length of light (1 cm).

2.5.1.1 UV-Vis measurement

UV measurements were performed on a Perkin Elmer Lambda 35 UV/VIS spectrometer using Hellma Analytics precision quartz cells with a 1 cm path length. Spectra were collected between 200 – 500 nm at the scan speed 240 nm/min with a data interval 0.1 nm. Air was used as a reference sample. Spectra for buffer solutions were acquired in the same way and manually subtracted from the protein spectra.

2.5.2 Fluorescence Spectroscopy

Fluorescence is the emission of a photon associated with transition of a molecule from a higher excited state to a lower ground state accompanied by radiation. When light of an appropriate energy is absorbed by a molecule with electrons of opposite spin, the electronic state of the molecule changes from the ground state singlet energy level, S_0 , to one of many vibrational levels in one of the excited states ($S_1, S_2 \dots$). All electrons are paired in singlet states so that their spins are in opposite directions. This is known as a singlet-singlet transition. Decay can occur via numerous process; non-radiative decay or radiative decay. Non-radiative decay occurs to the ground vibrational energy level, S_1 , generally in the form of heat. Electrons upon returning to the ground state, S_0 , via radiative decay from the first excited singlet state emit fluorescence. They can also move to a triplet state ($T_1, T_2 \dots$) having two unpaired electrons, known as intersystem crossing. The process of electrons returning to the ground state, S_0 , from a triplet state is known as phosphorescence. This is generally shifted to longer wavelengths (lower energy) relative to fluorescence.

Within each electronic state there are also vibrational energy levels more closely spaced than the electronic energy levels, due to the vibration of atoms within the molecule. Molecules are generally in their ground electronic state are in their lowest vibrational

energy level. Once excited to the next electronic energy level, they can be found in many different vibrational energy levels. The molecules rapidly decay (10^{-13} s or less) to the lowest vibrational level of the second electronic energy level by non-radiative decay. Electrons in this level will then decay to the ground state. The Perrin-Jablonski diagram (figure 2.9) is a useful way of visualising these processes. (Valeur 2002; Hammes 2005; Lakowicz 2006).

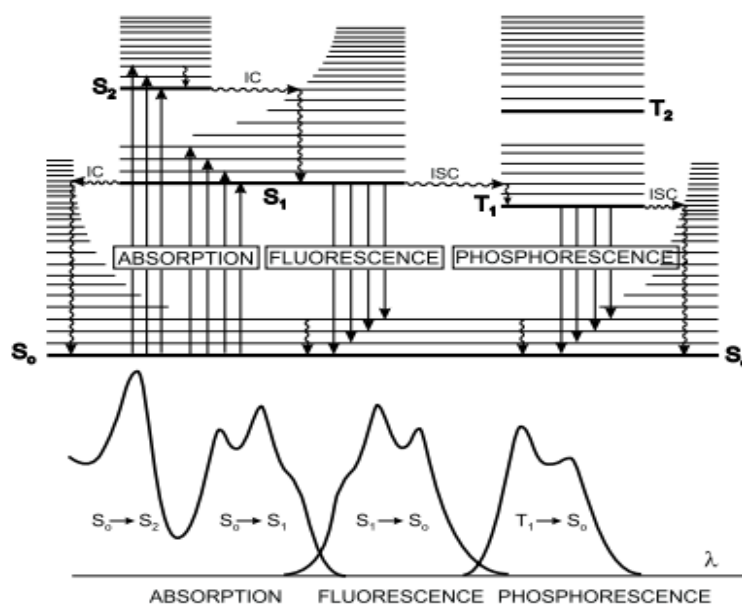


Figure 2.9: Perin-Jablonski diagram and illustration of the relative positions of absorption, fluorescence and phosphorescence spectra (Valeur 2002).

Fluorescence data is mostly presented as emission spectra. The emission spectrum depends on the chemical structure of the fluorophore, the surrounding environment and the wavelength of excitation (Lakowicz 2006). There are two classes of fluorophores; extrinsic fluorophores are added to a sample that doesn't display the desired spectral properties and intrinsic fluorophores that occur naturally. Proteins are examples of intrinsically fluorescent molecules. The fluorescence of proteins originates from the aromatic residues; phenylalanine, tyrosine, and tryptophan residues (figure 2.8). However fluorescence is dominated by the contribution of tryptophan and tyrosine residues as their absorbance at wavelength excitation and their quantum yield of emission are considerably greater than for phenylalanine. Emission of both tyrosine and tryptophan residues is observed when proteins are excited at 280 nm. Excitation at 295 nm results in an emission spectrum for tryptophan only. The emission spectra of intrinsic fluorophores are highly sensitive to the solvent/environmental conditions.

2.5.2.1 Fluorescence Measurements

Intrinsic fluorescence measurements were conducted using Molecular Devices SpectraMax M2e (Molecular Devices, USA) with Hellma Analytics fluorescent precision quartz cells. Protein solutions were excited at 280 nm (excitation of tyrosine and tryptophan residues) and 295 nm (tryptophan excitation mainly) and fluorescence emission spectra were recorded in the range 200 – 500 nm with a data interval of 1 nm. The blank was a buffer solution, which was manually subtracted from the protein spectra. From each fluorescence spectrum both the fluorescence intensity and lambda max (λ_{\max}) of fluorescence were recorded and used for analysis.

2.5.3 Derivative spectroscopy

Derivative spectroscopy is useful for extracting both qualitative and quantitative information from spectra composed of unresolved bands. It allows for the enhancement of the resolution of overlapping peaks, and the elimination or reduction of background or matrix absorption (Aligent Technologies, 2000 ; Rojas, Ojeda *et al.* 1988; Ojeda and Rojas 2004). A single absorbance band with a broad maximum and unresolved individual features occurs for proteins, due the overlap in absorbance in the 280 nm region for the aromatic three amino acids; tyrosine, tryptophan and phenylalanine. The differentiation of zero order spectra allows more spectral detail to be observed and results in more accurate determination of the wavelengths of broad peak maximum, for peaks which appear only as shoulders, as well as for the isolation of small peaks from an interfering large background absorption (Grum, Paine *et al.* 1972) (figure 2.10). Finite differences between adjacent data points can be determined by calculating the derivative of a spectrum (Saakov, Drapkin *et al.* 2012). A typical spectrum is expressed as absorbance A as a function of wavelength λ . The derivative of absorbance spectra are:

$$\text{Zero order} \quad A = f(\lambda) \quad 2.9$$

$$\text{First order} \quad \frac{dA}{d\lambda} = f'(\lambda) \quad 2.10$$

$$\text{Second order} \quad \frac{d^2A}{d\lambda^2} = f''(\lambda) \quad 2.11$$

The first derivative spectrum starts and ends at zero, it passes through zero at the same wavelength as λ_{\max} of the absorbance band. The second derivative spectrum reveals a

negative band that crosses the x-axis at the point of maximum slope and with a minimum the same wavelength as the maximum on the zero-order band which is a significantly sharper band (Mark and Workman 2007).

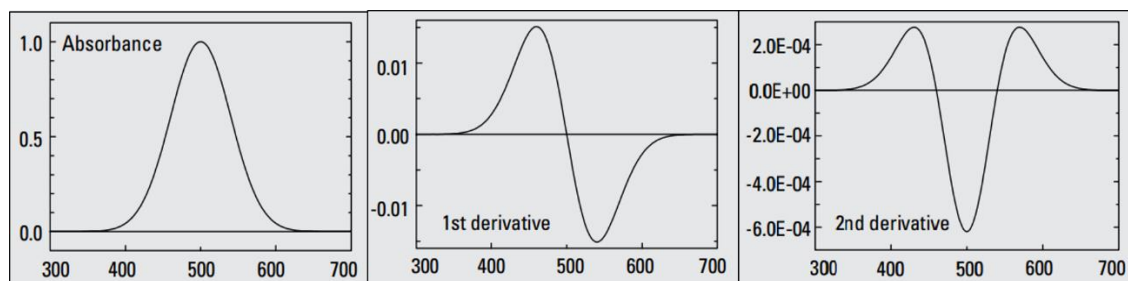


Figure 2.10: Simulation of the effects of first order and second order derivatisation on the absorbance spectra (Aligent Technologies, 2000).

Differentiation progressively increases the signal-to-noise ratio (S/N) of the original zero order spectrum with higher orders (Rojas, Ojeda *et al.* 1988). It is important to remove or significantly minimize the S/N, yet maintaining the underlying spectral data. To achieve this a certain amount of mathematical smoothing must be performed either before or after differentiation. This method of spectral averaging results in the reduction in random noise (figure 2.11). The Savitzky-Golay smoothing filters are the most frequently used digital smoothing filters in spectroscopy. The Savitzky-Golay filter method essentially performs a local polynomial regression to determine the smoothed value for each data point. This method is superior to adjacent averaging because it tends to preserve features of the data such as peak height and width, which are usually 'washed out' by adjacent averaging (Chen, Liu *et al.* 2004).

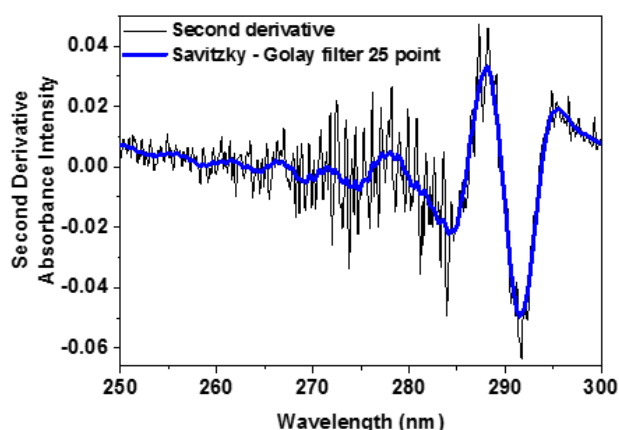


Figure 2.11: Smoothing of a second derivative absorbance spectra of BSA using the Savitzky – Golay filter method.

2.5.3.1 Second derivative absorbance spectroscopy measurement

Figure 2.12 illustrates how second derivative operations on the zero order spectrum allows deconvolution of this single band into distinct peaks representing each individual aromatic amino acid residue.

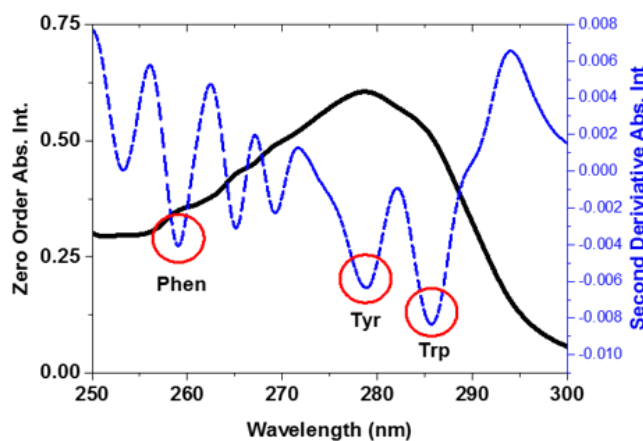


Figure 2.12: UV zero order absorbance spectrum of BSA (solid black line) indicating the single absorbance band attributed to the three aromatic amino acids. The corresponding second derivative spectrum of BSA (broken blue line) clearly showing the individual signals from the three aromatic amino acids.

Protein solutions had their zero order absorbance spectra recorded as per section 2.5.1.1. Data for each sample was analysed using Origin Pro 9.1. Second derivative operations were performed on each spectrum. The Savitzky-Golay algorithm was used for smoothing after differentiation. In the second derivative spectra, the distinctive peaks were assigned to phenylalanine (Phen), tyrosine (Tyr) and tryptophan (Trp). The presence of phenylalanine in proteins is generally recognised as weak ripples or shoulders within the absorbance spectra between 250 – 270 nm region (Schmid 2001). Due to this, phenylalanine shows as several negative bands within the second derivative spectrum within this region (Balestrieri, Colonna *et al.* 1978). However it is the strongest band at 259 nm that is monitored for evaluation purposes (Ichikawa and Terada 1977; Mach, Thomson *et al.* 1991; Lucas, Ersoy *et al.* 2006). For each of the three residue peaks the wavelength value at the peak minima was recorded.

2.5.3.2 Second derivative fluorescence spectroscopy

Fluorescence derivative methods can be used to attain structural and functional information about transitions taking place in the environments of aromatic residues (Mozo-Villarias 2002). It can be used as a technique to monitor and characterise relatively

small changes and subtle alterations that may not be seen in the zero order fluorescence spectra. When proteins are excited at 280 nm, the signal arises from both the tryptophan residue and tyrosine residue. Using second derivative operations on the zero order spectrum can allow deconvolution of this single fluorescence signal into the two individual peaks representing each individual amino acid residue (figure 2.13). Any overlapping peaks and shoulders can be resolved into distinct peaks.

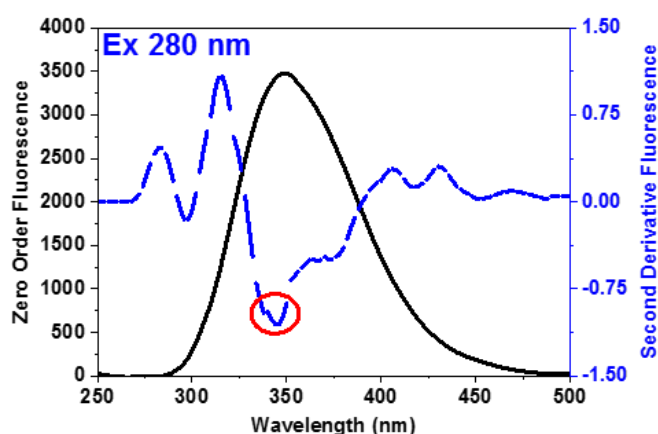


Figure 2.13: Intrinsic fluorescence zero order spectrum of BSA (solid line) indicating the single absorbance band attributed to the tyrosine and tryptophan residues. The corresponding second derivative spectrum of BSA (broken line).

All fluorescence measurements were performed as per section 2.5.2.1. Differentiation of the normalised spectrum was performed using Origin Pro 9.1. The Savitzky-Golay algorithm was used for smoothing after differentiation. The criteria for smoothing was that the overall shape and intensity of the zero order emission spectrum is not affected, and that the peak and shape of the bands in the second derivative is preserved and excess noise is removed. In the second derivative spectra, the distinctive peaks were assigned to both tyrosine (Tyr) and tryptophan (Trp). For each peak both the intensity and wavelength values of the minimum negative band were recorded.

2.6 High throughput assays

High throughput screening assays were carried out using a multi-platform plate reader, Molecular Devices SpectraMax M2e and are performed in a 96 well-plate format.

2.6.1 Turbidity assay

Changes in protein sample turbidity is in general associated with protein aggregation (provided other phase transitions associated with increased turbidity, including liquid-

liquid phase separation are not occurring). If the process of aggregation produces a relatively high concentration of small particles or a smaller number of larger particles, this leads to light scattering and opacity. The turbidity, or light scattering by precipitates can be detected at non-absorbing wavelengths (≥ 400 nm) to detect protein aggregates.

2.6.1.1 Protocol

Samples were measured in triplicate, in a 96 well plate format using UV-Star microplate 96 well, f-bottom, μ Clear (Greiner Bio-One). These plates have a specialised optical window that allows for measurement of samples within the UV range (down to 230 nm), ideal for measuring protein concentrations at 280 nm. No pre-treatment of the plate surface was carried out. Stock protein samples were diluted in the chosen buffer to a protein concentration range of 0.5 – 20 mg/ml. The total volume of samples was fixed at 200 μ l. Solution turbidity was determined spectrometrically by measuring the change in absorbance at 450 nm. The extent of turbidity was calculated using percentage (%) transmission. The variation in transmitted light through a turbid solution can be seen as an increase in sample absorbance. This change in absorbance can be related to the percentage of light transmitted through the sample *i.e.* the more turbid the sample, the higher the scattering and thus the % transmission is low. The absorbance at 450 nm is related to a percentage transmittance (% T) using the following relationship;

$$A = 2 - \log_{10}(\%T) \quad 2.12$$

$$\%T = 10^{(2-A)} \quad 2.13$$

Measurements were taken every 20 minutes until no further changes in % T were detected. Buffer solutions at the appropriate salt concentration and pH were used as blanks, and manually subtracted from the protein data. Average % T data from the three replicates was averaged and error calculated as the standard deviation.

2.6.2 Thioflavin T assay

Thioflavin T (ThT) is a benzothiazole dye traditionally used as an indicator for the presence of amyloid structures (figure 2.14). It exhibits enhanced fluorescence upon association amyloid fibrils (extrinsic fluorescence). Due to the specificity of ThT to associate amyloid fibrils, it has found many applications such as diagnosis of amyloid in tissue sections using fluorescence microscopy, monitoring extracted amyloid and in vitro

amyloid fibril formation using fluorescence spectroscopy (Khurana, Coleman *et al.* 2005).

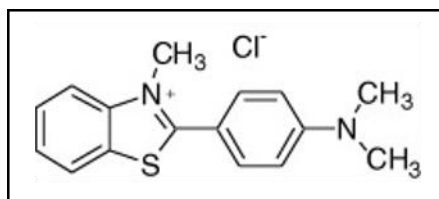


Figure 2.14: Chemical structure of Thioflavin T.

2.6.2.1 Protocol

ThT (Fisher) stock solutions were prepared at a concentration of 2 mM in Milli-Q water (3.18 mg in 5 ml), covered in foil and stored at 4°C for up to one month. ThT was mixed with protein and buffer (at appropriate concentrations) to give a total volume of 300 μl and a final ThT concentration of 200 μM in NUNC 96 black Polypropylene MicroWell™ plates. A sample containing ThT and protein in buffer was used as a reference. Excitation at 435 nm and emission between 465 nm and 565 nm or at a single wavelength of 485 nm were recorded (figure 2.15). In the presence of amyloid fibrils, the intensity of ThT fluorescence at 485 nm increases sharply. The average of three identical samples was determined and errors calculated as the standard deviation.

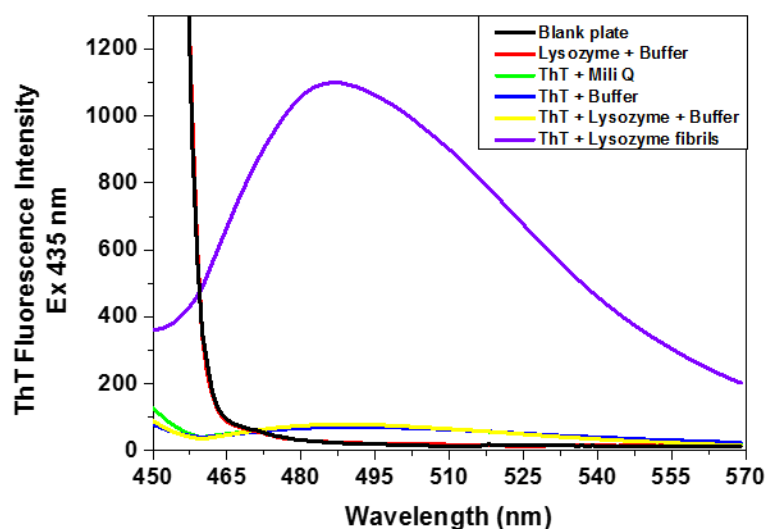


Figure 2.15: Thioflavin T assay spectrum indicating the increase in fluorescence peak at 485 nm attributed to ThT association with fibrillar material.

2.7 Phase separation

2.7.1 Removal of large (precipitated) aggregates

The formation of large aggregate particles during denaturation (or other aggregation pathway) is a common occurrence. To further characterise either the aggregates or the supernatant, samples were centrifuged (15,000 x rpm/20 minutes) to allow for the separation of precipitated aggregates from the bulk protein solution (the supernatant) (figure 2.16). The protein supernatant was then carefully pipetted from the aggregated pellet of protein and the concentration of the supernatant re-determined as per 2.5.1.1 to establish the loss in protein to aggregated particles.

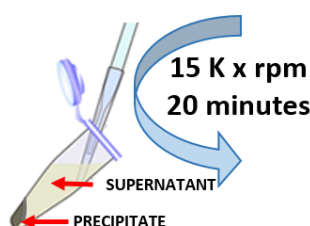


Figure 2.16: Separation of precipitated aggregates from bulk protein solution by centrifugation.

2.7.2 Characterisation of supernatant

Once the precipitated aggregates have been removed, the supernatant can be characterised using PAGE or HPLC. These techniques can be used to identify the quantity of monomeric protein retained after aggregation and/or the formation of small, oligomeric (soluble) aggregates in the supernatant. The protein remaining in the supernatant can also be assessed over a prolonged time period to evaluate the stability of the remaining protein. This was done using SE-HPLC.

2.8 Microscopy

As protein aggregates, particles can become large enough to be accessible to different microscopic techniques. Numerous microscopy methods can be used to image large protein aggregates (larger than $\sim 1 \mu\text{m}$). The major challenge in microscopy is isolation and preparation of the protein sample. It is important that these procedures don't impact on the structure of the particle. It is also important that a sample aliquot is representative of the sample as a whole (Mahler and Jiskoot 2012).

2.8.1 Light microscopy imaging protocol

In order to visualise protein particles within a solution, the sample was pipetted gently up and down to ensure mixing of the sample so an adequate representation of the sample could be obtained. Samples for analysis were prepared on microscope slides and a coverslip was placed on top. Where thick precipitated protein aggregates were formed, these were gently transferred onto the microscope slide with a spatula and a small (~10 µl) amount of buffer was pipetted over the sample to ensure it would not dry out. Protein aggregates were observed using an Olympus BX61 microscope, equipped with a digital imaging system with magnification to up 100X. Images of the viewed structures were viewed and recorded using Cell^F software. Microscopy images had their contrast and brightness adjusted so that the morphology of the structures could be best viewed and evaluated.

2.8.2 Polarised microscopy imaging protocol

Polarised light microscopy is used to generate contrast in birefringent sample components, indicative of structural order. For example, it allows for the determination of the optical crystallographic properties of crystals. The polariser functions to convert light from the light source, which vibrates in random directions, into linearly polarised light. When the linearly polarised light passes through a birefringent sample, it is divided into ordinary and extraordinary rays, with different speeds and refraction indexes. When the polariser and the analyser are aligned orthogonally only the light modified by birefringent sample can pass through. Polarising microscopy can be used for quantitative and qualitative analysis of protein aggregate species, providing information on properties such as morphology, refracting index, dichroism and interference figures (Weaver 2003). Protein aggregates were assessed for their ability to show birefringence using the same protocol as in section 2.8.1, using the same optical microscope equipped with a polariser (located before the sample) and an analyser (located between the sample and camera). Images of the viewed structures were viewed and recorded using Cell^F software. Microscopy images had their contrast and brightness adjusted so that the morphology of the structures could be best viewed and evaluated.

2.8.3 Transmission electron microscopy

Transmission electron microscopy (TEM) uses a high voltage beam of highly energetic electrons sent through a thin slice of sample to create an image of the sample, the resulting interference pattern. It allows for better resolution than light microscopy as the beam of electrons have a far smaller wavelength. The electron source is an electron gun (tungsten cathode) that emits an electron beam accelerated by an anode (40 – 400 keV). The beam is then focused by electrostatic and electromagnetic lenses and transmitted through the sample. Depending on the density of the material present, some of the electrons are scattered from the beam. Based on the degree of scattering and interaction with the sample, the electron beam carries information about the structure of the sample and is magnified by an objective lens. CCD cameras can be coupled to the device to take images (Mahler and Jiskoot 2012). The resolution limit of TEM is approximately in the range of 100 μm to 0.05 nm.

2.8.3.1 TEM Protocol

TEM work was carried out in The Royal College of Surgeons, Ireland with the help of Mr. Brenton Cavanagh. The instrument used was a Hitachi H-7650 Transmission Electron Microscope. Samples for examination were prepared using an ‘inverted droplet method’. This consisted of gently pipetting up and down the sample for analysis to ensure a homogenous mixture. The mixed sample was then pipetted ($\sim 20 \mu\text{l}$) onto parafilm and a metal grid (300 mesh thin bar copper grids; 3.05 mm diameter) was placed on top (figure 2.17). These were left for 30 minutes to allow sufficient time for sample to adhere to the grid while avoiding large protein aggregates sticking to and blocking up the grid. The grid was removed from sample and excess solution evaporated off. The grid was then inserted into the electron column in the microscope and into the beam path. The beam strength was 100 kV. Microscopy images had their contrast and brightness adjusted so that the morphology of the structures could be best viewed and evaluated.

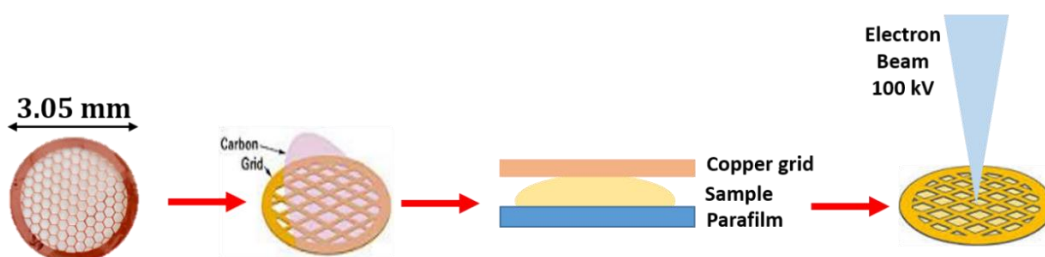


Figure 2.17: Method of sample preparation for TEM.

2.9 Isolation of monoclonal antibody IgG1

2.9.1 Expression and purification of IgG1 from CHO DP-12

Cell culture medium with expressed IgG1 was kindly provided by Professor Ian Marison's laboratory, Dublin City University. The antibody was expressed in the CHO (DP12v2) cell line. Two batches of protein were received. Batch 1 was grown in shake flasks and Batch 2 was grown in a 2.5 L bioreactor. The bioreactor conditions used were; temperature 37°C, pH: 7.2 and dissolved oxygen: ~95%. The cells were cultured in ExCell CHO CD3 media which was supplemented with 4 mM L-glutamine, 1 µM MTX, 10 mg/L insulin and 0.1% antifoam. The medium was stored at -80°C.

2.9.2 Purification of monoclonal antibody

IgG1 was purified from the cell culture medium using a 5 ml HiTrap™ Protein A HP column (GE Healthcare). Before purification, the cell culture medium was defrosted slowly at room temperature. Defrosted material was then centrifuged at 8,500 rpm for 2 hours to remove large particles and any remaining cellular material. The supernatant was then separated from the pellet. The supernatant was syringed through 0.22 µm filters prior to being loaded onto the separation column.

A series of buffers were prepared for the purification of the IgG1 from the medium;

1. Binding buffer – 20 mM sodium phosphate, pH 7.0
2. Elution buffer – 0.1 M citric acid, pH 3.5
3. Neutralisation buffer – 1 M Tris-HCl pH 9.0

Collection tubes were primed by adding 100 µl of neutralisation buffer per 9 ml of fraction to be collected. The HiTrap™ column was washed with 10 column volumes of binding buffer at a flow rate of 5 ml/min. Prepared sample was loaded on to the column at a flow rate of 1 ml/min. The column was again washed with at least 10 column volumes of binding buffer (or until no material appeared in the effluent) at a flow rate of 5 ml/min. Protein was eluted from the column with 2 to 5 column volumes of the elution buffer at a flow rate of 1 ml/min into the collection tubes. The column was then washed with 5 column volumes of 20% ethanol to prevent bacterial growth and for storage purposes. Peak fractions from the purified IgG1 were pooled and washed into a 50 mM phosphate buffer, pH 8 using Amicon ultra 0.5 ml 10 kDa cut off membrane centrifugal filter units (Merck Millipore). Purified antibody was aliquoted into microcentrifuge tubes and kept

at - 80°C. The concentration of purified IgG was measured by UV absorbance using the extinction coefficient 1.40 cm²/mg as a guide (Neergaard, Nielsen *et al.* 2014; Hoger, Mathes *et al.* 2015; Molloy, Fesinmeyer *et al.* 2015). This is to be officially determined.

The protein was analysed using SE-HPLC (as described in 2.3) to ensure purity. The column used for all IgG1 work was TSK-GEL® 3000 SWxl column (Tosoh Bioscience). The columns packing material is made up of 5 µm silica particles with 250Å pores, making it ideal for the separation of IgG1, a relatively large protein (150 kDa). The column resolution ranges from 10 – 500 kDa. The mobile phase used was 50 mM sodium phosphate, pH 8 at a flow rate of 1 ml/min with a 45 µl injection volume. The purity of the protein was also assessed using PAGE. Native PAGE was used to insure the correct molecular weight protein was isolated and purified. Reduced SDS-PAGE was used to separate the IgG into its heavy and light chain.

2.9.3 Identifying the glycoform of IgG1

Glycoanalysis was performed in the laboratory of Professor Pauline Rudd, NIBRT. All immunoglobulins are glycosylated. Glycosylation is post-translational modification that involves the attachment of sugars (glycans) to the antibody. The quantity and quality of glycosylation depends both on the protein itself and also on the cell type expressing the protein. Changes in the IgG glycoforms can alter their function and stability.

Sugars were enzymatically cleaved from the IgG1 using the enzyme PNGase F, capable of removing most *N*-Glycans and is the most reliable and easiest approach (Huhn, Selman *et al.* 2009). PNGase F is an amidase that cleaves the *N*-linked glycan from the Fc portion of IgG. It removes all asparagine-linked oligosaccharides with high yields by cleaving the bond between the GlcNAc and asparagine yielding a glycan with a free reducing terminus. This mechanism is not yet fully understood, but believed to occur via an initial β-elimination step prior to a reaction with hydrazine derivate (Merry, Neville *et al.* 2002).

2.9.3.1 Extraction of glycan from IgG1

The heavy chain fragment of the IgG was isolated using reducing SDS-PAGE, cut out and chopped into 1 mm sections using a sterile knife. 1 ml of acetonitrile was added to the gel pieces in a microcentrifuge tube and rotated/shaken for 10 minutes to cause the gel pieces to expand. All liquid was removed from gel pieces. 1 ml of 20 mM sodium

bicarbonate (NaHCO_3) was added and rotated/shaken for 10 minutes to remove the coomassie blue stain from the gel pieces. The stain contains SDS that will interfere with steps further in the process. These steps were repeated until all dye was removed. 2 μl of Peptide -N-Glycosidase F (PNGase F) enzyme and 48 μl of buffer (NaHCO_3) were added to the gel pieces. Gel pieces were left to swell and soak up the solution, and kept topped up with solution to prevent them drying out. Samples were then left in an incubator overnight (at least 16 hours) at 37°C . Samples were then washed using 200 μl of water, vortexed and sonicated for 15 minutes. All liquid was removed into labelled microcentrifuge tubes and retained. This process was repeated with 200 μl water followed with 200 μl acetonitrile two more times with all liquid retained (1000 μl total). Liquid samples were completely dried in a vacuum centrifuge (~ 50 μl liquid/ 30mins, total time ~10 hours). Once dry, samples were stored at -20°C .

2.9.3.2 Labelling of glycan

Glycans do not absorb in the UV region and don't contain any fluorophores. They require derivatization to add a chromophore or fluorophore, allowing for detection after separation (Ruhaak, Zauner *et al.* 2010). IgG1 was derivatized using the fluorescent label 2-aminobenzamide (2-AB). Samples were removed from the freezer (-20°C) and 5 μl of 2-aminobenzamide (2-AB) labelling mixture (LudgerTag 2AB labelling kit) is added to each tube, agitated for 5 minutes to ensure thorough mixing and incubated for 30 minutes at 65°C . Samples were agitated for a further 5 minutes and incubated for a minimum of 1.5 hours at 65°C . This approach is advantageous as it gives 1:1 stoichiometric attachment of one label per glycan, allowing accurate and direct quantification based on fluorescence or UV-absorbance intensity. This is carried out in excess to ensure complete derivatization of the released sugars.

2.9.3.3 Label clean up – paper chromatography

Paper chromatography was used after reductive amination to remove excess label. 3 mm Whatman chromatography paper was cut, washed three times in Milli-Q water (2 minutes/wash) and dried in an oven at 65°C . 5 μl of sample placed on the dried paper. Acetonitrile was used as the mobile phase. This was left for 30 minutes or until a clear separation could be seen with a UV lamp. The paper was dried for 15 minutes and the spot containing the sample was cut out. The paper piece was then washed into a microcentrifuge tube using a Millex Syringe Driven non-sterile, low protein binding,

hydrophilic filter (Millipore). First the paper was soaked in 500 μ l of Milli-Q water for 15 minutes, then 400 μ l Milli-Q water was added, and washed through. This process was repeated 3 times, waiting 15 minutes between each wash. Samples were placed into a Speed Vac to dry overnight.

2.9.3.4 Separation of glycoforms

Undigested Samples

After labelling, glycan samples were separated using ultra performance liquid chromatography (UPLC). The remaining pellet was diluted in 50 μ l Milli-Q water. 5 μ l of sample was added to acetonitrile in a LC vial and mixed. Samples were then run using ACQUITY® UPLC H-class 4 on an Acquity® UPLC BEH Glycan 1.7 μ m, 2.1 x 50 mm column (both Waters) designed for HILIC-mode separations of 2-AB labelled glycans. The mobile phase was 30% 50 mM NH_4HCOOH pH 4.4, 70% acetonitrile, flow rate 0.561 ml/min for 30 minutes, sample injection was 20 μ l, and Milli-Q water was used as a blank and a gradient separation with increasing amounts of aqueous phase was performed.

Digested Samples

20 μ l of sample was digested into fragments using specific enzymes known as exoglycosidases, and a denaturing buffer (10X; 0.5% SDS, 40 mM DTT) according to Table 2.6. Each enzyme cleaves a different part of a glycan structure breaking it down into individual sugar units to allow identification of the full glycan structure. The sample was mixed with the components and left in oven overnight at 37°C.

| | <i>10X buffer (μl)</i> | <i>ABS (μl)</i> | <i>BTG (μl)</i> | <i>BKF (μl)</i> | <i>GUH (μl)</i> | <i>JBM (μl)</i> | <i>Milli-Q H₂O (μl)</i> | <i>TOTAL (μl)</i> |
|---------------------------------|---|------------------------------------|------------------------------------|------------------------------------|------------------------------------|------------------------------------|---|--------------------------------------|
| UND | 1 | - | - | - | - | - | 9 | 10 |
| ABS | 1 | 1 | - | - | - | - | 8 | 10 |
| ABS+BTG | 1 | 1 | 2 | - | - | - | 6 | 10 |
| ABS+BTG+BKF | 1 | 1 | 2 | 1 | - | - | 5 | 10 |
| ABS+BTG+BKF +GUH | 1 | 1 | 2 | 1 | 2 | - | 3 | 10 |
| ABS+BTG+BKF +GUH+JBM | 1 | 1 | 2 | 1 | 2 | 1 | 2 | 10 |

Table 2.6: Reagents and enzymes used to digest the glycan structure *Arthrobacter ureafaciens* sialidase (ABS), *Bovine kidney alpha-fucosidase* (BKF), *Bovine testes β -galactosidase* (BTG), *B-N-acetylglucosaminidase* (GUH), *Jack Bean α (1-2,3,6)-Mannosidase* (JBM)

Figure 2.18 details the order of exoglycosidases used to digest our IgG1 glycan. Starting with our undigested sample, successive enzymes were added to cleave off different elements of the glycan structure to aid in identifying the glycan profile.

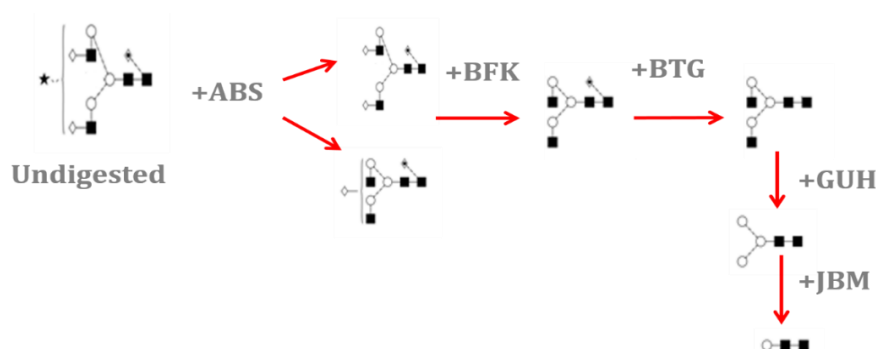


Figure 2.18: Exoglycosidases used to achieve glyco profiling of our IgG1 and where they specifically cleave.

After the glycan structure had been completely digested into individual sugar units, the remaining enzyme was neutralized by placing the sample into an oven at 65°C for 5 minutes. Nanosep 10 kDa MWCO centrifugal filters (OMEGA) (washed x2 with 500 μ l Milli-Q water at 13,000 rpm for 5 minutes) were used to wash the sample from the microcentrifuge tube. Addition of water to the tube and then to the filter is repeated three times to ensure all sample is removed. The sample was then placed into the Speed Vac and dried down.

The dried pellet was dissolved in 5 μ l Milli-Q water & added to 80 μ l acetonitrile plus 20 μ l of Milli-Q water (total 105 μ l). Samples were run on Normal Phase HPLC (Waters)

using a TSKgel Amide 3 μm (150 x 4.6 mm) column (Anachem). Mobile phase was 30% 50 mM NH_4HCO_2 pH 4.4, 70% acetonitrile, flow rate 0.4 ml/min, sample injection was 95 μl . A 2-AB labelled dextran ladder was also run as a standard. The dextran ladder is a characteristic ladder profile made up monomeric glucose. Each peak is expressed as a glucose unit (GU) used to assign GU values released in the glycan pool.

Elution properties were carried out by performing a standardization/calibration using a dextran hydrolysate ladder (Royle, Campbell *et al.* 2008). This is a characteristic ladder profile from monomeric glucose. Each peak is expressed as a glucose unit (GU) used to assign GU values released in a glycan pool. Elution peaks are expressed in terms of (GU). The GU for each glycan is then directly related to the number and linkage of its constituent oligosaccharides; the larger the glycan, the higher its GU. Therefore, the removal of sugars by exoglycosidase digestion results in movement of the peak to a lower GU.

2.9.3.5 Peak assignment

Interpretation of data and structural assignment of peaks from UPLC and NP-HPLC profiles was carried out via a web-based database bioinformatic tool-GlycoBase 3.0+ to establish the glycan composition of our IgG1.

Chapter 3 High-Throughput Approach to Measure Protein Aggregation

3.1 Introduction

3.1.1 Background

The pharmaceutical industry has invested billions of dollars in the development of new drug targets and leads over the last 20 years. Despite this large investment, the industry is in a decline. Since 2010 the number of FDA-approved new molecular entities (NMEs) has dramatically decreased and it has been argued that improved target selection is the key to tackling this challenge (Bunnage 2011). The growing number of biopharmaceutical therapeutic proteins coming to market are more challenging to analyse than small molecules because of their inherent complexity; numerous molecular weights, variable conformations, solubility and stability, and post-translational modifications (Crommelin, Storm *et al.* 2003). The manufacturing of such protein formulations requires a greater number of batches (>250 vs. <10), a larger number of product quality tests (>2000 vs. <100) and larger process data entries (>60,000 vs. <4000) (Rathore 2011). This means more powerful analytical techniques and complex approaches to screening are necessary to obtain detailed characterisation of the protein product (Staub, Guillarme *et al.* 2011). Forced degradation (FD) studies or accelerated tests are an essential step in formulation development. They are carried out early in the development stage to establish the most suitable conditions for protein formulations and define optimal conditions for long term storage (Hasija, Li *et al.* 2013). They also give insight into potential degradation pathways and ensure that analytical methods are able to detect degradation products for stability assays and to set up screening protocols (Hawe, Wiggernhorn *et al.* 2012). The international Conference on Harmonisation (ICH) provides scientific and regulatory guidelines on how to approach FD studies for the pharmaceutical industry in USA, Europe and Japan. The ICH Guideline Q5C on stability testing in biotechnological products advises that drug manufacturers provide information on product stability, including information on numerous external conditions that may affect the quality of the product (ICH 1998).

3.1.2 The problem

Currently, drug discovery and formulation stability measurements can be time-consuming, require high concentrations of pure product and can be costly. The high expense in producing protein formulations means manufacturers often do not commit to

the production of large quantities of material until positive screening results are achieved (Capelle, Gurny *et al.* 2007). The time cycle for introducing new targets now averages over 13 years from project inception to market costing as much as €1 billion. Roughly 4 years is spent in preclinical trials and on average it takes 9 years for a drug candidate to progress through the clinical phases of development (Bunnage 2011).

3.1.3 High-throughput as a solution

For this reason it is often advantageous to screen a large number of solution conditions in order to assess the quality/stability of the protein in solution. The main factors driving the competitive biopharmaceutical market are; ‘first to market’, product quality and cost-effectiveness of the product production. Phrases such as ‘hit generation’, lead generation’ and ‘lead optimization’ are used in the rapid discovery and selection of new candidates/products (Roberts 2001). The ability to measure numerous parameters in development processes in a rapid, cost-effective fashion to screen samples for production and to aid in design and formulation development is the key to achieving this (Capelle, Gurny *et al.* 2007). This is fuelling the need for improved high-throughput screening (HTS). HTS is a technology widely used in the pharmaceutical and biotechnology industry. Only in the past 10 years have HTS scientists been able to bring together sophisticated automated equipment, robust assays, validated targets with high quality compound libraries, useful data analysis tools and a better understanding of how to use HTS. These technologies are now the primary tool to monitor protein product degradation including aggregation and precipitation (Pereira and Williams 2007; Xia and Wong 2012). Essential to obtaining high-precision data is the use of fully automated platforms coupled with robotics with benefits such as miniaturisation and parallelisation. This benefits the cost-effectiveness of the process and run times. The challenge is in finding the right drug/protein formulation while concurrently optimising the screening of targets at a minimal cost per sample, through quality detection with a high degree of confidence (Mayr and Bojanic 2009; Xia and Wong 2012).

3.1.4 Advancements in high-throughput technologies

Various technologies have been developed that take advantage of assay miniaturisation, lab automation and robotics to aid in the development of HTS (Pereira and Williams 2007; Mayr and Bojanic 2009). Since the first microtitre plate was invented in 1951 by

Dr. Guyla Takatskay, a revolution in the technologies for HTS has occurred and has advanced rapidly with instruments now capable of preparing and/or analysing large numbers of samples with low or no personal assistance in relatively short time periods (Sundberg 2000). By the late 1990s, a shift from the assay systems and robotics that were capable of screening thousands of samples per day, saw the rise in the development of ultra high-throughput screening methods (uHTS). These have increased the number of assays per day from several thousands to 100,000 or more per day (Hertzberg and Pope 2000). The goals of HTS and uHTS are automation, miniaturisation and high-throughput with emphasis on the development of assay formats coupled to highly sensitive detection methods. The trend towards miniaturisation and automation was also driven by the need to reduce development and operating costs, thus reducing the amount of biological and chemical reagents used per assay, facilitating higher throughput assay formats by enabling the use of parallel sample approaches and multiplexed detection modes (Gomez-Hens and Aguilar-Caballos 2007). uHTS has evolved in the areas of assay automation, miniaturisation and detection methodologies. Assays while initially adopting a 96-well microplate format, have now advanced towards 384- and even 1536-well microplate assays. The next goal is in designing plate-based formats with higher well densities and smaller volumes. There is also a trend towards microfabricated devices designed to perform continuous flow assays and advanced chip technology capable of handling nanoliter samples (Sundberg 2000; Mayr and Bojanic 2009).

3.1.5 HTS methodology

Numerous automated processing steps are involved in in HTS including using liquid and power handling systems, barcode readers, stackers, shakers, cooling or heating systems, incubators, centrifuges, filtration stations, plate sealers, washers and a variety of detection systems. It is of vital importance that any of these systems do not induce any stress on the protein that might disrupt its stability. A variety of techniques can then be employed to analyse the protein quantitatively or qualitatively *i.e.* fluorescence, absorbance, luminescence (Gomez-Hens and Aguilar-Caballos 2007). HTS can be classified into two groups; invasive (calorimetry, chromatography, mass spectrometry) and non-invasive (UV-Vis absorbance, fluorescence and luminescence) depending on the technique used and the possibility of recovering the protein after characterisation. (Capelle, Gurny *et al.* 2007). Figure 3.1 introduces a schematic representation of several components involved

in HTS from sample preparation in a multi-well plate format using automated handling systems to detection and interpretation of results and how this can feed back into a loop for designing new experiments. Assays in microarray formats can be carried out to either assess the same target in multiple wells with different concentrations to assess its potency, or one sample per well gaining multiple measurements from a single sample to increase information about their properties (Gomez-Hens and Aguilar-Caballos 2007). Typically just a single measurement of each sample is obtained in an initial primary screen. This is evaluated for biological relevance by a counter screen and confirmed as a hit by a secondary screen that may test fewer samples. -Marginal hits on the first screen may fail to validate on the second screen because of measurement or procedural error (Malo, Hanley *et al.* 2006).

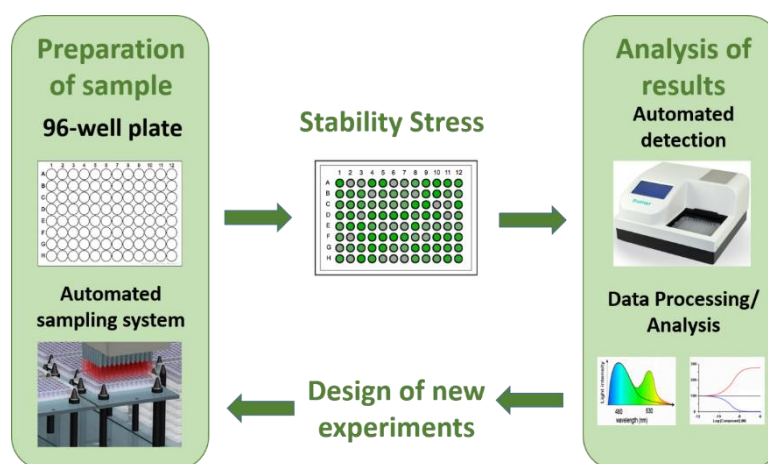


Figure 3.1: Schematic of HTS process from sample preparation, applying a stability stress to detection and analysis of results.

The most important factors in HTS are minimising the amount of sample and products needed, selecting the correct assay, integrating multiple automated techniques, continuous evaluation of the data and flexibility in designing new experiments (Capelle, Gurny *et al.* 2007).

3.1.6 Scepticism of HTS

However there are some that believe the balance between increasing the throughput to a maximum level of assays per day, versus higher quality screening is askew. Analysis has shown that in fact HTS have hindered the productivity of the pharmaceutical industry as it undermines creativity and critics of HTS feels it is not worth the return on investment (Lahana 2003; Garnier 2008). As there is a push for quantity and higher throughput, there

needs to be careful consideration that this is not at a downfall of higher quality data that can provide more information and leads that can become clinical candidates (Macarron, Banks *et al.* 2011).

A problem with HTS methodologies is that with increasing throughput there is usually a greater chance that good or bad properties of a sample will be wrongly categorised. The generation of “false-positives” (sample that wrongly pass a particular minimum requirement) and “false-negatives” (samples that should have passed) are a major point of concern for HTS. Roberts (Roberts 2001) states of these outcomes; false-negative are of more concern as de-selection during screening can lead to exclusion of promising candidates. The quality of analytical results depends critically on the precision and accuracy of the HTS device (Wolcke and Ullmann 2001). Sometimes a screen may be sufficient for one set of conditions, but not another. Microplates stored for prolonged time before measuring can result in sample evaporation. Water in sealed versus non-sealed microplates have shown increased evaporation at the corners and borders compared to the centre leading to potential errors in measurement (Berg, Undisz *et al.* 2001). This study also showed that while automated pipetting into microplates of higher density (384+) was more efficient than manual pipetting in time, effort and accuracy, the accuracy and precision of HTS liquid handling systems depended on many factors from the medium handled, the plate type used, the delivery methods and methods of detection (Rhode, Schulze *et al.* 2004). Bertrand *et al.* believe that while progress with HTS methods has been achieved, progress has moved too fast with automated systems being established before we have fully mastered the handling of the vast amounts of data produced (Bertrand, Jackson *et al.* 2000). There is a lack of universal procedures for processing and extracting knowledge from screens or any standardised method of the ‘hit’ identification process (Malo, Hanley *et al.* 2006). There is also a considerable lack of literature reporting the limitations and pitfalls that have been identified from HTS methods that might highlight errors and incorrect interpretation of data recovered.

3.1.7 The end goal

More interaction between chemists and biologists creating multidisciplinary teams with design plans can improve efficiency and quality. Improvements in both hit specificity and sensitivity need to be met by the technologies running and operating the systems, along with data analysis methods to achieve the full potential of HTS (Malo, Hanley *et al.*

2006). However the industry is still growing and improving the quality and validation of their targets and leads, industrialising and streamlining their HTS laboratories and increasing communication and integration with other drug discovery groups (Macarron 2006).

The combination of forced degradation studies with high-throughput assays is of great importance as it will aid in identifying more robust formulations while being more economic and less time consuming (Hawe, Wiggenhorn *et al.* 2012). However, careful consideration of the quality of data must be given to ensure that speed is not valued over quality and this will be explored in this chapter.

3.2 Aims of the study

Quantification of protein aggregation is regularly carried out in the pharmaceutical industry using expensive techniques such as the Coulter method, microscopic techniques or by light obscuration, that either need high concentrations of material, are time consuming, require manual counting of particles or limited by detection of particle size (Roberts and Wang 2010). Any process that can reduce the amount of time and resources involved while still reaching acceptable outcomes are preferred (Capelle, Gurny *et al.* 2007). Therefore optimisation of screening assays using low concentrations of protein, straightforward particle size characterisation and rapid turnaround times are advantageous.

The purpose of this investigation was to assess the effectiveness of a high-throughput approach to measure protein aggregation after forced degradation. Thermal denaturation (Mechanism 2) was used as a model stress as it is well known to induce elevated levels of aggregation and is one of the most common stress mechanisms (Bischof and He 2005; Hawe, Wiggenhorn *et al.* 2012). The mechanism of thermally induced aggregation in our model proteins is well established and studied; lysozyme (Raccosta, Manno *et al.* 2010) and BSA (Militello, Vetri *et al.* 2003). The aggregation of both proteins is known to be sensitive to solution conditions and so could be tuned to ensure protein aggregation occurred for both proteins at concentrations suitable for our high-throughput assay.

A protein high-throughput formulation platform was designed based on the use of microplates using specific methods developed to investigate physical and chemical

properties of the protein formulations. The aggregation of both proteins was assessed using a combination of characterisation techniques, including, visual changes in solution, turbidity measurements, and Thioflavin T (ThT) fluorescence utilising a 96 well multi-plate format and a microplate reader for detection. By detecting changes in solution turbidity and ThT fluorescence, an assessment of the degree and type of protein aggregation could be determined. These results were then compared to other more robust, higher resolution techniques to establish if the high-throughput approach was accurate in the detection and characterisation of protein aggregation. It is important that HTS assays are complemented with other techniques (often referred to as orthogonal techniques) typically applied to analyse protein stability. Ideally the strengths and weakness of any HTS method should be validated so that the method can then be applied when and where it will be most useful and predictive.

3.3 Results

3.3.1 High-throughput assay - Turbidity

3.3.1.1 Lysozyme

The HTS assays were conducted in a 96 well-plate format. Lysozyme (0.5 – 20 mg/ml) in the chosen buffer and within the pH range selected, is colloiddally stable *i.e.* no visible solution aggregation occurs in the absence of the applied stress, in this case, heat (figure 3.2A). After thermal denaturation of lysozyme at 80°C, a visible change in many of the samples could be seen (figure 3.2B). An opaque precipitate was observed in several of the plate wells.

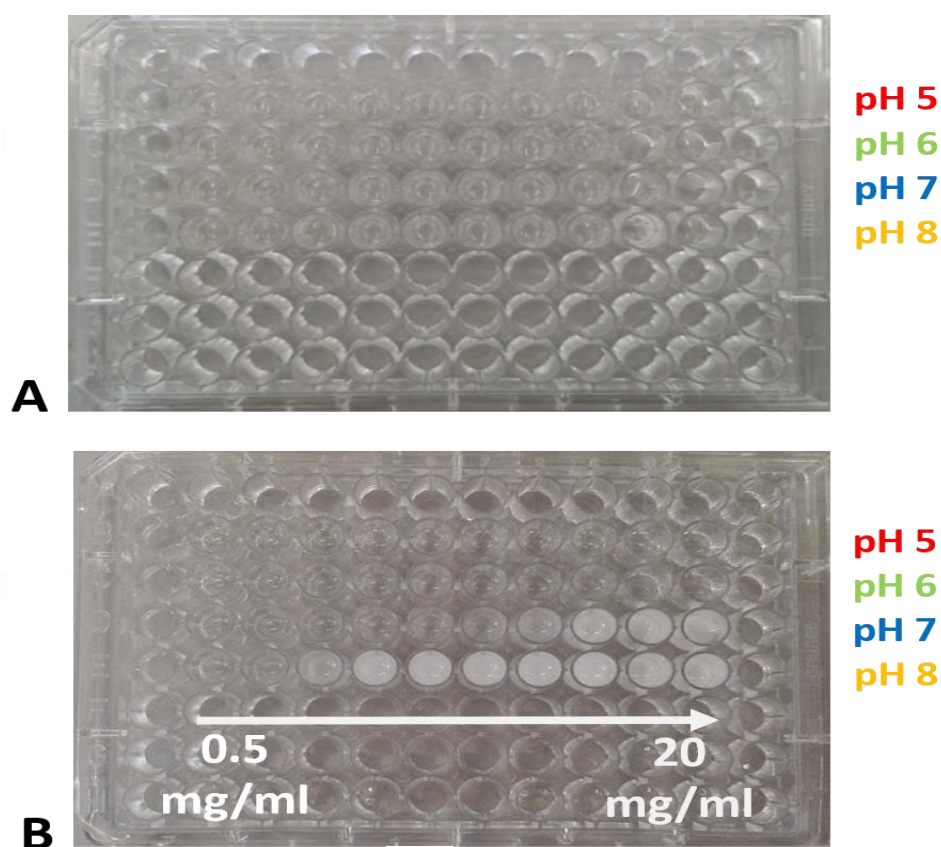


Figure 3.2: 96-multiwell plates A) before and B) after thermal denaturation of lysozyme in 50 mM sodium phosphate for a range of pHs and protein concentrations.

The change in solution turbidity was measured by % transmission at 450 nm before ($t = -20$ mins) and after ($t = 20$ mins) thermal denaturation (figure 3.3A). Based on the measured turbidity levels after denaturation and the visual appearance of the samples, a relative aggregation scale was created to establish a “phase diagram” as a way of indicating the level of precipitated material detected (figure 3.3B). Three divisions on the

relative scale were used; samples that showed high level of turbidity (visible particles), moderate levels of turbidity (subvisible particles, indicated by the absence of visible aggregates, but a decrease in % transmission) and samples that were both visually transparent and indicated no decrease in % transmission (free from large particles).

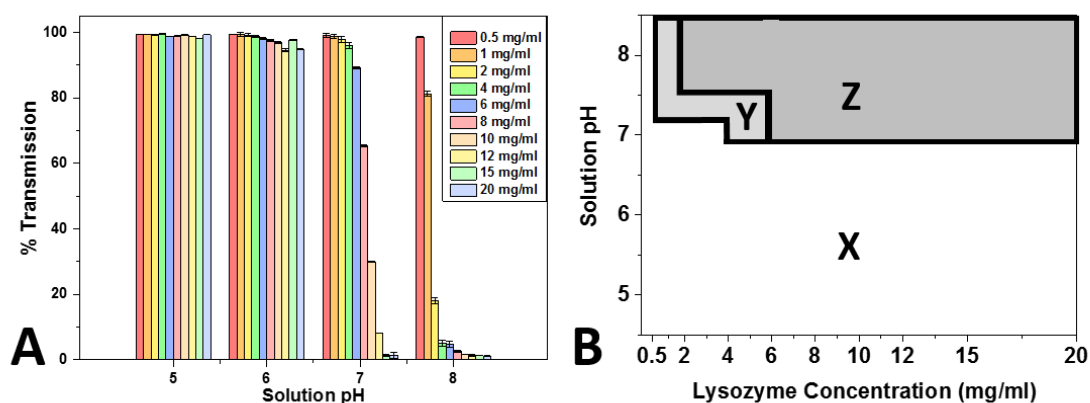


Figure 3.3: A) Turbidity measurement for thermally denatured lysozyme in solution pH 5, 6, 7 and 8 at $t = 20$ min. B) Phase diagram of lysozyme turbidity after thermal denaturation (X=No turbidity, Y=Moderate turbidity and Z=Visible turbidity).

Lysozyme in sodium phosphate buffer at pH 5 and pH 6, showed almost no decrease in % transmission, compared to lysozyme at pH 7 and pH 8, which aggregated to a significant degree. Only in the lower concentration range for both solution conditions (pH 7 < 2 mg/ml, pH 8 < 0.5 mg/ml), did the samples not indicate a measured change in turbidity after thermal denaturation.

3.3.1.2 BSA

The same assay and conditions were used for BSA. Again, before heating, all samples were colloiddally stable and after thermal denaturation at 80°C, a visible change in the protein solutions was observed (figure 3.4A and B). The % transmission was recorded for each sample (figure 3.4C). The turbidity measurements indicated that only at pH 5 did the BSA solution turn visibly turbid. The other solution conditions (pH 6, 7 and 8) showed no change in solution appearance or in % transmission. The samples appeared to be free of large particles which would give rise to turbidity. This is represented as a phase diagram (figure 3.4D).

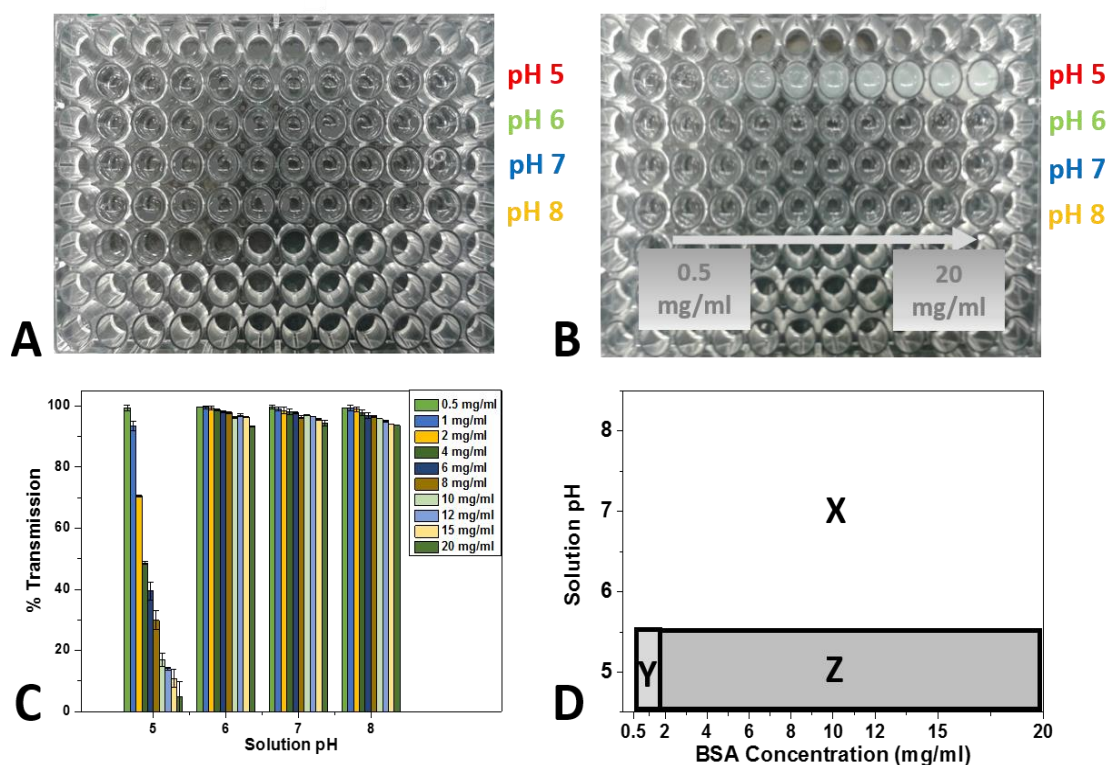


Figure 3.4: 96-multiwell plates A) before and B) after thermal denaturation of BSA in 50 mM sodium phosphate, pH 5, 6, 7 and 8. C) Turbidity measurement for thermally denatured BSA at $t = 20$ min. D) Phase diagram of BSA turbidity after thermal denaturation (A=No turbidity, B=Moderate turbidity and C=Visible turbidity).

3.3.2 Solution turbidity over time

After the initial measurement at $t = 20$ min, further absorbance measurements were made each hour thereafter until no further changes in turbidity were detected. Figure 3.5 shows the difference in % transmission values recorded at $t = 20$ min and $t = 180$ min for both lysozyme (figure 3.5A) and BSA (figure 3.5B). For both proteins the calculated difference in the % transmission recorded between the two time points varied, up to 12% in some cases. In the majority of cases the change in % transmission appears (strangely) to indicate that less turbidity is observed over time. The variation was greater for the samples had turned visually turbid (BSA, pH 5 and lysozyme pH 7 - 8).

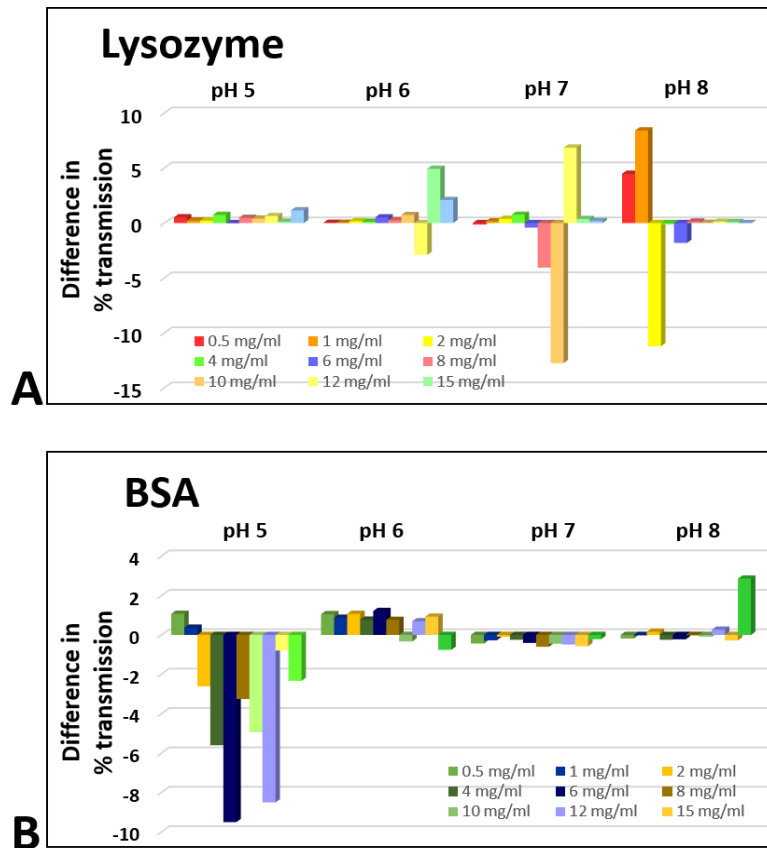


Figure 3.5: Change in % transmission between $t = 20$ mins and $t = 180$ mins. A) Lysozyme and B) BSA.

The question is then if there has been a genuine decrease in the level of protein aggregation or if the decrease has some other explanation. As Figure 3.6 shows, visually there was a difference in the sample appearance between $t = 20$ min and $t = 180$ min. Visual inspection of the sample indicates that all of the aggregated material sediments within the 180 minute timeframe. This leads to heterogeneity in the distribution of the particles at the bottom of the well, and therefore a difference in the % transmission measured for the same sample (which is only read at one point).

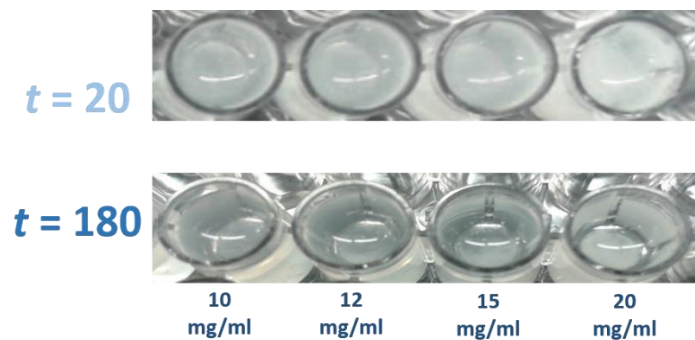


Figure 3.6: 96-multiwell plates containing BSA at pH 5, after thermal denaturation at $t = 20$ min and $t = 180$ min showing the sedimentation of the precipitated material.

3.3.3 Analysis of the supernatant

To determine the amount of protein material that had precipitated as a result of thermal denaturation for each solution condition, the precipitated material was separated from the bulk solution by centrifugation. The concentration of protein remaining in the supernatant was then determined. Table 3.1 shows the concentration of lysozyme (mg) in the supernatant after removal of the precipitated material.

| Initial Lysozyme conc (mg/ml - mg) | pH 5 | | pH 6 | | pH 7 | | pH 8 | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|
| | wt. % loss | wt. % loss | wt. % loss | wt. % loss | wt. % loss | wt. % loss | wt. % loss | wt. % loss |
| 0.5 - 0.1 | 0.098 | 2.00 | 0.098 | 2.00 | 0.088 | 12.00 | 0.086 | 14.00 |
| 1 - 0.2 | 0.196 | 2.00 | 0.194 | 3.00 | 0.172 | 14.00 | 0.162 | 19.00 |
| 2 - 0.4 | 0.394 | 1.50 | 0.396 | 1.00 | 0.310 | 22.50 | 0.316 | 21.00 |
| 4 - 0.8 | 0.796 | 0.50 | 0.794 | 0.75 | 0.596 | 25.50 | 0.594 | 25.75 |
| 6 - 1.2 | 1.196 | 0.33 | 1.192 | 0.67 | 0.982 | 18.17 | 0.954 | 20.50 |
| 8 - 1.6 | 1.592 | 0.50 | 1.594 | 0.38 | 1.250 | 21.88 | 1.282 | 19.88 |
| 10 - 2.0 | 1.990 | 0.50 | 1.992 | 0.40 | 1.460 | 27.00 | 1.588 | 20.60 |
| 12 - 2.4 | 2.394 | 0.25 | 2.384 | 0.67 | 1.966 | 18.08 | 1.820 | 24.17 |
| 15 - 3.0 | 2.980 | 0.67 | 2.994 | 0.20 | 2.330 | 22.33 | 2.172 | 27.60 |
| 20 - 4.0 | 3.990 | 0.25 | 3.980 | 0.50 | 3.042 | 23.95 | 2.806 | 29.85 |

Table 3.1: Concentrations of lysozyme (mg) in the supernatant after thermal denaturation for a range of solution pHs. The wt. % loss in protein is also shown.

At each pH, the proportion of protein lost to precipitation of protein aggregates was only weakly dependent on the protein concentration. At pH 5 and pH 6, the reduction in protein concentration was very small and unlikely to be due to aggregation, but rather a loss of protein due to non-specific surface absorption (Mechanism 5) on the micro-well plate polystyrene surface, and through material transfer steps, since the amount of material lost as a proportion of the total amount decreases with increasing protein concentration. The opposite behaviour would be observed for aggregation, as is observed at pH 7 and pH 8 where a larger % wt. loss in protein can be seen, consistent with the turbidity data.

However, the lower concentration samples at pH 7 (0.5 – 4 mg/ml) that appeared visually free from precipitates and showed no decrease in % transmission after thermal denaturation, also have a relatively high material loss (12 – 25%). This indicates the limit of sensitivity of the high-throughput assay to detect aggregation in either low

concentration samples, or in higher concentration samples with low levels of aggregation. Lysozyme at pH 7 at 6 mg/ml was the lowest concentration to show a measured change in solution turbidity (89% transmission), with a calculated 18% wt. loss in protein, accounting for a 218 μg total loss in protein to aggregates large enough to be detected by the turbidity assay, indicating the limit of detection for the turbidity assay for aggregation. While lysozyme 4 mg/ml at pH 7, with 95% transmission but had a 25% loss in protein accounting for a 204 μg total loss in protein to aggregates that the turbidity assay did not detect. Comparatively, lysozyme at pH 8, 1 mg/ml had a 38 μg loss of protein to aggregation and showed a % transmission of 81%. Thus indicating the assays inability to differentiate between small amounts of large aggregates or high amounts of small aggregates.

The assay also fails to fully quantify the degree of turbidity in highly aggregated samples. Lysozyme at pH 8, between 4 – 20 mg/ml all showed between 1 – 5% transmissions, yet the total loss of protein in the 4 mg/ml sample was 206 μg , for 10 mg/ml the protein loss was 412 μg , while at 15 and 20 mg/ml samples there was a loss of 828 μg and 1.19 mg respectively, of protein to aggregation. The assay is able to detect the change in samples due to aggregation, but is unable to differentiate the amount of protein lost per sample once a certain level of turbidity is reached. Similarly, BSA remaining in the sample supernatant (mg) was also calculated (Table 3.2) along with the wt. % loss of protein as a result of thermal denaturation.

| Initial BSA conc (mg/ml - mg) | pH 5 (mg) | wt. % loss | pH 6 (mg) | wt. % loss | pH 7 (mg) | wt. % loss | pH 8 (mg) | wt. % loss |
|-------------------------------------|--------------|---------------|--------------|---------------|--------------|---------------|--------------|---------------|
| 0.5 - 0.1 | 0.026 | 74.00 | 0.096 | 4.00 | 0.094 | 6.00 | 0.094 | 6.00 |
| 1 - 0.2 | 0.084 | 58.00 | 0.184 | 8.00 | 0.182 | 9.00 | 0.184 | 8.00 |
| 2 - 0.4 | 0.198 | 50.50 | 0.396 | 1.00 | 0.348 | 13.00 | 0.362 | 9.50 |
| 4 - 0.8 | 0.278 | 65.25 | 0.794 | 0.75 | 0.744 | 7.00 | 0.756 | 5.50 |
| 6 - 1.2 | 0.514 | 57.17 | 1.192 | 0.67 | 1.102 | 8.17 | 1.114 | 7.17 |
| 8 - 1.6 | 0.594 | 62.88 | 1.594 | 0.38 | 1.530 | 4.38 | 1.460 | 8.75 |
| 10 - 2.0 | 0.634 | 68.30 | 1.992 | 0.40 | 1.826 | 8.70 | 1.894 | 5.30 |
| 12 - 2.4 | 0.842 | 64.92 | 2.384 | 0.67 | 2.294 | 4.42 | 2.254 | 6.08 |
| 15 - 3.0 | 0.978 | 67.40 | 2.994 | 0.20 | 2.754 | 8.20 | 2.894 | 3.53 |
| 20 - 4.0 | 1.686 | 57.85 | 3.980 | 0.50 | 3.638 | 9.05 | 3.790 | 5.25 |

Table 3.2: Concentrations of BSA in the supernatant after thermal denaturation a range of solution pHs. The wt. % loss of protein is also shown.

For BSA again the amount of aggregate detected visually and with the turbidity assay did not always correlate with the wt. % loss established from the concentration of protein remaining in the protein supernatant, with less aggregation detected by % transmission, then by simply measuring the concentration of protein in the supernatant (which may of course still contain smaller aggregates). Samples that remained visually free from precipitated material (pH 5, 0.5 – 2 mg/ml) showed a decrease in protein concentration corresponding to a % wt. loss comparable to samples that had turned turbid (pH 5, 4 – 20 mg/ml). The 2 mg/ml sample at pH 5, was the last sample to show a measured change in % transmission, where again 202 µg loss to large aggregates was the limit of detection. Solutions at pH 6, 7 and 8 all show between 2 and 13% loss in protein. We know from visual inspection and the turbidity assay that this loss in protein was not through the formation of precipitated material. This loss accounts for up to 360 µg of total protein. This was mostly likely lost through the formation of aggregates not large enough to be detected by a change in % transmission but large enough to be removed through centrifugation at 15K rpm.

3.3.4 Further Screening

Both lysozyme and BSA samples, initially heat treated to 80°C and cooled were then treated to two further heat cycles as a method of secondary screening to assess if marginal

hits on the first screen failed to validate on the second screen *i.e.* to see if any more aggregation occurred upon the addition of further heat stress in samples that appeared stable initially. While certain protein formulations may have no immediate response to the stress condition applied, this doesn't mean stability has not been compromised. From the time point of the last turbidity measurement at heat cycle 1 ($t = 180$ min), without removing the precipitated material from the plate wells, changing or altering the samples, the multi-well plate was again heated for 20 minutes at 80°C. Figure 3.7 shows a multi-well plate containing the lysozyme samples after heat cycle 2 and heat cycle 3.

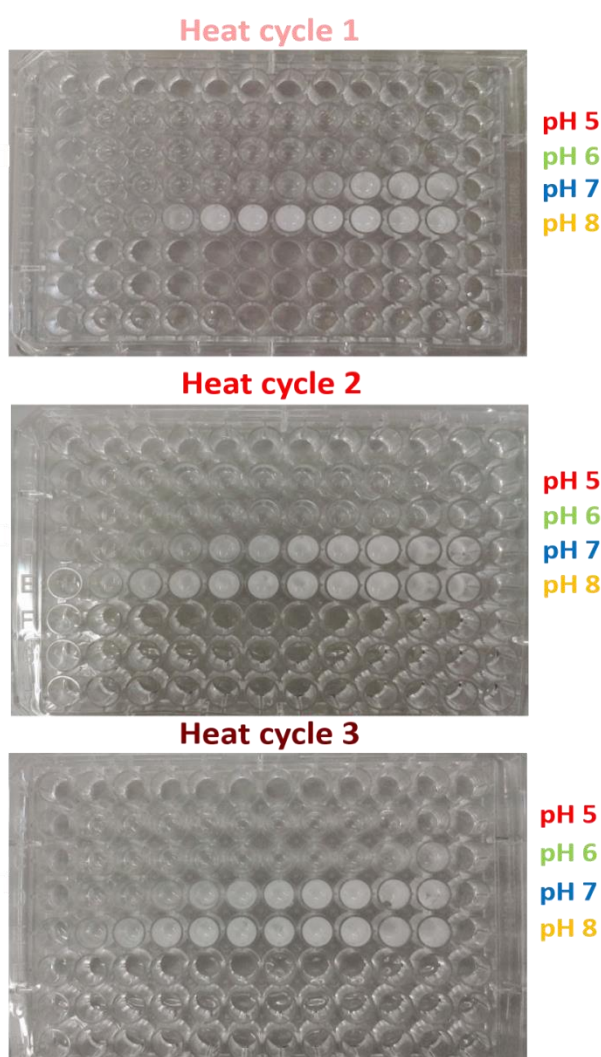


Figure 3.7: 96-multiwell plates showing thermal denaturation of lysozyme after heat cycles 1, 2 and 3.

For lysozyme, at pH 5 the overall turbidity of the solution remained low with minimal change in % transmission between successive heat cycles observed across the full concentration range (figure 3.8A). For lysozyme at pH 6, the lower concentration samples (0.5 – 6 mg/ml) only showed minimal change in % transmission with each successive

heat cycle (figure 3.8B). However the higher concentration samples (8 – 20 mg/ml) showed a noticeable increase in turbidity by cycle 2, and a more significant decrease in % transmission for the 15 and 20 mg/ml samples by cycle 3. Lysozyme pH 7 (0.5 – 12 mg/ml) and pH 8 (0.5 – 6 mg/ml), both showed a marked decrease in % transmission recorded between the first and second heat cycle, but none thereafter (figure 3.8C and D).

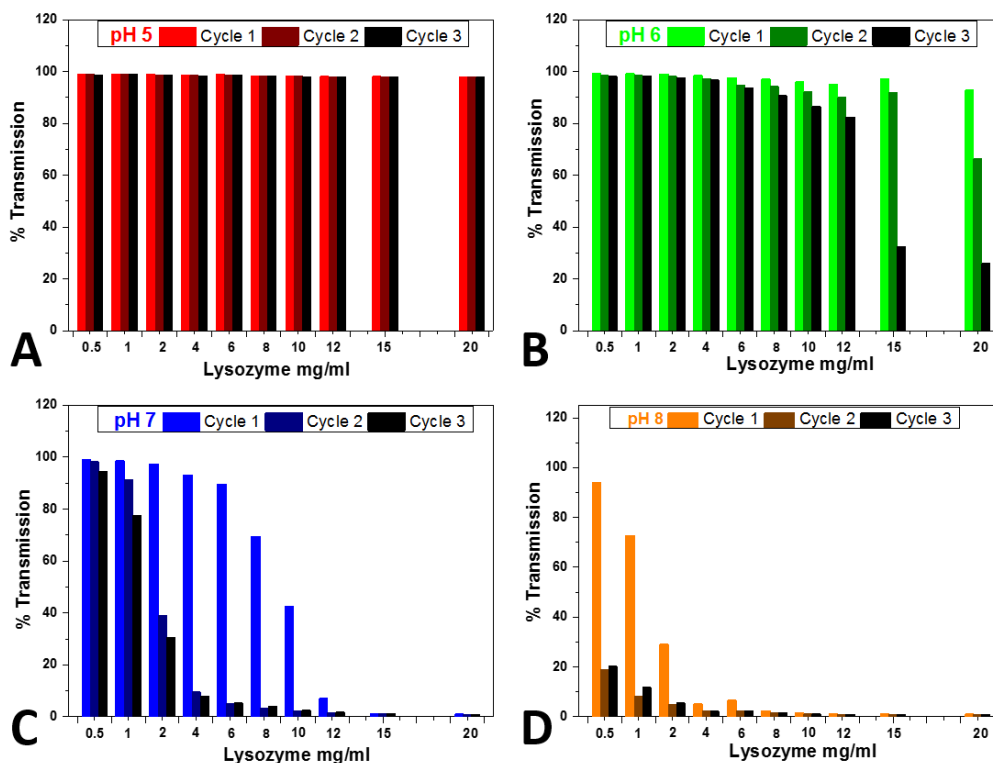


Figure 3.8: The % transmission of lysozyme with successive heat cycles.

A) pH 5. B) pH 6. C) pH 7. D) pH 8.

The BSA samples were also treated in this manner (microplates not shown). No substantial change in any of the samples could be visually observed after heat cycles 2 or 3, except for those that had already turned visually turbid after the first heat cycle. In these samples a slight increase in visual turbidity after the second and third heat cycles did occur.

The measured % transmission after successive heat cycles of BSA is shown in Figure 3.9. For BSA at pH 6 – 8, the level of turbidity remains low. A relatively small decrease in % transmission (0.5 – 5%) was measured after each heating cycle consistent across all BSA concentrations apart from the highest concentration (20 mg/ml) where % transmission decreases to 90% (figure 3.9 B – D). For BSA at pH 5, all samples showed a decrease in

% transmission with each successive heat cycle, especially in the 1 – 10 mg/ml samples after the second heat cycle, when compared to the initial heat cycle.

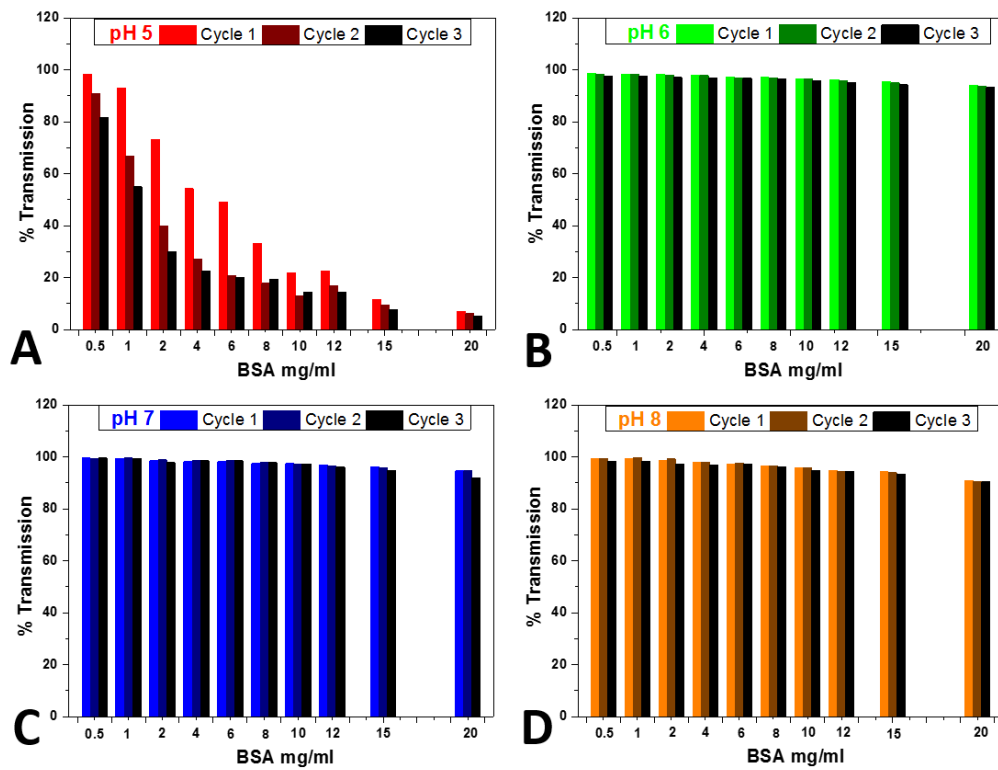


Figure 3.9: The % transmission of BSA with successive heat cycles.
 A) pH 5. B) pH 6. C) pH 7. D) pH 8.

If the data for both BSA and lysozyme are summarized based on the % transmission data alone at two concentrations, we see that for both proteins, the stability against thermal denaturation decreases as the protein approaches its pI (figure 3.10). This is consistent with previous work on both lysozyme (Bonincontro, De Francesco *et al.* 1998) and BSA (Tobitani and Murphy 1997). In the case of lysozyme it appears that the protein is thermally stable at pH 5 – 6, however between pH 7 – 8, stability decreases. While BSA remained virtually free from precipitation between pH 6 – 8, and precipitated at pH 5, indicating aggregation.

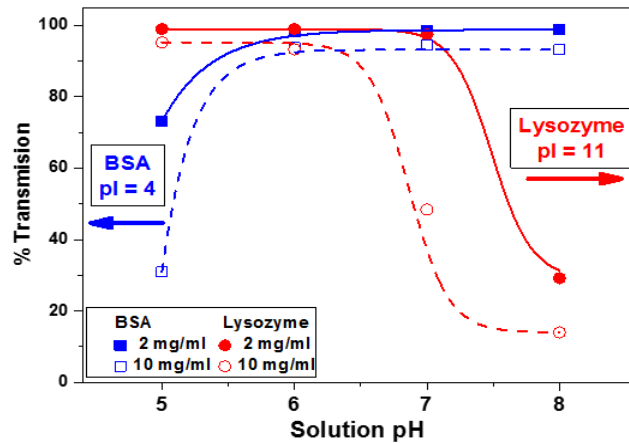


Figure 3.10: Turbidity assay data of lysozyme (red) and BSA (blue) at 2 mg/ml and 10 mg/ml indicating that both proteins are most stable against thermal denaturation at pHs far from their pI values.

3.3.5 Aggregate morphology

3.3.5.1 Appearance of the protein aggregates

Lysozyme formed an opaque, thick, dense, solid aggregate after denaturation that did not break up after the application of shear. The BSA opaque aggregate was a less solid material that remained free in suspension and could easily be pipetted. Figure 3.11 shows both the lysozyme and BSA aggregate pellet formed by thermal denaturation of a 20 mg/ml sample, after centrifugation.

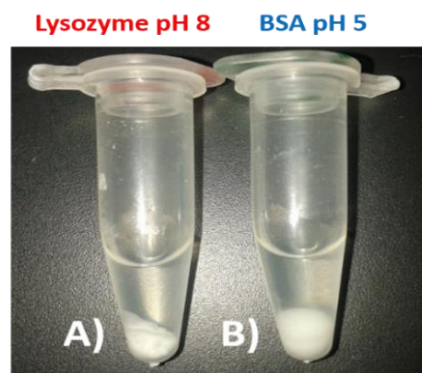


Figure 3.11: Thermally denatured aggregate pellet of A) lysozyme pH8, B) BSA pH 5.

The precipitated material for both lysozyme and BSA was examined in more detail using optical microscopy. This was to establish the morphology of the aggregated protein produced by thermal denaturation. The precipitated material from both proteins was visually different (figure 3.12). Lysozyme formed large clumps of spherical aggregated units stuck together, compared to free floating spherical structures in formed BSA.

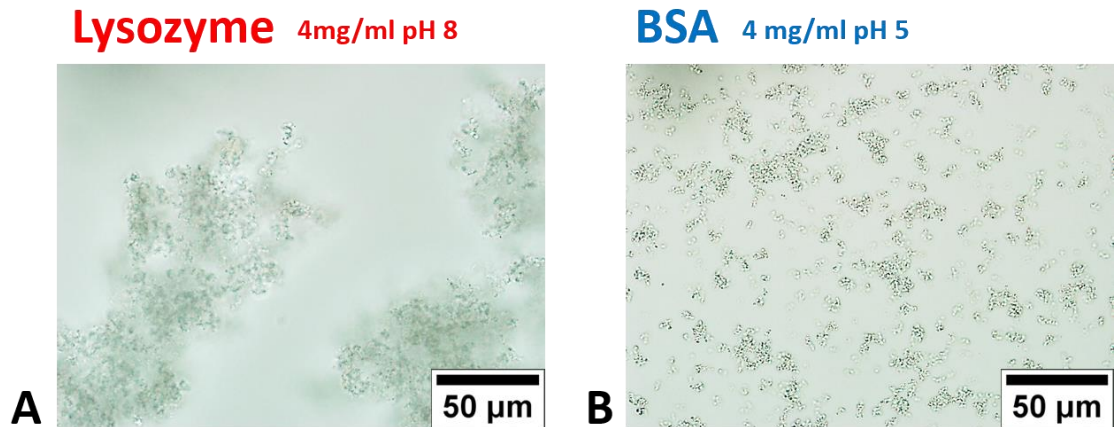


Figure 3.12: Optical micrograph of A) lysozyme and B) BSA for a 4 mg/ml thermally denatured aggregate.

3.3.5.2 Polarised light microscopy

Both amorphous aggregates and amyloid type material can be formed after thermal denaturation of proteins (Khurana, Gillespie *et al.* 2001; Vetri, Canale *et al.* 2007), and polarization light microscopy can be used to detect the presence of amyloid material, which can help to distinguish both of these aggregate types. The lysozyme aggregates indicated some degree of optical birefringence, suggesting that amyloid-type aggregates may have formed after denaturation (figure 3.13). This phenomenon was not observed with the BSA aggregates.

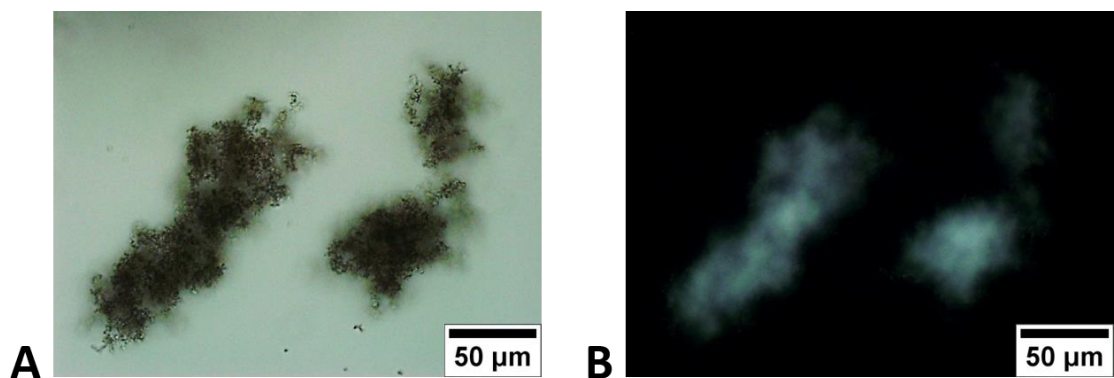


Figure 3.13: Micrographs of aggregated protein material produced after the thermal denaturation of lysozyme (4 mg/ml, pH 8). A) Phase contrast image B) Polarisation microscopy image.

3.3.6 High-throughput assay – Thioflavin T

The use of a Thioflavin T (ThT) assay is another convenient tool for high-throughput screening, based on the extrinsic fluorescence of a dye bound to an amyloid-type protein aggregate. The fluorescence intensity of Thioflavin T increases significantly when bound to the cross β -structure specific to amyloid fibrils. However small increases in fluorescence intensity can occur when bound to hydrophobic sites of amorphous aggregates and to amorphous aggregates showing β -cross structure but lacking fibril morphology (Chang, Liao *et al.* 2009; Sen, Fatima *et al.* 2009; Biancalana and Koide 2010). A ThT assay for both lysozyme and BSA was performed under the same conditions used for the % transmission measurements. ThT fluorescence was monitored over time for all protein samples after thermal denaturation.

For lysozyme in sodium phosphate buffer, pH 8, ThT fluorescence increased relative to protein samples at pH 5. Only in the higher protein concentration range (10 – 20 mg/ml) did the samples at pH 5, show an increase in ThT fluorescence (figure 3.14).

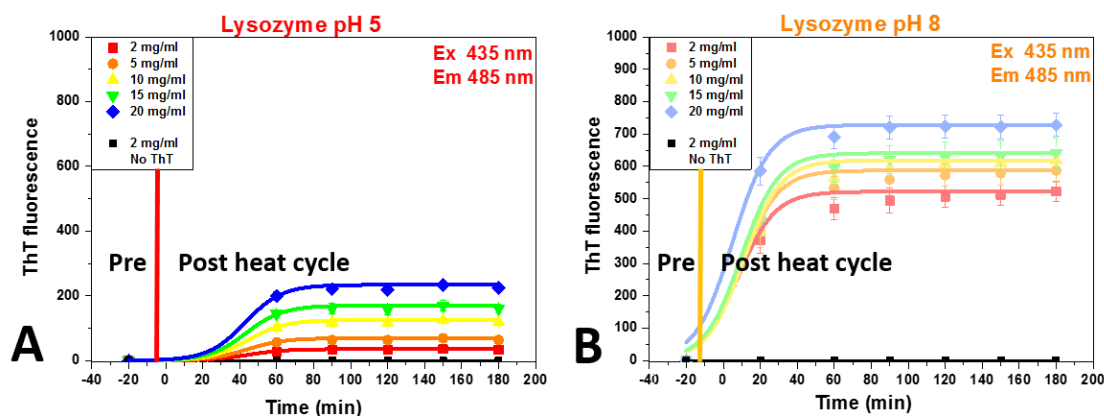


Figure 3.14: Thioflavin T assay measurements (normalised) of thermally denatured lysozyme. A) pH 5. B) pH 8.

BSA samples also showed increased ThT fluorescence after thermal denaturation at all protein concentrations (figure 3.15). Higher ThT fluorescence intensities were measured for protein at pH 5 than at pH 8. At pH 5, the measured ThT fluorescence intensity increases with increasing protein concentration. The data was normalized such that the intensity is divided by the protein concentration and represents the ThT fluorescence per milligram of protein, hence the increase in fluorescence intensity was related to the

association of ThT to aggregates beyond those expected by an increase in protein concentration.

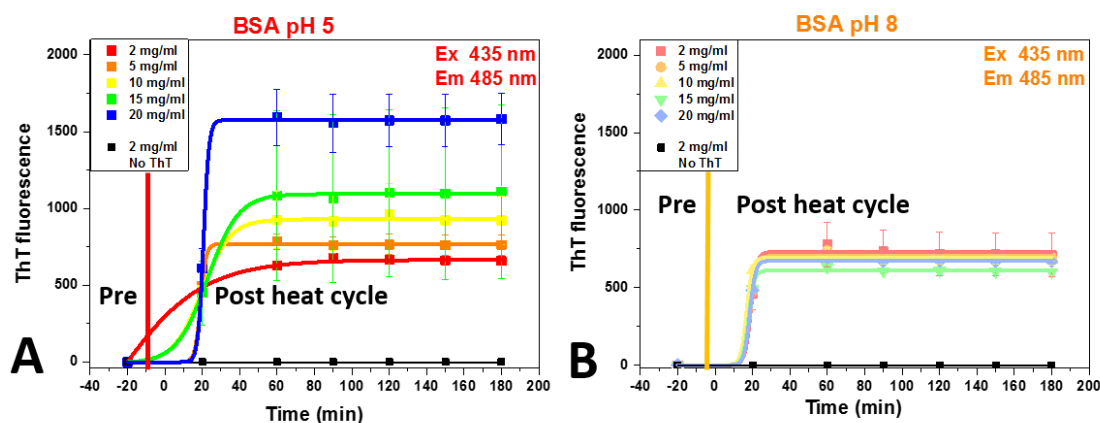


Figure 3.15: Thioflavin T assay measurements (normalised) for thermally denatured BSA. A) pH 5. B) pH 8.

Lysozyme and BSA both showed the highest increase in ThT fluorescence in the samples that also had the biggest change in % transmission (figure 3.16). This was also related to the solution condition and pI of each individual protein.

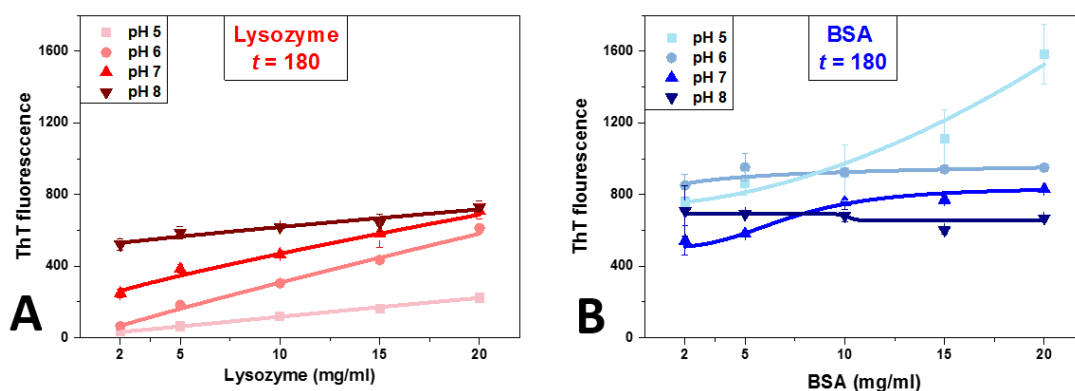


Figure 3.16: Thioflavin T fluorescence (normalised) of A) lysozyme and B) BSA, after thermal denaturation ($t = 180$ min).

3.3.7 SE-HPLC

SE-HPLC was chosen as a technique to compare the effectiveness of the high-throughput assays in characterising and detecting protein aggregation after thermal denaturation. The bulk precipitated aggregates were removed from the protein samples by pelleting them from the protein solution by centrifugation. The remaining protein in the supernatant was examined using SE-HPLC.

3.3.7.1 SE-HPLC of lysozyme

The supernatant of lysozyme samples after one heat cycle were subjected to SE-HPLC on day 1 (after separation from the pellet) and also on days 3 and 7, at solution pH 5 (figure 3.17A). At this pH, the protein was relatively colloidal stable and no aggregation after thermal denaturation was detected using either % transmission measurements or the ThT assay.

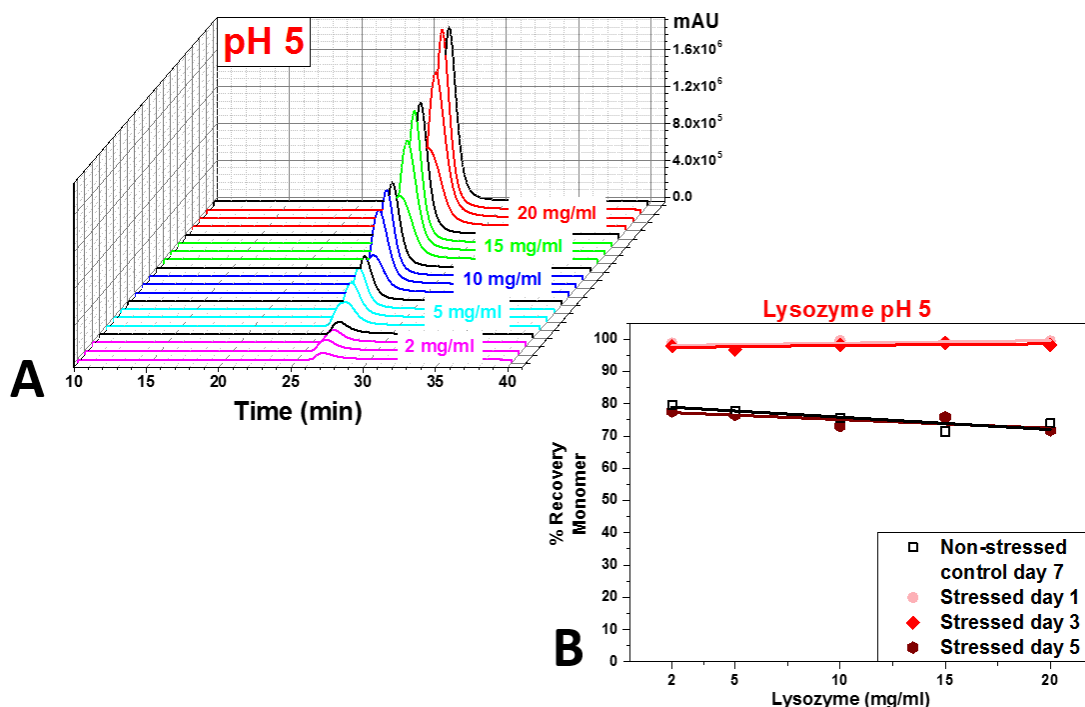


Figure 3.17: SE-HPLC data for the lysozyme supernatant (2, 5, 10, 15 and 20 mg/ml) at pH 5, after one cycle of thermal denaturation on day 1, 3 and 7. A) Lysozyme chromatograms showing the monomer peak and no indication of higher molecular weight species. B) Percentage recovery of lysozyme monomer.

Plotted alongside each thermally denatured sample is a corresponding lysozyme control sample (data in black). This represents a chromatogram for a lysozyme sample of equal concentration to the protein sample before thermal denaturation for comparison. No higher molecular weight species were detected on the column. A clear monomer peak was detected at 26.3 minutes in all samples. Figure 3.17B plots the percentage recovery of monomer (calculated from area under each chromatogram peak) on each day. Also plotted for comparison is the percentage recovery of lysozyme control solutions (pH 5) at each concentration that wasn't subjected to any sort of stress or denaturation on day 7. From this graph on day 1, for all concentrations, almost 100% of lysozyme at pH 5 was recovered from the supernatant (*i.e.* there is no precipitation) after thermal denaturation,

which is consistent with the data from the turbidity measurements. With time, the percentage of lysozyme monomer decreased but the same amount of material was lost in the non-stressed sample. This is likely to be as a result of a decrease in colloidal stability over time. This data indicates that after thermal denaturation, the recovery of lysozyme monomer at pH 5 was very high and that the stability of the recovered protein did not appear in any way to be diminished or affected by the imposition of a thermal stress. Table 3.3 shows the recovery of protein (mg/ml) in the supernatant (column 1) and the calculated recovery of monomer lysozyme (mg/ml) recovered (column 2; as per area under the SE-HPLC chromatogram). This confirms that the recovery of protein by SE-HPLC was consistent with what was recovered in the supernatant and that almost no lysozyme was lost to aggregation by thermal denaturation. A small amount of lysozyme was unaccounted for (column 3), but is within the experimental error (~1%).

| <i>Initial Lysozyme conc (mg/ml)</i> | <i>Lysozyme supernatant</i> | <i>SE-HPLC monomer</i> | <i>Lysozyme unaccounted/ > 600 kDa</i> |
|--|---------------------------------|----------------------------|---|
| 2 | 1.97 | 1.97 | 0.03 |
| 5 | 4.92 | 4.91 | 0.1 |
| 10 | 9.95 | 9.96 | 0.05 |
| 15 | 14.9 | 14.86 | 0.1 |
| 20 | 19.95 | 19.90 | 0.05 |

Table 3.3: Lysozyme pH 5, concentration of lysozyme monomer resolved by SE-HPLC and in the supernatant (per UV measurement) (mg/ml) after thermal denaturation day 1.

The supernatant of lysozyme at pH 8 was also analysed using SE-HPLC (figure 3.18). Both the turbidity and ThT measurements indicated significant protein aggregation at this pH, while a 15 – 30% loss to aggregation was measured by calculating the concentration of protein remaining in the supernatant. The lysozyme monomer can be identified as the main peak at 26.3 minutes, and again there was no indication of higher molecular weight species. The chromatograms in black represent control samples of lysozyme not subjected denaturation at this pH.

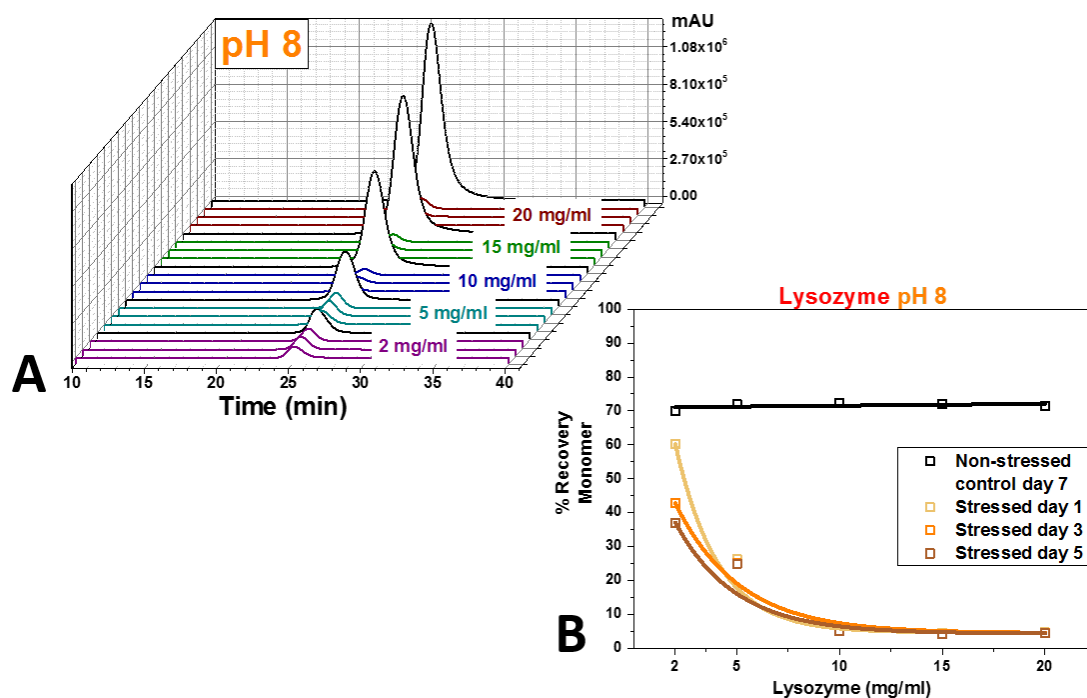


Figure 3.18: SE-HPLC data of lysozyme supernatant (2, 5, 10, 15 and 20 mg/ml) at pH 8 after one cycle of thermal denaturation on day 1, 3 and 7. A) Lysozyme chromatograms showing monomer peak and no indication of higher molecular weight species. B) Percentage recovery of monomer.

There was a significant loss in monomeric protein in the thermally stressed samples compared to non-stressed samples. The recovery of monomer was lower than expected since the concentration of protein in the supernatant indicated 70 - 85% recovery of protein. The quantity (mg/ml) of lysozyme recovered by SE-HPLC and in the supernatant did not match up. Table 3.4 shows the recovery of protein (mg/ml) in the supernatant (column 1) and the calculated recovery of monomer (mg/ml) recovered (column 2; as per area under the SE-HPLC chromatogram). The concentration difference suggested a population of lysozyme aggregates too large to be resolved on the column (> 600 kDa) but too small to be observed by the turbidity assay (< 1 μ m) or removed by centrifugation (column 3). The quantity of lysozyme precipitate was calculated as the difference between lysozyme recovered in the supernatant and the initial lysozyme concentration (column 4).

| <i>Initial Lysozyme conc</i> | <i>Lysozyme supernatant</i> | <i>SE-HPLC monomer</i> | <i>>600 kDa, <1 μm</i> | <i>Lysozyme precipitate > 1 μm</i> |
|------------------------------|-----------------------------|------------------------|---|--|
| 2 | 1.58 | 1.20 | 0.38 | 0.42 |
| 5 | 3.67 | 0.84 | 2.86 | 1.33 |
| 10 | 7.94 | 0.57 | 7.37 | 2.06 |
| 15 | 10.86 | 0.67 | 10.19 | 4.14 |
| 20 | 10.03 | 0.91 | 9.12 | 9.97 |

Table 3.4: Lysozyme pH 8, concentration of monomer resolved by SE-HPLC and in the supernatant (per UV measurement) (mg/ml) after thermal denaturation day 1.

3.3.7.2 SE-HPLC of BSA

BSA in its native, unstressed form has a SE-HPLC chromatogram with two peaks; a monomer peak with a retention time of 18.9 minutes (~88% of total protein content) and a small portion of dimer with a retention time of 16.39 minutes (~12% of total protein content) (figure 3.19).

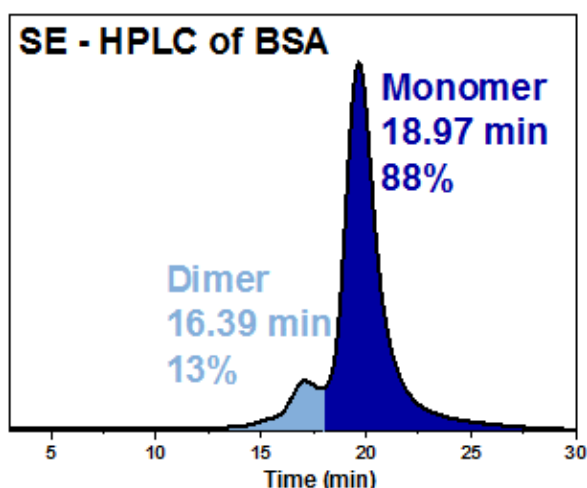


Figure 3.19: SE-HPLC of a BSA (10mg/ml, 50 mM NaPo₄, pH 8) sample showing the proportion of monomer to dimer.

SE-HPLC was performed on the supernatant of BSA samples after one heat cycle on day 1 (after separation from pellet) and also on day 3, day 5 and day 7, for samples at pH 5 (figure 3.20A). A non-stressed BSA sample is also plotted at each concentration (black chromatograms). All the BSA supernatant samples showed a decrease in monomer protein, very low levels of dimer species (less than 1%) and no evidence of higher molecular weight aggregates within the column resolution. The % recovery of monomer (calculated from the area under each chromatogram peak) for each day is shown in Figure 3.20 B. Also plotted for comparison is the percentage recovery of lysozyme control

solutions (pH 5) at each concentration that wasn't subjected to any sort of stress or denaturation on day 7.

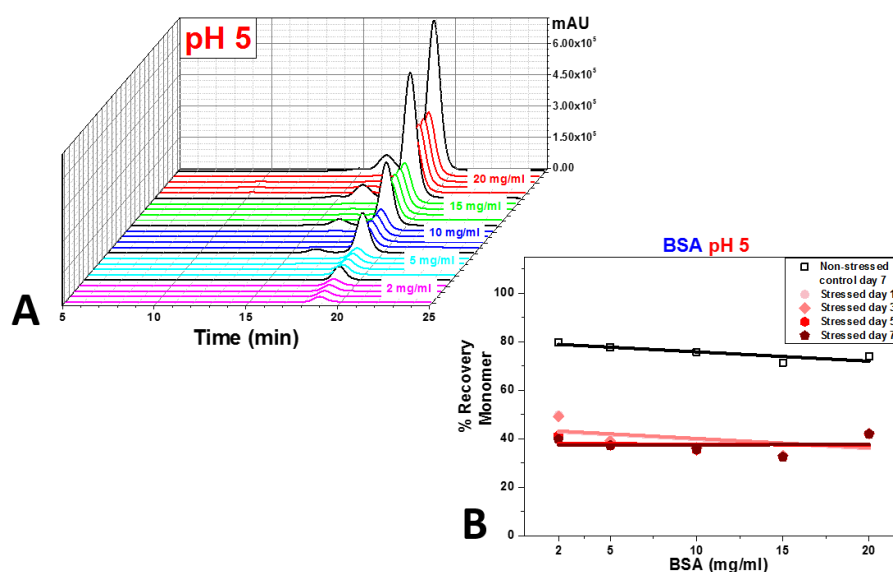


Figure 3.20: SE-HPLC data of BSA supernatant (2, 5, 10, 15 and 20 mg/ml) at pH 5, after one cycle of thermal denaturation on day 1, 3, 5 and 7. A) Chromatograms showing monomer peak and dimer peak at 16.39 minutes. B) Percentage recovery of monomer BSA.

The supernatant of BSA shows between 30 – 50% recovery of monomeric protein after thermal denaturation. There does not appear to be any further loss in the recovered monomer after separation from the bulk aggregated material up to 7 days later indicating good stability of the recovered monomer. The proportion of monomer loss was also consistent across all BSA concentrations. The amount of BSA recovered by SE-HPLC was compared to the amount recovered in the supernatant. Table 3.5 shows the calculated recovery (mg/ml) of BSA recovered in the supernatant (column 1) and the recovery of protein by SE-HPLC (column 2 and 3; as per area under the SE-HPLC chromatogram).

| <i>Initial BSA conc</i> | <i>BSA supernatant</i> | <i>SE-HPLC monomer</i> | <i>SE-HPLC ≥ dimer, < 600 kDa</i> | <i>>600 kDa, <1 μm</i> | <i>BSA precipitates</i> |
|-------------------------|------------------------|------------------------|--------------------------------------|------------------------------|-------------------------|
| 2 | 0.99 | 0.88 | 0 | 0.11 | 1.01 |
| 5 | 2.05 | 1.84 | 0.02 | 0.19 | 2.95 |
| 10 | 3.17 | 3.02 | 0.06 | 0.09 | 6.83 |
| 15 | 4.89 | 4.24 | 0.10 | 0.55 | 10.11 |
| 20 | 8.43 | 7.36 | 0.33 | 0.74 | 11.57 |

Table 3.5: BSA pH 8, concentration of lysozyme monomer resolved by SE-HPLC and in the supernatant (mg/ml) after thermal denaturation day 1.

From this table it appears that while SE-HPLC does detect very low levels of dimer, a very small proportion of BSA (< 0.55 mg/ml) has formed a population of aggregates outside the limits of detection for SE-HPLC and the detection range of the turbidity assay (column 4), with the greatest proportion of BSA at pH 5, going on to form large precipitated aggregates as detected by the turbidity assay (column 5).

BSA samples at pH 8 was analysed in the same manner (figure 3.21A). A relatively small amount of monomer was detected in the supernatant by SE-HPLC, while a range of higher molecular weight peaks were observed, up to the maximum size resolution of the column (600 kDa). As little as 10 - 15% monomer was measured in the supernatant at all BSA concentrations after thermal denaturation (figure 3.21B).

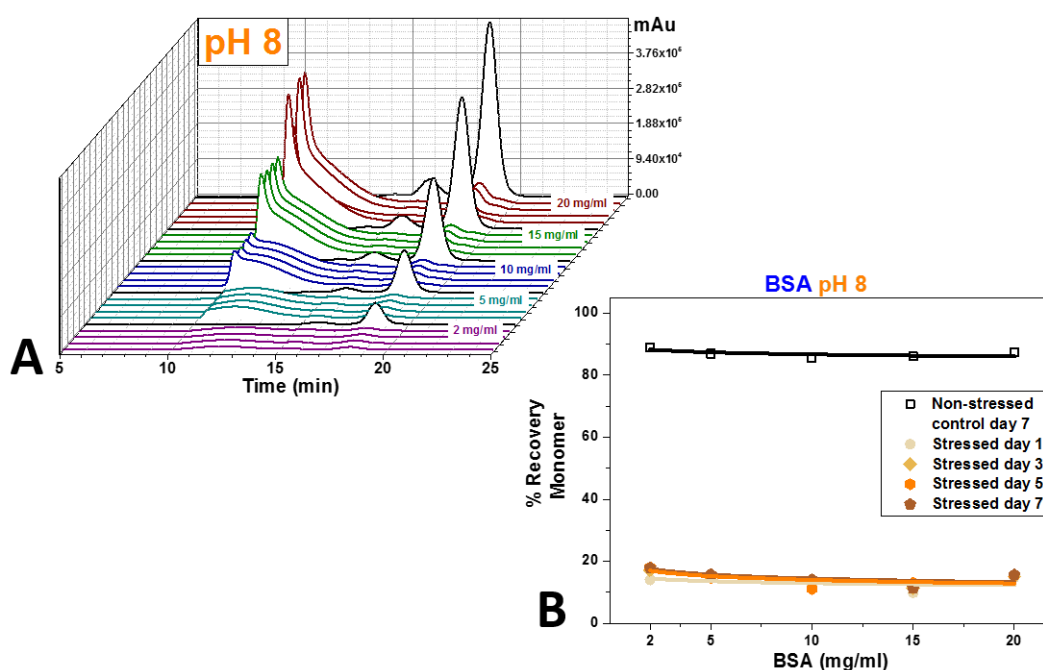


Figure 3.21: SE-HPLC data of BSA supernatant at pH 8 after one cycle of thermal denaturation on day 1, 3, 5 and 7. A) Chromatograms showing a monomer peak and higher molecular weight species to the left of the monomer. B) Percentage recovery of monomer.

Again the proportion of each BSA species (mg/ml) was calculated as per Table 3.6. The recovery of monomer and the proportion of higher molecular weight species (< 600 kDa) from SE-HPLC (column 2 and 3) was compared to the recovery of protein in the supernatant (column 1). Since no visual aggregation was detected it was assumed that the majority of the protein aggregates were smaller than ~1 μm accounting for a very small

proportion of aggregates out-side the detection limit of the column (column 4), but the majority of BSA in pH 8 that aggregated by thermal denaturation formed aggregates between 130 and 600 kDa (column 3).

| Initial BSA conc | BSA supernatant | SE-HPLC monomer | SE-HPLC ≥ dimer, < 600 kDa | >600 kDa, <1 μm |
|-----------------------------|----------------------------|----------------------------|--|----------------------------------|
| 2 | 1.81 | 0.24 | 1.04 | 0.58 |
| 5 | 4.83 | 0.65 | 4.07 | 0.11 |
| 10 | 9.47 | 1.04 | 7.17 | 1.26 |
| 15 | 14.47 | 1.27 | 11.88 | 1.22 |
| 20 | 18.95 | 1.57 | 14.63 | 2.75 |

Table 3.6: Lysozyme pH 8, concentration of lysozyme monomer resolved by SE-HPLC and in the supernatant (mg/ml) after thermal denaturation day 1.

3.4 Discussion

High-throughput screening assays used in this study enabled the testing of large numbers of protein samples and solution conditions, to gauge their response to a thermal stress. In this study we verified whether the high-throughput assays used here were capable of detecting the levels and types of aggregation induced by the thermal denaturation of two model protein, lysozyme and BSA.

The first assay assessed was the turbidity assay. After samples were subjected to heat denaturation (Mechanism 2), protein aggregation occurred. A number of the protein solutions turned visibly turbid, indicating the presence of large protein aggregates which was confirmed with optical microscopy. This change was measured as a marked decreases in % transmission coinciding with the samples that also showed a visual change in solution appearance. The degree to which the turbidity assay could differentiate between levels of turbidity was broken down into three categories; a large decreases in % transmission coinciding with visually high levels of turbid material (1 – 80% % *T*), a slight decreases in % transmission accounting for low levels of subvisible particles (80 – 98% % *T*) and no loss in % transmission indicating no change in solution appearance. This shows that the turbidity assay is useful as a HTS assay for establishing a “phase diagram” for a range of solution conditions.

The % transmission data for the range of solution conditions examined is consistent with previous work. This pH-dependent association of lysozyme observed has been well documented (Bruzzezi.Mr, Chiancon.E *et al.* 1965; Hampe 1972; Mine, Noutomi *et al.* 1990; Goldberg, Rudolph *et al.* 1991; Ibrahim, Higashiguchi *et al.* 1996; Price, Tsuchiya *et al.* 1999; Cao, Wang *et al.* 2002; Raccosta, Manno *et al.* 2010; James and McManus 2012).

Many aspects of BSA thermal denaturation have been examined, indicating solution-dependent unfolding related to solution conditions including ionic strength and pH (Ross, Finlayson *et al.* 1984; Shrake, Finlayson *et al.* 1984; Yamasaki, Yano *et al.* 1990; Michnik 2003; Michnik, Michalik *et al.* 2006). This work is in general consistent with our measurements.

However, the turbidity assay does not give any detailed information about the type, size or quantity of aggregation and was actually misleading for several samples (on the basis of further analysis). It was found that within the region where a large decrease in % transmission were recorded for lysozyme at pH 8, did not correlate well with the actual loss of protein to aggregation. Additionally, once the % transmission reaches zero (the saturation point), no further changes in aggregation behaviour can be detected, even when further aggregation continues to occur over time. Hence, the technique is only appropriate for a range of relatively low concentration samples. In fact, determination of the protein concentration in the sample supernatant after removal of bulk aggregates was much more informative than the turbidity assay. The measured values were a more accurate representation of the actual loss of protein to aggregation. Additionally, heterogeneity within the sample is not captured by the measurement, since % transmission is assessed at a single point.

The assay also seemed limited to only detecting large aggregates. A small proportion of large aggregates in lysozyme at 2 mg/ml was detected, but a higher concentration of aggregates in a 4 mg/ml indicated no aggregation by % transmission. The limits of the assay to differentiate between the sizes of aggregates was most apparent for BSA. Since the turbidity assay clearly had limits of detection, it was compared to other orthogonal techniques to further assess its sensitivity at detecting aggregation of protein samples. BSA at pH 6 – 8, that remained visually transparent even with almost 50% of the material was detected as small aggregates by SE-HPLC.

As a stand-alone method of detection, % transmission measurements would suggest that for lysozyme at pH 5 – 6, was resistant to aggregation relative to the thermally unstable pH range (pH 7 and 8). If this is coupled with the SE-HPLC analysis, we were able to establish that stability is in fact only maintained at pH 5. From the SE-HPLC analysis of both proteins it became obvious that not only does the turbidity assay not detect large amounts of small aggregates, it is unable to detect high levels of protein aggregates in higher concentration samples (*i.e.* once aggregation is sufficient to bring the % transmission to zero). Unlike the turbidity assay, SE-HPLC can quantitatively detect different forms/sizes of protein and when coupled to the measurement of protein concentration in the supernatant of separated samples, provides a good description of the overall level of aggregation within a sample.

BSA is a multi-domain protein and it has been proposed that these domains unfold independent of one another and at slightly different temperatures resulting in the formation of intermediate states from which aggregation is likely to occur (Fink 1998; Michnik 2003; Militello, Vetri *et al.* 2003). The SE-HPLC of BSA (pH range 6 – 8) did detect the formation of small oligomers likely formed as the thermally unfolded subunits of one molecule interacted with a matching exposed hydrophobic surface of a neighbouring molecule. As the solution pH moves towards the *pI* of BSA (pH 5), the charge distribution on the surface of BSA was reduced and the small soluble oligomers further assemble into more disordered aggregates with the loss of electrostatic effect stabilisation. Eventually the size of the aggregates exceeds the solubility limit and large visible aggregates are formed, as observed visually and with the turbidity assay, consistent with previous work (Clark, Judge *et al.* 1981; Clark, Kavanagh *et al.* 2001; Militello, Casarino *et al.* 2004; Vetri, Librizzi *et al.* 2007).

A second high-throughput assay utilising the dye Thioflavin T was assessed. The polarised light micrographs of the lysozyme aggregates indicated birefringence leading to the assumption some level of ordered structure (possible amyloid aggregates) exists within the aggregation, which the BSA aggregates did not show. The fluorescence intensity of the ThT with thermally denatured lysozyme did show an increase, indicative of association with amyloid structure (Krebs, Bromley *et al.* 2005; Biancalana and Koide 2010) consistent with those samples that indicated the presence of precipitated aggregates (at pH 8) in the turbidity assay. Taken together, this suggests that lysozyme aggregates are amyloid-like to some extent, but TEM images are required to confirm this. Human lysozyme is capable of forming amyloid fibrils in individuals suffering from non-neuropathic systemic amyloidosis as a result of a point mutation in the lysozyme gene. HEWL (hen egg white lysozyme), wild-type human and wild-type hen lysozyme have all been shown to form amyloid fibrils confirmed by EM, Congo red and ThT fluorescence (Krebs, Wilkins *et al.* 2000; Morozova-Roche, Zurdo *et al.* 2000; Cao, Hu *et al.* 2004; Frare, Polverino De Laureto *et al.* 2004). The formation of amyloid fibrils from lysozyme has only been reported using only extreme conditions of pH, temperature and incubation times, suggesting that the experimental conditions used in our study are likely not harsh enough to induce the conformational changes needed to form fibrils. As an example, amyloid fibrils of lysozyme were formed by incubating the protein at 81 °C for 45 minutes

in H₂O at pH 3.8, followed by lyophilisation. Samples were then stirred in 32 mM HCl, pH 1.5, at 45°C for up to ten days (Frare, Mossuto *et al.* 2006).

Another point against the assumption of amyloid fibril formation was the comparison to the analysis of BSA samples incubated with ThT. The BSA polarised microscopy images show that there is no presence of ordered structure in the aggregates formed. However the BSA samples in the ThT assay show a higher amount of ThT fluorescence upon thermal denaturation than for lysozyme. While the exact mechanism of ThT association with amyloid fibrils is still not fully understood it is believed to associate with, (but not exclusively), to the crossed- β structure in amyloid fibrils (Krebs, Bromley *et al.* 2005). Protein aggregated into coagulated gel structures is known to involve intermediate protein conformations with higher β structure content, compared to native forms (Kelly 1998). Thermally denatured irreversible aggregates of BSA have been shown to proceed by a conversion of the protein secondary structure from α -helix into well ordered β -sheet rich aggregates (Murayama and Tomida 2004). These types of aggregates have been shown to bind ThT and Congo red, both indicators of fibrillation and aggregates with amyloid properties, however the aggregates didn't possess the same structural rigidity as classical amyloid fibres (Holm, Jespersen *et al.* 2007). Heat denatured ovalbumin has been shown to induced opaque amorphous precipitates that developed cross- β structure as recognised by ThT (Azakami, Mukai *et al.* 2005). These possibility may account for why we see a rise in ThT fluorescence within our samples but it is not believed to be a result of fibril formation. Typical fibrillation has a sigmoidal profile usually commencing with an initial lag time followed by a relatively short elongation growth phase where the fibrils accumulate. The ThT intensity values measured by the ThT assay were minimal compared to ThT fluorescence intensities characteristic of fibril formation, further indicating that the ThT fluorescence is most likely as a result of association with aggregated material and not amyloid fibrils (LeVine 1999). Therefore the ThT assay is useful in detecting protein aggregates (of both amorphous and amyloid type) and does in fact detect aggregation that is not indicated by the % transmission assay.

3.5 Conclusions

High-throughput screening methods have both pros and cons. As a first round, positive/negative screening method for solution conditions where high levels of aggregation occur, our turbidity assay is suitable and reliable to establish a first-pass phase diagram. It can decipher between solution conditions where protein aggregates precipitated and those that did not. However analysis of the % transmission data was misleading about how much aggregation was occurring. It was also unable to differentiate between the different sizes of protein aggregates formed. For this more robust and sensitive techniques need to be applied with a sample-by-sample approach. The ThT assay for our purposes had inconsistencies, illustrating another downside of relying on high throughput screening. Using this assay as a standalone result, both lysozyme and BSA would test positive for the presence of amyloid fibrils. By increasing the methods of detection and analysis, and comparing the results of BSA to lysozyme a much more likely picture of the type of aggregated material for each protein can be concluded. Collectively the data for BSA and lysozyme does not point towards fibril formation. Hence high throughput technologies do have a role to play in formulation and screening, but the results must be viewed in context. An orthogonal approach using a number of different techniques (including sample-by-sample measurements) must also be conducted if all potentially useful formulations are to be identified.

Chapter 4 Oxidation as a Pathway to Protein Aggregation

4.1 Protein Oxidation

Protein oxidation occurs as a result of either direct “attack” by a reactive oxidative species (ROS) or indirectly through peroxidation of lipids that further degrade and attack proteins (Mirzaei and Regnier 2008). The rate of radical attack in proteins is of great interest, due to the linkage between oxidative stress and many pathological disorders, the process of aging and in biopharmaceuticals. Numerous studies point to the aromatic residues and sulphur containing amino acid residues (methionine and cysteine) being most sensitive to oxidative attack (Kim, Berry *et al.* 2001; Lasch, Petras *et al.* 2001; Guedes, Vitorino *et al.* 2009; Ji, Zhang *et al.* 2009; Le, Chaffotte *et al.* 2009).

Hydrogen peroxide is a well-studied oxidizing agent, known to be specific for oxidation of methionine and cysteine (Shechter, Burstein *et al.* 1975; Ji, Zhang *et al.* 2009). Understanding protein sensitivity to H₂O₂ is of significant value. Proteins are not only susceptible to H₂O₂ attack through cellular process and events but also during process and formulation during biopharmaceutical production. Protein formulations containing aged polysorbate 80 showed signs of oxidation as a result of H₂O₂ (Ha, Wang *et al.* 2002) while H₂O₂ is also used as an aseptic agent for the insulators used in the filling of sterile products, consequently increasing the proteins susceptibility to oxidative damage (Ji, Zhang *et al.* 2009).

Chemical modification (Mechanism 4) as a consequence of oxidation leads to alterations in the protein structure. The major problem with defining the chemistry arises from the complex chemical composition of the protein surface offering a wealth of targets for oxidation-induced reactions. Garrison *et al.* (Garrison 1987), Schuessler and Schilling (Schuessler and Schilling 1984), and Kopoldova and Liebster (Kopoldova and Liebster 1963; Liebster, Kopoldova *et al.* 1963), all pioneered work on the radical-mediated oxidation of proteins noting that it leads to oxidation of amino side chains, fragmentation of the polypeptide chain, and the generation of protein-protein cross linkages (Stadtman 2006). Oxidative modifications leads to a variety of structural conformational changes including the formation of higher molecular weight dimer products through mixed disulphides, inter- and intra-disulphides, and intramolecular abstraction at the α -carbon from the hydrogen of amino acid residues typically resulting in backbone fragmentation (Finnegan, Linley *et al.* 2010).

Garison *et al.* (Garrison 1987), proposed that oxidative cleavage of the protein backbone is initiated by the hydroxyl radical $\bullet\text{OH}$, formed during the radiolysis of water or the metal-catalysed cleavage of H_2O_2 as illustrated in Figure 4.1 (reaction a and b). This radical is capable of intramolecular α -hydrogen abstraction of amino acid residues, leading to the formation of α -amino carbon-centered radicals (reaction c). These radicals can undergo rapid addition of O_2 to form an alkylperoxy radical derivative (reaction d) that is readily converted to the peroxide (reaction f) and subsequently to the alkoxy derivatives (reaction h), that can then be converted to a hydroxyl protein derivative (reaction i) mediated by reactions with either $\text{HO}_2\bullet$ or Fe^{2+} (reaction e, g and j).

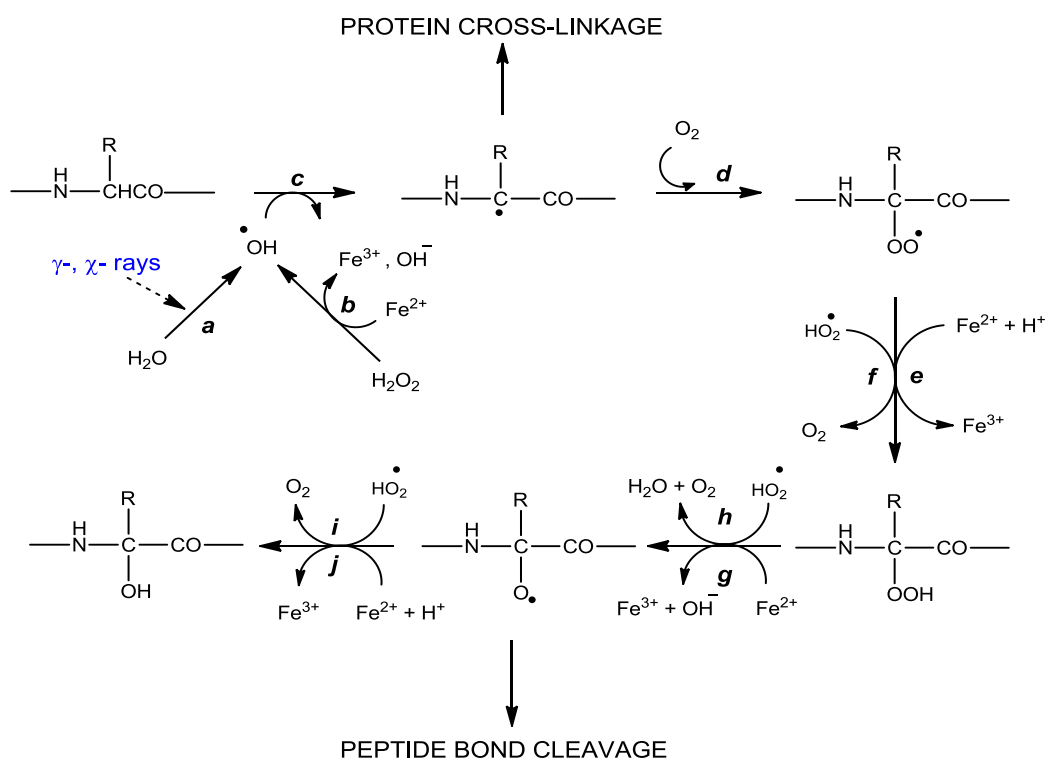


Figure 4.1: Role of ROS in the oxidation and cleavage of the protein backbone (Berlett and Stadtman 1997).

The generation of alkoxy radicals (reactions h and g) sets the stage for peptide bond cleavage by one of two pathways;

1. Diamide (figure 4.2, pathway a); the *C*-terminal amino acid of the peptide fragments derived from the *N*-terminal portion of the protein is present as a diamide derivative, and the *N*-terminal amino acid of the fragment derived from the *C*-terminal portion of the protein is present as an isocyanate derivative.

2. α -amidation (figure 4.2, pathway b); the C-terminal amino acid residue of the fragment derived from the N-terminal portion of the protein exists as an amide, and the N-terminal amino acid of the fragment derived from the C-terminal portion of the protein exists as an N- α -ketoacyl derivative.

With acid hydrolysis, peptide fragments obtained by the diamide pathway will yield CO₂, NH₃, and a free carboxylic acid, whereas hydrolysis of the fragment obtained by the α -amidation pathway yields NH₃ and a free α -ketocarboxylic acid.

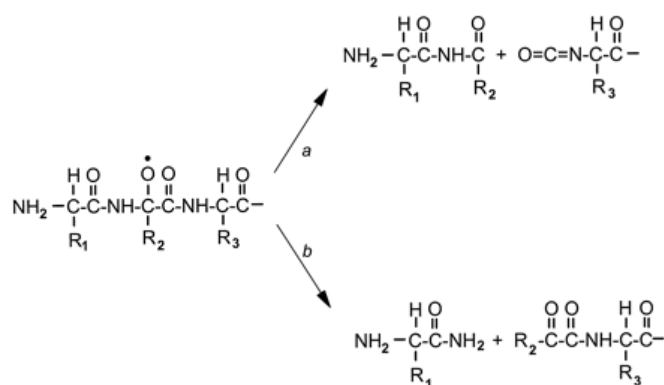


Figure 4.2: Cleavage of the protein backbone by the diamide pathway (a) and the α -amidation pathway (b) (Berlett and Stadtman 1997).

This idea of oxidative modification is supported by the work of Davies and co. (Davies, Lin *et al.* 1987). However the Schuessler and Puchalas model of oxidative modification of proteins proposes that the fragmentation of BSA involves site specific oxidative destruction of proline residues (Puchala and Schuessler 1993). Wolff *et al.* (Wolff, Garner *et al.* 1986) further explored the susceptibility of proline to oxidative attack showing that peptide bond hydrolysis occurs. However, Dean *et al.* re-assessed Schuesslers *et al.* fragment size data of BSA to state that it was more likely that glycine is the primary residue of attack over proline. They found that over 90% of the terminal amino acids after fragmentation by the OH• radical were glycine (Dean, Fu *et al.* 1997). While there appears to be conjecture over the exact mechanism of attack and whether or not it is site specific, it is apparent that the primary, secondary and tertiary structure of the protein under oxidative attack is modified and that this process can lead to the formation of protein aggregates.

The amino acid cysteine (Cys) is highly susceptible to oxidative stress; some consider it the most potent nucleophile of all amino acids. The cysteine residue is present in most proteins and often plays a fundamental role in the tertiary structure through the formation of the dimer, cystine. This disulphide (sulphur – sulphur) linkage involves the interconversion between thiol and disulphide groups in a redox reaction (figure 4.3); the thiol is the reduced state and the disulphide is the oxidised state. In the disulphide state each sulphur atom loses an H-bond and gains an S-bond.

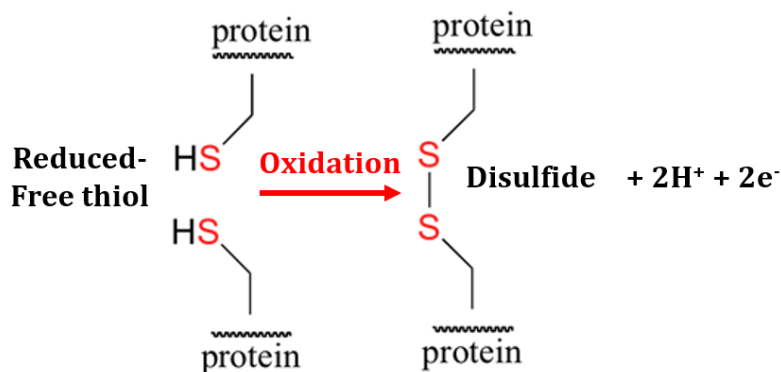


Figure 4.3: Disulphide bond formation by the coupling of two thiol groups with the formation of S–S bond.

The side group of Cys reacts rapidly with HO• in a one electron oxidation reaction to yield Cys thiol radicals (RS*), that can form the disulphide radical anions (RSSR*⁻) through a reaction with an additional thiol/thiolate, known as thiol-disulphide exchange:



It can also form the thiyl peroxy radicals (ROSS*) through the reversible reaction with O₂ and the reaction with ascorbate (Asc⁻);



In the case of protein oxidation, reactions (4.1) and (4.3) can be less efficient based on the access of CysS* to additional protein Cys residues and/or endogenous glutathione or ascorbate. Reaction (4.2) doesn't represent an efficient pathway for the removal of CysS* radicals, specifically at physiological concentrations of O₂. Hence competitive reactions of CysS* may become important in proteins (Schoneich 2008). Mild oxidation of cysteine can generate, inter- or – intra-molecular disulphides, protein mixed disulphides resulting in macroscopic variation of protein molecular mass, as well as aggregation. Numerous

low resolution techniques have been employed to monitor such structural changes as a result of oxidation such as mixtures of protein dimers (S-S linked) and larger aggregates including 2D gel electrophoresis (Molinari and Helenius 1999; Lepedda, Zinellu *et al.* 2013) and SE-HPLC (Mirzaei and Regnier 2008; Giustarini, Dalle-Donne *et al.* 2011).

Work by Puchala and Schuessler has added greatly to the understanding of oxidative modification to proteins. The aim of their work was to understand the role of oxygen in the radiolysis of proteins by comparing and contrasting the degradation and formation of products as a result of the process. They have compared four proteins; lactate dehydrogenase (Schuessler and Herget 1980), BSA (Schuessler and Schilling 1984), haemoglobin (Puchala and Schuessler 1993) and myoglobin (Puchala and Schuessler 1995) under the same radiolysis conditions using SDS-PAGE and HPLC. They found that for all proteins, while being fundamentally different in their structures (number of tryptophan residues, disulphide bonds, SH-groups) there was a loss of native protein through cleavage of the polypeptide chain, yielding fragments which could be separated by electrophoresis. They found that the fragmentation of the protein was not a random process, but produced specific fragments, while the degree of aggregated material between proteins varied. They also point out that the alteration to protein size varies; in the absence of O₂ substantial cross-linking occurs resulting in aggregation of the protein, while in its presence oxidative modification only yields small fragments.

However, the occurrence of aggregation specifically in the presence of O₂ does not appear to occur for all proteins. An overview of the literature on protein oxidation modifications and the resulting products is diverse and it becomes apparent that it is highly dependent on the protein under attack. Studies have shown that H₂O₂ induced modification of proteins can result in the both the formation of higher order aggregated species as a consequence of cysteine modification and also in protein fragmentation. Lactoperoxidase undergoes H₂O₂ dependent formation of dimeric and trimeric products (Lardinois, Medzihradzky *et al.* 1999). H₂O₂ has also been shown to increase protein cysteine disulphides in the lens protein $\gamma\beta$ -crystallin (Hanson, Chen *et al.* 1999). Evidence also supports the breakdown of protein as a result of H₂O₂ modification of cysteine. HSP27, S-thiolated during oxidative stress by H₂O₂, disaggregates from its multimeric native state (Eaton, Fuller *et al.* 2002). Davies and Delsignore have attempted to correlate oxidative modification to increased proteolytic susceptibility (Davies 1987; Davies and Delsignore

1987; Davies, Delsignore *et al.* 1987; Davies, Lin *et al.* 1987). They show that alterations to the primary and secondary structure of BSA are severely damaged by the presence of the superoxide anion radical, the hydroxyl radical and oxygen. BSA exposed to solely HO• underwent progressive crosslinking to form dimers, trimers and tetramers, while exposure to HO• and O₂⁻ or O₂ induced fragmentation of the polypeptide chain at the α-carbon instead of the peptide bonds. They also linked the formation of bityrosine within a protein (intramolecular bonding) or between proteins (intermolecular bonding) as one of the processes leading to covalent protein cross-linking.

4.2 Aims of the study

The aim of this chapter focuses on understanding how oxidation induced modifications to proteins affect protein aggregation. Hydrogen peroxide (H₂O₂) was used as the oxidant. Again, BSA and lysozyme were used as model proteins. Lysozyme has a relatively low cysteine content (Cys- 8) and bovine serum albumin (Cys- 35) has a comparatively high cysteine content. Both proteins have a low methionine content; BSA contains three, with only two in lysozyme. Lysozyme also has 6 tryptophan residues compared to 2 in BSA. Hence it may be possible to establish if there is a correlation between the levels of oxidation induced aggregation observed for each protein relative to the quantity of susceptible amino acids present. Oxidation was induced over a range of mild to extreme oxidative stress conditions (dependent on the oxidant concentration). The influence of pH on oxidation induced protein aggregation was also explored. The oxidative damage to the proteins was evaluated by SE-HPLC, gel electrophoresis, turbidity and ThT assays, while intrinsic fluorescence was used to explore changes to the protein structure.

The exact mechanism by which protein induced aggregation is related to oxidative stress is still poorly understood. This chapter will address several important questions; how easily do proteins aggregate when oxidized using H₂O₂? Does their amino acid content play a role and if so how to what extent? To what degree does aggregation occur through the formation of homogenous clusters of monomeric protein, or is it random aggregation via heterogeneous aggregates due to the assembly of protein fragments?

4.3 Results

4.3.1 Oxidation of Lysozyme

4.3.1.1 Oxidative modification of lysozyme secondary structure

Changes to the secondary structure of lysozyme as a consequence of oxidation was monitored by intrinsic fluorescence by measuring the loss in total protein fluorescence intensity (excitation at 280 nm and 295 nm). Lysozyme solutions with a fixed concentration of 10 mg/ml (pH 5 – 8) were incubated with increasing concentrations of H_2O_2 and their intrinsic fluorescence was measured 1 hour after exposure to H_2O_2 (figure 4.4A and B). H_2O_2 induced clear conformation alterations of both the tyrosine and tryptophan fluorescence when the total fluorescence intensity (figure 4.4C and D) and λ_{max} (figure 4.4E and F) were monitored with increasing H_2O_2 concentrations.

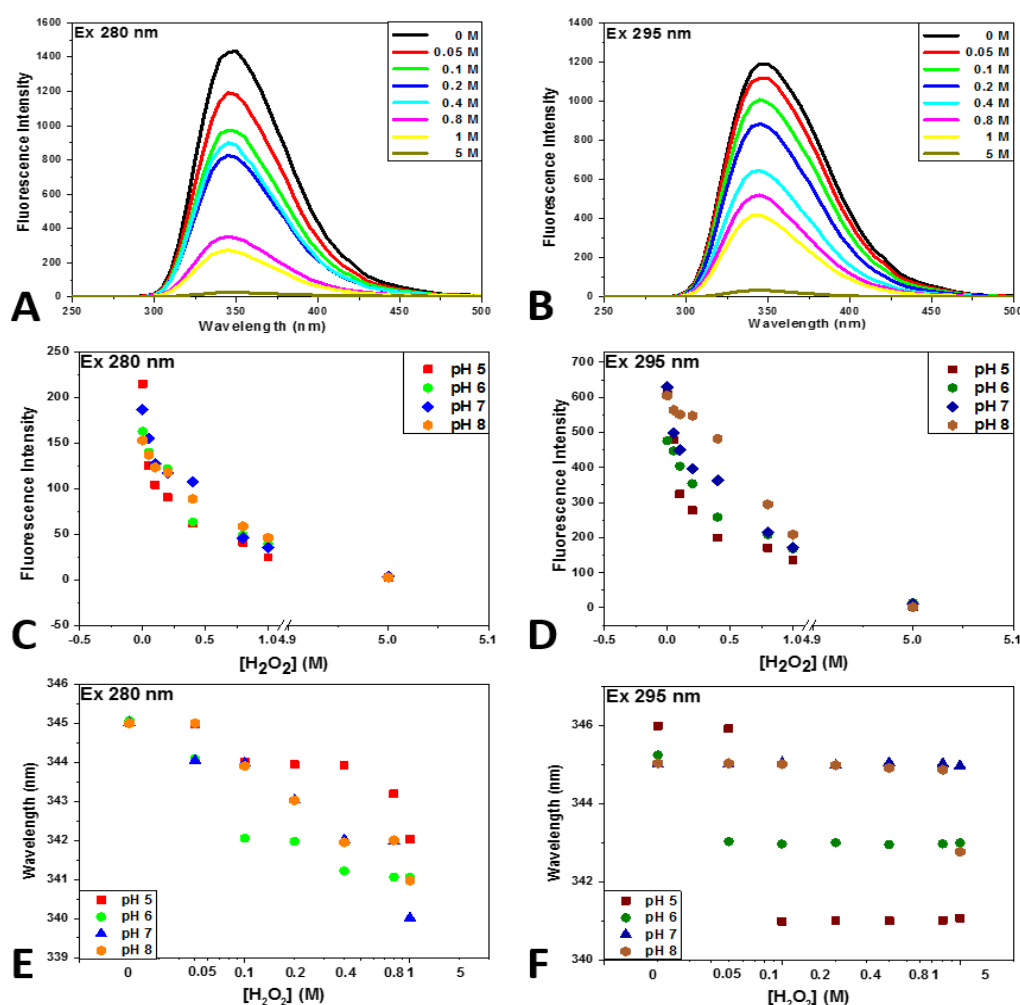


Figure 4.4: Intrinsic fluorescence spectra for lysozyme (pH 5) (A and B), the total fluorescence intensity (C and D) and the λ_{max} (E and F) of each individual fluorescence signal for lysozyme (pH 5 – 8) excited at 280 nm and 295 nm.

The loss in fluorescence intensity is an early phenomenon of radical-mediated protein oxidation and confirms alternations of the protein secondary structure (Chao, Ma *et al.* 1997; Campos, Lissi *et al.* 1999; Ma, Chao *et al.* 1999). The decrease of both the tryptophan and tyrosine intrinsic fluorescence indicates both residues are vulnerable to oxidative modification and are likely destroyed by the free radicals (Wu, Zhang *et al.* 2009). The fluorescence loss for tryptophan residues has been shown to be a direct consequence of oxidative degradation (Sun, Zhao *et al.* 2011). Table 4.1 shows the percentage recovery of fluorescence intensity across all solution conditions measured.

| <i>Ex 280 nm</i> % recovery | <i>pH 5</i> | <i>pH 6</i> | <i>pH 7</i> | <i>pH 8</i> | <i>Ex 295 nm</i> % recovery | <i>pH 5</i> | <i>pH 6</i> | <i>pH 7</i> | <i>pH 8</i> |
|---------------------------------------|-------------|-------------|-------------|-------------|---------------------------------------|-------------|-------------|-------------|-------------|
| <i>H₂O₂ (M)</i> | | | | | <i>H₂O₂ (M)</i> | | | | |
| 0.05 | 58.46 | 85.86 | 83.08 | 89.51 | 0.05 | 78.61 | 93.85 | 88.86 | 93.17 |
| 0.1 | 48.39 | 77.27 | 67.90 | 80.64 | 0.1 | 53.33 | 84.68 | 71.40 | 91.11 |
| 0.2 | 42.05 | 74.82 | 62.77 | 77.13 | 0.2 | 45.64 | 74.23 | 62.96 | 90.50 |
| 0.4 | 28.75 | 38.94 | 57.68 | 57.96 | 0.4 | 32.80 | 54.23 | 57.74 | 79.65 |
| 0.8 | 19.03 | 30.01 | 24.56 | 38.37 | 0.8 | 28.09 | 43.70 | 34.20 | 48.73 |
| 1 | 11.43 | 24.05 | 19.22 | 30.16 | 1 | 22.52 | 35.17 | 27.29 | 34.49 |
| 5 | 1.92 | 1.95 | 1.97 | 1.94 | 5 | 3.22 | 2.56 | 1.63 | 1.05 |

Table 4.1: Recovery of intrinsic fluorescence intensity for oxidised lysozyme (pH 5 – 8) calculated from the λ_{max} values.

These results suggest that extensive structural changes occurred for oxidised lysozyme. The fluorescence intensity was almost completely diminished for the high concentrations of H₂O₂ (1 – 5 M) for all solution conditions. The reduction in fluorescence intensity was comparable for pH 6 – 8, while lysozyme pH 5 showed a significantly larger reduction in fluorescence intensity, decreased by almost 40% in 0.05 M H₂O₂ compared to 15 – 20% for pH 6 – 8, indicating it's susceptibility to alteration was more acute. The blue shift in the wavelength position (figure 4.4) to shorter wavelengths for oxidised tryptophan is indicative of increased burial of the residues in a hydrophobic environment (Keerati-urai, Miriani *et al.* 2012). The result of this structural damage was monitored over time to assess if changes at the secondary structure due to oxidative damage would result in aggregation and/or fragmentation of lysozyme.

4.3.1.2 Macroscopic changes following oxidation

Exposure of lysozyme to H_2O_2 over long periods of time caused clear changes in the visual appearance of the protein solutions (figure 4.5).

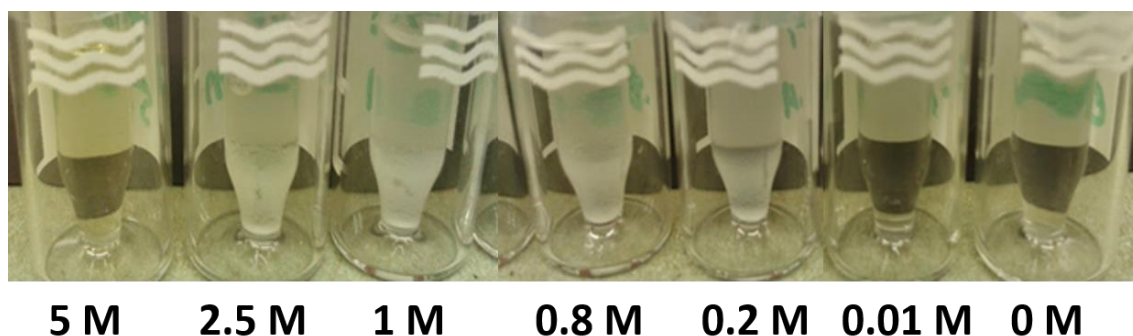


Figure 4.5: SE-HPLC glass vials containing oxidised lysozyme at pH 8, with different H_2O_2 concentrations after 7 days.

Lysozyme in 5 M H_2O_2 , resulted in a solution colour change from clear to an off yellow/orange that was visible within 24 hours and no evidence of visible aggregation (*i.e.* turbidity) was observed. From 0.2 – 2.5 M H_2O_2 , visible particles could be seen in the glass vials but without a colour change. Lysozyme exposed to lower concentrations of H_2O_2 (0.05 - 0.01 M) remained clear and no obvious particles were observed.

4.3.1.3 Lysozyme turbidity after oxidation

Exposure of lysozyme to H_2O_2 resulted in varying levels of turbidity, likely due to the formation of large aggregates. In order to discriminate between those samples that formed large aggregates and those that remained free from particles (at least sufficient in number to result in a measureable change in turbidity), % transmission was recorded in a multi-well plate over a 160 hour period (figure 4.6).

Highly concentrated solutions of H_2O_2 can react with plastic. The plates used are made from polystyrene. Chemical compatibility charts by both Thermo Scientific (Thermo Scientific, 2014) and JDR Enterprises Inc. (JDR Enterprises Incorporated, 2014) show polystyrene has excellent resistance, with “no damage after 30 days of constant exposure” to 30% hydrogen peroxide. For this reason it is believed the plates did not have any effect on the oxidation and subsequent aggregation of lysozyme.

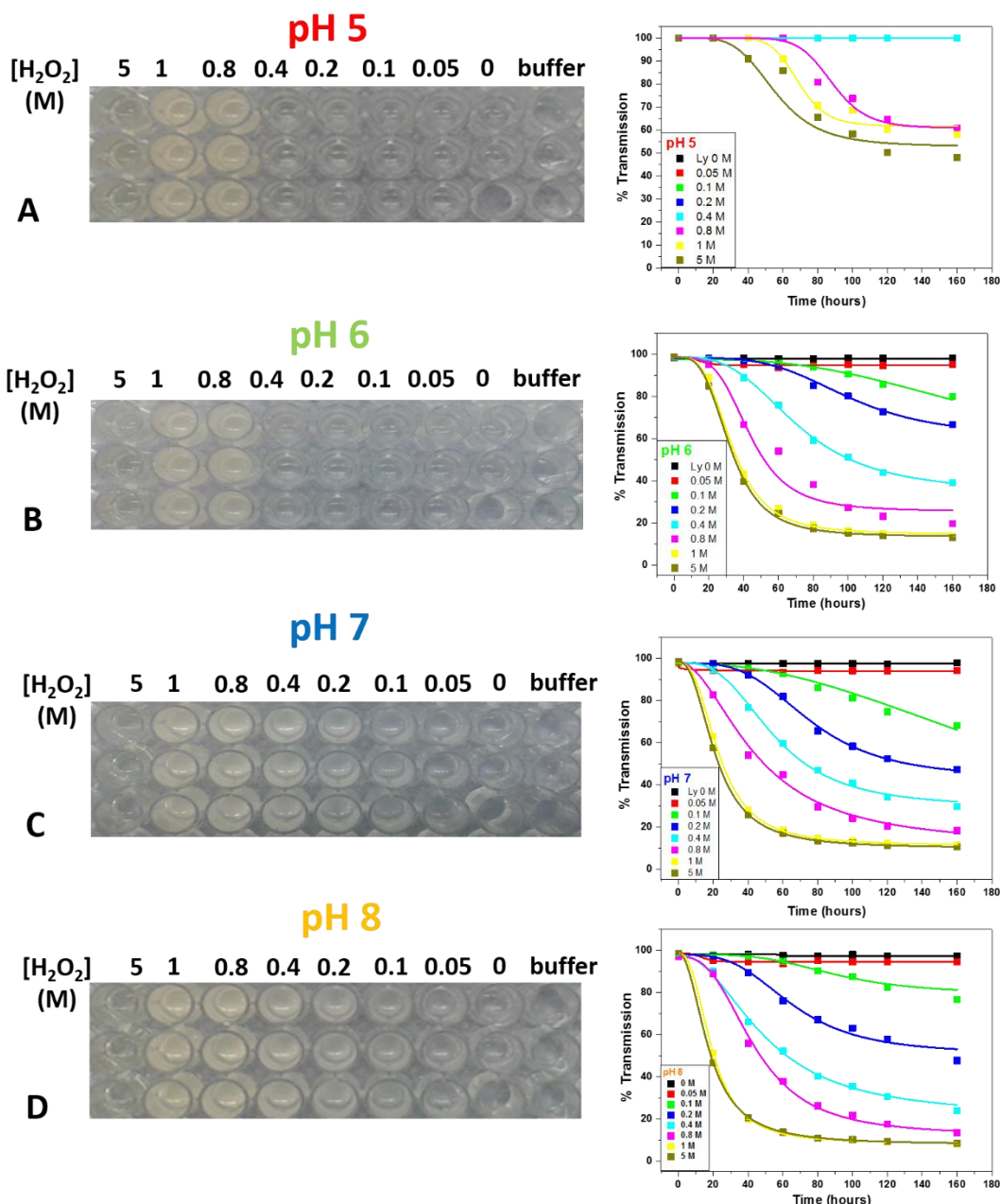


Figure 4.6: Images and turbidity analysis for lysozyme (10mg/ml) incubated with H_2O_2 over a 160 hour period. A) pH 5. B) pH 6. C) pH 7. D) pH 8. Multi-well plates contain an oxidised lysozyme solution after 160 hours.

The solution turbidity of some samples increased over time due to the formation of large aggregates, measured as a change in % transmission. At pH 6 – 8, aggregation was observed in all samples and increased with increasing H_2O_2 concentration. At high concentrations of H_2O_2 (0.8 – 5 M), aggregation was detected after ~20 hours, while the lower concentrations of H_2O_2 (0.1 – 0.4 M) showed a change in solution turbidity after longer incubation periods (60 – 80 hours). For lysozyme at pH 5, the lowest levels of turbidity were recorded. Only in the high concentrations of H_2O_2 (0.8 – 5 M) did lysozyme

indicate the formation of large aggregated material, and only after 50 – 60 hours of incubation. At the lower concentrations of H₂O₂ (0.1 – 0.4 M), no change in % transmission occurred. At 0.05 M H₂O₂, no change in % transmission occurred at any pH.

4.3.1.4 Characterisation of oxidised lysozyme precipitates

The lysozyme samples were also examined using light microscopy in order to visualise the particles formed. All microscopy was performed on samples of oxidised lysozyme on day 7 of incubation. Figure 4.7 shows micrographs of aggregated lysozyme at pH 5 and pH 8.

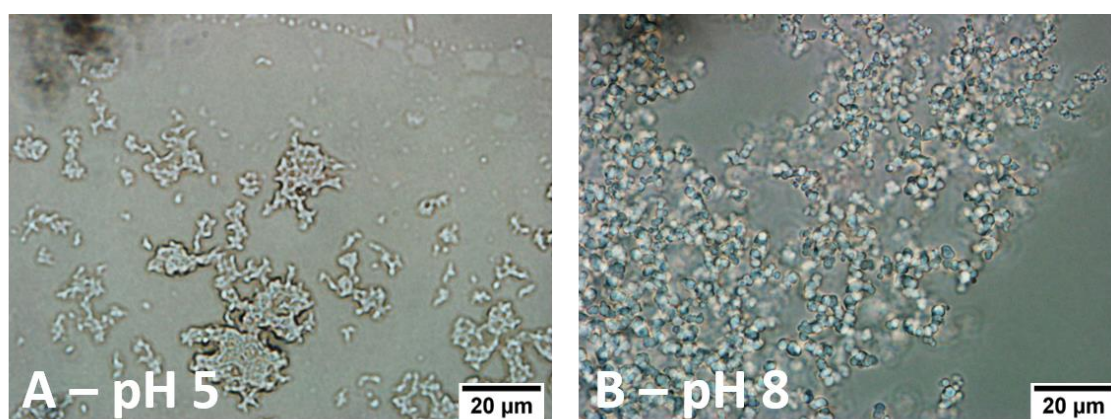


Figure 4.7: *Micrographs (light microscopy) of oxidised lysozyme aggregates at two different pH values.*

The aggregates formed at pH 8, appeared to be small clusters of spherical aggregates. The protein assemblies formed at pH 5, appeared smaller, less dense and more liable to break up. It is unclear from the optical microscopy images if there are differences between the aggregate types formed at each pH. It may simply be a concentration effect, since less aggregates form at pH 5.

Destabilisation of native lysozyme due to oxidative modification leads to increased amounts of protein with altered conformation as suggested with the intrinsic fluorescence analysis. Coupled with visual inspection, the turbidity assay and optical microscopy we have indicated that this leads to protein aggregation. However the light microscopy images cannot distinguish if these are amorphous aggregates or amyloid fibril material. The aggregate growth profiles determined by the % transmission measurements were similar to profiles expected for fibril formation (Sulatskaya, Turoverov *et al.* 2010). For this reason ThT fluorescence was used to assess if the lysozyme aggregates did have

amyloid fibril characteristics. Figure 4.8 shows the relationship between the level of ThT fluorescence intensity and solution pH after 180 hours (figure 4.8 A), along with the corresponding ThT fluorescence spectra for lysozyme at pH 8 after 180 hours (figure 4.8 B).

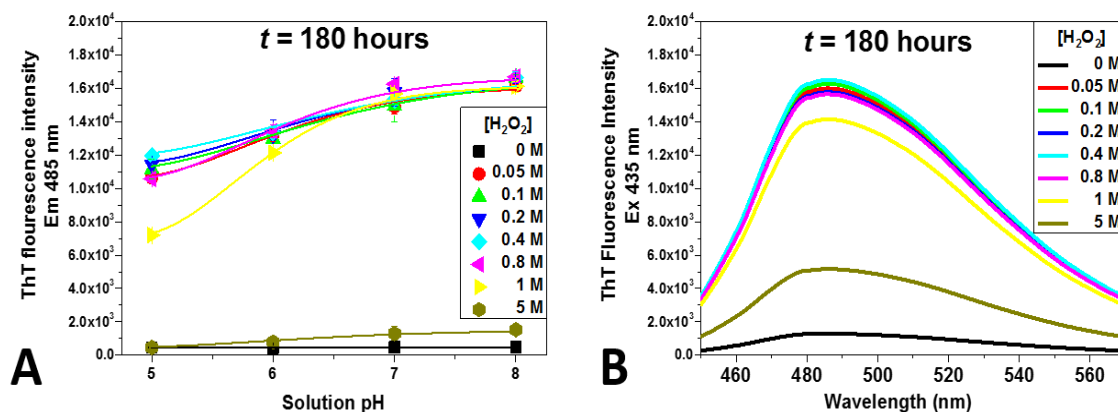


Figure 4.8: ThT assay of oxidised lysozyme solutions after 180 hours effect of solution pH on ThT fluorescence. A) Measured ThT fluorescence intensity(λ max 485 nm) for oxidised lysozyme pH 5 – 8. B) ThT fluorescence spectra for oxidised lysozyme at pH 8.0

The ThT fluorescence intensities recorded gradually increased with increasing pH. This is consistent with the % transmission data. At 5 M H_2O_2 and at all pH values, a small increase in ThT fluorescence was initially measured but decreased after 20 hours. The absence of ThT fluorescence correlates with the absence of precipitated material in the glass vials under these conditions. It appears that any aggregated particles capable of associating with ThT that do form, are quickly lost.

Plotting ThT fluorescence intensity as a function of time produced a sigmoidal plot (figure 4.9), with a short lag phase (not captured), a growth phase, and a termination phase, corresponding to initiation, elongation, and termination phases respectively, which is typical of a fibril forming kinetic mechanism (Biancalana and Koide 2010).

After oxidation, lysozyme solutions at pH 5 and pH 8, showed different characteristics in their ThT fluorescence curves and TEM images (figure 4.9 and 4.10). At pH 8, there is a rapid increase in fluorescence intensity after 20 minutes that reaches a maximum at 60 minutes for H_2O_2 concentrations between 0.1 – 1 M. At 0.05 M, a maximum is reached at 120 minutes. At the highest peroxide concentration (5 M), an initial increase in intensity is followed by a sharp decrease, reaching a minimum at 100 minutes. Transmission

electron microscopy was used to confirm the morphology of the oxidised lysozyme aggregates. TEM imaging confirmed the presence of fibrillar material for lysozyme at pH 8 (figure 4.9B).

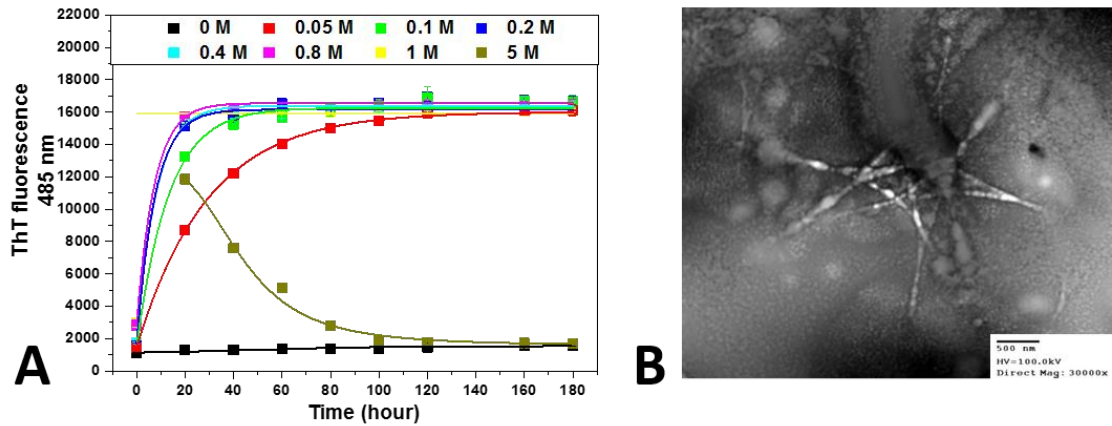


Figure 4.9: Oxidised lysozyme, pH 8. A) Increase in ThT fluorescence upon association with the aggregated material. B) TEM images indicating the presence of fibrillar material.

In contrast to this, the ThT assay of oxidised lysozyme at pH 5, indicates slightly lower levels of ThT fluorescence while TEM imaging did not indicate the formation of fibrils (figure 4.10B).

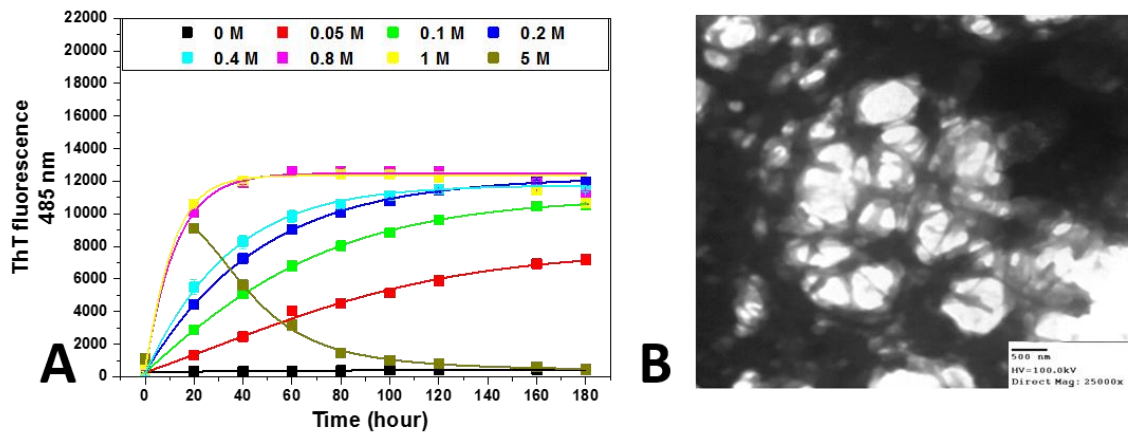


Figure 4.10: Oxidised lysozyme, pH 5. A) Increase in ThT fluorescence upon association with the aggregated material. B) TEM images indicating the presence of amorphous aggregated material.

As the TEM did not distinguish any distinct fibrils at this pH, it is more likely that amorphous aggregates are formed at pH 5 and that the ThT dye either binds to exposed hydrophobic regions of these aggregates or that they have a cross β -structure devoid of

fibril morphology. Both types of amorphous aggregates are known to bind ThT at significant levels (Chang, Liao *et al.* 2009; Sen, Fatima *et al.* 2009).

4.3.1.5 Analysis of small lysozyme aggregates

Oxidation of lysozyme clearly produces large aggregates (observed visually). Low concentrations of H₂O₂ (0.05 – 0.4 M) that showed no increase in solution turbidity or precipitated material but did cause ThT fluorescence. This would suggest the presence of a population of small soluble protein aggregates capable of associating with ThT, but not detected by the turbidity assay. We know from our analysis in Chapter 3 that the turbidity assay is not capable of detecting high proportions small aggregates. Therefore size exclusion HPLC was used to detect small lysozyme aggregates not detected by the turbidity assay and to monitor the loss in native monomeric lysozyme to possible fragmentation. Figure 4.11 shows chromatograms of lysozyme at pH 5 (10 mg/ml), incubated with increasing concentrations of H₂O₂.

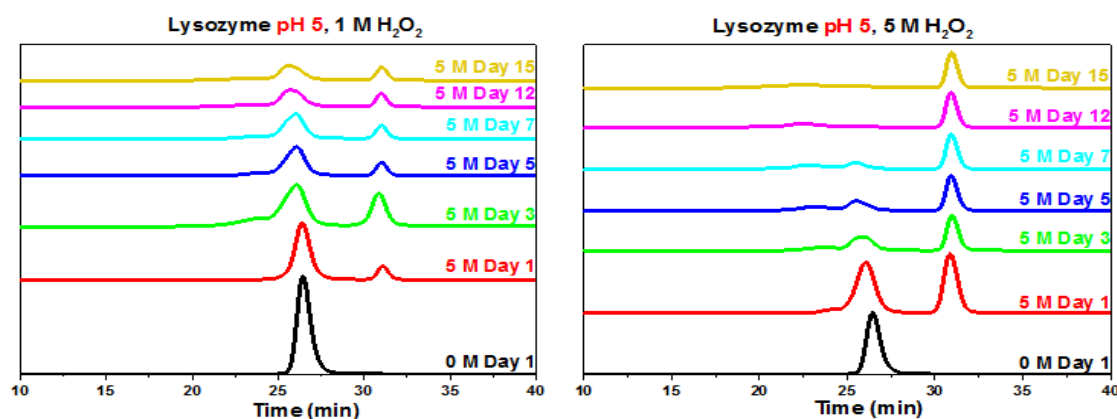


Figure 4.11: SE-HPLC chromatograms for oxidised lysozyme pH 5 (10 mg/ml) over 15 days.

The SE-HPLC chromatogram for lysozyme with H₂O₂ shows two main peaks. The peak at 26 minutes is the monomer of lysozyme and the peak at 31 minutes is H₂O₂. As the stability of the monomer becomes compromised as a result of oxidation over time, the area under the monomer peak of lysozyme decreases coupled with an increase in peaks at shorter retention times indicating higher molecular weight aggregates. The degree to which the peak shapes changed was related to the concentration of oxidant and time.

The overlaid SE-HPLC chromatograms for oxidised lysozyme at pH 5, on day 1 and day 15 are shown in Figure 4.12. The loss in lysozyme concentration was calculated on the

basis of a previously calibrated concentration series for lysozyme. Oxidation of the monomer occurs quickly. On day one there is an almost 50% loss in monomer concentration and a small proportion of higher molecular weight aggregates are observed at high oxidant concentrations (2.5 – 5 M). By day 15 all samples show a decrease in monomer concentration and an increase in the concentration of higher molecular weight aggregates. The decrease in monomer concentration over 15 days is shown in Figure 4.12 C. By day 15 the high concentrations of H₂O₂ (2.5 – 5 M) indicate that nearly 90% of monomeric protein has been lost to oxidation. None of the oxidised samples showed evidence of fragmentation (*i.e.* peaks associated with lower molecular weight protein fragments), a common product of protein oxidation processes (Guedes, Vitorino *et al.* 2009).

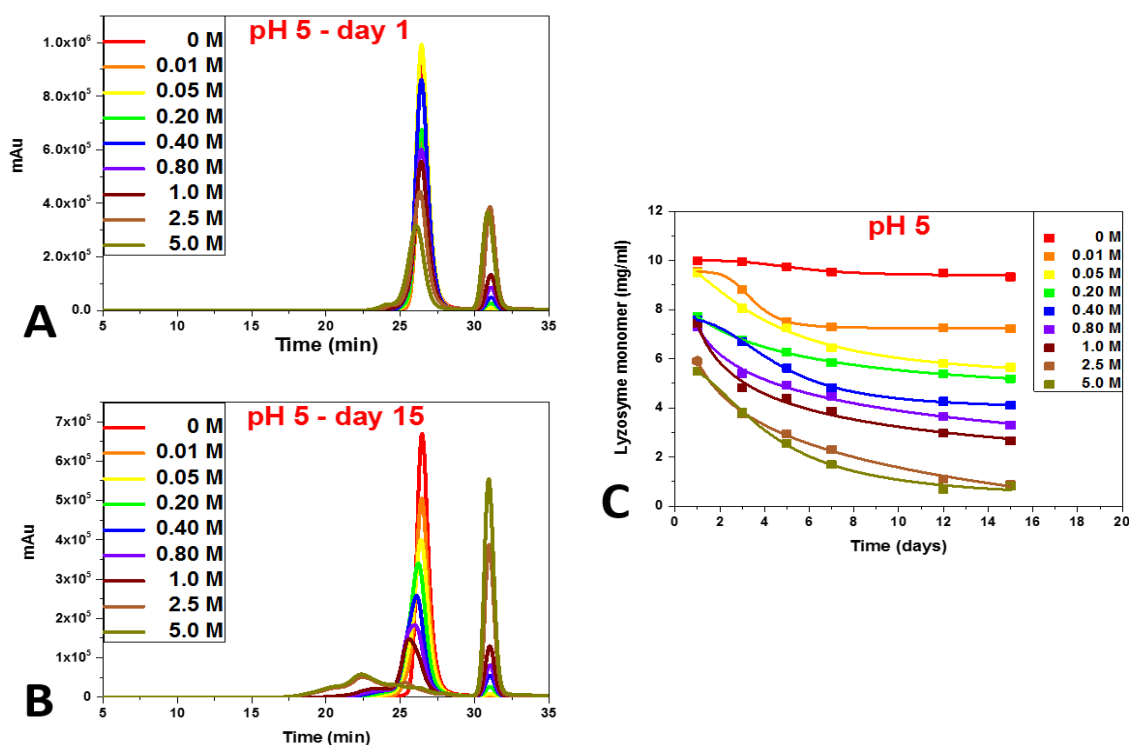


Figure 4.12: SE-HPLC chromatograms of oxidised lysozyme at pH 5 on A) day 1 and B) day 15 of incubation. C) Loss in lysozyme monomer concentration (mg/ml) as a function of time and H₂O₂ concentration.

Lysozyme at pH 8, treated in the same manner showed similar behaviour (figure 4.13). There was an overall reduction in the concentration of lysozyme monomer initially and a subsequent increase in higher molecular weight aggregates, which increased over time, and with increasing H₂O₂ concentrations (figure 4.13C).

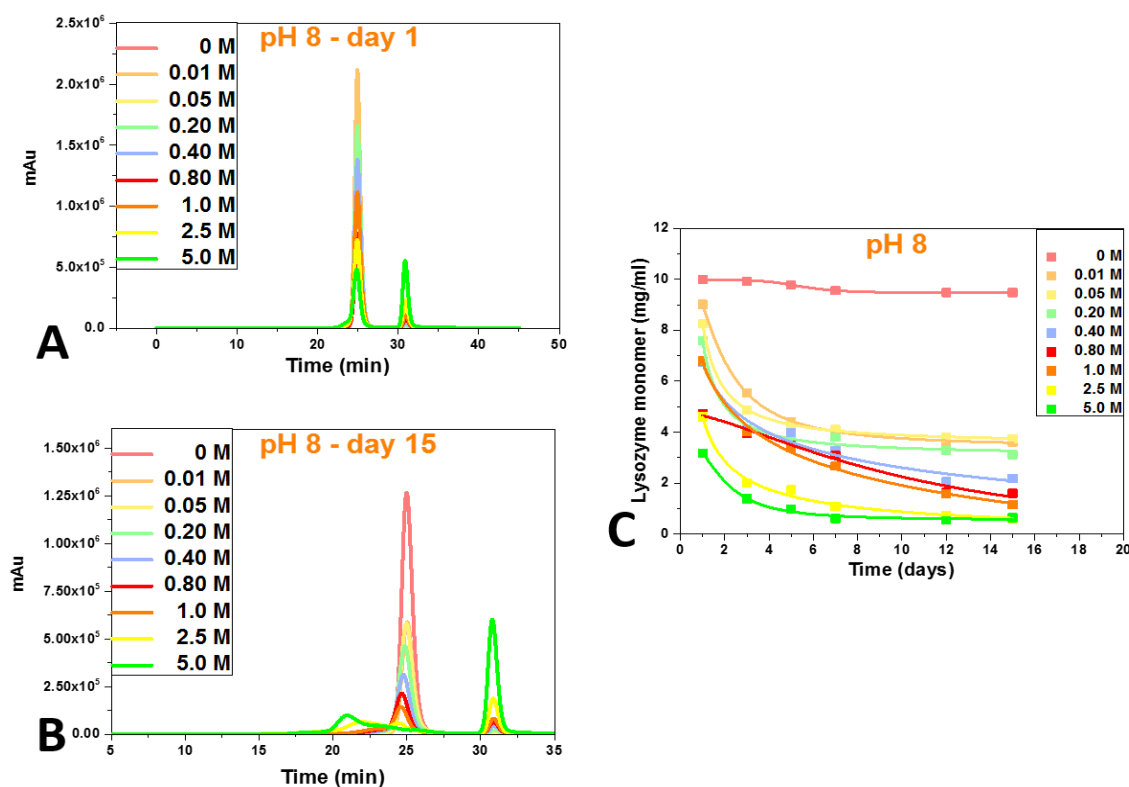


Figure 4.13: SE-HPLC chromatograms of oxidised lysozyme at pH 8 on A) day 1 and B) day 15 of incubation. C) Loss in lysozyme monomer concentration (mg/ml) as a function of time and H_2O_2 concentration.

The loss in native monomeric lysozyme (pH 5 and pH 8) into oxidation products was calculated after 7 days incubation (table 4.2 and 4.3). The initial concentration of lysozyme was 10 mg/ml. From the SE-HPLC chromatograms, the quantity of monomer (column 2) and higher molecular weight aggregates (≤ 600 kDa) resolved by HPLC (column 3) was calculated. Any precipitated aggregate material was phase separated and the concentration of protein remaining in the supernatant was calculated. This allowed us to establish the loss of native protein to large protein particles (column 4). The concentration of protein unaccounted for was thought to be either made up of fragments (so far undetected), or protein aggregates larger than the resolution of the column (> 600 kDa) but smaller than protein particles removed with the bulk precipitate separated by centrifugation (column 5).

| <i>H₂O₂ (M)</i> | <i>SE-HPLC monomer</i> | <i>SE-HPLC higher order (≤600 kDa)</i> | <i>Phase separation large aggregates</i> | <i>Not detected</i> |
|--|-------------------------------|---|---|----------------------------|
| <i>0.01</i> | 7.29 | 0.93 | 0.23 | <i>1.55</i> |
| <i>0.05</i> | 6.45 | 0.87 | 1.28 | <i>1.40</i> |
| <i>0.2</i> | 5.85 | 0.14 | 2.16 | <i>1.85</i> |
| <i>0.4</i> | 4.82 | 0.31 | 265 | <i>2.22</i> |
| <i>0.8</i> | 4.47 | 0.39 | 2.94 | <i>2.17</i> |
| <i>1.0</i> | 3.87 | 0.44 | 3.48 | <i>2.21</i> |
| <i>2.5</i> | 2.29 | 1.59 | 1.56 | <i>1.56</i> |
| <i>5.0</i> | 1.68 | 1.92 | 1.25 | <i>5.15</i> |

Table 4.2: Lysozyme at pH 5 (10 mg/ml), concentration of oxidation products (mg/ml) after 7 days exposure to H₂O₂.

| <i>H₂O₂ (M)</i> | <i>SE-HPLC monomer</i> | <i>SE-HPLC higher order (≤600 kDa)</i> | <i>Phase separation large aggregates</i> | <i>Not detected</i> |
|--|-------------------------------|---|---|----------------------------|
| <i>0.01</i> | 4.07 | 1.21 | 1.04 | 3.67 |
| <i>0.05</i> | 4.11 | 1.34 | 1.97 | 2.57 |
| <i>0.2</i> | 3.79 | 0.70 | 2.12 | 3.38 |
| <i>0.4</i> | 3.28 | 0.61 | 2.45 | 3.64 |
| <i>0.8</i> | 3.08 | 0.42 | 3.86 | 2.62 |
| <i>1.0</i> | 2.66 | 0.37 | 4.15 | 2.81 |
| <i>2.5</i> | 1.08 | 0.66 | 4.72 | 3.53 |
| <i>5.0</i> | 0.60 | 1.42 | 1.14 | 6.83 |

Table 4.3: Lysozyme pH 8, concentration of oxidation products (mg/ml) after 7 days exposure to H₂O₂.

SE-HPLC of the supernatant was only able to resolve low levels of molecular weight species < 600 kDa. The majority of lysozyme had either formed aggregates > 600 kDa or this was an indication that fragmentation was occurring. It was possible that the extent to which lysozyme was being fragmented could not be resolved by SE-HPLC because the monomer peak (25 min) eluted too close in proximity to the H₂O₂ peak (31 min) so overlapping at the point where the lysozyme peak ended and the H₂O₂ peak began occurred at the elution point where fragments may have been detected (figure 4.14).

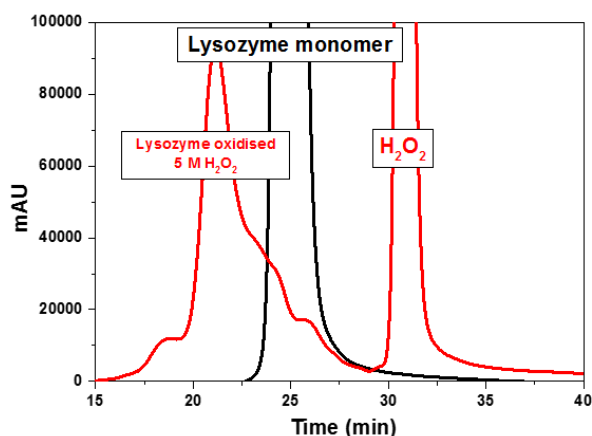


Figure 4.14: Lysozyme monomer peak along with oxidised lysozyme indicating the overlap between protein peak and H_2O_2 peak.

4.3.1.6 Fragmentation of oxidised lysozyme

Exposure of lysozyme to H_2O_2 as analysed by gel electrophoresis indicates fragmentation of the protein was occurring (figure 4.15). Non-denaturing PAGE conditions revealed both fragmentation and aggregation of lysozyme and that the relative concentration of either fragments or aggregates depends on the H_2O_2 concentration. No distinct bands could be resolved except for lysozyme monomer and dimer at 14.7 kDa and 29.4 kDa, respectively.

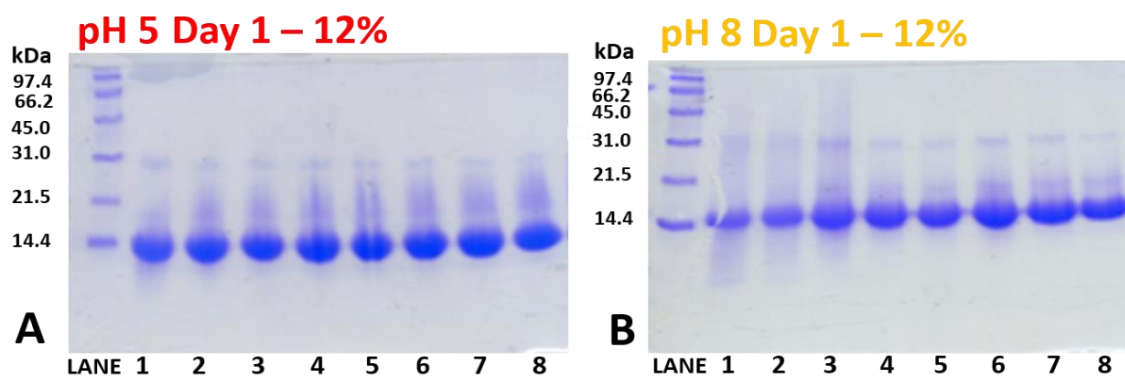


Figure 4.15: 12% SDS PAGE gel of lysozyme A) at pH 5 and B) at pH 8, with increasing concentrations of H_2O_2 day 1. Lane 1- 5 M, lane 2- 2.5 M, lane 3- 1 M, lane 4- 0.8 M, lane 5- 0.4 M, lane 6- 0.2 M, lane 7- 0.1 M, lane 8- 0.05 M.

Lysozyme at pH 5, lanes 1 – 2 only showed a small portion of fragments at the high concentration of H_2O_2 (2.5 – 5 M). Lysozyme at pH 8, at the same time point indicated fragmentation in more of the samples, lanes 1 – 6 (0.2 – 5 M H_2O_2). Higher molecular weight species were also observed. At pH 5 these species were no bigger than 30.1 kDa, while they were larger at pH 8 (~ 66 kDa) for higher concentrations of H_2O_2 (lanes 1 – 3). It therefore seems reasonable to suggest from the size range of higher molecular

weight aggregates observed, that the fragmented protein segments were aggregating together, or to remaining monomers.

Lysozyme at pH 5 and at pH 8, incubated with H₂O₂ showed increased levels of both fragmentation and aggregation with time (figure 4.16). For both solution conditions the amount of fragments and higher molecular weight species resolved was greater after 3 days of incubation. The monomer band can still be observed at 14.7 kDa.

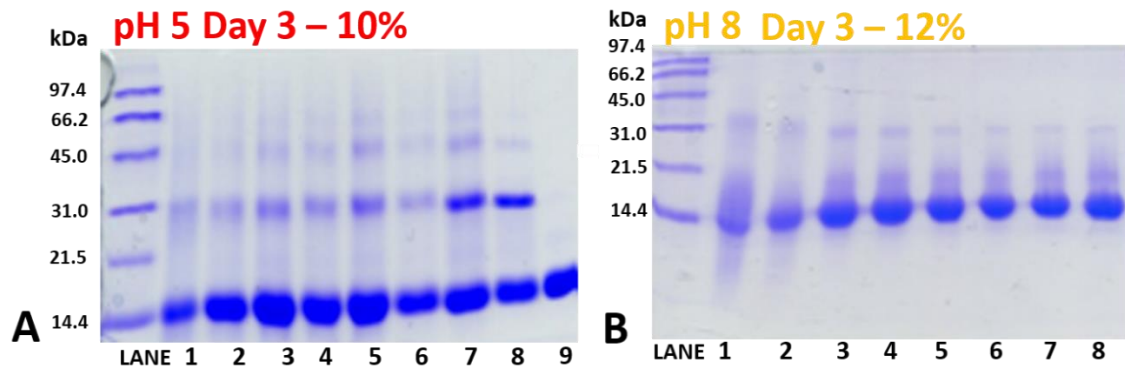


Figure 4.16: A) 10% SDS PAGE gel of lysozyme at pH 5 and B) 12% SDS PAGE gel of lysozyme at pH 8, with increasing concentrations of H₂O₂ day 3. Lane 1- 5 M, lane 2- 2.5 M, lane 3- 1 M, lane 4- 0.8 M, lane 5- 0.4 M, lane 6- 0.2 M, lane 7- 0.1 M, lane 8- 0.05 M, lane 9- 0 M.

By day 5 for both solution conditions (figure 4.17), almost no fragments were detected and there was a reduction in the amount of higher molecular weight species resolved. By this time point the many visible aggregates are observed and are too large to penetrate the gel network and instead accumulate at the top of the gel (not shown).

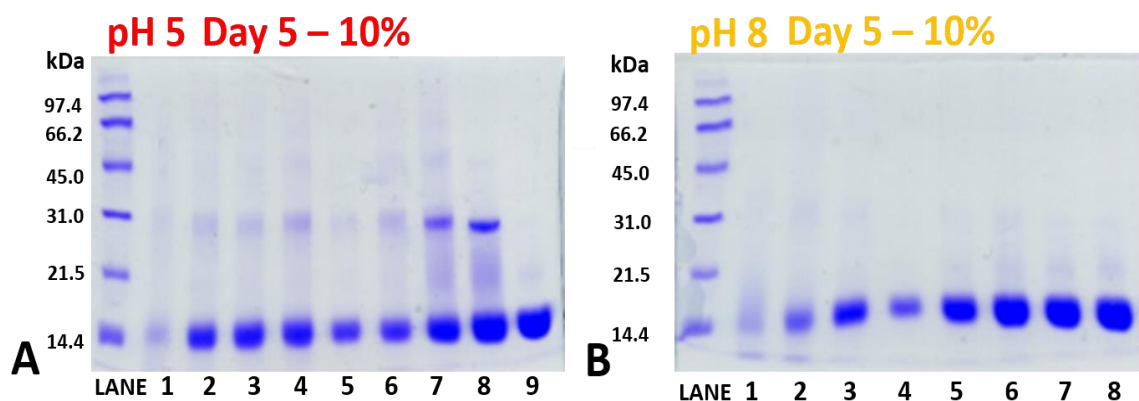


Figure 4.17: A) 10% SDS PAGE gel of lysozyme at pH 5 and B) 12% SDS PAGE gel of lysozyme with increasing concentrations of H₂O₂ day 5. Lane 1- 5 M, lane 2- 2.5 M, lane 3- 1 M, lane 4- 0.8 M, lane 5- 0.4 M, lane 6- 0.2 M, lane 7- 0.1 M, lane 8- 0.05 M, lane 9- 0 M.

4.3.2 Oxidation of Bovine serum albumin

Oxidation of BSA was studied following exposure to increasing concentrations of H_2O_2 in the same way as was described above for lysozyme. Initial analysis of oxidised BSA after one week of incubation did not indicate any change in solution appearance, in % transmission or in ThT fluorescence *i.e.* no aggregation resulting in large protein particles was observed. Therefore, analysis focused on a single solution condition; 50 mM sodium phosphate buffer, pH 7.

4.3.2.1 Oxidative modification of BSA secondary structure

Intrinsic fluorescence was used to examine the effect of oxidation on the secondary structure of BSA by measuring the loss in total protein fluorescence and λ_{max} . The change in intrinsic fluorescence for BSA with increasing H_2O_2 concentration is shown in Figure 4.18A and B. Exposure to H_2O_2 induced a clear decrease in both the tyrosine and tryptophan fluorescence intensity.

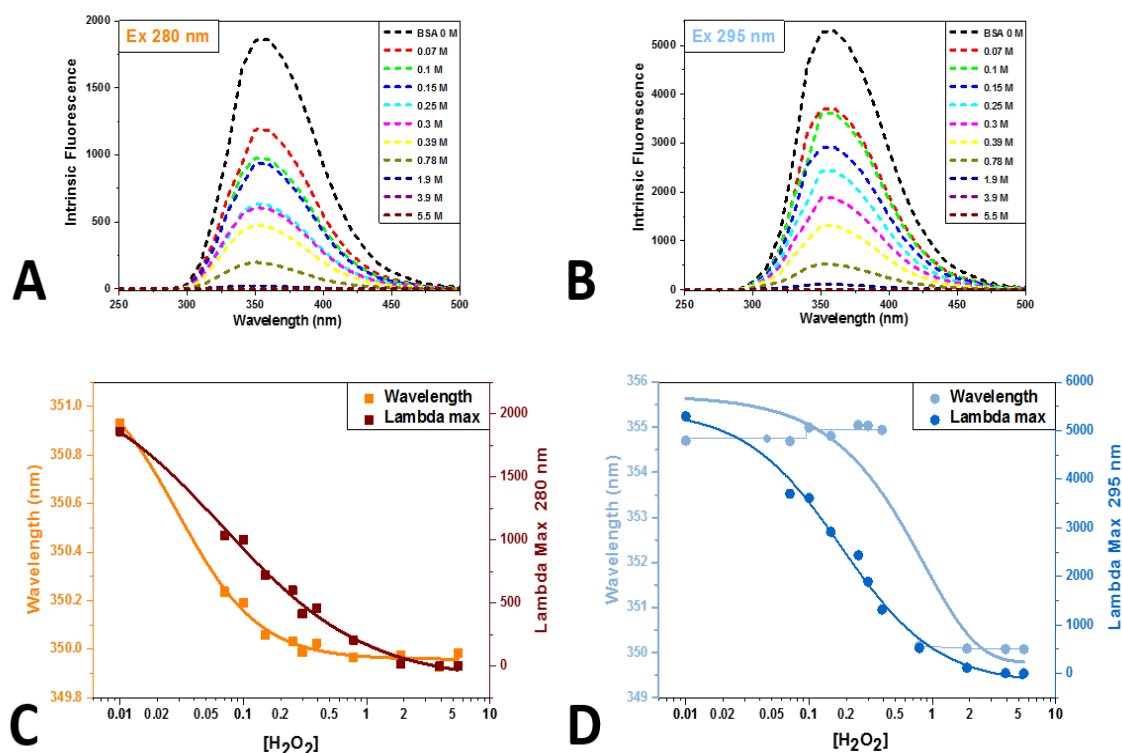


Figure 4.18: Intrinsic fluorescence spectra for oxidised BSA (pH 7) along with plots of the λ_{max} and wavelength shift of each individual fluorescence signal excited at 280 nm (A and C) and 295 nm (B and D).

Exposure to H₂O₂ caused a blue shift in the emission wavelength. The fluorescence emission spectrum of BSA in the absence of H₂O₂ (excitation at 280 nm) has a λ_{max} at 351 nm which becomes slightly blue shifted to 349 nm with increasing oxidant concentration. The λ_{max} after excitation at 295 nm had a more significant blue shift from 355 nm to 350 nm, indicating a substantial alteration to the tryptophan residues between 0.5 and 0.8 M H₂O₂ affecting their hydrophobicity. The oxidation of the BSA resulted in a progressive decrease in the intrinsic fluorescence signal after excitation at both 280 nm and 295 nm (table 4.4).

| <i>[H₂O₂] M</i> | <i>Ex 280 nm % loss</i> | <i>Ex 295 nm % loss</i> |
|---------------------------------------|-----------------------------|-----------------------------|
| 0.07 | 55.52 | 69.87 |
| 0.1 | 53.97 | 68.24 |
| 0.15 | 38.85 | 55.07 |
| 0.25 | 32.25 | 45.99 |
| 0.3 | 22.38 | 35.63 |
| 0.39 | 24.82 | 24.85 |
| 0.78 | 11.27 | 9.9 |
| 1.9 | 0.89 | 2.23 |
| 3.9 | 0.09 | 0.18 |
| 5.5 | 0.03 | 0.05 |

Table 4.4: Loss in intrinsic fluorescence intensity for oxidised BSA calculated from the λ_{max} values.

At the higher concentrations of H₂O₂ (1.9 – 5.5 M), the signal intensity was almost completely diminished indicating disruption of both the tyrosine and tryptophan residues. These observations showed that as a result of oxidative stress, there was significant conformational changes and alterations in the solvent environment in the vicinity of the tyrosine and tryptophan residues (Chen and Barkley 1998; Lichtman and Conchello 2005; Sun, Zhao *et al.* 2011).

4.3.2.2 Analysis of BSA aggregation and fragmentation

Monomeric BSA generally exists in equilibrium with dimeric and trimeric forms of the protein (figure 4.19). SDS PAGE resolves BSA as three separate bands; monomer (66.5 kDa), dimer (133 kDa) and a small amount of trimer (199.5 kDa). SE-HPLC only resolves both the monomer and dimer form (ratio is 87:13 monomer/dimer). Since SE-HPLC does not resolve the trimer, its concentration is less than 0.5% of the total protein content.

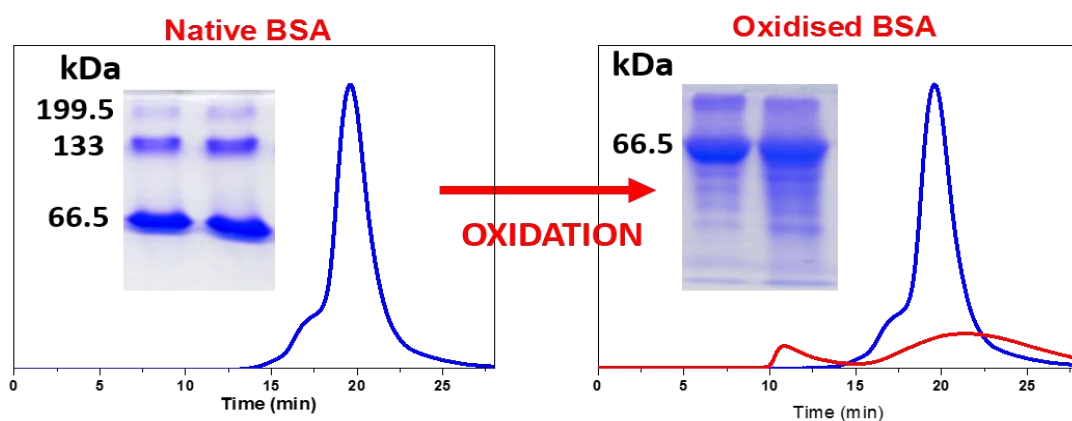


Figure 4.19: SE-HPLC and gel electrophoresis of BSA before and after addition of H_2O_2 .

Oxidation of BSA resulted in the addition of several new protein bands and a spreading of the chromatogram peak indicating both aggregation and fragmentation of BSA as a consequence of oxidative stress. To monitor changes to the protein, three categories were considered; monomer, aggregates (\geq dimer) and fragments ($<$ monomer). The concentration in each category was monitored and expressed as a percentage of the total protein present. This allowed us to monitor the evolution of the oxidation products of BSA over time.

A reduction in the area under the monomer peak was measured as the stability of BSA became compromised (figure 4.20). The decrease in BSA concentration was calculated on the basis of a previously calibrated concentration series for BSA. By day 2, over 50% of protein monomer was lost at the higher concentrations of H_2O_2 (0.78 – 5.5 M).

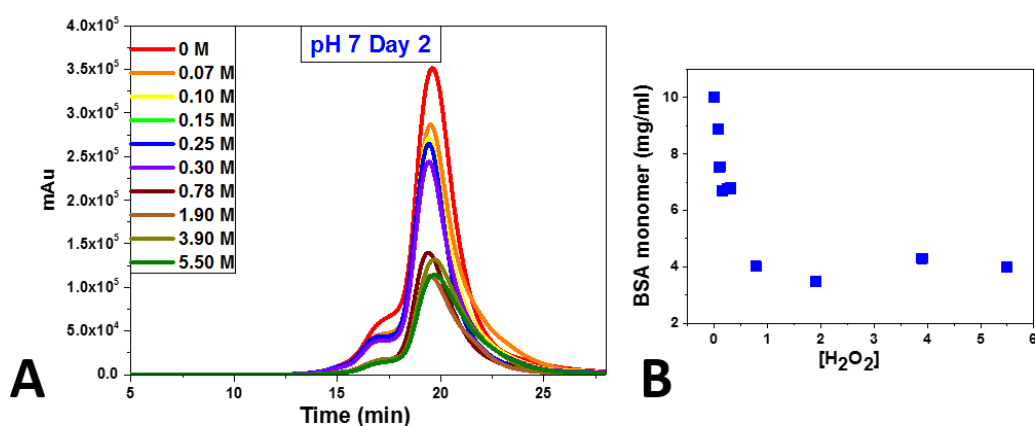


Figure 4.20: A) SE-HPLC chromatograms of oxidised BSA at pH 7 after 24 hours. B) Loss in BSA monomer concentration (mg/ml) as a function of H_2O_2 concentration.

However the loss in monomer did not correlate with the detected concentration of higher molecular weight aggregates or fragments (within the column resolution), suggesting significant aggregation and/or fragmentation which remained undetected was occurring since only monomer and dimer were observed (figure 4.21). Monitoring both of these species with increasing H_2O_2 concentration showed that at the low concentrations of H_2O_2 (0.07 – 0.3 M) the ratio of monomer to dimer remains consistent with native BSA. At higher concentrations (0.78 – 5.5 M), the increase in the relative amount of monomer (and corresponding decrease in dimer) suggested that the disulphide bridge between the BSA dimer was cleaved.

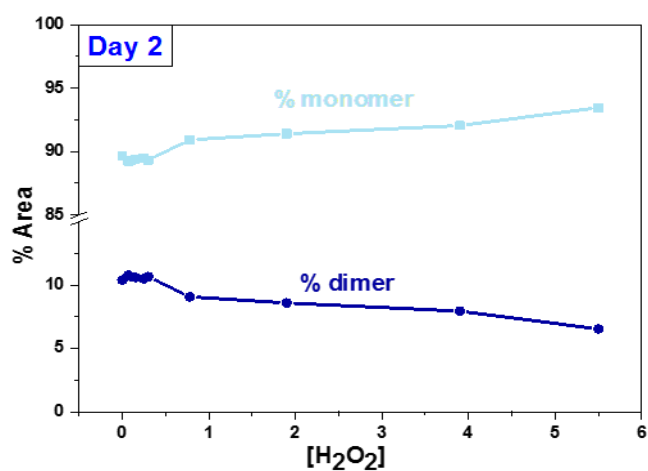


Figure 4.21: SE-HPLC data of BSA at pH 7, indicating the % area under the monomer peak, and dimer peaks of the chromatograms in Figure 4.20.

Non-denaturing PAGE of the same samples on day 2 showed a very small degree of fragmentation at the higher concentrations of H_2O_2 (lanes 3 – 6) (figure 4.22). The concentration of the dimer band had reduced, an indication that it has been cleaved as the HPLC suggests.

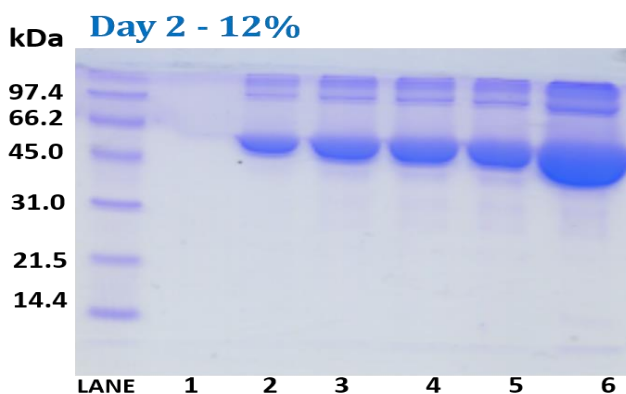


Figure 4.22: 12% SDS PAGE 12% gel of BSA with increasing concentrations of H_2O_2 day 2. Lane 1- empty, lane 2- 0.3 M, lane 3- 0.78 M, lane 4- 1.9 M, lane 5- 5.5 M, lane 6- 5.5 M (overloaded).

The fragmentation of BSA exposed to H_2O_2 increased with time. By day three of oxidation, fragmentation of the BSA structure had further increased relative to day 2 for the high concentrations of H_2O_2 (lanes 4 – 6), with the low concentrations of H_2O_2 also showing evidence of fragmentation (lanes 1 – 3) (figure 4.23).

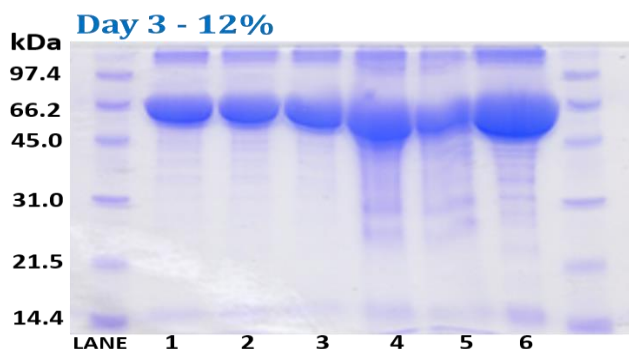


Figure 4.23: 12% SDS PAGE gel of BSA with increasing concentrations of H_2O_2 , day 3. Lane 1- 0.07, lane 2- 0.1 M, lane 3- 0.25 M, lane 4- 0.78 M, lane 5- 1.9 M, lane 6- 5.5 M. Samples were all overloaded so the lower concentrations of fragments present could be observed.

Progressive oxidation of BSA resulted in an almost complete loss of monomer. SE-HPLC analysis after 8 days of oxidation shows a decrease in monomer concentration at all H_2O_2 concentrations concurrent with the formation of fragments and higher molecular weight aggregates (figure 4.24).

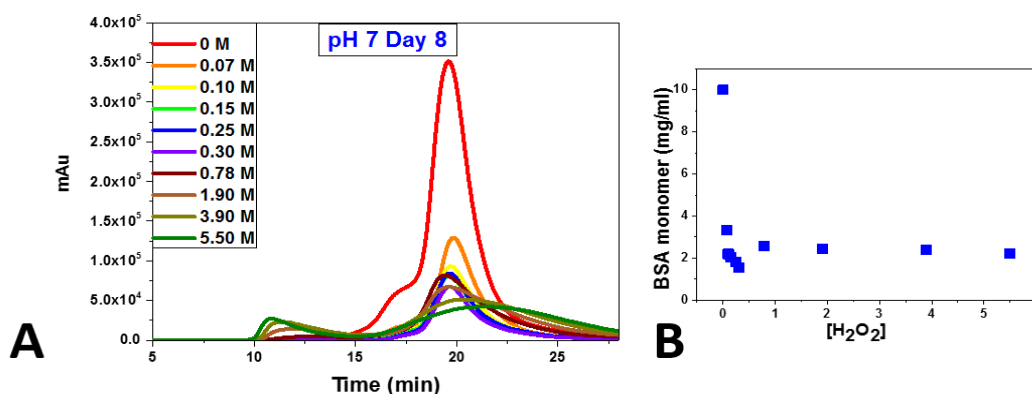


Figure 4.24: A) SE-HPLC chromatograms of oxidised BSA at pH 7 after 8 days. B) Loss in BSA monomer concentration (mg/ml) as a function of H_2O_2 concentration.

For the higher concentrations of H_2O_2 (0.78 – 5.5 M) after 8 days of oxidation, the monomer peak is replaced with a very broad peak ranging in molecular weight from 10 – 600 kDa (resolution of the SE-HPLC column), which would include fragments, monomers and aggregates. The low concentrations of H_2O_2 (0.07– 0.3 M) showed a reduction in the area under the monomer peak, along with a very slight increase in the

amount of higher molecular weight aggregates, believed to be an increase in dimer rather than new higher order species as the retention time of the peak was close to dimer retention time. However fragments were not detected at these low concentrations (figure 4.25).

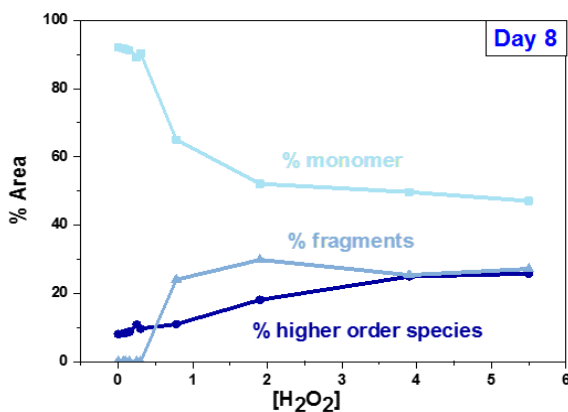


Figure 4.25: SE-HPLC data for BSA at pH 7, day 8 showing the % area under the monomer peak, fragment peaks and higher molecular weight aggregate peaks of the chromatograms in Figure 4.24.

PAGE of corresponding samples on day 8 does show a very limited amount of fragmentation of BSA at these concentrations lanes 5 - 9 (figure 4.26).

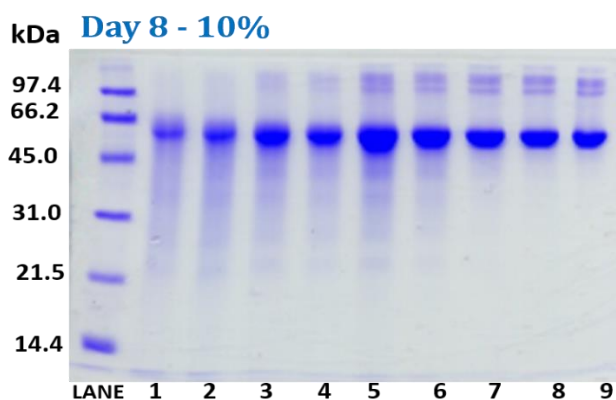


Figure 4.26: 10% SDS PAGE gel of BSA with increasing concentrations of H₂O₂, day 8. Lane 1- 5.5 M, lane 2- 3.9 M, lane 3- 1.9 M, lane 4- 0.78 M, lane 5- 0.3 M, lane 6- 0.25 M, lane 7- 0.15 M, lane 8- 0.1 M, lane 9- 0.07 M.

The size of the fragments between each sample again appeared to be quite uniform, since bands of identical molecular weight across all wells were observed at all H₂O₂ concentrations. This observation is indicative of site specific oxidation (Hunt, Bottoms *et al.* 1993). A reduction in the monomer band, along with a reduction in the dimer and trimer band could also be observed in lanes 1 – 4 compared to lanes 5 – 9. The H₂O₂ appeared to be breaking up the dimer and trimer forms of BSA at the high concentration of H₂O₂.

4.3.2.3 Quantification of oxidation products

The quantity of BSA that formed different oxidation products after 8 days of exposure to H₂O₂ was calculated (table 4.5). From SE-HPLC the quantity of remaining monomer (after removal of large aggregates by centrifugation) (column 2), higher order species (\leq 600 kDa) and fragments (column 3) detected was calculated. The concentration of protein not detected by this technique was assumed to be either fragments, which were undetected or protein aggregates, larger than the resolution of the column ($>$ 600 kDa) (column 4).

| <i>H₂O₂ (M)</i> | <i>SE-HPLC monomer</i> | <i>SE-HPLC higher order (\leq600 kDa)</i> | <i>SE-HPLC fragments</i> | <i>Not detected</i> |
|---------------------------------------|------------------------|--|--------------------------|---------------------|
| <i>0.07</i> | 2.93 | 0.26 | - | 6.81 |
| <i>0.1</i> | 1.93 | 0.18 | - | 7.89 |
| <i>0.15</i> | 1.79 | 0.17 | - | 8.03 |
| <i>0.25</i> | 1.70 | 0.23 | - | 8.07 |
| <i>0.30</i> | 1.76 | 0.15 | - | 8.09 |
| <i>0.78</i> | 2.00 | 0.38 | 1.01 | 6.61 |
| <i>1.9</i> | 1.94 | 0.75 | 2.24 | 5.07 |
| <i>3.9</i> | 1.75 | 0.98 | 2.08 | 5.19 |
| <i>5.5</i> | 1.43 | 1.07 | 2.54 | 4.96 |

Table 4.5: BSA at pH 7, concentration of oxidation products (mg/ml) after 8 days exposure to H₂O₂.

This showed that the majority of BSA had formed either fragmented products or protein aggregates that were larger than 600 kDa but not large enough to alter the appearance of the BSA solutions *i.e.* give rise to solution turbidity and that the total concentration of protein left in solution (and hence detected by HPLC and PAGE) was quite low. However we known from the gel electrophoresis that only a small proportion of fragments were resolved for the low concentrations of H₂O₂. It is therefore assumed that aggregation was the far more likely oxidation product at the low concentrations of oxidant, while at the higher concentrations of oxidant greater levels of both fragmentation and aggregation were occurring.

4.4 Discussion

Although BSA and lysozyme are very different in their structures, there were some similarities in the results of our oxidation study. In both cases under our oxidation conditions aggregation of the protein occurs.

The levels of aggregation for lysozyme after exposure to H_2O_2 was in stark contrast to BSA. Within 24 hours of lysozyme oxidation, large protein aggregates were detected, of which some showed evidence of fibril formation. The ability of H_2O_2 to aggregate lysozyme is broken down into three oxidation conditions; mild (0.01 – 0.1 M), severe (0.2 – 1 M) and extreme (5 M) oxidation based on the amount of monomer structure retained after oxidation. The formation of protein aggregates was observed across all three oxidation conditions by the changes in solution appearance (turbidity), ThT assay, SE-HPLC, and gel electrophoresis.

The most striking feature about the oxidation of BSA is that a large proportion of the protein aggregates (up to 80%), but that these aggregates are not large enough to produce sample turbidity. Hence, these aggregates are not detected by either of the high throughput methods, % transmission or by ThT assay. This again reiterates the point made in Chapter 3 about the sensitivity of certain high throughput screening methods to detect levels of aggregation or sizes of aggregates. The high throughput assays cannot detect the level of aggregation and fragmentation of BSA. Both SE-HPLC and electrophoresis indicate that for the BSA remaining in solution it is mostly composed of small soluble oligomers and aggregates. The amount of monomer structure retained is dependent on the concentration of H_2O_2 . The total amount of fragments produced as a result of oxidation was quite low, but again there was clearly a distinction between the H_2O_2 concentrations at which fragments were resolved (table 4.5). This allowed us to categorise the oxidation of BSA into two conditions; mild oxidation conditions ($< 0.3 \text{ M H}_2\text{O}_2$), and extreme oxidation conditions ($> 0.78 \text{ M H}_2\text{O}_2$).

The main difference between BSA and lysozyme (apart from the M_w) is the relative amino acid content; in particular, cysteine content, and tryptophan content. It is believed that the reason the oxidation products and size of aggregates formed between both proteins varies is due to where the radical appears to favour attack and how much this

alters the secondary structure of the protein. Oxidation of protein by H₂O₂ is believed to stem from the decomposition of H₂O₂ into the hydroxyl radical HO•, one of the most potent radicals. The generation of these radicals is the primary catalyst in the mechanism by which oxidative damage leads to protein fragmentation and aggregation (Reubsaet, Beijnen *et al.* 1998).

The identical molecular weight of the fragmented bands identified by gel electrophoresis in the early stages of oxidation of BSA, gives support to the assumption that the mode of oxidative attack of BSA is at a preferential site (Hunt, Bottoms *et al.* 1993). BSA showed evidence of fragmentation across all H₂O₂ concentrations. Due to their high abundance, regions near the 27 cysteine residues, but more specifically the disulphide bonds and thiol groups of BSA are likely to be a major target for the free radical HO•, disrupting both the tertiary and secondary structure of BSA and resulting in fragmentation as evidenced by SE-HPLC (Schuessler and Schilling 1984; Puchala and Schuessler 1993; Guedes, Vitorino *et al.* 2009). This may explain why in the initial stages of oxidation, we saw a decrease in the dimeric form of BSA with SE-HPLC; the HO• radical acts to disrupt the Cys-Cys bond between the two monomeric forms responsible for the dimer reducing it to monomer.

Detection of lysozyme fragments by gel electrophoresis or SE-HPLC was significantly lower than for BSA. Only limited fragmentation was detected by gel electrophoresis so site specific oxidation could not be confirmed or fully interpreted by comparing the molecular weight of the fragments. Evidence from gel electrophoresis would also suggest that any fragmented protein segments of lysozyme were either aggregating to the monomer remaining in the solution. Thus attempts to correlate fragmentation of lysozyme with a specific residue modification were challenging.

Both proteins responded to oxidation with a loss in intrinsic fluorescence. The tryptophan residues have been shown to be vulnerable to radical attack and loss in intrinsic fluorescence is an early occurrence of radical-mediated protein oxidation (Campos, Lissi *et al.* 1999; Ma, Chao *et al.* 1999; Wu, Zhang *et al.* 2009; Sun, Zhao *et al.* 2011). Both proteins had a completely diminished fluorescence signal under extreme oxidation conditions indicating severe conformation changes relative to mild oxidation conditions.

Of the 6 tryptophan residues in lysozyme Trp-62 and -108 make the largest contribution to the fluorescence emission (Imoto, Forster *et al.* 1972; Formoso and Forster 1975). Damage to the secondary structure by the HO• radical could have resulted in an opening of the folded native conformation of lysozyme exposing Trp-62 and -108 to oxidative damage accounting for the loss of fluorescence signal. BSA contains two tryptophan residues of which one is buried in a hydrophobic region, Trp-212, while Trp-134 is found within a hydrophobic pocket and is more surface exposed. Again damage to the secondary structure by radical attack would result in exposure of the buried Trp residue leaving it susceptible to attack. It must also be noted that the HO• radical might also be acting to directly chemically oxidise tryptophan, forming new molecular entities, accounting for the change in its intrinsic fluorescence properties, without inducing a conformational change. For the identification of tryptophan degradation compounds in proteins, it would be necessary to isolate and characterise them by enzymatic hydrolysis coupled with mass spectroscopy or RP-HPLC (Simat and Steinhart 1998). Such tryptophan degradation products detected include oxindolylalanine, kynurine and *N*-formylkynurine (Holt, Milligan *et al.* 1977; Kanner and Fennema 1987; Krogull and Fennema 1987; Itakura, Uchida *et al.* 1994; Simat and Steinhart 1998).

The tryptophan residue is a hydrophobic amino acid known to both red shift and blue shift with protein oxidation (Wu, Zhang *et al.* 2009). Protein hydrophobicity has been the object of numerous oxidation studies, where oxidation enhances protein surface hydrophobicity by exposure of hydrophobic amino acids, which are buried in native protein (Davies and Delsignore 1987; Liang 1999; Sante-Lhoutellier, Aubry *et al.* 2007; Wu, Zhang *et al.* 2009). Exposure of free hydrophobic tryptophan residues, would coincide with a red shift of emission with the changing of the surface hydrophobicity of both proteins. Yet we observed a blue shift in our intrinsic fluorescence indicating a more buried and nonpolar environment (Lakowicz 2006). It is therefore assumed that the exposed tryptophan residues, among others exposed, rapidly attach to other exposed hydrophobic residues via non-specific hydrophobic associations, inducing aggregation (by Mechanism 2) and further burial of the hydrophobic residues occurs. Aggregation of lysozyme and BSA results from the disruption and rearrangement of the secondary structure, exposing regions that favourably aggregate. Aggregation due to favourable aggregation of exposed hydrophobic regions as a consequence of oxidation has been

shown for myofibrils (Sante-Lhoutellier, Aubry *et al.* 2007) and in the formation of visible aggregates of peanut protein isolate (Ye, Liao *et al.* 2013).

BSA aggregated to a much lesser extent than lysozyme. Again this could be linked to the site of oxidative attack or just the overall number of sites available to attack in the peptide chain. BSA has 607 amino acid residues (538 in the secreted form) compared to 128 in lysozyme, thus BSA becomes more damaged than lysozyme due to the greater number of amino acids present and susceptible to attack. This is obvious when we compared the relative amounts of monomer of each protein after oxidation. BSA is mainly aggregated into small soluble aggregates with very little monomer structure retained. While lysozyme forms larger aggregates, there is actually less monomer lost to oxidation with similar oxidation conditions (table 4.4, 4.3 and 4.5).

For lysozyme, the solution pH dependence on the degree of oxidation and/or aggregation between solution pH 5 and pH 8, may be explained by the charge distribution on the surface of lysozyme. Several studies point to the mode of oxidative attack being pH dependent (Tien, Berlett *et al.* 1999; Clague 2013; Lee, Baek *et al.* 2013). The exact reason why solution pH influences the degree of oxidation is speculative but it does imply that solvent exposure of particular residues susceptible to oxidation is likely important in the mode of modification; residues that are fully solvent exposed will show a maximal degree of oxidation compared to residues that may be buried in the native conformation, as shown for the oxidation of methionine residues (Manning, Chou *et al.* 2010). It is also possible that the solution pH difference doesn't influence the action of H₂O₂ to modify lysozyme but it more likely regulates the formation and morphology of the aggregates that are formed as a consequence of the structural modifications. The loss in secondary structure appears to be greater for lysozyme at pH 5 as indicated by the greater loss in intrinsic fluorescence compared to the other solution pH values at equal H₂O₂ concentrations (table 4.1). This may account for the lower levels and types of aggregates detected at this solution pH, like BSA, the structure is more damaged and thus cannot aggregate to the same extent. The formation of lysozyme fibrils at pH 8 compared to the formation of amorphous aggregates at pH 5, may also indicate the level of structural damage between the two conditions is different. A study on the fibrillation of β -amyloid peptide found that differences in secondary structure after shock treatment play a key role in the growth of amyloid fibrils and the rate of conversion of monomer to β -sheet dimer

is an important step (Shen and Murphy 1995). The secondary structure of lysozyme at pH 8, is obviously less damaged allowing self-association of the β -domain, initiating fibril formation providing a template for further deposition of protein and develops a stable mainly β -sheet core structure of the fibril β -sheet formation. This mechanism for lysozyme amyloid fibril formation through self-assembly of β -domains has been shown (Booth, Sunde *et al.* 1997).

The most obvious case where the level of structural damage links to the amount and aggregation is the case of extreme oxidation of lysozyme (5 M H_2O_2) across all solution conditions. These samples showed almost complete loss of monomeric lysozyme by SE-HPLC and gel electrophoresis and also showed the highest amount of fragments. In the very early stages of oxidation the lysozyme solutions indicated the formation of higher order species capable of associating with ThT, but after 24 hours the ThT fluorescence decreased completely, with only small soluble aggregates and fragments being detected. This shows how severe oxidation conditions could actually impede lysozyme aggregation relative to the other oxidation conditions. The damage to the protein was more intense, so that any protein assemblies initially formed were degraded, the consequence of which might account of the high abundance of fragments detected by gel electrophoresis; it's not only the native monomer that was being fragmented but lysozyme aggregates initially formed were also being fragmented.

4.5 Conclusions

Both BSA and lysozyme are vulnerable to oxidative attack. Oxidation of both proteins was accompanied by a gradual loss in monomer structure to backbone fragmentation, and aggregation. By monitoring the chemical modification of both lysozyme and BSA via oxidative modification, we can provide insight into the mechanism of protein aggregation and also characterise of the types of aggregates produced via different aggregation pathways (amorphous versus fibrils). Protein oxidation has been demonstrated to induce structural changes to the secondary structure of each protein. The amount of damage caused was related to how much each protein aggregated. BSA was severely damaged resulting in small soluble protein aggregates, while lysozyme was comparatively less damaged and was capable of forming large protein particles, and in certain cases forming amyloid fibrils. These results demonstrate that with the same oxidation conditions, chemical modification can be protein specific in the types of protein aggregates formed *i.e.* large protein particles and fibrils in lysosome compared to particulate free solutions in BSA.

The difference in oxidation products formed between lysozyme and BSA is likely a result of accessibility/number of sites for attack. While the exact site of oxidative attack is still ambiguous, what is clear is that the extent of aggregation after oxidation is quite different and that different modes of detection are required to capture all of this information.

Chapter 5 Second Derivative Spectroscopy to Monitor Protein Unfolding

5.1 Introduction

5.1.1 Protein unfolding with chemical denaturants

Upon denaturation, proteins lose secondary and tertiary structure as they unfold resulting in exposure of hydrophobic residues to the external solvent environment (Roberts and Wang 2010). The mechanism by which a protein acquires its unfolded conformation is a major question and is still not fully understood despite several decades of research on this topic. The ensemble of unfolding states can be characterised by diverse structural and conformational properties. Any advancement in unravelling this mechanism will include detection and characterisation of the folded/unfolded conformation along with any partially folded states (intermediates) (Duy and Fitter 2006).

The presence of a chemical denaturant can modify the optical properties of the protein as it changes from its native structure to an unfolded state. These observed signal changes can vary from changes in intensity of emitted light, fluorescence lifetime, fluorescence anisotropy, quantum yield and emission wavelength position (Eftink 1994). By monitoring such changes with respect to environment changes, one can begin to reveal the protein unfolding pathway (Sancho 2013).

Chemical denaturants such as urea and guanidine hydrochloride (GdmHCl) play a central role in protein-unfolding experiments. GdmHCl is capable of unfolding proteins at substantially lower concentrations than urea (Greene and Pace 1974). At sufficiently high denaturant concentrations, almost all proteins can be unfolded while remaining soluble (for both of these denaturants) (Eaton, Munoz *et al.* 2000).

5.1.2 Spectroscopy and protein unfolding

Protein unfolding due to either chemical or thermal denaturation, is generally performed in bulk where the dynamics of the system, averaged over the whole population of molecules, is measured (Zocchi 1997). Fluorescence spectroscopy can be used to probe global kinetics, solvent exposure, motional anisotropy, the formation of specific tertiary contacts and distance correlation functions (Gruebele 1999). In bulk experiments, fluorescence intensity measures the ratio of the population in the native state to that in the unfolded state as environmental conditions are changed. Parameters such as fluorescence

intensities, fluorescence lifetime, spectral shift, and anisotropy can all be used to monitor conformational changes indicative of protein unfolding (Valeur 2002; Lakowicz 2006).

The intrinsic fluorescence of proteins arises from the aromatic residues; tryptophan, tyrosine and phenylalanine of which tryptophan is the most commonly used probe and which has a maximum in the excitation wavelength between 290 – 300 nm, whereas the maximum excitation wavelength for tyrosine occurs at 280 nm (Chen and Barkley 1998). The fluorescence spectrum of tryptophan is highly solvent dependent and thus can be used to detect the micro-environment of tryptophan residues. If excitation of tryptophan at 295 nm is performed, the resulting emission wavelength reflects the polarity of the local environment of the indole side chains (Mann, Royer *et al.* 1993). Changes in the emission spectra can reveal the location of tryptophan residues; an exposed surface residue will occur at longer wavelengths (red shifted) than that from an interior tryptophan residue. Tryptophan fluorescence shifts from $\lambda_{\max} \approx 300 - 350$ nm as it moves from the hydrophobic core, moving to a more polar exterior (Royer, Mann *et al.* 1993; Gruebele 1999). Therefore shifts in tryptophan spectra upon unfolding can indicate exposure of the residues to the solvent surrounding it making it a useful probe for monitoring unfolding reactions (Lakowicz 2006). In principle, if the structure and location of the intrinsically fluorescent residues of a protein are known, then changes in fluorescence properties can be interpreted and related to changes in protein conformation and structure (Chen and Barkley 1998).

Various spectroscopic techniques have been widely employed to study the changes within a protein due to chemical denaturation (Lakowicz 2006; Togashi and Ryder 2006; Togashi and Ryder 2008). These are advantageous in their sensitivity and selectivity to very low protein concentrations, simplicity of basic data acquisition, temporal resolution and the only major requisite is that equimolecular solutions of native and non-native/unfolded protein must have discernible differences in values of their signals upon comparison of their spectra (Eftink 1994; Sancho 2013).

5.1.3 Derivative spectroscopy to monitor protein unfolding

Monitoring structural or conformational changes of a protein as it unfolds in an easy, reliable, non-invasive way which has advanced with a new generation of spectrometers

equipped with simple and fast microprocessors, allowing interpretation of results from classical spectroscopic techniques (Mozo-Villarias 2002). Recently, derivative methods have begun to be used in structural studies taking advantage of established fluorescence and absorbance techniques (Mozo-Villarias 2002; Kumar, Sharma *et al.* 2005) and are sensitive and dependable tools for monitoring and characterizing the transitions taking place in the micro-environments of aromatic residues within the protein structure. These methods can monitor minute changes in the protein spectra that may not be as easily distinguishable with zero order spectra. Calculating the derivative of an absorbance spectrum enhances resolution and facilitates de-convolution of complex zero order spectra which are composed of multiple overlapping spectra arising from the absorbance of the three aromatic residues (phenylalanine, tyrosine and tryptophan) (Ichikawa and Terada 1979; Ojeda and Rojas 2004). Thus, zero order aromatic spectra can contain multiple components that can be isolated through the use of resolution-enhancement techniques such as derivative analysis (Kuelto, Ersoy *et al.* 2003). Numerous studies are emerging showing that second derivative spectroscopy of both absorbance and fluorescence spectra allow semiquantification of spectral changes and these can be related to environmental characteristics of aromatic residues in the proteins and therefore conformational changes within the protein (Donovan 1973; Solli and Herskovits 1973; Ragone, Colonna *et al.* 1984; Mach and Middaugh 1994; Mozo-Villarias 2002; Kuelto, Ersoy *et al.* 2003). Lucas *et al.* showed second derivative UV absorbance spectroscopy provided information regarding features of protein structure through small shifts in the magnitude and position of absorbance/fluorescence wavelengths that were influenced by protein charge, size and local environment (Lucas, Ersoy *et al.* 2006). Breydo *et al.* based their method for determining pKa values for tyrosine residues on the decrease in the absorption intensity in the second derivative UV spectrum (Breydo, Shevchenko *et al.* 1997). While Kumar *et al.* used second derivative Trp fluorescence spectroscopy to identify partially unfolded states of proteins during formulation (Pace 1986; Kumar, Sharma *et al.* 2005).

5.1.4 Unfolding of BSA

BSA is a multi-domain protein (three domains, each with two sub-domains) making measurements and analysis of the unfolding process more complex, since each domain is capable of dependent or independent unfolding behaviour (Das, Chitra *et al.* 2004; Batey,

Scott *et al.* 2006). Previous work has shown that the aromatic amino acid residues of BSA are not distributed equally throughout the 3 domain protein (Huang, Kim *et al.* 2004). The 27 phenylalanine and the 20 tyrosine residues are found in all three domains in the protein structure. There are only two tryptophan residues in the entire protein. One buried in a hydrophobic region, Trp-212 situated in subdomain IIA. This Trp is important in the hydrophobic interaction between subdomains IIA and IIIA. While Trp-134 is found in the second helix of domain I within a hydrophobic pocket that acts as a binding domain and is more surface exposed (figure 5.1) (He and Carter 1993; Ma, Nishikawa *et al.* 2006; Majorek, Porebski *et al.* 2012). Spectroscopic properties of tryptophan residues are known to be sensitive to their environment (Eftink 1994; Lakowicz 2006), therefore as only two of the three domains contain tryptophan, and one is surface exposed while the other is buried in a hydrophobic region, the analysis of the micro-environment of this amino acid can prove useful in understanding local changes induced by chemically unfolding the protein. The tryptophan residues of BSA has been known to blue shift, indicative of an increase in the hydrophobicity of the residues (Sulkowska, Bojko *et al.* 2003), while red shifts were attributed to the exposure of BSA residues to a more polar environment (Itri, Caetano *et al.* 2004).

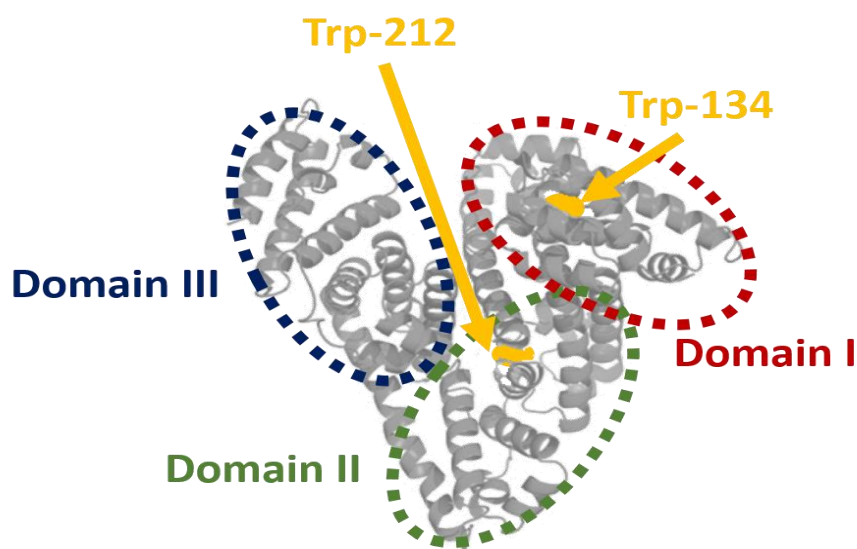


Figure 5.1: Structure of BSA showing the three domains and the location of the two tryptophan residues.

The unfolding of BSA and HSA using urea and GdmHCl as chemical denaturants has previously been well documented and shown to follow a multi-state transition, forming stable intermediates around 4 – 5 M urea and ~2 M GdmHCl (Khan, Agarwal *et al.* 1987; Flora, Brennan *et al.* 1998; Tayyab, Sharma *et al.* 2000; Gonzalez-Jimenez and Cortijo

2002; Sulkowska, Bojko *et al.* 2003; Das, Chitra *et al.* 2004; Ahmad, Ahmed *et al.* 2005; Sulkowska, Rownicka *et al.* 2005; Adel, Nadia *et al.* 2008; Togashi, Ryder *et al.* 2010; Anand, Jash *et al.* 2011).

Several studies point to domain I undergoing significant unfolding along with domain III during early urea denaturation (Tayyab, Sharma *et al.* 2000; Tayyab, Ahmad *et al.* 2002). While other studies conform to the idea that domain I of BSA does not play a major role in the initial unfolding of BSA, but it is in fact domain III which is the least stable domain and initially unfolds, followed by domain II and finally domain I (Khan, Agarwal *et al.* 1987; Tanaka, Nishizawa *et al.* 1997; Tayyab, Sharma *et al.* 2000; Ahmad, Ahmed *et al.* 2005; Togashi, Ryder *et al.* 2010). Therefore, despite many studies, using the most up-to-date techniques, some debate about the unfolding of BSA remains.

5.1.5 Unfolding of lysozyme

Lysozyme is a small two-domain protein (α - and β - domain), containing six tryptophan residues, two of which (Trp-62 and Trp-108) are partially exposed to the solvent, and with Trp-62 making the biggest contribution to its intrinsic fluorescence (Formoso and Forster 1975). Lysozyme has been extensively studied as a model for understanding the mechanism of protein stability, folding and denaturation by a combination of techniques including spectroscopic and physicochemical methods, CD, DSC and NMR (Evans, Topping *et al.* 1991; McKenzie and White 1991; Dobson, Evans *et al.* 1994; Fink, Calciano *et al.* 1994; Sugai and Ikeguchi 1994; Wilson, Ford *et al.* 1995; Blumlein and McManus 2013).

Unfolding of lysozyme from its native state into a randomly coiled conformation is often considered an example of a two-state mechanism under equilibrium conditions (Greene and Pace 1974; Sugai and Ikeguchi 1994; Pardon, Haezebrouck *et al.* 1995; Chang and Li 2002; Sassi, Onori *et al.* 2011);



The unfolding of lysozyme with GdmHCl has been shown to follow a two-state model, being fully unfolded by 6 M GdmHCl (Laurents and Baldwin 1997). The transition region for lysozyme unfolding was shown to occur between 1.5 – 5 M GdmHCl (Sirotkin and Winter 2010). Lysozyme with GdmHCl up to 5 M has been shown to retain a certain

degree of intramolecular order, with the α -helical structure remaining intact, thus not forming a completely random coil at 5 M and maintaining up to 70% of its enzymatic activity, although denaturation was believed to be complete (Tanford, Pain *et al.* 1966; Hedoux, Krenzlin *et al.* 2010; Huerta-Viga and Woutersen 2013).

For urea induced unfolding of lysozyme, a transition region has been shown to occur between 2.5 – 5 M urea (Timasheff and Xie 2003) while earlier reports showed with addition up to 4 M urea all major conformational changes involving the active site (location of Trp-62 and Trp-108) had occurred, with little further change up to 9 M (Barnes, Warren *et al.* 1972; Mcguire and Feldman 1975). Crystallographic studies of urea-lysozyme interactions also showed no significant conformational changes with increasing urea concentrations (0.7 – 5 M) (Pike and Acharya 1994). Urea has been shown to change the tertiary structure into a “softer” structure similar to the molten globule, without unfolding of the secondary structure, while comparable studies with GdmHCl showed it clearly induced conformational changes (Hedoux, Krenzlin *et al.* 2010). Lysozyme has been shown upon unfolding to blue shift reflecting burial of the tryptophan residues into a less polar environment where a more compact conformation is achieved, while a red shift with increasing unfolded/denatured structure indicating exposure of residues to a more polar environment have also been observed (Mcguire and Feldman 1975; Pastor, Ferrer *et al.* 2007; Wu, Ma *et al.* 2008; Balme, Guegan *et al.* 2013; Esquembre, Sanz *et al.* 2013).

5.2 Aims of the study

The aim of this work was to study the unfolding effects of two denaturants, urea and guanidine hydrochloride on the structure of two model proteins, BSA and lysozyme. Using UV absorbance, fluorescence spectroscopy and second derivative analysis, a system was developed to study the protein structure in solution and the stages involved in protein unfolding. The first objective of this study was to replicate previous studies on urea and GdmHCl induced unfolding of the model proteins using the well-characterised technique of intrinsic fluorescence. This was then used to compare the validity of second derivative UV absorbance spectroscopy.

We show how the use of second derivative UV spectroscopy allows the monitoring of the individual aromatic amino acids; tyrosine, tryptophan and phenylalanine, while second

derivative fluorescence spectroscopy monitors tryptophan and tyrosine residues. Direct de-convolution of overlapped spectra of proteins allows us to simultaneously and precisely monitor individual amino acids with simple techniques and uncomplicated data analysis. This allows for more informative details about the microenvironment of the individual residues as they undergo unfolding when compared to the zero order spectra. Using this technique we can monitor and characterize spectral changes at a single amino acid and relate these to unfolding events at specific regions/domains within the protein structure thereby gaining a more detailed picture of the unfolding mechanism of the protein. This study shows that second derivative UV spectroscopy is a sensitive and reliable technique and can elucidate further the mechanism by which urea and GdmHCl induced unfolding of BSA and lysozyme occurs. We find evidence of a step wise unfolding event with intermediate structures, between the fully folded and fully unfolded states, for BSA which is consistent with other studies. While a two-state model for lysozyme unfolding was also confirmed.

5.3 Results

5.3.1 Unfolding of BSA

5.3.1.1 Fluorescence spectroscopy of BSA unfolding

Intrinsic fluorescence was used initially to assess the unfolding of BSA with two denaturants, urea (0 – 10.5 M) and guanidine hydrochloride (0 – 7.5 M). Intrinsic fluorescence allowed us to monitor characteristic features related to the behaviour of both the tyrosine and tryptophan residues as BSA underwent unfolding with increasing concentrations of denaturant. The resulting spectra for BSA with increasing concentrations of urea are shown in Figure 5.2.

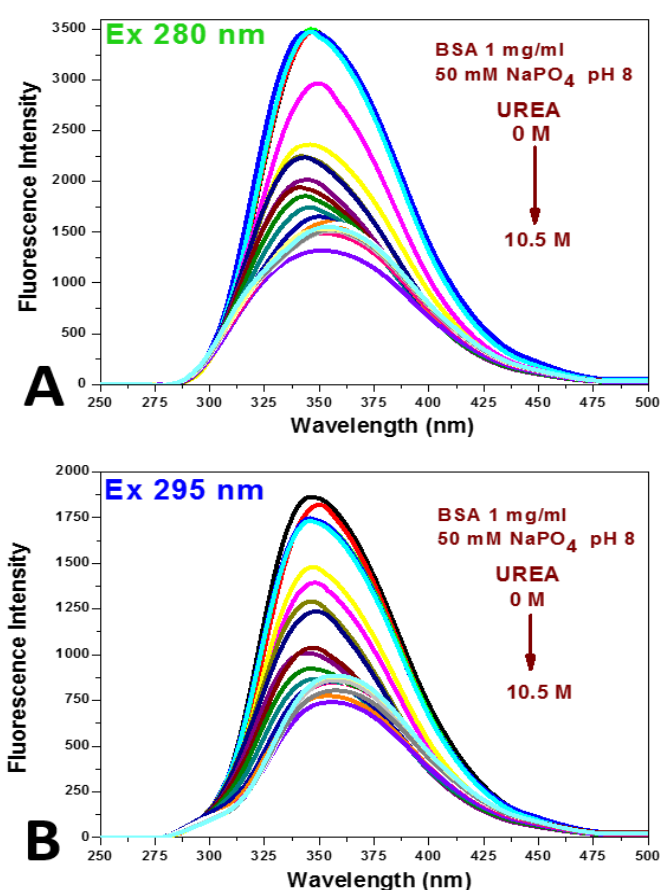


Figure 5.2: Intrinsic fluorescence spectra for BSA with increasing concentrations of urea. Excitation at A) 280 nm and B) 295 nm.

Native BSA (50 mM sodium phosphate buffer, pH 8) with excitation at 280 nm gave rise to a single peak associated with the fluorescence for tyrosine and tryptophan with a λ_{max} at 348 nm and excitation at 295 nm representing the fluorescence of tryptophan alone resulted in a λ_{max} at 350 nm. As the molar concentration of denaturant was increased, the

fluorescence properties of BSA were altered, implying structural alterations as a result of the protein unfolding and exposure of the aromatic residues to the surrounding solvent (figure 5.3). Both the intrinsic fluorescence intensity decreased and λ_{max} shifted with increasing denaturant concentration.

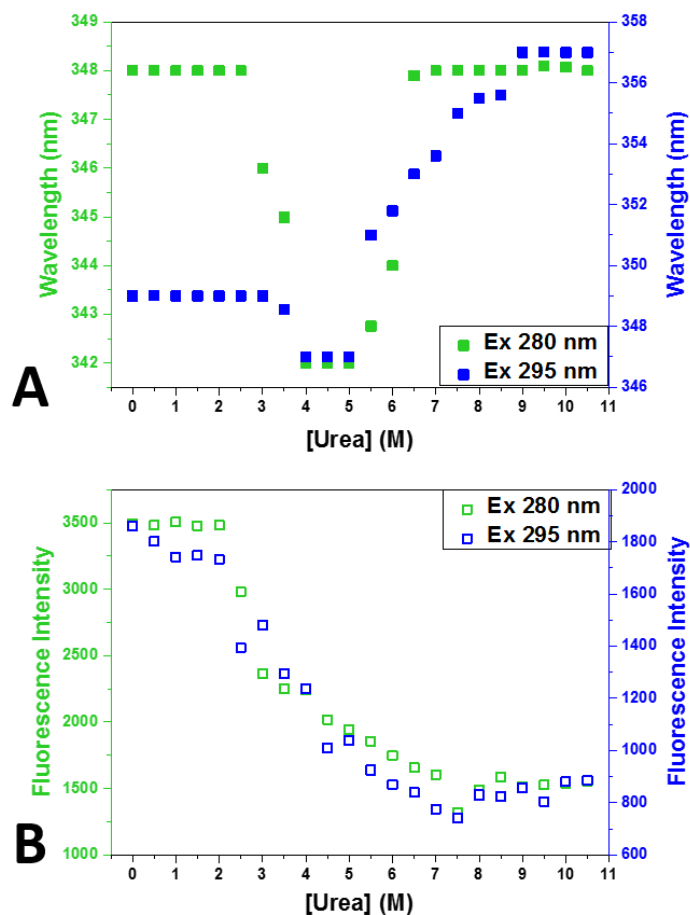


Figure 5.3: Change in BSA intrinsic fluorescence with increasing urea concentration. A) Wavelength shift. B) Intrinsic fluorescence intensity.

The emission peak (Ex at 280 nm) had an initial blue shift from 348 nm to a minimum at 342 nm between 2.5 – 4 M urea, and from 349 nm to 347 nm between 3 – 4 M when excited at 295 nm. These features are usually indicative of an increase in hydrophobicity, *i.e.* that the Trp residues have moved to a less polar inner core due to a partial rearrangement of the domains within BSA, or that alteration in the packing structure of the domains as the distance of separation between them changed, suggesting the presence of a transitional stage during the unfolding process (Sulkowska, Bojko *et al.* 2003). This was followed by a significant red shift between 4 – 5 M urea with an emission maximum at 356 nm (Ex 280) and 364 nm (Ex 295) suggesting the exposure of BSA Trp residues to a more polar environment (Itri, Caetano *et al.* 2004). As a result of further domain

rearrangement/unfolding, water molecules and/or other amino acids quench the fluorescence signal. This was seen as a decrease in the protein fluorescence emission intensity, initiated at 2.5 M urea. This indicated the onset of structural re-arrangement and at urea concentrations of 8.5 M and higher there was no further change in fluorescence intensity or λ_{\max} . At this point it is assumed that BSA is fully unfolded and the Trp residues are fully exposed to the solvent.

BSA exhibited similar intrinsic fluorescence properties when GdmHCl was used as the denaturant (figure 5.4). Both quenching of the fluorescence intensity of both residues and shifts in the emission peak wavelength (λ_{\max}) were observed.

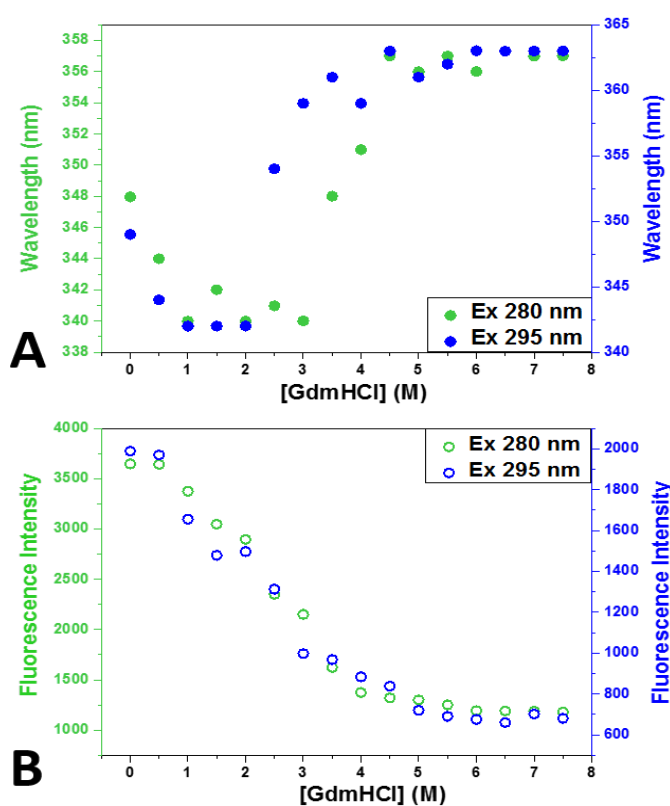


Figure 5.4: Change in BSA intrinsic fluorescence properties with increasing GdmHCl concentration. A) Wavelength shift. B) Intrinsic fluorescence intensity changes.

Again, with GdmHCl, both the fluorescence intensity and λ_{\max} change with increasing denaturation concentration, again consistent with an initial rearrangement of the BSA domain packing followed by exposure of the core amino acids to a more polar environment (Togashi, Ryder *et al.* 2010). Both emission peaks (from excitation at 280 nm and 295 nm) showed a steady blue shift from 348 nm to a minimum of 342 nm (Ex 280 nm) and from 349 nm to a minimum of 340 nm (Ex 295 nm), between 0.5 – 1 M as domain re-arrangement began. As the concentration of GdmHCl increased, there was a

red shift in the emission peak position between 2 – 5 M GdmHCl after which both the fluorescence intensity and λ_{max} remained constant, indicating that BSA was fully unfolded. This data is consistent with previous measurements using intrinsic fluorescence to measure GdmHCl induced BSA unfolding (Togashi, Ryder *et al.* 2010).

5.3.1.2 Second derivative fluorescence spectroscopy of BSA unfolding

Second derivative operations were performed on the intrinsic fluorescence spectra of BSA since previous work has shown that additional information about the solvent environment surrounding tyrosine and tryptophan residues can be observed from this analysis (Mozo-Villarias 2002). Figure 5.5 shows the overlapped second derivative spectra of the intrinsic fluorescence spectra for BSA with increasing urea concentrations (excitation at 280 nm and 295 nm).

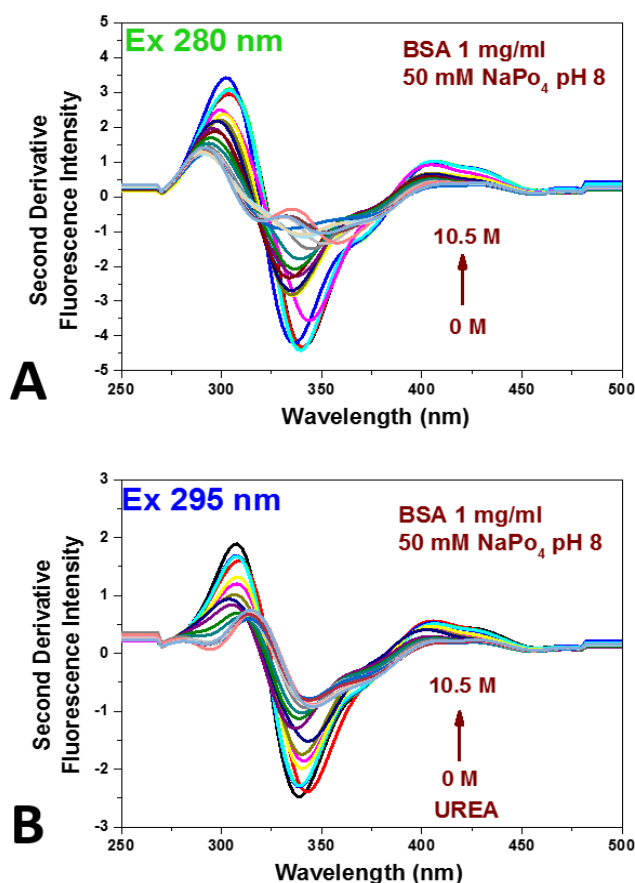


Figure 5.5: Second derivative intrinsic fluorescence spectra of BSA with increasing concentrations of urea. A) Ex 280 nm. B) Ex 295 nm.

In the second derivative spectrum, native BSA has a minimum at 340 nm (Ex 280 nm) and at 338 nm (Ex 295 nm). There was a relatively fixed maximum at 400 nm and a shoulder around 360 – 370 nm. The minima in the second order derivative spectra

correspond to peaks in the original spectrum (Gopalan, Golbik *et al.* 1997) and thus based on the position of the minima in the second derivative spectrum, conformational information about the protein can be extracted. The positions (red or blue shifts) and relative magnitudes of the minima depend on the solvent environment (Mozo-Villarias 2002; Kumar, Sharma *et al.* 2005). Previous work has shown that changes in the fluorescence emission at these wavelengths are the result of various microstates of tryptophan residues with protein unfolding (Gopalan, Golbik *et al.* 1997). The wavelength peak position and intensity of the peak minima were recorded as a means to monitor protein conformational changes during urea induced unfolding of BSA (figure 5.6). Comparison of the data obtained from the intrinsic fluorescence zero order data with the second derivative data indicate that little additional information about the protein unfolding pathway can be determined from the second derivative data.

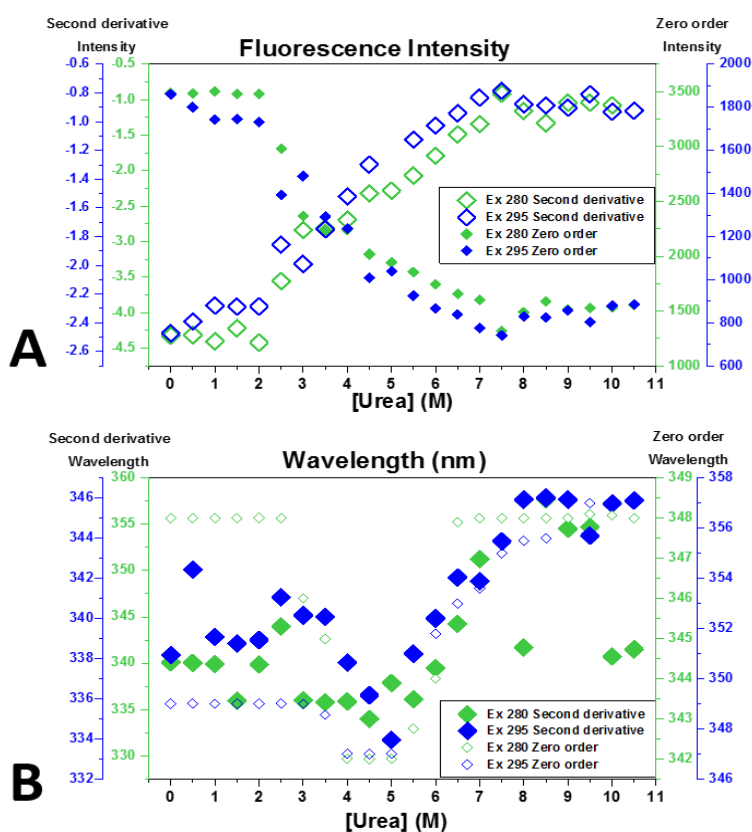


Figure 5.6: Change in BSA second derivative intrinsic fluorescence compared to the zero order intrinsic fluorescence with increasing urea concentration. A) Fluorescence intensity. B) Wavelength shift.

The change in fluorescence intensity in the second derivative intrinsic fluorescence spectra appeared to mirror that of the zero order; initial unfolding by 2 M urea, intermediate states from 2.5 – 7 M with no further change in the fluorescence intensity

indicating fully unfolded BSA by 8 M. The wavelength shift in the second derivative spectrum (for excitation at 295nm) was similar to the zero order data; the characteristic blue shift to shorter wavelengths by 3.5 M urea, followed by a red shift at 5 M urea, and a levelling off by 7.5 M where BSA was fully unfolded. The emission peak positions for excitation at 280 nm appeared more scattered compared to the zero order wavelength data.

In the second derivative intrinsic fluorescence spectra (excitation at 280nm), at the higher urea concentrations (6.5 – 10.5 M urea) where BSA is mostly likely fully unfolded there was a splitting of the peak minimum into two smaller peaks that successively increased with increasing urea concentration (figure 5.7). This change in the peak shape was not observed for any of the spectra recorded using 295 nm as the excitation wavelength.

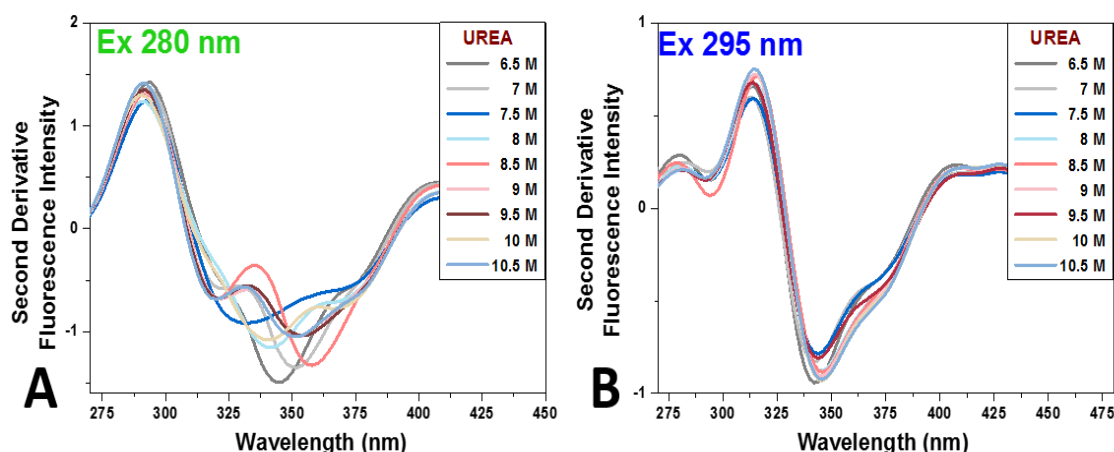


Figure 5.7: Second derivative fluorescence spectra of BSA showing the change in peak shape with increasing concentrations of urea. A) Ex 280 nm. B) Ex 295 nm.

The second derivative intrinsic fluorescence spectra of BSA unfolded with GdmHCl were also analysed in this manner. Again the second derivative fluorescence emission properties of BSA were comparable with the zero order intrinsic fluorescence spectral properties. The most striking feature of the second derivative analysis of BSA unfolding with GdmHCl, were the alterations to the spectra (Ex at 280 nm) for the high denaturant concentrations (5 – 7.5 M) where BSA was believed to be fully unfolded (figure 5.8).

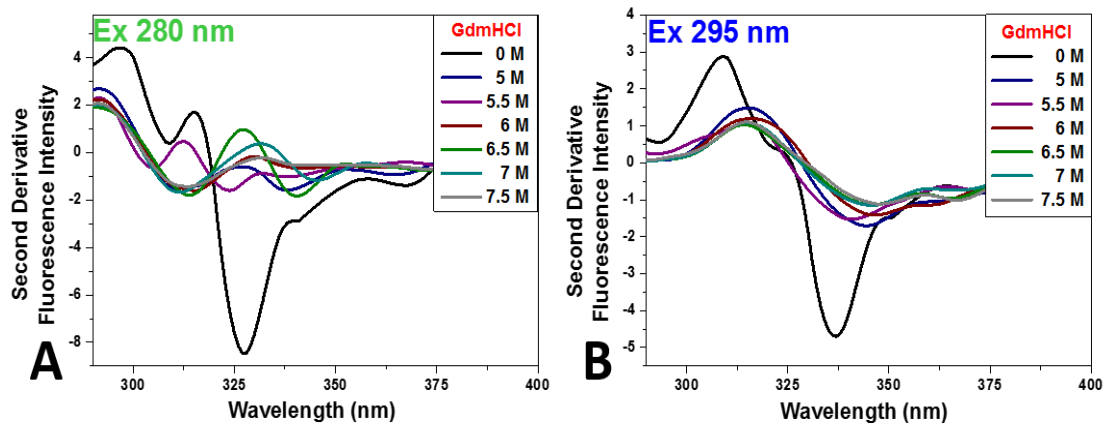


Figure 5.8: Second derivative fluorescence spectra of BSA showing the change in peak shape with increasing concentrations of GdmHCl. A) Ex 280 nm. B) Ex 295nm.

5.3.1.3 Zero order UV absorbance of BSA unfolding

The zero order UV absorbance of BSA with increasing concentrations of urea and GdmHCl were recorded. The normalised spectra overlaid for BSA for urea and GdmHCl are shown in Figure 5.9. The zero order UV absorbance spectra obtained show a single characteristic protein peak with a maximum at 280 nm. There is considerable overlap in the wavelengths at which the three aromatic amino acids absorb and they cannot be distinguished clearly in the zero order spectrum.

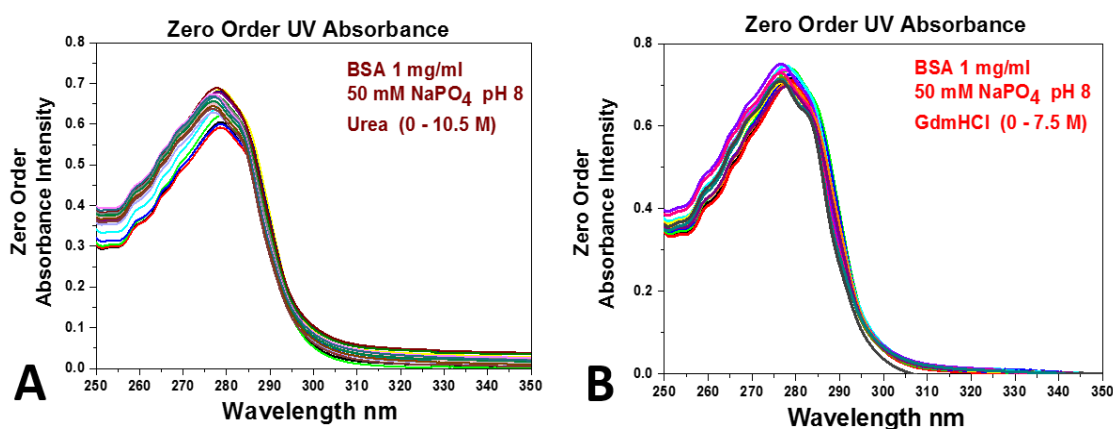


Figure 5.9: Zero order absorbance spectra of BSA with increasing concentrations of denaturant. A) GdmHCl. B) Urea.

Monitoring of the λ_{\max} intensity values did not reveal any information about the unfolding transitions of BSA except for a slight increase in the absorption values with increasing denaturant concentration. However a notable shift in λ_{\max} was observed (figure 5.10).

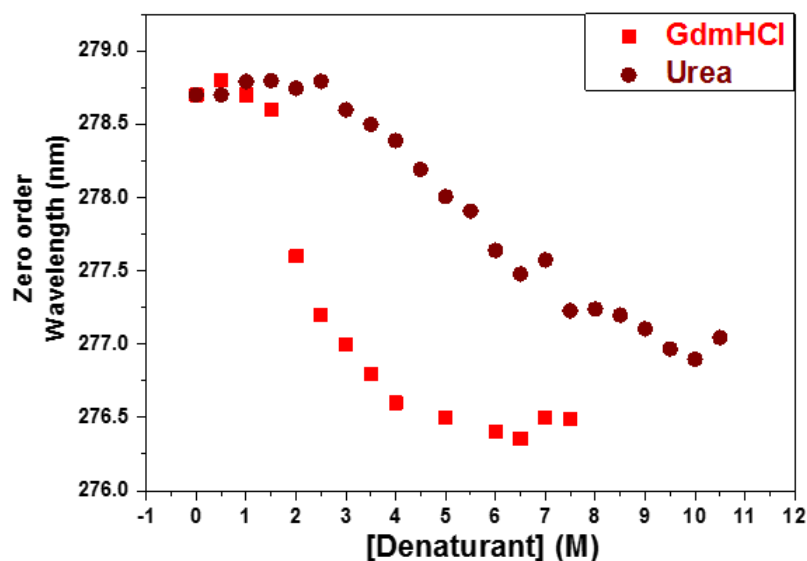


Figure 5.10: Shift in λ_{\max} from the zero order absorbance spectra of BSA with increasing concentrations of urea and GdmHCl.

In the absence of denaturant, the λ_{\max} for BSA occurs at 278.7 nm. With the addition of denaturant the λ_{\max} shifts to lower wavelengths. For GdmHCl this shift was initiated at 1 M GdmHCl and levelled off by 5 – 6 M. For urea induced denaturation the wavelength shift was observed at 3 M and was complete at 7.5 M urea. The molar concentrations of these events detected by zero order UV absorbance were consistent with the intrinsic fluorescence measurements. However no further insight into domain specific unfolding can be inferred from this analysis.

5.3.1.4 Second derivative UV absorbance during BSA unfolding

A second derivative operation was carried out on the zero order spectra of BSA with increasing concentrations of urea and GdmHCl. In contrast to the zero order spectrum, the second derivative spectrum de-convoluted the single protein associated peak at ≈ 280 nm into several peaks (Ichikawa and Terada 1979). Those associated with the three aromatic amino acids are shown in Figure 5.11. Tyrosine and tryptophan can be seen at strong negative peaks at ~ 280 nm and ~ 290 nm respectively. Phenylalanine in proteins manifests itself as weak ripples or shoulders within the zero order absorbance spectra between 250 – 270 nm region (Schmid 2001), resulting in several negative bands within the second derivative spectrum within this region (Balestrieri, Colonna *et al.* 1978). However it is the strongest band at 259 nm that is monitored for evaluation purposes

(Ichikawa and Terada 1977; Mach, Thomson *et al.* 1991; Lucas, Ersoy *et al.* 2006). For each of the three residue peaks the wavelength value at the peak minima was recorded

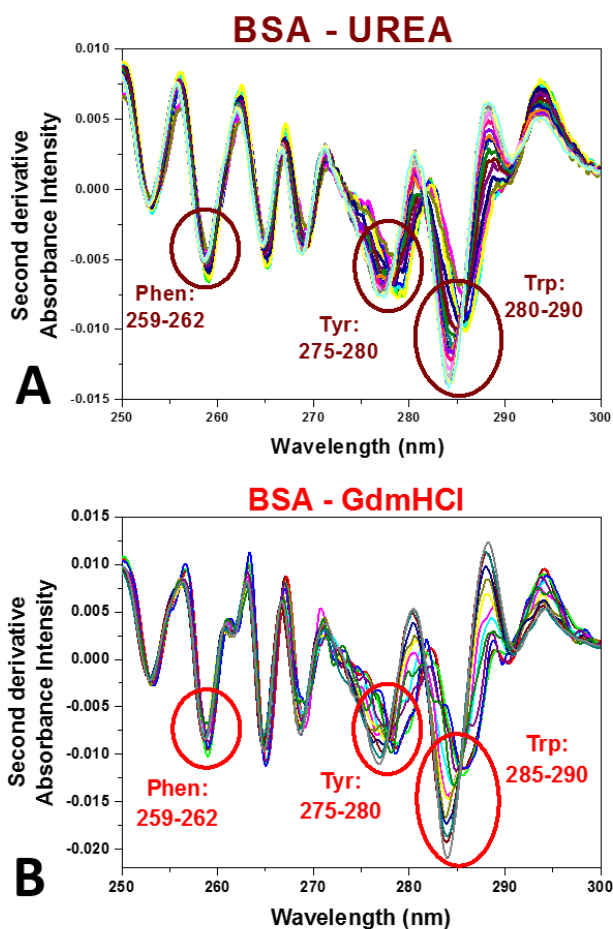


Figure 5.11: Second derivative absorbance spectra of BSA with increasing concentrations of denaturant. A) GdmHCl. B) Urea.

The second derivative UV absorbance spectra allow undefined peaks and shoulders from the zero order to be transformed into distinct negative peaks having minima centred at different wavelengths for each individual aromatic residue thus improving the resolution of the overlapping peak at 280 nm. Depending on the concentration of denaturant, the shape of the peak and the wavelength minima positions shifted; phenylalanine (259 – 262 nm), tyrosine (275 – 280 nm) and tryptophan (285 – 290 nm). These changes in the λ_{\min} with increasing urea concentration are shown in Figure 5.12.

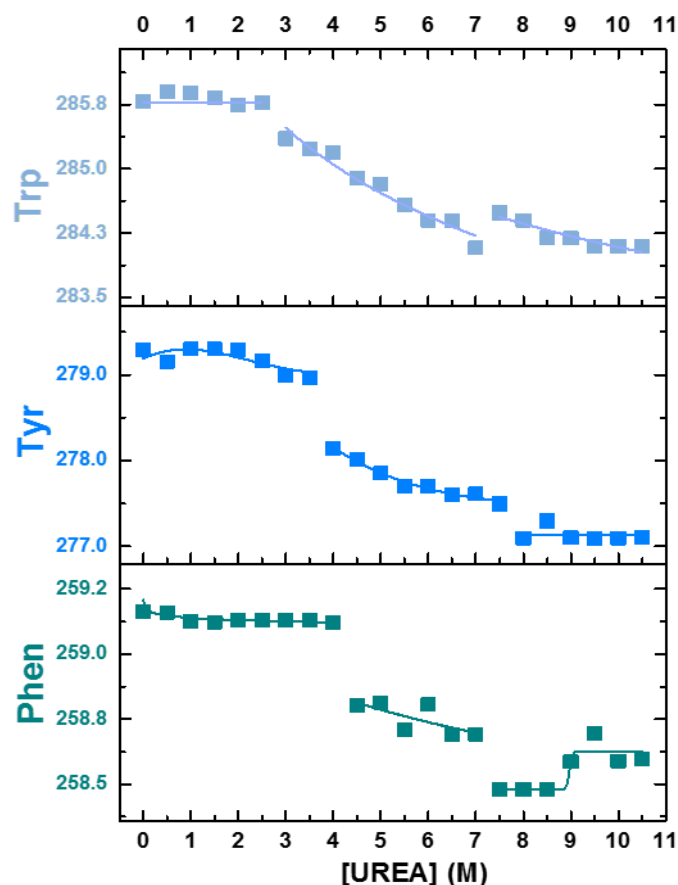


Figure 5.12: Effect of urea concentration on the second derivative UV wavelength position for the three aromatic amino acid residues in BSA.

The wavelength position for each residue blue shifted with increasing concentrations of urea, reflecting exposure of these residues to the solvent as a consequence of the unfolding process (Lucas, Ersoy *et al.* 2006). The second derivative also distinguishes the denaturant concentration at which each residue is affected by the structural transitions occurring during denaturation. This urea concentration at which this happened was not the same for each aromatic amino acid (table 5.1). Phenylalanine exhibits initial changes in λ_{\min} at a urea concentration of 4 M, with tryptophan and tyrosine indicating signs of unfolding at lower urea concentrations, 2.5 and 3.5 M respectively.

| <i>Amino acid</i> | <i>On set of unfolding</i> | <i>Partially unfolded intermediates</i> | <i>Fully unfolded protein</i> |
|-------------------|----------------------------|---|-------------------------------|
| <i>Trp</i> | 0 – 2.5 M | 3 – 7 M | 8 M |
| <i>Tyr</i> | 0 – 3.5 M | 4 – 7.5 M | 8 M |
| <i>Phen</i> | 0 – 4 M | 4.5 – 7 M | 8 M |

Table 5.1: Urea concentrations of domain unfolding for BSA

GdmHCl induced unfolding of BSA showed similar trends (figure 5.13 and table 5.2)

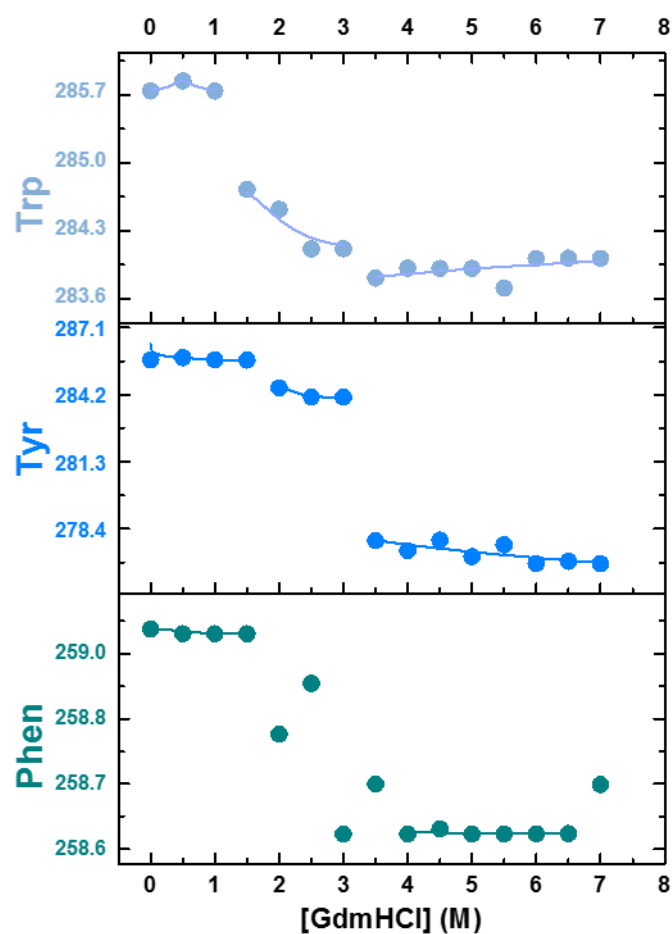


Figure 5.13: Effect of GdmHCl on the second derivative UV spectroscopic wavelength position for the three aromatic amino acid residues in BSA.

The tryptophan residue showed the earliest indication of structural alteration at 1 M GdmHCl. Unfolding was complete at 3.5 M GdmHCl. Phenylalanine and tyrosine signals indicated that unfolding was initiated at 1.5 M GdmHCl and was complete by 3.5 and 4 M GdmHCl respectively.

| <i>Amino acid</i> | <i>On set of unfolding</i> | <i>Partially unfolded intermediates</i> | <i>Fully unfolded protein</i> |
|-------------------|----------------------------|---|-------------------------------|
| <i>Trp</i> | 0 – 1 M | 1.5 – 3 M | 3.5 M |
| <i>Tyr</i> | 0 – 1.5 M | 2 – 3 M | 3.5 M |
| <i>Phen</i> | 0 – 1.5 M | 2 – 3.5 M | 4 M |

Table 5.2: GdmHCl concentrations of domain unfolding for BSA.

5.3.2 Unfolding of lysozyme

5.3.2.1 Fluorescence spectroscopy of lysozyme unfolding

Changes in the intrinsic fluorescence of lysozyme were monitored with increasing concentrations of urea (figure 5.14).

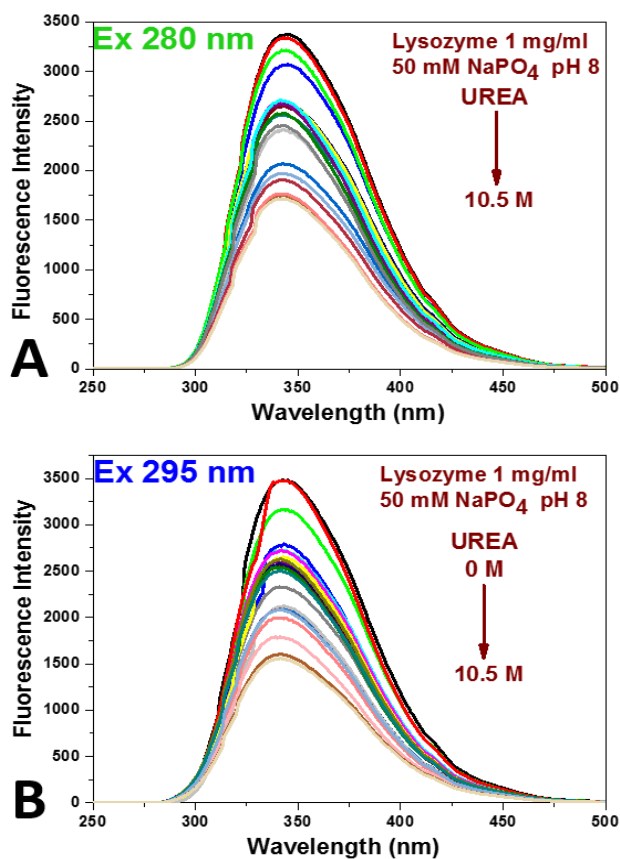


Figure 5.14: Intrinsic fluorescence spectra for lysozyme with increasing concentrations of urea A) Excitation at 280 nm. B) Excitation at 295 nm.

Native lysozyme (50 mM sodium phosphate buffer, pH 8) had a λ_{\max} at 345 nm when excited at 280 nm, and 343 nm when excited at 295 nm. With increasing concentrations of urea, shifts in the wavelength peak position and a loss in the fluorescence intensity were observed (figure 5.15).

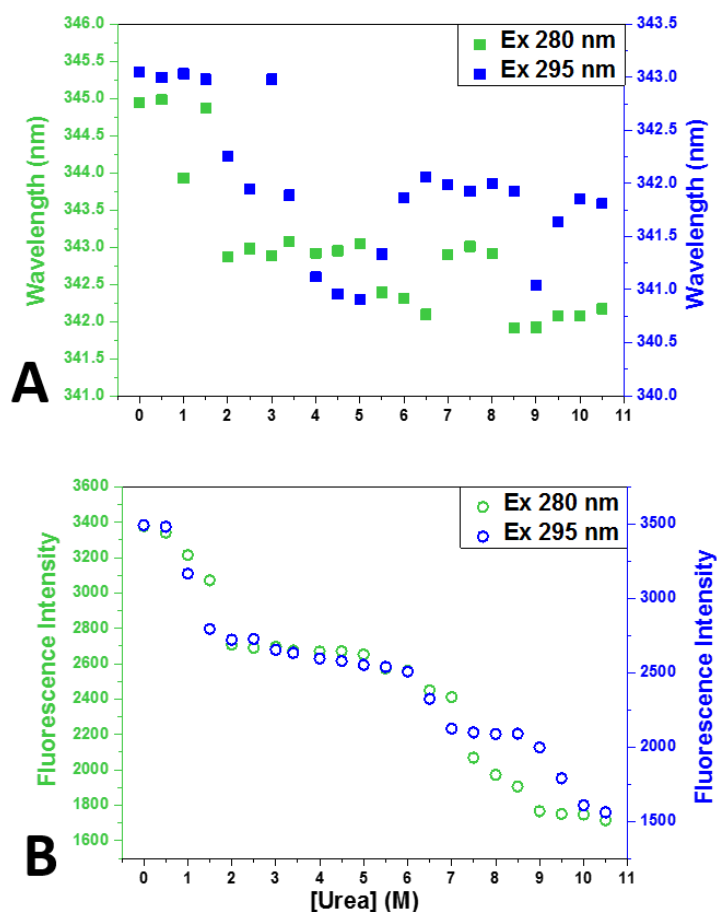


Figure 5.15: Change in lysozyme intrinsic fluorescence properties with increasing urea concentration. A) Wavelength shift. B) Intrinsic fluorescence.

For both excitation at 280 nm and 295 nm, the wavelength peak position initially blue shifts to shorter wavelengths attributed to the burial of the tryptophan side chains in an environment of lower polarity between 1 – 2 M urea (Ex 280 nm) to a minimum of 342 nm at 6.5 M urea, and at 1.5 M urea (Ex 295 nm) to a minimum at 341 nm at 5 M urea. For Ex 295 nm this initial blue shift was then followed by a red shift to a maximum at 342 nm at 6.5 M urea indicating increasing exposure of the residues to the solvent. For Ex 280 nm a slight red shift occurs between 7 – 8 M urea. Again both excitation wavelengths show a second blue shift between 8.5 – 10.5 M urea, the concentrations that indicate lysozyme is fully unfolded. The fluorescence spectra steadily decreased in intensity at both excitation wavelengths as lysozyme goes from its folded conformation to an unfolded state.

GdmHCl induced unfolding of lysozyme did not result in a significant initial blue shift in wavelength position, but had a steady decrease in fluorescence intensity as a function of denaturant concentration (figure 5.16).

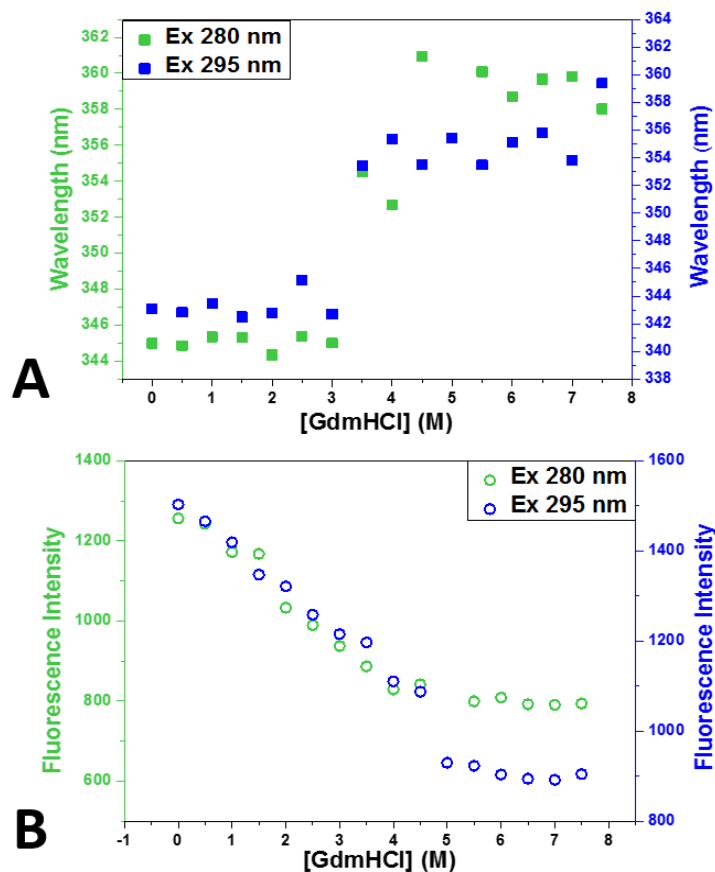


Figure 5.16: Change in lysozyme intrinsic fluorescence with increasing GdmHCl concentration. A) Wavelength shift. B) Intrinsic fluorescence intensity.

Both emission peaks indicate a significant red shift, at 3.5 M urea from 345 nm to a maximum of 361 nm at 4.5 M GdmHCl (Ex 280 nm) and from 343 nm to a maximum of 355 nm at 4 M GdmHCl (Ex 295 nm). No further structural alterations at higher GdmHCl concentrations were observed. The fluorescence intensity shows a steady decrease in fluorescence intensity that appears to level off by 4.5 or 5 M GdmHCl.

5.3.2.2 Second derivative fluorescence spectroscopy of lysozyme unfolding

The second derivative of lysozymes intrinsic fluorescence spectra are shown in Figure 5.17. Again, no additional spectral features were determined from the second order fluorescence spectra for either urea or GdmHCl induced unfolding when compared to the zero order fluorescence data (not shown).

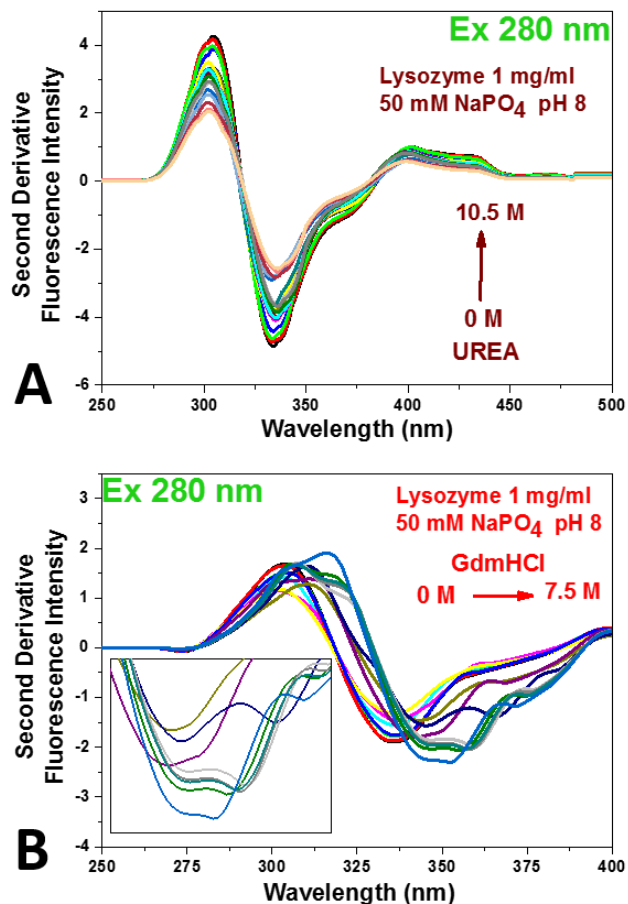


Figure 5.17: Second derivative intrinsic fluorescence spectra for lysozyme with increasing concentrations of denaturant (excitation at 280 nm). A) Urea. B) GdmHCl.

The unfolding of lysozyme with GdmHCl second derivative intrinsic fluorescence spectra did show a splitting of the Ex 280 nm peak into two minima at 4.5 M GdmHCl, indicative of a separation of the tyrosine and tryptophan residues as lysozyme unfolded, consistent with the zero order intrinsic fluorescence data.

5.3.2.3 Zero order UV absorbance during lysozyme unfolding

The second derivative UV spectrum for lysozyme was also investigated to see if it was possible to capture a two-state unfolding model with this technique. The zero order UV absorbance spectra for lysozyme were collected. The λ_{\max} for each zero order spectrum was plotted as a function of denaturant concentration in Figure 5.18. Native lysozyme (50 mM sodium phosphate buffer, pH 8) had a λ_{\max} at 281.5 nm. With increasing concentrations of denaturant only slight shifts (< 0.5 nm) in the λ_{\max} were recorded.

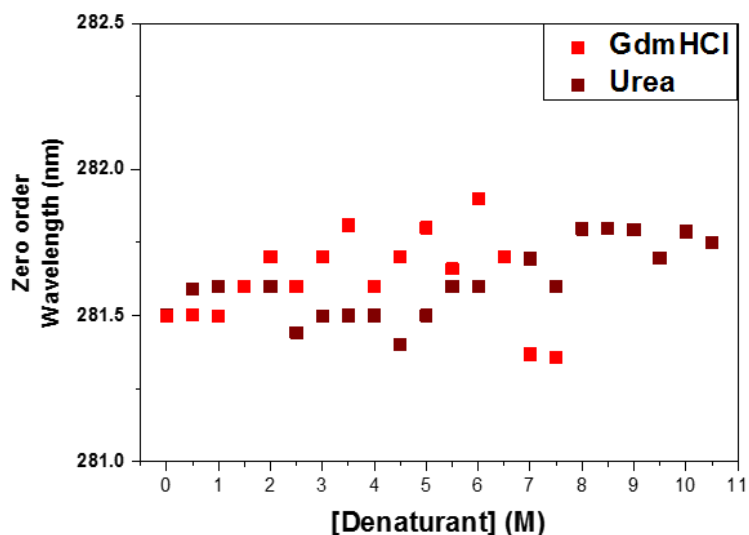


Figure 5.18: Shift in λ_{max} from the zero order absorbance spectra for lysozyme with increasing concentrations of urea and GdmHCl.

The unfolding transitions for lysozyme were not clearly distinguished from the zero order spectra when compared to the same data for BSA, where wavelength shifts of 2.3 nm were recorded.

5.3.2.4 Second derivative UV absorbance of lysozyme unfolding

The second derivative spectra for lysozyme with increasing concentrations of GdmHCl and urea were also analysed (figure 5.19). As for BSA, the single protein related absorbance peak of the zero order spectrum separates into the three individual aromatic residues in the derivative spectrum.

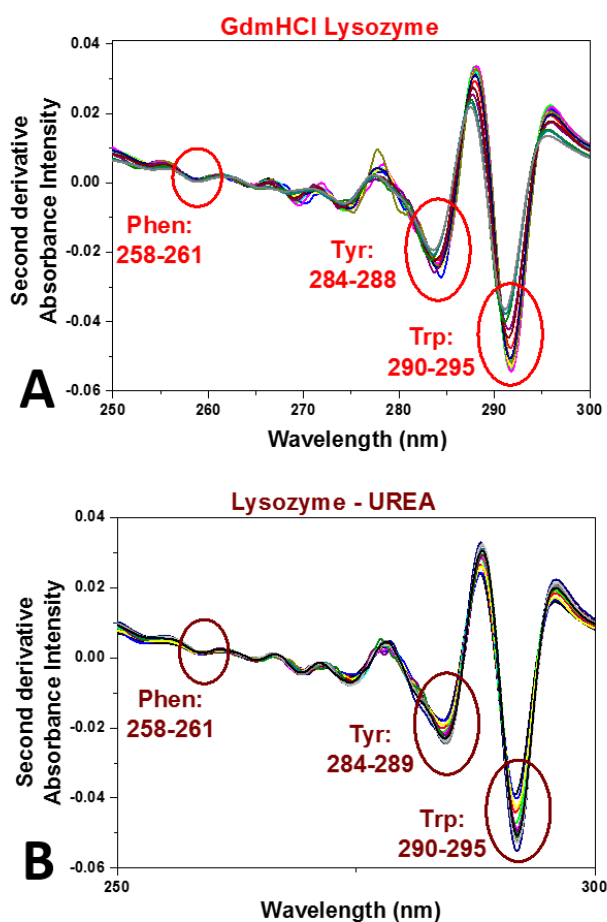


Figure 5.19: Second derivative absorbance spectra of lysozyme with increasing concentrations of denaturant. A) GdmHCl. B) Urea.

The shape and appearance of both the phenylalanine and tryptophan peaks appeared unperturbed by changes to their solvent environment in high denaturant concentrations; no significant in λ_{\min} (minimum of the peak) was observed. The tyrosine peak however did show a slight alteration at the peak minimum with increasing denaturant concentration (figure 5.20).

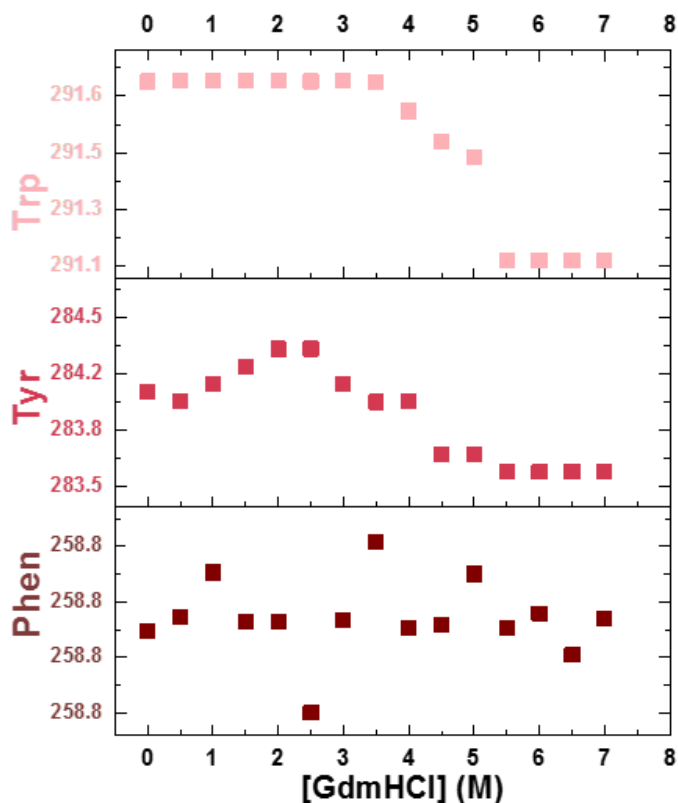


Figure 5.20: Effect of GdmHCl on the second derivative UV spectroscopic wavelength position of the three aromatic residues in lysozyme.

With increasing GdmHCl concentrations, changes in λ_{min} for both tyrosine and tryptophan were measured, but no significant trend was seen for phenylalanine. The λ_{min} for tyrosine and tryptophan reach a minimum value at 5.5 M GdmHCl (table 5.3), suggesting that lysozyme is fully unfolded at this GdmHCl concentration. However, the spectral shifts were not as defined as they were for the BSA unfolding curves. The onset of unfolding and fully unfolded states can be observed, any intermediate values are assumed to be related to the proportion of protein that have unfolded versus those that have not (consistent with the two-state unfolding model for lysozyme).

| <i>Amino acid</i> | <i>On set of unfolding</i> | <i>folded –v- unfolded</i> | <i>Fully unfolded protein</i> |
|-------------------|----------------------------|----------------------------|-------------------------------|
| <i>Trp</i> | 3.5 M | 4 – 5 M | 5.5 M |
| <i>Tyr</i> | 1 M | 1.5 – 5 M | 5.5 M |
| <i>Phen</i> | N/A | N/A | N/A |

Table 5.3: GdmHCl concentrations of unfolding events for lysozyme.

The urea induced unfolding of lysozyme was investigated by following the shifts in the wavelength positions for the three aromatic residues in the second derivative spectra (figure 5.21).

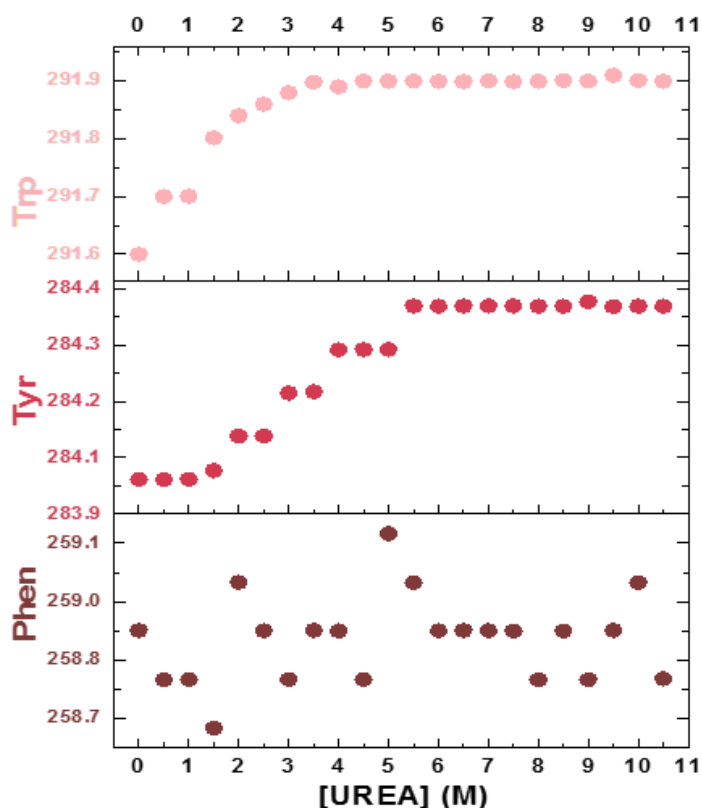


Figure 5.21: Effect of urea on the second derivative UV spectroscopic wavelength position of the three aromatic residues of lysozyme.

For urea induced unfolding of lysozyme, the phenylalanine residue in the second derivative again didn't give any further insight into the unfolding events. The λ_{\min} for tryptophan and tyrosine residues did change with increasing urea concentration (table 5.4). This data is consistent with two-state unfolding.

| <i>Amino acid</i> | <i>On set of unfolding</i> | <i>folded -v- unfolded</i> | <i>Fully unfolded protein</i> |
|-------------------|----------------------------|----------------------------|-------------------------------|
| <i>Trp</i> | 0.5 M | 1 – 4 M | 4.5 M |
| <i>Tyr</i> | 1.5 M | 2 – 5 M | 5.5 M |
| <i>Phen</i> | N/A | N/A | N/A |

Table 5.4: Urea concentrations of unfolding events for lysozyme.

5.4 Discussion

This study focused on the potential of second derivative UV absorbance spectroscopy to monitor protein unfolding for BSA and lysozyme using the denaturants, urea and GdmHCl. The three dimensional structure is known for these proteins, thus making it possible to correlate the changes in the environment around aromatic residues with alterations in their spectral properties.

5.4.1 Unfolding of BSA

Our initial study using intrinsic fluorescence analysis found that changes in the λ_{\max} and total fluorescence intensity occur as BSA is denatured. The unfolding is consistent with a two-step unfolding process via three transition states. The concentrations at which each unfolding event occurs for both urea and GdmHCl was previously documented and is consistent with our measurements (Sulkowska, Bojko *et al.* 2003; Itri, Caetano *et al.* 2004; Sulkowska, Rownicka *et al.* 2004; Sulkowska, Rownicka *et al.* 2005; Togashi, Ryder *et al.* 2010).

The zero order absorbance analysis during unfolding for BSA does provide similar data to the fluorescence measurements, which agrees with the proposed transitions for BSA unfolding; a two-step unfolding process via three transition states (Tayyab, Sharma *et al.* 2000; Sulkowska, Bojko *et al.* 2003; Das, Chitra *et al.* 2004; Adel, Nadia *et al.* 2008). But from the zero order UV absorbance analysis it is not clear which domain/s are involved in each unfolding step. It was however found that the second derivative of the UV absorbance spectra were more useful indicators of the unfolding states of BSA as the three individual aromatic residues could be monitored independently. Since the structure of BSA is known, changes in each of the three aromatic residues could be interpreted and related to changes in protein conformation and structure as individual domains begin to unfold.

Within the second derivative UV absorbance spectra, the positions of the wavelength minima for the three individual peaks associated with the aromatic amino acid residues changed with increasing concentration of denaturant. It was also evident that each amino acid residue responded independently to the increasing concentrations of denaturant. Therefore by using second derivative UV spectra the average local environment of each

specific amino acid was monitored to give insight into the effect the denaturant had on each separate residue. When this data was plotted against denaturant concentration, distinct sequences of events relating to the unfolding of BSA domains were observed (figure 5.12 and 5.13).

In our study of urea induced unfolding of BSA, the on-set of unfolding was observed at different concentrations for each amino acid. It is known from the literature that a urea concentration as low as 2 M is sufficient to instigate BSA unfolding from its native state (Tayyab, Sharma *et al.* 2000; Adel, Nadia *et al.* 2008). This was also indicated by our zero order UV absorbance spectra and intrinsic fluorescence analysis. Most researchers believe that the hydrophobic tryptophan residue (Trp-212) is the most sensitive as a marker of structural rearrangement.

Our second derivative tryptophan data shows the earliest signs of structural alterations at 2.5 M urea, suggesting that neither domain I nor domain II undergo any significant structural changes until after 2.5 M urea. Therefore domain III containing no tryptophan residue was the first of the domains to unfold, followed by rearrangement of domains I and II (indicated as three distinct regions in figure 5.1). The buried Trp-212 is involved in the hydrophobic interaction between subdomains IIA and IIIA. Therefore the unfolding of domain III and the subsequent shift in the tryptophan residue likely reflects a rearrangement within domain II. This can be confirmed with the intrinsic fluorescence data, where between 2.5 and 3 M urea, a blue shift in wavelength position is observed, indicative of partial burial of residues into a less polar inner core due to domain rearrangement, or alteration in the packing structure of the domains. Collectively all three amino acids indicate a significant structural alteration from 3 – 7 M urea, signalling the movement of Trp-212 and other residues out of the non-polar inner core as a consequence of domain II being unfolded. Again this coincides with a red shift observed in the intrinsic fluorescence data at 4 M urea, clear evidence that Trp-212 is more exposed to a polar environment. Lastly the tryptophan signal indicates the final structural alteration occurs by 7 M urea with domain I unfolding, beyond which no further structural changes are detected indicating the point where all three domains of BSA are fully unfolded and the protein adopts an extended conformation. This correlates with the concentration at which fully unfolded BSA is indicated by both the intrinsic fluorescence and zero order UV absorbance data.

Comparable data for the unfolding of BSA with GdmHCl as a denaturant was observed by solely examining the three individual aromatic residues in the second derivative UV absorbance spectra. The tryptophan residue was first to indicate structural change in the wavelength peak position, indicating an unfolding of domain III first at very low GdmHCl concentrations, resulting in domain rearrangement that the sensitive Trp-212 residue was the first to detect as it moved to a less polar inner core of the hydrophobic cavity of domain II, also evidenced by a blue shift observed in the intrinsic fluorescence measurements. This was followed by a second unfolding event of domain II at GdmHCl concentration greater than 1.5 M whereas tyrosine and phenylalanine don't indicate domain II unfolding until a higher GdmHCl concentration (> 2 M). Again this rearrangement of domain II at this GdmHCl concentration is consistent with a red shift observed for both the intrinsic fluorescence wavelength position of tyrosine and tryptophan from 2 – 3 M GdmHCl. Domain I only begins to unfold after domain II by 4 M GdmHCl, leaving a fully unfolded BSA by 6 M GdmHCl, again consistent with both our intrinsic fluorescence and zero order absorbance studies.

The unfolding behaviour of BSA with both urea and GdmHCl could also be observed with the second derivative fluorescence signals. At the denaturant concentrations at which it is likely that all three domains are unfolded (7 M urea and 5.5 M GdmHCl), a splitting of the single fluorescence peak occurs. This behaviour was likely due to independent signals arising from tyrosine and tryptophan. With excitation at 280 nm the fluorescence signal originates from both residues. As the protein becomes unfolded, the distance between the residues increases and their individual fluorescence signals became more evident in the second derivative spectra. This highlights the advantage of using derivative spectroscopy over zero order spectra where it is more difficult to discriminate between the tyrosine and tryptophan emissions by excitation at 280 nm.

The order of BSA domain unfolding observed with our second derivative residue peak behaviour is consistent with numerous other studies (Khan, Agarwal *et al.* 1987; Tanaka, Nishizawa *et al.* 1997; Tayyab, Sharma *et al.* 2000; Ahmad, Ahmed *et al.* 2005; Togashi, Ryder *et al.* 2010). Our analysis is indicative of a two-step, three state unfolding event with the formation of partially unfolded intermediate states as sequential BSA domains unfold, thereby confirming the known sequence for BSA unfolding and for the first time,

showing that second derivative UV spectroscopy can distinguish this series of events. It was also found that the structural alterations detected by the tryptophan residue, specifically Try-212, alone are more representative of the unfolding of BSA domains, rather than the global unfolding of the protein as observed with the zero order UV absorbance spectra. The second derivative unfolding curves for both phenylalanine and tyrosine while also indicative of a two-step, three state transition unfolding mechanism, the denaturant concentration of each domain unfolding was slightly greater than for tryptophan. Since these residues are more abundant and spread throughout BSAs structure they are less sensitive to slight structural alterations compared to tryptophan. Modelling of these unfolding events in their correct sequence is an interesting, but a difficult problem hence the advantage of being able to monitor the tryptophan residue alone, that is known to be a sensitive probe for the monitoring of protein unfolding.

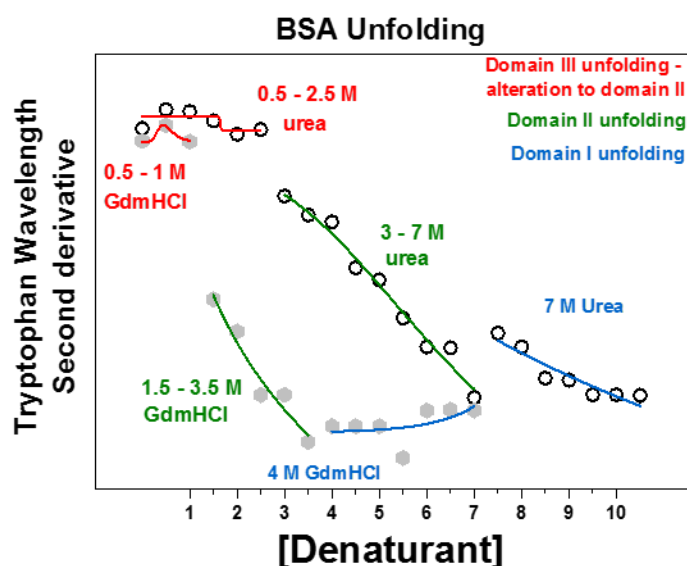


Figure 5.22: Schematic for the mechanism of urea induced unfolding of BSA as derived from second derivative wavelength position of the tryptophan residue

5.4.2 Unfolding of lysozyme

The unfolding curves of lysozyme in this study were derived from the structural analysis of denatured lysozyme with increasing concentrations of urea and GdmHCl. Comparison of the lysozyme data with that of BSA reveals that both proteins do not share a common mechanism during the process of unfolding. For BSA, unfolding is characterised by distinctive wavelength shifts for domain unfolding representative of a two-step unfolding process via three transition states. Unlike BSA, no obvious domain unfolding transitions

could be extracted from either the zero order or the second derivative UV absorbance spectra for unfolded lysozyme. The unfolding curves for lysozyme were all consistent with a two-state unfolding mechanism, thus confirming a two state model stated in the literature of lysozyme unfolding (Greene and Pace 1974; Sugai and Ikeguchi 1994; Chang and Li 2002; Sassi, Onori *et al.* 2011). The use of second derivative UV absorbance of lysozyme for creating unfolding curves was able to illustrate the concentrations where lysozyme begins to unfold and the point at which it is a fully unfolded structure. This was seen to be applicable to the tyrosine and tryptophan residues only. The signal arising from the phenylalanine residue did not appear to be sensitive to structural alterations nor did it indicate any changes within its solvent environment and was therefore difficult to discern clear trends in that data with respect to lysozyme unfolding.

As proposed for GdmHCl unfolding of lysozyme, all structural alterations have occurred by 6 M GdmHCl, the point at which no further alterations of the tryptophan or tyrosine residues are recorded in its second derivative wavelength position. The wavelength position of tryptophan remains steady up until 3.5 – 5.5 M GdmHCl where a large shift in the absorbance band was recorded. As proposed for lysozyme unfolding with GdmCl (Laurents and Baldwin 1997) the more surface exposed tryptophan active site residues (Trp-62 and 63) do not respond to lysozyme unfolding, instead it is the buried hydrophobic Trp residues (Trp- 123, 111 and 108) that indicate the unfolding of lysozyme. These three tryptophan residues do not show significant shifts during the unfolding process indicating they are tightly packed within their hydrophobic environment, and it's not until the global unfolding reaction occurs (> 3.5 M) that they become solvent exposed. This was also apparent in the intrinsic fluorescence of the same data, no blue shifts in the wavelength positions were recorded, but a large red shift was observed at 3 M GdmHCl, indicative of exposure of the hydrophobic residues to a more polar environment.

The second derivative unfolding curve of lysozyme with urea is not fully consistent with this opinion; while both the tyrosine and tryptophan residues propose a two-state unfolding model as seen in their unfolding curves (figure 5.20 and figure 5.21), they appear to respond during the initial stages of unfolding. This could be due to the interaction of urea with the surface exposed Trp residues. A steady shift in the tryptophan wavelengths is recorded between 0.5 – 4 M urea signifying the initiation unfolding as the

tight packing loosens and the hydrophobic residues gradually but simultaneously become more solvent exposed with the global unfolding event (fully unfolded conformation) occurring by 4 M urea. The fluorescence wavelength shift pattern appeared to fluctuate with an initial blue shift, followed by a red shift and another blue shift indicative of residue movement into and out of hydrophobic environments. Since lysozyme is neither fully folded nor unfolded within this concentration range it is possible the wavelength shift reflects differences in the populations of lysozyme structures.

The second derivative unfolding curves for urea induced unfolding, would suggest progressive expansion of the protein conformation toward its linear shape, while GdmHCl induced unfolding of lysozyme would suggest a tightly packed conformation with global unfolding event upon which hydrophobic residues become solvent exposed. Therefore GdmHCl studies capture the fully unfolded form of lysozyme while the urea studies appears to detect the process of unfolding for lysozyme. It is known that both urea and GdmHCl both induced unfolding of protein it different ways (Halim, Kadir *et al.* 2008; Lim, Rosgen *et al.* 2009; Hedoux, Krenzlin *et al.* 2010). Despite extensive studies, the exact mechanism by which urea or GdmHCl denature proteins is still divided between the “indirect” (Frank 1968; Bennion and Daggett 2003) and “direct” (Robinson and Jencks 1965; Wallqvist, Covell *et al.* 1998; Hua, Zhou *et al.* 2008) mechanisms. Based on the “direct” mechanism, it has been shown that both while denaturants directly bind to lysozyme, GdmHCl is believed to interact with the hydrophobic groups, with urea mainly interacting with the hydrophilic residues transforming the tertiary structure into a softer structure similar to a molten globule without unfolding of the secondary structure (Pike and Acharya 1994; Gao, She *et al.* 2010; Hedoux, Krenzlin *et al.* 2010). It’s possible that the second derivative data reflects this. It is also possible that because lysozyme is unfolding in a two-state manner (either unfolded or fully unfolded), the residues mainly reflect the most populated form, and for urea and GdmHCl induced unfolding of lysozyme this was different. This behaviour could be extracted out of the second derivative fluorescence spectra. For GdmHCl at the higher concentrations a splitting of the peaks occurs at the point which global unfolding of the protein is likely to have occurred. As the protein is fully unfolded, the individual tyrosine and tryptophan fluorescence signals became more evident in the second derivative spectra.

5.5 Conclusions

It has been demonstrated here that second derivative values extracted from zero order spectra are a meaningful parameters to characterise the unfolding states of proteins. We have shown how monitoring the second derivative UV absorbance spectra of a protein is a reliable method of indication of the conformational transitions of the protein as it is unfolded, establishing a connection between the three aromatic residues and the unfolding dynamics of a protein with respect to changes in their microenvironment. It has been shown to be highly sensitive to detecting minute changes in spectra characteristics that may go undetected in classical absorbance or fluorescence spectroscopy. It is also highly specific for resolving spectra with strongly overlapped emission/absorbance bands.

The second derivative spectrum serves as a useful indicator of the unfolding state of the three individual aromatic residues of BSA. The unfolding transitions occurred at different denaturant concentrations for each of the three aromatic residues. This made it possible to reconstruct which regions of the protein are more stable, and which are first disrupted during protein unfolding. The second derivative spectra were also capable of detecting the two-state unfolding events of lysozyme, with the tyrosine and tryptophan proving most valuable for monitoring the unfolding of lysozyme. It was found for both proteins specifically the hydrophobic tryptophan residues, are most sensitive for detecting unfolding transitions, and for BSA it allowed monitoring of specific domain unfolding. Second derivative UV absorbance spectra appears to be more sensitive to detecting changes to the solvent environment around the residues that undergo significant solvent changes *i.e.* movement of residues into and out of polar environments.

This study illustrates two important concepts. Firstly, it shows that measurement of the individual aromatic residues provides a means of following the transitions between the denatured and native states of proteins. Secondly, it shows how knowledge of protein structure coupled with a simple laboratory technique can give just as detailed information about the unfolding process of multi-domain proteins compared to more expensive and technically demanding techniques such as Raman spectroscopy, MS and AFM. The second derivative of UV absorbance spectra, which isolates the three aromatic residues allows excellent monitoring of protein unfolding-refolding process for BSA and lysozyme. The speed and efficiency with which we can collect, process and interpret

second derivative data from simple techniques like fluorescence and UV-absorbance using small sample volumes without expensive detection equipment highlights the advantage of this technique.

Chapter 6 Purification and Characterisation of IgG1

6.1 Introduction

The use of monoclonal antibodies (mAbs) as therapeutic agents is widespread in the biopharmaceutical industry. Humanised monoclonal antibodies are the fastest-growing category entering clinical study (Casi and Neri 2012; Scott, Allison *et al.* 2012). The global market for pharmaceuticals as of 2010 was \$597 billion, of which mAbs accounted for 7%. By this time 10 mAbs had acquired blockbuster status as they generated in excess of \$1 billion. The top five selling mAbs generated over €32 billion between them; bevacizumab (Avastin®) (16% market share), rituximab (Rituxan®) (16% market share), adalimumab (Humira®) (15% market share), infliximab (Remicade®) (15% market share), and trastuzumab (Herceptin®) (13% market share) (Elvin, Couston *et al.* 2013). As such, it is of vital importance that key features such as quality and safety is secured and approved to make them commercially viable and available for market. The most tedious challenge to this is ensuring antibody stability. This is mainly due to the choice of delivery, via subcutaneous injection (Jiskoot, Randolph *et al.* 2012; Lee, Perchiacca *et al.* 2013). Therapeutic formulations of mAbs have to be delivered in high concentration (> 100 mg) and in small volumes (< 2 ml) for intravenous administration. These highly concentrated formulations (50 mg/ml and higher) are susceptible to aggregation, that not only decreases the mAbs activity and efficiency, but may also be highly toxic and immunogenic to the patient (Harding, Stickler *et al.* 2010). It is widely accepted that injection of a foreign protein into humans can elicit an immune reaction (Rosenberg 2006; Zolls, Tantipolphan *et al.* 2012). Patient reactions can range from local skin irritation at the injection site, pyrexia and influenza-like symptoms, to acute anaphylaxis and systemic inflammatory response syndrome, which could be fatal (Hansel, Kropshofer *et al.* 2010; Jiskoot, Randolph *et al.* 2012). Therefore the development of these new types of therapies is highly dependent on understanding the molecular characteristics of the mAbs and what makes them unsafe. Despite great advancement in improving the safety of mAbs, adverse effects can still occur and steps to identify and minimize the risks are a necessity.

During the production and processing of mAbs designed as therapeutics aimed for market, concerns with the stability of the mAbs must be overcome. Problems such as aggregation due to physical instability, deamidation or oxidation due to chemical instability have to be addressed. One such proposed method is the engineering of

glycoforms expressed in the cell lines that produce the mAbs (Li, Vijayasankaran *et al.* 2010).

For therapeutic proteins it is important to understand both the biological function, typically defined by its amino acid component, and also the glycan component. A major proportion of biotherapeutic products are glycoproteins (Jefferis 2009). Control of glycosylation is of key concern during the development of therapeutic proteins because glycan chains play an important role in determining and maintaining the proteins conformation, activity and stability, in protein protection from proteolytic degradation, in intracellular trafficking and antigenicity (Bertozzi, Freeze *et al.* 2009; Roth and Khalaila 2012). Engineering glycoforms to increase the biological activity of mAbs allows optimisation of their roles and enhances the functional diversity of the proteins. Glycosylation achieves this as every aspect of glycosylation can be modified from the glycosidic linkage, the glycan composition, structure and shape (Teillaud 2005). For these reasons, the ability to determine the glycosylation pattern of antibodies is garnering increasing interest and importance as it is fundamentally linked to the activity, solubility, interactions, folding, stability, pharmacokinetics and immunogenicity of the antibody (Igawa, Tsunoda *et al.* 2011). Establishing connections between glycan structures and their functionality or monitoring glycosylation in disease diagnosis and prognosis are important research areas and give more control over protein quality (Li and d'Anjou 2009).

One of the key goals of recombinant glycoprotein production is to achieve consistent glycosylation. The cell lines used to produce IgG mAbs also play a significant role in the synthesis of the glycan chain. Therefore it is advantageous to make the cell lines able to produce recombinant IgG with a well-defined pattern of glycosylation in their Fc region, or with the ability to express excess levels of the glycoenzymes (glycosyltransferase, glycosylhydrolases) needed to perform glycosylation (Teillaud 2005). Glycans can be modified by sialylation, fucosylation, sulphatation, methylation or acetylation all of which can greatly affect their efficacy including their biological activity and clearance rate. Example include Chinese hamster ovary (CHO) cell lines that are engineered to express high levels of human β -1, 4-galactosyltransferase (GT) to reduce the rate of terminal GlcNac (Jones, Papac *et al.* 2007). Glycoproteins exposing glycans that end in Man, GlcNAc or Fuc residues are believed to be more rapidly removed from circulation. CHO cells were also modified to express ST6GalI, not naturally expressed in CHO cells,

to increase the α -2,6-linked sialic acid over α -2,3-linked sialic acid, a link that is known to increase sialylation of the oligosaccharides and shown to improve protein half-life (Bragonzi, Distefano *et al.* 2000).

No straightforward technique to sequence glycans has been developed but instead relies on the multi-step combination of electrophoresis to isolate the glycan linked protein, enzymatic cleavage to release the glycan, chromatography for separation and finally mass spectroscopy and chemical digestion to study individual glycans. Obtaining sufficient quantities of a given glycan for analysis, especially with structures that are in low abundance, is also a problem. Once a structure has been solved, duplicating that structure artificially is difficult as the synthesis of each structure has to be designed (Costa, Rodrigues *et al.* 2013).

As glycobiology aims to develop precise, robust and sensitive approaches for glycan production and analysis to meet the growing therapy demands, it needs to optimise and safeguard proper process control to ensure product efficacy and consistency. However, the problems faced include, heterogeneity of sugar structures that arise from the vast variety and diversity of glycan structures, sequences and linkages (Marino, Bones *et al.* 2010). A relatively simple set of sugars can form a huge number of complex structures and any defects in their structure can have disastrous consequences.

It must also be pointed out that while cell lines have been engineered to produce recombinant mAbs with a well-defined pattern of glycosylation in their Fc region, that the molecule being modified itself can show signs of unusual behaviour during production. Glycoforms can actually alter the structure and stability of the mAbs in a negative way (Zheng, Bantog *et al.* 2011). Focusing on the mAb IgG, numerous studies have looked at the solution solubility and stability of both glycosylated and aglycosylated mAbs to try find correlation between mAb glycosylation and aggregation (Arnold, Wormald *et al.* 2007; Wang, Antonsen *et al.* 2008; Kayser, Chennamsetty *et al.* 2011; Zheng, Bantog *et al.* 2011). The general consensus from the many reviews on the glycosylation of therapeutic proteins is that there are pros and cons to both methods. Sola and Griebenow have reviewed the effects that glycosylation has on the therapeutic efficacy of protein drugs and looks at both the instabilities of glycosylation on protein such as irreversible conformational changes and unfolding that can result in aggregation, but also highlights the effective long term stability glycosylation can bring by increasing

internal non-covalent forces and rigidifying the protein structure (Sola and Griebenow 2009). While recently Jung *et al.* (Jung, Kang *et al.* 2011) have highlighted the advantages and benefits of producing aglycosylated mAbs including simplified bioprocessing development, shortened running times and high yields in bacteria, overcoming issues associated with glycan heterogeneity. They also noted that aglycosylated mAbs can display novel therapeutic effector functions and the probability of them having a higher degree of conformational mobility than glycosylated mAbs.

6.2 Aims of the study

The objective of this chapter is to isolate, purify and characterise a monoclonal antibody, IgG1, from cell culture medium. Purified protein can be characterised by gel electrophoresis and SE-HPLC. Methods and procedures for both of these techniques need to be optimised to ensure accurate identification of the purified product. Two growth conditions were used in growing the cells producing the protein, Batch 1 (shake flasks) and Batch 2 (bio-reactor). It was important to analyse the different effects that these two culture conditions had on the production of the protein of interest. Initial stability studies on the purified protein were also conducted to assess the effect storage conditions had on the protein.

Changes in the growing conditions of the cell lines producing the protein can result in degradation of the protein and thus result in changes to the glycosylation pattern. Monitoring the glycoform can ensure product efficacy and consistency. There is enormous diversity in the type of glycan structures that can attach to IgG. Variation or change to the glycoform attached to a protein can lead to numerous alterations in the protein structure that in turn can affect the protein's stability & ability to function (Zheng, Bantog *et al.* 2011). The glycoform attached to the IgG1 was isolated and its fingerprint identified. Stability studies on the short and long term storage conditions were conducted to assess changes in the isolated glycoform.

While there is still some debate about the optimal glycoform required for stability, there is a general consensus that glycosylation does indeed lead to enhanced molecular stability and has therapeutic advantages for the pharmaceutical industry (Sola and Griebenow 2010). However, reliable techniques to fingerprint glycoforms for a particular mAb or

therapeutic protein are required and in-fact, this information is often not available during formulation, in particular. Hence, there remains a need for a comprehensive perspective on the subject and for increased understanding of this aspect of protein post translation modification to aid in the design of successful protein-based therapeutics.

6.3 Results

6.3.1 Purification of monoclonal antibody.

The monoclonal antibody IgG1 was expressed in CHO (DP12v2) cell line and grown in either shake flasks or a 2.5 L bioreactor by Ian Marison's group at DCU. The cell culture medium was kindly gifted to us. Purification of IgG1 from the cell medium was performed by affinity-purification using an immobilized Protein A column. Figure 6.1 shows the binding and elution of the protein to and from the column. Fractions are collected within the time indicated and pooled.

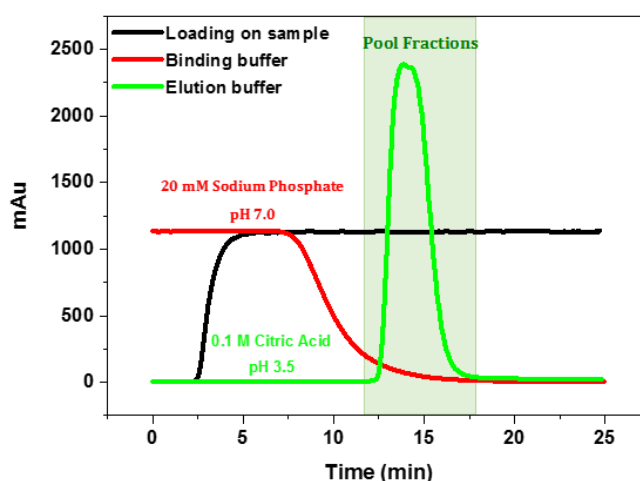


Figure 6.1: Elution profile of IgG1 after passage through a Protein A column. The black line represents loading the cell culture medium onto the column, the red line is the binding buffer while the green line is the elution of protein.

To ensure that only the desired protein, IgG1, was purified in its native monomeric form, the pooled fractions were subjected to SE-HPLC and gel electrophoresis as indicators of the protein purity obtained.

6.3.2 SE-HPLC

Depending on the protein class, the isoelectric point of IgG1 can range from 6.3 to 8.1 (Abraham, Podell *et al.* 1979). As the protein was eluted from the column, fractions are collected into tubes containing 1 M Tris-HCl, pH 9. This is to ensure that the final pH of the protein solution is neutral, since the elution buffer is an acidic buffer, 0.1 M citric acid, pH 4. The effects of leaving purified protein overnight in this neutral Tris-HCl buffer can be seen with SE-HPLC (Figure 6.2). The chromatogram shows several peaks corresponding to a range of molecular weights, the majority of which are lower in molecular weight than protein monomer. In fact, a protein monomer peak is not present. There appears to be excessive degradation of the protein.

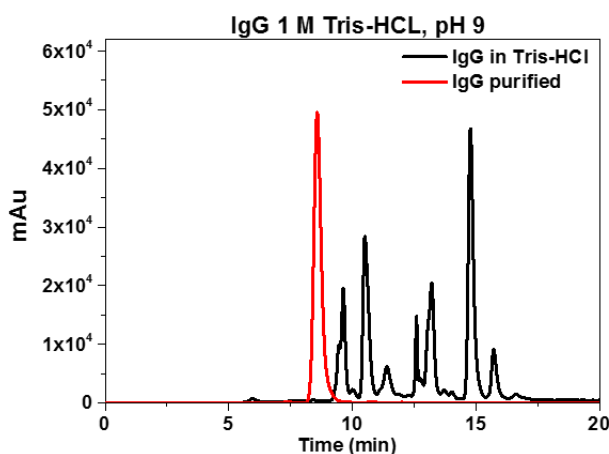


Figure 6.2: SE-HPLC of purified sample IgG1 left in 1 M Tris-HCl overnight compared to purified IgG1.

For this reason the pooled fractions had their buffer exchanged into 50 mM sodium phosphate, pH 8 using Amicon ultra 0.5ml 100K membrane centrifugal filter units immediately after collection. A chromatogram of IgG1 in this buffer condition is shown in Figure 6.3. The IgG1 in sodium phosphate shows a major peak at 8.58 minutes. From the calibration of the TSK 300 SWxl column used to separate the protein, a peak at this retention time corresponds to a molecular weight of 150 kDa, which is the expected molecular weight for an IgG1.

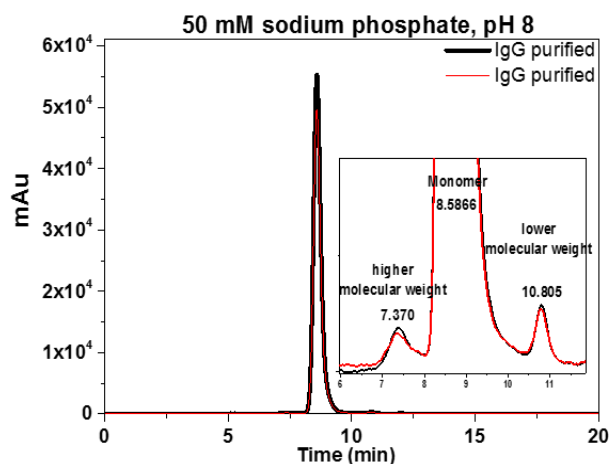


Figure 6.3: SE-HPLC of different purified samples of IgG1 in 50 mM sodium phosphate, pH 8, showing a single monomeric protein peak along with a small portion of higher and lower molecular weight species (less than 0.7%).

The chromatogram also shows a small fraction of higher and lower molecular weight species, less than 0.7% of the total protein in the sample. These peaks were attributed to a dimer species of the IgG with a molecular weight of 300 kDa and the presence of a lower molecular weight species of 50 kDa, possibly the heavy chain of IgG1.

6.3.3 Gel electrophoresis

Gel electrophoresis was used as a second method to confirm the purity of the IgG1 collected. To achieve this two gel types were utilized; Native and SDS-PAGE. Native gel electrophoresis allows for detection of proteins in their native form, in our case IgG1 as an intact 150 kDa monomeric protein, visualised as a single protein band at a molecular weight of 150 kDa, while reducing SDS-PAGE would allow us to break the IgG1 into its characteristic heavy and light chains via reduction of the disulphide linkages between light and heavy chains, another method of identifying if the correct protein had been purified.

6.3.3.1 Identifying intact monomeric IgG1

The preparation of IgG1 samples for gel electrophoresis had to be optimised. The progression of the protocol optimisation is shown in Figure 6.4. In our initial experiments the gels were prepared from 12% polyacrylamide solutions. This was reduced to 7% to prevent protein retention in the stacking gel. Concurrent with this there was aggregation of the protein upon mixing with the sample buffer. Originally, sample buffers were prepared with 0.5 M Tris-HCl, pH 6.8. While more protein was resolved with the 7%

SDS-PAGE, it resulted in gels with smeared bands. Different samples buffers were prepared with a range of pHs. Upon mixing with IgG1, a colour change from the normal blue, to yellow/green solution could be observed, a possible indication of aggregation of the sample. It was found that at pH 8, matching that of the protein buffer, no visual aggregation could be observed and the sample buffer retained its blue colour. It seems likely that pH 6.8 is close to the isoelectric point of the IgG1 and that resulted in aggregation. However on the 7% SDS-PAGE even with sample buffer at pH 8, the IgG still showed band smearing and protein aggregates trapped in the wells of the gel. SDS-PAGE gels were used with a non-reducing sample buffer (no 2-mercaptoethanol) to isolate the intact monomeric protein. However, it was also likely that the SDS present in the sample buffer and running buffer was possibly affecting the IgG1 stability. It was apparent that SDS-PAGE could not be used to isolate the intact monomeric IgG1. Instead, native gels were used, free from SDS to ensure that denaturation or aggregation did not occur. Using this method a single strong monomeric protein band at the correct molecular weight was observed.

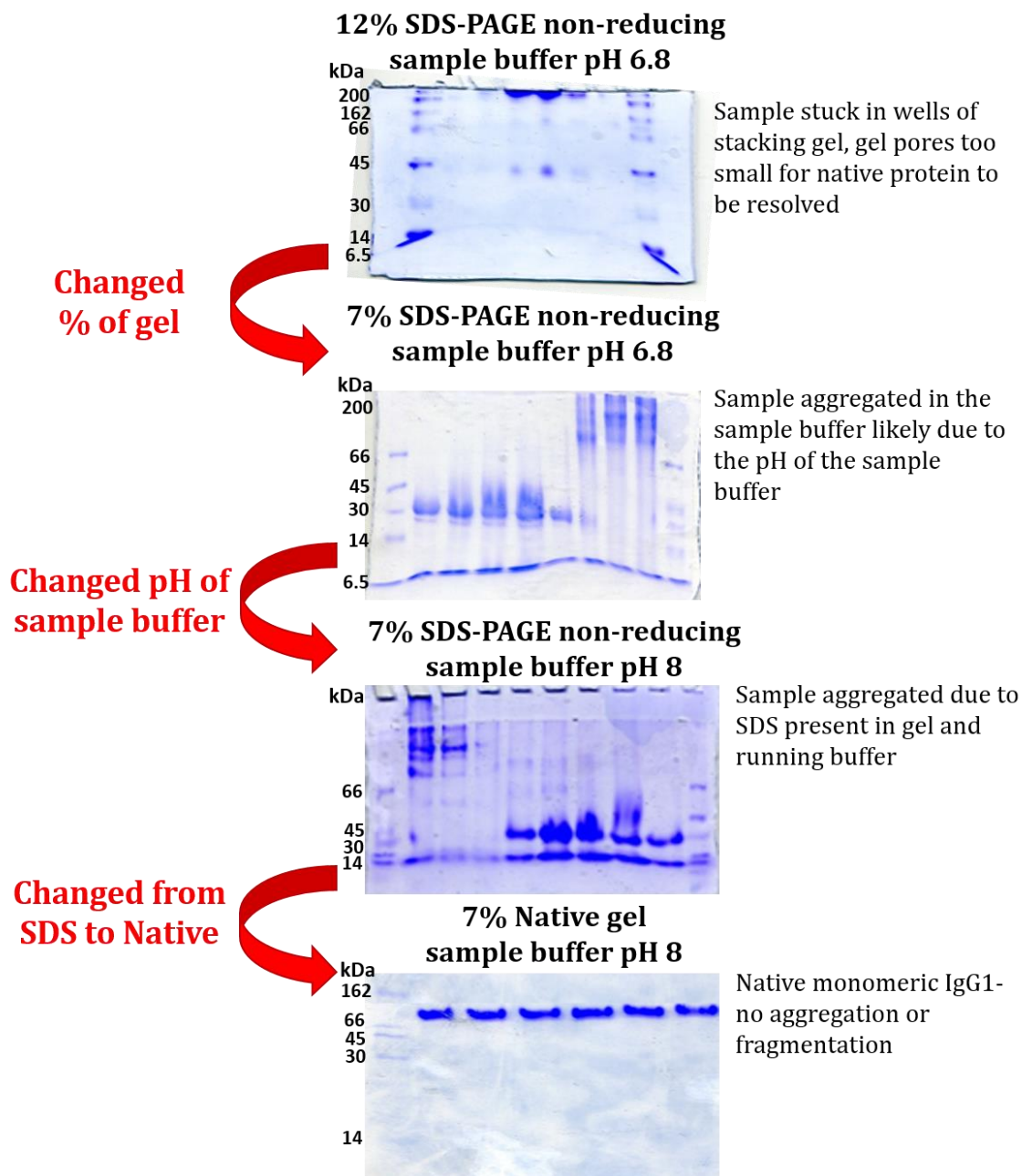


Figure 6.4: Isolation of intact IgG1 using gel electrophoresis. The evolution of gel preparation to obtain a clear monomeric band of IgG1 using native gels without any signs of aggregation or degradation to confirm the purity of the desired protein.

6.3.3.2 Identifying heavy and light chains of IgG1

IgG is made up of two heavy chains (50 kDa each) and two light chains (25 kDa each) that are held together by covalent cysteine bonds. Therefore SDS-PAGE with reducing conditions was used to separate and identify the heavy and light chains at the correct molecular weights, as another means to confirm the purification of IgG1 (figure 6.5). The pores in a 12% polyacrylamide gel were of sufficient size to resolve protein within the

molecular weight range of the heavy and light chains. The pH of the sample buffer was adjusted from the standard protocol of pH 6.8 to pH 8 to match that of the protein sample buffer to avoid aggregation of the IgG1. The sample buffer also contained 2-mercaptoethanol to break the cysteine bond between the two chain types.

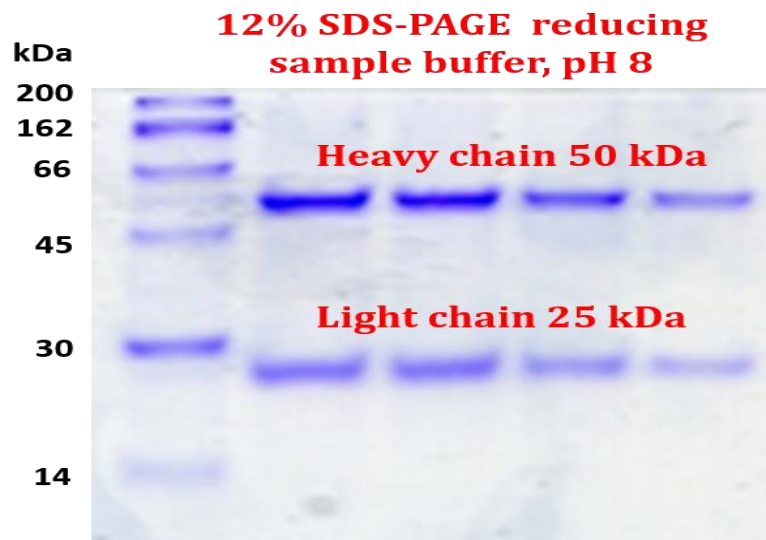


Figure 6.5: Separation of the heavy chain and light chain of IgG1 using gel electrophoresis. 12% SDS-PAGE gel using reducing conditions (2-mercaptoethanol) to break the cysteine bonds that hold the IgG1 intact. Two clear separate bands can be seen representing the heavy chain at 50 kDa and the light chain at 25 kDa.

6.3.4 Batch to batch variation

6.3.4.1 Protein yield

The electrophoresis confirmed that for both batches, IgG1 was isolated, but there were clear differences in the quality of product purified from each batch. This was analysed in greater detail. Table 6.1 illustrates the concentration of IgG retrieved after three separate purifications from each batch of IgG1 after buffer exchange in 50 mM sodium phosphate, pH8. Each batch had an initial volume of 50 ml for Batch 1 and 30 ml for batch 2. Batch 1 had a much lower yield compared to Batch 2.

| <i>Batch #</i> | <i>Purified IgG1 (mg/ml)</i> | <i>Volume (ml)</i> | <i>Total yield (mg)</i> |
|----------------|----------------------------------|------------------------|-----------------------------|
| <i>1</i> | 2.82 | 8.5 | 23.97 |
| <i>1</i> | 2.59 | 8 | 20.72 |
| <i>1</i> | 1.25 | 8.5 | 10.63 |
| <i>2</i> | 5.9 | 8.5 | 47.2 |
| <i>2</i> | 5.47 | 8.5 | 43.82 |
| <i>2</i> | 4.43 | 8.5 | 37.69 |

Table 6.1: Concentration analysis of purified IgG

6.3.4.2 Gel Electrophoresis

The variation between the batches was detected by gel electrophoresis. Gel electrophoresis on Batch 1 (flasks) showed that the purified IgG1 had aggregated and was unable to be resolved on the gel, even in the optimised conditions (Figure 6.6). It was likely that the protein had aggregated prior to purification as the same large band can be seen in the well of the gel in both of the native gels, before and after purification. Bands at the molecular weights consistent with heavy and light chains can also be observed in the unpurified sample indicating possible degradation of the native protein prior to purification. The protein was confirmed however as IgG1 through identification of the heavy and light chains on the reducing gel. This also confirms that the type of aggregate formed was by covalent attachment as the aggregates could be broken up by 2-mercaptoethanol.

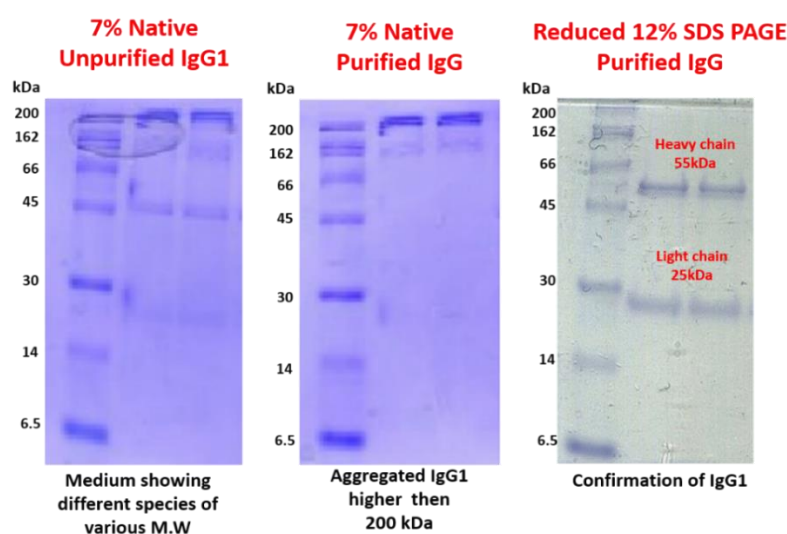


Figure 6.6: Batch 1 of IgG1 characterised by gel electrophoresis. The 7% native purified gel shows a faint band at 150 kDa the molecular weight of IgG1 but the majority of the material is aggregated in the wells of the gel. The reduced gel confirms IgG1 was purified.

Gel electrophoresis of Batch 2 (bioreactor) showed a strong, single intact band at 150 kDa in the native purified gel indicating purification of monomeric IgG1. This was also confirmed using a reducing SDS-PAGE showing the two characteristic bands, one for each of the heavy and light chains (Figure 6.7).

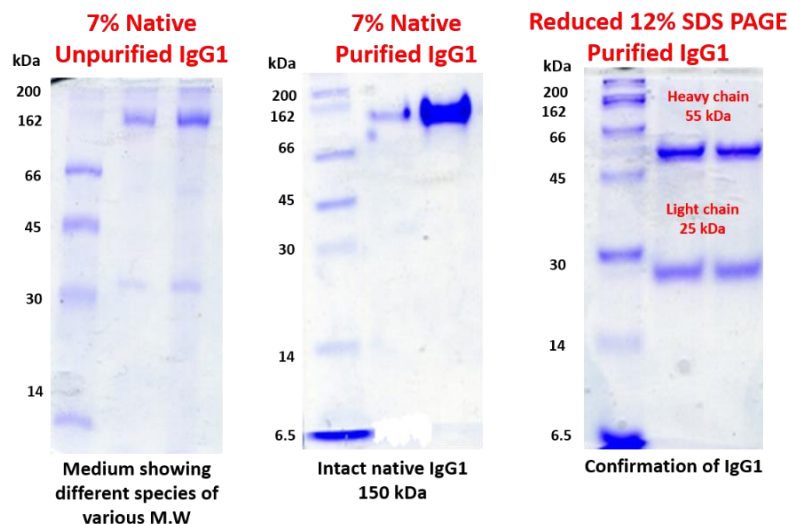


Figure 6.7: Batch 2 of IgG1 characterised by gel electrophoresis. The 7% native purified gel shows a strong band at 150 kDa the molecular weight of IgG1. The reduced gel confirms IgG1 was purified.

6.3.4.3 SE-HPLC

SE-HPLC also detected variation between the two protein batches (Figure 6.8). A 1 mg/ml solution of IgG1 from each batch was compared by SE-HPLC (injection volume = 40 μ l). Batch 1 had a much lower monomer yield compared to batch 2. Batch 2 had a 99% yield of monomeric IgG1 with only 1% of dimer and higher molecular weight species. In Batch 1 a high proportion of lower molecular weight species, nearly 88% of the total sample detected.

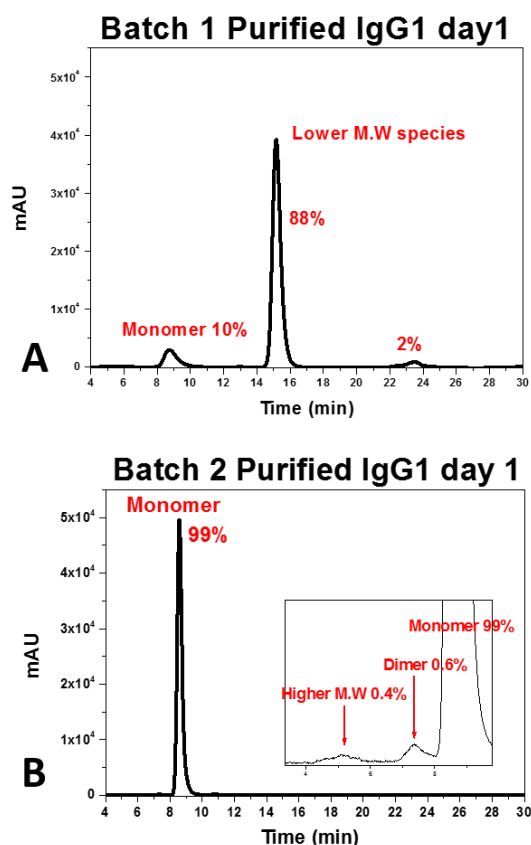


Figure 6.8: SE-HPLC analysis of Batch 1 (A) and Batch 2 (B) of purified IgG1. The amount of monomer protein between each batch varied greatly.

There was clearly significant differences in the protein material taken from shake flasks (Batch 1) versus the protein material grown in the bio-reactor (Batch 2). Further discussion with our colleagues at DCU revealed that cells were grown longer than is optimal during production in shake flasks (batch 1) and some cell death occurred. We speculate that this resulted in enzymatic damage to the IgG1 in the culture medium as a result of this cell death. Therefore, for our study, IgG1 from Batch 2 was used for further analysis. All of the electrophoresis experiments indicate a protein with a molecular weight consistent with IgG1 (150 kDa) and reducing conditions show the characteristic heavy and light chain bands. SE-HPLC indicates that the protein is very pure.

6.3.5 Stability of purified IgG1

The monomeric IgG1 from Batch 2 was used to evaluate the stability of the purified protein, in 50 mM sodium phosphate buffer at pH 8. The protein was monitored using SE-HPLC over a two week period. Two different storage conditions were examined. One sample set was stored at 24°C, and the other at 4°C. Figure 6.9 shows the chromatograms

of the two sample sets along with the positions of the three major peaks identified in the chromatogram in Figure 6.8.

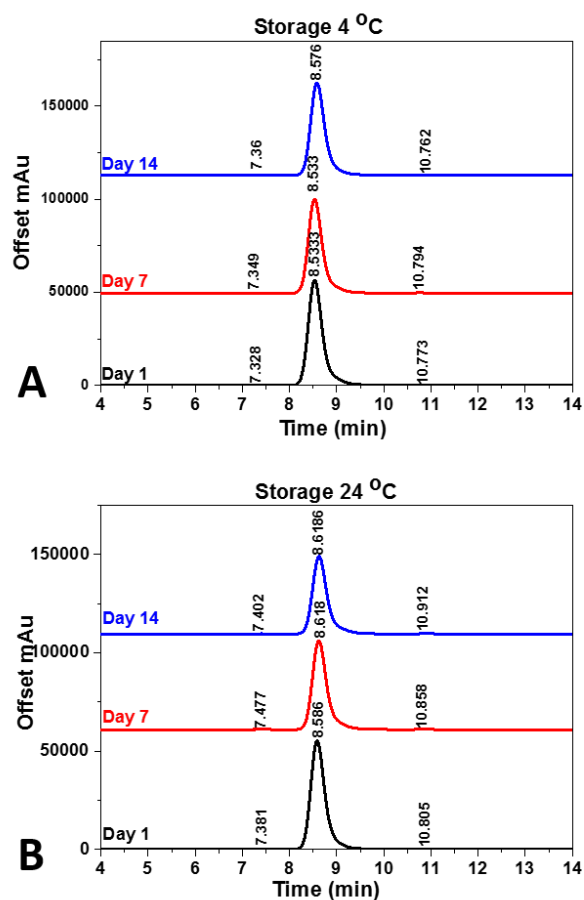


Figure 6.9: SE-HPLC of IgG1 in 50 mM sodium phosphate, pH 8. A) IgG1 stored at 4°C and B) at 24°C at day 1 just after purification, washed into 50 mM sodium phosphate, pH 8, on day 7 and on day 14.

There is no significant migration in peak positions, change in peak shape or the appearance of new peaks that could indicate aggregation or degradation (within the column resolution) of the protein due to a decrease in protein stability. However, there does appear to be a decrease in the area under the peak with increasing time. The area under each peak was monitored over time to see if there was any loss in protein that might indicate the formation of larger aggregates that would not have been observed on the SE-HPLC column (Figure 6.10).

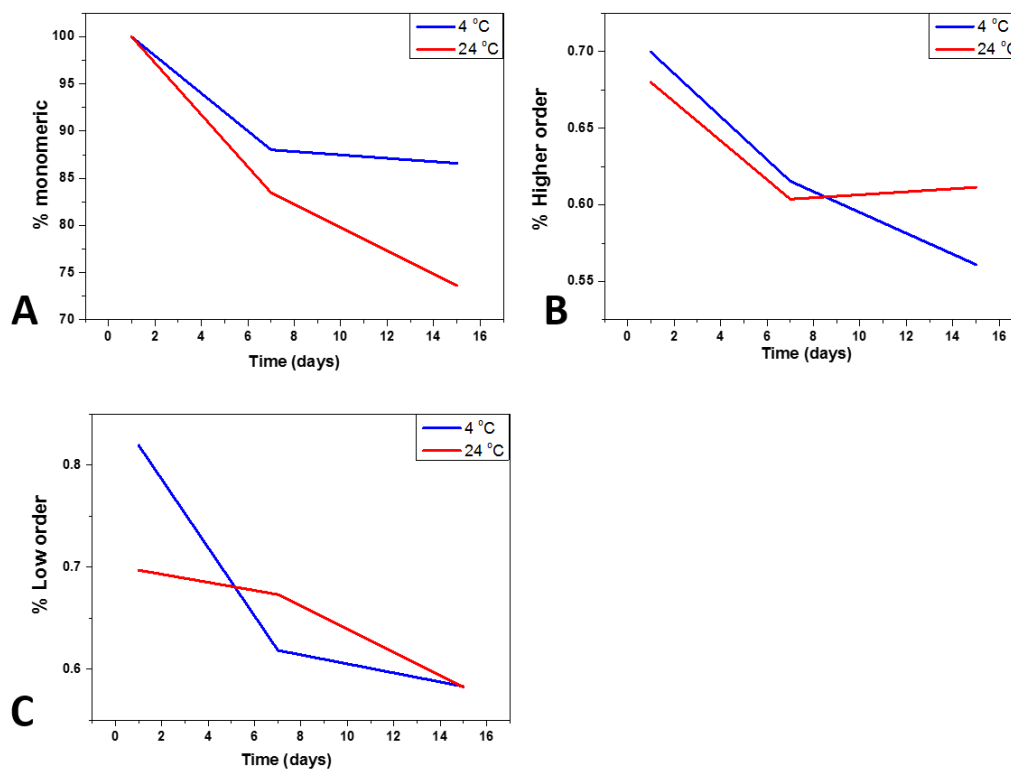


Figure 6.10: SE-HPLC calculated % area under the three main peaks of IgG stored at 4°C and 24°C over 14 days. A) Shows the change in monomer. B) Higher order species. C) Lower order species.

From Figure 6.11 we see that a significant decrease in protein monomer concentration and corresponding decrease in the already low proportions of the other species in solution. Between 12 - 18% of the monomer is lost by day 7, and this increases to nearly 30% by day 14. This loss is likely attributed to the formation of higher molecular weight aggregates, too large to be resolved on the SE-HPLC column. The column resolution is only up to 500 kDa, too low to resolve protein much larger than a trimer of IgG1. The way in which the samples were stored will also play a role. The IgG1 was transferred to a HPLC vial on day 1. If large aggregates form (by Mechanism 1, 2 or 4, and in the absence of external stress, mechanism 1 is most likely), they will sediment in the vial and will therefore, not be injected onto the column during subsequent runs. It is also probable that some portion of the material became irreversibly absorbed on the glass surface (by Mechanism 5), which would account for some portion of the loss of protein.

6.3.6 Identifying the glycoform of IgG1

The glycosylation pattern for this IgG1 was determined and the stability of the sugars against degradation was assessed for samples stored in different conditions and over a range of time scales. Four sets of sample conditions were examined:

1. IgG1 at time zero (Frozen -80°C directly after purification)
2. IgG1 stored at 24°C over 7 days
3. IgG1 stored at 4°C over 7 days
4. Year old sample of IgG1 stored at 4°C

Prior to enzymatic cleavage to release the glycan, the glycosylated heavy chain (50 kDa) of the purified IgG1 was isolated on a SDS reducing gel (Figure 6.11). Gel pieces were chopped up and incubated with the enzyme to allow sufficient release of the *N*-linked glycan.

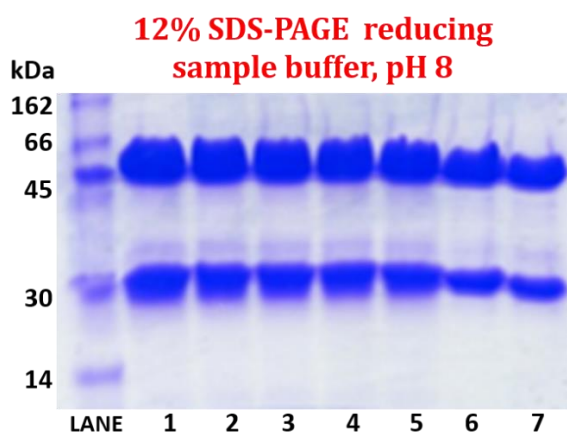


Figure 6.11: SDS-PAGE reducing gel of IgG1 for glycosylation analysis. Lane 1- IgG year old, lane 2-3 IgG stored at 4°C over 7 days, lane 4-5 IgG stored at 24°C over 7 days, lane 6-7 IgG at time zero.

N-glycans were then analysed using ultra-high pressure liquid chromatography (UPLC). This was chosen over HPLC as it offers increased separation efficiency, and has significantly reduced separation times and higher peak capacity (Tharmalingam, Adamczyk *et al.* 2013).

6.3.6.1 Undigested samples

Four IgG samples from a single batch (Batch 2) of purified protein had their undigested glycoform analysed. All four conditions were analysed in duplicate. Using UPLC and standardised against a dextran ladder. Once the elution profile for each sample was

obtained, the glucose unit (GU) value was determined for each identifiable peak (Figure 6.12). This involved comparing the elution of the dextran ladder (with each successive peak in the ladder adding a GU) to the profile of our sample. The retention times of the glycan peaks were compared to the retention times of the standard dextran ladder to assign a glucose unit. The assigned value could be compared to a database of known glycan structures to tentatively assign glycan structures, GlycoBase 3.0+.

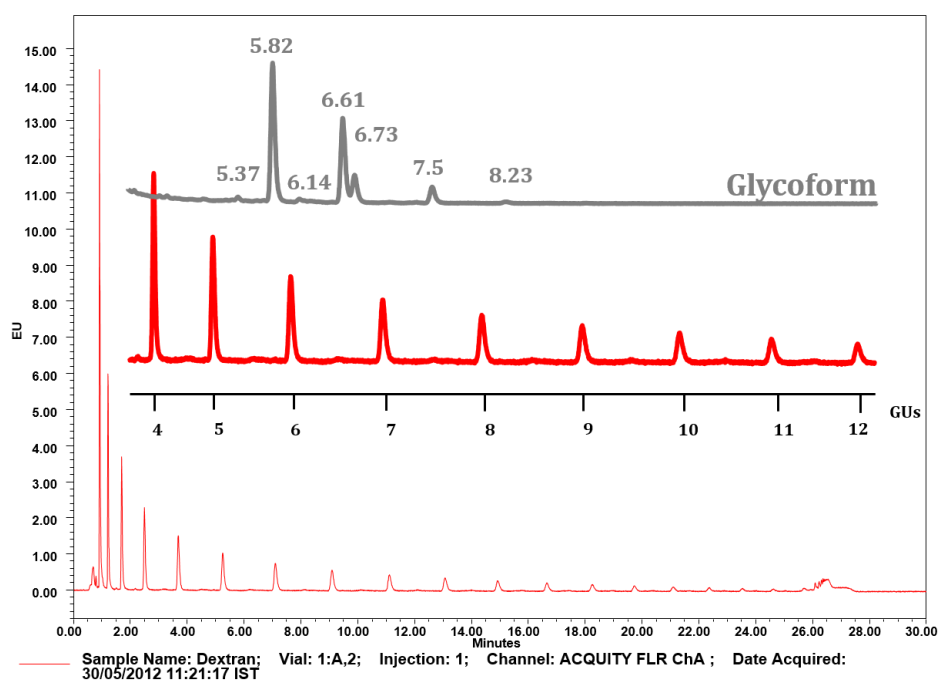


Figure 6.12: Dextran ladder standard profile obtained from separation using UPLC. Each peak represents the retention of a polysaccharide of glucose, where the first peak represents 1 glucose, the second 2 glucose units etc. This is then used to assign GU values to the glycoform of our IgG1.

Once GU values had been assigned, the % area under each peak was calculated for both samples under the same storage condition to ensure consistency in the glyco profile. Figure 6.13 shows the UPLC glycosylation profile of the different IgG1 sample sets. Each peak represents a type of oligosaccharide moiety present in the glycoform. The areas under each peak were integrated to show the relative amounts of a glycan type compared to that of the overall glycoform. Therefore by monitoring the % area of each individual peak, we can relate a % loss to degradation of that glycan in the glycoform.

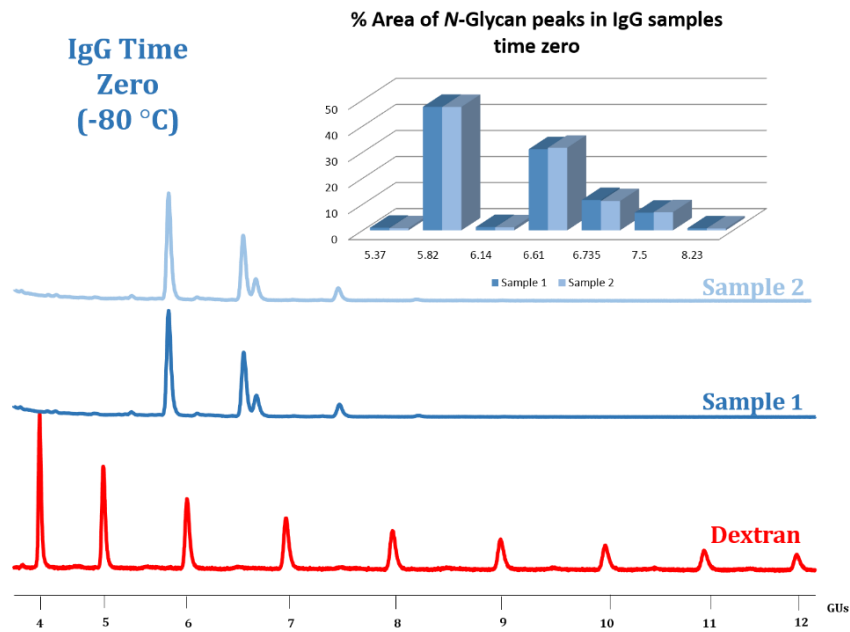


Figure 6.13: Elution profile for IgG time zero (-80°C) along with bar chart comparing the % area under each peak for duplicate samples.

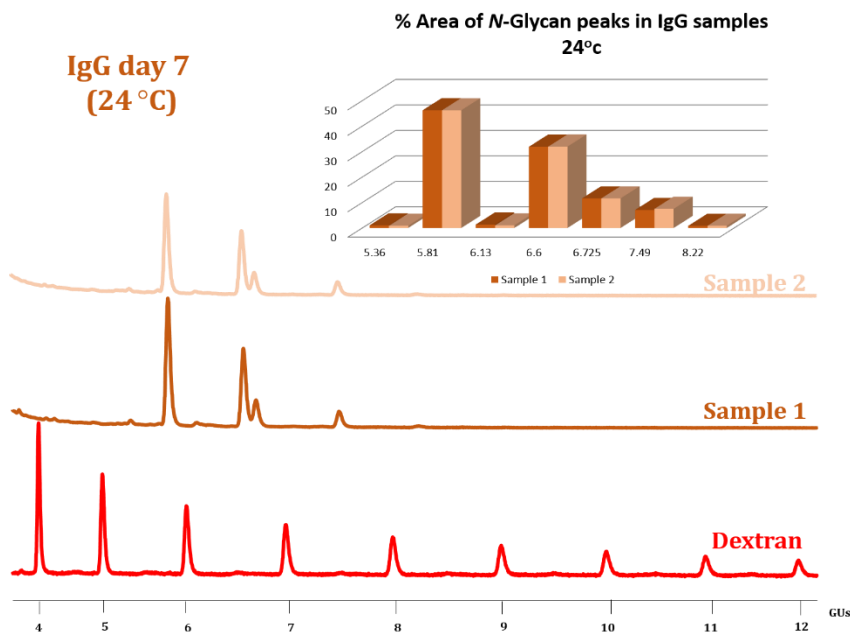


Figure 6.14: Elution profile for IgG day 7 (24°C) along with bar chart comparing the % area under each peak for duplicate samples.

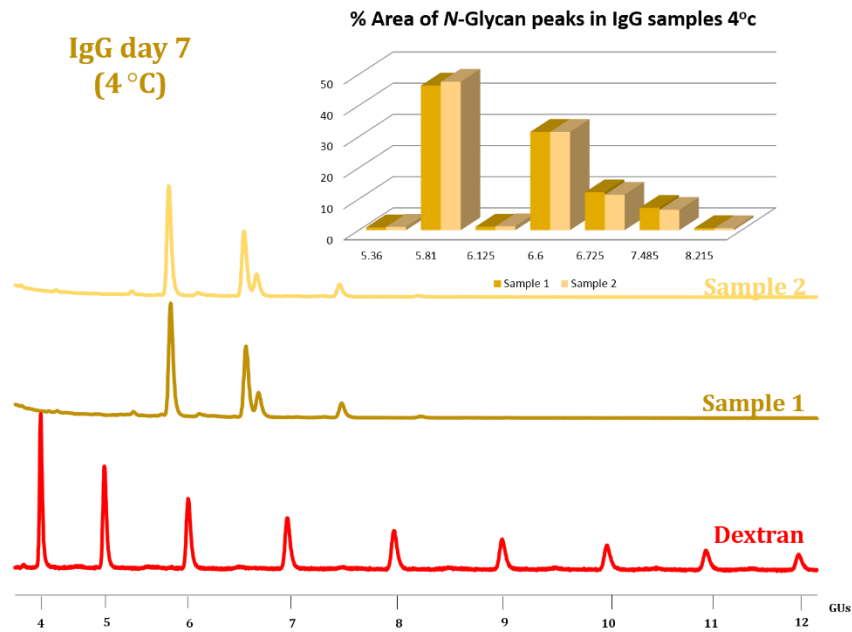


Figure 6.15: Elution profile for IgG day 7 (4°C) along with bar chart comparing the % area under each peak for duplicate samples.

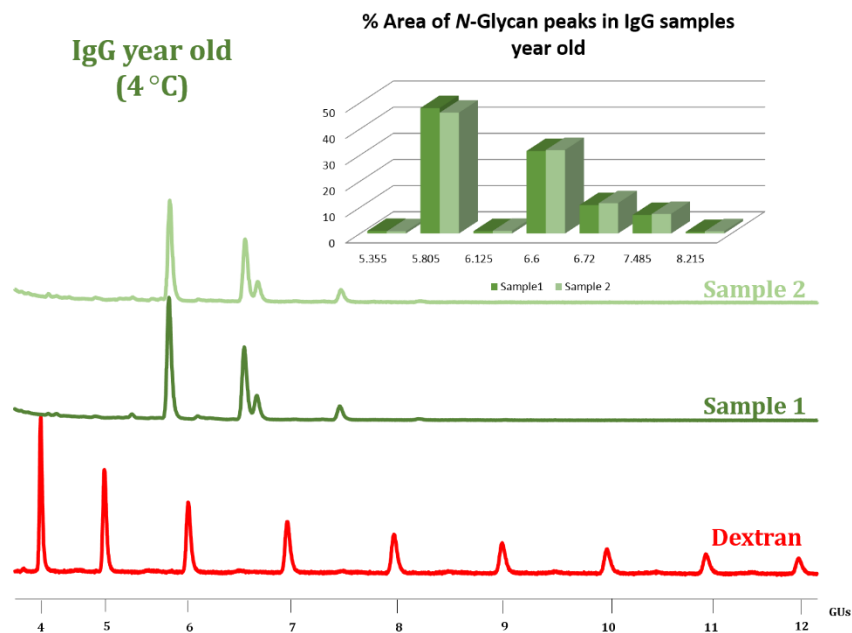


Figure 6.16: Elution profile for IgG year old (4°C) along with bar chart comparing the % area under each peak for duplicate samples.

It's important to ensure that similar glycan profiles for two identical samples are obtained. In each case there was little variation in GU values or peak % area between each sample per storage condition, thus indicating good release of the sugars from the protein and consistent profiling of the glycoform.

The stability of the glycoforms was examined by comparing the individual glyco profiles from the four storage conditions to each other (Figure 6.17). For each sample set, the glyco-profiles and the average % area under each peak were compared. This was to assess if the different storage conditions had an effect on the glycan profile.

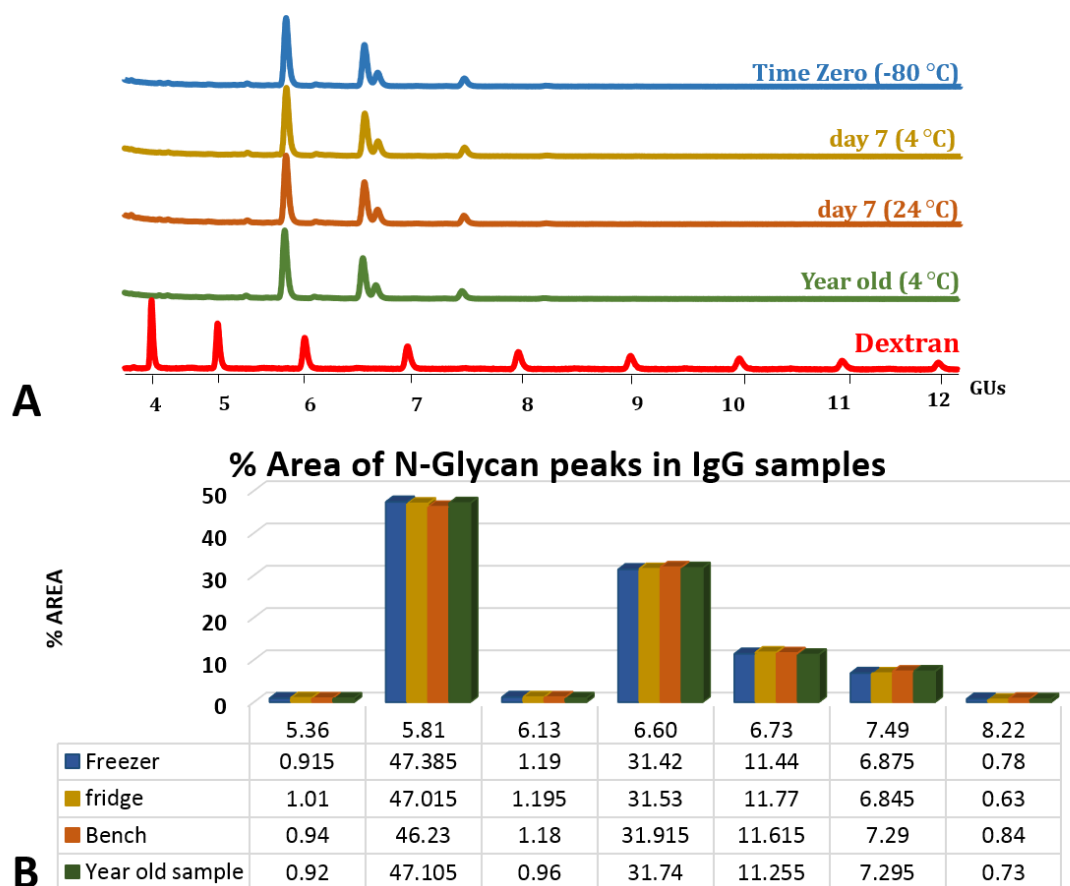


Figure 6.17: Comparison of the glycan profile from four different storage conditions. A) Elution profile for the glycoform distribution of the four storage conditions of IgG. B) Bar chart indicating the % area under each peak for the four different storage conditions.

From this data it appears that each of the four sample sets have considerable overlap and uniformity between their glyco profiles and in their retention behaviour, indicating that all samples still have the same glycoform attached to the Fc portion of the IgG1. The loss in % area between each sample is also minimal, signifying that the glycoform stored under the four different conditions remained completely stable with no degradation. This indicates that the storage conditions used did not impact on the glycoform structure. From this work it was established that the glycosylation of this particular IgG1 is very stable at different storage temperatures and over prolonged periods of time.

6.3.6.2 Digested samples

In order to fully understand the stability of the IgG protein, it was important to decipher its glyco form distribution. This detailed structural analysis involves the separation of a glycan pool to provide a signature profile that gives an overview of the individual glycan moieties present. Intact glycans released from purified IgG1, were broken down into a specific oligosaccharide units by enzymatic cleavage and analysed by NP-HPLC.

After release of the glycoform from the protein, prior to examination by chromatography, the samples underwent treatment with enzymes known as exoglycosidases. This step yields information about both the sugar sequence and linkages between the sugar units. Specific enzymes are chosen to systematically cleave off oligosaccharide units with defined linkages. After the initial analysis of the undigested samples, an array of different enzyme combinations were chosen to progressively break down the individual elements in the glyco pool. These enzymes do not cleave internal bonds between carbohydrates but remove terminal carbohydrates from the non-reducing end of a glycan. Using this techniques, the removal of glycan residues can be identified by linkage as well as sugar.

The 2AB labelled *N*-glycans were treated sequentially with exoglycosidase enzymes. The exoglycosidases and their role used in this digest include (Marino, Bones *et al.* 2010);

1. *Arthrobacter ureafaciens* sialidase (ABS) – releases $\alpha(2-6/3/8)$ –linked nonreducing terminal *N*- acetylneuraminic acid (NANA, Neu5Ac) and *N*-glycolylneuraminic acids (NANA, Neu5Gc)
2. Bovine kidney alpha-fucosidase (BFK) – releases $\alpha(1-2/6)$ –fucose linked nonreducing terminal fucose residues more efficiently than $\alpha(1-3/4)$ –linked fucose.
3. Bovine testes β -galactosidase (BTG) – hydrolyzes nonreducing terminal galactose with $\beta(1-3/4)$ linkages.
4. B-*N*-acetylglucosaminidase cloned from *S. pneumonia*, expressed in *Escherichia coli* – digests $\beta(1-4)$ –linked GlcNAc to mannose but not a bisecting GlcNAc $\beta(1-4)$ linked to mannose.
5. Jack Bean $\alpha(1-2,3,6)$ -Mannosidase (JBM) – releases non-reducing terminal $\alpha(1-2,6)$ -linked mannose residues more efficiently than $\alpha(1-3)$.

Data analysis of glyco profiles can be complex and difficult due to differences between the various platforms of analysis. Interpretation of results is a challenging task, so reliance on a comprehensive and accessible database is a necessity. For this reason this data was interpreted using a bioinformatic tool, GlycoBase 3.0+ to establish the glycan composition for the IgG1. This database is continually being developed at NIBRT and branched off from EurocarbDB. It is a novel HPLC/UPLC resource that contains elution positions for more than 650 2AB-labeled *N*-linked and *O*-linked glycan structures determined by a combination of HPLC, UPLC, exoglycosidase sequencing and mass spectrometry (Tharmalingam, Adamczyk *et al.* 2013).

Three different pieces of information allow the glycan structure to be predicted

1. Change in peak area – as parts of the glycan are cleaved off, the % area of the remaining peaks in the profile increases, while the overall number of peaks decreases. The % area in the UND sample peaks in Figure 6.18 is spread over 7 peaks totalling the whole glycan profile. Peaks disappear with successive digests indicating loss of that structure and the % area of the remaining peaks increases. The purple bar in Figure 6.18 shows that with that combination of exoglycosidases all of the glycan structures apart from one moiety have been cleaved off and so the single peak left accounts for 100% of the area resolved.
2. The shift in retention time – as parts of the glycan are cleaved off, you decrease the overall size of the glycan causing the retention time to shift slightly. Specific shifts in retention times in combination with the enzyme used, are characteristic of certain glycan structures *e.g.* cleavage of galactose can result in two changes: 6-linked Gal shows a peak shift to lower retention time compared to cleavage of a 3-linked Gal.
3. The GU values – Glycobase3.0+ has GU values assigned to known structures.

% Area of N-Glycan peaks in Digested IgG1 samples

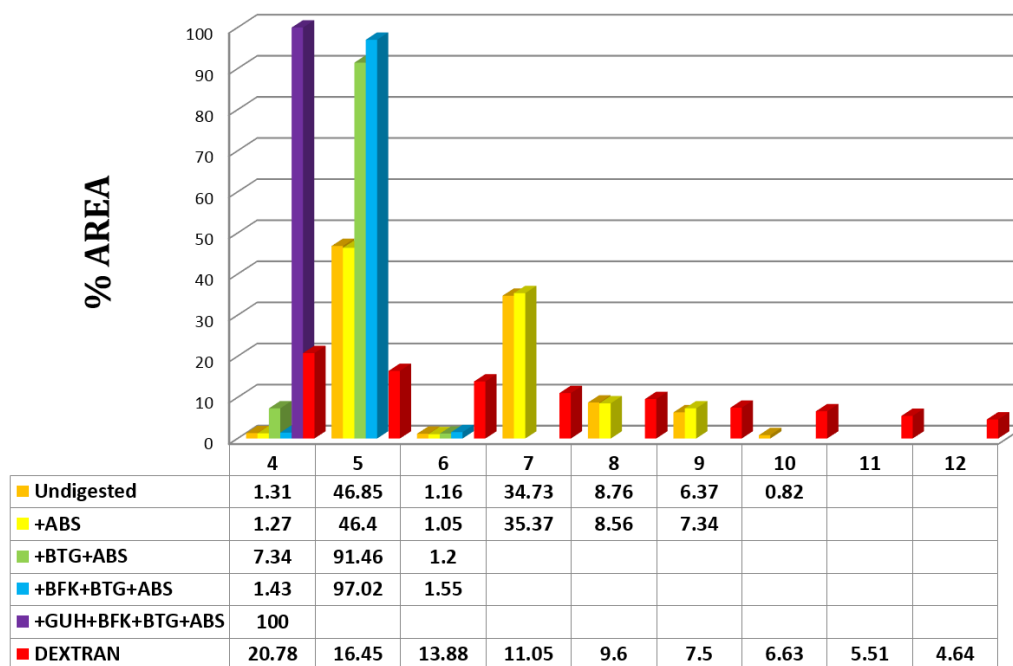


Figure 6.18: Exoglycosidase digest of the 2-AB labelled N- glycan of IgG1. % area under each peak of the glycan profiles after successive addition of enzymes.

The combination of the specificity of the exoglycosidase used and a knowledge of the incremental value of each moiety facilitates the analysis of the glycan sequence and linkage assignment. With these three factors in conjunction with the database it was possible to predict the glycan structures present in our IgG1 glyco profile. Figure 6.19 shows the NP-HPLC glycosylation profile of IgG1 in its undigested form (UND) along with the profiles of the glycans successively treated with different exoglycosidases. The confirmation of the structure using exoglycosidase digestions is important to confirm the heterogeneity of the glycans profiles obtained.

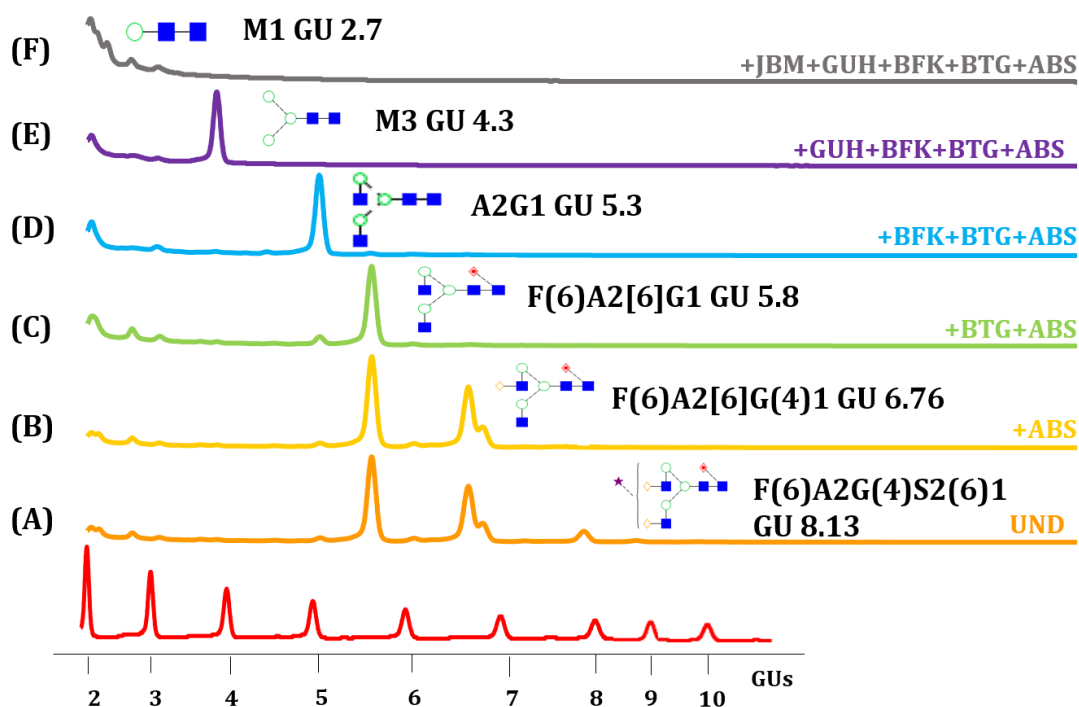


Figure 6.19: NP-HPLC Glycan profiles of 2AB labelled *N*-glycans from IgG1 (A) with (B) *Arthrobacter ureafaciens* sialidase (ABS) (C) *Bovine kidney alpha*-fucosidase (BFK) (D) *Bovine testes beta*-galactosidase (BTG) (E) *B-N*-acetylglucosaminidase (GUH) (F) *Jack Bean alpha*(1-2,3,6)-Mannosidase (JBM) along with their predicted glycan structure.

The major peak was assigned as the structure F(6)A2G(4)S2(6)1 at 8.13 GU, with all other structures shown in Table 6.2. With the addition of the ABS the peak was reduced to a major peak at 6.76 GU, which confirmed the presence sialic acids [S2(6)]. This peak was further reduced to just 5.8 GU with the addition of BTG, which confirmed the presence of terminal galactose [G(4)] on the glycans. BFK releases the terminal fucose [F(6)] to give a peak at 5.3 GU, while the GUH will digest BGlcNac [A2G1] to give a peak at 4.3 GU. Finally profile (F) shows the peak at 2.7 GU as JBK releases the mannose [M3-M1] to leave just the N-Acetyl-Glucosamine, GlcNac –GlcNac- mannose moiety.


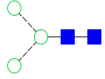
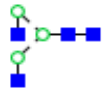
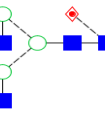
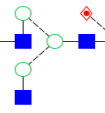
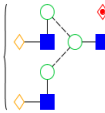
| <i>Nomenclature</i> | <i>Glucose units</i> | <i>Proposed glycan</i> | <i>GlycoBase I.D</i> |
|-------------------------|----------------------|---|----------------------|
| <i>M1</i> | 2.7 |  | 530 |
| <i>M3</i> | 4.3 |  | 597 |
| <i>A2G1</i> | 5.3 |  | 759 |
| <i>F(6)A2(6)G1</i> | 5.8 |  | 735 |
| <i>F(6)A2(6)G(4)1</i> | 6.76 |  | 634 |
| <i>F(6)A2G(4)S2(6)1</i> | 8.13 |  | 683 |

Table 6.2: Exoglycosidase digestions of IgG1. The Nomenclature abbreviations include M (mannose), A (antennae), G (galactose), F (fucose), and S (sialic acid). The structures contain sugars including GlcNAc (\square), Man (\circ), Gal (\diamond), Fuc (\diamond) and NeuAc (\star).

6.4 Discussion

Non-optimal conditions during protein expression, purification, handling and storage can severely alter the protein conformation, resulting in aggregation and loss of activity (Bondos and Bicknell 2003). Therefore optimisation of the purification process is an extremely important step in any protein purification work to guarantee isolation of the correct protein without adverse effects on its structure. Controlling the environment and purification strategies can minimise the extent of aggregation (Cromwell, Hilario *et al.* 2006). This was evident in our batch to batch variation analysis. We could see the effect that growth conditions had on the quality and stability of protein product. Batch 1 grown in shake flasks showed evidence of protein aggregation and degradation even before purification as evidenced by gel electrophoresis. It also gave much lower purified protein yields compared to Batch 2. It was difficult to get consistent data from this material. Upon consultation with Prof. Ian Morrison's group it appeared there were difficulties with growing the cells that produced the cell medium for Batch 1. This would account for the observed batch to batch variation.

Another step to control aggregation of protein involves optimising the purification process (Roberts and Wang 2010; Mach and Arvinte 2011). This also aids in optimising additional sample preparation steps such as buffer exchange and filtering, necessary to get protein samples ready for downstream analytical methods of characterisation. To achieve this, buffer conditions were controlled so that optimum purity and stability of IgG1 could be achieved. It was found that washing the protein into 50mM sodium phosphate, pH 8, buffer the protein could maintain its monomeric form, compared to being left in the purification protocol buffer, Tris-HCl. Excessive damage to the protein was detected by SE-HPLC when the protein was left in this buffer.

The buffer conditions play a significant role in the shelf-life of therapeutic proteins. Buffers are used to maintain the pH of the protein formulation and thereby help stabilise the protein. One of the first steps in the formulation process is to determine the pH and ionic strength where the protein is natively folded, and minimal amount of degradation/aggregation events will occur (Daugherty and Mersny 2006). Numerous studies have detailed both improved stability and aggregation of monoclonal antibodies under various storage buffers, temperatures and length scales (Sarciaux, Mansour *et al.*

1999; Arosio, Jaquet *et al.* 2012; Majumdar, Manikwar *et al.* 2013). Initial stability tests were performed assessing the long term storage of IgG1 in the buffer condition (50 mM sodium phosphate, pH 8), analysed over a two week period, at two different storage temperatures, 4 °C and 24 °C. It was observed that there was a significant loss of monomeric protein, presumed to be higher order aggregates (~30%) through either self-association, or surface induced aggregation. This was not explored further. Continuation or additional stability optimisation experiments could not be carried out due to a shortage of sample. It was not possible to obtain additional cell culture medium to continue the stability tests it was hoped could be performed.

There is enormous variability in the glycosylation site of IgG in that they can contain different patterns of glycans which could affect their biological function along with their stability (Raju, Briggs *et al.* 2000; Bertozzi, Freeze *et al.* 2009). In order to fully understand the stability of the IgG1 protein purified, it was important to identify its glycoform. Glycosylation is an important step in the recombinant expression of mAbs and is therefore a significant variable to monitor, especially where the stability of the mAbs is being considered (Sola and Griebenow 2010; Costa, Rodrigues *et al.* 2013). Glycan heterogeneity can be altered by many cell culture environmental factors including nutrient starvation, metabolic waste accumulation, culture viability, pH and temperature (Andersen, Bridges *et al.* 2000; Yang and Butler 2000; Kontoravdi, Asprey *et al.* 2007). Thus there is a requirement to monitor variables in the attachment of a glycan in order to ensure consistent glycosylation and glycoprotein formation.

In this study it was confirmed that the IgG1 purified had a high degree of homogeneity of *N*-glycans within its glycan pool. The structure of the carbohydrate chain of the FC region of all the IgG1 samples was found to contain N-Acetyl-Glucosamine (GlcNAc), mannose (Man), galactose (Gal), sialic acid (Sia) and fucose (Fuc) with no variations occurring. This shows there was batch-to-batch consistency in the protein formulation and glycosylation of the IgG1, signifying that the culture conditions and the culture medium used to grow the transfected cells produced well defined glycan chains. This glycoform was seen to be stable over a prolonged time period (up to 1 year). Storage temperature also had no apparent effect on the glycan profile.

6.5 Conclusions and future work

This study has highlighted the importance of cell culture conditions in the formulation of IgG1. We have shown that IgG1 grown in shake flasks caused undesirable damage to the protein believed to be caused by non-optimal growth conditions. IgG1 grown from bioreactor conditions was purified with a high monomer content. The buffer and storage conditions also play a role on maintaining the stability of protein and conditions need to be optimised at the early stages of purification to prevent unwarranted damage to the protein.

One of the original aims of this work was to assess the structural stability of the purified IgG1. It was hoped to evaluate the behaviour of the protein when subjected to different stress mechanisms by completing different forced degradation studies. Having a detailed characterisation of the purified protein product was a key achievement towards this goal. Future work would include further characterisation of the physicochemical properties of the protein including sequencing of the primary structure by mass spectroscopy. A more detailed assessment of buffer conditions where optimal stability is achieved giving potential to determine the isoelectric point of the protein would prove highly useful.

Homogeneous glycosylation of the IgG1 purified was confirmed. The glycosylation of the IgG1 also appeared to be stable over an extended time period with different storage conditions. The results of this study will enhance our understanding of the IgG1 protein. While the stability of the glycoform has been confirmed, its impact on the stability and the activity of this IgG1 is yet to be tested. This analysis would prove highly useful when exploring the stability of the protein.

Summary and Concluding Remarks

The aggregation of proteins formed by a number of distinct mechanisms has been examined using a range of chemical and biophysical techniques. Care must be taken when interpreting the results of such studies. We have shown that protein aggregates, resulting from a single stress (thermal or oxidation) can range from nanometres to many microns in size and can even result in fragmentation of the protein (oxidation) with these fragments also aggregating. We have shown that the type and quantity of aggregation can vary greatly with just a single solution condition, for example, the thermal denaturation of BSA at pH 5 (large protein assemblies) versus pH 6 (small oligomeric aggregates). It was also shown that the same stress mechanism can result in protein aggregates with different morphologies (oxidised lysozyme at pH 5 (amorphous) versus lysozyme at pH 8 (fibrils)). This shows how two proteins in the same solution conditions and with the same stress mechanism can react very differently.

Here in lies the challenge of analysing and characterising protein aggregation. It is the variety in the types of aggregates formed in addition to the range in particle sizes means that no single technique can monitor the whole aggregation pathway. Also it appears that no specific mechanism produces a specific aggregate type. This is of significant importance given the intensive drive towards high-throughput screening within the biopharmaceutical industry to speed up the generation of lead candidates as drug targets or therapeutic proteins. With the aim of improving the time cycle for introducing targets and reducing production costs, are they compromising at the cost of accuracy, efficiency and safety? As we have shown with the assessment of two high-throughput screening assays, while useful as a first pass screening platform, it also has major drawbacks. A more detailed assessment of the data obtained using high throughput methods can be misleading and not representative of the true aggregation levels. While the use of orthogonal high throughput techniques will be more discriminating than the limited number of assays used in this study, it is unlikely that in-depth sample-by-sample analysis will become obsolete, since it is still the best way to fully characterise protein aggregation. The use of second derivative UV absorbance was demonstrated to be a reliable tool for the characterisation of unfolding states of protein as a result of chemical denaturation. The second derivative unfolding curves for lysozyme with two denaturants, urea and

GdmHCl correctly indicated a two-state unfolding mechanism. For BSA, which is known to unfold via intermediate states, it was possible to identify the concentrations of denaturants at which specific domain unfolding occurred by monitoring the second derivative absorbance of tryptophan. The data agrees with previous work on the unfolding of BSA measured with a variety of techniques. This was a surprising level of sophistication for such a simple technique. Second derivative UV absorbance could be further explored by looking at other proteins with known unfolding mechanisms to further establish its usefulness and sensitivity at monitoring unfolding transitions.

The final objective of this work was to be able to compare the model proteins (lysozyme and BSA) to a more biologically relevant therapeutic glycoprotein, IgG. Due to the structural complexity of IgG, it is a challenge to preserve its biological activity and stability throughout upstream processing (formulation and purification) along with downstream processing (storage and administration). A more thorough understanding of the aggregation mechanisms and pathways for therapeutic proteins encountered during their development is needed to enhance their formulation. Initial studies assessing the stability of the IgG1 protein were performed, however no further optimisation experiments could be carried out due to a shortage of sample. However, significant progress was made towards this aim. Protocols were developed for the purification and characterisation of the protein and a glycosylation fingerprint was measured for the protein. Therefore a platform from which a more detailed analysis of the stability of this IgG1 against aggregation can be performed in the future.

References

- Aligent Technologies. "Fundamentals of modern UV-visible spectroscopy " Retrieved 3rd April 2014, 2014, from http://www.chem.agilent.com/Library/primers/Public/59801397_020660.pdf.
- Thermo Scientific. "Chemical Compatibility Guide." Retrieved 25th July, 2014, from <http://sevierlab.vet.cornell.edu/resources/Chemical-Resistance-Chart-Detail.pdf>.
- JDR Enterprises Incorporated. "Chemical Resistance Chart." Retrieved 25th June, 2014, from <http://www.j-drain.com/pdfs/crc.pdf>.
- Abraham, G. N., D. N. Podell, R. Wistar, S. L. Johnston and E. H. Welch** (1979). *Clinical and Experimental Immunology* **36**(1): 63-70.
- Adel, A., M. Nadia, O. Mohamed and G. Abdelhafidh** (2008). *Materials Science & Engineering C-Biomimetic and Supramolecular Systems* **28**(5-6): 594-600.
- Aghamohammadi, A., A. Farhoudi, M. Moin, Z. Pourpak, N. Rezaei, M. Nikzad, M. Movahedi, M. Gharagozlou, L. Atarod, A. Ahmadi Afshar, N. Bazargan, K. Abolmaali and M. Mahmoudi** (2003). *Iran J Allergy Asthma Immunol* **2**(3): 121-126.
- Ahmad, B., M. Z. Ahmed, S. K. Haq and R. H. Khan** (2005). *Biochimica Et Biophysica Acta* **1750**(1): 93-102.
- Amani, S. and A. Naeem** (2013). *Process Biochemistry* **48**(11): 1651-1664.
- Ami, D., A. Natalello, M. Lotti and S. M. Doglia** (2013). *Microb Cell Fact* **12**: 17.
- Amzel, L. M. and R. J. Poljak** (1979). *Annual Review of Biochemistry* **48**: 961-997.
- Anand, U., C. Jash and S. Mukherjee** (2011). *Physical Chemistry Chemical Physics* **13**(45): 20418-20426.
- Andersen, D. C., T. Bridges, M. Gawlitzek and C. Hoy** (2000). *Biotechnology and Bioengineering* **70**(1): 25-31.
- Anfinsen, C. B.** (1973). *Science* **181**(4096): 223-230.
- Annunziata, O., O. Ogun and G. B. Benedek** (2003). *Proceedings of the National Academy of Sciences of the United States of America* **100**(3): 970-974.
- Arnold, J. N., M. R. Wormald, R. B. Sim, P. M. Rudd and R. A. Dwek** (2007). *Annual Review of Immunology* **25**: 21-50.
- Arosio, P., B. Jaquet, H. Wu and M. Morbidelli** (2012). *Biophysical Chemistry* **168-169**: 19-27.
- Arvinte, T., C. Palais, E. Green-Trexler, S. Gregory, H. Mach, C. Narasimhan and M. Shameem** (2013). *MAbs* **5**(3): 491-500.
- Azakami, H., A. Mukai and A. Kato** (2005). *Journal of Agricultural and Food Chemistry* **53**(4): 1254-1257.
- Balestrieri, C., G. Colonna, A. Giovane, G. Irace and L. Servillo** (1978). *European Journal of Biochemistry* **90**(3): 433-440.
- Balme, S., R. Guegan, J. M. Janot, M. Jaber, M. Lepoitevin, P. Dejardin, X. Bourrat and M. Motelica-Heino** (2013). *Soft Matter* **9**(11): 3188-3196.
- Barnes, K. P., J. R. Warren and J. A. Gordon** (1972). *Journal of Biological Chemistry* **247**(6): 1708-1712.
- Batey, S., K. A. Scott and J. Clarke** (2006). *Biophysical Journal* **90**(6): 2120-2130.
- Beale, D. and A. Feinstein** (1976). *Quarterly Reviews of Biophysics* **9**(2): 135-&.
- Benedek, G. B.** (1997). *Investigative Ophthalmology & Visual Science* **38**(10): 1911-1921.
- Bennion, B. J. and V. Daggett** (2003). *Proceedings of the National Academy of Sciences of the United States of America* **100**(9): 5142-5147.
- Berg, M., K. Undisz, R. Thiericke, P. Zimmermann, T. Moore and C. Posten** (2001). *Journal of Biomolecular Screening* **6**(1): 47-56.
- Berlett, B. S. and E. R. Stadtman** (1997). *Journal of Biological Chemistry* **272**(33): 20313-20316.
- Bertozi, C. R., H. H. Freeze, A. Varki and J. D. Esko** (2009). *Glycans in Biotechnology and the Pharmaceutical Industry*. Cold Spring Harbor (NY).

Bertrand, M., P. Jackson and B. Walther (2000). *Eur J Pharm Sci* **11**(2): S61-72.

Bhak, G., Y. J. Choe and S. R. Paik (2009). *BMB Rep* **42**(9): 541-551.

Bhagal, N. (2010). *Curr Drug Saf* **5**(4): 293-307.

Biancalana, M. and S. Koide (2010). *Biochimica Et Biophysica Acta* **1804**(7): 1405-1412.

Biancalana, M., K. Makabe, A. Koide and S. Koide (2009). *Journal of Molecular Biology* **385**(4): 1052-1063.

Bischof, J. C. and X. M. He (2005). *Cell Injury: Mechanisms, Responses, and Repair* **1066**: 12-33.

Bloomfield, V. (1966). *Biochemistry* **5**(2): 684-689.

Blumlein, A. and J. J. McManus (2013). *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1834**(10): 2064-2070.

Bondos, S. E. and A. Bicknell (2003). *Analytical Biochemistry* **316**(2): 223-231.

Bonincontro, A., A. De Francesco and G. Onori (1998). *Colloids and Surfaces B-Biointerfaces* **12**(1): 1-5.

Booth, D. R., M. Sunde, V. Bellotti, C. V. Robinson, W. L. Hutchinson, P. E. Fraser, P. N. Hawkins, C. M. Dobson, S. E. Radford, C. C. F. Blake and M. B. Pepys (1997). *Nature* **385**(6619): 787-793.

Bos, O. J., J. F. Labro, M. J. Fischer, J. Wilting and L. H. Janssen (1989). *Journal of Biological Chemistry* **264**(2): 953-959.

Boye, J. I., I. Alli and A. A. Ismail (1996). *Journal of Agricultural and Food Chemistry* **44**(4): 996-1004.

Bragonzi, A., G. Distefano, L. D. Buckberry, G. Acerbis, C. Foglieni, D. Lamotte, G. Campi, A. Marc, M. R. Soria, N. Jenkins and L. Monaco (2000). *Biochimica Et Biophysica Acta* **1474**(3): 273-282.

Breydo, L. P., A. A. Shevchenko and O. A. Kost (1997). *Russian Chemical Bulletin* **46**(7): 1339-1343.

Bruzzesi, Mr, Chiancon, E and E. Antonini (1965). *Biochemistry* **4**(9): 1796-&.

Bryant, C. M. and D. J. McClements (1998). *Trends in Food Science & Technology* **9**(4): 143-151.

Buchner, G. S., R. D. Murphy, N. V. Buchete and J. Kubelka (2011). *Biochimica Et Biophysica Acta* **1814**(8): 1001-1020.

Bujacz, A. (2012). *Acta Crystallographica Section D-Biological Crystallography* **68**: 1278-1289.

Bunnage, M. E. (2011). *Nature Chemical Biology* **7**(6): 335-339.

Calamai, M., C. Canale, A. Relini, M. Stefani, F. Chiti and C. M. Dobson (2005). *Journal of Molecular Biology* **346**(2): 603-616.

Campos, A. M., E. A. Lissi, C. Vergara, M. E. Lanio, C. Alvarez, I. Pazos, V. Morera, Y. Garcia and D. Martinez (1999). *J Protein Chem* **18**(3): 297-306.

Canfield, R. E. (1963). *Journal of Biological Chemistry* **238**: 2698-2707.

Cao, A., D. Hu and L. Lai (2004). *Protein Science* **13**(2): 319-324.

Cao, A., G. Wang, Y. Q. Tang and L. H. Lai (2002). *Biochemical and Biophysical Research Communications* **291**(4): 795-797.

Cao, Y. and H. Li (2008). *Journal of Molecular Biology* **375**(1): 316-324.

Capelle, M. A. H., R. Gurny and T. Arvinte (2007). *European Journal of Pharmaceutics and Biopharmaceutics* **65**(2): 131-148.

Carter, D. C., X. M. He, S. H. Munson, P. D. Twigg, K. M. Gernert, M. B. Broom and T. Y. Miller (1989). *Science* **244**(4909): 1195-1198.

Casi, G. and D. Neri (2012). *Journal of Controlled Release* **161**(2): 422-428.

Chang, E. S. H., T. Y. Liao, T. S. Lim, W. Fann and R. P. Y. Chen (2009). *Journal of Molecular Biology* **385**(4): 1257-1265.

Chang, J. Y. and L. Li (2002). *Febs Letters* **511**(1-3): 73-78.

Chao, C. C., Y. S. Ma and E. R. Stadtman (1997). *Proc Natl Acad Sci U S A* **94**(7): 2969-2974.

Chen, H. X., N. A. Liu, L. F. Shu and R. W. Zong (2004). *Journal of Thermal Analysis and Calorimetry* **78**(3): 1029-1041.

Chen, Y. and M. D. Barkley (1998). *Biochemistry* **37**(28): 9976-9982.

- Chi, E. Y., S. Krishnan, T. W. Randolph and J. F. Carpenter** (2003). *Pharmaceutical Research* **20**(9): 1325-1336.
- Chi, E. Y., J. Weickmann, J. F. Carpenter, M. C. Manning and T. W. Randolph** (2005). *Journal of Pharmaceutical Sciences* **94**(2): 256-274.
- Chiti, F., P. Webster, N. Taddei, A. Clark, M. Stefani, G. Ramponi and C. M. Dobson** (1999). *Proc Natl Acad Sci U S A* **96**(7): 3590-3594.
- Chusainow, J., Y. S. Yang, Y. H. M. Yeo, P. C. Toh, P. Asvadi, N. S. C. Wong and M. G. S. Yap** (2009). *Biotechnology and Bioengineering* **102**(4): 1182-1196.
- Clague, M. J.** (2013). *Nature* **497**(7447): 49-50.
- Clark, A. H., F. J. Judge, J. B. Richards, J. M. Stubbs and A. Suggett** (1981). *Int J Pept Protein Res* **17**(3): 380-392.
- Clark, A. H., G. M. Kavanagh and S. B. Ross-Murphy** (2001). *Food Hydrocolloids* **15**(4-6): 383-400.
- Cohen, S. I., M. Vendruscolo, C. M. Dobson and T. P. Knowles** (2012). *Journal of Molecular Biology* **421**(2-3): 160-171.
- Constatinescu, D., C. Herrmann and H. Weingartner** (2010). *Physical Chemistry Chemical Physics* **12**(8): 1756-1763.
- Cooper, A.** (1999). *Curr Opin Chem Biol* **3**(5): 557-563.
- Costa, A. R., M. E. Rodrigues, M. Henriques, R. Oliveira and J. Azeredo** (2013). *Crit Rev Biotechnol*.
- Creighton, T. E., N. J. Darby and J. Kemmink** (1996). *Faseb Journal* **10**(1): 110-118.
- Crommelin, D. J. A., G. Storm, R. Verrijck, L. de Leede, W. Jiskoot and W. E. Hennink** (2003). *Int J Pharm* **266**(1-2): 3-16.
- Cromwell, M. E. M., E. Hilario and F. Jacobson** (2006). *Aaps Journal* **8**(3): E572-E579.
- Das, A., R. Chitra, R. R. Choudhury and M. Ramanadham** (2004). *Pramana-Journal of Physics* **63**(2): 363-368.
- Daugherty, A. L. and R. J. Mersny** (2006). *Adv Drug Deliv Rev* **58**(5-6): 686-706.
- Davies, K. J.** (1987). *Journal of Biological Chemistry* **262**(20): 9895-9901.
- Davies, K. J. and M. E. Delsignore** (1987). *Journal of Biological Chemistry* **262**(20): 9908-9913.
- Davies, K. J., M. E. Delsignore and S. W. Lin** (1987). *Journal of Biological Chemistry* **262**(20): 9902-9907.
- Davies, K. J., S. W. Lin and R. E. Pacifici** (1987). *Journal of Biological Chemistry* **262**(20): 9914-9920.
- Davies, K. J. and W. A. Pryor** (2005). *Free Radic Biol Med* **39**(10): 1263-1264.
- Dean, R. T., S. L. Fu, R. Stocker and M. J. Davies** (1997). *Biochemical Journal* **324**: 1-18.
- Demeule, B., R. Gurny and T. Arvinte** (2006). *European Journal of Pharmaceutics and Biopharmaceutics* **62**(2): 121-130.
- den Engelsman, J., P. Garidel, R. Smulders, H. Koll, B. Smith, S. Bassarab, A. Seidl, O. Hainzl and W. Jiskoot** (2011). *Pharm Res* **28**(4): 920-933.
- DeToma, A. S., S. Salamekh, A. Ramamoorthy and M. H. Lim** (2012). *Chem Soc Rev* **41**(2): 608-621.
- Dill, K. A. and D. Shortle** (1991). *Annual Review of Biochemistry* **60**: 795-825.
- Dobson, C. M.** (2004). *Semin Cell Dev Biol* **15**(1): 3-16.
- Dobson, C. M.** (1999). *Trends in Biochemical Sciences* **24**(9): 329-332.
- Dobson, C. M., P. A. Evans and S. E. Radford** (1994). *Trends in Biochemical Sciences* **19**(1): 31-37.
- Dobson, C. M., A. Sali and M. Karplus** (1998). *Angewandte Chemie-International Edition* **37**(7): 868-893.
- Donovan, J. W.** (1973). *Methods Enzymol* **27**: 497-525.
- Duy, C. and J. Fitter** (2006). *Biophysical Journal* **90**(10): 3704-3711.
- Eaton, P., W. Fuller and M. J. Shattock** (2002). *Journal of Biological Chemistry* **277**(24): 21189-21196.
- Eaton, W. A., V. Munoz, S. J. Hagen, G. S. Jas, L. J. Lapidus, E. R. Henry and J. Hofrichter** (2000). *Annu Rev Biophys Biomol Struct* **29**: 327-359.

- Eftink, M. R.** (1994). *Biophysical Journal* **66**(2): 482-501.
- Elvin, J. G., R. G. Couston and C. F. van der Walle** (2013). *Int J Pharm* **440**(1): 83-98.
- Esquembre, R., J. M. Sanz, J. G. Wall, F. del Monte, C. R. Mateo and M. L. Ferrer** (2013). *Physical Chemistry Chemical Physics* **15**(27): 11248-11256.
- Evans, P. A., K. D. Topping, D. N. Woolfson and C. M. Dobson** (1991). *Proteins-Structure Function and Genetics* **9**(4): 248-266.
- Ferrer, M. L., R. Duchowicz, B. Carrasco, J. G. de la Torre and A. U. Acuna** (2001). *Biophysical Journal* **80**(5): 2422-2430.
- Fink, A. L.** (1998). *Folding & Design* **3**(1): R9-R23.
- Fink, A. L., L. J. Calciano, Y. Goto, T. Kurotsu and D. R. Palleros** (1994). *Biochemistry* **33**(41): 12504-12511.
- Finnegan, M., E. Linley, S. P. Denyer, G. McDonnell, C. Simons and J. Y. Maillard** (2010). *J Antimicrob Chemother* **65**(10): 2108-2115.
- Flora, K., J. D. Brennan, G. A. Baker, M. A. Doody and F. V. Bright** (1998). *Biophysical Journal* **75**(2): 1084-1096.
- Formoso, C. and L. S. Forster** (1975). *Journal of Biological Chemistry* **250**(10): 3738-3745.
- Frank, S.** (1968). *Chemical Physics* **48**: 4748.
- Frare, E., M. F. Mossuto, P. Polverino de Laureto, M. Dumoulin, C. M. Dobson and A. Fontana** (2006). *Journal of Molecular Biology* **361**(3): 551-561.
- Frare, E., P. Polverino De Laureto, J. Zurdo, C. M. Dobson and A. Fontana** (2004). *Journal of Molecular Biology* **340**(5): 1153-1165.
- Freeman, M. L., M. J. Borrelli, M. J. Meredith and J. R. Lepock** (1999). *Free Radical Biology and Medicine* **26**(5-6): 737-745.
- Gao, M., Z. S. She and R. H. Zhou** (2010). *Journal of Physical Chemistry B* **114**(47): 15687-15693.
- Garbett, N. C., C. S. Mekmaysy, C. W. Helm, A. B. Jenson and J. B. Chaires** (2009). *Experimental and Molecular Pathology* **86**(3): 186-191.
- Garcia-Moreno, B.** (2009). *J Biol* **8**(11): 98.
- Garnier, J. P.** (2008). *Harv Bus Rev* **86**(5): 68-70, 72-66, 128.
- Garrison, W. M.** (1987). *Chemical Reviews* **87**(2): 381-398.
- George, A. and W. W. Wilson** (1994). *Acta Crystallogr D Biol Crystallogr* **50**(Pt 4): 361-365.
- Giustarini, D., I. Dalle-Donne, A. Milzani and R. Rossi** (2011). *Mechanisms of Ageing and Development* **132**(4): 141-148.
- Goldberg, M. E., R. Rudolph and R. Jaenicke** (1991). *Biochemistry* **30**(11): 2790-2797.
- Gomez-Hens, A. and M. P. Aguilar-Caballos** (2007). *Trac-Trends in Analytical Chemistry* **26**(3): 171-182.
- Gomez, J., V. J. Hilser, D. Xie and E. Freire** (1995). *Proteins* **22**(4): 404-412.
- Gonzalez-Jimenez, J. and M. Cortijo** (2002). *J Protein Chem* **21**(2): 75-79.
- Gopalan, V., R. Golbik, G. Schreiber, A. R. Fersht and S. Altman** (1997). *Journal of Molecular Biology* **267**(4): 765-769.
- Greene, R. F. and C. N. Pace** (1974). *Journal of Biological Chemistry* **249**(17): 5388-5393.
- Gruebele, M.** (1999). *Annu Rev Phys Chem* **50**: 485-516.
- Grum, F., D. Paine and L. Zoeller** (1972). *Appl Opt* **11**(1): 93-98.
- Guedes, S., R. Vitorino, R. Domingues, F. Amado and P. Domingues** (2009). *Rapid Commun Mass Spectrom* **23**(15): 2307-2315.
- Ha, E., W. Wang and Y. J. Wang** (2002). *Journal of Pharmaceutical Sciences* **91**(10): 2252-2264.
- Halim, A. A. A., H. A. Kadir and S. Tayyab** (2008). *Journal of Biochemistry* **144**(1): 33-38.
- Hames, B. D. and D. Rickwood** (1981). *Gel electrophoresis of proteins : a practical approach*. London, IRL Press.
- Hamilton, R. J. and P. A. Sewell** (1982). *Introduction to High Performance Liquid Chromatography*. London, Chapman and Hall.
- Hammes, G. G.** (2005). *Spectroscopy For The Biological Sciences*. USA, John Wiley & Sons.,
- Hampe, O. G.** (1972). *European Journal of Biochemistry* **31**(1): 32-&.

Hansel, T. T., H. Kropshofer, T. Singer, J. A. Mitchell and A. J. George (2010). *Nat Rev Drug Discov* **9**(4): 325-338.

Hanson, S. R. A., A. A. Chen, J. B. Smith and M. F. Lou (1999). *Journal of Biological Chemistry* **274**(8): 4735-4742.

Harding, F. A., M. M. Stickler, J. Razo and R. B. DuBridge (2010). *MAbs* **2**(3): 256-265.

Hasija, M., L. Li, N. Rahman and S. F. Ausar (2013). *Vaccine: Dev. Ther* **3**: 11-33.

Hawe, A., M. Wiggenhorn, M. van de Weert, J. H. O. Garbe, H. C. Mahler and W. Jiskoot (2012). *Journal of Pharmaceutical Sciences* **101**(3): 895-913.

He, X. M. and D. C. Carter (1993). *Nature* **364**(6435): 362-362.

Hedoux, A., S. Krenzlin, L. Paccou, Y. Guinet, M. P. Flament and J. Siepmann (2010). *Physical Chemistry Chemical Physics* **12**(40): 13189-13196.

Henkler, F., J. Brinkmann and A. Luch (2010). *Cancers (Basel)* **2**(2): 376-396.

Hertzberg, R. P. and A. J. Pope (2000). *Curr Opin Chem Biol* **4**(4): 445-451.

Hillier, R. M., R. L. J. Lyster and G. C. Cheeseman (1980). *Journal of the Science of Food and Agriculture* **31**(11): 1152-1157.

Hoger, K., J. Mathes and W. Friess (2015). *Journal of Pharmaceutical Sciences* **104**(1): 34-43.

Holm, N. K., S. K. Jespersen, L. V. Thomassen, T. Y. Wolff, P. Sehgal, L. A. Thomsen, G. Christiansen, C. B. Andersen, A. D. Knudsen and D. E. Otzen (2007). *Biochimica Et Biophysica Acta-Proteins and Proteomics* **1774**(9): 1128-1138.

Holt, L. A., B. Milligan, D. E. Rivett and F. H. Stewart (1977). *Biochimica Et Biophysica Acta* **499**(1): 131-138.

Hossler, P., S. F. Khattak and Z. J. Li (2009). *Glycobiology* **19**(9): 936-949.

Hou, L. M. and M. G. Zagorski (2004). *Biophysical Journal* **86**(1): 1-2.

Hua, L., R. H. Zhou, D. Thirumalai and B. J. Berne (2008). *Proceedings of the National Academy of Sciences of the United States of America* **105**(44): 16928-16933.

Huang, B. X., H. Y. Kim and C. Dass (2004). *Journal of the American Society for Mass Spectrometry* **15**(8): 1237-1247.

Huerta-Viga, A. and S. Woutersen (2013). *J Phys Chem Lett* **4**(20): 3397-3401.

Huhn, C., M. H. Selman, L. R. Ruhaak, A. M. Deelder and M. Wuhrer (2009). *Proteomics* **9**(4): 882-913.

Huie, R. E. and S. Padmaja (1993). *Free Radical Research Communications* **18**(4): 195-199.

Hunt, J. V., M. A. Bottoms and M. J. Mitchinson (1993). *Biochemical Journal* **291** (Pt 2): 529-535.

Ibrahim, H. R., S. Higashiguchi, L. R. Juneja, M. Kim and T. Yamamoto (1996). *Journal of Agricultural and Food Chemistry* **44**(6): 1416-1423.

ICH, G. (1998). *Safety of Biological Products Prepared from Mammalian Cell Culture* **93**: 211-219.

Ichikawa, T. and H. Terada (1979). *Biochimica Et Biophysica Acta* **580**(1): 120-128.

Ichikawa, T. and H. Terada (1977). *Biochimica Et Biophysica Acta* **494**(1): 267-270.

Igawa, T., H. Tsunoda, T. Kuramochi, Z. Sampei, S. Ishii and K. Hattori (2011). *MAbs* **3**(3): 243-252.

Imoto, T., L. S. Forster, J. A. Rupley and F. Tanaka (1972). *Proc Natl Acad Sci U S A* **69**(5): 1151-1155.

Itakura, K., K. Uchida and S. Kawakishi (1994). *Chem Res Toxicol* **7**(2): 185-190.

Itri, R., W. Caetano, L. R. S. Barbosa and M. S. Baptista (2004). *Brazilian Journal of Physics* **34**(1): 58-63.

Jackson, S. E. (1998). *Folding & Design* **3**(4): R81-R91.

James, S. and J. J. McManus (2012). *J Phys Chem B* **116**(34): 10182-10188.

Jefferis, R. (2009). *Nature Reviews Drug Discovery* **8**(3): 226-234.

Ji, J. A., B. Zhang, W. Cheng and Y. J. Wang (2009). *Journal of Pharmaceutical Sciences* **98**(12): 4485-4500.

Jiskoot, W., T. W. Randolph, D. B. Volkin, C. R. Middaugh, C. Schoneich, G. Winter, W. Friess, D. J. Crommelin and J. F. Carpenter (2012). *Journal of Pharmaceutical Sciences* **101**(3): 946-954.

- Jolles, J., J. Jauregui Adell, I. Bernier and P. Jolles** (1963). *Biochimica Et Biophysica Acta* **78**: 668-689.
- Jomova, K., D. Vondrakova, M. Lawson and M. Valko** (2010). *Molecular and Cellular Biochemistry* **345**(1-2): 91-104.
- Jones, A. J. S., D. I. Papac, E. H. Chin, R. Keck, S. A. Baughman, Y. S. Lin, J. Kneer and J. E. Battersby** (2007). *Glycobiology* **17**(5): 529-540.
- Jung, S. T., T. H. Kang, W. Kelton and G. Georgiou** (2011). *Curr Opin Biotechnol* **22**(6): 858-867.
- Kanner, J. D. and O. Fennema** (1987). *Journal of Agricultural and Food Chemistry* **35**(1): 71-76.
- Kaur, R., J. Kaur, J. Mahajan, R. Kumar and S. Arora** (2014). *Environ Sci Pollut Res Int* **21**(3): 1599-1613.
- Kayser, V., N. Chennamsetty, V. Voynov, K. Forrer, B. Helk and B. L. Trout** (2011). *Biotechnol J* **6**(1): 38-44.
- Keerati-u-rai, M., M. Miriani, S. Iametti, F. Bonomi and M. Corredig** (2012). *Colloids and Surfaces B-Biointerfaces* **93**: 41-48.
- Kelly, J. W.** (1998). *Curr Opin Struct Biol* **8**(1): 101-106.
- Keren, D. F.** (2003). *Protein Electrophoresis in Clinical Diagnosis*. Great Britain, Edward Arnold.
- Khan, M. Y., S. K. Agarwal and S. Hangloo** (1987). *Journal of Biochemistry* **102**(2): 313-317.
- Khurana, R., C. Coleman, C. Ionescu-Zanetti, S. A. Carter, V. Krishna, R. K. Grover, R. Roy and S. Singh** (2005). *Journal of Structural Biology* **151**(3): 229-238.
- Khurana, R., J. R. Gillespie, A. Talapatra, L. J. Minert, C. Ionescu-Zanetti, I. Millett and A. L. Fink** (2001). *Biochemistry* **40**(12): 3525-3535.
- Kim, Y. H., A. H. Berry, D. S. Spencer and W. E. Stites** (2001). *Protein Eng* **14**(5): 343-347.
- Kohler, G. and C. Milstein** (1975). *Nature* **256**(5517): 495-497.
- Kontoravdi, C., S. P. Asprey, E. N. Pistikopoulos and A. Mantalaris** (2007). *Computers & Chemical Engineering* **31**(5-6): 392-400.
- Kopito, R. R.** (2000). *Trends in Cell Biology* **10**(12): 524-530.
- Kopoldova, J. and J. Liebster** (1963). *Int J Appl Radiat Isot* **14**: 493-498.
- Krebs, M. R., D. K. Wilkins, E. W. Chung, M. C. Pitkeathly, A. K. Chamberlain, J. Zurdo, C. V. Robinson and C. M. Dobson** (2000). *Journal of Molecular Biology* **300**(3): 541-549.
- Krebs, M. R. H., E. H. C. Bromley and A. M. Donald** (2005). *Journal of Structural Biology* **149**(1): 30-37.
- Krebs, M. R. H., K. R. Domike and A. M. Donald** (2009). *Biochemical Society Transactions* **37**: 682-686.
- Krogull, M. K. and O. Fennema** (1987). *Journal of Agricultural and Food Chemistry* **35**(1): 66-70.
- Kueltzo, L. A., B. Ersoy, J. P. Ralston and C. R. Middaugh** (2003). *Journal of Pharmaceutical Sciences* **92**(9): 1805-1820.
- Kumar, V., V. K. Sharma and D. S. Kalonia** (2005). *Int J Pharm* **294**(1-2): 193-199.
- Kumaran, R. and P. Ramamurthy** (2011). *Journal of Fluorescence* **21**(4): 1499-1508.
- Lahana, R.** (2003). *Drug Discov Today* **8**(15): 655-656.
- Lakowicz, J. R.** (2006). *Principles of Fluorescence Spectroscopy*, springer.
- Lardinois, O. M., K. F. Medzihradzky and P. R. O. de Montellano** (1999). *Journal of Biological Chemistry* **274**(50): 35441-35448.
- Lasch, P., T. Petras, O. Ullrich, J. Backmann, D. Naumann and T. Grune** (2001). *Journal of Biological Chemistry* **276**(12): 9492-9502.
- Laurents, D. V. and R. L. Baldwin** (1997). *Biochemistry* **36**(6): 1496-1504.
- Le, H. T., A. F. Chaffotte, E. Demey-Thomas, J. Vinh, B. Friguet and J. Mary** (2009). *Journal of Molecular Biology* **393**(1): 58-66.
- Lee, C. C., J. M. Perchiacca and P. M. Tessier** (2013). *Trends in Biotechnology* **31**(11): 612-620.
- Lee, J. G., K. Baek, N. Soetandyo and Y. Ye** (2013). *Nat Commun* **4**: 1568.

- Leibiger, H., D. Wustner, R. D. Stigler and U. Marx (1999). *Biochemical Journal* **338** (Pt 2): 529-538.
- Lepedda, A. J., A. Zinellu, G. Nieddu, E. Zinellu, C. Carru, R. Spirito, A. Guarino, P. De Muro and M. Formato (2013). *Oxidative Medicine and Cellular Longevity*.
- LeVine, H., 3rd (1999). *Methods Enzymol* **309**: 274-284.
- Li, F., N. Vijayasankaran, A. Shen, R. Kiss and A. Amanullah (2010). *MAbs* **2**(5): 466-479.
- Li, H. J. and M. d'Anjou (2009). *Current Opinion in Biotechnology* **20**(6): 678-684.
- Liang, J. H. (1999). *Food Chemistry* **66**(1): 103-108.
- Lichtman, J. W. and J. A. Conchello (2005). *Nat Methods* **2**(12): 910-919.
- Liebster, J., J. Kopoldova and A. Babicky (1963). *Radiation Research* **19**: 551-558.
- Lim, W. K., J. Rosgen and S. W. Englander (2009). *Proceedings of the National Academy of Sciences of the United States of America* **106**(8): 2595-2600.
- Liu, Y., G. Fiskum and D. Schubert (2002). *Journal of Neurochemistry* **80**(5): 780-787.
- Lucas, L. H., B. A. Ersoy, L. A. Kuelzto, S. B. Joshi, D. T. Brandau, N. Thyagarajapuram, L. J. Peek and C. R. Middaugh (2006). *Protein Science* **15**(10): 2228-2243.
- Lumry, R. and H. Eyring (1954). *The Journal of Physical Chemistry* **58**(2): 110-120.
- Luo, J. J., F. G. Wu, J. S. Yu, R. Wang and Z. W. Yu (2011). *Journal of Physical Chemistry B* **115**(28): 8901-8909.
- Ma, S. F., M. Nishikawa, Y. Yabe, F. Yamashita and M. Hashida (2006). *Biological & Pharmaceutical Bulletin* **29**(9): 1926-1930.
- Ma, Y. S., C. C. Chao and E. R. Stadtman (1999). *Archives of Biochemistry and Biophysics* **363**(1): 129-134.
- Macarron, R. (2006). *Drug Discov Today* **11**(7-8): 277-279.
- Macarron, R., M. N. Banks, D. Bojanic, D. J. Burns, D. A. Cirovic, T. Garyantes, D. V. S. Green, R. P. Hertzberg, W. P. Janzen, J. W. Paslay, U. Schopfer and G. S. Sittampalam (2011). *Nature Reviews Drug Discovery* **10**(3): 188-195.
- Mach, H. and T. Arvinte (2011). *European Journal of Pharmaceutics and Biopharmaceutics* **78**(2): 196-207.
- Mach, H. and C. R. Middaugh (1994). *Analytical Biochemistry* **222**(2): 323-331.
- Mach, H., J. A. Thomson, C. R. Middaugh and R. V. Lewis (1991). *Archives of Biochemistry and Biophysics* **287**(1): 33-40.
- Mahler, H. C., W. Friess, U. Grauschopf and S. Kiese (2009). *Journal of Pharmaceutical Sciences* **98**(9): 2909-2934.
- Mahler, H. C. and W. Jiskoot (2012). *Analysis of Aggregates and Particles in Protein Pharmaceuticals*, John Wiley.
- Majhi, P. R., R. R. Ganta, R. P. Vanam, E. Seyrek, K. Giger and P. L. Dubin (2006). *Langmuir* **22**(22): 9150-9159.
- Majorek, K. A., P. J. Porebski, A. Dayal, M. D. Zimmerman, K. Jablonska, A. J. Stewart, M. Chruszcz and W. Minor (2012). *Mol Immunol* **52**(3-4): 174-182.
- Majumdar, R., P. Manikwar, J. M. Hickey, H. S. Samra, H. A. Sathish, S. M. Bishop, C. R. Middaugh, D. B. Volkin and D. D. Weis (2013). *Biochemistry* **52**(19): 3376-3389.
- Malo, N., J. A. Hanley, S. Cerquozzi, J. Pelletier and R. Nadon (2006). *Nature Biotechnology* **24**(2): 167-175.
- Mann, C. J., C. A. Royer and C. R. Matthews (1993). *Protein Science* **2**(11): 1853-1861.
- Manning, M. C., D. K. Chou, B. M. Murphy, R. W. Payne and D. S. Katayama (2010). *Pharm Res* **27**(4): 544-575.
- Manno, M., A. Emanuele, V. Martorana, D. Bulone, P. L. San Biagio, M. B. Palma-Vittorelli and M. U. Palma (1999). *Physical Review E* **59**(2): 2222-2230.
- Marino, K., J. Bones, J. J. Kattla and P. M. Rudd (2010). *Nature Chemical Biology* **6**(10): 713-723.
- Mark, H. and J. Workman (2007). *Chemometrics in Spectroscopy*: 339-350.
- Matsudomi, N., D. Rector and J. E. Kinsella (1991). *Food Chemistry* **40**(1): 55-69.
- Mayr, L. M. and D. Bojanic (2009). *Curr Opin Pharmacol* **9**(5): 580-588.
- Mcguire, R. and I. Feldman (1975). *Biopolymers* **14**(5): 1095-1102.
- Mckenzie, H. A. and F. H. White (1991). *Advances in Protein Chemistry* **41**: 173-315.

McManus, J. J., A. Lomakin, O. Ogun, A. Pande, M. Basan, J. Pande and G. B. Benedek (2007). *Proceedings of the National Academy of Sciences of the United States of America* **104**(43): 16856-16861.

Merry, A. H., D. C. Neville, L. Royle, B. Matthews, D. J. Harvey, R. A. Dwek and P. M. Rudd (2002). *Analytical Biochemistry* **304**(1): 91-99.

Michnik, A. (2003). *Journal of Thermal Analysis and Calorimetry* **71**(2): 509-519.

Michnik, A., K. Michalik and Z. Drzazga (2005). *Journal of Thermal Analysis and Calorimetry* **80**(2): 399-406.

Michnik, A., K. Michalik, A. Kluczevska and Z. Drzazga (2006). *Journal of Thermal Analysis and Calorimetry* **84**(1): 113-117.

Militello, V., C. Casarino, A. Emanuele, A. Giostra, F. Pullara and M. Leone (2004). *Biophysical Chemistry* **107**(2): 175-187.

Militello, V., V. Vetri and M. Leone (2003). *Biophysical Chemistry* **105**(1): 133-141.

Mine, Y., T. Noutomi and N. Haga (1990). *Journal of Agricultural and Food Chemistry* **38**(12): 2122-2125.

Mirzaei, H. and F. Regnier (2008). *J Chromatogr B Analyt Technol Biomed Life Sci* **873**(1): 8-14.

Molinari, M. and A. Helenius (1999). *Nature* **402**(6757): 90-93.

Moller, I. M., P. E. Jensen and A. Hansson (2007). *Annu Rev Plant Biol* **58**: 459-481.

Molloy, S., R. M. Fesinmeyer, T. Martinez, P. D. Murphy, M. E. Pelletier, M. J. Treuheit and G. R. Kleemann (2015). *Journal of Pharmaceutical Sciences* **104**(2): 508-514.

Morales, R., I. Moreno-Gonzalez and C. Soto (2013). *PLoS Pathog* **9**(9).

Moreno-Gonzalez, I. and C. Soto (2011). *Semin Cell Dev Biol* **22**(5): 482-487.

Morozova-Roche, L. A., J. Zurdo, A. Spencer, W. Noppe, V. Receveur, D. B. Archer, M. Joniau and C. M. Dobson (2000). *Journal of Structural Biology* **130**(2-3): 339-351.

Mozo-Villarias, A. (2002). *Journal of Biochemical and Biophysical Methods* **50**(2-3): 163-178.

Murayama, K. and M. Tomida (2004). *Biochemistry* **43**(36): 11526-11532.

Myers, J. K., C. N. Pace and J. M. Scholtz (1995). *Protein Science* **4**(10): 2138-2148.

Narhi, L. O., Y. J. Jiang, S. Cao, K. Benedek and D. Shnek (2009). *Current Pharmaceutical Biotechnology* **10**(4): 373-381.

Narhi, L. O., J. Schmit, K. Bechtold-Peters and D. Sharma (2012). *Journal of Pharmaceutical Sciences* **101**(2): 493-498.

Neergaard, M. S., A. D. Nielsen, H. Parshad and M. Van de Weert (2014). *Journal of Pharmaceutical Sciences* **103**(1): 115-127.

Nelson, A. L., E. Dhimolea and J. M. Reichert (2010). *Nat Rev Drug Discov* **9**(10): 767-774.

Nydegger, U. E. and M. Sturzenegger (1999). *Drug Saf* **21**(3): 171-185.

Ojeda, C. B. and E. S. Rojas (2004). *Analytica Chimica Acta* **518**(1-2): 1-24.

Owczarz, M. and P. Arosio (2014). *Biophysical Journal* **107**(1): 197-207.

Pace, C. N. (1986). *Methods Enzymol* **131**: 266-280.

Pardon, E., P. Haezebrouck, A. Debaetselier, S. D. Hooke, K. T. Fancourt, J. Desmet, C. M. Dobson, H. Vandael and M. Joniau (1995). *Journal of Biological Chemistry* **270**(18): 10514-10524.

Pastor, I., M. L. Ferrer, M. P. Lillo, J. Gomez and C. R. Mateo (2007). *J Phys Chem B* **111**(39): 11603-11610.

Pathak, y. and S. Benita (2012). *Antibody-Mediated Drug Delivery System*. New Jersey, Wiley.

Paul, W. E. (1993). *Fundamental Immunology*. New York, Raven Press.

Pereira, D. A. and J. A. Williams (2007). *Br J Pharmacol* **152**(1): 53-61.

Peskin, A. V., F. M. Low, L. N. Paton, G. J. Maghzal, M. B. Hampton and C. C. Winterbourn (2007). *Journal of Biological Chemistry* **282**(16): 11885-11892.

Peters, T. (1985). *Advances in Protein Chemistry* **37**: 161-245.

Philo, J. S. and T. Arakawa (2009). *Curr Pharm Biotechnol* **10**(4): 348-351.

Pike, A. C. W. and K. R. Acharya (1994). *Protein Science* **3**(4): 706-710.

Pisal, D. S., M. P. Kosloski and S. V. Balu-Iyer (2010). *Journal of Pharmaceutical Sciences* **99**(6): 2557-2575.

- Price, W. S., F. Tsuchiya and Y. Arata (1999). *Journal of the American Chemical Society* **121**(49): 11503-11512.
 . Retrieved 11/02/2015.
- Puchala, M. and H. Schuessler (1993). *International Journal of Radiation Biology* **64**(2): 149-156.
- Puchala, M. and H. Schuessler (1995). *International Journal of Peptide and Protein Research* **46**(3-4): 326-332.
- Raccosta, S., M. Manno, D. Bulone, D. Giacomazza, V. Militello, V. Martorana and P. L. San Biagio (2010). *Eur Biophys J* **39**(6): 1007-1017.
- Ragone, R., G. Colonna, C. Balestrieri, L. Servillo and G. Irace (1984). *Biochemistry* **23**(8): 1871-1875.
- Raju, T. S., J. B. Briggs, S. M. Borge and A. J. S. Jones (2000). *Glycobiology* **10**(5): 477-486.
- Rathore, A. S. (2011). *Comprehensive Biotechnology, Vol 3: Industrial Biotechnology and Commodity Products, 2nd Edition*: 521-529.
- Raymond, S. and L. Weintraub (1959). *Science* **130**(3377): 711-711.
- Reed, R. G., R. C. Feldhoff, O. L. Clute and T. Peters, Jr. (1975). *Biochemistry* **14**(21): 4578-4583.
- Reisler, E., J. Pouyet and Eisenber.H (1970). *Biochemistry* **9**(15): 3095-&.
- Reubsat, J. L. E., J. H. Beijnen, A. Bult, R. J. van Maanen, J. A. D. Marchal and W. J. M. Underberg (1998). *Journal of Pharmaceutical and Biomedical Analysis* **17**(6-7): 955-978.
- Rhode, H., M. Schulze, S. Renard, P. Zimmerman, T. Moore, G. A. Cumme and A. Horn (2004). *Journal of Biomolecular Screening* **9**(8): 726-733.
- Ripple, D. C. and M. N. Dimitrova (2012). *Journal of Pharmaceutical Sciences* **101**(10): 3568-3579.
- Roberts, C. J. (2007). *Biotechnology and Bioengineering* **98**(5): 927-938.
- Roberts, C. J. and W. Wang (2010). *Aggregation of Therapeutic Proteins*. New Jersey, John Wiley and Sons.
- Roberts, S. A. (2001). *Xenobiotica* **31**(8-9): 557-589.
- Robinson, D. R. and W. P. Jencks (1965). *Journal of the American Chemical Society* **87**(11): 2462-&.
- Rojas, F. S., C. B. Ojeda and J. M. C. Pavon (1988). *Talanta* **35**(10): 753-761.
- Rosenberg, A. S. (2006). *Aaps Journal* **8**(3): E501-E507.
- Ross, P. D., J. S. Finlayson and A. Shrake (1984). *Vox Sanguinis* **47**(1): 19-27.
- Roth, Z. and I. Khalaila (2012). *Mol Reprod Dev* **79**(7): 478-487.
- Royer, C. A., C. J. Mann and C. R. Matthews (1993). *Protein Science* **2**(11): 1844-1852.
- Royle, L., M. P. Campbell, C. M. Radcliffe, D. M. White, D. J. Harvey, J. L. Abrahams, Y. G. Kim, G. W. Henry, N. A. Shadick, M. E. Weinblatt, D. M. Lee, P. M. Rudd and R. A. Dwek (2008). *Analytical Biochemistry* **376**(1): 1-12.
- Ruhaak, L. R., G. Zauner, C. Huhn, C. Bruggink, A. M. Deelder and M. Wuhrer (2010). *Anal Bioanal Chem* **397**(8): 3457-3481.
- Saakov, V. S., V. Z. Drapkin, A. I. Krivchenko, E. V. Rozengart, Y. V. Bogachev and M. N. Knyazev (2012). *Derivative Spectrophotometry and Electron Spin Resonance (ESR) Spectroscopy for Ecological and Biological Questions*, Springer.
- Saluja, A., R. M. Fesinmeyer, S. Hogan, D. N. Brems and Y. R. Gokarn (2010). *Biophysical Journal* **99**(8): 2657-2665.
- San Biagio, P. L., V. Martorana, A. Emanuele, S. M. Vaiana, M. Manno, D. Bulone, M. B. Palma-Vittorelli and M. U. Palma (1999). *Proteins* **37**(1): 116-120.
- Sancho, J. (2013). *Archives of Biochemistry and Biophysics* **531**(1-2): 4-13.
- Sante-Lhoutellier, V., L. Aubry and P. Gatellier (2007). *Journal of Agricultural and Food Chemistry* **55**(13): 5343-5348.
- Sarciaux, J. M., S. Mansour, M. J. Hageman and S. L. Nail (1999). *Journal of Pharmaceutical Sciences* **88**(12): 1354-1361.
- Sassi, P., G. Onori, A. Giugliarelli, M. Paolantoni, S. Cinelli and A. Morresi (2011). *Journal of Molecular Liquids* **159**(1): 112-116.
- Schellekens, H. (2003). *Neurology* **61**(9 Suppl 5): S11-12.

- Schellman, J. A. (1994). *Biopolymers* **34**(8): 1015-1026.
- Schmid, F. X. (2001). *Biological Macromolecules: UV-visible Spectrophotometry*.
- Schoneich, C. (2008). *Chem Res Toxicol* **21**(6): 1175-1179.
- Schroeder, H. W. and L. Cavacini (2010). *Journal of Allergy and Clinical Immunology* **125**(2): S41-S52.
- Schuessler, H. and A. Herget (1980). *International Journal of Radiation Biology* **37**(1): 71-80.
- Schuessler, H. and K. Schilling (1984). *International Journal of Radiation Biology* **45**(3): 267-281.
- Schwalbe, H., S. B. Grimshaw, A. Spencer, M. Buck, J. Boyd, C. M. Dobson, C. Redfield and L. J. Smith (2001). *Protein Science* **10**(4): 677-688.
- Schwarz, F. and M. Aebi (2011). *Curr Opin Struct Biol* **21**(5): 576-582.
- Scott, A. M., J. P. Allison and J. D. Wolchok (2012). *Cancer Immun* **12**: 14.
- Sen, P., S. Fatima, B. Ahmad and R. H. Khan (2009). *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* **74**(1): 94-99.
- Shacter, E. (2000). *Drug Metab Rev* **32**(3-4): 307-326.
- Shechter, Y., Y. Burstein and A. Patchornik (1975). *Biochemistry* **14**(20): 4497-4503.
- Shen, C. L. and R. M. Murphy (1995). *Biophysical Journal* **69**(2): 640-651.
- Shrake, A., J. S. Finlayson and P. D. Ross (1984). *Vox Sanguinis* **47**(1): 7-18.
- Shukla, A. A. and J. Thommes (2010). *Trends in Biotechnology* **28**(5): 253-261.
- Sies, H. (1997). *Exp Physiol* **82**(2): 291-295.
- Simat, T. J. and H. Steinhart (1998). *Journal of Agricultural and Food Chemistry* **46**(2): 490-498.
- Sirotkin, V. A. and R. Winter (2010). *Journal of Physical Chemistry B* **114**(50): 16881-16886.
- Sola, R. J. and K. Griebenow (2009). *Journal of Pharmaceutical Sciences* **98**(4): 1223-1245.
- Sola, R. J. and K. Griebenow (2010). *Biodrugs* **24**(1): 9-21.
- Solli, N. J. and T. T. Herskovits (1973). *Analytical Biochemistry* **54**(2): 370-378.
- Stadtman, E. R. (2006). *Free Radic Res* **40**(12): 1250-1258.
- Stanley, P., H. Schachter and N. Taniguchi (2009). *N-Glycans*. Cold Spring Harbor (NY).
- Staub, A., D. Guillarme, J. Schappler, J. L. Veuthey and S. Rudaz (2011). *Journal of Pharmaceutical and Biomedical Analysis* **55**(4): 810-822.
- Street, T. O., N. Courtemanche and D. Barrick (2008). *Biophysical Tools for Biologists: Vol 1 in Vitro Techniques* **84**: 295-+.
- Sugai, S. and M. Ikeguchi (1994). *Adv Biophys* **30**: 37-84.
- Sugiyama, M., N. Fujii, Y. Morimoto, K. Itoh, K. Mori and T. Fukunaga (2010). *Chem Biodivers* **7**(6): 1380-1388.
- Sulatskaya, A. I., K. K. Turoverov and I. M. Kuznetsova (2010). *Spectroscopy-an International Journal* **24**(1-2): 169-172.
- Sulkowska, A., B. Bojko, J. Rownicka, D. Pentak and W. Sulkowski (2003). *Journal of Molecular Structure* **651**: 237-243.
- Sulkowska, A., J. Rownicka, B. Bojko, J. Pozyccka, I. Zubik-Skupien and W. Sulkowski (2004). *Journal of Molecular Structure* **704**(1-3): 291-295.
- Sulkowska, A., J. Rownicka, J. Pozyccka, B. Bojko and W. W. Sulkowski (2005). *Journal of Molecular Structure* **744**: 775-779.
- Sun, W., M. Zhao, B. Yang, H. Zhao and C. Cui (2011). *Meat Sci* **88**(3): 462-467.
- Sundberg, S. A. (2000). *Current Opinion in Biotechnology* **11**(1): 47-53.
- Tanaka, N., H. Nishizawa and S. Kunugi (1997). *Biochimica Et Biophysica Acta* **1338**(1): 13-20.
- Tanford, C., R. H. Pain and N. S. Otchin (1966). *Journal of Molecular Biology* **15**(2): 489-504.
- Tayyab, S., B. Ahmad, Y. Kumar and M. M. Khan (2002). *International Journal of Biological Macromolecules* **30**(1): 17-22.
- Tayyab, S., N. Sharma and M. Mushahid Khan (2000). *Biochem Biophys Res Commun* **277**(1): 83-88.
- Teillaud, J. L. (2005). *Expert Opinion on Biological Therapy* **5**: S15-S27.
- Tessier, P. M. and A. M. Lenhoff (2003). *Current Opinion in Biotechnology* **14**(5): 512-516.

- Tharmalingam, T., B. Adamczyk, M. A. Doherty, L. Royle and P. M. Rudd** (2013). *Glycoconj J* **30**(2): 137-146.
- Thusius, D.** (1975). *Journal of Molecular Biology* **94**(3): 367-383.
- Tien, M., B. S. Berlett, R. L. Levine, P. B. Chock and E. R. Stadtman** (1999). *Proceedings of the National Academy of Sciences of the United States of America* **96**(14): 7809-7814.
- Timasheff, S. N. and G. F. Xie** (2003). *Biophysical Chemistry* **105**(2-3): 421-448.
- Tobitani, A. and S. B. Murphy** (1997). *Macromolecules* **30**(17): 4845-4854.
- Togashi, D. M. and A. G. Ryder** (2008). *Journal of Fluorescence* **18**(2): 519-526.
- Togashi, D. M. and A. G. Ryder** (2006). *Journal of Fluorescence* **16**(2): 153-160.
- Togashi, D. M., A. G. Ryder and D. O'Shaughnessy** (2010). *Journal of Fluorescence* **20**(2): 441-452.
- Tovey, M. G. and C. Lallemand** (2011). *Ther Adv Drug Saf* **2**(3): 113-128.
- Valeur, B.** (2002). *Molecular Fluorescence: Principles and Applications*. Germany, Wiley-VCH.
- Vetri, V., C. Canale, A. Relini, F. Librizzi, V. Militello, A. Gliozzi and M. Leone** (2007). *Biophysical Chemistry* **125**(1): 184-190.
- Vetri, V., F. Librizzi, M. Leone and V. Militello** (2007). *European Biophysics Journal with Biophysics Letters* **36**(7): 717-725.
- Wallqvist, A., D. G. Covell and D. Thirumalai** (1998). *Journal of the American Chemical Society* **120**(2): 427-428.
- Wang, W., K. Antonsen, Y. J. Wang and D. Q. Wang** (2008). *Eur J Pharm Sci* **33**(2): 120-127.
- Wang, W., S. Nema and D. Teagarden** (2010). *Int J Pharm* **390**(2): 89-99.
- Wang, W., S. K. Singh, N. Li, M. R. Toler, K. R. King and S. Nema** (2012). *Int J Pharm* **431**(1-2): 1-11.
- Wang, Y., A. Lomakin, J. J. McManus, O. Ogun and G. B. Benedek** (2010). *Proceedings of the National Academy of Sciences of the United States of America* **107**(30): 13282-13287.
- Weaver, R.** (2003). *American Laboratory* **35**(20): 55-+.
- Weiner, L. M., R. Surana and S. Z. Wang** (2010). *Nature Reviews Immunology* **10**(5): 317-327.
- Weiss, W. F. t., T. M. Young and C. J. Roberts** (2009). *Journal of Pharmaceutical Sciences* **98**(4): 1246-1277.
- Westermeier, R.** (2001). *Electrophoresis in Practice*. Germany, John Wiley and Son.
- Wilson, G., S. J. Ford, A. Cooper, L. Hecht, Z. Q. Wen and L. D. Barron** (1995). *Journal of Molecular Biology* **254**(4): 747-760.
- Wolcke, J. and D. Ullmann** (2001). *Drug Discov Today* **6**(12): 637-646.
- Wolff, S. P., A. Garner and R. T. Dean** (1986). *Trends in Biochemical Sciences* **11**(1): 27-31.
- Wright, A. K. and M. R. Thompson** (1975). *Biophysical Journal* **15**(2 Pt 1): 137-141.
- Wu, C. and J. E. Shea** (2011). *Curr Opin Struct Biol* **21**(2): 209-220.
- Wu, L. Z., B. L. Ma, Y. B. Sheng and W. Wang** (2008). *Journal of Molecular Structure* **891**(1-3): 167-172.
- Wu, W., C. M. Zhang, X. Z. Kong and Y. F. Hua** (2009). *Food Chemistry* **116**(1): 295-301.
- Xia, X. and S. T. Wong** (2012). *Stem Cells* **30**(9): 1800-1807.
- Yamada, T.** (2011). *Keio J Med* **60**(2): 37-46.
- Yamasaki, M., H. Yano and K. Aoki** (1990). *International Journal of Biological Macromolecules* **12**(4): 263-268.
- Yang, M. and M. Butler** (2000). *Biotechnol Prog* **16**(5): 751-759.
- Yano, Y. F.** (2012). *Journal of Physics-Condensed Matter* **24**(50).
- Ye, L., Y. Liao, M. M. Zhao and W. Z. Sun** (2013). *Journal of Chemistry*.
- Yoshimura, Y., Y. Lin, H. Yagi, Y. H. Lee, H. Kitayama, K. Sakurai, M. So, H. Ogi, H. Naiki and Y. Goto** (2012). *Proc Natl Acad Sci U S A* **109**(36): 14446-14451.
- Zheng, K., C. Bantog and R. Bayer** (2011). *MAbs* **3**(6): 568-576.
- Zilg, H., H. Schneider and F. R. Seiler** (1980). *Safety of Biological Products Prepared from Mammalian Cell Culture* **48**: 31-42.
- Zocchi, G.** (1997). *Proc Natl Acad Sci U S A* **94**(20): 10647-10651.
- Zolls, S., R. Tantipolphan, M. Wiggenhorn, G. Winter, W. Jiskoot, W. Friess and A. Hawe** (2012). *Journal of Pharmaceutical Sciences* **101**(3): 914-935.